

Pipeline for searching and fitting instrumental glitches in LISA data

Martina Muratore^{1,*}, Jonathan Gair¹, Olaf Hartwig^{2,3}, Michael L. Katz⁴, and Alexandre Toubiana^{5,6}

¹*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), D-14476 Potsdam, Germany*

²*Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut),
Callinstraße 38, 30167 Hannover, Germany*

³*Leibniz Universität Hannover, Institut für Gravitationsphysik,
Callinstraße 38, 30167 Hannover, Germany*

⁴*NASA Marshall Space Flight Center, Huntsville, Alabama 35811, USA*

⁵*Dipartimento di Fisica “G. Occhialini,” Università degli Studi di Milano-Bicocca,
Piazza della Scienza 3, 20126 Milano, Italy*

⁶*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*



(Received 28 May 2025; accepted 29 July 2025; published 22 September 2025)

Instrumental artifacts, such as glitches, can significantly compromise the scientific output of the Laser Interferometer Space Antenna (LISA). Our methodology employs advanced Bayesian techniques, including reversible jump Markov chain Monte Carlo and parallel tempering to find and characterize glitches and astrophysical signals. The robustness of the pipeline is demonstrated through its ability to simultaneously handle diverse glitch morphologies and it is validated with a “Spritz”-type dataset from the LISA Data Challenge. Our approach enables accurate inference on massive black hole binaries, while simultaneously characterizing both instrumental artifacts and noise. These results present a significant development in strategies for differentiating between instrumental noise and astrophysical signals, which will ultimately improve the accuracy and reliability of source population analyses with LISA.

DOI: [10.1103/PhysRevD.112.063041](https://doi.org/10.1103/PhysRevD.112.063041)

I. INTRODUCTION

The space-based gravitational wave detector, the Laser Interferometer Space Antenna (LISA) mission [1,2] was adopted as a mission by the European Space Agency in January 2024 and is scheduled for launch in 2035. One of its main objectives is the first direct observation of coalescing massive black hole binaries (MBHBs) with gravitational waves (GWs). Such observations will transform our understanding of the astrophysical processes leading to the formation of massive black holes and driving their coevolution with their host galaxies [3–13]. Moreover, these observations will reach signal-to-noise ratios (SNRs) of thousands, allowing for exquisite tests of general relativity [14–22]. Finally, the possible observation of electromagnetic counterparts [23–25] would allow MBHBs to be used as “bright sirens” for the measurement of cosmological parameters, not only the Hubble constant, but possibly also the equation of state of dark energy, thanks to the wide redshift

range of these observations [26–28]. To achieve these science objectives, it is crucial to develop a robust data analysis pipeline for LISA.

Unlike current ground-based GW detectors such as LIGO [29], Virgo [30], and KAGRA [31,32]. LISA will be a signal-dominated detector with multiple GW signals overlapping in time and frequency. Indeed, in addition to MBHBs, signals from $\sim 10^7$ Galactic binaries are expected to be present in the data stream, of which $\sim 10^4$ will be individually resolved, and the rest will form a stochastic noise that will dominate over the instrumental noise around 1 mHz [33–38]. The extreme-mass-ratio inspirals (EMRIs) of stellar-origin compact objects into massive black holes in the centers of galaxies will also generate signals in the LISA data that could last for the whole mission duration. The rate of EMRI events is very uncertain, ranging from a few to tens of thousands [39–43], and at the upper end of this range EMRIs could perhaps even form a stochastic foreground [44,45]. Finally, the loudest and most massive of the stellar-origin black hole binary population being observed by ground-based detectors are also expected to be observable during their inspiral phase [46–52] and a variety of cosmological backgrounds could significantly contribute to the total noise (see [53] and references therein). Moreover, we will not have a direct and independent measurement of the instrument noise [54], meaning that we will have to estimate its properties simultaneously with

*Contact author: martina.muratore@aei.mpg.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

the parameters of the astrophysical sources. The simultaneous characterization of all the sources of multiple different types that are expected to be present in the LISA data, along with instrumental properties, is usually referred to as the LISA global fit [55–59].

