Topics in Compact Object Astrophysics and Fundamental Physics with Current and Future Gravitational Wave Observations

by
Zhenwei Lyu

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Physics

Guelph, Ontario, Canada

ABSTRACT

TOPICS IN COMPACT OBJECT ASTROPHYSICS AND FUNDAMENTAL PHYSICS WITH CURRENT AND FUTURE GRAVITATIONAL WAVE OBSERVATIONS

Zhenwei Lyu
University of Guelph, 2022

Advisor:
Dr. Huan Yang

The thesis is devoted to compact object astrophysics and fundamental physics with current and future gravitational wave (GW) observations. Since the first detection of GW event GW150914 by LIGO in 2015, there have been more than 90 transients observed by the LIGO, Virgo and KAGRA collaboration up till now. Additionally, more observations of stellar mass compact binaries would be detected by current and the second generation detectors as well as observations of Extreme Mass Ratio Inspirals (EMRIs) by future space-based gravitational wave detectors such as LISA, TianQin and TaiJi.

At the beginning (Chapter 1), a review of data analysis techniques on gravitational waves is introduced, followed by an introduction to the focused projects.

In Part I, we study the detections and constraints on potential effects postulated by Einstein's general relativity (GR) and beyond, such as probing tidal-induced i-mode in inspiraling binary neutron stars (BNSs) (Chapter 2), searching nuclear coupling of axions from the BNSs (Chapter 3), and constraining Einstein-dilation Gauss-Bonnet (EdGB) gravity from black hole neutron star (BHNS) events (Chapter 4).

In Part II, we propose a new hybrid waveform model, which illustrates comparable accuracy in characterizing tidal effects as the effective-one-body (EOB) waveform and numerical relativity simulation.

Part III investigates an important extreme mass ratio inspirals (EMRIs) formation channel driven by active galactic nuclei (AGN) accretion disk. Chapter 6 introduces the expectation of a higher rate of occurrence of EMRIs in AGN for spaceborne gravitational wave detectors; and in Chapter 7, we show that the AGN channel is much more promising to produce mass-gap EMRIs.

Dedication

To my mom *Suling Niu* and dad *Chengkun Lyu*

Acknowledgements

This work would not have been possible without the constant support, guidance, and patience of my advisor Dr. Huan Yang. His knowledge and passion toward this field has shaped my understanding of the role of researcher today. I have learnt a lot from him: not only the positive frame of mind, but also critical thinking skills during my four years of doctoral studies.

It is my best fortune to have such an opportunity to work with the amazing professors and colleagues in these four years. I am always grateful for their valuable advice. I would like to express my sincere gratitude to my committee, Eric Poisson, William East and Alexandros Gezerlis for their insightful comments and suggestions. Additionally, I would like to offer my special thanks to Daniel Siegel for his kindness in writing the recommendation letter.

I am honoured to work with Zhen Pan, Jun Zhang, Kent Yagi, Xuefeng Feng, Nan Jiang, Béatrice Bonga, Laura Sagunski, Mairi Sakellariadou, Junwu Huang, Matthew C. Johnson, Ńestor Ortiz. My excellent colleagues gave me countless encouragement and enthusiasm over these years. Specially, I would like to thank GGG group members at University of Guelph and strong gravity members at Perimeter Institute for Theoretical Physics, including but not limited to: Aman Agarwal, Dhruv Desai, Erik Schnetter, Haomiao Jiang, Hassan Khalvati, Luis Lehner, Maxence Corman, Niayesh Afshordi, Reed Essick, Sharaban Tahura, Soham Mukherjee. They had always been ready to offer me help and support. It has been my greatest pleasure to work with them.

Moreover, David Tsang, Andrea Passamonti, Katerina Chatziioannou, Claudia de Rham, Nicolás Yunes, Tanja Hinderer, Adam Pound, Hui-Min Fan, and Alberto Sesana had provided great assis-

tance for my research. Monica Yin Tan had offered me language support. This project is never able to be completed without them.

I would like to express my sincere gratitude to those for providing financial support without which this research would not have been possible. All of the work presented in the thesis was supported by the Natural Sciences and Engineering Research Council of Canada and in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities. This thesis is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org/), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain.

Work in Chapter 3 was supported by the European Union's Horizon 2020 Research Council grant 724659 MassiveCosmo ERC-2016-COG, by the National Science and Engineering Research Council through a Discovery grant, in part by the Science and Technology Facility Council (STFC), United Kingdom, under research grant No. ST/P000258/1. In working on Chapter 4, we acknowledge support from the Owens Family Foundation, NSF Grant PHY-1806776, NASA

Grant 80NSSC20K0523, a Sloan Foundation, the COST Action GWverse CA16104 and JSPS KAKENHI Grants No. JP17H06358. Work in Chapter 5 makes use of the Black Hole Perturbation Toolkit and was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grants No.XDA1502070401.

I would deeply appreciate the support from my family, especially the unconditional support and selfless love from my mom. I would thank her for the endless encouragement and support: it is she who kept my faith, and gave me courage to pursue my dream life.

Contents

List of Tables

List of Figures

List of Abbreviations

AGN Active Galactic Nuclei 2, 3, 38

BBH Binary Black Hole 37

BH Black Hole 1

BHNS Black Hole Neutron Star 1, 2, 29

BNS Binary Neutron Star 1, 2, 29, 31

CDF Cumulative Distribution Function 11

CMB Cosmic Microwave Background 4

EdGB Einstein-dilation Gauss-Bonnet 1, 2

EM ElectroMagneti 5

EMRIs Extreme Mass Ratio Inspirals 2, 3

EOB Effective-One-Body 1, 3

EoS Equation of State 7

FIM Fisher information matrix 21

GR General Relativity 2

GW Gravitational Waves 1, 2

ISCO Innermost Stable Circular Orbit 23

KAGRA KAmioka GRAvitational wave detector 9

LIGO Laser Interferometer Gravitational-wave Observatory 1, 9

LISA Laser Interferometer Space Antenna 2

MBH Massive Black Holes 2, 38

MCMC Markov-Chain Monte Carlo 9

MGO Mass-Gap Object 40

NS Neutron Star 10, 29

PBH Primordial Black Holes 6

PN Post-Newtonian 1, 23

ppE parameterized post-Einsteinian 23, 33

PSD Power Spectral Density 9, 14

PTAs Pulsar Timing Arrays 4

sBH Stellar-mass Black Hole 38

SDDR Savage-Dickey Density Ratio 16

SGWB Stochastic Gravitational-wave Background 6

SNR Signal-to-Noise Ratio 14

SPA Stationary Phase Approximation 25, 30

TaiJi a Chinese space-based gravitational wave observatory 2

TaylorF2 Taylor F2 approximant frequency domain waveform model 30

TianQin a Chinese space-based gravitational wave observatory 2

Virgo Virgo is an interferometric gravitational-wave antenna 1, 9

Chapter 1

Introduction

The whole PhD thesis consists of three parts which all focus on gravitational physics. First, I will begin with the background of gravitational wave (GW) astronomy and data analysis techniques of GWs used in the research, such as Bayesian inference, model selection, Fisher analysis and so on. Then, I will introduce the backgrounds and results of the projects studied in the research.

Part I focus on GW data analysis with inspiraling compact binaries. Benefitted from the detection of GWs by LIGO and Virgo [1–4], in Chapter 2, we search tidal-induced mode resonance in binary neutron stars (BNSs) event GW170817 [5]; Chapter 3 explores the probing of nuclear coupling of axions from GW170817 [6]; and in Chapter 4, we demonstrate the constraints on Einstein-dilation Gauss-Bonnet (EdGB) gravity from black hole neutron star (BHNS) events GW200105 and GW200115 [7].

In Part II we propose a new hybrid waveform model, which shows comparable accuracy to the effective-one-body (EOB) waveform and numerical relativity simulation in characterizing tidal effects. Previous studies on compact object tidal effects have been carried out in the post-Newtonian (PN) and EOB formalisms [8]. Within the black hole (BH) perturbation framework, we propose a frequency domain hybrid waveform model [9] in Chapter 5, It is used to characterize the tidal influence in the expansion of mass ratios, while higher-order PN corrections are naturally included.

This model shows comparable accuracy as the EOB waveform in characterizing the tidal effects, as calibrated by numerical relativity simulations.

Part III mainly explains the important extreme mass ratio inspirals (EMRIs) formation channel driven by active galactic nuclei (AGN) accretion disk we developed. There are plans for space-based detectors such as LISA [10], TianQin [11] and TaiJi [12] other than the ground-based detection of GWs. One of the targeting GW sources of space-based detectors is EMRIs: a stellar mass compact object orbiting around a massive black hole (MBH) with mass around $\sim 10^6 M_\odot$. We suggest that EMRIs in AGN may be more common for spaceborne GW detectors [13] (Chapter 6) and AGN channel is more promising to produce mass-gap EMRIs [14] (Chapter 7).

1.1 Gravitational Wave Astronomy and Physics

GWs are ripples of spacetime propagating in the speed of light. The existence of GWs is a prediction of Einstein's general relativity and also valid for all other theories of gravity that obey special relativity. GWs are quadrupole radiation from accelerating compact objects carrying energy and momentum information as well as intrinsic properties of the system, which are different from other astrophysical signal carriers such as photos and cosmic ray. Any important influences to the dynamics of the compact system will leave probably detectable signatures to the detected GW and it is possible to not only extract intrinsic information of GW emitters but also other signatures from possible effects by analysing the observed data stream.

As a matter of fact, any accelerating object with mass will generate GWs with an extremely small characteristic strength. We usually define GW strain as $h(t) := \delta L(t)/L$, where L is the separation between two reference free points and $\delta L(t)$ is the passing GW induced displacement. As for compact stellar mass binary coalescence such as BH-BH, BH-NS and NS-NS, the typical strain amplitude is about order 10^{-21} or less. For LIGO detectors with arm lengths $L = 4\,\mathrm{km}$, the displacement δL is around $\sim 10^{-17}\mathrm{m}$ or less, which is even smaller than the radius of a proton

(8.4×10^{-16}m). Enhanced Michelson interferometer could help us detect such a small effect and most of current and future GW detectors are base on Michelson interferometer technology, such as current ground-based detectors LIGO, Virgo and KAGRA [2–4], future ground-based ET [15] and CE [16], and future space-based LISA [10], TaiJi [12] and TianQin [11].

In this section, I will first present a brief overview of GW observational history, then some astrophysical and physical questions that have been or could be answered by GW observation, and prospects for the future.

1.1.1 History of Gravitational Wave Observation

The first indirect evidence of GWs is given by the observation of the orbital decay of the Hulse–Taylor binary pulsar, which shows exactly the same as the prediction from general relativity [17]. Hulse and Taylor were awarded the Nobel Prize in Physics for this discovery in 1993. As for direct observation, it is more challenging. Indeed, there are many attempts over the past half century. They are all failed due to both technology level and types of detectors, until on September 14 in 2015, the first direct GW detection is achieved by LIGO [18]. Currently, the most important GW detection collaboration (LIGO, Virgo and KAGRA) has reported up to 90 GW events from binary black hole (BBH), BNS and BHNS mergers [1–4] and three physicist (Barry Barish, Kip Thorne and Rainer Weiss) were awarded the 2017 Nobel Prize in Physics for leading this work.

In the near future, existing ground-based detectors LIGO and Virgo will upgrade to A+ and AdV+ [2, 19] with improved sensitivities by more than two levels to current designs, and a new aLIGO-type detector LIGo-india will join the detector network to enhance the combined sensitivity [19]. Besides, current operating GEO600 are concentrating on testing technologies for future detectors [20].

The next 3rd generation (3G) ground-based GW detectors will increase the sensitivity by more than 10 times than current detectors, which will largely increase the number of events detected

and the astrophysical distance that could reach. As defined previously, the GW strain $h(t) := \delta L(t)/L$, one straightforward way to increase the detector sensitivity is by lengthening the arm length. Therefore, future 3G detector designs (one is ET located in Europe with 10 km arm lengths, and the other one is CE in the USA with 40 km arm lengths) are both increasing the baseline several time than LIGO [19].

Future space-based detectors such as LISA with baseline lengths $\sim 2.5\times10^9$ km are designed to probe GW frequency ranging from 10^{-5}Hz to 10^{-1}Hz generated by EMRIs and MBH coalescence with mass $10^2 \sim 10^7 M_\odot$ [10]. Relative to ground-based detectors, the GW induced displacement δL in LISA is about 10^{-12}m or less, which is about 5 order of magnitude greater than the scale of ground-based detectors $\delta L \sim 10^{-17}$m. As we can see from previous comparison, the longer baselines, the higher GW induced displacement δL and the lower optimal response frequency band.

There are many important astrophysical sources such as (super-)MBH inspiral and merger, EMRIs, stellar mass binary inspiral and merger, pulsars and supernovae, which will generate GWs with frequencies ranging from 10^{-9}Hz to several thousands Hz. The frequency band below 10^{-9}Hz could be generated by scalar and tensor perturbations at the inflation stage of the universe leaving imprints to the cosmic microwave background (CMB), which may be probed by measuring the B-mode polarization of CMB by BICEP and Keck Array experiments [21]. Pulsar timing arrays (PTAs) are capable to detect GWs within range $10^{-9} \sim 10^{-6}$Hz by measuring pulse arrival times [22–25]. Space-based and ground-based detectors are designed to detect GWs ranging from $10^{-5} \sim 10^{-1}$Hz and $10^0 \sim 10^3$Hz, respectively.

In addition to previous mentioned detectors, there are several other methods, such as resonant mass antennas [26–28], atom interferometry [29, 30].

Multi-messenger Astronomy

Multi-messenger astronomy is referred as collecting more than one astrophysical messengers from the same object or event. For example, on August 17 in 2017, 1.74 seconds after detecting the BNS

event GW170817, many observatories observed signals along almost all electromagnetic (EM) frequency bands, including gamma rays, X-rays, ultraviolet, visible light and radio from galaxy NGC 4993 [31–33], which is the first and the most important multi-messenger observation involving GWs till now. This breakthrough multi-messenger observation provides a lot of knowledge of the understanding of the universe [32, 33]: the progenitors of short gamma-ray bursts are thought as BNS mergers; EM observation shows that kilonova is a result of BNS merger which will produce chemical elements heavier than iron through r-process nuclei; the speed of GWs is almost the same as the speed of EM; GWs together with EM signal provides an independent estimate of the Hubble constant. In the near future, we hope there may be many GW-EM multi-messenger observations, especially, when more powerful GW detectors are established and start observations.

1.1.2 Gravitational Wave Astrophysics, Cosmology and Fundamental Physics

With the discovery of the first BBH coalescence GW150914 [18], the era of GW science has started. In the next few decades, we expect to observe thousands of events from compact binary coalescence with 3G ground-based detectors and space-based detectors [34]. Since the observed data directly encode information about the sources, including the masses, spins, sky location, luminosity distance and other intrinsic and external properties of the compact objects, by extracting these useful information of the detected events, one could explore many critical problems in astrophysics, cosmology and fundamental physics.

Astrophysics

The mass distribution of observed GW events will be used to determine the formation rates and channels of compact objects [35] which could be astrophysical models (formed from supernova and mergers of compact objects) [36] and cosmological models (cosmic string and/or inflationary mechanisms) [37, 38]. As LIGO has observed several lower mass-gap object mergers (with mass

in $3 \sim 5M_{\odot}$), it is possible to interpret the underlying formation mechanism with the inferred mass and spin distributions, especially when we observed enough cases [39, 40]. No detection of BHs with mass $0.2 \sim 1M_{\odot}$ has already been perform to constrain primordial black holes (PBHs) since BH with mass lower than $3M\odot$ cannot be produced by stellar evolution process [41]. In Part III, we propose an important EMRI formation channel driven by AGN disk, which will be tested by future detection of EMRIs in LISA mission.

On the other hand, all random sub-threshold GW signals form stochastic gravitational-wave background (SGWB). The detection of SGWB will help us constrain and determine various formation models as well [42–44].

Cosmology

GW events can serve as **standard siren** by estimating the luminosity distance, especially with BNS events, we may observe the EM counterparts, which will help us reduce the uncertainties of Hubble constant [31]. Therefore, they have the potential to solve some crucial cosmological issues such as Hubble tension, dark mater and the existence of dark energy. Accurate estimates of Hubble constant by GWs is another independent method, which may resolve the Hubble tension problem to clarify whether it is measurement issues, unknown large-scale structures or new physics beyond the ΛCDM model[45–47].

Future 3G ground-based detectors are expected to observed compact binary coalescences to redshift $z = 3$ and LISA mission will provide observations of stellar mass compact binaries, EMRIs and super-massive BBHs up to redshift 0.2 [48], 1 [49] and 10 [50], respectively. With the high redshift observations, GWs as standard siren can also help us probe the universe expansion history, dark matter and the equation of state of dark energy, which may provide insights in explaining the accelerated expanding universe.

Fundamental Physics

GWs generated from the most extreme environment of gravity (strong field and high curvature) have huge potentials to explore fundamental physics including but not limited to probe equation of state (EoS) of NS and test GR and beyond GR theories. Currently, the estimation of tidal deformation from event GW170817 has excluded some stiffest EoS and constrained the radii to the range $9 \sim 13$km [51]. We also probe tidal excitation of resonant mode in Chapter 2 in BNS event GW170817 and predict that such signal may be observable with A+ and more likely with third-generation detectors ET and CE.

In addition to determine the EoS of NS, GW observations can be used to test the dispersion relation, which may be caused by non-zero graviton or violations of Lorentz invariance. Current constraints are still consistent with GR [52]. Besides testing GR, we can also probe modified theories of gravity, such as EdGB gravity as demonstrated in Chapter 4. we provide a stronger constraint on $\sqrt{\alpha_{\rm GB}} \lesssim 1.33\,\mathrm{km}$ by individual event GW200115.

GWs also provide a unique way to explore dark matter, including probing PBHs by formation channels and SGWB studies, exploring weakly interacting massive particles (WIMPs) and extremely light particles, such as axions [53, 54]. In Chapter 3, we explore light axions from GW170817, which is a unique prediction of string theory, as well as one of the most compelling candidates for dark matter.

With the upgrade of GW detectors, we expect to observed more and more compact binary coalescences, which will lead to an era of precision GW science. In the thesis, the next things to be clarified would be the general methods of data analysis.

1.2 Data Analysis Techniques on Gravitational Waves

1.2.1 Bayesian Inference

To unveil the basic information of compact binary systems behind GW events, one usually makes use of a reliable method – Bayesian inference [55, 56]. According to the Bayes' theorem, a **posterior** probability $p(\boldsymbol{\vartheta}|d,\mathcal{H})$ on a set of parameters $\boldsymbol{\vartheta}$ from data d under a given hypothesis $\mathcal{H}$ is given by:

$$p(\boldsymbol{\vartheta}|d,\mathcal{H}) = \frac{p(d|\boldsymbol{\vartheta},\mathcal{H})\ p(\boldsymbol{\vartheta}|\mathcal{H})}{p(d|\mathcal{H})}, \tag{1.1}$$

where $p(d|\boldsymbol{\vartheta},\mathcal{H})$ is the likelihood function while $p(\boldsymbol{\vartheta}|\mathcal{H})$ is the prior on $\boldsymbol{\vartheta}$ of hypothesis $\mathcal{H}$. $p(d|\mathcal{H})$ called **evidence** is the probability of the data d given hypothesis $\mathcal{H}$. Mathematically, one could evaluate the evidence (also known as marginal likelihood) by marginalizing the likelihood over all possible parameters $\boldsymbol{\vartheta}$ of the hypothesis:

$$p(d|\mathcal{H}) := \int d\boldsymbol{\vartheta}\ p(d|\boldsymbol{\vartheta},\mathcal{H})\ p(\boldsymbol{\vartheta}|\mathcal{H}), \tag{1.2}$$

With a stationary Gaussian noise assumption, the **log-likelihood** function $\log p(d|\boldsymbol{\vartheta},\mathcal{H})$ can be expressed as:

$$\log p(d|\boldsymbol{\vartheta},\mathcal{H}) = \log\bar{\alpha} - \frac{1}{2}\sum_k \langle d_k - h_k(\boldsymbol{\vartheta})\,|\,d_k - h_k(\boldsymbol{\vartheta})\rangle\,, \tag{1.3}$$

where the index k refers to different detectors and $\log\bar{\alpha}$ is the normalization factor while d_k and $h_k(\boldsymbol{\vartheta})$ are the data and waveform templates from given detectors. The noise-weighted **inner product** between complex functions a and b is defined as:

$$\langle a(t)\,|\,b(t)\rangle := 2\int_{f_{\rm low}}^{f_{\rm high}} \frac{\tilde{a}^*(f)\tilde{b}(f) + \tilde{a}(f)\tilde{b}^*(f)}{S_n(f)}df. \tag{1.4}$$

Here $*$ refers to a complex conjugate, $\tilde{a}(f)$ denotes the Fourier transform of $a(t)$, $f_{\rm low}$ and $f_{\rm high}$

are the low and high frequency cutoff of GW data respectively. For ground based detectors such as LIGO, Virgo and KAGRA, the low frequency detectability is about 20 Hz (may change with specific GW events) and the high frequency cutoff should be greater than the maximum frequency of the GW signal. $S_n(f)$ is the one-sided (only positive frequencies) **power spectral density** (PSD) of given detectors defined as:

$$\langle \tilde{n}^*(f)\, \tilde{n}(f') \rangle = \frac{1}{2}\delta(f - f')S_n(f)\,, \tag{1.5}$$

where $n(f)$ is the Fourier transform of detector observed time series $n(t)$ which is usually assumed to be stationary and Gaussian. The stationarity means the characteristic of the noise process do not change over time $\langle n(t) \rangle = 0$ and the assumption of Gaussianity indicates the probability density for noise is a Gaussian distribution.

A common way to estimate the posterior probability is to generate large enough sizes of samples from posterior probability $p(\vartheta|d, \mathcal{H})$ and then draw histograms and contour plots of parameters ϑ. Two of the well implemented methods are Markov Chain Monte Carlo (MCMC) and nested sampling [57, 58] techniques.

1.2.2 MCMC and Nested Sampling

MCMC is one of the most popular sampling methods used to generate equal weighted samples proportional to the posterior distribution. Whereas it is hard to estimate the evidence when we need to compare two models (see subsection 1.2.6), nested sampling proposed by Skilling [57, 58] is an alternative approach which is able to generate posterior samples as well as estimating the evidence. Afterall, it is time-consuming to put these in practice. Fortunately, there are several well-organized packages which have already employed MCMC and nested sampling, such as packages *emcee* [59], *dynesty* [60], *pymultinest* [61, 62] etc.

Both MCMC and nested sampling are effective in estimating the posterior distribution of multi-

parameter ϑ; specially, although it takes more tedious efforts, nested sampling is better to estimate the evidence in the process of generating unbiased posterior samples from multi-dimensional parameter space as well as estimating the Bayes factor.

The next topic would be the parameter space of the waveform models. There are more than 15 parameters for BBH waveform model (11 parameters for spin-aligned BBH waveform model) and 17 parameters for BNS waveform model by including tidal effects:

$$\vartheta = (\mathcal{M}, q, a_1, a_2, \theta_1, \theta_2, \phi_1, \phi_2, \Lambda_1, \Lambda_2, \alpha, \delta, \psi, \iota, \phi_{\rm ref}, t_c, D_L), \tag{1.6}$$

Here $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ is the detector frame chirp mass, $q = m_2/m_1 (< 1)$ is the mass ratio, a_A are the dimensionless spin magnitudes while (θ_A, ϕ_A) are the polar and azimuthal angles of the spin angular momentum of the Ath body, Λ_A is the dimensionless tidal deformability of Ath neutron star (NS), (α, δ) are the sky location of the binary (right ascension and declination), ψ is the polarization angle of GWs with respect to the earth-centered coordinates, ι is the inclination angle of the binary's orbital angular momentum relative to the detector's line of sight, for aligned or nonspining system this ι coincides with the angle between the total angular momentum and the detector's line of sight, $\phi_{\rm ref}$ is the reference phase at the reference frequency, t_c is the coalescence time, and D_L is the luminosity distance.

One should include more parameters while considering other GR effects such as spin-induced quadrupole moment [63], interface (i) modes [5, 64] as excited at the interface of the fluid core and solid crust of NS, pressure-gravity (p-g) mode tidal instability of NS [65] and other possible beyond GR effects would also introduce new parameters to the waveform models. For instance, axionlike particles sourced by neutron stars would introduce two new parameters (relative strength γ_a and axion mass m_a) [6] and EdGB gravity would bring in coupling constant $\alpha_{\rm GB}$ [7, 66, 67].

Due to the large number of parameters (15 in total), it is computationally expensive to calculate the evidence by marginalizing the likelihood over all parameters. Subsection 1.2.6 will

demonstrate how to estimate evidence as well as Bayes factor. Followings are the other important sampling methods used in this research.

1.2.3 Generating Samples from Probability Distributions

The Inverse Transform Sampling

For any normalized probability distribution $f(x)$ with cumulative distribution function (CDF) $F(x) = p(f(x') \leq x) = \int_{-\infty}^{x} f(x')dx'$ $(0 \leq F(x) \leq 1)$. $F(x)$ must be a strictly increasing function since $f(x) \geq 0$. If the inverse F^{-1} exits, $F^{-1}(u)$ follows $f(x)$ distribution where u is a distribution on Uniform$[0, 1]$. A simple proof is as follows, if U is a random variable on Uniform$[0, 1]$, then $p(U < y) = y$ for any $y \in [0, 1]$. The cumulative distribution of $F^{-1}(U)$ is:

$$p(F^{-1}(U) \leq x) = p(U \leq F(x)) = F(x) = p(f(x') \leq x) \tag{1.7}$$

which means if we generate a large amount of samples u randomly from uniform distribution on $[0, 1]$, then $F^{-1}(u)$ follows probability distribution $f(x)$. For example, $f(x) = sin(x)/2$ in the range of $[0, \pi]$, the CDF and inverse function are $F(x) = (1 - cos(x))/2$ and $F^{-1}(x) = acos(1 - 2x)$. We could generate samples u randomly from uniform distribution on $[0, 1]$, then $F^{-1}(u) = acos(1-2u)$ has a probability distribution the same as $f(x) = sin(x)/2$.

Acceptance-Rejection sampling

When it is hard to find the inverse of CDF $F(x)$ and there is another general sampling method called Acceptance-Rejection sampling though sometimes it is not that efficient. Given any normalized probability distribution $f(x)$ that we want to generate samples from, we have another probability distribution $g(y)$ where we know how to generate samples. We first choose a number $N \geq \frac{f(x)}{g(x)}$ for any x. Then generate a random number Y from $g(y)$ and another random number U from uniform

distribution Uniform[0, 1]. If $U < \frac{f(Y)}{Ng(Y)}$, we accept Y as one of our samples by setting $X = Y$. Otherwise, go back and generate another new pair of Y and U. When we have a large amount of random samples X, the histogram of X should follow the probability distribution $f(x)$. Here I will provide a simple proof. Let's consider the CDF of X:

$$\begin{aligned} p(X \leq x) &= p\left(Y \leq x \,|\, U \leq \frac{f(Y)}{Ng(Y)}\right), \\ &= \frac{p\left(Y \leq x,\, U \leq \frac{f(Y)}{Ng(Y)}\right)}{p\left(U \leq \frac{f(Y)}{Ng(Y)}\right)}, \end{aligned} \tag{1.8}$$

Note that in the last equality, we used the definition of conditional probability. For any realization of random variable $Y = y$, the numerator:

$$\begin{aligned} p\left(Y \leq x,\, U \leq \frac{f(Y)}{Ng(Y)}\right) &= \int p\left(Y \leq x,\, U \leq \frac{f(Y)}{Ng(Y)} \,|\, Y = y\right) g(y)dy, \\ &= \int p\left(y \leq x,\, U \leq \frac{f(y)}{Ng(y)}\right) g(y)dy, \\ &= \int_{-\infty}^{x} p\left(U \leq \frac{f(y)}{Ng(y)}\right) g(y)dy, \\ &= \int_{-\infty}^{x} \frac{f(y)}{Ng(y)}\, g(y)dy, \\ &= \frac{1}{N}\int_{-\infty}^{x} f(y)dy, \end{aligned} \tag{1.9}$$

Similarly, the denominator:

$$\begin{aligned} p\left(U \leq \frac{f(Y)}{Ng(Y)}\right) &= \int p\left(U \leq \frac{f(Y)}{Ng(Y)} \,|\, Y = y\right) g(y)dy, \\ &= \int p\left(U \leq \frac{f(y)}{Ng(y)}\right) g(y)dy, \\ &= \int \frac{f(y)}{Ng(y)}\, g(y)dy, \\ &= \frac{1}{N}\int f(y)dy \;= \frac{1}{N}, \end{aligned} \tag{1.10}$$

Putting all altogether we obtain:

$$p(X \leq x) = \frac{p\left(Y \leq x,\ U \leq \frac{f(Y)}{Ng(Y)}\right)}{p\left(U \leq \frac{f(Y)}{Ng(Y)}\right)} = \int_{-\infty}^{x} f(y)dy\,, \tag{1.11}$$

which means that the random variable X does have the density $f(x)$. In practice, we want to choose N as small as possible because a small N leads to a higher chance of accepting Y . To see this, note that the denominator $p\left(U \leq \frac{f(Y)}{Ng(Y)}\right) = \frac{1}{N}$. Thus, a small N leads to a large accepting probability.

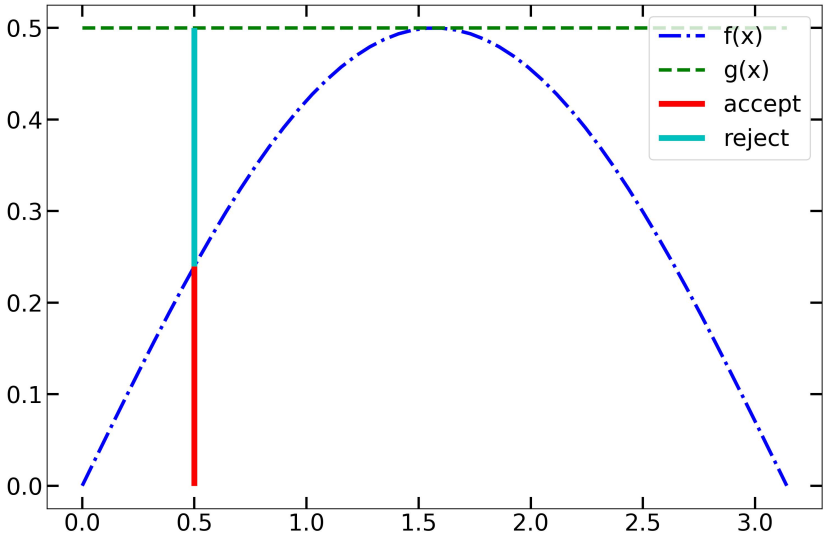

Figure 1.1: Illustration of reject sampling method from probability distribution $f(x) = sin(x)/2$. Note that $f(x)$ is normalized whereas $g(x)$ is not.

For example, probability distribution $f(x) = sin(x)/2$ in the range of $[0, \pi]$, the max value of $f(x)$ is 0.5. We naturally choose $g(x) = 0.5$ as a uniform distribution over range Uniform$[0, \pi]$ and $N = 1$. As shown in Fig. 1.1, the motivation is straightforward. Generating a random sample x from distribution $g(x)$ for example $x = 0.5$, then generating another random sample u from Uniform$[0, 1]$. If $u < f(0.5)/g(0.5) = 0.48$, we accept u as a sample from the distribution $f(x)$,

else we reject u. Repeat the process until we have a desired number of samples.

1.2.4 Detection of GW Events

One well-known technique to identify whether a GW event signal exists when a data stream is observed by LIGO is matched filtering: it searches for GW signals with known templates. With a given length of data stream, the filter has the ability to find the maximum signal-to-noise ratio (SNR) from a large bank of gravitational waveform templates generated with given waveform models. In other words, the waveform template corresponding to the maximum SNR has the best match for the observed data.

The **Optimal Network SNR** is defined as:

$$\left(\frac{S}{N}\right)_{\rm max} := \max_i \sum_k \frac{\langle h(\boldsymbol{\vartheta}_i) \,|\, d\rangle}{\sqrt{\langle h(\boldsymbol{\vartheta}_i) \,|\, h(\boldsymbol{\vartheta}_i)\rangle}}\,, \tag{1.12}$$

where $\langle ...\rangle$ is the inner product defined in Eq. (1.4). Index k refers to different detectors. d is the observed data (GW signal may or may not present), $h(\boldsymbol{\vartheta}_i)$ is the trail waveform template with parameters $\boldsymbol{\vartheta}_i$ from possible parameter space. The detection threshold depends on the cosmological model as well as the population distribution of the binary systems and a typical value of detection threshold is 8 [68, 69].

In Part III, we will estimate the detection rate of EMRIs from different formation channels. Since the LISA detector is yet to be launched (planned in early/mid 2030s), there is no observed EMRI data by now. Supposing $d = h(\boldsymbol{\vartheta}_i)$, then the SNR becomes:

$$\text{SNR} = \sqrt{\langle h(\boldsymbol{\vartheta}_i) \,|\, h(\boldsymbol{\vartheta}_i)\rangle} = \sqrt{4\int_{f_{\rm low}}^{f_{\rm high}} \frac{\tilde{h}^*(\boldsymbol{\vartheta}_i, f)\tilde{h}(\boldsymbol{\vartheta}_i, f)}{S_n(f)} df}\,, \tag{1.13}$$

where $S_{\rm n}(f)$ is the detector PSD of LISA.

It might be hard to determine the correct, or preferred, models when various waveform models

are suggested by different physical mechanisms. Frequentist and Bayesian perspectives provide several mode selection techniques, and the proper one adopted in this essay is the most popular Bayesian approach.

1.2.5 Bayesian Model Selection

To compare two models or hypotheses, we apply the Bayesian model selection method. For hypothesis $\mathcal{H}_1$ and $\mathcal{H}_2$ and observed data d, the **odds ratio** is defined as:

$$O_{12} := \frac{p(\mathcal{H}_1|d)}{p(\mathcal{H}_2|d)} = \frac{p(\mathcal{H}_1)}{p(\mathcal{H}_2)} \frac{p(d|\mathcal{H}_1)}{p(d|\mathcal{H}_2)} = \frac{p(\mathcal{H}_1)}{p(\mathcal{H}_2)} \mathcal{B}_{12} \tag{1.14}$$

where

$$\mathcal{B}_{12} := \frac{p(d|\mathcal{H}_1)}{P(d|\mathcal{H}_2)} \tag{1.15}$$

is **Bayes factor** between two different hypotheses $\mathcal{H}_1, \mathcal{H}_2$ and:

$$\begin{aligned} p(d|\mathcal{H}_1) &= \int d\vartheta \; p(d|\vartheta, \mathcal{H}_1) \; p(\vartheta|\mathcal{H}_1), \\ p(d|\mathcal{H}_2) &= \int d\vartheta \; p(d|\vartheta, \mathcal{H}_2) \; p(\vartheta|\mathcal{H}_2), \end{aligned} \tag{1.16}$$

Usually, we don't have any preference between hypotheses $\mathcal{H}_1$ and $\mathcal{H}_2$, assuming priors on $\mathcal{H}_1$ and $\mathcal{H}_2$ are equal $p(\mathcal{H}_1) = p(\mathcal{H}_2) = \frac{1}{2}$. Then we have $O_{12} = \mathcal{B}_{12}$.

According to the criteria in [70], larger Bayes factor $\mathcal{B}_{12}$ implies more preference of hypothesis $\mathcal{H}_1$ over hypothesis $\mathcal{H}_2$, and vice versa (see Table 1.1).

$\ln \mathcal{B}_{12}$	Strength of Evidence
-1.1 to 1.1	not prefer any
1.1 to 3	Positive
3 to 5	Strong
greater than 5	Very strong

Table 1.1: Criteria on Bayes Factor $\ln \mathcal{B}_{12}$. According to the justification in [70], if $-1.1 < \ln \mathcal{B}_{12} < 1.1$, the data does not prefer one model over the other; if $1.1 < \ln \mathcal{B}_{12} < 3$, there is positive support for hypothesis $\mathcal{H}_1$; if $3 < \ln \mathcal{B}_{12} < 5$, there is strong support for hypothesis $\mathcal{H}_1$ and if $\ln \mathcal{B}_{12} > 5$, the support is overwhelming.

1.2.6 Bayes Factor Estimation

As defined in the previous subsection 1.2.5, Bayes factor is the ratio of evidences from two hypotheses. While it is usually extremely hard to calculate the evidence analytically, there are multiple tools to estimate the Bayesian factor such as the thermodynamic integration method [71, 72], nested sampling method [57, 58, 60] and the Savage-Dickey Density Ratio (**SDDR**) method [73, 74].

Thermodynamic Integration Method

Let's first take a closer look at the thermodynamic integration method by redefining the unnormalized posterior $q(\vartheta|d)_\beta$ as a function of $\beta \in [0, 1]$:

$$q(\vartheta|d)_\beta := p(d|\vartheta)^\beta \; p(\vartheta)\,, \tag{1.17}$$

then we have the modified evidence:

$$Z(d)_\beta := \int d\vartheta \; q(\vartheta|d)_\beta\,, \tag{1.18}$$

when $\beta = 0$, we have $Z(d)_{\beta=0} = \int d\boldsymbol{\vartheta}\ p(d|\boldsymbol{\vartheta})^0\ p(\boldsymbol{\vartheta}) = \int d\boldsymbol{\vartheta}\ p(\boldsymbol{\vartheta}) = 1$ since the prior is normalized and we want to figure out $Z_{\beta=1}$ which is used to calculate the Bayes factor. Taking the derivative of $\ln Z_\beta$ with respect to β leads to:

$$\frac{\partial \ln Z_\beta}{\partial \beta} = \frac{1}{Z_\beta} \int d\boldsymbol{\vartheta}\ \frac{\partial q(\boldsymbol{\vartheta}|d)_\beta}{\partial \beta} = \int d\boldsymbol{\vartheta}\ \frac{\partial \ln q(\boldsymbol{\vartheta}|d)_\beta}{\partial \beta} \frac{q(\boldsymbol{\vartheta}|d)_\beta}{Z_\beta}\,, \tag{1.19}$$

since

$$\frac{\partial \ln q(\boldsymbol{\vartheta}|d)_\beta}{\partial \beta} = \frac{p(\boldsymbol{\vartheta})}{q(\boldsymbol{\vartheta}|d)_\beta} \frac{\partial p(d|\boldsymbol{\vartheta})^\beta}{\partial \beta} = \frac{p(\boldsymbol{\vartheta})\ p(d|\boldsymbol{\vartheta})^\beta\ \ln p(d|\boldsymbol{\vartheta})}{q(\boldsymbol{\vartheta}|d)_\beta} = \ln p(d|\boldsymbol{\vartheta})\,, \tag{1.20}$$

thus we have

$$\frac{\partial \ln Z_\beta}{\partial \beta} = \int d\boldsymbol{\vartheta}\ \ln p(d|\boldsymbol{\vartheta})\, \frac{q(\boldsymbol{\vartheta}|d)_\beta}{Z_\beta} = E_{\boldsymbol{\vartheta}}\left[\ln p(d|\boldsymbol{\vartheta})\right]_\beta\,, \tag{1.21}$$

where $E_{\boldsymbol{\vartheta}}[...]$ stands for the expectation value of log likelihood $\ln q(\boldsymbol{\vartheta}|d)_\beta$ with respect to normalized posterior $q(\boldsymbol{\vartheta}|d)_\beta \big/ Z_\beta$ at temperature $T = 1/\beta$ $(0 \le \beta \le 1)$. Taking the integration of $\partial \ln Z_\beta\ /\ \partial\beta$ respect to β from 0 to 1 leads to:

$$\ln Z_{\beta=1} - \ln Z_{\beta=0} = \int_0^1 \frac{\partial \ln Z_\beta}{\partial \beta} d\beta = \int_0^1 E_{\boldsymbol{\vartheta}}\left[\ln p(d|\boldsymbol{\vartheta})\right]\ |_\beta\ d\beta\,, \tag{1.22}$$

When we run MCMC samplings with a series of temperature T's, there are a large amount of posterior samples with log likelihood at different temperatures. The log evidence $\ln Z_{\beta=1}$ is then the numerical integration:

$$\ln Z_{\beta=1} = \int_0^1 E_{\boldsymbol{\vartheta}}\left[\ln p(d|\boldsymbol{\vartheta})\right]\ |_\beta\ d\beta\,, \tag{1.23}$$

Although thermodynamic integration method is a good way to evaluate the evidence, it is indeed computationally expensive since we have to run more than 10 MCMC's at different temperatures to get a lower error on the Bayes factor.

Nested Sampling Method

Different from the MCMC approach which tries to generate equal weighted samples directly from the posterior distribution, nested sampling estimates the evidence as well as generating weighted samples.

Let's rewrite the evidence defined in Eq. (1.2) as:

$$\mathcal{Z} := \int_{\Omega_{\vartheta}} \mathcal{L}(\vartheta)\, \pi(\vartheta) d\vartheta \,, \tag{1.24}$$

where $\mathcal{L}(\vartheta)$ is the likelihood function, $\pi(\vartheta)$ is the prior, and Ω_{ϑ} stands for the prior volume.

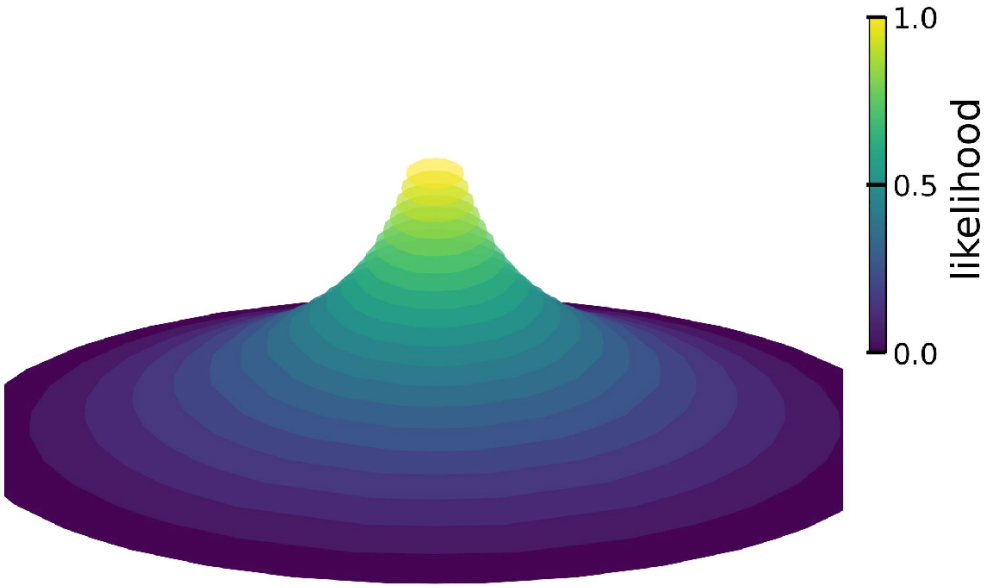

Figure 1.2: A graphical representation shows iso-likelihood contour with a 2D prior. The evidence is the sum of all contours weighted by the likelihood.

Since there are up to 15 parameters in GW waveform, it is indeed hard to perform the integration. Nested sampling provides a way by defining an **iso-likelihood contour**: as the likelihood increases, the prior volume decreases from 1 to 0. One could consider the integral as over the prior volume (one dimensional) instead of multi-parameters ϑ. The prior volume $X(\lambda)$ enclosed

by iso-likelihood contour with likelihood $\mathcal{L}(\vartheta) > \lambda$ defined as:

$$X(\lambda) := \int_{\Omega_\vartheta : \mathcal{L}(\vartheta) > \lambda} \pi(\vartheta) d\vartheta \,, \tag{1.25}$$

Within the scope of GW parameter estimation, the Gaussian likelihood implies $\mathcal{L}(\vartheta) \in (0, 1)$, as $\lambda \to 0$, $X(\lambda) \to 1$, and $\lambda \to 1$, $X(\lambda) \to 0$. As shown in Fig. 1.2, we could rewrite the evidence as the integral over the likelihood λ:

$$\mathcal{Z} = \int X(\lambda) d\lambda = \int \mathcal{L}(X)\, dX \,, \tag{1.26}$$

If the inverse of $X(\mathcal{L} = \lambda)$ exists, we could also integrate over the prior volume X instead of likelihood λ. When we run a nested sampling, as time goes on, the prior volume decreases from 1 to 0, and the log-likelihood increase from $-\infty$ to 0. The difficulty of nested sampling is to find the iso-likelihood contour and evaluate the prior volume.

The Savage-Dickey Density Ratio Method

The SDDR method directly estimates the Bayes Factor between two models instead of the evidence of each one. Consider two models $\mathcal{H}_0$ and $\mathcal{H}_1$, there are common parameters ϑ and one particular ϕ. ϕ is a free parameter in model $\mathcal{H}_1$ whereas in model $\mathcal{H}_0$, $\phi = \phi_0$ is fixed. We also call $\mathcal{H}_0$ a null hypothesis that is nested under hypothesis $\mathcal{H}_1$. In the nested models, as $\phi \to \phi_0$, model $\mathcal{H}_1$ reduces to $\mathcal{H}_0$, which is $p(\vartheta|\mathcal{H}_0) = p(\vartheta|\phi \to \phi_0, \mathcal{H}_1)$. The evidence under model $\mathcal{H}_0$ is:

$$\begin{aligned} p(d|\mathcal{H}_0) &= \int d\vartheta \; p(d|\vartheta, \mathcal{H}_0) \; p(\vartheta|\mathcal{H}_0) \,, \\ &= \int d\vartheta \; p(d|\vartheta, \phi = \phi_0, \mathcal{H}_1) \; p(\vartheta, \phi = \phi_0|\mathcal{H}_1) \,, \\ &= p(d|\phi = \phi_0, \mathcal{H}_1) \,, \end{aligned} \tag{1.27}$$

By adopting Bayes' theorem:

$$p(d|\mathcal{H}_0) = p(d|\phi = \phi_0, \mathcal{H}_1) = \frac{p(\phi = \phi_0|d, \mathcal{H}_1)\, p(d|\mathcal{H}_1)}{p(\phi = \phi_0|\mathcal{H}_1)}\,, \tag{1.28}$$

Then the Bayes factor is:

$$\begin{aligned} \mathcal{B}_{01} &= \frac{p(d|\mathcal{H}_0)}{P(d|\mathcal{H}_1)} = \frac{p(d|\phi = \phi_0, \mathcal{H}_1)}{P(d|\mathcal{H}_1)}\,, \\ &= \frac{p(\phi = \phi_0|d, \mathcal{H}_1)}{p(\phi = \phi_0|\mathcal{H}_1)}\,, \end{aligned} \tag{1.29}$$

which is the ratio between the posterior and prior at $\phi = \phi_0$ under model $\mathcal{H}_1$. As an example illustrated in Fig. 1.3, if model $\mathcal{H}_1$ degenerates to $\mathcal{H}_0$ when $\phi = 0$, the Bayes factor is:

$$\mathcal{B}_{01} = \frac{p(d|\mathcal{H}_0)}{P(d|\mathcal{H}_1)} = \frac{p(\phi = 0|d, \mathcal{H}_1)}{p(\phi = 0|\mathcal{H}_1)} = 0.01/0.033 \approx 0.3 \tag{1.30}$$

In other words, the Bayes factor $\ln \mathcal{B}_{10} = \ln(1/\mathcal{B}_{01}) = 1.20$ is greater than 1.1. From criteria shown in Table 1.1, there is a positive support for model $\mathcal{H}_1$ than the null hypothesis $\mathcal{H}_0$. As we can see in Fig. 1.3, if $\phi = 0$ is far from the peak of the posterior, the error on $p(\phi = 0|d, \mathcal{H}_1)$ will largely influence the estimated Bayes factor, which will result in an unreliable estimation of Bayes factor.

We will perform the SDDR method to estimate Bayes factor in our project [5], when $\delta\phi \to 0$, the hypothesis with mode resonance reduces to the one without. As well as in project [7], when $\sqrt{\alpha_{\rm GB}} \to 0$, EdGB gravity reduces to GR.

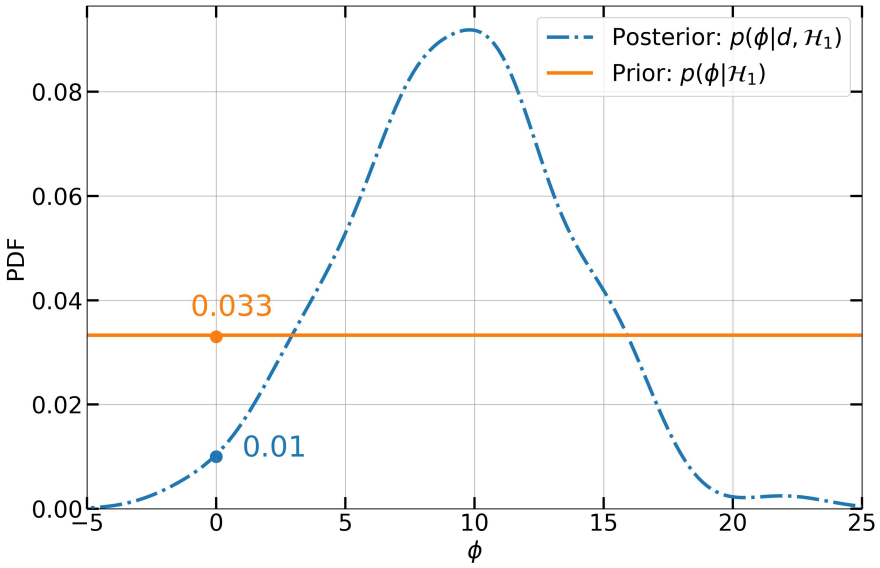

Figure 1.3: An example shows the prior and posterior on parameter ϕ. The Bayes factor is the ratio between the posterior and prior at $\phi = 0$

1.2.7 Fisher Analysis

We next explain another method for parameter estimation, namely the Fisher Information Matrix (**FIM**) [75–78], which is valid when the SNR is large and the noise is stationary and Gaussian. We begin by expanding the log-likelihood function at the maximum likelihood point $\boldsymbol{\vartheta}^{\rm ML}$ for a given hypothesis $\mathcal{H}$:

$$\begin{aligned}\log p(d|\boldsymbol{\vartheta}, \mathcal{H}) &\propto -\frac{1}{2}\sum_k \langle d_k - h_k(\boldsymbol{\vartheta})|d_k - h_k(\boldsymbol{\vartheta})\rangle\,, \\ &\propto -\frac{1}{2}\sum_k \Gamma^{(k)}_{ij}\,\Delta\boldsymbol{\vartheta}^i\,\Delta\boldsymbol{\vartheta}^j\,, \end{aligned} \tag{1.31}$$

where $\Delta\boldsymbol{\vartheta}^i = \boldsymbol{\vartheta}^{i,\mathrm{ML}} - \boldsymbol{\vartheta}^i$ is the error of a given parameter relative to the value at maximum likelihood point and $\Gamma_{ij}^{(k)}$ is the FIM evaluated at the maximum likelihood point $\boldsymbol{\vartheta}^{\mathrm{ML}}$:

$$\Gamma_{ij}^{(k)} = \left\langle \frac{\partial h(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}^i} \,\middle|\, \frac{\partial h(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}^j} \right\rangle \Big|_{\boldsymbol{\vartheta}^{\mathrm{ML}}}, \qquad \Gamma_{ij} = \sum_k \Gamma_{ij}^{(k)}, \tag{1.32}$$

where the inner product is given in Eq. (1.4) with the power spectral density $S_n^{(k)}$ for the kth detector. Notice that the elements of FIM are partial derivatives of the waveform template with respect to given parameters. Similar to the Bayseian inference, one can introduce a prior to find the posterior distribution on $\boldsymbol{\vartheta}$. The inverse of the FIM is an estimator of the error covariance matrix Σ_{ij}. The standard error is the square root of the diagonal elements of the covariance matrix. For a given parameter $\boldsymbol{\vartheta}^i$, the standard error can be expressed as:

$$\sqrt{\langle (\delta\boldsymbol{\vartheta}^i)^2 \rangle} = \sqrt{\Sigma_{ii}}, \quad \Sigma_{ij} = \left(\tilde{\Gamma}^{-1}\right)_{ij}. \tag{1.33}$$

We implement the Fisher analysis method to estimate $\sqrt{\alpha_{\mathrm{GB}}}$ in our EdGB project. See more details in Chapter 4.

1.3 Gravitational Waveform Models

GR is a theory of gravitation who has achieved a great success in not only explaining various astronomical phenomena but also predicting new phenomena such as GWs which have been direct observed by LIGO in 2015 [18]. In order to search GW signal in observed data, we need to construct waveform models (solve Einstein field equation) which could generate waveform templates with given parameters as mentioned in subsection 1.2.4. Since Einstein field equation $G_{\mu\nu} = 8\pi T_{\mu\nu}$ is a 4-dimentional second order differential tensor equation and the highly non-linearity makes it extremely hard to solve. Numerical relativity is a way to solve the Einstein equation directly, however, it is really computational expensive and can only solve limited parameter space in which

the equation is stably evolved. One way to speed up the generation of waveforms is adopting PN approximation by sacrificing the accuracy, while the approximated waveform is still accurate enough at the inspiral stage for current parameter estimation purpose (requires millions of waveform templates with various parameters). Since the field is strong at merger and ringdown stages[1], the PN waveforms will have a large deviation from the "real" GR waveforms. In this case, one can employ inspiral-merger-ringdown waveform models (EOB or Phenom formalism) to realize a better accuracy than PN approximation.

The waveform is naturally expressed in the frequency domain, which allows fast waveform evaluation. This is because matched filtering implemented in the frequency domain is computationally less expensive, whereas the time domain waveforms usually requires solving the equation of motion down to orbital timescales.

Throughout the research, we will focus on three waveform models. **TaylorF2** [80], a commonly used frequency domain model which gives an accurate approximation to the inspiral stage for large separation and slow velocities. **IMRPhenomXPHM** [81], a phenomenological frequency domain model including precessing effects and subdominant harmonic modes. **SEOBNRv4T** [8], a time domain EOB model including the effect of dynamical tides. IMRPhenomXPHM and SEOBNRv4T are both full waveform models including inspiral-merger-ringdown stage by combining PN theory and NR simulations [82]. Above three models have already been well implemented in the LALSimulation package [83].

We will go through the post-Newtonian (PN) formalism, then two families of full waveform models (EOB and phenomenological), and, at last, the parameterized post-Einsteinian (ppE) formalism.

[1]The binary coalescence usually splits into three stages: **inspiral stage**, when the field strength is comparable to Newtonian case and relative velocity is much small than the speed of light; **merger stage**, when the separation of the binary is around the Innermost Stable Circular Orbit (ISCO) ($r_{\rm ISCO} = 6M$) of Schwarzschild BHs with the same total mass, the corresponding frequency is called transition frequency; **ringdown stage**, the transition frequency from merger to ringdown is called ringdown frequency, see paper [79].

1.3.1 PN Approximants

At the early stage of compact binary coalescence, the separation is large with respect to their size (week field) and relative velocity of the objects is much smaller the the speed of light (slow motion):

$$\frac{GM}{c^2 R_c} \sim \frac{v_c^2}{c^2} \ll 1 \tag{1.34}$$

where M is the total mass of the system, R_c is the characteristic separation of the binary, v_c is the characteristic orbital velocity and c is the speed of light. In the domain of PN theory (week field and slow motion), it is safe to expand the Einstein field equations in powers of dimensionless velocity $v = (GM\Omega/c^3)^{1/3} = (\pi M f)^{1/3}$, here the orbital frequency Ω relates to the GW frequency f through $\Omega = \pi f$ and use $G = c = 1$ convention in the last equality. A nth-PN correction usually refers the expansion to v^{2n} and the zeroth-PN order is just the Newtonian theory without any corrections. We can include higher order corrections until the desired accuracy is achieved. To learn more on PN theory, please read the excellent textbook *Gravity: Newtonian, Post-Newtonian, Relativistic* by Poisson and Will [84].

Orbits of the binary system shrink as a result of energy loss (GW radiation) and orbital frequencies increase at the same time. In the adiabatic approximation (orbital frequency changes slowly) and for circular orbits, we expect the energy balance equation $P = -dE/dt$ holds where P and E are energy and energy flux of the system, respectively.

$$P = -dE/dt = -E'(v)\frac{dv}{dt}\,, \tag{1.35}$$

which is

$$\frac{P}{E'(v)} + \frac{dv}{dt} = 0\,, \tag{1.36}$$

Together with Kepler's third law $\Omega^2 r^3 = GM$, GW phase $\phi = 2\pi f t$ and the definition of dimen-

sionless characteristic velocity $v = (GM\Omega)^{1/3}$:

$$\frac{1}{2}\frac{d\phi}{dt} - \frac{v^3}{M} = 0\,, \tag{1.37}$$

where 1/2 comes from the fact that orbital phase is one half of GW phase. Together with differential Eqs. (1.36) and (1.37) (Equivalently the integral form shown in Eq. (1.40)), one could solve the evolution of any parameters such as v and ϕ as a function of t or f. Though the observed GW data is in time domain, it is indeed more efficient to work in frequency when we need to compute the inner product of time series defined in Eq.(1.4).

Let's consider a common used frequency domain waveform model TaylorF2. Within the stationary phase approximation (SPA) [76, 80], one can analytically integrate Eqs. (1.36) and (1.37) with respect to GW frequency f. The waveform in the Fourier space is then given by:

$$h(f) = \mathcal{A}(f)\exp\left[i\,\Psi(f)\right]\,, \tag{1.38}$$

Here $\mathcal{A}(f) \propto \mathcal{M}^{5/6} f^{-7/6} Q(\iota,\psi)/D_L$ is the amplitude. The phase $\Psi(f)$ can be calculated by:

$$\Psi(f) = 2\pi f\, t(f) - \phi(f) - \frac{\pi}{4}, \tag{1.39}$$

with

$$t(f) = t_c - \int_{f_c}^{f} \frac{1}{P}\left(\frac{dE}{df'}\right) df'\,, \qquad \phi(f) = \phi_c - \int_{f_c}^{f} \frac{2\pi f'}{P}\left(\frac{dE}{df'}\right) df', \tag{1.40}$$

where E and P are the binding energy and energy flux of the binary system respectively. Note that E and P are functions of r and Ω, which are related by the Kepler's law $r(\Omega)$, and the GW frequency f relates to the orbital frequency Ω through $\Omega = \pi f$. After integration, the phase upto

kth-PN order is given by [80]:

$$\Psi(f) = 2\pi f\, t_c - \phi_{ref} - \frac{\pi}{4} + \frac{3}{128\,\eta v^5} \sum_{i \le k} \alpha_{2i}\, v^{2i}\,, \tag{1.41}$$

where $v = (\pi M f)^{1/3}$ and $M = m_1 + m_2$ is the total mass of the system. t_c is the coalescence time and ϕ_{ref} is the reference phase at the reference frequency. $\eta = m_1 m_2 / M^2$ is the symmetric mass ratio. α_i are known functions of M and η.

For example, let's consider 0PN order binding energy E and energy flux P:

$$E = -\frac{1}{2} M \eta v^2 \qquad F = -\frac{32}{5} M \eta^2 v^{10}\,, \tag{1.42}$$

After integrating Eq. (1.40), one finds the 0PN order coefficient $\alpha_0 = 1$. One could employ higher PN orders of binding energy and energy flux to compute the coefficients of GW phase until the desired accuracy has been achieved. The state of art TaylorF2 has implemented up to 3.5PN order corrections in LALSimulation package.

If there are other GR effects as mentioned in section 1.2.2, which will introduce an extra phase corrections. When we search for i-mode resonance signature [5] from BNS event GW170817, we directly add an extra phase correction $\delta\Psi(f)$ to the TaylorF2 phase:

$$\Psi(f)_{\rm Tot} = \Psi(f)_{\rm TaylorF2} + \delta\Psi(f)_{\rm i-mode} \tag{1.43}$$

See section 1.5.1 and Chapter 2 for more details.

1.3.2 Full Waveform Models (EOB and Phenomenological)

PN approximants are only valid at the inspiral stage under the weak field and slow motion assumptions. As the separation of binary shrinks, the field strength and relative velocity increase gradually,

which leads PN approximants more and more inaccurate. However, the merger and ringdown frequency band is exactly observable by current ground based detectors. Therefore, for data analysis purpose we much find another way to achieve a better approximation to NR simulations at the merger-ringdown stage (it is still computational expensive to solve field equation directly), such as EOB and Phenomenological, which are two families of full inspiral-merger-ringdown waveform models by combining the PN theory at inspiral region, and the BH perturbation theory or EOB approach at merger-ringdown region, and calibrated with NR simulations.

