Cosmologie avec les sirènes sombres et populations de binaires de trous noirs avec les ondes gravitationnelles de LIGO-Virgo-KAGRA

Grégoire Pierra

Université Claude Bernard Lyon 1

THESE de DOCTORAT DE
L'UNIVERSITE CLAUDE BERNARD LYON 1

Ecole Doctorale N° 52
Physique et Astrophysique

Discipline : Physique

Soutenue publiquement le 03/10/2024, par :
Grégoire PIERRA

Dark Siren Cosmology and Binary Black Hole Populations with LIGO-Virgo-KAGRA Gravitational Waves

---

Devant le jury composé de :

Corinne, Augier	Professeure des Universités Université Claude Bernard Lyon 1	Présidente
Tamanini, Nicola	Directeur de Recherche CNRS, Laboratoire des deux infinis de Toulouse	Rapporteur
Lamberts, Astrid	Chargée de Recherche CNRS, Laboratoire Lagrange Nice	Rapporteur
Mastrogiovanni, Simone	Chercheur Université de Rome/INFN	Examinateur
Robinet, Florent	Chargé de Recherche CNRS, Laboratoire IJCLab, Paris-Saclay	Examinateur
Steer, Danièle	Professeure des Universités Université Paris Cité	Examinateur
Perriès, Stéphane	Professeur des Universités Université Claude Bernard Lyon 1	Directeur de thèse

"Who has seen the wind? Neither I nor you. But when the leaves hang trembling, the wind is passing through. Let the wind carry these wings to you."

– Hayao Miyazaki
(The Wind Rises)

RÉSUMÉ

UNIVERSITÉ CLAUDE BERNARD LYON 1

Thèse de Doctorat

Cosmologie avec les Sirènes Sombres et Populations de Binaires de Trous Noir avec les Ondes Gravitationnelles de LIGO-Virgo-KAGRA

par Grégoire PIERRA

Les ondes gravitationnelles constituent une nouvelle sonde pour explorer l'Univers et étudier des phénomènes cosmiques jusque-là inaccessibles. Cette thèse se base sur des données d'ondes gravitationnelles récoltées par les détecteurs de la collaboration scientifique LIGO-Virgo-KAGRA. La première partie de ce travail porte sur la cosmologie avec les sirènes sombres, une méthode qui utilise les fusions de trous noirs pour mesurer les distances cosmologiques et inférer la valeur de la constante de Hubble, sans nécessiter de contrepartie électromagnétique. Elle présente également icarogw, un code d'inférence bayésienne hiérarchique, utilisant les données d'ondes gravitationnelles et des modèles décrivant les propriétés astrophysiques des trous noirs, comme leurs masses, leurs distances ou encore leurs spins, pour mesurer la constante de Hubble. La deuxième partie de cette étude teste particulièrement la robustesse des sirènes sombres pour la cosmologie. Elle explore l'impact des modèles de population de trous noirs sur l'estimation de la constante de Hubble, tout particulièrement lorsque certains processus astrophysiques ne sont pas modélisés. La troisième partie s'attache à la recherche et à l'identification de sous-populations de binaires de trous noirs dans l'univers. Elle examine la manière dont différents canaux de formation peuvent influencer les caractéristiques intrinsèques de ces objets compacts, notamment à travers les corrélations potentielles entre leur masse et leur spin. L'existence de ces corrélations serait révélatrice de la présence de sous-populations de trous noirs, comme les trous noirs hiérarchiques issus de coalescences précédentes. Enfin, le manuscrit se termine par une étude sur l'utilisation de méthodes de machine learning pour améliorer la qualité des données de l'interféromètre Virgo et détecter la présence de bruits non-gaussiens. Ces travaux explorent également l'intégration potentielle des résultats d'iDQ dans les algorithmes de détection des signaux d'ondes gravitationnelles, visant ainsi à renforcer le niveau de confiance dans ces détections.

ABSTRACT

UNIVERSITÉ CLAUDE BERNARD LYON 1

Doctor of Philosophy

Dark Siren Cosmology and Binary Black Hole Populations with LIGO-Virgo-KAGRA Gravitational Waves

by Grégoire PIERRA

Gravitational waves constitute a new probe for exploring the Universe and studying cosmic phenomena that were previously inaccessible. This thesis is based on gravitational wave data collected by the LIGO-Virgo-KAGRA scientific collaboration detectors. The first part of this work focuses on cosmology with dark sirens, a method that uses binary black hole mergers to measure cosmological distances and infer the value of the Hubble constant, without requiring an electromagnetic counterpart. It also presents icarogw, a hierarchical Bayesian inference code that uses gravitational wave data and models describing the astrophysical properties of black holes, such as their masses, distances, and spins, to estimate the Hubble constant. The second part of this study tests in particular the robustness of the dark siren method for cosmology. It explores the impact of binary black hole population parameterization on the estimation of the Hubble constant, especially when certain astrophysical processes are not modelled. The third part focuses on the search and identification of subpopulations of binary black holes in the universe. It examines how different formation channels can influence the intrinsic characteristics of these compact objects, particularly through potential correlations between their mass and spin. The existence of these correlations would be indicative of the presence of subpopulations of black holes, such as hierarchical black holes, resulting from previous mergers. Finally, the manuscript concludes with a study on the use of machine learning methods to improve the quality of the Virgo interferometer data and detect the presence of non-Gaussian noises. This work also explores the potential integration of inferential data quality (iDQ) results into gravitational wave signal detection algorithms, thereby aiming to strengthen the confidence level in these detections.

REMERCIEMENTS

À mes rapporteurs, Astrid Lamberts et Nicola Tamanini, qui ont bien voulu s'atteler à la lecture de ce manuscrit (au volume non négligeable). Aux autres membres du jury: Florent Robinet, Danièle Steer, Simone Mastrogiovanni et Corinne Augier, pour avoir assisté à ma soutenance et pour tous vos retours. Merci pour votre temps et pour nos discussions, scientifiques ou non, que nous avons pu avoir durant ces trois années. Une pensée spéciale pour toi, Corinne, qui m'as guidé et inspiré à la fois humainement et scientifiquement, et qui m'as supporté (avec le sourire) pendant toutes ces heures de TP.

À l'équipe administrative et technique du laboratoire : Anne, Sybil, Bruno, Clément, Deborah et Céline, qui ont dû écouter mes incessantes demandes tout au long de ma thèse, mes problèmes informatiques, ainsi que mes ordres de mission compliqués de dernière minute. Merci pour votre aide et votre patience.

À Stéphane, car finalement, une thèse, c'est aussi et surtout une aventure entre un.e doctorant.e et son directeur.e. Tu es simplement une personne incroyable, sur le plan scientifique, évidemment, mais également humain. Je suis persuadé que ma "bonne humeur" et mon "optimisme" permanents (comme tu le disais souvent) sont intrinsèquement liés au fait que j'ai eu la chance de travailler à tes côtés tout ce temps. C'était pour moi un véritable plaisir de venir tous les jours au laboratoire en sachant que tu étais là, car, au final, nous avons, oui, travaillé, mais aussi bu beaucoup de cafés, discuté, mangé et surtout rigolé ensemble. Tout ça pour dire que, même si un simple remerciement ne sera jamais suffisant pour réellement exprimer ce que je pense, merci profondément.

À Viola, pour tes encouragements, ta bonne humeur et ta volonté implacable de vouloir me faire réussir. Tu as amplement contribué à mon bien-être, mon envie d'apprendre et de m'améliorer. Finalement, Stéphane et toi, vous avez été tout ce dont je pouvais rêver et avoir besoin... un profond merci.

À Simone, car même si officiellement tu n'étais pas mon directeur de thèse, tu m'as conseillé, guidé et aidé tout au long de mon doctorat. Merci de m'avoir pris sous ton aile et lancé dans ce passionnant domaine qu'est la cosmologie avec les ondes gravitationnelles. Tu as impacté ma vision et mon amour de la physique. Et tout aussi important, merci de m'avoir fait découvrir les fameuses "pizze al metro" de la Scuderia, lors de ma première Virgo Week à Pise, une vraie révélation.

À tous les autres membres de l'équipe Ondes Gravitationnelles de l'Institut de Physique des Deux Infinis : Morgan, Sébastien, Roberto, Patrice, ainsi que les non-permanents : Inès, Amazigh, Francesco, Elisa, Alexis, avec qui j'ai eu la chance de passer des moments super. Une pensée toute particulière pour vous, Elisa et Francesco, qui avez gentiment partagé votre bureau avec moi et qui m'avez supporté à longueur de journée, même si je parlais "trop fort" pendant mes appels.

Aux membres de la collaboration LIGO-Virgo-KAGRA, et particulièrement au groupe de l'ICRR de Tokyo, qui m'ont invité et permis de travailler à leurs côtés, je vous adresse mes sincères remerciements. Cette période passée au Japon fut l'une des plus belles de ma vie, et je vous en suis profondément reconnaissant.

À l'ensemble de ma famille et tout particulièrement ma maman (Anne), mon papa (Macaya), mon frère (Étienne) et ma sœur (Charlotte), vous êtes des personnes extraordinaires qui ne cesseront jamais de m'impressionner. Je réalise petit à petit, avec le temps qui passe, la chance que j'ai eue de grandir à vos côtés. Au fil de ces trois années, vous avez été un vrai soutien, moral et psychologique. Chaque membre, sans exception, est une personne qui me fascine, et je pourrais écrire des pages entières sans parvenir à en capturer toute la profondeur. Je vous aime, merci pour votre soutien inconditionnel.

Et désormais, sur un autre ton, la partie la plus compliquée des remerciements selon moi : celle des amis. Je vous prie d'ores et déjà de m'excuser pour ce qui va suivre, je vais évidemment oublier des personnes, mais il ne faut pas vous vexer car je vous aime tout aussi fort.

Tout commence dans ce petit bâtiment du DLST, dans la "meilleure" promo PCM-Minter, accompagné de personnes qui resteront à mes côtés pendant toutes ces années et qui deviendront très rapidement de vrais amis. Oui, je fais référence à vous: Prune, Thotho, Olivio, Alban, Alex (le petit), Théo, Alex (le frisé), Tiphaine, vous êtes des crèmes et je suis tellement content d'avoir croisé votre chemin. Sans déconner, ces premières années d'études sup avec vous ont été les plus belles, les plus funs et les plus épanouissantes. On a tout fait ensemble : aller en cours, sécher les cours, réviser des heures à la BU, boire des coups, aller au ski aux 7 Laux, faire la fête, les soirées à la coloc, jouer au volley à PPM, partir en vacances, rigoler jusqu'à finir par terre avec la mâchoire bloquée... et j'en passe (oui, on ne peut pas tout dire ici). Je me dois de rajouter obligatoirement Sophie, Solal et Mathis, qui n'étaient pas au DLST, mais que j'ai rencontré antérieurement, dans ce fameux lycée Champollion (cette galère). Sophie, à notre périple dans les Lofoten qui, de part sa météo toute particulière, restera gravé dans nos mémoire, et puis, félicitations à toi Dr. Clouard. Les amis, on se connaît depuis si longtemps et pourtant vous êtes toujours aussi importants à mes yeux. À vous tous, merci du fond du coeur, et je prie pour que l'on ne se perde jamais de vue.

La suite, ça se passe à Stockholm, là où j'ai eu cette révélation : « bah, pourquoi ne pas essayer de faire de la recherche plus tard? », alors que ça faisait 3 ans que je voulais être ingénieur (mdr). À Vishal et Kyle, pour tous ces moments que nous avons passés ensemble en Suède, vous m'avez tellement appris. Je commence donc une liste non exhaustive (que je ne finirai évidemment pas) des choses que vous m'avez apportées ou que nous avons pu faire ensemble pendant cette année de fou : commencer l'escalade, parler anglais (correctement), apprendre des recettes indiennes, le ferry trip avec plein de Russes, voir des aurores boréales, les dégustations de Gin, faire des soirées junk-food en regardant les WC de bloc, visiter la Suède... vous êtes la raison pour laquelle cette année fut l'une des plus belles de ma vie. Merci les amis, on se reverra c'est certain.

Et c'est l'heure du retour en France, en Master de physique, avec la même dream-team qu'au DLST, mais aussi avec de nouvelles additions : Guigui, Adri et Paulo. Mis à part les heures interminables de révisions de partiels, je crois que ce qui m'a le plus profondément marqué, ce sont tous nos moments ensemble, que ce soit le méga trip à vélo dans la Drôme provençale, notre confinement dans la maison dans le Vercors, les soirées Dofus, ou encore nos méga trips d'escalade à Ailefroide. Vous avez rendu ces deux années de Master tellement simples et agréables, merci beaucoup les amis.

Et enfin, on en arrive à Lyon, la thèse. Bon, cette partie n'est pas facile à écrire, car il y aurait beaucoup trop de choses à dire... je vais essayer d'être concis. Je commence donc par les gens avec qui j'ai passé le plus de temps et aussi les meilleurs moments de ma thèse, mes ami·es de doctorat: Flo (statut spécial), Yann (BigBoi), Elso, Sarah, Téo, Ben, Benjamin, (d)Inès, Alfred, Amine (et Léa), David, Élise, Charles. Merci à vous tous, pour avoir embelli mes journées, détruit mon bureau, égayé les repas à Domus, proposé plein de pauses et pour une multitude d'autres raisons que je ne saurai citer dans ce texte. Vous êtes toutes et tous des personnes magnifiques, merci énormément, vous avez fait de mes trois années de thèse une période que je n'oublierai jamais. Mention spéciale au "Booty Squad", pour toutes nos séances de sports, nos soirées "Burgers & Glee", les nouilles DanDan qui piquent pas assez, ou plus récemment les soirées avec "Shooooooooooooondell". Merci également à tous les nouveaux et anciens doctorants que j'ai pu rencontrer au fil des années, à l'IP2I ou ailleurs.

Mes années lyonnaises ont aussi étaient marquées par toutes ces séances d'escalade faites à Mroc Villeurbanne. À tous mes vrais potes de MV, a.k.a. les "Tartineur.es": Quentin, Féfé, Aymeric, Antho, Sacha, Teddy, Mathieu, Irène, Sevag, et j'en oublie certainement plein (désolé les amis, je vous kiffe aussi). Merci pour tout, j'ai passé tellement de bons moments en votre compagnie, tellement de soirées à grimper avec vous, ou bien juste à discuter devant le bar, tellement d'après-midi juste posés sur la terrasse de MV, et tellement de journées d'escalade en intérieur comme en extérieur. Et Quentin, merci de m'avoir fait découvrir le plus bel endroit sur Terre, le rêve de

tout grimpeur qui se respecte un peu : Magic Wood.

Enfin, je souhaite conclure ces remerciements par toi, Florencia (mujersita). Parmi toutes les personnes merveilleuses mentionnées précédemment, tu es celle qui se distingue des autres, et sans toi, ma thèse n'aurait pas été illuminée par un tel rayon de soleil. Depuis notre rencontre au Ninkasi en D1, où je t'aurais "allegedly" vomi dessus, jusqu'à nos moments de vie ensemble à Lyon, où nous avons partagé des rires, des séances de sport, des soirées au restaurant, tu n'as cessé de m'impressionner. Chaque instant passé à tes côtés m'a aidé à devenir un meilleur physicien, une meilleure personne et, plus important encore, tu as su me rendre extrêmement heureux. Je tiens aussi à souligner que, parallèlement à la thèse, nous avons réalisé plusieurs de nos rêves ensemble, comme partir en road trip dans les Dolomites, découvrir le Danemark ou encore vivre au Mexique, me créant des souvenirs que je n'oublierai jamais. Tu es une véritable source d'inspiration, et pour tout cela, merci.

Merci également à toutes les personnes dont la présence ne me vient pas immédiatement à l'esprit en écrivant ces lignes, mais qui comptent tout autant.

À vous tous que j'ai eu la chance de croiser sur mon chemin... un immense merci.

CONTENTS

INTRODUCTION

The history of gravitational wave physics is a lengthy and intricate one, beginning over a century ago with the groundbreaking work of one of the greatest scientists who ever lived, Albert Einstein. The year 1915 not only marks the publication of his theory of General Relativity but also a profound shift in our understanding of gravity. Until then, gravity was considered as one of the fundamental forces of the Universe, described as an invisible force acting between massive objects. Within the formalism of General Relativity, however, gravity is no longer a force but rather a consequence of a deeper and more general description of the Universe's fabric. In this framework, the fundamental structure of the Universe—spacetime—is a dynamic fabric whose geometry is affected by the presence of mass and energy. Gravity is perceived as the natural path that objects follow when the geometry of spacetime around them is curved. This theory, besides being one of the most important breakthroughs in modern physics, also predicted the existence of previously unforeseen astrophysical phenomena, such as black holes and gravitational waves. For decades, both black holes and gravitational waves remained purely theoretical constructs, born from a deeper comprehension of the Universe's structure, awaiting the technological advancements necessary for their observation. Nevertheless, even without direct detection, scientists recognized the unprecedented potential of these objects, as they promised to open a new window into the understanding of the Universe and the extraordinarily powerful astrophysical phenomena within it.

Black holes are cosmic entities that have continuously intrigued scientists. From the perspective of General Relativity, they are referred to as singularities, corresponding to regions of spacetime where the gravitational field is so intense that the geometry of spacetime is curved to infinity. These gravitational giants, primarily formed from the collapse of dying stars in powerful explosions known as supernovae, possess such gravitational field that even light particles cannot escape their grasp. Various kinds of black holes exist in the Universe, ranging from supermassive black holes often located at the centers of galaxies, to stellar-mass black holes, remnants of collapsed stars, and even hypothetical primordial black holes. The intrinsic properties of black holes, such as their masses and spins, depend entirely on the astrophysical processes they have undergone during their formation. Together, they constitute a "bestiary" of black holes, relics of past astrophysical phenomena. Occasionally, two black holes can form a gravitationally bound system known as a binary black hole, which is destined to merge in a highly energetic explosion after a long period of unstoppable spiraling. Due to the extreme gravitational fields and masses involved in such events, these systems are ideal sources of gravitational waves. These waves, intrinsic to the structure of spacetime,

provide further evidence of the dynamic nature of the Universe's geometry. The coalescence of binary black holes produces distinctive gravitational wave signatures that travel across the Universe at the speed of light, awaiting detection.

The quest to detect the first gravitational wave signal began in the 1960s with the innovative work of Joseph Weber. Ahead of his time, Weber developed an apparatus dedicated to the detection of gravitational waves, using resonant metallic bars. While he aimed to make the first direct detection of these tiny oscillations of spacetime, his experiments led to controversial results rather than unequivocal detection. Despite the significant skepticism from the scientific community regarding Weber's work, these pioneering efforts greatly stimulated the research field. In the early 1990s, the American project known as the Laser Interferometer Gravitational-Wave Observatory (LIGO) was approved, led by Kip Thorne, Rainer Weiss, and Ronald Drever. Alongside LIGO, the French-Italian Virgo project, spearheaded by Adalberto Giazotto and Alain Brillet, also moved forward. These large-scale Michelson interferometers, with kilometer-long arms, were designed to detect the tiniest variations in space caused by the passage of a gravitational wave. After more than twenty years of instrumental improvements, on September 14, 2015, the LIGO interferometers made history by detecting the very first confirmed gravitational wave signal, GW150914. This event, which later earned a Nobel Prize for the American team, is considered one of the greatest discoveries of the century. It was produced by the merger of two massive black holes 440 *Mpc* away from Earth, during which the equivalent of three solar masses of energy were radiated in the form of gravitational waves. The detection of GW150914 confirmed Einstein's predictions about the existence of black holes and gravitational waves in the Universe. In the following years, the field of gravitational wave astronomy rapidly expanded. The LIGO and Virgo Collaborations managed to detect more than 90 confirmed gravitational wave signals, with some of them originating from new sources such as binary neutron stars and neutron star-black hole binaries. Beyond the significant achievement of detecting gravitational wave signals, the scientific community began using this data to explore the physics of the Universe through a completely new lens. This new observational method provided insights into the astrophysics of black holes, the origins of gravitational waves, cosmology and more.

The word cosmology comes from the ancient Greek roots, *kósmos*" meaning *world* or *order* and "*logía*" meaning *treating of*, cosmology is then the study of the Universe as a whole, from its origin, its evolution, its dynamics, to its fundamental properties today. This field of physics attempts to describe our entire Universe as a single volume with all its structures at large and small scales. In a very general sense, cosmology is the science that tries to answer questions like: "What is the beginning of the Universe?", "Was there a beginning?", "How did it evolve in order to explain what we observe today?", and "What will be the future of the Universe?". Modern cosmology, as we know it today, is primarily based on observations, utilizing giant terrestrial antennas and telescopes or satellites that can probe regions of the Universe invisible from Earth.

From these various observations, scientists have been able to test and invalidate numerous models and theories, leading to what we now refer to as the "standard model of cosmology". In this standard model, the Universe has evolved from a hot Big Bang approximately 13.7 billion years ago, a stage when the temperature and density were so high that all the ordinary and non-ordinary matter, radiations, and other components were bound to one another. The Universe then entered a new stage of rapid expansion known as inflation. This period, which lasted a fraction of a second, allowed the entire Universe to cool down even more, and the flow of matter and radiation began to populate the cosmos. As the Universe expanded, its temperature continued to cool, and part of the radiation and matter started to free themselves. Over a long period, through physical interactions (nucleosynthesis) and gravitational attraction, entities and structures started forming, resulting in what we can observe today such as stars, planets, nebulae, black holes, neutron stars, galaxies. Even though the standard model of cosmology is highly effective in explaining and predicting the history and future of the Universe, several incompatibilities with observations remain. These discrepancies can be grouped into two broad categories: those affecting the "large scales", i.e. the determination of cosmological parameters like the Hubble constant, and those affecting the "small scales" like the overabundance of predicted cold dark matter sub-halos compared to satellite galaxies known to exist in the local group.

In light of this introduction, this manuscript aims to provide a clear and pedagogical exploration of several aspects of gravitational wave physics. These include advances in statistical techniques for estimating parameters such as the Hubble constant and characteristics of black hole populations, investigations of systematics in gravitational wave Standard Siren cosmology, the application of machine learning methods to enhance gravitational wave interferometer data quality, and the exploration of binary black hole subpopulations and hierarchical mergers.
This manuscript is structured as follows:

Chapter I serves as an introduction to the foundational concepts underpinning this manuscript. It covers a range of topics, starting with the physics of gravitational waves and their formalism inherited from the theory of General Relativity. It then explores the astrophysics of binary black holes, including their formation and properties as dense astrophysical entities. Finally, chapter I delves into the standard model of cosmology and its current problematics.

Chapter II offers an overview of the various methodologies for gravitational wave cosmology. It details the cosmological pipeline ICAROGW, a hierarchical Bayesian framework designed for inferring population and cosmological parameters from gravitational wave data. Furthermore, this chapter presents the incorporation of spin population models within ICAROGW, novel addition to the Standard Siren method for cosmological and population inference.

Chapter III conducts a comprehensive study of the systematic errors and biases involved in Standard Siren inference. It explores the complex relationship between the models used to parameterize the binary black hole population and the inferred cosmological parameters. This chapter aims to clarify the primary sources of biases impacting the Hubble constant, such as unmodelled mass features or unknown astrophysical effects.

Chapter IV introduces a study on binary black hole populations, aiming to unravel the relationship between spin and mass parameters. Employing novel population models sensitive to spin-mass correlations, this chapter explores the influence of binary black hole formation channels. It also seeks for evidence of the existence of subpopulation, through the presence of variations in the spin distributions of black holes.

Chapter V exposes an exploratory study focused on enhancing the data quality of the Virgo interferometer by identifying non-Gaussian noise (glitches). It explores the application of iDQ, a machine learning-based algorithm that detects glitches by establishing high-dimensional correlations between the gravitational wave-sensitive channel and safe auxiliary channels of the detector.

1 GRAVITATIONAL WAVES, BLACK HOLES AND COSMOLOGY

1.1 INTRODUCTION

Gravitational waves (GWs), predicted by Albert Einstein's general theory of relativity (GR) introduced in 1916, emerge as a testament to the dynamic nature of the cosmos [1]. These gravitational ripples offer a unique window into the Universe's most cataclysmic events, such as the violent collisions of black holes (BBH) and the merger of neutron star binary (BNS) systems. GWs carry the signatures of astrophysical phenomena that transcend traditional astronomical observations, serving as a direct confirmation of GR's predictions and opening new avenues for exploring the Universe's gravitational landscape.

Cosmology, the systematic study of the universe's origin, evolution, and fundamental properties, has continually evolved, from ancient Greek models to the advancements of Sir Isaac Newton's Principia Mathematica [2]. The true modernization of cosmology began with the theory of GR, providing a framework for understanding the universe's large-scale structure and dynamics. Central to contemporary cosmology is the Hubble tension, a discrepancy in measurements of the Hubble constant (H_0), which describes the rate of cosmic expansion [3]. GWs offer a promising avenue for resolving this tension by providing independent probes of cosmic distances

and the expansion rate of the Universe, potentially refining our understanding of H_0 [4].

Black holes, enigmatic compact objects predicted by the theory of GR, also occupy a central role in the study of both cosmology and GWs. These gravitational giants, formed from the collapse of massive stars, possess such intense gravitational fields that even light cannot escape from them [5]. As cosmic laboratories of extreme gravity, BHs represent ideal sources for generating detectable GWs [6]. The merger of two BHs, accompanied by the violent release of gravitational energy, produces the distinctive GW signatures observed by detectors such as LIGO, Virgo, and KAGRA (LVK) [7–11]. The shape of the GWs emitted during a coalescence is governed by the astrophysical properties of the merging black holes; their masses and spins impact the amplitude, frequency, and phase of the wave.

This chapter aims to discuss all the fundamentals needed to understand the field of GW physics, the astrophysics of spinning BHs and modern cosmology. It is constructed as follows. In Sec. 1.2, we provide an introduction to GWs physics, covering everything from the fundamental properties of GWs, the violent phenomena at their origin, to the current detector network used to detect them. In Sec. 1.3, we present an introduction to the astrophysics of stellar-mass BHs, starting from their formation as dying stars to the special case of spinning BHs and the complex processes leading to the formation of BBH systems. Sec. 1.4 introduces the foundations and basic principles of modern cosmology, particularly focusing on how contemporary cosmology describes the Universe mathematically. It delves into a fundamental issue often referred to as the Hubble tension, emphasizing its critical importance in understanding the intricacies of the Universe. In Sec. 1.5, we establish the connection between cosmology and GW physics, underscoring the significance of GWs as a powerful tool for probing cosmological phenomena. This section also gives an overview of the most recent results obtained in GW cosmology. Finally, Sec. 1.6 concludes.

1.2 GRAVITATIONAL WAVES

This section presents the current theoretical framework, obtained from the theory of GR, to describe GW as ripples of the space-time geometry. Starting from the Einstein field equations, we derive the GWs' equations. Moreover, we introduce the current GW detector network, known as LIGO-Virgo-KAGRA. In particular the final product that the collaboration produces, namely the GW transient catalogs (GWTCs).

1.2.1 Gravitational waves in a nutshell

In 1916, one year after the publication of his GR theory [12], Albert Einstein theoretically predicted the existence of GWs [13]. According to the GR theory, the structure of our Universe can be described by four dimensions, three of space and one of time, this space is often referred to as space-time. In the presence of energy or mass, space-

time can be distorted, and so Einstein predicted that accelerating masses can create waves of space-time, that propagates at the speed of light. These waves of space-time, seen as ripples of the structure of the Universe, were later called GWs. This prediction was not accepted by the scientific community up until the 1950s, when Pirani showed that such GWs could be seen as gravitational radiation within the Riemann tensor formalism [14].

Although indirect evidence for GWs was identified as early as the 1980s through the study of binary pulsar orbital decays [15, 16], so another thirty years after Pinari's work, the first detection of GWs occurred in 2015-almost a century after Einstein's initial prediction [6]. This historic observation was made by the advanced LIGO detectors (Laser Interferometers GW Observatory) situated in Hanford and Livingston in the United States of America [7]. The source that emitted the observed signal on Earth came from a pair of coalescing BHs, approximately 440 *Mpc* away, that were spiraling around each other before finally merging into a newborn BH. One important aspect of the GW signal emitted by a compact binary coalescence (CBC) is its ability to provide an estimation of the astrophysical properties of the source itself, and in particular the masses, the spins and the luminosity distance between the source and the observer. This feature makes GWs particularly valuable for astrophysical and cosmological studies. Measuring cosmological distances has always been challenging, but GWs offer a novel and independent method for obtaining direct distance measurements independent of the cosmic distance ladder used for cosmology with the type Ia supernovae. The rest of the manuscript will assume from now on that the discussed GW signals are emitted by CBC sources.

In the following, we present how the expressions of the GW signal (GW waveform) can be derived starting from the linearized Einstein field equations. From the GR theory, the Einstein field equations formalize how the geometry of space-time changes in the presence of matter and energy. Following [17], these equations can be written under the form

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

where $R_{\mu\nu}$ is the Ricci tensor representing the local curvature of space-time, R is the scalar curvature defined as $R = g^{\mu\nu} R_{\mu\nu}$,$T_{\mu\nu}$ is the stress-energy tensor and $g_{\mu\nu}$ is the metric tensor. In the linearized theory of gravity, which is a particular case of GR in the presence of a weak gravitational field, the metric tensor can be expressed as the metric of a flat space-time $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ plus a small perturbation of the weak gravitational field $h_{\mu\nu}$, such that

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \ , |h_{\mu\nu}| \ll 1. \tag{1.2}$$

Assuming that the gravitational field is weak, i.e. the metric tensor is close to being flat, and by choosing a harmonic gauge such that $\partial_\nu \bar{h}^{\mu\nu} = 0$, where $\bar{h}^{\mu\nu} = h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu} h$, the Einstein field equations can be written as

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}, \tag{1.3}$$

where $T_{\mu\nu}$ is the stress-energy tensor of a region in space-time, and $\Box = \eta_{\mu\nu}\partial^{\mu}\partial^{\nu}$ is the d'Alambertian operator. Eq. 1.3 has the typical form of a wave equation emitted by a source, here represented by the term on the right hand-side of the equation. In vacuum, the expression of the wave equation can be simplified in

$$\Box\bar{h}_{\mu\nu} = 0. \tag{1.4}$$

One possible solution to the wave equation (Eq. 1.4), directly inspired from the theory of electromagnetism, can be derived such that

$$\bar{h}_{\mu\nu} = A_{\mu\nu}e^{ik_{\alpha}x^{\alpha}}. \tag{1.5}$$

This peculiar solution, also called the plane wave solution, $A_{\mu\nu}$ is the polarization tensor that contains information about the respective amplitudes and polarizations of the GW, and k_{α} is the wave quadri-vector. The GW has the physical property of being uniform in all directions and traveling with a certain frequency and amplitude. Following a specific choice of gauge, the transverse-traceless gauge, the number of independent components of the metric perturbation $h_{\mu\nu}$ can be reduced to only two, corresponding to two polarizations named "plus" and "cross". With the specific choice of coordinates, the plane wave is transverse, and the solution to the wave equation found in Eq. 1.5 can only be described by two possible functions, one for each polarization.

GWs cause distortions in the space-time structure through which they travel, perpendicular to their direction of movement, resulting in a phenomenon known as strain. The strain corresponds to the physical deformation of space-time, under the passage of a GW. In simpler terms, if two test masses in free fall were placed some distance apart, a passing GW would alter the distance between them. This is the fundamental principle behind the detection of GW signals by laser interferometers. Similar to how light can be polarized, with its orientation expressed as a combination of two orthogonal polarizations, GWs also exhibit polarizations. The strain caused by the passing GW can be divided into two components, as shown by the mathematical formalism above: the plus-polarized component h_{+} and the cross-component $h_{\times}$, as illustrated in Fig. 1.1

Up to now, most direct detections of gravitational waves have been made using the terrestrial laser interferometer network LIGO-Virgo-KAGRA. However, another detection technique called Pulsar Timing Array, has also likely detected GW signals at low frequencies. The sensitivity of these detectors is not absolute nor isotropic. It implies that the strain of a GW signal in the detector will depend on the astrophysical properties of the source such as the masses of the merging objects, their distances to the detectors, their orientations, but also on the detector response. The response of a GW detector directly depends on the direction of propagation of the waves, in particular, its sensitivity is maximum when the wave crosses the detector plane perpendicularly.

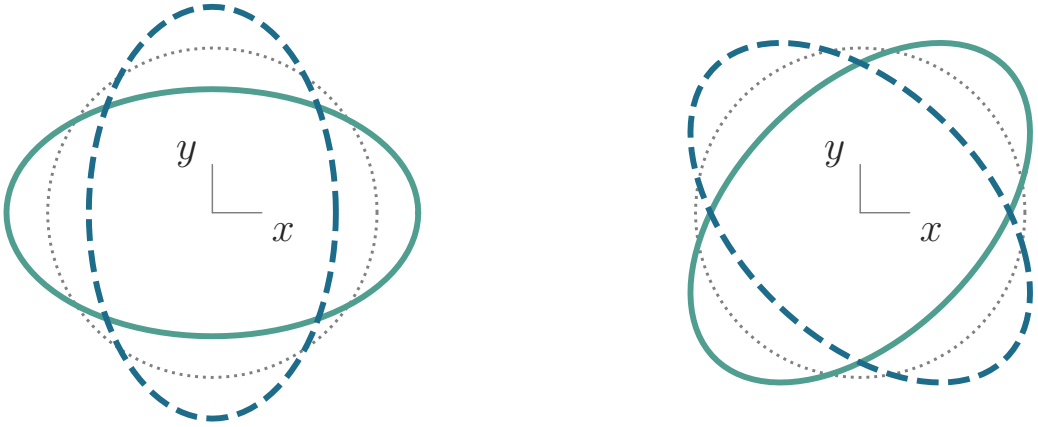

Figure 1.1: Effect of a GW on a free-falling ring of masses located in the x-y plane. The green plain line and the blue dashed lines correspond to the modified shapes of the ring of masses, at the maximum and minimum of the wave. **Left:** Effect of the plus polarization on a free-falling ring of masses. **Right:** Effect of the plus polarization on a free-falling ring of masses. The wave propagates along the z direction, and the unperturbed ring is represented as the thin dotted line. Figure taken from [18].

In the case of a single laser interferometer detector, the response function for each polarization are written as in [19]

$$\begin{aligned} F_+(\theta,\phi,\psi) &= \frac{1}{2}(1+cos^2\theta)cos2\phi cos2\psi - cos\theta sin2\phi sin2\psi, \\ F_\times(\theta,\phi,\psi) &= \frac{1}{2}(1+cos^2\theta)cos2\phi cos2\psi + cos\theta sin2\phi sin2\psi, \end{aligned} \tag{1.6}$$

where (θ,ϕ) are the sky position angles, ψ is the polarization angle, and F_+ and $F_\times$ the respective detector response functions for each polarization of the GW. Finally, the GW strain induced in the detector as a function of time $h(t)$ can be written following [20], as a linear combination of the two polarizations of the waves combined with the detector response functions such that

$$h(t) = F_+ h_+(t) + F_\times h_\times(t). \tag{1.7}$$

The strain, for each polarization can be expressed in a more detailed way, in order to have an explicit dependency on the astrophysical properties of the GW source. Assuming the quadrupole approximation [20], the plus-polarized strain takes the following form

$$h_+(t) = \frac{2\mathcal{M}_c G}{d_L c^2}(1+cos^2\nu)\left(\frac{5G}{256c^3}\frac{\mathcal{M}_c}{T-t}\right)^{1/4} cos\left(-2\left(\frac{(T-t)c^3}{5\mathcal{M}_c G}\right)^{5/8} + \Psi\right), \tag{1.8}$$

where d_L is the luminosity distance of the source to the observer (here the detector), ν is the inclination of the source w.r.t the observer, T is the time of coalescence, f is the

time-dependent frequency of the GW, Ψ is the phase of the system at $t = T$. Finally, $\mathcal{M}_c$ is the redshifted chirp mass expressed as

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}(1+z), \tag{1.9}$$

with m_1 and m_2 being the source frame masses of the two compact objects coalescing, and z the redshift of the source. The concept of redshift is detailed later in Sec. 1.4.1. The redshifted chirp mass encodes the mass dependency of the GW signal. The cross-polarized strain is similarly expressed such that

$$h_\times(t) = \frac{4\mathcal{M}_c G}{d_L c^2} cos\nu \left(\frac{5G}{256c^3}\frac{\mathcal{M}_c}{T-t}\right)^{1/4} sin\left(-2\left(\frac{(T-t)c^3}{5\mathcal{M}_c G}\right)^{5/8} + \Psi\right). \tag{1.10}$$

After the detection process, where the GW signal's strain is extracted from the detector noise, various physical properties specific to GWs can be inferred from the waveform. The amplitude of a GW, as described by Eq. 1.8 and Eq. 1.10, is directly proportional to the system's redshifted chirp mass and inversely proportional to its luminosity distance. Unlike electromagnetic waves, for which intensity decreases with distance squared, the amplitude of a GW decreases inversely with the luminosity distance. Accurately estimating the luminosity distance is challenging due to its partial degeneracy with the inclination angle ν of the binary system. However, observations from multiple detectors can help break this degeneracy, allowing for more precise luminosity distance estimates. Conversely, the redshifted chirp mass of the source can be estimated due to its dual degeneracy with the luminosity distance, as well as with the phase and phase evolution of the waveform.

The derivation of the strain of a GW seen by a laser interferometer, as shown earlier, was for the case of a single detector detection. However, the current detector network consists of four instruments, LIGO Hanford and Livingston, Virgo and KAGRA. When the same GW signal is observed by multiple detectors, the precision in measuring the intrinsic parameters of the source increases significantly [21]. Some degeneracies can be better resolved, leading to smaller uncertainties on parameters such as the luminosity distance as discussed above. A multi-detector detection also allows for triangulation of the signal's origin, providing an estimation of the sky localization of the source, Ω. In addition to the redshifted chirp mass, luminosity distance, and inclination of the source, a process called "parameter estimation" allow us to infer several extra astrophysical properties of the binary system from the detected GW signal [22]. These include the primary and secondary masses and the spins of the compact objects. Parameter estimation in GW physics involves extracting key properties of the source from the detected waveform. This is typically done using Bayesian inference, which compares the observed data to theoretical models to determine the most likely values of these parameters, along with their uncertainties.

As we will see later in the context of GW cosmology, in addition to the masses, distance, spins, and inclination of the source, the redshift must also be determined. However, even if the redshift information is embedded within the GW signal, it is

degenerated with the source-frame masses through the chirp mass term, $\mathcal{M}_c$. This interdependence leads to a degeneracy that blurs precise redshift determination from the GW signal alone. To effectively use GW for cosmological measurements, it is essential to find methods to break this redshift degeneracy. This can involve either disentangling the redshift from the source-frame mass in the GW data or independently determining the source's redshift through alternative methods. Such methods may include observing electromagnetic counterparts or identifying the host galaxy of the GW source.

1.2.2 The LIGO-Virgo-Kagra detector network

The international GW detector network, LVK, is currently composed of three terrestrial Michelson interferometers. Two of them are located in the United States of America (LIGO Hanford and Livingston), one French-Italian in Italy (Virgo) and the latest in Japan (KAGRA). In 2015 started the first observation run (O1) led by the LIGO detectors, who detected the first GW signal from a merging BBH (GW150914). In 2017, Virgo joined for the final months of the second observing run (O2), helping the other two LIGOs for several detections, which then led to the detection of the first BNS merger. The third observing run O3 lasted 11 months, between 2019 and 2020, and was the longest data taking period for the LIGO and Virgo detectors. Currently, we are in the middle of the fourth observing run (O4), for which the KAGRA detector joined. At the end of O3, we, as the LVK Scientific Collaboration, have detected 90 confident GW signals coming for the large majority from BBHs, but also some from neutron star black hole (NSBH) and BNSs.

Following each observing run, we collect the GW signals and summarize all relevant information about them in the so-called GW transient catalogs (GWTC) [24–27]. The main result of these catalogs is a list of these events and the estimated population parameters of their sources (masses, spins, distances, sky localizations, ...). From a data analysis point of view, these catalogs are the starting point of numerous kinds of studies, ranging from astrophysical studies of the sources, GW signal exploration, cosmological analyses, tests of GR theory, or even modification of gravity. In other words, GW detections are crucial probes for all kinds of field in physics. However, certain studies are limited by the incompleteness of the detections with the actual detectors, often referred to as selection effects. GWs are also characterized by their time varying frequencies, which are typically in the $[10 - 1000]\,Hz$ range for a vanilla BBH merger coming from stellar origin. Fig. 1.2 shows the frequency sensitivity curves of the advanced LIGO and Virgo detectors, compared to the frequency ranges of potential GW sources. From this figure, it is clear that the actual detectors are fitted to detect CBC mergers, which makes the network highly important for any study related to compact objects.

In particular, $[10 - 1000]\,Hz$ corresponds to the typical frequency range of merging stellar-mass BHs. Although, the entire population can not yet be seen, since far away or light systems produced GWs still too faint in amplitude. This manuscript, with the

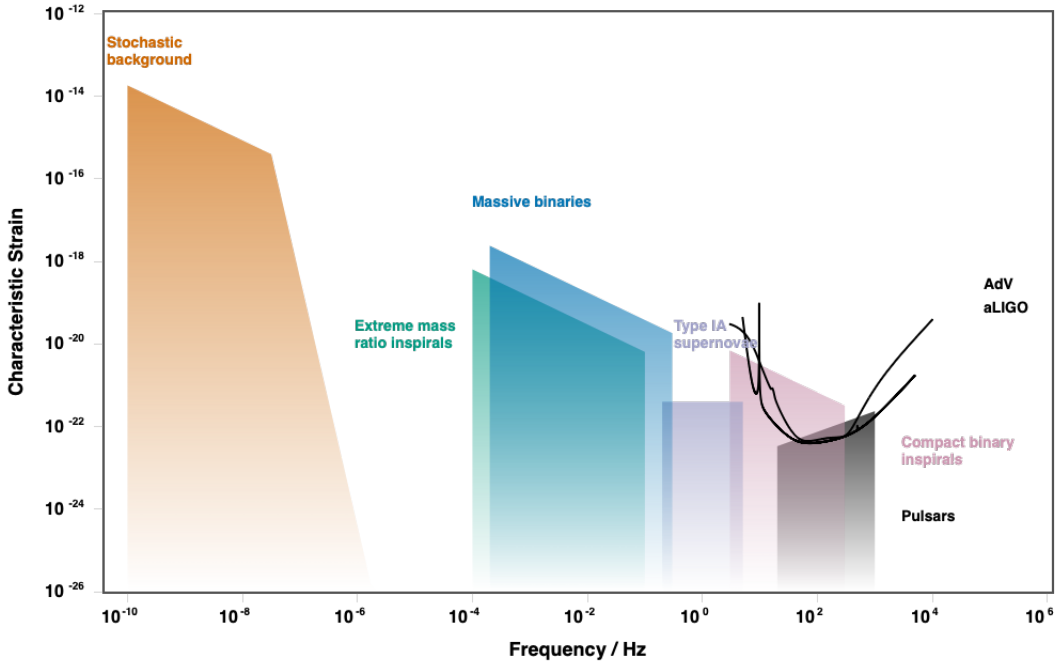

Figure 1.2: The sensitivity curve plot presents different type of potential GW sources, producing GWs from 10^{-10} Hz with the stochastic GW background up to 10^{3} Hz with CBCs. The two dark lines shows the sensitivity curves of the advanced Virgo and Advanced LIGO detectors, meaning every strains that fall above these lines can be detected. Figure based on [23]

help of sophisticated statistical inference methods, explores the populations of stellar-mass BHs, from an astrophysical point of view as well as from a cosmological one.

1.3 ASTROPHYSICS BLACK HOLES

In this section, we introduce the astrophysics of BHs. We give an overview of how stellar-mass BHs form in the Universe and how their formation channels can affect their astrophysical properties, such as their masses and spins. A focus is given to spinning BHs, which are described within the GR theory by the Kerr metric. And we conclude with a discussion about the impact of spin when considering a BBH system.

1.3.1 Black hole formation

Before discussing the different ways BBH systems can form, it is important to understand the origin of single BHs, and more especially stellar-mass BHs which correspond to the typical sources of GWs detected by the LVK network. A stellar-mass BH is a BH formed after the gravitational collapse of a star [28]. Although, not all dying stars will create a BH at the end of their life. For instance, the remnant of a dying star can also be a white dwarf or a neutron star, if its original mass was too light to overcome

the degeneracy pressure of electrons or neutrons in order to self-collapse into a BH [29]. For stars massive enough, the helium core of the star will collapse under its own gravity, overcoming the neutron and electron degeneracy pressures, and form a BH. This process is referred to as core-collapse supernovae [30]. However, it is important to note that core-collapse events can result in various outcomes: they may produce remnants such as black holes, neutron stars, or white dwarfs, but in some cases, they may also leave no remnant at all. The current astrophysical models and observational data show that stellar-mass BHs are expected to have masses ranging from a few solar masses to approximately $\mathcal{O} \sim 100\,M_\odot$ [31].

A BH has three intrinsic properties that are defined by the No-hair theorem, this theorem states that these objects can be entirely described by their mass, their electric charge and their angular momentum (also called spin) [32]. The stellar-mass BHs are generally neutral objects ($Q = 0$). The spin, is highly dependent on the angular momentum of the star from which it was formed and BHs with various spin values exist in the Universe. For single BH, we often refer to as the spin magnitude (speed of rotation). The range of mass for stellar-mass BHs is still very uncertain since the latter is probably affected by several astrophysical processes that are currently not well understood [33].

Significant uncertainties exist regarding the relationship between a BH's mass and the properties of its parent star. One of the most widely accepted astrophysical process which is expected to influence the mass distribution of BHs is the "Pair Instability gap of SuperNovae" (PISN) [34–36]. When the star is about to explode as a supernova, depending on the mass of its helium core m_{He}, it can collapse into a BH or be totally disrupted, hence leaving no compact remnant. Typically, stars having helium cores with a mass in the range of $64 < m_{He} < 135\,M_\odot$ can not form a stellar-mass BH, this translates into a depletion in the BH spectrum between $\sim 50^{+20}_{-10}\,M_\odot$ and $\sim 120\,M_\odot$. Pair instability supernovae happens when the pair production of electrons and positrons temporarily decreases the radiation pressure of the stars core, pressure which fight against the gravitational collapse. The large uncertainties on the lower edge of the gap are due to a lack of understanding of the physical processes ruling the end of life of a star. Other phenomena like stellar winds, mass transfer, and the core-collapse supernovae also influence the global mass spectrum of single BHs [37, 38].

The origins of a BH's spin are also unclear, and not as well measured as the mass. It is reasonable to assume that a compact object inherits the totality of the angular momentum of its parent star if the star collapses directly into a BH without a supernova explosion (direct collapse) [39, 40]. In contrast, when the star goes through several astrophysical processes, effects like mass ejection during the supernova explosion can significantly reduce the parent star's final spin. If the parent star's final spin is large, we expect large birth spins for the most massive BHs, which form from direct collapse, and low birth spins for less massive BHs, which form from successful supernova explosions.

1.3.2 Binary black hole formation

The straightforward question that arises, in the context of GWs, is how BBH systems form in the Universe ? Two massive stars are in a binary system, they eventually collapse both into BHs, forming a BBH system ? This ideal scenario can only be true if the system is large enough, i.e. the distance between the two objects is large enough so that its evolution is not affected by external interaction. In the case of CBCs, which lose angular momentum and merge through the emission of GWs, the typical distance between the two object is too small to avoid such processes, we talk about tight systems [41]. Currently, there are two main families of formation channel for BBHs with stellar-mass BHs: the isolated formation and the dynamical formation [33, 42].

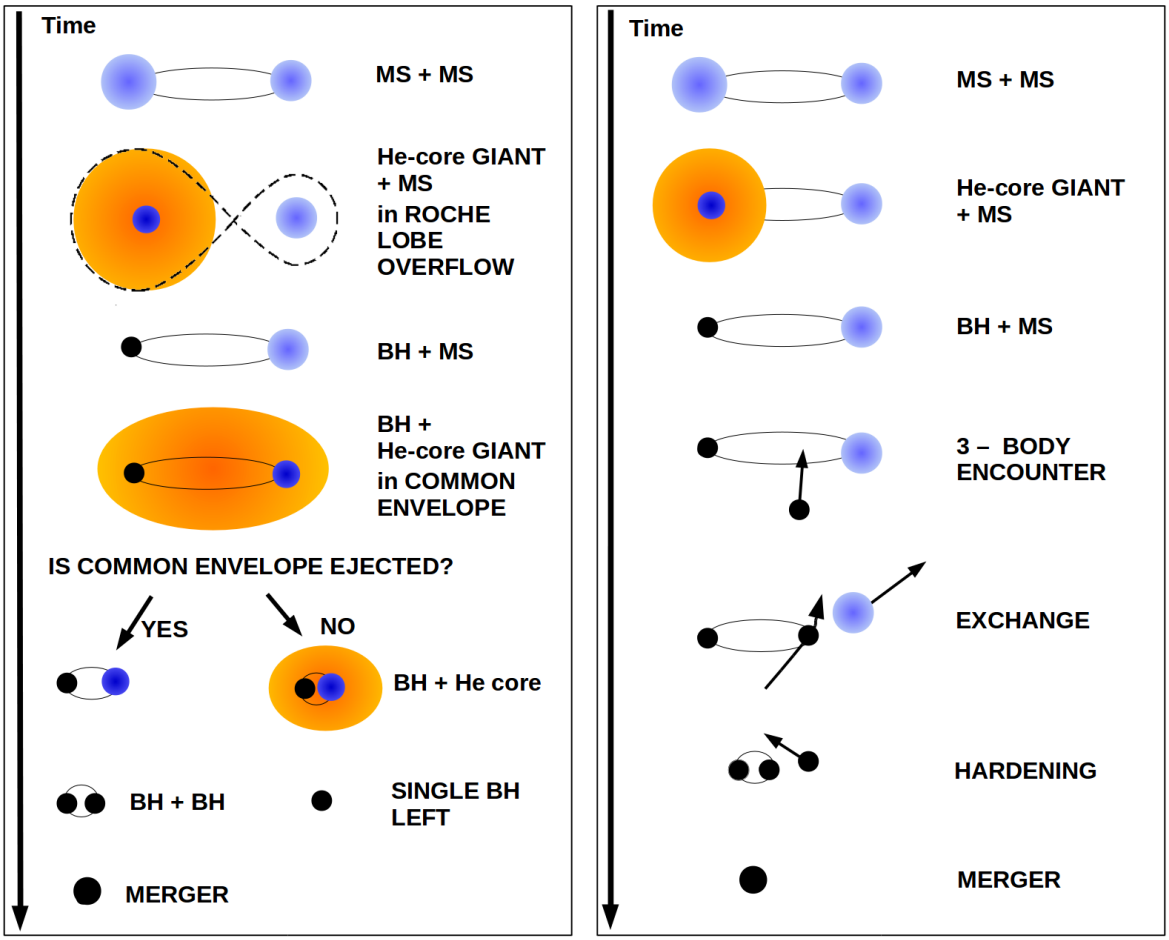

Figure 1.3: Left: Schematic view of the isolated formation channel through common envelope phase, starting from a binary system of massive stars to a BBH system. Right: Schematic view of the dynamical formation channel of a BBH system in a dense stellar cluster. Figure taken from [42].

The isolated formation channel of BBHs involves astrophysical processes such as mass transfer and common envelope phases, which significantly impact the evolution of tight binary stars. During these phases, a massive star can transfer mass to its companion or undergo a common envelope stage, which leads to the reduction in orbital

separation. This results in the formation of BBHs with short orbital periods, essential for merging within a Hubble time under GW emission, as depicted in Fig. 1.3. The typical masses in this channel range from approximately $3M_\odot$ to $45M_\odot$, with mass ratio close to unity. Additionally, the spins of these BHs tend to be aligned, and the systems exhibit near-zero eccentricity. Large uncertainties of the understanding of the common envelope dynamics and its impact of the BBH population are still at play.

The dynamical formation channel of BBHs mainly happens in high-density stellar clusters, where dynamical exchanges and three-body encounters rule the formation processes. In this environment, massive objects like BHs are efficient in forming new binaries through exchanges. Three-body encounters, particularly with tight binaries, can lead to dynamical hardening processes, accelerating the merger by reducing the distance between the two objects (see Fig. 1.3). The hardening of the binary system gravitationally locks the two BHs together, under the effect of a third party. Dynamically formed BBHs can extend to higher masses than the ones formed in isolated environments, and potentially populate the PISN gap and/or the intermediate mass range $[10^2 - 10^5]\,M_\odot$. Unlike the isolated BBHs, their spin distributions tend to be more isotropic due to the dynamical exchanges, and can exhibit non-zero eccentricity. A special formation channel considered "dynamical" is the hierarchical merger channel, where BBH systems are formed with BHs born from previous BBH mergers. The remnant BH formed after a BBH merger is expected to have a specific value of spin magnitude peaking around 0.7 [43–45]. These astrophysical characteristics offer unique signatures for distinguishing between the dynamical and isolated formation channels of BBHs, crucial for understanding their demographics and evolution. The later will be further investigated in Sec. 4.

1.3.3 The Kerr black hole

To address the question of spinning BHs in a more elaborate manner, it is crucial to introduce the notion of "*metric*". Formally, a metric is a distance function for a general space, that describes the separation of two events happening in that space. Usually, we consider two events that are infinitesimally close to one another so that the metric is true for every geometry of spaces, curved or flat. As the geometry of space-time around a non-spinning BH is given by the Schwarzschild metric, the mathematical description of a single spinning BH within the GR framework is done through the Kerr metric [46, 47]. The Kerr metric encapsulates how the geometry of space-time around a spinning and uncharged BH is changing. It is an exact solution of the EFEs of GR, discussed in Sec. 1.2.2. The Kerr metric is a generalized version to spinning object of the Schwarzschild metric for neutral, spherical and non-spinning BHs. One of the main prediction from GR about a Kerr body is called frame-dragging [48], which states that objects getting close to a spinning mass like a Kerr BH will start to rotate around it due to the swirling curvature of space-time created by the rotating BH. This region where objects (light particles included) start rotating due to the geometry of space-time is called the ergosphere of the spinning BH. The geometry of space-time

around a Kerr BH, given by the Kerr metric is completely characterized by the BH mass and its angular momentum. Typically, one can write the following line element of the

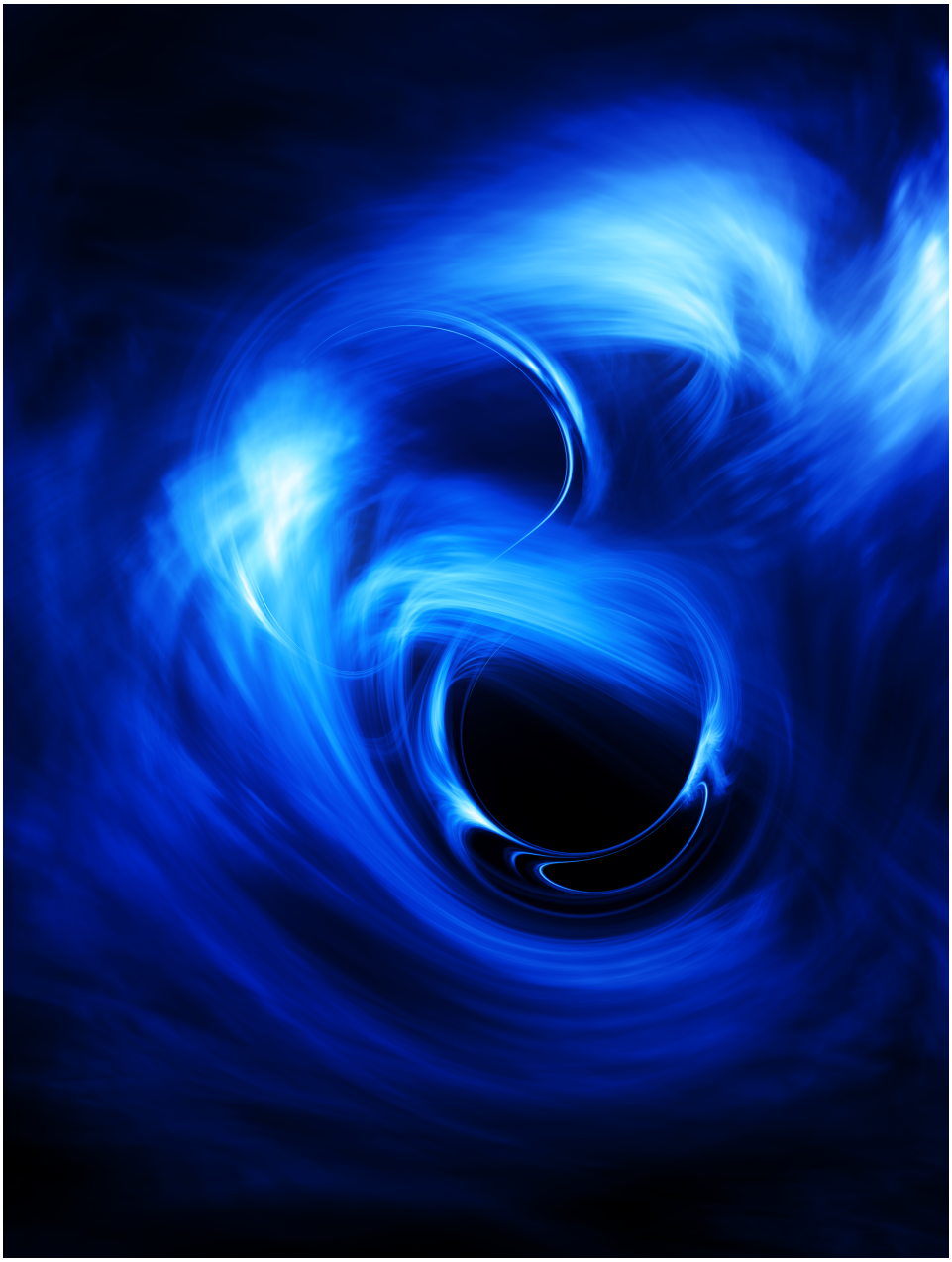

Figure 1.4: Simulation of a merging BBH system with two rapidly spinning stellar-mass BHs. This simulation includes derivation of the accretion disks and gravitational lensing effects of their respective horizons. This image has been created with a GR simulation of Kerr black-holes, by Alessandro Roussel in 2024. The mass ratio between the two BHs is $q = 0.55$ and the spin magnitudes are chosen to be $\chi \sim 0.9$.

Kerr metric using the Boyer-Lindquist coordinates where $x^{\alpha} = (ct, r, \theta, \phi)$, (extension of polar coordinates [49]) as

$$\begin{aligned} ds^2 &= -\left(1 - \frac{2Gmr}{c^2\rho^2}\right)c^2dt^2 - \frac{4maGrsin^2\theta}{c^2\rho^2}cdtd\phi + \frac{\rho^2}{\Delta}dr^2 \\ &+ \rho^2 d\theta^2 + \left(r^2 + a^2 + \frac{2ma^2Grsin^2\theta}{c^2\rho^2}\right)sin^2\theta d\phi^2, \end{aligned} \tag{1.11}$$

where $\Delta = r^2 - \frac{2Gm}{c^2} + a^2$ and $\rho = r^2 + a^2 cos^2\theta$ and a, the Kerr parameter, is a term that relates the mass m of the BH and its angular momentum J such that

$$a = \frac{J}{cm}. \tag{1.12}$$

The Kerr parameter can be redefined as $\bar{a}$, a dimensionless parameter. This dimensionless Kerr parameter is typically expressed such as

$$\bar{a} = \frac{c^2}{G}\frac{a}{m} = \frac{c}{G}\frac{J}{m^2}. \tag{1.13}$$

In terms of physical interpretation, there is a spot at the center of the Kerr BH where the space-time curvature diverges, also called singularity. In order to avoid a "naked" singularity, i.e. a BH with no event horizon, the dimensionless Kerr parameter $\bar{a}$ has to be bounded such that $|\bar{a}| \leq 1$. The dimensionless Kerr parameter of the BH impacts the size of the radius of the event horizon such that:

$$r_{horizon} = \left(1 + \sqrt{1 - \bar{a}^2}\right)\frac{Gm}{c^2}. \tag{1.14}$$

We note that the event horizon of a Kerr BH is directly influenced by its dimensionless Kerr parameter $\bar{a}$, and just like for a Schwarzschild BH, this radius marks the distance below which anything will be dragged into the singularity. If the dimensionless Kerr parameter $\bar{a}$ was allowed to go above unity, the event horizon of the Kerr BH would not exist, and the singularity would be naked. The upper limit on the dimensionless Kerr parameter is called the Kerr-bound. There exists other physical properties related to the geometry of space-time around a spinning BH, but the later will not be discussed here, we encourage the reader to check [50] for more detailed explanations.

1.3.4 Binary systems of spinning black holes

So far, we have explored how BHs and BBHs systems can form in the Universe, as well as the physics of spinning BHs, but in the context of this manuscript and of GW analysis, we must consider the case of spinning BBHs. Fig. 1.4 illustrates a GR simulation of what one could observe looking at a coalescence of two spinning BHs. As GR is a non-linear theory, it is not correct to superpose two Kerr solutions to the EFEs in order to have a spinning BBH solution. Generally, the physical properties of the emitted

GWs for such system are derived numerically, at least for the inspiral and the merger phase. An analytical approach, called Post-Newtonian approximation, also permits to model the inspiral phase of the GW emitted by binary merger. The observation of such systems by GW astronomy is truly a new lens through which Kerr BHs can be studied, even if the current constraints on the spin parameters of BBHs derived from GW detections are still limited by the detector sensitivities [31].

At the moment, the constraints on the BH spin parameters are essentially from the inspiral signal of the GW detection. The waveform generated by spinning BBHs is affected by the primary and secondary spins, through the dimensionless Kerr parameter, combined to the individual orientations of the spins w.r.t the orbital angular momentum of the system. From the post-Newtonian (PN) development of the inspiral waveform of GWs, one can demonstrate that the first dominant term of the expansion which includes spin parameters appears in the phase of the waveform and is a combination of each BH mass and spin. We often refer to as the "effective" spin parameter

$$\chi_{eff} = \frac{c}{G}\Big(\frac{\vec{S}_1}{m_1} + \frac{\vec{S}_2}{m_2}\Big).\frac{\hat{L}}{M_{tot}}, \tag{1.15}$$

where $\hat{L}$ is the normalized orbital angular momentum of the system, M_{tot} is the total mass, and $\vec{S}_i$ are the spin vectors of each black hole. We can also define a new spin parameter, the dimensionless spin magnitude χ_i following [51], such that

$$\chi_i = \frac{c}{G}\frac{S_i}{m_i^2}, \tag{1.16}$$

where $S_i = ||\vec{S}_i||$ is the norm of the spin vector of each black hole. Using Eq. 1.16, we can rewrite the effective spin parameter in an equivalent form as

$$\chi_{eff} = \frac{m_1\chi_1 cos\theta_1 + m_2\chi_2 cos\theta_2}{m_1 + m_2}, \tag{1.17}$$

where χ_i are the dimensionless spin for each component of the binary, each bounded between $[0, 1]$ and $cos\theta_i = \hat{L}.\vec{S}_i$ is the cosine of the tilt angle between the spin $\vec{S}_i$ and the orbital angular momentum $\hat{L}$ bounded in $[-1, 1]$.

Beside the projection of the spins onto the orbital angular momentum, one can also project the individual spins onto the orbital plane of the system. The amount of spin in the plane causes the binary system to precess. Similarly, as for χ_{eff}, the dominant term arising from the PN expansion of the waveform that encapsulates precession effects of the orbital plane is another combination of mass and spin and is called χ_p, the precession spin, defined as

$$\chi_p = max\Big[\chi_1 sin\theta_1; \frac{4q+3}{3q+4}q\chi_2 sin\theta_2\Big], \tag{1.18}$$

where $q = m_2/m_1$ is the mass ratio of the binary. Fig. 1.5 shows a summary of all spin parameters at play when considering a BBH system.

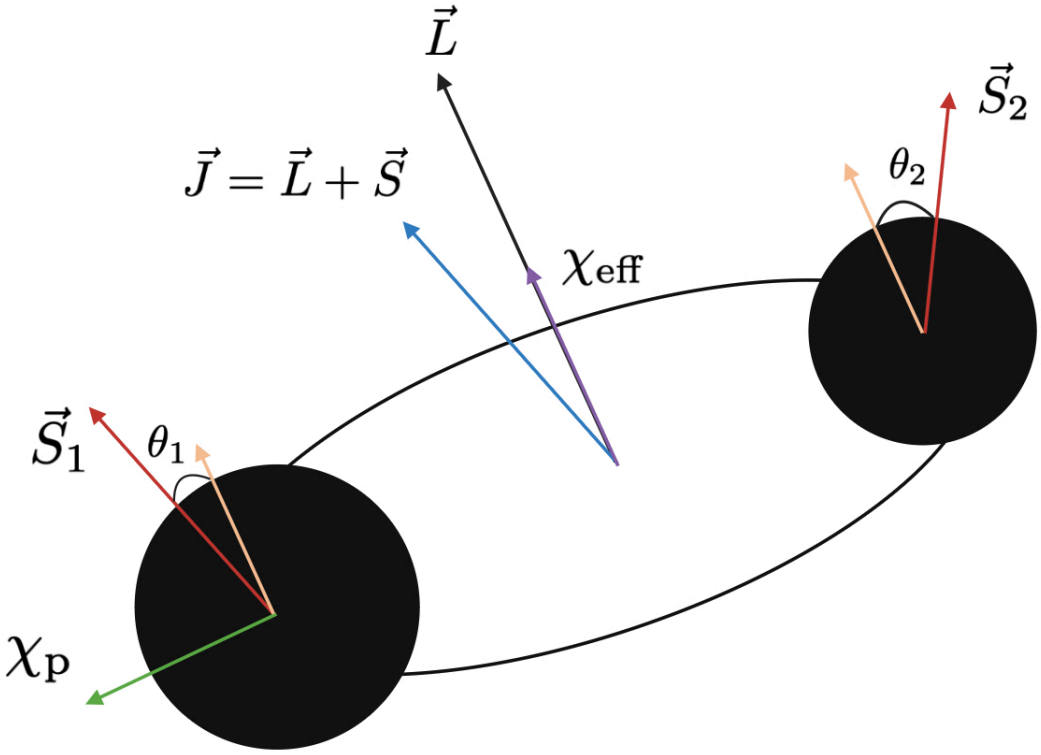

Figure 1.5: Scheme of a binary system, with spin components. The black vector is the orbital angular momentum $\vec{L}$, the blue vector is the total angular momentum $\vec{J}$, the red vectors are the spin vectors $\vec{S}_1$ and $\vec{S}_2$, the green arrow the precession spin parameter χ_p and the purple arrow is the effective spin parameter χ_{eff}. The two angles θ_1 and θ_2 are the tilt angles of the system, defined by the angle between the orbital angular momentum and the spin magnitudes of each object.

The effect of the spin on the GW signal emitted by the merging system can dramatically vary. As discussed above, the spin can have some impact on the phase of the GW, but this is not all. According to [52, 53], the alignment of the spin with respect to the angular orbital momentum is crucial too and plays an important role in the total energy radiated by the BBH merger. This increase of energy radiated (under the form of GWs) in directly linked to the duration of the inspiral phase of the merger. The duration of the merger, and the number of orbits that the two BHs go through are correlated with the spin configuration, such that if both spins are aligned with respect to the orbital angular momentum, the time before merger will be delayed, resulting in a greater energy radiation under GWs, respectively the anti-aligned spin configuration will make the merger happen quicker hence less energy are radiated [54, 55]. As explained in Sec. 2, which details the population models for the spins in the case of a

binary system, the total spin configuration can be deduced from the measurement of the effective and precession spin parameters.

1.4 THE STANDARD MODEL OF COSMOLOGY

This section sets the fundamentals of modern cosmology, from the cosmological principle, the Hubble-Lemaître law and the widely accepted ΛCDM model. We introduce the concepts and mathematical tools needed to describe the Universe through the lens of the GR, leading to the vastly accepted Big-Bang model and the current cosmology's state-of-the-art. Finally, we discuss one of the main problem that the standard model of cosmology can not reconcile so far, the Hubble tension.

1.4.1 Fundamental principles of modern cosmology

Modern cosmology relies on a concept known as the cosmological principle, which holds important implications for our understanding of the Universe. This principle is composed of two fundamental axioms. The first axiom states that our Universe is homogeneous, meaning that every region of the Universe exhibits the same physical properties and, is uniform. In a homogeneous Universe, no point in space holds a special place- each point is considered equivalent. The second axiom implies that our Universe is isotropic, indicating that there is no preferred direction in space. In other words, the Universe appears the same in all directions. Initially, the notion of homogeneity may seem wrong considering our observations, especially when considering the heterogeneous distribution of matter, such as galaxies and galaxy clusters, at smaller scales. However, the homogeneity hypothesis becomes relevant when examined on a sufficiently large scale ($\mathcal{O} \sim 250 Mpc$), far beyond the scope of individual galaxy clusters. At these larger scales, the Universe indeed appears uniform, with matter distributed evenly across space.

On top of the large scale homogeneity and isotropy, observations across the Universe suggest the Universe is expanding uniformly. In particular, galaxies seems to be moving away from Earth. This stands as one of the cornerstones of observational cosmology, at greater distances galaxies recede at faster rates. These observations can only be true, considering the cosmological principle, if the Universe itself is expanding uniformly in all directions. This expansion, which maintains homogeneity, implies that the mean rate of change in separation of pairs of objects like galaxies, with a separation d, is directly proportional to their receding velocities. Regardless of an observer's location within the Universe, they would perceive a similar pattern of galaxies moving away from them. Hubble and Lemaître, looking at the proper velocity of galaxies, formalized this idea through the renowned Hubble-Lemaître law [56–59]

$$v = H_0 d, \tag{1.19}$$

where d is the physical distance between any object and its observer, often given in *Mpc*. The receding velocity of the observed object is given by v, usually measured in $km.s^{-1}$, and H_0 is the Hubble constant, evaluated today. The typical unit for the Hubble constant is $km\,s^{-1}\,Mpc^{-1}$. The relation in Eq. 1.19 between the recession velocity of an object and its distance to the observer was found by Hubble in 1929 using a set of twenty-two galaxies. Nowadays, a significant number of new measurements have been added to the original Hubble diagram, as shown in Fig. 1.6. Observational

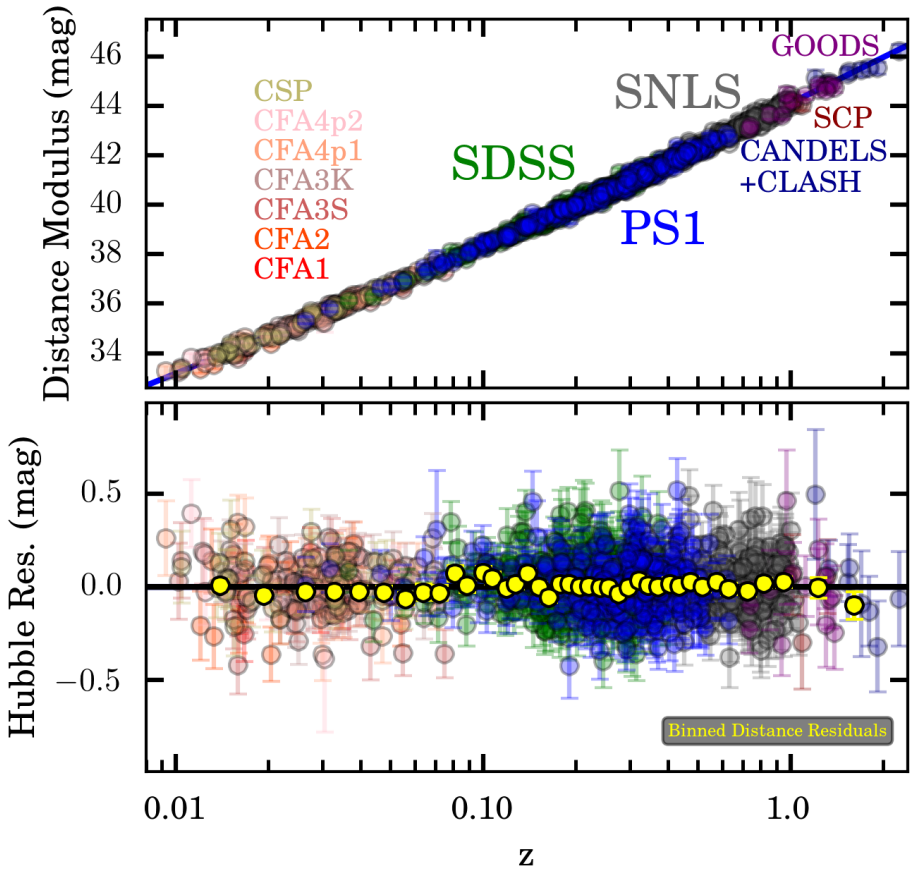

Figure 1.6: Hubble diagram obtained with the Pantheon sample of 1048 type Ia supernovae (2018). **Top:** Distance modulus for all the supernovae as function of the redshift. **Bottom:** Hubble residuals to the best fit cosmology. The supernovae displayed in the diagram are a collection from different surveys like Pan-STARRS (PS), Sloan Digital Sky Survey (SDSS), Supernovae Legacy Survey (SNLS) and the Carnegie Supernovae Project (CSP). Figure taken from [60].

evidence of the Universe's expansion indicates that it was much smaller in the past compared to its current size. This concept of a smaller, denser Universe serves as a foundational principle of the Big Bang Theory, which emerged in the late 1940s [61]. According to this theory, the expansion of the Universe implies that galaxies were much closer together in the past, leading to significantly higher densities and temper-

atures. The Big Bang model suggests that the Universe originated from an energetic explosion at a theoretical time point, and has since evolved into its present state. The prevailing cosmological model, known as the Λ-Cold-Dark-Matter (ΛCDM) model, begins with the Big Bang and describes the Universe as predominantly composed of cold (non-relativistic) dark matter, along with ordinary matter (baryonic matter). The inclusion of the cosmological constant, denoted by Λ, is essential in the Λ*CDM* model to account for observational data [62]. This cosmological constant represents the presence of dark energy, which seems to act as an "anti-gravitational" force, accelerating the expansion of the Universe. The Λ*CDM* model thus provides a framework that aligns with most of our observations and understanding of the Universe's evolution.

The ΛCDM model traces the history of the Universe from its early ages, approximately 13.7 billion years ago. The universe emerged from a singularity in an energetic event known as the Big Bang. Initially, the Universe was made of a hot, dense plasma filled with matter and photons. Within this primordial soup, small perturbations propagated as sound waves, creating regions of over- and under-densities known as Baryonic Acoustic Oscillations (BAO) [63]. As the universe expanded and cooled, atoms formed, marking the epoch of recombination. During this period, matter and light decoupled, allowing photons to travel freely through space. These ancient photons, remnants of the Universe's early stages, are detectable today as the cosmic microwave background (CMB) [64]. The fluctuations imprinted in the primordial plasma persisted as tiny temperature variations in the CMB. Over time, gravity acted upon the uneven distribution of matter, causing over-dense regions to contract and the gas to cool down. Within these denser regions of cold gas, stars and then galaxies eventually formed, giving rise to the cosmic structures we observe today. The universe evolved into a complex web of matter, characterized by clusters of galaxies interconnected by filaments. Remarkably, this large-scale structure retains the imprint of the BAO, providing a window into the Universe's early history. Despite the gravitational pull of matter, the expansion of the Universe continues unstopped. As time goes, the influence of dark energy becomes increasingly prominent. In the Λ*CDM* model, the dark energy acts as an unknown force, driving the acceleration of the Universe's expansion.