When designing a methodology to characterize instrumental noise, it is crucial to avoid systematic errors arising from noise misinterpretation or mismodeling. The LISA precursor mission, LISA Pathfinder (LPF) [60], which shares technology with LISA, was affected by spurious signals of unknown origin, referred to as glitches. Since various types of glitches were detected in LPF [61], it is reasonable to expect that, unless they are fully understood and mitigated through hardware requirements, similar glitches may appear in LISA. The LISA hardware, particularly the gravitational reference sensors (GRS), will be largely the same as in LPF, with only minor modifications to accommodate integration into the movable optical subassembly and the LISA spacecraft [62]. Therefore, we anticipate encountering glitches similar to those seen in LPF—especially force impulses—which could impact LISA’s sensitivity.

Glitches are also commonly seen in the data collected by ground-based gravitational wave detectors [63]. These glitches show a complex range of morphologies, in contrast to the glitches observed in LPF, which were all of a much simpler form [64]. These differences most likely arise from the difference in the technology of the instruments and in their different operating environments. However, given the differences in the mission environment and duration from LISA and LPF, we should remain open to the possibility that LISA may encounter previously unobserved or more complex types of glitches. Regarding signals observed in ground-based detectors, these are typically very short in duration compared to their typical separation in time. For this reason, segments of data flagged as having glitches are normally excluded from scientific analysis. Some events, for example, the first binary neutron star merger, GW170817 [65], were clearly identified as being astrophysical in origin, due to good data quality in one of the instruments, while having overlapping glitches in one or more of the other detectors. In those cases, the analysis either gated out the range of times when the glitch was present or fitted out the glitch using a flexible model [66]. In the case of LISA, signal durations will be typically much longer than the average time between glitches, and so only the latter strategies will be available to us.

In this paper, we present a possible solution to address these types of artifacts—particularly force glitches acting on the test masses—within the framework of the global fit development for LISA. Our analysis is performed by using the open-source sampler *Eryn*, taking advantage of its reversible jump (RJ) capabilities [67,68]. We demonstrate the feasibility of performing both searches and parameter estimation to fit multiple glitches of different shapes,

morphology, and SNR along with instrumental noise and an astrophysical source, in this case, an MBHB.

With this methodology, we aim to contribute to the ongoing efforts to develop robust data analysis techniques for LISA, ensuring its capacity to distinguish between astrophysical signals and instrumental artifacts. We therefore show the following achievements:

- (1) Effective search, parameter estimation and fit of multiple glitches while recovering the LISA instrumental noise.
- (2) Methodology to distinguish between a glitch and an MBHB signal.
- (3) Preservation of parameter estimation accuracy for an MBHB signal in noisy data in the presence of one or multiple glitches, injected at various times—either before, after, or during the time of maximum overlap between the MBHB signal and the glitch.

This paper is organized as follows: in Sec. II we describe the LPF main results in terms of measured and cataloged artifacts; in Sec. III we introduce the models used in this paper for the instrumental noise, glitches, and the MBHB signal; in Sec. IV, we provide details on the Bayesian formalism used, including the likelihood and the sampling methodology; in Sec. V, we present an example where we generate our own data containing only glitches and instrumental noise. This demonstrates the algorithm’s ability to identify and fit multiple artifacts with different SNRs at the same time, while distinguishing them from stationary instrumental noise. Then, in Sec. VI, we analyze the impact of glitches with different SNRs on MBHB parameter estimation when left unmodeled and assess the recovery of both glitch and MBHB parameters when performing a joint fit. Finally, in Sec. VII, we apply our methodology to a “light-Spritz” data challenge [69]. We address it as light Spritz because in this analysis we ignore the nonstationarities caused by gaps and Galactic binaries present in the dataset, as these are subjects of a follow-up study. Additionally, we replace the MBHB signal included in this dataset with one generated using the PhenomHM [70] waveform model from the BBH_x software package [71], in order to eliminate systematics arising from differences between the simulated signal and the model used for the fit. We conclude the paper in Sec. IX with a summary of our main results and perspectives for future work.

