Università degli Studi dell'Insubria - Como

Physics Department

The dissertation is submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Challenges in the stochastic sector of LISA data analysis

Supervisor
Prof. Francesco Haardt

Candidate
Federico Pozzoli

Co-Supervisor
Dr. Riccardo Buscicchio

Academic Year 2024 - 2025

Cycle XXXVIII

Abstract

The Laser Interferometer Space Antenna (LISA) is a pioneering space-based observatory designed to detect Gravitational Waves (GWs) in the millihertz frequency band. Over its mission lifetime, LISA is expected to probe a wide variety of astrophysical phenomena, ranging from mergers of massive black hole binaries in the early Universe to inspirals of stellar-origin compact objects within our own Galaxy. In particular, LISA will resolve more than ten thousand nearly monochromatic white dwarf binaries, providing unprecedented insights into the Galactic population of compact binaries. Beyond individually resolvable binaries, the majority of Galactic sources will remain unresolved, producing a confusion-limited signal known as the Galactic foreground, which is predicted to dominate the 0.5–3 mHz frequency range. In addition, astrophysical and cosmological processes are expected to contribute to the LISA datastreams, generating multiple Stochastic Gravitational Wave Backgrounds (SGWBs).

The LISA datastream will be signal-rich, with GW sources continuously emitting within the instrument frequency band. Such complexity motivates the development of the so-called *Global fit*. Among many proposed approaches, the most consolidated one is to use a Blocked Gibbs sampling algorithm to jointly analyze different GW source classes, including instrumental noise and SGWBs.

In this thesis, we focus on the modelling challenges for the stochastic and noise sectors of the Global fit. We mainly focus on astrophysical contributions, from White Dwarfs (WDs) binaries, Stellar-Origin Black Holes Binaries (SOBBHs), and Extreme Mass Ratio Inspirals (EMRIs).

The main goal in modeling a stochastic signal is to disentangle it from instrumental noise. This is a non-trivial task, particularly when accounting for uncertainties associated to both. Unlike ground-based detectors, LISA instrumental noise uncertainties stem from limited knowledge of the environment in which the satellites reside. Similarly, incomplete knowledge of astrophysical source populations or additional physics beyond GW emission in vacuum, must be accounted for.

To investigate the aspects above, we characterize the EMRI background by considering several population models and study the SOBBH one, including possible environmental effects. We explore systematically amplitude and spectral shape uncertainties in both. To account for them in a data analysis context, we develop a model based Expectation value of Gaussian Process (EGP), a flexible parametrization capable of capturing a wide variety of Power Spectral Density (PSD) shapes. Such uncertainties propagate into the spectral reconstruction and ultimately affect the effective detectability of the signal.

To rigorously quantify detectability, we leverage Bayesian evidence. This approach naturally incorporates marginalization over the uncertain noise properties of the instrument within the detectability criterion. Building on this, we propose a robust method to compute a *Bayesian sensitivity*, defined as the threshold Bayes factor above which a stochastic signal can be confidently identified.

Given the uncertainties and the overlapping nature of different SGWB processes, we explore additional, statistical properties to distinguish components from one another. For instance, the non-Gaussianity arising from the intrinsic Poissonian nature of the astrophysical SGWB may serve as a useful discriminator. To this end, we assess the level of non-Gaussianity in the EMRI background. We further investigate astrophysical anisotropies, leaving quasi-stationary signatures in the signal. This behavior is typical of the Galac-

tic foreground and of the SGWB produced by unresolved WD binary populations in the Milky Way satellites. To address this, we construct a fully Bayesian framework capable of handling anisotropies, extracting information from spectral properties but also on the morphology of the source.

Finally, we introduce `bahamas`, an algorithm designed to model the quasi-stationarity arising from the Galactic foreground in a Global Fit context. To this end, we develop a highly flexible time–frequency framework that can accommodate different setup configurations. In particular, by leveraging coarse-grained data and efficient sampling techniques, we demonstrate how to significantly suppress the computational cost. Moreover, the code infrastructure is designed to incorporate multiple source classes as well as mission-specific data gaps.

This thesis is organized as follows: in Chapter 1 we give an overview on the theory of GWs. We describe the GW emission from binaries systems and the specific peculiarities of eccentrics ones. We introduce SGWBs and their properties. In Chapter 2 we introduce the tools relevant to process statistically data recorded by GWs detectors. We describe relevant signal and noise parametrizations, leading to the definition of Signal-to-Noise Ratio (SNR) in a stochastic context. We then introduce Bayesian inference. In Chapter 3 we describe key aspects of LISA for our studies. We provide an overview of the mission design and we summarize the properties of expected astrophysical sources. We describe the instrument response to GWs and the dominant noise sources followed by an introduction to Time-Delay Interferometry (TDI).

In Chapter 4, we present results on the characterization of the astrophysical backgrounds from EMRIs and SOBBHs. We then introduce a weakly parametric method, the EGP, to flexibly model the PSD, explicitly accounting for uncertainties in the spectral shape. We also develop the concept of Bayesian sensitivity to assess the detectability of a stochastic signal in the presence of noise and its uncertainties.

In Chapter 5 we characterize typical statistical properties of SGWB processes of astrophysical origin. We begin by analyzing the non-Gaussianity of the EMRI background, then move on to study anisotropy and non-stationarity, evaluating their detectability in the context of the Milky Way and its satellites. Finally, we introduce a time–frequency approach as a Global-fit-viable approach to inference.

The results chapters (Chapters 4 and 5) conclude each main section with a *Caveats and Takeaways* subsection that summarizes key findings, highlights limitations and discusses implications for LISA science, and outlines directions for future work.

Chapter 6 synthesizes the key concepts and findings of this thesis.

Acronyms

AGN Active Galactic Nucleus

AAK Augmented Analytic Kludge

BH Black Hole

BBH Binary Black Hole

BPLS Bayesian Power-Law Sensitivity

CO Compact Object

EF Extragalactic Foreground

EGP Expectation value of Gaussian Process

EM Electromagnetic

EMRI Extreme Mass Ratio Inspiral

FT Fourier Transform

GF Galactic Foreground

GW Gravitational Wave

GP Gaussian Process

GR General Relativity

GSS Generalized Stepping Stone

HMC Hamiltonian Monte Carlo

IFE Iterative Foreground Estimation

JSD Jensen-Shannon Divergence

LMC Large Magellanic Cloud

LISA Laser Interferometer Space Antenna

LVK LIGO-Virgo-KAGRA

LPF LISA PathFinder

MBH Massive Black-Hole

MBHB Massive Black-Hole Binary

MCMC Markov Chain Monte Carlo

MLE Maximum Likelihood Estimation

MW Milky Way

NS Nested Sampling

NUTS No-U-Turn Sampler

OB Optical Bench

OMS Optical Metrology System

PN Post-Newtonian

PLS Power-Law Sensitivity

PSD Power Spectral Density

PTA Pulsar Timing Array

RPL Rational Power Law

RPLP Rational Power Law + Peak

SFR Star Formation Rate

SGWB Stochastic Gravitational Wave Background

SMC Small Magellanic Cloud

SNR Signal-to-Noise Ratio

SOBBH Stellar-Origin Black Hole Binary

SS Stepping Stone

SSB Solar System Barycenter

STFT Short Time Fourier Transforms

TI Thermodynamical integration

TM Test Mass

TDI Time-Delay Interferometry

TT Transverse-Traceless

WD White Dwarf

Notation

Units

Throughout the Thesis we adopt the *cgs* units system.

Unless otherwise specified, the speed of light and the Newton constant are set to $c = 3 \times 10^{10} \mathrm{cm} \cdot \mathrm{s}^{-1}$ and $G = 6.67 \times 10^{-8} g^{-1} \cdot \mathrm{cm}^3 \cdot \mathrm{s}^{-2}$, respectively. Where specified, we adopt geometric units where $G = c = 1$.

The solar mass is set to $M_\odot = 2 \times 10^{33} g$ and the typical astrophysical distances are set in units of *parsec* ($1pc = 3.09 \times 10^{18} \mathrm{cm}$).

Indices

Greek indices, such as α, β, ... or μ, ν, ..., take the values 0,..., 3, while spatial indices are denoted by Latin letters, i, j,... = 1, 2, 3.

Array and Vector

We denote array variables and vector quantities with bold symbols, such as $\boldsymbol{v}$. A hat symbol (e.g.,$\hat{\boldsymbol{v}}$) denotes a normalized (unit) vector.

Cosmology

We consider a flat ΛCDM cosmology ($H_0 = 70 \mathrm{Km/s/Mpc}$, $\Omega_{\mathrm{m},0} = 0.3$).

Contents

Chapter 1

Introduction to gravitational waves

In this chapter, we introduce in Sections 1.1 and 1.2 the theory of General Relativity (GR) and its linearization as theoretical framework for the description of GWs. In Section 1.3 we focus on the generation of GWs from binary systems.

1.1 Introduction to general relativity

In GR, the principle of relativity is extended to include non-inertial reference frames, i.e., frames that accelerate relative to one another. Following the *equivalence principle*, local effects of gravity are indistinguishable from those of acceleration.

Accordingly, GR describes gravity as the manifestation of the *curvature of spacetime*. In doing so, the flat Minkowski metric $\eta_{\mu\nu}$ from special relativity is replaced by a general metric tensor $g_{\mu\nu}$, which defines the spacetime geometry.

To determine the metric $g_{\mu\nu}$, one must solve the *Einstein field equations*, relating the curvature of spacetime to the distribution of matter and energy:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{1.1}$$

Here below, we do not aim to provide a formal derivation of Einstein field equations, rather build an intuition on their physical meaning and implications. To simplify the notation, we adopt natural units where $G = c = 1$. A natural starting point is the motion of a point-like test particle in curved spacetime. The trajectory of such a particle of mass m is determined by the principle of least action. The action is given by:

$$S = -m \int d\tau = -m \int \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}}\, d\tau, \tag{1.2}$$

where $d\tau$ is the proper time along the particle worldline. Minimizing the action yields the geodesic equation

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = 0, \tag{1.3}$$

where $\Gamma^\mu_{\alpha\beta}$ is the Christoffel symbol, defined in terms of the metric tensor $g_{\mu\nu}$ and its derivatives:

$$\Gamma^\alpha_{\beta\mu} = \frac{1}{2} g^{\alpha\sigma} (g_{\sigma\beta,\mu} + g_{\sigma\mu,\beta} - g_{\beta\mu,\sigma}), \tag{1.4}$$

In the Newtonian limit, we assume that the gravitational field is weak (i.e., $g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}$) and static (i.e., $\partial_0 g_{\mu\nu} = 0$). The tensor $h_{\mu\nu}$ is a small perturbation to the Minkowski metric. In this regime, the geodesic equation is approximated by:

$$\frac{d^2x^i}{d\tau^2} \approx -\Gamma^i_{00} = \frac{1}{2}\partial^i g_{00}. \tag{1.5}$$

Comparing Eq. 1.5 with the Newtonian equation of motion, we find that the g_{00} component of the metric tensor is related to the Newtonian gravitational potential ϕ by:

$$g_{00} = -1 - 2\phi, \tag{1.6}$$

which implies $h_{00} = -2\phi$. The potential ϕ is related to the mass density ρ by the Poisson equation:

$$\Delta\phi = 4\pi\rho, \tag{1.7}$$

where $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the Laplacian operator. Eq. 1.7 can be rewritten in terms of the metric tensor perturbation h_{00} as:

$$\Delta h_{00} = -8\pi\rho. \tag{1.8}$$

Eq. 1.8 relates the metric perturbation to the matter density. We want to promote this equation to a more general one, interpreting the right-hand-side as a source term. Specifically, we consider ρ related to the T_{00} component of the stress-energy tensor $T_{\mu\nu}$, describing the distribution of energy and momentum in spacetime. For the left-hand side, we need a rank-2 tensor, that contains information about spacetime curvature. This leads to the Riemann tensor $R^\alpha_{\beta\mu\nu}$, a rank-4 tensor characterizing spacetime geometry, which can be expressed as:

$$R^\alpha_{\beta\mu\nu} = \Gamma^\alpha_{\beta\nu,\mu} - \Gamma^\alpha_{\beta\mu,\nu} + \Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\beta\mu} - \Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\beta\nu}, \tag{1.9}$$

where the comma denotes the partial derivative with respect to the coordinates. For our purposes, we consider the Ricci tensor $R_{\mu\nu}$, which is a contraction of the Riemann tensor:

$$R_{\mu\nu} = R^\alpha_{\mu\alpha\nu} = g^{\alpha\sigma}R_{\mu\sigma\nu\alpha}. \tag{1.10}$$

Under static conditions—i.e., when the metric and matter distribution are time independent—the Ricci tensor can be algebraically manipulated to yield the following relation:

$$R_{00} \sim -\frac{1}{2}\Delta h_{00}, \tag{1.11}$$

such that we rewrite Eq. 1.8 as:

$$R_{00} \sim 4\pi T_{00}. \tag{1.12}$$

This suggests a more general relation between the curvature of spacetime and the stress-energy tensor:

$$R_{\mu\nu} \overset{?}{=} 4\pi T_{\mu\nu}, \tag{1.13}$$

Unfortunately, this equation is not generally valid. The stress-energy tensor $T_{\mu\nu}$ has to satisfy the conservation law, i.e., $\nabla^\mu T_{\mu\nu} = 0$ with ∇^μ being the contracted covariant derivative. The Ricci tensor $R_{\mu\nu}$ does not always satisfy such constraint in general, i.e. $\nabla^\mu R_{\mu\nu} \neq 0$. In general, its divergence is given by:

$$\nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R. \tag{1.14}$$

where $R = g^{\mu\nu}R_{\mu\nu}$ is the Ricci scalar, i.e., Ricci tensor trace. Thus, the Einstein tensor $G_{\mu\nu}$, defined as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \tag{1.15}$$

is divergence-free, i.e. $\nabla^{\mu}G_{\mu\nu} = 0$. This suggests that $G_{\mu\nu} = \alpha T_{\mu\nu}$. We set $\alpha = 8\pi$ by requiring consistency with Newtonian gravity. This gives us the Einstein field equations:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}R\, g_{\mu\nu} = 8\pi T_{\mu\nu}. \tag{1.16}$$

Eq. 1.16 is a set of sixteen scalar equations. The metric and stress-energy tensor are both symmetric tensors. Thus, Einstein equations can be written as a set of ten equations. Furthermore, they are not independent, since the Einstein tensor is divergence-free, i.e. $\nabla^{\mu}G_{\mu\nu} = 0$. This relation, also known as the Bianchi identity, represents four further constraints on the Einstein field equations. In this way, we get a total of six independent equations. In most cases, one has to solve the Einstein equations numerically. However, it is possible to find analytical solutions in special cases as the weak field approximation.

1.2 Linearization of Einstein equations

GWs are small perturbations of the spacetime metric $g_{\mu\nu}$, whose existence is predicted from the linearization of Einstein equations in the weak-field approximation. In what follows, we briefly review the weak field approximation and the derivation of the linearized Einstein equations.
For simplicity, we assume that far from matter and energy the metric is approximately flat, i.e., a perturbation over Minkowski spacetime:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} + \mathcal{O}(h^2), \tag{1.17}$$

with $h_{\mu\nu}$ a small spacetime perturbation. Eq. 1.17 can be inserted inside the Einstein equation, keeping only terms linear in the spacetime perturbation $h_{\mu\nu}$.
It is convenient to work with the trace-reversed perturbation:

$$\overline{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h, \tag{1.18}$$

where $h = \eta^{\mu\nu}h_{\mu\nu}$ is the perturbation trace. Being GR a gauge theory, we choose the harmonic gauge condition to further simplify the equations. The harmonic gauge condition reads

$$\partial_{\mu}\overline{h}^{\mu\nu} = 0. \tag{1.19}$$

Under this gauge, the weak-field Einstein equations become:

$$\Box\overline{h}_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu}, \tag{1.20}$$

where $\Box = \eta^{\mu\nu}\partial_{\mu}\partial_{\nu}$ is the d'Alembert operator, i.e., the wave operator in Minkowski spacetime.
In a region far away from any matter and energy source, the components of the stress energy tensor $T_{\mu\nu}$ vanish, leading to:

$$\Box\overline{h}_{\mu\nu} = 0. \tag{1.21}$$

This equation describes the propagation of GWs in vacuum. The solution of this equation can be written as a superposition of plane waves:

$$\overline{h}_{\mu\nu} = A_{\mu\nu} e^{ik_\alpha x^\alpha}, \tag{1.22}$$

with $A_{\mu\nu}$ representing the complex amplitude tensor and k_α the wave four-vector. The speed of the wave is given by the dispersion relation $k^\alpha k_\alpha = 0$, implying that the wave propagates at the speed of light.

The harmonic gauge adopted in the weak-field approximation must also be satisfied by the plane-wave ansatz, leading to an important constraint:

$$A_{\mu\nu} k^\nu = 0. \tag{1.23}$$

Eq. 1.23 means that $A_{\mu\nu}$ is orthogonal with respect k^ν and then GWs are transverse. The harmonic gauge does not fully determine the solution: transformations of the form

$$\overline{h}_{\mu\nu} \to \overline{h}_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu, \tag{1.24}$$

with $\Box \xi_\mu = 0$ and arbitrary ξ_μ, leave both the gauge condition and the linearized equations invariant. We further reduce the gauge freedom by imposing the Transverse-Traceless (TT) condition. With TT gauge, we set all four components $\overline{h}_{0\nu}$ to zero, leaving only the spatial components h_{ij}. We also impose vanishing trace, $\overline{h}^\mu_\mu = 0$, which means $\overline{h}_{\mu\nu} = h_{\mu\nu}$. In conclusion, we set

$$h^{0\mu} = 0, \qquad h^i_j = 0, \qquad \partial^i h_{ij} = 0. \tag{1.25}$$

For a single plane wave traveling along the z-axis, the perturbation tensor in TT gauge reads:

$$h^{TT}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{1.26}$$

where h_+ and $h_\times$ are the amplitudes of the plus and cross polarizations, respectively. Eq. 1.26 is the GW tensor in TT gauge, describing a plane GW traveling along the z-direction. This shows that GWs have two degrees of freedom (or polarization), and $h^{TT}_{\mu\nu}$ can therefore be decomposed as:

$$h^{TT}_{\mu\nu} = h_+ e^+_{\mu\nu} + h_\times e^\times_{\mu\nu}, \tag{1.27}$$

where $e^p_{\mu\nu}$ are the polarization tensors. The polarization tensors e_+ and $e_\times$ describe two independent distortions due to GWs: h_+ stretches and squeezes spacetime along orthogonal directions transverse to the propagation, and $h_\times$ along the diagonal axes. To understand this effect, we now consider two free particles, whose separation is described by the four-vector ξ^μ. The geodesic deviation equation describes the evolution of ξ^μ due to the curvature of spacetime induced by the gravitational field. The geodesic deviation equation reads:

$$\frac{d^2 \xi^\mu}{d\tau^2} = R^\mu_{\alpha\beta\gamma} u^\alpha u^\beta \xi^\gamma, \tag{1.28}$$

where $u^{\alpha,\beta} = d\xi^{\alpha,\beta}/d\tau$. Under the weak field approximation, Eq. 1.28 becomes:

$$\frac{d^2 \xi^i}{d^2 t} = \frac{1}{2} \eta^{ik} \frac{\partial^2 h_{kj}}{\partial t^2} \xi^j, \tag{1.29}$$

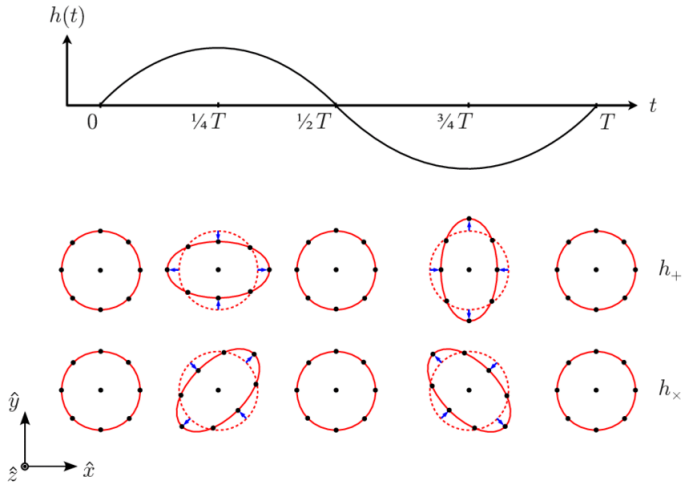

Figure 1.1: Evolution with time of the amplitude of the GW, h, (top). Effect of the two polarization $h_{+,\times}$, of a GW which propagates along the z-axis, on a ring of free fall test masses (bottom). Figure adapted from Ref. [117]

where we use only Latin index because GWs do not affect time component of spacetime (see Eq. 1.26). For simplicity, we consider the case of GW propagating along the z-axis, so that in TT gauge the only non-trivial components are:

$$\begin{aligned} h_+ = h_{xx} = -h_{yy} = 2A_+ \cos(\omega t - k_z z), \\ h_\times = h_{xy} = h_{yx} = 2A_\times \cos(\omega t - k_z z), \end{aligned} \tag{1.30}$$

where we use Eq. 1.22 with $k_\alpha = (\omega, 0, 0, \omega)$. Approximating $\xi^i = \xi^i_0 + \delta\xi^i(t)$, with ξ^i_0 constant, from Eq. 1.29 we obtain the variation in time of separation along each axis:

$$\begin{aligned} \delta\xi^x &= \xi^x_0 + \frac{1}{2}(h_{xy}\xi^y_0 + h_{xx}\xi^x_0), \\ \delta\xi^y &= \xi^y_0 + \frac{1}{2}(h_{yz}\xi^z_0 + h_{yy}\xi^y_0), \\ .\delta\xi^z &= \xi^z_0. \end{aligned} \tag{1.31}$$

Using Eq. 1.31, we observe the polarization effect, represented in Fig. 1.1. In a general coordinate system, the polarization tensors are defined as:

$$\begin{aligned} e^+_{ab} &= \hat{p}^a\hat{p}^b - \hat{q}^a\hat{q}^b, \\ e^\times_{ab} &= \hat{p}^a\hat{q}^b + \hat{p}^b\hat{q}^a, \end{aligned} \tag{1.32}$$

where $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$ form an orthonormal basis perpendicular to $\hat{\boldsymbol{k}}$. The propagation vector $\hat{\boldsymbol{k}}$ alone does not uniquely determine the polarization tensors, since any pair of orthonormal vectors can be chosen as the basis. To account for this freedom, we introduce the polarization angle ψ, which defines a rotation between the source frame $(\hat{\boldsymbol{p}}, \hat{\boldsymbol{q}})$ and an arbitrary

basis $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}})$. Under this rotated basis, the polarization tensors transform as:

$$\begin{aligned} e^{+}_{ab} &= \epsilon^{+}_{ab} \cos 2\psi - \epsilon^{\times}_{ab} \sin 2\psi, \\ e^{\times}_{ab} &= \epsilon^{+}_{ab} \sin 2\psi + \epsilon^{\times}_{ab} \cos 2\psi. \end{aligned} \tag{1.33}$$

1.2.1 Energy content of the GW

Although GWs are ripples in spacetime itself, they do carry energy and momentum. The energy transported by a GW can be quantified using the *Isaacson stress-energy tensor*, which arises in the context of general relativity when one considers perturbations on a background spacetime.

To derive the GW energy content, we expand the Einstein tensor to second order in a small perturbation $h_{\mu\nu}$ on a background metric $\bar{g}_{\mu\nu}$:

$$G_{\mu\nu} = G^{(0)}_{\mu\nu}[\bar{g}] + G^{(1)}_{\mu\nu}[h] + G^{(2)}_{\mu\nu}[h] + \dots \tag{1.34}$$

Here, $G^{(1)}_{\mu\nu}$ is linear in the perturbation and governs the wave propagation, as we have seen so far, while $G^{(2)}_{\mu\nu}$ contains terms quadratic in $h_{\mu\nu}$ that can be interpreted as describing the energy and momentum carried by waves themselves.

In vacuum, the full Einstein tensor must vanish. After averaging over several wavelengths (to smooth out rapid oscillations), the second-order terms contribute as an effective source to the background curvature:

$$G^{(0)}_{\mu\nu}[\bar{g}] = -\left\langle G^{(2)}_{\mu\nu}[h] \right\rangle . \tag{1.35}$$

This motivates the definition of the Isaacson effective stress-energy tensor for GWs as

$$T^{\mathrm{GW}}_{\mu\nu} \equiv \frac{1}{8\pi G} \left\langle G^{(2)}_{\mu\nu}[h] \right\rangle . \tag{1.36}$$

In the TT gauge, Eq. 1.36 reads:

$$T^{\mathrm{GW}}_{\mu\nu} = \frac{c^4}{32\pi G} \left\langle \partial_\mu h^{\mathrm{TT}}_{ij}\, \partial_\nu h^{\mathrm{TT}\,\alpha\beta} \right\rangle . \tag{1.37}$$

For a plane GW propagating in the z-direction, the energy flux is given by:

$$S^z = c\, T^{\mathrm{GW}}_{0z} = \frac{c^3}{32\pi G} \left\langle \dot{h}^{\mathrm{TT}}_{ij}\, \partial_z h^{\mathrm{TT}\,\alpha\beta} \right\rangle . \tag{1.38}$$

Using the fact that the wave travels at the speed of light, the spatial and time derivatives are related by

$$\partial_z h^{\mathrm{TT}}_{ij} = -\frac{1}{c} \dot{h}^{\mathrm{TT}}_{ij}, \tag{1.39}$$

so the energy flux carried away by GW becomes

$$\frac{dE}{dA dt} = \frac{c^2}{32\pi G} \left\langle \dot{h}^{\mathrm{TT}}_{ij}\, \dot{h}^{\mathrm{TT}}_{ij} \right\rangle = \frac{c^2}{16\pi G} \langle \dot{h}^2_{+} + \dot{h}^2_{\times} \rangle. \tag{1.40}$$

By considering the surface element of a sphere at distance r from the source, $dA = r^2 d\Omega$, where $d\Omega$ is the solid angle element, we express Eq. 1.40 in terms of power:

$$\frac{dE}{dt} = \frac{c^2 r^2}{16\pi G} \int d\Omega \langle \dot{h}^2_{+} + \dot{h}^2_{\times} \rangle. \tag{1.41}$$

Starting from Eq. 1.40, we further obtain the energy spectrum in frequency domain:

$$\frac{dE}{df} = \frac{\pi c^3}{2G} f^2 r^2 \int d\Omega \langle |\tilde{h}_+(f)|^2 + |\tilde{h}_\times(f)|^2 \rangle. \tag{1.42}$$

where $\tilde{h}_{+,\times}(f)$ are the Fourier Transforms (FTs) of the plus and cross polarizations.

1.3 GW from binaries

The general solution of Eq. 1.20 with a non-vanishing $T_{\mu\nu}$, at time t and position r, is given by:

$$\overline{h}_{\mu\nu}(t, r) = \frac{4G}{c^4 r} \int_V T_{\mu\nu}(t - r/c, x^i) d^3x. \tag{1.43}$$

This solution is simplified employing the conservation laws for $T_{\mu\nu}$:

$$\nabla_\nu T^{\mu\nu} = 0 \rightarrow \frac{\partial T^{\mu 0}}{\partial t} = -\frac{\partial T^{\mu k}}{\partial x^k}. \tag{1.44}$$

Integrating Eq. 1.44 over a source volume V and using Gauss' theorem, we rewrite the conservation law as:

$$\frac{1}{c}\frac{\partial}{\partial t} \int_V T^{\mu 0} d^3x = -\int_V \frac{\partial T^{\mu k}}{\partial x^k} d^3x = -\int_S T^{\mu k} dS_k, \tag{1.45}$$

where S is the surface bounding the volume V and dS_k is the surface element. The integral over the surface S can be neglected if we assume that the source is localized in a region of space, so that the surface integral vanishes at infinity, i.e. $\int_S T^{\mu k} dS_k = 0$. Therefore, the left-hand side of Eq. 1.45 vanishes and it follows that:

$$\overline{h}_{\mu 0} = 0, \tag{1.46}$$

in agreement with the fact GW does not influence the time component of the metric.
To obtain the spatial components of the metric perturbation, we consider the tensor-virial theorem, deriving from the spatial component of the conservation law. The theorem reads:

$$\frac{1}{c^2}\frac{\partial^2}{\partial^2 t} \int_V T^{00} x^k x^n d^3x = 2 \int_V T^{kn} d^3x. \tag{1.47}$$

We define the mass quadrupole moment as follows:

$$Q_{kn} \coloneqq \int_V \rho(t, x^i) x^k x^n d^3x = \frac{1}{c^2} \int_V T^{00}(t, x^i) x^k x^n d^3x. \tag{1.48}$$

In the slow-motion approximation, where internal velocities of the source are small compared to the speed of light, the dominant component of the energy-momentum tensor is $T_{00} = \rho c^2$. This justifies the second equality in Eq. 1.48. Therefore, combining Eqs. 1.43, 1.47, and 1.48, we write the solution for the spatial components of the metric perturbation as:

$$\overline{h}_{ij}(t, r) = \frac{2G}{c^4 r} \frac{d^2}{d^2 t} (Q_{ij}(t - r/c)). \tag{1.49}$$

Eq. 1.49 tells us that GWs propagation is possible when the second time derivative of Q_{ij} is non-vanishing. The mass quadrupole moment encodes how the mass is spread relative to the coordinate axes and its distribution asymmetry. It vanishes for spherically symmetric systems and remains constant for systems with purely static or symmetric motion, such as uniform rotation. Thus, only time-varying, non-spherical mass distributions contribute to gravitational radiation.
The factor $2G/c^4$ is extremely small (SI units: $1.7 \times 10^{-44}\,\mathrm{s}^2\,\mathrm{kg}^{-1}\,\mathrm{m}^{-1}$), reflecting the weak-perturbation nature of GWs. Consequently, the second time derivative of the mass quadrupole moment must be extremely large to produce a significant GW signal.

In Section 1.2, we saw that TT gauge is particularly useful. In this gauge, the mass quadrupole moment is defined as

$$Q_{ij}^{TT} = P_{ijmn} Q_{mn} - \frac{1}{3}\delta_{ij} Q_m^m, \tag{1.50}$$

where $Q_m^m = \sum_{i=1}^{3} Q_{ii}$ is the mass quadrupole moment trace, and P_{ijmn} is the TT projection operator, defined as

$$P_{ijmn} = P_{im}P_{jn} - \frac{1}{2}P_{ij}P_{mn}, \tag{1.51}$$

where $P_{ij} = \delta_{ij} - k_i k_j$ is the transverse operator, with propagation vector $\hat{\boldsymbol{k}}$. Thus, substituting Eq. 1.50 into Eq. 1.49, the spatial components of the metric perturbation become

$$h_{ij}^{TT}(t, r) = \frac{2G}{c^4 r}\frac{d^2}{d^2 t}(Q_{ij}^{TT}(t - r/c)). \tag{1.52}$$

1.3.1 GW from circular binaries

In the previous section, we derived a general expression for the GW emission from an arbitrary source We now turn our attention to an astrophysically relevant example: a circular binary system composed of two point-like masses.

The binary has component masses m_1 and m_2, with total mass $m = m_1 + m_2$, reduced mass $\mu = m_1 m_2/m$, and symmetric mass ratio $\eta = \mu/m$. The initial separation is $a = r_1 + r_2$, where $r_1 = m_2 a/m$ and $r_2 = m_1 a/m$ are the distances of the two objects from the center of mass.

In our approximation, the binary orbital frequency is the Keplerian one ω_k, related to the separation through Kepler's law:

$$\omega_k = \sqrt{\frac{Gm}{a^3}}. \tag{1.53}$$

Without loss of generality, we place the binary in the x-y plane. The coordinates of $m_{1,2}$ are then:

$$\begin{pmatrix} x_{1,2}(t) \\ y_{1,2}(t) \end{pmatrix} = \pm r_{2,1} \begin{pmatrix} \cos(\omega_k t) \\ \sin(\omega_k t) \end{pmatrix}, \tag{1.54}$$

Given this configuration, the binary energy density is

$$T_{00} = c^2 \sum_{i=1}^{2} m_i \delta(x - x_i)\delta(y - y_i)\delta z. \tag{1.55}$$

The non-vanishing components of the mass quadrupole moment are

$$Q_{ij} = \sum_{i=1}^{2} m_i \int_V x_i^i x_j^i \, \delta(x - x_i) \, \delta(y - y_i) \, \delta(z) \, dx \, dy \, dz. \tag{1.56}$$

Eq. 1.56 is evaluated by substituting the coordinates from Eq. 1.54:

$$\begin{aligned} Q_{xx} &= \frac{1}{2}\mu a^2 \cos(2\omega_k t), \\ Q_{yy} &= -\frac{1}{2}\mu a^2 \cos(2\omega_k t), \\ Q_{xy} &= \frac{1}{2}\mu a^2 \sin(2\omega_k t), \end{aligned} \tag{1.57}$$

Defining the matrix A_{ij} as:

$$A_{ij} = \begin{pmatrix} \cos(2\omega_k t) & \sin(2\omega_k t) & 0 \\ \sin(2\omega_k t) & -\cos(2\omega_k t) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.58}$$

we rewrite the mass quadrupole moment as:

$$Q_{ij} = \frac{\mu}{2} A_{ij} a^2. \tag{1.59}$$

To compute the GW in the TT gauge, we must project the mass quadrupole moment onto the TT plane. This projection is performed using the operator P_{ijmn} defined in Eq. 1.51. Specifically, for a wave propagating along $\hat{\boldsymbol{k}}$, the metric perturbation in the TT gauge is obtained as shown in Eq. 1.52. Finally, the h_+ and $h_\times$ polarization components of the metric perturbation in the TT gauge are written as:

$$\begin{aligned} h_+(t) &= A_{\mathrm{GW}} \frac{1 + \cos^2 \iota}{2} \cos \phi(t), \\ h_\times(t) &= A_{\mathrm{GW}} \cos \iota \sin \phi(t), \end{aligned} \tag{1.60}$$

with ι is the inclination angle between the line of sight and binary orbital angular momentum. When the binary is edge-on ($\iota = \pi/2$), the observer receives the plus polarization only. Instead, for a face-on binary ($\iota = 0$), both polarizations with equal magnitudes are observed. The amplitude A_{GW} and phase $\phi(t)$ are given by:

$$A_{\mathrm{GW}} = \frac{4}{r}(\pi f_{\mathrm{GW}})^{2/3} \frac{(G\mathcal{M}_c)^{5/3}}{c^4}, \tag{1.61}$$

$$\phi(t) = 2\pi f_{\mathrm{GW}} t + \phi_0, \tag{1.62}$$

where we express the result in terms of $f_{\mathrm{GW}} = \omega_{\mathrm{GW}}/(2\pi)$. In Eqs. 1.61 and 1.62, we introduced two key variables:

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \tag{1.63}$$

$$f_{\mathrm{GW}} = 2 f_k. \tag{1.64}$$

We see from Eq. 1.64 that the GW frequency is twice the Keplerian orbital frequency of the binary itself.
By substituting Eq. 1.60 into Eq. 1.41, we obtain the power radiated by the binary in GWs:

$$\frac{dE_{\rm GW}}{dt} = \frac{32}{5}\frac{G^{7/3}}{c^5}(\mathcal{M}_c \pi f_{\rm GW})^{10/3}. \tag{1.65}$$

The binding energy of the binary is $E_B = -\frac{Gm\mu}{2a}$. By applying energy conservation, the GW luminosity is related to the rate of change of the binding energy:

$$\frac{dE_{\rm GW}}{dt} = -\frac{dE_B}{dt}. \tag{1.66}$$

From Eq. 1.66, using Kepler's law in Eq. 1.53, we derive an evolution equation for the semi-major axis a:

$$\frac{da}{dt} = -\frac{64}{5}\frac{G\mu m^2}{c^5 a^3}. \tag{1.67}$$

The negative sign in Eq. 1.67 shows that the semi-major axis shrinks with the emission of GWs. The rate of change of the semi-major axis is used to estimate the time until the binary merges due to GWs emission. Using $f_{\rm GW} \propto a^{-3/2}$, the GW frequency evolution is

$$\frac{df_{\rm GW}}{dt} = \frac{96}{5}\pi^{8/3}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{5/3} f_{\rm GW}^{11/3}. \tag{1.68}$$

Eq. 1.68 shows that $\mathcal{M}_c$ controls how quickly the GW frequency rises, i.e., how fast the signal *"chirps"*, giving the chirp mass its name.

1.3.2 GW from eccentric binaries

In the previous section, we have seen how to characterize aGW signal from a circular binary. However, in the universe, binaries are often observed in eccentric orbits. Considering the generic case of an eccentric binary, Ref. [142] has shown how the orbit parameters evolve due to GWs:

$$\frac{da}{dt} = -\frac{64}{5}\frac{G^3\mu M^2}{c^5}a^{-3}F(e), \tag{1.69}$$

$$\frac{de}{dt} = -\frac{304}{15}\frac{G^3\mu M^2}{c^5}a^{-4}(1-e^2)^{-5/2}(1+\frac{121}{304}e^2), \tag{1.70}$$

where $F(e)$ is a function of the eccentricity:

$$F(e) = (1-e^2)^{-7/2}(1+\frac{73}{24}e^2+\frac{37}{96}e^4). \tag{1.71}$$

Eq. 1.69 shows, as for the circular binary case, that the binary tightens as a consequence of GWs emission. Eq. 1.70 shows in addition that the eccentricity decreases over time, circularizing the binary. The energy released during the inspiral of an eccentric binary can be expressed as:

$$\dot{E}_{\rm GW;ecc} = \dot{E}_{\rm GW;circ}F(e), \tag{1.72}$$

where $E_{\rm GW;circ}$ is given by Eq. 1.65. The function $F(e)$ increases rapidly with eccentricity, for instance, $F(e=0.9) \sim 300$ while $F(e=0.99) \sim 10^6$. Eq. 1.72 suggests that eccentric

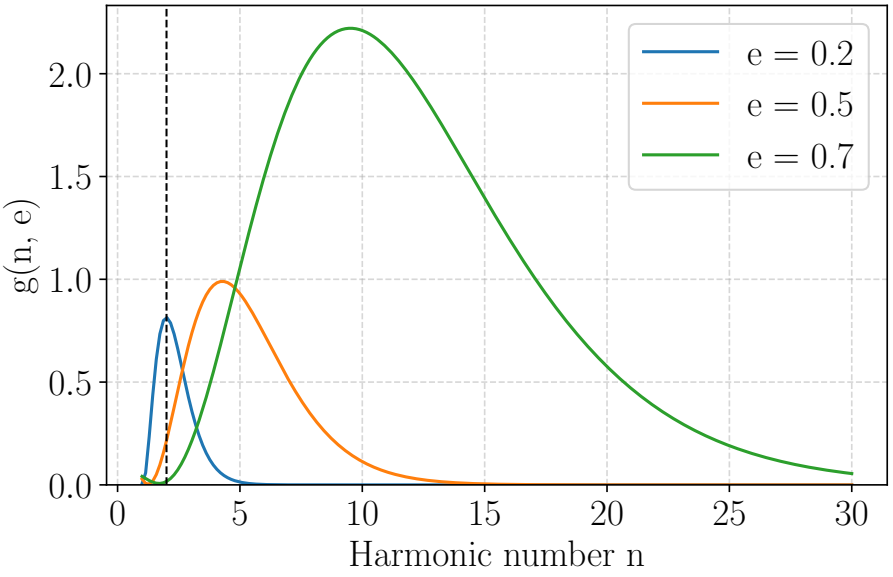

Figure 1.2: Function $g_n(e)$ for different values of eccentricity. The function quantifies how much power of the signal is radiated through each harmonic. As the eccentricity increases, the power gets spread towards higher harmonics.

binaries shrink much faster than circular systems with the same semi-major axis. Moreover, while the GW frequency is twice the orbital frequency for circular binaries, for eccentric binaries multiple harmonics are excited at $f_n = nf_{\rm k}$, with n being the harmonic number. Thus, the total emitted power is distributed over a broad spectrum of frequencies, with the power fraction in each harmonic given by

$$\dot{E}_n = \frac{32G^{7/3}}{5c^5}(2\pi\mathcal{M}_c f_k)^{10/3} g_n(e). \tag{1.73}$$

The dimensionless function $g_n(e)$ controls how much power signal goes in each harmonic through the relation:

$$\begin{aligned} g_n(e) =& \frac{n^4}{32}[(J_{n-2}(ne) - 2eJ_{n-1}(ne) + \frac{2}{n}J_n(ne) + \\ &+2eJ_{n+1}(ne) - J_{n+2}(ne))^2 + \\ &+(1-e^2)(j_{n-2}(ne) - 2J_n(ne) + J_{n+2}(ne))^2 + \\ &+\frac{4}{3n^2}J_n^2(ne)]. \end{aligned} \tag{1.74}$$

Here, J_n represents the n-th order of the first kind Bessel function. For circular binaries, the only non-vanishing harmonic is the second one, i.e. $g_n(e=0) = \delta_{n,2}$, where $\delta_{n,m}$ is the Kronecker delta function. In Fig. 1.2 we show the function $g_n(e)$ for different values of eccentricity.

1.3.3 Fast Emri Waveform

EMRIs are relativistic binaries composed of a stellar-mass compact object, such as a Black Hole (BH) or a neutron star, orbiting around a Massive Black-Hole (MBH). The typical mass-ratio of these systems is between 10^{-7} and 10^{-4}. The formation mechanism underlying these systems is yet not fully understood, though a viable process is the capture of the light object by the light one in dense stellar environments, such as globular clusters or nuclear star clusters [14]. Their eccentricity is expected to be significant, and thus the GW signal emitted by these systems is characterized by a broad spectrum of frequencies, as discussed in the previous section. Due to the extremely small mass ratio, the secondary Compact Object (CO) can be approximately treated as a test particle moving in the background spacetime of the primary BH. To account for the reaction of the secondary's own gravitational field on its motion, one can use the gravitational self-force approach, solving Einstein equations perturbatively in powers of the mass ratio [133]. This method is computationally expensive, making it impractical to generate a large number of accurate waveforms required for EMRI data analysis. To overcome this limit, approximate families of waveforms have been constructed to capture the key features essential for signal detectability. One such approach involves constructing Post-Newtonian (PN) waveforms, which are advantageous due to their analytic nature and ease of evaluation. PN theory provides a systematic method to approximate GR by solving the Einstein field equation perturbatively in powers of the small parameter (v/c). A special class of such approximations is represented by the so-called kludge waveforms. Three main types have been proposed in literature: the Analytic Kludge [34], Numerical Kludge [30], and the 5PN Augmented Analytic Kludge (AAK) [103]. The AAK model evolves the orbital parameters using a fifth-order PN expansions for small eccentricities. The (quadrupolar $l = m = 2$) amplitudes are constructed using a weak-field Peters and Matthews approximation [143] to the metric perturbation. In Sections 4.1 and 5.1, we will use the AAK waveform model as implemented in the `FastEMRIWaveforms` package, following Refs. [66, 34].

To fully characterize an EMRI waveform, 14 parameters must be specified. The central object is described by its mass M and dimensionless spin magnitude $a = |\boldsymbol{S}|c/(GM^2)$, where $\boldsymbol{S}$ is the spin angular momentum. The secondary compact object, of mass μ, is typically assumed to have negligible spin. The spin orientation of the MBH is described by the unit vector $\hat{\boldsymbol{S}}$, defined with respect to the Solar System Barycenter (SSB), and expressed in terms of polar and azimuthal angles (θ_K, ϕ_K) in ecliptic coordinates. The orbital angular momentum of the compact object is described by the vector $\boldsymbol{L}$, and the angle between $\boldsymbol{L}$and $\hat{\boldsymbol{S}}$ defines the orbital inclination ι. The orbits followed by the CO can be both eccentric and non-equatorial. In this context, the phases Φ_r, Φ_θ, and Φ_ϕ denote the initial radial, polar, and azimuthal phases of the motion. The eccentricity e and semi-latus rectum p characterize the orbital shape. From these, the pericenter and apocenter distances are given by $r_p = p/(1+e)$ and $r_a = p/(1-e)$, respectively. In addition to the geodesic parameters, the source position is specified by the polar and azimuthal angles (θ_S, ϕ_S) in ecliptic coordinates, along with its luminosity distance D_L, defined as the distance relating the intrinsic (source-frame) GW luminosity to the observed strain amplitude, accounting for the redshift and expansion of the Universe.

1.4 SGWB

While the previous section focused on the GW emitted by a single source, in this section we discuss the SGWB, a stochastic process generated by the incoherent superposition ofGWs. The SGWB can be generated by early universe processes, such as inflation, phase transitions, or cosmic strings [58, 27], or by the incoherent superposition of GWs emitted by a large number of unresolved astrophysical sources [81]. In the TT gauge, the SGWB metric perturbation is written as a linear superposition of plane waves, each characterized by a frequency f and a direction $\hat{\mathbf{k}}$:

$$h_{ij}(t,\mathbf{x}) = \sum_{p=+,\times} \int_{-\infty}^{\infty} df \int d\hat{\mathbf{k}}\, \tilde{h}_p(f,\hat{\mathbf{k}}) e^p_{ij}(\hat{\mathbf{k}}) e^{i2\pi f(t-\hat{\mathbf{k}}\cdot\mathbf{x}/c)}, \tag{1.75}$$

1.4.1 Statistical properties

The SGWB Fourier components, $\tilde{h}_p(f,\hat{\mathbf{k}})$, are *random fields*, meaning their values fluctuate randomly across frequency f, sky direction $\hat{\mathbf{k}}$, and polarization p. The *statistical properties* of this field are described by its *moments*, which are expectation values of products of the field. More generally, the *n-th moment* is defined as

$$M^{(n)}(\{(p_s,f_s,\hat{\mathbf{k}}_s)\}_{s\in S}) = \Big\langle \prod_{s\in S} \tilde{h}_{p_s}(f_s,\hat{\mathbf{k}}_s) \Big\rangle, \tag{1.76}$$

where S is an abstract finite index set with $|S| = n$. Without loss of generality, we assume the field has zero mean,

$$\Big\langle \tilde{h}_p(f,\hat{\mathbf{k}}) \Big\rangle = 0, \tag{1.77}$$

so that, in the case of a Gaussian background, the second moment ($n = 2$) fully characterizes its statistical properties. Specifically, the angle brackets denote the ensemble average, defined as the average over many realizations of the stochastic process. Since we observe only a single realization of the Universe, the ensemble average cannot be directly computed from data. In practice, under the ergodic hypothesis, the ensemble average is approximated by a time average over the observed data. In addition to ergodicity, a number of properties are often assumed to describe SGWBs:

- Stationarity: SGWB statistical properties do not change over time. A SGWB is expected to form over astrophysical or cosmological timescales, so that the time variation of the SGWB are negligible on human timescales. This implies that the two-point correlation $\langle h(t)h(t')\rangle$ in time depends only on the time lag $t-t'$. Consequently, in the frequency domain, different components are uncorrelated:

$$\langle \tilde{h}_p(f,\hat{\mathbf{k}}) \tilde{h}^*_{p'}(f',\hat{\mathbf{k}}')\rangle \propto \delta(f-f'). \tag{1.78}$$

- Isotropy: SGWB statistical properties do not depend on the direction in the sky. Similarly to the cosmic microwave background, the SGWB is expected to be approximately isotropic as well. Therefore, the two-point correlation function is independent of the direction $\hat{k}$, and therefore we write:

$$\langle \tilde{h}_p(f,\hat{\mathbf{k}}) \tilde{h}^*_{p'}(f',\hat{\mathbf{k}}')\rangle \propto \delta^2(\hat{\mathbf{k}},\hat{\mathbf{k}}'). \tag{1.79}$$

- Polarization: the SGWB is the incoherent, uncorrelated superposition of two processes, one for each polarization. This implies that the two-point correlation function is diagonal in the polarization space:

$$\langle \tilde{h}_p(f, \hat{\mathbf{k}}) \tilde{h}^*_{p'}(f', \hat{\mathbf{k}}') \rangle \propto \delta_{pp'}. \tag{1.80}$$

- Gaussianity: the SGWB is a Gaussian random field, which implies that all the higher-order moments can be expressed in terms of the two-point correlation function. Gaussianity is supported by the central limit theorem, which states that the sum of a large number of independent and identically distributed random variables tends to a Gaussian distribution, provided that the individual variables have finite variance.

As a direct consequence of the above assumptions, the SGWB is characterized only by its two-point correlation function, defined as:

$$\langle \tilde{h}_p(f, \hat{\mathbf{k}}) \tilde{h}^*_{p'}(f', \hat{\mathbf{k}}') \rangle = \frac{1}{4\pi} \delta_{pp'} \delta(f - f') \delta^2(\hat{\mathbf{k}}, \hat{\mathbf{k}}') \frac{1}{2} S_h(f), \tag{1.81}$$

where $S_h(f)$ is the one-sided PSD of the SGWB. PSD will be better defined in the next section in Eq. 2.10 and Eq. 2.11 The factor $1/2$ in Eq. 1.81 is due to the fact that we are considering a one-sided PSD, which is defined only for positive frequencies. The factor $1/4\pi$ is a normalization factor that ensures that the SGWB is normalized to unity when integrated over all directions in the sky. From a cosmological point of view, the SGWB is often described in terms of $\Omega_{\rm GW}(f)$, which is defined as:

$$\Omega_{\rm GW}(f) = \frac{1}{\rho_c} \frac{d\rho_{\rm GW}}{d \log f}, \tag{1.82}$$

where $\rho_c = 3H_0^2/(8\pi G)$ is the Universe critical energy density. Eq. 1.82 represents the critical energy density fraction of the Universe contained in GWs at a given frequency f. Leveraging the Isaacson relation for the energy density, we relate SGWB PSD to the energy density as:

$$S_h(f) = \frac{H_0^2}{4\pi^2 f^3} \Omega_{\rm GW}(f). \tag{1.83}$$

1.4.2 Practical application

So far, we have studied the emission and propagation of GWs under the assumption that the distance r between the source and the observer is sufficiently small such that the Universe expansion can be neglected. However, with the increasing sensitivity of current and upcoming GW detectors, we are now able to observe GWs emitted by sources at cosmological distances. The key cosmological effect on the GW frequency is encoded by the redshift z of the source, which stretches the observed waveform in time. Specifically, the observed frequency is related to the emitted frequency by the source f_s through

$$f = \frac{f_s}{1+z}, \tag{1.84}$$

where $1 + z$ is the ratio between the cosmic scale factor at signal reception and that at emission. Then, the total spectral energy density $\Omega_{\rm GW}(f)$ from coalescing binary systems in the universe can be expressed as [13, 146, 3]

$$\Omega_{\rm GW}(f) = \frac{f}{\rho_c H_0} \iint dz d\boldsymbol{\phi} \frac{R_{\rm GW}(z)}{(1+z)\mathcal{E}(z)} p(\boldsymbol{\phi}) \left. \frac{dE_{\rm GW}}{df_s}(\boldsymbol{\phi}) \right|_{f_s}, \tag{1.85}$$

where $\frac{dE_{\rm GW}}{df_s}(\boldsymbol{\phi})|_{f_s}$ is the source-frame energy spectrum radiated by a single source, evaluated at the source GW frequency f_s. In Eq. 1.85, the integration is performed over the distribution $p(\boldsymbol{\phi})$ of source parameters $\boldsymbol{\phi}$ (e.g., masses, spins, etc.). The quantity $R_{\rm GW}(z)$ denotes the comoving merger rate density of GW sources, measured per unit source-frame time and comoving volume, and $\mathcal{E}(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$. For a binary with source-frame component masses m_1 and m_2, the leading-order PN expression for the energy spectrum of the emitted gravitational radiation at source-frame frequency f_s is given by [146, 126]

$$\frac{dE_{\rm GW}}{df_s} \equiv \frac{\dot{E}_{\rm GW}}{\dot{f}_s} = \frac{\eta m^{5/3}\pi^{2/3}}{3f_s^{1/3}}, \tag{1.86}$$

where the chirp rate $\dot{f}_s$ is given by

$$\frac{df_s}{dt} = \frac{96}{5} f_s^{11/3} m^{5/3} \pi^{8/3} \eta. \tag{1.87}$$

From Eqs. 1.85 and 1.86, the frequency-independent contribution to the SGWB is absorbed into an overall amplitude $A_{\rm vac}$, yielding

$$\Omega_{\rm GW}(f) = A_{\rm vac} f^{2/3}. \tag{1.88}$$

Thus, the energy density per logarithmic frequency interval from a population of inspiral binaries is expected to follow a power-law with slope 2/3 ($-7/6$ in terms of PSD). This concludes our discussion of the theoretical framework for GWs generation and propagation; in the next chapter, we turn to the observational implications and techniques used to detect and characterize such signals.

Chapter 2

GW data analysis

In this chapter, we introduce in Section 2.1 the fundamental concepts of signal analysis, essential to understand how GWs are detected. Then, in Section 2.2, we discuss the Bayesian inference framework, which is extensively used in this thesis to analyse LISA data.

2.1 Fundamentals of signal analysis

In Section 1.2 we showed that the metric perturbation $h_{ij}(t, \mathbf{x})$ is decomposed into two independent polarizations: plus and cross (Eqs. 1.32 and 1.33). These polarizations distort the spacetime by squeezing and squashing it. In Section 1.2 we showed that the metric perturbation $h_{ij}(t, \mathbf{x})$ can be decomposed into two independent polarizations, plus and cross (Eqs. 1.32 and 1.33). These polarizations distort spacetime by periodically stretching and squeezing it in the plane orthogonal to the propagation direction. Interferometers are designed to measure the phase difference of light caused by this distortion. These measurements are fundamentally local: the detector responds to variations in the proper distance between freely falling test masses, as inferred from light travel times. In the TT gauge, this corresponds to variations in the spatial components h_{ij}, since the time and mixed components vanish ($h_{00} = h_{0i} = 0$). To describe the detector response, the polarization tensors must be projected onto the detector arms, whose geometric configuration is encoded in the detector tensor D_{ij}. The measured signal is thus:

$$h(t) = F_+(\hat{\mathbf{k}})\, h_+(t) + F_\times(\hat{\mathbf{k}})\, h_\times(t), \tag{2.1}$$

where F are the antenna pattern functions, defined as:

$$F_+(\hat{\mathbf{k}}) = D^{ij}\, e^+_{ij}(\hat{\mathbf{k}}), \tag{2.2}$$

$$F_\times(\hat{\mathbf{k}}) = D^{ij}\, e^\times_{ij}(\hat{\mathbf{k}}). \tag{2.3}$$

At any given time, the antenna pattern functions depend on the source sky position and the polarization basis orientation. In Section 3.2.2, we will discuss in detail how to project a signal onto LISA detector arms, taking into account the constellation motion.

2.1.1 Matched filtering

In a realistic setting, a GW signal projected onto a detector as in Eq. 2.1 is superimposed with some detector noise. The total output $d(t)$ reads:

$$d(t) = n(t) + h(t). \tag{2.4}$$

Instrumental noise originates from multiple sources—such as residual acceleration, optical path length fluctuations, and laser frequency variations—depending on the experimental setup and facility. In Section 3.3, we discuss the noise contributions relevant to LISA in more detail. Under idealised conditions, the detector noise is a zero mean, Gaussian, and stationary stochastic process. Hence, each time series is considered as a single realization drawn from the corresponding Gaussian distribution.

$$n(t) \sim \mathcal{N}(0, \Sigma(t, t')). \tag{2.5}$$

As mentioned in Section 1.4, stationarity implies that the two point correlation function depends only on the lag between two times:

$$\Sigma(t, t') = \langle n(t)n(t')\rangle = R(t - t') = R(\tau), \tag{2.6}$$

where $R(\tau)$ defines the autocorrelation function. The FT of $n(t)$ is given by

$$\tilde{n}(f) = \int_{-\infty}^{\infty} \mathrm{d}t\ n(t)e^{-2\pi\imath ft}. \tag{2.7}$$

Being the FT a linear operator, the process is readily cast in frequency domain as a circular Gaussian one:

$$\Re[n(f)], \Im[n(f)] \sim \mathcal{N}(0, S_n(f)), \tag{2.8}$$

where $S_n(f)$ denotes the PSD and $\Re, \Im$ denote the real and imaginary part of a complex number, respectively. For a stationary stochastic process, the PSD is defined as

$$S_n(f) = \lim_{T\to+\infty} \frac{1}{2T} \left\langle \left| \int_{-T}^{+T} \mathrm{d}t\ n(t)e^{-2\pi\imath ft} \right|^2 \right\rangle. \tag{2.9}$$

There is a well-defined relation between the autocorrelation function and the one-sided PSD for a stationary process:

$$S_n(f) = \int_0^{\infty} R(\tau)e^{-2\pi\imath f\tau} d\tau. \tag{2.10}$$

Eq. 2.10 is known as the Wiener-Khinchin theorem.

Following Eqs. 2.6 and 2.10, we express the covariance function in frequency domain as:

$$\langle \tilde{n}^*(f)\tilde{n}(f')\rangle = \frac{1}{2} S_n(f)\delta(f - f'), \tag{2.11}$$

For $f = f'$, the Dirac delta diverges in a distributional sense. However, from a realistic point of view, observations are made over a finite time interval T, which leads to a finite frequency resolution. Therefore, we approximate the Dirac delta as:

$$\delta_T(f) \approx \int_{-T/2}^{T/2} dt\, e^{-i2\pi ft} = \frac{\sin(\pi fT)}{\pi f}, \tag{2.12}$$

therefore resolving the apparent divergence. The goal of GW data analysis is to detect and infer on a GW signal $h(t)$ from the detector noise $n(t)$. This is achieved through a well-established framework known as matched filtering, which applies a filter function to the datastream. Here, we show that the filter function, maximising the SNR, matches the expected signal template. The SNR is generally defined as the ratio between the expectation value of a detection statistic under the signal-present hypothesis (H_1) and its root-mean-square under the noise-only hypothesis (H_0). A *detection statistic* is a scalar functional of observed data, quantifying the presence of a signal. Formally, for a detector output $d(t)$, a detection statistic $\hat{s}$ is defined as

$$\hat{s} = \hat{s}[d(t)]\,. \tag{2.13}$$

The corresponding SNR is then

$$\mathrm{SNR} = \frac{\mathbb{E}[\hat{s} \mid H_1]}{\sqrt{\mathrm{Var}[\hat{s} \mid H_0]}}\,. \tag{2.14}$$

For our target, we define $\hat{s}$ as the zero-lag cross-correlation between the detector output $s(t)$ and a filter function $K(t)$:

$$\hat{s} = \int d(t)\, K(t)\, dt\,. \tag{2.15}$$

The expectation value of $\hat{s}$ reads:

$$\begin{aligned}
\langle \hat{s} \rangle_{H_1} &= \left\langle \int_{-\infty}^{\infty} dt\, d(t) K(t) \right\rangle \\
&= \int_{-\infty}^{\infty} dt\, \langle d(t) \rangle K(t) && (2.16)\\
&= \int_{-\infty}^{\infty} dt\, \left[h(t) + \langle n(t) \rangle\right] K(t) && (2.17)\\
&= \int_{-\infty}^{\infty} dt\, h(t) K(t) && (2.18)\\
&= \int_{-\infty}^{\infty} df\, \tilde{h}(f) \tilde{K}^*(f), && (2.19)
\end{aligned}$$

where we have used the linearity in Eq. 2.16, the zero-mean condition for the noise in Eq. 2.18, and Parseval's theorem in Eq. 2.19. When the signal is not present, the root-mean-square of $\hat{s}$ is given by:

$$\begin{aligned}
\left[\left(\langle \hat{s}^2 \rangle - \langle \hat{s} \rangle^2\right)_{H_0}\right]^{1/2} &= \left[\langle \hat{s}^2 \rangle\right]^{1/2} && (2.20)\\
&= \left[\int_{-\infty}^{\infty} dt dt'\, K(t) K(t') \langle n(t) n(t') \rangle\right]^{1/2} \\
&= \left[\int_{-\infty}^{\infty} dt dt'\, K(t) K(t') \int_{-\infty}^{\infty} df df'\, e^{2\pi \imath f t - 2\pi \imath f' t'} \langle \tilde{n}^*(f) \tilde{n}(f') \rangle\right]^{1/2} && (2.21)\\
&= \left[\int_{-\infty}^{\infty} df\, \frac{1}{2} S_n(f) |\tilde{K}(f)|^2\right]^{1/2} && (2.22)
\end{aligned}$$

where we use the definition of PSD from Eq. 2.11. Given Eq. 2.19 and Eq. 2.20, we define the SNR as:

$$\begin{aligned}\text{SNR} &= \frac{\int_{-\infty}^{\infty} df\, \tilde{h}(f)\tilde{K}^*(f)}{\left[\int_{-\infty}^{\infty} df\, \frac{1}{2}S_n(f)|\tilde{K}(f)|^2\right]^{1/2}} \\ &= \frac{\left(\frac{1}{2}S_n(f)\tilde{K}(f)|\tilde{h}(f)\right)}{\left(\frac{1}{2}S_n(f)\tilde{K}(f)|\frac{1}{2}S_n(f)\tilde{K}(f)\right)^{1/2}},\end{aligned} \tag{2.23}$$

where we have used the inner product, defined as:

$$(A|B) = \Re \int_{-\infty}^{\infty} df\, \frac{\tilde{A}^*(f)\tilde{B}(f)}{\frac{1}{2}S_n(f)} = 4\Re \int_0^{\infty} df\, \frac{\tilde{A}^*(f)\tilde{B}(f)}{S_n(f)}, \tag{2.24}$$

The Wiener filter $\tilde{K}(f)$ is chosen to maximize the SNR. From the Cauchy-Schwarz inequality, the SNR is maximized when:

$$\tilde{K}(f) = \frac{\tilde{h}(f)}{S_n(f)}. \tag{2.25}$$

In order to construct the Wiener filter, we need to know the PSD of the noise $S_n(f)$ and a template $\tilde{h}(f)$ for the signal we are looking for. In a realistic setup, the signal model depends on uncertain parameters (e.g., masses, spins, sky location). In practice, we construct a bank of templates spanning the parameter space and search for the template that maximizes the SNR, thereby identifying both the presence of a signal and its most likely parameters. Plugging Eq. 2.25 into Eq. 2.23, we get

$$\text{SNR} = (\tilde{h}(f)|\tilde{h}(f))^{1/2} = \left[4\int_0^{\infty} df\, \frac{|\tilde{h}(f)|^2}{S_n(f)}\right]^{1/2}. \tag{2.26}$$

The SNR is commonly used in data analysis to assess whether a signal is present in the data. A signal is considered detected if the SNR exceeds a given threshold, which is set to achieve a desired false alarm probability—the probability of incorrectly claiming a detection when only noise is present. Sources with SNR values below this threshold remain undetected individually and accumulate, forming an SGWB.

2.1.2 Matched filtering for SGWB

We now turn our attention to the detection of a SGWB signal, as introduced in Section 1.4. As we shall see in Section 3.2.3, for a GW detector, the response to a SGWB is obtained by integrating the metric perturbation in Eq. 1.75 over all sky directions, weighted by the antenna pattern functions $F^p(f, \hat{\mathbf{k}})$ defined in Eq. 2.2 and Eq. 2.3. The resulting detector signal is given by:

$$h(t) = \int df \int d\hat{\mathbf{k}} \sum_{p=+,\times} \tilde{F}^p(f, \hat{\mathbf{k}})\tilde{h}_p(f, \hat{\mathbf{k}})e^{i2\pi f(t-\hat{\mathbf{k}}\cdot\mathbf{x}/c)}, \tag{2.27}$$

In the frequency domain, the signal becomes:

$$\tilde{h}(f) = \sum_{p=+,\times} \int d\hat{\mathbf{k}}\, \tilde{F}^p(f, \hat{\mathbf{k}})\tilde{h}_p(f, \hat{\mathbf{k}})e^{-i2\pi f\hat{\mathbf{k}}\cdot\mathbf{x}/c}. \tag{2.28}$$

As discussed in Section 1.4, a SGWB is characterized by its statistical properties rather than the deterministic signal content. In order to compare our model for the SGWB statistical properties against those measured from the data, we cross-correlate signals from multiple detectors. Considering two detectors with antenna pattern functions $\tilde{F}_1(f, \hat{\mathbf{k}})$ and $\tilde{F}_2(f, \hat{\mathbf{k}})$, the cross-correlation between the two signals is given by

$$\langle \tilde{h}_1(f) \tilde{h}_2^*(f') \rangle = \frac{1}{2} \delta(f - f') \Gamma_{12}(f) S_h(f), \tag{2.29}$$

where $S_h(f)$ is the SGWB PSD defined in Eq. 1.83, and $\Gamma_{12}(f)$ is the overlap reduction function, which is defined as:

$$\Gamma_{12}(f) = \frac{1}{8\pi} \sum_{p=+,\times} \int d\hat{\mathbf{k}}\, \tilde{F}_1^p(f, \hat{\mathbf{k}}) \tilde{F}_2^{p*}(f, \hat{\mathbf{k}}) e^{-i2\pi f \hat{\mathbf{k}} \cdot (\mathbf{x}_1 - \mathbf{x}_2)/c}, \tag{2.30}$$

Similarly to Section 2.1, we now define the matched filter for a SGWB signal. We define the outputs $d_{1,2}$ from the two detectors as in Eq. 2.4, where the detector noises $n_{1,2}$ are uncorrelated between each other, i.e. $\langle n_1(t) n_2(t') \rangle = 0$. The analogous statistic $\hat{s}$ for the SGWB signal is:

$$\hat{s} = \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt'\, d_1(t) d_2(t') K(t, t'), \tag{2.31}$$

where we consider a finite time interval T for the observation. The expectation value of $\hat{s}$ under H_1 hypothesis is given by:

$$\begin{aligned} \langle \hat{s} \rangle_{H_1} &= \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt'\, \langle d_1(t) d_2(t') \rangle K(t, t') \\ &= \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt'\, \langle h_1(t) h_2(t') \rangle K(t, t'). \end{aligned} \tag{2.32}$$

In contrast with Eq. 2.19, the SGWB signal is not a deterministic template, but a stochastic process. Therefore, we need to average over the different realizations. Since we are assuming in this section that both SGWB and noise are stationary, the best choice for the filter $K(t, t')$ will depend on the lag between the two times, i.e. $K(t, t') = K(t - t')$. Then, Eq. 2.32 equals:

$$\begin{aligned} \langle \hat{s} \rangle_{H_1} &= \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt'\, \langle h_1(t) h_2(t') \rangle K(t - t') \\ &= \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt' \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df'\, e^{i2\pi f t - i2\pi f' t'} \langle \tilde{h}_1(f) \tilde{h}_2^*(f') \rangle K(t - t') \\ &= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df'\, \langle \tilde{h}_1(f) \tilde{h}_2^*(f') \rangle \int_{-T/2}^{T/2} dt e^{i2\pi t (f - f')} \int_{-T/2}^{T/2} dt' K(t') e^{i2\pi f' t'} \\ &= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \langle \tilde{h}_1(f) \tilde{h}_2^*(f') \rangle \tilde{K}(f') \int_{-T/2}^{T/2} dt\, e^{i2\pi t (f - f')} \\ &= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df'\, \langle \tilde{h}_1(f) \tilde{h}_2^*(f') \rangle \tilde{K}(f') \delta_T(f - f') \\ &= T \int_{-\infty}^{\infty} df \Gamma_{12}(f) S_h(f) \tilde{K}(f). \end{aligned} \tag{2.33}$$

To get the final result, we consider Eq. 2.29 and Eq. 2.12 to approximate the Dirac delta function. This yields the factor T that appears in Eq. 2.33. The noise contribution, N, is given by

$$\begin{aligned}(\langle\hat{s}^2\rangle - \langle\hat{s}\rangle^2)_{H_0} &= \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt' \int_{-T/2}^{T/2} dt'' \int_{-T/2}^{T/2} dt''' \, \langle n_1(t) n_2(t') n_1(t'') n_2(t''')\rangle \times \\ &\times K(t-t')K(t''-t''') \\ &= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \int_{-\infty}^{\infty} df'' \int_{-\infty}^{\infty} df''' \, \langle \tilde{n}_1(f)\tilde{n}_1^*(-f'')\rangle \langle \tilde{n}_2(-f')\tilde{n}_2^*(f''')\rangle \times \\ &\times \tilde{K}(f')\tilde{K}(f''')\delta(f-f')_T \delta(f''-f''')_T \\ &= \frac{1}{4}\int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \, \delta_T^2(f-f') S_{1,n}(f) S_{2,n}(f') \tilde{K}(f)\tilde{K}^*(f') \\ &= \frac{T}{4}\int_{-\infty}^{\infty} df \, S_{1,n}(f) S_{2,n}(f) |\tilde{K}(f)|^2, \end{aligned} \tag{2.34}$$

where we replaced one of the finite-time delta functions $\delta_T(f-f')$ by an ordinary Dirac delta function, and evaluated the other at $f = f'$ to obtain the last line. The functions $S_{1,n}(f)$ and $S_{2,n}(f)$ are the noise PSDs of the two detectors. Thus, we define the SNR as:

$$\begin{aligned}\mathrm{SNR} &= \frac{T\int_{-\infty}^{\infty} df \Gamma_{12}(f) S_h(f) \tilde{K}(f)}{\left[\frac{T}{4}\int_{-\infty}^{\infty} df \, S_{1,n}(f) S_{2,n}(f) |\tilde{K}(f)|^2\right]^{1/2}} \\ &= \frac{T\left(\tilde{K}(f) S_{1,n} S_{2,n} | \Gamma_{12} S_h(f)\right)}{\frac{\sqrt{T}}{2}\left(\tilde{K}(f) S_{1,n}(f) S_{2,n}(f) | \tilde{K}(f) S_{1,n}(f) S_{2,n}(f)\right)^{1/2}}, \end{aligned} \tag{2.35}$$

where, similarly to Eq. 2.24, we have used the inner product defined as:

$$(A|B) = \Re \int_{-\infty}^{\infty} df \, \frac{\tilde{A}^*(f)\tilde{B}(f)}{S_{n,1}(f) S_{n,2}(f)} = 2\Re \int_0^{\infty} df \, \frac{\tilde{A}^*(f)\tilde{B}(f)}{S_{n,1}(f) S_{n,2}(f)}. \tag{2.36}$$

The Wiener filter $\tilde{K}(f)$ is chosen again to maximize the SNR:

$$\tilde{K}(f) = \frac{\Gamma_{12}(f) S_h(f)}{S_{1,n}(f) S_{2,n}(f)}. \tag{2.37}$$

Plugging Eq. 2.37 into Eq. 2.35, we get:

$$\mathrm{SNR} = \left[4T \int_0^{\infty} df \, \frac{(\Gamma_{12}(f) S_h(f))^2}{S_{1,n}(f) S_{2,n}(f)}\right]^{1/2}. \tag{2.38}$$

The square-root dependence on T suggests that, even if the signal PSD is smaller than the noise PSD in each detector, it is still possible to overcome a given detection threshold by integrating the correlated output over a sufficiently long observation time. Moreover, Eq. 2.38 can be generalized to a network of p detectors as follows:

$$\mathrm{SNR} = \left[4T \sum_{i,j=1}^{p} \int_0^{\infty} df \, \frac{(\Gamma_{ij}(f) S_h(f))^2}{S_{i,n}(f) S_{j,n}(f)}\right]^{1/2}. \tag{2.39}$$

We remark that Eq. 2.39 represents the optimal statistic for a detector network under the assumption of Gaussian, stationary, and mutually uncorrelated noise.