EOB approach is trying to solve the dynamical evolution of compact binary system by mapping two-body problem to an effective-one-body problem which is a test particle moving in an effective external metric [85]. It is a time domain model calibrated with NR simulations by adopting a variety of analytical approximations. In Chapter 5, we will compute the tidal phase including dynamical tides with SEOBNRv4T model [8] which have already included the effect of dynamical tides.

Phenomenological (IMRPhenom) model is another approach to deal with the merger-ringdown stage in frequency domain with the BH perturbation theory, as well calibrated with NR simulations. We adopt state-of-the-art IMRPhenomXPHM [81] model in Chapter 4 with appropriate modification (see next subsection), since the model incorporates both precessing effects and subdominant harmonic modes.

1.3.3 Modification to GR Waveform Models

There may be other beyond GR effects as mentioned in section 1.2.2. One import question is how to construct the corresponding waveform models? Yunes and Pretorius [86] proposed the parameterized post-Einsteinian (ppE) formalism for modifying waveform models. It is a way to include beyond GR effects to existing well constructed frequency-domain GR models according to the leading PN order phase corrections predicted by beyond GR theories. See papers [86–88] for

more details.

In working on projects searching nuclear coupling of axions [6] and constraining on EdGB gravity [7], we incorporate the ppE formalism via including phase corrections in the inspiral regime to the corresponding waveforms, which is similar to Eq. (1.43). See sections 1.5.2 and 1.5.3 for more details.

1.4 Tidal Interaction from Compact Binary Objects

If at least one of the compact objects in the binary is a neutron star, tide-induced neutrons star deformation has to be included into the binary dynamics. This effect was first computed in [89] for the leading order term in the waveform, with higher order PN corrections worked out in [90]. Later on these PN tidal corrections were incorporated in the EOB framework, for both the equilibrium tide [91–93] and the dynamic tide [8].

In the adiabatic limit, the star's induced quadrupole moment Q_{ab} is directly proportional to the tidal gravitational field E_{ab}, with the proportionality constant given by the tidal Love number λ:

$$Q_{ab} = -\lambda E_{ab} \,, \tag{1.44}$$

In the equilibrium tide approximation, λ is assumed to be a constant; with dynamical tide included, λ can be thought as a function of the orbital frequency. Additional subtlety comes in if the orbital evolution cross one or more mode resonances, where residual free mode oscillations will be present after these resonances and Eq. (1.44) breaks down [94]. For the purpose of study in Chapter 5, since the primary mode (f-mode) generally has frequency higher than the inspiral frequency, we will assume that the adiabatic approximation holds in the entire inspiral frequency range.

In addition to adiabatic tides, tidal interaction can excite internal modes of neutron stars as the binary sweeps through the inspiral frequency range. The pressure (p-) and fundamental (f-)modes

[95] will not be fully excited as their frequencies are generally higher than the inspiral frequency, although it has been suggested that early excitation of f-modes may be observed in the late inspiral stage [96]. Gravity modes may be fully excited, but their couplings to tidal gravitational fields are so small that the induced phase shifts are $O(10^{-3})$ or smaller [97, 98]. Resonance of rotational modes has also been investigated assuming a rotational frequency of a few $\times 10^2$ Hz [99–102], whereas the fastest rotating pulsar known in a binary neutron star system has a frequency of ~ 60 Hz [103, 104].

The interface (i-)modes [64], excited at the interface of the fluid core and solid crust, have frequencies around several tens to a few hundred Hertz, depending on the star's equation of state and prescription of the crust. The resonance of i-modes was proposed to explain precursors of short gamma-ray bursts due to possible crust failures [105]. We observe that through excitation of i-modes, the crustal material actually reaches its elastic limit well before the mode resonance. After reaching this threshold the crust undergoes an elastic-to-plastic transition and the tidal driving starts to heat up the crust. The whole process ends with the meltdown of the crust in tens of cycles. We search such an i-mode with GW event GW170817 in Chapter 2.

1.5 Research on Astrophysics and Fundamental Physics with Inspiraling Compact Bianries

This section talks about three projects based on the detected GW events: GW170817 [106], GW200105 and GW200115 [107]. The first one is on tidal induced interface (i) mode of NS, we will try to search for such a i-mode signal in observed BNS events GW170817 (Chapter 2); the second project is to search for axionkile particles from GW170817 (Chapter 3) and the third is about constraints on EdGB gravity from BHNS events GW200105 and GW200115 (Chapter 4).

The intrinsic information of inspiraling compact binaries and the surrounding environment will

influence the orbital motion hence change both the phase and amplitude of waveforms. However, when we search for possible signals (i-mode, axion and EdGB) from GW events, we only consider the inspiral stage of the waveform and for any possible modification to the waveform model, we just add an extra phase change to the total phase since matched-filtering techniques are particularly sensitive to the phase of GW signals [108]. Though the GW amplitude will indeed change in some sense, the estimation uncertainty on luminosity distance D_L is much larger than the changes due to other effects (i-mode, axion and EdGB). Therefore we neglect any corrections to amplitude of the waveform model.

We adopt the frequency domain waveform model TaylorF2 in Chapter 2 and Chapter 3, waveform model IMRPhenomXPHM in Chapter 4. By employing the SPA, the TaylorF2 model is expressed in PN expansion, hence it is only valid at the late-inspiral stage $f_{\rm high} = 0.018/M$ [109].

1.5.1 Probing Crust Meltdown in Inspiraling Event GW170817 (Chapter 2)

Neutron stars have been one of the most mysterious objects in the universe since they were first hypothesized nearly a century ago. Recent measurements of star deformability and radius enable a better understanding of their stellar structure. However, these measurements are only sensitive to the core, i.e., the most dense part of the star. In this work, we report an eigenmode excited on the interface of the core and crust during the inspiraling stage of a neutron star binary. This mode oscillation causes a elastic-to-plastic transition of the crust, heat up and finally melt the crust, thereby imprinting a detectable signature on the gravitational waveform. Once detected, it will be a unique probe of the crust structure of neutron stars.

After the melting process, part of the binary orbital energy is converted to the mode and thermal energy resulting in a phase shift of the gravitational waveform. Similar to the discussion in [97, 98, 110] for mode resonances, for the binary neutron star waveform $h(f) = A(f)e^{i\Psi(f)}$, its phase is

modified as

$$\delta\Psi(f) = \sum_{i=1,2} \delta\phi_i \left(1 - \frac{f}{f_i}\right) \Theta(f - f_i)$$
$$\approx \delta\phi_a \left(1 - \frac{f}{f_a}\right) \Theta(f - f_a) \tag{1.45}$$

where Θ is the Heaviside function and f_i is the melting frequency of each star. The melting process decreases the coalescence phase by $\delta\phi_i$ and the coalescence time by $\delta\phi_i/2\pi f_i$. In the second line we introduced $\delta\phi_a = \sum_i \delta\phi_i$ and $\delta\phi_a/f_a = \sum_i \delta\phi_i/f_i$ to reduce the number of extra parameters in this model, which simplifies the parameter estimation process. Notice that if energy transfers from the orbit to the mode (or heat in this case) during resonance, $\delta\phi$ is positive; if energy transfers from the mode to the orbit, as expected in some of the r-mode resonances [101], $\delta\phi$ is negative.

As discussed in section 1.3 and 1.2.2, we plug the induced phase $\delta\Psi(f)$ to pure GR waveform model TaylorF2 and perform MCMC simulation to search the crust melting signature. Though we don't find the this signal in GW170817, possibly due to the SNR is relatively small, we expect to observe the signal with GW detector A+ and more likely with third-generation GW detectors. Furthermore, by estimating Bayes factor (see section 1.2.6 for more details) between different possible signatures, we find it is possible to distinguish the i-mode from other possible origins, such as tidal-p-g mode [111] and axion filed around NSs [6].

See more details on this work in Chapter 2.

1.5.2 Constraints on Nuclear Coupling of Axions from Event GW170817 (Chapter 3)

In this project, we report on a new search for certain axions using GW170817, the GWs from a BNS inspiral detected by LIGO and Virgo [106][2]. We focus on axions that couple to nuclear matter in

[2] Although the scenario of GW170817 being a neutron star-black hole merger can not be ruled out, the astrophysical processes to produce a black hole with neutron-star mass are generally considered to be exotic [112].

the same way as the QCD axion, but with masses that are relatively light [113, 114]. Such axions can be sourced by NSs and affect the dynamics of binary NS coalescence, leaving potentially detectable fingerprints in the inspiral waveform [115, 116]. Given the fact that the observed GWs do not deviate significantly from GR, we expect anything induced by the scalar axion must be small. We calculate the axion induced phase shift by taking the effect as a perturbation to GR. First, Let's find out the modified Kepler's law $r(\Omega)$. The modified Lagrangian is $\mathcal{L} = \mathcal{L}_{\rm GR} - V_a$, where V_a is the leading correction to the GR binding energy [116] ($c = 1$):

$$V_a = -8 q_a M \eta \, \frac{e^{-m_a r}}{r} \,, \tag{1.46}$$

After plugging the modified Lagrangian in Euler equation for circular orbit we have:

$$(GM\Omega)^2 = \left(\frac{GM}{r}\right)^3 \left[1 + 8 q_a (1 + m_a r) e^{-m_a r}\right] \,, \tag{1.47}$$

Then the deviation of orbital frequency and separation between the two NSs from GR are:

$$\delta\Omega^2 = \frac{GM}{r^3} \, 8 q_a (1 + m_a r) \, e^{-m_a r} \,, \qquad \delta r = \frac{8}{3} q_a r (1 + m_a r) \, e^{-m_a r} \,, \tag{1.48}$$

Here Ω is the orbital frequency, r denotes the dimensionless separation (devide by GM) between the two NSs of masses M_1 and M_2, and $M = M_1 + M_2$ is the total mass. $q_a = q_1 q_2/(\eta M^2)$ is dimensionless axion charge, where q_1 and q_2 are axion charge for the two NSs. m_a is the axion mass.

By taking the modified Kepler's law, the axion induced leading order corrections to the GR binding energy and radiation power are:

$$\delta E = \frac{\partial E_{\rm GR}}{\partial r} \delta r + V_a \,, \qquad \delta P = \frac{\partial P_{\rm GR}}{\partial r} \delta r + P_a \,, \tag{1.49}$$

where P_a is the axion dipole radiation [116]. Then Eq. (1.40) becomes:

$$\delta t(f) = \int_{f_c}^{f} \left(\frac{\delta P}{P_{\rm GR}^2} \frac{dE_{\rm GR}}{df'} - \frac{1}{P_{\rm GR}} \frac{d\delta E}{df'} \right) df' \,,$$
$$\delta \phi(f) = \int_{f_c}^{f} 2\pi f' \left(\frac{\delta P}{P_{\rm GR}^2} \frac{dE_{\rm GR}}{df'} - \frac{1}{P_{\rm GR}} \frac{d\delta E}{df'} \right) df' \,, \tag{1.50}$$

After integrating above equations, the axion induced phase $\Psi(f)_a = 2\pi f \delta t - \delta\phi$ is shown in Appendix 3.A.

To search for such axions, we plug axion phase $\Psi_a(f)$ into TaylorF2 waveform model (see section 1.3) and perform Bayesian analysis (see section 1.2.2) of GW170817 taking into account the possible dephasing caused by the axions. The posterior distribution over the waveform parameters suggest no significant evidence for such axion fields. As shown in Fig. 3.1, this null result excludes a large region of the axion parameter space, much of which has not been probed by existing experiments. Importantly, our constraints are independent of the assumption that axions are the dark matter, which is required for the constraint from big bang nucleosynthesis (BBN) [117].

See more details on this work in Chapter 3.

1.5.3 Constraints on EdGB gravity from BH-NS events (Chapter 4)

Scalar Gauss-Bonnet gravity theory is described by a scalar field ϕ non-minimally coupled to the Gauss-Bonnet invariant $\mathcal{R}_{\rm GB} = R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho} - 4R_{\mu\nu}R^{\mu\nu} + R^2$ [118, 119]. The corresponding action is:

$$S = \int d^4x \sqrt{-g} \left[\frac{R}{16\pi} - \frac{1}{2}(\nabla\phi)^2 + \alpha_{\rm GB} f(\phi) \mathcal{R}^2_{\rm GB} \right] + S_m \,. \tag{1.51}$$

where g is the determinant for the metric $g_{\mu\nu}$. R is the Ricci scalar on the 4-dimensional manifold with metric $g_{\mu\nu}$. $\alpha_{\rm GB}$ is the coupling parameter and $f(\phi)$ is an arbitrary function of the scalar field that determines how it is coupled to the metric. S_m donates the matter action.

For compact binary inspirals within the parameterized post-Einsteinian (ppE) framework, the

parameter $\alpha_{\rm GB}$ is usually introduced to describe the non-Einsteinian effects [87, 88, 120]. We have derived the higher order phase corrections upto 2PN order by maping the result from Jordan frame to Einstein frame, see Appendix 4.A. To construct a modified GW waveform model, we introduce these phase correction as a perturbation to the existing GR waveform model IMRPhenomXPHM [67, 81, 121].

In EdGB gravity, dimensionless BH charges $s_i = 2(\sqrt{1-\chi_i^2} - 1 + \chi_i^2)/\chi_i^2$, where $\chi_i = \vec{S}_i \cdot \hat{L}/m_i^2$ are dimensionless spins of BHs. $s_i = 0$ for ordinary stars like NS. As reported in LVC papers [107, 122], the $(l, m) = (3, 3)$ mode is non negligible. Therefore we include $(3, 3)$ mode in all of our MCMC analyses except for event GW151226.

Since we only derive the EdGB effect upto post-inspiral stage and the events we employed are inspiral stage dominant, we only consider inspiral waveform modification by setting a high frequency cutoff $f_{\rm high} = 0.018/M$ [109] (for the total mass M in a unit of second). The modified waveform model IMRPhenomXPHM doesn't include tidal effect though it includes both higher order modes and precession effects. Since no waveform model includes all of the three effects currently, and tidal phase is O(1) which is relatively small compared to EdGB phase correction O(10), we are safe to perform MCMC simulation with waveform model IMRPhenomXPHM.

In order to constrain $\sqrt{\alpha_{\rm GB}}$, we first perform Fisher analysis (in section 1.2.7) to determine a rough estimation and then perform Bayesian inference (in section 1.2.2) to find the posterior distributions on $\sqrt{\alpha_{\rm GB}}$ with GW events GW151226, GW190814(BBH), GW190814(NSBH), GW200105 and GW200115. The constraints on $\sqrt{\alpha_{\rm GB}}$ are shown on Table 4.1 and Fig. 4.1. Constraint on GW151226 is consistent with Perkins' work on Fig. 6 [67], which also means our work is reliable. We give a stronger constraint $\sqrt{\alpha_{\rm GB}} \lesssim 1.33\,\text{km}$ by individual event GW200115 rather than stacking individual events from GWTC-1 and GWTC-2 catalogs which gives $\sqrt{\alpha_{\rm GB}} \lesssim 1.7\,\text{km}$.

Further more, we adopt the stacking technique introduced in paper [67] by combining marginalized posteriors from individual events GW200105, GW200115 and GW190814(NSBH) and Perkin's combined posterior together to give a more stringent constraint on $\sqrt{\alpha_{\rm GB}}$. The combined constraints

are $\sqrt{\alpha_{\rm GB}} \lesssim 0.59\,\mathrm{km}$ by Fisher analysis and $\sqrt{\alpha_{\rm GB}} \lesssim 1.18\,\mathrm{km}$ by Bayesian inference.

Besides constraints from leading order -1PN phase correction, we perform MCMC simulation from higher PN order phase corrections with GW events GW200105 and GW200115, and show the 90% upper bounds are improved by 14.5% from GW200105 and 6.9% from GW200115 respectively, A posterior distribution comparison between -1PN phase correction and higher PN phase corrections are shown in Fig. 4.2.

In the future, we would happy to see if there are more higher mass ratio events with higher SNR, which will strongly improve the constraint on $\sqrt{\alpha_{\rm GB}}$. Or in the other way, we could stack all individual events together, if there are N statistical independent events with the similar SNR, which could improve the constraint by a factor of $\sqrt{N}$ in principle.

See more details on this work in Chapter 4.

1.6 The Hybrid Waveform for inspiraling Compact Binaries (Chapter 5)

As mentioned in section 1.4, tidal interaction could introduce an extra phase shift to GW phase. The tide-induced phase shift can be expanded in the velocity $v = \sqrt{M/r}$ (M is the total mass and r is the orbital separation) within the PN formalism:

$$\psi_{\rm PN} = \lambda v^5 (\psi_{0\rm PN} + v^2 \psi_{1\rm PN} + v^3 \psi_{1.5\rm PN} + v^4 \psi_{2\rm PN} + ...)\,. \tag{1.52}$$

The leading PN-order coefficient $\psi_{0\rm PN}$ is:

$$\delta\psi = -\frac{9}{16}\frac{v^5}{\mu M^4}\left(\frac{11 m_2}{m_1} + \frac{M}{m_1}\right) \tag{1.53}$$

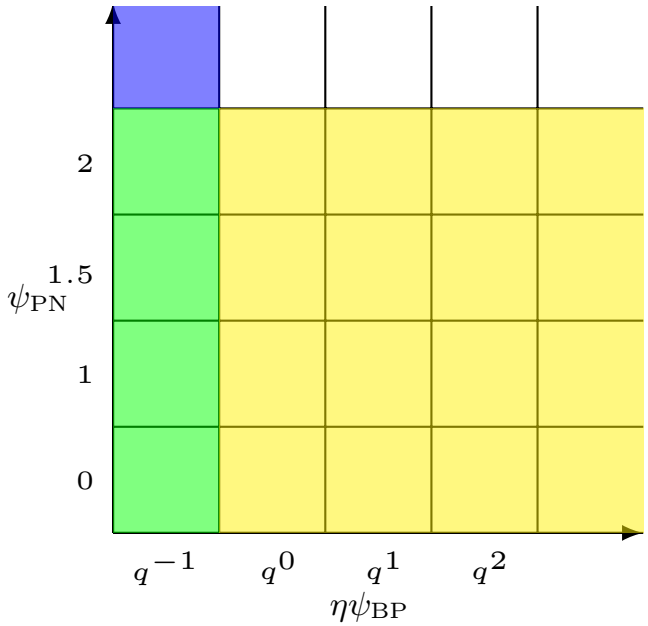

Figure 1.4: $\psi_{\rm BP}$ and $\psi_{\rm PN}$ approximately obtained in truncated expansions in the Black Hole Perturbation Theory and Post-Newtonian Formalism. The green and blue parts denote $\psi_{\rm BP}$, the green and yellow parts denote $\psi_{\rm PN}$, and then the overlap of $\psi_{\rm BP}$ and $\psi_{\rm PN}$ is green part.

with μ being the reduced mass $\mu = m_1 m_2/M = \eta M$ and M is the total mass $M = m_1 + m_2$. This motivates us to write down the tide-induced phase shift contributed by the less massive star (star "1") as

$$\eta\psi_{\rm BP,1} \approx \lambda_1 (q^{-1}\psi_{\rm BP}^{(-1)} + \psi_{\rm BP}^{(0)} + \sum_{n\geq 1} \eta^n \psi_{\rm BP}^{(n)}) \,. \tag{1.54}$$

which naturally includes all PN corrections, with the subscript "BP" denoting "Black Hole Perturbation". In particular, the $\psi_{\rm BP}^{(-1)}$ term can be obtained considering the tidal deformation of the neutron star due to the background Kerr spacetime of the primary black hole, and $\psi_{\rm BP}^{(0)}$ corresponds to the extra tidal deformation induced by h_1.

Theoretically speaking, after summing over all PN terms in Eq. (1.52) and all mass ratio terms in Eq. (1.54), $\psi_{\rm BP}$ and $\psi_{\rm PN}$ should agree. In practice, $\psi_{\rm BP}$ and $\psi_{\rm PN}$ are approximately obtained

in truncated expansions in the Black Hole Perturbation Theory and Post-Newtonian Formalism respectively, as illustrated in Fig. 1.4. In order to better capture the tidal effect with these two independent expansions, we propose to construct a hybrid waveform by using

$$\psi_{\rm hyd} = \psi_{\rm PN} + \psi_{\rm BP} - \psi_{\rm ovp}\,, \tag{1.55}$$

where $\psi_{\rm ovp}$ denotes the contribution from the overlap regime of the Post-Newtonian and Black Hole Perturbation methods (the green regime in Fig. 1.4). As a result, the difference between this hybrid waveform and the "true" waveform come from the blank space in Fig. 1.4. As the expansion orders in Post-Newtonian and Black Hole Perturbation methods increase, the blank space shrinks and we shall obtain a better approximated waveform. Notice that this construction applies not only to BNS and BHNS binaries discussed here, but binary black holes BBHs as well. It will be interesting to perform the exercise combining the EMRI-inspired waveform with the PN waveform for BBHs, and compare with other resumed waveforms such as the EOB templates[123–127].

In this work we truncate the series with only $\psi_{\rm BP}^{(-1)}$ in $\psi_{\rm BP}$ and up to $\psi_{\rm 2PN}$ in $\psi_{\rm PN}$. The accuracy of the resulting hybrid waveform is comparable to the state of the art EOB waveform for the tidal correction, for the numerical waveforms that we have used for comparison. The systematic error is understood as the blank space in the phase diagram as in Fig. 1.4. The waveform model is also easily extendible when higher order correction terms in $\psi_{\rm BP}$ and $\psi_{\rm PN}$ are available. We plan to update the hybrid waveform with $\psi_{\rm BP}^{(0)}$ in the future, and possibly with inspiral-to-plunge corrections and higher multipoles if necessary.

1.6.1 Tide-induced Phase Comparison

Numerical waveforms downloaded from SXS Gravitational Waveform Database [128] are all in time domain and are expressed as $h(t) = h_+(t) - ih_{\rm x}(t)$ with two polarizations $h_+(t)$ and $h_{\rm x}(t)$. In order to compare tidal induced phase shift in frequency domain, we should take a time-frequency

domain transformation. We define the GW phase and frequency as:

$$\phi(t) = -arg(h(t))\,, \qquad f(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt}\,, \tag{1.56}$$

Combining above two equations we could find $\phi(f)$ as of function of GW frequency. The tidal induced phase is then $\delta\psi(f) = \phi_{\rm BHNS}(f) - \phi_{\rm BBH}(f)$. As mentioned in paper [129], one could also define ϕ as an argument of the time derivative of $h(t)$. We have checked they are almost the same and doesn't influence our comparisons as shown in Fig. 5.2 and Fig. 5.3, which indicate the hybrid method has a comparable accuracy as the EOB waveform and numerical relativity simulation in characterizing the tidal effects and the hybrid waveform is more accurate if mass ratio is small. We also compare the spin influence to the tidal phase in Fig. 5.4 which shows the discrepancy is less than 10% even at the end of the inspiral stage.

See more details on this work in Chapter 5.

1.7 Formation and Rate Study of EMRIs in AGN (Chapter 6 and 7)

One of the important sources for spaceborne GW detectors is EMRIs [130] and previous studies have focus on the "loss cone" channel (also called "dry EMRIs"), where stellar-mass black holes (sBHs) are gravitationally captured by a MBH, following multi-body scatterings within the nuclear cluster [131–139]. As previously advanced by Zhen Pan and Huan Yang [140], there is another comparable formation channel in AGN driven by accretion disk (called "wet EMRIs"). In this project we will compare these two main channels of EMRI formation by estimating LISA detectable rate respectively and we forecast that wet EMRI channel is evidently an important or even dominant channel for spaceborne GW detectors.

In the dry channel, a sBH is scattered by stars in the nuclear stellar cluster and gravitation-

ally captued by the MBH. There are two relevant timescales in the long-term dynamics: the GW emission timescale $t_{\rm gw}$ on which the sBH orbit shrink, and the diffusion timescale t_J on which the orbital angular momentum changes by successive scatterings [131, 134, 141–143]. If GW emission is more efficient with $t_{\rm gw} < t_J$, the sBH becomes a stable EMRI with continually shrinking orbit until the final coalescence with the MBH. If scatterings are more efficient, i.e., $t_J < t_{\rm gw}$, the sBH is randomly scattered towards or away from the MBH. SBHs that are scattered into the MBH without losing much energy via GW emission are called prompt infalls [131]. As the GW emission timescale is much shorter for eccentric orbiters ($e \to 1$), EMRIs in the dry channel are highly eccentric at formation.

A fraction $\mathcal{O}(10^{-2}-10^{-1})$ of MBHs in the universe (referred as AGNs) are actively accreting gas with an accretion disk [144, 145]. The presence of an accretion disk introduces new interactions affecting the motion of sBHs in the stellar cluster. For a sBH embedded in the AGN disk, its periodic motion produces density waves [146–149] that in turn drive the sBH to migrate inward, damp its orbital eccentricity and its inclination w.r.t. the disk plane. For a sBH on a highly inclined orbit, the effects of density waves becomes subdominant to dynamical friction [150, 151] arising from the relative motion of the sBH and the surrounding gas as it passes through the disk. As a result, sBHs are first captured onto the disk driven by dynamical fraction and density waves, and then migrate inward driven by the density waves upon reaching the vicinity of the MBH, where GW emission become dominant. We refer this type of EMRIs as wet EMRIs. Because density waves are very efficient at damping out the eccentricity, wet EMRIs are essentially circular. Dry and wet EMRIs can be easily distinguished from each other by measuring their eccentricities using spaceborne gravitational wave detectors [137–139]. In addition, the imprints of an accretion disk on the EMRI waveform may also be detectable [152, 153].

In order to compare these two channels. We first construct kinds of models by choose reasonable functions and parameters from each model, for examples, we consider two redshift-independent MBH mass functions as shown in Eq. (6.24). Then we generate random samples from all models

(see section 6.5 for more details). For each EMRIs, we generate its time-domain waveform $h_{+,\times}(t)$ using the Augment Analytic Kludge (AAK) [154–156] with the conservative Schwarzschild plunge condition, then calculate the SNR using Eq. (1.13) with LISA sensitivity. As shown in Fig. 6.2 and Table 6.1, the expected LISA detectable EMRI rates (SNR$\geq$ 20) from wet channel is evidently an important or even dominant channel for all the models we have considered.

Besides EMRIs with a sBH, we also studied EMRIs with a lower mass gap object (MGO) ($\sim (2.5-5)M_\odot$). In the dry channel, sBHs are generally closer to the central MBH than MGOs due to the mass segregation, so that their EMRI rate should be larger. In the wet channel, the EMRI formation rates of both sBHs and MGOs mainly depend on the capture rate onto the AGN disk and the rate of migration along the disk. Both rates are proportional to the object mass, therefore the EMRI rate of MGOs is lower than that of sBHs by a factor of their mass ratio at most. We find the wet EMRI rate per AGN is generally higher than the dry EMRI rate per MBH by $O(10^1-10^3)$ for sBHs, and by $O(10^3-10^4)$ for MGOs. Taking into account the AGN fraction $f_{\rm AGN}=O(10^{-2}-10^{-1})$, the wet channel turns out to be primary way in producing mass-gap EMRIs. If there is no gap in the mass spectrum of supernova remnants and roughly equal number of MGOs and sBHs are born in SN explosions, we expect LISA to detect $\lesssim 1$ dry mass-gap EMRIs per year, and $O(1-10^2)\times(f_{\rm AGN}/1\%)$ wet mass-gap EMRIs per year. In addition, future detections of dry and wet mass-gap EMRIs have interesting implications of MGO formation. An excess of mass-gap EMRI detection could be a signature of MGOs of more exotic origin, e.g., PBHs. The relative fraction of mass-gap EMRIs to sBH EMRIs is a sensitive probe for the mass spectrum of SN remnants and therefore the SN explosion mechanism.

See more details on this work in Chapter 6 and Chapter 7.

Part I

Gravitational Wave Data Analysis with Inspiraling Compact Bianries

Chapter 2

Probing Crust Meltdown in Inspiraling Binary Neutron Stars

Zhen Pan, Zhenwei Lyu, Béatrice Bonga, Néstor Ortiz, Huan Yang

Phys. Rev. Lett. 125, 201102 (Nov. 2020), arXiv:2003.03330

My specific contributions: Modify TaylorF2 waveform model by adding i-mode phase $\delta\Psi(f)$ discussed in section 2.4 to the total phase and perform MCMC simulation to search such a specific signal in GW event GW170817 in section 2.5. After searching i-mode in observed data, I generate GW170817-like mock data with a i-mode at $f_a = 60$Hz and $\delta\phi = 0.3$, then search the i-mode with A+ and third-generation detectors. I also try to search tidal p-g mode and scalar filed using the same mock data, then calculate the corresponding Bayes factors in appendix 2.B.

Thanks to recent measurements of tidal deformability and radius, the nuclear equation of state and structure of neutron stars are now better understood. Here, we show that through resonant tidal excitations in a binary inspiral, the neutron crust generically undergoes elastic-to-plastic transition, which leads to crust heating and eventually meltdown. This process could induce $\sim O(0.1)$ phase

shift in the gravitational waveform. Detecting the timing and induced phase shift of this crust meltdown will shed light on the crust structure, such as the core-crust transition density, which previous measurements are insensitive to. A direct search using GW170817 data has not found this signal, possibly due to limited signal-to-noise ratio. We predict that such signal may be observable with Advanced LIGO Plus and more likely with third-generation gravitational-wave detectors such as the Einstein Telescope and Cosmic Explorer.

2.1 Introduction

Inspiraling neutron stars deform under mutual tidal interactions. In the adiabatic limit, the star's induced quadrupole moment is directly proportional to the tidal gravitational field, with the proportionality constant given by the tidal Love number. Deformed neutron stars orbit each other differently from black holes with the same masses, and the phase difference can be used to measure the tidal Love number [157], as shown in the analysis of GW170817 [158]. Together with neutron star radius measurements [159], maximum mass estimates [160] and possibly post-merger electromagnetic signals [161], the star's equation of state (EoS) is now better constrained.

As discussed in sections 1.4 and 1.5.1, we will search the tidal resonance mode from GW170817 in this project.

2.2 Crust heating up and melting down

The outer part of the crust is commonly described by a Coulomb lattice with shear modulus μ [162]. The inner crust may have nonuniform structures associated with the "nuclear pasta" phase [163, 164], which is not considered in this study. Simulations of molecular dynamics [165] have shown that the lattice responds elastically under small applied stress; once the induced strain exceeds the breaking strain ($\epsilon_b \sim 0.1$), plastic deformation starts to develop. Assuming an applied stress σ,

the plastic deformation rate $\dot{\epsilon}_{\rm pl}$ is exponentially small if $\sigma < \sigma_b = \mu\epsilon_b$, and becomes exponentially large if $\sigma > \sigma_b$. Mathematically, it is well approximated by [165]

$$\dot{\epsilon}_{\rm pl} = \frac{n_i Z^2 e^2}{a} \frac{\omega_{\rm p}}{\mu \bar{N} \Gamma} e^{(-18.5\bar{\sigma}_b + \bar{\sigma}\bar{N})\Gamma} , \tag{2.1}$$

where the dot denotes a time derivative, $\omega_{\rm p}$ is the plasma frequency, $\bar{N} = 500/(\Gamma - 149) + 18.5$, $\bar{\sigma} = \sigma/(n_i Z^2 e^2/a)$ and $\Gamma = Z^2 e^2/aT$ is the melting parameter with e the electron charge, Ze the total charge per ion, a the lattice spacing, n_i the ion density and T the temperature. The elastic part of the strain $\epsilon_{\rm el}$ satisfies $\sigma = \mu\epsilon_{\rm el}$ and the total strain is simply $\epsilon = \epsilon_{\rm el} + \epsilon_{\rm pl}$.

With the plastic deformation, mode energy dissipates into thermal energy, heating up the crust with a rate [166]

$$n_i \dot{e}_i = \sigma \dot{\epsilon}_{\rm pl}(\sigma, T) , \tag{2.2}$$

where n_i is the ion number density, e_i is the thermal energy per ion, and $de_i = c_V dT$ with c_V the specific heat capacity for $T < T_{\rm melt}$ [167]. Once the melting temperature $T_{\rm melt}$ is reached, the crustal material still needs an extra amount of latent heat ($\sim kT_{\rm melt}$ per ion) to be melted [168]. As a result, the total energy per ion needed to melt the crust from its initial cold state is roughly $e_{\rm melt} = \int_0^{T_{\rm melt}} c_V dT + kT_{\rm melt}$. In this work we have ignored contributions from dripped neutrons as their specific heat may be suppressed by superfluidity.

2.3 Mode Analysis

In the linear approximation, the stellar response to the tidal force is specified by the Lagrangian displacement $\boldsymbol{\xi}(\boldsymbol{r}, t)$ of a fluid element from its equilibrium position. The displacement can be decomposed into eigenmodes, $\boldsymbol{\xi}(\boldsymbol{r}, t) = \sum_\alpha a_\alpha(t)\boldsymbol{\xi}_\alpha(\boldsymbol{r})$, where α denotes the quantum number of an eigenmode. In the context of this paper, we only consider i-modes driven by the leading quadrupole term of the tidal force, so that $\boldsymbol{\xi}_m(\boldsymbol{r}) = [U(r)\hat{r} + rV(r)\nabla]Y_{2m}(\theta, \phi)$, where $Y_{2m}(\theta, \phi)$ is the $l = 2$

spherical harmonic. The displacement behavior is governed by the linear pulsation equation [64]

$$[\mathcal{L}(r;\mu) - \omega_0^2]\boldsymbol{\xi}_m = 0\,, \tag{2.3}$$

with $\mathcal{L}$ being an operator specifying the restoring force inside the star (see 2.A for the explicit expression).

For the example star with $M_\star = 1.3M_\odot, R_\star = 11.7$ km assuming SLy4 EoS [169, 170] and a core-crust baryon transition density $n_{\rm b,cc} = 0.065$ fm^{-3}, we obtain an i-mode frequency $f_0 = \omega_0/2\pi = 190$ Hz [64, 105] and the tidal coupling coefficient (a measure quantifying the overlap between the waveform and the tidal field)

$$Q = \frac{1}{M_\star R_\star^2}\int d^3x\rho\ \boldsymbol{\xi}_m^* \cdot \nabla[r^2 Y_{2m}(\theta,\phi)] = 0.018, \tag{2.4}$$

with the normalization $\langle\boldsymbol{\xi}_m|\boldsymbol{\xi}_{m'}\rangle := \int d^3x\rho\ \boldsymbol{\xi}_m \cdot \boldsymbol{\xi}_{m'}^* = \delta_{mm'}M_\star R_\star^2$, where ρ is the mass density [1].

The evolution of the mode amplitude $a_m(t)$ is governed by [97]

$$\ddot{a}_m + \gamma(t)\dot{a}_m + \omega_0^2(t)a_m = \frac{GM'W_{2m}Q}{D^3}e^{-im\Phi(t)}\,, \tag{2.5}$$

where the right-hand side is the leading quadrupole term of the tidal driving force with $M' = qM_\star$ the companion star mass, D the binary seperation, $\Phi(t)$ the orbital phase and W_{2m} is a coefficient of $O(1)$ (see Eq. (2.4) in Ref. [97]). On the left-hand side, $\gamma(t)\dot{a}_m$ is a damping term capturing the plastic deformation induced dissipation with $\gamma(t)$ defined as the ratio between the mode energy dissipation rate and two times the mode kinetic energy, i.e.,

$$\gamma(t) = \frac{\int_{\rm crust} n_i\dot{e}_i\ d^3x}{M_\star R_\star^2\ \sum_m |\dot{a}_m|^2}\,, \tag{2.6}$$

[1] In Ref. [105], a factor $\sqrt{l(l+1)}$ was missed in the normalization calculation.

where the numerator is the crust heating rate (which is equal to the mode energy dissipation rate), and the mode kinetic energy is $\frac{1}{2}\int d^3x\rho\dot{\boldsymbol{\xi}}(\boldsymbol{r},t)\cdot\dot{\boldsymbol{\xi}}^*(\boldsymbol{r},t) = \frac{1}{2}M_\star R_\star^2\sum_m|\dot{a}_m|^2$. The mode frequency $\omega_0(t)$ to leading order is determined by (see Eq. (2.3))

$$\omega_0^2(t) = \frac{\langle\boldsymbol{\xi}_m|\mathcal{L}(r;\mu_{\rm avg})\boldsymbol{\xi}_m\rangle}{\langle\boldsymbol{\xi}_m|\boldsymbol{\xi}_m\rangle}\,, \tag{2.7}$$

where $\mu_{\rm avg}$ is the average shear modulus which decreases as the crust is heated and we find the mode frequency is roughly proportional to the square root of the average shear modulus [171].

Given the mode amplitude $a_m(t)$, it is straightforward to calculate the fluid element displacement $\boldsymbol{\xi}(\boldsymbol{r},t) = \sum_m a_m(t)\boldsymbol{\xi}_m(\boldsymbol{r})$ and the corresponding strain $\epsilon_{\rm el}$. From equation (2.1), the plastic deformation rate $\dot{\epsilon}_{\rm pl}$ has an exponential dependence on the local strain $\epsilon_{\rm el}$ for $\epsilon_{\rm el}\gtrsim 0.1$, so does the energy dissipation rate $\sigma\dot{\epsilon}_{\rm pl}$. Physically, the dissipated energy comes from the local elastic energy, therefore the energy dissipation rate cannot exceed its replenishment rate $\frac{\mathcal{A}}{2}\mu\epsilon_{\rm el}^2 f_{\rm GW}$, where $f_{\rm GW}$ is the frequency of both the tidal force and the GW emission and $\mathcal{A}$ is a coefficient of $O(1)$. Here we take $\mathcal{A}=2$ as an example. As for the initial condition, we choose $T_i = 0.02T_{\rm melt}$, where $T_{\rm melt}\sim 1$ MeV is the melting temperature of the ion crystal at the crust base [162]. Using the 4th-order Runge-Kutta scheme, we evolve Equations (2.1, 2.2, 2.5) on the two-dimensional surface of the crust base, i.e., we only trace the thermal evolution of the crust base considering its dominant role in the crust heat capacity.

As the neutron star binary spirals inward, the tidal field increases and so does the i-mode amplitude $a_{m=0,\pm2}$, as shown in Fig. 2.1. At a certain binary separation (with corresponding gravitational wave frequency $f_{\rm GW,melt} < f_0$), part of the crust reaches the yield limit ϵ_b due to the i-mode excitation and plastic deformation starts. Heating first takes place at the equator where the strain maximizes. As the crust heats up, it softens so that i-mode frequency f_0 decreases and the mode amplitude a_m increases. As a result, the crust yields on larger and larger areas, extending from the equator to the poles, and finally the whole crust is melted. The crust melting takes about 20

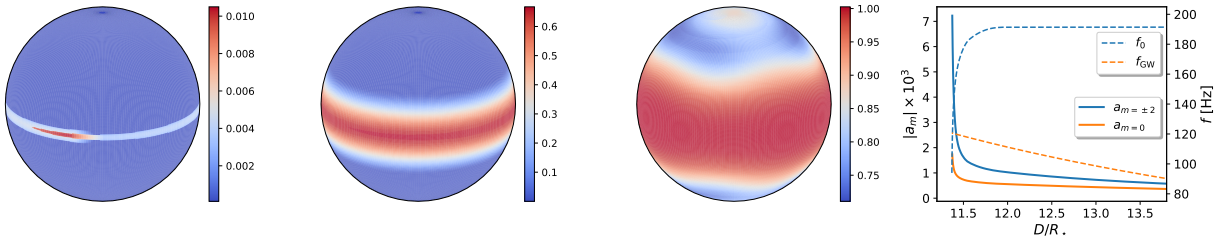

Figure 2.1: Left three panels are the heat maps $e_i/e_{\rm melt}$ of the neutron star crust (within a $1.3M_\odot$ + $1.3M_\odot$ binary) at binary separations $D = 12.0/11.6/11.4\ R_\star$, respectively. In the rightmost panel, dashed lines denote the evolution of the i-mode frequency f_0 and the gravitational wave (GW) frequency $f_{\rm GW}$, and solid lines denote the evolution of mode amplitude a_m with $m = \pm 2, 0$.

orbit periods and a total amount of energy $E_{\rm melt} \simeq 1.1 \times 10^{47}$ ergs. Notice that this mode treatment is approximate once the plastic motion turns on, where a more accurate description requires 3-dimensional dynamical modeling of crustal motions. A 2-dimensional consistent evolution was implemented in [166] to reveal yield patterns of magnetar crust under strong magnetic stress.

2.4 Waveform signature

As discussed in section 1.5.1, the search and forecast presented below for crust melting apply equally for generic mode resonances, and we will use 'mode resonance signature' and 'crust melting signature' interchangeably.

For each neutron star, $\delta\phi$ depends on its mass $M_\star$, the mass ratio of the companion q (with the companion mass being $qM_\star$), the melting energy $E_{\rm melt}$ and the melting frequency $f_{\rm GW,melt}$ as follows [97]

$$\delta\phi = \frac{2\omega_{\rm orb}E_{\rm melt}}{P_{\rm GW}} \simeq \frac{0.1}{q^2}\left(\frac{1+q}{2}\right)^{2/3} E_{47}M_{1.3}^{-10/3}f_{70}^{-7/3}\,, \tag{2.8}$$

where $\omega_{\rm orb} = \pi f_{\rm GW,melt}$ is the orbital angular frequency, $P_{\rm GW}$ is the energy loss rate due to GW emission, and $E_{47} = E_{\rm melt}/10^{47}$ergs, $M_{1.3} = M_\star/1.3M_\odot$, $f_{70} = f_{\rm GW,melt}/70$Hz. From Equation (2.8), we immediately see that the phase shift increases if the melting process happens earlier (lower

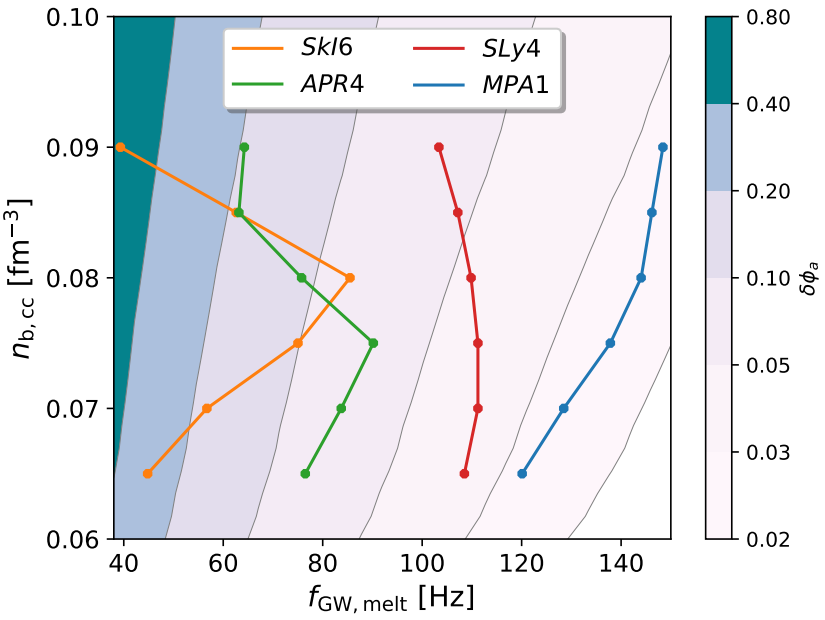

Figure 2.2: The crust melting induced phase change $\delta\phi_a$ in GWs of a BNS merger with each star of $M_\star = 1.3M_\odot$ and $R_\star = 12.5/11.3/11.7/12.7$ km for the EoS SkI6/APR4/SLy4/MPA1, respectively, which are not ruled out by the LIGO tidal measurement with GW170817. Note that the core-crust transition density $n_{\mathrm{b,cc}}$ is subject to a large uncertainty in each EoS instead of being an accurately predicted value, so we take the transition density as a free parameter.

$f_{\mathrm{GW,melt}}$) in the inspiral phase. In Fig. 2.2, we show the total phase change $\delta\phi_a$ for an equal-mass binary neutron star (BNS) merger with $M_\star = 1.3M_\odot$, where $\delta\phi_a$ varies from 0.03 to 0.6 depending on the star's EoS and the core-crust transition baryon density $n_{\mathrm{b,cc}}$. The melting energy increases substantially with increasing $n_{\mathrm{b,cc}}$ (commonly assumed to be within $0.06-0.1$ fm^{-3} [172–174]), whereas the i-mode frequency and the associated melting frequency are non-monotonic functions of $n_{\mathrm{b,cc}}$. We also note that since the mode calculation presented here is Newtonian with the Cowling approximation [175], the fully relativistic mode frequencies may be different (for examples, the frequencies of p- and f-modes are smaller with the metric perturbation included [176, 177]). If there are also more unpaired neutrons present within the star, as suggested by the cooling measurement in [178], the melting energy may be significantly boosted and the internal

mode spectrum may be modified as well. Therefore the search of mode resonance signatures may also help probe the superfluid composition of neutron stars. The effects of nuclear pastas on the melting energy budget and the mode frequency determination also need to be better understood. Nevertheless, the measurement of f_a and $\delta\phi_a$ will convey useful information about the core-crust transition density and the star's EoS around that density.

2.5 Search with GW170817

We now present the first search for mode resonance effects (including crust melting) in binary neutron star systems with data from GW170817 with Equation (1.45) implemented. A similar search for tidal-p-g instability is discussed in [111] using different $\delta\Psi(f)$. The Markov-Chain Monte Carlo (MCMC) parameter estimation is performed with PyCBC [179], for which we assume the source distance and sky location are known as the electromagnetic counterpart of this source has been identified [31]. We use the TaylorF2 waveform [180] as the background binary neutron waveform. We present the posterior distributions of chirp mass, mass ratio, $\delta\phi_a$ and f_a in Fig. 2.3. The marginal distribution of ϕ_a indicates that there is no evidence for mode resonance in GW170817, as $\delta\phi_a < 1.5$ at 95% confidence level. A similar conclusion can be drawn from a Bayesian model comparison framework. We denote $\mathcal{H}_a$ as the hypothesis with mode resonance and $\mathcal{H}_0$ as the one without, the Bayes factor can be defined as

$$\mathcal{B}^a_0 = \frac{P(\text{GW170817}|\mathcal{H}_a)}{P(\text{GW170817}|\mathcal{H}_0)} \tag{2.9}$$

which measures the relative probability of these two hypotheses. We have computed the Bayes factor using both the method of thermodynamic integration [71] and the Savage-Dickey Density Ratio method [73], which both suggest consistent values of $\log \mathcal{B}^a_0$ in the range of $[-0.6, -0.3]$. This means that these two hypotheses are essentially indistinguishable with this set of gravitational

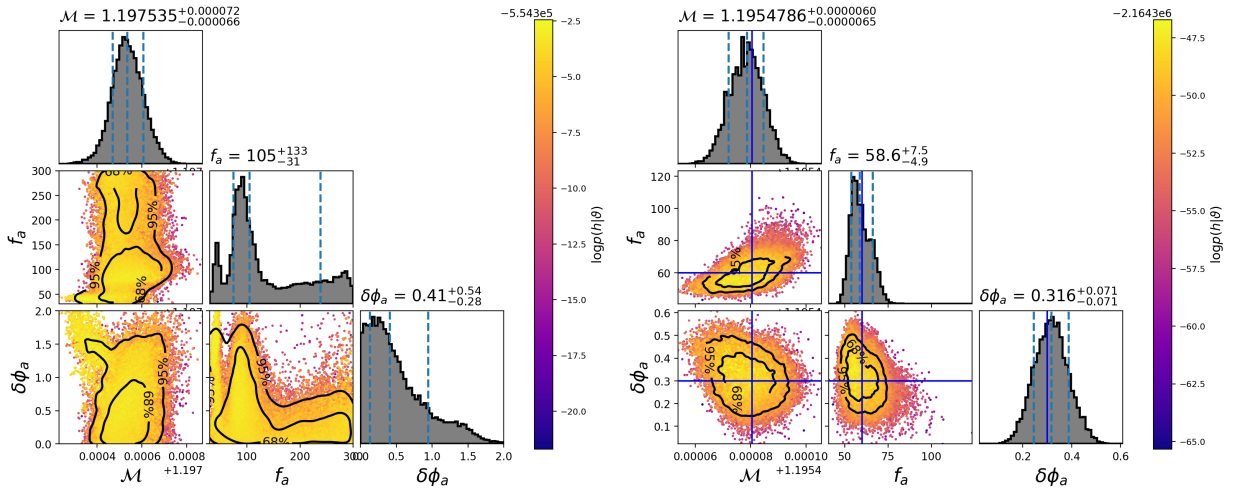

Figure 2.3: Posterior distribution of chirp mass $\mathcal{M}$, mass ratio q, phase shift $\delta\phi_a$ and melting frequency f_a obtained with PyCBC, where the prior for f_a is set to be $[30, 300]$Hz and $[0, 2]$ for $\delta\phi_a$. Left Plot: the search using data from GW170817. Right Plot: a search obtained assuming LIGO A+ sensitivity and an mode resonance injection at $f_a = 60$Hz and $\delta\phi_a = 0.3$.

wave data [70].

It is natural to expect observations with higher signal-to-noise ratios as the sensitivity of gravitational wave detectors improves. In the mid-2020s the upgrade of Advanced LIGO, LIGO A+, is expected to start its constrcution [2]. Assuming LIGO A+ design sensitivity for both detectors at Hanford and Livingston, and Advanced Virgo with its full sensitivity, we may observe GW170817-like events with signal-to-noise ratios beyond 100. In the right panel of Fig. 2.3, we present a sample search with an injected signal with $\delta\phi_a = 0.3$, $f_a = 60$Hz (for a GW170817-type system) into simulated detector noises consistent with the aforementioned LIGO A+ network sensitivity. A MCMC analysis of the mock data successfully recovers the injected values of f_a and ϕ_a with small uncertainties. So it is possible that we observe the crust melting signature in gravitational waves with LIGO A+.

Stacking different events may also improve detectability, as is the case for subdominant modes

[2] https://dcc.ligo.org/LIGO-G1601435/public

in black hole ringdowns [181]. However, we have no prior information on $\delta\phi_a$ and f_a, which are distinct for each binary neutron star system. If we have an underlying or phenomenological model that predicts or characterizes $\delta\phi_a$, f_a as a function of star mass, core-crust transition density and star compactness (which depends on the EoS), the hyper-parameters in this model may be constrained from different events. Certainly the posterior distribution of the hyper-parameters from different events can be multiple together to form the joint probability distribution. This is something worth to pursue in future studies.

If a mode resonance signature is indeed detected (i.e. preferred over the null hypothesis), it is still necessary to compare to other possible origins, such as tidal-p-g coupling [111], dynamical scalarization and vectorization [182, 183], scalar modes associated to certain GR extensions [184] and extensions to standard particle physics [185], that predict different $\delta\Psi(f)$. Since the crust melting is nearly instant (Fig. 2.1), its impact on the waveform boils down to shifting the coalescence time and the coalescence phase, i.e., $\delta\Psi(f) = \delta\phi_a \times (1 - f/f_a)$ with $\delta\phi_a$ being a constant. For other processes with continuous orbital energy draining, e.g., the tidal-p-g coupling extending the whole frequency range once the nonlinear instability is turned on, the waveform signature can be formulated in a similar way except with a frequency dependent phase shift $\delta\phi_a(f)$ which encodes the details of orbital energy draining. To simulate this, we inject a mode resonance signal ($\delta\phi_a = 0.3, f_a = 60$Hz) into detector noise corresponding to the LIGO A+ network, and perform the Bayesian model selection between our model resonance waveform and the tidal-p-g waveform. We find a Bayes factor $\log \mathcal{B}^a_{p-g} = 2.7 \pm 0.3$, suggesting that it is also possible to determine the correct model if a positive detection occurs (see 2.B for more details of the Bayesian analysis). The comparison will be much sharper with third-generation gravitational wave detectors. Similarly for the scalarized neutron stars proposed in scalar tensor theories or other particle physics considerations, there are also effects, such as dipole scalar radiation, that will be effective during the whole frequency range once turned on [185]. We also perform a model selection between the mode resonance and an example model of BNSs with scalar dipole radiation using the same mock

data above, and we find the Bayes factor is $\log \mathcal{B}^{a}_{\text{dipole}} = 11.7 \pm 2.2$ (see 2.B).

2.6 Conclusions and Discussion

Resonant tidal excitations in a neutron star binary induce a phase shift $\delta\phi_a$ in the gravitational wave signal by melting its crust. In calculating the crust heating rate, we have used the fitting formula $\dot{\epsilon}_{\text{pl}}(\sigma)$ [Eq. (2.1)] which is a result of molecular dynamics simulations [165]. If this simulations result does not accurately apply to the NS crust with a breaking strain ϵ_b different from 0.1, the crust melting frequency $f_{\text{GW,melt}}$ will also change. For a smaller breaking strain $\epsilon_b = 0.05$, we find the melting frequency $f_{\text{GW,melt}}$ decreases by $\sim 25\%$ and the phase shift $\delta\phi_a$ increases by a factor ~ 2. All the predicted phase shifts corresponding to different EoSs are still well consistent with the constraint $\delta\phi_a < 1.5$ (95% confidence level) from GW170817. LIGO A+ may already be able to detect such induced phase shifts. A 3rd-generation detector network with Cosmic Explorer [16] sensitivity at the LIGO detectors and Einstein Telescope [15] sensitivity at the Virgo detector is able to limit $\delta\phi_a$ with uncertainty ~ 0.01 and f_a below 1%. This will not only allow high-confidence detection of the crust melting effect, but also precisely measure crustal and EoS properties as shown in Fig. 2.2.

We do not expect significant energy release to the neutron star magnetosphere associated with crustal failure, as the magnetic fields ($\sim 10^{12}$G) assumed are too weak to efficiently transfer energy by sending out Alfvén waves. However, if the star is a magnetar with field $\sim 10^{15}$ G, this emission mechanism may excite star magnetospheres and power precursor gamma-ray bursts [166, 186, 187]. Interestingly, the recent LIGO observation of a heavy neutron-star binary (GW190425 [188]) may indicate the existence of a fast-merging channel to form binary neutron stars. Such systems may have short-enough lifetime $\sim 10^4$ years to allow active magnetars in the binary coalescence stage [189].

2.A Appendix: Pulsation Equations

The motion of a mass element inside a star is governed by the continuity equation, the momentum equation and the Possion equation

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) &= 0 \,, \\ \frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} &= \frac{1}{\rho}\nabla \cdot \boldsymbol{\sigma} - \nabla\Phi \,, \\ \nabla^2\Phi &= 4\pi G\rho \,, \end{aligned} \tag{2.10}$$

where $\boldsymbol{\sigma}$ is the stress tensor. In the equilibrium state where $\boldsymbol{v}^{(0)} = 0$, the stress tensor is simply $\sigma_{ij}^{(0)} = -p\delta_{ij}$ with p being the pressure.

The linear pulsation equations can be derived assuming the Lagrangian displacement $\boldsymbol{\xi}(\boldsymbol{r}, t) = \boldsymbol{\xi}(\boldsymbol{r})e^{i\omega_0 t}$ and the potential perturbation $\delta\Phi(\boldsymbol{r}, t) = \delta\Phi(\boldsymbol{r})e^{i\omega_0 t}$ with $\{\boldsymbol{\xi}(\boldsymbol{r}), \delta\Phi(\boldsymbol{r})\}$ and ω_0 being the to-be-determined eigenfunctions and eigenfrequency, respectively. Consequently, we obtain $\boldsymbol{v} = \partial\boldsymbol{\xi}(\boldsymbol{r}, t)/\partial t + (\boldsymbol{v}^{(0)} \cdot \nabla)\boldsymbol{\xi}(\boldsymbol{r}, t) = \partial\boldsymbol{\xi}(\boldsymbol{r}, t)/\partial t$ and $\sigma_{ij} = \sigma_{ij}^{(0)} + \Gamma_1 p^{(0)}\epsilon_{kk}\delta_{ij} + 2\mu(\epsilon_{ij} - \frac{1}{3}\epsilon_{kk}\delta_{ij})$, where $\Gamma_1 = (d\ln p/d\ln\rho)_{\rm ad}$ is the adiabatic index, $\epsilon_{ij} = \frac{1}{2}(\xi_{i,j} + \xi_{j,i})$ is the strain tensor and μ is the shear modulus. Plugging them into Eq. (2.10), we obtain the linear pulsation equation $[\mathcal{L}(r;\mu) - \omega_0^2]\boldsymbol{\xi}(\boldsymbol{r}) = 0$ with [64, 190]

$$\begin{aligned} \mathcal{L}(r;\mu)\boldsymbol{\xi} = &-\nabla\left(\frac{\Gamma_1 p}{\rho}\nabla\cdot\boldsymbol{\xi}\right) - \nabla\left(\frac{1}{\rho}\boldsymbol{\xi}\cdot\nabla p\right) + \nabla\delta\Phi \\ &+ \frac{1}{\rho}\left[\nabla\left(\frac{2}{3}\mu\nabla\cdot\boldsymbol{\xi}\right) - (\nabla\mu\cdot\nabla)\,\boldsymbol{\xi} - \nabla(\boldsymbol{\xi}\cdot\nabla\mu)\right] \\ &+ \frac{1}{\rho}\left[(\boldsymbol{\xi}\cdot\nabla)\nabla\mu - \mu\left(\nabla^2\boldsymbol{\xi} + \nabla(\nabla\cdot\boldsymbol{\xi})\right)\right] \,, \end{aligned} \tag{2.11}$$

and the linear Possion equation

$$\nabla^2\delta\Phi = -4\pi G(\boldsymbol{\xi}\cdot\nabla\rho + \rho\nabla\cdot\boldsymbol{\xi}) \,. \tag{2.12}$$

For spheroidal modes (for example, the i-mode), the displacement vector can be written as a variable-seperation form

$$\begin{aligned}
\boldsymbol{\xi}^r(\boldsymbol{r}) &= U(r)Y_{lm}(\theta,\phi)\,, \\
\boldsymbol{\xi}^\theta(\boldsymbol{r}) &= V(r)\frac{\partial Y_{lm}}{\partial\theta}(\theta,\phi)\,, \\
\boldsymbol{\xi}^\phi(\boldsymbol{r}) &= \frac{V(r)}{\sin\theta}\frac{\partial Y_{lm}}{\partial\phi}(\theta,\phi)\,, \\
\delta\Phi(\boldsymbol{r}) &= S(r)Y_{lm}(\theta,\phi)\,.
\end{aligned} \tag{2.13}$$

Plugging them into Eqs. (2.11,2.12), we obtain the governing equations of $\{U(r), V(r), S(r)\}$ [64, 190]

$$\begin{aligned}
\rho\omega_0^2 U =& \rho\frac{d\hat{\chi}}{dr} - \frac{d}{dr}\left(\frac{1}{3}\mu\hat{\alpha}\right) + \frac{d\mu}{dr}\left(\hat{\alpha} - 2\frac{dU}{dr}\right) \\
&- \mu\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dU}{dr}\right) - \frac{l(l+1)}{r^2}U + \frac{2l(l+1)}{r^2}V\right]\,, \\
\rho\omega_0^2 V =& \rho\frac{\hat{\chi}}{r} - \frac{1}{3}\frac{\mu\hat{\alpha}}{r} - \frac{d\mu}{dr}\left(\frac{dV}{dr} - \frac{V}{r} + \frac{U}{r}\right) \\
&- \mu\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dV}{dr}\right) - \frac{l(l+1)}{r^2}V + \frac{2U}{r^2}\right]\,, \\
\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dS}{dr}\right) - \frac{l(l+1)}{r^2}S =& 4\pi G\left(U\frac{d\rho}{dr} + \hat{\alpha}\rho\right)\,,
\end{aligned} \tag{2.14}$$

with

$$\begin{aligned}
\hat{\alpha} &= \frac{1}{r^2}\frac{d}{dr}(r^2U) - \frac{l(l+1)}{r}V\,, \\
\hat{\chi} &= -\frac{\Gamma_1 p}{\rho}\hat{\alpha} - \frac{1}{\rho}\frac{dp}{dr}U + S\,.
\end{aligned}$$

For simplicity, we take the Cowling approximation assuming $\delta\Phi = 0$ and solve the i-mode eigenvalue problem following Ref. [64].