II. GLITCHES AS INSTRUMENTAL ARTIFACTS

LPF successfully demonstrated the feasibility of maintaining free-falling test masses with residual acceleration noise below the stringent requirements for LISA [72]. Operating from 2016 to 2017, LPF achieved unprecedented noise levels; however, it also showed that measured(-experimental) noise could deviate quite significantly with respect to the derived noise models. In particular, the measured test-mass (TM) acceleration noise at frequencies below 10^{-4} is still not fully understood [72,73]. Also, transient acceleration

glitches were observed during the mission [74,75]. They spanned a wide amplitude range—transferring impulses from femtonewton seconds ($\text{fN} \cdot \text{s}$) to nanonewton seconds ($\text{nN} \cdot \text{s}$) to the test masses—and exhibited diverse durations, from seconds to hours. These glitches fell into two main categories: rapid transients in the interferometric readout and longer-lasting forces acting on the test masses. Glitches in the second category arose from spurious impulsive forces of unknown origin acting on the test masses, which occurred approximately once per day, following a Poisson distribution. These impulses exhibited a distinctive time evolution, with a peak followed by an exponential decay over several hours [61]. A detailed analysis of these phenomena showed that the observed signals could be accurately described using a fitted model [74], with minimal residual deviations. The strength of these impulsive events can be quantified through their SNR, which measures the detectability of a signal and will be formally defined later in the text in Sec. III E 1.

Out of the 432 events detected, 98 occurred during “ordinary runs,” characterized by standard operational conditions (22 or 11 °C). The SNR distribution of glitches that occurred during ordinary runs peaked around moderate values of 10. In contrast, more glitches (334) were observed during the “cold runs” performed in May 2017, during which the instrument was operated at about 0°. All glitches observed during the cold runs were short and intense and had higher SNR values (see Fig. 7.12 in [74]). We refer the reader to [74] for a more detailed discussion of the possible origin of this phenomenon.

In this paper, we focus primarily on glitches with morphology similar to those observed in the ordinary runs since LISA will operate at standard operational conditions. These exhibit a well defined and predictable pattern. Indeed, the analysis of LPF glitches leveraged phenomenological modeling, as single or double exponentials with a certain decay time and amplitude [74]. Recently, [76] has shown that shapelet functions could be used to fit LPF glitches and, therefore, characterize their properties. These models, even if not directly connected to physical models as the ones derived in [74,75], laid the groundwork for future efforts to mitigate noise artifacts in LISA that are agnostic to the physical properties of the signals, such as the glitch decay time and impulse (amplitude), which determine their shape.

To conclude this section, we note that one of the key challenges in GW data analysis is to develop methods that can robustly separate astrophysical signals from instrumental artifacts. To address this, we first need to derive a fiducial glitch model to use in our analysis. Therefore, in the next section, we present the shapelet function used to model the glitches and we relate it to the physical glitch model used to fit the observed LPF glitches. We also describe the waveform model used to describe the MBHB signal and the time-delay interferometry (TDI) technique, which is applied to the raw data as a preprocessing step to

reduce laser frequency noise before parameter estimation can be performed.

III. TIME-DELAY INTERFEROMETRY, GLITCHES, MASSIVE BLACK HOLE BINARIES, AND NOISE MODELS

A. Time-delay interferometry

TDI is used in LISA to suppress laser frequency noise, which would otherwise exceed the GW signal by 8 orders of magnitude, due to the inequality of the interferometer arms. The LISA mission comprises three identical spacecraft arranged in a triangular configuration, separated by 2.5×10^6 km and linked by six active laser connections. While the LISA satellite formation is designed to remain roughly equilateral, gravitational and orbital dynamics cause the arm lengths to fluctuate by approximately 1.5% annually, with relative velocity drifts between spacecraft reaching up to 10 m/s [1,2]. These time-varying and unequal arm lengths introduce laser frequency noise into the measurement channels. TDI achieves laser frequency noise reduction by forming specific linear combinations of time-delayed interspacecraft phase measurements [77,78]. Many combinations can be constructed [79]; however, for this work, we use the three noise-orthogonal channels A , E , and T [80], which are constructed from the Michelson TDI combinations X , Y , and Z , as these are the data streams available in Spritz.

