



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN ASTROFISICA

CICLO 36

Settore Concorsuale: 02/C1 - ASTRONOMIA, ASTROFISICA, FISICA DELLA TERRA E DEI PIANETI

Settore Scientifico Disciplinare: FIS/05 - ASTRONOMIA E ASTROFISICA

Unveiling the Expansion History of the Universe with Cosmic Chronometers and Gravitational Waves

Presentata da: **Nicola Borghi**

Coordinatore Dottorato
Andrea Miglio

Supervisore
Michele Moresco

Co-supervisore
Andrea Cimatti

Esame finale anno 2024

To Alice.

*“E chi mi impenna, e chi mi scalda il core?
Chi non mi fa temer fortuna o morte?
Chi le catene ruppe e quelle porte,
Onde rari son sciolti ed escon fore?
L'etadi, gli anni, i mesi, i giorni e l'ore
Figlie ed armi del tempo, e quella corte
A cui né ferro, né diamante è forte,
Assicurato m'han dal suo furore.
Quindi l'ali sicure a l'aria porgo;
Né temo intoppo di cristallo o vetro,
Ma fendo i cieli e a l'infinito m'ergo”*

Giordano Bruno
De l'infinito, universo et Mondi

Abstract

This Thesis explores two novel and independent cosmological probes, Cosmic Chronometers (CCs) and Gravitational Waves (GWs), to measure the expansion history of the Universe. CCs provide direct and cosmology-independent measurements of the Hubble parameter $H(z)$ up to $z \sim 2$. In parallel, GWs provide a direct measurement of the luminosity distance without requiring additional calibration, thus yielding a direct measurement of the Hubble constant $H_0 = H(z = 0)$. This Thesis extends the methodologies of both these probes to maximize their scientific yield. This is achieved by accounting for the interplay of cosmological and astrophysical parameters to derive them jointly, study possible degeneracies, and eventually minimize potential systematic effects. As a legacy value, this work also provides interesting insights into galaxy evolution and compact binary population properties. The first part presents a detailed study of intermediate-redshift passive galaxies as CCs, with a focus on the selection process and the study of their stellar population properties using specific spectral features. From their differential aging, we derive a new measurement of the Hubble parameter $H(z)$ and thoroughly assess potential systematics. In the second part, we develop a novel methodology and pipeline to obtain joint cosmological and astrophysical population constraints using GWs in combination with galaxy catalogs. This is applied to GW170817 to obtain a measurement of H_0 . We then perform realistic forecasts to predict joint cosmological and astrophysical constraints from black hole binary mergers for upcoming gravitational wave observatories and galaxy surveys. Using these two probes we provide an independent reconstruction of $H(z)$ with direct measurements of H_0 from GWs and $H(z)$ up to $z \sim 2$ from CCs and demonstrate that they can be powerful independent probes to unveil the expansion history of the Universe.

Contents

Introduction	1
1 An expanding Universe	5
1.1 Fundamentals of General Relativity	7
1.2 Friedmann–Lemaître–Robertson–Walker cosmology	8
1.2.1 Hubble–Leimaître law and definition of redshift	8
1.2.2 Distances and geometries	10
1.2.3 Friedmann equations	10
1.2.4 Fluids of the Universe	11
1.3 The Standard Cosmological Model	14
1.3.1 Issues of the Standard Cosmological Model	15
1.3.2 Proposed solutions for the Hubble tension	16
1.4 Observational Probes	19
1.4.1 Standard Candles (SNe)	19
1.4.2 Standard Rulers (BAO)	21
1.4.3 Cosmic Microwave Background (CMB)	22
1.4.4 Emerging probes	24
1.5 Open questions in Cosmology	28
1.5.1 Aim of the Thesis	29
2 Cosmology with Cosmic Chronometers	31
2.1 Framework: theory and observations	33
2.1.1 The cosmic chronometers method	33
2.1.2 Passively evolving galaxies	34
2.1.3 Measuring the differential ages	35
2.1.4 Main systematics	37
2.2 Selection of passive galaxies	39
2.2.1 The passive sample	39
2.2.2 <code>PyLick</code> : a new tool to measure spectral features	43
2.2.3 Measuring spectral features with <code>PyLick</code>	47
2.2.4 Observed spectral features	50
2.3 Physical properties of passive galaxies at intermediate z	53
2.3.1 Stellar population models	53
2.3.2 MCMC Analysis	54
2.3.3 The H/K ratio as a stellar population diagnostic	57

2.3.4	Physical parameters versus σ_* and M_*	59
2.3.5	Physical parameters versus z	63
2.3.6	Analysis with different sets of absorption features	65
2.3.7	Analysis with different models	67
2.4	A new measurement of $H(z)$ at $z \sim 0.7$	71
2.4.1	The direct measurement	71
2.4.2	Assessing systematic uncertainties	71
2.4.3	Fit to the age–redshift relations	74
2.5	Further studies with cosmic chronometers	78
2.6	Summary	80
3	Cosmology with Gravitational Waves	83
3.1	Framework: theory and observations	85
3.1.1	Fundamentals of gravitational wave theory	85
3.1.2	Gravitational wave detectors	91
3.1.3	Gravitational wave sources	96
3.1.4	Gravitational wave cosmology	98
3.1.5	Hierarchical Bayesian Inference	102
3.2	Extending the standard sirens method	105
3.2.1	Population prior	106
3.2.2	Ideal galaxy catalog	109
3.2.3	Realistic galaxy catalog	109
3.2.4	Selection effects	112
3.2.5	Full form of the likelihood	115
3.2.6	CHIMERA : a new code for standard sirens analyses	115
3.3	Best-localized observed GW events	126
3.3.1	The case of GW170817	126
3.3.2	The case of GW190814	127
3.4	The sample	129
3.4.1	Parent Galaxy Catalog	129
3.4.2	Sample of GW events	130
3.5	Results for O4 and O5-like scenarios	135
3.5.1	Full O4- and O5-like scenarios	135
3.5.2	Spectroscopic vs. photometric galaxy catalog	139
3.5.3	Spectral-only analysis	140
3.5.4	Hyperparameters correlations	142
3.5.5	Assessing potential biases	143
3.6	Summary	146
4	Conclusions	149
5	Future perspectives	155

A	Appendix	159
A.1	Validation of PyLick	159
A.2	Validation of CHIMERA	163
A.2.1	Validation of the spectral siren mode using external code	163
A.2.2	Tests on one-dimensional parameter space	164
A.3	Full corners for the O4 and O5 standard sirens analyses	167

List of Figures

1.1	Evolution of the cosmological fluids	13
1.2	Current tension on H_0	17
1.3	SN Ia: Hubble diagram and $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane constraints	20
1.4	BAO: $H(z)$ and $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane constraints	22
1.5	CMB: Power Spectrum observed by Planck	23
1.6	Redshift coverage of cosmological probes	25
1.7	Current cosmological constraints from emerging probes	27
2.1	CC: Illustration of various methods to constrain galaxy's ages	36
2.2	CC: Key observational properties of passive galaxies in LEGA-C	40
2.3	CC: Median composite spectrum of the passive sample	42
2.4	CC: Example of the measurement of spectral features with PyLick	47
2.5	CC: Example of the measurement of the H/K feature	48
2.6	CC: Distribution of key spectral features	49
2.7	CC: Redshift evolution of key spectral features	50
2.8	CC: Example of MCMC constraints	55
2.9	CC: Four diagnostic diagrams color-coded by the H/K feature	60
2.10	CC: Scaling relations of individual galaxies	61
2.11	CC: Formation redshifts as a function of galaxy stellar mass	63
2.12	CC: Redshift trends of age, metallicity, and α -enhancement	64
2.13	CC: Comparison of the results using two different models	68
2.14	CC: Example of the age–star formation timescale degeneracy	69
2.15	CC: Derivation of $H(z)$	72
2.16	CC: Results obtained by varying the set of spectral features	73
2.17	CC: Theoretical age–redshift relations	75
2.18	CC: Constraints from age–redshift relations	77
2.19	CC: Median binned age– z relations using two methodologies	79
2.20	CC: Final constraints on $H(z)$ from cosmic chronometers	82
3.1	GW: Effect on test masses	86
3.2	GW: Coordinate systems	88
3.3	GW: Scheme of an interferometer	92
3.4	GW: Observation timeline	94
3.5	GW: Characteristic strain of astrophysical sources and detectors	96
3.6	GW: Mass function models	106
3.7	GW: Rate evolution models	108

3.8	GW: Detection probability model	113
3.9	GW: Workflow of CHIMERA	116
3.10	GW: 3D representation of the workflow of CHIMERA	117
3.11	GW: CHIMERA mask completeness	119
3.12	GW: CHIMERA interpolation of GLADE+	122
3.13	GW: CHIMERA interpolation of MICEv2	123
3.14	GW: CHIMERA likelihood computation	124
3.15	GW: CHIMERA 1D selection biases	125
3.16	GW: Analysis of GW170817	127
3.17	GW: Analysis of GW190814	128
3.18	GW: Properties of the O4- and O5-like catalogs	134
3.19	GW: Percentage errors on parameters for O4 and O5 analyses	137
3.20	GW: Corner plot for O4 and O5 with spectroscopic catalog	138
3.21	GW: Predictive posterior distribution of the mass functions	139
3.22	GW: Marginalized 1D posteriors for O4 and O5 analyses	140
3.23	GW: The $H_0 - \mu_g$ degeneracy	141
3.24	GW: Correlations among the population parameters	142
3.25	GW: Constraints on H_0 for multiple O5 realizations	143
3.26	GW: P-P plot and bias for multiple O5 realizations	144
3.27	GW: Final constraints on H_0 with BBH as standard sirens	148
4.1	CC & GW: Current and future constraints on $H(z)$ obtained in this Thesis	153
A.1	App: Validation of PyLick with LEGA-C data	160
A.2	App: Validation of CHIMERA with MGCosmoPop	163
A.3	App: 1D posteriors of O4 vs. O5	164
A.4	App: 1D posteriors for O5 varying the KDE bandwidth	165
A.5	App: Redshift reshuffling inconsistency	166
A.6	App: Full O4 and O5 corner plot including spectroscopic catalog	168
A.7	App: Full O4 and O5 corner plot including photometric catalog	169
A.8	App: Full O4 corner plot including spectroscopic or photometric catalog	170
A.9	App: Full O5 corner plot including spectroscopic or photometric catalog	171

Introduction

The discovery of the accelerated expansion of the Universe (Riess et al., 1998; Perlmutter et al., 1999) has been a major revolution in modern cosmology. This phenomenon is currently explained by invoking a new unknown form of energy – *dark energy* – which accounts for about 70% of the total matter-energy budget of the Universe. Similarly, the formation of large-scale structures and the dynamics of astrophysical objects require a new form of matter which interacts very weakly with standard matter particles – *dark matter* – constituting about 25%. Dark energy and dark matter are two fundamental assumptions of the standard Λ CDM model, which provides the simplest, yet most effective, description of a wealth of observations.

However, current tensions in cosmology, in particular the emerging discrepancy between the Hubble constant H_0 measured in the local Universe with type-Ia supernovae calibrated with Cepheids (Riess et al., 2022) and the value inferred from the CMB assuming a Λ CDM model (Planck Collaboration, 2020), hint at new physics at play beyond the standard cosmological model. In this context, the reconstruction of the expansion history of the Universe $H(z)$ using new and independent cosmological probes can play a crucial role (see Abdalla et al., 2022; Moresco et al., 2022). Among these, cosmic chronometers and gravitational waves are particularly worth exploring, as they can provide cosmology independent estimates of $H(z)$ and H_0 , respectively.

Cosmic Chronometers (CCs) have proven to be very promising probes to obtain direct measurements of $H(z)$ up to $z \sim 2$. This method, first introduced by Jimenez & Loeb (2002), consists in using massive and passive galaxies as tracers of the aging of the Universe dt_U as a function of z under the minimal assumption of an FLRW metric, $H(z) = -(1+z)^{-1} dz/dt_U$. Given an accurate selection of CCs, their aging dt , can replace dt_U in the equation above. The redshifts can be precisely measured (up to 0.1% accuracy from spectroscopic observations), but age dating galaxies is challenged by the complex reconstruction of their star formation history and the inherent degeneracies within stellar population parameters, such as stellar age, formation timescale, and chemical composition (Moresco et al., 2020). It is therefore crucial to select the purest samples of CCs, measure their ages in a cosmology independent way (e.g., no priors based on the age of the Universe), and thoroughly evaluate potential systematics.

Gravitational Waves (GWs) produced by compact binaries can be used as *standard sirens*, since they provide the luminosity distance d_L to the source without requiring any additional calibration beyond general relativity. Unfortunately, determining the redshift z with current GW data alone is not possible due to its inherent degeneracy with binary masses, and external information is required to obtain cosmological constraints. A first possibility is to observe an electromagnetic counterpart (*bright sirens*; Schutz, 1986; Holz & Hughes, 2005). However,

the detection probability decreases with increasing z , and most of the events observed so far are binary black holes (BBHs), from which such a signal is not expected. Specifically, after the first three observing runs (O1, O2, O3), the catalog of confident GW detections released by the LIGO-Virgo-Kagra (LVK) Collaboration contains about 90 events, mostly BBH, two binary neutron star (BNS), and two neutron star-black hole (NSBH) mergers (Abbott et al., 2023b). For this reason, a statistical approach must be used to infer the redshift information. In particular, one can either correlate GWs with galaxy catalogs (*dark sirens*; Schutz, 1986; Del Pozzo, 2012; Chen et al., 2018), or use the presence of features in the astrophysical source-frame mass distribution as well as tidal effects in neutron star mergers to break the mass-redshift degeneracy (*spectral sirens*; e.g. Chernoff & Finn, 1993; Taylor et al., 2012; Messenger & Read, 2012).

In this Thesis, we work toward unveiling the expansion history of the Universe using CCs and GWs, with the ultimate goal of optimizing their scientific yield. The main focus for both GWs & CCs is the interplay of cosmological and astrophysical parameters to derive them together, study possible degeneracies, and eventually minimize possible systematic effects.

In Borghi et al. (2022a) and Borghi et al. (2022b), we study the stellar population properties of a CC sample selected from the LEGA-C survey and derive a new measurement of $H(z = 0.75) = 98.8 \pm 33.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$. This result is obtained after a careful study of the ages of 140 individual passive galaxies using calibrated spectral features (Lick indices), which were measured with PyLick, a software that we publicly release.¹ In a subsequent analysis of the same CC sample, we estimate galaxy ages using an alternative method, full-spectrum fitting, which allows us to better assess the impact of a longer star formation history on the final $H(z)$ value, obtaining an excellent agreement with our previous results (Jiao et al., 2023). The same methodology is applied to a separate CC sample selected from the VANDELs survey, yielding a new value of $H(z)$ at $z \sim 1.26$ and demonstrating the validity of the CC approach up to $z \gtrsim 1$ (Tomasetti et al., 2023).

In Borghi et al. (2024), we study a novel methodology to jointly constrain cosmological and astrophysical population parameters using GWs and galaxy catalogs. We also release a new software, CHIMERA², to accurately perform standard siren analyses across different regimes. These regimes encompass cases where the EM redshift information is highly constraining (bright sirens), more loosely constraining (dark sirens), or absent (spectral sirens). Substantial efforts have been dedicated to prioritize computational efficiency to align with the demands of next-generation GW observatories (e.g., Einstein Telescope Punturo et al. 2010, LISA Amaro-Seoane et al. 2017, Cosmic Explorer Reitze et al. 2019) and upcoming galaxy surveys and facilities (e.g., Euclid Laureijs et al. 2011, Vera Rubin Observatory Ivezić et al. 2019, Nancy Grace Roman Space Telescope Spergel et al. 2015).

By combining results from GWs and CCs we provide an independent reconstruction of the cosmic expansion history with direct measurements of H_0 from GWs and $H(z)$ up to $z \sim 2$ from CCs. We also present forecasts for upcoming GW detectors and EM telescopes, demonstrating the compelling potential of GWs and CCs to unveil the expansion history of the Universe.

¹Documentation available at: <https://pylick.readthedocs.io>

²Documentation available at: <https://chimera-gw.readthedocs.io/>

The main stages of this Thesis are outlined below.

- **Chapter 1** lays the mathematical and physical foundations of modern cosmology, starting with general relativity and introducing the Friedmann equations, which govern the evolution of a homogeneous, isotropic universe. It discusses the standard cosmological model and the tensions motivating extensions to it. This chapter reviews established probes, such as supernovae, baryon acoustic oscillations, and the cosmic microwave background, as well as emerging ones, including CCs and GWs. Finally, it presents the main open questions in cosmology and how this Thesis aims to tackle them.
- **Chapter 2** describes the CC method and the main results of this Thesis. First, it presents the selection of a pure CC sample at intermediate redshift, illustrating the survey from which it is selected and the criteria adopted to minimize the presence of outliers. Subsequently, it presents `PyLick`, a new Python tool developed in this work, used to measure specific spectral features on galaxy spectra. The methodology employed to constrain the physical properties of each galaxy is then described. Finally, this chapter presents the astrophysical and cosmological results, including a new measurement of $H(z)$ at $z \sim 0.7$.
- **Chapter 3** describes the GW standard sirens method and the main results of this Thesis. First, it introduces the GW theory, detectors, sources, and state-of-the-art standard sirens techniques for GW cosmology. Then, it presents a novel methodology developed in this Thesis to obtain joint cosmological and astrophysical constraints with GWs and galaxy catalogs. This pipeline is included in `CHIMERA`, a new code developed within this work that is publicly released. The chapter then presents the analysis of two well-localized GW events, GW170817 and GW190814, the former being the event that better constrains H_0 . Subsequently, it describes the generation of the mock galaxy catalog and the extraction and simulation of GW events from it. Finally, it presents forecasts for upcoming GW detectors and galaxy surveys.
- To conclude, **Chapter 4** provides a comprehensive summary of the results of this Thesis, presenting new constraints on $H(z)$. It then discusses the path for future work, where the acquired knowledge and tools will be applied to current and upcoming galaxy surveys and GW facilities to unveil the expansion history of the Universe.

An expanding Universe

Cosmology traces its modern roots to [Einstein \(1915\)](#) theory of general relativity. The concept of an expanding Universe, a cornerstone of modern cosmology, was established by the pioneering observations of [Hubble \(1929\)](#) and theoretical work of [Lemaître \(1927\)](#), providing the formulation of a linear relationship between the speed at which galaxies recede and their distance from Earth. This aligned with [Friedmann \(1922\)](#) solutions to Einstein's field equations, predicting a dynamic evolution of the Universe, in this case, a never-ending expansion. Extrapolating this expansion backward, leads to an epoch – the *Big Bang* – when all matter and energy in the Universe was in a much hotter and denser state. A few years later, [Zwicky \(1933\)](#) realized that a significant amount of *dark matter* had to exist to reconcile the observed motions of galaxies within the Coma Cluster with the estimated quantity of luminous matter. The existence of a matter that does not interact with standard matter but with gravitation is backed by various observations, such as the rotation curve of galaxies ([Rubin & Ford, 1970](#)), X-ray emission in galaxy clusters, gravitational lensing, mass-luminosity ratio in galaxies, and large scale structure, yet its nature remains unknown.

In the mid-twentieth century, the Big Bang paradigm was the subject of growing interest in the scientific community due to two precise predictions. The first is the Big Bang Nucleosynthesis (BBN) of primordial elements (H, ^4He , D, ^3He and traces of Li, Be, and B), introduced in the famous $\alpha\beta\gamma$ paper ([Alpher et al., 1948](#)), that remains one of the earliest and most reliable probe of the primordial Universe. The second comes from ([Dicke et al., 1965](#)), predicting the existence of a thermal bath of photons coming from *last scattering* surface, which represents the set of points in space-time where matter and radiation, initially coupled due to high temperatures, gradually decoupled. It was serendipitously discovered by [Penzias & Wilson \(1965\)](#) while trying to find the cause of excessive noise measured on a telecommunications antenna. The first precise measurement of its black-body spectrum has been carried out with the COBE satellite, determining it to be $T_0 = 2.726$ K ([Fixsen, 2009](#)). Another significant achievement with COBE was the observation of CMB anisotropies at the level of $\delta T/T \approx 10^{-5}$ ([Smoot et al., 1992](#)). These fluctuations, which serve as the seeds for the large-scale structures observed in the present-day Universe, have been subsequently measured with remarkable precision by the WMAP ([Hinshaw et al., 2013](#)) and Planck ([Planck Collaboration, 2016](#)) missions.

By the late 1990s, two independent studies utilizing Type Ia supernovae (SNIa) provided evidence that the universe is undergoing accelerated expansion ([Riess et al., 1998](#); [Perlmutter et al., 1999](#)). This revolutionary discovery implies the presence of some form of dark energy driving the acceleration, such as a cosmological constant Λ or possible modifications of general relativity. A decade later, the observations of the large-scale structure, in particular

galaxy redshift surveys, provided new insights. The Sloan Digital Sky Survey (SDSS), became the first survey to detect baryon acoustic oscillations (BAOs) in the galaxy distribution (Eisenstein et al., 2005), a signature of primordial sound waves imprinted in the large-scale structure.

Since the first detection of the accelerated expansion of the Universe, several alternative cosmological probes have been proposed and studied. A key goal has been to precisely measure the expansion rate of the Universe today, quantified by the Hubble constant H_0 . However, as the data sets have become more constraining, tensions have emerged between different measurements of H_0 (Verde et al., 2019). Most notably, the value of H_0 inferred from CMB observations within the concordance cosmological model disagrees with the measurement of H_0 from SNIa. In general, this discrepancy arises between several early- and late-universe probes (Abdalla et al., 2022) and current debate, driving an ongoing debate in cosmology.

Recently, the groundbreaking discovery of gravitational waves (GWs) by the LIGO collaboration (Abbott et al., 2016b) has opened a new era in cosmology and has established the era of multi-messenger astrophysics thanks to the first coincident detection of the GW and electromagnetic signal with the GW170817 and GRB170817A events (Abbott et al., 2017a).

1.1 Fundamentals of General Relativity

In modern physics, General Relativity (GR) stands as the most reliable theory for describing gravity. The essence of GR is that gravity admits a geometrical interpretation. The geometry of space-time is linked by its matter and energy content through Einstein's field equations (Einstein, 1915, 1916):

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} , \quad (1.1)$$

where the left-hand side is a measure of the space-time geometry, with the curvature tensor $R_{\mu\nu}$, the metric tensor $g_{\mu\nu}$, and the scalar curvature R , while the right-hand side represents the distribution of energy and momentum with the energy-momentum tensor $T_{\mu\nu}$. This tensor is symmetric ($T^{\mu\nu} = T^{\nu\mu}$) and satisfies the continuity equations, $\nabla_{\mu}T^{\mu\nu} = 0$.

The space-time interval between two infinitesimally close events is defined as

$$ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu} , \quad \mu, \nu = (0, 1, 2, 3) , \quad (1.2)$$

where the line element ds^2 is invariant under inertial coordinate transformations, and $x^{\mu} = (ct, x_1, x_2, x_3)$ are the 1D (time) + 3D (space) coordinates. The signature convention for the metric $g_{\mu\nu}$ is $(-, +, +, +)$. For example, in the case of flat space, the metric is given by:

$$\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1) . \quad (1.3)$$

The shortest interval between two events is called *geodesic* and is found by minimizing Eq. (1.2), i.e. $\delta \int ds = 0$. In the absence of external non-gravitational forces, any particle moves along geodesics. The parametrized path $x^{\mu}(\tau)$ is given by the geodesic equation:

$$\frac{d^2x^{\mu}}{d\tau^2} + \Gamma_{\nu\rho}^{\mu}(x) \frac{dx^{\nu}}{d\tau} \frac{dx^{\rho}}{d\tau} = 0 , \quad (1.4)$$

where the Christoffel symbol $\Gamma_{\mu\nu}^{\rho} = \frac{1}{2}g^{\rho\sigma} (\partial_{\mu}g_{\sigma\nu} + \partial_{\nu}g_{\sigma\mu} - \partial_{\sigma}g_{\mu\nu})$

The Eq. (1.1) only admits expanding or contracting universe solutions. In 1917, Einstein (1917) introduced the cosmological constant Λ to obtain static solutions that matched the prevailing view at that time,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} . \quad (1.5)$$

The cosmological constant has now taken on a new role, as it is used to explain the observed accelerated expansion of the universe (Riess et al., 1998; Perlmutter et al., 1999). In particular, different interpretations are possible depending on its position in the above equation:

- Left-hand side: geometrical modification of gravity in the GR framework;
- Right-hand side: additional energy component, as an exotic fluid with negative pressure (see Section 1.2.4).

Alternative routes to explain the observations include the modification of GR, or the introduction of completely new theories of gravity.

1.2 Friedmann–Lemaître–Robertson–Walker cosmology

The Standard Cosmological Model rests upon two fundamental pillars, the cosmological principle¹ and the theory of general relativity, that are encapsulated into the *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric:

$$ds^2 = -c^2 dt^2 + a^2(t) \left[\frac{dr^2}{1 - \kappa r^2} + r^2 (\sin^2 \theta d\phi^2 + d\theta^2) \right], \quad (1.6)$$

where the *scale factor* $a(t)$ is determined by the Einstein equations, (t, r, θ, ϕ) are the *comoving coordinates*, and the *curvature parameter* k can take values $k = 0$ (spatially flat universe), $k > 0$ (closed universe), or $k < 0$ (open universe). If a test mass is at rest in the comoving frame, it will follow the so-called *Hubble flow* driven by the expansion of the Universe, but its comoving coordinates will remain unchanged.

1.2.1 Hubble-Lemaître law and definition of redshift

The *proper distance* d_p is defined as the distance between two objects that follow the Hubble flow evaluated instantaneously ($dt = 0$). Given homogeneity and isotropy, it is possible to set the first object at the origin and the second at comoving radial distance r without loss of generality and Eq. (1.6) gives

$$d_p(t) \equiv a(t) \int_0^r \frac{dr}{\sqrt{1 - \kappa r^2}}. \quad (1.7)$$

It follows that the proper distance between two objects following the Hubble flow depends on cosmic time through the relation

$$d_p(t) = \frac{a(t)}{a(t_0)} d_p(t_0), \quad (1.8)$$

valid for any arbitrary reference time t_0 . From here on, the subscript $_0$ refers to quantities evaluated today, and we follow the common practice of setting $a(t_0) = 1$. The radial velocity between these objects can be computed as the time derivative of Eq. (1.8). This gives

$$v_p(t) = \frac{dd_p}{dt} = H(t) d_p, \quad (1.9)$$

where the *Hubble parameter* is defined as

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)}, \quad (1.10)$$

with units of $[s^{-1}]$, but is usually expressed in more observationally convenient units of $[km s^{-1} Mpc^{-1}]$. When evaluated in today's Universe, Eq. (1.10) is the Hubble-Lemaître law

¹The cosmological principle asserts that the Universe appears the same, in terms of its physical properties, to observers regardless of their position or direction. This approximation is valid on sufficiently large scales (> 1 Gpc).

(Lemaître, 1927; Hubble, 1929). The parameter $H_0 \equiv H(t_0)$ is called *Hubble constant*, referring to the fact that *today, every point* of the Universe has the same value of H_0 . It is often expressed in dimensionless form as $h \equiv H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

The *redshift* is defined as the relative difference between the wavelength λ_{em} of a signal emitted by a distant source and the λ_0 received today by an observer,

$$z \equiv \frac{\lambda_0 - \lambda_{\text{em}}}{\lambda_{\text{em}}}. \quad (1.11)$$

or analogously for frequencies, using the relation $\nu = c/\lambda$. This quantity is positive when $\lambda_0 > \lambda_{\text{em}}$, in the opposite case it takes the name of *blueshift*. Consider a source at comoving distance r that emits two signals at t_{em} and $t_{\text{em}} + \delta t_{\text{em}}$ which travel at the speed of light and are later received by an observer located at $r = 0$ at t_0 and $t_0 + \delta t_0$. Since the source is at fixed r and the signals follow the null geodesic ($ds = 0$), Eq. (1.6) gives

$$\int_{t_{\text{em}}}^{t_0} \frac{c dt}{a(t)} = \int_0^r \frac{dr}{\sqrt{1 - kr^2}} = \int_{t_{\text{em}} + \delta t_{\text{em}}}^{t_0 + \delta t_0} \frac{c dt}{a(t)}. \quad (1.12)$$

Then by taking the difference between the left and the right terms, to linear order in δt_{em} , one finds $\delta t_0/a(t_0) = \delta t_{\text{em}}/a(t_{\text{em}})$. Therefore, since $\delta t = 1/\nu$ and $\lambda = c/\nu$, for an observer located at the present time and a generic source at t , we obtain

$$1 + z = \frac{1}{a(t)}, \quad (1.13)$$

with the convention $a(t_0) = 1$, and

$$dt_{\text{obs}} = (1 + z) dt. \quad (1.14)$$

Therefore the redshift can be also thought of as a measure of how much the scale factor has changed since the signal was emitted. By taking the time derivative of Eq. (1.13), it is possible to relate time intervals dt and redshift intervals dz as follows

$$dt = -\frac{dz}{(1+z)H(z)}. \quad (1.15)$$

To close the overview on the kinematic aspects of the FLRW metric, we introduce the *deceleration parameter*. By performing a Taylor expansion of $a(t)$ around the present epoch $t = t_0$, we get

$$\frac{a(t)}{a_0} = 1 + H_0(t - t_0) - \frac{1}{2}q_0 H_0^2(t - t_0)^2 + \dots \quad (1.16)$$

where the *deceleration parameter* is

$$q \equiv -\frac{\ddot{a}a}{\dot{a}^2}, \quad (1.17)$$

negatively defined because historical models favored a decelerated expansion of the Universe. To date, the measured value of h_0 is close to 0.7, with differences at the level of a few percent, while q_0 is between -0.5 and -0.7 (e.g., Abdalla et al., 2022).

1.2.2 Distances and geometries

Observational cosmology is challenged by the fact that we live in an expanding Universe. Here the concept of “distance” is not straightforward and several notions may be introduced (see Hogg, 1999).

The *comoving distance* between us ($z = 0$) and an object at redshift z_{em} is defined as the light travel time between us and the object. For a spatially flat universe ($\kappa = 0$), Eqs. (1.6) and (1.15) give

$$d_C(z_{\text{em}}) = \int_{t_{\text{em}}}^{t_0} \frac{c dt}{a(t)} = \int_0^{z_{\text{em}}} \frac{c dz}{H(z)}. \quad (1.18)$$

By its nature, the comoving distance remains constant in cosmological time and is linked to the proper distance Eq. (1.7) through the scale factor, $d_p(t) = a(t) d_C$. When the integral is evaluated between $t_{\text{em}} = 0$ and a generic time t , Eq. (1.18) becomes the *comoving horizon* and represents the maximum comoving distance that a photon can travel since the *Big Bang*. Regions at larger distances are not in causal connection. The general form of Eq. (1.18) in case of nonzero curvature (see Section 1.2.3) can be written as

$$d_C(z_e) = \lim_{\Omega'_\kappa \rightarrow \Omega_\kappa} \frac{c}{H_0 \sqrt{\Omega'_\kappa}} \sinh \left(\sqrt{\Omega'_\kappa} \int_0^{z_e} \frac{H_0}{H(z)} dz \right). \quad (1.19)$$

The *luminosity distance* is defined as the distance at which a source with known luminosity L produces the observed flux $f = L/(4\pi d_L^2)$. The expansion leads to time dilation, cosmological redshift, and variation of the geodesic, resulting in the following relation:

$$d_L(z) = (1 + z) d_C(z). \quad (1.20)$$

The *angular (diameter) distance* is defined as the distance at which a source with known diameter D is seen with an angular separation $\delta\theta = D/d_A$, and it can be shown that

$$d_A(z) = \frac{1}{1 + z} d_C(z). \quad (1.21)$$

Finally, the *comoving volume* element dV_C , given the solid angle $d\Omega$ and depth dz , is defined as follows:

$$dV_C = d_C^2 d\Omega dd_C = \frac{c^3}{H(z)} \left[\int_0^z \frac{dz}{H(z)} \right]^2 dz d\Omega. \quad (1.22)$$

1.2.3 Friedmann equations

The time evolution of the scale factor $a(t)$ is governed by GR and thus depends on the energy-matter content of the Universe. At the level of background evolution, the energy-momentum tensor is taken to be that of a perfect fluid, $T_{\mu\nu} = \text{diag}(\rho, p, p, p)$, with energy density ρ and pressure p . Plugging this into Eq. (1.1) and using Eq. (1.6), we obtain the **Friedmann (1922)**

equations

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{\kappa c^2}{a^2} = \frac{8\pi G}{3}\rho, \quad (1.23)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right). \quad (1.24)$$

A third useful, but not independent equation derives from the energy-momentum conservation $\nabla_\mu T^{\mu\nu}$ and reads

$$\dot{\rho} + \frac{3\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) = 0. \quad (1.25)$$

The pressure p and density ρ can be redefined to accommodate the effect of a possible “repulsive” component, necessary to ensure an accelerated expansion ($\ddot{a} > 0$, see Eq. (1.24)). A special case is obtained with the following substitutions

$$p \rightarrow p - \frac{\Lambda c^4}{8\pi G}; \quad \rho \rightarrow \rho + \frac{\Lambda c^2}{8\pi G}, \quad (1.26)$$

where Λ is called *cosmological constant* and can be interpreted as the energy density of the vacuum. Formally, this is equivalent as for Eq. (1.5).

Equation 1.23 can be rearranged as

$$H^2 \left(1 - \frac{\rho}{\rho_{\text{crit}}}\right) = -\frac{\kappa c^2}{a^2}. \quad (1.27)$$

where $\rho_{\text{crit}}(t) \equiv 3H^2(t)/8\pi G$ is the *critical density* for the Universe to be spatially flat. Now, the density of each i -th component can be expressed using the *density parameter* Ω_i , extensively used in literature

$$\Omega_i \equiv \frac{\rho_i}{\rho_{\text{crit}}}. \quad (1.28)$$

From equation 1.27 evaluated at t_0 , it can be seen that the total density parameter of today’s universe, $\Omega_{\text{tot},0} = \sum_i \Omega_{i,0}$, constrains the sign of κ (and therefore the geometry of the universe) at all cosmic times. In particular, for a spatially flat Universe ($\kappa = 0$), their sum must equal 1.

1.2.4 Fluids of the Universe

To solve the Friedmann equations, one needs to specify the equation of state of each component of the cosmic fluid, which is taken to be of the form

$$p(t) = w(t) \rho c^2, \quad (1.29)$$

where $w(t)$ is, in general, a function of time, and in ordinary physics it can assume values $0 \leq w < 1$. Combining Eq. (1.29) with Eq. (1.25), gives the evolution of ρ as a function of redshift,

$$\rho(z) = \rho_0 \exp \left\{ 3 \int_0^z \frac{1+w(z)}{1+z} dz \right\}, \quad (1.30)$$

where ρ_0 is the present value of the energy density. In a cosmological epoch when only one component dominates, $w(t) \simeq w$ can be approximated as constant. Then, the above equation reduces to

$$\rho(z) = \rho_0 (1+z)^{3(1+w)} = \rho_0 a^{-3(1+w)}. \quad (1.31)$$

Plugging these expression in Eq. (1.23) and integrating we find

$$a(t) \propto \begin{cases} t^{1/2} & (\text{r}), \\ t^{2/3} & (\text{m}), \\ e^{Ht} & (\Lambda). \end{cases} \quad (1.32)$$

where the subscripts (r), (m), and (Λ) refer to radiation ($w = 1/3$, e.g., photons and early-universe neutrinos), matter ($w = 0$, e.g., cold dark matter, baryonic matter, and non-relativistic neutrinos), and cosmological constant ($w = -1$, this is a special case in the family of dark energy models, in fact, the accelerated expansion can be produced for any $w < -1/3$). The evolution of these three components, as supported by current observations, is shown in Fig. 1.1.

Equation (1.23) can be rewritten in terms of z , separating the contributions of individual fluid components

$$H(z) = H_0 \sqrt{\Omega_{r,0}(1+z)^4 + \Omega_{m,0}(1+z)^3 + \Omega_{\text{DE},0}(1+z)^{3(1+w)} + \Omega_{\kappa,0}(1+z)^2}, \quad (1.33)$$

where $\Omega_{r,0}$, $\Omega_{m,0}$, and $\Omega_{\text{DE},0}$ are the density parameters of radiation, matter, and dark energy evaluated in the present-day Universe, while $\Omega_{\kappa} \equiv -\kappa c^2 / (Ha)^2$ is the curvature parameter for which $1 = \sum_i \Omega_i + \Omega_{\kappa}$. For a flat universe ($\Omega_{\kappa} = 0$) with a cosmological constant Λ , the above equation becomes

$$H(z) = H_0 \sqrt{\Omega_{r,0}(1+z)^4 + \Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}}, \quad (1.34)$$

which depends on two parameters only, since the flatness assumption implies $\Omega_{r,0} + \Omega_{m,0} + \Omega_{\Lambda,0} = 1$. At later cosmic epochs ($z \lesssim 10$), the relation can be further simplified as the radiation component becomes negligible (see Fig. 1.1), obtaining

$$H(z) = H_0 \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}}. \quad (1.35)$$

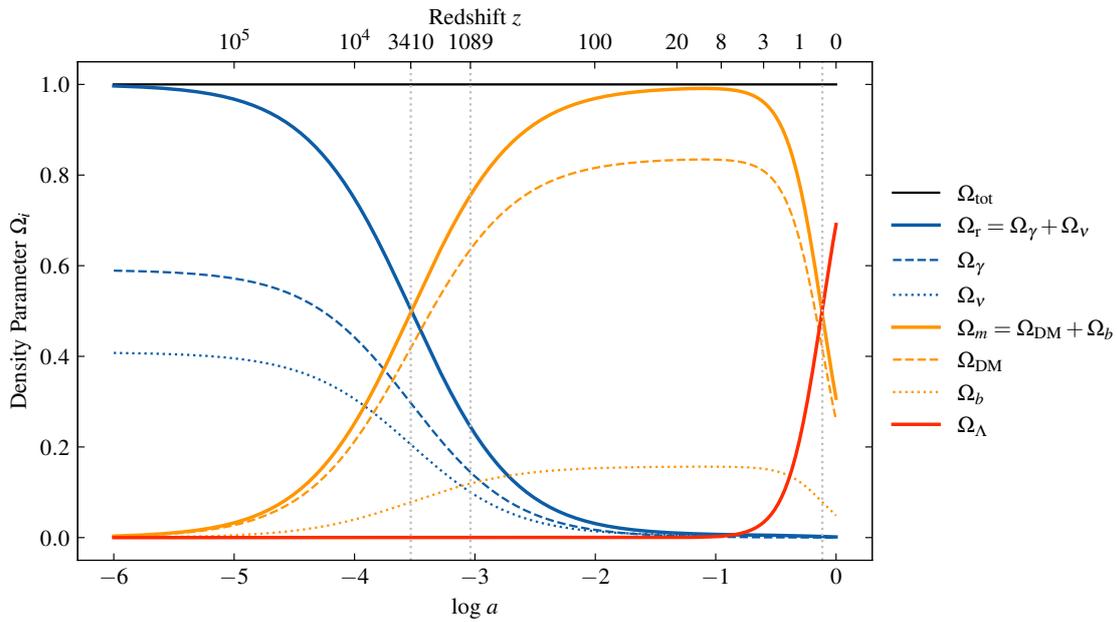


FIGURE 1.1: Evolution of density parameters Ω_i for radiation (r) in blue, matter (m) in orange, and cosmological constant (Λ) in red as a function of the scale factor a and z . Dashed lines show the contribution of photons (γ) and relativistic neutrinos (ν) to radiation, and dark matter (DM) and baryons (b) to matter. Vertical lines indicate three key epochs: radiation-matter equivalence ($z \approx 3410$), photon-baryon decoupling ($z \approx 1089$), and matter-cosmological constant equivalence ($z \approx 0.3$). The plot has been produced assuming a flat Λ CDM cosmology with parameters from [Planck Collaboration \(2020\)](#).

1.3 The Standard Cosmological Model

The cosmological model that best describes current observations is the flat Λ CDM model, with a cosmic fluid consisting of radiation, matter, and the cosmological constant. Specifically, the present-day energy-matter budget is dominated by dark energy ($\approx 70\%$), while the contributions from dark matter and baryonic matter are approximately 25% and 5%, respectively. The cosmological constant Λ is the simplest candidate for dark energy. Other possibilities are not excluded; in fact, the quality of observational data achievable today allows constraints to be placed on non-standard models that predict different forms of dark energy. Dark matter must be cold (hence the acronym CDM, Cold Dark Matter) to promote the structure formation in the early Universe and reproduce the observed distribution of matter. An example of warm dark matter particles are $m_\nu c^2 \lesssim 0.2$ eV neutrinos. Their velocity dispersion in the present-day Universe is too high to form gravitationally bound structures on galactic scales and thus cannot be the primary constituents of dark matter.

A primordial epoch of *inflation* (Guth, 1981; Linde, 1982; Albrecht et al., 1982) has been introduced to address certain issues of the standard model (see Section 1.3.1). It also represents one of the most plausible mechanisms for generating the initial density perturbations at the origin of the currently observed large-scale structure. The power spectrum of perturbations on the scale k is parameterized as a power law $\mathcal{P}(k) \equiv A_s (k/k_*)^{n_s-1}$, where A_s and n_s are the amplitude and spectral index of the primordial fluctuations, and k_* is an arbitrary reference scale.² The measurement of n_s places stringent constraints on inflation models, and current values indicate $n_s \approx 1$, meaning the fluctuations have (or nearly have) no preferred scales. Two other fundamental epochs predicted by the standard model are decoupling (see Section 1.4.3), and re-ionization.

Main ingredients of *vanilla* Λ CDM

The Λ CDM model in its minimal (or *vanilla*) form has six free parameters:

1. Baryon density: $\Omega_{b,0} h^2$;
2. Cold dark matter density: $\Omega_{c,0} h^2$;
3. Amplitude of the primordial power spectrum: A_s , conventionally evaluated at a scale $k_* = 0.05 \text{ Mpc}^{-1}$ and expressed as $\ln(10^{10} A_s)$;
4. Tilt of the primordial power spectrum: n_s , conventionally evaluated at a scale $k_* = 0.05 \text{ Mpc}^{-1}$;
5. Angular scale of the sound horizon at decoupling: θ_s ;
6. Optical depth associated with the epoch of reionization: $\tau(z_{\text{reion}})$.

The radiation density is well measured from the CMB temperature and is therefore fixed a priori to the FIRAS value of $T_0 = 2.726$ K (Fixsen, 2009). The other potential parameters are set to their “natural” values (e.g., $\Omega_{\text{tot}} = 1$, $w_{\text{DE}} = -1$). An often-analyzed extension of this model includes additional parameters such as the effective number of neutrino species N_{eff} and the sum of the neutrino masses $\sum m_\nu$, and/or the dark energy equation of state

²The power spectrum $\mathcal{P}(k)$ quantifies the excess power in an interval of width dk centered at k .

parameter w_{DE} . These models are frequently constrained by combining observations of the CMB with large-scale structure data. Finally, it should be noted that the value of the Hubble parameter H_0 , expressed here through h , and the various density parameters are determined indirectly.

1.3.1 Issues of the Standard Cosmological Model

As previously discussed, one of the major cosmological questions is understanding the nature of dark matter and dark energy. This issue can be addressed by searching for particles beyond the Standard Model of particle physics and/or formulating new theories of gravity on large scales, while simultaneously utilizing multiple independent probes to place stringent constraints on the rate of Universe expansion. Here, we provide a brief overview of the key challenges within the Standard Cosmological Model:

Origin (Big Bang). A quantum formulation of gravity is needed to describe the singularity at $t = 0$ and the first $\approx 10^{-43}$ s.

Flatness. To achieve $\Omega_{\text{tot}} \simeq 1$ in the present Universe (supported by various observational evidence, even to the third decimal place), the primordial Universe required $\Omega_{\text{tot}} = 1$ with a precision of 10^{-60} . This represents a clear case of *fine-tuning* of the model. The most accepted solution is the introduction of an inflationary epoch characterized by exponential expansion capable of generating this flatness.

Horizon. Arises from the challenge of explaining the observed homogeneity in the CMB on scales larger than the cosmological horizon (and thus not causally connected). Resolvable with the inflation paradigm.

Magnetic Monopoles. The Grand Unification Theory (GUT), used to describe the primordial Universe, predicts the existence of magnetic monopoles that should have a density comparable to that of baryonic matter today. Their absence can be justified by introducing an inflationary epoch.

Cosmological Constant. One of the immediate interpretations of the cosmological constant is, through quantum considerations, the zero-point energy of the vacuum. The difference between the vacuum energy density at the Planck scale and its measured value today is around 120 orders of magnitude, which requires a fine-tuning of the models. Moreover, the cosmological constant becomes non-negligible only in the recent Universe ($z \lesssim 0.3$) (e.g., see Fig. 1.1). If this had occurred slightly earlier, galaxies would not have formed, leading to the so-called *coincidence problem*. Currently, these two issues lack widely accepted solutions.

Small-Scale Dark Matter Anomalies. Numerical simulations of structure formation predict a distribution of dark matter on scales of ~ 1 kpc that differs from observations (e.g., mass profiles of halos, satellite structure counts). It is not yet clear whether these differences can be solely attributed to the effect of baryonic matter or if modifications to the physics of dark matter are necessary (see [Bullock & Boylan-Kolchin, 2017](#), for a review).

Large-Scale CMB Anomalies. Observations of the CMB at large angular scales have revealed several unexpected features with a typical 2 to 3 σ significance (e.g., [Planck Collaboration et al., 2020](#)). These include the lack of power on large angular scales, an unusually cold spot, alignment of low multipoles, and asymmetry in power between two hemispheres. While some models involving non-standard physics have been proposed, it is still possible that they may be caused by unknown systematics or random fluctuations in the data. Larger CMB data sets are needed to fully understand these anomalies.

S_8 Tension. The S_8 tension refers to the $\sim 2-3\sigma$ disagreement between the value of the parameter $S_8 \equiv \sigma_8(\Omega_m/0.3)^{1/2}$ inferred by the *Planck* CMB data [Planck Collaboration \(2020\)](#) and low redshift probes such as weak lensing and galaxy clustering measurements [Di Valentino et al. \(2021b\)](#); [Perivolaropoulos & Skara \(2022\)](#):

$$S_8 = (0.834 \pm 0.016) \quad (\textit{Planck} + \Lambda\text{CDM}) \quad (1.36)$$

$$S_8 \approx (0.76 - 0.78) \quad (\text{Lensing/clustering}) \quad (1.37)$$

H_0 Tension. As mentioned in the introduction to this Chapter, precision measurements of the Hubble constant H_0 are showing a 4 – 5 σ tension between local/late-time and CMB-related probes. An updated summary of these measurements is presented in [Fig. 1.2](#). An emblematic case is given by the disagreement between the value of the Hubble constant H_0 inferred by [Planck Collaboration \(2020\)](#) assuming a Λ CDM model (see [Section 1.4.3](#)) and the latest value measured by the SH0ES Collaboration [Riess et al. \(2022\)](#) based on supernovae calibrated by cepheids (see [Section 1.4.1](#))

$$H_0 = (67.27 \pm 0.60) \text{ km s}^{-1} \text{ Mpc}^{-1} \quad (\textit{Planck} + \Lambda\text{CDM}) \quad (1.38)$$

$$H_0 = (73.04 \pm 1.04) \text{ km s}^{-1} \text{ Mpc}^{-1} \quad (\text{SH0ES}) \quad (1.39)$$

The difference of $6.67 \text{ km s}^{-1} \text{ Mpc}^{-1}$ can be explained by extending the base Λ CDM model and modifying the physics of the primordial Universe (see [Abdalla et al., 2022](#), for an extensive review). An alternative explanation is that either [\(1.38\)](#) and/or [\(1.39\)](#) measurements are subject to unaccounted systematic effects. However, several comprehensive studies suggest that this hypothesis is currently unlikely. In light of this, it is essential to explore this tension with multiple probes.

1.3.2 Proposed solutions for the Hubble tension

To account for potentially discordant observations, a wide range of Λ CDM extensions have been proposed in the literature (see [Di Valentino et al., 2021a](#); [Abdalla et al., 2022](#); [Schöneberg et al., 2022](#), for a review). In this case, one assumes that the systematic effects are not the culprit of the tensions and invokes new physics at play during the expansion history of the Universe. The models can be broadly classified into two main categories.

Early-Universe proposals. One of the simplest extensions consists of adding extra relativistic degrees of freedom at recombination that can be captured by an increase in the effective neutrino number N_{eff} . Within Λ CDM, N_{eff} has a reference value of 3.044 ([Gariazzo et al., 2019](#)) for the three known neutrino families. From observations of

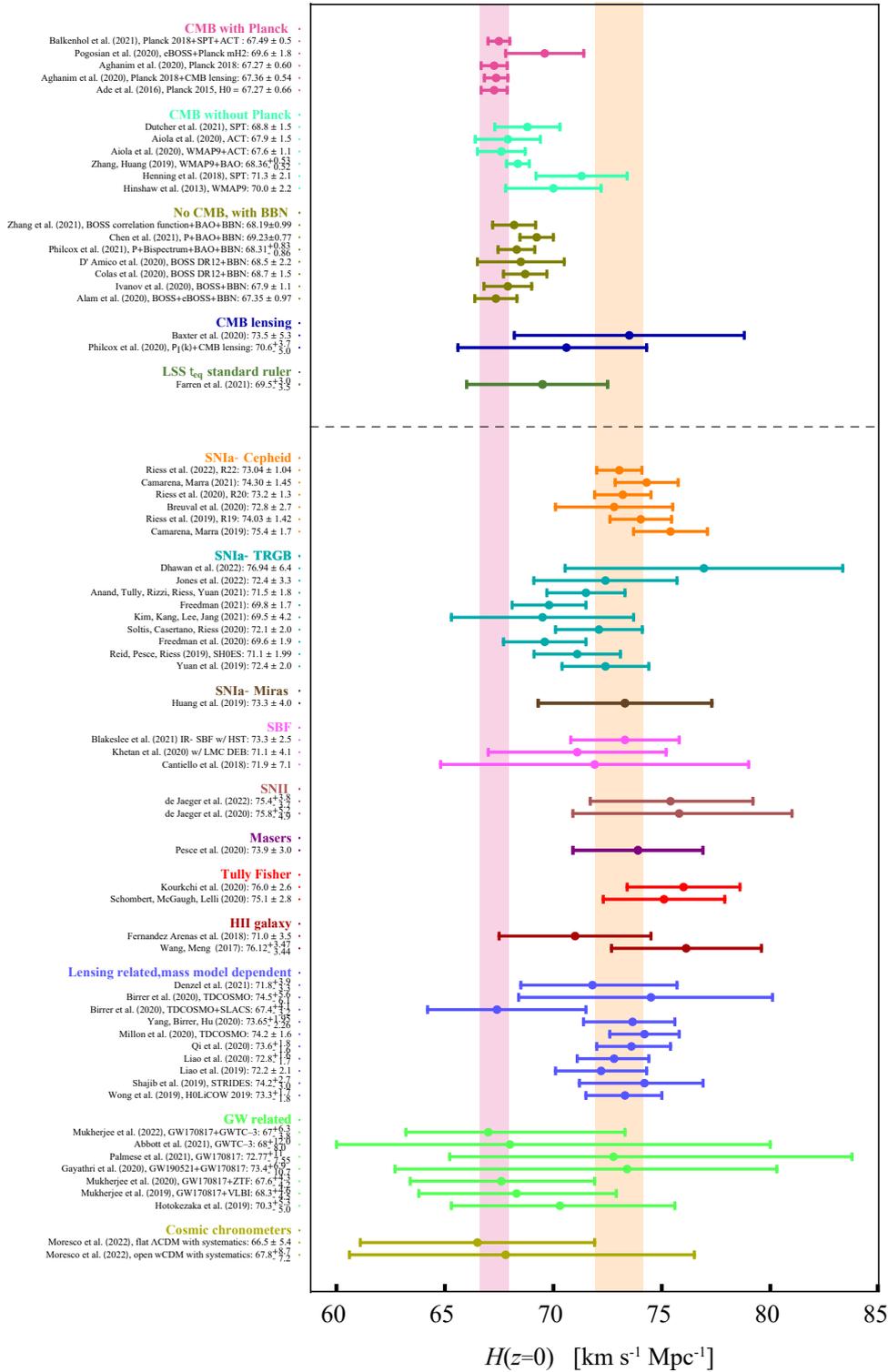


FIGURE 1.2: Current 68% CL constraints on H_0 from multiple cosmological probes. Adapted from Abdalla et al. (2022).

Table 1.1. Test of models against the (Planck + BAO + Pantheon) observed data sets. Adapted from [Schöneberg et al. \(2022\)](#).

Name	ΔN_{par}	Tension	ΔAIK	Podium
ΛCDM	0	4.4σ	0.0	
SIDR	1	3.2σ	-7.57	
Majoron	3	3.0σ	-9.49	
primordial B	1	3.5σ	-9.42	
varying m_e	1	2.9σ	-10.27	
varying $m_e + \Omega_k$	2	2.0σ	-13.26	
EDE	3	3.6σ	-15.98	
NEDE	3	3.1σ	-12.93	
EMG	3	3.7σ	-12.56	

CMB anisotropies ([Planck Collaboration, 2020](#)), there is a well-known degeneracy, where enhancing the radiation density at early times can lead to higher H_0 values, with $\Delta H_0 \simeq 5.9 \Delta N_{\text{eff}}$ ([Abdalla et al., 2022](#)). Therefore, a solution is to include extra relativistic dark relics (*dark radiation*), either streaming or self-interacting (SIDR). With future surveys such as CMB-S4 it will be possible to detect deviations of ΔN_{eff} within 0.06 (at 95% CL) corresponding to $\Delta H_0 \simeq 0.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Another proposal invokes the presence of a $\sim\text{eV}$ -scale majoron (e.g., [Escudero & Witte, 2020](#)), a weakly-coupled pseudo-Goldstone boson that could emerge from extensions to neutrino mass generation models. Alternatively, there are early-universe solutions that do not involve dark radiation; these include primordial magnetic fields (e.g., [Jedamzik & Pogosian, 2020](#)), varying fundamental constants like the electron mass m_e or the fine structure constant (e.g., [Franchino-Viñas & Mosquera, 2021](#)), varying m_e and assuming non-zero curvature Ω_k (e.g., [Sekiguchi & Takahashi, 2021](#)), a scalar field behaving as early dark energy (EDE [Karwal & Kamionkowski, 2016](#)), and alternative new dark energy (NEDE) consisting in first-order transition of a vacuum energy component near recombination (e.g., [Niedermann & Sloth, 2021](#)).

Late-Universe proposals In this case, the idea is to modify the cosmological expansion history at late times (i.e. well after recombination). In particular, they attempt to balance an increase of H_0 , hence an increase of the energy density today, with a decrease of the energy density at earlier times (such as through phantom dark energy $w < -1$). Proposals include the Chevallier-Polarski-Linder (CPL) parameterization of the dark energy equation of state ([Chevallier & Polarski, 2001](#); [Linder, 2003](#)), phenomenological emergent dark energy (e.g., [Li & Shafieloo, 2019](#)), a fraction of cold dark matter decaying into dark radiation (e.g., [Poulin et al., 2016](#)).

A systematic analysis comparing how different models are capable of explaining current observations has been presented in [Schöneberg et al. \(2022\)](#) and is shown in Table 1.1. In general, late-time solutions struggle to solve the tensions ([Schöneberg et al., 2022](#); [Abdalla et al., 2022](#)).

1.4 Observational Probes

Several probes have been proposed during the last two decades to measure the expansion history of the Universe, the growth of structures within, and to precisely constrain cosmological parameters, as this could give further hints in understanding the nature of dark energy and dark matter. This Section presents the main observational cosmological probes. The underlying method consists of using standard reference objects whose properties remain unchanged as a function of z , allowing us to probe the evolution. These properties are intrinsic luminosity (“standard candles”; e.g. supernovae, Section 1.4.1), size (“standard rulers”; e.g. BAO, Section 1.4.2), or intrinsic fluctuations in the cosmic microwave background (Section 1.4.3). In the end, Section 1.4.4 presents more novel and emerging cosmological probes to constrain the expansion history of the Universe.

1.4.1 Standard Candles (SNe)

The method of standard (or standardizable) candles relies on knowledge of the intrinsic luminosity of an observed object (see Eq. 1.20). The most emblematic example is provided by Type Ia supernovae (SN Ia), typically produced by a carbon-oxygen white dwarf that passes the Chandrasekhar mass limit ($\sim 1.4 M_{\odot}$). This event triggers a thermonuclear explosion during which radioactive elements (mainly ^{56}Ni) are synthesized. Their decay produces a light curve with a characteristic profile (peak, duration, and color) that makes them standardizable candles (Phillips, 1993; Riess et al., 1998; Perlmutter et al., 1999). However, at least two intermediate steps are required to derive the luminosity distance d_L , as the supernovae represent only the last step of the so-called *distance ladder*, which involves:

1. Measurement of the distance using geometric techniques (e.g., parallax, binary stars, masers) for a calibrator (e.g., Cepheids and the tip of the red giant branch);
2. Observation of galaxies hosting SNe Ia, sufficiently nearby ($z \lesssim 0.01$) to utilize the calibrator to measure their distance;
3. Observation of a large number of galaxies hosting SNe Ia at higher redshifts to minimize peculiar velocity contributions.

For $z \ll 1$, Eq. (1.20) reduces to the Hubble-Lemaître relation, $d_L(z) \approx cz/H_0$. Therefore, the Hubble constant H_0 can be measured directly. This method was used in the original work of Hubble (1929) and in more recent measurements of H_0 (e.g., Riess et al., 2021; Freedman et al., 2019). At higher redshifts, it is possible to constrain the other cosmological parameters, as first presented in Riess et al. (1998) and Perlmutter et al. (1999).

In general, the expected distance modulus for an SN at redshift z in a universe described by cosmological parameters $\lambda_c = \{H_0, \Omega_{m,0}, \Omega_{\Lambda,0}, \dots\}$ is defined as

$$\mu \equiv m - M = 5 \log \left(\frac{d_L(z, \lambda_c)}{10 \text{ pc}} \right), \quad (1.40)$$

where m and M are the observed and absolute magnitudes of the SN, respectively. From an observational point of view, μ can be estimated by parameterizing the light curve of each SN

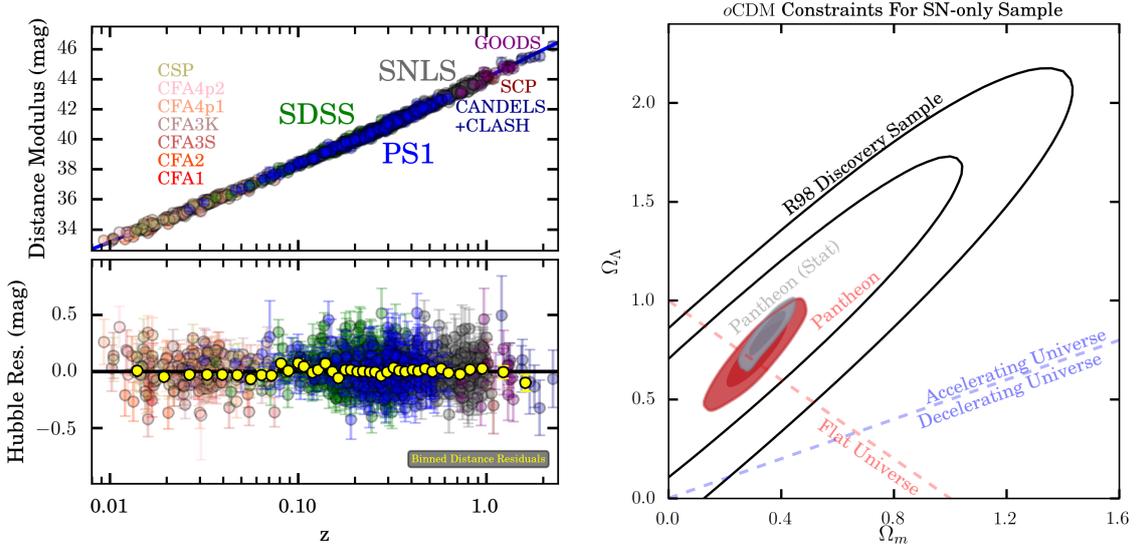


FIGURE 1.3: *Left*: Hubble diagram obtained with various samples of supernovae. The lower panel shows the residuals between the observational data and the best-fit Λ CDM model, with yellow points representing binned data; *Right*: 68% and 95% CL constraints to the $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane considering only statistical (grey) and both statistical and systematic uncertainties (red). Historical results from Riess et al. (1998) are shown in black. Reproduced from Scolnic et al. (2018).

as follows:

$$\mu = m_B^* + \alpha s - \beta c - M_B + \Delta_{host}, \quad (1.41)$$

where m_B^* is the observed magnitude of the peak in the rest-frame B band, s and c are the stretch and color of the observed SN, respectively, which are assumed to be related to the absolute magnitude through coefficients α and β . M_B is the absolute magnitude in the B band of a simulated SN with $s = 0$ and $c = 0$. Finally, Δ_{host} represents a correction factor, which depends on the host galaxy mass. The constraints on the cosmological parameters λ_c are then obtained by comparing 1.40 and 1.41.

Systematic errors are typically included as an additional contribution to the covariance matrix associated with the observed SNIa μ . The most significant contribution arises from the photometric calibration process (e.g., Conley et al., 2011), both for individual SNe and for different catalogs that are combined and used simultaneously. Other systematic effects may be due to selection effects, redshift evolution, light curve modeling, extinction due to interstellar dust, peculiar velocities (important at low z), and uncertainties in correlations with host galaxies. In Fig. 1.3 is shown an example of the standard candle method, presenting results obtained from the Pantheon sample (Scolnic et al., 2018).

The most recent cosmological constraints come from the Pantheon+ analysis of 1550 distinct SNe Ia located at $0.001 < z < 2.26$ (Brout et al., 2022). The authors find $\Omega_{m,0} = 0.334 \pm 0.018$ for a flat Λ CDM model and $w_0 = -0.90 \pm 0.14$ by relaxing the constant equation of state parameter of dark energy w_0 , which is compatible with the cosmological constant ($w_0 = -1$). By combining their results with CMB data from Planck Collaboration

(2020) and BAO measurements from the SDSS and CMASS surveys (Anderson et al., 2014; Ross et al., 2015), they obtain:

$$w_0 = -0.978_{-0.031}^{+0.024}, \quad (1.42)$$

which is one of the tightest constraints on dark energy and is consistent with the cosmological constant. This improvement is possible because the degeneracies between the parameters $\Omega_{m,0}$ and w_0 are lifted in the CBM + BAO + SNe combination, as they are nearly orthogonal.

Other astrophysical objects can also be used as standard candles (e.g., Type II supernovae, quasars, radio galaxies, and Gamma-Ray Bursts; see Section 1.4.4); however, they have not been thoroughly analyzed yet or provide less competitive cosmological constraints.

1.4.2 Standard Rulers (BAO)

In the early Universe, photons and baryons were tightly coupled, and density perturbations propagated as acoustic waves. These oscillations imprint a characteristic scale on the clustering of matter observed today, providing a standard cosmological *ruler* that can be measured in maps of large-scale structures. Their existence was predicted by Sunyaev & Zeldovich (1970) and Peebles & Yu (1970), but definitive detections were made possible only by the increased survey volume and number density of galaxies achieved in the SDSS and 2dF redshift surveys (Eisenstein et al., 2005; Cole et al., 2005; Percival et al., 2007). Observationally, this phenomenon leads to an excess probability of observing galaxies separated by a characteristic scale known as the BAO scale, which is set by the radius of the sound horizon at the drag epoch z_d when photons and baryons are decoupled (see also Eq. 1.18),

$$r_d = \int_0^{t_d} \frac{c_s(t) dt}{a(t)} = \int_{z_d}^{\infty} \frac{c_s(z) dz}{H(z)}, \quad (1.43)$$

where c_s is the speed of sound in the baryon-photon fluid. The precise length of this ruler depends on the baryon density and the total matter density: the former modifies the plasma inertia and hence c_s , the latter influences cosmic expansion, and thus t_* and the physical scale of BAO at different redshifts. The parameters used to calibrate r_d are typically derived with high precision from the morphology of the peaks in the CMB power spectrum 1.4.3.

In essence, BAOs can be used as *standard rulers* to determine d_A (Eq. 1.20) and, at the same time, the Hubble parameter $H(z)$ as a function of redshift. The first is obtained by measuring the *transverse modes*, which involve the angular separation $\delta\theta_s = r_d/d_A(z)$. The parameter $H(z)$ is obtained by measuring the *radial modes*, which involve the differences in the cosmological redshifts $\delta z_s = H(z) r_d/c$. The initial surveys had high statistical errors, allowing only for the estimation of the generalized distance $d_V^3(z) = (1+z)^2 d_A(z) cz/H(z)$, and the degeneracy between $d_A(z)$ and $H(z)$ prevented the direct measurement of the cosmic expansion rate. This changed with more recent surveys (e.g., Fig. 1.4a).

The advantage of this method lies in the detailed knowledge of the physics of acoustic oscillations and the possibility to use various tracers beyond galaxies, such as galaxy clusters, AGN, Ly α forest, and cosmic voids (see Bassett & Hlozek, 2010, for a review). To date, this method has been applied by observing galaxies up to $z \sim 2.2$ and the Ly- α forest in quasar spectra up to $z \sim 3.5$.

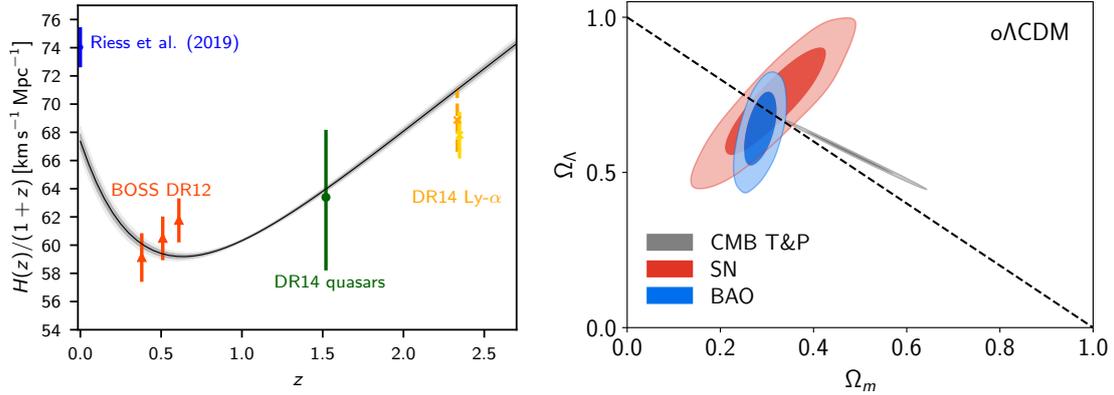


FIGURE 1.4: *Left*: Constraints to $H(z)$ using BAOs measurements from BOSS DR12 galaxies (red), BOSS DR14 quasars (green), and Ly α forest for higher- z BOSS DR14 quasars (yellow) (Alam et al., 2017; Zarrouk et al., 2018; de Sainte Agathe et al., 2019). For comparison, the constraints from Planck Collaboration (2020) data assuming the base Λ CDM model and Riess et al. (2019) SNe are shown in gray and blue, respectively. *Right*: 68% and 95% CL constraints to the $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane from the final eBOSS analysis (blue). For comparison, the contours obtained with Planck and SNe are shown in gray and red, respectively. Reproduced from Alam et al. (2021).

1.4.3 Cosmic Microwave Background (CMB)

In the primordial Universe, the temperature was sufficiently high to ensure thermal coupling between matter and radiation through scattering processes. The formation of neutral helium and hydrogen atoms only occurred when the temperature dropped to around 10^4 K due to expansion. This favored the process of recombination ($z \approx 1100$). When the plasma became devoid of electrons, the scattering process became inefficient, and photons decoupled from matter ($z_{dec} \simeq 1089$, Fig. 1.1). From the last-scattering surface, the cosmic microwave background radiation has been able to reach us almost unobstructed, providing us with an opportunity to get a snapshot of the primordial Universe.

The energy spectrum of the CMB is a black body, whose temperature today is measured with great precision as $T_0 = 2.726 \pm 0.001$ K (Fixsen, 2009). The weak fluctuations ($\delta T/T \sim 10^{-5}$) first observed by the COBE satellite (Smoot et al., 1992) indicate that the Universe is *approximately*, but not *perfectly*, homogeneous. It is within these anisotropies that the main cosmological information is contained. A quantity that provides a statistical measurement of these anisotropies is the angular power spectrum $\mathcal{C}(l)$, which is defined as follows:

$$\mathcal{C}(l) = \frac{1}{2l+1} \sum_{m=-l}^l |a_{lm}|^2, \quad (1.44)$$

where a_{lm} are the expansion coefficients of the spherical harmonics into which the sky is decomposed. Essentially, $\mathcal{C}(l)$ quantifies the power of fluctuations on the angular scale $\approx \pi/l$, where l is the multipole moment. Figure 1.5 shows the power spectrum measured by the Planck satellite.

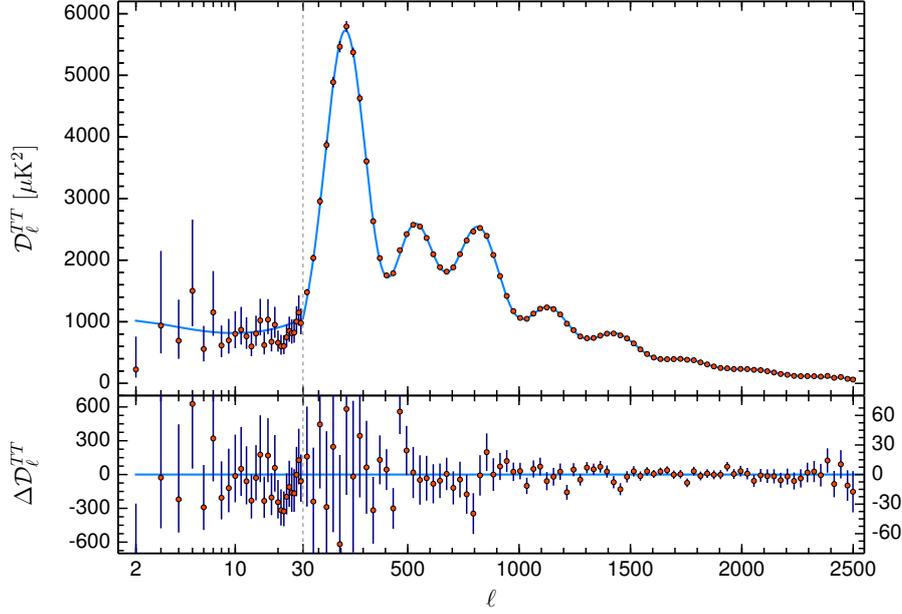


FIGURE 15: Cosmic Microwave Background (CMB) power spectrum observed by the Planck satellite (data points with $\pm 1\sigma$ error bars) compared to the ΛCDM model (blue line) and resulting residuals (*lower panel*). Note: $\mathcal{D}_l^{TT} \equiv l(l+1)\mathcal{C}(l)/(2\pi)$. Reproduced from [Planck Collaboration \(2020\)](#).

The shape of the power spectrum of the CMB is governed by various effects ([Lesgourgues et al., 2013](#)), which are typically described in terms of the six parameters presented in Section 1.3:

1. The position of the first peak depends on the angular scale of the sound horizon at decoupling $\theta_s = r_d(z_*)/\chi(z_*)$ (see 1.18 and 1.43). Its position is mainly sensitive to curvature and its shape is fixed by the density of baryons and the total matter density. Observational data strongly suggest that the Universe is spatially flat. Furthermore, the value of $\chi(z_*)$ at $z \lesssim 1$ depends on the dark energy density, which can then be constrained from the CMB.
2. The difference between the amplitude of odd and even peaks depends on Ω_b/Ω_γ , hence tracing Ω_b , as Ω_γ is fixed by the CMB temperature;
3. The amplitude of all peaks is governed by the expansion history between matter-radiation equality and decoupling that primarily depends on Ω_m (and thus Ω_c once Ω_b is known);
4. High- l (small-scale) anisotropies are damped due to diffusion effects, adding further dependence on Ω_b and Ω_m ;
5. The overall amplitude is proportional to the primordial amplitude A_s ;
6. The overall slope is proportional to the primordial slope that measures n_s ;

Table 1.2. Constraints on cosmological parameters for a base Λ CDM cosmology 1.3. The second column presents Planck measurements (including CMB and lensing), the third column include measurements from BAO (adapted from [Planck Collaboration, 2020](#)).

Parameter	Planck TT+TE+EE+lowE+lensing	+BAO
$\Omega_{b,0}h^2$	0.02237 ± 0.00015	0.02242 ± 0.00014
$\Omega_{c,0}h^2$	0.1200 ± 0.0012	0.1193 ± 0.0009
$100 \theta_s$	1.0409 ± 0.0003	1.0410 ± 0.0003
n_s	0.965 ± 0.004	0.966 ± 0.004
τ	0.054 ± 0.007	0.056 ± 0.007
$\ln(10^{10} A_s)$	3.044 ± 0.014	3.047 ± 0.014

7. Low- l (large-scale) slope also depends on the gravitational redshift effect between CMB and us observers (integrated Sachs–Wolfe effect) that depends on the expansion history and thus the dark energy density and h ;
8. The global amplitude at $l \ll 40$ and $l > 40$ depends on the optical depth at the epoch of reionization $\tau(z_{reion})$.

Table 1.2 lists the constraints on the six parameters of the base Λ CDM model obtained by [Planck Collaboration \(2020\)](#). It is important to note that the CMB can be measured with extreme precision and, although it cannot directly constrain $H(z)$ (which is why w_{DE} is fixed to -1 in the base model), the degeneracies among the cosmological parameters are often orthogonal to those from other probes. Therefore, by combining external data, it is possible to obtain very tight constraints on dark energy parameters (see (1.42)).

1.4.4 Emerging probes

Over the last two decades, extensive research has been done on the main probes discussed in the previous section, and it has become evident that relying on a single probe is not sufficient to obtain precise and accurate cosmological constraints ([Verde et al., 2019](#)). The combination of multiple probes allows us to better understand the systematics of the specific physical processes, which is at the base of each method, and untangle parameter degeneracies. The main emerging probes are described below (see [Moresco et al., 2022](#), for a comprehensive review). Figure 1.6 shows the redshift range mapped by each probe, while the current constraints are shown in Fig. 1.7.

Cosmic Chronometers (CC). Very massive and passively evolving galaxies can be used to directly measure the Hubble parameter $H(z)$ by measuring their aging $dt(z)$ (see Eq. 1.15, [Jimenez & Loeb 2002](#)). This method provides a direct determination of the expansion history of the Universe without any further cosmological assumption. These passive galaxies must be selected with strict criteria to minimize star-forming contaminants. Their aging is estimated

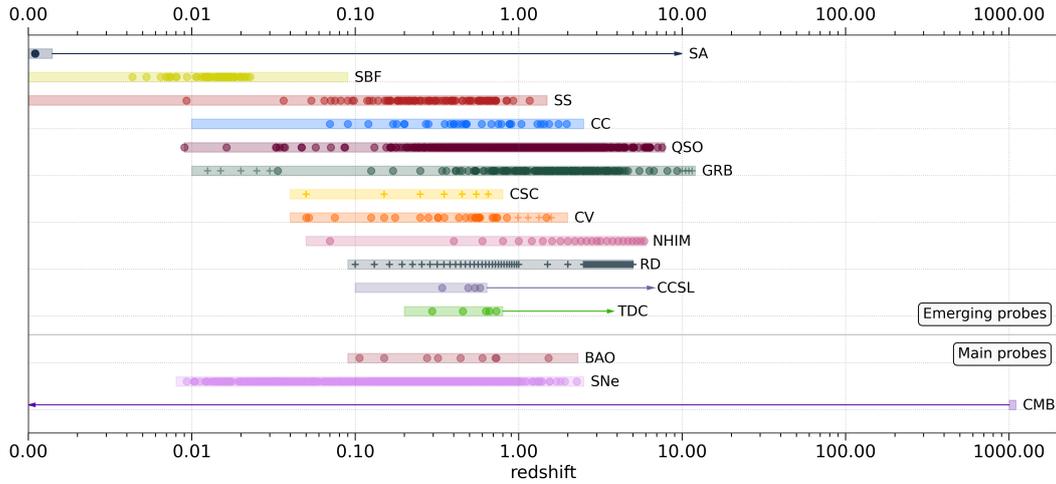


FIGURE 1.6: Redshift coverage of cosmological probes (see main text for acronyms). Horizontal bands show the redshift coverage with current (circle dots) and future (cross signs) measurements. Arrows indicate when a probe carries integrated information from higher redshift. Reproduced from [Moresco et al. \(2022\)](#).

from full-spectrum fitting or using specific spectral features. The CC method is studied in this Thesis and presented in more detail in Chapter 2.

