

SISSA

Scuola
Internazionale
Superiore di
Studi Avanzati

International School for Advanced Studies
Physics Area - PhD course in Astroparticle and Gravitational Physics

Probing the astrophysical environment with millihertz Gravitational Waves

Candidate:

Lorenzo Copparoni

Supervisor:

Enrico Barausse

Academic Year 2025/2026



Abstract

With the expected launch of LISA in 2035, gravitational waves in the millihertz band will be detected. One of the most promising avenue which will be opened is the study of the astrophysical environment surrounding massive black holes and stellar-mass compact objects through the wave signal alone. Given the complexity and the many possible physical cases resulting in the interplay between matter and black hole binaries, it is of the utmost importance to develop pipelines and formalisms that are able to correctly capture the presence of an environment in an agnostic way, and to minimize any possible bias coming from the neglect of these additional physical properties. This thesis works in this direction, mainly focusing on two classes of sources in this band: extreme mass ratio inspirals, which are the most promising sources for the detection of beyond-GR physics, and massive black hole binaries. In the first case, we did a systematic study through the injection of realistic signals, in which the environmental contribution is generated through numerical simulations including stochastic features, against agnostic templates. We found that little to no systematics is present even with the simple agnostic templates. In a similar direction we also developed a consistency test for long lived signals able to detect missing physical properties. In the massive black hole binary case, instead, we performed a population study on different scenarios, and we assessed the ability of LISA to measure deviations coming from the astrophysical environment. We also assessed LISA's ability to distinguish between a modified GR and an environment injection for a full population catalogue, finding that the most preferred population description is always the vacuum one.

Preface

List of Publications

These work produced during my PhD are discussed in this work:

1. Lorenzo Copparoni, Enrico Barausse, Lorenzo Speri, Laura Sberna, Andrea Derdzinski, Implications of stochastic gas torques for asymmetric binaries in the LISA band, *Phys. Rev. D*, 111(10):104079, 2025.
2. Lorenzo Copparoni, Enrico Barausse. Environmental effects vs. modified gravity in the LISA massive black hole binary population, *arXiv:2607.00845*
3. Lorenzo Copparoni, Rohit S. Chanramouli, Enrico Barausse. When vacuum breaks: a self-consistency test for astrophysical environments in extreme mass ratio inspirals, *Phys. Rev. Lett.*, 137(2):021405, 2026.

Additional works which I have been done or contributing to is currently in preparation:

1. Rahul Srinivasan , Lorenzo Copparoni, Chantal Pitte, Enrico Barausse, Roberto Trotta. Low Latency Alerts of Massive Black Hole Binaries in LISA, (*in Preparation.*).
2. Lorenzo Copparoni, Rohit S. Chanramouli, Francisco Duque. Environmental properties recovery from agnostic templates in relativistic Accretion disks, (*in Preparation.*).

Since both of these work are near completion they are cited here but are not discussed further in the text.

Conventions

Throughout this work the collaborative term *we* is used, due to the personal preferences of the author. We assume $G = c = 1$ except if outwardly stated.

Contents

1	Gravitational Waves	3
1.1	Linearized Theory of Gravity	3
1.2	Realistic Waveforms: Inspiral Merger Ringdown	6
1.3	Gravitational Waves Detectors	8
1.4	LISA	10
1.5	LISA Sources	13
2	Gravitational Waves Data Analysis	19
2.1	Bayesian Inference	19
2.2	Gravitational Wave Systematics	23
2.3	Environmental Effects on Binary Mergers	26
3	Stochastic Torques in EMRIs	29
3.1	Disk driven Migration	29
3.2	Injected signals and templates	31
3.3	Results	35
4	Consistency Test for Extreme Mass Ratio Inspirals	45
4.1	Analysis Setup	45
4.2	Inconsistent inference across observation duration.	46
4.3	Further Validation Tests	50
5	Massive Black Hole Binaries, Environment vs Modified gravity	55
5.1	Modifications at -4 PN order	55
5.2	Population models and single-event analysis	57
5.3	Population-level inference	63
6	Conclusions and Future Prospects	71
6.1	Conclusion	71
6.2	Future Perspective	73
	Bibliography	77

1 Gravitational Waves

This chapter provides an introduction to the physical framework of gravitational waves (GW). In Section 1.1, we describe a toy model for the formation, deformation, and propagation of gravitational waves in simple astrophysical sources. In Section 1.2, we provide an overview of the inspiral-merger-ringdown framework that governs gravitational wave signal evolution. We sketch the different approximants and techniques used to generate waveforms and construct templates for data analysis. In Section 1.3, we describe the experimental methodology employed in gravitational wave detection. This includes the physics underlying gravitational wave detection through laser interferometry, the characteristic response imprinted by gravitational wave signals, and the design principles of gravitational wave interferometers. We conclude with a brief overview of current and upcoming experiments.

1.1 Linearized Theory of Gravity

1.1.1 Gravitational Waves propagation in Vacuum

The interplay between gravity and matter in general relativity is well described by the Einstein equation [106]:

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

where $R_{\mu\nu}$ is the Ricci tensor and $R = R_{\mu\nu}g^{\mu\nu}$ is the Ricci scalar, which describe the structure of the spacetime manifold. The right-hand side contains the energy-momentum tensor $T_{\mu\nu}$, which encodes the presence and motion of matter and energy. The underlying theory is highly nonlinear, as evident from the definition of the Ricci tensor in terms of second derivatives of the metric. While the full theory is extremely complex, in most astrophysical applications, linearized approximations are sufficient to describe the observed physical phenomena.

We expand Equation (1.1) around vacuum as $g_{\mu\nu} \rightarrow \eta_{\mu\nu} + h_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric and $h_{\mu\nu}$ encodes the perturbation to the spacetime. Discarding terms of order $\mathcal{O}(h^2)$, the left-hand side takes the following form:

$$\frac{1}{2}(\partial_\lambda \partial_\nu h_\mu^\lambda + \partial^\lambda \partial_\mu h_{\nu\lambda} - \square h_{\mu\nu} - \partial_\mu \partial_\nu h_\lambda^\lambda - \eta_{\mu\nu} \partial_\lambda \partial_\rho h^{\lambda\rho} + \eta_{\mu\nu} \square h_\lambda^\lambda) \quad (1.2)$$

General relativity is a diffeomorphism invariant theory, meaning it is invariant under coordinate transformations. At linear order, the effect of a coordinate transformation $x_\mu \rightarrow x_\mu + \xi_\mu(x)$ is given by $h_{\mu\nu} \rightarrow h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu$. It is common to exploit this freedom and work in a specific gauge. In this specific case the theory is greatly

simplified by the choice of a Lorenz gauge with respect to the trace reversed metric tensor $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^\lambda_\lambda$:

$$\partial^\mu \bar{h}_{\mu\nu} = 0 \quad (1.3)$$

Inserting it in 1.2 the only remaining term on the left hand side is the d'alambertian of the trace reversed metric perturbation tensor:

$$-\frac{1}{2}\square \bar{h}_{\mu\nu} = 8\pi T_{\mu\nu} \quad (1.4)$$

Moreover applying the gauge transformation to the gauge condition itself (1.3) we get a residual condition $\square \xi_\mu = 0$. Before directly solving the full PDE in the presence of matter we solve it in vacuum to gain an understanding on how these gauge conditions act on the propagating waves.

1.1.2 Wave propagation

Setting $T_{\mu\nu} = 0$ and imposing the real nature of the metric perturbation equation (1.4) is solved by:

$$\bar{h}_{\mu\nu} = A_{\mu\nu}e^{ik_\alpha x^\alpha} + \text{c.c} \quad \xi_\mu = B_\mu e^{ik_\alpha x^\alpha} + \text{c.c} \quad (1.5)$$

where k_α is the wave propagation while $A_{\mu\nu}$ is the wave polarization tensor, both of which are constant. Applying the gauge condition (1.3) to the solution we have that the polarization tensor is orthogonal to the wave propagation direction so that, assuming for simplicity that the wave propagates only along the z direction, we have $A_{0\mu} = -A_{3\mu}$. Through a gauge transformation we can then set the A_{0i} components to 0 by choosing a suitable gauge transformation such that:

$$A_{0i} \rightarrow A_{0i} - ik_0 B_i - ik_i B_0 = 0 \quad (1.6)$$

which, given the transverse condition results in $A_{0\mu} = -A_{3\mu} = 0$. This fix does not exhaust the residual gauge freedom and we are able to also eliminate the trace component of the polarization tensor:

$$A^\lambda_\lambda \rightarrow A^\lambda_\lambda - 2ik^\lambda B_\lambda = 0 \quad (1.7)$$

Through the use of these gauge conditions, we have moved from the full ten degrees of freedom of the symmetric metric tensor to transverse traceless polarization with 2 degrees of freedom. The choice of transverse-traceless (TT) polarization, also referred as TT gauge, is not merely a gauge choice; following the argument by Eddington [105] and replicated by Hughes and Flanagan [113], one can show that in a fully gauge-invariant formulation, that the only propagating degrees of freedom are indeed the transverse-traceless ones.

1.1.3 Quadrupole Formula

We now consider the interplay of matter in gravitational wave production and we solve Equation (1.1) using the retarded green function for the d'Alambertian operator. As

discussed before, the only propagating degrees of freedom are the tensor transverse-traceless ones. While we obtained this result in vacuum, a similar result can be derived for the Helmholtz decomposition when matter is included [113]. As such we can restrict ourself to the TT gauge by introducing the projection operator for the spatial component $P_{ij} = \delta_{ij} - \frac{\partial_i \partial_j}{\nabla^2}$, so that equation solution is given by:

$$h_{ij}^{\text{TT}} = 4 \int d^3 x' \left(P_i^k P_j^l - \frac{1}{2} P_{ij} P^{kl} \right) \frac{T_{kl}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \quad (1.8)$$

For astrophysical systems, we can approximate $|\mathbf{x} - \mathbf{x}'| \simeq D_L$, where D_L is the luminosity distance between the source and the observer. Applying this far-field approximation to the projector we can move it outside the integral as, we obtain $P_{ij} = \delta_{ij} - \frac{x_i x_j}{D_L^2}$. Using energy-momentum conservation in flat space-time $\partial_\mu T^{\mu\nu} = 0$, we express the spatial components of the energy-momentum tensor as:

$$T_{ij} = \frac{1}{2} \left(\partial_t^2 (T_{00} x_i x_j) - \partial_k \partial_l (T^{kl} x_i x_j) + 2 \partial^k (T_{ik} x_j + T_{jk} x_i) \right). \quad (1.9)$$

Substituting this into Equation (1.8), all divergence terms can be expressed as surface integrals which, assuming $T_{\mu\nu} = 0$ outside the compact object, leaves:

$$h_{ij}^{\text{TT}} = \frac{2}{r} \left(P_i^k P_j^l - \frac{1}{2} P_{ij} P^{kl} \right) \partial_t^2 \int d^3 x' T_{00}(t - D_L, \mathbf{x}') x'_k x'_l \quad (1.10)$$

as we see the gravitational wave signal is proportional to the second time derivative of the mass quadrupole.

We consider now a simple toy model to illustrate the typical gravitational wave emission from astrophysical sources. Consider two point masses m_1 and m_2 in a quasi-circular Keplerian orbit at separation r . Assuming the binary rotation axis coincides with the z direction and inserting the energy momentum for the point particle system in (1.10), we see that the metric perturbation is given by:

$$h_{ij} = \frac{4}{D_L} M^{2/3} \mu \omega_s^{2/3} \begin{pmatrix} \left(\frac{1+\cos^2 \iota}{2} \right) \cos 2\omega_s(t - D_L) & \cos \iota \sin 2\omega_s(t - D_L) & 0 \\ \cos \iota \sin 2\omega_s(t - D_L) & -\left(\frac{1+\cos^2 \iota}{2} \right) \cos 2\omega_s(t - D_L) & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (1.11)$$

where $M = m_1 + m_2$ and $\mu = \frac{m_1 m_2}{M}$ are the total and reduced mass, $\omega_s = \sqrt{\frac{M}{r^3}}$ is the orbital frequency, and ι is the angle between the binary orbital plane and the observer's line of sight.

The production of gravitational waves causes energy loss from the binary system, inserting the wave tensor into the Landau-Lifshitz pseudo-tensor [161] we can compute the energy flux as [204]:

$$\frac{dE}{dt} = -\frac{32}{5} \frac{M^2 \mu^2 \omega_s^2}{r^2} \quad (1.12)$$

As a result, the binary's orbital separation decreases and the two point masses spiral inward, gradually drawing closer together.

1.2 Realistic Waveforms: Inspiral Merger Ringdown

The toy model described above is not well suited to provide a full description of astrophysical systems, and it is only used to give insight into the gravitational wave mechanism. First of all, a series of assumptions were made in the derivation, such as the conservation of energy-momentum using the Minkowski metric, low velocities, and neglecting the gravitational potential of the particles, which are not well representative of a realistic situation. While the formula above is still correct for a low velocity system, as it can be computed in a more rigorous way taking into consideration the self-gravitation of the binary [113] giving the same result, in most instances where the gravitational radiation is detectable the particle velocities are well relativistic.

In waveform modelling two different approximation strategies are used, the PN expansion and the adiabatic expansion using the self force formalism. In this formalism, the action of point particle and the Einstein equations are expanded in powers of the parameter $\frac{v^2}{c^2}$, see [48] for a review, where v is the particle speed. Including higher PN terms in the action reveals features of binary dynamics completely absent from the Newtonian quadrupole formula, such as spin-orbit interactions [153].

The computation of the full radiation reaction contribution at different PN orders is complex and computationally expensive. The current state of the art is the 4.5 PN for the radiation reaction [49, 50]. For high mass ratio systems, where the PN expansion

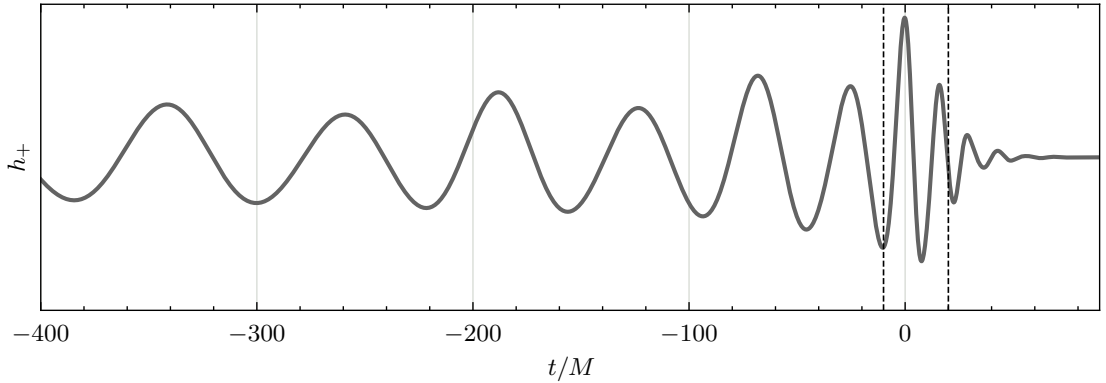


Figure 1.1: Inspiral-Merger-Ringdown gravitational wave emission. The transitions between inspiral and merger, and between merger and ringdown, are marked by dotted lines

breaks down, the post-adiabatic expansion is more appropriate [97, 137]. This uses the framework of perturbation theory and the self-force mechanism as an expansion in the mass ratio q . The accuracy decreases for comparable-mass systems, but at second-order self-force, the inspiral phase accurately replicates numerical relativity results for mass ratios of $q = 0.1$ [159].

The PN expansion becomes increasingly unreliable as binary separation decreases and higher-PN terms grow in importance. For high mass ratio systems, the transition from inspiral to plunge can be studied [43, 44] through post-adiabatic expansions. The merger

stage has not been modelled analytically, but there efforts in the direction of highly asymmetric binaries [116]. At this stage, the only reliable method is to solve the Einstein equations directly through numerical relativity. Given the extreme computational cost, it is impractical to generate solutions dynamically during data analysis. Only a limited number of simulations can be performed, providing a restricted set of templates [225].

Post-merger, the system enters the ringdown phase, where the remnant black hole radiates away residual perturbations through gravitational waves, asymptotically approaching equilibrium. The perturbed spacetime is described by a Schrödinger-like PDE [216, 251], whose solutions at infinity are superpositions of damped sinusoids, known as quasi-normal modes. By the no-hair theorem, quasi-normal mode frequencies depend only on the Kerr black hole parameters: mass and spin¹. Thus, ringdown analysis and the measurement of deviations from the expected results, also referred as black hole spectroscopy, represents one of the most promising stages for testing general relativity [46].

Waveform Approximants For data analysis to successfully extract information from data the whole evolution of a binary inspiral merger and ringdown need to be computed many times in an efficient way. In this context various approximants have been developed. *IMRPhenom* waveforms are phenomenological approximants based on PN expansions with calibration terms obtained from numerical relativity fits. The current most accurate models are *IMRPhenomX* [128, 212], which generates waveforms in the frequency domain with higher modes and eccentricity effects, and *IMRPhenomT* [107], which generates waveforms in the time domain. Both of these approximants include description of the spin precession and for the latter model an eccentric extensions also exists [208].

The *effective one-body (EOB)* approach is highly successful and accurate. In this method, the binary is mapped to an effective system of a test particle with reduced mass μ orbiting a larger black hole of total mass M [56]. The dynamics are governed by an effective Hamiltonian incorporating modifications to spacetime curvature, while radiation reaction is added as a non-conservative term. In comparison to the Phenom models the equation of motion resulting from the hamiltonian for the binary inspiral must be solved on the fly. Two template families exist: *TEOB* [13, 192] and *SEOB* [117, 209], both of which can probe eccentricity and spin precession. Both the Phenom and EOB models are calibrated against numerical relativity simulations to increase their accuracy in the inspiral phase and in the merger-ringdown phase.

Fully numerical relativity-based templates have been implemented as *NRSur* (Numerical Relativity Surrogates) [265]. These consist of interpolated waveforms spanning the full inspiral-merger-ringdown evolution from complete numerical simulations. Given computational cost, the regions of parameter space in which these surrogates are valid is limited by the number of available orbits and cannot probe early inspiral stages. Despite these limitations, for LIGO data analysis of equal-mass binaries, high-mass stellar-mass black hole systems, and comparable-mass mergers, these surrogates represent the most accurate models currently available.

¹Electric charge is expected to be negligible for astrophysical black holes

For strongly asymmetric sources these waveforms models which are generally based on PN expansions are not well equipped. While it has been shown that the EOB model can be well equipped for intermediate mass ratio inspirals [14] the most efficient approximant for EMRIs is `FastEMRIWaveforms` [69, 72, 147, 237]. This approximant can generate equatorial eccentric inspirals with relativistic waveforms and generic orbits with the Augmented Analytic Kludge method [73]. Extensions with spherical precessing inspirals and highly eccentric inspirals are being tested for the fully relativistic models with generic inspirals in development.

1.3 Gravitational Waves Detectors

1.3.1 Gravitational Wave Strain and Detection Principle

We consider spacetime perturbations in the transverse-traceless (TT) gauge and compute the geodesic deviation induced by the perturbation. Using Eq. (1.11), the typical perturbation scale created by astrophysical compact objects is of order $\mathcal{O}(10^{-20} - 10^{-16})$. In the weak-field limit, expressing coordinates as $x + \xi$, where ξ encodes the gravitational perturbation, we have:

$$\frac{d^2 \xi_i}{dt^2} = -R_{i0j0} \xi^j \simeq \frac{1}{2} \ddot{h}_{ij}^{\text{TT}} x^j, \quad (1.13)$$

where we neglect terms $\mathcal{O}(x^2/\lambda_{\text{GW}}^2)$, effectively assuming that the wavelength of the gravitational wave perturbation is much longer than the detector. It is then easy to solve equation (1.13) and observe that the relative change in length produced by the passage of gravitational waves is proportional to the gravitational strain itself.

The first technical proposal to directly measure gravitational waves experimentally was made by Joseph Weber in the 1960s using solid resonant-mass detectors [269, 271]. Equation (1.13) can be interpreted as a force term; when applied to a rigid bar, it induces oscillatory motion resonating at frequencies $\frac{n\pi v_s}{L}$, where v_s is the sound speed in the bar material and L is the bar size.

Measuring the small displacements induced by gravitational waves presents significant technical challenges, particularly since realistic gravitational perturbations are comparable to thermal noise. Newer resonant-bar detectors were built in the 1980s [180], operating in the kHz frequency range to detect gravitational wave bursts from supernova explosions. Despite initial detection claims [270], these experiments did not confirm gravitational wave detections.

A different class of experiments, which comprises most of the gravitational wave experiments currently operating on Earth, relies on the use of interferometers. In a typical Michelson interferometer a laser beam is split between two arms of length L_1, L_2 , and each beam is reflected back. The change in length induced by the passage of a gravitational wave is measured through the change in phase of the lasers interference. In the long-wavelength approximation $\lambda_{\text{GW}} \gg L$, the phase shift induced by the passage of the gravitational wave h_{ij} can be readily computed as:

$$\Delta\phi = 4\pi\nu D_{ij} h_{ij} \quad (1.14)$$

where ν is the laser frequency, while the term D_{ij} encodes the detector geometry. As the frequency of the gravitational wave increases the long-wavelength approximation does not hold. To relax this approximation, we can compute the actual time needed for the light beam to cross a given arm, using the full machinery of general relativity. Assuming, for simplicity, that the gravitational wave is a monochromatic wave $h = A \cos(\omega_{\text{GW}}t - \phi_0)$ the change in armlength is given by:

$$\frac{\delta L}{L} = \frac{1}{2} h \text{sinc}(\omega_{\text{GW}}L) \quad (1.15)$$

the additional sinc term is also known as the detector transfer function. In the long wavelength approximation this term is equal to 1 so that the phase shift in the detector coincides with Eq. (1.14), while for $\omega_{\text{GW}} \gg \frac{1}{L}$ the sensitivity decreases as ω_{GW}^{-1} . Like bar detectors, the central challenge is the extremely small effect size: to achieve useful sensitivity for gravitational waves in the Hz frequency range requires interferometer arms of length $\mathcal{O}(100 \text{ km})$. At this scale, achieving the required displacement sensitivity demands extremely sophisticated engineering to suppress seismic noise, thermal noise, magnetic noise, and to maintain laser and mirror free-fall.

1.3.2 GW Experiments

Currently active gravitational wave experiments include the LIGO-Virgo-KAGRA (LVK) collaboration and various pulsar timing array (PTA) collaborations.

The LVK collaboration is a joint effort of NSF, the European Space Agency, and Japanese space agencies, operating four gravitational wave interferometers. Two LIGO interferometers operate in the United States: one in Livingston, Louisiana, and one in Hanford, Washington [1]. The Virgo interferometer [10], is located in Cascina, Italy. The KAGRA interferometer, operated by Japanese agencies [12], is located in Japan.

All employ advanced Michelson interferometer designs with arm lengths of 3-4 kilometers in equal-arm configurations. These lengths are below the theoretical optimum for detecting astrophysical gravitational waves, since constructing a 100 km-scale detector on Earth is impractical. Instead, Fabry-Pérot cavities replace the arms, effectively increasing the arm length by a factor of $\mathcal{O}(100)$ [183].

Ground-based detectors lose sensitivity below approximately 1 Hz [223] due to seismic noise and at high frequencies due to shot noise [114]. Given current size constraints, the LVK detectors operate within the frequency range of 10-20 Hz to 4000 Hz with a sampling rate of 16384 samples per second in the last runs [4]. Within this band, primary sources are stellar-mass black holes [7], binary neutron stars [8], and neutron star-black hole binaries [9]. Throughout the first four observing runs, the LVK collaboration has detected approximately 390 events [6], with increasing detection rates reflecting improvements in detector sensitivity and hardware upgrades.

Pulsar timing arrays provide a complementary gravitational wave detection method in the low-frequency regime. This approach detects gravitational waves by measuring timing variations in radio pulses from known pulsars. Observations of millisecond pulsars enable detection of stochastic quadrupole radiation [126, 133]. Several international

collaborations operate timing arrays with different telescope networks: the European Pulsar Timing Array (EPTA) [110], the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) [139], the Indian Pulsar Timing Array (InPTA) joined in the international pulsar timing array collaboration [134]. From analysis by EPTA and NANOGrav, a first detection of the gravitational wave background has been claimed [11, 19], along with constraints on ultralight dark matter [233] and the supermassive black hole population [20].

Updates and improvement to the current detectors have been proposed and discussed. On the ground based side future detectors include the LIGO India, which will join the LVK collaboration [164] and the Einstein Telescope [5] and Cosmic Explorer [217] which are planned to be built in Europe and US respectively. The latter two will improve the sensitivity to GW in the LVK band by one order of magnitude, allowing for the detection of thousands of binary mergers [55]. To probe the millihertz frequency range space based detectors have been proposed: LISA which has been adopted by the European Space Agency, TianQuin [172] and Taiji [136]. A detailed description of the LISA mission is given in section 1.4

1.4 LISA

The detection of low-frequency gravitational waves detection is to overcome a significant technical challenge on Earth. Seismic noise and the order of magnitude increase in detector arm length make the detection of gravitational waves below 1 Hz almost unfeasible. The idea to move the interferometer to space as a way to bypass this issue has emerged in the 1980s. This concept which has become the LISA mission. Throughout the years the proposal has been changed and improved and, in 2024 it was accepted as an ESA mission with scheduled launch date in 2035.

The LISA experimental setup consists of three free-flying satellites arranged in a triangular configuration with arm length of 2.5×10^6 km, orbiting the Sun in a heliocentric orbit; see Fig. 1.2 for a graphical representation. The orbital dynamics around the Sun have been carefully studied to maintain equilibrium and preserve the triangular configuration throughout the mission duration. The scientific requirements for the LISA mission require the satellite to be active and measure data for a period of at least 4 yr, with a possible extended total lifetime of 10 yr, beyond which the ability of the satellite to maintain the triangular shape is not guaranteed.

For the good success of the LISA mission the laser and detectors must be maintained in free-fall for years. This capability was demonstrated by the LISA Pathfinder [182] mission, launched in 2015 with the specific objective of verifying the ability of the LISA spacecraft to maintain test masses in free-fall. This experiment was an incredible success, with results exceeding the LISA scientific requirements [23].

An important consideration for the LISA mission is that, due to high costs and technical engineering requirements, only one LISA interferometer will be deployed. While proposals for additional gravitational wave detectors such as TianQin and Taiji have been made. This does not significantly undermine the scientific returns of the LISA mission.

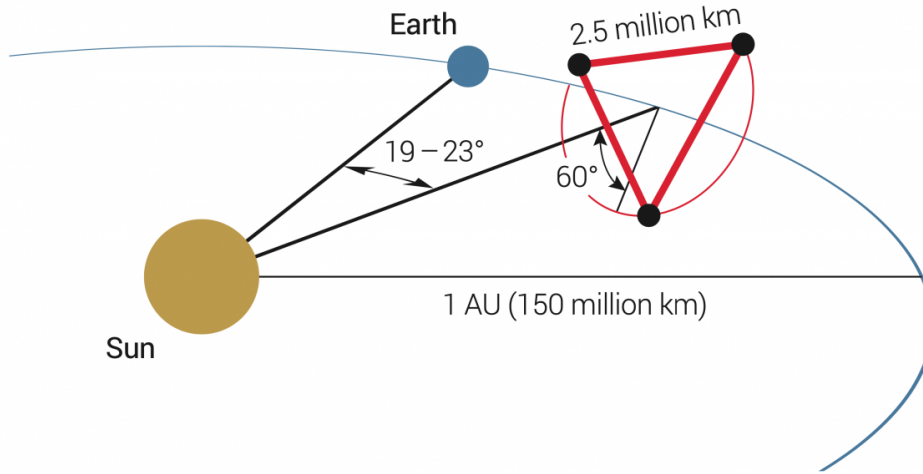


Figure 1.2: Schematic picture of the LISA orbital configuration (not to scale). The barycentre of the triangular satellite structure follows a heliocentric orbit. The constellation plane is inclined by 60° with respect to the ecliptic plane. Courtesy of the LISA Mission website.

For ground-based interferometers, multiple detectors are needed to break sky position degeneracy; however, gravitational wave signals in the mHz band persist for extended durations—days to months to years—allowing the spacecraft’s orbital motion alone to break the degeneracy.²⁰

1.4.1 Time Delay Interferometry

The three satellites orbiting the Sun in a triangular constellation are not connected in a rigid structure and relative changes in arm lengths up to $\mathcal{O}(10^{-2})$ are expected [16, 75]. In a Michelson interferometer the noise component due to laser frequency fluctuations is proportional to $f(L_1 - L_2)$, and it cancels out for an equal-arm-length interferometer. For the laser used in the LISA mission the amplitude of this noise component can completely overwhelm the astrophysical signals being order of magnitude larger than the scientific requirement of $10^{-20}/\sqrt{\text{Hz}}$. This makes the use of a standard Michelson interferometer unfeasible; however, a different experimental setup focused on canceling the noise fluctuations was proposed [254]: the time delay interferometry. In this setup rather than making the lasers reflected by mirrors 1 and 2 interfere with each other, the reflected signals directly interfere with the photodetector; specifically the ray emitted at the instant t which travels through an arm length of size L gets reflected by the mirror and interferes with a ray emitted at $t + 2L$. In this way each arm length produces a different observable y_i in which the fluctuation noise are present as $C(t) - C(t + 2L)$. As the noise fluctuations are common between the two detectors using the combination of

the single-link data streams:

$$X = y_1(t) - y_2(t) - y_1(t - 2L_2) + y_2(t - 2L_1) \quad (1.16)$$

one can get rid of laser fluctuations.

When translating this method to the LISA experiment two important considerations are necessary. The first obvious issue is that, given the LISA arm length of 2.5×10^6 km it is impossible for a laser to maintain coherence and bounce back to the original emitter. This issue is solved by neglecting the mirror altogether and having a laser emitted at link i towards link j interferes with the laser emitted at link j ; thus we have 6 single-link observables $\eta_{12}, \eta_{13}, \eta_{21}, \eta_{23}, \eta_{31}, \eta_{32}$ which can then be combined into TDI channels. The choice of TDI combination is not unique; we refer to [262] for a more accurate description, and different combination can be used to probe different effects. The natural choice is given by the Sagnac combination $(\alpha, \beta, \gamma, \zeta)$, which are four non-independent combinations [255]; this combination though can be affected significantly by laser failures. To ensure the operability of LISA during glitches and error in one of the links the unequal-Michelson combination, also known as XYZ combination is used [255].

The specific framework used to obtain combination were developed in the early 2000s, increasingly tackling more features of the LISA constellation. The first examples were TDI 1.0 (or first-generation TDI), which considered the motion of LISA with unequal arm lengths and TDI 1.5 which also took account of the Sagnac effect. As noted in [81, 228] these methods were not enough to fully cancel out the frequency fluctuations as they ignored the time dependence on the LISA arm length, and the resulting laser fluctuation would still be well above the signals generated by gravitational waves, thus requiring the development of second-generation TDI (or TDI 2.0).