In summary, the Λ*CDM* model give a good representation of the current observations: the web of galaxies, galaxy clusters, and cosmic voids that form the structure of the Universe. Thanks to the gravitational interaction between dark matter and ordinary matter, the model manages to elucidate the formation and evolution of these cosmic structures. Observational studies and simulations support this explanation, highlighting excellent agreement between predicted and observed galaxy clustering models. Furthermore, the Λ*CDM* model responds in a relevant way to the observed abundances of primordial light elements, such as hydrogen, helium and lithium. By integrating information from primordial nucleosynthesis and studies of the interstellar medium, the model provides a robust framework for understanding the production and distribution of these building blocks. Observations of primordial abundances closely match the model's predictions, thus reaffirming its explanatory

prowess. Based on observational evidences, and in particular from the statistical tem-

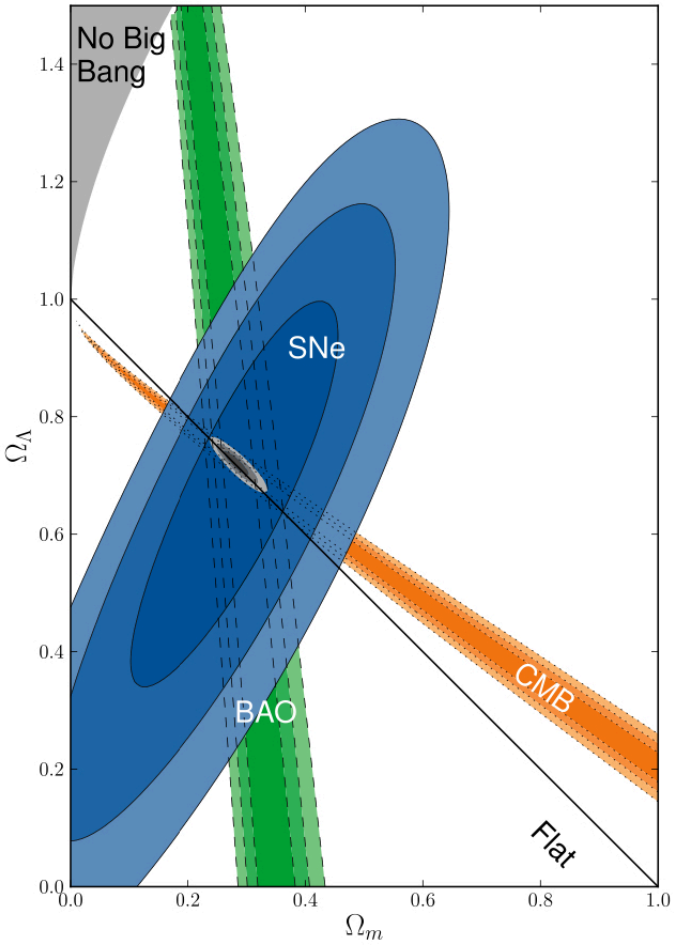

Figure 1.7: 68.3%, 95.4% and 99.7% confidence regions of the $(\Omega_m, \Omega_\Lambda)$ plane, from the supernovae type Ia (blue), CMB (orange) and BAO (green). This result includes both statistical and systematics errors estimations. The gray region is the middle is the overlapping contour between the three probes. Figure taken from [65].

perature fluctuations of the CMB, the BAO and the supernovae type Ia (see Fig. 1.7), the proportion of dark matter, baryonic matter and dark energy in the Universe today can be estimated. Results published in 2018 from the Planck collaboration shows that the actual Universe is made of $\sim$ 69.2% of dark energy, $\sim$ 25.8% of dark matter, $\sim$ 4.84% of baryonic matter and close to $\sim$ 0% of radiation [61]. Based on these results, the current state of the Universe is thought to be dominated by the dark energy, causing it to be in an accelerated expansion.

1.4.2 One metric to describe them all

Just like we used a metric to describe the geometry of space-time around a Kerr BH, we also need a metric to understand the question of an expanding Universe. In the case of a Universe, where the cosmological principle is verified, this peculiar metric

is named the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, after the physicists that found this solution in the late 1920s [56, 66–68]. The FLRW metric is an exact solution of the EFEs of GR. The EFEs describe how the space-time geometry relates to the distribution of mass and energy in that space-time, the FLRW exact solution tells us how space-time is curved in the presence of matter and energy[69]. The cosmological principle implies that, for any Universe's geometry, the metric of the universe must be of the form

$$ds^2 = -c^2dt^2 + a^2(t)ds_3^2, \tag{1.20}$$

where ds^2 is the distance, in terms of space-time interval, between two events. The second term in Eq. 1.20, $-c^2dt^2$, corresponds to the "time" separation between these events, and the last term $a^2(t)ds_3^2$ to the "space" separation. The space separation is highly dependent of the geometry of the Universe, so that ds_3^2 exact expression will change depending on whether we consider a flat, spherical or hyperbolic geometry. The latter will be discussed in the following sections.

$a(t)$ is usually called the scale factor or the expansion factor, this parameter is of a high importance because it relates the proper (physical) distance and the comoving distance. The comoving distance is the distance defined in the set of coordinates called comoving coordinates. The comoving coordinates evolve as the Universe is expanding, leaving the objects with a set of fixed coordinates. Even when stationary, two events in space will experience a change in physical distance between them as time progresses, owing to the expansion of the universe. This expansion is quantified using the scale factor $a(t)$, which is conventionally set to one at the present time $a(t_0) = 1$, serving as a reference point for measuring the evolving distance between the events. $a(t)$ determines how large-scale distances in space change with time.

From the FLRW metric, the relation between the proper distance d(t) at any time t, the scale factor, and the proper distance today $d(t_0)$ is given by

$$d(t) = \frac{a(t)}{a(t_0)}d(t_0). \tag{1.21}$$

Because the scale factor is defined to be one at the present day, Eq. 1.21 simplifies in

$$d(t) = a(t)d(t_0), \tag{1.22}$$

and this form directly tells us that the proper distance between $d(t)$ to events in space-time at any time t, is equal to their proper distance today $d(t_0)$ (also referred to as comoving distance), multiplied by the scale factor. From Eq. 1.22, one can take the time derivative of it, giving

$$\dot{d}(t) = \dot{a}(t)d(t_0). \tag{1.23}$$

And by injecting the expression of $d(t_0)$ from Eq. 1.22 inside Eq. 1.23, one can derive the Hubble-Lemaître law seen in Sec. 1.2.1, such that:

$$\dot{d}(t) = \frac{\dot{a}(t)}{a(t)}d(t) = H(t)d(t), \tag{1.24}$$

where $\dot{a}(t)/a(t)$ is the Hubble parameter $H(t)$, $\dot{d}(t)$ is again the proper distance and $d\dot{(t)}$ is the velocity of an object. Eq. 1.24 holds significant importance as it establishes a connection between the empirical parameter $H(t)$, as discovered by Hubble, and the expansion parameter within the Friedmann equation. The Hubble term in Eq. 1.24 is denoted $H(t)$ instead of H_0, signifying that this relationship between Hubble's parameter and the scale factor $a(t)$ remains valid not only at the present moment but at any given time.

As mentioned above, the FLRW metric is a solution of the EFEs, so by using the FLRW solution inside the EFEs, one can obtain the so-called Friedmann equations [66, 67]. These are two differential equations describing how the scale factor $a(t)$ evolves with time, they can be written as

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2(t)} + \frac{\Lambda c^2}{3}, \tag{1.25}$$

and

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3}. \tag{1.26}$$

In Eq. 1.25 and Eq. 1.26, $\rho = \rho(t)$ is the density and $p = p(t)$ the pressure of the Universe. The Universe's curvature is denoted by k, G is the gravitational constant and Λ is the cosmological constant. In the above equation, a, p and ρ depends on the time, but the spatial curvature of the Universe k does not, in fact it can take three different values, that give rise to three distinct geometries

$$k = \begin{cases} +1 & \text{Spherical Universe} \\ 0 & \text{Flat Universe} \\ -1 & \text{Hyperbolic Universe.} \end{cases} \tag{1.27}$$

These three possible values for the curvature govern the shape of the Universe at very large scale, as shown in Fig. 1.8. The spherical universe is often called the "closed" universe, in opposition to the "open" geometry of the hyperbolic Universe.

The cosmological constant in the Friedmann equations acts as a form of energy with a negative pressure that drives the expansion of the Universe. The pressure $p(t)$ is created by the matter content in the Universe and $\rho(t)$ can be seen as the energy density of the Universe.

Before introducing the next section, it is important to discuss a few parameters that frequently appear in the Friedmann equations, starting with the previously derived Hubble parameter $H(t)$. In Eq. 1.25, the Hubble parameter is defined as

$$H(t) = \frac{\dot{a}(t)}{a(t)}, \tag{1.28}$$

for any time t, but usually most of the measurements are happening at the present time t_0, all parameters estimated at this time will be given the subscript 0. In the same manner, we define the Hubble parameter at t_0, called the Hubble constant:

$$H_0 = H(t_0) = \frac{\dot{a}(t_0)}{a(t_0)} = \frac{\dot{a}_0}{a_0}. \tag{1.29}$$

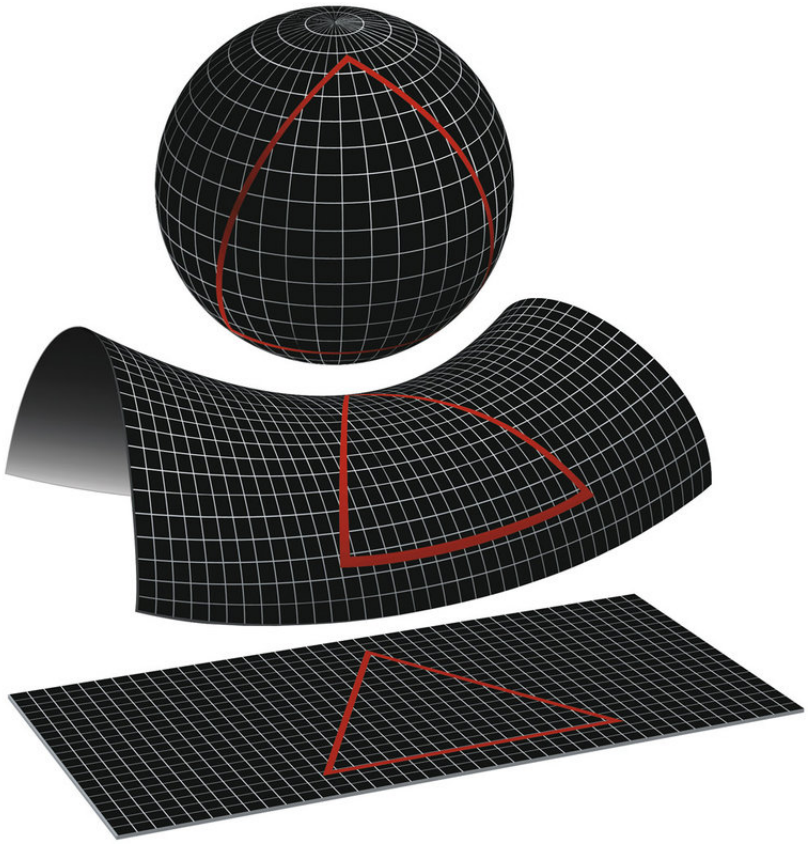

Figure 1.8: Illustration of the three possible geometry of the Universe at very large scale, represented for a 2D space, given for $k = +1$ (top), $k = -1$ (middle) and $k = 0$ (bottom). The red line is a schematic figure of a triangle seen on a spherical, hyperbolic and flat space-time. Credit: NASA / WMAP Science Team

The Hubble constant is not actually a constant, as its value changes over time since our present time is continuously evolving. However, it can be considered effectively constant on cosmological scales. For the sake of cosmological inference with GW data, which is one of the main topic of this manuscript, it is important to derive an equation where the Hubble constant, the redshift (z) and the luminosity distance (d_L) are linked to one another [70, 71]. To obtain a relationship of H_0, z and d_L, we can modify the Friedmann equations and introduce the density parameters. Eq. 1.25 can be written such as

$$H^2(t) = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2(t)} + \frac{\Lambda c^2}{3}, \tag{1.30}$$

and dividing both sides by $H^2(t)$, we obtain

$$1 = \frac{8\pi G}{3H^2(t)}\rho - \frac{kc^2}{a^2(t)H^2(t)} + \frac{\Lambda c^2}{3H^2(t)}. \tag{1.31}$$

In Eq. 1.31, three density parameters can be identified, namely the curvature density parameter Ω_k, the matter density parameter Ω_m and the dark energy density parameter Ω_Λ. The matter density Ω_m is defined as the density of the Universe ρ, over the

critical density ρ_c that the Universe should have in order to be flat ($k = 0$) and deprived of dark energy ($\Lambda = 0$). Hence, looking at Eq. 1.31, the matter density parameter can be expressed as

$$\Omega_m = \frac{8\pi G}{3H^2(t)}\rho = \frac{\rho}{\rho_c}, \tag{1.32}$$

where the critical density is

$$\rho_c = \frac{3H^2(t)}{8\pi G}. \tag{1.33}$$

Similarly, the curvature density and the dark energy density are defined as

$$\Omega_k = \frac{-kc^2}{a^2H^2(t)}, \tag{1.34}$$

and

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H^2(t)}. \tag{1.35}$$

Finally, replacing all density parameters inside Eq. 1.31 lead to the following equation

$$1 = \Omega_m + \Omega_k + \Omega_\Lambda. \tag{1.36}$$

These densities have a direct dependency on the Hubble parameter, therefore their values are changing with time. A similar description of Eq. 1.31 can be written, when considering the values of the density parameter at the time t_0, time at which $H(t) = H_0$

$$\frac{H^2}{H_0^2} = \frac{\rho}{\rho_0}\Omega_{m,0} + \frac{1}{a^2(t)}\Omega_{k,0} + \Omega_{\Lambda,0}, \tag{1.37}$$

where today's density parameters are equal to:

$$\Omega_{m,0} = \frac{8\pi G}{3H_0^2}\rho_0, \quad \Omega_{k,0} = \frac{-kc^2}{H_0^2}, \quad \Omega_{\Lambda,0} = \frac{\Lambda c^2}{3H_0^2}. \tag{1.38}$$

Now that the density parameters as well as the Hubble constant are properly introduced, within the context of the standard model of cosmology and the Friedmann equations. We can relate them to the quantities of interest for GW cosmology, namely the Hubble constant, the cosmological redshift and the luminosity distance. To clarify, the cosmological redshift z is the redshift which originates only from the expansion of the Universe, and not the peculiar velocity of the object. A signal redshifted by the cosmological expansion of the Universe will see its wavelength stretched as it propagates, the further the signal emission from an observer, the more the signal's frequency will change. One can show that the cosmological redshift relates to the scale factor such that

$$\frac{a(t_o)}{a(t_e)} = 1 + z, \tag{1.39}$$

where the scale factors are taken when the light from the source is emitted ($a(t_e)$) and when the light is observed ($a(t_o)$). Usually, the observation time t_o corresponds to the present day, hence Eq. 1.39 can be simplified, since by definition $a(t_0) = 1$:

$$a(t) = (1+z)^{-1}. \tag{1.40}$$

In a matter dominated universe, the ratio of the universe's density over the present day density is proportional to $\rho/\rho_0 \propto a(t)^{-3}$. Under this assumption and by substituting Eq. 1.40 inside Eq. 1.37, a new quantity called the dimensionless Hubble parameter, $E(z)$, can be expressed as

$$E(z) = \frac{H(z)}{H_0} = \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0}}, \tag{1.41}$$

which directly relates the Hubble constant to the cosmological redshift. In order to connect it to the luminosity distance d_L, it is easier to assume a geometrically flat Universe such that $\Omega_{k,0} \sim 0$, this assumption stays reasonable since the state-of-the-art measurement of the density parameters support such value. For a flat Universe, the luminosity distance directly relates to the comoving distance D_c such that

$$d_L = (1+z)D_c = (1+z)\frac{c}{H_0}\int_0^z \frac{dz'}{E(z')}. \tag{1.42}$$

Eq. 1.42 connects the luminosity distance of a source, to the Hubble constant, the density parameters of the Universe and the cosmological redshift. The comoving distance D_c by definition is the distance that factors out the effect of the expansion of the Universe, hence defining a distance that does not change across time under the Hubble flow [71]. Alongside the comoving distance, we define the so-called comoving volume V_c, which corresponds to the volume in which the number densities of non-evolving objects locked in the Hubble flow are constant with redshift- it corresponds to the volume in comoving coordinates in which the density of objects are the same regardless of how far away we look. Following the cosmological principle stating that the Universe is isotropic and homogeneous, the distribution of matter should be uniform when looked at large enough scales, this leads to the expression of the differential of the comoving volume w.r.t the cosmological redshift for a uniform in comoving volume matter distribution:

$$\frac{dV_c}{dz} = \frac{c^3}{H_0^3}\left(\int_0^z \frac{dz'}{E(z')}\right)^2. \tag{1.43}$$

For GW cosmology purposes, Eq. 1.42 is central, as we will see in the following Sec. 1.5, the luminosity distance of a GW source can be directly inferred from the detected signal and the redshift from direct or statistical methods, leading to a measurement of the Hubble constant H_0. But before diving into the intricacies of GW cosmology, Sec. 1.4.3 will present the current state-of-the-art of H_0 measurements.

1.4.3 A cosmological problem: The Hubble tension

In this section we present the most recent measurements of the Hubble constant H_0, coming from diverse experiments, but especially comparing the early and late Universe measurements [3, 72]. We discuss the different cosmological probes, and their results. We insist in particular on an actual cosmological problem, also known as the Hubble tension. This tension emerges from independent measurements of the local value of the Hubble parameter, that show a strong disagreement.

Currently, the ΛCDM cosmological model constitutes the standard framework for understanding the large-scale structure and evolution of the universe. Its success lies in its ability to reconcile theoretical predictions with a large amount of observational data at various cosmic epochs. The main achievement of the ΛCDM model is its precise agreement with observations of the CMB, the relic radiation, emitted approximately 380000 years after the Big Bang. Experiments such as those carried out by the Planck satellite have carefully measured temperature fluctuations in the CMB. These observations align impeccably with the model's predictions, affirming its fidelity in describing the beginnings of the universe.

Gravitational lensing, a phenomenon predicted in Einstein's theory of GR, provides additional evidence in favor of the ΛCDM model. Strong and weak gravitational lensing phenomena, resulting from the bending of light by massive structures, agree perfectly with the model predictions. Lensing observations, particularly of galaxy clusters and large-scale structures, reflect the expected gravitational signatures, thus consolidating the model's position. In summary, the ΛCDM cosmological model appears to be a robust and complete framework, harmonizing theoretical predictions with a large amount of observational data.

However, despite some agreement with the current observations, the measured values of the Hubble constant H_0 are in strong tension. As depicted in Fig. 1.9, the measurements of H_0 are often categorized into "early" and "late" measurements. The early measurements refer to high redshift epochs, estimating the Hubble constant at a different time and recovering the local value through the ΛCDM model. The late ones correspond to small redshifts measurements, directly inferring the local value of the Hubble parameter. The leading result from the early Universe has been obtained with the Planck satellite, with a method based on the anisotropies in the CMB. The Planck satellite measured the Hubble constant by analyzing the CMB radiation. In a perfectly homogeneous and isotropic Universe (cosmological principle), the CMB would be uniform. However, some small quantum fluctuations in the early Universe's plasma density occurred. When the Universe cooled and the recombination happened, these fluctuations were imprinted on the CMB. The Planck satellite observed these fluctuations and their angular dependencies on the sky. By precisely estimating the size of the anisotropies in the CMB, Planck provided data on the Universe's composition and expansion. The scale of these variations allowed the Planck collaboration to estimate the Hubble constant at $H_0 = 67.27 \pm 0.60\ km\,s^{-1}\,Mpc^{-1}$ in 2018 [61].

The leading late Universe measurement of the Hubble constant has been and still is obtained from the observation of type Ia supernovae. These objects have the intrinsic

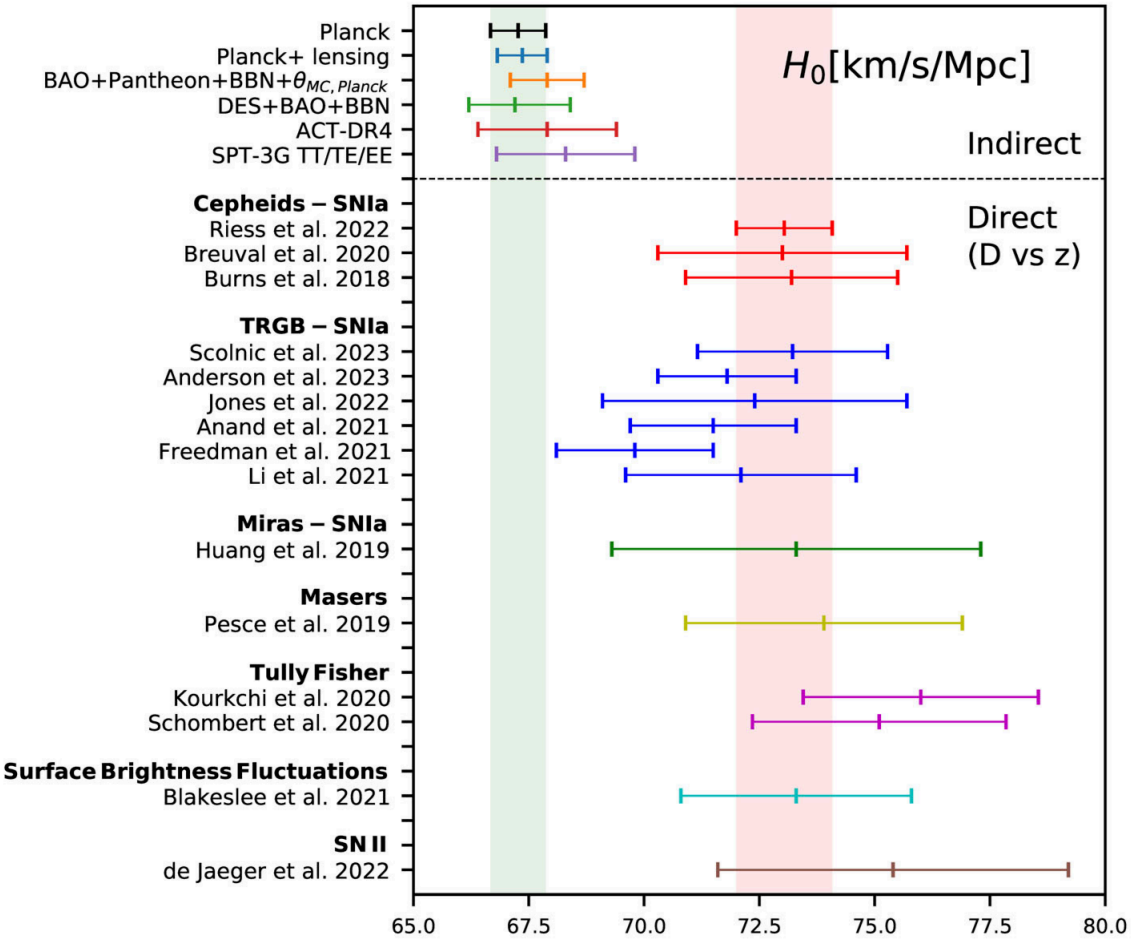

Figure 1.9: Graphic taken from [73, 74], illustrating the most recent different values of the Hubble constant from various experiments listed on the left side of the figure. The upper half are characterized as "indirect" (early universe) and the bottom half as "direct" (late universe). The two vertical bands are the best estimation of the Hubble constant in both families, having a 4.5σ disagreement.

properties of being "standardizable" candles, i.e. they have a known absolute magnitude, which combined to an estimate of their apparent magnitude, lead to a measure of their distances. Usually, type Ia supernovae are calibrated using Cepheid variable stars, which give them the name of "standardizable" candles. The redshift of the type Ia is obtained by localizing the host galaxy of the source. In 2020, the Equation of State of Dark Energy (SHOES) collaboration estimated the local value of the Hubble constant at $H_0 = 73.2 \pm 1.3\ km\,s^{-1}\,Mpc^{-1}$ [75]. SHOES used an expanded sample of 75 Cepheid stars to recalibrate the extragalactic distance ladder, and their best H_0 estimation is found with a 4.2σ tension respectively to the Planck 2018 value discussed above. In a similar fashion, other experiments, both looking at the early and late Universe managed to measure the Hubble constant, as shown in Fig. 1.9.

From the state-of-the-art results shown in Fig. 1.9, there is a clear discrepancy in different values of H_0, often referred as the Hubble tension, highlighting the 4.2σ de-

viation between Planck and SHOES [3]. While this tension is not new and keeps getting stronger and stronger over the years, there seems to be no obvious explanation to it. So far, the literature seems to agree on two possibilities, the presence of unknown systematics, or a proof that the ΛCDM model is incomplete. For the early Universe measurement, it is likely that systematics errors are not predominant, but the model dependency of the method makes it sensitive to the ΛCDM hypothesis. For the late Universe case, the method is more prone to be affected by systematic errors, since several measured quantities with imperfect astrophysical models have to be combined to obtain H_0. But the current attempts to modify the ΛCDM model to solve the Hubble tension such as the modified gravity models, the early dark energy models or the extra radiation models do not seem to provide an answer without introducing other problems. So the question remains, is the ΛCDM model the correct description of the Universe ?

1.5 GRAVITATIONAL WAVE COSMOLOGY

Previously, we briefly mentioned that GW cosmology analysis can only be performed if the redshift of the GW source is somehow determined. In 1986, Schutz mentioned for the first time the use of GWs as "*Standard Sirens*" for cosmology, proposing two methods based on BNS mergers [70]. GWs are called *Standard Sirens*, as a reference to the supernovae type Ia (standard candles) used as a tool to measure distances in cosmology, we make a distinction in the name because supernovae type Ia are "seen" while GWs are "heard". In the case where the BNS merger is perfectly localized and its electromagnetic counterpart (EMC) is detected, the Hubble constant H_0 could be estimated through the identification of the host galaxy and its redshift. If the EMC is not detected, all galaxies within the localization volume are considered, and the true redshift of the source can be statistically inferred from the redshifts of these galaxies. By stacking redshift information and H_0 estimates, the true value of the Hubble constant would emerge. More than two decades later, Del Pozzo developed a method based on Schutz's idea, using Bayesian inference to estimate H_0, regardless of whether an EMC of a BNS is detected [76]. Following this, several new methods emerged for determining the redshift associated with the GW source. These include using galaxy catalogs, which do not require an EMC, or methods based solely on GW data where the redshift is inferred by breaking the degeneracy between the source-frame masses and the detector-frame masses [77–81]. Today, new methods based on BNS, BBH, or neutron star-BH (NSBH) mergers can infer the Hubble constant. In this section, we present a review of the three main GW-based methods for cosmology: the EMC, the galaxy catalog, and the *dark siren*.

The EMC method, also called "*bright siren*", in analogy with the light emitted at the moment of the merger has been the first one able to give a precise measurement of the Hubble constant, based on a GW detection [82–84]. The BNS merger GW170817, seen by the Advanced LIGO and Virgo detectors on the 17th of August 2017 was simul-

taneously observed in the X-rays by the Fermi satellite [85–87]. During the neutron star merger, part of the energy was dissipated under the form of GWs, but also as EM emission that is usually called a kilonovae. The GW source, was estimated to be at around 40 Mpc from Earth and well localized by the LIGO-Virgo network, and 1.7s after the GW detection, a prompt gamma-ray-burst, also emitted by the BNS merger, was detected and led to the precise identification of the host galaxy NGC4993. The first ever value of the Hubble constant, from the combined detection of a GW and an EMC, was estimated at $H_0 = 70^{+12}_{-8}\, km\, s^{-1}\, Mpc^{-1}$, in agreement with both main measurement defining the Hubble tension (see Fig. 1.10) [88, 89]. The uncertainties of this first cosmological measurement are mainly driven by the estimation of the luminosity distance from the GW signal, since d_L presents a strong degeneracy with the orbital inclination angle of the system, whereas the redshift of the galaxy NGC4993 is precisely estimated after correction for peculiar velocities effects. After this date, the excitement about future BNS detections for cosmology became stronger and stronger, and some forecast studies showed that a 1% estimation of the Hubble constant could be reached with less than one hundred GW events like GW170817 [90–92]. Such precision, with similar order of magnitude as the ones obtained with the supernovae or the CMB, would resolve the Hubble tension puzzle. Up to now, after first three observing runs of the LVK scientific collaboration (O1, O2, O3), GW170817 remains the only merger seen through both GWs and EMC. Improving the chance to detect such incredible GW event has been at the heart of recent GW detector upgrades.

Out of the $\mathcal{O} \sim 100$ GW events detected since the start of O1 in 2015, only a few of them (including GW170817) contained at least one neutron star among their components, making the EMC-based method less applicable [24–27]. And binary mergers of two BHs or mixed systems with a BH and a NS are not expected to produce any strong EM emission. As a result, we have increasingly focused on other GW cosmology methods where an EMC is not needed, and the redshift of the source is estimated differently. As Schutz suggested, one can use external catalogs of galaxies with pre-measured redshifts, covering large portions of the sky. The GW detection provides a measurement of the luminosity distance and a sky localization volume, and all galaxies within this volume are considered as potential hosts. This type of method can be very convenient since it can be applied to all kinds of mergers (BBHs, NSBHs, BNSs), but it also has pitfalls. The first is that current galaxy surveys are not complete at high redshift, meaning some GW sky localizations could fall in regions where not all galaxies have been observed. The second issue is the inverse situation, where too many galaxies fall into the detection volume, making the statistical inference of the true redshift less precise. Recently, several studies have used the galaxy catalog method to infer H_0 from the GW catalogs [79, 93]. The galaxy catalog method has been utilized with GW events observed across all three runs of the LIGO, Virgo and KAGRA observing runs [4], using 47 CBC mergers and the GLADE+ galaxy catalog. The value of $H_0 = 67^{+13}_{-12}\, km\, s^{-1}\, Mpc^{-1}$ was estimated from the 47 events, and combined to the H_0 measurement obtained with GW170817, this led to the best estimation of the Hubble constant from GW observations $H_0 = 68^{+8}_{-6}\, km\, s^{-1}\, Mpc^{-1}$, as shown in Fig. 1.10.

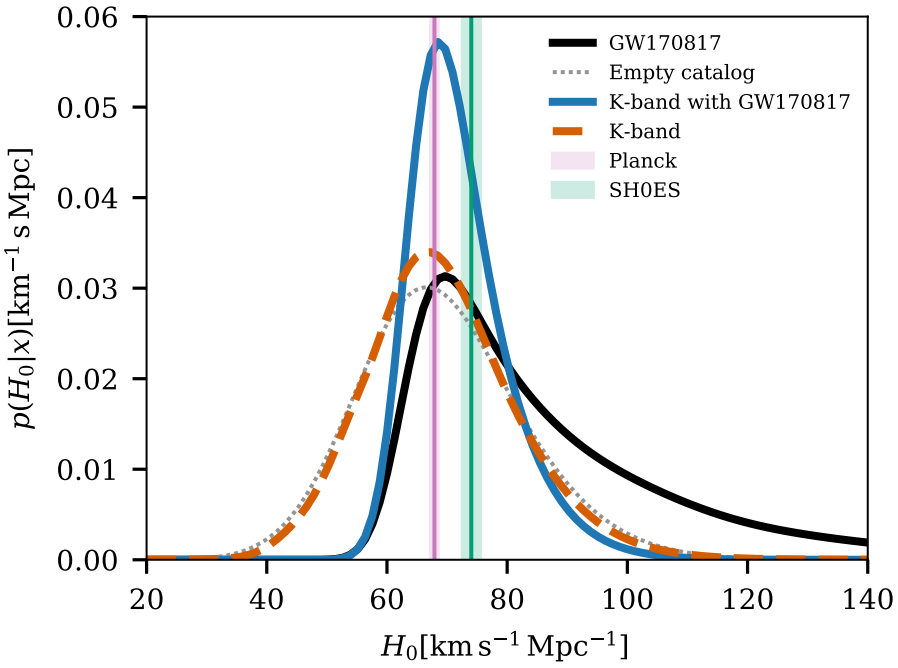

Figure 1.10: Figure taken from [4], of the Hubble constant posteriors obtained from several methods. The orange dotted line: H_0 posterior using the galaxy catalog method with GLADE+ in the K-band, with 47 GW detections and a fixed population model. The black line: H_0 posterior obtained with the EMC method one GW170817. The gray dotted line: H_0 posteriors from the dark sirens analysis with no information from galaxy catalog. The blue line: combined posterior between the galaxy and EMC methods. The green and pink vertical lines are the value of the Hubble constant obtained with the supernovae type 1a and the CMB respectively.

The use of the two combined methods showed a 40% improvements compared to the EMC method alone with one event. As explained in Sec. 2 and in the following paragraph, the mass spectrum parameterization of the mergers plays a crucial role during the inference process of H_0, as well as the merger rate. We note here that the result called "empty catalog" in Fig. 1.10 is solely driven by the population assumptions, i.e. the fixed values of the parameters used to describe the mass spectrum and merger rate of the GW mergers. This choice can have a strong effect on the Hubble constant estimation, which is discussed in [4, 94], and explore more generally in Sec. 3. Recent developments have shown that the H_0 inference using GW data without EMC must also jointly infer the population parameters in order to reduce the effect of systematics [79, 93].

The previous two methods, namely the EMC and the galaxy catalog, have proven to be efficient to estimate the Hubble constant, but both have also shown pitfalls. The first one is precise, but the rarity of GW event with a detectable EMC makes it difficult to use, and the second one highly depends on the completeness of the galaxy catalog. As seen before, the redshift of the source is the key to GW cosmology, but can not be

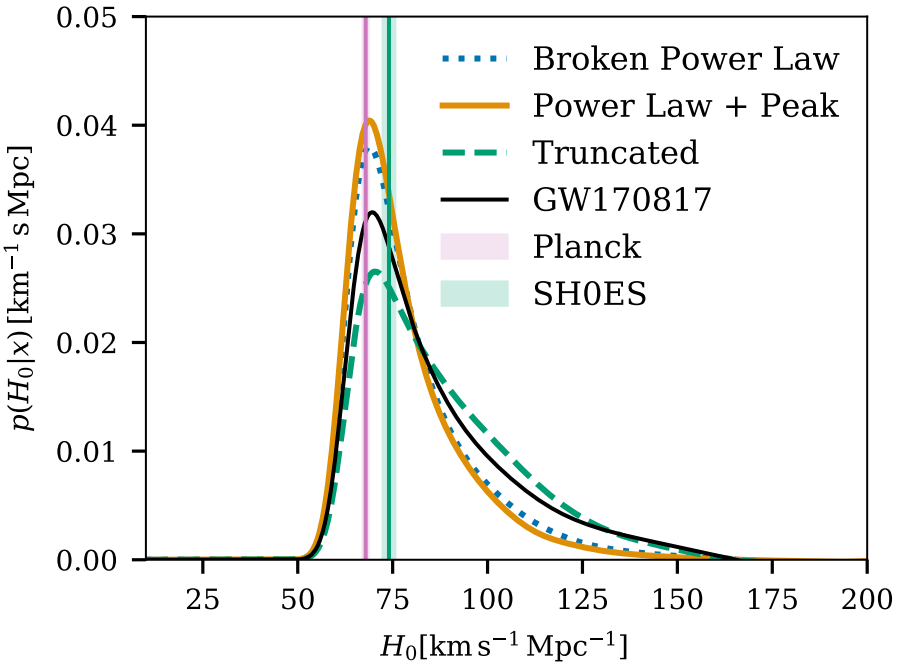

Figure 1.11: Figure taken from [4], of the Hubble constant posteriors obtained from a population only analysis of 42 BBHs using three different mass parameterizations and combined with the result of the EMC method on GW170817. The green and pink vertical lines are the value of the Hubble constant obtained with the supernovae type 1a and the CMB respectively.

obtained directly from the GW signal if no EMC or galaxy catalog is used. However, while GW detections do not provide direct information on the redshift of the source, the signal is directly proportional to the redshifted masses, as seen from Eq. 1.8 and Eq. 1.10. The detector frame masses $m_{\rm det}$ are fully degenerate with the redshift and the source frame masses m_s such that

$$m_{\rm det} = (1+z)m_s, \tag{1.44}$$

which means that the redshift of the source can be derived if the source frame masses are somehow determined. In order to break the mass-redshift degeneracy, parametric model describing the source frame mass spectrum and the merger rate of CBCs have to be passed to the analysis. In particular, some studies have demonstrated that it is possible to jointly infer the Hubble constant alongside the population parameters for the mass and the merger rate [95]. This method is often referred to as the population-only inference, or the *dark siren* analysis. The dark siren inference is crucial for GW cosmology, since no external data are needed, but one has to be aware that the choice of the parameterization of the astrophysical population can have an impact on the inferred H_0. In the latest cosmology paper from the LVK collaboration [4], 42 BBHs from the third transient catalog are analyzed to jointly infer the Hubble constant and the population parameters. In this work, three different mass models were tested: a power-law, a broken power-law and a more complex model made of a

power-law plus a Gaussian peak. More details about these models and their effects on the inferred H_0 will be given later in Sec. 3. Additionally, we found out that a simple power-law was highly disfavored as the correct mass population model. The Hubble constant was estimated at $H_0 = 44^{+52}_{-24}\, km\, s^{-1}\, Mpc^{-1}$ using the broken power-law and $H_0 = 50^{+37}_{-30}\, km\, s^{-1}\, Mpc^{-1}$ with the power-law plus peak. Combining this result with the BNS GW170817 led to the value of $H_0 = 68^{+12}_{-8}\, km\, s^{-1}\, Mpc^{-1}$, which is close to 17% better than the previous results based on GW events from the second observing run. The dark sirens analysis's constraining power on H_0, is closely linked to the mass spectrum structures, in fact it was shown that the current values obtained in [4] were mainly driven by the reconstructed peak around $35\, M_\odot$, that was told to "set a scale" for the redshift distribution of BBHs.

Additionally, for binary systems where at least one of the two object is a neutron star, the mass-redshift degeneracy can be resolved with the use of the neutron star equation of state via the tidal deformability parameter that is degenerated with the Hubble constant [96].

The GWs-only method discussed above provide a new approach for cosmology studies, which does not need extra information other than the GW detections. But as we will demonstrate in Sec. 3, this method is susceptible to systematics errors induced by the population models. Just like the presence of features in the mass spectrum of BBHs are the origin of the constraining power of such method, they can also be the source of biases, if not accounted properly in during the inference.

1.6 CONCLUSIONS

Exploring the cosmos, from the vast structure of the universe to the subtle vibrations of GWs and their sources, is a profound journey towards understanding the fundamental physics of our Universe. This introductory chapter combines basics of cosmology, GW physics, and BH astrophysics, to provide a coherent picture of our universe.

GWs, vibrations of space-time predicted by Einstein's general theory of relativity, are a revolutionary tool in this search. They provide a unique window into the most violent and energetic events in the Universe, such as BH mergers and neutron star collisions. The discovery and analysis of these waves not only confirm theoretical predictions, but also give new insights into the behavior of matter and energy in extreme conditions, as well as a new probe to test the standard model of cosmology.

The study of BH mergers, particularly the complex dynamics of rotating Kerr BHs, enriches our understanding of the universe. These compact objects, with their intense gravitational fields, are the primary sources of known gravitational wave emissions. The physics governing their formation, evolution, and interaction in binary systems sheds light on fundamental forces in the universe, although it is not entirely understood.

Modern cosmology is based on a standard model that describes the expansion of the Universe from the Big Bang to its present state. Built on principles such as the

cosmological principle and the dynamics embodied in the Friedman equations, this model provides a basis for understanding the large-scale structure and evolution of the cosmos. Nevertheless, ongoing challenges such as the Hubble tension emphasize the need for innovative observation methods to improve our measurements and theories.

Together, these elements form an integrated approach to our Universe understanding. The intersection of cosmology, GW astronomy and BH physics not only solves existing issues, but also opens up new paths for discoveries. As observational methods advance and theoretical models evolve, we are approaching the discovery of the deepest secrets of the Universe, from its origin to its final destination.

2 BAYESIAN INFERENCE FOR COSMOLOGY AND BINARY BLACK HOLE POPULATION

The majority of the content in this chapter relies on the work in a published paper, in which I am the second author. This article is a method paper of the cosmological analysis pipeline ICAROGW [77], that was published at the same time as the public release of the code ICAROGW 2.0 (ICAROGW). I also deeply contributed to the official LVK review of ICAROGW 2.0. The Sec. 2.3 and the Sec. 2.4 correspond the method described in the paper. The Sec. 2.6 taken from the spin section of the same paper, and the Sec. 2.5 has been taken from another paper in which I am the first author [94].

2.1 INTRODUCTION

In this chapter, we introduce the hierarchical Bayesian inference scheme and methodologies used to infer the astrophysical properties of BBH populations alongside with the cosmological parameters, based on gravitational waves data from compact binary mergers. Such analysis is required for population inference since the number of observations is limited and has to be corrected for a selection bias. The selection bias, also referred as the Malmquist bias [97], obstructs the observation of the entire binary black hole merger population, hence biases our sample if not accounted in our analysis. Moreover, GW data used for population inference studies often exhibit measurement uncertainties of their source parameters due to the fact that such signals are found deep inside instrumental noise of the detectors, making it significantly more difficult to extract accurate information on the source population. As a result, the inference of BBH population properties and of cosmological parameters requires to have advanced numerical and statistical techniques in order to handle incomplete and noisy data.

This chapter is organized as follows. Sec. 2.2 gives an introduction to the basic principles of Bayesian inference and in particular hierarchical Bayesian inference in the context of GWs. In Sec. 2.3, we present ICAROGW, a python package for inference of astrophysical population properties of noisy and incomplete observations. In Sec. 2.4, we go over the different methodologies to perform such analysis, using solely the gravitational wave data (Spectral Siren), adding extra information from galaxy catalogs, or combining the gravitational wave result with its electromagnetic counterpart. In Sec. 2.5, we introduce a new addition to the Bayesian inference, allowing for the estimation of spin parameters of binary black holes, via two distinct spin models. We bring forth the new parameterization for the hierarchical inference, as well as a discussion on the potential impact of spins for cosmology inference. Finally, Sec. 2.7 concludes, and summarize the possibilities of such analysis for GW cosmology purposes.

2.2 BAYESIAN INFERENCE OF POPULATION PROPERTIES WITH GRAVITATIONAL WAVE DATA

2.2.1 Bayesian inference

Following the detection of a GW passing through our detectors, we start studying the individual properties of this peculiar source, such as its astrophysical properties or its origin, these make each signal unique. But as the number of detections increases with the observing time and the enhanced sensitivity of the detectors, it becomes interesting to switch our way of looking. Instead of exploring the physics of individual GW sources, it can be crucial to look at them as an ensemble, a population of sources. Studying the entire population, as a whole, may help us understand what are their common characteristics, what are their differences, and even allow us to perform complex studies as GW cosmology. However, one have to be careful while trying to infer

common properties of a population based on detected data only, because of the limited sensitivity of the detectors. In fact, all sources are not equally detectable and for various reasons. As discussed in Sec. 1.2.2, this induces selection effects that have to be taken care of during the analysis, in order to get unbiased results. The purpose of Bayesian and hierarchical Bayesian inference is to infer the population properties of the entire population of sources, while deconvolving the selection effects.

More formally, in the context of GWs, the Bayesian inference concept can be explained like that. An experiment or a detector gives us access to a set of data $\{x\}$, representing information on the individual sources of the detected GWs. We want to use these data, to extract new information about the overall population of GW sources. Usually, the data $\{x\}$ are explained via model M which is governed by a set of model parameters θ. This model, if correct, should be able to make scientific predictions. Based on the data $\{x\}$, the model M and its intrinsic parameters θ, we can construct the likelihood probability

$$P(\{x\}|\theta, M), \tag{2.1}$$

that tell us what is the probability of observing the data $\{x\}$, given the model M and a specific set of values for the parameters θ. The value of this probability changes for different values of the parameters θ, and this give us insights on which choice of parameters would be best to explain the data $\{x\}$. In the context of GW analysis, we already "know" the data $\{x\}$, we are rather interested in the opposite of Eq. 2.1

$$P(\theta|\{x\}, M), \tag{2.2}$$

which tell us the probability of having certain values for our set of parameter θ, given a model M and some data $\{x\}$. Using probability laws, one can relate the so-called likelihood $P(\{x\}|\theta, M)$ and the probability $P(\theta|\{x\}, M)$ such that

$$P(\theta|\{x\}, M)P(\{x\}|M) = P(\theta, \{x\}|M) = P(\{x\}|\theta, M)P(\theta|M). \tag{2.3}$$

By flipping Eq. 2.3, we obtain a new form that is referred to as the Bayes' Theorem [98]

$$P(\theta|\{x\}, M) = \frac{P(\{x\}|\theta, M)P(\theta|M)}{P(\{x\}|M)}. \tag{2.4}$$

The likelihood probability $P(\{x\}|\theta, M)$ is often written as $\mathcal{L}(\{x\}|\theta, M)$, so we will keep this notation for the rest of the manuscript. The term $P(\theta|M)$ is called the prior probability, the prior corresponds to the prior knowledge we have about the choice of parameter θ, before any new detections. The prior probability of the parameters can take any form, but is often chosen uniform in order to keep the analysis as agnostic as possible. The denominator $P(\{x\}|M)$ is named the evidence, sometimes noted $\mathcal{Z}$, it corresponds to the likelihood probability marginalized over all the parameters θ. The evidence is used to understand how well a model is able to fit the data, after integrating over all combinations of θ, or perform model comparison with the Bayes factor between the derived evidences of two different models on the same set of data. $P(\theta|\{x\}, M)$ is called the posterior probability, this term gives us the probability of having certain value given the data and the model.

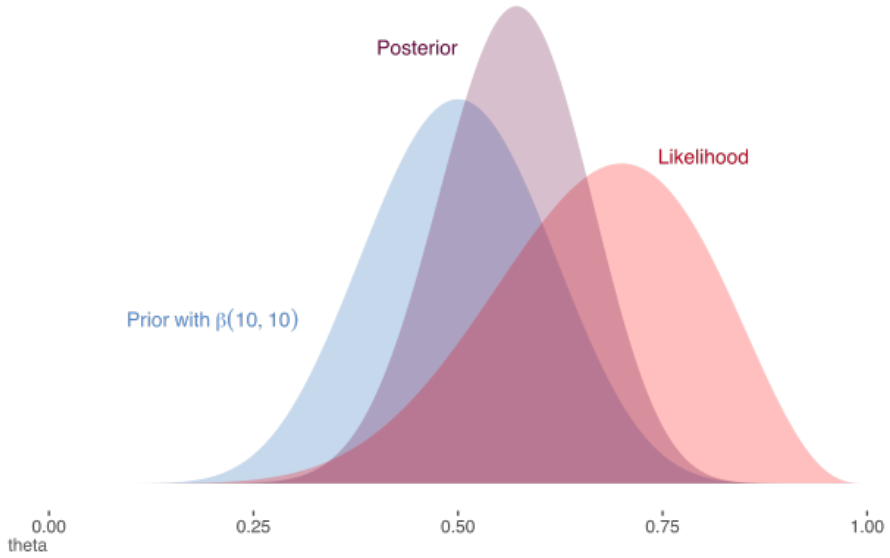

Figure 2.1: Schematic view of the density probabilities at play in the Bayes theorem. In blue is the prior probability $P(\theta|M)$ drawn from a Beta distribution $\beta(10, 10)$, in red is the likelihood probability $P(\{x\}|\theta, M)$ and in purple is the resulting posterior distribution $P(\theta|\{x\}, M)$, that is the combination of the two. The evidence here is not represented. The model used is governed by only one parameter θ in $[0, 1]$. Figure taken from *m − clark*.

Fig. 2.1 and Fig. 2.2 illustrates how the prior distribution impacts the resulting posterior distribution, for the same likelihood distribution. A more peaked prior will have a greater influence of the resulting posterior than a rather flat prior.

The estimation of the posterior distribution is usually done to either compare models or make scientific predictions, by inferring constraints on the population parameters θ. In most cases, the model does not depend on a single parameter θ, but a set of several parameters $\{\theta_i\}$. Out of all parameters, some are considered as noise and other as parameters of interest, the noisy one should be marginalized over. If the total posterior is "known", the marginalization can be done such that

$$P(\theta_{int}|\{x\}, M) = \int P(\{\theta\}|\{x\}, M) d\theta_{noise}, \tag{2.5}$$

where θ_{int} are the parameters of interest to be constrained from the posterior, and θ_{noise} are the set of non-interesting parameters to be marginalized over. In the end, the crucial quantities in Bayesian inference are not the full posterior distribution, but rather integrals of this distribution. In the specific case of ICAROGW, the Bayesian framework introduced in Sec. 2.4, the posterior is approximated using Markov-chain-Monte-Carlo methods [99].

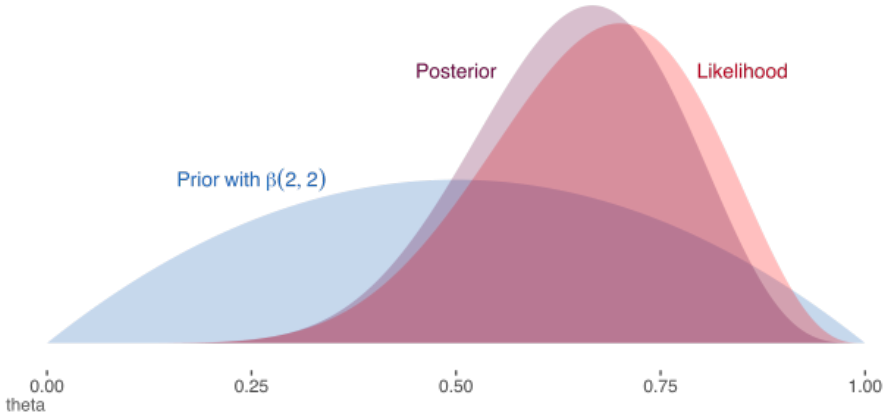

Figure 2.2: Schematic view of the density probabilities at play in the Bayes theorem. In blue is the prior probability $P(\theta|M)$ drawn from a Beta distribution $\beta(2,2)$, in red is the likelihood probability $P(\{x\}|\theta,M)$ and in purple is the resulting posterior distribution $P(\theta|\{x\},M)$, that is the combination of the two. The evidence here is not represented. The model used is governed by only one parameter θ in $[0,1]$. Figure taken from *m − clark*.

2.2.2 Hierarchy and selection effects

In GW population or cosmology studies, we often refer to Bayesian inference as being hierarchical, meaning that the parameters at plays can be organized in a "cascading" manner. We have a set of intrinsic parameters for each GW source (θ = {masses, spins, ...}), they are the population-level parameters. And we have the common parameters which governs the shape of the distributions of the population-level parameters. For example, the distribution of the mass could be chosen to be a Gaussian with two parameters, in this case the population parameter would be the mass, and the mean and standard deviation of the Gaussian the second level of parameter. The combination of the population-level parameters with the parameters that governs the shape of the distributions are called the hyperparameters Λ, to which we add the cosmological parameters (H_0, Ω_m,...). The term hierarchical is used here to highlight the fact that we use a set of detections from a population to infer the hyperparameters Λ, in which we have two levels of inference. We estimate the parameter θ for each of the sources, and then use them to infer the parameters on the population level.

As mentioned before, the detected GW signals represent only a subset of the entire population of sources, and this effect can be an important source of bias. In the hierarchical Bayesian framework, this is formulated with a question: for a set of population parameters θ, is the source detectable with the current GW detector network and sensitivity ? And because a detector is not stationary, meaning that its noise and sensitivity

are not the same over time, the probability of detecting the same source evolves over time. So far, this effect is mitigated by computing the detection probability of a source with parameters θ, for multiple noise realizations, through a ranking statistic like the signal-to-noise ratio (SNR). We manage to compute such probability over the entire parameter space by injecting simulated GW signals with all combination of parameters θ into real realizations of the detector noise, and using analysis pipeline searches to rank them as detectable or not. This entire process is called "injection search", and allows us to estimate the detection probability of GW signals.

2.3 ICAROGW: INFERRING COSMOLOGY AND ASTROPHYSICS WITH OBSERVATIONS OF GRAVITATIONAL WAVES

This section presents the foundations of ICAROGW [77], a python based code developed for the inference of population and cosmological properties of compact binary coalescence via gravitational wave data. ICAROGW is one of the two official pipelines of the LIGO-Virgo-KAGRA (LVK) Collaboration for cosmology studies, and has been used during the previous observing runs; its development is still ongoing today [4].

2.3.1 A hierarchical Bayesian inference framework

The purpose of this Bayesian framework is to infer the cosmological and population parameters, Λ, that governs the CBC rate as a function of the intrinsic GW parameters θ. This CBC rate is usually denoted

$$\frac{dN_{CBC}}{dt_s d\theta}(\Lambda). \tag{2.6}$$

For cosmology purposes, the GW parameters θ are usually the sources frame masses of both black holes, and their redshift denoted (m_1^s, m_2^s, z), but the intrinsic spin parameters can also be at play here. For instance, in Sec. 2.4.3, the implementation of spin models is discussed. For simplicity, a summary table of all parameters stated in this section is shown in Table. 2.1.

The detection of GW events can be described as an inhomogeneous Poisson process in the presence of selection biases [100, 101]. For a set of N_{obs} GW signals detected, over an observation time T_{obs}, the probability of having a specific GW dataset $\{x\}$ given some population hyperparameters Λ and described by their source parameters θ can be written under the form of the following hierarchical likelihood:

$$\mathcal{L}(\{x\}|\Lambda) \propto e^{-N_{exp}(\Lambda)} \prod_i^{N_{obs}} T_{obs} \int d\theta \times \mathcal{L}_{GW}(x_i|\theta,\Lambda) \frac{1}{1+z} \frac{dN_{CBC}}{d\theta dt_s}(\Lambda). \tag{2.7}$$

The hyperparameters Λ describe the BBH population, but also include the cosmological parameters. For a chosen flat ΛCDM cosmology, the cosmological parameters are

Parameter	Description
H_0	Hubble constant, i.e. the local expansion rate of the Universe today.
N_{CBC}	Compact binary coalescence merger rate
N_{obs}	Number of detected (observed) real gravitational wave signals.
N_{det}	Number of detected signal out of the total number of generated injections.
N_{exp}	Expected number of gravitational wave events.
$\{x\}$	Gravitational wave data set.
θ	Intrinsic gravitational wave parameters of the source.
Λ	Hyperparameters of the binary black hole population.
t_s, t_d	Source frame time and the detector frame time.
R_0	Local value of the compact binary merger rate density, per giga parsecs cube, per year.
p_{det}	Detection probability of GW events.
$N_{eff,i}$	Effective number of posterior samples per GW event.
$N_{eff,inj}$	Effective number of injections.
(m_1^s, m_2^s)	Primary and secondary masses of the compact binary merger in the source frame.
z	Cosmological redshift of the CBC source.
$\vec{\chi}$	Vector of the normalized spin parameters of the CBC source.
$w_{i,j}$	Weight representing the number of CBC mergers per unit of time (see Sec. 2.3.2).
s_j	Weight proportional to the number of CBC mergers detected per detector frame time (see Sec. 2.3.2).

Table 2.1: Summary table of the parameters and quantities used in the Bayesian framework of ICAROGW.

the Hubble constant H_0 and the fraction of matter density today $\Omega_{m,0}$. All other parameters are population parameters governing the shape of the BBH mass spectrum, as well as the spin spectrum or the merger rate spectrum (see Sec. 2.4.3 and Sec. 3.3.1). In Eq. 2.7, the term $\frac{1}{1+z}$ accounts for the time dilatation between the source frame time and the detector frame time. An alternative form for Eq. 2.7, that allows to suppress the information about the rate of mergers can be derived by assuming a "scale-free" prior $\pi(N_{exp}) \propto 1/N_{exp}$ on the expected number of detection [102, 103]. The equivalent form of Eq. 2.7 with a "scale-free" prior can be expressed as

$$\mathcal{L}(\{x\}|\Lambda) \propto \prod_{i=1}^{N_{obs}} \frac{\int \mathcal{L}_{GW}(x_i|\theta,\Lambda)\frac{dN_{CBC}}{dt_s d\theta}(\Lambda)d\theta}{\int p_{det}(\theta,\Lambda)\frac{dN_{CBC}}{dt_s d\theta}(\Lambda)d\theta}. \tag{2.8}$$

Assuming a scale-free prior can be highly useful, especially while doing inference on simulated data, it allows us to forget about the local CBC merger rate density R_0 in the inference, without putting to risk the overall result. In section. 3, all the results are obtained with a scale-free prior. In Eq. 2.7 and Eq. 2.8, several important terms are at play, such as the individual likelihood $\mathcal{L}_{GW}(x_i|\theta,\Lambda)$ of each GW event, that gauges the error on the estimation of intrinsic parameters θ for a specific GW event $\{x_i\}$, or the detection probability $p_{det}(\theta,\Lambda)$ which determines the probability of detecting a certain CBC at a redshift z, characterized by its parameters θ and Λ.

2.3.2 Monte-Carlo integration and numerical stability estimators

As shown above, the first term that requires a numerical evaluation is the single event likelihood $\mathcal{L}_{GW}(x_i|\theta,\Lambda)$, which is related to the measurement error of the GW parameters θ. When dealing with detected GW signals, the single event likelihood is not directly given by these data, but by a set of posterior samples (PEs), defined by

$$p(\theta|x_i,\Lambda) \propto \mathcal{L}_{GW}(x_i|\theta,\Lambda)\pi_{PE}(\theta|\Lambda). \tag{2.9}$$

The PEs are obtained after the detection of each single event. The quantity $\pi_{PE}(\theta|\Lambda)$ is the prior with which the posterior samples were drawn. Hence, $p(\theta|x_i,\Lambda)$ are the PEs of the source frame parameters for each detected GW event $\{x_i\}$. Within ICAROGW, the integral of the individual likelihood in Eq. 2.7 and the integral of the probability of detection in Eq. 2.8 are approximated using numerical Monte-Carlo integral methods, summing over the PEs [104]. Taking the single likelihood for example, it is possible to re-write it such as:

$$\int \mathcal{L}_{GW}(x_i|\theta,\Lambda)\frac{dN_{CBC}}{dt_s d\theta}(\Lambda)d\theta \approx \frac{1}{N_{s,i}}\sum_{j=1}^{N_{s,i}}\frac{1}{\pi_{PE}(\theta_{i,j}|\Lambda)}\frac{dN_{CBC}}{dt_s d\theta}(\Lambda)\Big|_{i,j} \tag{2.10}$$

$$\approx \frac{1}{N_{s,i}}\sum_{j=1}^{N_{s,i}} w_{i,j}. \tag{2.11}$$

The indices i and j point respectively to the event and the posterior samples. The quantity $w_{i,j}$ can be seen as a weight in Eq. 2.11, representing the number of CBC mergers per unit of time. As the integral of the single likelihood is approached via a sum over a finite number of posterior samples j per GW event i, numerical instabilities could arise if the number of effective posterior samples used to evaluate the integral is too small. As proposed in [105],

$$N_{eff,i} = \frac{(\sum_j^{N_{s,i}} w_{i,j})^2}{\sum_j^{N_{s,i}}(w_{i,j})^2}, \tag{2.12}$$

is a new quantity used to monitor how many samples per event are truly helping the Monte-Carlo integration. Before running a joint cosmological and population analysis,

a threshold value is chosen, if for a specific GW event and population model, this condition is not fulfilled, the value of associated likelihood will automatically be put to zero. In general, for cosmological and population study, the value of N_{eff} is chosen around 10, in order to have at least 10 posterior samples per event per model. The effective number of posterior samples, N_{eff}, plays a crucial role in determining the numerical stability of the likelihood during inference. If N_{eff} is too low, the inference process can overly rely on a small subset of samples, leading to an unstable likelihood estimation and unreliable final results. Currently, there is no universally accepted rule for selecting the optimal value of N_{eff}; instead, it is typically determined through empirical testing and experience gained from previous inferences.

Beside the single event likelihood discussed above, the expected number of detection $N_{exp}(\Lambda)$ also needs Monte-Carlo integration. This term can be expressed as a function of the probability of detection $p_{det}(\theta, \Lambda)$ following Eq. 2.13 below:

$$N_{exp}(\Lambda) = T_{obs} \int p_{det}(\theta, \Lambda) \frac{dN_{CBC}}{dt_d d\theta} d\theta. \tag{2.13}$$

$N_{exp}(\Lambda)$ is directly related to the estimation of the selections biases within the Bayesian framework. Moreover, the expected number of detections can further be derived, by developing the probability of detection in the integral such that

$$p_{det}(\theta, \Lambda) = \int_{x \in detectable} \mathcal{L}(x_i|\theta, \Lambda) dx, \tag{2.14}$$

where the probability of detection is the integral over all individual GW data $\{x_i\}$, of each single event likelihoods. Nevertheless, from the GW data alone, there is no easy way to get an analytical form for the probability of detection. To address this matter, Monte-Carlo simulations are the most widely accepted way for selection biases computation. In the field of GW analysis, it is often referred to as "injections", to speak about simulated GW events. The simulated events are injected in the real noise of the detectors, and detected as real signals in order to estimate the detectable parameter space volume. Said differently, an intrinsic property of detectors is their limitation to detect all signals, in fact, there exists a volume in the parameter space (defined by all hyperparameters in Λ), where signals can not be detected, but still exist in the Universe. Using injections is how the Bayesian framework manages to correct for this selection bias. The frequency of detected injections, which varies proportionally with $p_{det}(\theta, \Lambda)$, along with the population models employed for their generation, can serve as criteria for assessing selection bias.

To evaluate the selection bias via Monte-Carlo simulations, from a set of N_{gen} simulated GW signals drawn from a prior $\pi_{inj}(\theta)$, only the subset of detected GW signals

N_{det} is given to ICAROGW. Similarly, as in Eq. 2.11, it follows that Eq. 2.13 can be computed with a Monte-Carlo integration as:

$$N_{exp}(\Lambda) = T_{obs} \int p_{det}(\theta, \Lambda) \frac{dN_{CBC}}{dt_d d\theta} d\theta \tag{2.15}$$

$$\approx \frac{T_{obs}}{N_{gen}} \sum_{j=1}^{N_{det}} \frac{1}{\pi_{inj}(\theta_j)} \frac{dN_{CBC}}{dt_d d\theta}\Big|_j \tag{2.16}$$

$$\equiv \frac{T_{obs}}{N_{gen}} \sum_{j=1}^{N_{det}} s_j. \tag{2.17}$$

The new weight s_j occurring in Eq. 2.17 is directly proportional to the number of CBC mergers per detector frame time. The prior $\pi_{inj}(\theta)$ used to draw the injections must be correctly normalized (its integral being equal to unity) in order to compute an acceptable value of expected detections. A wrong normalization of the prior would result in a wrong number of expected detections, which would influence the final correction of the selection effects. As for the effective number of posterior samples, it is possible to define an effective number of injections threshold, below which the likelihood would automatically drop to zero, in order to avoid numerical instabilities. In agreement with [106], the effective number of injections can be defined as

$$N_{eff,inj} = \frac{(\sum_j^{N_{det}} s_j)^2}{(\sum_j^{N_{det}} s_j^2 - N_{gen}^{-1}(\sum_j^{N_{det}} s_j)^2)}. \tag{2.18}$$

In [106], they show that there exists a criterion linking the effective number of injection and the number of observed GW events used in the analysis. They found that a conservative value for $N_{eff,inj}$ is given by:

$$N_{eff,inj} > 4N_{obs}. \tag{2.19}$$

Finally, it is possible to re-write the complete hierarchical likelihood defined in Eq. 2.7, using the results of Eq. 2.11 and Eq. 2.17, derived with the Monte-Carlo integrations. The logarithm of Eq. 2.7 takes the following form:

$$ln(\mathcal{L}(\{x\}|\Lambda)) \simeq -\frac{T_{obs}}{N_{gen}} \sum_{j=1}^{N_{det}} s_j + \sum_i^{N_{obs}} ln\Big(\frac{T_{obs}}{N_{s,i}} \sum_{j=1}^{N_{s,i}} w_{i,j}\Big). \tag{2.20}$$

The final form of the hierarchical likelihood written in Eq. 2.20 is the one evaluated by ICAROGW, from a set of detected GW data and a set of simulated injections.

2.3.3 Overall architecture of ICAROGW

The Bayesian framework of ICAROGW relies on several "modules", important for the joint population and cosmological inference. This section aims to clarify how the

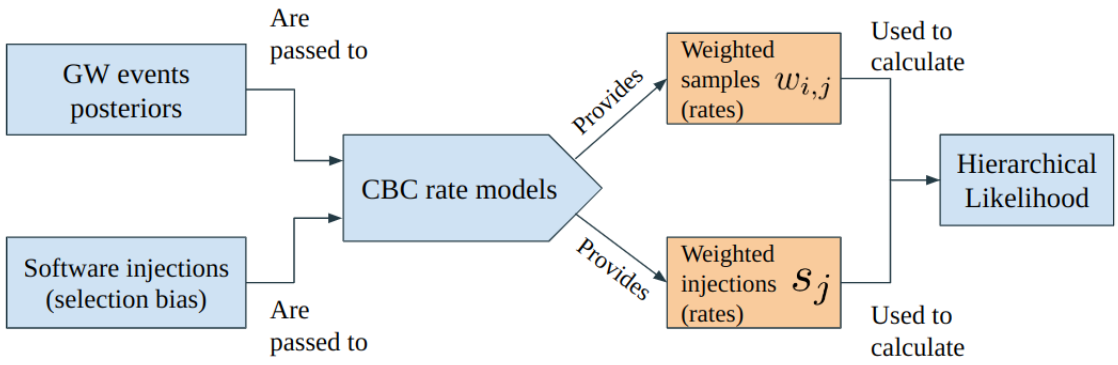

Figure 2.3: Scheme of the simplified workflow of ICAROGW. The blue boxes are the modules used for handling the injection sets, the GW posterior samples, and building the hierarchical likelihood derived in Eq. 2.20. The two yellow boxes on the other hand correspond to the evaluation of the weights shown in Eq. 2.12 and Eq. 2.19.

different modules work together and what is the role of each one of them. Fig. 2.3 presents a simplified scheme of the different modules that make the core of ICAROGW. Two distinct modules are responsible for the processing of the GW posterior samples and the injections, and once given to ICAROGW, they will both be used to construct the CBC rate models. The CBC rate models are the central entities of the code, more specifically they are classes (to be understood as python objects) used to build the merger rates defined in Eq. 2.6. Here we talk about merger rates, plural, due to the different methodologies available for the population inference. These methodologies are further explained in the Sec. 2.4. Each rate model has different intrinsic parameters θ and different hyperparameters Λ.

An important point to be aware of in ICAROGW is the frame in which lies the GW parameters θ, when going through the workflow presented in Fig. 2.3. As a matter of fact, the GW parameters θ that are provided as entries to the code lie in the detector frame, i.e. θ_d, but the GW parameters used to compute the weights $w_{i,j}$ and s_j are in the source frame (θ_s). The source frame corresponds to the "true" parameters of the source of GWs, and the detector frame parameters are the parameters obtained on Earth. These two sets of parameters θ_s and θ_d do not have to be equal, the fact that the GW travels through the Universe while it being in expansion introduces a relation between the two, that depends on the cosmology. The conversion from detector frame to the source frame of these quantities is handled by the CBC rate classes and the hierarchical likelihood so that the weights can be computed. The final step of the inference, i.e. the computation of the hierarchical likelihood and the Monte-Carlo process is taken care by a widely used python package for Bayesian inference: BILBY [107–109].

2.4 THREE METHODOLOGIES, THREE CBC MERGER RATES

In the context of cosmological and population inference from GW data, ICAROGW is able to perform the joint estimation of these properties with three methodologies. In this section, we present in Sec. 2.4.1 the method based on BBH merger solely, called the Spectral Siren method. The galaxy catalog approach is developed in Sec. 2.4.2 and finally, Sec. 2.4.3 highlights the electromagnetic counterpart method.

From the detected GW signals alone, we saw in Sec. 1.2 that the primary and secondary detector frame masses (m_d^1, m_d^2) can be estimated, as well as the luminosity distance of the source d_L. In reality, more parameters related to the source of GWs are inferred: an estimate of the sky position Ω along with the spin parameters $\vec{\chi}$ of both objects are measured. We denote the normalized spin parameters of the binary system by the vector $\vec{\chi} = (\chi_1, \chi_2)$, as introduced in Sec. 1.3. Consequently, the CBC rate as defined in Eq. 2.6, can be expanded by developing all the dependencies of θ_d. This form of the rate is now a function of the masses, the spins, the sky localization and the luminosity distance of the source, written as:

$$\frac{dN_{CBC}}{dd_L d\Omega dm_1^d dm_2^d d\vec{\chi}_1 d\vec{\chi}_2 dt_d}. \tag{2.21}$$

Note here that the CBC merger rate written in Eq. 2.21 is in the detector frame, and is consequently independent of the cosmological parameters, and in particular independent of the expansion rate of the Universe H_0. The actual dependency on the cosmological model appears when this rate in converted into the source frame merger rate, where some of its parameters such as the masses and the luminosity distance are changed. For the masses, there exists a simple relation between the detector frame and the source frame parameters, this relation as seen in Sec. 1.2 is written such that:

$$m_{1,2}^d = (1+z) m_{1,2}^s, \tag{2.22}$$

where z is the redshift of the source. The luminosity distance of the source is also related to the redshift, but through the cosmological model and parameters values as in shown in Sec. 1.4.

Expressing the relation between the detector CBC merger rate and the source frame CBC merger rate allows us to put in evidence the cosmological model dependency. Using a change of variable, the detector frame merger rate can be expressed as:

$$\frac{dN_{CBC}}{d\theta_d dt_d} = \frac{dN_{CBC}}{d\theta_s dt_s} \frac{dt_s}{dt_d} \frac{1}{detJ_{d\to s}} \tag{2.23}$$

$$= \frac{dN_{CBC}}{d\theta_s dt_s} \frac{1}{1+z} \frac{1}{detJ_{d\to s}}, \tag{2.24}$$

where the differential of the source frame time with respect to the detector frame time is expressed as $1/(1+z)$. In Eq. 2.24, the change of variable is contained within the

Jacobian of the transformation. Since only the luminosity distance and the masses are affected by this change of frame, the Jacobian is equal to:

$$\frac{1}{detJ_{d\to s}} = \frac{\partial d_L}{\partial z}(1+z)^2. \tag{2.25}$$

The remaining term in Eq. 2.25 is the differential of the luminosity distance w.r.t cosmological redshift. This term highly depends on the choice of the cosmological model. Following Sec. 1.4, the luminosity distance of a source is related to the redshift via the relation:

$$d_L = \frac{c(1+z)}{H_0}\int_0^z \frac{dz'}{E(z')}, \tag{2.26}$$

where c is the speed of light in vacuum, and $E(z)$ is the dimensionless Hubble function. Combining Eq. 2.26 and Eq. 2.25, we can write the differential of the luminosity distance:

$$\frac{\partial d_L}{\partial z} = \frac{d_L(z)}{1+z} + c\frac{(1+z)}{H_0}\frac{1}{E(z)}. \tag{2.27}$$

Given a certain cosmological model, it is then possible to fully express the CBC merger rate in the source frame, as a function of the source frame masses, spins, redshift. From this point, three methodologies of population inference are available, depending on the type of analysis wanted. For each analysis depicted below, the source frame CBC merger rate will change its parameterization, adding or suppressing population parameters depending on the type of data needed/available.

2.4.1 The Spectral sirens case

The first parameterization of the source frame CBC merger rate is for the Spectral siren method. This method is the simplest, in terms of types of data involved, since is solely based on the GW data detected. One of the first realization of Spectral siren analysis was performed in [80, 110], where they managed to jointly infer the cosmological parameters and the population properties of the CBC merger rate. For the Spectral Siren analysis, the set of detector frame parameters θ_d governing the CBC rate is $\theta_d = (m_1^d, m_2^d, d_L, \vec{\chi}_1, \vec{\chi}_2)$ and in source frame $\theta_s = (m_1^s, m_2^s, z, \vec{\chi}_1, \vec{\chi}_2)$. Note here that the spin parameters of the compact objects are not affected by the change of frame, in fact they are equal in both the source and detector frame. The Spectral siren CBC merger rate is then parameterized as

$$\frac{dN_{CBC}}{d\theta_d dt_d} = \frac{dN_{CBC}}{dd_L dt_d d\vec{\chi} dm_1^d dm_2^d} \tag{2.28}$$

$$= \frac{dN_{CBC}}{dz dt_s d\vec{\chi} dm_1^s dm_2^s}\frac{1}{1+z}\frac{1}{detJ_{d\to s}} \tag{2.29}$$

$$= \frac{dN_{CBC}}{dV_c dt_s d\vec{\chi} dm_1^s dm_2^s}\frac{dV_c}{dz}\frac{1}{1+z}\frac{1}{detJ_{d\to s}}, \tag{2.30}$$

where the last term in Eq. 2.30 is the differential of the comoving volume w.r.t redshift. This term depends on the cosmological model chosen through the following relation:

$$\frac{dV_c}{dz} = \left(\frac{c}{H_0}\right)^3 \left(\int_0^z \frac{dz'}{E(z')}\right)^2. \tag{2.31}$$

From Eq. 2.30, it is now possible to explicitly write the dependency of the CBC merger rate as a function of all population models describing the masses, the spins, and the redshift. The final form of the Spectral siren CBC merger rate is:

$$\frac{dN_{CBC}}{d\theta_d dt_d} = R_0 \psi(z;\Lambda) p_{pop}(m_1^s, m_2^s|\Lambda) p_{pop}(\vec{\chi}|\Lambda) \frac{dV_c}{dz} \frac{1}{1+z} \frac{1}{detJ_{d\to s}}. \tag{2.32}$$

In Eq. 2.32, R_0 is the local CBC merger rate density per time per comoving volume, the unit commonly used for this parameter is $[Gpc^{-3}yr^{-1}]$ (local here means at $z = 0$). $\psi(z;\Lambda)$ is a phenomenological function that models the evolution of the rate in redshift, based on the shape of the star formation rate. $p_{pop}(m_1^s, m_2^s|\Lambda)$ is the probability density function (PDF) of the phenomenological model used for the source frame masses, it is the model of the joint distribution of BBH primary and secondary masses. $p_{pop}(\vec{\chi}|\Lambda)$ is also a PDF, that governs the spin distribution, for a given spin model. Within ICAROGW, there exist several mass, spin and redshift models, that can be combined to construct the Spectral siren CBC merger rate and used for population analysis. The combined set of parameters that governs the shape of the distributions for θ_s are the one inferred during the Spectral siren analysis, jointly with the parameters of the cosmological model. In Eq. 2.32, we assume that the merger rate of CBCs can be factorized into the product of three independent terms, describing the mass, rate, and spin of the CBCs. However, more complex parameterization can be constructed, depending on the population models used for these parameters. For instance, in Sec. 4, we investigate models that account for correlations between BH spin magnitudes and masses. Additionally, recent studies have suggested a potential redshift evolution in the mass spectrum of BBHs, which could be modeled in Eq. 2.32 through a joint distribution between the merger rate and mass terms.

2.4.2 The galaxy catalog case

The galaxy catalog method can be seen as an extension of the spectral siren analysis, since the parameterization of the CBC merger rate is constructed similarly. In the galaxy catalog approach, extra information on the redshift of the source is added to the analysis. This information, on the potential redshift of the source, is statistically estimated from the support of external galaxy catalogs [70, 81, 111, 112]. Likewise, the galaxy catalog approach aims to infer the parameters of the populations models and cosmological model used to construct the source frame CBC merger rate. For the galaxy catalog method, the detector frame set of parameter θ_d and source frame θ_s are composed of the masses, the luminosity distance of the source (or the redshift), the spins, and an extra parameter Ω denoting the sky localization of the GW event

(in squared radian). Hence, the new sets of parameters are $\theta_d = (m_1^d, m_2^d, d_L, \vec{\chi}, \Omega)$ and $\theta_s = (m_1^s, m_2^s, z, \vec{\chi}, \Omega)$. Since this method for GW cosmology and population inference is not the main subject of this manuscript, fewer details about the derivation of this specific CBC rate will be given here, instead we refer the reader to [78, 79] for a complete description of the galaxy catalog approach.

Two major assumptions are made in order to derive the new CBC merger rate, the first one is that all CBC mergers happen inside galaxies, and the second one is that the number of these mergers per galaxy is proportional to the absolute magnitude M and redshift z of the galaxy. The absolute magnitude of a galaxy can be understood as its luminosity. Instead of the comoving volume, the CBC merger rate is parameterized in terms of the merger per galaxy, such that:

$$\frac{dN_{CBC}}{d\theta_s dt_s dz d\vec{m}_s d\Omega} = \int dM \frac{dN_{CBC}}{d\theta_s dt_s dz d\vec{m}_s d\Omega dM} \tag{2.33}$$

$$= \int dM \frac{dN_{CBC}}{dN_{gal} d\theta_s d\vec{m}_s dt_s} \frac{dN_{gal}}{dz d\Omega dM}, \tag{2.34}$$

where N_{gal} is the number of CBC mergers per galaxy. From Eq. 2.34, it is assumed that N_{gal} is only a function of the redshift, the absolute magnitude and the sky localization. The first part of Eq. 2.34 is very similar to the Spectral siren CBC merger rate, in fact it can be expressed correspondingly such that:

$$\frac{dN_{CBC}}{dN_{gal} d\theta_s d\vec{m}_s dt_s} = R^*_{gal,0} \Psi(z, M; \Lambda) p_{pop}(m_1^s, m_2^s | z, M; \Lambda) p_{pop}(\theta_s | z, M; \Lambda), \tag{2.35}$$

where similar terms are at play: $R^*_{gal,0}$ is the local CBC merger rate density per galaxy per year, $\Psi(z, M; \Lambda)$ is the function parameterizing the evolution of the number of CBC mergers per galaxy and the last terms are the population models for the masses and potential other population parameters like the spins. For more details on the parameterization of the evolution function, we refer the reader to [78, 79].

For the second part of Eq. 2.34, namely the number density of galaxies, we have to take into account the fact that the current galaxy catalogs do not contain all the galaxies in the Universe, hence galaxy density has to be the sum of two terms: the galaxy density from the catalog and the galaxy density which is not in the catalog. This second term is usually called the "completeness correction". Therefore, the galaxy density rate of Eq. 2.34 is given by

$$\frac{dN_{gal}}{dz d\Omega dM} = \frac{dN_{gal,cat}}{dz d\Omega dM} + \frac{dN_{gal,out}}{dz d\Omega dM}, \tag{2.36}$$

where $N_{gal,cat}$ and $N_{gal,out}$ are respectively the number of galaxy inside the catalog and the number of galaxy outside the catalog. For a detailed explanation of how the completeness corrections are computed, see [79].