Standard Sirens (SS). Merging binaries detected in gravitational waves directly measure d_L Eq. (1.20) via the strain amplitude without requiring a distance ladder as for SNe ([Schutz, 1986](#)). The peculiarity of GW cosmology is that determining the redshift with GW data alone is not possible because of its inherent degeneracy with binary masses. External information is required to provide cosmological constraints, including information on z from the EM counterpart, from galaxy catalogs, or from assumptions on the source frame population. The detection of GW170817 event and its EM counterpart has demonstrated the strength of this approach, providing a $\sim 15\%$ constraint on H_0 . The SS method is studied in this Thesis and is presented in more detail in Chapter 3.

Quasars (QSO). Quasars are among the most luminous objects in the Universe, reaching $L \approx 10^{44} - 10^{48} \text{ erg s}^{-1}$ in the x-ray-ultraviolet range. The non-linear relationship between L_{UV} and L_X makes them standardizable candles (e.g., [Watson et al., 2011](#); [Risaliti & Lusso, 2015](#)). The advantage is that, due to their luminosities, QSO can reach deeper z with respect to SNe. Current measurements reach up to $z \sim 7$.

Gamma-Ray Bursts (GRB). The objects are detectable at very high z thanks to the enormous energy they release in X/gamma rays (with isotropic radiated energy up to $E_{iso} \simeq 10^{54}$). Correlations between distance-dependent quantities and rest-frame observables allow us to standardize them as distance indicators ([Amati et al., 2008](#)). In particular, the $E_{p,i} - E_{iso}$ (“Amati”) correlation has been used to construct the Hubble diagram out to $z \sim 8$.

Time Delay Cosmography (TDC). The relative arrival time of multiple lensed signals can be used to measure the time-delay distance (Refsdal, 1964; Birrer et al., 2020; Wong et al., 2020). With the additional measurement of the source and lens redshifts, together with lens mass modeling, it is possible to constrain the cosmological parameters. Current measurements use a limited sample of ~ 10 lensed quasars.

Cosmography with Cluster Strong Lensing (CCSL). This method is similar to TDC but uses galaxy clusters, which produce more multiple images (10–100 with respect to 2–4, e.g., Postman et al. 2012) of background sources as the lensing cross section is much larger, but have more complex mass model parameterization. Current measurements include a limited sample of ~ 5 cluster lenses.

Cosmic Voids (CV). As large underdense regions, cosmic voids can trace dark energy and modified gravity without relying on baryonic physics. The most studied observables for cosmological applications are the void size function and the void density profile (e.g., Contarini et al., 2019; Verza et al., 2019).

Neutral Hydrogen Intensity Mapping (NHIM). Spatial maps of the redshifted 21cm HI signal provide large-scale structure information without requiring galaxy identification. Foreground mitigation techniques aim to recover the cosmological signal. Key challenges are foreground removal and instrumental calibration uncertainties.

Surface Brightness Fluctuations (SBF). Fluctuations in galaxy surface brightness, due to the granularity of resolved stars, decrease with respect to the average surface brightness at increasing distances. This information can be used to obtain cosmological constraints (e.g., Blakeslee et al., 2021).

Stellar Ages (SA). Age measurements of the oldest objects as old globular clusters can provide cosmological constraints through the look-back time relation, (e.g., Jimenez et al., 2019; Cimatti & Moresco, 2023). Improving these constraints requires high-resolution spectroscopic observations and more comprehensive stellar population modeling.

Secular Redshift Drift (RD). By observing the variations in the redshift of celestial objects over time, it is possible to directly probe the cosmic expansion (Sandage, 1962). The major challenge is that the measuring instrument must have a laser-metrology stability of 10^{-10} over years that is beyond the reach of current technology.

Clustering of Standard Candles (CSC). This method uses the clustering of standard candles, such as SNe Ia, around tracer galaxies to statistically obtain redshifts and improve distance estimates (e.g., Quartin et al., 2014; Zumalacárregui & Seljak, 2018).

The various emerging cosmological probes presented offer different and complementary ways to study the expansion history of the Universe. Each probe has its own specific redshift range (see Fig. 1.6), as well as its own strengths, weaknesses, and susceptibility to systematic

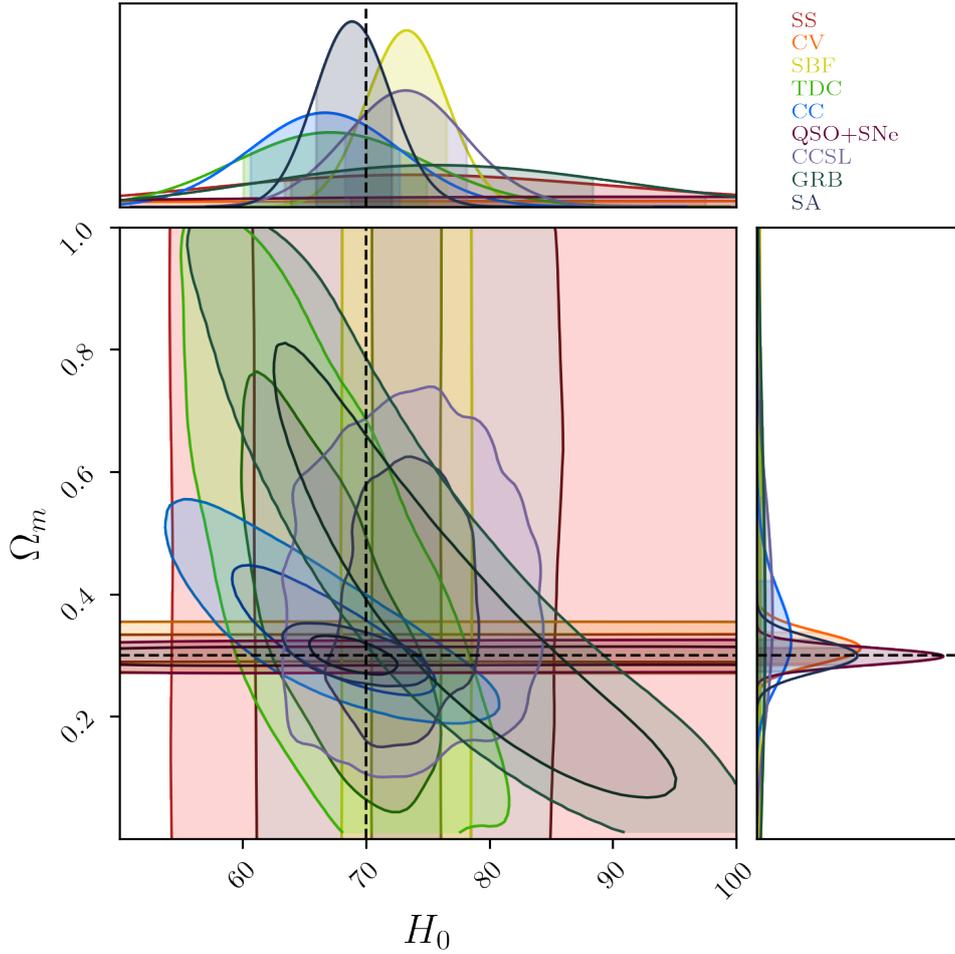


FIGURE 1.7: Current $H_0 - \Omega_{m,0}$ plane constraints from emerging cosmological probes (see text for acronyms and explanation). The dashed lines indicate a fiducial flat Λ CDM cosmology with $H_0 = 70$ km/s/Mpc and $\Omega_{m,0} = 0.3$.
Reproduced from [Moresco et al. \(2022\)](#).

uncertainties. Cosmic chronometers and stellar ages probe the differential age evolution of passively evolving galaxies to constrain the Hubble parameter and the cosmic time-redshift relation, respectively. Quasars and gamma-ray bursts take advantage of empirical correlations to serve as standardizable candles out to high redshift. Cluster strong lensing, cosmic voids, intensity mapping, and clustering of standard candles contain independent information about geometry, growth of structure, and baryon acoustic oscillations. Surface brightness fluctuations use galaxies themselves as standard rulers. Together with redshift drift, these probes provide multiple independent paths to measure cosmic acceleration and test cosmological models. The combination of these emerging probes with more established techniques like SNe, CMB, and spectroscopic galaxy surveys promises an exciting future for precision cosmology.

1.5 Open questions in Cosmology

As presented in previous sections, the discovery of the accelerating expansion of the Universe (Riess et al., 1998; Perlmutter et al., 1999) has led to significant changes in the understanding of the Universe’s dynamics, prompting considerable interest within the scientific community. In the last 30 years, progress has been remarkable: large networks of ground-based telescopes and dedicated space observatories have ushered in the era of precision cosmology. As always in science, many questions have found answers, yet in the process new and more profound ones have arisen.

Nature of dark energy. The accelerated expansion of the Universe can be explained by assuming that it is currently dominated by a dark energy component, consisting of around 70% of the total matter-energy budget of the Universe. The simplest theoretical explanation is a cosmological constant Λ , though this suffers from fine-tuning issues and its nature remains unknown. Alternative proposals for dynamical DE such as scalar fields or modifications of general relativity also aim to explain the observed cosmic acceleration, but more (and more precise) data are needed to make progress.

Nature of dark matter. Observations of large-scale structures and gravitational dynamics extending back to the 1930s provided evidence for the existence of a component of dark matter, which does not interact with ordinary matter except through gravity, comprising around 25% of the total matter-energy budget. However, despite extensive searches, the nature of DM remains unknown. Well-motivated candidates include weakly interacting massive particles (WIMPs) predicted in extensions to the standard model of particle physics, though experiments have yet to definitively detect DM in the laboratory. Understanding DM interactions and the origin of its abundance represents a key open problem at the frontier of astroparticle physics and cosmology.

Expansion history and growth of structures. Precision measurements of the cosmic microwave background and large-scale structure have helped establish the standard Lambda Cold Dark Matter (Λ CDM) cosmological model. This 6-parameter model provides an excellent fit to observations, with the cosmic expansion rate $H(z)$ increasing at early times due to radiation and matter before transitioning to acceleration driven by the cosmological constant Λ at late times. However, tensions have emerged at the level $3 - 4\sigma$ when comparing probes of the expansion history such as H_0 with those of structure growth such as σ_8 . Resolving these discrepancies may require new physics to modify the early or late-time evolution from the baseline Λ CDM scenario. Continued observational progress promises to reveal whether the standard model requires small extensions or potentially new paradigms to fully describe our Universe.

Several observational facilities aim to shed new light on these questions. The Euclid mission (Laureijs et al., 2011) launched in 2023 will precisely measure billions of galaxies to map the expansion history and large-scale structure. Euclid is optimized for two primary cosmological probes: weak gravitational lensing, consisting of measuring the distortions of galaxy images

caused by mass inhomogeneities along the line of sight, and baryonic acoustic oscillations, imprinted in the clustering of galaxies (see Section 1.4.3). The Legacy Survey of Space and Time (Ivezić et al., 2019) is another vast census of galaxies carried out at the Vera C. Rubin Observatory planned to begin in 2025. Two years later, the Roman Space Telescope (Akeson et al., 2019) will be launched with the end goal of measuring the effects of dark energy. Meanwhile, next-generation gravitational wave detectors, such as the Einstein Telescope (Punturo et al., 2010) and Cosmic Explorer (Reitze et al., 2019), will significantly expand the field of GW astronomy and cosmology. The new observations will significantly improve statistics and cosmological volume probed with respect to the one obtained up to now by the LIGO-Virgo-Kagra Collaboration. The multi-band, multi-messenger data from these facilities will provide unprecedented tests of cosmology and general relativity, tackling mysteries like the nature of dark matter and dark energy, the cause of cosmic acceleration, the validity of general relativity on large scales, and more. Each experiment offers new capabilities that together will drive progress on the biggest open questions in cosmology.

1.5.1 Aim of the Thesis

The aim of this Thesis is to explore constraints to the expansion history of the Universe. Motivated by the open questions and tension, we decide to focus on two emerging probes, independent from the main ones, which have the advantage of providing cosmological-model-independent constraints to $H(z)$. On the one hand, cosmic chronometers (CC, Chapter 2) allow one to directly measure $H(z)$ up to $z \sim 2$, while, on the other hand, gravitational waves (GW, Chapter 3) allow one to obtain direct measurement of d_L without needing an additional calibrator, thus providing a direct measurement of $H_0 = H(z = 0)$. In particular, we explore new methodologies for the analysis of each probe to maximize its scientific yield. The main focus for both GWs & CCs has been the interplay of cosmological and astrophysical parameters to derive them jointly, study possible degeneracies, and eventually minimize potential systematic effects. Ultimately, the aim of this Thesis is to place new constraints on $H(z)$ and to explore future constraints in the context of upcoming GW observatories and galaxy surveys.

Cosmology with Cosmic Chronometers

Cosmic chronometers (CC) have proven to be very promising probes to obtain direct measurements of $H(z)$ up to $z \sim 2$. This method, first introduced by Jimenez & Loeb (2002), consists of using massive and passive galaxies as tracers of the aging of the Universe under the minimal assumption of an FLRW metric. While the redshifts can be precisely measured, age-dating galaxies is challenged by the complex reconstruction of their star formation history and inherent degeneracies within stellar population parameters, like stellar age, formation timescale, and chemical composition. In this work, our aim is to address this issue by studying an optimal selection of a pure sample of CCs, measure their ages in a cosmology-independent way together with their chemical abundances, derive a new measurement of $H(z)$ at intermediate z , and thoroughly evaluate potential systematics. Section 2.1 provides a brief introduction to the CC method, while the subsequent sections cover the original contributions of this Thesis.

The main reference articles for the analyses presented in this chapter are:

- **Nicola Borghi**, Michele Moresco, Andrea Cimatti, Alexandre Huchet, Salvatore Quai, & Lucia Pozzetti, *Toward a Better Understanding of Cosmic Chronometers: Stellar Population Properties of Passive Galaxies at Intermediate Redshift*, *The Astrophy. J.*, **927** (2022) no. 2, 164 [[arXiv:2106.14894](#)]
- **Nicola Borghi**, Michele Moresco, & Andrea Cimatti, *Toward a Better Understanding of Cosmic Chronometers: A New Measurement of $H(z)$ at $z \sim 0.7$* , *The Astrophy. J. Lett.*, **928** (2022) no. 1, L4 [[arXiv:2110.04304](#)]
- Kang Jiao, **Nicola Borghi**, Michele Moresco, & Tong-Jie Zhang, *New Observational $H(z)$ Data from Full-Spectrum Fitting of Cosmic Chronometers in the LEGA-C Survey*, *The Astroph. J., Supp. Ser.*, **265** (2023) no. 2, 48 [[arXiv:2205.05701](#)]
- Elena Tomasetti, Michele Moresco, **Nicola Borghi**, Kang Jiao, Andrea Cimatti, Lucia Pozzetti, Adam C. Carnall, Ross J. McLure, and Laura Pentericci. *A new measurement of the expansion history of the Universe at $z = 1.26$ with cosmic chronometers in VANDELS*, *Astron. Astrophys.* (2023) [[arXiv:2305.16387](#)]
- Elcio Abdalla, et al. (incl. **Nicola Borghi**), *Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies*, *J. High Energy Phys.*, **34** (2022), 49 [[arXiv:2203.06142](#)]

- **Nicola Borghi.** *Toward an independent reconstruction of the expansion history of the Universe*, ESO Hypatia Colloquium (2022) [DOI: 10.5281/zenodo.7104538]

Along with this work we developed and publicly released a flexible Python tool to measure spectral features in galaxy spectra:

 **PyLiCK**: Available at: <https://pylick.readthedocs.io>

2.1 Framework: theory and observations

This Section provides an introduction to the cosmic chronometers method (§ 2.1.1) and to passive galaxies (§ 2.1.2), which are the best CC candidates. Then it discusses the methodologies to measure their ages (§ 2.1.3) and, ultimately, presents the main systematics associated with the CC method (§ 2.1.4).

2.1.1 The cosmic chronometers method

Under the minimal assumption of a Friedmann–Lemaître–Robertson–Walker (FLRW) metric, the Hubble parameter $H(z)$ can be expressed as a function of the differential time evolution of the universe dt_U in a given redshift interval dz , as provided by Eq. (1.15):

$$H(z) = -\frac{1}{1+z} \frac{dz}{dt_U}. \quad (2.1)$$

This means that by measuring the differential aging of the Universe between two redshift bins it is possible to obtain a direct measurement of the expansion rate $H(z)$. However, dt_U is not a direct observable.

The idea of using a homogeneous population of astrophysical objects to trace dt_U , i.e. *cosmic chronometers* (CC), came from Jimenez & Loeb (2002), who proposed massive passively evolving galaxies as ideal candidates. Many observational probes have shown that these galaxies build up their mass at high redshift ($z \gtrsim 2$) over short timescales (< 1 Gyr) exhausting almost completely their gas reservoir in the very first stages of their life and hence evolve passively to the present age (e.g., Cimatti et al., 2004; Treu et al., 2005; Renzini, 2006; Pozzetti et al., 2010; Thomas et al., 2010).

The valuable aspect of this method, in comparison to other probes, is the possibility of directly measuring $H(z)$ without relying on cosmological assumptions beyond the requirement of an FLRW metric. Consequently, CC constraints can be used to test a wide range of cosmological predictions.

In summary, the key ingredients of the CC method are the following:

1. Selection of an optimal CC sample. The CC method is based on the use of a homogeneous population of tracers at different redshifts. It is, therefore, crucial to maximize the completeness and purity of the passive galaxy samples (see Section 2.1.2), by minimizing the presence of star-forming outliers.
2. Robust measurement of dt . Although redshifts can be measured with a high level of precision and accuracy (typically 0.1% for spectroscopic observations), it is important to find accurate methodologies to obtain an unbiased measurement of dt (see Section 2.1.3).
3. Evaluation of systematic effects. Like any other cosmological probe, it is crucial to assess potential systematic effects that could introduce measurement biases (see Section 2.1.4).

2.1.2 Passively evolving galaxies

Since the initial work of [Hubble \(1936\)](#), the subdivision of galaxies into morphological classes (e.g., ‘early-type’ ellipticals and lenticulars versus ‘late-type’ spirals and irregulars) has been a good starting point for deducing their physical properties.

Overall, *early-type* galaxies are more massive, and at a fixed mass, they appear redder, contain an older stellar population, and exhibit less ongoing star formation (e.g., [Strateva et al., 2001](#); [Kauffmann et al., 2003](#); [Baldry et al., 2004](#)). Based on these characteristics, various classifications can be found in the literature, including *red*, *red and dead*, *quiescent*, and *passive*. In particular, the term *passive* refers to galaxies where star formation can be considered suppressed or, at most, negligible. However, despite the aforementioned traits (morphology, color, age, and star formation), these properties have been shown to not perfectly overlap [Moresco et al. \(2013\)](#).

Passive galaxies are excellent laboratories for studying cosmology. While the Λ CDM model predicts a hierarchical growth of structures, there is substantial evidence that galaxy evolution is driven by galaxy mass, with more massive galaxies forming earlier and over shorter timescales compared to less massive galaxies. This scenario is known as *downsizing* (see e.g., [Cowie et al., 1996](#); [Cimatti et al., 2006](#); [Thomas et al., 2010](#)). Evidence from the local Universe comes from archaeological studies of ages and chemical abundances. Analyses of spectral indices for large samples of galaxies find that local massive elliptical galaxies exhibit systematically older ages and higher metallicities compared to less massive galaxies ([Gallazzi et al., 2005](#); [Thomas et al., 2005, 2010](#); [McDermid et al., 2015](#)). They also show super-solar $[\alpha/\text{Fe}]$ ratios, indicating very short star formation timescales.

Given these old formation ages $z_f > 2$ over short timescales, passive galaxies are expected to exist already at $z > 1$. Their comoving number density should also remain approximately constant at low redshifts. These predictions are confirmed by observations of passive galaxies out to high redshifts $z > 2$ ([Cimatti et al., 2004](#); [Belli et al., 2015](#)) up to $z = 4.658$ ([Carnall et al., 2023](#)) when the Universe was less than 1.5 billion years old. The existence of such objects demonstrates that the first onset of galaxy quenching occurred no later than 1 Gyr after the Big Bang, tightly constraining galaxy evolution and cosmological models.

Formation scenarios

Traditionally, two main models have competed to explain the formation of these galaxies (see [Cimatti et al., 2019](#)):

- Monolithic collapse model (e.g., [Eggen et al., 1962](#); [Dekel et al., 2009](#)): Galaxies form rapidly through the dissipative collapse of primordial gas clouds at high redshift $z \gtrsim 2$. Intense star formation is quickly followed by passive evolution, as gas is converted into stars or expelled by feedback processes. Globular clusters also form during this early dissipative collapse stage.
- Hierarchical merging model (e.g., [Blumenthal et al., 1984](#)): Galaxies grow through successive mergers and accretion of smaller systems over cosmic time. Mergers play a key role in the transformation of disks into spheroids. Galaxy mergers are also thought to be responsible for triggering starbursts and AGN activity observed in many systems.

For this model to reproduce early-type galaxies, the progenitors must either already be gas-poor or the mergers must be sufficiently dissipationless.

The currently favored scenario presents aspects of both models: an initial phase where mass assembly occurs through cold filaments leading to early in situ star formation between $2 \lesssim z \lesssim 6$ and a subsequent phase at $z \lesssim 2$ where new stellar populations are accreted through mergers and can significantly contribute to the total stellar content of present-day galaxies (Oser et al., 2010). This two-phase model has been able to successfully explain many observed properties of massive early-type galaxies across cosmic time, such as their age distributions, colors, and sizes.

2.1.3 Measuring the differential ages

Age-dating stellar populations present several challenges. While the age information is encoded in their spectral energy distribution (SED), retrieving an accurate measurement is hindered by the complex reconstruction of the star formation history and degeneracies with the chemical composition. The age-metallicity degeneracy is the most well known (Worthey et al., 1994): a similar SED can be obtained by modifying the ages and metallicities $[Z/H]$ in a ratio $\Delta\text{age}/\Delta[Z/H] \sim 3/2$. Another critical degeneracy is the age-star formation timescale τ (Gavazzi et al., 2002): a similar SED can be obtained by modifying ages and τ by $\Delta\text{age}/\Delta\tau \sim 3$ (Borghi et al., 2022b). Other issues such as the presence of dust is typically negligible for passive populations in an accurate selection is performed in advance (e.g., Pozzetti & Mannucci, 2000).

The methods adopted to constrain dt may vary according to the available data and stellar models. A key quantity to characterize observed and modeled SED is the spectral resolution R defined as

$$R \equiv \frac{\lambda}{\delta\lambda}, \quad (2.2)$$

where $\delta\lambda$ is the minimum interval that separates two λ samples. Photometric data have typical $R_{\text{phot}} \sim 5$, while current spectroscopic observations of local galaxies can reach up to $R_{\text{spec}} \sim 15000$. Although photometric observations can cover a wide range of wavelengths (typically UV to far IR), spectroscopy is performed on narrower ranges. It is then useful to introduce the spectral coverage as

$$\Delta\lambda = \lambda_{\text{max}} - \lambda_{\text{min}}. \quad (2.3)$$

Finally, another relevant property to characterize observed data is the signal-to-noise ratio,

$$\text{SNR} \equiv \frac{F}{\sigma_F}, \quad (2.4)$$

where F is the observed flux and σ_F the associated uncertainty. Typically, $F = F(\lambda)$ is a function of wavelength. These quantities can define the four main methodologies used to study differential ages that are presented in the subsequent paragraphs and illustrated in Fig. 2.1.

Full-spectral fitting. When high resolution and SNR spectra are available we would like to use the whole information available. This can be achieved with full-spectral fitting

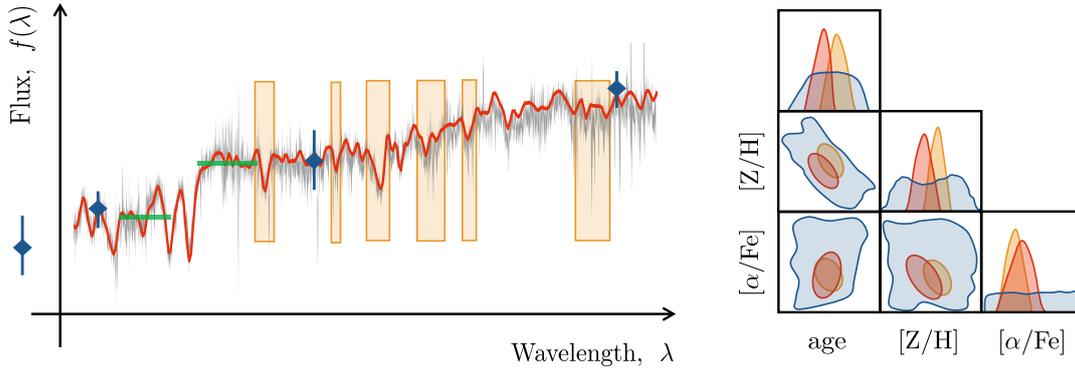


FIGURE 2.1: Illustration of various methods to constrain galaxy's ages: photometry fitting (blue), D4000 (green), absorption features (orange), full-spectral-fitting (red). Left: Galaxy spectrum, Right: Resulting contours.

techniques. The spectrum of a galaxy is usually simulated by convolving the spectrum of single-burst simple stellar populations (SSP) with a parameterized star formation history (SFH = SFR(t_U)) by summing over each i -th cosmic time bin and j -th population (see e.g., [Carnall et al., 2018](#)),

$$F(\lambda) = \sum_i \sum_j F_{\text{SSP}}(\lambda; t_i, [Z/H]_j) \text{SFR}_j(t_{U,i}) \Delta t_i, \quad (2.5)$$

where t_i is the age of the population and $t_{U,i}$ the age of the Universe. Eventually, photometric data points can be constrained simultaneously with the spectrum. Once the ages of a statistically meaningful sample of CC are measured, it is possible to compute dt by dividing the data into different z bins. Successful CC analyses have been conducted with this method by several groups ([Simon et al., 2005](#); [Stern et al., 2010](#); [Zhang et al., 2014](#); [Ratsimbazafy et al., 2017](#)), including the recent joint spectroscopic and photometric analyses by [Jiao et al. \(2023\)](#) and [Tomasetti et al. \(2023\)](#).

Absorption features. An alternative approach to full-spectral fitting consists of using only specific features that are known to be very sensitive to specific parameters, e.g., Balmer lines for age, iron lines for $[Z/H]$, and magnesium lines for $[\alpha/Fe]$ (e.g., see [Korn et al., 2005](#)). Being based on integrated measurements, this approach is less sensitive to noise fluctuations and spectral resolution and therefore can perform better in the presence of lower-quality data. Moreover, the selection of a specific region, while increasing statistical error (less information is used), may minimize systematics related to specific assumptions on full-spectrum modeling.

D4000. As introduced by [Moresco et al. \(2011\)](#), a single spectral feature can be used to trace the differential ages of CCs. This is the case of the discontinuity at 4000 Å (D4000), which can be easily detected and measured in early-type galaxies. The correlation between D4000 and age can be modeled as a linear relationship of the form $\text{D4000}(Z, \dots) \propto A(Z, \dots) \cdot \text{age}$, where the coefficient A accounts for the effects of varying stellar population parameters (such as metallicity). Then Eq. (2.1) can be

rewritten as (see [Moresco et al., 2011](#)),

$$H(z) = -A(Z, \dots) \frac{1}{1+z} \frac{dz}{dD4000}. \quad (2.6)$$

To estimate $H(z)$, the following steps are necessary: (1) measure the median (or mean) D4000– z relation from observational data and estimate the quantity $dz/d4000$; (2) calibrate the D4000–age relation using stellar population models to quantify the coefficient A ; (3) derive $H(z)$ and study potential systematic effects, such as the variation of model parameters. As D4000 is the strongest absorption feature of passive galaxies in the UV–IR range, it can be measured even on lower SNR and R data with respect to other absorption features. Moreover, since it requires only a narrow spectral window ($\delta\lambda \sim 300 \text{ \AA}$) it can be the optimal choice for future wide-field grism surveys, such as Euclid ([Laureijs et al., 2011](#)) and Rubin [Ivezić et al. \(2019\)](#). Most of the CC measurements so far have been obtained with this method ([Moresco et al., 2012b](#); [Moresco, 2015](#); [Moresco et al., 2016](#)).

Photometry. Photometric information alone is not sufficient to break the age-metallicity degeneracy to accurately constrain individual galaxy ages. However, as recently demonstrated by [Jimenez et al. \(2023\)](#), Machine Learning techniques can be used to provide CC constraints, when accurately trained on well-determine ages and applied to large ($\sim 10^4$ sample of galaxies). For this reason, this can be a promising approach for Euclid and Rubin photometric surveys.

2.1.4 Main systematics

The full covariance matrix for the CC method has been formalized by [Moresco et al. \(2022\)](#) and can be expressed as:

$$\begin{aligned} \text{Cov}_{ij}^{\text{tot}} = & \text{Cov}_{ij}^{\text{stat}} + \\ & + \text{Cov}_{ij}^{\text{young}} + \\ & + \text{Cov}_{ij}^{\text{SFH}} + \text{Cov}_{ij}^{\text{IMF}} + \text{Cov}_{ij}^{\text{stlib.}} + \text{Cov}_{ij}^{\text{SPS}}, \end{aligned} \quad (2.7)$$

where “stat” refers to the statistical error in the dt or $dD4000$ measurement, “young” to the young component contamination¹ ([Moresco et al., 2016, 2018](#)), while the other terms correspond to the ingredients of the chosen model, namely: star formation history (SFH), initial mass function (IMF), stellar library (st.lib.), stellar population model (SPS). The subsequent paragraphs provide a more thorough description of the main systematics.

Residual Star Formation. After the selection process, the sample may still be contaminated by galaxies with residual or with recent bursts of star formation. Although the young population may not dominate in terms of mass, it can significantly alter the overall galaxy emission (e.g., [Maraston, 2005](#)), compromising the accurate modeling of their age and potentially introducing a bias in the dt measurement. To overcome this issue, [Moresco et al. \(2018\)](#) proposed a combination of multiple selection criteria, with an

¹An additional term $\text{Cov}_{ij}^{\text{met}}$ is needed in the D4000 approach as the metallicity is fixed from external data.

initial photometric selection, followed by spectral feature-based cuts that trace different phases of ongoing or recent stellar evolution. The author found that contamination on previous CC samples (Moresco et al., 2012b, 2016) is consistent with zero and can at most introduce a $0.4 - 1\%$ bias (at 1σ) in $H(z)$.

Stellar Population Model. The choice of the stellar population model can significantly affect the measurements of $H(z)$. This has been extensively explored in the recent work by Moresco et al. (2020), where they calculated the contribution to the systematic error covariance matrix from different parameters of various stellar population models. The total contribution is dominated by the choice of the model used, which can introduce systematic errors of up to $\sim 5 - 9\%$. Furthermore, an error in determining the metallicity of about 5% results in a systematic error of approximately 4% in the estimation of $H(z)$. These contributions can be reduced with higher-quality spectra and future improvements in stellar population models.

Progenitor Bias. The progenitors of younger early-type galaxies can be missing at higher z in galaxy catalogs due to observational constraints (Bender & Davies, 1996; van Dokkum et al., 2000). This may lead to a flattening of the age-redshift relation and thus a bias on $H(z)$. By deriving $H(z)$ from the upper envelope of the oldest galaxies, this effect is mitigated. This has been used in Moresco et al. (2012b) to estimate the impact of progenitor bias, finding that it could contribute at most to 1% uncertainty to $H(z)$, which translates to $\lesssim 2\%$ in $\Omega_{m,0}$ and $\Omega_{DE,0}$ and $\lesssim 10\%$ in w_0 (Moresco et al., 2012a).

2.2 Selection of passive galaxies

This section presents the selection of a sample of passive galaxies in LEGA-C (§ 2.2.1) and the measurements of specific spectral features (§ 2.2.1). The observed quantities (§ 2.2.3) allow to both characterize the sample and provide a first assessment of redshift trends of the main physical properties of these galaxies.

2.2.1 The passive sample

The sample of galaxies used in this study is sourced from the second data release (DR2) of LEGA-C, an ESO Public Spectroscopic Survey that focused on approximately 3000 K_s -band-selected galaxies within the redshift range $0.6 \lesssim z \lesssim 1$ in the COSMOS field. The observations were conducted using the Visible Multi-Object Spectrograph (VIMOS) on the Very Large Telescope at the Paranal Observatory. The flux-calibrated spectra cover the wavelength range of $6300 < \lambda < 8800 \text{ \AA}$ with a spectral resolution of $R \sim 3500$, and a median signal-to-noise ratio (SNR) of around 15 per pixel (0.6 \AA). Spectra were acquired using $1''$ wide slits, corresponding to approximately 7 kpc at these redshifts.

The publicly available catalog for DR2 includes reduced galaxy spectra, spectroscopic redshifts, flux measurements for the main emission lines (Balmer lines, [O II] λ 3727, [O III] λ 4959, [O III] λ 5007), velocity dispersion determinations, and measurements of a selection of 14 Lick/IDS indices (Straatman et al., 2018). In our analysis, we use galaxy spectra, as well as the measurements of redshift (z), observed stellar velocity dispersion (σ_*), and [O II] λ 3727 emission-line flux. For the main analysis, we do not use the spectral indices measurements provided in the catalog, but instead determine our own line strengths from the spectra after matching the resolution to the one of the adopted stellar population models. This also allows us to extend measurements to redder indices, reaching up to $\sim 5000 \text{ \AA}$.

We cross-match the LEGA-C sample with the COSMOS 2015 catalog (Laigle et al., 2016) employing a search radius of $1''$ to complement spectroscopic data with photometric information, including NUV, r , and J bands, as well as stellar masses (M_*) and specific star formation rates (sSFR = SFR/ M_*) derived through SED fitting. Additionally, we incorporate morphological data from the Zurich Estimator of Structural Types catalog (Scarlata et al., 2007), based on principal component analysis of surface brightness profiles. Our parent sample is selected by requiring high-quality spectra (see Straatman et al. 2018 for spectra quality flags) and the availability of NUV, r , and J absolute magnitudes, resulting in a sample of 1622 sources.

Selection criteria

Various methodologies have been proposed in the literature for distinguishing between “passive” and “star-forming” galaxies. These methods encompass a range of approaches, such as the morphological selection of spheroidal systems (following the original separation by Hubble, 1936), cuts on color-color diagrams (e.g., UVJ Williams et al. 2009 and NUVrJ Ilbert et al. 2013) or on color-mass diagrams (e.g., Peng et al., 2010), SED fitting techniques (e.g., Ilbert et al., 2009), and sSFR criteria (e.g., Pozzetti et al., 2010).

However, these criteria do not completely align (Renzini, 2006). In fact, a single criterion is not stringent enough, resulting in 10% – 30% contamination from star-forming outliers.

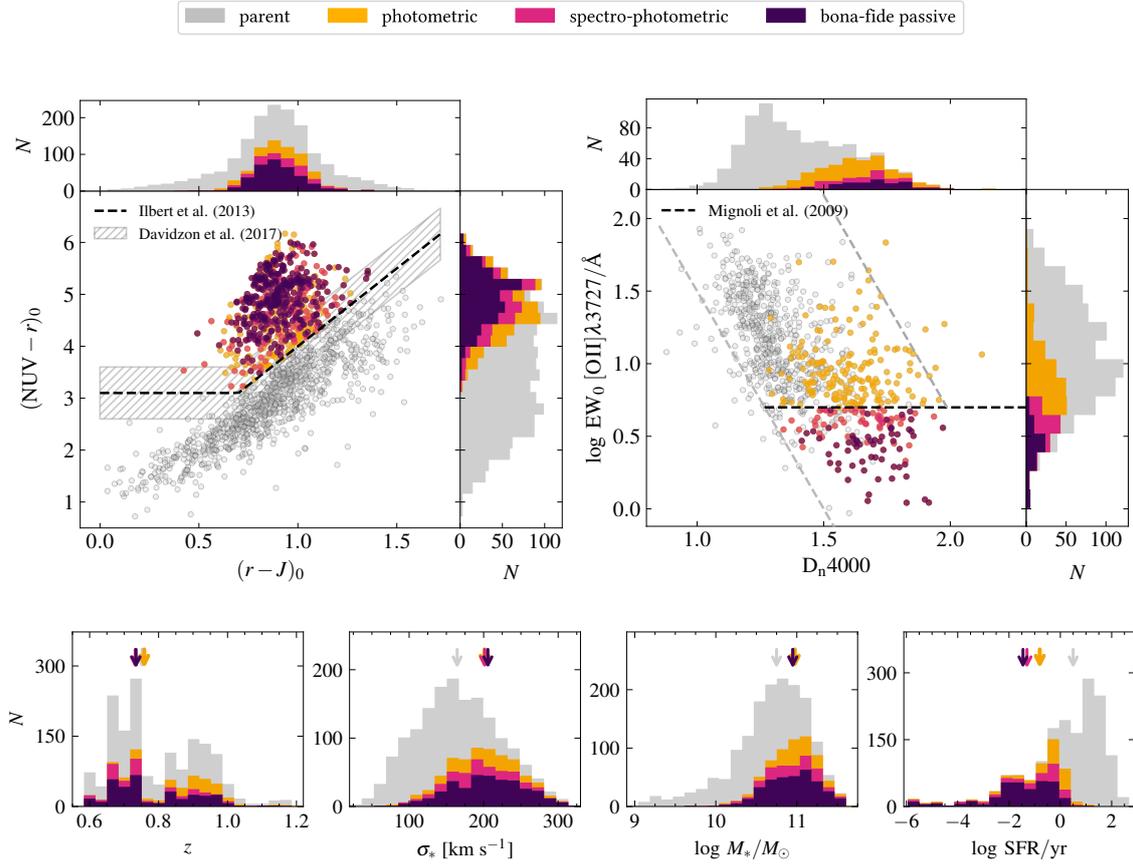


FIGURE 2.2: Distribution of key observational properties in subsequently refined samples of passive galaxies in LEGA-C. *Upper panels:* NUVr and $\text{EW}[\text{O II}]-D_{n4000}$ diagnostic diagrams with selection criteria from Ilbert et al. (2013) and Mignoli et al. (2009). The hatched region (± 0.05 mag from the NUVr cut) indicates the location of green-valley galaxies (Davidzon et al., 2017). We use black colors for the criteria also adopted in this study, while gray lines are for illustrative purposes only. Note that the $[\text{O II}]\lambda 3727$ line is detected for only 65/350 passive galaxies. *Lower panels:* redshift (z), observed stellar velocity dispersion (σ_*), stellar mass (M_*), and star formation rate (SFR) distributions. Arrows represent the median values of the different subsamples.

Table 2.1. Median properties of different subsamples of passive galaxies selected from LEGA-C.

Sample	N	$\langle z \rangle$	$\langle \sigma_{\star} \rangle$ km s $^{-1}$	$\langle \log M_{\star} \rangle$ M $_{\odot}$	$\langle \log \text{SFR} \rangle$ M $_{\odot}$ yr $^{-1}$	$\langle \text{H/K} \rangle$	$\langle D_n4000 \rangle$	$\langle \text{H}\delta_A \rangle$ Å	$\langle \text{G4300} \rangle$ Å	$\langle \text{Fe4383} \rangle$ Å	$\langle \text{C}_24668 \rangle$ Å
<i>parent</i>	16220	0.752	165	10.75	0.50	1.166	1.426	2.941	2.357	2.583	3.256
<i>photo</i>	658	0.758	201	10.98	-0.80	0.969	1.654	0.692	4.286	3.594	4.935
<i>spectrophoto</i>	485	0.732	202	10.95	-1.30	0.957	1.681	0.462	4.387	3.741	4.996
<i>bona fide</i>	350	0.735	206	10.95	-1.46	0.957	1.690	0.347	4.443	3.832	5.097

Note. — Data from: LEGA-C DR2 (col. 3–4, [Straatman et al. 2018](#)); COSMOS2015 (col. 5–6, [Laigle et al. 2016](#)); this work (col. 7–12, Section 2.2.3).

The combination of multiple criteria, leveraging both photometric and spectroscopic data, is more effective in ensuring a pure sample ([Franzetti et al., 2007](#); [Moresco et al., 2013](#)).

In our analysis, where the goal is to obtain the purest possible sample of cosmic chronometers, we combine various complementary selection cuts, as detailed below:

- **NUVrJ Selection.** We identify photometric passive galaxies using the rest-frame NUV – r and $r - J$ colors, as defined by [Ilbert et al. \(2013\)](#). This criterion requires $(\text{NUV} - r) > 3.1$ and $(\text{NUV} - r) > 3(r - J) + 1$. These color combinations can effectively distinguish objects with recent (1–100 Myr) star formation episodes (bluer NUV – r colors) even if they are dust-obscured (redder $r - J$ colors). For this reason, they have been extensively used in the literature to separate quiescent and star-forming galaxies ([Arnouts et al., 2007](#); [Ilbert et al., 2015](#); [Davidzon et al., 2017](#)). Applying this cut results in 658 sources, referred to as the *photometric* passive sample.
- **Emission-line cut.** We remove galaxies with a strong [O II] λ 3727 emission line, typically associated to ongoing star formation.² In particular, we apply the following threshold $\text{EW}[\text{O II}] > 5 \text{ \AA}$, as previously done in the literature (e.g., [Mignoli et al., 2009](#)). This yields 485 galaxies, referred to as the *spectrophotometric* passive sample.
- **Visual inspection.** The sample is further refined by visually inspecting all the remaining spectra. We exclude galaxies with strong [O II] λ 3727 and/or [O III] λ 5007 lines, ensuring typical $\text{SNR} < 3$ in their equivalent widths (EWs). This step is particularly essential to spectroscopically characterize $z \lesssim 0.65$ galaxies, where [O II] is unavailable in LEGA-C spectra. At the end, we obtain a final sample of 350 *bona fide* passive galaxies.

²While low-ionization nuclear emission-line regions (LINERs) and ionization from old stars can also produce [O II] and [O III] emission lines (e.g., [Yan et al., 2006](#); [Singh et al., 2013](#); [Cimatti et al., 2019](#)), we opt for strict purity and exclude these sources as well.

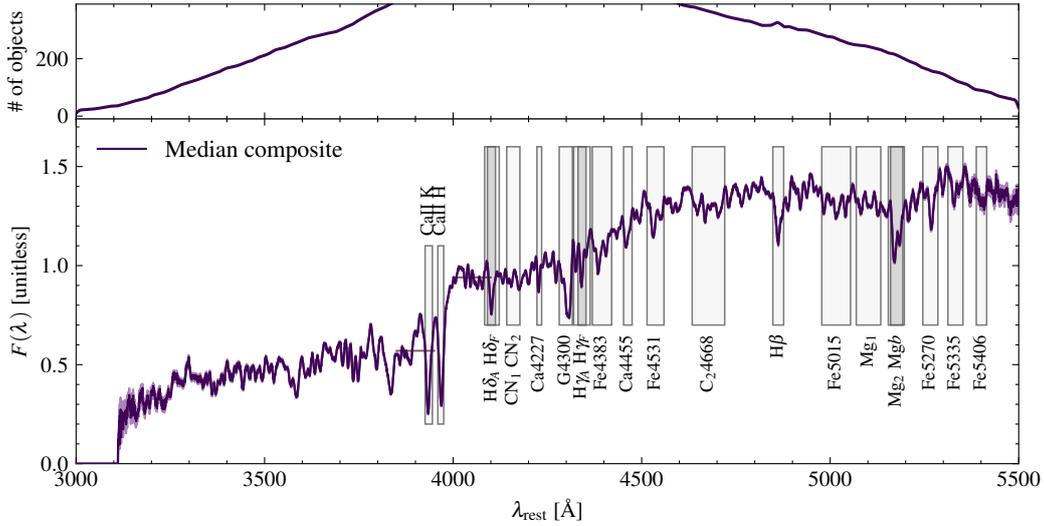


FIGURE 2.3: Median composite spectrum of 350 passive galaxies selected in LEGA-C. *Upper panel*: Number of stacked galaxies at each wavelength; *Lower panel*: Median composite spectrum with 1σ error. Gray boxes represent the central regions of the 22 Lick and pseudo-Lick indices measured in this study; narrower indices in overlapping regions ($H\delta_F$, $H\gamma_F$, and $Mg\ b$) have darker shadings. Horizontal segments indicate the blue and red bandpasses used for computing D_n4000 .

Figure 2.2 illustrates the distribution of LEGA-C galaxies in two diagnostic diagrams (NUV_rJ and $EW[O\ II]-D_n4000$), as well as their redshift, σ_* , M_* , and SFR values, across the *parent*, *photometric*, *spectrophotometric*, and *bona fide* passive subsamples. It’s important to note that objects without $[O\ II]$ detection are not displayed in the upper-right diagram. Table 2.1 provides the median values of key parameters.

The LEGA-C galaxies exhibit two distinct peaks in their redshift distribution, centered around $z \sim 0.7$ and $z \sim 0.9$, with a limited number of galaxies at $z > 1$. In the NUV_rJ plot, they form two distinct groups: a “blue sequence” which reaches low ($NUV - r$) $\simeq 1$ values and a “red cloud” which constitutes the *photometric* passive sample. While a NUV_rJ -only criterion significantly reduces the presence of star-forming systems in the sample ($\langle \log sSFR/yr \rangle = -11.8$), about one-third of objects still exhibit significant $[O\ II]$ emission. Therefore, the subsequent spectroscopic selection on the EW of the $[O\ II]$ line is crucial to ensure the purity of the sample. This second cut not only eliminates the tail of bluer ($NUV - r$) galaxies in the green-valley region, but also systems with redder colors. The final *bona fide* passive sample has a median redshift of $\langle z \rangle = 0.735$. Passive galaxies tend to occupy the high- σ_* and $\log M_*/M_\odot$ tails of the parent distribution. Specifically, the median σ_* ($\log M_*/M_\odot$) increases from 164.5 km s^{-1} (10.75) to 205.7 km s^{-1} (10.95) moving from the *parent* to the *bona fide* passive sample. Compared to the *spectrophotometric* sample, the *bona fide* passive sample exhibits a lower SFR by 0.16 dex, with a median uncertainty of 0.18 dex. As a final remark, the passive sample has a median specific star formation rate of $\langle \log sSFR/yr \rangle = -12.1$, with only 15 galaxies (4%) exceeding the threshold > -11 , a common criterion for classifying “passive” galaxies (see Pozzetti et al., 2010).

To further evaluate the robustness of the selection, the composite spectrum of the 350

bona fide passive sample is constructed. First of all, each rest-frame spectrum is normalized to the median flux within the 4200–4400 Å range. This minimizes potential biases toward galaxies with higher SNRs. Each spectrum is then interpolated onto a common grid spanning 3000–5500 Å, using a resolution of 0.35 Å per pixel. Pixels with bad spectral flags are excluded. To mitigate the impact of potential sky emission-line residuals and noise fluctuations, a σ -clipping process is applied, removing flux values deviating by more than 4 standard deviations from the mean. In each i -th pixel, the composite spectrum flux F_i is computed as the median flux of the stacked spectra, while the associated uncertainty as the normalized median absolute deviation (NMAD, see Hoaglin et al., 1983) divided by the square root of the number of objects, i.e. $\sigma_i = 1.483 \times \langle |F_i - \langle F_i \rangle| \rangle / \sqrt{N_i}$. The resulting spectrum shown in Figure 2.3 has a high SNR ~ 230 per pixel at ~ 4000 Å rest frame.

The spectrum exhibits characteristic features consistent with a passively evolving stellar population: prominent red continuum, strong 4000 Å discontinuity, no Balmer break at 3650 Å, Ca II K absorption line deeper than Ca II H (see Section 2.2.3), and presence of several metallic indices (e.g., G4300, Fe4383, Fe4531, and Mg₂). Importantly, even by stacking the spectra, no emission lines become detectable.

In summary, the analysis of stellar masses, SFRs, and the composite spectrum, confirms the strength of the selection criteria adopted: these galaxies show no evidence of recent star formation. The presence of possible underlying young stellar components must be further assessed by studying spectral absorption features.

2.2.2 PyLick: a new tool to measure spectral features

This section presents **PyLick** , a new Python tool that we develop and publicly release to measure spectral features from galaxy spectra.³ Currently, 80 indices from the near-UV to the near-IR are included and new ones can be easily introduced. In Appendix A.1 the code is validated against the available LEGA-C measurements. The code has been recently used to obtain the spectroscopic measurements catalog for the VANDELS survey (Talia et al., 2023). **PyLick** is currently based on five main modules:

- **io.py**: implements the built-in methods for spectra I/O and preliminary analysis;
- **indices.py**: loads the index library with pass-bands definitions;
- **measurements.py**: measures multiple indices on a single spectrum;
- **analysis.py**: contains two main classes: **Galaxy**, optimized to analyze a single spectrum (wavelength, flux, flux error, [mask]) and **Catalog**, optimized to perform the analysis of a catalog of spectra;
- **plot.py**: plotting routines.

PyLick implements several spectral features definition: Lick indices, discontinuities, bumps, and specifically defined features (see Table A.1). Here we list the main equations depending on the **units** value in the index table in parenthesis,

³Documentation available at: <https://pylick.readthedocs.io>

$$I_a = \int_{\lambda_{c1}}^{\lambda_{c2}} \left(1 - \frac{F(\lambda)}{F_c(\lambda)} \right) d\lambda, \quad [\text{\AA}] \quad (2.8)$$

$$I_m = -2.5 \log \left(\frac{1}{\lambda_{c2} - \lambda_{c1}} \int_{\lambda_{c1}}^{\lambda_{c2}} \frac{F(\lambda)}{F_c(\lambda)} d\lambda \right), \quad [\text{mag}] \quad (2.9)$$

$$D4000 = \frac{\lambda_{b2} - \lambda_{b1} \int_{\lambda_{r1}}^{\lambda_{r2}} \lambda^2 F(\lambda) d\lambda}{\lambda_{r2} - \lambda_{r1} \int_{\lambda_{b1}}^{\lambda_{b2}} \lambda^2 F(\lambda) d\lambda}, \quad [\text{break_nu}] \quad (2.10)$$

$$B4000 = \frac{\lambda_{b2} - \lambda_{b1} \int_{\lambda_{r1}}^{\lambda_{r2}} F(\lambda) d\lambda}{\lambda_{r2} - \lambda_{r1} \int_{\lambda_{b1}}^{\lambda_{b2}} F(\lambda) d\lambda}, \quad [\text{break_lb}] \quad (2.11)$$

$$\text{Mg}_{\text{UV}} = \frac{2 \int_{2625}^{2725} F(\lambda) d\lambda}{\int_{2525}^{2625} F(\lambda) d\lambda + \int_{2725}^{2825} F(\lambda) d\lambda}, \quad [\text{bump}] \quad (2.12)$$

where $F(\lambda)$ and $F_c(\lambda)$ are the spectrum flux and the local pseudo-continuum, which can be computed through linear interpolation or by weighting the fraction of covered pass-bands. In particular, `PyLick` handles three measuring methods:

- `int` (Default): the spectrum is interpolated with a fine bin size of \AA ;
- `exact`: the spectrum is interpolated with `scipy.interpolate.interp1d` (0th order). The resulting function can be evaluated at each wavelength, therefore the integrals are more precise at the edges of pass-band windows;
- `wei`: the flux of pixels is weighted by the fraction covered by the pass-bands definition.

The complete Lick system comprises 25 indices, encompassing the rest-frame wavelength range of 4000 – 6000 \AA (Worthey & Ottaviani, 1997; Trager et al., 1998). Each index requires the identification of a central region (λ_{c1} , λ_{c2}), as well as two regions toward the red (λ_{r1} , λ_{r2}) and blue (λ_{b1} , λ_{b2}) of the central one that are employed to estimate a reference pseudo-continuum level Eqs. (2.8) and (2.9).

The 4000 \AA discontinuity (D4000, eq. 2.10) is one of the strongest features in passive galaxy spectra. It arises from the accumulation of multiple spectral lines from G0-type and cooler stars at $\lambda < 4000 \text{\AA}$. If young stars are present, the discontinuity appears shallower because of their strong flux at these wavelengths. This property makes it a good age tracer (e.g., Kauffmann et al., 2003; Moresco et al., 2011). The D4000 index is defined as the ratio of the average flux density in two wavelength regions, one redwards and one bluewards of 4000 \AA . The original definition by Bruzual (1983) sets $\lambda_{b1}, \lambda_{b2} = 3750, 3950 \text{\AA}$ and $\lambda_{r1}, \lambda_{r2} = 4050, 4250 \text{\AA}$. Alternatively, Balogh et al. (1999) sets narrower intervals (D_n4000) with $\lambda_{b1}, \lambda_{b2} = 3850, 3950 \text{\AA}$ and $\lambda_{r1}, \lambda_{r2} = 4000, 4100 \text{\AA}$, to minimize reddening effects.

More specific methods are directly defined in the `measure` module. This is the case for the Mg_{UV} index that traces the absorption bump present at 2640–2850 \AA (Daddi et al., 2005) and for the calcium triplet index $\text{CaT}^* = \text{CaT} - 0.93 \text{ PaT}$, which traces the strength of the CaII lines (CaT) corrected from the contamination by Paschen lines (PaT), as presented in detail in Cenarro et al. (2001).

Typical workflows

The example below shows the typical usage of `PyLick`. As a first step, it may be useful to load the full index table to visualize the available indices (see Table A.1), or eventually define a new one.

```
from pylick.indices import IndexLibrary
lib = IndexLibrary()
```

CODE EXAMPLE 2.1: Table of spectral features available in `PyLick`.

In this way, it is possible to specifically select a set of features to analyze by adding their IDs to a `index_keys` list. The next examples show the analysis of a single galaxy spectrum. Steps include: importing a `.fits` spectrum, loading the table of spectral features with the keys to be measured, at instantiating the `Galaxy` class to perform the measurement.

```
import pylick.io as io
from pylick.indices import IndexLibrary
from pylick.analysis import Galaxy

spectrum = io.load_spec_fits(dir_spec, filename,
                             colnames=['lambda', 'flux', 'flux_err'])
ind_library = IndexLibrary(index_keys)

ind_measured = Galaxy(ID, index_list, spec_wave=spectrum[0],
                      spec_flux=spectrum[1], spec_err=spectrum[2], z=z)
vals, errs = ind_measured.vals, ind_measured.errs
```

CODE EXAMPLE 2.2: Analysis of a galaxy spectrum with `PyLick`.

`PyLick` also includes a specific module for the full catalog analysis. Steps include: defining a function to load the spectra from a catalog folder, loading the table of spectral features with the keys to be measured, and instantiating the `Catalog` class.

```
import pylick.io as io
from pylick.indices import IndexLibrary
from pylick.analysis import Catalog

def load_spec(ID):
    ...
    return [wave, flux, ferr, mask]

IDs = [...]
ind_library = IndexLibrary(index_keys)

ind_measured = Catalog(IDs, load_spec, index_keys, z=zs,
                       do_plot=True, verbose=True)
```

CODE EXAMPLE 2.3: Analysis of a galaxy catalog with `PyLick`.

Index errors are evaluated following the signal-to-noise method proposed by [Cardiel et al. \(1998\)](#). The code handles bad pixels and the user can choose a bad-to-total pixel ratio (defaulting to $BPR=0.15$) above which the measurement is not performed. Otherwise, a zeroth- or first-order interpolation is done over the bad pixels prior to the measurement. An example of the measurement on one spectrum is shown in Fig. 2.4.

Table 2.2. Spectral Indices Properties for the 350 Passive Galaxies

Index	$\langle C(\sigma) \rangle$	$z_{\text{exp}}^{(a)}$	Range ^(b)	$\langle I \rangle^{(c)}$	$\langle \text{SNR} \rangle^{(c)}$
Ca II K	1.000	> 0.65	4.928÷9.157	7.167	22.40
Ca II H	1.000	> 0.65	5.472÷8.340	6.891	40.93
D4000	-0.001	> 0.65	1.732÷2.028	1.881	236.61
D _n 4000	0.008	> 0.65	1.517÷1.846	1.690	172.13
H δ_{A}	1.121	All	-1.438÷3.826	0.347	3.44
H δ_{F}	1.204	All	0.630÷3.516	1.709	10.37
CN ₁	0.014	All	-0.039÷0.098	0.035	5.34
CN ₂	0.015	All	-0.005÷0.145	0.074	9.21
Ca4227	1.154	All	0.290÷1.701	0.922	7.69
G4300	1.053	All	2.605÷5.818	4.443	18.29
H γ_{A}	0.957	All	-4.225÷0.431	-2.729	9.86
H γ_{F}	1.102	All	-1.121÷2.348	-0.105	3.64
Fe4383	1.065	All	1.675÷5.778	3.832	11.87
Ca4455	1.363	< 0.96	0.219÷2.366	1.395	8.54
Fe4531	1.106	< 0.92	1.438÷4.599	2.958	11.16
C ₂ 4668	1.003	< 0.88	1.423÷7.925	5.097	13.94
H β	1.027	< 0.77	1.175÷3.490	1.989	13.72
Fe5015	1.189	< 0.74	1.641÷7.171	4.737	13.98
Mg ₁	0.002	< 0.68	0.036÷0.145	0.077	23.94
Mg ₂	0.003	< 0.68	0.079÷0.280	0.205	46.88
Mg <i>b</i>	1.049	< 0.69	1.789÷4.827	3.274	17.08
Fe5270	1.158	< 0.64	0.628÷4.054	2.549	15.96
Fe5335	1.271	< 0.63	0.425÷3.553	2.448	13.02
Fe5406	1.184	< 0.62	0.144÷2.462	1.573	8.67

Note. — (a) Expected redshift coverage within $0.6 < z < 1$ of the VIMOS HR red spectrograph; (b) Computed between 5th – 95th percentiles; (c) Median index value and signal-to-noise ratio. Index units are angström for all indices except: CN₁, CN₂, Mg₁, Mg₂ (mag); D4000, D_n4000 (dex).

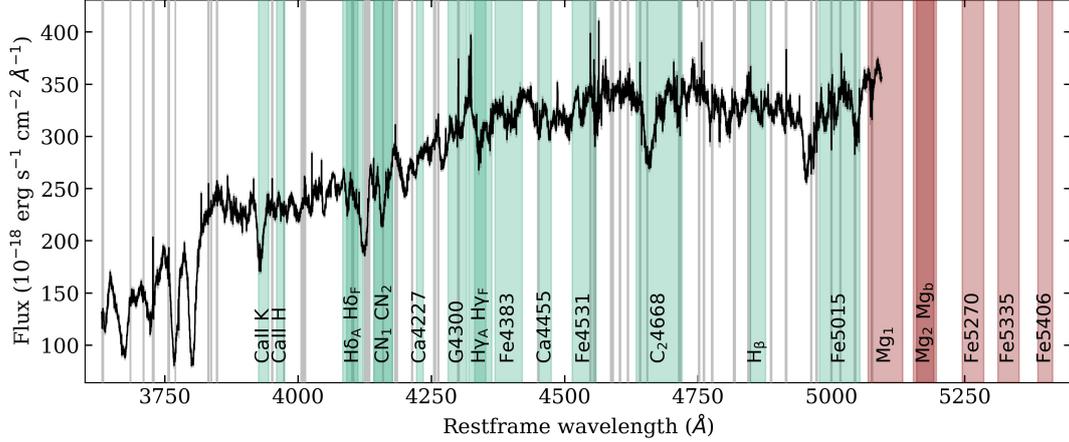


FIGURE 2.4: Example of the measurements of spectral features on a galaxy spectrum with `PyLick`. Vertical bands identify the central regions of indices that have been successfully measured (green) or indices that could not be measured due to lack of spectral coverage (red).

2.2.3 Measuring spectral features with `PyLick`

We measure spectral absorption features of LEGA-C galaxies using `PyLick`. Specifically, we firstly degrade each rest-frame spectrum ($\text{FWHM}_{\text{spec}} \approx 1.3 \text{ \AA}$ at $z = 0.7$) to match the resolution of the models employed ($\text{FWHM}_{\text{mod}} = 2.5 \text{ \AA}$). This is achieved by applying a Gaussian kernel to the observed data using a standard deviation of

$$\sigma = \frac{\sqrt{\text{FWHM}_{\text{mod}}^2 - \text{FWHM}_{\text{spec}}^2}}{2.355}. \quad (2.13)$$

After the measurement, `PyLick` calibrates indices to zero-velocity dispersion following the approach described by [Carson & Nichol \(2010\)](#). The dependence of an index I on σ_* is expressed through a coefficient C , defined differently for atomic and molecular indices:⁴

$$C_{I_a}(\sigma_*) = \frac{I_a(\sigma_* = 0)}{I_a(\sigma_*)}; \quad C_{I_m}(\sigma_*) = I_m(\sigma_* = 0) - I_m(\sigma_*). \quad (2.14)$$

The relationships $C_I(\sigma_*)$ are determined by fitting the indices measured on the [Bruzual & Charlot \(2003\)](#), specifically the 2016 version) models at different σ_* values using fourth-order polynomials (Huchet A. et al, *in prep*). By inverting the relationships, it is possible to bring the measurements of any observed index to zero-velocity dispersion. At the end of this process, we obtain measurements for a set of 24 spectral indices, namely:

- Balmer indices: $\text{H}\delta_A$, $\text{H}\delta_F$, $\text{H}\gamma_A$, $\text{H}\gamma_F$, $\text{H}\beta$;
- Iron-dominated indices: Fe4383, Fe4531, Fe5015, Fe5270, Fe5335;
- Molecular indices: CN_1 , CN_2 , Mg_1 , Mg_2 ;

⁴For D4000 and D_n4000 , the molecular indices convention is used.

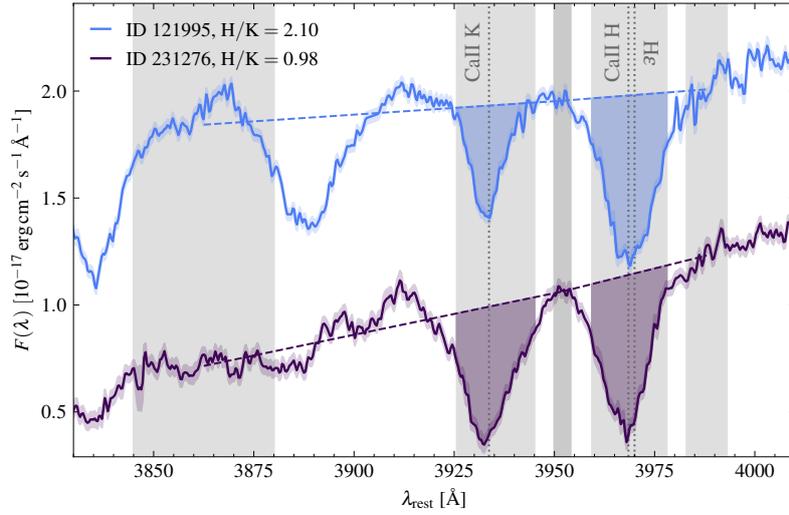


FIGURE 2.5: Measurement of pseudo-Lick indices Ca II K and Ca II H of a *bona fide* passive galaxy (violet) and a NUVrJ star-forming galaxy (blue, with a vertical offset of +0.35). Vertical bands show the regions where these indices are defined (red pseudo-continuum of Ca II H and blue pseudo-continuum of Ca II K are overlapped). Vertical dotted lines are the centroids of Ca II K, Ca II H, and H ϵ lines. Dashed segments show the estimated pseudo-continuum flux used as a reference to compute the indices values (filled regions).

- 4000 Å discontinuity: D4000, D_n4000;
- Additional Lick indices not available in LEGA-C Ca4227, G4300, Fe4531, C₂4668, Mg *b*;
- Two additional pseudo-Lick indices: Ca II K, Ca II H (Section 2.2.3).

This dataset extends LEGA-C public catalog from H β to Fe5406 indices and, in particular, Mg ones, commonly used as proxies to study the α enhancement. Table 2.2 presents the median correction coefficient $C(\sigma)$, expected redshift coverage, 5th–95th percentile range, median value, and median SNR for each measured index.

The study of the spectral features of this sample is limited by the unavailability of bluer (redder) indices at lower (higher) z . A limited set of features in the 3700–4900 Å range is available for all of the galaxies. About 200 galaxies (57%) commonly span the 3700–4900 Å wavelength range, limiting redshift coverage to $z \lesssim 0.9$. The typical SNR for most indices is above 10, except for those with negligible signals (Balmer and CN indices) or defined in narrow central regions (Ca4227 and Ca4455).

The H/K ratio

As firstly proposed by Rose (1984), the Ca II H over Ca II K ratio can be used as an age diagnostic, being particularly sensitive to starburst ages in post-starburst galaxies. In practice, the H ϵ absorption line ($\lambda = 3970.1$ Å), which is stronger in the presence of young A- and B-type stars, overlaps to Ca II H ($\lambda = 3968.5$ Å), while Ca II K ($\lambda = 3933.7$ Å) remains

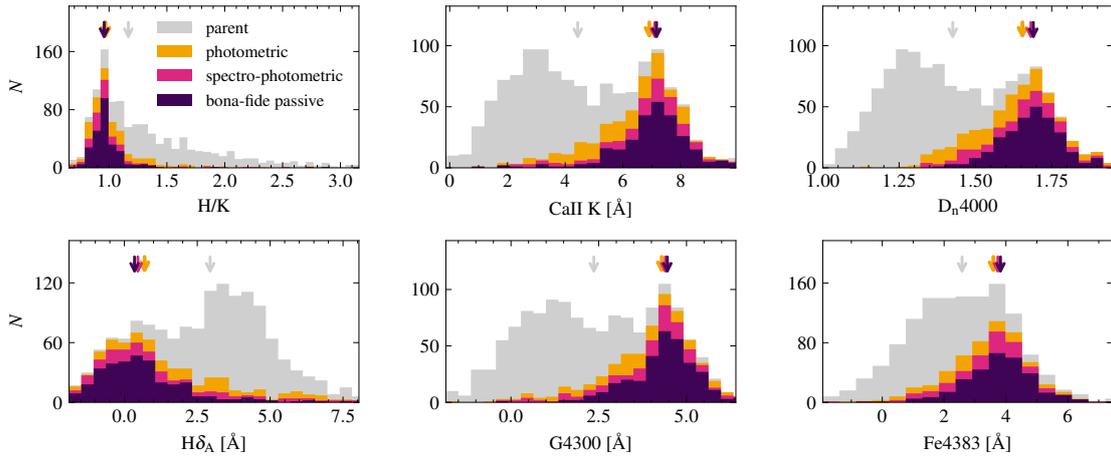


FIGURE 2.6: Distribution of LEGA-C galaxies as a function of the main spectral indices measured in this work. Different colors are used for the *parent* (gray), *photometric* (yellow), *spectrophotometric* (pink), and *bona fide passive* (violet) subsamples. Arrows represent the median values.

relatively uncontaminated. Therefore, the compact notation H/K actually represents the ratio $(\text{Ca II H} + \text{H}\epsilon)/\text{Ca II K}$. This diagnostic is particularly interesting as it is sensitive to even a small fraction of young ($\lesssim 1$ Gyr) stars (see Longhetti et al., 1999; Lonoce et al., 2014; Moresco et al., 2018). In the literature, it has been usually measured as the ratio of minimum fluxes in H and K lines (e.g., Rose, 1985; Leonardi & Rose, 1996; Longhetti et al., 1999; Lonoce et al., 2014; Moresco et al., 2018):

$$|\text{H/K}|_{\min} = \frac{F_{\min}(\text{H})}{F_{\min}(\text{K})}. \quad (2.15)$$

In this thesis, we adopt an alternative method for computing H and K values, which consists in measuring H and K values as pseudo-Lick indices:

$$\text{H/K} = \frac{I_{\text{H}}}{I_{\text{K}}}, \quad (2.16)$$

with the index regions defined in Fanfani (2019) that have also been included in PyLick, i.e. [3925.65 – 3945, 3845 – 3880, 3950 – 3954] for Ca II K and [3959.40 – 3978, 3950 – 3954, 3983 – 3993] for Ca II H. With this integrated approach, the H/K measurement is less affected by noise fluctuations with respect to flux minima. An illustrative example of the H/K measurement for a passive and a star-forming LEGA-C galaxy is presented in Figure 2.5. For a typical passive population, the H line is less deep than the K line, therefore $|\text{H/K}|_{\min} > 1$. The so-called H/K inversion can already arise for a contribution of $\sim 5\%$ in mass of a young stellar population (with ages < 200 Myr; Moresco et al. 2018). The two quantities defined in Eqs. 2.15 and 2.16 have an inverse relationship, but there is not a strict one-to-one correspondence. From the parent sample, we find an equivalent inversion value for the H/K between 1.2 and 1.5; therefore, the passive regime corresponding to $|\text{H/K}|_{\min} > 1$ can be safely defined at $\text{H/K} < 1.2$.

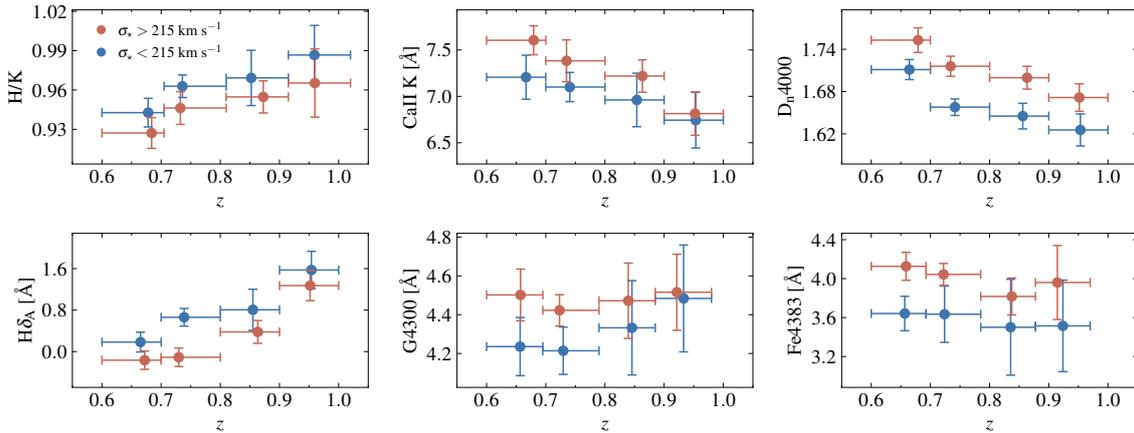


FIGURE 2.7: Binned relationships between the main spectral indices and redshift for the *bona fide* passive sample divided into two velocity dispersion regimes ($\sigma_{\star} = 215 \text{ km s}^{-1}$ as threshold). Each bin contains on average $N \sim 30$ galaxies. Error bars on the x -axis represent the bin width, while those on the y -axis are errors associated to the mean values.

2.2.4 Observed spectral features

The distribution of key absorption features in different subsamples of passive galaxies is shown in Figure 2.6, and median values are listed in Table 2.1.

As can be seen, *bona fide* passive galaxies distinctly separate from the *parent* sample median. They typically exhibit H/K ratios of 0.96 ± 0.08 , with few exceeding $\text{H/K} > 1.1$, whereas parent sample galaxies often exceed 1.2. Notably, even without imposing a H/K ratio cutoff, our selection criteria effectively exclude the majority of high H/K outliers, resulting in a sample consistent with a pure passive population. This underscores the robustness of our selection and minimal contamination from young stellar populations. For the first time, we report a clear bimodality in the Ca II K, with passive galaxies characterized by $\text{Ca II K} > 5 \text{ \AA}$. Hence, Ca II K is likely a primary contributor to the D4000 bimodality since it falls within its blue passband. This is not the case for Ca II H because, as mentioned above, the He line strengthens the index in younger populations. The well-known D_n4000 and $\text{H}\delta_A$ bimodalities have been extensively studied in the local universe (e.g., Kauffmann et al., 2003; Siudek et al., 2017) and in LEGA-C (Wu et al., 2018), with photometrically selected passive galaxies already populating high- D_n4000 and low- $\text{H}\delta_A$ tails. The addition of the spectroscopic criteria further reduces low- D_n4000 and high- $\text{H}\delta_A$ galaxies: *bona fide* passive galaxies, only 11 (3%) have $D_n4000 < 1.5$, and only 28 (8%) have $\text{H}\delta_A > 2.5 \text{ \AA}$. We observe a slight bimodality also for G4300 sensitive to carbon abundance (Tripicco & Bell, 1995; Korn et al., 2005), with *bona fide* passive galaxies characterized by a relatively high $\text{G4300} \gtrsim 2.5 \text{ \AA}$. Finally, they have relatively high $\text{Fe4383} \sim 3.8 \text{ \AA}$, a primary diagnostic of the average stellar metallicity.

We delve deeper into the *bona fide* passive sample by examining index- z - σ_{\star} trends. For this purpose, galaxies are firstly divided into two bins based on their stellar velocity dispersion using $\sigma_{\star} = 215 \text{ km s}^{-1}$ as a threshold. Then, each σ_{\star} bin is further divided into four or three z bins depending on the redshift coverage, considering intervals $\Delta z \sim 0.08 - 0.1$.

The final bins have $N \sim 30$ objects and in all cases $N > 10$. We calculate the mean and associated error for each bin and present the resulting trends in Figure 2.7. Similar trends are observed with an alternative binning scheme, optimizing the number of galaxies per bin by adjusting the redshift intervals. However, this approach comes at the cost of reduced z leverage due to the uneven redshift distribution of LEGA-C. Additionally, the trends within individual galaxies are quantified using the nonparametric Spearman rank correlation test.

In general, we observe moderate correlations between Balmer indices and z , with a Spearman $\rho \sim 0.4$ (p -value 10^{-8}). At fixed z , galaxies with higher σ_* have weaker Balmer and H/K features. On the other hand, D_n4000 exhibits a moderate anticorrelation with z , with a Spearman $\rho \sim -0.3$ (p -value 10^{-5}). Interestingly, a clear separation is found between the two σ_* regimes. It is worth noting that these relationships have been, and can be, employed within the cosmic chronometers framework to constrain $H(z)$, once the D4000–age relation is accurately calibrated (Moresco et al., 2011). Temporarily neglecting the effects of metallicity, these observed trends align perfectly with the mass-downsizing concept, i.e. more massive galaxies formed earlier and in relatively short timescales, and subsequently underwent passive evolution. This idea is further supported by the iron indices (e.g. Fe4383), which appear stronger in high- σ_* galaxies but do not exhibit significant evolution over ~ 2 Gyr of cosmic time. This finding supports the idea that the population of selected passive galaxies has exhausted its gas reservoir, rendering it incapable of undergoing substantial changes in its metallic content, and is instead undergoing passive evolution as a function of cosmic time.

Given the previous remarks, the absence of a redshift correlation for C_24668 and $G4300$ suggests that these indices are not reliable age indicators. However, the segregation in σ_* is still consistent with the idea that they are sensitive to stellar population age in the very first Gyr after the formation. Alternatively, the segregation could be attributed to variations in metal abundances. Even in this scenario, these indices could play a crucial role in indirectly determining galaxy ages by breaking the age–metallicity degeneracy.