Generally, when we refer to first-generation TDI [77], this means measurement combinations that are constructed to cancel laser frequency noise for a static, but unequal arm length, constellation. In contrast, for second-generation TDI, time-varying and unequal arm lengths are taken into account [81,82]. In this paper, while we consider both generations, we assume in both cases equal and nonvarying arm lengths. Specifically, the first-generation TDI is used for generating and analyzing frequency data and modeling the noise and glitches in Sec. V. Whereas, the second-generation TDI is used here for modeling the noise, signal, and glitches in the analysis of the light-Spritz Data Challenge [69] in Sec. VII.

Note that by assuming equal-arm lengths, uncorrelated noise, and identical optical metrology system (OMS) and TM acceleration noise terms, the combinations X , Y , and Z can be further transformed into three noise-orthogonal channels A , E , and T . This orthogonality simplifies the likelihood computation—in the frequency domain—by not having cross-correlation terms between channels. We will further discuss in the Conclusion (Sec. IX) the effect if these assumptions are dropped.

The channels A , E , and T [83,84] that will be used here are

$$A = \frac{Z - X}{\sqrt{2}}, \quad E = \frac{X - 2Y + Z}{\sqrt{6}}, \quad T = \frac{X + Y + Z}{\sqrt{3}}. \quad (1)$$

In the case of first-generation TDI, the channel X is defined as

$$\begin{aligned} X_1 = & (1 - D_{12}D_{21}) \times (\eta_{13} + D_{13}\eta_{31}) \\ & - (1 - D_{13}D_{31}) \times (\eta_{12} + D_{12}\eta_{21}), \end{aligned} \quad (2)$$

and the same variable in second generation is defined as

$$\begin{aligned} X_2 = & (1 - D_{12}D_{21} - D_{12}D_{21}D_{13}D_{31} \\ & + D_{13}D_{31}D_{12}D_{21}D_{12}D_{21}) \times (\eta_{13} + D_{13}\eta_{31}) \\ & - (1 - D_{13}D_{31} - D_{13}D_{31}D_{12}D_{21} \\ & + D_{12}D_{21}D_{13}D_{31}D_{13}D_{31}) \times (\eta_{12} + D_{12}\eta_{21}). \end{aligned} \quad (3)$$

The Y and Z variables differ from X by cyclic permutations of the three satellites. The operator D_{pq} is a delay operator, corresponding to a constant time shift by L_{pq} and thus in frequency to multiplication by $\mathcal{F}\{D_{pq}\} = e^{-i2\pi f L_{pq}}$ with L_{pq} representing the light travel time for a beam received at spacecraft p and emitted from spacecraft q , in seconds [85]. In this work we will assume the six arm lengths to be equal and constant, and therefore $L_{pq} = L$. The η_{pq} are the so-called intermediary variables [86]. They represent a synthesized interferometric TM to TM measurement, here expressed as a fractional frequency shift between the interfering laser beams. The first index p indicates the spacecraft where the measurement is performed at time t , and the second index q indicates the distant spacecraft from which light was emitted at time $t - L_{pq}$. The intermediary variables η_{pq} contain the sum of the contributions from GW signals, glitches, and instrumental noises, and so the TDI transfer function applies linearly to these different components.

B. Shapelet model

In this section, we describe how glitches are modeled and the LISA response to them. We model glitches using the exponential shapelet functions following [76,87]. This is a family of exponentially damped functions that are eigenwave-functions of the normalized one-dimensional (1D) hydrogen atom. They can be defined in the time domain as follows:

$$\psi_n(t) = 2 \frac{t}{n} e^{-\frac{t}{n}} L_{n-1}^1 \left(2 \frac{t}{n} \right) \Theta(t), \quad (4)$$

with L_n^α denoting the generalized Laguerre polynomials and $\Theta(t)$ is the Heaviside step function. A glitch perturbation is then modeled by a linear combination of shapelets,

$$g(t) = \sum_{j=0}^P A_j \psi_{n_j} \left(\frac{t - \tau_j}{\beta_j} \right), \quad (5)$$

with $P + 1$ being the number of shapelets. The parameters to be estimated for each shapelet component are the scale