2.B Appendix: Bayesian Parameter Estimation and Model Test

For the search of possible mode resonance in GW170817, we have incorporated $\delta\phi_a, f_a$ plus all the binary parameters (except for the source distance and sky location which are known from elec-

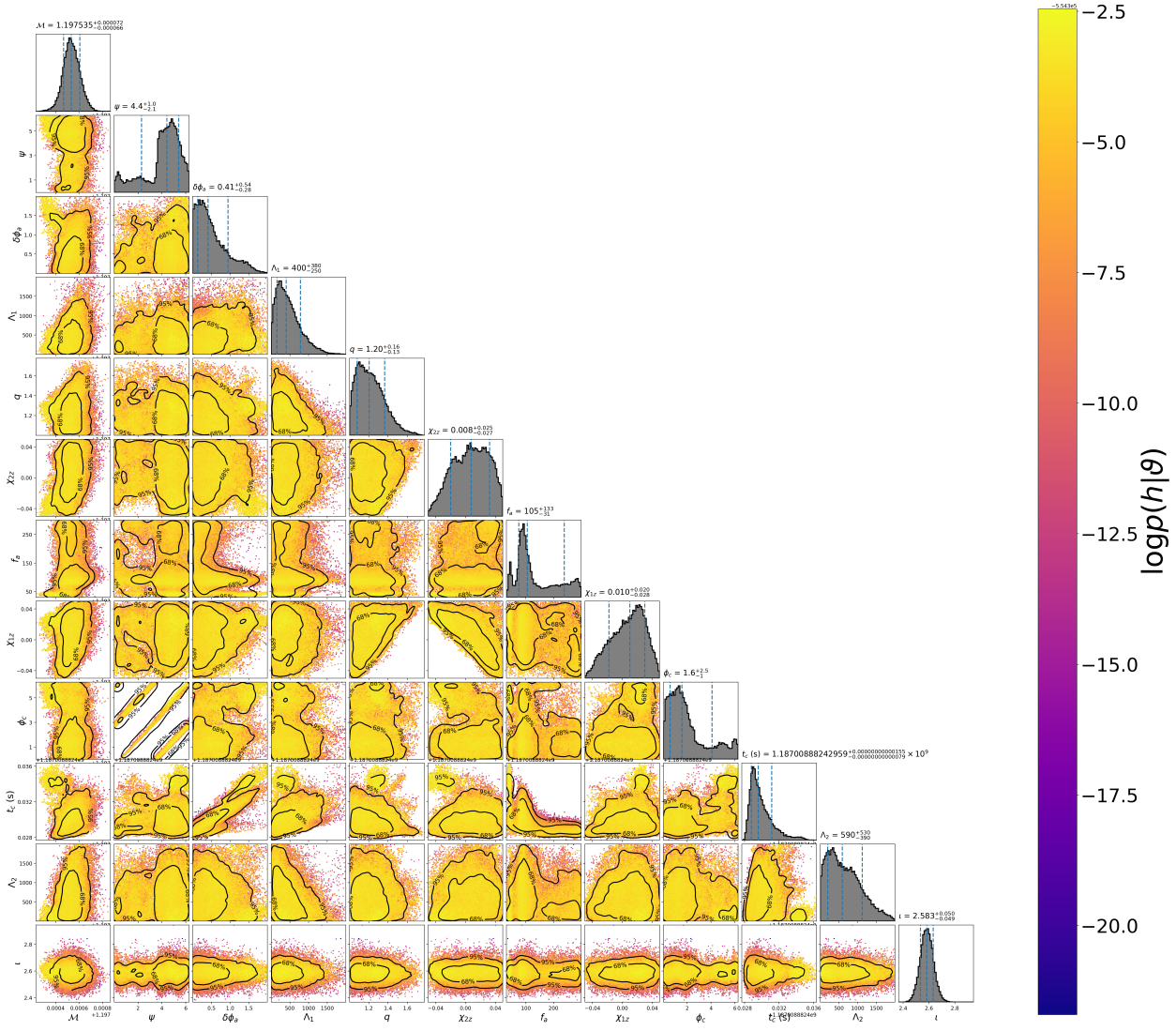

Figure 2.4: The posterior distribution of all parameters in the search of mode resonance presented in Fig. 3a in the maintext with data from GW170817.

tromagnetic counterparts), including chirp mass $\mathcal{M}$, mass ratio q, inclination angle ι, polarization phase ψ_c, coalescence phase ϕ_c, coalescence time t_c, tidal Love numbers of both stars $\Lambda_{1,2}$ and parallel spins of both stars $\chi_{1,2z}$. The priors of the spin are set to be $|\chi_{1,2z}| < 0.05$. The full posterior distribution of parameters and the Markov-Chain Monte-Carlo samples are presented in Fig. 2.4. In general, the accuracy of the search result not only depends on the event signal-to-noise ratio, but also on the melting frequency. If the melting frequency is too small, even if it is still in the LIGO band, the imbalance of the waveform signal-to-noise ratio before and after the melting process still degrades the search accuracy. For GW170817, given that the low-frequency sensitivity of the LIGO detectors in O2 is significantly worse than O3, we find that it is beneficial to set the lower bound of the frequency range to be at least 40 Hz to allow SNR ~ 5 in the waveform before the resonance. This situation will be greatly improved as LIGO reaches design sensitivity when the low-frequency performance is much better, and definitely for LIGO A+ and 3rd-generation detectors, which is important as crust melting may happen before 40 Hz.

We have applied Bayesian model selection formalism in the search for a resonance signature in the data of GW170817, in which case $\mathcal{H}_1$ is the model including the resonance effect and the null hypothesis $\mathcal{H}_0$ is the one without. We obtain $\log \mathcal{B}^1_0 \sim [-0.6, -0.3]$, so that there is no evidence of mode resonance in the parameter range we searched for in the strain data of GW170817.

The model selection method also applies to distinguish possible origins of the signal. For example, if we detect a signal by searching with our mode resonance waveform, it may also show a positive signal if we had searched for this signal with waveforms motivated by other reasons. To illustrate this, we injected a mode resonance signal ($\delta\phi_a = 0.3$, $f_a = 60$Hz) to simulated detector noise compatible with LIGO A+, and searched it with both our mode resonance waveform and the waveform for tidal-p-g coupling [111]:

$$\delta\Psi(f) = -\frac{2C}{3B^2(3-n_0)(4-n_0)}\left\{\Theta\left(\frac{f}{f_{\rm ref}}\right)^{n_0-3}\right.$$

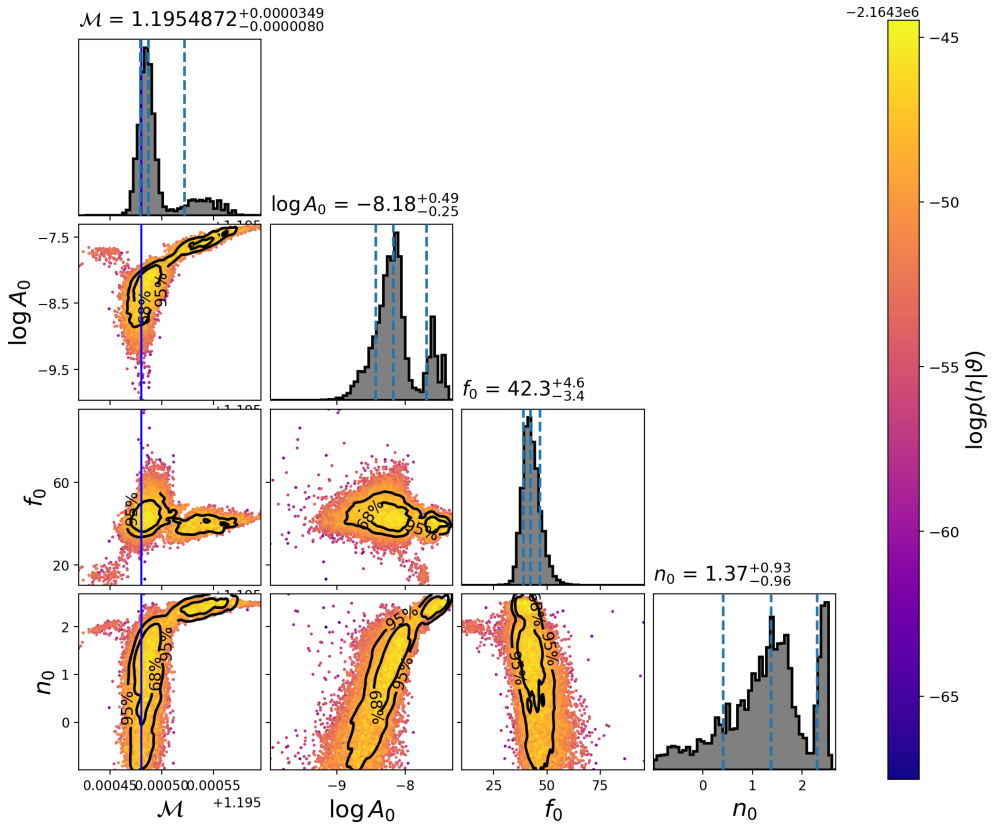

Figure 2.5: The posterior distribution of chirp mass $\mathcal{M}$, and A_0, n_0, f_0 as we try to fit the mock data with the tidal-p-g mode waveform.

$$+(1-\Theta)\left(\frac{f_0}{f_{\rm ref}}\right)^{n_0-3}\left[(4-n_0)-(3-n_0)\left(\frac{f}{f_0}\right)\right]\Bigg\}\,, \tag{2.15}$$

where $\Theta = \Theta(f - f_0)$, $f_{\rm ref} = 100$Hz, $C = A_0[(2m_1)^{2/3} + (2m_2)^{2/3}]/M^{2/3}$, and $B = (32/5)(G\mathcal{M}\pi f_{\rm ref}/c^3)^{5/3}$. The corresponding posterior distributions of parameters are shown in Fig. 2.5. The fitting with tidal-p-g coupling does not generate a compact posterior distribution of the parameters of this model, A_0, f_0 and n_0, although the distribution of $\log A_0$ is significantly different from the lower bound of its prior, which is -10. As we compare the two models, the Bayes factor $\log \mathcal{B}^{\rm res}_{\rm pg}$ is 2.7 ± 0.3, which shows a preference for the mode resonance model. This means that it is still possible to distinguish these two models when we detect a mode resonance signal

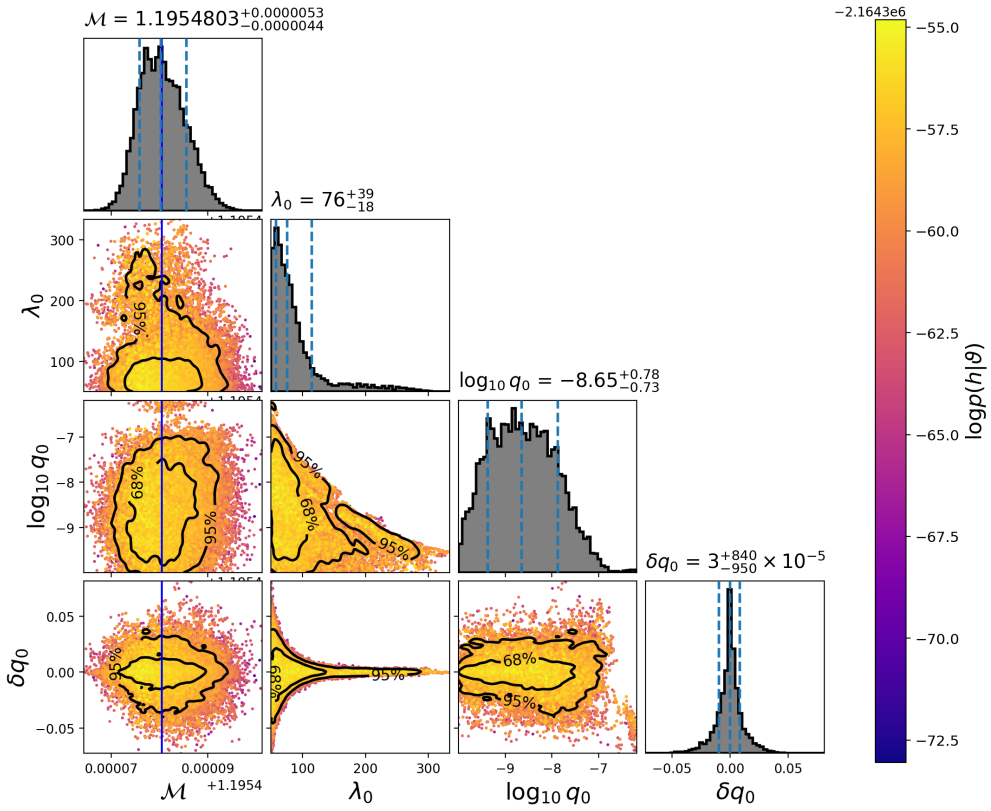

Figure 2.6: The posterior distribution of chirp mass $\mathcal{M}$, and $\lambda_0, q_0, \delta q_0$ as we try to fit the mock data with the waveform of scalar dipole radiation.

with LIGO A+.

In the case of double NSs carrying scalar (e.g., axions with mass m_s [185]) charge q_1 and q_2 , the BNS evolution would be modified by both the extra force mediated by the scalar and the extra scalar dipole radiation. To the leading order, the extra force can be described in term of Yukuwa potential $V_s(r) = -8q_1q_2e^{-m_s r}/r$ and scalar dipole emission power is $P_s(r) = \frac{1}{12}\frac{(q_1m_2-q_2m_1)^2}{M^2M^2_{\rm Planck}}\left(1-\frac{m_s^2}{\Omega^2}\right)^{3/2} r^2\Omega^4\Theta(\Omega^2-m_s^2)$, where $M = m_1 + m_2$ is the total mass of the BNS system, $\Omega(r)$ is the orbital frequency. For convenience, we define symmetry charge $q_0 := q_1q_2/m_1m_2$, anti-symmetry charge $\delta q_0 := q_1/m_1 - q_2/m_2$ and dimensionless variable $\lambda_0 := (GMm_s)^{-1}$. We find the GW phase shift driven

the extra scalar degree of freedom is $\delta\Psi(f) = [2\pi f \delta t_{\rm s}(f) - \delta\phi_{\rm s}(f)]\,\Theta(f - \frac{1}{\pi G M \lambda_0})$, with

$$
\begin{aligned}
\delta\phi_{\rm s}(f) = {} & \frac{5 q_0 \lambda_0^{5/2}}{12\eta}\Bigg[6\Gamma\left(\frac{5}{2}, \frac{1}{\lambda_0 v^2}\right) + 2\Gamma\left(\frac{7}{2}, \frac{1}{\lambda_0 v^2}\right) \\
& - \Gamma\left(\frac{9}{2}, \frac{1}{\lambda_0 v^2}\right)\Bigg] + \frac{25(\delta q_0)^2}{1344\eta v^7}\,{}_2F_1\left(-\frac{1}{2}, \frac{7}{6}, \frac{13}{6}, \frac{1}{\lambda_0^2 v^6}\right) \\
& - \frac{25(\delta q_0)^2}{2496\eta\lambda_0^2 v^{13}}\,{}_2F_1\left(-\frac{1}{2}, \frac{13}{6}, \frac{19}{6}, \frac{1}{\lambda_0^2 v^6}\right) - C_1\,, \\
\delta t_{\rm s}(f) = {} & \frac{5 q_0 \lambda_0^4 v^3}{24\eta\Omega}\Bigg[6\Gamma\left(4, \frac{1}{\lambda_0 v^2}\right) + 2\Gamma\left(5, \frac{1}{\lambda_0 v^2}\right) \\
& - \Gamma\left(6, \frac{1}{\lambda_0 v^2}\right)\Bigg] + \frac{5(\delta q_0)^2}{768\eta\Omega v^7}\,{}_2F_1\left(-\frac{1}{2}, \frac{5}{3}, \frac{8}{3}, \frac{1}{\lambda_0^2 v^6}\right) \\
& - \frac{25(\delta q_0)^2}{6144\eta\Omega\lambda_0^2 v^{13}}\,{}_2F_1\left(-\frac{1}{2}, \frac{8}{3}, \frac{11}{3}, \frac{1}{\lambda_0^2 v^6}\right) - C_2\,,
\end{aligned}
\tag{2.16}
$$

where $v(f) := (GM\Omega)^{1/3} = (GM\pi f)^{1/3}$, $\Gamma(a, z) := \int_z^\infty t^{a-1} e^{-t} dt$ is the gamma function, ${}_2F_1$ is the hypergeometric function, and $C_{1,2}$ are two integration constants enabling vanishing $\delta t_{\rm s}$ and $\delta\phi_{\rm s}$ at $f = \frac{1}{\pi G M \lambda_0}$. To illustrate the power of LIGO A+ distinguishing the scalar dipole radiation from the mode resonance, we also constrain the scalar radiation model using the same mock data above (Fig. 2.6) and we find the Bayes factor $\ln \mathcal{B}^{\rm res}_{\rm dipole} = 11.7 \pm 2.2$.

Chapter 3

First Constraints on Nuclear Coupling of Axionlike Particles from the Binary Neutron Star Gravitational Wave Event GW170817

Jun Zhang, Zhenwei Lyu, Junwu Huang, Matthew C. Johnson

Laura Sagunski, Mairi Sakellariadou, Huan Yang

Phys. Rev. Lett. 127, 161101 (Oct. 2021), arXiv:2105.13963

My specific contributions: Calculate axion induced phase together with Jun in appendix 3.A. I construct the waveform model by adding axion induced phase $\Psi_a(f)$ discussed in section 3.3 to the total phase and perform Bayesian inference in section 3.4 to search axion signal in event GW170817. I analysis the MCMC result and draw a violin plot shown in Fig. 3.2. Though we don't find such signal in GW170817, we discuss and explain the MCMC simulation result in section 3.5.

Light axion fields, if they exist, can be sourced by neutron stars due to their coupling to nuclear

matter, and play a role in binary neutron star mergers. We report on a search for such axions by analysing the gravitational waves from the binary neutron star inspiral GW170817. We find no evidence of axions in the sampled parameter space. The null result allows us to impose constraints on axions with masses below 10^{-11}eV by excluding the ones with decay constants ranging from 1.6×10^{16}GeV to 10^{18}GeV at 3σ confidence level. Our analysis provides the first constraints on axions from neutron star inspirals, and rules out a large region in parameter space that has not been probed by the existing experiments.

3.1 Introduction

Axions are hypothetical scalar particles that generally appear in many fundamental theories. An archetypal example is the QCD axion, a pseudo-scalar field proposed to solve the strong CP problem [191–194]. Light axions are also a unique prediction of string theory [195, 196], as well as one of the most compelling candidates for dark matter [197–199].

Axions have been constrained by measuring the energy loss and energy transport in various astrophysical objects, such as stars [200–202] and supernova 1987A [203, 204]. Further constraints can be imposed if axions make up all of the dark matter in our universe [117, 205–208]. In addition, axions with weak self-interactions could lead to black hole superradiance, and hence are constrained by the black hole spin measurements [209–214], the polarimetric observations [215], and the gravitational waves (GWs) emitted by the superradiance cloud [216–222]. Bosonic fields may also form compact objects that have GW implications [223].

In this Letter, as discussed in section 1.5.2, we search such axions using GW events GW170817 and use the conventions $\hbar = c = 1$.

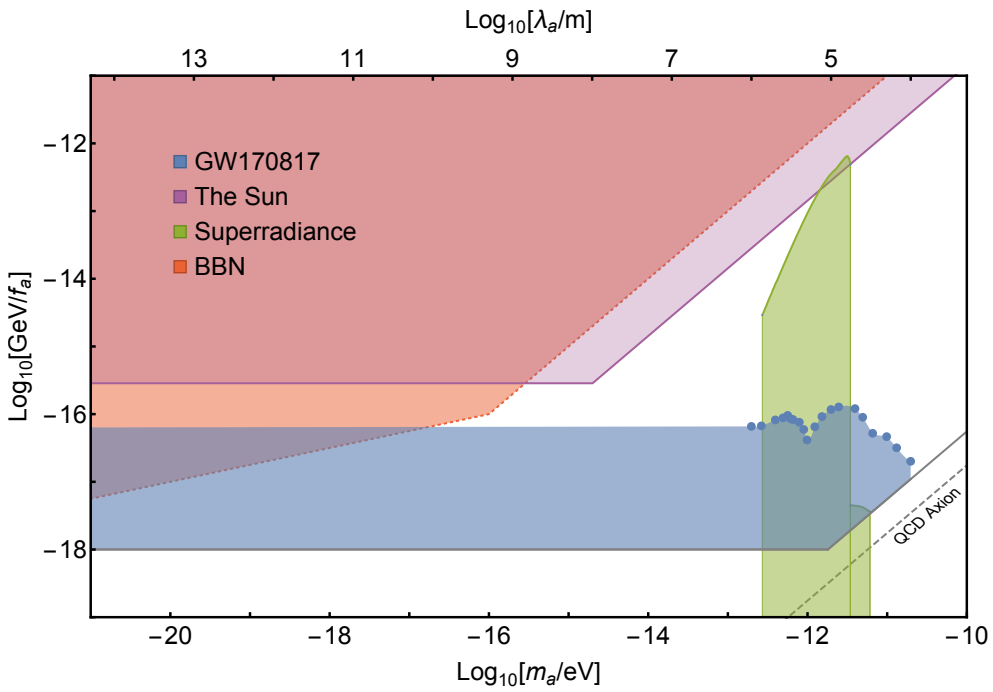

Figure 3.1: Constraints on the axion parameter space. m_a is the mass of axion field and f_a is the axion decay constant. The blue dots show the masses of axion fields that are sampled in this work, and the corresponding 3σ constraints on f_a from GW170817. For $f_a \gtrsim 1.6 \times 10^{16}$GeV, the GW170817 data cannot distinguish waveforms with $m_a \lesssim 10^{-13}$eV, allowing us to extrapolate the constraints on small m_a to the massless limit and to exclude the existence of axions in the blue regime. Axions in the purple region could also be significantly sourced by the Earth and the Sun, and hence are excluded [115]. In addition, we show the 3σ constraint from the spin measurements of the stellar mass black holes (Cyg X-1 and GRS 1915+105) [214] (in green), as well as the 1σ constraint on axion dark matter from the BBN [117] (in red).

3.2 Neutron Stars with Axions

We consider axions that couple to nuclear matter in a similar way as the QCD axion. The low energy effective potential is [115]

$$V(a) = -m_\pi^2 f_\pi^2 \epsilon \sqrt{1 - \frac{4 m_u m_d}{(m_u + m_d)^2} \sin^2\left(\frac{a}{2 f_a}\right)}, \tag{3.1}$$

where f_a is the axion decay constant, m_π and f_π are the pion mass and decay constant, and $m_{u,d}$ stands for the mass of the up, down quarks. Assuming $\epsilon < 0.1$, the mass of the axions is $m_a \simeq \sqrt{\epsilon}\, m_\pi f_\pi / f_a$, and is lighter than the mass of the QCD axion. In vacuum, the axion field is expected to stay at the minimum of its potential $a = 0$. Inside a dense object, such as a NS, the axion potential receives finite density corrections [224]

$$V(a) = -m_\pi^2 f_\pi^2 \left[\left(\epsilon - \frac{\sigma_N n_N}{m_\pi^2 f_\pi^2} \right) \left| \cos\left(\frac{a}{2f_a} \right) \right| + \mathcal{O}\left(\left(\frac{\sigma_N n_N}{m_\pi^2 f_\pi^2} \right)^2 \right) \right] \quad , \tag{3.2}$$

where n_N is the number density of nucleons, and $\sigma_N \approx 59\,\mathrm{MeV}$ [225][1]. For $\epsilon < 1$, the axion potential inside the dense object can change sign while perturbation theory is still valid. If the radius of the dense object is larger than the critical radius

$$R_{\rm crit} \equiv \frac{2f_a}{\sqrt{\sigma_N n_N - \epsilon m_\pi^2 f_\pi^2}}, \tag{3.3}$$

a phase transition occurs, shifting the vacuum expectation value of the axion field from 0 to $\pm\pi f_a$ inside the dense object. Assuming NSs have a radius on the order of 10km, this phase transition generally happens inside NSs for axions with $f_a \lesssim 10^{18}$GeV. As a result, the NS develops an axion profile, interpolating from $\pm\pi f_a$ near the NS surface to 0 at spatial infinity.

In this case, the axion field mediates an additional force between NSs, with strength that could in principle be as strong as gravity. The axion force cannot be sourced by nuclei (as nuclei are too small to trigger the phase transition), and can therefore avoid fifth force constraints in laboratories.

[1] Specific mechanisms that suppress the axion masses [113, 114] might also change the period of this low energy effective potential. However, the axion profile and subsequent analysis is determined exclusively by the finite density effect inside the NS, with period $2\pi f_a$. Therefore, our analysis applies to the light axions in [113, 114].

At leading order, the axion force between two NSs is

$$\mathbf{F}_a = -\frac{Q_1 Q_2}{4\pi r^2}\,(1 + m_a r)\exp[-m_a r]\,\hat{\mathbf{r}}\,, \tag{3.4}$$

where $Q_{1,2}$ is the axion charge carried by each NS and is related to the NS radius $R_{1,2}$ through

$$Q_{1,2} = \pm 4\pi^2 f_a R_{1,2}\,. \tag{3.5}$$

The axion force can be either attractive or repulsive, depending on whether the axion field values are of the same or opposite sign on the surfaces of the two NSs. Moreover, the axion force is only "turned on" if the two NSs are within the axion's Compton wavelength $\lambda_a \equiv 1/m_a$.

If such NSs form binaries, the axion field might also radiate axion waves during binary coalescence. For circular orbits, the leading order radiation power is

$$P_a = \frac{(Q_1 M_2 - Q_2 M_1)^2}{12\pi\,(M_1 + M_2)^2} r^2 \Omega^4 \left(1 - \frac{m_a^2}{\Omega^2}\right)^{3/2}\,, \tag{3.6}$$

where Ω is the orbital frequency, r denotes the separation between the two NSs of masses M_1 and M_2. According to Eq. (3.6), the axion radiation is turned on only when the orbital frequency is larger than the axion mass. The axion force as well as the axion radiation power are calculated to the next-to-leading order in [116].

3.3 Waveform Template

Inspirals in the presence of a generic massive scalar field have been studied in [116, 226–232], among which corrections of the scalar field on the GR waveforms are calculated to the first post-Newtonian (PN) order in [116]. The waveform template cannot be written in a closed analytic

form, and cannot be described by the usual PN templates, e.g., the one used in [233]. In our analysis, the waveform is generated by a modified TaylorF2 template, in which the frequency domain waveform is given by

$$h(f) \simeq H(f) \exp\left[i\Psi(f)\right]. \tag{3.7}$$

Since the existing analyses of GW170817 [106, 158] show good agreement with GR, the axion charges, if non-zero, must be very small, which allows us to expand Ψ as

$$\Psi = \Psi_{\rm GR} + \Psi_a + O(Q_{1,2}^4) + O(Q_{1,2}^2 v^2). \tag{3.8}$$

Here $\Psi_{\rm GR}$ is the phase in the usual TaylorF2 template in the PyCBC package [234], Ψ_a is the leading order phase correction caused by the axion field, and v^2 counts the PN order. The expression of Ψ_a can be found in the Supplementary Material. In practice, we only consider the leading order correction caused by the axion field, which is justified by the necessary smallness of the axion charge.

Generally, taking into account the leading correction from a massive scalar field introduces three parameters in the waveform template, i.e., the scalar charge of each star and the mass of the scalar field. In our case, the two charges Q_1 and Q_2 are given by Eq. (3.5) and hence are not independent. Thus, we define

$$\gamma_a \equiv \frac{Q_1 Q_2}{4\pi G M_1 M_2}, \tag{3.9}$$

a dimensionless parameter that characterizes the relative strength of the axion and gravitational force between the two NSs. The effects of the axion field are then parameterized by m_a and γ_a. In order to obtain each charge $Q_{1,2}$ from γ_a, we first use the universal $\Lambda - C$ relation [235–237] to compute the compactness and hence the radius of each NS. Then with Eq. (3.5) we compute

Q_1/Q_2, and eventually obtain the two charges Q_1 and Q_2 that are used to generate the waveform.

Moreover, we assume the two NSs obey the same equation of state (EOS), in which case their tidal deformabilities Λ_1 and Λ_2 are related. Following [158], we consider that the symmetric tidal deformability $\Lambda_s \equiv (\Lambda_2 + \Lambda_1)/2$, the antisymmetric tidal deformability $\Lambda_a \equiv (\Lambda_2 - \Lambda_1)/2$ and the mass ratio of the binary $q \equiv m_2/m_1 \leq 1$ are related through an EOS-insensitive relation $\Lambda_a(\Lambda_s, q)$ [238, 239]. In Bayesian analysis, we sample uniformly in the symmetric tidal deformability $\Lambda_s \in [0, 2000]$, while Λ_a and hence Λ_1 and Λ_2 are obtained using the EOS-insensitive relation $\Lambda_a(\Lambda_s, q)$ which is tuned to a large set of EOS models [240, 241].

3.4 Bayesian Inference

To search for axions, we scan the parameter space by sampling axion fields with different masses (see Fig. 3.1 for the masses). In addition, we also consider the massless limit $m_a = 0$. For each mass, we perform a Bayesian analysis of GW170817, taking into account the possible dephasing caused by the axion field in the inspiral waveform. In particular, we consider a set of parameters $\boldsymbol{\vartheta} = (\gamma_a, \boldsymbol{\vartheta}_{\rm NS})$, and evaluate the posterior probability density function $p(\boldsymbol{\vartheta}|d)$ given the GW170817 data d. Here $\boldsymbol{\vartheta}_{\rm NS}$ includes chirp mass $\mathcal{M}$, mass ratio q, coalescence time, coalescence phase, polarization, inclination, spins of two NSs, and symmetric tidal deformability which are defined in the usual TaylorF2 waveform template. In these analyses, we fix the luminosity distance $D_L = 40.7$Mpc [242] and the sky localization (RA, Dec) = (197.450374, −23.381495) [243] for GW170817, as they have been accurately measured independently.

In order to determine the posterior distribution over the parameters $\boldsymbol{\vartheta}$, we make use of the Markov-chain Monte Carlo algorithm as implemented in the PyCBC package [234]. For the likelihood calculation, we use GW170817 data version 3 released by the LIGO and Virgo scientific collaboration on GW open science center [244], and assume a Gaussian noise model with a low

frequency cutoff of 20Hz. We only use LIGO Hanford and Livingston data, since the signal to noise ratio (SNR) of Virgo data is far smaller [106].

The priors on γ_a are chosen to be $(-0.1,\ 0.1)$. The sign of γ_a indicates whether the axion force is attractive or repulsive. In principle, the probability of an attractive or repulsive axion force can be different, depending on the formation history of the binary. Nevertheless, we assume the same prior on positive and negative γ_a for simplicity.

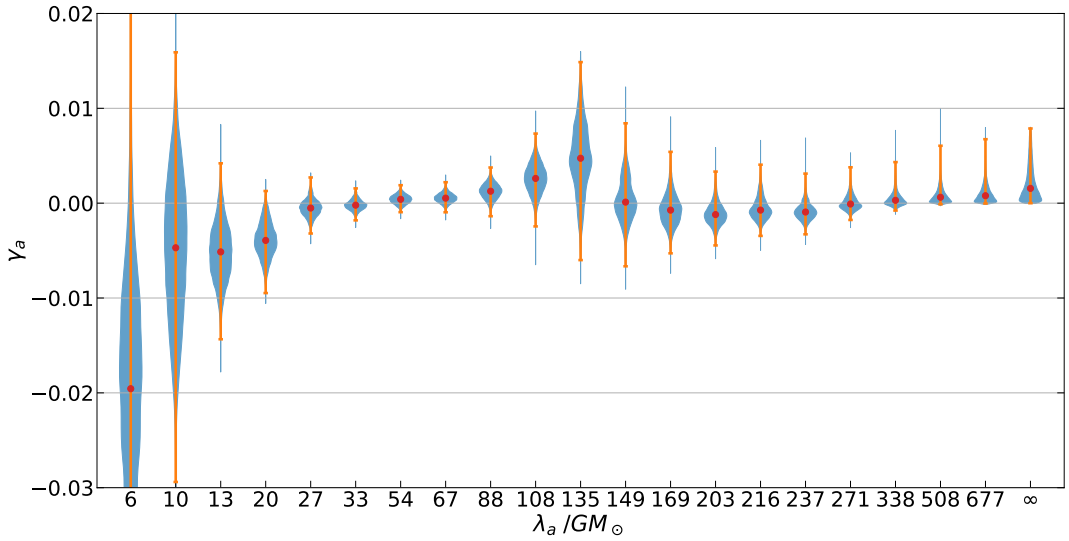

Figure 3.2: Posteriors on γ_a of axions with different masses. $\lambda_a \equiv 1/m_a$ is the Compton wavelength. The horizontal bars mark the 3σ standard deviations. The deviation increases rapidly at small λ_a as the axion effects is suppressed by m_a. The deviation is large around $\lambda_a \sim 135GM_\odot$ due to the degeneracy between γ_a and the chirp mass $\mathcal{M}$. This degeneracy is partially broken (especially for $\gamma_a < 0$) at large λ_a by the axion radiation. The posteriors eventually approach to that in massless limit (noted as $\lambda_a = \infty$) since the GW170817 data is insensitive to waveform with $|\gamma_a| \leq 10^{-2}$ at large λ_a.

3.5 Results

We focus on the posteriors over γ_a, which are shown in Fig. 3.2 as a function of the axion Compton wavelength. The posteriors show no significant evidence for non-zero γ_a, and are compatible with $\gamma_a = 0$ at 3σ confidence level over the full range of axion masses sampled.

The standard deviation of γ_a increases dramatically when λ_a becomes smaller than $10\,GM_\odot$, in which case λ_a becomes less than the NS radii and the effects of the axion field is suppressed. On the other hand, for $\lambda_a \gtrsim r_{\rm cut}$, where $r_{\rm cut} \simeq 110\,GM_\odot$ is the separation between the two NSs corresponding to the 20Hz low frequency cutoff, the axion force (3.4) behaves like a Newtonian force during the whole observed inspiral stage. Without axion radiation, γ_a would be highly degenerate with the chirp masses $\mathcal{M}$. This is indeed the case for $\lambda_a \sim 135GM_\odot$, where the axion radiation is still not significant, and the standard deviation of γ_a is large due to the degeneracy between γ_a and $\mathcal{M}$. As λ_a increases, the axion radiation becomes significant and breaks this degeneracy, especially for negative γ_a. For positive γ_a, the charge difference is small if the radii of the two NSs are comparable, thus the axion radiation is always weaker than for negative γ_a. This is also why the constraints on negative γ_a are better than those for positive γ_a at large λ_a. The degeneracy can also be partially broken by considering the induced charge effect studied in [116], which could improve the constraints at large λ_a by a factor of roughly two.

The posteriors on γ_a become stable for $\lambda_a > 338GM_\odot$. This is because $|\gamma_a|$ is constrained to be smaller than 10^{-2} for axions with mass $\lambda_a > 338GM_\odot$. With such a small $|\gamma_a|$, the phase difference is less than $O(1)$, and hence the GW170817 data has no distinguishing power. Indeed, we find that the posteriors with $\lambda_a > 338GM_\odot$ eventually approach the posterior in the massless limit. The insensitivity of posterior on large λ_a allows us to also impose constraints on axions with $\lambda_a > 338GM_\odot$.

To draw conclusions about axion fields, we project the 3σ constraints on γ_a onto the decay constant f_a. We combine the constraints from positive and negative γ_a by picking the weakest one.

As shown in Fig. 3.1, the constraints on γ_a translate to a constraint of $f_a < 1.6\times10^{16}$GeV on axions with $m_a \leq 10^{-11}$eV. On the other hand, for axions with $f_a > 10^{18}$GeV, the critical radius is so large that NSs cannot trigger the phase transition. In other words, axions with $f_a > 10^{18}$GeV cannot be sourced by NSs even if they exit, and are free from the NS inspiral constraint. Therefore, our analysis indicates that GW170817 imposes constraint on axions with masses below 10^{-11}eV by excluding the ones with decay constants $1.6 \times 10^{16}\text{GeV} < f_a < 10^{18}\text{GeV}$.

3.6 Conclusions and Discussion

Our analysis provides the first constraint on axions from NS inspirals, and excludes a large parameter space that has not been probed by existing experiments. As a comparison, in Fig. 3.1 we show the 3σ constraint from the spin measurements of the stellar mass black holes (in green) [214], as well as the 1σ constraint on axion dark matter from BBN (in red) [117]. In addition, axions can be constrained by the absence of GWs emitted by the superradiance cloud around stellar mass black holes [218–222]. Since this constrain is also based on the absence of the superradiance, it excludes a similar parameter space as the one that is excluded by the spin measurements of stellar mass black holes. We emphasize that our analysis imposes constraint on parameter space that cannot be covered by the existing experiments. For example, superradiance can only be used to probe axions whose Compton wavelength is comparable or slightly larger than the size of black holes. Therefore, the superradiance constraints, from both black hole spin measurements and the GWs emitted by superradiance clouds, cannot probe axions with very small masses due to the lack of the corresponding heavy black holes or the low superradiance efficiency. Moreover, our analysis does not rely on the assumption that the axions make up the dark matter, which is required for the BBN [117] and the neutron electric dipole moment (nEDM) [206, 207] constraints. Especially, the kinetic energy and momentum of axions with $f_a \lesssim 10^{17}$GeV would change by more than $O(1)$

near the Earth due to the finite density corrections, therefore most of the constraint from the Earth based nEDM experiments are in question. Beside the above constraints, axions with smaller f_a (in purple region in Fig. 3.1) can be sourced by the Earth and the Sun for the same reason, and hence are excluded [115]. Also see [245, 246] for constraints from pulsars.

We did not consider the induced charge effect, whose relative magnitude is $v^2 \exp(-m_a r)/C_{1,2}$ comparing to the axion effects considered in this Letter (see Supplementary Material). This effect could become important at the late inspirals for axions with small masses, and could potentially extend the excluded region to $10^{16}\text{GeV} < f_a < 10^{18}\text{GeV}$ for $m_a \lesssim 10^{-14}\text{eV}$. However, including of this effect requires further understanding on how the induced charges affect the axion radiations, and is beyond the scope of this work.

Constraint from binary NS inspirals can be further improved if the SNR of the merger event is enhanced, for example by stacking multiple binary NS merger events or with the next generation GW detectors. We expect the constraint on f_a to improve by a factor of $\sqrt{N}$ if the SNR is improved by a factor of N. In addition, assuming a similar SNR as GW170817, the constraint could also be improved by roughly two orders of magnitude if we observe a NS-black hole merger, as in which case the axion radiation is not suppressed by the small charge difference and there is no degeneracy between parameters for axions with small masses. A joint analysis of the events GW190425 [188] and GW190814 [247], which may contain NSs, is left for future work.

3.A Appendix: Supplementary Material

Flowing section 1.5.2, in the presence of the axions, the leading corrections to the GR binding energy and radiation power are given by

$$V_a = -\frac{Q_1 Q_2}{4\pi}\frac{e^{-m_a r}}{r}\left(1 - 16 p\, GM \frac{e^{-m_a r}}{r}\right) \tag{3.10}$$

and

$$P_a = \frac{\left(\bar{Q}_1 M_2 - \bar{Q}_2 M_1\right)^2}{12\pi} r^2 \Omega^4 \left(1 - \frac{m_a^2}{\Omega^2}\right)^{3/2} \tag{3.11}$$

with

$$p \equiv \frac{1}{M}\left(\frac{Q_1}{Q_2} p_2 + \frac{Q_2}{Q_1} p_1\right) \tag{3.12}$$

and

$$\bar{Q}_{1,2} \equiv Q_{1,2}\left(1 - 8G p_{2,1} \frac{e^{-m_a r}}{r}\right). \tag{3.13}$$

Here Ω is the orbital frequency, r denotes the separation between the two NSs of masses M_1 and M_2, and $M \equiv M_1 + M_2$ is the total mass. Comparing to Eqs. (3.4) and (3.6), we also include terms proportional to $p_{1,2}$ that could in principle arise due to the present of a generic scalar field. The value of $p_{1,2}$ is model dependent. For axions, these terms characterize the induced charge effect, and $p = (R_1 + R_2)/16GM$ when $m_a = 0$. Thus, we expect that the induced charge effect could become important at the late stage of inspirals for axions with small masses. However, taking into account this effect requires further studies on how $p_{1,2}$ relates to the parameters of the neutron stars and the axion field. Therefore, we neglected the induced charge effect in our analysis.

Given the fact that the axion charge $Q_{1,2}$ must be small, we neglect terms of $O(Q_{1,2}^4)$, $O(Q_{1,2}^2 v^2)$ and higher when we evaluate Eq. (1.40). In this case, the phase $\Psi(f)$ is given by Eq. (3.8) with

$$\Psi_a = \Psi_a^E + \begin{cases} \Psi_a^{P_>} & x > \alpha \\ \Psi_a^{P_<} & x \leq \alpha \end{cases}, \tag{3.14}$$

where we have defined

$$\Psi_a^E = \frac{5}{64}\frac{\gamma_a e^{-\frac{\alpha}{x^{2/3}}}}{\eta x^{5/3}}\left[-4-\frac{32x^{2/3}}{\alpha}-\frac{138x^{4/3}}{\alpha^2}-\frac{360x^2}{\alpha^3}+\frac{360x^{8/3}\left(e^{\frac{\alpha}{x^{2/3}}}-1\right)}{\alpha^4}-\frac{21\sqrt{\pi}x^{5/3}e^{\frac{\alpha}{x^{2/3}}}\,\mathrm{Erf}\left(\frac{\alpha^{1/2}}{x^{1/3}}\right)}{\alpha^{5/2}}\right], \tag{3.15}$$

$$\begin{aligned}\Psi_a^{P_>} = \frac{5}{254951424}\frac{\delta q^2}{\eta x^{16/3}}\Bigg[&-\sqrt{x^2-\alpha^2}\left(-822640\alpha^2+\frac{227089x^6}{\alpha^4}+\frac{261342x^4}{\alpha^2}+671304x^2\right)\\&+\frac{140049x^7}{\alpha^4}\,{}_2F_1\left(-\frac{5}{6},\frac{1}{2};\frac{1}{6};\frac{\alpha^2}{x^2}\right)+320x\left(1183\alpha^2+\frac{512x^6}{\alpha^4}-\frac{684x^4}{\alpha^2}-741x^2\right){}_2F_1\left(-\frac{1}{2},-\frac{1}{3};\frac{2}{3};\frac{\alpha^2}{x^2}\right)\\&+960x\left(-1183\alpha^2-\frac{80x^6}{\alpha^4}+\frac{684x^4}{\alpha^2}+741x^2\right){}_2F_1\left(-\frac{1}{3},\frac{1}{2};\frac{2}{3};\frac{\alpha^2}{x^2}\right)\Bigg],\end{aligned} \tag{3.16}$$

$$\Psi_a^{P_<} = \frac{25\sqrt{\pi}}{1536}\frac{\delta q^2}{\alpha^{10/3}\eta}\frac{\Gamma\left(\frac{5}{3}\right)\Gamma\left(\frac{11}{3}\right)x-\Gamma\left(\frac{7}{6}\right)\Gamma\left(\frac{25}{6}\right)\alpha}{\Gamma\left(\frac{11}{3}\right)\Gamma\left(\frac{25}{6}\right)}. \tag{3.17}$$

Here Erf is the Gauss error function, ${}_2F_1$ is the hypergeometric function, and

$$\begin{aligned}&\alpha \equiv GMm_a, \quad \eta \equiv \frac{M_1M_2}{M^2}, \quad x \equiv \pi GMf\,,\\&\delta q \equiv \frac{1}{4\sqrt{2\pi G}}\left(\frac{Q_1}{M_1}-\frac{Q_2}{M_2}\right).\end{aligned} \tag{3.18}$$

Chapter 4

Constraints on Einstein-dilation-Gauss-Bonnet gravity from black hole-neutron star gravitational wave events

Zhenwei Lyu, Nan Jiang, Kent Yagi

Phys. Rev. D 105, 064001 (March 2022), arXiv:2201.02543

My specific contributions: Construct EdGB waveform model through ppE formalism, which is including EdGB corrections to GR waveform model IMRPhenomXPHM in section 4.2.2. Then perform all MCMC simulation in section 4.3 and analysis the MCMC result in section 4.4. The draft is written by me and polished by Kent.

Recent gravitational wave observations allow us to probe gravity in the strong and dynamical field regime. In this paper, we focus on testing Einstein-dilation Gauss-Bonnet gravity which is mo-

tivated by string theory. In particular, we use two new neutron star black hole binaries (GW200105 and GW200115). We also consider GW190814 which is consistent with both a binary black hole and a neutron star black hole binary. Adopting the leading post-Newtonian correction and carrying out a Bayesian Markov-chain Monte Carlo analyses, we derive the 90% credible upper bound on the coupling constant of the theory as $\sqrt{\alpha_{\rm GB}} \lesssim 1.33\,{\rm km}$, whose consistency is checked with an independent Fisher analysis. This bound is stronger than the bound obtained in previous literature by combining selected binary black hole events in GWTC-1 and GWTC-2 catalogs. We also derive a combined bound of $\sqrt{\alpha_{\rm GB}} \lesssim 1.18\,{\rm km}$ by stacking GW200105, GW200115, GW190814, and selected binary black hole events. In order to check the validity of the effect of higher post-Newtonian terms, we derive corrections to the waveform phase up to second post Newtonian order by mapping results in scalar-tensor theories to Einstein-dilation Gauss-Bonnet gravity. We find that such higher-order terms improve the bounds by 14.5% for GW200105 and 6.9% for GW200115 respectively.

4.1 Introduction

Recent updates of the gravitational-wave (GW) catalog (GWTC-3) [1–4] reports, in total, 90 gravitational wave events from binary black hole (BBH), binary neutron star (BNS), and neutron star black hole (NSBH) mergers (see [248–250] for the previous catalogs). These events have been used to obtain implications on astrophysics, cosmology, nature of black holes (BHs) and nuclear physics (see studies on e.g. population properties of compact objects [251], Hubble tension [252], stochastic GW background [253], black hole spectroscopy [254], equations of state of neutron stars (NSs) [255, 256], and possible mode instabilities driven by NS tidal effects [257–259]). GW events are also ideal sources to probe strong/dynamical fields of gravity [233, 260–263] that are difficult to access through other experiments/observations, including table-top and solar system experiments, or binary pulsar and cosmological observations. For example, they have been used to

probe the mass of the graviton [260–262], scalar-tensor theories (Brans-Dicke theory, those with scalarization phenomena proposed by Damour and Esposito-Farèse, screened modified gravity, and the time dependence of the scalar field) [261, 264, 265], light axion fields sourced by neutron stars [6], and dynamical Chern-Simons gravity [66, 67, 261, 266, 267]).

Scalar Gauss-Bonnet (sGB) gravity [268–271] is another theory beyond General Relativity (GR) that has been studied extensively. In the action, a dynamical scalar field is coupled to a Gauss-Bonnet (GB) invariant (consisting of a certain combination of curvature-squared scalars) with a coupling constant $\alpha_{\rm GB}$ that has a dimension of length squared. Depending on what kind of coupling one considers, one recovers a shift-symmetric theory (linear coupling) [272, 273], Einstein-dilation Gauss-Bonnet (EdGB) gravity [274–277] (exponential coupling) motivated by string theory and inflation [278, 279], and a theory admitting spontaneous scalarization of BHs and NSs (quadratic coupling is an example) [280–283].

EdGB gravity has been constrained by GWs from BBHs that is summarized in Table 4.1, together with other astrophysical constraints from a BH low-mass x-ray binary (LMXB) and NS observations. The current upper bound on the coupling constant $\sqrt{\alpha_{\rm GB}}$ is $\sim$ 1km. For example, Perkins et al. [67] combined bounds on $\sqrt{\alpha_{\rm GB}}$ from 6 selected BBH events from the GW catalogs GWTC-1 and GWTC-2 and found the bound $\sqrt{\alpha_{\rm GB}} \lesssim 1.7\,{\rm km}$. These GW bounds are obtained by taking into account the leading correction to the gravitational waveform phase that enters at -1 post-Newtonian (PN) order relative to GR due to the scalar dipole radiation [261, 272]. Such a correction is derived within the small coupling approximation, where the coupling constant $\alpha_{\rm GB}$ is assumed to be much smaller than the characteristic curvature scale of a system (e.g. the mass for a BH) and one keeps only to $O(\alpha_{\rm GB}^2)$. Under this approximation, EdGB gravity effectively reduces to shift-symmetric GB gravity with a linear coupling between the scalar field and the GB invariant.

In this paper, we derive new bounds on EdGB gravity through GWs from NSBH binaries. Some forecasts on constraining the theory with such systems were made in [286] based on a Fisher analysis. The authors showed that the existing bounds can be improved further for NSBH binaries

	LMXB	NS	GW (BBH)		**GW (NSBH) (this work)**	
			O1–O2	O1–O3	**GW200115**	**combined**
$\sqrt{\alpha_{\rm GB}}$ [km]	1.9 [269]	1.29 [284]	5.6 [66], 1.85 [285], 4.3 [88]	1.7 [67], 4.5 [121], (0.4) [121]	1.33	1.18

Table 4.1: Astrophysical bounds on EdGB gravity. We show bounds from a LMXB, NSs ($\sim 2M_\odot$ NSs), GWs from BBHs, and NSBHs (this work). The one in brackets comes from GW190814 assuming that it is a BBH, which has some uncertainty. For NSBH, we present the bound from GW200115 and that by combining NSBHs (GW200115, GW200105, and GW190814; assuming the last one as a NSBH is a conservative choice) and BBHs from [67].

with a sufficiently small BH mass. We here derive new bounds through a Bayesian analysis using GW200105 and GW200115 [107]. We also consider GW190814, which is consistent with BBH or NSBH, and find bounds on EdGB gravity for the BBH and NSBH assumptions separately. We perform Bayesian inference to analyze the above events by adopting IMRPhenomXPHM waveform [81, 287, 288] (a phenomenological inspiral-merger-ringdown waveform for precessing BBHs in GR) as our base GR waveform and include EdGB corrections to the inspiral phase. We set a high frequency cutoff as $f_{\rm high} = 0.018/M$ [109] (for the total mass M in a unit of second) on the strain data, since the EdGB modifications to the waveform within the PN expansion is only valid for the phase at the inspiral stage. We also carry out independent Fisher analyses for cross-checking the results.

We improve previous analyses by deriving and including EdGB corrections to the waveform phase to higher PN orders. Recently, Shiralilou et al. [119, 289] derived the waveform valid to 1PN order higher than the leading tensor/scalar non-dipole and scalar dipole emission respectively. We update this by taking the waveform in scalar-tensor theories (in the Jordan frame) valid to 2PN relative to the leading for each of dipole and non-dipole contributions [290]. We apply a conformal transformation in scalar-tensor theories to go from the Jordan frame to the Einstein frame, find the mapping between the scalar fields in scalar-tensor theories and EdGB gravity, and use the scalar charges for BHs and NSs in the latter theory. We checked that this correctly reproduces the leading

−1PN correction in EdGB gravity known previously [261, 272].

We find the following results. First, using the leading EdGB correction to the phase, we find the 90% credible upper bound on $\sqrt{\alpha_{\rm GB}}$ as $\sqrt{\alpha_{\rm GB}} \lesssim 1.33\,{\rm km}$ for GW200115. This bound is stronger than the bound $\sqrt{\alpha_{\rm GB}} \lesssim 1.7\,{\rm km}$ in [67] obtained by combining selected BBHs from GWTC-1 and GWTC-2 catalogs. We also derive combined bounds by stacking posterior distributions on $\sqrt{\alpha_{\rm GB}}$ from GW200105, GW200115, and GW190814 (conservatively assuming it is a NSBH), and BBHs considered in [67], and find $\sqrt{\alpha_{\rm GB}} \lesssim 1.18\,{\rm km}$. These results are also summarized in Table 4.1. We next study the effect of including higher PN corrections. We find that such corrections do not make a significant difference on the bound on $\sqrt{\alpha_{\rm GB}}$ from the case with the leading correction, but improve the bound by 14.5% for GW200105 and 6.9% for GW200115 respectively. Such a finding is consistent with the analysis in [67].

This paper is organized as follows. We first review EdGB gravity and corrections to the waveform phase in Sec. 4.2. We next explain in Sec. 4.3 two methods of data analysis adopted in this paper, namely Bayesian inference through Markov-chain Monte Carlo (MCMC) and a Fisher analysis. In Sec. 4.4, we present our results and conclude in Sec. 4.5. We use the convention $G = c = 1$ throughout the paper.

4.2 Einstein-dilation Gauss-Bonnet Gravity

Let us first review EdGB gravity within the context of sGB theory and explain corrections to the gravitational waveform from GR.

4.2.1 Theory

We begin by presenting the action for sGB gravity [268–271]:

$$S = \int d^4x\sqrt{-g}\Big[\frac{R}{16\pi} - \frac{1}{2}(\nabla\phi)^2 + \alpha_{\rm GB} f(\phi)\mathcal{R}^2_{\rm GB}\Big] + S_m\,. \tag{4.1}$$

Here g is the determinant for the metric $g_{\mu\nu}$, R is the Ricci scalar, ϕ is a scalar field, $\alpha_{\rm GB}$ is the coupling constant between the scalar field and the metric, S_m is the matter action, and

$$\mathcal{R}^2_{\rm GB} = R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho} - 4R_{\mu\nu}R^{\mu\nu} + R^2\,, \tag{4.2}$$

is the GB invariant. $f(\phi)$ is an arbitrary function of the scalar field that determines how it is coupled to the metric. EdGB gravity is realized by choosing $f(\phi) = e^{-\gamma\phi}$ for a constant γ. As shown in [291, 292], this theory can be written in a second-order, hyperbolic form that is well-posed for numerical relativity evolution within a range of parameter space.

String theory predicts even higher order curvature terms in the action that we do include in the analysis. To justify this and treat the theory as an effective field theory, we work in the small coupling approximation scheme (or reduced-order scheme) where we assume that the GR contribution is dominant and handle EdGB corrections as small perturbations. In particular, we define a dimensionless coupling constant

$$\zeta \equiv \frac{16\pi\alpha^2_{\rm GB}}{L^4}\,, \tag{4.3}$$

where L is the characteristic length of the system and assume $\zeta \ll 1$. This technique has been used to find scalar charges of compact objects [272, 293, 294], corrections to the GW phase at the inspiral stage [272], and to carry out numerical simulations of BBH mergers [295].

Let us study the theory within the small coupling approximation scheme in more detail. We perturb field equations in $\alpha_{\rm GB}$ and solve them order by order. Then, $\phi = O(\alpha_{\rm GB})$ and one can expand $f(\phi)$ in small ϕ as:

$$f(\phi) = f(0) + f'(0)\phi + O(\phi^2)\,. \tag{4.4}$$

The first term is a constant and this does not change the field equations from the GR ones as the GB invariant is a topological term and can be rewritten as a total derivative. Thus, the leading effect comes from the second term where the scalar field is linearly coupled to the GB invariant. For this

reason, we consider the following action in this paper:

$$S = \int d^4x \sqrt{-g}\left[\frac{R}{16\pi} - \frac{1}{2}(\nabla\phi)^2 + \alpha_{\rm GB}\phi\mathcal{R}^2_{\rm GB}\right] + S_m\,, \tag{4.5}$$

where we have absorbed $f'(0)$ into $\alpha_{\rm GB}$. In this theory, BHs can have non-vanishing scalar charges [272, 293] while NSs do not [294].

Current astrophysical bounds on $\sqrt{\alpha_{\rm GB}}$ are summarized in Table 4.1. Besides these, one could use electromagnetic radiation emitted by gas or stars orbiting BHs. For example, simulations of the reflection spectrum of thin accretion disks with present and future X-ray missions show that current missions cannot distinguish BHs in GR and those in sGB gravity, while next-generation missions may be able to distinguish them [296]. Another possibility is to use Solar System experiments, though they are weaker than the astrophysical bounds in Table 4.1 by six orders of magnitude [269, 297] as the curvature of spacetime in the vicinity of the Sun is much smaller than that of BHs and NSs.

4.2.2 Gravitational Waveforms

We next find EdGB corrections to the gravitational waveform phase. Given that most of the signal-to-noise ratios (SNRs) for GWs from NSBHs and (small mass) BBHs come from the inspiral portion, we focus on the inspiral stage in our analysis. The leading correction to the phase at the inspiral stage enters at -1PN order due to the scalar dipole radiation and was derived in [269]. Some of the higher PN corrections were recently derived in [119, 289]. Here, we identified even higher PN corrections using the waveforms in scalar-tensor theories [290] (see Appendix 4.A for details of the derivation).

Within the stationary phase approximation [76, 80], the waveform in the Fourier space is given by:

$$h(f) = A(f)\exp\left[i\,\Psi(f)\right]\,, \quad \Psi(f) = \Psi_{\rm GR}(f) + \delta\Psi(f)\,. \tag{4.6}$$

Here $A(f)$ is the amplitude, $\Psi_{\rm GR}$ is the GR phase, and the EdGB correction to the phase $\delta\Psi$ (up to $O(\alpha_{\rm GB}^2)$) is given in a form

$$\delta\Psi = \sum_i \delta\Psi_{i\,\rm PN} = \frac{\alpha_{\rm GB}^2}{M^4} \sum_i c_i\, v^{-5+2i}\,. \tag{4.7}$$

Here $v = (\pi M f)^{1/3}$ is the relative velocity of the binary constituents with GW frequency f and the total mass $M = m_1 + m_2$, where m_1 and m_2 are the masses of the primary and secondary objects of the system. The coefficients c_i up to 2PN order can be found in Appendix 4.A. We note that corrections at 1.5PN and 2PN terms contain terms that have not been computed yet and are thus not fully complete.

4.3 Data Analysis

In this paper, we carry out two independent analyses to find constraints on $\sqrt{\alpha_{\rm GB}}$. The first method is a MCMC analysis based on Bayesian inference by using the publicly-available GW data. The second method is a simpler Fisher analysis that can be used to obtain rough bounds on $\sqrt{\alpha_{\rm GB}}$ to cross check the results from the first method.

Which GW events shall we consider? Since the EdGB corrections to the phase are proportional to $\alpha_{\rm GB}^2/M^4$, such corrections become larger for systems with smaller total masses. If the data is consistent with GR, this translates to a stronger bound on EdGB gravity. Furthermore, the leading scalar dipole radiation is proportional to the square of the difference in the scalar charges between two objects. This means that we expect to find stronger bounds on $\sqrt{\alpha_{\rm GB}}$ for systems with smaller mass ratios ($q = m_2/m_1 < 1$). For these reasons, we will consider the two NSBH events, GW200105 and GW200115, from O3a, (whose total masses are $10.9 M_\odot$ and $7.1 M_\odot$, and mass ratios are 0.22 and 0.26, respectively [107, 298]). We also employ GW190814 [122] whose mass ratio is small (0.11) and the secondary mass is $m_2 \approx 2.6 M_\odot$. The system is consistent with both BBH and NSBH, though the probability of a NS with $2.6\,M_\odot$ may be small [299–301]. Given

the uncertainty in the nature of the secondary object, we consider both possibilities of GW190814 being a BBH and a NSBH. We also use GW151226, a BBH with a relatively small mass, to check our results against those found previously [66, 67].

4.3.1 Bayesian Inference

As reviewed in section 1.2.1, we set $f_{\rm high} = 0.018/M$ [109] which is the approximate maximum frequency at the inspiral stage. Notice that $f_{\rm high}$ is not a fixed number but varies among different MCMC realizations. For our analysis, the parameters are those in GR plus the EdGB coupling constant $\sqrt{\alpha_{\rm GB}}$.

We find posterior distributions on all parameters $\boldsymbol{\vartheta}$ for GW events taken from Gravitational Wave Open Science Center (GWOSC) [302] as follows. We perform MCMC samplings through the PyCBC package [303, 304] and emcee_pt sampler [305] with 500 walkers and 3 temps. We analyze 32 s of data for GW200105 and 64 s of data for GW200115. Regarding the low frequency cutoff, we set $f_{\rm low} = 20\,{\rm Hz}$ except for LIGO Livingston for GW200115, where $f_{\rm low} = 25\,{\rm Hz}$ was used to avoid some excess noise localized at low frequency [107]. Regarding priors, we assume a uniform distribution on $\sqrt{\alpha_{\rm GB}}$ with $[0, 5]\,{\rm km}$ for GW200105, GW200115 and GW190814 (BBH), and $[0, 15]\,{\rm km}$ for GW190814 (NSBH). As for spin priors, we adopt isotropic spin distribution on (θ_A, ϕ_A) with a high-spin prior on magnitude, a_1 and $a_2 \lesssim 0.99$, for all of the MCMC analyses.