Morphology

We study the morphological properties of the *bona fide* passive sample using the Zurich Estimator of Structural Types (ZEST) classification (Scarlata et al., 2007), based on principal component analysis of the surface brightness profiles. Most galaxies (71%) are E/S0 types, 27% fall into the intermediate category, and only 2% are irregular, with no late-type galaxies. Interestingly, a similar percentage of E/S0 types (72%) has also been found by Moresco et al. (2013) in their sample of ~ 17000 zCOSMOS galaxies. The two works share 127 galaxies (37% of the *bona fide* passive sample) but Moresco et al. (2013) use different selection (both color and spectroscopic [O II] and $H\alpha$) cuts.

The presence of passive systems with non-purely early-type morphologies has already been discussed in the literature (e.g., Dressler et al., 1999; Pozzetti et al., 2010) and can be explained by the existence of a class of objects where the morphological transformation occurs over a more extended timescale than changes in stellar populations, i.e. galaxy colors redden before the galaxy reaches an early-type morphology. It is therefore improper to classify galaxies without an early-type morphology as contaminants. In this work, the intermediate and irregular galaxies are still included in the analysis as they have passed the selection cuts and have spectra typical of old stellar populations. Further analysis has also confirmed that excluding the irregulars does not have a significant impact on the final results.

In conclusion, SED fitting (Figure 2.2) and observed spectral properties (Figures 2.3, 2.6, 2.7) confirm a passive galaxy population characterized by reduced star formation, older stellar populations, and higher metallicities compared to the parent galaxy population, consistently with the mass-downsizing scenario. These observational data also support that the *bona fide* selection adopted is able to maximize the purity of the sample, providing a sample of massive and passive galaxies, with negligible (if any) contamination by star-forming outliers. Additional insights will be gained through the analysis of each individual galaxy, providing a more granular understanding of the overall picture.

2.3 Physical properties of passive galaxies at intermediate z

This work extends the cosmic chronometer method by using multiple spectral features to constrain the ages of galaxies. This approach opens up the possibility of astrophysical studies. In particular, we provide age, $[Z/H]$, and $[\alpha/Fe]$ constraints for a large set of individual passive galaxies at higher z with respect to current literature. We begin by presenting the adopted stellar models (§ 2.3.1) and the statistical framework (§ 2.3.2). Next, we discuss the initial results obtained for the H/K ratio as a stellar population diagnostic tool (§ 2.3.3). We then delve into the study of scaling relations (§ 2.3.4) and examine the redshift trends of age, $[Z/H]$, and $[\alpha/Fe]$ (Section 2.3.5), which are crucial for the cosmic chronometer study. Finally, we repeat the analysis by varying the set of spectral features (Section 2.3.6) to investigate potential systematics.

2.3.1 Stellar population models

We adopt simple stellar populations (SSP) models by [Thomas, Maraston & Johansson \(2011\)](#) (hereafter [TMJ11](#)) that provide Lick indices estimates by varying stellar ages (age), metallicities ($[Z/H]$), and α -element enhancements ($[\alpha/Fe]$). These models have been widely employed in similar Lick indices studies (e.g., [Johansson et al., 2012](#); [Jørgensen & Chiboucas, 2013](#); [Onodera et al., 2015](#); [Scott et al., 2017](#); [Lonoce et al., 2020](#)). They are based on the evolutionary synthesis code of [Maraston \(2005\)](#) and element response functions from [Korn et al. \(2005\)](#), and are carefully calibrated against data from galactic globular clusters. The main ingredients are single-burst SFH, [Salpeter \(1955\)](#) initial mass function, MILES empirical stellar libraries ([Sánchez-Blázquez et al., 2006](#)), and [Cassisi et al. \(1997\)](#) stellar evolutionary tracks. Specifically, we use the models provided at the MILES resolution of 2.5 Å ([Beifiori et al., 2011](#)).

The original parameter grid spans $0.1 < \text{age/Gyr} < 15$, $-2.25 < [Z/H] < 0.67$, and $-0.3 < [\alpha/Fe] < 0.5$ with 480 grid points, each one corresponding to the prediction for a single SSP. We perform a three-dimensional linear interpolation of the grid obtaining a resolution of $\Delta \text{age} = 0.1$ Gyr, and $\Delta[Z/H] = \Delta[\alpha/Fe] = 0.01$ dex. This process allows to achieve a higher numerical precision in parameter estimation while maintaining the best-fit values fully compatible with each other within 1σ , as verified in our final results.

It is worth noting that the SSP model assumption of a single instantaneous burst of star formation might not capture the full complexity of individual galaxy formation pathways. While this assumption could potentially lead to a significant underestimation of the global stellar age in mixed populations, it is crucial to highlight that our selection criteria were chosen to obtain a sample with short formation timescales and to minimize the contamination from a significant residual star formation, as discussed in the analysis of the various spectral features in Section 2.2.4. In particular, when a young stellar population is present, mass-weighted age - which can be derived with full spectral fitting codes - can provide a better estimate of the integrated galaxy SFH. From the analysis of the H/K ratio (Sect. 2.2.3) we can exclude contamination from young (200 Myr–1 Gyr) stars in our *bona fide* passive galaxies, therefore also the difference between light- and mass-weighted is expected to be negligible. Finally, for this population we expect the SFHs to be extremely coeval, with very

short timescales ($\tau \lesssim 0.3$ Gyr, if modeled with a delayed exponential SFH), as corroborated from a parallel analysis.

2.3.2 MCMC Analysis

To compare the measured absorption features to [TMJ11](#) models, we develop a fully Bayesian analysis pipeline. We assume that the uncertainties on indices are well-determined, Gaussianly distributed, and independent. A set of model indices, which are a function of parameters $\boldsymbol{\theta} = (\text{age}, [\text{Z}/\text{H}], [\alpha/\text{Fe}])$, can therefore be fitted to the observed ones using the log-likelihood function

$$\ln(\mathcal{L}) = k - \frac{1}{2} \sum_i^{\text{N}_{\text{ind}}} \left(\frac{I_i - I_i^{\text{mod}}(\boldsymbol{\theta})}{\sigma_i} \right)^2, \quad (2.17)$$

where k is a constant, $I_i^{\text{mod}}(\boldsymbol{\theta})$ the model prediction for the i th observed index I_i , and σ_i its uncertainty. The posterior probability distributions of $\boldsymbol{\theta}$ are explored using the affine-invariant ensemble sampler [emcee](#) ([Foreman-Mackey et al., 2019](#)). Chains are initialized with 200 walkers randomly scattered around the center of the parameter space. Each walker performs at least 2000 model realizations. Parameters and uncertainties are defined as the median, and 16th and 84th percentiles of the marginalized posterior distributions.

We use flat priors that span the entire parameter space allowed by the models. An important point that we stress here is that we do not assume any cosmological prior for galaxy ages. This is a crucial point, and a difference with respect to other similar works, to avoid introducing cosmological biases in the age determination and keep the results cosmological-independent. After the analysis, we carefully assess the convergence of each chain.

Assessing convergence of MCMC posterior distributions

In a Bayesian analysis, it is crucial to determine whether MCMC chains are reproducing with sufficient accuracy the target posterior distribution. However, there is not an established standard to assess the convergence ([Hogg & Foreman-Mackey, 2018](#); [Roy, 2020](#)). One possibility involves an examination of the autocorrelation time τ_{int} , which quantifies how many steps are needed to obtain independent samples.

We consider a chain to be formally converged when τ_{int} for each parameter is greater than 1/100th of the chain size. In this way, the analyzed galaxies typically require ~ 7000 steps. Nonetheless, sample medians and associated errors computed from formally converged chains could still be unreliable. This is the case when the posterior distributions:

1. are skewed toward the priors (in this work, the whole parameter space allowed from the models);
2. are not constraining (e.g., due to high age-metallicity degeneracy);
3. are multimodal.

While (2) and (3) may be partially alleviated by choosing a different estimator with respect to the sample median, (1) requires a deeper study on potential observational issues and/or the capability of stellar population models to accurately reproduce all types of galaxies.

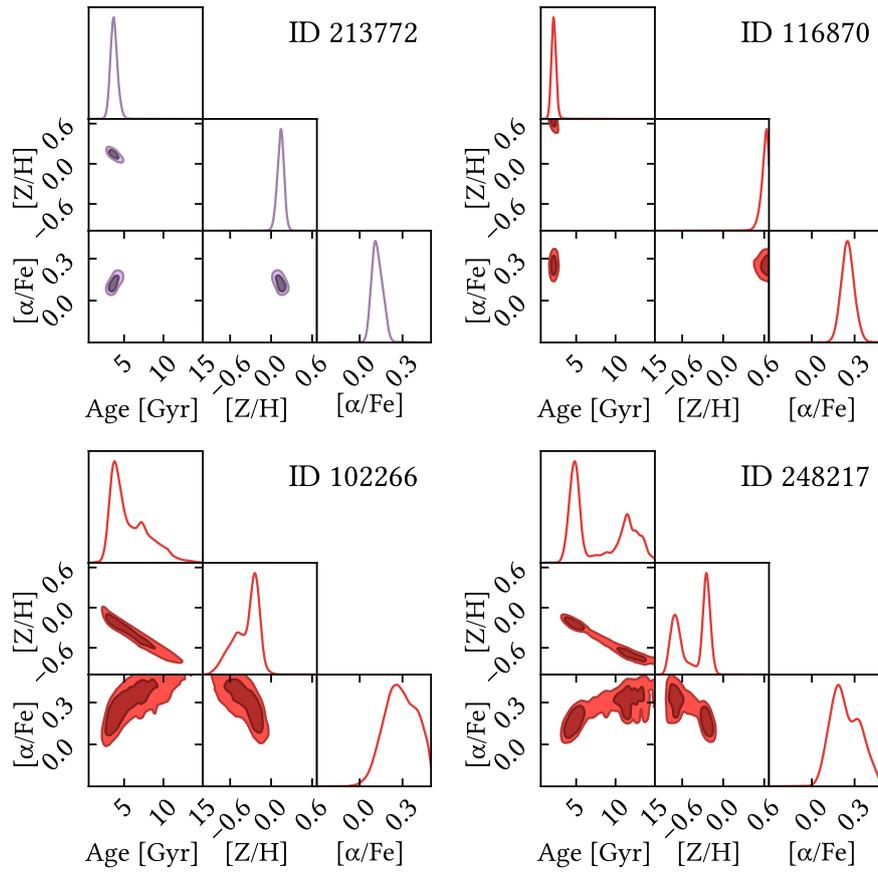


FIGURE 2.8: Examples of four contour plots that may generate reliable (violet) and unreliable (red) results. The axes span the full parameter space allowed by the stellar evolution models adopted. The contours enclose 1 and 2σ confidence regions.

Figure 2.8 shows an example for each category and a typical contour plot for a good fit. Notably, the age-metallicity degeneracy adheres to the "3/2 rule" (Worthey, 1998): a twofold increase (decrease) in log age, coupled with a threefold decrease (increase) in stellar metallicity ($[Z/H]$), can reproduce the same observed indices. The multimodal distributions occur in less than 4% of the galaxies always along the age-metallicity degeneracy direction and vanish when using different index sets. Therefore they should be considered a special case of (2) and attributed to intrinsic degeneracies rather than genuine complexity in the stellar population.

After checking joint and marginal distributions of all the analyzed galaxies, those falling into (1–3) categories are flagged and excluded. Around 85% of the excluded galaxies have a low overall spectral S/N (<25). However, the application of this cut *a priori* would have reduced the final sample by half, also excluding well-fitted cases. After a more thorough analysis, we find that the SNR of each individual index used in the MCMC analysis is an effective indicator for distinguishing between included and excluded galaxies. This applies specifically to the higher SNR indices, as listed in Table 2.2: G4300, Fe4383, Fe4531, and C₂4668. Approximately 95% of the excluded galaxies possess at least one index with $S/N \lesssim 10$. It is worth noting that the precise value for this threshold is not universal, as it may vary for different combinations of indices.

In summary, we find that the inspection of MCMC posterior distributions is an efficient procedure to detect unreliable constraints—mostly due to low-SNR indices—while maximizing the number of galaxies for which we obtain reliable constraints.

Index combination

The choice of the index set to use in the analysis must be carefully addressed knowing that, given the number of the measured indices, there are more than one million possible combinations. Different spectral features have different sensitivity to stellar population parameters and abundances (Tripicco & Bell, 1995; Korn et al., 2005; Lee et al., 2009). The Balmer and D_n4000 features are better suited to constrain ages, Fe-dominated indices to measure the Fe abundance and total stellar metallicity, and Mg indices to estimate the α elements abundance. It is therefore expected that different index sets could provide slightly different constraints.

To ensure an unbiased age – z relation, we consistently apply a fixed set of indices for the analysis of all galaxies, constraining the available wavelength range to approximately 3600–4900 Å; see Figure 2.3 and Table 2.2 for details. Simultaneously, it is crucial to ensure that galaxies span a broad redshift span to effectively trace the differential age evolution of the population for the application of the CC method, in this case $0.6 < z < 0.9$. The combined trade-off reduces our measured galaxy count by approximately one-third but guarantees a homogeneous analysis.

The baseline set of indices used in this study is:

$$H\delta_A, CN_1, CN_2, Ca4227, G4300, H\gamma_A, H\gamma_F, Fe4383, Fe4531, \text{ and } C_24668$$

. These indices are chosen among those calibrated against globular cluster data in TMJ11, excluding those redder than $H\beta$ to ensure a statistically meaningful sample and $H\beta$ due to potential biases from residual emission components (see Concas et al., 2017). They span a narrow optical wavelength range from 4000 to 4800 Å. We have tested the inclusion of Mg b and found it does not significantly alter the results. Extensive tests with different index sets

have been conducted (see Appendix 2.1.4), and this set optimizes both spectral coverage and the number of galaxies for obtaining constraints. It also offers a good balance between age, metallicity, and α -sensitivity.

Ultimately, we derive robust constraints for 140 out of 199 analyzed galaxies after removing those with nonconverging fits (see Sect. 2.3.2). An extract of the catalog is presented in Table 2.3.2. These galaxies are at $\langle z \rangle = 0.70$ with a median spectral $\langle \text{SNR} \rangle = 26.4$ per pixel. Typical uncertainties are 0.33 Gyr in age and 0.05 dex in $[\text{Z}/\text{H}]$ and $[\alpha/\text{Fe}]$.

In the subsequent sections, the results are structured as follows: Section 2.3.3 discusses the H/K ratio and its correlation with commonly used diagnostics; Section 2.3.4 explores scaling relations of stellar population parameters versus σ_* and M_* ; Finally, Section 2.3.5 presents trends with redshift, with a specific focus on the age–redshift relation. Here are the key points to bear in mind when comparing our analysis results with previous studies:

1. Our galaxies are passive. Similar studies typically focus on morphological early-type galaxies, which may include galaxies with some level of ongoing star formation (around $\sim 20\%$ at the current stellar masses; see Moresco et al. 2013).
2. We do not employ cosmological priors. We have verified that, particularly for low-S/N galaxies, imposing a cosmological upper limit on galaxy ages can lead to an apparent convergence of MCMC chains towards the prior at higher ages, potentially introducing a bias into the sample (see Section 2.3.2).

Before applying the cosmic chronometer method, it is essential to carefully consider (1) and (2). However, we emphasize that the method relies on differential - not absolute - galaxy ages. Therefore, a more extended SFH for the entire passive galaxy population (i.e. a vertical offset in the age – z plane), does not affect the ultimate $H(z)$ measurement (see Section 2.3.1).

2.3.3 The H/K ratio as a stellar population diagnostic

In Figure 2.9 we show the distribution of the entire parent sample in four widely used diagnostic diagrams, namely NUVrJ, UVJ, EW[O II]– D_n4000 , and SFR– M_* , color-coded by the H/K ratio. To better capture the mean trends of H/K, we perform locally weighted regression (LOESS). We use the LOESS package of Cappellari et al. (2013) based on the two-dimensional algorithm of Cleveland & Devlin (1988), with a regularization factor $f = 0.5$. The selected *bona fide* passive galaxies are highlighted with black borders. The Spearman rank correlation coefficients are computed to quantify the trends.

The diagrams reveal a clear correlation between the H/K ratio and other diagnostic tools despite its narrow dynamic range ($0.8 \lesssim \text{H}/\text{K} \lesssim 2$, and a median error of ~ 0.14). In more detail, we observe a strong correlation with (NUV– r) ($\rho = -0.72$; p -value = 10^{-164}) and with (U – V) ($\rho = -0.72$; p -value = 10^{-178}). The correlation with (r – J) and (V – J) colors, while weaker, remains statistically significant ($\rho = -0.25$; p -value = 10^{-19} and 10^{-15} , respectively). Interestingly, we find that a selection based on the threshold value $\text{H}/\text{K} < 1.1$ can reproduce the NUVrJ selection with 17% incompleteness and 19% contamination.

As shown in the third panel, implementing an H/K cut does not exclude galaxies with a significant [O II] emission. Specifically, 19% of the plotted galaxies have $\text{H}/\text{K} < 1$ and

Table 2.3. Resulting stellar population parameters for 140 passive galaxies selected in LEGA-C. The measurements and uncertainties are reported as the median and 16–84th percentile range.

ID	RA (J2000)	DEC (J2000)	age [Gyr]	[Z/H]	[α /Fe]
133240	150 09 41.4	2 21 44.9	$3.08^{+0.64}_{-0.64}$	$+0.26^{+0.12}_{-0.09}$	$+0.16^{+0.06}_{-0.06}$
133783	150 16 44.5	2 22 06.3	$1.90^{+0.07}_{-0.08}$	$+0.33^{+0.05}_{-0.03}$	$+0.14^{+0.03}_{-0.03}$
134169	150 17 56.0	2 22 20.7	$2.79^{+0.36}_{-0.36}$	$+0.16^{+0.08}_{-0.07}$	$+0.24^{+0.05}_{-0.05}$
139772	150 16 00.7	2 26 21.9	$3.05^{+0.14}_{-0.12}$	$+0.12^{+0.02}_{-0.02}$	$+0.23^{+0.03}_{-0.03}$
205742	150 05 19.5	2 27 38.3	$3.07^{+0.39}_{-0.27}$	$-0.01^{+0.06}_{-0.04}$	$+0.04^{+0.03}_{-0.03}$
206573	150 16 01.8	2 28 11.2	$3.30^{+0.29}_{-0.19}$	$-0.02^{+0.02}_{-0.03}$	$+0.28^{+0.03}_{-0.03}$
207825	150 17 40.8	2 29 19.3	$1.94^{+0.12}_{-0.07}$	$+0.16^{+0.01}_{-0.05}$	$+0.16^{+0.02}_{-0.02}$
...					

Note. — The full table is released in [Borghi et al. \(2022b\)](#).

$\text{EW}[\text{O II}] > 5 \text{ \AA}$. Unfortunately, the absence of other spectral features as $\text{H}\alpha$, $[\text{N II}]$, and $[\text{S II}]$ due to the restricted wavelength coverage of the current data set does not allow us to investigate the nature of these sources. For these systems, the combination of multiple indicators remains essential to obtain a pure sample of passive galaxies.

A strong correlation is also observed for the sSFR ($\rho = 0.70$; $p\text{-value} = 10^{-161}$). We find that a threshold on H/K values of $\text{H}/\text{K} < 1.1$ can reproduce a $\log \text{sSFR}/\text{yr} < -11$ cut with 15% incompleteness and 16% contamination.

These are remarkable results considering the wavelength coverage requirements for the different selections. On one hand, wide photometric coverage is needed for a reliable estimate of a NUVrJ diagram, SFR, or M_* (typically from the UV to the near-IR), with accuracy increasing with the number of available photometric points; it is therefore not always available in many surveys. In contrast, the H/K feature is defined over a much narrower wavelength window of about 150 \AA , and it necessitates deep rest-frame optical spectroscopy. The H/K can therefore play a key role in the selection of pure samples of passive galaxies in future wide-field grism surveys such as Euclid ([Laureijs et al., 2011](#)) and the Roman Space Telescope ([Spergel et al., 2015](#)). Another advantage of this diagnostic is the mild dependence on the spectral resolution. Differences in H/K values obtained on 8 and 2.5 \AA FWHM spectra are $\lesssim 4\%$, while the same analysis on individual Ca II K and Ca II H indices yields 10% differences.

Concerning the 140 *bona fide* passive galaxies, we note here that no correlation between stellar population parameters (especially age) and H/K is present. This result is expected given the fact that no significant contribution from a young stellar component is present, i.e. the H ϵ line deepening effect becomes negligible in the passive regime.

In summary, young ($\sim 200 \text{ Myr}$) stellar populations whose light is predominantly due to A- and B-type stars ($\text{H}/\text{K} > 1.1$) are characterized by higher UV fluxes, lower D4000, a

higher EW[O II], and dominate the star formation main sequence of LEGA-C galaxies. The H/K lowers when approaching the quiescence criteria, but the study of timescales and the interplay with stellar population parameters for the whole population of galaxies will require further assessment.

2.3.4 Physical parameters versus σ_* and M_*

Figure 2.10 shows the derived stellar population parameters as a function of the observed stellar velocity dispersion σ_* for passive galaxies at $z \sim 0.7$. As a local reference, we utilize the SDSS/MOSES dataset from Thomas et al. (2010), consisting of morphologically selected early-type galaxies at $z \sim 0.05$. The choice is based on the similarity in models and analysis techniques employed. The local dataset, being approximately 20 times larger, covers a broader range in $\log \sigma_*$ (1.7–2.5). To ensure a correct comparison, the local sample is limited to comparable σ_* values, i.e. $2.1 < \log \sigma_* [\text{km s}^{-1}] < 2.5$. Finally, the catalog is supplemented with M_* estimates from MPA-JHU DR 8 (Kauffmann et al., 2003). To study \log age, $[Z/H]$, and $[\alpha/\text{Fe}]$ versus $\log \sigma_*$ and $\log M_*$ we perform robust linear regression with the least trimmed squares (LTS) algorithm (Rousseeuw, 1984) and measure their Spearman coefficients. Results are quoted in Table 2.4.

In general, the observed σ_* trends align with passive evolution from $z \sim 0.7$ to $z \sim 0.05$. At $z \sim 0.7$, we find shallower scaling relations, which may be ascribed to the smaller statistics and/or different selection criteria adopted. Additionally, the dynamic range in age naturally decreases with increasing redshift because the Universe gets younger.

We start by examining SSP-equivalent stellar ages. Remarkably, the 5.5 Gyr difference between the two samples aligns perfectly with the age evolution of the Universe, confirming that these stellar populations underwent a purely passive evolution under the assumption of the reference cosmology. Contamination from young stellar populations in the high- z (low- z) sample would result in a larger (smaller) offset. We find mild correlations between \log age vs. $\log \sigma_*$ (with a slope of 0.5 ± 0.1) and \log age vs. $\log M_*$ (with a slope of 0.19 ± 0.04). To facilitate comparisons with other studies, we convert galaxy ages into formation redshifts z_{form} and investigate their relation with M_* , obtaining:

$$z_{\text{form}} = (0.40 \pm 0.05) \log_{10} \left(\frac{M_*}{10^{11} M_{\odot}} \right) + (1.46 \pm 0.02), \quad (2.18)$$

with an intrinsic scatter of approximately 0.24. This means that galaxies with higher stellar mass ($\log M_*/M_{\odot} = 11.3$) formed their stars at $z_{\text{form}} \sim 1.6$, while less massive ones ($\log M_*/M_{\odot} = 10.7$) formed their stars at $z_{\text{form}} \sim 1.3$. Intriguingly, the formation epoch of these passive galaxies occurred shortly after the peak of cosmic star formation rate density at $z \sim 2$ (Madau & Dickinson, 2014). We also identify a small subset (22 galaxies, roughly 16% of the passive sample) of very massive ($\log M_*/M_{\odot} > 11$) galaxies with $z_{\text{form}} > 2.5$ up to 5.

In Figure 2.11, we compare our results to similar studies in the literature. We closely match the findings of Jørgensen & Chiboucas (2013), who studied ~ 80 early-type cluster galaxies at $z = 0.5 - 0.9$ with TMJ11 models and found formation redshifts of around $z_{\text{form}} \approx 1.24$ and 1.95 for stellar masses of $\log M_*/M_{\odot} \approx 10.6$ and 11.4, respectively. We also find very good agreement with Choi et al. (2014) and Gallazzi et al. (2014). In more detail, Choi et al. (2014)

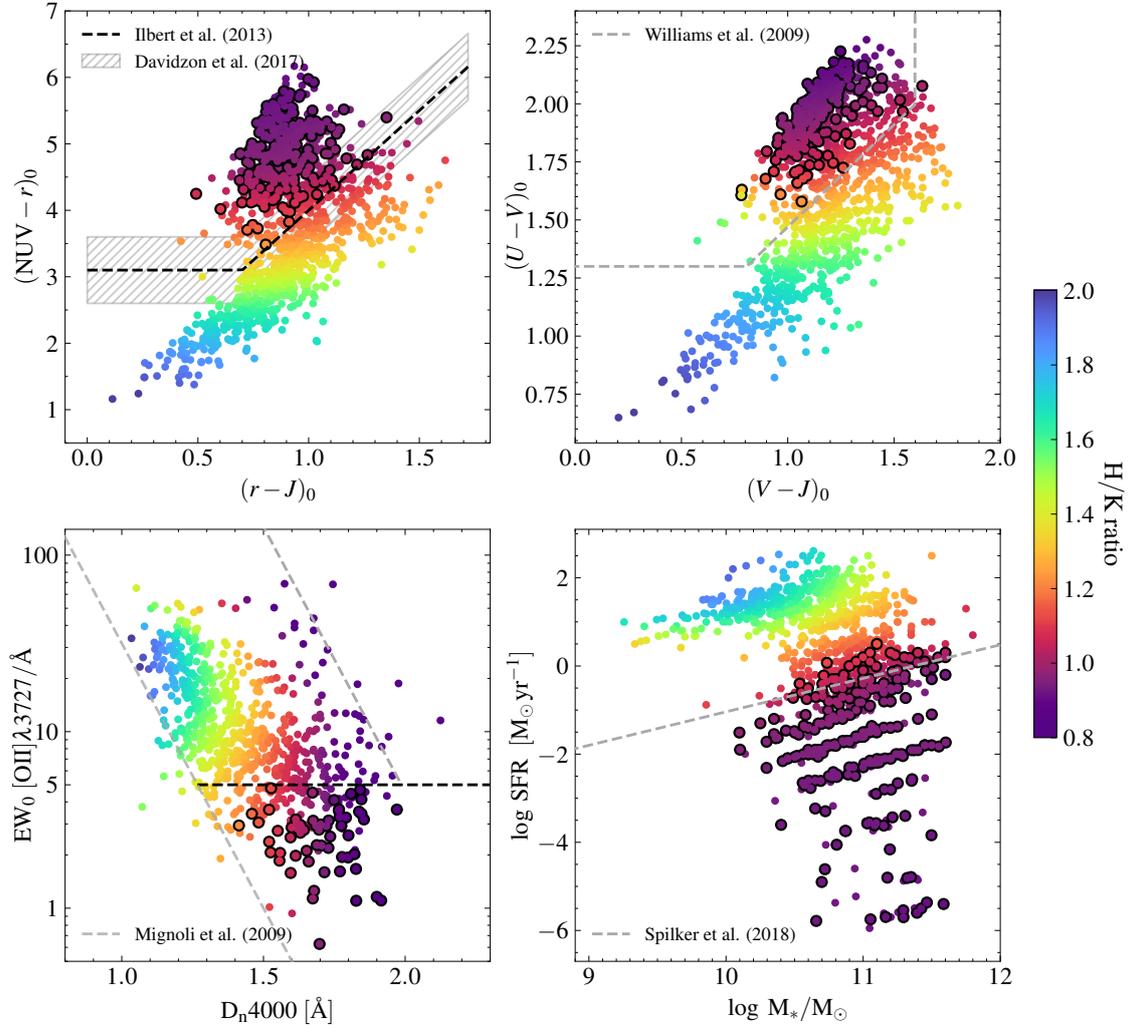


FIGURE 2.9: Four diagnostic diagrams color-coded by the H/K ratio using LOESS smoothing (see text). Each panel shows the *parent sample* galaxies derived from LEGA-C DR2 with a good H/K measurement, $\text{SNR}(\text{H}/\text{K}) > 3$. The black borders identify the selected *bona fide* passive galaxies. The dashed lines are the criteria to separate star-forming from passive galaxies taken from the literature (Ilbert et al., 2013; Williams et al., 2009; Mignoli et al., 2009; Spilker et al., 2018, respectively). Black colors indicate the criteria adopted in this study, while gray lines are for illustrative purposes only.

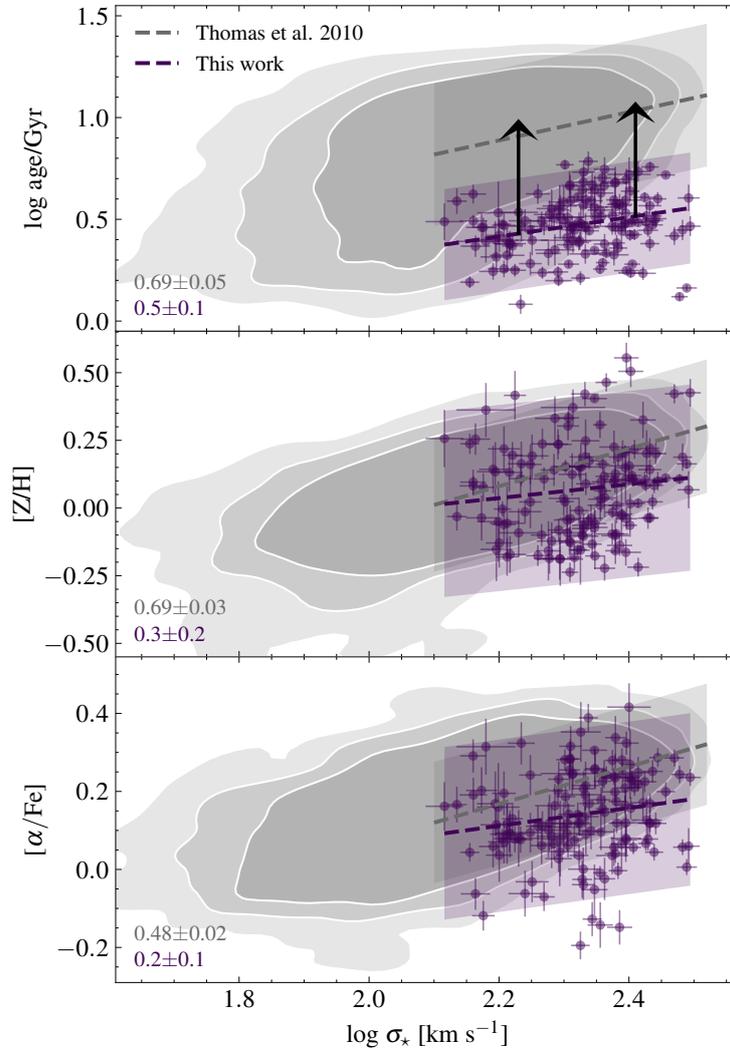


FIGURE 2.10: Distribution of mean stellar ages, metallicities, and α enhancements as functions of stellar velocity dispersion for $z \sim 0.7$ passive galaxies selected in LEGA-C (violet points). Gray contours represent early-type galaxies in the local Universe from SDSS/MOSES (Thomas et al., 2010), enclosing 1σ , 2σ , and 5σ regions. Dashed lines and shaded regions are robust linear fits and associated 2σ scatter regions, for $2.1 < \log \sigma_*$ < 2.5 galaxies. The resulting slopes and their uncertainties are annotated in the bottom left. Black arrows represent the age evolution expected for a passive evolution.

Table 2.4. Scaling relations for the $z \sim 0.7$ passive galaxies selected in LEGA-C.

y	$a \pm \text{err}(a)$	$b \pm \text{err}(b)$	rms	ρ (p -value)
$y = a \log \sigma_* + b$				
log age/Gyr	0.5 ± 0.1	-0.6 ± 0.3	0.1	0.2 (0.02)
[Z/H]	0.3 ± 0.2	-0.5 ± 0.4	0.2	0.2 (0.15)
$[\alpha/\text{Fe}]$	0.2 ± 0.1	-0.4 ± 0.3	0.1	0.2 (0.03)
$y = a \log(M_*/10^{11}M_\odot) + b$				
log age/Gyr	0.19 ± 0.04	0.47 ± 0.01	0.1	0.3 (< 0.01)
[Z/H]	0.03 ± 0.05	0.07 ± 0.02	0.2	0.0 (0.63)
$[\alpha/\text{Fe}]$	0.02 ± 0.03	0.14 ± 0.01	0.1	0.1 (0.50)

analyzed stacked spectra of sSFR-selected passive galaxies at comparable z , finding typical SSP formation epochs of $z_{\text{form}} \sim 1.5$ using full spectral fitting. Gallazzi et al. (2014) analyzed ages and stellar metallicities for ~ 70 between star-forming and quiescent galaxies at $z \sim 0.7$ using age-sensitive D_n4000 , $H\beta$, and $H\delta_A + H\gamma_A$, and metal-sensitive $[\text{Mg}_2\text{Fe}]$ and $[\text{MgFe}]'$ indices. Finally, by rewriting Equation 2.18 in terms of formation time t_{form} versus $\log M_*$, we obtain a trend of -1.26 ± 0.27 Gyr per decade in mass in excellent agreement with Carnall et al. (2019) (-1.48 ± 0.37 Gyr per decade in mass), despite their different methodology (full spectral fitting) and star formation history (double power-law) assumptions.

We now move to the analysis of mean stellar metallicities. Our results show no significant evolution in $[\text{Z}/\text{H}]$ for passive galaxies since $z \sim 0.7$, with a median offset of 0.05 dex comparable to the median uncertainty, supporting and strengthening earlier findings at similar z Gallazzi et al. (2014) and at earlier epochs (Onodera et al., 2012, 2015; Citro et al., 2016; Estrada-Carpenter et al., 2019). On the other hand, recent studies report a significant evolution of $\Delta[\text{Z}/\text{H}] > 0.1$ dex. An example is the work by Beverage et al. (2021) based on a sample of 65 LEGA-C quiescent galaxies analyzed with a full spectral fitting code. The authors found no evolution in $[\text{Mg}/\text{Fe}]$ values, but a $\Delta[\text{Fe}/\text{H}]$ (hence $\Delta[\text{Z}/\text{H}]$) of about 0.2 dex with respect to local $\log(M_*/M_\odot) = 11$ quiescent stacks. When examining $[\text{Z}/\text{H}]$ against $\log \sigma_*$ and $\log M_*$, we find no significant correlations. This can be attributed to our strict selection criteria for the most passive systems, which tend to exhibit shallower metallicity-mass relations (see Peng et al., 2015; Gallazzi et al., 2014, 2021). Indeed, when splitting them into two σ_* groups, we observe that systems with $\sigma_* > 215 \text{ km s}^{-1}$ typically have $[\text{Z}/\text{H}]$ values 0.1 dex higher than those with lower σ , consistent with the downsizing scenario.

Finally, we analyze mean stellar α enhancements. Our galaxies show $[\alpha/\text{Fe}]$ values about 0.1 dex lower than the local sample. However, as discussed in Section 2.1.4, this could be completely explained by the fact that we could not use Mg indices in our baseline analysis. For the first time, we observe a positive correlation between $[\alpha/\text{Fe}]$ and $\log \sigma_*$ with a slope of 0.2 ± 0.1 among individual passive galaxies at $z \sim 0.7$. This finding aligns with trends seen in stacks of quiescent galaxies at $0.1 < z < 0.7$ by Choi et al. (2014). However, we do not detect a significant trend with $\log M_*$. On one hand, local early-type galaxies show stronger

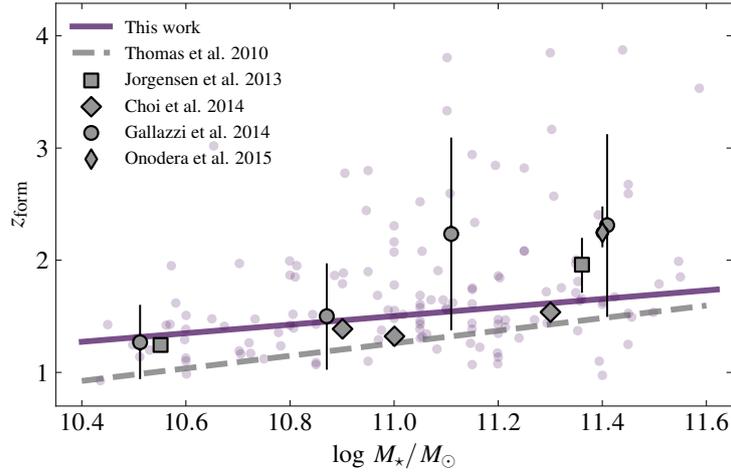


FIGURE 2.11: Formation redshifts as a function of galaxy stellar mass for 140 *bona fide* passive galaxies selected in LEGA-C DR2 (violet points). We compare our measurements with literature data of massive quiescent galaxies in the local (Thomas et al., 2010) and intermediate-redshift (Jørgensen & Chiboucas, 2013; Choi et al., 2014; Gallazzi et al., 2014; Onodera et al., 2015) Universe. Lines are obtained by performing robust linear fits (see text).

correlations with the gravitational potential well (of which σ_* is a tracer) than with M_* (e.g., Thomas et al., 2005; Barone et al., 2018). Nevertheless, uncertainties in M_* derivation may play a role in washing out these relations.

2.3.5 Physical parameters versus z

Figure 2.12 shows the derived stellar population parameters as a function of redshift. Interestingly, despite not having imposed any cosmological prior on the age of the galaxies, our derived ages are in all cases in agreement with a generic cosmological model, never exceeding the age of the Universe at any redshift. Moreover, the upper envelope of the distribution follows the decrease expected from the aging of the Universe. Across all redshifts, only a few passive galaxies exhibit ages below $\lesssim 2$ Gyr due to the stringent selection criteria. The median age value is $\langle \text{age} \rangle = 3.01$ Gyr, with a 1σ scatter of ± 0.97 Gyr. This implies a formation time of around $t_{\text{form}} \sim 4$ Gyr after the Big Bang, corresponding to a formation redshift of $z_{\text{form}} \sim 1.5$, as previously discussed in Section 2.3.4.

Stellar metallicities have solar or slightly supersolar values, $\langle [Z/H] \rangle = 0.08$ dex with a ± 0.18 1σ scatter, spanning a very narrow range if compared to the initial parameter space ($-2.25 < [Z/H] < 0.67$). Differently from stellar ages, they show no sign of evolution even within the redshift range explored in this work.

We find slightly supersolar $[\alpha/\text{Fe}]$ values, $\langle [\alpha/\text{Fe}] \rangle = 0.13$ dex with a ± 0.11 1σ scatter. In particular, 124 (89%) galaxies have $[\alpha/\text{Fe}] > 0$, pointing to very short formation timescales, i.e. before Type Ia supernova explosions can pollute the interstellar medium with a relatively high amount of iron-peak elements. As for the metallicities, these star formation timescale tracers do not significantly evolve over the probed z range.

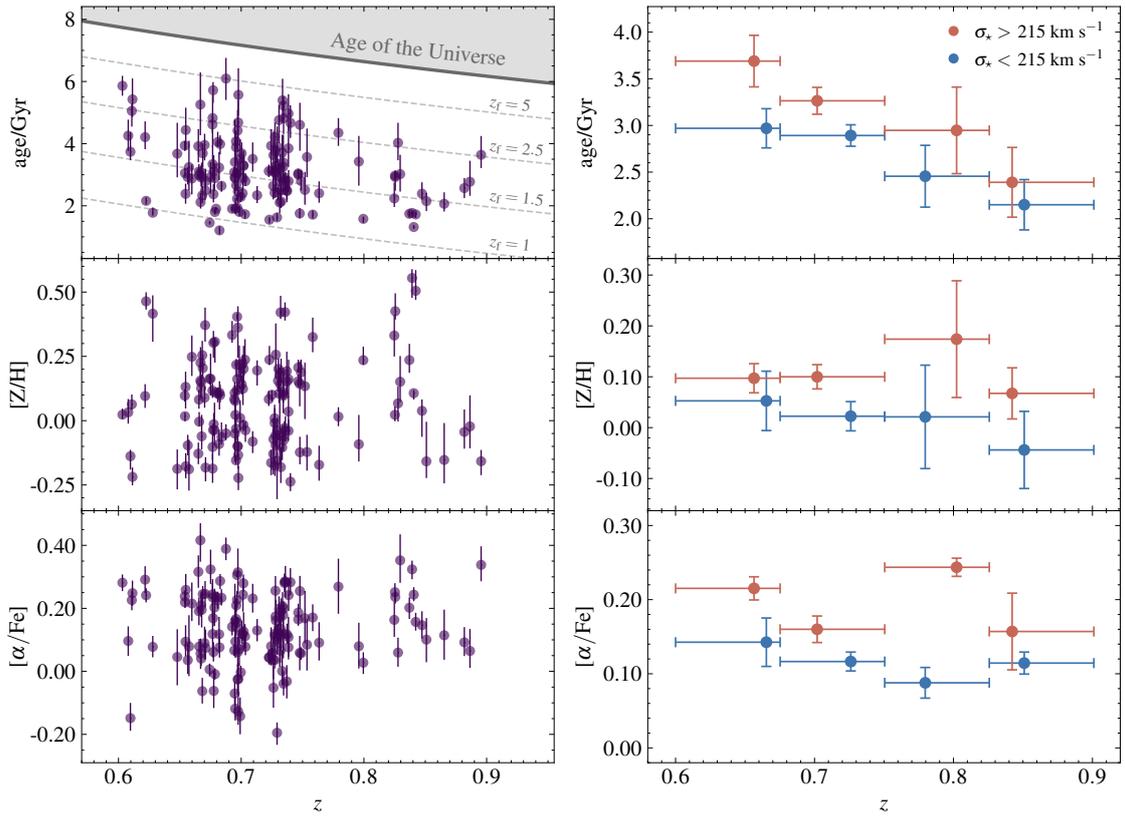


FIGURE 2.12: Left panels: distribution of single-burst stellar ages, metallicities, and α -enhancements as a function of z for 140 *bona fide* passive galaxies selected in LEGA-C (violet points). Error bars represent the 16th and 84th percentiles of the marginalized posterior distributions. In the age – z panel, we shade the parameter space not allowed for a general cosmology (gray solid), as well as the formation redshift assuming a pure passive evolution (dotted lines). Right panels: median binned relations obtained by dividing the sample into two σ_* regimes, with $\sigma_* = 215 \text{ km s}^{-1}$ as a threshold. Each bin contains 5–40 objects. Error bars on the y -axes indicate median value errors, while those on x -axes are the bin widths.

Median binned relations

To provide a comprehensive picture of the results discussed in Sections 2.3.4 and 2.3.5, we bin age, $[Z/H]$, and $[\alpha/Fe]$ in σ_* and in z . Binning in σ_* instead of M_* has the benefit of avoiding model-dependent effects introduced by SED-fit modeling, relying instead on an observational quantity. The galaxies are firstly binned into two σ_* using $\langle\sigma_*\rangle = 215 \text{ km s}^{-1}$ (approximately equivalent to $\langle\log M_*/M_\odot\rangle \approx 11$) as a threshold, then into four equally spaced redshift bins, with $\Delta z = 0.075$ from $z = 0.6$ to $z = 0.9$. Bins have from ~ 5 (40) objects, at higher (lower) z . To each bin, we assign a mean z value and a median (age, $[Z/H]$, $[\alpha/Fe]$) with associated uncertainty. Results are shown in the right panels of Figure 2.12.

Clearly, the median properties of the analyzed passive galaxies follow a downsizing pattern. At each cosmic epoch, stellar populations hosted in galaxies with higher mass are older, more metal-rich and α -enhanced. This suggests that their formation occurred at earlier times, with a difference of $\Delta\text{age} \approx 0.5 \text{ Gyr}$, and on shorter time scales with respect to less massive ones. We note that these trends were already qualitatively confirmed from the analysis of the main absorption indices (Section 2.2.3).

Last but not least, it is remarkable that for each σ_* regime we find a clear, almost parallel, age–redshift relation. The study of their differential evolution allows us to perform cosmological studies using the CC approach.

Overall, the sample of selected passive galaxies at $z \sim 0.7$ shows trends in age, $[Z/H]$, and $[\alpha/Fe]$ versus σ_* in agreement with those expected from a passively evolving population. The age offset, as well as the lack of a significant offset between the typical values of $[Z/H]$ and $[\alpha/Fe]$ is evidence that these systems should have formed their stars on short timescales, depleting the great majority of their gas reservoirs, and experienced passive evolution since then. These results confirm the passive evolution within the $\sim 1.6 \text{ Gyr}$ interval of cosmic time explored in this work shedding light on the granularity of the physical properties and SFHs. The uniformly small scatter in $[Z/H]$ and $[\alpha/Fe]$, $\lesssim 0.20 \text{ dex}$ at fixed σ_* or z , confirms the large homogeneity of the sample and puts strong constraints on the duration of the chemical assembly of these systems.

2.3.6 Analysis with different sets of absorption features

In this final section, we illustrate the results obtained with different index sets. We explore the huge number of viable combinations (~ 1 million) following three approaches:

- *Maximize the information to be fitted.* Using a higher number of indices should provide more stable results. However, indices should be calibrated, and able to disentangle degeneracies giving equal weight to each model parameter.
- *Use already-proposed index combinations.* In particular, we include redder Mg and Fe indices as done in previous works.
- *Use a small, essential set to break the existing degeneracy between parameters.* Spectral indices are sensitive to variations in age, $[Z/H]$, $[\alpha/Fe]$, but the relative sensitivity to these parameters is not identical. One can choose a small combination of four to six indices based upon their different sensitivity (e.g., for element abundances see Tripicco & Bell, 1995; Korn et al., 2005; Lee et al., 2009).

Table 2.5. Results of the additional tests using different sets of spectral features.

ID	H δ_A	H δ_F	CN ₁	CN ₂	Ca4227	G4300	H γ_A	H γ_F	Fe4383	Ca4455	Fe4531	C ₂ 4668	H β	Fe5015	Mg ₁	Mg ₂	Mg b	Fe5270	Fe5385	Fe5406	N	Δage (σ)	$\Delta[\text{Z}/\text{H}]$ (σ)	$\Delta[\alpha/\text{Fe}]$ (σ)
baseline	■		■	■	■	■	■	■	■	■	■	■									140 (140)	–	–	–
1	■				■	■	■	■	■	■	■	■									105 (95)	0.04 (0.1)	-0.04 (0.4)	-0.05 (0.6)
2	■				■	■	■	■	■	■	■	■									131 (115)	-0.42 (0.6)	0.03 (0.3)	-0.06 (0.8)
3	■		■		■	■	■	■	■	■	■	■					■				39 (39)	0.05 (0.1)	0.00 (0.0)	0.08 (0.9)
4	■				■	■	■	■	■	■	■	■		■	■	■	■	■	■	■	11 (8)	0.25 (0.4)	-0.01 (0.1)	0.09 (1.2)
5				■	■	■	■	■	■	■	■										133 (98)	-0.17 (0.3)	0.01 (0.1)	-0.13 (1.6)
6				■	■	■			■	■		■									119 (102)	-1.11 (1.5)	0.11 (1.0)	-0.06 (0.7)

Note. — For each set we report the included indices, number of constrained galaxies (in parenthesis, the number in common with the *baseline* analysis), simple and 1σ differences in the derived parameters with respect to the *baseline* combination.

In the baseline analysis, we use an optimal combination that maximizes the spectral coverage, the number of constrained galaxies, and the precision of the constraints: H δ_A , CN₁, CN₂, Ca4227, G4300, H γ_A , H γ_F , Fe4383, Fe4531, and C₂4668 (hereafter *baseline*). In Table 2.5, we report the other main sets of indices analyzed and the differences in the derived parameters with respect to the *baseline* combination. We also analyzed many other sets (~ 50) that do not add significant information to this study.

We find no significant systematic differences in the derived parameters when small changes to the baseline combination are applied, i.e. by adding or removing one to two indices. In particular, we focus here on the removal of CN indices, as the nitrogen abundance is not a free parameter in TMJ11 models (Combo 1), and on the removal of those indices that sample twice the same spectral region (Combo 2). In the first case, we obtain constraints for fewer galaxies with respect to the *baseline* combination, but with an overall excellent agreement. In the second case, we constrain about the same number of galaxies, obtaining lower ages but still in agreement with the *baseline* set. This is likely due to the removal of H γ_F that reduces the weight of age-sensitive features, therefore producing less reliable age estimates.

Index combinations discussed above lack redder indices, such as the Mg indices, traditionally used as α -abundance indicators. Therefore, we repeat the analysis including Mg b (Combo 3). Among the 59 with such relatively large spectral coverage, we obtain constraints for 39 galaxies. While ages and [Z/H] are in overall good agreement, the inclusion of another α -sensitive index suggests that [α /Fe] values derived with Combo 1 may be underestimated by ~ 0.1 dex. However, given the small statistical significance, we do not correct for this offset. A similar discussion can be done by extending the analysis at all the redder indices (Combo 4, which is also the same combination used in Onodera et al. 2015).

Finally, we note here that although minimal sets of $N = 4$ indices in the 4000–4600 Å region would enable the analysis of a larger sample of galaxies (~ 300), we do not find a relevant set to effectively constrain more than one-third of it. The situation is improved when $N = 5 - 7$ (e.g., Combo 5), but results show an overall stronger age–metallicity degeneracy

with respect to $N \geq 8$. It is also interesting to note that even by removing the age-sensitive Balmer indices (Combo 6), we still obtain constraints for a significant number of galaxies.

The trends presented in Sections 2.3.4 and 2.3.5 are qualitatively confirmed for Combo 1, 2, 5, and 6, while Combo 3 and 4 do not provide sufficient statistics, for a more thorough we refer to (Borghi et al., 2022a). These results are used in Section... and contribute to the systematics in the $H(z)$ derivation.

In summary, we find that a blind choice of index combinations can lead to less robust results. This happens when a combination is unbalanced toward one or more parameters of the fit, but also if indices are measured on spectra where the sky subtraction was imperfect. After an extensive analysis, we demonstrate that within the limited statistics and wavelength coverage of current data, results do not show significant systematic differences.

2.3.7 Analysis with different models

To verify the dependence of our results on the assumed stellar population synthesis (SPS) model, we repeat the entire analysis by adopting the α -MILES models by Vazdekis et al. 2015 (hereafter V15). Similarly to TMJ11, they are generated with variable age, $[Z/H]$, $[\alpha/Fe]$ parameters, and use an updated version of the same empirical stellar library (MILES, Falc3n-Barroso et al., 2011), but are based on corrections from theoretical stellar spectra and assume different stellar isochrones (BaSTI, Pietrinferni et al., 2006). We note that with respect to TMJ11, one of the drawbacks of V15 models is that they allow a poorer exploration of the parameter space, having, in particular, a smaller sampling of $[\alpha/Fe] = 0, 0.4$. This introduces some limitations in their use, as will be discussed below, and is one of the reasons why we adopted TMJ11 models as our reference. However, they give us the possibility to test the assumption of a more extended star exponentially declining formation history

$$\text{SFR}(t) \propto e^{-(\text{age}-t)/\tau}, \quad (2.19)$$

where τ is the characteristic star formation timescale.

We generate synthetic spectra with variable age, $[Z/H]$, $[\alpha/Fe]$, and τ covering the wavelength range $3550 < \lambda/\text{\AA} < 5500$ at a resolution of 2.5 \AA FWHM and measure the main spectral indices with `pyLick`. The original grid, spanning the following parameter space: $0.1 < \text{age}/\text{Gyr} < 14$ (14 points), $0.01 < \tau/\text{Gyr} < 3$ (7 points), $-2.25 < [Z/H] < 0.40$ (7 points), and only two $[\alpha/Fe]$ points (0 and 0.4), is interpolated to a resolution of 0.2 Gyr in age and τ , and 0.02 dex in $[Z/H]$ and $[\alpha/Fe]$. We use the baseline set of spectral features for a fair comparison. The indices measured on modeled spectra, which are a function of $\theta = (\text{age}, \tau, [Z/H], [\alpha/Fe])$, are compared to the ones measured on the selected passive galaxies Section 2.2.1. In the MCMC analysis, we explore the entire parameter space allowed from the models and - we emphasize here - no cosmological priors are used to derive galaxy ages.

At the end of this process, we obtain two datasets describing the stellar population properties of the 140 cosmic chronometers using the V15 models:

1. V15 – SSP: single-burst SFH ($\tau \equiv 0$);
2. V15 – τ -decl.: exponentially declining SFH (Eq. (2.19)).

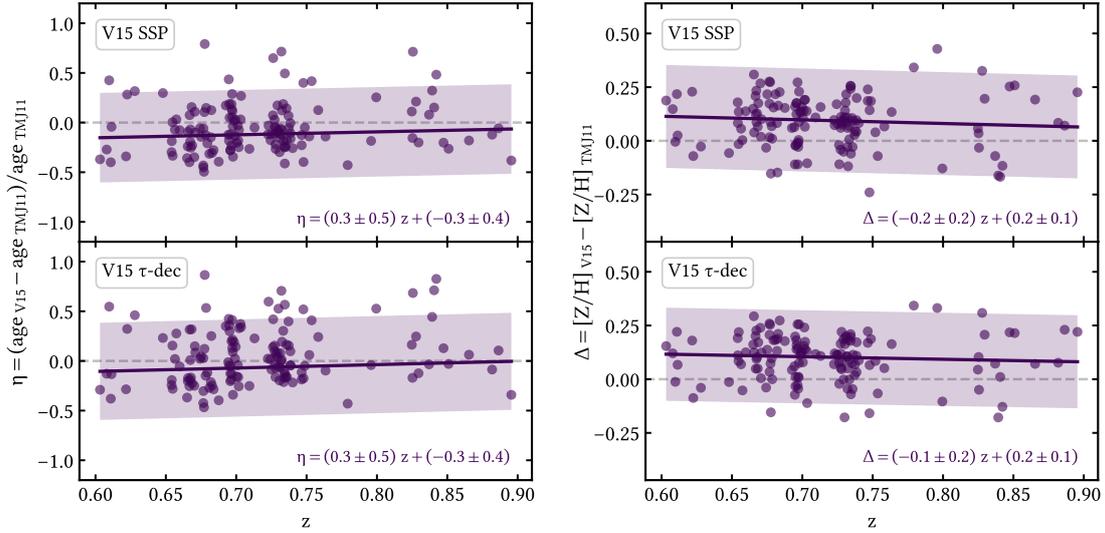


FIGURE 2.13: Differences as a function of redshift between stellar ages and metallicities $[Z/H]$ of 140 LEGA-C passive galaxies obtained with [Vazdekis et al. \(2015\)](#) single-burst (SSP, upper panels) and exponentially declining (τ -decl., lower panels) star formation histories versus [Thomas et al. \(2011\)](#) SSP models. Violet lines and shaded regions are robust linear fits and associated 2σ scatter regions, respectively.

In the following section, these will be compared the [TMJ11](#) results. For the purpose of this study, we are interested in detecting any possible variation in the trends with redshift. In particular, we will discuss percentage differences in ages and absolute differences in $[Z/H]$ (already expressed in log units) as a function of z :

$$\eta = \text{age}_{\text{V15}}/\text{age}_{\text{TMJ11}} - 1, \quad (2.20)$$

$$\Delta = [Z/H]_{\text{V15}} - [Z/H]_{\text{TMJ11}}. \quad (2.21)$$

We note here that the interpolation between the two available $[\alpha/\text{Fe}]$ points is not optimal to capture the granularity of this parameter. Indeed, for almost all the galaxies ($> 90\%$), we obtain typical $[\alpha/\text{Fe}] \sim 0$, which is also the grid point nearest to the values obtained with [TMJ11](#) models (~ 0.13 dex). We have also checked that the baseline index combination is not optimal to capture $[\alpha/\text{Fe}]$ variations with the current models. While further analysis with models with denser $[\alpha/\text{Fe}]$ grid points is needed to better understand these differences and study any possible trend of $[\alpha/\text{Fe}]$ with redshift, the analysis of age and $[Z/H]$ —which can be strongly degenerate—is sufficient to explore systematic effects on the final $H(z)$ value.

Different models with the same (single burst) SFH

When galaxies are modeled as SSPs, we obtain typical values of $\langle \text{age} \rangle = 2.65 \pm 0.46$ Gyr and $\langle [Z/H] \rangle = 0.16 \pm 0.27$ dex, differing by -0.36 Gyr and $+0.08$ dex, respectively, from the [TMJ11](#) results. Even if these differences are consistent within 1σ , it is interesting to note that they follow the trend expected from the age-metallicity degeneracy, i.e. younger ages and higher metallicities. However, one of the main advantages of the cosmic chronometer method

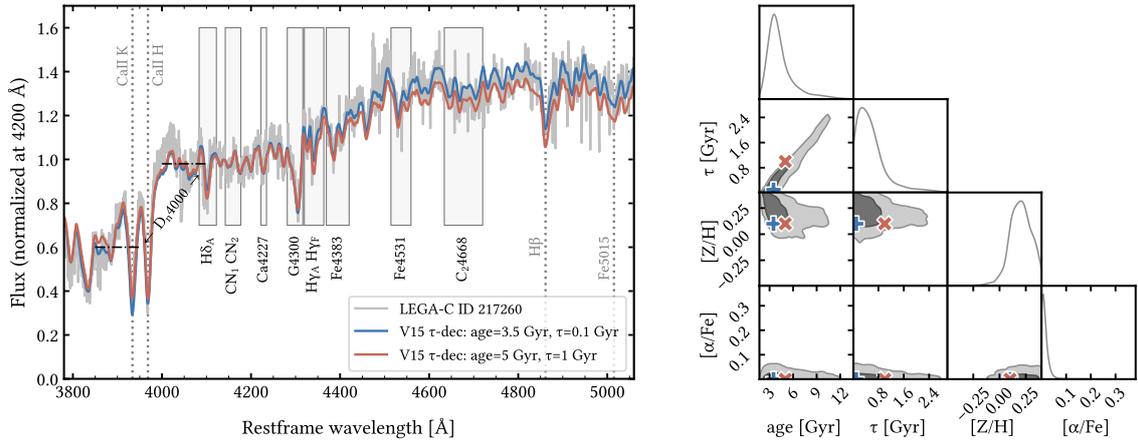


FIGURE 2.14: Example of the age–star formation timescale degeneracy for a LEGA-C galaxy (ID 217260, $S/N \simeq 22 \text{ pix}^{-1}$) compared to the [Vazdekis et al. \(2015\)](#) τ declining models. Left: observed (gray) and synthetic (blue and red) spectra normalized at 4200 Å. The blue spectrum corresponds to the model closer to the best-fit parameters, while the red one is taken at the edge of the 1σ confidence region. Right: corner plot for stellar age, star formation timescale τ , stellar metallicity $[Z/H]$, and $[\alpha/Fe]$. The contours enclose 1σ (darker shade) and 2σ (lighter shade) confidence regions. Blue and red symbols indicate the values at which the synthetic spectra are generated.

is that it is insensitive to any systematic offset of the age redshift relation (Eq. (2.1)). We thus explore systematics by studying the evolution of differences over redshift (Figure 2.13, upper panels).

Remarkably, we find no significant deviations as a function of z in the redshift range of interest, with typical differences ranging between $-0.15 < \eta < -0.07$ (with 0.22 rms scatter) and $0.07 < \Delta < 0.12$ (with 0.12 rms scatter). This means that the mean trends of this population of galaxies do not significantly deviate from those observed with [TMJ11](#) models. As done in the baseline analysis, we compute age–redshift relations for the lower and the higher σ_* subsamples in 4 redshift bins. The final $H(z)$ measurements using median and mean as estimators differ by only 0.6σ from the baseline. When also testing different binning schemes (including the systematic effects already estimated in Section 2.1.4), we obtain measurements consistent within 0.7σ .

Different models with a more extended SFH

In the main analysis, we adopt a single-burst star formation history. Even if our selection criteria were chosen to obtain a sample of galaxies with very short SFH, the single single-burst approximation is not realistic. However, it is important to stress that any constant star formation time scale for the entire population of these galaxies leads to a vertical shift of the age–redshift relation, therefore, the final $H(z)$ measurement would not be affected. Again, we want to test whether there is any trend of τ with z , which could in principle introduce a bias in the $H(z)$ measurement. At the same time, this analysis gives us the possibility to test how well our assumption of single-burst SFH fits with the observed sample.

Despite the wide range of τ adopted (0.01–3 Gyr), we obtain typical values of $\langle\tau\rangle = 0.24 \pm 0.21$ Gyr with only 23% of the galaxies having $\tau > 1$ Gyr. This is an important confirmation that the stellar components of the bulk of these systems formed in very short episodes. We also find no significant dependence on z , suggesting minor systematic effects on the final $H(z)$ value. From a more detailed analysis of their posterior distributions, we find that especially systems with $\tau \gtrsim 1$ Gyr suffer from strong degeneracy between age and τ . This degeneracy is well known in the literature (e.g., [Gavazzi et al., 2002](#)) and together with the age-metallicity degeneracy is one of the major obstacles in the accurate reconstruction of galaxy star formation histories. Quantitatively, from the analysis of the posterior distributions of our dataset, we find:

$$\Delta\tau/\Delta\text{age} \simeq 0.3 \quad (2.22)$$

i.e., the same set of indices can be reproduced if a galaxy is 1 Gyr older and its star formation time scale extends by 0.3 Gyr. This aspect should be carefully considered when ages from different samples with different SFH assumptions are compared. The age–star formation timescale degeneracy is generally (partially) broken by placing a cosmological prior in the form of an upper limit on galaxies’ ages depending on the redshift of observation. However, we shall not use cosmological assumptions in our analysis, as it would introduce a circularity: the retrieved $H(z)$ constraints would be driven by the priors assumed. A possible solution could come from the detailed modeling of Ca II H and K features, which have proven to be good diagnostics of underlying young stellar populations (see, e.g., [Moresco et al., 2018](#); [Borghi et al., 2022a](#)). This can also be seen in [Figure 2.14](#) where we show a typical galaxy for which the best fit (blue curve, with $\tau = 1$ Gyr and age = 5 Gyr) provides similar results with respect to the solution with $\tau = 0.01$ Gyr and age = 3.5 Gyr (red curve). The spectral indices are insensitive to any difference in the flux normalization, therefore, the normalization of models adopted in the figure (currently chosen in the range $4180 < \lambda/\text{\AA} < 4220$) is only used for a visual comparison of the models. On the contrary, the difference in the Ca II H and K lines (not used in this analysis), could be a viable option, preferring the solution with lower τ . This diagnostic will be further explored in future work using the full spectral fitting technique, which allows more flexibility and extensibility to study galaxy SFHs. In this work, we repeat the analysis by fixing an upper prior of $\tau < 0.5$ Gyr corresponding to the upper 1σ value of the entire population. We verified that this prior does not significantly modify the shape of the age–redshift relation.

We obtain typical values of $\langle\text{age}\rangle = 2.88 \pm 0.61$ Gyr, $\langle\tau\rangle = 0.17 \pm 0.09$ Gyr and $\langle[Z/H]\rangle = 0.21 \pm 0.24$ dex, differing by +0.23 Gyr, +0.17 Gyr, and +0.05 dex, respectively, from the SSP results. In [Figure 2.13](#) (lower panels) we show the evolution of age and $[Z/H]$ differences over redshift.

Again, it is remarkable that there are no significant deviations as a function of z in the redshift range of interest, with typical differences ranging between $-0.11 < \eta < -0.01$ (with 0.24 rms scatter) and $0.08 < \Delta < 0.12$ (with 0.11 rms scatter). As in the main analysis, we have computed median age–redshift relations for the lower and the higher σ_* subsamples in four redshift bins. The final $H(z)$ measurements using median and mean as estimators differ by 0.8σ from the ones obtained with [V15](#) assuming a single-burst SFH and by 0.4σ from the baseline ([TMJ11](#), single-burst SFH). Also in this case, we test different binning schemes, obtaining measurements consistent within 0.7σ .

2.4 A new measurement of $H(z)$ at $z \sim 0.7$

This Section presents the cosmological constraints from the 140 cosmic chronometers selected in LEGA-C. Their median age–redshift relation is used to derive a new direct measurement of $H(z)$ with the CC approach (§ 2.4.1), with a detailed studied of the involved systematics (§ 2.4.2). In parallel, we also propose a simple framework to derive H_0 from the age–redshift relations of passive galaxies in the context of galaxy downsizing (§ 2.4.3).

2.4.1 The direct measurement

In the CC approach, the Hubble parameter $H(z)$ can be derived directly and without any cosmological assumptions. In this work, we use the median binned age– z relation derived in Section 2.3.5 for a population of 140 passive galaxies in LEGA-C. We recall that the galaxies are divided into two stellar velocity dispersion subsamples using their median $\langle\sigma_\star\rangle = 215 \text{ km s}^{-1}$. For each σ_\star regime, we evaluate the median age in four narrow redshift bins (see fig. 2.15, upper panel). The constant bin width $\Delta z \simeq 0.075$ corresponds to ~ 0.4 Gyr difference in cosmic time, which is also the average age uncertainty. The two resulting age– z relations for the higher and lower σ_\star regime are approximately parallel and with an offset of $\Delta\text{age} \simeq 0.5$ Gyr. This is consistent with the mass-downsizing scenario, for which more massive galaxies formed earlier and faster. We estimate $\tilde{H}(z)$ (see Eq. (2.1)) as

$$H(z_{\text{eff}}) = -\frac{1}{1 + z_{\text{eff}}} \frac{z_{i+2} - z_i}{t_{i+2} - t_i}, \quad (2.23)$$

where the quantities $\Delta z = z_{i+2} - z_i$ and $\Delta\text{age}_{\text{CC}} = t_{i+2} - t_i$ are measured between each non-consecutive bins and $H(z)$ is defined at an effective redshift $z_{\text{eff}} = (z_i + z_{i+2})/2$. The choice to use alternate points is to ensure that the evolution in age over the assumed redshift intervals Δz (~ 0.4 Gyr of cosmic time) is larger than the statistical scatter, but at the same time sufficiently small to minimize possible systematic effects (see Moresco et al., 2012b). With this bin choice, we obtain four independent $H(z)$ estimates, two for each mass bin Table 2.6.

We find that the results for lower and higher σ_\star regimes are in very good agreement, with their mean values being within 0.1σ , confirming the idea that these two subpopulations are tracing the same underlying cosmology. Since all four measurements are independent from each other, we combine them using a variance-weighted average, obtaining $H(z = 0.75) = 98.8 \pm 24.8$ (*stat*) $\text{km s}^{-1} \text{Mpc}^{-1}$ at 68% C.L. (Figure 2.15, lower panel). Our measurement is perfectly consistent with the values estimated with different CC datasets and methods. In particular, the most comparable measurements at this redshift are both from Moresco et al. (2012b) using the D4000 method. Our value lies in between their $H(z = 0.68) = 91.6 \pm 8.0 \text{ km s}^{-1} \text{Mpc}^{-1}$ and $H(z = 0.78) = 104.5 \pm 12.2 \text{ km s}^{-1} \text{Mpc}^{-1}$, differing only by $+0.3\sigma$ and -0.2σ , respectively.

2.4.2 Assessing systematic uncertainties

Dependence on the Lick indices set. In Section 2.1.4, we studied how different combinations of absorption features impact the constraints on age, $[Z/H]$, and $[\alpha/\text{Fe}]$. Here, we use

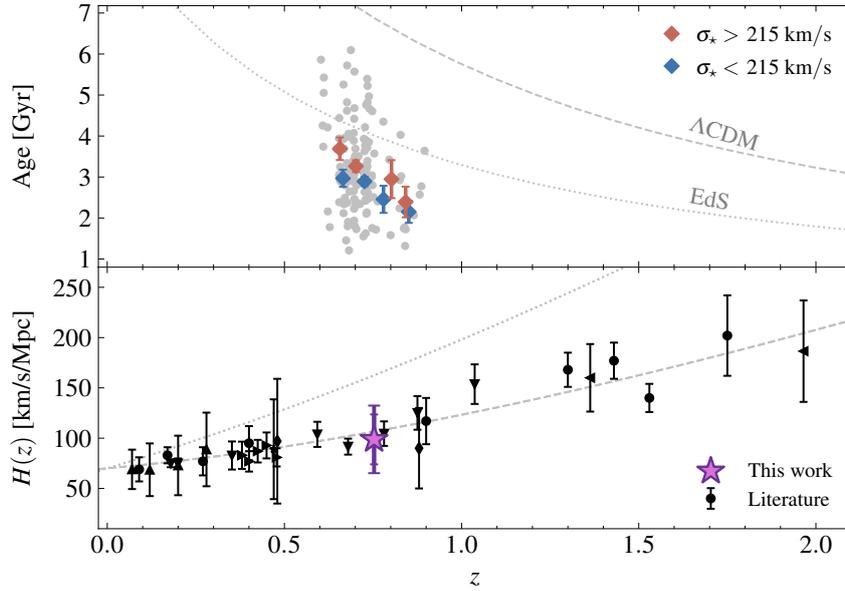


FIGURE 2.15: Upper panel: median binned age–redshift relation for 140 passive galaxies analyzed in [Borghi et al. \(2022a\)](#) (gray points) divided into higher (red) and lower (blue) σ_* regimes. Lower panel: $H(z)$ measurement (violet star) with statistical (inner error bar) and total (outer error bar) uncertainties. Black points are literature data from: [Simon et al. 2005](#) (\bullet), [Stern et al. 2010](#) (\blacklozenge), [Moresco et al. 2012b](#) (\blacktriangledown), [Zhang et al. 2014](#) (\blacktriangle), [Moresco 2015](#) (\blacktriangleleft), [Moresco et al. 2016](#) (\blacktriangleright), and [Ratsimbazafy et al. 2017](#) (\blacksquare). Gray lines are theoretical relations for a standard Λ CDM (dashed) and Einstein–de Sitter (dotted) cosmology.

Table 2.6. Hubble Parameter Measurements.

Sample	Bins	N. of gal.	z_{eff}	Δz	$\Delta \text{age}_{\text{CC}}$ [Gyr]	$H(z_{\text{eff}})$ [km/s/Mpc]	σ_{stat} [km/s/Mpc]
Lower σ_*	1 & 3	20	0.723	0.114	-0.514	126.3	96.4
	2 & 4	50	0.789	0.125	-0.742	92.0	36.3
Higher σ_*	1 & 3	21	0.729	0.145	-0.741	111.0	80.7
	2 & 4	49	0.772	0.149	-0.874	88.6	40.6
Joint	all	140	0.753 ^a	–	–	98.8 ^a	24.8 ^a

^aJoint results are the error-weighted average of the four Hubble parameter values and are defined at the average effective redshift.

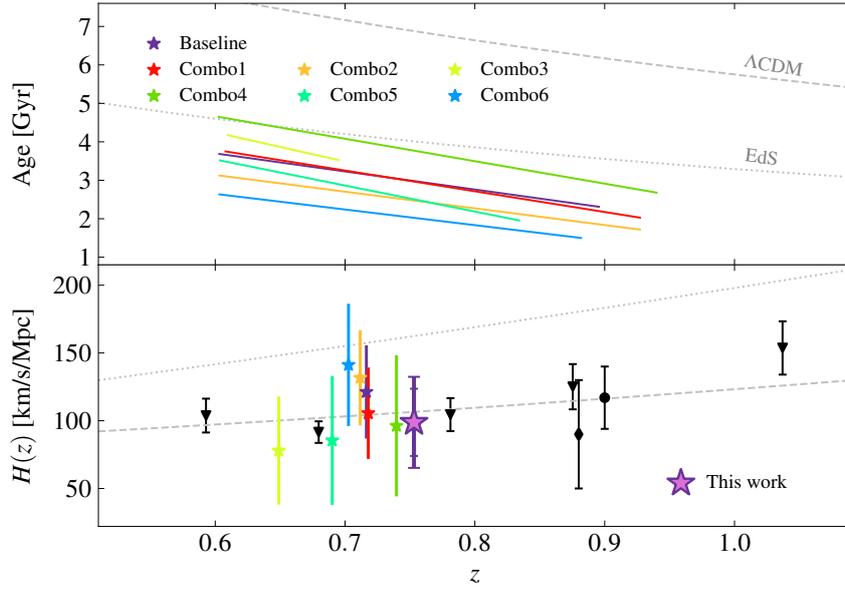


FIGURE 2.16: Same as Figure 2.15, but showing the results obtained with linear fits to the unbinned age- z relations for different combinations of indices used in the analysis (as presented in Borghi et al. 2022a). Gray lines are theoretical relations for a standard Λ CDM (dashed) and Einstein-de Sitter (dotted) cosmology and are shown for visual inspection purposes only. We note that the EdS model predicts a much flatter slope compared to the ones of the data, which are instead more compatible with a Λ CDM scenario, as also confirmed by the $H(z)$ measurements.

these results to study the effect of using different indices on the age-redshift slope and the final $H(z)$ value. The baseline combination is devised to maximize the number of indices to be measured given the redshift and wavelength coverage of the various galaxies; moreover, any other index set provides age constraints for fewer objects (down to a dozen for the worst case) and binning them is not always an option. For this reason, to assess this systematic effect, we estimate $\Delta z / \Delta \text{age}_{\text{CC}}$ and its associated uncertainty from the inverse slope of the age-redshift relation obtained with a simple linear regression. Results are shown in Figure 2.16 (see tab. A.1 for their definitions).

Different combinations of indices can provide systematically different absolute age estimates, ranging within ± 1 Gyr. However, we find that the $H(z)$ estimates are consistent with each other and with the more statistically rigorous value obtained with the median binning within 0.4σ , on average. These results clearly highlight the advantages of CC being a differential approach; in other words, the absolute age calibration that might be obtained in different analyses does not significantly affect the final $H(z)$ value, but only the normalization of the age- z relation.

Dependence on the binning. We verify that our result is robust against different redshift binning schemes and adopted estimators. In particular, by using from two up to six redshift intervals, or/and the mean instead of the median age, $H(z)$ results are on average within 0.5σ with respect to the baseline. We do not use weighted averages because in Borghi et al.

(2022a) we found that the stellar population analysis intrinsically yields higher uncertainties for older galaxies and this would bias the final age–redshift slope. Finally, we also repeat the analysis using equally populated redshift bins (about 20 objects per bin). This method improves the statistics of single bins at the expense of smaller leverage in redshift. Even in this case, we obtain values in good agreement, with an average difference of 0.5σ .

Dependence on the SPS model. The choice of the stellar population model plays a major role in the overall systematics of the CC approach. Quantitatively, Moresco et al. (2020) measured an average contribution of $\sim 7\%$ on the final uncertainty of $H(z)$ using the D4000 method. To assess this effect in our work, we repeat the entire analysis by adopting the α -MILES models by Vazdekis et al. (2015). The detailed analysis is presented in Section 2.3.7. We find that the $H(z)$ measurements obtained with the assumption of a different SPS model are consistent with the baseline within 0.6σ , on average.

Dependence on a more extended SFH. In all the previous analyses, we assume single-burst star formation histories (SFH). This is justified by the accurate selection of passive systems which maximizes the presence of galaxies with coeval SFH concentrated at early cosmic epochs. Here, we assess the effect of assuming a more extended SFH $\propto \exp(t/\tau)$, i.e. exponentially declining with a characteristic timescale τ . The detailed analysis is presented in Section 2.3.7. As expected from the sample selection, we confirm very short SFHs with a typical $\tau \lesssim 0.4$ Gyr. By analyzing the slopes of the age–redshift relations obtained with these more extended SFHs, we find that the final $H(z)$ measurements differ by only 0.4σ , on average, with respect to the baseline.

2.4.3 Fit to the age–redshift relations

In this Section, we propose a simple scheme to derive cosmological parameters from the age–redshift relations of different subsamples of CCs binned by their stellar velocity dispersion σ_* in the context of a downsizing evolution.