2.1.3 Sensitivity

The expression in Eq. 2.38 highlights that a detector sensitivity to a signal depends not only on its noise level, but also on how it couples to the signal. To account for both contributions, it is useful to introduce an effective PSD, $S_n^{\rm eff}(f)$. For a single detector, we define the *sky- and polarization-averaged response function* as $R(f) = \Gamma_{11}(f)$. The effective PSD is then given by:

$$S_n^{\rm eff}(f) = \frac{S_{n,1}(f)}{R(f)}. \tag{2.40}$$

In the case of a single deterministic source, this effective PSD provides a meaningful measure of sensitivity: the squared SNR is directly proportional to the area under the ratio between the signal PSD and $S_n^{\rm eff}(f)$. In other words, the frequency domain contribution to the SNR can be read from the area where the signal PSD stands above the effective noise PSD. However, in the case of a SGWB, Eq. 2.40 does not fully capture the detector sensitivity, since the SNR explicitly depends on the observation time, as shown in Eq. 2.38. We therefore introduce an *integrated sensitivity curve* to characterize the detector capability to observe a certain SGWB. Specifically, we consider a SGWB modeled by a power-law spectrum of the form:

$$\Omega_{\rm GW}(f) = \Omega_0 \left(\frac{f}{f_0}\right)^{\alpha}, \tag{2.41}$$

where Ω_0 is the amplitude at the reference frequency f_0, and α is the spectral index or slope. As discussed in Section 1.4.2, such a power-law shape naturally arises from a population of inspiraling compact binaries, but it also provides a useful model for various other sources, including cosmological backgrounds. Similarly to Eq. 2.40, and considering the relation in Eq. 1.83, we define the effective $\Omega_n^{\rm eff}(f)$ the one implicitly defined by:

$$\mathrm{SNR} = \left[4T \int_0^{\infty} df \left(\frac{\Omega_{\rm GW}(f)}{\Omega_n^{\rm eff}(f)}\right)^2\right]^{1/2}. \tag{2.42}$$

For a set of slopes α, the amplitude $\Omega_{0,\alpha}$ associated to a given SNR threshold ρ is expressed as

$$\Omega_{0,\alpha} = \frac{\rho}{\sqrt{4T}} \left[\int_0^{\infty} df \left(\frac{(f/f_0)^{\alpha}}{\Omega_n^{\rm eff}(f)}\right)^2\right]^{-1/2}. \tag{2.43}$$

Each pair $(\alpha, \Omega_{0,\alpha})$ defines a corresponding power-law model $\Omega_{\rm GW}(f)$. By taking the envelope of these power-law curves, we construct the power-law integrated sensitivity curve, also known as Power-Law Sensitivity (PLS):

$$\Omega_{\rm PLS}(f) = \max_{\alpha} \left[\Omega_{0,\alpha} \left(\frac{f}{f_0}\right)^{\alpha}\right] \tag{2.44}$$

As an example, we show the PLS together with the effective PSD in Fig. 2.1 for a LISA-type detector network. Details on the response function and instrumental noise PSD are provided in Sections 3.2.3 and 3.3. The interpretation of the PLS is now well grounded: any power-law SGWB that lies above the PLS curve is expected to produce a SNR exceeding the chosen threshold ρ.

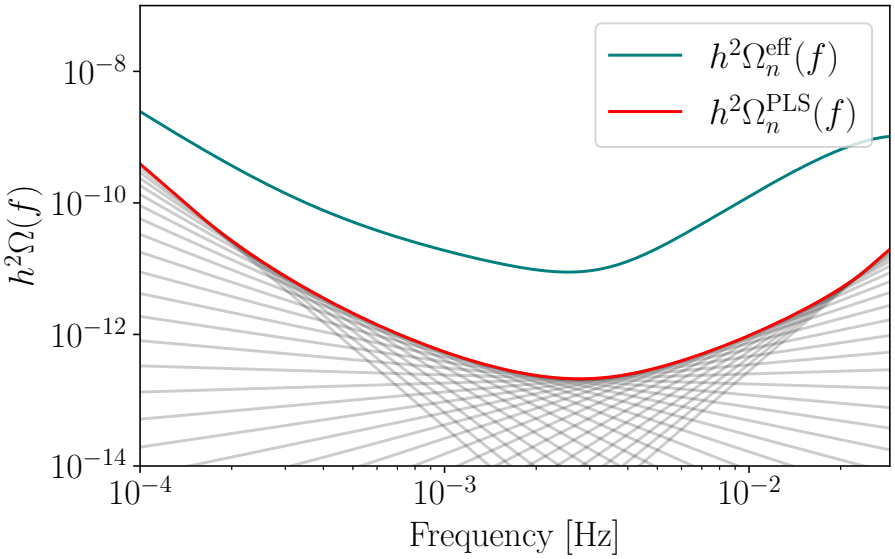

Figure 2.1: PLS (red line) and effective PSD (teal line) for a LISA-type detector. The PLS is computed for an observation time of $T = 4$yr and a SNR threshold $\rho = 10$. The black lines represent the family of power-law spectra used to construct the PLS as their upper envelope. As expected, the PLS lies below the effective PSD, reflecting the enhanced sensitivity obtained by integrating over the observation time and by taking the envelope across different spectral slopes.

2.2 Bayesian inference

The matched filtering technique described in Section 2.1 is a powerful method to identify optimal statistics in a purely frequentist context, assuming the presence of a known signal in the data. However, in practice, we do not know *a priori* whether a signal is present, nor the values of its parameters. Moreover, since GW observations are non-repeatable single events, a Bayesian statistical framework is particularly well-suited for inference and parameter estimation in GW astronomy. Below, we summarize the key concepts of Bayesian inference. Before observing new data, we encode our existing knowledge about the system in a *prior* distribution—for example, a probability distribution over source parameters $\boldsymbol{\theta}$. Once data d become available, we update the prior using a *likelihood* function, which quantifies how likely the observed data are, as a function of the model free parameters. Bayesian inference combines prior knowledge with observations to produce the *posterior* distribution, which encapsulates our updated beliefs about the parameters $\boldsymbol{\theta}$. This update is performed using Bayes' theorem:

$$p(\boldsymbol{\theta}|d) = \frac{\mathcal{L}(d|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})}{\mathcal{Z}(d)}, \tag{2.45}$$

where:

- $\pi(\boldsymbol{\theta})$ is the prior distribution: it encodes our beliefs about the parameters $\boldsymbol{\theta}$ before observing the data from previous observations, or theoretical considerations.
- $\mathcal{L}(d|\boldsymbol{\theta})$ is the likelihood function.
- $\mathcal{Z}(d)$ is the normalization constant of the posterior distribution, known as the *evidence* (or marginalized likelihood), given by:

$$\mathcal{Z}(d) = \int \mathcal{L}(d|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,d\boldsymbol{\theta}. \tag{2.46}$$

2.2.1 Likelihood

To compute the likelihood function, we need to define the model for the data. As mentioned in Section 2.1, we consider the noise to be a stationary Gaussian process described by the PSD $S_n(f)$ (see Eq. 2.11). The probability density function of a noise realization n_0 reads:

$$p(n_0) = \mathcal{N} \exp\left\{-\frac{1}{2}4\int_0^\infty \frac{|\tilde{n}_0(f)|^2}{S_n(f)}\,df\right\}, \tag{2.47}$$

where $\mathcal{N}$ is a normalization factor which depends on the PSD. With the inner product introduced in Eq. 2.24, the probability is equivalently recast as:

$$p(n_0) = \mathcal{N} \exp\left\{-\frac{1}{2}(n_0|n_0)\right\}, \tag{2.48}$$

If a signal described by a template $h(\boldsymbol{\theta})$ is present in the data, produced by a source with true parameters $\boldsymbol{\theta}_t$, then the detector output d is given by Eq. 2.4:

$$d = h(\boldsymbol{\theta}_t) + n_0. \tag{2.49}$$

In this case, the likelihood function for d can be written by substituting $n_0 = d - h(\boldsymbol{\theta})$ in Eq. 2.48:

$$\begin{aligned}\mathcal{L}(d|\boldsymbol{\theta}) = p(d - h(\boldsymbol{\theta})) &= \mathcal{N} \exp\left\{-\frac{1}{2}(d - h(\boldsymbol{\theta})|d - h(\boldsymbol{\theta}))\right\} \\ &= \mathcal{N} \exp\left\{(h(\boldsymbol{\theta})|d) - \frac{1}{2}h(\boldsymbol{\theta})|h(\boldsymbol{\theta}) - \frac{1}{2}(d|d)\right\},\end{aligned} \tag{2.50}$$

where the term $(d|d)$ is constant with respect $\boldsymbol{\theta}$ and can be ignored when performing parameter estimation. This function is known as the *Whittle* likelihood [137]. In particular, taking the logarithm of Eq. 2.50 is convenient, as it reduces to a linear combination of inner products:

$$\log \mathcal{L}(d|\boldsymbol{\theta}) \propto (h(\boldsymbol{\theta})|d) - \frac{1}{2}(h(\boldsymbol{\theta})|h(\boldsymbol{\theta})). \tag{2.51}$$

Now, we consider the scenario where multiple sources overlap with each other and/or are too faint to be detected individually. In this case, it happens that we are not able to distinguish all the sources or resolve them. Then a stochastic signal, such as the SGWB, emerges in the data, and we decompose them as follows:

$$d = \sum_{i=1}^{n_d} h_i(\boldsymbol{\theta}_{i,t}) + n_0 + h_{0,s}, \tag{2.52}$$

where $h_i(\boldsymbol{\theta}_{i,t})$ denotes the i-th source template with true parameters $\boldsymbol{\theta}_{i,t}$, n_0 is the noise, and $h_{0,s}$ is the SGWB signal. As discussed in Section 1.4, for stochastic signals we instead characterize their statistical properties. Thus, we introduce the template $S_h(f, \boldsymbol{\theta}_h)$ to describe the PSD in Eq. 1.83. Since we are effectively modeling the variance in the Gaussian probability distribution, we can no longer neglect $\mathcal{N}$ and $(d|d)$ in Eq. 2.50. Moreover, we also want to simultaneously infer on some uncertain noise properties from the data. This is particularly relevant in experiments like LISA, where the true noise level cannot be determined in advance. Thus, for a finite set of frequencies, we define the likelihood function as

$$\begin{aligned}\log \mathcal{L}(d|\boldsymbol{\theta}) = -\frac{1}{2} \sum_{f=f_{min}}^{f_{max}} \sum_{i=1}^{n_d} \Bigg(&\frac{|\tilde{d}(f) - \tilde{h}_i(f, \boldsymbol{\theta}_{i,t})|^2}{S_n(f, \boldsymbol{\theta}_n) + R(f)\, S_h(f, \boldsymbol{\theta}_h)} \\ &+ \log(S_n(f, \boldsymbol{\theta}_n) + R(f)\, S_h(f, \boldsymbol{\theta}_h)) + \log(2\pi) \Bigg),\end{aligned} \tag{2.53}$$

where $R(f)$ is the response function defined in Eq. 2.40 for a single detector.

As we will discuss in Sections 3.5 and 3.6, this is a realistic scenario for the LISA mission, where multiple sources—both stochastic and individually resolvable—are expected to be present. Disentangling all these components is known as the *global fit* problem [123, 181, 104, 73]. This constitutes an extremely challenging task, mostly because the number of resolvable sources, n_d, is unknown, and it must be treated as a free parameter in the inference process. For the purpose of this thesis, we assume that the number of resolvable sources is known. These sources are characterized and subtracted using an iterative subtraction procedure, as described in Section 4.1.3. Therefore, we assume that all non-stochastic sources have been perfectly subtracted, and we set $h_i(f, \boldsymbol{\theta}_{i,t}) = 0$ in Eq. 2.51.

For a network of p detectors, we denote their time-domain datastreams as $\mathbf{d} = \{d_1, \ldots, d_p\}$. These datastreams represent outputs from a variety of detectors, which for GW detectors could be TDI channels from space-based interferometers [186] (discussed in Section 3.4), pulsar timing residuals [193], or strain measurements from ground-based detectors [5]. The likelihood is generalized to model the correlations between datastreams:

$$\log \mathcal{L}(\mathbf{d}|\boldsymbol{\theta}) = -\frac{1}{2} \sum_{f=f_{min}}^{f_{max}} \left(\mathbf{d}^\dagger(f) C^{-1}(f, \boldsymbol{\theta}) \mathbf{d}(f) - \frac{1}{2} \log |C(f, \boldsymbol{\theta})| - \frac{p}{2} \log(2\pi) \right), \tag{2.54}$$

where $\mathbf{d}(f)$ is the vector of FT datastreams, and $C(f, \theta)$ is the covariance matrix defined as

$$C(f, \theta) = C_n(f, \theta_n) + C_{\mathrm{GW}}(f, \theta_{\mathrm{GW}}). \tag{2.55}$$

Computing the likelihood can be computationally expensive, especially for SGWB, as it involves matrix inversions and handling large datasets due to the need to characterize the signal across the entire frequency band of interest. For this reason, it is often advantageous to work with averaged data rather than the full-resolution frequency series. We divide the total observation time into n_c adjacent segments, each of duration $T = T_{\mathrm{obs}}/n_c$. We assume the datastreams to be uniformly sampled, with a cadence Δt, and denote each segment by $\mathbf{d}_{(c)} = (d_{(c)1}, \ldots, d_{(c)p})$. We use the segments to construct coarse-grained data, and define the averaged periodograms in the frequency domain as

$$Y(f) = \frac{1}{n_c} \sum_{c=1}^{n_c} \tilde{\mathbf{d}}_{(c)}(f) \tilde{\mathbf{d}}^\dagger_{(c)}(f), \tag{2.56}$$

the matrix $Y(f)$ follows a complex Wishart distribution

$$\mathcal{L}(Y(f)|\Gamma(f)) \propto \frac{|Y(f)|^{n_c - p}}{|\Gamma(f)|^{n_c}} \exp\left[-\mathrm{tr}(\Gamma^{-1}(f) Y(f)) \right], \tag{2.57}$$

with n_c degrees of freedom and scale matrix

$$\Gamma(f) = \frac{1}{n_c} \left(C_n(f) + C_{\mathrm{GW}}(f) \right), \tag{2.58}$$

A lower-limit on n_c must also be enforced: $n_c > p - 1$, a defining property of the complex Wishart distribution. In the opposite regime, data in full frequency resolution can be employed, accounting for the multivariate normal distribution describing their likelihood.

2.2.2 Parameter estimation

We now want to estimate the most probable $\boldsymbol{\theta}$. A common approach is to find the values that maximize the posterior probability. Assuming a flat prior distribution, this is equivalent to maximizing the likelihood function. This approach is known as Maximum Likelihood Estimation (MLE). The MLE is defined as:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \mathcal{L}(d|\boldsymbol{\theta}). \tag{2.59}$$

The MLE provides point estimates of parameters, but it does not quantify the uncertainty associated. To do that, we employ the Fisher information matrix. It is defined as the

second derivative of log-likelihood function, evaluated at the MLE $\hat{\theta}$:

$$\mathcal{F}_{ij} = -\left.\frac{\partial^2 \log \mathcal{L}(d|\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right|_{\hat{\boldsymbol{\theta}}=\hat{\boldsymbol{\theta}}}. \tag{2.60}$$

In the high-SNR limit, the Fisher information matrix approximates the inverse covariance matrix of parameter estimates, since the likelihood function is well described by a Gaussian distribution centered on the MLE. The diagonal elements provide the inverse variances of each parameter, while the off-diagonal terms quantify their mutual correlations. Although this approach is computationally efficient, it remains accurate only under certain conditions. For example, estimating the SGWB is particularly challenging, as its amplitude is typically small compared to the noise level. Likewise, the Fisher matrix approximation can break down in more complex scenarios, such as when the likelihood is multi-modal or when parameters are non-linearly correlated. In these cases, more robust methods are necessary. For instance, in Chapters 4 and 5 we use two different methods to explore stochastically the likelihood function:

Nested Sampling: This technique was first proposed in [175]. The goal of the algorithm is to compute the evidence defined in Eq. 2.46. This becomes particularly challenging when posteriors exhibit complex features such as multimodality, non-Gaussian shapes, or high dimensionality, and especially when the likelihood concentrates in a tiny fraction of the prior volume. Direct Monte Carlo integration becomes prohibitively inefficient in these scenarios. Nested Sampling (NS) overcomes these difficulties by transforming the multidimensional integral Eq. 2.46 into a one-dimensional integral over the likelihood:

$$\mathcal{Z}(d) = \int_0^1 \mathcal{L}(X)\, dX, \tag{2.61}$$

where $X(\lambda)$ is the prior volume enclosed by the likelihood contour $\mathcal{L}(d|\boldsymbol{\theta}) > \lambda$:

$$X(\lambda) = \int_{\mathcal{L}(d|\boldsymbol{\theta})>\lambda} \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}. \tag{2.62}$$

To evaluate this numerically, the algorithm begins by sampling a set of points, called *live points*, from the prior distribution. At each iteration, the point with the lowest likelihood λ is removed and replaced by a new sample drawn from the prior subject to the constraint that the new sample has likelihood greater than λ (i.e., sampling from the prior restricted to the region where $\mathcal{L}(d|\boldsymbol{\theta}) > \lambda$). This process continues iteratively until a predefined stopping criterion is met. Since the integrand in Eq. 2.62 concentrates on regions of parameter space with high likelihood, the algorithm naturally generates samples that provide an accurate representation of the posterior distribution. We show in Fig. 2.2 the basic idea behind the NS algorithm. In the analysis carried out in chapters 4 and 5, we use NS as it is implemented in the `nessai` package [197].

Hamiltonian Monte Carlo: This Markov Chain Monte Carlo (MCMC) method exploits Hamiltonian dynamics to efficiently explore the parameter space, offering significant computational advantages over NS, particularly in high-dimensional problems. Specifically, we define the system phase space with position variables $\mathbf{q}$, which represent the parameters of interest, and momentum variables $\mathbf{p}$. The auxiliary momentum variables are sampled from a multivariate normal distribution with zero mean and covariance matrix M, named the *mass matrix*. This matrix rescales and decorrelates the momentum variables, effectively applying a linear transformation to the parameter space. When M^{-1} approximates

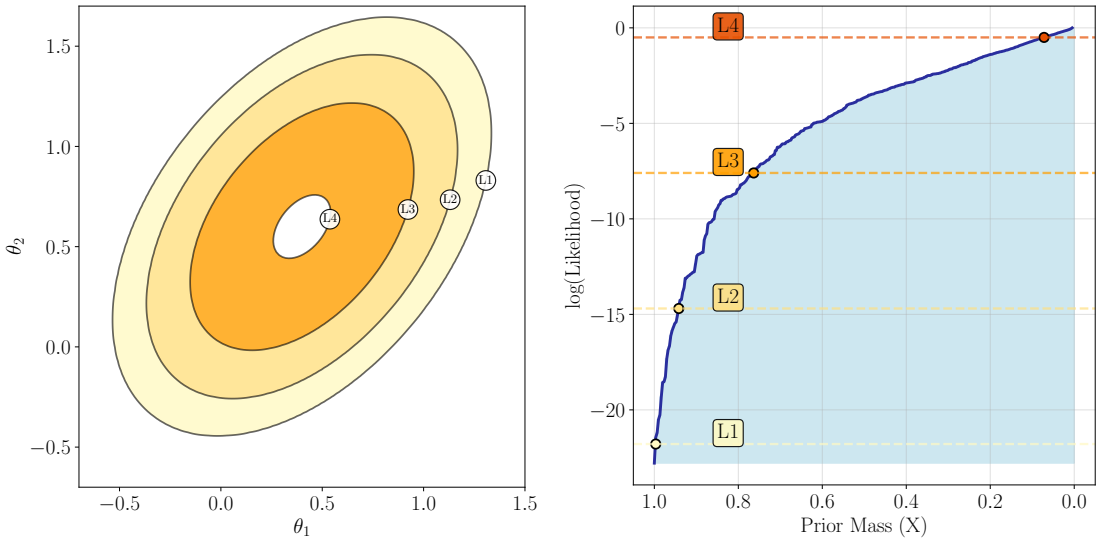

Figure 2.2: Illustration of NS algorithm. This example has been performed with `nessai`. (*Left Panel*) Posterior probability distribution for a two-dimensional problem, highlighting contours of equal likelihood. (*Right Panel*) Transformed $\mathcal{L}(X)$ curve, where each prior volume X_i is associated with its corresponding relative likelihood.

the covariance structure of the posterior, strongly correlated parameters become more independent, dramatically improving sampling efficiency. The system potential energy is defined as the negative of the posterior distribution. In this way, the Hamiltonian function is defined as

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &= U(\mathbf{q}) + K(\mathbf{p}) \\ &= -\log \mathcal{L}(d \mid \mathbf{q}) - \log p(\mathbf{q}) + \frac{1}{2}\mathbf{p}^\top M^{-1}\mathbf{p}. \end{aligned} \tag{2.63}$$

The dynamics of the system are governed by Hamilton's equations:

$$\frac{d\mathbf{q}}{dt} = \frac{\partial H}{\partial \mathbf{p}} = M^{-1}\mathbf{p}, \tag{2.64}$$

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{q}} = \nabla_{\mathbf{q}} \log \mathcal{L}(d \mid \mathbf{q}) + \nabla_{\mathbf{q}} \log p(\mathbf{q}). \tag{2.65}$$

Eqs. 2.64 and 2.65 are integrated numerically using the leapfrog method, a symplectic integrator that preserves phase-space volume and guarantees reversibility. Proposed states are then accepted or rejected according to the Metropolis–Hastings criterion, with acceptance probability:

$$\alpha = \min\left(1, \exp\left[H(\mathbf{q}_{\text{old}}, \mathbf{p}_{\text{old}}) - H(\mathbf{q}_{\text{new}}, \mathbf{p}_{\text{new}})\right]\right). \tag{2.66}$$

In Fig. 2.3, we illustrate the basic idea behind the Hamiltonian Monte Carlo (HMC) algorithm. In practice, we employ the Hamiltonian Monte Carlo (HMC) algorithm provided by `NumPyro` [144], which uses the No-U-Turn Sampler (NUTS) [96], an adaptive extension of HMC. Standard HMC requires manual specification of two parameters: the step size

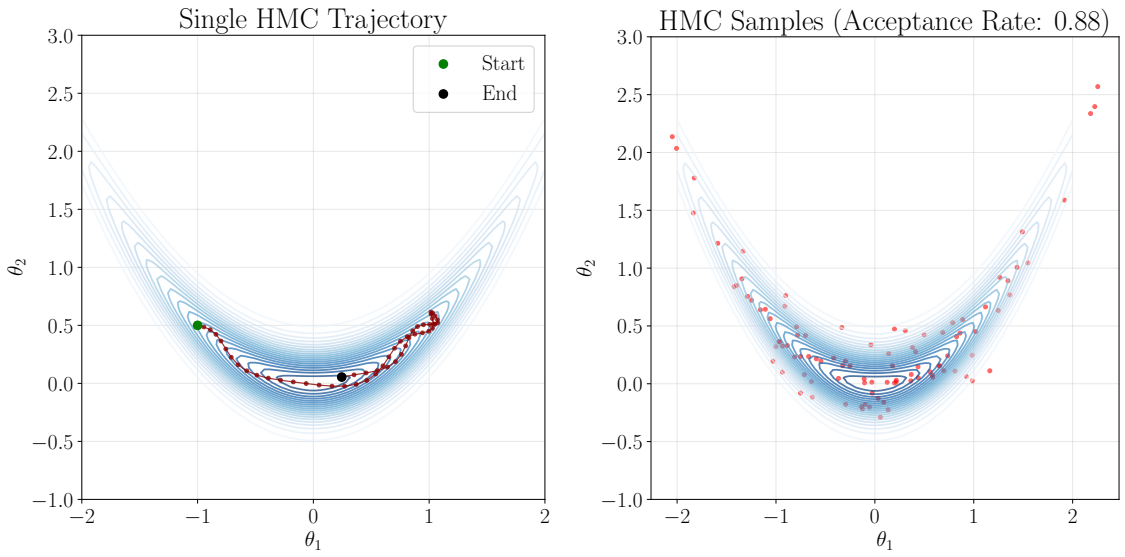

Figure 2.3: (*Left Panel*) Illustration of the Hamiltonian Monte Carlo algorithm. The left panel shows the trajectory of a single chain in the parameter space. (*Right Panel*) Samples obtained after running the algorithm for a given number of iterations. This example has been performed with `NumPyro`.

ε and the trajectory length, typically expressed as either a number of leapfrog steps L or a simulation time εL. The step size controls the discretization error in the leapfrog integrator—too large leads to inaccurate dynamics and low acceptance rates, while too small wastes computation. The trajectory length is even more problematic: too short results in random walk behavior, while too long causes the trajectory to loop back on itself, wasting computation by revisiting already-explored regions.

NUTS eliminates the need to set the trajectory length by automatically determining when to stop the simulation. It starts with a single leapfrog step, then progressively doubles the trajectory length (1, 2, 4, 8, ... steps) in randomly chosen directions (forward or backward in time) until stopping. The algorithm stops when it detects a "U-turn". Mathematically, this occurs when $(\boldsymbol{q}^+ - \boldsymbol{q}^-) \cdot \mathbf{p}^- < 0$ or $(\boldsymbol{q}^+ - \boldsymbol{q}^-) \cdot \mathbf{p}^+ < 0$, where $\boldsymbol{q}^{-,+}$ are the positions at the two ends of the trajectory and $\mathbf{p}^{-,+}$ are the corresponding momentum vectors. The next sample is then drawn from among all valid states visited, with appropriate weights to preserve detailed balance. The step size ε is automatically tuned during the warm-up phase to achieve a target acceptance rate. The `NumPyro` implementation additionally adapts a mass matrix M, estimating M from the sample covariance during warm-up. Crucially, HMC and NUTS require gradient information $\nabla_{\boldsymbol{q}} \log \mathcal{L}(d|\boldsymbol{q})$ at each leapfrog step. `NumPyro` leverages automatic differentiation through `jax` [46].

2.2.3 Evidence and model selection

The evidence in Eq. 2.46 is a scalar value, specifically a function solely dependent on the data. While it acts as a constant in parameter estimation, it is a crucial quantity

in Bayesian model selection. The goal of model selection is to determine which model is statistically preferred by the data and to what extent. The posterior support for one model M_1 over another M_2 is quantified by the odds ratio:

$$\frac{P(M_1|d)}{P(M_2|d)} = \mathcal{B}_{12}\,\frac{\pi(M_1)}{\pi(M_2)}, \tag{2.67}$$

Here, $\mathcal{B}_{12}$ represents the Bayes factor, corresponding to the evidence ratio between $\mathcal{Z}_{M_1}$ and $\mathcal{Z}_{M_2}$ for the two competing models. Without assuming any prior preference for either model, we set $\pi(M_1) = \pi(M_2) = 1$, so that the posterior odds ratio is equal to Bayes factor. In this regime, a $\log_{10}(\mathcal{B}_{12}) > 1$ is considered strong evidence in favor of model M_1 [99].

A practical application of the Bayes factor is to test the presence of a signal against the null hypothesis, which assumes no signal is present. If we denote with $\hat{\boldsymbol{\theta}}$ the maximum likelihood estimator of parameters we approximate the likelihood through the Laplace approximation:

$$\log \mathcal{L}(d|\boldsymbol{\theta}) \approx \log \mathcal{L}(d|\hat{\boldsymbol{\theta}}) - \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathcal{H}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}), \tag{2.68}$$

where the Hessian matrix $\mathcal{H}(\hat{\boldsymbol{\theta}})$ is equivalent to the Fisher information matrix defined in Eq. 2.60: Using Eq. 2.68, the log-evidence for the signal model is approximated by

$$\log \mathcal{Z}_1(d) \approx \log \mathcal{L}(d|\hat{\boldsymbol{\theta}}) + \frac{1}{2}\log\left(\frac{(2\pi)^k}{|\mathcal{H}(\hat{\boldsymbol{\theta}})|}\right), \tag{2.69}$$

where k is the number of free parameters in the model. The term

$$\mathcal{O}_c = \frac{1}{2}\log\left(\frac{(2\pi)^k}{|\mathcal{H}(\hat{\boldsymbol{\theta}})|}\right) \tag{2.70}$$

is often referred to as the Occam factor: it penalizes models with larger parameter spaces, favoring simpler explanations when the data do not support sufficiently additional complexity. For null hypothesis, we adopt a normalization choice where $\mathcal{Z}_0(d) = 1$, measuring the signal evidence relative to the noise-only baseline. Under this convention, the log-Bayes factor for signal versus noise is given by:

$$\begin{aligned}\log \mathcal{B}_{10} &= \log \mathcal{Z}_1(d) = \log \mathcal{L}(d|\hat{\boldsymbol{\theta}}) + \mathcal{O}_c \\ &= (\hat{h}|h_t) - \frac{1}{2}(\hat{h}|\hat{H}) + (O)_c \\ &= \frac{1}{2}(\hat{h}|\hat{h}) + \mathcal{O}_c.\end{aligned} \tag{2.71}$$

From the SNR definition in Eq. 2.26, it follows that the log-Bayes factor is proportional to the squared SNR. Although this derivation provides a Bayesian interpretation of the SNR, it is important to note that the SNR is not well-defined in Bayesian statistics, as it does not account for the prior information. Indeed, by deriving Eq. 2.71, we have assumed that the likelihood is strongly peaked around $\hat{\boldsymbol{\theta}}$. This approximation is valid in the high SNR regime, where data strongly constrain the parameters and the likelihood becomes increasingly concentrated around the true parameter values. In the low-SNR limit, the likelihood is broader and this approximation breaks down. Furthermore, evidence

integration is generally non-trivial, since the posterior may be dominated by the prior when the data are uninformative or when the prior heavily penalizes certain regions of parameter space.

As described in Section 2.2.2, NS algorithms are designed to compute the evidence directly. MCMC methods, instead, do not directly provide an estimate of the evidence; however, several techniques have been proposed to extract it from MCMC posterior chains, e.g., *power posterior method*. These methods use a series of intermediate distributions that gradually change the prior into the posterior. The most common way to do this is through the geometric path, which defines the power posterior as.

$$p_\beta(\boldsymbol{\theta}|d) = \frac{\mathcal{L}(d|\boldsymbol{\theta})^\beta \pi(\boldsymbol{\theta})}{\mathcal{Z}_\beta(d)}, \tag{2.72}$$

where β is a parameter that varies between 0 and 1. In particular, $\beta = 0$ corresponds to the prior distribution, while $\beta = 1$ corresponds to the posterior distribution. The normalization constant $\mathcal{Z}_\beta(d)$ is defined as:

$$\mathcal{Z}_\beta(d) = \int \mathcal{L}(d|\boldsymbol{\theta})^\beta \, \pi(\boldsymbol{\theta}) \, d\boldsymbol{\theta}. \tag{2.73}$$

Here, we summarize the most common methods.

Thermodynamic integration: Thermodynamical integration (TI) estimates the model evidence $\mathcal{Z}(d)$ by integrating along a path of intermediate distributions defined by the power posterior (Eq. 2.72). Specifically, the log-evidence reads:

$$\ln \mathcal{Z}(d) = \int_0^1 \mathbb{E}_\beta \left[\ln \mathcal{L}(d|\boldsymbol{\theta})\right] d\beta, \tag{2.74}$$

where $\mathbb{E}_\beta[\cdot]$ denotes the expectation value over the power posterior at β. This method is inspired by concepts from statistical thermodynamics: the parameter β is equivalent to the inverse temperature, and the likelihood raised to the power β corresponds to a Boltzmann factor. Then, the model evidence $\mathcal{Z}(d)$ plays as a partition function.

Stepping Stone: Stepping Stone (SS) relies on the same sampling scheme required by TI, but it typically requires fewer intermediate steps along the annealing path and produces a less biased estimate of evidence. Given a set of β_k values, where $0 = \beta_0 < \ldots < \beta_{K-1} < \beta_K = 1$, the evidence is expressed as a product of evidence ratios at consecutive β_k values:

$$\mathcal{Z}(d) = \frac{\mathcal{Z}_{\beta=1}(d)}{\mathcal{Z}_{\beta=0}(d)} = \prod_{k=1}^{K-1} \frac{\mathcal{Z}_{\beta_k}(d)}{\mathcal{Z}_{\beta_{k-1}}(d)}. \tag{2.75}$$

Each ratio is obtained using samples from the power posterior at β_{k-1}:

$$\frac{\mathcal{Z}_{\beta_k}(d)}{\mathcal{Z}_{\beta_{k-1}}(d)} = \frac{\int \mathcal{L}(d|\boldsymbol{\theta})^{\beta_k} \pi(\boldsymbol{\theta}) \, d\boldsymbol{\theta}}{\mathcal{Z}_{\beta_{k-1}}(d)} \tag{2.76}$$

$$= \frac{\int \mathcal{L}(d|\boldsymbol{\theta})^{(\beta_k - \beta_{k-1})} \mathcal{L}(d|\boldsymbol{\theta})^{\beta_{k-1}} \pi(\boldsymbol{\theta}) \, d\boldsymbol{\theta}}{\mathcal{Z}_{\beta_{k-1}}(d)} \tag{2.77}$$

$$= \frac{\mathcal{Z}_{\beta_{k-1}}(d) \int \mathcal{L}(d|\boldsymbol{\theta})^{(\beta_k - \beta_{k-1})} p_{\beta_{k-1}}(\boldsymbol{\theta}|d) \, d\boldsymbol{\theta}}{\mathcal{Z}_{\beta_{k-1}}(d)} \tag{2.78}$$

$$= \mathbb{E}_{\beta_{k-1}} \left[\mathcal{L}(d|\boldsymbol{\theta})^{\beta_k - \beta_{k-1}}\right], \tag{2.79}$$

where in Eq. 2.78 we have used the Bayes theorem. Thus

$$\mathcal{Z}(d) = \prod_{k=1}^{K-1} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(d|\boldsymbol{\theta}^i_{\beta_{k-1}})^{\beta_k - \beta_{k-1}}, \tag{2.80}$$

Generalized stepping stone: [201] Similarly to SS, the Generalized Stepping Stone (GSS) method estimates the evidence by computing a series of intermediate distributions that interpolate between the posterior distribution and a carefully chosen reference distribution π_0. In particular, π_0 is introduced in power posterior in Eq. 2.72 as follows:

$$p_\beta(\boldsymbol{\theta}|d) \propto (\mathcal{L}(d|\boldsymbol{\theta})\pi(\boldsymbol{\theta}))^\beta \pi_0(\boldsymbol{\theta})^{1-\beta}, \tag{2.81}$$

Once again, for $\beta = 0$, the reference distribution π_0 is equivalent to the prior distribution, while for $\beta = 1$, it corresponds to the posterior distribution. If $\pi_0 = \pi$ is the prior distribution, then GSS reduces to SS. Similarly to the SS, the evidence is computed as product of ratios which is defined as follows:

$$\frac{\mathcal{Z}_{\beta_k}(d)}{\mathcal{Z}_{\beta_{k-1}}(d)} = \mathbb{E}_{\beta_{k-1}} \left[\frac{\mathcal{L}(d|\theta)^{\beta_k - \beta_{k-1}}}{\pi_0} \right]. \tag{2.82}$$

To make GSS efficient, the reference distribution π_0 is chosen to be close to the posterior distribution, so that the expectation value is estimated with a small number of samples.

Chapter 3

Laser Interferometer Space Antenna

In this chapter, we discuss aspects of the LISA to address mission relevant for our studies. In Section 3.1, we give an overview of the mission design. In Section 3.2, we describe the LISA response to GWs. In Section 3.3, we discuss the main sources of instrumental noise. In Section 3.4, we introduce the TDI technique. In Section 3.5, we discuss GW sources observable by LISA.

3.1 Mission overview

LISA is a space-based gravitational wave observatory, selected as an L2-class large mission under the ESA Cosmic Vision program. It is designed to detect and measure GWs in the low-frequency band between 0.1mHz and 1Hz, otherwise inaccessible to ground-based detectors. With a nominal mission duration of 4 years, LISA consists of three identical spacecraft following heliocentric orbits, flying in a quasi-equilateral triangular configuration. As shown in Fig. 3.1, the constellation center of mass trails the Earth orbit at an angle of 20° while the constellation plane is inclined by around 60° with respect to the ecliptic plane.

LISA spacecrafts perform a clockwise cartwheeling motion with an orbital period of one year. The distances between the spacecrafts have an average value of 2.5×10^6 km and fluctuate by about $\sim 1\%$ over each orbit. The large armlength, combined with isolation from terrestrial seismic and Newtonian noise, allows LISA probing GWs at millihertz frequencies. To see this explicitly, an interferometer is most sensitive to gravitational waves whose wavelength $\lambda_{\rm GW}$ is comparable to its armlength L. For LISA case, we have

$$\lambda_{\rm GW} \sim L \sim 2.5 \times 10^9 \text{ m} \quad \Rightarrow \quad f_{\rm GW} \sim \frac{c}{\lambda_{\rm GW}} \sim \frac{3 \times 10^8 \text{ m/s}}{2.5 \times 10^9 \text{ m}} \sim 0.1 \text{ Hz}. \tag{3.1}$$

The orbits provide a constant angle of the constellation plane with respect to the Sun, allowing for a thermally stable environment and, more importantly, constant sunlight exposure to charge onboard equipment batteries. Furthermore, the heliocentric orbit keeps

the distance to Earth approximately constant, simplifying radio communication and data downlink by avoiding large variations in signal travel time and required transmitter power.

Each LISA spacecraft hosts two Optical Benches (OBs). Each OB contains a laser operating at a wavelength of 1064nm, an optical readout system, and a 40cm-diameter telescope used to transmit and receive light between the distant spacecraft. The OBs also contain a test mass—a 46mm platinum–gold cube with a mass of 2kg—designed to remain in near-perfect free fall along the laser propagation direction, shielded from external disturbances. In addition, the spacecrafts are equipped with phasemeters to analyse the readout signals and an ultra-stable oscillator that serves as a local clock for synchronization and timestamping. Micro-thrusters continuously adjust the spacecrafts position to follow the test masses, maintaining their free-fall motion their along their translational degrees of freedom.

To validate the mission design, the ESA launched the LISA PathFinder (LPF) mission in 2015 to test the technology required to place test masses in near-perfect free fall, where motion is influenced solely by gravity. LPF consisted of a single spacecraft implementing one of the LISA interferometer arms in a shortened configuration [23]. Operating around the Lagrange point L1 between December 2015 and July 2017, it showed that non-gravitational forces on the test masses could be suppressed to levels well below those required for LISA, exceeding the mission targets by more than an order of magnitude.[22, 24].

3.2 LISA response to GW

3.2.1 Convention

In this section, we introduce the conventions used to describe the LISA response to GWs. We define the *source frame*, previously introduced in Section 1.3, with an orthonormal Cartesian basis (x_S, y_S, z_S) (see left panel of Fig. 3.2). It is convenient to express vectors in spherical coordinates (θ_S, ϕ_S) relative to this basis, with associated unit vectors $(\hat{\boldsymbol{e}}_{S,r}, \hat{\boldsymbol{e}}_{S,\theta}, \hat{\boldsymbol{e}}_{S,\phi})$. The unit vector opposite the line of sight, expressed in the Cartesian basis, is

$$\hat{\boldsymbol{k}} = (\sin\theta_S \cos\phi_S, \sin\theta_S \sin\phi_S, \cos\theta_S). \tag{3.2}$$

Next, we introduce a polarization basis vectors:

$$\hat{\boldsymbol{p}} = \hat{\boldsymbol{e}}_{S,\theta}, \quad \hat{\boldsymbol{q}} = \hat{\boldsymbol{e}}_{S,\phi}. \tag{3.3}$$

Together with $\hat{\boldsymbol{k}}$, the vectors $(\hat{\boldsymbol{p}}, \hat{\boldsymbol{q}}, \hat{\boldsymbol{k}})$ form the radiation frame triad, used to define the polarization tensors as in Eq. 1.32.

Since the LISA constellation orbits the Sun, it is convenient to define a rigid reference frame for describing both the detector's motion and the source position. This is the SSB, aligned with the ecliptic plane and defined by the orthonormal basis vectors (x, y, z) (as illustrated in the right panel of Fig. 3.2). We adopt standard spherical coordinates (θ, ϕ) in the SSB frame, along with the corresponding spherical basis vectors $(\hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_\phi)$. The source sky location is parameterized by the *ecliptic latitude* $\beta = \pi/2 - \theta$ and *ecliptic longitude* $\lambda = \phi$. In these coordinates, the GW propagation vector $\hat{\boldsymbol{k}}$ is given by:

$$\hat{\boldsymbol{k}} = (-\cos\beta\cos\lambda, -\cos\beta\sin\lambda, -\sin\beta). \tag{3.4}$$

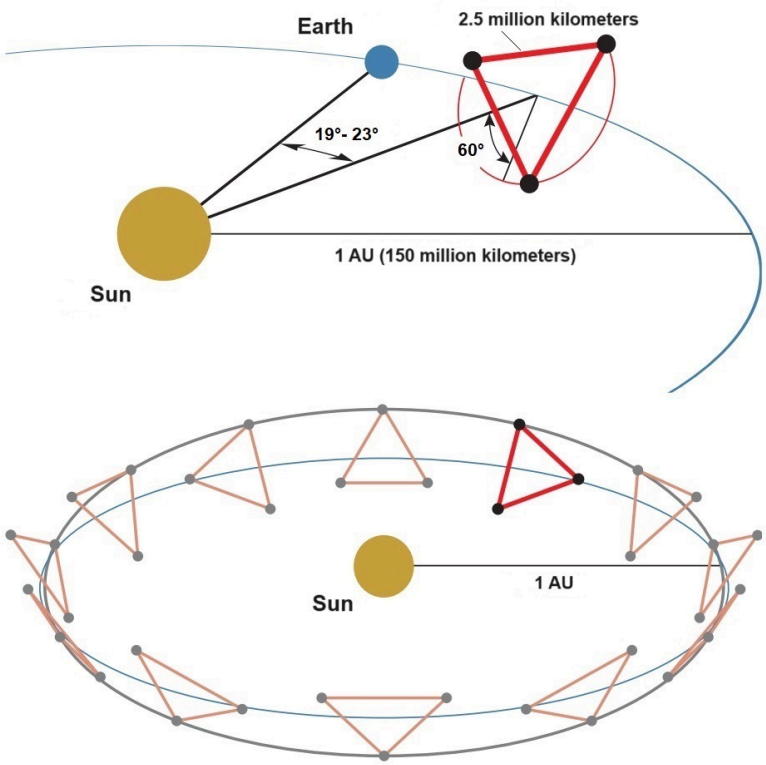

Figure 3.1: LISA constellation in heliocentric orbit. The spacecrafts are arranged in a triangular formation with an average distance of 2.5×10^6km between them. The constellation orbits around the Sun, trailing the Earth at an angle between 19° and 23°. The constellation plane is inclined by about 60° with respect to the ecliptic plane and performs a clockwise cartwheeling motion with the orbital frequency of one cycle per year. Figure adapted from [71].

We also define the reference polarization vectors in the SSB frame as

$$\hat{\boldsymbol{u}} = -\hat{\boldsymbol{e}}_\theta, \quad \hat{\boldsymbol{v}} = -\hat{\boldsymbol{e}}_\phi, \tag{3.5}$$

so that $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}, \hat{\boldsymbol{k}})$ form a right-handed orthonormal triad. The remaining degree of freedom between the frames is a rotation around the line of sight $\hat{\boldsymbol{k}}$, characterized by the polarization angle ψ. We define it as the rotation angle around $\hat{\boldsymbol{k}}$ corresponding to the map $\hat{\boldsymbol{u}} \to \hat{\boldsymbol{p}}$. The polarization basis tensors associated with $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}})$ are then related to those defined by $(\hat{\boldsymbol{p}}, \hat{\boldsymbol{q}})$ via a rotation by ψ, as in Eq. 1.33. The corresponding representation of the strain in the SSB frame is

$$h_{ab}^{\mathrm{SSB}}(t) = [h_+(t)\cos 2\psi + h_\times(t)\sin 2\psi]\epsilon_{ab}^{+} + [-h_+(t)\sin 2\psi + h_\times(t)\cos 2\psi]\epsilon_{ab}^{\times}. \tag{3.6}$$

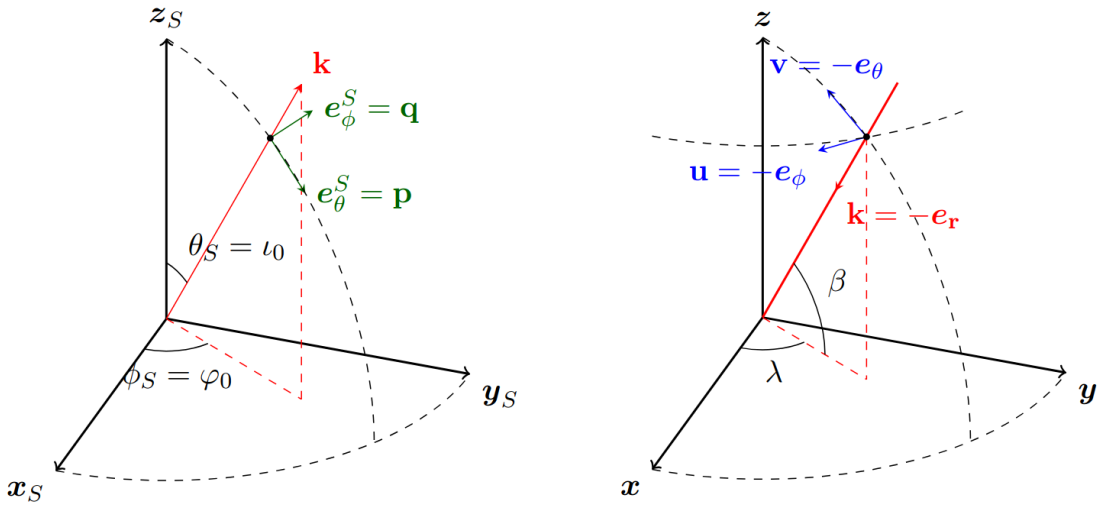

Figure 3.2: (*Left panel*) Source frame, defined by the unit vectors (x_S, y_S, z_S) and the spherical coordinates (θ_S, ϕ_S). (*Right panel*) the SSB frame, defined by the unit vectors (x, y, z) and the spherical coordinates (θ, ϕ). The GW propagation vector $\hat{\boldsymbol{k}}$ is shown in both frames. The polarization angle ψ is defined as the angle between the $\hat{\boldsymbol{u}}$ vector and the $\hat{\boldsymbol{p}}$ vector. Image adapted from [183]

3.2.2 Single-arm response to GWs

As shown in Section 1.2, GWs are weak perturbations of the spacetime metric. In the SSB frame, we describe the metric in the vicinity of the LISA constellation as

$$ds^2 = -dt^2 + (\delta_{ab} + h_{ab}(\mathrm{t}, \hat{\boldsymbol{k}}))dx^a dx^b, \tag{3.7}$$

where x^a denote the spatial coordinates introduced in Section 3.2.1. The dependence of Eq. 3.7 on the retarded time $\mathrm{t} = t - \hat{\boldsymbol{k}} \cdot \boldsymbol{x}$ is motivated by Eq. 1.49. Here, we adopt $c = 1$.

We now introduce λ as the affine parameter along the photon trajectory. Consider two spacetime events corresponding to photon emission and reception, with coordinates

$$x_0^\mu = x^\mu(\lambda = 0) = (t_0, x_0^a), \quad x_1^\mu = x^\mu(\lambda = 1) = (t_1, x_1^a), \tag{3.8}$$

respectively. The four-vector separation between these events is $r^\mu = x_1^\mu - x_0^\mu$. The phase evolution experienced by the photon is:

$$\Delta\phi = \int_{\text{path}} \omega_L dt = \int_0^1 \omega_L \dot{t} d\lambda. \tag{3.9}$$

Here, ω_L denotes the laser frequency, and the dot is the derivative with respect to the affine parameter λ. Assuming a constant laser frequency, Eq. 3.9 is equivalently expressed in terms of an optical path length as

$$L = \frac{\Delta\phi}{\omega_L} = \int_0^1 \dot{t} d\lambda \,. \tag{3.10}$$

The photon travels along a null geodesic, given by:

$$\dot{t}^2 = [\eta_{ab} + h_{ab}(\lambda)]\, \dot{x}^a \dot{x}^b, \tag{3.11}$$

where $\dot{x}^a = r^a + \delta r^a(\lambda)$. Inserting Eq. 3.11 in Eq. 3.9, we get:

$$L = \int_0^1 \left[x^a x_a + h_{ab}(\lambda) \dot{x}^a \dot{x}^b \right]^{1/2} d\lambda \tag{3.12}$$

$$= L_0 + \frac{L_0}{2} \hat{r}^a \hat{r}^b \int_0^1 h_{ab}(\lambda) d\lambda + \mathcal{O}(h^2) \tag{3.13}$$

$$= L_0 + \frac{1}{2} \frac{\hat{r}^a \hat{r}^b}{1 - \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}} \int_{\mathfrak{t}^0}^{\mathfrak{t}^1} h_{ab}(\mathfrak{t}') d\mathfrak{t}' + \mathcal{O}(h^2). \tag{3.14}$$

Thus, the armlength variation δL at linear order in h is

$$\Delta L(t) = \frac{1}{2} \frac{\hat{r}^a(t) \hat{r}^b(t)}{1 - \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}(t)} \int_{\mathfrak{t}^0}^{\mathfrak{t}^1} h_{ab}(\mathfrak{t}') d\mathfrak{t}', \tag{3.15}$$

where $\mathfrak{t}^1 = t - \hat{\boldsymbol{k}}_1 \cdot \boldsymbol{x}$ and $\mathfrak{t}^0 = (t - L_0) - \hat{\boldsymbol{k}}_0 \cdot \boldsymbol{x}$. Eq. 3.15 constitutes the single-link response to GWs in the time domain, where the arm direction $\hat{\boldsymbol{r}}$ is changing due to the spacecrafts motion.

We obtain a frequency-domain response by applying the FT to Eq. 3.15.

$$\int_{\mathfrak{t}^0}^{\mathfrak{t}^1} h_{ab}(\mathfrak{t}') d\mathfrak{t}' = \int_{\mathfrak{t}^0}^{\mathfrak{t}^1} d\mathfrak{t}' \int_{-\infty}^{\infty} df \tilde{h}_{ab}(f) e^{2\pi i f(\mathfrak{t}')} \tag{3.16}$$

$$= \int_{-\infty}^{\infty} df \frac{\tilde{h}_{ab}(f)}{2\pi i f} e^{2\pi i f(\mathfrak{t}^1 - \mathfrak{t}^0)} \tag{3.17}$$

Therefore, the response between sender j and receiver i spacecraft reads

$$\delta_{ij}(t, \hat{\boldsymbol{k}}) = \frac{\Delta L_{ij}(t)}{L_{ij}} = \frac{1}{2} \hat{r}_{ij}^a \hat{r}_{ij}^b \int_{-\infty}^{\infty} \tilde{h}_{ab}(f) \tau_{ij}(f, t, \hat{\boldsymbol{k}}) e^{2\pi i f(t - \Delta t)} df, \tag{3.18}$$

where

$$\Delta t = \frac{1}{2} \left[\hat{\boldsymbol{k}} \cdot (\boldsymbol{x}_i + \boldsymbol{x}_j) + L_{ij} \right], \tag{3.19}$$

$$\tau_{ij}(f, t, \hat{\boldsymbol{k}}) = \text{sinc}(\pi L_{ij} f [1 - \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}_{ij}(t)]). \tag{3.20}$$

3.2.3 Response to GWs superposition

We describe the stochastic strain as a linear superposition of plane tensor-waves incoming from each direction of the sky. Therefore we define an integrated response to each polarization $p = +, \times$, as follows:

$$\delta^p_{ij}(t) = \int_{\hat{\boldsymbol{k}}} \delta^p_{ij}(t, \hat{\boldsymbol{k}}) d\hat{\boldsymbol{k}}, \tag{3.21}$$

where

$$\delta^p_{ij}(t, \hat{\boldsymbol{k}}) = \frac{1}{2} \hat{r}^a_{ij} \hat{r}^b_{ij} \int_{-\infty}^{\infty} \tilde{h}_p(f, \hat{\boldsymbol{k}}) \epsilon^p_{ab}(\hat{\boldsymbol{k}}) \tau_{ij}(f, t, \hat{\boldsymbol{k}}) e^{2\pi i f(t-\Delta t)} df \tag{3.22}$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} \tilde{h}_p(f, \hat{\boldsymbol{k}}) F^p(\hat{\boldsymbol{k}}, \hat{\boldsymbol{r}}_{ij}) \tau_{ij}(f, t, \hat{\boldsymbol{k}}) e^{2\pi i f(t-\Delta t)} df. \tag{3.23}$$

In Eq. 3.21 we considered the antenna pattern functions $F^p(\hat{\boldsymbol{k}}, \hat{\boldsymbol{r}}_{ij})$, introduced in Eq. 2.2 and Eq. 2.3. We rewrite Eq. 3.22 equivalently as:

$$\delta^p_{ij}(t, \hat{\boldsymbol{k}}) = \int_{-\infty}^{\infty} \tilde{h}_p(f, \hat{\boldsymbol{k}}) \overline{G}^p_{ij}(f, t, \hat{\boldsymbol{k}}) e^{2\pi i f t} df, \tag{3.24}$$

where $\overline{G}^p_{ij}(f, t, \hat{\boldsymbol{k}})$ reads

$$\overline{G}^p_{ij}(f, t, \hat{\boldsymbol{k}}) = \frac{1}{2} F^p(\hat{\boldsymbol{k}}, \hat{\boldsymbol{r}}_{ij}) e^{-i\pi f(L_{ij} + \hat{\boldsymbol{k}} \cdot (\boldsymbol{x}_i + \boldsymbol{x}_j))} \tau_{ij}(f, t, \hat{\boldsymbol{k}}). \tag{3.25}$$

Notably, $\overline{G}^p_{ij}(f, t, \hat{\boldsymbol{k}})$ is a complex quantity unlike $\tau_{ij}(f, t, \hat{\boldsymbol{k}})$. Equivalently, we recast Eq. 3.25 as a response in fractional frequency domain

$$G^p_{ij}(f, t, \hat{\boldsymbol{k}}) = 2\pi i L_{ij} f \overline{G}^p_{ij}(f, t, \hat{\boldsymbol{k}}). \tag{3.26}$$

The FT of Eq. 3.24 is then

$$\begin{aligned} \delta^p_{ij}(f, \hat{\boldsymbol{k}}) &= \int_{-\infty}^{\infty} dt e^{-2\pi i f t} \delta^p_{ij}(t, \hat{\boldsymbol{k}}) \\ &= \int_{-\infty}^{\infty} dt e^{-2\pi i f t} \int_{-\infty}^{\infty} \tilde{h}_p(f', \hat{\boldsymbol{k}}) \overline{G}^p_{ij}(f', t, \hat{\boldsymbol{k}}) e^{2\pi i f' t} df' \\ &= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} \tilde{h}_p(f', \hat{\boldsymbol{k}}) \overline{G}^p_{ij}(f', t, \hat{\boldsymbol{k}}) e^{-2\pi i t (f - f')} df' \\ &= \int_{-\infty}^{\infty} df' \overline{G}^p_{ij}(f', f - f', \hat{\boldsymbol{k}}) \tilde{h}_p(f', \hat{\boldsymbol{k}}), \end{aligned} \tag{3.27}$$

where we have defined

$$\overline{G}^p_{ij}(f', f, \hat{\boldsymbol{k}}) = \int_{-\infty}^{\infty} e^{-2\pi i f t} \overline{G}^p_{ij}(f', t, \hat{\boldsymbol{k}}) dt. \tag{3.28}$$

For an isotropic, stationary, zero-mean background, fully characterized by its PSD, we implicitly define a response matrix for two given links, ij and $i'j'$, as

$$\Sigma_{ij,i'j'}(f) = \left\langle \delta^p_{ij}(f) \delta^{*p}_{i'j'}(f) \right\rangle = \int_{\hat{\boldsymbol{k}}} d\hat{\boldsymbol{k}} \int_{-\infty}^{\infty} df' \overline{G}^p_{ij}(f', f - f', \hat{\boldsymbol{k}}) \overline{G}^{*p}_{i'j'}(f', f - f', \hat{\boldsymbol{k}}) S_h(f'). \tag{3.29}$$

Assuming an isotropic distribution of sources, we integrate the product of the $\overline{G}$ functions in Eq. 3.29 and evaluate it at a reference time t_0—without loss of generality, due to the assumption of perfect stationarity—to obtain:

$$\Sigma^p_{ij,i'j'}(f) = G^p_{ij,i'j'}(f,t_0)S_h(f), \tag{3.30}$$

with

$$G_y(f,t_0) = G^p_{ij,i'j'}(f,t_0) = \int_{\hat{\boldsymbol{k}}} d\hat{\boldsymbol{k}}\overline{G}^p_{ij}(f,t_0,\hat{\boldsymbol{k}})\overline{G}^{*p}_{i'j'}(f,t_0,\hat{\boldsymbol{k}}). \tag{3.31}$$

Summing over the polarizations, we obtain the single link response matrix, which we denote with a 6 by 6 matrix $\mathcal{R}(f,t_0)$.

$$\Sigma_{ij,i'j'}(f) = \sum_p G^p_{ij,i'j'}(f,t_0)S_h(f) = \mathcal{R}_{ij,i'j'}(f,t_0)S_h(f). \tag{3.32}$$

3.3 Instrumental noise

The sensitivity of LISA is limited by several noise sources, some astrophysical in origin. Here, we focus on the instrumental noise, leaving astrophysical sources for a detailed discussion in Section 3.6. One main source is expected to dominate the raw LISA interferometric data:

Laser frequency noise : by far the dominant noise for LISA, it is due to laser frequency fluctuations. In an ideal case, a laser operates at a single frequency with zero linewidth. In the real world, however, a laser has a finite linewidth due to phase fluctuations, which cause instantaneous frequency shifts away from the central frequency, given by $\delta\nu(t) = \frac{1}{2\pi}\frac{d\phi}{dt}$. Unlike ground-based detectors, the frequency fluctuations do not cancel naturally thanks to common signal propagation across different arms, thus they appear in the LISA data stream. In order to suppress the laser frequency noise, a post-process technique called TDI is employed, which we describe in Section 3.4.

Assuming laser frequency noise is effectively suppressed, two secondary noise sources remain unmitigated by TDI:

Test Mass (TM) noise : test masses onboard each spacecraft are shielded in vacuum chambers from outer disturbances. However, as observed in LPF, free-fall can be achieved only partially. External spurious forces push the test masses away from their geodesics, resulting in a non-inertial residual accelerations. A number of different phenomena contribute to such residual forces, including: residual local gravity gradients result in inertial forces that can mimic tidal accelerations from GW. Additionally, cosmic charged particles can deposit charge on the test masses, which results in electrostatic forces. Also, the thermal noise from the surrounding molecules can cause fluctuations in the test mass position. Overall, we define a PSD model describing the test mass noise in a single link measurement ij as:

$$S_{ij,\mathrm{tm}}(f) = A^2_{ij,\mathrm{tm}}(2\pi cf)^{-2}\left[1+\left(\frac{0.4\mathrm{mHz}}{f}\right)^2\right]\left[1+\left(\frac{f}{8\mathrm{mHz}}\right)^4\right]\mathrm{Hz}^{-1}. \tag{3.33}$$

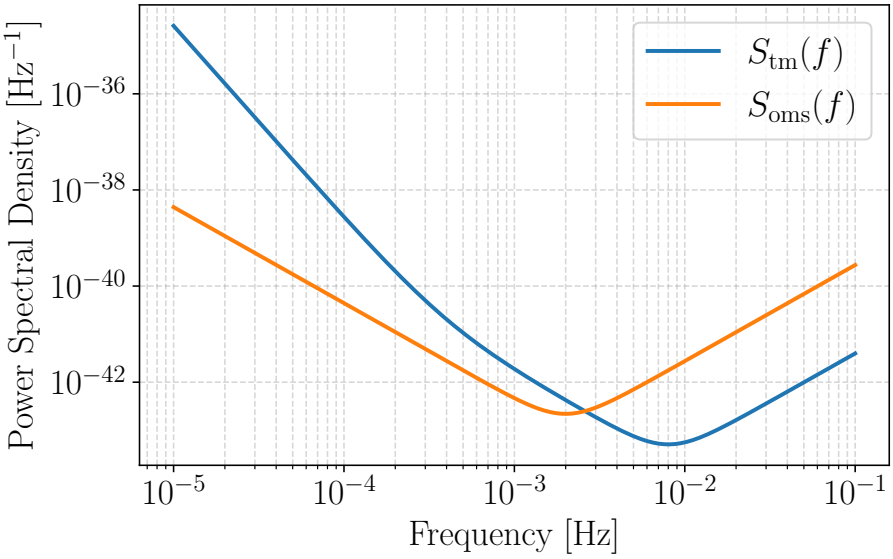

Figure 3.3: PSDs of the instrumental noises affecting LISA measurements. The blue (orange) line represents the TM (OMS) component, dominant at high (low) frequencies. From the superposition of the two noise sources, we observe that the overall noise is minimum around 3 mHz.

Optical Metrology System (OMS) noise the OMS noise enters the LISA datastream via the photodetectors, responsible for measuring the light power received from each distant spacecraft. The OMS noise is due to fluctuations in the laser power, which is caused by several factors. The shot noise is an example of such noise, originating from the quantum nature of light: as the laser beam is composed by photons emitted randomly in time, the photon count in a given time interval obeys to a Poisson distribution, which results in fluctuations in the measured power. The OMS PSD noise model, in a single link measurement ij, is given by:

$$S_{ij,\rm oms}(f) = A^2_{ij,\rm oms}\left[1+\left(\frac{2\rm mHz}{f}\right)^4\right]\left(\frac{2\pi f}{c}\right)^2 \rm Hz^{-1}. \tag{3.34}$$

In a single link measurement, the secondary noises enter as:

$$n_{ij}(t) = n_{ij,\rm oms}(t) + n_{ij,\rm tm}(t) + n_{ji,\rm tm}(t - L_{ij}), \tag{3.35}$$

where $n_{ji,\rm tm}(t - L_{ij})$ is the TM noise from the other spacecraft, which is delayed by the light travel time L_{ij}. Assuming that individual noise terms are stationary, zero-mean, and uncorrelated, and using the single-link noise definition in Eq. 3.35, we compute the non-zero entries of the covariance matrix for single links:

$$\Sigma^n_{ij,ij}(f) = S_{ij,\rm tm}(f) + S_{ij,\rm oms}(f) + S_{ji,\rm tm}(f) \tag{3.36}$$

$$\Sigma^n_{ij,ji}(f) = \exp\left[2\pi i f L_{ji}\right] S_{ij,\rm tm}(f) + \exp\left[-2\pi i f L_{ij}\right] S_{ji,\rm tm}(f) \tag{3.37}$$

It is important to notice that in principle amplitude for TM and OMS noise are different for different single arm measurements. In this thesis, we consider a central values $A_{\rm tm} = 2.4\times10^{-15}{\rm m\,s^2\,Hz^{-0.5}}$ and $A_{\rm oms} = 7.9\times10^{-12}{\rm m\,Hz^{-0.5}}$ are considered respectively. Overall, the OMS noise dominates at frequencies above 3 mHz, as shown in Fig. 3.3, while TM noise dominates at lower frequencies.

3.4 Time delay interferometry

In conventional equal arm Michelson interferometers, the laser frequency fluctuations affecting the two arms cancel out at the photodetector [191]. This is because the two arms have the same length and the laser frequency fluctuations undergo the same delay. However, in LISA, the arms are not equal in length. To quantify this effect, we consider a simplified interferometer with two arms of different lengths L_1 and L_2, and denote by $n^l(t)$ the laser frequency fluctuations. Since measurements at the photodetector occur after a round-trip time of $2L_i/c$ along each arm, the differential laser noise is given by

$$\Delta n^l(t) = n^l\left(t - \frac{2L_1}{c}\right) - n^l\left(t - \frac{2L_2}{c}\right). \tag{3.38}$$

Under the low-frequency approximation $f \ll c/L_i$, the FT of Eq. 3.38 is

$$\Delta n^l(f) \approx 4\pi f \frac{L_1 - L_2}{c} |n^l(f)|. \tag{3.39}$$

For LISA, the armlength difference is approximately 1%, i.e. $|L_1 - L_2|/L \sim 0.01$, corresponding to $(L_1 - L_2)/c \sim 0.083$ s. Assuming typical laser frequency noise $n^l(f = 1\ {\rm mHz}) \sim 10^{-13}/\sqrt{\rm Hz}$ [164], the resulting differential noise is $\Delta n^l \sim 1.04 \times 10^{-16}/\sqrt{\rm Hz}$ at $f = 1$ mHz. Since the LISA sensitivity requirement at this frequency is $\sim 10^{-20}/\sqrt{\rm Hz}$, laser frequency fluctuations would exceed the target by four orders of magnitude, completely overwhelming the GWs signal. To mitigate this effect, we employ the TDI technique [186], a post-processing method that synthesizes virtual equal-arm interferometers by combining suitably delayed linear combinations of measurements from the three spacecraft.

3.4.1 A gentle introduction to TDI

In this section, we present a simplified example of the TDI application. We consider a Michelson interferometer with two arms of different lengths L_1 and L_2 and laser frequency noise $n^l(t)$. We denote the signals and the overall photodetector measurements as $s_{1,2}(t)$ and $y_{1,2}(t)$, respectively. The latter are computed as:

$$y_1(t) = s_1(t) + n^l(t - 2L_1) - n^l(t), \tag{3.40}$$
$$y_2(t) = s_2(t) + n^l(t - 2L_2) - n^l(t). \tag{3.41}$$

The pure datastreams subtraction $(y_1(t) - y_2(t))$ does not remove the laser phase noise $n^l(t)$. The crucial point is to delay the photodetector signals by the round-trip time of the other arm (for example $y_1(t)$ by $2L_2$).

$$y_1(t - 2L_2) = s_1(t - 2L_2) + n^l(t - 2L_1 - 2L_2) - n^l(t - 2L_2), \tag{3.42}$$
$$y_2(t - 2L_1) = s_2(t - 2L_1) + n^l(t - 2L_2 - 2L_1) - n^l(t - 2L_1). \tag{3.43}$$

The term $n^l(t-2L_1-2L_2)$ appears in both equations and is therefore eliminated by taking the difference $y_1(t-2L_2)-y_2(t-2L_1)$. In this combination, the laser phase noise enters exactly in the same form of $(y_1(t)-y_2(t))$. Thus, we define the variable $x(t)$ as:

$$x(t) = [y_1(t-2L_2) - y_2(t-2L_1)] - [y_1(t) - y_2(t)]\,, \tag{3.44}$$
$$= y_1(t-2L_2) - y_2(t-2L_1) - y_1(t) + y_2(t). \tag{3.45}$$

Eq. 3.45 emphasizes the basic idea of TDI. The aim is to construct laser-noise-free datastreams by properly adding and time-shifting the interferometric photodetector measurements. The new variables are designed to retain the GW signal.