For the base waveform model in GR, we adopt IMRPhenomXPHM (that is also used in [121]) from LALSimulation package [83], which is a phenomenological model in the frequency domain that includes spin precession and higher order multipole radiation modes. As the $(l, m) = (3, 3)$ mode is found to be non-negligible for GW200105, GW200115, and GW190814 [107, 122], we include this mode in these events while we only consider the dominant $(l, m) = (2, 2)$ mode for GW151226. We adopt IMRPhenomXPHM model that was constructed for BBHs. As for NSBHs, the tidal effects were found to be negligible [107] for the events considered in this paper, and thus

it is safe to adopt the same waveform model.

4.3.2 Fisher analysis

Fisher analysis method is discussed in section 1.2.7. We follow [306] and impose a Gaussian prior, for simplicity, with a standard deviation $\sigma^{(0)}_{\vartheta^i}$ on each parameter. FIM then becomes

$$\tilde{\Gamma}_{ij} = \frac{1}{\left(\sigma^{(0)}_{\vartheta^i}\right)^2}\delta_{ij} + \Gamma_{ij}\,, \tag{4.8}$$

the standard error can be expressed as:

$$\sqrt{\langle(\delta\vartheta^i)^2\rangle} = \sqrt{\Sigma_{ii}}\,, \quad \Sigma_{ij} = \left(\tilde{\Gamma}^{-1}\right)_{ij}\,. \tag{4.9}$$

Regarding the base waveform in GR, we follow [261] and use IMRPhenomD instead of IMRPhenomXPHM that was used for the Bayesian inference analysis (as explained in Sec. 4.3.1). The former is a simpler version of the latter in the sense that it is valid only for spin-aligned systems (i.e. no spin precession) and includes only the dominant mode. This simplification is justified as we only use the FIM analysis to cross check the results from the Bayesian inference which is more robust. Moreover, Perkins et al. [67] showed that the difference in the waveform models between IMRPhenomPv2 (a precessing model similar to IMRPhenomXPHM but only includes the dominant mode) and IMRPhenomD changes the bound on $\sqrt{\alpha_{\rm GB}}$ only by $\sim 20\%$. For simplicity, we use a sky-averaged waveform (and rescale the amplitude so that the SNR matches with the observed one) and the parameters for this second method are as follows:

$$\vartheta = \left(\mathcal{M}, q, a_1, a_2, \phi_{\rm ref}, t_c, D_L, \alpha^2_{\rm GB}\right)\,. \tag{4.10}$$

Notice that we take $\alpha^2_{\rm GB}$ as our EdGB parameter instead of $\sqrt{\alpha_{\rm GB}}$. This is because the former is

	GW200105	GW200115	GW190814		GW151226	combined
	NSBH	NSBH	NSBH	BBH	BBH	
Fisher	1.55	0.91	7.39	0.90	4.19 (2.51 [67])	0.59
Bayesian	1.90	1.33	2.72	0.37 (0.4 [121])	3.43 (4.4 [67])	1.18
small coupl. limit	4.40	2.94	11.4	1.27	3.81	—

Table 4.2: Constraints on $\sqrt{\alpha_{\rm GB}}$ [km] at 90% credible level with Fisher analysis and Bayesian inference from selected NSBH and BBH events. For GW190814, we consider both NSBH and BBH possibilities due to the uncertainty in the nature of the secondary object. These constraints are derived by using the leading phase correction at -1PN order, which are improved by approximately 7–15% if we include higher PN corrections. Our results for GW190814 (BBH) and GW151226 are consistent with those found in previous work shown in brackets. The last column shows the bound by combining posteriors from GW200105, GW200115, GW190814 (NSBH), and the combined posterior from selected BBHs from GWTC-1 and GWTC-2 catalogs obtained in [67]. The last row shows the upper limits on $\sqrt{\alpha_{\rm GB}}$ that is valid within the small coupling approximation (Eq. (4.11)). Observe that all the bounds from the Fisher and Bayesian analyses are within these upper limits, showing the validity of our results.

what enters directly in the waveform and if one chooses to use the latter, the Fisher matrix becomes singular when we take the fiducial value as $\alpha_{\rm GB} = 0$ (for the fiducial values of other parameters, we use those reported by LVC and set $\phi_{\rm ref} = t_c = 0$). We impose a Gaussian prior [306] with the standard deviation of $\sigma_{a_1} = \sigma_{a_2} = 1$ and $\sigma_{\phi_{\rm ref}} = \pi$.

4.4 Results

4.4.1 Leading Correction

We now present our results. Constraints on $\sqrt{\alpha_{\rm GB}}$ from various GW events with Bayesian and Fisher analyses are summarized in Table 4.2[1]. Here, we only included the leading -1PN correction to the waveform phase. Observe that the bounds from the two analyses for each GW event agree within a factor of ~ 3. Since the phase corrections are derived within the small coupling approximation, we need to check whether the bounds presented here satisfies this approximation. Following [67], we require

$$16\pi\frac{\alpha_{\rm GB}^2}{m^4} \leq 0.5\,, \tag{4.11}$$

where m is the smallest length scale in the binary. We choose $m = m_2$ (the mass of the smaller BH) for BBH while $m = m_1$ (the mass of the BH) for NSBH[2]. We present in Table 4.2 the upper limit on $\sqrt{\alpha_{\rm GB}}$ that satisfies the above bound. Notice that all the Fisher and Bayesian bounds satisfy the small coupling approximation and thus are reliable. Notice also that our Fisher and Bayesian results for GW151226 and GW190814 (BBH) are consistent with those in [67, 121][3]. Our results are also roughly consistent with the forecast made in [286] for bounds on $\sqrt{\alpha_{\rm GB}}$ with NSBHs derived through a Fisher analysis. For example, the bound for a BH mass of $8M_\odot$ and an SNR of 8 (similar to GW200115 where the BH mass is $5.7M_\odot$ and an SNR of 11.4 [107]) was found to be $\sqrt{\alpha_{\rm GB}} \lesssim 0.4$km with advanced LIGO's design sensitivity which has a slightly different shape for the noise curve than that with O3 detectors.

The most stringent constraint comes from GW190814 (BBH) though the event is still consistent

[1] Notice that there are some differences in Bayesian and Fisher analyses, such as the waveform modeling (PhenomXPHM vs IMRPhenomD), sGB parameter ($\sqrt{\alpha}_{\rm sGB}$ vs $\alpha^2_{\rm sGB}$) and its prior (uniform vs Gaussian). This may explain why Fisher bounds are weaker than the Bayesian ones in some cases.

[2] For simplicity, we use the mass estimates found by LVC assuming GR while Ref. [67] used the median values of the masses from posterior distributions including $\sqrt{\alpha_{\rm GB}}$.

[3] Perhaps a small discrepancy in the results for GW190814 (BBH) is due to the fact that we vary the coalescence time t_c in our Bayesian inference while it seems that Ref. [121] fixed this parameter (at least the posterior distribution on this parameter is not shown in Appendix A of [121]).

with NSBH and thus such a bound may not be robust. The reason why the bound on $\sqrt{\alpha_{\rm GB}}$ is stronger for BBH than NSBH for GW190814 can be understood as follows. First, notice that the leading correction to the phase is proportional to $(m_1^2 s_2 - m_2^2 s_1)^2/M^4$ (see Eq. (4.18)). Second, let us consider the case $m_1 \gg m_2$ for simplicity. In this case, we find $c_{-1} \propto 1$ for BBH while $c_{-1} \propto q^4$ for NSBH (the scalar charge s_2 is 0 for a NS). Thus, the EdGB correction can be much larger for BBH than NSBH.

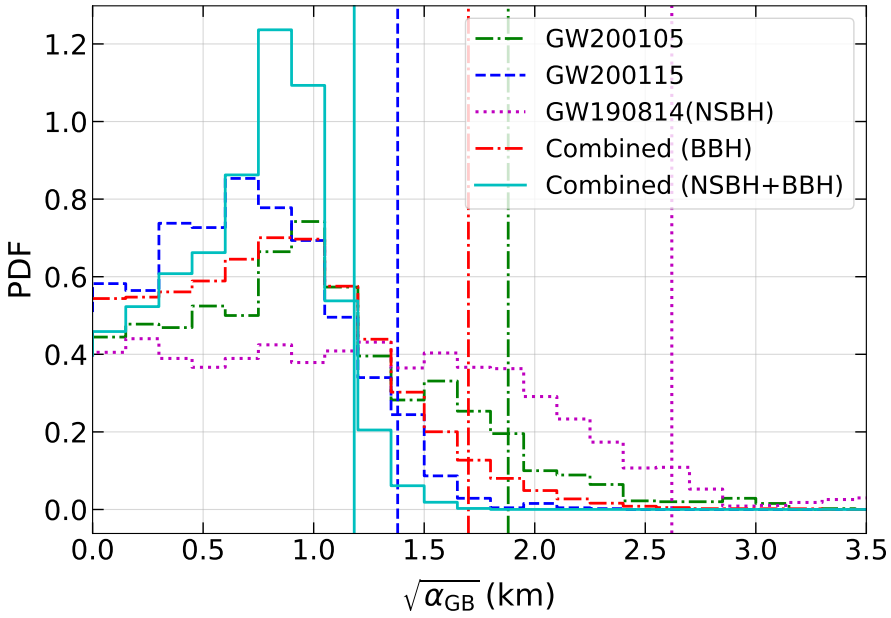

Figure 4.1: Posterior probability distributions for $\sqrt{\alpha_{\rm GB}}$ from selected GW events. We also show an upper bound on $\sqrt{\alpha_{\rm GB}}$ at 90% credible level for each event as vertical lines, which indicates the result is consistent with GR. The posteriors are found by including only the leading EdGB correction to the phase at -1PN order.

Besides constraints from the events GW151226 and GW190814 (BBH) which have already been derived in the previous works [67, 121], we here derived bounds from NSBHs (GW200105, GW200115, and GW190814) for the first time. We present the posterior distributions for $\sqrt{\alpha_{\rm GB}}$ for these events in Fig. 4.1. The bound from GW200115 is $\sqrt{\alpha_{\rm GB}} \lesssim 1.33\,$km, which is stronger than the

bound obtained in [67] by stacking several BBHs from GWTC-1 and GWTC-2 catalogs ($\sqrt{\alpha_{\rm GB}} \lesssim$ 1.7 km). Observe that the posterior distributions are quite different from Gaussian centered at $\sqrt{\alpha_{\rm GB}} = 0$, which partially explains the difference between the Fisher and Bayesian results (see also TABLE II, FIG. 2, and FIG. 3 in [67]).

Furthermore, we derive combined bounds by multiplying normalized posterior histograms on $\sqrt{\alpha_{\rm GB}}$[4] from GW200105, GW200115, GW190814 (with the NSBH assumption that gives us a more conservative bound), and combined BBH bounds in [67]. We found a stringent bound of $\sqrt{\alpha_{\rm GB}} \lesssim 1.18$ km through the Bayesian analysis as shown in Table 4.2 and Fig. 4.1.

4.4.2 Effects of Higher PN Corrections

We next study the effect of higher PN corrections to the waveform phase by including PN corrections up to 2PN as presented in Appendix 4.A. Perkins et al. [67] carried out a similar analysis though such higher PN corrections were not available at that time. Thus, the authors considered three different ways to parameterize the unknown 0PN correction (which is 1PN higher than the leading -1PN correction) based on the functional forms at 1PN order in GR and the leading -1PN EdGB corrections. They then marginalized over such a parameter and concluded that higher PN corrections do not affect the results much and the bounds derived with the leading correction are robust. We check this outcome by using explicit forms of the higher PN corrections in EdGB gravity.

Figure 4.2 presents posteriors on $\sqrt{\alpha_{\rm GB}}$ for GW200105 and GW200115 with and without higher PN corrections, while Fig. 4.3 shows corresponding corner plots on $\sqrt{\alpha_{\rm GB}}$, $\mathcal{M}$ and q. Notice that the inclusion of the higher PN corrections does not affect the posteriors much, especially for GW200115. The 90% credible upper bound on $\sqrt{\alpha_{\rm GB}}$ improves from the case with the leading correction by 14.5% for GW200105 and 6.9% for GW200115 respectively. These findings are

[4]This corresponds to the second method discussed in Sec. IIIE of [67] for obtaining combined bounds.

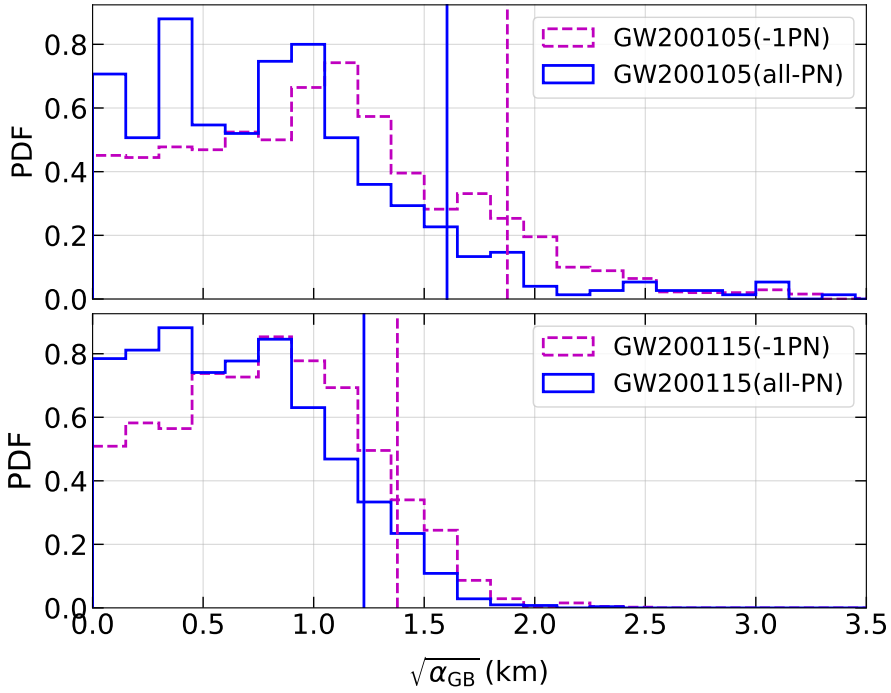

Figure 4.2: A comparison of the posteriors on $\sqrt{\alpha_{\rm GB}}$ from the leading -1PN correction and those including higher PN corrections (up to 2PN) for GW200105 (top) and GW200115 (bottom). Observe that the 90% upper bounds on $\sqrt{\alpha}_{\rm GB}$ are improved by 14.5% for GW200105 and 6.9% for GW200115 respectively.

consistent with those in [67] and a very recent work [307] that investigated the improvement one obtains when including higher PN order terms.

4.5 Conclusions and Discussion

In this paper, we derived bounds on EdGB gravity using GWs from NSBH binaries. Using the leading PN correction, we found $\sqrt{\alpha_{\rm GB}} \lesssim 1.33\,\text{km}$ as a 90% credible limit from GW200115, which is stronger than the bound in [67] found by combining selected BBHs from GWTC-1 and GWTC-

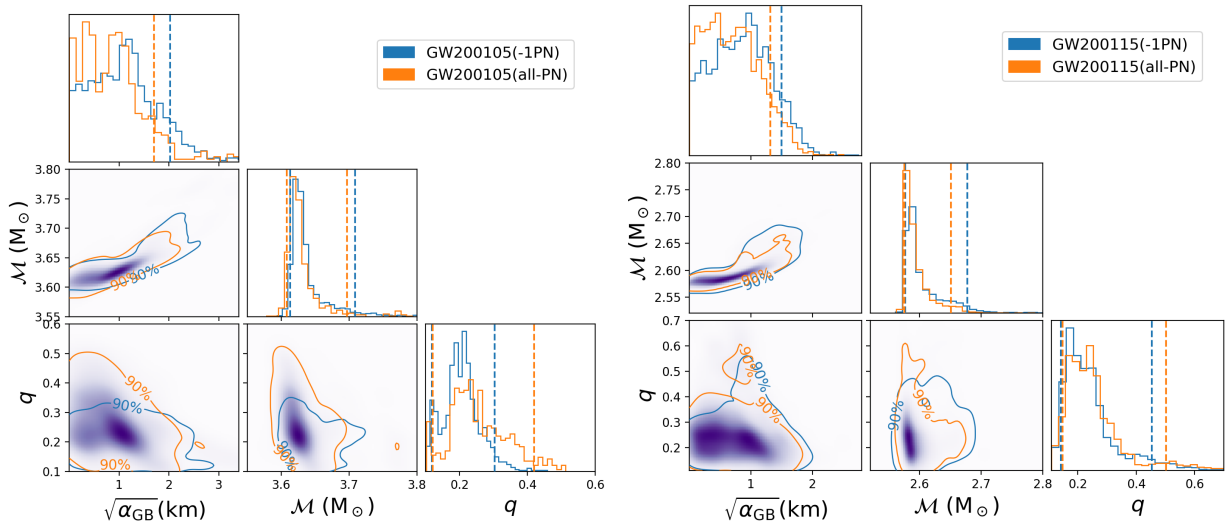

Figure 4.3: Posterior probability distributions for the EdGB coupling constant $\sqrt{\alpha_{\rm GB}}$, the chirp mass $\mathcal{M}$, and the mass ratio q from GW200105 (left) and GW200115 (right). We compare the marginal posterior distributions for the case with the leading EdGB correction at -1PN order (blue) and the case including higher PN orders up to 2PN (orange). The purple shaded regions indicate the posterior probabilities of the latter case and the solid lines represent the 90% credible regions for the two cases. The vertical dashed lines show the one-sided 90% confidence interval for $\sqrt{\alpha_{\rm GB}}$ and the two-sided 90% credible intervals for $\mathcal{M}$ and q.

2 catalogs. We also derived combined bounds by stacking posterior distributions on $\sqrt{\alpha_{\rm GB}}$ from GW200105, GW200115, GW190814 and the combined posteriors from selected BBHs in [67], and found $\sqrt{\alpha_{\rm GB}} \lesssim 1.18\,\text{km}$. We further derived higher PN corrections in the waveform phase up to 2PN order from the results in scalar-tensor theories [290]. Using these, we improved bounds on $\sqrt{\alpha_{\rm GB}}$ for GW200105 and GW200115 from the case with leading PN correction alone by 14.5% and 6.9% respectively.

The analysis carried out here can easily be extended to probe other theories of gravity. We looked at constraining dynamical Chern-Simons gravity [308], which is a parity-violating quadratic gravity whose leading PN correction to the phase is derived in [309]. Similar to the case with BBHs [66, 67, 261], we were not able to find meaningful bounds that satisfy the small coupling approxi-

mation. For future work, one could consider e.g. sGB gravity with the coupling function $f(\phi) \propto \phi^2$ or $f(\phi) \propto 1 - e^{-6\phi^2}$ that admits spontaneous scalarization of BHs [280, 282].

4.A Appendix: EdGB Corrections to Gravitational Waveforms

In this appendix, we explain how to map the waveform (for non-spinning BBHs) in scalar-tensor theories [290] to that in EdGB gravity. The former is valid to 2PN order higher than the leading for each of tensor and scalar emission.

The waveform in scalar-tensor theories is derived in the Jordan frame while EdGB gravity is in the Einstein frame. Therefore, we first turn the former into the Einstein frame. This can be realized by using the mapping provided in Appendix A of [290]. After this transformation, the waveform is given in terms of the scalar charge α_A and its derivative β_A for the Ath body.

The next step is to find these charges in EdGB gravity and substitute this into the waveform. We can compute these following [310] which uses a slightly different convention for sGB gravity:

$$S = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\Big[R - 2(\nabla\varphi)^2 + \alpha_{\rm GB}\bar{f}(\varphi)\mathcal{R}^2_{\rm GB}\Big] + S_m\,. \tag{4.12}$$

One can perform the following rescaling in the scalar field φ and the identification of the function $\bar{f}(\varphi)$ to recover the action in Eq. (4.5):

$$\bar{f}(\varphi) = 2\sqrt{16\pi}\varphi\,, \quad \varphi = \frac{\sqrt{16\pi}}{2}\phi\,. \tag{4.13}$$

From this, α_A and β_A for a non-rotating BH to leading order in $\alpha_{\rm GB}$ are given by:

$$\alpha_A^{\rm BH} = -\frac{\alpha_{\rm GB}\bar{f}'(\varphi_0)}{2m_A^2} = -\frac{\sqrt{16\pi}\alpha_{\rm GB}}{m_A^2}\,, \tag{4.14}$$

$$\beta_A^{\rm BH} = -\frac{\alpha_{\rm GB}^2\bar{f}'(\varphi_0)^2}{2m_A^2} = -\frac{32\pi\alpha_{\rm GB}^2}{m_A^2}\,, \tag{4.15}$$

where φ_0 is the asymptotic value of the scalar field φ at infinity. When substituting these into the waveform expression, the terms with β_A enter at $O(\alpha_{\rm GB}^4)$ and are negligible. For α_A, we add the spin dependence as:

$$\alpha_A^{\rm BH} = -\frac{\sqrt{16\pi}s_A\alpha_{\rm GB}}{m_A^2}\,, \tag{4.16}$$

where the spin dependent factor is given by: [261, 294]

$$s_A = 2\frac{\sqrt{1-\chi_A^2}-1+\chi_A^2}{\chi_A^2}\,. \tag{4.17}$$

This reduces to $s_A^{\rm BH} \to 1$ in the limit $\chi_A \to 0$. For NSs, $\alpha_A^{\rm NS} = O(\alpha_{\rm GB}^3)$ and is negligible while $\beta_A^{\rm NS}$ has not been computed. Though we expect the $\alpha_{\rm GB}$ dependence to be the same as BH and ignore such terms in the waveform.

Using these charge expressions in the dominant harmonics ($\ell = m = 2$) of the waveform and keeping only to $O(\alpha_{\rm GB}^2)$, EdGB corrections to the waveform can be expressed as in Eq. (4.7) with the coefficients given as follows:

$$c_{-1} = -\frac{5\pi}{448}\frac{(m_1^2 s_2 - m_2^2 s_1)^2}{\eta^5 M^4}\,, \tag{4.18}$$

$$c_0 = -\frac{5\pi}{43008}\frac{(659+728\eta)(m_1^2 s_2 - m_2^2 s_1)^2}{\eta^5 M^4} + \frac{5\pi}{8600}\frac{s_1 s_2}{\eta^3}\,, \tag{4.19}$$

$$c_{0.5} = \frac{75\pi^2}{448}\frac{(m_1^2 s_2 - m_2^2 s_1)^2}{\eta^5 M^4}\,, \tag{4.20}$$

$$c_1 = -\frac{5\pi}{48384}\frac{(m_1^2 s_2 + m_2^2 s_1)^2\,(535+924\eta)}{\eta^5 M^4} - \frac{5\pi}{2016}\frac{s_1 s_2(743+924\eta)}{\eta^3}$$

$$- \frac{25\pi}{576}\frac{(m_1^2 s_2 - m_2^2 s_1)^2}{\eta^5 M^4}\left[\frac{12497995}{1016064} - \frac{11(m_1-m_2)(m_1^2 s_2 + m_2^2 s_1)}{2M(m_1^2 s_2 - m_2^2 s_1)} + \frac{15407\eta}{1440} + \frac{165\eta^2}{16}\right], \tag{4.21}$$

$$c_{1.5} = \frac{\pi^2}{2}\frac{(m_1^2 s_2 - m_2^2 s_1)^2}{\eta^5 M^4} - \frac{3 f_3^{\rm GB}}{32\eta}\,, \tag{4.22}$$

$$\begin{aligned} c_2 = \frac{5\pi}{32514048}\frac{1}{\eta^5 M^5}\Big[&(m_1^5 s_2^2 + m_2^5 s_1^2)(-4341025 + 65553264\eta - 684432\eta^2) \\ &+ \eta M^2 (m_1^3 s_2^2 + m_2^3 s_1^2)(20044511 + 65553264\eta - 684432\eta^2) \\ &+ 42\eta^2 M^5 s_1 s_2 (1029619 - 36387504\eta - 7970256\eta^2)\Big] - \frac{15 f_4^{\rm GB}}{64\eta}. \end{aligned} \tag{4.23}$$

Here $\eta \equiv m_1 m_2/M^2$ is the symmetric mass ratio while $f_3^{\rm GB}$ and $f_4^{\rm GB}$ represent our ignorance of the correction to the tensor non-dipole emission in EdGB gravity at 1.5PN and 2PN orders[5]. The above corrections can be mapped to the parameterized post-Einsteinian (PPE) framework [87, 88, 120] of

$$\delta\Psi = \sum_i \beta_i^{\rm PPE} v^{-5+2i}\,, \tag{4.24}$$

with

$$\beta_i^{\rm PPE} = \frac{\alpha_{\rm GB}^2}{M^4} c_i\,. \tag{4.25}$$

The leading -1PN term (c_{-1} or $\beta_{-1}^{\rm PPE}$) derived here agrees with those found in [261, 272].

Figure 4.4 presents each PN correction term in the phase against the GW frequency f for GW200115, together with the leading GR term. We chose $\sqrt{\alpha_{\rm GB}} = 1.33$km that is the 90% credible limit found through our Bayesian inference in Table 4.2. Notice that the EdGB corrections are subdominant to GR by at least an order of magnitude. Notice also that the leading EdGB correction

[5]We have replaced $f_i^{\rm ST}$ in [290] to $(\alpha_{\rm GB}^2/M^4) f_i^{\rm GB}$ for $i = 3, 4$.

at -1PN order dominates higher PN contributions at $f \lesssim 200$Hz and the latter becomes only important when the frequency becomes high (though the noise becomes larger as the frequency becomes higher), which explains why higher PN corrections do not affect the bound on $\sqrt{\alpha_{\rm GB}}$ much. It is interesting to note that for $f \gtrsim 200$Hz, the EdGB phase is dominated by the contribution at 1.5PN order, though the phase is still incomplete at this order (we have set the unknown contributions $f_3^{\rm GB}$ and $f_4^{\rm GB}$ to 0 in Fig. 4.4).

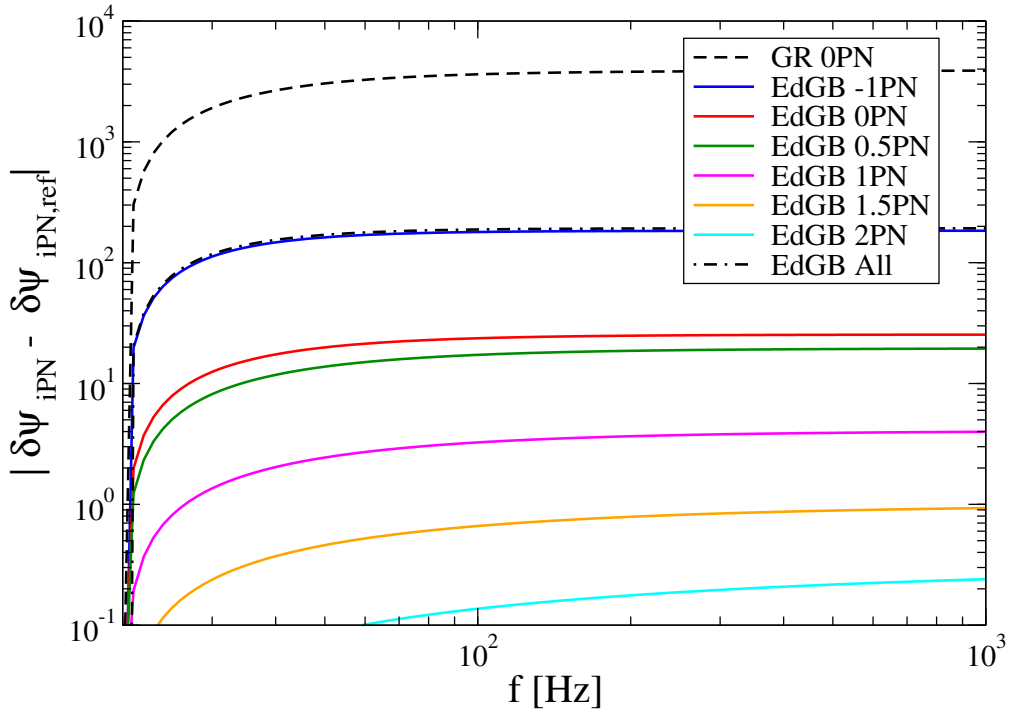

Figure 4.4: Comparison of EdGB corrections to the phase entering at different PN orders as a function of the GW frequency. We also present the leading phase in GR and the contribution from all of the EdGB corrections combined. For each contribution, we show the phase relative to that at a reference frequency chosen to be 20Hz. We chose $(m_1, m_2) = (5.9, 1.4)M_\odot$ and $(\chi_1, \chi_2) = (0.31, 0)$, corresponding to GW200115, and $\sqrt{\alpha_{\rm GB}} = 1.33\,$km that is the 90% credible limit found through our Bayesian inference (see Table 4.2).

Let us comment on up to which PN order the above waveform corrections are complete. The $\alpha_{\rm GB}$ dependence in the above corrections enter only through the scalar charges α_A. There are other

contributions to the waveform where $\alpha_{\rm GB}$ appears explicitly though such contributions enter at 3PN order and are negligible for our purpose [6]. For non-spinning binaries, they are complete up to 1PN order. The expressions at 1.5PN and 2PN include currently unknown $f_3^{\rm GB}$ and $f_4^{\rm GB}$ but they also have other missing contributions, such as the scalar dipole radiation at 1.5PN and 2PN orders (which correspond to 2.5PN and 3PN relative to the leading -1PN contribution) and the correction to the binding energy or Kepler's law at 3PN that couples to the -1PN dipole radiation and enter at 2PN in the waveform. For spinning binaries, the waveform is complete only up to 0PN order as the effect of spins are only included through the scalar charges α_A. Missing contributions include e.g. a spin-orbital coupling in the binding energy at 1.5PN order that couples with the leading dipole radiation to enter at 0.5PN in the waveform.

We end by comparing the 0PN corrections found here with different functional forms considered in [67]. Using Eqs. (4.18) and (4.19), the 0PN correction to the phase can be expressed as:

$$\delta\Psi_{\rm 0PN} = \frac{659+728\eta}{96} v^2 \delta\Psi_{\rm -1PN} + \frac{5\pi}{16}\frac{s_1 s_2}{\eta^3}\frac{\alpha_{\rm GB}^2}{M^4} v^{-5} \,. \tag{4.26}$$

The first term is similar to one of the functional forms considered in [67]:

$$\delta\Psi_{\rm 0PN}^{\rm (PNSY,1)} = \frac{5}{756}(743+924\eta)\gamma\, u^2\, \delta\Psi_{\rm -1PN}\,, \tag{4.27}$$

where γ is a constant that does not depend on binary parameters, $u \equiv (\pi\mathcal{M}f)^{1/3}$ and the η dependence is taken from that in the phase at 1PN order in GR. The η dependence in the two expressions,however, are different. The second term in Eq. (4.26) is similar to another functional form considered in [67]:

$$\Psi_{\rm 0PN}^{\rm (PNSY,2)} = 16\pi\frac{\alpha_{\rm GB}^2}{M^4}\gamma u^{-5}\,, \tag{4.28}$$

though again, the expressions are different. This is because if one maps the second term in Eq.

[6] We count the PN order in powers of v/c while Shilarirou et al. [119, 289] counts in powers of $1/c$. With the latter counting, the $\alpha_{\rm GB}$ dependence other than scalar charges enters at 1PN.

(4.26) to Eq. (4.28), γ depends on binary parameters through η and s_A.

Part II

The Hybrid Waveform for Compact Binaries

Chapter 5

The Hybrid Waveform

Xuefeng Feng, Zhenwei Lyu, Huan Yang

accepted by PRD, arXiv:2104.11848

My specific contributions: Calculate all tidal phases of EOB waveform model and numerical relativity simulation in section 5.4.2. Discuss neutron star tidal effect and equation of state with Xuefeng.

We consider the motion of nonspinning, compact objects orbiting around a Kerr black hole with tidal couplings. The tide-induced quadrupole moment modifies both the orbital energy and outgoing fluxes, so that over the inspiral timescale there is an accumulative shift in the orbital and gravitational wave phase. Previous studies on compact object tidal effects have been carried out in the Post-Newtonian (PN) and Effective-One-Body (EOB) formalisms. In this work, within the black hole perturbation framework, we propose to characterize the tidal influence in the expansion of mass ratios, while higher-order PN corrections are naturally included. For the equatorial and circular orbit, we derive the leading order, frequency dependent tidal phase shift which agrees with the Post-Newtonian result at low frequencies but deviates at high frequencies. We also find

that such phase shift has weak dependence ($\leq 10\%$) on the spin of the primary black hole. Combining this black hole perturbation waveform with the Post-Newtonian waveform, we propose a frequency-domain, hybrid waveform that shows almost comparable accuracy as the EOB waveform in characterizing the tidal effects, in a comparison with limited number of numerical relativity simulations. Further improvement is expected as the next-to-leading order in mass ratio and the higher-PN tidal corrections are included. This hybrid approach is also applicable for generating binary black hole waveforms.

5.1 Introduction

Inspiraling and coalescing compact-object binary systems, including black holes and/or neutron stars, are important sources of ground-based gravitational waves (GW) detectors, e.g. LIGO[311] and Virgo[312]. Up to the O3 observation run, Advanced LIGO and Virgo have detected more than thirty binary black hole mergers, two binary neutron star mergers and one possible black hole-neutron star merger. The number of events is expected to increase significantly as Advanced LIGO and Virgo reach their design sensitivities.

Constructing GW waveform models is crucial for efficiently detecting these binary systems as well as accurately estimating their source properties based on the observation data. Since it is computationally expensive to numerically solve Einstein's equation (and associated hydrodynamical equations if a neutron star is involved) for the binary evolution across the entire observation frequency band, especially with the large parameter space needed to characterize these binaries, several (semi)-analytical or phenomenological methods[313–315] have been developed to complement the information from numerical simulations and generate reliable waveforms [316–319].

These methods generally follow different avenues of analytical approximations in modelling the binary black hole inspiral waveform. For example, the low-frequency inspiral dynamics and associated waveform are treated within the Post-Newtonian (PN) framework in the "Phenom" wave-

form series [316, 320]. At higher frequency certain calibrations with numerical waveforms are performed to bridge the gap between the PN inspiral description with the black hole ringdown. On the other hand, the PN expansion is restructured in the Effective-One-Body formalism [123] through a mapping to an effective spacetime of the relative motion, so that the resumed PN results may be better attached to the strong-gravity regime. Calibration with numerical relativity data has also been used to improve the accuracy of Effective-One-Body (EOB) waveforms.

When the mass ratio between the secondary and the primary black hole is small, we can view the smaller black hole as a particle moving in a perturbed spacetime of the primary black hole, where the metric perturbation and associated dynamical effects can be evaluated in a systematic expansion in the mass ratio. This black-hole-perturbation approach is the leading solution to produce waveforms of extreme mass-ratio inspirals (EMRIs), which are important sources for space-borne GW detectors such as LISA [321]. Given this expansion scheme, it is then natural to ask what is its regime of applicability in mass ratios? Interestingly, recent studies [319, 322–334] on this question have revealed a rather surprising result: the EMRI-based waveform may be even applicable for equal-mass binaries. In particular, for the equatorial and circular orbit, the GW phase can be written as the post-adiabatic expansion[319]

$$\psi(\omega) = \frac{\psi_{-1}(\omega)}{\eta} + \psi_0(\omega) + \eta\psi_1(\omega) + \dots , \tag{5.1}$$

where ω is the orbital angular frequency and $\eta = m_1 m_2/(m_1 + m_2)^2$ is the symmetric mass ratio, the function $\psi_i(\omega)$ is the coefficient of the order η^i term. When the mass ratio is extreme, the symmetric mass ratio is almost the same as the mass ratio $q = m_1/m_2 \leq 1$. The comparison with numerical relativity waveforms shows that, across the entire inspiral frequency range, high order terms (starting from ψ_2 in the expansion) only contribute ≤ 2 radians phase shift even for equal-mass black hole binaries (with $\eta = 1/4$) for most of the frequency range, except near the transition

regime from inspiral to plunge [1]. This observation indicates that Eq. (5.1) may be a fast-converging series even for equal-mass binaries, so that the first several terms may suffice to produce accurate waveforms.

In this work, we adopt the black hole perturbation point of view, and evaluate the induced quadrupole moment of a neutron star moving in a perturbed spacetime of the primary black hole. In the local rest frame (or more precisely, within the "asymptotically Cartesian and mass centered" coordinates [336] [2]) of the neutron star.

In the black hole perturbation picture, the metric perturbation generated by the less massive black hole can be expanded in power laws of the mass ratio $h = h_1 q + h_2 q^2 + ...$, with $q = m_1/m_2$, and the less massive black hole m_1 can be viewed as moving along geodesics of the spacetime with metric $g_{\rm Kerr} + h$ [338]. This mass ratio expansion justifies the mass ratio expansion of ψ in Eq. (5.1). When the less massive object is a neutron star, its motion can be viewed as a perturbed geodesic of the spacetime $g_{\rm Kerr} + h$. This deviation from geodesic mainly comes from multipole interaction between the star and its environmental tidal field, while h is sourced by the monopole ("the point-mass" piece), quadrupole, and all higher order multipole parts of the stress-energy tensor. For simplicity, we truncate the multipole expansion at the quadrupole order and use the Mathisson-Papapetrou-Dixon prescription [339] to construct the stress-energy tensor of the star. To the linear order in λ, the tidal energy of the object and the tidal induced gravitational radiation flux are all $1/q$ or $1/\eta$ order lower than those of a point mass, so that the correction to the gravitational phase starts at q^{-2} or η^{-2} order. Both q and η are eligible choices of expansion parameters in the small mass ratio limit, but they will give rise to rather different result as we truncate the series and apply it in the comparable mass ratio limit. For binary black hole waveforms it seems η is a more efficient expansion parameter [319], but for tidal corrections the optimal choice is yet to be determined.

[1] It is expected that an additional correction of order $\eta^{-1/5}$ must be introduced to account for the transition effects [124, 335].

[2] In the multipole expansion picture discussed in [337], the central object can be fully relativistic. As the multipole moments are derived in the asymptotic zone, Eq. (1.44) can be viewed as the definition for the relativistic Love number λ.

As decribrd in section 1.6, if the companion is also a neutron star, its tidal contribution to the waveform can be obtained by replacing q by q^{-1}, λ_1 by λ_2 and keeping η to be the same in Eq. (1.54). As a result, the total tidal correction is

$$\begin{aligned}\eta\psi_{\rm BP} =&\eta\psi_{\rm BP,1} + \eta\psi_{\rm BP,2}\\ =&\lambda_1(q^{-1}\psi_{\rm BP}^{(-1)} + \psi_{\rm BP}^{(0)} + \sum_{n\geq 1}\eta^n\psi_{\rm BP}^{(n)})\\ &+ \lambda_2(q\psi_{\rm BP}^{(-1)} + \psi_{\rm BP}^{(0)} + \sum_{n\geq 1}\eta^n\psi_{\rm BP}^{(n)})\,. \end{aligned} \tag{5.2}$$

Strictly speaking, if both compact objects are neutron stars, there is no horizon absorption of the gravitational wave flux. Such effect enters the dynamics at 2.5 relative PN order for rotating black holes and 4PN for non-rotating black holes [340]. The overall contribution to the phase is less than 0.1 rad for the point-mass motion terms, which means for the tidal correction it should be even smaller[341]. We shall neglect this effect in the waveform construction. Notice that $\psi_{\rm BP}^{(0)}(\omega)$ becomes the leading order term for star "2". In fact, it can be evaluated by computing the deformation of a star by an orbiting point mass, and then determining the extra energy change and gravitational wave flux due to the star deformation. This offers an alternative (and likely easier) way to compute $\psi_{\rm BP}^{(0)}(\omega)$. Such calculation will be carried out in a follow-up study of this work, where more systematic comparison with numerical waveforms will be provided.

The paper is organized as follows. In Section II, we derive the explicit equations of motion of an extended body with nonzero quadrupole moment moving on a circular and equatorial orbit in the Kerr spacetime. A series of conserved quantities discussed here. In Section III, we review the Teukolsky formalism where the asymptotic behavior of the homogeneous solution, waveforms and fluxes, and the quadrupole source term are shown. In Section IV, we construct the hybrid waveform and compare it with numerical relativity waveforms, as well as the EOB waveform. We summarize in Section V. Throughout this paper, we adopt geometrical units, $G = c = 1$, where G denotes the

gravitational constant and c the speed of light, respectively. The metric signature is $(-,+,+,+)$

5.2 Conservative orbital motion

In this section, we consider a nonspinning body (with nonzero quadrupolar moment) moving in the Kerr spacetime, focusing on the case of circular, equatorial orbits. Without including the gravitational radiation reaction, the orbital motion is conservative and easily solvable. We focus on the conservative piece of motion in this section, and leave the discussion on radiative effects to Sec. 5.3.

The Boyer-Lindquist coordinates (t, r, θ, ϕ) are used in the analysis, in which the Kerr metric takes the following form:

$$\begin{aligned} ds^2 &= -\left(1-\frac{2Mr}{\Sigma}\right)dt^2 - \frac{4aMr\sin^2\theta}{\Sigma}\,dtd\phi \\ &\quad +\frac{\Sigma}{\Delta}\,dr^2 + \Sigma\,d\theta^2 \\ &\quad +\sin^2\theta\left(r^2+a^2+\frac{2a^2Mr\sin^2\theta}{\Sigma}\right)d\phi^2, \end{aligned} \tag{5.3}$$

where M is the mass of black hole, a is the spin parameter with $|a| \leq M$, and

$$\Delta = r^2 - 2Mr + a^2, \quad \Sigma = r^2 + a^2\cos^2\theta. \tag{5.4}$$

The Kerr spacetime has two Killing vector fields given by ∂_t and ∂_ϕ.

5.2.1 Equations of motion

The motion of a test body with multipolar structure is discussed in detail in [342]. Following the same formalism, considering the influence of quadrupole moment-curvature coupling, the equation

of motion of a spinning extended body reads

$$\frac{Dp_a}{d\tau} = -\frac{1}{2}R_{abcd}u^b S^{cd} - \frac{1}{6}\nabla_a R_{bcde}J^{bcde}, \tag{5.5}$$

$$\frac{DS^{ab}}{d\tau} = 2p^{[a}u^{b]} + \frac{4}{3}R^{[a}{}_{cde}J^{b]cde} \tag{5.6}$$

where u^a denotes the 4-velocity of the body along its world line (normalized to $u^a u_a = 1$), τ is an affine parameter of the orbit, R_{abcd} denotes the Riemann tensor of a Kerr spacetime, p^a is the momentum, and J^{abcd} is the quadrupole tensor which obeys the following symmetries:

$$J^{abcd} = J^{[ab][cd]} = J^{cdab}, \tag{5.7}$$

$$J^{abcd} + J^{bcad} + J^{cabd} = 0. \tag{5.8}$$

If we only consider the gravito-electric tidal field, neglecting the gravito-magnetic tidal field and quadrupole deformations induced by the spin, the induced quadrupole moment is:

$$Q^{ab} = -\lambda E^{ab}, \tag{5.9}$$

where λ is the tidal Love number and $E_{ab} = \frac{1}{m_1{}^2}R_{acbd}p^c p^d$ is the tidal tensor of the spacetime. In addition, the tidal quadrupole deformations J^{abcd} is related to Q^{ab} by

$$J^{abcd} = -\frac{3m_0}{m_1^3}p^{[a}Q^{b][c}p^{d]}, \tag{5.10}$$

where

$$m_1^2 = -p^a p_a\,,$$

$$m_0 = -p_a u^a\,,$$

In this paper, we suppose that the extended body has no spin, then the 4-momentum can be obtained from (5.6):

$$p^a = m_0 u^a + \frac{4}{3} u_b R^{[a}{}_{cde} J^{b]cde}. \tag{5.11}$$

The difference between m_0 and m_1 is at higher multipole order than the quadrupole[342]. As a result, we shall not distinguish m_0 from m_1 in this work, as we only consider effects by the quadrupole moment. The stress-energy tensor of the test body can be written in the following form:

$$\begin{aligned} T^{ab} = \quad & \int d\tau \Big[u^{(a} p^{b)} \frac{\delta_{(4)}}{\sqrt{-g}} + \frac{1}{3} R_{cde}{}^{(a} J^{b)edc} \frac{\delta_{(4)}}{\sqrt{-g}} \\ & - \frac{2}{3} \nabla_d \nabla_c (J^{d(ab)c} \frac{\delta_{(4)}}{\sqrt{-g}}) \Big]. \end{aligned} \tag{5.12}$$

5.2.2 Conserved Quantities

A test particle moving in the Kerr spacetime has four conserved quantities: energy, angular momentum along the symmetry axis, the Carter constant and its rest mass. As a result, its motion is integrable for generic geodesic orbits. When the internal quadrupole moment is included, we can still construct conserved quantities for extended bodies in the Kerr spacetime based on the Killing vector fields. According to [343, 344], the quantity

$$Q_\xi = p_a \xi^a \tag{5.13}$$

is conserved if ξ^a is a Killing vector, $\nabla_{(b}\xi_{a)} = 0$. We then decompose energy and angular momentum as $E = E_0(r) + E_t(r)$ and $J = J_0(r) + J_t(r)$, where $E_t(r)$ and $J_t(r)$ are proportional to the tidal Love number λ. As only the first order tidal effects included, we just need to substitute $E_0(r)$ and $J_0(r)$ into Eq. (5.10) to obtain the momentum p and quadrupole moment J^{dabc} .For the Kerr

spacetime, there are two killing fields $\partial_t, \partial_\phi$, which lead to

$$\begin{aligned} E =& - p_t \\ =& \frac{1}{r}(2Mau^\phi - 2Mu^t + ru^t) - \frac{9M^2\lambda}{m_0^5 r^{10}}(J_0 - E_0 a)\times \\ & (E_0 J_0 - E_0^2 a + m_0^2 a)(2J_0^2 - 4E_0 J a + 2E_0^2 a^2 + m_0^2 r^2)\,, \end{aligned} \tag{5.14}$$

$$\begin{aligned} J =& \, p_\phi \\ =& \frac{1}{r}[2Mau^t - u^\phi\left(2Ma^2 + ra^2 + r^3\right)] \\ & + \frac{9M^2\lambda}{m_0^5 r^{10}}(J_0 - E_0 a)\{2J_0^4 - 6E_0 J_0^3 a + m_0^2(a^2 + r^2) \\ & + J_0^2[6E_0^2 a^2 + m_0^2(2a^2 + 3r^2)] \\ & - J_0[2E_0^3 a^3 + E_0 m_0^2 a(4a^2 + 5r^2)]\}. \end{aligned} \tag{5.15}$$

As both E, J are conserved and the geodesic contributions E_0, J_0 are not, one can obtain E_0, J_0 at any stage of the orbits as functions of E, J from the above equations. Notice that both m_0 and m_1 are no longer constant with the presence of quadrupole deformation. In fact, as shown in [344], the following mass-like quantity μ as

$$\mu = m_0 + \frac{\lambda}{4} E_{ab} E^{ab} + O(\lambda^2) \tag{5.16}$$

is approximately constant if we neglect the second order tidal effects. It is straightforward to show that Eq. (5.16) implies

$$\frac{m_0}{\mu} = 1 - \frac{3M^2\lambda}{2m_0^5 r^{10}}[m_0^4 r^4 + 3m_0^2 r^2 (J_0 - E_0 a)^2$$

$$+ 3(J_0 - E_0 a)^4] + O(\lambda^2)\,. \tag{5.17}$$

5.2.3 Orbital description

In the Kerr spacetime, the motion of a generic test body with internal quadrupolar moment is no longer separable as there are only three conserved quantities: E, J and μ. However, for equatorial orbits the inclination angle being a constant: $\theta = \pi/2$, and the motion in r and ϕ directions are still separable. In particular, if the orbit is circular, all conserved quantities can be expressed as functions of r. With this understanding, we shall explicitly write down the orbital equation of motion up to linear order in λ for equatorial orbits. According to the expressions for E, J in (5.14), (5.15) and $m_1^2 = m_0^2 = -p^a p_a$, they are

$$\begin{aligned}\left(\frac{dt}{d\tau}\right) &= \frac{E}{m_0 r^2}\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\right] + \frac{aJ}{m_0 r^2}\left(1 - \frac{r^2+a^2}{\Delta}\right) \\ &\quad + F_{0t}(r,a,E,J) \\ &:= F_{00}(r,a,E,J) + F_{0t}(r,a,E,J),\end{aligned} \tag{5.18}$$

$$\begin{aligned}\left(\frac{d\phi}{d\tau}\right) &= \frac{J}{m_0 r^2} + \frac{aE}{m_0 r^2}\left(\frac{r^2+a^2}{\Delta} - 1\right) - \frac{a^2 J}{m_0 r^2 \Delta} \\ &\quad + F_{3t}(r,a,E,J) \\ &:= F_{30}(r,a,E,J) + F_{3t}(r,a,E,J),\end{aligned} \tag{5.19}$$

$$\begin{aligned}\left(\frac{dr}{d\tau}\right)^2 &= \frac{\left[E(r^2+a^2) - aJ\right]^2}{m_0^4 r^4} - \frac{\Delta[r^2 + (J-aE)^2]}{m_0^4 r^4} + F_{1t}(r,a,E,J) \\ &:= F_{10}(r,a,E,J) + F_{1t}(r,a,E,J),\end{aligned} \tag{5.20}$$

where

$$F_{0t}(r,a,E,J) = -\frac{9M^2\lambda}{m_0^6 r^{11}\Delta}\Big(J-Ea\Big)\Big[2J^2-4EJa+2E^2a^2+m_0^2r^2\Big]\Big[2J^2Ma+E^2a(2Ma^2+a^2r+r^3)$$
$$-EJ(4Ma^2+a^2r+r^3)-m_0^2ar(a^2-2Mr+r^2)\Big], \tag{5.21}$$

$$F_{3t}(r,a,E,J) = \frac{9M^2\lambda}{m_0^6 r^{11}\Delta}\Big(J-Ea\Big)\Big[2J^2-4EJa+2E^2a^2+m_0^2r^2\Big]\Big[-2E^2Ma^2+EJa(4M-r)$$
$$+J^2(-2M+r)+m_0^2r(a^2-2Mr+r^2)\Big], \tag{5.22}$$

$$F_{1t}(r,a,E,J) = \frac{18M^2\lambda}{m_0^7 r^{13}}\Big(J-Ea\Big)^2\Big[2J^2-4EJa+2E^2a^2+m_0^2r^2\Big]\Big[-4EJMa+J^2(2M-r)-m_0^2r\Delta$$
$$+E^2(2Ma^2+a^2r+r^3)\Big]. \tag{5.23}$$

The terms F_{00}, F_{10}, F_{30} represent the geodesic motion in the Kerr spacetime, and F_{0t}, F_{1t}, F_{3t} account for the leading-order tidal correction [3]. Strictly speaking, the adiabatic tide approximation (Eq. (5.9)) breaks down for eccentric orbits as the environmental tidal tensor E_{ab} varies on the orbital timescale. The f-mode excitation and evolution have to be included into the equations of motion [345]. However, as the main purpose of this paper is to generate waveforms for circular orbits, where the adiabatic approximation still holds, we can view Eq. (5.18), Eq. (5.19) and Eq. (5.20) as effective equations of motions that are introduced as intermediate steps to find the circular orbits.

In the remaining part of the paper, for the sake of convenience, we introduce the following dimensionless variables:

$$r \to \frac{r}{M},\ J \to \frac{J}{M\mu},\ E \to \frac{E}{\mu} \tag{5.24}$$

[3] There are no F_{20}, F_{2t} terms here as the motion in the θ direction is not present for equatorial orbits.

to replace the unnormalized variables. In this convention, we can rewrite Eq. (5.20) in the form:

$$\left(\frac{1}{r}\frac{dr}{d\tau}\right)^2 = \alpha E^2 - 2\beta\frac{J}{r}E + \gamma\frac{J^2}{r^2} - \delta\frac{m_0^2}{\mu^2}, \tag{5.25}$$

where

$$\begin{aligned} \alpha &= 1 + \frac{a^2(2+r)}{r^3}, \quad \beta = \frac{2a}{r^2} \\ \delta &= 1 + \frac{a^2 - 2r}{r^2}, \quad \gamma = -1 + \frac{2}{r}. \end{aligned} \tag{5.26}$$

Therefore, we know that Eq. (5.25) describes a one-dimensional motion within a potential well. For circular orbits, we require the radial velocity to be zero at the equilibrium radius and the radial acceleration to be zero at the same location. Based on these two requirements, we can obtain the conserved E, J as functions of the equilibrium radius r:

$$\begin{aligned} E(r) = \quad & \frac{1 - 2v^2 + av^3}{\sqrt{1 - 3v^2 + 2av^3}} \\ & + \lambda\left(4r\sqrt{1 - 3v^2 + 2av^3}\right)^{-1}\Big[a^2 g(r) + 2a f(r)v \\ & + (-2 + r)(2f(r) + g(r)r)\Big], \end{aligned} \tag{5.27}$$

$$\begin{aligned} J(r) = \quad & \frac{1 - 2v^2 + av^3}{\sqrt{1 - 3v^2 + 2av^3}} \\ & + \lambda\Big[4(1 - 3v^2 + 2av^3)\Big]^{-3/2}\frac{1}{r^{5/2}}\Big\{2a^4 g(r) \\ & + a^3 v[3g(r)(r-1)r + 4f(r)] \\ & + a^2[g(r)r((r-1)r - 4) + 2f(r)(r-7)] \\ & + 3a[g(r)(r-2)(r-1)r + 4f(r)]\sqrt{r} \\ & + (r-3)r^2[g(r)(r-2)r + 2f(r)]\Big\}, \end{aligned} \tag{5.28}$$

where

$$
\begin{aligned}
v = \quad & \sqrt{\frac{1}{r}}, \\
f(r) = \quad & \frac{3\lambda}{r^8\,(2av + r - 3)^2}\Big[-3a^4 + \frac{6a^3}{v} - a^2 r(3r+1) \\
& + \frac{2a}{v^5} - r^2\left(r^2 - 3r + 3\right)\Big], \qquad (5.29) \\
g(r) = \quad & -\frac{18\lambda}{r^9\,(2av + r - 3)^2}\Big[-5a^4 + \frac{12a^3}{v} - 2a^2 r(2r+3) \\
& + \frac{4a}{v^5} - r^2\left(r^2 - 2r + 2\right)\Big]. \qquad (5.30)
\end{aligned}
$$

In order to compute the gravitational wave fluxes, we also need to evaluate the orbital frequency (only prograde orbits are considered here):

$$
\begin{aligned}
\Omega_\phi = \frac{d\phi}{dt} = \frac{d\phi/d\tau}{dt/d\tau} = \frac{1}{r^{3/2} + a}\Big\{1 + \\
\frac{F_{3t}(r,a,E,J)}{F_{30}(r,a,E,J)} - \frac{F_{0t}(r,a,E,J)}{F_{00}(r,a,E,J)} - \\
\frac{(E_t J_0 - E_0 J_t) r^2 \Delta}{[2aE_0 + J_0(-2+r)][-2aJ_0 + E_0 r^3 + a^2 E_0 (2+r)]}\Big\}.
\end{aligned} \qquad (5.31)
$$

If we substitute Eq. (5.31) into Eq. (5.27) and Eq. (5.28), we can obtain $E(\Omega) = E_0(\Omega) + E_t(\Omega)$ and $J = J_0(\Omega) + J_t(\Omega)$, where $E_0(\Omega)$, $E_t(\Omega)$, $J_0(\Omega)$ and $J_t(\Omega)$ are the geodesics and tidal parts of energy and angular momentum respectively.

We have incorporated these explicit tidal corrections in Eq. (5.27) and Eq. (5.31) in an open source Teukolsky code "Gremlin" within the "Black Hole Perturbation Toolkit" project[346], which provides many useful toolboxes for describing the motion and wave emission of EMRIs. This tide-modified Gremlin package [347] allows us to evolve the trajectory of a point particle in a Kerr spacetime, while counting for the tide-induced corrections. In Sec. 5.3 we use the same code

to compute the gravitational radiation associate with the particle motion.

5.2.4 Dynamic tide

In the low frequency limit, the stars answer to the adiabatic environmental tidal fields by deforming themselves according to Eq. (5.9), with λ being a constant. This scenario is often referred as the "equilibrium tide"[348]. In the late part of the inspiral, although the orbital frequency Ω is still lower than the frequency ω_f of the $\ell = 2$ f-mode, the gradual excitation of the f-mode in the pre-resonance stage is no longer negligible. In fact, as shown in [8], effectively we need to replace the constant (dimensionless) Love number

$$\lambda = \frac{2k_2R^5}{3G} \tag{5.32}$$

by

$$k_l^{\rm dyn} = k_l\Big[a_l + \frac{b_l}{2}\Big(\frac{Q^{DT}_{m=l}}{Q^{AT}_{m=l}} + \frac{Q^{DT}_{m=-l}}{Q^{AT}_{m=-l}}\Big)\Big] \tag{5.33}$$

where

$$\begin{aligned}\frac{Q^{DT}_m}{Q^{AT}_m} &= \frac{\omega_f^2}{\omega_f^2 - (m\Omega)^2} + \frac{\omega_f^2}{2(m\Omega)^2\epsilon_f\Omega'_f(\phi-\phi_f)} \\ &\pm \frac{i\omega_f^2}{(m\Omega)^2\sqrt{\epsilon_f}} e^{\pm i\Omega'_f\epsilon_f(\phi-\phi_f)^2}\int_{-\infty}^{\sqrt{\epsilon_f}(\phi-\phi_f)} e^{\mp i\Omega'_f s^2} ds\,.\end{aligned} \tag{5.34}$$

where the coefficients $a_2 = 1/4$, $b_2 = 3/4$(only $\ell = 2$ is considered here), $\Omega^2 = M/r^3$ and ϵ_f is the ratio between the orbital timescales and the gravitational radiation reaction timescales, $\Omega'_f = 3/8$ is a rescaled derivative in frequency, ϕ is the orbital phase and ϕ_f denotes the orbital phase evaluation

at $\omega = \omega_f$. These quantities can be written as a function of r:

$$\epsilon_f = \frac{256\, M^{2/3}\omega_f^{5/3}\mu}{5\,|m|^{5/3}}\,, \tag{5.35}$$

$$\phi - \phi_f = \frac{1}{32M^{3/2}\mu}\left[\left(\frac{\sqrt{M}|m|}{\omega}\right)^{5/3} - r^{5/2}\right], \tag{5.36}$$

which can be found in [8]. In the above two equations, we do not use the dimensionless variables defined in Eq. (5.24) in order to express them explicitly.

Note here the star still oscillates at the same frequency of the external tidal force, which is why a frequency-dependent Love number can be introduced here. If the f-mode frequency were within the inspiral frequency range, the post-resonance star also oscillates with a frequency component $\omega = \omega_f$ [94]. Such free f-mode oscillations have been observed in numerical simulations of eccentric binary neutron stars [345].

In Sec. 5.4 we show the performance of hybrid waveform models with the dynamic tide effect implemented. The dynamic tide model generally fits better with the numerical waveforms in the late inspiral stage, as demonstrated in [8].

5.3 Radiation

Neutron stars develop nonzero quadrupole moments because of the gravitational tidal fields from their companions. As a result, the stress-energy of the star is modified by the tidal deformation Eq. (5.12). This extra piece of stress energy also generates additional gravitational wave radiation, which in turn affects the orbital evolution. In this section we first review the relevant Teukolsky formalism and then compute the tide-induced gravitational wave radiation.