parameter β_j , which represents the characteristic damping time, the starting (or injection) time of the shapelet τ_j , and the shapelet amplitude A_j . Note that Eq. (5) is given in units of acceleration, but considering that the phase meter in LISA outputs phase or frequency, we must convert it to the appropriate units. To achieve this, we utilize the generating function (see [88] for reference and Appendix A 1) and integrate Eq. (5)—after substituting Eq. (4)—to derive the fractional frequency expression for a single exponential shapelet with $n_0 = 1$ as

$$\begin{aligned} \frac{\Delta \nu_g}{\nu_0}(t, \tau_0, \beta_0, A_0) \\ = \frac{2A_0\beta_0}{c} \left(1 - \frac{e^{-\frac{t+\tau_0}{\beta_0}}(t + \beta_0 - \tau_0)}{\beta_0} \right) \Theta \left(\frac{t - \tau_0}{\beta_0} \right), \end{aligned} \quad (6)$$

where we have considered that for our case $\alpha = 1$ [76] and ν_0 is the nominal laser frequency equal to 2.816×10^{14} Hz. The shapelet model zero order depends thus on three parameters $\phi = \{\tau_0, A_0, \beta_0\}$. We note that Eq. (6) is the same as the one derived in [76]. This function is similar to a step function starting from a value of zero and increasing to

$$\lim_{t \rightarrow \infty} \frac{\Delta \nu_g}{\nu_0}(t, \tau_0, \beta_0, A_0) = \frac{2A_0\beta_0}{c}. \quad (7)$$

The shapelet model with $j = 0$ thus reduces to the way glitches were modeled in LPF. Although, since the most significant impact of glitches during the LPF mission was an effective overestimation of the measured LPF acceleration noise, they had been fitted in terms of acceleration [Eq. (7.1) in [74] and Appendix A 2] using the following model:

$$h_g(t) = \frac{\Delta v}{\tau^2} t' e^{-\frac{t'}{\tau}} \Theta(t'), \quad t' = t - t_0 \quad (8)$$

which, integrated in time to get the fractional frequency change, gives

$$\int_0^t \frac{h_g(t')}{c} dt' = \frac{\Delta v}{c} \left(\frac{e^{-\frac{t-t_0}{\tau}}(-t - \tau + t_0)}{\tau} + 1 \right) \Theta(t - t_0). \quad (9)$$

Δv is the total transferred impulse per unit mass, t_0 is the injection time, and τ is the decay time. Therefore, comparing the shapelet model [Eq. (6)] with the glitch model [Eq. (9)], the shapelet amplitude A_0 can be related to the glitch impulse Δv according to $\Delta v = 2A_0\beta_0$, while β_0 equals the glitch τ , and τ_0 equals the glitch injection time t_0 . These relationships will be important in Sec. V, where we will use them to define the priors on the shapelet for fitting the glitches present in the data. Specifically, the prior for the shapelet amplitude will be chosen based on this relationship such that smaller amplitude values will be allowed compared to the nominal LPF glitch amplitude.