Since the channels X, Y, Z are computed by combining the different single-link responses η_{ij} , the resulting noise will be correlated between the channels. In the approximation of equal arm lengths and equal noise; more detail in [214, 243], it is possible to rotate the covariance matrix and obtain an orthogonal channel combination A, E and T :

$$A_n = \frac{1}{\sqrt{2}}(Z_n - X_n) \quad E_n = \frac{1}{\sqrt{6}}(X_n - 2Y_n + Z_n) \quad T_n = \frac{1}{\sqrt{3}}(X_n + Y_n + Z_n) \quad (1.17)$$

where n indicates the TDI generation. The T channel is often neglected in analysis as it is a null channel in which the gravitational wave signal is suppressed with respect to noise at low frequencies [178], in which regime LISA can be approximated as a set of 2 independent Michelson interferometers. It has been discussed whether this channel, while containing little information for most astrophysical sources, may be used as a direct probe of the constellation noise; however, recent work [191] has shown that a different TDI combination ξ is better suited. It is important to notice that the orthogonality requires no arm-flexing; as such when LISA small cross-channel correlations are still present. As such, while the A, E and T channels are still being used to assess the scientific objectives of LISA through figures of merit and case studies, in the global fit, either the full X, Y, Z Michelson combination or the Sagnac combination $\alpha, \beta, \gamma, \zeta$ will be used.

1.5 LISA Sources

LISA will be sensitive to gravitational waves in the mHz band and, over its lifetime, LISA will observe an extremely diverse population of sources. Below is a summary of the main source classes expected in the LISA band; we highlight key features relevant to their detection and analysis.

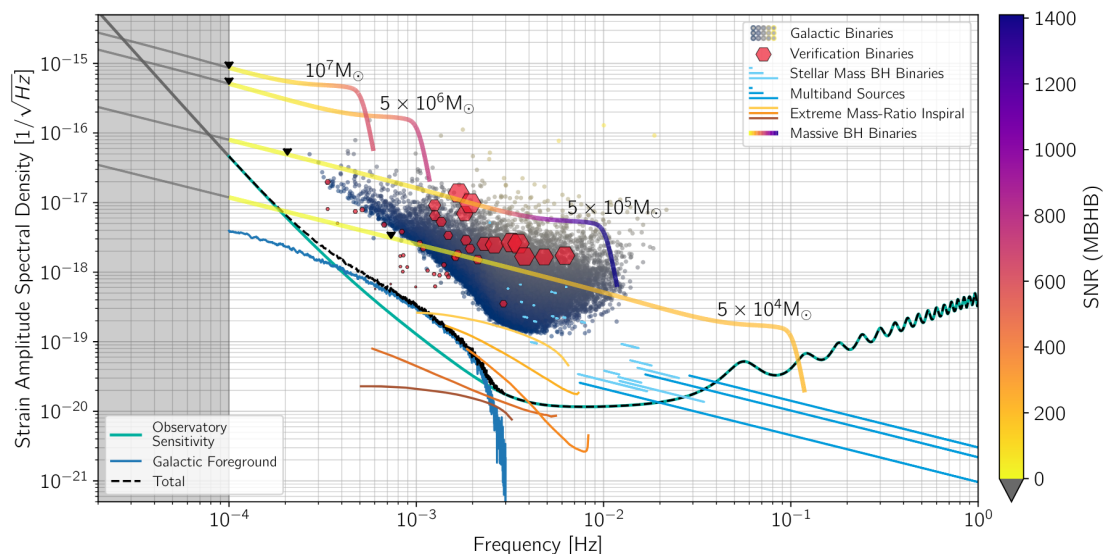


Figure 1.3: Figure taken from the LISA Redbook [75], courtesy of the LISA consortium). LISA sensitivity curve, compared against gravitational waves signals in the frequency domain. Solvable galactic binaries are indicated by dots, with the verification binaries found through EM observations [240], while the combined contribution due to the unresolved binaries is reported. For MBHBs the dominant track and the accumulated SNR is reported for four different representative masses. For EMRIs and SOBH binaries the various tracks are reported.

1.5.1 Galactic Binaries

The most abundant sources in the LISA band are galactic binaries (GBs). This subset includes binary systems composed of stellar-mass black holes, neutron stars, and white dwarfs, as well as ultra-compact binaries with orbital periods ranging from hours to minutes. Crucially, the frequency evolution of these sources will not change significantly during the LISA observation period and most are expected to be quasi-monochromatic; for sources with frequency $f > 3$ mHz also the first time derivative of the frequency can be probed [218]. In these regimes, the dominant effect inducing changes in the binary separation can be both the radiation reaction or the interaction with the astrophysical environment; as such the emission frequencies and its first and second derivative will be treated as independent parameters in the analysis.

Of the estimated 10^7 galactic binaries emitting in the mHz band [160, 200] only a small fraction accumulates enough SNR throughout the LISA mission and can thus be resolved. The remaining sources produce a noise foreground [142]; since the sources are localized on the galactic plane the amplitude of the foreground depends on the LISA position in the solar system reference frame [59], thus creating a cyclostationary noise component.

From the study of galactic binaries several information about the shape of the galaxy [53] and the formation rate of white dwarfs and neutron stars can be probed. Critically this would require a population analysis combining the information from the resolvable sources with the background [91, 238, 259].

1.5.2 Massive Black Hole Binaries

The second major source class in the LISA data stream is the merger of comparable-mass black hole binaries (MBHBs). These systems are expected to merge with masses ranging from 10^2 – $10^8 M_\odot$, with the observable portion of the signal varying significantly across this mass range. For higher-mass sources with $M > 10^4 M_\odot$, the full gravitational waveform can be observed, including both the inspiral and merger-ringdown phases. Even in this subset the morphology of the signals can drastically change, as some events will spend only hours to days in the data stream, while others can spend months or even years in band. Conversely, for lower-mass systems with total mass below $10^4 M_\odot$, the merger occurs outside the LISA band and thus only the inspiral phase can be observed.

Different theoretical models for the formation of supermassive black holes have been proposed in the literature. The predictions for event rates and mass distributions depend critically on key assumptions about the initial conditions and growth mechanisms. The primary factors governing supermassive black hole formation are the seed mass at high redshift and the accretion growth mechanisms by which these objects acquire mass throughout cosmic history. By varying these fundamental assumptions, widely different model predictions can be obtained.

For light seed scenarios, such as Population III seed models, the expected number of sources can reach hundreds per year, with typical masses generally below $10^6 M_\odot$ [156]; conversely, heavy seed models predict higher masses and total event rates that can easily exceed 1000 per year. Additionally the spin configuration of supermassive black holes depends strongly on the accretion mechanism through which they form. In mergers occurring in gas-rich environments, such as those involving accretion disks or circumbinary disks, the spin of the black hole aligns with the orbital angular momentum due to the Bardeen effect [36]. In contrast, within gas-poor environments, such as dynamical capture mergers, the spin orientation is expected to be random. Since LISA will be able to measure spin precession [118, 213], we can expect to distinguish between different formation channels for supermassive black holes and potentially disentangle formation mechanisms for subsets of mergers.

1.5.3 Stellar-Origin Binary Black Holes

By observing the mHz gravitational wave band it is possible to probe the early inspiral stages of stellar-mass black hole binaries (SOBHs), which are typically observed at merger by ground-based interferometers in the LVK collaboration [155, 227]. These sources live in the higher frequency range of the LISA band, between 10^{-2} – 10^{-1} Hz and can remain in the band for timescales spanning from half a year up to the entire mission duration. Since these sources are typically extremely faint [190] and we will likely only detect a fraction of the total population observed by LIGO. Multi-band observations allow the study of gravitational wave emission across different evolutionary stages: the early inspiral phase in the LISA band and the merger and ringdown in future ground-based detectors such as the Einstein Telescope or Cosmic Explorer. This allows the direct measurements of kick velocity [215], probing of line-of-sight acceleration, and environmental effects [224, 257]. Moreover a multiband observations allows for precise measurement of the spin precession [Gerosa:2013laaFarr:2017uvj] and eccentricity at the early stages of the inspiral [54] which are offer a strong insight into the formation channel mechanism of stellar mass binaries.

1.5.4 Extreme Mass Ratio Inspirals

Strongly asymmetric binary systems are generally considered separately from MBHBs; in the LISA band strongly asymmetric configurations can be created as systems composed of a supermassive black hole and a stellar-mass black hole orbiting the primary. Given the extension of the possible ranges of the primary these events are classified in the literature as extreme mass ratio inspirals (EMRIs) for $q < 10^{-4}$ and intermediate mass ratio inspirals (IMRIs) for $10^{-4} < q < 0.1$. Additional work has also proposed more asymmetric events (XMRI), with $q < 10^{-6}$ composed of sub-solar-mass objects orbiting the center of the Milky Way [15]. These sources are expected to spend years in the LISA band and to complete $\mathcal{O}(10^5)$ cycles while in the strong-gravity regime, allowing the resolution of the binary's property with exquisite precision. As such from an extreme mass ratio inspiral we are able to resolve the properties of the source with exquisite precision and these sources are the prime candidates for tests of GR.

Similarly, to the MBHB case the event rate of extreme mass-ratio inspirals remains largely uncertain. Although estimates based on population synthesis models of supermassive black holes have been proposed [25], the predicted detection rate spans a wide range, from tens of events per year to no detectable events over the entire mission lifetime depending on the population model. Two main formation channels have been considered in the literature. The first is the loss-cone (or dry) formation channel, in which compact objects are scattered onto highly eccentric orbits around the central supermassive black hole; this scenario generally predicts an isotropic distribution of orbital inclinations. Initial estimates of the eccentricity for EMRIs in the loss-cone mechanism predicted a uniform distribution up to $e = 0.2$ and a tail for higher eccentricities; recent work though [176] has shown that extreme eccentricity, up to $e = 0.99$ while entering the LISA band, are possible. The second main mechanism is the disk-capture (or wet) formation

channel, in which the secondary object is captured by the accretion disk surrounding the supermassive black hole [95, 203]. As a consequence of the dissipative interaction with the disk, these systems are expected to be on nearly equatorial, prograde orbits with very low eccentricities, possibly approaching quasi-circular configurations. Recent studies suggest that the importance of the wet formation channel may have been substantially underestimated in previous event-rate estimates, which considered only the dry channel [202].

1.5.5 Stochastic Backgrounds

Additionally LISA will be able to detect gravitational wave backgrounds, which are stochastic in nature, generated uniformly in the universe. Different possible background sources have been identified either astrophysical in nature, such as the unresolvable galactic binaries, stellar-origin black hole binaries [24, 162] and EMRIs [206, 211] or cosmological such as phase transitions, primordial gravitational waves, cosmic strings.

1.5.6 LISA Global Fit Analysis

Based on the populations model of massive black holes we can expect multiple binaries, both equal mass (MBHBs) or asymmetric (IMRIs/EMRIs) to be present in the data stream at the same time. Additionally long-lived signals such as galactic binaries, backgrounds and a fraction of the stellar-origin black hole binaries will be present throughout the LISA mission. To avoid biases arising from the incorrect subtraction of individual sources all the signals must be modeled and sampled consistently; as such the dimensionality of the Bayesian problem grows exponentially with respect to current ground-based detectors. Although different classes of sources are generally expected to be weakly correlated, this assumption does not necessarily hold for sources belonging to the same class. For example, two MBHBs merging at the same time, or several galactic binaries emitting within the same frequency band, can exhibit significant correlations. To address this high-dimensional inference problem, the method proposed for LISA data analysis is the global fit approach [80].

Exploiting the approximate independence between different source classes, each class is analyzed separately, with the inferred signals successively subtracted from the data stream in an iterative, cyclic procedure, as illustrated in Fig. 1.4

It should also be noted that, unlike ground-based detectors, LISA never observes a period containing pure instrumental noise; consequently, there is no dedicated noise calibration phase prior to the analysis. Instead, the noise model must be inferred jointly with the astrophysical signals as part of the global fit and, at each iteration of the analysis, the noise parameters are re-estimated and sampled simultaneously with the source parameters.

The development of the LISA global fit is an ongoing project with various official pipelines in development: the ETH pipeline [241] **Erebor** at the Albert Einstein Institute [148], **Gee-moo** [92], at CNRS and **Glass** at NASA [169]. To gradually prepare for the LISA launch in the 2030s data challenges with increasing complexity are being

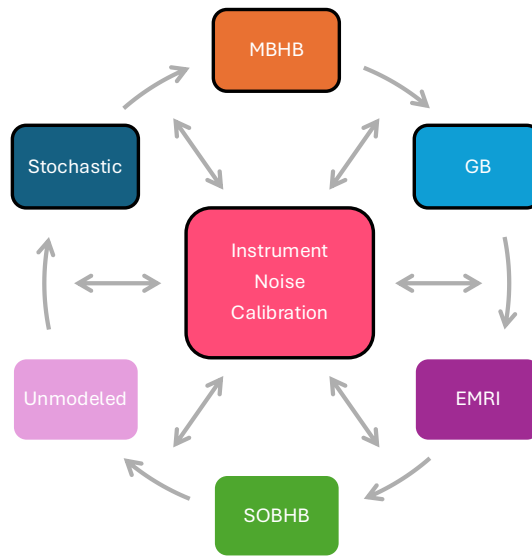


Figure 1.4: A schematic representation of the global fit setup. Each block, represented by a circle, corresponds to a different block of sources in the LISA data stream, and the arrows indicate the movement from sampling one specific block to another. At each step, the noise properties of the system — the noise properties of LISA — must be re-sampled, as indicated by the arrow pointing toward the central blob, representing the noise properties of the system.

tackled by the different pipelines to ensure that the scientific objectives of LISA will be met. The current release is *Mojito* which contains all the expected LISA source classes, non-rigid arms for the detector and, in its heavy version, data gaps. Future challenges *Long Island Ice Tea* and *Aunt Roberta* will increase the problem complexity, represented by the cocktail's alcohol content, by adding noise defects, gaps and increasing the number of sources bringing them closer to what may realistically be present in LISA.

2 Gravitational Waves Data Analysis

In this section we review the main topics in gravitational waves data analysis. In section 2.1 we describe the basics of Bayesian Inference and the observables used in Gravitational Wave Astronomy. In section 2.2 we describe some results in waveform systematics and in section 2.3 we give a review on Environmental Effects for gravitational waves.

2.1 Bayesian Inference

In most branches of astrophysics, given the complexity of the data, high dimensionality of the models and impossibility to replicate an observation, the framework used is the Bayesian one. The generic objective for parameter estimation is to compute the probability that, under a specific hypothesis $\mathcal{H}$, a set of parameters θ describes the observed data, this probability $p(\theta|D, \mathcal{H})$ is typically referred as posterior. To estimate the posterior distribution, one makes use of Bayes theorem:

$$p(\theta|D, \mathcal{H}) = \frac{p(D|\theta, \mathcal{H})p(\theta|\mathcal{H})}{p(D|\mathcal{H})} \quad (2.1)$$

On the right-hand side, the term $p(D|\theta, \mathcal{H})$ is the likelihood, which measures how well the model, for fixed parameters θ , replicates the observed data. The term $p(\theta|\mathcal{H})$ is the prior distribution, which encodes information that we have based on the model itself regardless of the observed data. The choice of a correct prior is a subtle point of the utmost importance, as a different prior can bring significant differences in the inferred properties and conclusion of an analysis. Lastly the probability that the data itself is represented by the hypothesis $p(D|\mathcal{H})$, can be computed by marginalizing over the parameters θ :

$$p(D|\mathcal{H}) = \int d\theta p(D|\theta, \mathcal{H})p(\theta|\mathcal{H}). \quad (2.2)$$

As we are ignorant of the specific mechanism generating the data, we may have multiple different competing hypotheses which more or less describe the phenomena, and we need a metric for selecting a preferred model. An important aspect of model selection is assessing how well a model reproduces the observed data while avoiding unnecessary complexity as one may incur an overfitting problem. A classic example is fitting N data points with a polynomial of degree $N - 1$. Such a model will always provide an excellent fit to the available data, but it may capture features that are merely statistical fluctuations rather than genuine properties of the underlying physical system. For this reason, model selection must take into account the principle of Occam's razor, namely,

given two models that explain the data with comparable accuracy, the one with fewer free parameters should be preferred.

In the Bayesian framework, model comparison is typically performed using either the posterior odds ratio or the Bayes factor. Since Bayesian inference assigns probabilities to hypotheses, models can only be compared relative to one another. Specifically, one compares the probability of the observed data under hypothesis A versus the probability of the same data under hypothesis B . This ratio is given by obtain the ratio of the probability of two different hypotheses as:

$$O_{12} = \frac{p(\mathcal{H}_1|D)}{p(\mathcal{H}_2|D)} = \frac{p(D|\mathcal{H}_1)p(\mathcal{H}_1)}{p(D|\mathcal{H}_2)p(\mathcal{H}_2)} = \frac{p(\mathcal{H}_1)}{p(\mathcal{H}_2)}\mathcal{B}_{12} \quad (2.3)$$

where O_{12} is known as the odds ratio, while the ratio of the evidences $\mathcal{B}_{12}$ is known as the Bayes factor while $p(\mathcal{H}_i)$ is the prior probability of the hypothesis itself. An odds ratio above 1 indicates that $\mathcal{H}_1$ is the preferred one, while the opposite is true for $0 < O_{12} < 1$. Assigning a probability to the model is non-trivial so in most cases the different hypotheses are considered equally likely, and, rather than using the odds-ratio the comparison is done directly through the Bayes factor. An important feature of this quantity is that the Occam penalty is naturally incorporated into the evidence integral. Adding additional parameters increases the volume of the parameter space, and unless these parameters are supported by the data, the evidence is diluted, automatically penalizing more complex models. However, this argument hides an important subtlety: the evidence depends on the choice of the prior distribution. If the prior volume is chosen to be excessively large compared with the region occupied by the posterior distribution, the evidence can be artificially suppressed, making a model appear less favorable than it actually is. Conversely, an overly restrictive prior may artificially increase the evidence; careful and physically motivated prior choices are essential when performing Bayesian model comparison.

A practical indication of the preference is given by Jeffreys' scale [138]:

- $0 < \log_{10} O_{12} < \frac{1}{2}$: Both hypotheses are effectively valid
- $\frac{1}{2} < \log_{10} O_{12} < 1$: Substantial preference for $\mathcal{H}_1$
- $1 < \log_{10} O_{12} < \frac{3}{2}$: Strong preference for $\mathcal{H}_1$
- $\frac{3}{2} < \log_{10} O_{12} < 2$: Very strong preference for $\mathcal{H}_1$
- $\log_{10} O_{12} > 2$: Decisive preference for $\mathcal{H}_1$

and vice-versa for a negative odds-ratio it translates to the preference for the hypothesis $\mathcal{H}_2$.

2.1.1 Gravitational wave Observables

In this subsection we offer a quick description of the metrics for event detection commonly used in gravitational wave experiments; while the focus of this work is on LISA these concepts can be promptly applied to current and future ground-based detectors.

The presence of noise in any experiment is inevitable, specifically, in gravitational wave experiments, the amplitude of the noise in a datastream is expected, for realistic astrophysical signals, to be comparable to, or significantly bigger than, the amplitude of the gravitational wave signal. Nevertheless, it is still possible to obtain information on the gravitational wave if we have an adequate template of the gravitational wave signal, through the matched filtering procedure. We refer to [57, 174] for a more detailed description. For this procedure, we define an inner product that given a data stream d containing a signal h and a noise component n , based on the Weiner filter [84, 261] which maximizes the signal to noise ratio. For a signal in time domain the inner product is defined as:

$$\langle a|b \rangle = \sum_{ij} \iint dt dt' a_i(t) C_{ij}^{-1}(t, t') b_j(t') \quad (2.4)$$

where the indices i, j indicate the detector channels and $C_{ij}(t, t')$ is the noise covariance matrix estimated as the ensemble average of the instrument noise. When the noise is stationary the matrix C takes a Toeplitz structure, so that $C(t, t') = \text{ACF}(|t - t'|)$ where ACF is the autocorrelation function of the noise. Moving to frequency domain the fourier transform of the covariance matrix is diagonal, so that in the Fourier domain the inner product (2.4) can be expressed as:

$$\langle a|b \rangle = 4 \sum_{ij} \text{Re} \int_0^{+\infty} df \frac{\tilde{a}_i^*(f) \tilde{b}_j(f)}{S_{ij}(f)} \quad (2.5)$$

where S_{ij} is the one-sided noise power spectral density (PSD). As discussed before this approximation does not hold for the LISA mission, as the noise will present non-stationary features due to the unresolved galactic binary foreground; still the approximation can be used to assess its ability to fulfill the scientific objectives. In this work we never considered the non-stationary contribution of the noise; as such when referring to the inner product throughout this work we specifically refer to Eq. (2.5).

The optimal SNR is defined as the square root of the noise-weighted inner product of a template with itself:

$$\rho_{\text{opt}} = \sqrt{\langle h|h \rangle} \quad (2.6)$$

Realistically we never have access to a clean signal from the detector, and to infer the signal which is actually extracted from the datastream we use the detected SNR:

$$\rho_{\text{det}} = \frac{\langle d|h \rangle}{\sqrt{\langle h|h \rangle}}, \quad (2.7)$$

This is also known as the optimal (matched-filter) statistic; it should be noted that this metric is invariant with respect to the distance used in the template, indeed this parameter is generally used in the search phase, as maximizing over the template parameters yields the best-fitting template for a given event. In the analysis of coalescing binaries the common choice is the Gaussian likelihood:

$$p(d|\theta) = \exp \left(-\frac{1}{2} \langle d - h | d - h \rangle \right) \sqrt{\det 2\pi S_{ij}} \quad (2.8)$$

also known as the Whittle likelihood; sometimes in the literature the normalizing term $\sqrt{\det 2\pi S_{ij}}$ is often neglected, as they refer to analysis in which the noise is known and thus the term is constant throughout the parameter space. Additional and more flexible likelihood models have been proposed, such as heavy-tailed likelihoods [144, 222] as a way to handle the presence of glitches and other unexpected features in the data.

It is then easy to see that, since the likelihood is peaked around the true value, the Bayes factor between pure noise hypothesis and transient signal hypothesis is approximately $\propto e^{-\rho_{\text{det}}^2/2}$. Indeed, the ρ_{det} is often used as the optimal detection statistic, and is used as an indicator of the significance of a given signal for transient events. Depending on the kind of source and the experiment, the threshold used to claim a detection can change based on the specific experiment's false alarm rate. In LIGO, typically for transient sources, the detection threshold is $\rho = 8$, a similar threshold is used for MBHBs in space based interferometers. For other long-lived sources, the choice of threshold differs: for example, for galactic binaries the usual choice is typically between 5 and 7 [75], while for EMRIs the typical threshold is around 20 [25].

A further set of quantities, not commonly employed in the analysis of real data but widely used in waveform modeling and in the study of systematics, are the match and the mismatch. The match, also known as overlap, measures the similarity between two templates, normalized by their respective amplitudes, and is given by:

$$\mathcal{O}(h_1, h_2) = \frac{\langle h_1 | h_2 \rangle}{\sqrt{\langle h_1 | h_1 \rangle \langle h_2 | h_2 \rangle}}. \quad (2.9)$$

We can promptly notice that this is the cosine of the angle between the two signals¹. By construction it lies between -1 and $+1$: a match of 1 indicates that the two waveforms are identical (same template at different distances), a match of -1 indicates they are perfectly anti-phased, and a match of 0 indicates they are orthogonal. Since the interest in systematics is to minimize the signal left in the datastream rather than the match it is often reported the mismatch:

$$\mathcal{M} = 1 - \mathcal{O} \quad (2.10)$$

which is often minimized over the phase and the time of merger, since as these parameters correspond to translations in time and phase and do not include any information on the underlying physics. The mismatch is itself a quasi-indicator of how well a template represents a given data stream. This can be shown by taking Eq. (2.8) and, assuming $\langle h_1 | h_1 \rangle \sim \langle h_2 | h_2 \rangle$ evaluating it in the noise-free limit to have that $\log p(\theta | D) \simeq -\rho_{\text{det}}^2 \mathcal{M}$.

2.1.2 Methods

Computing the values of the likelihood and posterior distributions directly becomes impractical for many physical problems and the only way to infer the posteriors is through the use of stochastic sampling methods. The standard approach is Markov Chain Monte Carlo (MCMC), in which a chain is initialized and propagates through

¹Both signals here are considered noiseless

parameter space based on posterior values. A typical implementation is the Metropolis-Hastings algorithm [129] in which a proposed jump is accepted or rejected based on the ratio of posterior values at the two locations. MCMC is a powerful method, but it encounters difficulties with multimodal posteriors which are common in gravitational wave astronomy, to address these issues, parallel tempering has been developed for MCMC. In this approach, multiple chains run at different temperatures, effectively sampling smoothed versions of the posterior at higher temperatures. Swaps between chains are allowed, enabling efficient exploration of complex posterior landscapes.

Computing the Bayesian evidence from MCMC samples for model comparison is challenging and subject to high computational uncertainty. Various augmentations and post-processing techniques have been proposed, such as thermodynamic integration [61], which uses the temperature ladder to sample the full posterior shape, reversible jump MCMC (RJMCMC), which allows the sampling chain to jump between different models. Lastly, machine learning methods such as flow-matching can be used to directly compute the evidence from posterior samples [239].

A completely different method fully dedicated to evidence computation is nested sampling [231], in which random points are generated in parameter space to compute the volume enclosed by hypersurfaces of fixed likelihood values. This allows to compute a Lebesgue integral which gives as a result the evidence with the posterior samples produced as a byproduct. Nested sampling is extremely effective and robust against multimodal posteriors and has been used extensively in gravitational wave astronomy with LIGO, however, it can become computationally expensive and is known to scale poorly with higher dimensionality.

In recent years, significant attention has been devoted to likelihood-free inference methods, such as simulation-based inference (SBI). In this approach, a neural network is trained on examples of noise and template realizations to learn to compute posteriors directly from data. Different implementations of amortized methods have been proposed for the LVK collaboration [90, 219, 234], both of which require extensive training time but are able to produce posterior samples in seconds, being extremely effective for long-lived signals such as binary neutron stars [89]. Alternatively, a network can be trained to solve a single-event, similar to standard MCMC and nested sampling methods, through sequential and dynamic SBI. In these approaches, the network is trained to match a specific event, zooming in on the posterior structure. Compared to traditional methods, sequential SBI can require significantly fewer waveform computations [47] to produce a posterior.

2.2 Gravitational Wave Systematics

The quality and success of parameter inference is strongly correlated with the quality of the templates used in the analysis. The choice of inaccurate templates, either due to some missing physics or mischaracterization of the source properties can result in biased estimation.

It should be noted that a small presence of systematics when inferring gravitational

wave signals is inevitable as the commonly used gravitational wave models are set up in perturbative expansions whose truncation order is selected based on the expected detector sensitivity. Additionally any analysis is limited by the quality of the numerical implementation such as interpolation errors in the computation of the binaries trajectory [150]. In this context even numerical relativity surrogates which are in principle including the full GR picture are susceptible to systematics.

The effect of missing features can be quantified, for small template deviation, by Taylor expanding the likelihood function around the injection point. Then the expected shift in the maximum likelihood point is given by the Cutler-Vallisneri [86] formula:

$$\delta\theta_i = \Gamma_{ij}^{-1} \langle d - h(\theta) | \partial_j h(\theta) \rangle \quad (2.11)$$

where Γ_{ij} is the Fisher matrix:

$$\Gamma_{ij} = \langle \partial_i h(\theta) | \partial_j h(\theta) \rangle \quad (2.12)$$

computed at the injection point. The Cutler-Vallisneri formula itself is not a perfect indicator as, similar to the Fisher matrix, it is a linear approximation of the full underlying Bayesian posterior and it breaks down for dephasing above one radian. A different metric used to assess the quality of the template against a fiducial waveform model is the mismatch (2.10). Close to the maximum likelihood point the logarithm of a gaussian likelihood follows a χ_D^2 distribution with D being the number of free parameters in the template. It is then easy to map the mismatch to the confidence interval with which the two waveforms are in agreement. A standard criterion used in the literature which derives from this simple computation is the Lindblom criterion [167]:

$$\mathcal{M} < \frac{D}{2\rho^2} \quad (2.13)$$

which is obtained using the approximation $\mathcal{M} \simeq \frac{\log p(\theta|D)}{\rho_{\text{det}}^2}$ and requiring that, for a given signal the true values lies well within the posterior. The agreement between two templates strongly depends on the SNR; for fainter signals systematical effects can be significantly less important than the statistical error while for louder signals increased precision is needed.

2.2.1 Systematics from Waveform Modeling in LISA

Here we give a quick review on the status of waveform systematics for pure GR binaries; throughout the years various of these works have assessed the relevance of neglecting different effects in the analysis. Various of these work have become superseded as through their contribution waveform models have become more accurate. We mainly focus here on works which are more pertinent for the development of waveform models for the next generation of detectors.

Most works have been focused on comparable-mass black hole binaries, as these kind of events are the main observable in both present and future detectors. Recent works have

highlighted the importance of the use of higher modes in space-based interferometers. An extensive analysis has been brought up for the mode component in the ringdown analysis [63], estimating the minimum number of ringdown modes to obtain an unbiased estimation of the mode amplitudes throughout the LISA parameter space. A similar work has studied the higher-mode contribution for full inspiral merger ringdown waveforms. In the current LIGO observations, where the main contribution of higher modes is the breaking of distance-inclination degeneracy [168], while in LISA, given the higher SNR for massive black hole binaries, the neglect of higher modes can result in strong bias in both intrinsic and extrinsic parameters, still the current mode content in IMR and EOB waveform models may be insufficient for unbiased estimations [275, 276].

Additional bias may arise in the waveform calibration stage. Waveform models are calibrated against a set of numerical relativity simulations to improve their accuracy. This calibration though is limited by the small number of available waveform due to the high computational cost. Hence the calibration coefficients may induce systematical error for waveforms which do not coincide or are far enough from the numerical relativity simulations. The use of Gaussian regression models to compensate for the statistical errors in the calibration stage have been proposed [188, 189, 210]. It has also been proposed that error in these calibration may result in false claim of modified GR detection; this has been tested for ground-based detectors in which case though the necessary SNR to include biases is generally above what expected for LIGO and ET [185].

In asymmetric binaries the study of systematics is an extremely active field, especially since the release of the first fully relativistic fast waveform templates with FEW [68, 146]. Various avenue of systematic errors are present in these waveform models, neglecting the contribution to the radiation reaction fluxes at high l can induce biases for highly spinning primaries [150], moreover the current accuracy in modeling 0 post-adiabatic or first-order self-force, is not accurate enough for intermediate mass ratio inspirals [39, 58]. Specifically at adiabatic order the spin of the secondary does not contribute to the inspiral dynamics and it enters at higher post-adiabatic order. The importance of the spin contribution has been shown to be negligible for low mass ratios and equatorial orbits [58], while for spherical orbits the mismatch between a template including the spin contribution and one can reach values of 0.8 for polar orbits [232].

2.2.2 Systematics from Unmodelled physics

Other sources of systematics come from unexpected features of the system that are generally not accounted for, such as the presence of modified gravity or environmental effects. Such features can result in biases in parameter estimation.

Deviations from general relativity in modified gravity theories have been extensively studied in the literature. Different works have highlighted possible contributions at different evolutionary stages of binary mergers.

In the inspiral stage, a classification method has been introduced [244, 278] as a way to parameterize deviations from general relativity within a single model-agnostic framework, considering only the dominant correction. This assumption is valid as long as secondary corrections are negligible with respect to the leading-order terms. Studies

including higher-order corrections show improved results. This method, known as the parametrized post-Einsteinian (ppE) formalism, maps a given deviation effect to a specific post-Newtonian correction applied to either the amplitude or the phase of the gravitational wave signal in the inspiral stage:

$$h_{\text{ppE}}(f) = h_{\text{GR}}(f)(1 + \alpha u^a)e^{i\beta u^b} \quad (2.14)$$

where $u = (\pi\mathcal{M}_c f)^{1/3}$ and h_{GR} is the pure GR waveform, while the coefficients a, b encode the ppE order through which can be mapped to specific theories, while α, β encode the magnitude of the effect acting on the amplitude and the phase respectively. These two parameters can be used to recover the specific properties of the studied modification. Since gravitational wave signals are primarily sensitive to variations in the phase rather than variations in the amplitude, the amplitude modification is often neglected in analysis and only the phase component is considered. Recent work has also proposed the use of neural networks to quickly identify both ppE and unmodeled deviations [273, 274].