Finally, by plugging Eq. 2.35 and Eq. 2.36 into Eq. 2.34, the source frame CBC merger rate inside ICAROGW is given by:

$$\begin{aligned}\frac{dN_{CBC}}{dzdt_s d\vec{\chi} dm_1^s dm_2^s d\Omega} &= \int dM \Big[R^*_{gal,0} \Psi(z, M; \Lambda) p_{pop}(m_1^s, m_2^s | z, M; \Lambda) p_{pop}(\theta_s | z, M; \Lambda) \times \\ &\times \left(\frac{dN_{gal,cat}}{dzd\Omega dM} + \frac{dN_{gal,out}}{dzd\Omega dM} \right) \Big] \end{aligned} \tag{2.37}$$

The main limitation of this method is the completeness of the catalogs, especially at high redshift. For distant GW events, with poor sky localizations, it can become very difficult to assign any galaxy as potential host, and this brings a poor estimation of the redshift. In this situation, where the galaxy catalog does not give enough information, the population assumptions (Spectral siren method) kicks in and the results of the inference are driven by the previous method.

2.4.3 The electromagnetic counterpart case

Finally, the last method, and currently the most precise one to perform GW cosmology is the electromagnetic counterpart (E.M.C.) analysis. Just like with galaxy catalogs, the CBC merger rate takes extra information on the sky position of the source, but this time from the associated E.M.C emitted during the merger of the compact objects. Just like for the galaxy catalog approach, the information of the redshift of the GW source is given by the redshift of its host, through its sky position. As emphasized in Sec. 1.5, this way of computing the redshift of the source is highly accurate since the host galaxy redshift is found with great precision, making this method competitive for GW cosmology.

The posterior samples obtained from the GW detection give estimates of the sky localization, the masses and the luminosity distance of the source, while the E.M.C gives an independent (w.r.t the GW data) and more accurate sky localization and redshift estimates of the same source. Assuming that both measurements are independent, a new likelihood described by the GW and E.M.C data has to be constructed such as $\mathcal{L}_{GW+EMC}(x_i|z, \Omega, m_1^s, m_2^s, \vec{\chi})$. And since we assume that the two measurements are independent, the likelihood can be decomposed as

$$\mathcal{L}_{GW+EMC}(x_i|z, \Omega, m_1^s, m_2^s, \vec{\chi}) \propto \mathcal{L}_{EMC}(x_i|z, \Omega) \mathcal{L}_{GW}(x_i|z, \Omega, \vec{m}_s, \vec{\chi}). \tag{2.38}$$

The E.M.C and the GW data do not carry the same amount of information about the GW source, but they both help for the estimation of the sky localization Ω and the redshift z. Consequently, the hierarchical likelihood used to perform the cosmological and population inference shown in Eq. 2.8 changes, such that the integral at the numerator becomes

$$\begin{aligned} I &= \int \mathcal{L}_{EMC}(x_i|z, \Omega) \mathcal{L}_{GW}(x_i|z, \Omega, \vec{m}_s, \vec{\chi}) \times \\ &\times \frac{dN_{CBC}}{dzd\Omega d\vec{m}_s d\vec{\chi} dt_s} \frac{1}{1+z} d\vec{\chi} d\vec{m}_s dz d\Omega. \end{aligned} \tag{2.39}$$

Moreover, the integral at denominator of Eq. 2.8 has to change too. This term, accounting for the selection bias within the ICAROGW analysis, has to be corrected with the new selection effects related to the E.M.C. Following [79, 93], it is assumed that the probability of detection for an E.M.C is always one, meaning that if there is an associated emission of light to the GW production, and that the GWs are detected, the E.M.C will always be detected too. That hypothesis permits us to avoid using a new parameterization that could model the detection probability of the E.M.C. Therefore, the denominator of Eq. 2.8, even when using the E.M.C counterpart method, only consider the GW detection for its selection effects. That way of handling the selection bias related to the E.M.C. stays valid as long as the GW detector sensitivity range is strictly lower than the one of E.M. telescopes and observatories. This assumption will probably have to be updated in the near future with the increasing sensitivity of GW detectors, as a wrong calculation of the selection effects can result in a biased value of the population parameters and also the Hubble constant [113].

2.5 THE PHENOMENOLOGICAL MASS MODELS FOR CBCS

In this section, we dive in further details into the phenomenological mass models used to describe the population of BBHs in terms of source frame masses. As explained in Sec. 1.5, the mass models play a key role in the Bayesian inference with Spectral Siren, or galaxy catalogs, they are the cornerstone to the estimation of the redshift of the GW sources. This section aims to present the different source mass models currently used in ICAROGW and more generally in recent GW cosmology analysis. Currently, Dark Siren based methods that use parametric mass models incorporate simple and flexible parameterizations to assess the distribution of source frame masses of CBCs, as in [4]. In the following subsections, the main mass models used for population inference are discussed, namely the BROKEN POWER LAW (BPL), the POWERLAW PLUS PEAK (PLP) and the MULTIPEAK (MLTP). First, the origin of each model is presented, as well as some astrophysical motivations to such models. Then we provide details about the probability density function for each one of them.

2.5.1 The Broken Powerlaw

The BROKEN POWER LAW (BPL), introduced in [114], is one of the simplest available model inside ICAROGW, and offers a basic approach to understanding various phenomena within astrophysical contexts. The left plot in Fig. 2.4 displays a schematic view of the typical PDF obtained with the BPL mass model. At its core, the BPL model comprises a power law, characterized by a smooth tapering at the lower mass edge, coupled with a distinctive break occurring at a critical mass point, denoted b. The high mass smoothing is introduced to incorporate the effect of stellar progenitor metallicity [114], which plays a crucial role in shaping the dynamics of black hole formation, hence the mass distribution of black holes in the Universe. By introducing a smooth-

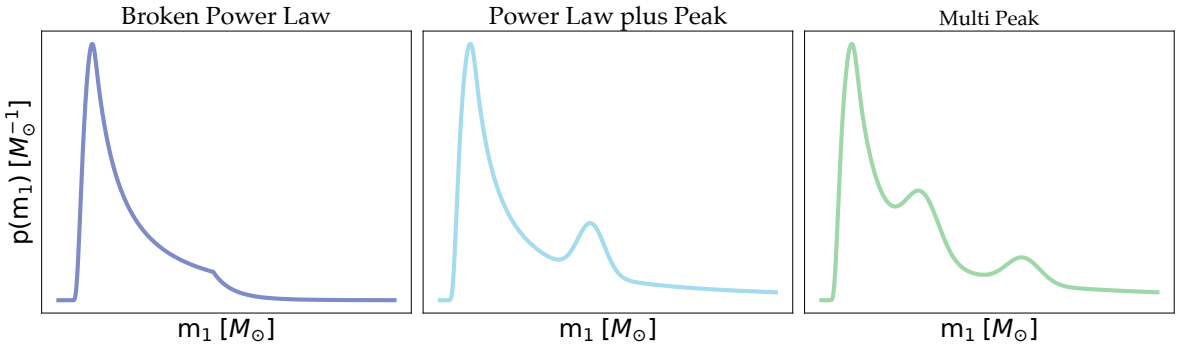

Figure 2.4: Schematic view of the main three phenomenological source mass models incorporated in ICAROGW: the *broken power law* (left), the *powerlaw plus peak* (middle) and the *multi peak* (right). The set of parameter chosen to construct these distributions are not displayed, these schemes only highlight the mass structures of each model.

ing effect, the model should capture the interplay between metallicity and black hole production.

The break in the power law serves to mirror the left boundary of the PISN gap [115]. This break reflects the different production rate of certain BBHs at peculiar source frame masses. Furthermore, the introduction of a second power law, after the breaking point, aims to mimic the existence of a secondary population of BBHs within the pair-instability supernovae gap, possibly arising from dynamically formed BBHs. This dual power law model presents the compelling aspect of being motivated by astrophysical processes that rule the BBH formation and evolution, in different astrophysical environments.

As explained above, the two power laws are attached at a breaking point b defined as

$$b = m_{min} + (m_{max} - m_{min})f, \tag{2.40}$$

where f is a scalar that belongs to $[0,1]$. When $f = 0$, the breaking point b is equal to the minimum mass, and respectively equal to the maximum mass when $f = 1$. The PDF for the primary source frame mass m_1^s is defined as:

$$\pi(m_1^s|m_{min}, m_{max}, \alpha) = \frac{1}{N}\left[\mathcal{P}(m_1^s|m_{min}, b, -\alpha_1) + \frac{\mathcal{P}(b|m_{min}, b, -\alpha_1)}{\mathcal{P}(b|b, m_{max}, -\alpha_2)}\mathcal{P}(m_1^s|m_{max}, b, -\alpha_2)\right], \tag{2.41}$$

where $\mathcal{P}$ is a truncated power law constructed as

$$\mathcal{P}(x|x_{min}, x_{max}, \alpha) = \begin{cases} \frac{1}{N_{PL}}x^{\alpha}, & (x_{min} \leqslant x \leqslant x_{max}), \\ 0, & \text{otherwise}, \end{cases} \tag{2.42}$$

where N_{PL} is the normalization factor defined as :

$$N_{PL} = \begin{cases} ln(\frac{x_{max}}{x_{min}}), & (if\ \alpha = -1), \\ \frac{1}{\alpha+1}(x_{max}^{\alpha+1} - x_{min}^{\alpha+1}), & \text{otherwise}. \end{cases} \tag{2.43}$$

The normalization factor of the BROKEN POWER LAW N is defined by

$$N = 1 + \frac{\mathcal{P}(b|m_{min}, b, -\alpha_1)}{\mathcal{P}(b|b, m_{max}, -\alpha_2)}. \tag{2.44}$$

The secondary mass distribution is parameterized as a function of the primary mass. It ensures that the secondary mass is smaller than m_1^s. The secondary mass is then given by

$$\pi(m_2^s|m_{min}, m_1^s, \beta) = \mathcal{P}(m_2^s|m_{min}, m_1^s, \beta). \tag{2.45}$$

A more detailed explanation of each of the parameter used in the PLP mass model is given in App. A.1.1.

2.5.2 The Powerlaw plus peak

The POWER LAW PLUS PEAK (PLP) mass model, initially introduced in [116], has gained popularity in various population and cosmological analysis related to GWs. Notably, it has emerged as the preferred source mass model according to the LVK Collaboration, as indicated by our recent work [4], characterizing the entire BBH population based on the GWTC-3 catalog [27]. This model integrates a power law component with a similar low mass smoothing as the BPL model, while additionally incorporating a Gaussian peak, as depicted on the middle plot in Fig. 2.4. The inclusion of this peak is motivated by the potential accumulation of BBHs just preceding the PISN gap [116]. What sets the PLP mass model apart is its enhanced flexibility: both the position (mean) and the width (standard deviation) of the Gaussian peak can vary independently of the rest of the model. This feature allows the PLP to capture high mass events or sub-structures in the mass spectrum, without altering the lower mass range of the spectrum. The parameter λ_{peak} governs the fraction of GW events that fall within the Gaussian peak.

The probability density functions that parameterize the distributions of m_1^s and m_2^s are given by

$$\begin{aligned} \pi(m_1^s|m_{min}, m_{max}, \alpha) &= (1 - \lambda_{peak})\mathcal{P}(m_1^s|m_{min}, m_{max}, -\alpha) + \\ &+ \lambda_{peak}\mathcal{G}(m_1^s|\mu_g, \sigma_g), \end{aligned} \tag{2.46}$$

with $0 \leq \lambda_{peak} \leq 1$, and

$$\pi(m_2^s|m_{min}, m_{max}, \beta) = \mathcal{P}(m_2^s|m_{min}, \beta, m_1^s). \tag{2.47}$$

As for the BPL mass model, the secondary mass is also defined with a power law conditioned on the primary mass. In Eq. 2.46, the Gaussian peak $\mathcal{G}(m_1^s|\mu_g, \sigma_g)$ is expressed as

$$\mathcal{G}_{[a,b]}(x|\mu_g, \sigma_g) = \begin{cases} \frac{1}{N_G}\frac{1}{\sigma_g\sqrt{2\pi}}\exp\left[-\frac{(x-\mu_g)^2}{2\sigma_g^2}\right], & a \leqslant x \leqslant b, \\ 0, & \text{otherwise}, \end{cases} \tag{2.48}$$

where N_G is also the normalization factor given by

$$N_G = \int_b^a \frac{1}{\sigma_g \sqrt{2\pi}} e^{-\frac{(x-\mu_g)^2}{2\sigma_g^2}} dx. \tag{2.49}$$

A more detailed explanation of each of the parameter used in the PLP mass model is give in App. A.1.2

2.5.3 The Multi peak

The *Multi peak* (MLTP) mass model, is an extension of the PLP model, used for the first time in [114]. It incorporates the same low mass smoothing, and is also built around a power law. The particularity of this mass model is the second Gaussian peak available for higher masses. The MLTP mass model is the combination of a power law plus two Gaussian peaks, as shown in the right plot of Fig. 2.4. The motivations for such model are BBHs systems that could arise from second-generation mergers, called hierarchical mergers. These binary systems, formed in dynamical environments are born from already merged black holes, hence could reach higher masses compared to first generation BBHs. The MLTP mass model is then a more flexible than the PLP model, but can also reduce into a PLP distribution if there is no support found in the data for a second Gaussian peak.

The parameterization of the MLTP mass model is the direct extension to the PLP parameterization, where the primary mass is defined as

$$\begin{aligned} \pi(m_1^s|m_{min}, m_{max}, \alpha) &= (1-\lambda)\mathcal{P}(m_1^s|m_{min}, m_{max}, -\alpha) + \\ &+ \lambda\lambda_{low}\mathcal{G}(m_1^s|\mu_g^{low}, \sigma_g^{low}) + \\ &+ \lambda(1-\lambda_{low})\mathcal{G}(m_1^s|\mu_g^{high}, \sigma_g^{high}), \end{aligned} \tag{2.50}$$

and the secondary mass m_2^s is one more time defined as a conditional power law on the primary mass

$$\pi(m_2^s|m_{min}, m_{max}, \beta) = \mathcal{P}(m_2^s|m_{min}, \beta, m_1^s). \tag{2.51}$$

A more detailed explanation of each of the parameter used in the PLP mass model is give in App. A.1.3

2.6 THE ADDITION OF CBC BINARY BLACK HOLE SPINS

In this section, we bring the attention on the addition on two phenomenological spin models within the cosmological and population inference of ICAROGW, namely the DEFAULT spin model (see Sec. 2.6.2) and the GAUSSIAN spin model (see Sec. 2.6.3) as introduced in [117, 118] and used in [119]. This addition to the code allowed us to produce

the first ever results of the joint inference between the Hubble constant H_0 and the spin parameters of the GW sources.

Prior to this work, the spin parameters of the BBHs were not used at all in GW cosmology analysis and only the source frame masses, the redshift and the cosmological parameters were of interest. The reason being that the Spectral Siren analysis is based on the degeneracy between the source frame masses and the redshift as explained in Sec. 1.5, making these parameters crucial for the joint inference. The addition of spins inside the framework of ICAROGW could allow us to look for correlations between the spin parameters and the cosmological ones. In particular, if the spins of CBC correlate with the mass of the BBHs, or the merger rate, this could enhance the constraining power of the GW based methods looking at measuring the Hubble constant.

2.6.1 The spin of binary systems

Referring to Fig. 2.5, below we discuss the main spin parameters at play when dealing with binary system of compact objects for GW analysis. By definition, the z axis (normal vector to the plan of the binary) is aligned with the orbital angular momentum $\vec{L}$. The spin properties of a binary system of compact objects can be completely characterized by 6 degrees of freedom (3 for each object): two normalized spin magnitudes (χ_1, χ_2), two tilt angles (θ_1, θ_2) and two azimuthal angles (ϕ_1, ϕ_2). The dimensionless spin magnitude $\chi_{1,2}$ encode how fast each object are spinning, and they are defined from the normalized Cartesian components of the spin vectors in Fig. 2.5 such that:

$$\chi_1 = \sqrt{s_{1,x}^2 + s_{1,y}^2 + s_{1,z}^2}, \tag{2.52}$$

$$\chi_2 = \sqrt{s_{2,x}^2 + s_{2,y}^2 + s_{2,z}^2}. \tag{2.53}$$

The tilt angles θ_1 and θ_2, are given by the angle between $\vec{L}$ and the spin magnitude vectors projected along the z axis. Often, we refer to them as the cosine of the tilt angle, since they are defined as

$$\cos\theta_1 = \frac{s_{1,z}}{\chi_1}, \tag{2.54}$$

$$\cos\theta_2 = \frac{s_{2,z}}{\chi_2}, \tag{2.55}$$

where $s_{1,z}$ and $s_{2,z}$ are the components along the z axis of the spin magnitudes vectors. The azimuthal angles $\phi_{1,2}$ are not depicted on Fig. 2.5, we will assume that these two angles have a uniform distribution.

Finally, we define the effective spin parameter χ_{eff} and the precession spin parameter χ_p. Using only these two parameters, four of the six spin degrees of freedom are

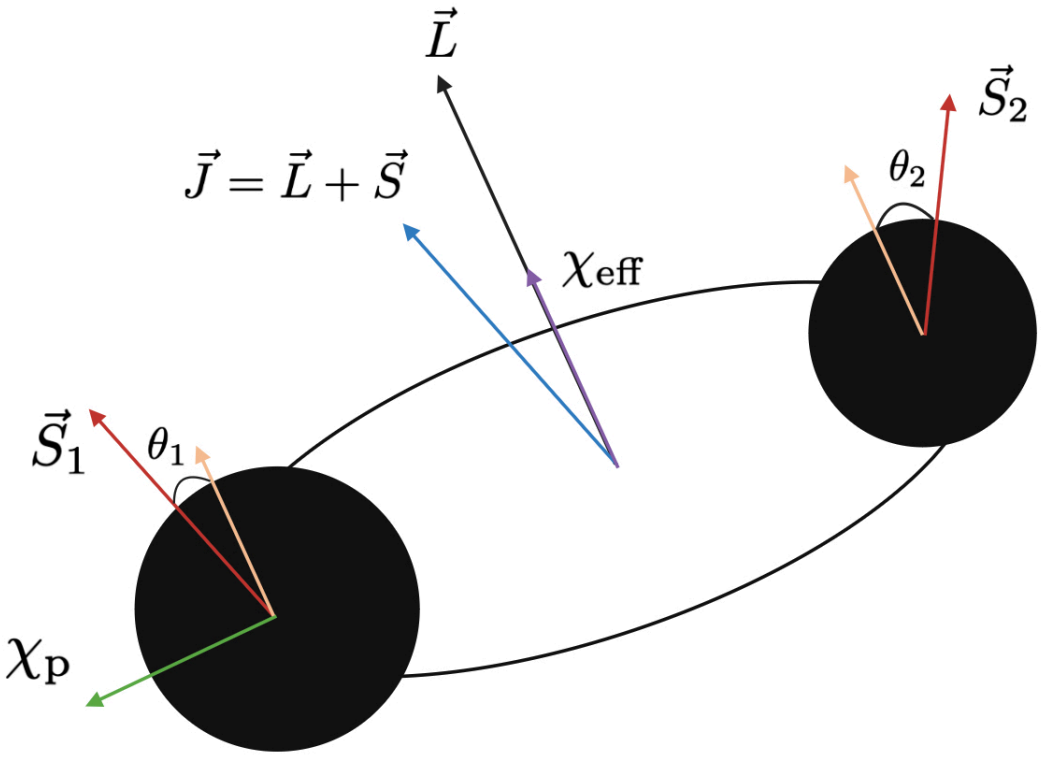

Figure 2.5: Scheme of a binary system, with spin components. The black vector is the orbital angular momentum $\vec{L}$, the blue vector is the total angular momentum $\vec{J}$, the red vectors are the spin vectors $\vec{S}_1$ and $\vec{S}_2$, the green arrow the precession spin parameter χ_p and the purple arrow is the effective spin parameter χ_{eff}. The two angles θ_1 and θ_2 are the tilt angles of the system, defined by the angle between the orbital angular momentum and the spin magnitudes of each object.

characterized, because both χ_{eff} and χ_p are combinations of the original spin parameters introduced above. Following Sec. 1.3, the two parameters are defined by:

$$\chi_{eff} = \frac{\chi_1 \cos\theta_1 + q\chi_2 \cos\theta_2}{1+q} = \frac{s_{1,z} + q s_{2,z}}{1+q}, \tag{2.56}$$

$$\chi_p = max\left[\chi_1 \sin\theta_1; \left(\frac{4q+3}{3q+4}\right) q\chi_2 \sin\theta_2\right]. \tag{2.57}$$

In Eq. 2.57, the letter q refers to the mass ratio of the binary and is taken equal to m_2/m_1, where m_1 is the primary mass assumed to be greater than m_2. From their expressions in Eq. 2.57, the effective spin parameter χ_{eff} is bounded between $[-1, 1]$ and the precession spin parameter between $[0, 1]$.

The origin of these two parameters comes from the approximate analytical methods used to describe mathematically the inspiral phase of a CBC merger, i.e. PN formal-

ism [120]. For the effective spin, it first appears at the 2.5 PN order in the equation of $\dot{L}$, which indicates that the phase of the GW waveform is influenced by the combined effect of the BHs spins, specifically their average spin weighted by the mass and projected onto the direction of the orbital angular momentum. These parameters are referred to as the effective and precession spins discussed above. For a random configuration of the spin parameters $(\chi_{1,2}, \theta_{1,2}, \phi_{1,2})$, the dominant signature of the spin on the waveform is hence associated with the effective spin and the precession spin. But since these two parameters are combinations of the individual spin parameters, they do not constraint the individual objects.

For instance, if we were to observe $|\chi_{eff}| << 1$, this scenario could arise from two distinct configurations of the spins: both black holes being minimally spinning ($\chi_1 << 1$ and $\chi_2 << 1$) but aligned or anti-aligned with the direction of the orbital angular momentum, or this could result from a binary system with significantly spinning black holes, but both spins being inside the orbital plane of the system. For values close to ± 1, the situation is different, it means that the black holes are highly spinning and aligned with $\vec{L}$ if $\chi_{eff} = 1$ or anti-aligned if $\chi_{eff} = -1$. The precession spin χ_p, on the contrary, tells how much spin is contained inside the orbital plane, causing the binary system to precess. Values close to 1 mean that the system is highly precessing and respectively slowly precessing for values reaching 0. Compared to the effective spin, the effect on the GW waveform of the precession spin is sub-dominant, since it appears at higher PN order [121].

2.6.2 The Default spin model

The Default spin model was originally used for population only studies in [119], but was proposed two years prior by [117, 118]. This model is based on the original spin parameters, namely the two dimensionless spin magnitudes $\chi_{1,2}$ and the two cosine of the tilt angle $cos\theta_{1,2}$. The two azimuthal angles are assumed flat, and are not considered in the Default spin model. In this model, the prior of the individual spin magnitudes of both black holes are given by:

$$\pi(\chi_1) = B(\chi_1|\alpha, \beta)\,, \tag{2.58}$$

$$\pi(\chi_2) = B(\chi_2|\alpha, \beta), \tag{2.59}$$

where B refers to the beta distribution [122]. The two merging objects are assumed to be described by the same two parameters α and β. There exists a relation linking the parameters of the beta distribution to the mean and standard deviation of $\pi(\chi_{1,2})$, this second parameterization of the Default spin model is expressed as:

$$\alpha = \left(\frac{1-\mu_\chi}{\sigma_\chi^2} - \frac{1}{\mu_\chi}\right)\mu_\chi^2 \geq 1\,, \tag{2.60}$$

$$\beta = \alpha\left(\frac{1}{\mu_\chi} - 1\right) \geq 1. \tag{2.61}$$

The condition $(\alpha, \beta) \geq 1$ is imposed by the model, this suppresses any non-singular asymptotic behavior around 0 or 1 for the beta distribution, behaviors that would be nonphysical. The choice of the beta distribution to model the spin magnitude is linked to its flexibility, the beta distribution can take multiple different shapes and is bounded between 0 and 1, with only two parameters.

On the other hand, the cosine of the tilt angles $cos\theta_{1,2}$ are drawn from a mixture model between a Gaussian centered around 1 and truncated on $[-1, 1]$, and a uniform distribution over $[-1, 1]$, following [31]:

$$\pi(\cos\theta_1|\zeta, \sigma_t) = \xi\mathcal{G}_{[-1,1]}(\cos\theta_1|1, \sigma_t) + \frac{1-\xi}{2}, \tag{2.62}$$

$$\pi(\cos\theta_2|\zeta, \sigma_t) = \xi\mathcal{G}_{[-1,1]}(\cos\theta_2|1, \sigma_t) + \frac{1-\xi}{2}. \tag{2.63}$$

In Eq. 2.63, $\mathcal{G}_{[-1,1]}$ is the truncated Gaussian. The two parameters governing the shape of the distribution for the tilt angles are σ_t and ξ. σ_t is the standard deviation of the truncated Gaussian and ξ is the mixing fraction between the Gaussian and the uniform distribution. The use of a mixing parameter for the tilt angle between the orbital angular momentum and the spin magnitude directions allows to monitor the proportion of binary system having isotropically oriented spins (if the uniform distribution is dominant) against aligned spins (if $cos\theta_{1,2}$ is centered around 1). This proportion can have a huge impact on our understanding of the formation channels of BBH systems, and this matter will be further explained in Sec. 4. From Eq. 2.59 and Eq. 2.63, the total population prior distribution of the Default spin model is written as:

$$\begin{aligned}\pi(\chi_1, \chi_2, \cos\theta_1, \cos\theta_2) = {} & Beta(\chi_1|\alpha, \beta)\pi(\cos\theta_1, \cos\theta_2|\xi, \sigma_t) \times \\ & Beta(\chi_2|\alpha, \beta).\end{aligned} \tag{2.64}$$

The overall distribution for the spin parameters of the Default spin model, defined in Eq. 2.64 is truly the spin population $p_{pop}(\chi|\Lambda)$ of the CBC merger rate, seen in Sec. 2.5. Fig. 2.6 presents a corner plot of the marginal distribution for each parameter drawn from a generic population using icarogw.

2.6.3 The Gaussian spin model

The second spin model implemented inside icarogw is the so-called Gaussian spin model. This parameterization aims to measure the joint distribution of the effective and precession spin parameters χ_{eff} and χ_p. This model was first introduced in [123], and then re-used by the LIGO-Virgo-KAGRA Collaboration as one of the main spin model for population study of CBC mergers. As mentioned in Sec. 2.6.1, the effective and precession are the best measured spin parameters. They result from combined spin effects, such as the overall alignment of each spins with the orbital angular momentum or the amount of spin contained in the orbital plan of the binary. The model parameterizes the joint distribution of (χ_{eff}, χ_p) with a two-dimensional multivariate

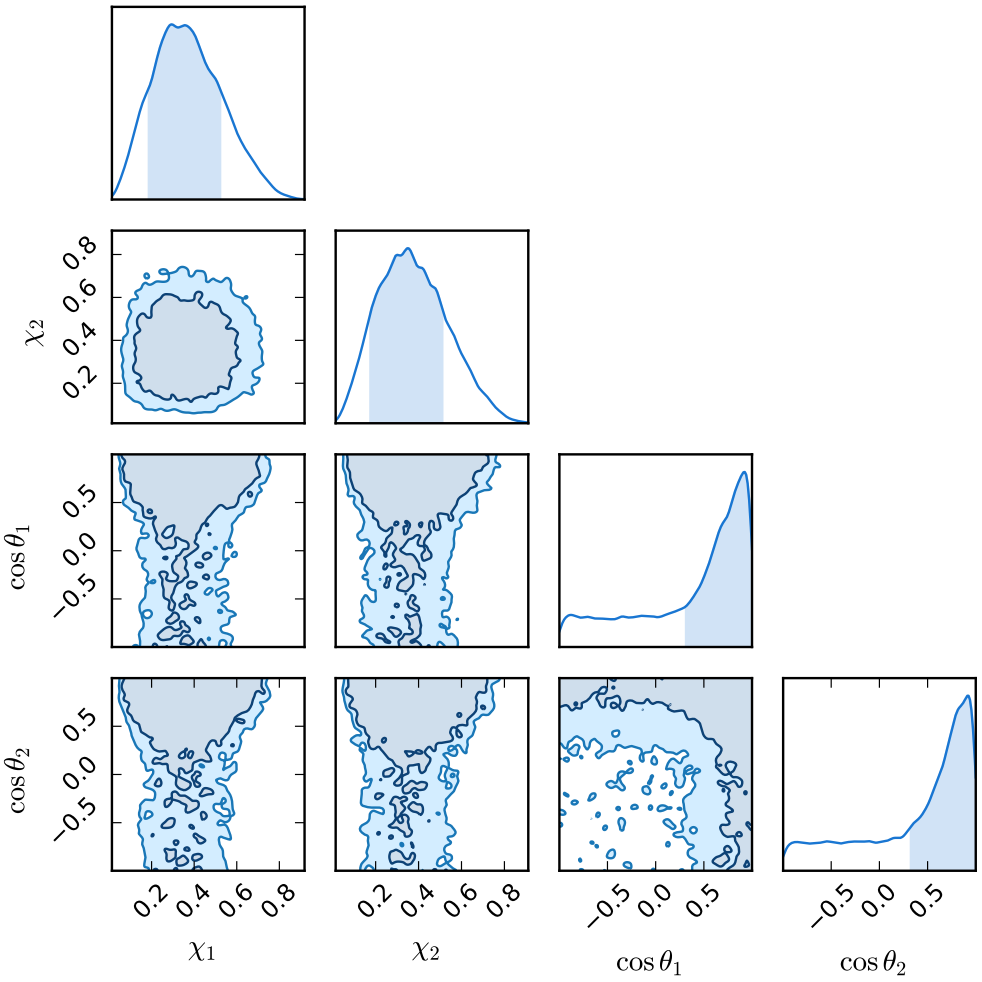

Figure 2.6: Corner plot of the joint and marginal distributions of $\chi_{1,2}$ and $\theta_{1,2}$, using the DEFAULT spin model of ICAROGW. These distributions have been drawn from the following set of parameters: $\alpha = 3$, $\beta = 5$, $\sigma_t = 0.3$ and $\xi = 0.5$. The colored contours correspond to the 90% confidence level (C.L.).

truncated Gaussian with five parameters: $\mu_{\chi_{eff}}, \sigma_{\chi_{eff}}, \mu_{\chi_p}, \sigma_{\chi_p}, \rho$. The covariance of that Gaussian is defined as:

$$cov_{[\chi_{eff},\chi_p]}(\sigma_{\chi_{eff}}, \sigma_{\chi_p}, \rho) = \rho\sigma_{\chi_{eff}}\sigma_{\chi_p}. \tag{2.65}$$

The multivariate Gaussian distribution, truncated on $[-1,1]$ for the χ_{eff} dimension and on $[0,1]$ for the χ_p dimension, can be written such as:

$$\pi(\chi_{eff}, \chi_p|\mu_{\chi_{eff}}, \sigma_{\chi_{eff}}, \mu_{\chi_p}, \sigma_{\chi_p}, \rho) = \mathcal{G}^{2D}_{[-1,1],[0,1]}(\chi_{eff}, \chi_p|\mu, \Sigma), \tag{2.66}$$

where μ and Σ are the mean and the covariance matrix of the joint distribution expressed as

$$\mu = (\mu_{\chi_{eff}}, \mu_{\chi_p}), \tag{2.67}$$

$$\Sigma = \begin{pmatrix} \sigma^2_{eff} & cov_{[\chi_{eff},\chi_p]} \\ cov_{[\chi_{eff},\chi_p]} & \sigma^2_p \end{pmatrix}. \tag{2.68}$$

Although, the implemented distribution inside ICAROGW is slightly different from Eq. 2.66, since the previous parameterization of the multidimensional distribution is computationally demanding. We then choose the following equivalent factorization for the multidimensional Gaussian of the GAUSSIAN spin model:

$$\begin{aligned} \pi(\chi_{eff}, \chi_p \mid & \mu_{\chi_{eff}}, \sigma_{\chi_{eff}}, \mu_{\chi_p}, \sigma_{\chi_p}, \rho) = \\ & \mathcal{G}_{[-1,1]}(\chi_{eff}|\mu_{\chi_{eff}}, \sigma_{\chi_{eff}}) \mathcal{G}_{[0,1]}(\chi_p|\mu_*, \sigma_*), \end{aligned} \tag{2.69}$$

where the two parameters μ_* and σ_* are defined as

$$\mu_* = \mu_{\chi_p} + \frac{cov_{[\chi_{eff},\chi_p]}}{\sigma^2_{\chi_{eff}}} (\chi_{eff} - \mu_{\chi_{eff}}), \tag{2.70}$$

$$\sigma_* = \frac{\sigma_{\chi_p} cov_{[\chi_{eff},\chi_p]}}{\sigma^2_{\chi_{eff}}}. \tag{2.71}$$

Fig. 2.7 shows the joint 2D distribution of the Gaussian spin model, for a generic set of parameters. The central plot depicts the joint distribution, while the two densities are the marginalized probability density functions for each parameter.

2.6.4 Spin for gravitational wave population and cosmology

The implementation of the DEFAULT and GAUSSIAN spin models within ICAROGW marks an advancement in GW-based cosmology analysis. This addition may enable an improved inference of the cosmological parameters of our Universe and the mass spectrum of BBHs. Both the Spectral Siren method and the galaxy catalog method can utilize these new population models, as seen in Sec. 2.5, enhancing the flexibility of the analysis. By integrating spin parameters into ICAROGW cosmological and population inference, we can now explore potential correlations between various astrophysical properties, such as the mass and spin of BBHs, cosmology and merger rate. Neglecting such correlations could lead to biased estimations of parameters like the Hubble constant H_0. Therefore, incorporating spin properties may be imperative for achieving more robust estimations of the cosmological parameters. Note that the use of spin population models for the analysis is not mandatory, one can perform a full population inference with only the mass and merger rate population models.

Furthermore, the inclusion of spin parameters in the inference framework could serve to constrain cosmology further, and enhance the population inference. Notably,

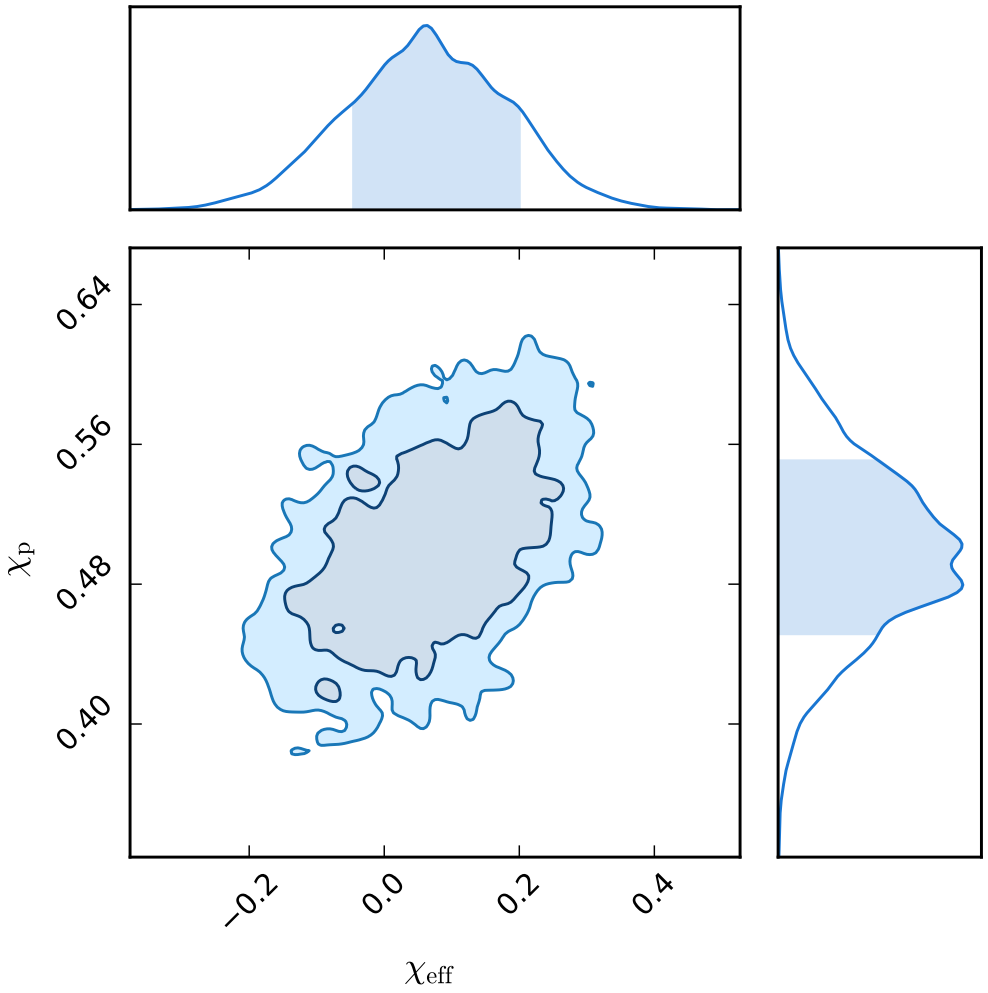

Figure 2.7: Corner plot of the joint and marginal distributions of χ_{eff} and χ_p, using the Gaussian spin model of icarogw. This distribution has been drawn from the following set of parameters: $\mu_{\chi_{eff}} = 0.06$, $\sigma_{\chi_{eff}} = 0.12$, $\mu_{\chi_p} = 0.5$,$\sigma_{\chi_p} = 0.05$ and $\rho = 0.5$. The colored contours correspond to the 90% confidence level (C.L.).

non-uniform distribution of BBH spins across the mass spectrum may help in estimating the mass model and thereby enhance constraints on the Hubble constant. Additionally, adding spin population models within the Bayesian inference of icarogw is crucial in understanding the formation channels of BBHs. Spins serve as significant signatures of the formation channels, influenced by factors such as isolated or dynamical formation([124, 125]). The latter is explored in Sec. 4.

In conclusion, with the incorporation of population spin models, icarogw becomes the first code capable of jointly inferring cosmological parameters and the spin properties of BBHs. Sec. 4 presents multiple analyses regarding spins, including extensions of the Default and Gaussian spin models for population purposes.

2.7 CONCLUSION

This chapter presents the hierarchical Bayesian framework of ICAROGW, which utilizes detected GW signals from CBC mergers in our Universe, in order to estimate the population properties of the sources, jointly with the cosmological parameters. Three methodologies are highlighted, the Spectral Siren analysis, based solely on GW signals, the galaxy catalog method that incorporates extra information from galaxy surveys in order to statistically estimates the redshift of the host galaxy of the source and finally the E.M.C method based on the joint detection of a GW signal and its electromagnetic counterpart.

Sec. 2.3 introduces the main ingredients of the Bayesian analysis, as the hierarchical likelihood constructed for a set of GW signals detected over a certain observation time, in presence of selection bias. This term shown in Eq. 2.7, incorporates another crucial quantity referred as the CBC merger rate, that encapsulate all the population models for the masses, the spins and the redshift distributions of CBCs. Sec. 2.3 also discussed how the computation of certain integrals in the Bayesian framework are handled numerically within ICAROGW, as well as some estimators needed to monitor the stability of the inference.

In Sec. 2.4, we explore the details of the three main methodologies of the GW-based methods for cosmology and population inference. The CBC merger rate of each of them is derived, as well as the underlying astrophysical assumptions necessary for the analysis. In this chapter, we propose an in-depth look at the parametric population models used in ICAROGW, Sec. 2.5 presents three models used to describe the primary and secondary masses of CBCs, as well as their astrophysical origin and motivations. In particular, we highlight the importance of certain population parameters, that will have a critical role in the following chapters.

Finally, Sec. 2.6 introduces the addition of phenomenological spin population models, the DEFAULT and GAUSSIAN spin models. It details the specifics of each model and how they are incorporated in ICAROGW's framework. This section also discusses the importance of including spin parameters to the method, alongside the new opportunities in terms of physics that are now achievable with the help of the spins.

Such method to estimate the Hubble constant with GW data is very new, and a lot of tests have to be done in order to assess its robustness. In particular, since the method is highly dependent on the choice of the models for describing the mass population of CBCs, it is of crucial importance to understand if such models can introduce systematic errors that could lead to a bias measurement of the Hubble constant. This leads directly to the next chapter of this manuscript, that will answer the following question: What are the main systematic related to population models that can influence the Hubble constant estimate with Spectral Siren, and in particular systematics related to the mass spectrum parameterization ?

3 STUDY OF SYSTEMATICS ON THE COSMOLOGICAL INFERENCE OF THE HUBBLE CONSTANT FROM GW STANDARD SIRENS

The majority of this chapter is based on the work of a published paper, in which I am the first author [94]. This paper presents a detailed analysis on the systematics errors that can impact the estimation of the Hubble constant when using GWs Spectral sirens methods. This chapter follows the same overview of the paper, from the methods we employ to do the analysis, to the sanity checks for the framework validation and finally the results. The subsections on the hierarchical analysis and phenomenological mass models have been removed because already discussed in Sec. 2. Moreover, the Sec. 3.2.2 and Sec. 3.2.3 have been added, these two sections are based on another project of mine. Most of the context of the paper has be re-written in order to better fit the overall manuscript.

3.1 INTRODUCTION

Since 2015, more than 90 confident detections of GW signals have been made by the LVK collaboration, over three observing runs (O1, O2 and O3) [25–27]. These de-

tected GW signals, as explained in Sec. 1.5, enable us to directly measure the luminosity distance d_L of the sources without making any assumption on the cosmological model, nor using the distance ladder. However, unless an electromagnetic counterpart (EMC) emitted by the same GW source is observed, the GW signals do not directly provide an estimate of the redshift of the source. The combined measurement of the redshift and the luminosity distance are necessary in order to estimate the Universe's expansion rate H_0. Because of the current status on the measurement of the Hubble constant, also known as the Hubble tension, GWs independent methods offer an interesting prospect for a new approach to assess the value of the local expansion rate of the Universe [126]. With the increasing number of GW detections by the LVK collaboration, a new era called "GW cosmology" has begun, and a totally new field of physics is unveiled.

To exploit the tens of GW sources observed without an EMC, complex statistical methods have been developed in order to get an estimate of the source redshift. In this chapter, one particular method is discussed, this method called "*Spectral Siren*" [80] is based solely on GW data, a more detailed description can be found in Sec. 1.5, Sec. 2.3 and Sec. 2.4. The Spectral siren method relies on the intrinsic degeneracy between the redshift z of a GW source and its detector frame mass m^{det} and source frame mass m^s, through $m^{det} = (1+z)m^s$.

The redshift of the sources is implicitly estimated by measuring the detector masses and jointly fitting the source mass distribution, CBC merger rate and cosmological model [77, 79, 127]. The Spectral siren method deeply relies on the choice of the phenomenological population models for the BBH source mass distribution and their capacity to describe the true underlying BBH population [78, 80, 81, 95, 103, 127–132]. The galaxy catalog method, described in Sec. 2.4.2, that incorporates extra information on the redshift of the source via the help of galaxy surveys data, also relies on the population models assumptions. This is true when the galaxy catalog is not complete enough to statistically estimate the redshift, and in this scenario, the Spectral siren method take over [70, 78, 81, 103]. In this chapter, we explore how the phenomenological population models, in particular source mass models, are able to fit complex and astrophysically motivated BBH populations in the context of the Spectral siren analysis. The goal is to understand how an inaccurate reconstruction of the source mass distribution could translate in a biased estimation of the Hubble constant, and to assess how robust or subject to systematic errors GW based cosmology methods are when facing modelled effects in the source mass distribution of BBHs. To make this study, we start with simulated GW data in order to validate our framework and then move on a BBH population generated synthetically through the modelling of crucial astrophysical processes responsible for the BBH population properties. We then perform the complete hierarchical Bayesian inference scheme developed in Sec. 2 to measure the Hubble parameter H_0 jointly with the population distribution of the source masses and CBC merger rate.

This chapter is organized as follows, in Sec. 3.2 we describe in details how the GW detections are simulated and the specific framework developed to do so. In Sec. 3.3,

we present an extensive set of sanity checks to test the validity of the inference and simulation of the GW events. At the same time, we investigate the robustness of the Spectral sirens' method when a wrong mass model is used for the inference and when the source mass spectrum incorporates modelled mass feature or even a redshift evolution. In Sec. 3.4, we perform the population and cosmological inference using a realistic catalog of BBH mergers. In this section we also investigate and discuss the potential sources of the biases, and peculiarly the effect of redshift evolution of the mass spectrum when inferring the cosmological parameters. Conclusions are drawn in Sec. 3.5.

3.2 SIMULATION OF GRAVITATIONAL WAVE OBSERVATIONS

This section introduces our simulation procedure and framework used to generate GW observations from BBH mergers. The detection criteria for a GW signal are based on the signal-to-noise ration (SNR) and the maximum frequency detectable by the current interferometers.

3.2.1 A fast generator of GW observations

Simulating realistic GW detection data is a complex task, typically involving the generation of authentic GW waveforms and their injection into noise detector realizations. Specialized algorithms, known as "search pipelines", are then employed to identify potential detections of these waveforms [133–135]. Subsequently, a crucial step called parameter estimation (PE) becomes necessary, aiming to estimate the population parameters of the GW sources from the extracted waveforms amidst detector noise [136]. However, for certain analyses, such as the investigation of systematic errors in Spectral Sirens analysis, utilizing the complete process of waveform generation, detection, and parameter estimation proves itself impractical and time-consuming. To address this, we present a Python framework capable of rapidly generating GW observations. These observations differ from real simulated GW events, as they do not involve GW waveforms, search pipelines, or parameter estimation to determine the population parameters of the GW source. In the following sections of this chapter, we will call a GW observation, a set of three parameters describing a detected GW source, namely the two detected detector frame masses (m_1^{det}, m_2^{det}) and the luminosity distance d_L. The workflow to efficiently generate these GW observations is the following: From a population of BBH mergers, described by the two source frame masses and the redshift of the merger, we compute an approximate value of the optimal SNR. Instead of computing the SNR of each event with the usual approach using a search pipeline on the full GW waveform and performing a match filtering search [137]. This proxy provides a realistic approximation that accounts for the primary sources of noise affecting the SNR estimation in real GW events. It incorporates factors such as signal recovery, multi-detector detections, parameter estimation uncertainties, and general

selection effects. However, this approximation could be further refined by including second-order effects that may influence the SNR. For instance, the influence of black hole spins in BBH systems could be introduced as a second-order correction to the SNR. We use Eq. 3.1 as the proxy for the optimal SNR [102, 127]

$$\rho = \delta \times \rho^* \left[\frac{\mathcal{M}_c}{\mathcal{M}_c^*}\right]^{\frac{5}{6}} \left[\frac{d_L^*}{d_L}\right], \tag{3.1}$$

where $\mathcal{M}_c$ is the binary detector chirp mass, computed from a combination of the two detector frame masses as

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}(1+z). \tag{3.2}$$

In Eq. 3.1, δ is a projection factor and ρ is the optimal SNR, i.e. the SNR obtained if the binary system was optimally oriented with respect to the detectors. The proxy in Eq. 3.1 is known to produce a good estimate of the optimal SNR ρ, with a slight underestimation of GW events for chirp masses above $30M_\odot$, but it is enough to generate a good estimation of detected GW events. The luminosity distance and the detector chirp mass are computed assuming a flat ΛCDM cosmological parameterization with the cosmological parameters fixed to the Planck 15 values measured from the CMB, namely $H_0 = 67.8\ km\,s^{-1}\,Mpc^{-1}$ and $\Omega_{m,0} = 0.308$ [138]. The detection sensitivity of the interferometer network is modelled by the star parameters with the superscript $*$, i.e. a GW signal with a projection factor $\delta = 1$, localized at a luminosity distance d_L^*, with a chirp mass of $\mathcal{M}_c^*$ would produce an optimal SNR of ρ^*. To mimic an $O4$ like sensitivity for a network made of three terrestrial interferometers, we set the star parameters to be $d_L^* = 1.5\ Gpc$, $\mathcal{M}_c^* = 25\ M_\odot$ and $\rho^* = 9$ [7]. The projection factor δ in Eq. 3.1 encodes multiple detections effects that can make the optimal SNR vary, it combines the effects of the GW polarization and the sky localization obtained with a three detector network. For one GW observation, its value is drawn from a cumulative density function introduced in [139].

Furthermore, to mimic the presence of noise in the detector, the optimal SNR is corrected into a "detected" SNR ρ^{det}. This detected SNR is drawn from a χ^2 distribution with six degrees of freedom, two for each detector contained in the network. The new SNR is then defined such as

$$\rho^{det} \propto \mathcal{L}_{\chi^2}(\rho, n = 6), \tag{3.3}$$

where n in Eq. 3.3 is the number of the degrees of freedom of the χ^2 distribution. In our simulation framework, a GW event is said to be detected if its associated ρ^{det} is greater than a detection threshold of 12. To further push the detection simulation process close to a realistic one, a frequency cut is also applied to each GW event. This cut is placed on the maximum GW frequency of each GW, defined as twice the frequency corresponding to the innermost stable circular orbit (ISCO)

$$f_{GW}^{max} = 2f_{ISCO}, \tag{3.4}$$

where f_{ISCO} is approximated following [140]

$$f_{ISCO} = \frac{1}{2\pi}\frac{1}{6^{3/2}}\frac{c^3}{G}\left(\frac{1M_\odot}{M^d_{tot}}\right)10^3 \quad Hz. \tag{3.5}$$

In Eq. 3.5, G is the gravitational constant, c is the velocity of light in vacuum and M^d_{tot} is the total mass of the binary in detector frame, computed as the sum of the primary and secondary mass. The threshold for a valid detection is chosen to be $f^{max}_{GW} \geq 15 \quad Hz$, where every GW event having a maximum GW frequency below this selection criteria are flagged as "non-detected", even if their detected SNR is above twelve. Adding a cut on the GW frequency has two beneficial effects, the first one being that real GW detectors have a frequency band, generally around $[10, 1000]Hz$ in which they are sensitive to GW signals, and outside which even massive and close by events can not be seen [10, 141, 142]. This cut eliminates non-physical detections that could negatively impact our analysis, specifically removing GW events from extremely distant sources with masses on the order of thousands of solar masses. Such outlier events should not be detectable with the O4-like sensitivity, even if their SNR exceeds the selected threshold. This is because the frequencies of these mergers fall outside the detectable range of the current GW detector network.

To summarize, the framework described above is able to generate the so-called GW observations. From a catalog of triplets containing the source frame masses and redshift of simulated GW events, it simulates a detection process for a chosen sensitivity and interferometer network, assigning to each of them a detected SNR based on the chirp mass and the luminosity distance of the source, leaving us with a subset of the entire catalog containing only the detected GW events, described by $(m^{det}_1, m^{det}_2, d_L)$. The original GW catalog in source frame, containing all GW events, detectable and non-detectable is generally generated using the population mass models and CBC merger rate models implemented inside ICAROGW and described in Sec. 2. This framework is very efficient and is able to generate GW observations quickly, on the other hand no errors are simulated for the detector frame masses and the luminosity distance of each source, and that is the reason why we call them GW observations and not GW detections data. Although, it is possible to simulate errors in order to have PE-like data for each GW detected merger, and this will be explained in the following Sec. 3.2.2. Finally, we choose not to generate errors on the detector frame masses and luminosity distance. In other word, we assume that they are perfectly detected by the network of interferometers from the GW signal. We make this choice for two reasons. The first one is that not generating errors greatly ease the computational load of the analysis, and allow us to test a large variety of test cases with thousands of GW detections. The second is that we want to understand the systematics of the Spectral siren analysis, and not generating the errors will maximize the potential effects of the systematics arising from the phenomenological models, systematics that could be hidden in the error budget of the masses. Although we opted not to introduce errors in the following analysis for the reasons listed above, the following subsections will explain how to generate errors on the detector frame masses and luminosity distance for this type

of simulation. Additionally, a simple mock data challenge is presented to test the simulation.

3.2.2 Simulation of the errors for the GW observations

In this section, we detailed the process to generate the errors on the masses and luminosity distance from the GW observations. The idea is to mimic the parameter estimation usually done on real GW events in order to estimate the measured values of the population parameters from the detected signal. This step is called the simulation of the posterior samples, and is done for each of the detected GW observations produced with the framework of Sec. 3.2.1. This procedure has already been applied in several analyses to quickly generate GW like posterior samples [95, 102, 127]. From the GW observations of each event, the posteriors samples for the measured chirp mass in detector frame $\mathcal{M}_c^{det}$ as well as for the measured mass ratio q_{det} are drawn from two Gaussian distributions such that

$$\mathcal{M}_c^{det} \propto \mathcal{N}\Big(\mathcal{M}_c, 10^{-3}\mathcal{M}_c\frac{10}{\rho^{det}}\Big), \tag{3.6}$$

and

$$q^{det} \propto \mathcal{N}\Big(q, 0.25q\frac{10}{\rho^{det}}\Big). \tag{3.7}$$

The specific parameterization of the mean and the standard deviations of the Gaussian in Eq. 3.6 and Eq. 3.7 are chosen to approximate the typical errors the of a full parameter estimation on a real GW signal [127]. So from the detector frame masses of detected signals, posterior samples on the detector frame chirp mass and the mass ratio are simulated from Eq. 3.6 and Eq. 3.7. In general, the posterior samples needed for a GW cosmology analysis in ICAROGW are given in terms of primary and secondary detector frame masses, hence the chirp mass and mass ratio a converted backward using the following equations:

$$m_1^{det} = \mathcal{M}_c\frac{(1+q)^{1/5}}{q^{3/5}} \tag{3.8}$$

and

$$m_2^{det} = qm_1^{det}. \tag{3.9}$$

This specific way to generate the posterior samples on the masses, drawing first the chirp mass and mass-ratio and then (m_1^{det}, m_2^{det}), permits accounting for the degeneracy during a real PE process between the two masses. Similarly, as for the measured chirp mass and measured mass ratio, posterior samples of the measured projection factor δ^{det} are drawn from the following normal distribution

$$\delta^{det} \propto \mathcal{N}(\delta, 0.3\frac{10}{\rho^{det}}). \tag{3.10}$$

Following this methodology, the posterior samples on the luminosity distance are obtained by inverting the Eq. 3.1 for d_L and injected in the already produced posterior samples on the chirp mass and the projection factor. This approach to generate posterior samples on the primary and secondary masses as well as the luminosity distance enables the production of samples that are conformed to the selection effects accounted for in the typical Spectral sirens analysis. As a result, characteristic uncertainties for the luminosity distance and masses range from 40% to 60% and 20% to 50%, respectively, at the 90% confidence level.

3.2.3 Illustration of a mock catalog of binary black hole mergers

We make use of the simulation framework presented in Sec 3.2.1 and Sec 3.2.2 to generate a mock catalog of detected BBHs. We start from a set of source frame masses and redshifts of potential detectable sources, to finish with the full cosmological and population inference on the simulated GW events. Here, in addition to the detection process via the SNR proxy and the frequency cutoff, the errors budget (posterior samples) for the masses and the luminosity distance are also generated.

First, we choose a flat ΛCDM cosmological model with $H_0 = 67.7\,\mathrm{km\,s^{-1}Mpc^{-1}}$ and $\Omega_{\mathrm{m},0} = 0.307$. The primary and secondary source frame masses of the entire simulated BBH population are generated using the PLP mass model (see Sec 2.5.2), and the values of the parameters used to construct the PLP are listed in Table 3.1. The redshifts assigned to each merger are drawn from the so-called Madau & Dickinson-like CBC merger rate model, with $\kappa = 3$, $\gamma = 3$, and $z_p = 2$ [143]. This parameterization of the CBC merger rate is the standard (or "vanilla") model in GW analysis when describing the merger rate in the Universe, assuming that it follows the star formation

Parameter name	Description	Value
α	Index of the primary power law.	3.5
β	Index of the secondary power law.	1
δ_{m}	Smoothing parameter.	6
μ_{g}	Mean of the Gaussian peak.	$31\,M_\odot$
σ_{g}	Standard deviation of the Gaussian peak.	$4\,M_\odot$
λ_{peak}	Fraction of events in the Gaussian peak.	0.2
$\mathrm{m}_{\mathrm{min}}$	Minimum mass.	$5\,M_\odot$
$\mathrm{m}_{\mathrm{max}}$	Maximum mass.	$130\,M_\odot$

Table 3.1: Table listing the population parameters governing the power law plus peak mass model, used to generate the primary and secondary masses of the entire BBH population for the mock catalog. The name, description, and chosen value for each parameter are shown in the table.

rate (SFR) model of Madau & Dickinson. Using the source frame mass model and the CBC merger rate model, we draw 200000 GW events, each composed of the triplet (m_1^s, m_2^s, z). The number of simulated GW events is not intended to be astrophysically motivated and does not correspond to a physical CBC rate. Instead, it is chosen to ensure a reasonable number of detections for illustrating the simulation of errors. Of the 200000 simulated GW events, only 102 were flagged as "detected" by our framework, which assumes O4-like sensitivity and a three-detector GW network. This detection process uses an optimal SNR threshold of 12 and a frequency cutoff of 15 Hz. The set of 102 detected GW events correspond to the so-called GW observations. From the GW observations, we simulate the error budget for the masses and the luminosity distance of each event (see Sec 3.2.2), and an example of the posterior samples simulated is given in Fig. 3.1. From Fig. 3.1, we can see that the true values of the masses

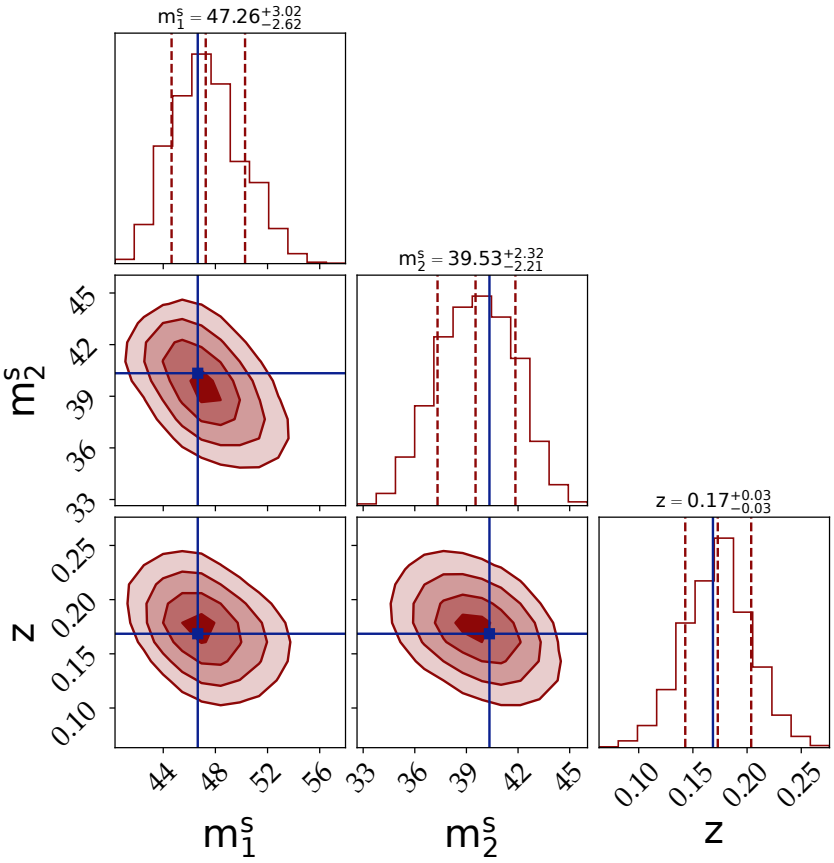

Figure 3.1: In source frame: Corner plot of the primary mass, the secondary mass, and the redshift for one of the detected GW observations for which the error budget has been simulated. The blue points and lines represent the true values of the masses and redshift from which the error budget has been generated. The red histograms are the simulated posterior samples, and the red contours are the 2D plot highlighting potential correlations between parameters. The red dotted lines represent the 1σ errors.

and the redshift are well within the simulated error budget. Moreover, the posterior distributions, i.e., the histograms of the 4000 samples generated for the primary mass,

secondary mass, and redshift, exhibit a close-to-Gaussian behavior. This feature is crucial as spiky posterior distributions would not be physically correct.

Using the BBH catalog composed of 100 simulated GW events, we validate this framework through a full population and cosmological inference with ICAROGW Spectral sirens analysis. We perform the inference using the same mass model and the same CBC merger rate from which the GW events have been simulated, placing us in the best case scenario where the correct population is known. The matter density parameter is not inferred here, since this kind of analysis struggles to put constraints on it [4]. The true value of all mass and CBC merger rate population parameters are recovered by the Spectral siren inference within the 90% C.L., in particular the Hubble constant is found at $H_0 = 64^{+15.0}_{-14.9} km\, s^{-1}\, Mpc^{-1}$. Fig. 3.2 shows the corner plot of some population parameters of interest for this inference. From Fig. 3.2, we demonstrate that

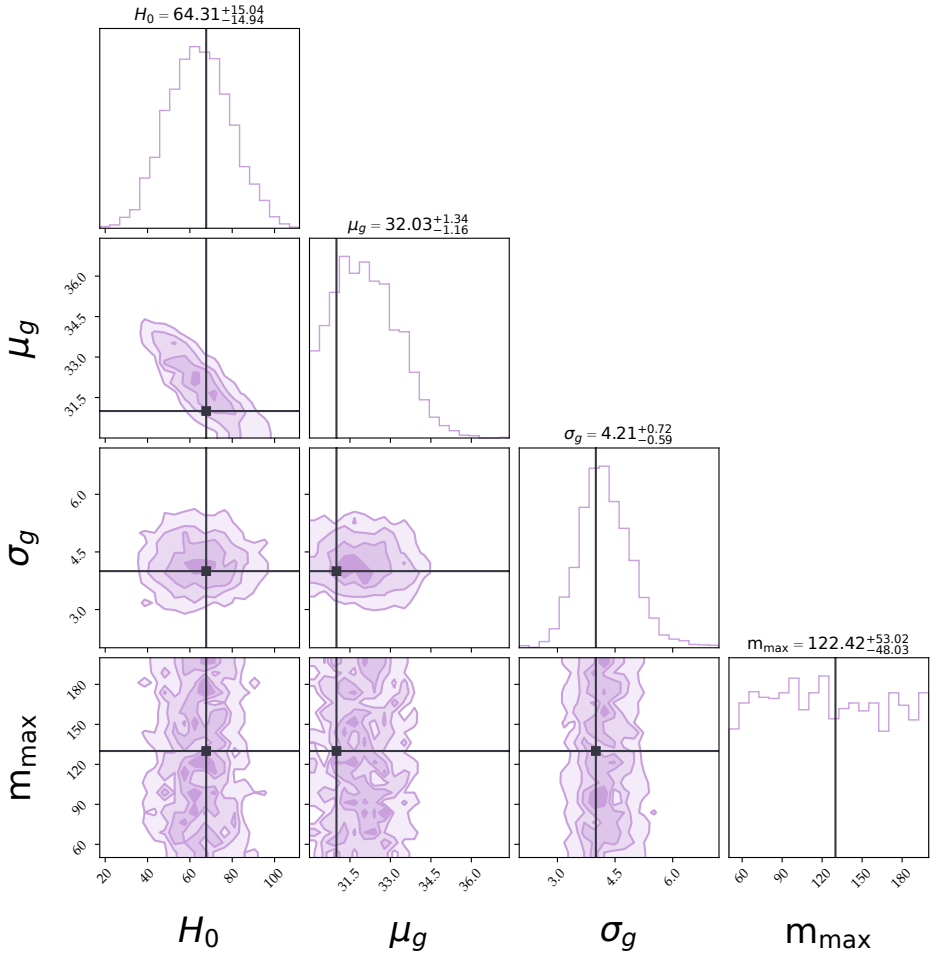

Figure 3.2: Corner plot of the inferred population and cosmological parameters estimated from the 100 mock GW data generated with our framework. From left to right, the Hubble constant, the mean of the Gaussian peak, the standard deviation of the Gaussian peak and the maximum mass. The true values used to simulate the GW data are shown with the black dots.

the framework developed to simulate GW data as well as the Spectral siren method to

infer population and cosmological parameters is validated. Moreover, since that kind of analysis is actually inferring the parameters on a hyper-posterior, the corner plot is a very nice tool to study correlations between parameters. For instance, the position of the Gaussian peak in the mass distribution is strongly correlated to the Hubble constant. This feature will play a key role in the following section.

3.3 APPLICATION TO ANALYTICAL BINARY BLACK HOLE POPULATION

This section considers several scenarios and populations of BBHs generated from the phenomenological models presented in Sec. 2.6, to study how the PLP, the BPL and the MLTP mass models are able to infer H_0. The spin parameters of the BBHs will not be considered in this section, and we will only work with GW observations (i.e. no error are simulated). In Sec. 3.3.1, the generation of the GW observations as well as the population inference are performed using the same mass models. In Sec. 3.3.2, the GW observations are again generated from one of the three mass models, but the inference using the Spectral siren analysis is done using the other two. And in Sec. 3.3.3, we simulate the GW events with a more complex version of the PLP mass model that incorporates a redshift evolution of its Gaussian peak. This allows us to test the response of the Spectral siren analysis to mass distributions evolving with the redshift, in the context where the mass model used for the inference does not account for such evolution. The main purpose of it being the study of the H_0 bias. In the remaining sections of this chapter, the H_0 bias will be quantified according to what confidence interval (C.I.) the true value of H_0 is found from the inferred posterior. Generally, biases are quantified using the so-called parameter-parameter (PP) plots that indicate what is the fraction of time that the true value of H_0 is found in a given C.I. for the posterior. For reference, if H_0 is inferred without any bias, it would be found in 50% of the C.I. for 50% of the case and so on, which would result in a diagonal curve on the PP plot. In this chapter, we will consider that an analysis biased following two key points, first if the true value of the Hubble constant is excluded at 2σ or more from the posterior of the estimated H_0. Second from the resulting PP plots generated for each analysis.

In order to better visualize the result from the Spectral siren population inference, we will use the posterior predictive checks (PPCs) to assess how well the PLP, BPL and MLTP models are able to fit the mass spectrum. The reconstructed distributions can be understood in source frame (astrophysical posterior predictive distribution) or in detector frame with the selection effects (detected posterior predictive distribution). The astrophysical posterior predictive distribution, given the redshift and the source frame masses is given by

$$p_{pop}(z, \mathrm{m}_1^s, \mathrm{m}_2^s|\{x\}) = \int d\Lambda \, p_{pop}(z, \mathrm{m}_1^s, \mathrm{m}_2^s|\Lambda) p(\Lambda|\{x\}), \tag{3.11}$$

and the detected posterior predictive distributions are defined as

$$p_{obs}(d_L, \mathrm{m}_1^d, \mathrm{m}_2^d|\{x\}) = \int d\Lambda\, p_{pop}(d_L, \mathrm{m}_1^d, \mathrm{m}_2^d|\Lambda) \times p(\Lambda|\{x\}) P_{det}(d_L, \mathrm{m}_1^d, \mathrm{m}_2^d). \tag{3.12}$$

3.3.1 Using the same mass model

We simulate three populations of BBH mergers, with the three source mass models described above and the Madau&Dickinson CBC merger rate for all of them. The precise values of the parameters used to draw these populations are summarized in App. B.1, in tables B.1 to B.3. From the three BBH populations, we use the GW observation framework to simulated 2000 detected GW events that will be used for the Spectral siren analysis. Typically, around $\mathcal{O} \sim 10^3$ GW detections are needed so that the posterior of the Hubble constant is truly data-informed within the chosen prior range $[20 - 140] km\, s^{-1}\, Mpc^{-1}$. Finally, for these three sets of detected GW observations, a full hierarchical Bayesian inference is done on the mass, CBC merger rate and cosmological parameter H_0. Fig. 3.3 presents the marginalized posterior distributions of H_0 estimated with the Spectral siren analysis using each of the three source mass models, and for all analyses, the true value of H_0 is recovered within the 90% C.I. Alongside with the Hubble constant, we also found that all other population parameters inferred contain their true value within the 90% C.I. From Fig. 3.3, it is shown that the posterior of H_0 tends to be more constrained when the PLP and the MLTP models are used. This effect is due to the fact that the PLP and MLTP models are built with sharper mass features in their source frame spectrum. In opposition, the BPL model gives a smoother posterior of H_0. Fig. 3.4 displays the PPCs for the three inferences over the three simulated BBH populations. This result shows that all the structures in the primary mass, secondary mass and CBC merger rate spectrum are perfectly reconstructed through the Spectral siren analysis. The sharp features of the mass spectra, produced by the Gaussian peak and also the maximum and minimum masses, are correctly inferred, both in the astrophysical frame and the detected frame. In summary, from this first set of tests, when using the correct mass model for the population inference, both the true cosmological and true population parameters are found unbiased. Moreover, this test illustrates the importance of mass features to constraint the Hubble constant.

Results consistent with Figure 3.3 have been replicated across approximately 20 independent population realizations. However, while these results provide initial validation, a more stringent assessment of bias in parameter inference requires the use of Probability-Probability (PP) plots, as introduced in Sec 3.3. Generating accurate PP plots presents challenges. Ideally, population models should be drawn from the hyper-priors utilized in the inference process. This approach ensures that the population realizations covers the full range of parameter space. Unfortunately, this is not always feasible, especially when dealing with pre-simulated catalogs. In cases where drawing from hyper-priors is restricted, such as with parameters intrinsic to the observed catalog, alternative strategies must be employed. While achieving diagonal PP

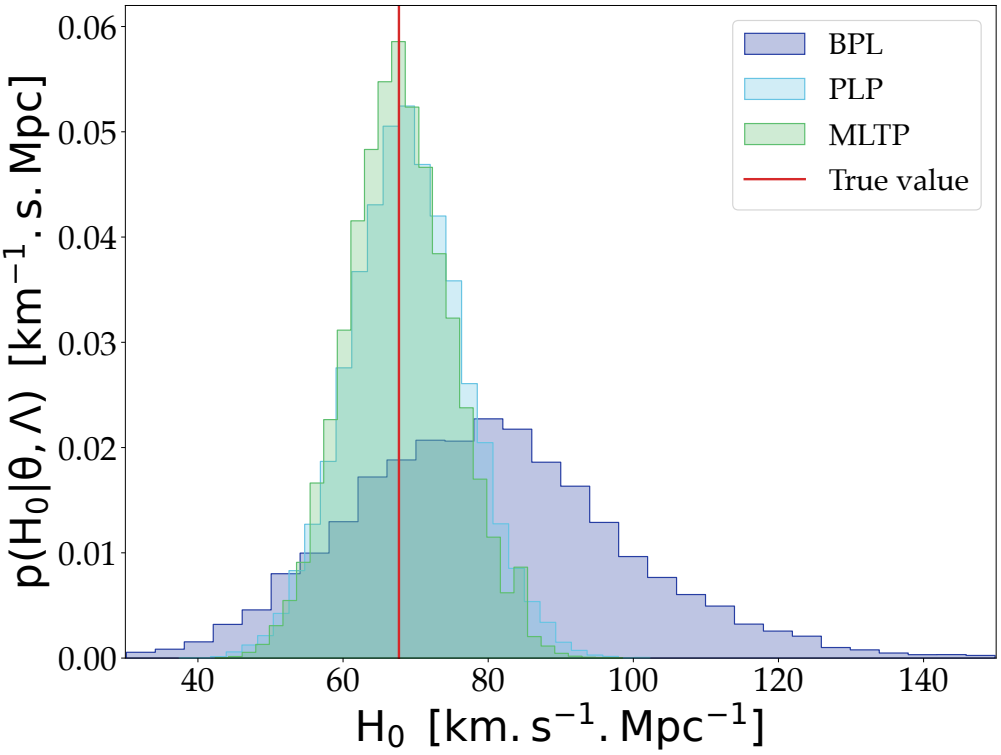

Figure 3.3: Marginalized posterior of the Hubble constant obtained from the Spectral siren analysis of 2000 detected GW observations using the PLP, BPL and MLTP mass models. The red line is the true value injected in the simulation, fixed to the Planck15 of $H_0 = 67.7 km\,s^{-1}\,Mpc^{-1}$.

plots without drawing from hyper-priors is possible, it necessitates stringent conditions. Specifically, all inferred posteriors must be gaussianized and entirely contained within their prior ranges, which may not always be achievable. Fig. 3.5 depicts the PP plot for the Spectral sirens inferences of 2000 GW observations simulated with the MLTP mass model and estimated with the same model (identical scenario as the green histogram in Fig. 3.3). From the several realizations with the same analysis set up, it is clear that the inference of the Hubble constant stays unbiased. The remaining two PP plots can be found in App. B.2.1.

3.3.2 Using different mass models

As an extension of the tests shown in Sec. 3.3.1, we generate three BBH catalogs with 2000 GW observations using the BPL, PLP and MLTP mass models. With these three populations of CBCs, we infer the population and cosmological parameter via the Spectral siren analysis, with the two other mass models that were not used for the simulation of each population. Fig. 3.6 presents all six marginalized posterior distri-

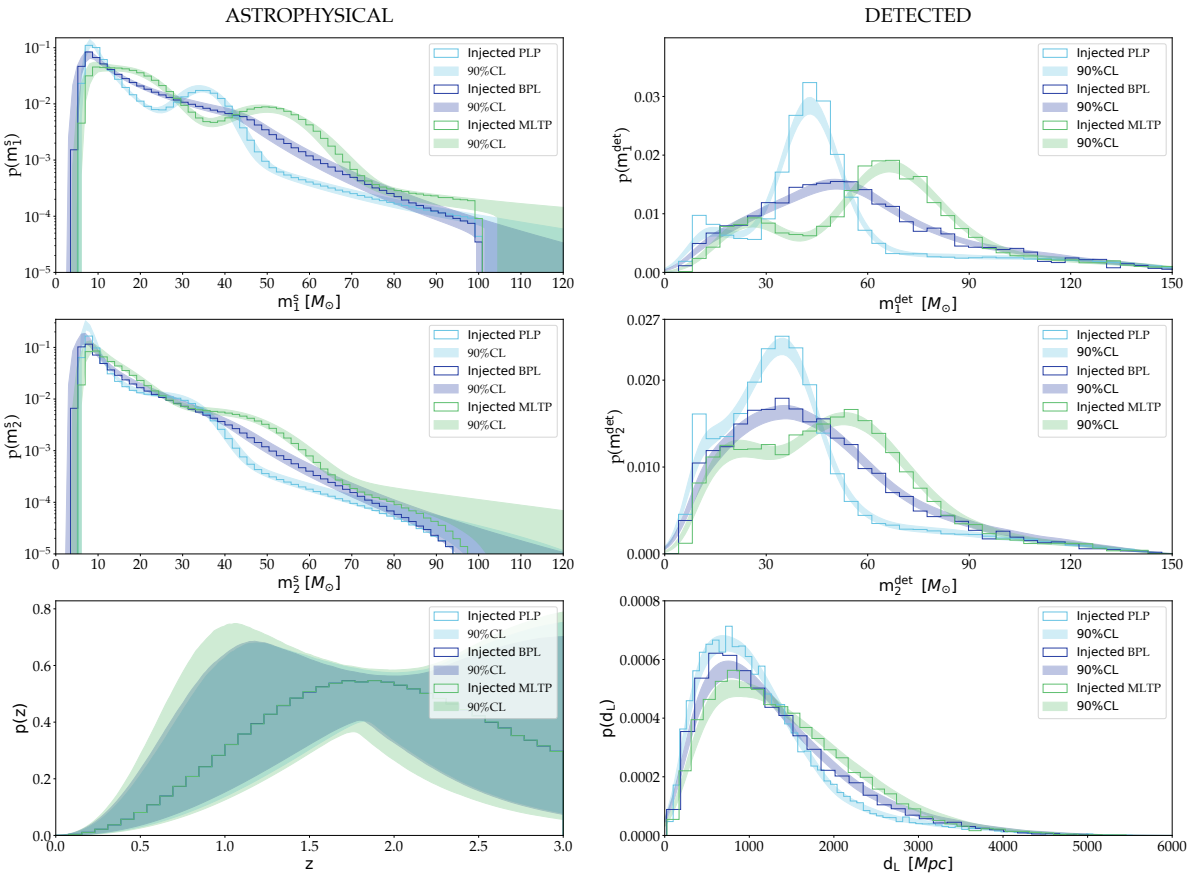

Figure 3.4: Posterior predictive checks obtained from the Spectral siren analysis of three BBH populations with 2000 detected GW events simulated with the PLP (in light blue), BPL (in blue) and MLTP (in green) mass models. The left column is the source frame population, from top to bottom m_1^s,m_2^s and z. The right column is the detector frame population, from top to bottom m_1^{det},m_2^{det} and d_L. The inferred population from the analysis is shown with the colored contours, as the 90% confidence level spectrum. And the injected populations are the colored histograms.

butions of the Hubble constant obtained from this analysis. We observe that when the GW observations are simulated using the BPL model, and the estimation of the population and cosmological parameters is done with the PLP or the MLTP models, the injected H_0 value is included in the 68% C.I. level despite having the incorrect source mass model in the Spectral siren analysis. Whereas when the GW events are generated with both the MLTP and the PLP models, and inferred with the BPL model, we observe that the true value of H_0 is excluded at 98.7% C.I. level. We interpret this bias as being the product of wrong reconstruction of the mass spectrum by the BPL model, the sharp mass features contained in the PLP and the MLTP models are missed by the BPL model. The mismatching of the source mass features results in a systematic mismatch of the redshifts of the GW sources which is inducing the bias on the Hubble constant.