The model

The age of the Universe as a function of redshift, $t_U(z)$, can be predicted from cosmological models. With the minimal assumption of an FLRW metric:

$$t_U(z) = \frac{1}{H_0} \int_z^\infty \frac{dz'}{(1+z')E(z')}, \quad (2.24)$$

where $E(z)$ is the normalized Hubble parameter. Here we assume that the late-time expansion history is described by a flat $w_0 w_a$ CDM universe, where the dark-energy equation of state varies with cosmic time under the CPL parameterization, $w(z) = w_0 + w_a(z/(1+z))$ (Chevallier & Polarski, 2001; Linder, 2003), therefore:

$$E(z) = \sqrt{\Omega_{m,0}(1+z)^3 + (1 - \Omega_{m,0})f(z)}, \quad (2.25)$$

$$f(z) = (1+z)^{3(1+w_0+w_a)} e^{-3w_a \frac{z}{1+z}} \quad (2.26)$$

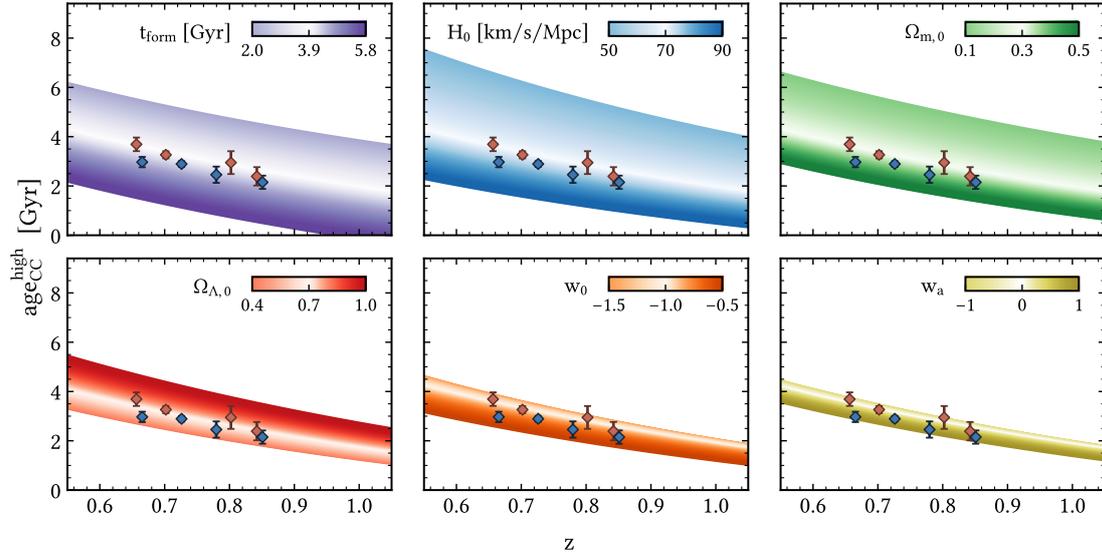


FIGURE 2.17: Theoretical age–redshift relations for the high- σ_* subsample of cosmic chronometers. Each panel shows the effect of varying each parameter to the labeled values. Red diamonds are median binned data for the high- σ_* subsample of cosmic chronometers from [Borghi et al. \(2022a\)](#). For illustrative purposes, we also show the data points of the lower σ_* subsample (blue diamonds), which are about 0.5 Gyr younger, on average. When not varied, the parameters are set to the following fiducial values: $t_f = 3.9$ Gyr, $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{m,0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$, $w_0 = -1$, and $w_a = 0$.

where radiation is not considered since its contribution is negligible in the late Universe. The function $f(z)$ describes the dark-energy contribution and for a flat Λ CDM model ($w_0 = -1$, $w_a = 0$) it becomes $f(z) = 1$.

Given the inverse relationship between $t_U(z)$ and H_0 , lower limits on $t_U(z)$ from the ages of the oldest objects would determine upper limits on the local H_0 value. Recently, this method has been applied by [Vagnozzi et al. \(2022\)](#) to obtain constraints on H_0 from galaxies and quasars observed up to $z \sim 8$.

Galaxy formation occurs after a time t_{form} from the big bang, which could in principle vary with redshift depending on the considered sample. However, CCs are a population of objects selected to be very coeval in formation time. Therefore, their age– z relation can be written as:

$$\text{age}_{\text{CC}}(z) = t_U(z) - t_{\text{form}}. \quad (2.27)$$

According to the downsizing scenario, galaxy mass is a main driver of galaxy formation and evolution, with more massive galaxies forming their stars at earlier cosmic epochs with respect to less massive ones. For this reason, multiple parallel age–redshift relations for different σ_* populations are expected (and actually visible in the current dataset). Therefore, we use both the lower and higher σ_* subsamples as homogeneous tracers of the age of the Universe by assuming a constant offset in formation time Δt_{form} computed as the mean age difference. We take as a reference the higher σ_* age– z relation $\text{age}_{\text{CC}}^{\text{high}}(z)$, so that $\text{age}_{\text{CC}}^{\text{low}}(z) = \text{age}_{\text{CC}}^{\text{high}}(z) - \Delta t_{\text{form}}$.

Preliminary tests

In Figure 2.17, we illustrate the dependence of $\text{age}_{\text{CC}}^{\text{high}}(z)$ on the typical formation time and the cosmological parameters by varying one parameter at a time.

As expected from Equation 2.24, similar age– z trends are found by increasing t_{form} (hence t_U) and decreasing H_0 (and vice versa). A less evident anti-correlation is observed between $H_0 - \Omega_{m,0}$ and $\Omega_{m,0} - \Omega_{\Lambda,0}$. The latter is orthogonal to the degeneracy that is present in CMB-only data, so that the combination of these two independent probes can eventually provide more stringent constraints on cosmological parameters (see, e.g., Moresco et al., 2016; Vagnozzi et al., 2021). Finally, it is clear that with the current data it is not possible to set strong constraints on the dark energy equation of state parameters w_0 and w_a because of their smaller effect on $\text{age}_{\text{CC}}(z)$.

In our analysis, we therefore assume a flat Λ CDM universe ($\Omega_{\Lambda,0} = 1 - \Omega_{m,0}$, $w_0 = -1$, $w_a = 0$), so that the final model is described by three parameters, $\theta = (t_{\text{form}}, H_0, \Omega_{m,0})$. We constrain these parameters by using the affine-invariant Markov Chain Monte Carlo sampler `emcee` (Foreman-Mackey et al., 2019), assuming a Gaussian likelihood function $\propto e^{-\chi^2/2}$. Priors are chosen to be large and uninformative, with $H_0 \sim \mathcal{U}(0, 150)$ km s⁻¹ Mpc⁻¹, $\Omega_{m,0} \sim \mathcal{U}(0.01, 0.99)$, and $t_{\text{form}} \sim \mathcal{U}(1, 10)$ Gyr. The final values and associated uncertainties are defined as the cumulative mean and 1σ values of the marginalized posterior distributions.

Results

The results are shown in Figure 2.18. We obtain $H_0 = 72_{-19}^{+27}$ km s⁻¹ Mpc⁻¹, $\Omega_{m,0} = 0.38_{-0.23}^{+0.36}$, and $t_{\text{form}} = 3.2_{-1.3}^{+1.8}$ Gyr. Given the large uncertainties and the small redshift range sampled, our current result is in agreement with both early- and late-Universe H_0 determinations; indeed, this method is limited by the intrinsic degeneracies between the parameters shown in Figure 2.17. We note, however, that these constraints can be significantly improved by increasing the redshift leverage and accuracy of the data, as, for example, could be done by analyzing massive and passive galaxies from proposed spectroscopic missions such as the ATLAS probe (Wang et al., 2019). Differently from the standard CC method presented in Section 2.4.1, the analysis of $\text{age}_{\text{CC}}(z)$ relies on absolute ages estimates and therefore requires an accurate calibration of galaxies' ages and SFHs and a homogeneous analysis between different samples.

We also repeat the analysis assuming a Gaussian prior on $\Omega_{m,0} \sim \mathcal{N}(0.316, 0.007)$ based on (Planck Collaboration, 2020) TT,TE,EE+lowE+lensing results. In this case, we obtain $H_0 = 77_{-17}^{+20}$ km s⁻¹ Mpc⁻¹, and $t_{\text{form}} = 3.0_{-1.2}^{+1.7}$ Gyr with a significant degeneracy between the two parameters.

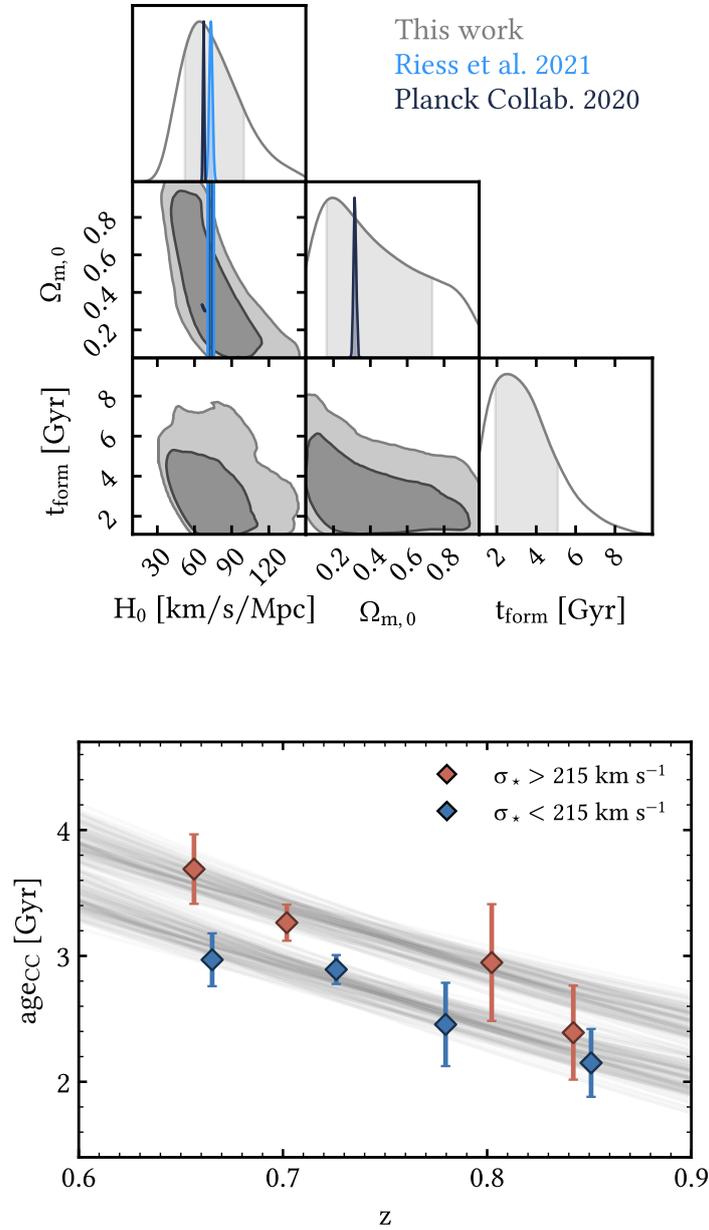


FIGURE 2.18: Constraints from age–redshift relations assuming a flat Λ CDM model. Top: corner plot for H_0 , $\Omega_{m,0}$, t_f , showing 1σ (darker shade) and 2σ (lighter shade) regions. Vertical shaded bands are the cumulative 1σ confidence regions. Our results are compared with those from [Riess et al. \(2021\)](#) and [Planck Collaboration \(2020, TT,TE,EE+lowE+lensing\)](#). Bottom: resulting fits (gray lines) to the observed age–redshift relations (diamonds).

2.5 Further studies with cosmic chronometers

This work was followed up by two other studies with the approach to test for the first time the same data set of passive galaxies using a different age measurement method and extend the analysis to higher- z passive galaxies.

In [Jiao et al. \(2023\)](#), we carry out a full-spectrum fitting analysis on the 350 massive and passive galaxies selected in LEGA-C. In particular, we use [BAGPIPES Carnall et al. \(2018\)](#) with a modification that allows us to remove cosmological priors from the estimation of the age. Extensive tests are performed to assess the dependence of the derived ages on possible contribution of dust, calibration of spectrum noise and flux, and use of photometric data in addition to spectral data. Both spectroscopic and photometric data (21 bands) are fit jointly, which helps break degeneracies between parameters. Convergence diagnostics based on modality, truncation, and H/K line ratios are defined to flag potentially unreliable fits. Contrary to the analysis presented in this thesis, we do not require a uniform spectral range to be analyzed for all galaxies. This, combined with the aforementioned diagnostics, allows one to increase the percentage of galaxies analyzed to 335. The age-redshift relation follows the expected cosmological decline, with a clear downsizing trend, with more massive galaxies being formed earlier ($z_{\text{form}} \sim 2.5$) compared to less massive ones ($z_{\text{form}} \sim 2$). The metallicities are subsolar with no strong evolution and the star-formation timescales are short (typically 0.36 Gyr).

A key result is the comparison to the CC analysis of ([Borghi et al., 2022b](#)) on the same LEGA-C sample (see Fig. 2.19). The differential ages dz/dt agree remarkably well between the two independent methods, despite the different assumed SFHs. This demonstrates the robustness of the CC technique. The final $H(z)$ is

$$H(z = 0.8) = 113.1 \pm 15.2 \text{ (stat.) } {}_{-4.0}^{+24.2} \text{ (syst.) km s}^{-1} \text{ Mpc}^{-1}. \quad (2.28)$$

in very good agreement with the Lick index method results ([Borghi et al., 2022b](#)), with a statistical error smaller by a factor of $\sqrt{2}$ due to the larger ($\times 2$) sample size.

In [Tomasetti et al. \(2023\)](#), we perform CC analysis with data from the VANDELS survey ([McLure et al., 2018](#); [Pentericci et al., 2018](#); [Garilli et al., 2021](#)), a deep VIMOS spectroscopic survey of the CANDELS UDS and CDFS fields, with spectra which span 4800 – 9800 Å at a resolution $R \sim 650$ and integration times up to 80 hours. The CC analysis is performed on a sample of 39 massive and passive galaxies at $1 < z < 1.5$ obtained by applying strict photometric (UVJ) and spectroscopic ([O II] and Ca H/K) selection cuts. With this sample, we obtain a new measurement of

$$H(z = 1.26) = 135 \pm 65 \text{ (stat. + syst.) km s}^{-1} \text{ Mpc}^{-1}, \quad (2.29)$$

including both statistical and systematic uncertainties, with the first dominating the error budget. Despite the large errorbars due to the limited statistics, this is a remarkable result, as it is the furthest CC measurement obtained with full-spectral fitting.

In summary, these analyses further validate the cosmic chronometer method as a powerful probe of the expansion history. The consistency obtained between different spectroscopic dating methods is very encouraging, and prospects are good for reducing uncertainties with future large spectroscopic surveys.

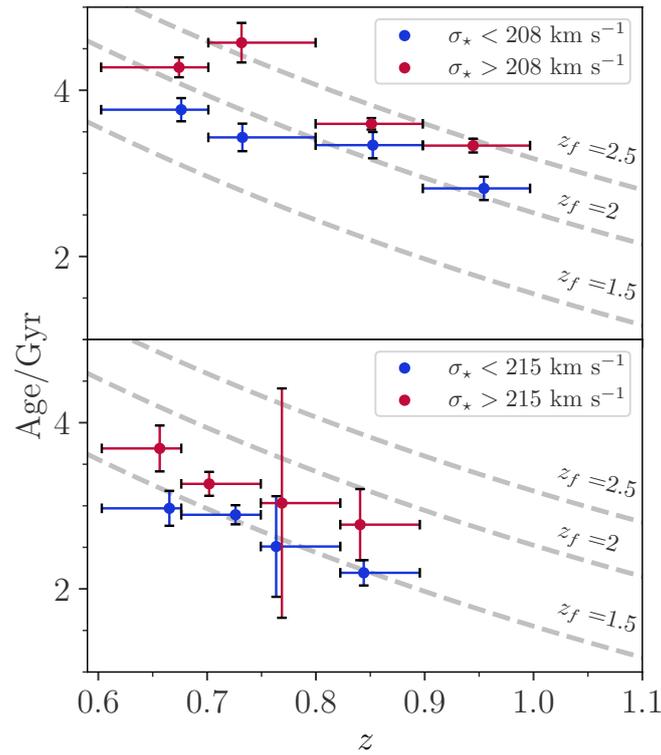


FIGURE 2.19: Median binned age– z relations of LEGA-C passive galaxies obtained with full-spectral fitting (*top panel*, Jiao et al., 2023) and Lick indices (*bottom panel*, Borghi et al., 2022a). The blue and red points represent lower and higher σ_* bins, divided using the median value of each sample as a threshold. For illustrative purposes only, the gray dashed line indicate the formation redshift assuming a flat Λ CDM cosmology from Planck Collaboration (2020).

Reproduced from Jiao et al. (2023).

2.6 Summary

In this Thesis, we extend the cosmic chronometer method and jointly study astrophysical and cosmological parameters, exploring synergies between galaxy evolution and cosmology. In particular, we take advantage of the second Data Release of the LEGA-C spectroscopic survey to place constraints on the stellar population properties of individual massive and passive galaxies at $0.6 < z < 0.9$. Based on a robust spectral analysis of Lick indices, we characterize this population and explore the reliability of using these galaxies as cosmic chronometers. The main results are outlined below.

1. We select a pure sample of 350 passive galaxies at $z \sim 0.7$ combining a photometric NUVrJ selection, a spectroscopic EW[O II] cut, and a careful visual inspection of individual spectra to further remove galaxies with significant emission lines (Figure 2.2). As confirmed by the stacked spectrum (Figure 2.3), no underlying emission line components are present in the sample, confirming its high purity. Selected passive galaxies have a median observed velocity dispersion of $\langle \sigma_* \rangle = 206 \text{ km s}^{-1}$, stellar mass of $\langle \log M_*/M_\odot \rangle = 10.95$, and very low specific star formation rate $\langle \log \text{sSFR/yr} \rangle = -12.1$. Most of them have an early-type morphology, but there is also a non-negligible percentage of systems (about one-third) with an intermediate morphology.
2. We develop and publicly release `PyLick`, a flexible Python tool to measure absorption features, implementing several absorption feature definitions.⁵ This allows us to measure spectral indices over a wide wavelength range in LEGA-C data, extending the current public catalog of Lick indices by [Straatman et al. \(2018\)](#) and allowing a more detailed exploration of the dependence of our results on different index combinations.
3. We introduce the H/K ratio, a new diagnostic feature defined as the ratio of pseudo-Lick indices Ca II K and Ca II H (Figure 2.5). We verify that it is an excellent tracer of potential contamination of the sample due to star-forming or young populations, confirming that our sample is compatible with no or negligible contamination, with $\langle \text{H/K} \rangle = 0.96 \pm 0.08 (1\sigma)$. Moreover, a selection based on $\text{H/K} < 1.1$ is found to reproduce a NUVrJ selection ([Ilbert et al., 2013](#)) or a $\log \text{sSFR/yr} < -11$ cut (Figure 2.9) with a small percentage of incompleteness ($\sim 15\%$) or contamination ($\sim 15\%$) while requiring a much narrower wavelength range.
4. Using an optimized combination of Lick indices (namely, $\text{H}\delta_{\text{A}}$, CN_1 , CN_2 , $\text{Ca}4227$, $\text{G}4300$, $\text{H}\gamma_{\text{A}}$, $\text{H}\gamma_{\text{F}}$, $\text{Fe}4383$, $\text{Fe}4531$, and C_24668), we measure single-burst stellar age, $[\text{Z}/\text{H}]$, and $[\alpha/\text{Fe}]$ for 140 passive galaxies, without assuming cosmological priors on the maximum value of age as a function of redshift. We also performed an extended analysis to assess the impact of different choices of indices, verifying that our findings are robust against the choice of a different combination of indices.
5. We find trends between $\log \text{age}$, $[\text{Z}/\text{H}]$, $[\alpha/\text{Fe}]$, and the stellar velocity dispersion consistent with those expected from a passively evolving population, with slopes of (0.5 ± 0.1) , (0.3 ± 0.2) , and (0.2 ± 0.1) , respectively (Figure 2.10). This analysis shows,

⁵Documentation available at: <https://pylick.readthedocs.io>

for the first time using individual galaxies, that a relation between $[\alpha/\text{Fe}]$ (a star formation timescale proxy) and σ_* is already in place at $z \sim 0.7$. Furthermore, the age difference of 5.5 Gyr between our sample and local ETG can be fully explained by pure passive evolution (Figure 2.10). Assuming a standard Λ CDM cosmology, the relation between formation redshifts and galaxy stellar masses is found to agree with several previous analyses (Figure 2.11), confirming that this population of massive galaxies forms at $z_{\text{form}} \sim 1.3(1.6)$ at masses $\log M_*/M_\odot = 10.7(11.3)$, after the peak of the cosmic star formation rate density.

6. Even if we do not impose any cosmological prior to the age of the population, the obtained age–redshift evolution is consistent with a Λ CDM universe (Figure 2.12). The stellar $[\text{Z}/\text{H}]$ and $[\alpha/\text{Fe}]$ do not evolve significantly over z and their values are slightly supersolar, $\langle [\text{Z}/\text{H}] \rangle = 0.08 \pm 0.18$ (1σ) dex, and supersolar, $\langle [\alpha/\text{Fe}] \rangle = 0.13 \pm 0.11$ (1σ) dex, compatibly with their local counterparts.
7. Finally, the analysis of median binned relations confirms the downsizing scenario and the passive nature of this population. Remarkably, we obtain two clear nearly parallel age – z relations for both the higher ($\sigma_* \approx 230 \text{ km s}^{-1}$) and the lower ($\sigma_* \approx 200 \text{ km s}^{-1}$) mass regimes. This difference of $\Delta \text{age} \approx 0.5 \text{ Gyr}$ can be interpreted as a delay in formation time between the two, with later formation epochs for the population of less massive galaxies.

Overall, our analysis of individual galaxies confirms the existence of a population of passively evolving galaxies at intermediate redshift that follows a downsizing pattern. This enables the study, for the first time at these redshifts, of the detailed stellar population properties of passive galaxies and their underlying cosmology jointly. We build upon this analysis of stellar population parameters of 140 individual passive galaxies at intermediate redshift to derive cosmological constraints using the cosmic chronometer approach (Borghi et al., 2022b). The main results are summarized below and illustrated in Fig. 2.20.

1. We derive a new direct and cosmology-independent estimate of the Hubble parameter $H(z = 0.75) = 98.8 \pm 33.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$, including both statistical and systematic uncertainties. The latter are obtained by varying the indices adopted to estimate mean stellar ages, the binning scheme, and by assuming different stellar population synthesis models and star formation histories. The accuracy is dominated at the moment by the limited statistics of the sample of cosmic chronometers studied, but nevertheless provides interesting perspectives in light of future large spectroscopic surveys.
2. We propose a simple model to analyze age–redshift relations of cosmic chronometers at different regimes of stellar velocity dispersion σ_* . By assuming a flat Λ CDM universe, we obtain $H_0 = 72_{-19}^{+27} \text{ km s}^{-1} \text{ Mpc}^{-1}$ and a typical formation time of $t_{\text{form}} = 3.2_{-1.3}^{+1.8} \text{ Gyr}$ after the big bang for the high σ_* ($> 215 \text{ km s}^{-1}$) subsample. In this second approach, it will be crucial to improve the reliability of galaxies' absolute ages using very high-quality spectra combined with up-to-date stellar population models.

Figure 2.20 highlights the main contributions of this Thesis. Overall, this work demonstrates that it is possible to extend the cosmic chronometer approach by performing a detailed

study of the stellar populations of individual galaxies with spectral indices, providing at the same time information on galaxy evolution and cosmology.

In view of the extremely interesting constraints to H_0 from gravitational waves (e.g., GW170817; Abbott et al., 2017a) and of the improvements expected in the near future, an important step forward will be the combination of CC and GW analyses to reconstruct for the first time a cosmology-independent measurement of the expansion history of the Universe from $z \sim 0$ to $z \sim 2$.

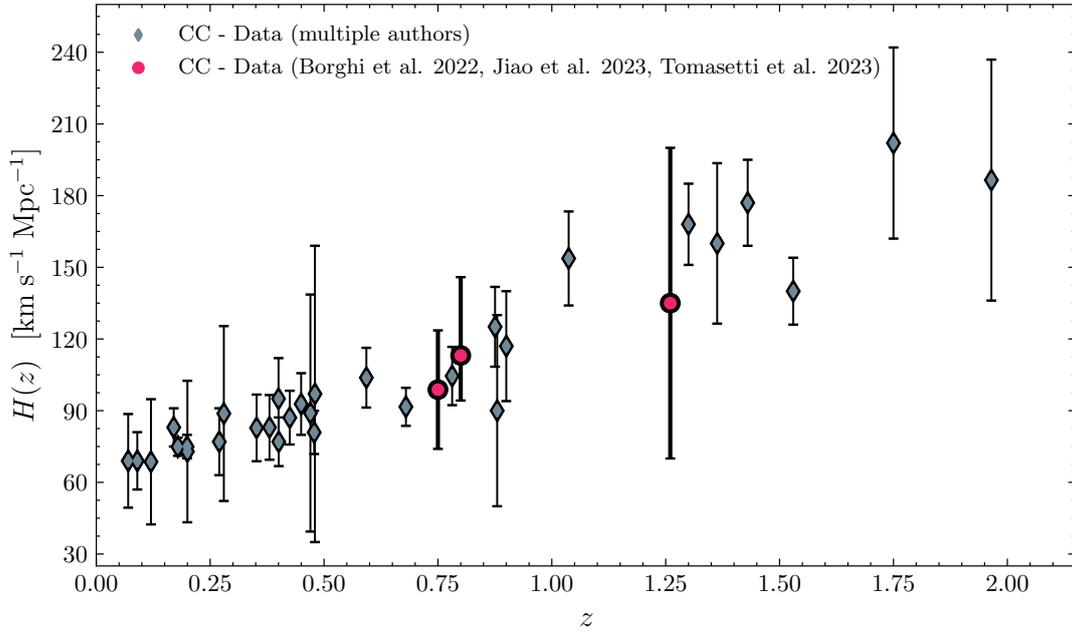


FIGURE 2.20: Constraints on the expansion history of the Universe from cosmic chronometers. The magenta points show results obtained in this Thesis or to which I contributed, while blue points are taken from previous studies in the literature (see references in Section 2.1.1).

Cosmology with Gravitational Waves

The first detection of gravitational waves (GWs) by LIGO interferometers in 2015 (Abbott et al., 2016b) opened a new era of astronomy. However, it was not until two years later, with the detection of the first binary neutron star and its electromagnetic counterpart, that the field of GW cosmology was also revolutionized. In fact, GWs produced from compact binaries can be used as *standard sirens*, as their luminosity distance can be directly measured from the signal alone, requiring no additional calibration. With the additional information on z , it is then possible to constrain cosmological parameters via the luminosity distance–redshift relation (Eq. (1.20)). In particular, since $H_0 \approx cz/d_L$ at low z , GW can provide a direct local measurement of the expansion history of the Universe. However, contrary to the cosmic chronometer case, determining z from current GW data alone is not possible due to its inherent degeneracy with binary masses. When an electromagnetic counterpart of a GW event is detected and its host galaxy is identified, the redshift can be measured directly from the host (e.g., Schutz, 1986; Holz & Hughes, 2005). GW170817 is the only known example so far and currently provides the most stringent constraint to H_0 from GW (Abbott et al., 2017a). The detection rates for these events for ongoing and upcoming observing runs of current GW facilities are quite uncertain, but in general not optimistic (Colombo et al., 2022). When the counterpart is too faint to be detected or absent, two main methodologies have been proposed to obtain redshift information: using the distribution of potential hosts within the GW localization volume as a prior (e.g., Schutz, 1986; Del Pozzo, 2012; Fishbach et al., 2018; Soares-Santos et al., 2019; Gray et al., 2020; Palmese et al., 2020; Finke et al., 2021; Gray et al., 2022), or modeling intrinsic astrophysical properties as the source-frame mass distribution to break the mass–redshift degeneracy (e.g., Chernoff & Finn, 1993; Taylor et al., 2012; Farr et al., 2019; Mastrogiovanni et al., 2021b; Ezquiaga & Holz, 2022).

In this work, we develop a unified approach, enabling joint cosmological and astrophysical parameter inference with gravitational waves and galaxy catalogs. We introduce **CHIMERA**, a novel Python code designed to perform standard sirens analyses in this joint framework. We validate **CHIMERA** using an independent external code and by comparing H_0 constraints obtained for well-localized GW events, including GW170817. In this new framework, we provide forecasts on the precision achievable on H_0 , mass function, and rate evolution of the binary black hole population for the ongoing and upcoming observing runs of the current GW facilities in combination with photometric and spectroscopic galaxy surveys. Section 3.1 provides an introduction to GW theory and cosmology, while the subsequent sections cover the original contributions of this Thesis.

The main reference articles for the analyses presented in this chapter are:

- **Nicola Borghi**, Michele Mancarella, Michele Moresco, Matteo Tagliazucchi, Francesco Iacovelli, Andrea Cimatti, & Michele Maggiore, *Cosmology and Astrophysics with Standard Sirens and Galaxy Catalogs in View of Future Gravitational Wave Observations*, *The Astroph. J.*, accepted (2024) [[arXiv:2312.05302](https://arxiv.org/abs/2312.05302)]
- Michele Mancarella, **Nicola Borghi**, Edwin Genoud-Prachex, Francesco Iacovelli, Michele Maggiore, Michele Moresco, Matteo Schulz, *Gravitational-wave cosmology with dark sirens: state of the art and perspectives for 3G detectors.*, *Proceedings of the 41st International Conference on High Energy physics* [[arXiv:2211.15512](https://arxiv.org/abs/2211.15512)]
- **Nicola Borghi**. *Toward an independent reconstruction of the expansion history of the Universe*, *ESO Hypatia Colloquium* (2022) [[DOI:10.5281/zenodo.7104538](https://doi.org/10.5281/zenodo.7104538)]

Along with this work, we developed and publicly released a new code to perform standard sirens analyses with galaxy catalogs:

 **CHIMERA**: Available at: <https://chimera-gw.readthedocs.io/>

3.1 Framework: theory and observations

This Section serves as an introduction to the gravitational wave theory (§ 3.1.1), describing the main GW sources (§ 3.1.3), and briefly how a GW detector works (§ 3.1.2). Then, it presents the Hierarchical Bayesian Inference formalism, which is the state-of-the-art methodology for standard sirens analysis (§ 3.1.5). Ultimately, it discusses current cosmological constraints obtained with gravitational waves.

3.1.1 Fundamentals of gravitational wave theory

The existence of gravitational waves can be derived in the *weak-field approximation* by adding a small perturbation $h_{\mu\nu}$ to the flat space-time metric $\eta_{\mu\nu}$ (see eq. 1.3):

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad |h_{\mu\nu}| \ll 1. \quad (3.1)$$

Now, thanks to the gauge freedom of the Einstein equations (see eq. 1.1) due to the huge symmetry group of general relativity, i.e. coordinate transformations $x^\mu \rightarrow x'^\mu(x)$, four of its ten components can be constrained with the choice of a gauge condition. The most common choice is the Lorentz gauge, which is defined by

$$\partial^\nu \bar{h}_{\mu\nu} = 0, \quad \text{with} \quad \bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^\alpha{}_\alpha, \quad (3.2)$$

In this gauge, the Einstein Eqs. 1.1 become (see Maggiore, 2007, for a detailed derivation):

$$\square \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}. \quad (3.3)$$

where $\square \equiv -(1/c^2)\partial_t^2 + \nabla^2$ is the flat space D’Alambertian. Observe that Eqs. (3.2) and (3.3) together translate into the conservation of energy and momentum in linearized theory $\partial^\nu T_{\mu\nu} = 0$.

Propagation of GWs

The Eq. (3.3) is a wave equation that describes how GWs are generated by their source $T_{\mu\nu}$. To study how they propagate, it is convenient to evaluate it outside of the matter distribution (i.e., in the vacuum), where $T_{\mu\nu} = 0$. This leads to

$$\square \bar{h}_{\mu\nu} = 0 \quad (\text{outside of the source}). \quad (3.4)$$

By considering the D’Alambertian notation, this implies that GWs propagate at the speed of light. Furthermore, within the Lorentz gauge we can always choose a coordinate system such that $\bar{h}_{\mu\nu} = h_{\mu\nu}$ and $\bar{h}_{0i} = 0$. By doing this, we are imposing 4 conditions (which reduce the degrees of freedom of $h_{\mu\nu}$ from six to two), which define the so-called Transverse-Traceless (TT) gauge. In this gauge, only two components are different from zero

$$h_+ \equiv h_{11} = -h_{22} \quad h_\times \equiv h_{12} = h_{21}. \quad (3.5)$$

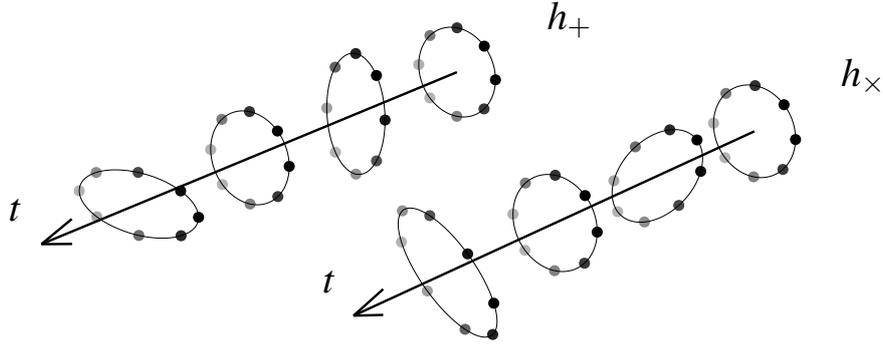


FIGURE 3.1: Deformation of a ring of test masses due to the h_+ and h_\times polarization. The snapshots are taken at $\omega t = 0, \pi/2, \pi, 3/2\pi$.

It can be seen that Eq. (3.4) has plane wave solutions, $h_{ij}^{\text{TT}}(x) = e_{ij}(\mathbf{k}) e^{ik_\mu x^\mu}$, where $k^\mu = (\omega/c, \mathbf{k})$, $|\mathbf{k}| = \omega/c$, and $e_{ij}(\mathbf{k})$ is called polarization tensor. Choosing a single-plane wave along the z axis, the above conditions give

$$h_{ab}^{\text{TT}}(t, z) = \begin{pmatrix} h_+ & h_\times \\ h_\times & -h_+ \end{pmatrix}_{ab} \cos[\omega(t - z/c)], \quad (3.6)$$

where $a, b = 1, 2$ are indices in the transverse plane (x, y) and h_+ and h_\times are defined as the amplitudes of the *plus* and *cross* polarization of the wave, and represent the two degrees of freedom of the gravitational wave.

Effects of GWs on test masses

To evaluate the effect of the passage of a GW, consider two test masses, $A(x_A, y_A, z_A) = (0, 0, 0)$ and $B(x_B, y_B, z_B)$. The proper distance L (see Eqs. (1.2) and (1.7)) in the *weak-field approximation* and TT gauge is

$$L^2 = g_{\mu\nu} (x_B^\mu - x_A^\mu) (x_B^\nu - x_A^\nu) = (\delta_{ij} + h_{ij}^{\text{TT}}) x_B^i x_B^j, \quad (3.7)$$

where δ_{ij} is the Kronecker delta, while $x_B^i = L_0 n^i$, where L_0 is the initial distance between the masses. This equation can be rewritten as

$$L = L_0 \sqrt{1 + h_{ij}^{\text{TT}} n^i n^j}. \quad (3.8)$$

Given the condition on $|h_{\mu\nu}| \ll 1$, the above equation can be approximated as

$$L \simeq L_0 \left(1 + \frac{1}{2} h_{ij}^{\text{TT}} n^i n^j \right) \implies \frac{\delta L}{L_0} \sim \frac{h}{2}, \quad (3.9)$$

where $\delta L \equiv L - L_0$ and h is the amplitude of the GW, also known as *strain amplitude*. Figure 3.1 shows the effect of the individual h_+ and h_\times polarizations of a GW propagating through a ring of test masses.

Sources of GWs

The formal solution for Eq. (3.3) can be described in terms of retarded Green functions with retarded time t_r , thus yielding

$$h_{ij}^{\text{TT}}(t, \mathbf{x}) = \frac{4G}{c^4} \Lambda_{ij,kl}(\hat{\mathbf{x}}) \int \frac{T_{kl}(t_r, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad \text{with} \quad t_r = t - \frac{|\mathbf{x} - \mathbf{x}'|}{c}, \quad (3.10)$$

where the Lambda tensor $\Lambda_{ij,kl}(\hat{\mathbf{x}})$ projects the metric in the TT gauge. This result is general within the weak field approximation, which is valid whenever the fields are sufficiently weak to assume the background to be flat. It is possible to evaluate Eq. (3.10) under two additional conditions: (1) *far zone*, the source, with typical size D , is assumed to be far from the detector, $r \gg D$; (2) *slow velocities*, typical velocities v inside the source are small, $v \ll c$. Through a multipole expansion, it can be seen that the leading term in $\mathcal{O}(v/c)$ is the quadrupole. This yields the quadrupole formula, derived for the first time by Einstein (1918) (see Maggiore, 2007, for the full derivation). The formula reads

$$[h_{ij}^{\text{TT}}(t, \mathbf{x})]_{\text{quad}} = \frac{1}{r} \frac{2G}{c^4} \ddot{Q}_{ij}^{\text{TT}}(t - r/c), \quad (3.11)$$

where the quadrupole moment is defined as

$$Q^{ij} \equiv \int d^3x \rho(t, \mathbf{x}) (x^i x^j - \frac{1}{3} r^2 \delta^{ij}). \quad (3.12)$$

The total radiated power (or, “gravitational luminosity”) in the quadrupole approximation reads

$$P \equiv \frac{dE_{\text{gw}}}{dt} = \frac{G}{5c^5} \langle \ddot{Q}_{ij}^{\text{TT}} \ddot{Q}_{ij}^{\text{TT}} \rangle, \quad (3.13)$$

where $\langle \dots \rangle$ denotes an average over several wavelengths. The observations of the orbital decay of PSR B1913+16 (Hulse, 1994), accurately described with energy loss in the form of gravitational wave radiation, marked a historic confirmation of the validity of Eq. (3.13).

To better understand Eqs. (3.11) and (3.13) it is useful to introduce the following approximations

$$\ddot{Q} \sim \epsilon \frac{M D^2}{T^3} \quad \text{and} \quad \ddot{Q} \sim \epsilon \frac{M D^2}{T^2}, \quad (3.14)$$

where $\epsilon (< 1)$ is a mass distribution asymmetry factor, M the source mass, D its characteristic size and T its characteristic time scale. Substituting the characteristic speed of the source $v = D/T$ and its Schwarzschild radius $R_S = 2GM/c^2$, Eqs. (3.11) and (3.13) become

$$P \sim \epsilon^2 \left(\frac{R_S}{D} \right)^2 \left(\frac{v}{c} \right)^6 \quad \text{and} \quad h \sim \frac{G M v^2}{c^4 r}. \quad (3.15)$$

This formula indicates that an optimal GW source must be asymmetric ($\epsilon \sim 1$), compact ($R_S/D \sim 1$), and relativistic ($v/c \sim 1$). Note that the quadrupole approximation would not apply to these regimes. The amplitude of GWs, which is the observable, decreases linearly with distance r .

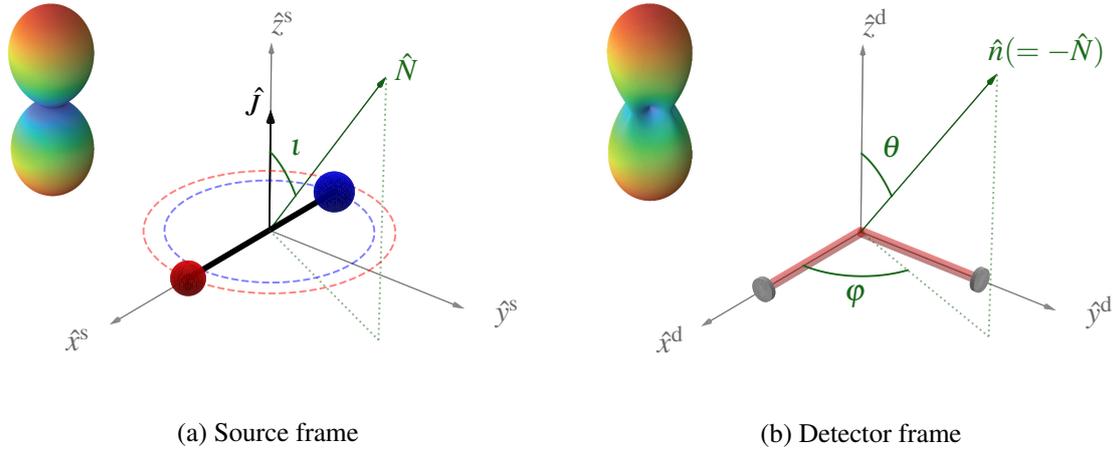


FIGURE 3.2: (a) *Source frame*: defines the coordinate system in which the intrinsic parameters of the binary are defined: masses, spins, and phase. The z axis is anchored to the orbital angular momentum \hat{J} . In this case, the phase of the binary $\phi_c = 0$. (b) *Detector frame*: the two orthogonal arms of the interferometer form the x and y axes. Angles θ and ϕ denote the polar and azimuth angles of the binary in the sky measured in the detector frame (for astronomers: $\theta = \pi/2 - \text{DEC}$ and $\phi = \text{RA}$). Color indicates increasing sensitivity/power from blue to red.

GWs from compact binary coalescence

Compact binaries, as binary black holes (BBHs), binary neutron stars (BNSs), and neutron star–black hole binaries (NSHBs) are relatively simple, yet insightful GW sources. When orbiting in close proximity they are capable of radiating away significant orbital energy through GW emission, so that their orbital separation shrinks and the two objects coalesce over realistic astrophysical timescales. In the so-called quasi-Newtonian limit, expressions are analytical. More precise analytic and semi-analytic post-Newtonian expressions (up to higher orders), which provide a more accurate description up to the merger phase, are known and well studied (e.g., [Buonanno et al., 2009](#)). Starting from the merger phase, full non-linear GR simulations are indeed needed to accurately describe the GW emission.

Consider two point-like masses m_1 and m_2 , with positions \mathbf{r}_1 and \mathbf{r}_2 , and orbital separation $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$, with $r = |\mathbf{r}|$ in Newtonian gravity. The mass density is $\rho(\mathbf{x}) = m_1 \delta(\mathbf{x} - \mathbf{r}_1) + m_2 \delta(\mathbf{x} - \mathbf{r}_2)$, the total mass $m = m_1 + m_2$ and the reduced mass $\mu = m_1 m_2 / m$. The orbital frequency ω_s in the case of circular orbits is described by Kepler’s third law

$$\omega_s^2 = \frac{G m}{r^3}. \quad (3.16)$$

In particular, considering a circular orbit in the $z = 0$ plane (in the center of mass frame), for which $\mathbf{r} = r (\cos(\omega_s t + \pi/2), \sin(\omega_s t + \pi/2), 0)$, it is possible to easily compute the components of the quadrupole moment via Eq. (3.12) and use Eq. (3.11) to get the two GW polarizations (see [Maggiore, 2007](#)). An observer located at distance d with an inclination ι with respect to the orbital angular momentum (see Fig. 3.2a) can write the received GW

signal as

$$\begin{aligned} h_+(t, \iota, \phi_c) &= \frac{4 G \mu \omega_s^2 r^2}{d c^4} \left(\frac{1 + \cos^2 \iota}{2} \right) \cos(2\omega_s t_r + 2\phi_c), \\ h_\times(t, \iota, \phi_c) &= \frac{4 G \mu \omega_s^2 r^2}{d c^4} \cos \iota \sin(2\omega_s t_r + 2\phi_c). \end{aligned} \quad (3.17)$$

where the coalescence phase ϕ_c is the most common arbitrary reference phase. Notice that the GW frequency is twice the orbital frequency, $\omega_{\text{gw}} = 2\omega_s$. When seen edge-on ($\iota = \pi/2$), h_\times vanishes and the GW is linearly polarized. When seen face-on ($\iota = 0$), h_+ and h_\times have the same amplitude and are out of phase by $\pi/2$, therefore the radiation is circularly polarized. At intermediate values of ι the radiation has an elliptic polarization. The angular distribution of the radiated power, calculated with Eq. (3.13), is illustrated in Fig. 3.2a. Note that radiation is maximum at $\iota = 0$, but never zero, since there is always a component of the source motion orthogonal to the line of sight. Using Eq. (3.16) and defining a key quantity called the *chirp mass*

$$\mathcal{M} = \mu^{3/5} m^{2/5} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \quad (3.18)$$

it is possible to rewrite Eq. (3.17) as

$$\begin{aligned} h_+(t, \iota, \phi_c) &= \frac{4}{d} \left(\frac{G \mathcal{M}}{c^2} \right)^{5/3} \left(\frac{\pi f_{\text{gw}}}{c} \right)^{2/3} \left(\frac{1 + \cos^2 \iota}{2} \right) \cos(2\pi f_{\text{gw}} t_r + 2\phi_c), \\ h_\times(t, \iota, \phi_c) &= \frac{4}{d} \left(\frac{G \mathcal{M}}{c^2} \right)^{5/3} \left(\frac{\pi f_{\text{gw}}}{c} \right)^{2/3} \cos \iota \sin(2\pi f_{\text{gw}} t_r + 2\phi_c), \end{aligned} \quad (3.19)$$

where $f_{\text{gw}} = \omega_{\text{gw}}/(2\pi)$, with $\omega_{\text{gw}} = 2\omega_s$. Thus, in this lowest-order Newtonian approximation, the GW signal depends exclusively on \mathcal{M} , not on the individual masses m_1 and m_2 . Hence, this is the mass combination that is best measured in a GW observation. With the same definitions, applying Eq. (3.13), it is possible to compute the total power radiated in GWs

$$P_{\text{gw}} = \frac{32 G}{5 c^5} \left(\frac{\pi G \mathcal{M} f_{\text{gw}}}{c^3} \right)^{10/3}. \quad (3.20)$$

Because GW emission costs energy, the orbital separation r must decrease over time. But then, since $\omega_s^2 \propto r^{-3}$, the frequency of the GW increases, and so does the power as a consequence, thus inducing a further decrease on r . This runaway process ends with coalescence. When $\dot{\omega}_s \ll \omega_s$, the orbits can still be approximated with circular motions. In this regime of quasi-circular motion, valid for actual astrophysical sources in the inspiral phase, it is possible to easily compute the orbit and GW frequency evolution over time. By equating Eq. (3.20) with the orbital energy loss, $P = -dE_{\text{orb}}/dt$, with the binary orbital energy $E_{\text{orb}} = -G m_1 m_2 / (2r)$, one gets

$$\dot{f}_{\text{gw}} = \frac{96}{5} \pi^{8/3} \left(\frac{G \mathcal{M}}{c^3} \right)^{5/3} f_{\text{gw}}^{11/3}, \quad (3.21)$$

whose integration gives

$$f_{\text{gw}}(\tau) = \frac{1}{\pi} \left(\frac{5}{256} \frac{1}{\tau} \right)^{5/3} \left(\frac{G\mathcal{M}}{c^3} \right)^{-5/8}, \quad (3.22)$$

where $\tau \equiv t_c - t$ is the time to coalescence. This increase in frequency as the binary inspirals inward is commonly known as the “chirp” of the gravitational wave. For real astrophysical sources, the maximum frequency before the merger is taken to be at the inner-most stable circular orbit (ISCO) of the Schwarzschild spacetime, $r_{\text{ISCO}} = 6Gm/c^2$. When the ISCO is approached, the dynamic is dominated by strong field effects. The phase before this event is known as *inspiral*, and the subsequent *merger*. Rewriting Eq. (3.21) in astrophysically convenient units, one finds

$$\langle h \rangle \simeq 10^{-23} \left(\frac{\mathcal{M}}{1.2 M_\odot} \right)^{5/3} \left(\frac{f_{\text{gw}}}{100 \text{ Hz}} \right)^{2/3} \left(\frac{d}{100 \text{ Mpc}} \right)^{-1}, \quad (3.23)$$

$$f_{\text{gw}} \simeq 134 \text{ Hz} \left(\frac{\mathcal{M}}{1.2 M_\odot} \right)^{-5/8} \left(\frac{\tau}{1 \text{ s}} \right)^{-3/8}, \quad (3.24)$$

$$\tau = \frac{f_{\text{gw}}}{\dot{f}_{\text{gw}}} \simeq 2.2 \text{ s} \left(\frac{\mathcal{M}}{1.2 M_\odot} \right)^{-5/3} \left(\frac{f_{\text{gw}}}{100 \text{ Hz}} \right)^{-8/3}, \quad (3.25)$$

where $\langle \cdot \rangle$ indicates an average over detector and source orientations, while the reference value of $\mathcal{M}_c = 1.2 M_\odot$ is set to match the one of a BNS with masses $m_1 = m_2 = 1.4 M_\odot$. In this case, at $f_{\text{gw}} = 10 \text{ Hz}$ (the lowest accessible with current interferometers), one has access to the last $\sim 17 \text{ min}$ of the coalescence, while for $f_{\text{gw}} = 100 \text{ Hz}$ and $f_{\text{gw}} = 1 \text{ kHz}$, one gets $\tau \sim 2.2 \text{ s}$ and $\tau \sim 4 \text{ ms}$, respectively. For these same frequencies, the corresponding orbital separation is $r \simeq 720, 155, 33 \text{ km}$. Note that, only compact objects (e.g., BHs and NSs) can achieve these small orbital separations.

GW signal modeling

In the general case of quasi-circular orbits, the parameters needed to describe the GW signal from a compact binary coalescence are 15 for a BBH, 17 for a BNS, and 16 for a NSBH (e.g., see Maggiore, 2007),

$$\boldsymbol{\theta} = \{ \mathcal{M}_c, \eta, d_L, \theta, \varphi, \iota, \psi, t_c, \Phi_c, \chi_{1,x}, \chi_{2,x}, \chi_{1,y}, \chi_{2,y}, \chi_{1,z}, \chi_{2,z}, \Lambda_1, \Lambda_2 \}, \quad (3.26)$$

where \mathcal{M}_c is the detector–frame chirp mass, η the symmetric mass ratio, d_L the luminosity distance to the source, θ and ϕ are the sky coordinates defined as $\theta = \pi/2 - \text{DEC}$ and $\phi = \text{RA}$ (with RA and DEC right ascension and declination, respectively), ι the inclination angle of the binary orbital angular momentum with respect to the line of sight, ψ the polarization angle, Φ_c the phase at coalescence, $\chi_{i,j}$ the dimensionless spin of the object $i = \{1, 2\}$ along the axis $j = \{x, y, z\}$, and Λ_i the dimensionless tidal deformability, which is zero for a BH. Here is a brief summary of current techniques used to simulate GW signal from compact binary coalescence:

- **Post-Newtonian (PN) approximation:** Provides analytical predictions of the waveform phase and amplitude as a series expansion in v/c from Eq. (3.10). Valid in the early inspiral regime. Restricted PN simplifies the PN approach by omitting specific corrections for easier analytical and numerical implementation at the cost of accuracy. An example is the `TaylorF2_RestrictedPN` waveform model (Buonanno et al., 2009; Ajith, 2011; Mishra et al., 2016).
- **Numerical Relativity (NR):** Directly solves Einstein’s equations numerically on high-performance computers. Provides “exact” waveforms in the late inspiral, merger, and ringdown but is highly computationally intensive, limiting their full applicability. (see Bishop & Rezzolla, 2016).
- **Effective One-Body (EOB):** Maps the two-body problem to an effective particle orbiting an effective BH. Predicts full inspiral-merger-ringdown waveforms including spin effects. Constructed by calibrating analytical effective one body dynamics to numerical relativity (see Taracchini et al., 2014).
- **Inspiral–Merger–Ringdown (IMR):** Combines PN inspiral with numerical relativity results in the late-inspiral and merger. Provides a continuous waveform across all regimes using phenomenological fits. The `IMRPhenomD` model (Husa et al., 2016; Khan et al., 2016) is one of the most widely used for BBHs, while its tidal extension `IMRPhenomD_NRTidalv2` better suits for BNSs. To date, the `IMRPhenomHM` full inspiral–merger–ringdown model (London et al., 2018; Kalaghatgi et al., 2020) is the most accurate for BBHs, as it into account not only the quadrupole of the signal, but also the subdominant multipoles $(l, m) = (2, 1), (3, 2), (3, 3), (4, 3),$ and $(4, 4)$. Finally, `IMRPhenomNSBH` (Pannarale et al., 2015; Dietrich et al., 2019) targets NSBH analyses, including tidal effects and higher-order spin terms calibrated to NR waveforms.

3.1.2 Gravitational wave detectors

The passage of a GW can be observed by probing changes in the proper distance between test masses (Eq. (3.9)). The pioneering idea of using a light signal between two freely gravitating test masses is already present in the work by Pirani (1956). However, for several decades, starting from (Weber, 1960), the experimental searches were carried out mainly using resonant mass bar detectors. These aluminum or niobium bars are kept at cryogenic temperatures and are attached to sophisticated vibration isolation systems. The idea is to transduce the GW strain into vibrations that are read electronically. In 1972, R. Weiss carried out a feasibility study proposing to use Michelson-Morley interferometers as GW detectors (R., 1972) and laying the foundations for current detectors. The global network of GW detectors now includes five interferometers: the two Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors in Handford and Livingston in the United States; Virgo in Italy; Kamioka Gravitational Wave Detector (KAGRA) in Japan; and GEO600 in Germany.

Earth-based interferometers

A diagram of the structure of a typical GW interferometer is shown in Fig. 3.3. In brief, it is constituted by two perpendicular arms (L-shaped) of the same length (4 km for LIGO, 3 km

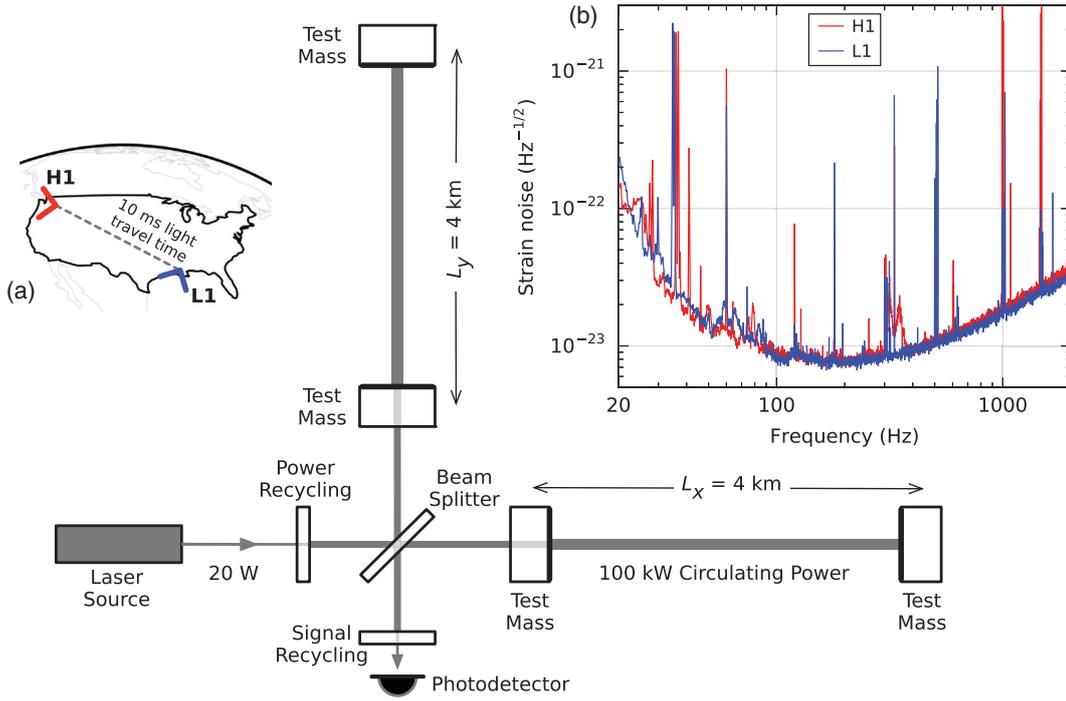


FIGURE 3.3: Simplified scheme of an Advanced LIGO detector. (a) Location and orientation of the LIGO detectors at Hanford (H1) and Livingston (L1); (b) The instrument noise for each detector near the time of the first detection of GW150914. Reproduced from [Abbott et al. \(2016b\)](#).

for Virgo and KAGRA) with two mirrors in free fall at their ends acting as test masses. A monochromatic laser beam is split in the two arms, which travels back and forth multiple times, then recombines at the beam-splitter, and is later recorded with a photo-detector. With no GWs, the recombination is set to produce destructive interference. When a GW crosses the interferometer, it changes the relative light travel time between the two arms, so that the signal does not combine destructively anymore. Note that the typical strain for a BBH system with masses of $m_{1,2} = 30 M_{\odot}$ located at 100 Mpc is $h \sim 10^{-21}$ (see Eq. (3.9)). Therefore, Eq. (3.9) implies that detectors with arms of the order of kilometers have

$$\delta L \sim 10^{-18} \left(\frac{h}{10^{-21}} \right) \left(\frac{L}{\text{km}} \right) \text{ m}. \quad (3.27)$$

To achieve this enormous precision, modern interferometers feature key components. First of all, they operate in vacuum, with the mirrors and beam splitter isolated via multistage suspensions to attenuate seismic noise. The light bounces back and forth in each arm using a Fabry-Pérot cavity to increase the effective optical path L (up to 1200 km in the case of LIGO). Then, a partially transmissive power-recycling mirror is placed at the input and at the output to resonantly enhance the laser power. A 1064 nm laser illuminates LIGO, Virgo, and KAGRA interferometers after amplitude, frequency, and beam shape stabilization. The most recent KAGRA detector is built underground and uses cryogenic mirrors to lower its thermal noise and improve sensitivity.

Antenna pattern

The tensor $h_{ij}^{\text{TT}}(t, \mathbf{x})$ becomes a scalar quantity $h(t)$ when projected onto the detector. The value of $h(t)$ depends on the geometry of the detector and its relative position with respect to the source (θ, φ) . This effect is encoded by the antenna pattern functions $F_{+, \times}$

$$h(t) = F_+(\theta, \varphi) h_+(t) + F_\times(\theta, \varphi) h_\times(t) \quad (3.28)$$

In Fig. 3.2 is shown the antenna pattern of a typical L-shaped interferometer, averaged over the two polarizations. The detector has good coverage of the sky, except for four blind spots along the directions defined by the arms. As Eq. (3.28) depends on four variables (h_+ , h_\times , θ , φ), with a single detector it is impossible to locate the source in the sky. Using two interferometers, one has three measurements, the strains $s_1(t)$, $s_2(t)$, and the time delay between them τ_{12} ; which are enough to solve Eq. (3.28). With three interferometers, one has in fact five measurements, three strains and two independent delay times, so that Eq. (3.28) can be solved.

Noise characterization

The general expression for the total output $s(t)$ of the detector can be described as the sum of a noise term $n(t)$ and the GW-induced strain $h(t)$,

$$s(t) = n(t) + h(t), \quad (3.29)$$

hence the detection problem consists of the separation between $h(t)$ and $n(t)$ contributions. By accurately knowing $h(t)$ from GR it is possible to extract the GW signal even in the situation where $|h(t)| \ll |n(t)|$ via the *matched filtering* technique. This process involves comparing the measured data with various expected signal templates in order to maximize the signal-to-noise ratio (SNR). When doing this, the fact that the detector sensitivity changes at different frequencies must be taken into account. Under the assumption that $n(t)$ is stationary in time, the matched filter SNR in a single detector is given by

$$\text{SNR}^2 = 4 \int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)} df, \quad (3.30)$$

where $\tilde{h}(f)$ is the GW strain in the Fourier domain. The term $S_n(f)$ is the *power spectral density* (PSD) of the detector, a commonly adopted metric to describe noise, expressed in units of Hz^{-1} . Alternatively, one can employ the *amplitude spectral density* (ASD), corresponding to $\sqrt{S_n(f)}$, with units of $\text{Hz}^{-1/2}$. Lastly, an additional quantity that is useful when studying the signal expected from sources, although not directly measurable, is the *characteristic strain* of the noise $\langle n \rangle$ averaged in time.

The main noise sources of a detector can be classified into two general classes: *displacement noises*, which work by directly moving the test masses, and *phase noises*, which cause phase fluctuations of the optical field used to measure the GW strain (Adhikari, 2014). As a guideline, Fig. 3.3b illustrates the noise ASD near the time of the first GW detection. Displacement noises are concentrated at lower f and include thermal noise from optical elements, seismic vibrations, Newtonian gravity noise (including density fluctuations in the atmosphere and

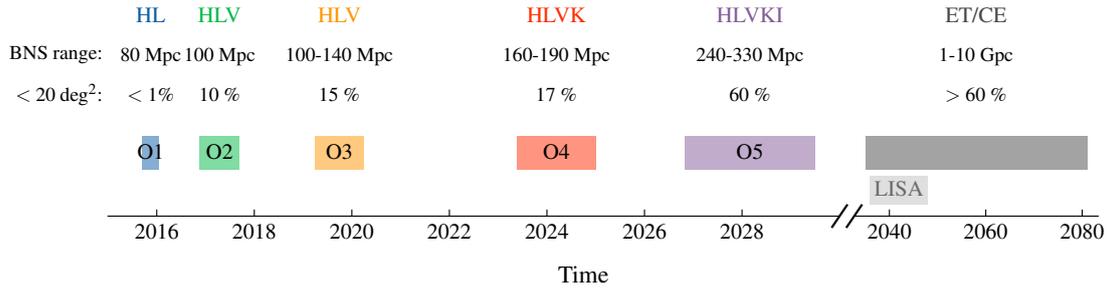


FIGURE 3.4: Timeline of the past, present, and future GW observing runs. The nomenclature used here is H=Hanford, L=Livingston, V=Virgo, K=KAGRA, and I=IndIGO. The second and third rows indicate the predicted BNS horizon for a single-detector SNR threshold of 8 and the percentage of events with localization uncertainties less than 20 deg². A five-year gap separates second-generation (2G) and third-generation (3G) detectors.

surface waves on the ground), and electromagnetic coupling (including cosmic rays, ambient magnetic fields, and surface charges on the instruments). In particular, these noises impose a low-frequency limit of a few Hz for all terrestrial laser interferometers. Phase noises include quantum vacuum fluctuations, scattering from the residual gas, and backscatter from mirror imperfections. Quantum vacuum fluctuations are a fundamental limit to the detection of optical phase shifts and constitute the dominant effect at large $f > 150$ Hz. Finally, a typical noise ASD includes narrow-band features, in the case of LIGO detectors in Fig. 3.3b, these are: calibration lines (33–38, 330, and 1080 Hz), vibrational modes of suspension fibers (500 Hz and harmonics), and 60 Hz electric power grid harmonics.

Past, present and future observations

Figure 3.4 presents a timeline of the GW observations. The first observing run (O1) was conducted by the LIGO and Virgo collaboration with the two LIGO detectors in Hanford and Livingston. On 14 September 2015, the landmark detection of GW150914 was made, representing the first direct observation of GWs. It was produced by a nearly equal mass BBH of about 36 and 29 M_{\odot} located at about 440 Mpc, demonstrating that these systems can form in nature and merge within a Hubble time. This discovery was awarded in 2017 the Nobel Prize in Physics.

Virgo joined the network for the second observing run (O2), significantly improving the sky localization capabilities. On 17 August 2017, the network detected the first BNS merger, GW170817, opening the multimessenger astronomy era. The event lasted ~ 100 s and was followed, 1.7 s later, by a short gamma-ray burst (GRB 170817A) observed with Fermi and INTEGRAL satellites. After a few hours, the host galaxy was identified with NGC 4993 and a UV/optical/IR kilonova emission (Pian et al., 2017) was detected. Radio (Hallinan et al., 2017) and X-ray (Troja et al., 2017) emissions attributed to the GRB afterglow were observed in the same region, as also confirmed by high-resolution radio observation 207 days after the event (Ghirlanda et al., 2019). At the end of O2, significant technological advances were made for the LIGO and Virgo detectors, resulting in a $\sim 50\%$ improvement in the range within which signals can be detected.

After the last completed observing run (O3) the catalog of confident GW detections (GWTC-3 [Abbott et al., 2023b](#)) now includes 90 GW events. Among them, there are very interesting systems. For example, GW190521 is the event with the heaviest BHs, with masses of about 85 and 66 M_{\odot} and a final estimated mass of 150 M_{\odot} ([Abbott et al., 2020a](#)). At least one of them falls into the so-called upper mass gap (about 60 – 120 M_{\odot}), where theoretical models of single-star evolution predict no BH (see e.g., [Spera & Mapelli, 2017](#)). GW190814 is an event characterized by highly asymmetric masses, with $m_1 = 2.6 M_{\odot}$ and $m_2 = 23 M_{\odot}$. The lighter object remains enigmatic, as it could be the heaviest NS or the lightest BH ever observed, thus falling in the so-called lower mass gap (about 2.5 – 5 M_{\odot}). Both of these events pose considerable challenges for standard stellar evolutionary models. Finally, GW200105_162426 and GW200115_042309 represent the first ever observed NSBH mergers. At the time of writing, the O4 run is well underway, with KAGRA joining the network, and 50 significant detection candidates in five months. Detailed planning for the post-O4 period and a possible post-O5 phase is still in progress and may result in significant changes to both target sensitivities and dates. In this phase, a third LIGO detector is planned to be built in India, and the target sensitivity of the LIGO detectors should reach the design sensitivity.

Figure 3.5 presents the characteristic noise strain of LIGO in its initial configuration (aLIGO-O1) and its target sensitivity (aLIGOD), as well as the target sensitivity for future GW detectors. Current ground-based interferometers are most sensitive in the 10 – 1000 Hz range. Future third-generation detectors aim to further improve the sensitivity and widen the observable bands. Among these, one of the most ambitious projects is the Einstein Telescope (ET, [Punturo et al., 2010](#); [Maggiore et al., 2020](#); [Branchesi et al., 2023](#)), a future underground detector, which will be provided with quantum technologies to reduce light fluctuations, mechanisms to reduce local noise, and a sophisticated cryogenic system. This would provide conditions for observing frequencies of 10 – 10^4 Hz. Another proposed third-generation detector is the Cosmic Explorer (CE, [Reitze et al., 2019](#); [Evans et al., 2021](#)), consisting of two L-shaped interferometers, one with 40 km and the other with 20 km long arms. The idea for CE is to further improve LIGO technologies to greatly improve the sensitivity in the LIGO band.

Space-based detectors operate on similar principles as ground-based interferometers, but are capable of sensing GWs at much lower frequencies. This is possible because they can both have much larger arms and be unaffected by seismic noise. The Laser Interferometer Space Antenna ([Amaro-Seoane et al., 2017](#)) would consist of a constellation of three satellites with arms of length $L = 2.5$ million km working in the frequency range of 10^{-4} – 10^{-1} Hz. The LISA Pathfinder spacecraft already tested the possibility of achieving the desired accuracy on the laser interferometry between two test masses in perfect free fall. The noise for LISA can be categorized as acceleration noise (e.g., Brownian motion, thermal radiation pressure, and magnetic forces), mainly contributing at low frequencies, and position noise (e.g., laser shot noise), mainly dominating at high frequencies.

To conclude, it is worth mentioning the International Pulsar Timing Array (IPTA), a collaboration made by the European Pulsar Timing Array (EPTA), the NANO-hertz Observatory for Gravitational Waves (NANOgrav), the Indian Pulsar Timing Array Project (InPTA), the Parkes Pulsar Timing Array (PPTA), and the Chinese PTA (CPTA). These observatories aim to detect nano-Hz GWs through the measurement of changes in the time-of-arrival on Earth of millisecond pulsars.

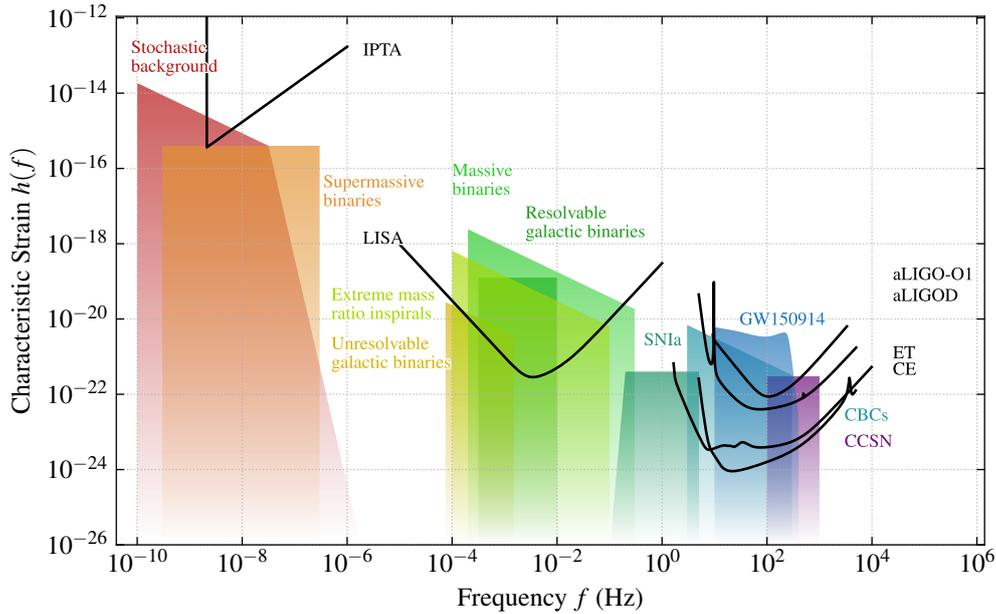


FIGURE 3.5: Characteristic strain as a function of frequency for various detectors and sources. The area between the detector’s curve and the top of the a source box indicates the loudness of the expected signal. Created using [nicoborghi/streamlit-gwplotter](https://github.com/nicoborghi/streamlit-gwplotter) and references therein.

3.1.3 Gravitational wave sources

GW sources emit over a broad range of frequencies, ranging from thousands of Hz to very low frequencies (down to 10^{-16} Hz). The expected strain as a function of frequency is presented in Fig. 3.5.

Stellar collapse. The collapse of high-mass stars can contribute to GWs in two ways, as they can both generate direct GW emission and lead to the formation of compact remnants, which may become binaries and then GW sources. There are two primary mechanisms that trigger supernova (SN) explosions. The first involves the core collapse of massive stars (CCSNs), historically classified as type Ib, Ic, or II depending on their observed spectra. This process typically yields a neutron star or a black hole as a remnant. The second mechanism involves the thermonuclear explosion of a carbon–oxygen white dwarf in a binary system (SNIa). In the most commonly accepted scenario, known as *single-degenerate*, the companion is a red giant (or even a main-sequence star if the system is close enough). Alternatively, in the *double-degenerate* scenario, the companion is a second white dwarf. In contrast to CCSN, SNIa leave no compact remnants. The SNIa explosion occurs when a specific criterion is met, i.e., when the white dwarf mass reaches the Chandrasekhar mass of $1.4 M_{\odot}$. Therefore, their luminosity is, at first approximation, constant. As discussed in Section 1.4.1, this makes them ideal standardizable candles. The rate of SNe explosions in Milky Way-like galaxies is 0.1-0.01 per year. Due to the complexity of modeling stellar collapse and its evolution, assessing how GWs are produced from SNe demands detailed numerical simulations. In a typical SN, simulations

Table 3.1. Number of confident GW events and inferred local rate \mathcal{R}_0 taken from the LIGO-Virgo KAGRA GWTC-3 catalog (Abbott et al., 2023b).

Type	N_{ev}	\mathcal{R}_0 [$\text{Gpc}^{-3} \text{yr}^{-1}$]
BBH	85	18-44
BNS	2	10-1700
NSBH	4	8-140
BH-gap	2	<25

suggest that GWs may extract about $10^{-7\div-5} M_{\odot}$ of the total mass-energy budget in the form of a burst with frequency 100 – 1000 Hz.

Compact binary coalescences. As presented in Section 3.1.1 compact binaries are optimal GW sources and the only type of sources that have been detected to date. After the three observing runs (O1–O3), the current catalog includes 93 events, mostly BBHs, but also includes two BNSs and two neutron star–black hole (NSBH) binaries (Abbott et al., 2023b). Due to the higher statistics, BBHs are the best-studied population. Multiple studies find that the preferred model to describe the black holes mass distribution is a multi-peak model with peaks at ~ 10 and $35 M_{\odot}$ (Leyde et al., 2022; Mancarella et al., 2022a). The rate of BBH mergers is observed to increase with z at a rate proportional to $(1+z)^{\gamma}$ with $\gamma = 2.9_{-1.8}^{+1.7}$. The observed black hole spins are small, with half of the spin magnitudes below $\chi \approx 0.25$. The NS mass distribution is usually assumed to be flat between 1 and $3 M_{\odot}$. From Eqs. (3.23) and (3.24) it can be estimated that these sources, when observed in their last tens of seconds of inspiral, emit in the 10 – 300 Hz range.

Galactic binaries. This category comprises primarily white dwarfs, but also neutron stars and black holes of stellar origin in various combinations. These systems are a target for future space detectors such as LISA (Amaro-Seoane et al., 2017). For systems that can be detected, the orbital periods $\omega = 2/f_{\text{gw}}$ can often be measured with high accuracy, thus their properties can be inferred. An astrophysical background may be present due to the superposition of many unresolvable binaries. In general, the frequency range of galactic binaries is about $3 \times 10^{-4} - 10^{-2}$ Hz.

Spinning neutron stars. Since the detection of the first pulsating radio source PSR B1919+21 by Jocelyn Bell (Hewish et al., 1968), and the association with rapidly rotating neutron stars (Pacini, 1968), pulsars have shed light on strong-field gravity, the equation of state of nuclear matter, evolutionary paths for massive binary systems, and much more. Neutron stars are predicted to have small asymmetric deformations from the ideal spherical shape due to their rapid rotation (approaching 1 kHz) and intense internal magnetic fields (Bonazzola & Gourgoulhon, 1996). These deformations cause the neutron star to possess a time-varying

quadrupole moment. According to Eq. (3.11) this will generate gravitational waves. This radiation takes the form of long-duration monochromatic signals or *continuous gravitational waves*.

Massive black-hole binaries. The origin of massive black holes at the centers of today's galaxies is unknown. Current studies predict the masses of their seeds in the interval between about $10^3 - 10^5 M_\odot$ and formation redshifts $10 < z < 15$ (Volonteri, 2010). If these binaries shrink the binary orbit up to the pc scale, detectable GWs may be produced. In their hierarchical growth, the mergers inevitably cross the LISA frequency spectrum, with a signal that may last hours up to months (see Eq. (3.25)). Overall, observations of massive and supermassive black hole binaries spanning nano- and megahertz frequencies will illuminate the hierarchical growth of structure in the early Universe through precision black hole physics and multiband gravitational wave astronomy.

Extreme Mass Ratio Inspirals. Extreme Mass Ratio Inspirals (EMRIs) describe the long-lasting inspiral (from months to a few years) and merger of stellar origin BHs ($10 - 60 M_\odot$) with intermediate massive black holes ($10^{5-6} M_\odot$). The orbits of EMRIs become highly relativistic, and the large number of orbital cycles enables precise measurements of the parameters of the system and gravity around massive objects. The frequency width of the box is uncertain, as EMRI can in principle occur around a black hole of any mass, and hence emit at any frequency. The box in Fig. 3.5 describes the characteristic strain expected from a binary with $m_1 = 10^6 M_\odot$ and $m_2 = 10 M_\odot$ located at 1 Gpc.