It should be noted that any test using the ppE formalism is subject to the quality of the general relativity baseline waveform. The use of incorrect general relativity waveform assumptions—for example, neglecting spin precession [66], or eccentricity [194, 229] can result in false claims of modified gravity.

A similar approach, frequently used in LIGO analysis, is the *est* Infrastructure for General Relativity (TIGER) [163] and the Flexible Theory-Independent (FTI) [184] framework, in which deviations of the coefficients describing the radiation reaction and phase evolution at each post-Newtonian order in the gravitational waveform approximant are measured. While the only applications with real data of these tests have been done for the LVK collaboration [2, 3] the ability of LISA [207] and ET [45] to assess deviations through these tests has been studied.

2.3 Environmental Effects on Binary Mergers

The discussion above, even in the context of modified gravity, only includes modification to the inspiral trajectory of the binary system due to radiation reaction from gravity. Any possible additional interaction with other astrophysical sources, has been ignored up to now. Still, in a realistic physical situation, we may expect the binary system undergoing merger not to be isolated and taking into consideration the effect of the interaction of the binary system with its astrophysical environment may be necessary. Typically, the importance of this effect becomes less relevant close to merger, since the system is in the strong gravity regime, and the dominant effect is expected to be gravitational radiation. In general, for most systems in band, we tend to expect that the effect of interaction with the environment is subdominant with respect to the radiation reaction. This assumption may not always hold [157]. Nonetheless, it is a good initial assumption to provide the description of the environmental effects. A good catalog, comparing and giving a back-of-the-envelope computation on the importance of various classes of effect, is provided in [31]. Here we intend to give a quick rundown of what we expect will be

the most important environmental effects in future experiments, focusing specifically on the latest developments.

In a gaseous environment particles are expected to fall into the black hole through accretion. This can be easily modeled as a time-dependent mass term inside the waveform, as computed by Caputo, entering as a -4PN effect, in which the contribution to the dephasing at leading order is extremely subdominant with respect to other effects. This is typically because the accretion timescale is given by the Salpeter time, and we can take the LISA observation time to be much shorter than the accretion timescale. In this regime, the effect of accretion is typically very faint: the waveform deviation it produces is typically small, as the change in mass needs to be significant for its effect to be measurable. The change in mass induced by accretion needs to be comparable to or larger than the statistical uncertainty on the total mass of the system. The measurability of accretion has been studied both in the LISA band and the LVK band, for stellar-mass black holes, and in the LISA band, for supermassive black holes. In the former case, there is the possibility of unveiling the modification to the inspiral due to accretion, thanks to multiband observations, while in the latter case, only the ability to infer the accretion rate is still open to discussion. From current observations on the LVK collaborations the constraints obtained on the time-dependent Newton constants [280] can be promptly mapped to constraints on the accretion rate for these events are well above what is realistic astrophysical values.

The motion of the black hole in the gaseous environment induces a gravitational wake which, interacting back with the black hole itself result in a force term known as dynamical friction [67]. The initial computation, carried out by Chandrasekhar, was used in the context of Newtonian calculations, and additional work has since improved the accuracy, going beyond the Newtonian limit up to post-Newtonian order and relativistic order [34, 242] for circular orbits [154, 201] and eccentric orbits with mutual wakes [108]. For dynamical friction to take place, the black hole must not completely disrupt the gas structure, in this context, the typical binaries affected by dynamical friction are not SMBHs, which would disrupt the gas environment, but rather SOBHBs binaries embedded in AGN disks.

In the LISA band the imprint left by gas is strong enough to be detected for Stellar origin black holes merging in a migration trap [257], with the precision on the measurement of the environment properties greatly improving with a multi-band observation. While only a small fraction of SOBHBs binaries will be detectable by LISA the effect of dynamical friction can still be observed through the background produced by non resolvable binaries [71]. These analyses have been also done for ground based detectors with current experiments requiring gas densities $\rho \geq 10 \text{ g/cm}^3$ for current detectors and $\rho \geq 10^{-3} \text{ g/cm}^3$ for future ground based detectors such as ET and CE [62]. When increasing the mass of the binaries and moving toward the massive black hole binary case, many of the computations performed here can still be used, at least as an order-of-magnitude estimate. The physical configuration of the system, however, is extremely different, since the binary is not enveloped in a gaseous environment, but is instead surrounded by a circumbinary disk, whose structure is strongly influenced by the motion of the binary itself. As such,

order-of-magnitude similarity computations have been carried out [120, 281]. All of these assumptions need to be validated against full hydrodynamical simulations, whose computational cost is extremely high. Recently, dynamical computations that include the gravitational wave reaction are being carried out to describe the early stages of binary evolution [121, 122].

A specific class of dynamical friction interaction happens when we take into consideration the size of the disk for strongly asymmetric systems. In this case, typically known as planetary migration, the interaction gets amplified with respect to dynamical friction, since the wakes get reflected by the disk boundary [123, 247] and, differently from vanilla dynamical friction in which the dominant effect is -5.5 PN, in planetary migration it strongly depends on the disk profile. Since this effect is one of the major topics in chapter 3, it will be discussed in more detail in Section 3.1. The ability of LISA to resolve the imprint left by planetary migration has been studied in [151, 236] for Shakura-Sunayev disk, with additional work which has assessed the importance of small eccentricity in the inference of the disk properties [101]. All of these work are based on the Newtonian modelling done for planet formation, recently these results have been extended to the relativistic regime, which is more realistic for the description of binary in the strong gravity. When the relativistic regime is taken into consideration the exorted torque can be at least one order of magnitude bigger with respect to the Newtonian computation [102, 131, 132], and for louder event these model can infer the disk properties without the need for additional electromagnetic signals [109]. While the interplay between spin of the black hole and the gas has been ignored a recent work has considered the interplay between the gaseous environment and the spin of the secondary in extreme mass ratio inspiral showing that it can improve the secondary spin reconstruction [103, 171].

This same effect can be computed for generic fluids, the dynamical friction effect has been studied also for dark matter densities. While typical dark matter haloes are not expected to leave a big enough imprint, the density can greatly increase in matter spikes at the center of galaxies [60, 195]. Since the spike is centred on the supermassive black hole, the case of interest is that of extreme-mass-ratio inspirals. This effect has been studied extensively, both in the Newtonian and in the relativistic regime [74, 266], while other have studied the effect of generic boson clouds [104, 151].

Moving away from the presence of gas, an additional interaction with the astrophysical environment is given by the gravitational interaction between the binary itself and a third body. Depending on the mass configuration of the three bodies different effect can take place. For a SOBH binary affected by a massive black hole different configuration can happen if the binary is merging in a migration trap the line of sight acceleration can induce significant dephasing [224] and induce Kozai-Lidov oscillation [170, 193]. For strongly asymmetric system the presence of a far third body can induce resonances in the binary evolution [125].

3 Stochastic Torques in EMRIs

Most of the work done to assess the ability of future detectors to measure deviations from GR due to the presence of an environment has usually relied on using the same model for injection and recovery. However, we do not expect a given model to contain all the possible physical information, but only to be a simplification of it. Specifically, in realistic accretion disks, the interaction due to the secondary can easily be stochastic in nature. In this chapter, we describe the work done in testing the use of agnostic templates for recovery with realistic disk injections. This chapter is based on Ref. [77], and we use the same notation used in the original work referring to the binary masses as M_1 , M_2 , and the primary spin as a/M_1 . The chapter is organized as follows: in Section 3.1, we introduce disk-driven migration and the torque models used throughout the chapter, in Section 3.2, we describe the injected signals and the templates used for recovery, in Section 3.3, we present the results of the parameter estimation and we discuss the astrophysical inference and its implications.

3.1 Disk driven Migration

The gravitational field generated by a compact object moving in a gaseous environment produces spiral density perturbations through Lindblad and corotation resonances [123], resulting in an exchange of angular momentum between the satellite and the disk [124]. This effect produces an inward, or outward, migration of the secondary object, and has been extensively studied in the context of planetary formation [42, 198].

The case in which the density perturbations do not strongly modify the disk structure is commonly referred to in the literature as Type I migration. This regime is relevant when the mass of the satellite is much smaller than the mass of the primary surrounded by the disk. Under these assumptions, one can neglect any effect due to the radial motion of the satellite when studying the torque exchange between the disk and the secondary [41]. The torque has been modelled at linear order in 2 and 3-dimensional isothermal disks in Refs. [124, 248]. In the case of a satellite on a circular orbit migrating in an isothermal disk, the torque is given by [247]

$$\dot{L}_I = \Sigma(r)r^4\Omega^2q^2\mathcal{M}^2 \quad (3.1)$$

where Ω is the angular velocity of the satellite, $\mathcal{M}$ is the Mach number of the disk, and Σ is the unperturbed surface density profile. For realistic disks, we expect additional coefficients multiplying the linear torque, depending on the temperature and density gradient [246, 249]. In this work, we ignore those additional terms. This expression reproduces the results of numerical hydrodynamical simulations [158].

When considering the migration of more massive objects, it is possible for the torque exerted by the satellite on the disk to become larger than the viscous torque, which is responsible for disk spreading [166, pag. 749-835]. The density wave perturbations launched by the satellite are then able to carve a low density annular region, also known as gap, around the orbit [165]. In this case, the system is said to be undergoing Type II (or gap opening) migration.

It should be noted that gap formation alone does not automatically imply that the system undergoes Type II migration. It is possible for the gap to close too quickly to provide a significant contribution to the satellite dynamics. The conditions separating the two regimes depend non-trivially on the disk parameters and the mass ratio of the system [175]. In more detail, the following conditions are necessary for a system to undergo type II migration:

- the secondary perturbation pushing away the gas cannot be compensated by its own gravitational attraction and the viscosity of the disk [82];
- the gap formation time must be smaller than the time necessary for the secondary to cross its horseshoe radius [268];
- the gap formation time must be smaller than the viscous diffusion timescale [120].

While the first condition concerns the gap formation itself, the last two ensure that the gap is preserved during the migration and is not refilled by viscous diffusion.

There have been numerous numerical studies on planetary evolution in the gap formation regime [100, 197]. Various empirical approximations describing the torque exchange have also been proposed, such as the one in Ref. [141],

$$\dot{L}_{II} = \frac{\dot{L}_I}{1 + 0.04q^2\mathcal{M}^5\alpha^{-1}} \quad (3.2)$$

where α is the viscosity of the disk. These empirical formulae, however, fail to reproduce the full non-linear behavior observed in numerical simulations, particularly in the parameter space of high-viscosity, geometrically thin disks. The full non-linear contribution can indeed lead to a sign change in the net exchanged torque [94, 99]. While a fully satisfactory analytic description has not been put forward yet, a common feature observed in hydrodynamical simulations is a dampening of the mean torque exerted on the disk when compared with the Type I equation (3.1).

Moreover, in realistic astrophysical situations, stochastic features are generically expected in the evolution of proto-planetary systems and in the motion of stellar mass and intermediate mass black holes in AGN disks. Fluctuations arise both from the influence of the secondary and from the turbulent nature of the disk. Turbulence becomes even more important for asymmetric mass ratio systems undergoing Type I migration [40, 196], where it can give rise to *chaotic* migration. As the mass ratio grows and the migration enters the gap-opening regime, the chaotic turbulence contribution becomes instead less significant [22, 272].

3.2 Injected signals and templates

In this section, we describe the setup used to generate the injection, and the parameter estimation framework. First, in 3.2.1 we briefly review the simulations of Ref. [93], and describe how we implement the numerical torque in the injected waveforms. Then, in 3.2.2 we describe the template family used in the parameter estimation.

3.2.1 Injection

Our review of Type II migration suggests that, to obtain a realistic estimate of the disk torques, it is necessary to perform fully numerical hydrodynamical simulations. In this work, we use the numerical results obtained by Derdzinski et al. in [93]. In the latter, they considered a thin, isothermal Shakura-Sunyaev accretion disk, with constant Mach number $\mathcal{M} = (H/r)^{-1}$ and constant α -law prescription for the viscosity. Imposing these conditions in the equations describing the thin disk structure in the Newtonian limit [115], it follows that the surface density profile is $\Sigma \propto r^{-1/2}$.

The secondary was modeled as a point mass with a smoothed gravitational potential, which inspirals on quasicircular equatorial orbits under gravitational radiation reaction (computed using the quadrupole formula). The effect of the environmental torque exerted on the secondary was neglected in the trajectory evolution. The simulations are scale invariant with respect to both the total mass of the system and the surface density of the accretion disk. Because the simulations are two-dimensional, they require that the smoothing of the secondary potential is of order a scale height of the disk in order to properly model vertically-integrated forces. Specifically, it is set to be one half of the disk scale height, as adopted in other studies [179, 247]. For the case of a secondary of $10^3 M_\odot$, this corresponds to a scale of ~ 275 Schwarzschild radii¹, below which the gas morphology is not resolved.²

Here, we instead consider quasicircular and equatorial prograde orbits around a *spinning* primary, evolving adiabatically under the effect of GW fluxes [97, 249, 251] down to the Kerr innermost stable circular orbit (ISCO) [37]. We do not consider the inspiral-to-plunge and ringdown GW emission, as they contain negligible power for very asymmetric mass ratio binaries. We rescale the simulations, which were performed in a scale-free Newtonian setting, so that the disk's inner edge matches the Kerr ISCO. We also interpolate the simulation torque in orbital separation r rather than time. This choice is mainly dictated by the fact that we are using adiabatic GW fluxes, while in the numerical simulations the inspiral was modeled using only the GW quadrupole flux. Finally, we add the environmental torque to the GW fluxes and evolve the trajectory with the **FastEMRIWaveform** (FEW) package.

We consider three cases associated with the simulations of [93], which vary by mass ratio and disk viscosity α :

¹This corresponds to $0.8 \left(\frac{r}{10M_1} \right)$ Hill radii

²The variability of the torques measured in this study is due to the gas flow outside of this region, and persists with higher resolution tests [94]

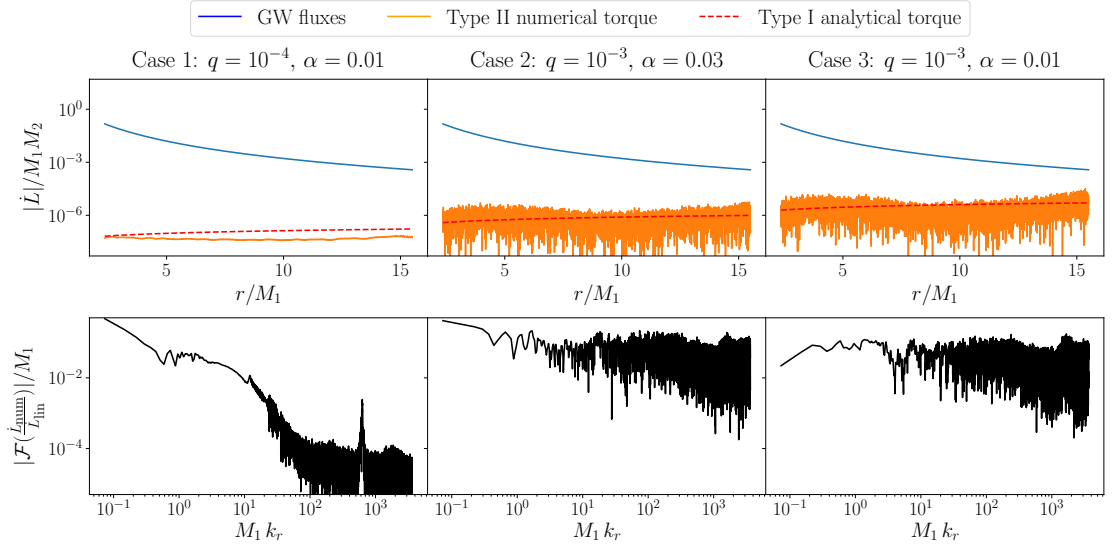


Figure 3.1: Impact of stochastic torques on the binary evolution. In the top panels we report the different components responsible for the binary evolution: the torque due to the GW fluxes (blue) and the disk torque (orange). The red dotted lines represent the analytic estimate of the linear (Type I) disk torque, Eq. (3.1). Note that, in all cases studied here, the environmental effect is subleading compared to GW emission. In the bottom panels, we plot the absolute value of the Fourier transform of the ratio between the numerical and analytic disk torques.

	Case 1	Case 2	Case 3
$M_1/M_\odot$	2.2×10^6	10^6	2×10^7
q	10^{-4}	10^{-3}	10^{-3}
a/M_1	0.9	0.9	0.9
r_0/M_1	15.10	15.487	15.487
Φ_S	1.43	2.30	0.39
θ_S	1.88	2.85	0.54
Φ_K	1.59	1.84	5.67
θ_K	3.059	0.95	0.04
Φ_0	5.45	3.59	2.69
Σ_0 (kg/m ²)	1.40×10^7	1.8×10^8	4.6×10^7
α	0.01	0.03	0.01
T_{obs} (yr)	4.0	0.2	4.0
$\mathcal{M}$	20.0	20.0	20.0

Table 3.1: Parameters for the three different systems that we consider. Σ_0 is the surface density of the disk at $r = 10M$.

- **Case 1:** $q = 10^{-4}$, $\alpha = 0.01$,
- **Case 2:** $q = 10^{-3}$, $\alpha = 0.03$,
- **Case 3:** $q = 10^{-3}$, $\alpha = 0.01$.

In all cases, the Mach number is $\mathcal{M} = 20$. The complete list of our systems' parameters is reported in Table 3.1. In all cases, we select the initial separation r_0 so that the secondary plunges within the observation time. We choose a random sky position and spin orientation, while we choose the distance to set the signal-to-noise ratio (SNR) of the signals to the values given in the next Section.

To generate the injected waveforms, we use the augmented analytic kludge (AAK) algorithm [73] implemented in FEW, together with an adiabatic trajectory modified for the purpose of this work as in [236]. The AAK algorithm is a refinement of the analytic kludge model [28], which approximates the secondary orbits as Keplerian ellipses, and calculates the waveform using the Peters-Mathews mode sum approximation [205]. The AAK algorithm improves the accuracy by using the Kerr orbital frequencies in place of their Keplerian approximations.

Following Ref. [93], we select Σ_0 such that the environmental contribution is subdominant relative to the GW fluxes. Note that this results in realistic values for the density, $\Sigma_0 \sim 10^7 - 10^8 \text{ kg/m}^2$ [115]. In Fig. 3.1, in the top panels, we compare the contributions to the torque arising from the GW fluxes and the environment. As can be seen, the analytic expression for Type I torques [Eq. (3.1)], represented by a red dashed line, cannot capture the stochastic behavior of the numerical torque (in orange), which is dominant for $q = 10^{-3}$.

Also shown in the bottom panels of Fig. 3.1, is the absolute value of the Fourier transform of the ratio between the numerical and the analytic disk torques. As can

be seen, the stochastic torques (middle and left panels) have a rather flat spectrum. The torque for case 1, which shows no stochastic features, has a peak at wavenumber $Mk_r \sim 10^3$, resulting from small oscillations barely visible in the top panel. The low frequency bump in the left panel is instead due to the torque's average slow evolution across the plotted radial range. The lower panel of Fig. 3.1 also shows that the smaller viscosity in case 3 suppresses the average torque by almost two orders of magnitude compared to case 2, see the low-frequency end, despite similarities in their stochastic fluctuations.

Note that for cases 1 and 2, we choose masses $M_1 = 2.2 \times 10^6 M_\odot$ and $M_1 = 10^6 M_\odot$ respectively, while for case 3 we consider a more massive primary $M_1 = 2 \times 10^7 M_\odot$. This latter choice allows for enhancing the stochastic torque, which depends linearly on $M_1 \Sigma_0$ [93], while keeping the injected signal marginally within the LISA band. This enhancement of the stochastic torque is not necessarily unrealistic. Indeed, Ref. [93] can only simulate disks with Mach numbers $\mathcal{M} \simeq 30$, corresponding to thicker geometries than realistic AGN disks, for which $\mathcal{M}$ can reach ~ 100 . If realistic disk aspect ratios could be simulated, they could possibly present more pronounced stochastic torque components, if the trend in the simulation data continues to the regime of highly supersonic flows. We note that the simulations discussed only consider laminar disks, finding that torque variability is due to the instabilities in the gas flow around the secondary. Realistic AGN disks are highly magnetized and turbulent, and this may lead to an additional source of torque stochasticity [40, 272, 282]. However, more research is needed to investigate this regime for the AGN IMRI case, in which there is likely an interplay between the nature of the turbulence, the characteristics of the disk and the BH mass.

3.2.2 Template and Parameter estimation

To perform parameter estimation, we build our templates using the AAK algorithm and Kerr circular equatorial orbits, supplemented with an environmental torque. For the latter, we adopt a simple parametric model inspired by the analytic Type I expression (3.1) [236], i.e. we supplement the GW flux with a power-law term given by

$$\dot{L} = \dot{L}_{\text{GW}} A \left(\frac{r}{10M_1} \right)^n. \quad (3.3)$$

Note that because the torque is normalized to the GW flux $\dot{L}_{\text{GW}} = -32/5 q(r/M_1)^{-7/2}$, the Post-Newtonian (PN) order at which environmental effects appear is given by $-n$. For consistency with the reference radius appearing in (3.3), in the following we will denote by Σ_0 the surface density at $r = 10M_1$.

To perform Bayesian parameter estimation, we inject the signal in the parallel tempering Monte Carlo Markov chain (MCMC) sampler **eryn** [143]. We use 4 temperatures and 16 walkers, and we stop the sampler only if the chain length is at least 50 times the mean autocorrelation time evaluated across chains. We sample over the intrinsic parameters ($\log M_1/M_\odot$, $\log M_2/M_\odot$, a/M_1 , $r_0/M_1, \Phi_{\phi 0}$), with a being the primary spin and $r_0, \Phi_{\phi 0}$ the initial separation and phase for ϕ ; the extrinsic parameters (D_L ,

$\cos \theta_S, \Phi_S, \cos \theta_K, \Phi_K$), with D_L being the luminosity distance, θ_S, Φ_S being the polar and azimuthal sky position in the Solar System barycenter frame, θ_K, Φ_K being the polar and azimuthal angle describing the orientation of the primary spin; and the power-law parameters (A, n) . We adopt the standard Gaussian likelihood $\log \mathcal{L} = -1/2 \sum_{X \in (\text{I}, \text{II})} \langle s_X - h_X(\lambda) | s_X - h_X(\lambda) \rangle$, with s the data in the two LISA data channels in the long-wavelength approximation [83], and $h(\lambda)$ the GW template with parameters λ . The inner product is defined as [84]:

$$\langle a | b \rangle = 4 \text{Re} \int_0^{+\infty} df \frac{\tilde{a}^*(f) \tilde{b}(f)}{S_n(f)}, \quad (3.4)$$

where $S_n(f)$ is LISA’s noise power spectral density (PSD) [250], while the tilde indicates the Fourier transform.

3.3 Results

In this section, we review the parameter estimation results for the three systems mentioned above. For all three systems, we will consider $n = 4$ as the “expected value” of the power law slope, as predicted by Eq. (3.1) for the disk model of the simulations. To estimate the “expected value” of the power-law amplitude, we use the torque averaged over the radial range of the simulations, assuming the expected slope $n = 4$. This procedure results in a smaller amplitude than would be predicted by the Type I expression (3.1), by a factor of ~ 0.28 , ~ -0.22 and ~ -0.015 for cases 1, 2 and 3, respectively. It was observed in [93] that in the last two cases the net migration is outward rather than inward, even in the last stages of the inspiral.

In Fig. 3.2 we show the residual signal left in the data after subtracting the best-fit template. In all the cases studied, the residual SNR is negligible ($\lesssim 1$). As such, it should not affect significantly the LISA global fit [92, 148, 169, 241].

3.3.1 Case 1 ($q = 10^{-4}$)

First, we consider the system with mass ratio $q = 10^{-4}$. Here, the stochastic features are subdominant relative to the mean contribution, i.e. they can be modeled as small perturbations to the mean torque, which has approximately power-law radial dependence, see Fig. 3.1. We set the primary mass to $M_1 = 2.2 \times 10^6 M_\odot$, the disk surface density to $\Sigma_0 = 1.40 \times 10^7 \text{ kg/m}^2$, so that we expect the power-law amplitude to be $A = 2 \times 10^{-5}$. Using the parameters indicated in Table 3.1 and an observation time of 4 yr, we fix the SNR for this event to 100 by setting the luminosity distance to 2.26 Gpc.

We show the posterior distribution of $(\log M_1/M_\odot, \log M_2/M_\odot, a/M_1, A, n)$ in Fig. 3.3. The posterior of the vacuum parameters does not exhibit biases at 90% confidence level. We also find that the posterior of the power-law parameters is compatible at one σ with the expected values of A and n .

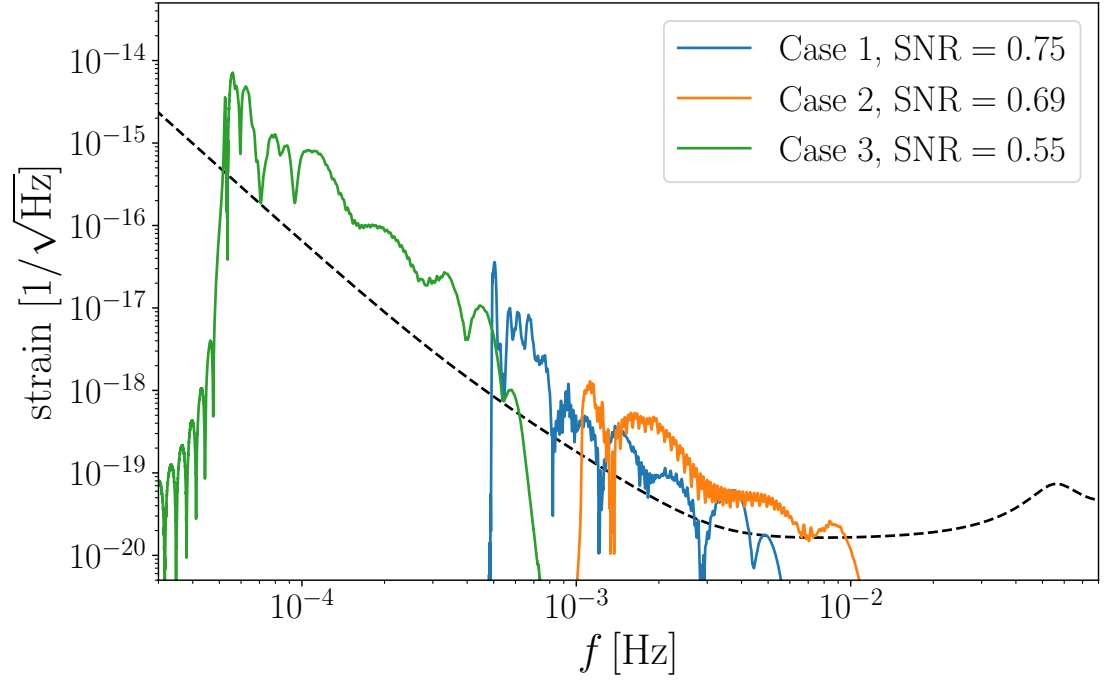


Figure 3.2: Residual signal left by subtracting the best-fit waveform obtained with analytic-torque templates from a signal produced with simulated torques. The SNR of the residuals is reported in the legend for the three systems we consider, summarised in Table 3.1.

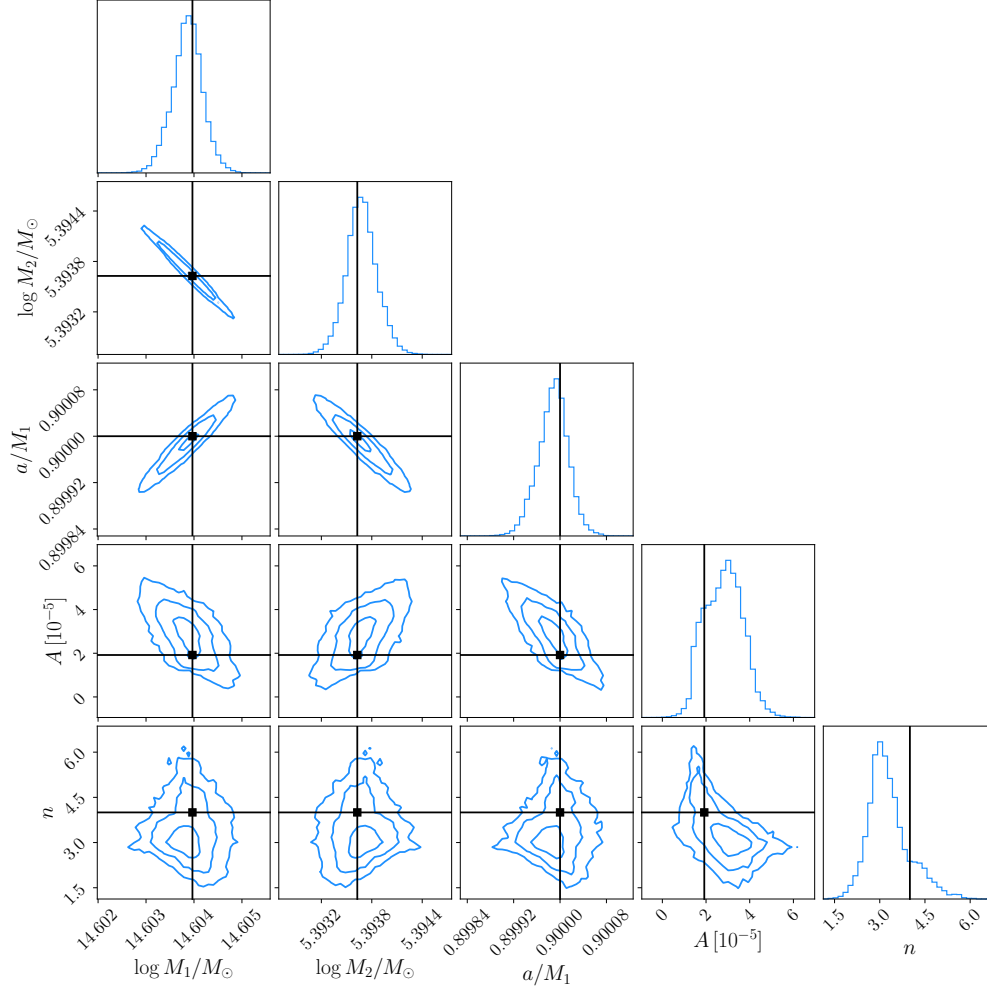


Figure 3.3: Posteriors of the intrinsic parameters for case 1 ($q = 10^{-4}$). In black we show the true values of the masses and primary spin, and the expected power law amplitude and slope as estimated from the numerical simulations used for the injection [93]. We see that using an analytic power law template, one can correctly recover the masses and the spin of the primary.