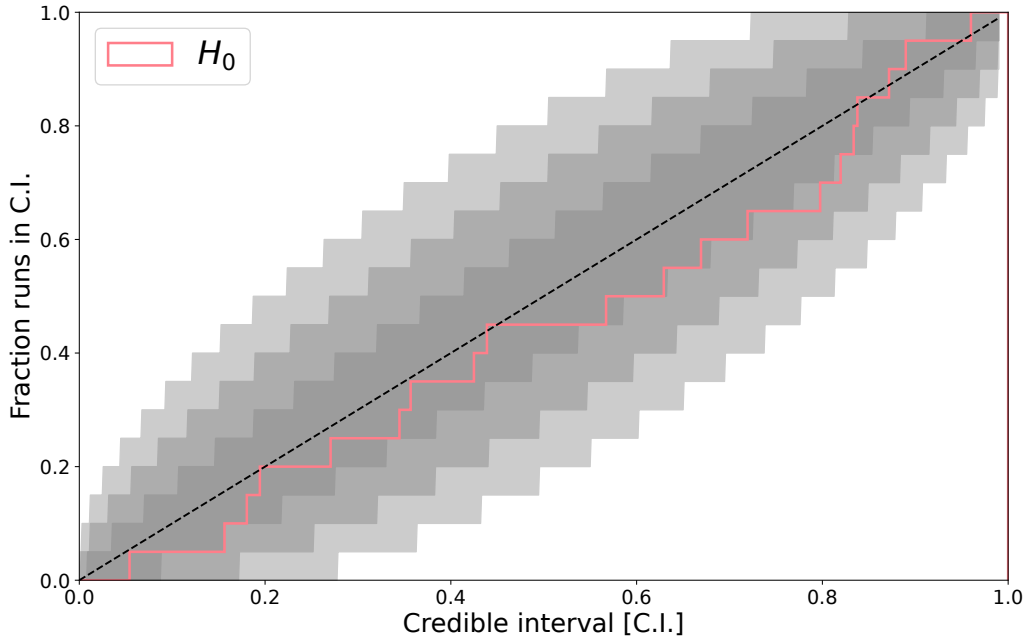

Figure 3.5: PP plot obtained from 20 independent population realization simulated with the MLTP mass model, and inferred with the MLTP model too. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

As in Sec. 3.3.1, the detailed understanding of the cosmological results has to be interpreted from the PPC plot. Fig. 3.7 shows the PPC obtained from the Spectral analysis of a BBH population simulated using the MLTP mass model and inferred with the PLP and BPL models. Based on the reconstructed source frame spectrum for m_1^s and m_2^s, it is clear that with the PLP model, the inference is able to reconstruct a peak located around $50M_\odot$. While with the BPL model, the same mass feature, important for the cosmology inference, is missed and underestimated. The observed deviation in H_0 can be interpreted as follows: when H_0 is decreased, GWs events tend to be assigned lower redshifts, resulting in higher masses in the source frame. By reducing the value of H_0, the BPL model attempts to accommodate the clustering of high-mass events generated by the MLTP model. In terms of the simulations employing the most intricate model, the MLTP, it is evident that both the PLP and BPL models struggle to accurately reproduce the lower end of the mass distribution. Nonetheless, this discrepancy does not reach a level where it induces a bias in H_0 during the reconstruction with the PLP. Remarkably, the CBC merger rate across redshifts is consistently reconstructed across all scenarios. In essence, the more pronounced mass feature within the BBH mass distribution can induce a substantial bias in the estimation of H_0.

As for the previous test case, this analysis has been repeated ~ 50 times on independent population realizations to ensure stability in our findings, and not just a statistical

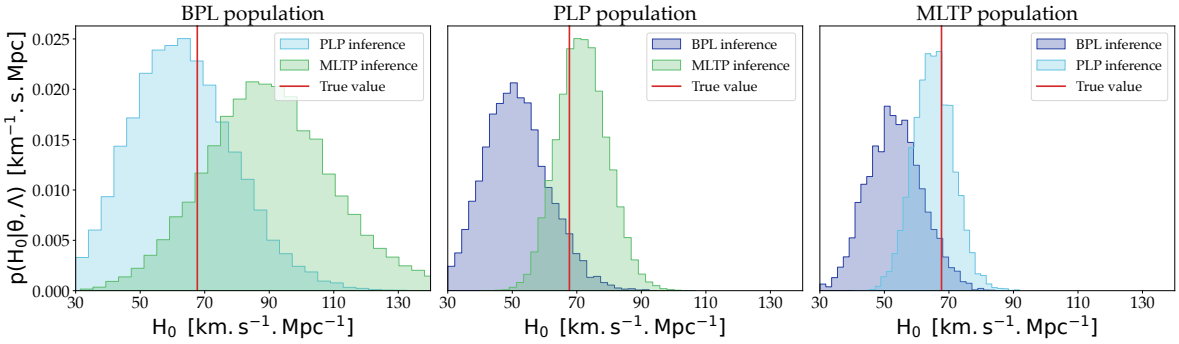

Figure 3.6: Left: Marginal distribution of H_0 inferred with the PLP and the MLTP mass models, on a BBH population simulated with the BPL model. Middle: Marginal distribution of H_0 inferred with the BPL and the MLTP mass models, on a BBH population simulated with the PLP model. Right: Marginal distribution of H_0 inferred with the BPL and the PLP mass models, on a BBH population simulated with the MLTP model. On all three plots, the true value of the Hubble constant is indicated by the red vertical line and is equal to the Planck15 value.

fluctuation. Fig. 3.8 presents the PP plot corresponding to the middle plot of Fig. 3.6, where the GW observations are simulated with the PLP mass model and the Spectral siren analysis is performed with the BPL model. From the PP plot behavior, laying below the diagonal line, there is undoubtable support for bias on the inference of the Hubble constant. The PP plot for H_0, being below the diagonal shows that most of the time, the true value of the Hubble constant is being underestimated, as shown on Fig. 3.6. All other PP plots are shown in App. B.2.2.

3.3.3 Impact of a redshift evolution of the mass spectrum

Sec. 3.3.3 aims to examine the robustness of redshift independent mass models to a possible evolution of the mass spectrum in redshift. For instance, the PLP, BPL and MLTP models are population models that are fixed in redshift, and currently used to performing GW cosmology studies within the LVK collaboration and external groups. To test this assumption, we assume a linear evolution of the features of the mass spectrum with respect to the cosmological redshift. In this section, we modify the PLP source mass model such that its Gaussian peak is able to evolve linearly with the redshift. In particular, the evolution is encoded in the position of the Gaussian peak, while its standard deviation remains the same. This pattern of evolution might emerge from a combination of factors, including the PISN mass threshold and the time intervals between the formation and merging of binary systems, as indicated in references [131, 144]. These factors notably constrain the mass distribution above the PISN threshold (the peak in the Gaussian).

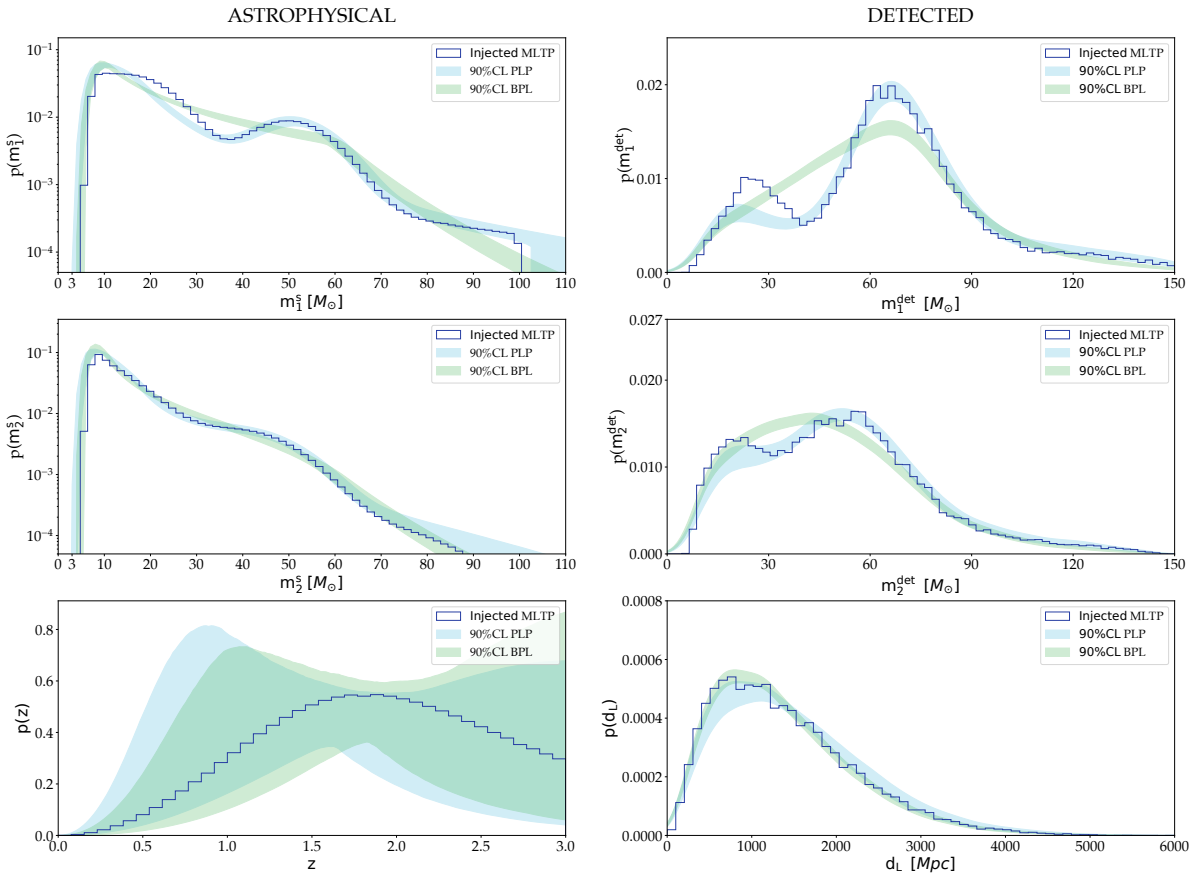

Figure 3.7: Posterior predictive checks when the wrong mass model is used for the inference, it corresponds to the right plot on Fig. 3.6. The blue histogram is the injected BBH population, that has been generated from the MLTP mass mode, using the same population parameters as for Sec. 3.3.1. The two colored contours are the 90% C.L. inferred from the Spectral analysis, using in green the BPL mass model and in light blue the PLP mass model.

Here instead, we do not suppress the mass spectrum above the PLP model Gaussian peak, but we assume a linear dependency,

$$\mu_g(z) = \mu_g^0 + z\left(\mu_g^1 - \mu_g^0\right), \tag{3.13}$$

where $\mu_g^0 = \mu_g(z=0)$ and $\mu_g^1 = \mu_g(z=1)$ are the positions of the Gaussian peak at $z = 0$ and $z = 1$. We fix $\mu_g^0 = 30\,M_\odot$, and we considered seven different values for μ_g^1 from $25M_\odot$ to $35\,M_\odot$. The specific point of $\mu_g^1 = 30\,M_\odot$ corresponds to zero evolution of the position of the peak. The primary motivation for shifting only the Gaussian peak of the mass distribution is that most of the constraining power on the Hubble constant comes from accurately reconstructing this mass feature, and because this constraining power is tied to the peak, potential biases could arise from its incorrect reconstruction if the peak of the true mass distribution evolves with redshift. This parameterization also represents the simplest modification of the model and some hints of such evolution have been observed in non-parametric study [145]. Future work could explore more complex models with additional mass features evolving with redshift, i.e. the

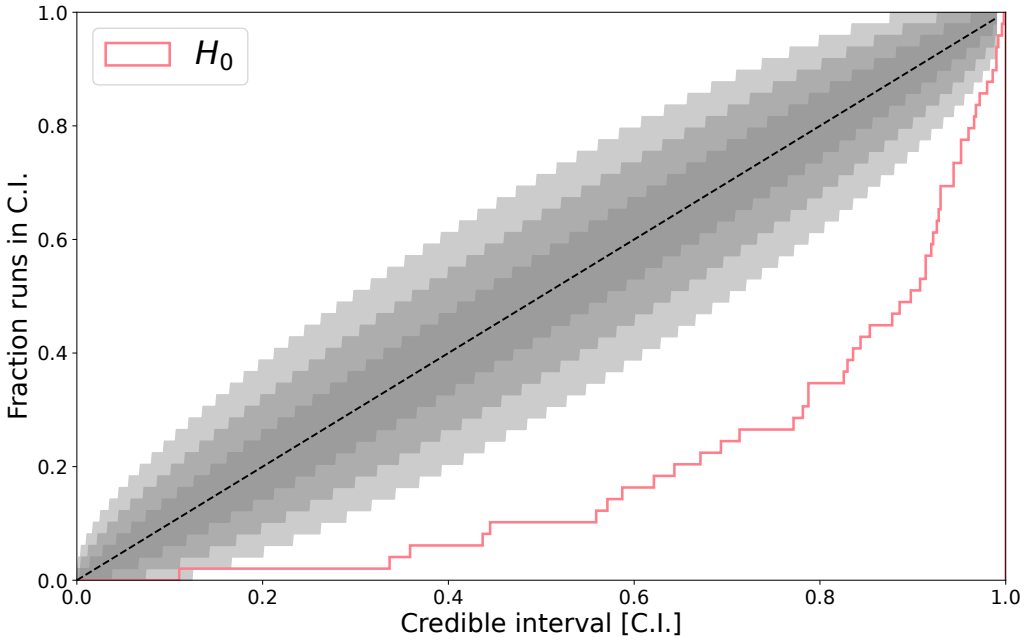

Figure 3.8: PP plot obtained from 50 independent population realizations simulated with the PLP mass model, and inferred with the BPL model too. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

minimum and maximum mass or the slope of the power law. However, from an astrophysical perspective, it remains unclear how the true BBH mass spectrum varies across different redshift ranges.

We simulated seven scenarios of BBH populations with the modified PLP model, with redshift evolution shifts ranging from $-5\,M_\odot$ to $+5\,M_\odot$, and for each one of them, sets of 2000 GW observations are generated and then pass down to ICAROGW for a full Spectral siren analysis. For the Bayesian inference, we use the non-evolving PLP mass model. The estimation of H_0 with a 68.3% confidence interval is presented in Fig. 3.9. This figure illustrates how the inferred value of H_0 changes in relation to the value of μ_g^1. To illustrate, if the Gaussian peak shifts by approximately five solar masses from redshift 0 to redshift 1, the estimated value of H_0 falls outside the 99.7% confidence interval. Fig. 3.9 demonstrates how the evolution of structures in the mass spectrum with redshift can lead to a bias in the inferred Hubble constant. Even a modest evolution of the mass spectrum across redshifts results in a noteworthy bias in the estimated H_0. Additionally, it's worth noting that the systematic bias observed in H_0 is directly proportional to the extent of the redshift-dependent evolution in the mass spectrum feature. This bias arises from inaccuracies in reconstructing the source frame mass spectrum, as elaborated in the following paragraph.

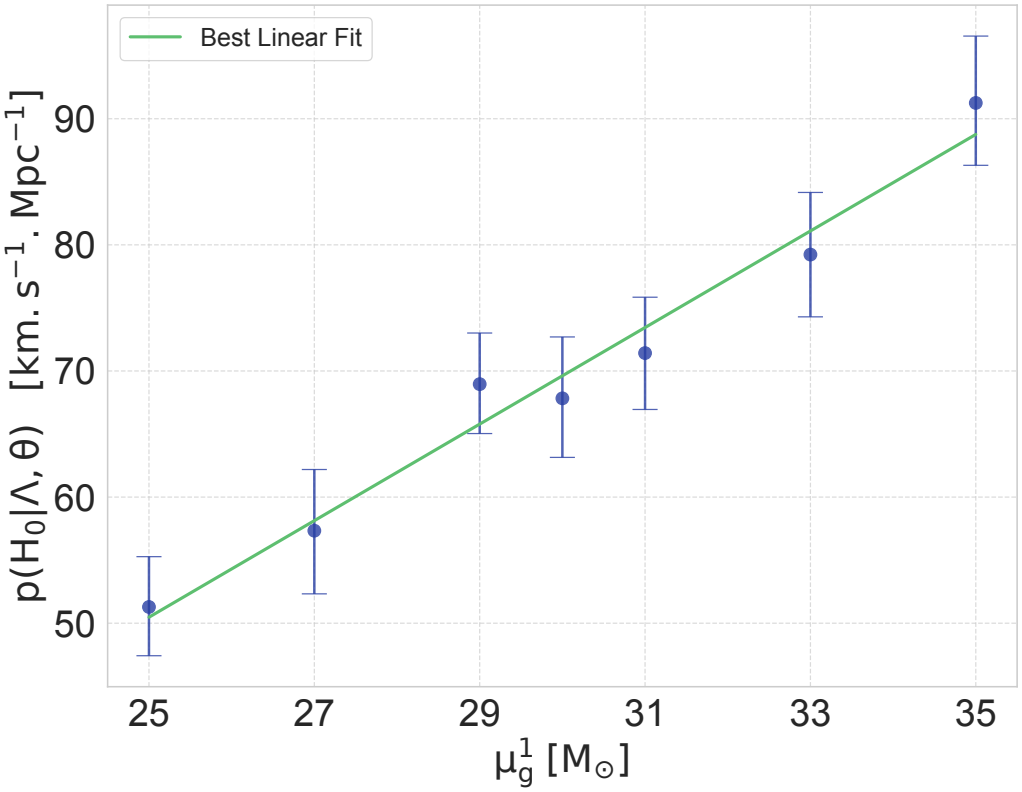

Figure 3.9: Scatter plot showing the evolution of the inferred H_0 value with respect to the position of the Gaussian peak of the PLP mass model used for the simulation of the GW observations at $z = 1$. The blue points correspond to the median of the H_0 posteriors, for each of the seven Spectral sirens analysis. The point of reference (no redshift evolution) is $\mu_g^1 = 30\,M_\odot$. The blue errors bars are the estimated $1\,\sigma$ intervals of the Hubble constant posteriors and the green line is the best linear fit to fit all the points.

Fig. 3.10 provides a summary of the PPCs for this specific test scenario. In the right column, the detected mass distributions from three out of the seven simulations are depicted. In the detector frame, these three simulated populations exhibit similar distributions for the masses and luminosity distance of the detected GW events. These distributions align with those assumed by the non-evolving PLP models. However, when observed in the source frame, these simulations demonstrate markedly different mass spectra, attributed to redshift evolution of the Gaussian peak. The PLP model fails to capture this evolution, resulting in an inaccurate reconstruction of the true mass spectrum. Notably, the non-evolving PLP model consistently reconstructs a peak located at $30\,M_\odot$. This outcome stems from the fact that approximately 80% of the GW detections occur at low redshifts ($z < 0.3$), where even in models with redshift evolution, the peak of the Gaussian component typically resides around $\mu_g^0 = 30\,M_\odot$.

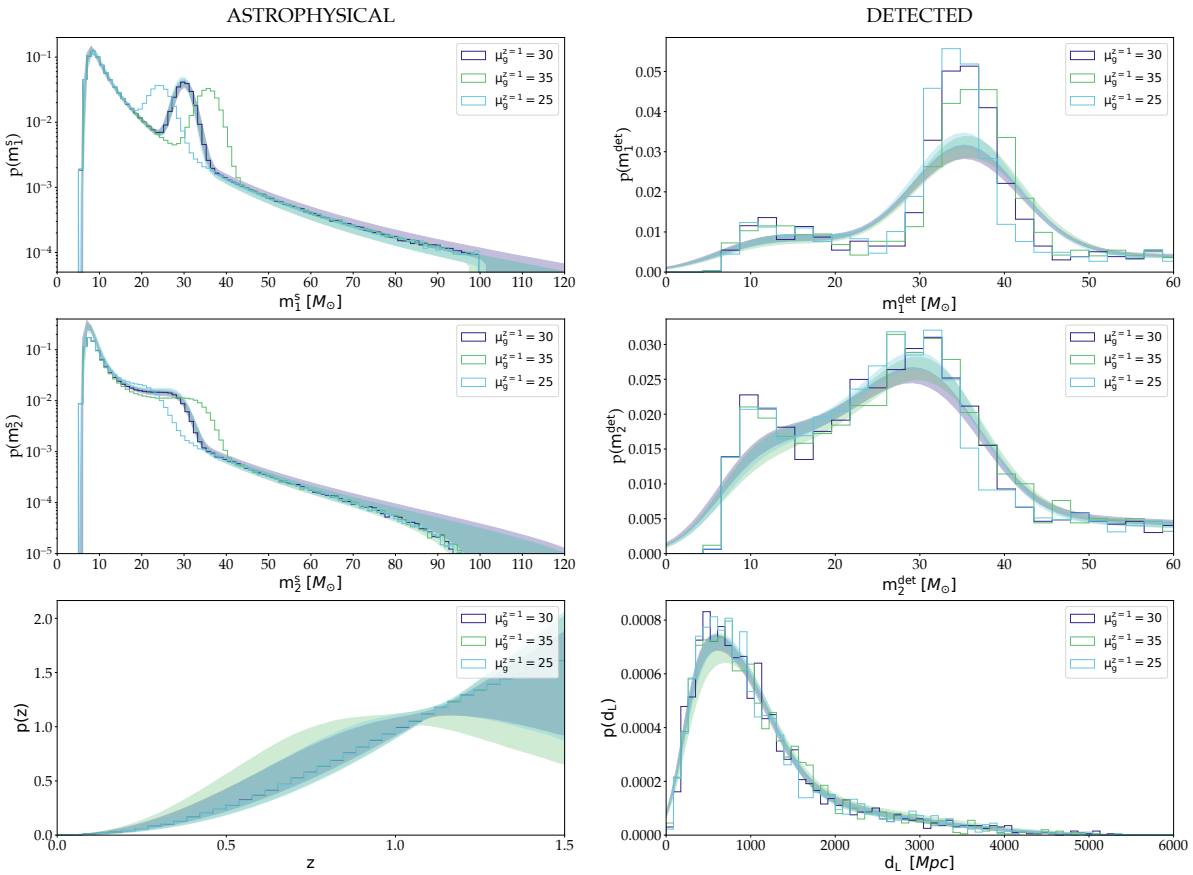

Figure 3.10: Posterior predictive checks when the wrong mass model is used for the inference, it corresponds to the right plot on Fig. 3.6. The blue histogram is the injected BBH population, that has been generated from the MLTP mass mode, using the same population parameters as for Sec. 3.3.1. The two colored contours are the 90% C.L. inferred from the Spectral analysis, using in green the BPL mass model and in light blue the PLP mass model.

The emergence of an H_0 bias stems from the misalignment of the Gaussian component for events occurring at higher redshifts. When $\mu_g^1 > \mu_g^0$ and the PLP model fails to adequately fit this discrepancy, we observe an H_0 bias towards higher values. This bias arises because, at higher H_0 values, GW events are situated at higher redshifts, resulting in lower source masses. By assigning these events at lower source masses, the PLP model can encompass them within the peak typically centered around $30\,M_\odot$ in the source frame. Conversely, when $\mu_g^1 < \mu_g^0$, we note a bias towards lower values of H_0. In this scenario, the GW events are placed at lower redshifts, thereby possessing higher source masses that align with the peak at $30\,M_\odot$. This simplified test case underscores the potential for redshift evolution of a feature in the mass spectrum to introduce significant bias in the estimation of H_0 based on GWs.

Consistency in results is confirmed by conducting several independent simulation runs for each data point, underscoring the reliability of the findings depicted in Fig. 3.9. Fig. 3.11, from these independent realizations, shows the PP plot for inference of H_0, when the Gaussian peak is moving to $35\,M_\odot$ at $z = 1$. The pink histogram railing to

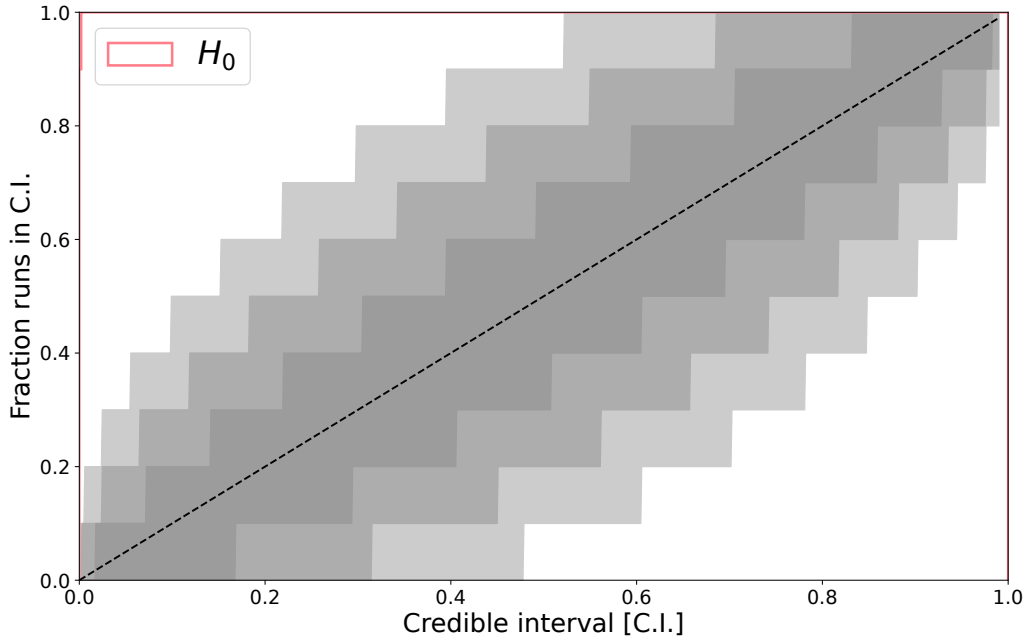

Figure 3.11: PP plot obtained from ~ 20 independent population realizations simulated with the modified PLP model, when the Gaussian peak is evolving from $30\,M_\odot$ at $z = 0$ up to $35\,M_\odot$ when $z = 1$. The spectral siren inference is done with the regular non evolving PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

the edge of the canvas, above the diagonal line, is a strong evidence that the Hubble constant measurement is biased toward higher values, as shown above. Respectively, PP plots railing below the diagonal have been found when in the simulated GW observations, the Gaussian peak is moving at lower masses when the redshift increases (see App. B.2.3).

3.3.4 Discussion and conclusions

In essence, the bias observed in the estimation of H_0 is intricately linked to the model's capacity to accurately capture and reconstruct the intricate features inherent in the CBCs mass spectrum. This bias can be influenced by a multitude of factors, each playing a crucial role in shaping the final inferred value of Hubble constant. Our analysis reveals that when the true underlying population of CBCs exhibits sharp, distinctive mass features that diverge from the assumptions of the reconstruction model, a substantial bias in H_0 emerges. This discrepancy underscores the importance of ensuring that the model accounts for the full complexity of the underlying mass distribution to avoid introducing systematic errors in the estimation process. Moreover, our investigations highlight the significance of considering redshift-dependent evolutions in the

mass spectrum's features. Even subtle shifts in these features, on the order of a few solar masses, can lead to the introduction of bias in the estimation of H_0. This sensitivity underscores the need for rigorous examination and refinement of models to account for any evolving dynamics in the mass spectrum across different redshifts.

It is important to note that the choice of 2000 detected events for this study case, while useful for generating informative and nearly Gaussian-like posteriors for H_0, is not entirely realistic for O4 sensitivity. This number roughly corresponds to about 10 years of data taking, beyond what is expected for the O4 observation run. However, this large sample size allows us to clearly illustrate the impact of biases in the reconstruction model. Looking ahead, as GW detections continue to accumulate, these results highlight a critical challenge for future cosmological analyses. If the mass spectrum indeed evolves with redshift, such biases could have a significant effect on H_0 estimates. Even in scenarios where the number of detected events is smaller, and the H_0 posterior is less informative, these biases could still persist, only masked by the broader error budget. As the precision on H_0 improves, these biases will become increasingly visible and impactful. This scenario is likely to emerge during O5, and even more so with next-generation detectors that will yield thousands of GW detections. Thus, addressing the evolving dynamics of the mass spectrum is essential to mitigate these biases and ensure accurate cosmological measurements.

In summary, our findings emphasize the intricate interplay between the underlying population characteristics of CBCs, the reconstruction model's capacity in capturing these features, and the potential impact of redshift evolution on biasing the estimation of H_0. Addressing these complexities is crucial for ensuring the accuracy and reliability of H_0 estimates derived from GW observations.

3.4 APPLICATION TO A COMPLEX BINARY BLACK HOLE POPULATION

In this section, we delve into an examination of how the BPL, PLP and MLTP source mass models interact with a population of BBHs mergers generated through synthetic astrophysical simulations. Our analysis utilizes a BBH catalog named "A03," comprising BBH mergers originating from diverse formation channels as detailed in [124, 125]. Subsequently, we present the results of our hierarchical Bayesian analysis conducted with these mass models in Sec. 3.4.2. Finally, in Sec. 3.4.3, we investigate potential sources of bias affecting the estimation of the Hubble constant.

3.4.1 General description of the *A*03 synthetic catalog

For this investigation, we utilize the A03 catalog from [125], comprising four distinct channels: isolated binary evolution and the dynamical assembly of BBHs within young stellar clusters, globular stellar clusters, and nuclear star clusters. These channels are combined based on their respective redshift-dependent rates as outlined in

[125]. The dynamical formation channels are characterized by an initial BH mass function derived from the MOBSE population synthesis code [146]. In model A03, they adopt the delayed model proposed in [38] for core-collapse supernovae, incorporate the pair-instability supernova formalism from [40], and account for the influence of progenitor star metallicity on the BH mass function, as detailed in [146]. Subsequently, BHs dynamically pair with others and undergo hardening due to interactions within their host star clusters [124], while hierarchical mergers of BBHs are also permitted, as described in [125]. The isolated binary evolution channel comprises systems evolved using the MOBSE binary population synthesis code, which considers the evolution of binary stars [146]. These simulations incorporate mechanisms such as mass transfer, common envelope, tidal evolution, natal kicks, and GW decay as outlined in [146]. A comprehensive discussion of the astrophysical processes embedded within model A03 and their associated uncertainties falls outside the scope of this study. Here, we aim to investigate the potential for nontrivial structures within the BBH mass distribution to introduce bias into the estimation of H_0. Fig. 3.12 illustrates the distribution

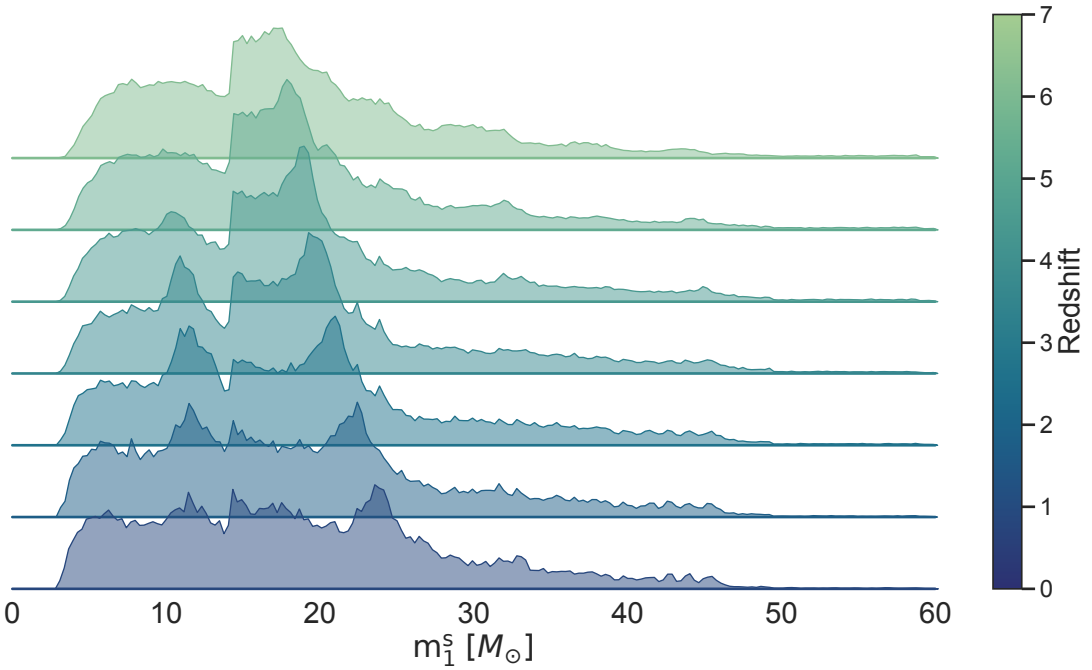

Figure 3.12: Mass spectrum of the primary masses in source frame from the *A*03 BBH catalog as a function of the redshift. Each plot corresponds to a slice in redshift, going from $z = 0$ up to $z = 7$. This figure shows several important features regarding the mass spectrum of BBHs inside the *A*03 catalog: redshift evolution, appearance of peaks, shifts of peaks.

of primary masses m_1^s across various redshift bins. The BBH mass spectrum reveals distinct features: progressing from lower to higher masses, there is a feature around $10M_\odot$, a "valley" spanning from $13M_\odot$ to $15M_\odot$, and a subsequent prominent peak at $16M_\odot$. Additionally, in the higher mass regime of the spectrum, a localized overabundance of BBHs is evident around $25M_\odot$, along with several less pronounced peaks extending up to $90M_\odot$. These structural characteristics are also discernible in the dis-

tribution of secondary masses m_2^s. Furthermore, the *A*03 catalog exhibits a significant redshift-dependent evolution in both primary and secondary source frame masses. As depicted in Fig. 3.12, it's evident that the features within the mass spectrum evolve as redshift increases. Notably, the feature at $20M_\odot$ is absent for mergers below a redshift of approximately 2, gradually emerging while transitioning to lower masses in higher redshift bins.

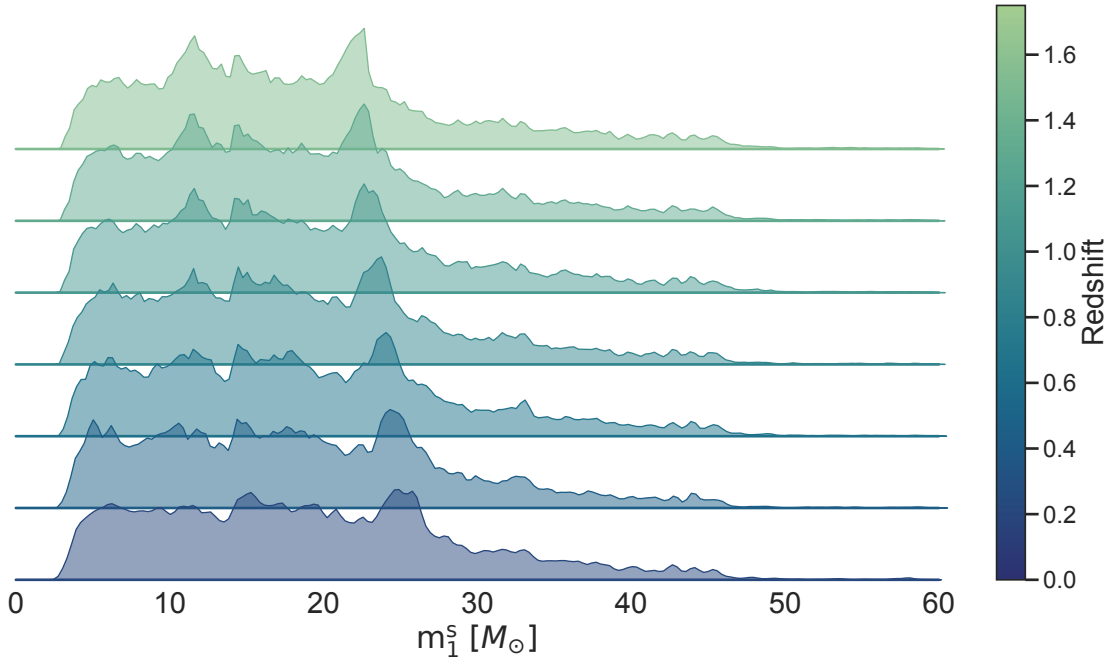

Figure 3.13: Low redshift mass spectrum of the primary masses in source frame from the *A*03 BBH catalog as a function of the redshift. Each plot corresponds to a slice in redshift, going from $z = 0$ up to $z = 1.75$. This figure shows several important features regarding the mass spectrum of BBHs inside the *A*03 catalog: redshift evolution, appearance of peaks, shifts of peaks.

Fig. 3.13 showcases the BBH mass spectrum within a redshift range of $0 < z < 1.75$, corresponding to the typical detection range of our simulations of BBH mergers. The evolution of the BBH mass spectrum at low redshifts appears more subdued compared to higher redshifts. Nevertheless, peaks around $11M_\odot$ emerge while transitioning to higher masses, while the structure around $25M_\odot$ shifts towards lower masses. Although real GWs events at high redshifts ($z > 1$) are challenging to detect in networks akin to the O4 observing run sensitivity, this effect was considered to have a subdominant impact on population and cosmological as proposed in [147]. However, if these structural changes are indeed evolving with redshift, as discussed in [148–152], it's plausible that mismatches between the mass models and these evolving features could introduce bias in the estimation of H_0, as detailed in Sec. 3.4.3.

3.4.2 Vanilla analysis of the *A*03 catalog

Utilizing the A03 catalog of BBHs and the framework outlined in Sec. 3.2, we simulate our own GW catalogs comprising 2000 detected GW events. Subsequently, for each of these catalogs, we undertake a comprehensive hierarchical Bayesian inference, yielding posterior distributions for H_0, alongside other population parameters such as the CBC merger rate and the three redshift-independent BBH mass models. The prior ranges employed for the Bayesian inference are detailed in tables A.1 to A.3. Fig. 3.14 illustrates the marginal posterior distributions of the estimated Hub-

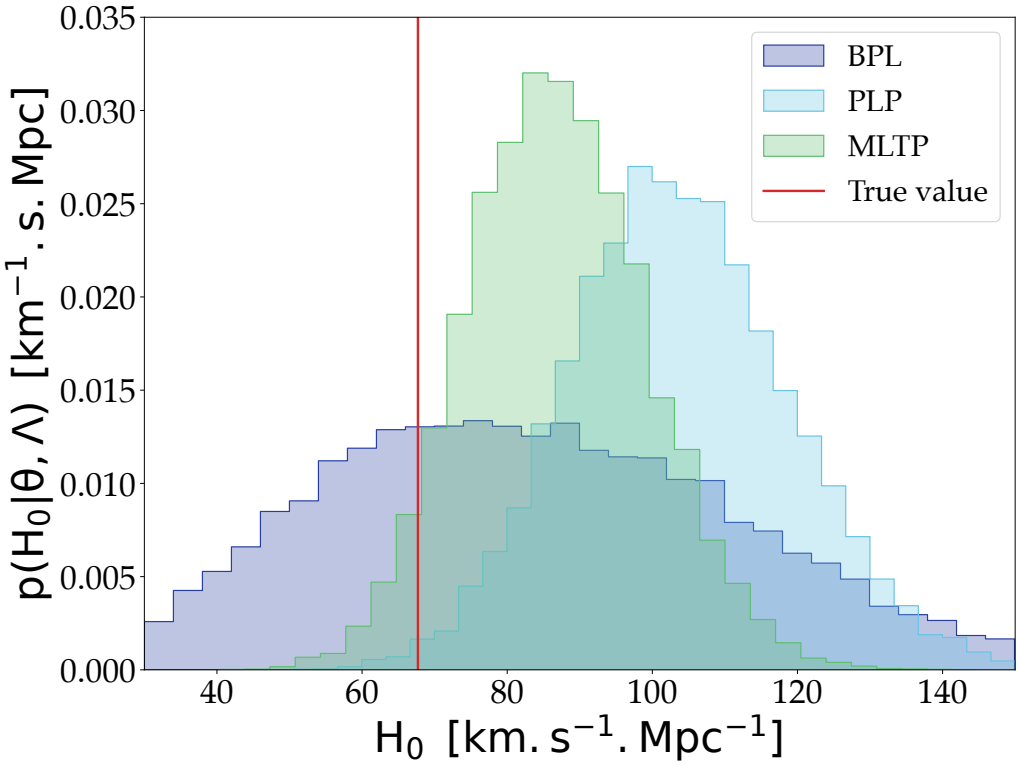

Figure 3.14: Marginal posteriors of the inferred Hubble constant H_0 resulting from the Spectral siren analysis of 2000 GW observations simulated from the *A*03 BBH catalog. In blue using the BPL model, in light blue with the PLP and the in green with the MLTP mass model. The red vertical line is the true value of H_0 in the catalog, fixed at $H_0 = 67.7\,km\,s^{-1}\,Mpc^{-1}$.

ble constant H_0. Notably, when employing the PLP and MLTP mass models, we observe a bias towards higher values of H_0. Specifically, for the PLP model, we estimate $H_0 = 103^{+11}_{-10}\,km\,s^{-1}\,Mpc^{-1}$ and for the MLTP model, H_0 is estimated to be $H_0 = 87^{+9}_{-8}\,km\,s^{-1}\,Mpc^{-1}$ at a 68.3% confidence interval. The true value of H_0 falls outside the 99.7% confidence interval for the PLP model and the 95% confidence interval for the MLTP model. Furthermore, it's discernible from Fig. 3.14 that the BPL mass model yields an uninformative posterior distribution for H_0. We obtain a value

of $H_0 = 82^{+20}_{-19}\, km\, s^{-1}\, Mpc^{-1}$, which encompasses the injected H_0 within the 68% confidence interval. As we will discuss further, the BPL model retrieves a less informative posterior on H_0, attributed to the absence of strong mass features. It is worth noting

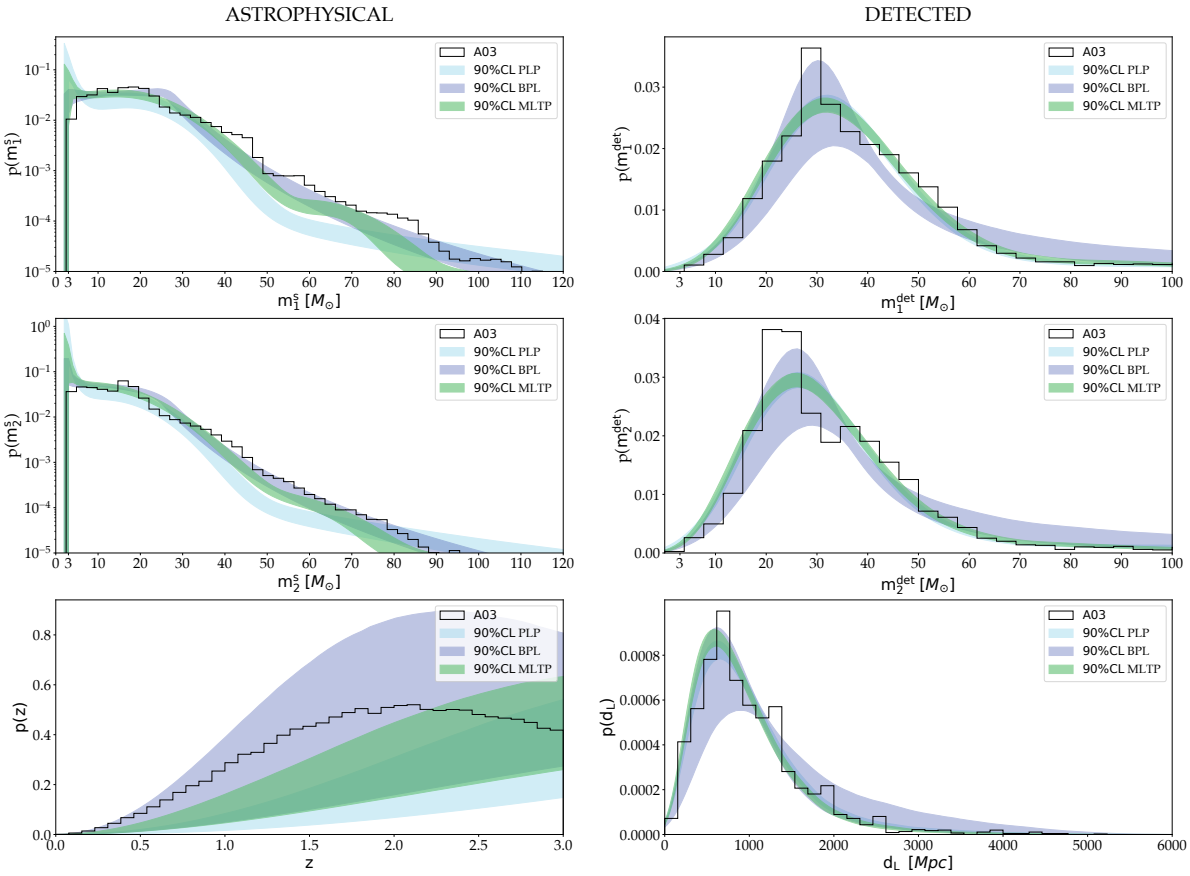

Figure 3.15: Posterior predictive checks of the Spectral siren analysis on all population and cosmological parameters performed on 2000 GW observations simulated from the *A*03 BBH catalog. The plain black histograms are the true population contained in the A03 catalog, in the source frame (left) and in the detector frame (right). The light blue contours are the inferred population with the PLP model, in darker blue with the BPL model and in green with the MLTP model. These contours are the 90% C.L.

that while the BPL model does not exhibit an H_0 bias for 2000 GW events, it may potentially manifest as more GW detections are used. Importantly, we made sure that this bias is not introduced by any of the other population parameters exceeding their prior ranges, behavior usually called "railing". In fact, we have verified that all parameters are well constrained within their respective prior ranges.

The PPC plot depicted in Fig. 3.15 provides valuable insights into the origins of the H_0 bias. While the reconstructed distributions of detector frame masses and luminosity distance exhibit no significant deviations between the injected population and the recovered one, noticeable differences emerge in the reconstructed source population. Specifically, in the source frame, the BPL model correctly reconstructs the CBC merger rate within the 90%C.L. Conversely, the PLP and MLTP models notably underestimate

the presence of BBHs within the mass range of $40M_\odot$ to $80M_\odot$. Consequently, based on the observed data, the PLP and MLTP models predict a considerably higher number of BBHs at lower source masses. This discrepancy is addressed by adjusting the value of the Hubble constant to higher values, thereby positioning events at higher redshifts with, on average, lower source frame masses. Another corroborating indicator supporting this conclusion is the reconstruction of the CBC merger rate across redshifts. As depicted in the bottom-left plot of Fig. 3.15, the PLP and MLTP models reconstruct a biased merger rate, favoring higher values of redshifts for BBH mergers. In contrast, the inference provided by the BPL model accurately captures the overall trend of the *A*03 source frame population. Specifically, the BPL model achieves superior reconstruction of source frame masses within the range of $40M_\odot$ to $80M_\odot$.

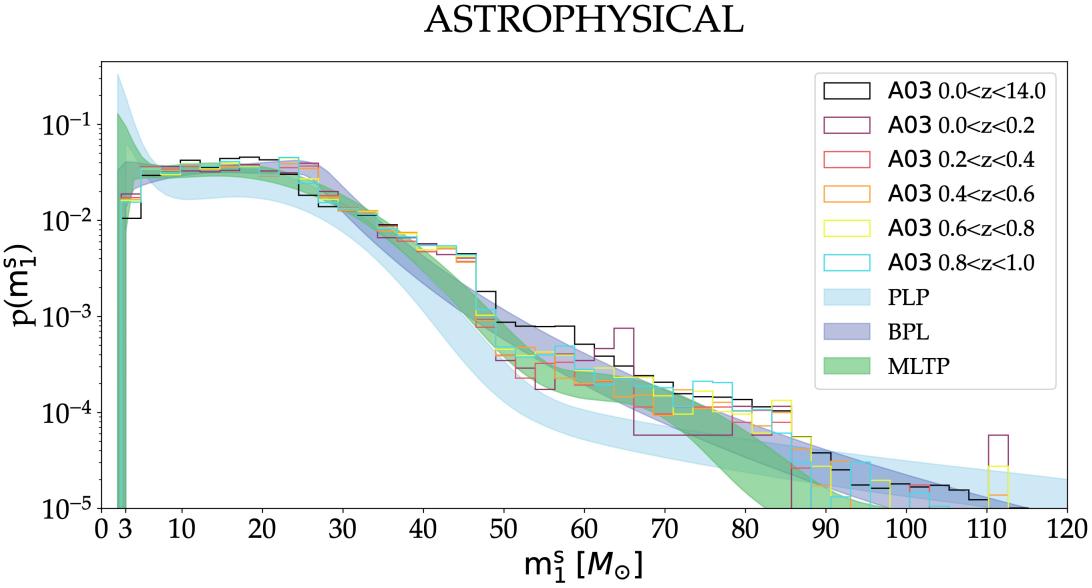

Figure 3.16: Posterior predictive check for the Spectral siren analysis of 2000 GW observations simulated from the *A*03 BBH catalog. This figure shows similar results as Fig. 3.15, except that only the primary source frame mass and the true population is displayed, and has been decomposed into several redshift bins (colored histograms) from $z = 0$ to $z = 2$. Similarly, the colored contours are the 90% C.L. reconstructed from the Bayesian inference.

Despite having sufficient degrees of freedom to approximate a smooth distribution of masses akin to those present in the *A*03 catalog, both the PLP and MLTP models still fall short in accurately reconstructing the mass spectrum. Understanding why these models fail to correctly capture the mass spectrum of the *A*03 catalog and consequently introduce bias in H_0 estimation is not straightforward. Initially, one hypothesis suggested that the PLP and MLTP models might be capturing local peaks in the mass spectrum at a fixed redshift that evolve with redshift. Indeed, from the posterior distributions, it was observed that these models identified local peaks at low masses (approximately $17M_\odot$) with a large standard deviation, along with a higher peak around $30M_\odot$. To explore this hypothesis, a PPC was constructed, as depicted in Fig. 3.16, by dividing the true population into redshift bins. Unfortunately, the PPC

does not clearly indicate any discrepancy between the population reconstruction and any of the mass spectra for the respective redshift bins. Given the inconclusive nature of this test in elucidating the origin of the H_0 bias, we investigated additional studies, which are further discussed below.

Consistent results have been replicated across approximately $\sim$ 15 distinct simulation realizations for each of the three posterior distributions depicted in Fig. 3.14. Fig. 3.17 presents the PP plot for one of the three inferences of the Hubble constant from the A03 BBH catalog, using the PLP mass model. From the PP plot, again there is a strong support for a biased estimation of the Hubble constant, since the histogram goes way above the expected fluctuations intervals for the Spectral siren analysis. And in agreement with the posterior found in Fig. 3.14, the Hubble constant measurement with the PLP mass model is overestimated. The remaining two PP plots, based on the BPL and MLTP mass models are displayed in the App. B.2.4.

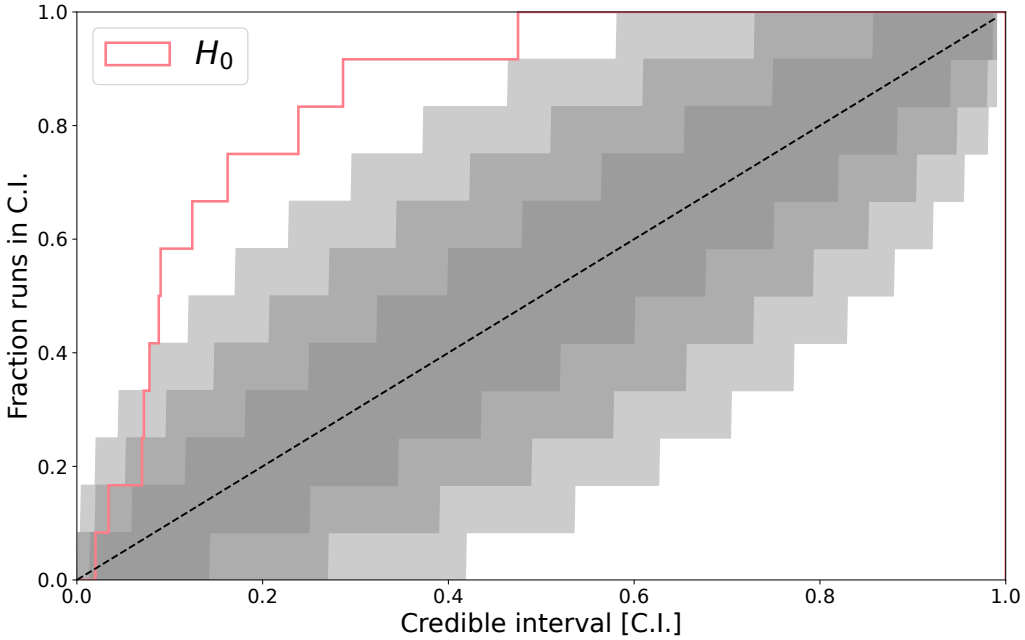

Figure 3.17: PP plot obtained from $\sim$ 15 independent population realizations simulated from the A03 BBH catalog. The spectral siren inference is done with the PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

3.4.3 Investigating the sources of the H_0 bias: Blinding the mass-redshift relation

To explore the potential impact of redshift evolution of the mass spectrum on the introduction of an H_0 bias, we conducted a simulation aimed at isolating the effects of redshift on the mass spectrum. In this simulation, we "blinded" the mass spectra of

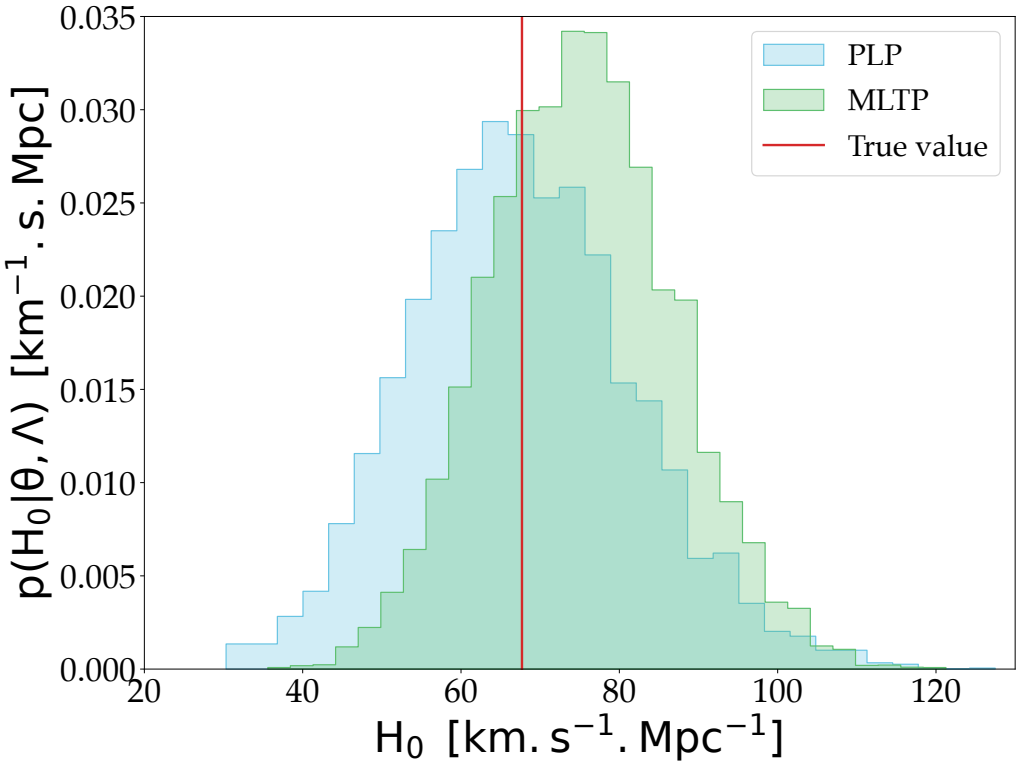

Figure 3.18: Marginal posteriors of the inferred Hubble constant H_0 resulting from the Spectral siren analysis of 2000 GW observations simulated from the *A*03 BBH catalog. In blue using the BPL model, in light blue with the PLP and the in green with the MLTP mass model. The red vertical line is the true value of H_0 in the catalog, fixed at $H_0 = 67.7\,km\,s^{-1}\,Mpc^{-1}$.

the *A*03 catalog to redshift evolution. This was achieved by randomly shuffling the pairs of BBH merger redshifts and masses. By doing so, we artificially eliminated the redshift dependency of the mass spectrum while preserving its inherent shape and complexity. For the Bayesian inference process, we focused solely on PLP and MLTP mass models, as these models exhibited a notable bias in H_0. This selective approach allowed us to concentrate on the models demonstrating significant deviations and explore the effects of redshift evolution in a controlled setting.

The result of the joint inferences conducted with the PLP and MTLP mass model is depicted in Fig. 3.18. For the PLP model, we ascertain Hubble constant to be $H_0 = 67^{+9}_{-8}\,km\,s^{-1}\,Mpc^{-1}$, while with the MLTP model, yields H_0 to be $H_0 = 75^{+7}_{-8}\,km\,s^{-1}\,Mpc^{-1}$, with the true value of H_0 falling within the 68.3% confidence interval of the posterior distribution. Fig. 3.19 displays the PPC plot for this test. The plot overlays the reconstructions for the original *A*03 catalog and the catalog after the redshift-blinding procedure. From the perspective of the detector frame, the two reconstructed distributions exhibit no significant disparities. However, when viewed from the source frame standpoint, the PLP model, when applied to the redshift-blinded *A*03 catalog,

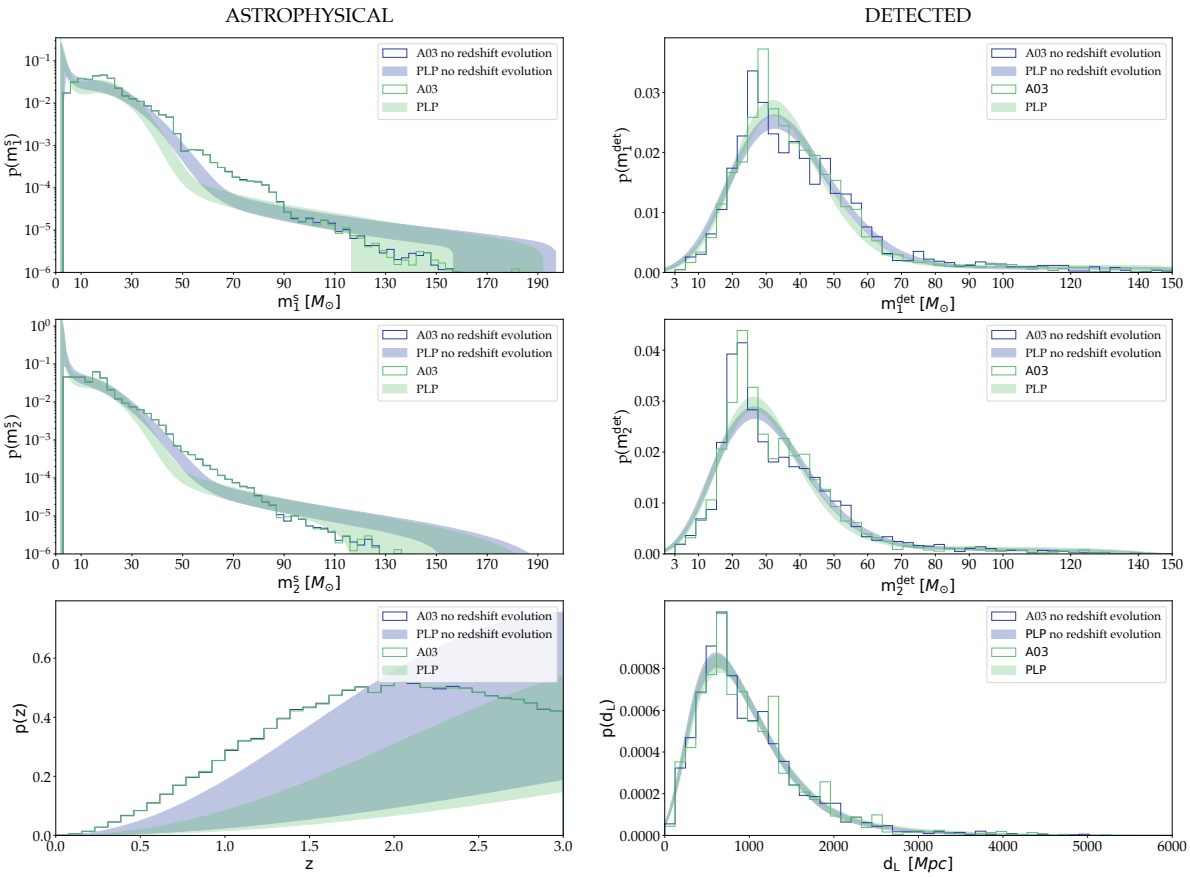

Figure 3.19: Posterior predictive check for the Spectral siren analysis of 2000 GW observations simulated from the *A*03 BBH catalog. This figure shows the same results as Fig. 3.15, except for the true population of source frame masses, that have been decomposed into several redshift bins (colored histograms) from $z = 0$ to $z = 2$. Similarly, the colored contours are the 90% C.L. reconstructed from the Bayesian inference.

demonstrates improved reconstruction of BBH masses within the range of $30M_\odot$ to $60M_\odot$. Though the reconstruction of the mass spectrum above $60M_\odot$ remains subpar, the PLP model for the non-redshift evolving *A*03 catalog manages to encompass the true value of H_0. Consequently, the reconstruction of the BBH CBC merger rate as a function of redshift exhibits slight enhancement, resulting in an underestimation of approximately 20% to 30% of the true CBC merger rate. Similar results have been obtained for the MLTP model as well.

It is crucial to highlight that even after neutralizing the redshift evolution of the *A*03 mass spectrum, the PLP and MLTP models remain unable to fully reconstruct the source frame distributions. Consequently, we anticipate the persistence of a systematic bias on H_0, even though concealed by the large statistical uncertainties evident in fig. 3.18. Identifying the root cause of these systematic biases is a complex endeavor. However, based on the tests conducted, we can reasonably infer that the redshift evolution of the mass spectrum likely serves as the predominant source of the bias in H_0. Despite our efforts to mitigate its influence, the residual impact of this evolutionary

aspect underscores the intricate challenges involved in accurately estimating fundamental cosmological parameters from GW observations.

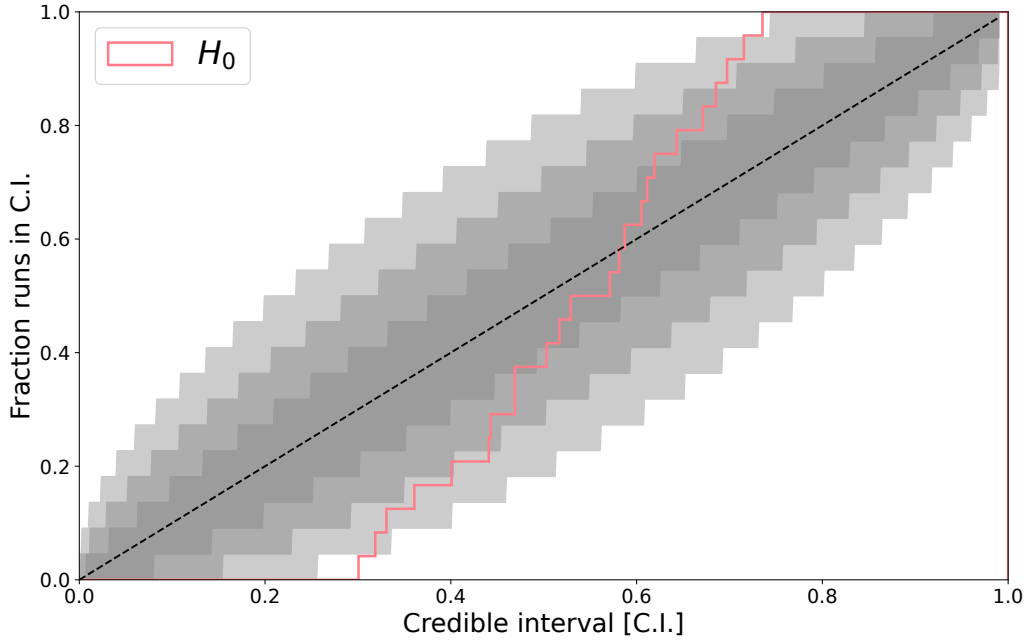

Figure 3.20: PP plot obtained from ~ 20 independent population realizations simulated from the A03 BBH catalog, where the redshift evolution of the mass spectrum has been artificially removed. The spectral siren inference is done with the PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

Through simulations conducted with 20 different population realizations, it has been verified that the value of H_0 obtained by the PLP and MLTP models from the redshift-blinded *A*03 catalog consistently remains unbiased. Fig. 3.20 shows the PP plot of these repeated analysis with the PLP mass model. In contrast with Fig. 3.17, where the PP plot was showing a clear deviation from the diagonal, now the histogram stays close to it and most of the curve is contained in the statistical fluctuations contours. The behavior of the histogram on this PP plot is the one of an unbiased inference, but where the final posterior is not fully constrained in the prior range (which is the case for the Hubble constant). The PP plot for the analysis with the MLTP mass model can be found in App. B.2.5.

3.5 CONCLUSION

In this chapter, we have delved into the intricate relationship between systematic biases affecting the estimation of the Hubble constant (H_0) and the reconstruction of

the BBH mass spectrum within the framework of Spectral Siren Cosmology. Through a series of simulations utilizing BBH merger catalogs generated with three prevalent mass models from the literature, alongside a synthetically simulated BBH catalog, we have explored various factors contributing to potential biases in H_0 estimation.

In Sec. 3.2, we demonstrated the efficiency of our python framework to quickly generate GW observations from a set of source frame masses and redshift. From these GW observations, we validated the framework simulating error budgets for each of the parameters and passing them through the Spectral siren inference of ICAROGW. We successfully recovered all population parameters from this inference, as well as the Hubble constant.

In Sec. 3.3, our investigations unveiled that simple phenomenological models devoid of sharp features could introduce biases on H_0, particularly if the true BBH population includes local overdensities of sources. Furthermore, we elucidated how the redshift evolution of source frame mass features can induce biases in H_0 estimation if not adequately accounted for in the mass models. Specifically, we demonstrated that the Hubble constant acts as a compensatory parameter, shifting to higher values if the true source mass spectrum is underestimated, and vice versa if overestimated.

Moving to Sec. 3.4, utilizing the *A*03 BBH catalog, we showcased that even when the detector frame is well-reconstructed by the inference, a biased value of H_0 may still occur. Notably, deviations of up to 2σ and 3σ from the injected value were observed for the MLTP and PLP model inferences, respectively. The discrepancy between the inferred populations in the source frame among different models was found to be related to their unique responses to GW events, prompting adjustments in the inferred populations aligned with observed data. Moreover, we demonstrated that the H_0 bias obtained with the PLP and MLTP mass models vanishes when the redshift evolution of the source mass spectrum is removed from the population.

Our findings demonstrate the limitations of commonly used mass models, namely the PLP, BPL, and MLTP models, in conjunction with the spectral siren analysis for GW cosmology, especially in scenarios where the BBH population exhibits even slight redshift evolution of its mass spectrum. Additionally, we highlight the potential for biased H_0 estimation arising from mismatches in the BBH mass distribution's sharp structures. These insights caution against the indiscriminate use of non-evolving parametric mass models, emphasizing the importance of incorporating models capable of absorbing effects such as redshift evolution and possessing adequate flexibility to accommodate unforeseen mass features. Looking ahead, future GW cosmology studies employing the spectral siren method should adopt source mass models with enhanced capabilities to mitigate biases. This entails accounting for redshift evolution effects and possessing sufficient degrees of freedom to adapt to variations in the true distribution of BBHs in the Universe. Failure to address these considerations may compromise the accuracy of future Hubble constant measurements, particularly in scenarios involving numerous detected GW events.

4 A STUDY OF BINARY BLACK HOLE POPULATIONS

This chapter is based on the work of a published article, in which I am the first author [36]. The chapter follows a similar overview as the original article. In this paper, we explore the relation between the spin magnitude and the mass of binary black holes using gravitational wave data from the third observing run of the LVK scientific collaboration. We developed new population models, that allow for different type of evolution of the spin parameters with respect to the source frame masses. We looked at the possibility of a linear evolution of the dimensionless spin magnitude with the mass, a mass transition between two distinct spin distributions and the superposition of two completely independent subpopulations, with different spin distributions.

4.1 INTRODUCTION

The astrophysical mechanisms governing the formation of stellar-mass binary black holes are still very uncertain, and numerous models and theories are still under debate today. As discussed in Sec. 1.3, BBHs are believed to emerge from three main formation channels: the isolated evolution of stellar binaries, the dynamical assembly, or the hierarchical merger [33, 41, 42, 153–156]. These channels can significantly influence the various intrinsic properties of BBH systems such as their masses, their spins magnitudes, orientations, and eccentricity. The impact of formation channels on the properties of BBHs can be so important that the presence of correlation between the spin and the mass of BBHs has been proposed as a smoking gun for the existence of such subpopulations of black holes. These subpopulations could originate from the distinct formation channels mentioned above. As an illustration, 1^{st}-generation BHs formed from isolated stellar binaries are expected to have masses $\leq 50\,M_{\odot}$, so before the PISN gap [115], and relatively small spins aligned to their orbital angular momentum. However, n^{th}-generation BHs born from previous mergers and binaries formed in dense stellar environments, are expected to have misaligned and higher spin magnitudes ($\chi \sim 0.7$) [157], induced by the pre-merger orbital angular momentum.

Since the first detection of GWs in 2015, emitted by the merger of two stellar-mass BHs [6], we have gained a new powerful probe to study these systems in great details. GWs have emerged as an invaluable means to directly examine and understand the astrophysical characteristics of BBH populations and the intricate influence of their environments on formation processes. The release of the largest catalogs of GW events to date, GWTC-2.1 and GWTC-3, in 2021 by the LVK collaboration has helped us further advanced our understanding of BBH populations [7, 27]. The latest catalog includes almost 90 detected coalescences, with a large majority of them being BBH events. The large amount of GW detection led to several recent population studies, utilizing both parametric and non-parametric Bayesian inferences [101, 136, 147, 158–161]. The latest population results using data from the GWTC-3 catalog find that: (i) The dimensionless spin magnitude distribution of BBHs favors lower values $\chi \leq 0.4$ [147]. (ii) There is no compelling evidence for the existence of a subpopulation of BBHs with zero spins [162–166]. (iii) The BBH spin component aligned to the orbital angular momentum does not significantly evolve with the mass of the compact object [147]. (iv) Higher spin seems to correlate with asymmetric mass binaries [167–169].

In this chapter, we focus on the potential correlations between the spin magnitude of BBHs and their masses, and the implications such correlations would mean for the formation channels of these systems. We study GW data from the third observing run, in particular the GWTC-2.1 and GWTC-3 catalogs, with new parametric models developed to explore the possibility of spin-mass correlation. The chapter is structured as follows. In Sec. 4.2, we present hints from the GWTC-3 catalog that may indicate the presence of a correlation between the spin and the mass of binary black holes.

We introduce in Sec. 4.3, three families of new phenomenological parametric models developed for population studies within a Bayesian inference framework. Sec. 4.4 presents the reconstructed spin distributions inferred with the population parameters. In Sec. 4.5, we validate our study through two additional analyses, a mock data challenge in which we simulate fake GW signals and an inference on real but modified GW signals. We conclude in Sec. 4.8.

4.2 FIRST HINTS FROM THE GWTC-3 CATALOG

For this analysis, we select a subset of 59 confident GW events from the third observing run, having a false alarm rate (IFAR) $IFAR \geq 1yr$. These events have been picked from the official catalog GWTC-2.1 and GWTC-3, released publicly in 2021 [1]. The estimated values of the spin magnitudes and source frame masses of these GW events are depicted in Fig. 4.1, with their respective errors obtained from the parameter estimation samples, provided by [27]. From Fig. 4.1, the data from the third observing run suggest the presence of a correlation between the dimensionless spin magnitude and the source mass of the BBHs selected. The correlation seems to indicate that low mass events are more susceptible to have low spin magnitudes, while more massive events have higher spin magnitudes. However, based only on this figure, this hint of the existence of a correlation in the data might not be real nor physical. Since the spin magnitudes and masses in Fig. 4.1 correspond to the ones of the detected GW events, this correlation may be induced solely by selection effects. For example, the current detectors could be less sensitive to systems having massive BHs and low spins, hence the lack of GW events in this part of the parameter space.

In this picture, we choose to explore the use of new phenomenological parametric population models (introduced in the following section), in order to model this correlation while also doing the deconvolution of the possible presence of selection biases. We estimate the parameters governing the population properties of BBHs, including their masses, CBC merger rate and spins, based on this set of 59 detected GW events. This analysis is conducted within the hierarchical Bayesian inference framework of ICAROGW. To estimate the selection effects within the hierarchical Bayesian inference, we utilize the public LVK set of detected injections [2], covering the entire parameter space of interest [147, 170]. Even if the analysis framework is the same as in Sec. 3, since we here focus on the interplay between population parameters and not the cosmology, we fix the cosmological parameters to the Planck 2015 measurements [138].

1 The C01:mixed samples we used can be found at https://zenodo.org/records/5546676.
2 The injection set we used can be found at https://doi.org/10.5281/zenodo.7890398.

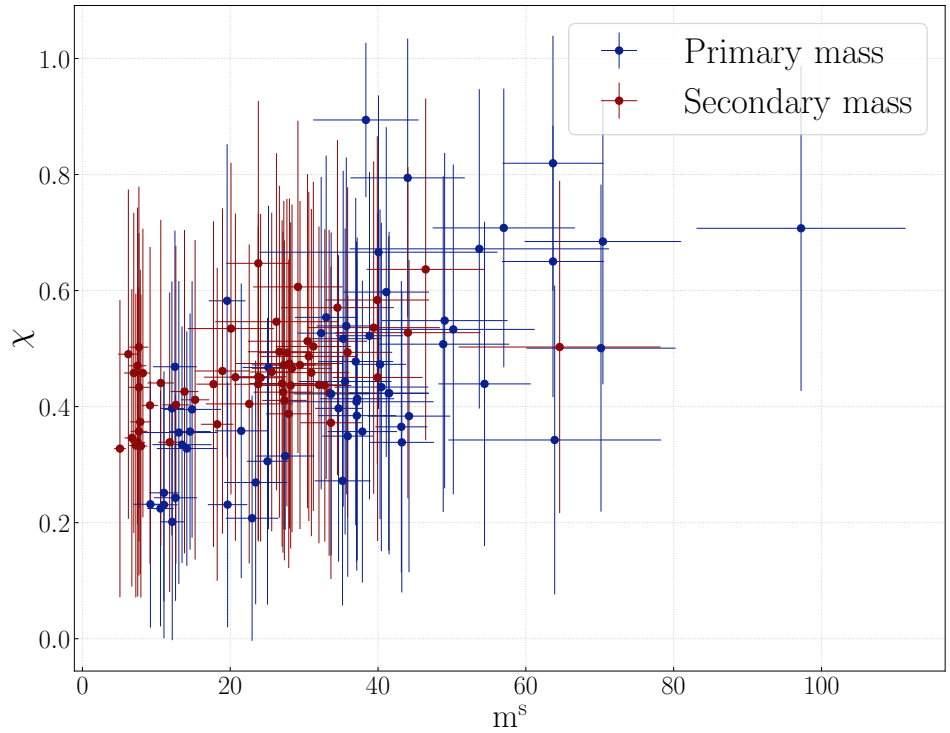

Figure 4.1: Scatter plot of the GW events used in the analysis. They correspond to the BBH GW events from GWTC-2.1 and GWTC-3 catalogs, selected with an $IFAR \geq 1yr$. The x-axis shows the source frame masses m^s and the y-axis displays the dimensionless spin magnitude χ. The error bars are the 1σ uncertainties of the official LVK parameter estimations C01:mixed samples. The blues crosses indicate the primary mass and spin parameters, while the red ones are for the secondary component.

4.3 NEW PHENOMENOLOGICAL MODELS FOR BBH SUB-POPULATIONS SEARCH

To characterize the interplay between mass and spin, we construct three families of parametric models that will be used within the hierarchical Bayesian inference framework of ICAROGW. The three families are namely, the EVOLVING models, the TRANSITION models and the MIXTURE models. The three classes of models have a singular way to include correlation between the spin and mass parameters, either via a direct evolution, a mass transition or the existence of two independent populations. Hence, each class of model has a specific parameterization for the BBH merger rate, from which the hierarchical likelihood is constructed. The following sections present the characteristics of each class of model, as implemented in our analysis pipeline.

4.3.1 The vanilla model

In order to assess the capabilities of the new phenomenological models, it is crucial to have a reference setup. The VANILLA model does not incorporate any sort of spin-mass correlation, and is built from the currently used parametric model within the LVK collaboration. This model will help us understand what a regular inference would produce, and also set a baseline result to perform model selection study through Bayes factor estimation.