5.3.1 The Teukolsky equation

The wave emission by an extended body moving in the Kerr spacetime can be described by the Teukolsky equation [349], which is separable in the frequency domain. In particular, consider the variable

$$\psi_4 = \frac{1}{(r - ia\cos\theta)^4}\int_{-\infty}^{\infty} d\omega \sum_{lm} R_{lm\omega}(r)\,{}_{-2}S^{a\omega}_{lm}(\theta)e^{im\phi - i\omega t} \tag{5.37}$$

which is a Newmann-Penrose quantity defined by contracting the Weyl tensor C_{abcd} with tetrad vectors: $\psi_4 = -C_{abcd}n^a\bar{m}^b n^c \bar{m}^d$. The Kinnersley tetrad components are being used [350]

$$\begin{aligned} n_a &= \frac{1}{2}(\frac{\Delta}{\Sigma}, 1, 0, -\frac{a\Delta\sin^2\theta}{\Sigma}) \\ \bar{m}_a &= \frac{\rho}{\sqrt{2}}(ia\sin\theta, 0, \Sigma, -i(r^2+a^2)\sin\theta) \end{aligned} \tag{5.38}$$

At any given frequency ω, the wave equation is separable. In particular, the eigen-solution of the angular part of the Teukolsky equation defines the spin-weighted spheroidal harmonic ${}_{-2}S^{a\omega}_{lm}(\theta)$, which is normalized by

$$\int_0^\pi |{}_{-2}S^{a\omega}_{lm}(\theta)|^2 \sin\theta d\theta = 1. \tag{5.39}$$

We have listed relevant properties of the spin-weighted spheroidal harmonics and their derivatives in Appendix A. The radial function $R_{lm\omega}(r)$ obeys the radial Teukolsky equation:

$$\Delta^2 \frac{d}{dr}\left(\frac{1}{\Delta}\frac{dR_{lm\omega}}{dr}\right) - V(r)R_{lm\omega}(r) = -\mathcal{T}_{lm\omega}(r). \tag{5.40}$$

where

$$V(r) = -\frac{K^2 + 4i(r-M)K}{\Delta} + 8i\omega r + \lambda_{lm}\ , \tag{5.41}$$

and $\Delta = r^2 - 2Mr + a^2$, $K = (r^2 + a^2)\omega - ma$, $\lambda_{lm} \equiv \mathcal{A}_{lm} - 2am\omega + a^2\omega^2 - 2$, where A_{lm} is the eigenvalue of the angular Teukolsky equation.

The radial Teukolsky equation is an ordinary differential equation, which can be solved by using the Green function method. To achieve this goal, one needs to first identify two independent solutions of the homogeneous Teukolsky equation: $R^H_{lm\omega}$ and $R^\infty_{lm\omega}$, which have the following asymptotic behaviour:

$$\begin{aligned} R^H_{lm\omega} &= B^{\text{trans}}_{lm\omega}\Delta^2 e^{-ip_m r^*}, \qquad r \to r_+ \\ R^H_{lm\omega} &= B^{\text{ref}}_{lm\omega} r^3 e^{i\omega r^*} + \frac{B^{\text{inc}}_{lm\omega}}{r} e^{-i\omega r^*}, \qquad r \to \infty \end{aligned} \tag{5.42}$$

and

$$\begin{aligned} R^\infty_{lm\omega} &= C^{\text{up}}_{lm\omega} e^{ip_m r^*} + C^{\text{ref}}_{lm\omega}\Delta^2 e^{-ip_m r^*}, \qquad r \to r_+ \\ R^\infty_{lm\omega} &= C^{\text{trans}}_{lm\omega} r^3 e^{i\omega r^*}, \qquad r \to \infty\ . \end{aligned} \tag{5.43}$$

where $p_m = \omega - ma/2Mr_+$ and the tortoise coordinate r^* is:

$$r^*(r) = r + \frac{2Mr_+}{r_+ - r_-}\ln\frac{r - r_+}{2M} - \frac{2Mr_-}{r_+ - r_-}\ln\frac{r - r_-}{2M},$$

where the outer and inner horizon radii are $r_\pm = M \pm \sqrt{M^2 - a^2}$. Based on the Green's functions method, the general solution of the Teukolsky equation with a source can be written in this form:

$$R_{lm\omega}(r) = Z^H_{lm\omega}(r)R^\infty_{lm\omega}(r) + Z^\infty_{lm\omega}(r)R^H_{lm\omega}(r)\ , \tag{5.44}$$

where

$$Z^H_{lm\omega}(r) = \frac{B^{\rm trans}_{lm\omega}}{2i\omega B^{\rm inc}_{lm\omega} C^{\rm trans}_{lm\omega}} \int_{r_+}^{r} dr' \frac{R^H_{lm\omega}(r')\mathcal{T}_{lm\omega}(r')}{\Delta(r')^2} \ ,$$
$$Z^\infty_{lm\omega}(r) = \frac{1}{2i\omega B^{\rm inc}_{lm\omega}} \int_{r}^{\infty} dr' \frac{R^\infty_{lm\omega}(r')\mathcal{T}_{lm\omega}(r')}{\Delta(r')^2} \ . \tag{5.45}$$

As the neutron star we consider here moves along circular and equatorial orbits, there is only one frequency in this setup $\omega = \Omega_\phi$. In particular, the mth harmonic has a frequency of

$$\omega_m = m\Omega_\phi. \tag{5.46}$$

Then we have

$$Z^{H,\infty}_{lm\omega} = \sum_m \delta(\omega - \omega_m) Z^{H,\infty}_{lm\omega_m} \tag{5.47}$$

The energy fluxes going out to infinity and black hole horizon can be obtained as:

$$\left(\frac{dE}{dt}\right)^{r\to\infty}_{GW} = \sum_{lm} \frac{|Z^\infty_{lm\omega_m}|^2}{4\pi\omega_m^2}. \tag{5.48}$$
$$\left(\frac{dE}{dt}\right)^{r\to r_+}_{GW} = \sum_{lm} \alpha_{lm} \frac{|Z^H_{lm\omega_m}|^2}{4\pi\omega_m^2} \tag{5.49}$$

where

$$\alpha_{lm} = \frac{256(2Mr_+)^5 p_m (p_m^2 + 4\varepsilon^2)(p_m^2 + 16\varepsilon^2)\omega_m^3}{|C_{lm}|^2} \ , \tag{5.50}$$

with $\varepsilon = \sqrt{M^2 - a^2}/4Mr_+$ and

$$|C_{lm}|^2 = \left[(\lambda_{lm} + 2)^2 + 4a\omega_m - 4a^2\omega_m^2\right][\lambda_{lm}^2 + 36ma\omega_m$$
$$- 36a^2\omega_m^2] + (2\lambda_{lm} + 3)\left[6a^2\omega_m^2 - 48ma\omega_m\right]$$

$$+ 144\omega_m^2(M^2 - a^2). \tag{5.51}$$

As mentioned earlier, the boundary condition for gravitational waves on the star's surface is different from the one for black hole horizon. As a result, the horizon flux should not be accounted for if both objects are neutron stars in the binary system. However, it is a 4PN effect for Schwarzschild black holes and 2.5PN effect for Kerr black holes, and the associated phase shift is less than 0.1 even for the point mass motion. Therefore in characterizing the tidal effect and the *extra* gravitational wave emission associated with tidal deformation, we shall not consider the issue of the horizon fluxes in our paper.

5.3.2 Source term

In order to obtain the energy flux, we need to evaluate the source term in Eq. (5.45). It is explicitly given by [351]:

$$\mathcal{T}_{lm\omega}(r) = 4\int d\Omega\, dt\, \frac{\Sigma}{\rho^4}\,(B_2' + B_2^{*\prime})_{-2}\,S_{lm}^{a\omega}(\theta)e^{-im\phi}e^{i\omega t}\ , \tag{5.52}$$

where the functions B_2' and $B_2^{*\prime}$ are

$$\begin{aligned}
B_2' =& -\frac{\rho^8\bar{\rho}}{2}L_{-1}\left[\rho^{-4}L_0\left(\rho^{-2}\bar{\rho}^{-1}T_{nn}\right)\right] \\
&+ \frac{\Delta^2\rho^8\bar{\rho}}{2\sqrt{2}}L_{-1}\left[\rho^{-4}\bar{\rho}^2 J_+\left(\rho^{-2}\bar{\rho}^{-2}\Delta^{-1}T_{n\bar{m}}\right)\right], \\
B_2^{*\prime} =& \frac{\Delta^2\rho^8\bar{\rho}}{2\sqrt{2}}J_+\left[\rho^{-4}\bar{\rho}^2\Delta^{-1}L_{-1}\left(\rho^{-2}\bar{\rho}^{-2}T_{n\bar{m}}\right)\right] \\
&- \frac{\Delta^2\rho^8\bar{\rho}}{4}J_+\left[\rho^{-4}J_+\left(\rho^{-2}\bar{\rho}T_{\bar{m}\bar{m}}\right)\right]\ .
\end{aligned} \tag{5.53}$$

Here, $\rho = -1/(r - ia\cos\theta)$, $\bar{\rho} = -1/(r + ia\cos\theta)$. The differential operators J_+ and L_s are

$$J_+ = \partial_r + \frac{iK(r)}{\Delta}\ ,$$

$$L_s = \partial_\theta + m\csc\theta - a\omega\sin\theta + s\cot\theta\,,$$
$$L_s^\dagger = \partial_\theta - m\csc\theta - a\omega\sin\theta + s\cot\theta\,. \tag{5.54}$$

The stress-energy tensor for an extended body moving in the Kerr spacetime, as described in Eq. (5.12), is given by

$$\begin{aligned} T^{ab}(x) &= \int d\tau\Big[u^{(a}p^{(b)}\delta_{(4)} + \frac{1}{3}R_{cde}{}^{(a}J^{b)edc}\delta_{(4)} \\ &\quad - \frac{2}{3}\nabla_d\nabla_c(J^{d(ab)c}\delta_{(4)})\Big]. \qquad (5.55) \\ &= \int d\tau\Big[u^{(a}p^{(b)} + \frac{1}{3}R_{cde}{}^{(a}J^{b)edc} - J^{daec}\partial_c\Gamma^b_{de} \\ &\quad + \Gamma^a_{de}\Big(\Gamma^d_{cf}J^{febc} + \Gamma^e_{cf}J^{dfbc}\Big) \\ &\quad + \Gamma^b_{de}\Big(\Gamma^d_{cf}J^{faec} + \Gamma^a_{cf}J^{dfec}\Big)\Big]\frac{\delta[x - z(\tau)]}{\sqrt{-g}} \\ &\quad + \partial_d\Big(\Gamma^d_{cf}J^{fabc} + \Gamma^a_{cf}J^{dfbc} + \Gamma^b_{cf}J^{cafd}\Big)\frac{1}{\sqrt{-g}} \\ &\quad + \partial_d\partial_c\left\{J^{dabc}\delta[x - z(\tau)]\right\}\frac{1}{\sqrt{-g}} \qquad (5.56) \\ &:= \int d\tau\, A^{ab}\frac{\delta[x - z(\tau)]}{\sqrt{-g}} + \partial_d\left\{B^{bab}\delta[x - z(\tau)]\right\}\frac{1}{\sqrt{-g}} \\ &\quad + \partial_d\partial_c\left\{J^{dabc}\delta[x - z(\tau)]\right\}\frac{1}{\sqrt{-g}}, \qquad (5.57) \end{aligned}$$

where we have converted the covariant derivatives into coordinate partial derivatives with Christoffel symbols, which are more convenient for numerical evaluation. Here the delta function $\delta[x-z(\tau)]$ is defined as

$$\delta[x - z(\tau)] = \delta[t - t(\tau)]\delta[r - r(\tau)]\delta[\theta - \theta(\tau)]\delta[\phi - \phi(\tau)]. \tag{5.58}$$

J^{dabc} in Eq. (5.55) is a tensor, then we have also introduced additional notations for J^{dabc} to account for various pieces of the source terms, as modified by the tidal field

$$A_{\{nn\}} := A^{ab} n_a n_b, \tag{5.59}$$

$$B^{\{\alpha\}}_{\{nn\}} := B^{\alpha ab} n_a n_b, \quad J^{\{\alpha\beta\}}_{\{nn\}} := J^{\alpha ab\beta} n_a n_b. \tag{5.60}$$

where $\alpha, \beta = t, r, \theta, \phi$, and

$$\begin{aligned}
B^{\{r\}}_{\{drnn\}} &:= B^{rab}\partial_r(n_a n_b)\,, \quad B^{\{\theta\}}_{\{d\theta nn\}} := B^{\theta ab}\partial_\theta(n_a n_b),\\
J^{\{tr\}}_{\{drnn\}} &:= J^{tabr}\partial_r(n_a n_b)\,, \quad J^{\{t\theta\}}_{\{d\theta nn\}} := J^{tab\theta}\partial_\theta(n_a n_b),\\
J^{\{r\phi\}}_{\{drnn\}} &:= J^{rab\phi}\partial_r(n_a n_b)\,, \quad J^{\{\theta\phi\}}_{\{d\theta nn\}} := J^{\theta ab\phi}\partial_\theta(n_a n_b),\\
J^{\{rr\}}_{\{ddrnn\}} &:= J^{rabr}\partial_r\partial_r(n_a n_b)\,, \quad J^{\{\theta\theta\}}_{\{dd\theta nn\}} := J^{\theta ab\theta}\partial_\theta\partial_\theta(n_a n_b).\\
J^{\{r\theta\}}_{\{drnn\}} &:= J^{rab\theta}\partial_r(n_a n_b)\,, \quad J^{\{r\theta\}}_{\{d\theta nn\}} := J^{rab\theta}\partial_\theta(n_a n_b)\\
J^{\{r\theta\}}_{\{drd\theta nn\}} &:= J^{r\theta ab\theta}\partial_r\partial_\theta(n_a n_b).
\end{aligned} \tag{5.61}$$

Apart from $n_a n_a$, we can also define components for J^{dabc} by contracting J^{dabc} with $n_{(a}\bar{m}_{b)}$ and $\bar{m}_a\bar{m}_b$ following similar convention as the above equations. The explicit forms of these components are given in the Appendix B.

5.3.3 Sample evolution

We incorporated the additional tide-related source terms into the Gremlin code, and evaluate the gravitational wave energy flux as a function of the orbital frequency. Formally we can write the total power as

$$P = P_{\rm pm} + P_{\rm tide} = \eta^2 \left(\frac{dE}{dt}\right)^{\rm pm} + \eta \frac{\lambda}{m_0^5}\left(\frac{dE}{dt}\right)^{\rm tide}. \tag{5.62}$$

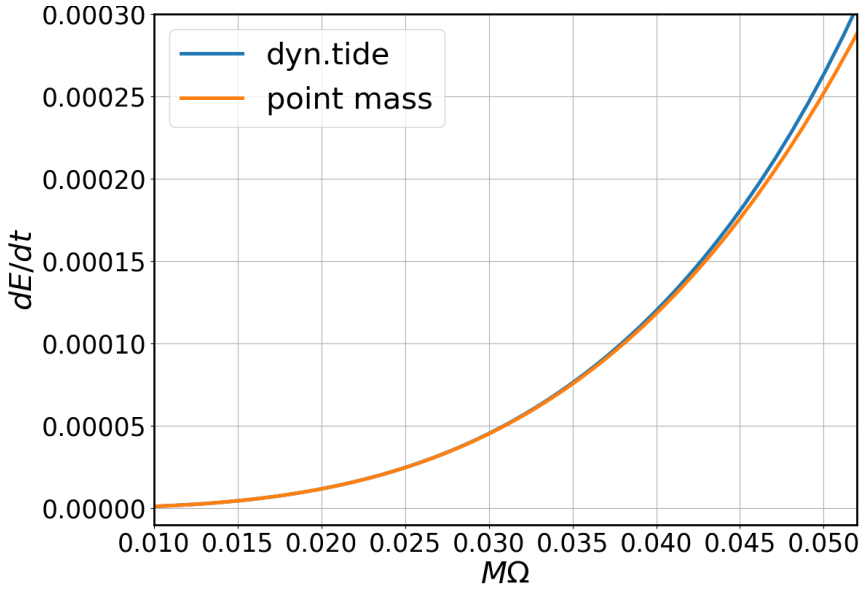

Figure 5.1: The energy flux computed for an equal-mass, black hole-neutron star system with $k_2(\Omega = 0) = 0.07346$, $m_0 = 1.4M_\odot$, $M\omega_f = 0.1349(1+q)/2$ and $\Gamma = 2$ polytropic equation of state, and a similar binary black hole system with the same mass and starting from the same initial location. The flux dE/dt is normalized by η^2.

The η^2 factor within the point mass term is related to the fact that metric perturbation generated by the point mass is proportional to the mass ratio, so that the flux is proportional to η^2. The tidal correction of the gravitational wave flux is generated by the beating of the wave generated by the point mass with the additional wave generated by the quadrupole deformation of the star. Both $\dot{E}^{\rm pm}$ and $\dot{E}^{\rm tide}$ can be computed given the initial conditions of the system. The values can be used in other systems with different η and λ.

In Fig. 5.1, we plot the total power versus the point mass power for a non-spinning, equal-mass black hole-neutron star system . The same type of system is also used in Sec. 5.4 for waveform comparison. The additional energy flux contributed by the tidal deformation (Eq. (5.62)) becomes more important at higher frequencies. Although the fluxes are computed within the extreme-mass-ratio limit, the results are applied in the comparable mass ration limit for the waveform construction.

5.4 Waveform Construction

With the preparation in Sec. 5.2 and Sec. 5.3 on the conservative and dissipative pieces of the tidal effects, we are ready to present the tidal correction to the gravitational waveform. We shall focus on the gravitational wave phase as it is the most sensitively measured quantity within a parameter estimation process. In general, amplitude corrections for binary neutron star and Black hole-neutron star waveforms are not negligible [352, 353].

Assuming adiabatic circular orbit evolution, the motion at any instantaneous moment can be approximately viewed as a circular orbit with frequency Ω. The gravitational wave phase, as a function of the orbital frequency, follows

$$\frac{d^2\psi}{d\Omega^2} = 2\frac{dE/d\Omega}{P} \,. \tag{5.63}$$

As we are interested in the tidal correction, we shall write the total phase ψ as $\psi^{\rm pm} + \psi^{\rm tide}$, the total energy as $E = E^{\rm pm} + E^{\rm tide}$, and expand Eq. (5.63) so that only linear order terms in λ are kept:

$$\frac{d^2\psi^{\rm tide}}{d\Omega^2} = 2\left(\frac{dE^{\rm tide}/d\Omega}{P^{\rm pm}} - P^{\rm tide}\frac{dE^{\rm pm}/d\Omega}{(P^{\rm pm})^2}\right), \tag{5.64}$$

where we plug in $E^{\rm tide}$ and $P^{\rm tide}$ evaluated in Sec. 5.2 and Sec. 5.3. In the Post-Newtonian theory, $E^{\rm tide}$ and $P^{\rm tide}$ can be computed to various PN orders, which lead to the PN tide waveform at different orders [90]. Notice that the gravitational wave phase increases twice as fast as the orbital phase, because we focus on the dominant piece of the waveform with $\ell = 2, m = 2$.

5.4.1 Hybrid waveform

The black hole perturbation calculation discussed in Sec. 5.2 and Sec. 5.3 gives rise to an EMRI-inspired waveform, which is fully capable of describing the gravitational wave emission in the highly relativistic regime. On the other hand, the PN tide waveform, although being less accurate in the strong-gravity regime, does not require an expansion in the mass ratio. In order to combine the merits of these two different approaches, we have proposed a hybrid version of the waveform, as explained in Eq. (1.55) and depicted in Fig. 1.4. By definition, this hybrid waveform is accurate if the mass ratio is small *or* if the binary separation is large. Similar to the spirit of the EOB construction, we anticipate that by ensuring matching at small mass ratio and weak gravity limit, the hybrid method still provides reasonably accurate description for comparable mass-ratio systems in the strong gravity regime. This point has to be checked with numerical relativity waveforms, as discussed in Sec. 5.4.2.

In constructing the hybrid waveform one needs to subtract the waveform contribution in the overlap regime, as explained in Fig. 1.4. In fact, it also serves as a sanity check that the PN waveform taking a mass ratio expansion should agree with the EMRI-inspired waveform taking a PN expansion. In light of Eq. (5.64), it suffices to show that $E^{\rm tide}$ and $P^{\rm tide}$ obtained in the PN theory

have the same small mass limit as their counterparts found in Sec. 5.2 and Sec. 5.3, expanded in various PN orders. Such a consistency check is explicitly performed in Appendix. 5.C.

5.4.2 Numerical comparison

In order to evaluate the performance of the black hole perturbation and hybrid methods in constructing waveforms, we adopt an equal mass, black-hole neutron-star waveform from the SXS waveform catalog [354]. For this particular waveform, the neutron stars have a polytropic equation of state $P = K\rho^{\Gamma}$, with $\Gamma = 2$, $K = 101.45$. The neutron star mass is $m = 1.4M_{\odot}$ and the radius is $R = 14.4$km. The phase error is approximately ~ 1 rad at the peak of the strain [8].

For comparison purpose, we also compute the EOB prediction of the tidal phase correction, with dynamic tide included, in addition to the black hole perturbation result. As shown in Fig. 5.2, the hybrid waveform that integrates both the black hole perturbation and 2PN methods, performs significantly better than the black hole perturbation result alone. The 2PN expansion of the tidal effects are from Eq.(5) of [355](”ad 2PN”). This hybrid waveform also has shown almost comparable accuracy as the EOB-dynamic tide waveform, using this polytropic star waveform. In Fig. 5.3, we consider a black hole-neutron star system with mass ratio 2:1 and the property of the neutron star is the same as Fig. 5.1 and Fig. 5.2. We observe slightly better agreement with the numerical wavefrom for the hybrid waveform is in this case, but the difference is within the phase uncertainty of the numerical waveform. Apart from these two scenarios, more detailed and systematic comparison and characterization are needed to address the phase error of the hybrid waveform.

This hybrid waveform is naturally expressed in the frequency domain, which is convenient for fast waveform evaluation. To further improve the waveform accuracy to meet the requirements of third-generation gravitational wave detectors, high-order corrections (q^0 and η^1) in the black hole perturbation method should be evaluated to reduce the empty space in Fig. 1.4. As numerical waveforms are required for validation and calibration purposes, we also likely require future numerical

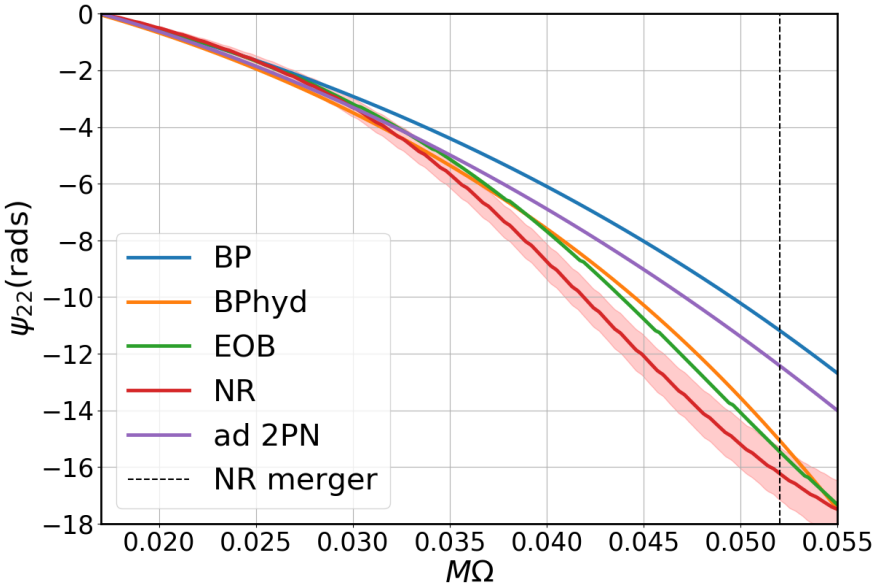

Figure 5.2: Tidal phases of gravitational perturbation theory, the hybrid method, the EOB framework and numerical relativity simulation for an equal-mass, binary neutron star system. Here $M\Omega$ is the orbital frequency and "ad 2PN" stands for adiabatic 2PN tides. The properties of neutron star are the same as those described in Fig. 5.1. The vertical dashed line shows the merger frequency of the system and the red shaded region indicates the estimated phase error from the numerical simulation[358].

waveforms with $O(0.1)$ phase error, i.e., a factor of ten improvement from current waveforms.

Interestingly, the black hole perturbation approach also offers straightforward evaluation of the spin-dependence of the tidal terms, which are absent in the current PN or EOB waveforms. According to Fig. 5.4, the influence of the spin parameter of the black hole on the tide-induced phase shift is less than 10% in the entire inspiral range. Such additional phase shift may be less important for binary neutron stars, as they are generally believed to be slowly spinning according to the observation of galactic pulsar binaries [356]. Nevertheless they should be relevant for black hole-neutron star binaries if we want to control the waveform phase error to be below 0.1, especially for the ones with a low-mass black hole [357].

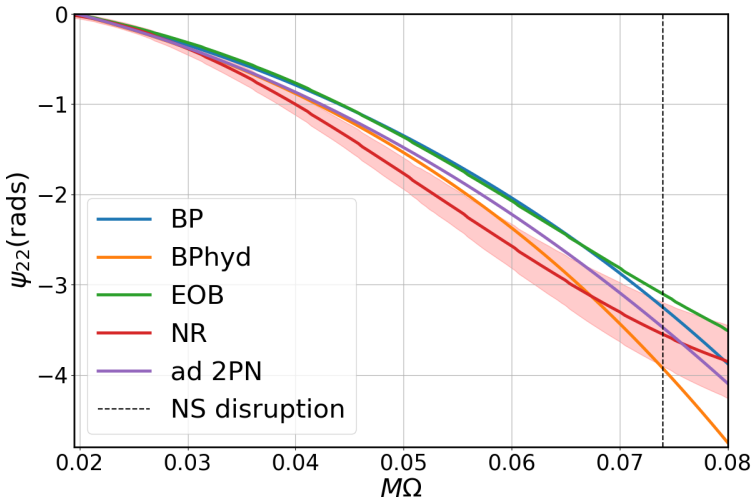

Figure 5.3: Tidal phases of gravitational perturbation theory, the hybrid method, the EOB framework and numerical relativity simulation for black hole-neutron star system with mass ratio 2:1. The properties of the neutron star are the same as those described in Fig. 5.1. The vertical dashed line shows the frequency of tidal disruption.

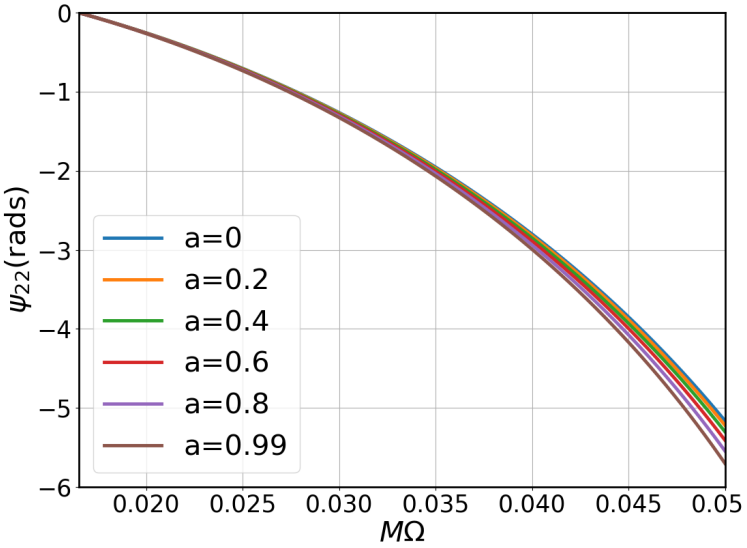

Figure 5.4: Tidal phases in the black hole perturbation waveform with spin ranging from $a = 0$ to $a = 0.99$ for six equal-mass, black hole-neutron star systems. The property of the neutron star is the same as Fig. 5.1.

5.5 Conclusions and Discussion

A recent program in connecting scattering amplitude calculations to two-body problems in General Relativity has triggered an evolution in Post-Newtonian and Post-Minkowski (PM) Theory [359–361]. Higher order PN and PM corrections to the equations of motion have been discovered with this new approach. On the other hand, the development of second-order (in mass ratio) gravitational self force is being carried out and implemented in circular orbits in Schwarzschild [362]. It is expected to correct phase error on the $O(1)$ order, which is on the comparable level of the environmental effects [363–366]. The hybrid approach proposed here naturally integrates these two independent expansions to achieve a better description of binary motions in the comparable mass ratio, strong gravity and high velocity regime. In this work we have incorporated the PN expansion of the tidal correction up to 2PN order and the leading-order term in the mass-ratio expansion, which gives rise to a hybrid waveform with almost comparable accuracy to the state of the art EOB waveform in our comparison with limited number of numerical waveforms. Although in this work we only consider the equal-mass and mass-ratio 2:1 system for numerical comprisons, our model applies for arbitrary mass ratios. This tidal phase can be added to the phase of binary black hole waveforms to obtain the waveforms for binary neutrons tars and black hole-neutron star systems[314].

Moving forward, it should be straightforward to include the $\psi_{2.5\mathrm{PN}}$ term [355, 367, 368] and $\psi^{(0)}_{1\mathrm{BP}}$ corrections. In particular, as $\psi^{(0)}_{1\mathrm{BP}}$ is the leading-order tidal term for the more massive object, it is easier to consider the problem of a point mass orbiting around a star to evaluate $\psi^{(0)}_{1\mathrm{BP}}$. In Fig. 5.4, we observe that the discrepancy between the black hole perturbation waveform and the numerical waveform monotonically increases as the binary evolves. The inclusion of $\psi_{2.5\mathrm{PN}}$ term and $\psi^{(0)}_{1\mathrm{BP}}$ may help alleviate the disagreement. In the future, it is feasible to also work out the $\psi^{(1)}_{1\mathrm{BP}}$ and beyond-2PN corrections to achieve better accuracy.

In [319], it is argued that for the comparison between the numerical relativity binary black

hole waveform and the EMRI-inspired waveform, the discrepancy at large orbital frequency might come from the breakdown of the adiabatic approximation, so that the inspiral-to-plunge transition has to be taken into account. However, in the comparison performed here for the tidal effects, the discrepancy never displays a sudden rise near the merger. Therefore we do not expect the inspiral-to-plunge transition to be the main reason of disagreement found here. Nevertheless, we may still include the transition in our future investigation to see how it affects the waveform construction.

On the other hand, higher-order in mass ratio terms may be obtained by calibration with a set of numerical waveforms [319]. Let us consider Eq. (5.2) as an example. If both $\psi^{(-1)}_{\rm 1BP}$ and $\psi^{(0)}_{\rm 1BP}$ are known through black hole perturbation calculations, we may truncate the summation up to $n \geq 1$ and fit $\psi^{(1)}_{\rm 1BP}, ..., \psi^{(n)}_{\rm 1BP}$ by comparing to a series numerical waveforms with different mass ratios. The obtained fitting formula and the associated waveform can be tested with another independent set of numerical waveforms. The accuracy of this method relies crucially on the accuracy of the calibration waveforms. We plan to perform this analysis using more binary neutron star and black hole-neutron star waveforms.

As Advanced LIGO continues to improve its sensitivity and especially with the third-generation gravitational-wave detectors [369, 370], we should expect to observe a set of high signal-to-noise-ratio (SNR) events, which will allow many important applications of precise gravitational wave astronomy. The gain in SNR also poses strict requirements on the modeling error of the waveforms, so that the waveform systematic error is smaller than the statistical error of these events. It has been shown that for third generation detectors the mismatch error for numerical relativity waveforms has to improve by one order of magnitude. For semi-analytical waveforms an improvement of three orders of magnitude is necessary [371]. Significant new developments are required to bridge such a large gap, and hopefully the hybrid method proposed here will provide one avenue for exploration.

5.A Appendix: Spheroidal harmonics

Even though some derivatives of the spin-weighted spheroidal harmonics can be found in [372], we need some other derivatives which we state them as follows:

$$\begin{aligned}\partial_{\theta -1}Y_{km}(\theta) &= (m\csc\theta - \cot\theta)_{-1}Y_{km}(\theta)\\ &\quad -[k(k+2)]^{1/2}{}_0Y_{km}(\theta),\end{aligned} \tag{5.65}$$

$$\begin{aligned}\partial_{\theta 0}Y_{km}(\theta) &= m\csc\theta_0 Y_{km}(\theta)\\ &\quad -[k(k+2)]^{1/2}{}_1Y_{km}(\theta),\end{aligned} \tag{5.66}$$

$$\begin{aligned}\partial_{\theta 1}Y_{km}(\theta) &= m\csc\theta_1 Y_{km}(\theta)\\ &\quad -[(k-1)(k+2)]^{1/2}{}_2Y_{km}(\theta),\end{aligned} \tag{5.67}$$

$$S = \sum_{k=l_{min}}^{\infty} b_k \partial_{\theta -2}Y_{km}, \tag{5.68}$$

$$\begin{aligned}\partial_\theta S = \quad & (m\cot\theta\csc\theta + 2\csc^2\theta)S\\ & - \sum_{k=l_{min}}^{\infty} b_k[(k-1)(k+2)]^{1/2}{}_{-1}Y_{km}(\theta),\end{aligned} \tag{5.69}$$

$$\begin{aligned}\partial_\theta\partial_\theta S =& (-m\cot\theta\csc\theta + 2\csc^2\theta)S\\ &+ (m\cot\theta\csc\theta + 2\csc^2\theta)\partial_\theta S\end{aligned}$$

$$- \sum_{k=l_{min}}^{\infty} b_k[(k-1)(k+2)]^{1/2}\partial_{\theta-1}Y_{km}(\theta), \tag{5.70}$$

$$\begin{aligned}\partial_\theta L_1^\dagger L_2^\dagger S =& \sum_{k=l_{min}}^{\infty} b_k[(k-1)k(k+1)(k+2)]^{1/2}\partial_{\theta 0}Y_{km}(\theta) \\ &+ 2a\omega\cos\theta L_2^\dagger S + 2a\omega\sin\theta\partial_\theta L_2^\dagger S \\ &- 2(a\omega\sin\theta)a\omega\cos\theta S - (a\omega\sin\theta)^2\partial_\theta S,\end{aligned} \tag{5.71}$$

$$L_2^\dagger S = a\omega\sin\theta S - \sum_{k=l_{min}}^{\infty} b_k[(k-1)(k+2)]^{1/2}{}_{-1}Y_{km}(\theta) \tag{5.72}$$

$$\begin{aligned}\partial_\theta L_2^\dagger S =& a\omega\cos\theta S + a\omega\sin\theta\partial_\theta S \\ &- \sum_{k=l_{min}}^{\infty} b_k[(k-1)(k+2)]^{1/2}\partial_{\theta-1}Y_{km}(\theta),\end{aligned} \tag{5.73}$$

$$\begin{aligned}L_1^\dagger L_2^\dagger S =& \sum_{k=l_{min}}^{\infty} b_k[(k-1)k(k+1)(k+2)]^{1/2}{}_0Y_{km}(\theta) \\ &+ 2a\omega\sin\theta L_2^\dagger S - (a\omega\sin\theta)^2 S\end{aligned} \tag{5.74}$$

$$\begin{aligned}\partial_\theta\partial_\theta L_2^\dagger S =& (-a\omega\sin\theta + 2a\omega\sin\theta)S \\ &+ (a\omega\cos\theta + a\omega m - 2a\omega\cos\theta)\partial_\theta S \\ &- (a\omega\cos\theta - m\cot\theta\csc\theta + \csc^2\theta)\times \\ &\sum_{k=l_{min}}^{\infty} b_k[(k-1)(k+2)]^{1/2}{}_{-1}Y_{km}(\theta) \\ &- (a\omega\sin\theta + m\csc\theta - \cot\theta)\times\end{aligned}$$

$$
\begin{aligned}
&\sum_{k=l_{min}}^{\infty} b_k[(k-1)(k+2)]^{1/2}{}_{-1}Y_{km} \\
&+\sum_{k=l_{min}}^{\infty} b_k[(k-1)k(k+1)(k+2)]^{1/2}{}_{0}Y_{km}, \qquad (5.75)
\end{aligned}
$$

$$
\begin{aligned}
\partial_\theta\partial_\theta L_1^\dagger L_2^\dagger S \;=&(-m\cot\theta\csc\theta)\times \\
&\sum_{k=l_{min}}^{\infty} b_k[(k-1)k(k+1)(k+2)]^{1/2}{}_{0}Y_{km}(\theta) \\
&+ m\csc\theta\sum_{k=l_{min}}^{\infty} b_k[(k-1)(k+2)]^{1/2}\partial_{\theta 0}Y_{km}(\theta) \\
&-\sum_{k=l_{min}}^{\infty} b_k[(k-1)k^2(k+1)^2(k+2)]^{1/2}\partial_{\theta 1}Y_{km}(\theta) \\
&-2a\omega\sin\theta L_2^\dagger S \;+2a\omega\cos\theta\partial_\theta L_2^\dagger S \\
&+2a\omega\cos\theta\partial_\theta L_2^\dagger S \;+2a\omega\sin\theta\partial_\theta\partial_\theta L_2^\dagger S \\
&-2a^2\omega^2(\cos^2\theta-\sin^2\theta)S \\
&-2a^2\omega^2\sin\theta\cos\theta\partial_\theta S \;-2a^2\omega^2\sin\theta\cos\theta\partial_\theta S \\
&-(a\omega\sin\theta)\partial_\theta\partial_\theta S. \qquad (5.76)
\end{aligned}
$$

5.B Appendix: Source terms

Because we consider the first order tidal effects, the we can substitute E_0 and J_0 into the tensors B^{dab} and J^{dabc}. The concrete components are

$$
\begin{aligned}
A_{\{nn\}} =&\frac{\left[Ja-E(r^2+a^2)\right]^2}{4mr^4}+\frac{\lambda(a^2+(-2+r)r)^2}{4r^{11}(-3+2av+r)^3}\Big[-18a^4+\frac{a^3}{v} \\
&-2a^2r(31+3r)+\frac{4a(3+4r)}{v^3}+r^2(9-12r+r^2)\Big] \qquad (5.77)
\end{aligned}
$$

$$A_{\{n\bar{m}\}} = -\frac{i(J-Ea)\left[Ja-E\left(r^2+a^2\right)\right]}{2\sqrt{2}mr^3}+\frac{i\lambda}{2\sqrt{2}\,(2va+r-3)^3}\left(\frac{1}{r}\right)^{19/2}\Big[18va^7$$
$$+\frac{(33r+17)a^5}{v}+r(133-97r)a^4+\frac{\left(17r^2+48r-139\right)a^3}{v^3}+r^2\left(-39r^2+84r-17\right)a^2$$
$$+\frac{\left(2r^3+21r^2-83r+66\right)a}{v^5}+r^3\left(-2r^3+r^2+15r-18\right)-60a^6\Big], \tag{5.78}$$

$$A_{\{\bar{m}\bar{m}\}} =\frac{E\left[(r+2)a^2+r^3\right]-2Ja}{mr\left[(r-2)r+a^2\right]}-\frac{\lambda}{r^9\,(2va+r-3)^3}\Big[\frac{30a^5}{v}-r(12r+7)a^4+\frac{4(11r-20)a^3}{v^3}$$
$$-r^2\left(r^2+46r-104\right)a^2+\frac{2\left(4r^2-2r-15\right)a}{v^5}+r^3\left(r^3-10r^2+21r-9\right)-9a^6\Big], \tag{5.79}$$

$$B^{\{t\}}_{\{nn\}} =\frac{3\lambda\left[(r-2)r+a^2\right]}{8r^9\,(2va+r-3)^2}\Big[6va^5+6ra^4+\frac{(3r-25)a^3}{v}$$
$$+2r\left(3r^2-5r+5\right)a^2+\frac{3\left(r^2-5r+4\right)a}{v^3}+2r^3\left(r^2-5r+6\right)\Big], \tag{5.80}$$

$$B^{\{t\}}_{\{n\bar{m}\}} =-\frac{3i\lambda}{8\sqrt{2}r^8\,(2va+r-3)^2}\Big[6va^6+6(r-1)a^5-\frac{(3r+19)a^4}{v}$$
$$+2r\left(3r^2-2r+13\right)a^3-\frac{2(4r+5)a^2}{v^3}+2r^2\left(r^3-2r^2+3\right)a+\frac{r^2-5r+6}{v^7}\Big], \tag{5.81}$$

$$B^{\{t\}}_{\{\bar{m}\bar{m}\}} = 0, \tag{5.82}$$

$$B^{\{r\}}_{\{nn\}} =\frac{3\lambda\left[(r-2)r+a^2\right]^2}{8r^{10}\,(2va+r-3)^2}\Big[\frac{6a^3}{v}-r(3r+1)a^2+\frac{2a}{v^5}-r^2\left(r^2-3r+3\right)-3a^4\Big], \tag{5.83}$$

$$B^{\{r\}}_{\{n\bar{m}\}} = \frac{3i\lambda v^{17}\left[(r-2)r+a^2\right]}{8\sqrt{2}\,(2va+r-3)^2}\Big[6va^5+\frac{2(3r+4)a^3}{v} + r(15-13r)a^2+\frac{2\left(r^2-3\right)a}{v^3}-3r^2\left(r^2-3r+2\right)-18a^4\Big], \tag{5.84}$$

$$B^{\{r\}}_{\{\bar{m}\bar{m}\}} = -\frac{3\lambda}{4r^8\,(2va+r-3)^2}\Big[\frac{12a^5}{v}-r(3r+10)a^4+\frac{(11r-15)a^3}{v^3} - r^2\left(r^2+9r-25\right)a^2+\frac{\left(3r^2-5r-6\right)a}{v^5}+r^3\left(-2r^2+6r-3\right)-3a^6\Big], \tag{5.85}$$

$$B^{\{r\}}_{\{drnn\}} = \frac{3\lambda\left(r-a^2\right)\left[(r-2)r+a^2\right]}{2r^{11}\,(2va+r-3)^2}\Big[\frac{6a^3}{v}-r(3r+1)a^2 + \frac{2a}{v^5}-r^2\left(r^2-3r+3\right)-3a^4\Big], \tag{5.86}$$

$$B^{\{r\}}_{\{drn\bar{m}\}} = -\frac{3i\lambda}{8\sqrt{2}r^{10}\,(2va+r-3)^2}\Big[-\frac{42a^6}{v}+12r(2r-1)a^5-\frac{9(3r-7)a^4}{v^3} + 2r^2\left(6r^2-20r-1\right)a^3+\frac{2\left(r^2+5r-3\right)a^2}{v^5}+2r^3\left(r^3-10r^2+21r-12\right)a + \frac{3\left(r^2-3r+2\right)}{v^9}+18a^7\Big], \tag{5.87}$$

$$B^{\{r\}}_{\{dr\bar{m}\bar{m}\}} = -\frac{3\lambda}{2r^9\,(2va+r-3)^2}\Big[-\frac{6a^5}{v}+r(3r-2)a^4-\frac{2a^3}{v^5} + r^2\left(r^2-9r+19\right)a^2+\frac{4(r-3)a}{v^5}+r^3\left(-2r^2+6r-3\right)+3a^6\Big], \tag{5.88}$$

$$B^{\{\theta\}}_{\{nn\}} = B^{\{\theta\}}_{\{n\bar{m}\}} = B^{\{\theta\}}_{\{\bar{m}\bar{m}\}} = B^{\{\theta\}}_{\{d\theta nn\}} = B^{\{\theta\}}_{\{d\theta n\bar{m}\}} = B^{\{\theta\}}_{\{d\theta\bar{m}\bar{m}\}} = 0, \tag{5.89}$$

$$B^{\{\phi\}}_{\{nn\}} = \frac{3\lambda v^{19}\left[(r-2)r+a^2\right]}{8\left(2va+r-3\right)^2}\Big[r(9r-25)a^2 - \frac{2(r-5)a}{v^3} + r^2\left(3r^2-13r+12\right)+6a^4\Big], \tag{5.90}$$

$$B^{\{\phi\}}_{\{n\bar{m}\}} = -\frac{3i\lambda}{8\sqrt{2}r^8\left(2va+r-3\right)^2}\Big[6va^5 + \frac{(9r-19)a^3}{v} + 2r(13-4r)a^2 + \frac{\left(3r^2-5r-10\right)a}{v^3} - 2(r-3)r^2 - 6a^4\Big], \tag{5.91}$$

$$B^{\{\phi\}}_{\{\bar{m}\bar{m}\}} = 0, \tag{5.92}$$

$$J^{\{tt\}}_{\{nn\}} = -\frac{3\lambda\left[(r-2)r+a^2\right]}{16r^8\left(2va+r-3\right)^2}\Big[\frac{2(r+5)a^3}{v} + 2r\left(r^2+2r-6\right)a^2 + \frac{2(r-3)a}{v^5} + r^4(2r-3) - a^4\Big], \tag{5.93}$$

$$J^{\{tt\}}_{\{n\bar{m}\}} = -\frac{3i\lambda a\left(r^2+a^2\right)}{8\sqrt{2}r^6}, \tag{5.94}$$

$$J^{\{tt\}}_{\{\bar{m}\bar{m}\}} = -\frac{3\lambda}{8r^6\left(2va+r-3\right)^2\left((r-2)r+a^2\right)}\Big[-\frac{2(r+4)a^5}{v} + r\left(-2r^2+6r+15\right)a^4 + \frac{4\left(r^2-2r-2\right)a^3}{v^3} + r^3\left(-2r^2+5r+4\right)a^2 + \frac{2(3r-8)a}{v^9} + r^5(9-4r) + a^6\Big], \tag{5.95}$$

$$J^{\{tr\}}_{\{nn\}} = -\frac{3\lambda\left(va+r\right)}{8r^8\left(2va+r-3\right)^2}\Big[\frac{2a^5}{v} + r(15-8r)a^4 + \frac{4(r-2)a^3}{v^3} + \frac{2(r-2)^2a}{v^5} + r^2\left(-7r^2+26r-24\right)a^2 - (r-2)^2r^3(2r-3) - 3a^6\Big], \tag{5.96}$$

$$J^{\{tr\}}_{\{n\bar{m}\}} = -\frac{3i\lambda\left(r^{3/2}+a\right)(va-1)}{8\sqrt{2}r^7(2va+r-3)^2}[(r-2)r+a^2]\Big[-\frac{2a}{v}+r(2r-3)+3a^2\Big], \tag{5.97}$$

$$J^{\{tr\}}_{\{\bar{m}\bar{n}\}} = 0, \tag{5.98}$$

$$J^{\{tr\}}_{\{drnn\}} = -\frac{3\lambda\left(r-a^2\right)(va+r)}{4r^9(2va+r-3)^2}[(r-2)r+a^2]\Big[\frac{2a}{v}+r(3-2r)-3a^2\Big], \tag{5.99}$$

$$J^{\{tr\}}_{\{drn\bar{n}\}} = \frac{3i\lambda\left(r^{3/2}+a\right)(va+1)\left[(r-2)r+a^2\right]\left[-\frac{2a}{v}+r(2r-3)+3a^2\right]}{8\sqrt{2}r^8(2va+r-3)^2}, \tag{5.100}$$

$$J^{\{tr\}}_{\{dr\bar{m}\bar{m}\}} = 0, \tag{5.101}$$

$$J^{\{t\theta\}}_{\{nn\}} = 0, \tag{5.102}$$

$$J^{\{t\theta\}}_{\{n\bar{m}\}} = -\frac{3\lambda(va+r)\left((r-2)r+a^2\right)\left(\frac{4a}{v}-r^2-3a^2\right)}{8\sqrt{2}r^7(2va+r-3)^2}, \tag{5.103}$$

$$J^{\{t\theta\}}_{\{\bar{m}\bar{n}\}} = -\frac{3i\lambda\left(r^{3/2}+a\right)(va-1)\left(-\frac{4a}{v}+r^2+3a^2\right)}{4r^6(2va+r-3)^2}, \tag{5.104}$$

$$J^{\{t\theta\}}_{\{d\theta nn\}} = 0, \tag{5.105}$$

$$J^{\{t\theta\}}_{\{d\theta n\bar{m}\}} = \frac{3i\lambda a\,(va+r)\left[(r-2)r+a^2\right]\left(\frac{4a}{v}-r^2-3a^2\right)}{8\sqrt{2}r^8\,(2va+r-3)^2}, \tag{5.106}$$

$$J^{\{t\theta\}}_{\{d\theta\bar{m}\bar{m}\}} = -\frac{3\lambda a\left(r^{3/2}+a\right)(va-1)\left(-\frac{4a}{v}+r^2+3a^2\right)}{2r^7\,(2va+r-3)^2}, \tag{5.107}$$

$$J^{\{t\phi\}}_{\{nn\}} = \frac{3\lambda v^{15}\left[(r-2)r+a^2\right]\left[va^3+(r-10)a^2+\frac{\left(r^2-6r+12\right)a}{v}+r^2(3-2r)\right]}{8\,(2va+r-3)^2}, \tag{5.108}$$

$$J^{\{t\phi\}}_{\{n\bar{m}\}} = -\frac{3i\lambda\left(r^2+2a^2\right)}{8\sqrt{2}r^6}, \tag{5.109}$$

$$J^{\{t\phi\}}_{\{\bar{m}\bar{m}\}} = \frac{3\lambda\left[(r-8)a^4\sqrt{r}-r\left(r^2-2r+15\right)a^3+\frac{2(r+4)a^2}{v^3}-r^3\left(r^2-3r+3\right)a+\frac{r-2}{v^9}-a^5\right]}{4r^6\,(2va+r-3)^2\left[(r-2)r+a^2\right]}, \tag{5.110}$$

$$J^{\{rr\}}_{\{nn\}} = \frac{3\lambda\left[2a\sqrt{r}+r(3-2r)-3a^2\right]\left[(r-2)r+a^2\right]^3}{16r^9\,(2va+r-3)^2}, \tag{5.111}$$

$$\begin{aligned} J^{\{rr\}}_{\{n\bar{m}\}} = &-\frac{3i\lambda v^{15}\,(va-1)}{8\sqrt{2}\,(2va+r-3)^2}\Bigg[\frac{2a^5}{v}+r(15-8r)a^4+\frac{4(r-2)a^3}{v^3} \\ &+r^2\left(-7r^2+26r-24\right)a^2+\frac{2(r-2)^2a}{v^5}-(r-2)^2r^3(2r-3)-3a^6\Bigg], \end{aligned} \tag{5.112}$$

$$J^{\{rr\}}_{\{\bar{m}\bar{m}\}} = -\frac{3\lambda\,(va-1)^2\left[2a\sqrt{r}+r(3-2r)-3a^2\right]\left[(r-2)r+a^2\right]}{8r^6\,(2va+r-3)^2}, \tag{5.113}$$

$$J^{\{rr\}}_{\{drnn\}} = \frac{3\lambda\left(r-a^2\right)}{4r^{10}\left(2va+r-3\right)^2}\Bigg[\frac{2a^5}{v}+r(15-8r)a^4+\frac{4(r-2)a^3}{v^3} + r^2\left(-7r^2+26r-24\right)a^2+\frac{2(r-2)^2a}{v^5}-(r-2)^2r^3(2r-3)-3a^6\Bigg], \tag{5.114}$$

$$J^{\{rr\}}_{\{drn\bar{m}\}} = \frac{3i\lambda\left[2a\sqrt{r}+r(3-2r)-3a^2\right]\left[(r-2)r+a^2\right]}{8\sqrt{2}r^9\left(2va+r-3\right)^2}\Bigg[-\frac{a^2}{v} + (r-4)ra+\frac{1}{v^5}+3a^3\Bigg], \tag{5.115}$$

$$J^{\{rr\}}_{\{dr\bar{m}\bar{n}\}} = -\frac{3\lambda\left(r-a^2\right)\left[2a\sqrt{r}+r(3-2r)-3a^2\right]\left[(r-2)r+a^2\right]}{4r^8\left(2va+r-3\right)^2}, \tag{5.116}$$

$$J^{\{rr\}}_{\{ddrnn\}} = -\frac{3\lambda\left[2a\sqrt{r}+r(3-2r)-3a^2\right]\left[(r-2)r+a^2\right]}{4r^{11}\left(2va+r-3\right)^2}\Bigg[-3(r-4)ra^2 + 2(r-3)r^2-5a^4\Bigg], \tag{5.117}$$

$$J^{\{rr\}}_{\{ddrn\bar{m}\}} = \frac{3i\lambda a\left[-a\sqrt{r}+(r-6)r+6a^2\right]\left[(r-2)r+a^2\right]\left[-2a\sqrt{r}+r(2r-3)+3a^2\right]}{4\sqrt{2}r^{10}\left(2va+r-3\right)^2}, \tag{5.118}$$

$$J^{\{rr\}}_{\{ddr\bar{m}\bar{n}\}} = -\frac{3\lambda\left(r+3a^2\right)\left[2a\sqrt{r}+r(3-2r)-3a^2\right]\left[(r-2)r+a^2\right]}{4r^9\left(2va+r-3\right)^2}, \tag{5.119}$$

$$J^{rab\theta} = 0, \tag{5.120}$$

$$J^{\{r\phi\}}_{\{nn\}} = -\frac{3\lambda v^{17}}{8\,(2va+r-3)^2}\Big[2a^5\sqrt{r} + r(15-8r)a^4 + \frac{4(r-2)a^3}{v^3} + \frac{2(r-2)^2 a}{v^5} + r^2\left(-7r^2+26r-24\right)a^2 - (r-2)^2 r^3(2r-3) - 3a^6\Big], \tag{5.121}$$

$$J^{\{r\phi\}}_{\{n\bar{m}\}} = \frac{3i\lambda\,(va-1)\left[2a\sqrt{r} + r(3-2r) - 3a^2\right]\left[(r-2)r+a^2\right]}{8\sqrt{2}r^7\,(2va+r-3)^2}, \tag{5.122}$$

$$J^{\{r\phi\}}_{\{\bar{m}\bar{m}\}} = 0, \tag{5.123}$$

$$J^{\{r\phi\}}_{\{drnn\}} = -\frac{3\lambda v^{19}\left(r-a^2\right)\left[2a\sqrt{r} + r(3-2r) - 3a^2\right]\left[(r-2)r+a^2\right]}{4\,(2va+r-3)^2}, \tag{5.124}$$

$$J^{\{r\phi\}}_{\{drn\bar{m}\}} = -\frac{3i\lambda\,(va+1)\left(2a\sqrt{r} + r(3-2r) - 3a^2\right)\left[(r-2)r+a^2\right]}{8\sqrt{2}r^8\,(2va+r-3)^2}, \tag{5.125}$$

$$J^{\{r\phi\}}_{\{dr\bar{m}\bar{m}\}} = 0, \tag{5.126}$$

$$J^{\{\theta\theta\}}_{\{nn\}} = -\frac{3\lambda}{16r^9\,(2va+r-3)^2}\Big[\frac{4a^5}{v} + r(12-7r)a^4 + \frac{8(r-2)a^3}{v^3} + r^2\left(-5r^2+16r-12\right)a^2 + \frac{4(r-2)^2 a}{v^5} - (r-2)^2 r^4 - 3a^6\Big], \tag{5.127}$$

$$J^{\{\theta\theta\}}_{\{n\bar{m}\}} = -\frac{3i\lambda\left(a-\sqrt{r}\right)\left[(r-2)r+a^2\right]\left[-4a\sqrt{r} + r^2 + 3a^2\right]}{8\sqrt{2}r^8\,(2va+r-3)^2}, \tag{5.128}$$

$$J^{\{\theta\theta\}}_{\{\bar{m}\bar{m}\}} = \frac{3\lambda\,(va-1)^2\left(4a\sqrt{r}-r^2-3a^2\right)}{8r^6\,(2va+r-3)^2}, \tag{5.129}$$

$$J^{\{\theta\theta\}}_{\{d\theta nn\}} = 0, \tag{5.130}$$

$$J^{\{\theta\theta\}}_{\{d\theta n\bar{m}\}} = -\frac{3\lambda a\left(a-\sqrt{r}\right)\left[(r-2)r+a^2\right]\left(-4a\sqrt{r}+r^2+3a^2\right)}{8\sqrt{2}r^9\,(2va+r-3)^2}, \tag{5.131}$$

$$J^{\{\theta\theta\}}_{\{d\theta\bar{m}\bar{m}\}} = -\frac{3i\lambda a\,(va-1)^2\left(4a\sqrt{r}-r^2-3a^2\right)}{4r^7\,(2va+r-3)^2}, \tag{5.132}$$

$$\begin{aligned} J^{\{\theta\theta\}}_{\{dd\theta nn\}} = &-\frac{3\lambda v^{21}a\,(va-1)}{4\,(2va+r-3)^2}\Big[-4a^5\sqrt{r}+r(7r-12)a^4-\frac{8(r-2)a^3}{v^3} \\ &+r^2\left(5r^2-16r+12\right)a^2-\frac{4(r-2)^2a}{v^5}+(r-2)^2r^4+3a^6\Big], \end{aligned} \tag{5.133}$$

$$J^{\{\theta\theta\}}_{\{dd\theta n\bar{m}\}} = \frac{3i\lambda\left[(r-2)r+a^2\right]\left(4a\sqrt{r}-r^2-3a^2\right)\left[6a^2\sqrt{r}-r(r+2)a+r^{5/2}-4a^3\right]}{8\sqrt{2}r^{10}\,(2va+r-3)^2}, \tag{5.134}$$

$$J^{\{\theta\theta\}}_{\{dd\theta\bar{m}\bar{m}\}} = \frac{3\lambda\,(va-1)^2\left(r^2+3a^2\right)\left(-4a\sqrt{r}+r^2+3a^2\right)}{4r^8\,(2va+r-3)^2}, \tag{5.135}$$

$$J^{\{\theta\phi\}}_{\{nn\}} = 0, \tag{5.136}$$

$$J^{\{\theta\phi\}}_{\{n\bar{m}\}} = \frac{3\lambda\left(3va^2 + r^{3/2} - 4a\right)\left[(r-2)r + a^2\right]}{8\sqrt{2}r^7\left(2va + r - 3\right)^2}, \tag{5.137}$$

$$J^{\{\theta\phi\}}_{\{\bar{m}\bar{m}\}} = -\frac{3i\lambda\left(va - 1\right)\left(-4a\sqrt{r} + r^2 + 3a^2\right)}{4r^6\left(2va + r - 3\right)^2}, \tag{5.138}$$

$$J^{\{\theta\phi\}}_{\{d\theta nn\}} = 0, \tag{5.139}$$

$$J^{\{\theta\phi\}}_{\{d\theta n\bar{m}\}} = -\frac{3i\lambda a\left(3va^2 + r^{3/2} - 4a\right)\left[(r-2)r + a^2\right]}{8\sqrt{2}r^8\left(2va + r - 3\right)^2}, \tag{5.140}$$

$$J^{\{\theta\phi\}}_{\{d\theta\bar{m}\bar{m}\}} = \frac{3\lambda a\left(va - 1\right)\left(4a\sqrt{r} - r^2 - 3a^2\right)}{2r^7\left(2va + r - 3\right)^2}, \tag{5.141}$$

$$J^{\{\phi\phi\}}_{\{nn\}} = \frac{3\lambda\left[(r-2)r + a^2\right]\left[2(2r-5)a\sqrt{r} + r\left(r^2 - 8r + 12\right) + a^2\right]}{16r^8\left(2va + r - 3\right)^2}, \tag{5.142}$$

$$J^{\{\phi\phi\}}_{\{n\bar{m}\}} = -\frac{3i\lambda a}{8\sqrt{2}r^6}, \tag{5.143}$$

$$J^{\{\phi\phi\}}_{\{\bar{m}\bar{m}\}} = -\frac{3\lambda\left[4(r-2)a^3\sqrt{r} + r\left(r^2 - 10r + 15\right)a^2 + 4(r-2)ar^{3/2} - (r-2)r^3 + a^4\right]}{8r^6\left(2va + r - 3\right)^2\left[(r-2)r + a^2\right]}. \tag{5.144}$$

5.C Appendix: Overlap regime of PN and BP method

To obtain the hybrid waveform between Post-Newtonian theory and Black hole perturbation method, we need to check the consistency within the overlap regime of these two methods. In other words, the PN waveform taking the mass ratio expansion should agree with the BP waveform taking the PN expansion, to the relevant orders. Technically it suffices to compare the tide-induced energy and energy flux, which we explicitly show here up to the q^{-1} and 1.5 PN order. In order to accomplish this goal, we need to expand the components in Appendix 5.B, as well as the homogeneous solutions of the Teukolsky equation with the ingoing boundary condition for $l = 2, 3$ and incident amplitudes which can be found in [373].

$$\begin{aligned}\omega R^{in}_{2m\omega} =& \frac{z^4}{30} + \frac{iz^5}{45} - \frac{11z^6}{1260} - \frac{iz^7}{420} + \frac{23z^8}{45360} + \frac{iz^9}{11340} \\ &+ \epsilon\left(-\frac{z^3}{15} - \frac{iz^4}{60} - \frac{41z^5}{3780} - \frac{31iz^6}{3780}\right) + \epsilon^2\left(\frac{z^2}{30} + \frac{iz^3}{60}\right) \qquad (5.145)\\ \omega R^{in}_{3m\omega} =& \frac{z^5}{630} + \frac{iz^6}{1260} - \frac{z^7}{3780} - \frac{iz^8}{16200} + \epsilon\left(-\frac{z^4}{252} - \frac{iz^5}{756}\right) \qquad (5.146)\\ B^{in}_{2m\omega} =& \frac{i}{8\omega^2}\left\{1 - \epsilon\frac{\pi}{2} + i\epsilon\left(\frac{5}{4} - \gamma - \log 2\right) + O(\epsilon^2)\right\} \qquad (5.147)\\ B^{in}_{3m\omega} =& -\frac{i}{8\omega^2}\left\{1 - \epsilon\frac{\pi}{2} + i\epsilon\left(\frac{13}{6} - \gamma - \log 2\right) + O(\epsilon^2)\right\} \qquad (5.148)\end{aligned}$$

where $z = \omega r$ and $\epsilon = 2M\omega$. With these equations and components in Appendix 5.B, we can obtain the energy flux up to the 1.5PN order from Eq. (5.48):

$$E = \eta\,(M\Omega)^{2/3}\left[-\frac{1}{2} + \frac{3}{8}(M\Omega)^{2/3} + \frac{9}{2}\lambda(M\Omega)^{10/3} + \frac{33}{4}\lambda(M\Omega)^4\right] \qquad (5.149)$$

$$\frac{dE}{dt} = -\frac{5}{32\eta^2\,(M\Omega)^{10/3}}\left[1 - \frac{1247}{336}(M\Omega)^{2/3} + 18\lambda(M\Omega)^{10/3} - \frac{704}{28}\lambda(M\Omega)^4\right] \qquad (5.150)$$

which are same as the corresponding PN result by keeping only the η^{-1} order term[90]. According to Eq. (5.63), we know that in the overlap regime the Post-Newtonian and Black Hole Perturbation

methods are consistent.