3.4.2 TDI combinations

We now show how to apply TDI to the LISA constellation. First, we introduce the notation defining the LISA constellation geometry and link labeling. The spacecrafts are labeled 1, 2, 3, and they are clockwise oriented. The links between the spacecrafts are labeled L_{ij}, where i refers to the local spacecraft and j refers to the distant spacecraft. Similarly, the LISA measurements associated to each OB are denoted as y_{ij}. The GW signal appears in y_{ij} as described in Eq. 3.18. Left panel of Fig. 3.4 shows the LISA setup and the notation used in this section. We now define the delay operator $\mathcal{D}_{ij}$ as follows:

$$\mathcal{D}_{ij}y(t) = y(t - L_{ij}). \tag{3.46}$$

For constant armlengths, the delay operator is self-commutative, and repeated applications yield:

$$\mathcal{D}_{i_1j_1}\ldots\mathcal{D}_{i_nj_n}y(t) = y\left(t - \sum_{k=1}^{n} L_{i_kj_k}\right). \tag{3.47}$$

We now define the first-generation TDI variables, which fully suppress the laser frequency noise in a constellation with constant arm lengths, either equal or unequal. The relative Michelson X variable is defined as:

$$X_1 = (1-\mathcal{D}_{13}\mathcal{D}_{31})(y_{12}+\mathcal{D}_{12}y_{21}) + (\mathcal{D}_{12}\mathcal{D}_{21}-1)(y_{13}+\mathcal{D}_{13}y_{31}). \tag{3.48}$$

A schematic representation of the X variable is shown in the right panel of Fig. 3.4.
The laser frequency noise in more realistic scenarios requires other synthetic loops in the photon path to be included in the TDI variables. Unequal and non-constant arm lengths require second-generation TDI variables, where the X variable is now defined as:

$$X_2 = (1-\mathcal{D}_{31}\mathcal{D}_{13}\mathcal{D}_{12}\mathcal{D}_{21})\mathrm{X}_1, \tag{3.49}$$

The Michelson variables represent one possible family of TDI combinations. Other families, such as the Sagnac variables [173], can be constructed from different combinations of the six time-delayed single-link measurements y_{ij}. From each family, an additional set of combinations is defined, denoted by A, E, T. For instance, given the Michelson variables XYZ, these read

$$A = \frac{Z-X}{\sqrt{2}}\,, \qquad E = \frac{X-2Y+Z}{\sqrt{6}}\,, \qquad T = \frac{X+Y+Z}{\sqrt{3}}\,. \tag{3.50}$$

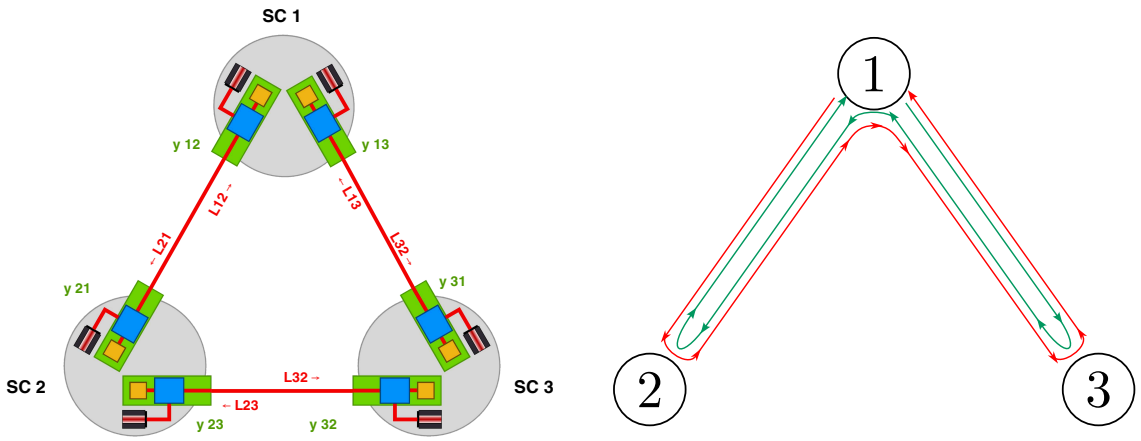

Figure 3.4: (*Left panel*) Schematic representation of the LISA constellation. The spacecrafts are labeled 1, 2, 3, and they are oriented clockwise. The links between the spacecrafts are labeled as L_{ij}, where j refers to the sender spacecraft and i refers to the receiver spacecraft. The LISA measurements associated to each OB are denoted as y_{ij}. Figure adapted from [32]. (*Right panel*) Scheme of the first generation Michelson X - TDI variable. It can be visualized as the difference between two sums of measurements (green and red), each corresponding to a specific light path from the laser onboard spacecraft 1.

The *AET* variables are *quasi*-orthogonal, meaning that the secondary noise in the three channels is uncorrelated in the equal-arms limit and for identical noise levels. This is significant because, in Eq. 2.54, the likelihood factorizes into three separate terms rather than forming a fully symmetric noise covariance matrix. Under the same assumptions, the T channel is also significantly less sensitive to the GW signal, allowing an instrumental noise measurement in a GW-quiet datastream. These features make the AET variables highly advantageous for LISA data analysis.

It is also useful to define the TDI variables in the frequency domain. We define the vector of FT single link measurements as $\tilde{\boldsymbol{y}}(f) = \{\tilde{y}_{ij}(f)\}_{i\neq j}$. We call $\tilde{\boldsymbol{d}}(f)$ the Fourier transformed TDI variables, e.g., $\tilde{\boldsymbol{d}}(f) = \{\tilde{X}(f), \tilde{Y}(f), \tilde{Z}(f)\}$. The delay operator turns into a multiplication by a phase factor in the frequency domain:

$$\mathcal{D}_{ij}\tilde{y}(f) = e^{-2\pi i f L_{ij}}\tilde{y}(f). \tag{3.51}$$

Applying multiple delay operators is then equivalent to summing the phases. Thus, the TDI FT variables are expressed as:

$$\tilde{\boldsymbol{d}}(f) = \mathcal{M}(f)\tilde{\boldsymbol{y}}(f), \tag{3.52}$$

where $\mathcal{M}(f)$ is a $[3 \times 6]$ matrix, which projects the single link measurements into the TDI variables. Through $\mathcal{M}(f)$, we can now propagate the matrix $\mathcal{R}(f, t_0)$, introduced in Section 3.2.3, to the TDI level.

$$R(f, t_0) = \mathcal{M}(f)\mathcal{R}(f, t_0)\mathcal{M}^{\dagger}(f). \tag{3.53}$$

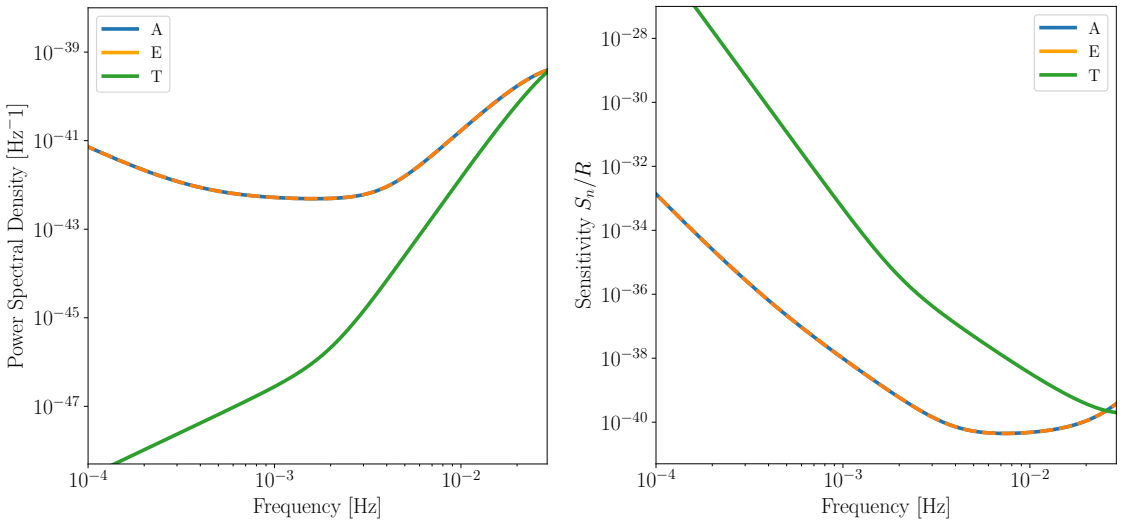

Figure 3.5: (*Left panel*) PSD of the noise affecting the LISA AET channels. (*Right panel*) Relative sensitivity curve. Both panels refer to the first-generation TDI observables. As discussed in the text, the T channel is primarily sensitive to instrumental noise rather than the gravitational wave signal, resulting in a significantly reduced sensitivity.

Then, the covariance matrix in TDI variables reads

$$C_{\rm GW}(f) = R(f, t_0) S_h(f). \tag{3.54}$$

Similarly, we now want to project the noise component described Eqs. 3.36 and 3.37 into the TDI variables.

$$C_n(f) = \mathcal{M}(f)\Sigma^n(f)\mathcal{M}^\dagger(f), \tag{3.55}$$

Given the single link noise definition Eq. 3.35, the TM and OMS components have different correlation properties, which will ultimately cause the two components to have different transfer functions in the various TDI channels. In terms of the XYZ variables, the noise covariance reads

$$S^{\rm n}_{aa}(f) = F_{\rm gen}(f)\, 8 \sin^2\left(\frac{2\pi f L}{c}\right) \left\{ \left[3 + \cos\left(\frac{4\pi f L}{c}\right)\right] S_{\rm tm}(f) + S_{\rm oms}(f) \right\}, \tag{3.56}$$

$$S^{\rm n}_{ab}(f) = -F_{\rm gen}(f)\, 2 \sin\left(\frac{2\pi f L}{c}\right) \sin\left(\frac{4\pi f L}{c}\right) \left[4 S_{\rm tm}(f) + S_{\rm oms}(f)\right], \tag{3.57}$$

where $a, b \in \{X, Y, Z\}$, $S_{\rm tm}$ and $S_{\rm oms}$ are the test-mass and optical metrology system noise spectra (see Eqs. 3.33 and 3.34), and $F_{\rm gen}(f)$ is a factor accounting for the TDI generation considered. In particular, we have

$$F_1(f) = 1, \tag{3.58}$$

$$F_2(f) = 4 \sin^2\left(\frac{4\pi f L}{c}\right), \tag{3.59}$$

where the difference from unity in F_2 arises from the extra delay factors in the second-generation TDI combinations (Eq. 3.49). Instead, in the A, E, T channels, the noise covariance entries are

$$\begin{aligned} S^{\mathrm{n}}_{AA}(f) &= S^{\mathrm{n}}_{EE}(f) \\ &= F_{\mathrm{gen}}(f)\, 8 \sin^2\left(\frac{2\pi f L}{c}\right) \Big\{ 4\big[1 + \cos\left(\frac{2\pi f L}{c}\right) + \cos^2\left(\frac{2\pi f L}{c}\right)\big] S_{\mathrm{tm}}(f) \\ &\quad + \big[2 + \cos\left(\frac{2\pi f L}{c}\right)\big] S_{\mathrm{oms}}(f) \Big\}, \end{aligned} \tag{3.60}$$

$$\begin{aligned} S^{\mathrm{n}}_{TT}(f) &= F_{\mathrm{gen}}(f)\, 8 \sin^2\left(\frac{2\pi f L}{c}\right) \Big\{ 4\big[1 - \cos\left(\frac{2\pi f L}{c}\right)\big]^2 S_{\mathrm{tm}}(f) \\ &\quad + 2\big[1 - \cos\left(\frac{2\pi f L}{c}\right)\big] S_{\mathrm{oms}}(f) \Big\}. \end{aligned} \tag{3.61}$$

In Fig. 3.5, we show the corresponding PSD and sensitivity curves for the AET channels. The effective sensitivity, defined in Eq. 2.40, is computed as the ratio between the noise PSD given in Eq. 3.60 and the sky- and polarization-averaged response from Eq. 3.53, evaluated along a single channel.

3.5 LISA sources

Here we give a brief overview of the main sources of GWs that LISA is expected to observe. The science goals of LISA are extremely broad. They include observing the formation, growth, and mergers of Massive Black-Hole Binaries (MBHBs). The mission will also survey stellar populations, study dynamics in galactic nuclei, and explore the structure of the Galaxy. Finally, it may detect primordial GWs from the early universe. The main expected GW sources in the LISA band are:

- **Massive Black Hole Binaries.** LISA is expected to detect the MBHB mergers with masses between 10^5 and $10^7 \mathrm{M}_\odot$ at redshifts up to $z \sim 15$. Different formation channels—such as stellar hardening through three-body interactions [172], gas-driven migration [77], or evolution via interactions with other massive black holes [45]—predict distinct properties and merger rates. LISA observations of coalescing MBHB will enable us to infer these mechanisms and constrain their evolutionary pathways [190]. These systems are also candidate for multimessenger astronomy [33]. Since these systems form and evolve in gas-rich environments, they are anticipated to emit Electromagnetic (EM) radiation, primarily in the X-ray and optical bands. Tracking the EM emission can be used to identify potential MBHBs candidate in the LISA band [128]. Furthermore, the merger of MBHBs is expected to be followed by strong prompt shocks in the circumbinary disk. These should produce a bright afterglow. Constraining sky location and merger time days before the final coalescence is crucial to trigger EM follow-up observations. In Fig. 3.6, MBHBs with different masses are represented by color coded lines, denoting the accumulated SNR up to the given frequency.

- **Galactic binaries**: binary systems consisting of two COs, mainly WDs, in the Milky Way (MW). LISA is expected to detect individually up to tens of thousands

of these systems, providing insightful information about the population of COs in the Galaxy. GW signals from these systems are almost monochromatic, therefore are represented as grey dots in Fig. 3.6. At the same time, the vast majority of them will remain unresolved. Then, the GWs signal emitted by millions of these systems overlap incoherently, producing a confusion noise, known as Galactic Foreground (GF). This will result in a dominant source of noise in the LISA band masking signals emitted by other systems. The GF effectively limits the sensitivity of LISA to other sources of GWs between 0.3mHz and 3mHz (black line in Fig. 3.6). Furthermore, several verification binaries with known sky position and their orbital period (so the GWs frequency) through EM observations, providing guaranteed multimessenger sources (red hexagons) [82]. Their name comes from the fact that they might be used to verify the LISA performance and sensitivity, knowing already their physical parameters [124].

- **Stellar Origin Black Hole Binaries**: Mergers of SOBBH are currently observed by ground-based detectors in the frequency band between 10 Hz and 1000 Hz. By detecting extragalactic SOBBHs at much lower frequencies, LISA will enable the measurement of properties that are difficult to access from the ground, such as the binary's orbital eccentricity, or the influence of a surrounding dense gaseous disk. Some of these systems will be observed by LISA during the final years of their inspiral and will later merge within the sensitivity band of ground-based detectors, offering a unique opportunity for multi-band gravitational wave astronomy [171]. However, the detection rate of these systems is expected to be low, up to a few tens per year [50]. The SOBBHs are represented by the blue lines in Fig. 3.6.

- **Extreme Mass Ratio Inspirals**: These objects are relativistic binaries composed by CO slowly inspiraling around a central MBH. They are characterised by an extremely tiny mass ratio $q = m_2/m_1$, with $m_1 > m_2$, ranging between $10^{-4} - 10^{-7}$. The formation channels remain uncertain; however, the systems they produce are expected to exhibit significant eccentricity. Due to their very small mass ratio, EMRIs evolve slowly, completing $\sim 10^4 - 10^6$ cycles in the LISA frequency range before eventually plunging onto the central MBH [95, 142]. A large number of cycles allow measuring binary parameters with exquisitely high precision. Therefore, EMRI are ideal sources to map MBH spacetime [163, 35], perform tests of general relativity [88], and possibly detect the presence of gas around the central MBH [40, 39]. Measuring the properties of a population of EMRI signals could additionally provide information on the mass distribution of MBHs [86] and their host stellar environment [15]. More recently, EMRIs have been proposed as potential targets for strong gravitational lensing [187]. Figure 3.6 includes the tracks of five harmonics of an EMRI source, whose signals are emitted simultaneously.

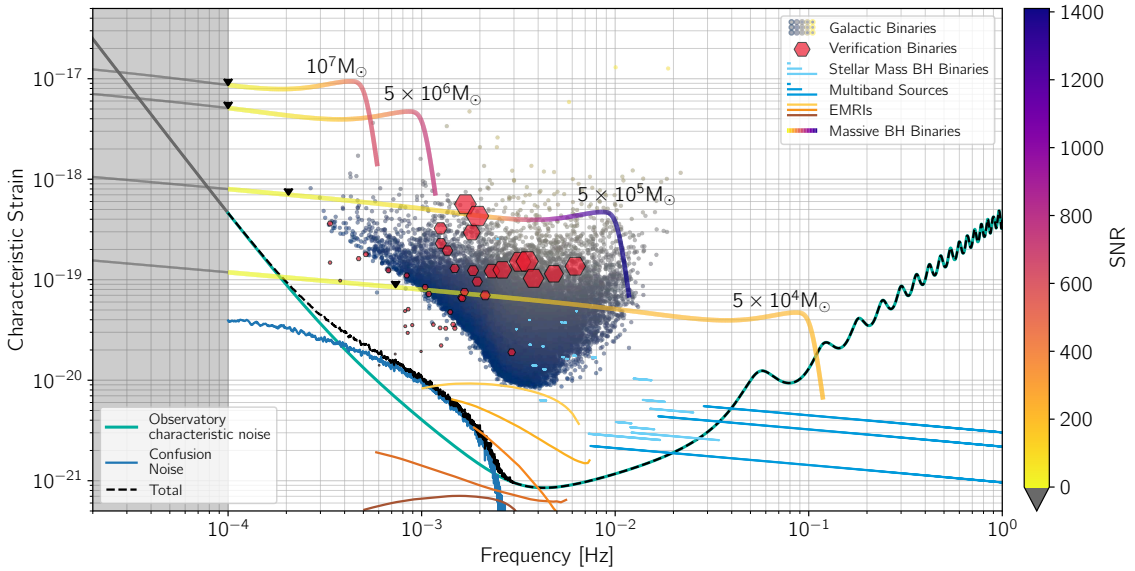

Figure 3.6: Expected Astrophysical Gravitational Wave Sources in the LISA Frequency Band. The figure presents four MBHB sources at varying masses, with color coding based on SNR accumulation. Below, five curves correspond to the EMRI harmonic. Galactic WDs binaries are represented by the dots, while the straight blue curves indicate the SOBBH sources, including multiband signals. Solid teal, solid blue, and dashed black curves represent the instrumental noise, unresolved gravitational wave foreground, and their combined contribution, respectively. The upward displacement of the resolvable sources corresponds to an SNR detection threshold of 7. The grey shaded region indicates the extrapolated LISA instrumental noise below 0.1 mHz. Figure adapted from [67].

3.6 SGWB in LISA

In addition to the detection of individual sources, LISA is also expected to observe SGWB. This can be either cosmological or astrophysical in origin, and we summarize below the main sources of SGWBs in LISA. Two brief, diagrammatic and spectral overviews of SGWB components are also provided in Figs. (3.7) and (3.8), respectively. Cosmological sources are related to early universe processes, some intrinsically stochastic. The main contributions are

- **Phase transitions**: These are processes in which the universe undergoes a change of state, such as the transition from radiation domination to matter domination. Typically associated to spontaneous breaking of fundamental theories symmetry groups, these processes can produce SGWBs through the emission of GWs from bubble collisions and sound waves in the plasma.
- **Cosmic strings**: These are topological defects that can form during phase transi-

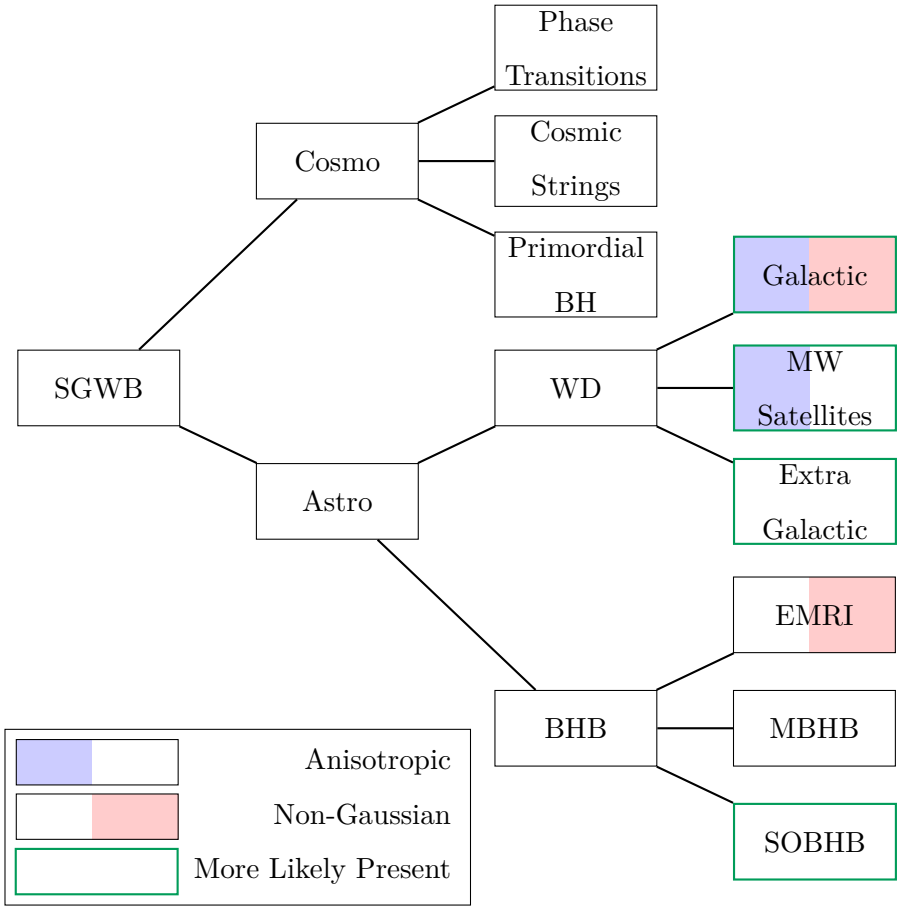

Figure 3.7: Classification of SGWB sources relevant to LISA. Sources are grouped according to their cosmological or astrophysical origin. Green borders indicate signals that are more likely to be present in the LISA data. Blue shading highlights anisotropic spatial distributions, while red shading denotes departures from Gaussian statistics

tions in the early universe. They usually evolve into a network of interacting strings and loops. The field theory dynamics, driving their evolution over cosmological timescales, leads to the production of aSGWBs through the GW emission due to strings oscillations and recombination.

- **Primordial black holes**: These are black holes that formed in the early universe, before the formation of stars and galaxies. They typically arise from the gravitational accretion and collapse of early Universe density perturbation according to its (unknown) initial quantum state. A population of such primordial BHs, if confirmed, may provide a SGWBs through mergers and matter accretion.

The interested reader may see [26] for a detailed review on cosmological SGWBs. Being cosmological, SGWBs generated by these processes are expected to be isotropic and stationary following the assumption of the *cosmological principle*. However, the expected amplitudes are still uncertain, and the prospects for detectability of such signals are still

debated.

In this thesis, we focus on astrophysical sources of SGWBs. They are generated by the emission of GWs from a large number of unresolved astrophysical sources, either Galactic or extragalactic. For this reason, the anisotropic assumption is not always valid, as well as the non-Gaussianity one if the number of overlapping sources is not large enough to leverage the central limit theorem. Moreover, our knowledge of the source population remains uncertain, and the expected amplitudes are still debated. The main astrophysical sources of SGWBs in LISA are:

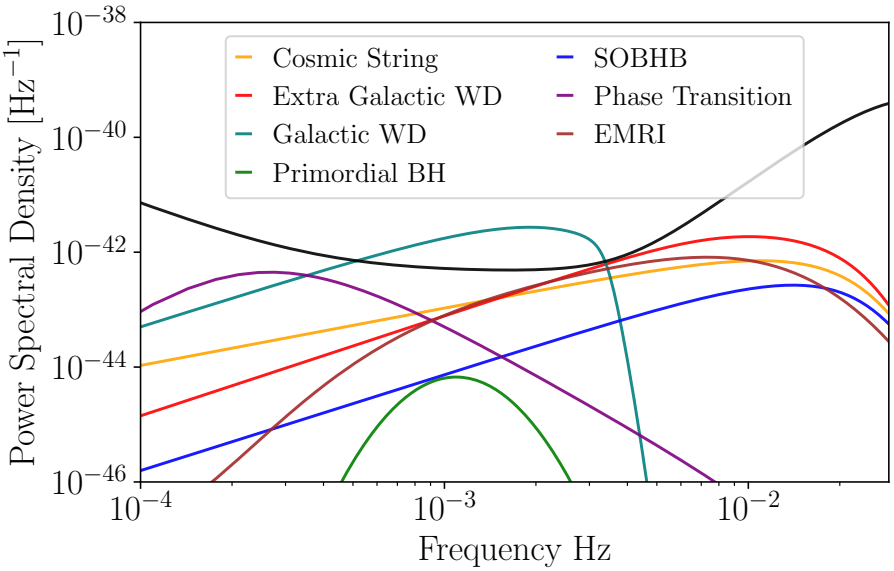

Figure 3.8: PSD of SGWB in LISA. Shown are signals in the first generation TDI of A channel. The instrumental noise is displayed in black. Astrophysical sources include the Galactic Foreground, Extra-Galactic Foreground, Stellar-Origin Black Hole Binaries , and Extreme Mass Ratio Inspirals. Cosmological contributions comprise Phase Transitions, Cosmic Strings, and Primordial Black Holes, modeled as a Gaussian spectral feature. Distinct spectral features enable the separation of overlapping components.

- **White Dwarfs Binaries**:

  Galactic Foreground. As anticipated in Section 3.5, the vast majority of WDs in the MW are individually unresolved by LISA, leading to a stochastic confusion noise [139]. This foreground dominates over the instrumental noise in the frequency range between roughly 0.5 and 3 mHz. Due to LISA periodic orbital motion, the detector response to this anisotropy induces a characteristic modulation in the observed signal. This modulation can be exploited to infer the spatial morphology of the Galaxy [47], encompassing components such as the disk, bulge, halo, and tidal streams. In addition, the GF spectral shape depends on the astrophysical properties of the underlying WDs binary population and on LISA sensitivity. Theoretical models predict two distinct spectral features: a low-frequency tail (below ~ 1 mHz)

and a high-frequency turnover or "knee" around a few millihertz. Assuming that the binary evolution is driven solely by GW emission, the low-frequency part of the spectrum follows a power-law behavior with a characteristic index of 2/3 in energy density, as discussed in Section 1.4.2. The high-frequency knee, however, is shaped by the transition from unresolved to resolvable sources and thus depends on several observational and astrophysical factors, including the total mission duration, the SNR threshold for resolvability, and the binary population properties such as metallicity. A phenomenological model has been proposed in [101], where the overall PSD is defined by:

$$S_h(f) = \frac{A}{2} f^{-7/3} \exp\left[-\left(\frac{f}{f_1}\right)^{\alpha}\right] \left[1 + \tanh\left(\frac{f_{\rm kn} - f}{f_2}\right)\right], \tag{3.62}$$

where A is the amplitude normalization, and the spectral breaks f_1 and $f_{\rm kn}$ depend on the observation time $T_{\rm obs}$ through the parametrization:

$$\log_{10} f_1 = a_1 \log_{10}(T_{\rm obs}) + b_1, \tag{3.63}$$

$$\log_{10} f_{\rm kn} = a_{\rm kn} \log_{10}(T_{\rm obs}) + b_{\rm kn}, \tag{3.64}$$

with a_i and b_i being calibrated model parameters. It is important to emphasize that this model is a simplified picture intended to reproduce the general structure of the foreground. In practice, the detailed spectral shape remains uncertain, not only due to limited knowledge of the WDs binary population, but also because of non-GW physical effects such as tidal interactions [188].

Furthermore, it has been recently shown that the GF deviates from Gaussian behavior, particularly near the knee frequency $f_{\rm kn}$ [51]. This non-Gaussianity increases the signal complexity and carries significant implications for data analysis. Altogether, these characteristics render the GF analysis both exceptionally fascinating and technically demanding. In Sections 5.2 and 5.3, we focus on analyzing the GF, with particular emphasis on exploiting its intrinsic anisotropy.

Milky Way Satellites backgrounds. Recent simulations indicate that besides the MW, other nearby dwarf galaxies such as the Large Magellanic Cloud (LMC), Small Magellanic Cloud (SMC), Sagittarius, Sculptor, and Fornax also host WDs binaries that are potentially resolvable as individual sources by LISA [158, 105]. The capacity to resolve WDs binaries depends on factors such as the mass, distance, and star formation history of the dwarf galaxy, as noted in [111]. Nonetheless, similarly to the MW, most WD binaries within dwarf galaxies are expected to remain unresolved. Therefore, it is likely that the undetectable WD binaries might collectively contribute to a SGWB detectable by LISA. In Section 5.2, we explore the potential contribution of these systems to the SGWB.

Extragalactic Foreground. Extragalactic WDs binaries are also expected to contribute to the astrophysical SGWB. However, significant uncertainties remain regarding the spectral power, as this population is currently inaccessible to direct observation by existing instruments. Specifically, a recent study by [97] examined the influence of different astrophysical assumptions—such as stellar metallicity and binary evolution prescriptions—demonstrating that the SGWB amplitude may vary by up to a factor of five. Despite this uncertainty, the extragalactic WDs background consistently emerges as the dominant astrophysical source expected to be detected by future

missions such as LISA—outshining contributions from binary black holes and neutron star binaries. This makes the extragalactic WD binaries background a virtually guaranteed LISA source [41, 180], with important implications for signal confusion and sensitivity, especially regarding the potential detection of cosmological SGWBs. The spectrum is roughly described by a broken power-law with an exponential cut-off, as follows,

$$S^{\rm EF}(f) = A_{\rm EF}\left(\frac{f}{f_{\rm break}}\right)^{\gamma_1}\left[1+\left(\frac{f}{f_{\rm break}}\right)^{4.15}\right]^{\gamma_2}\exp\left[-\left(\frac{f}{f_{\rm cut}}\right)^{3}\right], \tag{3.65}$$

with $f_{\rm break} = 7.2$mHz and $f_{\rm cut} = 40.2$mHz. This model is motivated by the physical properties of the WD binaries population. At low frequencies, the background exhibits a power-law behavior, characteristic of a superposition of binary inspirals. At higher frequencies, however, the finite size —much larger than black-hole binaries of WDs suppress efficient emission, leading to a depletion of sources and a corresponding spectral turnover. This transition introduces a smooth break in the spectrum, reflecting the distribution of physical parameters—such as masses and orbital separations.

- **Black Hole Binaries**:

  SOBBH. Given the recent constraints from the LIGO-Virgo-KAGRA (LVK) collaboration following the first three observation runs, we anticipate a large population of such systems contributing to the LISA data stream [1]. Most of these sources will form a SGWB: (i) they are either too faint or too distant to be individually resolved, and (ii) they generate long-lived, overlapping signals in the time domain. As detailed in Section 1.4.2, the resulting spectral shape is expected to follow a power-law, with a spectral index of approximately $\alpha \sim 2/3$ in terms of $\Omega_{\rm GW}(f)$. However, in Section 4.2, we explore how environmental effects may alter this background, potentially introducing significant deviations from the canonical spectral shape associated to emission in vacuum.

  EMRI. As previously mentioned in Section 1.3.2, EMRI formation mechanisms remain subject to debate, leading to significant uncertainty in the predicted event rates. Due to the individual waveform complexity and their extragalactic origin, most EMRIs are expected to remain unresolved throughout the mission, possibly generating a SGWB. Additionally, their typically high eccentricities make the prediction of the spectrum shape particularly challenging. In Section 4.1, we present a detailed analysis of the EMRI contribution to the SGWB. Furthermore, in Section 5.1, we demonstrate that such SGWBs may exhibit non-Gaussian features.

  MBHB. At nano-Hz frequencies, the first evidence for an unresolved GW signal of astrophysical and/or cosmological origin [19, 9] has recently been reported by the European pulsar timing array [79], NANOGrav [12], the Parkes Pulsar timing array [156], and the Chinese pulsar timing array [199]. While the detection prospects for individually resolvable MBHBs in LISA have been extensively characterized, the stochastic background arising from the unresolved population has remained largely unquantified. Recent analysis [56] demonstrates that, depending on the underlying MBHB formation scenario, this component may constitute a significant astrophysical foreground.

Chapter 4

Parameterizing and detecting the astrophysical SGWB

This chapter presents results from recent publications [148, 63, 151, 153]. We begin with detailed characterizations of two astrophysical backgrounds accessible to LISA: Section 4.1 analyzes the SGWB generated by EMRIs, while Section 4.2 focuses on the SGWB produced by SOBBHs. These complementary analyses illustrate how uncertainties in amplitude and spectral shape arise from fundamentally different physical origins.

For this reason, in Section 4.3 we introduce a flexible parameterization framework based on EGPs, implemented within the `balrog` pipeline. In Section 4.4, we also address the fundamental question of detectability in the presence of modeling uncertainties, proposing a novel Bayesian framework that enables rigorous assessment of detection prospects while naturally marginalizing over uncertainties in both signal and noise properties.

4.1 Characterization of EMRI background

The EMRIs population represents one of the most distinctive GW sources in the LISA band. As discussed in Section 3.6, the vast majority of EMRIs are not expected to be individually detectable with LISA, either because they are too distant or too far from the final plunge onto MBH. These unresolved sources collectively generate a SGWB whose properties reflect the underlying astrophysics of EMRI formation and evolution. Understanding this background is critical for two reasons: first, in the worst-case scenario, the SGWB could even exceed the instrumental noise level, potentially hindering the detection of other sources; second, its statistical properties encode valuable information about EMRI populations that cannot be accessed through individual detections alone. However, significant uncertainties plague our understanding of EMRI formation channels—particularly regarding the MBH mass function, spin distribution, cusp erosion following MBHB mergers, and the ratio of plunges to EMRIs. In this section, we characterize the expected EMRI background by generating realistic populations based on the models presented in Ref. [31], which are constructed to bracket the expected range of EMRI event rates. We use the AAK waveforms introduced in Section 1.3.3 and propagate the signals through the

full LISA response detailed in Sections 3.2.2 and 3.4.2 to evaluate the final level of the resulting stochastic background. These results will inform the flexible parametric analysis in Section 4.3.

4.1.1 Population model

Introducing the mass distribution of MBHs in the Universe $d^3N/(dM\,dz\,da)$, we create an effective MBH mass distribution for potential EMRI hosts

$$\left(\frac{d^3N}{dM\,dz\,da}\right)_{\text{eff}} = \frac{d^3N}{dM\,dz\,da}\,p_0(M,z)\,R(M,a). \tag{4.1}$$

The observed MBHs in present-day galaxies are thought to originate from initial BH seeds existing at high redshifts. This assumption is essential, as there is currently no known physical mechanism capable of directly forming a billion-solar-mass compact object. A comprehensive overview of this topic is provided by Ref. [169].
After the formation of a seed BH, a significant mass growth is required to evolve them into MBH. Three main mechanisms contribute to such mass growth. Firstly, mass can increase through the dynamical capture and subsequent accretion of stars or compact objects. Concurrently, gas accretion can transpire from gas clouds or disk-like structures. Lastly, the growth can result from mergers with others MBHs.
Taking into account the variety of available models, two distinct MBH mass functions have been employed by [31], capturing the uncertainties in the lower-end of the mass distribution. The first, labeled Model PopIII, assumes light MBH seeds originating from PopIII stars while accounting for delays between MBH and galaxy mergers [108]. In the relevant LISA mass range, the density mass function is approximated as follows:

$$\frac{dn}{d(\log(M_\bullet))} = 0.005\left(\frac{M_\bullet}{3\cdot 10^6 M_\odot}\right)^{-0.3}\ \text{Mpc}^{-3}. \tag{4.2}$$

The other model is a purely phenomenological one, aligned with the current observational constraints on the MBH mass function [87], and it is given by

$$\frac{dn}{d(\log(M_\bullet))} = 0.002\left(\frac{M}{3\cdot 10^6 M_\odot}\right)^{0.3}\ \text{Mpc}^{-3}. \tag{4.3}$$

In both cases, the redshift dependence in the MBH mass and the relevant redshift dependence for EMRIs is negligible. However, in Eq. 4.3 the function exhibits an increase with mass at the lower end, making it less suitable for EMRI events that fall within the LISA frequency band. Additionally, another key element characterizing the population is the MBHs spin distribution: the PopIII model traces the spin evolution of MBHs through accretion and mergers. The majority of MBHs relevant to LISA are expected to have maximal spin. This is due to the requirement that seeds must accrete sufficient mass to enter the LISA frequency band. Simulations [170] have demonstrated that for $M_\bullet < 10^7\,M_\odot$, the MBHs consistently align with the surrounding matter. Consequently, the angular momentum released during accretion accelerates their spin. Astrophysical models have considered a distribution with a median value of $a = 0.98$. However, for comparative and exploratory purposes, two additional scenarios have been explored: one with a flat spin distribution $0 < a < 1$, and another featuring non-spinning MBHs $a = 0$.

Model	Mass Function	MBH spin	Cusp Erosion	$M-\sigma$	N_p	M_{CO} $[M_\odot]$
M1	Barausse12	a98	yes	Gultekin09	10	10
M2	Barausse12	a98	yes	KormendyHo13	10	10
M3	Barausse12	a98	yes	GrahamScott13	10	10
M4	Barausse12	a98	yes	Gultekin09	10	30
M5	Gair10	a98	no	Gultekin09	10	10
M6	Barausse12	a98	no	Gultekin09	10	10
M7	Barausse12	a98	yes	Gultekin09	0	10
M8	Barausse12	a98	yes	Gultekin09	100	10
M9	Barausse12	aflat	yes	Gultekin09	10	10
M10	Barausse12	a0	yes	Gultekin09	10	10
M11	Gair10	a0	no	Gultekin09	10	10
M12	Barausee12	a98	yes	Gultekin09	0	10

Table 4.1: EMRI models considered in this study, developed in Ref. [31]. Column 1 denotes the label of each model. For each one, we specify the MBH mass distribution (column 2), the MBH spin model (column 3), whether we consider the effect of cusp erosion following MBHB mergers (column 4), the $M-\sigma$ relation (column 5), the ratio of plunges to EMRIs (column 6), and the mass of the COs (column 7). Finally, the total number of EMRIs occurring in a year up to z = 4.5 is shown in column 8.

To make the formation of EMRIs an efficient process, a core prerequisite is the presence of a dense distribution of stellar objects surrounding the MBH, usually referred to as a *cusp*. However, galactic mergers, including those involving MBHs, tend to carve out low-density cores, preventing the formation or disrupting cusps, thus making EMRI formation more inefficient. The factor $p_0(M,z)$ in Eq. 4.1 accounts for this effect. In particular, it is defined as the probability from a Poisson distribution that an individual MBH with a mass of M, observed at redshift z, undergo zero merger during its cusp regrowth time $t_{\rm cusp}$. To model this distribution, the mean number of mergers has to be estimated, integrating along $t_{\rm cusp}$ the semianalytical MBH evolution models as detailed in [20]. At first approximation $t_{cusp} \propto t_{rlx}$. The *two-body relaxation* time t_{rlx} captures the typical duration required for a star to exchange an amount of energy comparable to the average energy, or equivalently for its velocity to change by 90°. Through a detailed analysis [179], the expression for the relaxation time reads as follows

$$t_{rlx} \approx 1.8 \times 10^8 \,\mathrm{yr} \left(\frac{\sigma}{100\mathrm{km/s}}\right)^3 \left(\frac{10 M_\odot}{m_{CO}}\right) \left(\frac{10^6 M_\odot/\mathrm{pc}^3}{\overline{m} n}\right). \tag{4.4}$$

Here, n denotes the local density of stars, $\overline{m}$ the star mass-distribution average, and m_{CO} the mass of the CO. so the adopted $M_\bullet - \sigma$ relation significantly affects t_{cusp}. Different prescriptions are explored for a Milky-Way-like MBH, including: (i) Ref. [94] (referred to

as "Gultekin09") which results in a cusp regrowth time of approximately 6 billion years; (ii) Ref. [110] (referred to as "KormendyHo13") predicting a t_{cusp} of around 10 billion years; (iii) Ref. [92] (referred to as "GrahamScott13") suggesting a t_{cusp} of about 2 billion years through the development of an optimistic model.

Finally, R in Eq. 4.1 specifies the rate at which COs are captured by the central MBH. We refer to the intrinsic rate in Ref. [16], which accounts for mass segregation

$$R_0 = 300\,\mathrm{Gyr}^{-1}\left(\frac{M}{10^6 M_\odot}\right)^{-0.19}. \tag{4.5}$$

COs can also directly plunge into the MBH without emitting significant GWs, affecting the plunge-to-EMRI ratio. Thus, the effective rate R accounts also for the number of direct plunges N_p. With these models, three different regimes are explored, with N_p values of 0, 10, and 100. varying N_p mainly adjusts the normalization of the EMRI rate, in particular the more COs are allowed to directly plunge, the lower the EMRI rate becomes.

In summary, the rate of EMRIs depends on several factors, which we summarize below:

1. The MBH mass function, with considerations of the Barausse12 and Gair10 models.
2. The distribution of MBH spins, is explored through the configuration of near-maximally spinning distribution (a98), flat spin distribution (aflat), and nonspinning MBHs (a0).
3. The $M_\bullet - \sigma$ relation, defining the characteristics of the stellar distribution surrounding the MBH, the regrowth time following MBHB erosion, and the EMRI duty cycle. Three $M_\bullet - \sigma$ relations are studied: Gultekin09, KormendyHo13, and GrahamScott13. Additionally, an alternative model assuming the Gultekin09 relation without cusp erosion is considered.
4. The ratio of plunges to EMRIs, with assumptions of $N_p = 0, 10,$ and 100.
5. The characteristic mass of COs, accounting for both $m = 10M_\odot$ and $m = 30M_\odot$.

The impacts of each assumption are examined by varying them individually, leading to a total of 12 models denoted as "Mx" with x = 1,. . . , 12. The default model is labeled as M1, and the essential properties of each model constructed by [31] are outlined in Table 4.1. For each model, a Monte Carlo approach is employed to construct the population of EMRIs by sampling from the distribution Eq. 4.1, yielding a comprehensive catalog of EMRI events, including their masses (M, μ), redshift z, and MBH spin a. In particular, 10 Monte-Carlo realizations of the expected population of EMRIs plunging in 1 year were generated. Therefore, we effectively build a library including all EMRI events occurring in the Universe, during a 10-year observation time, for each of the 12 models.

4.1.2 Catalogs production

In Section 1.3.3 we showed that 14 parameters are required to characterize the EMRI waveform: $M, \mu, a, p, e, \cos(\iota), D_L, \theta_S, \phi_S, \theta_K, \phi_K, \Phi_{r,0}, \Phi_{\theta,0}, \Phi_{\phi,0}$. For each EMRI catalog, we assign a cosine of the inclination angle $\cos\iota$ by randomly drawing values from a uniform distribution between $[-1, 1]$, under the assumption that EMRIs form within a spherical bulge rather than a disk-like structure. This is a convenient parameterization, in that ι

spans the full interval $[0, \pi]$, with prograde (retrograde) orbits corresponding to $0 \leq \iota \leq \pi/2$ ($\pi/2 \leq \iota \leq \pi$). To determine the initial eccentricity and semi-latus, we follow [31]: therein, it has been shown that the eccentricity distribution at plunge is nearly flat in the range $0 < e_{\rm pl} < 0.2$. Overall, to draw each system parameters, we follow closely the strategy outlined in [43]:

- We use our knowledge on M and MBH spin a to estimate the radius at the last stable orbit. For each event, we draw the eccentricity at the last stable orbit e_p from a flat distribution in the range $[0; 0.2]$.
- We integrate the event's orbital elements backward in time for $T_{\rm back}$ years, employing Eq. 1.69, Eq. 1.70.
- We randomly sample $N_{\rm back} = {\rm int}(T_{\rm back}/10)$ points within the interval $[0; T_{\rm back}]$ to represent different evolutionary points of the EMRI. The division by 10 is performed because we collect 10 catalogs of EMRIs for every model, each representing one year of observation, as mentioned in Section 4.1.1. $T_{\rm back}$ is computed as $T_{\rm back} = 20(\frac{M_\bullet}{10^4 {\rm M}_\odot})$yr, the time to cover the same range in gravitational radii scales linearly with the mass of the system. Therefore, EMRIs featuring low-mass primaries predominantly emit GWs within the LISA band during the final years of their inspiral, whereas those involving larger MBH masses emit within the LISA band for longer time.
- For each of the $N_{\rm back}$ points, we store the semi-major axis a and eccentricity, and compute the semi-latus rectum using the definition $p_0 = a(1 - e_0^2)$.

This approach creates $N_{\rm back}$ copies of each EMRI from the catalogs in Table 4.1, sharing the same redshift, masses, and spin. Despite this limitation, the 10-year EMRI catalogs we are using contain thousands of events covering the relevant range of MBH mass and redshift. Therefore, we expect that this procedure will not introduce any significant bias in the foreground computation. We assume isotropic distributions for sky positions (θ_S, ϕ_S) and spin orientations (θ_K, ϕ_K). Similarly, we assume uniform distributions in $[0, 2\pi]$ for the three initial phases $\Phi_{r,0}$, $\Phi_{\theta,0}$, and $\Phi_{\phi,0}$.

The full EMRI populations we generate contain between several tens of thousands and several million sources. Evaluating the SGWB contribution from the entire population would be computationally prohibitive, particularly given the non-negligible cost of generating EMRI waveforms. To reduce the computational cost, we select only the sources that significantly contribute to the SGWB. For this purpose, we compute a preliminary SNR for each EMRI using an inclination- and polarization-averaged version of the analytic kludge waveform, following the same approach as in [43]. The EMRI evolution is truncated at the Schwarzschild last stable orbit, assuming an observation time of 4 years. We then discard all sources with $\rho < 1$, a threshold that removes the vast majority of sources while retaining the dominant contribution to the background. After this cut, we observe a ${\rm SNR}_{\rm SGWB}$ loss of only $\sim 5\%$, as reported in Table 4.2.

In addition, we apply two further selection criteria: (i) a cut on the initial eccentricity and (ii) a cut on the initial semi-latus rectum. The number of harmonics required to construct the waveform increases steeply with eccentricity, significantly raising the computational cost. Therefore, we remove all sources with $e_0 > 0.9$ from the catalogs. We also discard sources with an initial semi-latus rectum $p_0 < 10$, as the accuracy of the AAK waveform is no longer guaranteed at small separations. This latter cut eliminates only about $\sim 1\%$ of

Model	$\%\mathrm{SNR}_{\mathrm{tot}}$	$\%\mathrm{SNR}_{\rho}$	$\%\mathrm{SNR}_{\mathrm{p}}$	$\%\mathrm{SNR}_{\mathrm{e}}$
M1	20	7	11	2
M2	18	4	11	3
M3	28	5	19	4
M4	22	4	13	6
M5	26	4	18	4
M6	43	8	30	5
M7	20	4	11	5
M8	16	3	12	1
M9	16	4	9	3
M10	14	4	8	2
M11	49	9	39	1
M12	29	4	18	7

Table 4.2: Percentage of SNR after different sources selections. Column 2 is the total percentage lost, column 3 are referred after removing the dimmest source with SNR < 1, while column 4 and 5 are respectively the SNR after deleting sources with $p < 10$ and $e > 0.9$.

the sources, but it results in a non-negligible reduction of the total background $\mathrm{SNR}_{\mathrm{SGWB}}$. This is due to the fact, that EMRIs close to the merger (i.e. small semi-latus) are overall more efficient in emitting GWs. As shown in Table 4.2, the final $\mathrm{SNR}_{\mathrm{SGWB}}$ after all selections is approximately 20–30% lower compared to that obtained from the full population. To summarize, the final EMRI catalogs include only those sources with $e_0 < 0.9$, $p_0 > 10$, and a preliminary SNR greater than 1, as estimated with a simplified analytic waveform. Despite these caveats, we emphasize the efficiency of this selection strategy: by retaining on average just $\sim 3\%$ of the sources, we recover nearly 80% of the total signal, significantly reducing the computational cost while preserving the bulk of the information.

4.1.3 Background evalutation

Using populations described in the previous section, we generate the AAK waveform for each EMRI, as introduced in Section 1.3.3, and project the signals into the *AET* TDI channels, assuming an observation time of four years. For each population model, we compute the initial SGWB as the incoherent sum of all signals in A and E channels. SNR threshold of 20 is used to identify resolvable sources, which ensures a sufficiently low false alarm probability, as discussed in [28].
To improve the estimate of resolvable sources, we apply the IFE algorithm developed by [101]. The core idea of the method is to iteratively estimate the PSD of the total noise in each channel, accounting for both instrumental noise and the contribution from the SGWB. At each iteration, the algorithm identifies and subtracts individually detectable sources, updates the noise estimate, and repeats the process until convergence—typically

Model	$\#_{\rm res}$ before IFE	$\#_{\rm res}$	$\rm SNR_{SGWB}$	$\#_{\rm res}$ ([43])	$\rm SNR_{SGWB}$ ([43])
M1	714	197	218	366	460
M2	266	76	86	386	352
M3	1436	348	606	1123	839
M4	2217	541	547	1713	750
M5	821	232	353	622	535
M6	50	17	17	35	27
M7	6333	621	3192	4619	3960
M8	88	28	33	53	51
M9	599	176	281	426	419
M10	549	171	247	406	408
M11	0	0	0.86	0	2.32
M12	7610	590	4013	5580	4533

Table 4.3: Number of resolvable EMRIs and the corresponding SNR of the background from unresolved sources for each model. The second column shows the number of sources before applying the Iterative Foreground Estimation (IFE) algorithm, while the third column shows the number of sources after applying IFE. The fourth column reports the SNR of the unresolved background, while columns 5 and 6 provide a comparison with the results from [43].

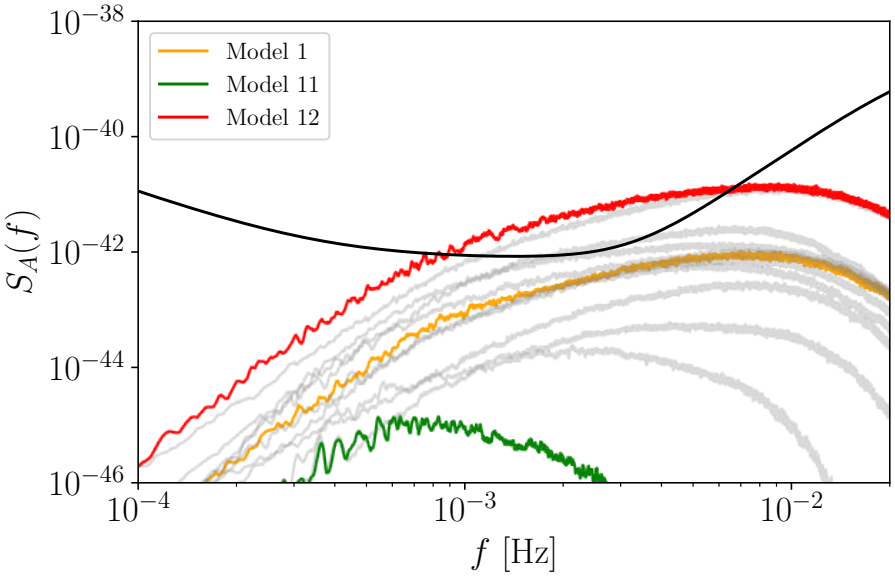

Figure 4.1: SGWB from differen EMRI model populations. The red line corresponds to the optimistic model (M12), the green line to the fiducial scenario (M1), while the orange is the pessimistic one (M11). The black line corresponds to the LISAinstrumental noise, while the shaded grey line are the remaining models.

within one to five iterations.
Fig. 4.1 shows the SGWB spectra for different EMRI population models. The red curve corresponds to the optimistic scenario (M12), the green to the fiducial model (M1), and the orange to the pessimistic case (M11).Table 4.3 summarizes the number of resolvable sources and the SNR of the unresolved background for each model. Uncertainties in the EMRI background span approximately three orders of magnitude, consistent with the rate uncertainties reported in [31]. Notably, several models predict a background comparable to or exceeding the LISA instrumental noise, implying that this astrophysical contribution cannot be neglected in studies of source detectability [43]. For most models, the unresolved SGWB achieves an SNR exceeding 100, making it expected to be detectable. However, background from models M2, M6, and M8 yields a moderate SNR and might be more challenging to observe, whereas M11 produces a negligible signal. A direct comparison with the results of [43] is provided by matching the SNR listed in column 4 and 6 in reported in Table 4.3. These latter values are computed following the methodology of [43], but using our updated source selection criteria described earlier. We find that our SNR estimates are systematically lower by a factor $\lesssim 2$, depending on the model. This discrepancy likely arises from two main differences in methodology. First, our SNR calculations include both instrumental and astrophysical EMRI noise in the denominator, whereas the reference analysis in [43] neglects the astrophysical component. Including the SGWB in the noise budget increases the total noise level, thereby lowering the SNR. Furthermore, neglecting the SGWB during source subtraction leads to inaccurate identification of resolvable sources. Second, the waveform modeling in [43] uses inclination- and polarization-averaged fluxes, whereas our approach directly injects each signal into the TDI channels using its actual

inclination relative to the LISA frame. This allows us to produce a more realistic estimate of the resulting signal.
Table 4.3 also reports the number of resolvable sources subtracted before and after applying the IFE algorithm. Clearly, including the SGWB contribution in the SNR estimation significantly reduces the number of detectable sources. We compare our results with the only two estimates currently available in the literature, namely [43] and [31]. Column 5 of Table 4.3 reports the number of EMRIs detected over a four-year mission, computed following the same procedure as in [43], but updated to reflect our selection criteria. Before applying IFE, our detection rates are broadly consistent with those previously reported. However, after accounting for the IFE, the number of resolvable EMRIs decreases by more than a factor of three in all models—except for M11, where no EMRIs are detectable.

4.1.4 Caveats and takeaways

This study presented a comprehensive assessment of the SGWB generated by unresolved EMRIs for LISA, incorporating the full detector response and realistic waveform models. Our key findings were:

Main results. Except for the pessimistic M11 model, all astrophysical scenarios yielded a detectable background with accumulated SNR ranging from ~ 20 to over 1000 after four years of observation. The signal primarily affected the frequency range 10^{-3}–10^{-2} Hz, with optimistic models (e.g., M12) predicting a background exceeding LISA instrumental noise in this band. In addition, LISA was expected to resolve several hundred individual EMRI systems in optimistic scenarios, although our exclusion of sources with $p_0 < 10$ rendered these estimates conservative.

LISA implications. The EMRI SGWB could significantly affect the total noise level near the LISA sensitivity minimum, potentially compromising the detectability of high-redshift MBHB mergers and SOBBH. This astrophysical foreground has to be carefully accounted for in studies of source detectability and in the design of detection pipelines.

Methodological considerations. Discrepancies with previous literature [43, 31]—including systematically lower SNR estimates and reduced detection counts—arose from two key differences: (i) we included the EMRI SGWB contribution in the noise budget during source subtraction, yielding more realistic identification of resolvable sources, and (ii) we injected signals with their actual inclinations relative to the LISA frame rather than using inclination- and polarization-averaged fluxes. The substantial uncertainties in both amplitude (spanning three orders of magnitude) and spectral shape underscored the need for model-agnostic detection strategies that exploited distinctive statistical signatures such as non-Gaussianity or anisotropies.

4.2 Characterization of environmental-effects on SOBBH background

LISA is well suited to detect the dynamical signatures induced by astrophysical environments [37, 168, 120, 25]. The primary reason for this is that environmental effects are typically more significant earlier in the inspiral. In the LISA band, EMRI binaries are typical probes of the strong dynamics induced by the astrophysical environment [40, 200, 86, 38, 72, 60, 61, 178, 69]. Yet, the resolvable SOBBH that form in gas-rich environments, e.g., in the disks of Active Galactic Nuclei (AGNs), are also potentially sensitive to environmental effects [189, 60, 167, 114]. However, the largest majority of SOBBH will not be detectable in the LISA band, resulting in the build-up of a SGWB, as mentioned in Section 3.6.

Environmental effects typically GN additional energy dissipation that are negative PN relative to the leading point-particle GW flux [38]. Consequently, when the additional energy loss due to the environment dominates over the GW flux at low frequencies, there will be a significant drop in the SGWB relative to the vacuum case. A similar scenario has been proposed to explain Pulsar Timing Array (PTA) measurements [11, 8, 10, 17, 18, 198], where the observed SGWB may originate from a population of supermassive binary black holes influenced by environmental effects [64, 44, 90]. The main environmental effects expected to affect SOBBHs in gas-rich environments are *dynamical friction* and *accretion* [38, 60, 189, 167]. Dynamical friction is the result of the gravitational interaction of each BH with the density wake produced by its motion through a fluid, collisionless [62] or collisional [107, 36, 38]. Accretion affects the binary because the infalling gas carries energy and momentum, which are transferred to the BH and change its mass and momentum [40, 60, 189].

Treating dynamical friction and accretion effects as perturbations on the Keplerian orbital motion, we derive their effect on the spectrum of the stochastic background from a population of SOBBHs consistent with the observational constraints from LVK third observing run (O3) [3, 4]. Specifically, the redshift-dependent merger rate is assumed to follow the cosmic star formation rate [125], weighted by metallicity [116] and convolved with a distribution of time delays [57]. The mass distribution follows the Power Law + Peak model [182] with negligible spins, consistent with LVK observations [2, 135]. With these models and the posterior distributions of their parameters inferred from Refs.[3, 4], we numerically evaluate Eq. 1.85 via Monte Carlo integration to generate posterior predictions for the SGWB. The median SGWB posterior prediction and the corresponding model parameters are adopted as fiducial values. A comparison with the O3-based forecast is shown in Appendix A, demonstrating consistency across the relevant frequency range.

In what follows, we analyze separately the effect of dynamical friction (Section 4.2.1) and accretion (Section 4.2.2) on SGWB spectrum, and we develop phenomenological analytic models for both.

4.2.1 Dynamical friction

Density wakes arise as each BH moves through the accretion disk. When the disk density is higher in the region trailing the BH than in the leading region, an opposing force acts on the BH, giving rise to dynamical friction [62, 140, 107, 36, 39, 38]. Such a force causes

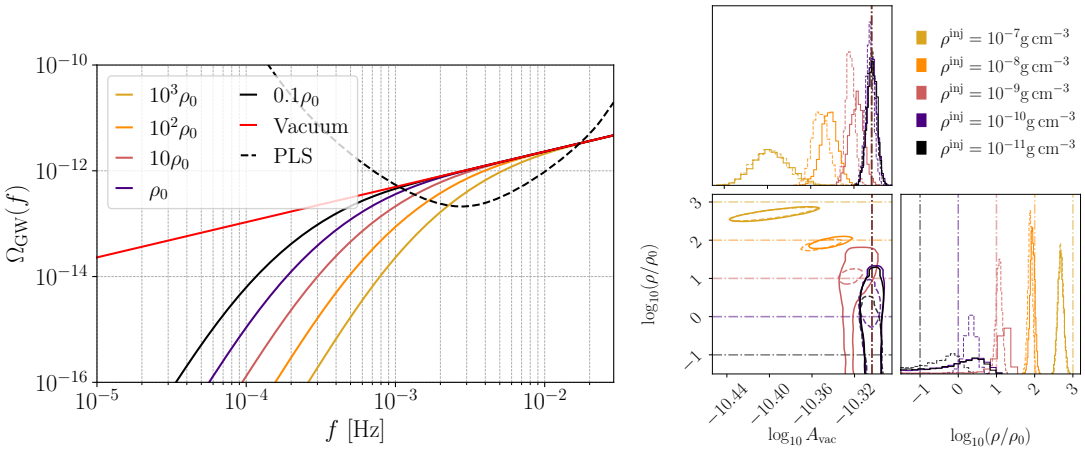

Figure 4.2: (*Left panel*) The SGWB spectra $\Omega_{\rm GW}(f)$ for various disk densities ρ, with the normalization $\rho_0 = 10^{-10}\,{\rm g\,cm^{-3}}$. The solid lines correspond to different values of ρ, with the vacuum case corresponding to $\rho = 0$. The gray dashed (black dotted) curves denotes the LISA PLS, assuming a signal-to-noise ratio of 10. (*Right panel*) Marginalized posteriors for the dynamical friction model across various matter density regimes. Solid (dashed) contours and histograms correspond to analyses with (without) the inclusion of the GF. Two-dimensional contours correspond to 90% credible regions. Dash-dot represent the true value for ρ and the single asymptotic value for $A_{\rm vac}$ (as listed in Table 4.4).

the Binary Black Hole (BBH) system to transfer binding energy to the gas. Thus, in addition to GW emission, energy dissipates through another channel. When such effect dominates over GW emission, a drop in the GW energy spectrum, hence in the SGWB, is expected. To compute it, we choose the center-of-mass frame and assume that the disk is co-moving with the center-of-mass. We model the dynamical friction using a Newtonian approximation, relativistic effects being negligible corrections at the frequencies of interest for this study. Dynamical friction on a BH depends on its mass m_A, the local disk density ρ and speed of sound v_s, their relative velocity $\vec{v}_A$, and the Coulomb logarithm I_A. Doing so, the dynamical friction $\vec{F}_{\rm DF,A}$ on the A-th body is expressed as [107, 140, 62]

$$\vec{F}_{{\rm DF},A} = -\frac{4\pi\rho m_A^2}{v_A^2} I_A \hat{v}_A. \tag{4.6}$$

Further, we assume that each BH is moving at highly supersonic speeds relative to the local sound speed of the disk, i.e. yielding a Mach number $\mathcal{M}_A \equiv v_A/v_s \gg 1$. For a typical environment where the SOBBH is embedded in the AGN disk of a supermassive BH, for a wide range of systems, the center of mass velocity is small compared to the stellar binary velocity [189].
For circular orbits, we only need the azimuthal component of the force, and for supersonic motion, the azimuthal Coulomb logarithm I_A (based on [107]) to leading order in $\mathcal{M}_A \gg 1$ (neglecting $\mathcal{O}(1/\mathcal{M}_A)$ corrections) can then be expressed as

$$I_A = \ln\left(\frac{100 r_A}{11\mathcal{M}_A r_{{\rm min},A}}\right), \tag{4.7}$$

with $r_{{\rm min},A} = 2m_A$ [189]. Since $r_{{\rm min},A}$ is effectively the smallest length scale in the system arising from the regularization of the dynamical friction integrals [107], as long as $r_{{\rm min},A} \ll r_A$, the precise parameterization of $r_{{\rm min},A}$ is not important.

The resulting (outgoing) energy flux lost due to dynamical friction is given by $\dot{E}_{\rm DF} = -\sum_A \vec{F}_{{\rm DF},A} \cdot \vec{v}_A$, becoming

$$\dot{E}_{\rm DF} = \frac{4\pi^{2/3}\rho\, m_1^2 m^{2/3}}{m_2 f_s^{1/3}} \ln\left(\frac{f_1^*}{f_s}\right) + (m_1 \leftrightarrow m_2), \tag{4.8}$$

where $f_A^* = 50 v_s/(11\pi m_A)$. From the energy-balance law, the rate of change of binding energy $E_{\rm b}$ is given by $\dot{E}_{\rm b} = -\dot{E}_{\rm GW} - \dot{E}_{\rm DF}$, yielding an additional contribution to $\dot{f}$:

$$\frac{df_s}{dt} = \left(\frac{df_s}{dt}\right)_{\rm vac} + \frac{12\rho}{m^3\eta^2}\left[m_1^3 \ln\left(\frac{f_1^*}{f_s}\right) + (m_1 \leftrightarrow m_2)\right], \tag{4.9}$$

where $(df_s/dt)_{\rm vac}$ is given by Eq. 1.87.

Thus, the GW energy spectrum is modified in the presence of dynamical friction and reads

$$\left(\frac{dE_{\rm GW}}{df_s}\right)_{\rm DF} = \left(\frac{dE_{\rm GW}}{df_s}\right)_{\rm vac} \left\{1 + \frac{5\rho f_s^{-11/3}}{8\pi^{8/3}\eta^3 m^{14/3}} \times \left[m_1^3 \ln\left(\frac{f_1^*}{f_s}\right) + (m_1 \leftrightarrow m_2)\right]\right\}^{-1}, \tag{4.10}$$

where $(dE_{\rm GW}/df_s)_{\rm vac}$ is given by Eq. 1.86. The energy flux due to dynamical friction is a -5.5PN relative to the GW flux, owing to the $f_s^{-11/3}$ scaling. Consequently, at low frequencies dynamical friction will dominate the orbital evolution and energy loss, and can deplete the SGWB.

At a critical turning point frequency $f_{\rm turn,DF}$, the flux due to dynamical friction becomes comparable to that of GW emission. For $f_s < f_{\rm turn,DF}$, the amplitude of the GW spectrum will start to decrease as more energy flows through the dynamical friction channel than the GW emission channel. We compute $f_{\rm turn,DF}$ from $|\dot{E}_{\rm DF}| = |\dot{E}_{\rm GW}|$, which amounts to setting the ρ-dependent term in the denominator of Eq. 4.10 to 1.

To gain insight onto $f_{\rm turn,DF}$, let us note that typical astrophysical systems yield $f_{A,*} \sim 10^{-1}$Hz, much higher than the observed frequencies that we want to probe. Observing that $f_{A,*} \gg f_{\rm turn,DF}$ and focusing on nearly equal mass systems, we obtain

$$\frac{f_{\rm turn}}{(\ln[1{\rm Hz}/{\rm f_{turn}}])^{3/11}} \approx (1.69\times10^{-3}{\rm Hz})\left(\frac{m}{50M_\odot}\right)^{-5/11} \times\left(\frac{\rho}{10^{-10}{\rm g\,cm^{-3}}}\right)^{3/11}, \tag{4.11}$$

where we have scaled the estimate for typical astrophysical thin disk densities of $\rho \sim 10^{-10}{\rm g\,cm^{-3}}$ and typical stellar masses of $m \sim 50M_\odot$.

In the left panel of Fig. 4.2, we show the SGWB computed with Eq. 4.10. As for the vacuum case, we perform a three-dimensional Monte Carlo population integral. SGWB are bracketed by a few reference ρ values. As ρ is increased, the turning point in the spectrum does increase as expected from Eq. 4.11. Further, the spectral index of the SGWB asymptotes to a value of 13/3 at low frequencies and 2/3 at high frequencies, the latter being the vacuum result.

4.2.2 Gas accretion

The gas from the surrounding disk will accrete onto the two BHs. We model the accretion of the A-th body as $\dot{m}_{A,\rm Edd} = L_{A,\rm Edd}/\zeta$, where $L_{A,\rm Edd}$ is the Eddington luminosity and ζ is the radiative efficiency [38]. We pick a conservative value for the radiative efficiency of $\zeta = 0.1$ and the resulting accretion rate is $\dot{m}_{A,\rm Edd} \simeq 2.2\times10^{-8}(m_A/M_\odot)M_\odot{\rm yr}^{-1}$ [38, 60]. For simplicity, we parameterize the accretion rate of both BHs by the same Eddington ratio $f_{\rm Edd} \equiv \dot{m}_A/\dot{m}_{A,\rm Edd}$. Doing so, the mass as a function of time reads [60]

$$m_A(t) = m_{A,0}e^{f_{\rm Edd}t/\tau}, \tag{4.12}$$

where $\tau = 4.5\times10^7{\rm yr}$ is known as the Salpeter time scale, and $m_{A,0}$ is the initial mass of the A-th BH of binaries.

For a slowly accreting binary, the component masses evolve adiabatically, satisfying the condition $\dot{m}_A \ll m_A\omega_s/(2\pi)$. In addition to the adiabatic mass increase of each body, the accreted material will additionally carry some angular momentum, which results in a hydrodynamic drag torque imparted on each body. We parameterize this drag force as:

$$\vec{F}_{{\rm Acc},A} = -\xi\dot{m}_A\vec{v}_A,, \tag{4.13}$$

where $\xi \sim \mathcal{O}(1)$ is the linear hydrodynamic drag coefficient that captures the effect of momentum transferred from the accreted material [189, 38, 86, 40]. Since angular momentum is an invariant under an adiabatic change in the masses [115], we can use angular

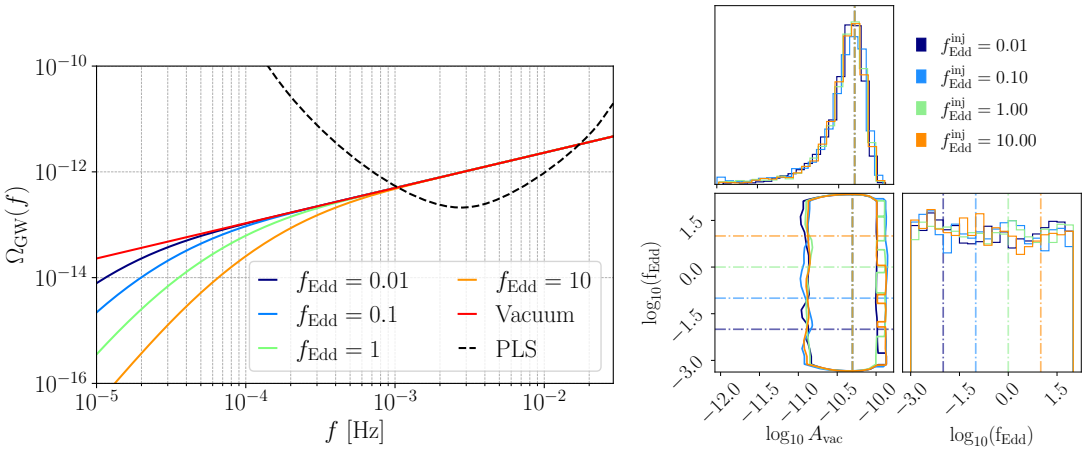

Figure 4.3: (*Left panel*) The SGWB spectra $\Omega_{\rm GW}(f)$, driven by gas accretion, for a range of Eddington ratios $f_{\rm Edd}$. The solid lines correspond to different values of $f_{\rm Edd}$, with the vacuum case corresponding to $f_{\rm Edd} = 0$. The gray dashed and black dotted curves indicate the LISA PLS, assuming a signal-to-noise ratio of 10. (*Right panel* Marginalized posterior probability for the gas accretion model across different Eddington ratios. Dash-dot lines represent the true value of $f_{\rm Edd}$ and the asymptotic value of $A_{\rm vac}$ (as listed in Table 4.4). Two-dimensional contours correspond to 90% credible regions. Observe that the posteriors of $f_{\rm Edd}$ are identical to the prior, implying gas accretion effects cannot be inferred.)

momentum balance law to obtain the back-reaction $\dot{f}_s$ under gas accretion [60]. In doing so, we obtain

$$\left(\frac{df_s}{dt}\right)_{\rm Acc} = (5+3\xi)\frac{f_{\rm Edd}}{\tau} f_s. \tag{4.14}$$

Thus, using Eq. 4.14 together with Eqs. 1.86 and 1.87, we obtain the energy spectrum of gravitational waves in the presence of accretion as

$$\left(\frac{dE_{\rm GW}}{df_s}\right)_{\rm Acc} = \left(\frac{dE_{\rm GW}}{df_s}\right)_{\rm vac} \times \left[1 + \frac{5(f_{\rm Edd}/\tau)(5+3\xi)}{96\pi^{8/3} m_0^{5/3}\eta} f_s^{-8/3}\right]^{-1}, \tag{4.15}$$

where $(dE_{\rm GW}/df_s)_{\rm vac}$ is given Eq. 1.86 and evaluated with the initial value of the masses.

Similar to the case of dynamical friction, the stochastic GW signal will have a turning point due to additional energy dissipation in the presence of accretion. At this turning point, the energy flux from gas accretion becomes comparable to that from gravitational-wave emission, i.e. $|\dot{E}_{\rm Acc}| = |\dot{E}_{\rm GW}|$, which is equivalent to setting the $f_{\rm Edd}$-dependent term in the denominator of Eq. 4.15 to 1.