For the VANILLA model, the BBH merger rate function is parameterized as

$$\frac{dN_{CBC}}{d\theta dz dt_s}(\Lambda) = R(z;\Lambda)\frac{dV_c}{dz}p_{pop}(\vec{m}_s|\Lambda)\pi(\vec{\chi},\cos\vec{\theta}|\Lambda), \tag{4.1}$$

where the vectors indicate each component of the two binary masses, spin magnitudes and tilt angles. This parameterization is identical to the regular "spectral siren" setup presented in Sec. 2.5. The sets of priors are listed in table. C.1. The rate function $\mathcal{R}(z;\Lambda)$ is modelled after the Madau&Dickinson star formation rate, following [143]

$$R(z;\Lambda) = R_0[1+(1+z_p)^{-\gamma-k}]\frac{(1+z)^\gamma}{1+\left(\frac{1+z}{1+z_p}\right)^{\gamma+k}}, \tag{4.2}$$

and the mass distribution is parameterized with the preferred GWTC-3 mass model, the POWER LAW + PEAK model such that

$$\begin{aligned}\pi(m_{1,s}|m_{min},m_{max},\alpha) &= (1-\lambda)\mathcal{P}(m_{1,s}|m_{min},m_{max},-\alpha) + \\ &\quad \lambda\mathcal{G}(m_{1,s}|\mu_g,\sigma)\,, \quad (0\le\lambda\le 1) &(4.3)\\ \pi(m_{2,s}|m_{min},m_{1,s},\beta) &= \mathcal{P}(m_{2,s}|m_{min},m_{1,s},\beta)\,, &(4.4)\end{aligned}$$

where $\mathcal{P}$ is the truncated powerlaw and $\mathcal{G}$ the extra Gaussian component. Finally, the spin distribution is chosen to be a DEFAULT spin model following Sec. 2.6:

$$\begin{aligned}\pi(\vec{\chi},\cos\vec{\theta}|\Lambda) &= Beta(\chi_1|\alpha,\beta)Beta(\chi_2|\alpha,\beta)\times \\ &\quad \pi(\cos\vec{\theta}|\xi,\sigma_t), &(4.5)\end{aligned}$$

with

$$\pi(\cos\theta_{1,2}|\zeta,\sigma_t) = \xi\mathcal{G}_{[-1,1]}(\cos\theta_{1,2}|1,\sigma_t) + \frac{1-\xi}{2}, \tag{4.6}$$

where $\mathcal{G}_{[-1,1]}(\cos\theta_i|1,\sigma_t)$ is a truncated Gaussian between -1 and 1.

With this parameterization, any correlation between the spin and the mass parameter if it exists, is not taken into account by the model. The mass, spin and redshift distributions are independent of one another, hence they would not be sensitive to correlations between them if present in the data.

4.3.2 The evolving model

Our first new class of models is called EVOLVING model, which looks for the existence of a linear correlation between the spin magnitude and the mass of the primary and secondary components in a binary. It describes the spin magnitude as a truncated Gaussian distribution between 0 and 1, with its mean and variance that evolve linearly with the value of the source frame mass. The distribution of the tilt angles follows the DEFAULT spin model as introduced in Sec. 2.6, where a fraction of the population has nearly aligned spins with the orbital angular momentum and the other fraction is isotropic. The new BBH merger rate function is given by

$$\frac{dN_{CBC}}{d\theta dz dt_s}(\Lambda) = R(z;\Lambda)\frac{dV_c}{dz}p_{pop}(\vec{m}_s|\Lambda)\pi(\vec{\chi},\cos\vec{\theta}|\vec{m},\Lambda), \tag{4.7}$$

The difference with the VANILLA model is hidden in the spin distribution, which is now conditioned on the value of the source mass. The total spin distribution is here factorized as

$$\pi(\vec{\chi},\cos\vec{\theta}|\vec{m},\Lambda) = \pi(\vec{\chi}|\vec{m},\Lambda)\pi(\cos\vec{\theta}|,\Lambda). \tag{4.8}$$

The spin magnitude $\pi(\vec{\chi}|\vec{m},\Lambda)$ has a direct dependency on the source mass, while the angular distribution $\pi(\cos\vec{\theta}|,\Lambda)$ is chosen independent as in the DEFAULT spin model. In addition, we assume that the two components for the spin magnitudes can be factorized such that

$$\pi(\vec{\chi}|\vec{m},\Lambda) = \pi(\chi_1|m_1,\Lambda)\pi(\chi_2|m_2,\Lambda). \tag{4.9}$$

The spin magnitude for this class of models is written as

$$\pi(\chi|m,\Lambda) = \mathcal{G}_{[0,1]}(\chi|\mu(m),\sigma(m)), \tag{4.10}$$

where $\mathcal{G}_{[0,1]}$ is the truncated Gaussian distribution, χ the spin magnitude and $(\mu(m),\sigma(m))$ are the mean and variance approximated with a linear first order Taylor's expansion such that:

$$\mu(m) = \mu_0 + \dot{\mu}m \tag{4.11}$$
$$\sigma(m) = \sigma_0 + \dot{\sigma}m. \tag{4.12}$$

For this model, the additional population parameters inferred are: μ_0, σ_0, $\dot{\mu}$ and $\dot{\sigma}$. The rate $\mathcal{R}(z;\Lambda)$ and mass $p_{pop}(\vec{m}_s|\Lambda)$ models for the EVOLVING class are kept to be a Madau&Dickinson and a POWER LAW + PEAK as in the VANILLA parameterization. The prior ranges are summarized in table. C.4, table. C.8 and table. C.6. The EVOLVING model is an extension of the VANILLA model, most of the intrinsic properties are modeled with the same distributions, except the dimensionless spin magnitude.

4.3.3 The transition model

The second class is named TRANSITION model. They incorporate a mass transition between two distinct populations with separate spin magnitude distributions. Within

this class, we construct two different models, the BETA TO BETA and the BETA TO GAUSSIAN. In the first case the two spin magnitude distributions are modeled by two Beta distributions, while the second by a Beta distribution to a Gaussian distribution. The transition between the two distribution is handled using a logistic function (window), whose midpoint and steepness are described by new parameters, added to the inference. The BBH merger rate function, mass distribution and merger function are, for both TRANSITION models, the same used for the EVOLVING model. However, the dimensionless spin magnitude is built differently.

The BETA TO GAUSSIAN model parameterizes the spin magnitude as a mass transition between a distribution described by a $Beta(\chi|\alpha,\beta)$, into a truncated Gaussian $\mathcal{G}_{[0,1]}(\chi|\mu,\sigma)$. The total spin magnitude distribution is given by

$$\begin{aligned}\pi(\chi|m,\Lambda) &= W(m;m_t,\delta m_t)B(\chi|\alpha,\beta)\\ &\quad +(1-W(m;m_t,\delta m_t))\mathcal{G}_{[0,1]}(\chi|\mu,\sigma),\end{aligned} \tag{4.13}$$

where $W(m;m_t,\delta m_t)$ is the logistic function which smoothly transitions from 1 to 0 as the source mass increases. The parameter m_t indicates the value in source mass at which the transition between the spin distributions happens, and δm_t is the steepness of this mass-transition. This function is defined as a regular window function such that

$$W(m;m_t,\delta m_t) = \frac{1}{1+e^{\frac{m-m_t}{\delta m_t}}}. \tag{4.14}$$

The BETA TO BETA is very similar to the BETA TO GAUSSIAN, the only difference is the second spin magnitude distribution in which the model transitions to, modeled as

$$\begin{aligned}\pi(\chi|m,\Lambda) &= W(m;m_t,\delta m_t)B(\chi|\alpha_1,\beta_1)\\ &\quad +(1-W(m;m_t,\delta m_t))B(\chi|\alpha_2,\beta_2),\end{aligned} \tag{4.15}$$

where the logistic function is the same as in Eq. 4.14. The prior ranges chosen for the Bayesian inference are listed in table. C.4, table. C.9, table. C.10 and table. C.6. This family of models aims to evaluate the presence of two formation channels populating the overall BBH spectrum.

4.3.4 The mixture model

The last family of models is called MIXTURE models. These models parameterize the overall population as the sum of two independent subpopulations [171, 172]. The two subpopulations, population 1 (Pop1) and population 2 (Pop2), are combined using a mixing fraction parameter. The CBC merger rate of the MIXTURE family is hence described by a superposition of two distinct CBC merger rates, one for each population:

$$\frac{dN_{CBC}}{d\theta dz dt_s} = \lambda_{pop}\frac{dN_{CBC}}{d\theta dz dt_s}(\Lambda^{Pop1}) + (1-\lambda_{pop})\frac{dN_{CBC}}{d\theta dz dt_s}(\Lambda^{Pop2}), \tag{4.16}$$

where λ_{pop} is the mixture fraction between the two subpopulations. The parameters Λ^{Pop1} and Λ^{Pop2} represent the hyperparameters for each subpopulation respectively, they are independent of one another. Each population has uncorrelated mass, spin and redshift distributions, in opposition to the EVOLVING and TRANSITION families. All models from the MIXTURE family have their CBC merger rate functions that follow the Madau&Dickinson parameterization and their spin distributions the DEFAULT spin model. Within the MIXTURE family, we construct three population models: the MIXTURE VANILLA, the MIXTURE PAIRED and the MIXTURE PEAK. The prior ranges for all three flavors of the MIXTURE are summarized in table. C.2 table. C.3, table. C.7 and table. C.5.

MIXTURE VANILLA: The population 1 of the MIXTURE VANILLA model is constructed with a POWER LAW + PEAK for the masses, a DEFAULT spin model for both the spin magnitudes and tilt angles, and Madau&Dickinson rate for the redshift distribution. The population 2, on the other hand, incorporates a simple truncated POWER LAW for the masses. The spin and rate distributions have the same parameterization as the population 1.

MIXTURE PEAK: For the MIXTURE PEAK model, the population 1 is modeled similarly as the population 1 of the MIXTURE VANILLA, with a DEFAULT spin model and the Madau & Dickinson rate. However, the primary and secondary masses of the population 1 are now given by a POWER LAW. For the population 2, the masses are parameterized with a Gaussian distribution of mean μ_g and standard deviation σ_g. We also add an extra constraint to ensure that the primary and secondary masses respect the restriction $m_1 > m_2$, as in the POWER LAW + PEAK model. The spins and redshift distributions are modeled with the same distribution as the population 1. The MIXTURE PEAK is built especially to look for a different spin distribution within the main feature of the mass spectrum. A recent study based on non-parametric inference, has found light support for the existence of a different spin distribution, located in this mass feature around $35\,M_\odot$, with a slightly different effective spin distribution [173]. With the MIXTURE PEAK, we aim to test this result.

MIXTURE PAIRED: The MIXTURE PAIRED is a variation of the MIXTURE VANILLA model. The distributions of the spins, redshift and primary masses of the two subpopulations are still parameterized as in the MIXTURE VANILLA, with the difference that the secondary masses are now forced to have the same distribution as the primary masses. Hence, if we model the distribution of the primary masses with a POWER LAW + PEAK, the secondary masses will also be distributed with a POWER LAW + PEAK. Mathematically, the mass distribution is generalized such that:

$$p_{pop}(m_1, m_2|\Lambda) \propto p_{pop}(m_1|\Lambda) p_{pop}(m_2|\Lambda) \left[\frac{m_2}{m_1}\right]^{\beta} \Theta(m_1 - m_2), \tag{4.17}$$

where $\Theta(m_1 - m_2)$ is the Heaviside step function forcing $m_1 > m_2$, and an extra weight which depends on the two masses is included. The secondary masses are no longer conditioned as a simple POWER LAW on the primary mass.

4.4 THE SPIN MAGNITUDE OF BLACK HOLES EVOLVES WITH THEIR MASSES

In this section, we use the new parametric population models described in Sec. 4.3 to infer the population properties of 59 GW events with $IFAR \geq 1yr$, taken from the third observing run of the LVK detector network and presented in Fig. 4.1.

4.4.1 Models selection

Table. 4.1 reports the Bayes factors and the maximum of log-likelihood ratios estimated between the baseline analysis (VANILLA model) for which the spins are not correlated to the masses and the new parametric population models presented in Sec. 4.3. We

Model	$\log_{10} \mathcal{B}$	$\log_{10} \mathcal{L}_{max}$
EVOLVING GAUSSIAN	-0.48	2.94
BETA TO GAUSSIAN	2.36	3.77
BETA TO BETA	2.55	3.91
MIXTURE VANILLA	2.78	4.66
MIXTURE PEAK	1.64	1.25
MIXTURE PAIRED	3.78	5.99

Table 4.1: Base 10 logarithm of the Bayes factors (second column) and the logarithm of the maximum likelihood ratio (third column), for the six models discussed in the section above, compared to the reference model. The reference model is the vanilla non evolving analysis.

recall that the baseline analysis inference has been done using a POWER LAW + PEAK mass model, a DEFAULT spin model and the Madau & Dickinson rate model, following state-of-the-art literature [147].

The log_{10} Bayes factors reveal that all the models that incorporate some spin-mass correlation, either as a mass transition between two subpopulations or via superposition of two subpopulations are strongly preferred against the non-evolving model, and this preference is still true despite the increased dimensionality of the fit. The new models that parameterized a spin-mass correlation have a significantly higher number of population parameters than the non-evolving one. The value of the Bayes factors hence highlight that our models are truly favored by the data, the maximum log-likelihood ratios also support this finding. For the EVOLVING GAUSSIAN model, which parameterize the spin-mass interplay as a continuous and linear evolution, the close to zero Bayes factor indicated that this model is not preferred nor excluded, the current GW data at our disposable are not sufficient to decide.

To understand if the values for the Bayes factor can be trusted, we performed several tests which are discussed in more details in Sec. 4.5. First, we made sure that the

values of the Bayes factor are actually driven by the inclusion of a spin-mass relation in our models, we found that analysis performed with only the mass information induce Bayes factor values that can not discriminate between the models. Secondly, we verified that our spin-mass models can be confidently excluded compared to the baseline analysis if the GW data present no spin-mass correlation. To do so we simulated a population of BBHs without any spin-mass correlation (see Sec. 4.5.1). Lastly, we also found that the Bayes factor become inconclusive when blinding real data to possible spin-mass correlation (see Sec. 4.5.2).

4.4.2 Evolution as a transition in mass between subpopulations

We divide the analysis in two main sections, the first one being the results of the EVOLVING and TRANSITION models, and the second one the results from the MIXTURE models. Fig. 4.2 shows the reconstructed dimensionless spin magnitude distributions for the EVOLVING GAUSSIAN, BETA TO GAUSSIAN and BETA TO BETA models. The median values of the distributions are displayed with plain lines and the 90% and 98% C.L. with the colored contours. We binned the reconstruction of the spin magnitude in ten source mass slices, ranging from $10\,M_\odot$ to $100\,M_\odot$, in order to visualize the presence of any form of evolution. We find that all the models reconstruct a transition from a population described by a low-spin magnitude distribution to a population described by a higher and wider spin magnitude distribution. All three models infer a lowly spinning population (around $\chi \sim 0.2$) of compact objects at low masses transitioning around $40 - 50 M_\odot$ to another population for which the spin magnitude is surely higher than $\chi = 0.5$. We also note that the high mass distribution of the spin magnitude is less constrained than the low mass part, this is due to the lack of GW events in this mass range. However, even if the distribution presents larger uncertainties, the evolution of the spin magnitude is still significant.

Fig. 4.3 displays the corner plot of the population parameters governing the evolution of the spin magnitude for this model. Either the mean of the Gaussian peak does not evolve with the mass, but the width of the distribution increases. This situation is given by values of $\dot{\mu}$ close to zero and non-zero value of $\dot{\sigma}$. Or the mean of the Gaussian is shifting as the mass increases ($\dot{\mu} > 0$), and its width stays nearly fixed with $\dot{\sigma} \sim 0$. The EVOLVING GAUSSIAN model describes the mass-spin evolution in two possible ways that can not be truly disentangled from the current data. The important result is that, either way, the distribution of the spin magnitude is changing as the mass range increases.

The results from the BETA TO GAUSSIAN and BETA TO BETA models that parameterize a transition between two subpopulations (middle and right column of Fig. 4.2) indicate the existence of a transition between two distinct spin distributions, happening around $40 - 50\,M_\odot$. Fig. 4.4 shows the corner plot of the two estimated population parameters sensitive to the mass transition for the BETA TO BETA and BETA TO GAUSSIAN

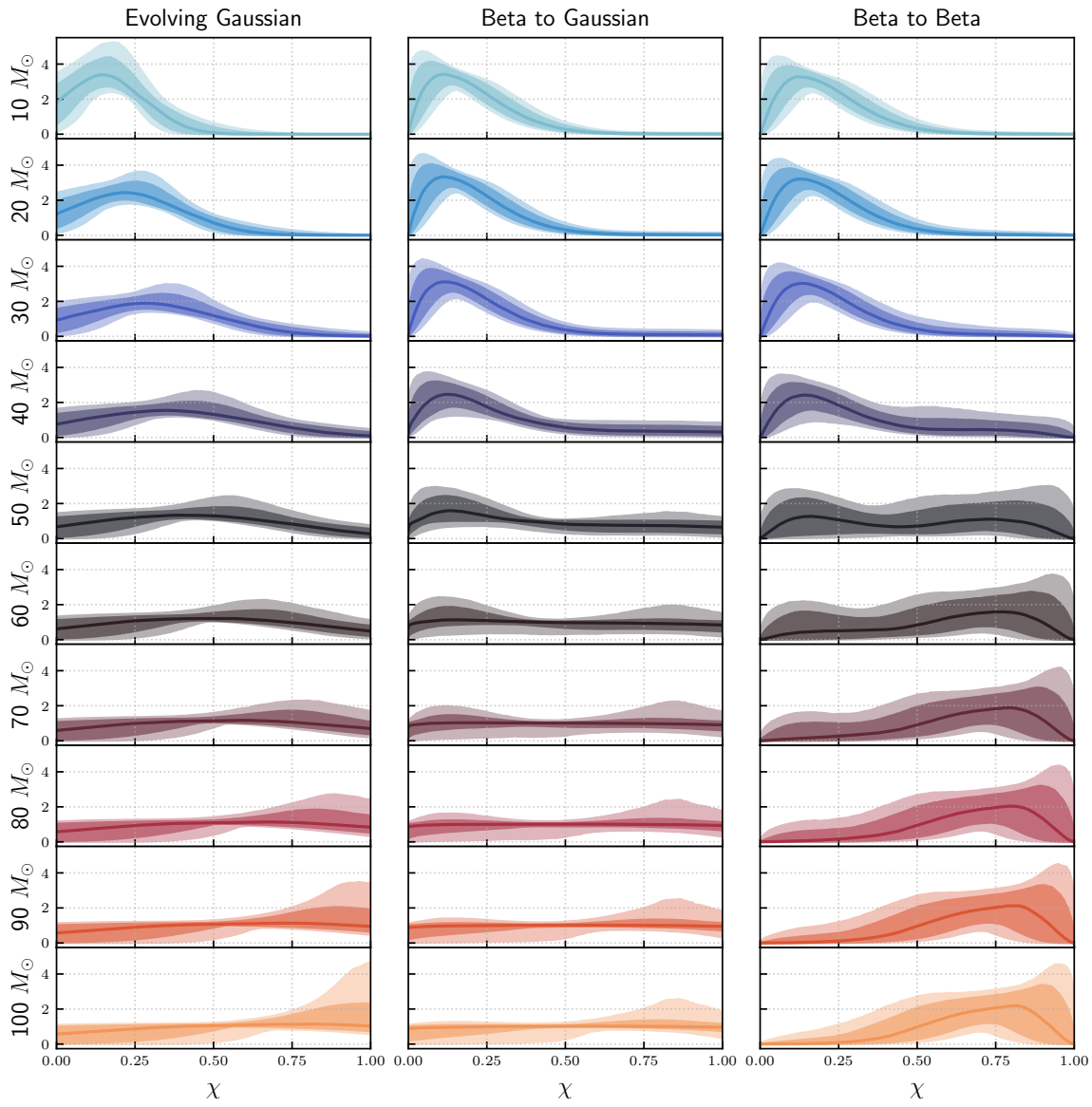

Figure 4.2: Joy plot of the probability density functions of the dimensionless spin magnitudes χ, reconstructed from the population inference on the 59 BBHs with $IFAR \geq 1yr$ from the GWTC-2.1 and GWTC-3 catalogs. The dimensionless spin magnitudes were reconstructed using the EVOLVING GAUSSIAN model (left column), the BETA TO GAUSSIAN model (middle column) and the BETA TO BETA model (right column). Each row corresponds to a slice in source frame mass, from $10\,M_\odot$ to $100\,M_\odot$, in order to highlight the spin magnitude evolution and transitions across different mass ranges. The plain lines are the median value of the spin magnitudes and the colored contours are the 90% and 98% C.L.

models. The density of the contours for both models tends to be more concentrated at low m_t and low δm_t. The reconstruction of the logistic function $W(m; m_t, \delta m_t)$ in Fig. 4.5 also support this fast and steep transition between the two spin magnitude distributions for both models. From the 59 GW events used in the Bayesian analysis, there is a preference for a steep transition around $40\,M_\odot$, rather than a smoother

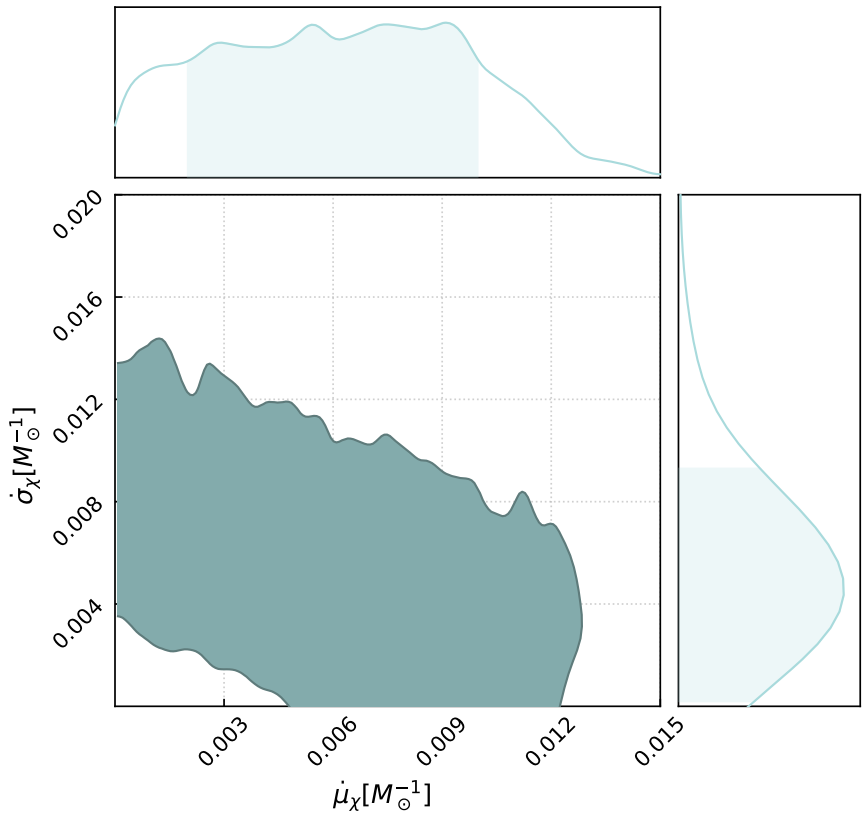

Figure 4.3: Corner plot of the population parameters $\dot{\mu}$ and $\dot{\sigma}$ of the EVOLVING GAUSSIAN model, which represent the slope parameters of the evolving spin distribution of μ and σ. They are obtained from the population inference of 59 GW events from the GWTC-2.1 and GWTC-3 catalogs. The contours are the 90% C.L.

and wider transition at higher masses $\sim 55\,M_\odot$. Even if the other transition scenario can not be completely ruled out, both models indicate a transition in spin magnitudes around $40 - 55\,M_\odot$ from slower to more rapidly spinning objects.

Beside the population parameter governing the spin magnitude distribution, we also estimate all the other population parameters, for the masses, the CBC merger rate and the tilt angles of the spins. We do that for the following reason, from Sec. 3 we saw that Bayesian inference with parametric models can be subject to biases if some parameters are wrongly estimated. The spin-mass correlation that we find in Fig. 4.2 may be induced by a wrong reconstruction of the BBH mass spectrum with our models for example. With all three models, we reconstruct a mass distribution in excellent agreement with our baseline analysis and with the one inferred in [147], that uses uncorrelated spin-mass models. The reconstructed mass spectrum with the different analysis is shown in Fig. C.2 of App. C.1. In addition, the reconstructed CBC merger rate and angular spin distribution of the tilt angle, for all three models, are also found to be in good agreement with our baseline analysis.

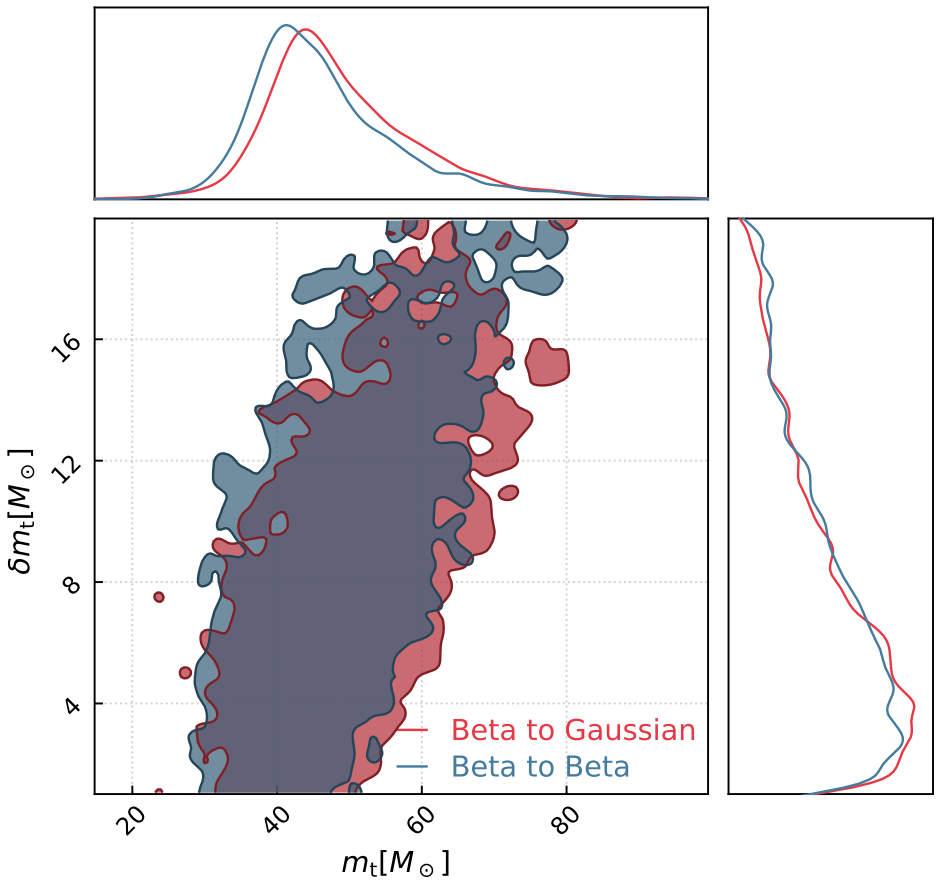

Figure 4.4: Corner plot of the mass transition point m_t and the transition steepness δm_t population parameters from the BETA TO BETA model in purple and BETA TO GAUSSIAN model in blue. These results have been obtained from the population inference of 59 GW events from the GWTC-2.1 and GWTC-3 catalogs.

In summary, both the EVOLVING and TRANSITION models used in the Bayesian inference found significant support for an evolution of the spin magnitude distribution. Since the TRANSITION models are strongly preferred by the Bayes factor, we rather describe this spin-mass correlation as a transition between two spin distributions: one at low masses with low spin magnitudes and a second at higher masses with spin magnitude higher than 0.5. We find evidence that the transition should happen quickly, around $45\,M_\odot$.

4.4.3 Evolution as the mixing of two independent subpopulations

The results from the previous section support a spin-mass interplay induced by a transition happening around $40-55\,M_\odot$ between two populations with different spin mag-

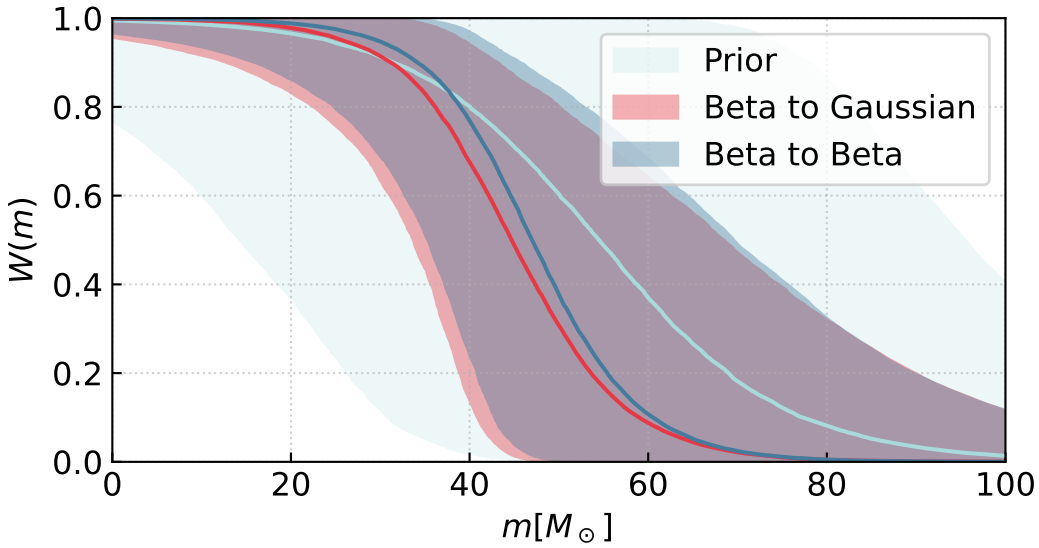

Figure 4.5: Reconstruction of the window function $W(m; m_t, \delta m_t)$ as a function of the source mass, obtained from the population inference of 59 GW events with $IFAR \geq 1yr$ from the GWTC-2.1 and GWTC-3 catalogs with the BETA TO BETA (pink) and BETA TO GAUSSIAN (blue) models. The light pink denotes is the prior used for the inference.

nitude distributions. From the Bayesian inference with the MIXTURE models, we want to understand if this evolution could be consistent with the overlap of two independent subpopulations of BBHs with separate and uncorrelated mass, spins and redshift distributions. These subpopulations could bring new evidence of the existence and impact of several formation channels for BBHs systems.

Fig. 4.6 depicts the reconstructed spin magnitude and inclination angles inferred from the three flavors of the MIXTURE models. We show the individual distribution estimated for the population 1 and 2, as well as the combined population. For the MIXTURE VANILLA and MIXTURE PAIRED inferences, we set the primary population to describe the BBHs with masses $\leq 40 - 60\,M_{\odot}$, based on the results obtained the EVOLVING and TRANSITION models. For the MIXTURE PEAK on the other hand, the secondary population is set to be a Gaussian in the $20 - 50\,M_{\odot}$ region, as motivated by [173].

The common result we find among all MIXTURE models is that we recover a primary population of BBHs (Pop1), characterized by low masses and small values of spin magnitude, peaking around $\chi \sim 0.1$. This population 1 represents a large fraction of the total population. The second population (Pop2) is found at higher masses and supports distributions of the dimensionless spin magnitude peaking around $\chi \sim 0.7$. As shown in Fig. 4.7, for all three MIXTURE models the data indicate that the population 1 accounts for almost 98% of the entire population. In terms of model selection, the MIXTURE PAIRED and the MIXTURE VANILLA seems to be favored according compared to the MIXTURE PEAK to the Bayes factor. In other words, the distributions of the GW data are

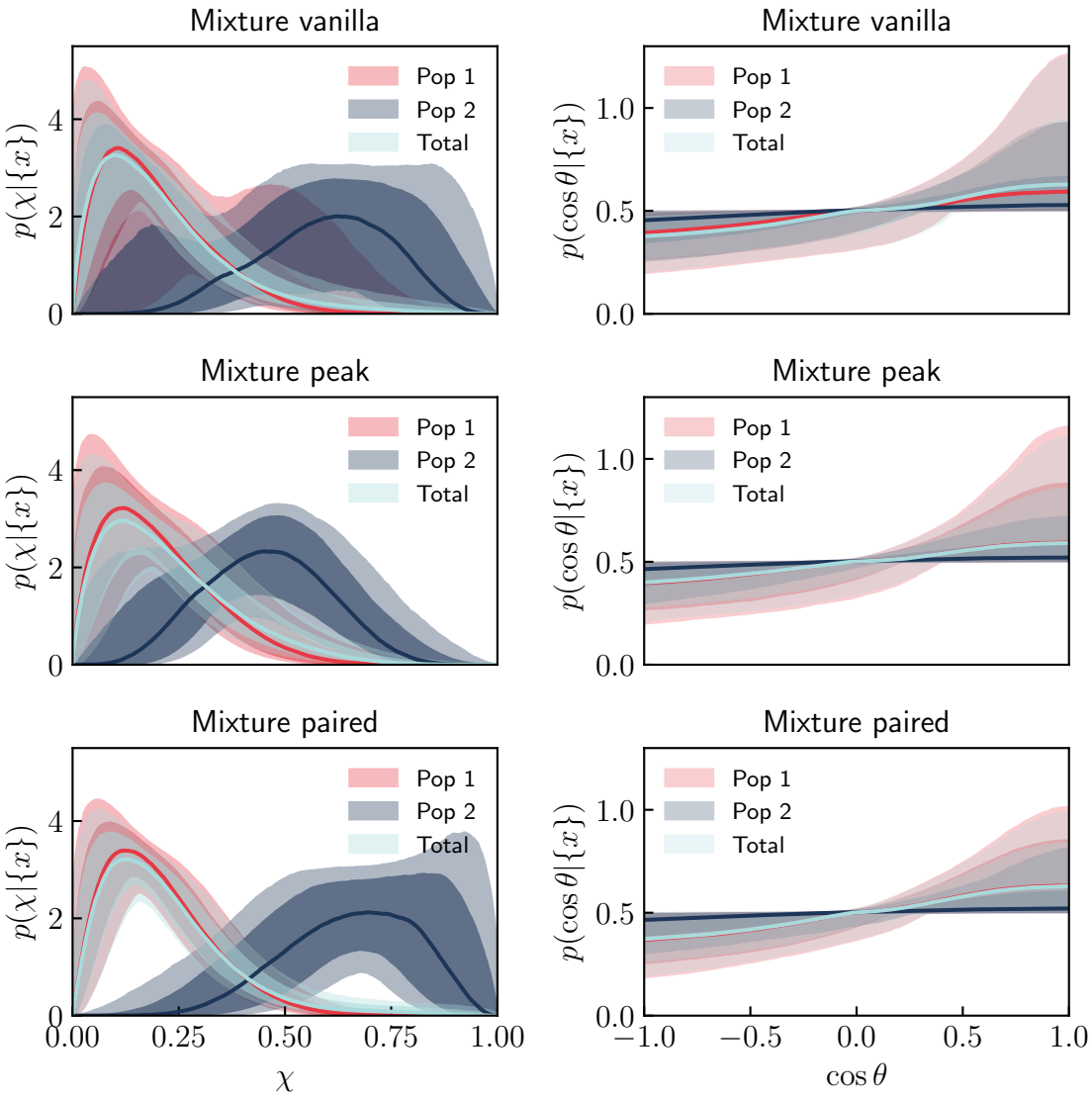

Figure 4.6: Reconstructed spectra of the spin magnitude distributions χ and the cosine of the tilt angle *cost* obtained with all three flavors of the MIXTURE models, namely MIXTURE VANILLA (top row), MIXTURE PEAK (middle row) and MIXTURE PEAK (bottom row). The population inference has been done on the 59 GW events with $IFAR \geq 1yr$ from the GWTC-2.1 and GWTC-3 catalogs. The red curves (Pop1) and blue curves (Pop2) are the estimated distributions for each population. The light blue curve (Total) shows the combined population when Pop1 and Pop2 are summed using the mixing fraction parameter λ_{pop}. The colored contours are the 90% and 98% C.L.

better explained with a secondary population with a different spin magnitude located above $40\,M_\odot$, rather than solely in the excess of BBHs around $35\,M_\odot$.

For each flavor of the MIXTURE model family, we also reconstruct the spin's tilt angle distribution with respect to the orbital angular momentum. We find that the inferred distributions for the population 1 (low-mass) weakly prefer spins aligned with the

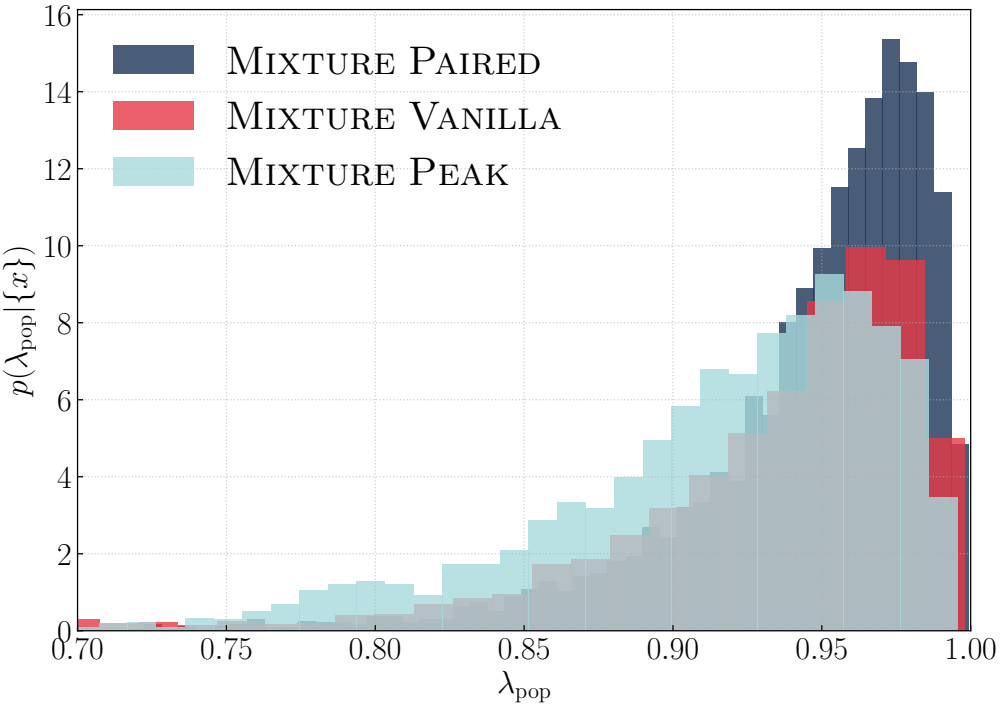

Figure 4.7: Histograms of the estimated posterior distribution of the mixing parameter λ_{pop} for the three flavors of the MIXTURE model family. The inferred value is obtained on 59 GW events from the GWTC-2.1 and GWTC-3 catalogs. The MIXTURE PAIRED is in blue, the MIXTURE VANILLA in red and the MIXTURE PEAK in light blue.

orbital angular momentum ($cos\theta \sim 1$), while the population 2(high-mass) weakly prefers a more isotropic distribution of the tilt angles. However, the uncertainties of the spin's tilt angles are for now still very uncertain to draw any robust conclusion about their distributions.

The mixing parameter λ_{pop}, which denotes the proportion of the population 1 with respect to the population 2, can exhibit a slight bi-modality with the minimum mass of the population 2 m_{min}^{pop2}, in particular when the second population is allowed to arrive down to $2\,M_\odot$. Fig. 4.8 shows the inferred posterior for these two parameters and their 2D correlations. This bi-modality, only visible with the MIXTURE VANILLA model, suggests that there is some support for a lower value of the mixing parameter combined with a minimum mass for the second population close to $3-4\,M_\odot$. This minimum mass corresponds to the one of the reconstructed first population. Nonetheless, the mode described by $\lambda_{pop} \sim 0.98$ and $m_{min}^{pop2} \sim 30 M_\odot$ is strongly preferred by the analysis, hence favoring the hypothesis that a subpopulation of BBHs with a different spin distribution could arise after $30-35\,M_\odot$.

In a similar fashion as for the EVOLVING and TRANSITION analysis, we also reconstruct the mass and CBC merger rate spectra. We find that they are in perfect agreement

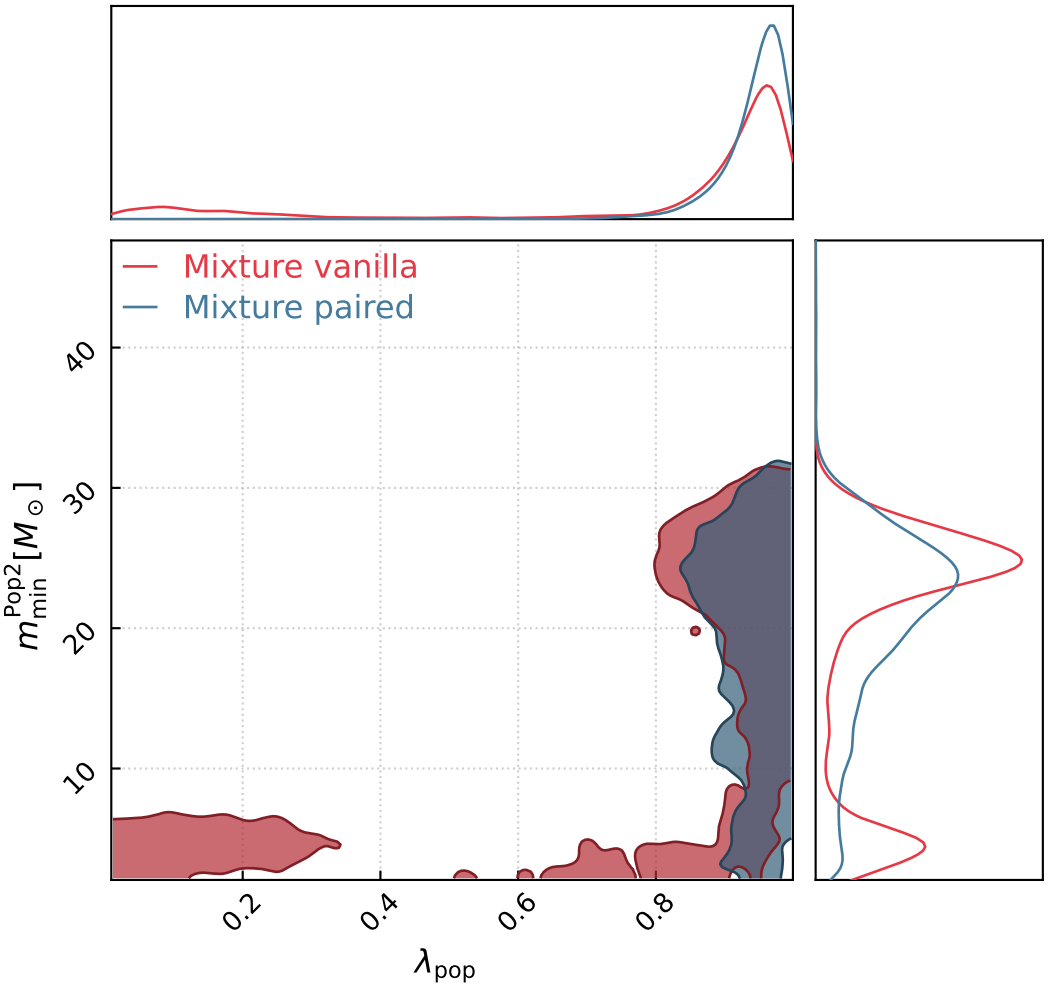

Figure 4.8: Corner plot of the mixing parameter λ_{pop} and the minimum mass of the second population m_{min}^{pop2}, inferred with the MIXTURE VANILLA (red) and MIXTURE PAIRED (blue) models from the population inference of 59 GW events with $IFAR \geq 1yr$ from the GWTC-2.1 and GWTC-3 catalogs. The colored contours are the 90% C.L. inferred.

with the reconstruction of the baseline analysis, with the distributions obtained with the EVOLVING and TRANSITION models, and with other independent non-evolving population inferences. The reconstructed mass spectra are shown in Fig. C.2 in App. C.1. Fig. 4.9 displays the estimated posterior distributions of γ, a CBC merger rate population parameter. All the marginalized posteriors of γ seem to have a similar trend, when comparing between the population 1 and population 2, but also comparing with the baseline analysis. In other words, if the two subpopulations corresponds to distinct formation channels of BBHs, then their time-delay distributions between the BBH formation and merger should be similar.

As an extra validation step of our results with the MIXTURE models, we run the same population inference, but removing the spin distributions from the analysis (estimating the masses and CBC merger rate only). The new Bayes factors between the base-

line analysis and the three MIXTURE models without spin are now close to 1- this shows the importance of the inclusion of spin populations models to disentangle between the two populations. The sole inference of the mass and rate spectra does not bring enough constraining power to the analysis to separate two mixing population of BBHs. Moreover, when removing the spin parameters from the MIXTURE population inference, the mixing parameter λ_{pop} is also less constrained and now in agreement with 1; i.e. only one population is supported by the data.

In short, the MIXTURE model inferences are strongly favored compared to canonical non-evolving analysis. With these models, we find evidence of the presence of two BBHs subpopulations in the GW data: one at low masses, characterized by a spin magnitude peaking at $\chi \sim 0.1$ and a second one at higher masses $\geq 35\,M_\odot$ with a spin magnitude distribution significantly larger ($\chi \sim 0.7$). The second population of BBHs, may originate from a different formation channel than the first one, and represent close to 2% of the total population.

4.4.4 Discussion

This analysis reveals novel compelling evidences for the existence of a correlation between the BBHs spin magnitudes and their mass. This trend persists across all our phenomenological population models and can either be understood as a mass-dependent transition between two spin distributions or as an overlap of two independent BBHs subpopulations with uncorrelated spins, masses and redshift distribution. Moreover, all models infer a lowly spinning population of BBHs at low masses and another population of BBHs at higher masses, with a different spin magnitude distribution characterized with more rapidly spinning objects.

Previous studies have already tried to inspect a possible spin-mass relation at the population level of BBHs using GW detections [147, 167, 174–178]. In [147, 178], it is argued that the absolute value of the spin projection $|s_z|$ over the orbital angular momentum does not significantly evolve with the chirp mass $\mathcal{M}_c$ of the system. This result is not in contrast with our findings about the spin magnitude evolution, as the spin tilt angles are poorly constrained in terms of masses. Therefore, the spin projection over the orbital angular momentum shows no particular correlation with the mass, although the spin magnitude can display a significant correlation. Fig. 4.10 shows the reconstructed distribution of the spin magnitudes aligned with the orbital angular momentum $|s_z|$, as a function of the detector chirp mass. We display the 50% and 90% credible upper bounds for $|s_z|$. We find that the upper bounds seem to evolve to a more rapidly regime for the BBHs, above $40\,M_\odot$, evolution which is consistent with our previous results that support a change in the spin distribution around the same mass range. The main difference between our result and [147] is that while they explain it as being due to a weak constraint on the spin for massive BHs, here we obtain this correlation as a consequence of a spin-mass interplay. In fact, we do not observed

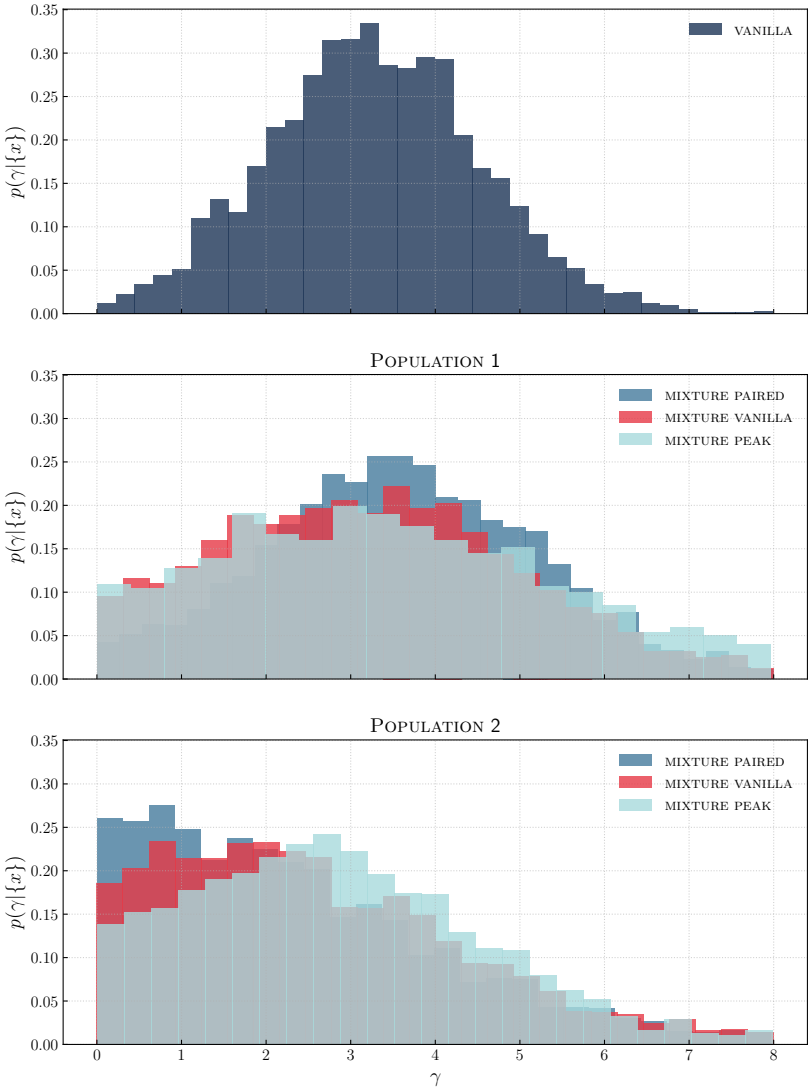

Figure 4.9: Histograms of the inferred posterior distributions of the CBC merger rate parameter γ, estimated with four different inferences on 59 GW events with $IFAR \geq 1yr$ from the GWTC-2.1 and GWTC-3 catalogs. Top panel: VANILLA analysis with a mono and non-evolving population model. Middle panel: Posteriors of γ for the population 1 CBC merger rate with three flavors of the MIXTURE model. Bottom panel: Posteriors of γ for the population 2 CBC merger rate with three flavors of the MIXTURE model.

such evolution of $|s_z|$ if we remove the spin mass-relation from the data. The later test is presented in more details in Sec. 4.5.2.

These results obtained with the MIXTURE PEAK model are consistent with the findings of [173]. Unlike our work, which focuses on spin magnitude, they examined the effective spin χ_{eff} and precession spin χ_p parameters. They employed a binned non-parametric inference scheme on a similar set of GW detections and suggested the existence of a subpopulation of BBHs with a different spin distribution in the mass range $30 - 50, M_\odot$. Although our model is not the preferred one, our results provide mild evidence supporting this scenario.

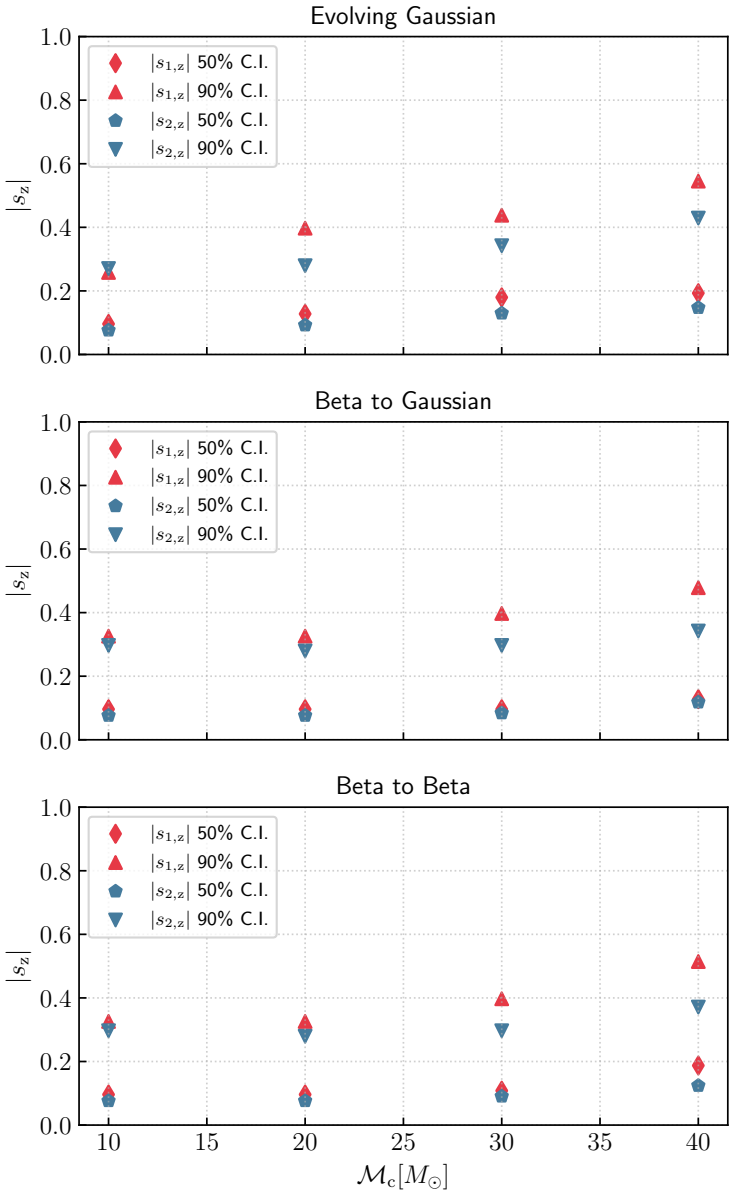

Figure 4.10: Scatter plot of the evolution of the aligned component of the spin magnitude $|s_z|$ with respect to the detector chirp mass $\mathcal{M}_c$, obtained from the population inference of 59 GW events with $IFAR \geq 1yr$ from the GWTC-2.1 and GWTC-3 catalogs, with three different population models. Top panel: EVOLVING GAUSSIAN model. Middle panel: BETA TO GAUSSIAN model. Bottom panel: BETA TO BETA model.

In [179], they argue that the BBHs present in GWTC-2 (second observing run) may arise from a population composed of $1^{st} - 1^{st}$ generation (both being borne from star collapse processes), $1^{st} - 2^{nd}$ generation and $2^{nd} - 2^{nd}$ generation of BHs. The mass and spin distributions of BBHs for the $1^{st} - 1^{st}$ generation of binaries are fitted using phenomenological models like the ones employed in this analysis, while the $1^{st} - 2^{nd}$ and $2^{nd} - 2^{nd}$ mass and spin distributions are estimated with transfer functions (defined in [162]) from the $1^{st} - 1^{st}$ generation of binaries. In our study of BBHs spin magnitude evolution, we go beyond the use of a transfer function calibrated on hierarchical formation channel only, we demonstrate that there is a correlation between spins and masses in BBH systems which is possibly introduced by the transition in mass between two subpopulations described by different spin distributions. Nevertheless, the general results that we obtained do not go against the conclusions of [179] that argues about the possible presence of $1^{st} - 2^{nd}$ and $2^{nd} - 2^{nd}$ generations of binaries. In fact, these types of mergers belong to the second population inferred by our MIXTURE MODELS, which encompasses both $1^{st} - 2^{nd}$ and $2^{nd} - 2^{nd}$ generations of binaries.

4.4.5 Astrophysical implications

In the light of these findings, we discuss the astrophysical interpretation and implications of our results regarding the BBHs formation channels. One of the most accepted theories for compact object's formation is that BHs from isolated stellar binaries can not be formed beyond $[45 - 60]\,M_{\odot}$. This mass scale is identified as the lower edge of the PISN gap [115, 152, 180, 181]. In this picture, the PISN mass scale would mark a transition between a population of 1^{st} generation BHs formed by their stellar progenitors to a population of n^{th} generation BHs dynamically assembled into binaries [33, 179] in dense stellar environments. The population of 1^{st} generation BHs is predicted to have relatively small spins, aligned with the orbital angular momentum, due to the various astrophysical processes expected to happen during the stellar binary evolution [33]. The population of n^{th} generation BHs, on the other hand, is expected to have spins magnitude centered around $\chi = 0.7$ (inherited from the pre-merger binary) and nearly isotropically distributed tilt angles [43–45, 182]. In addition, according to the latest BBH synthesis simulations, $1^{st} - 1^{st}$ generation are expected to form 97.5%-98% of the total population of BBHs, while n^{th} generations the rest, all formation channels combined [124, 183]. In this study, we do not include the contributions of BHs formed by population III stars [184], which are likely to be located at very high redshifts (not accessible with the current detectors sensitivities) and compose a tiny fraction of the overall astrophysical population.

These results bring evidences for this scenario. The transition between the subpopulations of BBHs that we observe around $40 - 55\,M_{\odot}$ could be linked to the PISN mass scale. The lower mass population displays a clear preference for low spin magnitude ($\chi \sim 0.1$) as expected for BHs formed in isolated stellar binaries. Although,

the spin distributions for the higher mass population is not as strongly constrained as the low mass one, we know that it is surely different from the spin distribution of the low mass population, and we find new evidence from the MIXTURE models that the spin magnitude of this second population supports values around $\chi \sim 0.7$. Moreover, the MIXTURE models infer that the high-mass (n^{th} generation BHs) population should compose only 2% of the astrophysical BBHs.

Another astrophysically relevant result highlighted by the MIXTURE model analysis is that the BBH merger rate as a function of redshift increases in the same way for the low and high mass subpopulations (see Fig. 4.9). If we identify the former population as 1^{st} generation BHs and the second as n^{th} generation BHs, this result would imply that the time scales over which hierarchical mergers happen are cosmologically small.

The correlations we observe between the spin magnitudes and masses are evidence of the existence of two subpopulations of BBHs, transitioning around $40 - 55\,M_{\odot}$ and described by different spin distributions. These findings provide support for the existence of n^{th} generation BBH mergers originating from the hierarchical merger formation channel. However, a definitive evidence should be reached with a better reconstruction of the spin tilt distributions that could be obtained with future GW observations.

4.5 A MOCK DATA CHALLENGE

As mentioned in the previous section, beside the main analysis on the 59 GW events from the GWTC-2.1 and GWTC-3 catalogs, we additionally tested and validated our new parametric population models and the population inference robustness. In particular, we performed two side analysis: a mock data challenge (MDC) on simulated GW detections and a "*blurred*" inference on real GW events.

The aim of the MDC is to understand how our new population models that incorporate spin-mass interplay react to a population of BBHs that does not actually include any spin-mass correlation. At the same time, the MDC analysis allows us to compute new Bayes factors and check that our models are correctly excluded with respect to simpler models that are "sufficient" when no spin-mass correlation is present in the data. We simulated sets of 50 detected GW events, characterized by their detector frame masses, luminosity distances and spin parameters. The simulated populations of BBHs are generated without spin-mass correlation. These events are drawn from the injection set used to estimate the selection effects in the main analysis of this chapter, and so they are representative of the sensitivity reached by the detector for the GWTC-2.1 and GWTC-3 catalogs. In the same fashion as in Sec. 3, we assume that the detector frame parameters of the sources are perfectly measured during the detection process (masses, luminosity distances, spins), in order to maximize the precision on the population parameters inferred with the different models.

Mock Data Challenge

Model	$\log_{10}\mathcal{B}$	$\log_{10}\mathcal{L}_{max}$
EVOLVING GAUSSIAN	-6.07	-1.46
BETA TO GAUSSIAN	-0.61	1.86
BETA TO BETA	-0.21	0.43
MIXTURE VANILLA	-1.36	2.53
MIXTURE PEAK	-3.69	-0.88
MIXTURE PAIRED	-3.26	3.16

Table 4.2: Base 10 logarithm of the Bayes factors (second column) and the logarithm of the maximum likelihood ratio (third column), for the six models discussed in this analysis, compared to the baseline non-evolving population model.

4.5.1 Models selection

The set of detected GW events are simulated following a Madau & Dickinson BBH rate model, a mass distribution based on a POWER LAW + PEAK model and a spin distribution with a preference for lowly spinning BBHs, nearly aligned with the orbital angular momentum. All simulated distributions—masses, spins, and merger rates—are represented by black dashed lines in the figures shown below.

Table. 4.2 report the summary of the Bayes factors and maximum likelihood ratios obtained from the population inference on the simulated GW events with our six parametric models which incorporate spin-mass interplay. Again, the model assessment is done comparing these models with a baseline analysis with a non-evolving population model. From the Bayes factors obtained for the MDC analysis, it is clear that the canonical non-evolving model is always strongly preferred for the MIXTURE and the EVOLVING models. With respect to the TRANSITION models, the baseline model is slightly preferred. This result demonstrates that the spin-mass correlation which is picked up in the main analysis, is the one driving the model selection shown with the set of real GW events. Moreover, this test shows that our models can confidently be excluded by Bayes factor comparison, if the spin-mass correlation is not present in the data.

4.5.2 Considerations on EVOLVING and TRANSITION models

Fig. 4.11 shows the reconstruction of the spin magnitudes obtained with the population inference using the EVOLVING GAUSSIAN, the BETA TO BETA and BETA TO GAUSSIAN models on the simulated GW data, which is represented with the black dashed line. The inference spin magnitude shows excellent agreement with the simulated population, with all three models. The deviations from the injected population in the posterior predictive checks above the mass slice of $70\,M_\odot$ are not to be worried about.

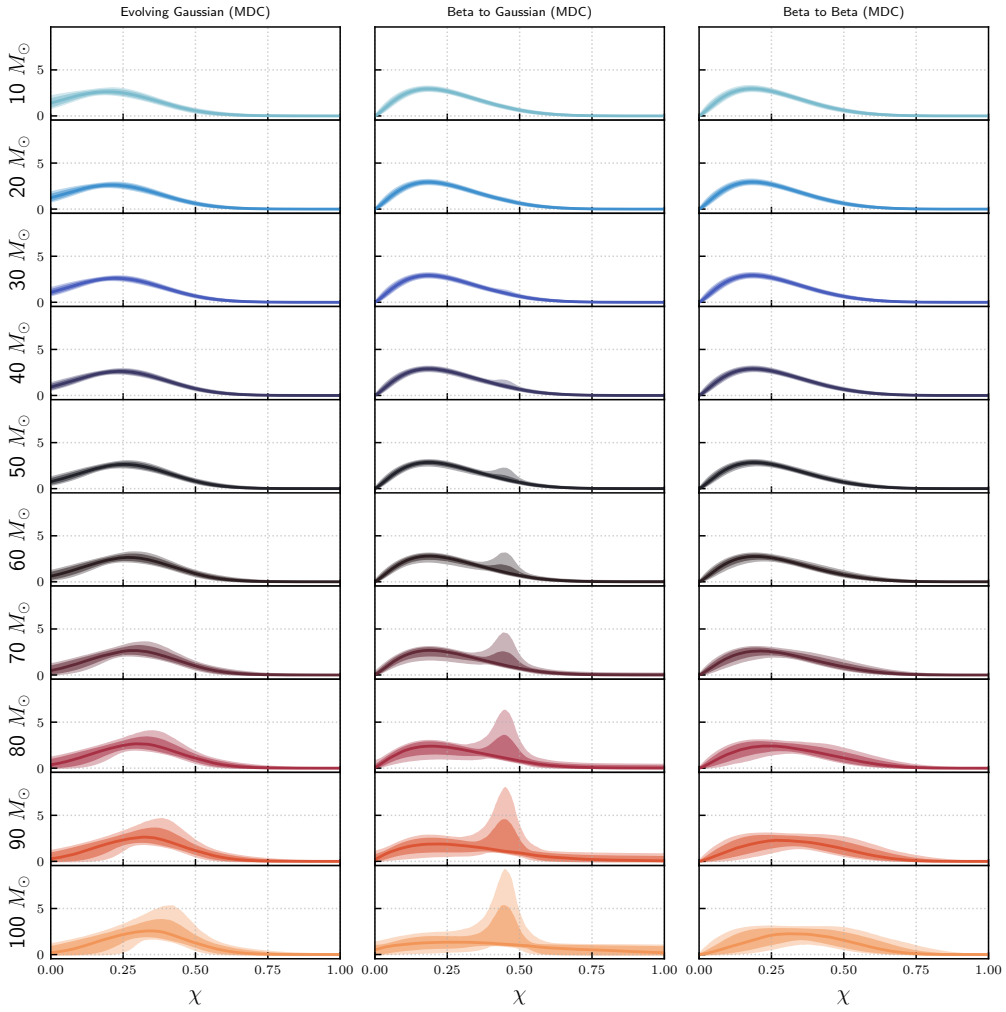

Figure 4.11: Joy plot of the probability density functions of the dimensionless spin magnitudes χ, reconstructed from the population inference on simulated GW data (MDC). The dimensionless spin magnitudes were reconstructed using the EVOLVING GAUSSIAN model (left column), the BETA TO GAUSSIAN model (middle column) and the BETA TO BETA model (right column). Each row corresponds to a slice in source frame mass, from $10\,M_\odot$ to $100\,M_\odot$, in order to highlight the spin magnitude evolution and transitions across different mass ranges. The black dashed lines are the simulated population from the MDC. The plain lines are the median value of the spin magnitudes and the colored contours are the 90% and 98% C.L.

They are in fact natural and due to the choice of prior ranges for the population parameters that govern the mass transition. For the BETA TO BETA and BETA TO GAUSSIAN inferences, the prior ranges on m_t and δm_t in charge of the window function for the mass transition are the limiting parameters from which the deviations arise. Since

there is no transition between two spin distributions in the simulated GW events, the inferred posterior of m_t does not converge within its prior range, but rather rails on its upper limit ($100\,M_\odot$). The upper plot of Fig. 4.12 shows the posterior predictive check for the reconstructed window function. As a consequence, the parameters of the window function being forced to converge on the upper bound of the prior range make the posterior predictive checks find "unwanted" support for a mass transition below $100\,M_\odot$. The behavior of the posterior predictive check with the EVOLVING GAUSSIAN is similarly explained. Considering the deviations of the reconstruction of the spin magnitude with respect to the simulated population, it is clear that this has no impact of the results we found with the real GW events. The marginalized posteriors of the population parameters of interest are correctly contained within their prior ranges, which avoid any behavior that could be interpreted as an astrophysical result.

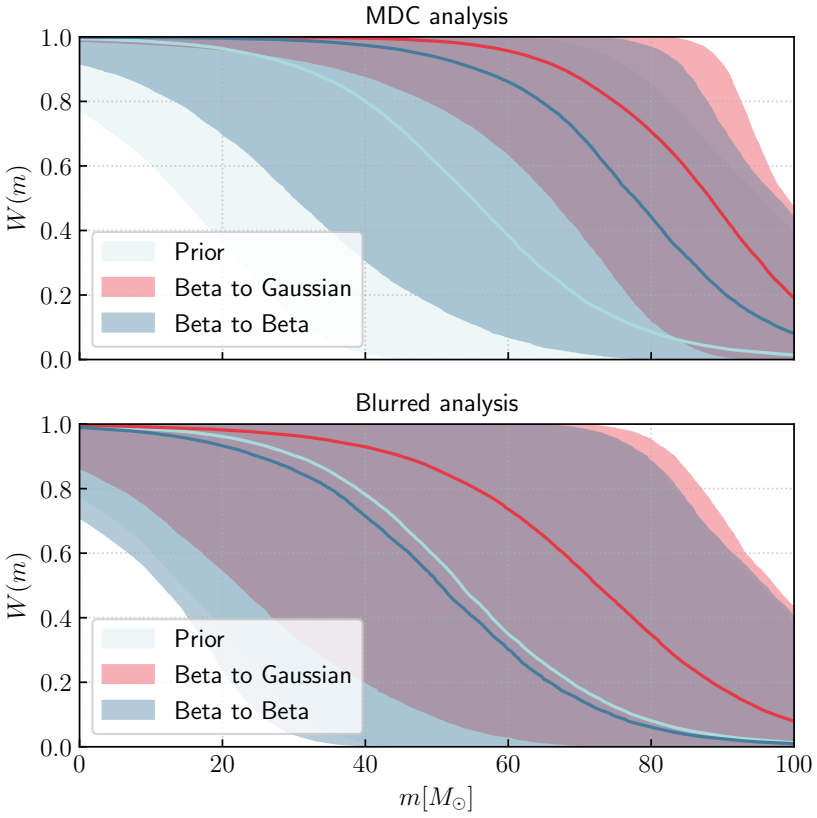

Figure 4.12: Reconstructed window function for the MDC (top) and the blurred (bottom) analysis, with the BETA TO GAUSSIAN and BETA TO BETA models. The colored contours are the 90% C.L. inferred spectra.

Beside the reconstruction of the spin magnitude, in analogy to the main analysis we also replicated the results on the aligned spin magnitude evolution with respect to the chirp mass. Fig. C.3 in App. C.1 shows that when no spin-mass correlation is present in the GW data, no evolutionary trend is recovered between these two parameters.

Concerning the EVOLVING GAUSSIAN model, we also compared the inferred population parameters that govern the evolution of the spin distribution as a function of the source mass, Fig. 4.13 shows the corner plot of $\dot{\mu}$ and $\dot{\sigma}$ for the MDC analysis. We observe that, contrary to the main analysis, the model now correctly supports a non-evolution of the spin distribution with the mass, the parameters $\dot{\mu}$ and $\dot{\sigma}$ are pushed to very small values, close to zero. Finally, the inferred mass spectrum with the EVOLVING GAUSSIAN model and the two TRANSITION models are in excellent agreement with the simulated population, as depicted in Fig. C.4 in App. C.1.

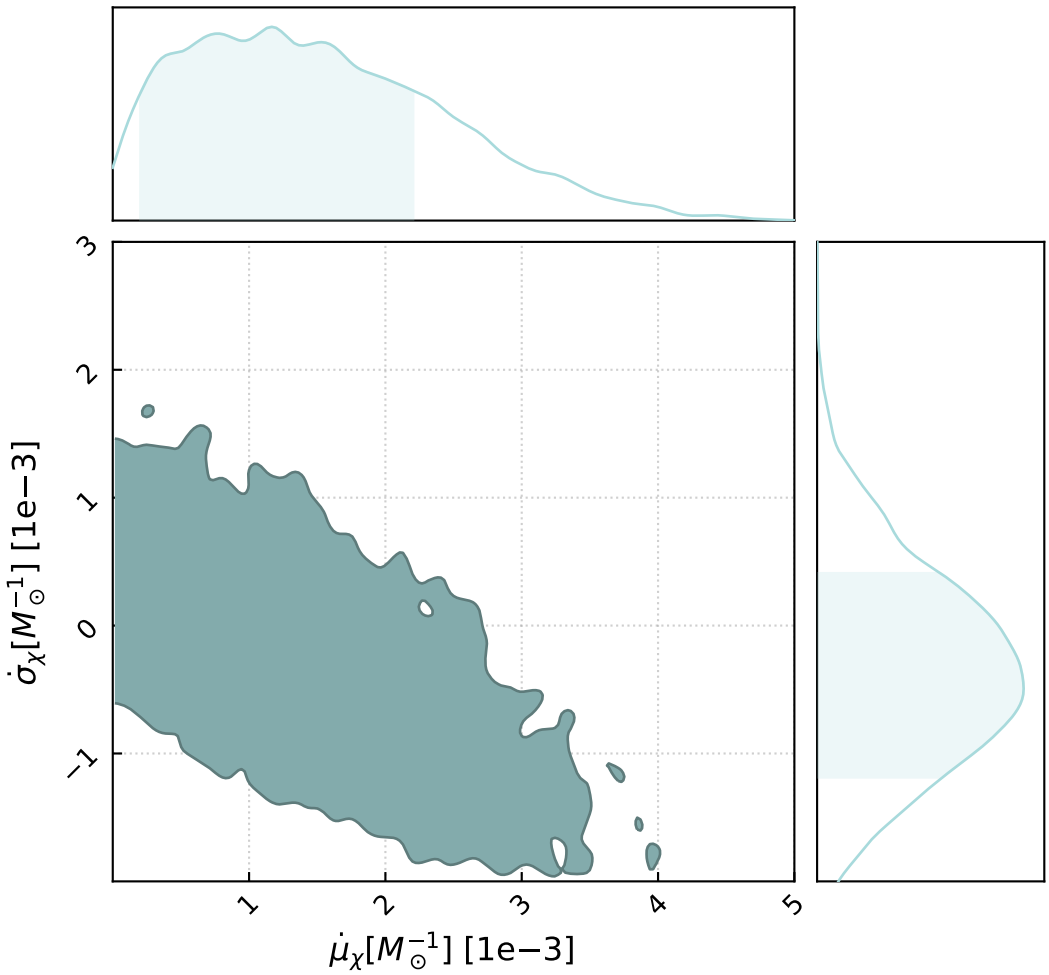

Figure 4.13: Corner plot of the inferred population parameters governing the evolution of the mean and width of the Gaussian distribution modelling the spin magnitude in the EVOLVING GAUSSIAN model, namely $\dot{\mu}_\chi$ and $\dot{\sigma}_\chi$. These results have been obtained for the MDC analysis.

4.5.3 Considerations on MIXTURE models

For the population inferences which use the MIXTURE models, Fig. 4.14 shows the reconstructed posterior predictive checks of the spin magnitudes and tilt angles, in analogy with the result of the main analysis. The simulated spin population is displayed with the black dashed line. We find that the inferred spin magnitudes and tilt angles distribution, obtained with all three MIXTURE models are in perfect agreement with the simulated BBH population. The total population reconstructed matches with the simulated one. This result is proof for the correct behavior of our MIXTURE models when no spin-mass correlation is present in the data, they are all able to reduce to a single spin population to reconstruct the true spin distributions.

As in the main analysis, we also looked at the inferred posterior distribution of the mixing parameter λ_{pop} and its correlation with the minimum mass of the second population m_{min}^{pop2}. The top panel of Fig. 4.15 and Fig. 4.16 show the inferred marginalized posterior of λ_{pop} and the corner plot of this parameter with m_{min}^{pop2} for the MDC analysis. We find that for the MIXTURE VANILLA and MIXTURE PAIRED analysis, the inference is not able to disentangle the presence of two populations of BBHs as the posterior of the mixing fraction are highly uninformative. Estimating a flat posterior for this population parameter means that our model does not find evidence for the existence of a second population with a different spin magnitude, which is the expected behavior when no spin-mass correlation are simulated like in this MDC. For the MIXTURE PEAK model, the mixture fraction is still constrained as the Gaussian peak for the spin magnitude is used to fit the BBHs within the central mass feature of the POWER LAW + PEAK distribution used for the simulation. The mixture fraction only estimates the fraction of BBH events in this contained in this mass feature, but still infer the same spin magnitude distribution as in the rest of the population.

In addition to these results, the BBH mass spectrum and the CBC merger rate reconstructed with the MIXTURE models are also nicely estimated with respect to the simulated population. The bottom half of Fig. C.4 present the inferred mass spectrum obtained with the MIXTURE models, that are all in very good agreement with the population simulated with a POWER LAW + PEAK. For the CBC merger rate, Fig. C.1 shows the inferred values of the population parameter γ, for both subpopulations. We find that they are all consistent with the injected value of 2.7.

Based on the results obtained for the MDC analysis, we can confidently say that EVOLVING, TRANSITION and MIXTURE models present the expected behavior when faced with data in which no spin-mass correlation or subpopulations are introduced. We correctly inferred the spin distributions (both spin magnitude and tilt angles) used for the simulation of the GW data. The other population distributions such as the mass spectrum and the CBC merger rate were also inferred properly and in good agreement with the simulated population of BBHs.

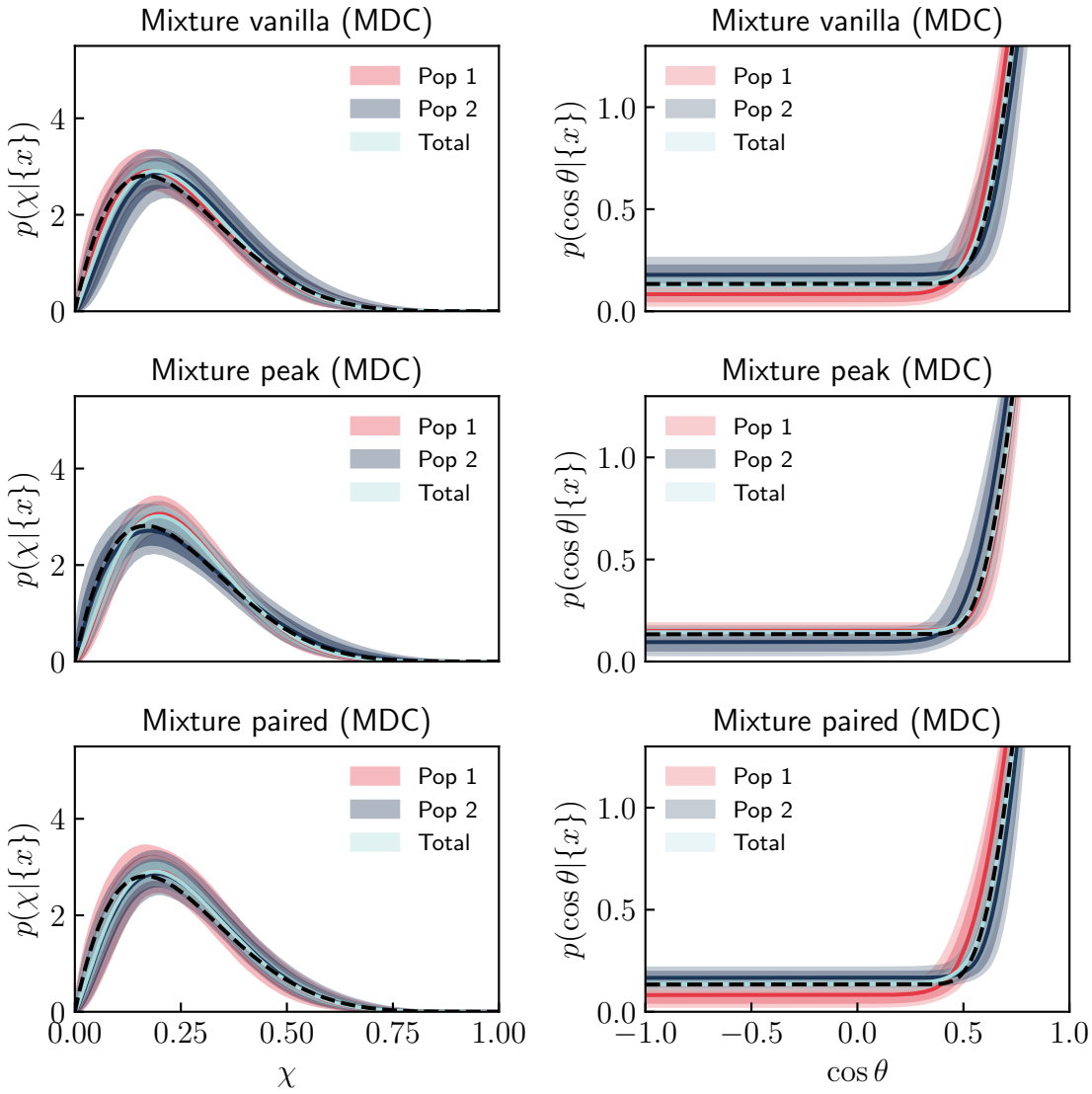

Figure 4.14: Reconstructed spectra of the spin magnitude distributions χ and the cosine of the tilt angle *cost* obtained with all three flavors of the MIXTURE models, namely MIXTURE VANILLA (top row), MIXTURE PEAK (middle row) and MIXTURE PEAK (bottom row). The population inference has been done on 50 detected GW events simulated for the MDC analysis. The red curves (Pop1) and blue curves (Pop2) are the estimated distributions for each population. The light blue curve (Total) shows the combined population when Pop1 and Pop2 are summed using the mixing fraction parameter λ_{pop}. The colored contours are the 90% and 98% C.L.