Part III

Formation and Rate Study of Extreme Mass Ratio Inspirals

Chapter 6

Wet Extreme Mass Ratio Inspirals

Zhen Pan, Zhenwei Lyu, Huan Yang

Phys. Rev. D 104, 063007 (Sept. 2021), arXiv:2104.01208

My specific contributions: Calculate EMRI rates for all models considered in Table 6.1. Generate MBH samples from Eqs. (6.24) and EMRI samples from Eqs. (6.25), then Calculate SNR of all models with LISA sensitivity and evaluate LISA detectable rate shown in Table 6.1. Demonstrate LISA detectable dry and wet EMRI rates per mass bin per year in Fig. 6.2.

Extreme Mass Ratio Inspirals (EMRIs) can be classified as dry EMRIs and wet EMRIs based on their formation mechanisms. Dry (or the "loss-cone") EMRIs, previously considered as the main EMRI sources for the Laser Interferometer Space Antenna, are primarily produced by multi-body scattering in the nuclear star cluster and gravitational capture. In this work, we highlight an alternative EMRI formation channel: (wet) EMRI formation assisted by the accretion flow around accreting galactic-center massive black holes (MBHs). In this channel, the accretion disk captures stellar-mass black holes that are intially moving on inclined orbits, and subsequently drives them to migrate towards the MBH - this process boosts the formation rate of EMRIs in such galaxies

by orders of magnitude. Taking into account the fraction ($O(10^{-2} - 10^{-1})$) of active galactic nuclei where the MBHs are expected to be rapidly accreting, we forecast that wet EMRIs will contribute an important or even dominant fraction of all detectable EMRIs by spaceborne gravitational wave detectors.

6.1 Introduction

The primary astrophysical sources for space-based gravitational wave detectors, such as Laser Interferometer Space Antenna (LISA) [10] and TianQin [11], include massive black hole (MBH) bianaries, extreme mass ratio inspirals (EMRIs) [130], galactic binaries and stellar-mass black hole (sBH) binaries. Other systems, e.g. intermediate mass ratio inspirals [374, 375], extremely large mass ratio inspirals [376, 377] and cosmic strings [378], may also be detectable, *albeit* with larger uncertainties. Among these sources, EMRIs provide unique opportunities in testing the Kerr spacetime [379, 380], probing the galactic-center cluster distribution [132, 134, 140], understanding the astrophysical environmental effects [381–383], and inferring the growth history of MBHs [384–386]. Loud EMRIs can serve as dark standard sirens for measuring the Hubble constant H_0 and the dark energy equation of state [387].

EMRI formation mechanism can be classified into two main channels. In the "dry EMRI" channel, an EMRI may be produced after a sBH is gravitationally captured by a MBH, following the multi-body scatterings within the nuclear cluster [135, 137, 388] (other processes involving tidal disruption or tidal capture of binary sBHs, or tidal stripping of giant stars [389–392] may also contribute a fraction of dry EMRIs). There are two characteristic timescales [131, 133] in this process: the GW emission timescale $t_{\rm gw}$ on which the sBH orbit shrinks and the relaxation timescale t_J on which the orbital angular momentum of the sBH changes, due to scatterings by stars and other sBHs. If $t_{\rm gw} > t_J$, the sBH will likely be randomly scattered either into or away from the MBH (sBHs scattered into the MBH are known as prompt infalls). If $t_{\rm gw} < t_J$, the sBH

orbit gradually spirals into the MBH to form an EMRI while random scatterings are negligible. The generic rate can be obtained by solving the Fokker-Planck equation or by N-body simulations [132, 134, 140], subject to assumptions on the initial distributions of stars and sBHs in the nuclear cluster. In addition to the generic rate per MBH, the EMRI rate density in the universe also depends on the mass function of MBHs, the fraction of MBHs living in stellar cusps and the relative abundance of sBHs in stellar clusters. Taking into account these astrophysical uncertainties, Babak et al. [137, 388] and Fan et al. [139] forecasted that there will be a few to thousands of detectable (dry) EMRIs per year by LISA and TianQin, respectively. In a recent paper [142], Zwick et al. reanalyzed the GW emission timescales of inspiraling eccentric binaries and realized Post-Newtonian (PN) corrections to the commonly used Peters' formula [393] are necessary. With PN corrections implemented, the dry EMRI rate decreases by approximately at least one order of magnitude [394].

Wet EMRIs come from MBHs in gas-rich environments, where the distributions of nearby stars and sBHs are significantly affected by the accretion flow. About 1% low-redshift ($z \lesssim 1$) galaxies and 1% – 10% high-redshift ($1 \lesssim z \lesssim 3$) galaxies are active [144, 145] and known as active galactic nuclei (AGNs), in which galactic MBHs are believed to be rapdidly accreting gas in a disk configuration. In the presence of an accretion disk, the periodic motion of a sBH generally generates density waves which in turn affect the sBH's motion by damping both the orbital inclination with respect to the disk plane and the orbital eccentricity, and driving the sBH's migration in the radial direction [146–149]. As long as the sBH is captured onto the disk, the density waves together with other disk-sBH interactions, e.g., head wind [152, 153], accelerate its inward migration until the vicinity of the MBH where GW emissions become prevalent. In addition to sBHs captured onto the disk, star formation and subsequent birth of sBHs in the AGN disk may also contribute to wet EMRI formation [395–397]. In this paper, we show that an accretion disk usually boosts the EMRI intrinsic rate per individual MBH by orders of magnitude compared with

the loss-cone channel [1]. In particular, we suggest that wet EMRI formation is an important or even dominant channel for all observable EMRIs by spaceborne GW detectors.

The remaining part of this paper is organized as follows. In Sec. 6.2, we summarize the interactions of AGN disks with sBHs and stars. In Sec. 6.3, we introduce the Fokker-Planck equation which governes the evolution of sBHs and stars in a cluster with or without the presence of an AGN disk. In Sec. 6.4, we present the generic dry EMRI rate per MBH and the wet EMRI rate per AGN. In Sec. 6.5, we calculate the LISA detectable EMRI rate from both channels, and we discuss the applications of wet EMRIs in Sec. 6.6.

Throughout this paper, we will use geometrical units $G = c = 1$ and assume a flat ΛCDM cosmology with $\Omega_{\rm m} = 0.307, \Omega_\Lambda = 1 - \Omega_{\rm m}$ and $H_0 = 67.7$ km/s/Mpc.

6.2 Disk-sBH and Disk-Star Interactions

In addition to the gravitational forces from the MBH and the stars/sBHs in the cluster, the orbital motion of a sBH around an accreting MBH is influenced by disk-sBH interactions: density waves, head wind [146–149, 152] and other sub-dominant interactions including dynamic friction [398, 399] and heating torque [400, 401].

As a sBH orbits around the MBH, its periodic motion excites density waves consisting of three components [148, 149]: regular density waves arising from the circular motion, eccentricity waves arising from the non-circular motion and bending waves arising from the motion normal to the disk. The density waves in turn affect the motion of the sBH: the regular density waves exert a (type-I) migration torque on the sBH and drives its migration in the radial direction on the timescale $t_{\rm mig,I}$; the eccentricity and bending density waves damp the orbit eccentricity and the inclination with respect to the disk plane on the timescale $t_{\rm wav}$. Previous analytic studies [148, 149] calibrated with

[1] More details can be found in a companion paper [140]

numerical simulations [402] show that the type-I migration torque can be formulated as

$$\dot{J}_{\rm mig,I} = C_{\rm I}\frac{m_{\rm bh}}{M}\frac{\Sigma}{M}\frac{r^4\Omega^2}{h^2}\ , \tag{6.1}$$

where $m_{\rm bh}$ is sBH mass, and $M = M(< r)$ is the total mass of the MBH, stars, sBHs and the disk within radius r; the prefactor $C_{\rm I} = -0.85 + d\log\Sigma/d\log r + 0.9\ d\log T_{\rm mid}/d\log r$ depends on the disk profile; $\Sigma(r), T_{\rm mid}(r), h(r), \Omega(r)$ are the disk surface density, the disk middle plane temperature, the disk aspect ratio and the sBH angular velocity, respectively. The corresponding migration timescale and damping timescale are

$$t_{\rm mig,I} = \frac{J}{|\dot{J}_{\rm mig,I}|} \sim \frac{M}{m_{\rm bh}}\frac{M}{\Sigma r^2}\frac{h^2}{\Omega}, \quad t_{\rm wav} = \frac{M}{m_{\rm bh}}\frac{M}{\Sigma r^2}\frac{h^4}{\Omega}\ , \tag{6.2}$$

where $J = r^2\Omega$ is the specific angular momentum of the sBH, and $t_{\rm wav} \approx t_{\rm mig,I}h^2$, i.e., the eccentricity/inclination damping is faster than the migration by a factor h^2. Therefore the orbit should become circular long before the sBH migrate into the LISA band. A gap in the disk opens up if the sBH is so massive that its tidal torque removes surrounding gas faster than the gas replenishment via viscous diffusion. After a gap is opened, the type-I migration turns off and the sBH is subject to type-II migration driven by a type-II migration torque $\dot{J}_{\rm mig,II}$ [403].

For a sBH embedded in the gas disk, surrounding gas in its gravitational influence sphere flows towards it. Considering the differential rotation of the disk, the inflow gas generally carries nonzero angular momentum relative to the sBH, so that the inflow tends to circularize and form certain local disk or buldge profile around the sBH. Depending on the radiation feedback and magnetic fields, a major part of captured materials may escape in the form of outflow and only the remaining part is accreted by the sBH [404, 405]. Because of the circularization process, it is reasonable to expect that the outflow carries minimal net momentum with respect to the sBH. As a result, the head wind in the influence sphere of the sBH is captured, and the momentum carried by the wind eventually

transfers to the sBH. Therefore the specific torque exerted on the sBH from the head wind is

$$\dot{J}^{\rm id}_{\rm wind} = -\frac{r\delta v_\phi \dot{m}_{\rm gas}}{m_{\rm bh}} \,, \tag{6.3}$$

where the upper index "id" denotes in-disk objects, $\delta v_\phi := v_{\phi,\rm gas} - v_{\phi,\rm bh}$ is the head wind speed, and $\dot{m}_{\rm gas}$ is the amount of gas captured per unit time (see [140] for detailed calculation).

In summary, the migration timescales of in-disk (id) sBHs and those outside (od) are

$$t^{\rm bh,id}_{\rm mig} = \frac{J}{|\dot{J}_{\rm mig,I,II} + \dot{J}_{\rm gw} + \dot{J}_{\rm wind}|}, \quad t^{\rm bh,od}_{\rm mig} = \frac{J}{|\dot{J}_{\rm mig,I} + \dot{J}_{\rm gw}|} \,, \tag{6.4}$$

where $\dot{J}_{\rm mig,I,II} = \dot{J}_{\rm mig,I}$ or $\dot{J}_{\rm mig,II}$ and $\dot{J}_{\rm wind} = \dot{J}^{\rm id}_{\rm wind}$ [Eq. (6.3)] or 0, depending on whether a gap is open. The specific torque arising from GW emissions is [2]

$$\dot{J}_{\rm gw} = -\frac{32}{5}\frac{m_{\rm bh}}{M}\left(\frac{M}{r}\right)^{7/2} \,. \tag{6.5}$$

The damping timescale of sBH orbital inclination and eccentricity is given by Eq. (6.2)

$$t^{\rm bh,od}_{\rm wav} = \frac{M}{m_{\rm bh}}\frac{M}{\Sigma r^2}\frac{h^4}{\Omega} \,. \tag{6.6}$$

The above discussion of disk-sBH interactions also equally applies to stars in the cluster, except stars are usually lighter ($m_{\rm star} < m_{\rm bh}$), and the head wind impact on stars is weak ($\dot{J}^{\rm star}_{\rm wind} \approx 0$) considering that the wind could be largely suppressed in the presence of star radiation feedback and solar wind [406, 407]. Becuse the structure of AGN disks has not been fully understood, we consider three commonly used AGN disk models: α-disk, β-disk [408] and TQM disk [409] in this work.

[2] GW emission turns out to has little influence on the wet EMRI rate which is determined by the capture and the migration of sBHs at large separations where GW emission is negligible. We include the GW emission in the calculation because in some special cases the migration torque $\dot{J}_{\rm mig,I}$ changes its sgin near the MBH and the GW emission works to overcome the would-be migration trap [140].

6.3 Fokker-Planck Equation

Statistical properties of stars and sBHs in the stellar cluster are encoded in their distribution functions $f_i(t, E, R)$ (i = star/bh) in the phase space, where

$$E := \phi(r) - v^2(r)/2\,, \quad R := J^2/J_c^2(E) \tag{6.7}$$

are the specific orbital (binding) energy and the normalized orbital angular momentum, respectively. Here $\phi(r)$ is the (positive) gravitational potential, v is the orbital speed, and $J_c(E)$ is the specific angular momentum of a circular orbiter with energy E. Given initial distributions $f_i(t = 0, E, R)$, the subsequent evolution is governed by the orbit-averaged Fokker-Planck equation. In the case of no gas disk, the Fokker-Planck equation (for both stars and sBHs) is formulated as [410–412]

$$C\frac{\partial f}{\partial t} = -\frac{\partial}{\partial E}F_E - \frac{\partial}{\partial R}F_R\,, \tag{6.8}$$

where $f = f_i(t, E, R)$, $C = C(E, R)$ is a normalization coefficient, and $F_{E,R}$ is the flux in the E/R direction:

$$\begin{aligned} -F_E &= \mathcal{D}_{EE}\frac{\partial f}{\partial E} + \mathcal{D}_{ER}\frac{\partial f}{\partial R} + \mathcal{D}_E f\,, \\ -F_R &= \mathcal{D}_{RR}\frac{\partial f}{\partial R} + \mathcal{D}_{ER}\frac{\partial f}{\partial E} + \mathcal{D}_R f\,, \end{aligned} \tag{6.9}$$

where the diffusion coefficients $\{\mathcal{D}_{EE}, \mathcal{D}_{ER}, \mathcal{D}_{RR}\}_i$ and the advection coefficients $\{\mathcal{D}_E, \mathcal{D}_R\}_i$ are functions of $f_i(t, E, R)$ [410–412]. From flux $\{F_E, F_R\}_{\rm bh}$, we can compute the EMRI rate via the lose cone mechanism as

$$\Gamma_{\rm lc}(t) = \int_{E>E_{\rm gw}} \vec{F}\cdot d\vec{l}\,, \tag{6.10}$$

where $\vec{F} = (F_E, F_R)$, $d\vec{l} = (dE, dR)$ is the line element along the boundary of the loss cone, and $E_{\rm gw}$ is a characteristic energy scale above which the sBH GW emission is dominant with $t_{\rm gw} < t_J$ [131, 132, 134, 140, 142, 394].

In the presence of an AGN disk, stars and sBHs settle as two components: a cluster component and a disk component. We expect the distribution functions of cluster-component stars and sBHs acquire some dependence on the orbital inclination as interacting with the disk. For convenience, we choose to integrate out the inclination and work with the inclination-integrated distribution functions $f_i(t, E, R)$ of the cluster-component stars and sBHs. Considering the density waves excited on the disk to damp the orbital inclinations and eccentricities of orbiters, and to drive the orbiters' inward migration together with head winds and GW emissions, we rewrite the Fokker-Planck equation as

$$C\frac{\partial f}{\partial t} = -\frac{\partial}{\partial E}F_E - \frac{\partial}{\partial R}F_R + S \ , \tag{6.11}$$

where flux $F_{E,R}$ are defined in Eq. (6.9), with the advection coefficients modified by disk-star/sBH interactions as

$$\mathcal{D}_{E,\rm bh} \to \mathcal{D}_{E,\rm bh} - C\frac{E}{t_{\rm mig}^{\rm bh,od}} \ , \qquad \mathcal{D}_{R,\rm bh} \to \mathcal{D}_{R,\rm bh} - C\frac{1-R}{t_{\rm wav}^{\rm bh,od}} \ ,$$
$$\mathcal{D}_{E,\rm star} \to \mathcal{D}_{E,\rm star} - C\frac{E}{t_{\rm mig}^{\rm star,od}} \ , \qquad \mathcal{D}_{R,\rm star} \to \mathcal{D}_{R,\rm star} - C\frac{1-R}{t_{\rm wav}^{\rm star,od}} \ ,$$

and the negative source term $S = S_i(t, E, R)$ arising from spherical-component stars/sBHs captured onto the disk is parameterized as

$$S_{\rm bh} = -\mu_{\rm cap} C\frac{f_{\rm bh}}{t_{\rm mig}^{\rm star,id}} \ , \quad S_{\rm star} = -\mu_{\rm cap}\frac{m_{\rm star}}{m_{\rm bh}} C\frac{f_{\rm star}}{t_{\rm mig}^{\rm star,id}} \ , \tag{6.12}$$

with $\mu_{\rm cap} \in [h, 1]\frac{m_{\rm bh}}{m_{\rm star}}$ a phenomenological parameter quantifying the disk capture efficiency (see [140] for more details). A new EMRI forms if a sBH is captured onto the disk and migrate to the vicinity of the MBH within the disk lifetime $T_{\rm disk}$, therefore the EMRI rate assisted by the AGN disk is given by

$$\Gamma_{\rm disk}(t; T_{\rm disk}) = \int \int_{t_{\rm mig}^{\rm bh,id} < T_{\rm disk}} -S_{\rm bh}(t, E, R)\, dE dR \ . \tag{6.13}$$

6.4 EMRI Rate per MBH/AGN: dry and wet

6.4.1 Dry EMRIs

Given initial distributions of stars and sBHs in the stellar cluster, one can evolve the system according to the Fokker-Planck equation (6.8) and calculate the EMRI rate in the loss cone channel using Eq. (6.10). As shown in Refs. [132, 134, 140], the EMRI rate mainly depends on the total number of stars within the MBH influence radius, which determines the relaxation timescale and the relative abundance of sBHs in the stellar cluster. Following Ref. [137], the time-averaged EMRI rate per MBH can be parameterized as

$$\Gamma_{\rm dry}(M_\bullet; N_p) = C_{\rm dep}(M_\bullet; N_p) C_{\rm grow}(M_\bullet; N_p) \Gamma_{\rm lc}(M_\bullet) \,, \tag{6.14}$$

with

$$\Gamma_{\rm lc}(M_\bullet) = 30 \left(\frac{M_\bullet}{10^6 M_\odot} \right)^{-0.19} \ {\rm Gyr}^{-1} \,, \tag{6.15}$$

where N_p is the average number of prompt infalls per EMRI; $C_{\rm dep}$ and $C_{\rm grow}$ are correction factors accounting for possible depletion of sBHs in the cusp as sBHs accreted by the MBH and capping the maximum MBH growth via accreting sBHs, respectively, and the loss-cone EMRI rate in Eq. (6.15) is lower than previous calculations [132, 134, 137, 140] by one order of magnitude because these previous results were based on the Peters' formula [393] which underestimate the GW emission timescales of eccentric binaries and the true EMRI rate should be lower by approximately at least one order of magnitude [142, 394].

Following Ref. [137], we explain the two corrections $C_{\rm dep}(M_\bullet; N_p)$ and $C_{\rm grow}(M_\bullet; N_p)$ to the generic dry EMRI rate. Consider a MBH with mass $M_\bullet$, whose influence sphere ($r < r_c = 2M_\bullet/\sigma^2$) encloses a number of sBHs with total mass $\Sigma m_{\rm bh} \simeq 0.06 M_\bullet$, and these sBHs will be depleted by

the MBH via EMRIs and prompt infalls on a timescale

$$T_{\rm dep}(r_c) = \frac{\sum m_{\rm bh}}{(1+N_p)\Gamma_{\rm lc}(M_\bullet)m_{\rm bh}} = \frac{200}{1+N_p}\left(\frac{m_{\rm bh}}{10M_\odot}\right)^{-1}\left(\frac{M_\bullet}{10^6 M_\odot}\right)^{1.19}\ {\rm Gyr}\ , \tag{6.16}$$

where N_p is the average number of prompt infalls per EMRI. On the influence sphere, the relaxation timescale of the star cluster is approximately [412]

$$T_{\rm rlx}(r_c) \simeq \left(\frac{\sigma}{20{\rm km/s}}\right)\left(\frac{r_c}{1{\rm pc}}\right)^2\ {\rm Gyr}\ , \tag{6.17}$$

where the velocity dispersion is related to the MBH mass by the famous $M_\bullet - \sigma$ relation [413]. The depletion correction $C_{\rm dep}$ is defined as

$$C_{\rm dep}(M_\bullet; N_p) := {\rm min.}\left\{\frac{T_{\rm dep}}{T_{\rm rlx}}, 1\right\}\ , \tag{6.18}$$

where

$$\frac{T_{\rm dep}}{T_{\rm rlx}} \simeq \frac{12}{1+N_p}\left(\frac{m_{\rm bh}}{10M_\odot}\right)^{-1}\left(\frac{M_\bullet}{10^6 M_\odot}\right)^{0.06}\ . \tag{6.19}$$

The growth correction

$$C_{\rm grow} := {\rm min.}\left\{e^{-1}\frac{M_\bullet}{\Delta M_\bullet}, 1\right\} \tag{6.20}$$

arises from requiring the MBH grows no more than e^{-1} via accreting sBHs, where

$$\Delta M_\bullet = m_{\rm bh}(1+N_p)C_{\rm dep}(M_\bullet; N_p)\Gamma_{\rm lc}(M_\bullet)T_{\rm emri}(M_\bullet)\ , \tag{6.21}$$

is the MBH growth via accreting sBHs, and

$$T_{\rm emri}(M_\bullet) = \int dz \frac{dt}{dz} C_{\rm cusp}(M_\bullet, z) \tag{6.22}$$

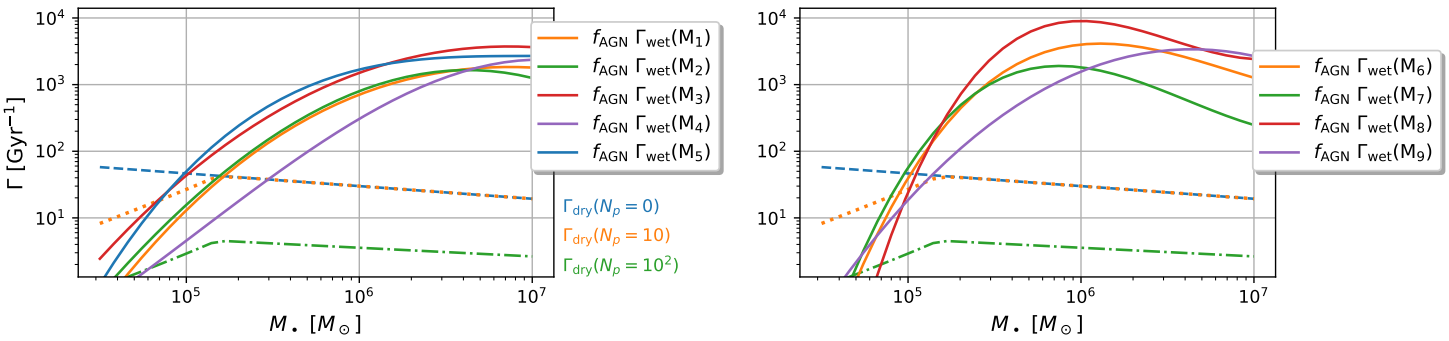

Figure 6.1: Average EMRI rates per MBH in the loss cone channel $\Gamma_{\rm dry}(M_\bullet; N_p)$ and per AGN in the disk channel $\Gamma_{\rm wet}(M_\bullet; \mathbb{M})$, where N_p is the number of prompt infalls per EMRI, and $\mathbb{M}$ consists of all model parameters of initial condition of stellar clusters, AGN duty cycles and AGN disk model, where the AGN fraction is $f_{\rm AGN} = 1\%$.

is the effective growth time when the MBH lives in a stellar cusp.

In Fig. 6.1, we show 3 sample models of dry EMRIs with $N_p = \{0, 10, 10^2\}$, where $\Gamma_{\rm dry}(N_p = 0)$ is the same as the generic rate [Eq. (6.15)] in the mass range of interest, $\Gamma_{\rm dry}(N_p = 10)$ is capped by the accretion growth limit $C_{\rm grow}$ for light MBHs, and $\Gamma_{\rm dry}(N_p = 10^2)$ is further reduced by the sBH depletion $C_{\rm dep}$ across the entire mass range.

6.4.2 Wet EMRIs

More technical complications are involved in calculating the wet EMRI rate due to the uncertainties in AGN accretion history, AGN accretion disks and initial conditions of stellar clusters, which we outline as follows in accordance with our previous work [140].

In this paper, we conservatively assume a constant AGN fraction $f_{\rm AGN} = 1\%$ throughout the universe, though it can be 10 times higher [144, 145]. Being consistent with the AGN fraction, the total duration of active phases of an AGN is about 10^8 yr [414], therefore MBHs are in quiet phase most of the time. Another complication is that AGN accretion is likely episodic [415, 416], i.e., a MBH may become active for multiple times during its whole life. Without detailed knowledge

of the duty cycle of an MBH, we simplify it as a long quiet phase of $T_0 = 5$ Gyr followed by a short active phase of $T_{\rm disk} = 10^7$ or 10^8 yr. For $T_{\rm disk} = 10^8$ yr, there is only one active phase. On the other hand, there are on average 10 active phases for models with $T_{\rm disk} = 10^7$ yr, and we only consider the following two extremal cases. If a (low-redshift) AGN that has gone through all the 10 active phases and the relaxation between different active phases is not expected to substantially change the sBH distribution, the average EMRI rate is approximately same to that in the case of $T_{\rm disk} = 10^8$ yr. If a (high-redshift) AGN that has gone through only 1 active phase, the duty cycle is simply a long quiet phase with duration $T_0 = 5$ Gyr followed by a short active phase with duration $T_{\rm disk} = 10^7$ yr. That is to say, $T_{\rm disk}$ in our model is approximately the total duration of *all* active phases an AGN has gone through.

The structure of AGN disks has not been fully understood either, partially due to the large range over which an AGN disk extends: from an inner radius of a few gravitational radii of the MBH to the outer radius of parsec scale where the AGN disk connects to the galactic gas disk. Three commonly used AGN disk models: α-disk, β-disk [408] and TQM disk [409], are different in their prescriptions of disk viscosity and/or disk heating mechanism which lead to large differences in predicted disk structures. Each disk model is specified by two model parameters, the MBH accretion rate $\dot{M}_\bullet$ and a viscosity parameter [140]: an α parameter which prescribes the ratio between the viscous stress and the local total/gas pressure in the α/β-disk and a X parameter which prescribes the ratio between the radial gas velocity and the local sound speed in the TQM disk. For calculating the wet EMRI rate, we consider both an α-disk model [408] with the viscosity parameter $\alpha = 0.1$ and accretion rate $\dot{M}_\bullet = 0.1\dot{M}_\bullet^{\rm Edd}$, and a TQM disk [409] with the viscosity parameter $X = 0.1$ and accretion rate $\dot{M}_\bullet = 0.1\dot{M}_\bullet^{\rm Edd}$ (β-disk is different from α-disk only in the inner region where radiation pressure dominates over gas pressure, and this difference has little impact on the wet EMRI rate).

For calculating the wet EMRI rate, we also need to specify the initial distributions of stars and sBHs in the stellar cluster, which we assume the commonly used Tremaine's cluster model [417]

with a power-law density profile $n_{\rm star}(r) \sim r^{-\gamma}$ deep inside the influence sphere of the MBH and $n_{\rm star}(r) \sim r^{-4}$ far outside, and sBHs are of the same density profile with a relative abundance δ.

Given initial distributions of stars and sBHs in the stellar cluster, we first evolve the system for time T_0 according to the Fokker-Planck equation (6.8), then turn on an accretion disk and continue the evolution for time $T_{\rm disk}$, according to the modified Fokker-Planck equation (6.11). In the active phase, the disk assisted EMRI rate is computed using Eq. (6.13). We show the time-averaged EMRI rate per AGN

$$\Gamma_{\rm wet}(M_\bullet; \mathbb{M}) = \frac{1}{T_{\rm disk}} \int_{T_0}^{T_0+T_{\rm disk}} \Gamma_{\rm disk}(t, M_\bullet; \mathbb{M}) dt \,, \tag{6.23}$$

for different models $\mathbb{M}$ in Fig. 6.1, where $\mathbb{M}$ denotes models parameterizing initial distributions of stars and sBHs in the cluster, duty cycles of MBHs and AGN disk model (see Table 6.1 for model parameters for all the 9 models considered in this work).

Because sBHs are captured onto the disk and migrate inward efficiently, and the sBH loss via prompt infalls is negligible ($N_p \ll 1$), the wet EMRI rate is mainly limited by the number of sBHs available in the stellar cluster. As a result, we find the presence of an AGN disk usually boosts the EMRI formation rate by orders of magnitude [140, 418] regardless of the variations of different disk models considered.

6.5 Total and LISA Detectable EMRI Rates

For calculating the total EMRI rate, we consider two redshift-independent MBH mass functions in the range of $(10^4, 10^7) M_\odot$,

$$\begin{aligned} f_{\bullet,-0.3} &: \frac{dN_\bullet}{d\log M_\bullet} = 0.01 \left(\frac{M_\bullet}{3\times 10^6 M_\odot}\right)^{-0.3} \ {\rm Mpc}^{-3} \,, \\ f_{\bullet,+0.3} &: \frac{dN_\bullet}{d\log M_\bullet} = 0.002 \left(\frac{M_\bullet}{3\times 10^6 M_\odot}\right)^{+0.3} \ {\rm Mpc}^{-3} \,, \end{aligned} \tag{6.24}$$

where the former one is approximate to the mass function as modelled in Refs. [419–422] assuming MBHs were seeded by Population III stars and accumulated mass via mergers and gas accretion along cosmic history, and the latter one is a phenomenological model [385]. The differential EMRI rates (in observer's frame) in the two formation channels are written as

$$\begin{aligned}\frac{d^2\mathcal{R}_{\rm dry}}{dM_\bullet dz} &= \frac{1}{1+z}\frac{dN_\bullet}{dM_\bullet}\frac{dV_{\rm c}(z)}{dz}C_{\rm cusp}(M_\bullet,z)\Gamma_{\rm dry}(M_\bullet;N_p)\ ,\\ \frac{d^2\mathcal{R}_{\rm wet}}{dM_\bullet dz} &= \frac{f_{\rm AGN}}{1+z}\frac{dN_\bullet}{dM_\bullet}\frac{dV_{\rm c}(z)}{dz}C_{\rm cusp}(M_\bullet,z)\Gamma_{\rm wet}(M_\bullet;\mathbb{M})\ ,\end{aligned} \tag{6.25}$$

where the factor $1/(1+z)$ arises from the cosmological redshift, $V_{\rm c}(z)$ is the comoving volume of the universe up to redshift z, $C_{\rm cusp}(M_\bullet,z)$ is the fraction of MBHs living in stellar cusps which are supposed to be evacuated during mergers of binary MBHs and re-grow afterwards [419–422]. For cases with mass function $f_{\bullet,+0.3}$, we use the same $C_{\rm cusp}$ function as in [137] and we simply take $C_{\rm cusp}=1$ for cases with phenomenological mass function $f_{\bullet,+0.3}$ [3].

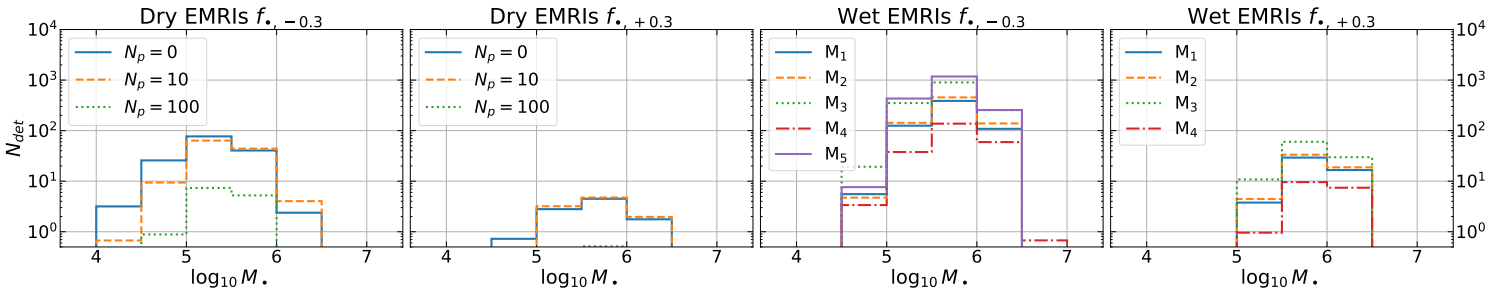

Figure 6.2: Forecasted LISA detectable dry and wet EMRI rates $N_{\rm det}$ per mass bin ($M_\bullet[M_\odot]$) per year for different models, where $f_{\bullet,\pm0.3}$ are the two different MBH mass functions [Eq. (6.24)], N_p is the number of prompt infalls per EMRI in the loss cone channel, and the wet EMRI model parameters of $\mathrm{M}_{1,\ldots,5}$ are detailed in Table 6.1.

In order to calculate the LISA detectable EMRI rate in each channel, we construct a population of EMRIs with sBH mass $m_{\rm bh}=10M_\odot$, MBH spin $a=0.98$, and MBH masses and redshifts ran-

[3] There was some misplot of the $C_{\rm cusp}(M_\bullet,z)$ function in [137], and we thank Alberto Sesana for kindly providing the correct one.

domly sampled according to the differential EMRI rates [Eq. (6.25)]. For each individual EMRI, we need 10 more parameters to uniquely specify its binary configuration at coalescence and its gravitational waveform [154–156]: sky localization $\hat{\boldsymbol{n}}$, MBH spin direction $\hat{\boldsymbol{a}}$, 3 phase angles, coalescence time t_0, inclination angle ι_0 and eccentricity e_0 at coalescence. For both dry and wet EMRIs, we assume the sky locations and the MBH spin directions are isotropically distributed on the sphere, 3 phase angles are uniformly distributed in $[0, 2\pi]$, coalescence times are randomly sampled from $[0, 2]$ yr, and cosines of inclination angles are randomly sampled from $[-1, 1]$. Distributions of eccentricity e_0 are different: uniform disbribution of e_0 in $[0, 0.2]$ for dry EMRIs v.s. $e_0 = 0$ for wet EMRIs.

For each EMRI, we compute its time-domain waveform $h_{+,\times}(t)$ using the Augment Analytic Kludge (AAK) [154–156] with the conservative Schwarzschild plunge condition, because the PN corrections used for constructing the AAK waveform model are increasingly inaccurate as the orbital separation decreases. Extending the waveform to the Kerr last stable orbit likely leads to an overestimate of the signal-to-noise ratio (SNR) [137, 156].

The SNR is calculated as a noise weighted inner product in the frequency domain [423]

$$\mathrm{SNR} = \sqrt{4 \int_0^\infty \frac{h_+(f) h_+^*(f) + h_\times(f) h_\times^*(f)}{S_\mathrm{n}(f)}\, df}\,, \tag{6.26}$$

where $S_\mathrm{n}(f)$ is the sky-averaged detector sensitivity of LISA [137, 424]. The expected LISA detectable EMRI rates (SNR$\geq$ 20) of different models in each mass bin are shown in Fig. 6.2, and the total event rates and the LISA detectable rates are collected in Table 6.1. From Fig. 6.2 and Table 6.1, wet EMRI formation is evidently an important or even dominant channel for all the models we have considered.

Table 6.1: Comparison of dry and wet EMRI rates in different models, where $f_\bullet$ is the MBH mass function. The last two columns are the total EMRI rate in the redshift range of $0 < z < 4.5$ and the corresponding LISA detectable (SNR$\geq$ 20) rate.

Dry	$f_\bullet$	N_p				Total [yr^{-1}]	LISA detectable [yr^{-1}]
	$f_{\bullet,-0.3}$	0				3500	150
		10				1300	120
		10^2				150	14
	$f_{\bullet,+0.3}$	0				160	10
		10				130	10
		10^2				15	1
Wet	$f_\bullet$	$\mathbb{M}$:	(γ, δ)	$\mu_{\rm cap}$	$(T_{\rm disk}$ [yr], $f_{\rm AGN})$	AGN Disk	Total [yr^{-1}]	LISA detectable [yr^{-1}]
	$f_{\bullet,-0.3}$	$\rm M_1$:	(1.5, 0.001)	1	$(10^8, 1\%)$	α-disk	11000	600
		$\rm M_2$:	(1.5, 0.001)	0.1			11000	760
		$\rm M_3$:	(1.5, 0.002)	1			24000	1500
		$\rm M_4$:	(1.8, 0.001)	1			8100	240
		$\rm M_5$:	(1.5, 0.001)	1	$(10^8, 1\%)$	TQM disk	23000	1900
		$\rm M_6$:	(1.5, 0.001)	1	$(10^7, 1\%)$	α-disk	39000	4200
		$\rm M_7$:	(1.5, 0.001)	0.1			21000	3000
		$\rm M_8$:	(1.5, 0.002)	1			80000	9800
		$\rm M_9$:	(1.8, 0.001)	1			22000	1400
	$f_{\bullet,+0.3}$	$\rm M_1$:	(1.5, 0.001)	1	$(10^8, 1\%)$	α-disk	2100	49
		$\rm M_2$:	(1.5, 0.001)	0.1			2000	57
		$\rm M_3$:	(1.5, 0.002)	1			4300	100
		$\rm M_4$:	(1.8, 0.001)	1			1900	18

6.6 Applications of wet EMRIs

Due to the high LISA sensitivity to the EMRI eccentricity whose value at coalescence can be measured with typical uncertainty as low as 10^{-5} [137], wet EMRIs can be distinguished from dry ones via eccentricity measurements, as wet EMRIs are expected to be circular in the LISA band as a result of the efficient eccentricity damping by the density waves ($t_{\rm wav} \ll t_{\rm mig}$), while dry EMRIs from the loss-cone channel are highly eccentric as entering the LISA band and remain mildly eccentric at coalescence [131, 137]. Another subdominant dry EMRI channel involving tidal stripping of giant stars seems unlikely to produce such circular EMRIs either [391], while the prediction of the channel involving tidal disruption of binary sBHs is more uncertain [389, 392]. The disk-environmental effects may produce measurable phase shift in the EMRI waveform [152, 153, 425].

EMRIs have unprecedented potential to probe fundamental laws of gravity and the nature of dark matter [379, 380, 426, 427]. In previous studies, such tests using EMRIs have been implicitly assumed in vacuum without any environmental contamination. However, as we have shown here, wet EMRIs are possibly more common in the universe, for which the environmental effects on the EMRI waveform are inevitable. The possible degeneracy calls for a systematic framework for searching new fundamental physics with EMRIs, with astrophysical environmental effects taken into account.

In the context of wet EMRIs, AGN jet physics and accretion physics are promising realms where LISA and next-generation Event Horizon Telescope (ngEHT [4]) may synergize. According to the estimate in [386], a fraction of low-redshift ($z \lesssim 0.3$) EMRIs can be traced back to their host galaxies with LISA observations alone, and host AGNs of $\sim 50\%$ of low-redshift ($z \lesssim 0.5$) wet EMRIs can be identified considering the much lower density of AGNs. Combining GW observations of wet EMRIs with radio obervations of AGN jets by, e.g. ngEHT, one can simultaneously

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

measure the MBH mass $M_\bullet$, the MBH spin $\hat{\boldsymbol{a}}$, the rotation direction of the accretion disk $\hat{\boldsymbol{L}}$, the jet power $\dot{E}_{\rm jet}$ and the jet direction $\hat{\boldsymbol{n}}_{\rm jet}$. This set of observables provide unprecedented opportunites to probe the AGN jet physics. For example, an ensemble of events with $\{\hat{\boldsymbol{n}}_{\rm jet}\cdot\hat{\boldsymbol{a}}, \hat{\boldsymbol{n}}_{\rm jet}\cdot\hat{\boldsymbol{L}}\}$ data may help us to constrain various jet launching models, i.e., powered by the rotating energy of the MBH [428] or by the accretion disk [429]. In addition, certain disk properties are directly constrained with GW observations via the disk environmental effects on the EMRI waveform [152, 153, 425], and accretion physics of AGN disks is also one of the primary targets of ngEHT.

Wet EMRIs with host AGNs identified are ideal "bright sirens" for constraining the late time cosmology (e.g., the Hubble constant and the equation of state of dark energy), because the luminosity distance and the redshift can be measured from GW and electromagnetic observations, respectively. It will be interesting to compare the sensitivity of this method to other approaches, with the predicted wet EMRI rate from this study.

Wet EMRIs encode additional information of MBH growth in their orbital inclination angles ι_0 with respective to the MBH spin. If all MBHs grow up via coherent gas accretion where gas feeds are from a fixed direction, orbital inclination angles of wet EMRIs at coalescence should be $\iota_0 \approx \pi/2$. If MBHs grow up via chaotic gas accretion from a random direction in each active phase, a fraction of wet EMRIs form before the MBH spin direction $\hat{\boldsymbol{a}}$ is aligned with the disk rotation direction $\hat{\boldsymbol{L}}$ via the Bardeen-Petterson mechanism [430], and their orbital inclinations are approximately $\iota_0 \approx \cos^{-1}(\hat{\boldsymbol{a}}\cdot\hat{\boldsymbol{L}})$. In a similar way, MBH growth via different merger channels also imprints differently on the inclinations of wet EMRIs.

Chapter 7

Mass-gap Extreme Mass Ratio Inspirals

Zhen Pan, Zhenwei Lyu, Huan Yang

Phys. Rev. D 105, 083005 (April 2022), arXiv:2112.10237

My specific contributions: Calculate EMRI rates for all models considered in Table 7.2. Generate EMRI samples and Calculate SNR of all models with LISA sensitivity and evaluate LISA detectable rate shown in Table 7.2.

In this work, we propose a new subclass of extreme-mass-ratio-inspirals (EMRIs): mass-gap EMRIs, consisting of a compact object in the lower mass gap $\sim (2.5-5)M_\odot$ and a massive black hole (MBH). The mass-gap object (MGO) may be a primordial black hole or produced from a delayed supernova explosion. We calculate the formation rate of mass-gap EMRIs in both the (dry) loss-cone channel and the (wet) active galactic nucleus disk channel by solving Fokker-Planck-type equations for the phase-space distribution. In the dry channel, the mass-gap EMRI rate is strongly suppressed compared to the EMRI rate of stellar-mass black holes (sBHs) as a result of mass segregation effect. In the wet channel, the suppression is roughly equal to the mass ratio of sBHs over MGOs, because the migration speed of a compact object in an active galactic

nucleus disk is proportional to its mass. We find that the wet channel is much more promising to produce mass-gap EMRIs observable by spaceborne gravitation wave detectors. (Non-)detection of mass-gap EMRIs may be used to distinguish different supernova explosion mechanisms and constrain the abundance of primordial black holes around MBHs.

7.1 Introduction

Observations of Galactic X-ray binaries have indicated a dearth of compact objects around $\sim (2.5-5)M_\odot$ in the mass spectrum [e.g., 431–433]. Whether this mass gap is a result of observational selection effects or underlying supernova (SN) explosion mechanism has been an open question for more than a decade [432, 434–436]. More recently new events detected with gravitational waves (GWs) and time-domain astronomy suggest the existence of mass-gap objects (MGOs) in compact object binaries and in Galactic non-interacting binaries - the compact objects of $\sim (2.6-2.8)M_\odot$ in compact binary coalescence events GW190814 and GW200210 [437, 438], a compact object of $\sim 3.3M_\odot$ as a non-interacting companion of a giant [439], and a compact object of $\sim 3.0M_\odot$ as a non-interacting companion of a red giant [440] (see [441–443] for candidate MGOs in non-interacting binaries and as dark lens in the Milky Way). These identified MGOs provide evidences of a population of compact objects lying in the mass gap, or even a more extreme possibility that the mass gap itself does not exist. A natural question is that, if indeed a population of MGOs is present, what should be their origin? One viable option is MGOs are born in delayed SN explosions [436], which also provides an explanation to the merger rate of GW190814-like events [444, 445]. A more exotic possibility is that MGOs are primordial black holes (PBHs) [112, 446–448], which have been intensively discussed in the context of compact binary mergers detected by LIGO. With the upgrading of LIGO and the coming era of third-generation detectors, more GW190814-like events are expected to be detected. However, as there is already a large number of proposed formation channels [e.g., 449–458], many of which are still subject to large theoretical uncertainties (see

[459] for a recent review), it is unclear whether we will be able to nail down the the origin of MGOs with only the detection of stellar-mass binaries.

To better answer these questions, we investigate the possiblity that MGOs appear as components of some extreme mass ratio inspirals (EMRIs), to form mass-gap EMRIs. We propose that future detections of EMRIs and especially mass-gap EMRIs by spaceborne GW detectors, may be a useful probe for the origins of MGOs, thanks to the limited number of EMRI formation channels and distinct signatures of EMRI sources from different channels. In this paper we show that wet channel is a primary way in producing mass-gap EMRIs.

The remaining part of this paper is organized as follows. In Sec. 7.2 and 7.3, we introduce the two EMRI formation channels and the formation rates of MGOs and sBHs in these two channels. In Sec. 7.4, we forecast the detection prospects of these EMRIs by LISA. We summarize this work in Sec. 7.5. Throughout this paper, we use geometrical units $G = c = 1$, and assume a flat ΛCDM cosmology with $\Omega_{\rm m} = 0.307, \Omega_\Lambda = 1 - \Omega_{\rm m}$ and $H_0 = 67.7$ km/s/Mpc ($h = 0.677$).

7.2 Dry loss-cone channel

In this section, we will first briefly introduce the loss-cone mechanism along with the Fokker-Planck equation governing the evolution of stars in a stellar cluster around a MBH. After that we calculate the EMRI formation rates of both sBHs and MGOs via the loss-cone mechanism.

7.2.1 Loss-cone mechanism

Consider a star orbiting around a MBH, with specific binding energy $E := \phi - v^2/2$ and specific angular momentum J, where $\phi(r)$ is the (positive) gravitational potential and $v^2/2$ is the kinetic energy. Its orbital motion is affected by two main effects: GW emission which shrinks the orbit on a timescale $t_{\rm gw}$ and gravitational scatterings by other stars in the stellar cluster which changes the orbital angular momentum by order of unity on a timescale t_J. For a star on a tight and eccentric

orbit where the GW emission is more efficient with $t_{\rm gw} < t_J$, the orbit is stable against random scatterings and the star becomes an stable EMRI [131]. On the other hand, for a star on a wide and/or circular orbit where the GW emission is less efficient with $t_{\rm gw} > t_J$, the star is expected to be scattered into a random direction: away from, towards or even directly into the central MBH.

In the phase space, a region of low angular momentum $J < J_{\rm lc}(E)$ is usually referred as the loss cone, where a star ususally promptly falls into the MBH within one orbital period $P(E)$ if its angular momentum is not altered much by gravitational scatterings. As a result, the loss cone region satisfying $P(E) < t_J$ is unpopulated (empty regime) and the loss cone region satisfying $P(E) > t_J$ is populated (full regime). For the problem we are investigating, relevant orbits are of low energy ($E \approx 0$) with semi-major axis length $a \gg M_\bullet$ and the boundary of the loss cone is [460]

$$J_{\rm lc}(E \approx 0) = 4M_\bullet \,. \tag{7.1}$$

7.2.2 Fokker-Planck equation

Statistical properties of stars can be described by their distribution functions $f_i(t, \vec{r}, \vec{v})$ in the $(\vec{r}, \vec{v})$ phase space, where i labels different star species. Following Refs. [410, 411], we approximate the distribution functions as $f_i \approx f_i(t, E, R)$, where $R := J^2/J_c^2(E)$ is the normalized orbital angular momentum with $J_c(E)$ being the maximum orbital angular momentum of a star with energy E. In order to relate the distribution function $f(E, R)$ to the number density $n(r)$, and derive the Fokker-Planck equation, it is necessary to understand the properties of star orbits in given potential field $\phi(r)$, for which we summarize as follows [411]. The definition of energy suggests that

$$2(\phi - E) = v^2 = v_t^2 + v_r^2 = \frac{J^2}{r^2} + v_r^2 \,, \tag{7.2}$$

where v_t and v_r are the tangential velocity and the radial velocity respectively. For a circular orbit of energy E, its orbit radius $r_c(E)$ and angular momentum $J_c(E)$ are determined by

$$\begin{aligned} J_c^2(E) &= -r_c^3 \phi'(r_c) \,, \\ 2(\phi(r_c) - E) &= \frac{J_c^2}{r_c^2} \,. \end{aligned} \tag{7.3}$$

For a general non-circular orbit with parameters (E, R), its turning points (apsis/periapsis) $r_\pm$ are determined by

$$2(\phi(r_\pm) - E) = \frac{J^2}{r_\pm^2} \,, \tag{7.4}$$

and its orbit period $P(E, R)$ is

$$P(E, R) = 2 \int_{r_-}^{r_+} \frac{dr}{v_r} \,. \tag{7.5}$$

Defining the particle number density in the (E, R) phase space as $N(E, R)$, with $N(E, R) dE dR := \int_{r_-}^{r_+} d^3 r d^3 v f(E, R)$, we have [410, 411]

$$\begin{aligned} N(E, R) &= 4\pi^2 P(E, R) J_c^2(E) f(E, R) \\ &:= C(E, R) f(E, R) \,. \end{aligned} \tag{7.6}$$

The position-space particle number density $n(r)$ can be expressed by the distribution function $f(E, R)$ by [411]

$$n(r) = \frac{2\pi}{r^2} \int_0^{\phi(r)} dE J_c^2(E) \int_0^{R_{\max}} \frac{dR}{v_r} f(E, R) \,, \tag{7.7}$$

where $R_{\max}(r, E) = 2r^2(\phi(r) - E)/J_c^2(E)$, and $v_r(r, E, R) = 2(\phi - E) - J^2/r^2 = (R_{\max} - R) J_c^2(E)/r^2$. In the case of thermal distribution $f = f(E)$, the above equation simplifies as [461]

$$n(r) = 4\pi \int_0^{\phi(r)} dE \sqrt{2(\phi(r) - E)} f(E) \,. \tag{7.8}$$

With all these orbital properties, the Fokker-Planck equation governing the phase space evolu-

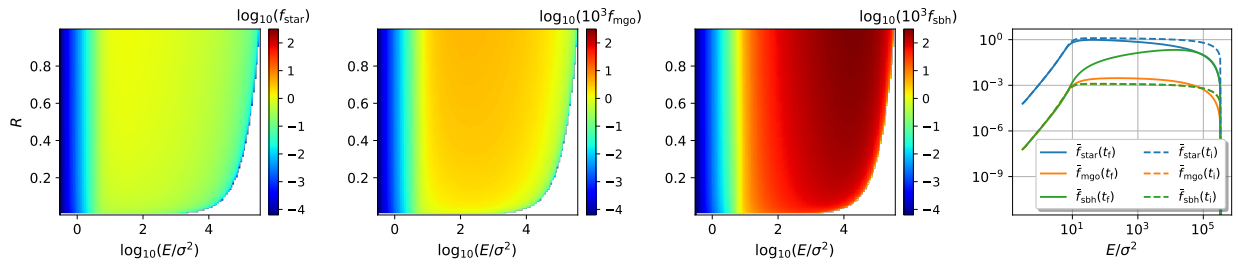

Figure 7.1: In the fiducial stellar cluster around a MBH with $M_\bullet = 4\times10^6 M_\odot$, distribution functions $f_i(E,R)$ ($i = \text{star}, \text{mgo}, \text{sbh}$) at $t_{\rm f} = 2$ Gyr are shown in the first 3 panels. The 4th panel shows R-integrated distribution functions $\bar{f}_i$ at $t_{\rm i} = 0$ (dashed lines) and at $t_{\rm f} = 2$ Gyr (solid lines). All the distribution functions are shown in units of $10^5 \text{pc}^{-3}/(2\pi\sigma^2)^{3/2}$.

tion is written in the form of [133, 410–412, 462]

$$C\frac{\partial f}{\partial t} = -\frac{\partial}{\partial E}F_E - \frac{\partial}{\partial R}F_R\,, \tag{7.9}$$

with $C = 4\pi^2 J_c^2 P$ being the weight function defined in Eq. (7.6) and $F_{E,R}$ being the flux in the E/R direction:

$$\begin{aligned} -F_E &= C\left(D_{EE}\frac{\partial f}{\partial E} + D_{ER}\frac{\partial f}{\partial R} + D_E f\right)\,,\\ -F_R &= C\left(D_{RR}\frac{\partial f}{\partial R} + D_{ER}\frac{\partial f}{\partial E} + D_R f\right)\,. \end{aligned} \tag{7.10}$$

The diffusion and advection coefficients are functions of the distribution functions and are derived in Appendix 7.A.

Given initial condition $f(t = 0, E, R)$, we evolve the cluster according to the Fokker-Planck equation (7.9) subject to following boundary conditions. On the $E \to 0$ boundary, we fix the distributions to their initial values, i.e.,

$$f(t,E,R)|_{E\to 0} = f(t=0,E,R)|_{E\to 0}\,, \tag{7.11}$$

considering the long relaxation timescale there. On the $R = 1$ boundary, the flux in the R direction should vanish,

$$F_R|_{R\to 1} = 0\ . \tag{7.12}$$

On the loss cone boundary $R = R_{\rm lc}(E) := J_{\rm lc}^2/J_c^2(E)$, there are two different regimes: full loss cone regime where

$$y_{\rm lc} := \frac{R_{\rm lc}}{(D_{RR}/R)_{R\to 0}P} < 1\ , \tag{7.13}$$

and empty regime where $y_{\rm lc} > 1$. In the empty regime, stars are expected to fall into the MBH within one orbital period P (that's why the phase space is empty). In the full regime, stars are in general scattered into/out of the loss cone mulitiple times within one orbital period, therefore the phase space is full of stars and the rate of stars falling into the MBH is low. Quantitatively, the flux in the R direction was obtained in Ref. [410] as

$$-\frac{F_R}{C} = \left(\frac{D_{RR}}{R}\right)\Big|_{R\to 0}\frac{f(R_0)}{\ln(R_0/R_{\rm lc}) + \mathcal{F}(y_{\rm lc})}\ , \tag{7.14}$$

where R_0 is any small R in the range of $R_{\rm lc} \le R \ll 1$, $\mathcal{F}(y_{\rm lc}) \sim 1/y_{\rm lc}$ for $y_{\rm lc} \lesssim 1$ and $\mathcal{F}(y_{\rm lc}) \simeq 0.824 y_{\rm lc}^{-1/2}$ for $y_{\rm lc} \gtrsim 1$. At $R_0 = R_{\rm lc}$, the above equation simplifies as $F_R(R_{\rm lc}) = 0$ in the full regime, and $f(R_{\rm lc}) = 0$ in the empty regime.

As a result, the EMRI rate and the promp infall rate per MBH via loss cone is given by

$$\begin{aligned}\Gamma_{\rm emri,lc} &= \int_{E_{\rm gw}}^{+\infty}\vec{F}\cdot d\vec{l},\\ \Gamma_{\rm infl,lc} &= \int_{\sigma^2}^{E_{\rm gw}}\vec{F}\cdot d\vec{l},\end{aligned} \tag{7.15}$$

where $\vec{F} = (F_E, F_R)$, $d\vec{l} = (dE, dR)$ is the line element along the boundary of the loss cone, and $E_{\rm gw}$ is the critical energy where $t_{\rm gw} = t_J$. To calculate the GW emission timescale $t_{\rm gw}$, we use a recently corrected version of Peters' time-scale that accounts for eccentricity evolution and post-Newtonian

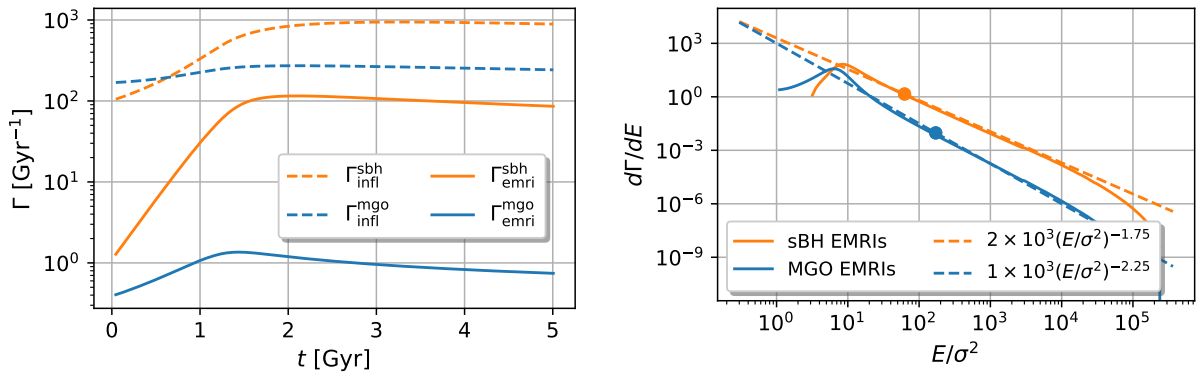

Figure 7.2: Left panel: EMRI rates $\Gamma_{\rm emri,lc}$ and prompt infall rates $\Gamma_{\rm infl,lc}$ of both sBHs and mass-gap objects in the loss-cone channel. Right panel: differential rates $d\Gamma/dE$ at $t = 2$ Gyr, and the two dots are the critical energy $E_{\rm gw}^{\rm mgo,sbh}$.

corrections [141–143, 463], with

$$t_{\rm gw} = \frac{5a^4}{256M_\bullet^2 m}\frac{(1-e^2)^{7/2}}{1+\frac{73}{24}e^2+\frac{37}{96}e^4}8^{1-\sqrt{1-e}}e^{\frac{5M_\bullet}{a(1-e)}}\,, \tag{7.16}$$

where m is the mass of the star orbiting around the MBH, a and e are the orbital semi-major axis and the eccentricity, respectively. For calculating the diffusion timescale t_J in the J-direction, we use the approximation [131]

$$t_J \approx \frac{J^2}{J_c^2(E)}t_E(E,R) = \frac{J^2}{J_c^2(E)}\frac{E^2}{2D_{EE}(E,R=0)}\,. \tag{7.17}$$

7.2.3 EMRI rate and prompt infall rate

We initialize the system with Tremaine's MBH+stellar cluster model [417, 464], assuming 3 star species in the stellar cluster: stars with mass $m_{\rm star}$, mass-gap objects with mass $m_{\rm mgo}$ and heavy sBHs with mass $m_{\rm sbh}$. The total star/mgo/sBH mass in the cluster are $M_{\rm star}$, $M_{\rm mgo}$ and $M_{\rm sbh}$, respec-

tively. Their number densities in the Tremaine's cluster model are specified by

$$n_{\rm star}(r) = \frac{M_{\rm star}}{m_{\rm star}} \frac{3-\gamma}{4\pi} \frac{r_a}{r^\gamma (r+r_a)^{4-\gamma}}, \\ n_i(r) = \delta_i \times n_{\rm star}(r) \ , \tag{7.18}$$

with i the index labelling different star species, r_a the density transition radius, γ the density scaling power index, and δ_i the abundance of species i relative to stars.