3.3.2 Cases 2 and 3 ($q = 10^{-3}$)

In the systems with mass ratio $q = 10^{-3}$, the stochastic component dominates over the mean torque, i.e. the environmental contribution is mainly stochastic. Still, we conjecture that it is possible to separate the migration torque into a mean evolution, which follows a power law, and a dominant noise term, which averages to zero.

We look first at case 2, with primary mass $10^6 M_\odot$ and the remaining intrinsic parameters reported in Table 3.1. Since we are interested in exploring the impact of stochastic effects on the parameter estimation accuracy, we consider an event of relatively large SNR=200 with an observation time of 0.2 yr, which corresponds to a luminosity distance $D_L = 1.375$ Gpc. The duration of this event is significantly shorter than that of typical EMRIs discussed in the literature. Indeed, this system has larger q and thus larger GW flux (which is proportional to q), which results in a shorter signal duration. We choose a disk surface density $\Sigma_0 = 1.8 \times 10^8$ kg/m², which leads to an expected power-law amplitude $A = -9.0 \times 10^{-5}$.

Like in case 1, the posterior of all vacuum parameters, does not exhibit biases. The parameters of the power-law environmental model are also very well compatible with the expected values, see Fig. 3.4.

As a further check, we consider a system with the same parameters, but with environmental torque exactly given by the Type I equation (3.1), rather than by the numerical simulation. For consistency, we choose the parameters to match the average torque of the numerical simulation. The posteriors obtained from the numerical and analytic injections are similar, i.e. we can estimate the parameters with a similar precision. We show this in Fig. 3.4. We therefore conclude that for a signal undergoing Type II migration, stochastic features in the orbital evolution are likely unimportant, i.e. one cannot distinguish an analytic injection from a stochastic one.

For case 3, the average of the numerical torque is two orders of magnitude smaller than expected from the analytic expression, see Fig. 3.1. Because of this strong suppression of the mean torque, the environmental effect will only be detectable by LISA for very dense disks, or for signals with large SNRs and long inspirals in band. To avoid unrealistically dense disks, we consider a very massive binary with $M_1 = 2 \times 10^7 M_\odot$, observed for 4 yr; by requiring an SNR of 100, we set the event at $D_L = 2.13$ Gpc. We choose an average amplitude of the environmental torque of $A = -2.8 \times 10^{-5}$ by setting $\Sigma_0 = 4.61 \times 10^7$ kg/m²; see Table 3.1 for the remaining parameters.

As for the other systems, no significant bias can be observed in the posterior of the vacuum parameters, shown in Fig. 3.5. Even with our choice of a very massive primary and dense disk, the best-fit torque amplitude is smaller than the error inferred by the parameter estimation. Therefore, the posteriors are compatible with vacuum (i.e. the amplitude A is compatible with zero).

Although denser disks would give larger environmental effects, one must be cautious because above a certain density threshold the torques may dominate over the gravitational fluxes in the early inspiral, thus violating one of the assumptions of the simulations of [93].

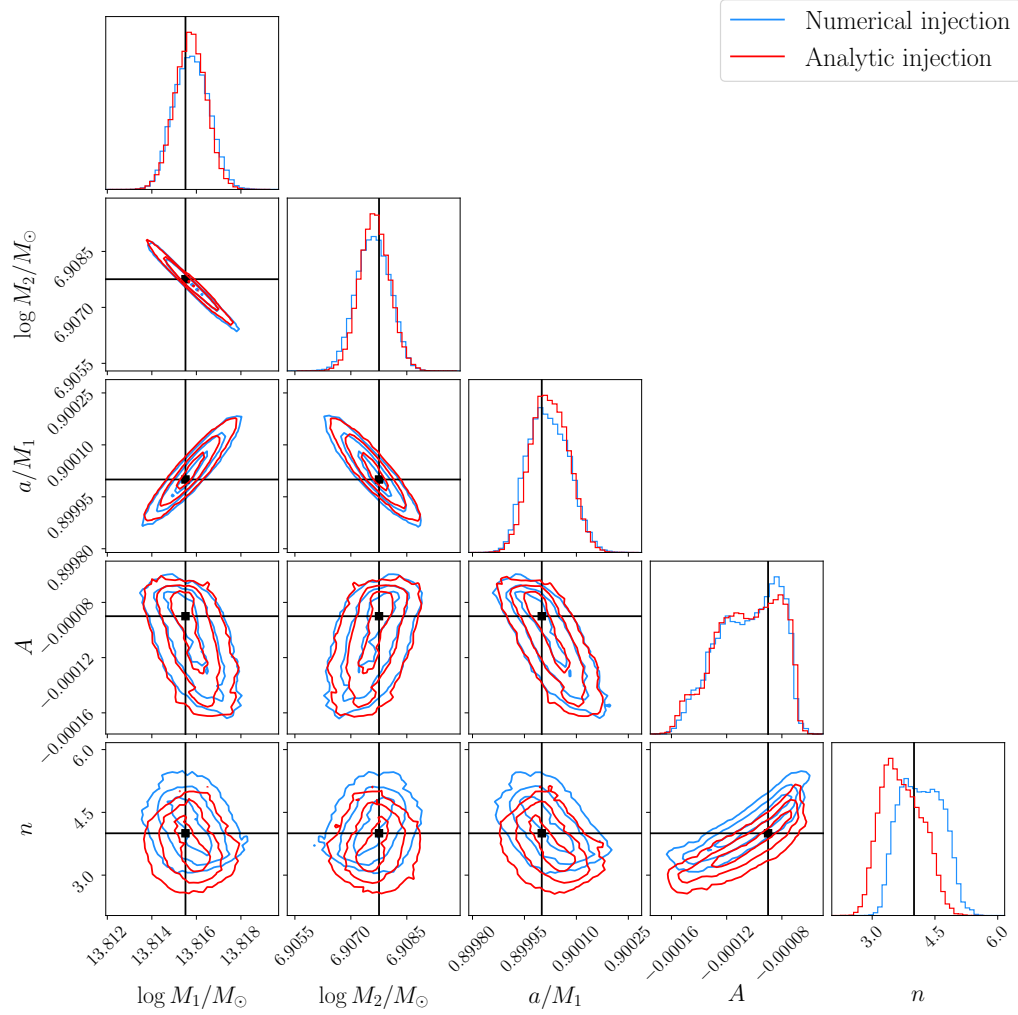


Figure 3.4: We compare the posteriors of two injections: one in which the IMRI experiences the full stochastic torque of case 2, and one in which the environmental torque is given by the power-law of Eq. (3.3), with amplitude matching the (averaged) numerical torque and slope $n = 4$. The posteriors have a similar size and are compatible at 90% confidence level.

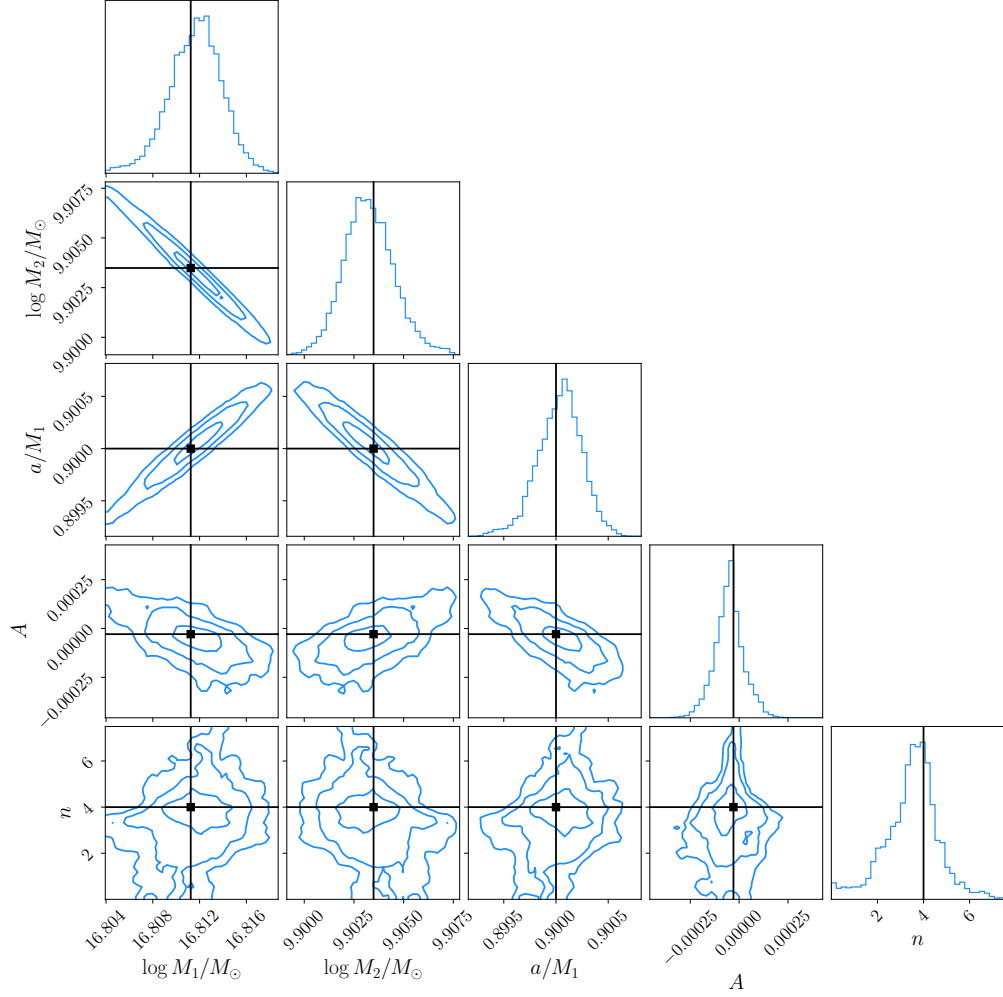


Figure 3.5: Posterior of the intrinsic parameters for the case 3 ($q = 10^{-3}$) injection. We show the injection values (black) of the masses and the primary spin, and the power law amplitude and slope estimated from the hydrodynamical simulations used in the injection [93]. Although the vacuum parameters priors are well reconstructed in this case, the stochastic features in the environmental torque wash away the power-law contribution.

3.3.3 Recovery of the Astrophysical environment properties

The parameters of our agnostic model can then be mapped to a specific disk model, to estimate its properties. This inference of the disk parameters, however, can itself be biased, as a result of two competing effects: *(i)* the biased estimation of the torques amplitude and slope; and/or *(ii)* the use of the wrong disk model in the inference.

To isolate these two contributions, in this section we will attempt to reconstruct the disk properties in two ways: *(i)* from the posteriors for A and n obtained for an injection featuring stochastic torques, but utilizing a disk model with constant Mach number (as assumed in the simulations of Ref. [93]); *(ii)* from the posteriors for A and n obtained for an injection featuring analytic torques with no stochastic component, but utilizing a disk model with constant opacity (as assumed e.g. in Ref. [115]). We stress that both disk models are solutions of the (Newtonian) hydrodynamics equations, and are therefore both viable theoretical models to use in the reconstruction of the disk's properties.

The slope of the power-law, n , is related to the disk model, as well as to the nature of the environmental effects that modify the inspiral trajectory [157]. In all the cases that we studied, this parameter does not show biases larger than 2σ , implying that the disk model could be, in principle, accurately constrained. The power-law amplitude A depends instead, at linear order, on a combination of disk viscosity α and the Eddington fraction f_{Edd} [236]. However, unless other observations (in addition to the GW ones) are available to break the degeneracy between α and f_{Edd} , extracting the local properties of the disk (e.g. its surface density) will be challenging for circular orbits. This degeneracy can be broken when eccentricity is included, cf. e.g. [101]. Alternatively, one may measure the Eddington ratio from electromagnetic observations of the host AGN, provided that the latter can be identified in the sky errorbox returned by GW observations [236]. Besides being degenerate, the relation between A and α , f_{Edd} depends sensitively on the disk model, which is subject to systematic uncertainties. This relation is especially hard to model accurately in the Type II regime, as discussed in Sec. 3.1. This can cause additional biases in the astrophysical inference.

We model the amplitude of the torque assuming a thin isothermal disk with the α viscosity prescription and assuming negligible radiative cooling. Solving the equilibrium equations for a Newtonian disk [115], we find that an opacity prescription $\kappa \propto \rho T_c^{-1}$ leads to a steady state disk with constant Mach number:

$$\Sigma = 1.5 \times 10^6 M_6^{4/15} \alpha_{0.01}^{-11/15} f_{0.1}^{3/5} r_{10}^{-1/2} \text{ kg/m}^2 \quad (3.5)$$

$$H = 4.2 \times 10^8 M_6^{13/15} f_{0.1}^{1/5} r_{10} \text{ m}, \quad (3.6)$$

where $M_6 = M_1/10^6 M_\odot$, $\alpha_{0.01} = \alpha/0.01$, $f_{0.1} = (f_{\text{Edd}}/0.1)$, $r_{10} = r/10 M_1$. For constant opacity $\kappa = 0.22 \text{ cm}^2/\text{g}$ we find instead

$$\Sigma = 1.1 \times 10^5 M_6^{1/5} \alpha_{0.01}^{-4/5} f_{0.1}^{38/5} r_{10}^{-3/5} \text{ kg/m}^2 \quad (3.7)$$

$$H = 9.6 \times 10^7 M_6^{1/10} f_{0.1}^{1/5} r_{10}^{21/20} \text{ m} \quad (3.8)$$

As already mentioned, this latter model is slightly different from the one used in the simulations from which we extracted the torque [93, 94], as the resulting Mach number is

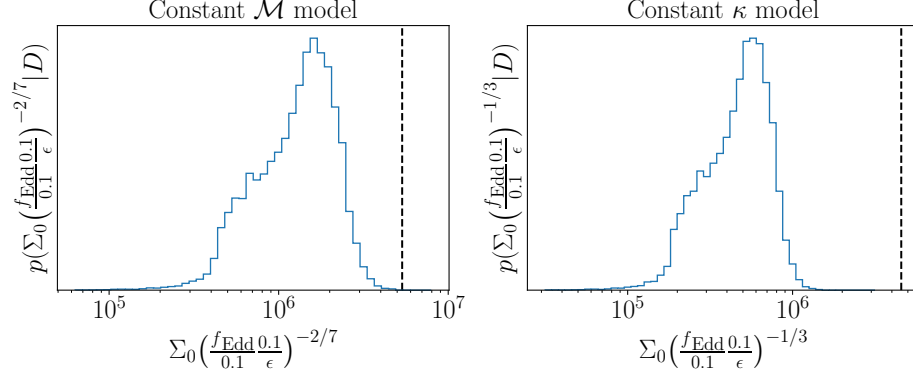


Figure 3.6: Disk properties inferred from gravitational wave observations for case 1, using two different disk models. In the *left panel*, we report the posterior obtained using a constant Mach number $\mathcal{M}$ disk model, matching the one used in the simulations. In black, we show the true value, obtained from the density Σ_0 and radiative efficiency ϵ used in the simulations and f_{Edd} obtained from Eq. (3.5) ($f_{\text{Edd}} = 287$). Even when using the correct disk model, we are unable to accurately infer the properties of the disk due to the stochastic fluctuations in the torque. In the *right panel*, we repeat the analysis with a constant opacity κ disk model. While this choice is compatible with the torque slope inferred from the signal, it does not match the disk model used in the simulations, thus introducing further a bias.

not constant across the disk, but rather evolves as $\mathcal{M} \propto r^{-1/20}$, and the expected value of the power-law slope is $n = 22/5 \simeq 4.4$ rather than $n = 4$. However, these values are still well within the 2σ posteriors, and thus very well compatible with the results of the analysis.

Based on the GW parameter estimation alone, assuming the analytic torque amplitude given by Eq. (3.1) and the disk models described above, one can infer the posterior of $\Sigma_0(f_{\text{Edd}}/\epsilon)^{-2/7}$ for the constant Mach number disk model, and $\Sigma_0(f_{\text{Edd}}/\epsilon)^{-1/3}$ for the constant opacity one. We compare the posteriors inferred for case 1 with the values used in the simulation ($\alpha = 0.01$ with $\epsilon = 0.1$). The true value of f_{Edd} instead is obtained from Eq. (3.5).

As one can see in the left panel of Fig. 3.6, for the constant Mach number disk model the recovered posterior already presents a bias of over two sigmas. This happens because in our analysis we are unable to confidently distinguish events undergoing stochastic type II migration from ones undergoing simple Type I migration. For case 2, for example, we have seen that the posteriors produced by injecting a model undergoing Type I migration and one undergoing gap-opening migration are not significantly different. Since the posteriors on A depend, in the stochastic case, on the averaged net torque, for which an exact relation linking it to the disk properties is unknown, the inference of the environment properties can result in biases even when knowing the correct disk model.

In the right panel of Fig. 3.6, we also report the disk properties inferred using the constant κ disk model. We see here that the bias, already present due to the stochastic

features in the torque, is amplified by the choice of the wrong disk model.

4 Consistency Test for Extreme Mass Ratio Inspirals

Environmental effects are expected to be generally stronger during the early inspiral (i.e., at lower frequencies) [31] and are suppressed, in relative terms, closer to the plunge/merger. If that is the case, the last portion of the signal will be well described by a vacuum EMRI waveform, while inconsistencies/biases arising from neglecting environmental effects will build up from the early inspiral. By constructing data segments of increasing duration, we develop a self-consistency test that compares the vacuum EMRI parameter posteriors obtained from each segment. When the self-consistency test fails, our method indicates the presence of missing, unmodeled effects in the vacuum EMRI template. We apply this test to a fiducial EMRI signal containing an environmental effect (disk-driven migration), and quantify the significance of the inconsistency in the posteriors from different observation durations. In our fiducial system, we show that for observations of more than two years, the presence of an environmental effect can be robustly identified, irrespective of the noise realization. This chapter is based on [78], we use M for primary mass, μ for secondary mass, a for the primary spin and r_0 as the orbital separation.

4.1 Analysis Setup

We inject such an EMRI signal into the time delay interferometry (TDI) [256] channels of LISA using `fastlisaresponse` [145]. Using vacuum EMRI templates, we perform Bayesian parameter estimation of $\theta = \{\log_{10} M/M_\odot, \log_{10} \mu/M_\odot, a, r_0/M, D_L/\text{Gpc}, \cos \theta_S, \phi_S, \cos \theta_K, \phi_K, \Phi_{\phi_0}\}$. Note that a is the dimensionless spin of the primary, r_0/M and Φ_{ϕ_0} are the initial separation and orbital phase, D_L/Gpc is the luminosity distance, θ_S and ϕ_S are the azimuthal and longitudinal sky position in the solar system barycenter frame, and $\cos \theta_K, \phi_K$ are the primary spin orientation relative to the line of sight.

For the injection, we use vacuum EMRI parameters $\theta_{\text{inj}} = \{6.0, 1.69897, 0.9, 14.382, 3.16, 0.751, 0.236, 0.555, 0.628, 0.720\}$, which for a 3 yr observation ¹ yields a signal-to-noise ratio (SNR) of 100 and plunges about 12600 s before the end of the observation window. We note that this is reasonable SNR based on astrophysical population models [25, 149, 202, 203]. We use uniform priors on all parameters except the luminosity distance, for which we use $p(D_L) \propto D_L^2$. Assuming stationary instrumental Gaussian noise, we write the joint log-likelihood considering only the TDI channels $\{A, E\}$. The T channel is not included in our analysis, as it is an antisymmetric TDI combination and is thus mostly dominated by noise [178]. We obtain corresponding posterior samples using

¹The injection value for $r_0/M, \Phi_{\phi_0}$ are the initial separation and phase corresponding to the 3 yr observation. These injected values change with the observation window.

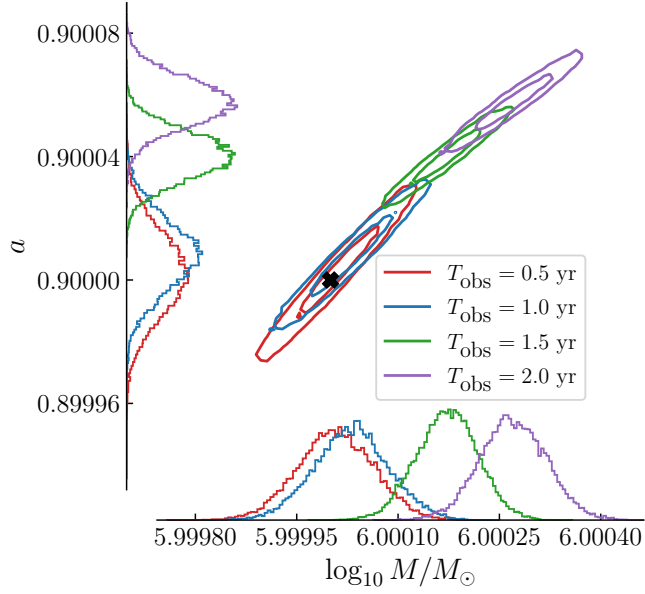


Figure 4.1: Two dimensional marginalized posteriors (with contours at 50% and 90% confidence level) for the primary mass M and spin a , for an injection with disk migration torques but with vacuum EMRI templates, for different observation durations. Significant inconsistencies in the posteriors reveal the neglected environmental effect. For clearness sake we indicate the true value of the parameters with a black cross.

the parallel tempered Markov Chain Monte Carlo sampler implemented in `eryn` [143]. We use the `scirdv1` version of the LISA PSD with second generation TDI [253], to which we add the confusion noise from unresolved galactic binaries [79]. We also consider specific Gaussian noise realizations in the injected data, to assess their impact on the parameter estimation results.

4.2 Inconsistent inference across observation duration.

To perform a self-consistency check of the vacuum EMRI templates, we compare the parameter posteriors resulting from Bayesian inference, with different signal durations. Specifically, we consider signal segments with $T_{\text{obs}} \in \{0.5, 1, 1.5, 2, 2.5, 3\}$ yr, all ending at the same time (right after the plunge). Since most of the SNR is in the last part of the waveform, this ensures that we obtain informative posteriors for the vacuum EMRI parameters for all observation times. Given the injected EMRI signal with Type-I α -disk migration torques, we show in Fig. 4.1 the inferred one-dimensional and two-dimensional marginalized posteriors of the primary mass M and dimensionless spin a , for different T_{obs} . With increasing T_{obs} , the inferred posteriors become more and more inconsistent. In more detail, the posteriors obtained with the vacuum EMRI model for $T_{\text{obs}} = 2$ yr show minimal overlap with those for $T_{\text{obs}} = 0.5$ yr. This reveals the presence of a missing effect

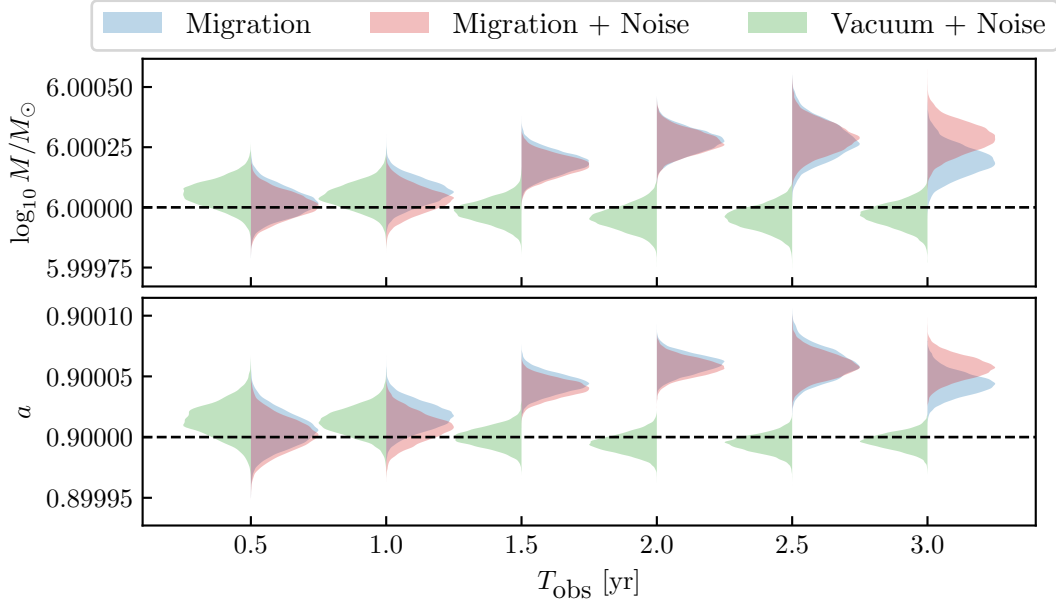


Figure 4.2: One dimensional marginalized posteriors of $\log_{10} M/M_{\odot}$, a for different injections, namely: a vacuum EMRI with noise (green); a migration-affected EMRI with (red) and without (blue) noise. The inconsistency in the posteriors due to the environment can be distinguished from the effect of the noise realization.

in the model—disk-driven migration in this example—that has accumulated significantly between $T_{\text{obs}} = 0.5$ and 2 yr of observation.

To better understand the role of the environmental effect (disk-driven migration) on the EMRI signal, we compute the resulting orbital dephasing (relative to the vacuum case), for the same injected vacuum EMRI parameters. In the top panel of Fig. 4.3, we show the orbital dephasing as a function of both the observation duration and (equivalently) the orbital separation r_0/M . Because the disk-driven migration (in a Type-I α -disk) considered in our work is a -8PN effect, Fig. 4.3 shows a significant orbital dephasing $\gtrsim 50$ rad for a 3 yr observation. Conversely, for a shorter observation of $\lesssim 1$ yr, the dephasing is $\lesssim 3$ rad, the inference inconsistencies become more significant with longer observation duration, because the environmental effects are dominant at lower frequencies.

We also report, in the bottom panel of Fig. 4.3, the squared residual SNR corresponding to the MLE [66, 167, 264], given by $\rho^2_{\text{res}} \equiv -\log \mathcal{L}(d|\theta_{\text{MLE}})$, which we calculated for the noiseless injections with migration torques. As expected, for longer observations, an increasing amount of signal is lost in the analysis, due to the missing physics in the template. The signal loss is never large enough to prevent detection, but in the context of the LISA global fit, the residual SNR from many sources may accumulate and potentially cause biases and/or missed detections [18].

To contrast the role of instrumental noise with that of the environmental effect, we consider a vacuum EMRI injection with a Gaussian noise realization, and compare it

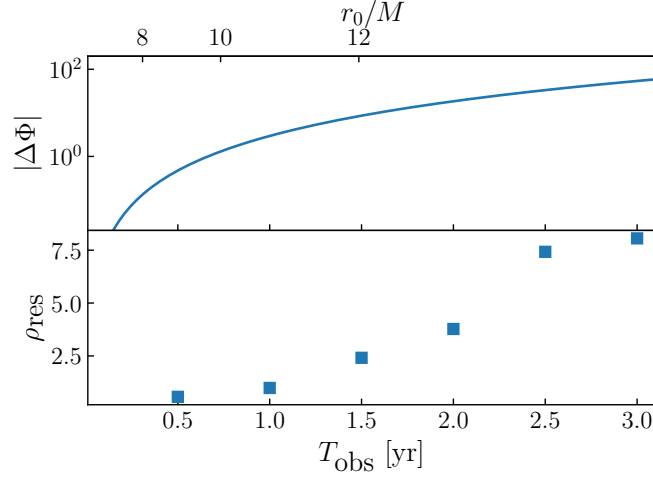


Figure 4.3: Top panel shows the orbital dephasing (as a function of duration and orbital separation r_0/M) between a noiseless EMRI with migration torques and the corresponding vacuum case. For this injection, the bottom panel shows the residual SNR obtained from the maximum likelihood vacuum EMRI waveform.

with injections featuring disk-driven migration with and without such a noise realization. In Fig. 4.2, we compare the resulting one-dimensional marginalized posteriors for $\log_{10} M/M_\odot$, a from these different injections. When injecting a vacuum EMRI signal, we observe that the (green) posteriors are statistically self-consistent, with shifts induced by the noise realization that have no significant trend with observation duration. However, the (blue) posteriors resulting from the environment-affected EMRI injection without noise clearly show significant inconsistencies growing with the observation duration. Note that given the high SNR, the posteriors obtained from this noiseless injection are a good approximation for the ensemble average (over noise realizations) of the posteriors obtained with noisy injections [199, 263]. Importantly, we also observe good agreement between these ‘average’ posteriors and those obtained for an environment-affected EMRI with a specific noise realization (red). We find that for short observation durations $T_{\text{obs}} \leq 1$ yr, the inconsistency due to the missing environmental effect is comparable with the effect of the instrumental noise realization. For longer observation durations, the inconsistency is apparent and becomes distinguishable from the effects of the noise realization.

We will now quantify the statistical significance of the inference inconsistency, shown in Fig. 4.1 and Fig. 4.2, using both a *mismatch* criterion and a *relative systematic bias*, which we describe below. For a given waveform model (including the LISA response) $h(\theta)$, the mismatch $\mathcal{M}(\theta_1, \theta_2)$ quantifies the distance between two points θ_1 and θ_2 in parameter space, with the mismatch $\mathcal{M}(\theta_1, \theta_2)$. Following [187], the match is maximized over overall phase and time shifts, which are needed to align the waveforms. When computing the mismatch across all channels, denoted as $\mathcal{M}_{\text{net}}$, we sum over each inner product as discussed in [85].

Using the maximum likelihood estimate (MLE) $\theta_{\text{MLE,ref}}$ obtained with $T_{\text{obs}} = 0.5$ yr

as a reference, we compute the matches (using the vacuum EMRI waveform templates) to $\boldsymbol{\theta}_{\text{MLE}}$ as obtained from other observation durations. We denote these mismatches by $\mathcal{M}_{\text{net}, T_{\text{obs}}} \equiv \mathcal{M}_{\text{net}}(\boldsymbol{\theta}_{\text{MLE, ref}}, \boldsymbol{\theta}_{\text{MLE}})$ for convenience. In each case, to compute the match, we generate the waveforms using an observation duration of $T_{\text{obs}} = 0.5, \text{yr}$, thus ensuring that their duration matches that of the reference. Based on [26, 66, 70, 112, 167, 258], we consider the Bayesian inference to be significantly inconsistent when the mismatch satisfies

$$\mathcal{M}_{\text{net}, T_{\text{obs}}} > \frac{\chi_{D, 90\%}^2}{2\rho_{\text{ref}}^2}, \quad (4.1)$$

with $\rho_{\text{ref}} = (\sum_i \|h_i(\boldsymbol{\theta}_{\text{MLE, ref}})\|^2)^{1/2}$ being the optimal SNR [85] of the reference waveform, and $\chi_{D, 90\%}^2$ being the 90% quantile for the chi-square distribution with $D = \dim(\boldsymbol{\theta})$ degrees of freedom. For $D = 10$, we have $\rho_{\text{ref}} = 99$ and $\chi_{D, 90\%}^2 = 16$, resulting in threshold mismatch of 8.3×10^{-4} . We also assess the inconsistency in the MLEs by computing the relative systematic bias as

$$\delta\theta^\alpha \equiv \frac{|\theta_{\text{ML, ref}}^\alpha - \theta_{\text{ML}}^\alpha|}{\Sigma_{90\%}^\alpha}, \quad (4.2)$$

where α is the parameter index, and $\Sigma_{90\%}^\alpha$ is the parameter's 90% credible interval for the reference case $T_{\text{obs}} = 0.5\text{yr}$. With a large mismatch satisfying Eq. (4.1), one would typically expect significant biases in all parameters, with $\delta\theta^\alpha > 1$. However, $\delta\theta^\alpha > 1$ for a particular parameter does not imply that Eq. (4.1) is satisfied. For this reason, we primarily use Eq. (4.1) as the main quantitative metric to identify inconsistent inferences, and report the corresponding relative systematic biases for a few key intrinsic parameters such as $\log_{10} M$ and a .