4.6 A BLURRED ANALYSIS

In addition to the MDC analysis presented in the section above, we perform another sanity check to ensure the robustness of our results. The idea is to repeat all of our population inference using our phenomenological models on a spin-mass blinded data set of real GW detections, we named this sanity check the "*blurred*" analysis. We selected

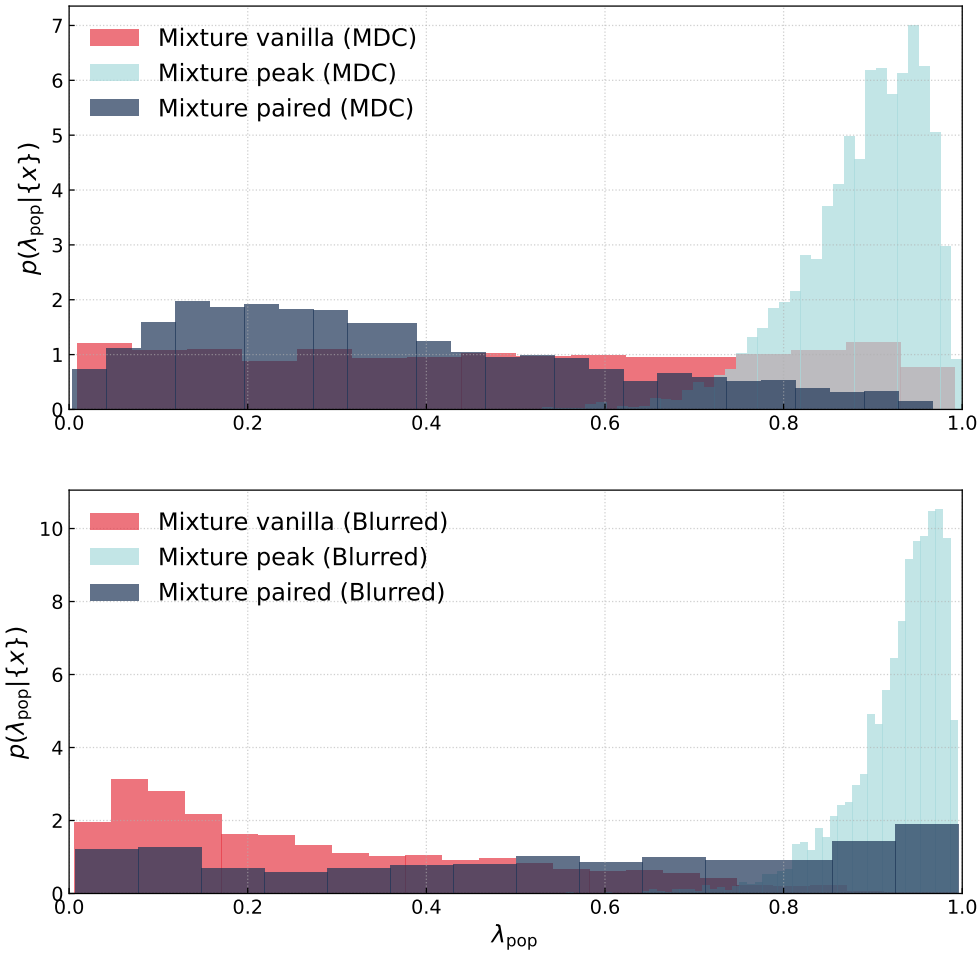

Figure 4.15: (Top panel) Posteriors distribution of the inferred λ_{pop} from the MDC analysis using the three flavors of the MIXTURE model. (Bottom panel) Posteriors distribution of the inferred λ_{pop} from the blurred analysis using the three flavors of the MIXTURE model.

the same set of 59 GW events as in the main analysis, from the GWTC-2.1 and GWTC-3 catalogs, with $IFAR \geq 1yr$. But for the blurred analysis, we permute among them their inferred spin values from the detection. Each GW events are then assigned random values of for their spin parameters, picked from other GW detections in the same set of data. In the limit that the determination of the other GW parameters, and the selection biases, are not strongly related to the spin magnitudes, this procedure artificially blinds the dataset to any spin-mass correlation.

With only 59 GW detections used for the population inference, we are placed in a low number of detections regime. In this regime, if there is really a spin-mass correlation in the data, shuffling the inferred spin values might not blind the totality of the

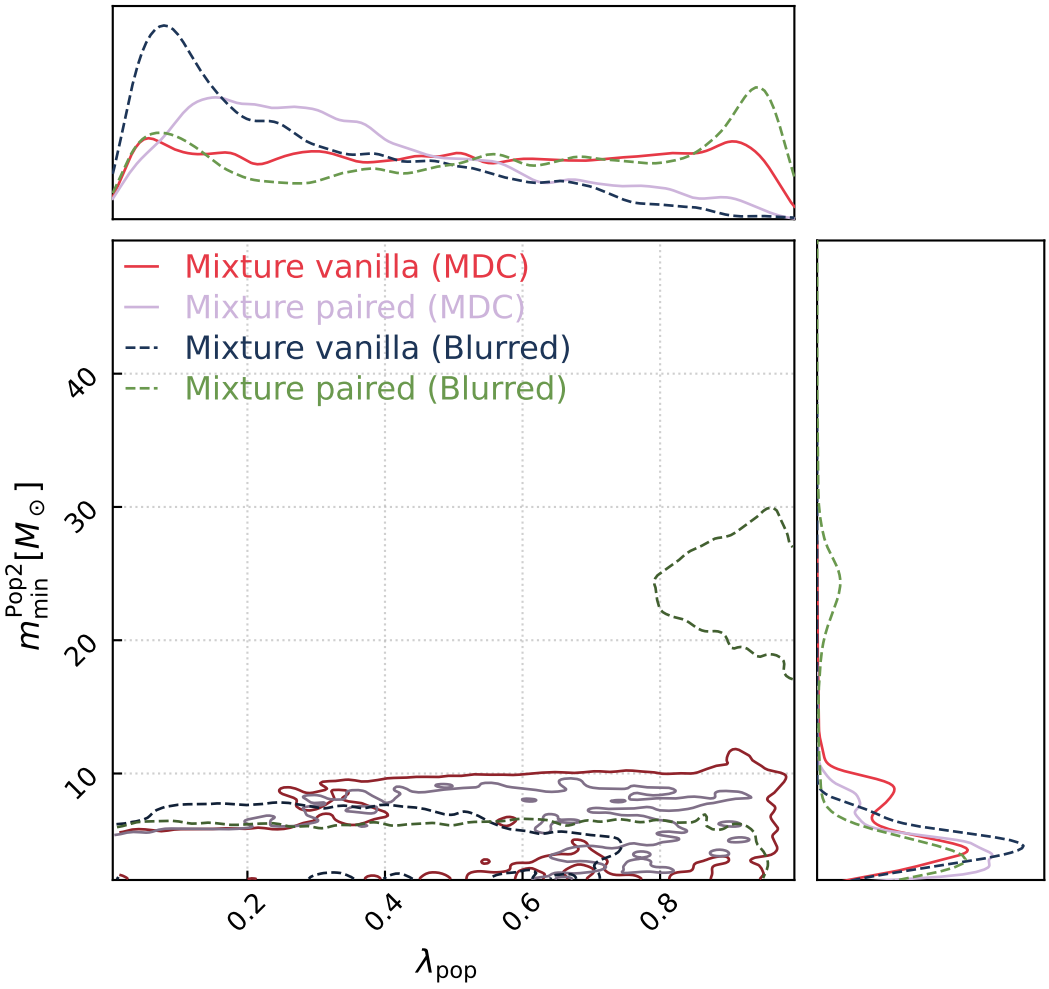

Figure 4.16: Overlapped corner plots of the mixing parameter λ_{pop} and the minimum mass of the secondary population m_{min}^{pop2}, obtained from the MIXTURE VANILLA and MIXTURE PAIRED models from the MDC and the blurred analysis.

data from a spin-mass correlation. The motivation is that low-mass events are more numerous than high-mass, therefore there is a higher probability that the spin of a low-mass event will be reassigned to a low-mass event, hence conserving the hidden spin-mass relation. However, the shuffling of the spin values is expected for sure to blind any spin-mass correlation that would be present at high masses since they are fewer. We refer to this type of sanity check analysis as *"blurred"* analysis to indicate that we are not able to ensure that 100% of the spin-mass correlation has been removed from our 59 GW events. The new distributions of source frame masses and spin magnitude for this set of events is shown in Fig. 4.17 with their estimated errors from the parameter estimation process. In comparison to the scatter plot with the real 59 GW

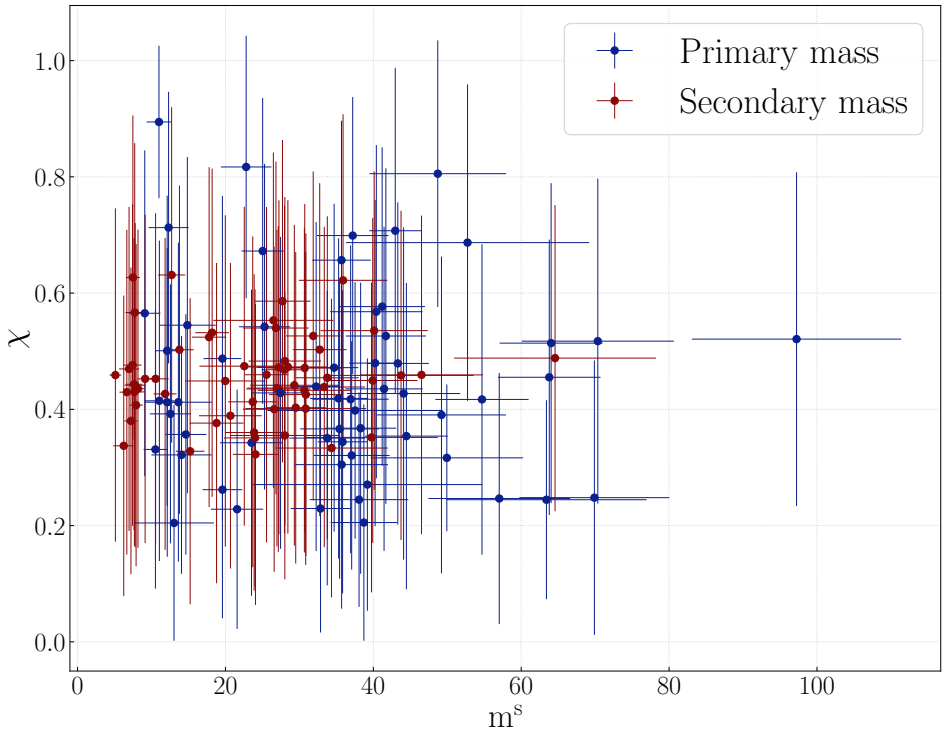

Figure 4.17: Scatter plot of the 59 BBH GW events from the GWTC-2.1 and GWTC-3 catalogs, selected with an $IFAR \geq 1yr$, and used for the blurred analysis. The estimated values of spins are shuffled between each event to blind the presence of any spin-mass correlation in the data set. The x-axis shows the source frame masses m^s and the y-axis displays the dimensionless spin magnitude χ. The error bars are the 1σ uncertainties of the official LVK parameter estimations C01:mixed samples. The blues crosses indicate the primary mass and spin parameters, while the red ones are for the secondary component.

events used in the main analysis, the blurred distribution of the spin magnitude versus the source frame masses clearly indicate no more visible evolution. The over density of GW events around $30\,M_\odot$ is still present, but the associated spin magnitudes do not seem to increase as the mass increases.

4.6.1 Models selections

In analogy with the MDC analysis of Sec. 4.5, we also want to ensure that the derived Bayes factors between our models and the baseline analysis are consistent with our expectations. Since the blurred set of GW events is supposed to have no spin-mass correlation, our complex phenomenological population models should not be preferred with respect to the more simple non-evolving population model.

Blurred analysis

Model	$\log_{10} \mathcal{B}$	$\log_{10} \mathcal{L}_{max}$
Evolving Gaussian	-3.34	-0.23
Beta to Gaussian	-0.28	0.18
Beta to Beta	-0.06	0.40
Mixture vanilla	0.90	0.92
Mixture peak	0.76	0.40
Mixture paired	1.11	2.24

Table 4.3: Base 10 logarithm of the Bayes factors (second column) and the logarithm of the maximum likelihood ratio (third column) for the blurred analysis. The six populations models are discussed in the above section, and are compared to the baseline non-evolving population model.

Table. 4.3 presents the logarithm in base 10 of the Bayes factors and the maximum likelihood ratios between each of our models compared to the canonical analysis. All evolving, transition and mixture populations models report inconclusive Bayes factors, meaning that no preference is found between our models and the baseline analysis from the point of view of the inference. A similar interpretation can be made about the likelihood ratios, that shows no significant preference. These values correspond to what one would expect using an overcomplicated population model, with multiple degrees of freedom that are not useful and do not help the overall fit of the population. Since our new models are complex extensions of the baseline population model, uninformative Bayes factor like the ones in table. 4.3 are reasonable, especially when compared to values obtained with the same set of GW events, unblurred.

4.6.2 Considerations on evolving and transition models

We can now verify that the population distributions (spin, mass and CBC merger rate) inferred with icarogw for the blurred analysis are in agreement with a population of BBHs that has no spin-mass correlation, nor subpopulations. Fig. 4.18 shows the posterior predictive check of the spin magnitudes for the evolving gaussian, beta to beta and beta to gaussian models. In comparison with the real analysis, we still observe that in the low mass region, the analysis reconstructs a spin magnitude distribution preferring low spin values ≤ 0.4. The spin magnitude reconstruction at higher masses is still consistent with slowly spinning BHs, but the overall uncertainties increase too. This is due to the fact that low mass events are more numerous than high mass ones, and that the spin-mass relation in the data is not completely blinded by the shuffling of the spin values. However, the spin-mass relation is diluted enough for the Bayes

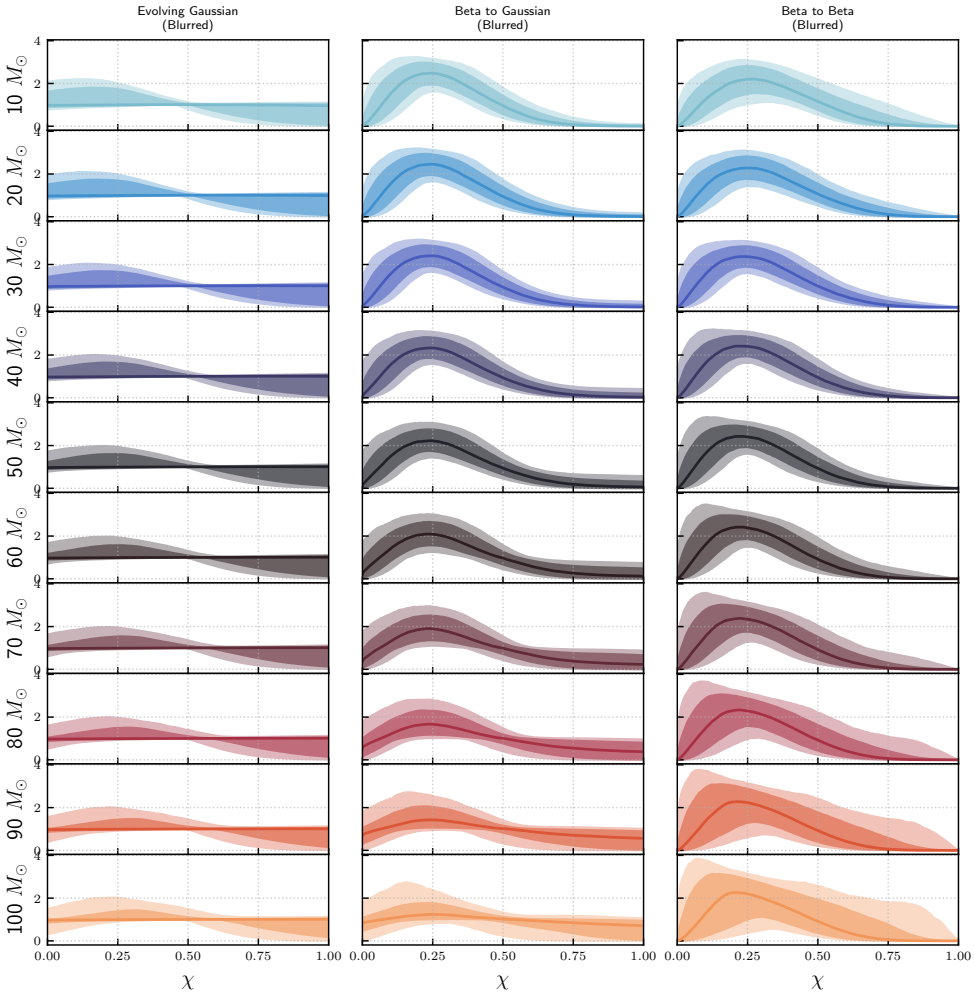

Figure 4.18: Joy plot of the probability density functions of the dimensionless spin magnitudes χ, reconstructed from the population inference on the real GW for which we blurred the spin-mass relation. The dimensionless spin magnitudes were reconstructed using the EVOLVING GAUSSIAN model (left column), the BETA TO GAUSSIAN model (middle column) and the BETA TO BETA model (right column). Each row corresponds to a slice in source frame mass, from $10\,M_\odot$ to $100\,M_\odot$, in order to highlight the spin magnitude evolution and transitions across different mass ranges. The black dashed lines are the simulated population from the MDC. The plain lines are the median value of the spin magnitudes and the colored contours are the 90% and 98% C.L.

factors on model selection to be inconclusive. Regarding the presence of an evolution of the spin magnitude, the posterior predictive checks for all three models show no support for a change in the distribution as the mass increases, as expected from the blurred population of BBHs.

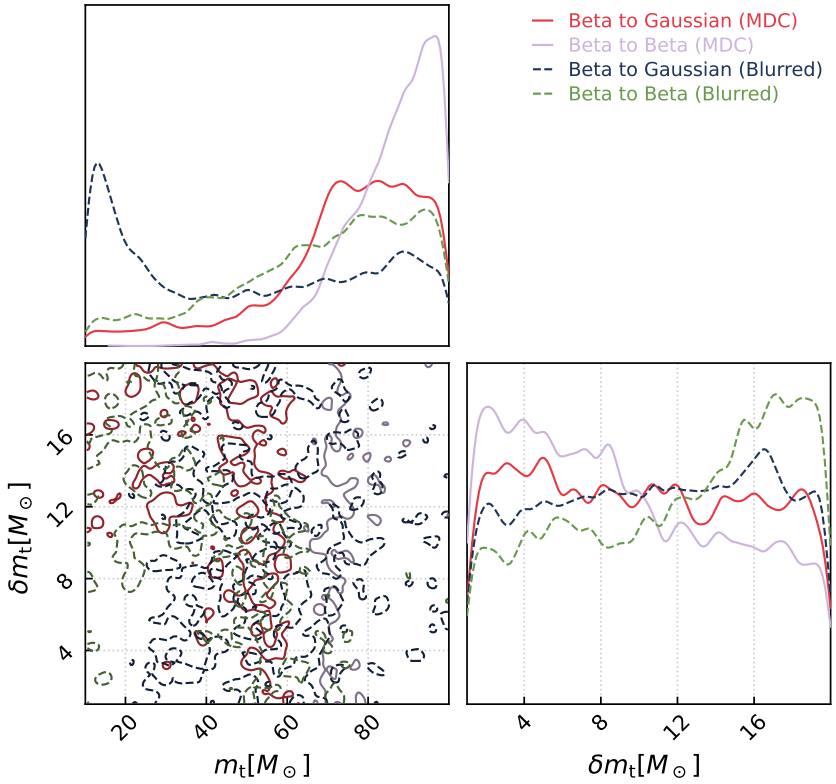

Figure 4.19: Overlapped corner plots of the population parameters governing the spin transition as a function of the mass, namely m_t and δ_{m_t}; for the two TRANSITION models in the case of the MDC and the blurred analysis.

Concerning the BETA TO BETA and BETA TO GAUSSIAN models, Fig. 4.12 (bottom panel) and Fig. 4.19 display the posterior predictive check for the reconstructed window function and the population parameters posteriors that govern it. The results are both consistent with no significant transition between two spin distributions. In opposition with the real analysis, the reconstructed window function shows a light support for a smoother transition at higher masses. The inferred population parameters that govern the window function are also consistent with no transition, as argued for the MDC analysis.

However, the result is not as stringent as for the MDC, which is a hint of the fact that the spin shuffling procedure to remove the spin-mass correlation has partially worked. We also check the evolution of the aligned spin component of χ with respect to the detector chirp mass. From Fig. C.5 in App. C.1, we find no evidence of any sort of evolution of the aligned spin component $|s_z|$ with respect to the chirp mass, as the one found in the real analysis. Regarding the EVOLVING GAUSSIAN analysis, again we find that there no support for a continuous evolution of the distribution of the spin magnitude with the source mass. Fig. 4.20 depicts the posterior distributions of

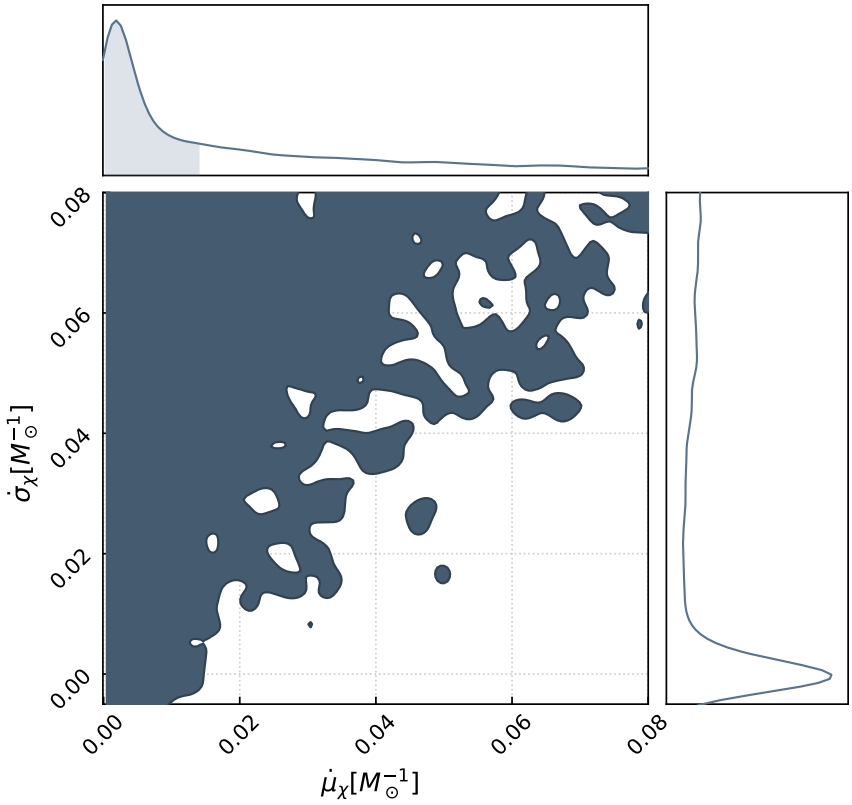

Figure 4.20: Corner plot of the inferred population parameters governing the evolution of the mean and width of the Gaussian distribution modelling the spin magnitude in the EVOLVING GAUSSIAN model, namely $\dot{\mu}_\chi$ and $\dot{\sigma}_\chi$ for the blurred analysis. These results have been obtained for the blurred analysis.

the parameters governing the evolution of the mean and standard deviation of the Gaussian spin distribution, and both posteriors are peak around zero, meaning that the spin distribution position and width are found fixed as the mass increases.

The reconstructed mass spectra obtained with the EVOLVING and TRANSITION models for the blurred are also in good agreement with the ones of the real analysis. Fig. C.6 in App. C.1 shows the inferred spectrum for the primary and secondary masses.

4.6.3 Considerations on MIXTURE models

We now finally examine the result obtained with the MIXTURE models for the blurred analysis. Fig. 4.21 show the posterior predictive checks for the spin magnitude and the tilt angle distributions reconstructed with the MIXTURE VANILLA, MIXTURE PAIRED and MIXTURE PEAK models. In this scenario, where most of the spin-mass correlations are removed, both subpopulations support low spin values, the spin magnitude distributions for the magnitude and the tilt angles are very similar. The second population,

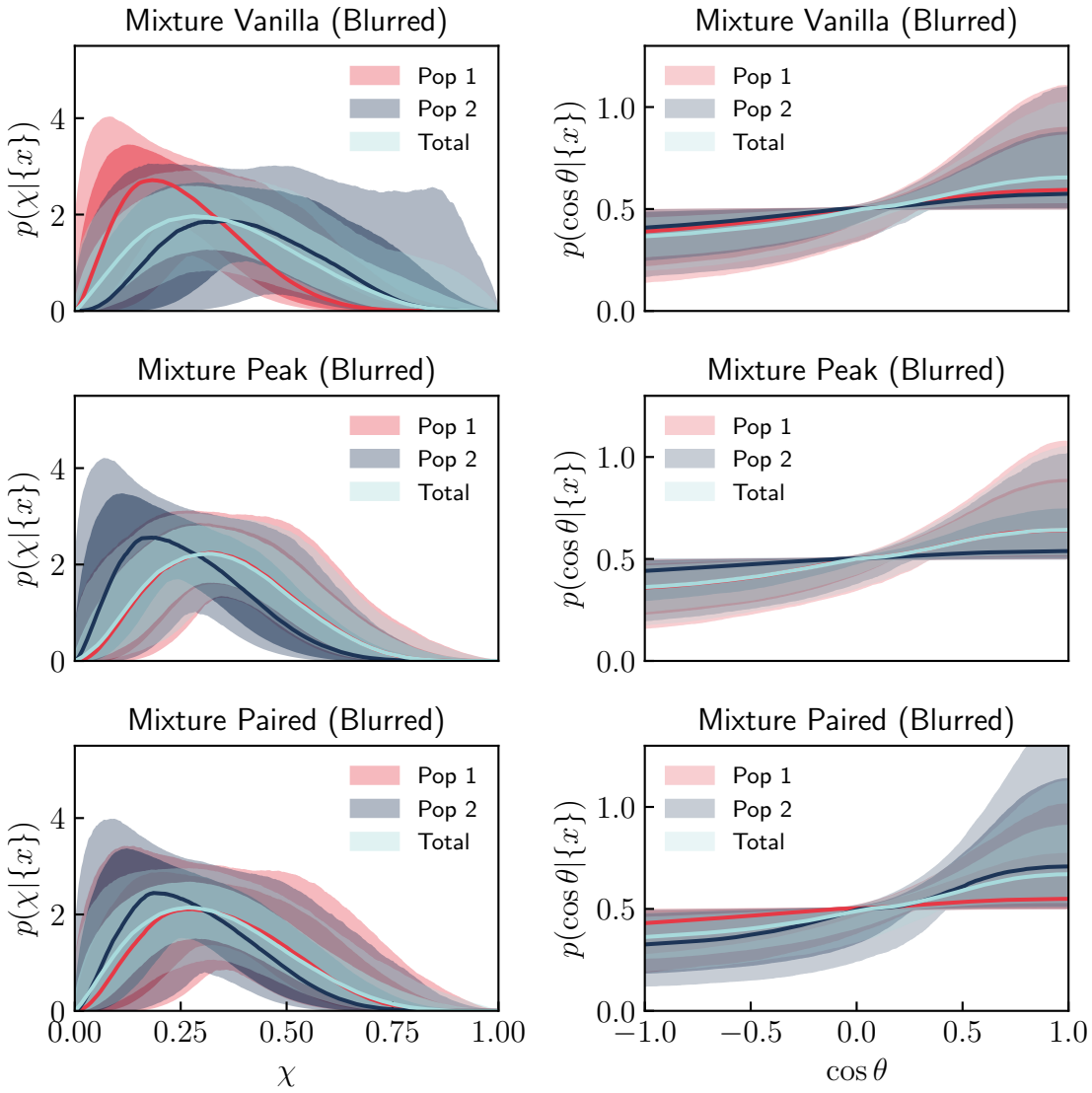

Figure 4.21: Reconstructed spectra of the spin magnitude distributions χ and the cosine of the tilt angle *cost* for the blurred analysis, obtained with all three flavors of the MIXTURE models, namely MIXTURE VANILLA (top row), MIXTURE PEAK (middle row) and MIXTURE PEAK (bottom row). The population inference has been done on 59 detected GW events from GWTC-2.1 and GWTC-3 catalogs, for the blurred analysis. The red curves (Pop1) and blue curves (Pop2) are the estimated distributions for each population. The light blue curve (Total) shows the combined population when Pop1 and Pop2 are summed using the mixing fraction parameter λ_{pop}. The colored contours are the 90% and 98% C.L.

for more massive binaries has slightly more uncertain spin reconstructions, due to the lack of events in this part of the parameter space. Similar to the results of the MDC, for the blurred analysis the mixing fraction λ_{pop}, shown in bottom half of Fig. 4.15, is not able to pinpoint the presence of any subpopulation of BBHs as opposed to what we observe with unblinded data.

The reconstruction of the mass spectrum, displayed in Fig. C.6 of App. C.1, is still in very good agreement with the distributions obtained in the real analysis, thus indicating that the spin-mass relation is not an artifact created by the fit of the mass spectrum. This point is crucial for the robustness of our results, since we have demonstrated in Sec. 3 that a wrong reconstruction of some population parameters can lead to significant biases. The mass spectra for all flavors of the MIXTURE models show also perfect agreement with the distribution obtained with the non-evolving model. This result also extend to the CBC merger rate presented in right column of Fig. C.1. The estimated posterior distributions of the population parameter γ are all consistent with the ones from the real analysis, as well as the value obtained from the baseline non-evolving model in the blurred analysis.

In conclusion, when analyzing real GW data devoid of spin-mass correlation, like in our blurred analysis, the MIXTURE, EVOLVING and TRANSITION models successfully reconstruct the correct mass, CBC merger rate and spin distributions, while indicating no support for any spin evolution. Additionally, we verified that our models are naturally disfavored by the Bayes factor model selection, when a more simple parameterization is sufficient. This second sanity check, the blurred analysis, allows us to further believe the astrophysical results about the presence of a subpopulation in the entire BBH population. In comparison to the results of the real analysis, we identify a significantly different behavior of our models regarding the spin distributions, while conserving the same reconstructed mass and merger rate spectra. We argue that the evidences for the presence of a subpopulation is then driven by the more detailed inference of the spin distribution.

4.7 NUMERICAL STABILITY ANALYSIS

The last non-physical possibility that could explain our results is problem of numerical stability of the hierarchical Bayesian inference. As explained in Sec. 2.4, ICAROGW has two stability estimators, that allows us to track the stability of the inference. These two estimators are the effective number of injections N_{eff}^{inj} and the effective number of posterior samples N_{eff}^{PE}. In the main analysis of this chapter, we applied cuts on these two parameters, in order to set the hierarchical likelihood to zero if one of the two conditions is not fulfilled at some point during the inference process. They were chosen to be $N_{eff}^{inj} = 4N_{GW}$ and $N_{eff}^{PE} = 10$, meaning that four times the number of used GW events were required for the effective injections and ten effective posterior samples minimum.

We performed two more analysis, using the BETA TO BETA and MIXTURE VANILLA models, on the same 59 GW events with $IFAR \geq 1yr$, but this time applying not cuts on the two stability estimators. Fig. 4.22 and Fig. 4.23 present some inferred distributions of key population parameters that drive the distributions of spins, alongside the estimated distributions of the two numerical stability for the BETA TO BETA and MIXTURE VANILLA models respectively. First, we do not observe any significant correlation be-

tween the population parameters and the stability estimators of ICAROGW, indicating that the spin evolution part of the parameter space explore by the MCMC in not influenced or dependent on N_{eff}^{inj} nor N_{eff}^{PE}. Furthermore, removing these cuts did not cause the hierarchical likelihood to shift to a different part of the parameter space, as all posterior distributions remain very similar with and without applying cuts on the stability estimators. Hence, we can safely conclude that the results obtained in the main analysis are stable with respect to numerical stability of the hierarchical Bayesian inference.

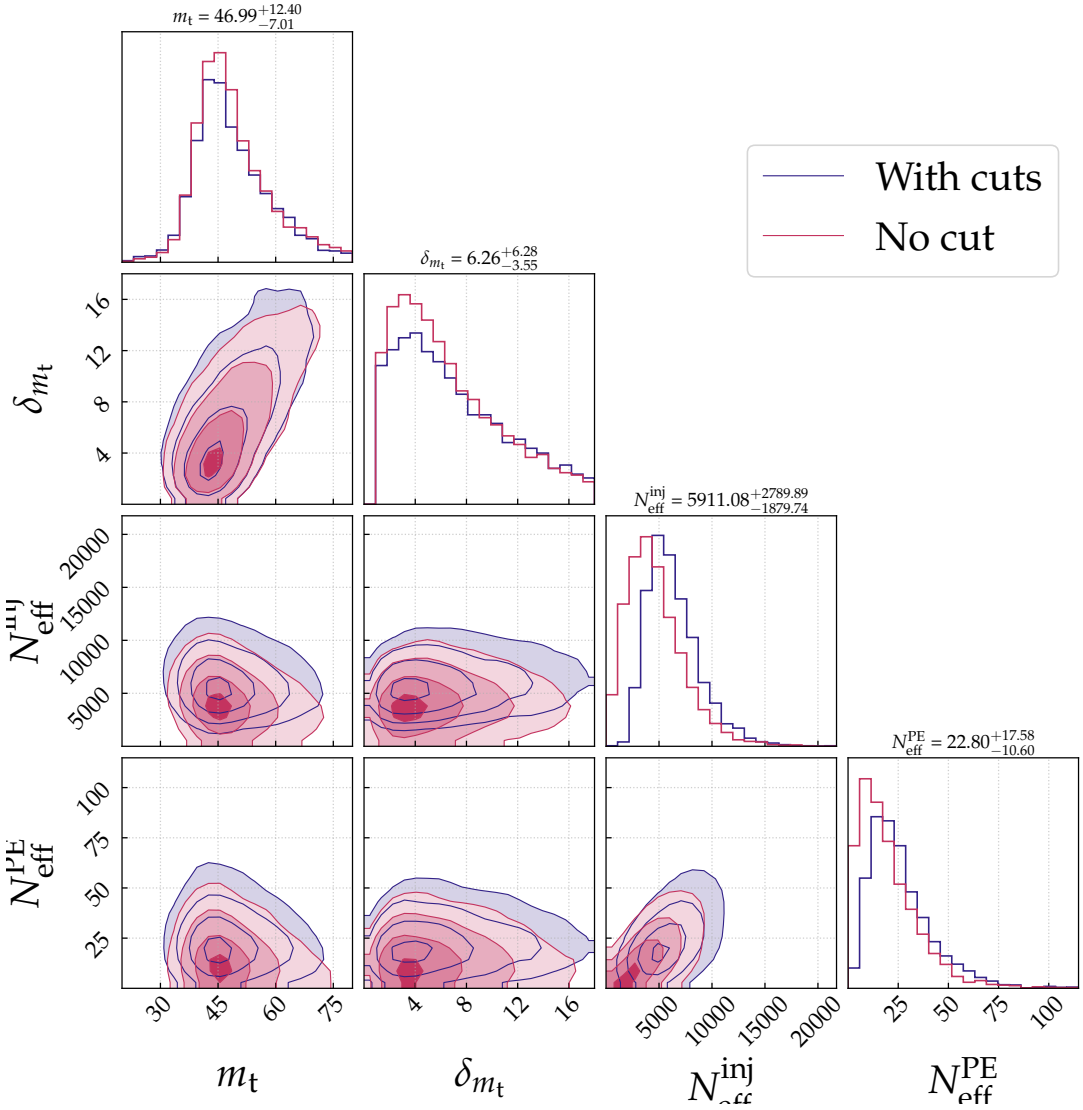

Figure 4.22: Corner plot of the population parameters (m_t, δ_{m_t}) and the stability estimator $(N_{eff}^{inj}, N_{eff}^{PE})$, obtained with the BETA TO BETA evolving model on the 59 GW events with $IFAR \geq 1yr$. The purple contours and histograms were estimated while putting the minimum value of $N_{eff}^{inj} = 4N_{GW}$ and $N_{eff}^{PE} = 10$. The pink contours and histograms were estimated while placing not cuts on these estimators.

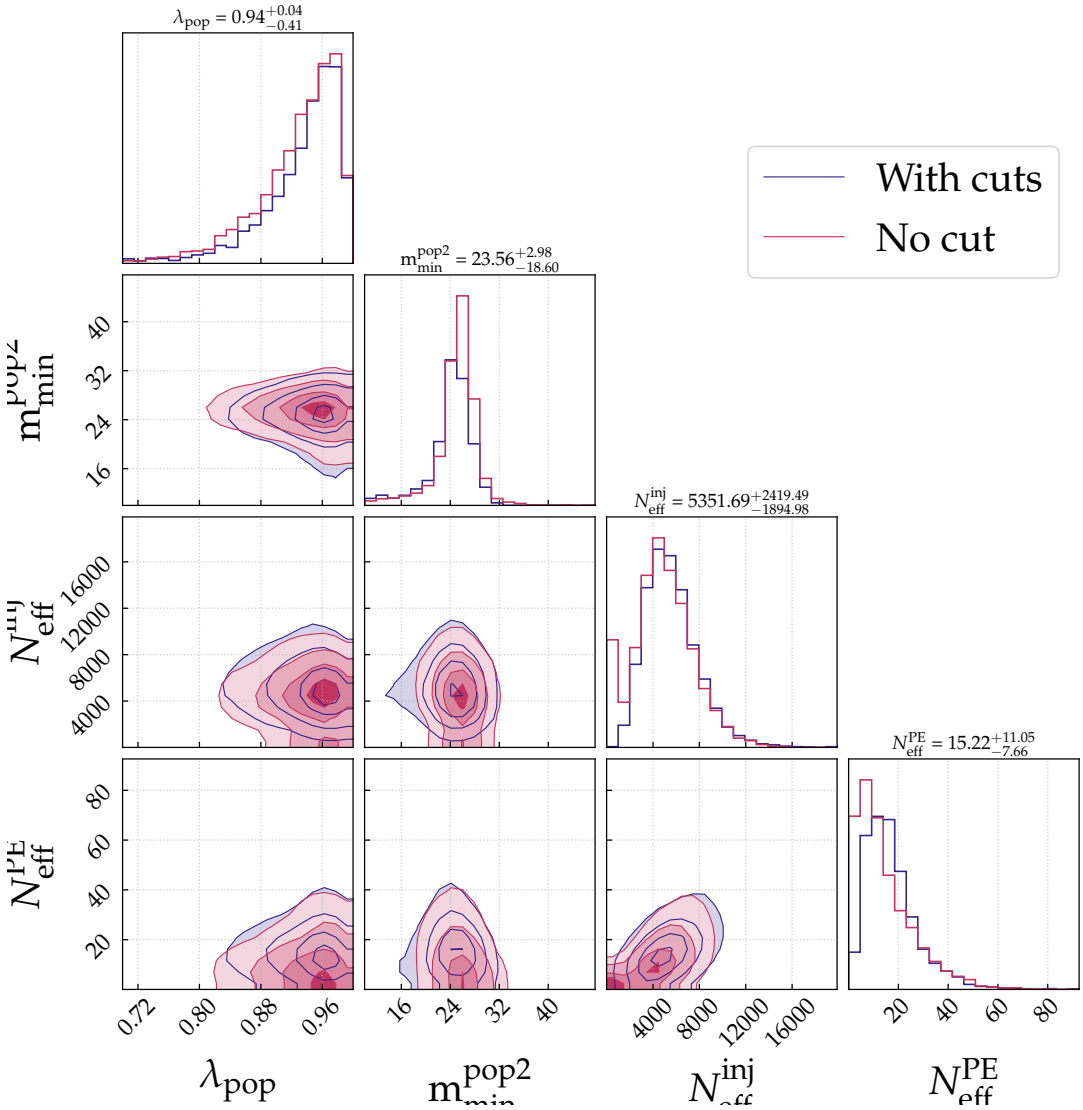

Figure 4.23: Corner plot of the population parameters (m_t, δ_{m_t}) and the stability estimator $(N_{eff}^{inj}, N_{eff}^{PE})$, obtained with the MIXTURE VANILLA evolving model on the 59 GW events with $IFAR \geq 1yr$. The purple contours and histograms were estimated while putting the minimum value of $N_{eff}^{inj} = 4N_{GW}$ and $N_{eff}^{PE} = 10$. The pink contours and histograms were estimated while placing not cuts on these estimators.

4.8 CONCLUSION

The comprehensive analysis of BBH populations presented in this study has yielded several significant findings. Through the application of three distinct families of parametric models - EVOLVING, TRANSITION, and MIXTURE models - within the hierarchical Bayesian inference framework of ICAROGW, we have gained valuable insights into the characteristics and potential correlations within the BBH population.

Our investigation revealed strong evidence for the presence of a subpopulation within the overall BBH population, primarily driven by the detailed inference of the spin distribution. The MIXTURE models, particularly the MIXTURE VANILLA and MIXTURE PAIRED variants, consistently indicated a preferred mode with $\lambda_{pop} \sim 0.98$ and $m^{pop2}_{min \sim 30\,M_\odot}$. This suggests the existence of a distinct subpopulation of BBHs with different spin distributions emerging above $30 - 35\,M_\odot$. The reconstructed mass and CBC merger rate spectra across all models showed remarkable consistency with canonical reconstructions from baseline analysis and other non-evolving population inferences. This consistency lends credibility to our findings and supports the robustness of our methodological approach.

A key observation from our analysis is the apparent evolution of black hole spin magnitudes with their masses. The BETA TO BETA evolving model and the MIXTURE models all pointed towards this conclusion, with the spin magnitude distributions showing distinct characteristics for different mass ranges. To validate our results and ensure they were not artifacts of our analysis methods, we conducted several sanity checks. A blurred analysis, using GW data devoid of spin-mass correlation, demonstrated that our models successfully reconstruct correct mass, CBC merger rate, and spin distributions while indicating no support for spin evolution when it is absent. This check reinforced the validity of our astrophysical results regarding the presence of a subpopulation in the entire BBH population.

Furthermore, we performed a numerical stability analysis to rule out any potential issues with the hierarchical Bayesian inference. By comparing results with and without cuts on the stability estimators (N^{inj}_{eff},N^{PE}_{eff}), we confirmed that our findings were not influenced by numerical instabilities in the inference process. The Bayes factor model selection consistently favored our more complex models over the simpler non-evolving baseline model when analyzing real GW data. Conversely, when applied to the blurred dataset without spin-mass correlation, the Bayes factors appropriately disfavored the more complex models, aligning with our expectations.

In conclusion, this study provides compelling evidence for the existence of distinct subpopulations within the BBH population, characterized by different spin distributions and potentially representing different formation channels. The observed correlation between spin magnitudes and masses offers new insights into the formation and evolution of binary black hole systems. These findings have significant implications for the understanding of BBH populations and their astrophysical origins, paving the way for further investigations and refinements in gravitational wave astronomy.

5 VIRGO DATA QUALITY AND MACHINE LEARNING ALGORITHM

This chapter is based on an exploratory study that I conducted with the Detector Characterization research group of the LIGO-Virgo-KAGRA Collaboration. My mission was to implement and test the performances of iDQ for the Virgo interferometer, for the O4 observing run of the LVK detector network.

5.1 INTRODUCTION

The journey to the detection of GWs began during the 1960s with the pioneering efforts of Joseph Weber, who proposed the first experimental apparatus to make a direct detection of GWs, predicted by the theory of general relativity [185]. He used an aluminum bar equipped with piezoelectric crystals to measure its expansion and contraction, induced by the passage of GWs. Despite Weber's claim to have discovered GWs in 1969, subsequent attempts by other researchers failed to reproduce his results, highlighting technical defects in his experiment. The fault was probably with Weber

and his apparatus, not able to thoroughly quantify the noise background, especially non-gaussian noises. However, Weber's fundamental work stimulated further efforts. In 1974, Russell Hulse and Joseph Taylor provided the first indirect evidence of the existence of GWs through the discovery of the binary pulsar PSR 1913+16. They found that the pulsar orbital period was varying repetitively, meaning that it was in orbit with a star. The evolution of the orbiting period was measured with high precision, and showed a decreasing trend over time. This decreased time of the orbital period was the sign that the binary system was merging, loosing its angular momentum. The analysis of the orbital period data confirmed the predictions of the GR theory for the emission of GWs by accelerating bodies [186]. This discovery revived the search for a direct detection, leading to the development of the first laser interferometer, originally proposed by Gertsenshtein and Pustovoid in 1962 [187]. The first prototypes were built in the early 1970s by Robert L. Forward and his colleagues at Hughes Research Laboratories, Rainer Weiss at MIT, Heinz Billing in Germany, and Ronald Driver and James Howe in Scotland [188–190].

In the early 2000s, the first generation of large-scale GW detectors, known as "Initial LIGO", were put into operation in Hanford and Livingston. Although these detectors did not report direct detections, they opened the way for future advances [8, 191, 192]. At the same time, in Europe, the Virgo detector was under construction in Cascina, Italy, and became later operational in 2003 [9]. In 2015, the upgraded Advanced LIGO detectors made significant progress in their sensitivities, allowing the achievement of the first confident direct detection of a GW signal from a black hole merger: GW150914 [6, 191, 193–206]. This milestone was achieved thanks to the common work between the LIGO and Virgo scientific collaborations. More recently, a new terrestrial laser interferometer, the Japanese KAGRA detector, designed to use new cryogenic technology, joined the global detection network [11, 207]. Looking ahead, new generations of detectors, such as Cosmic Explorer, LISA, and Einstein Telescope, aim to further enhance the sensitivity of GW detections, leading to a new era with significantly more frequent detections [208–212]. Pushing the limits of the current sensitivity, these new detectors, referred to as third generation (3G), aim to unveil new GW sources that are at the moment only predicted by the theory and beyond the reach of the current detectors [80, 213, 214].

These large scale detectors are designed to detect space distortions as small as $\mathcal{O} \sim 10^{-19}$ m. This incredible sensitivity, however, makes them highly susceptible to both instrumental and natural noise. As a result, a considerable amount of effort has been dedicated in developing advanced algorithms to clean the data before it can be used for scientific analyses. A major challenge in this process is the presence of "glitches." Glitches are transient noises within the detector, that can closely resemble genuine GW signals. Glitches introduce several issues in the search of GW detections, they can mask the presence of a real GW signal if superimposed with it, they can bias the parameter estimation results of the source properties, or lead to a wrong estimation of

the detector sensitivity. These glitches can originate from various sources and might correlate with the auxiliary channels—sensors placed throughout the detector that monitor environmental and instrumental conditions. Given the large number of these auxiliary channels, and the complexity of the relationships between these noises and the GW data stream, classical algorithms struggle to process the data and identify correlations between the auxiliary channels and the GW detection channel, $h(t)$. To address this, new techniques based on machine learning (ML) have been developed. These advanced methods aim to enhance the data quality (DQ) of the interferometer by accurately identifying and flagging glitches. ML algorithms can analyze large datasets more efficiently and potentially identify patterns that classical methods might miss. By doing so, they help in isolating genuine GW signals from noise, thereby improving the reliability of detections and the overall performances of the detectors.

This chapter is constructed as follows: In Sec. 5.2 we present the detection principles of GWs based on terrestrial laser interferometry techniques, from the basics of GW detection, to the impact of instrumental noise on the sensitivity and the recording of this noise. In Sec. 5.3 we review the different families of noise and the properties. Sec. 5.4 presents an exploratory work using ML based strategy to improve the data quality of the Virgo detector and its potential adaptation to the search algorithms for GW signal extraction.

5.2 GRAVITATIONAL WAVE DETECTION PRINCIPLE

This section aims to present the laser interferometer instrument and its detection principle. Without going too deep into the details, we discuss the optical layout of such detector and give an overview of how noise can impact the detector sensitivity. Lastly, we introduce Virgo's auxiliary channels and sensors, that are the main input for the ML analyses that we perform later with the iDQ algorithm.

5.2.1 Virgo laser interferometry

The fundamental detection principle of laser interferometry is based on the intrinsic property of GWs to distort the structure of space-time itself. The current detectors of GWs are large scale Michelson interferometers, that make use of this phenomenon to detect passing GWs. A GW interferometer is characterized by two perpendicular arms that are several kilometers long ($\mathcal{O} \sim 1km$), in which a powerful laser beam, separated at the crossing point of the two arms, travels along the arms back and forth. The laser, reflected by large mirrors placed at the beginning and ending points of the arms, will see its travel time along the arms modified by the change in separation of the mirrors due to the passage of a GW. Fig. 5.1 shows the optical layout of the Advanced Virgo interferometer [9]. The light source, typically a $1064nm$ laser beam, enters the cavities and is divided in two by the beam splitter mirror, separating the original laser into

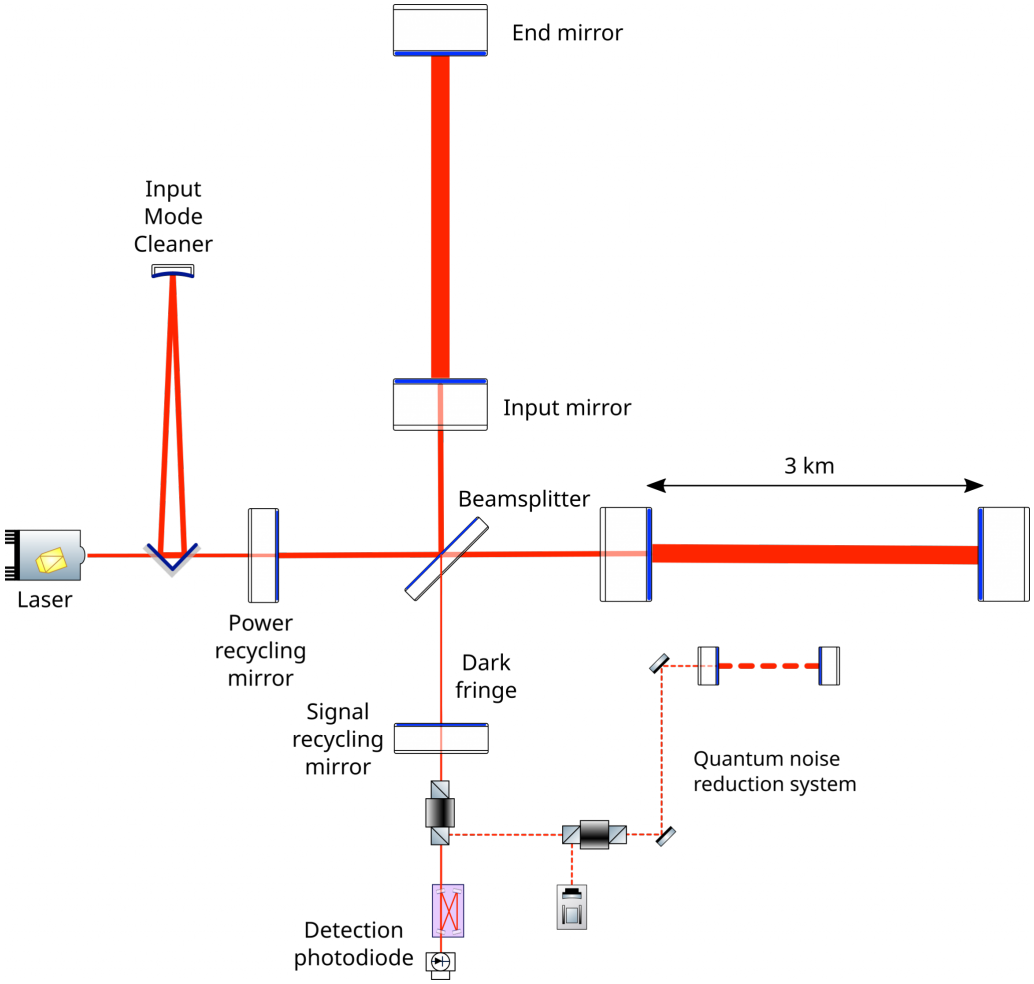

Figure 5.1: Optical layout of Advanced Virgo during the third observing run O3. This layout includes several recent upgrades of the detectors like the power recycling mirror and the signal recycling mirror, as well as the frequency dependent squeezing (quantum noise reduction system). Credits: The Virgo Collaboration.

two beams with equal intensities. Each beam is sent into each arm and then reflected by the west and north end mirrors of the detector. The lasers are then recombined by the beam splitter, and observed via a detection photodiode. The arm's lengths are set in such a way that when no GW goes through the detector, the recombined lasers are in opposition of phase, also called dark fringe. When a GW passes through the detector, it impacts the time travel of the light between each mirror, and the arm length variations induced by the GW can be written such that

$$\Delta L = \frac{h(t)L}{2}, \tag{5.1}$$

where L is the arm length when no GW passes through, ΔL is the absolute length variation of the arm, and $h(t)$ is the strain of the GW. This variation causes a phase shift of the two laser beams at the recombination point, creating light interferences, and the end photodiode detects the change in intensity of the recombined laser. The mirrors

of the interferometer are crucial components of the detector, in order to maximally reduce the potential disruptions of the mirrors, they are isolated and suspended in special columns called superattenuators, that reduce (by twelve order of magnitude) the vibrations of the ground, making the mirrors behave as if they were in "free-fall" with respect to the GWs, on the horizontal plane of the detector.

One important characteristic of the giant GW interferometer Virgo is the presence of Fabry-Pérot cavities inside both arms [9]. These cavities, made out of an extra mirror at the beginning of each arm (see Fig. 5.1), trap the laser beams, which go back and forth in between those extra optics and the end mirrors. This setup induces an increase of the laser power in the cavities, which increases the effect of even tiny phase shifts, making the interferometer more sensitive to weak GW signals. The added Fabry-Pérot cavities artificially lengthen the arms of the detector, hence the strain produced by a GW. Thanks to this complex optical set up, GW detectors like Virgo, with $3km$ long arms, are able to detect differential variations of the length of their arms of the order of $\mathcal{O} \sim 10^{-23}$.

Other innovative instrumental setups have been developed and added to the Virgo detector since its creation, like the power recycling and signal recycling cavities, or the frequency dependent squeezing of the laser. The power recycling cavities goal is to increase the power of the laser traveling inside the arms of the interferometer: since more power means larger amplitude of the signal, this results in an enhanced sensitivity to GW signals. The signal recycling cavity is used to enlarge the sensitivity band of the detector, making it sensible to potentially further sources or astrophysical phenomena. Finally, the frequency dependent squeezing is used to reduce the noise produced by the light particles of the laser when hitting the mirrors, called quantum noise.

Due to the shape of the interferometer, and the physical properties of GWs seen in Sec. 1.3, the sensitivity of such detector is not isotropic. They have specific responses to GW signals, which depend on the direction of arrival of the GW with respect to the orientation of the interferometer. As mentioned in Sec. 1.2, the detector sensitivity is at its maximum when the GW signal arrives perpendicular to the plane of the interferometer, depending on which polarization of the GW we consider, the sensitivity pattern will change too. Fig. 5.2 presents the antenna pattern response (see Eq. 1.6) for an L shaped Michelson interferometer like Virgo.

5.2.2 Sensitivity and noise impact

The main output of current GW detectors is a time-series called $h(t)$, which is the combination of the pure GW strain denoted $s(t)$ and the instrumental noise $n(t)$ as

$$h(t) = s(t) + n(t). \tag{5.2}$$

We call a GW detection when the strain $s(t)$ is successfully extracted from the noise $n(t)$. The later can be highly challenging when the amplitude of the GW strain is very weak and/or when the noise is deeply covering the real signal. We quantify the

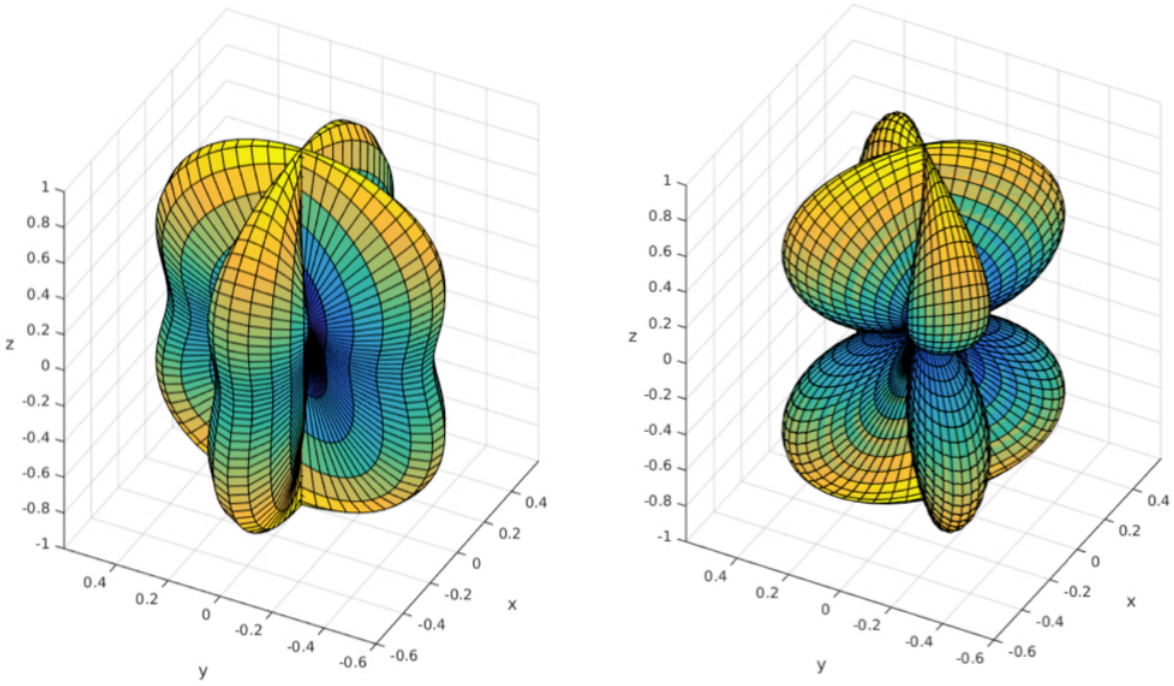

Figure 5.2: Antenna response pattern of an L shaped Michelson interferometer, whose arms are parallel to the x-y plane. The color shows the change in sensitivity, from blue (low) to yellow (high). **Left:** Plus antenna response pattern F_+. **Right:** Cross antenna response pattern $F_\times$. Figure taken from [215].

amount of noise as a function of the frequency through a quantity known as the power spectral density (PSD) denoted $S_n(f)$. The PSD, given in Hz^{-1}, tells the amplitude of the oscillations of a signal at a specific frequency. When we talk about detector sensitivity, it is characterized by the PSD curves, which highlights the frequency boundaries in which a GW with a certain strain can be detected. Following [216], in the Fourier domain, if on average the noise time-serie is equal to zero, the PSD relates to the time average of the noise $n(t)$ squared such that

$$< n^2(t) >= \int_0^\infty S_n(f)df, \tag{5.3}$$

where $S_n(f)$ is the PSD and <> is a time average. When discussing noise detector, we also often refer to the amplitude spectral density (ASD), which is defined as the square root of the PSD

$$A_n(f) = \sqrt{S_n(f)}. \tag{5.4}$$

Fig. 5.3 displays the sensitivity curves of the LIGO Hanford, LIGO Livingston and Virgo detectors, seen through the ASD during the second half of the third observing run (O3b). The frequency bandwidth in which the current detector is sensitive is between $[20 - 4000]\,Hz$, but over this large frequency range, the sensitivity can be affected by various noises. The specific noise sources are clearly visible on the ASD curves, the spikes in Fig. 5.3, also called spectral lines, depict the loss in sensitivity of

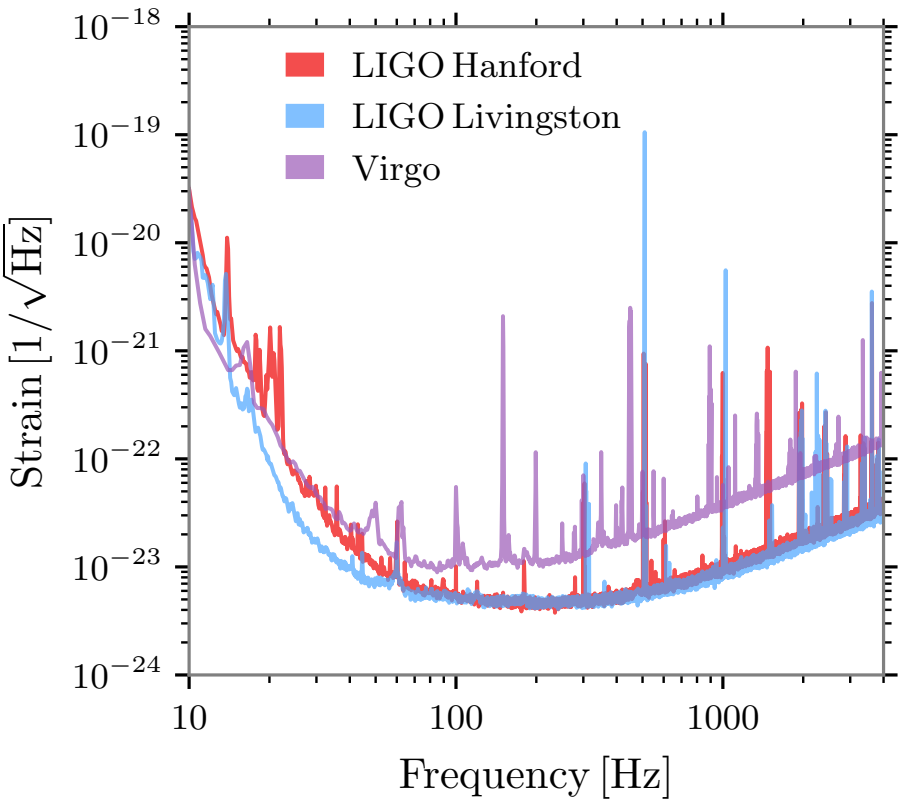

Figure 5.3: Amplitude spectral density (ASD) of the LIGO Hanford, Livingston and Virgo interferometers during the second half of the third observing run. ASD estimated on the 4th of January 2020 at 02:53:42 UTC for LIGO and on the 9th of February 2020 at 01:16:00 UTC for Virgo. Figure taken from [27].

a detector at a certain frequency due to specific noises. Depending on the frequency range considered, different type of noise will be dominant, affecting in various manner the sensitivity of a detector. Noise can have different origins, and some of them are very well characterized like the seismic noise or the electric power noise, but a significant number of noises are still unknown or uncontrolled.

5.2.3 Auxiliary channels

Besides measuring the strain $h(t)$, terrestrial interferometers such as LIGO or Virgo record about $\mathcal{O} \sim 10^5$ other data time-series per detector. These other data channels are called "auxiliary" channels. They aim to measure everything that could impact the sensitivity of the detectors, from the internal state of the interferometer to the external environment around it. The auxiliary channels are used to quantify the quality of the data as a function of time during the observing periods. Beside the state of the interferometer, they can also sense noises which could be present in the GW strain channel, like scattered light on the mirrors or fluctuations of the laser power. Ideally,

the main sources of noise which are witnessed by the auxiliary channels can then be subtracted to the strain, if identified.

Often, the main auxiliary channels are used to give alerts or flag periods of time when the interferometer strain channel is known to be affected by strong noise. This procedure called "vetoing", helps to identify times when the detector data can not be trusted, allowing analysts to know when they should or should not use the produced data. The auxiliary channels of an interferometer are classified into two categories: the "safe" and "unsafe" channels. A channel is considered "safe" if it remains unaffected by the passage of a GW through the detector, meaning it does not register or sense the GW signal. Channels that might be sensitive to a GW signal are flagged as "unsafe" to prevent any confusion between real GW events and instrumental noise. During the vetoing process, only the "safe" auxiliary channels are used to rule out noise artifacts. This ensures that genuine GW signals are not mistakenly vetoed due to a misinterpretation of noise as a GW event.

5.3 A BESTIARY OF NOISE SOURCES

Noise is the result of sources of no astrophysical origin that can produce in the detector output an effect similar to a strain. These noises can have several origins that can be divided into fundamental noises such as quantum noise or thermal noise and environmental noises as the Newtonian noise or the seismic noise or technical noise. In this section, we propose a review of the different sources of noise which impact the sensitivity of the detector. In particular, we discuss glitches which are characteristic by they short time duration and high statistic.

Fig. 5.4 presents the simplified noise budget of the Advanced LIGO Hanford interferometer ASD during the O3 run. The noise budget is a good way to quantify the different noise contributions and sources, to the total noise of the detector.

5.3.1 Fundamental noise

Any physical effect that can impact the position or stability of the mirror can be a source of noise. These noises, directly related to the physical principle of detection with a Michelson interferometer are called fundamental noises. In the frequency range in which GW are supposed to be, the main contributions to the fundamental noise are the thermal noises and the quantum noises.

Thermal noise is a large class of noises which is characterized by the random thermal agitation of the atoms and molecules constituting any part of the detector, from the mirrors themselves, to the suspensions or the cavities. The basic reasoning of thermal fluctuations as being a source of noise is the following, the higher the temperature and the greater the mechanical losses (internal friction), the greater is the agitation/noise. Virgo, specifically, is operating at room temperature, which is not the obvious choice when trying to reduce the thermal noise. The strategy so far has been to use only high

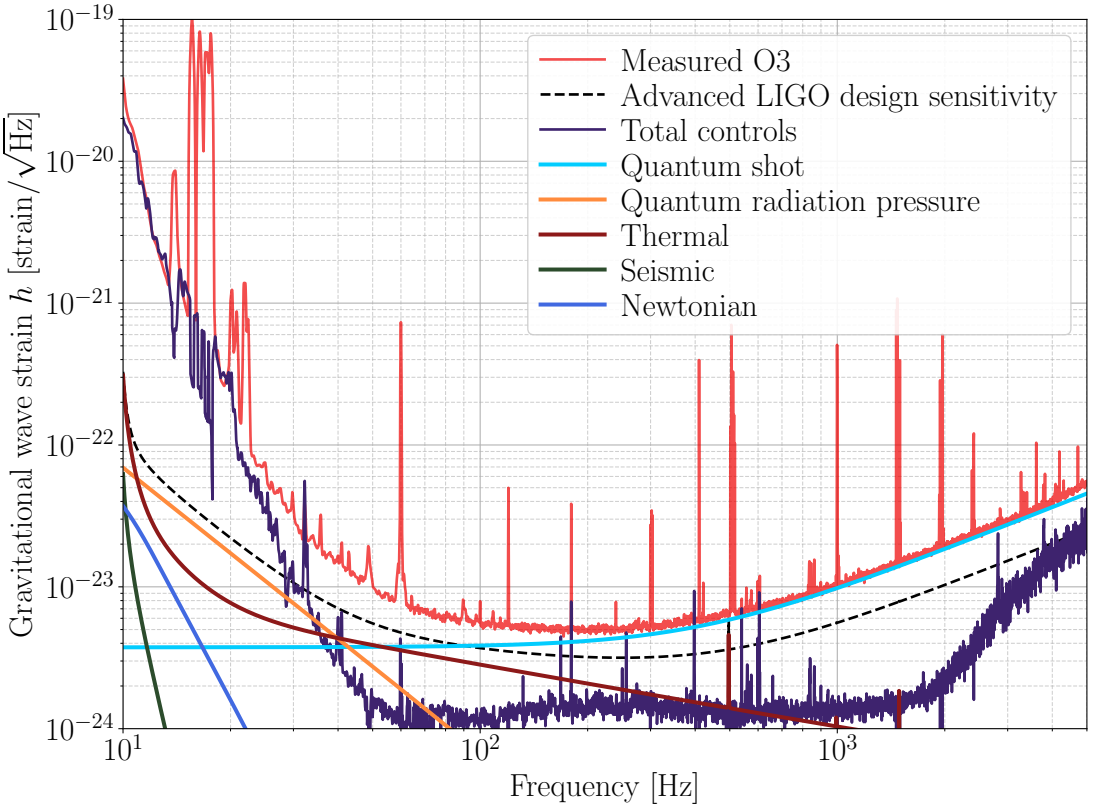

Figure 5.4: ASD of the O3 LIGO Hanford interferometer simplified noise budget. The red curve is the measured sensitivity to GWs, the dashed black line is the Advanced LIGO design sensitivity and all the other colored curves are the estimated noise contributions to the total noise. Only the main known sources of noise are depicted in this plot. Figure taken from [217].

quality materials, with very low internal friction. In Virgo, the main contribution to the thermal noise is coming from the coating of the mirrors and their suspensions in the arm cavities. This noise results in a random vibration of the mirrors surfaces, crucial pieces of the detector. The coating Brownian thermal noise in particular, limits the detector's sensitivity over a large frequency range $[40 - 200]Hz$, the noise originates from the mechanical friction of the different layers making up the coating of the mirrors due to thermal fluctuations. The displacement of the coating layers impacts the phase shift of the laser beam and then affects the GW detection efficiency. In addition to the material with low internal friction, the laser beam is usually focused on a large spot on the mirrors, so that the random fluctuations are spread out over a bigger area and averaged away. On Fig. 5.4, the thermal noise contribution is given by the brown curve.

The quantum noise, represented by the light blue and oranges curves on Fig. 5.4 is one of the most important limits of the total sensitivity, and in particular at medium and high frequency. The total quantum noise is the combination of two quantum

mechanisms, the radiation pressure noise (amplitude fluctuations) at low frequencies and the quantum shot noise (phase fluctuations) at high frequencies. Both of them are the result of the intrinsic property of the light being made of discrete energy particles called photons. The photodetector which quantifies the light power counts these photons, and these photons also hit the mirrors when the laser beam is reflected on them. The radiation pressure noise emerges from the fluctuations of mirror position due to the radiation pressure of the photons. The displacement of the mirror created by this pressure induces a potential phase shift which is detected by the photodiode. The shot noise on the other hand, is due to the variations of the arrival time of the photon on the photo diode, arrival time that follows a Poisson process. The lower the power of the laser, the higher is the effect of shot noise in the detector, but the higher is the contribution of the radiation pressure noise. This relation is also called the standard quantum limit, understood as an effect of the Heisenberg uncertainty principle. The perfect balance between the two quantum noises has to be found in order to maximize the sensitivity of the detector, at both low and high frequencies. For a specific GW frequency, there exists a better configuration of the input laser power which minimize the combined contribution of both quantum noises. At the end of the day, the quantum noise is the main limitation for all frequencies above $\sim 30\,Hz$.

5.3.2 Environmental noise

On top of the fundamental noises arising from sources internal to the GW interferometer, the natural environment in which the detector is located is also a source of noise on its own. We talk about environmental noise sources, such a seismic motion, earthquakes, electromagnetic noise or even the change in the local gravitational field. As an example, the Virgo site, located near the city of Pisa in Italy, is regularly shaken by the motion of the ground on which it is constructed. Due to the extremely high sensitivity of the detector, there is no need for high magnitude earthquakes for the detector to be disturbed by the Earth motion. The slight, but constant vibration of the ground impacts the steadiness of the suspended mirrors. To fight against the effect of seismic noise, the mirrors are suspended in a tower by several pendulums. These columns manage to reduce the effect of seismic noise by a factor 10^{12}, for all frequencies above $10\,Hz$, below this threshold value the suspensions are not able to reduce the seismic disturbances. We call this limit the seismic wall, which prevent any detection of GWs below $10\,Hz$.

Extra contributions to the environmental noise can also be attributed to the weather changes, like high velocity winds or sea activity like powerful tides that can create seismic noises (sea microseisms). The sea microseisms contribute to the low frequency range noise, they can also disturb the instrument components. The GW detectors are also sensitive to lightning strikes and bad weather, these can affect the instrument via electromagnetic interferences or pressure waves in the ground.

Finally, the Newtonian noise, present at low frequencies, is one additional noise that can be classified as environmental. This noise is created by the local fluctuations

over time of the gravitational field near the mirrors of the interferometer. These local variations of the gravitational field induce a Newtonian interaction which can push the mirrors hence reducing the sensitivity. These local variations of the gravitational field can arise from atmospheric density fluctuations, pressure waves, temperature variations or density changes due to seismic waves. The Newtonian noise is solely affecting the 1 to 10 H_z region, but since it is a direct interaction of gravity, there is no known shielding that can counterbalance this effect.

5.3.3 Technical noise

The technical noise, or technical "control" noise arises from all the control systems and sensors that are used to keep the interferometer optical cavities on resonance (the lock position). As shown in Fig. 5.4, this noise greatly impacts the total noise budget of the interferometer at low frequencies, but also at high frequencies. Below 60 Hz, the technical noise is the dominant source [217]. The working state of an interferometer is not a natural stable state, the optical layout has several degrees of freedom, longitudinal and angular, that prevent the system to be on resonance mode on its own. These complex control systems are used to reduce the real motion of the optics (displacement noise) or even the proper noise of the output photodiode (sensor noise). The sensor noise is hardly identifiable from the displacement noise, but dominates above 50 Hz. The optimal resonance state of an interferometer is very unstable, and multiple controllers are employed to hold the optimal position in the most permanent way.

The control noise is predominant at low frequencies, the purple curve on Fig. 5.4 represents the sum of the noise produced to suppress the displacement noise and the sensor noise.

5.3.4 Transient noise

Glitches is the name given to peculiar type of non-Gaussian noise characterized by its short durations (transient), ranging from a few tens of a second up to several seconds. They can mislead astrophysical searches by producing even high signal-to-noise ratio (SNR) triggers [218–221]. The origin of such glitches can be various, but they are usually associated with environmental disturbances and control systems noises. These short transient noises can affect directly the sensitivity of the detectors, limiting the sensitivity at different frequency ranges. However, they also affect the ranking statistics of the GW detections like the false alarm rate (FAR). The FAR corresponds to the expected rate of signals/triggers that comes from background sources. As glitches can resemble real GW events, the associated FAR of certain signals in the detector will be affected by them.

As depicted in Fig. 5.5, glitches come in various types, usually they are categorized in families (Blip, Tomte, Helix, ...) based on their shape seen in the time-frequency domain, and their SNR. The presence of glitches in the GW data stream and their suppression is a complicated task, which is typically done based on the data streams

of the auxiliary channels of the detectors (see Sec. 5.2.3 for more details on auxiliary channels) that monitor numerous control sensors, environmental sensors or instrumental sensors. The process of identifying glitches in the GW data stream is called "noise-hunting". If a glitch appears in the GW data stream channel, its origin could be coupled to some other auxiliary channels of the detector. The auxiliary channel that witnessed the glitch is then used to remove it from the GW strain. There exists a wide range of techniques used to identify/remove glitches, like band-pass filters for specific frequencies or data quality vetoes ("gating") to remove chunks of data over a time period polluted by noise and sets to zero the data [222]. All these techniques are able to remove part of the glitches in the data stream, but they have several downsides. First, processes like gating or vetoes are conservatives methods, the time period around a supposed glitch is removed from any analysis, which can slightly reduce the total observation time. Second, the entire noise-hunting process for glitches can be very tedious, with numerous steps and incremental vetoes to be run over the data stream. Thirdly, glitches can overlap real GW signal. Lastly, not all the glitches are found with these methods.

Since the advanced detector era, with the increase number of detections and sensitivity of the detector, the number of triggers due to the presence of glitches has not especially increased but is more likely to be superimposed with a real astrophysical trigger and multiple new methods based on ML were developed. The rest of this chapter focuses on one of these methods for glitch identification, both as a tool for a new data quality veto and its potential impact in the GW search pipelines.