As an example, we initialize a stellar cluster with three different star species with $m_i = (1, 3, 10)M_\odot$, $\delta_i = (1, 10^{-3}, 10^{-3})$ around a MBH with $M_\bullet = 4 \times 10^6 M_\odot$. Note if there is no gap in the mass spectrum of SN remnants with the power-law mass spectrum $dN/dm \propto m^{-2.35}$ [465] holding in the whole mass range $3M_\odot \leq m \leq 50M_\odot$, we expect nearly equal number of MGOs and sBHs produced in SN explosions. The total star mass $M_{\rm star} = 20M_\bullet$, the density transition radius $r_a = 4r_{\rm h} = 4M_\bullet/\sigma^2$ and the density power index $\gamma = 1.5$, where the star velocity dispersion σ satisfies the $M_\bullet - \sigma$ relation [413, 466]

$$M_\bullet = 1.53 \times 10^6 \left(\frac{\sigma}{70 \ {\rm km/s}} \right)^{4.24} \ , \tag{7.19}$$

We evolve the cluster according to the Fokker-Planck equation (7.9) (see [467] for detailed numerical algorithm). In first 3 panels of Fig. 7.1, we show the distribution functions $f_i(t, E, R)$ at $t = 2$ Gyr, and in the 4th panel, we show the R-integrated functions $\bar{f}_i(E) = \int_0^1 dR f_i(t, E, R)$ at $t = 0$ and $t = 2$ Gyr, respectively. From the 4th panel, we see sBHs (which are the most massive star component) concentrate around the MBH as a result of mass segregation, yielding a large increase in the distribution function for sBHs at small radii/large binding energy E with time [132, 134, 468], while little concentration is found for less massive MGOs or stars.

In the upper panel of Fig. 7.2, we show the EMRI rates and the prompt infall rates of both sBHs and MGOs, where $\Gamma^{\rm mgo}_{\rm emri}$ is lower than $\Gamma^{\rm sbh}_{\rm emri}$ by a factor of $O(10^2)$, as a result of the stronger

mass segregation and shorter GW emission timescale $t_{\rm gw}$ for sBHs. In contrast, the prompt infall rates are less affected by the mass segregation, because the prompt infall rate depends on the star density at lower energy ($< E_{\rm gw}$), while the EMRI rate depends on the star density at higher energy ($> E_{\rm gw}$) (Eq. 7.15), and the latter is more sensitive to the mass segregation (Fig. 7.1). As a result, we find the number of prompt infalls per EMRI $N_{\rm p} := \Gamma_{\rm infl}/\Gamma_{\rm emri}$ are $N_{\rm p}^{\rm sbh} \approx 10, N_{\rm p}^{\rm mgo} \approx 250$ for the fiducial model.

For an analytic understanding of these results, we re-estimate the number of prompt infalls per EMRI $N_{\rm p}$ using previous analytic formula [Eq. (17) and (26) in [131]],

$$\begin{aligned}\Gamma_{\rm emri,lc} &= \int_0^{a_{\rm gw}} \frac{4\pi a^2 n(a)}{\ln(R_{\rm lc}) t_{\rm rlx}(a)}\, da \sim a_{\rm gw}^{1.5-2p}\,, \\ \Gamma_{\rm infl,lc} &= \int_{a_{\rm gw}}^{a_{\rm max}} \frac{4\pi a^2 n(a)}{\ln(R_{\rm lc}) t_{\rm rlx}(a)}\, da \sim a_{\rm max}^{1.5-2p}\,,\end{aligned} \tag{7.20}$$

where $n(a) \sim a^{-1.5-p}$ is the number density, $t_{\rm rlx}(a) \sim a^p$ is the local relaxation timescale, $a_{\rm max}$ is a characteristic radius of the stellar cluster, and $a_{\rm gw}$ is the critical radius where $t_{\rm gw} = t_J$ [Eqs.(7.16,7.17)]. For a single-species cluster filled with stars of mass m, the dependence of $a_{\rm gw}$ on mass m and on the MBH mass $M_\bullet$ is [131, 469],

$$\frac{a_{\rm gw}}{M_\bullet/\sigma^2} \propto m^{\frac{2}{3-2p}} M_\bullet^0\,. \tag{7.21}$$

For comparison with numerical results, we formulate the above analytic results in the phase space, with the analytic differential rates

$$d\Gamma/dE = d\Gamma/da \times da/dE \sim E^{2p-2.5}\,, \tag{7.22}$$

which can be directly compared wtih the numerical results of the fiducial model at $t = 2$ Gyr (lower panel of Fig. 7.2). From the comparison, we see the power laws $d\Gamma_{\rm mgo}/dE \sim E^{-2.25}$ and $d\Gamma_{\rm sbh}/dE \sim E^{-1.75}$ are a good approximation at $E \gtrsim 10\sigma^2$, i.e., $2p_{\rm mgo} \approx 0.25, 2p_{\rm sbh} \approx 0.75$. In

terms of the differential rates, the EMRI rate and the prompt infall rates are written as

$$\begin{aligned}\Gamma_{\rm emri,lc} &= \int_{E_{\rm gw}}^{\infty} d\Gamma/dE\, dE\,,\\ \Gamma_{\rm infl,lc} &= \int_{\sigma^2}^{E_{\rm gw}} d\Gamma/dE\, dE\,,\end{aligned} \tag{7.23}$$

where the critical energy are numerically found as $E_{\rm gw}^{\rm sbh} = 63\sigma^2, E_{\rm gw}^{\rm mgo} = 171\sigma^2$ and they are consistent with the analytic expectation $E_{\rm gw}^{\rm sbh}/E_{\rm gw}^{\rm mgo} \approx (m_{\rm sbh}/m_{\rm mgo})^{\frac{2}{3-2p}}$ [Eq. (7.21)]. Using the power-law approximations to the differential rates, we have

$$N_{\rm p} \approx (E_{\rm gw}/E_{\rm min})^{1.5-2p}, \tag{7.24}$$

where $E_{\rm min} \in (1, 10)\sigma^2$ is an effective minimum energy (lower panel of Fig. 7.2). As a result, we obtain an analytic estimate $N_{\rm p}^{\rm sbh} \in (4, 22)$ and $N_{\rm p}^{\rm mgo} \in (35, 600)$.

Similar to the fiducial model, we initialize the stellar cluster around a MBH with mass in the range of $(10^5, 10^7)M_\odot$, then evolve the system for $T_0 = 5$ Gyr, and summarize the time-averaged EMRI rates $\bar{\Gamma}_{\rm emri}^{\rm sbh,mgo}$ and prompt infall rates $\bar{\Gamma}_{\rm infl}^{\rm sbh,mgo}$ in Fig. 7.3. We find the EMRI rates peak around $M_\bullet = 10^6 M_\odot$, because the rates are limited by the longer relaxation timescale of the stellar cluster around a heavier MBH, while the rates are limited by the lower number of sBHs and MGOs in the stellar cluster around a lighter MBH [467]. We find the time-averaged EMRI rates are $\bar{\Gamma}_{\rm emri}^{\rm sbh} = O(10 - 10^2)$ Gyr^{-1} and $\bar{\Gamma}_{\rm emri}^{\rm mgo} = O(1)$ Gyr^{-1}. For longer evolution time T_0, the rates decrease further because of the depletion of sBHs and MGOs in the stellar cluster. On average, the number of prompt infalls per EMRI $N_{\rm p}$ are similar to in the fiducial model with $N_{\rm p}^{\rm sbh} \approx 10, N_{\rm p}^{\rm mgo} \approx 250$, except $N_{\rm p}(M_\bullet = 10^5 M_\odot)$ is lower by a factor ~ 2. The nearly independence of $N_{\rm p}$ on the MBH mass $M_\bullet$ comes from the independence of $a_{\rm gw}$ (or equavilently $E_{\rm gw}$) on $M_\bullet$ [Eq. (7.21,7.24)], while $N_{\rm p}^{\rm mgo,sbh}(M_\bullet = 10^5 M_\odot)$ are lower simply because MGOs/sBHs around the lighter MBH are of lower number and are quickly depleted via the loss cone, consequently the critical energy $E_{\rm gw}$ decreases

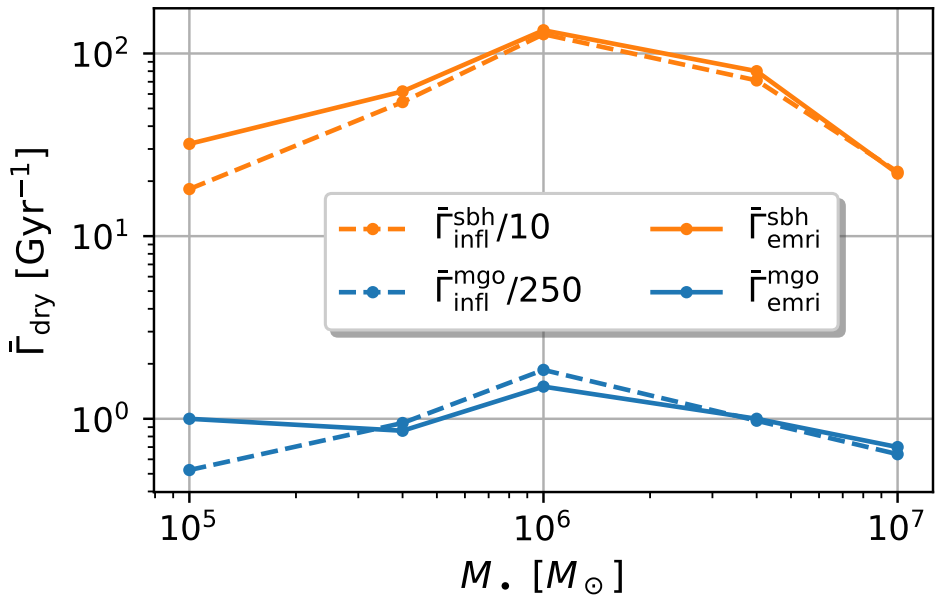

Figure 7.3: Average EMRI rates $\bar{\Gamma}_{\rm emri,dry}$ and average prompt infall rates $\bar{\Gamma}_{\rm infl,dry}$ of both sBHs and mass-gap objects in the loss-cone channel.

and $N_{\rm p}$ is reduced.

7.3 Wet AGN disk channel

In the presence of an accretion disk around a MBH, the distributions of all different orbiting object species are affected by the disk. As a result, the spherical symmetry is broken and the distribution function $f(E, R, \mu)$ generally acquires dependence on the orbital inclination ι w.r.t. the disk plane, where we have defined $\mu := \cos\iota = \hat{\boldsymbol{J}} \cdot \hat{\boldsymbol{J}}_{\rm disk}$, with $\hat{\boldsymbol{J}}$ and $\hat{\boldsymbol{J}}_{\rm disk}$ being the unit direction vectors of the star orbital angular momentum and the disk angular momentum, respectively.

For the problem we are considering, all stars can be conveniently decomposed as a cluster

component and a disk component, i.e.,

$$f(E, R, \mu) \to f(E, \mu) + g(E)\delta(\mu - 1)\delta(R - 1) \,, \tag{7.25}$$

where we have approximate the cluster-component distribution as R-independent and approximate the disk component as circular orbiters lying on the equator with $\iota = 0$ ($\mu = 1$), because the orbital eccentricity damping timescale is in general much shorter than the migration timescale (see subsection 7.3.1 for details). With this decomposition, we have number density $n(r, \theta)$ of the cluster-component stars and surface number density $\Sigma(r)$ of the disk-component stars as

$$\begin{aligned} n(r, \theta) &= 4\pi \int_0^{\phi(r)} dE \sqrt{2(\phi(r) - E)} \bar{f}(E, \theta) \,, \\ \Sigma(r) &= 2\pi^2 r E \sqrt{2(\phi(r) - E)} g(E) \Big|_{r=r_c(E)} \,, \end{aligned} \tag{7.26}$$

with

$$\bar{f}(E, \theta) = \frac{1}{2\pi} \int_0^{2\pi} d\eta f(E, \mu = \sin\theta \cos\eta) \,, \tag{7.27}$$

where θ is the polar angle w.r.t. to the $\hat{\boldsymbol{J}}_{\rm disk}$ direction.

In the remaining part of this section, we will first summarize the important interactions between stars and the accretion disk, then incorporate these interactions into the Fokker-Planck equation, and finally evolve the stellar cluster to calculate the wet EMRI rates.

7.3.1 Star-disk interactions

Interactions of an accretion disk with stars, MGOs and sBHs are similar in aspects of density waves and dynamical friction. In term of gas accretion onto stars, MGOs and sBHs within the AGN disk, the star size makes a difference. Compact objects of relatively small sizes only grow mildly within the AGN disk, while the stellar evolution is expected to be impacted by gas accretion onto stars of much larger sizes [470, 471]. In this work, we are not intended to model the star evolution in

detail, and we simply assume no mass change of all the star species during the evolution period.

For illustration purpose, we take sBHs as an example. As a sBH orbits around the central MBH surrounded by an accretion disk, its periodic motion generates density waves, which in turn drive the star to migrate inward, damp its orbital eccentricity e and its orbital inclination ι. For a highly inclined orbiter, the density wave effects become subdominant with respect to dynamical friction as it goes through the gas disk. The two effects (density waves and dynamical friction) together contribute to the advection in (E, μ) space as [150, 151]

$$\begin{aligned}\langle \Delta\mu \rangle_t^{\rm dsk} &= (1-\mu^2)\frac{\iota}{\sin\iota} \times \min.\left\{\frac{0.544}{t_{\rm wav}}, \frac{1.46}{t_{\rm wav}}\frac{h^4}{\iota \sin^3(\iota/2)}\right\} , \\ \langle \Delta E \rangle_t^{\rm dsk} &= E \times \min.\left\{\frac{2.7+1.1\alpha_s}{t_{\rm mig,I}}, \frac{8.8}{t_{\rm mig,I}}\frac{h^2}{\sin(\iota)\sin(\iota/2)}\right\} ,\end{aligned} \tag{7.28}$$

with $\langle \Delta X \rangle_t := \frac{\Delta X}{\Delta t}|_{\Delta t \to 0}$, $\alpha_s := d\ln\Sigma / d\ln r$ and

$$t_{\rm mig,I} = \frac{M}{m}\frac{M}{\Sigma r^2}h^2 \, , t_{\rm wav} = t_{\rm mig,I} h^2 \, , \tag{7.29}$$

where m is the mass of the orbiter, M is the total mass within radius r, h and Σ are disk scale height and the disk surface density [1].

For a sBH captured into the disk, its orbital eccentricity will be damped by the eccentricity density waves on timescale $t_{\rm wav}$, which is generally much shorter than all other relevant timescales, including the migration timescale $t_{\rm mig,I}$ and two diffusion timescales, E^2/D_{EE} and $(1-\mu^2)/D_{\mu\mu}$. As a result, sBHs in the AGN disk are generally moving in circular orbits. For a sBH embedded in the gas disk, surrounding gas tends to flow towards it nearly in the radial direction at large distances, in the rest frame of the sBH. Due to the differential rotation of the gas disk, the inflowing gas generally carries non-zero angular momentum relative to the sBH, consequently circularizes and

[1] In the previous work [467], we have approximated the effect of density waves as inclination-independent and have neglected the contribution from dynamical friction. As a result, the previous approximation is an over estimate for highly inclined orbiters.

forms a disk around the sBH. The gas inflow rate at the outer boundary is usually super-Eddington and a strong outflow naturally emerges. As a result, a major part of the inflowing gas may escape as outflow and only the remaining part is accreted by the sBH (see Ref. [470] and references therein for detailed modeling). Because of the circularization process, it is reasonable to expect that the outflow carries minimal net momentum with respect to the sBH, and the momentum carried by the inflow eventually transfers to the sBH. The head wind contributes to the advection in the E-direction as

$$\langle \Delta E \rangle_t^{\rm wnd} = \frac{2J}{\dot{J}_{\rm wnd}} \quad \text{(for in-disk orbiters)}, \tag{7.30}$$

where $\dot{J}_{\rm wnd}$ is sBH angular momentum loss rate due to the head wind (see Ref. [467] for calculation details).

GW emission only becomes important when the orbiter is very close to the MBH and it drives an advection in the E direction for a circular orbiter as [463]

$$\langle \Delta E \rangle_t^{\rm gw} = \frac{64}{5} \frac{M^2 m}{a^4} E \, . \tag{7.31}$$

Accretion disk structure of AGNs has not been well understood especially in the outer parts, where both disk heating mechanism and the angular momentum transport mechanism are not clear. Three commonly used AGN disk models (α/β disks [408], and TQM disk proposed by Thompson, Quataert, and Murray [409]) have been numerically solved and compared in our previous work [467]. In this work, we will use α and β disks with accretion rate $\dot{M}_\bullet = 0.1 \dot{M}_{\bullet,\rm Edd}$ as fiducial disk models (Fig. 5 in [467]). In Fig. 7.4, we show the migration timescales $E/\langle \Delta E \rangle_t := E/(\langle \Delta E \rangle_t^{\rm dsk} + \langle \Delta E \rangle_t^{\rm wnd} + \langle \Delta E \rangle_t^{\rm gw})$ of a sBH with $m_{\rm sbh} = 10 M_\odot$ embedded in the two fiducial accretion disks. The two disks only differs in inner parts where the radiation pressure dominates over the gas pressure and the two migration timescales only differs where $E/\sigma^2 \gtrsim 10^3$. We do not include the TQM disk model in this work because a much more efficient angular momentum transport mechanism

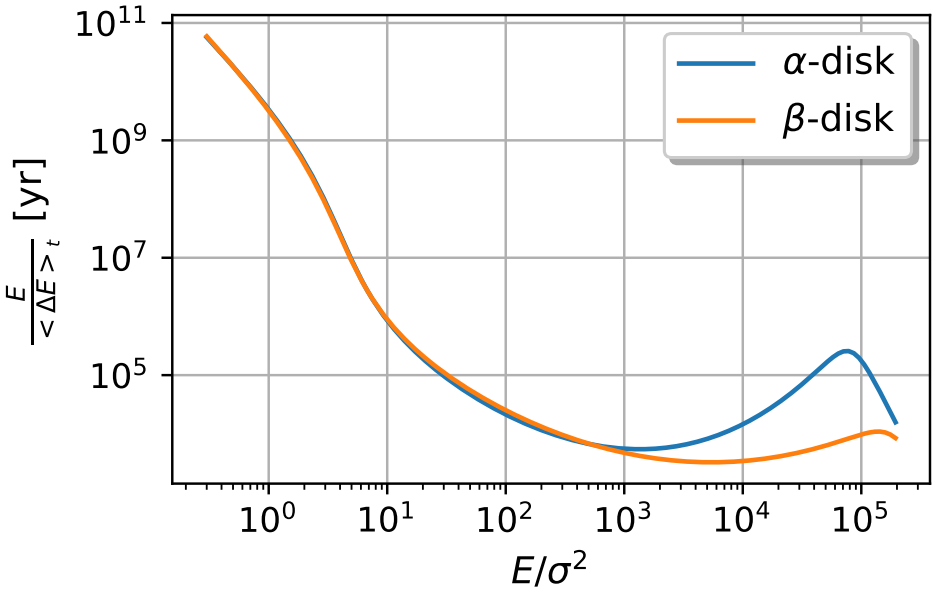

Figure 7.4: The migration timescales of a $10M_\odot$ BH embedded in the fiducial α and β disks, respectively, where the GW emission becomes dominant at $E/\sigma^2 \sim 10^5$ or equivalently $a \sim 10^2 M_\bullet$.

is assumed in the TQM disk model, which is inconsistent with the turbulence viscosity driven by magnetorotational instability in inner parts of the accretion disk [472–474]. Due to the high efficiecy of the angular momentum transport assumed in the TQM disk model, TQM disks are in general of lower surface density, therefore longer migration timescale $t_{\rm mig,I}$, which hinders sBHs and MGOs from migrating to the vicinity of the central MBH and forming EMRIs if $T_{\rm disk} < t_{\rm mig,I}$. If $T_{\rm disk} > t_{\rm mig,I}$, the EMRI rates in TQM disks are similar to those in α/β-disks (see [467] for details).

7.3.2 Fokker-Planck equation

For cluster-component stars, the Fokker-Planck equation takes the form

$$C_\mu \frac{\partial f}{\partial t} = -\frac{\partial}{\partial E} F_E - \frac{\partial}{\partial \mu} F_\mu \,, \tag{7.32}$$

with flux

$$
\begin{aligned}
-F_E &= C_\mu \left(D_{EE} \frac{\partial f}{\partial E} + D_{E\mu} \frac{\partial f}{\partial \mu} + D_E f \right) , \\
-F_\mu &= C_\mu \left(D_{\mu\mu} \frac{\partial f}{\partial \mu} + D_{E\mu} \frac{\partial f}{\partial E} + D_\mu f \right) ,
\end{aligned}
\tag{7.33}
$$

and the weight function

$$
C_\mu(E) := \frac{1}{2} \int_0^1 C(E, R) dR \, , \tag{7.34}
$$

where the factor 2 comes from $\int_{-1}^{1} d\mu$. All the coefficients of the Fokker-Planck equation (7.32) are contributed by star interactions with the accretion disk and by scatterings with both the cluster-component stars and the disk-component stars, where the first only contributes to the advection coefficients as

$$
D_E = -\left(\langle \Delta E \rangle_t^{\rm dsk} + \langle \Delta E \rangle_t^{\rm gw} \right) , \; D_\mu = -\langle \Delta \mu \rangle_t^{\rm dsk} \; . \tag{7.35}
$$

and the latter two contributions are given in Eqs.(7.78,7.79), respectively.

With proper initial conditions, we evolve the system according to Eq. (7.32) with the following boundary conditions. On the $E \to 0$ boundary, we again fix the distribution, i.e.,

$$
f(t, E, R)|_{E \to 0} = f(t = 0, E, R)|_{E \to 0} \, . \tag{7.36}
$$

On the $E \to E_{\rm max}$ boundary, where the evolution of distribution function is dominated by GW emission, we set

$$
F_E|_{E \to E_{\rm max}} = -C_\mu D_E f \, . \tag{7.37}
$$

On the $\mu = -1$ boundary, we use the zero-flux condition

$$
F_\mu|_{\mu \to -1} = 0 \, . \tag{7.38}
$$

On the $\mu = 1$ boundary, where the evolution of distribution function is dominated by the inclination

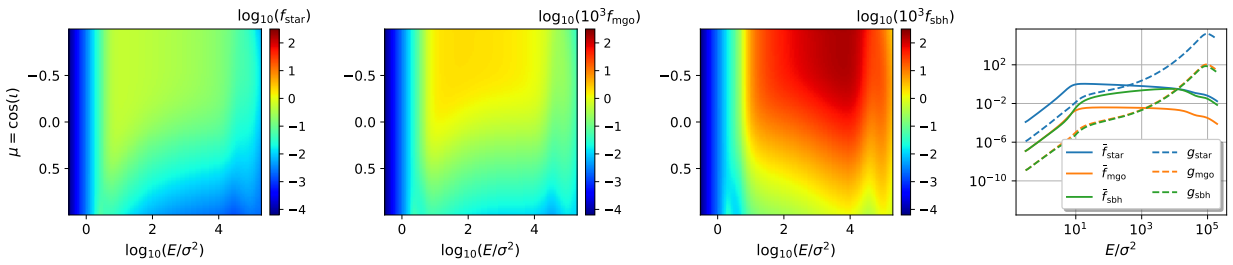

Figure 7.5: In the fiducial stellar cluster around a MBH with $M_\bullet = 4 \times 10^6 M_\odot$, the cluster-component distribution functions $f_i(E, \mu)$ ($i = \text{star}, \text{mgo}, \text{sbh}$) at $t = 10^7$ yr are shown in the first 3 panels. The 4th panel shows μ-integrated cluster-component distribution functions $\bar{f}_i(E)$ (solid lines) and disk-component $g_i(E)$ (dashed lines). All the distribution functions are shown in units of $10^5 \text{pc}^{-3}/(2\pi\sigma^2)^{3/2}$.

damping arising from the normal density waves, we set

$$F_\mu|_{\mu\to 1} = -C_\mu D_\mu f \ . \tag{7.39}$$

For disk-component stars, the Fokker-Planck equation reduces to be 1-dimensional,

$$C_\mu \frac{\partial g}{\partial t} = \frac{\partial}{\partial E}\left[C_\mu\left(D_{EE}\frac{\partial g}{\partial E} + D_E g\right)\right] + F_\mu(E, \mu = 1) \ , \tag{7.40}$$

with the source term arising form disk-component stars captured by the disk. We also find the evolution of disk-component stars is dominated by the star-disk interactions and GW emission [Eqs. (7.28,7.30,7.31)], so we simply neglect the contributions from scatterings, i.e.,

$$D_E = -\left(\langle \Delta E\rangle_t^{\rm dsk} + \langle \Delta E\rangle_t^{\rm gw} + \langle \Delta E\rangle_t^{\rm wnd}\right) \ , \quad D_{EE} = 0 \ . \tag{7.41}$$

Now Eq. (7.40) is an first-order differential equation, which requires only one boundary condition,

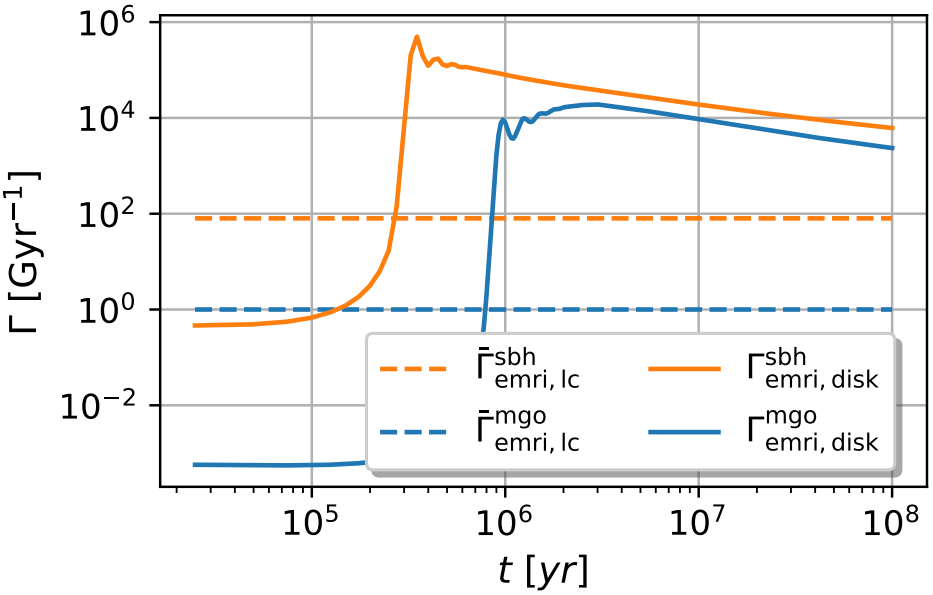

Figure 7.6: Wet EMRI rates $\Gamma_{\text{emri,disk}}(t)$ of sBHs and MGOs. For comparision, we also plot the average dry EMRI rates $\bar{\Gamma}_{\text{emri,lc}}$ during the quiet phase as horizontal lines (Fig. 7.3).

and we choose it as

$$g(t, E, R)|_{E\to 0} = g(t = 0, E, R)|_{E\to 0} \,. \tag{7.42}$$

The wet EMRI rate is determined by the flux in the E direction at the E_{max} boundary, i.e.,

$$\Gamma_{\text{emri,disk}} = -C_\mu D_E g|_{E=E_{\text{max}}} \,. \tag{7.43}$$

Strictly speaking, the cluster-component contribution should also be included in addition to the disk-component contribution. As we will see later, the disk-component constribution at the E_{max} boundary turns out to be dominant.

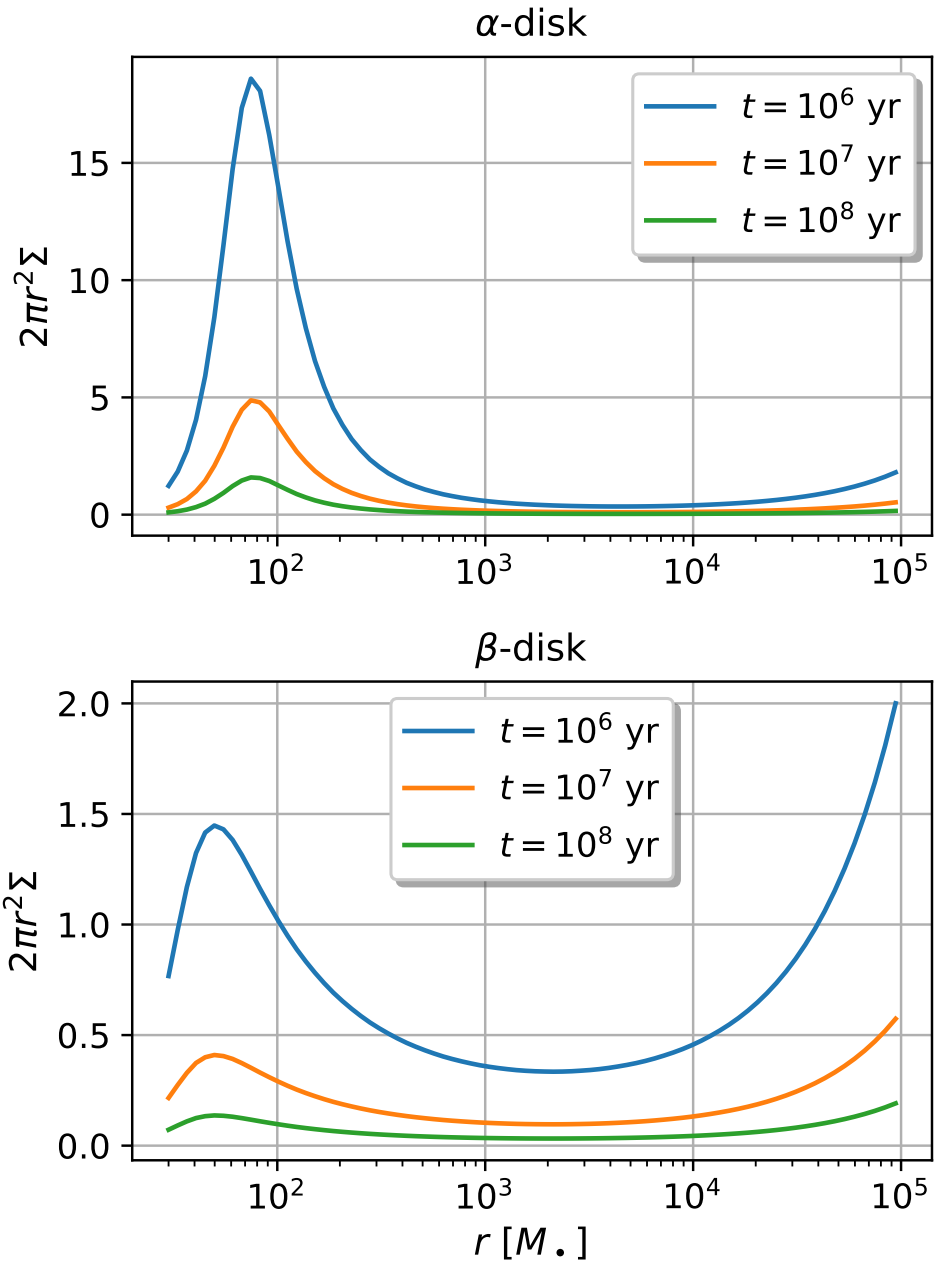

Figure 7.7: The surface number density $\Sigma(r)$ of disk-component sBHs at different times for the two different AGN disk models.

7.3.3 Wet EMRI rates

Considering that the accretion history of MBHs is likely episodic [415, 416] and the active phase of an MBH is in general much shorter that its quiet phase [144, 145, 414], we simplify the duty cycle of an AGN as a long quiet phase of $T_0 = 5$ Gyr followed by a short active phase of duration $T_{\rm disk} \in \{10^6, 10^7, 10^8\}$ yr. This simplied picture holds if the cluster evolution driven by purely two-body scatterings is negligible during a quiet phase between two active episodes, so that adjacent episodes can be effectively glued together as we understand the hisotory of evolution.

For calculating the wet EMRI rate, we initialize the stellar cluster around a MBH and evolve the system for T_0 in the same way as in the previous section, then turn on the accretion disk and initialize the cluster-component and disk-component distributions as

$$\begin{aligned} f_i(t=0,E,\mu) &= 0.99 \int_0^1 f_i(t=T_0,E,R)\, dR\,, \\ g_i(t=0,E) &= 0.01 \int_0^1 f_i(t=T_0,E,R)\, dR\,, \end{aligned} \tag{7.44}$$

where the integrand $f_i(t = T_0, E, R)$ is the distribution function at the end of the quiet phase, and the initial fraction 0.01 is the typical disk aspect ratio [467], the exact value of which does not matter because the disk-component star densities are mostly determined by subsequent capture and migration processes. With this initialization, we continue the evolution of $f_i(t, E, \mu)$ and $g_i(t, E)$ according to Eqs.(7.32,7.40) for a duration $T_{\rm disk}$.

As a fiducial model of the wet channel, we again consider a fiducial model with a MBH and a stellar cluster same to in the previous section, and a fiducial α-disk with lifetime $T_{\rm disk} = 10^8$ yr (Fig. 7.4). In the first 3 panels of Fig. 7.5, we show the cluster-component distribution functions $f_i(E, \mu)$ at $t = 10^7$ yr, where we see low-inclination ($\mu \to 1$) orbiters have been captured into the disk, therefore the distribution function $f_i(E, \mu \to 1)$ is relatively lower. In the 4th panel, we show the μ-integrated distribution functions $\bar{f}_i(E) = \int_{-1}^1 d\mu f_i(E, \mu)$ and the disk-component functions

$\bar{g}_i(E)$ at $t = 10^7$ yr. The disk component dominates over the cluster component in the vicinity of the MBH (large E), and the disk component peaks around $E \sim 10^5\sigma^2$ where the migration timescale peaks (Fig. 7.4). It is interesting to note that the disk-component distribution functions $g_i(E)$ are proportional to their abundance δ_i with no dependence on different star masses m_i, simply because both the rate of stars captured onto the disk and the migration rate of stars along the disk are proportional to their masses m_i, and the local density $g_i(E)$ is determined by the ratio of the two rates in the equilibrium state.

In Fig. 7.6, we show the wet EMRI rates for both sBHs and MGOs as functions of time. Initally both of them are low because it takes sometime for sBHs and MGOs that are captured by the disk to migrate to the MBH; and the two rates peaks around $t = 3 \times 10^5$ yr and $t = 10^6$ yr, respectively, because the migration timescale is inversely proportional to the star mass m_i. After the peak time, the rate of stars captured by the disk is in equilibrium with the corresponding EMRI rate, and both of them steadily decrease with time $\propto t^{-0.5}$, which is the typical behavior of diffusion processes with absorbing boundary conditions. We find the wet EMRI rate of sBHs in the equilibrium state is higher than the dry EMRI rate by $O(10^2 - 10^3)$ for the fiducial model, which is consistent with the result of the slow disk capture scenario in our previous work (Fig. 10 of Ref. [467]), though the detailed time dependence does not match exactly because we approximated the effect of density waves as inclination-independent and neglected the contribution from dynamical friction in the previous work, while we have used more refined prescription of star-disk interaction [Eq. (7.28)] and self-consistent calculation of the disk capture rate [Eq. (7.40)] in this work. In the loss-cone channel, the EMRI rate of MGOs is largely suppressed compared to sBHs because MGOs benefit little from the mass segregation (see Figs. 7.1 and 7.2), while in the disk channel, the EMRI rate of MGOs is lower than that of sBHs by a factor of $\sim m_{\rm sbh}/m_{\rm mgo}$ in the equilibrium state (Fig. 7.6). As a result, we find the wet EMRI rate of MGOs in the equilibrium state is higher than the dry rate by $O(10^3 - 10^4)$ for the fiducial model.

As shown in Fig. 7.4, the migration timescale of a sBH in an AGN disk peaks around $r \sim$

$10^2 M_\bullet$, where the GW emission start to dominate over other processes. As a result, a number of sBHs aggregate around this radius due to a traffic jam (4th panel of Fig. 7.5). For clarity, we show the surface number density of disk-component sBHs Σ in Fig. 7.7 at different times. Compared with the β-disk, more sBHs aggregate around $r \sim 10^2 M_\bullet$ in the fiducial α-disk because of slower migration speed. A similar result (with slightly higher surface number densities) was obtained in our previous work [470] though the disk capture rate was not calculated from first principles.

Similar to the fiducial model, we initialize a stellar cluster of a MBH with mass in the range of $(10^5 - 10^7)\ M_\odot$, evolve the MBH+cluster system for $T_0 = 5$ Gyr, then turn on an accretion disk around the MBH and initialize the cluster-component and disk-component distributions, and continue the evolution for $T_{\rm disk}$. The average wet EMRI rates of sBHs and MGOs

$$\bar{\Gamma}_{\rm wet} = \frac{1}{T_{\rm disk}} \int_0^{T_{\rm disk}} \Gamma_{\rm emri,disk} dt \ , \tag{7.45}$$

are summarized in Table 7.1. For long disk lifetime $T_{\rm disk} \gtrsim 10^7$ yr, the wet EMRI rates of sBHs and MGOs are higher for heavier MBHs because more sBHs and MGOs are available around heavier MBHs. For short disk lifetime $T_{\rm disk} = 10^6$ yr, the trend reverses for $M_\bullet \gtrsim 10^6 M_\odot$ because longer migration timescale in AGN disks ($\gtrsim T_{\rm disk}$) around heavier MBHs hinders the sBHs and MGOs captured by the AGN disk from migrating to the vicinity of the MBH and becoming EMRIs. In comparison with the dry channel, we find the wet EMRI rate of sBHs is usually higher by $O(10^1 - 10^3)$ and the wet EMRI rate of MGOs is usually higher by $O(10^3 - 10^4)$.

7.4 Detection Prospects

In addition to the generic EMRI rates per MBH/AGN obtained in the previous two sections, a few extra pieces of information are needed for calculating the LISA detectable EMRI rate: the mass function of MBHs $dN_\bullet/dM_\bullet$ and the fraction of MBHs living a stellar cusp which is supposed to

Table 7.1: Average EMRI rates (Gyr^{-1}) of sBHs and MGOs per AGN, assuming sBHs and MGOs are produced from SN explosion with equal numbers.

AGN disk	$M_\bullet/M_\odot$	$(\bar{\Gamma}^{\rm mgo}_{\rm wet}, \bar{\Gamma}^{\rm sbh}_{\rm wet})_{T_{\rm disk}=10^6{\rm yr}}$	$(\bar{\Gamma}^{\rm mgo}_{\rm wet}, \bar{\Gamma}^{\rm sbh}_{\rm wet})_{T_{\rm disk}=10^7{\rm yr}}$	$(\bar{\Gamma}^{\rm mgo}_{\rm wet}, \bar{\Gamma}^{\rm sbh}_{\rm wet})_{T_{\rm disk}=10^8{\rm yr}}$
α-disk	1×10^7	$(0.1\times10^3, 4.1\times10^4)$	$(1.2\times10^4, 4.9\times10^4)$	$(6.7\times10^3, 1.7\times10^4)$
	4×10^6	$(0.7\times10^3, 9.0\times10^4)$	$(1.2\times10^4, 3.7\times10^4)$	$(4.7\times10^3, 1.2\times10^4)$
	1×10^6	$(1.0\times10^4, 4.4\times10^4)$	$(7.4\times10^3, 1.3\times10^4)$	$(2.5\times10^3, 4.3\times10^3)$
	4×10^5	$(1.1\times10^4, 1.7\times10^4)$	$(4.3\times10^3, 4.5\times10^3)$	$(1.5\times10^3, 1.4\times10^3)$
	1×10^5	$(2.7\times10^3, 1.6\times10^3)$	$(1.1\times10^3, 0.6\times10^3)$	$(0.5\times10^3, 0.3\times10^3)$
β-disk	1×10^7	$(4.6\times10^3, 9.3\times10^4)$	$(1.3\times10^4, 5.1\times10^4)$	$(6.7\times10^3, 1.7\times10^4)$
	4×10^6	$(7.3\times10^3, 1.1\times10^5)$	$(1.3\times10^4, 3.8\times10^4)$	$(4.7\times10^3, 1.2\times10^4)$
	1×10^6	$(1.5\times10^4, 4.8\times10^4)$	$(7.6\times10^3, 1.4\times10^4)$	$(2.5\times10^3, 4.3\times10^3)$
	4×10^5	$(1.2\times10^4, 1.8\times10^4)$	$(4.4\times10^3, 4.9\times10^3)$	$(1.5\times10^3, 1.4\times10^3)$
	1×10^5	$(3.7\times10^3, 3.2\times10^3)$	$(1.1\times10^3, 0.8\times10^3)$	$(0.5\times10^3, 0.3\times10^3)$

Table 7.2: Forecasted Total and LISA detectable (with SNR$\geq$ 20) EMRI rates of sBHs and MGOs in the redshift range $0 < z < 4.5$ assuming sBHs and MGOs are produced from SN explosions with equal numbers. For the wet channel, we have assumed a conservative AGN fraction $f_{\rm AGN} = 1\%$ throughout the universe.

Wet EMRIs	$f_\bullet$	AGN disk	$T_{\rm disk}$ [yr]	Total rates of (MGOs, sBHs) [yr^{-1}]	LISA detectable rates of (MGOs, sBHs) [yr^{-1}]
	$f_{\bullet,-0.3}$	α-disk	10^6	(1900, 6400)	(50, 480)
			10^7	(1400, 2500)	(24, 130)
			10^8	(540, 860)	(10, 54)
		β-disk	10^6	(2700, 8200)	(65, 530)
			10^7	(1400, 2600)	(24, 150)
			10^8	(540, 860)	(10, 54)
	$f_{\bullet,+0.3}$	α-disk	10^6	(110, 1000)	(3, 34)
			10^7	(180, 470)	(1, 10)
			10^8	(71, 160)	(< 1, 3)
		β-disk	10^6	(200, 1300)	(5, 38)
			10^7	(190, 500)	(2, 11)
			10^8	(71, 160)	(< 1, 3)
Dry EMRIs	$f_\bullet$			Total rates of (MGOs, sBHs) [yr^{-1}]	LISA detectable rates of (MGOs, sBHs) [yr^{-1}]
	$f_{\bullet,-0.3}$			(79, 1300)	(1, 120)
	$f_{\bullet,+0.3}$			(3, 130)	(< 1, 10)

be destroyed during a MBH merger following a previous galaxy merger.

Following Ref. [137], we consider two MBH mass functions in the range of $(10^4, 10^7)M_\odot$,

$$
\begin{aligned}
f_{\bullet,-0.3} : \frac{dN_\bullet}{d\log M_\bullet} &= 0.01\left(\frac{M_\bullet}{3\times10^6 M_\odot}\right)^{-0.3}\ \mathrm{Mpc}^{-3}\ , \\
f_{\bullet,+0.3} : \frac{dN_\bullet}{d\log M_\bullet} &= 0.002\left(\frac{M_\bullet}{3\times10^6 M_\odot}\right)^{+0.3}\ \mathrm{Mpc}^{-3}\ ,
\end{aligned}
\tag{7.46}
$$

where $f_{\bullet,-0.3}$ is an approximation to the mass function in the model assuming MBHs are seeded by Pop-III stars and grow via accretion and mergers [419], and $f_{\bullet,+0.3}$ is purely a phenomenological model [385].

In the frame of observers on the earth, the differential dry and wet EMRI rates are written as

$$
\begin{aligned}
\frac{d^2\mathcal{R}_{\rm dry}}{dM_\bullet dz} &= \frac{1}{1+z}\frac{dN_\bullet}{dM_\bullet}\frac{dV_{\rm c}(z)}{dz}C_{\rm cusp}(M_\bullet,z)\bar{\Gamma}_{\rm dry}(M_\bullet;N_p)\ , \\
\frac{d^2\mathcal{R}_{\rm wet}}{dM_\bullet dz} &= \frac{f_{\rm AGN}}{1+z}\frac{dN_\bullet}{dM_\bullet}\frac{dV_{\rm c}(z)}{dz}C_{\rm cusp}(M_\bullet,z)\bar{\Gamma}_{\rm wet}(M_\bullet;\mathbb{M})\ ,
\end{aligned}
\tag{7.47}
$$

where z is cosmological redshift, $V_{\rm c}(z)$ is the comoving volume of the universe up to redshift z, $C_{\rm cusp}(M_\bullet, z)$ is the fraction of MBHs embedded in a stellar cusp, where we use the same $C_{\rm cusp}(M_\bullet, z)$ function as in [137] for models with MBH mass function $f_{\bullet,-0.3}$ and take $C_{\rm cusp}(M_\bullet, z) = 1$ for models with MBH mass function $f_{\bullet,+0.3}$. $\bar{\Gamma}_{\rm wet}(M_\bullet;\mathbb{M})$ is the average wet EMRI rate (Table 7.1) and we conservatively take the AGN fraction as $f_{\rm AGN} = 1\%$.

In consistent with Ref. [137], we parametrize the average dry EMRI rate of sBHs as

$$
\bar{\Gamma}^{\rm sbh}_{\rm dry}(M_\bullet;N_{\rm p}) = C_{\rm dep}(M_\bullet;N_p)C_{\rm grow}(M_\bullet;N_p)\Gamma^{\rm sbh}_{\rm lc}(M_\bullet)\ ,
\tag{7.48}
$$

where we take the number of prompt infalls per EMRI as $N_{\rm p} = 10$ (Fig. 7.2), the generic rate

$$
\Gamma^{\rm sbh}_{\rm lc}(M_\bullet) = \Gamma_0\left(\frac{M_\bullet}{10^6 M_\odot}\right)^{-0.19}\ ,
\tag{7.49}
$$

with $\Gamma_0 \in (30, 300)$ Gyr^{-1} [137, 475], and two correction factors are correction from possible depletion of sBHs available $C_{\rm dep}(M_\bullet; N_p)$ and correction capping the MBH mass growth via accreting sBHs (from both prompt infalls and EMRIs) $C_{\rm grow}(M_\bullet; N_p)$, respectively. Though there is uncertainty of a factor of $O(10)$ in the generic rate $\Gamma^{\rm sbh}_{\rm lc}$, we will see that the uncertainty does not propagate to the average rate $\bar{\Gamma}^{\rm sbh}_{\rm dry}$ with the two corrections. The depletion correction is formulated as

$$C_{\rm dep} = \text{min.}\left\{\frac{T_{\rm dep}}{T_{\rm rlx}}, 1\right\} , \tag{7.50}$$

where $T_{\rm dep}$ is the depletion timescale of sBHs residing in the MBH influence sphere ($r_c = 2M_\bullet/\sigma^2$)

$$T_{\rm dep} = \frac{\Sigma m_{\rm sbh}}{(1+N_{\rm p})\Gamma^{\rm sbh}_{\rm lc} m_{\rm sbh}} , \tag{7.51}$$

assuming the total mass of sBHs in the influence sphere is $\Sigma m_{\rm sbh} = 0.06 M_\bullet$, and the relaxation timescale at $r = r_c$ is

$$T_{\rm rlx} = \left(\frac{\sigma}{20\ {\rm km/s}}\right)\left(\frac{r_c}{1\ {\rm pc}}\right)^2\ {\rm Gyr} . \tag{7.52}$$

The MBH growth correction comes from requiring the MBH mass grows no more than $1/e$ via accreting sBHs,

$$C_{\rm grow} = \text{min.}\left\{e^{-1}\frac{M_\bullet}{\Delta M_\bullet}, 1\right\} , \tag{7.53}$$

with

$$\Delta M_\bullet = m_{\rm sbh}(1+N_{\rm p}) C_{\rm dep}(M_\bullet, N_{\rm p}) \Gamma^{\rm sbh}_{\rm lc}(M_\bullet, N_{\rm p}) T_{\rm emri}(M_\bullet) , \tag{7.54}$$

and

$$T_{\rm emri} = \int dt \frac{dt}{dz} C_{\rm cusp}(M_\bullet, z) \tag{7.55}$$

is the total duration when a MBH lives in a stellar cusp.

With these two corrections, we find the average EMRI rate of sBHs is well fitted by

$$\begin{aligned}&\bar{\Gamma}^{\rm sbh}_{\rm dry}(M_\bullet; N_{\rm p}=10)\\&=\min.\left\{26\left(\frac{M_\bullet}{10^5 M_\odot}\right), 30\left(\frac{M_\bullet}{10^6 M_\odot}\right)^{-0.19}\right\}\ {\rm Gyr}^{-1}\ .\end{aligned} \tag{7.56}$$

with little dependence on the generic rate Γ_0 as long as it is higher than 30 Gyr^{-1}, because the average rate is in fact determined by the number of sBHs available around MBHs and the MBH growth limit via accreting sBHs. For the average EMRI rate of MGOs, we simply take it as $\bar{\Gamma}^{\rm mgo}_{\rm dry}(M_\bullet) \approx 1$ Gyr^{-1} (Fig. 7.3).

With all the elements for calculating the differential EMRI rates ready [Eq. (7.47)], we calcuate the total EMRI rates of MGOs and sBHs from the two channels and the LISA detectable EMRI rates. We first sample the EMRI sources according to the differential rates [Eq. (7.47)], then compute the EMRI waveform using the Augment Analytic Kludge [476–478] and the expected signal-to-ratio (SNR) by the LISA detector (see all the source sampling and SNR computation details in the previous work [475]). The forecast results are listed in Table 7.2. For the well-motivated MBH mass function $f_{\bullet,-0.3}$, we expect LISA to detect ~ 1 mass-gap EMRIs, $O(10^2)$ sBH EMRIs from the dry channel, $O(10-10^2)\times(f_{\rm AGN}/1\%)$ mass-gap EMRIs, and $O(10^2-10^3)\times(f_{\rm AGN}/1\%)$ sBH EMRIs from the wet channel per year. For the less optimistic MBH mass function $f_{\bullet,+0.3}$, the expected detection numbers are overall lower by a factor of $O(10)$.

7.5 Conclusions and Discussion

7.5.1 Summary

In the dry EMRI formation channel, the formation rate of mass-gap EMRIs is strongly suppressed compared with EMRIs of sBHs, because sBHs are heavier and accumulate closer to MBH due to

the mass segregation effect and therefore easier to form EMRIs (Fig. 7.2). In the wet channel, the EMRI formation turns out to be much more efficient than in the dry channel because the capture of compact objects onto the accretion disk and subsequent inward migration along the disk are highly efficient in transporting compact objects [467, 475]. Both the capture rate onto to disk and the migration speed along the disk are linearly proportional to the compact object mass, so that the formation rate of wet mass-gap EMRIs in the equilibrium state is suppressed by a factor of $\sim m_{\rm mgo}/m_{\rm sbh}$ assuming their abundances are equal ($\delta_{\rm sbh} = \delta_{\rm mgo}$). As a result, we find the wet EMRI rate of sBHs per AGN is higher than the dry rate per MBH by $O(10^1 - 10^3)$, and the wet EMRI rate of MGOs per AGN is higher than the dry rate per MBH by $O(10^3 - 10^4)$. Accounting for the AGN fraction $f_{\rm AGN} = O(10^{-2} - 10^{-1})$, the wet channel turns out to the dominant channel of mass-gap EMRI formation. As for the LISA detection prospects, we expect LISA to detect no more than $\sim$ 1 dry mass-gap EMRIs, and $O(10 - 10^2) \times (f_{\rm AGN}/1\%)$ wet mass-gap EMRIs per year for the physically motivated MBH mass function $f_{\bullet,-0.3}$. For the less optimistic MBH mass function $f_{\bullet,+0.3}$, the expected detection numbers are lower by $O(10)$ (Table 7.2).

7.5.2 Discussion

As shown above, the expected number of EMRI detections (denoted as $D_{\rm sbh}$ and $D_{\rm mgo}$) are sensitive to the unknown MBH mass function, while the ratio $D_{\rm sbh}/D_{\rm mgo}$ is not, which can be used as a more robust probe to the MGO abundance and origin.

If LISA detects $D_{\rm mgo}$ dry mass-gap EMRIs and $D_{\rm sbh}$ dry EMRIs of sBHs per year, we can infer the relative abundance of MGOs and sBHs within nuclear stellar clusters as

$$\frac{\delta_{\rm mgo}}{\delta_{\rm sbh}} \approx \frac{D_{\rm mgo}/D_{\rm sbh}}{R^{\rm mgo}_{\rm sbh}}\,, \tag{7.57}$$

where $R^{\rm mgo}_{\rm sbh} \approx 1/120$ is the ratio of expected detection numbers of two different EMRIs assuming MGOs and sBHs are of the same abundance (Table 7.2). In a similar way making use of Eq. (7.57),

one can again infer the relative abundance of MGOs $\delta_{\rm mgo}/\delta_{\rm sbh}$ from detections of wet EMRIs, where $R_{\rm sbh}^{\rm mgo} = (1/10 - 1/5)$ (Table 7.2) varies little over all different model parameters for the parameter space we considered. The inferred relative abundance $\delta_{\rm mgo}/\delta_{\rm sbh}$ can be used to constrain the SN explosion mechanisms, where the delayed SN explosion mechanism predicts $\delta_{\rm mgo}/\delta_{\rm sbh} \to 1$ while the rapid explosion mechanism predicts $\delta_{\rm mgo}/\delta_{\rm sbh} \to 0$ [436].

An excess of mass-gap EMRI detection by LISA is a signature of MGOs of exotic origins (e.g., PBHs [446–448, 479, 480]). If these MGOs are of primordial origin, their abundance around MBHs may be further used to constrain the mass fraction of PBHs in dark matter (DM) $f_{\rm mgo} := \Omega_{\rm mgo}/\Omega_{\rm DM}$. This constraint sensitively depends on the DM distribution around MBHs, which is poorly understood theoretically. We consider two extremal cases: (1) DM around MBHs traces baryons with the DM to baryon ratio $\Omega_{\rm DM}/\Omega_{\rm B}$; (2) the DM density around MBHs follows the NFW profile [481].

In case (1), the abundance of MGOs relative to stars is

$$\frac{\delta_{\rm mgo}}{\delta_{\rm star}} = \frac{f_{\rm mgo}\Omega_{\rm DM}}{(1-f_{\rm gas})\Omega_{\rm B}}\frac{m_{\rm star}}{m_{\rm mgo}}\,, \tag{7.58}$$

where $f_{\rm gas}$ is the mass fraction of gas in baryons and $f_{\rm mgo}$ is formulated as

$$\begin{aligned} f_{\rm mgo} &= (1-f_{\rm gas})\frac{\delta_{\rm sbh}}{\delta_{\rm star}}\frac{D_{\rm mgo}/D_{\rm sbh}}{R_{\rm sbh}^{\rm mgo}}\frac{\Omega_{\rm B}}{\Omega_{\rm DM}}\frac{m_{\rm mgo}}{m_{\rm star}} \\ &= 6\times10^{-4}(1-f_{\rm gas})\frac{\delta_{\rm sbh}/\delta_{\rm star}}{10^{-3}}\frac{D_{\rm mgo}/D_{\rm sbh}}{R_{\rm sbh}^{\rm mgo}}\frac{\Omega_{\rm B}/\Omega_{\rm DM}}{0.2}\frac{m_{\rm mgo}/m_{\rm star}}{3} \end{aligned} \tag{7.59}$$

where we have used Eq.(7.57).

In case (2), the DM abundance around MBHs is usually much lower, with the total DM mass with the MBH influence radius $M_{\rm DM}(<r_c) \approx 0.3\% M_\bullet$ (see Appendix 7.C for details). For comparison, the total star mass is $\Sigma m_{\rm star}(<r_c) \approx 2M_\bullet$ and the total mass of astrophysical MGOs is $\Sigma m_{\rm mgo}(r<r_c) \approx 0.6\% M_\bullet \times (\delta_{\rm mgo}/10^{-3})$, i.e., the DM abundance around MBHs is comparable

with that of astrophysical MGOs. In this case, it is unlikely to observe excess of mass-gap EMRIs. Therefore, an excess of mass-gap EMRI detection by LISA would disfavor the NFW distribution of MGOs as DM.

Many other proposals to explain GW190814-like events involve hierachical mergers, e.g., in young stellar clusters [452], triple systems [455], AGN disks [457], etc. If the abundance of MGO production can be realibly estimated in these scenarios, the rate of mass gap EMRIs may also be used to test these models. For example, as young stellar clusters are not expected to host massive black holes, if they are the only places that MGOs are produced, we should expect the mass-gap EMRI rate to be minimized.

One working assumption we used is no mass change of MGOs in AGN disks. As shown in Fig. 7.4, the typical migration timescale of MGOs is $\sim 10^6$ yr, which is much shorter than the Salpeter timescale 5×10^7 yr (mass doubling timescale with the Eddington accretion rate). But the accretion rate of MGOs/sBHs in AGN disks is uncertain [470, 482]. If gas accretion onto MGOs was highly super-Eddington and largely increased their masses, identifying the distorted mass gap is less straightforward, and the commonly used techniques of searching for mass-gap features in the mass spectrum of LIGO events [438] would also be valuable for our purpose.

7.A Appendix: Diffusion and advection coefficients in the Fokker-Planck equation (7.9)

In Refs. [410, 411, 462], the diffusion and the advection coefficients of a single-species cluster have been derived in detail. Following Refs. [133, 412], we extend the them to multi-species cases. We

first define auxiliary functions:

$$
\begin{aligned}
F_0^{(j)}(E,r) &= (4\pi m_j)^2 \ln\Lambda \int_{-\infty}^{E} dE' \bar{f}_j(E') \,, \\
F_n^{(j)}(E,r) &= (4\pi m_j)^2 \ln\Lambda \int_{E}^{\phi(r)} dE' \left(\frac{\phi - E'}{\phi - E}\right)^{n/2} \bar{f}_j(E') \,,
\end{aligned}
\tag{7.60}
$$

where $n \geq 1$, j is the index labelling different star species, $\ln\Lambda$ the Coulomb's logarithm which we take as $\ln\Lambda = 10$, and

$$
\bar{f}_j(E) := \int_0^1 f_j(E,R) dR \,. \tag{7.61}
$$

With these auxiliary functions, the coefficients are written as

$$
\begin{aligned}
D_{EE}^{(i)} &= \sum_j \frac{2}{3P} \int_{r_-}^{r_+} \frac{dr}{v_r} v^2 (F_0^{(j)} + F_3^{(j)}) \,, \\
D_{E}^{(i)} &= \sum_j -\frac{2}{P} \int_{r_-}^{r_+} \frac{dr}{v_r} F_1^{(j)} \times \frac{m_i}{m_j} \,, \\
D_{ER}^{(i)} &= \sum_j \frac{4}{3P} R \int_{r_-}^{r_+} \frac{dr}{v_r} \left(\frac{v^2}{v_c^2} - 1\right) (F_0^{(j)} + F_3^{(j)}) \,, \\
D_{RR}^{(i)} &= \sum_j \frac{4}{3P} \frac{R}{J_c^2} \int_{r_-}^{r_+} \frac{dr}{v_r} \left\{ 2\frac{r^2}{v^2} \left[v_t^2 \left(\frac{v^2}{v_c^2} - 1\right)^2 + v_r^2 \right] F_0^{(j)} \right. \\
&\qquad \left. + 3\frac{r^2}{v^2} v_r^2 F_1^{(j)} + \frac{r^2}{v^2} \left[2v_t^2 \left(\frac{v^2}{v_c^2} - 1\right)^2 - v_r^2 \right] F_3^{(j)} \right\} \,, \\
D_{R}^{(i)} &= \sum_j -\frac{4}{P} \frac{R}{v_c^2} \int_{r_-}^{r_+} \frac{dr}{v_r} \left(1 - \frac{v_c^2}{v^2}\right) F_1^{(j)} \times \frac{m_i}{m_j} \,,
\end{aligned}
\tag{7.62}
$$

where i (similar to j) is also the star species index, $v_t = J/r$ is the tangential velocity, and $v_c(E)$ is the velocity of a circular orbit with energy E.