Cosmological background. Ongoing theoretical and observational work aims to realize the discovery potential of the gravitational wave cosmological background as a novel tool to study the early Universe (see Caprini & Figueroa, 2018, for a review). These GW signals are difficult to predict, as they may probe unknown physics. However, one of the major challenges in this context would be to model and subtract the astrophysical background of unresolved sources. Recently, the pulsar timing collaboration (NANOGrav, EPTA, IPTA, CCPTA) detected a signal compatible with a stochastic GW signal, with a confidence level of $\sim 4\sigma$ (e.g., see Agazie et al., 2023). Whether it is produced by the superposition of unresolved astrophysical sources or the cosmological background is still under debate.

3.1.4 Gravitational wave cosmology

The concept of using gravitational waves as standard distance indicators comes from Schutz (1986). The work discusses how the GW signal contains enough information to determine the distance d from a binary, independently of any assumptions on the masses of the components. This is due to the exact inverse mass scaling of h and $\tau = f/\dot{f}$ (see Eqs. 3.23 and 3.25), which are both observable, so that by measuring the ratio h/τ it is possible to determine d . However, as also noted in the original paper, the measurement is not straightforward, as Eq. (3.19) implies that the observed strain is a function of the orientation and position between the source and the detector. This can be partially solved provided that a network of interferometers observes the same event, since knowing the degree of polarization of the wave h_+/h_\times it is possible to constrain the ι , and thus break the distance-inclination degeneracy.

Another possibility to break this degeneracy is when the binary has a precessing spin that causes a modulation of the amplitude that can disentangle ι .

GWs propagation in GR

A formalization of the evolution of tensor perturbations over FLRW in GR can be obtained from Eqs. (1.6) and (3.4). When written in Fourier space and as a function of conformal time η , with $d\eta \equiv dt/a(t)$, it reads

$$\tilde{h}''_{+,x} + 2\mathcal{H}\tilde{h}'_{+,x} + c^2k^2\tilde{h}_{+,x} = 0, \quad (3.31)$$

where $\tilde{h}_{+,x}(\eta, \mathbf{k})$ are the Fourier-transformed GW amplitude for each polarization, the prime denotes the derivative with respect to η , $a(\eta)$ is the FLRW scale factor and $\mathcal{H} \equiv a'/a$. This corresponds to a wave equation with a friction term due to cosmic expansion. This ‘‘Hubble friction’’ results in a redshift of the frequencies $f \rightarrow f/(1+z)$ and a rescaling of the GW amplitude $h \rightarrow h/a$. The formulae derived for compact binary coalescence (eqs. 3.19–3.25) are still valid if the chirp mass is substituted with the *redshifted chirp mass* \mathcal{M}_z and the physical distance $d \cdot a$ with the luminosity distance d_L , defined as

$$\mathcal{M}_z = (1+z)\mathcal{M}_c \quad \text{and} \quad d_L = (1+z) \int_0^z \frac{c}{H(z)} dz. \quad (3.32)$$

For example, using Eq. (3.19), the gravitational radiation received by a detector on Earth takes the following form

$$h_{+,x} \propto \frac{\mathcal{M}_z^{5/3} f_{\text{gw}}^{2/3}}{d_L} F_{+,x}(\text{angles}) \cos \Phi(t), \quad (3.33)$$

where the function F encodes the relative orientation of the binary system with respect to the detector.

GWs propagation in Modified Gravity

A further generalization of Eq. (3.31) is the following (Belgacem et al., 2018b)

$$\tilde{h}''_{+,x} + 2\mathcal{H}[1 - \delta(\eta)]\tilde{h}'_{+,x} + c^2k^2\tilde{h}_{+,x} = 0, \quad (3.34)$$

where the function $\delta(\eta)$ encodes the deviation from GR in tensor perturbation. In general, $\delta(\eta)$ can describe a large class of modified gravity models which imply a ‘‘friction term’’ in the tensor perturbation.¹ In this case, the quantity obtained from the GW data is no longer the standard luminosity distance d_L^{em} as for electromagnetic data, but a ‘‘GW luminosity

¹A modification of the $k^2\tilde{h}_{+,x}$ is also possible, inducing a speed of GWs, c_{gw} , different from that of light. However, the observation of GW170817 and GRB 170817A constrained $|c_{\text{gw}} - c|/c < \mathcal{O}(10^{-15})$ (see Abbott et al., 2017d)

distance” d_L^{gw} that is related to the former through the relation.

$$d_L^{\text{gw}}(z) = d_L^{\text{em}}(z) \exp \left\{ - \int_0^z \frac{\delta(z')}{1+z'} dz' \right\}. \quad (3.35)$$

A convenient parameterization of the ratio $d_L^{\text{gw}}/d_L^{\text{em}}$ in terms of two parameters (Ξ_0, n) , as proposed by [Belgacem et al. \(2018b\)](#), is

$$\frac{d_L^{\text{gw}}(z)}{d_L^{\text{em}}(z)} = \Xi_0 + \frac{1 - \Xi_0}{(1+z)^n} \quad \text{and} \quad \delta(z) = \frac{n(1 - \Xi_0)}{1 - \Xi_0 + \Xi_0(1+z)^n}, \quad (3.36)$$

with $\Xi_0 = 1$ (for all n) in GR and can reach 1.80 for some modified gravity models (e.g., [Maggiore, 2014](#)). This parameterization reproduces the fact that $d_L^{\text{gw}}/d_L^{\text{em}} \rightarrow 1$, hence $\delta(0) = n(1 - \Xi_0)$, in the local Universe, as at small distances there can be no effect from modified propagation. On the other hand, as $z \rightarrow \infty$, $d_L^{\text{gw}}/d_L^{\text{em}}$ saturates to the constant value Ξ_0 . The analysis of GWTC-3 ([Abbott et al., 2023b](#)) BBHs with the spectral sirens method yields $\Xi_0 = 1.2 \pm 0.7$ at 68% CL ([Mancarella et al., 2022b](#)).

GWs as standard sirens

Being able to determine d_L directly from the observations is remarkable in itself. From the $d_L - z$ relation, GW events in the local Universe provide a direct measurement of $H_0 = cz/d_L + \mathcal{O}(z^2)$, while at higher z they also allow exploring the dark energy sector, through the study of the dark energy equation of state and to constrain the phenomenon of “modified GW propagation” ([Belgacem et al., 2018a, 2019](#)). However, due to the inherent degeneracy between redshift and binary masses (see Eq. 3.32), determining z from GW data is not possible and external information is required. Current gravitational wave cosmology techniques can be categorized as follows:

- **Bright sirens.** When an electromagnetic (EM) counterpart is detected and its host galaxy identified, the redshift can be directly measured by spectroscopy ([Schutz, 1986](#); [Holz & Hughes, 2005](#); [Nissanke et al., 2010](#); [Chen et al., 2018](#)). The single example to date, GW170817, provides a measurement of $H_0 = 70.0_{-8.0}^{+12.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$ ([Abbott et al., 2017a,c,b](#)). As the redshift can be measured with spectroscopic accuracy, the uncertainty on H_0 is typically dominated by the uncertainty on d_L measured from GWs that is in turn strongly correlated with the inclination of the binary system. Moreover, peculiar velocities of hosts may introduce systematics, especially for nearby events ([Howlett & Davis, 2020](#); [Mukherjee et al., 2021](#)). It has been estimated that $\mathcal{O}(50)$ of these events are needed to obtain percent level measurements on H_0 with current GW facilities (e.g., [Del Pozzo, 2012](#); [Chen et al., 2018](#); [Feeney et al., 2019](#)). However, EM counterparts are rare, as they typically require mergers involving at least one neutron star. Indeed, after three observing runs (O1, O2, O3), the catalog of confident GW detections released by the LIGO-Virgo-Kagra (LVK) collaboration includes about 90 events, the vast majority being binary black holes (BBHs), two binary neutron stars (BNSs), and two neutron star–black hole (NSBH) mergers [Abbott et al. \(2023b\)](#). Another main challenge in detecting the EM counterparts is to have precise localizations ($< 50 \text{ deg}^2$) and to organize extensive follow-up campaigns.

- **Dark sirens.** When counterparts are too faint or absent, the redshift can still be statistically inferred from the distribution of potential hosts within the volume of localization of the GW (Schutz, 1986; Del Pozzo, 2012; Fishbach et al., 2018; Soares-Santos et al., 2019; Gray et al., 2020; Palmese et al., 2020; Finke et al., 2021; Gray et al., 2022; Muttoni et al., 2023). Combining the redshift of each potential host with the d_L measured with GWs provides separate constraints to the cosmological parameters.² In this way, by stacking together the information from multiple events, the true parameters will statistically prevail. This method becomes more effective with a smaller GW localization volume, up to the limit when only a single galaxy is present, resembling the situation of bright sirens. A key challenge for this method is to obtain the most complete catalog of potential hosts within this volume and also properly account for catalog incompleteness in the analysis (Chen et al., 2018; Gray et al., 2020; Finke et al., 2021). The latest LVK analysis yields $H_0 = 67_{-12}^{+13}$ km/s/Mpc (excluding GW170817, Abbott et al., 2023b).
- **Spectral sirens.** Additionally, by modeling intrinsic astrophysical properties as the source-frame mass distribution, the degeneracy between mass and redshift can be broken, thus enabling cosmological analyses (Chernoff & Finn, 1993; Taylor et al., 2012; Messenger & Read, 2012; Fishbach et al., 2018; Farr et al., 2019; Mastrogiovanni et al., 2021b; Ezquiaga & Holz, 2022). To fully realize its potential, the source-frame mass distribution must contain features beyond a simple power law, such as breaks, peaks, or changes in slope (Ezquiaga & Holz, 2022). A caveat of this method is that it requires the assumption of a mass distribution function. Therefore, it is crucial to evaluate the potential systematics on cosmological inference (Abbott et al., 2023a; Mancarella et al., 2022a).

Multiple pipelines have been publicly released, for dark sirens (`gwcsmo`, Gray et al. 2020; `DarkSirensStat`, Gray et al. 2020; `cosmolisa`, Laghi & Del Pozzo 2020) and spectral sirens (`icarogw`, Mastrogiovanni et al. 2021a; `MGCosmoPop`, Mancarella et al. 2022b). The latest cosmological results presented in Abbott et al. (2023c) are obtained either by studying the population of binaries without catalog information with `icarogw`, or by using catalog information while fixing the astrophysical population parameters using `gwcsmo`.

At the time of the beginning of this work, no public codes were available to join the dark and spectral sirens methods. Only very recently, two updated pipelines were presented in Mastrogiovanni et al. (2023) and Gray et al. (2023), namely `icarogw2.0` and `gwcsmo2.0`, building on the homonyms aforementioned codes.

The current state-of-the-art statistical approach for standard sirens analyses is hierarchical Bayesian inference. In this framework, population-level parameters describing the source properties and cosmology are hyperparameters of a hierarchical model, which also includes selection effects and measurement uncertainties. The next section sets the groundwork for hierarchical Bayesian inference, which is later extended in this thesis work.

²A similar approach consists in the spatial cross-correlation of GWs with galaxy catalogs (Oguri, 2016; Mukherjee et al., 2020).

3.1.5 Hierarchical Bayesian Inference

Astrophysical analyses often involve studying parameters of a distribution of a population of sources using a set of observations sampled from it. Measurement uncertainties and selection biases are two effects that can simultaneously complicate this process. A Bayesian approach to address this problem has been presented by [Loredo \(2004\)](#) and more recently reworked by [Mandel et al. \(2019\)](#) and [Vitale et al. \(2022\)](#).

Consider a population of events, individually described by event-level parameters $\boldsymbol{\theta}$ that globally follow a distribution,

$$\frac{dN}{d\boldsymbol{\theta}}(\boldsymbol{\lambda}) = N p_{\text{pop}}(\boldsymbol{\theta}|\boldsymbol{\lambda}), \quad (3.37)$$

described by hyperparameters $\boldsymbol{\lambda}$. This distribution is sampled by observers who draw a set of N_{ev} events, each characterized by a likelihood function $p(\mathbf{d}_i, \boldsymbol{\theta}_i)$, for $i \in [1, N_{\text{ev}}]$, that relates the measured data set $\{\mathbf{d}_i\}$ to the event parameters $\{\boldsymbol{\theta}_i\}$. The key goal of population inference is to determine $\boldsymbol{\lambda}$, however, this is challenged by limited sampling, selection biases, and measurement uncertainties. One can adopt a Bayesian approach to determine the posterior probability of these parameters given the available observations via the Bayes theorem:

$$p(\boldsymbol{\lambda}|\{\mathbf{d}_i\}) = \frac{p(\{\mathbf{d}_i\}|\boldsymbol{\lambda}) \pi(\boldsymbol{\lambda})}{p(\{\mathbf{d}_i\})}, \quad (3.38)$$

where $p(\{\mathbf{d}_i|\boldsymbol{\lambda}\})$ is the likelihood of observing the data set given $\boldsymbol{\lambda}$, $\pi(\boldsymbol{\lambda})$ the prior on $\boldsymbol{\lambda}$, and the evidence $p(\{\mathbf{d}_i\})$ is the integral of the numerator over all possible $\boldsymbol{\lambda}$. The posterior probability $p(\boldsymbol{\lambda}|\{\mathbf{d}_i\})$ can be sampled with a variety of Monte Carlo methods. The most popular are the Markov chain Monte Carlo (MCMC) [Metropolis et al. \(1953\)](#), in which particles or “walkers” generate a sequence of random draws by moving through the probability space according to a proposal distribution. An alternative is nested sampling ([Skilling, 2004](#)), originally designed to calculate the evidence term. In this case, a sequence of *live points* is recursively updated by replacing the lowest likelihood points with new points at higher likelihood. Both of these methods require as input the likelihood function $p(\{\mathbf{d}_i|\boldsymbol{\lambda}\})$, which we derive following [Mandel et al. \(2019\)](#); however, see also [Loredo & Lamb \(2002\)](#); [Thrane & Talbot \(2019\)](#); [Vitale et al. \(2022\)](#).

First, in the absence of measurement uncertainties, the observations directly provide the event parameters $\{\boldsymbol{\theta}_i\}$, for $i \in [1, N_{\text{ev}}]$. The probability of obtaining them from a set of independent observations is given by

$$p(\{\boldsymbol{\theta}_i\}|\boldsymbol{\lambda}) = \prod_{i=1}^{N_{\text{ev}}} \frac{p_{\text{pop}}(\boldsymbol{\theta}_i|\boldsymbol{\lambda})}{\int p_{\text{pop}}(\boldsymbol{\theta}'|\boldsymbol{\lambda}) d\boldsymbol{\theta}'}, \quad (3.39)$$

where the normalization factor accounts for the probability of making an observation given the population parameters $\boldsymbol{\lambda}$. Each i -th event may also have a higher or lower probability of being observed depending on its parameters $\boldsymbol{\theta}$. For instance, higher-mass mergers are

relatively easier to detect than lower-mass ones. This effect is encoded in the detection probability term $P_{\text{det}}(\boldsymbol{\theta})$ and Eq. (3.39) is updated as follows,

$$p(\{\boldsymbol{\theta}_i\}|\boldsymbol{\lambda}) = \prod_{i=1}^{N_{\text{ev}}} \frac{p_{\text{pop}}(\boldsymbol{\theta}_i|\boldsymbol{\lambda}) P_{\text{det}}(\boldsymbol{\theta}_i)}{\int p_{\text{pop}}(\boldsymbol{\theta}'|\boldsymbol{\lambda}) P_{\text{det}}(\boldsymbol{\theta}') d\boldsymbol{\theta}'}. \quad (3.40)$$

where by definition, $P_{\text{det}}(\boldsymbol{\theta}_i) = 1$ since the i -th event has actually been detected.

In general, a measurement is noisy and it is not possible to directly determine $\boldsymbol{\theta}$. This also implies that the detection probability $P_{\text{det}}(\boldsymbol{\theta})$ is not a binary 0 or 1, but may vary depending on the noise realization. Under the assumption that the detectability is deterministic (e.g., SNR threshold), the detection probability can be defined as the integral over all the detectable data sets given a set of event parameters,

$$P_{\text{det}}(\boldsymbol{\theta}) = \int_{\mathbf{d} \in \text{det}} p(\mathbf{d}'|\boldsymbol{\theta}) d\mathbf{d}'. \quad (3.41)$$

Note that, by construction, this function includes effects from both intrinsic (e.g., masses of the binary system) and extrinsic (e.g., sky location, orientation) event-level parameters $\boldsymbol{\theta}$, as well as marginalization over instrumental noise fluctuations. Using Eq. (3.41) and the Bayes theorem, the probability of observing one event in the presence of both measurement uncertainty and selection effects becomes

$$p(\mathbf{d}_i|\boldsymbol{\lambda}) = \frac{\int p(\mathbf{d}_i|\boldsymbol{\theta}') p_{\text{pop}}(\boldsymbol{\theta}'|\boldsymbol{\lambda}_i) d\boldsymbol{\theta}'}{\int_{\mathbf{d} \in \text{det}} d\mathbf{d}' \int p(\mathbf{d}'|\boldsymbol{\theta}') p_{\text{pop}}(\boldsymbol{\theta}'|\boldsymbol{\lambda}) d\boldsymbol{\theta}'}. \quad (3.42)$$

The denominator represents the fraction of events in the Universe that are detectable given a population model characterized by hyperparameters $\boldsymbol{\lambda}$ and can be redefined with

$$\xi(\boldsymbol{\lambda}) \equiv \int_{\mathbf{d} \in \text{det}} d\mathbf{d}' \int p(\mathbf{d}'|\boldsymbol{\theta}') p_{\text{pop}}(\boldsymbol{\theta}'|\boldsymbol{\lambda}) d\boldsymbol{\theta}' = \int P_{\text{det}}(\boldsymbol{\theta}') p_{\text{pop}}(\boldsymbol{\theta}'|\boldsymbol{\lambda}) d\boldsymbol{\theta}'. \quad (3.43)$$

Thus, combining Eqs. (3.38), (3.40) and (3.42), the posterior on the population parameters $\boldsymbol{\lambda}$ becomes

$$p(\boldsymbol{\lambda}|\{\mathbf{d}_i\}) = \frac{\pi(\boldsymbol{\lambda})}{\xi(\boldsymbol{\lambda})^{N_{\text{ev}}}} \prod_{i=1}^{N_{\text{ev}}} \int p(\mathbf{d}_i|\boldsymbol{\theta}_i) p_{\text{pop}}(\boldsymbol{\theta}_i|\boldsymbol{\lambda}_i) d\boldsymbol{\theta}_i. \quad (3.44)$$

This derivation does not take into account the overall normalization N of Eq. (3.37). When the expected number N_{exp} can be predicted, the probability of observing N_{ev} events is given by the Poisson distribution $p(N_{\text{ev}}|N_{\text{exp}}) = e^{-N_{\text{exp}}} (N_{\text{exp}})^{N_{\text{ev}}}$. Then, multiplying p_{pop} by N in Eq. (3.44), the full posterior with the rate included is

$$p(\boldsymbol{\lambda}, N|\{\mathbf{d}_i\}) = e^{-N_{\text{exp}}} (N_{\text{exp}})^{N_{\text{ev}}} \frac{\pi(\boldsymbol{\lambda}) \pi(N)}{\xi(\boldsymbol{\lambda})^{N_{\text{ev}}}} \prod_{i=1}^{N_{\text{ev}}} \int p(\mathbf{d}_i|\boldsymbol{\theta}_i) p_{\text{pop}}(\boldsymbol{\theta}_i|\boldsymbol{\lambda}_i) d\boldsymbol{\theta}_i. \quad (3.45)$$

where

$$N_{\text{exp}}(\boldsymbol{\lambda}) \equiv \int_{\mathbf{d} \in \text{det}} d\mathbf{d}' d\boldsymbol{\theta}' p(\mathbf{d}'|\boldsymbol{\theta}') \frac{dN}{d\boldsymbol{\theta}'}(\boldsymbol{\lambda}) = N \xi(\boldsymbol{\lambda}) \quad (3.46)$$

With a prior $\pi(N) \propto 1/N$ on the intrinsic event rate N , Eq. (3.45) can be marginalized over N , recovering Eq. (3.44) modulo a normalization constant that depends only on N_{ev} and would not affect the inference of λ (Fishbach et al., 2018).

In this Thesis, we aim to extend the hierarchical Bayesian framework to incorporate dark and spectral sirens methods, thus enabling a joint cosmological and astrophysical parameters inference with galaxy catalogs. We present and release CHIMERA, a novel Python code that we develop for these analyses. In Section 3.2 we derive the new methodology and discuss the implementations in CHIMERA. In Section 3.4 we present the sample of the O4- and O5-like configurations, which are then analyzed in Section 3.5. This work aims to tackle the challenges of astrophysical and cosmological analysis for next-generation gravitational wave observatories such as ET (Punturo et al., 2010) and LISA (Amaro-Seoane et al., 2017), and upcoming wide-field galaxy surveys such as Euclid (Laureijs et al., 2011), the Nancy Grace Roman Space Telescope (Akeson et al., 2019), WFIRST (Akeson et al., 2019), and LSST (LSST Science Collaboration et al., 2009).

3.2 Extending the standard sirens method

We start by considering a population of GW sources, individually described by source-frame parameters $\boldsymbol{\theta}$ which globally follow a probability distribution $p_{\text{pop}}(\boldsymbol{\theta}|\boldsymbol{\lambda})$ described by hyperparameters $\boldsymbol{\lambda}$ (Eq. (3.44)). We elaborate on the hierarchical Bayesian inference formalism to infer $\boldsymbol{\lambda}$ given a catalog of GW detections and a catalog of their potential hosts. Given a set of $\boldsymbol{d}^{\text{GW}} = \{\boldsymbol{d}_i^{\text{GW}}\}$ ($i = 1, \dots, N_{\text{ev}}$) independent GW events, the population likelihood is (Mandel et al., 2019; Vitale et al., 2022),

$$p(\boldsymbol{d}^{\text{GW}}|\boldsymbol{\lambda}) \propto \frac{1}{\xi(\boldsymbol{\lambda})^{N_{\text{ev}}}} \prod_{i=1}^{N_{\text{ev}}} \int p(\boldsymbol{d}_i^{\text{GW}}|\boldsymbol{\theta}_i) p_{\text{pop}}(\boldsymbol{\theta}_i|\boldsymbol{\lambda}) d\boldsymbol{\theta}_i, \quad (3.47)$$

where $p(\boldsymbol{d}_i^{\text{GW}}|\boldsymbol{\theta}_i)$ is the individual source likelihood and

$$\begin{aligned} \xi(\boldsymbol{\lambda}) &\equiv \iint_{\boldsymbol{d}^{\text{GW}} \in \text{det}} p(\boldsymbol{d}^{\text{GW}}|\boldsymbol{\theta}) p_{\text{pop}}(\boldsymbol{\theta}|\boldsymbol{\lambda}) d\boldsymbol{d}^{\text{GW}} d\boldsymbol{\theta} \\ &\equiv \int P_{\text{det}}(\boldsymbol{\theta}, \boldsymbol{\lambda}_c) p_{\text{pop}}(\boldsymbol{\theta}|\boldsymbol{\lambda}) d\boldsymbol{\theta}, \end{aligned} \quad (3.48)$$

is the selection function, which measures the overall fraction of detectable events given $\boldsymbol{\lambda}$ and a particular GW network configuration. We separate the hyperparameters describing the underlying cosmology $\boldsymbol{\lambda}_c$ from those describing the astrophysical population of GW sources, i.e. mass distribution $\boldsymbol{\lambda}_m$ and redshift distribution $\boldsymbol{\lambda}_z$, so that $\boldsymbol{\lambda} = \{\boldsymbol{\lambda}_c, \boldsymbol{\lambda}_m, \boldsymbol{\lambda}_z\}$. Under the assumption that the mass function does not evolve with cosmic time (justified by current data, e.g. see Abbott et al. 2023a), the population function can be split as follows:

$$p_{\text{pop}}(\boldsymbol{\theta}|\boldsymbol{\lambda}) = p(m_1, m_2|\boldsymbol{\lambda}_m) p(z, \hat{\Omega}|\boldsymbol{\lambda}_c, \boldsymbol{\lambda}_z). \quad (3.49)$$

We consider the following set of source parameters $\boldsymbol{\theta} = \{z, \hat{\Omega}, m_1, m_2\}$, where z is the redshift of the binary, $\hat{\Omega}$ the sky localization (measured in square degrees), and m_1, m_2 are the primary and secondary source-frame masses. However, current GW observations do not provide information on source-frame parameters, but on detector-frame quantities $\boldsymbol{\theta}^{\text{det}} = \{d_L, \hat{\Omega}, m_1^{\text{det}}, m_2^{\text{det}}\}$, which are related to the former by the underlying cosmology:

$$z = z(d_L|\boldsymbol{\lambda}_c) \quad (3.50)$$

$$m_{1,2} = m_{1,2}^{\text{det}}/[1 + z(d_L|\boldsymbol{\lambda}_c)] \quad (3.51)$$

Here, we have highlighted the fact that the redshift of a particular source, given a measured luminosity distance d_L , is dependent on the assumed cosmology through the parameters $\boldsymbol{\lambda}_c$. Note also that the likelihood is a normalized PDF on the data, not on the parameters, so in the transformation from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}^{\text{det}}$ it remains unchanged and we can simply write $p(\boldsymbol{d}_i^{\text{GW}}|\boldsymbol{\theta}_i) d\boldsymbol{\theta}_i = p(\boldsymbol{d}_i^{\text{GW}}|\boldsymbol{\theta}_i^{\text{det}}) d\boldsymbol{\theta}_i^{\text{det}}$. We do not have direct access to the likelihood, as the LVK data products consist of a set of samples drawn from the posterior distribution $p(\boldsymbol{\theta}_i^{\text{det}}|\boldsymbol{d}_i^{\text{GW}})$ obtained with priors $\pi(\boldsymbol{\theta}_i)$. In particular, the LVK posterior chains are obtained assuming flat priors in detector-frame masses and a prior on luminosity distance

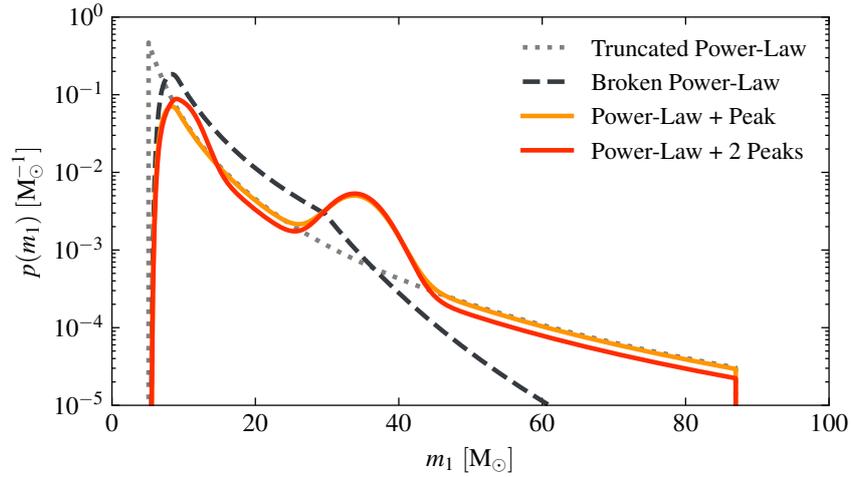


FIGURE 3.6: Examples of typical BBH mass function models for the primary mass using parameters from [Mancarella et al. \(2022a\)](#).

$\pi(d_L) \propto d_L^2$. Eventually, Eq. (3.47) becomes

$$p(\mathbf{d}^{\text{GW}}|\boldsymbol{\lambda}) \propto \frac{1}{\xi(\boldsymbol{\lambda})^{N_{\text{ev}}}} \prod_{i=1}^{N_{\text{ev}}} \int \frac{p(\boldsymbol{\theta}_i^{\text{det}}|\mathbf{d}_i^{\text{GW}})}{\pi(\boldsymbol{\theta}^{\text{det}})} \left| \frac{d\boldsymbol{\theta}_i^{\text{det}}}{d\boldsymbol{\theta}_i} \right| p_{\text{pop}}(\boldsymbol{\theta}_i|\boldsymbol{\lambda}) d\boldsymbol{\theta}_i. \quad (3.52)$$

3.2.1 Population prior

Let's consider the mass term of Eq. (3.49). It can be factorized as

$$p(m_1, m_2|\boldsymbol{\lambda}_m) = p(m_1|\boldsymbol{\lambda}_m) p(m_2|m_1, \boldsymbol{\lambda}_m). \quad (3.53)$$

were $p(m_1|\boldsymbol{\lambda}_m)$ is the primary mass distribution and $p(m_2|m_1, \boldsymbol{\lambda}_m)$ the secondary mass distribution conditioned on the former. Here we assume the convention that $m_1 > m_2$. For BBHs, the primary BH mass distribution can be described by different phenomenological models (see the definitions in [Abbott et al., 2023a](#)), which are listed below in increasing order of features.

- (i) **Truncated Power-Law:** a power-law with index $-\alpha$ and sharp lower and upper mass cutoffs (m_{low} and m_{high}). The primary mass distribution for this model is:

$$p(m_1|\alpha, m_{\text{low}}, m_{\text{high}}) \equiv \mathcal{P}(m_1) \propto \begin{cases} m_1^{-\alpha} & m_{\text{low}} < m_1 < m_{\text{high}} \\ 0 & \text{otherwise} \end{cases}, \quad (3.54)$$

while the mass ratio $q \equiv m_2/m_1$ follows a power-law with spectral index β ,

$$p(q|\beta, m_{\text{low}}, m_1) \propto \begin{cases} q^\beta & m_{\text{low}} < m_2 < m_1 \\ 0 & \text{otherwise} \end{cases}, \quad (3.55)$$

- (ii) **Broken Power-Law:** a smoothed power-law with two different slopes and break mass. The second slope is motivated by the potential tapering of the primary mass distribution at high masses, while the smoothing prevents a sharp cutoff at low masses. The primary mass distribution for this model is:

$$p(m_1|\alpha_1, \alpha_2, b, \delta_m, m_{\text{low}}, m_{\text{high}}) \propto \begin{cases} m_1^{-\alpha_1} \mathcal{S}(m_1|m_{\text{low}}, \delta_m) & m_{\text{low}} < m_1 < m_{\text{break}} \\ m_1^{-\alpha_2} \mathcal{S}(m_1|m_{\text{low}}, \delta_m) & m_{\text{break}} < m_1 < m_{\text{high}} \\ 0 & \text{otherwise} \end{cases} \quad (3.56)$$

where $m_{\text{break}} \equiv m_{\text{low}} + b(m_{\text{high}} - m_{\text{low}})$ is the break mass which is positioned at a fraction b between the minimum and maximum mass and $\mathcal{S}(m) \in [0, 1]$ is a smoothing piece-wise function with a tapering parameter δ_m (described in detail in [Abbott et al. 2021](#), appendix B). The smoothing term is also included in the conditional mass ratio distribution,

$$p(q|\beta, m_{\text{low}}, m_1) \propto q^\beta \mathcal{S}(q m_1|m_{\text{low}}, \delta_m). \quad (3.57)$$

- (iii) **Power-law + Peak:** a smoothed power-law with the addition of one Gaussian peak $\mathcal{G}(m) \propto \mathcal{N}(\mu_g; \sigma_g^2)$ to model position and width of the observed overdensity of masses at $\approx 34 M_\odot$. This model is motivated by the idea that the mass loss undergone by pulsational pair-instability supernovae may result in a pileup of BBH events before the pair-instability gap ([Talbot & Thrane, 2018](#)). The primary mass distribution for this model is:

$$p(m_1|\alpha, \delta_m, m_{\text{low}}, m_{\text{high}}, \mu_g, \sigma_g, \lambda_g) \propto [(1 - \lambda_g) \mathcal{P}(m_1) + \lambda_g \mathcal{G}(m_1)] \mathcal{S}(m_1), \quad (3.58)$$

where λ_g regulates the relative contribution of \mathcal{G} and \mathcal{P} . The conditional mass ratio distribution follows Eq. (3.57).

- (iv) **Power-law + 2 Peaks:** as (iii) with the addition of a second Gaussian peak to capture the apparent overdensity at about $10 M_\odot$ ([Abbott et al., 2023a](#)). The primary mass distribution for this model is:

$$p(m_1|\alpha, \delta_m, m_{\text{low}}, m_{\text{high}}, \mu_{g,1}, \sigma_{g,1}, \mu_{g,2}, \sigma_{g,2}, \lambda_g, \lambda_{g1}) \propto \propto [(1 - \lambda_g) \mathcal{P}(m_1) + \lambda_g \lambda_{g1} \mathcal{G}_1(m_1) + \lambda_g (1 - \lambda_{g1}) \mathcal{G}_2(m_1)] \mathcal{S}(m_1), \quad (3.59)$$

where the parameters λ_g and λ_{g1} model the fraction of binaries in any Gaussian component and in the lower-mass Gaussian components, respectively. The statistical significance of \mathcal{G}_2 in current data is still under discussion (e.g., [Edelman et al., 2022](#)).

An example of these mass functions is given in Fig. 3.6.

Now we take a closer look at the second term of Eq. (3.49), that we factorize as follows

$$p(z, \hat{\Omega}|\boldsymbol{\lambda}_c, \boldsymbol{\lambda}_z) = \frac{p_{\text{gal}}(z, \hat{\Omega}|\boldsymbol{\lambda}_c) p_{\text{rate}}(z|\boldsymbol{\lambda}_z)}{\int p_{\text{gal}}(z, \hat{\Omega}|\boldsymbol{\lambda}_c) p_{\text{rate}}(z|\boldsymbol{\lambda}_z) dz d\hat{\Omega}} \quad (3.60)$$

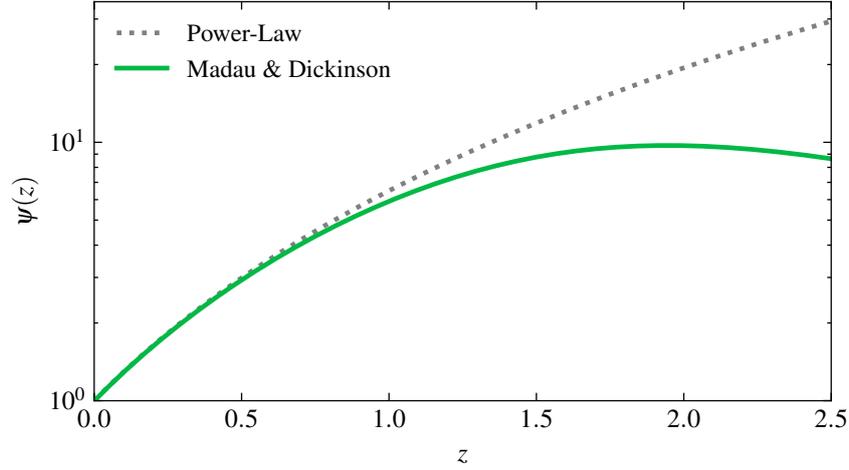


FIGURE 3.7: Examples of rate evolution functions using parameters from [Mancarella et al. \(2022a\)](#). The functions are renormalized so that $\psi(z = 0) = 1$.

where p_{gal} is the probability that there is a galaxy at $(z, \hat{\Omega})$ and p_{rate} the probability of a galaxy at redshift z to host a GW event. This takes into account that the probability for a galaxy to host a merger has a non-trivial redshift dependence, that can be parametrized as

$$p_{\text{rate}}(z|\boldsymbol{\lambda}_z) \propto \frac{\psi(z; \boldsymbol{\lambda}_z)}{(1+z)} \quad (3.61)$$

where $\psi(z; \boldsymbol{\lambda}_z)$ is the source-frame merger rate and the term $(1+z)^{-1}$ takes into account the conversion between source and detector time. The typically adopted models for $\psi(z)$ are outlined below.

- (i) **Power-law:** a simple power-law with index γ ,

$$\psi(z) \propto (1+z)^\gamma. \quad (3.62)$$

- (ii) **Madau-Dickinson:** rate evolution from [Madau & Dickinson \(2014\)](#) that can be parametrized as (see [Callister et al., 2020](#)),

$$\psi(z; \boldsymbol{\lambda}_z) = \frac{(1+z)^\gamma}{1 + \left(\frac{1+z}{1+z_p}\right)^{\gamma+\kappa}}, \quad (3.63)$$

where $\psi(z) \propto (1+z)^\gamma$ at low z , then reaches its peak near z_p , and subsequently declines as $\psi(z) \propto (1+z)^{-\kappa}$.

An example of these rate evolution functions is given in Fig. 3.7. The model (ii) assumes that the GW event rate follows the cosmic star formation history. Sometimes a delay-time between the two is adopted in the analyses (e.g., [Vitale et al., 2019](#)). With current GW events, it is not possible to place stringent constraints on the rate evolution above redshift $z = 1$, therefore the power-law remains a good approximation ([Abbott et al., 2023a](#)).

3.2.2 Ideal galaxy catalog

As a starting point, we assume that the galaxy catalog is *complete*, i.e. contains all the potential host galaxies. In this case, $p_{\text{gal}}(z, \hat{\Omega} | \boldsymbol{\lambda}_c) = p_{\text{cat}}(z, \hat{\Omega} | \boldsymbol{\lambda}_c)$, which is computed as a sum over the contribution of each galaxy. Given a set of $\mathbf{d}^{\text{EM}} = \{\mathbf{d}_g^{\text{EM}}\}$ ($g = 1, \dots, N_{\text{gal}}$) EM observations of galaxies (i.e., photometric or spectroscopic redshifts) we have

$$p_{\text{cat}}(z, \hat{\Omega} | \boldsymbol{\lambda}_c) = \frac{\sum_g w_g p(z | \mathbf{d}_g^{\text{EM}}, \boldsymbol{\lambda}_c) \delta(\hat{\Omega} - \hat{\Omega}_g)}{\sum_g w_g}, \quad (3.64)$$

where w_g weights the probability of each galaxy to host a GW event (e.g., by the galaxy luminosity), $p(z | \mathbf{d}_g^{\text{EM}}, \boldsymbol{\lambda}_c)$ is the galaxy's redshift posterior distribution that we want to use as a prior, and $\delta(\hat{\Omega} - \hat{\Omega}_g)$ is a Dirac delta distribution of each galaxy's sky localization that can be treated as errorless.

The galaxy catalog contains redshift measurements \tilde{z}_g and associated uncertainties $\tilde{\sigma}_{z,g}$ for each observed galaxy, i.e. $\mathbf{d}_g^{\text{EM}} = \{\tilde{z}_g, \tilde{\sigma}_{z,g}\}$. From these quantities, we construct the likelihoods, which we assume to be Gaussian, $p(\tilde{z}_g | z) = \mathcal{N}(z; \tilde{z}_g, \tilde{\sigma}_{z,g}^2)$. This is a probability distribution over the observed values \tilde{z}_g . We note that in the case that the full posterior distribution of each redshift source in the catalog is provided, this can substitute the Gaussian approximants. To get $p(z | \mathbf{d}_g^{\text{EM}}, \boldsymbol{\lambda}_c)$ we need to multiply it by a prior on the redshift distribution, which in the absence of other information is naturally chosen as uniform in comoving volume (Gair et al., 2023). Using Bayes' theorem, we get

$$p(z | \mathbf{d}_g^{\text{EM}}, \boldsymbol{\lambda}_c) = \frac{\mathcal{N}(z; \tilde{z}_g, \tilde{\sigma}_{z,g}^2) \frac{dV_c}{dz}(z; \boldsymbol{\lambda}_c)}{\int \mathcal{N}(z; \tilde{z}_g, \tilde{\sigma}_{z,g}^2) \frac{dV_c}{dz}(z; \boldsymbol{\lambda}_c) dz} \quad (3.65)$$

where dV_c/dz is the differential comoving volume element in a flat universe. With this definition, Eq. (3.64) is normalized so that if $p(\tilde{z}_g | z) = \delta(z - \tilde{z}_g)$, and in case of uniform weights, we get $n_{\text{cat}}(z) = (1/V_c) \sum_g \delta(z - \tilde{z}_g)$ which is consistent with Eq. 3.39 of Finke et al. (2021), i.e. the comoving density of galaxies $n_{\text{cat}}(z)$ is estimated by counting the objects in the catalog and dividing by the total volume. If, instead, the likelihood is completely uninformative, $p(\tilde{z}_g | z) = 1$, we get a comoving density constant in redshift, which also is consistent.

3.2.3 Realistic galaxy catalog

Real galaxy catalogs suffer from completeness issues, meaning that there is a probability different from zero that true GW hosts are missing from catalogs due to observational effects. In general, this effect is z -dependent: at higher z , fewer galaxies are observed due to the Malmquist bias, hence the chance that the true GW host is missed is higher. For this reason, if not taken into account, this effect would bias the final constraints.

Definition of completeness

In this context, the *completeness* inside a given region \mathcal{S} is defined as the ratio of observed-to-total galaxies with physical properties \mathcal{G} ,

$$P_{\text{compl}}(\mathcal{S}, \mathcal{G}) \equiv \frac{N_{\text{cat}}(\mathcal{S}, \mathcal{G})}{N_{\text{gal}}(\mathcal{S}, \mathcal{G})}, \quad (3.66)$$

where N_{cat} is the number of galaxies in the catalog, while the total number of galaxies N_{gal} needs to be modeled. The region \mathcal{S} should be sufficiently large to include a representative number of galaxies and sufficiently small to resolve the cosmic structures. To compute the reference background distribution of N_{gal} one can assume that they globally follow a distribution with constant number density per comoving volume, therefore,

$$P_{\text{compl}}(\mathcal{S}, \mathcal{G}) = \frac{N_{\text{cat}}(\mathcal{S}, \mathcal{G})}{\bar{n}_{\text{gal}}(\mathcal{G}) V_c(\mathcal{S})}, \quad (3.67)$$

where a typical value is $\bar{n}_{\text{gal}} \sim 0.1 \text{ Mpc}^{-3}$ (Conselice et al., 2016). Note that the notion of completeness is quasi-local and already requires a cosmological model for $V_c(\mathcal{S})$, but as discussed in Finke et al. (2021), we expect the cosmology dependence of the completion model to be weak.

The region \mathcal{S} can be defined by a voxel (3D pixel) identified with redshift z and sky position $\hat{\Omega}$. Inside a region \mathcal{S} the number of catalog n_{cat} and missing n_{miss} galaxies must satisfy the following relation:

$$\frac{1}{V_c(\mathcal{S})} \int_{\mathcal{S}} [n_{\text{cat}}(z, \hat{\Omega}) + n_{\text{miss}}(z, \hat{\Omega})] dV_c(z) = \bar{n}_{\text{gal}}. \quad (3.68)$$

To estimate n_{miss} , hence “complete” the catalog, one can adopt two natural approaches discussed in detail in (Finke et al., 2021). The *homogeneous completion* consists in assuming that n_{miss} is uniform within \mathcal{S} , hence from Eq. (3.68), $n_{\text{miss}}(z, \hat{\Omega}, \mathcal{G}) = \bar{n}_{\text{gal}}(\mathcal{G}) - N_{\text{cat}}(\mathcal{S}, \mathcal{G})/V_c(\mathcal{S})$, and in terms of the completeness we have

$$n_{\text{miss}}^{\text{HOM}}(z, \hat{\Omega}, \mathcal{G}) = \bar{n}_{\text{gal}}(\mathcal{G}) - [1 - P_{\text{compl}}(\mathcal{S}, \mathcal{G})]. \quad (3.69)$$

It can be shown that this approach correctly yields an average density \bar{n}_{gal} with a variance corresponding to the one present in the catalog. Alternatively, the *multiplicative completion* consists in assuming that galaxies are more likely to be missed near high-density regions, therefore

$$n_{\text{miss}}^{\text{MULT}}(z, \hat{\Omega}, \mathcal{G}) = b(\mathcal{S}, \mathcal{G}) n_{\text{cat}}(z, \hat{\Omega}), \quad (3.70)$$

where the multiplicative $b(\mathcal{S}, \mathcal{G})$ factor is related to the completeness as follows (see Eq. (3.68)),

$$b(\mathcal{S}, \mathcal{G}) = \frac{1 - P_{\text{compl}}(\mathcal{S}, \mathcal{G})}{P_{\text{compl}}(\mathcal{S}, \mathcal{G})}. \quad (3.71)$$

This second approach produces a variance that is higher by a factor $1/P_{\text{compl}}^2$. When the completeness is 1 we fall back to the homogeneous completion case, but when it is lower, the

clustering of missing galaxies is higher with respect to homogeneous completion, up to the limit of empty catalog, where this definition diverges. A more physically motivated approach may reside in between these two, distributing the missing galaxies according to a correlation length. In this work, as done in most analyses presented so far, we adopt the homogeneous completion.

In Eq. (3.67) we assume that the number density of galaxies is constant in comoving volume. In practice, the evolution of the number density of galaxies, is by itself a fundamental question for galaxy evolution. A detailed study has been presented in [Conselice et al. \(2016\)](#), where the authors compare and combine observed mass functions from a wide range of previous studies considering surveys up to redshift $z \sim 8$, providing important constraints on theoretical models of galaxy formation across cosmic time. Studying \bar{n}_{gal} for different galaxy masses, they find:

$$\begin{aligned}\bar{n}_{\text{gal};[6,12]} &\approx 0.1 \text{ Mpc}^{-3} \\ \bar{n}_{\text{gal};[9,12]} &\approx 0.01 \text{ Mpc}^{-3} \\ \bar{n}_{\text{gal};[10,12]} &\approx 0.005 \text{ Mpc}^{-3}\end{aligned}\tag{3.72}$$

where the subscripts refer to the mass interval inside which the number density is computed with the notation $[\log_{10} m_{\text{low}}/M_{\odot}, \log_{10} m_{\text{high}}/M_{\odot}]$. These approximations hold up to $z \sim 1$, where a potential z evolution is within the scatter of the different studies analyzed. Therefore, these approximations are reasonable for the current GW horizon and SNRs, while will require a more in-depth study for third-generation detectors.

The completeness correction

The correction of the completeness presented in Eqs. (3.68) to (3.70) relies on the definition of a region $\mathcal{R}_{\mathcal{H}}$ sufficiently large to include all the detectable GW events given the network of detectors considered.³ In this context, the p_{gal} term in Eq. (3.60) is split into the sum over the catalog galaxies p_{cat} and those that have been missed p_{miss} as follows:

$$p_{\text{gal}}(z, \hat{\Omega} | \boldsymbol{\lambda}_{\text{c}}) = f_{\mathcal{R}_{\mathcal{H}}} p_{\text{cat}}(z, \hat{\Omega} | \boldsymbol{\lambda}_{\text{c}}) + (1 - f_{\mathcal{R}_{\mathcal{H}}}) p_{\text{miss}}(z, \hat{\Omega} | \boldsymbol{\lambda}_{\text{c}}),\tag{3.73}$$

with

$$p_{\text{miss}}^{\text{HOM}}(z, \hat{\Omega} | \boldsymbol{\lambda}_{\text{c}}) = \frac{1 - P_{\text{comp}}(z, \hat{\Omega})}{(1 - f_{\mathcal{R}_{\mathcal{H}}}) V_{\text{c}}(\boldsymbol{\lambda}_{\text{c}})} \frac{dV_{\text{c}}}{dz}(z; \boldsymbol{\lambda}_{\text{c}}),\tag{3.74}$$

$$p_{\text{miss}}^{\text{MULT}}(z, \hat{\Omega} | \boldsymbol{\lambda}_{\text{c}}) = \frac{f_{\mathcal{R}_{\mathcal{H}}}}{1 - f_{\mathcal{R}_{\mathcal{H}}}} \frac{1 - P_{\text{comp}}(z, \hat{\Omega})}{P_{\text{comp}}(z, \hat{\Omega})} p_{\text{cat}}(z, \hat{\Omega} | \boldsymbol{\lambda}_{\text{c}}),\tag{3.75}$$

and

$$f_{\mathcal{R}_{\mathcal{H}}} \equiv \frac{1}{V_{\text{c}}(\boldsymbol{\lambda}_{\text{c}})} \int P_{\text{comp}}(z, \hat{\Omega}) dV_{\text{c}},\tag{3.76}$$

which weights the probability P_{comp} of observing a galaxy by encompassing all the volume $\mathcal{R}_{\mathcal{H}}$ where we can observe GW events (see [Finke et al., 2021](#)). When the weights w_g of

³Beyond the GW horizon, the selection effect term would remove any contribution from incompleteness.

Eq. (3.64) are set to 1, we estimate P_{compl} as in Eq. (3.67). Alternatively, in the case of *luminosity weighting*, the average luminosity of galaxies observed in each \mathcal{S} region is compared to an expected luminosity density. This is obtained assuming a Schechter function

$$\Phi(L) dL = \Phi^* \left(\frac{L}{L^*} \right)^\alpha e^{-L/L^*} d \left(\frac{L}{L^*} \right), \quad (3.77)$$

where Φ^* , α , and L^* are the Schechter parameters in a given wavelength band. When a lower luminosity $L > L_{\text{cut}}$ cut is applied, the average luminosity density can be easily computed as

$$\bar{l}_{\text{cut}} = \Phi^* L^* \Gamma(\alpha + 2, L_{\text{cut}}/L^*). \quad (3.78)$$

where Γ is the incomplete Gamma function. The optimal choice of L_{cut} is a compromise between higher statistics or higher completeness. This value also depends on the adopted parameters of the Schechter function and therefore on galaxy types, masses, and observed wavelength. Varying L_{cut} in current standard sirens analyses produces negligible effects concerning the statistical errors (e.g., [Fishbach et al., 2019](#)), however this may not hold true for next-generation GW observatories and galaxy surveys.

3.2.4 Selection effects

In this framework, it is crucial to accurately compute the selection effect (or selection bias) term in Eq. (3.48). For instance, in the case of measuring H_0 , at fixed (measured) d_L , increasing H_0 has the effect of increasing the inferred z (remember that $H_0 = cz/d_L + \mathcal{O}(z^2)$ at low z). On the other side, assuming that galaxies are uniformly distributed in comoving volume, the redshift prior increases more steeply with z , i.e. $p(z) \sim dV_c/dz \sim z^2$ at low z . As a result, the ‘‘convolution’’ between the GW kernel \mathcal{K}_{GW} and the prior p_z necessarily steepen at increasing H_0 . The selection bias term ξ actually compensates for this effect. A quantitative estimate can be easily obtained for a one-dimensional analysis with a single parameter. At higher dimensions, an analytical description is not convenient and one usually resorts to Monte Carlo integrals. We start by describing the first case and later discuss the Monte Carlo approach which is currently adopted in our pipeline.

Computation of $\xi(H_0)$

A simple detection model has been firstly presented in [Chen et al. \(2018\)](#) and applied to a simple population of sources (e.g., BNS with masses $1.4 + 1.4 M_\odot$). Considering a complete and homogeneous galaxy catalog, then Eqs. (3.49), (3.60) and (3.64) give

$$p_{\text{pop}}(z|H_0) = \frac{1}{V_c(\mathcal{R}_{\mathcal{H}})} \frac{dV_c}{dz}(z; H_0). \quad (3.79)$$

We then assume a Gaussian likelihood model with a luminosity-distance error σ_{d_L} described by $\sigma_{d_L} = A d_L(z, H_0)$, with A a constant fractional error ([Gair et al., 2023](#)). We deem the GW event as detected if their measured luminosity distance \hat{d}_L is smaller than a threshold

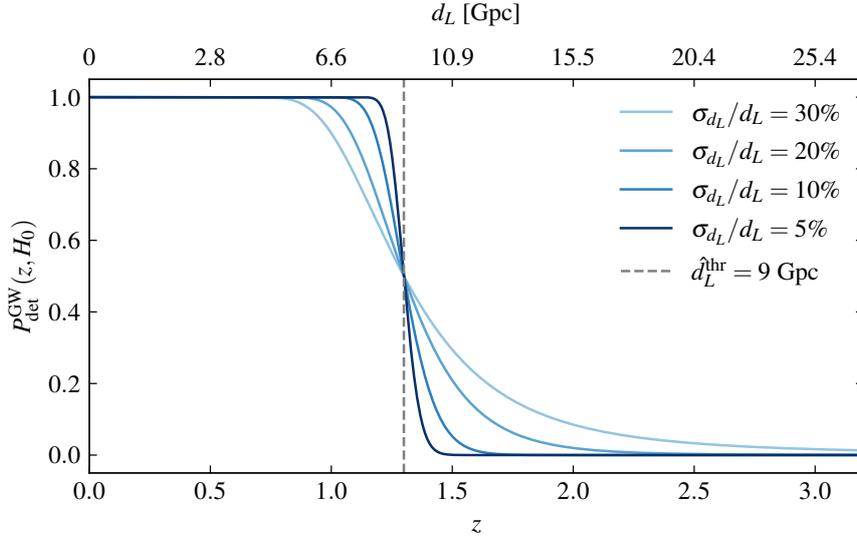


FIGURE 3.8: Simple detection probability model as a function of redshift from Eq. (3.80). The vertical dashed line shows the detection threshold $\hat{d}_L^{\text{thr}} = 9$ Gpc, equivalent to $z \simeq 1.3$ in a flat Λ CDM cosmology with $H_0 = 70$ km/s/Mpc.

\hat{d}_L^{thr} . Then, from Eq. (3.41) we obtain

$$\begin{aligned}
 P_{\text{det}}^{\text{GW}}(z, H_0) &= \int_{-\infty}^{+\infty} \Theta(\hat{d}_L; \hat{d}_L^{\text{thr}}) p(\hat{d}_L | d_L(z, H_0)) d\hat{d}_L \\
 &= \int_{-\infty}^{\hat{d}_L^{\text{thr}}} \mathcal{G}(\hat{d}_L; A d_L) d\hat{d}_L \\
 &= \frac{1}{2} \left[1 + \text{erf} \left(\frac{\hat{d}_L^{\text{thr}} - d_L(z, H_0)}{\sqrt{2} A d_L(z, H_0)} \right) \right],
 \end{aligned} \tag{3.80}$$

where $\Theta(x)$ is the Heaviside theta function, $\mathcal{G}(\mu, \sigma)$ a Gaussian with mean μ and standard deviation σ , and $\text{erf}(x)$ the error function. Figure 3.8 shows the resulting detection probability by varying the factor A assuming a detection threshold for the measured luminosity distance of $\hat{d}_L^{\text{thr}} = 9$ Gpc and $H_0 = 70$ km/s/Mpc. This toy model has been discussed in (Gair et al., 2023) to stress the idea that $P_{\text{det}}^{\text{GW}}(z, H_0)$ cannot be approximated by a step function unless the GW luminosity distance is known with high precision: the estimation of the Hubble constant would be biased to lower values.

In the limit when $\sigma_{d_L} \rightarrow 0$, the bias for H_0 can be simply computed from Eq. (3.48) as follows,

$$\begin{aligned}
\xi(H_0) &= \int P_{\text{det}}(z, H_0) p_{\text{pop}}(z) dz \\
&= \int \Theta\left(z^{\text{thr}}(\hat{d}_L^{\text{thr}}; H_0) - z\right) \frac{1}{V_c(\mathcal{R}_{\mathcal{H}})} \frac{dV_c}{dz} dz \\
&= \frac{\int_0^{z^{\text{thr}}(\hat{d}_L^{\text{thr}}; H_0)} \frac{dV_c}{dz} dz}{\int_0^{z_{\mathcal{R}_{\mathcal{H}}}} \frac{dV_c}{dz} dz}.
\end{aligned} \tag{3.81}$$

As $\mathcal{R}_{\mathcal{H}}$ must be taken to encompass all possible galaxies from which GW events are generated (see the discussion for Eq. (3.73)), we have that $z_{\mathcal{R}_{\mathcal{H}}} > z^{\text{thr}}(\hat{d}_L^{\text{thr}}; H_0)$. Both the numerator and the denominator depend on H_0^{-3} through the comoving volume element. However, the numerator has an additional dependency through the upper limit of the integral, $z^{\text{thr}}(\hat{d}_L^{\text{thr}}; H_0)$. Recalling that $dV_c/dz \sim z^2$ and $z \sim H_0 d_L$ at low z , it follows that:

$$\xi(H_0) \sim H_0^3 \quad \text{at low } z. \tag{3.82}$$

It is important to stress that this is an ideal scenario. This approximation is only valid for zero uncertainties and at low z and does not account for potential correlation among other population parameters $\boldsymbol{\lambda}$. Realistic analyses must take into account the potential correlation between all the parameters as we will discuss in the next section.

Computation of $\xi(\boldsymbol{\lambda})$

In our pipeline, the selection bias term $\xi(\boldsymbol{\lambda})$ in Eq. (3.48) is computed by using the injections approach. We simulate posteriors of the hyperparameters $\boldsymbol{\lambda}$ from a reference population using `gwfast`⁴ (Iacovelli et al., 2022a). We deem an event as detected when its SNR is higher than a given threshold, then compute the Monte Carlo integral over the detected events as follows:

$$\xi(\boldsymbol{\lambda}) = \frac{1}{N_{\text{inj}}} \sum_{i=1}^{N_{\text{det}}} \frac{p_{\text{pop}}(\boldsymbol{\theta}_i | \boldsymbol{\lambda})}{p_{\text{draw}}(\boldsymbol{\theta}_i)}, \tag{3.83}$$

where N_{inj} is the total number of injections and N_{det} the number of detected ones. We note here that, once computed, the selection effect term is a function only of the population hyperparameters $\boldsymbol{\lambda}_i$, and therefore can be evaluated only once for an entire catalog of events at given set of $\boldsymbol{\lambda}$ (see Eq. (3.47)).

To ensure the numerical stability when computing Eq. (3.83), it is possible to adopt a threshold on the so-called the *effective number of samples* (Farr, 2019). This value quantifies how many samples per event are contributing to the Monte Carlo integral. It is computed as follows,

$$N_{\text{eff},i} = \frac{\left(\sum_{i=1}^{N_{s,i}} w_{i,j}\right)^2}{\sum_{i=1}^{N_{s,i}} w_{i,j}^2}, \tag{3.84}$$

⁴<https://github.com/CosmoStatGW/gwfast/>

$w_{i,j}$ the weight associated to the j -th posterior sample of the i -th event and $N_{s,i}$ the total number of samples of the event. The default threshold to ensure numerical stability in CHIMERA is $N_{\text{eff}} > 5 N_{\text{det}}$, but this number can be adjusted by the user. When this threshold is not met, the code print a warning and returns a negative-infinite log-likelihood.

3.2.5 Full form of the likelihood

By putting together Eqs. 3.47–3.76 we obtain the following set of equations:

$$\begin{aligned}
 p(\mathbf{d}^{\text{GW}}|\boldsymbol{\lambda}) &\propto \frac{1}{\xi(\boldsymbol{\lambda})^{N_{\text{ev}}}} \prod_{i=1}^{N_{\text{ev}}} \iint \mathcal{K}_{\text{gw},i}(z, \hat{\Omega}|\boldsymbol{\lambda}_{\text{c}}, \boldsymbol{\lambda}_{\text{m}}) \frac{p_{\text{gal}}(z, \hat{\Omega}|\boldsymbol{\lambda}_{\text{c}}) p_{\text{rate}}(z|\boldsymbol{\lambda}_{\text{z}})}{\int p_{\text{gal}}(z, \hat{\Omega}|\boldsymbol{\lambda}_{\text{c}}) p_{\text{rate}}(z|\boldsymbol{\lambda}_{\text{z}}) dz d\hat{\Omega}} dz d\hat{\Omega}, \\
 \mathcal{K}_{\text{gw},i}(z, \hat{\Omega}|\boldsymbol{\lambda}_{\text{c}}, \boldsymbol{\lambda}_{\text{m}}) &\equiv \iint \frac{p(\boldsymbol{\theta}_i|\mathbf{d}_i^{\text{GW}}, \boldsymbol{\lambda}_{\text{c}})}{\pi(\boldsymbol{\theta}^{\text{det}})} \left| \frac{d\boldsymbol{\theta}_i}{d\boldsymbol{\theta}_i^{\text{det}}} \right| p(m_1, m_2|\boldsymbol{\lambda}_{\text{m}}) dm_1 dm_2, \\
 p_{\text{gal}}(z, \hat{\Omega}|\boldsymbol{\lambda}_{\text{c}}) &= \frac{\int P_{\text{comp}}(z, \hat{\Omega}) dV_{\text{c}}}{V_{\text{c}}(\boldsymbol{\lambda}_{\text{c}})} p_{\text{cat}}(\hat{\Omega}, z|\boldsymbol{\lambda}_{\text{c}}) + \frac{[1 - P_{\text{comp}}(z, \hat{\Omega})]}{V_{\text{c}}(\boldsymbol{\lambda}_{\text{c}})} \frac{dV_{\text{c}}}{dz}(z; \boldsymbol{\lambda}_{\text{c}}), \\
 p_{\text{rate}}(z|\boldsymbol{\lambda}_{\text{z}}) &= \frac{\psi(z; \boldsymbol{\lambda}_{\text{z}})}{(1+z)}, \\
 \xi(\boldsymbol{\lambda}) &= \int P_{\text{det}}(\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{c}}) p(m_1, m_2|\boldsymbol{\lambda}_{\text{m}}) p_{\text{gal}}(z, \hat{\Omega}|\boldsymbol{\lambda}_{\text{c}}) p_{\text{rate}}(z|\boldsymbol{\lambda}_{\text{z}}) d\boldsymbol{\theta},
 \end{aligned} \tag{3.85}$$

where the *GW kernel* $\mathcal{K}_i(z, \Omega|\boldsymbol{\lambda}_{\text{c}}, \boldsymbol{\lambda}_{\text{m}})$ is the marginalization of the posterior over m_1, m_2 reweighted to be proportional to the adopted mass function. The kernels \mathcal{K}_i are evaluated for each set of mass function parameters $\boldsymbol{\lambda}_{\text{m}}$ and cosmological parameters $\boldsymbol{\lambda}_{\text{c}}$ (i.e. for each step in the inference), by computing for all the posterior samples the redshift and source–frame masses, and interpolating the posterior samples in $(z, \text{RA}, \text{Dec})$ with weights corresponding to $p(m_1, m_2|\boldsymbol{\lambda}_{\text{m}})/(d_L^2 (1+z)^2 \partial d_L/\partial z)$.

3.2.6 CHIMERA: a new code for standard sirens analyses

This section presents **CHIMERA** ⁵ (Combined Hierarchical Inference Model for Electromagnetic and gRavitational wave Analyses) a new Python code that allows to jointly fit cosmological and astrophysical source populations parameters leveraging information from galaxy catalogs. The ultimate goal of CHIMERA is to compute the result of Eq. (3.85) in an efficient and accurate way. The code is designed to be accurate for different scenarios, encompassing bright, dark, and spectral sirens methods, and computationally efficient in view of next-generation GW observatories and upcoming galaxy surveys. While it draws some functionality from two previously released and thoroughly tested codes, **DarkSirensStat** (Finke et al., 2021) and **MGCosmoPop** (Mancarella et al., 2022b), CHIMERA is a fully new code, that extends both previous ones by providing a complete and different likelihood modeling, and different treatments of various aspects. Figure 3.9 shows the workflow of CHIMERA. The key aspects are presented in the following paragraph, while a more detailed description of the code validation is given in Appendix A.2.

⁵Available at <https://chimera-gw.readthedocs.io/>

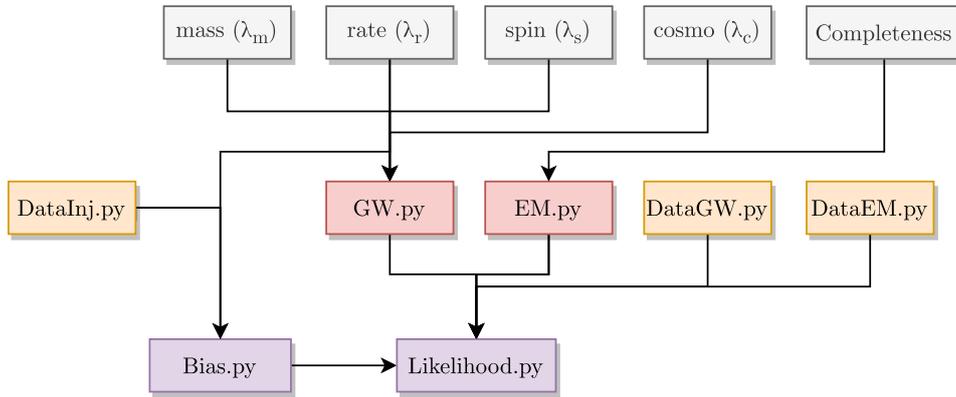


FIGURE 3.9: Workflow of **CHIMERA**. The main modules include functions related to the full likelihood computation (violet), computation of probabilities (red), data file I/O (yellow), and model functions (gray).

The core modules of **CHIMERA** are `Likelihood.py` and `Bias.py`. In particular, the file `Likelihood.py` contains classes to compute the product of the integrals in Eq. (3.85), while `Bias.py` allows one to calculate the selection effects expressed in the term $\xi(\boldsymbol{\lambda})$. The computation of the likelihood includes functions to analyze the gravitational wave (`GW.py`) and electromagnetic (`EM.py`) information. These modules contain specific methods to perform all the preliminary computations that do not change during the likelihood evaluation (a more extended discussion can be found in the next paragraphs). Data are loaded with tailored modules present in `DataGW.py` (e.g., `DataGWMock` and `DataLVK`) and `DataEM.py` (e.g., `MockGalaxiesMICEv2` and `GLADEPlus`), respectively.

Likelihood

We start by discussing the implementation of the likelihood in **CHIMERA**. An illustration of the pipeline is presented in Fig. 3.10, with the main aspects elaborated in the following paragraphs.

3D GW kernel. The term \mathcal{K}_{gw} is a smooth interpolation of the GW posteriors obtained with a weighted kernel density estimate (KDE) using the weights discussed in Section 3.2.5. Compared to Gaussian approximation adopted in the skymaps files released by LVK, directly evaluating the posteriors using a kernel in the $(z(d_L|\boldsymbol{\lambda}_c), \text{RA}, \text{Dec})$ space results in a more accurate representation of the d_L posteriors (see Gray et al., 2020). We take advantage of the KDE weighting to untangle: (i) the priors of the *source-frame* mass distribution, (ii) the priors adopted in the GW event analysis (e.g., prior on d_L^{-2}), and (iii) the conversion of the posterior samples from detector- to source-frame. Figure 3.10 shows an illustration of \mathcal{K}_{gw} in the $(z, \text{RA}, \text{Dec})$ assuming a fiducial flat Λ CDM cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

Angular integration and pixelization. The integral in $d\Omega$ is performed by pixelizing the GW sky localization region and averaging each pixel contribution, while the integral in dz is

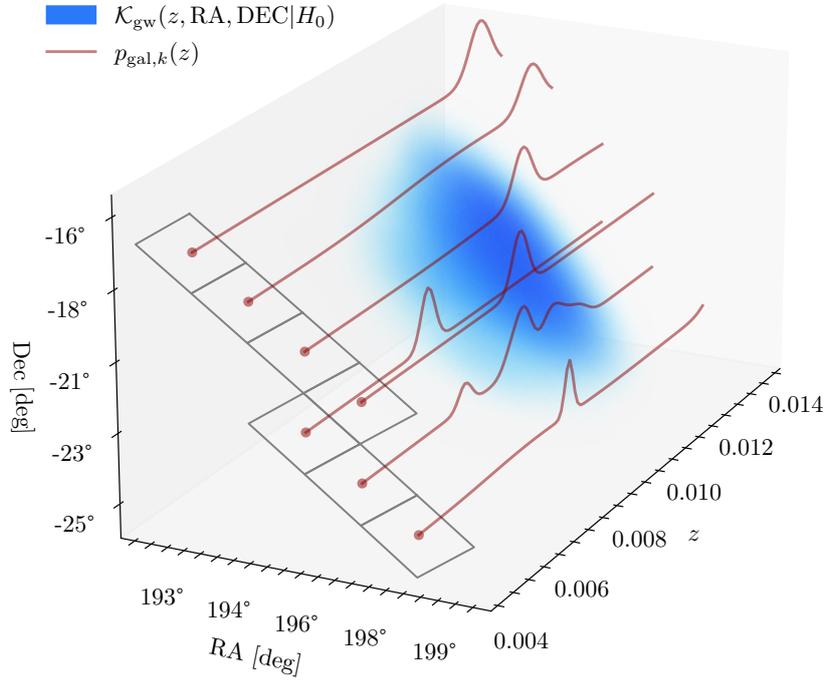


FIGURE 3.10: Visual representation of the underlying workflow of **CHIMERA**. The GW probability is approximated by a three-dimensional KDE (blue cloud), while the galaxy probabilities p_{gal} (red lines) are evaluated by summing the contribution of galaxies enclosed within each pixel (black lines) in the GW sky localization area.

Table 3.2. Various Healpix Pixelizations schemes.

n_{side}	n_{pix}	pix_area [deg ²]
4	192	214.9
8	768	53.72
16	3072	13.43
32	12288	3.36
64	49152	0.84
128	≈ 200 k	0.21
256	≈ 800 k	0.05
512	≈ 3 M	0.013
1024	≈ 13 M	0.003

done by numerical integration. In particular, CHIMERA implements the equal-area pixelization scheme of `Healpix` (Górski et al., 2005; Zonca et al., 2019). The $\hat{\Omega} = (\varphi, \theta)$ parameters are mapped into $(nside, pix)$ parameters, where $nside$ controls the number of pixels $npix = 12 nside^2$ the skymap is divided into, and pix identifies the individual pixel of a given pixelization scheme. CHIMERA includes an adaptive pixelization procedure which allows to fix the desired number of pixels into which each GW event is divided, homogenizing the pixelization scheme in presence of events with very diverse localization areas. As a reference, Table 3.2 presents the skymap resolution that can be obtained for various pixelization schemes. Current GW analyses implement values of $nside = 32, 64$ (e.g., Gray et al., 2022, 2023) so that each pixel covers $1 - 3 \text{ deg}^2$. The quantity $p_{\text{gal}}(z, \hat{\Omega} | \boldsymbol{\lambda}_c)$ is then computed for each pixel by considering the contribution of all the galaxies enclosed within it. This method has three advantages: (1) *improved computational efficiency*: by grouping potentially thousands of galaxies into each pixel, the computational cost can be significantly reduced. For example, in a typical event from GWTC-3, this leads to a factor ~ 100 improvement; (2) *direction-dependent catalog completeness*: compared to assuming a uniform completeness correction in Ω , this method provides a more accurate estimate of P_{comp} (e.g., consider the presence of the Milky Way which limits galaxy detectability). This approach has been proposed and successfully applied in previous dark sirens analyses (Finke et al., 2021; Gray et al., 2022).

Mask completeness. The completeness is computed following the mask method presented in Finke et al. (2021). First of all, pixels are grouped together in N_m masks by applying the agglomerative clustering algorithm of `sklearn` (Pedregosa et al., 2011). The feature chosen for the grouping is the total number of galaxies inside each pixel. The number/luminosity density $n_m(z)$ is then computed inside each mask in N_z redshift bins of a given width. The parameters N_m and N_z must be chosen carefully to ensure a statistically significant number of galaxies in each bin. To smooth out fluctuations of the galaxy structure, a smoothing Gaussian filter is applied to $n_m(z)$. The resulting function is then divided by the target number density or luminosity density. The $P_{\text{compl}}(z)$ is defined by this ratio, so it can be lower/higher than 1 because of under/over-densities in the galaxy distribution. To preserve the notion of completeness, we set $P_{\text{compl}}(z < z_*) = 1$, where z_* is the larger redshift for which P_{compl} crosses 1. This is equivalent to deeming the catalog as complete up to the last over-density. Figure 3.11 shows an example of this implementation.

Efficiency improvements. The normalization factors in p_{pop} that depend solely on hyperparameters are not computed, as they cancel out between the likelihood and the selection effects (see Eq. (3.83)). Specifically, this applies to the $V_c(\boldsymbol{\lambda}_c)$ terms and to the integral of Eq. (3.60). Additionally, the quantity $p_{\text{gal}}(\Omega, z | \boldsymbol{\lambda}_c)$ is cosmology-dependent, but not on H_0 , since this constant cancels out between the numerator and the denominator of Eq. (3.65). Therefore, when considering the parameter H_0 , the pixelized quantity p_{gal} can be computed once for all, given a specific galaxy catalog.

Computational efficiency of the cosmological library. The cosmological functions are called multiple times in the pipeline, so it is important to avoid any bottleneck in their evaluation while, at the same time, maintaining a flexible framework for easy distribution. We perform a thought study of the computational efficiency of multiple cosmological codes. In particular,

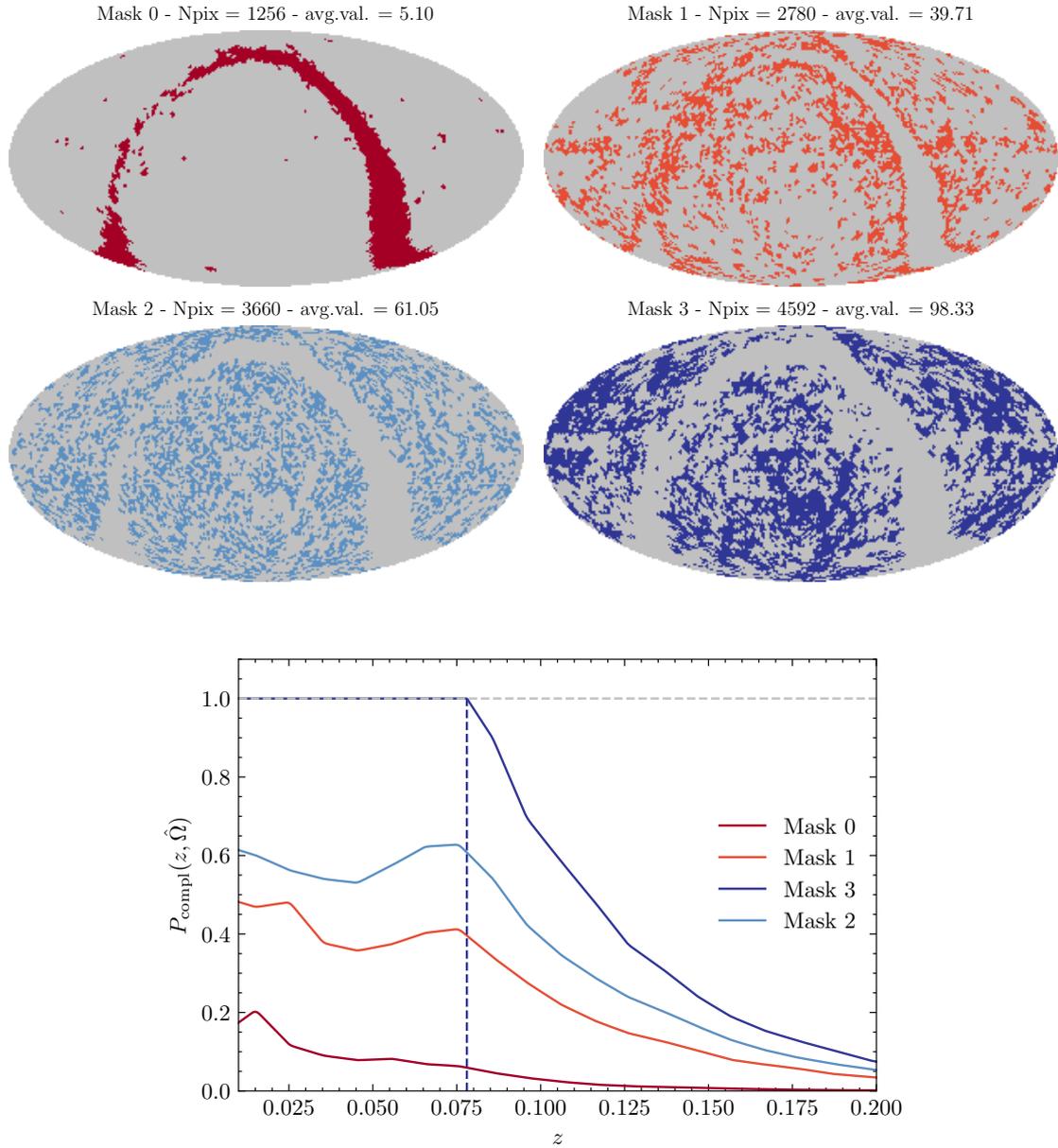


FIGURE 3.11: Example of mask completeness approach using $n_{\text{side}} = 32$ and $N_{\text{m}} = 4$ masks with the GLADE+ galaxy catalog (Dályá et al., 2022). *Upper panels:* Mollview projection of the pixels belonging to the four masks. The titles include the number of pixels and the average number of galaxies in each pixel. *Lower plot:* completeness as a function of redshift for each mask. The completeness is computed using K -band luminosity, assuming a reference Schechter (with parameters $\alpha = -1.02$; $\Phi^* = 5.5 \times 10^{-3}$; $L_* = 1.1 \times 10^{11} L_{\odot}$), and applying a cut $L > 0.6 L_*$.