Table 4.1: Mismatches and systematic biases relative to the reference MLE at $T_{\text{obs}} = 0.5\text{yr}$. The SNR for 3 yr of observation is 100.

T_{obs} (yr)	$\mathcal{M}_{\text{net}, T_{\text{obs}}}$	$\delta \log_{10} M/M_\odot$	δa
1.0	9.0×10^{-5}	0.10	0.10
1.5	4.8×10^{-4}	1.60	1.60
2.0	1.4×10^{-3}	2.58	2.60
2.5	1.6×10^{-3}	3.13	3.17
3.0	3.2×10^{-3}	3.10	3.14

In Table 4.1, we show the mismatches with the reference MLE computed with Eq. (2.10), along with the relative systematic biases $\delta \log_{10} M/M_\odot$ and δa for each observation duration. We find that for $T_{\text{obs}} \leq 1.5\text{yr}$, the mismatch is below the threshold (as given in Eq. 4.1). For a longer duration $T_{\text{obs}} = 2\text{yr}$, the mismatch exceeds the threshold, indicating a significant inconsistency. For $T_{\text{obs}} \geq 2.5\text{yr}$, the inconsistency is so significant that the probability that it is induced by noise is $\leq 5 \times 10^{-4}$, which corresponds to $\delta \log_{10} M/M_\odot > 3$ and $\delta a > 3$. Our quantitative analysis with Eqs. (4.1)–(4.2) therefore confirms the results shown in Fig. 4.1–4.2.

4.3 Further Validation Tests

As a validation of our self-consistency test, we also injected a vacuum EMRI signal with noise. We considered the same source parameters θ_{inj} and observation durations as the one discussed in section 4.1, and inferred posteriors on θ . The results are reported in Table 4.2. We see that the mismatch criterion given by (4.1) is always well satisfied. Note that noise can cause a relative systematic bias $\delta\theta^\alpha > 1$ in the one-dimensional marginalized posteriors, while the mismatch criterion is still satisfied. Indeed, since the mismatch criterion takes into consideration the whole parameter space, it is a more reliable diagnostic of the presence of environmental effects than the one-dimensional posteriors.

We have checked that different noise realizations do not yield statistically different results for both the vacuum and migration injections. In Table 4.3, we show the mismatches obtained with $T_{\text{obs}} = 2.5$ yr and $T_{\text{obs}} = 3.0$ yr observations against the reference $T_{\text{obs}} = 0.5$ yr observation. We considered two different noise realizations for both the vacuum and migration injections, using the same source parameters as those described in section 4.1. We find that the mismatches obtained for given T_{obs} are comparable across noise realizations, and agree with Tables 4.2 and 4.1. Additionally, we tested whether an inconsistent inference across observations can be caused by a slight mismodeling of the noise PSD, instead of a mismodeling of a physical effect. We considered two vacuum injections in which the PSD used to produce the noise and the one in the likelihood calculation differ. In the first case we assume that the acceleration component of the LISA noise curve is underestimated during the global fit. Specifically we increase the acceleration noise component by 10% in the PSD used in the noise generation compared to the one used in the data analysis pipeline. In the second case we consider a slight mismodel in the galactic binaries foreground. Using the empirical description for unresolved galactic binaries we simulate a situation in which the number of extracted GBs is overestimated. Specifically we use the empirical fits introduced in [142] and we use an observation time which is shorter by 2 months in the PSD used for noise generation compared to the one used for parameter estimation. In both cases no inconsistency between the posteriors has been found and the mismatch values obtained were comparable with the ones reported in table 4.3 for different noise realizations. We therefore conclude that our results are robust to different noise realizations and to noise mismodeling.

We also explored how our self-consistency test performs for injections containing a different environmental effect, or corresponding to different source parameters. For the same source parameters mentioned in the 4.1, which corresponds to an SNR of 100 with a 3 yr observation, we modified the environmental effect by choosing different values of n (keeping the relative torque magnitude A fixed). For event A, we set $n = 4$ (that contributes at -4PN relative to the GW flux), which corresponds to environmental effects such as gas accretion and center-of-mass acceleration, or modification to gravity such as a time dependent Newton's constant [64, 224, 257, 277, 279]. For events B and C, we chose $n = 5.5$ (that contributes at -5.5PN relative to the GW flux), which corresponds to non-

Table 4.2: Mismatches and systematic biases relative to the reference MLE with $T_{\text{obs}} = 0.5$ yr, for a vacuum injection with noise.

T_{obs} (yr)	$\mathcal{M}_{\text{net}, T_{\text{obs}}}$	$\delta \log_{10} M$	δa
1.0	5.8×10^{-5}	0.07	0.04
1.5	1.5×10^{-4}	0.59	0.49
2.0	3.2×10^{-4}	0.92	0.76
2.5	6.0×10^{-4}	1.16	0.87
3.0	4.8×10^{-4}	0.99	0.74

Table 4.3: Mismatch for different noise realizations, for the vacuum and migration injections described in section 4.1. Note that the vacuum mismatches are always below the threshold.

T_{obs} (yr)	Vacuum		Migration	
	Noise 1	Noise 2	Noise 1	Noise 2
2.5	4.8×10^{-5}	8.7×10^{-5}	1.3×10^{-3}	1.3×10^{-3}
3.0	7.0×10^{-5}	5.7×10^{-5}	1.0×10^{-3}	8.8×10^{-4}

relativistic dynamical friction in a uniform collisionless medium [64, 67, 201]. We also chose different source parameters for event C, setting $M = 5 \times 10^5 M_{\odot}$, $\mu = 30 M_{\odot}$, $a = 0.7$, which yields an SNR of 60 for a 3yr observation. In Tables 4.4 and 4.2, we show the results of the self-consistency test for each event.

For events A and B, the posteriors become inconsistent with the reference MLE for $T_{\text{obs}} \geq 2.5$ yr, in agreement with the accumulated dephasing relative to the vacuum template. Moreover, at fixed T_{obs} , event B exhibits larger mismatches than event A. We attribute this to the larger dephasing in event B at given T_{obs} , since the environmental effect enters at a more negative PN order. For event C, we find that the mismatch threshold is 2.2×10^{-3} , due to its lower SNR. We observe that event C is fully incompatible with vacuum at $T_{\text{obs}} = 2.0$ yr. Also note that for a longer observation time $T_{\text{obs}} = 2.5$ yr, the mismatch is comparable to the threshold, and then exceeds it at $T_{\text{obs}} = 3.0$ yr. Even though the behavior of the mismatch as a function of T_{obs} is not strictly monotonic for event C, unlike for events A and B, the fact that it becomes comparable to and then exceeds the threshold for $T_{\text{obs}} \geq 2$ yr signals an inconsistency due to using the vacuum template.

Since the main focus of this work is on the detection of unmodeled contribution coming from the astrophysical environment we considered negative PN modification to the gravitational fluxes. However, positive PN modifications to the flux can instead arise due to modified gravity [278]. We consider here a 1PN effect that would correspond to a propagation effect due to massive gravity [278], in order to contrast against the systematics of negative PN effects. To see the impact on the self-consistency test due to neglecting such a 1PN correction, we pick the fiducial values $A = 2 \times 10^{-4}$ and $n = -1$

Table 4.4: Mismatches, systematic biases relative to the reference MLE at $T_{\text{obs}} = 0.5 \text{ yr}$ and dephasing from the vacuum trajectory, for different injections. Events A and B have the same source parameters as the source in Fig. 4.1, but a different power law for the environmental torque. Event C has different source parameters ($M = 5 \times 10^5 M_{\odot}$, $\mu = 30 M_{\odot}$, $a = 0.7$, $\text{SNR} = 60.0$) and the same torque as event B.

Event	T_{obs} (yr)	$\mathcal{M}_{\text{net}, T_{\text{obs}}}$	$\delta \log_{10} M/M_{\odot}$	δa	$\Delta \Phi$
A	2.0	2.5×10^{-5}	1.22	1.00	10.7
	2.5	6.5×10^{-4}	1.76	1.77	15.4
	3.0	1.3×10^{-3}	2.38	2.43	20.7
B	2.0	4.6×10^{-4}	1.55	1.53	12.8
	2.5	1.2×10^{-3}	2.69	2.73	19.9
	3.0	1.8×10^{-3}	3.10	3.15	28.7
C	2.0	3.6×10^{-3}	1.93	2.00	34.3
	2.5	2.0×10^{-3}	1.48	1.51	53.2
	3.0	2.2×10^{-3}	1.50	1.54	76.2

for the injection, together with the same EMRI parameters considered in 4.1. In Fig. 4.4, we show the posteriors for the primary mass obtained with a vacuum EMRI model, corresponding to a 0.5 yr (red histogram) and 3 yr (blue histogram) observation. We see that there is significant inconsistency between the two observations, thus indicating that the self-consistency test also works for positive PN effects. However, we also note that the inference with 0.5 yr observation is biased relative to the true injected value, unlike what we saw with negative PN order environmental effects. This is somewhat expected as the modification becomes more significant close to merger, while in the previous cases the late inspiral modification was negligible. The additional bias between the observation close to merger and the one with the full inspiral is due to the accumulated dephasing between observation durations, which is why this applies to both positive and negative PN effects.

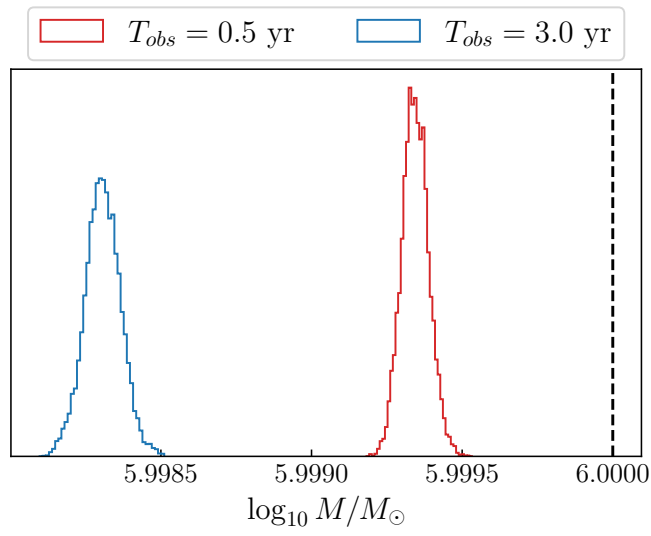


Figure 4.4: Comparison of the posterior for $\log_{10} M$ obtained from an observation time of 0.5 yr and 3.0 yr. In black, we report the injection value. The injection coincides with the one in section 4.1 for the vacuum parameters while here we have set $A = 2.0 \times 10^{-4}$ and $n = -1.0$.

5 Massive Black Hole Binaries, Environment vs Modified gravity

Environmental effects can bias the tests of GR that LISA MBHBs are expected to enable with unprecedented precision. Accretion and gas drag enter the waveform phase at *negative* post-Newtonian (PN) orders, and so do many of the best-motivated modifications of GR. Both classes of effects therefore affect the same low-frequency, early-inspiral portion of the signal. Examples of such GR modifications include a secular variation of Newton’s constant [244, 279], scalar-dipole radiation, and violations of the strong equivalence principle [35]. In this chapter, based on the work done on [76] we focus on a realization of this overlap: a GR modification and a set of environmental effects that all enter the inspiral phase at exactly the same -4 PN order, and are therefore formally indistinguishable on a single source. As the GR-modification channel we adopt a secular drift of the gravitational constant, $\dot{G}/G$ [244, 279]. Current bounds on the fractional drift rate range from $|\dot{G}/G| \lesssim 10^{-12} - 10^{-13} \text{ yr}^{-1}$ in the quasi-static Solar-System regime [27] to much weaker limits in the strong-field, dynamical regime accessible to ground-based GW observations of SOBHBs [280]. In the environmental sector, we consider the two expected leading effects of a gaseous circumbinary disk. The first is gas accretion onto the components, which slowly increases their masses and, by angular-momentum conservation, induces a -4 PN dephasing [65]. The second is gas-driven migration, i.e. the torque exerted on the binary by the disk, which contributes at the same -4 PN order, with the same frequency dependence as accretion but a different overall prefactor. At leading order, all three effects are therefore captured by a single -4 PN deformation of the inspiral phase, and the signal of an *individual* MBHB cannot distinguish among them. The two classes of effects, however, behave differently across a population of sources: environmental contributions are expected to vary in magnitude from one binary to another, whereas a putative $\dot{G}$ drift is universal. A hierarchical analysis of the full *population* of LISA MBHBs can therefore in principle still disentangle the two, in analogy with what has been proposed for EMRIs [149].

5.1 Modifications at -4 PN order

Rather than introducing a dedicated template for each candidate effect, we adopt the parameterized post-Einsteinian (ppE) framework [278], in which a single parameter captures the dominant phase deformation at a given PN order. For comparable-mass binaries near merger—the regime probed most precisely by current LVK tests of GR—negative-PN corrections of the kind considered here are subdominant to radiation reaction. However, at the lower orbital frequencies accessible to LISA MBHBs, these corrections

become a natural target. We specialize the ppE template to the $\dot{G}$ case in Sec. 5.1.1, to accretion and viscous migration in Sec. 5.1.2, and present the resulting Fisher-matrix parameter-estimation results across parameter space in Sec. 5.2.2.

5.1.1 Gravitational constant variation

We adopt here a phenomenological modified-gravity model in which the gravitational constant is time-dependent, as would be the case, e.g., in scalar-tensor theories [17, 87, 88, 260]. We parametrize this time variation as $G \approx G_N [1 + G_1 (t - t_0)]$, where t_0 is the present time and G_N is the present-day value of the Newton constant. Note that when computing the effect on a binary inspiral, keeping only the first-order term $G_1(t - t_0) \ll 1$ is sufficient. One can neglect the modification to the waveform amplitude, and consider only the dephasing contribution, which enters at -4 PN order with respect to the leading-order quadrupole phase [244, 279]:

$$\delta\varphi = \Psi_{\text{TF2}} C_{G_1} \frac{M}{M_\odot} \eta^{-1} \frac{G_1}{\text{yr}^{-1}} (\pi M f)^{-8/3}, \quad (5.1)$$

where M is the total mass of the binary, $\eta = m_1 m_2 / M^2$ is the symmetric mass ratio, f is the wave frequency, $C_{G_1} = -3.52 \times 10^{-15}$ and $\Psi_{\text{TF2}} = (3/128\eta)(\pi M f)^{-5/3}$ is the leading-order phase of the TaylorF2 approximant. As discussed in Sec. ??, the most stringent current bounds on G_1 come from quasi-static Solar-System experiments, $|G_1| \lesssim 10^{-12} - 10^{-13} \text{ yr}^{-1}$ [27], while strong-field GW observations from LVK comparable-mass binaries give limits several orders of magnitude weaker, $|G_1| \lesssim 10^2 \text{ yr}^{-1}$ [280].

5.1.2 Environmental dephasing: accretion and viscous migration

About 10% of observed supermassive black holes are active galactic nuclei (AGN) [130, 230], a duty cycle that we will use in Sec. 5.3 as order of magnitude for the active-fraction hyperparameter λ of the population model. Galaxy mergers generically drive gas inflows toward the galactic nucleus [38, 135, 186], so MBHBs can plausibly be embedded in a gas-rich environment in which the surrounding material settles into a circumbinary disk [181]. This disk drives the accretion and migration effects discussed below. In the presence of such a gaseous environment, the two black holes accrete and grow in mass. The accretion timescale is expected to be comparable to the Salpeter timescale $\tau = 4.5 \times 10^7 \text{ yr}$ [221], so that, assuming that both components accrete at the same rate, the total mass of the system evolves as $M(t) = M_0 \exp(f_{\text{Edd}}(t - t_0)/\tau) \simeq M_0 [1 + f_{\text{Edd}}(t - t_0)/\tau]$, where M_0 is the total mass at t_0 . Since the LISA observation time is much shorter than τ , we keep only the linear order. Inserting this time-dependent mass into the quadrupole formula and imposing angular-momentum conservation [65], one finds that the induced dephasing is

$$\delta\varphi = \Psi_{\text{TF2}} C_{\text{acc}} \frac{M}{M_\odot} \eta^{-1} \frac{f_{\text{Edd}}}{0.1} (\pi M f)^{-8/3}, \quad (5.2)$$

where $C_{\text{acc}} = 1.75 \times 10^{-23}$.

Another effect that can arise in a gaseous astrophysical environment is the migration torque exerted by the disk onto the binary. In this case, which we will refer to as viscous migration, the dephasing enters at the same -4 PN order [119, 120]:

$$\delta\varphi = \Psi_{\text{TF2}} C_{\text{mig}} \frac{M}{M_{\odot}} \eta^{-2} \frac{\xi f_{\text{Edd}}}{0.1} (\pi M f)^{-8/3}, \quad (5.3)$$

where $C_{\text{mig}} = 1.0 \times 10^{-22}$ and ξ encodes the efficiency of the disk-binary coupling; extrapolating the results of numerical simulations to thin disks, ξ may reach values of $\mathcal{O}(10)$ [96]. The extra factor $\eta^{-1}\xi$ in Eq. (5.3) relative to Eq. (5.2) makes migration more efficient than accretion at the same Eddington ratio, especially for asymmetric mass ratios.

5.2 Population models and single-event analysis

The parameter space spanned by LISA MBHBs covers several orders of magnitude in both chirp mass and distance, and the number of detected events in any given region of this space depends sensitively on the assumed astrophysical history of the MBHB population. We therefore rely on a suite of semi-analytic population models, each encoding different choices for the seeding mechanism, the delay time between the galaxy merger and the subsequent MBH merger, and the impact of supernova (SN) feedback on MBH accretion. The catalogs produced by these models are the basis for the detectability and bias estimates presented in the rest of this section.

5.2.1 Population models

The formation history of MBHBs is not well constrained observationally, and plausible evolutionary sequences spanning high redshift to the present yield substantially different catalog realizations. We adopt the semi-analytic model for the formation and evolution of galaxies and MBHs originally introduced in [29]. Additions and improvements to specific aspects of the model were subsequently introduced in [21, 33, 51, 52, 226]. The catalogs are publicly available as part of the BACH data release [30].

We consider five population models, three HS and two LS. In the HS class, the black-hole seeds arise from the collapse of proto-galactic disks induced by bar instabilities [267], with typical seed masses of the order of $\sim 10^5 M_{\odot}$:

- HS-nod-SN (from [33]): no delay between the galaxy merger and the MBH merger (except for the dynamical-friction timescale, including tidal effects between dark-matter halos), and including the effect of SN feedback on MBH accretion;
- HS-nod-SN-high-accr (from [32, 33]): as above, but with the MBH accretion rate boosted by a factor of ~ 4 ;
- Q3-nod (from [156]): an earlier HS realization with no delay and no SN feedback on accretion.

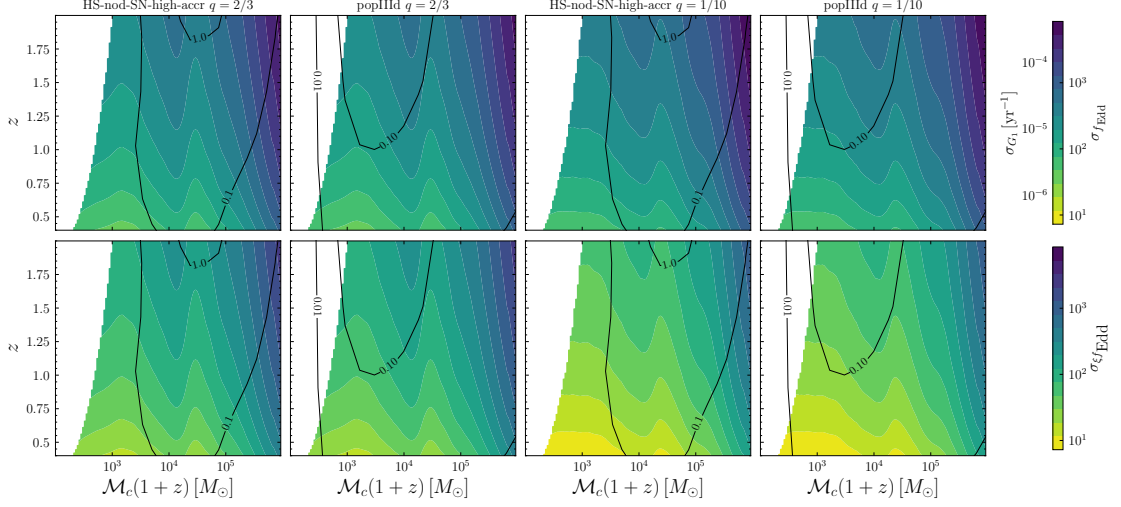


Figure 5.1: Sky-averaged 1σ uncertainty on the three -4 PN parameters across the (detector-frame) chirp-mass-redshift plane, for 10 yr of LISA observation and mass ratios $q = 2/3$ (left) and $q = 1/10$ (right). Top panels: uncertainty on the Eddington fraction f_{Edd} (accretion) and on G_1 (modified gravity). Bottom panels: uncertainty on the combination ξf_{Edd} relevant for viscous migration. Contours show the merger-rate density $d^3N/(d \log_{10} \mathcal{M}_{c,z} dz dt_c) \times 1 \text{ yr}$ for the HS-nod-SN-high-accr and popIII-d models, which bracket the population scenarios considered in this work. Points with sky-averaged $\text{SNR} < 8$ across the three TDI channels are masked out.

In the LS class, the seeds form instead from the remnants of population III stars in high-redshift, low-metallicity systems [173]:

- LS-nod-SN (from [33]): no delay between the galaxy merger and the MBH merger (except for the dynamical-friction timescale, including tidal effects, between dark-matter halos), and including the effect of SN feedback on MBH accretion;
- popIII-d (from [156]): an earlier LS realization that includes, on top of the dynamical-friction timescale with tidal effects between dark-matter halos, further delay contributions from stellar hardening, triple-MBH interactions, and gas-driven migration.

Together, these five population models span a representative range of astrophysically motivated assumptions about the seeding mechanism and accretion history of MBHs across cosmic time.

5.2.2 LISA sensitivity across parameter space

We now compare the sensitivity of LISA to the effects described in Sec. 5.1 across the parameter space. Since all of these effects enter at the same PN order, we consider an agnostic dephasing term $\delta\varphi = \Psi_{\text{TF2}} C_0 (\pi M f)^{-8/3}$, which allows one to estimate the LISA sensitivity to generic -4 PN phase terms. The model-specific dependence on the intrinsic and beyond-GR source parameters is encoded in the agnostic coefficient C_0 , and

we can recover information on f_{Edd} or G_1 via the physical mapping defined by Eqs. (5.2) and (5.3). Note that the Jacobian of this transformation is essentially constant over the posterior, so that a uniform prior on C_0 maps to an effectively uniform prior on f_{Edd} (or G_1).

To map out LISA’s reach across parameter space, we evaluate the Fisher matrix on a grid over the plane of (detector-frame) chirp mass and redshift $(\mathcal{M}_{c,z}, z)$, with $\mathcal{M}_c \equiv (m_1 m_2)^{3/5} / (m_1 + m_2)^{1/5}$ and $\mathcal{M}_{c,z} = \mathcal{M}_c(1+z)$, at four representative values of the mass ratio $q = 2/3, 1/3, 1/5, 1/10$. For each point in the $(\mathcal{M}_{c,z}, z)$ grid we compute the Fisher matrix over 500 Monte Carlo realizations of the extrinsic parameters (sky position, inclination, and polarization angle), while fixing the mass ratio and spins to fiducial values. The nominal LISA mission duration is 4 yr, with an extended-phase option reaching 10 yr [75]; throughout this work we adopt the optimistic 10 yr figure. For accretion and viscous migration under realistic astrophysical assumptions, MBHBs in the LISA band have $C_0 \sim 10^{-15}$. Since the dephasing depends linearly on C_0 , and this additional contribution is strongly subdominant compared to the radiation-reaction-driven phase evolution, the Fisher matrix is approximately independent of C_0 .

In Fig. 5.1 we show the sky-averaged statistical error on the physical parameter characterizing each of the three -4 PN channels, for mass ratios $q = 2/3$ and $q = 1/10$, together with the merger-rate distribution of two of the population models discussed in Sec. 5.2.1. We mask out the points for which the sky-averaged signal-to-noise ratio (SNR) across the three LISA Time-Delay Interferometry (TDI) channels is below 8. The top panels show the 1σ uncertainty on the Eddington fraction f_{Edd} (accretion) and on the fractional drift rate G_1 (modified gravity); the bottom panels show the corresponding uncertainty on the combination ξf_{Edd} relevant for viscous migration. The figures are overlaid with the merger-rate $\frac{d^3 N}{d \log_{10} \mathcal{M}_{c,z} dz dt_c} \times 1 \text{ yr}$ contours for the HS-nod-SN-high-accr and popIII-d models, which bracket the range of formation models explored above. In agreement with previous single-event forecasts [120], the sensitivity to all three -4 PN deformations is concentrated at low chirp mass and low redshift, where the early inspiral spends months to years in band; conversely, the massive sources that dominate the HS-nod-SN-high-accr merger rate fall in the poorly-constrained top-right corner of each panel. Even in the optimal cases—long observation time and low redshift and mass—the expected average statistical error on the environmental parameters is generally well above $f_{\text{Edd}} = 1$ for both the accretion and the viscous-migration case.

Naively one might expect LS models to be the optimal case for the inference of these negative-PN effects. We find instead that the HS models deliver more mergers per year with low statistical error on the -4 PN physical parameters. The underlying reason is the difference in seed mass between the two classes. LS mergers typically have chirp masses in the range $500\text{--}1000 M_\odot$, with the merger—where most of the SNR is accumulated—outside the LISA band; despite their long in-band inspiral, such events are therefore faint and only the closest to us are detectable. SN feedback [127] further reduces the detectable population by stunting seed growth [33, 127], so that the merger rate at $\mathcal{M}_c \sim 10^4\text{--}10^5 M_\odot$ is strongly suppressed in LS-nod-SN relative to a popIII-remnant scenario without SN feedback. In the HS models, by contrast, the seeds are already

born at $\sim 10^5 M_\odot$, i.e. directly in the LISA chirp-mass window, so the merger rate at $\mathcal{M}_c \sim 10^4\text{--}10^5 M_\odot$ is largely insensitive to the SN-feedback prescription and is typically an order of magnitude higher than in LS-nod-SN (Fig. 5.1).

5.2.3 Single-event inference

The Fisher-matrix parameter-estimation results of Fig. 5.1 describe the sky-averaged reach of LISA as a function of the intrinsic source parameters. To move from this sky-averaged picture to more realistic per-catalog forecasts, we now draw specific realizations of each MBHB population and analyze them event by event. For each population model the intrinsic parameters (component masses, spins, and luminosity distance) are drawn from the associated merger-rate distribution, while the extrinsic parameters (sky position, time and phase of coalescence, polarization, and inclination) are drawn from uninformative (uniform or isotropic) distributions. We retain only events exceeding the detection threshold $\text{SNR} > 8$.

We analyze each catalog under two contrasting assumptions for how the parameter estimation treats potential -4 PN contributions to the signal. In the *first scenario*, the waveform templates include a free -4 PN phase deformation parametrized by C_0 , whose posterior we estimate from the Fisher information matrix computed over the intrinsic parameters *and* C_0 . Figure 5.2 summarizes the outcome for the accretion and viscous-migration templates, respectively: for each 10 yr-catalog, we report the number and fraction of detectable events with $\sigma_{f_{\text{Edd}}}$ below 50, 100 (accretion) or 10, 50, 100 (migration). The tighter threshold is not reported for the accretion case, because in most catalogs no event satisfies $\sigma_{f_{\text{Edd}}} \lesssim 10$.

In a second scenario, we assume that parameter estimation is performed with vacuum GR templates, and any unmodeled negative-PN contribution induces a systematic bias on the recovered parameters. We quantify the per-event bias with the Cutler–Vallisneri formula [86],

$$\Delta\theta_j = \Gamma_{ij}^{-1} \langle d - h(\theta) | \partial_j h(\theta) \rangle, \quad (5.4)$$

where Γ_{ij} is now the vacuum-template Fisher matrix, evaluated at the true value of the parameters. The residual between the data and the vacuum template can be approximated as

$$d - h(\theta) = h(\theta)(e^{i\delta\varphi} - 1) \simeq i\delta\varphi h(\theta), \quad (5.5)$$

which, inserted into Eq. (5.4), shows that the bias scales linearly with f_{Edd} (or with G_1). It is therefore sufficient to compute the bias at a fiducial f_{Edd} or G_1 , and to rescale to the desired value. We have verified this by comparing the bias computed with the approximated and with the full residual formulae.

The dominant effect of neglecting a negative-PN contribution is a shift in the recovered chirp mass, while the recovered spins and mass ratio remain essentially unaffected with respect to the statistical errors, even for super-Eddington injections. Moreover, the chirp-mass shift itself is negligible for the vast majority of events, consistent with the limited detectability of generic -4 PN corrections across the parameter space discussed in

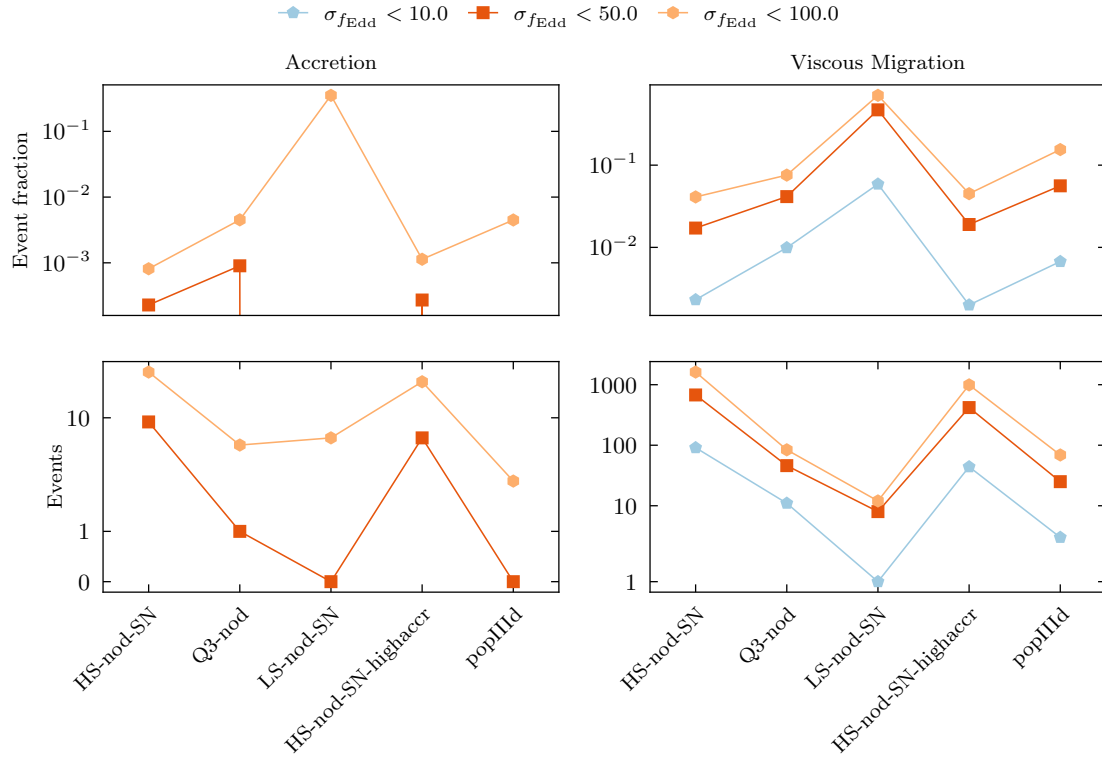


Figure 5.2: Number and fraction of events detectable in 10 yr of LISA observation whose Fisher-predicted single-event uncertainty $\sigma_{f_{\text{Edd}}}$ falls below a threshold, for various population models. Top panels: fraction; bottom panels: absolute number. *Left:* accretion injection, with thresholds $\sigma_{f_{\text{Edd}}} < 50$ (red) and < 100 (orange). *Right:* viscous-migration injection, with thresholds $\sigma_{f_{\text{Edd}}} < 10$ (blue), < 50 (red) and < 100 (orange).