5.4 MACHINE LEARNING ALGORITHM FOR DATA QUALITY

In this section, we present an exploratory study based on a machine learning algorithm (iDQ) to improve the data quality of the Virgo interferometer data stream [224]. This algorithm has already been used in low latency at both LIGO Hanford and Livingston during the third observing run, but never for the Virgo interferometer. We propose tests to quantify the efficiency of iDQ to identify glitches in the GW data stream, comparing its performances with methods currently used by the LIGO and Virgo collaborations. Finally, we propose a method to include the data quality output of iDQ into the GW search pipelines.

5.4.1 iDQ: A supervised learning algorithm

iDQ is a supervised learning algorithm that incorporates a framework for the statistical inference of glitches. As mentioned in the previous sections, an interferometer has on top of the main GW channel, several auxiliary channels ($\mathcal{O} \sim 10^5$), that record the state of the interferometer at all time with a high frequency rate. These channels can be the witnesses of numerous noise sources, noises that may also couple to the GW data stream, creating problems for the GW searches analyses. Out of the total number

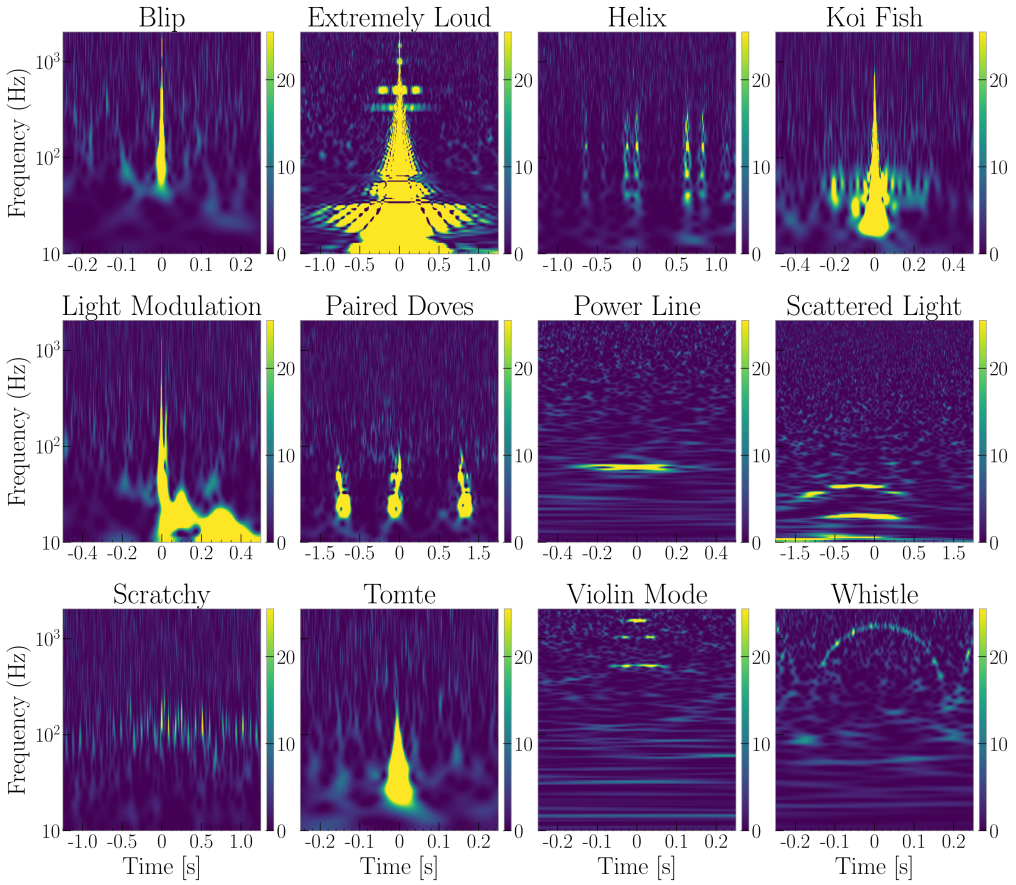

Figure 5.5: Time-frequency visualization of 12 different type of glitches from various origins, that can be observed in the GW data stream. The color gradient shows the associated signal-to-noise ratio (SNR). Figure taken from [223].

of auxiliary channels, some of them are safe, i.e. they are not affected by the passage of a real GW signal. This property makes them very suitable to do noise-hunting, by looking at correlations between the safe auxiliary channels and the main one. Because of the very large number of channels, and the complexity of the interplay between these channels, classical algorithms can have a hard time disentangling the presence of glitches. The goal of iDQ is to produce probabilistic quantities about the amount of noise across time, in the interferometer data stream, quantities that could translate to a probability of having a glitch at a certain time.

Before diving into how the iDQ framework is built, we clarify some nomenclature and technical terms specific to this search, following [225]. First, we call the target channel $h(t)$, the channel supposed to contain both the real GW signals and the pollution due to the presence of non-gaussian noise that we aim to flag. The target channel

can be any proxy of the GW data stream channel. Second, the safe auxiliary channels $\vec{a}(t)$, refers to the list of channels beside $h(t)$ that are not sensitive to GW signal, hence safe to use for flagging the presence of coupled noise between them and $h(t)$. The safe channels are determined through an injection campaign, power excess are injected into the interferometer to simulate the effect of a GW signal and the channels sensitive to these injections are categorized as unsafe. iDQ framework only uses safe channels in its inference. From a more mathematical approach, iDQ assumes that the target channel can be decomposed into three terms

$$h(t) = n(t) + s(t) + g(t), \tag{5.5}$$

where $n(t)$ is the typical Gaussian noise made out of all the noise sources presented in Sec. 5.3, $s(t)$ are the real GW signal and $g(t)$ are the short transient non-Gaussian noises. The goal of iDQ can be summarized in distinguish $g(t)$ from $n(t)$. The target channel is the only data stream that can really be monitored, the other can only be modeled- the Gaussian noise is supposed stationary, meaning that it can be completely characterized by the ASD, but not $s(t)$ and $g(t)$. iDQ works with a two-class classification, where G defines all time samples with possible non-Gaussian noise and C all clean sample times, such that

$$p(G) + p(C) = 1 \forall t. \tag{5.6}$$

iDQ, based on a set of data, estimates the probability $p_G(t)$ which denotes the presence of non-Gaussian noise in $h(t)$, based on correlations with the safe auxiliary channels. This probability can be expressed as:

$$p_G(t) = p(G|\vec{a}(t)) = \frac{p(\vec{a}(t))p(G)}{p(\vec{a}(t)|G)p(G) + p(\vec{a}(t)|C)p(C)}, \tag{5.7}$$

where $p(\vec{a}(t)|G)$ and $p(\vec{a}(t)|C)$ are inferred through the following supervised learning process.

The data stream from the interferometer auxiliary channels are discretely sampled time series, iDQ extracts from these time series a set of features from each channel, typically $\mathcal{O}(10^5)$ per channel, and creates a high dimensional array based on those, representing the auxiliary state of the interferometer. For the two-class classification of iDQ, the Neyman-Pearson lemma [226] stipulates that an optimal classification should be done based on the likelihood ratio

$$\Lambda^G_C(\vec{a}(t)) = \frac{p(\vec{a}(t)|G)}{p(\vec{a}(t)|C)}. \tag{5.8}$$

Unfortunately the analytical forms of these probability functions are not known and the high dimensionality of the safe auxiliary channel array $\vec{a}(t)$ ($\mathcal{O} \sim 10^4$) makes the estimation of likelihood ratio from the observed sampled complex. In this context, the supervised learning algorithm is of great help since it can produce a map from a high dimensional space to the one dimensional unit space $[0, 1]$. From this inference,

elements of G should be mapped close to 1 and elements of C close to 0. Following [225], the likelihood ratio in Eq. 5.8 transforms in a new one such that

$$p_G(t) = \frac{p(\mathcal{M}(\vec{a}(t))|G)p(G)}{p(\mathcal{M}(\vec{a}(t))|G)p(G) + p(\mathcal{M}(\vec{a}(t))|C)p(C)}, \tag{5.9}$$

where $\mathcal{M}(\vec{a}(t))$ is the mapping from high dimension to the unit space. The new probabilities are then estimated from the labeled C and G samples. Beside the likelihood ratio, iDQ is also able to produce other ranking statistics as a false alarm probability (FAP) or efficiency.

5.4.2 iDQ's supervised learning workflow

The workflow of iDQ inference can be decomposed into four stages, the training, the evaluation, the calibration and the output production, and some of them happen simultaneously in order to communicate updated models. iDQ takes as input three separate sets of data, Omicron triggers on the target channel $h(t)$ from which potential times with non-Gaussian noises will be flagged, the list of safe auxiliary channels for a specific interferometer and an external source of features giving the time localization and SNR of non-Gaussian noises in channels. In our exploratory work, these external features are taken to be Omicron triggers [227]. An Omicron trigger is a signal identified by the Omicron software after applying a signal-to-noise ratio threshold to a spectrogram generated through a Q-transform of the time series data, indicating a potential event or feature in the input signal. A pre-processing to flag important features helps the iDQ classification, and allows for a first classification of time segments in $h(t)$ into "clean" and "noisy" segments. Each segments with labels C or G are then used for training, and evaluation.

The second stage is the training phase, where iDQ trains the different classifiers to distinguish C and G samples, the training is only made on the auxiliary channels. iDQ can use different types of classifiers, in the following results, we choose to use the already incorporated classifier OVL [228]. From the auxiliary channels, iDQ manages to generate models, to be understood as different mapping from the high dimensional array to a scalar $\in [0, 1]$ called the rank. The trained models from the classifier also save the information on the range of data used for the training, this information will be used in the evaluation stage.

The evaluation phase quantifies the classifier performances. As said previously, the different phases of iDQ's workflow can happen simultaneously (cross-evaluation), the data are separated into several subsets, where one of them will be sent for training and the other for evaluation. In the iDQ setup we used to analyze the Virgo data stream, we chose the "round-robin" method for subdivision. The data are divided into bins and then segments, iDQ will then train on time segments in some bins, to generate models which will evaluate the data in some other bins. This cross-validation is iterated, until all bins are used.

The last stage before the iDQ production of the final output is the calibration. The calibration consists of the transition between ranks to probabilistic quantities about the non-Gaussian noise content in the targeted channel. This is done by estimating the conditioned probabilities seen in Eq. 5.9 for each model's rank. The combination of $p(\mathcal{M}|G)$ and $p(\mathcal{M}|C)$ are called calibration maps.

Finally, the iDQ workflow generates time series of probabilistic quantities from the training and calibration phases. These probabilistic time series are the final product of iDQ, and also the ones that should be incorporated into GW searches. Beside $p_G(t)$, iDQ also estimates quantities such as the likelihood ratio Λ^G_C, the false alarm probability, which are sampled at 128 Hz. Physically, $p_G(t)$ is the probability of having a glitch in the target channel, based on correlations with the auxiliary channels of the interferometer. Λ^G_C indicates the condition of the auxiliary channels of the interferometer for both noisy and clean samples, large values of the ratio are more likely to be associated with noisy times and small values with clean times. The false alarm probability, cumulative integral of $p(\vec{a}(t)|C)$, shows the probability of wrongly labeling clean times as noisy times. Low FAP values indicate confident triggers, times with a high probability of having no glitches in the target channel.

5.4.3 Machine learning versus classical methods

The main advantage of using ML methods instead of more "classical" approaches is not merely their ability to handle large amounts of data or to do so more quickly with respect to computational resources. The primary benefit of ML methods like iDQ lies in their capacity to predict complex, non-linear relationships between the auxiliary channels and the target channel, whereas classical algorithms are generally limited to identifying linear or quadratic relationships [229].

The assertion that ML methods can handle large amounts of data is mainly a computational consideration, as both ML and classical algorithms can be designed to process large volumes of data, either correctly or incorrectly. Moreover, from a computational perspective, ML methods employing neural networks with multiple layers and numerous nodes—each performing regression statistical processes—are significantly more complex than classical algorithms in terms of both calculation time and memory usage. The key issue is which method produces better results, specifically, in our case, which one identify more glitches, with the best efficiency.

The fundamental advantage of ML algorithms over classical methods is their ability to identify and analyze complex patterns in the data, especially non-linear ones. Non-Gaussian noises present in the auxiliary channels do not necessarily have a linear coupling to the targeted channel. Additionally, various combinations of small but numerous glitches in the auxiliary channels can collectively create an effect visible in the target channel. Conversely, loud glitches in one of the auxiliary channels might have no impact on the target channel. Classical noise-hunting algorithms struggle to handle such situations, but ML methods can address them effectively. A neural net-

work provides enough versatility and flexibility to predict these cases by learning the patterns and relationships between glitches in the auxiliary channels and the target channel through training on known cases.

This is why the training phase in ML methods is critically important; the performance of the algorithm directly depends on the quality of the training set. With GW data from terrestrial interferometers like Virgo, we have numerous identified glitches available for training. Through repetition, iDQ learns to predict glitches similar to the tens of thousands it has previously seen and analyzed. A subtle aspect of working with interferometers is that the total noise is considered non-stationary, meaning that patterns change significantly over time as the state of the detector changes. Therefore, it is not appropriate to train iDQ over very long periods but rather over short ones, using it to make predictions while the detector's state remains relatively unchanged.

In summary, the advantage of using ML methods to hunt non-Gaussian noises lies in their ability to uncover complex relationships that classical approaches cannot detect. The iterative training process of iDQ is crucial for its effectiveness in identifying glitches in an interferometer's data set, which changes constantly over time, rendering simpler methods ineffective.

5.5 IMPROVING THE VIRGO DATA QUALITY WITH IDQ

During the third observing run of the LVK detector network, the Virgo interferometer used several techniques to improve the data quality for analysis purposes. As we mentioned previously, these gating techniques, also called vetoes, were introduced in order to clean the data stream from most of the unwanted noise, by flagging time segments as "unusable" for data analysis purposes. In Virgo, all these flagging techniques are referred to as CAT1 vetoes. CAT1 veto labels are given to periods of time when parts of the interferometer are not functioning properly. Although, CAT1 vetoes are not efficient against glitches, during the long period where the state of the interferometer is in"science" mode, the detector can still be affected by the presence of non-Gaussian noises coupling to the GW channel. In this section, we explore the possibility of using iDQ, first in high latency (post-processing of the data), and as an additional veto (CAT2) for the Virgo interferometer data quality. The supervised learning method of iDQ could increase the confidence level of GW-events by flagging periods of time with large non-Gaussian noise activity, and globally improve the data quality and give insight for the commissioning with its products on the auxiliary channels that may have witnessed the glitches.

5.5.1 Test on the Virgo O3 observing run

To assess the performance of iDQ, we ran an offline iDQ analysis, using the OVL classifier included in iDQ, over four weeks of data produced by the Virgo interferometer. These four weeks, between April 1st and May 4th 2019, correspond to a period

of time when Virgo witnessed numerous noise and transient signals, making it an interesting period to test iDQ performances. These four weeks are defined between $GPS_{start} = 1238166016.000$ and $GPS_{end} = 1241000000.000$.

Up to now, iDQ has been solely utilized for the LIGO interferometers, so we also had to implement a new version designed for the Virgo interferometer, configured with its specific set of safe auxiliary channels and data stream. We successfully implemented a Virgo configuration of iDQ, first at the computing center of Cascina (at the Virgo site) in Italy and then locally at the CCIN2P3 in Lyon, FRANCE where more auxiliary channel data streams are stored. We find that during the analyzed period, on average, close to ~1.5% of the time was identified by iDQ as "noisy" time, with a probability of having a glitch in the proxy of the GW channel was greater than $p(G|\vec{d}(t)) \geq 0.9$. This arbitrary threshold for the probability of 0.9 denotes the periods of time when iDQ finds a high support for the presence of non-Gaussian noise in the detector. Fig. 5.6 shows the report of the statistical time series sampled at 128 Hz, produced over ~2.3 days of data. Each row corresponds to one of the probabilistic quantities of iDQ, namely $p(G|\vec{d}(t))$,

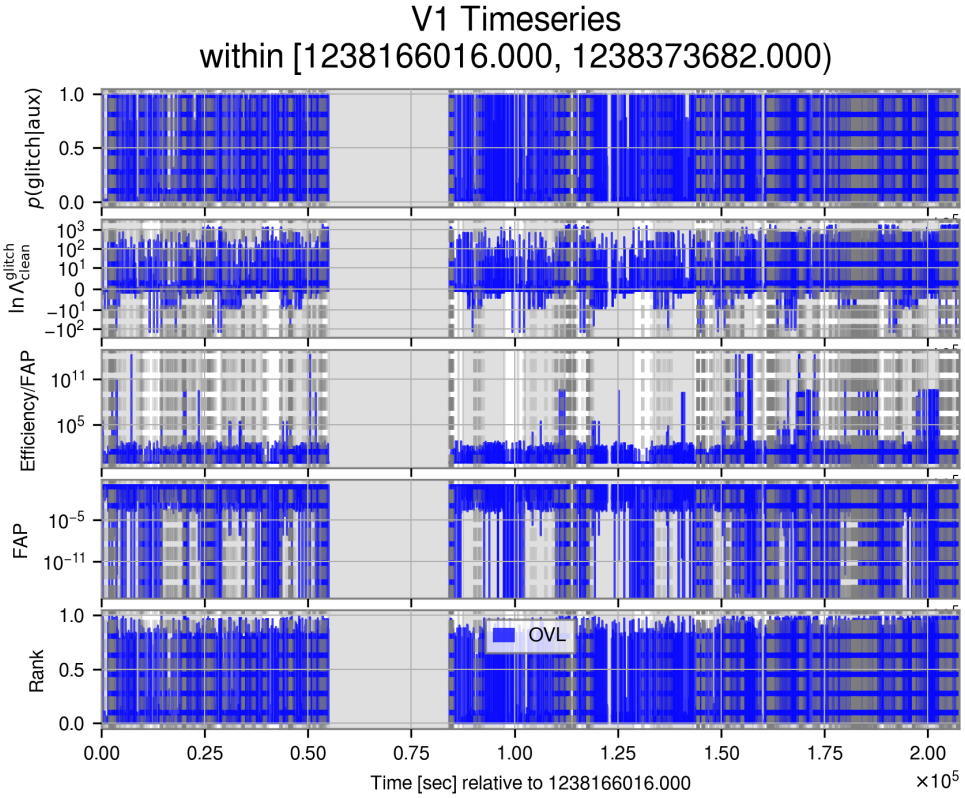

Figure 5.6: Report of the probabilistic quantities produced by iDQ inference over 2.3 days of Virgo data, during the third observing run. The vertical gray area corresponds to a time period when the interferometer was not in science mode (not taking data).

$ln(\Lambda^G_C)$, FAP and the rank. These quantities can then be used as extra information for the GW searches.

In addition to the ranking figures of merit discussed above, iDQ also generates a "feature importance" summary. This summary ranks the different safe auxiliary channels that may have witnessed glitches from most likely to less likely. Fig. 5.7 displays the feature importance summary for a subset of time analyzed in the third observing run of Virgo. From this result, important information about the source of the transient noises can be deduced. For instance, at a given time t, we can look at which auxiliary channels seem to be the most correlated with the presence of non-Gaussian noise in the target channel. In the example shown, the channel "$V1 : ENV\_WI\_ACC\_X$" presents significant correlation with the presence of a glitch. This channel corresponds to an environmental sensor on the west input tower, measuring acceleration in the direction perpendicular to the laser beam of the detector. Two diagnosis are possible, either there was already some evidence for the presence of a glitch at this time, and the feature importance can help us track to possible origin of noise- or the trigger was thought to be from astrophysical origin and iDQ can help us flag this period of time as likely noisy. Therefore, this product of iDQ may be of interest as an additional tool for commissioning and for noise-hunting purposes, but also to increase the confidence of detected signal with a more in-depth analysis around the time of the trigger.

Beside looking at specific time periods for the presence of glitches in the target channel, it is also interesting to look at the feature importance plot around periods when the global sensitivity of the detector drops. Between the 28th and the 29th of April 2019, Virgo's BNS range was very unstable and an important decrease of the range was observed around midnight (see Fig. 5.8), from $\sim 40Mpc$ to $\sim 20Mpc$. Based on iDQ's inference around the same time, we managed to track possible auxiliary channels of the interferometer that show high probability for non-Gaussian noise correlating with the $h(t)$ channel, namely $V1 : SDB2\_B5\_56MHz\_I$ and $V1 : SDB2\_B1s2\_DC$. Even if a more robust case could be made with a more detailed analysis, iDQ already shows potential as a new tool for data quality and glitch hunting. As for this drop in the BNS range, other algorithms also found evidence for an excess of non-Gaussian noise perturbing the detector sensitivity.

5.5.2 Inclusion of iDQ for gravitational wave searches

Another prospect for the iDQ inference, is to incorporate some probabilistic quantities about the data quality from iDQ, into the GW transient search algorithms. Up to now, this type of analysis has been only partially studied in the literature [230]. In this section, we explore the effect of a post-processing of the SNR assigned to single triggers obtained with the PyCBC GW search pipeline [135, 231] on O3 data. Currently, data quality inputs are either incorporated after a search pipeline has produced a list of trigger candidates, or prior to the analysis. Periods of time labeled as noisy are typically excluded, and all triggers falling within these periods are then removed from the analysis. Alternatively, before a search analysis, time segments labeled as unsafe are "gated" and not analyzed by the GW searches at all. Hence, an important caveat

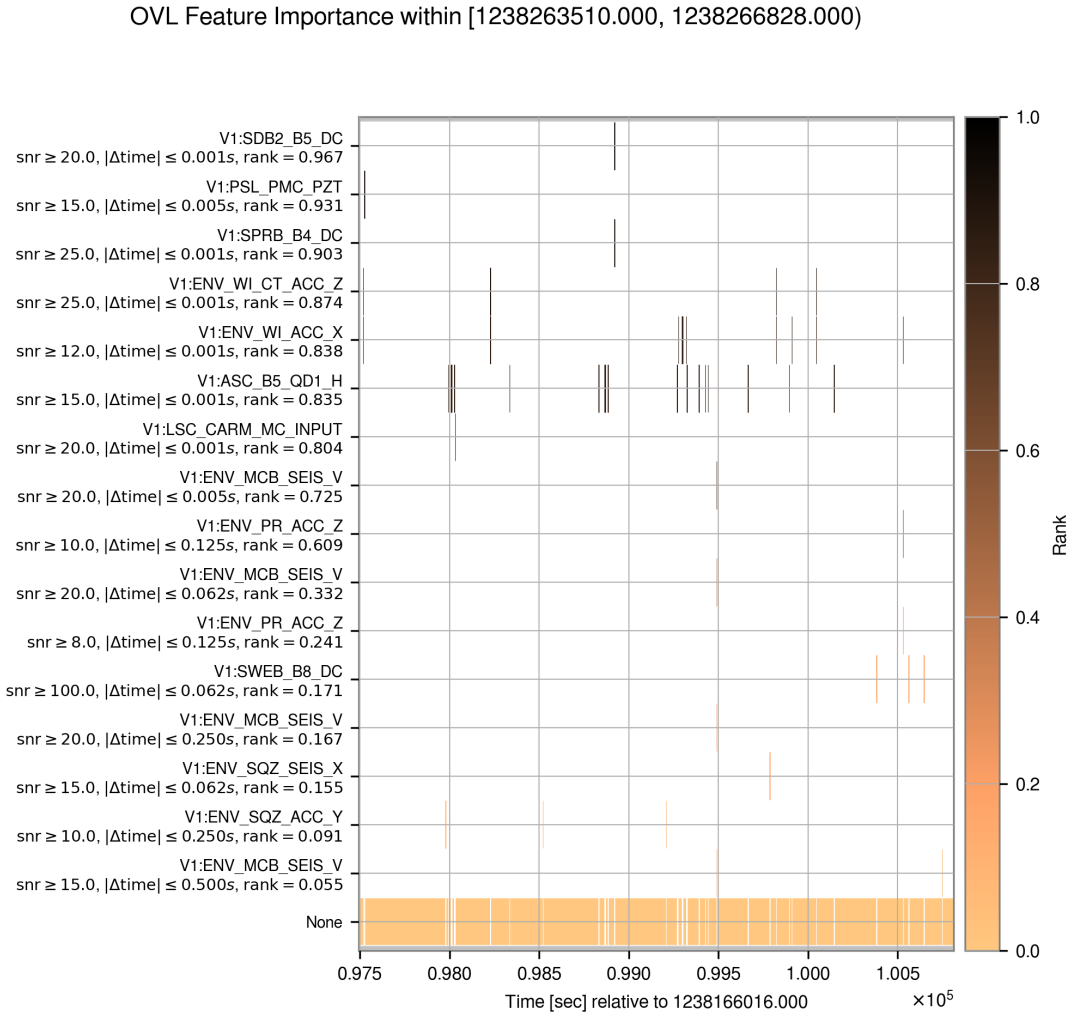

Figure 5.7: Feature importance summary plot produced by iDQ. Each row corresponds to one auxiliary channel, identified as the one more likely to have witnessed the presence of the non-Gaussian noise in the data. The color gradient shows the importance of each channel, from 1 being highly probable, to 0 highly unlikely.

of vetoes is its intrinsic binary formulation, if a time segment is not considered clean, it is then flagged as noisy and thrown away.

We propose here to include a more continuous data quality estimation to a GW search, using the likelihood ratio inferred by iDQ. The likelihood ratio Λ^G_C, as defined in Eq. 5.8, will take large values when a time is considered noisy, and small ones when the presence of a glitch has no support. In other terms, it gives an insight on the probability of having a glitch at a specific time. A GW search usually produces a list of GW candidates with an associated ranking statistic which determines the confidence of this candidate being from an astrophysical source. Generally, each trigger is given an SNR value measurement from match filtering techniques [51]. Other ranking statistics than the SNR are often derived, like the false alarm rate or p_{astro}. In this exploratory study, we only focus on the SNR of each candidate, we test the effect of re-weighting

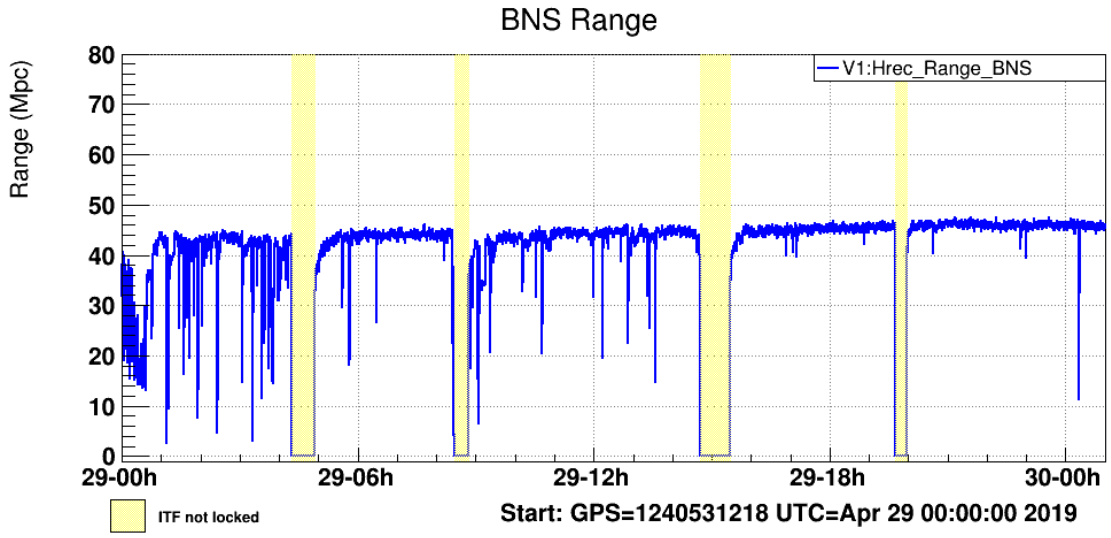

Figure 5.8: Binary neutron star range in Mpc of the Virgo interferometer on the 29th of April 2019. The blue curve shows the BNS range as a function of time and the yellow contours the periods when the interferometer was lock locked. Credits: The Virgo scientific collaboration

the SNR of each candidate based on the data quality information that iDQ produces. We apply the following relation to each trigger:

$$SNR_w = \sqrt{SNR_t^2 - 2ln(\Lambda_C^G)}, \tag{5.10}$$

where SNR_t is the SNR of the trigger computed with the match filtering, Λ_C^G is the likelihood ratio from iDQ and SNR_w is the reweighed SNR. Because the likelihood ratio is sampled at $128Hz$, we average its value around the time of the trigger. The data quality information is applied a posteriori to the GW search, here we do not reject any candidates, but rather increase or decrease the confidence we have about them being real astrophysical signal or excess power induce by the presence of a glitch.

We performed a cross-comparison between real single trigger produced by the LVK GW search pipeline PyCBC [135, 231], and injected fake signals recovered by another GW search pipeline MBTA [133, 232]. Fig. 5.9 shows the SNR values of the recovered single triggers from the PyCBC analysis with an SNR detection threshold of 5 (left) and from MBTA on injected signals with an SNR detection threshold of 8 (right). For both searches, the analysis was done on the chunk number four of the third observing run, on approximately $700000\,s$ of data. The sample of PyCBC triggers is very large with almost $\sim$ 60000 candidates, and most of them are supposed to be associated with background noises (not real GW events). On the other hand, the MBTA triggers are much less numerous ($\sim$ 900) and correspond to fake GW signals randomly injected in the instrumental noise from real data of the same chunk in O3. If iDQ is able to correctly identify times when multiple transient signals polluted the data stream, the reweighed SNR based on the likelihood ratio should on average affect more the PyCBC triggers than the MBTA triggers. Since the PyCBC candidates arise from noise and the MBTA are not injected at specific noisy-time.

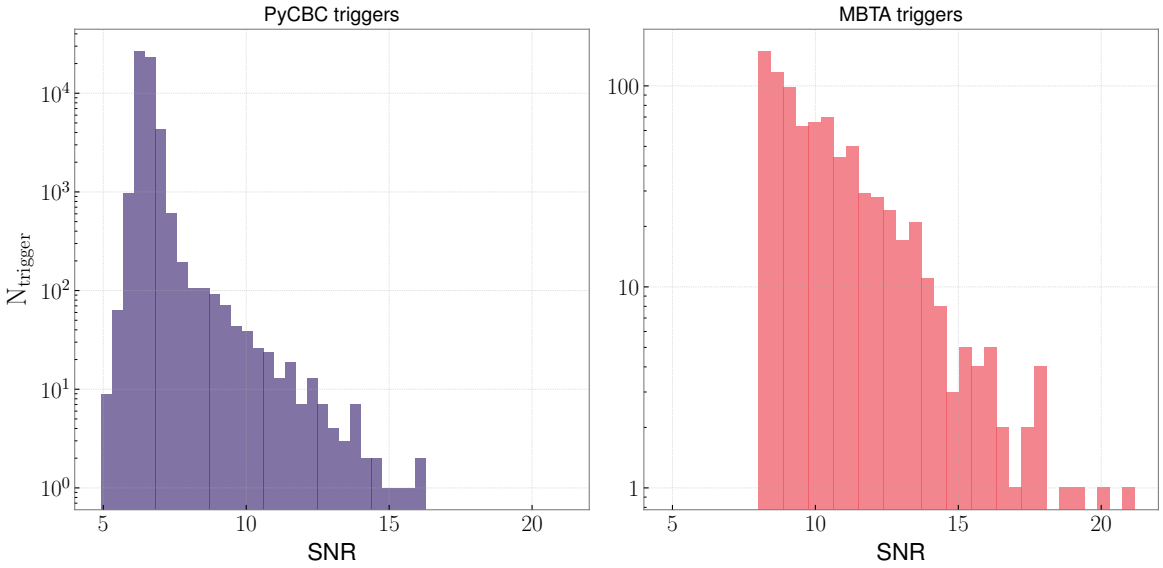

Figure 5.9: Histograms of the single trigger SNR for the chunk 4 of the third observing run. **Left:** SNR of real triggers discovered by the PyCBC pipeline with an SNR cut of 5. **Right:** SNR of injected signal recovered from the MBTA pipeline with an SNR cut of 8.

Fig. 5.10 shows the reweighted SNR_w against the measurement SNR_t of each trigger for both PyCBC and MBTA. As expected, we observe that the new SNR of the PyCBC triggers are significantly lower after weighting them with the likelihood ratio of iDQ. This result supports the fact that a large proportion of these real triggers fall in periods of time when the interferometer was highly affected by non-Gaussian noise. Additionally, we observe the opposite behavior for the SNR of MBTA's triggers, which are downgraded by an amount consistent with the hypothesis of no correlation with iDQ. Similarly, it supports the idea that the MBTA triggers were injected at random times, so some of them fell during noisy time, but the majority during clean time. To further verify that the impact on the SNR is not a statistical effect, we tested the same procedure, but this time randomly shifting in time the positions of the PyCBC and MBTA triggers. We find that the effect on the PyCBC triggers is drastically reduced, in agreement with the fact that these triggers were indeed associated with noisy time before we shifted them. The result on the MBTA triggers did not change, this corroborates the fact these triggers were already associated with random times.

From this analysis, we demonstrate that iDQ is correctly able to flag periods of time when the interferometer is affected by the presence of glitches. Real noise triggers are efficiently down ranked in terms of SNR, using the likelihood ratio produced as an output by the supervised learning inference. It is important to note that some triggers are also re ranked with a new SNR higher than the one estimated by the search pipeline, meaning that iDQ can also label periods as clean, increasing the confidence of some rare triggers. However, it is unclear if this up ranking is a desirable and robust effect.

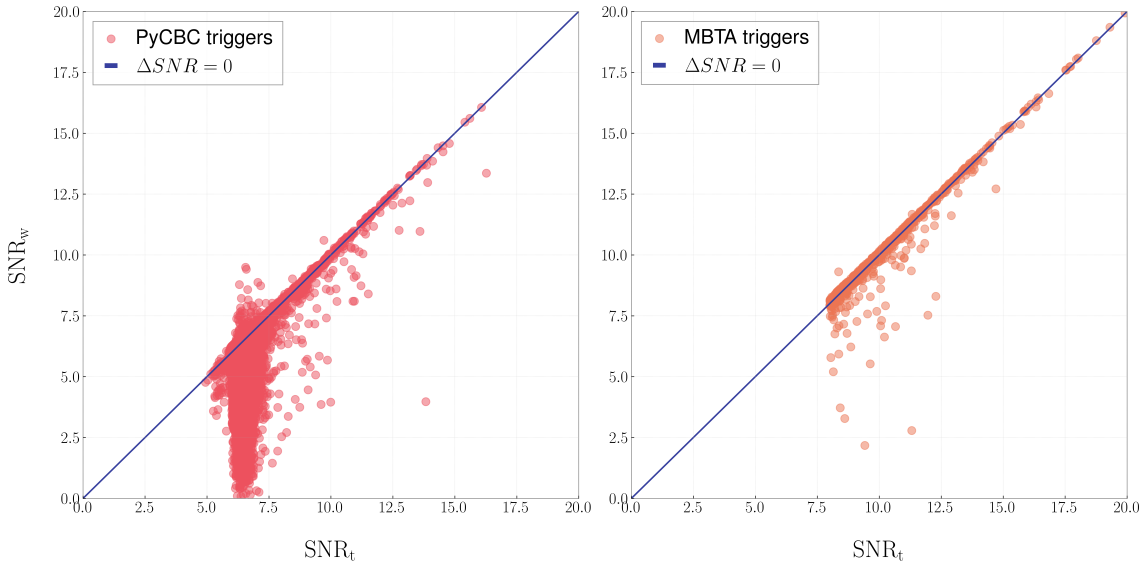

Figure 5.10: Scatter plot of the reweighted SNR_w as a function of SNR_t, for the PyCBC single triggers (left) and the MBTA single triggers (right). The blue line corresponds to the zone where $\Delta SNR = SNR_w - SNR_t = 0$.

5.5.3 Performances of iDQ

Beside the inclusion of iDQ statistical outputs into GW searches, we also tried to quantify the performances of its inference (in terms of amount of glitches found over a certain time) in comparison to other types veto algorithms built for noise-hunting. Over the chunk four of the third observing run, we compared iDQ to Use-percentage-veto (UPV) algorithm (UPV), using the VetoPerf analysis tool [233, 234]. The VetoPerf analysis is a tool used to compare the performances of a data quality vetoes, which flags segment of time as noisy. VetoPerf will count the number of triggers in the target channel and compare it to the number of segments flagged by the noise-hunting algorithm.

We ran both UPV and iDQ analysis between $GPS_{start} = 1240011371$ and $GPS_{end} = 1240639381$, using the same target channel. A time segment is considered as "noisy" if the figure of merit of each algorithm is greater than 0.4. For iDQ, we choose to use $p(\vec{d}(t)|G)$ as the statistical quantity to decide if a segment contains glitches. Fig. 5.11 shows the result of the iDQ analysis. We find that iDQ manages to consistently flag $\sim$ 25% of the glitches, for triggers with $SNR > 7$, $SNR > 8$ and $SNR > 10$. Compared to the fraction of veto clusters found by UPV, it corresponds to an increase of +6.4%. In addition, we tested iDQ inference against two other types of veto streams, one sensitive to light scattered glitches and the other to Schuman resonances (lightning glitches) [234]. Similarly, we also observe that iDQ is able to flag more glitches, using the likelihood ratio or $p(\vec{d}(t)|G)$ as a figure of merit.

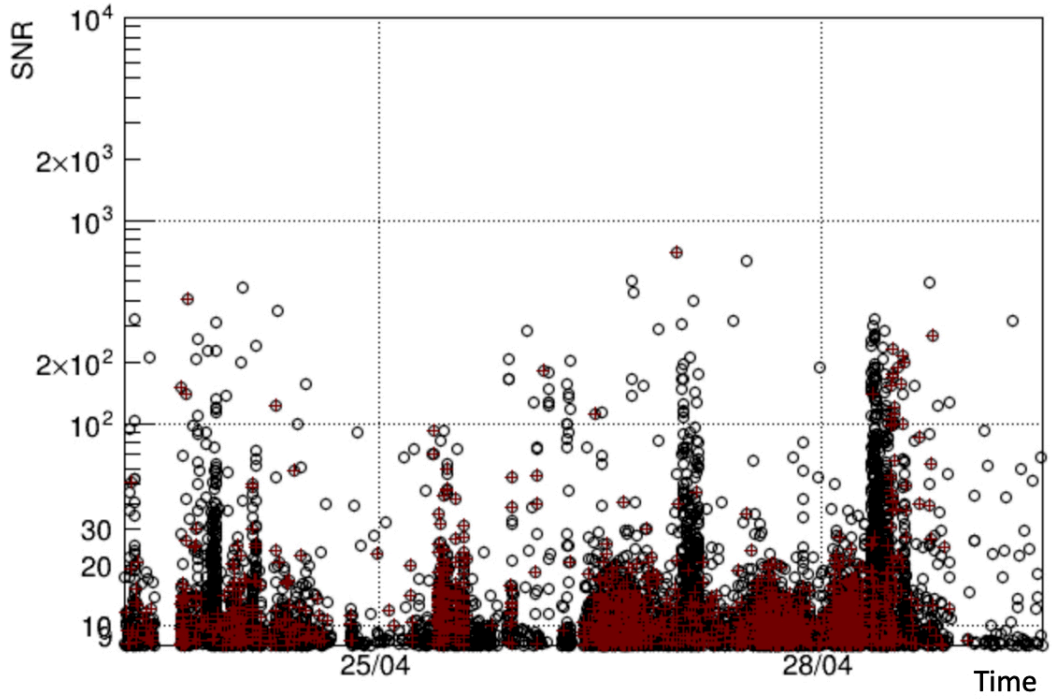

Figure 5.11: Scatter plot of the SNR associated with a trigger as a function of time. The black circles correspond to the list of known trigger and the red crosses are the triggers recovered by the iDQ inference.

5.6 CONCLUSION

In this chapter, we explored the use of a machine learning algorithm called iDQ to improve the data quality of the Virgo interferometer for GW searches. Until now, the data quality flags of the Virgo interferometer were mainly based on methods like gating, which can impact the overall quantity of data available for GW analysis. iDQ is a specialized supervised learning algorithm that incorporates a framework for the statistical inference of non-Gaussian noises (glitches) in the interferometer data stream. We used four weeks of data from the third observing run of Virgo to test iDQ's abilities and noise-hunting performance. Of all types of noise, glitches have the highest impact on the confidence level we assign to GW candidates because of their similarities with real GW signals.

We found that iDQ can produce reliable probabilistic quantities about the amount of noise over time in the interferometer data by efficiently searching for correlations between a target channel sensitive to GW and numerous auxiliary channels monitoring the state of the detector. These quantities can be used to flag periods when the detector is affected by glitches. The main quantities are the probability of having a glitch at time t, $p(\vec{d}|G)$, the likelihood ratio Λ^G_C or the false alarm probability.

We demonstrated that applying iDQ's data quality information about the likelihood ratio to GW search pipelines can significantly lower the SNR of triggers associated with noisy periods, while leaving triggers associated with clean periods unaffected. This exploratory study also shows the potential impact of iDQ as a post-analysis tool for GW searches, to increase or decrease the confidence level in GW candidates. Compared to classical noise-hunting algorithms like UPV, iDQ can consistently flag a larger fraction of glitches, with an increase of 6.4% for triggers with $SNR > 7$. Additionally, promising results in flagging glitches associated with scattered light glitches and lightning noises were found.

Besides the probabilistic quantities provided by iDQ output products, we also highlight the potential of feature importance summaries that give crucial insights into the auxiliary channels most correlated with the presence of non-Gaussian noise in the target channel. This information can help diagnose the source of transient noises and guide the interferometer commissioning process.

In conclusion, this study demonstrates the potential of using machine learning-based algorithms like iDQ as a new tool to improve the data quality of the Virgo interferometer and possibly enhance the performance of GW searches. By incorporating iDQ's probabilistic outputs into the search pipelines and leveraging its feature importance analysis, we can increase confidence in GW event detections and gain valuable insights into the interferometer's noise sources. Based on the results of this study, the products of iDQ will be produced in high latency for the current observing run of the LVK collaboration, O4. The production will be organized at the CCIN2P3 computing center in FRANCE, where the Virgo data streams are stored. All products will be available to the GW searches.

CONCLUSION

"What is the future of the Universe?", "How did black holes originate?", and "What confidence do we have in our measurements?", these are long-standing questions introduced after the pioneering work of Albert Einstein and his theory of General Relativity, or Edwin Powell Hubble and Georges Henri Lemaître with their novel understanding of the Universe's expansion. This manuscript contributes to the scientific efforts to answer these questions, aided by the novel probe of gravitational waves. Gravitational waves present a revolutionary tool for fundamental physics, offering insights into a vast array of phenomena. They provide a deeper understanding of the formation, evolution, and dynamics of astrophysical sources, such as black holes and neutron stars. Moreover, they offer a unique perspective on cosmology, allowing us to explore the early and late moments of the Universe, and track its subsequent expansion. Recent gravitational wave observations of binary black hole and neutron star mergers have opened a new window into the Universe, leading to significant discoveries. These include direct evidence of black holes, insights into their astrophysical formation, evolution, and intrinsic properties, as well as the discovery of the stochastic gravitational wave background using pulsar timing arrays. Gravitational wave physics has also unveiled the production of short gamma-ray bursts during neutron star coalescence and provided the ability to measure the Universe's expansion rate. Through these diverse applications, gravitational waves have the transformative potential to revolutionize our comprehension of the Universe's fundamental workings.

In addition to the review of fundamental physics concerning gravitational waves and the standard model of cosmology presented in Chapter I, this manuscript details our understanding of binary black hole astrophysics, their formation processes, and how these systems can aid in measuring the dynamics of the Universe's expansion. The capability to estimate cosmological parameters, such as the Hubble constant, using gravitational wave signals alone is innovative, as this probe is intrinsically independent of other approaches like Type Ia supernovae or cosmic microwave background methods. The *Standard Siren* method introduced in Chapter II which uses binary black hole mergers to measure cosmological distances and estimate the Hubble constant, will become increasingly accurate as the number of detected events naturally grows with the detectors' advancements. This is especially true with the advent of projects like the Laser Interferometer Space Antenna (LISA), the Einstein Telescope (ET), and Cosmic Explorer (CE), which are expected to detect thousands of signals per week from various astrophysical sources. The *Standard Siren* analysis for gravitational wave cosmology shows good agreement with other approaches, indicating that our Universe is still rapidly expanding today, at a rate of $H_0 = 68^{+12}_{-8}\, km\, s^{-1}\, Mpc^{-1}$ for the

best estimate. However, this approach is not without flaws and can be subject to significant systematic errors inherent to its methodology, as shown in Chapter III. We have demonstrated that the effects stemming from unknown astrophysical processes intrinsic to gravitational wave sources could introduce biases in our estimation of both cosmological parameters and the binary black hole population. Confidence in the measurements will require extensive studies of the systematics, particularly as we approach $\mathcal{O}(\sim 1\%)$ precision for cosmological parameters. The *Standard Siren* inference is affected by unmodeled astrophysical phenomena, such as the redshift evolution of the binary black hole mass spectrum, which this method has demonstrated to be highly sensitive to. Nonetheless, these unknown astrophysical processes also provide opportunities for new discoveries. Chapter IV, which explores novel models describing the spin-mass interplay of binary black hole populations, investigates the effects of black hole formation channels and uncovers evidence for the existence of hierarchical mergers. These mergers exhibit distinct spin distributions, with their spin magnitudes close to $\chi \sim 0.7$. Compelling evidence suggests that a small fraction of black holes (approximately 2%) originate from previously merged binaries. These black holes are characterized by higher source-frame masses and spin magnitudes compared to first-generation black holes. In the future, a major limitation of such analyses will be the purity of gravitational wave detections. Most methods assume that the data originate from real astrophysical sources and not from complex noise sources that mimic them. Therefore, exploring high-end machine learning techniques, as discussed in Chapter V, to identify glitches and thereby increase the confidence in gravitational wave detections will become crucial. Concerning the advancements in data quality for the Virgo interferometer, the machine learning work with iDQ shows promising results both for glitch hunting and in managing the confidence level of our detections.

Prospects and hopes for new discoveries in the field of gravitational wave physics are high, driven by both technological breakthroughs and fundamental explorations of the Universe's secrets. As the LIGO-Virgo-KAGRA Collaboration completes its fourth observing run with unprecedented sensitivity, preparations for the next generation of detectors have already begun, ensuring that gravitational wave astronomy will play a major role in physics in the coming decades. In the near future, the number of applications related to gravitational waves is destined to increase. Scientists eagerly await the moment when post-merger signals from neutron star mergers become visible, which will help constrain the equation of state of the densest astrophysical objects made of ordinary matter. The cosmological prospects are even more promising: with more precise localization and increased detections, constraints on the matter density parameter and the Hubble constant could achieve sub-percent accuracy. This would allow us to compete with other measurements and potentially resolve the persistent problem of the Hubble tension. Additionally, more exotic searches lie ahead for gravitational wave astronomy, including deviations from the ΛCDM cosmological model, tests of General Relativity, gravitational wave propagation friction, and the exploration of extra dimensions of space-time. Regarding the future of population studies, the advent

of third-generation detectors will be highly valuable. LISA is anticipated to detect gravitational waves from new sources such as massive binaries and extreme mass ratio inspirals, which will enable us to explore the formation and astrophysical properties of unseen objects. With CE and ET, we expect to resolve nearly the entire population of stellar-mass black holes, which will provide a comprehensive understanding of their mass and spin distributions. Consequently, it is essential for the gravitational wave scientific community to continue pursuing these efforts. The future of the field will heavily rely on the performance of new algorithms capable of fast and accurate parameter estimations of gravitational wave sources, the correct identification of noisy periods in the detectors, the development of new waveform and population models that can resolve a broad spectrum of new sources, and robust inference frameworks suited to high statistical regimes. In this context, this manuscript represents one step forward, contributing to the advancement of methods and understanding what will underpin the next generation of discoveries in gravitational wave astronomy.

EPILOGUE

Besides making a personal contribution to the field of gravitational wave physics, these past three years (2021-2024) as a PhD student and a member of the LVK Collaboration have taught me a lot. Out of all the opportunities this work gave me, I would like to highlight some of them. I am particularly grateful for the chance to participate directly in the preparation of the cosmological analysis for the fourth observing run of LIGO-Virgo-KAGRA. This involvement included developments in the ICAROGW pipeline, organizing analysis shifts during the detection periods, engaging in long scientific discussions during the *cosmological pipeline development* calls as one of the co-chairs, and planning and writing the O4 cosmology paper. Additionally, working within the LVK Collaboration gave me the opportunity to experience life in Japan for a month, collaborating with colleagues from the KAGRA side of the collaboration in Tokyo and meeting new people to discuss physics. Finally, I recall the *Virgo weeks*, held several times a year next to the Virgo detector in Italy, which were invaluable experiences.

A APPENDIX A

A.1 SOURCE FRAME POPULATION MASS MODELS

In this appendix, we present the supplementary material of Sec. 3. The summary tables for each of the three source mass models used for the hierarchical Bayesian framework in ICAROGW. These three mass models are called: the *Broken Powerlaw*, the *Powerlaw plus peak* and the *Multi peak* mass models.

A.1.1 The BROKEN POWERLAW model

Table. A.1 presents the parameters that control the BPL mass distribution, as well as the prior ranges used to perform the joint inference on the cosmology and the BBH populations.

Parameter	Description	Prior
α_1	Power Law index number 1 primary mass.	$\mathcal{U}(-4, 10)$
α_2	Power Law index number 2 primary mass.	$\mathcal{U}(-4, 10)$
β	Power Law index secondary mass.	$\mathcal{U}(-4, 10)$
m_{min}	Minimum value of the source mass $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
m_{max}	Maximum value of the source mass $[M_\odot]$.	$\mathcal{U}(100M_\odot, 200M_\odot)$
δ_m	Smoothing parameter $[M_\odot]$.	$\mathcal{U}(0M_\odot, 10M_\odot)$
b	Breaking point $[M_\odot]$.	$\mathcal{U}(0M_\odot, 1M_\odot)$

Table A.1: Summary table of the population parameters governing the source BPL mass model, alongside their description. On the right hand side of the table are reported the prior ranges typically used for the Spectral sirens analysis.

A.1.2 The POWERLAW PLUS PEAK model

Table. A.2 presents the parameters that control the PLP mass distribution, as well as the prior ranges used to perform the joint inference on the cosmology and the BBH populations.

Parameter	Description	Prior
α	Power Law index primary mass	$\mathcal{U}(1, 10)$
β	Power Law index secondary mass	$\mathcal{U}(-4, 10)$
m_{min}	Minimum minimum value of the source mass $[M_\odot]$.	$\mathcal{U}(1M_\odot, 12M_\odot)$
m_{max}	Maximum value of the source mass $[M_\odot]$.	$\mathcal{U}(50M_\odot, 200M_\odot)$
λ_g	Fraction of the model in the Gaussian component.	$\mathcal{U}(0, 1)$
μ_g	Mean of the Gaussian peak $[M_\odot]$.	$\mathcal{U}(10M_\odot, 40M_\odot)$
σ_g	Standard deviation of the Gaussian peak $[M_\odot]$.	$\mathcal{U}(6M_\odot, 17M_\odot)$
δ_m	Range of mass tapering at the lower end of the mass distribution $[M_\odot]$.	$\mathcal{U}(0M_\odot, 12M_\odot)$

Table A.2: Summary table of the population parameters governing the source PLP mass model, alongside their description. On the right hand side of the table are reported the prior ranges typically used for the Spectral sirens analysis.

A.1.3 The MULTI PEAK model

Table. A.3 presents the parameters that control the MLTP mass distribution, as well as the prior ranges used to perform the joint inference on the cosmology and the BBH populations.

Parameter	Description	Prior
α	Power Law index primary mass.	$\mathcal{U}(1, 10)$
β	Power Law index secondary mass.	$\mathcal{U}(-4, 10)$
m_{min}	Minimum value of the source mass $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
m_{max}	Maximum value of the source mass $[M_\odot]$.	$\mathcal{U}(50M_\odot, 200M_\odot)$
δ_m	Smoothing parameter $[M_\odot]$.	$\mathcal{U}(0M_\odot, 10M_\odot)$
μ_g^{low}	Mean of the lower gaussian peak $[M_\odot]$.	$\mathcal{U}(11M_\odot, 30M_\odot)$
μ_g^{high}	Mean of the higher gaussian peak $[M_\odot]$.	$\mathcal{U}(40M_\odot, 80M_\odot)$
σ_g^{low}	Standard deviation of the higher gaussian peak $[M_\odot]$.	$\mathcal{U}(6M_\odot, 17M_\odot)$
σ_g^{high}	Standard deviation of the higher gaussian peak $[M_\odot]$.	$\mathcal{U}(6M_\odot, 17M_\odot)$
λ_g	Proportion of events in the peaks.	$\mathcal{U}(0, 1)$
λ_g^{low}	Proportion of events in the lower peak.	$\mathcal{U}(0, 1)$

Table A.3: Summary table of the population parameters governing the source MLTP mass model, alongside their description. On the right hand side of the table are reported the prior ranges typically used for the Spectral sirens analysis.

B APPENDIX B

In this appendix, we present the supplementary material of Sec. 3.

B.1 INJECTED VALUES FOR THE POPULATION DISTRIBUTIONS

B.1.1 The BROKEN POWERLAW model

Parameter	Description	Injected value
α_1	Power Law index number 1 primary mass.	1.5
α_2	Power Law index number 2 primary mass.	5.5
β	Power Law index secondary mass.	1.4
m_{min}	Minimum value of the source mass $[M_\odot]$.	$5M_\odot$
m_{max}	Maximum value of the source mass $[M_\odot]$.	$100M_\odot$
δ_m	Smoothing parameter $[M_\odot]$.	$5M_\odot$
b	Breaking point $[M_\odot]$.	$0.4M_\odot$

Table B.1: Summary of the values injected to construct the BPL mass distribution of the BBH populations analyzed in Sec. 3.3.

B.1.2 The POWERLAW PLUS PEAK model

Parameter	Description	Injected value
α	Spectral index for the PL of the primary mass distribution.	2
β	Spectral index for the PL of the mass ratio distribution.	1
m_{min}	Minimum mass of the primary mass distribution $[M_\odot]$.	$5M_\odot$
m_{max}	Maximum mass of the primary mass distribution $[M_\odot]$.	$100M_\odot$
λ_g	Fraction of the model in the Gaussian component.	0.1
μ_g	Mean of the Gaussian in the primary mass distribution $[M_\odot]$.	$35M_\odot$
σ_g	Width of the Gaussian in the primary mass distribution $[M_\odot]$.	$5M_\odot$
δ_m	Range of mass tapering at the lower end of the mass distribution $[M_\odot]$.	$5M_\odot$

Table B.2: Summary of the values injected to construct the PLP mass distribution of the BBH populations analyzed in Sec. 3.3.

B.1.3 The MULTI PEAK model

Parameter	Description	Injected value
α	Power Law index primary mass.	2
β	Power Law index secondary mass.	1
m_{min}	Minimum value of the source mass $[M_\odot]$.	$5M_\odot$
m_{max}	Maximum value of the source mass $[M_\odot]$.	$100M_\odot$
δ_m	Smoothing parameter $[M_\odot]$.	$5M_\odot$
μ_g^{low}	Mean of the lower gaussian peak $[M_\odot]$.	$16M_\odot$
μ_g^{high}	Mean of the higher gaussian peak $[M_\odot]$.	$50M_\odot$
σ_g^{low}	S.t.d of the higher gaussian peak $[M_\odot]$.	$8M_\odot$
σ_g^{high}	S.t.d of the higher gaussian peak $[M_\odot]$.	$8M_\odot$
λ_g^{high}	Proportion of events in the peaks.	0.7
λ_g^{low}	Proportion of events in the lower peak.	0.8

Table B.3: Summary of the values injected to construct the MLTP mass distribution of the BBH populations analyzed in Sec. 3.3.

B.2 PERCENT-PERCENT PLOTS

B.2.1 Using the correct model

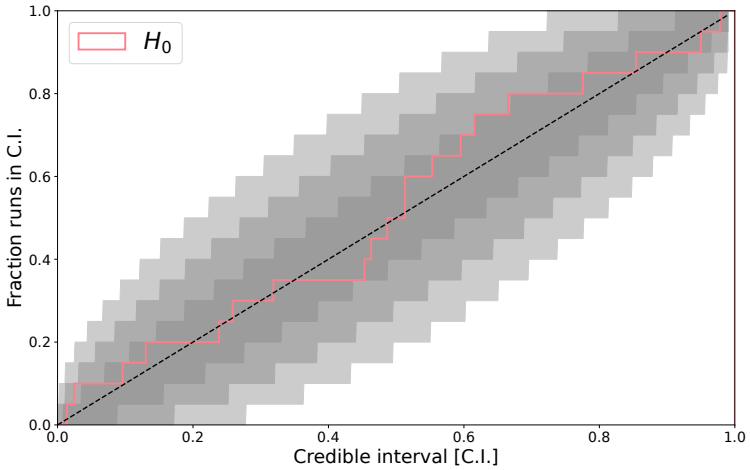

Figure B.1: PP plot obtained from ∼ 20 independent population realizations simulated using the PLP mass model. The Spectral Siren inference is done with the PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

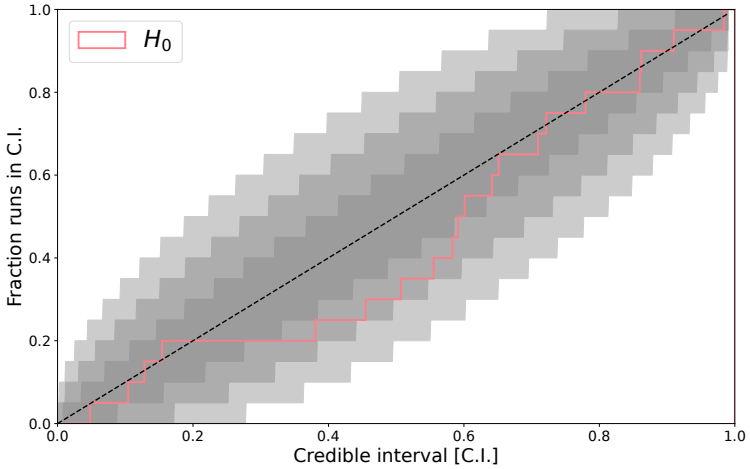

Figure B.2: PP plot obtained from ∼ 20 independent population realizations simulated using the BPL mass model. The Spectral Siren inference is done with the BPL mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

B.2.2 Using the wrong model

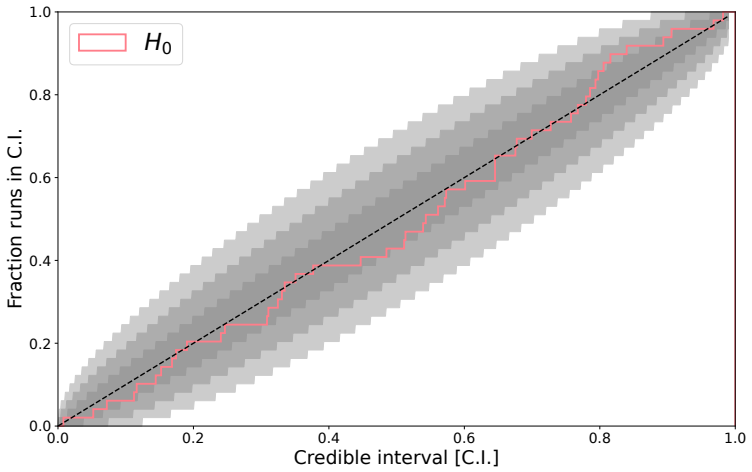

Figure B.3: PP plot obtained from ~ 20 independent population realizations simulated using the PLP mass model. The Spectral Siren inference is done with the MLTP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

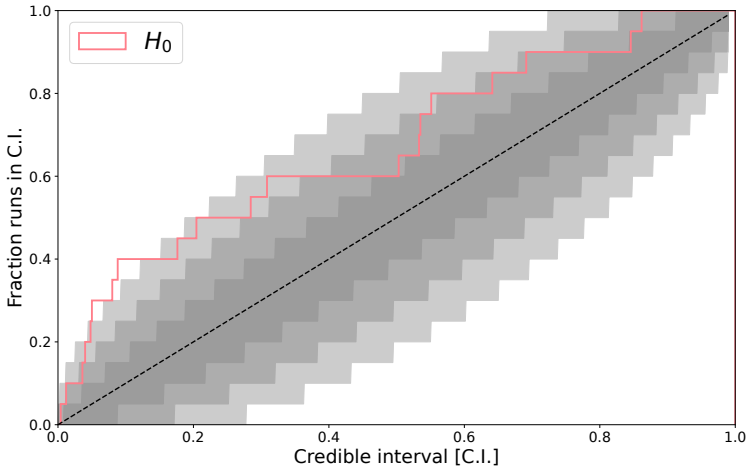

Figure B.4: PP plot obtained from ~ 20 independent population realizations simulated using the MLTP mass model. The Spectral Siren inference is done with the PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

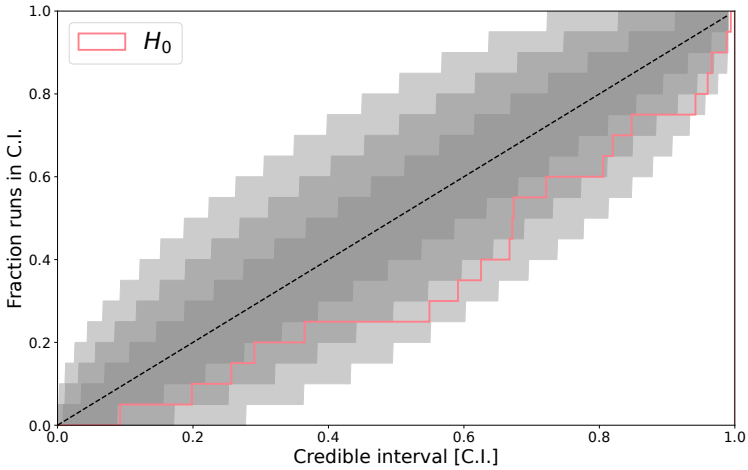

Figure B.5: PP plot obtained from ∼ 20 independent population realizations simulated using the MLTP mass model. The Spectral Siren inference is done with the BPL mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

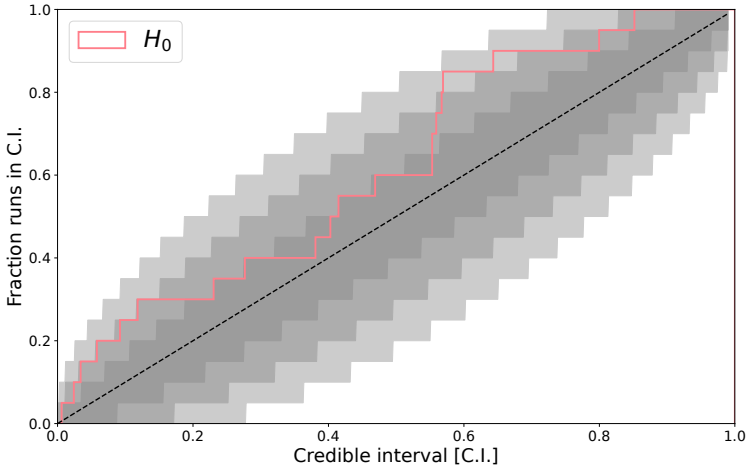

Figure B.6: PP plot obtained from ∼ 20 independent population realizations simulated using the BPL mass model. The Spectral Siren inference is done with the PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

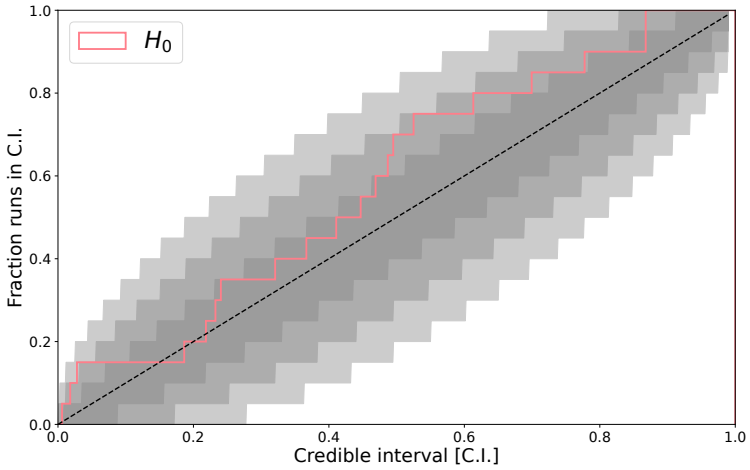

Figure B.7: PP plot obtained from $\sim$ 20 independent population realizations simulated using the BPL mass model. The Spectral Siren inference is done with the MLTP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

B.2.3 Redshift evolution

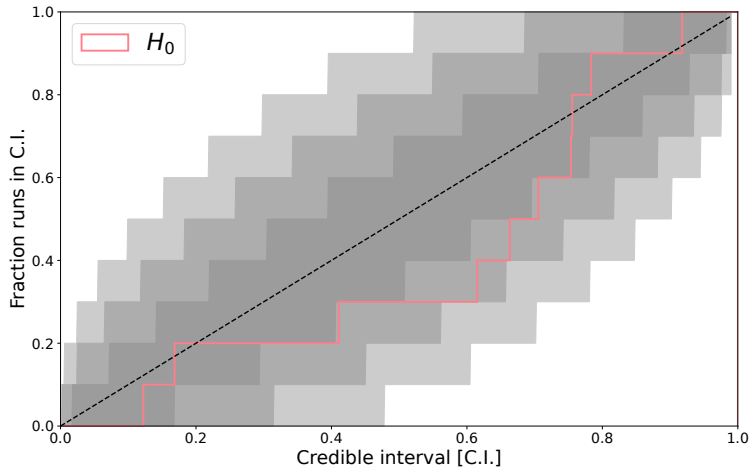

Figure B.8: PP plot obtained from ~ 20 independent population realizations simulated with the modified PLP model, when the Gaussian peak is evolving from $30\,M_\odot$ at $z = 0$ up to $30\,M_\odot$ when $z = 1$. The Spectral Siren inference is done with the regular non evolving PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

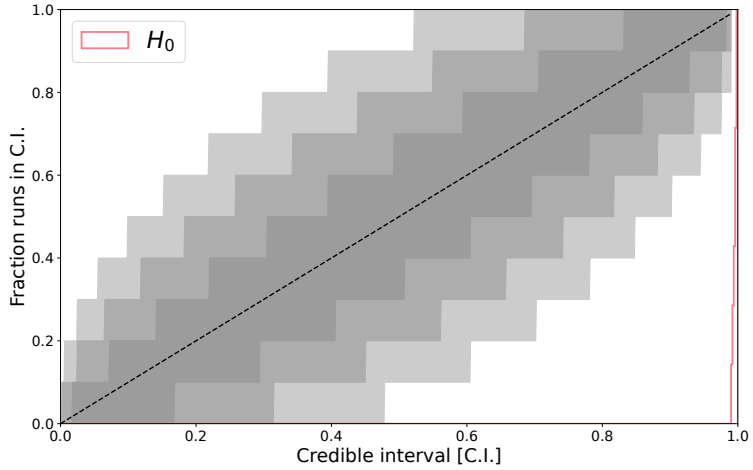

Figure B.9: PP plot obtained from ~ 20 independent population realizations simulated with the modified PLP model, when the Gaussian peak is evolving from $30\,M_\odot$ at $z = 0$ up to $25\,M_\odot$ when $z = 1$. The Spectral Siren inference is done with the regular non evolving PLP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

B.2.4 A03 vanilla

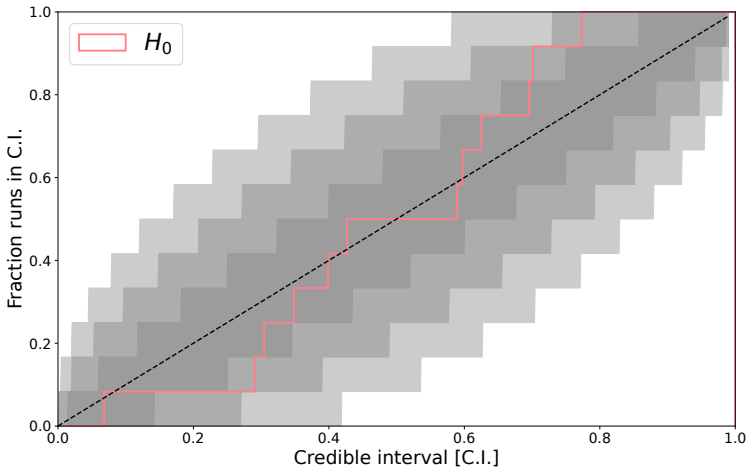

Figure B.10: PP plot obtained from ∼ 15 independent population realizations simulated from the A03 BBH catalog. The Spectral Siren inference is done with the BPL mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

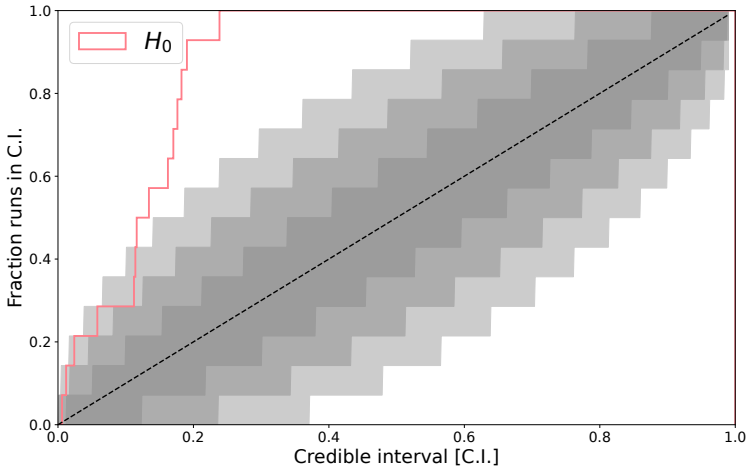

Figure B.11: PP plot obtained from ∼ 15 independent population realizations simulated from the A03 BBH catalog. The Spectral Siren inference is done with the MLTP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

B.2.5 Blinded mass-redshift relation

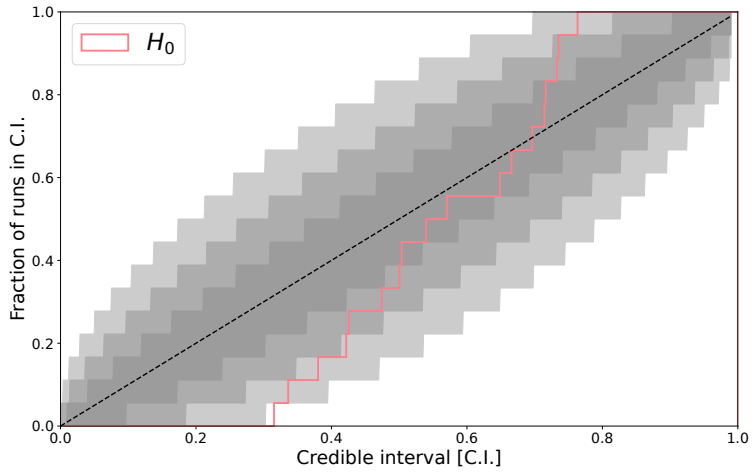

Figure B.12: PP plot obtained from $\sim$ 20 independent population realizations simulated from the A03 BBH catalog, where the redshift evolution of the mass spectrum has been artificially removed. The Spectral Siren inference is done with the MLTP mass model. The gray contours are the 99.7%, 95.0% and 68.0% C.I. respectively from lighter to darker, showing the expected fluctuations due to the finite size of realizations.

C APPENDIX C

C.1 POPULATION RESULTS

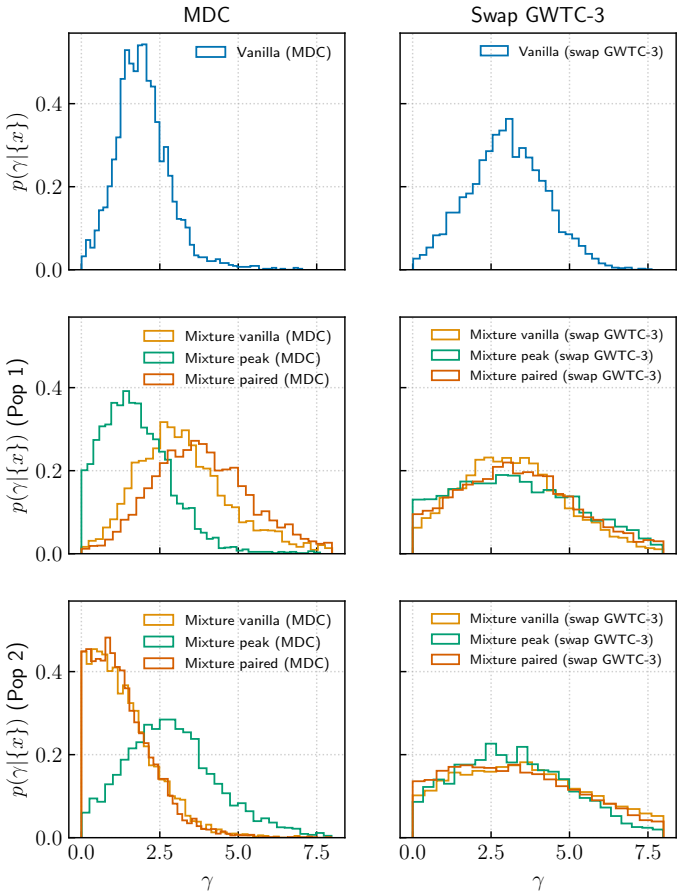

Figure C.1: Inferred posterior of the γ parameter from the CBC merger rate population model. Left plot: Estimated γ from the MDC with the non evolving canonical model (top), first population with MIXTURE models (middle) and second population with the MIXTURE models (bottom). Right plot: Estimated γ from the blurred analysis with the non evolving canonical model (top), first population with MIXTURE models (middle) and second population with the MIXTURE models (bottom).

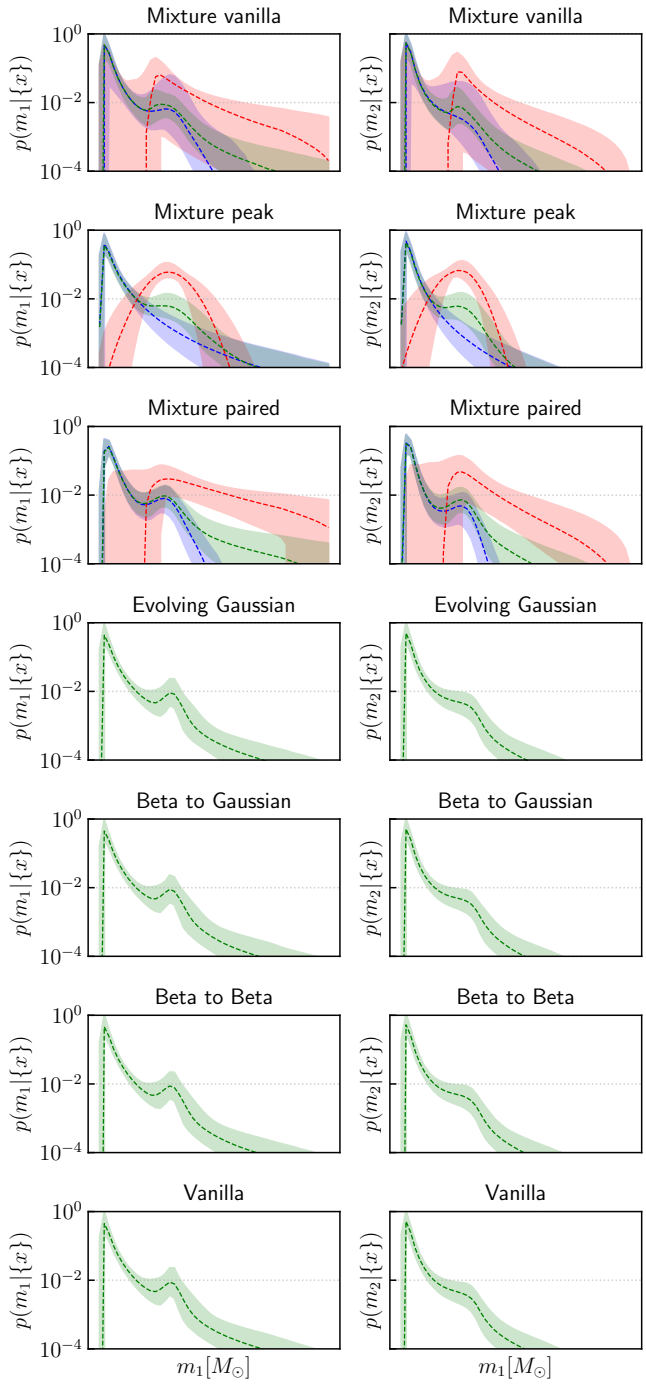

Figure C.2: Reconstructed mass spectra of the primary (left column) and secondary (right column) masses obtained from population inference of the 59 GW events from GWTC-2.1 and GWTC-3, obtained with the MIXTURE VANILLA model (first row), MIXTURE PEAK (second row), MIXTURE PAIRED (third row), EVOLVING GAUSSIAN (fourth row), BETA TO GAUSSIAN (fifth row), BETA TO BETA (sixth row) and the canonical VANILLA model (seventh row). The dotted lines are the median values and the colored contours the 90% C.L. inferred.