We consider typical stellar binaries embedded in thin disks and estimate the turning point analytically as which

$$f_{\rm turn} \approx 10^{-4}{\rm Hz} \left(\frac{f_{\rm edd}}{1}\right)^{3/8} \left(\frac{m}{50\,M_\odot}\right)^{-5/8} \times \left(\frac{5+3\xi}{8}\right)^{3/8} \left(\frac{\tau}{4.5\times10^7\,{\rm yr}}\right)^{-3/8}. \tag{4.16}$$

In addition, as the total mass increases, the turning point shifts toward lower frequencies. For typical values of $f_{\rm Edd} \sim 1$, we find that $f_{\rm turn} \sim 10^{-4}$Hz, implying that the effects are significant only at the lower end of the LISA sensitivity band. Only with much higher values of $f_{\rm Edd} \sim 10^3$, which may not be astrophysically realistic, we have $f_{\rm turn} \sim 10^{-3}$Hz, which is comparable to that of dynamical friction. As expected, the effect of gas accretion occurs at -4PN order.

The key reason that gas accretion has an overall lower turning point than dynamical friction is its higher (less negative) PN order relative to GW emission.

In the left panel of Fig. 4.3, we show the SGWB computed with Eq. 4.15 for different astrophysical values of $f_{\rm Edd}$, and we set $\xi = 1$ for simplicity. As expected, the signal is suppressed at low frequencies due to the additional energy loss channel, and asymptotes to the vacuum SGWB at high frequencies. Although the stochastic signals shown in Fig. 4.3 lie above the PLS curve for the frequency range of $f \gtrsim 10^{-3}$Hz, suggesting that they are detectable, the turning points are outside the LISA SGWB sensitivity band for the range of astrophysical $f_{\rm Edd}$ values considered here. Thus, we do not expect to clearly measure $f_{\rm Edd}$ with LISA and distinguish gas accretion from vacuum.

4.2.3 Parametrisation and inference

The stochastic signals described in Sections 4.2.1 and 4.2.2 are computationally expensive to evaluate for data analysis purposes. In order to analyze the effects of dynamical friction

Parameter	**Dynamic Friction**	**Accretion**
α	ρ	$(5+3\xi)\, f_{\rm Edd}/\tau$
β	$-11/3$	$-8/3$
κ	1	0
Asymptotic matching parameters		
$A_{\rm vac}$	4.97×10^{-11}	4.97×10^{-11}
A_m	-2.73×10^{-7}	4.35×10^{2}
Gaussian spectral correction parameters		
A_g	$1.02 - 0.0262 \log_{10}(\alpha/\alpha_0)$	0.904
$f_{\rm peak}$	$0.0269\,(\alpha/\alpha_0)^{0.262}$	$9.73\,(\alpha/\alpha_0)^{0.374}$
σ	0.222	0.321

Table 4.4: The mapping parameters under different environmental effects. The first three rows correspond to the environmental coefficient and respective spectral indices. The asymptotic matching and Gaussian correction parameters under different environmental effects are given in the next set of rows. All coefficients are scaled to convenient SI units as follows: $A_{\rm vac}$ in $\mathrm{Hz}^{-2/3}$; A_m in $\mathrm{kg}^{-1}\mathrm{m}^3\mathrm{s}^{-11/3}$ (dynamical friction) and $\mathrm{s}^{-5/3}$ (accretion); $f_{\rm peak}$ in Hz. Here, α denotes the mapping parameter, normalized by $\alpha_0 = \mathrm{kg\,m}^{-3}$ (dynamical friction) and s^{-1} (accretion).

and accretion, we develop "ready-to-use" phenomenological models. As a first step, we construct a Rational Power Law (RPL) model $\Omega_{\rm RPL}$ given by

$$\Omega_{\rm RPL} = \frac{A_{\rm vac} f^{\gamma}}{1 + A_{\rm m} \alpha f^{\beta} [\ln(f/1\mathrm{Hz})]^{\kappa}}, \tag{4.17}$$

where $A_{\rm vac}$ is the vacuum amplitude, γ the vacuum spectral index, α is a phenomenological coefficient, $\{\beta, \kappa\}$ are spectral indices parameterizing each environmental effect, and A_m controls the fractional change in the SGWB amplitude due to the environment. We obtain $A_{\rm vac}$ and A_m using asymptotic matching at low and high frequencies. Specifically, we expand both $\Omega_{\rm RPL}$ and $\Omega_{\rm GW}$ to leading order for $f_s \ll f_{\rm turn}$ and for $f_s \gg f_{\rm turn}$ and solve for $A_{\rm vac}$ and A_m. Values for $A_{\rm vac}$ and A_m are listed in Table 4.4. At intermediate frequencies $f_s \sim f_{\rm turn}$, the asymptotic expansions break down and $\Omega_{\rm RPL}$ is not a sufficiently accurate approximation. To improve the model accuracy at intermediate frequencies, we include a Gaussian correction to the RPL model. The resulting Rational Power Law + Peak (RPLP) model $\Omega_{\rm RPLP}$ performs well across all frequencies. The model $\Omega_{\rm RPLP}$ is given by

$$\Omega_{\rm RPLP} = \frac{\Omega_{\rm RPL}}{1 + \mathcal{G}(f, \alpha)}, \tag{4.18}$$

where the Gaussian correction $\mathcal{G}(f,\alpha)$ captures deviations near the turning point correction and reads

$$\mathcal{G}(f,\alpha)=A_g(\alpha)\exp\left\{-\frac{[\log_{10}(f/f_{\rm peak}(\alpha))]^2}{2\sigma^2}\right\}. \tag{4.19}$$

Here $A_g(\alpha)$, $f_{\rm peak}(\alpha)$ and σ control the amplitude, central log-frequency, and width of the correction to $\Omega_{\rm RPL}$. In Table 4.4, we show the mapping between model parameters and the specific cases of dynamical friction and gas accretion.

Afterwards, we perform several injection-recovery studies to assess the measurability of the environmental effects. Specifically, we assess separately the measurability of dynamical friction and gas accretion parameters. To quantify the impact of the GF, we perform a separate set of analyses where its parameters are assumed perfectly known. Environmental effects are typically measurable when the recovered one-dimensional marginal posterior is well constrained relative to the prior. To quantify the measurement precision for a parameter θ, we quote the (symmetric, fractional) statistical uncertainty corresponding to the 90% credible interval given by

$$\frac{\delta\theta}{\theta^{\rm inj}}=\left|\frac{\theta^{95\%}-\theta^{5\%}}{2\theta^{\rm inj}}\right|, \tag{4.20}$$

where $\theta^{\rm inj}$ denotes the true injected parameter value. We further assess the distinguishability from a vacuum signal by computing the Bayes factor between the non-vacuum and vacuum hypotheses.

We model LISA data in the AET combinations. Under the assumption of static arms and equal link noises (see Sections 3.3 and 3.4), we do not consider TDI correlation. Moreover, we do not account for additional features beyond a stationary Gaussian spectrum such as anisotropy, non-Gaussianity, and non-stationarity.

In each TDI channel, we perform inference on frequency-domain data that have been coarse-grained over N neighboring points, resulting in $D = 5096$ segments. The likelihood associated with the coarse-grained spectrum is expected to follow a complex Wishart distribution, as shown in Eq. 2.57. However, under the above conditions, the likelihood simplifies to a Gamma distribution [21], where the covariance matrix is diagonal for each frequency bin in the TDI domain.

In the most general case, PSD includes contributions from instrumental noise, the GF, and the superposition of multiple SGWBs, possibly accounting for environmental effects. We adopt a two-parameter model for the instrumental noise S_n, describing the spectral amplitudes of the TM and OMS noises, respectively, with parameters $\boldsymbol{\theta_n}$ (Eq. 3.60).

We adopt the phenomenological model for the GF, given in Eq. 3.62 with $\boldsymbol{\theta_{\rm Gal}} = \{\alpha_{\rm Gal}, A_{\rm Gal}, f_{\rm kn}, f_1, f_2\}$.

We simulate data for $S_{\rm GW,SOBBH}$ using $\Omega_{\rm GW,SOBBH}$ models introduced in Sections 4.2.1 and 4.2.2. Instead, for parameter estimation, we use the RPLP models introduced in Eq. 4.17 and Eq. 4.18, parameterized by $\boldsymbol{\theta_{\rm SOBBH}} = \{A_{\rm vac}, \gamma, \alpha, \beta, \kappa, A_m\}$. Thus, in the most general case, the inference parameter space is $\boldsymbol{\theta} = \boldsymbol{\theta_n} \cup \boldsymbol{\theta_{\rm Gal}} \cup \boldsymbol{\theta_{\rm SOBBH}}$.

Dynamical friction

We first consider the effects of dynamical friction by generating injected SGWB data (using methods of Section 4.2.1) with $\rho^{\rm inj} \in \{10^{-7}, 10^{-8}, 10^{-9}, 10^{-10}, 10^{-11}\}\,{\rm g\,cm^{-3}}$. For the

RPLP model parameters $\theta_{\rm SOBBH}$, we are mainly interested in recovering the vacuum amplitude $A_{\rm vac}$ and the environmental parameter α, which in this case is just the disk density ρ. Since $A_{\rm vac}$ and γ, are strongly correlated, we fix $\gamma = 2/3$ in our inference. Likewise, we also set $\{A_m, \beta, \kappa\}$ to the values listed in Table 4.4, due to the strong correlation with α. These parameters are therefore not sampled in our parameter estimation experiments. We use log-uniform priors $\log_{10} A_{\rm vac} \sim \mathcal{U}(-12, -8)$ to be agnostic regarding the order of magnitude of the vacuum amplitude. Based on astrophysically motivated values for the disk density [39, 38], we take a conservative upper bound on $\rho \lesssim 10^{-6}\mathrm{g\,cm^{-3}}$, consistent with the perturbative regime of dynamical friction with respect to the Keplerian motion. We set log-uniform priors $\log_{10}(\rho/\rho_0) \sim \mathcal{U}(-2, 4)$, with $\rho_0 = 10^{-10}\mathrm{g\,cm^{-3}}$.

We checked, by further lowering the lower prior bound, that posteriors and evidences are not affected. Therefore, we consider our results robust in the limit of $\rho \to 0$.

For the GF parameters, we use the following uniform priors: $\alpha_{\rm vac} \sim \mathcal{U}(0, 2.5)$, $\log_{10} A_{\rm Gal} \sim \mathcal{U}(-47, -40)$, $\log_{10} f_{\rm kn}/\mathrm{Hz} \sim \mathcal{U}(-4, -2)$, $\log_{10} f_1/\mathrm{Hz} \sim \mathcal{U}(-4, -2)$, $\log_{10} f_2/\mathrm{Hz} \sim \mathcal{U}(-4, -2)$.

In the right panel of Fig. 4.2, we show the one-dimensional and two-dimensional marginalized posteriors for $A_{\rm vac}$ and ρ, inferred with fixed (dashed curves) or free (solid curves) $\theta_{\rm Gal}$. Overall, we find that the recovered posteriors are well constrained relative to the prior, and the effect of dynamical friction is indeed measurable.

For known GF parameters, we observe a trend in the posteriors for increasing $\rho^{\rm inj}$. When $\rho^{\rm inj} \gtrsim 10^{-9}\mathrm{g\,cm^{-3}}$, the recovered posteriors of ρ do not overlap, implying that we can better resolve between different orders of magnitude of the disk density. We quantify this further by computing the statistical uncertainty $\delta\rho/\rho$ using Eq. 4.20. As listed in Table 4.5, with increasing $\rho^{\rm inj}$, we find that $\delta\rho$ decreases. Specifically, we find that $\delta\rho \lesssim \rho^{\rm inj}$ when $\rho^{\rm inj} \gtrsim 10^{-9}\mathrm{g\,cm^{-3}}$, with known GF parameters. In all cases, we can infer the vacuum amplitude $A_{\rm vac}$ to $\mathcal{O}(1)\%$ precision. The fact that we can strongly measure the vacuum amplitude is consistent with the results of [29], which also investigated the detectability of SGWB produced by SOBBHs in vacuum. While the one-dimensional marginalized posterior of ρ shrinks with increasing $\rho^{\rm inj}$, the one-dimensional marginalized posterior of $A_{\rm vac}$ widens instead due to the correlation between the two parameters.

We now discuss the impact of the GF parameters on the marginalized posteriors shown in Fig. 4.2. As expected, the simultaneous inference on GF parameters results in wider marginal posteriors. For $\rho^{\rm inj} \lesssim 10^{-9}\mathrm{g\,cm^{-3}}$, however, fixing the GF parameters $\theta_{\rm Gal}$ causes an underestimation of $A_{\rm vac}$ and ρ compared to inferring on them. At higher densities, the influence of GF becomes negligible. This is related to how the SNR is affected by the GF for different densities.

In Fig. 4.4, we show the cumulative SNR with and without including the GF parameters, for different values of $\rho^{\rm inj}$. We find that as $\rho^{\rm inj}$ increases, the total SNR decreases, consistent with the behavior of the SGWB shown in Fig. 4.2. Further, including the GF lowers the SNR, which along with the additional five GF parameters, contributes to wider posteriors. From Fig. 4.4, we also observe that with increasing $\rho^{\rm inj}$, the difference in SNR with and without GF decreases. This is because as $\rho^{\rm inj}$ increases, the turning point shifts to frequencies above $\sim 5 \times 10^{-3}\mathrm{Hz}$, where the GF is suppressed. Consequently, the posteriors of ρ with and without GF overlap better as $\rho^{\rm inj}$ is increased, with virtually no difference in the case of $\rho^{\rm inj} \gtrsim 10^{-8}\mathrm{g\,cm^{-3}}$. Additionally, the recovered SNR (with and without GF) typically accumulates above 50 only at $f \gtrsim (\text{few}) \times 10^{-3}\mathrm{Hz}$. Thus, the measurements of signal parameters are typically informed by this specific frequency regime.

For each injection, we performed a separate set of runs with a vacuum SGWB model

$\rho^{\rm inj}/\rho_0$	$\delta\rho/\rho^{\rm inj}$		$\delta A_{\rm vac}/A^{\rm inj}_{\rm vac}$		$\log_{10}\mathcal{B}^{\rm non-vac}_{\rm vac}$	
	With	Without	With	Without	With	Without
10^{-1}	37.85	8.218	0.02188	0.01727	1.424	1.010
1	4.248	1.4670	0.02075	0.01490	1.146	3.965
10	1.118	0.4042	0.02114	0.01430	1.777	8.240
10^2	0.3362	0.2033	0.03404	0.02478	1.281	10.16
10^3	0.1606	0.1340	0.05334	0.0493	1.427	7.570

Table 4.5: Posterior statistical uncertainties (from 90% quantiles) on ρ and $A_{\rm vac}$ together with the Bayes factor between the non-vacuum and vacuum hypotheses. For each injection, results are listed with uncertain (With) or fixed (Without) GF parameters. We observe that the precision of ρ improves with increasing $\rho^{\rm inj}$, while that of $A_{\rm vac}$ worsens. Although the Bayes factor decreases significantly when inferring on the GF, in all cases the non-vacuum model is strongly preferred.

by setting $\alpha = \rho = 0$ in $\Omega_{\rm RPLP}$. We compute the Bayes factor $\mathcal{B}^{\rm non-vac}_{\rm vac}$ in favor of the non-vacuum model ($\rho \neq 0$), which we list in Table 4.5. For all injections, we find that $\mathcal{B}^{\rm non-vac}_{\rm vac} > 10$, even when including the GF. Thus, despite the marginalized posteriors for $A_{\rm vac}$ and ρ displaying biases, the RPLP model can detect dynamical friction effects by strongly disfavoring the vacuum model. We also checked that the posterior predictive from the RPLP model accurately describe the injected signal within statistical errors, consistent with our Bayes factor results.

If we detect a SGWB consistent with vacuum, our RPLP model can also be used to place upper bounds on the environmental parameters. In Fig. 4.5, we show the constraints on ρ, with and without including the GF parameters in our model. The constraints, as given by the 90% one-sided credible interval, are informative, because the prior extends to $\rho = 10^{-6}\,{\rm g\,cm^{-3}}$. When excluding the GF, we find an upper bound of $\rho < 8.51 \times 10^{-11}\,{\rm g\,cm^{-3}}$. Owing to the increase in dimensionality when including GF, the constraint weakens slightly to $\rho < 7.58 \times 10^{-10}\,{\rm g\,cm^{-3}}$. To explore systematic biases on the vacuum SGWB induced by neglecting environmental effects, we inject a SGWB containing dynamical friction effects, and analyze it with the vacuum model, i.e. $\Omega_{\rm RPL}$ with $\alpha = 0$. In Fig. 4.6, for the injected values of $\rho^{\rm inj} \in \{10^{-7}, 10^{-8}, 10^{-9}, 10^{-10}, 10^{-11}\}\,{\rm g\,cm^{-3}}$, we show the recovered two- and one-dimensional marginalized posteriors for the vacuum amplitude $A_{\rm vac}$ and the vacuum spectral index γ. When assessing the systematic bias in $A_{\rm vac}$ and γ, we compare the maximum posterior point to the asymptotic vacuum values (vertical dash-dotted lines) listed in Table 4.4. As expected, with increasing $\rho^{\rm inj}$, the systematic biases increase in significance.

In more details, γ is biased to values larger than the asymptotic value of $\gamma = 2/3$, because the SGWB containing dynamical friction effects has a steeper slope, with an asymptotic value of 13/3. Due to the expected positive correlation with $A_{\rm vac}$ (given the functional

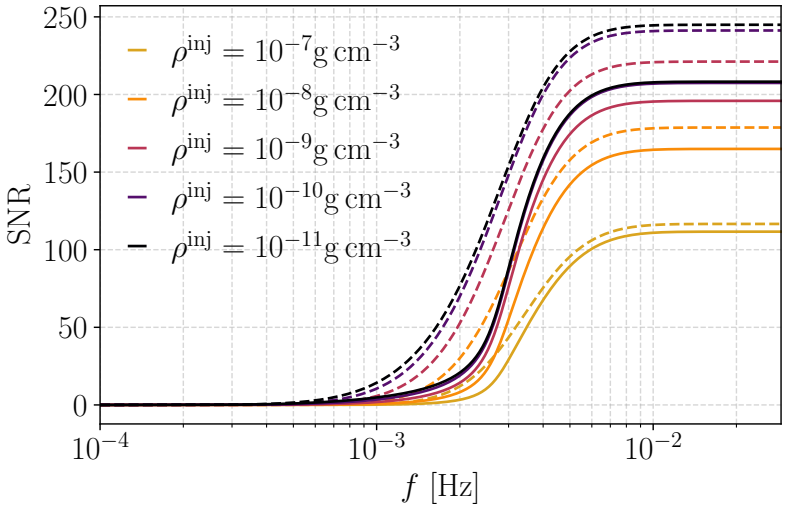

Figure 4.4: Cumulative SNR of the best-fit recovered SGWB model as a function of frequency. The solid and dashed lines represent the cases with and without the GF respectively, while the different colored lines correspond to different $\rho^{\rm inj}$.

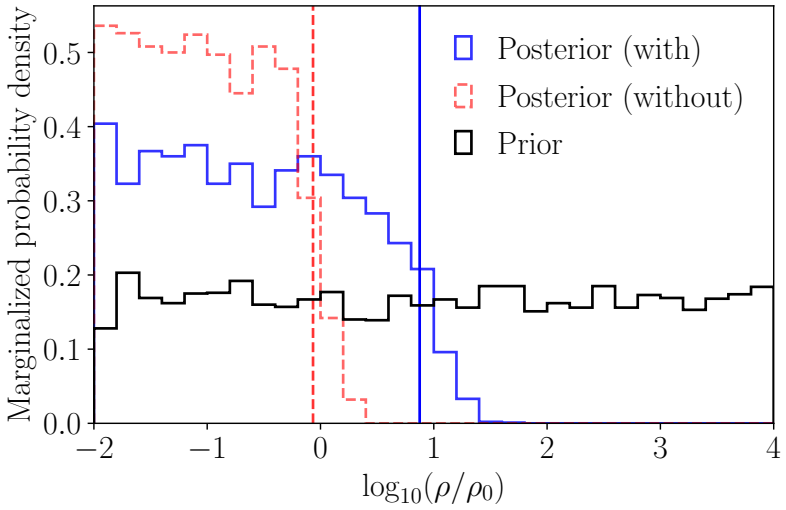

Figure 4.5: Marginalized posterior probability for $\log_{10}(\rho/\rho_0)$ shown for the cases with (blue histogram) and without (red dashed histogram) GF parameters included. The vertical lines indicate the one-sided 90% credible intervals, and the black histogram indicates the uniform prior on $\log_{10}(\rho/\rho_0)$. Observe that the constraint is slightly weaker when including the Galactic parameters.

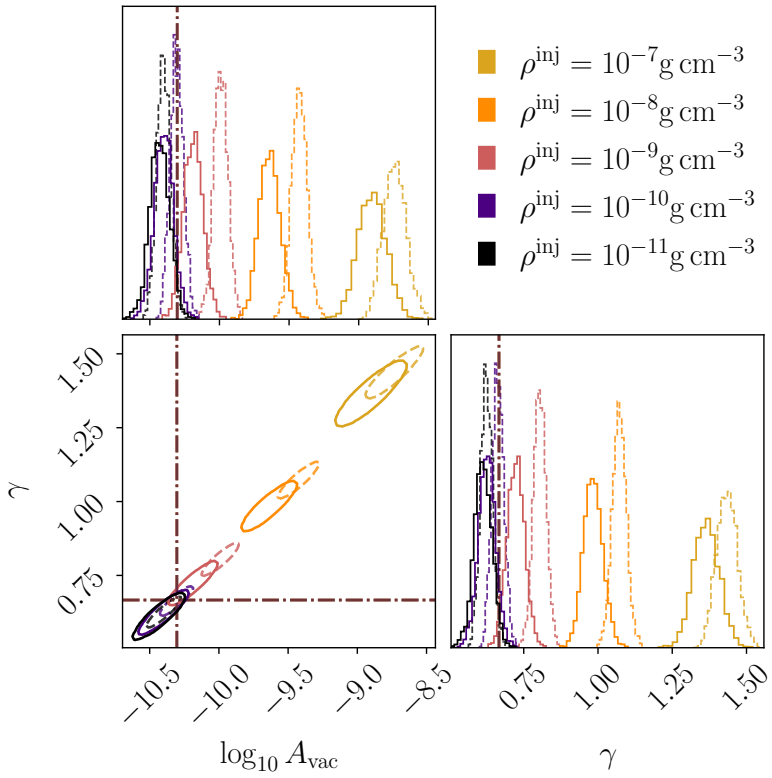

Figure 4.6: Marginalized posterior probability for the vacuum model across different matter density regimes. Solid and dashed histograms represent the inference results with and without the inclusion of the GF, respectively. Dash-dotted lines indicate the asymptotic value of $A_{\rm vac}$ (from Table 4.4) with $\gamma = 2/3$. Two-dimensional contours correspond to 90% credible regions.

form of the SGWB), we find that $A_{\rm vac}$ is biased to larger values. When simultaneously inferring the GF parameters, the reduced constraining power on dynamical friction, shown in Fig. 4.2, leads to correspondingly milder biases on $A_{\rm vac}$ and γ, as expected.

Gas accretion

To characterize gas accretion measurability, we inject SGWB data (using models in Section 4.2.2) with $f_{\rm Edd}^{\rm inj} \in \{0.01, 0.10, 1.00, 10\}$. We use a log-uniform prior given by $\log_{10} f_{\rm Edd} \sim \mathcal{U}(-3, 2)$, where the upper bound reflects astrophysical expectations of how large the Eddington ratio can be. We have checked that our results are robust to the exact choice of the lower bound, which we cannot set exactly to zero due to the log-uniform prior. We carry out the Bayesian inference just like we did for dynamical friction by fixing $\{\gamma, A_m, \beta, \kappa\}$.

In the right panel of Fig. 4.3, we show the two-dimensional and one-dimensional marginalized posteriors of $A_{\rm vac}$ and $f_{\rm Edd}$. We accurately measure the vacuum amplitude across all injections, thus finding that marginalizing over gas accretion does not impact the measurability of the vacuum stellar SGWB. However, we do not obtain informative pos-

teriors of Eddington ratios for any of the injected values, even with $f_{\rm Edd}^{\rm inj} = 10$ marginally recovering the prior. The lack of constraining power on $f_{\rm Edd}$ is due to the turning point of the SGWB occurring at frequencies lower than the LISA sensitive band. Given our findings, we do not perform any additional Bayesian analysis for gas accretion effects.

4.2.4 Caveats and takeaways

We investigated the detectability of environmental signatures—dynamical friction and gas accretion—in the SOBBH SGWB, developing phenomenological templates suitable for rapid parameter estimation.

Main results. For typical disk densities $\rho \sim 10^{-10}$–$10^{-9}\,{\rm g\,cm^{-3}}$, dynamical friction effects were measurable with LISA: the non-vacuum model was strongly favored over the vacuum hypothesis even when marginalizing over Galactic foreground parameters, with Bayes factors exceeding 10. In contrast, gas accretion remained undetectable even for Eddington ratios $f_{\rm Edd} = 10$, as the characteristic frequency turnover lay below the LISA sensitive band. Neglecting environmental effects when present induced significant systematic biases in the inferred vacuum amplitude and spectral index, potentially mischaracterizing the underlying SOBBH population.

Model limitations. Our analysis assumed a monochromatic environmental model—all non-vacuum SOBBHs experienced identical disk densities and Eddington ratios. Realistically, a distribution of environmental parameters across the population would modify the resulting SGWB. We further neglected orbital eccentricity, which suppressed the SGWB at low frequencies [64] and could be degenerate with environmental effects. Moreover, dynamical friction and gas accretion themselves influenced eccentricity evolution [40, 86, 61], motivating further study of how environmental processes affected the eccentricity distribution in non-vacuum SOBBH populations.

LISA implications. The measurability of dynamical friction signatures provided a novel avenue for constraining the astrophysical environments of SOBBH formation, particularly in AGN disks. However, confusion from overlapping SGWB components—including EMRI and extragalactic white dwarf backgrounds—could complicate detection. Optimistically, a ground-based detection of the SOBBH SGWB could inform priors for LISA analyses, enhancing detectability despite foreground contamination.

4.3 Flexible parametric approach

Disentangling the various components of the SGWB from instrumental noise is one of the key challenges in LISA data analysis. Several approaches have been proposed in the literature [7, 42, 59, 32, 165], falling into two broad categories: some rely on predefined templates for both noise and SGWB spectra, while others relax these assumptions for one of the components. The analyses in Sections 4.1 and 4.2 reveal significant uncertainties in both the amplitude and spectral shape of astrophysical SGWB sources. For EMRIs, the uncertainty is primarily due to the lack of constraints on population properties, with predictions spanning approximately three orders of magnitude; for SOBBHs, additional uncertainties are introduced by possible environmental effects that can substantially modify the spectral shape at low frequencies. These findings motivate the development of inference methods that do not rely on strong template assumptions. In this section, we introduce a weakly-parametric approach by implementing an inference based on stochastic Gaussian Processes (GPs), which can be applied with great flexibility in the SGWB spectral shape. The framework is implemented within `balrog` [176, 154, 177, 55, 109, 82], and we demonstrate its capabilities by applying it to the EMRI background characterized in Section 4.1. The method ability to capture complex spectral features without strong prior assumptions provides an important complement to the parametric environmental models explored in Section 4.2.

4.3.1 Model definition

A GP is a stochastic model, formally describing distributions over functions $g(x)$. It is parameterized by a mean function $m(x)$ and positive definite covariance function $k(x, x')$. Realizations from such process are denoted as follows

$$g(x) \sim \mathcal{GP}(m(x), k(x, x')), \tag{4.21}$$

The defining properties of GPs is that for any finite subset $X = \{x_1 \dots x_n\}$ of the domain for x, the marginal distribution is a multivariate Gaussian

$$\boldsymbol{g}(X) \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma), \tag{4.22}$$

with mean vector and covariance matrix defined by $\mu = m(x_i)$ and $\Sigma_{ij} = k(x_i, x_j)$, respectively. Typically, GP are used to flexibly incorporate some observed data in the distribution over $\boldsymbol{g}(X)$ and make predictions on new domain points X_* as $\boldsymbol{g}(X_*)$. The joint probability distribution $p(\boldsymbol{g}(X_*), \boldsymbol{g}(X))$ is a multivariate normal distribution. In order to employ this formalism as a regression model, one needs the conditional probability $p(\boldsymbol{g}(X_*)|\boldsymbol{g}(X))$. Upon suitable marginalization, the conditional distribution is also a multivariate normal distribution

$$p(\boldsymbol{g}(X_*)|\boldsymbol{g}(X)) = \mathcal{N}(\boldsymbol{\mu}\left(X_*|X\right), \Sigma\left(X_*|X\right)), \tag{4.23}$$

where

$$\boldsymbol{\mu}\left(X_*|X\right) = \boldsymbol{\mu}(X_*) + \Sigma\left(X_*, X\right)\Sigma\left(X, X\right)^{-1}\left(\boldsymbol{g}(X) - \boldsymbol{\mu}(X)\right) \tag{4.24}$$

and

$$\Sigma\left(X_*|X\right) = \Sigma\left(X_*, X_*\right) + \Sigma\left(X_*, X\right)\Sigma^{-1}\left(X, X\right)\Sigma\left(X, X_*\right), \tag{4.25}$$

are specified by the single kernel function $k(x, x')$. The matrix Σ is also referred to as *kernel matrix*, and it models the covariance between each pair of its two arguments through the definition of the bivariate function k. In order to be a valid covariance for the multivariate Gaussian in Eq. 4.23, the kernel matrix must be symmetric and positive definite. A variety of kernels are available in literature, to capture different processes peculiarities. In this work, we consider the radial basis kernel function (RBF) defined by

$$\Sigma_{ij}^{\mathrm{RBF}}(X, Y) = k_{\mathrm{RBF}}(x_i, y_j) = \exp\left(-\frac{|x_i - y_j|^2}{2\sigma^2}\right), \tag{4.26}$$

where σ is a model positive hyperparameter.

In the present study, we exploit GPs in an unconventional way: we use each EGP from Eq. 4.24 as a proposed spectrum, and set $\boldsymbol{g}(X)$ as free parameters. The GP covariance is not used in the inference model. The dimensionality and domain location of X are to be considered hyperparameters.

Henceforth, we model the logarithm of noise and signal PSDs as two independent EGPs over the logarithmic frequencies, i.e. $X = \{\log_{10} f_1, \ldots, \log_{10} f_n\}$. We illustrate both in the following sections. We will refer to the parameters X and to the mean function $m(X)$ as *knots* and *baseline*, respectively.

Signal

We assume the response matrix $R(f, t_0)$ in Eq. 3.54 to be exactly known, allowing us to model the signal directly at the link level through $S_h(f)$.

We use an EGP, with baseline and knots left free to vary simultaneously. We do so to be able to capture both global spectral shapes and fine structure in specific frequency regions: the choice of baseline influences globally the spectral-shape proposals, while knots control the fine-local structure. Signal detection is robust against variations in the baseline shape. The knots variability can compensate easily for misidentification of the truth baseline family. In our inference, we use a power-law baseline family for $\log_{10} S_h(f)$, parametrized by logarithmic amplitude $\log_{10} A$ at a reference frequency $f_\star = 10^{-2.5}$Hz and slope γ. Thus, $\boldsymbol{\mu}(X_*)$ in Eq. 4.24 reads $\log_{10} A + \gamma \log_{10}(f/f_\star)$. Fine-structure deviations with respect to the baseline are parameterized through knots $\boldsymbol{\delta}^h$, defined by

$$\begin{aligned} g(X)_k - \mu(X)_k &= \log_{10}(S_h(f_k) \cdot \delta_k^h) - \log_{10} S_h(f_k) \\ &= \log_{10} \delta_k^h \end{aligned} \tag{4.27}$$

The number of knots is fixed in each inference, and we choose the associated frequencies f_k equally log-spaced within the data frequency range. Overall,the model is summarized by functions from the parametric family

$$\log_{10} S_h(f; \log_{10} A, \gamma, \boldsymbol{\delta}^h) = \log_{10} A + \gamma \log_{10}(f/f_\star) + k(f, f_j) k(f_j, f_k)^{-1} \log_{10} \delta_k^h, \tag{4.28}$$

Noise

We follow a similar approach to parameterize and infer upon the noise uncertainties. Again, we assume a well known TDI transfer matrix M_{TDI} in Eq. 3.55, so the modelling freedom is left for the single-link noise PSD. Specifically, we introduce a set of parameter $\log_{10} \boldsymbol{\delta}^n$.

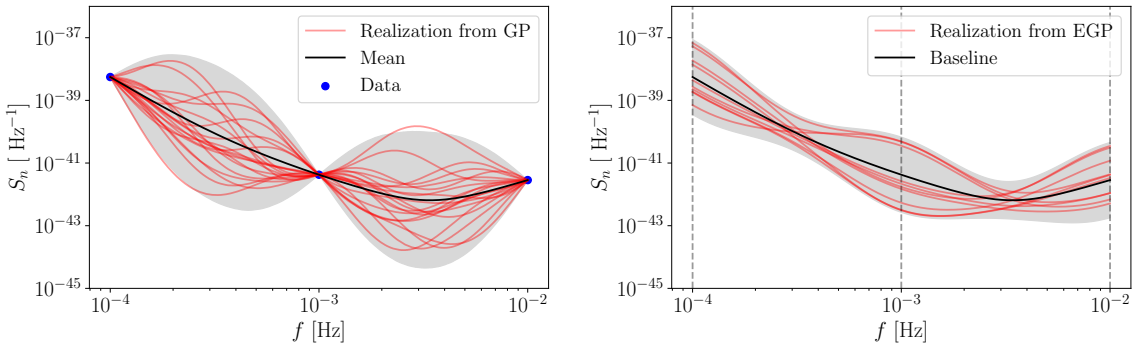

Figure 4.7: Noise realizations from GP and EGPs model. (*Left panel*) Realizations from GP. The gray shaded region represents the 90% credible interval. (*Right panel*) Realizations from EGP. The gray area represents the prior 90% credible interval. While GP allows for straightforward inclusion of observed data, EGPs smoothly explores a large set of spectral shapes, with the *data* (i.e. knots amplitudes at chosen locations shown as grey dashed vertical lines) effectively proposed as inference parameters.

Unlike the signal case, we fix the baseline to the nominal PSD level, using the reference amplitude values for the TM and OMS noise components given in Eq. 3.33 and Eq. 3.34:

$$S_n^0(f) = 2S_{\rm TM}(f) + S_{\rm OMS}(f). \tag{4.29}$$

This choice is consistent with the assumption of identical and independent link noises. Thus, the final noise model reads:

$$\log_{10} S_n(f;\boldsymbol{\delta}^n) = \log_{10} S_n^0(f) \; + k(f,f_j)k(f_j,f_k)^{-1} \log_{10} \delta_k^n, \tag{4.30}$$

With this parameterization, we augment the noise reference model with considerable flexibility to vary across a large functional space. In each inference, the injected noise in the data is generated according to the fixed baseline, so we anticipate recovering zero values for $\log_{10} \boldsymbol{\delta}^n$. In Fig. 4.7, we show a comparison between GP- and EGP-like instrumental noise modelling, illustrating the different degree of flexibility each one exhibits.

In the next section, we present parameter estimation results performed using the model introduced. First, we use a simple toy model to control the correct recovery of a SGWB with known spectral shape (Section 4.3.2). Then, we apply our formalism to the parameter estimation of SGWB from a population of unresolved EMRIs, as generated in Section 4.1. In this study, we model the data using the Whittle likelihood, as introduced in Eq. 2.53. We choose uniform priors for each parameter in $\boldsymbol{\theta}$, and the following prior ranges:

- Power-law amplitude $\log_{10} A$: $[-70, -35]$
- Power-law slope γ: $[-5, 5]$
- Signal knots amplitudes $\log_{10} \boldsymbol{\delta}^h$: $[-2, 2]$

- Noise knots amplitudes $\log_{10} \boldsymbol{\delta}^n$: $[-0.5, 0.5]$

We highlight once again that the number of knots $\boldsymbol{\delta}^h$ and $\boldsymbol{\delta}^n$ can be further optimized, as much as the remaining hyperparameters (i.e. the kernel length scale, its functional form, and the knots locations). This represents a potential high-degree of flexibility to our algorithm, which can be leveraged to find unexpected spectral features upon inference. In this study, we focus on a finite set of kernel lengths σ and number of knots. Noise parameter estimation is instead fixed to three knots, given the baseline choice matching the noise model used for injection in the data.

In the absence of a trans-dimensional sampling framework [93] in our codebase, our study is constrained to perform inference for each choice of hyperparameters, and select the most suitable via Bayesian model selection. In order to compare the different choices, we use the log-Bayes factor

$$\log_{10} \mathcal{B}^{\sigma,n}_{\sigma',n'} = \log_{10} \mathcal{Z}_{\sigma,n} - \log_{10} \mathcal{Z}_{\sigma',n'}, \tag{4.31}$$

where (σ, n) and (σ', n') are labels identifying competing models with different kernel length scale and number of knots, respectively. The evidences $\mathcal{Z}$ and the posterior distributions are computed using `nessai`.

4.3.2 Application to toy Model

We conduct a test retrieving EGPs model with a power-law baseline and non-zero knot amplitudes. Although the injected signals lack an astrophysical interpretation, they serve as a compelling test for our method capability of recovering complex spectral shapes. We inject separately two SGWBs, with SNR of 209 and 5, respectively. We choose the following parameters for each signal:

- Power-law baseline with $\log_{10} A = -43, \gamma = -2$ (SNR=209) and $\log_{10} A = -44.5, \gamma = -1.5$ (SNR=5);
- Knots amplitudes $\log_{10} \boldsymbol{\delta}^h = [0.6, -0.2, -0.3, 0.4]$, identical in both cases.

We choose the parameters as representative of two SGWBs, above and below the nominal PLS curve at SNR= 10 after 4 years of observation. We emphasize once more that the injected model is an EGPs with local features arising with respect to a simple power-law. In this specific context, the PLS is not a faithful indicator of a signal detectability, but should be interpreted as a rough reference level.

Spectral inference results and parameter posteriors are shown in Figs. (4.8) and (B.1) in Appendix B, respectively. The EGPs model is able to capture the spectral shape of the high SNR injected signal, and posterior distributions are consistent with the injected values within their 90% credible intervals. The posterior distributions for noise parameters are consistent with zero, as expected. If the signal SNR is too low, as for the second simulated SGWB (SNR=5) the posterior distribution for its parameters is effectively an upper limit, only.

4.3.3 Application to astrophysical case

We further apply our method to the astrophysical case of a SGWB from EMRIs. We inject a background signal chosen from a set available, discussed earlier in Section 4.1. Resulting

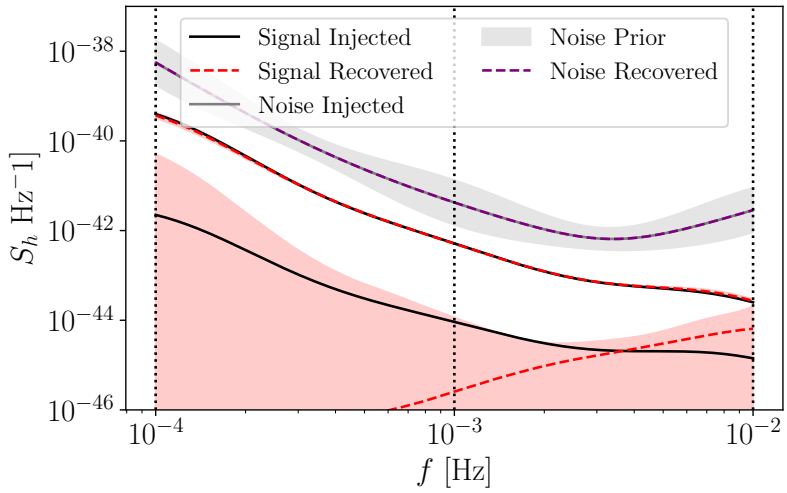

Figure 4.8: Spectral reconstruction of two toy models SGWBs. Injected SGWBs are denoted by black solid lines. They have SNR of 209 and 5, respectively. This places them above and below the nominal PLS curve (solid blue line) to a SGWB with SNR of 10 after 4 years of observation. Noise injection and prior 90% confidence intervals are shown as gray solid and purple dashed lines, respectively. Posterior median and 90% confidence intervals on each signal –analysed independently– are denoted by red dashed lines and light-red shared areas. The EGPs flexibility captures features in the PSD shape.

SGWBs exhibit a wide range of amplitudes, though the majority of them provides a SGWB potentially detectable with LISA over 4 years of observation time.

In this study, we use M1 as our fiducial model, as it provides an intermediate SGWB amplitude across the populations studied. Because of uncertainties on spectral morphology and population distribution, our flexible inference method is suitable to analyse such signal. To coherently inject the signal into `balrog`, we first divide the smoothed A channel signal realization obtained in Section 4.1 by the response function. The resulting spectrum is then smoothed to define a reference model PSD. Finally, we re-generate the data for each TDI channel and use it for inference.

Exploring hyperparameters space

We perform a set of Bayesian inferences, exploring two hyperparameters for the signal model: the number of knots, $n = 3, 4, \ldots, 8$, and the kernel length-scale, as integer m_σ multiples of $\sigma = 0.6$, with $m_\sigma = \{1, 2, 4, 8, 16, 32\}$.

We evaluate each evidence through stochastic NS. Results are presented in Fig. 4.9 (top panel), as Bayes factors with respect to the highest-evidence model, i.e. $(n, m_\sigma) = (8, 32)$. They reveal a prevailing trend: as m_σ increases, the number of knots become less effective at influencing each model evidence, resulting in Bayes factors increasingly close to each as function of m_σ.

For completeness, in Fig. 4.9 (bottom panel) we also provide the computational cost of each inference, measured in CPU core kilohours (kh). The observed trend is the result of two competing factors: as the number of parameters (i.e., the number of knots n) increases, the inference process would take longer. However, if the increased model complexity (e.g., shorter kernel length-scale) is not required by the simulated data, exploration of parameter space is typically fast. As a result, we observe a pattern in computational cost broadly similar to the log Bayes factors: as the number of knots (m_σ) increases (decreases), the computational cost diminishes.

As a further check, we perform inference for the null-hypothesis $\mathcal{H}_0$, where we exclusively model the data as instrumental noise. Results reveal decisive evidence, in favour of any signal-including model (n, m_σ) with respect to the noise-only model $\mathcal{H}_0$, with log-Bayes factors largely greater than 10^3. As expected, the spectral noise reconstruction in Fig. 4.10 shows significant biases to compensate for the unmodelled SGWB signal in the data.

Spectral reconstruction

We discuss here in greater detail the signal spectral reconstruction of each model. Our findings are best described by a comparison between the inferences with (n, m_σ) equal to $(8, 1)$ and $(8, 32)$, shown in the left and right panels of Fig. 4.11, respectively. The reconstruction of fiducial EMRI background is displayed both with and without the inclusion of the T channel data, in the top and bottom panels of Fig. 4.11, respectively. We observe however that the spectral reconstructions with and without the T channel are broadly consistent, with broader posteriors when the T channel is excluded, since its inclusion contributes to a more precise characterization of the instrumental noise. This is due to the common noise modelling of the single link PSD spectrum, which is then propagated coherently to each channel through $M_{\rm TDI}$ Eq. 3.55. Relaxing the exact knowledge of the transfer matrix is expected to significantly broaden the noise PSD posterior.

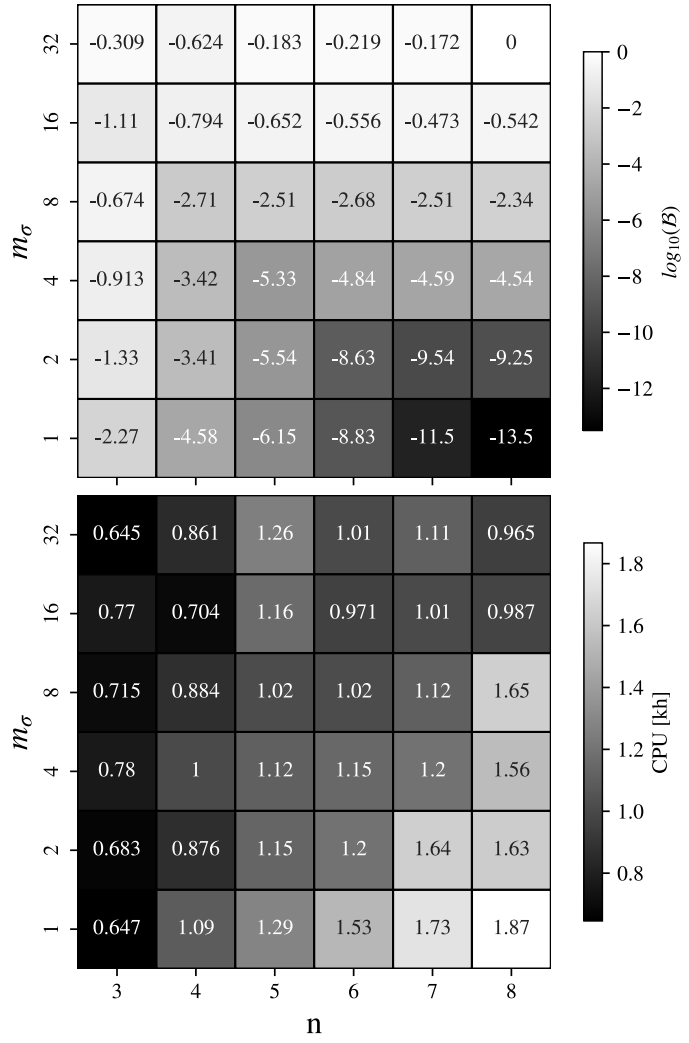

Figure 4.9: (*Top panel*) Grid of logarithmic Bayes factor for different pairs of number of signal knots (n) and integers of Kernel length scale σ. The numbers are computed in relation to the model that has the highest evidence, specifically $(n, m_\sigma) = (8, 32)$. We observe a diagonal pattern: as m_σ decreases and n increases, the model becomes progressively less preferred. (*Bottom panel*) Computational Cost (i.e. Sampling time as reported by `nessai`) of parameter estimations as function of integer multiple of σ, m_σ (with $\sigma = 0.6$) and the number of knots of the signal model, n. The numbers are in unit of kilohours Each inference is run in multi-threading over 40 cores.

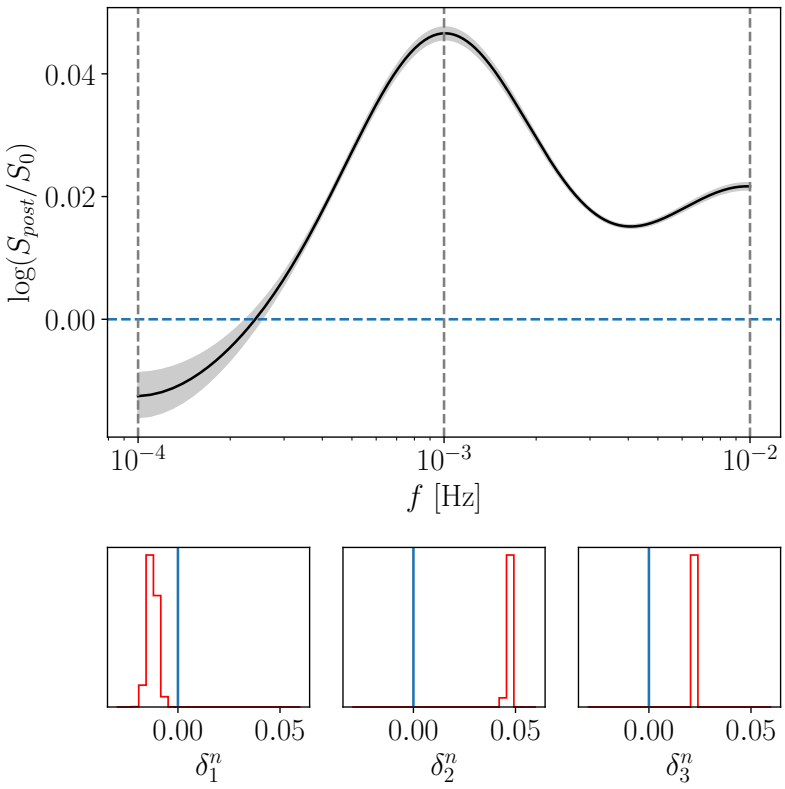

Figure 4.10: Posterior distribution for the inference on EMRIs background, with a noise-only model. The noise estimation employs 3 knots, with frequencies corresponding to vertical dashed gray lines in the main panel. Bottom sub-panels show posteriors on individual knot amplitudes. The injected spectral noise model corresponds to $\boldsymbol{\delta}^n = \mathbf{0}$, and is denoted by solid blue lines. The noise model can compensate with significant biases the presence of a SGWB in the data. Nonetheless, its evidence is significantly lower when compared to the noise and signal models, as illustrated in Fig. 4.9.

A comparison between left and right panels reveals an important result, originating from the competition in Bayes factors between excess model complexity and the effectiveness of our chosen parameterization. The inference presented in this study is inherently integrated across a large frequency band: we annotate the injected SGWB with blue values for cumulative SNR with respect to the true noise level on a set of representative frequencies. It is readily apparent that 80% of the SNR is accumulated between 1 and 4mHz: in this frequency range, the reconstruction is consistent with the injected spectrum. The dashed red lines and red shaded area denote the posterior median and 90% confidence interval, respectively.

At the lowest frequencies, the preferred model (top panel, $n, m_\sigma = 8, 32$) exhibits larger biases as compared to the disfavoured one (bottom panel, $n, m_\sigma = 8, 1$). This is due to the narrowband ($m_\sigma = 1$) fluctuations being highly disfavoured by the data in the high SNR region, where the spectrum is close to a power-law model, a shape already described by the baseline parameters. On the contrary, broad-band ($m_\sigma = 32$) fluctuations can be absorbed by an adjustment of the baseline parameters, while keeping an accurate spectral reconstruction. However, the preference for $m_\sigma = 32$ has an important side effect: m_σ is fixed and equal for every knot in each analysis, including the four at the lowest frequencies. Therefore, the model with the highest evidence is not capable of introducing local fluctuations to capture significant tilts with respect to the baseline powerlaw. These are instead provided by the disfavoured model.

Figs. (B.2) and (B.3) in appendix B further support our interpretation, where we compare the posterior distributions for the two runs shown in Fig. 4.11 (red contours) with two more differing only by a broader prior for the knot amplitudes $\boldsymbol{\delta}^h$, uniform in $[-5, 5]$ (blue contours). The inference results for the noise are largely unaffected and uncorrelated to the signal parameters (and hyperparameters). However, we find that relaxing the knot priors affects the posterior on the baseline parameters: a sign of strong correlation. Moreover, posteriors on $\boldsymbol{\delta}^h$ for the preferred runs (Fig. B.2) are prior dominated, while the likelihood becomes informative in the disfavoured models $(n, m_\sigma) = (8, 1)$ (Fig. B.3). In conclusion, we argue that in a realistic scenario models more flexible than EGPs should be developed, alongside astrophysically motivated priors to avoid introducing –even in an evidence based model selection– unwanted biases.

T channel influence

Given its reduced sensitivity to the SGWB, we further investigate the impact of the T channel on parameter estimation. We anticipate that excluding the T channel from data analysis will result in less informative posterior distributions (as depicted in bottom panels of Fig. 4.11). To be more specific, we seek to assess whether hyperparameterization modeling remains influential, as depicted in Fig. 4.9.

To achieve this, we examine two models characterized by the highest and lowest values of evidence (or hyperlikelihood), denoted as $(n, m_\sigma) = (8, 32)$ and $(n, m_\sigma) = (8, 1)$, respectively. The calculated logarithmic Bayes factor stands at -10.95, underscoring a substantial strength of evidence in favor of one model over the other. In other words, the outcome suggests that the trend observed in Fig. 4.9 is robust and not significantly affected by variations in our ability to estimate the noise level.

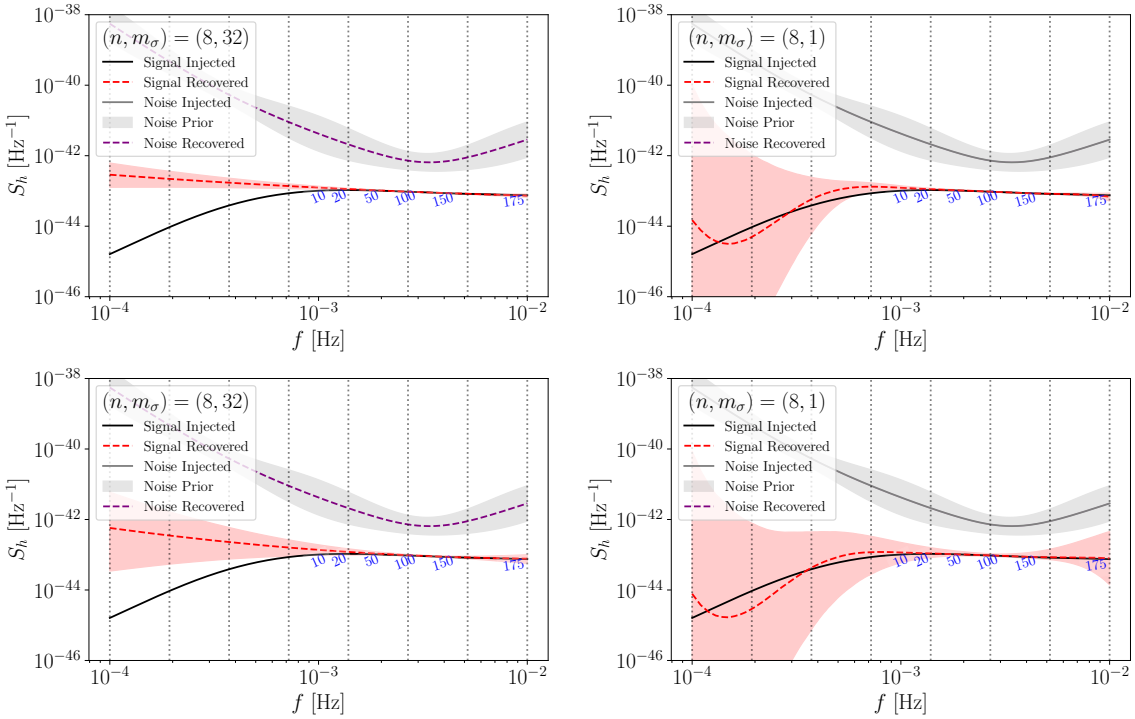

Figure 4.11: Signal reconstruction of the M1 EMRI background carried out using the EGPs parametrization. In both figures, the noise estimation employs 3 knots. (*Left panel*) The model features 8 knots for the signal and a kernel length scale that is 32 times larger than σ. This configuration aligns with the model exhibiting the highest evidence. (*Right panel*) the model features 8 knots for the signal, and the kernel length scale is set equal to σ. This configuration corresponds to the model with the lowest evidence. Blue numbers denote the cumulative SNR integrated from 10^{-4}Hz up to the frequency they are located at.

4.3.4 Caveats and takeaways

We introduced a weakly-parametric method for SGWB inference using EGP, enabling reconstruction of spectra with uncertain shape.

Main results. The EGP framework successfully captured complex spectral morphologies, as demonstrated through toy models and application to the fiducial EMRI background (M1). For high-SNR signals, the method accurately recovered both baseline power-law parameters and fine-scale spectral structure. Bayesian model selection effectively distinguished between competing hyperparameter choices (kernel length-scale σ and number of knots n), though computational costs scaled with model complexity.

Methodological caveats. While the EGP formalism provided considerable flexibility, it introduced non-locality: inference in high-SNR frequency regions influenced spectral reconstruction in low-SNR regimes through correlated baseline parameters. Specifically, models favored by the data in the sensitive band (1–4 mHz) exhibited systematic biases at lower frequencies where SNR accumulation was minimal. This behavior arose because the kernel length-scale σ was held constant across all frequencies, preventing localized adjustments. Evidence-based model selection, while principled, therefore favored reduced complexity in high-SNR regions at the expense of accuracy elsewhere.

Future developments. Realistic LISA analyses should incorporate astrophysically motivated priors on spectral parameters to mitigate biases from excessive model flexibility. Trans-dimensional sampling frameworks [93] would enable simultaneous optimization of hyperparameters within a single inference, avoiding the computational burden of separate runs for each hyperparameter choice. The EGP method versatility extends beyond SGWB characterization; Ref. [141] successfully applied EGPs to tests of general relativity, demonstrating its broad utility for LISA science.

4.4 Bayesian sensitivity

The uncertainties showed in Sections 4.1 and 4.2 pose fundamental challenges for assessing SGWB detectability. The traditional PLS curve introduced in Section 2.1.3 provides a useful heuristic based on SNR thresholds. However, as shown in Section 2.1.2, the SNR definition corresponds to a statistic assuming perfect knowledge of both signal and noise spectral shapes and noise level, neglecting their respective prior uncertainties. Several studies in the literature have focused on relaxing such assumptions [100, 32, 138], and the flexible parametric analysis in Section 4.3 demonstrates how model uncertainties can affect spectral reconstruction. In this section, we propose a novel formalism that systematically addresses these limitations by employing the Bayes factor as our fundamental detectability metric, as introduced in Section 2.2.3. By leveraging conjugate priors on the covariance matrix describing the total data, we derive semi-analytic expressions that naturally marginalize over uncertainties in both signal and noise properties. This allows us to rapidly evaluate model evidences and construct Bayesian Power-Law Sensitivity (BPLS) curves following the standard PLS methodology of Ref. [184]. We apply this framework to LISA under various assumptions about noise uncertainty, providing a robust approach to quantify detection prospects for the SGWB in the presence of the modeling uncertainties identified throughout this chapter. This methodology offers a principled way to assess detectability when the astrophysical uncertainties documented in Sections 4.1 and 4.2 and the spectral flexibility explored in Section 4.3 must be systematically incorporated into the detection framework.

4.4.1 Conjugate priors

Here, we adopt the same assumptions outlined in Section 2.2.1, working with the time-domain datastreams from p detectors. We segment the data to construct coarse-grained representations and define the averaged periodograms $Y(f)$ as in Eq. 2.56. These are modeled using the complex Wishart likelihood given in Eq. 2.57, with n_c degrees of freedom and a scale matrix $\Gamma(f)$, as defined in Eq. 2.58. This approach significantly reduces the computational cost of the analysis, but at the expense of limiting the frequency resolution, particularly at low frequencies.

To evaluate model evidences, priors must be specified and marginalized over as follows,

$$\mathcal{Z}(Y) = \int d\Gamma \mathcal{L}(Y \mid \Gamma)\pi(\Gamma), \tag{4.32}$$

where we drop the frequency dependence of the matrices for brevity. We choose priors conjugate to the likelihood in Eq. 2.57 for two main reasons: first, upon inference the posterior preserves the same functional form of the prior; second, as we shall see below, the free parameters available yield enough flexibility to choose a reference expectation value and variance for Γ. We assume a complex inverse-Wishart prior $\mathcal{CW}^{-1}$ on Γ, with ν degrees of freedom and scale matrix Ψ.

$$\Gamma \mid \Psi, \nu \sim \mathcal{CW}^{-1}(\Psi, \nu) \tag{4.33}$$

$$\pi(\Gamma|\Psi, \nu) \propto |\Psi|^{\nu}|\Gamma|^{-(\nu+p)} \exp\left[-\mathrm{tr}(\Psi\Gamma^{-1})\right]. \tag{4.34}$$

The chosen prior yields an expectation value for Γ

$$E[\Gamma] = \frac{\Psi}{(\nu - p)} \tag{4.35}$$

for $\nu > p+1$. Therefore, we set the scale-matrix parameter of the prior to

$$\Psi = \frac{\nu - p}{n_c}\left(\Sigma_n^0 + \Sigma_{\rm GW}(\vec{\lambda})\right), \tag{4.36}$$

where Σ_n^0 denotes a fiducial estimate of the noise covariance matrix, and $\vec{\lambda}$ are the parameters describing the SGWB power spectral density (e.g., the amplitude and slope for a power-law signal). As a consequence, the mean of the prior distribution is the expected Γ in Eq. 4.35. The value of ν can be adjusted to represent the scale of our prior uncertainty on Γ, since the variances $\mathrm{Var}(\Gamma_{ii})$ are proportional to $\Psi_{ii}^2/\left[(\nu-p)^2(\nu-p-1)\right]$. For instance, $\nu \sim 10$ represents almost 10% uncertainty for $p = 3$.

In realistic data-analysis scenarios, the parameter $\vec{\lambda}$ must be inferred, hence a prior on it must be placed. On the contrary, following the PLS construction, here we will explore it systematically. The posterior distribution for Γ follows a complex inverse-Wishart distribution with parameters updated according to the observations, which are described by the amount of data observed, n_c, and the average periodogram, Y, from Eq. 2.56:

$$\Gamma \mid Y, \Psi, \nu \sim \mathcal{CW}^{-1}(\Psi + Y, \nu + n_c), \tag{4.37}$$

$$p(\Gamma|Y,\Psi,\nu) \propto \frac{|\Psi|^\nu |Y|^{n_c-p}}{|\Gamma|^{n_c+\nu+p}} \exp\left[-\mathrm{tr}((Y+\Psi)\Gamma^{-1})\right] \tag{4.38}$$

Upon marginalization, the evidence in Eq. 4.32 reads

$$\mathcal{Z}(Y|\Psi,\nu) = C\frac{|\Psi|^\nu |Y|^{n_c-p}}{|\Psi+Y|^{(\nu+n_c)}}, \tag{4.39}$$

where C is a suitable normalization constant, independent on Ψ, ν. Notably, the evidence retains a dependence on the prior hyperparameters Ψ and ν. Reinstating explicitly the dependence on the frequency, the Bayes factor for a given bin—between the presence (H_1) and absence (H_0) of GW signal hypotheses— can be written as:

$$\begin{aligned}\log \mathcal{B}(f|\nu, n_c, \vec{\lambda}, Y) = (n_c+\nu)\log\left|\frac{\Psi(f|H_0)+Y(f)}{\Psi(f|H_1)+Y(f)}\right| + \\ + \nu \log\left|\frac{\Psi(f|H_1)}{\Psi(f|H_0)}\right|,\end{aligned} \tag{4.40}$$

where the scale matrices under the two hypotheses read:

$$\Psi(f|H_1) = \frac{\Sigma_n^0(f) + \Sigma_{\rm GW}(f,\vec{\lambda})}{n_c}(\nu - p), \tag{4.41}$$

$$\Psi(f|H_0) = \frac{\Sigma_n^0(f)}{n_c}(\nu - p). \tag{4.42}$$

Under the assumption that the signal and noise are stationary, the overall Bayes factor is obtained by taking the product over all frequencies. Eq. 4.40 is the central result of this work, as it allows for quick evaluation of Bayes factors assuming *a priori* uncertain knowledge of the noise and signal. From Eq. 4.40, we now want to identify the values of $\vec{\lambda}$ yielding a specified Bayes factor. Typically, we are interested in the contour corresponding to $\log_{10}(\mathcal{B}) \geq 1$ as fiducial threshold of strong evidence in favor of H_1.

Since Eq. 5.26 depends on the specific observed realization of Y, we compute the expectation value of $\log_{10}(\mathcal{B})$ over multiple realisations. Two viable approaches are possible here: numerical and analytical approximation. For the former, given each $\vec{\lambda}$, we draw realizations of Y from the Wishart distribution and compute the average of the logarithm in Eq. 5.26. For the latter, we first rewrite Eq. 4.40 as

$$\begin{aligned}\ln \mathcal{B} = \nu \mathrm{tr}(\ln \Psi_1 - \ln \Psi_0) + \\ + (n_c + \nu)\mathrm{tr}(\ln(\mathbb{I} + Y^{-1}\Psi_0) + \\ - \ln(\mathbb{I} + Y^{-1}\Psi_1)).\end{aligned} \tag{4.43}$$

using the identity $\ln |M| = \mathrm{tr}(\ln M)$. We now use the expansion

$$\ln\left(\mathbb{I} + Y^{-1}\Psi\right) = \left(\sum_{k=1}^{\infty}(-1)^{k+1}(Y^{-1}\Psi)^k/k\right), \tag{4.44}$$

which holds under the assumption that $Y^{-1}\Psi \ll 1$. Since $\Psi \propto (\nu - p)/n_c$, this condition corresponds to the limit where $\nu - p$ is small and n_c is large. Practically, this represents a scenario characterized by significant uncertainty in the process and a long observation period. At leading order, the log Bayes factor reads

$$\begin{aligned}\ln \mathcal{B} = \nu \mathrm{tr}(\ln \Psi_1 - \log \Psi_0) + \\ + (\nu + n_c)\mathrm{tr}(Y^{-1}\Psi_1 - Y^{-1}\Psi_0) + \mathcal{O}(Y^{-1}\Psi),\end{aligned} \tag{4.45}$$

and taking its expectation value

$$\begin{aligned}E[\ln \mathcal{B}] &= \nu \mathrm{tr}(\ln \Psi_1 - \ln \Psi_0) + \\ &+ \frac{(\nu + n_c)}{n_c - p}\mathrm{tr}(\Gamma^{-1}\Psi_1 - \Gamma^{-1}\Psi_0) + \mathcal{O}(Y^{-1}\Psi) \qquad (4.46) \\ &\approx \nu \mathrm{tr}\left[\ln(\Sigma_n^0 + \Sigma_{\mathrm{GW}}) - \ln \Sigma_n^0\right] + \\ &+ \frac{(\nu + n_c)(\nu - p)}{n_c - p}\mathrm{tr}\left[(\Sigma_n^0 + \Sigma_{\mathrm{GW}})^{-1}\Sigma_n^0 - \mathbb{I}\right] \qquad (4.47)\end{aligned}$$

where we used the relation $E[\mathrm{Y}^{-1}] = \Gamma^{-1}/(n_c - p)$ [162]. Even though Eq. 4.45 can be expanded by including higher order moments from the complex inverse-Wishart distribution [127], we provide here results at leading-order only. As expected, in the case of $\Sigma_n^0 \gg \Sigma_{\mathrm{GW}}$, the expectation value of the log-Bayes factor goes to 0. Instead for $\Sigma_n^0 \ll \Sigma_{\mathrm{GW}}$, the expectation value is

$$\nu \mathrm{tr}\left[\ln(\Sigma_{\mathrm{GW}}(\Sigma_n^0)^{-1})\right] - \frac{(\nu + n_c)(\nu - p)}{n_c - p}\mathrm{tr}[\mathbb{I}]\ . \tag{4.48}$$

4.4.2 Application to LISA

We show here a concrete example by considering the SGWB detection problem for LISA. Specifically, we consider a four-year mission duration. To reduce the computational cost, we restrict our analysis to the frequency range $[10^{-4}, 10^{-2}]$ Hz. We assume three datastreams, i.e., $p = 3$, each corresponding to a TDI variable. Specifically, we assume equal and constant armlengths, using the AET variables. In doing so, the model for Γ simplifies to a

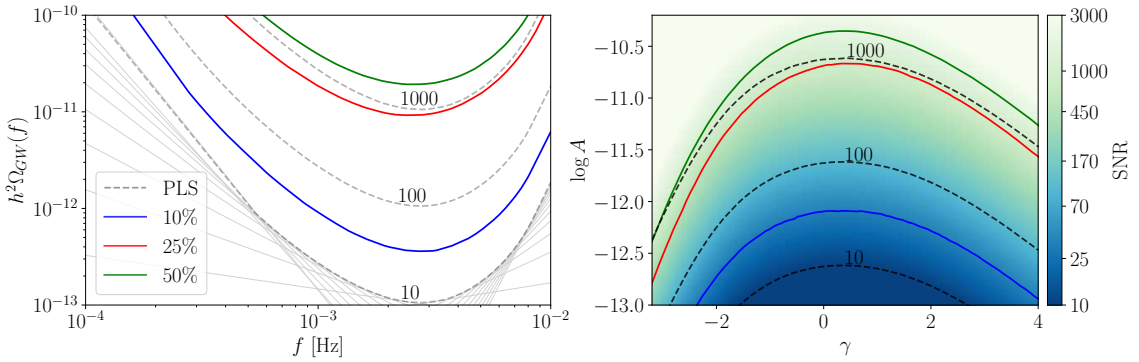

Figure 4.12: SGWB PLS with Bayes factor threshold of 10, for three reference noise uncertainties $(10\%, 25\%, 50\%)$. (*Left panel*) Bayesian power-law sensitivity corresponding to a log-Bayes factor threshold of one. The colors represent different numbers of prior degrees of freedom, ν, reflecting varying levels of uncertainty. The dashed curves represent the power-law sensitivity for different SNR thresholds (10, 100, 1000). For reference, we show as solid light grey lines the power-law spectra used to construct the PLS at SNR 10. (*Right panel*) The solid curves denote contours corresponding to $E\left[\log_{10}(\mathcal{B})\right] = 1$ in the parameter space of the amplitude and slope of a power-law spectrum. The dashed lines represent contours of fixed SNR within the same parameter space.

diagonal matrix. Under these assumptions, the SNR of a signal with parameters $\vec{\lambda}$ is given by Eq. 2.39, combined with SGWB response and PSD noises as described in Sections 3.2.3, 3.3, and 3.4.2. It is important to note that Eq. 2.23 is valid in the regime where the noise dominates over the signal. Otherwise, more robust statistics need to be considered, as in [160].

We consider power-law shaped spectra, characterized by their amplitude $A = \Omega_{\rm GW}(f = 10^{-2.5}\text{Hz})$ and slope γ, as in Eq. 2.41 We evaluate the expectation value of the Bayes factor as a function of (A, γ), for three reference values of ν: $\nu_{10} = 10$, $\nu_{25} = 8$, and $\nu_{50} = 7$, corresponding to 10%, 25% and 50% prior uncertainties on the covariance in Eq. 2.58.

For each ν, we construct the contour for $\vec{\lambda}$ corresponding to the level $E\left[\log_{10}(\mathcal{B})\right] = 1$. This is shown in the right panel of Fig. 4.12. The contour $(A(s), \gamma(s))$ describes a collection of power-law signals detectable with a fixed, Bayesian significance. By identifying at each frequency the one yielding the highest $\Omega_{\rm GW}(f)$, we define the envelope $(A_\nu(f), \gamma_\nu(f))$

$$\Omega_{\rm GW}^{\rm BPLS}(f) = A_\nu(f) \left(\frac{f}{10^{-2.5}\text{Hz}}\right)^{\gamma_\nu(f)} \tag{4.49}$$

$$:= \max_s \left[A(s) \left(\frac{f}{10^{-2.5}\text{Hz}}\right)^{\gamma(s)}\right]. \tag{4.50}$$

This is the BPLS. We show the resulting curve in the left panel of Fig. 4.12. For a direct comparison, we over-plot the standard SNR-based PLS, following Ref. [184].