Sec. 5.2.2. To quantify the shift, we normalize it to the Fisher-predicted 1σ uncertainty on $\mathcal{M}_c$,

$$\frac{\delta\mathcal{M}_c}{f_{\text{Edd}}} \equiv \frac{\Delta\mathcal{M}_c}{f_{\text{Edd}} \sigma_{\mathcal{M}_c}}, \quad (5.6)$$

and count, for each 10 yr- catalog, the number of events for which this normalized shift exceeds 0.01, 0.1, and 1 (assuming at $f_{\text{Edd}} = 1$). The results are reported in Table 5.1. No event in any catalog is shifted by more than 10% of the statistical uncertainty: accretion at physically plausible rates leaves essentially no imprint on the vacuum-template chirp-mass measurement, and only at unrealistically high accretion rates, $f_{\text{Edd}} \sim 10^2\text{--}10^3$, does the shift become comparable to the statistical error.

catalog	$\frac{\delta\mathcal{M}_c}{f_{\text{Edd}}} > 0.01$	$\frac{\delta\mathcal{M}_c}{f_{\text{Edd}}} > 0.1$
HS-nod-SN-high-accr	33	0
HS-nod-SN	41	0
LS-nod-SN	11	0
popIII-d	15	0
Q3-nod	5	0

Table 5.1: Average number of sources observed in 10 yr whose vacuum-template chirp-mass bias, normalized to the Fisher 1σ uncertainty, exceeds the thresholds $\Delta\mathcal{M}_c/(f_{\text{Edd}} \sigma_{\mathcal{M}_c}) > 0.01, 0.1$, assuming every event undergoes accretion at $f_{\text{Edd}} = 1$, for various population models.

Repeating the same exercise for a viscous-migration injection at $f_{\text{Edd}} = 1$, we find a substantially stronger effect (Table 5.2). This is expected: the migration dephasing scales as η^{-2} rather than the η^{-1} of accretion [Eqs. (5.2)–(5.3)], and is therefore amplified for asymmetric mass ratios. Already at $f_{\text{Edd}} = 1$, individual sources in several catalogs show a chirp-mass shift comparable to the statistical uncertainty. Viscous migration can be further enhanced or suppressed by the internal structure of the disk through the efficiency factor ξ : current numerical simulations give $\xi = \mathcal{O}(1)$ [98], but those simulations do not resolve realistically thin disks because of their computational cost, and it has been argued that for a thin circumbinary disk ξ may be as large as $\mathcal{O}(10)$ [96, 119]. Under this more optimistic assumption, Table 5.2 implies that up to a few events per year in the HS catalogs could exhibit chirp-mass shifts at or above the statistical error.

catalog	$\frac{\delta\mathcal{M}_c}{f_{\text{Edd}}} > 0.01$	$\frac{\delta\mathcal{M}_c}{f_{\text{Edd}}} > 0.1$	$\frac{\delta\mathcal{M}_c}{f_{\text{Edd}}} > 1$
HS-nod-SN-high-accr	1132	66	1
HS-nod-SN	1812	117	2
LS-nod-SN	15	3	1
popIII-d	139	5	0
Q3-nod	89	11	0

Table 5.2: Same as Table 5.1, but for a viscous-migration injection at $f_{\text{Edd}} = 1$ and fiducial $\xi = 1$.

Together, these per-event uncertainties and biases on the template parameters are the input to the hierarchical population analysis of Sec. 5.3, where they are combined to yield joint population constraints on $\dot{G}$ and on the environmental parameters.

5.3 Population-level inference

For single events, the precision to which a negative-PN contribution can be measured is limited. The different physical effects considered in this work leave essentially the same imprint: they all induce a -4 PN dephasing and are therefore degenerate at the level of an individual source. Information at the *population* level, however, can break part of this degeneracy. A putative drift of Newton’s constant is expected to act in the same way on every source, whereas environmental effects act only on the fraction of binaries embedded in a gaseous environment, with disk properties that vary from event to event [140].

Combining many noisy single-event measurements into a joint constraint on the population-level hyperparameters is naturally handled by hierarchical Bayesian inference, which we briefly review below before applying it to our problem.

5.3.1 Hierarchical Bayesian framework

We consider a set of N independent events with data $\mathbf{D} = \{D_1, D_2, \dots, D_N\}$. The full set of parameters is split into the single-event parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, \dots, \theta_N\}$ where θ_i is the set of parameters for the event i , and the population-level hyper-parameters $\boldsymbol{\Lambda}$, the latter characterizing the probability distribution of $\boldsymbol{\theta}$ across the population. The joint posterior then reads

$$p(\boldsymbol{\theta}, \boldsymbol{\Lambda} | \mathbf{D}) \propto p(\mathbf{D} | \boldsymbol{\theta}) p(\boldsymbol{\theta} | \boldsymbol{\Lambda}) p(\boldsymbol{\Lambda}), \quad (5.7)$$

and marginalizing over the single-event parameters yields the posterior on the hyper-parameters.

Directly inferring the full posterior is pretty involved, especially for an high-dimensional posterior. In situations where the posteriors $p(D_i | \theta_i)$ are uncorrelated with each other the hyperposterior can be obtained through importance reweighting [252]:

$$\begin{aligned} p(\mathbf{D} | \boldsymbol{\Lambda}) &= \prod_i \int d\theta_i p(D_i | \theta_i) p(\theta_i | \boldsymbol{\Lambda}) \\ &= \prod_i \int d\theta_i \frac{p(\theta_i | D_i) p(D_i)}{p(\theta_i)} p(\theta_i | \boldsymbol{\Lambda}) \\ &\simeq \prod_i p(D_i) \mathbb{E} \left[\frac{p(\theta_i | \boldsymbol{\Lambda})}{p(\theta_i)} \right] \end{aligned} \quad (5.8)$$

where the index i labels the different events in the catalog, $p(\theta_i)$ are the agnostic priors (independent of the hyperparameters), $p(D_i)$ is the evidence for the single event computed with these agnostic priors, and

$$\mathbb{E}[f(\theta_i)] = \frac{1}{N_s} \sum_{\theta_i \in p(\theta_i | D_i)} f(\theta_i) \quad (5.9)$$

where N_s is the number of samples extracted from the posterior $p(\theta_i|D_i)$.

Because the hyper-parameters control the underlying distribution of events, a naive combination of the single-event likelihoods would overweight high-SNR signals and largely ignore marginally detectable ones. To avoid this bias, the hyper-posterior must be normalized by a selection factor, which measures the fraction of parameter space hosting detectable sources for a given choice of $\mathbf{\Lambda}$ [111, 177, 220]. This correction is essential when inferring the mass or redshift distribution of the population, since the SNR depends strongly on these parameters. In our case, however, we are not interested in the binary distribution itself but only in the deviation from vacuum GR. Since all the effects considered here enter only through a small dephasing with negligible impact on the SNR, the selection function is effectively constant across the hyper-parameter space and can be dropped when computing Bayes factors.

5.3.2 Population constraints on time dependent newton constant

For a time-dependent Newton constant, we can assume that the physical parameter G_1 is common to all sources. As such, the informed prior takes the following form:

$$p(\boldsymbol{\theta}|G_1) = \prod_i^N \delta\left(G_1 - \frac{C_{0,i}\eta_i}{C_G M_i}\right) \quad (5.10)$$

where we use the index i to specify the event in the catalog. Inserting this into Eq. (5.7), the delta function in the informed prior fixes $C_0^i = (C_G M_i / \eta_i) G_1$ for each event, so that each event contributes a Gaussian posterior on G_1 once the vacuum parameters are marginalized over. The hyperposterior on the universal G_1 is then the product of these single-event Gaussians, and is therefore Gaussian by construction with variance

$$\sigma_{G_1} = \left(\sum_i^N \Gamma_{G_1 G_1}^i - \Gamma_{k G_1}^i C_{kj}^i \Gamma_{j G_1}^i \right)^{-1/2}, \quad (5.11)$$

where the Latin indices k, j run over the vacuum parameters, the index i labels the event in the catalog, and C^i is the inverse of the vacuum-only Fisher submatrix for event i . Equivalently, the term in parentheses is the inverse-variance sum $\sum_i 1/\sigma_{G_1,i}^2$, where $\sigma_{G_1,i}$ is the marginal single-event uncertainty on G_1 , obtained by inverting the full per-event Fisher matrix over $(\theta_{\text{vac}}^i, G_1)$ and reading off the (G_1, G_1) diagonal of the resulting covariance matrix.

In Fig. 5.3 we report the constraint on G_1 for different population models and observation times. Since the detection rate for the LS-nod-SN model is $\mathcal{O}(1 \text{ yr}^{-1})$, the total statistical error depends strongly on the specific catalog realization. In the plot we report the average statistical error on G_1 across 19 catalogue realizations for the LS-nod-NS model. For the remaining models, we report only one catalogue realization, as the results do not change significantly among realizations.

In the HS models and in the popIII-d model, the improvement in resolution after the first few years of observation is smooth, indicating that the rate at which new

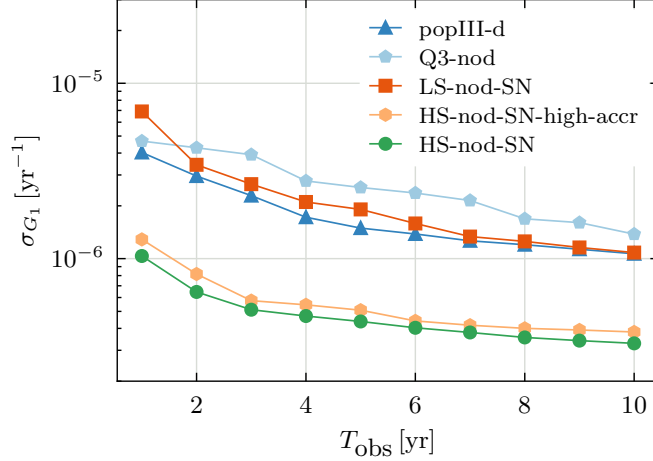


Figure 5.3: Population-level statistical error σ_{G_1} on the time derivative of Newton’s constant, as a function of the LISA mission duration. For the LS-nod-SN we report the mean across the different catalogues.

events contribute appreciably to the hyperparameter constraint is roughly constant. In none of the population models considered does the statistical error on G_1 drop below $\mathcal{O}(10^{-7} \text{ yr}^{-1})$. While this is significantly better than the constraints obtained with LVK, it is a few orders of magnitude weaker than what can be probed with other long-lived sources in the LISA band (SOBHs, EMRIs).

Note that, at linear order in G_1 , the uncertainty σ_{G_1} is independent of the fiducial value of G itself: the leading-order correction to any gravitational-wave observable enters as $\delta O \propto G_1 t$, with no dependence on the absolute normalization of G . This facilitates comparison between different hypotheses.

As the posteriors are Gaussian, it is then straightforward to compute the evidence for the modified-gravity model as

$$p(\mathbf{D}) = \frac{\sqrt{2\pi}\sigma_{G_1}}{V_{G_1}} e^{-\mathcal{L}_{\text{max}}} \prod_i \frac{\sqrt{(2\pi)^d \det(C^i)}}{V_{\text{vac}}} \quad (5.12)$$

where d is the number of vacuum parameters, V_{vac} is the prior volume for the vacuum parameters, V_{G_1} is the prior volume of G_1 , and:

$$\mathcal{L}_{\text{max}} = \sum_i \frac{\tilde{\Gamma}^i G_{1i}^2}{2} - \sigma_{G_1}^2 \frac{(\sum_i \tilde{\Gamma}^i G_{1i})^2}{4} \quad (5.13)$$

with $\tilde{\Gamma} = \Gamma_{G_1 G_1} - \Gamma_{k G_1} C_{kj} \Gamma_{j G_1}$, while G_{1i} is the true value of G_1 for each event in the catalog. We validated the analytical result of Eq. (5.12) against a numerical computation based on Eq. (5.8) and a nested-sampling algorithm [231]; with G_1 having uniform prior in $[-5 \times 10^{-6}, 5 \times 10^{-6}] \text{ yr}^{-1}$.

5.3.3 Population constraints on environmental effects

For an astrophysical population, the value of f_{Edd} is expected to vary from source to source. We assume that the Eddington ratio follows a power-law distribution with fixed upper cutoff $f_{\text{max}} = 10$ and, in line with the AGN duty cycle quoted in Sec. 5.1.2, that only a fraction of binaries are active. We therefore introduce two population-level hyperparameters: the active fraction λ and the power-law slope α ,

$$p(f_{\text{Edd}} | \lambda, \alpha) = \lambda \alpha \frac{f_{\text{Edd}}^{\alpha-1}}{f_{\text{max}}^{\alpha}} + (1 - \lambda) \delta(f_{\text{Edd}}). \quad (5.14)$$

We consider three recovery hypotheses:

- $\mathcal{H}_V$: a vacuum template, with no further population-level assumption;
- $\mathcal{H}_E$: a -4 PN template for the single events and, at the population level, a fraction λ of binaries undergoing viscous migration while the rest merge in vacuum;
- $\mathcal{H}_{\dot{G}}$: a -4 PN template for the single events and, at the population level, a universal time-dependent Newton constant.

Under $\mathcal{H}_E$ we adopt uniform priors on λ in $[0, 1]$ and on α in $[0.01, 2]$; under $\mathcal{H}_{\dot{G}}$ we adopt a uniform prior on G_1 in $[-5 \times 10^{-6}, 5 \times 10^{-6}] \text{ yr}^{-1}$. To compare two recovery hypotheses $\mathcal{H}_A$ and $\mathcal{H}_B$ we compute the Bayes factors. For $\mathcal{H}_V$ and $\mathcal{H}_{\dot{G}}$ the evidences can be computed analytically, giving

$$\mathcal{B}_V^{\dot{G}} = \exp \left(\frac{\sigma_{G_1}^2 (\sum_i \tilde{\Gamma}^i G_{1i})^2}{4} \right) \frac{\sqrt{2\pi} \sigma_{G_1}}{V_{G_1}}. \quad (5.15)$$

For $\mathcal{H}_E$, instead, the evidence does not admit a closed-form expression and is evaluated numerically with the nested-sampling algorithm implemented in `dynesty` [235].

To speed up the likelihood evaluation under the environmental hypothesis we retain only events with $\sigma_{f_{\text{Edd}}} < 20$. The discarded events contribute the same multiplicative factor to the evidence of both $\mathcal{H}_V$ and $\mathcal{H}_E$ within our priors, and therefore drop out of the Bayes factor. After this cut, no catalog retains usable events in the pure-accretion case, so we drop that comparison; for the viscous-migration injections most catalogs are left with fewer than 10 events over 10 yr of observation, and we restrict the analysis to the HS models, where the count is larger.

Figure 5.4 compares the recovered (λ, α) posteriors for three injections: a vacuum injection, and two viscous-migration injections with $\alpha = 0.5$ and active fractions $\lambda = 0.2$ and $\lambda = 0.5$. For the $\lambda = 0.2$ injection the posterior is essentially indistinguishable from the vacuum-injection one. The $\lambda = 0.5$ posterior visibly shifts away from vacuum, and the injected value lies well outside the posterior recovered from the vacuum injection. Even in this most favorable case, however, the posteriors are not informative enough to pin down the astrophysical properties of the population: we recover only an upper limit on the active fraction λ and a 2σ lower limit on the slope α , and all posteriors are fully

compatible with vacuum. Consistent with this, Table 5.3 — which reports the pairwise base-10 logarithms of the Bayes factors among $\mathcal{H}_V$, $\mathcal{H}_E$ and $\mathcal{H}_{\dot{G}}$ for the six injections considered — favors $\mathcal{H}_V$ over $\mathcal{H}_E$ for both environmental injections. Increasing the active fraction from $\lambda = 0.2$ to $\lambda = 0.5$ shrinks $|\log_{10} \mathcal{B}_V^E|$ as expected, but the additional prior volume associated with the environmental hyperparameters still acts as an Occam’s-razor penalty that the data are not informative enough to overcome.

Table 5.3: Base-10 logarithms of Bayes factors between the vacuum ($\mathcal{H}_V$), environmental ($\mathcal{H}_E$) and modified-gravity ($\mathcal{H}_{\dot{G}}$) hypotheses, for the six representative injections discussed in the text. Positive values favor the numerator hypothesis.

injection	$\log_{10} \mathcal{B}_V^E$	$\log_{10} \mathcal{B}_G^E$	$\log_{10} \mathcal{B}_V^{\dot{G}}$
Migration ($\lambda = 0.5$)	−0.46	0.07	−0.52
Migration ($\lambda = 0.2$)	−1.71	0.04	−1.75
Modified GR ($G_1 = 2 \times 10^{-6} \text{ yr}^{-1}$)	−1.79	−2.77	0.97
Modified GR ($G_1 = 10^{-6} \text{ yr}^{-1}$)	−1.78	−0.63	−1.15
Modified GR ($G_1 = 10^{-12} \text{ yr}^{-1}$)	−1.73	0.13	−1.85
Vacuum	−1.73	0.13	−1.84

We repeat the analysis for modified-GR injections; the posteriors are shown in Fig. 5.5. Since the statistical error on G_1 is $\mathcal{O}(10^{-7} \text{ yr}^{-1})$, an injection compatible with the Solar-System bound, $G_1 = 10^{-12} \text{ yr}^{-1}$, is indistinguishable from pure GR. We therefore consider larger injections that still respect the LVK bound on G_1 . From Eq. (5.15), the modified-gravity Bayes factor scales as $\exp(G_1^2/\sigma_{G_1}^2)$, so doubling the injected G_1 from 10^{-6} yr^{-1} to $2 \times 10^{-6} \text{ yr}^{-1}$ raises $\mathcal{B}_G^V$ by roughly two orders of magnitude. Only the $G_1 = 2 \times 10^{-6} \text{ yr}^{-1}$ case yields a clear preference for the modified-gravity hypothesis.

We now compare $\mathcal{H}_E$ and $\mathcal{H}_{\dot{G}}$ directly. When the data are well described by vacuum, the marginalized evidence under $\mathcal{H}_E$ is dominated by the region near $\lambda = 0$, where the environmental model reduces to $\mathcal{H}_V$; the integral therefore equals the vacuum evidence multiplied by an Occam factor (the ratio of the posterior width near $\lambda = 0$ to the (λ, α) prior volume). The same happens for $\mathcal{H}_{\dot{G}}$ near $G_1 = 0$. For our priors the two Occam factors are numerically similar, so the Bayes factor between $\mathcal{H}_E$ and $\mathcal{H}_{\dot{G}}$ stays close to unity — as confirmed by the $\log_{10} \mathcal{B}_G^E$ column of Table 5.3, which gives $|\log_{10} \mathcal{B}_G^E| \leq 0.13$ for the four vacuum-compatible injections. The approximation breaks for the two strong modified-gravity injections ($G_1 = 10^{-6}$ and $2 \times 10^{-6} \text{ yr}^{-1}$), where the evidence of $\mathcal{H}_{\dot{G}}$ is genuinely enhanced over vacuum and $\log_{10} \mathcal{B}_G^E$ drops to −0.63 and −2.77, respectively.

As a final cross-check, we fit the environmental hypothesis $\mathcal{H}_E$ to data generated under modified-GR injections. The recovered (λ, α) posteriors are fully consistent with $\lambda = 0$, so the environmental recovery does not falsely claim a viscous-migration signal even when the data contain a real $\dot{G}$ signal. Table 5.3 bears this out from two angles: the $\log_{10} \mathcal{B}_V^E$ column is essentially unchanged across the three modified-GR injections (it stays near −1.75), confirming that the V-vs-E balance does not respond to the true G_1 ; in parallel, $\log_{10} \mathcal{B}_V^{\dot{G}}$ moves from −1.85 at $G_1 = 10^{-12} \text{ yr}^{-1}$ to +0.97 at $G_1 = 2 \times 10^{-6} \text{ yr}^{-1}$, showing

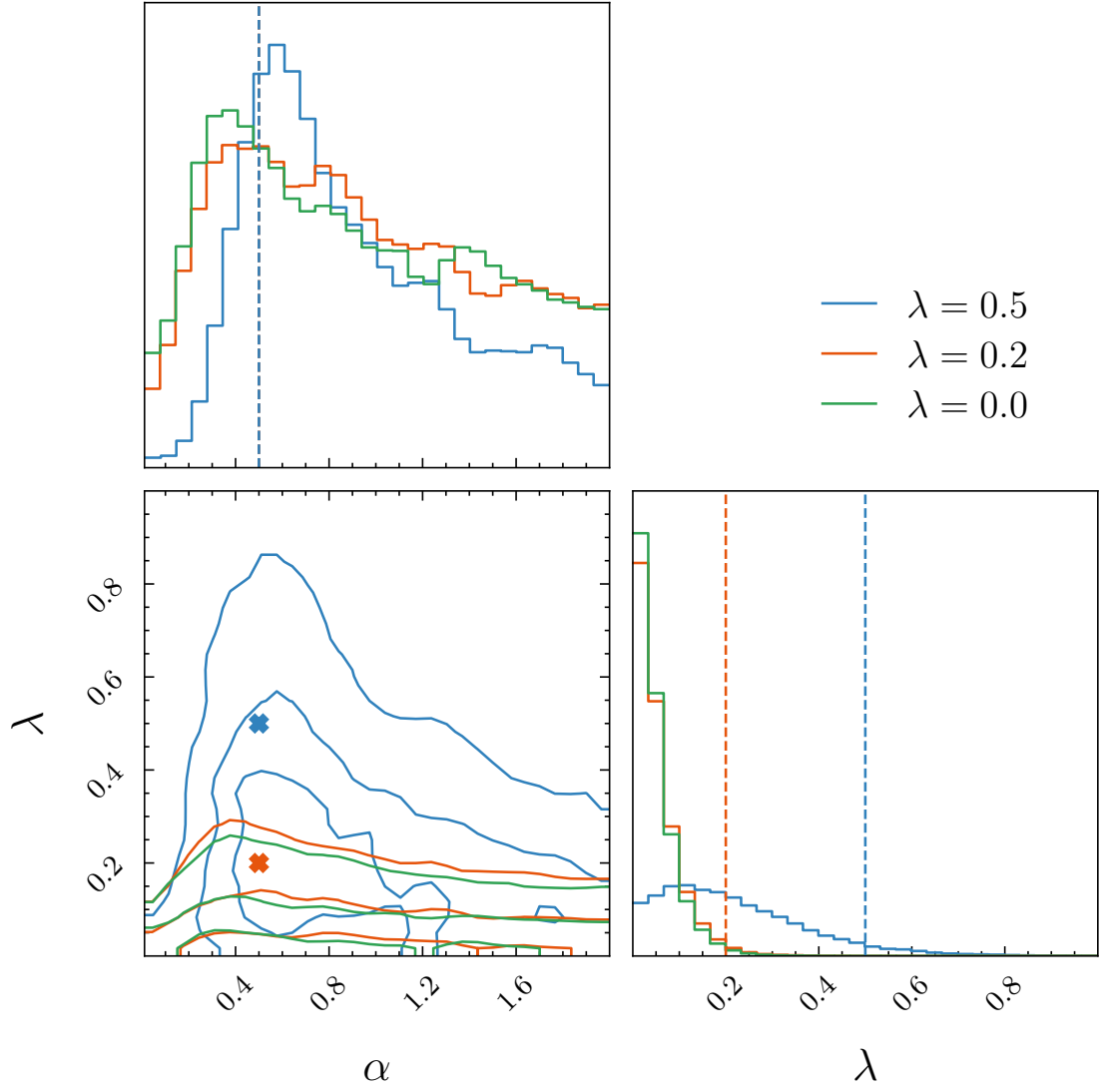


Figure 5.4: Population-level posteriors on the hyperparameters (λ, α) for three injections analyzed under the environmental hypothesis: a vacuum injection (green), and two viscous-migration injections with $\alpha = 0.5$ and active fractions $\lambda = 0.2$ (red) and $\lambda = 0.5$ (blue). In all three cases the recovered posteriors are fully compatible with vacuum.

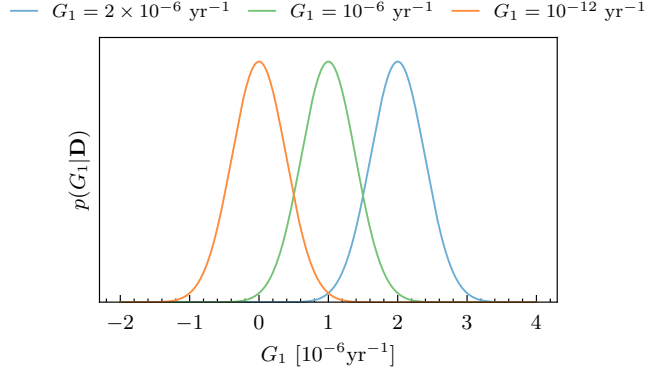


Figure 5.5: Population-level posteriors on G_1 for three modified-GR injections: $G_1 = 2 \times 10^{-6} \text{ yr}^{-1}$, $G_1 = 10^{-6} \text{ yr}^{-1}$, and $G_1 = 10^{-12} \text{ yr}^{-1}$ (in blue, green and orange respectively with the last one being compatible with Solar-System constraints). The Solar-System-compatible case is effectively centered at $G_1 = 0$.

that the $\dot{G}$ analysis correctly responds to the injection strength.

6 Conclusions and Future Prospects

The work carried out in this thesis has been directed toward the study of environmental effects in the mHz band, which will be probed by the LISA experiment. The main focus of this work has been the study of extreme mass ratio inspirals, and the importance of environmental effects and systematics in this study. We now present the conclusions for each of the separate works then we outline future perspectives.

6.1 Conclusion

6.1.1 Stochastic Torques in EMRIs

We investigated the robustness of simple analytic models of gas torques to perform parameter estimation for GW signals from EMRIs/IMRIs in accretion disks, as detectable by LISA. Using hydrodynamic simulations of accretion disk torques as injections, we investigated the effect of stochastic torque variability on the inference of both binary and environmental parameters using analytic templates. Our parameter estimation studies demonstrate that, while simple torque models without stochastic variability allow for accurate recovery of binary parameters, they can bias the inferred environmental effects and accretion disk properties.

For moderate stochastic time variability, the posteriors for the torque amplitude are approximately centered on the averaged value of the simulated torque. However, for significant stochastic variability, the estimated amplitude may effectively reduce to zero, thus obscuring the presence of torques. This result underscores the subtleties that one has to face when inferring accretion disk properties from LISA observations using simplified torque models. While these models are sufficient for binary parameter extraction, a more sophisticated approach may be needed to reliably constrain the properties of the astrophysical environment.

Indeed, even in the presence of unbiased estimation of the torque amplitude, we have shown that converting the latter to estimates of the disk's physical properties is subject to additional uncertainties, due to the possible mismatch between the disk model used in the analysis and the one describing the real astrophysical system (or, in our case, the simulations). Moreover, even when the bias in the recovered torque amplitudes is absent, the stochastic components of the torques can enhance the residual power left in the data. Although these residuals are small (with $\text{SNR} \lesssim 1$) and should therefore not pose a problem for the LISA global fit, their combined effect if thousands of EMRI sources are present in the data [25] may warrant further investigation.

6.1.2 Consistency Test for Extreme Mass Ratio Inspirals

We have developed a self-consistency check solely based on vacuum GR EMRI waveforms. Our approach tests inconsistencies/biases in the posteriors as inferred from different portions of the signal, which can reveal missing physics in the templates, due e.g. to environmental effects and/or possible beyond-GR effects. For an injected EMRI signal affected by disk-driven migration, we showed that the parameter posteriors are significantly inconsistent, when inferred from different portions of the signal, for observations of more than two years.

Crucially, our model relies on the successful subtraction of louder sources in the LISA global fit. Indeed, residual power from imperfect subtraction of overlapping sources (e.g. Galactic binaries [152]) may be flagged by our test. In the Supplemental Material, we consider a slight mismodeling of the galactic noise, and find that it does not result in significant inconsistencies across observation times. We stress that our method can also be viewed as a consistency test for the “goodness” of the LISA global fit.

6.1.3 Massive Black Hole Binaries, Environment vs Modified gravity

We investigated whether the astrophysical environment of LISA MBHBs could leave a detectable imprint at the population level, and whether it could mimic—and thereby bias—LISA’s tests of GR. We focused on three ~ 4 PN deviations from vacuum templates: gas accretion onto the components, viscous migration in a circumbinary disk, and a secular drift of Newton’s constant. All three are formally degenerate for a single source, but the degeneracy is broken at the population level: a $\dot{G}$ drift is universal across the catalog, whereas the environmental effects act only on a fraction of binaries and with a magnitude that varies from event to event.

Across the five population models considered, the most informative sources are low-mass MBHBs at low redshift, which are rare irrespective of the seeding model. Even at the optimistic LISA 10 yr mission duration, no event is individually sensitive to accretion at realistic Eddington ratios: at $f_{\text{Edd}} = 1$, vacuum-template recovery shifts no event’s chirp mass by more than 10% of the statistical error. Viscous migration is a more promising channel because of its stronger η^{-2} mass-ratio scaling: up to ~ 8 events per year in the HS models admit a single-event measurement with $\sigma_{f_{\text{Edd}}} < 10$ at fiducial $\xi = 1$, and at the optimistic thin-disk value $\xi \sim 10$ a few HS-catalog events per year can exhibit a vacuum-template chirp-mass bias at or above the statistical error. This handful of relatively noisy detections is, however, not enough to drive a population-level preference for an environmental signature.

At the population level we compared three recovery hypotheses: vacuum, a hierarchical environmental model parameterized by the active fraction λ and the power-law slope α of the Eddington-ratio distribution, and a universal time-dependent Newton constant. For $\dot{G}$, the population-level statistical error settles at $\sigma_{G_1} \sim 10^{-7} \text{ yr}^{-1}$ across all the catalogs considered—significantly better than current bounds from ground-based GW observations, but several orders of magnitude weaker than the Solar-System limit and than what can be reached with longer-lived LISA sources such as EMRIs and SOBHBs.

For the environmental hypothesis, all the injections considered—including a viscous-migration injection with active fraction $\lambda = 0.5$ and a super-Eddington Eddington-ratio tail ($f_{\text{max}} = 10$)—yield population posteriors that are fully compatible with vacuum, although they do produce a non-trivial upper limit on λ and a 2σ lower bound on α . A marked Bayes-factor preference for the modified-gravity hypothesis emerges only for injected G_1 several orders of magnitude above the Solar-System limit; such values are not yet excluded in the strong-field, dynamical regime probed by ground-based GW observations, but no realistic theoretical motivation singles them out.