Table 3.3. Relative execution time for the evaluation of $d_L(z; \boldsymbol{\lambda}_c)$ in different cosmologies. The values are normalized with respect to the execution time obtained with **astropy**.

Code (Language/framework)	fACDM	fwCDM	fw ₀ w _a CDM
astropy (Python+Cython)	1	1	1
CHIMERA (Python+XLA)	0.3	0.8	0.7
C-LAL libs (C+Cython)	1.2	0.2	0.1
CLASS (C)	0.8	0.1	0.1

Note. — The test are performed with a 4 GHz processor and 16 GB RAM system, averaging 100 evaluations of a vector of 10^6 elements in the redshift range $0 < z < 2$.

we consider the widely used **astropy** (Astropy Collaboration et al., 2018), LAL cosmological libraries (LIGO Scientific Collaboration, 2018), and **CLASS** code (Lesgourgues, 2011). We also develop a new cosmological library in **CHIMERA**, based on the hypergeometric solution of Baes et al. (2017) and the just-in-time compilation using **XLA** (Accelerated Linear Algebra) in **JAX** (Bradbury et al., 2018). Table 3.3 shows the results of the execution time for the evaluation of $d_L(z)$ normalized with respect to **astropy**. In the case of a flat Λ CDM model, the implementation in **CHIMERA** is a factor of 3 faster than the other tested alternatives. Similar results are obtained also for dd_L/dz and dV_c/dz . When more extended cosmological models are used, C-based libraries outperform Python-based codes by a factor of 10. Further work should focus on continuing to optimize cosmological functions by, e.g., porting C libraries, developing approximate analytical methods, or leveraging machine learning techniques to achieve additional speedups while maintaining accuracy. This will be important to scale up the analysis to handle growing data sets in the future.

In **CHIMERA**, the Likelihood class is initialized as follows:

```
from CHIMERA.Likelihood import MockLike

like = MockLike(model_cosmo, model_mass, model_rate,
                data_GW, data_GW_smooth, data_GAL_dir,
                nside_list, npix_event, sky_conf,
                z_int_H0_prior, z_int_sigma, z_int_res)
```

CODE EXAMPLE 3.1: Inizialization of the Likelihood in **CHIMERA**.

The algorithm proceeds as follows:

1. First of all, the class **MockLike** stores all the population models (**model_***), GW data (**data_GW***), galaxy data (**data_GAL***), pixelization parameters (**nside_list**, **npix_event**, **sky_conf**), and integration parameters (**z_int_H0_prior**, **z_int_sigma**, **z_int_res**).

2. The **GW** class is initialized and the pixelization and redshift grids are pre-computed. To optimize the computation in case of large galaxy catalog analysis, the code not only restricts the sky localization, but also the redshift integration grid. The first task is done starting from posterior distributions in RA and Dec. The user can choose the approximate number of pixels desired for each event (**npix_event**) to be found within a confidence level ellipse (**sky_conf**), given a list of possible pixelizations (**nside_list**). **CHIMERA** optimizes the pixelization of each event (in particular, the *nside* parameter of HealPix) to obtain the number of pixels closest to **npix_event**. The second task must take into account that by varying the cosmology, also \mathcal{K}_{gw} varies, so to avoid biases, the integration grid must encompass all the redshift range explored during the inference. This is obtained starting from the GW posteriors on d_L and defining the range inside **z_int_sigma** standard deviations at a resolution of **z_int_res** spanning all the range⁶ of H_0 explored (**z_int_H0_prior**, see Fig. 3.14).
3. The **EM** class is initialized, and the quantity $p_{\text{cat}}(z, \hat{\Omega} | \boldsymbol{\lambda}_c)$ (Eqs. 3.64 and 3.65) is pre-computed pixel-by-pixel on the redshift grids. At this point, based on the included catalog (e.g., GLADE+, MICEv2), it is possible to activate catalog-related tasks (e.g., luminosity cut, or associate user-defined redshift uncertainties). Similarly, if defined, the completeness is computed, and the pixelized completeness function $P_{\text{compl}}(z)$ (see the discussion in the above paragraph) is stored in the class to be accessible during the inference.

The pixelization approach and the pre-computation of p_{gal} are essential for next-gen galaxy surveys. The first allows us to reduce the dimension, combining the probability of $10 - 10^3$ galaxies in one single pixel ensuring granularity in the sky analysis. The second allows us to avoid the computation of p_{gal} at each step of the MCMC inference. While this last approximation is not a problem for H_0 inference since both the numerator and denominator of Eq. (3.65) have a H_0^3 dependence, the impact of a possible bias on Ω_m should be carefully assessed when future analyses using events at higher z will be carried out.

Bias

We now move to the selection bias term and its implementation in **CHIMERA**. In general, following Section 3.2.4, its evaluation is done with the injections approach, relying on an external catalog of injections and computing Eq. (3.83) at varying $\boldsymbol{\lambda}$. The main aspects are elaborated in the subsequent paragraphs.

Inclusion of the galaxy distribution in the selection effects. The most important part is the evaluation of the $p_{\text{pop}}(\boldsymbol{\theta}_i | \boldsymbol{\lambda})$ term, which would require the computation of the completeness for every injected event, making this step very computationally inefficient. However, we can take advantage of the fact that over the large number of injections required to accurately estimate the Monte Carlo integral for the selection bias term, the effect of local inhomogeneities smooths out, replicating on average the overall galaxy distribution $\bar{p}_{\text{cat}}(z)$. Therefore, when

⁶In practice the grid starts at $z(d_L - N \sigma_{d_L} | H_{0,\text{min}})$ and ends at $z(d_L + N \sigma_{d_L} | H_{0,\text{max}})$, where N is the number of standard deviations (**z_int_sigma**) and $[H_{0,\text{min}}, H_{0,\text{max}}]$ the interval of H_0 spanned in the inference.

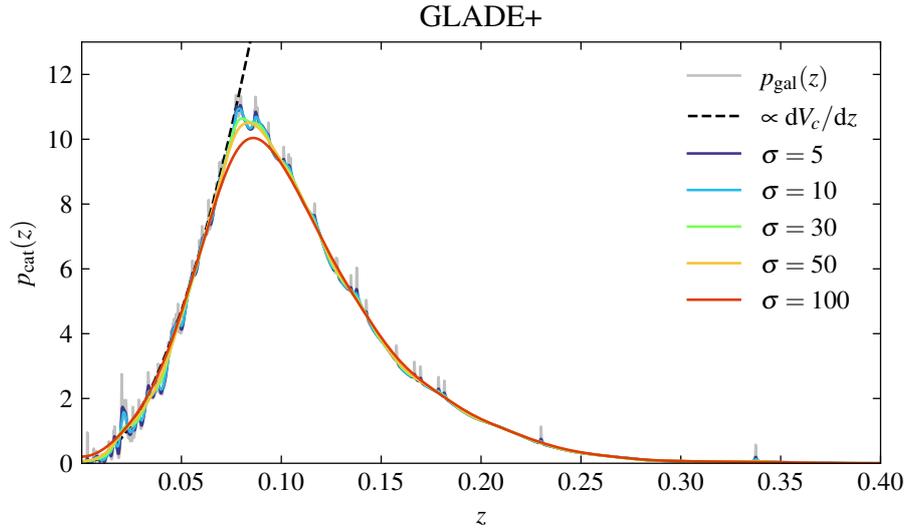


FIGURE 3.12: Example of the computation of the average interpolant of GLADE+ (Dályá et al., 2022). The line represents the original redshift distribution (gray), a uniform in comoving volume distribution (dashed black), and the profiles resulting after a Gaussian smoothing varying the standard deviation σ .

using galaxies, we compute the selection bias term using the interpolant of the whole catalog $\bar{p}_{\text{cat}}(z)$ inside $p_{\text{pop}}(\theta_i|\lambda)$. The interpolants are obtained by computing Eq. (3.64) for the whole catalog, then smoothing the resulting profile with a Gaussian filter with standard deviation σ . Two examples for GLADE+ (Dályá et al., 2022) and MICEv2 (Fosalba et al., 2015b) catalogs are shown in Figs. 3.12 and 3.13. We find that a good choice for σ is 30 for GLADE+ and 50 for MICEv2.

In CHIMERA, the Bias class is initialized as follows:

```
from CHIMERA.Bias import Bias

bias = Bias(model_cosmo, model_mass, model_rate,
            file_inj, snr_th)
```

CODE EXAMPLE 3.2: Initialization of the selection bias term in CHIMERA.

The algorithm proceeds as follows:

1. First of all, the class `Bias` stores all the population models (`model_*`), the directory of the GW injection catalog data (`file_inj`), and the SNR threshold to be applied. Optional arguments also include the catalog interpolant function. If not given, the bias is evaluated for a uniform in comoving volume galaxy distribution (i.e., $\propto dV_C/dz$).
2. The injection catalog is loaded by applying the chosen SNR cut. It is important to ensure that this cut is equivalent to the one adopted when creating the catalog of GW events to analyze.

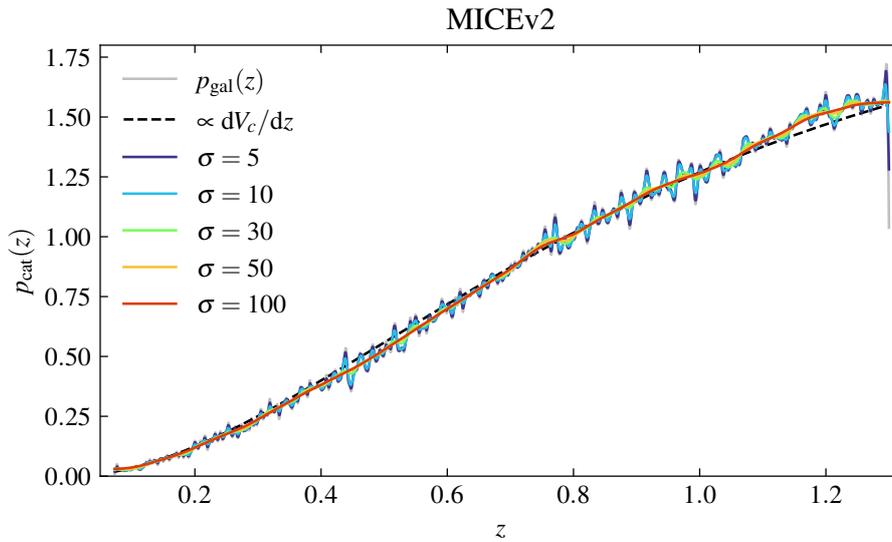


FIGURE 3.13: As Fig. 3.12, but for the MICEv2 (Fosalba et al., 2015b) galaxy catalog.

Full likelihood

Finally, we discuss how the *full likelihood* is computed in CHIMERA. Following Eq. (3.47), it can be conveniently derived by first calculating the log-likelihood for all the events and then subtracting the log-bias term, which is computed only once and multiplied by the number of events. The computation of the likelihood and selection bias is performed by calling the `.compute()` methods or, in logarithmic form, the `.compute_ln()` methods. In the latter case, the models should also be given in logarithmic form. The full likelihood is then computed as:

```
lnlike = like.compute_ln(lambda_cosmo, lambda_mass, lambda_rate) - \
        like.Nevents * bias.compute_ln(lambda_mass, lambda_cosmo,
        lambda_rate)
```

CODE EXAMPLE 3.3: Computation of the full likelihood in CHIMERA.

The algorithm proceeds as follows:

1. The only arguments that need to be passed to the `.compute()` methods are the population hyperparameters: `lambda_cosmo` (λ_c), `model_mass` (λ_m), and `model_rate` (λ_r). Eventually, the user can set the `inspect` variable to True to save the intermediate results ($p_{\text{gal}}(z)$, $p_{\text{rate}}(z)$, ...).
2. The likelihood computation is performed taking advantage of the pre-computed $p_{\text{gal}}(z)$ that does not change during the inference since CHIMERA works in redshift space. This is not the case for \mathcal{K}_{gw} , which varies for two main factors: (i) λ_c affecting the $d_L \rightarrow z$ conversion of the GW posterior samples and (ii) λ_c and λ_m on the KDE weighting, as discussed in Section 3.2.5. A working example is shown in Fig. 3.14. The “convolution” of the two terms provides the log-likelihood for one event. This process is repeated for all the events and the log-likelihoods are summed.

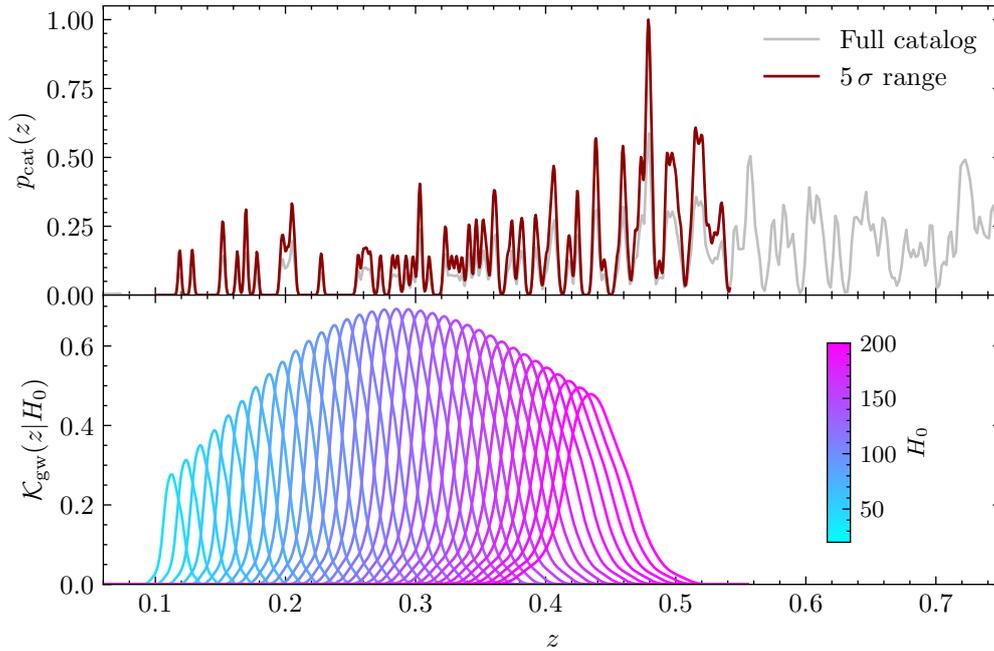


FIGURE 3.14: Illustration of the algorithm to compute the likelihood in CHIMERA. *Upper panel:* galaxy distribution p_{gal} evaluated on the 5σ grid spanning all the parameter space explored in the inference (see text). *Lower panel:* GW kernel \mathcal{K}_{gw} at varying cosmology.

3. The selection bias computation is carried out as explained in Section 3.2.4 using the pre-loaded injections. As a working example, in Fig. 3.15 we show the behavior of the selection bias term, for the various hyperparameters included in the analysis that is described in Section 3.5. We note that these one-dimensional tests are primarily intended for visualization purposes. In the MCMC inference, the selection bias term is computed within the full n -dimensional space of the hyperparameters.

Included models

Finally, we list here below the population models currently included in CHIMERA.

1. Mass distributions (`mass.py`): `logpdf_TPL` (Truncated Power Law), `logpdf_BPL` (Broken Power Law), `logpdf_PLP` (Power Law + Peak), `logpdf_PL2P` (Power Law + 2 Peaks). See Fig. 3.6.
2. Rate evolutions (`rate.py`): `logphi_PL` (Power Law), `logphi_MD` (Madau-like). See Fig. 3.7.
3. Spin distributions (`spin.py`): `logpdf_G` (Gaussian), `logpdf_U` (uniform).
4. Cosmological models (`cosmo.py`): `flCDM`, `flCDM modified gravity`

All the functions accept parameters in the form of dictionaries, an example for the cosmological parameters is `lambda_cosmo = {'H0':70.0, 'Om0':0.3}`.

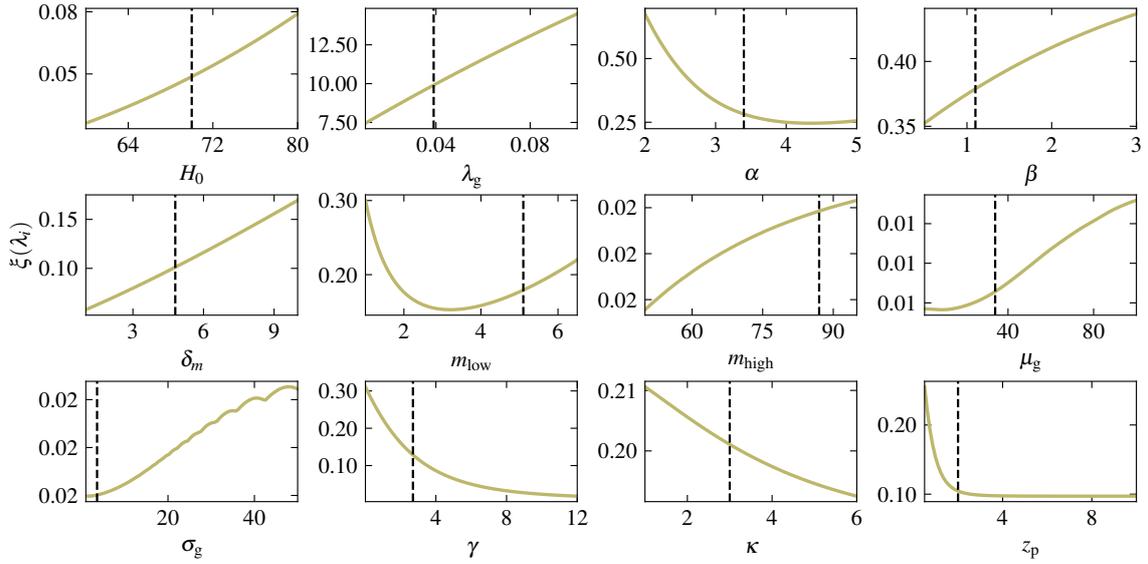


FIGURE 3.15: One-dimensional selection effect functions $\xi(\lambda_i)$ for the twelve population hyperparameters studied in this work (see Table 3.4.2) and computed for an O5-like LVK configuration with a SNR threshold of 25.

We perform extensive validation tests using an external code and high SNR events to assess potential biases in the recovery of the hyperparameters, with a specific focus on H_0 . These efforts are presented in detail in Appendix A.2, but we summarize here the main conclusions. Overall, we verify that the results obtained with **CHIMERA** in spectral sirens mode (no galaxy catalog) are in almost perfect agreement with the ones from **MGCosmoPop** (Mancarella et al., 2022b), despite the differences in the likelihood. In the most informative mode, i.e. bright sirens, **CHIMERA** also performs well, as demonstrated by the results obtained for GW170817 and GW190814, which are in agreement with the literature (see Section 3.3). At the moment of the writing of this Thesis, it was not possible to test the code against alternative dark+spectral sirens codes since they were not available. In the near future, we are planning a full comparison against the recently released **icarogw** and **gwcsmo**, which will, however, require the development of dedicated Mock Galaxy catalogs. In this work, we validated the code by performing one-dimensional posterior analyses (Appendix A.2) and rigorous statistical tests to assess possible biases (Section 3.5.5).

3.3 Best-localized observed GW events

In this Section, we test CHIMERA by analyzing the two best-localized events of the Third Gravitational Wave Catalog (GWTC-3) released by LVK (Abbott et al., 2023b). We focus on determining the Hubble constant H_0 , fixing the other hyperparameters. As a reference galaxy catalog, we use the GLADE+ galaxy catalog (Dályá et al., 2022), which includes data from multiple galaxy surveys (GWGC, 2MPZ, 2MASS XSC, HyperLEDA, WISEXSCOSPZ, and SDSS-DR16Q quasars) and is largely employed in standard sirens analyses (e.g., Abbott et al., 2023c). As done in previous analyses, to obtain a more complete catalog of potential hosts, we select galaxies with a luminosity above a certain threshold L_K^* . The implicit assumption is that the true host is selected with the considered threshold. This is well motivated by the fact that the luminosity is a reasonable proxy for the stellar mass (in K-band) and the star formation rate (in B-band), therefore more luminous galaxies have a higher chance to host mergers. This is currently a common assumption in standard siren analysis; in the future, a detailed study would be beneficial for the field, to more accurately assess the correlations with galaxy properties and potential biases in standard sirens analyses.

The completeness is computed with the mask method, as outlined in Section 3.2.6. The sky is pixelized with a size of 0.83 deg^2 and pixels are organized into nine distinct masks based on the galaxy counts in each pixel. For each mask, the completeness fraction $f_{\mathcal{R}_H}$ (Eq. (3.76)) is computed by comparing the number density to a reference Schechter function with parameters from Kochanek et al. (2001). We fix the population parameters and study the posterior on H_0 .

3.3.1 The case of GW170817

When considering the GW data alone, GW170817 stands out as the best localized source to date. It also serves as an ideal benchmark to apply our method and assumptions, since its host galaxy, NGC 4993, has been uniquely identified (Abbott et al., 2017a). Initially, we consider each galaxy within the 90% GW sky localization area (about 28 deg^2) as a potential host (Fig. 3.16a). We adopt the following population model: flat mass distribution between 1 and $3 M_\odot$ and power-law rate evolution with slope $\gamma = 2.7$. In Fig. 3.16b we show the resulting GW kernel and p_{gal} evaluated inside each pixel. We repeat the analysis assuming the identification of the host galaxy with $z = 0.0100 \pm 0.0005$ (Abbott et al., 2017a). In this case, we assign a weight $w = 1$ to NGC 4993 and $w = 0$ to all the other galaxies (see Eq. (3.64)). The resulting posteriors on H_0 are shown in Fig. 3.16c.

We obtain a value of $H_0 = 73_{-22}^{+58} \text{ km/s/Mpc}$ in the dark siren case, in good agreement with (Fishbach et al., 2019). The primary contribution to the posterior arises from galaxies at approximately $z \sim 0.01$. This value combined with the GW measurement of $d_L \sim 40 \text{ Mpc}$, implies $H_0 \sim 70 \text{ km/s/Mpc}$. The galaxy “groups” at $z \sim 0.02$ (0.04) provide a shallower contribution to $H_0 \sim 150$ (300) km/s/Mpc due to the presence of selection effects which disfavor high values of H_0 (we remind here that $\xi(H_0) \sim H_0^3$). The assumption of a population model for GW170817 provides no significant improvements to the constraints on H_0 , as expected given the featureless mass function adopted. In the bright siren case, we obtain a value of $H_0 = 69_{-8}^{+15} \text{ km/s/Mpc}$, in very good agreement with Abbott et al. (2017a). The

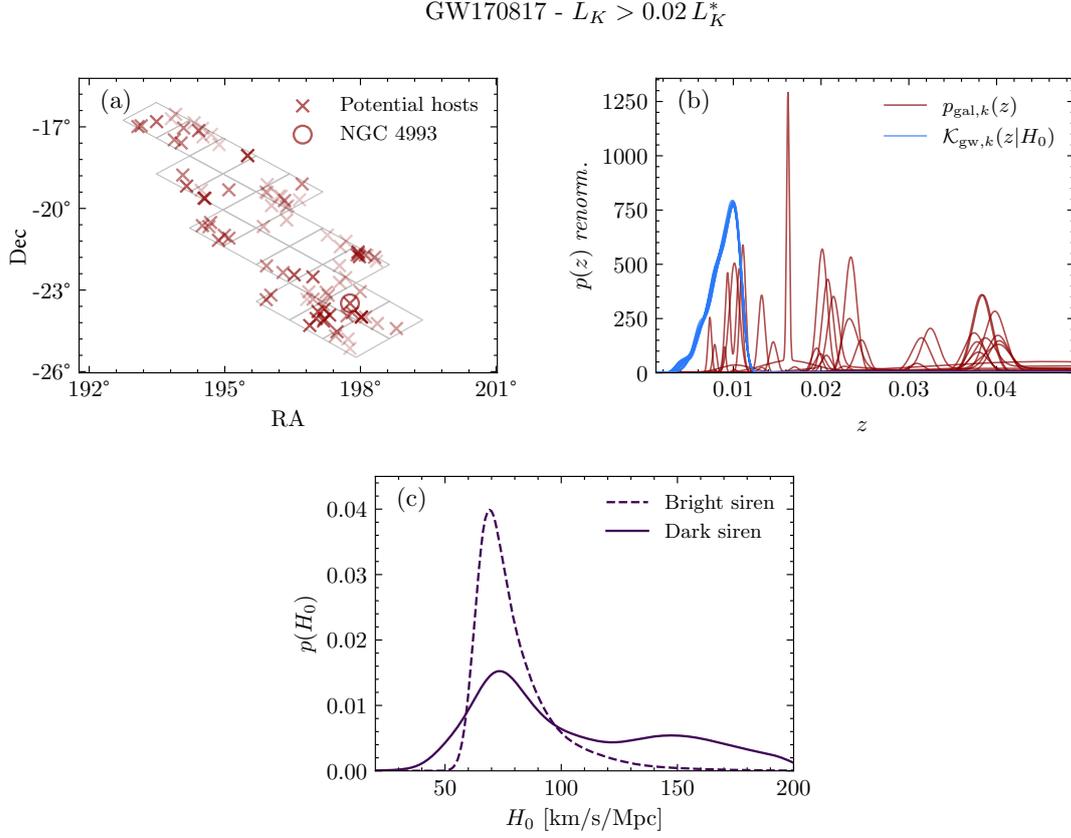


FIGURE 3.16: Analysis of GW170817 with **CHIMERA** using GLADE+ galaxies with a luminosity $L_K > 0.02 L_K^*$. (a) Sky distribution of the potential host galaxies (red crosses) contained in the 90% credible pixels (gray lines). The true host NGC4993 is identified with a circle; (b) Redshift distribution of the GW kernel (blue) assuming $H_0 = 70$ km/s/Mpc and galaxy probability p_{gal} (red) both plotted for each pixel; (c) Posterior distributions on H_0 assuming that the true host is or is not observed.

non-gaussianity of the peak is due to the degeneracy between distance and inclination that results in extended tails at lower d_L (and, as a consequence, in a higher H_0).

3.3.2 The case of GW190814

The second most interesting event is GW190814, located at $d_L \sim 280$ Mpc and with a 90% credible interval sky area of about 20 deg^2 (Abbott et al., 2020b). This event triggered a debate on the nature of the secondary object, which, with a mass of $2.6M_\odot$, is either the lightest BH or the heaviest NS in a binary system. In this analysis, we treat GW190814 as a NSBH merger with a flat secondary mass distribution between 1 and $3M_\odot$ and a PLP primary mass distribution with parameters from Table 3.4.2. In Fig. 3.17b we show the GW kernel and p_{gal} evaluated inside each pixel. The resulting posterior on H_0 is shown in Fig. 3.17c.

We obtain a maximum a posteriori value of $H_0 = 64^{+45}_{-28}$ km/s/Mpc (68% C.L.) consistent with (Gray et al., 2023). The primary contribution to the main peak arises from galaxies

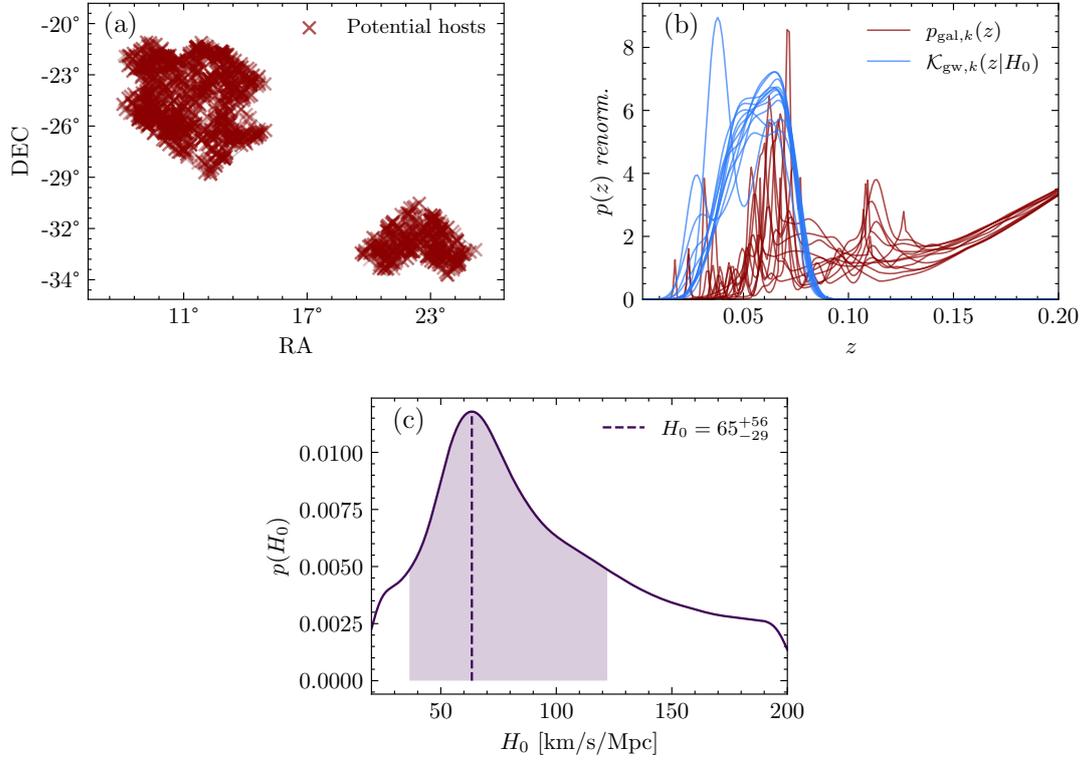
GW190814 - $L_K > 0.6 L_K^*$ 

FIGURE 3.17: Same as Fig. 3.16 for GW190814 binary black hole merger with a more conservative luminosity cut.

at approximately $z \sim 0.06$, while at higher redshifts, there is a noticeable decrease in completeness.

An extensive study of the full GWTC-3 catalog (Abbott et al., 2023b) will be subject of a future investigation, with a specific focus on the systematic effects arising from completeness assumptions. As anticipated, this work will be devoted to exploring standard sirens forecasts for the joint cosmological and astrophysical population parameters. In particular we will study O4-like and O5-like detector configurations in combination with photometric and spectroscopic galaxy surveys.

3.4 The sample

This section presents the procedure to build a realistic mock galaxy and GW catalogs. These will be used to both validate the code and provide forecasts for O4-like and O5-like detector configurations. It is crucial to ensure that in this process of data generation, the procedures are consistent with the assumptions made in the methodology.

3.4.1 Parent Galaxy Catalog

We generate our mock galaxy catalog (hereafter, *parent sample*) starting from the MICE Grand Challenge light-cone simulation (v2)⁷, which populates one octant of the sky (close to 5157 deg²) and is complete for DES-like (Dark Energy Survey) surveys up to an observed magnitude $i < 24$ at redshift $z < 1.4$ (Fosalba et al., 2015b,a; Croce et al., 2015; Carretero et al., 2015; Hoffmann et al., 2015). MICE assumes a flat Λ CDM cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{m,0} = 0.25$, $\Omega_{\Lambda,0} = 0.75$.

While we ideally require a complete catalog with high number density, as a simplifying assumption in this paper we consider only galaxies with stellar masses $\log M_*/M_\odot > 10.5$. This cut is consistent with the idea that the binary merger rate is traced by stellar mass, as also adopted in current standard sirens analysis via absolute magnitude cuts and luminosity weighting (Fishbach et al., 2019; Finke et al., 2021; Gray et al., 2022; Abbott et al., 2023c; Mastrogiovanni et al., 2023; Gray et al., 2023). A similar cut in mass is also considered in (Muttoni et al., 2023) in the context of simulations for the Einstein Telescope.

We subsample the MICEv2 catalog to reproduce the density for the cut described above extracting the galaxies to get a uniform in comoving volume distribution. In the end, we obtain a parent sample of about 1.6 million massive galaxies.

For the redshift uncertainties, we consider two cases. First of all, we explore the possibility of maximizing the galaxy catalog information by having a spectroscopic catalog. This could be done by expanding the currently available catalogs (GLADE+ Dalya et al., 2022) in the future by exploiting the information provided by the next large spectroscopic surveys. As an example, the ESA mission *Euclid* (Laureijs et al., 2011) will provide an all-sky map of spectroscopic redshift in the range $0.9 < z < 1.8$, with an accuracy of $\sigma_z/(1+z) \lesssim 0.001$, and the Dark Energy Spectroscopic Instrument (DESI DESI Collaboration et al., 2016) is planned to observe $\sim 14000 \text{ deg}^2$ covering the redshift range $0.4 < z < 2.1$. Second, we study how the information extracted changes when using photometric redshift, assuming an uncertainty $\sigma_z/(1+z) = 0.05$. This is currently easily accessible with current ongoing surveys like the Dark Energy Survey (e.g., DES has reached $\sigma_z \sim 0.01$, Myles et al. 2021, and this limit can be pushed to $\sigma_z \sim 0.007$ with improved techniques, e.g. Buchs et al. 2019) on a smaller area, and in future surveys like *Euclid* and Rubin Observatory (Ivezic et al., 2019) are planned to extend it to the entire sky to a depth of $H_E \sim 24$ ($H_E \sim 26$ in the Deep Survey) and an expected uncertainty $\sigma_z/(1+z) \lesssim 0.05$.

⁷Available at <https://cosmohub.pic.es/home>

Therefore, two regimes of photometric (hereafter, z_{phot}) and spectroscopic (hereafter, z_{spec}) redshift uncertainties are considered,

$$\sigma_z = \begin{cases} 0.001 (1 + z) & (z_{\text{spec}}) \\ 0.05 (1 + z) & (z_{\text{phot}}) \end{cases} \quad (3.86)$$

3.4.2 Sample of GW events

We generate mock GW events from the parent sample by fixing cosmological hyperparameters λ_c and astrophysical population hyperparameters λ_z and λ_m . We describe the redshift and mass distributions in turn.

For the source-frame merger rate, we assume a Madau-Dickinson evolution (see Eq. 3.63) with $\gamma = 2.7$ consistent with the LVK GWTC-3 results (Abbott et al., 2023a). The limited detection range of current gravitational wave detectors restricts our ability to determine the merger rate at higher redshifts, therefore we assume $z_p = 2$, $\kappa = 3$, consistent with the idea that $\psi(z; \lambda_z)$ follows the galaxy’s star formation rate density with parameters from (Madau & Dickinson, 2014). The catalog of potential sources is then obtained by sampling the parent sample using a weight proportional to the detector-frame merger rate $\psi(z; \lambda_z)/(1 + z)$. For the BBH mass distribution, we adopt a phenomenological “PowerLaw+Peak” (PLP) model (see Eq. 3.58) following the LVK GWTC-3 results (Abbott et al., 2023a). With these assumptions, the cosmological and astrophysical hyperparameters to be studied are:

$$\lambda_c = \{H_0, \Omega_{m,0}\} \quad (3.87)$$

$$\lambda_z = \{\gamma, k, z_p\} \quad (3.88)$$

$$\lambda_m = \{\alpha, \beta, \delta_m, m_{\text{low}}, m_{\text{high}}, \mu_g, \sigma_g, \lambda_g\}. \quad (3.89)$$

The fiducial values and the prior ranges chosen for this work are reported in Table 3.4.2. For the cosmological parameters, we will consider the value of the matter density to be fixed to its fiducial value.

GW data generation

To generate the GW posterior samples we use **gwfast**⁸ (Iacovelli et al., 2022a,b), a state-of-the-art Fisher-matrix (FIM) code specifically designed to study parameter estimation capabilities of GW networks under different configuration assumptions. This provides an approximation to a full Bayesian parameter estimation, with the likelihood of the GW parameters in Eq. 3.90 assumed to be given by a multivariate Gaussian distribution. While this is not true in general, it is a good approximation for the high SNR events that we are considering in this work. The advantage of the FIM approach is the greatly reduced computational cost, which is crucial in particular when producing the bias term through the MC integral in Eq. 3.83. Being developed to handle large catalogs, **gwfast** is then a convenient choice.

⁸Available at: <https://gwfast.readthedocs.io/>

Table 3.4. Summary of the population hyperparameters and priors used in this work.

λ_i	Description	Fiducial	Prior
Cosmology (flat ΛCDM)			
H_0	Hubble constant [$\text{km s}^{-1} \text{Mpc}^{-1}$]	70.0	$\mathcal{U}(10.0, 200.0)$
$\Omega_{\text{m},0}$	Matter energy density	0.25	Fixed
Rate evolution (Madau-like)			
γ	Slope at $z < z_p$	2.7	$\mathcal{U}(0.0, 12.0)$
κ	Slope at $z > z_p$	3	$\mathcal{U}(0.0, 6.0)$
z_p	Peak redshift	2	$\mathcal{U}(0.0, 4.0)$
Mass distribution (PowerLaw+Peak)			
α	(Primary) slope of the power law	3.4	$\mathcal{U}(1.5, 12.0)$
β	(Secondary) slope of the power law	1.1	$\mathcal{U}(-4.0, 12.0)$
δ_m	(Primary) smoothing parameter	4.8	$\mathcal{U}(0.01, 10.0)$
m_{low}	Lower value [M_\odot]	5.1	$\mathcal{U}(2.0, 50.0)$
m_{high}	Upper value [M_\odot]	87.0	$\mathcal{U}(50.0, 200.0)$
μ_g	(Primary): mean of the Gaussian peak [M_\odot]	34.0	$\mathcal{U}(2.0, 50.0)$
σ_g	(Primary): standard deviation of the Gaussian peak	3.6	$\mathcal{U}(0.4, 10.0)$
λ_g	(Primary): fraction of the Gaussian peak	0.039	$\mathcal{U}(0.01, 0.99)$

We assume quasi-circular non-precessing BBH systems. Their waveform is characterized by the detector-frame parameters

$$\boldsymbol{\theta}^{\text{det}} = \{\mathcal{M}_c, \eta, \chi_{1,z}, \chi_{2,z}, d_L, \theta, \phi, \iota, \psi, t_c, \Phi_c\} \quad (3.90)$$

where \mathcal{M}_c is the detector-frame chirp mass, η is the symmetric mass ratio, $\chi_{1/2,z}$ are the adimensional spin parameters along the direction of the orbital angular momentum, d_L is the luminosity distance, $\theta = \pi/2 - \text{Dec}$ and $\phi = \text{RA}$ are the sky position angles, ι refers to the inclination angle of the binary's orbital angular momentum with respect to the line of sight, ψ is the polarization angle, t_c is the coalescence time, and Φ_c is the phase at coalescence.

First of all, we generate a population of GW events by following the prescriptions given in Section 3.4.2, so that each source is characterized by a set of source-frame parameters $\boldsymbol{\theta}$. We assume the following distributions: the sky position angles are uniform on the octant covered by the parent sample, the inclination angles have a uniform distribution in $\cos \iota$ in the range $[0, \pi]$, the polarization angle and coalescence phase have a uniform distribution in $[0, \pi]$ and $[0, 2\pi]$, respectively, and the time of coalescence, given as Greenwich mean sidereal time, is uniform in the $[0, 1]$ day. Instead, for the spin parameters, we adopt a flat distribution between $[-1, 1]$ for the components aligned with the orbital angular momentum.

For each source we simulate GW emission using the `IMRPhenomHM` (London et al., 2018; Kalaghatgi et al., 2020) waveform approximant, which includes the contribution of subdominant modes to the signal, that are of great relevance to describe in particular the merger

phase of BBH systems. We consider two network configurations. The first, denoted as *O4-like*, is composed of a network of the two LIGO interferometers at Hanford and Livingston, USA, the Virgo interferometer in Cascina, Italy, and the KAGRA interferometer in Japan. For the second configuration, denoted as *O5-like* the network includes the two LIGO, Virgo, and KAGRA instruments, as well as a LIGO detector located in India. We assume sensitivity curves representative of the O4 and O5 runs of the LVK Collaboration, with public sensitivity curves from [Abbott et al. \(2016a\)](#).⁹ We assume a 100% duty cycle in all cases.

Then, we select samples of GW events as follows,

- *O4-like*: 100 events with a network SNR > 12 ;
- *O5-like*: 100 events with a network SNR > 25 .

These cuts are designed to yield the 100 best events for each configuration over approximately one year of observation. We determined these numbers simulating with **GWFAST** a one-year observing run for each of the two scenarios, with a population of BBHs with overall merger rate calibrated on the latest LVK constraints (see e.g., [Iacovelli et al., 2022b](#)). We note that the fact that the simulation is performed on one octant of the sky is irrelevant in our case, as we do not constrain direction-dependent hyperparameters.

The main properties of the galaxy and GW catalogs are summarized in [Fig. 3.18](#). The top panels show the GW skymaps of the events detected in the two configurations overlaid to the galaxy sky distribution. The central panels present the scatter plots of the sky localization area versus the error on d_L (with a color scale giving information on the SNR). Finally, the bottom panels show the redshift and mass distribution of the detected GW events, as well as the distribution of the number of galaxies found within their localization volumes.

The *O4-like* events are at redshift $z \lesssim 0.9$ with sky areas between a few and a few hundreds deg^2 . This typically results in more than ~ 500 potential host galaxies for the 90% of the events ([Fig. 3.18](#), bottom-right panel). A particularly lucky event is present with small localization area ($\sim 2 \text{ deg}^2$) and high SNR (~ 55 , see the mid-left panel of [Fig. 3.18](#)). While this event represents an outlier of the distribution that can be ascribed to a statistical fluctuation, it still represents a possibility with this configuration. For the *O5-like* events, the redshift distribution remains limited to $z \lesssim 1$ as a consequence of the higher SNR cut, while the larger and more sensitive detector network substantially improves the localization capabilities. This typically results in more than ~ 50 potential host galaxies for the 90% of the events ([Fig. 3.18](#), bottom-right panel). In this configuration, there is a significant tail of events with just few tens of galaxies, corresponding to sky localization regions of less than one degree with SNR that can exceed 100. Overall, while the number of galaxies in the localization volume depends on the assumption on the galaxy catalog employed, it is interesting to observe the $10\times$ reduction in the number of potential host galaxies between the O4 and O5 networks. This improvement, following from the much smaller localization volumes, is a key factor to obtain improved dark sirens constraints as discussed in [Section 3.5](#).

Ultimately, for the computation of the selection bias, we generate injection sets for both the *O4-like* and the *O5-like* scenarios with **GWFAST**, adopting the same selection cuts as for

⁹The amplitude spectral densities can be found at <https://dcc.ligo.org/LIGO-T2000012/public>. For O4, we use **kagra_3Mpc** for KAGRA. For O5, we use **AplusDesign** for the three LIGO detectors, **avirgo_05low_NEW** for Virgo, and **kagra_80Mpc** for KAGRA.

the GW catalogs. The injections cover the same sky area as the catalogs and a distance range that arrives up to the detector horizon for the given SNR cuts. The injections set is made of $N_{\text{inj}} = 2 \times 10^7$ and 4×10^7 events, resulting in about 1.5×10^6 and 1×10^6 detected events in the *O4-like* and *O5-like* scenario, respectively. These are then used to estimate the selection bias as in Eq. 3.83.

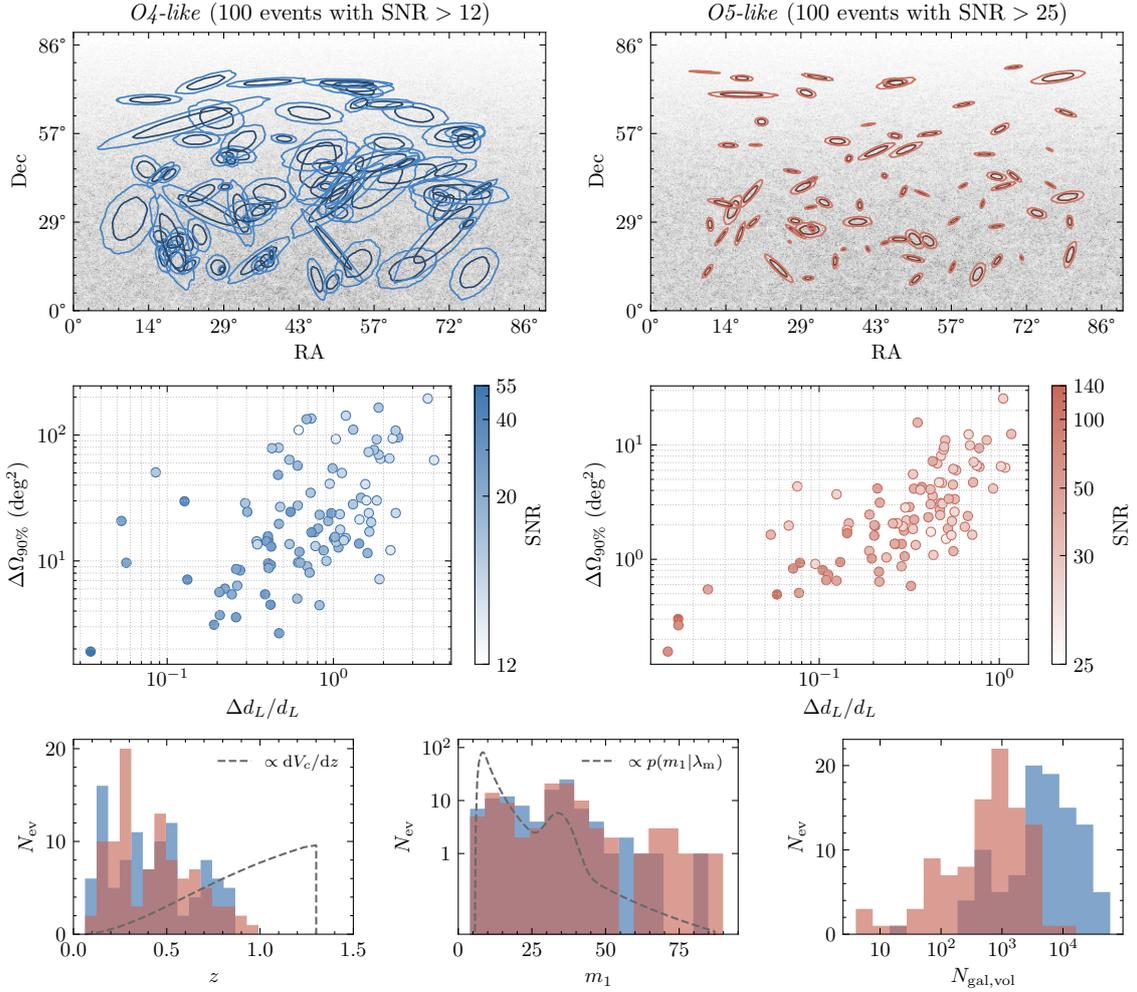


FIGURE 3.18: Main properties of the simulated *O4-like* (blue) and *O5-like* (red) GW catalogs. *Upper panels*: GW sky localization areas at 1 and 2σ overlaid to their potential host galaxies (gray) extracted from MICEv2. *Middle panels*: distribution of the relative uncertainty on the luminosity distance and sky localization area as a function of the network S/N. *Lower panels*: distribution of the GW events as a function of (*left to right*): redshift, primary mass, and number of galaxies contained in the GW localization volume.

3.5 Results for O4 and O5-like scenarios

In this section, we report the results of the analyses of the *O4*- and *O5-like* configurations. For both detector networks, we consider three distinct analysis setups:

- Full (z_{spec}): Full analysis using the parent catalog of 1.6 million galaxies derived from MICEv2 adopting spectroscopic redshift uncertainties as in Eq. 3.86.
- Full (z_{phot}): Full analysis as above, adopting broader (photometric) redshift uncertainties as in Eq. 3.86.
- Spectral: No galaxy catalog information is incorporated and instead galaxies are assumed to be uniformly distributed in comoving volume. The constraints are based only on the assumption of the source-frame mass distribution.

In this way, we obtain a total of six configurations. We adopt wide uniform priors for all the hyperparameters (Table 3.4.2). The posterior distribution is sampled with affine-invariant Markov Chain Monte Carlo sampler `emcee` (Foreman-Mackey et al., 2013) and the convergence is assessed ensuring that the number of samples is at least 50 larger than the integrated autocorrelation time for all the parameters. Results are reported using the median statistic with symmetric 68.3% credible levels.

The final constraints on all individual parameters are reported in Table 3.5. To facilitate the comparison among different configurations, we report the percentage uncertainty on each parameter in Table 3.6 and in Fig. 3.19. Finally, in Appendix A.3 we show examples of the full corner plots. For all the six configurations, with CHIMERA we recover the fiducial values with a typical deviation of 0.2σ , with fluctuations that can be ascribed to the specific realizations of the data sets.

3.5.1 Full O4- and O5-like scenarios

We start by discussing the results for the best-case scenario, consisting of a complete galaxy catalog with spectroscopic redshift measurements. The marginal posterior distributions for the selection of parameters $\{H_0, \alpha, \mu_g, \sigma_g, \gamma\}$ are shown in Fig. 3.20.

We find that in one year of *O4-like* configuration, the LVK interferometers are able to constrain H_0 with 7% uncertainty (at the 1σ level) from BBH in a combined cosmological and astrophysical population inference. This is a remarkable improvement with respect to the current constraints, as the analysis of the 42 BBHs at $\text{SNR} > 12$ observed so far yields a 46% measurement of H_0 (Mastrogiovanni et al., 2023). To arbitrate the Hubble tension, however, percent-level measurements would be required. If we assume that the uncertainty on H_0 scales as $1/\sqrt{N}$, in the O4 configuration it is not possible to reach 1% in the planned schedule of about two years.

In contrast, with the *O5-like* configuration it is possible to reach 1.1% uncertainty in about one year of observation just considering 100 high SNR ($\text{SNR} > 25$) GW events. We stress that this is a best-case scenario, relying on having a complete catalog of potential hosts up to $z \sim 1$ and on a network comprising LIGO+Virgo+Kagra+India. In general, the completeness varies among different galaxy surveys and galaxy types. For example, with Euclid it would change between the photometric or the spectroscopic survey mode and the north

Table 3.5. Median and 68% CL interval of the hyperparameters resulting from spectral only and dark+spectral sirens analyses for both O4-like and O5-like data set and for both spectroscopic and photometric errors on the galaxies redshift.

λ_i	O4-like (100 events with SNR > 12)			O5-like (100 events with SNR > 25)		
	Spectral	Full (z_{phot})	Full (z_{spec})	Spectral	Full (z_{phot})	Full (z_{spec})
H_0	64^{+32}_{-23}	76^{+16}_{-12}	$75.3^{+5.2}_{-4.9}$	55^{+20}_{-16}	$73.3^{+7.2}_{-6.3}$	$70.24^{+0.76}_{-0.81}$
α	$3.60^{+0.37}_{-0.32}$	$3.68^{+0.40}_{-0.36}$	$3.72^{+0.33}_{-0.32}$	$3.24^{+0.26}_{-0.26}$	$3.28^{+0.26}_{-0.29}$	$3.27^{+0.29}_{-0.27}$
β	$2.51^{+0.78}_{-0.72}$	$2.43^{+0.85}_{-0.72}$	$2.55^{+0.74}_{-0.74}$	$1.90^{+0.56}_{-0.52}$	$1.96^{+0.58}_{-0.57}$	$1.81^{+0.57}_{-0.52}$
δ_m	$7.1^{+1.9}_{-3.2}$	$7.0^{+2.1}_{-3.1}$	$7.3^{+1.8}_{-2.5}$	$3.1^{+2.5}_{-1.8}$	$2.8^{+2.3}_{-1.6}$	$3.5^{+2.3}_{-1.7}$
m_{low}	$4.35^{+0.75}_{-0.66}$	$4.28^{+0.73}_{-0.69}$	$4.25^{+0.57}_{-0.51}$	$5.96^{+0.73}_{-1.06}$	$5.82^{+0.69}_{-1.05}$	$5.58^{+0.70}_{-0.97}$
m_{high}	102^{+52}_{-21}	95^{+47}_{-17}	94^{+59}_{-15}	102^{+23}_{-13}	$91.5^{+27.8}_{-7.2}$	$93.2^{+35.3}_{-7.1}$
μ_g	$34.2^{+3.5}_{-4.8}$	$32.9^{+2.4}_{-3.5}$	$32.8^{+1.7}_{-2.7}$	36.5 ± 2.7	$34.1^{+1.3}_{-1.4}$	$34.4^{+1.0}_{-1.1}$
σ_g	$5.3^{+2.6}_{-1.8}$	$5.0^{+2.0}_{-1.6}$	$4.9^{+2.2}_{-1.4}$	$4.47^{+1.21}_{-0.84}$	$4.09^{+0.90}_{-0.65}$	$4.24^{+0.82}_{-0.71}$
λ_g	$0.03^{+0.03}_{-0.01}$	$0.03^{+0.03}_{-0.01}$	$0.03^{+0.02}_{-0.01}$	$0.07^{+0.04}_{-0.03}$	$0.07^{+0.04}_{-0.03}$	$0.07^{+0.04}_{-0.03}$
γ	$1.57^{+1.26}_{-0.93}$	$1.79^{+1.09}_{-0.85}$	$1.77^{+1.38}_{-0.92}$	$2.2^{+1.9}_{-1.1}$	$2.32^{+1.63}_{-0.94}$	$2.26^{+1.26}_{-0.95}$

Table 3.6. Same as in Table 3.5, with the percentage uncertainties.

λ_i	O4-like (100 events with SNR > 12)			O5-like (100 events with SNR > 25)		
	Spectral	Full (z_{phot})	Full (z_{spec})	Spectral	Full (z_{phot})	Full (z_{spec})
H_0	43 %	18 %	7 %	32 %	9 %	1 %
α	10 %	10 %	9 %	8 %	8 %	9 %
β	29 %	32 %	29 %	28 %	29 %	30 %
δ_m	34 %	37 %	29 %	69 %	70 %	57 %
m_{low}	16 %	16 %	13 %	15 %	15 %	15 %
m_{high}	35 %	33 %	38 %	18 %	19 %	23 %
μ_g	12 %	9 %	7 %	7 %	4 %	3 %
σ_g	42 %	35 %	38 %	23 %	19 %	18 %
λ_g	74 %	82 %	62 %	50 %	46 %	49 %
γ	67 %	54 %	62 %	68 %	56 %	50 %

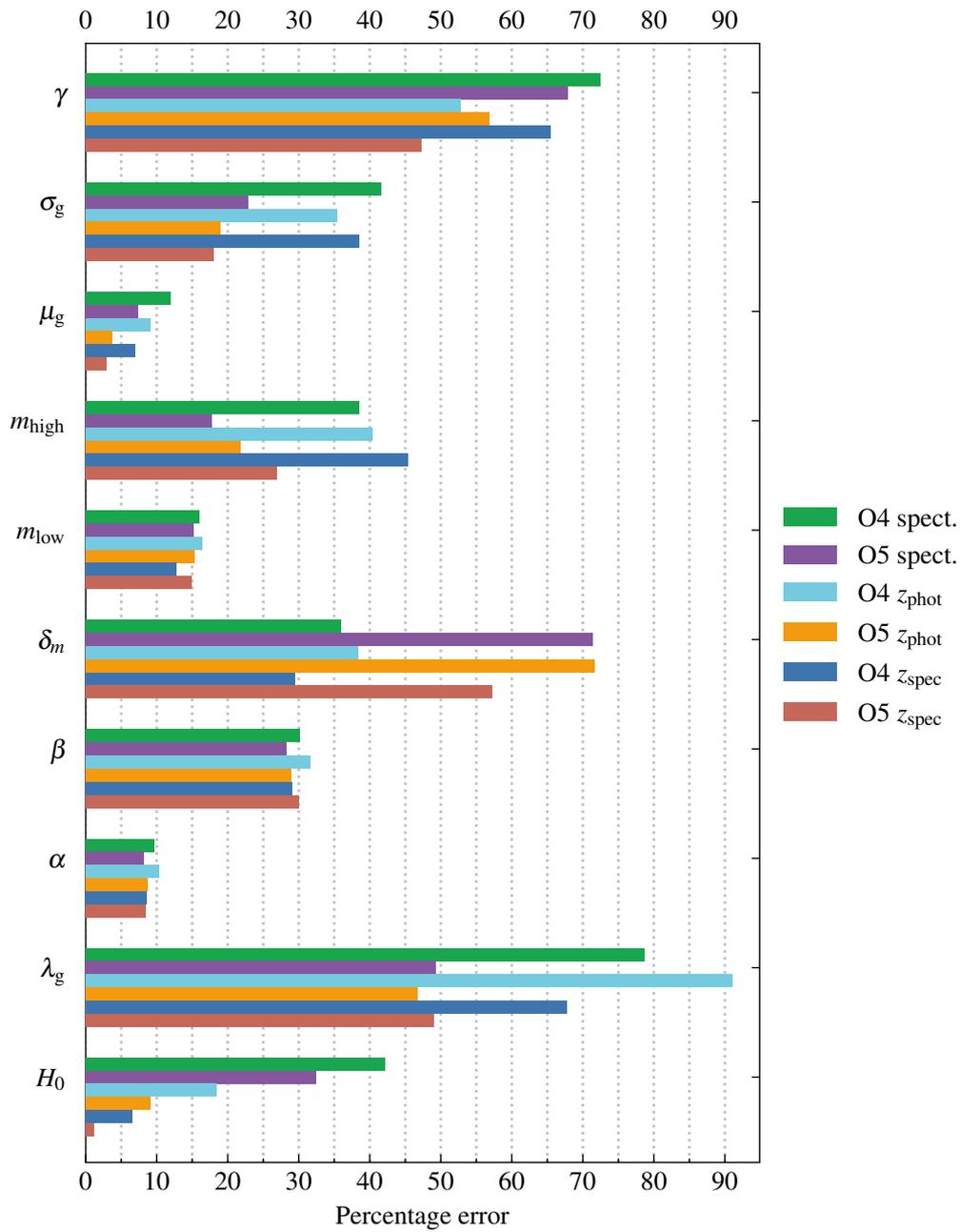


FIGURE 3.19: Comparison of percentage errors on cosmological and population parameters obtained in O4- and O5-like configurations for the spectral, Full (z_{phot}), and Full (z_{spec}) analyses.

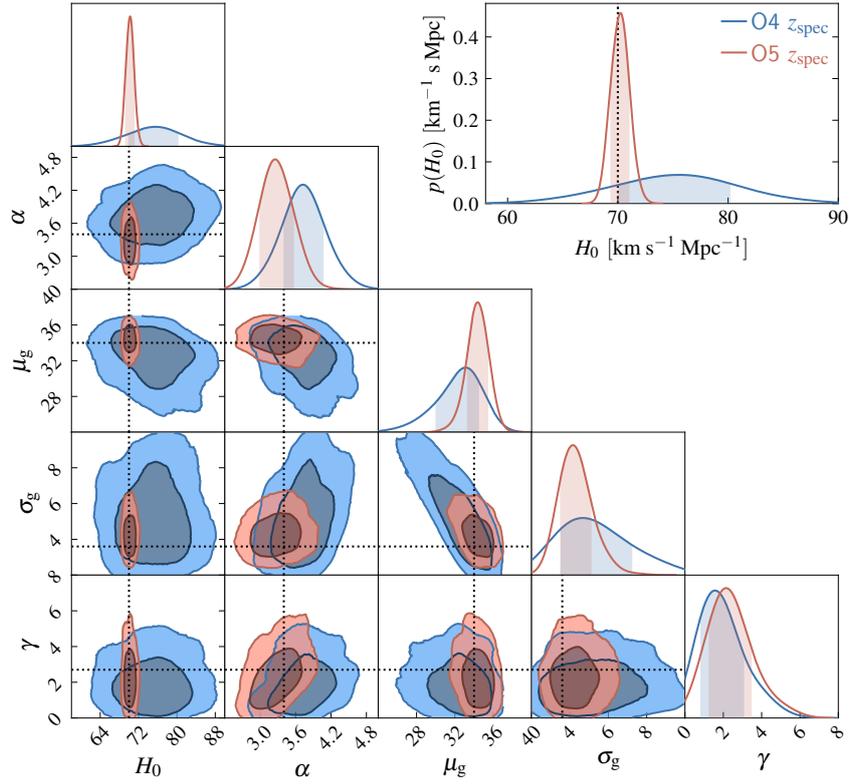


FIGURE 3.20: Joint cosmological and astrophysical constraints from the full standard sirens analysis of 100 BBHs in the *O4-like* (blue) and *O5-like* (red) configurations. Here we show, for illustrative purposes, the most relevant parameters, namely $\lambda_c = \{H_0\}$, the BBH mass function parameters $\lambda_m = \{\mu_g, \sigma_g\}$, and the rate evolution parameter $\lambda_z = \{\gamma\}$, while in App. A.3 we show for completeness the entire distribution of parameters. The contours represent the 1 and 2 σ confidence levels. The dotted lines indicate the fiducial values adopted.

or south direction, requiring a more detailed assessment in a future study. Moreover, even if the astrophysical parameters are not fixed, the results still depend on the assumed shape of the mass function and its independence on redshift. On the other side, this analysis is based only on the BBH population, further improvements can be obtained including NSBH and BNS events and their potential electromagnetic counterparts.

We now move to population hyperparameters. Even if in these cases the cosmology would be typically fixed (e.g., Abbott et al., 2023a), it is interesting to study how well the fiducial models are recovered and potential degeneracies with cosmology hyperparameters. In Fig. 3.21 we show the reconstructed primary mass distribution. For both O4 and O5 scenarios, the Gaussian peak at $34 M_\odot$ is clearly visible and its mean value μ_g is recovered with a precision of 7% and 3%, respectively. The second best-constrained mass parameter is the slope α of the primary BBH mass distribution that is recovered with fractional uncertainties of 15% and 13%, respectively. Overall, for the mass function parameters there will be a gain of a factor 2 to 4 in the determination of mass function parameters, i.e. $\sigma_{\mu_g}/\mu_g \sim 13\%$ and $\sigma_\alpha/\alpha \sim 11\%$ (see Mastrogiovanni et al., 2023). Finally, we note that the constraints on the

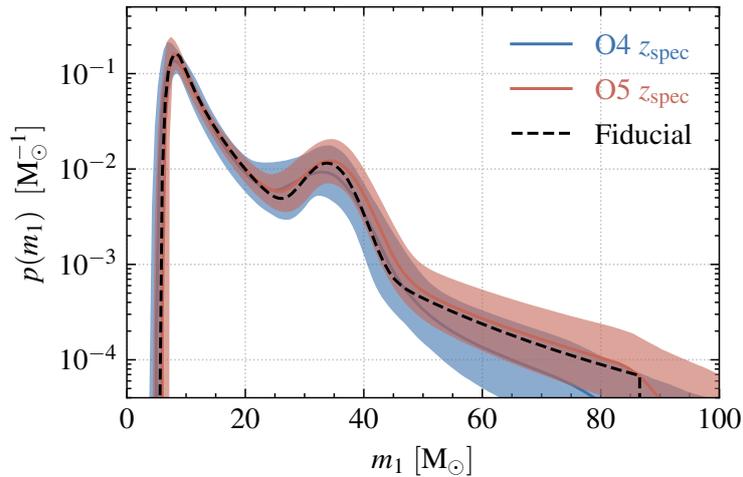


FIGURE 3.21: Posterior on the Power Law + Peak BBH primary mass function from the full standard analysis of 100 BBH in the *O4-like* (blue) and *O5-like* (red) configurations. The colored lines and bands represent the median and 90% credible level. The black dashed line is the mass function evaluated at the fiducial hyperparameters.

rate parameter γ remain essentially unchanged between O4- and O5-like configuration. This can be explained by the higher SNR threshold adopted in the O5 catalog, resulting in GW events that map a redshift range comparable to that of O4.

In conclusion, we recall that our results are based on the full astrophysical and cosmological analysis of the best 100 GW events detectable in 1 yr for each configuration. Population studies typically benefit from the inclusion of all confident events detected and are carried out by fixing the cosmological parameters (e.g., [Abbott et al., 2023a](#)). In this sense, our population constraints should not be taken as representative of the overall performance of O4 and O5.

3.5.2 Spectroscopic vs. photometric galaxy catalog

While obtaining a complete spectroscopic catalog poses challenges and awaits future facilities, ongoing surveys such as DES and Euclid are already building extensive photometric galaxy catalogs. The full z_{phot} and z_{spec} configurations for both O4 and O5 are compared in Fig. 3.22.

When using photometric redshifts, the constraint on H_0 for O4 is notably less accurate, with a measurement uncertainty that is three times greater ($\sigma_{H_0}/H_0 \approx 18\%$) compared to the spectroscopic approach.

With photometric redshifts, the constraint on H_0 for O4 is notably less accurate, with a measurement uncertainty that is three times greater ($\sigma_{H_0}/H_0 \approx 18\%$) compared to the spectroscopic approach. In the case of O5 this factor increases to 9 ($\sigma_{H_0}/H_0 \approx 9\%$). Interestingly, this shows that from 100 O5 events at $S/N > 25$ it is not possible to achieve percent-level precision on H_0 using a photometric catalog, even under the assumption of completeness. Of course, one may consider lowering the S/N threshold to include more events; however, such events would also have much larger localization volumes, and thus a

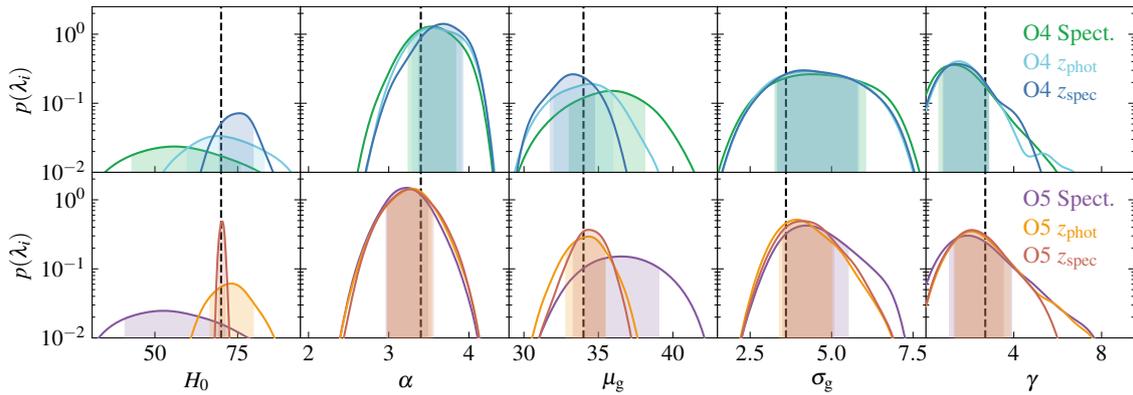


FIGURE 3.22: Marginalized posterior distributions of a selection of hyperparameters in the *O4-like* (upper panels) and *O5-like* (lower panels) detector networks. The analysis is performed with CHIMERA in three setups, namely: spectral-only (Spect.) and spectral with the inclusion of a photometric (z_{phot}) or spectroscopic (z_{spec}) galaxy catalog.

large number of potential hosts. This would likely limit their additional constraining power. Some information can be still retrieved for the mass distribution, whose reconstruction benefits from a larger sample; however, in the next section we will show that the constraints obtained in the absence of a galaxy catalog with our sample are only marginally worse than what can be obtained from a sample with a much lower S/N cut. This seems to suggest that it would be difficult to reach percent-level accuracy.

Another interesting result concerns the comparison of the Full (z_{phot}) *O5-like* and Full (z_{spec}) *O4-like* configurations (see Fig. 3.19). We find that considering a galaxy catalog with spectroscopic redshift uncertainties in the *O4-like* scenario, we are able to achieve a more precise constraint on H_0 compared to having a larger LVK detector network at O5 design sensitivity with a photometric galaxy catalog. This occurs despite the factor 10 improvement in the GW localization volume (see Fig. 3.19). Overall, these findings underline the importance of mapping the GW localization volume - at least for well-localized events - with dedicated spectroscopic surveys.

3.5.3 Spectral-only analysis

In the absence of the galaxy catalog information (spectral sirens case), the cosmological constraints are determined by the capability of the source-frame mass function to break the mass-redshift degeneracy. For this reason, the shape of the mass function plays an important role: more features would provide more stringent constraints (e.g., Ezquiaga & Holz, 2022), while a potential evolution in redshift (e.g., Rinaldi et al., 2023) may strongly affect the cosmological inference.

In our spectral siren analysis, the Hubble parameter is recovered at $\sigma_{H_0}/H_0 \approx 43\%$ in *O4-like* and $\sigma_{H_0}/H_0 \approx 32\%$ in *O5-like* configuration. Our results are in good agreement with those of Mancarella et al. (2022b); Leyde et al. (2022). In particular, Leyde et al. (2022) find 38% (24%) uncertainty on H_0 using $\text{SNR} > 12$ spectral siren events (obtaining a total of 87 and 423 events for O4 and O5, respectively). When comparing with our results for O5,

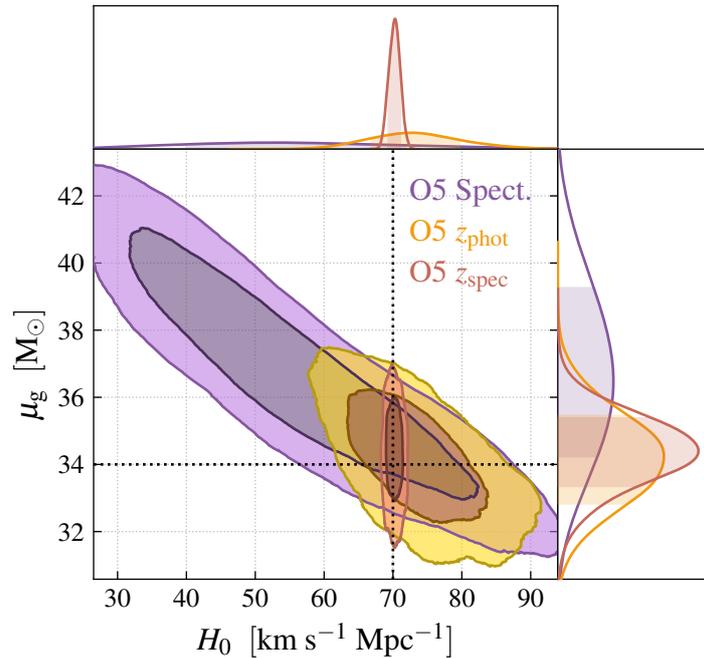


FIGURE 3.23: Constraints on the $H_0 - \mu_g$ plane for the O5-like scenario in case of spectral, Full (z_{phot}), Full (z_{spec}) analyses with CHIMERA.

we must consider that we applied a higher SNR > 25 threshold, resulting in a factor of 4 difference in the number of detected events.

In general, we conclude that the analysis of 100 well-localized GW events with a complete galaxy catalog with spectroscopic measurements, can provide better constraints with respect to a pure spectral siren analysis under the assumed mass functions. This result holds true also when comparing with the more optimistic results for five years of O5 spectral sirens by (Farr et al., 2019), providing $\sigma_{H_0}/H_0 \approx 3\%$. A more detailed assessment of the optimal analysis settings to understand the number of events required for the spectral-only method to outperform the catalog analysis will be carried out in future work. However, this analysis suggests that the synergies between GW observational campaigns and galaxy surveys will play a crucial role in gravitational-wave cosmology.

Finally, there is a clear degeneracy between H_0 and μ_g , resulting in uncertainties on μ_g that are larger by a factor of 2 compared to the spectroscopic catalog case (see Fig. 3.22). This effect is related to the fact that μ_g is the mass feature that mostly acts as a standard ruler, breaking the mass-redshift degeneracy and thus enabling cosmological constraints. Smaller, but still interesting $1.2 - 1.3\times$ precision improvements are also seen for δ_m , σ_g , and γ . As these results do not depend on the quality of the GW data, these enhanced constraints likely emerge from the breaking of correlations between parameters thanks to the improved H_0 precision. A quantification of these correlations is provided in section 3.5.4. Eventually, this effect is negligible, for the remaining astrophysical population parameters.

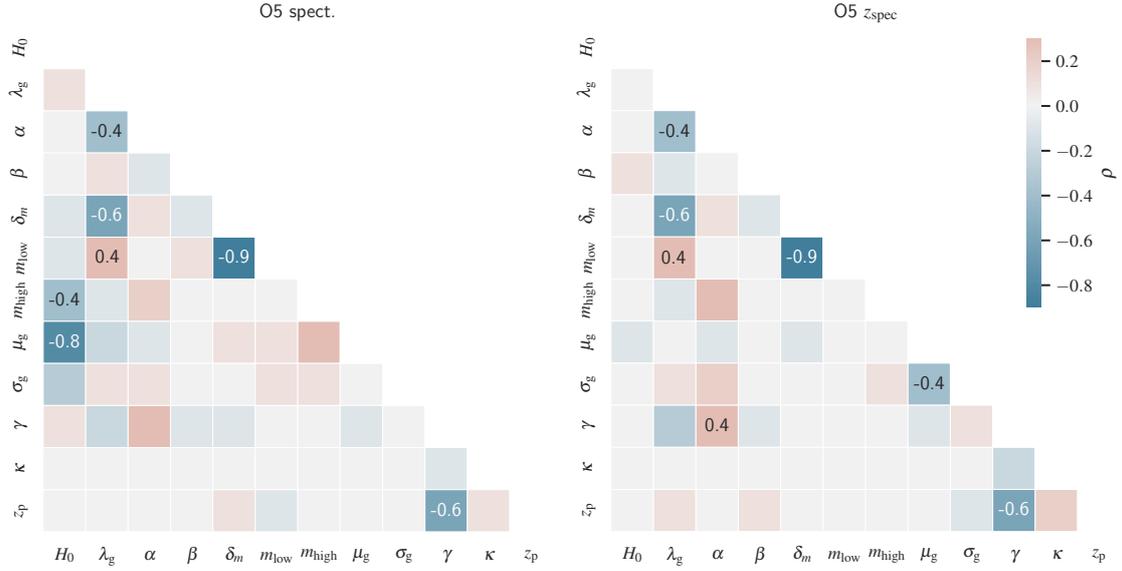


FIGURE 3.24: Correlations among the population parameters evaluated with the Spearman’s coefficient. Coefficients’ values are shown only for $|\rho| \geq 0.4$.

3.5.4 Hyperparameters correlations

To evaluate the correlations between couples of parameters, we compute Spearman’s rank correlation coefficient from the MCMC chains. This provides a quantification of the degeneracies among the posteriors by assessing the extent of their monotonic relationships. It is important to consider that these results strongly depend on the assumption of the population functions. An illustration of the correlations for the spectral and Full (z_{spec}) cases is shown in Fig. 3.24.

A relatively strong anti-correlation ($\rho \sim -0.7$) is observed between the position of the mass peak and H_0 in the Spectral and Full (z_{phot}) analyses for both O4 and O5 configurations. In fact, a higher inferred H_0 would move the events at higher z , requiring smaller source-frame masses to reproduce the observed data (see Eq. 3.51), hence a smaller μ_g value. This anti-correlation is observed also in the latest GWTC-3 analyses (Abbott et al., 2023c). Figure 3.23 shows the constraints on the $H_0 - \mu_g$ plane in the O5-like scenario. Having a complete spectroscopic galaxy catalog allows us to break the $H_0 - \mu_g$ degeneracy by constraining H_0 with high precision. No other population parameters show meaningful correlations with H_0 in our data.