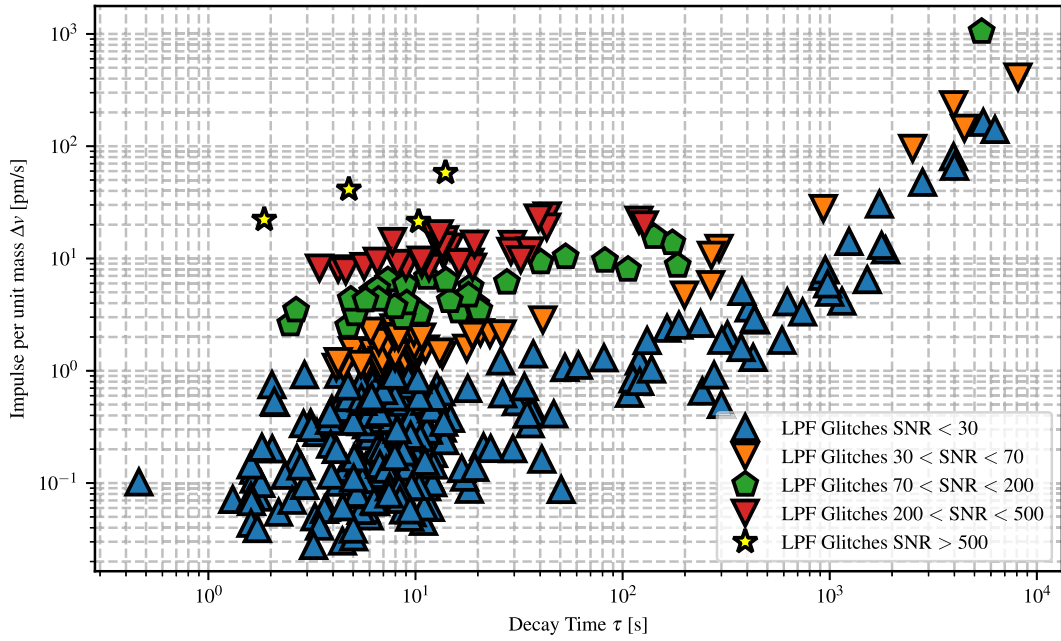


FIG. 1. Impulse versus decay time of LPF glitches at different SNRs computed considering the LISA noise power spectral densities and the three TDI channels A , E , and T .

Figure 1 shows the distribution of glitch parameters observed for LPF glitches. It is clear that relatively short glitches have the highest SNR. This is because short glitches have the majority of the power concentrated in the middle of the LISA frequency band, whereas long glitches concentrate the power at very low frequencies, where LISA is less sensitive. We note that SNRs for LISA are higher with respect to LPF (Fig. 1 versus Fig. 8.8 in [74]) for two reasons: the first is the different acceleration noise transfer functions between LISA and LPF, in particular, LISA shows a factor 4 enhancement for the X channels in the TM acceleration noise power spectral density (PSD) (Fig. 8.8 in [74]); the second is that we compute the SNR for LISA assuming the three TDI channels (A , E , and T) such that the SNR squared is doubled with respect to that in a single channel.¹

We also show in Fig. 2 two examples of shapelets with $j = 0$ and $j = 9$.

The shapelet expression [Eq. (6)] can be written directly in the frequency domain as follows² (see Appendix A 1):

$$\frac{\Delta\tilde{\nu}_g}{\nu_0}(f, \tau_0, \beta_0, A_0) = \frac{iA_0\beta_0 e^{-2if\pi\tau_0}}{cf\pi(i - 2f\pi\beta_0)^2}. \quad (10)$$

¹The TDI channel T , or in general any channel null to GWs, being less sensitive to test-mass acceleration noise is also less sensitive to acceleration glitches. Therefore, their contribution to the SNR is negligible.

²We have tested that using the analytical frequency-domain function is sufficiently equivalent to performing a numerical fast Fourier transform of the time-domain data.

This expression agrees with the physical model used to fit glitches in LPF data but expressed in fractional frequency units [Eq. (7.3) in [74] and Appendix A 2].