6.2 Future Perspective

The field of environmental effects is currently expanding, with a plethora of topics that could still be tackled, providing interesting results for the development of future experiments. A significant open question, tangentially touched upon and discussed in this work, is how to handle environmental effects within the global fit, or as a post-processing step. In the LISA data stream, most events will happen concurrently, and the events arguably most interesting for the study of environmental effects will likely span a long time, perhaps even the whole duration of the mission. It is therefore necessary to understand how to extract this information at the level of the global fit results. An important open question is whether the introduction of additional terms in the gravitational wave templates is necessary for sources such as EMRIs, since this can mitigate the bias arising from the use of an incomplete gravitational wave template. In this direction, we are working on a project, currently in preparation, on the parameter estimation of environmental effects using agnostic templates tested against more refined injections. Specifically, we focus on a relativistic Lindblad torque for which we know the environmental torque to change sign throughout the inspiral [102, 131, 132]. We are testing the applicability of agnostic templates, and whether they can be used during the global fit, assessing the importance of the signal lost in the datastream and how to apply post-processing to recover the actual physical properties of the astrophysical environment.

One important caveat is in order: most of our analysis is restricted to quasi-circular inspirals, while recent works [245, 283] have shown that environmental dephasing can be substantially amplified in eccentric orbits. An untouched avenue is the interplay between eccentricity and the physical environment; while this topic has been discussed in some works and models, no work has assessed in a rigorous way their importance. Specifically a source class for which the presence of eccentricity and environment is expected is that of SOBHBs in the LISA band. We intend to use and/or extend current environmental interaction templates as injections and assess systematically our ability to infer their properties across the parameter space with a strong focus on possible biases induced by the use of wrong templates injections and assess systematically our ability to infer their properties across the parameter space with a strong focus on possible biases induced by the use of wrong templates.

Acknowledgments

When looking back at the past four years I am surprised by the amount of amazing people I have met and helped me in this troubling time. I wish to thank all of them for the support and joy they brought me. A first and huge thank goes to Enrico, who has guided me in my research and whose patience I put to the test in many occasions through my awful writing skills and my stubbornness. I thank all the amazing people in the group: Marcelo, Jam, Adrien, Lodovico, Clemente, Ramiro, Rahul, Massimo, Rohit, Chantal, Emma, Lidia, Ran, Thomas, Gianmarco, Antonia, Kallol and Erika, with which I would have loved to spend more time. I thank Laura and Lorenzo, for helping and support they have given me.

I deeply thank all the other friends I have met in these years, Claudio, Francesco, Francesco, Carlos, Anant, Martina, Emma, Leonardo and Francescopaolo and all the other which I may have forgotten. I deeply thank my parents, my sister for the support given from much before the start of this work and throughout all these years spent studying. The biggest thank goes to Chiara, to which this work is dedicated, thanks for staying at my side even through the distance, none of the words that I can write could express the depth of recognition that I have for you enduring my stupid jokes, and my endless discussion about EMRIs and LISA.

To conclude,

So long, and thanks for all the trips

Bibliography

- [1] J. Aasi et al. (LIGO Scientific), “Advanced LIGO”, *Class. Quant. Grav.* **32**, 074001 (2015).
- [2] A. G. Abac et al. (LIGO Scientific, VIRGO, KAGRA), “GWTC-4.0: Tests of General Relativity. II. Parameterized Tests”, (2026).
- [3] A. G. Abac et al. (LIGO Scientific, VIRGO, KAGRA), “GWTC-5.0: Tests of General Relativity”, (2026).
- [4] A. G. Abac et al. (LIGO Scientific, Virgo, KAGRA), “Open Data from LIGO, Virgo, and KAGRA through the First Part of the Fourth Observing Run”, *Astrophys. J.* **1004**, 232 (2026).
- [5] A. Abac et al. (ET), “The Science of the Einstein Telescope”, *JCAP* **03**, 081 (2026).
- [6] N. Abac et al. (LIGO Scientific, VIRGO, KAGRA), “Open Data from LIGO, Virgo, and KAGRA through the Second Part of the Fourth Observing Run”, (2026).
- [7] B. P. Abbott et al. (LIGO Scientific, Virgo), “Observation of Gravitational Waves from a Binary Black Hole Merger”, *Phys. Rev. Lett.* **116**, 061102 (2016).
- [8] B. P. Abbott et al. (LIGO Scientific, Virgo), “GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral”, *Phys. Rev. Lett.* **119**, 161101 (2017).
- [9] R. Abbott et al. (LIGO Scientific, KAGRA, VIRGO), “Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences”, *Astrophys. J. Lett.* **915**, L5 (2021).
- [10] F. Acernese et al. (VIRGO), “Advanced Virgo: a second-generation interferometric gravitational wave detector”, *Class. Quant. Grav.* **32**, 024001 (2015).
- [11] G. Agazie et al. (NANOGrav), “The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background”, *Astrophys. J. Lett.* **951**, L8 (2023).
- [12] T. Akutsu et al. (KAGRA), “KAGRA: 2.5 Generation Interferometric Gravitational Wave Detector”, *Nature Astron.* **3**, 35–40 (2019).
- [13] S. Albanesi, R. Gamba, S. Bernuzzi, J. Fontbuté, A. Gonzalez, and A. Nagar, “Effective-one-body modeling for generic compact binaries with arbitrary orbits”, *Phys. Rev. D* **112**, L121503 (2025).

- [14] A. Albertini, R. Gamba, A. Nagar, and S. Bernuzzi, “Effective-one-body waveforms for extreme-mass-ratio binaries: Consistency with second-order gravitational self-force quasicircular results and extension to nonprecessing spins and eccentricity”, *Phys. Rev. D* **109**, 044022 (2024).
- [15] P. Amaro-Seoane, “Extremely large mass-ratio inspirals”, *Phys. Rev. D* **99**, 123025 (2019).
- [16] P. Amaro-Seoane et al., “Laser Interferometer Space Antenna”, *arXiv e-prints*, [arXiv:1702.00786](#), [arXiv:1702.00786](#) (2017).
- [17] D. Anderson, N. Yunes, and E. Barausse, “Effect of cosmological evolution on Solar System constraints and on the scalarization of neutron stars in massless scalar-tensor theories”, *Phys. Rev. D* **94**, 104064 (2016).
- [18] A. Antonelli, O. Burke, and J. R. Gair, “Noisy neighbours: inference biases from overlapping gravitational-wave signals”, *Monthly Notices of the Royal Astronomical Society* **507**, 5069–5086 (2021).
- [19] J. Antoniadis et al. (EPTA, InPTA:), “The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals”, *Astron. Astrophys.* **678**, A50 (2023).
- [20] J. Antoniadis et al. (EPTA, InPTA), “The second data release from the European Pulsar Timing Array - IV. Implications for massive black holes, dark matter, and the early Universe”, *Astron. Astrophys.* **685**, A94 (2024).
- [21] F. Antonini, E. Barausse, and J. Silk, “The Coevolution of Nuclear Star Clusters, Massive Black Holes, and their Host Galaxies”, *Astrophys. J.* **812**, 72 (2015).
- [22] Y. Aoyama and X.-N. Bai, “Three-dimensional global simulations of type-ii planet–disk interaction with a magnetized disk wind. i. magnetic flux concentration and gap properties”, *The Astrophysical Journal* **946**, 5 (2023).
- [23] M. Armano et al., “Beyond the required lisa free-fall performance: new lisa pathfinder results down to 20 μHz ”, *Phys. Rev. Lett.* **120**, 061101 (2018).
- [24] S. Babak, C. Caprini, D. G. Figueroa, N. Karnesis, P. Marcoccia, G. Nardini, M. Pieroni, A. Ricciardone, A. Sesana, and J. Torrado, “Stochastic gravitational wave background from stellar origin binary black holes in LISA”, *JCAP* **08**, 034 (2023).
- [25] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. L. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, “Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals”, *Phys. Rev. D* **95**, 103012 (2017).
- [26] E. Baird, S. Fairhurst, M. Hannam, and P. Murphy, “Degeneracy between mass and spin in black-hole-binary waveforms”, *Phys. Rev. D* **87**, 024035 (2013).
- [27] C. Bambi, M. Giannotti, and F. L. Villante, “Response of primordial abundances to a general modification of G N and/or of the early universe expansion rate”, *Physical Review D* **71**, 123524 (2005).

- [28] L. Barack and C. Cutler, “LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy”, *Physical Review D* **69**, 082005 (2004).
- [29] E. Barausse, “The evolution of massive black holes and their spins in their galactic hosts”, *Mon. Not. Roy. Astron. Soc.* **423**, 2533–2557 (2012).
- [30] E. Barausse, *Black holes across cosmic history*, (2023) https://people.sissa.it/~barausse/catalogs/models_LISA_after_PTAs.
- [31] E. Barausse, V. Cardoso, and P. Pani, “Can environmental effects spoil precision gravitational-wave astrophysics?”, *Phys. Rev. D* **89**, 104059 (2014).
- [32] E. Barausse, K. Dey, M. Crisostomi, A. Panayada, S. Marsat, and S. Basak, “Implications of the pulsar timing array detections for massive black hole mergers in the LISA band”, *Physical Review D* **108**, arXiv:2307.12245 [astro-ph], 103034 (2023).
- [33] E. Barausse, I. Dvorkin, M. Tremmel, M. Volonteri, and M. Bonetti, “Massive Black Hole Merger Rates: The Effect of Kiloparsec Separation Wandering and Supernova Feedback”, *Astrophys. J.* **904**, 16 (2020).
- [34] E. Barausse and L. Rezzolla, “The Influence of the hydrodynamic drag from an accretion torus on extreme mass-ratio inspirals”, *Phys. Rev. D* **77**, 104027 (2008).
- [35] E. Barausse, N. Yunes, and K. Chamberlain, “Theory-Agnostic Constraints on Black-Hole Dipole Radiation with Multiband Gravitational-Wave Astrophysics”, *Phys. Rev. Lett.* **116**, 241104 (2016).
- [36] J. M. Bardeen and J. A. Petterson, “The Lense-Thirring Effect and Accretion Disks around Kerr Black Holes”, *The Astrophysical Journal Letters* **195**, L65 (1975).
- [37] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, “Rotating Black Holes: Locally Nonrotating Frames, Energy Extraction, and Scalar Synchrotron Radiation”, en, *The Astrophysical Journal* **178**, 347 (1972).
- [38] J. E. Barnes and L. Hernquist, “Transformations of Galaxies. 2. Gasdynamics in Merging Disk Galaxies”, *Astrophys. J.* **471**, 115 (1996).
- [39] S. Barsanti, O. Burke, A. Maselli, T. P. Sotiriou, A. Spiers, and N. Warburton, “The significance of first post-adiabatic contributions for scalar charge measurements with intermediate and extreme mass ratio inspirals”, (2026).
- [40] C. Baruteau and D. N. C. Lin, “Protoplanetary migration in turbulent isothermal disks”, *The Astrophysical Journal* **709**, 759 (2010).
- [41] C. Baruteau, A. Crida, S.-J. Paardekooper, F. Masset, J. Guilet, B. Bitsch, R. P. Nelson, W. Kley, and J. C. B. Papaloizou, “Planet-Disc Interactions and Early Evolution of Planetary Systems”, in *Protostars and planets vi* (University of Arizona Press, 2014).

- [42] C. Baruteau and F. Masset, “Recent developments in planet migration theory”, in *Tides in astronomy and astrophysics* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013), pp. 201–253.
- [43] D. R. Becker and S. A. Hughes, “Transition from adiabatic inspiral to plunge for eccentric binaries”, *Phys. Rev. D* **111**, 064003 (2025).
- [44] D. R. Becker, S. A. Hughes, and G. Khanna, “Gravitational waves from the late inspiral, transition, and plunge of small-mass-ratio eccentric binaries”, *Phys. Rev. D* **113**, 084046 (2026).
- [45] A. Begoni, W. Del Pozzo, M. Pegorin, J. Pomper, and A. Ricciardone, “Tests of General Relativity with Einstein Telescope”, (2025).
- [46] E. Berti et al., “Black hole spectroscopy: from theory to experiment”, *Class. Quant. Grav.* **43**, edited by E. Berti, V. Cardoso, and G. Carullo, 123001 (2026).
- [47] U. Bhargwaj, J. Alvey, B. K. Miller, S. Nissanke, and C. Weniger, “Sequential simulation-based inference for gravitational wave signals”, *Phys. Rev. D* **108**, 042004 (2023).
- [48] L. Blanchet, “Post-Newtonian Theory for Gravitational Waves”, *Living Rev. Rel.* **17**, 2 (2014).
- [49] L. Blanchet, G. Faye, E. Serrail, and D. Trestini, “Gravitational radiation reaction for compact binary systems at the fourth-and-a-half post-Newtonian order in harmonic coordinates”, *Class. Quant. Grav.* **43**, 105009 (2026).
- [50] L. Blanchet, G. Faye, and D. Trestini, “Gravitational radiation reaction for compact binary systems at the fourth-and-a-half post-Newtonian order”, *Class. Quant. Grav.* **42**, 065015 (2025).
- [51] M. Bonetti, A. Sesana, E. Barausse, and F. Haardt, “Post-Newtonian evolution of massive black hole triplets in galactic nuclei – III. A robust lower limit to the nHz stochastic background of gravitational waves”, *Mon. Not. Roy. Astron. Soc.* **477**, 2599–2612 (2018).
- [52] M. Bonetti, A. Sesana, F. Haardt, E. Barausse, and M. Colpi, “Post-Newtonian evolution of massive black hole triplets in galactic nuclei – IV. Implications for LISA”, *Mon. Not. Roy. Astron. Soc.* **486**, 4044–4060 (2019).
- [53] K. Breivik, C. M. F. Mingarelli, and S. L. Larson, “Constraining Galactic Structure with the LISA White Dwarf Foreground”, *Astrophys. J.* **901**, 4 (2020).
- [54] K. Breivik, C. L. Rodriguez, S. L. Larson, V. Kalogera, and F. A. Rasio, “Distinguishing Between Formation Channels for Binary Black Holes with LISA”, *Astrophys. J. Lett.* **830**, L18 (2016).
- [55] F. S. Broekgaarden, S. Banagiri, and E. Payne, “Visualizing the Number of Existing and Future Gravitational-wave Detections from Merging Double Compact Objects”, *Astrophys. J.* **969**, 108 (2024).

- [56] A. Buonanno and T. Damour, “Effective one-body approach to general relativistic two-body dynamics”, *Phys. Rev. D* **59**, 084006 (1999).
- [57] O. Burke, S. Marsat, J. R. Gair, and M. L. Katz, “Addressing data gaps and assessing noise mismodeling in LISA”, *Phys. Rev. D* **111**, 124053 (2025).
- [58] O. Burke, G. A. Piovano, N. Warburton, P. Lynch, L. Speri, C. Kavanagh, B. Wardell, A. Pound, L. Durkan, and J. Miller, “Assessing the importance of first postadiabatic terms for small-mass-ratio binaries”, *Phys. Rev. D* **109**, 124048 (2024).
- [59] R. Buscicchio, F. Pozzoli, D. Chirico, and A. Sesana, “The first year of LISA Galactic foreground”, (2025).
- [60] R. Caiozzo, G. Bertone, P. Ullio, R. Vicente, B. J. Kavanagh, and D. Gaggero, “Dark matter mounds from the collapse of supermassive stars: A general-relativistic analysis”, *Phys. Rev. D* **113**, 063052 (2026).
- [61] B. Calderhead and M. Girolami, “Estimating bayes factors via thermodynamic integration and population mcmc”, *Computational Statistics and Data Analysis* **53**, 4028–4045 (2009).
- [62] G. Caneva Santoro, S. Roy, R. Vicente, M. Haney, O. J. Piccinni, W. Del Pozzo, and M. Martinez, “First Constraints on Compact Binary Environments from LIGO-Virgo Data”, *Phys. Rev. Lett.* **132**, 251401 (2024).
- [63] L. Capuano, M. Vaglio, R. S. Chandramouli, C. L. Pitte, A. Kuntz, and E. Barausse, “Systematic bias in LISA ringdown analysis due to waveform inaccuracy”, *Phys. Rev. D* **112**, 104031 (2025).
- [64] A. Caputo, L. Sberna, A. Toubiana, S. Babak, E. Barausse, S. Marsat, and P. Pani, “Gravitational-wave detection and parameter estimation for accreting black-hole binaries and their electromagnetic counterpart”, *Astrophys. J.* **892**, 90 (2020).
- [65] A. Caputo, L. Sberna, A. Toubiana, S. Babak, E. Barausse, S. Marsat, and P. Pani, “Gravitational-wave detection and parameter estimation for accreting black-hole binaries and their electromagnetic counterpart”, *The Astrophysical Journal* **892**, arXiv:2001.03620 [astro-ph, physics:gr-qc], 90 (2020).
- [66] R. S. Chandramouli, K. Prokup, E. Berti, and N. Yunes, “Systematic biases due to waveform mismodeling in parametrized post-Einsteinian tests of general relativity: The impact of neglecting spin precession and higher modes”, *Phys. Rev. D* **111**, 044026 (2025).
- [67] S. Chandrasekhar, “Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction.”, *The Astrophysics Journal* **97**, 255 (1943).
- [68] C. E. A. Chapman-Bird et al., “Efficient waveforms for asymmetric-mass eccentric equatorial inspirals into rapidly spinning black holes”, *Phys. Rev. D* **112**, 104023 (2025).

- [69] C. E. A. Chapman-Bird et al., “The Fast and the Frame-Dragging: Efficient waveforms for asymmetric-mass eccentric equatorial inspirals into rapidly-spinning black holes”, arXiv preprint (2025).
- [70] K. Chatziioannou, A. Klein, N. Yunes, and N. Cornish, “Constructing Gravitational Waves from Generic Spin-Precessing Compact Binary Inspirals”, *Phys. Rev. D* **95**, 104004 (2017).
- [71] R. Chen, R. S. Chandramouli, F. Pozzoli, R. Buscicchio, and E. Barausse, “Environmental effects in the LISA stochastic signal from stellar-mass black hole binaries”, *Phys. Rev. D* **112**, 084053 (2025).
- [72] A. J. K. Chua, M. L. Katz, N. Warburton, and S. A. Hughes, “Rapid generation of fully relativistic extreme-mass-ratio-inspiral waveform templates for LISA data analysis”, *Phys. Rev. Lett.* **126**, 051102 (2021).
- [73] A. J. K. Chua, C. J. Moore, and J. R. Gair, “Augmented kludge waveforms for detecting extreme-mass-ratio inspirals”, *Phys. Rev. D* **96**, 044005 (2017).
- [74] P. S. Cole, G. Bertone, A. Coogan, D. Gaggero, T. Karydas, B. J. Kavanagh, T. F. M. Spieksma, and G. M. Tomaselli, “Distinguishing environmental effects on binary black hole gravitational waveforms”, *Nature Astron.* **7**, 943–950 (2023).
- [75] M. Colpi et al., “LISA Definition Study Report”, en, *arXiv.org* (2024).
- [76] L. Copparoni and E. Barausse, “Environmental effects vs. modified gravity in the LISA massive black hole binary population”, (2026).
- [77] L. Copparoni, E. Barausse, L. Speri, L. Sberna, and A. Derdzinski, “Implications of stochastic gas torques for asymmetric binaries in the LISA band”, *Phys. Rev. D* **111**, 104079 (2025).
- [78] L. Copparoni, R. S. Chandramouli, and E. Barausse, “When Vacuum Breaks: A Self-Consistency Test for Astrophysical Environments in Extreme Mass Ratio Inspirals”, *Phys. Rev. Lett.* **137**, 021405 (2026).
- [79] N. Cornish and T. Robson, “Galactic binary science with the new LISA design”, *Journal of Physics: Conference Series* **840**, 012024 (2017).
- [80] N. J. Cornish and J. Crowder, “LISA data analysis using MCMC methods”, *Phys. Rev. D* **72**, 043005 (2005).
- [81] N. J. Cornish and R. W. Hellings, “The Effects of orbital motion on LISA time delay interferometry”, *Class. Quant. Grav.* **20**, 4851–4860 (2003).
- [82] A. Crida, A. Morbidelli, and F. Masset, “On the width and shape of gaps in protoplanetary disks”, *Icarus* **181**, 587–604 (2006).
- [83] C. Cutler, “Angular resolution of the LISA gravitational wave detector”, en, *Physical Review D* **57**, 7089–7102 (1998).
- [84] C. Cutler and É. E. Flanagan, “Gravitational waves from merging compact binaries: How accurately can one extract the binary’s parameters from the inspiral waveform?”, en, *Physical Review D* **49**, 2658–2697 (1994).

- [85] C. Cutler and É. E. Flanagan, “Gravitational waves from merging compact binaries: How accurately can one extract the binary’s parameters from the inspiral waveform?”, en, *Physical Review D* **49**, 2658–2697 (1994).
- [86] C. Cutler and M. Vallisneri, “Lisa detections of massive black hole inspirals: parameter extraction errors due to inaccurate template waveforms”, *Phys. Rev. D* **76**, 104018 (2007).
- [87] T. Damour and K. Nordtvedt, “Tensor - scalar cosmological models and their relaxation toward general relativity”, *Phys. Rev. D* **48**, 3436–3450 (1993).
- [88] T. Damour and K. Nordtvedt, “General relativity as a cosmological attractor of tensor scalar theories”, *Phys. Rev. Lett.* **70**, 2217–2219 (1993).
- [89] M. Dax, S. R. Green, J. Gair, N. Gupte, M. Pürrer, V. Raymond, J. Wildberger, J. H. Macke, A. Buonanno, and B. Schölkopf, “Real-time inference for binary neutron star mergers using machine learning”, *Nature* **639**, 49–53 (2025).
- [90] M. Dax, S. R. Green, J. Gair, J. H. Macke, A. Buonanno, and B. Schölkopf, “Real-Time Gravitational Wave Science with Neural Posterior Estimation”, *Phys. Rev. Lett.* **127**, 241103 (2021).
- [91] F. De Santi, A. Santini, A. Toubiana, N. Karnesis, and D. Gerosa, “Inferring the population properties of Galactic binaries from LISA’s stochastic foreground”, *Phys. Rev. D* **113**, 123006 (2026).
- [92] S. Deng, S. Babak, M. Le Jeune, S. Marsat, É. Plagnol, and A. Sartirana, “Modular global-fit pipeline for LISA data analysis”, *Phys. Rev. D* **111**, 103014 (2025).
- [93] A. Derdzinski, D. D’Orazio, P. Duffell, Z. Haiman, and A. MacFadyen, “Evolution of gas disc-embedded intermediate mass ratio inspirals in the *LISA* band”, en, *Monthly Notices of the Royal Astronomical Society* **501**, 3540–3557 (2021).
- [94] A. M. Derdzinski, D. D’Orazio, P. Duffell, Z. Haiman, and A. MacFadyen, “Probing gas disc physics with LISA: simulations of an intermediate mass ratio inspiral in an accretion disc”, en, *Monthly Notices of the Royal Astronomical Society* **486**, 2754–2765 (2019).
- [95] A. Derdzinski and L. Mayer, “In-situ extreme mass ratio inspirals via sub-parsec formation and migration of stars in thin, gravitationally unstable AGN discs”, en, *Monthly Notices of the Royal Astronomical Society* **521**, 4522–4543 (2023).
- [96] A. J. Dittmann and G. Ryan, “A survey of disc thickness and viscosity in circumbinary accretion: binary evolution, variability, and disc morphology”, *Monthly Notices of the Royal Astronomical Society* **513**, 6158–6176 (2022).
- [97] S. Drasco and S. A. Hughes, “Gravitational wave snapshots of generic extreme mass ratio inspirals”, en, *Physical Review D* **73**, 024027 (2006).
- [98] P. C. Duffell, D. D’Orazio, A. Derdzinski, Z. Haiman, A. MacFadyen, A. L. Rosen, and J. Zrake, “Circumbinary disks: Accretion and torque as a function of mass ratio and disk viscosity”, *The Astrophysical Journal* **901**, 25 (2020).

- [99] P. C. Duffell, Z. Haiman, A. I. MacFadyen, D. J. D’Orazio, and B. D. Farris, “The Migration of Gap-Opening Planets is not Locked to Viscous Disk Evolution”, en, *The Astrophysical Journal* **792**, L10 (2014).
- [100] P. C. Duffell and A. I. MacFadyen, “Gap opening by extremely low-mass planets in a viscous disk”, *The Astrophysical Journal* **769**, 41 (2013).
- [101] F. Duque, S. Kejriwal, L. Sberna, L. Speri, and J. Gair, “Constraining accretion physics with gravitational waves from eccentric extreme-mass-ratio inspirals”, *Phys. Rev. D* **111**, 084006 (2025).
- [102] F. Duque, L. Sberna, A. Spiers, and R. Vicente, “Extreme-mass-ratio inspirals in relativistic accretion discs”, *Phys. Rev. D* **113**, 084028 (2026).
- [103] C. Dyson, J. Redondo-Yuste, M. van de Meent, and V. Cardoso, “Relativistic aerodynamics of spinning black holes”, en, (2024).
- [104] C. Dyson, T. F. M. Spieksma, R. Brito, M. van de Meent, and S. Dolan, “Environmental Effects in Extreme-Mass-Ratio Inspirals: Perturbations to the Environment in Kerr Spacetimes”, *Phys. Rev. Lett.* **134**, 211403 (2025).
- [105] A. S. Eddington, “The propagation of gravitational waves”, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **102**, 268–282 (1922).
- [106] A. Einstein, “The foundation of the general theory of relativity.”, *Annalen Phys.* **49**, edited by J.-P. Hsu and D. Fine, 769–822 (1916).
- [107] H. Estellés, A. Ramos-Buades, S. Husa, C. García-Quirós, M. Colleoni, L. Haegel, and R. Jaume, “Phenomenological time domain model for dominant quadrupole gravitational wave signal of coalescing binary black holes”, *Phys. Rev. D* **103**, 124060 (2021).
- [108] G. Eytan, V. Desjacques, and Y. B. Ginat, “Harmonic-decomposition approach to dynamical friction for eccentric orbits”, [10.48550/arXiv.2509.15632](https://arxiv.org/abs/10.48550/arXiv.2509.15632) (2025).
- [109] F. Fantoccoli, F. Duque, and J. Gair, “Gotta light? Illuminating AGN disks with LISA EMRIs”, (2026).
- [110] R. D. Ferdman, R. van Haasteren, C. G. Bassa, M. Burgay, I. Cognard, A. Corongiu, N. D’Amico, G. Desvignes, J. W. Hessels, G. H. Janssen, et al., “The european pulsar timing array: current efforts and a leap toward the future”, *Classical and Quantum Gravity* **27**, 084014 (2010).
- [111] M. Fishbach, D. E. Holz, and W. M. Farr, “Does the black hole merger rate evolve with redshift?”, *The Astrophysical Journal Letters* **863**, L41 (2018).
- [112] E. E. Flanagan and S. A. Hughes, “Measuring gravitational waves from binary black hole coalescences: 2. The Waves’ information and its extraction, with and without templates”, *Phys. Rev. D* **57**, 4566–4587 (1998).
- [113] E. E. Flanagan and S. A. Hughes, “The Basics of gravitational wave theory”, *New J. Phys.* **7**, 204 (2005).

- [114] R. L. Forward, “Wide Band Laser Interferometer Gravitational Radiation Experiment”, *Phys. Rev. D* **17**, 379–390 (1978).
- [115] J. Frank, A. R. King, and D. Raine, *Accretion power in astrophysics*, eng, 3. ed (Cambridge Univ. Press, Cambridge, 2002).
- [116] A. M. Frassino, D. C. Lopes, and J. V. Rocha, “Extreme mass ratio head-on collisions of black holes in Einstein-scalar-Gauss-Bonnet theory”, (2026).
- [117] A. Gamboa et al., “Accurate waveforms for generic planar-orbit binary black holes: The multipolar effective-one-body model SEOBNRv6EHM”, (2026).
- [118] C. García-Quirós, S. Tiwari, and S. Babak, “GPU-accelerated LISA parameter estimation with full time-domain response”, *Phys. Rev. D* **112**, 064017 (2025).
- [119] M. Garg, A. Derdzinski, S. Tiwari, J. Gair, and L. Mayer, “Measuring eccentricity and gas-induced perturbation from gravitational waves of LISA massive black hole binaries”, *Mon. Not. Roy. Astron. Soc.* **532**, 4060–4074 (2024).
- [120] M. Garg, A. Derdzinski, L. Zwick, P. R. Capelo, and L. Mayer, “The imprint of gas on gravitational waves from LISA intermediate-mass black hole binaries”, en, *Monthly Notices of the Royal Astronomical Society* **517**, arXiv:2206.05292 [astro-ph, physics:gr-qc], 1339–1354 (2022).
- [121] M. Garg, A. Franchini, and A. Lupi, “Gas-induced perturbations on the gravitational wave in-spiral of live post-Newtonian LISA massive black hole binaries: 0.1 disk aspect ratio”, (2026).
- [122] M. Garg, A. Franchini, A. Lupi, M. Bonetti, and L. Mayer, “Gas-induced Perturbations on the Gravitational Wave Inspiral of Live Post-Newtonian LISA Massive Black Hole Binaries”, *Astrophys. J.* **993**, 145 (2025).
- [123] P. Goldreich and S. Tremaine, “The excitation of density waves at the Lindblad and corotation resonances by an external potential.”, *The Astrophysical Journal* **233**, 857–871 (1979).
- [124] P. Goldreich and S. Tremaine, “Disk-satellite interactions.”, *The Astrophysical Journal* **241**, 425–441 (1980).
- [125] P. Gupta, B. Bonga, A. J. K. Chua, and T. Tanaka, “Importance of tidal resonances in extreme-mass-ratio inspirals”, *Phys. Rev. D* **104**, 044056 (2021).
- [126] R. van Haasteren, Y. Levin, P. McDonald, and T. Lu, “On measuring the gravitational-wave background using Pulsar Timing Arrays”, *Mon. Not. Roy. Astron. Soc.* **395**, 1005 (2009).
- [127] M. Habouzit, M. Volonteri, and Y. Dubois, “Blossoms from black hole seeds: properties and early growth regulated by supernova feedback”, *Mon. Not. Roy. Astron. Soc.* **468**, 3935–3948 (2017).
- [128] E. Hamilton et al., “Improved gravitational wave model linking precessing inspirals and numerical-relativity-calibrated merger-ringdown”, *Phys. Rev. D* **113**, 084055 (2026).