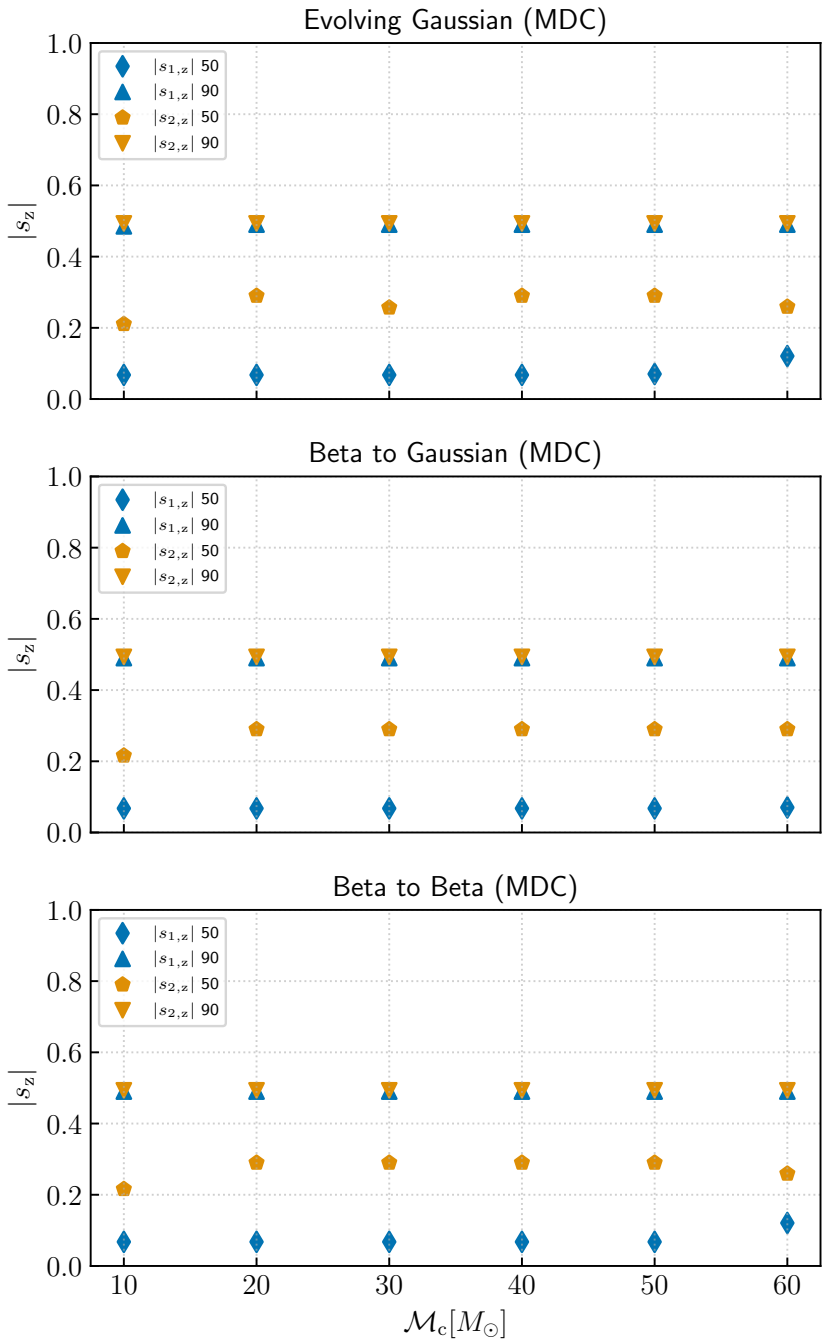

Figure C.3: Scatter plot representing the evolution of the aligned component of the spin magnitude s_z with respect to the chirp mass $\mathcal{M}_c$, obtained from the population inference of simulated GW data (MDC) using the Evolving Gaussian (top), the Beta to Gaussian (middle) and Beta to Beta (bottom) models.

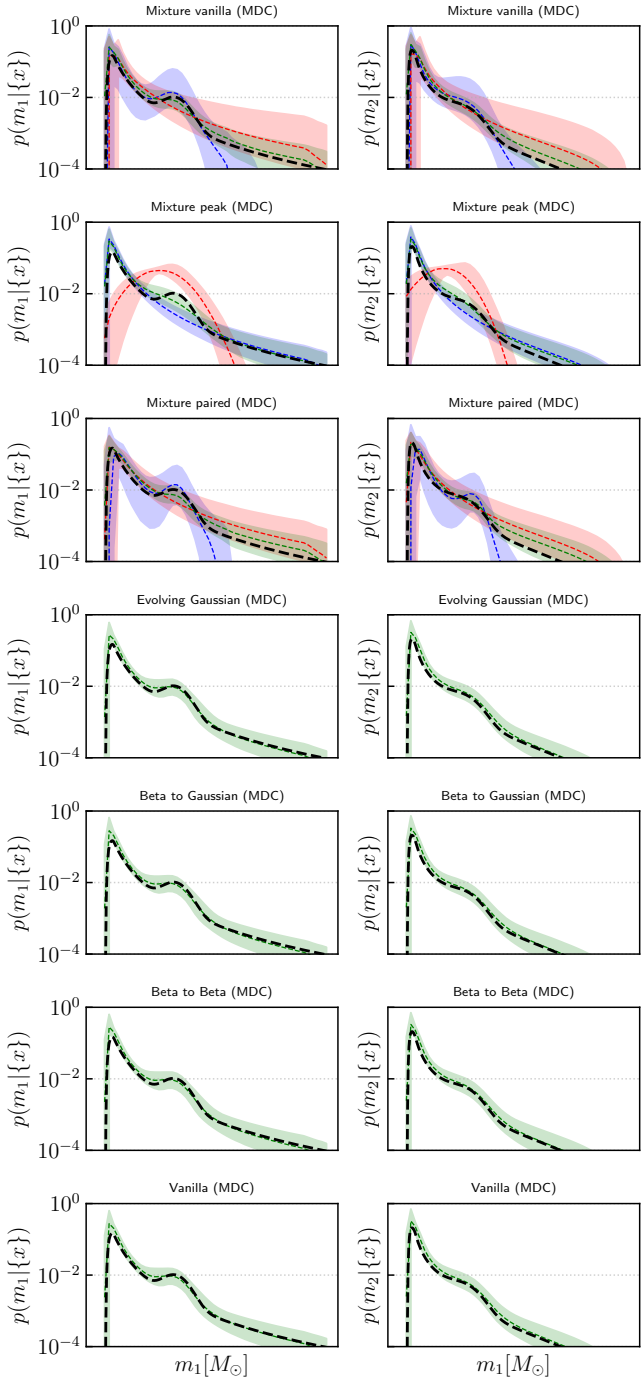

Figure C.4: Reconstructed mass spectra of the primary (left column) and secondary (right column) masses obtained from MDC population inference of the MDC with the MIXTURE VANILLA model (first row), MIXTURE PEAK (second row), MIXTURE PAIRED (third row), EVOLVING GAUSSIAN (fourth row), BETA TO GAUSSIAN (fifth row), BETA TO BETA (sixth row) and the canonical VANILLA model (seventh row). The dotted lines are the median values and the colored contours the 90% C.L. inferred.

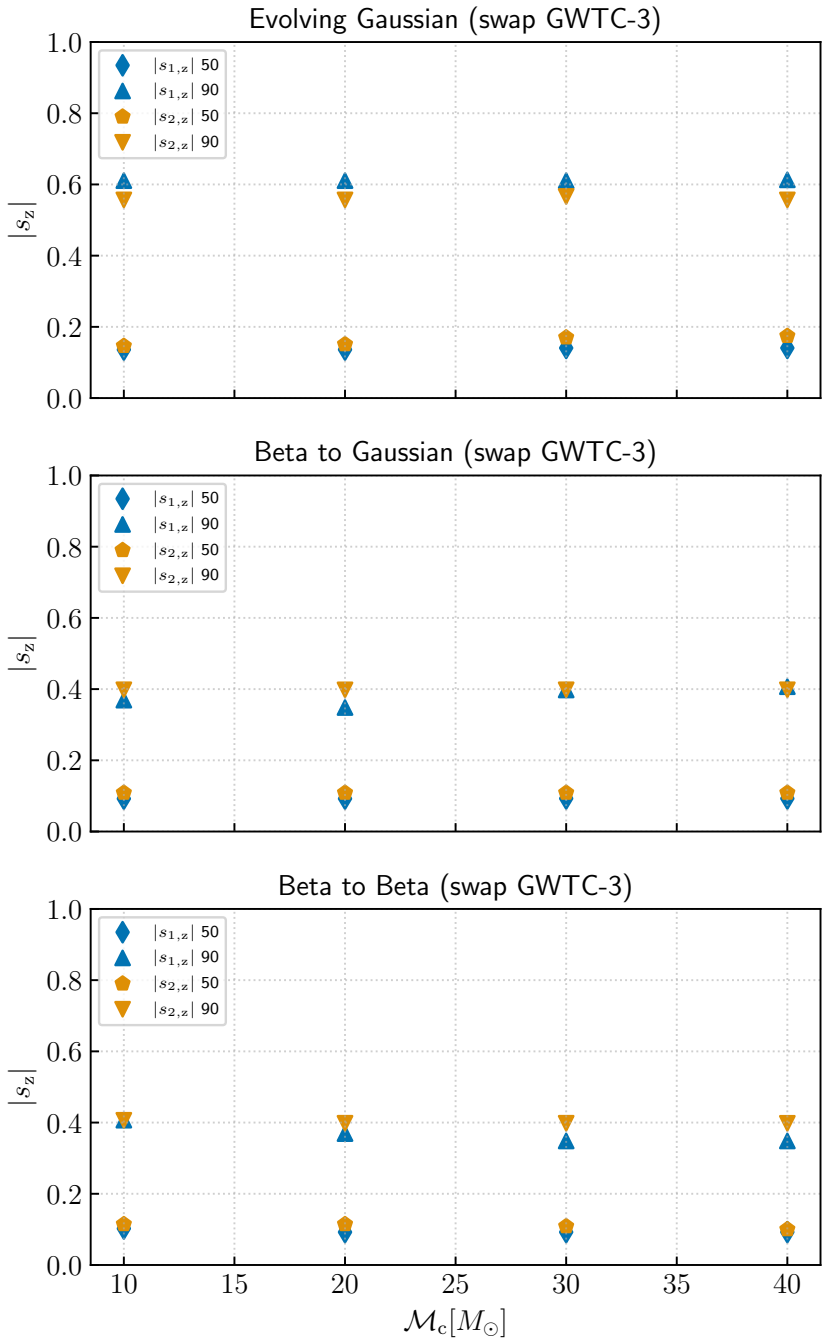

Figure C.5: Scatter plot representing the evolution of the aligned component of the spin magnitude s_z with respect to the chirp mass $\mathcal{M}_c$, obtained from the population inference on the blurred set of 59 GW event, using the Evolving Gaussian (top), the Beta to Gaussian (middle) and Beta to Beta (bottom) models. These results have been obtained for the blurred analysis.

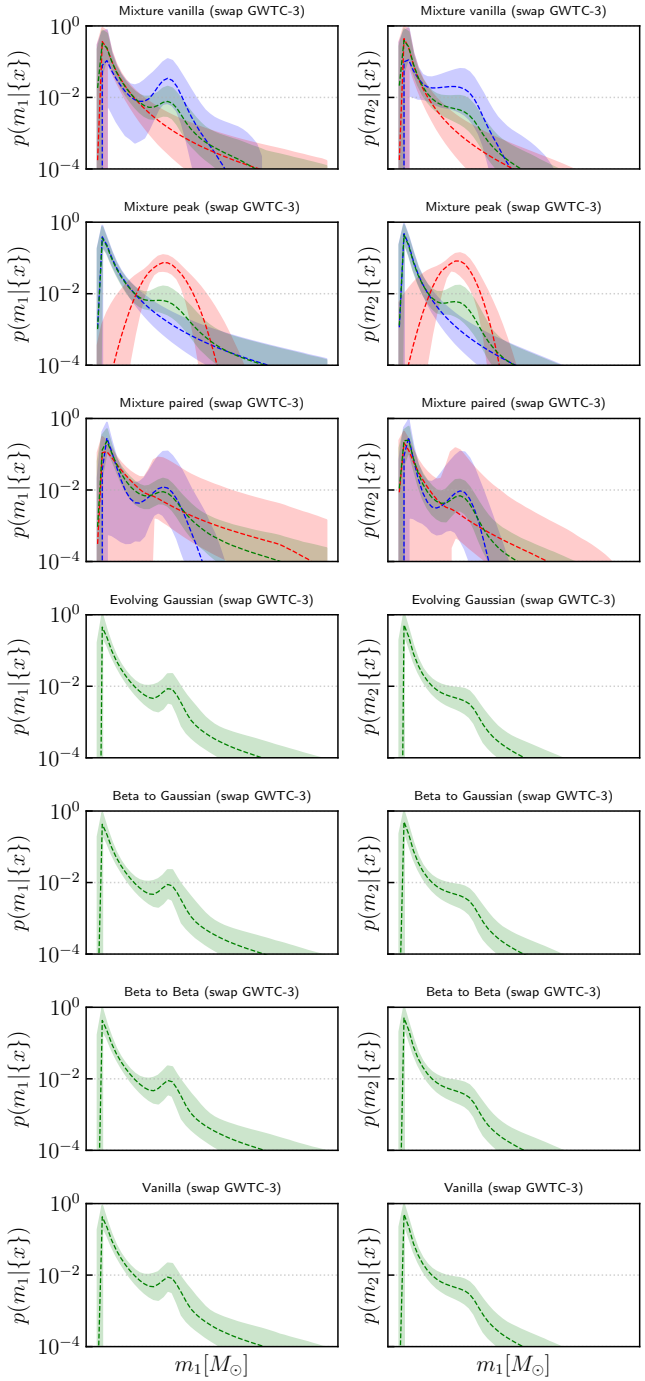

Figure C.6: Reconstructed mass spectra of the primary (left column) and secondary (right column) masses obtained from the blurred analysis with the Mixture vanilla model (first row), Mixture peak (second row), Mixture paired (third row), Evolving gaussian (fourth row), Beta to Gaussian (fifth row), Beta to Beta (sixth row) and the canonical Vanilla model (seventh row). The dotted lines are the median values and the colored contours the 90% C.L. inferred.

C.2 POPULATION MODELS AND PRIOR RANGES

Parameter	Description	Prior
	Masses: Power Law plus Peak	
α	Spectral index for the PL of the primary mass distribution.	$\mathcal{U}(0, 8)$
β	Spectral index for the PL of the mass ratio distribution.	$\mathcal{U}(-1, 10)$
m_{min}	Minimum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 8M_\odot)$
m_{max}	Maximum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(70M_\odot, 130M_\odot)$
λ_g	Fraction of the model in the Gaussian component.	$\mathcal{U}(0, 1)$
μ_g	Mean of the Gaussian in the primary mass distribution $[M_\odot]$.	$\mathcal{U}(20M_\odot, 50M_\odot)$
σ_g	Width of the Gaussian in the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
δ_m	Range of mass tapering at the lower end of the mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
	Spins: Default model	
α_χ^{pop1}	First parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
β_χ^{pop1}	Second parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
σ_t^{pop1}	Standard deviation of the truncated gaussian for the cosine tilt angle distribution.	$\mathcal{U}(0, 5)$
ξ^{pop1}	Mixing parameter for the cosine of the tilt angle distribution.	$\mathcal{U}(0, 1)$
	Rate: MD	
γ	Slope of the power law regime before the point z_p.	$\mathcal{U}(0, 8)$
k	Slope of the power law regime after the point z_p.	$\mathcal{U}(0, 8)$
z_p	Redshift turning point between the power law regimes.	$\mathcal{U}(0, 8)$
$\mathcal{R}_0$	Local value of the CBC merger rate, at $z = 0$.	$\mathcal{U}(0,100)$

Table C.1: Prior ranges and population parameters for the non evolving Vanilla analysis, construct with a Powerlaw + Peak for the masses, a Default spin model and a Madau & Dickinson CBC merger rate.

Parameter	Description	Prior
	Pop1: Power Law plus Peak	
α^{pop1}	Spectral index for the PL of the primary mass distribution.	$\mathcal{U}(0, 8)$
β^{pop1}	Spectral index for the PL of the mass ratio distribution.	$\mathcal{U}(-1, 8)$
m_{min}^{pop1}	Minimum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 8M_\odot)$
m_{max}^{pop1}	Maximum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(45M_\odot, 60M_\odot)$
λ_g^{pop1}	Fraction of the model in the Gaussian component.	$\mathcal{U}(0, 1)$
μ_g^{pop1}	Mean of the Gaussian in the primary mass distribution $[M_\odot]$.	$\mathcal{U}(20M_\odot, 50M_\odot)$
σ_g^{pop1}	Width of the Gaussian in the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
δ_m^{pop1}	Range of mass tapering at the lower end of the mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
	Pop2: Power Law	
α^{pop2}	Spectral index for the PL of the primary mass distribution.	$\mathcal{U}(0, 8)$
β^{pop2}	Spectral index for the PL of the mass ratio distribution.	$\mathcal{U}(-1, 8)$
m_{min}^{pop2}	Minimum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 8M_\odot)$
m_{max}^{pop2}	Maximum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(80M_\odot, 130M_\odot)$
δ_m^{pop2}	Range of mass tapering at the lower end of the mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
	Common parameters	
λ_{pop}	Fraction of the population *pop*1 w.r.t to the overall population.	$\mathcal{U}$(0,1)

Table C.2: Prior ranges used the mass distribution of the MIXTURE VANILLA and MIXTURE PAIRED models.

Parameter	Description	Prior
	Pop1: Power Law plus Peak	
α^{pop1}	Spectral index for the PL of the primary mass distribution.	$\mathcal{U}(0, 8)$
β^{pop1}	Spectral index for the PL of the mass ratio distribution.	$\mathcal{U}(-1, 8)$
m_{min}^{pop1}	Minimum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 8M_\odot)$
m_{max}^{pop1}	Maximum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(80M_\odot, 130M_\odot)$
δ_m^{pop1}	Range of mass tapering at the lower end of the mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
	Pop2: Power Law	
μ^{pop2}	Mean of the gaussian	$\mathcal{U}(20M_\odot,50M_\odot)$
σ^{pop2}	Standard deviation of the gaussian	$\mathcal{U}(2M_\odot,10M_\odot)$
	Common parameters	
λ_{pop}	Fraction of the population $pop1$ w.r.t to the overall population.	$\mathcal{U}(0,1)$

Table C.3: Prior ranges used for the mass distribution of the MIXTURE peak model.

Parameter	Description	Prior
	Power Law plus Peak	
α	Spectral index for the PL of the primary mass distribution.	$\mathcal{U}(0, 8)$
β	Spectral index for the PL of the mass ratio distribution.	$\mathcal{U}(-1, 10)$
m_{min}	Minimum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 8M_\odot)$
m_{max}	Maximum mass of the primary mass distribution $[M_\odot]$.	$\mathcal{U}(70M_\odot, 130M_\odot)$
λ_g	Fraction of the model in the Gaussian component.	$\mathcal{U}(0, 1)$
μ_g	Mean of the Gaussian in the primary mass distribution $[M_\odot]$.	$\mathcal{U}(20M_\odot, 50M_\odot)$
σ_g	Width of the Gaussian in the primary mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$
δ_m	Range of mass tapering at the lower end of the mass distribution $[M_\odot]$.	$\mathcal{U}(1M_\odot, 10M_\odot)$

Table C.4: Prior ranges used for the mass distribution for the Evolving and transition models.

Parameter	Description	Prior
	Pop1: Madau&Dickinson rate	
γ^{pop1}	Slope of the power law regime before the point z_p.	$\mathcal{U}(0, 8)$
k^{pop1}	Slope of the power law regime after the point z_p.	$\mathcal{U}(0, 8)$
z_p^{pop1}	Redshift turning point between the power law regimes.	$\mathcal{U}(0, 8)$
	Pop2: Madau&Dickinson rate	
γ^{pop2}	Slope of the power law regime before the point z_p.	$\mathcal{U}(0, 8)$
k^{pop2}	Slope of the power law regime after the point z_p.	$\mathcal{U}(0, 8)$
z_p^{pop2}	Redshift turning point between the power law regimes.	$\mathcal{U}(0, 8)$
	Common parameters	
$\mathcal{R}_0$	Local value of the CBC merger rate, at $z = 0$.	$\mathcal{U}(0,100)$

Table C.5: Prior ranges of the CBC merger rate models used for all the three flavors of the MIXTURE models.

Parameter	Description	Prior
	Madau&Dickinson rate	
γ	Slope of the power law regime before the point z_p.	$\mathcal{U}(0, 8)$
k	Slope of the power law regime after the point z_p.	$\mathcal{U}(0, 8)$
z_p	Redshift turning point between the power law regimes.	$\mathcal{U}(0, 8)$
$\mathcal{R}_0$	Local value of the CBC merger rate, at $z = 0$.	$\mathcal{U}(0,100)$

Table C.6: Prior ranges of CBC merger rate model used for the EVOLVING model and the TRANSITION models.

Parameter	Description	Prior
	Pop1: Default spin	
α_χ^{pop1}	First parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
β_χ^{pop1}	Second parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
σ_t^{pop1}	Standard deviation of the truncated gaussian for the cosine tilt angle distribution.	$\mathcal{U}(0, 5)$
ξ^{pop1}	Mixing parameter for the cosine of the tilt angle distribution.	$\mathcal{U}(0, 1)$
	Pop2: Default spin	
α_χ^{pop2}	First parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
β_χ^{pop2}	Second parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
σ_t^{pop2}	Standard deviation of the truncated gaussian for the cosine tilt angle distribution.	$\mathcal{U}(0, 5)$
ξ^{pop2}	Mixing parameter for the cosine of the tilt angle distribution.	$\mathcal{U}(0, 1)$

Table C.7: Prior ranges for the spin models used for all the three flavors of the MIXTURE models.

Parameter	Description	Prior
	Evolving Gaussian	
μ_χ	Zero order parameter expansion of the mean of the gaussian for the spin magnitude	$\mathcal{U}(0, 1)$
σ_χ	Zero order parameter expansion of the standard deviation of the gaussian for the spin magnitude	$\mathcal{U}(10^{-3}, 2)$
$\dot{\mu}_\chi$	First order parameter expansion of the mean of the gaussian for the spin magnitude.	$\mathcal{U}(0, 0.1)$
$\dot{\sigma}_\chi$	First order parameter expansion of the standard deviation of the gaussian for the spin magnitude.	$\mathcal{U}(-0.1, 0.1)$
σ_t	Standard deviation of the truncated gaussian for the cosine tilt angle distribution.	$\mathcal{U}(0, 5)$
ξ	Mixing parameter for the cosine of the tilt angle distribution.	$\mathcal{U}(0, 1)$

Table C.8: Spin parameters and prior ranges for the EVOLVING model

Parameter	Description	Prior
	Beta to Gaussian	
α_χ	First parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
β_χ	Second parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1,10)$
μ_χ	Mean of the gaussian for the spin magnitude.	$\mathcal{U}(0, 1)$
σ_χ	Standard deviation of the gaussian for the spin magnitude.	$\mathcal{U}(10^{-3}, 2)$
σ_t	Standard deviation of the truncated gaussian for the cosine tilt angle distribution.	$\mathcal{U}(0, 5)$
ξ	Mixing parameter for the cosine of the tilt angle distribution.	$\mathcal{U}(0, 1)$
m_t	Critical mass at which the window function is equal to 0.5 (transition point).	$\mathcal{U}(10M_\odot,$ $100M_\odot)$
δ_{m_t}	Steepness of the window function.	$\mathcal{U}(1, 20)$
f_{mix}	Starting value of the window function.	$\mathcal{U}(0, 1)$

Table C.9: Spin parameters and prior ranges for the Evolving Beta to Gaussian

Parameter	Description	Prior
	Beta to Beta	
α_χ^{low}	First parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
β_χ^{low}	Second parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1,10)$
α_χ^{high}	First parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1, 10)$
β_χ^{high}	Second parameter of the Beta distribution for the spin magnitude.	$\mathcal{U}(1,10)$
σ_t	Standard deviation of the truncated gaussian for the cosine tilt angle distribution.	$\mathcal{U}(0, 5)$
ξ	Mixing parameter for the cosine of the tilt angle distribution.	$\mathcal{U}(0, 1)$
m_t	Critical mass at which the window function is equal to 0.5 (transition point).	$\mathcal{U}(10M_\odot,$ $100M_\odot)$
δ_{m_t}	Steepness of the window function.	$\mathcal{U}(1, 20)$
f_{mix}	Starting value of the window function.	$\mathcal{U}(0, 1)$

Table C.10: Spin parameters and prior ranges for the Evolving Beta to Beta

BIBLIOGRAPHY

[1] A. Einstein. "The foundation of the general theory of relativity." *Annalen Phys.*, 49(7):769–822, 1916. doi:10.1002/andp.19163540702.

[2] I. Newton. *Philosophiæ Naturalis Principia Mathematica*. England, 1687.

[3] L. Verde, et al. "Tensions between the Early and the Late Universe". *Nature Astron.*, 3:891, 2019. doi:10.1038/s41550-019-0902-0. 1907.10625.

[4] R. Abbott et al. (LIGO Scientific, Virgo, KAGRA). "Constraints on the Cosmic Expansion History from GWTC–3". *Astrophys. J.*, 949(2):76, 2023. doi:10.3847/1538-4357/ac74bb. 2111.03604.

[5] J. D. Bekenstein. "Black holes and entropy". *Phys. Rev. D*, 7:2333–2346, 1973. doi:10.1103/PhysRevD.7.2333.

[6] B. P. Abbott et al. (LIGO Scientific, Virgo). "Observation of Gravitational Waves from a Binary Black Hole Merger". *Phys. Rev. Lett.*, 116(6):061102, 2016. doi:10.1103/PhysRevLett.116.061102. 1602.03837.

[7] B. P. Abbott et al. (KAGRA, LIGO Scientific, Virgo, VIRGO). "Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA". *Living Rev. Rel.*, 21(1):3, 2018. doi:10.1007/s41114-020-00026-9. 1304.0670.

[8] J. Aasi et al. (LIGO Scientific). "Advanced LIGO". *Class. Quant. Grav.*, 32:074001, 2015. doi:10.1088/0264-9381/32/7/074001. 1411.4547.

[9] F. Acernese et al. (VIRGO). "Advanced Virgo: a second-generation interferometric gravitational wave detector". *Class. Quant. Grav.*, 32(2):024001, 2015. doi:10.1088/0264-9381/32/2/024001. 1408.3978.

[10] T. Akutsu et al. (KAGRA). "Overview of KAGRA: Detector design and construction history". *PTEP*, 2021(5):05A101, 2021. doi:10.1093/ptep/ptaa125. 2005.05574.

[11] Y. Aso, et al. (KAGRA). "Interferometer design of the KAGRA gravitational wave detector". *Phys. Rev. D*, 88(4):043007, 2013. doi:10.1103/PhysRevD.88.043007. 1306.6747.

[12] A. Einstein. "Die Feldgleichungen der Gravitation". *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pages 844–847, 1915.

[13] A. Einstein. "Näherungsweise Integration der Feldgleichungen der Gravitation". *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pages 688–696, 1916.

[14] F. A. E. Pirani. "On the Physical significance of the Riemann tensor". *Acta Phys. Polon.*, 15:389–405, 1956. doi:10.1007/s10714-009-0787-9.

[15] J. H. Taylor et J. M. Weisberg. "A new test of general relativity: Gravitational radiation and the binary pulsar PS R 1913+16". *Astrophys. J.*, 253:908–920, 1982. doi:10.1086/159690.

[16] J. H. Taylor et J. M. Weisberg. "Further experimental tests of relativistic gravity using the binary pulsar PSR 1913+16". *Astrophys. J.*, 345:434–450, 1989. doi:10.1086/167917.

[17] A. Einstein. "The Field Equations of Gravitation". *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915:844–847, 1915.

[18] M. Isi. "Parametrizing gravitational-wave polarizations". *Class. Quant. Grav.*, 40(20):203001, 2023. doi:10.1088/1361-6382/acf28c. 2208.03372.

[19] W. G. Anderson, et al. "An Excess power statistic for detection of burst sources of gravitational radiation". *Phys. Rev. D*, 63:042003, 2001. doi:10.1103/PhysRevD.63.042003. gr-qc/0008066.

[20] L. S. Finn et D. F. Chernoff. "Observing binary inspiral in gravitational radiation: One interferometer". *Phys. Rev. D*, 47:2198–2219, 1993. doi:10.1103/PhysRevD.47.2198. gr-qc/9301003.

[21] S. A. Usman, et al. "Constraining the Inclinations of Binary Mergers from Gravitational-wave Observations". *Astrophys. J.*, 877(2):82, 2019. doi:10.3847/1538-4357/ab0b3e. 1809.10727.

[22] J. Veitch et al. "Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library". *Phys. Rev. D*, 91(4):042003, 2015. doi:10.1103/PhysRevD.91.042003. 1409.7215.

[23] C. J. Moore, et al. "Gravitational-wave sensitivity curves". *Class. Quant. Grav.*, 32(1):015014, 2015. doi:10.1088/0264-9381/32/1/015014. 1408.0740.

[24] B. P. Abbott et al. (LIGO Scientific, Virgo). "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs". *Phys. Rev. X*, 9(3):031040, 2019. doi:10.1103/PhysRevX.9.031040. 1811.12907.

[25] R. Abbott et al. (LIGO Scientific, Virgo). "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run". *Phys. Rev. X*, 11:021053, 2021. doi:10.1103/PhysRevX.11.021053. 2010.14527.

[26] R. Abbott et al. (LIGO Scientific, VIRGO). "GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run". *arXiv*, 2021. 2108.01045.

[27] R. Abbott et al. (KAGRA, VIRGO, LIGO Scientific). "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run". *Phys. Rev. X*, 13(4):041039, 2023. doi:10.1103/PhysRevX.13.041039. 2111.03606.

[28] S. W. Hawking. "Black hole explosions". *Nature*, 248:30–31, 1974. doi:10.1038/248030a0.

[29] F. Kaeppeler, et al. "The s Process: Nuclear Physics, Stellar Models, Observations". *Rev. Mod. Phys.*, 83:157, 2011. doi:10.1103/RevModPhys.83.157. 1012.5218.

[30] A. Burrows et D. Vartanyan. "Core-Collapse Supernova Explosion Theory". *Nature*, 589(7840):29–39, 2021. doi:10.1038/s41586-020-03059-w. 2009.14157.

[31] R. Abbott et al. (KAGRA, VIRGO, LIGO Scientific). "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3". *Phys. Rev. X*, 13(1):011048, 2023. doi:10.1103/PhysRevX.13.011048. 2111.03634.

[32] W. Israel. "Event horizons in static vacuum space-times". *Phys. Rev.*, 164:1776–1779, 1967. doi:10.1103/PhysRev.164.1776.

[33] M. Mapelli. *Formation Channels of Single and Binary Stellar-Mass Black Holes*. 2021. doi:10.1007/978-981-15-4702-7_16-1. 2106.00699.

[34] K. Belczynski et al. "The Effect of Pair-Instability Mass Loss on Black Hole Mergers". *Astron. Astrophys.*, 594:A97, 2016. doi:10.1051/0004-6361/201628980. 1607.03116.

[35] S. E. Woosley. "Pulsational Pair-Instability Supernovae". *Astrophys. J.*, 836(2):244, 2017. doi:10.3847/1538-4357/836/2/244. 1608.08939.

[36] G. Pierra, et al. "The spin magnitude of stellar-mass binary black holes evolves with the mass: evidence from gravitational wave data". 2024. 2406.01679.

[37] C. L. Fryer. "Mass limits for black hole formation". *Astrophys. J.*, 522:413, 1999. doi:10.1086/307647. astro-ph/9902315.

[38] C. L. Fryer, et al. "Compact Remnant Mass Function: Dependence on the Explosion Mechanism and Metallicity". *Astrophys. J.*, 749:91, 2012. doi:10.1088/0004-637X/749/1/91. 1110.1726.

[39] K. Belczynski et al. "Evolutionary roads leading to low effective spins, high black hole masses, and O1/O2 rates for LIGO/Virgo binary black holes". *Astron. Astrophys.*, 636:A104, 2020. doi:10.1051/0004-6361/201936528. 1706.07053.

[40] M. Mapelli, et al. "Impact of the Rotation and Compactness of Progenitors on the Mass of Black Holes". *Astrophys. J.*, 888:76, 2020. doi:10.3847/1538-4357/ab584d. 1909.01371.

[41] P. Marchant, et al. "A new route towards merging massive black holes". *Astron. Astrophys.*, 588:A50, 2016. doi:10.1051/0004-6361/201628133. 1601.03718.

[42] M. Mapelli. "Binary Black Hole Mergers: Formation and Populations". *Front. Astron. Space Sci.*, 7:38, 2020. doi:10.3389/fspas.2020.00038. 2105.12455.

[43] E. Berti et M. Volonteri. "Cosmological black hole spin evolution by mergers and accretion". *Astrophys. J.*, 684:822–828, 2008. doi:10.1086/590379. 0802.0025.

[44] D. Gerosa et E. Berti. "Are merging black holes born from stellar collapse or previous mergers?" *Phys. Rev. D*, 95(12):124046, 2017. doi:10.1103/PhysRevD.95.124046. 1703.06223.

[45] M. Fishbach, et al. "Are LIGO's Black Holes Made From Smaller Black Holes?" *Astrophys. J. Lett.*, 840(2):L24, 2017. doi:10.3847/2041-8213/aa7045. 1703.06869.

[46] R. P. Kerr. "Gravitational field of a spinning mass as an example of algebraically special metrics". *Phys. Rev. Lett.*, 11:237–238, 1963. doi:10.1103/PhysRevLett.11.237.

[47] C. W. Misner, et al. *Gravitation*. W. H. Freeman, San Francisco, 1973. ISBN 978-0-7167-0344-0, 978-0-691-17779-3.

[48] L. F. O. Costa et J. Natário. "Frame-Dragging: Meaning, Myths, and Misconceptions". *Universe*, 7(10):388, 2021. doi:10.3390/universe7100388. 2109.14641.

[49] R. H. Boyer et R. W. Lindquist. "Maximal analytic extension of the Kerr metric". *J. Math. Phys.*, 8:265, 1967. doi:10.1063/1.1705193.

[50] C. S. Reynolds. "Observational Constraints on Black Hole Spin". *Ann. Rev. Astron. Astrophys.*, 59:117–154, 2021. doi:10.1146/annurev-astro-112420-035022. 2011.08948.

[51] M. Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments*. Gravitational Waves. OUP Oxford, 2008. ISBN 9780198570745. URL https://books.google.fr/books?id=mk-1DAAAQBAJ.

[52] M. Maggiore. *Gravitational Waves. Vol. 2: Astrophysics and Cosmology*. Oxford University Press, 2018. ISBN 978-0-19-857089-9.

[53] V. Tiwari, et al. "Constraining black-hole spins with gravitational wave observations". *Astrophys. J.*, 868(2):140, 2018. doi:10.3847/1538-4357/aae8df. 1809.01401.

[54] M. Campanelli, et al. "Accurate evolutions of orbiting black-hole binaries without excision". *Phys. Rev. Lett.*, 96:111101, 2006. doi:10.1103/PhysRevLett.96.111101. gr-qc/0511048.

[55] M. Campanelli, et al. "Spinning-black-hole binaries: The orbital hang up". *Phys. Rev. D*, 74:041501, 2006. doi:10.1103/PhysRevD.74.041501. gr-qc/0604012.

[56] G. Lemaitre. "A Homogeneous Universe of Constant Mass and Growing Radius Accounting for the Radial Velocity of Extragalactic Nebulae". *Annales Soc. Sci. Bruxelles A*, 47:49–59, 1927. doi:10.1007/s10714-013-1548-3.

[57] E. Hubble. "A relation between distance and radial velocity among extragalactic nebulae". *Proc. Nat. Acad. Sci.*, 15:168–173, 1929. doi:10.1073/pnas.15.3.168.

[58] E. Hubble et M. L. Humason. "The Velocity-Distance Relation among Extra-Galactic Nebulae". *Astrophys. J.*, 74:43–80, 1931. doi:10.1086/143323.

[59] G. Lemaitre. "The expanding universe". *Annales Soc. Sci. Bruxelles A*, 53:51–85, 1933. doi:10.1023/A:1018855621348.

[60] D. M. Scolnic et al. (Pan-STARRS1). "The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample". *Astrophys. J.*, 859(2):101, 2018. doi:10.3847/1538-4357/aab9bb. 1710.00845.

[61] N. Aghanim et al. (Planck). "Planck 2018 results. VI. Cosmological parameters". *Astron. Astrophys.*, 641:A6, 2020. doi:10.1051/0004-6361/201833910. [Erratum: Astron.Astrophys. 652, C4 (2021)], 1807.06209.

[62] P. J. E. Peebles et B. Ratra. "The Cosmological Constant and Dark Energy". *Rev. Mod. Phys.*, 75:559–606, 2003. doi:10.1103/RevModPhys.75.559. astro-ph/0207347.

[63] B. A. Bassett et R. Hlozek. "Baryon Acoustic Oscillations". 2009. 0910.5224.

[64] A. A. Penzias et R. W. Wilson. "A Measurement of excess antenna temperature at 4080-Mc/s". *Astrophys. J.*, 142:419–421, 1965. doi:10.1086/148307.

[65] N. Suzuki et al. (Supernova Cosmology Project). "The Hubble Space Telescope Cluster Supernova Survey: V. Improving the Dark Energy Constraints Above z>1 and Building an Early-Type-Hosted Supernova Sample". *Astrophys. J.*, 746:85, 2012. doi:10.1088/0004-637X/746/1/85. 1105.3470.

[66] A. Friedman. "On the Curvature of space". *Z. Phys.*, 10:377–386, 1922. doi:10.1007/BF01332580.

[67] A. Friedmann. "On the Possibility of a world with constant negative curvature of space". *Z. Phys.*, 21:326–332, 1924. doi:10.1007/BF01328280.

[68] A. G. Walker. "On Milne's Theory of World-Structure". *Proc. Lond. Math. Soc. s*, 2-42(1):90–127, 1937. doi:10.1112/plms/s2-42.1.90.

[69] A. H. Guth. "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems". *Phys. Rev. D*, 23:347–356, 1981. doi:10.1103/PhysRevD.23.347.

[70] B. F. Schutz. "Determining the Hubble Constant from Gravitational Wave Observations". *Nature*, 323:310–311, 1986. doi:10.1038/323310a0.

[71] D. W. Hogg. "Distance measures in cosmology". 1999. astro-ph/9905116.

[72] E. Di Valentino et al. "Snowmass2021 - Letter of interest cosmology intertwined II: The hubble constant tension". *Astropart. Phys.*, 131:102605, 2021. doi:10.1016/j.astropartphys.2021.102605. 2008.11284.

[73] E. Di Valentino. "A combined analysis of the H_0 late time direct measurements and the impact on the Dark Energy sector". *Mon. Not. Roy. Astron. Soc.*, 502(2):2065–2073, 2021. doi:10.1093/mnras/stab187. 2011.00246.

[74] A. G. Riess et L. Breuval. "The Local Value of H0". *IAU Symp.*, 376:15–29, 2022. doi:10.1017/S1743921323003034. 2308.10954.

[75] A. G. Riess, et al. "Cosmic Distances Calibrated to 1% Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with ΛCDM". *Astrophys. J. Lett.*, 908(1):L6, 2021. doi:10.3847/2041-8213/abdbaf. 2012.08534.

[76] W. Del Pozzo. "Inference of the cosmological parameters from gravitational waves: application to second generation interferometers". *Phys. Rev. D*, 86:043011, 2012. doi:10.1103/PhysRevD.86.043011. 1108.1317.

[77] S. Mastrogiovanni, et al. "ICAROGW: A python package for inference of astrophysical population properties of noisy, heterogeneous, and incomplete observations". *Astron. Astrophys.*, 682:A167, 2024. doi:10.1051/0004-6361/202347007. 2305.17973.

[78] J. R. Gair et al. "The Hitchhiker's Guide to the Galaxy Catalog Approach for Dark Siren Gravitational-wave Cosmology". *Astron. J.*, 166(1):22, 2023. doi: 10.3847/1538-3881/acca78. 2212.08694.

[79] S. Mastrogiovanni, et al. "Joint population and cosmological properties inference with gravitational waves standard sirens and galaxy surveys". *Phys. Rev. D*, 108(4):042002, 2023. doi:10.1103/PhysRevD.108.042002. 2305.10488.

[80] J. M. Ezquiaga et D. E. Holz. "Spectral Sirens: Cosmology from the Full Mass Distribution of Compact Binaries". *Phys. Rev. Lett.*, 129(6):061102, 2022. doi: 10.1103/PhysRevLett.129.061102. 2202.08240.

[81] R. Gray, et al. "A pixelated approach to galaxy catalogue incompleteness: improving the dark siren measurement of the Hubble constant". *Mon. Not. Roy. Astron. Soc.*, 512(1):1127–1140, 2022. doi:10.1093/mnras/stac366. 2111.04629.

[82] N. Dalal, et al. "Short grb and binary black hole standard sirens as a probe of dark energy". *Phys. Rev. D*, 74:063006, 2006. doi:10.1103/PhysRevD.74.063006. astro-ph/0601275.

[83] B. S. Sathyaprakash, et al. "Cosmography with the Einstein Telescope". *Class. Quant. Grav.*, 27:215006, 2010. doi:10.1088/0264-9381/27/21/215006. 0906.4151.

[84] S. Nissanke, et al. "Exploring short gamma-ray bursts as gravitational-wave standard sirens". *Astrophys. J.*, 725:496–514, 2010. doi:10.1088/0004-637X/725/1/496. 0904.1017.

[85] B. P. Abbott et al. (LIGO Scientific, Virgo). "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral". *Phys. Rev. Lett.*, 119(16):161101, 2017. doi:10.1103/PhysRevLett.119.161101. 1710.05832.

[86] B. P. Abbott et al. (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT). "Multi-messenger Observations of a Binary Neutron Star Merger". *Astrophys. J. Lett.*, 848(2):L12, 2017. doi: 10.3847/2041-8213/aa91c9. 1710.05833.

[87] M. Soares-Santos et al. (DES, Dark Energy Camera GW-EM). "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/Virgo GW170817. I. Discovery of the Optical Counterpart Using the Dark Energy Camera". *Astrophys. J. Lett.*, 848(2):L16, 2017. doi:10.3847/2041-8213/aa9059. 1710.05459.

[88] B. P. Abbott et al. (LIGO Scientific, Virgo, 1M2H, Dark Energy Camera GW-E, DES, DLT40, Las Cumbres Observatory, VINROUGE, MASTER). "A gravitational-wave standard siren measurement of the Hubble constant". *Nature*, 551(7678):85–88, 2017. doi:10.1038/nature24471. 1710.05835.

[89] C. Guidorzi et al. "Improved Constraints on H_0 from a Combined Analysis of Gravitational-wave and Electromagnetic Emission from GW170817". *Astrophys. J. Lett.*, 851(2):L36, 2017. doi:10.3847/2041-8213/aaa009. 1710.06426.

[90] H.-Y. Chen, et al. "A two per cent Hubble constant measurement from standard sirens within five years". *Nature*, 562(7728):545–547, 2018. doi:10.1038/s41586-018-0606-0. 1712.06531.

[91] S. M. Feeney, et al. "Prospects for resolving the Hubble constant tension with standard sirens". *Phys. Rev. Lett.*, 122(6):061105, 2019. doi:10.1103/PhysRevLett.122.061105. 1802.03404.

[92] D. J. Mortlock, et al. "Unbiased Hubble constant estimation from binary neutron star mergers". *Phys. Rev. D*, 100(10):103523, 2019. doi:10.1103/PhysRevD.100.103523. 1811.11723.

[93] R. Gray et al. "Joint cosmological and gravitational-wave population inference using dark sirens and galaxy catalogues". *JCAP*, 12:023, 2023. doi:10.1088/1475-7516/2023/12/023. 2308.02281.

[94] G. Pierra, et al. "Study of systematics on the cosmological inference of the Hubble constant from gravitational wave standard sirens". *Phys. Rev. D*, 109(8):083504, 2024. doi:10.1103/PhysRevD.109.083504. 2312.11627.

[95] W. M. Farr, et al. "A Future Percent-Level Measurement of the Hubble Expansion at Redshift 0.8 With Advanced LIGO". *Astrophys. J. Lett.*, 883(2):L42, 2019. doi:10.3847/2041-8213/ab4284. 1908.09084.

[96] C. Messenger et J. Read. "Measuring a cosmological distance-redshift relationship using only gravitational wave observations of binary neutron star coalescences". *Phys. Rev. Lett.*, 108:091101, 2012. doi:10.1103/PhysRevLett.108.091101. 1107.5725.

[97] K. G. Malmquist. "On some relations in stellar statistics". *MeLuF*, 100:1–52, 1922.

[98] T. Bayes, Rev. "An essay toward solving a problem in the doctrine of chances". *Phil. Trans. Roy. Soc. Lond.*, 53:370–418, 1764. doi:10.1098/rstl.1763.0053.

[99] W. K. Hastings. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications". *Biometrika*, 57:97–109, 1970. doi:10.1093/biomet/57.1.97.

[100] I. Mandel, et al. "Extracting distribution parameters from multiple uncertain observations with selection biases". *MNRAS*, 486(1):1086–1093, 2019. doi:10.1093/mnras/stz896. 1809.02063.

[101] S. Vitale, et al. "Inferring the Properties of a Population of Compact Binaries in Presence of Selection Effects". 45, 2022. doi:10.1007/978-981-15-4702-7_45-1.

[102] M. Fishbach, et al. "Does the Black Hole Merger Rate Evolve with Redshift?" *Astrophys. J. Lett.*, 863(2):L41, 2018. doi:10.3847/2041-8213/aad800. 1805.10270.

[103] R. Gray. *Gravitational wave cosmology : measuring the Hubble constant with dark standard sirens*. Thèse de doctorat, University of Glasgow, 2021. doi:10.5525/gla.thesis.82438.

[104] W. J. Morokoff et R. E. Caflisch. "Quasi-monte carlo integration". *Journal of Computational Physics*, 122(2):218–230, 1995. ISSN 0021-9991. doi:https://doi.org/10.1006/jcph.1995.1209. URL https://www.sciencedirect.com/science/article/pii/S0021999185712090.

[105] C. Talbot et J. Golomb. "Growing pains: understanding the impact of likelihood uncertainty on hierarchical Bayesian inference for gravitational-wave astronomy". *MNRAS*, 526(3):3495–3503, 2023. doi:10.1093/mnras/stad2968. 2304.06138.

[106] W. M. Farr. "Accuracy Requirements for Empirically Measured Selection Functions". *RNAAS*, 3(5):66, 2019. doi:10.3847/2515-5172/ab1d5f. 1904.10879.

[107] G. Ashton, et al. "BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy". *ApJS*, 241(2):27, 2019. doi:10.3847/1538-4365/ab06fc. 1811.02042.

[108] G. Ashton, et al. "Bilby: A User-Friendly Bayesian Inference Library". Zenodo, 2019. doi:10.5281/zenodo.2602178.

[109] G. Ashton et C. Talbot. "BILBY-MCMC: an MCMC sampler for gravitational-wave inference". *MNRAS*, 507(2):2037–2051, 2021. doi:10.1093/mnras/stab2236. 2106.08730.

[110] S. Mastrogiovanni, et al. "Cosmology in the dark: How compact binaries formation impact the gravitational-waves cosmological measurements". *PoS*, EPS-HEP2021:098, 2022. doi:10.22323/1.398.0098. 2205.05421.

[111] W. Del Pozzo. "Inference of cosmological parameters from gravitational waves: Applications to second generation interferometers". *Phys. Rev. D*, 86:043011, 2012. doi:10.1103/PhysRevD.86.043011. URL https://link.aps.org/doi/10.1103/PhysRevD.86.043011.

[112] R. Gray, et al. "Cosmological inference using gravitational wave standard sirens: A mock data analysis". *Phys. Rev. D*, 101:122001, 2020. doi:10.1103/PhysRevD.101.122001. URL https://link.aps.org/doi/10.1103/PhysRevD.101.122001.

[113] H.-Y. Chen. "Systematic Uncertainty of Standard Sirens from the Viewing Angle of Binary Neutron Star Inspirals". *PhRvL*, 125(20):201301, 2020. doi:10.1103/PhysRevLett.125.201301. 2006.02779.

[114] R. Abbott et al. (LIGO Scientific, Virgo). "Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog". *Astrophys. J. Lett.*, 913(1):L7, 2021. doi:10.3847/2041-8213/abe949. 2010.14533.

[115] R. Farmer, et al. "Mind the gap: The location of the lower edge of the pair instability supernovae black hole mass gap". *arXiv*, 2019. doi:10.3847/1538-4357/ab518b. 1910.12874.

[116] C. Talbot et E. Thrane. "Measuring the binary black hole mass spectrum with an astrophysically motivated parameterization". *Astrophys. J.*, 856(2):173, 2018. doi:10.3847/1538-4357/aab34c. 1801.02699.

[117] C. Talbot, et al. "Parallelized Inference for Gravitational-Wave Astronomy". *Phys. Rev. D*, 100(4):043030, 2019. doi:10.1103/PhysRevD.100.043030. 1904.02863.

[118] D. Wysocki, et al. "Reconstructing phenomenological distributions of compact binaries via gravitational wave observations". *Phys. Rev. D*, 100:043012, 2019. doi:10.1103/PhysRevD.100.043012. URL https://link.aps.org/doi/10.1103/PhysRevD.100.043012.

[119] R. Abbott, et al. "Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog". *ApJL*, 913(1):L7, 2021. doi:10.3847/2041-8213/abe949. 2010.14533.

[120] L. Blanchet et T. Damour. "Post-newtonian generation of gravitational waves". Dans "Annales de l'IHP Physique théorique", tome 50, pages 377–408. 1989.

[121] P. Schmidt, et al. "Towards models of gravitational waveforms from generic binaries II: Modelling precession effects with a single effective precession parameter". *Phys. Rev. D*, 91(2):024043, 2015. doi:10.1103/PhysRevD.91.024043. 1408.1810.

[122] A. K. Gupta et S. Nadarajah. *Handbook of beta distribution and its applications*. CRC press, 2004.

[123] S. Miller, et al. "The low effective spin of binary black holes and implications for individual gravitational-wave events". *The Astrophysical Journal*, 895(2):128,

2020. doi:10.3847/1538-4357/ab80c0. URL https://dx.doi.org/10.3847/1538-4357/ab80c0.

[124] M. Mapelli et al. "Hierarchical black hole mergers in young, globular and nuclear star clusters: the effect of metallicity, spin and cluster properties". *Mon. Not. Roy. Astron. Soc.*, 505(1):339–358, 2021. doi:10.1093/mnras/stab1334. 2103.05016.

[125] M. Mapelli, et al. "The cosmic evolution of binary black holes in young, globular, and nuclear star clusters: rates, masses, spins, and mixing fractions". *Mon. Not. Roy. Astron. Soc.*, 511(4):5797–5816, 2022. doi:10.1093/mnras/stac422. 2109.06222.

[126] W. L. Freedman. "Measurements of the Hubble Constant: Tensions in Perspective". *Astrophys. J.*, 919(1):16, 2021. doi:10.3847/1538-4357/ac0e95. 2106.15656.

[127] S. Mastrogiovanni, et al. "On the importance of source population models for gravitational-wave cosmology". *Phys. Rev. D*, 104(6):062009, 2021. doi:10.1103/PhysRevD.104.062009. 2103.14663.

[128] S. R. Taylor, et al. "Hubble without the Hubble: Cosmology using advanced gravitational-wave detectors alone". *Phys. Rev. D*, 85:023535, 2012. doi:10.1103/PhysRevD.85.023535. 1108.5161.

[129] D. Wysocki, et al. "Reconstructing phenomenological distributions of compact binaries via gravitational wave observations". *Phys. Rev. D*, 100(4):043012, 2019. doi:10.1103/PhysRevD.100.043012. 1805.06442.

[130] M. Mancarella, et al. "Cosmology and modified gravitational wave propagation from binary black hole population models". *Phys. Rev. D*, 105(6):064030, 2022. doi:10.1103/PhysRevD.105.064030. 2112.05728.

[131] S. Mukherjee. "The redshift dependence of black hole mass distribution: is it reliable for standard sirens cosmology?" *Mon. Not. Roy. Astron. Soc.*, 515(4):5495–5505, 2022. doi:10.1093/mnras/stac2152. 2112.10256.

[132] K. Leyde, et al. "Current and future constraints on cosmology and modified gravitational wave friction from binary black holes". *JCAP*, 09:012, 2022. doi:10.1088/1475-7516/2022/09/012. 2202.00025.

[133] F. Aubin et al. "The MBTA pipeline for detecting compact binary coalescences in the third LIGO–Virgo observing run". *Class. Quant. Grav.*, 38(9):095004, 2021. doi:10.1088/1361-6382/abe913. 2012.11512.

[134] K. Cannon, et al. "Gstlal: A software framework for gravitational wave discovery". *SoftwareX*, 14:100680, 2021. ISSN 2352-7110. doi:https://doi.org/10.1016/

j.softx.2021.100680. URL https://www.sciencedirect.com/science/article/pii/S235271102100025X.

[135] S. A. Usman et al. "The PyCBC search for gravitational waves from compact binary coalescence". *Class. Quant. Grav.*, 33(21):215004, 2016. doi:10.1088/0264-9381/33/21/215004. 1508.02357.

[136] G. Ashton et al. "BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy". *Astrophys. J. Suppl.*, 241(2):27, 2019. doi:10.3847/1538-4365/ab06fc. 1811.02042.

[137] B. J. Owen et B. S. Sathyaprakash. "Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement". *Physical Review D*, 60(2):022002, 1999.

[138] P. A. R. Ade et al. (Planck). "Planck 2015 results. XIII. Cosmological parameters". *Astron. Astrophys.*, 594:A13, 2016. doi:10.1051/0004-6361/201525830. 1502.01589.

[139] M. Dominik, et al. "Double Compact Objects III: Gravitational Wave Detection Rates". *Astrophys. J.*, 806(2):263, 2015. doi:10.1088/0004-637X/806/2/263. 1405.7016.

[140] B. Moore, et al. "Gravitational-wave phasing for low-eccentricity inspiralling compact binaries to 3PN order". *Phys. Rev. D*, 93(12):124061, 2016. doi:10.1103/PhysRevD.93.124061. 1605.00304.

[141] A. Buikema et al. (aLIGO). "Sensitivity and performance of the Advanced LIGO detectors in the third observing run". *Phys. Rev. D*, 102(6):062003, 2020. doi:10.1103/PhysRevD.102.062003. 2008.01301.

[142] F. Acernese et al. (Virgo). "Increasing the Astrophysical Reach of the Advanced Virgo Detector via the Application of Squeezed Vacuum States of Light". *Phys. Rev. Lett.*, 123(23):231108, 2019. doi:10.1103/PhysRevLett.123.231108.

[143] P. Madau et M. Dickinson. "Cosmic Star Formation History". *Ann. Rev. Astron. Astrophys.*, 52:415–486, 2014. doi:10.1146/annurev-astro-081811-125615. 1403.0007.

[144] C. Karathanasis, et al. "GWSim: Python package for creating mock GW samples for different astrophysical populations and cosmological models of binary black holes". *A&A*, 677:A124, 2023. doi:10.1051/0004-6361/202245216. 2210.05724.

[145] S. Rinaldi, et al. "Evidence of evolution of the black hole mass function with redshift". *Astron. Astrophys.*, 684:A204, 2024. doi:10.1051/0004-6361/202348161. 2310.03074.

[146] N. Giacobbo, et al. "Merging black hole binaries: the effects of progenitor's metallicity, mass-loss rate and Eddington factor". *Mon. Not. Roy. Astron. Soc.*, 474(3):2959–2974, 2018. doi:10.1093/mnras/stx2933. 1711.03556.

[147] R. Abbott et al. (KAGRA, VIRGO, LIGO Scientific). "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3". *Phys. Rev. X*, 13(1):011048, 2023. doi:10.1103/PhysRevX.13.011048. 2111.03634.

[148] K. Belczynski, et al. "A Comprehensive study of binary compact objects as gravitational wave sources: Evolutionary channels, rates, and physical properties". *Astrophys. J.*, 572:407–431, 2001. doi:10.1086/340304. astro-ph/0111452.

[149] M. Spera et M. Mapelli. "Very massive stars, pair-instability supernovae and intermediate-mass black holes with the SEVN code". *Mon. Not. Roy. Astron. Soc.*, 470(4):4739–4749, 2017. doi:10.1093/mnras/stx1576. 1706.06109.

[150] M. Mapelli, et al. "The cosmic merger rate of stellar black hole binaries from the Illustris simulation". *Mon. Not. Roy. Astron. Soc.*, 472(2):2422–2435, 2017. doi:10.1093/mnras/stx2123. 1708.05722.

[151] S. Vitale, et al. "Measuring the star formation rate with gravitational waves from binary black holes". *Astrophys. J. Lett.*, 886(1):L1, 2019. doi:10.3847/2041-8213/ab50c0. 1808.00901.

[152] M. Renzo, et al. "Sensitivity of the lower-edge of the pair instability black hole mass gap to the treatment of time dependent convection". *Mon. Not. Roy. Astron. Soc.*, 493(3):4333–4341, 2020. doi:10.1093/mnras/staa549. 2002.08200.

[153] L. Wen. "On the eccentricity distribution of coalescing black hole binaries driven by the Kozai mechanism in globular clusters". *Astrophys. J.*, 598:419–430, 2003. doi:10.1086/378794. astro-ph/0211492.

[154] I. Mandel et A. Farmer. "Merging stellar-mass binary black holes". *Phys. Rept.*, 955:1–24, 2022. doi:10.1016/j.physrep.2022.01.003. 1806.05820.

[155] Y. Bouffanais, et al. "New insights on binary black hole formation channels after GWTC-2: young star clusters versus isolated binaries". *Mon. Not. Roy. Astron. Soc.*, 507(4):5224–5235, 2021. doi:10.1093/mnras/stab2438. 2102.12495.

[156] I. Bartos, et al. "Rapid and Bright Stellar-mass Binary Black Hole Mergers in Active Galactic Nuclei". *Astrophys. J.*, 835(2):165, 2017. doi:10.3847/1538-4357/835/2/165. 1602.03831.

[157] D. Gerosa et M. Fishbach. "Hierarchical mergers of stellar-mass black holes and their gravitational-wave signatures". *Nature Astron.*, 5(8):749–760, 2021. doi:10.1038/s41550-021-01398-w. 2105.03439.

[158] E. Thrane et C. Talbot. "An introduction to Bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models". *Publ. Astron. Soc. Austral.*, 36:e010, 2019. doi:10.1017/pasa.2019.2. [Erratum: Publ.Astron.Soc.Austral. 37, e036 (2020)], 1809.02293.

[159] S. Rinaldi et W. Del Pozzo. "(H)DPGMM: a hierarchy of Dirichlet process Gaussian mixture models for the inference of the black hole mass function". *Mon. Not. Roy. Astron. Soc.*, 509(4):5454–5466, 2021. doi:10.1093/mnras/stab3224. 2109.05960.

[160] A. Q. Cheng, et al. "What You Don't Know Can Hurt You: Use and Abuse of Astrophysical Models in Gravitational-wave Population Analyses". *Astrophys. J.*, 955(2):127, 2023. doi:10.3847/1538-4357/aced98. 2307.03129.

[161] A. M. Farah, et al. "No need to know: astrophysics-free gravitational-wave cosmology". 2024. 2404.02210.

[162] C. Kimball, et al. "Black Hole Genealogy: Identifying Hierarchical Mergers with Gravitational Waves". *Astrophys. J.*, 900(2):177, 2020. doi:10.3847/1538-4357/aba518. 2005.00023.

[163] T. A. Callister, et al. "No Evidence that the Majority of Black Holes in Binaries Have Zero Spin". *Astrophys. J. Lett.*, 937(1):L13, 2022. doi:10.3847/2041-8213/ac847e. 2205.08574.

[164] H. Tong, et al. "Population properties of spinning black holes using the gravitational-wave transient catalog 3". *Phys. Rev. D*, 106(10):103019, 2022. doi:10.1103/PhysRevD.106.103019. 2209.02206.

[165] M. Mould, et al. "Which black hole formed first? Mass-ratio reversal in massive binary stars from gravitational-wave data". *Mon. Not. Roy. Astron. Soc.*, 517(2):2738–2745, 2022. doi:10.1093/mnras/stac2859. 2205.12329.

[166] C. Adamcewicz, et al. "Which Black Hole Is Spinning? Probing the Origin of Black Hole Spin with Gravitational Waves". *Astrophys. J. Lett.*, 964(1):L6, 2024. doi:10.3847/2041-8213/ad2df2. 2311.05182.

[167] T. A. Callister, et al. "Who Ordered That? Unequal-mass Binary Black Hole Mergers Have Larger Effective Spins". *Astrophys. J. Lett.*, 922(1):L5, 2021. doi:10.3847/2041-8213/ac2ccc. 2106.00521.

[168] C. Adamcewicz et E. Thrane. "Do unequal-mass binary black hole systems have larger χeff? Probing correlations with copulas in gravitational-wave astronomy". *Mon. Not. Roy. Astron. Soc.*, 517(3):3928–3937, 2022. doi:10.1093/mnras/stac2961. 2208.03405.

[169] C. Adamcewicz, et al. "Evidence for a Correlation between Binary Black Hole Mass Ratio and Black Hole Spins". *Astrophys. J.*, 958(1):13, 2023. doi:10.3847/1538-4357/acf763. 2307.15278.

[170] R. Abbott et al. (KAGRA, VIRGO, LIGO Scientific). "Open Data from the Third Observing Run of LIGO, Virgo, KAGRA, and GEO". *Astrophys. J. Suppl.*, 267(2):29, 2023. doi:10.3847/1538-4365/acdc9f. 2302.03676.

[171] M. Zevin, et al. "Constraining Formation Models of Binary Black Holes with Gravitational-Wave Observations". *Astrophys. J.*, 846(1):82, 2017. doi:10.3847/1538-4357/aa8408. 1704.07379.

[172] M. Zevin, et al. "One Channel to Rule Them All? Constraining the Origins of Binary Black Holes Using Multiple Formation Pathways". *Astrophys. J.*, 910(2):152, 2021. doi:10.3847/1538-4357/abe40e. 2011.10057.

[173] A. Ray, et al. "Searching for binary black hole sub-populations in gravitational wave data using binned gaussian processes". 2024. 2404.03166.

[174] S. Biscoveanu, et al. "The Binary Black Hole Spin Distribution Likely Broadens with Redshift". *Astrophys. J. Lett.*, 932(2):L19, 2022. doi:10.3847/2041-8213/ac71a8. 2204.01578.

[175] M. Fishbach, et al. "Limits on Hierarchical Black Hole Mergers from the Most Negative χ_{eff} Systems". *Astrophys. J. Lett.*, 935(2):L26, 2022. doi:10.3847/2041-8213/ac86c4. 2207.02924.

[176] J. Godfrey, et al. "Cosmic Cousins: Identification of a Subpopulation of Binary Black Holes Consistent with Isolated Binary Evolution". 2023. 2304.01288.

[177] Y.-J. Li, et al. "Resolving the stellar-collapse and hierarchical-merger origins of the coalescing black holes". 2023. 2303.02973.

[178] V. Tiwari. "Exploring Features in the Binary Black Hole Population". *Astrophys. J.*, 928(2):155, 2022. doi:10.3847/1538-4357/ac589a. 2111.13991.

[179] C. Kimball et al. "Evidence for Hierarchical Black Hole Mergers in the Second LIGO–Virgo Gravitational Wave Catalog". *Astrophys. J. Lett.*, 915(2):L35, 2021. doi:10.3847/2041-8213/ac0aef. 2011.05332.

[180] R. Farmer, et al. "Constraints from gravitational wave detections of binary black hole mergers on the $^{12}C\,(\alpha,\gamma)^{16}O$ rate". *Astrophys. J. Lett.*, 902(2):L36, 2020. doi:10.3847/2041-8213/abbadd. 2006.06678.

[181] C. Karathanasis, et al. "Binary black holes population and cosmology in new lights: signature of PISN mass and formation channel in GWTC-3". *Mon. Not. Roy. Astron. Soc.*, 523(3):4539–4555, 2023. doi:10.1093/mnras/stad1373. 2204.13495.

[182] J. T. Galvez Ghersi et L. C. Stein. "A fixed point for black hole distributions". *Class. Quant. Grav.*, 38(4):045012, 2021. doi:10.1088/1361-6382/abcfd2. 2007.11578.

[183] G.-P. Li. "Constraining hierarchical mergers of binary black holes detectable with LIGO-Virgo". *Astron. Astrophys.*, 666:A194, 2022. doi:10.1051/0004-6361/202244257. 2208.11894.

[184] H. E. Bond. "Where is population III ?" *ApJS*, 248:606–611, 1981. doi:10.1086/159186.

[185] J. Weber. "Evidence for discovery of gravitational radiation". *Phys. Rev. Lett.*, 22:1320–1324, 1969. doi:10.1103/PhysRevLett.22.1320.

[186] R. A. Hulse et J. H. Taylor. "Discovery of a pulsar in a binary system". *Astrophys. J. Lett.*, 195:L51–L53, 1975. doi:10.1086/181708.

[187] V. I. Pustovoit. "On the direct detection of gravitational waves". *Phys. Usp.*, 59(10):1034–1051, 2016. doi:10.3367/UFNe.2016.03.037900.

[188] R. L. Forward. "Wide Band Laser Interferometer Gravitational Radiation Experiment". *Phys. Rev. D*, 17:379–390, 1978. doi:10.1103/PhysRevD.17.379.

[189] P. R. Saulson et al. "INTERFEROMETRIC GRAVITATIONAL WAVE DETECTION AT MIT". *Conf. Proc. C*, 861214:15–17, 1986.

[190] D. Shoemaker, et al. "Noise Behavior of the Garching 30-meter Prototype Gravitational Wave Detector". *Phys. Rev. D*, 38:423–432, 1988. doi:10.1103/PhysRevD.38.423.

[191] B. P. Abbott et al. (LIGO Scientific, Virgo). "GW150914: The Advanced LIGO Detectors in the Era of First Discoveries". *Phys. Rev. Lett.*, 116(13):131103, 2016. doi:10.1103/PhysRevLett.116.131103. 1602.03838.

[192] B. P. Abbott et al. "Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy". *Phys. Rev. D*, 93(11):112004, 2016. doi:10.1103/PhysRevD.93.112004. [Addendum: Phys.Rev.D 97, 059901 (2018)], 1604.00439.

[193] B. P. Abbott et al. (LIGO Scientific, Virgo). "Directly comparing GW150914 with numerical solutions of Einstein's equations for binary black hole coalescence". *Phys. Rev. D*, 94(6):064035, 2016. doi:10.1103/PhysRevD.94.064035. 1606.01262.

[194] B. P. Abbott et al. (LIGO Scientific, Virgo). "Properties of the Binary Black Hole Merger GW150914". *Phys. Rev. Lett.*, 116(24):241102, 2016. doi:10.1103/PhysRevLett.116.241102. 1602.03840.

[195] S. Adrian-Martinez et al. (ANTARES, IceCube, LIGO Scientific, Virgo). "High-energy Neutrino follow-up search of Gravitational Wave Event GW150914 with ANTARES and IceCube". *Phys. Rev. D*, 93(12):122010, 2016. doi:10.1103/PhysRevD.93.122010. 1602.05411.

[196] T. D. Abbott et al. (LIGO Scientific, Virgo). "Improved analysis of GW150914 using a fully spin-precessing waveform Model". *Phys. Rev. X*, 6(4):041014, 2016. doi:10.1103/PhysRevX.6.041014. 1606.01210.

[197] B. P. Abbott et al. (LIGO Scientific, Virgo). "The basic physics of the binary black hole merger GW150914". *Annalen Phys.*, 529(1-2):1600209, 2017. doi:10.1002/andp.201600209. 1608.01940.

[198] B. P. Abbott et al. (LIGO Scientific, Virgo). "Effects of waveform model systematics on the interpretation of GW150914". *Class. Quant. Grav.*, 34(10):104002, 2017. doi:10.1088/1361-6382/aa6854. 1611.07531.

[199] B. P. Abbott et al. (LIGO Scientific, Virgo). "The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914". *Astrophys. J. Lett.*, 833(1):L1, 2016. doi:10.3847/2041-8205/833/1/L1. 1602.03842.

[200] B. P. Abbott et al. (LIGO Scientific, Virgo). "Tests of general relativity with GW150914". *Phys. Rev. Lett.*, 116(22):221101, 2016. doi:10.1103/PhysRevLett.116.221101. [Erratum: Phys.Rev.Lett. 121, 129902 (2018)], 1602.03841.

[201] B. P. Abbott et al. (LIGO Scientific, Virgo). "Observing gravitational-wave transient GW150914 with minimal assumptions". *Phys. Rev. D*, 93(12):122004, 2016. doi:10.1103/PhysRevD.93.122004. [Addendum: Phys.Rev.D 94, 069903 (2016)], 1602.03843.

[202] B. P. Abbott et al. (LIGO Scientific, Virgo, ASKAP, BOOTES, DES, Fermi GBM, Fermi-LAT, GRAWITA, INTEGRAL, iPTF, InterPlanetary Network, J-GEM, La Silla-QUEST Survey, Liverpool Telescope, LOFAR, MASTER, MAXI, MWA, Pan-STARRS, PESSTO, Pi of the Sky, SkyMapper, Swift, C2PU, TOROS, VISTA). "Localization and broadband follow-up of the gravitational-wave transient GW150914". *Astrophys. J. Lett.*, 826(1):L13, 2016. doi:10.3847/2041-8205/826/1/L13. 1602.08492.

[203] B. P. Abbott et al. (LIGO Scientific, Virgo). "GW150914: Implications for the stochastic gravitational wave background from binary black holes". *Phys. Rev. Lett.*, 116(13):131102, 2016. doi:10.1103/PhysRevLett.116.131102. 1602.03847.

[204] B. P. Abbott et al. (LIGO Scientific, Virgo). "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914".

Class. Quant. Grav., 33(13):134001, 2016. doi:10.1088/0264-9381/33/13/134001. 1602.03844.

[205] B. P. Abbott et al. (LIGO Scientific, Virgo). "Astrophysical Implications of the Binary Black-Hole Merger GW150914". *Astrophys. J. Lett.*, 818(2):L22, 2016. doi:10.3847/2041-8205/818/2/L22. 1602.03846.

[206] B. P. Abbott et al. (LIGO Scientific). "Calibration of the Advanced LIGO detectors for the discovery of the binary black-hole merger GW150914". *Phys. Rev. D*, 95(6):062003, 2017. doi:10.1103/PhysRevD.95.062003. 1602.03845.

[207] T. Akutsu et al. (KAGRA). "Overview of KAGRA: Calibration, detector characterization, physical environmental monitors, and the geophysics interferometer". *PTEP*, 2021(5):05A102, 2021. doi:10.1093/ptep/ptab018. 2009.09305.

[208] B. P. Abbott et al. (LIGO Scientific). "Exploring the Sensitivity of Next Generation Gravitational Wave Detectors". *Class. Quant. Grav.*, 34(4):044001, 2017. doi:10.1088/1361-6382/aa51f4. 1607.08697.

[209] K. Chamberlain et N. Yunes. "Theoretical Physics Implications of Gravitational Wave Observation with Future Detectors". *Phys. Rev. D*, 96(8):084039, 2017. doi:10.1103/PhysRevD.96.084039. 1704.08268.

[210] S. Dwyer, et al. "Gravitational wave detector with cosmological reach". *Phys. Rev. D*, 91(8):082001, 2015. doi:10.1103/PhysRevD.91.082001. 1410.0612.

[211] R. Essick, et al. "Frequency-dependent responses in third generation gravitational-wave detectors". *Phys. Rev. D*, 96(8):084004, 2017. doi:10.1103/PhysRevD.96.084004. 1708.06843.

[212] D. Reitze et al. "Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO". *Bull. Am. Astron. Soc.*, 51(7):035, 2019. 1907.04833.

[213] K. K. Y. Ng, et al. "Probing multiple populations of compact binaries with third-generation gravitational-wave detectors". *Astrophys. J. Lett.*, 913(1):L5, 2021. doi:10.3847/2041-8213/abf8be. 2012.09876.

[214] F. Iacovelli, et al. "Forecasting the Detection Capabilities of Third-generation Gravitational-wave Detectors Using GWFAST". *Astrophys. J.*, 941(2):208, 2022. doi:10.3847/1538-4357/ac9cd4. 2207.02771.

[215] F. Aubin. *Recherche à faible latence d'ondes gravitationnelles émises lors de coalescences de binaires compactes durant la troisième période d'observation de Advanced Virgo et Advanced LIGO*. Theses, Université Savoie Mont Blanc, 2020. URL https://theses.hal.science/tel-03336046.

[216] M. Maggiore. *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford University Press, 2007. ISBN 978-0-19-171766-6, 978-0-19-852074-0. doi:10.1093/acprof:oso/9780198570745.001.0001.

[217] C. Cahillane et G. Mansell. "Review of the Advanced LIGO Gravitational Wave Observatories Leading to Observing Run Four". *Galaxies*, 10(1):36, 2022. doi:10.3390/galaxies10010036. 2202.00847.

[218] E. Cuoco, et al. "Strategy for signal classification to improve data quality for Advanced Detectors gravitational-wave searches". *Nuovo Cim. C*, 40(3):124, 2017. doi:10.1393/ncc/i2017-17124-4.

[219] J. Glanzer et al. "Data quality up to the third observing run of advanced LIGO: Gravity Spy glitch classifications". *Class. Quant. Grav.*, 40(6):065004, 2023. doi:10.1088/1361-6382/acb633. 2208.12849.

[220] S. Mukherjee, et al. "Classification of glitch waveforms in gravitational wave detector characterization". *J. Phys. Conf. Ser.*, 243:012006, 2010. doi:10.1088/1742-6596/243/1/012006.

[221] J. Powell, et al. "Classification methods for noise transients in advanced gravitational-wave detectors II: performance tests on Advanced LIGO data". *Class. Quant. Grav.*, 34(3):034002, 2017. doi:10.1088/1361-6382/34/3/034002. 1609.06262.

[222] B. P. Abbott et al. (LIGO Scientific, Virgo). "A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals". *Class. Quant. Grav.*, 37(5):055002, 2020. doi:10.1088/1361-6382/ab685e. 1908.11170.

[223] E. Cuoco et al. "Enhancing Gravitational-Wave Science with Machine Learning". *Mach. Learn. Sci. Tech.*, 2(1):011002, 2021. doi:10.1088/2632-2153/abb93a. 2005.03745.

[224] R. Essick, et al. "idq: statistical inference of non-gaussian noise with auxiliary degrees of freedom in gravitational-wave detectors". *Machine Learning: Science and Technology*, 2, 2020. doi:10.1088/2632-2153/abab5f.

[225] R. Essick, et al. "iDQ: Statistical Inference of Non-Gaussian Noise with Auxiliary Degrees of Freedom in Gravitational-Wave Detectors". 2020. 2005.12761.

[226] J. Neyman et E. S. Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". *Phil. Trans. Roy. Soc. Lond. A*, 231(694-706):289–337, 1933. doi:10.1098/rsta.1933.0009.

[227] F. Robinet, et al. "Omicron: a tool to characterize transient noise in gravitational-wave detectors". *SoftwareX*, 12:100620, 2020. doi:10.1016/j.softx.2020.100620. 2007.11374.

[228] R. Essick, et al. "Optimizing Vetoes for Gravitational-Wave Transient Searches". *Class. Quant. Grav.*, 30:155010, 2013. doi:10.1088/0264-9381/30/15/155010. 1303.7159.

[229] K. Hornik, et al. "Multilayer feedforward networks are universal approximators". *Neural Networks*, 2(5):359–366, 1989. ISSN 0893-6080. doi:https://doi.org/10.1016/0893-6080(89)90020-8. URL https://www.sciencedirect.com/science/article/pii/0893608089900208.

[230] D. Davis, et al. "Incorporating information from LIGO data quality streams into the PyCBC search for gravitational waves". *Phys. Rev. D*, 106(10):102006, 2022. doi:10.1103/PhysRevD.106.102006. 2204.03091.

[231] C. M. Biwer, et al. "PyCBC Inference: A Python-based parameter estimation toolkit for compact binary coalescence signals". *Publ. Astron. Soc. Pac.*, 131(996):024503, 2019. doi:10.1088/1538-3873/aaef0b. 1807.10312.

[232] N. Andres et al. "Assessing the compact-binary merger candidates reported by the MBTA pipeline in the LIGO–Virgo O3 run: probability of astrophysical origin, classification, and associated uncertainties". *Class. Quant. Grav.*, 39(5):055002, 2022. doi:10.1088/1361-6382/ac482a. 2110.10997.

[233] T. Isogai (LIGO Scientific, Virgo). "Used percentage veto for LIGO and virgo binary inspiral searches". *J. Phys. Conf. Ser.*, 243:012005, 2010. doi:10.1088/1742-6596/243/1/012005.

[234] F. Acernese et al. (Virgo). "Virgo detector characterization and data quality: tools". *Class. Quant. Grav.*, 40(18):185005, 2023. doi:10.1088/1361-6382/acdf36. 2210.15634.