An unambiguous mapping between PLS and BPLS is possible only in the regime where

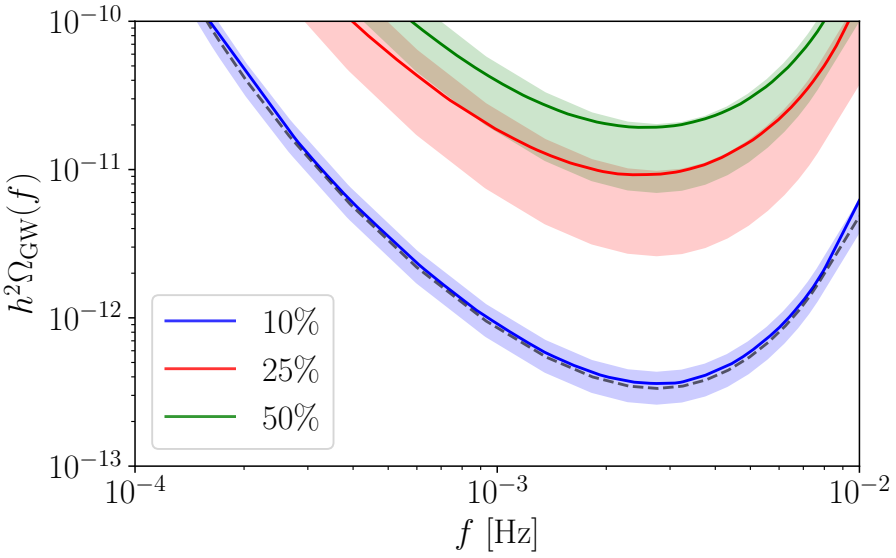

Figure 4.13: Mapping between BPLS and standard PLS. The solid curves represent the same BPLS curves shown in the left panel of Fig. 4.12. The shaded region encompasses the PLS curves intersecting the BPLS at least once. Therefore, SNRs associated with the PLS curves in each band correspond to SGWB signals detectable at $\log_{10}\mathcal{B} = 1$ or higher. For reference, the dashed curve denotes a PLS closely matching the BPLS associated with ν_{10}, and corresponds to an SNR = 35 over 4yr of observation. Each color corresponds to a different level of a priori uncertainties, as described by the parameter ν.

the posterior distribution is dominated by the likelihood, such that $\log_{10}(\mathcal{B}) \propto \mathrm{SNR}^2$[70], as also shown in Section 2.2.3. In general, such a comparison is non-trivial. Notably, we find that the relationship between the two quantities becomes increasingly degenerate as the prior uncertainty in the covariance matrix grows. We illustrate this effect in Fig. 4.13. Therein, we identify the set of PLSs (shown as shaded areas) intersecting a given BPLS at least once. Each intersection corresponds to a SGWB signal yielding a certain SNR and $\log_{10}\mathcal{B} = 1$. As shown in the figure, to higher prior uncertainties correspond larger PLS bands, hence a larger set of SNRs. SNRs yielding a detection significance $\log_{10}\mathcal{B} = (0.5, 1, 2)$ —substantial, strong, and decisive, respectively— may range in $([20, 30], [25, 40], [37, 60])$ for ν_{10}, in $([248, 917], [253, 920], [263, 924])$ for ν_{25}, and in $([664, 1897], [673, 1900], [690, 1905])$ for ν_{50}.

In the analysis presented so far, we assumed a nominal mission duration of four years [67]. To explore the dependence of the BPLS on the observation time, we vary $T_{\rm obs}$ by adjusting the number of chunks n_c at fixed chunk duration, while keeping the prior uncertainty level fixed at ν_{10}. We focus on the dependency on $T_{\rm obs}$ of the BPLS at a single significance $\log_{10}\mathcal{B} = 1$, which closely matches the PLS with SNR = 35 over the nominal mission duration $T_{\rm obs} = 4$yr. We illustrate our findings in the left panel of Fig. 4.14. Notably, the PLS shows a stronger dependence on time than BPLS. At fixed frequency and SNR, the PLS level is inversely proportional to the square root of the ob-

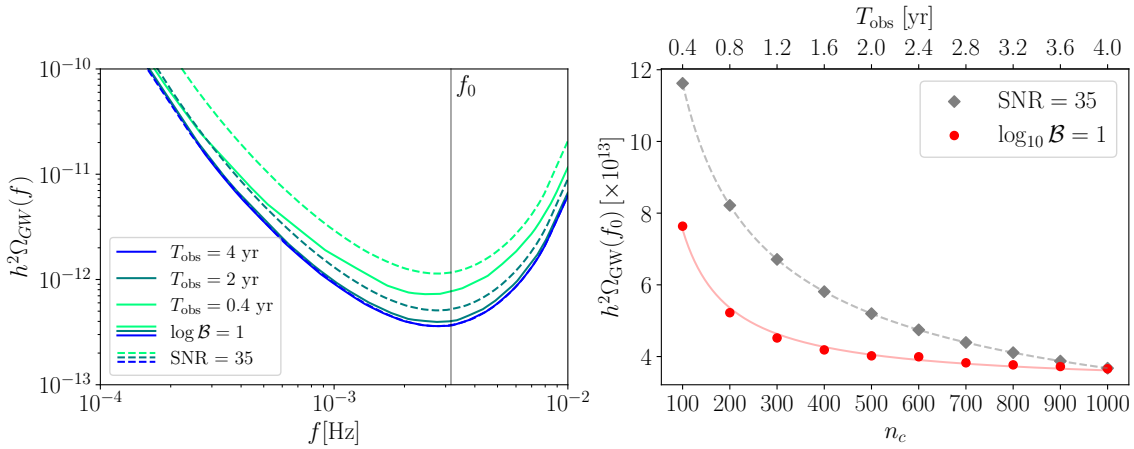

Figure 4.14: Dependency of BPLS on the observing time. (*Left panel*) The dashed lines refer to PLSs constructed for a target SNR $= 35$, while solid ones denote the BPLSs at $\log_{10} \mathcal{B} = 1$ considering 10% prior uncertainties in the covariance matrix. Colors denote three fiducial observation times, $T_{\rm obs} = 0.4\text{yr}, 2\text{yr}, 4\text{yr}$, as indicated in the legend. (*Right panel*) Dependency of the PLS (grey) and BPLS (red) value at a reference frequency $f_0 = 10^{-2.5}$Hz as a function of the number of chunks, i.e., the observation time. Dashed (solid) curve represent the best fit of 10 PLS (BPLS) values with as expected from Eq. 2.23 and Eq. 4.47, respectively.

servation time, i.e., $\propto 1/\sqrt{T_{\rm obs}}$. Instead, the BPLS depends non-trivially on $T_{\rm obs}$: from the expansion in Eq. 4.47 and considering $n_c = T_{\rm obs}/T$, we observe that at leading order $E[\ln \mathcal{B}] \propto (\nu + n_c)/(n_c - p) = (\nu T + T_{\rm obs})/(T_{\rm obs} - pT)$. A fit of this form agrees well with the numerical data (see right panel of Fig. 4.14). The BPLS does in fact show the expected $1/\sqrt{T_{\rm obs}}$ behavior, but only in the regime $n_c \ll \nu$. This can be derived by considering the $\nu \to \infty$ limit of Eq. 5.26. The physical interpretation is that we are able to measure the total stochastic component in the data with a certain precision, which increases as the number of segments increases. For small observation times, the uncertainty in the measured background is large compared to the prior uncertainty, and so the latter is irrelevant: the BPLS behaves like the PLS in this regime. As the number of segments increases, the precision on the total stochastic component of the data improves, and there is a transition into a regime in which the prior uncertainty dominates. From this point onward, the background constraint can no longer improve: at leading order, the background level observable at a fixed $\mathcal{B}$ exhibits a saturation effect.

4.4.3 Caveats and takeaways

We developed a Bayesian framework for assessing SGWB detectability systematically accounting for uncertainties in both signal and noise properties, generalizing the standard PLS concept.

Main results. By employing conjugate priors on the model covariance matrix, we obtained semi-analytic expressions for the Bayes factor comparing models with and without a signal, enabling efficient construction of BPLS curves. When applied to LISA with a 4-year observation time, we found that although the mapping between Bayes factor and SNR is not one-to-one, SGWB signals with a 10% prior uncertainty on the noise level typically require SNR $\sim$ 25–40 to reach $\log_{10} \mathcal{B} = 1$, with the closest PLS occurring at SNR $\simeq 35$.

Extensions and limitations. Our idealized LISA analysis assumed uncorrelated TDI channels and power-law signal morphologies. Realistic scenarios involving correlated datastreams would increase model complexity but were naturally accommodated by the Wishart likelihood framework. The BPLS concept generalized straightforwardly to arbitrary spectral shapes by constructing $(n-1)$-dimensional surfaces at fixed Bayes factor in the parameter space, then enveloping these surfaces at each frequency. This versatile methodology applies to any GW observatory targeting stochastic backgrounds, providing a principled alternative to SNR-based detectability assessments when model uncertainties are non-negligible.

Chapter 5

Uncovering statistical signatures in the astrophysical SGWB

In this chapter, we present results published in [147, 150, 149]. In the previous chapter, we focused on the analysis of stationary, isotropic, and Gaussian signals. We now turn to cases where these assumptions break down for certain astrophysical SGWB components. Rather than treating these deviations as complications, we aim to exploit them to enhance the extraction and characterization of the signal.

In Section 5.1, we focus on quantifying the non-Gaussian and non-stationary features of EMRIs background. In Section 5.2, we explore the concept of cyclostationarity and its application to modeling the MW foreground and background from its satellite galaxies. Finally, in Section 5.3, we present an alternative methodology based on time-frequency analysis, implemented within the `bahamas` pipeline, specifically designed to accommodate the non-stationarity of the Galactic foreground in the context of a global fit analysis for LISA.

5.1 Non-Gaussianity

The characterization of the EMRI background presented in Section 4.1 reveals significant uncertainties spanning approximately three orders of magnitude in the predicted background amplitude. Under the assumptions of stationarity, isotropy, and Gaussianity, the flexible parametric reconstruction in Section 4.3 demonstrates that spectral features alone may be insufficient to extract the SGWB signal when astrophysical uncertainties are significant. To improve signal reconstruction and enhance our ability to disentangle overlapping SGWB components, we aim to characterize additional statistical features beyond the simple power spectral description. Notably, EMRIs are typically highly eccentric systems—with eccentricities ranging from 0.1 to 0.9—and emit GWs across a broad frequency range [143], resulting in a complex spectral structure that may violate the Gaussianity. Moreover, the discrete nature of the source population—particularly in pessimistic formation scenarios—suggests that the central limit theorem may not hold in all frequency bands, leading to detectable deviations from Gaussianity and stationarity. For this reason, in this work,

we investigate thoroughly the contribution of individual emissions to the collective SGWB, and their effect on its stationarity and Gaussianity. These findings motivate the need for extended likelihood models that incorporate non-Gaussian features.

5.1.1 Rayleigh test

We now want to find a statistical test to assess the Gaussianity and stationarity of a SGWB. We take inspiration from the literature available for ground-based detectors [83, 80], where such a test is employed for data-quality diagnostics, e.g., to detect the presence of glitches. The null-hypothesis we test against is that of an ergodic, zero-mean stationary GP. We further assume the signal to be second-order stationary. Thus, given Eq. 2.8, the absolute value of the complex variable $x(f)$ is therefore distributed as follows

$$|x(f)| := \sqrt{\Re x(f)^2 + \Im x(f)^2} \sim \text{Rayleigh}(\sigma(f)) \tag{5.1}$$

Similarly, the squared norm of the Fourier variable is distributed as follows

$$|x(f)|^2 \sim \Gamma(1, 2\sigma(f)) \tag{5.2}$$

where $\Gamma(k, \theta)$ denotes the Gamma distribution and k, θ are the positive shape and scale parameters, respectively. The Gamma distribution has mean $k\theta$ and variance $k\theta^2$, respectively. The test is constructed the following Eq. 5.2 by taking the ratio between the standard deviation and mean of the Gamma distribution. Under the null-hypothesis, the variable is identically one across all frequencies,

$$\rho_{[x]}(f) = \frac{\sqrt{\left\langle \left(|x(f)|^2 - \left\langle |x(f)|^2 \right\rangle \right)^2 \right\rangle}}{\left\langle |x(f)|^2 \right\rangle} = \frac{2\sigma(f)}{2\sigma(f)} = 1 \tag{5.3}$$

where ρ denotes the operator acting on the process x. We will denote the expected value one as, $\rho_{[x],\text{exp}}$. The Rayleigh test is then a statistical test of the stationarity and Gaussianity hypotheses that the *coefficient of variation* (standard deviation over mean) of the signal FT has the value expected for the corresponding Rayleigh random variable. An equivalent statistic is that implemented by LIGO and Virgo where the squared norm, distributed like a Gamma variable [75], is considered. The test is practically carried out by constructing estimators for the random quantities in Eq. 5.3. To obtain multiple samples, we leverage the process ergodicity and suitably chunk the data. The denominator is evaluated through Welch's PSD estimator [196] while the numerator is obtained through FT, and represents a measure of the signal's statistical properties variability. Critical values for the test statistic can be obtained under the null-hypothesis, i.e., for a perfectly Gaussian, stationary signal of the same (finite) duration of our SGWB datastreams. Asymptotically, for a finite number N of samples (which in this context are to be regarded as the chunks), the test is distributed as a $\mathcal{N}(1, 1/2\sqrt{N})$. Ref. [54] presented a simplified example demonstrating how violations of stationarity and Gaussianity are detected using such test.

5.1.2 Updating EMRI catalogs

We build a selection of EMRI population signals following closely Refs. [43, 148]. Further details are presented in Sections 4.1.1 and 4.1.2. We construct a family of populations

leading to three representative models out of the twelve presented in [31]. Our choice is driven by the findings of Section 4.1.3, showing that the EMRI SGWB amplitudes are bracketed by M8 and M12 as extreme scenarios, with M1 providing a fiducial, intermediate case. For each one, we construct a catalog of EMRI systems in a simulated Universe. A list of primary MBH masses M and spins a, along with a redshift z for each event is given. In Section 4.1, the mass μ of the CO was fixed to 10 or $30\mathrm{M}_{\odot}$. Here, we instead sample it from a distribution inspired by the population of compact binary mergers observed by GW ground-based detectors. In reality, it is likely that COs in EMRIs follow a more top-heavy mass distribution compared to that of coalescing compact binaries. This is because EMRIs form in galactic nuclei, where the more massive COs are expected to cluster toward the center due to dynamical processes such as relaxation and mass segregation [85, 98]. However, in the absence of robust observational constraints on the CO mass distribution in galactic nuclei, we just opt here to introduce a reasonable scatter in the underlying mass distribution to explore its role. A more realistic population choice goes beyond the scope of this work and is left for future investigation.

In the third Gravitational Wave Transient Catalog (GWTC-3) observations of the Advanced LIGO [118] and Virgo [194] detectors from the first three observing runs (O1, O2, and O3, respectively) are collected [1]. Among those, 69 confident BBH events are identified based on their significativity to perform population inference [6] with a variety of mass-distribution models. As a reference for our CO mass distribution, we choose the point-wise median of the marginal population posterior on μ PowerLaw+Peak model. We highlight here that our extension remains in excellent agreement with the former choice of fixed-CO mass in Ref. [31], as shown in Fig. 5.1. The median of our chosen CO mass distribution (solid line) is very close to the nominal value of $10M_{\odot}$, and is in good agreement with the variability inferred in Ref. [6] (dashed lines). For reference, we also show the overall population posterior (orange shaded regions) and our chosen reference model (orange solid line).

The remaining parameters are sampled in the same way as described in Section 4.1.2. In particular, we highlight that the sky positions (θ_S, ϕ_S) are drawn uniformly over the celestial sphere, ensuring isotropy.

The EMRI background signals are then computed, following the procedure outlined in Section 4.1.3, as follows:

1. For each source in the catalog (after the initial selection, with a total count of N_{final}), we compute the two gravitational-wave polarizations $h_+(t)$ and $h_\times(t)$ using the 5PN-AAK waveform model.

2. We generate the noise-orthogonal TDI variables AET.

3. We compute the accurate SNR ρ for each individual source and classify sources as unresolvable if $\rho < 20$. This threshold is consistent with values adopted in the literature to define detectability of EMRIs [31, 43, 148, 49].

4. For each TDI channel, we sum the contributions of all unresolvable sources to construct a composite data stream representing the unresolved EMRI background.

Following this procedure, we generate the EMRI backgrounds corresponding to models M1, M8, and M12 in the time domain, yielding three global time series for each catalog. We consider LISA mission duration of $T_{\mathrm{obs}} = 4$ years, resulting in a frequency resolution of

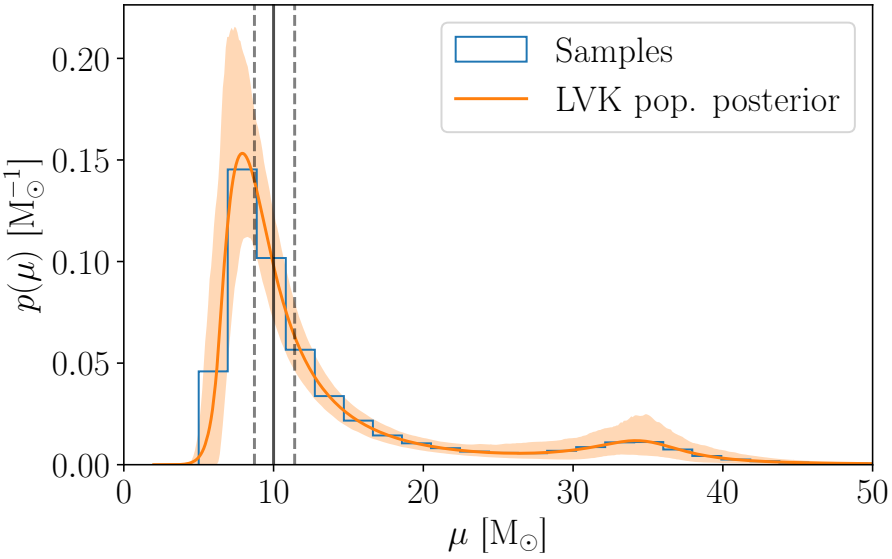

Figure 5.1: Distribution of stellar-mass black hole based on current gravitational wave detections [6]. The orange-shaded area represents the overall posterior distribution, while the solid orange line shows our selected reference model. Solid and dashed black lines denote the posterior on the median, $\mu_{\rm med} = 10^{+1.4}_{-1.3} M_\odot$, which is in large agreement with the fixed $\mu = 10 M_\odot$ model employed in Ref. [6]. In blue are highlighted the samples from the chosen distribution used in the analysis.

$\Delta f = 1/T_{\rm obs} \approx 8 \times 10^{-9}$ Hz. The cadence is set to $\Delta t = 20$ s, corresponding to a maximum non-aliased frequency of $f_{\rm max} = 1/(2\Delta t) = 0.025$ Hz. While a realistic data stream will likely have a higher sampling rate of approximately 0.25 s, we note that for $f > 0.01$ Hz, the LISA noise budget begins to increase, and the EMRI foreground brightness starts to decrease [148]. Thus, our choice of $f_{\rm max} = 0.025$Hz allows us to characterize the statistical properties within a region of primary interest for LISA observations.

5.1.3 Non-gaussianity and background Evaluation

To assess the stationarity and Gaussianity of the EMRI backgrounds, we employ the Rayleigh test introduced in Section 5.1.1. Our catalog construction enforces isotropically distributed sources. Therefore, a GW detector receives an overall signal (and couples to it) whose spectral contributions arise equally from each direction. This assumption is expected to hold for cosmological backgrounds but may be partially or entirely violated for astrophysical ones. Stochastic signals are often modeled as the superposition of infinitely many signals overlapping in time or frequency domain. However, for this limit to be a suitable approximation, the central limit theorem must hold, justifying the use of a Gaussian likelihood model for inference [159]: for each given frequency and sky-location intervals, the number of sources contributing individually must be $\gg 1$. Failure in satisfying this constrain implies different degrees of deviations from Gaussianity and stationarity. Of-

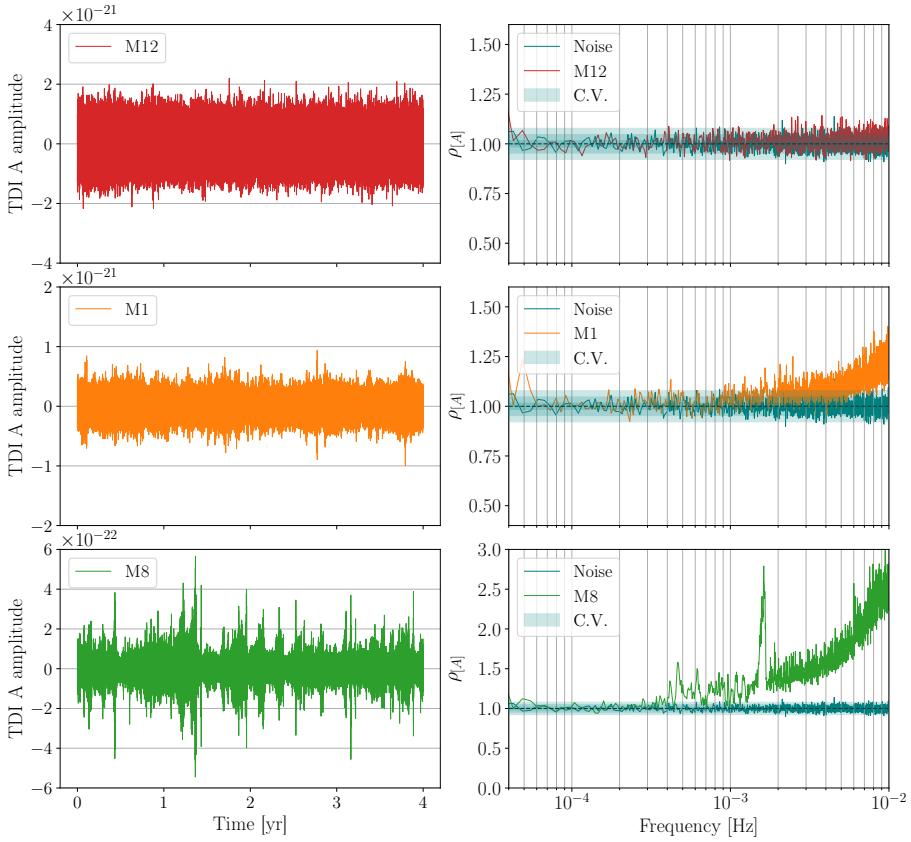

Figure 5.2: Rayleigh test applied to the constructed EMRI foregrounds. (*Left Column*) EMRI SGWB in the time domain A channel for models M12 (top), M1 (middle), and M8 (bottom), to notice the different scales on the y-axis. (*Left panel*) Rayleigh statistics as a function of frequency in logarithmic scale. The test applied to a stationary and Gaussian stochastic time series is shown by the teal region. Within this region, the expected value $\rho_{[A],\mathrm{exp}} = 1$ is shown as a black dashed line, while inner (middle, outer) teal-shaded bands denote the critical values for the null-hypothesis rejection, at 1σ (2σ, 3σ) confidence. As mentioned in Section 5.1.1 these are asymptotic values constructed for a finite number of realizations of a stationary GP.

ten, the resulting signal exhibits a "popcorn-like" behaviour [161]. While the frequency content is non-trivial to characterize due to the broadband nature of each signal, the sky distribution can be readily tested via catalogs inspection. We do so by performing Upon cross-checking our hypothesis, we observe through a simple angular multipole decomposition and find that M8 (M1) yields SGWB power at $\ell > 1$ relative to the monopole ($\ell = 0$) 100 (10) times larger than M12. This is confirmed by the increasing degree of popcorn-like behaviour for the timeseries we construct below and present in Fig. 5.2.

Following the common practice for statistical tests on time series when only a single

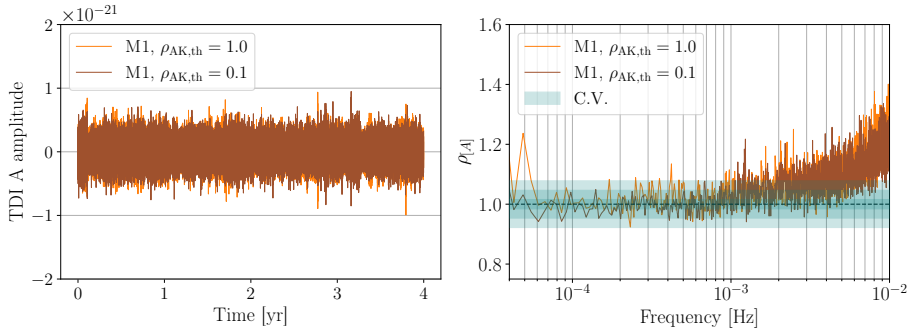

Figure 5.3: Rayleigh test applied to the constructed EMRI foregrounds model M1 with varying downselection SNR thresholds $\rho_{\mathrm{AK,tr}} = 1 \quad (0.1)$ in orange (brown). (*Left panel*) EMRI SGWB in the time domain A channel. (*Left panel*) Rayleigh statistics as a function of frequency in logarithmic scale. In the right panel, the expected value $\rho_{[A],\mathrm{exp}} = 1$ is shown as a black dashed line, and the shaded blue area represents the Rayleigh test's confidence intervals. Adding faint sources does not significantly change the test statistics, as the loudest sources dominate the time series.

realization is available, we invoke ergodicity (i.e., the equivalence in distribution between ensemble averages and time averages) as a working hypothesis. We divide the time series into 3000 chunks, resulting in a frequency resolution of $\Delta f \sim 10^{-5}$Hz. This value is small enough to consider individual chunks' spectra approximately stationary, provides suitable resolution in the mHz band, and yields a large enough set of samples for ensemble average estimators with small variances. Fig. 5.2 (right column) shows the results of the Rayleigh test statistic for the three models considered. We present the test statistics $\rho_{[A]}$ as a function of frequency for the TDI channel A; depending on the considered EMRI catalogs, approaching higher frequencies ($f \gtrsim 1$mHz), the test shows varying degrees of deviation from the expected value $\rho_{[A],\mathrm{exp}} = 1$ (shown as a black line in each panel), suggesting the presence of either non-Gaussianities or non-stationarities at such frequencies.

Upon closer inspection of each model, we discern the test response to variations in the number of sources in the catalogs:

- M8 (Fig. 5.2, bottom row): the time-domain SGWB shows a time-modulation due to the low number of sources in the catalog contributing to the foreground ($N_{\mathrm{final}} = 3209$). This is expected to be a source of deviation from the null hypothesis, i.e. stationarity and Gaussianity. The test exhibits small fluctuations at low frequencies and more substantial ones at $f \gtrsim 1$mHz. See also Ref. [54] for a discussion on non-Gaussian signals as non-stationarity mimickers for tests involving ergodicity.

- M1 (Fig. 5.2, middle row): the larger number of sources ($N_{\mathrm{final}} = 26932$) corresponds to a smaller time-modulation compared to M8. The test fluctuations at low frequencies disappear, but deviations at high frequencies exceeding ~ 1 mHz are still noticeable, indicating a stationarity or Gaussianity violation, albeit much smaller than M8.

- M12 (Fig. 5.2, top row): having the largest number of sources in the catalog ($N_{\rm final} = 319309$), the high-frequency range of the SGWB is more densely populated, making the test violation almost negligible at the target highest significance of 3σ.

We additionally investigate the robustness of our findings on model M1, relaxing the threshold for source removal described in Section 4.1.2 from $\rho_{\rm AK,th} = 1$ to 0.1. Therefore, the number of sources effectively contributing to the SGWB evaluation is increased from $N_{\rm final} = 26932$ to 209072. Fig. 5.4, illustrates how the EMRI foreground for M1 changes when adding sources with starting approximated SNR in the range $0.1 < \rho_{\rm AK} < 1$. Non-coalescing EMRIs with $\rho < 1$ accumulate below 1mHz, contributing to the low-frequency component of the SGWB, with only a modest increase in SNR of about 2.5%, from $\rm SNR_{SGWB} = 311$ to 319. This is in line with our expectations: as coalescing binaries are typically resolvable with SNR = 20 up to $z = 1$, EMRIs would have an SNR above 1 should they be placed at the largest redshift in our integration range, $z < 4.5$, except for low-mass systems ($M \approx 10^4 M_\odot$) contributing only very little to the SGWB. Conversely, a large number of non-coalescing EMRIs with $\rho < 1$ accumulates below 1mHz, hence the observed contribution SGWB. Moreover, low-frequency non-coalescing EMRIs are expected to be highly eccentric, which suppresses GW emission in the cross-polarization. This aligns with

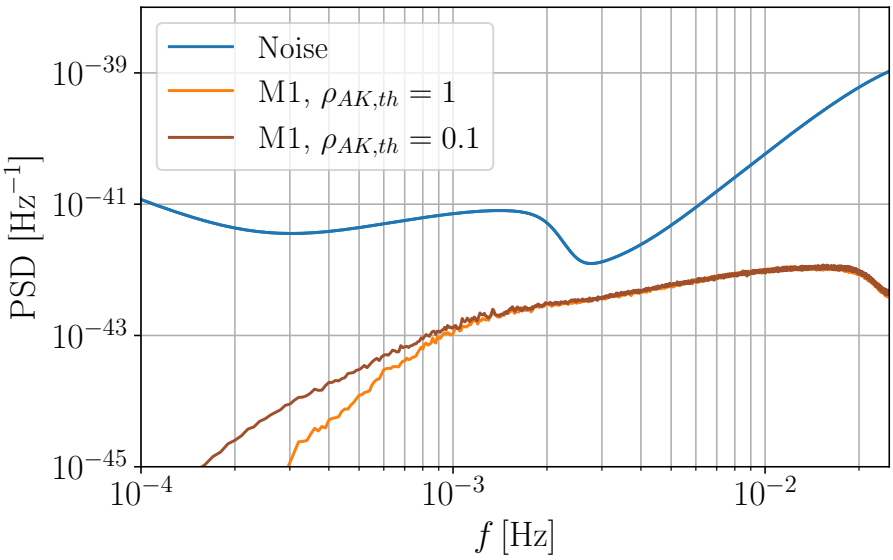

Figure 5.4: SGWB PSD in the LISA A channel for model M1 with two different preliminary cuts in the populations, $\rho_{AK,\rm th} = 1$ ($\rho_{AK,\rm th} = 0.1$) solid orange line (solid brown line). Solid blue line represents the LISA instrumental noise, together with the GF. The inclusion of fainter sources in the analysis populates the low-frequency band.

findings from Ref. [148], which shows a consistent 5% reduction in background SNR after excluding the faintest sources across all examined models. Fig. 5.3 (right panel), shows that the Rayleigh test results remain largely unchanged. The addition of faint sources to the catalog does not substantially alter the foreground properties investigated. In fact, including low SNR sources appears to primarily affect the signal PSD in Fig. 5.4 only at $f < 1$mHz, whereas the Rayleigh test in Fig. 5.3 shows deviation at and above $f > 1$mHz.

Therefore, the choice of considering only sources with $\rho_{AK} > 1$ suffices to support our conclusions.

The threshold for detectability ($\rho = 20$), could similarly influence the SGWB stationarity and its overall brightness. Establishing realistically EMRI detectability is heavily dependent on (i) an accurate detector characterization and (ii) a precise definition of the detection statistics. In absence of those, we opt to leave this for future work.

5.1.4 Caveats and takeaways

We performed a statistical analysis of EMRI backgrounds to quantify deviations from the core assumptions of Gaussianity and stationarity that underlie standard LISA inference methods.

Main results. The Rayleigh test revealed varying degrees of non-Gaussianity and non-stationarity, with violations scaling inversely with the number of unresolved sources contributing to each foreground model. Model M12 ($N_{\rm final} \approx 3.2\times10^5$ sources) showed negligible deviations at 3σ significance across the LISA band, whereas M8 ($N_{\rm final} \approx 3.2 \times 10^3$) exhibited substantial departures from the null hypothesis at $f \gtrsim 1\,\mathrm{mHz}$. The fiducial M1 model ($N_{\rm final} \approx 2.7 \times 10^4$) displayed intermediate behavior, with detectable but moderate violations.

LISA implications. The observed violations have dual implications for LISA science. First, non-Gaussian and non-stationary signatures provide discriminating features to disentangle overlapping SGWB components through extended likelihood models [166, 150]. Second, and more critically, misestimation of the PSD under the Gaussian approximation introduces systematic biases in parameter reconstruction for resolvable sources. As demonstrated in [166], overestimating the PSD through Gaussian likelihood assumptions leads to overestimation of source luminosity distances, potentially producing systematic offsets in population-level inference.

5.2 Anisotropies from Milky Way and its satellites

As mentioned in Section 3.6, the GF [139] is expected to dominate LISA observations, particularly in the 0.5–3 mHz range. Unlike isotropic EMRI and SOBBH background characterized in Sections 4.1 and 4.2, the bulk of GF arises from systems located toward the Galactic center, leading to a strongly anisotropic signal. Combined with the annual orbital motion of LISA, this anisotropy produces a time-domain modulation of the signal, resulting in a cyclostationary stochastic process. The relevance of this statistical feature is twofold: it allows better characterization of the foreground itself, and it facilitates the study and identification of other overlapping signals through degeneracy breaking. Previous works [134, 76] have studied this phenomenon, although typically using empirical formulas. Additionally, similarly to Section 5.1, Ref. [53] has shown that the GF exhibits significant non-Gaussianity in the 2–10 mHz range, suggesting that multiple statistical signatures may need to be jointly modeled for accurate foreground characterization. As discussed in Section 3.6, we also expect highly anisotropic stochastic signals from unresolved sources in MW satellite galaxies, similar to the GF. These satellite contributions offer valuable probes of underlying astrophysical populations—providing insights into total stellar mass, metallicity, and star formation history [47, 89]. A detailed analysis of the SGWB from the LMC has been presented in Ref. [157], where anisotropy is modeled via spherical harmonic decomposition—a technique widely adopted for cosmological sources. In this section, we introduce an alternative analytic framework to characterize cyclostationary SGWBs in LISA through frequency-domain modeling of the autocorrelation function. This approach provides a natural parameterization in terms of sky distribution properties and detector response, enabling signal recovery and parameter estimation without requiring pixelation or spherical harmonic expansions. The method will be further extended to the time-frequency domain in Section 5.3 to address practical limitations in global fit scenarios requiring sub-annual noise updates.

5.2.1 Cyclostationary processes

Cyclostationary processes are stochastic processes whose statistical properties are periodic in time. In this work, we focus on those exhibiting periodicity in their second-order statistics (e.g., the autocorrelation function) [192]. In particular, a continuous stochastic process $X(t)$ having finite second-order moments is said to be wide-sense cyclostationary with period T if the expectation values

$$E\left[X(t)\right] = m(t) = m(t+T), \tag{5.4}$$

$$E\left[X(t')X(t)\right] = \Sigma(t',t) = \Sigma(t'+T, t+T) \tag{5.5}$$

are periodic functions with period T, for $(t',t) \in \mathbf{R}^2$. Defining $\tau_d = t' - t$, then $\Sigma(t',t)$ can be equivalently represented as a Fourier series:

$$\Sigma(t',t) = \sum_{n=-\infty}^{+\infty} B_n(t'-t)e^{-2\pi i n t/T}, \tag{5.6}$$

where $B_n(\tau_d)$ is

$$B_n(\tau_d) = \frac{1}{T}\int_0^T dt \Sigma(t+\tau_d, t)e^{-2\pi i n t/T}. \tag{5.7}$$

The FT of Eq. 5.7 corresponds to the so-called *cyclic spectrum*. We contextualize the above formalism to populations of unresolved WD binaries in the MW and its satellites, as seen by LISA. For an ensemble of N WD binaries, the total signal $s(t)$ is given by

$$s(t) = \sum_{i=0}^{N} X_{\rm GW}(t, \theta), \tag{5.8}$$

where θ is the vector of parameters characterizing the GW signal $X_{\rm GW}$. We list them here for completeness: the initial phase of the signal ϕ_0, the inclination angle ι of the binary system's angular momentum relative to the line of sight, the polarization angle ψ, the distance D to the binary, the ecliptic coordinates (λ, β) describing the source locations in the SSB, the chirp masses $\mathcal{M}_c$, and the angular frequencies ω_s. $X_{\rm GW}$ is the signal in one of the three LISA noise-orthogonal channel [185]. We will consider only sources in their early inspiral, such that in the low-frequency approximation, each monochromatic signal in $X_{\rm GW}$ is expanded as

$$X_{\rm GW}(t) = 2C(t)(1+\cos^2\iota)\cos(\omega_s t + \phi_0)A + 4S(t)\cos\iota\sin(\omega_s t + \phi_0)A, \tag{5.9}$$

where $C(t)$ and $S(t)$ depend on the motion of the LISA satellites constellation. Here, A is the GW amplitude, introduced in Eq. 1.61.

The largest fraction of Galactic WD binaries are roughly approximated to share the same distance D (this is even more true for MW satellites). Thus, to compute the autocorrelation function $\Sigma(t', t)$, we assume that the population distribution $p(\theta)$ is separable into five distributions, $p_1(\phi_0)$, $p_2(\iota)$, $p_3(\psi)$, $p_4(\beta, \lambda)$, and $p_5(\mathcal{M}_c, \omega_s)$. Specifically, we assume probability distributions for ψ (ϕ_0, $\cos\iota$) uniform in $[0, 2\pi]$ ($[0, 2\pi]$, $[-1, 1]$). In what follows, we refer for brevity to whole WD populations in MW satellites as "sources". Then, we compute the autocorrelation function as

$$\Sigma(t', t) = N \int_0^{2\pi} d\phi_0 \int_0^{2\pi} d\psi \int_{-1}^{1} d\iota \int_{V_4} dV_4 \frac{p_4}{16\pi^2} \int_{V_5} dV_5 p_5 \, X_{\rm GW}(t') X_{\rm GW}(t), \tag{5.10}$$

where V_4 (V_5) is the differential volumes of the ecliptic latitude and longitude β, λ (chirp mass, orbital angular frequency $\mathcal{M}_c, \omega_s$) parameter space, respectively. By inserting Eq. 5.9 in Eq. 5.10 and performing the integral over ψ, $\cos\iota$ and ϕ_0, we obtain:

$$\Sigma(t', t) = \frac{N}{8} \int_{V_4} dV_4 p_4 \int_{V_5} dV_5 p_5 A^2 \cos\left(\omega_s (t' - t)\right) \left[\frac{56}{15} C(t') C(t) + \frac{8}{3} S(t') S(t)\right]. \tag{5.11}$$

Introducing the additional variable $\tau_s = \frac{t+t'}{2}$, we define $B(\tau_s, \tau_d) = \Sigma(\tau_s + \tau_d/2, \tau_s - \tau_d/2)$. Therefore, the autocorrelation $B(\tau_s, \tau_d)$ is explicitly periodic in τ_s for any fixed τ_d. Thus, similarly to Eq. 5.6 we decompose it in Fourier series

$$B(\tau_s, \tau_d) = \frac{N}{8} \int_{V_5} dV_5 p_5 \sum_{n=-8}^{n=8} B_n(\tau_d) \exp\left[2\pi i \frac{n\tau_s}{T}\right] A^2 \cos(\omega_s \tau_d). \tag{5.12}$$

Note that the limitation to $|n| \leq 8$ is a known result for the Fourier series decomposition of product of oscillating functions at harmonic frequencies (see e.g. [91]). Further assuming that $p_5(\mathcal{M}_c, \omega_s)$ is only mildly frequency dependent – or equivalently that τ_d is such that $\tau_d/T \ll 1$– we can approximate $B_n(\tau_d) = B_n(\tau_d = 0)$, as in Ref. [78]. Therefore, we will

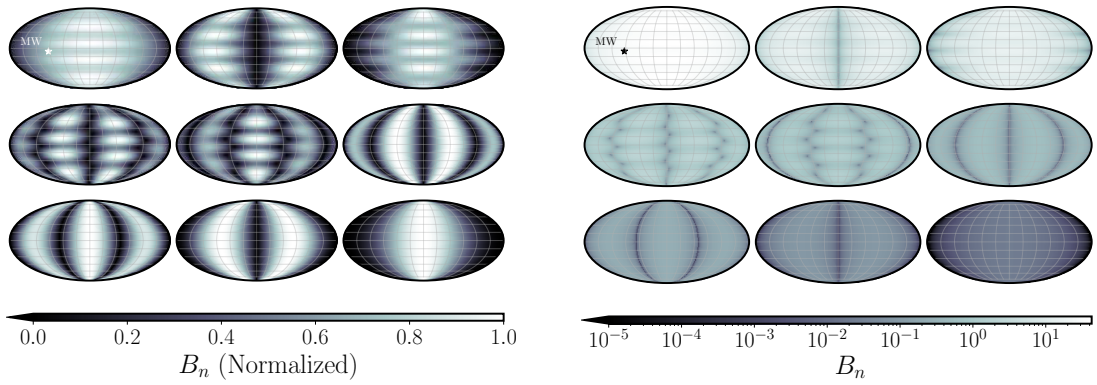

Figure 5.5: Harmonic decomposition of LISA modulation. Each point in the sky corresponds to a source with the angular size of the LMC ($\sigma \approx 0.015$ rad). In each panel, the plots are associated with harmonics in increasing order of n from top left ($n = 0$) to bottom right ($n = 8$). The projections are expressed in ecliptic coordinates. (*Left panel*) Each heatmap represents the magnitude of the Fourier Coefficient normalized to the relative maximum value in each plot. (*Right panel*) The absolute magnitude of the Fourier coefficient is shown in logarithmic scale.

refer to it in what follows as B_n. We express the autocorrelation function in Eq. 5.12 as the product of two functions

$$B(\tau_s, \tau_d) = R(\tau_s)\mathcal{C}(\tau_d), \tag{5.13}$$

where

$$R(\tau_s) = \sum_{n=-8}^{n=8} B_n e^{2\pi i \frac{n\tau_s}{T}}, \tag{5.14}$$

$$\mathcal{C}(\tau_d) = \frac{N}{8} \int_{V_5} dV_5 p_5 A^2 \cos(\omega_s \tau_d). \tag{5.15}$$

By Fourier transforming over τ_d and τ_s, we obtain

$$C(f, f') = \sum_{n=-8}^{n=8} B_n S_h \left(\frac{f' + f}{2}\right) \delta \left(f - f' + \frac{n}{T}\right), \tag{5.16}$$

where δ is the Dirac delta function. In a discretized form (which is the natural domain for an observed digital timeseries), $C(f, f')$ corresponds to a band matrix C_{ij} with $2n+1 = 17$ non-zero diagonals spaced $|i - j| = T_{\rm obs}/T$ apart from each other. Here, the periodicity of the correlation function is T and the frequency resolution is limited by the observation time (i.e., the LISA mission duration) $T_{\rm obs}$. Each diagonal is proportional to the one-sided PSD S_h, which depends on the intrinsic astrophysical properties of the source population.

The proportionality constants B_n are the n-th harmonic Fourier coefficients of the time-domain signal modulation, and they depend solely on the extrinsic properties of the source population, e.g., the sky distribution. This is a crucial advantage of our approach, as it

directly relates the signal modulation to the source distribution. As compared to other approaches in literature, it does not require either pixelation or spherical harmonics decomposition. Once a suitable parameterization for C_{ij} is introduced for a specific detector, it allows for a straightforward interpretation of the inference results.

5.2.2 Modulation characterization

Following the detailed mathematical derivation in Ref. [53], reported for completeness in Appendix C, we construct the LISA modulation for a bivariate Gaussian distribution in the sky-coordinates. Therein, the contribution of different harmonics to the overall time domain modulation are listed. In particular, the Fourier coefficient B_n, are described by five independent parameters: the ecliptic coordinates of the distribution expectation value $(\sin\beta, \lambda)$, and the two principal component axes variances (σ_1, σ_2), along with their rotation angle ψ with respect to the coordinate system. Given the modest angular size of MW satellites, we choose to enforce the constraint $\sigma_1 = \sigma_2$, which reduces the number of free parameters to four, with ψ being completely degenerate. Therefore, the only parameters left are $\sigma \equiv \sigma_1, \sin\beta, \lambda$, and our inference model is therefore specified by $R(\tau_s; \sin\beta, \lambda, \sigma)$.

In what follows, we explore the satellite modulation as a function of its free parameters. In Fig. 5.5, we show the magnitude of the Fourier coefficients across the sky, each panel corresponding to a single harmonic, for a reference $\sigma \approx 0.015$ rad. We highlight that due to the modulation being a real function of time, $B_{-n} = B_n^*$. The zero-th mode, which corresponds to the main diagonal in the covariance matrix, is dominant compared to the others. For a given sky position, the higher-order modes are increasingly subdominant. The odd (even) harmonics share common maxima (minima) along the directions $(\sin\beta, 0)$, although we point out that the specific value depends crucially on the chosen initial LISA position along its orbit. Overall, far from the ecliptic plane, the LISA modulation becomes less relevant: if a source is located close to the ecliptic poles, the Doppler effect induced by the LISA orbital motion is negligible. Conversely, the ecliptic longitude appears to have a less significant influence on the modulation.

In Fig. 5.6, we additionally show the magnitude of each Fourier coefficient for a set of known MW satellites (together with Andromeda), whose parameters are listed in Table 5.1 (for more details, see Ref. [158] and references therein). These are the sources we will consider in this study. We observe that the size of a source may also play a role on its modulation. In particular, smaller objects result in larger coefficients. However, this effect is less significant compared to that of the ecliptic latitude and longitude. Sagittarius and the MW are relatively close to each other in the sky, but some B_ns of the former are larger than those of the latter (e.g., $n = 4$). On the contrary LMC, located near the ecliptic South Pole, yields much weaker modulation than the MW even if it has a comparable angular size to Sagittarius.

We further investigate the impact of the source angular size in Fig. 5.7. In the left panel, we show the relative uncertainties in the time modulation for the SMC and the MW, varying their σ by $\pm 50\%$. For the Galactic case, relative deviations reach nearly 10% during a LISA orbit, as opposed to a maximum of 2% for SMC. Similarly, in the right panel of Fig. 5.7, the same variability is illustrated in frequency domain. Ranges for each MW modulation coefficient are much larger than the corresponding ones for the SMC. Hence, the modulation phenomenology is only mildly influenced by the satellite angular size, thus only marginally affecting the SGWB detection. A more detailed discussion is provided in

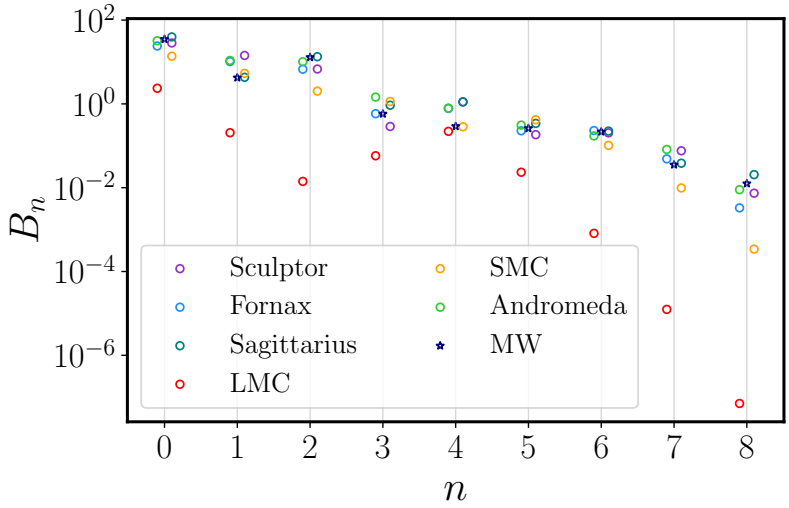

Figure 5.6: Magnitude of Fourier coefficients in logarithmic scale at various harmonics for the MW, Andromeda and the set of satellites considered in this study. Higher harmonics are generally less significant. The values for the LMC are lower compared to other sources due to its proximity to the ecliptic South Pole. Individual source parameters are listed in Table 5.1.

Section 5.2.4.

5.2.3 Spectrum characterization

We now aim to provide a framework to compute the effective astrophysical signal. We present below two simple models to characterize the spectra S_h in Eq. 2.9 for the MW, its satellites, and other nearby Galaxies. In Fig. 5.8 (left panel), we show the spectra for the LMC, and SMC, Sculptor, Fornax, and Sagittarius, as well as Andromeda and the MW foreground. The SNR values are computed for each source, assuming a LISA mission duration of 4 years, and are shown for the loudest (LMC, SNR = 3.46) and faintest (Sculptor, SNR = 0.002). For simplicity, we show only the upper envelope of each modulated stochastic signal. The spectra are accompanied by the individual modulations (right panel), as described in Section 5.2.2.

Satellites and nearby Galaxies

Upon Fourier transforming $\mathcal{C}$ in Eq. 5.15 we describe the spectrum of the signal as

$$S_h(f) = \frac{N}{8} \int d\mathcal{M}_c \int df_s p(\mathcal{M}_c) p(f_s) \frac{(G\mathcal{M}_c)^{10/3}}{(c^4 D)^2} (\pi f_s)^{4/3} \delta(f - f_s), \tag{5.17}$$

where for brevity we relabelled $f = \frac{f+f'}{2}$. The integral in Eq. 5.17 corresponds to the isotropic spectrum amplitude evaluated within a spherical shell of radius D, as computed in Ref. [145]. This formulation is particularly convenient, as the sky integration is already incorporated in the coefficients B_n.

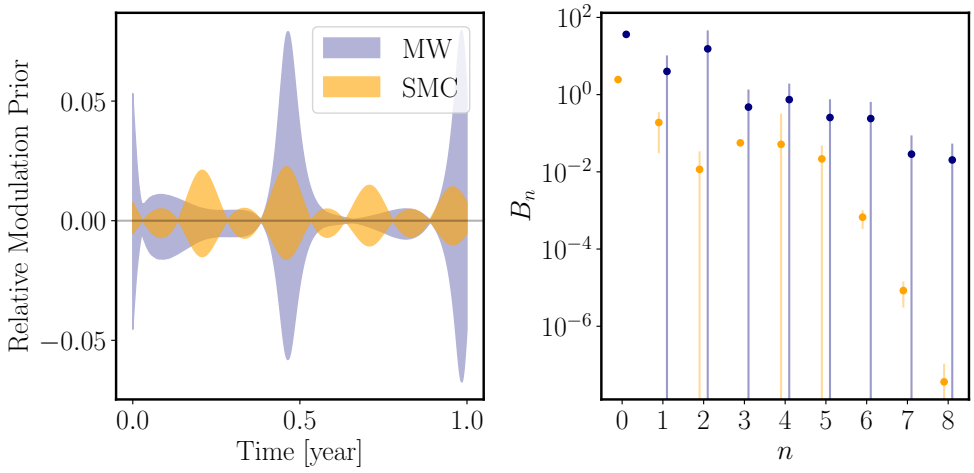

Figure 5.7: The Impact of varying σ: Each panel shows the modulation calculated by varying σ by $\pm 50\%$ of its original value, while keeping the true position of the relative object fixed. (*Left panel*) The 90% confidence interval of the relative time modulation. 8*Right panel*) The 90% confidence intervals for different harmonic modes. Overall, the SMC modulation shows less variability compared to the MW when exploring the σ prior.

We recast the frequency dependence in the integral in terms of $x = a/(Kt_0)^{1/4}$. The variable x represents the separation normalized to that of a binary with time to merger t_0, and a is related to f_s through Kepler's law and

$$K = \frac{256G^3}{(5c^5)}(m_1 + m_2)m_1m_2. \tag{5.18}$$

Thus, Eq. 5.17 becomes

$$S_h(f) = \frac{N}{8}\int d\mathcal{M}_c \int \left|\frac{dx}{df_s}\right| df_s p(\mathcal{M}_c)p(x(f_s))\,\frac{(G\mathcal{M}_c)^{10/3}}{(c^4D)^2}\,(\pi f_s)^{4/3}\,\delta(f - f_s). \tag{5.19}$$

The Jacobian in Eq. 5.19 reads

$$\left|\frac{dx}{df_s}\right| = \frac{2}{3}\frac{G^{1/3}}{\pi^{1/3}\bar{K}^{1/4}}\frac{(m_1+m_2)^{1/12}}{(m_1m_2)^{1/4}}f_s^{-5/3}, \tag{5.20}$$

and upon substitution in Eq. 5.19 and integrating with the Dirac delta, we get

$$S_h(f) \propto \int d\mathcal{M}_c \mathcal{M}_c^{35/12} f^{-1/3} p(\mathcal{M}_c)p(x(f_s = f)) \tag{5.21}$$

This is a convenient parametrization of the integral, as it matches closely the one adopted in previous population studies (see, e.g., Ref. [129, 130, 111]). In particular, we adopt the following parametrization

$$p(x) \propto \begin{cases} x^{4+\alpha}\left[(1+x^{-4})^{\frac{\alpha+1}{4}} - 1\right], & \text{if } \alpha \neq -1 \\ x^3 \ln\left(1+x^{-4}\right), & \text{if } \alpha = -1 \end{cases}. \tag{5.22}$$

where α is the power-law slope describing the separation distribution at WD formation and the star formation is assumed to be constant (for detailed derivation, see Ref. [129]). In our analysis, we will consider $\alpha = -1.3$ following results obtained in Ref. [130], based on two complementary large, multi-epoch, spectroscopic samples: the Sloan Digital Sky Survey (SDSS), and the Supernova Ia Progenitor survey (SPY). When $\alpha \leq -1$, in the limit of $x \ll 1$, $p(x)$ can be approximated by $x^{4+\alpha}$, so that the final model for $S_h(f)$ is

$$S_h(f) \propto f^{-(9+2\alpha)/3} \int d\mathcal{M}_c \mathcal{M}_c^{5(3-\alpha)/12} p(\mathcal{M}_c). \tag{5.23}$$

Eq. 5.23 provides a semi-analytical power-law describing the SGWB spectrum for a generic MW satellite with constant star formation. This may be an optimistic assumption for lower-mass satellites, which typically stop forming stars early, but it represents a good approximation for satellites like the SMC and LMC that continue forming stars. Its amplitude depends on the integral in the chirp mass, which we perform numerically from the distribution on m_1 and m_2. We assume that m_1 follows the mass distribution of single white dwarfs. Following Ref. [111], we consider a univariate Gaussian mixture with three components for m_1 with weights $w = (0.66, 0.24, 0.1)$, means $\mu = (0.55, 0.72, 0.3)\mathrm{M}_\odot$, and standard deviations $\sigma = (0.07, 0.3, 0.078)M_\odot$, as measured in Ref. [106] based on large spectroscopic sample. Observations show that at the main-sequence stage, the secondary stars follow a mass-ratio distribution that is approximately flat, rather than the same mass distribution as the primary (e.g., [136]) Therefore, in our model we draw m_2 from a flat distribution between $0.15\mathrm{M}_\odot$, the minimum mass of observed extremely-low-mass (ELM) white dwarfs, and m_1. For a small number of cases in which we get $m_1 < 0.25M_\odot$, i.e., falling into the ELM category, we draw m_2 from the range $[0.2, 1.2]M_\odot$ with equal probability. Note that in the latter case, the definitions of m_1 and m_2 are swapped.

The amplitudes of the SGWB at $f = 10^{-3.5}$Hz are reported for different satellites in Table 5.1. Therefore, a satellite spectrum with assumed mass distribution has only two free parameters: we reparameterize them with the power-spectrum amplitude at a reference frequency $A(f = 10^{-3.5}\mathrm{Hz})$ and its slope $\gamma = -(9 + 3\alpha)/3$. Notably, when $\alpha = -1$, the slope simplifies to $-7/3$, which is the well-known value associated with a SGWB originating solely from inspiral signals.

For modeling the GF spectrum, we adopt the phenomenological approach introduced in Ref. [102], with its explicit formulation given in Eq. 3.62.

System	λ	β	Sky Area [deg^2]	D [kpc]	Stellar Mass [$\mathrm{M}_\odot$]	$\log_{10} A(f = 10^{-3.5}\mathrm{Hz})$
Andromeda	27.8°	33.3°	3.11	765.0	$\sim 10^{10}$	-41.96
Fornax	22.5°	−46.9°	0.17	139.0	2.0×10^7	-43.18
LMC	−47.7°	−85.4°	77	50.0	1.5×10^9	-40.42
SMC	−47.9°	−64.6°	13	60.6	4.6×10^8	-41.10
Sagittarius	−78.1°	−7.6°	37	26.7	2.1×10^7	-41.73
Sculptor	−1.6°	−36.5°	0.076	86.0	2.3×10^6	-43.70

Table 5.1: Properties of selected systems including MW satellites and Andromeda. Masses, distances, and sky locations are from Refs. [132, 68, 174]. The last column shows the logarithmic amplitude of the SGWB as defined in Section 5.2.3.

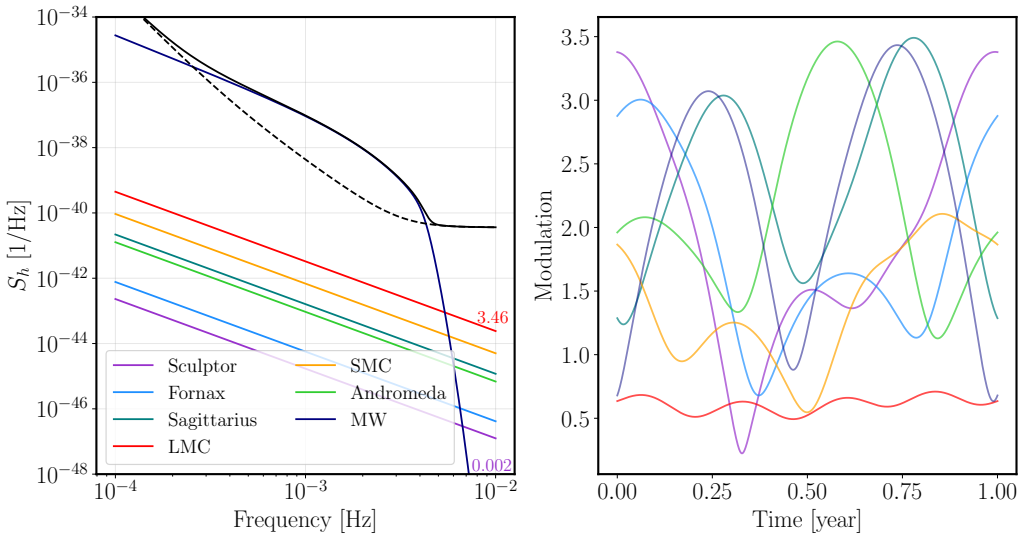

Figure 5.8: (*Left panel*) The astrophysical SGWB spectrum is computed using a semi-analytical model from Eq. 5.23. The amplitude level mainly depends on the distance and size of the source, in terms of stellar mass (or the number of WD binaries within it). The slope of the power-law is $\gamma = -(9 + \alpha)/3$, where α, a parameter related to the initial WD separation distribution, is α. The black dashed line refer to the LISA instrumental noise. *Right panel*) The modulation in time as seen by LISA during one year for different astrophysical sources, assuming a two-dimensional Gaussian distribution of unresolved WD binaries in the sky for each of them.

Data	Model	Parameters	References
SMC	Cyclo	(A, γ, σ)	Fig. 5.9 (Panel 1)
LMC	Cyclo	(A, γ, σ)	Fig. 5.9 (Panel 2)
Andromeda	Cyclo	(A, γ, σ)	Fig. 5.9 (Panel 3)
Sagittarius	Cyclo	(A, γ, σ)	Fig. 5.9 (Panel 4)
MW + SMC	Cyclo + Cyclo	$(\alpha, A, f_2, \psi, \sigma_1, \sigma_2) + (A, \gamma, \sigma)$	Fig. D.1
MW + LMC	Cyclo + Cyclo	$(\alpha, A, f_2, \psi, \sigma_1, \sigma_2) + (A, \gamma, \sigma)$	Fig. D.2
MW + Andromeda	Cyclo + Cyclo	$(\alpha, A, f_2, \psi, \sigma_1, \sigma_2) + (A, \gamma, \sigma)$	Fig. D.3
MW + Sagittarius	Cyclo + Cyclo	$(\alpha, A, f_2, \psi, \sigma_1, \sigma_2) + (A, \gamma, \sigma)$	Fig. D.4
LMC (realistic)	Cyclo	(A, γ, σ)	Fig. D.5
LMC (realistic)	Stat	(A, γ)	Fig. D.5
MW + LMC-like	Cyclo + Cyclo	$(\alpha, A, f_2, \psi, \sigma_1, \sigma_2) + (A, \gamma, \lambda, \sin\beta, \sigma)$	Fig. D.6

Table 5.2: Summary of parameter estimation. The first column lists injected sources (all including LISA noise), the second describes the signal model, the third specifies the inferred parameters, and the last column points to the corresponding figure showing posteriors.

5.2.4 Inference result

In this study, we focus exclusively on the parameter estimation of satellites' and MW SGWBs together with the instrumental noise. We proceed with the assumption that non-stochastic GWs sources have been subtracted perfectly. Additionally, we consider the absence of data gaps, resulting in the generation of an idealized residual dataset. Throughout the analysis, we consider 4 years of observation time of LISA. Data are expressed in terms of TDI variables [185]. Specifically, we work with the FT of the A and E channels. We assume that data are distributed according to the Gaussian likelihood. In the most general case, we include in the likelihood the contribution from the GF, the satellite SGWB and the instrumental noise:

$$\log \mathcal{L}(\tilde{d}|\boldsymbol{\theta} = \{\boldsymbol{\theta}_{\rm MW}, \boldsymbol{\theta}_{\rm sat}, \boldsymbol{\theta}_{\rm n}\}) \propto -\sum_{i=A,E} \left(\frac{1}{2} \log(\det C_i) + \frac{1}{2} \tilde{d}_i^T C_i^{-1} \tilde{d}_i \right), \tag{5.24}$$

$$C_i = [C_{\rm MW}(\boldsymbol{\theta}_{\rm MW}) + C_{\rm sat}(\boldsymbol{\theta}_{\rm sat}) + C_{\rm n}(\boldsymbol{\theta}_{\rm n})]_i \tag{5.25}$$

where $C_{\rm MW}$ and $C_{\rm sat}$ correspond to the MW and satellite covariance matrices, described earlier. C_n is the diagonal noise covariance matrix with PSD, as described in Eq. 3.60. The parameters for each process are:

- $\boldsymbol{\theta}_{\rm MW} = \{\mathcal{A}_{\rm MW}, \alpha, f_{\rm knee}, f_2, f_1, \lambda, \sin\beta, \sigma_1, \sigma_2, \psi\}$;
- $\boldsymbol{\theta}_{\rm sat} = \{\mathcal{A}_{\rm sat}, \gamma, \lambda, \sin\beta, \sigma\}$;
- $\boldsymbol{\theta}_n = \{\mathcal{P}_{\rm tm}, \mathcal{P}_{\rm oms}\}$.

Calligraphic letters indicate the $\log_{10}$ of the corresponding quantity, e.g., $\mathcal{A}_{\rm MW} = \log_{10} A_{\rm MW}$. We choose uniform priors for each parameter listed. To compare results obtained from cyclostationary and stationary inference, we use the log-Bayes factor:

$$\log_{10} \mathcal{B}^{\rm cyclo}_{\rm stat} = \log_{10} \mathcal{Z}_{\rm cyclo} - \log_{10} \mathcal{Z}_{\rm stat}, \tag{5.26}$$

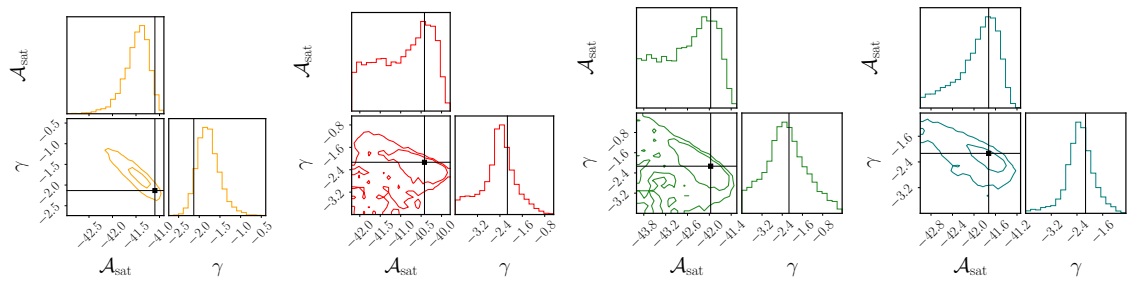

Figure 5.9: Marginalized probability distribution of the amplitude and slope SGWB parameters for SMC, LMC, Andromeda, and Sagittarius (from left to right). In this analysis, the GF is excluded from both the injection and parameter estimation stages. Additionally, we assume that the sky locations of the satellites are known throughout the inference process.

.

We now present our inference results. First, we examine LISA capability to observe sources in the presence of instrumental noise alone. Next, we assess the impact of including the MW foreground in both simulated data and inference model. In addition, we investigate the detectability of the MW largest nearby Galaxy, Andromeda.

The characteristic parameters of galaxies and satellites we investigate are reported in Table 5.1. Table 5.2 provides a summary of the parameter estimation discussed in this section and the following ones.

Milky Way satellites and Andromeda

To determine which MW satellite might be observable by LISA, we first study its detectability in presence of instrumental noise only, excluding the MW foreground from the data and parameter estimation. Specifically, we inject idealized satellites one by one, with their spectra and modulation described in Sections 5.2.2 and 5.2.3, along with their corresponding amplitude, sky location, and angular radius.

Initially, we also fit the extrinsic parameters of the SGWB source, specifically the size and sky location. The results indicate that we can clearly recover both the spectrum amplitude and slope only for the SMC. Regarding other parameters, the $\sin\beta$ value closely matches the true value, while the ecliptic longitude appears difficult to estimate accurately. This inaccuracy is likely due to the SMC's proximity to the South Pole, where the partial degeneracy in longitudes becomes stronger. The size σ is not constrained due to its mild influence on the modulation with respect to the latitude, as explained in Section 5.2.2.

From this initial analysis, we report a central finding: the ability to detect a satellite is influenced by the interplay between the astrophysical spectrum and the modulation effects in reconstructing this type of SGWB. From an astrophysical point of view, the LMC is favored due to its mass and distance, which makes it more probable to detect individual WD binaries [158]. However, from the perspective of SGWB detection, the SMC proves to be easier to detect due to its location.

Despite the SGWB signal from satellites still being unknown, these objects have already been observed through electromagnetic emission. Therefore, in what follows, we assume perfect prior knowledge of their locations to determine if this improves the spectrum parameter estimation. By doing so, the spectral reconstruction improves for LMC, Sagittarius and Andromeda. The SGWB amplitude is well constrained only for SMC. The marginalized posterior probability of their slope and amplitude are shown in Fig. 5.9. Although LISA is more sensitive to the sky positions of Fornax and Sculptor compared to the LMC or SMC, the spectra from these satellites are too faint to be detected, even if their positions (hence their signal modulation) are known.

Next, we consider the case where the MW foreground is included in the data, alongside the instrumental noise and each satellite signal. The Galactic signal is generated using the spectrum given by Eq. 3.62 with the optimal fit parameters found in Ref. [102], considering 4 years of LISA observation time and SNR threshold of 7 for WD detection. For the modulation, we adopt the same strategy used by Ref. [53], where it is generated with $\lambda = -93.16°$, $\beta = -5.53°$, $\sigma_1 = 0.042$, $\sigma_2 = 0.138$, and $\psi = -56.72°$. These parameters result from the best fit of the interpolation between a rotated two-dimensional Gaussian distribution and a realistic unresolved MW catalog of WD binaries. This approach is an approximation that only partially captures the more complex structure of the MW (as depicted in Fig.5 of Ref. [53]). A more sophisticated description of the MW time-variability is deferred to future work, where we employ a mixture of distributions to better reconstruct the global structure of the MW. In this scenario, as expected, the detection of MW satellites is highly compromised. The posterior distribution for SMC, LMC, Andromeda, and Sagittarius are presented in the Appendix D. Using informative priors on sky location (both for the satellite and MW), we are able to put an upper limit on the amplitude of the SGWB spectrum, while the slope is in general easier to constrain. Nonetheless, a good reconstruction is guaranteed for both the MW foreground and the instrumental noise. In particular, for the MW, we infer only A, f_2, and α, which are the most relevant parameters from an astrophysical point of view. We fix the remaining parameters to their injected values due to computational limitations. By doing so, the NS algorithm will converge more quickly without introducing significant biases in the analysis.

We emphasize that, unlike what is observed for the satellites, we are indeed able to recover the MW size. This is due to the fact that the MW angular size is sufficiently large that modulation effects become sensitive to fluctuations around the true value in the prior parameter space. This aligns with our expectations from Fig. 5.7.

After applying our cyclostationary model to ideal scenarios, we aim to test it in the presence of realistic datasets. Therefore, we inject a realization of the SGWB generated from a realistic catalog of individual WDss. We consider the case of the LMC, since its source catalog has been already used in previous work [113, 112].