- [129] W. K. Hastings, “Monte carlo sampling methods using markov chains and their applications”, *Biometrika* **57**, 97–109 (1970).
- [130] T. M. Heckman and P. Best, “The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe”, *Ann. Rev. Astron. Astrophys.* **52**, 589–660 (2014).
- [131] A. Hegade K. R., C. F. Gammie, and N. Yunes, “Relativistic treatment of accretion disk torques on extreme mass ratio inspirals around spinning black holes”, *Phys. Rev. D* **112**, 124068 (2025).
- [132] A. Hegade K. R., C. F. Gammie, and N. Yunes, “Relativistic treatment of accretion disk torques on extreme mass-ratio inspirals around nonspinning black holes”, *Phys. Rev. D* **112**, 124012 (2025).
- [133] R. W. Hellings and G. S. Downs, “Upper limits on the isotropic gravitational radiation background from pulsar timing analysis.”, *The Astrophysics Journal Letters* **265**, L39–L42 (1983).
- [134] G. Hobbs, A. Archibald, Z. Arzoumanian, D. Backer, M. Bailes, N. Bhat, M. Burgay, S. Burke-Spolaor, D. Champion, I. Cognard, et al., “The international pulsar timing array project: using pulsars as a gravitational wave detector”, *Classical and Quantum Gravity* **27**, 084013 (2010).
- [135] P. F. Hopkins and E. Quataert, “How Do Massive Black Holes Get Their Gas?”, *Mon. Not. Roy. Astron. Soc.* **407**, 1529–1564 (2010).
- [136] W.-R. Hu and Y.-L. Wu, “The taiji program in space for gravitational wave physics and the nature of gravity”, *National Science Review* **4**, 685–686 (2017).
- [137] S. A. Hughes, S. Drasco, E. E. Flanagan, and J. Franklin, “Gravitational radiation reaction and inspiral waveforms in the adiabatic limit”, *Phys. Rev. Lett.* **94**, 221101 (2005).
- [138] H. Jeffreys, *The theory of probability*, Oxford Classic Texts in the Physical Sciences (OUP Oxford, 1998).
- [139] F. Jenet, L. Finn, J. Lazio, A. Lommen, M. McLaughlin, I. Stairs, D. Stinebring, J. Verbiest, A. Archibald, Z. Arzoumanian, et al., “The north american nanohertz observatory for gravitational waves”, arXiv preprint arXiv:0909.1058 (2009).
- [140] M. L. Jones, R. C. Hickox, C. S. Black, K. N. Hainline, M. A. DiPompeo, and A. D. Goulding, “The intrinsic eddington ratio distribution of active galactic nuclei in star-forming galaxies from the sloan digital sky survey”, *The Astrophysical Journal* **826**, 12 (2016).
- [141] K. D. Kanagawa, H. Tanaka, and E. Szuszkiewicz, “Radial migration of gap-opening planets in protoplanetary disks. I. The case of a single planet”, en, *The Astrophysical Journal* **861**, 140 (2018).
- [142] N. Karnesis, S. Babak, M. Pieroni, N. Cornish, and T. Littenberg, “Characterization of the stochastic signal originating from compact binary populations as measured by LISA”, *Phys. Rev. D* **104**, 043019 (2021).

- [143] N. Karnesis, M. L. Katz, N. Korsakova, J. R. Gair, and N. Stergioulas, “Eryn : A multi-purpose sampler for Bayesian inference”, *Monthly Notices of the Royal Astronomical Society* **526**, 4814–4830 (2023).
- [144] N. Karnesis, A. Sasli, R. Buscicchio, and N. Stergioulas, “Characterization of non-Gaussian stochastic signals with heavier-tailed likelihoods”, *Phys. Rev. D* **111**, 022005 (2025).
- [145] M. L. Katz, J.-B. Bayle, A. J. K. Chua, and M. Vallisneri, “Assessing the data-analysis impact of LISA orbit approximations using a GPU-accelerated response model”, *Phys. Rev. D* **106**, 103001 (2022).
- [146] M. L. Katz, A. J. K. Chua, L. Speri, N. Warburton, and S. A. Hughes, “Fast extreme-mass-ratio-inspiral waveforms: New tools for millihertz gravitational-wave data analysis”, *Phys. Rev. D* **104**, 064047 (2021).
- [147] M. L. Katz, A. J. K. Chua, L. Speri, N. Warburton, and S. A. Hughes, “FastEM-RIWaveforms: New tools for millihertz gravitational-wave data analysis”, *Physical Review D* **104**, 064047 (2021).
- [148] M. L. Katz, N. Karnesis, N. Korsakova, J. R. Gair, and N. Stergioulas, “Efficient GPU-accelerated multisource global fit pipeline for LISA data analysis”, *Phys. Rev. D* **111**, 024060 (2025).
- [149] S. Kejriwal, E. Barausse, and A. J. K. Chua, “Hierarchical modeling of gravitational-wave populations for disentangling environmental and modified-gravity effects”, *Phys. Rev. D* **113**, 064001 (2026).
- [150] H. Khalvati, P. Lynch, O. Burke, L. Speri, M. van de Meent, and Z. Nasipak, “Systematic errors in fast relativistic waveforms for extreme mass ratio inspirals”, *Phys. Rev. D* **113**, 084042 (2026).
- [151] H. Khalvati, A. Santini, F. Duque, L. Speri, J. Gair, H. Yang, and R. Brito, “Impact of relativistic waveforms in LISA’s science objectives with extreme-mass-ratio inspirals”, *Phys. Rev. D* **111**, 082010 (2025).
- [152] S. Khukhlaev and S. Babak, “”Chat” between extreme mass ratio inspirals and Galactic binaries in LISA data”, arXiv preprint (2025).
- [153] L. E. Kidder, C. M. Will, and A. G. Wiseman, “Spin effects in the inspiral of coalescing compact binaries”, *Phys. Rev. D* **47**, R4183(R)–R4187(R) (1993).
- [154] H. Kim and W.-T. Kim, “Dynamical Friction of a Circular-Orbit Perturber in a Gaseous Medium”, *Astrophys. J.* **665**, 432–444 (2007).
- [155] A. Klein et al., “The last three years: multiband gravitational-wave observations of stellar-mass binary black holes”, (2022).
- [156] A. Klein et al., “Science with the space-based interferometer eLISA. I: Supermassive black hole binaries”, *Physical Review D* **93**, arXiv:1511.05581 [astro-ph, physics:gr-qc], 024003 (2016).

- [157] B. Kocsis, N. Yunes, and A. Loeb, “Observable Signatures of EMRI Black Hole Binaries Embedded in Thin Accretion Disks”, *Phys. Rev. D* **84**, 024032 (2011).
- [158] D. G. Korycansky and J. B. Pollack, “Numerical Calculations of the Linear Response of a Gaseous Disk to a Protoplanet”, *Icarus* **102**, 150–165 (1993).
- [159] L. Küchler, G. Compère, and A. Pound, “Self-force framework for merger-ringdown waveforms”, *Class. Quant. Grav.* **43**, 015018 (2026).
- [160] A. Lamberts, S. Garrison-Kimmel, P. Hopkins, E. Quataert, J. Bullock, C.-A. Faucher-Giguère, A. Wetzel, D. Keres, K. Drango, and R. Sanderson, “Predicting the binary black hole population of the Milky Way with cosmological simulations”, *Mon. Not. Roy. Astron. Soc.* **480**, 2704–2718 (2018).
- [161] L. Landau and E. Lifshits, *The classical theory of fields: volume 2*, Course of theoretical physics (Elsevier Science, 1975).
- [162] L. Lehoucq, I. Dvorkin, R. Srinivasan, C. Pellouin, and A. Lamberts, “Astrophysical uncertainties in the gravitational-wave background from stellar-mass compact binary mergers”, *Mon. Not. Roy. Astron. Soc.* **526**, 4378–4387 (2023).
- [163] T. G. F. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio, “Towards a generic test of the strong field dynamics of general relativity using compact binary coalescence”, *Phys. Rev. D* **85**, 082003 (2012).
- [164] *LIGO-M1100296-v2: LIGO-India, Proposal of the Consortium for Indian Initiative in Gravitational-wave Observations (IndIGO)*, [Public], <https://dcc.ligo.org/ligo-m1100296/public> (visited on 08/10/2026).
- [165] D. N. C. Lin and J. Papaloizou, “On the tidal interaction between protoplanets and the primordial solar nebula. ii - self-consistent nonlinear interaction”, *The Astrophysical Journal* **307**, 395 (1986).
- [166] D. N. C. Lin and J. C. B. Papaloizou, “On the tidal interaction between protostellar disks and companions”, in *Protostars and planets iii*, edited by E. H. Levy and J. I. Lunine (Arizona University Press, 1993).
- [167] L. Lindblom, B. J. Owen, and D. A. Brown, “Model Waveform Accuracy Standards for Gravitational Wave Data Analysis”, *Phys. Rev. D* **78**, 124020 (2008).
- [168] T. B. Littenberg, J. G. Baker, A. Buonanno, and B. J. Kelly, “Systematic biases in parameter estimation of binary black-hole mergers”, *Phys. Rev. D* **87**, 104003 (2013).
- [169] T. B. Littenberg and N. J. Cornish, “Prototype global analysis of LISA data with multiple source types”, *Phys. Rev. D* **107**, 063004 (2023).
- [170] B. Liu and D. Lai, “Hierarchical Black-Hole Mergers in Multiple Systems: Constrain the Formation of GW190412, GW190814 and GW190521-like events”, *Mon. Not. Roy. Astron. Soc.* **502**, 2049–2064 (2021).

- [171] L. Lui, L. V. Drummond, and A. Torres-Orjuela, “Pitching Cosmic Curveballs: Environmental Effects on Extreme-Mass-Ratio Inspirals with Spinning Secondaries”, (2026).
- [172] J. Luo, L.-S. Chen, H.-Z. Duan, Y.-G. Gong, S. Hu, J. Ji, Q. Liu, J. Mei, V. Milyukov, M. Sazhin, et al., “Tianqin: a space-borne gravitational wave detector”, *Classical and Quantum Gravity* **33**, 035010 (2016).
- [173] P. Madau and M. J. Rees, “Massive black holes as Population III remnants”, *Astrophys. J. Lett.* **551**, L27–L30 (2001).
- [174] M. Maggiore, “7 data analysis techniques”, in *Gravitational waves: volume 1: theory and experiments* (Oxford University Press, Oct. 2007).
- [175] M. Malik, F. Meru, L. Mayer, and M. Meyer, “On the gap-opening criterion of migrating planets in protoplanetary disks”, *The Astrophysical Journal* **802**, 56 (2015).
- [176] D. Mancieri, L. Broggi, M. Vinciguerra, A. Sesana, and M. Bonetti, “Eccentricity distribution of extreme mass ratio inspirals”, *Phys. Rev. D* **113**, 043062 (2026).
- [177] I. Mandel, W. M. Farr, and J. R. Gair, “Extracting distribution parameters from multiple uncertain observations with selection biases”, *Mon. Not. Roy. Astron. Soc.* **486**, 1086–1093 (2019).
- [178] S. Marsat, J. G. Baker, and T. Dal Canton, “Exploring the Bayesian parameter estimation of binary black holes with LISA”, *Phys. Rev. D* **103**, 083011 (2021).
- [179] F. S. Masset, “The co-orbital corotation torque in a viscous disk: Numerical simulations”, *Astronomy and Astrophysics* **387**, 605–623 (2002).
- [180] E. Mauceli, Z. K. Geng, W. O. Hamilton, W. W. Johnson, S. Merkowit, A. Morse, B. Price, and N. Solomonson, “The ALLEGRO gravitational wave detector: Data acquisition and analysis”, *Phys. Rev. D* **54**, 1264–1275 (1996).
- [181] L. Mayer, S. Kazantzidis, P. Madau, M. Colpi, T. R. Quinn, and J. Wadsley, “Rapid Formation of Supermassive Black Hole Binaries in Galaxy Mergers with Gas”, *Science* **316**, 1874–1877 (2007).
- [182] P. McNamara, S. Vitale, and K. Danzmann (LISA), “LISA Pathfinder”, *Class. Quant. Grav.* **25**, edited by S. M. Scott and D. E. McClelland, 114034 (2008).
- [183] B. J. Meers, “Recycling in Laser Interferometric Gravitational Wave Detectors”, *Phys. Rev. D* **38**, 2317–2326 (1988).
- [184] A. K. Mehta, A. Buonanno, R. Cotesta, A. Ghosh, N. Sennett, and J. Steinhoff, “Tests of general relativity with gravitational-wave observations using a flexible theory-independent method”, *Phys. Rev. D* **107**, 044020 (2023).
- [185] S. Mezzasoma, C.-J. Haster, and N. Yunes, “Impact of numerical-relativity waveform calibration on parametrized post-Einsteinian tests”, *Phys. Rev. D* **113**, 124043 (2026).

- [186] J. C. Mihos and L. Hernquist, “Gasdynamics and starbursts in major mergers”, *Astrophys. J.* **464**, 641 (1996).
- [187] B. Moore, T. Robson, N. Loutrel, and N. Yunes, “Towards a Fourier domain waveform for non-spinning binaries with arbitrary eccentricity”, *Class. Quant. Grav.* **35**, 235006 (2018).
- [188] C. J. Moore, C. P. L. Berry, A. J. K. Chua, and J. R. Gair, “Improving gravitational-wave parameter estimation using Gaussian process regression”, *Phys. Rev. D* **93**, 064001 (2016).
- [189] C. J. Moore and J. R. Gair, “Novel Method for Incorporating Model Uncertainties into Gravitational Wave Parameter Estimates”, *Phys. Rev. Lett.* **113**, 251101 (2014).
- [190] C. J. Moore, D. Gerosa, and A. Klein, “Are stellar-mass black-hole binaries too quiet for LISA?”, *Mon. Not. Roy. Astron. Soc.* **488**, L94–L98 (2019).
- [191] M. Muratore, D. Vetrugno, and S. Vitale, “Revisitation of time delay interferometry combinations that suppress laser noise in LISA”, *Class. Quant. Grav.* **37**, 185019 (2020).
- [192] A. Nagar, P. Rettegno, R. Gamba, S. Albanesi, A. Albertini, and S. Bernuzzi, “Analytic systematics in next generation of effective-one-body gravitational waveform models for future observations”, *Phys. Rev. D* **108**, 124018 (2023).
- [193] S. Naoz, “The eccentric kozai-lidov effect and its applications”, *Annual Review of Astronomy and Astrophysics* **54**, 441–489 (2016).
- [194] P. Narayan, N. K. Johnson-McDaniel, and A. Gupta, *Effect of Ignoring Eccentricity in Testing General Relativity with Gravitational Waves*, (Aug. 29, 2023) <http://arxiv.org/abs/2306.04068> (visited on 08/11/2026), pre-published.
- [195] J. F. Navarro, C. S. Frenk, and S. D. M. White, “The Structure of cold dark matter halos”, *Astrophys. J.* **462**, 563–575 (1996).
- [196] R. P. Nelson, “On the orbital evolution of low mass protoplanets in turbulent, magnetised disks”, en, *Astronomy & Astrophysics* **443**, 1067–1085 (2005).
- [197] R. P. Nelson, J. C. B. Papaloizou, F. Masset, and W. Kley, “The migration and growth of protoplanets in protostellar discs”, en, *Monthly Notices of the Royal Astronomical Society* **318**, 18–36 (2000).
- [198] R. P. Nelson, “Planetary Migration in Protoplanetary Disks”, en, in *Handbook of Exoplanets*, edited by H. J. Deeg and J. A. Belmonte (Springer International Publishing, Cham, 2018), pp. 2287–2317.
- [199] S. Nissanke, D. E. Holz, S. A. Hughes, N. Dalal, and J. L. Sievers, “Exploring short gamma-ray bursts as gravitational-wave standard sirens”, *Astrophys. J.* **725**, 496–514 (2010).
- [200] S. Nissanke, M. Vallisneri, G. Nelemans, and T. A. Prince, “Gravitational-wave emission from compact Galactic binaries”, *Astrophys. J.* **758**, 131 (2012).

- [201] E. C. Ostriker, “Dynamical friction in a gaseous medium”, *Astrophys. J.* **513**, 252 (1999).
- [202] Z. Pan, Z. Lyu, and H. Yang, “Wet extreme mass ratio inspirals may be more common for spaceborne gravitational wave detection”, *Phys. Rev. D* **104**, 063007 (2021).
- [203] Z. Pan and H. Yang, “Formation Rate of Extreme Mass Ratio Inspirals in Active Galactic Nuclei”, *Phys. Rev. D* **103**, 103018 (2021).
- [204] P. C. Peters, *Phys. Rev.* **136**, B1224–B1232 (1964).
- [205] P. C. Peters and J. Mathews, “Gravitational Radiation from Point Masses in a Keplerian Orbit”, en, *Physical Review* **131**, 435–440 (1963).
- [206] M. Piarulli, R. Buscicchio, F. Pozzoli, O. Burke, M. Bonetti, and A. Sesana, “Test for LISA foreground Gaussianity and stationarity: Extreme mass-ratio inspirals”, *Phys. Rev. D* **111**, 103047 (2025).
- [207] M. Piarulli, S. Marsat, E. M. Sanger, A. Buonanno, J. Steinhoff, and N. Tamanini, “Parametrized test of general relativity for LISA massive black hole binary inspirals”, *Phys. Rev. D* **112**, 124044 (2025).
- [208] M. d. L. Planas, A. Ramos-Buades, C. Garca-Quiros, H. Estelles, S. Husa, and M. Haney, “Time-domain phenomenological multipolar waveforms for aligned-spin binary black holes in elliptical orbits”, *Phys. Rev. D* **113**, 024006 (2026).
- [209] L. Pompili et al., “Laying the foundation of the effective-one-body waveform models SEOBNRv5: Improved accuracy and efficiency for spinning nonprecessing binary black holes”, *Phys. Rev. D* **108**, 124035 (2023).
- [210] L. Pompili, A. Buonanno, and M. Purrer, “Accounting for numerical-relativity calibration uncertainty in gravitational-wave modeling and inference”, *Phys. Rev. D* **112**, 044062 (2025).
- [211] F. Pozzoli, S. Babak, A. Sesana, M. Bonetti, and N. Karnesis, “Computation of stochastic background from extreme-mass-ratio inspiral populations for LISA”, *Phys. Rev. D* **108**, [Erratum: *Phys.Rev.D* 110, 049903 (2024)], 103039 (2023).
- [212] G. Pratten, S. Husa, C. Garcia-Quiros, M. Colleoni, A. Ramos-Buades, H. Estelles, and R. Jaume, “Setting the cornerstone for a family of models for gravitational waves from compact binaries: The dominant harmonic for nonprecessing quasicircular black holes”, *Phys. Rev. D* **102**, 064001 (2020).
- [213] G. Pratten, P. Schmidt, H. Middleton, and A. Vecchio, “Precision tracking of massive black hole spin evolution with lisa”, *Phys. Rev. D* **108**, 124045 (2023).
- [214] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, “The LISA optimal sensitivity”, *Phys. Rev. D* **66**, 122002 (2002).
- [215] S. Ranjan, K. Jani, A. H. Nitz, K. Holley-Bockelmann, and C. Cutler, “A Tale of Two Black Holes: Multiband Gravitational-wave Measurement of Recoil Kicks”, *Astrophys. J.* **983**, 27 (2025).

- [216] T. Regge and J. A. Wheeler, “Stability of a Schwarzschild singularity”, *Phys. Rev.* **108**, 1063–1069 (1957).
- [217] D. Reitze et al., “Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO”, *Bull. Am. Astron. Soc.* **51**, 035 (2019).
- [218] E. Roebber et al., “Milky Way Satellites Shining Bright in Gravitational Waves”, *Astrophys. J. Lett.* **894**, L15 (2020).
- [219] J. Roulet, M. Crisostomi, L. M. Thomas, and K. Chatziioannou, “Domain-optimized machine-learning tool for gravitational-wave inference”, *Phys. Rev. D* **114**, 024080 (2026).
- [220] J. Sadiq, K. Dey, T. Dent, and E. Barausse, “Reconstructing the LISA massive black hole binary population via iterative kernel density estimation”, *Phys. Rev. D* **111**, 063051 (2025).
- [221] E. E. Salpeter, “Accretion of Interstellar Matter by Massive Objects”, *Astrophys. J.* **140**, 796–800 (1964).
- [222] A. Sasli, N. Karnesis, and N. Stergioulas, “Heavy-tailed likelihoods for robustness against data outliers: Applications to the analysis of gravitational wave data”, *Phys. Rev. D* **108**, 103005 (2023).
- [223] P. R. Saulson, “TERRESTRIAL GRAVITATIONAL NOISE ON A GRAVITATIONAL WAVE ANTENNA”, *Phys. Rev. D* **30**, 732–736 (1984).
- [224] L. Sberna et al., “Observing GW190521-like binary black holes and their environment with LISA”, *Phys. Rev. D* **106**, 064056 (2022).
- [225] M. A. Scheel et al., “The SXS collaboration’s third catalog of binary black hole simulations”, *Class. Quant. Grav.* **42**, 195017 (2025).
- [226] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi, “Linking the spin evolution of massive black holes to galaxy kinematics”, *Astrophys. J.* **794**, 104 (2014).
- [227] A. Sesana, “Prospects for Multiband Gravitational-Wave Astronomy after GW150914”, *Phys. Rev. Lett.* **116**, 231102 (2016).
- [228] D. A. Shaddock, M. Tinto, F. B. Estabrook, and J. W. Armstrong, “Data combinations accounting for LISA spacecraft motion”, *Phys. Rev. D* **68**, 061303 (2003).
- [229] M. A. Shaikh, S. A. Bhat, and S. J. Kapadia, “A study of the Inspiral-Merger-Ringdown Consistency Test with gravitational-wave signals from compact binaries in eccentric orbits”, *Physical Review D* **110**, 024030 (2024).
- [230] F. Shankar, D. H. Weinberg, and J. Miralda-Escude, “Self-Consistent Models of the AGN and Black Hole Populations: Duty Cycles, Accretion Rates, and the Mean Radiative Efficiency”, *Astrophys. J.* **690**, 20–41 (2009).
- [231] J. Skilling, “Nested sampling for general bayesian computation”, *Bayesian Analysis* **1**, 833–859 (2006).

- [232] V. Skoupý, G. A. Piovano, and V. Witzany, “Spherical inspirals of spinning bodies into Kerr black holes”, *Phys. Rev. D* **112**, 124054 (2025).
- [233] C. Smarra et al. (European Pulsar Timing Array), “Second Data Release from the European Pulsar Timing Array: Challenging the Ultralight Dark Matter Paradigm”, *Phys. Rev. Lett.* **131**, 171001 (2023).
- [234] A. Spadaro, J. Gair, D. Gerosa, S. R. Green, R. Buscicchio, N. Gupte, R. Tenorio, S. Clyne, M. Pürrer, and N. Korsakova, “Accurate and efficient simulation-based inference for massive black-hole binaries with LISA”, *Phys. Rev. D* **114**, 023038 (2026).
- [235] J. S. Speagle, “dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences”, *Mon. Not. Roy. Astron. Soc.* **493**, 3132–3158 (2020).
- [236] L. Speri, A. Antonelli, L. Sberna, S. Babak, E. Barausse, J. R. Gair, and M. L. Katz, “Probing Accretion Physics with Gravitational Waves”, *Phys. Rev. X* **13**, 021035 (2023).
- [237] L. Speri, M. L. Katz, A. J. K. Chua, S. A. Hughes, N. Warburton, J. E. Thompson, C. E. A. Chapman-Bird, and J. R. Gair, “Fast and Fourier: Extreme Mass Ratio Inspirational Waveforms in the Frequency Domain”, [10.3389/fams.2023.1266739](https://arxiv.org/abs/10.3389/fams.2023.1266739) (2023).
- [238] R. Srinivasan, E. Barausse, N. Korsakova, and R. Trotta, “Simulation-based population inference of LISA’s Galactic binaries: Bypassing the global fit”, *Phys. Rev. D* **112**, 103043 (2025).
- [239] R. Srinivasan, M. Crisostomi, R. Trotta, E. Barausse, and M. Breschi, “Bayesian evidence estimation from posterior samples with normalizing flows”, *Phys. Rev. D* **110**, 123007 (2024).
- [240] A. Stroeer and A. Vecchio, “The LISA verification binaries”, *Class. Quant. Grav.* **23**, edited by M. Diaz and S. Mohanty, S809–S818 (2006).
- [241] S. H. Strub, L. Ferraioli, C. Schmeltzbach, S. C. Stähler, and D. Giardini, “Global analysis of LISA data with Galactic binaries and massive black hole binaries”, *Phys. Rev. D* **110**, 024005 (2024).
- [242] D. Syer, “Relativistic dynamical friction in the weak scattering limit”, *Mon. Not. Roy. Astron. Soc.* **270**, 205 (1994).
- [243] J. Sylvestre and M. Tinto, “Noise characterization for LISA”, *Phys. Rev. D* **68**, 102002 (2003).
- [244] S. Tahura and K. Yagi, “Parameterized Post-Einsteinian Gravitational Waveforms in Various Modified Theories of Gravity”, *Phys. Rev. D* **98**, [Erratum: *Phys. Rev. D* **101**, 109902 (2020)], 084042 (2018).
- [245] J. Takátsy, L. Zwick, K. Hendriks, P. Saini, G. Fabj, and J. Samsing, “The construction and use of dephasing prescriptions for environmental effects in gravitational wave astronomy”, *Class. Quant. Grav.* **42**, 215006 (2025).

- [246] H. Tanaka and K. Okada, “Three-dimensional interaction between a planet and an isothermal gaseous disk. iii. locally isothermal cases”, *The Astrophysical Journal* **968**, 28 (2024).
- [247] H. Tanaka, T. Takeuchi, and W. R. Ward, “Three-dimensional Interaction between a Planet and an Isothermal Gaseous Disk. I. Corotation and Lindblad Torques and Planet Migration”, en, *The Astrophysical Journal* **565**, 1257–1274 (2002).
- [248] H. Tanaka and W. R. Ward, “Three-dimensional Interaction between a Planet and an Isothermal Gaseous Disk. II. Eccentricity Waves and Bending Waves”, en, *The Astrophysical Journal* **602**, 388–395 (2004).
- [249] A. Taracchini, A. Buonanno, G. Khanna, and S. A. Hughes, “Small mass plunging into a kerr black hole: anatomy of the inspiral-merger-ringdown waveforms”, *Phys. Rev. D* **90**, 084025 (2014).
- [250] L. S. S. Team, “Lisa science requirements document”,
- [251] S. A. Teukolsky, “Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations”, *Astrophys. J.* **185**, 635–647 (1973).
- [252] E. Thrane and C. Talbot, “An introduction to bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models”, *Publications of the Astronomical Society of Australia* **36**, e010 (2019).
- [253] M. Tinto, S. Dhurandhar, and D. Malakar, “Second-generation time-delay interferometry”, *Phys. Rev. D* **107**, 082001 (2023).
- [254] M. Tinto and S. V. Dhurandhar, “Time-delay interferometry”, *Living Rev. Rel.* **24**, 1 (2021).
- [255] M. Tinto, F. B. Estabrook, and J. W. Armstrong, “Time delay interferometry for LISA”, *Phys. Rev. D* **65**, 082003 (2002).
- [256] M. Tinto, F. B. Estabrook, and J. W. Armstrong, “Time delay interferometry with moving spacecraft arrays”, *Phys. Rev. D* **69**, 082001 (2004).
- [257] A. Toubiana et al., “Detectable environmental effects in GW190521-like black-hole binaries with LISA”, *Phys. Rev. Lett.* **126**, 101105 (2021).
- [258] A. Toubiana and J. R. Gair, “Indistinguishability criterion and estimating the presence of biases”, arXiv preprint (2024).
- [259] A. Toubiana and J. R. Gair, “A framework for LISA population inference”, *JCAP* **07**, 025 (2026).
- [260] J.-P. Uzan, “Varying Constants, Gravitation and Cosmology”, *Living Rev. Rel.* **14**, 2 (2011).
- [261] L. Vainshtein and V. Zubakov, *Extraction of signals from noise: by l.a. wainstein and v.d. zubakov* (Dover, 1970).
- [262] M. Vallisneri, “Geometric time delay interferometry”, *Phys. Rev. D* **72**, [Erratum: *Phys.Rev.D* 76, 109903 (2007)], 042003 (2005).

- [263] M. Vallisneri, “Beyond Fisher: exact sampling distributions of the maximum-likelihood estimator in gravitational-wave parameter estimation”, *Phys. Rev. Lett.* **107**, 191104 (2011).
- [264] M. Vallisneri, “Testing general relativity with gravitational waves: a reality check”, *Phys. Rev. D* **86**, 082001 (2012).
- [265] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, “Surrogate models for precessing binary black hole simulations with unequal masses”, *Phys. Rev. Research*. **1**, 033015 (2019).
- [266] R. Vicente, T. K. Karydas, and G. Bertone, “Fully Relativistic Treatment of Extreme Mass-Ratio Inspirals in Collisionless Environments”, *Phys. Rev. Lett.* **135**, 211401 (2025).
- [267] M. Volonteri, G. Lodato, and P. Natarajan, “The evolution of massive black hole seeds”, *Mon. Not. Roy. Astron. Soc.* **383**, 1079 (2008).
- [268] W. R. Ward and K. Hourigan, “Orbital migration of protoplanets - the inertial limit”, *The Astrophysical Journal* **347**, 490 (1989).
- [269] J. Weber, “Gravitational-wave-detector events”, *Phys. Rev. Lett.* **20**, 1307–1308 (1968).
- [270] J. Weber, “Evidence for discovery of gravitational radiation”, *Phys. Rev. Lett.* **22**, 1320–1324 (1969).
- [271] J. Weber, “Gravitational radiation experiments”, *Phys. Rev. Lett.* **24**, 276–279 (1970).
- [272] Y. Wu, Y.-X. Chen, and D. N. C. Lin, “Chaotic Type I migration in turbulent discs”, *Mon. Not. Roy. Astron. Soc.* **528**, L127–L132 (2023).
- [273] Y. Xie, D. Chatterjee, G. Narayan, and N. Yunes, “Neural post-Einsteinian framework for efficient theory-agnostic tests of general relativity with gravitational waves”, *Phys. Rev. D* **110**, 024036 (2024).
- [274] Y. Xie, G. Narayan, and N. Yunes, “Neural Post-Einsteinian Test of General Relativity with the Third Gravitational Wave Transient Catalog”, *Phys. Rev. Lett.* **136**, 181402 (2026).
- [275] S. Yi, F. Iacovelli, E. Berti, R. S. Chandramouli, S. Marsat, D. Wadekar, and N. Yunes, “Systematic biases in parameter estimation on LISA binaries. II. The effect of excluding higher harmonics for spin-aligned, high-mass binaries”, *Phys. Rev. D* **114**, 024017 (2026).
- [276] S. Yi, F. Iacovelli, S. Marsat, D. Wadekar, and E. Berti, “Systematic biases in parameter estimation on LISA binaries: The effect of excluding higher harmonics for non-spinning binaries”, *Phys. Rev. D* **112**, 124063 (2025).
- [277] N. Yunes, M. Coleman Miller, and J. Thornburg, “The Effect of Massive Perturbors on Extreme Mass-Ratio Inspiral Waveforms”, *Phys. Rev. D* **83**, 044030 (2011).

- [278] N. Yunes and F. Pretorius, “Fundamental Theoretical Bias in Gravitational Wave Astrophysics and the Parameterized Post-Einsteinian Framework”, *Phys. Rev. D* **80**, 122003 (2009).
- [279] N. Yunes, F. Pretorius, and D. Spergel, “Constraining the evolutionary history of Newton’s constant with gravitational wave observations”, *Phys. Rev. D* **81**, 064018 (2010).
- [280] N. Yunes, K. Yagi, and F. Pretorius, “Theoretical Physics Implications of the Binary Black-Hole Mergers GW150914 and GW151226”, *Phys. Rev. D* **94**, 084002 (2016).
- [281] L. Zwick, “Limits of vacuum-template subtraction for LISA massive black hole binary sources in realistic environments”, arXiv e-prints (2026).
- [282] L. Zwick, A. Derdzinski, M. Garg, P. R. Capelo, and L. Mayer, “Dirty waveforms: multiband harmonic content of gas-embedded gravitational wave sources”, en, *Monthly Notices of the Royal Astronomical Society* **511**, 6143–6159 (2022).
- [283] L. Zwick, K. Hendriks, P. Saini, J. Takátsy, C. Rowan, J. Samsing, and J. Stegmann, “Environmental effects in stellar mass gravitational wave sources II: Joint detections of eccentricity and phase shifts in binary sub-populations”, arXiv e-prints (2025).