Other strong correlations are seen within the mass function parameters. In particular, by increasing the smoothing parameter δ_m it is possible to reproduce a decrease of m_{low} , or a decrease of the relative contribution of the Gaussian peak λ_g . Their degeneracies are not solved when including the galaxy catalog information, instead they require higher-quality GW data. For example the correlation coefficient for $\delta_m - m_{\text{low}}$ decreases from $\rho = -0.9$ for O4 to $\rho = -0.7$ for O5.

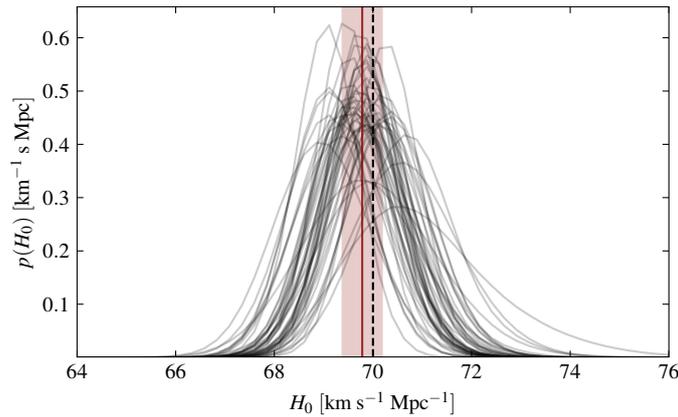


FIGURE 3.25: Posterior distributions for H_0 from 50 realizations of 100 independent O5-like events using CHIMERA. The dashed line indicates the fiducial value at which the catalog is generated. The red line and band represent the mean and the standard deviation of the recovered values.

3.5.5 Assessing potential biases

To show more quantitatively that the statistical method we developed does not present any significant bias, we resample multiple independent GW data sets from the catalog presented in Section 3.4.2. In particular, we focus on the configuration that provides the most precise constraints on H_0 , i.e., O5-like configuration in combination with spectroscopic galaxy catalog. Firstly, we perform one-dimensional tests on H_0 using independent sets of events extracted from the same GW catalog, then in we re-generate the GW catalog assuming different cosmologies and study our ability to recover H_0 independently of its value.

Independent realizations from the same GW catalog

Starting from the O5-like catalog of 4×10^5 GW events, we apply the same $\text{SNR} > 25$ cut as done in our main analysis, obtaining a catalog of about 5000 events. From this, we extract 50 independent sets of 100 events each and perform a one-dimensional analysis on H_0 fixing all the other parameters to their fiducial values as in Table 3.4.2. We decided to proceed this way because otherwise it would have been computationally unfeasible to run 50 independent MCMC chains. While we notice that this is an approximation of the true posterior, since it represents a slice of the full posterior along the direction of a single hyperparameter, we verified with the example for which we obtained a MCMC that the results are representative of the true distribution. The results are shown in Fig. 3.25.

First of all, every independent realization provides a constraint on H_0 which is fully compatible with the fiducial value adopted within the statistical error. We verify that the constraining power of different realizations can be explained by fluctuations in the number of very well-localized events, which play a major role in the determination of the H_0 constraint. These regions typically encompass tens of galaxies and have a significant contribution to the final constraints.

Overall, we find that all H_0 determinations is well compatible with the fiducial value, with best-fit values showing a deviation of only $\Delta H_0 = 0.2 \pm 0.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ from

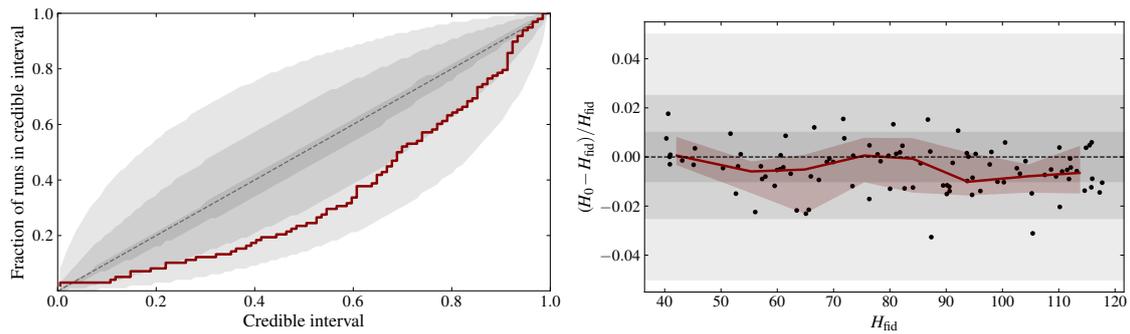


FIGURE 3.26: P–P plot showing the results of the dark+spectral analysis in **CHIMERA** performed on 100 O5-like realizations using the spectroscopic galaxy catalog. The shaded regions show 1%, 0.5%, 0.1% displacement intervals in decreasing order of opacity. Relative difference from the fiducial value as a function of H_0 .

the fiducial value. This would introduce a negligible bias of less than 0.3% with respect to our current constraints. To better understand this effect, in the next section we extend the statistical analysis using a more robust Bayesian framework.

Independent realizations varying the true H_0

A common tool to validate Bayesian analyses is by studying posterior quantiles from simulated data sets. Here, we use the percentile–percentile (P–P) plot (Cook et al., 2006) to test whether the correct proportion of injected values is recovered within a given confidence interval for a specific prior distribution. These tests are particularly useful when dealing with Gaussian likelihoods, as the fraction of events within a given confidence interval should be uniformly distributed. For an ideal, unbiased case for $x\%$ of the realizations, the true population parameters should lie within the $x\%$ confidence interval (CI). This methodology has been recently adopted in Gair et al. (2023) to extensively study possible errors related to data generation or mismodelling in dark sirens analyses. Differently from the test performed in the previous section, for a more comprehensive Bayesian analysis, here we generate 50 independent realizations of 100 GW events, but varying the fiducial H_0 value (hereafter, H_{fid}).

To construct the P–P plot, we first compute the cumulative distribution function (CDF) of the H_0 posterior for the i -th realization. Then we evaluate and record the CDF at the fiducial H_{fid} , $\text{CDF}(H_{\text{fid}})_i$, obtaining a description of how much posterior falls at or below H_{fid} . This process is iterated over all the i -th realizations. In the end, we plot the cumulative sum of the distribution of $\{\text{CDF}(H_{\text{fid}})_i\}$, representing the fraction of runs that constrain H_0 within a CI as a function of CI. For an ideal case, this line would lie along the diagonal. In parallel, to understand the effects of possible systematics, we simulate the P–P plot of 100 events characterized by an increasing offset of 1%, 0.5%, 0.1% from H_{fid} . This allows us to illustrate three “displacement intervals” on the P–P plot. Figure 3.26 shows the P–P plot obtained with these prescriptions.

We find that H_0 is consistently recovered in our Bayesian pipeline within an offset of about 0.45%. This effect may also be related to the fact that our posterior slightly deviate from a perfect Gaussian (see Gair & Moore, 2015).

In conclusion, from the analysis of multiple realizations of the O5-like catalog varying the true H_0 , we find that the recovered values are consistent within $< 0.5\%$, demonstrating the robustness of our approach for precision gravitational wave cosmology. Analysis of GW data from future facilities may require a deeper assessment of this effect, which may become non-negligible with more precise constraints. In particular, the efforts may include the development of semi-analytical approaches to perturb the GW posterior around their true values, providing a more robust statistical approach. Another possibility in this direction, which remains computationally expensive for large catalogs, would be to perform full MCMC parameter estimations of simulated GW signals.

3.6 Summary

In this Thesis, we extend the standard sirens method to jointly study cosmological and astrophysical population parameters including the information of galaxy catalogs. In this new framework, we provide forecasts on the precision achievable on H_0 , mass function, and rate evolution of the astrophysical population, for current and upcoming gravitational wave facilities and galaxy surveys.

The main results are outlined below and illustrated in Fig. 3.27.

1. We extend the hierarchical Bayesian framework, providing an analytical formulation for the likelihood to obtain joint constraints to cosmological and gravitational-wave population parameters from standard sirens and galaxy catalogs. This approach allows us to take into account all the uncertainties of the population model while including information from the distribution of potential hosts from a galaxy catalog.
2. We develop and release **CHIMERA**¹⁰, a novel Python code for the joint analysis of GW transient catalogs and galaxy catalogs. The code is designed to be accurate for different scenarios, from the spectral siren (population inference), to the dark siren (galaxy catalog) and the bright siren (redshift precisely measured after the identification of an electromagnetic counterpart) methods. This provides a crucial upgrade with respect to previously available software, enabling a joint analysis of cosmological and astrophysical population parameters using galaxy catalogs. Key features include:
 - *3D GW kernel.* The entire pipeline operates in redshift space. Initially, the GW posterior distribution is smoothed using the kernel density estimate (KDE) approach in a 3D space defined by the parameters (d_L , RA, Dec), then are converted in z during the inference under a specific cosmological assumption.
 - *Pixelization.* We implement the pixelization approach for the likelihood as presented in (Gray et al., 2020), including an extension for adaptive pixelization, and for the completeness computation following (Finke et al., 2021). This leads to a factor 10–100× time improvement in the calculation of the galaxy distribution, hence the likelihood, for the analyses carried out in this work.
 - *Efficiency improvements.* In view of future GW facilities and large galaxy surveys, we further optimize the likelihood evaluation in several aspects, including the optimization of cosmological libraries.
 - *Selection effects.* We implement an efficient approach based on the interpolation of galaxy catalogs to evaluate the selection effect term using the injection approach.
 - *Flexibility.* We design the code to facilitate extensive analyses with different configurations, such as varying population models, data thresholds, pixelization and integration parameters, assumptions regarding galaxy catalogs, and completeness settings.
3. We analyze the best-localized GW events, GW170817 and GW190814, using LIGO-Virgo posteriors and the GLADE+ galaxy catalog (Dály et al., 2022). When considering all the galaxies included in the GW localization volume as potential hosts

¹⁰Available at <https://chimera-gw.readthedocs.io/>

(dark sirens method), we obtain $H_0 = 73_{-22}^{+58}$ km/s/Mpc (68% C.L.). When restricting our analysis to the identified host NGC 4993 (bright sirens method), we get $H_0 = 69_{-8}^{+15}$ km/s/Mpc. Finally, from the dark siren analysis of GW19814 by fixing the astrophysical population parameters, we obtain $H_0 = 64_{-28}^{+45}$ km/s/Mpc. These results are in agreement with previous measurements found in the literature (Fishbach et al., 2019; Abbott et al., 2023c; Mastrogiovanni et al., 2023; Gray et al., 2023).

4. We generate two catalogs of BBH events for detector networks representative of the LIGO-Virgo-KAGRA O4 and O5 observing runs. We start from a parent catalog of galaxies extracted from the MICEv2 mock galaxy catalog, which we select with a cut in stellar mass $\log M_*/M_\odot > 10.5$. From this catalog, we generate BBH events with the **GWFAST** code, considering the 100 best events that can be detected over approximately 1 yr of observation, namely, $S/N > 12$ BBHs for the *O4-like* configuration and 100 $S/N > 25$ BBHs for the *O5-like* configuration. We then associate the galaxy redshift catalog with two different uncertainties on the redshift, representative of the case of having a catalog of photometric and spectroscopic redshifts.
5. We find that in the best-case scenario of having a complete spectroscopic galaxy catalog, the Hubble constant H_0 can be constrained to 7% (O4) and 1% (O5) at 68% confidence level in about a year of observations from a joint cosmological and astrophysical standard siren analysis. This remarkable precision demonstrates the potential of joint standard sirens + galaxy catalog analyses to arbitrate the Hubble tension with a novel independent probe.
6. In the case of a photometric galaxy catalog, the constraints on H_0 are notably weaker. For O4, the uncertainty increases by a factor of ~ 3 for O4 and ~ 9 for O5 compared to the spectroscopic case, respectively. In general, we find that joint analysis can provide better constraints on H_0 compared to a spectral-only siren analysis, even when considering optimistic scenarios of 5 years of O5 observations. This highlights the importance of spectroscopic follow-up campaigns to map the galaxy distribution within the GW localization volumes, at least for the best-localized events.
7. Interestingly, we find that the *O4-like* configuration with a spectroscopic catalog provides a more precise measurement of H_0 than the *O5-like* configuration with photometric redshifts. This points to the importance of having spectroscopic redshift measurements for GW cosmology, in the absence of which the potential of a factor ~ 10 improvement in the localization capabilities between *O4-* and *O5-like* GW detector networks could be completely lost.
8. We study constraints on the astrophysical population parameters, finding that a strong correlation between H_0 and the mass peak μ_g , which acts as a standard property that enables cosmological constraints. As a consequence, more precise H_0 measurements obtained using spectroscopic catalogs improve the constraints on μ_g by factor of ~ 2 with respect to spectral-only analyses. This is not the case for the other astrophysical population parameters, which would be better served by additional and/or higher quality GW data rather than more precise galaxy redshift measurements.

9. Finally, we carried out tests to validate that our Bayesian methodology does not introduce significant biases. Analyzing multiple realizations of O5-like data varying the true H_0 , we find that the recovered values are consistent within $< 0.5\%$, demonstrating the robustness of our approach for precision gravitational wave cosmology.

Figure 3.27 shows the main contributions of this Thesis to unveil the expansion history of the Universe with gravitational waves. GW170817 is the GW event that better constrain H_0 , providing already to date an independent local hook to $H(z)$. With the LIGO-Virgo-KAGRA-LIGO India network at the design sensitivity and a complete spectroscopic survey of potential host galaxies, it will be possible to obtain percent-level constraints on H_0 with 100 BBH events, potentially addressing the current tension.

In conclusion, this Thesis demonstrates that joint analyses of gravitational wave catalogs and galaxy surveys will be crucial to harness the full cosmological potential of the standard sirens method. The development of the hierarchical Bayesian framework, CHIMERA code, and mock data challenges pave the way for robust cosmological studies in the upcoming era of precision gravitational wave cosmology.

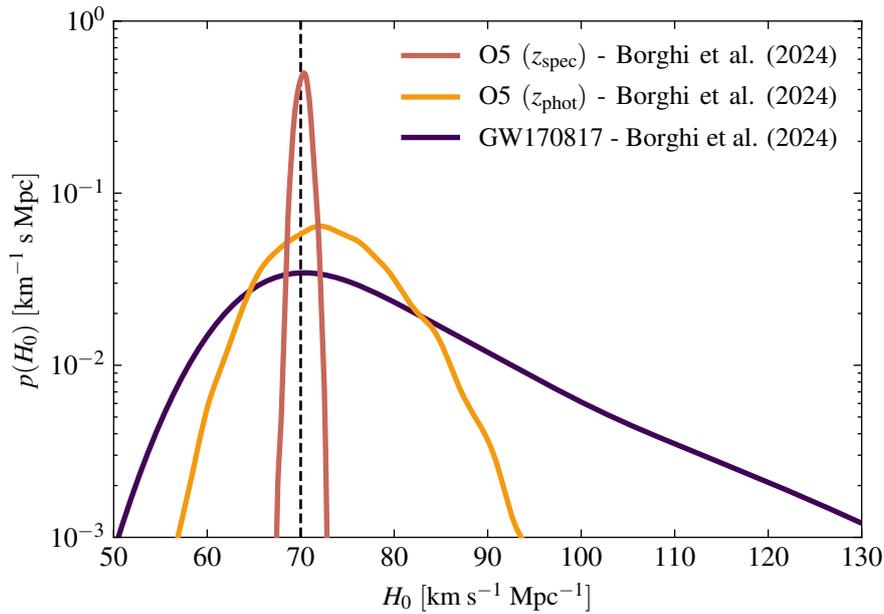


FIGURE 3.27: Current and future constraints on H_0 presented in this work. In particular, we show the most constraining observed event GW170817 (Abbott et al., 2017a) analyzed as a bright siren and forecasts for 100 BBH events from O5-like GW network configuration including a photometric (z_{phot}) or spectroscopic (z_{spec}) galaxy catalog. The vertical dashed line shows the fiducial value used for the simulation.

Conclusions

The discovery of the accelerated expansion of the Universe (Riess et al., 1998; Perlmutter et al., 1999) has been a major revolution in modern cosmology. This phenomenon is currently explained by invoking a new unknown form of energy (the *dark energy*) constituting about 70% of the total matter-energy budget of the Universe. Similarly, the formation of Large-Scale Structures and the dynamics of astrophysical objects require a new form of matter that interacts very weakly with standard matter particles (the *dark matter*) constituting about 25%. Dark energy and dark matter are two fundamental assumptions of the standard Λ CDM model, which provides the simplest, yet effective description of a wealth of observations.

In the last three decades, few key cosmological probes have been proposed and extensively refined to obtain precise and accurate measurements of cosmological parameters and, in parallel, consistent theoretical modeling. These include the cosmic microwave background, type Ia supernovae, and baryonic acoustic oscillations. However, some hints in the data are suggesting that Λ CDM should be extended to fully account for emerging discrepancies among early and late cosmological epochs (Verde et al., 2019; Di Valentino et al., 2021a; Abdalla et al., 2022).

Motivated by these fundamental questions and the emerging tensions, constraining the expansion history of the Universe with new independent probes became one of the main goals in cosmology, as it could provide fundamental insights into understanding the nature of dark energy and gravity (Moresco et al., 2022).

In this Thesis, we explored two cosmological probes: cosmic chronometers (CC) and gravitational waves (GW). The reason behind this choice lies in the possibility of obtaining, with CCs, direct and cosmological-model independent measurements of $H(z)$ up to $z \sim 2$, and, with GWs, direct measurement of d_L without additional calibrators (as for SNe Ia), thus providing a direct measurement of H_0 . Moreover, both probes show promising prospects in view of upcoming galaxy surveys and next-generation GW observatories.

The main focus for both GWs & CCs has been the development of an extended methodology to maximize its scientific yield and account for the interplay of cosmological and astrophysical parameters. This allows us to derive them jointly, study possible degeneracies, and eventually minimize potential systematic effects. As a legacy value, it also enables interesting insights into galaxy evolution and properties of the compact binary populations.

Cosmology with Cosmic Chronometers. Cosmic chronometers (CC) have proven to be very promising probes to obtain direct measurements of $H(z)$ up to $z \sim 2$. This method, first introduced by Jimenez & Loeb (2002), consists of using massive and passive galaxies as tracers of the aging of the Universe under the minimal assumption of a metric. In Chapter 2,

we explored and extended the CC method, obtaining new constraints on $H(z)$ using, for the first time, the absorption features of passive galaxies (Lick indices) as age indicators and obtaining, as a by-product, very interesting insights into galaxy evolution. The main results are reviewed below.

- Extension of the CC method, consisting of constraining differential CC ages using specific spectral features calibrated to be sensitive to specific stellar population parameters. This enables synergies with galaxy evolution studies. A specific Bayesian MCMC code has been developed and models have been extended and calibrated, to derive robust and accurate constraints on stellar age, metallicity, and α -enhancement from the analysis of absorption features.
- Development and release of **PyLick**¹, a flexible Python tool for measuring absorption features in galaxy spectra, implementing several definitions of absorption features.
- Selection of a CC sample from the LEGA-C galaxies at $0.6 < z < 0.9$, maximizing the purity by combining photometric and spectroscopic criteria. Study of their observational features to exclude significant star-forming outliers, which could bias the cosmological measurement of $H(z)$.
- Study of a new spectral feature, CaII H/K, as a diagnostic of recent star formation episodes. For the first time, a detailed correlation between this index and other more widely extended star formation diagnostic has been quantitatively demonstrated, finding that the H/K is capable of reproducing color-color and specific star formation rate selections with good reliability, whilst being significantly less expensive in terms of data needed (compared to wide photometry or extended spectroscopic wavelength coverage).
- Identification of an optimized set of spectral features (Lick indices) to study the age, metallicity ($[Z/H]$), and α -element enhancement ($[\alpha/Fe]$) without cosmological assumptions. Extension of this analysis to other sets of spectral features to assess possible systematics in the final $H(z)$ measurement.
- Study of the stellar populations of LEGA-C galaxies, finding results in agreement with the mass-downsizing scenario, i.e. more massive galaxies are older and contain more metals. Overall, the selected CCs show, already at $z \sim 0.7$, $[Z/H]$ and $[\alpha/Fe]$ values compatibles with local passive galaxies. Within the redshift range of LEGA-C, the metal content of these systems does not evolve significantly. This suggests a scenario for which this population is already in place already at $z \sim 0.7$, further supporting the use of massive and passive galaxies as optimal CCs. Their age-redshift relation is binned to obtain from this population a new CC measurement.
- Study of degeneracies among H_0 and the formation timescale of CCs by fitting the age-redshift relation under the assumption of a cosmological model. This contribution is also included among the systematics in the final $H(z)$ measurement.

¹Available at: <https://pylick.readthedocs.io>

- Measurement of the Hubble parameter at intermediate redshift with the CC method in LEGA-C, obtaining $H(z = 0.75) = 98.8 \pm 33.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (68% CL) including both statistical and systematic uncertainties. This is the first measurement obtained with the Lick indices method.

On the galaxy evolution front, this work confirms the downsizing scenario whereby massive quiescent galaxies formed earlier and on shorter timescales. On the cosmological side, it provides a new $H(z)$ measurement obtained for the first time with Lick indices, with an extended study of possible systematics. Overall, this Thesis opens a new window into the formation histories of massive galaxies through spectral modeling of individual systems. The results highlight the interplay between galaxy archaeology and cosmology using cosmic chronometers.

Cosmology with Gravitational Waves. Gravitational waves (GW) produced from compact binaries can be used as *standard sirens*, as they provide the luminosity distance d_L to the source, requiring no further calibration. However, determining z from current GW data alone is not possible because of its inherent degeneracy with binary masses, therefore external information is required to provide cosmological constraints. Current methodologies include the use of features in the source-frame mass distribution (spectral sirens), redshift priors from galaxy catalogs (dark sirens), or identifying the host galaxy from an electromagnetic counterpart and measuring its redshift (bright sirens). In Chapter 3, we developed a novel methodology that encompasses together all three approaches, to obtain a comprehensive framework of cosmological and astrophysical population properties. The main results are reviewed below.

- Extension of the hierarchical Bayesian framework, providing an analytical formulation for the likelihood to obtain joint constraints of cosmological and astrophysical population parameters from standard sirens and galaxy catalogs. This approach allows us to take into account all the uncertainties of the population model while including information from the distribution of potential hosts from a galaxy catalog.
- Development and release of **CHIMERA**², a novel and Python code for the joint analysis of GW transient and galaxy catalogs. **CHIMERA** provides a crucial upgrade with respect to previously available software, enabling a joint analysis of cosmological and astrophysical population parameters including the information of galaxy catalogs.
- Analysis of the two best-localized events, GW170817 and GW190814, using LIGO-Virgo data and the GLADE+ galaxy catalog (Dálya et al., 2022). Using GW170817 as a bright siren, we constrain the Hubble constant to $H_0 = 69^{+15} - 8 \text{ km/s/Mpc}$. In parallel to this, several other tests have been carried out to verify and validate **CHIMERA**.
- Development of a framework to generate a mock GW catalog and its associated galaxy catalog, not only to validate **CHIMERA** but also to exploit its potential in determining cosmological and astrophysical population parameters. The study is performed considering two, O4- and O5-like, GW detector networks, and three analysis scenarios: spectral sirens (no galaxy catalogs), dark+spectral with photometric galaxy catalogs, and dark+spectral with spectroscopic galaxy catalogs.

²Available at: <https://chimera-gw.readthedocs.io/>

- Forecast for future measurements of H_0 , finding that the LIGO-Virgo-KAGRA-LIGO India network at design (O5) sensitivity can constrain H_0 to 1% when including a complete galaxy catalog with spectroscopic redshift uncertainties and to 9% with a photometric galaxy catalog. These results are obtained with 100 binary black hole events at $\text{SNR} > 25$, corresponding to about one year of observations. A detailed evaluation of the changes in science output for the different configurations outlined above is presented, highlighting the importance of having a spectroscopic catalog to obtain percent-level precision on H_0 without observing electromagnetic counterparts.
- Study of the main correlations among the cosmological and astrophysical population parameters, discussing the degeneracy between H_0 and a feature in the BBH mass function (the Gaussian peak), acting as a standard ruler that enables cosmological studies in the absence of a galaxy catalog.
- Validation of the pipeline from the statistical analysis of multiple realizations of GW data sets. The tests show a negligible $< 0.5\%$ bias in the recovery of H_0 , demonstrating the robustness of our approach for precision gravitational wave cosmology, but requiring further attention for future studies based on next-generation GW facilities.

Figure 4.1 summarizes the main contributions of this Thesis on the current and future constraints to the expansion history of the Universe. First, we provided an extension to both the CC and the GW framework by developing methods to simultaneously constrain cosmology and astrophysical properties. We obtained new measurements of $H(z)$ from CCs and a measurement of H_0 from GW170817. We also provided forecasts of the precision achievable on H_0 , jointly with other astrophysical population parameters, from the analysis of BBH in O4 and O5-like configurations. In the figure, we include, for completeness, other CC measurements and another determination of GW170817 from the literature. We also include CC forecasts from Moresco et al. (2022) featuring 10 $H(z)$ points at $0.2 < z < 0.8$ measured at 1% (e.g., BOSS DR16 Ahumada et al. 2020) and 5 $H(z)$ points at $1..5 < z < 2.1$ measured at 5% (e.g., Euclid Laureijs et al. 2011 and ATLAS Probe Wang et al. 2019). This shows the power of studying GWs and CCs in combination, since it allows us to reconstruct as precisely as possible the expansion history of the Universe with a few assumptions as possible. In the future, this will open a window not only in providing new hints to address cosmic tensions, but also provide fundamental data to tackle pressing questions about the nature of the dark components in our Universe.

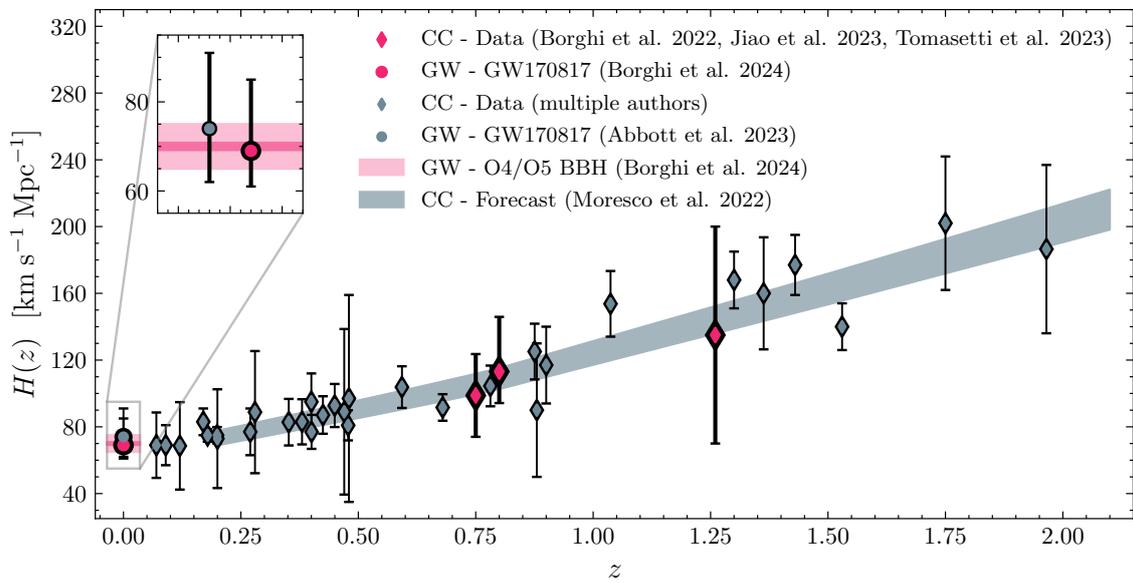


FIGURE 4.1: Current and future (next decade) constraints on $H(z)$ with cosmic chronometers (CC, diamonds) and gravitational waves (GW, points) obtained in this Thesis. The magenta colors represent the constraints obtained in this Thesis or to which I contributed, while blue ones are taken from previous studies in the literature (see [Moresco et al. 2022](#) for CC and [Abbott et al. 2023c](#) for GW measurements).

Future perspectives

Although we have made significant headway in exploring both probes, we are certainly far from finished. Exciting times are ahead for both cosmic chronometers and gravitational waves. Current and future surveys carried out with *Euclid* (Laureijs et al., 2011), DESI (DESI Collaboration et al., 2016), and Rubin (Ivezić et al., 2019), as well as proposed missions such as the ATLAS probe (Wang et al., 2019) and WST (Pasquini et al., 2016), may revolutionize our understanding of the nature of dark matter and dark energy, the cause of cosmic acceleration, the validity of general relativity on large scales, and more. In parallel, next-generation gravitational wave detectors such as the Einstein Telescope (Punturo et al., 2010), the space-based LISA (Amaro-Seoane et al., 2017), and the Cosmic Explorer (Reitze et al., 2019) will significantly increase the statistic and distance of GW events, exploring unknown territories in terms of wavelength and redshift probed. This will foster synergies with black-hole galaxy coevolution, studies of primordial black holes, and cosmological background. Together, these advancements will surely drive progress on the biggest open questions in cosmology. The subsequent paragraphs present future steps and challenges that we plan to tackle with cosmic chronometers and gravitational waves.

Cosmic chronometers.

1. *Extension of current data sets.* The precision of the cosmic chronometer measurements is dominated at the moment by the limited statistics of the sample of cosmic chronometers studied. Currently, there are no dedicated surveys specifically targeting cosmic chronometers (such as for SNe and BAO). This would be a major step forward, as these systems are typically found in a limited percentage in most of the current surveys due to the strict criteria required to minimize star-forming and young contaminants. From this point of view, the advent of DESI (DESI Collaboration et al., 2016), and the future project of ATLAS probe (Wang et al., 2019) and WST (Pasquini et al., 2016) will represent a true revolution in the field, providing unprecedented statistics.
2. *Measurement of differential ages.* The work presented in Borghi et al. (2022b) and (Jiao et al., 2023) provided the first example of a cosmic chronometer analysis carried out on the same data set with two completely different methodologies, i.e. Lick indices and full-spectral fitting. Despite \sim Gyr offsets in the absolute ages, the two methods found $H(z)$ values in very good agreement. Extending these tests to the other data sets may provide more robust cross-validations, further strengthening the method.

3. *Study of stellar population synthesis models* Stellar population modeling currently represents the main source of systematic error in the cosmic chronometer method (Moresco et al., 2022). Higher-resolution models extending over a larger wavelength range (e.g., Maraston et al., 2020) will allow for better assessment of these systematic effects and potentially find new features to break the degeneracy between the age and metal content of these galaxies, thus allowing more precise age measurements.
4. *Impact of the star formation histories.* Another major systematic effect is the assumption of the galaxy star formation histories (SFH). As demonstrated in Moresco et al. (2022) and Borghi et al. (2022b) this is based on the existing degeneracy between age and SFH duration. Studying the connection between the models and high-resolution data from local old stellar populations such as old globular clusters, may shed more light on the formation episodes of these old populations. At the same time, insights may come from galaxies at higher redshift taking advantage of James Webb Space Telescope data. Nonparametric SFH models may also help to study potential systematics related to specific SFH assumptions (e.g., Leja et al., 2019).

Gravitational waves.

1. *Completeness.* As demonstrated by the results obtained in this work, including the galaxy catalog information in standard sirens analysis will be of pivotal importance to obtain a precise measurement of H_0 . The best-case scenario presented here relies on the LIGO-Virgo-KAGRA-LIGO India network at design sensitivity, coupled with the inclusion of a complete catalog of potential hosts with spectroscopic redshift uncertainties. Together, these can provide a percent-level measurement of H_0 considering only 100 high-SNR BBH events. However, obtaining a complete census of potential hosts up to $z \sim 1$ in the next decades is a significant challenge. In the meanwhile, a more thorough assessment of the effects of incompleteness is required, and large mock galaxy simulations may be the ideal testbed to conduct these studies. Once completeness has been evaluated, there may be different ways to account for missing galaxies in the inference, including homogeneous, multiplicative, or a variance completion (Finke et al., 2021; Dalang & Baker, 2023). These advancements will be fundamental for accurate standard sirens cosmology in the absence of electromagnetic counterparts.
2. *Simulation of GW data.* In (Borghi et al., 2024), we simulated GW data using the Fisher matrix approach, which approximates the likelihood for the population parameters (Eq. 3.90) as a multivariate Gaussian distribution. While this has proven to be a good approximation for higher SNR events, including those analyzed in this work (e.g., Iacovelli et al., 2022b), a detailed assessment of the impact of this assumption has to be carried out. A possibility, which remains computationally expensive for large catalogs would be to carry out full MCMC parameter estimation of simulated GW signals. In parallel, there are possible ways forward in improving the Fisher matrix approach, such as implementing a semi-analytical method to account for the statistical noise around the true simulated parameters. We do not explicitly account for this effect; nevertheless, as we do not observe significant biases, it is likely to have only minor or no impact on our results. A more quantitative assessment is needed for a more robust statistical

analysis especially when future GW data will be considered (as the Einstein Telescope and Cosmoc Explorer) pushing uncertainties to the sub-percent level.

3. *Connection with galaxy properties.* Most of the current standard sirens analyses simply include galaxies as points, or at maximum weight them by their luminosity as tracer of their mass or star formation history. However, ongoing efforts are underway to investigate the coevolution of compact binary mergers and their host galaxies across cosmic time (e.g., Santoliquido et al., 2022). This extension may follow two parallel paths: investigating improvements in current standard siren constraints under more specific connection with galaxy properties, or using next-generation GW surveys with horizons well beyond the peak of galaxy formation to study how constraints may be improved with this additional information. In this second case, instead of generating GW catalogs from phenomenological mass and redshift distributions, it would be interesting to pursue a more realistic approach consisting of evolving binary populations inside in large cosmological simulations (e.g., Pakmor et al., 2023), providing an ideal environment to test connections with galaxy properties.
4. *Forecasts for upcoming surveys.* The analysis presented in Borghi et al. (2024) can be extended to next-generation GW detectors (e.g., ET Punturo et al. 2010) and upcoming galaxy surveys (e.g. Euclid, Laureijs et al. 2011). This would allow for the exploration of the constraints on the cosmological parameters determining the evolution of $H(z)$ at higher z with respect to current data and modified gravity theories (e.g., Belgacem et al., 2018b). This would also open up the possibility to study more explicit synergies with GW & CC in the redshift regime where these two probes overlap. The advancements of (1) and (2) would greatly improve the scientific yield of this analysis, providing a more realistic case and enabling synergies with galaxy evolution. Moreover, of particular interest for 3rd generation GW detector networks will be the very well-localized events, for which it could be possible to obtain pseudo-bright sirens constraints, by performing deep drilling galaxy surveys.
5. *Analysis of current data sets.* The improvements in the completeness modeling and connection with astrophysical properties will enable a more accurate study of the completeness assumption and possible systematics on GWTC-3 dark sirens constraints (Abbott et al., 2023c). This analysis can also benefit from some well-localized events that are being observed in the O4 run.
6. *Evolution of the mass function.* The spectral sirens approach is based on the assumption that the source-frame mass distribution can be used as a standard property to break the mass-redshift degeneracy. However, recent studies found evidence for a redshift evolution (e.g., Rinaldi et al., 2023). While by modeling this evolution it can still be *standardized*, in this scenario the cosmological constraints from the spectral siren approach could significantly worsen. It would be interesting to quantify this effect based on realistic GW catalogs discussed in task (2).

As entities living in a limited amount of space-time, we had to make our choice among many emerging observational probes. This Thesis was devoted to cosmic chronometers and gravitational waves for a simple reason, both may be used to directly measure the expansion history of the Universe ($H(z)$ and H_0 , respectively), without relying on cosmological assumptions and intermediate distance calibrators. However, it is crucial to acknowledge that a comprehensive understanding will ultimately be achieved through the synergy and combination of various cosmological probes working together to unveil the expansion history of the Universe.

Appendix

A.1 Validation of PyLick

To validate PyLick, we compare the measured indices and errors obtained on the LEGA-C spectra I_{PyLick} , with those released in the LEGA-C DR2 catalog $I_{\text{LEGA-C}}$ (Straatman et al., 2018). The comparison is performed within the passive sample to minimize differences due to emission-line subtraction performed in LEGA-C DR2 (see Straatman et al., 2018). Differences are computed as

$$\eta(I) = \frac{I_{\text{PyLick}} - I_{\text{LEGA-C}}}{|I_{\text{LEGA-C}}|}. \quad (\text{A.1})$$

The same analysis is performed for uncertainties of the indices (the notation $\eta(\sigma)$ will be used). For a fair comparison, we multiplied our uncertainties for the same coefficients applied in the LEGA-C DR2 pipeline (see Straatman et al., 2018, Tab. 3). Figure A.1 shows the results for values of the indices, sorting each index available in LEGA-C DR2 by increasing wavelength.

Overall, we find excellent agreement with existing data, with a typical $\eta(I)$ of $\sim 5 \times 10^{-5}$ ($\sim 10^{-4}$ 1σ scatter). Differences are lower for higher S/N indices G4300, Fe4383, Fe4531, and C₂4668 ($|\eta(I)| < 10^{-5}$), and outliers are mostly galaxies with lower-quality spectra. There is no trend between relative differences and the indices values, and the distribution is qualitatively Gaussian. We find good agreement also for the measured uncertainties, with a typical $\eta(\sigma)$ of $\sim 10^{-2}$ ($\sim 10^{-1}$ 1σ scatter). Larger deviations are seen for indices with values ~ 0 (Balmer indices, CN₁), for which measured uncertainties are $\sim 10\%$ lower. This difference can be due to different methods to estimate formal errors.

We note that at the current resolution ($R \sim 3500$), the method used to interpolate the spectra can introduce higher discrepancies than those observed before. In particular, using 0th order interpolation the typical scatter in $\eta(I)$ increases up to $\sim 10^{-2}$. Overall, these results confirm the reliability of PyLick to measure indices values and formal errors from observed spectra. Table A.1 reports the full table of 80 spectral features currently available in PyLick.

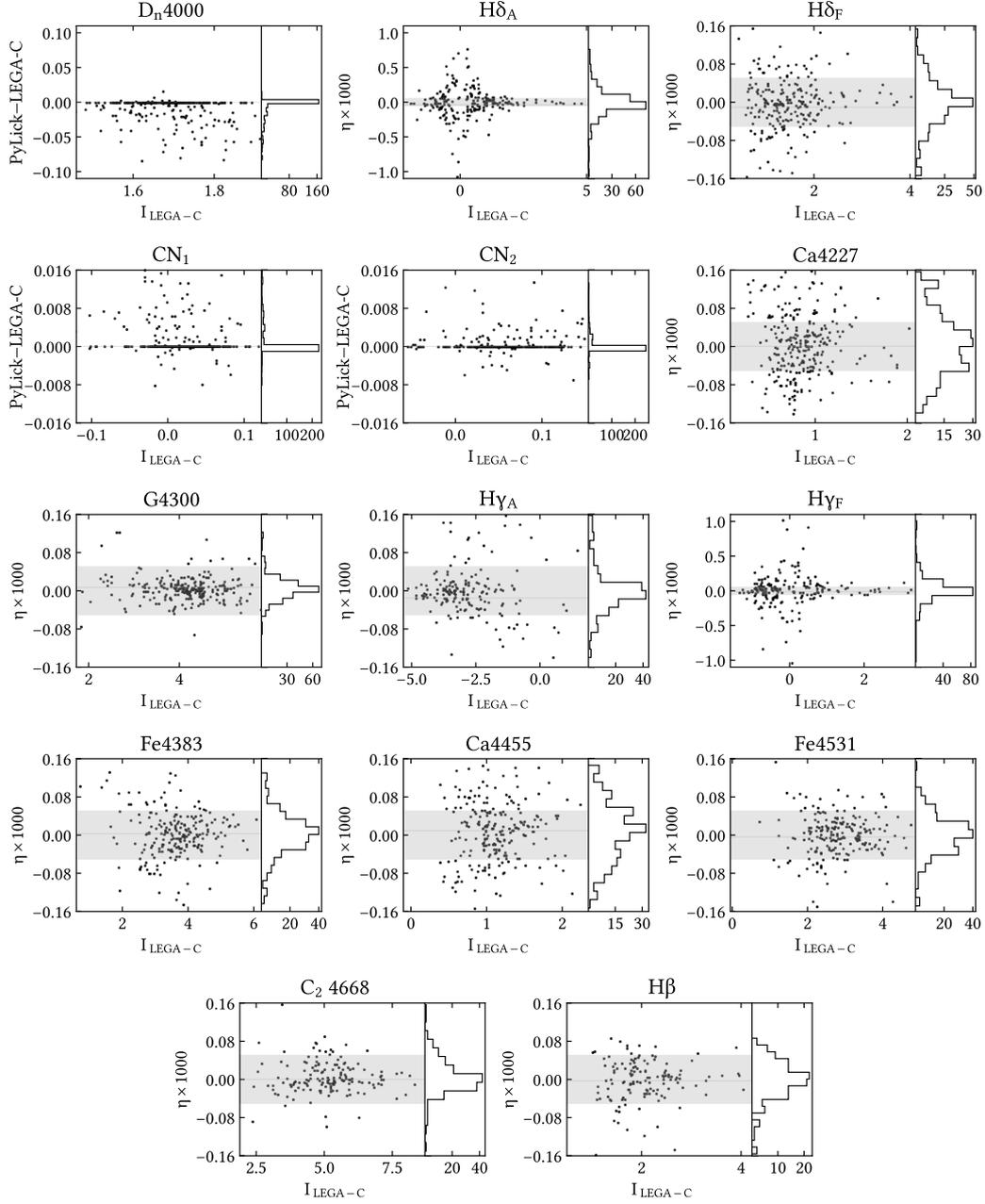


FIGURE A.1: Differences between the indices values measured with PyLick and those published in LEGA-C DR2 (Straatman et al., 2018), quantified as $\eta = (I_{\text{PyLick}} - I_{\text{LEGA-C}})/I_{\text{LEGA-C}}$, except for the D_n4000 break and the molecular indices CN₁ and CN₂, for which the absolute differences ($I_{\text{PyLick}} - I_{\text{LEGA-C}}$) are displayed. Gray bands correspond to the $\eta = \pm 5 \times 10^{-4}$ regions. Note: y -axis limits are set to the 5th–95th percentile range.

Table A.1. Table of the indices available in *PyLick* (see Section 2.2.2 for the definitions).

ID	Name	units	$(\lambda_{c1}, \lambda_{c2})$	$(\lambda_{b1}, \lambda_{b2})$	$(\lambda_{r1}, \lambda_{r2})$	Ref.
000	BL ₁₃₀₂	Å	1292.000 – 1312.000	1270.000 – 1290.000	1345.000 – 1365.000	[3]
001	SiIV	Å	1387.000 – 1407.000	1345.000 – 1365.000	1475.000 – 1495.000	[3]
002	BL ₁₄₂₅	Å	1415.000 – 1435.000	1345.000 – 1365.000	1475.000 – 1495.000	[3]
003	FeI ₄₅₃	Å	1440.000 – 1466.000	1345.000 – 1365.000	1475.000 – 1495.000	[3]
004	C _{IV} ^A	Å	1530.000 – 1550.000	1500.000 – 1520.000	1577.000 – 1597.000	[3]
005	CIV	Å	1540.000 – 1560.000	1500.000 – 1520.000	1577.000 – 1597.000	[3]
006	C _{IV} ^E	Å	1550.000 – 1570.000	1500.000 – 1520.000	1577.000 – 1597.000	[3]
007	BL ₁₆₁₇	Å	1604.000 – 1630.000	1577.000 – 1597.000	1685.000 – 1705.000	[3]
008	BL ₁₆₆₄	Å	1651.000 – 1677.000	1577.000 – 1597.000	1685.000 – 1705.000	[3]
009	BL ₁₇₁₉	Å	1709.000 – 1729.000	1685.000 – 1705.000	1803.000 – 1823.000	[3]
010	BL ₁₈₅₃	Å	1838.000 – 1868.000	1803.000 – 1823.000	1885.000 – 1915.000	[3]
011	FeII ₂₄₀₂	Å	2382.000 – 2422.000	2285.000 – 2325.000	2432.000 – 2458.000	[3]
012	BL ₂₅₃₈	Å	2520.000 – 2556.000	2432.000 – 2458.000	2562.000 – 2588.000	[3]
013	FeII ₂₆₀₉	Å	2596.000 – 2622.000	2562.000 – 2588.000	2647.000 – 2673.000	[3]
014	B(2640)	break_lb	0.000 – 0.000	2600.000 – 2630.000	2645.000 – 2675.000	[7]
015	MgII	Å	2784.000 – 2814.000	2762.000 – 2782.000	2818.000 – 2838.000	[3]
016	MgI	Å	2839.000 – 2865.000	2818.000 – 2838.000	2906.000 – 2936.000	[3]
017	Mg _{UV}	bump	2625.000 – 2725.000	2525.000 – 2625.000	2725.000 – 2825.000	[4]
018	Mg _{wide}	Å	2670.000 – 2870.000	2470.000 – 2670.000	2930.000 – 3130.000	[3]
019	B(2900)	break_lb	0.000 – 0.000	2855.000 – 2885.000	2915.000 – 2945.000	[7]
020	FeI	Å	2965.000 – 3025.000	2906.000 – 2936.000	3031.000 – 3051.000	[3]
021	BL ₃₀₉₆	Å	3086.000 – 3106.000	3031.000 – 3051.000	3115.000 – 3155.000	[3]
022	CaII K	Å	3925.650 – 3945.000	3845.000 – 3880.000	3950.000 – 3954.000	[9]
023	CaII H	Å	3959.400 – 3975.000	3950.000 – 3954.000	3983.000 – 3993.000	[9]
024	D ₄₀₀₀	break_nu	0.000 – 0.000	3750.000 – 3950.000	4050.000 – 4250.000	[5]
025	D _n ₄₀₀₀	break_nu	0.000 – 0.000	3850.000 – 3950.000	4000.000 – 4100.000	[6]
026	H δ _A	Å	4083.500 – 4122.250	4041.600 – 4079.750	4128.500 – 4161.000	[2]
027	H δ _F	Å	4091.000 – 4112.250	4057.250 – 4088.500	4114.750 – 4137.250	[2]
028	CN ₁	mag	4142.125 – 4177.125	4080.125 – 4117.625	4244.125 – 4284.125	[1]
029	CN ₂	mag	4142.125 – 4177.125	4083.875 – 4096.375	4244.125 – 4284.125	[1]
030	Ca ₄₂₂₇	Å	4222.250 – 4234.750	4211.000 – 4219.750	4241.000 – 4251.000	[1]
031	G ₄₃₀₀	Å	4281.375 – 4316.375	4266.375 – 4282.625	4318.875 – 4335.125	[1]
032	H γ _A	Å	4319.750 – 4363.500	4283.500 – 4319.750	4367.250 – 4419.750	[2]
033	H γ _F	Å	4331.250 – 4352.250	4283.500 – 4319.750	4354.750 – 4384.750	[2]
034	Fe ₄₃₈₃	Å	4369.125 – 4420.375	4359.125 – 4370.375	4442.875 – 4455.375	[1]
035	Ca ₄₄₅₅	Å	4452.125 – 4474.625	4445.875 – 4454.625	4477.125 – 4492.125	[1]
036	Fe ₄₅₃₁	Å	4514.250 – 4559.250	4504.250 – 4514.250	4560.500 – 4579.250	[1]
037	C ₂ ₄₆₆₈	Å	4634.000 – 4720.250	4611.500 – 4630.250	4742.750 – 4756.500	[1]
038	H β	Å	4847.875 – 4876.625	4827.875 – 4847.875	4876.625 – 4891.625	[1]
039	Fe ₅₀₁₅	Å	4977.750 – 5054.000	4946.500 – 4977.750	5054.000 – 5065.250	[1]
040	Mg ₁	mag	5069.125 – 5134.125	4895.125 – 4957.625	5301.125 – 5366.125	[1]
041	Mg ₂	mag	5154.125 – 5196.625	4895.125 – 4957.625	5301.125 – 5366.125	[1]

Table A.1 (cont'd)

ID	Name	units	$(\lambda_{c1}, \lambda_{c2})$	$(\lambda_{b1}, \lambda_{b2})$	$(\lambda_{r1}, \lambda_{r2})$	Ref.
042	Mg _b	Å	5160.125 – 5192.625	5142.625 – 5161.375	5191.375 – 5206.375	[1]
043	Fe5270	Å	5245.650 – 5285.650	5233.150 – 5248.150	5285.650 – 5318.150	[1]
044	Fe5335	Å	5312.125 – 5352.125	5304.625 – 5315.875	5353.375 – 5363.375	[1]
045	Fe5406	Å	5387.500 – 5415.000	5376.250 – 5387.500	5415.000 – 5425.000	[1]
046	Fe5709	Å	5696.625 – 5720.375	5672.875 – 5696.625	5722.875 – 5736.625	[1]
047	Fe5782	Å	5776.625 – 5796.625	5765.375 – 5775.375	5797.875 – 5811.625	[1]
048	NaD	Å	5876.875 – 5909.375	5860.625 – 5875.625	5922.125 – 5948.125	[1]
049	TiO ₁	mag	5936.625 – 5994.125	5816.625 – 5849.125	6038.625 – 6103.625	[1]
050	TiO ₂	mag	6189.625 – 6272.125	6066.625 – 6141.625	6372.625 – 6415.125	[1]
051	CaT*	general	0.000 – 0.000	0.000 – 0.000	0.000 – 0.000	[8]
052	CN0.93	Å	9138.000 – 9465.000	9040.000 – 9070.000	9500.000 – 9530.000	[10]
053	FeH0.99	Å	9880.000 – 9935.000	9830.000 – 9860.000	9955.000 – 9980.000	[10]
054	CN1.10	Å	10920.00 – 11012.000	10845.00 – 10860.000	11124.00 – 11170.000	[10]
055	NaI1.14	Å	11362.00 – 11420.000	11305.00 – 11325.000	11453.00 – 11480.000	[10]
056	FeI1.16	Å	11593.00 – 11640.000	11540.00 – 11570.000	11705.00 – 11740.000	[10]
057	MgI1.18	Å	11810.00 – 11850.000	11705.00 – 11740.000	11905.00 – 11935.000	[10]
058	Paβ1.28	Å	12790.00 – 12845.000	12760.00 – 12780.000	12855.00 – 12875.000	[10]
059	All1.31	Å	13105.00 – 13160.000	13045.00 – 13075.000	13170.00 – 13195.000	[10]
060	CN1.42	Å	14070.00 – 14250.000	14015.00 – 14050.000	14315.00 – 14350.000	[10]
061	CN1.46	Å	14500.00 – 14750.000	14460.00 – 14485.000	14800.00 – 14840.000	[10]
062	MgI1.50	Å	15005.00 – 15044.000	14935.00 – 14985.000	15097.00 – 15120.000	[10]
063	CO1.56	Å	15545.00 – 15620.000	15445.00 – 15485.000	15640.00 – 15670.000	[10]
064	CO1.58	Å	15750.00 – 15810.000	15705.00 – 15735.000	15835.00 – 15860.000	[10]
065	CO1.60	Å	15970.00 – 16020.000	15920.00 – 15950.000	16045.00 – 16120.000	[10]
066	CO1.64	Å	16370.00 – 16434.000	16330.00 – 16355.000	16480.00 – 16505.000	[10]
067	CO1.66	Å	16600.00 – 16680.000	16565.00 – 16585.000	16780.00 – 16800.000	[10]
068	All1.67	Å	16710.00 – 16770.000	16565.00 – 16585.000	16780.00 – 16800.000	[10]
069	CO1.68	Å	16815.00 – 16890.000	16780.00 – 16800.000	16905.00 – 16945.000	[10]
070	CaI1.95	Å	19405.00 – 19530.000	19350.00 – 19385.000	19545.00 – 19580.000	[10]
071	CaI1.99a	Å	19730.00 – 19800.000	19660.00 – 19695.000	20005.00 – 20090.000	[10]
072	CaI1.99b	Å	19830.00 – 19885.000	19660.00 – 19695.000	20005.00 – 20090.000	[10]
073	CaI1.99c	Å	19905.00 – 19980.000	19660.00 – 19695.000	20005.00 – 20090.000	[10]
074	NaI2.21	Å	22025.00 – 22114.000	21930.00 – 21975.000	22130.00 – 22185.000	[10]
075	FeI2.24	Å	22360.00 – 22480.000	22315.00 – 22350.000	22500.00 – 22545.000	[10]
076	CaI2.26	Å	22590.00 – 22680.000	22500.00 – 22545.000	22715.00 – 22765.000	[10]
077	CO2.30	Å	22915.00 – 23015.000	22850.00 – 22895.000	23090.00 – 23170.000	[10]
078	CO2.32	Å	23190.00 – 23300.000	23090.00 – 23170.000	23340.00 – 23425.000	[10]
079	CO2.35	Å	23440.00 – 23630.000	23340.00 – 23425.000	23660.00 – 23710.000	[10]

Note. — References: [1] Trager et al. 1998; [2] Worthey & Ottaviani 1997; [3] Maraston et al. 2009; [4] Daddi et al. 2005; [5] Bruzual 1983; [6] Balogh et al. 1999; [7] Spinrad et al. 1997; [8] Cenarro et al. 2001; [9] Fanfani 2019; [10] Eftekhari et al. 2021.

A.2 Validation of CHIMERA

This Section expands on the validation tests performed for CHIMERA.

A.2.1 Validation of the spectral siren mode using external code

We validate CHIMERA in the spectral case (population analysis) with MGCosmoPop.¹ The code is presented and used in Mancarella et al. (2022b) to provide joint cosmological and astrophysical constraints, including deviation from general-relativity, in the empty catalog case. The codes feature similar model functions implementations, except for the cosmology class, which has been rewritten in CHIMERA to improve computational efficiency. We compare these functions across a wide range of parameters, finding machine-level differences.

On the contrary, the two codes feature distinct implementations of the likelihood evaluation. In MGCosmoPop the integral in Eq. 3.47 is performed with a Monte Carlo approach, while in CHIMERA we pixelize the sky, evaluate the redshift prior inside each pixel, and perform the integral on the redshift grid (see Fig. 3.10). While this approach has the advantage of improve the computational efficiency with large galaxy catalogs (Gray et al., 2022) and the evaluation of their completeness (Finke et al., 2021), it is crucial to perform a comparative analysis to better assess potential biases.

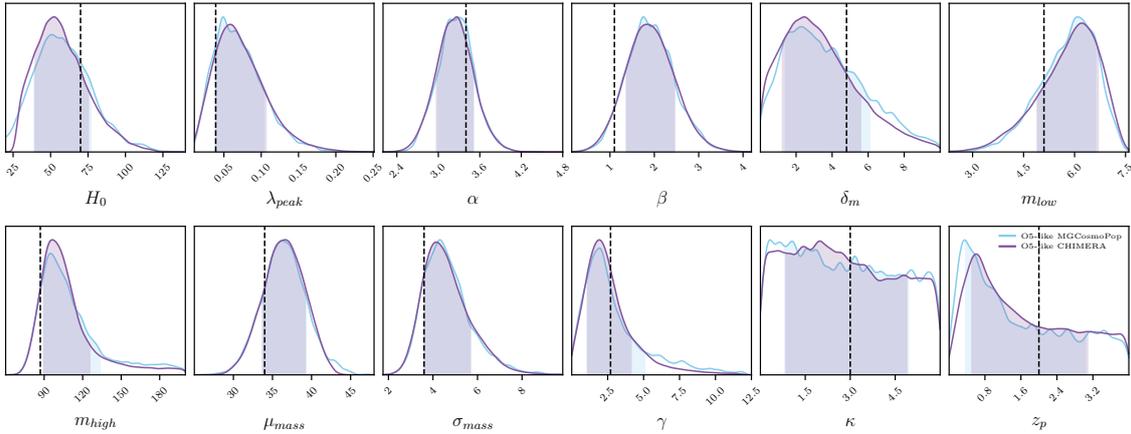


FIGURE A.2: Posterior distributions obtained in the case of the Spectral Siren analysis with CHIMERA and MGCosmoPop using the O5-like catalog.

We use the set of 100 BBH events at $\text{SNR} > 25$ presented in Section 3.4. We assume a flat Λ CDM cosmology, PLP mass distribution, and Madau-like rate evolution, with parameters and priors reported in Table 3.4.2. In CHIMERA, we set `npix_event` to 15 and `nside` to $[2^n]$ with $n = (3, \dots, 9)$ so that the pixelization of each event is automatically adjusted to have approximately 15 pixels in the 90% credible sky area. Then, we perform full MCMC analyses with CHIMERA and MGCosmoPop. The resulting posteriors are shown in Fig. A.2. The posteriors obtained with the two codes are in excellent agreement.

¹<https://github.com/CosmoStatGW/MGCosmoPop>

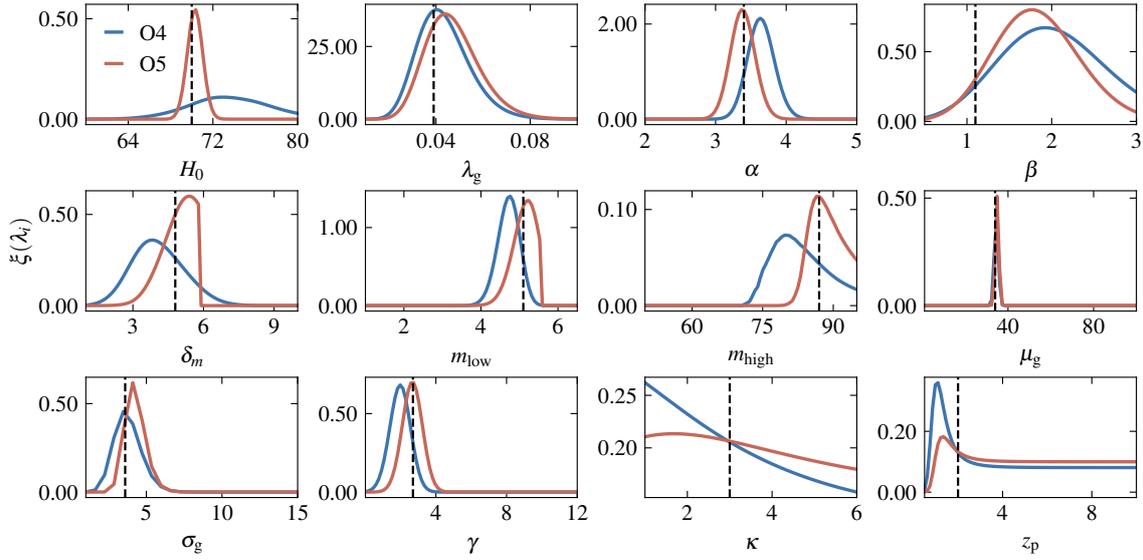


FIGURE A.3: One dimensional posteriors for population the parameters explored in the O4- and O5-like configurations presented in the main analysis in Section 3.5.

A.2.2 Tests on one-dimensional parameter space

Extensive validation tests are performed in CHIMERA during the development of the code. This involves conducting computationally efficient one-dimensional analyses for each individual hyperparameter while keeping the others fixed at their fiducial value. In this way, it is possible to detect, at first order, potential flaws in the generation of the data sets, data modeling, or implementation of the analysis pipeline.

One-dimensional posterior analysis

Figure A.3 shows the 1D tests carried out on the main analysis configuration presented in Section 3.5. Already from these analysis we recover fiducial values well within 1σ difference. The least constrained are rate evolution parameters k , representing the high- z slope and z_p (see Section 3.4). This result is expected as the GW events do not fully map $\varphi(z)$ beyond $z \sim 1$. The non-flat posteriors indicate that there may be a certain degree of correlation with other parameters, which contribute to inform final posteriors. This effect is seen also when performing the spectral siren analysis (no galaxy catalog), also with external code (see Appendix A.2.1).

KDE bandwidth

Figure A.4 shows a 1D test carried out varying the KDE smoothing parameters. In fact, when converting the posterior distribution of GW data in a three-dimensional function $f = f(d_L, \text{RA}, \text{Dec})$ using the KDE, one must provide a smoothing bandwidth bw . The optimal choice for bw is a value which correctly smooths out the posteriors without generating spurious sub-structures, while at the same time, avoiding to artificially broaden the constraints. We find that the Scott's rule (Scott, 1979) to estimate bw , provided as a default

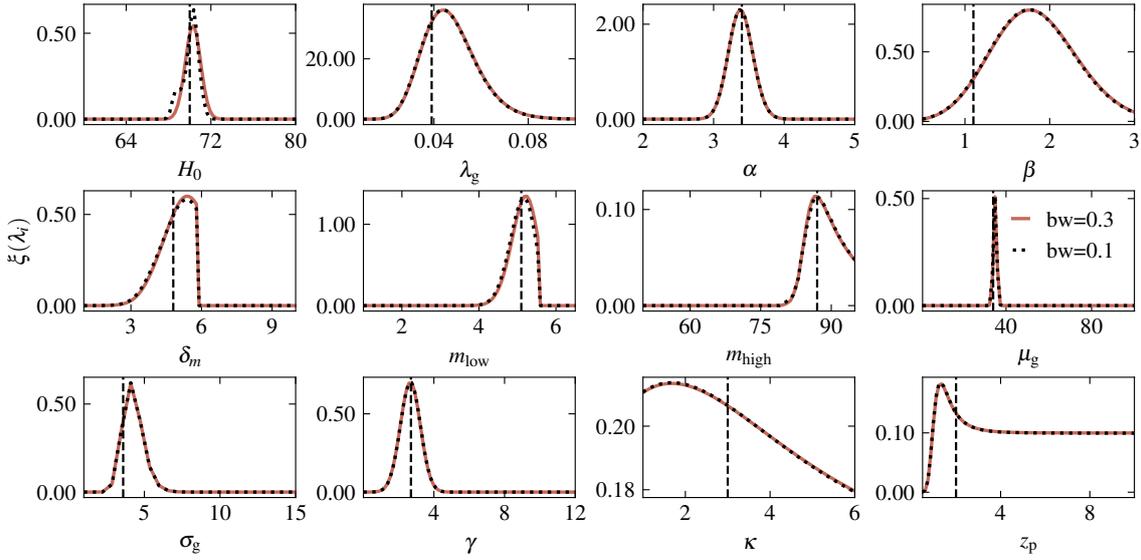


FIGURE A.4: One dimensional posteriors for O5 varying the KDE bandwidth parameter.

in many codes (e.g. `scipy` Virtanen et al., 2020), does not always fulfill the second requirement. Therefore the `bw` parameter must be determined by a more careful analysis and on a catalog-by-catalog basis. In particular, we use the `GridSearchCV` estimator logarithm to find the optimal bandwidth on generated GW data. We perform the test on a logarithmic grid spanning values of $10^{-3} < \text{bw} < 10$, finding that an optimal KDE bandwidth is $\text{bw} = 0.3$. Larger value would artificially broaden the constraints. However, Fig. A.4 also shows that the constraints do not significantly vary with values below 0.3, up to $\text{bw} = 0.1$, when spurious peaks would appear in the H_0 posteriors.

Shuffling (or not) of the observed galaxy redshifts

Inconsistency between the statistical framework presented in Section 3.2 and the generation of mock data Section 3.5 may result in biases in the estimation of the parameters, in particular H_0 . Recently, Gair et al. (2023) explored the main inconsistencies that may arise in dark sirens analyses. Here, we discuss the issue of not re-shuffling the true galaxy redshifts in the catalog. In fact, we must take into account that we are not perfectly able to measure the true redshifts, but we have measured values (see the discussion in Section 3.2.2). Therefore, for a consistent implementation of the methodology, we must reshuffle each galaxy redshift according to the adopted uncertainty. Figure A.5 shows the effects of this inconsistency for three analysis configurations: O4 network with photometric (z_{phot}) and spectroscopic (z_{spec}) galaxy catalog and O5 network with spectroscopic galaxy catalog. We find that in the two extreme cases of having the minimum information from galaxy distribution (O4 with photometric catalog) or the maximum information (O5 with spectroscopic catalog), the result show no significant bias. However, in the intermediate case of O4 GW data with a spectroscopic galaxy catalog, this inconsistency may lead to a $\sim 6 \text{ km s}^{-1} \text{ Mpc}^{-1}$ bias, highlighting the importance of correctly implement the statistical assumptions.

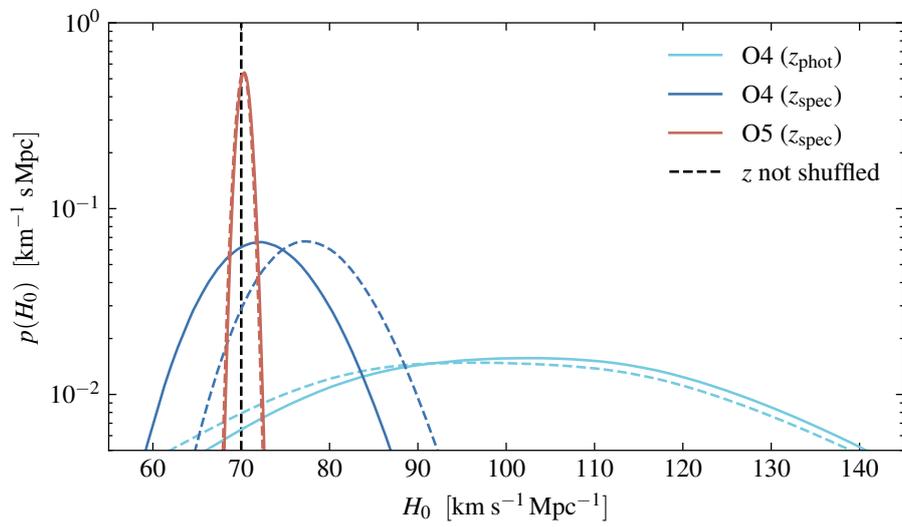


FIGURE A.5: Posteriors on H_0 derived for three analysis configurations (see text) and effects of the inconsistency of not perturbing galaxy redshifts around their true value.

A.3 Full corners for the O4 and O5 standard sirens analyses

In this section, we show the full 12-parameters marginalized 1D and 2D posteriors for the standard sirens analysis configurations presented in Section 3.5. For all the configurations, we are able to recover the fiducial values with a typical deviation of 0.2σ . The merger rate evolution parameters k and z_p remain practically unconstrained. This is expected as the GW events do not fully map $\varphi(z)$ beyond $z \sim 1$ (see also Fig. 3.18). The other parameters show fluctuations around the fiducial values that can be ascribed to the specific realizations of the data sets.

In particular, we focus on two types of comparisons:

- *Improvement of GW data:* O4- and O5-like scenarios including a complete galaxy catalog with spectroscopic (Fig. A.6) or photometric (Fig. A.7) redshift uncertainties.
- *Improvement of EM data:* photometric and spectroscopic galaxy catalogs in the O4- (Fig. A.8) or O5-like (Fig. A.9) detector configurations.

In general, at fixed EM data, the more advanced GW detector configuration of O5 provides better constraints on almost all the constrained parameters. This is not the case for the rate parameter γ as the higher SNR cut adopted in the O5 catalog yields GW events which map the same redshift range as in O4. To date, a detailed study of the impact on the final constraints by varying the SNR cut of GW data has not been done yet and requires a future dedicated study. In fact, even if bad localized events cannot directly constrain H_0 , they can still provide significant improvement to astrophysical population constraints and thus indirectly inform H_0 by breaking degeneracies with mass parameters.

At fixed GW data, having a spectroscopic instead of a photometric galaxy catalog greatly improves the constraints on H_0 and as discussed in Section 3.5 it is needed to reach a percent level determination. As expected, the constraints on the astrophysical population parameters are mostly unaffected. However, as noted in Section 3.5, a small $\sim 25\%$ improvement is seen for the uncertainty on the position of the Gaussian peak of the BBH mass function μ_g .

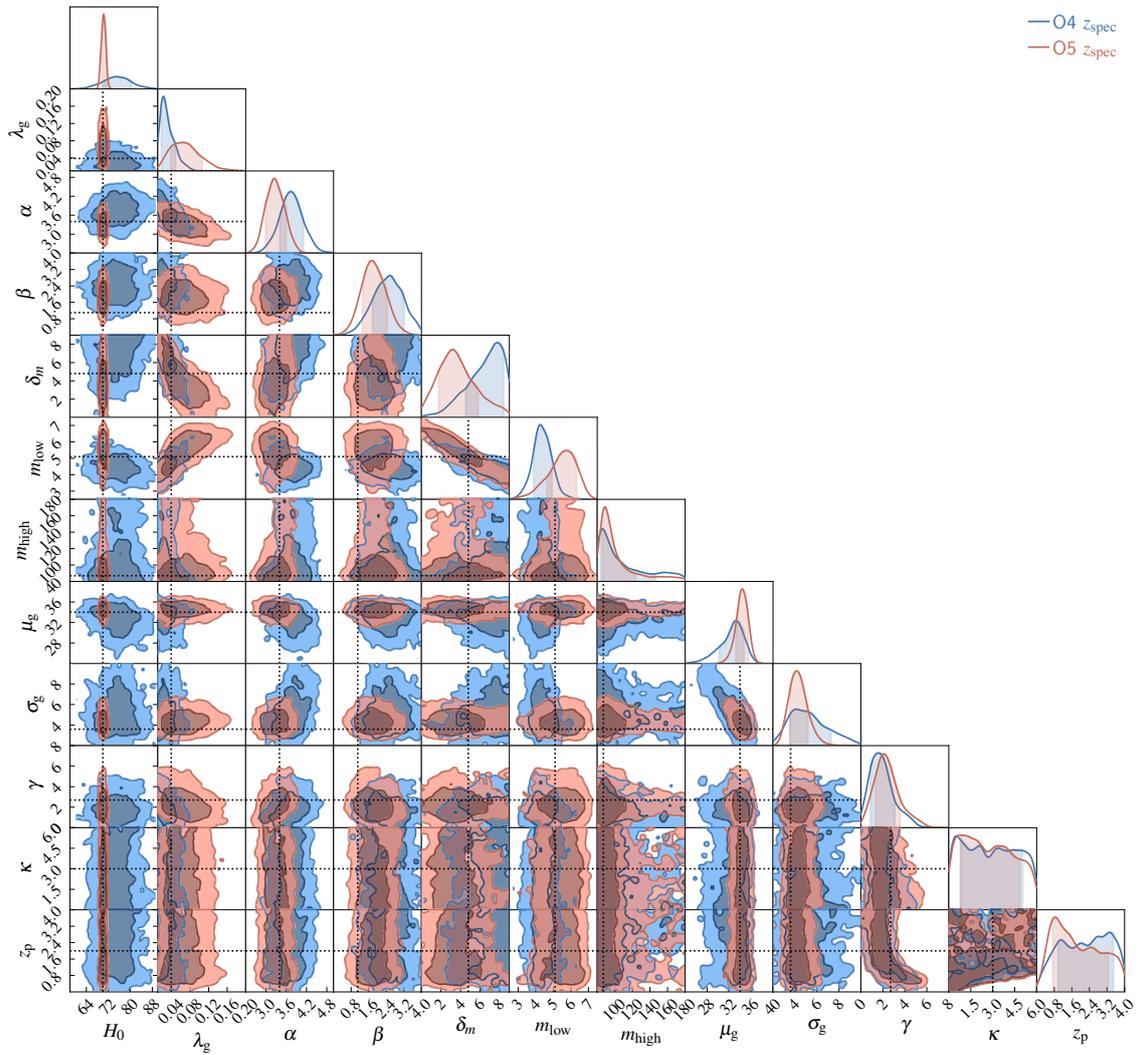


FIGURE A.6: Marginalized 1D and 2D posteriors for the O4- and O5-like configurations in case of a complete galaxy catalog with spectroscopic redshift measurements.

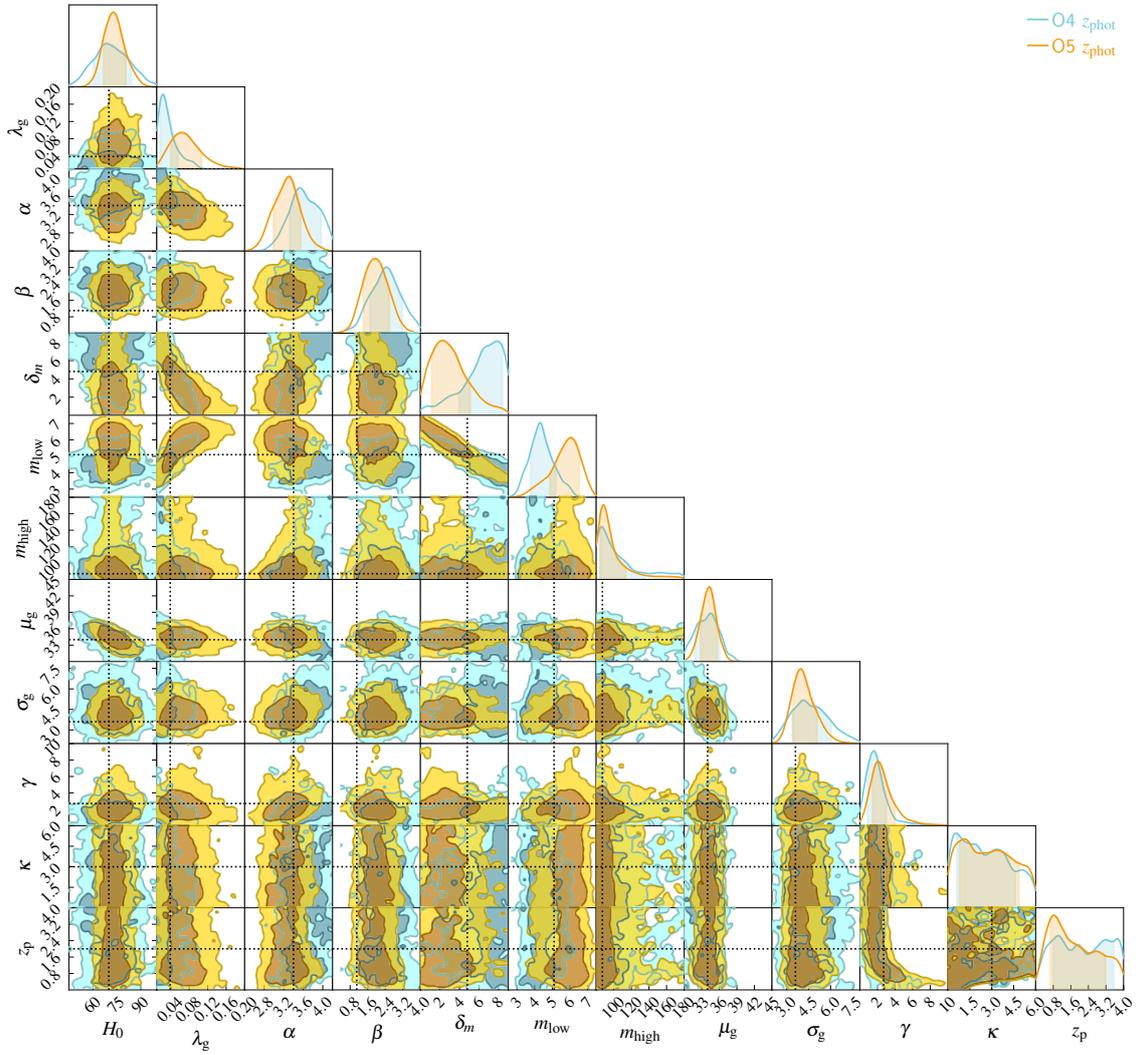


FIGURE A.7: Marginalized 1D and 2D posteriors for the O4- and O5-like configurations in case of a complete galaxy catalog with photometric redshift measurements.