Similarly to the previous section, we fix the satellite position in our model. The results align well with those obtained using mock data. We successfully constrain the spectral parameters, particularly the slope. However, the satellite size has a minimal impact on the model, making it difficult to effectively constrain this parameter. We report the posterior distribution in Fig. D.5 in Appendix D. We also infer the signal under the assumption of stationarity, modeling the background signal using a power-law. It is important to note that the amplitude here directly refers to the TDI domains. Additionally, since the T channel is consistently excluded from our analysis, we simplify the model by using a single amplitude parameter associated with both the A and E channels. To compare the

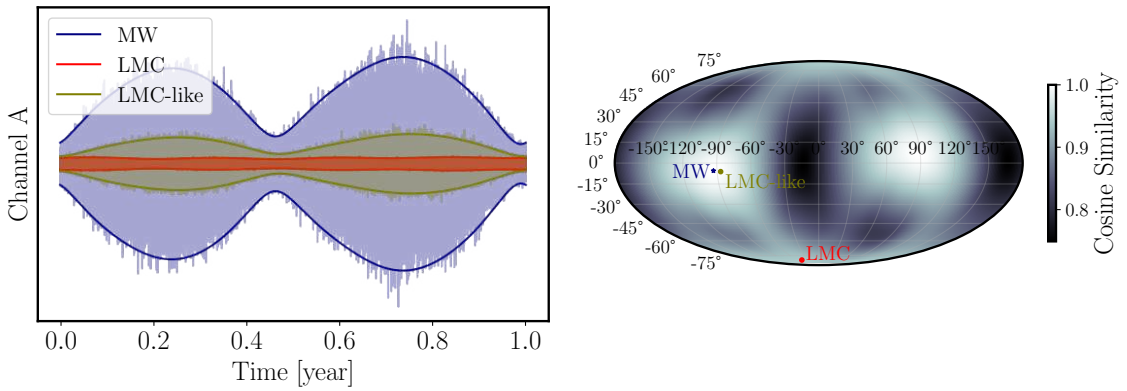

Figure 5.10: (*Left panel*) Simulated signals for MW (teal) and LMC (red), and shifted LMC within the Galactic disk (light blue) field of view over the course of one year. The corresponding modulations are represented as envelopes superimposed on the top and bottom of the signals. (*Right panel*) Cosine similarity map between the MW and the LMC-like satellite modulations. The three dots represent: the center of the MW (navy), the true position of the LMC (red), and the hypothetical position of the LMC-like satellite in the Galactic disk (gold).

amplitude between two models, we divide the TDI amplitude in the stationary case by the zero order Fourier coefficient, computed using the real satellite parameters (size and sky position). It is interesting to note that in both models, the slope of the background is consistent with our simplified theoretical model described in Section 5.2.3.

We compute the Bayes factor as described in Eq. 5.26. We find that $\log_{10} \mathcal{B}^{\text{cyclo}}_{\text{stat}} \lesssim -0.5$, which does not indicate a strong evidence in favor of one of the two hypothesis. This is due to the faintness of the signal modulation. The off-diagonal elements in the covariance matrix are not significant enough to help distinguish the signal from the noise in the main diagonal. As a result, the stationary model can extract the same amount of information. Instead, for the Galactic signal, exploiting the off-diagonal elements will be crucial. However, as previously mentioned, a detailed analysis of realistic data for the MW foreground will be addressed in future work.

Discovering hidden satellite

Unlike electromagnetic emission, which is obstructed by dust and gas in the Zone of Avoidance, GWs are not hindered by such obstructions. Thus, with LISA, we have the potential to observe beyond the Galactic plane. According to Ref. [158], detecting an excess of ~ 100 WDs binaries within the Galactic disk would be considered a statistically significant overdensity. In an ideal scenario, if a satellite were located behind the MW disk, we could attempt to detect its SGWB. For this reason, we inject a LMC like SGWB (i.e., with the same slope and amplitude), in a region close to the Galactic center. For simplicity, we consider the same σ. In Fig. 5.10, we display the simulated signals for the MW

(navy), LMC (red), and the hypothetical LMC behind the MW (green olive). We observe that moving the LMC away from the pole enhances LISA sensitivity to the corresponding SGWB. This adjustment causes the modulation of the signal to become similar to the MW one. Consequently, this can lead to a degenerate covariance matrix structure between the two processes, potentially resulting in inaccurate signal reconstruction or, in the worst case, failure to detect the SGWB. However, this does not occur. As shown in Fig. D.6, we successfully disentangle the SGWB from both the LISA instrumental noise and the MW foreground. Additionally, we reanalyze the same dataset assuming LMC-like SGWB as stationary. In this case, we not only obtain a smaller evidence ($\log_{10} \mathcal{B}^{\rm cyclo}_{\rm stat} \gtrsim 1000$) but also introduce biases in both the foreground and background parameters. This finding reinforces our earlier point regarding the realistic LMC dataset analysis: as the signal becomes more modulated, the importance of cyclostationarity increases for accurately resolving the background, leading to stronger evidence. However, it is important to note that, contrary to the realistic LMC analysis, the dataset is generated under ideal conditions.

5.2.5 Caveats and takeaways

We developed an analytic framework to exploit the cyclostationary properties of SGWB signals from unresolved WD binaries in the MW, its satellite galaxies, and nearby galaxies, enabling signal recovery and parameter estimation through time-dependent modulation induced by LISA orbital motion.

Main results. Satellite detectability depended critically on the interplay between spectral amplitude and sky location. Despite being astrophysically fainter, the SMC proved more readily detectable than the LMC due to its ecliptic latitude, which yielded stronger LISA modulation sensitivity. With perfect prior knowledge of satellite positions, we successfully constrained spectral parameters for SMC, LMC, Sagittarius, and Andromeda. In the presence of the dominant MW foreground, satellite detection became severely challenged, yielding primarily upper limits on amplitudes while slopes remained moderately constrained. For realistic LMC catalog data, both cyclostationary and stationary models recovered consistent spectral slopes.

Discovery potential. We demonstrated the potential for discovering hidden satellite galaxies through GW observations. By injecting an LMC-like signal near the Galactic center—a region obscured electromagnetically—we successfully disentangled it from both instrumental noise and the MW foreground using the cyclostationary framework. This finding illustrated how cyclostationarity provided essential degeneracy-breaking power as signals became increasingly modulated, offering a unique avenue for probing stellar populations inaccessible to electromagnetic astronomy.

Framework generality and limitations. Our method modeled time-dependent modulation analytically in the frequency domain via a band-diagonal covariance matrix structure arising from the Fourier decomposition of the autocorrelation function. While we adopted a simplified bivariate Gaussian sky distribution, the framework could naturally extend to more complex spatial distributions—e.g., Gaussian mixtures or physically motivated models discretized over numerical sky grids. However, capturing the full cyclostationary correlation structure required accumulating at least one full LISA orbit ($T = 1\,\mathrm{yr}$) to achieve sufficient frequency resolution for modeling off-diagonal covariance elements. This temporal requirement could have limited applicability in low-latency detection scenarios or during early-mission global fit analyses, where frequent noise updates with shorter in-

tegration times were critical. Section 5.3 explore quasi-stationary approximations suitable for sub-annual timescales.

5.3 Time-frequency analysis

The cyclostationary framework developed in Section 5.2 provides a method to account for non-stationarities arising from astrophysical anisotropies in LISA, fully formulated in the frequency domain through a band-diagonal covariance matrix structure. While we demonstrated the effectiveness of this approach for characterizing the MW and satellite galaxy backgrounds, we also emphasized a fundamental limitation: achieving sufficient frequency resolution to model the covariance matrix off-diagonal elements requires accumulating at least one year of LISA data. This requirement may limit applicability in global fit analyses or low-latency searches where frequent updates to the noise estimate are critical on timescales much shorter than the annual LISA orbital period. The GF estimate is influenced by two time-dependent effects that must be captured on sub-annual timescales. First, as LISA observes longer, more WDs binaries are individually resolved and subtracted from the overall confusion noise, inducing a spectral variation consistent with a gradual drift toward lower frequencies, as captured by the phenomenological template in Eq. 3.62. Second, the cyclostationary nature of the foreground discussed in Section 5.2 manifests as a time-dependent modulation on annual timescales. To better account for both effects simultaneously while enabling sub-annual inference, we extend our approach to the time-frequency domain using the Short Time Fourier Transforms (STFT). This representation allows us to treat the signal as locally stationary within short time segments, while capturing the long-term spectral evolution and modulation pattern across the full mission. We note that this representation is not unique; alternative time-frequency methods—such as those based on wavelets—have also been explored in the literature. In this section, we extend the likelihood framework introduced in Chapter 4 to the time-frequency domain by incorporating the cyclostationary nature of the MW foreground through a quasi-stationary approach. This is achieved by leveraging the modulation model developed in Sections 5.2 and 5.2.2 and applying it locally to each time segment. The resulting framework, implemented in the `bahamas` pipeline, provides a flexible and computationally efficient approach suitable for integration into global fit analyses where frequent noise updates with sub-annual resolution are essential. We demonstrate its effectiveness on both synthetic datasets and realistic simulations from the *Yorsh* data challenge [121].

5.3.1 STFT representation

To capture the time-varying frequency content of the cyclostationary GF signal , we divide the full data stream into $N_{\rm chunk}$ non-overlapping segments and perform a STFT on each. This time-frequency decomposition allows us to construct a likelihood that accounts for the local stationarity of the signal and noise within each chunk.

The total log-likelihood is expressed as the sum of the log-likelihoods for each chunk:

$$\log \mathcal{L}(\boldsymbol{d}|\theta) = \sum_{c=1}^{N_{\rm c}} \log \mathcal{L}^{\rm chunk}_{W,G}(\boldsymbol{d}^c|\theta), \tag{5.27}$$

where $\boldsymbol{d}^c$ is the data segment corresponding to the c-chunk, and $\mathcal{L}^{\rm chunk}_{W,G}$ represent the two likelihood models considered for each single chunk, conditioned on the full set of model parameters θ, and described below. Each chunk has a duration T, sampling interval dt, and sampling frequency $f_s = 1/dt$, so the number of points per chunk is $N = T/dt$. For a

given frequency range $[f_{\rm min}, f_{\rm max}]$, we define n_f as the total number of frequency bins. We adopt the following index notation:

- c — time chunk, $c = 1, \ldots, N_{\rm chunk}$
- j — TDI channel, $j \in A, E$
- k — frequency bin, $k = 1, \ldots, n_f$
- l — time sample within a chunk, $l = 1, \ldots, N$

In what follows, we will consider equal-duration chunks. However, our approach can be readily applied to unequal chunk-lengths, allowing to allocate time or frequency resolution where needed.

Full-resolution data. First, we consider each chunk at full frequency resolution, i.e., through the discrete FT of time-domain data, defined as

$$\tilde{d}^c_{jk} = \sqrt{\frac{2}{W f_s}} \sum_{l=1}^{N} w_l \, d^c_{jl} \, \exp\left[-\frac{2\pi\imath}{N} k \, l\right], \tag{5.28}$$

where w_l is a window function. The factor $W = \sum_{l=1}^{N} |w_l|^2$ accounts for the window normalization. In this work, we generate synthetic data directly in the frequency domain, so windowing effects do not arise during preprocessing. For normalization consistency, we choose $W = N$. However, in Section 5.3.5 we use time-domain, simulated data, hence we apply the Kaiser window with shape parameter equal to 30. Assuming perfect Gaussianity, the single-chunk *Whittle* log-likelihood reads

$$\log \mathcal{L}^{\rm chunk}_W(\vec{\boldsymbol{d}}^c|\theta) = -\frac{1}{2} \sum_{j\in\{\rm A,E\}} \sum_{k=1}^{n_f} \left(\frac{|\tilde{d}^c_{jk}|^2}{S^c_{jk}(\theta)} + \log S^c_{jk}(\theta) + \log 2\pi \right). \tag{5.29}$$

Here, the factor $2/(N f_s)$, which converts the discrete Fourier amplitudes into a one-sided power spectral density (PSD) with units of $\rm Hz^{-1}$ [195], is implicitly accounted for in the definition of d^c_{jk} in Eq. 5.28.

Average Periodograms. Alternatively, we consider data preprocessed into averaged periodograms

$$\hat{P}^c_{jm} = \frac{1}{n_{b_m}} \sum_{k=1}^{n_{b_m}} |\tilde{d}^c_{jk}|^2. \tag{5.30}$$

to reduce the data volume, and the likelihood computation time. Here, the initial n_f frequency values within each chunk c are compressed into n_g coarse-grained points, indexed by $m = 1, \ldots, n_g$. We adopt the following data compression scheme (where all frequencies are expressed in Hz):

$$\Delta = \frac{\log_{10}(f_{\rm max}) - \log_{10}(f_{\rm min})}{n_g}, \tag{5.31}$$

$$\log_{10} f_m = \log_{10} f_{\rm min} + m\Delta, \tag{5.32}$$

$$\mathcal{F}_k = \{m | f_m < f_k < f_{m+1}\}, \tag{5.33}$$

$$n_{b_m} = |\mathcal{F}_m|, \tag{5.34}$$

where Δ denotes the logarithmic-interval width, $\mathcal{F}_m$ is the set of frequencies falling within the m-th interval, and n_{b_m} its cardinality. In our analyses we choose $f_{\rm min} = 0.1$ mHz, $f_{\rm max} = 2.9$ mHz, and $n_g = 1000$.

Being $\hat{P}^c_{jm}$ the sum of uncorrelated, squared, circular Gaussian complex random variables, they follow a *Gamma* distribution. Accordingly, the single-chunk likelihood reads

$$\log \mathcal{L}_G^{\rm chunk}(\boldsymbol{d} = \hat{\boldsymbol{P}}^c|\theta) = -\sum_{j\in\{\rm A,E\}}\sum_{m=1}^{n_g}\left[\log\Gamma\left(\frac{n_{b_m}}{2}\right) + \frac{n_{b_m}}{2}\log S^c_{jm}(\boldsymbol{\theta}) + \right.$$
$$\left. - \left(\frac{n_{b_m}}{2} - 1\right)\log \hat{P}^c_{jm} + \frac{\hat{P}^c_{jm}}{S^c_{jm}(\boldsymbol{\theta})}\right]. \tag{5.35}$$

Power spectral density model. We model the PSD in each chunk as the sum of three components:

$$S^c_{jk}(\theta) = \overline{M}^c_j(\theta_{\rm GF_{II}})S^{\rm GF}_k(\theta_{\rm GF_I}) + S^{(n)}_{jk}(\theta_n) + R_{jk}S^{\rm EF}_k(\theta_{\rm EF}), \tag{5.36}$$

We adopt the phenomenological spectral model for the Galactic foreground $S^{\rm GF}_k(\theta_{\rm GF_I})$ proposed in [101], shown in Eq. 3.62. The spectral-tilt frequencies f_1 and $f_{\rm knee}$ are functions of the observation time, as defined in Eqs. 3.63 and 3.64. We simulate data assuming the following parameters: $\log_{10} A = -43.9$, $\alpha = 1.8$, $a_1 = -0.25$, $b_1 = -2.7$, $a_{\rm knee} = -0.27$, $b_{\rm knee} = -2.47$, $\log_{10} f_2 = -3.5$.

The factor $\overline{M}^c_j$ captures the time-averaged square of the LISA modulation function $M_j(t)$ during chunk i:

$$\overline{M}^c_j = \frac{1}{t^c_E - t^c_S}\int_{t^c_S}^{t^c_E} dt M^2_j(t|\theta_{\rm GF_{II}}), \tag{5.37}$$

where t^c_S and t^c_E denote the start and end times of the chunk. We distinguish between $\theta_{\rm GF_I}$ and $\theta_{\rm GF_{II}}$ to emphasize that the first describe the signal spectrum $\theta_{\rm GF_I} = \{\mathcal{A} = \log_{10} A, \log_{10} f_{\rm knee}, \log_{10} f_1, \log_{10} f_2, \}$, while the latter includes all parameters affecting the time-dependent LISA modulation. For a bivariate Gaussian source distribution over the sky, the parameter space is defined by $\theta_{\rm GF_{II}} = \{\sin\beta, \lambda, \sin\psi, \sigma^2_1, \sigma^2_2\}$. A detailed discussion of such parametrization is provided in Section 5.2.2. Here, $\sin\beta$ and λ describe the position of the distribution center in Ecliptic coordinates, whose principal axes are rotated by an angle ψ with respect to the Ecliptic latitude direction, and yield variances $\sigma^2_{1,2}$. In what follows, we inject a signal consistent with the following values: $\sin\beta = -0.096$, $\lambda = -1.62$rad, $\sin\psi = -0.83$, $\sigma^2_1 = 0.04{\rm rad}^2$, $\sigma^2_2 = 0.14{\rm rad}^2$.

The last term in Eq. 5.36 describes the SGWB of extragalactic origin considered in Section 5.3.4. We model it as stationary and isotropic signal by applying the chunk-independent response R_{jk} [32, 152], described in Section 3.2.3. We adopt, for simulation and inference, the template spectral model from [97, 180], which is consistent with a broken power law with an exponential cut-off at the highest frequencies, as defined in Eq. 3.65. In this work, we choose to infer on the parameters $\theta_{\rm EF} = \{\mathcal{A}_{\rm EF} = \log_{10} A_{\rm EF}, \gamma_1, \gamma_2\}$ and simulated data assuming $\log_{10} A_{\rm EF} = -10.76$, $\gamma_1 = 0.741$, $\gamma_2 = -0.255$. We emphasize that the cutoff observed in the Extragalactic Foreground (EF) component arises purely from physical considerations, rather than from DWD subtraction, by contrast with the GF case. Therefore, the spectral parameters are not a function of $T_{\rm obs}$. This cutoff is expected to occur at approximately 40.2mHz.

T_{chunk}	Inf.	Gaps	Lkl.	Sampler	Model	Ref.
2 wks	Seq.	✗	Whittle	HMC, NS	Quasi-stat. Stat.	Figs. (5.11), (5.13), and (E.2)
			Gamma	HMC, NS	Quasi-stat. Stat.	Figs. (5.13), (E.1), (E.2), and (E.4)
1 wk	Seq.	✓, sch.	Gamma	NS	Quasi-stat. Stat.	Fig. E.4
1 wk	Seq.	✓, unsch.	Gamma	NS	Quasi-stat. Stat.	Fig. E.4
2 wks	Diff.	✗	Gamma	NS	Quasi-stat.	Figs. (5.15) and (E.3)
2 wks	Seq.	✗	Gamma	NS	EGF + GF (cyclo) + noise GF (cyclo) + noise EGF + GF (stat) + noise	Figs. (5.16), (5.17), and (E.5)

Table 5.3: Summary of parameter estimation (PE) runs performed over the full 1-year dataset, including configurations with the extragalactic foreground.

Finally, the instrumental noise model adopted in this work accounts for the two dominant contributions that remain after TDI post-processing: OMS noise and TM noise. Both components are projected into the TDI channels through the corresponding transfer functions, following the implementation described in [155]. The PSD of these noise terms are given in Eqs. 3.33 and 3.34. Throughout this analysis, we inject fixed amplitude values for both components and do not introduce additional non-stationarity by varying them across segments.

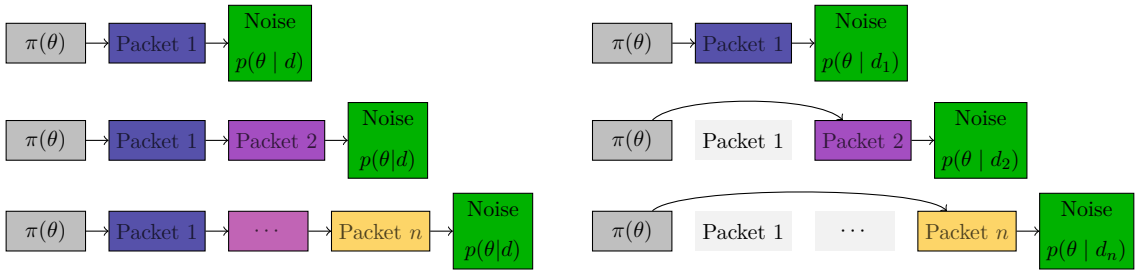

Figure 5.11: (*Left*) Sequential analysis scheme: data are analyzed packet by packet over the observation time. (*Right*) Differential analysis scheme: pairs of adjacent segments are analyzed independently.

In Fig. 5.12, we illustrate the evolution of the GF spectrum across different time chunks (top left panel) and as a function of the total observation time (top right panel), for seven reference times across a year. For reference, we underplot the instrumental noise and EF spectra, both assumed stationary. In addition, we show the modulation squared amplitude (bottom panel), highlighting the same reference times.

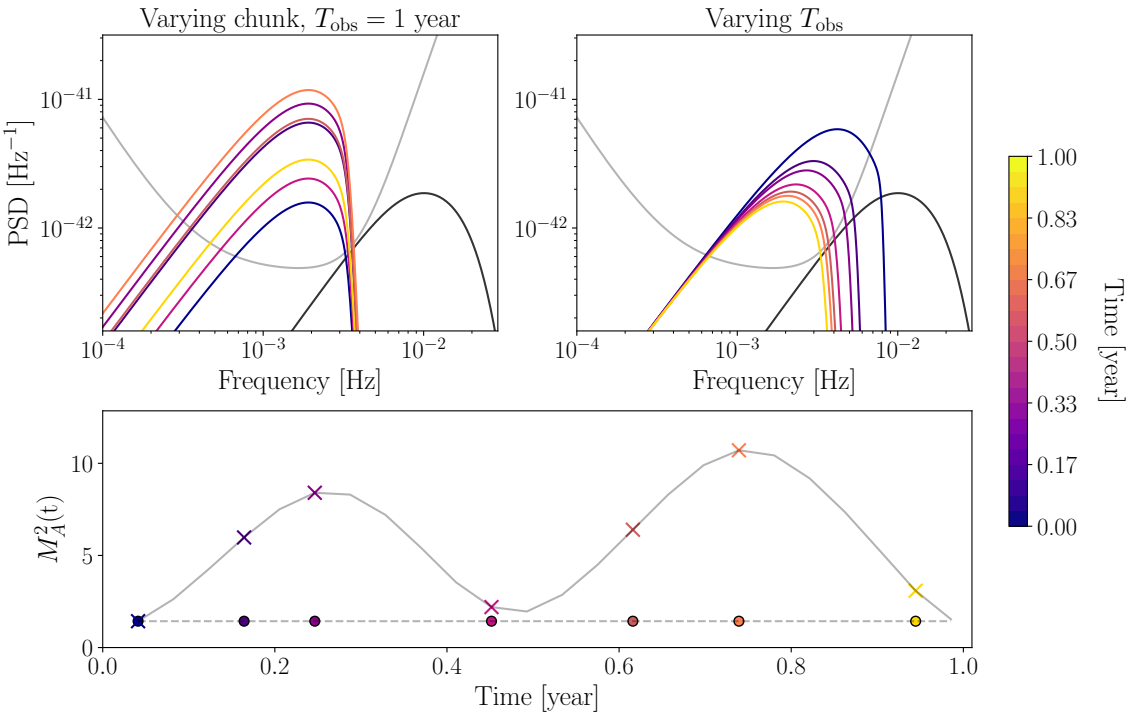

Figure 5.12: (*Top left*) Evolution of the Galactic foreground spectrum across different time chunks, with the corresponding modulation amplitudes shown in the bottom subplot as colored crosses. The spectrum level is obtained for a fixed total observation time $T_{\rm obs}$ = 1 yr. (*Top right*) Evolution of the spectrum as the observation time increases. By contrast to the left panel, the modulation amplitude is fixed to that of the first chunk, shown in the bottom subplot as a colored dark blue circle. The overall effect on the spectrum is a shift toward lower frequencies as $T_{\rm obs}$ increases. Instrumental noise and EF background are represented in gray and black, respectively. (*Bottom*) Time evolution of the modulation amplitude $M^2A(t)$ over one year, where the colored markers correspond to the time chunks used in the top panels. For simplicity, all quantities refer to the A channel.

In a global fit scheme, Galactic foreground estimates have to be continuously updated as the individual source-content evolves during Gibbs-sampling steps—sources are added, removed, or their parameters changed— and as more data are accumulated during the mission. In the absence of a full global fit pipeline, we use an iterative source subtraction scheme to approximate the GF evolution, following [101]. Therefore, we adopt the following procedure: we generate random Gaussian samples consistent with the PSD model defined in Eq. 3.62, and sequentially, at each iteration we increase the number of time segments, adjusting new segments spectra for their corresponding observation time, as described by Eqs. 3.63 and 3.64. We emphasize once again that the GF PSD level in each segment (or *packet*) accounts for the corresponding modulation factor, too. Therefore,

the GF PSD evolution reflects both the gradual subtraction of simulated individual DWD signals—resulting in a drift toward lower frequencies—and the modulation-induced variations, resulting in an up-and-down drift of the spectrum. This behavior is illustrated in Fig. 5.12. In the left panel, we show the evolution of the spectrum over a one-year observation period, highlighting the impact of modulation. By contrast, in the right panel, we illustrate the effect of binary subtraction at fixed modulation amplitude.

Code infrastructure. The model described above is implemented in `bahamas` code, flexible enough to support a variety of operational configurations [149].

First, parameter estimation can be performed using either NS or the NUTS, introduced in Section 2.2.2, as implemented in `nessai` [197] and `NumPyro` [144], respectively. While the former yields directly marginal-likelihood estimates, the latter is a more suitable candidate for deployment in a global-fit infrastructure. The typically large data volume at each likelihood evaluation makes HMC particularly appealing, as native support for automatic differentiation and accelerated hardware in `Numpyro` offsets dramatically the posterior exploration time. Second, `bahamas` hosts implementations of both the Whittle likelihood on STFT data and the Gamma likelihood on averaged power spectra, as described in Eqs. 5.29 and 5.35. Finally, `bahamas` can flexibly simulate and analyse segments of arbitrary, heterogeneous lengths. For simplicity, we will consider equal-length segments of 1 and 2 weeks, both yielding frequency content well below 0.1 mHz.

Our results are organized as follows. In Section 5.3.2 we first present inferences, on simulated GF and instrumental noise, only. We show the equivalence of the two likelihood models introduced in Eqs. 5.29 and 5.35, and quantify the evidence in favour of a cyclostationary GF model during the first year of LISA operations. We argue for the suitability of our approach, by showing the equivalence of posteriors obtained through nested sampling and HMC. Then, we investigate the impact of data gaps in Section 5.3.3, and of an additional EF component in Section 5.3.4. Finally, we apply our methodology to the more realistic *Yorsh* data challenge [121] in Section 5.3.5. A summary of inferences performed is provided in Table 5.3. Additional, auxiliary plots are shown in Appendix E. Overall, our model parameter space consists of 14 real numbers and we adopt uniform priors over $\theta_{\mathrm{GF_I}}$, $\theta_{\mathrm{GF_{II}}}$, θ_{EF}, and A, P.

5.3.2 Galactic Foreground across the first year

We consider a first analysis scheme mimicking a *global-fit-like* approach, as illustrated in the left panel of Fig. 5.11. Henceforth, in Table 5.3 and Fig. 5.11, we refer to it as the *sequential* analysis. At each iteration we produce posterior distributions for both the instrumental noise and the GF for increasingly longer, segmented datasets.

Posterior evolution and comparison. We first inject and recover the GF and LISA instrumental noise, only. Marginal posteriors for the former at each iteration are shown in Fig. E.1, where we highlight the progressive improvement in parameter estimation as more data are accumulated.

As expected, the Whittle and Gamma likelihood yield consistent posterior distributions [84]. To illustrate our finding, we compute the Jensen-Shannon Divergence (JSD) between the posterior distributions obtained from the two likelihoods. This metric quan-

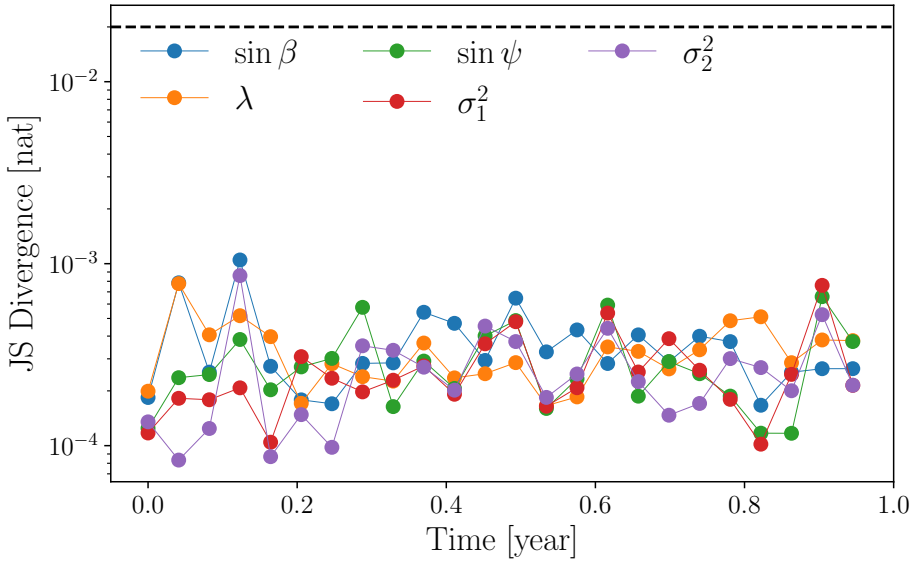

Figure 5.13: Jensen–Shannon divergence between the posterior distributions of the modulation parameters inferred from the Gamma and Whittle likelihoods, as a function of time. Each value remains below 0.02 nat (black dashed line), which is the typical threshold used to claim indistinguishability between two distributions.

tifies the similarity between two distributions, P and Q, as [119]:

$$\mathrm{JSD}(P||Q) = \tfrac{1}{2}D(P||M) + \tfrac{1}{2}D(Q||M), \tag{5.38}$$

$$D(P||M) = \int_X p(x) \log\left(\frac{p(x)}{m(x)}\right) dx, \tag{5.39}$$

where $M = \frac{1}{2}(P+Q)$. For clarity, we display in Fig. 5.13 only results on divergences for $\theta_{\mathrm{GF_{II}}}$, for each iteration of the sequential analysis. The observed values over time suggest that the posterior distributions derived from the two likelihoods are largely compatible. Likewise, different stochastic samplers yield very similar outcomes. We illustrate consistency across samplers and likelihoods in the full parameter space with a representative corner plot in Fig. E.2, obtained after 34 weeks of observation: we present three joint posterior distributions, obtained from NS and Whittle likelihood, NS with Gamma likelihood, and NUTS with Gamma likelihood. The agreement across the different methods is largely satisfactory over the full parameter space.

Yielding the same posteriors under the same prior assumptions, both likelihoods are expected to produce consistent values of marginal-likelihood. While NS algorithms naturally compute the evidence as part of their inference process [175], Monte Carlo Markov-chain methods such as NUTS do not provide direct estimates of it. Nevertheless, several approaches have been developed to infer the evidence from samples drawn from power posterior distributions, described in Section 2.2.3. We verify that GSS algorithm [201] returns evidence estimates consistent with those obtained via NS, with a relative discrepancy smaller 1% on $\log \mathcal{Z}$.

In Fig. 5.14, we show the performance gain achieved under different setups. Specifically,

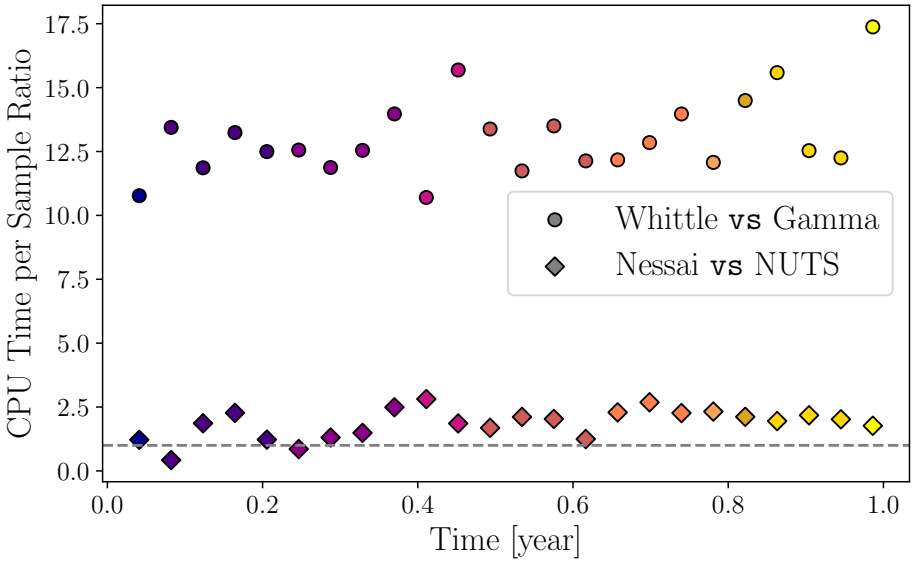

Figure 5.14: Time evolution of the gain factor across different setup configurations available in `bahamas`. Circles indicate the ratio of CPU time required to obtain one sample using the Whittle and Gamma likelihoods, showing that the latter is approximately 15 times faster. Diamonds represent the analogous comparison between NS and NUTS. Each marker is colored according to the total observation time considered, matching those in Fig. E.1.

posterior sampling with NUTS—fully relying on `jax` for reverse-mode differentiation—results in a speed-up factor of approximately 2.5. In this work, we do not exploit GPU acceleration (natively supported by `jax`), so an even larger gain is expected when doing so. Using the Gamma likelihood, a significant reduction in computational cost occurs. This depends on the exploitation of power spectra averages, which allows to reduce the number of computations from n_f to $n_g \ll n_f$. The adopted averaging and binning scheme introduced in Section 5.3 results in additional speed-up factor of about 15. Overall, a total computational gain of roughly a factor $\times 30$ is achieved. Based on our findings, we henceforth adopt the Gamma likelihood, as it is significantly faster to evaluate while providing unbiased results with respect to the full-frequency resolution Whittle.

Bayes Factor time evolution We now assess the statistical suitability of the quasi-stationary model, as compared to a simpler, stationary one. We compute the (log-)Bayes factor, defined as the (log-)ratio of marginal likelihoods between the quasi-stationary hypothesis and the stationary one

$$\log \mathcal{B}_S^{QS} = \log \mathcal{Z}_{QS} - \log \mathcal{Z}_S \,, \tag{5.40}$$

where the stationary assumption corresponds to setting $\overline{M}_j^i = 1$ in Eq. 3.62. We show in the left panel of Fig. 5.15 the evolution of $\mathcal{B}$ over time for both the Whittle and Gamma likelihoods. As anticipated in Section 5.3.2, the two approaches are equivalent. The evidence in favour of the quasi-stationary hypothesis increases significantly over time. After

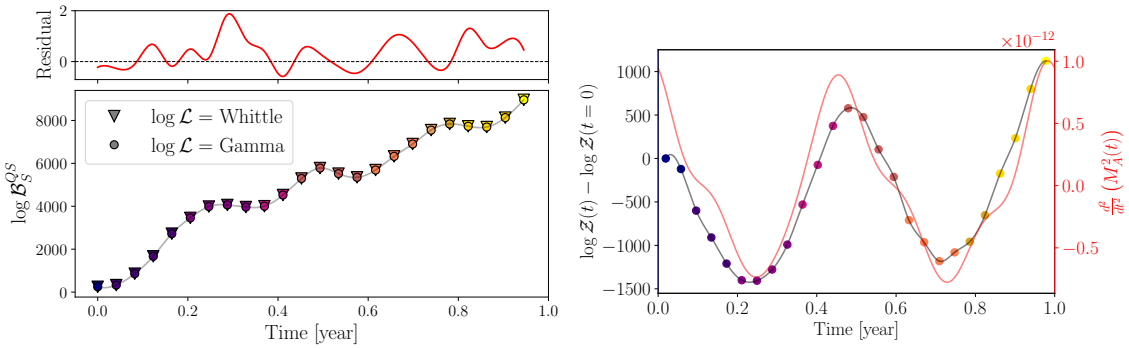

Figure 5.15: (*Left Panel*) Temporal evolution of the log Bayes factor between the quasi-stationary and stationary hypotheses for the Galactic foreground, obtained from a sequential analysis. The residual plot above shows the consistency between results from the Gamma and Whittle likelihoods. (*Right Panel*) Evolution of the evidence for the quasi-stationary model over time. Each dot corresponds to the evidence computed from a segment containing two consecutive weeks of data. The red and to the second time derivative of the modulation in the A channel. The plot highlights how the evidence closely follows the behavior of the second derivative of $M_A^2(t)$. Each marker is colored according to the total observation time considered, matching those in Fig. E.1

the first few weeks of data are accumulated, the quasi-stationary hypothesis is overwhelmingly favored. A more complex, oscillating and upwards drifting structure, emerges over longer $T_{\rm obs}$, likely tracking the shape of the Galactic modulation.

We investigate the observed trend, performing a 'differential' analysis (as referred to in Table 5.3 and Fig. 5.11, right panel), where the year of simulated data is divided into two weeks segments. Unlike the previous analysis, we now examine each pair of consecutive weeks independently using the quasi-stationary model. By tracking the single-segment evidence over time, we aim to identify which segment yields the largest constraining power, likely driven by the varying modulation pattern. The evolution of the model evidence over time is shown in the right panel of Fig. 5.15. Interestingly, the evidence does not track the modulation directly; rather, it appears to follow the second time derivative of the squared modulation amplitude M^2. The smooth, rapid variations in the modulation help break degeneracies between adjacent chunks, constrain the fundamental harmonics in the modulation model, allowing the QS model to better constrain the signal over time. This behavior has important implications for parameter reconstruction, which we illustrate in Fig. E.3, focusing on three distinct periods of increasing log-evidence throughout the year. We omit the reconstruction of $(\sin\psi, \sigma_1^2, \sigma_2^2)$, as these parameters are poorly constrained when inferred from two-weeks intervals, only. We observe that the reconstruction of the source sky distribution progressively and steadily improves. In particular, after 3 months, the inference is primarily driven by the spectral shape parameters, which are better constrained in the early stages as compared to $\sin\beta$ and λ. Conversely, the differential inference on the last segment yield tight constrains on the quasi-stationarity of the signal, while those on the spectral parameters are not significantly improved. As expected, an intermediate behavior is observed after a few months of data.

5.3.3 Impact of data gaps

We now quantify the effect of gapped datastream on the reconstruction of the Galactic foreground, as presented in Section 5.3.2. Periods of no usable data due to maintenance, antenna-repointing, or unforeseen exogenous events, gaps can be broadly classified either as *scheduled* or *unscheduled* ones.

In contrast to other studies [48], in `bahamas` we do not explicitly model for the presence of data gaps *in the data*. Instead, assuming the start and end times of each gap is known, we exclude the corresponding data segments in our inference.

To explore their effect on our results, we simulate both types of gaps, following the approach in [74]:

- we consider regular blind periods of either 3.5 or 7 hours, occurring every 1 or 2 weeks, respectively, as representative of *scheduled* gaps;

- we assume a fixed duration of 3 days per *unscheduled* gap, and model interval ΔT between consecutive gaps as an exponential distribution:

$$p(\Delta T \mid \lambda) = \lambda \exp\left[-\lambda \Delta T\right]. \tag{5.41}$$

  where the rate parameter λ is chosen to achieve an expect mission duty cycle of approximately 70%.

With this setup, we first analyze data corresponding to a total observation time of $T_{\rm obs} = 1$ year, divided into 2 weeks-long segments, and accounting only for scheduled gaps occurring every two weeks, each lasting for $T_{\rm gap} = 7$hr.

In a second, more realistic scenario, we include both scheduled and unscheduled gaps. The former occur weekly with $T_{\rm gap} = 3.5$hr, while realisations of unscheduled gaps are generated according to the distribution in Eq. 5.41. As a consequence, the data segments have variable durations due to the irregular occurrence of unscheduled gaps. We adopt a conservative approach and retain only chunks with an effective duration of $T_{\rm chunk} = 1$ week, discarding all other data. An example of such data segmentation is shown in the subplot of Fig. E.4. Therein, (i) we illustrate the chosen unscheduled gap realisation to achieve a target duty cycle close to 70% and (ii) compare posterior distributions obtained in absence of gaps, with scheduled gaps only, and with both scheduled and unscheduled gaps. For both gapped scenarios, their presence does not compromise the reconstruction of either the Galactic foreground or the instrumental noise. We highlight that $f_{\rm knee}$ and f_1 take different injected values compared to the no-gap case, since the presence of gaps reduces the effective observation time, leading to modified values in Eqs. 3.63 and 3.64. We highlight that for the instrumental noise model, we make the simplifying assumption of same amplitude level across before and after gaps. This is, of course, a simplification, as the LISA noise could vary following periods with gaps.

5.3.4 Extragalactic foreground

We now study how a putative additional EF in the datastreams influences the time–frequency reconstruction of the GF. As discussed in Section 5.3.1, we inject it as an additional stationary, isotropic, Gaussian noise component. The corresponding PSD model is given by Eq. 3.65. We perform inference using the sequential approach and two-weeks long segments. Specifically, we consider three alternative hypotheses:

$\mathrm{GF}_{\rm QS} \cup \mathrm{EF}$: we model consistently the stationary EF the quasi-stationary GF;

$\mathrm{GF}_{\rm S} \cup \mathrm{EF}$: we model both EF and GF as stationary.

$\mathrm{GF}_{\rm QS}$: we neglect the EF and model the GF as quasi-stationary.

For brevity, we omit from all three model names the instrumental noise, which is nonetheless jointly inferred upon. The time evolution of the reconstruction for the $\mathrm{GF}_{\rm QS} \cup \mathrm{EF}$ model is illustrated by the ridgeplot in Fig. E.5.

As discussed in Section 5.3.1, the cutoff frequency of the EF spectrum occurs at approximately 40.2mHz, whereas the Galactic knee, following the parametrization in Eq. 3.64, spans from about 10mHz down to 3.4mHz over the course of a year. This clear spectral separation keeps the two components distinguishable. For this reason, we do not further infer on $f_{\rm cut}$. The posterior on the EF becomes informative starting from packet 7, indicating that after roughly 3 months we may be able to detect it with large confidence. We leave a more detailed study of the EF detectability as a function of the underlying assumptions to future work.

We observe that the presence of an EF does not compromise the reconstruction of the GF, provided it is suitably incorporated into the data model. This is demonstrated in the left panel of Fig. 5.16: we show the relative fractional error of the reconstructed modulation derived from the posterior distribution across data packets, and averaged over them. When

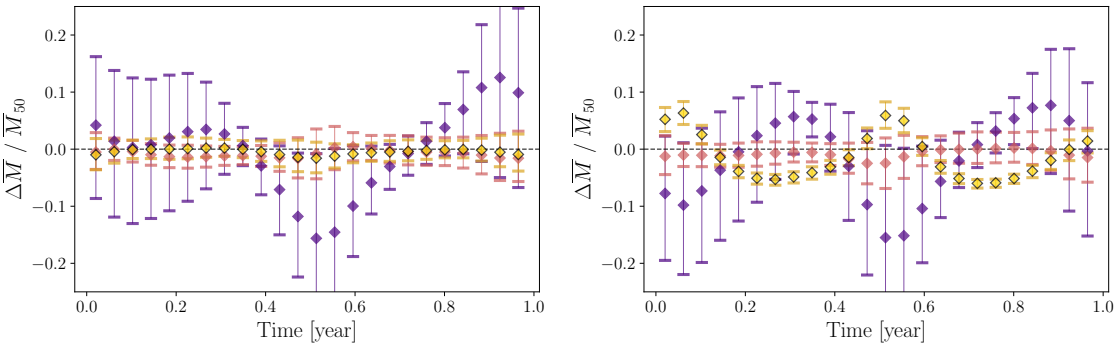

Figure 5.16: (*Left panel*) Reconstructed residual average modulation obtained from the posterior distributions of packets 5 (purple), 12 (salmon), and 24 (yellow), projected over one year. The posteriors are obtained including the EF in the model hypothesis. Diamonds and error bars represent the median and the 90% credible interval of the relative fractional error between the injected and reconstructed modulation, averaged over all packets analysed up to that time. (*Right panel*) Same as left panel, but with EF neglected in the model hypothesis.

only the first five packets (purple points) are analysed, the modulation amplitude posterior predictive distribution is both poorly constrained and biased. However, after roughly six months, an unbiased GF reconstruction emerges with nearly the same level of precision obtained after a full year of observations, namely at the percent level. Achieving such high precision in the measurement of the modulation is crucial for future prospects of constraining key parameters describing the Galactic geometry, such as the bulge and disk sizes.

Analogously, the right panel of Fig. 5.16 shows the same quantity obtained from inferences where the EF is neglected. As in the previous case, the modulation is initially poorly constrained. As more segments are collected, the reconstruction becomes nearly unbiased. However, as GF and EF spectra builds up signal-to-noise ratio over time, their cross-contamination becomes comparable to their respective posterior uncertainties. Therefore, after one year of observation, the modulation is severely biased if the EF is not accounted for. Our findings illustrate how cyclostationarity provides strong degeneracy breaking power, allowing the modulation—and thus potentially the structure of the Milky Way inferred from unresolved sources—to be characterized even after the first few months of observation. Nevertheless, considering predictions over a full year of data —thereby leveraging the complete information encoded in the LISA response to anisotropic background—is safer to achieve unbiased results, and fully breaking parameter degeneracies.

We conclude illustrating in Fig. 5.17 the Bayes factors between the three hypotheses introduced above. Circled and starred markers denote the log-evidence comparison when the GF cyclostationarity and the EF are ignored against the full model, respectively. Noteworthily, the log-evidence loss due to neglecting the GF quasi-stationarity is systematically larger over that of ignoring the EF. From a statistical perspective, it is therefore preferable to neglect the EF rather than sacrifice the GF quasi-stationarity. As shown by the

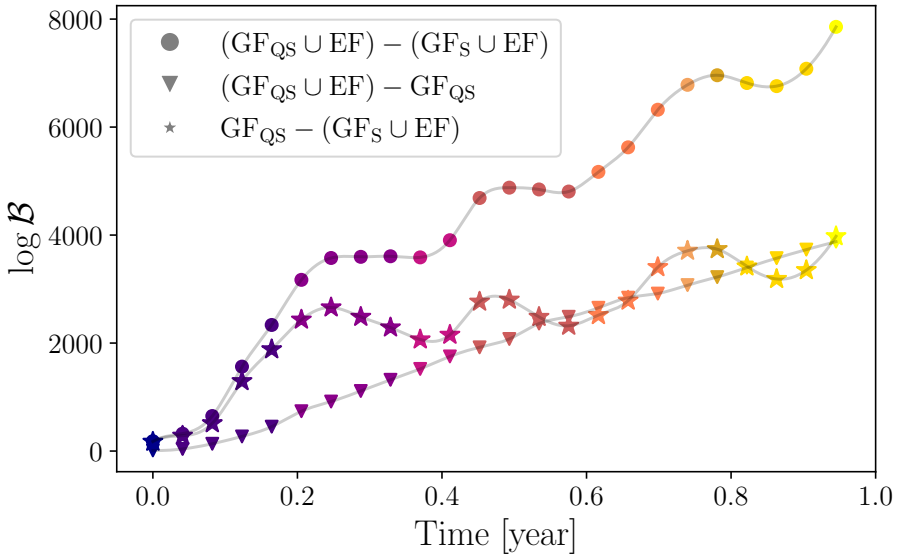

Figure 5.17: Evolution of the log Bayes factor between different hypotheses. The injected data include instrumental noise, GF, and EF. The blue line shows the $\log \mathcal{B}$ between models describing the GF as quasi-stationary or stationary, both including the EF. The orange line corresponds to the comparison where the EF is neglected in the second hypothesis, while the green line represents the remaining model comparison. All three hypotheses also account for instrumental noise. Each marker is colored according to the total observation time considered, matching those in Fig. E.1

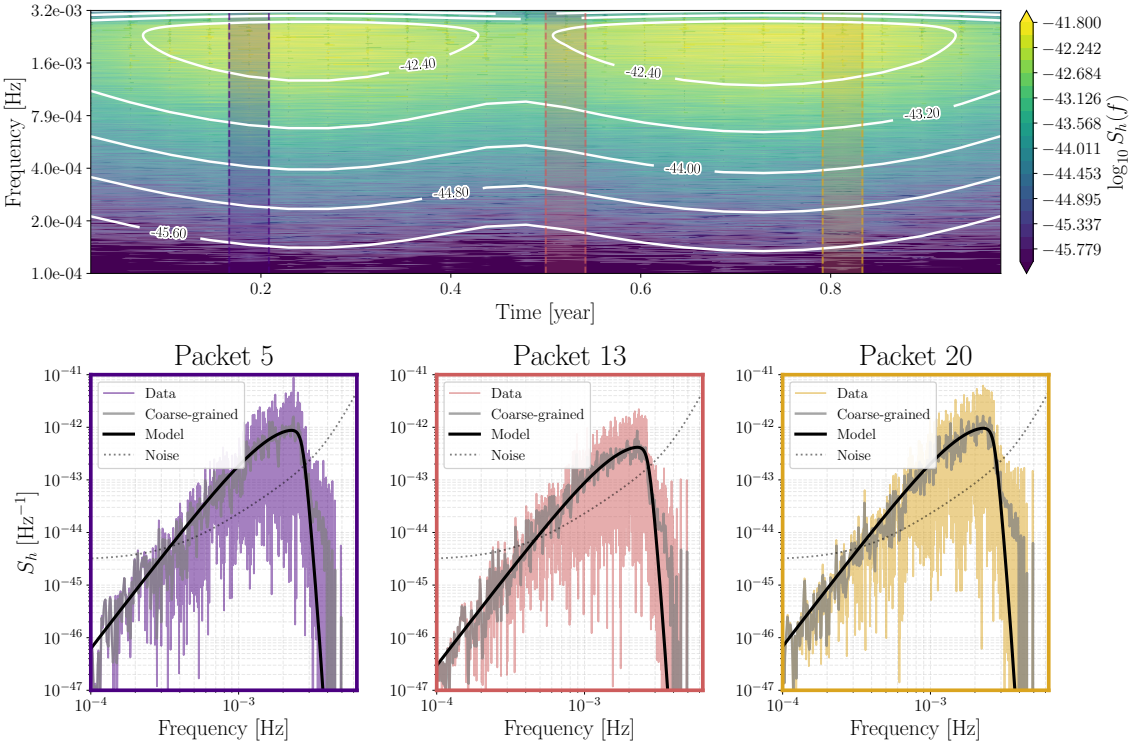

Figure 5.18: Spectrogram of the first year of the *Yorsh* 1b challenge, using only GF data. The white contours represent the median power reconstruction from the posterior distribution of the quasi-stationary model. The colored window highlights a zoomed-in segment, with the corresponding PSD shown in the bottom panels. Specifically, the solid black line represents the median posterior predictive of the model, while the colored lines indicate the PSD levels from the data within the selected segment. The gray lines correspond to the PSD obtained after coarse-graining the data.

triangular markers, the difference is linear in $T_{\rm obs}$, as both competing models account for the GF cyclostationarity.

5.3.5 Application to *Yorsh*

As a final validation step, we apply our framework to realistic observational data. Specifically, we analyze the Yorsh 1b dataset [121], which contains two years of time-domain data modelled generated as second-generation TDI variables (X, Y, Z), downsampled to a 5-second cadence. While this dataset was originally designed for testing the reconstruction of stellar-origin black hole binaries, its lego-like let us isolate data from instrumental noise and GF, and analyse them.

We show in Fig. 5.18 the GF reconstruction: in the top panel the GF data spectrogram is overlaid with contours of constant PSD, as posterior medians inferred by our model

over each segment. This visualization demonstrates our framework ability to recover the characteristic time-frequency structure of the galactic signal. In addition, we focus on the spectral reconstruction of posterior predictive PSDs from the quasi-stationary model for packets 5 (13, 20) after 10 (26, 40) weeks of observation. Although residual spurious power at frequencies above 3 mHz is not capture — likely originating from near-threshold unresolved DWDs— the PSD reconstruction remains at lower frequencies is robust and uncompromised, despite the increased complexity of the data with respect to our data model.

5.3.6 Caveats and takeaways

We extended SGWB inference to the time-frequency domain using STFT representations, enabling joint characterization of the evolving GF spectrum and its cyclostationary modulation on sub-annual timescales.

Main results. The modular `bahamas` pipeline provided a flexible framework for reconstructing the GF spectral evolution, driven by the progressive subtraction of WDs binaries, as well as its time-dependent modulation induced by LISA annual motion. After approximately six months of observation, modulation reconstruction reached percent-level precision, with posterior predictive distributions converging to unbiased estimates. Both the Whittle likelihood on full-resolution STFT data and the Gamma likelihood on coarse-grained periodograms produced statistically indistinguishable posteriors, although the latter offered a $\sim 15\times$ computational speed-up. Implementing automatic differentiation via `NumPyro` further accelerated sampling by a factor of ~ 2.5.

Model discrimination. The quasi-stationary hypothesis was overwhelmingly favored over the stationary model, with $\log \mathcal{B}_{\mathrm{S}}^{\mathrm{QS}}$ increasing systematically over time and exhibiting structure correlated with the second time derivative of the squared modulation amplitude. Differential analysis of two-week segments revealed that rapid, smooth modulation variations—rather than peak amplitudes—provided maximal constraining power by breaking degeneracies between adjacent chunks. When the EF component was present, spectral separation between the GF knee frequency and the EF cutoff enabled joint reconstruction. Critically, neglecting GF cyclostationarity incurred larger evidence penalties than neglecting the EF itself, establishing cyclostationarity as the statistically dominant feature. However, failure to model the EF introduced systematic biases in the modulation reconstruction after one year of observation, emphasizing the necessity of complete models for unbiased long-term inference.

Robustness and practical considerations. The framework accommodated both scheduled and unscheduled data gaps by excluding affected segments rather than explicitly modeling gaps in the likelihood. Simulations incorporating realistic duty cycles ($\sim 70\%$) demonstrated robust parameter recovery without significant bias. Application to the *Yorsh* 1b challenge confirmed the method's effectiveness on a realistic dataset, successfully recovering the time-frequency structure of the GF despite residual power from near-threshold unresolved DWDs above 3 mHz.

Limitations and future directions. Our analysis assumed perfect subtraction of resolvable sources and neglected instrumental noise non-stationarity across data gaps. These simplifications would need to be relaxed for realistic global fit integration. The bivariate Gaussian sky distribution, while analytically tractable, provided only a first-order approximation to the MW complex morphology; future work should employ Gaussian mixture

models or numerical catalogs to capture finer structural details, as already anticipated in appendix C.2.

Chapter 6

Conclusion and Outlook

In this thesis, we explored the landscape of GW signals in the millihertz regime, as observed by LISA. In particular, we focused on the characterization of SGWB arising from unresolved astrophysical populations. From a data analysis perspective, the detection of such signals presents numerous challenges. Throughout this thesis, we tackled several of them.

Both SGWB signal and instrumental noise suffer from significant uncertainties, due to our limited physical understanding and prior predictive ability on both. This translates into poor constraints on stochastic components in LISA datastream, to be suitably propagated onto Bayesian inference methods. In Sections 4.1 and 4.2, we characterized the SGWB from EMRIs and SOBBHs, respectively. For EMRIs, we explored a range of population scenarios, while for SOBBHs, we incorporated the effects of environmental interactions—specifically, dynamical friction and accretion—on the background, informed by current constraints from the LVK. In both cases, we observed that different modeling assumptions lead to significantly different amplitudes and spectral shapes for the reconstructed SGWB.

Focusing on the SOBBH signal, we investigated the detectability of environmental effects using parametric models. Although non-vacuum effects are more significant in frequency bands where instrumental noise exceeds the signal by several orders of magnitude, employing parametric templates enhances both sensitivity and information extraction. However, we also find that parameter estimation may be biased due to the imperfect nature of the adopted models.

To mitigate the limitations arising from incomplete physical knowledge in parametric templates, in Section 4.3, we introduced a more flexible, weakly parametric approach known as the EGP. This method proves highly effective in exploring a wide variety of spectral shapes. Nonetheless, when applied to the detection of the EMRI background, the benefits of this flexibility are diminished due to the Occam factor: in low-SNR regimes, the method favors simpler models that better balance fit quality and model complexity, thus limiting its ability to fully explore more intricate spectral features.

While the EGP approach enhances model flexibility, it primarily addresses limitations in spectral modeling. To further account for uncertainties in both signal and noise characterization, we further developed in Section 4.4 a Bayesian framework to systematically assess SGWB detectability while marginalizing over uncertainties in both signal and noise

properties. We derived semi-analytic expressions for the Bayes factor between signal-present and signal-absent hypotheses, enabling efficient construction of BPLScurves that generalize the standard SNR-based PLS.

In the current global fit pipelines, Bayesian inference is performed under idealized assumptions about the statistical properties of stochastic components— Gaussianity being ubiquitous. However, this assumption may not hold in the case of astrophysical backgrounds, where the number of unresolved overlapping sources may be insufficient for the central limit theorem to hold. Using an incorrect statistical model (i.e. inappropriate likelihood function), can introduce significant biases in the inference process.

In Section 5.1, we investigated the presence and impact of non-Gaussianity in the specific case of an EMRI-generated background. Despite the large uncertainties associated with the EMRI population and its overall amplitude, we found that non-Gaussian features can be non-negligible in a variety of scenarios. Moreover, although not explicitly discussed in our study, we expect that non-Gaussianity will naturally emerge in a global-fit context due to residuals left from the subtraction of other signals.

Our findings motivate the need to go beyond the traditional Gaussian likelihood and develop more general statistical frameworks. In future work, we aim to explore methodologies inspired by the Cosmic Microwave Background analysis, which offer effective tools for capturing non-Gaussian signatures in cosmological data.

Anisotropies are expected to arise from unresolved sources concentrated in well-defined regions of the sky, such as the MW and its satellite galaxies. When computing the overall response to such anisotropic distributions, the periodic motion of the LISA constellation over the course of a year induces a cyclostationary signature in the SGWB data. This effect was addressed in Section 5.2, where we extended the `balrog` framework to include a frequency-domain prescription for cyclostationary signals. Accounting for anisotropies is not only crucial to avoid biases in signal reconstruction, but also potentially beneficial, as it may allow to uncover previously unknown structures in the sky, as found in Section 5.2.4.

In this work, we implemented a method to compute the LISA response to an anisotropic distribution using a simplified sky distribution. In the future, we aim to extend this methodology incorporating more complex and realistic sky distributions, as partially demonstrated in Appendix C.2.

Finally, we note that this thesis does not explore anisotropies arising from the cosmological redshift dependence of the compact object (CO) population. According to the current literature, such effects are not expected to significantly impact LISA observations, but may become relevant in future investigations as our understanding and sensitivity improve.

Both the stochastic signal and the instrumental noise in LISA are expected to exhibit non-stationary features. From an astrophysical perspective, we have already highlighted the cyclostationary nature of the GF and potential contributions from MW satellites. A similar effect is likely to manifest in the instrumental noise as well. In realistic scenarios, the LISA arm lengths are not constant over time, leading to time-varying delays and modulations. This orbital motion induces a nonstationary imprint on the noise, similar to those observed in the Galactic foreground.

As discussed in Section 3.3, the noise associated with individual links is expected to be unequal and may vary over the course of the mission. Additional sources of transient non-stationarity include mission-related artifacts such as glitches and data gaps.

The time-frequency domain provides a natural framework to analyze locally-stationary processes. For this reason, in Section 5.3, we have extended the GF analysis within the

`bahamas` pipeline to represent data in terms of STFT. In future developments, we plan to incorporate more advanced methods into the pipeline to address the various sources of non-stationarity outlined above.

Challenges in SGWB and noise characterization are numerous and must be addressed to achieve a successful LISA global fit. In this thesis, we introduced tools, methodologies, and codebases laying the foundations for future developments.

Appendix A

Merger Rate and Population Model

The local merger rate of BBH near redshift $z = 0$ is relatively well constrained [4]. However, its evolution at higher redshifts remains highly uncertain [3, 4]. We assume that the merger rate density of GW sources, $R_{\rm GW}(z)$, as being proportional to the cosmic Star Formation Rate (SFR) $R_{\rm SFR}$ [125], expressed as

$$R_{\rm SFR}(z) = \frac{R_0}{\mathcal{C}} \frac{(1+z)^{\lambda_1}}{1+\left(\frac{1+z}{1+z_p}\right)^{\lambda_1+\lambda_2}}. \tag{A.1}$$

Here, R_0 denotes the local rate of binary systems at redshift $z = 0$, while $\mathcal{C}$ is a normalization constant ensuring $R_{\rm GW}(0) = R_0$.

As SOBBH formation is expected to be more efficient in low-metallicity environments, we weigh the SFR by the fraction of stars formed with metallicity below a critical threshold, $F\left(Z < Z_{\rm thresh}, z\right)$. This fraction follows the fitting formula of Ref. [116], and we adopt a more stringent cutoff $Z_{\rm thresh} = 0.1 Z_\odot$ [65, 131]. Moreover, black holes are expected to undergo a range of evolutionary time delays between formation and their eventual binary mergers. We assume these time delays follow a log-uniform distribution, $p\left(t_d\right) \sim t_d^{-1}$, within the range 10 Myr $\leq t_d \leq$ 13.5 Gyr [57].

The unnormalized merger rate is then obtained by convolving this distribution with the metallicity-weighted star formation rate:

$$\tilde{R}_{\rm GW}(z) = \int dt_d R_{\rm SFR}\left(z_f\right) F\left(Z < Z_{\rm thresh}, z_f\right) p\left(t_d\right), \tag{A.2}$$

where $z_f \equiv z_f(z, t_d)$ denotes the redshift corresponding to the formation time of the BBH. Finally, we normalize the merger rate as

$$R_{\rm GW}(z) = R_0 \frac{\tilde{R}_{\rm GW}(z)}{\tilde{R}_{\rm GW}(0)}. \tag{A.3}$$

Recent observations indicate that the SOBBH population exhibits low effective spins [2, 135]. Consequently, when computing the SGWB, we assume that SOBBH have negligible

spin and focus solely on the mass distribution, which is described by the Power Law + Peak mode [182]. Under this assumption, the source parameters $\boldsymbol{\phi}$ in Eq. 1.85 contain the primary mass m_1 and mass ratio q, and the distribution $p(\boldsymbol{\phi})$ reduces to joint probability density function $p(m_1, q|\ \Lambda_m)$, where $\Lambda_m = \{\alpha_{\rm pop}, \beta_{\rm pop}, m_{\rm min}, m_{\rm max}, \delta_{\rm m}, \lambda_{\rm peak}, \mu_{\rm m}, \sigma_{\rm m}\}$ represents the set of hyperparameters governing the mass distribution.

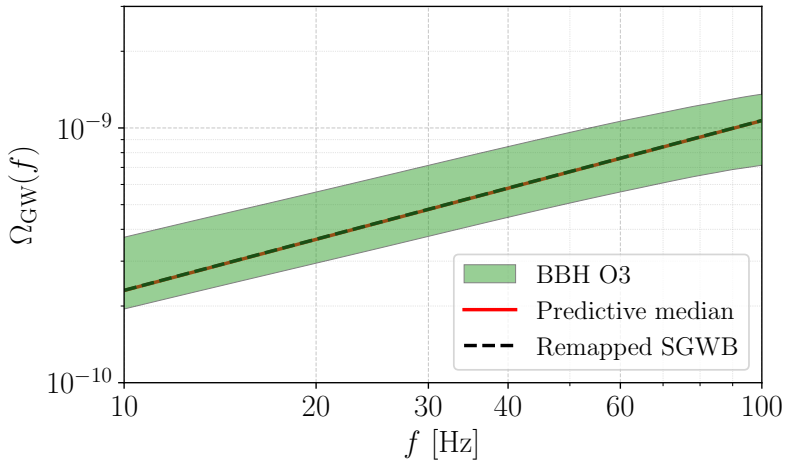

Figure A.1: The predictive median SGWB spectrum generated by our adopted model of SOBBH mergers (red solid) and the remapped spectrum based on fiducial parameter values (black dotted), compared with the predicted 90% credible spectrum band (green) from the O3 observational data [4] in the frequency band of 10-100 Hz.

We numerically evaluate Eq. 1.85 via Monte Carlo integration, using the corresponding posterior distributions of merger rate and mass distribution models inferred from [3, 4] and generate posterior predictions for the SGWB. We obtain the predictive posterior median by evaluating, at each frequency, the median of the SGWB realizations from the posterior samples. By computing the ℓ^2-norm between each sampled SGWB spectrum and the median curve, we select the sample that minimizes this distance and the corresponding parameters are adopted as fiducial values in this work. The merger rate parameters are taken as $R_0 = 21\ {\rm Gpc}^{-3}{\rm yr}^{-1}$, $\lambda_1 = 1.5$, $\lambda_2 = 3.7$, and $z_p = 3.0$. The hyperparameters of mass distribution are $\alpha_{\rm pop} = 3.6$, $\beta_{\rm pop} = 3.4$, $m_{\rm min} = 4.6M_\odot$, $m_{\rm max} = 84M_\odot$, $\delta_{\rm m} = 4.5M_\odot$, $\lambda_{\rm peak} = 0.018$, $\mu_{\rm m} = 29, M_\odot$, and $\sigma_{\rm m} = 8.3, M_\odot$. In Fig. A.1, we demonstrate the consistency between the median predicted SGWB curve generated by our adopted model and the forecast based on the O3 observational data [4]. Further, we observe that the spectral indices match well at low frequencies corresponding to the inspiral regime. However, at higher frequencies, higher order PN effects become relevant as the spectrum transitions toward merger and ringdown regimes. Since we use a leading PN approximation to compute the energy spectrum given by Eq. 1.86, we find that the spectral index is overestimated relative to the O3 inference that uses the higher PN terms. In the lower-frequency bands relevant to LISA, which is our primary focus, the energy spectrum of SOBBHs is well described by the leading PN expression of Eq. 1.86, and thus we are justified in neglecting higher PN terms for this work.

Appendix B

Corner from EGP inference

This appendix presents the full posterior parameter distributions obtained from the EGP inference analyses. Figure B.1 corresponds to the toy-model validation case presented in Sec. 4.3.2, demonstrating the recovery of injected parameters under idealized conditions. Figures B.2 and B.3 display the posteriors for the M1 EMRI SGWB reconstructions shown in Fig. 4.11, corresponding to different choices of the noise model complexity parameter m_σ. These distributions complement the spectral reconstructions by revealing the underlying parameter space structure and the degree of correlation between signal and noise parameters.

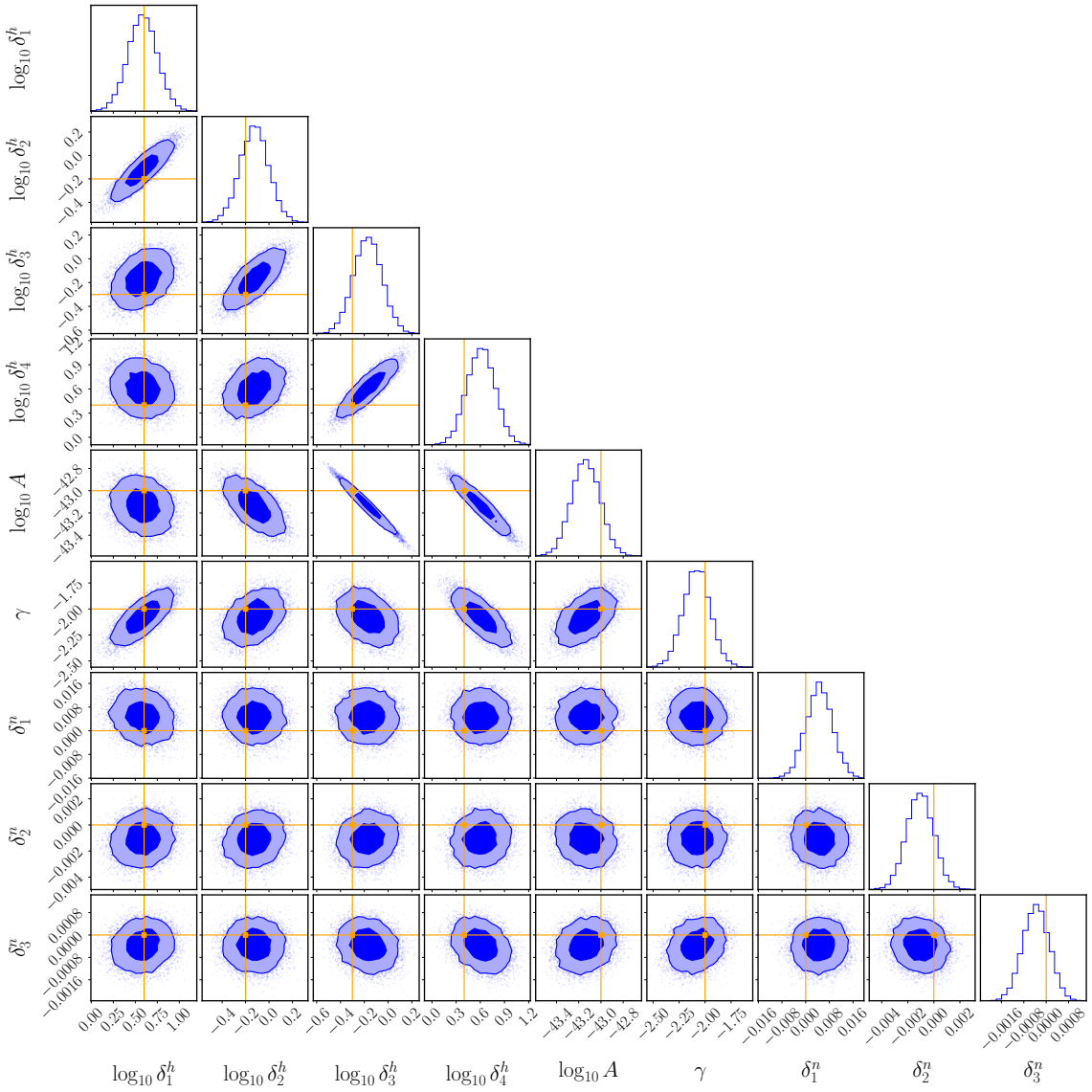

Figure B.1: Posterior parameter distribution for the toy-model SGWB with SNR=209 described in Sec. 4.3.2, obtained with an EGP model inference. The signal injected parameters are $\left\{\log_{10}\boldsymbol{\delta}^h, \log_{10}A, \gamma\right\} = \{[0.6, -0.2, -0.3, 0.4], -43, -2\}$. The injected noise level matches the nominal LISA sensitivity curve [122]. Darker (lighter) shaded areas denote 90% (50%) credible regions, and solid orange lines indicate the injected values. Posteriors are consistent with the injected values at the 90% credible level, and the noise parameters are consistent with zero with accuracies between 10^{-3} and 10^{-2}.