7.B Appendix: Diffusion and advection coefficients in the Fokker-Planck equation (7.32)

Following Refs. [133, 410–412], we derive the diffusion and advection coefficients in the Fokker-Planck equation (7.32) in this section. Consider a star m with orbital energy E and velocity $\boldsymbol{v}$ changes its energy by ΔE and velocity by $\Delta\boldsymbol{v}$ due to a scattering with a field star m_a. In the following orthonormal basis,

$$\begin{aligned}\hat{\boldsymbol{v}} &= \boldsymbol{v}/v\,,\\ \hat{\boldsymbol{J}} &= \boldsymbol{J}/J = \boldsymbol{r}\times\boldsymbol{v}/J\,,\\ \hat{\boldsymbol{w}} &= \boldsymbol{v}\times\boldsymbol{J}/|\boldsymbol{v}\times\boldsymbol{J}|\,,\end{aligned} \tag{7.63}$$

the velocity change is written as

$$\Delta\boldsymbol{v} = \Delta v_{\parallel}\hat{\boldsymbol{v}} + \Delta\boldsymbol{v}_{\perp} = \Delta v_{\parallel}\hat{\boldsymbol{v}} + \Delta v_J\hat{\boldsymbol{J}} + \Delta v_w\hat{\boldsymbol{w}}\,. \tag{7.64}$$

As a result, the changes in energy and angular momentum are

$$\Delta E = -\frac{1}{2}(\Delta v_{\parallel})^2 - \frac{1}{2}(\Delta v_{\perp})^2 - v\Delta v_{\parallel}\,, \tag{7.65}$$

and

$$\begin{aligned}\Delta\boldsymbol{J} &= \boldsymbol{r}\times\Delta\boldsymbol{v}\\ &= J\left(\frac{\Delta v_{\parallel}}{v} - \frac{v_r}{v_t}\frac{\Delta v_w}{v}\right)\hat{\boldsymbol{J}} + J\frac{\Delta v_J}{v}\left(\frac{v_r}{v_t}\hat{\boldsymbol{w}} - \hat{\boldsymbol{v}}\right)\\ &:= \Delta\boldsymbol{J}_{\parallel} + \Delta\boldsymbol{J}_{\perp}\,,\end{aligned} \tag{7.66}$$

with $\Delta J_{\perp} = J(\Delta v_J/v)\sqrt{1+(v_r/v_t)^2} = J(\Delta v_J/v_t) = r\Delta v_J$ and

$$\Delta J := |\boldsymbol{J}+\Delta\boldsymbol{J}| - J = J\frac{\Delta v_{\parallel}}{v} - rv_r\frac{\Delta v_w}{v} + \frac{1}{2}\frac{r^2(\Delta v_J)^2}{J}\,. \tag{7.67}$$

For $\mu = \cos\iota = J_z/J$ with J_z the z-component angular momentum, its change is

$$\Delta\mu = \frac{\Delta J_z}{J} - \mu\frac{\Delta J}{J} = \frac{\Delta v_J}{v}\left(\frac{v_r}{v_t}\hat{\boldsymbol{w}} - \hat{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{z}} - \frac{\mu}{2}\frac{r^2(\Delta v_J)^2}{J^2} \,. \tag{7.68}$$

Defining $\langle\Delta X\rangle_t := \frac{\langle\Delta X\rangle}{\Delta t}|_{\Delta t\to 0}$ and $\langle\Delta X\Delta Y\rangle_t := \frac{\langle\Delta X\Delta Y\rangle}{\Delta t}|_{\Delta t\to 0}$ (where $\langle\rangle$ is ensemble average over scatterings a star has experienced), it is straightforward to see

$$\begin{aligned}
\langle\Delta E\rangle_t &= -\frac{1}{2}\langle(\Delta v_\parallel)^2\rangle_t - \frac{1}{2}\langle(\Delta v_\perp)^2\rangle_t - v\langle\Delta v_\parallel\rangle_t \,, \\
\langle(\Delta E)^2\rangle_t &= v^2\langle(\Delta v_\parallel)^2\rangle_t \,, \\
\langle\Delta\mu\rangle_t &= \langle\frac{\Delta v_J}{v}\rangle_t\left(\frac{v_r}{v_t}\hat{\boldsymbol{w}} - \hat{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{z}} - \frac{\mu}{2}\frac{r^2}{J^2}\langle(\Delta v_J)^2\rangle_t \,, \\
\langle(\Delta\mu)^2\rangle_t &= \langle\left(\frac{\Delta v_J}{v}\left(\frac{v_r}{v_t}\hat{\boldsymbol{w}} - \hat{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{z}}\right)^2\rangle_t \\
&= \langle\frac{(\Delta v_J)^2}{v^2}\left(\frac{v_r^2}{v_t^2} + 1\right)\frac{\sin^2\iota}{2}\rangle_t \\
&= \frac{1-\mu^2}{2}\frac{r^2}{J^2}\langle(\Delta v_J)^2\rangle_t \,, \\
\langle\Delta E\Delta\mu\rangle_t &= -\left(\frac{v_r}{v_t}\hat{\boldsymbol{w}} - \hat{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{z}}\,\langle\Delta v_\parallel\Delta v_J\rangle_t \,,
\end{aligned} \tag{7.69}$$

accurate to quadractic order in Δv.

In the case of spherical symmetry, coefficients $\langle\Delta\boldsymbol{v}\rangle_t$ and $\langle\Delta\boldsymbol{v}\Delta\boldsymbol{v}\rangle_t$ has been derived by [412] (assuming the field stars m_a are symmetrically distributed in the azimuthal direction in the rest frame of particle m) as

$$\begin{aligned}
\langle\Delta v_\parallel\rangle_t &= -\kappa\frac{m+m_a}{m_a}\int_0^v dv_a\frac{v_a^2}{v^2}f_a(v_a) \,, \\
\langle(\Delta v_\parallel)^2\rangle_t &= \frac{2}{3}\kappa\left(\int_0^v dv_a\frac{v_a^4}{v^3}f_a(v_a) + \int_v^\infty dv_a v_a f_a(v_a)\right) \\
\langle(\Delta v_\perp)^2\rangle_t &= \frac{2}{3}\kappa\left(\int_0^v dv_a\left(\frac{3v_a^2}{v} - \frac{v_a^4}{v^3}\right)f_a(v_a) + 2\int_v^\infty dv_a v_a f_a(v_a)\right)
\end{aligned} \tag{7.70}$$

$\langle \Delta \boldsymbol{v}_\perp \rangle_t = 0$ and $\langle \Delta v_\parallel \Delta v_J \rangle_t = 0$, with $\kappa = (4\pi m_a)^2 \ln\Lambda$. It is straightforward to extend the above results to the non-spherical symmetry case, with the replacement $f_a(v_a) \to \bar{f}_a(v_a, \theta)$, where

$$\bar{f}_a(v_a, \theta) := \frac{1}{2\pi} \int_0^{2\pi} d\eta f_a(v_a, \mu = \sin\theta\cos\eta) d\eta \,, \tag{7.71}$$

and θ is the polar angle w.r.t. the z-axis.

With $\langle \Delta \boldsymbol{v} \rangle_t$ and $\langle \Delta \boldsymbol{v} \Delta \boldsymbol{v} \rangle_t$ ready, the derivation of coefficients in the Fokker-Planck equation (7.32) is parallel to the previous section and we outline it as follows. Define functions

$$\begin{aligned} F_0^{(j)}(E, \theta) &= (4\pi m_j)^2 \ln\Lambda \int_{-\infty}^{E} dE' \bar{f}_j(E', \theta), \\ F_n^{(j)}(E, r, \theta) &= (4\pi m_j)^2 \ln\Lambda \int_{E}^{\phi(r)} dE' \bar{f}_j(E', \theta) \left(\frac{\phi - E'}{\phi - E}\right)^{n/2}, \end{aligned} \tag{7.72}$$

with $n \geq 1$ and

$$\bar{f}_j(E, \theta) := \frac{1}{2\pi} \int_0^{2\pi} d\eta f_j(E, \mu = \sin\theta\cos\eta) d\eta \,. \tag{7.73}$$

With these auxiliary functions, the local diffusion/advection coefficients $\hat{D}(E, R, r, \theta)$ (which depend on energy E, normalized angular momentum R, and spatial coordinates r and θ) are

$$\begin{aligned} \hat{D}_{EE}^{(i)} &= \sum_j \frac{1}{3} v^2 (F_0^{(j)} + F_3^{(j)}) \,, \\ \hat{D}_{E}^{(i)} &= \sum_j -F_1^{(j)} \times \frac{m_i}{m_j} \,, \\ D_{\mu\mu}^{(i)} &= \sum_j \frac{1-\mu^2}{4} \frac{1}{3} v_t^{-2} (2F_0^{(j)} + 3F_2^{(j)} - F_4^{(j)}) \,, \\ D_{\mu}^{(i)} &= \sum_j \frac{1-\mu^2}{4} \frac{1}{3} \frac{\partial}{\partial\mu} v_t^{-2} (2F_0^{(j)} + 3F_2^{(j)} - F_4^{(j)}) \,, \\ D_{\mu E}^{(i)} &= 0 \,. \end{aligned} \tag{7.74}$$

In general, the orbital coordinates (r, θ) of a star orbit with semi-major axis $a(E)$ and eccentric-

ity e are specified by

$$r = \frac{a(1-e^2)}{1+e\cos\psi'}\,, \quad \cos\theta = \cos(\psi'+\phi)\sin\iota\,, \tag{7.75}$$

with $\phi \in [0, 2\pi]$ specifying the pericenter location in the azimuthal direction, and $\psi' \in [0, 2\pi]$ is the orbital phase w.r.t the pericenter. With orbit average, we obtain

$$\bar{D}(E, R, \mu, \phi) = \frac{2}{P}\int_{r_-}^{r_+} \frac{dr}{v_r}\hat{D}(E, R, r, \theta)\,. \tag{7.76}$$

After ensemble average over R and ϕ, we arrive at the final form $D(E,\mu) = \langle \bar{D}(E, R, \mu, \phi)\rangle_{R,\phi}$.

In fact, we find circular orbits is a good approximation in calculating the coefficients. For circular orbits, the orbital equation is simplified as

$$r = r_c(E), \quad \cos\theta = \cos\psi'\sin\iota \tag{7.77}$$

and the diffusion/advection coefficients arising from scatterings with the cluster-component stars

$$\begin{aligned}
D^{(i)}_{EE} &= \sum_j \frac{1}{3\pi}\int_0^\pi d\psi' v^2 (F_0^{(j)} + F_3^{(j)})\,,\\
D^{(i)}_{E} &= \sum_j -\frac{1}{\pi}\int_0^\pi d\psi' F_1^{(j)} \times \frac{m_i}{m_j}\,,\\
D^{(i)}_{\mu\mu} &= \sum_j \frac{1-\mu^2}{4}\frac{1}{3\pi}\int_0^\pi d\psi' v_t^{-2}(2F_0^{(j)} + 3F_2^{(j)} - F_4^{(j)})\,,\\
D^{(i)}_{\mu} &= \sum_j \frac{1-\mu^2}{4}\frac{1}{3\pi}\frac{\partial}{\partial\mu}\int_0^\pi d\psi' v_t^{-2}(2F_0^{(j)} + 3F_2^{(j)} - F_4^{(j)})\,,\\
D^{(i)}_{\mu E} &= 0\,.
\end{aligned} \tag{7.78}$$

With the same approximation, the diffusion/advection coefficients arising from scatterings with

the disk-component stars are

$$
\begin{aligned}
D^{(i)}_{EE} &= \sum_j \frac{1}{3}\epsilon v^2 (G^{(j)}_0 + G^{(j)}_3)\,, \\
D^{(i)}_{E} &= \sum_j -\epsilon G^{(j)}_1 \times \frac{m_i}{m_j}\,, \\
D^{(i)}_{\mu\mu} &= \sum_j \frac{1-\mu^2}{4}\frac{1}{3}\epsilon v_t^{-2}(2G^{(j)}_0 + 3G^{(j)}_2 - G^{(j)}_4)\,, \\
D^{(i)}_{\mu} &= \sum_j \frac{1-\mu^2}{4}\frac{1}{3}\frac{\partial}{\partial\mu}\epsilon v_t^{-2}(2G^{(j)}_0 + 3G^{(j)}_2 - G^{(j)}_4)\,, \\
D^{(i)}_{\mu E} &= 0\,.
\end{aligned}
\tag{7.79}
$$

where $\epsilon(h,\iota) = \min\{1, \frac{2}{\pi}\arcsin(\frac{h}{\iota})\} \approx \frac{2}{\pi}\arcsin(\frac{h}{h+\iota})$ is fraction of the orbit lying inside the disk component, and

$$
\begin{aligned}
G^{(j)}_0(E) &= (4\pi m_j)^2 \ln\Lambda \int_{-\infty}^{E} dE' g_j(E'), \\
G^{(j)}_n(E,r) &= (4\pi m_j)^2 \ln\Lambda \int_{E}^{\phi(r)} dE' g_j(E') \left(\frac{\phi - E'}{\phi - E}\right)^{n/2}\,,
\end{aligned}
\tag{7.80}
$$

with $n \geq 1$.

7.C Appendix: Dark Matter NFW profile

If DM density around a MBH follows the NFW profile [481]

$$
\rho_{\rm DM}(r) = \frac{\rho_s}{\frac{r}{R_s}\left(1+\frac{r}{R_s}\right)^2},
\tag{7.81}
$$

the total DM mass within radius r is written as $M_{\rm DM}(<r) = 4\pi\rho_s R_s^3 \mathcal{G}(c)$, with concentration $c := r/R_s$ and $\mathcal{G}(c) = \ln(1+c) - c/(1+c)$, where ρ_s and R_s are the characteristic density and radius, respectively. For relating ρ_s and R_s to the MBH mass $M_\bullet$, we need the aid of a commonly used cutoff radius within which the average DM density is 200 times the critical density of the universe

$\rho_{\rm crit}$, i.e.,

$$M_{200} = 4\pi\rho_s R_s^3 \mathcal{G}(c_{200}), \quad \frac{\rho_s}{\rho_{\rm crit}} = \frac{200}{3}\frac{c_{200}^3}{\mathcal{G}(c_{200})}\,, \tag{7.82}$$

As found in Ref. [483], M_{200} and $M_\bullet$ are correlated with

$$\frac{M_\bullet}{10^7 M_\odot} \approx \left(\frac{M_{200}}{10^{12} M_\odot}\right)^{1.65}\,. \tag{7.83}$$

For low redshift, the concentration c_{200} has a weak dependence on the mass M_{200} [484]

$$\log_{10} c_{200} = 0.905 - 0.101 \log_{10}(M_{200}/10^{12} h^{-1} M_\odot)\,. \tag{7.84}$$

Combining the above three equations, we find the total mass of DM within the influence radius r_c is $M_{\rm DM}(< r_c) \approx 0.3\% M_\bullet$ for $M_\bullet \in (10^5, 10^7) M_\odot$.

References

[1] The LIGO Scientific Collaboration, The Virgo Collaboration, and The KAGRA Collaboration, 10.48550/ARXIV.2111.03606 (2021).

[2] B. P. Abbott et al., Living Reviews in Relativity **23**, 1 (2020).

[3] F. Acernese et al., Class. Quant. Grav. **32**, 024001 (2014).

[4] J. Aasi et al., Class. Quant. Grav. **32**, 074001 (2015).

[5] Z. Pan et al., Phys. Rev. Lett. **125**, 201102 (2020).

[6] J. Zhang et al., Phys. Rev. Lett. **127**, 161101 (2021).

[7] Z. Lyu, N. Jiang, and K. Yagi, Phys. Rev. D **105**, 064001 (2022).

[8] T. Hinderer et al., Phys. Rev. Lett. **116**, 181101 (2016).

[9] X. Feng, Z. Lyu, and H. Yang, Phys. Rev. D **105**, 104043 (2022).

[10] J. Baker et al., 10.48550/ARXIV.1907.06482 (2019).

[11] J. Mei et al., Progress of Theoretical and Experimental Physics **2021**, 05A107 (2020).

[12] W.-H. Ruan et al., International Journal of Modern Physics A **35**, 2050075 (2020).

[13] Z. Pan, Z. Lyu, and H. Yang, Phys. Rev. D **104**, 063007 (2021).

[14] Z. Pan, Z. Lyu, and H. Yang, Phys. Rev. D **105**, 083005 (2022).

[15] M. Punturo et al., Class. Quant. Grav. **27**, 194002 (2010).

[16] B. P. Abbott et al., Class. Quant. Grav. **34**, 044001 (2017).

[17] J. M. Weisberg and J. H. Taylor, 10.48550/ARXIV.ASTRO-PH/0407149 (2004).

[18] B. P. Abbott et al., Phys. Rev. Lett. **116**, 061102 (2016).

[19] The LIGO Scientific Collaboration, 10.48550/ARXIV.1904.03187 (2019).

[20] K. L. Dooley et al., Class. Quant. Grav. **33**, 075009 (2016).

[21] P. A. R. Ade et al. (BICEP/Keck Collaboration), Phys. Rev. Lett. **127**, 151301 (2021).

[22] M. A. McLaughlin, Class. Quant. Grav. **30**, 224008 (2013).

[23] J. P. W. Verbiest et al., Mon. Not. Roy. Astron. Soc. **458**, 1267 (2016).

[24] M. Kramer and D. J. Champion, Class. Quant. Grav. **30**, 224009 (2013).

[25] R. N. Manchester et al., Publications of the Astronomical Society of Australia **30**, 10.1017/pasa.2012.017 (2013).

[26] A. de Waard et al., Class. Quant. Grav. **22**, S215 (2005).

[27] J. L. Levine, Physics in Perspective **6**, 42 (2004).

[28] O. D. Aguiar, Research in Astronomy and Astrophysics **11**, 1 (2010).

[29] S. Dimopoulos et al., Physics Letters B **678**, 37 (2009).

[30] D.-F. Gao, J. Wang, and M.-S. Zhan, Communications in Theoretical Physics **69**, 37 (2018).

[31] B. Abbott et al., Phys. Rev. Lett. **119**, 161101 (2017).

[32] B. P. Abbott et al., ApJ **848**, L12 (2017).

[33] B. P. Abbott et al., ApJ **848**, L13 (2017).

[34] A. Lamberts et al., Mon. Not. Roy. Astron. Soc. **490**, 5888 (2019).

[35] B. P. Abbott et al., ApJ **818**, L22 (2016).

[36] R. Penrose, Riv. Nuovo Cim. **1**, 252 (1969).

[37] B. J. Carr, 10.48550/ARXIV.ASTRO-PH/0511743 (2005).

[38] T. Damour and A. Vilenkin, Phys. Rev. D **71**, 063510 (2005).

[39] M. Safarzadeh et al., ApJ **888**, L3 (2019).

[40] V. D. Luca et al., Phys. Rev. Lett. **126**, 051101 (2021).

[41] B. Abbott et al., Phys. Rev. Lett. **123**, 161102 (2019).

[42] B. Abbott et al., Phys. Rev. Lett. **118**, 121101 (2017).

[43] B. Abbott et al., Phys. Rev. Lett. **116**, 131102 (2016).

[44] B. Abbott et al., Phys. Rev. D **100**, 061101 (2019).

[45] E. Abdalla et al., 10.48550/ARXIV.2203.06142 (2022).

[46] E. D. Valentino et al., Class. Quant. Grav. **38**, 153001 (2021).

[47] A. G. Riess et al., ApJ **876**, 85 (2019).

[48] W. D. Pozzo, A. Sesana, and A. Klein, Mon. Not. Roy. Astron. Soc. **475**, 3485 (2018).

[49] C. L. MacLeod and C. J. Hogan, Phys. Rev. D **77**, 043512 (2008).

[50] N. Tamanini et al., Journal of Cosmology and Astroparticle Physics **2016**, 002 (2016).

[51] B. Margalit and B. D. Metzger, ApJ **850**, L19 (2017).

[52] The LIGO Scientific Collaboration et al., 10.48550/ARXIV.2112.06861 (2021).

[53] V. Cardoso and P. Pani, Living Reviews in Relativity **22**, 10.1007/s41114-019-0020-4 (2019).

[54] R. Brito et al., Phys. Rev. D **96**, 064050 (2017).

[55] E. Thrane and C. Talbot, Publications of the Astronomical Society of Australia **36**, e010 (2019).

[56] R. Smith et al., Phys. Rev. Lett. **127**, 081102 (2021).

[57] J. Skilling, in Bayesian inference and maximum entropy methods in science and engineering: 24th international workshop on bayesian inference and maximum entropy methods in science and engineering, Vol. 735, edited by R. Fischer, R. Preuss, and U. V. Toussaint, American Institute of Physics Conference Series (Nov. 2004), pp. 395–405.

[58] J. Skilling, Bayesian Analysis **1**, 833 (2006).

[59] D. Foreman-Mackey et al., Publications of the Astronomical Society of the Pacific **125**, 306 (2013).

[60] J. S. Speagle, Mon. Not. Roy. Astron. Soc. **493**, 3132 (2020).

[61] J. Buchner et al., A&A **564**, A125 (2014).

[62] F. Feroz, M. P. Hobson, and M. Bridges, MNRAS **398**, 1601 (2009).

[63] E. Poisson, Phys. Rev. D **57**, 5287 (1998).

[64] P. N. McDermott, H. Van Horn, and C. Hansen, ApJ **325**, 725 (1988).

[65] B. Abbott et al., Phys. Rev. Lett. **122**, 061104 (2019).

[66] R. Nair et al., Phys. Rev. Lett. **123**, 191101 (2019).

[67] S. E. Perkins et al., Phys. Rev. D **104**, 024060 (2021).

[68] The LIGO Scientific Collaboration and The Virgo Collaboration, 10.48550/ARXIV.1203.2674 (2012).

[69] L. S. Finn, Phys. Rev. D **53**, 2878 (1996).

[70] R. E. Kass and A. E. Raftery, Journal of the American Statistical Association **90**, 773 (1995).

[71] N. Lartillot and H. Philippe, Systematic Biology **55**, 195 (2006).

[72] P. Liu et al., Water Resources Research **52**, 734 (2016).

[73] J. M. Dickey, The Annals of Mathematical Statistics **42**, 204 (1971).

[74] E.-J. Wagenmakers et al., Cognitive Psychology **60**, 158 (2010).

[75] L. S. Finn, Phys. Rev. D **46**, 5236 (1992).

[76] C. Cutler and É. E. Flanagan, Phys. Rev. D **49**, 2658 (1994).

[77] C. Cutler, Phys. Rev. D **57**, 7089 (1998).

[78] M. Vallisneri, Phys. Rev. D **77**, 042001 (2008).

[79] E. Berti, V. Cardoso, and A. O. Starinets, Class. Quant. Grav. **26**, 163001 (2009).

[80] A. Buonanno et al., Phys. Rev. D **80**, 084043 (2009).

[81] G. Pratten et al., Phys. Rev. D **103**, 104056 (2021).

[82] S. Isoyama, R. Sturani, and H. Nakano, in *Handbook of gravitational wave astronomy* (Springer Singapore, 2021), pp. 1–49.

[83] LIGO Scientific Collaboration, *LIGO Algorithm Library - LALSuite*, (2018) `https://git.ligo.org/lscsoft/lalsuite`.

[84] E. Poisson and C. M. Will, *Gravity: newtonian, post-newtonian, relativistic* (Cambridge University Press, 2014).

[85] A. Buonanno and T. Damour, Phys. Rev. D **59**, 084006 (1999).

[86] N. Yunes and F. Pretorius, Phys. Rev. D **80**, 122003 (2009).

[87] K. Chatziioannou, N. Yunes, and N. Cornish, Phys. Rev. D **86**, 022004 (2012).

[88] S. Tahura and K. Yagi, Phys. Rev. D **98**, 084042 (2018).

[89] É. É. Flanagan and T. Hinderer, Phys. Rev. D **77**, 021502 (2008).

[90] J. Vines, E. E. Flanagan, and T. Hinderer, Phys. Rev. D **83**, 084051 (2011).

[91] A. Nagar et al., Phys. Rev. D **98**, 104052 (2018).

[92] T. Damour and A. Nagar, Phys. Rev. D **81**, 084016 (2010).

[93] D. Bini, T. Damour, and G. Faye, Phys. Rev. D **85**, 124034 (2012).

[94] H. Yang, Phys. Rev. D **100**, 064023 (2019).

[95] K. D. Kokkotas and B. G. Schmidt, Living Reviews in Relativity **2**, 2 (1999).

[96] G. Pratten, P. Schmidt, and T. Hinderer, Nature Communications **11**, 2553 (2020).

[97] D. Lai, MNRAS **270**, 611 (1994).

[98] H. Yu and N. N. Weinberg, Mon. Not. Roy. Astron. Soc. **464**, 2622 (2017).

[99] W. C. G. Ho and D. Lai, Mon. Not. Roy. Astron. Soc. **308**, 153 (1999).

[100] D. Lai and Y. Wu, Phys. Rev. D **74**, 024007 (2006).

[101] E. E. Flanagan and E. Racine, Phys. Rev. D **75**, 044001 (2007).

[102] E. Poisson, Phys. Rev. D **101**, 104028 (2020).

[103] M. Burgay et al., Nature **426**, 531 (2003).

[104] J. J. Andrews and I. Mandel, ApJL **880**, L8 (2019).

[105] D. Tsang et al., Phys. Rev. Lett. **108**, 011102 (2012).

[106] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **119**, 161101 (2017).

[107] R. Abbott et al., ApJL **915**, L5 (2021).

[108] B. Allen et al., Phys. Rev. D **85**, 122006 (2012).

[109] B. P. Abbott et al., Phys. Rev. D **100**, 104036 (2019).

[110] É. É. Flanagan and É. Racine, Phys. Rev. D **75**, 044001 (2007).

[111] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **122**, 061104 (2019).

[112] H. Yang, W. E. East, and L. Lehner, ApJ **856**, [Erratum: Astrophys.J. 870, 139 (2019)], 110 (2018).

[113] A. Hook, Phys. Rev. Lett. **120**, 261802 (2018).

[114] L. D. Luzio et al., Journal of High Energy Physics **2021**, 184 (2021).

[115] A. Hook and J. Huang, Journal of High Energy Physics **06**, 036 (2018).

[116] J. Huang et al., Phys. Rev. D **99**, 063013 (2019).

[117] K. Blum et al., Phys. Lett. B **737**, 30 (2014).

[118] P. Kanti et al., Phys. Rev. D **54**, 5049 (1996).

[119] B. Shiralilou et al., Class. Quant. Grav. **39**, 035002 (2022).

[120] N. Yunes and F. Pretorius, Phys. Rev. D **80**, 122003 (2009).

[121] H.-T. Wang et al., Phys. Rev. D **104**, 024015 (2021).

[122] R. Abbott et al., ApJ **896**, L44 (2020).

[123] A. Buonanno and T. Damour, Phys. Rev. D **59**, 084006 (1999).

[124] A. Buonanno and T. Damour, Phys. Rev. D **62**, 064015 (2000).

[125] T. Damour, P. Jaranowski, and G. Schaefer, Phys. Rev. D **62**, 084011 (2000).

[126] T. Damour, Phys. Rev. D **64**, 124013 (2001).

[127] T. Damour, P. Jaranowski, and G. Schäfer, Phys. Rev. D **91**, 084024 (2015).

[128] SXS, *Sxs gravitational waveform database*, `https://data.black-holes.org/waveforms/index.html`.

[129] M. Boyle et al., Phys. Rev. D **78**, 104020 (2008).

[130] P. Amaro-Seoane et al., Class. Quant. Grav. **24**, R113 (2007).

[131] C. Hopman and T. Alexander, ApJ **629**, 362 (2005).

[132] M. Preto and P. Amaro-Seoane, ApJL **708**, L42 (2010).

[133] B. Bar-Or and T. Alexander, ApJ **820**, 129 (2016).

[134] P. Amaro-Seoane and M. Preto, Class. Quant. Grav. **28**, 094017 (2011).

[135] P. Amaro-Seoane, Living Reviews in Relativity **21**, 4 (2018).

[136] P. Amaro-Seoane, 10.48550/ARXIV.2011.03059 (2020).

[137] S. Babak et al., Phys. Rev. D **95**, 103012 (2017).

[138] J. R. Gair et al., J. Phys. Conf. Ser. **840**, edited by D. Giardini and P. Jetzer, 012021 (2017).

[139] H.-M. Fan et al., Phys. Rev. D **102**, 063016 (2020).

[140] Z. Pan and H. Yang, Phys. Rev. D **103**, 103018 (2021).

[141] L. Zwick et al., MNRAS **495**, 2321 (2020).

[142] L. Zwick et al., Mon. Not. Roy. Astron. Soc. **506**, 1007 (2021).

[143] V. Vázquez-Aceves et al., Mon. Not. Roy. Astron. Soc. **510**, 2379 (2021).

[144] A. Galametz et al., ApJ **694**, 1309 (2009).

[145] M. Macuga et al., ApJ **874**, 54 (2019).

[146] P. Goldreich and S. Tremaine, ApJ **233**, 857 (1979).

[147] P. Goldreich and S. Tremaine, ApJ **241**, 425 (1980).

[148] H. Tanaka, T. Takeuchi, and W. R. Ward, ApJ **565**, 1257 (2002).

[149] H. Tanaka and W. R. Ward, ApJ **602**, 388 (2004).

[150] L. Arzamasskiy, Z. Zhu, and J. M. Stone, MNRAS **475**, 3201 (2018).

[151] Z. Zhu, MNRAS **483**, 4221 (2019).

[152] B. Kocsis, N. Yunes, and A. Loeb, Phys. Rev. D **84**, 024032 (2011).

[153] N. Yunes et al., Phys. Rev. Lett. **107**, 171103 (2011).

[154] L. Barack and C. Cutler, Phys. Rev. D **69**, 082005 (2004).

[155] A. J. Chua and J. R. Gair, Class. Quant. Grav. **32**, 232002 (2015).

[156] A. J. Chua, C. J. Moore, and J. R. Gair, Phys. Rev. D **96**, 044005 (2017).

[157] E. E. Flanagan and T. Hinderer, Phys. Rev. D **77**, 021502 (2008).

[158] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **121**, 161101 (2018).

[159] M. C. Miller et al., ApJL **887**, L24 (2019).

[160] L. Rezzolla, E. R. Most, and L. R. Weih, ApJ **852**, L25 (2018).

[161] D. Radice et al., ApJ **852**, L29 (2018).

[162] T. Strohmayer et al., ApJ **375**, 679 (1991).

[163] D. G. Ravenhall, C. J. Pethick, and J. R. Wilson, Phys. Rev. Lett. **50**, 2066 (1983).

[164] M.-a. Hashimoto, H. Seki, and M. Yamada, *Shape of nuclei in the crust of neutron star*, Vol. 71, 2 (Oxford University Press, 1984), pp. 320–326.

[165] A. I. Chugunov and C. J. Horowitz, MNRAS **407**, L54 (2010).

[166] C. Thompson, H. Yang, and N. Ortiz, ApJ **841**, 54 (2017).

[167] G. Chabrier, ApJ **414**, 695 (1993).

[168] S. L. Shapiro and S. A. Teukolsky, *Cooling of neutron stars* (John Wiley & Sons, Ltd, 1983) Chap. 6-11, pp. 127–334.

[169] F. Douchin and P. Haensel, Astronomy & Astrophysics **380**, 151 (2001).

[170] J. S. Read et al., Phys. Rev. D **79**, 124032 (2009).

[171] A. Passamonti and N. Andersson, MNRAS **419**, 638 (2012).

[172] C. J. Horowitz and J. Piekarewicz, Phys. Rev. Lett. **86**, 5647 (2001).

[173] J. Xu et al., ApJ **697**, 1549 (2009).

[174] C. C. Moustakidis et al., Phys. Rev. C **81**, 065803 (2010).

[175] T. G. Cowling, MNRAS **101**, 367 (1941).

[176] S. Yoshida and Y. Kojima, MNRAS **289**, 117 (1997).

[177] C. Chirenti, G. H. de Souza, and W. Kastaun, Phys. Rev. D **91**, 044034 (2015).

[178] E. F. Brown et al., Phys. Rev. Lett. **120**, 182701 (2018).

[179] C. M. Biwer et al., PASP **131**, 024503 (2019).

[180] A. Buonanno et al., Phys. Rev. D **80**, 084043 (2009).

[181] H. Yang et al., Phys. Rev. Lett. **118**, 161101 (2017).

[182] C. Palenzuela et al., Phys. Rev. D **89**, 044024 (2014).

[183] L. Annulli, V. Cardoso, and L. Gualtieri, Phys. Rev. D **99**, 044038 (2019).

[184] R. F. P. Mendes and N. Ortiz, Phys. Rev. Lett. **120**, 201104 (2018).

[185] J. Huang et al., Phys. Rev. D **99**, 063013 (2019).

[186] M. Ackermann et al. (Fermi GBM/LAT), Nature **462**, 331 (2009).

[187] E. Troja, S. Rosswog, and N. Gehrels, ApJ **723**, 1711 (2010).

[188] B. P. Abbott et al., ApJL **892**, L3 (2020).

[189] H. Yang and Y.-C. Zou, ApJ **893**, L31 (2020).

[190] P. N. McDermott et al., ApJL **297**, L37 (1985).

[191] R. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).

[192] R. Peccei and H. R. Quinn, Phys. Rev. D **16**, 1791 (1977).

[193] S. Weinberg, Phys. Rev. Lett. **40**, 223 (1978).

[194] F. Wilczek, Phys. Rev. Lett. **40**, 279 (1978).

[195] P. Svrcek and E. Witten, Journal of High Energy Physics **06**, 051 (2006).

[196] A. Arvanitaki et al., Phys. Rev. D **81**, 123530 (2010).

[197] J. Preskill, M. B. Wise, and F. Wilczek, Phys. Lett. B **120**, edited by M. A. Srednicki, 127 (1983).

[198] L. F. Abbott and P. Sikivie, Phys. Lett. B **120**, edited by M. A. Srednicki, 133 (1983).

[199] M. Dine and W. Fischler, Phys. Lett. B **120**, edited by M. A. Srednicki, 137 (1983).

[200] G. G. Raffelt, Phys. Rev. D **33**, 897 (1986).

[201] G. G. Raffelt, Phys. Rept. **198**, 1 (1990).

[202] V. Anastassopoulos et al. (CAST), Nature Phys. **13**, 584 (2017).

[203] G. Raffelt and D. Seckel, Phys. Rev. Lett. **60**, 1793 (1988).

[204] J. H. Chang, R. Essig, and S. D. McDermott, Journal of High Energy Physics **09**, 051 (2018).

[205] P. Sikivie, Phys. Rev. Lett. **51**, edited by M. A. Srednicki, [Erratum: Phys. Rev. Lett. 52, 695 (1984)], 1415 (1983).

[206] C. Abel et al., Phys. Rev. X **7**, 041034 (2017).

[207] N. Du et al. (ADMX), Phys. Rev. Lett. **120**, 151301 (2018).

[208] S. Sibiryakov, P. Sørensen, and T.-T. Yu, Journal of High Energy Physics **12**, 075 (2020).

[209] A. Arvanitaki, M. Baryakhtar, and X. Huang, Phys. Rev. D **91**, 084011 (2015).

[210] A. Gruzinov, 10.48550/ARXIV.1604.06422 (2016).

[211] H. Davoudiasl and P. B. Denton, Phys. Rev. Lett. **123**, 021102 (2019).

[212] M. J. Stott, 10.48550/ARXIV.2009.07206 (2020).

[213] K. K. Ng et al., Phys. Rev. Lett. **126**, 151102 (2021).

[214] M. Baryakhtar et al., Phys. Rev. D **103**, 095019 (2021).

[215] Y. Chen et al., Phys. Rev. Lett. **124**, 061102 (2020).

[216] A. Arvanitaki et al., Phys. Rev. D **95**, 043001 (2017).

[217] S. J. Zhu et al., Phys. Rev. D **102**, 063020 (2020).

[218] R. Brito et al., Phys. Rev. Lett. **119**, 131101 (2017).

[219] R. Brito et al., Phys. Rev. D **96**, 064050 (2017).

[220] L. Tsukada et al., Phys. Rev. D **99**, 103015 (2019).

[221] C. Palomba et al., Phys. Rev. Lett. **123**, 171101 (2019).

[222] L. Sun, R. Brito, and M. Isi, Phys. Rev. D **101**, [Erratum: Phys. Rev. D 102, 089902 (2020)], 063020 (2020).

[223] J. C. Bustillo et al., Phys. Rev. Lett. **126**, 081101 (2021).

[224] T. D. Cohen, R. J. Furnstahl, and D. K. Griegel, Phys. Rev. C **45**, 1881 (1992).

[225] J. M. Alarcon, J. Martin Camalich, and J. A. Oller, Phys. Rev. D **85**, 051503 (2012).

[226] L. Sampson, N. Yunes, and N. Cornish, Phys. Rev. D **88**, [Erratum: Phys. Rev. D 88, 089902 (2013)], 064056 (2013).

[227] D. Croon et al., ApJL **858**, L2 (2018).

[228] L. Sagunski et al., Phys. Rev. D **97**, 064016 (2018).

[229] S. Alexander et al., Class. Quant. Grav. **35**, 235012 (2018).

[230] J. Kopp et al., Journal of High Energy Physics **11**, 096 (2018).

[231] B. C. Seymour and K. Yagi, Class. Quant. Grav. **37**, 145008 (2020).

[232] B. C. Seymour and K. Yagi, Phys. Rev. D **102**, 104003 (2020).

[233] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **123**, 011102 (2019).

[234] C. M. Biwer et al., Publ. Astron. Soc. Pac. **131**, 024503 (2019).

[235] K. Yagi and N. Yunes, Phys. Rept. **681**, 1 (2017).

[236] A. Maselli et al., Phys. Rev. D **88**, 023007 (2013).

[237] M. Urbanec, J. C. Miller, and Z. Stuchlik, Mon. Not. Roy. Astron. Soc. **433**, 1903 (2013).

[238] K. Yagi and N. Yunes, Science **341**, 365 (2013).

[239] K. Yagi and N. Yunes, Phys. Rev. D **88**, 023009 (2013).

[240] K. Yagi and N. Yunes, Class. Quant. Grav. **33**, 13LT01 (2016).

[241] K. Chatziioannou, C.-J. Haster, and A. Zimmerman, Phys. Rev. D **97**, 104036 (2018).

[242] M. Cantiello et al., ApJL **854**, L31 (2018).

[243] M. Soares-Santos et al. (DES, Dark Energy Camera GW-EM), ApJL **848**, L16 (2017).

[244] Rich Abbott et al., SoftwareX **13**, 100658 (2021).

[245] T. Kumar Poddar, S. Mohanty, and S. Jana, Phys. Rev. D **101**, 083007 (2020).

[246] T. K. Poddar and S. Mohanty, Phys. Rev. D **102**, 083029 (2020).

[247] R. Abbott et al. (LIGO Scientific, Virgo), ApJL **896**, L44 (2020).

[248] B. P. Abbott et al., Phys. Rev. X **9**, 031040 (2019).

[249] R. Abbott et al., Phys. Rev. X **11**, 021053 (2021).

[250] The LIGO Scientific Collaboration et al., 10.48550/ARXIV.2108.01045 (2021).

[251] R. Abbott et al., ApJL **913**, L7 (2021).

[252] B. P. Abbott et al., ApJ **909**, 218 (2021).

[253] B. P. Abbott et al., Phys. Rev. D **100**, 061101 (2019).

[254] A. Ghosh, R. Brito, and A. Buonanno, Phys. Rev. D **103**, 124041 (2021).

[255] B. P. Abbott et al., Phys. Rev. Lett. **121**, 161101 (2018).

[256] F. Hernandez Vivanco et al., Phys. Rev. D **100**, 103009 (2019).

[257] B. P. Abbott et al., Phys. Rev. Lett. **122**, 061104 (2019).

[258] Z. Pan et al., Phys. Rev. Lett. **125**, 201102 (2020).

[259] S. Y. Lau and K. Yagi, Phys. Rev. D **103**, 063015 (2021).

[260] B. P. Abbott et al. (Virgo, LIGO Scientific), Phys. Rev. Lett. **116**, 221101 (2016).

[261] N. Yunes, K. Yagi, and F. Pretorius, Phys. Rev. D **94**, 084002 (2016).

[262] B. P. Abbott et al., Phys. Rev. D **100**, 104036 (2019).

[263] R. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. D **103**, 122002 (2021).

[264] J. Zhao et al., Phys. Rev. D **100**, 064034 (2019).

[265] R. Niu et al., ApJ **921**, 149 (2021).

[266] M. Okounkova et al., Phys. Rev. D **101**, 104016 (2020).

[267] M. Okounkova et al., 10.48550/ARXIV.2101.11153 (2021).

[268] S. Nojiri, S. D. Odintsov, and M. Sasaki, Phys. Rev. D **71**, 123509 (2005).

[269] K. Yagi, Phys. Rev. D **86**, 081504 (2012).

[270] G. Antoniou, A. Bakopoulos, and P. Kanti, Phys. Rev. D **97**, 084037 (2018).

[271] G. Antoniou, A. Bakopoulos, and P. Kanti, Phys. Rev. Lett. **120**, 131102 (2018).

[272] K. Yagi et al., Phys. Rev. D **85**, [Erratum: Phys.Rev.D 93, 029902 (2016)], 064022 (2012).

[273] E. Barausse and K. Yagi, Phys. Rev. Lett. **115**, 211105 (2015).

[274] P. Kanti et al., Phys. Rev. D **54**, 5049 (1996).

[275] T. Torii, H. Yajima, and K.-i. Maeda, Phys. Rev. D **55**, 739 (1997).

[276] K.-i. Maeda, N. Ohta, and Y. Sasagawa, Phys. Rev. D **80**, 104032 (2009).

[277] M. Herrero-Valea, Journal of High Energy Physics **2022**, 75 (2022).

[278] S. Odintsov, V. Oikonomou, and F. Fronimos, Nuclear Physics B **958**, 115135 (2020).

[279] V. K. Oikonomou, Class. Quant. Grav. **38**, 195025 (2021).

[280] D. D. Doneva and S. S. Yazadjiev, Phys. Rev. Lett. **120**, 131103 (2018).

[281] D. D. Doneva and S. S. Yazadjiev, Journal of Cosmology and Astroparticle Physics **2018**, 011 (2018).

[282] H. O. Silva et al., Phys. Rev. Lett. **120**, 131104 (2018).

[283] H. O. Silva et al., Phys. Rev. D **99**, 064011 (2019).

[284] A. Saffer and K. Yagi, Phys. Rev. D **104**, 124052 (2021).

[285] K. Yamada, T. Narikawa, and T. Tanaka, PTEP **2019**, 103E01 (2019).

[286] Z. Carson, B. C. Seymour, and K. Yagi, Class. Quant. Grav. **37**, 065008 (2020).

[287] G. Pratten et al., Phys. Rev. D **102**, 064001 (2020).

[288] C. García-Quirós et al., Phys. Rev. D **102**, 064002 (2020).

[289] S. E. Perkins et al., Phys. Rev. D **104**, 024060 (2021).

[290] N. Sennett, S. Marsat, and A. Buonanno, Phys. Rev. D **94**, 084003 (2016).

[291] W. E. East and J. L. Ripley, Phys. Rev. Lett. **127**, 101102 (2021).

[292] W. E. East and J. L. Ripley, Phys. Rev. D **103**, 044040 (2021).

[293] K. Yagi, L. C. Stein, and N. Yunes, Phys. Rev. D **93**, 024010 (2016).

[294] E. Berti, K. Yagi, and N. Yunes, Gen. Rel. Grav. **50**, 46 (2018).

[295] M. Okounkova, Phys. Rev. D **102**, 084046 (2020).

[296] H. Zhang et al., Phys. Rev. D **95**, 104043 (2017).

[297] L. Amendola, C. Charmousis, and S. C. Davis, Journal of Cosmology and Astroparticle Physics **2007**, 004 (2007).

[298] R. Abbott et al., Phys. Rev. X **11**, 021053 (2021).

[299] I. Tews et al., ApJ **908**, L1 (2021).

[300] D.-S. Shao et al., Phys. Rev. D **102**, 063006 (2020).

[301] A. Nathanail, E. R. Most, and L. Rezzolla, ApJL **908**, L28 (2021).

[302] R. Abbott et al., SoftwareX **13**, 100658 (2021).

[303] A. Nitz et al., *Gwastro/pycbc: pycbc release 1.18.1*, version v1.18.1, May 2021.

[304] C. M. Biwer et al., Publications of the Astronomical Society of the Pacific **131**, 024503 (2019).

[305] D. Foreman-Mackey et al., PASP **125**, 306 (2013).

[306] E. Berti, A. Buonanno, and C. M. Will, Phys. Rev. D **71**, 084025 (2005).

[307] S. Perkins and N. Yunes, 10.48550/ARXIV.2201.02542 (2022).

[308] S. Alexander and N. Yunes, Phys. Rept. **480**, 1 (2009).

[309] K. Yagi, N. Yunes, and T. Tanaka, Phys. Rev. Lett. **109**, [Erratum: Phys.Rev.Lett. 116, 169902 (2016), Erratum: Phys.Rev.Lett. 124, 029901 (2020)], 251105 (2012).

[310] F.-L. Julié and E. Berti, Phys. Rev. D **100**, 104061 (2019).

[311] G. González, Class. Quant. Grav. **21**, S691 (2004).

[312] F. Acernese et al., Class. Quant. Grav. **21**, S709 (2004).

[313] G. Pratten et al., Phys. Rev. D **102**, 064001 (2020).

[314] T. Dietrich, S. Bernuzzi, and W. Tichy, Phys. Rev. D **96**, 121501 (2017).

[315] T. Dietrich et al., Phys. Rev. D **100**, 044003 (2019).

[316] P. Ajith et al., Phys. Rev. Lett. **106**, 241101 (2011).

[317] Y. Pan et al., Phys. Rev. D **89**, 084006 (2014).

[318] S. E. Field et al., Phys. Rev. X **4**, 031006 (2014).

[319] M. van de Meent and H. P. Pfeiffer, Phys. Rev. Lett. **125**, 181101 (2020).

[320] S. Khan et al., Phys. Rev. D **93**, 044007 (2016).

[321] P. Amaro-Seoane et al., 10.48550/ARXIV.1702.00786 (2017).

[322] A. Le Tiec et al., Phys. Rev. Lett. **107**, 141101 (2011).

[323] M. Favata, S. A. Hughes, and D. E. Holz, ApJL **607**, L5 (2004).

[324] A. Le Tiec et al., Phys. Rev. D **88**, 124027 (2013).

[325] A. Le Tiec, International Journal of Modern Physics D **23**, 1430022 (2014).

[326] A. Zimmerman, A. G. Lewis, and H. P. Pfeiffer, Phys. Rev. Lett. **117**, 191101 (2016).

[327] M. Van De Meent, Phys. Rev. Lett. **118**, 011101 (2017).

[328] A. Le Tiec and P. Grandclément, Class. Quant. Grav. **35**, 144002 (2018).

[329] N. E. Rifat et al., Phys. Rev. D **101**, 081502 (2020).

[330] P. Anninos et al., Phys. Rev. D **52**, 4462 (1995).

[331] M. J. Fitchett and S. Detweiler, Mon. Not. Roy. Astron. Soc. **211**, 933 (1984).

[332] U. Sperhake et al., Phys. Rev. D **84**, 084038 (2011).

[333] A. Le Tiec, E. Barausse, and A. Buonanno, Phys. Rev. Lett. **108**, 131103 (2012).

[334] A. Nagar, Phys. Rev. D **88**, 121501 (2013).

[335] A. Ori and K. S. Thorne, Phys. Rev. D **62**, 124022 (2000).

[336] K. S. Thorne, Phys. Rev. D **58**, 124031 (1998).

[337] K. S. Thorne, Reviews of Modern Physics **52**, 299 (1980).

[338] S. Detweiler, Phys. Rev. Lett. **86**, 1931 (2001).

[339] W. G. Dixon, Il Nuovo Cimento (1955-1965) **34**, 317 (1964).

[340] E. Poisson, Phys. Rev. D **70**, 084044 (2004).

[341] H. Tagoshi, S. Mano, and E. Takasugi, Progress of theoretical physics **98**, 829 (1997).

[342] J. Steinhoff and D. Puetzfeld, Phys. Rev. D **81**, 044019 (2010).

[343] J. Ehlers and E. Rudolph, General Relativity and Gravitation **8**, 197 (1977).

[344] J. Steinhoff and D. Puetzfeld, Phys. Rev. D **86**, 044033 (2012).

[345] H. Yang et al., Phys. Rev. D **98**, 044007 (2018).

[346] *Black Hole Perturbation Toolkit*, (http://bhptoolkit.org/).

[347] *Mathematica notebook and C++ code for studying the tidal effects of compact-object binary systems*, (https://zenodo.org/).

[348] T. Damour and A. Nagar, Phys. Rev. D **80**, 084035 (2009).

[349] S. A. Teukolsky, ApJ **185**, 635 (1973).

[350] W. Kinnersley, Journal of Mathematical Physics **10**, 1195 (1969).

[351] R. A. Breuer, *Gravitational perturbation theory and synchrotron radiation.* Vol. 44 (1975).

[352] J. E. Thompson et al., Phys. Rev. D **101**, 124059 (2020).

[353] F. Pannarale et al., Phys. Rev. D **92**, 084050 (2015).

[354] M. Boyle et al., Class. Quant. Grav. **36**, 195006 (2019).

[355] M. Agathos et al., Phys. Rev. D **92**, 023012 (2015).

[356] J. J. Andrews and I. Mandel, ApJL **880**, L8 (2019).

[357] H. Yang, W. E. East, and L. Lehner, ApJ **856**, 110 (2018).

[358] F. Foucart et al., Phys. Rev. D **99**, 044008 (2019).

[359] T. Damour, Phys. Rev. D **102**, 024060 (2020).

[360] D. Bini, T. Damour, and A. Geralico, Phys. Rev. D **102**, 024061 (2020).

[361] D. Bini, T. Damour, and A. Geralico, Phys. Rev. D **102**, 024062 (2020).

[362] A. Pound et al., Phys. Rev. Lett. **124**, 021101 (2020).

[363] H. Yang and M. Casals, Phys. Rev. D **96**, 083015 (2017).

[364] B. Bonga, H. Yang, and S. A. Hughes, Phys. Rev. lett. **123**, 101103 (2019).

[365] N. Yunes et al., Phys. Rev. lett. **107**, 171103 (2011).

[366] E. Barausse, V. Cardoso, and P. Pani, Phys. Rev. D **89**, 104059 (2014).

[367] Q. Henry, G. Faye, and L. Blanchet, Phys. Rev. D **101**, 064047 (2020).

[368] Q. Henry, G. Faye, and L. Blanchet, Phys. Rev. D **102**, 044033 (2020).

[369] S. Hild et al., Class. Quant. Grav. **28**, 094013 (2011).

[370] D. Reitze et al., 10.48550/ARXIV.1907.04833 (2019).

[371] M. Pürrer and C.-J. Haster, Phys. Rev. Research **2**, 023151 (2020).

[372] S. A. Hughes, Phys. Rev. D **61**, 084004 (2000).

[373] M. Shibata et al., Phys. Rev. D **51**, 1646 (1995).

[374] M. Arca-Sedda and R. Capuzzo-Dolcetta, MNRAS **483**, 152 (2019).

[375] M. A. Sedda, P. A. Seoane, and X. Chen, A&A **652**, A54 (2021).

[376] P. Amaro-Seoane, Phys. Rev. D **99**, 123025 (2019).

[377] P. Amaro-Seoane, 10.48550/ARXIV.2011.03059 (2020).

[378] Z. Khakhaleva-Li and C. J. Hogan, 10.48550/ARXIV.2006.00438 (2020).

[379] K. Glampedakis and S. Babak, Class. Quant. Grav. **23**, 4167 (2006).

[380] L. Barack and C. Cutler, Phys. Rev. D **75**, 042003 (2007).

[381] B. Bonga, H. Yang, and S. A. Hughes, Phys. Rev. Lett. **123**, 101103 (2019).

[382] H. Yang et al., Phys. Rev. D **100**, 124056 (2019).

[383] E. Barausse, V. Cardoso, and P. Pani, Phys. Rev. D **89**, 104059 (2014).

[384] E. Berti and M. Volonteri, Astrophys. J. **684**, 822 (2008).

[385] J. R. Gair, C. Tang, and M. Volonteri, Phys. Rev. D **81**, 104014 (2010).

[386] Z. Pan and H. Yang, Astrophys. J. **901**, 163 (2020).

[387] D. Laghi et al., Mon. Not. Roy. Astron. Soc. **508**, 4512 (2021).

[388] J. R. Gair et al., J. Phys. Conf. Ser. **840**, edited by D. Giardini and P. Jetzer, 012021 (2017).

[389] M. C. Miller et al., ApJL **631**, L117 (2005).

[390] X. Chen and W.-B. Han, Communications Physics **1**, 53 (2018).

[391] Y. Y. Wang et al., ApJL **886**, L22 (2019).

[392] Y. Raveh and H. B. Perets, MNRAS **501**, 5012 (2021).

[393] P. C. Peters and J. Mathews, Physical Review **131**, 435 (1963).

[394] V. Vázquez-Aceves et al., Mon. Not. Roy. Astron. Soc. **510**, 2379 (2021).

[395] Y. Levin, 10.48550/ARXIV.ASTRO-PH/0307084 (2003).

[396] G. Sigl, J. Schnittman, and A. Buonanno, Phys. Rev. D **75**, 024034 (2007).

[397] Y. Levin, Mon. Not. Roy. Astron. Soc. **374**, 515 (2007).

[398] S. Chandrasekhar, ApJ **97**, 255 (1943).

[399] E. C. Ostriker, ApJ **513**, 252 (1999).

[400] F. S. Masset, MNRAS **472**, 4204 (2017).

[401] A. M. Hankla, Y.-F. Jiang, and P. J. Armitage, ApJ **902**, 50 (2020).

[402] S. .-. Paardekooper et al., MNRAS **401**, 1950 (2010).

[403] D. Syer and C. J. Clarke, MNRAS **277**, 758 (1995).

[404] X.-H. Yang et al., ApJ **780**, 79 (2014).

[405] J. C. McKinney et al., MNRAS **441**, 3177 (2014).

[406] A. Gruzinov, Y. Levin, and C. D. Matzner, MNRAS **492**, 2755 (2020).

[407] X. Li et al., MNRAS **494**, 2327 (2020).

[408] E. Sirko and J. Goodman, MNRAS **341**, 501 (2003).

[409] T. A. Thompson, E. Quataert, and N. Murray, ApJ **630**, 167 (2005).

[410] H. Cohn and R. M. Kulsrud, ApJ **226**, 1087 (1978).

[411] H. Cohn, ApJ **234**, 1036 (1979).

[412] J. Binney and S. Tremaine, *Galactic dynamics* (1987).

[413] K. Gültekin et al., ApJ **698**, 198 (2009).

[414] A. Soltan, MNRAS **200**, 115 (1982).

[415] A. King and C. Nixon, MNRAS **453**, L46 (2015).

[416] K. Schawinski et al., MNRAS **451**, 2517 (2015).

[417] S. Tremaine et al., Astronomical Journal **107**, 634 (1994).

[418] H. Tagawa, Z. Haiman, and B. Kocsis, ApJ **898**, 25 (2020).

[419] E. Barausse, MNRAS **423**, 2533 (2012).

[420] A. Sesana et al., ApJ **794**, 104 (2014).

[421] F. Antonini, E. Barausse, and J. Silk, ApJL **806**, L8 (2015).

[422] F. Antonini, E. Barausse, and J. Silk, ApJ **812**, 72 (2015).

[423] L. S. Finn, Phys. Rev. D **46**, 5236 (1992).

[424] P. Amaro-Seoane et al., `10.48550/ARXIV.1702.00786` (2017).

[425] E. Barausse, V. Cardoso, and P. Pani, Phys. Rev. D **89**, 104059 (2014).

[426] O. A. Hannuksela et al., Nature Astron. **3**, 447 (2019).

[427] J. Zhang and H. Yang, Phys. Rev. D **101**, 043020 (2020).

[428] R. D. Blandford and R. L. Znajek, MNRAS **179**, 433 (1977).

[429] R. D. Blandford and D. G. Payne, Mon. Not. Roy. Astron. Soc. **199**, 883 (1982).

[430] J. M. Bardeen and J. A. Petterson, ApJL **195**, L65 (1975).

[431] C. D. Bailyn et al., ApJ **499**, 367 (1998).

[432] F. Özel et al., ApJ **725**, 1918 (2010).

[433] F. Özel et al., ApJ **757**, 55 (2012).

[434] L. Kreidberg et al., ApJ **757**, 36 (2012).

[435] K. Belczynski et al., ApJ **757**, 91 (2012).

[436] C. L. Fryer et al., ApJ **749**, 91 (2012).

[437] LIGO Scientific Collaboration and Virgo Collaboration, ApJL **896**, L44 (2020).

[438] The LIGO Scientific Collaboration et al., 10.48550/ARXIV.2111.03606 (2021).

[439] T. A. Thompson et al., Science **366**, 637 (2019).

[440] T. Jayasinghe et al., MNRAS **504**, 2577 (2021).

[441] B. Giesers et al., MNRAS **475**, L15 (2018).

[442] B. Giesers et al., A&A **632**, A3 (2019).

[443] Ł. Wyrzykowski and I. Mandel, A&A **636**, A20 (2020).

[444] M. Zevin et al., ApJL **899**, L1 (2020).

[445] P. Drozda et al., 10.48550/ARXIV.2009.06655 (2020).

[446] B. Carr et al., Physics of the Dark Universe **31**, 100755 (2021).

[447] S. Clesse and J. Garcia-Bellido, 10.48550/ARXIV.2007.06481 (2020).

[448] K. Jedamzik, Phys. Rev. Lett. **126**, 051302 (2021).

[449] M. Safarzadeh et al., ApJL **888**, L3 (2020).

[450] G. Fragione, A. Loeb, and F. A. Rasio, ApJL **895**, L15 (2020).

[451] Y. Yang et al., ApJL **901**, L34 (2020).

[452] S. Rastello et al., MNRAS **497**, 1563 (2020).

[453] T. Broadhurst, J. M. Diego, and G. F. Smoot, 10.48550/ARXIV.2006.13219 (2020).

[454] A. S. Hamers et al., MNRAS **506**, 5345 (2021).

[455] W. Lu, P. Beniamini, and C. Bonnerot, MNRAS **500**, 1817 (2021).

[456] B. Liu and D. Lai, MNRAS **502**, 2049 (2021).

[457] H. Tagawa et al., ApJ **908**, 194 (2021).

[458] M. Arca Sedda, ApJL **908**, L38 (2021).

[459] I. Mandel and F. S. Broekgaarden, Living Reviews in Relativity **25**, 1 (2022).

[460] C. Cutler, D. Kennefick, and E. Poisson, Phys. Rev. D **50**, 3816 (1994).

[461] D. F. Chernoff and M. D. Weinberg, ApJ **351**, 121 (1990).

[462] S. L. Shapiro and A. B. Marchant, ApJ **225**, 603 (1978).

[463] P. C. Peters, Phys. Rev. **136**, B1224 (1964).

[464] W. Dehnen, MNRAS **265**, 250 (1993).

[465] R. Perna et al., ApJL **878**, L1 (2019).

[466] S. Tremaine et al., ApJ **574**, 740 (2002).

[467] Z. Pan and H. Yang, Phys. Rev. D **103**, 103018 (2021).

[468] T. Alexander and C. Hopman, ApJ **697**, 1861 (2009).

[469] S. Naoz et al., ApJL **927**, L18 (2022).

[470] Z. Pan and H. Yang, ApJ **923**, 173 (2021).

[471] A. S. Jermyn et al., ApJ **914**, 105 (2021).

[472] S. A. Balbus and J. F. Hawley, ApJ **376**, 214 (1991).

[473] S. A. Balbus and J. F. Hawley, Reviews of Modern Physics **70**, 1 (1998).

[474] R. G. Martin et al., New Astronomy **70**, 7 (2019).

[475] Z. Pan, Z. Lyu, and H. Yang, Phys. Rev. D **104**, 063007 (2021).

[476] L. Barack and C. Cutler, Phys. Rev. D **69**, 082005 (2004).

[477] A. J. Chua and J. R. Gair, Class. Quant. Grav. **32**, 232002 (2015).

[478] A. J. Chua, C. J. Moore, and J. R. Gair, Phys. Rev. D **96**, 044005 (2017).

[479] H.-K. Guo, J. Shu, and Y. Zhao, Phys. Rev. D **99**, 023001 (2019).

[480] S. Barsanti et al., Phys. Rev. Lett. **128**, 111104 (2022).

[481] J. F. Navarro, C. S. Frenk, and S. D. M. White, ApJ **462**, 563 (1996).

[482] H. Tagawa et al., ApJ **927**, 41 (2022).

[483] L. Ferrarese, ApJ **578**, 90 (2002).

[484] A. A. Dutton and A. V. Macciò, MNRAS **441**, 3359 (2014).