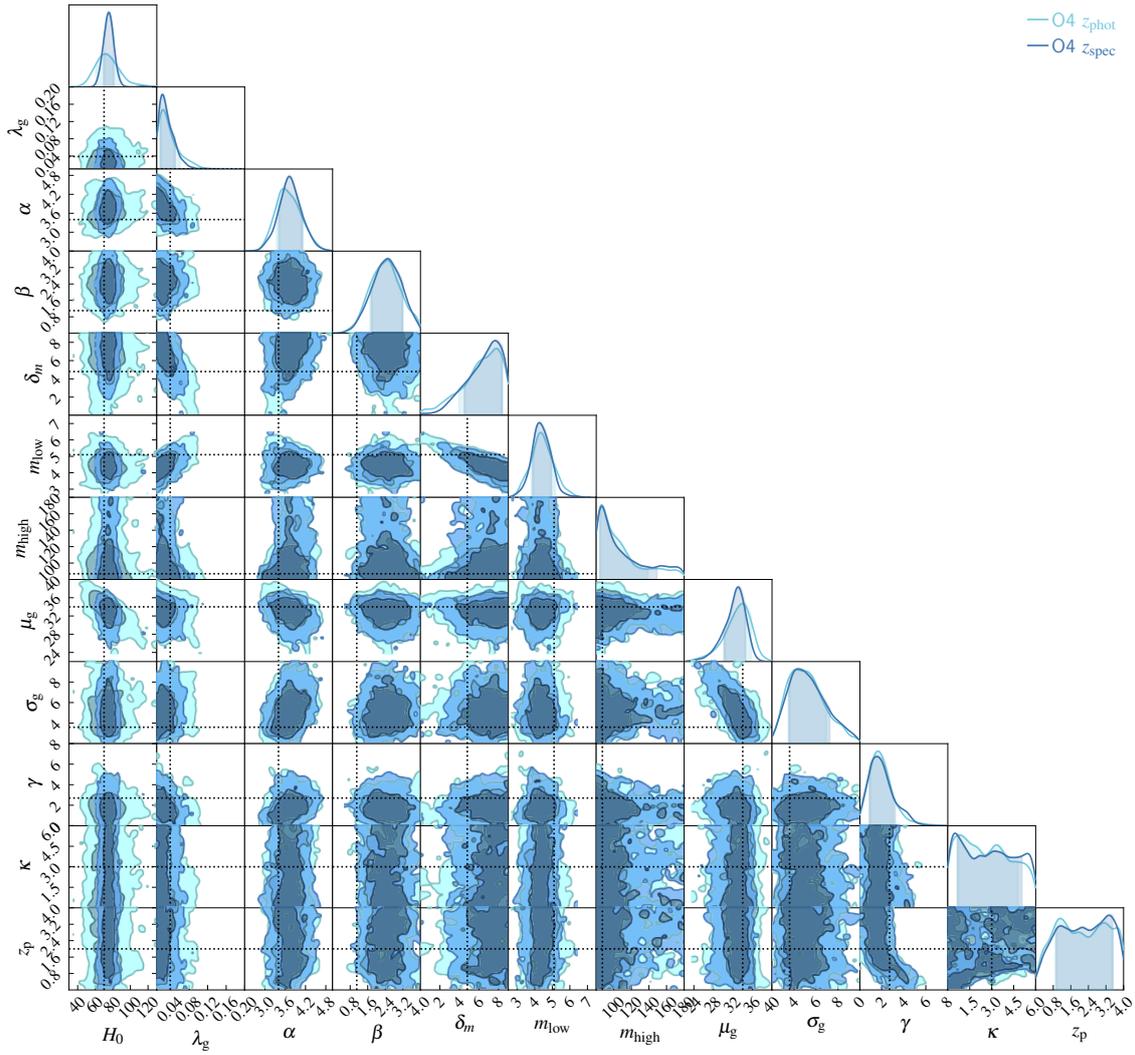


FIGURE A.8: Marginalized 1D and 2D posteriors for the photometric and spectroscopic galaxy catalogs in the O4-like detector configuration.

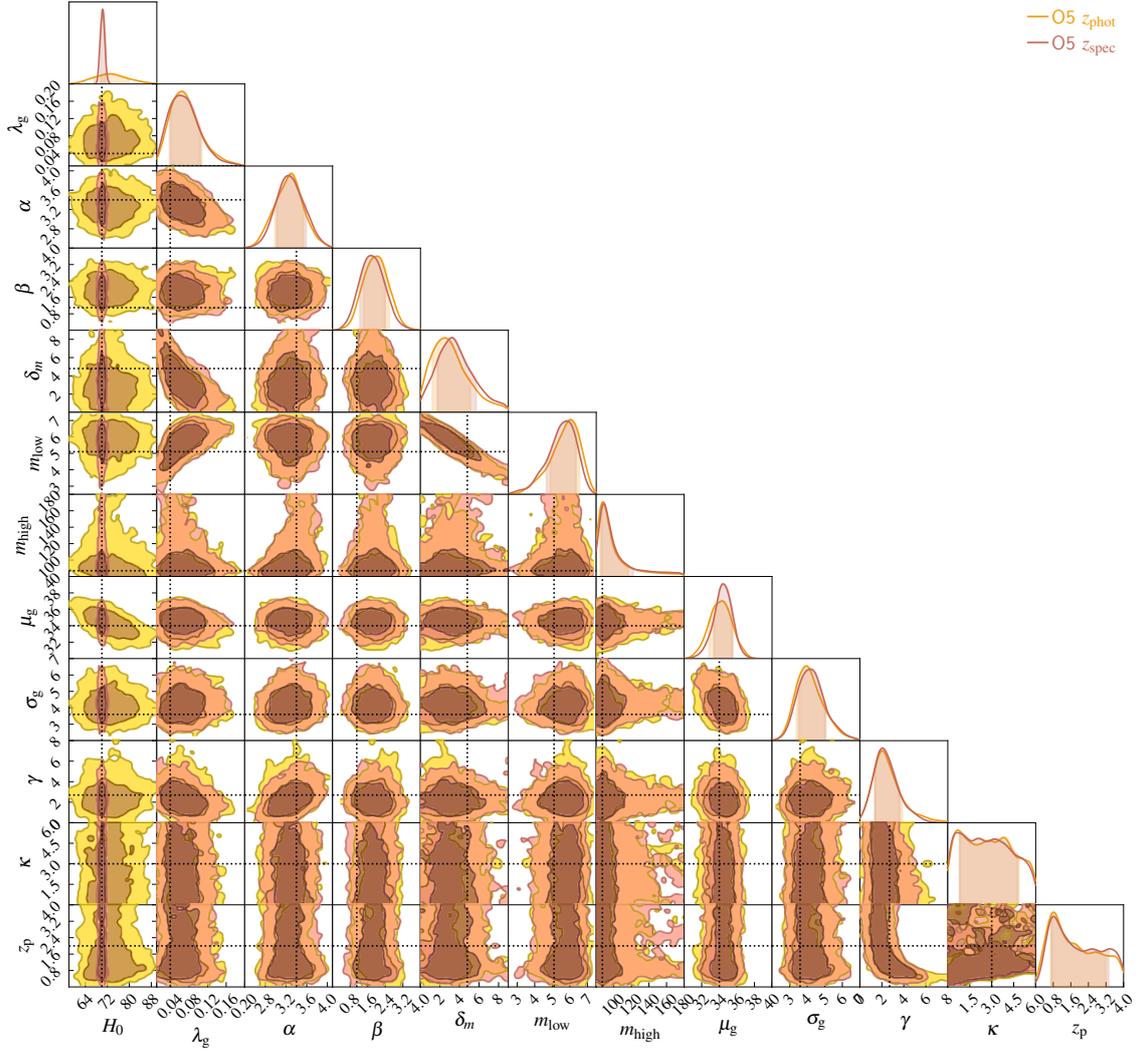


FIGURE A.9: Marginalized 1D and 2D posteriors for the photometric and spectroscopic galaxy catalogs in the O5-like detector configuration.

Bibliography

- Abbott B. P., et al., 2016a, “Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO and Advanced Virgo” *Living Reviews in Relativity*, 19, 1
- Abbott B. P., et al., 2016b, “Observation of Gravitational Waves from a Binary Black Hole Merger” *Phys. Rev. Lett.*, 116, 061102 ([arXiv:1602.03837](#))
- Abbott B. P., et al., 2017a, “GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral” *Phys. Rev. Lett.*, 119, 161101 ([arXiv:1710.05832](#))
- Abbott B. P., et al., 2017b, “A gravitational-wave standard siren measurement of the Hubble constant” *Nature*, 551, 85 ([arXiv:1710.05835](#))
- Abbott B. P., et al., 2017c, “Multi-messenger Observations of a Binary Neutron Star Merger” *ApJ*, 848, L12 ([arXiv:1710.05833](#))
- Abbott B. P., et al., 2017d, “Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A” *ApJ*, 848, L13 ([arXiv:1710.05834](#))
- Abbott R., et al., 2020a, “GW190521: A Binary Black Hole Merger with a Total Mass of $150 M_{\odot}$ ” *Phys. Rev. Lett.*, 125, 101102 ([arXiv:2009.01075](#))
- Abbott R., et al., 2020b, “GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object” *ApJ*, 896, L44 ([arXiv:2006.12611](#))
- Abbott R., et al., 2021, “Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog” *ApJ*, 913, L7 ([arXiv:2010.14533](#))
- Abbott R., et al., 2023a, “Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3” *Physical Review X*, 13, 011048 ([arXiv:2111.03634](#))
- Abbott R., et al., 2023b, “GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run” *Physical Review X*, 13, 041039 ([arXiv:2111.03606](#))
- Abbott R., et al., 2023c, “Constraints on the Cosmic Expansion History from GWTC-3” *ApJ*, 949, 76 ([arXiv:2111.03604](#))
- Abdalla E., et al., 2022, “Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies” *Journal of High Energy Astrophysics*, 34, 49 ([arXiv:2203.06142](#))

- Adhikari R. X., 2014, “Gravitational radiation detection with laser interferometry” *Reviews of Modern Physics*, 86, 121 ([arXiv:1305.5188](#))
- Agazie G., et al., 2023, “The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background” *ApJ*, 951, L8 ([arXiv:2306.16213](#))
- Ahumada R., et al., 2020, “The 16th Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra” *ApJS*, 249, 3 ([arXiv:1912.02905](#))
- Ajith P., 2011, “Addressing the spin question in gravitational-wave searches: Waveform templates for inspiralling compact binaries with nonprecessing spins” *Phys. Rev. D*, 84, 084037 ([arXiv:1107.1267](#))
- Akeson R., et al., 2019, “The Wide Field Infrared Survey Telescope: 100 Hubbles for the 2020s” *arXiv e-prints*, p. arXiv:1902.05569 ([arXiv:1902.05569](#))
- Alam S., et al., 2017, “The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample” *MNRAS*, 470, 2617 ([arXiv:1607.03155](#))
- Alam S., et al., 2021, “Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory” *Phys. Rev. D*, 103, 083533 ([arXiv:2007.08991](#))
- Albrecht A., Steinhardt P. J., Turner M. S., Wilczek F., 1982, “Reheating an Inflationary Universe” *Phys. Rev. Lett.*, 48, 1437
- Alpher R. A., Bethe H., Gamow G., 1948, “The Origin of Chemical Elements” *Physical Review*, 73, 803
- Amaro-Seoane P., et al., 2017, “Laser Interferometer Space Antenna” *arXiv e-prints*, p. arXiv:1702.00786 ([arXiv:1702.00786](#))
- Amati L., Guidorzi C., Frontera F., Della Valle M., Finelli F., Landi R., Montanari E., 2008, “Measuring the cosmological parameters with the $E_{p,i}$ - E_{iso} correlation of gamma-ray bursts” *MNRAS*, 391, 577 ([arXiv:0805.0377](#))
- Anderson L., et al., 2014, “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples” *MNRAS*, 441, 24 ([arXiv:1312.4877](#))
- Arnouts S., et al., 2007, “The SWIRE-VVDS-CFHTLS surveys: stellar mass assembly over the last 10 Gyr. Evidence for a major build up of the red sequence between $z = 2$ and $z = 1$ ” *A&A*, 476, 137 ([arXiv:0705.2438](#))
- Astropy Collaboration et al., 2018, “The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package” *AJ*, 156, 123 ([arXiv:1801.02634](#))

- Baes M., Camps P., Van De Putte D., 2017, “Analytical expressions and numerical evaluation of the luminosity distance in a flat cosmology” *MNRAS*, 468, 927 ([arXiv:1702.08860](#))
- Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, “Quantifying the Bimodal Color-Magnitude Distribution of Galaxies” *The Astrophysical Journal*, 600, 681
- Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1999, “Differential Galaxy Evolution in Cluster and Field Galaxies at $z \sim 0.3$ ” *ApJ*, 527, 54 ([arXiv:astro-ph/9906470](#))
- Barone T. M., et al., 2018, “The SAMI Galaxy Survey: gravitational potential and surface density drive stellar populations – I. early-type galaxies” *ApJ*, 856, 64 ([arXiv:1802.04807](#))
- Bassett B., Hlozek R., 2010, in Ruiz-Lapuente P., ed., , Dark Energy: Observational and Theoretical Approaches. arXiv, p. 246, [doi:10.48550/arXiv.0910.5224](#), <https://ui.adsabs.harvard.edu/abs/2010deot.book..246B>
- Beifiori A., Maraston C., Thomas D., Johansson J., 2011, “On the spectral resolution of the MILES stellar library” *A&A*, 531, A109 ([arXiv:1012.3428](#))
- Belgacem E., Dirian Y., Foffa S., Maggiore M., 2018a, “Gravitational-wave luminosity distance in modified gravity theories” *Phys. Rev. D*, 97, 104066 ([arXiv:1712.08108](#))
- Belgacem E., Dirian Y., Foffa S., Maggiore M., 2018b, “Modified gravitational-wave propagation and standard sirens” *Phys. Rev. D*, 98, 023510 ([arXiv:1805.08731](#))
- Belgacem E., et al., 2019, “Testing modified gravity at cosmological distances with LISA standard sirens” *J. Cosmology Astropart. Phys.*, 2019, 024 ([arXiv:1906.01593](#))
- Belli S., Newman A. B., Ellis R. S., 2015, “Stellar Populations from Spectroscopy of a Large Sample of Quiescent Galaxies at $Z > 1$: Measuring the Contribution of Progenitor Bias to Early Size Growth” *ApJ*, 799, 206 ([arXiv:1409.7088](#))
- Bender R., Davies R. L., eds, 1996, Measuring the evolution of the M/L ratio from the Fundamental Plane in CL 0024+16 at $z=0.39$ IAU Symposium Vol. 171. arXiv ([arXiv:astro-ph/9603029](#)), [doi:10.48550/arXiv.astro-ph/9603029](#), <https://ui.adsabs.harvard.edu/abs/1996IAUS..171..233F>
- Beverage A. G., Kriek M., Conroy C., Bezanson R., Franx M., van der Wel A., 2021, “Elemental Abundances and Ages of $z \sim 0.7$ Quiescent Galaxies on the Mass–Size Plane: Implication for Chemical Enrichment and Star Formation Quenching” *ApJ*, 917, L1 ([arXiv:2105.12750](#))
- Birrer S., et al., 2020, “TDCOSMO. IV. Hierarchical time-delay cosmography - joint inference of the Hubble constant and galaxy density profiles” *A&A*, 643, A165 ([arXiv:2007.02941](#))
- Bishop N. T., Rezzolla L., 2016, “Extraction of gravitational waves in numerical relativity” *Living Reviews in Relativity*, 19, 2 ([arXiv:1606.02532](#))

- Blakeslee J. P., Jensen J. B., Ma C.-P., Milne P. A., Greene J. E., 2021, “The Hubble Constant from Infrared Surface Brightness Fluctuation Distances” *ApJ*, 911, 65 ([arXiv:2101.02221](#))
- Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, “Formation of galaxies and large-scale structure with cold dark matter.” *Nature*, 311, 517
- Bonazzola S.,ourgoulhon E., 1996, “Gravitational waves from pulsars: emission by the magnetic-field-induced distortion.” *A&A*, 312, 675 ([arXiv:astro-ph/9602107](#))
- Borghi N., Moresco M., Cimatti A., Huchet A., Quai S., Pozzetti L., 2022a, “Toward a Better Understanding of Cosmic Chronometers: Stellar Population Properties of Passive Galaxies at Intermediate Redshift” *ApJ*, 927, 164 ([arXiv:2106.14894](#))
- Borghi N., Moresco M., Cimatti A., 2022b, “Toward a Better Understanding of Cosmic Chronometers: A New Measurement of $H(z)$ at z 0.7” *ApJ*, 928, L4 ([arXiv:2110.04304](#))
- Borghi N., Mancarella M., Moresco M., Tagliazucchi M., Iacovelli F., Cimatti A., Maggiore M., 2024, “Cosmology and Astrophysics with Standard Sirens and Galaxy Catalogs in View of Future Gravitational Wave Observations” *Accepted to ApJ* ([arXiv:2312.05302](#))
- Bradbury J., et al., 2018, JAX: composable transformations of Python+NumPy programs, <http://github.com/google/jax>
- Branchesi M., et al., 2023, “Science with the Einstein Telescope: a comparison of different designs” *J. Cosmology Astropart. Phys.*, 2023, 068 ([arXiv:2303.15923](#))
- Brout D., et al., 2022, “The Pantheon+ Analysis: Cosmological Constraints” *ApJ*, 938, 110 ([arXiv:2202.04077](#))
- Bruzual G., 1983, “Spectral evolution of galaxies. I. Early-type systems.” *ApJ*, 273, 105
- Bruzual G., Charlot S., 2003, “Stellar population synthesis at the resolution of 2003” *MNRAS*, 344, 1000 ([arXiv:astro-ph/0309134](#))
- Buchs R., et al., 2019, “Phenotypic redshifts with self-organizing maps: A novel method to characterize redshift distributions of source galaxies for weak lensing” *MNRAS*, 489, 820 ([arXiv:1901.05005](#))
- Bullock J. S., Boylan-Kolchin M., 2017, “Small-Scale Challenges to the Λ CDM Paradigm” *ARA&A*, 55, 343 ([arXiv:1707.04256](#))
- Buonanno A., Iyer B. R., Ochsner E., Pan Y., Sathyaprakash B. S., 2009, “Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors” *Phys. Rev. D*, 80, 084043 ([arXiv:0907.0700](#))
- Callister T., Fishbach M., Holz D. E., Farr W. M., 2020, “Shouts and Murmurs: Combining Individual Gravitational-wave Sources with the Stochastic Background to Measure the History of Binary Black Hole Mergers” *ApJ*, 896, L32 ([arXiv:2003.12152](#))

- Cappellari M., et al., 2013, “The ATLAS^{3D} project - XV. Benchmark for early-type galaxies scaling relations from 260 dynamical models: mass-to-light ratio, dark matter, Fundamental Plane and Mass Plane” *MNRAS*, 432, 1709 ([arXiv:1208.3522](#))
- Caprini C., Figueroa D. G., 2018, “Cosmological backgrounds of gravitational waves” *Classical and Quantum Gravity*, 35, 163001 ([arXiv:1801.04268](#))
- Cardiel N., Gorgas J., Cenarro J., Gonzalez J. J., 1998, “Reliable random error estimation in the measurement of line-strength indices” *A&AS*, 127, 597 ([arXiv:astro-ph/9706116](#))
- Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, “Inferring the star formation histories of massive quiescent galaxies with BAGPIPES: evidence for multiple quenching mechanisms” *MNRAS*, 480, 4379 ([arXiv:1712.04452](#))
- Carnall A. C., et al., 2019, “The VANDELS survey: the star-formation histories of massive quiescent galaxies at $1.0 < z < 1.3$ ” *MNRAS*, 490, 417 ([arXiv:1903.11082](#))
- Carnall A. C., et al., 2023, “A massive quiescent galaxy at redshift 4.658” *Nature*, 619, 716 ([arXiv:2301.11413](#))
- Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, “An algorithm to build mock galaxy catalogues using MICE simulations” *MNRAS*, 447, 646 ([arXiv:1411.3286](#))
- Carson D. P., Nichol R. C., 2010, “The age-redshift relation for luminous red galaxies in the Sloan Digital Sky Survey” *MNRAS*, 408, 213 ([arXiv:1006.2830](#))
- Cassisi S., Castellani M., Castellani V., 1997, “Intermediate-age metal deficient stellar populations: the case of metallicity $Z=0.00001$.” *A&A*, 317, 108 ([arXiv:astro-ph/9603023](#))
- Cenarro A. J., Cardiel N., Gorgas J., Peletier R. F., Vazdekis A., Prada F., 2001, “Empirical calibration of the near-infrared Ca ii triplet - I. The stellar library and index definition” *MNRAS*, 326, 959 ([arXiv:astro-ph/0109157](#))
- Chen H.-Y., Fishbach M., Holz D. E., 2018, “A two per cent Hubble constant measurement from standard sirens within five years” *Nature*, 562, 545 ([arXiv:1712.06531](#))
- Chernoff D. F., Finn L. S., 1993, “Gravitational Radiation, Inspiral Binaries, and Cosmology” *ApJ*, 411, L5 ([arXiv:gr-qc/9304020](#))
- Chevallier M., Polarski D., 2001, “Accelerating Universes with Scaling Dark Matter” *International Journal of Modern Physics D*, 10, 213 ([arXiv:gr-qc/0009008](#))
- Choi J., Conroy C., Moustakas J., Graves G. J., Holden B. P., Brodwin M., Brown M. J. I., van Dokkum P. G., 2014, “The Assembly Histories of Quiescent Galaxies since $z = 0.7$ from Absorption Line Spectroscopy” *ApJ*, 792, 95 ([arXiv:1403.4932](#))
- Cimatti A., Moresco M., 2023, “Revisiting the Oldest Stars as Cosmological Probes: New Constraints on the Hubble Constant” *ApJ*, 953, 149 ([arXiv:2302.07899](#))

- Cimatti A., et al., 2004, “Old galaxies in the young Universe” *Nature*, 430, 184 ([arXiv:astro-ph/0407131](https://arxiv.org/abs/astro-ph/0407131))
- Cimatti A., Daddi E., Renzini A., 2006, “Mass downsizing and “top-down” assembly of early-type galaxies” *A&A*, 453, L29 ([arXiv:astro-ph/0605353](https://arxiv.org/abs/astro-ph/0605353))
- Cimatti A., Fraternali F., Nipoti C., 2019, Introduction to Galaxy Formation and Evolution: From Primordial Gas to Present-Day Galaxies. Cambridge University Press, <https://ui.adsabs.harvard.edu/abs/2019igfe.book.....C>
- Citro A., Pozzetti L., Moresco M., Cimatti A., 2016, “Inferring the star-formation histories of the most massive and passive early-type galaxies at $z < 0.3$ ” *A&A*, 592, A19 ([arXiv:1604.07826](https://arxiv.org/abs/1604.07826))
- Cleveland W. S., Devlin S. J., 1988, “Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting” *Journal of the American Statistical Association*, 83, 596
- Cole S., et al., 2005, “The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications” *MNRAS*, 362, 505 ([arXiv:astro-ph/0501174](https://arxiv.org/abs/astro-ph/0501174))
- Colombo A., Salafia O. S., Gabrielli F., Ghirlanda G., Giacomazzo B., Perego A., Colpi M., 2022, “Multi-messenger Observations of Binary Neutron Star Mergers in the O4 Run” *ApJ*, 937, 79 ([arXiv:2204.07592](https://arxiv.org/abs/2204.07592))
- Concas A., Pozzetti L., Moresco M., Cimatti A., 2017, “On the robustness of the $H\beta$ Lick index as a cosmic clock in passive early-type galaxies” *MNRAS*, 468, 1747 ([arXiv:1702.07348](https://arxiv.org/abs/1702.07348))
- Conley A., et al., 2011, “Supernova Constraints and Systematic Uncertainties from the First Three Years of the Supernova Legacy Survey” *ApJS*, 192, 1 ([arXiv:1104.1443](https://arxiv.org/abs/1104.1443))
- Conselice C. J., Wilkinson A., Duncan K., Mortlock A., 2016, “The Evolution of Galaxy Number Density at $z < 8$ and Its Implications” *ApJ*, 830, 83 ([arXiv:1607.03909](https://arxiv.org/abs/1607.03909))
- Contarini S., Ronconi T., Marulli F., Moscardini L., Veropalumbo A., Baldi M., 2019, “Cosmological exploitation of the size function of cosmic voids identified in the distribution of biased tracers” *MNRAS*, 488, 3526 ([arXiv:1904.01022](https://arxiv.org/abs/1904.01022))
- Cook S. R., Gelman A., Rubin D. B., 2006, “Validation of Software for Bayesian Models Using Posterior Quantiles” *Journal of Computational and Graphical Statistics*, 15, 675
- Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, “New Insight on Galaxy Formation and Evolution From Keck Spectroscopy of the Hawaii Deep Fields” *AJ*, 112, 839 ([arXiv:astro-ph/9606079](https://arxiv.org/abs/astro-ph/9606079))
- Crocce M., Castander F. J., Gaztañaga E., Fosalba P., Carretero J., 2015, “The MICE Grand Challenge lightcone simulation - II. Halo and galaxy catalogues” *MNRAS*, 453, 1513 ([arXiv:1312.2013](https://arxiv.org/abs/1312.2013))

- DESI Collaboration et al., 2016, “The DESI Experiment Part I: Science, Targeting, and Survey Design” *arXiv e-prints*, p. [arXiv:1611.00036](#) ([arXiv:1611.00036](#))
- Daddi E., et al., 2005, “Passively Evolving Early-Type Galaxies at $1.4 < z < 2.5$ in the Hubble Ultra Deep Field” *ApJ*, 626, 680 ([arXiv:astro-ph/0503102](#))
- Dalang C., Baker T., 2023, “The clustering of dark siren host galaxies” *arXiv e-prints*, p. [arXiv:2310.08991](#) ([arXiv:2310.08991](#))
- Dálya G., et al., 2022, “GLADE+ : an extended galaxy catalogue for multimessenger searches with advanced gravitational-wave detectors” *MNRAS*, 514, 1403 ([arXiv:2110.06184](#))
- Davidzon I., et al., 2017, “The COSMOS2015 galaxy stellar mass function . Thirteen billion years of stellar mass assembly in ten snapshots” *A&A*, 605, A70 ([arXiv:1701.02734](#))
- Dekel A., Sari R., Ceverino D., 2009, “Formation of Massive Galaxies at High Redshift: Cold Streams, Clumpy Disks, and Compact Spheroids” *ApJ*, 703, 785 ([arXiv:0901.2458](#))
- Del Pozzo W., 2012, “Inference of cosmological parameters from gravitational waves: Applications to second generation interferometers” *Phys. Rev. D*, 86, 043011 ([arXiv:1108.1317](#))
- Di Valentino E., et al., 2021a, “In the Realm of the Hubble tension – a Review of Solutions” *Class. Quantum Grav.* 38, 153001 (2021), 38, 153001 ([arXiv:2103.01183](#))
- Di Valentino E., et al., 2021b, “Cosmology Intertwined III: $f\sigma_8$ and S_8 ” *Astroparticle Physics*, 131, 102604 ([arXiv:2008.11285](#))
- Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, “Cosmic Black-Body Radiation.” *ApJ*, 142, 414
- Dietrich T., Samajdar A., Khan S., Johnson-McDaniel N. K., Dudi R., Tichy W., 2019, “Improving the NRTidal model for binary neutron star systems” *Phys. Rev. D*, 100, 044003 ([arXiv:1905.06011](#))
- Dressler A., Smail I., Poggianti B. M., Butcher H., Couch W. J., Ellis R. S., Oemler Augustus J., 1999, “A Spectroscopic Catalog of 10 Distant Rich Clusters of Galaxies” *ApJS*, 122, 51 ([arXiv:astro-ph/9901263](#))
- Edelman B., Doctor Z., Godfrey J., Farr B., 2022, “Ain’t No Mountain High Enough: Semi-parametric Modeling of LIGO-Virgo’s Binary Black Hole Mass Distribution” *ApJ*, 924, 101 ([arXiv:2109.06137](#))
- Eftekhari E., Vazdekis A., La Barbera F., 2021, “Fingerprints of stellar populations in the near-infrared: an optimized set of spectral indices in the JHK bands” *MNRAS*, 504, 2190 ([arXiv:2104.01051](#))
- Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, “Evidence from the motions of old stars that the Galaxy collapsed.” *ApJ*, 136, 748
- Einstein A., 1915, “Zur allgemeinen Relativitätstheorie” *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pp 778–786

- Einstein A., 1916, “Die Grundlage der allgemeinen Relativitätstheorie” *Annalen der Physik*, 354, 769
- Einstein A., 1917, “Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie” *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pp 142–152
- Einstein A., 1918, “Über Gravitationswellen” *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pp 154–167
- Eisenstein D. J., et al., 2005, “Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies” *ApJ*, 633, 560 ([arXiv:astro-ph/0501171](https://arxiv.org/abs/astro-ph/0501171))
- Escudero M., Witte S. J., 2020, “A CMB search for the neutrino mass mechanism and its relation to the Hubble tension” *European Physical Journal C*, 80, 294 ([arXiv:1909.04044](https://arxiv.org/abs/1909.04044))
- Estrada-Carpenter V., et al., 2019, “CLEAR. I. Ages and Metallicities of Quiescent Galaxies at $1.0 < z < 1.8$ Derived from Deep Hubble Space Telescope Grism Data” *ApJ*, 870, 133 ([arXiv:1810.02824](https://arxiv.org/abs/1810.02824))
- Evans M., et al., 2021, “A Horizon Study for Cosmic Explorer: Science, Observatories, and Community” *arXiv e-prints*, p. arXiv:2109.09882 ([arXiv:2109.09882](https://arxiv.org/abs/2109.09882))
- Ezquiaga J. M., Holz D. E., 2022, “Spectral Sirens: Cosmology from the Full Mass Distribution of Compact Binaries” *Phys. Rev. Lett.*, 129, 061102 ([arXiv:2202.08240](https://arxiv.org/abs/2202.08240))
- Falcón-Barroso J., Sánchez-Blázquez P., Vazdekis A., Ricciardelli E., Cardiel N., Cenarro A. J., Gorgas J., Peletier R. F., 2011, “An updated MILES stellar library and stellar population models” *A&A*, 532, A95 ([arXiv:1107.2303](https://arxiv.org/abs/1107.2303))
- Fanfani V., 2019, *mathesis*, University of Bologna, <https://amslaurea.unibo.it/17900/>
- Farr W. M., 2019, “Accuracy Requirements for Empirically Measured Selection Functions” *Research Notes of the American Astronomical Society*, 3, 66 ([arXiv:1904.10879](https://arxiv.org/abs/1904.10879))
- Farr W. M., Fishbach M., Ye J., Holz D. E., 2019, “A Future Percent-level Measurement of the Hubble Expansion at Redshift 0.8 with Advanced LIGO” *ApJ*, 883, L42 ([arXiv:1908.09084](https://arxiv.org/abs/1908.09084))
- Feeney S. M., Peiris H. V., Williamson A. R., Nissanke S. M., Mortlock D. J., Alsing J., Scolnic D., 2019, “Prospects for Resolving the Hubble Constant Tension with Standard Sirens” *Phys. Rev. Lett.*, 122, 061105 ([arXiv:1802.03404](https://arxiv.org/abs/1802.03404))
- Finke A., Foffa S., Iacovelli F., Maggiore M., Mancarella M., 2021, “Cosmology with LIGO/Virgo dark sirens: Hubble parameter and modified gravitational wave propagation” *J. Cosmology Astropart. Phys.*, 2021, 026 ([arXiv:2101.12660](https://arxiv.org/abs/2101.12660))

- Fishbach M., Holz D. E., Farr W. M., 2018, “Does the Black Hole Merger Rate Evolve with Redshift?” *ApJ*, 863, L41 ([arXiv:1805.10270](#))
- Fishbach M., et al., 2019, “A Standard Siren Measurement of the Hubble Constant from GW170817 without the Electromagnetic Counterpart” *ApJ*, 871, L13 ([arXiv:1807.05667](#))
- Fixsen D. J., 2009, “The Temperature of the Cosmic Microwave Background” *ApJ*, 707, 916 ([arXiv:0911.1955](#))
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, “emcee: The MCMC Hammer” *PASP*, 125, 306 ([arXiv:1202.3665](#))
- Foreman-Mackey D., et al., 2019, “emcee v3: A Python ensemble sampling toolkit for affine-invariant MCMC” *Journal of Open Source Software*, 2 4(43), 1864 (2019), 4, 1864 ([arXiv:1911.07688](#))
- Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2015a, “The MICE Grand Challenge light-cone simulation - III. Galaxy lensing mocks from all-sky lensing maps” *MNRAS*, 447, 1319 ([arXiv:1312.2947](#))
- Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015b, “The MICE grand challenge lightcone simulation - I. Dark matter clustering” *MNRAS*, 448, 2987 ([arXiv:1312.1707](#))
- Franchino-Viñas S. A., Mosquera M. E., 2021, “The cosmological lithium problem, varying constants and the H_0 tension” *arXiv e-prints*, p. arXiv:2107.02243 ([arXiv:2107.02243](#))
- Franzetti P., et al., 2007, “The VIMOS-VLT deep survey. Color bimodality and the mix of galaxy populations up to $z \sim 2$ ” *A&A*, 465, 711 ([arXiv:astro-ph/0607075](#))
- Freedman W. L., et al., 2019, “The Carnegie-Chicago Hubble Program. VIII. An Independent Determination of the Hubble Constant Based on the Tip of the Red Giant Branch” *ApJ*, 882, 34 ([arXiv:1907.05922](#))
- Friedmann A., 1922, “Über die Krümmung des Raumes” *Zeitschrift für Physik*, 10, 377
- Gair J. R., Moore C. J., 2015, “Quantifying and mitigating bias in inference on gravitational wave source populations” *Phys. Rev. D*, 91, 124062 ([arXiv:1504.02767](#))
- Gair J. R., et al., 2023, “The Hitchhiker’s guide to the galaxy catalog approach for gravitational wave cosmology” *The Astronomical Journal*, 166, 22 ([arXiv:2212.08694](#))
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, “The ages and metallicities of galaxies in the local universe” *MNRAS*, 362, 41 ([arXiv:astro-ph/0506539](#))
- Gallazzi A., Bell E. F., Zibetti S., Brinchmann J., Kelson D. D., 2014, “Charting the Evolution of the Ages and Metallicities of Massive Galaxies since $z = 0.7$ ” *ApJ*, 788, 72 ([arXiv:1404.5624](#))

- Gallazzi A. R., Pasquali A., Zibetti S., Barbera F. L., 2021, “Galaxy evolution across environments as probed by the ages, stellar metallicities, and $[\alpha / \text{Fe}]$ of central and satellite galaxies” *MNRAS*, 502, 4457 ([arXiv:2010.04733](#))
- Gariazzo S., de Salas P. F., Pastor S., 2019, “Thermalisation of sterile neutrinos in the early universe in the 3+1 scheme with full mixing matrix” *J. Cosmology Astropart. Phys.*, 2019, 014 ([arXiv:1905.11290](#))
- Garilli B., et al., 2021, “The VANDELS ESO public spectroscopic survey. Final data release of 2087 spectra and spectroscopic measurements” *A&A*, 647, A150 ([arXiv:2101.07645](#))
- Gavazzi G., Bonfanti C., Sanvito G., Boselli A., Scodreggio M., 2002, “Spectrophotometry of Galaxies in the Virgo Cluster. I. The Star Formation History” *ApJ*, 576, 135
- Ghirlanda G., et al., 2019, “Compact radio emission indicates a structured jet was produced by a binary neutron star merger” *Science*, 363, 968 ([arXiv:1808.00469](#))
- Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, “HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere” *ApJ*, 622, 759
- Gray R., et al., 2020, “Cosmological inference using gravitational wave standard sirens: A mock data analysis” *Phys. Rev. D*, 101, 122001 ([arXiv:1908.06050](#))
- Gray R., Messenger C., Veitch J., 2022, “A pixelated approach to galaxy catalogue incompleteness: improving the dark siren measurement of the Hubble constant” *Monthly Notices of the Royal Astronomical Society*, 512, 1127 ([arXiv:2111.04629](#))
- Gray R., et al., 2023, “Joint cosmological and gravitational-wave population inference using dark sirens and galaxy catalogues” *J. Cosmology Astropart. Phys.*, 2023, 023 ([arXiv:2308.02281](#))
- Guth A. H., 1981, “Inflationary universe: A possible solution to the horizon and flatness problems” *Phys. Rev. D*, 23, 347
- Hallinan G., et al., 2017, “A radio counterpart to a neutron star merger” *Science*, 358, 1579 ([arXiv:1710.05435](#))
- Hewish A., Bell S. J., Pilkington J. D. H., Scott P. F., Collins R. A., 1968, “Observation of a Rapidly Pulsating Radio Source” *Nature*, 217, 709
- Hinshaw G., et al., 2013, “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results” *ApJS*, 208, 19 ([arXiv:1212.5226](#))
- Hoaglin D. C., Mosteller F., Tukey J. W., 1983, Understanding robust and exploratory data analysis. WILEY SERIES in PROBABILITY and STATISTICS: PROBABILITY and STATISTICS SECTION Series, Wiley, <https://ui.adsabs.harvard.edu/abs/1983ured.book.....H>

- Hoffmann K., Bel J., Gaztañaga E., Crocce M., Fosalba P., Castander F. J., 2015, “Measuring the growth of matter fluctuations with third-order galaxy correlations” *MNRAS*, 447, 1724 ([arXiv:1403.1259](#))
- Hogg D. W., 1999, “Distance measures in cosmology” *arXiv e-prints*, pp astro-ph/9905116 ([arXiv:astro-ph/9905116](#))
- Hogg D. W., Foreman-Mackey D., 2018, “Data Analysis Recipes: Using Markov Chain Monte Carlo” *ApJS*, 236, 11 ([arXiv:1710.06068](#))
- Holz D. E., Hughes S. A., 2005, “Using Gravitational-Wave Standard Sirens” *ApJ*, 629, 15 ([arXiv:astro-ph/0504616](#))
- Howlett C., Davis T. M., 2020, “Standard siren speeds: improving velocities in gravitational-wave measurements of H_0 ” *MNRAS*, 492, 3803 ([arXiv:1909.00587](#))
- Hubble E., 1929, “A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae” *Proceedings of the National Academy of Science*, 15, 168
- Hubble E., 1936, *The Realm of the Nebulae*. Yale University Press, New Haven
- Hulse R. A., 1994, “The discovery of the binary pulsar” *Reviews of Modern Physics*, 66, 699
- Husa S., Khan S., Hannam M., Pürrer M., Ohme F., Forteza X. J., Bohé A., 2016, “Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal” *Phys. Rev. D*, 93, 044006 ([arXiv:1508.07250](#))
- Iacovelli F., Mancarella M., Foffa S., Maggiore M., 2022a, “GWFAST: A Fisher Information Matrix Python Code for Third-generation Gravitational-wave Detectors” *ApJS*, 263, 2 ([arXiv:2207.06910](#))
- Iacovelli F., Mancarella M., Foffa S., Maggiore M., 2022b, “Forecasting the Detection Capabilities of Third-generation Gravitational-wave Detectors Using GWFAST” *ApJ*, 941, 208 ([arXiv:2207.02771](#))
- Ilbert O., et al., 2009, “Cosmos Photometric Redshifts with 30-Bands for 2-deg²” *ApJ*, 690, 1236 ([arXiv:0809.2101](#))
- Ilbert O., et al., 2013, “Mass assembly in quiescent and star-forming galaxies since $z = 4$ from UltraVISTA” *A&A*, 556, A55 ([arXiv:1301.3157](#))
- Ilbert O., et al., 2015, “Evolution of the specific star formation rate function at $z < 1.4$ Dissecting the mass-SFR plane in COSMOS and GOODS” *A&A*, 579, A2 ([arXiv:1410.4875](#))
- Ivezić Ž., et al., 2019, “LSST: From Science Drivers to Reference Design and Anticipated Data Products” *ApJ*, 873, 111 ([arXiv:0805.2366](#))
- Jedamzik K., Pogosian L., 2020, “Relieving the Hubble Tension with Primordial Magnetic Fields” *Phys. Rev. Lett.*, 125, 181302 ([arXiv:2004.09487](#))

- Jiao K., Borghi N., Moresco M., Zhang T.-J., 2023, “New Observational $H(z)$ Data from Full-spectrum Fitting of Cosmic Chronometers in the LEGA-C Survey” *ApJS*, 265, 48 ([arXiv:2205.05701](#))
- Jimenez R., Loeb A., 2002, “Constraining Cosmological Parameters Based on Relative Galaxy Ages” *ApJ*, 573, 37 ([arXiv:astro-ph/0106145](#))
- Jimenez R., Cimatti A., Verde L., Moresco M., Wandelt B., 2019, “The local and distant Universe: stellar ages and H_0 ” *J. Cosmology Astropart. Phys.*, 2019, 043 ([arXiv:1902.07081](#))
- Jimenez R., Moresco M., Verde L., Wandelt B. D., 2023, “Cosmic Chronometers with Photometry: a new path to $H(z)$ ” *arXiv e-prints*, p. arXiv:2306.11425 ([arXiv:2306.11425](#))
- Johansson J., Thomas D., Maraston C., 2012, “Chemical element ratios of Sloan Digital Sky Survey early-type galaxies” *MNRAS*, 421, 1908 ([arXiv:1112.0322](#))
- Jørgensen I., Chiboucas K., 2013, “Stellar Populations and Evolution of Early-type Cluster Galaxies: Constraints from Optical Imaging and Spectroscopy of $z = 0.5-0.9$ Galaxy Clusters” *AJ*, 145, 77 ([arXiv:1301.3177](#))
- Kalaghatgi C., Hannam M., Raymond V., 2020, “Parameter estimation with a spinning multimode waveform model” *Phys. Rev. D*, 101, 103004 ([arXiv:1909.10010](#))
- Karwal T., Kamionkowski M., 2016, “Dark energy at early times, the Hubble parameter, and the string axiverse” *Phys. Rev. D*, 94, 103523 ([arXiv:1608.01309](#))
- Kauffmann G., et al., 2003, “Stellar masses and star formation histories for 10^5 galaxies from the Sloan Digital Sky Survey” *MNRAS*, 341, 33 ([arXiv:astro-ph/0204055](#))
- Khan S., Husa S., Hannam M., Ohme F., Pürrer M., Forteza X. J., Bohé A., 2016, “Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era” *Phys. Rev. D*, 93, 044007 ([arXiv:1508.07253](#))
- Kochanek C. S., et al., 2001, “The K-Band Galaxy Luminosity Function” *ApJ*, 560, 566 ([arXiv:astro-ph/0011456](#))
- Korn A. J., Maraston C., Thomas D., 2005, “The sensitivity of Lick indices to abundance variations” *A&A*, 438, 685 ([arXiv:astro-ph/0504574](#))
- LIGO Scientific Collaboration 2018, LIGO Algorithm Library - LALSuite, free software (GPL), [doi:10.7935/GT1W-FZ16](https://doi.org/10.7935/GT1W-FZ16)
- LSST Science Collaboration et al., 2009, “LSST Science Book, Version 2.0” *arXiv e-prints*, p. arXiv:0912.0201 ([arXiv:0912.0201](#))
- Laghi D., Del Pozzo W., 2020, cosmolisa, <https://github.com/wdpozzo/cosmolisa>
- Laigle C., et al., 2016, “The COSMOS2015 Catalog: Exploring the $1 < z < 6$ Universe with Half a Million Galaxies” *ApJS*, 224, 24 ([arXiv:1604.02350](#))

- Laureijs R., et al., 2011, “Euclid Definition Study Report” *arXiv e-prints*, p. arXiv:1110.3193 ([arXiv:1110.3193](#))
- Lee H.-c., et al., 2009, “Stellar Population Models and Individual Element Abundances. II. Stellar Spectra and Integrated Light Models” *ApJ*, 694, 902 ([arXiv:0812.1563](#))
- Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, “How to Measure Galaxy Star Formation Histories. II. Nonparametric Models” *ApJ*, 876, 3 ([arXiv:1811.03637](#))
- Lemaître G., 1927, “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*, 47, 49
- Leonardi A. J., Rose J. A., 1996, “The Ages of Starbursts in Post-Starburst Galaxies” *AJ*, 111, 182 ([arXiv:astro-ph/9510033](#))
- Lesgourgues J., 2011, “The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview” *arXiv e-prints*, p. arXiv:1104.2932 ([arXiv:1104.2932](#))
- Lesgourgues J., Mangano G., Miele G., Pastor S., 2013, *Neutrino Cosmology*. Cambridge University Press, doi:10.1017/cbo9781139012874, <https://ui.adsabs.harvard.edu/abs/2013neco.book.....L>
- Leyde K., Mastrogiovanni S., Steer D. A., Chassande-Mottin E., Karathanasis C., 2022, “Current and future constraints on cosmology and modified gravitational wave friction from binary black holes” *J. Cosmology Astropart. Phys.*, 2022, 012 ([arXiv:2203.11680](#))
- Li X., Shafieloo A., 2019, “A Simple Phenomenological Emergent Dark Energy Model can Resolve the Hubble Tension” *ApJ*, 883, L3 ([arXiv:1906.08275](#))
- Linde A. D., 1982, “A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems” *Physics Letters B*, 108, 389
- Linder E. V., 2003, “Exploring the Expansion History of the Universe” *Phys. Rev. Lett.*, 90, 091301 ([arXiv:astro-ph/0208512](#))
- London L., et al., 2018, “First Higher-Multipole Model of Gravitational Waves from Spinning and Coalescing Black-Hole Binaries” *Phys. Rev. Lett.*, 120, 161102 ([arXiv:1708.00404](#))
- Longhetti M., Bressan A., Chiosi C., Rampazzo R., 1999, “Star formation history of early-type galaxies in low density environments. V. Blue line-strength indices for the nuclear region” *A&A*, 345, 419 ([arXiv:astro-ph/9903315](#))
- Lonoce I., Longhetti M., Saracco P., Gargiulo A., Tamburri S., 2014, “Spectral detection of multiple stellar populations in $z \sim 1$ early-type galaxies” *MNRAS*, 444, 2048 ([arXiv:1409.1029](#))
- Lonoce I., Maraston C., Thomas D., Longhetti M., Parikh T., Guarnieri P., Comparat J., 2020, “Stellar population properties of individual massive early-type galaxies at $1.4 < z < 2$ ” *MNRAS*, 492, 326 ([arXiv:1912.00084](#))

- Loredo T. J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, American Institute of Physics Conference Series Vol. 735, AIP Conference Proceedings. AIP, pp 195–206 ([arXiv:astro-ph/0409387](#)), [doi:10.1063/1.1835214](#), [https://ui.adsabs.harvard.edu/abs/2004AIPC...735..195L](#)
- Loredo T. J., Lamb D. Q., 2002, “Bayesian analysis of neutrinos observed from supernova SN 1987A” *Phys. Rev. D*, 65, 063002 ([arXiv:astro-ph/0107260](#))
- Madau P., Dickinson M., 2014, “Cosmic Star-Formation History” *ARA&A*, 52, 415 ([arXiv:1403.0007](#))
- Maggiore M., 2007, Gravitational Waves: Volume 1: Theory and Experiments. Oxford University Press/Oxford, [doi:10.1093/acprof:oso/9780198570745.001.0001](#)
- Maggiore M., 2014, “Phantom dark energy from nonlocal infrared modifications of general relativity” *Phys. Rev. D*, 89, 043008 ([arXiv:1307.3898](#))
- Maggiore M., et al., 2020, “Science case for the Einstein telescope” *J. Cosmology Astropart. Phys.*, 2020, 050 ([arXiv:1912.02622](#))
- Mancarella M., Borghi N., Foffa S., Genoud-Prachex E., Iacovelli F., Maggiore M., Moresco M., Schulz M., 2022a, in Proceedings of 41st International Conference on High Energy physics - PoS(ICHEP2022). Sissa Medialab, [doi:10.22323/1.414.0127](#)
- Mancarella M., Genoud-Prachex E., Maggiore M., 2022b, “Cosmology and modified gravitational wave propagation from binary black hole population models” *Phys. Rev. D*, 105, 064030 ([arXiv:2112.05728](#))
- Mandel I., Farr W. M., Gair J. R., 2019, “Extracting distribution parameters from multiple uncertain observations with selection biases” *MNRAS*, 486, 1086 ([arXiv:1809.02063](#))
- Maraston C., 2005, “Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies” *MNRAS*, 362, 799 ([arXiv:astro-ph/0410207](#))
- Maraston C., Nieves Colmenárez L., Bender R., Thomas D., 2009, “Absorption line indices in the UV. I. Empirical and theoretical stellar population models” *A&A*, 493, 425 ([arXiv:0811.0619](#))
- Maraston C., et al., 2020, “Stellar population models based on the SDSS-IV MaStar library of stellar spectra - I. Intermediate-age/old models” *MNRAS*, 496, 2962 ([arXiv:1911.05748](#))
- Mastrogiovanni S., et al., 2021a, “On the importance of source population models for gravitational-wave cosmology” *Phys. Rev. D*, 104, 062009 ([arXiv:2103.14663](#))
- Mastrogiovanni S., Haegel L., Karathanasis C., Magaña Hernandez I., Steer D. A., 2021b, “Gravitational wave friction in light of GW170817 and GW190521” *J. Cosmology Astropart. Phys.*, 2021, 043 ([arXiv:2010.04047](#))

- Mastrogiovanni S., et al., 2023, “Joint population and cosmological properties inference with gravitational waves standard sirens and galaxy surveys” *Phys. Rev. D*, 108, 042002 ([arXiv:2305.10488](#))
- McDermid R. M., et al., 2015, “The ATLAS^{3D} Project - XXX. Star formation histories and stellar population scaling relations of early-type galaxies” *MNRAS*, 448, 3484 ([arXiv:1501.03723](#))
- McLure R. J., et al., 2018, “The VANDELS ESO public spectroscopic survey” *MNRAS*, 479, 25 ([arXiv:1803.07414](#))
- Messenger C., Read J., 2012, “Measuring a Cosmological Distance-Redshift Relationship Using Only Gravitational Wave Observations of Binary Neutron Star Coalescences” *Phys. Rev. Lett.*, 108, 091101 ([arXiv:1107.5725](#))
- Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, “Equation of State Calculations by Fast Computing Machines” *J. Chem. Phys.*, 21, 1087
- Mignoli M., et al., 2009, “The zCOSMOS redshift survey: the three-dimensional classification cube and bimodality in galaxy physical properties” *A&A*, 493, 39 ([arXiv:0810.2245](#))
- Mishra C. K., Kela A., Arun K. G., Faye G., 2016, “Ready-to-use post-Newtonian gravitational waveforms for binary black holes with nonprecessing spins: An update” *Phys. Rev. D*, 93, 084054 ([arXiv:1601.05588](#))
- Moresco M., 2015, “Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at $z \gtrsim 2$.” *MNRAS*, 450, L16 ([arXiv:1503.01116](#))
- Moresco M., Jimenez R., Cimatti A., Pozzetti L., 2011, “Constraining the expansion rate of the Universe using low-redshift ellipticals as cosmic chronometers” *J. Cosmology Astropart. Phys.*, 2011, 045 ([arXiv:1010.0831](#))
- Moresco M., Verde L., Pozzetti L., Jimenez R., Cimatti A., 2012a, “New constraints on cosmological parameters and neutrino properties using the expansion rate of the Universe to $z \sim 1.75$ ” *J. Cosmology Astropart. Phys.*, 2012, 053 ([arXiv:1201.6658](#))
- Moresco M., et al., 2012b, “Improved constraints on the expansion rate of the Universe up to $z \sim 1.1$ from the spectroscopic evolution of cosmic chronometers” *J. Cosmology Astropart. Phys.*, 2012, 006 ([arXiv:1201.3609](#))
- Moresco M., et al., 2013, “Spot the difference. Impact of different selection criteria on observed properties of passive galaxies in zCOSMOS-20k sample” *A&A*, 558, A61 ([arXiv:1305.1308](#))
- Moresco M., et al., 2016, “A 6% measurement of the Hubble parameter at $z \sim 0.45$: direct evidence of the epoch of cosmic re-acceleration” *J. Cosmology Astropart. Phys.*, 2016, 014 ([arXiv:1601.01701](#))

- Moresco M., Jimenez R., Verde L., Pozzetti L., Cimatti A., Citro A., 2018, “Setting the Stage for Cosmic Chronometers. I. Assessing the Impact of Young Stellar Populations on Hubble Parameter Measurements” *ApJ*, 868, 84 ([arXiv:1804.05864](#))
- Moresco M., Jimenez R., Verde L., Cimatti A., Pozzetti L., 2020, “Setting the Stage for Cosmic Chronometers. II. Impact of Stellar Population Synthesis Models Systematics and Full Covariance Matrix” *ApJ*, 898, 82 ([arXiv:2003.07362](#))
- Moresco M., et al., 2022, “Unveiling the Universe with emerging cosmological probes” *Living Reviews in Relativity*, 25, 6 ([arXiv:2201.07241](#))
- Mukherjee S., Wandelt B. D., Silk J., 2020, “Probing the theory of gravity with gravitational lensing of gravitational waves and galaxy surveys” *MNRAS*, 494, 1956 ([arXiv:1908.08951](#))
- Mukherjee S., Lavaux G., Bouchet F. R., Jasche J., Wandelt B. D., Nisanke S., Leclercq F., Hotokezaka K., 2021, “Velocity correction for Hubble constant measurements from standard sirens” *A&A*, 646, A65 ([arXiv:1909.08627](#))
- Muttoni N., Laghi D., Tamanini N., Marsat S., Izquierdo-Villalba D., 2023, “Dark siren cosmology with binary black holes in the era of third-generation gravitational wave detectors” *Phys. Rev. D*, 108, 043543 ([arXiv:2303.10693](#))
- Myles J., et al., 2021, “Dark Energy Survey Year 3 results: redshift calibration of the weak lensing source galaxies” *MNRAS*, 505, 4249 ([arXiv:2012.08566](#))
- Niedermann F., Sloth M. S., 2021, “New early dark energy” *Phys. Rev. D*, 103, L041303
- Nisanke S., Holz D. E., Hughes S. A., Dalal N., Sievers J. L., 2010, “Exploring Short Gamma-ray Bursts as Gravitational-wave Standard Sirens” *ApJ*, 725, 496 ([arXiv:0904.1017](#))
- Oguri M., 2016, “Measuring the distance-redshift relation with the cross-correlation of gravitational wave standard sirens and galaxies” *Phys. Rev. D*, 93, 083511 ([arXiv:1603.02356](#))
- Onodera M., et al., 2012, “Deep Near-infrared Spectroscopy of Passively Evolving Galaxies at $z > 1.4$ ” *ApJ*, 755, 26 ([arXiv:1206.1540](#))
- Onodera M., et al., 2015, “The Ages, Metallicities, and Element Abundance Ratios of Massive Quenched Galaxies at $z = 1.6$ ” *ApJ*, 808, 161 ([arXiv:1411.5023](#))
- Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, “The Two Phases of Galaxy Formation” *ApJ*, 725, 2312 ([arXiv:1010.1381](#))
- Pacini F., 1968, “Rotating Neutron Stars, Pulsars and Supernova Remnants” *Nature*, 219, 145
- Pakmor R., et al., 2023, “The MillenniumTNG Project: the hydrodynamical full physics simulation and a first look at its galaxy clusters” *MNRAS*, 524, 2539 ([arXiv:2210.10060](#))
- Palmese A., et al., 2020, “A Statistical Standard Siren Measurement of the Hubble Constant from the LIGO/Virgo Gravitational Wave Compact Object Merger GW190814 and Dark Energy Survey Galaxies” *ApJ*, 900, L33 ([arXiv:2006.14961](#))

- Pannarale F., Berti E., Kyutoku K., Lackey B. D., Shibata M., 2015, “Aligned spin neutron star-black hole mergers: A gravitational waveform amplitude model” *Phys. Rev. D*, 92, 084050 ([arXiv:1509.00512](#))
- Pasquini L., Delabre B., Ellis R., de Zeeuw T., 2016, in Hall H. J., Gilmozzi R., Marshall H. K., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9906, Ground-based and Airborne Telescopes VI. SPIE, p. 99063C ([arXiv:1606.06494](#)), [doi:10.1117/12.2232289](#)
- Pedregosa F., et al., 2011, “Scikit-learn: Machine Learning in Python” *Journal of Machine Learning Research*, 12, 2825
- Peebles P. J. E., Yu J. T., 1970, “Primeval Adiabatic Perturbation in an Expanding Universe” *ApJ*, 162, 815
- Peng Y.-j., et al., 2010, “Mass and Environment as Drivers of Galaxy Evolution in SDSS and zCOSMOS and the Origin of the Schechter Function” *ApJ*, 721, 193 ([arXiv:1003.4747](#))
- Peng Y., Maiolino R., Cochrane R., 2015, “Strangulation as the primary mechanism for shutting down star formation in galaxies” *Nature*, 521, 192 ([arXiv:1505.03143](#))
- Pentericci L., et al., 2018, “The VANDELS ESO public spectroscopic survey: Observations and first data release” *A&A*, 616, A174 ([arXiv:1803.07373](#))
- Penzias A. A., Wilson R. W., 1965, “A Measurement of Excess Antenna Temperature at 4080 Mc/s.” *ApJ*, 142, 419
- Percival W. J., Cole S., Eisenstein D. J., Nichol R. C., Peacock J. A., Pope A. C., Szalay A. S., 2007, “Measuring the Baryon Acoustic Oscillation scale using the Sloan Digital Sky Survey and 2dF Galaxy Redshift Survey” *MNRAS*, 381, 1053 ([arXiv:0705.3323](#))
- Perivolaropoulos L., Skara F., 2022, “Challenges for Λ CDM: An update” *New A Rev.*, 95, 101659 ([arXiv:2105.05208](#))
- Perlmutter S., et al., 1999, “Measurements of Ω and Λ from 42 High-Redshift Supernovae” *ApJ*, 517, 565 ([arXiv:astro-ph/9812133](#))
- Phillips M. M., 1993, “The Absolute Magnitudes of Type IA Supernovae” *ApJ*, 413, L105
- Pian E., et al., 2017, “Spectroscopic identification of r-process nucleosynthesis in a double neutron-star merger” *Nature*, 551, 67 ([arXiv:1710.05858](#))
- Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2006, “A Large Stellar Evolution Database for Population Synthesis Studies. II. Stellar Models and Isochrones for an α -enhanced Metal Distribution” *ApJ*, 642, 797 ([arXiv:astro-ph/0603721](#))
- Pirani F. A. E., 1956, “On the physical significance of the Riemann tensor” *Acta Physica Polonica*, 15, 389
- Planck Collaboration 2016, “Planck 2015 results. XIII. Cosmological parameters” *A&A*, 594, A13 ([arXiv:1502.01589](#))

- Planck Collaboration 2020, “Planck 2018 results. VI. Cosmological parameters” *A&A*, 641, A6 ([arXiv:1807.06209](#))
- Planck Collaboration et al., 2020, “Planck 2018 results. VII. Isotropy and statistics of the CMB” *A&A*, 641, A7 ([arXiv:1906.02552](#))
- Postman M., et al., 2012, “The Cluster Lensing and Supernova Survey with Hubble: An Overview” *ApJS*, 199, 25 ([arXiv:1106.3328](#))
- Poulin V., Serpico P. D., Lesgourgues J., 2016, “A fresh look at linear cosmological constraints on a decaying Dark Matter component” *J. Cosmology Astropart. Phys.*, 2016, 036 ([arXiv:1606.02073](#))
- Pozzetti L., Mannucci F., 2000, “Extremely red galaxies: age and dust degeneracy solved?” *MNRAS*, 317, L17 ([arXiv:astro-ph/0006430](#))
- Pozzetti L., et al., 2010, “zCOSMOS - 10k-bright spectroscopic sample. The bimodality in the galaxy stellar mass function: exploring its evolution with redshift” *A&A*, 523, A13 ([arXiv:0907.5416](#))
- Punturo M., et al., 2010, “The Einstein Telescope: a third-generation gravitational wave observatory” *Classical and Quantum Gravity*, 27, 194002
- Quartin M., Marra V., Amendola L., 2014, “Accurate weak lensing of standard candles. II. Measuring σ_8 with supernovae” *Phys. Rev. D*, 89, 023009 ([arXiv:1307.1155](#))
- R. W., 1972, “Electromagnetically Coupled Broadband Gravitational Antenna” Quarterly Progress Report, Research Laboratory of Electronics, MIT, 105
- Ratsimbazafy A. L., Loubser S. I., Crawford S. M., Cress C. M., Bassett B. A., Nichol R. C., Väisänen P., 2017, “Age-dating luminous red galaxies observed with the Southern African Large Telescope” *MNRAS*, 467, 3239 ([arXiv:1702.00418](#))
- Refsdal S., 1964, “On the possibility of determining Hubble’s parameter and the masses of galaxies from the gravitational lens effect” *MNRAS*, 128, 307
- Reitze D., et al., 2019, “Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO” *Bulletin of the AAS*, 51, 35 ([arXiv:1907.04833](#))
- Renzini A., 2006, “Stellar Population Diagnostics of Elliptical Galaxy Formation” *ARA&A*, 44, 141 ([arXiv:astro-ph/0603479](#))
- Riess A. G., et al., 1998, “Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant” *AJ*, 116, 1009 ([arXiv:astro-ph/9805201](#))
- Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, “Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond Λ CDM” *ApJ*, 876, 85 ([arXiv:1903.07603](#))

- Riess A. G., Casertano S., Yuan W., Bowers J. B., Macri L., Zinn J. C., Scolnic D., 2021, “Cosmic Distances Calibrated to 1% Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with Λ CDM” *ApJ*, 908, L6 ([arXiv:2012.08534](#))
- Riess A. G., et al., 2022, “A Comprehensive Measurement of the Local Value of the Hubble Constant with $1 \text{ km s}^{-1} \text{ Mpc}^{-1}$ Uncertainty from the Hubble Space Telescope and the SH0ES Team” *ApJ*, 934, L7 ([arXiv:2112.04510](#))
- Rinaldi S., Del Pozzo W., Mapelli M., Medina A. L., Dent T., 2023, “Evidence for the evolution of black hole mass function with redshift” *arXiv e-prints*, p. arXiv:2310.03074 ([arXiv:2310.03074](#))
- Risaliti G., Lusso E., 2015, “A Hubble Diagram for Quasars” *ApJ*, 815, 33 ([arXiv:1505.07118](#))
- Rose J. A., 1984, “Spectral anomalies in the Hyades and Pleiades and in field stars with active chromospheres.” *AJ*, 89, 1238
- Rose J. A., 1985, “Constraints on stellar populations in elliptical galaxies.” *AJ*, 90, 1927
- Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, “The clustering of the SDSS DR7 main Galaxy sample - I. A 4 per cent distance measure at $z = 0.15$ ” *MNRAS*, 449, 835 ([arXiv:1409.3242](#))
- Rousseeuw P. J., 1984, “Least Median of Squares Regression” *Journal of the American Statistical Association*, 79, 871
- Roy V., 2020, “Convergence Diagnostics for Markov Chain Monte Carlo” *Annual Review of Statistics and Its Application*, 7, 387 ([arXiv:1909.11827](#))
- Rubin V. C., Ford W. Kent J., 1970, “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions” *ApJ*, 159, 379
- Salpeter E. E., 1955, “The Luminosity Function and Stellar Evolution.” *ApJ*, 121, 161
- Sánchez-Blázquez P., et al., 2006, “Medium-resolution Isaac Newton Telescope library of empirical spectra” *MNRAS*, 371, 703 ([arXiv:astro-ph/0607009](#))
- Sandage A., 1962, “The Change of Redshift and Apparent Luminosity of Galaxies due to the Deceleration of Selected Expanding Universes.” *ApJ*, 136, 319
- Santoliquido F., Mapelli M., Artale M. C., Boco L., 2022, “Modelling the host galaxies of binary compact object mergers with observational scaling relations” *MNRAS*, 516, 3297 ([arXiv:2205.05099](#))
- Scarlata C., et al., 2007, “COSMOS Morphological Classification with the Zurich Estimator of Structural Types (ZEST) and the Evolution Since $z = 1$ of the Luminosity Function of Early, Disk, and Irregular Galaxies” *ApJS*, 172, 406

- Schöneberg N., Abellán G. F., Sánchez A. P., Witte S. J., Poulin V., Lesgourgues J., 2022, “The H_0 Olympics: A fair ranking of proposed models” *Phys. Rep.*, 984, 1 ([arXiv:2107.10291](#))
- Schutz B. F., 1986, “Determining the Hubble constant from gravitational wave observations” *Nature*, 323, 310
- Scolnic D. M., et al., 2018, “The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample” *ApJ*, 859, 101 ([arXiv:1710.00845](#))
- Scott D. W., 1979, “On optimal and data-based histograms” *Biometrika*, 66, 605
- Scott N., et al., 2017, “The SAMI Galaxy Survey: global stellar populations on the size-mass plane” *MNRAS*, 472, 2833 ([arXiv:1708.06849](#))
- Seiguchi T., Takahashi T., 2021, “Early recombination as a solution to the H_0 tension” *Phys. Rev. D*, 103, 083507 ([arXiv:2007.03381](#))
- Simon J., Verde L., Jimenez R., 2005, “Constraints on the redshift dependence of the dark energy potential” *Phys. Rev. D*, 71, 123001 ([arXiv:astro-ph/0412269](#))
- Singh R., et al., 2013, “The nature of LINER galaxies: Ubiquitous hot old stars and rare accreting black holes” *A&A*, 558, A43 ([arXiv:1308.4271](#))
- Siudek M., et al., 2017, “The VIMOS Public Extragalactic Redshift Survey (VIPERS). Star formation history of passive red galaxies” *A&A*, 597, A107 ([arXiv:1605.05503](#))
- Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, American Institute of Physics Conference Series Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering. pp 395–405, [doi:10.1063/1.1835238](#), <https://ui.adsabs.harvard.edu/abs/2004AIPC..735..395S>
- Smoot G. F., et al., 1992, “Structure in the COBE Differential Microwave Radiometer First-Year Maps” *ApJ*, 396, L1
- Soares-Santos M., et al., 2019, “First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo Binary–Black-hole Merger GW170814” *The Astrophysical Journal*, 876, L7
- Spera M., Mapelli M., 2017, “Very massive stars, pair-instability supernovae and intermediate-mass black holes with the `sevn` code” *MNRAS*, 470, 4739 ([arXiv:1706.06109](#))
- Spergel D., et al., 2015, “Wide-Field Infrared Survey Telescope–Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report” arXiv e-prints, p. [arXiv:1503.03757](#) ([arXiv:1503.03757](#))
- Spilker J., et al., 2018, “Molecular Gas Contents and Scaling Relations for Massive, Passive Galaxies at Intermediate Redshifts from the LEGA-C Survey” *ApJ*, 860, 103 ([arXiv:1805.02667](#))

- Spinrad H., Dey A., Stern D., Dunlop J., Peacock J., Jimenez R., Windhorst R., 1997, “LBDS 53W091: An Old, Red Galaxy at $z = 1.552$ ” *ApJ*, 484, 581 ([arXiv:astro-ph/9702233](#))
- Stern D., Jimenez R., Verde L., Kamionkowski M., Stanford S. A., 2010, “Cosmic chronometers: constraining the equation of state of dark energy. I: $H(z)$ measurements” *J. Cosmology Astropart. Phys.*, 2010, 008 ([arXiv:0907.3149](#))
- Straatman C. M. S., et al., 2018, “The Large Early Galaxy Astrophysics Census (LEGA-C) Data Release 2: Dynamical and Stellar Population Properties of $z \lesssim 1$ Galaxies in the COSMOS Field” *ApJS*, 239, 27 ([arXiv:1809.08236](#))
- Strateva I., et al., 2001, “Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data” *AJ*, 122, 1861 ([arXiv:astro-ph/0107201](#))
- Sunyaev R. A., Zeldovich Y. B., 1970, “Small-Scale Fluctuations of Relic Radiation” *Ap&SS*, 7, 3
- Talbot C., Thrane E., 2018, “Measuring the Binary Black Hole Mass Spectrum with an Astrophysically Motivated Parameterization” *ApJ*, 856, 173 ([arXiv:1801.02699](#))
- Talia M., et al., 2023, “The VANDELS ESO public spectroscopic survey: The spectroscopic measurements catalogue” *A&A*, 678, A25 ([arXiv:2309.14436](#))
- Taracchini A., et al., 2014, “Effective-one-body model for black-hole binaries with generic mass ratios and spins” *Phys. Rev. D*, 89, 061502 ([arXiv:1311.2544](#))
- Taylor S. R., Gair J. R., Mandel I., 2012, “Cosmology using advanced gravitational-wave detectors alone” *Phys. Rev. D*, 85, 023535 ([arXiv:1108.5161](#))
- Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, “The Epochs of Early-Type Galaxy Formation as a Function of Environment” *ApJ*, 621, 673 ([arXiv:astro-ph/0410209](#))
- Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, “Environment and self-regulation in galaxy formation” *MNRAS*, 404, 1775 ([arXiv:0912.0259](#))
- Thomas D., Maraston C., Johansson J., 2011, “Flux-calibrated stellar population models of Lick absorption-line indices with variable element abundance ratios” *MNRAS*, 412, 2183 ([arXiv:1010.4569](#))
- Thrane E., Talbot C., 2019, “An introduction to Bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models” *PASA*, 36, e010 ([arXiv:1809.02293](#))
- Tomasetti E., et al., 2023, “A new measurement of the expansion history of the Universe at $z=1.26$ with cosmic chronometers in VANDELS” *A&A*, p. arXiv:2305.16387 ([arXiv:2305.16387](#))
- Trager S. C., Worthey G., Faber S. M., Burstein D., González J. J., 1998, “Old Stellar Populations. VI. Absorption-Line Spectra of Galaxy Nuclei and Globular Clusters” *ApJS*, 116, 1 ([arXiv:astro-ph/9712258](#))

- Treu T., et al., 2005, “The Assembly History of Field Spheroidals: Evolution of Mass-to-Light Ratios and Signatures of Recent Star Formation” *ApJ*, 633, 174 ([arXiv:astro-ph/0503164](#))
- Tripicco M. J., Bell R. A., 1995, “Modeling the LICK/IDS Spectral Feature Indices Using Synthetic Spectra” *AJ*, 110, 3035
- Troja E., et al., 2017, “The X-ray counterpart to the gravitational-wave event GW170817” *Nature*, 551, 71 ([arXiv:1710.05433](#))
- Vagnozzi S., Loeb A., Moresco M., 2021, “Eppur è piatto? The Cosmic Chronometers Take on Spatial Curvature and Cosmic Concordance” *ApJ*, 908, 84 ([arXiv:2011.11645](#))
- Vagnozzi S., Pacucci F., Loeb A., 2022, “Implications for the Hubble tension from the ages of the oldest astrophysical objects” *Journal of High Energy Astrophysics*, 36, 27 ([arXiv:2105.10421](#))
- Vazdekis A., et al., 2015, “Evolutionary stellar population synthesis with MILES - II. Scaled-solar and α -enhanced models” *MNRAS*, 449, 1177 ([arXiv:1504.08032](#))
- Verde L., Treu T., Riess A. G., 2019, “Tensions between the early and late Universe” *Nature Astronomy*, 3, 891 ([arXiv:1907.10625](#))
- Verza G., Pisani A., Carbone C., Hamaus N., Guzzo L., 2019, “The void size function in dynamical dark energy cosmologies” *J. Cosmology Astropart. Phys.*, 2019, 040 ([arXiv:1906.00409](#))
- Virtanen P., et al., 2020, “SciPy 1.0: fundamental algorithms for scientific computing in Python” *Nature Methods*, 17, 261 ([arXiv:1907.10121](#))
- Vitale S., Farr W. M., Ng K. K. Y., Rodriguez C. L., 2019, “Measuring the Star Formation Rate with Gravitational Waves from Binary Black Holes” *ApJ*, 886, L1 ([arXiv:1808.00901](#))
- Vitale S., Gerosa D., Farr W. M., Taylor S. R., 2022, in , *Handbook of Gravitational Wave Astronomy*. Springer, p. 45, doi:10.1007/978-981-15-4702-7_45-1, <https://ui.adsabs.harvard.edu/abs/2022hgwa.bookE..45V>
- Volonteri M., 2010, “Formation of supermassive black holes” *A&A Rev.*, 18, 279 ([arXiv:1003.4404](#))
- Wang Y., et al., 2019, “ATLAS probe: Breakthrough science of galaxy evolution, cosmology, Milky Way, and the Solar System” *PASA*, 36, e015 ([arXiv:1802.01539](#))
- Watson D., Denney K. D., Vestergaard M., Davis T. M., 2011, “A New Cosmological Distance Measure Using Active Galactic Nuclei” *ApJ*, 740, L49 ([arXiv:1109.4632](#))
- Weber J., 1960, “Detection and Generation of Gravitational Waves” *Physical Review*, 117, 306
- Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, “Detection of Quiescent Galaxies in a Bicolor Sequence from $Z = 0-2$ ” *ApJ*, 691, 1879 ([arXiv:0806.0625](#))

- Wong K. C., et al., 2020, “H0LiCOW - XIII. A 2.4 per cent measurement of H_0 from lensed quasars: 5.3σ tension between early- and late-Universe probes” *MNRAS*, 498, 1420 ([arXiv:1907.04869](#))
- Worthey G., 1998, “Abundance Ratio Trends and Nucleosynthesis in Elliptical Galaxies and Spheroids” *PASP*, 110, 888
- Worthey G., Ottaviani D. L., 1997, “ $H\gamma$ and $H\delta$ Absorption Features in Stars and Stellar Populations” *ApJS*, 111, 377
- Worthey G., Faber S. M., Gonzalez J. J., Burstein D., 1994, “Old Stellar Populations. V. Absorption Feature Indices for the Complete Lick/IDS Sample of Stars” *ApJS*, 94, 687
- Wu P.-F., et al., 2018, “Stellar Populations of over 1000 $z \sim 0.8$ Galaxies from LEGA-C: Ages and Star Formation Histories from D_n4000 and $H\delta$ ” *ApJ*, 855, 85 ([arXiv:1802.06799](#))
- Yan R., Newman J. A., Faber S. M., Konidaris N., Koo D., Davis M., 2006, “On the Origin of [O II] Emission in Red-Sequence and Poststarburst Galaxies” *ApJ*, 648, 281 ([arXiv:astro-ph/0512446](#))
- Zarrouk P., et al., 2018, “The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2” *MNRAS*, 477, 1639 ([arXiv:1801.03062](#))
- Zhang C., Zhang H., Yuan S., Liu S., Zhang T.-J., Sun Y.-C., 2014, “Four new observational $H(z)$ data from luminous red galaxies in the Sloan Digital Sky Survey data release seven” *Research in Astronomy and Astrophysics*, 14, 1221 ([arXiv:1207.4541](#))
- Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Gorski K., 2019, “healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python” *The Journal of Open Source Software*, 4, 1298
- Zumalacárregui M., Seljak U., 2018, “Limits on Stellar-Mass Compact Objects as Dark Matter from Gravitational Lensing of Type Ia Supernovae” *Phys. Rev. Lett.*, 121, 141101 ([arXiv:1712.02240](#))
- Zwicky F., 1933, “Die Rotverschiebung von extragalaktischen Nebeln” *Helvetica Physica Acta*, 6, 110
- de Sainte Agathe V., et al., 2019, “Baryon acoustic oscillations at $z = 2.34$ from the correlations of $Ly\alpha$ absorption in eBOSS DR14” *A&A*, 629, A85 ([arXiv:1904.03400](#))
- van Dokkum P. G., Franx M., Fabricant D., Illingworth G. D., Kelson D. D., 2000, “Hubble Space Telescope Photometry and Keck Spectroscopy of the Rich Cluster MS 1054-03: Morphologies, Butcher-Oemler Effect, and the Color-Magnitude Relation at $Z = 0.83$ ” *ApJ*, 541, 95 ([arXiv:astro-ph/0002507](#))