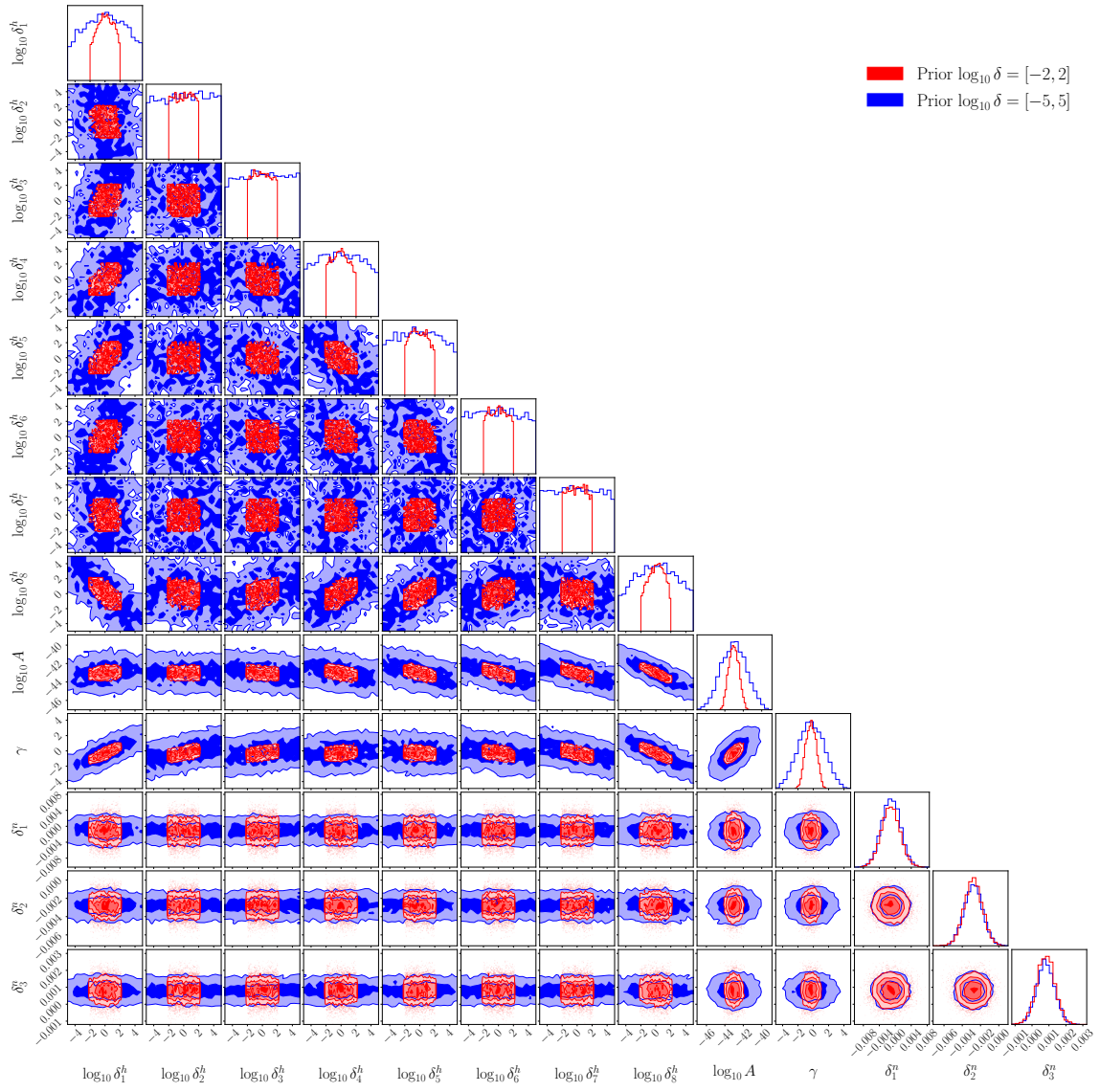

Figure B.2: Posterior parameter distribution for M1 EMRI SGWB for model $(n, m_\sigma) = (8, 32)$ using a larger and narrower prior on $\boldsymbol{\delta}^h$. Posteriors shown in this plot correspond to the spectral reconstructions in left panels of Fig. 4.11. They support the result interpretation provided in Sec. 4.3.3.

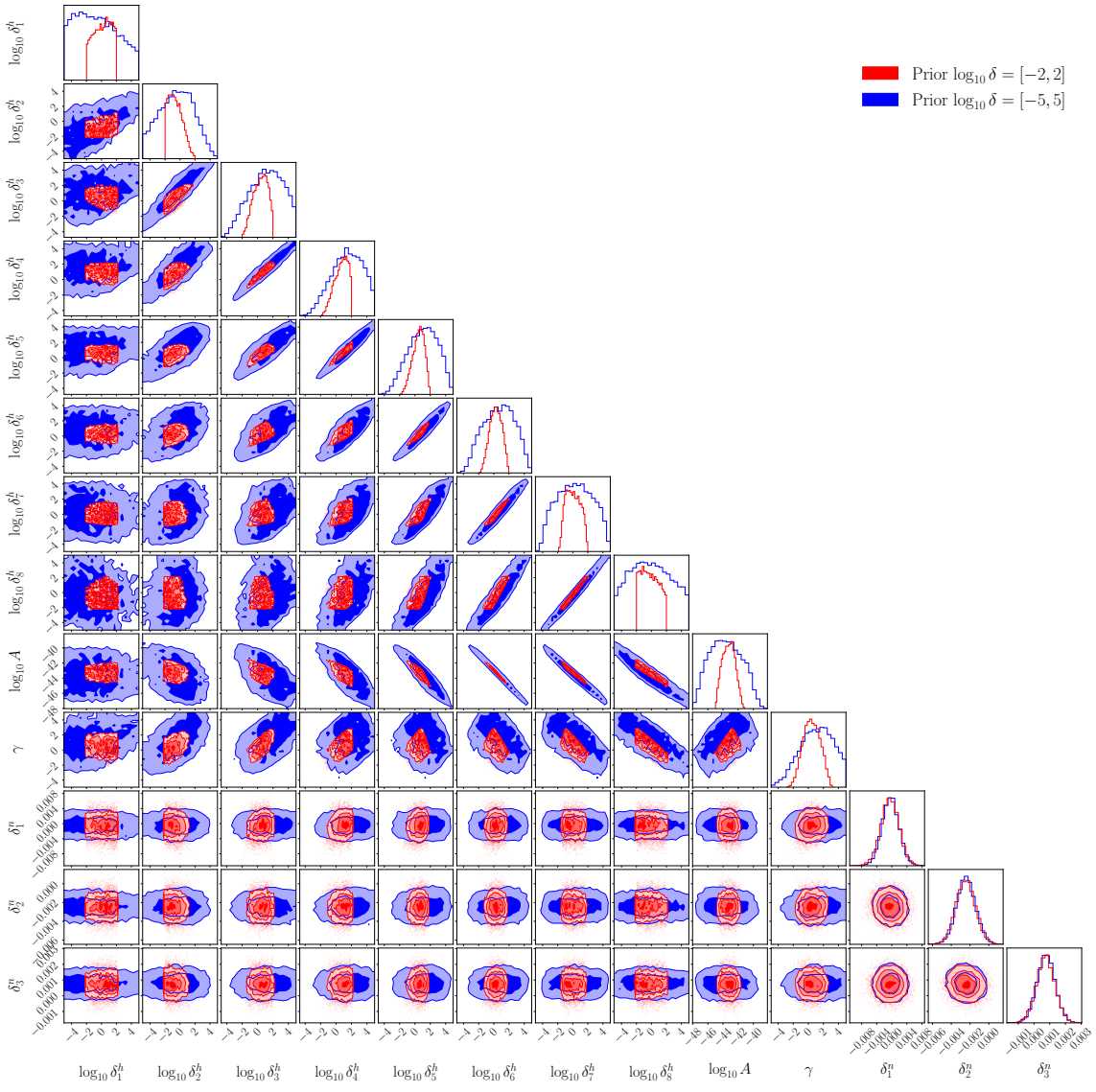

Figure B.3: Posterior parameter distribution for M1 EMRI SGWB for model $(n, m_\sigma) = (8, 1)$ using a larger and narrower prior on $\boldsymbol{\delta}^h$. Posteriors shown in this plot correspond to the spectral reconstructions in right panels of Fig. 4.11. They support the result interpretation provided in Sec. 4.3.3.

Appendix C

Stochastic signals envelopes

In order to derive the envelope of the LISA response to unresolved Galactic sources, we start by writing the LISA response to a single binary source in the low-frequency approximation. The spacecraft orbits can be described by the following Keplerian orbits:

$$P_j = r_j \begin{pmatrix} \cos i \cos\beta_j \cos v_j - \sin\beta_j \sin v_j \\ \cos i \sin\beta_j \cos v_j + \cos\beta_j \sin v_j \\ -\sin i \cos v_j \end{pmatrix}, \tag{C.1}$$

$$r_j = \frac{R\,(1-e^2)}{1+e\cos v_j}, \tag{C.2}$$

$$\beta_j = \frac{2\pi j}{3} + \beta_0, \tag{C.3}$$

where P_j is the position of spacraft $j \in \{1,2,3\}$ with semimajor axis $R = 1\mathrm{AU}$, true anomaly v_j, inclination i, and eccentricity e.

The mean anomaly l_j increases linearly with time t, and is related to the true anomaly by

$$l_j = \Omega t + \alpha_0 - \beta_j = u_j - e\sin u_j, \tag{C.4}$$

$$\tan\frac{u_j}{2} = \sqrt{\frac{1-e}{1+e}}\tan\frac{v_j}{2}, \tag{C.5}$$

with mean orbital angular frequency $\Omega = 2\pi/\mathrm{yr}$, mean anomaly l_j, and eccentric anomaly u_j.

In the following, we will use $\alpha_0 = \beta_0 = 0$ to simplify the derivation. The signal in the low frequency approximation from a binary with sky angles (β, λ), with an arbitrary α_0 and β_0, can then be computed with

$$\begin{aligned} h_A(\beta,\lambda) &= \cos 2(\beta_0-\alpha_0) h_{A,0}(\beta, \lambda-\alpha_0) \\ &\quad - \sin 2(\beta_0-\alpha_0) h_{E,0}(\beta,\lambda-\alpha_0), \end{aligned} \tag{C.6}$$

$$\begin{aligned} h_E(\beta,\lambda) &= \sin 2(\beta_0-\alpha_0) h_{A,0}(\beta, \lambda-\alpha_0) \\ &\quad + \cos 2(\beta_0-\alpha_0) h_{E,0}(\beta,\lambda-\alpha_0), \end{aligned} \tag{C.7}$$

where $h_{A,0}$ and $h_{E,0}$ have been computed assuming $\alpha_0 = \beta_0 = 0$.

C.1 Response to an individual binary

We assume that the GW signal is described in the SSB frame by

$$h_{ab}(t) = -\frac{A}{2}\left(1+\cos^2\iota\right) e^{+}_{ab}\cos\phi(t) + A\cos\iota\; e^{\times}_{ab}\sin\phi(t), \tag{C.8}$$

where e^{p}_{ab} are the polarization tensors defined in Eq. 1.33, A is the signal amplitude given in Eq. 1.61, $\phi(t)$ is the phase evolution defined in Eq. 1.62, and ι is the inclination angle of the binary.

In the low-frequency approximation, the LISA response to this wave can be modelled by two noise-independent GW detectors with signals h_A and h_E described by

$$h_A(t) = \frac{1}{\sqrt{2}}\left[h_Z(t) - h_X(t)\right], \tag{C.9}$$

$$h_E(t) = \frac{1}{\sqrt{6}}\left[h_X(t) - 2h_Y(t) + h_Z(t)\right], \tag{C.10}$$

$$h_X(t) = \left[\hat{L}^a_3(t)\hat{L}^b_3(t) - \hat{L}^a_2(t)\hat{L}^b_2(t)\right] h_{ab}(u), \tag{C.11}$$

$$h_Y(t) = \left[\hat{L}^a_1(t)\hat{L}^b_1(t) - \hat{L}^a_3(t)\hat{L}^b_3(t)\right] h_{ab}(u), \tag{C.12}$$

$$h_Z(t) = \left[\hat{L}^a_2(t)\hat{L}^b_2(t) - \hat{L}^a_1(t)\hat{L}^b_1(t)\right] h_{ab}(u), \tag{C.13}$$

$$u = t - \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{P}}(t), \tag{C.14}$$

$$\hat{\boldsymbol{P}}(t) = R\left(\cos\Omega t, \sin\Omega t, 0\right), \tag{C.15}$$

where $R = 1$ AU is the distance from the Sun to the detector barycenter, $\Omega = 2\pi/\mathrm{yr}$ is the detector barycenter angular orbital frequency, and $\hat{\boldsymbol{k}}$ is given by Eq. 3.4 The directions of the LISA arms $\hat{\boldsymbol{L}}_{\boldsymbol{i}}(t)$ can be described by

$$\hat{\boldsymbol{L}}_{\boldsymbol{1}} = -\frac{1}{2}\hat{\boldsymbol{x}} - \frac{\sqrt{3}}{2}\hat{\boldsymbol{y}}, \tag{C.16}$$

$$\hat{\boldsymbol{L}}_{\boldsymbol{2}} = \hat{\boldsymbol{x}}, \tag{C.17}$$

$$\hat{\boldsymbol{L}}_{\boldsymbol{3}} = -\frac{1}{2}\hat{\boldsymbol{x}} + \frac{\sqrt{3}}{2}\hat{\boldsymbol{y}}, \tag{C.18}$$

where $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ are part of a triad tied to the detector arms together with $\hat{\boldsymbol{z}}$, which are expressed in a fixed ecliptic frame by

$$\hat{\boldsymbol{x}} = \left[\frac{1}{4}(3-\cos 2\Omega t), -\frac{1}{4}\sin 2\Omega t, \frac{\sqrt{3}}{2}\cos\Omega t\right], \tag{C.19}$$

$$\hat{\boldsymbol{y}} = \left[-\frac{1}{4}\sin 2\Omega t, \frac{1}{4}(3+\cos 2\Omega t), \frac{\sqrt{3}}{2}\sin\Omega t\right], \tag{C.20}$$

$$\hat{\boldsymbol{z}} = \left[-\frac{\sqrt{3}}{2}\cos\Omega t, -\frac{\sqrt{3}}{2}\sin\Omega t, \frac{1}{2}\right]. \tag{C.21}$$

In order to model the envelope of a signal constituting of the sum of a large number of signals, we start by computing the inclination and polarization averages of the signal from an individual source. The structure of h_A and h_E can be expressed by setting $A = 1$,

$$h_{A,E}(t) = 2C_{A,E}(t)\left(1+\cos^2\iota\right)\cos 2\phi(t) + 4S_{A,E}(t)\cos\iota\;\sin 2\phi(t). \tag{C.22}$$

This allows us to compute the inclination, polarization, and orbital averages

$$\langle h_{A,E}(\lambda,\beta)\rangle^2_{(\iota)} = \frac{1}{2\pi}\int_0^\pi d\psi \int_{-1}^1 d\cos\iota\ \frac{1}{T}\int_0^T dt h_{A,E}(t)^2 \tag{C.23}$$

$$= \sum_{n=-8}^{8}\Big[h^2_{(c;A,E;\iota,\psi;n)}\cos n\Omega t + h^2_{(s;A,E;\iota,\psi;n)}\sin n\Omega t\Big]. \tag{C.24}$$

We can write the result using

$$h_S^2 = \langle h_A\rangle^2_{(\iota,\psi)} + \langle h_E\rangle^2_{(\iota,\psi)} \tag{C.25}$$

$$= \sum_{n=0}^{4} h^2_{S;(n)}(\beta)\cos(n\Omega t - n\lambda), \tag{C.26}$$

$$h_D^2 = \langle h_A\rangle^2_{(\iota,\psi)} - \langle h_E\rangle^2_{(\iota,\psi)} \tag{C.27}$$

$$= \sum_{n=0}^{8} h^2_{D;(n)}(\beta)\cos\left(n\Omega t - n\lambda + 4(\lambda + \frac{\pi}{12})\right), \tag{C.28}$$

with coefficients given by

$$h^2_{S;(0)} = \frac{9}{320}\left(328 + 152\cos^2\beta - 37\cos^4\beta\right), \tag{C.29}$$

$$h^2_{S;(1)} = -\frac{9\sqrt{3}}{40}\cos\beta\sin\beta\left(52 + 5\cos^2\beta\right), \tag{C.30}$$

$$h^2_{S;(2)} = \frac{81}{80}\cos^2\beta\left(10 - \cos^2\beta\right), \tag{C.31}$$

$$h^2_{S;(3)} = -\frac{27\sqrt{3}}{40}\sin\beta\cos^3\beta, \tag{C.32}$$

$$h^2_{S;(4)} = \frac{81}{320}\cos^4\beta, \tag{C.33}$$

$$h^2_{D;(4)} = -\frac{81}{320}\left(8 - 40\cos^2\beta + 35\cos^4\beta\right), \tag{C.34}$$

$$h^2_{D;(3)} = -\frac{81\sqrt{3}}{80}\sin\beta\cos\beta\left(4 - 7\cos^2\beta\right), \tag{C.35}$$

$$h^2_{D;(5)} = \frac{27\sqrt{3}}{80}\sin\beta\cos\beta\left(4 - 7\cos^2\beta\right), \tag{C.36}$$

$$h^2_{D;(2)} = -\frac{243}{160}\cos^2\beta\left(6 - 7\cos^2\beta\right), \tag{C.37}$$

$$h^2_{D;(6)} = -\frac{27}{160}\cos^2\beta\left(6 - 7\cos^2\beta\right), \tag{C.38}$$

$$h^2_{D;(1)} = -\frac{243\sqrt{3}}{80}\sin\beta\cos^3\beta, \tag{C.39}$$

$$h^2_{D;(7)} = \frac{9\sqrt{3}}{80}\sin\beta\cos^3\beta, \tag{C.40}$$

$$h^2_{D;(0)} = -\frac{729}{640}\cos^4\beta, \tag{C.41}$$

$$h^2_{D;(8)} = -\frac{9}{640}\cos^4\beta. \tag{C.42}$$

C.2 Sky Distribution

Having computed the response averaged over the orbit, polarization, and inclination angles in Eq. C.23, we now compute the average response to a large number of sources distributed over the sky according to a specified spatial distribution.

$$H^2_{S,D}(t,\beta,\lambda|\boldsymbol{\theta}) = \int \cos\beta d\beta \int d\lambda\; p(\lambda,\beta|\boldsymbol{\theta}) h^2_{S,D}(t), \tag{C.43}$$

We consider two possible approaches: an analytical integration and a numerical one. The analytical method, adopted in Section 5.2, is derived in full detail in [52]. In this approach, we assume that the sources follow a two-dimensional probability distribution $p(\lambda,\beta|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes the set of parameters defining the distribution. For instance, in the case of a 2D Gaussian, $\boldsymbol{\theta} = \{\lambda_M, \beta_M, \psi, \sigma_1, \sigma_2\}$, where (λ_M, β_M) is the center of the distribution, ψ is the rotation angle with respect to the ecliptic coordinate frame, and (σ_1, σ_2) are the standard deviations along the principal axes.

$$p(\beta,\lambda) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\theta^T \mathbb{M} \theta\right], \tag{C.44}$$

$$\theta = \begin{pmatrix} \beta - \beta_M \\ \lambda - \lambda_M \end{pmatrix}, \tag{C.45}$$

$$\mathbb{M} = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} \sigma_1^{-2} & 0 \\ 0 & \sigma_2^{-2} \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}, \tag{C.46}$$

For a generic random variable x distributed according to $p(x)$, we define the characteristic function $\varphi_x(m)$ as

$$\varphi_x(m) = \int dx\, p(x)\, e^{imx}. \tag{C.47}$$

This corresponds to a well-known analytical form for many standard distributions. By leveraging the characteristic function, we can analytically compute the averaged response function in Eq. C.43. This approach generalizes naturally to more complex distributions, such as Gaussian mixtures. As an example, in Fig. C.1, we illustrate the modulation of the Galactic foreground for different choices of sky distributions.

We emphasize that the probability distribution used here is an idealization: Galactic binaries are unlikely to follow a Gaussian spatial distribution on the sky, and the probability density is normalized over $(\lambda,\beta) \in \mathbb{R}^2$, rather than over the physically meaningful sky domain $\lambda \in [0, 2\pi]$, $\beta \in [-\pi/2, \pi/2]$. However, with a suitable choice of parameters, the impact of these limitations can be significantly reduced.

For these reasons, we also propose a more accurate approach in this thesis, based on a numerical integration of Eq. C.43. Specifically, we define a grid (or cube) over sky position and distance, which we transform into Galactic coordinates in order to evaluate a probability density, which depends on physically motivated parameters. We then transform back to ecliptic coordinates and perform the integration over distance and sky position numerically. Although this approach is more computationally expensive, it allows us to directly incorporate physically motivated source distributions in the sky. For example, we

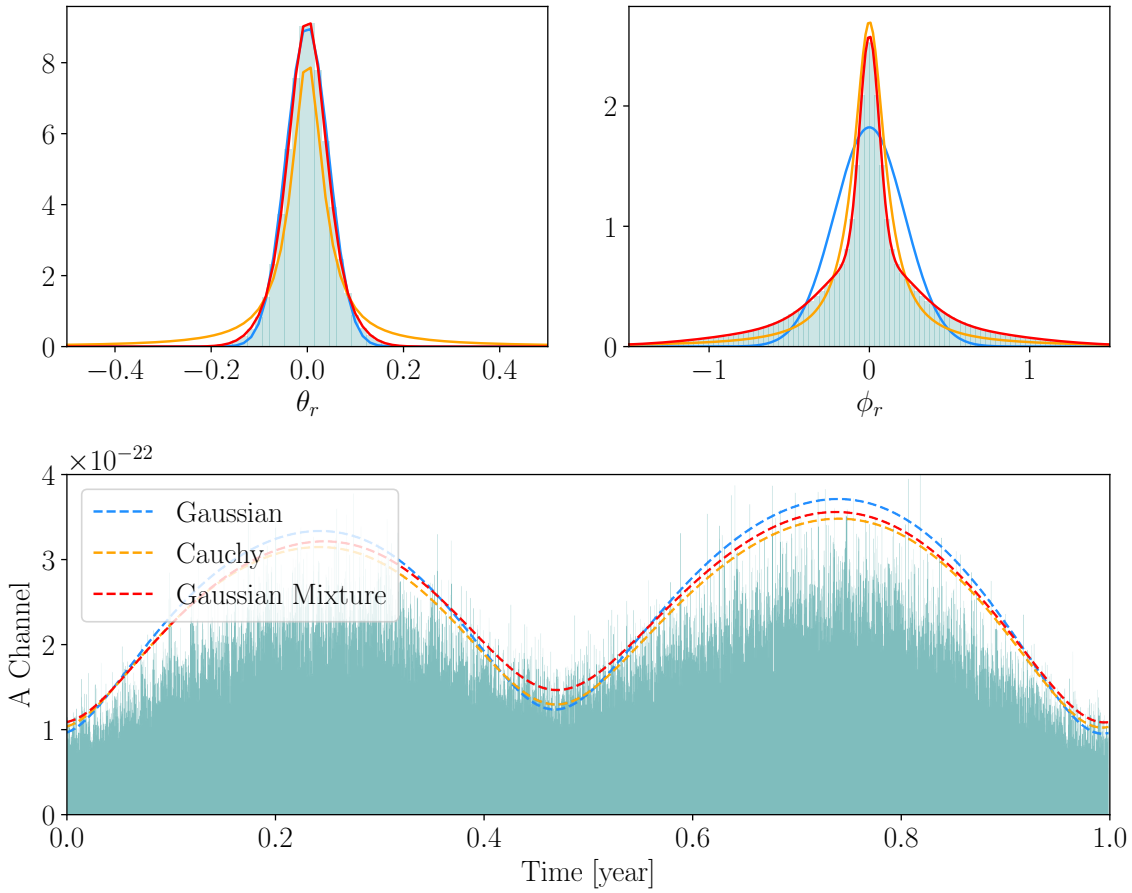

Figure C.1: (*Top panel*) Fit of rotated coordinates with various probability distributions. (*Bottom Panel*) Galactic confusion noise over one year of observation, incorporating time modulation from different probability distributions.

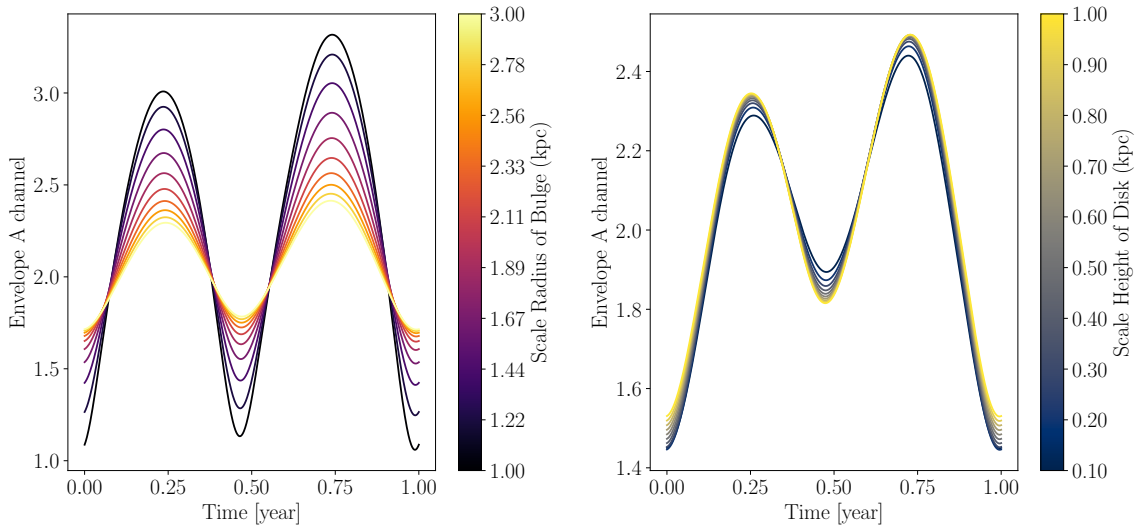

Figure C.2: (*Left panel*) Modulation of the LISA response to unresolved Galactic sources assuming a bulge-like distribution, as defined in Eq. C.48. (*Right Panel*) Modulation of the LISA response assuming a disk-like distribution, defined in Eq. C.49. In both cases, the modulation is shown for the TDI A channel, averaged over one year of observation. For the bulge model, we fix $r_0 = 0.075$ kpc and $q = 0.5$, and explore different values of R_b. For the disk model, we fix $R_d = 3.25$ kpc and vary z_d to explore its effect on the modulation.

consider a bulge-like and a disk-like distribution, defined respectively as:

$$p_{\rm bulge}(R, z; R_b, r_0, q) = p_0 \exp\left(-\frac{R}{R_b}\right)\left(1 + \frac{\sqrt{R^2 + (z/q)^2}}{r_0}\right)^{-1} \tag{C.48}$$

$$p_{\rm disk}(R, z; R_d, z_d) = p_0 \exp\left(-\frac{R}{R_d}\right)\exp\left(-\frac{|z|}{z_d}\right), \tag{C.49}$$

where p_0 is an appropriate normalization constant in each case. We consider distribution as given in [47].

In Fig. C.2, we show the modulation in the TDI A channel resulting from the numerical integration using the distributions given in Eqs. C.48 and C.49.

Appendix D

Corner from Cyclostationary Model

Corner plots in Figs. D.1–D.4, illustrate the posterior distributions analyzing mock data containing instrumental noise, SGWB, and Galactic Foreground. Both the MW and SMC signals are modeled as cyclostationary, following the formalism in Section 5.2. The true positions of the source centers are assumed known. In general, the MW signal is well constrained both in terms of its spectrum $(\alpha, \mathcal{A}_{\rm MW}, f_2)$ and modulation parameters $(\psi, \sigma_1, \sigma_2)$. The presence of the MW let us set an upper bound only on the SGWB amplitude, while the spectral slope remains accurately constrained. Fig. D.5 refers to realistic LMC data and compares the stationary and cyclostationary models. Figure D.6 shows the reconstruction of the SGWB from an hypothetical satellite located behind the Galactic disk, in the presence of the MW foreground and instrumental noise.

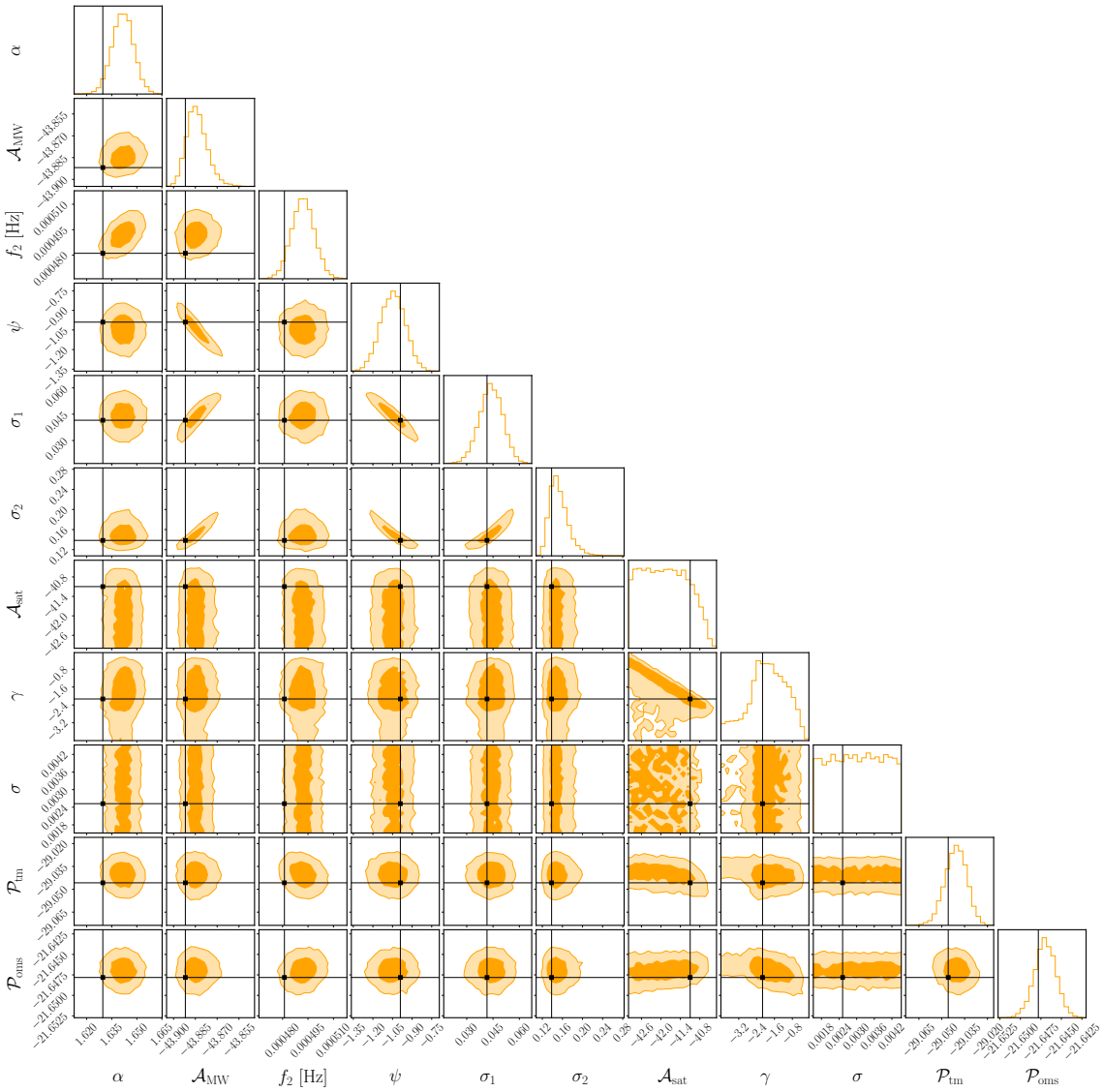

Figure D.1: Posterior distribution on SMC SGWB in the presence of noise and MW foreground. Darker (lighter) shaded areas denote 90% (50%) credible regions and black lines indicate the true values.

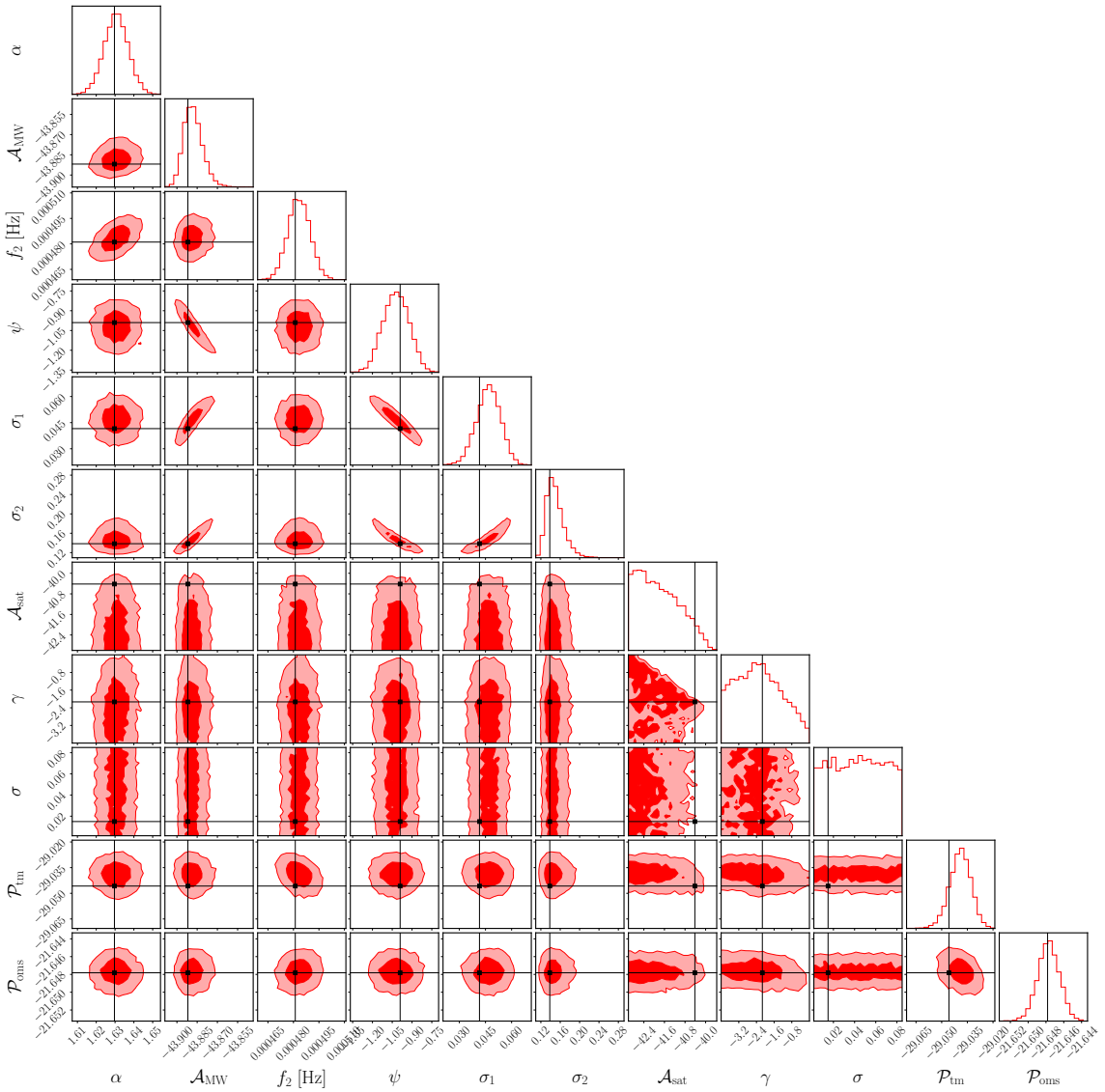

Figure D.2: Posterior distribution on LMC SGWB in presence of noise and MW foreground. Darker (lighter) shaded areas denote 90% (50%) credible regions and black lines indicate the true values.

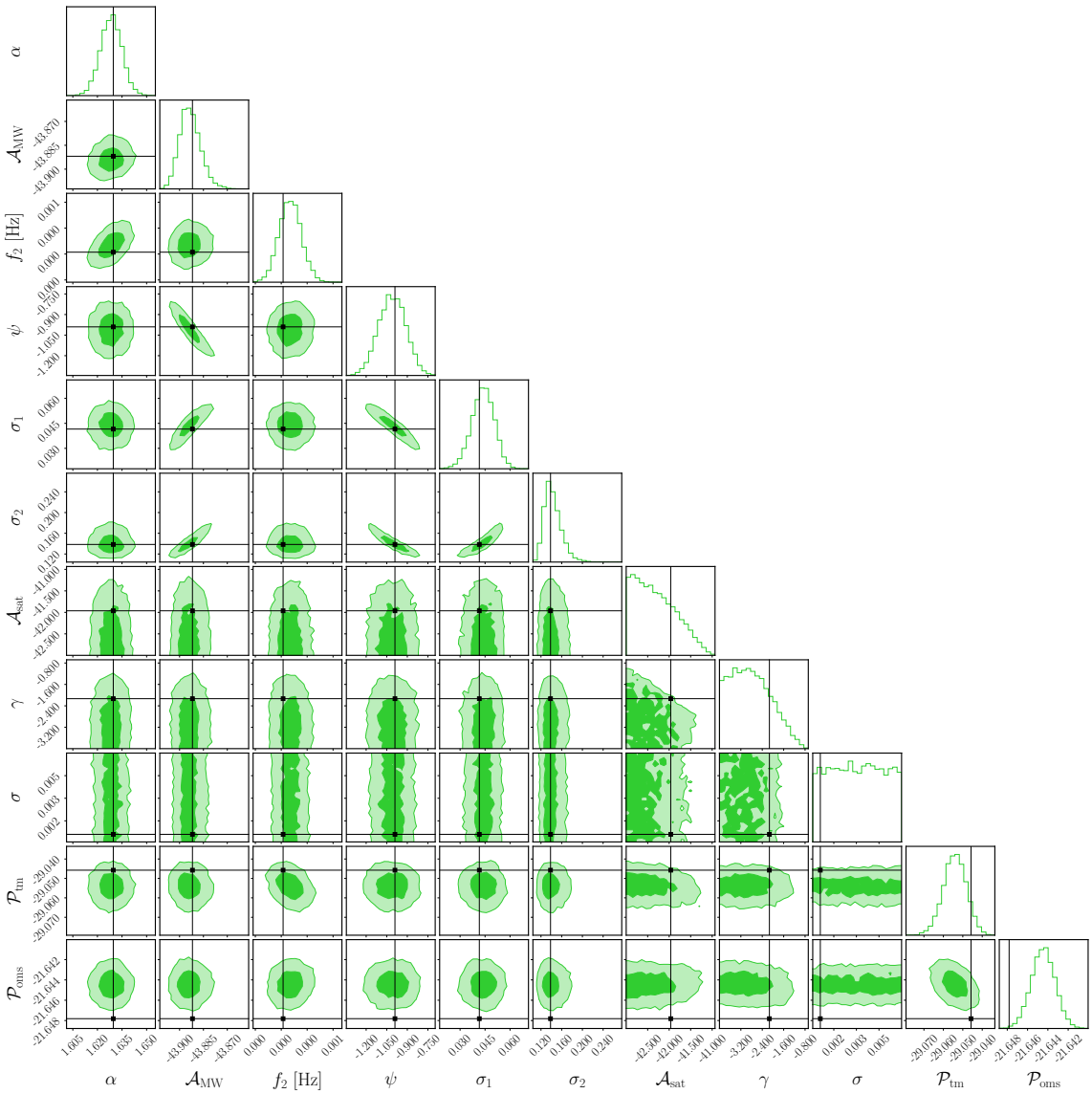

Figure D.3: Posterior distribution on Andromeda's SGWB in presence of noise and MW foreground. Darker (lighter) shaded areas denote 90% (50%) credible regions and black lines indicate the true values.

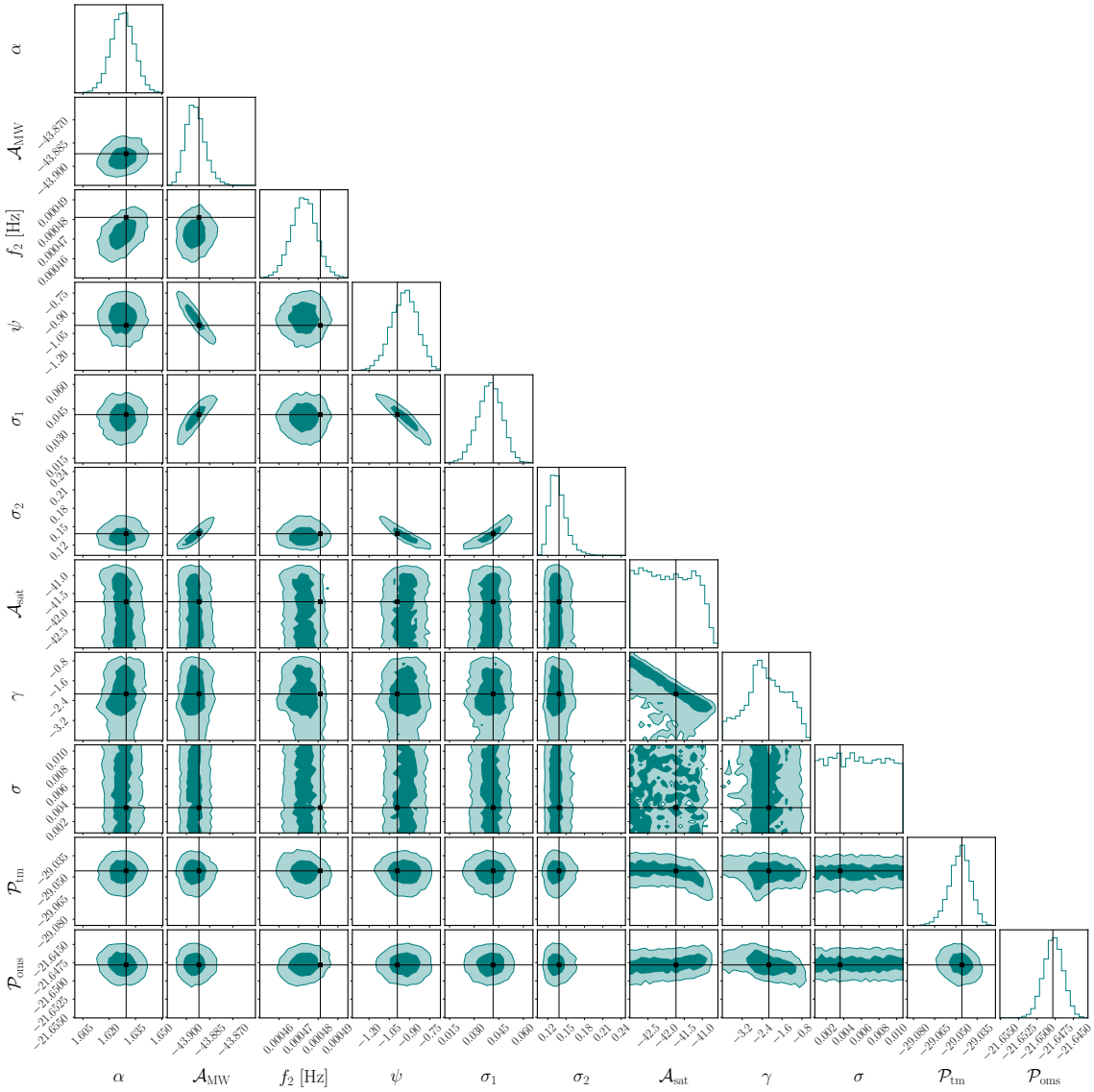

Figure D.4: Posterior distribution on Sagittarius' SGWB in presence of noise and MW foreground. Darker (lighter) shaded areas denote 90% (50%) credible regions and black lines indicate the true values.

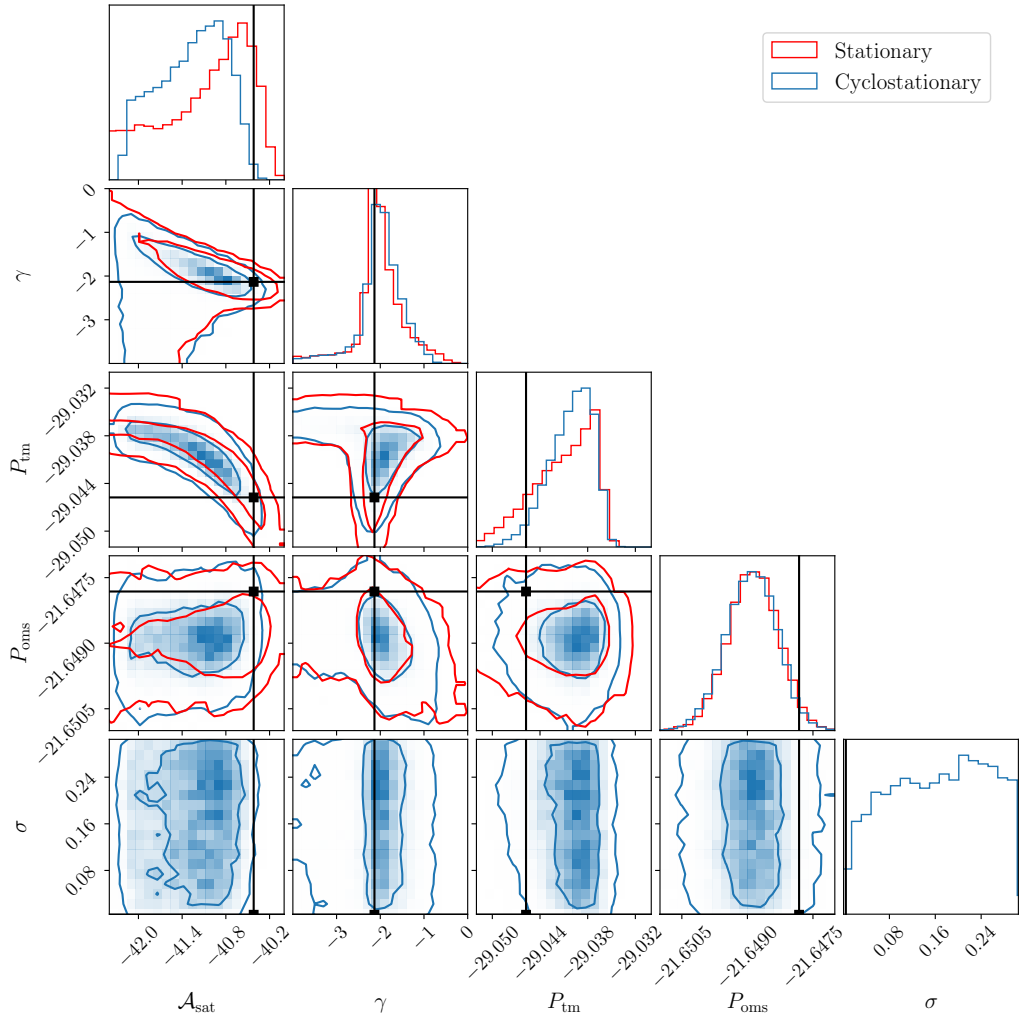

Figure D.5: The posterior distribution for noise and LMC signal's parameters, under the stationary (red) and cyclostationary (blue) hypotheses, as discussed in Sec. 5.2.4. The SGWB data are generated using a realistic catalog of unresolved sources. In the stationary model, the SGWB amplitude is directly modelled in TDI domain, whereas in the cyclostationary model, it refers to the astrophysical strain. To facilitate comparison between the two approaches, the amplitude of the stationary model in the plot has been rescaled using the modulation coefficient of the 0-th harmonic. Darker (lighter) shaded areas denote 90% (50%) credible regions and black lines indicate the true values.

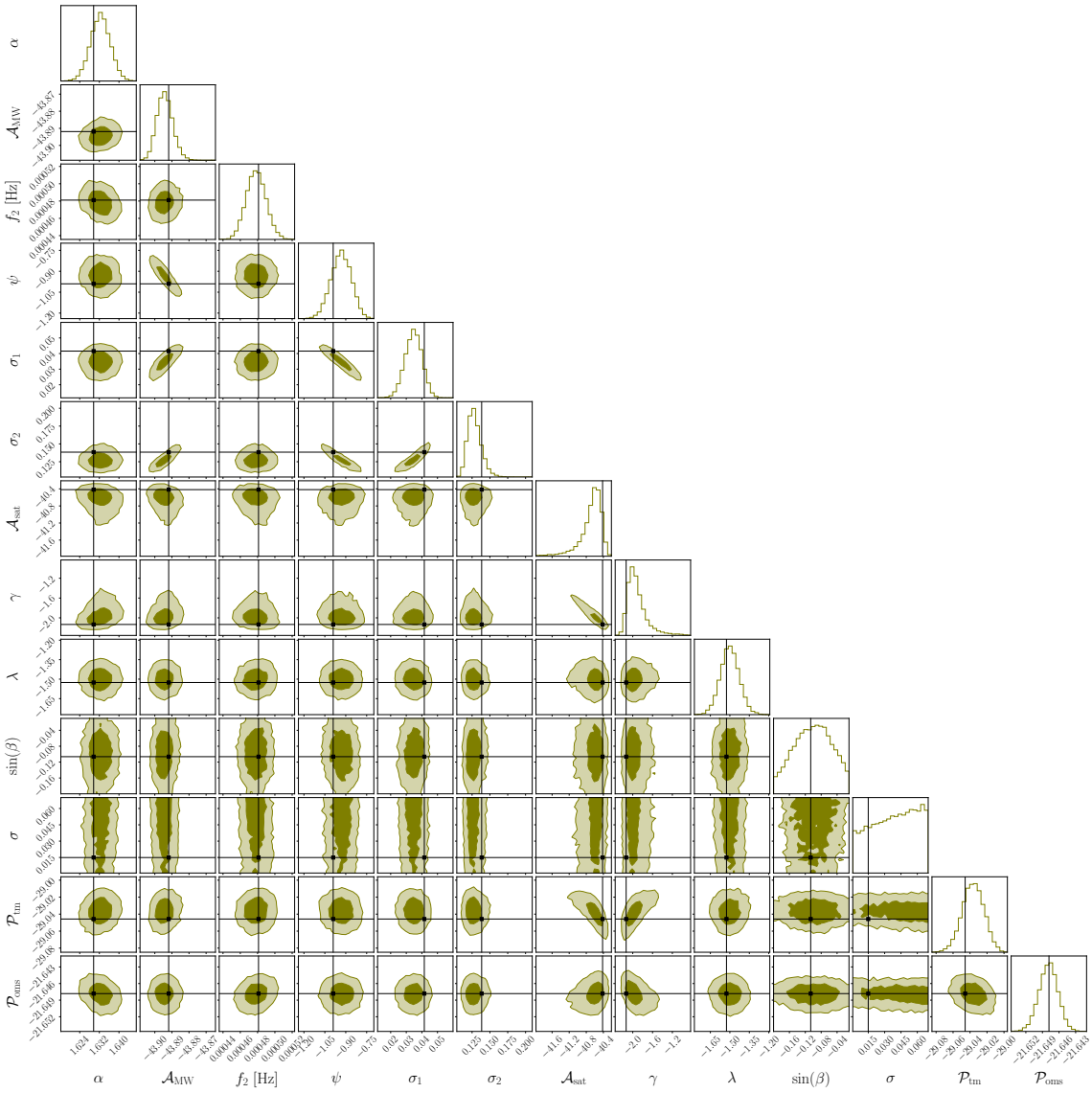

Figure D.6: Posterior distribution of MW foreground, noise, and an LMC-like mock signal emitted from the Galactic disk. The plot is discussed in Sec. 5.2.4. Darker (lighter) shaded areas denote 90% (50%) credible regions and black lines indicate the true values.

Appendix E

Plot from Bahamas

This appendix provides supplementary visualizations from the `bahamas` analysis framework, illustrating the temporal evolution of parameter constraints, the robustness of inference across different algorithmic configurations, and the impact of observational gaps on parameter recovery. Results are discussed in Section 5.3.

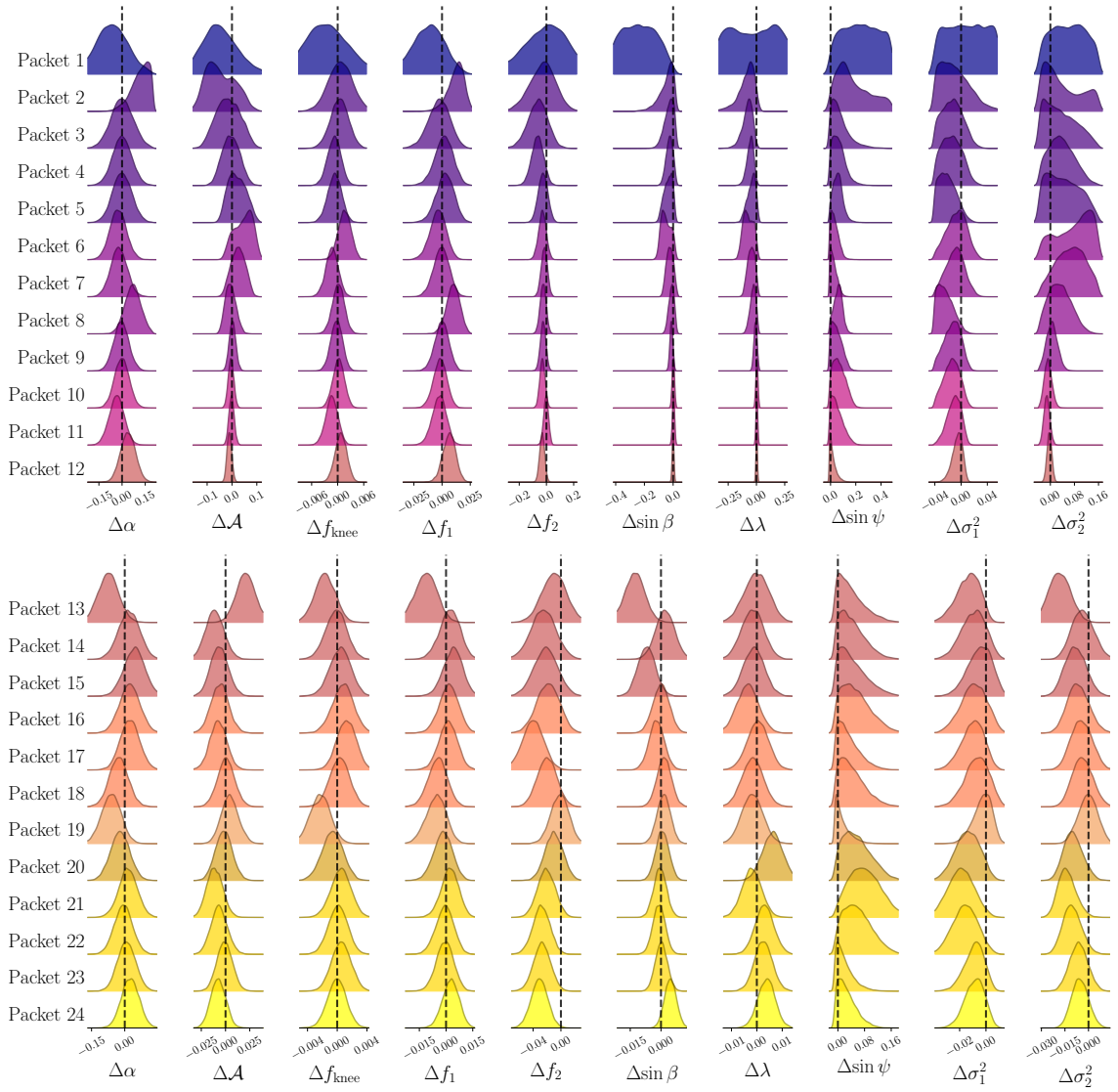

Figure E.1: Ridgeline plot of the posterior distributions of Galactic foreground parameters obtained from a sequential analysis over one year of observation. Posteriors are centered on the injected values, indicating unbiased reconstruction. From top to bottom, the distributions narrow as more data accumulates, with parameters describing the Galactic modulation becoming increasingly well-constrained by leveraging consistency across different segments.

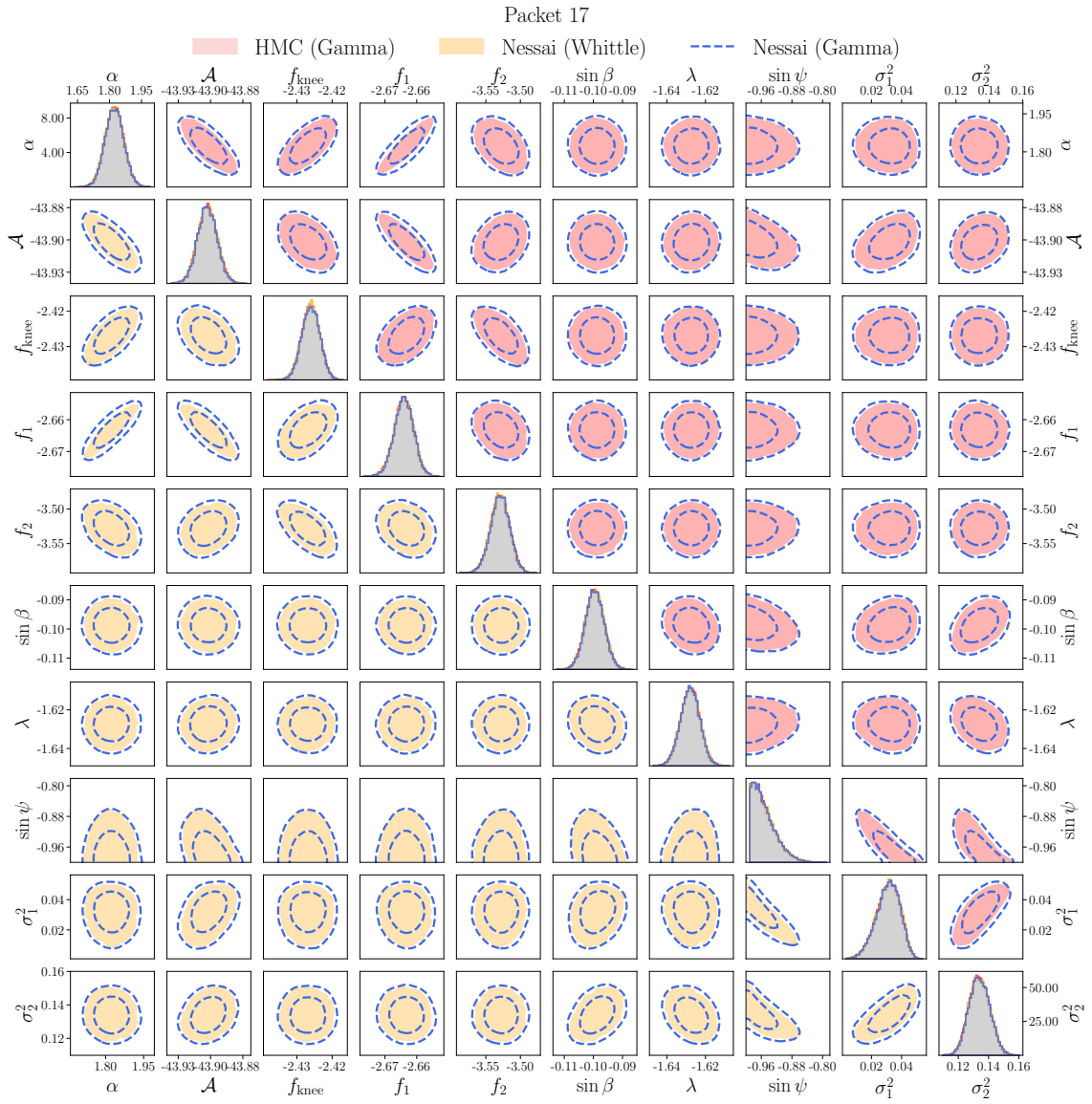

Figure E.2: Posterior distributions of the parameters obtained with different `bahamas` setup configurations. The upper triangle plot compares the results from NS and NUTS using the Gamma likelihood. The lower triangle plot shows a comparison between two likelihood choices, Whittle and Gamma, both employing NS algorithm. Overall, no significant differences appear in the distributions, demonstrating consistency across all the proposed setups. The posterior samples are taken from packet 17.

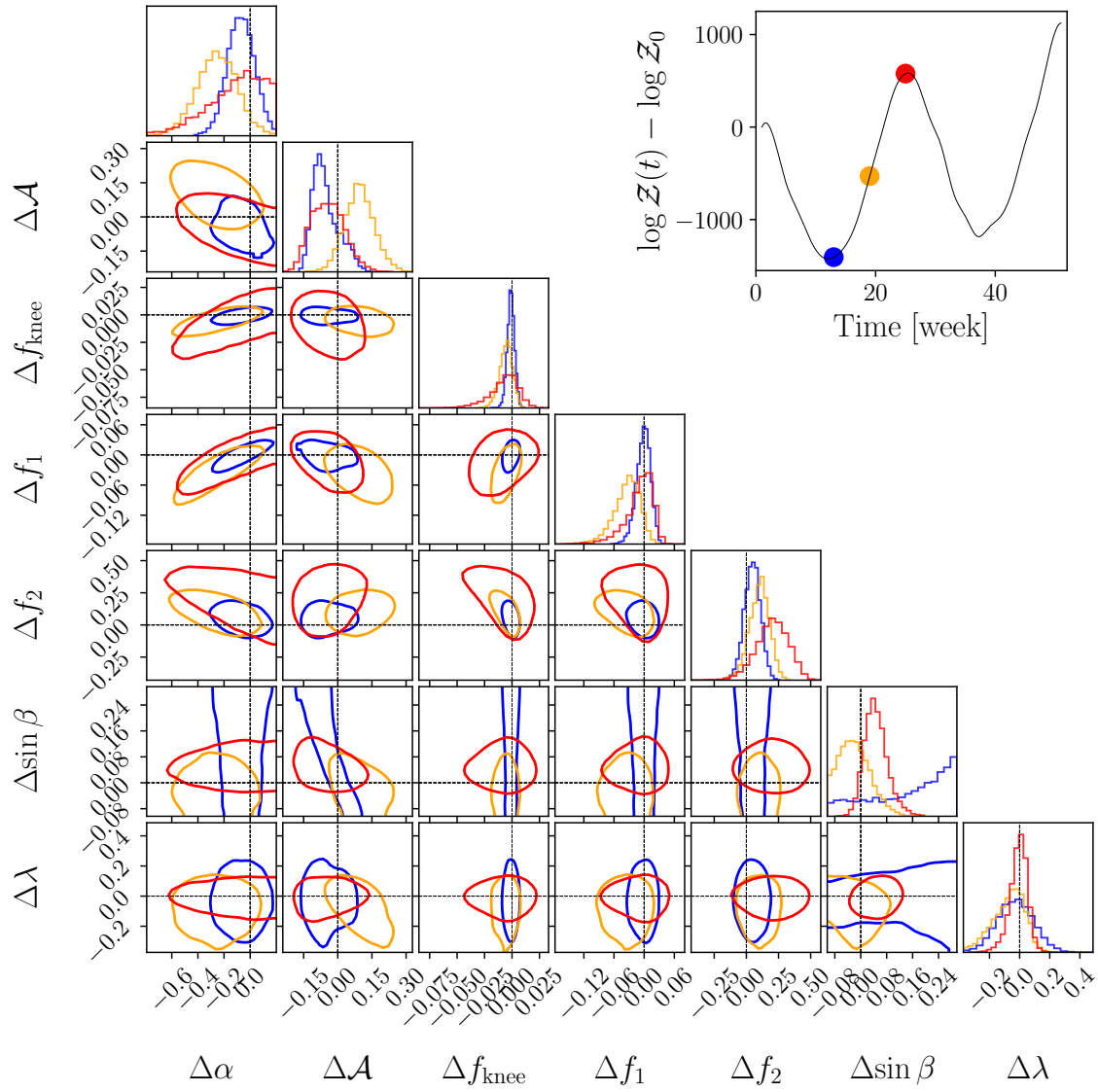

Figure E.3: Posterior distributions of the spectral parameters of the Galactic foreground and the sky coordinates of the Gaussian center describing its spatial distribution. The posteriors are shown relative to their true values. The color-coded posteriors correspond to three different values of the evidence for the quasi-stationary model, increasing from blue to yellow to red, as indicated in the top-right subplot. The plot illustrates how the reconstruction of the sky position improves with higher evidence, while the precision on the spectral parameters degrades.

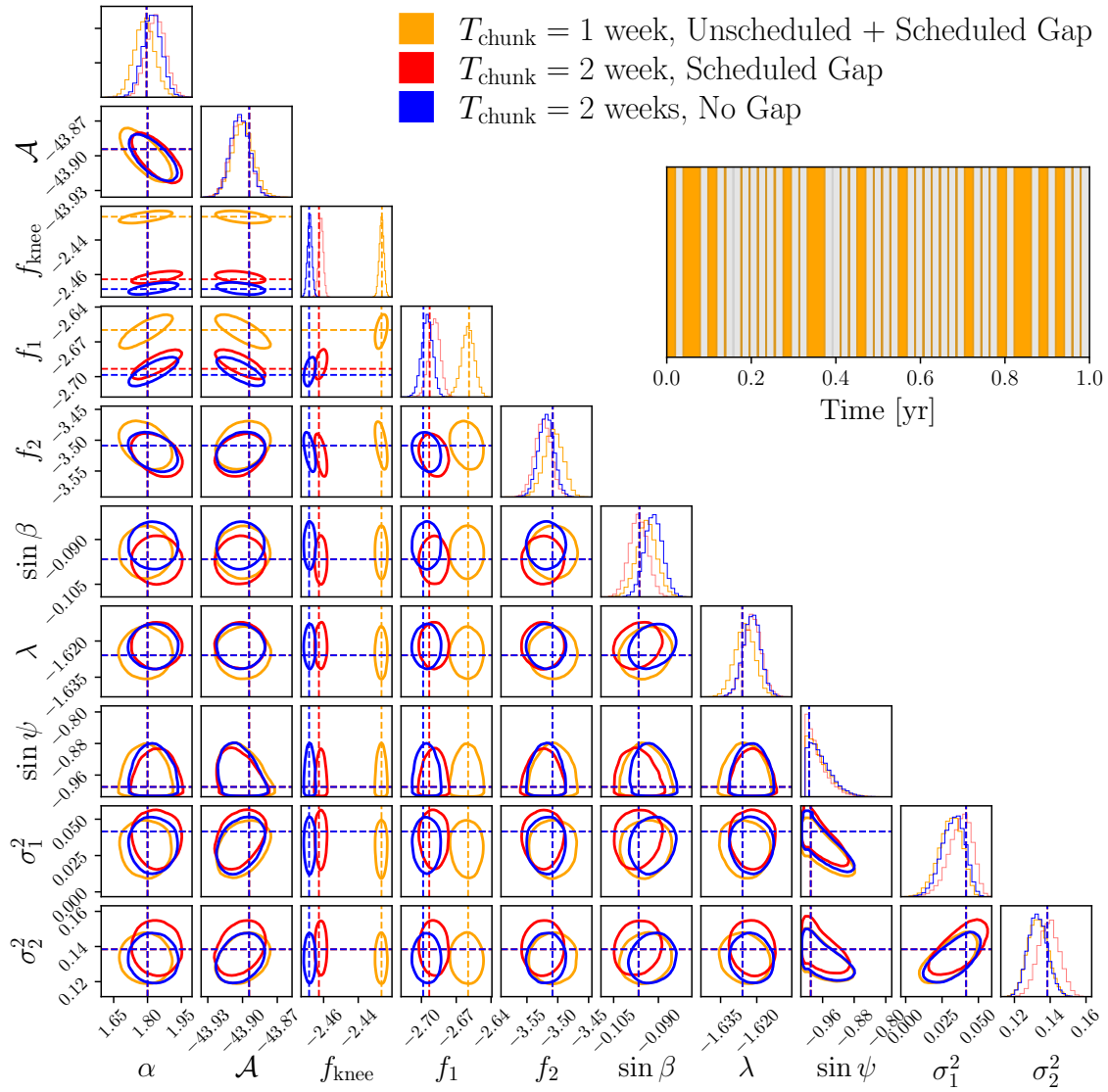

Figure E.4: Posterior distribution of the Galactic foreground under different data–gap configurations. Blue contours correspond to the reference case with full data and no gaps. Red contours include scheduled gaps of 7 hours every two weeks, while orange contours represent the most pessimistic case, with both scheduled and unscheduled gaps. The lower panel shows the temporal placement of the gaps. In this latter scenario, the overall duty cycle is about 70%. The dashed lines indicate the injected values. Note that $\log_{10} f_{\rm knee}$ and $\log_{10} f_1$ differ among the cases, as they depend on the effective observation time $T_{\rm obs}$, which is reduced in the presence of gaps.

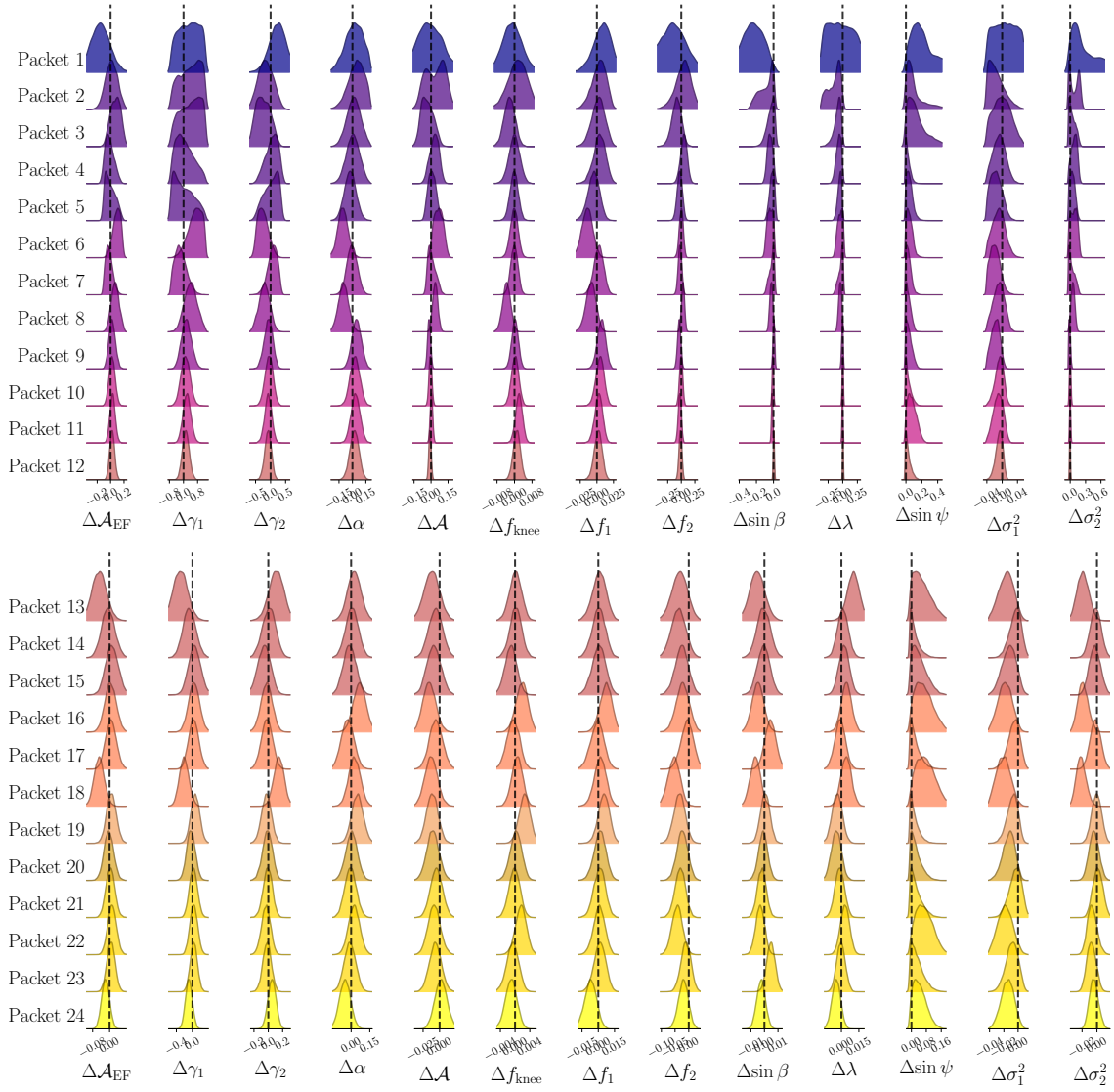

Figure E.5: Ridgeline plot of the posterior distributions of extragalactic and Galactic foreground parameters from a sequential one-year analysis. Posteriors are centered on the injected values, indicating unbiased reconstruction. The extragalactic foreground becomes well constrained starting from Packet 7, corresponding to roughly 3–4 months of observation.

Bibliography

Citing pages are listed after each reference.

[1] B. P. Abbott et al. "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run". In: *Physical Review X* 13.4, 041039 (Oct. 2023), p. 041039. DOI: 10.1103/PhysRevX.13.041039. arXiv: 2111.03606 [gr-qc] (cit. on pp. 62, 105).

[2] R. Abbott et al. "Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog". In: *Astrophys. J. Lett.* 913.1 (2021), p. L7. DOI: 10.3847/2041-8213/abe949. arXiv: 2010.14533 [astro-ph.HE] (cit. on pp. 72, 144).

[3] R. Abbott et al. "Upper limits on the isotropic gravitational-wave background from Advanced LIGO and Advanced Virgo's third observing run". In: *Phys. Rev. D* 104.2 (2021), p. 022004. DOI: 10.1103/PhysRevD.104.022004. arXiv: 2101.12130 [gr-qc] (cit. on pp. 25, 72, 144, 145).

[4] R. Abbott et al. "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3". In: *Phys. Rev. X* 13.1 (2023), p. 011048. DOI: 10.1103/PhysRevX.13.011048. arXiv: 2111.03634 [astro-ph.HE] (cit. on pp. 72, 144, 145).

[5] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, et al. "Upper limits on the isotropic gravitational-wave background from Advanced LIGO and Advanced Virgo's third observing run". In: *Phys. Rev. D* 104.2, 022004 (July 2021), p. 022004. DOI: 10.1103/PhysRevD.104.022004. arXiv: 2101.12130 [gr-qc] (cit. on p. 37).

[6] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, V. B. Adya, C. Affeldt, D. Agarwal, et al. "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3". In: *Physical Review X* 13.1, 011048 (Jan. 2023), p. 011048. DOI: 10.1103/PhysRevX.13.011048. arXiv: 2111.03634 [astro-ph.HE] (cit. on pp. 105, 106).

[7] M. R. Adams and N. J. Cornish. "Detecting a stochastic gravitational wave background in the presence of a galactic foreground and instrument noise". In: *Phys. Rev. D* 89.2, 022001 (Jan. 2014), p. 022001. DOI: 10.1103/PhysRevD.89.022001. arXiv: 1307.4116 [gr-qc] (cit. on p. 85).

[8] A. Afzal et al. "The NANOGrav 15 yr Data Set: Search for Signals from New Physics". In: *Astrophys. J. Lett.* 951.1 (2023), p. L11. DOI: 10.3847/2041-8213/acdc91. arXiv: 2306.16219 [astro-ph.HE] (cit. on p. 72).

[9] A. Afzal, G. Agazie, A. Anumarlapudi, A. M. Archibald, and et al. "The NANOGrav 15 yr Data Set: Search for Signals from New Physics". In: *Astrophys. J. Lett.* 951.1, L11 (July 2023), p. L11. DOI: 10.3847/2041-8213/acdc91. arXiv: 2306.16219 [astro-ph.HE] (cit. on p. 62).

[10] G. Agazie et al. "The NANOGrav 15 yr Data Set: Constraints on Supermassive Black Hole Binaries from the Gravitational-wave Background". In: *Astrophys. J. Lett.* 952.2 (2023), p. L37. DOI: 10.3847/2041-8213/ace18b. arXiv: 2306.16220 [astro-ph.HE] (cit. on p. 72).

[11] G. Agazie et al. "The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background". In: *Astrophys. J. Lett.* 951.1 (2023), p. L8. DOI: 10.3847/2041-8213/acdac6. arXiv: 2306.16213 [astro-ph.HE] (cit. on p. 72).

[12] G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, and et al. "The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background". In: *Astrophys. J. Lett.* 951.1, L8 (July 2023), p. L8. DOI: 10.3847/2041-8213/acdac6. arXiv: 2306.16213 [astro-ph.HE] (cit. on p. 62).

[13] B. Allen and J. D. Romano. "Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities". In: *Phys. Rev. D* 59 (1999), p. 102001. DOI: 10.1103/PhysRevD.59.102001. arXiv: gr-qc/9710117 (cit. on p. 25).

[14] P. Amaro-Seoane. "Relativistic dynamics and extreme mass ratio inspirals". In: *Living Reviews in Relativity* 21.1, 4 (Dec. 2018), p. 4. DOI: 10.1007/s41114-018-0013-8. arXiv: 1205.5240 [astro-ph.CO] (cit. on p. 23).

[15] P. Amaro-Seoane, J. R. Gair, M. Freitag, M. C. Miller, I. Mandel, C. J. Cutler, and S. Babak. "TOPICAL REVIEW: Intermediate and extreme mass-ratio inspirals—astrophysics, science applications and detection using LISA". In: *Classical and Quantum Gravity* 24.17 (Sept. 2007), R113–R169. DOI: 10.1088/0264-9381/24/17/R01. arXiv: astro-ph/0703495 [astro-ph] (cit. on p. 57).

[16] P. Amaro-Seoane and M. Preto. "The impact of realistic models of mass segregation on the event rate of extreme-mass ratio inspirals and cusp re-growth". In: *Classical and Quantum Gravity* 28.9, 094017 (May 2011), p. 094017. DOI: 10.1088/0264-9381/28/9/094017. arXiv: 1010.5781 [astro-ph.CO] (cit. on p. 66).

[17] J. Antoniadis et al. "The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals". In: *Astron. Astrophys.* 678 (2023), A50. DOI: 10.1051/0004-6361/202346844. arXiv: 2306.16214 [astro-ph.HE] (cit. on p. 72).

[18] J. Antoniadis et al. "The second data release from the European Pulsar Timing Array - V. Search for continuous gravitational wave signals". In: *Astron. Astrophys.* 690 (2024), A118. DOI: 10.1051/0004-6361/202348568. arXiv: 2306.16226 [astro-ph.HE] (cit. on p. 72).

[19] J. Antoniadis, P. Arumugam, S. Arumugam, P. Auclair, and et al. "The second data release from the European Pulsar Timing Array: V. Implications for massive black holes, dark matter and the early Universe". In: *arXiv e-prints* (June 2023). DOI: 10.48550/arXiv.2306.16227. arXiv: 2306.16227 [astro-ph.CO] (cit. on p. 62).

[20] F. Antonini, E. Barausse, and J. Silk. "The Coevolution of Nuclear Star Clusters, Massive Black Holes, and Their Host Galaxies". In: *Astrophys. J.* 812.1, 72 (Oct. 2015), p. 72. DOI: 10.1088/0004-637X/812/1/72. arXiv: 1506.02050 [astro-ph.GA] (cit. on p. 65).

[21] T. Appourchaux. "On Maximum Likelihood Estimation of averaged power spectra". In: *Astronomy & Astrophysics* 412 (Dec. 2003), pp. 903–904. DOI: 10.1051/0004-6361:20034401 (cit. on p. 79).

[22] M. Armano and others. "Sub-Femto-g Free Fall for Space-Based Gravitational Wave Observatories: LISA Pathfinder Results". In: *Phys. Rev. Lett.* 116 (23 June 2016), p. 231101. DOI: 10.1103/PhysRevLett.116.231101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.231101 (cit. on p. 45).

[23] M. Armano and others. "LISA Pathfinder". In: *arXiv e-prints*, arXiv:1903.08924 (Mar. 2019), arXiv:1903.08924. DOI: 10.48550/arXiv.1903.08924. arXiv: 1903.08924 [astro-ph.IM] (cit. on p. 45).

[24] M. Armano and others. "In-depth analysis of LISA Pathfinder performance results: Time evolution, noise projection, physical models, and implications for LISA". In: *Phys. Rev. D* 110.4, 042004 (Aug. 2024), p. 042004. DOI: 10.1103/PhysRevD.110.042004. arXiv: 2405.05207 [astro-ph.IM] (cit. on p. 45).

[25] K. G. Arun et al. "New horizons for fundamental physics with LISA". In: *Living Rev. Rel.* 25.1 (2022), p. 4. DOI: 10.1007/s41114-022-00036-9. arXiv: 2205.01597 [gr-qc] (cit. on p. 72).

[26] P. Auclair et al. "Cosmology with the Laser Interferometer Space Antenna". In: *Living Reviews in Relativity* 26.1, 5 (Dec. 2023), p. 5. DOI: 10.1007/s41114-023-00045-2. arXiv: 2204.05434 [astro-ph.CO] (cit. on p. 59).

[27] P. Auclair, J. J. Blanco-Pillado, D. G. Figueroa, A. C. Jenkins, M. Lewicki, M. Sakellariadou, S. Sanidas, L. Sousa, D. A. Steer, J. M. Wachter, S. Kuroyanagi, and LISA Cosmology Working Group. "Probing the gravitational wave background from cosmic strings with LISA". In: *J. Cosmology Astropart. Phys.* 2020.4, 034 (Apr. 2020), p. 034. DOI: 10.1088/1475-7516/2020/04/034. arXiv: 1909.00819 [astro-ph.CO] (cit. on p. 24).

[28] S. Babak and the Challenge 3 partecipants. "The Mock LISA Data Challenges: from challenge 3 to challenge 4". In: *Classical and Quantum Gravity* 27.8, 084009 (Apr. 2010), p. 084009. DOI: 10.1088/0264-9381/27/8/084009. arXiv: 0912.0548 [gr-qc] (cit. on p. 68).

[29] 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". In: *JCAP* 08 (2023), p. 034. DOI: 10.1088/1475-7516/2023/08/034. arXiv: 2304.06368 [astro-ph.CO] (cit. on p. 80).

[30] S. Babak, H. Fang, J. R. Gair, K. Glampedakis, and S. A. Hughes. ""Kludge" gravitational waveforms for a test-body orbiting a Kerr black hole". In: *Phys. Rev. D* 75.2, 024005 (Jan. 2007), p. 024005. DOI: 10.1103/PhysRevD.75.024005. arXiv: gr-qc/0607007 [gr-qc] (cit. on p. 23).

[31] 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". In: *Phys. Rev. D* 95.10, 103012 (May 2017), p. 103012. DOI: 10.1103/PhysRevD.95.103012. arXiv: 1703.09722 [gr-qc] (cit. on pp. 63–67, 70, 71, 105).

[32] Q. Baghi, N. Karnesis, J.-B. Bayle, M. Besançon, and H. Inchauspé. "Uncovering gravitational-wave backgrounds from noises of unknown shape with LISA". In: *J. Cosmology Astropart. Phys.* 2023.4, 066 (Apr. 2023), p. 066. DOI: 10.1088/1475-7516/2023/04/066. arXiv: 2302.12573 [gr-qc] (cit. on pp. 54, 85, 96, 127).

[33] J. Baker, Z. Haiman, and others. "Multimessenger science opportunities with mHz gravitational waves". In: *BAAS* 51.3, 123 (May 2019), p. 123. DOI: 10.48550/arXiv.1903.04417. arXiv: 1903.04417 [astro-ph.HE] (cit. on p. 56).

[34] L. Barack and C. Cutler. "LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy". In: *Phys. Rev. D* 69.8, 082005 (Apr. 2004), p. 082005. DOI: 10.1103/PhysRevD.69.082005. arXiv: gr-qc/0310125 [gr-qc] (cit. on p. 23).

[35] L. Barack and C. Cutler. "Using LISA extreme-mass-ratio inspiral sources to test off-Kerr deviations in the geometry of massive black holes". In: *Phys. Rev. D* 75.4, 042003 (Feb. 2007), p. 042003. DOI: 10.1103/PhysRevD.75.042003. arXiv: gr-qc/0612029 [gr-qc] (cit. on p. 57).

[36] E. Barausse. "Relativistic dynamical friction in a collisional fluid". In: *Mon. Not. Roy. Astron. Soc.* 382 (2007), pp. 826–834. DOI: 10.1111/j.1365-2966.2007.12408.x. arXiv: 0709.0211 [astro-ph] (cit. on p. 72).

[37] E. Barausse et al. "Prospects for Fundamental Physics with LISA". In: *Gen. Rel. Grav.* 52.8 (2020), p. 81. DOI: 10.1007/s10714-020-02691-1. arXiv: 2001.09793 [gr-qc] (cit. on p. 72).

[38] E. Barausse, V. Cardoso, and P. Pani. "Can environmental effects spoil precision gravitational-wave astrophysics?" In: *Phys. Rev. D* 89.10 (2014), p. 104059. DOI: 10.1103/PhysRevD.89.104059. arXiv: 1404.7149 [gr-qc] (cit. on pp. 72, 75, 80).

[39] E. Barausse, V. Cardoso, and P. Pani. "Environmental Effects for Gravitational-wave Astrophysics". In: *J. Phys. Conf. Ser.* 610.1 (2015). Ed. by G. Ciani, J. W. Conklin, and G. Mueller, p. 012044. DOI: 10.1088/1742-6596/610/1/012044. arXiv: 1404.7140 [astro-ph.CO] (cit. on pp. 57, 72, 80).

[40] E. Barausse and L. Rezzolla. "The Influence of the hydrodynamic drag from an accretion torus on extreme mass-ratio inspirals". In: *Phys. Rev. D* 77 (2008), p. 104027. DOI: 10.1103/PhysRevD.77.104027. arXiv: 0711.4558 [gr-qc] (cit. on pp. 57, 72, 75, 84).

[41] G. Boileau, T. Bruel, A. Toubiana, A. Lamberts, and N. Christensen. "Gravitational-Wave Background from Extragalactic Double White Dwarfs for LISA". In: *arXiv e-prints*, arXiv:2506.18390 (June 2025), arXiv:2506.18390. DOI: 10.48550/arXiv.2506.18390. arXiv: 2506.18390 [gr-qc] (cit. on p. 62).

[42] G. Boileau, N. Christensen, C. Gowling, M. Hindmarsh, and R. Meyer. "Prospects for LISA to detect a gravitational-wave background from first order phase transitions". In: *J. Cosmology Astropart. Phys.* 2023.2, 056 (Feb. 2023), p. 056. DOI: 10.1088/1475-7516/2023/02/056. arXiv: 2209.13277 [gr-qc] (cit. on p. 85).

[43] M. Bonetti and A. Sesana. "Gravitational wave background from extreme mass ratio inspirals". In: *Phys. Rev. D* 102.10, 103023 (Nov. 2020), p. 103023. DOI: 10.1103/PhysRevD.102.103023. arXiv: 2007.14403 [astro-ph.GA] (cit. on pp. 67, 69–71, 104, 105).

[44] 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". In: *Mon. Not. Roy. Astron. Soc.* 477.2 (2018), pp. 2599–2612. DOI: 10.1093/mnras/sty874. arXiv: 1709.06095 [astro-ph.GA] (cit. on p. 72).

[45] 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". In: *Mon. Not. R. Astron. Soc.* 486.3 (July 2019), pp. 4044–4060. DOI: 10.1093/mnras/stz903. arXiv: 1812.01011 [astro-ph.GA] (cit. on p. 56).

[46] J. Bradbury and others. *JAX: composable transformations of Python+NumPy programs.* Version 0.3.13. 2018. URL: http://github.com/google/jax (cit. on p. 40).

[47] K. Breivik, C. M. F. Mingarelli, and S. L. Larson. "Constraining Galactic Structure with the LISA White Dwarf Foreground". In: *Astrophys. J.* 901.1, 4 (Sept. 2020), p. 4. DOI: 10.3847/1538-4357/abab99. arXiv: 1912.02200 [astro-ph.GA] (cit. on pp. 60, 111, 155).

[48] O. Burke, S. Marsat, J. R. Gair, and M. L. Katz. "Addressing data gaps and assessing noise mismodeling in LISA". In: *Phys. Rev. D* 111.12, 124053 (June 2025), p. 124053. DOI: 10.1103/5jr8-k2ss. arXiv: 2502.17426 [gr-qc] (cit. on p. 134).

[49] 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". In: *Phys. Rev. D* 109.12, 124048 (June 2024), p. 124048. DOI: 10.1103/PhysRevD.109.124048. arXiv: 2310.08927 [gr-qc] (cit. on p. 105).

[50] R. Buscicchio, J. Torrado, C. Caprini, G. Nardini, and et al. "Stellar-mass black-hole binaries in LISA: characteristics and complementarity with current-generation interferometers". In: *J. Cosmology Astropart. Phys.* 2025.1, 084 (Jan. 2025), p. 084. DOI: 10.1088/1475-7516/2025/01/084. arXiv: 2410.18171 [astro-ph.HE] (cit. on p. 57).

[51] R. Buscicchio, A. Klein, V. Korol, F. Di Renzo, C. J. Moore, D. Gerosa, and A. Carzaniga. "A test for LISA foreground Gaussianity and stationarity. I. Galactic white-dwarf binaries". In: *arXiv e-prints*, arXiv:2410.08263 (Oct. 2024), arXiv:2410.08263. DOI: 10.48550/arXiv.2410.08263. arXiv: 2410.08263 [astro-ph.HE] (cit. on p. 61).

[52] R. Buscicchio, A. Klein, V. Korol, F. Di Renzo, C. J. Moore, D. Gerosa, and A. Carzaniga. "A test for LISA foreground Gaussianity and stationarity. I. Galactic white-dwarf binaries". In: *arXiv e-prints*, arXiv:2410.08263 (Oct. 2024), arXiv:2410.08263. DOI: 10.48550/arXiv.2410.08263. arXiv: 2410.08263 [astro-ph.HE] (cit. on p. 153).

[53] R. Buscicchio, A. Klein, V. Korol, F. Di Renzo, C. Moore, D. Gerosa, and A. Carzaniga. "A test for LISA foreground Gaussianity and stationarity. I. Galactic white-dwarf binaries". In: *arXiv e-prints*, arXiv: (Sept. 2024), arXiv: arXiv: 2410.08263 [astro-ph.HE] (cit. on pp. 111, 114, 121).

[54] R. Buscicchio, A. Klein, V. Korol, F. Di Renzo, C. J. Moore, D. Gerosa, and A. Carzaniga. "A test for LISA foreground Gaussianity and stationarity. I. White-dwarf binaries". In: *arXiv e-prints*, arXiv: (Sept. 2024), arXiv: arXiv: {} [astro-ph] (cit. on pp. 104, 108).

[55] R. Buscicchio, A. Klein, E. Roebber, C. J. Moore, D. Gerosa, E. Finch, and A. Vecchio. "Bayesian parameter estimation of stellar-mass black-hole binaries with LISA". In: *Phys. Rev. D* 104.4, 044065 (Aug. 2021), p. 044065. DOI: 10.1103/PhysRevD.104.044065. arXiv: 2106.05259 [astro-ph.HE] (cit. on p. 85).

[56] M. Çalışkan, N. A. Kumar, M. Kamionkowski, and S. Cheng. "A Fog Over the Cosmological SGWB: Unresolved Massive Black Hole Binaries in the LISA Band". In: *arXiv e-prints*, arXiv:2506.18965 (June 2025), arXiv:2506.18965. DOI: 10.48550/arXiv.2506.18965. arXiv: 2506.18965 [astro-ph.CO] (cit. on p. 62).

[57] T. Callister, L. Jenks, D. E. Holz, and N. Yunes. "New probe of gravitational parity violation through nonobservation of the stochastic gravitational-wave background". In: *Phys. Rev. D* 111.4 (2025), p. 044041. DOI: 10.1103/PhysRevD.111.044041. arXiv: 2312.12532 [gr-qc] (cit. on pp. 72, 144).

[58] C. Caprini, M. Chala, G. C. Dorsch, M. Hindmarsh, S. J. Huber, T. Konstandin, J. Kozaczuk, G. Nardini, J. M. No, K. Rummukainen, P. Schwaller, G. Servant, A. Tranberg, and D. J. Weir. "Detecting gravitational waves from cosmological phase transitions with LISA: an update". In: *J. Cosmology Astropart. Phys.* 2020.3, 024 (Mar. 2020), p. 024. DOI: 10.1088/1475-7516/2020/03/024. arXiv: 1910.13125 [astro-ph.CO] (cit. on p. 24).

[59] C. Caprini, D. G. Figueroa, R. Flauger, G. Nardini, M. Peloso, M. Pieroni, A. Ricciardone, and G. Tasinato. "Reconstructing the spectral shape of a stochastic gravitational wave background with LISA". In: *J. Cosmology Astropart. Phys.* 2019.11, 017 (Nov. 2019), p. 017. DOI: 10.1088/1475-7516/2019/11/017. arXiv: 1906.09244 [astro-ph.CO] (cit. on p. 85).

[60] 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". In: *Astrophys. J.* 892.2 (2020), p. 90. DOI: 10.3847/1538-4357/ab7b66. arXiv: 2001.03620 [astro-ph.HE] (cit. on pp. 72, 75, 77).

[61] R. S. Chandramouli and N. Yunes. "Ready-to-use analytic model for gravitational waves from a hierarchical triple with Kozai-Lidov oscillations". In: *Phys. Rev. D* 105.6 (2022), p. 064009. DOI: 10.1103/PhysRevD.105.064009. arXiv: 2107.00741 [gr-qc] (cit. on pp. 72, 84).

[62] S. Chandrasekhar. "Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction." In: *Astrophys. J.* 97 (Mar. 1943), p. 255. DOI: 10.1086/144517 (cit. on pp. 72, 74).

[63] R. Chen, R. S. Chandramouli, F. Pozzoli, R. Buscicchio, and E. Barausse. "Muffled Murmurs: Environmental effects in the LISA stochastic signal from stellar-mass black hole binaries". In: *arXiv e-prints*, arXiv:2507.00694 (July 2025), arXiv:2507.00694. DOI: 10.48550/arXiv.2507.00694. arXiv: 2507.00694 [gr-qc] (cit. on p. 63).

[64] S. Chen, A. Sesana, and W. Del Pozzo. "Efficient computation of the gravitational wave spectrum emitted by eccentric massive black hole binaries in stellar environments". In: *Mon. Not. Roy. Astron. Soc.* 470.2 (2017), pp. 1738–1749. DOI: 10.1093/mnras/stx1093. arXiv: 1612.00455 [astro-ph.CO] (cit. on pp. 72, 84).

[65] M. Chruslinska, G. Nelemans, and K. Belczynski. "The influence of the distribution of cosmic star formation at different metallicities on the properties of merging double compact objects". In: *Mon. Not. Roy. Astron. Soc.* 482.4 (2019), pp. 5012–5017. DOI: 10.1093/mnras/sty3087. arXiv: 1811.03565 [astro-ph.HE] (cit. on p. 144).

[66] 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". In: *Phys. Rev. Lett.* 126.5, 051102 (Feb. 2021), p. 051102. DOI: 10.1103/PhysRevLett.126.051102. arXiv: 2008.06071 [gr-qc] (cit. on p. 23).

[67] M. Colpi, K. Danzmann, M. Hewitson, K. Holley-Bockelmann, et al. "LISA Definition Study Report". In: *arXiv e-prints* (Feb. 2024). DOI: 10.48550/arXiv.2402.07571. arXiv: 2402.07571 [astro-ph.CO] (cit. on pp. 58, 100).

[68] D. O. Cook, D. A. Dale, B. D. Johnson, L. Van Zee, J. C. Lee, R. C. Kennicutt, D. Calzetti, S. M. Staudaher, and C. W. Engelbracht. "The Spitzer Local Volume Legacy (LVL) global optical photometry". In: *Mon. Not. R. Astron. Soc.* 445.1 (Nov. 2014), pp. 881–889. DOI: 10.1093/mnras/stu1580. arXiv: 1408.1130 [astro-ph.GA] (cit. on p. 117).

[69] L. Copparoni, L. Speri, L. Sberna, A. Derdzinski, and E. Barausse. "The implications of stochastic gas torques for asymmetric binaries in the LISA band". In: (Feb. 2025). arXiv: 2502.10087 [gr-qc] (cit. on p. 72).

[70] N. Cornish, L. Sampson, N. Yunes, and F. Pretorius. "Gravitational wave tests of general relativity with the parameterized post-Einsteinian framework". In: *Phys. Rev. D* 84.6, 062003 (Sept. 2011), p. 062003. DOI: 10.1103/PhysRevD.84.062003. arXiv: 1105.2088 [gr-qc] (cit. on p. 100).

[71] K. Danzmann and A. Rüdiger. "LISA technology - concept, status, prospects". In: *Classical and Quantum Gravity* 20.10 (May 2003), S1–S9. DOI: 10.1088/0264-9381/20/10/301 (cit. on p. 46).

[72] B. Deme, B.-M. Hoang, S. Naoz, and B. Kocsis. "Detecting Kozai–Lidov Imprints on the Gravitational Waves of Intermediate-mass Black Holes in Galactic Nuclei". In: *Astrophys. J.* 901.2 (2020), p. 125. DOI: 10.3847/1538-4357/abafa3. arXiv: 2005.03677 [astro-ph.HE] (cit. on p. 72).

[73] S. Deng, S. Babak, M. Le Jeune, S. Marsat, É. Plagnol, and A. Sartirana. "Modular global-fit pipeline for LISA data analysis". In: *Phys. Rev. D* 111.10, 103014 (May 2025), p. 103014. DOI: 10.1103/PhysRevD.111.103014. arXiv: 2501.10277 [gr-qc] (cit. on p. 36).

[74] K. Dey, N. Karnesis, A. Toubiana, E. Barausse, N. Korsakova, Q. Baghi, and S. Basak. "Effect of data gaps on the detectability and parameter estimation of massive black hole binaries with LISA". In: *Phys. Rev. D* 104.4, 044035 (Aug. 2021), p. 044035. DOI: 10.1103/PhysRevD.104.044035. arXiv: 2104.12646 [gr-qc] (cit. on p. 134).

[75] F. Di Renzo. "Characterisation and mitigation of non-stationary noise in Advance Gravitational Wave Detectors". PhD dissertation. University of Pisa, 2020. URL: https://etd.adm.unipi.it/t/etd-06192020-081853 (cit. on p. 104).

[76] M. C. Digman and N. J. Cornish. "LISA Gravitational Wave Sources in a Time-varying Galactic Stochastic Background". In: *Astrophys. J.* 940.1, 10 (Nov. 2022), p. 10. DOI: 10.3847/1538-4357/ac9139. arXiv: 2206.14813 [astro-ph.IM] (cit. on p. 111).

[77] 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". In: *Astrophys. J.* 901.1, 25 (Sept. 2020), p. 25. DOI: 10.3847/1538-4357/abab95. arXiv: 1911.05506 [astro-ph.SR] (cit. on p. 56).

[78] J. A. Edlund, M. Tinto, A. Królak, and G. Nelemans. "White-dwarf white-dwarf galactic background in the LISA data". In: *Phys. Rev. D* 71.12, 122003 (June 2005), p. 122003. DOI: 10.1103/PhysRevD.71.122003. arXiv: gr-qc/0504112 [gr-qc] (cit. on p. 112).

[79] EPTA Collaboration, InPTA Collaboration, J. Antoniadis, P. Arumugam, and et al. "The second data release from the European Pulsar Timing Array. III. Search for gravitational wave signals". In: *Astron. Astrophys.* 678, A50 (Oct. 2023), A50. DOI: 10.1051/0004-6361/202346844. arXiv: 2306.16214 [astro-ph.HE] (cit. on p. 62).

[80] F Acernese, M. Agathos, A. Ain, S. Albanesi, A. Allocca, A. Amato, T. Andrade, N. Andres, M. Andrés-Carcasona, T. Andrić, et al. "Virgo detector characterization and data quality: tools". In: *Classical and Quantum Gravity* 40.18, 185005 (Sept. 2023), p. 185005. DOI: 10.1088/1361-6382/acdf36. arXiv: 2210.15634 [gr-qc] (cit. on p. 104).

[81] A. J. Farmer and E. S. Phinney. "The gravitational wave background from cosmological compact binaries". In: *Mon. Not. R. Astron. Soc.* 346.4 (Dec. 2003), pp. 1197–1214. DOI: 10.1111/j.1365-2966.2003.07176.x. arXiv: astro-ph/0304393 [astro-ph] (cit. on p. 24).

[82] E. Finch, G. Bartolucci, D. Chucherko, B. G. Patterson, V. Korol, A. Klein, D. Bandopadhyay, H. Middleton, C. J. Moore, and A. Vecchio. "Identifying LISA verification binaries among the Galactic population of double white dwarfs". In: *Mon. Not. R. Astron. Soc.* 522.4 (July 2023), pp. 5358–5373. DOI: 10.1093/mnras/stad1288. arXiv: 2210.10812 [astro-ph.SR] (cit. on pp. 57, 85).

[83] L. Finn, G. Gonzalez, and P. Sutton. "Rayleigh Monitor (Gaussianity test for LIGO data)". In: (2007). DOI: . eprint: . URL: http://gallatin.physics.lsa.umich.edu/~keithr/lscdc/sutton_may01.html (cit. on p. 104).

[84] G. Franciolini, M. Pieroni, A. Ricciardone, and J. D. Romano. "Likelihoods for Stochastic Gravitational Wave Background Data Analysis". In: *arXiv e-prints*, arXiv:2505.24695 (May 2025), arXiv:2505.24695. DOI: 10.48550/arXiv.2505.24695. arXiv: 2505.24695 [gr-qc] (cit. on p. 130).

[85] M. Freitag, P. Amaro-Seoane, and V. Kalogera. "Stellar Remnants in Galactic Nuclei: Mass Segregation". In: *Astrophys. J.* 649.1 (Sept. 2006), pp. 91–117. DOI: 10.1086/506193. arXiv: astro-ph/0603280 [astro-ph] (cit. on p. 105).

[86] J. R. Gair, E. E. Flanagan, S. Drasco, T. Hinderer, and S. Babak. "Forced motion near black holes". In: *Phys. Rev. D* 83 (2011), p. 044037. DOI: 10.1103/PhysRevD.83.044037. arXiv: 1012.5111 [gr-qc] (cit. on pp. 57, 72, 75, 84).

[87] J. R. Gair, C. Tang, and M. Volonteri. "LISA extreme-mass-ratio inspiral events as probes of the black hole mass function". In: *Phys. Rev. D* 81.10, 104014 (May 2010), p. 104014. DOI: 10.1103/PhysRevD.81.104014. arXiv: 1004.1921 [astro-ph.GA] (cit. on p. 64).

[88] J. R. Gair, M. Vallisneri, S. L. Larson, and J. G. Baker. "Testing General Relativity with Low-Frequency, Space-Based Gravitational-Wave Detectors". In: *Living Reviews in Relativity* 16.1, 7 (Dec. 2013), p. 7. DOI: 10.12942/lrr-2013-7. arXiv: 1212.5575 [gr-qc] (cit. on p. 57).

[89] M. Georgousi, N. Karnesis, V. Korol, M. Pieroni, and N. Stergioulas. "Gravitational waves from double white dwarfs as probes of the milky way". In: *Mon. Not. R. Astron. Soc.* 519.2 (Feb. 2023), pp. 2552–2566. DOI: 10.1093/mnras/stac3686. arXiv: 2204.07349 [astro-ph.GA] (cit. on p. 111).

[90] A. Ghoshal and A. Strumia. "Probing the Dark Matter density with gravitational waves from super-massive binary black holes". In: *JCAP* 02 (2024), p. 054. DOI: 10.1088/1475-7516/2024/02/054. arXiv: 2306.17158 [astro-ph.CO] (cit. on p. 72).

[91] G. Giampieri and A. G. Polnarev. "Detecting an anisotropic gravitational wave background with a space-borne interferometer". In: *Mon. Not. R. Astron. Soc.* 291.1 (Oct. 1997), pp. 149–161. DOI: 10.1093/mnras/291.1.149 (cit. on p. 112).

[92] A. W. Graham and N. Scott. "The M $_{BH}$-L $_{spheroid}$ Relation at High and Low Masses, the Quadratic Growth of Black Holes, and Intermediate-mass Black Hole Candidates". In: *Astrophys. J.* 764.2, 151 (Feb. 2013), p. 151. DOI: 10.1088/0004-637X/764/2/151. arXiv: 1211.3199 [astro-ph.CO] (cit. on p. 66).

[93] P. J. Green. "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination". In: *Biometrika* 82.4 (1995), pp. 711–732. ISSN: 00063444. URL: http://www.jstor.org/stable/2337340 (visited on 11/13/2023) (cit. on pp. 88, 95).

[94] K. Gültekin, D. O. Richstone, K. Gebhardt, T. R. Lauer, S. Tremaine, M. C. Aller, R. Bender, A. Dressler, S. M. Faber, A. V. Filippenko, R. Green, L. C. Ho, J. Kormendy, J. Magorrian, J. Pinkney, and C. Siopis. "The M-σ and M-L Relations in Galactic Bulges, and Determinations of Their Intrinsic Scatter". In: *Astrophys. J.* 698.1 (June 2009), pp. 198–221. DOI: 10.1088/0004-637X/698/1/198. arXiv: 0903.4897 [astro-ph.GA] (cit. on p. 65).

[95] T. Hinderer and É. É. Flanagan. "Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime: Orbital motion". In: *Phys. Rev. D* 78.6, 064028 (Sept. 2008), p. 064028. DOI: 10.1103/PhysRevD.78.064028. arXiv: 0805.3337 [gr-qc] (cit. on p. 57).

[96] M. D. Hoffman and A. Gelman. "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo". In: *arXiv e-prints*, arXiv:1111.4246 (Nov. 2011), arXiv:1111.4246. DOI: 10.48550/arXiv.1111.4246. arXiv: 1111.4246 [stat.CO] (cit. on p. 39).

[97] S. Hofman and G. Nelemans. "Uncertainty of the white dwarf astrophysical gravitational wave background". In: *Astron. Astrophys.* 691, A261 (Nov. 2024), A261. DOI: 10.1051/0004-6361/202451510. arXiv: 2407.10642 [astro-ph.HE] (cit. on pp. 61, 127).

[98] C. Hopman and T. Alexander. "The Effect of Mass Segregation on Gravitational Wave Sources near Massive Black Holes". In: *Astrophys. J. Lett.* 645.2 (July 2006), pp. L133–L136. DOI: 10.1086/506273. arXiv: astro-ph/0603324 [astro-ph] (cit. on p. 105).

[99] H. Jeffreys. *The Theory of Probability*. Oxford Classic Texts in the Physical Sciences. OUP Oxford, 1998. ISBN: 9780191589676. URL: https://books.google.it/books?id=vh9Act9rtzQC (cit. on p. 41).

[100] N. Karnesis, M. Lilley, and A. Petiteau. "Assessing the detectability of a stochastic gravitational wave background with LISA, using an excess of power approach". In: *Classical and Quantum Gravity* 37.21, 215017 (Nov. 2020), p. 215017. DOI: 10.1088/1361-6382/abb637. arXiv: 1906.09027 [astro-ph.IM] (cit. on p. 96).

[101] 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". In: *Phys. Rev. D* 104.4, 043019 (Aug. 2021), p. 043019. DOI: 10.1103/PhysRevD.104.043019. arXiv: 2103.14598 [astro-ph.IM] (cit. on pp. 61, 68, 127, 129).

[102] 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". In: *Phys. Rev. D* 104.4, 043019 (Aug. 2021), p. 043019. DOI: 10.1103/PhysRevD.104.043019. arXiv: 2103.14598 [astro-ph.IM] (cit. on pp. 117, 121).

[103] 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". In: *Phys. Rev. D* 104.6, 064047 (Sept. 2021), p. 064047. DOI: 10.1103/PhysRevD.104.064047. arXiv: 2104.04582 [gr-qc] (cit. on p. 23).

[104] M. L. Katz, N. Karnesis, N. Korsakova, J. R. Gair, and N. Stergioulas. "Efficient GPU-accelerated multisource global fit pipeline for LISA data analysis". In: *Phys. Rev. D* 111.2, 024060 (Jan. 2025), p. 024060. DOI: 10.1103/PhysRevD.111.024060. arXiv: 2405.04690 [gr-qc] (cit. on p. 36).

[105] M. A. Keim, V. Korol, and E. M. Rossi. "The large magellanic cloud revealed in gravitational waves with LISA". In: *Mon. Not. R. Astron. Soc.* 521.1 (May 2023), pp. 1088–1098. DOI: 10.1093/mnras/stad554. arXiv: 2207.14277 [astro-ph.GA] (cit. on p. 61).

[106] S. O. Kepler, I. Pelisoli, D. Koester, G. Ourique, S. J. Kleinman, A. D. Romero, A. Nitta, D. J. Eisenstein, J. E. S. Costa, B. Külebi, S. Jordan, P. Dufour, P. Giommi, and A. Rebassa-Mansergas. "New white dwarf stars in the Sloan Digital Sky Survey Data Release 10". In: *Mon. Not. R. Astron. Soc.* 446.4 (Feb. 2015), pp. 4078–4087. DOI: 10.1093/mnras/stu2388. arXiv: 1411.4149 [astro-ph.SR] (cit. on p. 117).

[107] H. Kim and W.-T. Kim. "Dynamical Friction of a Circular-Orbit Perturber in a Gaseous Medium". In: *Astrophys. J.* 665 (2007), pp. 432–444. DOI: 10.1086/519302. arXiv: 0705.0084 [astro-ph] (cit. on pp. 72, 74).

[108] A. Klein, E. Barausse, A. Sesana, A. Petiteau, E. Berti, S. Babak, J. Gair, S. Aoudia, I. Hinder, F. Ohme, and B. Wardell. "Science with the space-based interferometer eLISA: Supermassive black hole binaries". In: *Phys. Rev. D* 93.2, 024003 (Jan. 2016), p. 024003. DOI: 10.1103/PhysRevD.93.024003. arXiv: 1511.05581 [gr-qc] (cit. on p. 64).

[109] A. Klein, G. Pratten, R. Buscicchio, P. Schmidt, C. J. Moore, E. Finch, A. Bonino, L. M. Thomas, N. Williams, D. Gerosa, S. McGee, M. Nicholl, and A. Vecchio. "The last three years: multiband gravitational-wave observations of stellar-mass binary black holes". In: *arXiv e-prints*, arXiv:2204.03423 (Apr. 2022), arXiv:2204.03423. DOI: 10.48550/arXiv.2204.03423. arXiv: 2204.03423 [astro-ph.HE] (cit. on p. 85).

[110] J. Kormendy and L. C. Ho. "Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies". In: *Ann. Rev. Astron. Astrophys.* 51.1 (Aug. 2013), pp. 511–653. DOI: 10.1146/annurev-astro-082708-101811. arXiv: 1304.7762 [astro-ph.CO] (cit. on p. 66).

[111] V. Korol, N. Hallakoun, S. Toonen, and N. Karnesis. "Observationally driven Galactic double white dwarf population for LISA". In: *Mon. Not. R. Astron. Soc.* 511.4 (Apr. 2022), pp. 5936–5947. DOI: 10.1093/mnras/stac415. arXiv: 2109.10972 [astro-ph.HE] (cit. on pp. 61, 116, 117).

[112] V. Korol, A. Igoshev, S. Toonen, N. Karnesis, C. Moore, E. Finch, and A. Klein. *Neutron Star - White Dwarf Binaries: Probing Formation Pathways and Natal Kicks with LISA*. Zenodo, Mar. 2024. DOI: 10.5281/zenodo.10854469. URL: https://doi.org/10.5281/zenodo.10854469 (cit. on p. 121).

[113] V. Korol, A. P. Igoshev, S. Toonen, N. Karnesis, C. J. Moore, E. Finch, and A. Klein. "Neutron star - white dwarf binaries: probing formation pathways and natal kicks with LISA". In: *Mon. Not. R. Astron. Soc.* 530.1 (May 2024), pp. 844–860. DOI: 10.1093/mnras/stae889. arXiv: 2310.06559 [astro-ph.HE] (cit. on p. 121).

[114] A. Kuntz and K. Leyde. "Transverse Doppler effect and parameter estimation of LISA three-body systems". In: *Phys. Rev. D* 108.2 (2023), p. 024002. DOI: 10.1103/PhysRevD.108.024002. arXiv: 2212.09753 [gr-qc] (cit. on p. 72).

[115] L. D. Landau and E. M. Lifshitz. *Mechanics*. 3rd. Vol. 1. Course of Theoretical Physics. Oxford: Pergamon Press, 1976. ISBN: 9780750628969 (cit. on p. 75).

[116] N. Langer and C. A. Norman. "On the collapsar model of long gamma-ray bursts: constraints from cosmic metallicity evolution". In: *Astrophys. J. Lett.* 638 (2006), pp. L63–L66. DOI: 10.1086/500363. arXiv: astro-ph/0512271 (cit. on pp. 72, 144).

[117] A. Le Tiec and J. Novak. "Theory of Gravitational Waves". In: *An Overview of Gravitational Waves: Theory*. 2017, pp. 1–41. DOI: 10.1142/9789813141766_0001 (cit. on p. 16).

[118] LIGO Scientific Collaboration et al. "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7, 074001 (Apr. 2015), p. 074001. DOI: 10.1088/0264-9381/32/7/074001. arXiv: 1411.4547 [gr-qc] (cit. on p. 105).

[119] J. Lin. "Divergence measures based on the Shannon entropy". In: *IEEE Transactions on Information Theory* 37.1 (1991), pp. 145–151. DOI: 10.1109/18.61115 (cit. on p. 131).

[120] LISA Consortium Waveform Working Group, N. Afshordi, S. Akçay, P. Amaro Seoane, and et al. "Waveform Modelling for the Laser Interferometer Space Antenna". In: *arXiv e-prints*, arXiv:2311.01300 (Nov. 2023), arXiv:2311.01300. DOI: 10.48550/arXiv.2311.01300. arXiv: 2311.01300 [gr-qc] (cit. on p. 72).

[121] *LISA Data Challenge 1b: Yorsh*. https://sbgvm-151-90.in2p3.fr/challenge1b (cit. on pp. 125, 130, 138).

[122] LISA Science Study Team. *LISA Science Requirements Document*. Tech. rep. 1.0. ESA, May 2018. URL: https://www.cosmos.esa.int/web/lisa/lisa-documents (cit. on p. 147).

[123] T. B. Littenberg and N. J. Cornish. "Prototype global analysis of LISA data with multiple source types". In: *Phys. Rev. D* 107.6, 063004 (Mar. 2023), p. 063004. DOI: 10.1103/PhysRevD.107.063004. arXiv: 2301.03673 [gr-qc] (cit. on p. 36).

[124] T. B. Littenberg and A. K. Lali. "Have any LISA verification binaries been found?" In: *arXiv e-prints*, arXiv:2404.03046 (Apr. 2024), arXiv:2404.03046. DOI: 10.48550/arXiv.2404.03046. arXiv: 2404.03046 [astro-ph.HE] (cit. on p. 57).

[125] P. Madau and M. Dickinson. "Cosmic Star Formation History". In: *Ann. Rev. Astron. Astrophys.* 52 (2014), pp. 415–486. DOI: 10.1146/annurev-astro-081811-125615. arXiv: 1403.0007 [astro-ph.CO] (cit. on pp. 72, 144).

[126] M. Maggiore. *Gravitational Waves. Vol. 1: Theory and Experiments*. Also available as ISBN 978-0-19-171766-6. Oxford University Press, 2007. ISBN: 978-0-19-852074-0. DOI: 10.1093/acprof:oso/9780198570745.001.0001 (cit. on p. 26).

[127] D. Maiwald and D. Kraus. "On moments of complex Wishart and complex inverse Wishart distributed matrices". In: *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 5. 1997, 3817–3820 vol.5. DOI: 10.1109/ICASSP.1997.604712 (cit. on p. 98).

[128] A. Mangiagli, C. Caprini, M. Volonteri, S. Marsat, S. Vergani, N. Tamanini, and H. Inchauspé. "Massive black hole binaries in LISA: Multimessenger prospects and electromagnetic counterparts". In: *Phys. Rev. D* 106.10, 103017 (Nov. 2022), p. 103017. DOI: 10.1103/PhysRevD.106.103017. arXiv: 2207.10678 [astro-ph.HE] (cit. on p. 56).

[129] D. Maoz, C. Badenes, and S. J. Bickerton. "Characterizing the Galactic White Dwarf Binary Population with Sparsely Sampled Radial Velocity Data". In: *Astrophys. J.* 751.2, 143 (June 2012), p. 143. DOI: 10.1088/0004-637X/751/2/143. arXiv: 1202.5467 [astro-ph.SR] (cit. on pp. 116, 117).

[130] D. Maoz, N. Hallakoun, and C. Badenes. "The separation distribution and merger rate of double white dwarfs: improved constraints". In: *Mon. Not. R. Astron. Soc.* 476.2 (May 2018), pp. 2584–2590. DOI: 10.1093/mnras/sty339. arXiv: 1801.04275 [astro-ph.SR] (cit. on pp. 116, 117).

[131] M. Mapelli, N. Giacobbo, F. Santoliquido, and M. C. Artale. "The properties of merging black holes and neutron stars across cosmic time". In: *Mon. Not. Roy. Astron. Soc.* 487.1 (2019), pp. 2–13. DOI: 10.1093/mnras/stz1150. arXiv: 1902.01419 [astro-ph.HE] (cit. on p. 144).

[132] A. W. McConnachie. "The Observed Properties of Dwarf Galaxies in and around the Local Group". In: *Astron. J.* 144.1, 4 (July 2012), p. 4. DOI: 10.1088/0004-6256/144/1/4. arXiv: 1204.1562 [astro-ph.CO] (cit. on p. 117).

[133] M. van de Meent. "Modelling EMRIs with gravitational self-force: a status report". In: *Journal of Physics: Conference Series* 840.1 (May 2017), p. 012022. DOI: 10.1088/1742-6596/840/1/012022. URL: https://doi.org/10.1088/1742-6596/840/1/012022 (cit. on p. 23).

[134] G. Mentasti, C. R. Contaldi, and M. Peloso. "Probing the galactic and extragalactic gravitational wave backgrounds with space-based interferometers". In: *J. Cosmology Astropart. Phys.* 2024.6, 055 (June 2024), p. 055. DOI: 10.1088/1475-7516/2024/06/055. arXiv: 2312.10792 [gr-qc] (cit. on p. 111).

[135] S. Miller, T. A. Callister, and W. Farr. "The Low Effective Spin of Binary Black Holes and Implications for Individual Gravitational-Wave Events". In: *Astrophys. J.* 895.2 (2020), p. 128. DOI: 10.3847/1538-4357/ab80c0. arXiv: 2001.06051 [astro-ph.HE] (cit. on pp. 72, 144).

[136] M. Moe and R. Di Stefano. "Mind Your Ps and Qs: The Interrelation between Period (P) and Mass-ratio (Q) Distributions of Binary Stars". In: *Astrophys. J. Supp.* 230.2, 15 (June 2017), p. 15. DOI: 10.3847/1538-4365/aa6fb6. arXiv: 1606.05347 [astro-ph.SR] (cit. on p. 117).

[137] P. artist Moran and P. Whittle. "Hypothesis Testing in Time Series Analysis." In: 1951. URL: https://api.semanticscholar.org/CorpusID:125739077 (cit. on p. 36).

[138] M. Muratore, J. Gair, and L. Speri. "Impact of the noise knowledge uncertainty for the science exploitation of cosmological and astrophysical stochastic gravitational wave background with LISA". In: *Phys. Rev. D* 109.4, 042001 (Feb. 2024), p. 042001. DOI: 10.1103/PhysRevD.109.042001. arXiv: 2308.01056 [gr-qc] (cit. on p. 96).

[139] G. Nelemans. "The Galactic gravitational wave foreground". In: *Classical and Quantum Gravity* 26.9, 094030 (May 2009), p. 094030. DOI: 10.1088/0264-9381/26/9/094030. arXiv: 0901.1778 [astro-ph.SR] (cit. on pp. 60, 111).

[140] E. C. Ostriker. "Dynamical friction in a gaseous medium". In: *Astrophys. J.* 513 (1999), p. 252. DOI: 10.1086/306858. arXiv: astro-ph/9810324 (cit. on pp. 72, 74).

[141] C. Pacilio and R. Buscicchio. "Functional inference on deviations from General Relativity". In: *arXiv e-prints*, arXiv:2507.13454 (July 2025), arXiv:2507.13454. DOI: 10.48550/arXiv.2507.13454. arXiv: 2507.13454 [gr-qc] (cit. on p. 95).

[142] P. C. Peters. "Gravitational Radiation and the Motion of Two Point Masses". In: *Physical Review* 136.4B (Nov. 1964), pp. 1224–1232. DOI: 10.1103/PhysRev.136.B1224 (cit. on pp. 21, 57).

[143] P. C. Peters and J. Mathews. "Gravitational Radiation from Point Masses in a Keplerian Orbit". In: *Physical Review* 131.1 (July 1963), pp. 435–440. DOI: 10.1103/PhysRev.131.435 (cit. on pp. 23, 103).

[144] D. Phan, N. Pradhan, and M. Jankowiak. "Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro". In: *arXiv e-prints*, arXiv:1912.11554 (Dec. 2019), arXiv:1912.11554. DOI: 10.48550/arXiv.1912.11554. arXiv: 1912.11554 [stat.ML] (cit. on pp. 39, 130).

[145] E. S. Phinney. "A Practical Theorem on Gravitational Wave Backgrounds". In: *arXiv e-prints*, astro-ph/0108028 (Aug. 2001), astro–ph/0108028. DOI: 10.48550/arXiv.astro-ph/0108028. arXiv: astro-ph/0108028 [astro-ph] (cit. on p. 115).

[146] E. Phinney. "A practical theorem on gravitational wave backgrounds". In: *arXiv preprint astro-ph/0108028* (2001) (cit. on pp. 25, 26).

[147] M. Piarulli, R. Buscicchio, F. Pozzoli, O. Burke, and et al. "Test for LISA foreground Gaussianity and stationarity: Extreme mass-ratio inspirals". In: *Phys. Rev. D* 111.10, 103047 (May 2025), p. 103047. DOI: 10.1103/nfn4-pgr5. arXiv: 2410.08862 [astro-ph.HE] (cit. on p. 103).

[148] F. Pozzoli, S. Babak, A. Sesana, M. Bonetti, and N. Karnesis. "Computation of stochastic background from extreme-mass-ratio inspiral populations for LISA". In: *Phys. Rev. D* 108.10, 103039 (Nov. 2023), p. 103039. DOI: 10.1103/PhysRevD.108.103039. arXiv: 2302.07043 [astro-ph.GA] (cit. on pp. 63, 104–106, 109).

[149] F. Pozzoli, R. Buscicchio, A. Klein, and D. Chirico. "Bahamas: BAyesian inference with HAmiltonian Montecarlo for Astrophysical Stochastic background". In: *arXiv e-prints*, arXiv:2506.22542 (June 2025), arXiv:2506.22542. DOI: 10.48550/arXiv.2506.22542. arXiv: 2506.22542 [astro-ph.IM] (cit. on pp. 103, 130).

[150] F. Pozzoli, R. Buscicchio, A. Klein, V. Korol, A. Sesana, and F. Haardt. "Cyclostationary signals in LISA: a practical application to Milky Way satellites". In: *arXiv e-prints*, arXiv:2410.08274 (Oct. 2024), arXiv:2410.08274. DOI: 10.48550/arXiv.2410.08274. arXiv: 2410.08274 [astro-ph.GA] (cit. on pp. 103, 110).

[151] F. Pozzoli, R. Buscicchio, C. J. Moore, F. Haardt, and A. Sesana. "Weakly parametric approach to stochastic background inference in LISA". In: *Phys. Rev. D* 109.8, 083029 (Apr. 2024), p. 083029. DOI: 10.1103/PhysRevD.109.083029. arXiv: 2311.12111 [astro-ph.CO] (cit. on p. 63).

[152] F. Pozzoli, R. Buscicchio, C. J. Moore, F. Haardt, and A. Sesana. "Weakly parametric approach to stochastic background inference in LISA". In: *Phys. Rev. D* 109.8, 083029 (Apr. 2024), p. 083029. DOI: 10.1103/PhysRevD.109.083029. arXiv: 2311.12111 [astro-ph.CO] (cit. on p. 127).

[153] F. Pozzoli, J. Gair, R. Buscicchio, and L. Speri. "Is your stochastic signal really detectable?" In: (Dec. 2024). arXiv: 2412.10468 [astro-ph.IM] (cit. on p. 63).

[154] G. Pratten, A. Klein, C. J. Moore, H. Middleton, N. Steinle, P. Schmidt, and A. Vecchio. "LISA science performance in observations of short-lived signals from massive black hole binary coalescences". In: *Phys. Rev. D* 107.12, 123026 (June 2023), p. 123026. DOI: 10.1103/PhysRevD.107.123026. arXiv: 2212.02572 [gr-qc] (cit. on p. 85).

[155] D. Quang Nam, J. Martino, Y. Lemière, A. Petiteau, J.-B. Bayle, O. Hartwig, and M. Staab. "Time-delay interferometry noise transfer functions for LISA". In: *Phys. Rev. D* 108.8, 082004 (Oct. 2023), p. 082004. DOI: 10.1103/PhysRevD.108.082004. arXiv: 2211.02539 [gr-qc] (cit. on p. 128).

[156] D. J. Reardon, A. Zic, R. M. Shannon, G. B. Hobbs, and et al. "Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array". In: *Astrophys. J. Lett.* 951.1, L6 (July 2023), p. L6. DOI: 10.3847/2041-8213/acdd02. arXiv: 2306.16215 [astro-ph.HE] (cit. on p. 62).

[157] S. Rieck, A. W. Criswell, V. Korol, M. A. Keim, M. Bloom, and V. Mandic. "A stochastic gravitational wave background in LISA from unresolved white dwarf binaries in the Large Magellanic Cloud". In: *Mon. Not. R. Astron. Soc.* 531.2 (June 2024), pp. 2642–2652. DOI: 10.1093/mnras/stae1283. arXiv: 2308.12437 [astro-ph.IM] (cit. on p. 111).

[158] E. Roebber, R. Buscicchio, A. Vecchio, C. J. Moore, A. Klein, V. Korol, S. Toonen, D. Gerosa, J. Goldstein, S. M. Gaebel, and T. E. Woods. "Milky Way Satellites Shining Bright in Gravitational Waves". In: *Astrophys. J. Lett.* 894.2, L15 (May 2020), p. L15. DOI: 10.3847/2041-8213/ab8ac9. arXiv: 2002.10465 [astro-ph.GA] (cit. on pp. 61, 114, 120, 122).

[159] J. D. Romano and N. J. Cornish. "Detection methods for stochastic gravitational-wave backgrounds: a unified treatment". In: *Living Reviews in Relativity* 20.1, 2 (Dec. 2017), p. 2. DOI: 10.1007/s41114-017-0004-1. arXiv: 1608.06889 [gr-qc] (cit. on p. 106).

[160] P. A. Rosado, A. Sesana, and J. Gair. "Expected properties of the first gravitational wave signal detected with pulsar timing arrays". In: *Mon. Not. R. Astron. Soc.* 451.3 (Aug. 2015), pp. 2417–2433. DOI: 10.1093/mnras/stv1098. arXiv: 1503.04803 [astro-ph.HE] (cit. on p. 99).

[161] R. Rosati and T. B. Littenberg. "Prototype Stochastic Gravitational Wave Background Recovery in the LISA Global Fit Residual". In: *arXiv e-prints*, arXiv:2410.17180 (Oct. 2024), arXiv:2410.17180. DOI: 10.48550/arXiv.2410.17180. arXiv: 2410.17180 [gr-qc] (cit. on p. 107).

[162] D. von Rosen. "Moments for the Inverted Wishart Distribution". In: *Scandinavian Journal of Statistics* 15.2 (1988), pp. 97–109. ISSN: 03036898, 14679469. URL: http://www.jstor.org/stable/4616090 (visited on 11/12/2024) (cit. on p. 98).

[163] F. D. Ryan. "Gravitational waves from the inspiral of a compact object into a massive, axisymmetric body with arbitrary multipole moments". In: *Phys. Rev. D* 52.10 (Nov. 1995), pp. 5707–5718. DOI: 10.1103/PhysRevD.52.5707 (cit. on p. 57).

[164] C. Salomon, D. Hils, and J. Hall. "Laser stabilization at the millihertz level". In: *Journal of the Optical Society of America B* 5 (Aug. 1988), pp. 1576–1587. DOI: 10.1364/JOSAB.5.001576 (cit. on p. 52).

[165] A. Santini, M. Muratore, J. Gair, and O. Hartwig. "A flexible, GPU-accelerated approach for the joint characterization of LISA instrumental noise and Stochastic Gravitational Wave Backgrounds". In: *arXiv e-prints*, arXiv:2507.06300 (July 2025), arXiv:2507.06300. DOI: 10.48550/arXiv.2507.06300. arXiv: 2507.06300 [gr-qc] (cit. on p. 85).

[166] A. Sasli, N. Karnesis, and N. Stergioulas. "Heavy-tailed likelihoods for robustness against data outliers: Applications to the analysis of gravitational wave data". In: *Phys. Rev. D* 108.10, 103005 (Nov. 2023), p. 103005. DOI: 10.1103/PhysRevD.108.103005. arXiv: 2305.04709 [gr-qc] (cit. on p. 110).

[167] L. Sberna et al. "Observing GW190521-like binary black holes and their environment with LISA". In: *Phys. Rev. D* 106.6 (2022), p. 064056. DOI: 10.1103/PhysRevD.106.064056. arXiv: 2205.08550 [gr-qc] (cit. on p. 72).

[168] P. A. Seoane et al. "Astrophysics with the Laser Interferometer Space Antenna". In: *Living Rev. Rel.* 26.1 (2023), p. 2. DOI: 10.1007/s41114-022-00041-y. arXiv: 2203.06016 [gr-qc] (cit. on p. 72).

[169] A. Sesana. "A practical guide to the massive black hole cosmic history". In: *Adv. Astron.* 2012 (2012), p. 805402. DOI: 10.1155/2012/805402. arXiv: 1110.6445 [astro-ph.CO] (cit. on p. 64).

[170] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi. "Linking the Spin Evolution of Massive Black Holes to Galaxy Kinematics". In: *Astrophys. J.* 794.2, 104 (Oct. 2014), p. 104. DOI: 10.1088/0004-637X/794/2/104. arXiv: 1402.7088 [astro-ph.CO] (cit. on p. 64).

[171] A. Sesana. "Prospects for Multiband Gravitational-Wave Astronomy after GW150914". In: *Phys. Rev. Lett.* 116.23, 231102 (June 2016), p. 231102. DOI: 10.1103/PhysRevLett.116.231102. arXiv: 1602.06951 [gr-qc] (cit. on p. 57).

[172] A. Sesana and F. M. Khan. "Scattering experiments meet N-body - I. A practical recipe for the evolution of massive black hole binaries in stellar environments". In: *Mon. Not. R. Astron. Soc.* 454.1 (Nov. 2015), pp. L66–L70. DOI: 10.1093/mnrasl/slv131. arXiv: 1505.02062 [astro-ph.GA] (cit. on p. 56).

[173] D. A. Shaddock. "Operating LISA as a Sagnac interferometer". In: *Phys. Rev. D* 69.2, 022001 (Jan. 2004), p. 022001. DOI: 10.1103/PhysRevD.69.022001. arXiv: gr-qc/0306125 [gr-qc] (cit. on p. 53).

[174] J. D. Simon. "The Faintest Dwarf Galaxies". In: *Ann. Rev. Astron. Astrophys.* 57 (Aug. 2019), pp. 375–415. DOI: 10.1146/annurev-astro-091918-104453. arXiv: 1901.05465 [astro-ph.GA] (cit. on p. 117).

[175] J. Skilling. "Nested Sampling". In: *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Ed. by R. Fischer, R. Preuss, and U. V. Toussaint. Vol. 735. American Institute of Physics Conference Series. AIP, Nov. 2004, pp. 395–405. DOI: 10.1063/1.1835238 (cit. on pp. 38, 131).

[176] A. Spadaro, R. Buscicchio, D. Izquierdo–Villalba, D. Gerosa, A. Klein, and G. Pratten. "Stars or gas? Constraining the hardening processes of massive black-hole binaries with LISA". In: *Phys. Rev. D* 111.2, 023004 (Feb. 2025), p. 023004. DOI: 10.1103/PhysRevD.111.023004. arXiv: 2409.13011 [astro-ph.HE] (cit. on p. 85).

[177] A. Spadaro, R. Buscicchio, D. Vetrugno, A. Klein, D. Gerosa, S. Vitale, R. Dolesi, W. J. Weber, and M. Colpi. "Glitch systematics on the observation of massive black-hole binaries with LISA". In: *Phys. Rev. D* 108.12, 123029 (Dec. 2023), p. 123029. DOI: 10.1103/PhysRevD.108.123029. arXiv: 2306.03923 [gr-qc] (cit. on p. 85).

[178] L. Speri, A. Antonelli, L. Sberna, S. Babak, E. Barausse, J. R. Gair, and M. L. Katz. "Probing Accretion Physics with Gravitational Waves". In: *Phys. Rev. X* 13.2 (2023), p. 021035. DOI: 10.1103/PhysRevX.13.021035. arXiv: 2207.10086 [gr-qc] (cit. on p. 72).

[179] J. Spitzer Lyman and M. H. Hart. "Random Gravitational Encounters and the Evolution of Spherical Systems. I. Method". In: *Astrophys. J.* 164 (Mar. 1971), p. 399. DOI: 10.1086/150855 (cit. on p. 65).

[180] S. Staelens and G. Nelemans. "Likelihood of white dwarf binaries to dominate the astrophysical gravitational wave background in the mHz band". In: *Astron. Astrophys.* 683, A139 (Mar.

2024), A139. DOI: 10.1051/0004-6361/202348429. arXiv: 2310.19448 [astro-ph.HE] (cit. on pp. 62, 127).

[181] S. H. Strub, L. Ferraioli, C. Schmelzbach, S. C. Stähler, and D. Giardini. "Global analysis of LISA data with Galactic binaries and massive black hole binaries". In: *Phys. Rev. D* 110.2, 024005 (July 2024), p. 024005. DOI: 10.1103/PhysRevD.110.024005. arXiv: 2403.15318 [gr-qc] (cit. on p. 36).

[182] C. Talbot and E. Thrane. "Measuring the binary black hole mass spectrum with an astrophysically motivated parameterization". In: *Astrophys. J.* 856.2 (2018), p. 173. DOI: 10.3847/1538-4357/aab34c. arXiv: 1801.02699 [astro-ph.HE] (cit. on pp. 72, 145).

[183] L. D. C. Team. *LISA Data Challenges Manual: Sangria*. Accessed: 2025-06-18. 2025. URL: https://lisa-ldc.lal.in2p3.fr/static/data/pdf/LDC-manual-Sangria.pdf (cit. on p. 47).

[184] E. Thrane and J. D. Romano. "Sensitivity curves for searches for gravitational-wave backgrounds". In: *Phys. Rev. D* 88.12, 124032 (Dec. 2013), p. 124032. DOI: 10.1103/PhysRevD.88.124032. arXiv: 1310.5300 [astro-ph.IM] (cit. on pp. 96, 99).

[185] M. Tinto and S. V. Dhurandhar. "Time-Delay Interferometry". In: *Living Reviews in Relativity* 8.1, 4 (Dec. 2005), p. 4. DOI: 10.12942/lrr-2005-4 (cit. on pp. 112, 119).

[186] M. Tinto and S. V. Dhurandhar. "Time-delay interferometry". In: *Living Reviews in Relativity* 24.1, 1 (Dec. 2021), p. 1. DOI: 10.1007/s41114-020-00029-6. arXiv: gr-qc/0409034 [gr-qc] (cit. on pp. 37, 52).

[187] M. Toscani, O. Burke, C. Liu, N. B. Zamel, N. Tamanini, and F. Pozzoli. "Strongly lensed extreme mass-ratio inspirals". In: *Phys. Rev. D* 109.6, 063505 (Mar. 2024), p. 063505. DOI: 10.1103/PhysRevD.109.063505. arXiv: 2307.06722 [astro-ph.CO] (cit. on p. 57).

[188] A. Toubiana, N. Karnesis, A. Lamberts, and M. C. Miller. "The interacting double white dwarf population with LISA: Stochastic foreground and resolved sources". In: *Astron. Astrophys.* 692, A165 (Dec. 2024), A165. DOI: 10.1051/0004-6361/202450174. arXiv: 2403.16867 [astro-ph.SR] (cit. on p. 61).

[189] A. Toubiana et al. "Detectable environmental effects in GW190521-like black-hole binaries with LISA". In: *Phys. Rev. Lett.* 126.10 (2021), p. 101105. DOI: 10.1103/PhysRevLett.126.101105. arXiv: 2010.06056 [astro-ph.HE] (cit. on pp. 72, 74, 75).

[190] A. Toubiana, K. W. K. Wong, S. Babak, E. Barausse, E. Berti, J. R. Gair, S. Marsat, and S. R. Taylor. "Discriminating between different scenarios for the formation and evolution of massive black holes with LISA". In: *Phys. Rev. D* 104.8, 083027 (Oct. 2021), p. 083027. DOI: 10.1103/PhysRevD.104.083027. arXiv: 2106.13819 [gr-qc] (cit. on p. 56).

[191] M. Vallisneri. "Synthetic LISA: Simulating time delay interferometry in a model LISA". In: *Phys. Rev. D* 71.2, 022001 (Jan. 2005), p. 022001. DOI: 10.1103/PhysRevD.71.022001. arXiv: gr-qc/0407102 [gr-qc] (cit. on p. 52).

[192] N. Van Kampen. "Chapter III - Stochastic Processes". In: *Stochastic Processes in Physics and Chemistry (Third Edition)*. Ed. by N. Van Kampen. Third Edition. North-Holland Personal Library. Amsterdam: Elsevier, 2007, pp. 52–72. DOI: https://doi.org/10.1016/B978-044452965-7/50006-4. URL: https://www.sciencedirect.com/science/article/pii/B9780444529657500064 (cit. on p. 111).

[193] J. P. W. Verbiest and G. M. Shaifullah. "Measurement uncertainty in pulsar timing array experiments". In: *Classical and Quantum Gravity* 35.13, 133001 (July 2018), p. 133001. DOI: 10.1088/1361-6382/aac412 (cit. on p. 37).

[194] Virgo Scientific Collaboration et al. "Advanced Virgo: a second-generation interferometric gravitational wave detector". In: *Classical and Quantum Gravity* 32.2, 024001 (Jan. 2015), p. 024001. DOI: 10.1088/0264-9381/32/2/024001. arXiv: 1408.3978 [gr-qc] (cit. on p. 105).

[195] P. Welch. “The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms”. In: *IEEE Transactions on Audio and Electroacoustics* 15.2 (1967), pp. 70–73. DOI: 10.1109/TAU.1967.1161901 (cit. on p. 126).

[196] P. D. Welch. “The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms”. In: *IEEE Trans. Audio & Electroacoust* 15 (Jan. 1967), pp. 70–73 (cit. on p. 104).

[197] M. J. Williams, J. Veitch, and C. Messenger. “Nested sampling with normalizing flows for gravitational-wave inference”. In: *Phys. Rev. D* 103.10, 103006 (May 2021), p. 103006. DOI: 10.1103/PhysRevD.103.103006. arXiv: 2102.11056 [gr-qc] (cit. on pp. 38, 130).

[198] H. Xu et al. “Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I”. In: *Res. Astron. Astrophys.* 23.7 (2023), p. 075024. DOI: 10.1088/1674-4527/acdfa5. arXiv: 2306.16216 [astro-ph.HE] (cit. on p. 72).

[199] H. Xu, S. Chen, Y. Guo, J. Jiang, and et al. “Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I”. In: *Research in Astronomy and Astrophysics* 23.7, 075024 (July 2023), p. 075024. DOI: 10.1088/1674-4527/acdfa5. arXiv: 2306.16216 [astro-ph.HE] (cit. on p. 62).

[200] N. Yunes, M. Coleman Miller, and J. Thornburg. “The Effect of Massive Perturbers on Extreme Mass-Ratio Inspiral Waveforms”. In: *Phys. Rev. D* 83 (2011), p. 044030. DOI: 10.1103/PhysRevD.83.044030. arXiv: 1010.1721 [astro-ph.GA] (cit. on p. 72).

[201] E. M. Zahraoui, P. Maturana-Russel, W. van Straten, R. Meyer, and S. Gulyaev. “Generalized stepping-stone sampling: efficient marginal likelihood estimation in gravitational wave analysis of pulsar timing array data”. In: *Mon. Not. R. Astron. Soc.* 540.4 (July 2025), pp. 3818–3825. DOI: 10.1093/mnras/staf953. arXiv: 2411.14736 [astro-ph.IM] (cit. on pp. 43, 131).