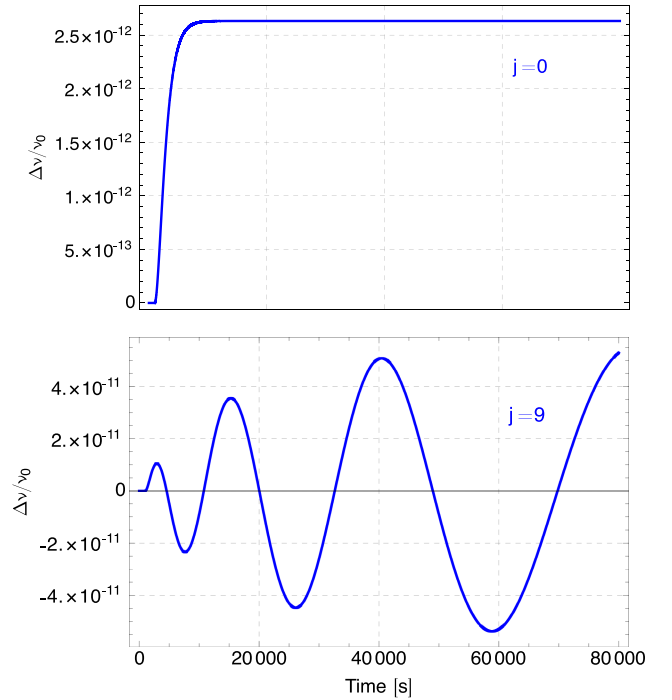


FIG. 2. Shapelets with $n_0 = 1$ (upper plot) and $n_9 = 10$ (lower plot), where we have considered for both $A_0 = 5 \times 10^{-7}$, $\beta_0 = 1000$ and $\tau_0 = 1000$. For the high-order shapelet we assume all A_j , β_j , and τ_j to be equal to A_0 , β_0 , and τ_0 , respectively.

C. TDI transfer function for glitches and shapelets in frequency domain

We can express the single link measurement $\tilde{\eta}_{pq}(\omega)$ in Eqs. (2) and (3) as

$$\tilde{\eta}_{pq}(f) = \Delta\tilde{\nu}_{g_{qp}}(f)e^{-2i\pi fL} + \Delta\tilde{\nu}_{g_{pq}}(f), \quad (11)$$

where the term $\Delta\tilde{\nu}_{g_{pq}}(f)$ indicates the contribution of a glitch (shapelet) acting on the test mass pq , and we have assumed constant and equal-arm lengths. Considering the expected glitch distributions from LPF, it is highly unlikely that multiple glitches happen simultaneously. We therefore assume that one of the two terms [$\Delta\tilde{\nu}_{g_{qp}}(f)$ or $\Delta\tilde{\nu}_{g_{pq}}(f)$] vanishes and that the other one contains a single glitch.

Considering now Eq. (2), we can model the TDI applied to a glitch (shapelet) as a product between the glitch (shapelet) expression in the frequency domain [Eq. (10)] and the transfer functions $\mathcal{T}_{X,Y,Z}^g(f)$ which take into account TDI, i.e., $\tilde{g}_{X,Y,Z}(f) = \mathcal{T}_{X,Y,Z}^g(f)\Delta\tilde{\nu}_g(f)/\nu_0$, where $\tilde{g}_{X,Y,Z}(f)$ is the signal of a glitch (shapelet) in the TDI frequency-domain data (and we have omitted the constant glitch parameters in the argument of the expression).

In this work we will always apply glitches on TM_{12} . Denoting by \mathcal{F} the Fourier transform operator and applying it to the delay operators gives $\mathcal{F}D = e^{-2i\pi fL}$. The transfer function for the X , Y , and Z first-generation TDI variables of a glitch on TM_{12} is

$$\mathcal{T}_{X_1}^g(f) = -1 + e^{-8iL\pi f} \quad (12a)$$

$$\mathcal{T}_{Y_1}^g(f) = 4ie^{-4iL\pi f} \sin(2L\pi f) \quad (12b)$$

$$\mathcal{T}_{Z_1}^g(f) = 0. \quad (12c)$$

The same type of derivation can be done for second-generation TDI using Eq. (3),

$$\mathcal{T}_{X_2}^g(f) = e^{-16iL\pi f}(-1 + e^{8iL\pi f})^2 \quad (13a)$$

$$\mathcal{T}_{Y_2}^g(f) = 8e^{-8iL\pi f} \sin(4L\pi f) \sin(2L\pi f) \quad (13b)$$

$$\mathcal{T}_{Z_2}^g(f) = 0. \quad (13c)$$

Note that, in the case of multiple glitches happening at the same time in different test masses, their effects sum up linearly. Expressions for the transfer functions for glitches occurring on the other test masses can be found in Appendix A 2.

D. Waveform model for massive black hole binaries

To model the full inspiral-merger-ringdown signal of an MBHB, we use the PhenomHM [70] waveform model from the BBHx software package [71]. This is an analytic frequency-domain model for quasicircular and spin-aligned

binaries that includes harmonics beyond the dominant (2,2) mode. The contribution of these harmonics increases with the asymmetry of the system (i.e., unequal mass ratios) and the magnitude of the spins. We do expect it to be possible to observe large mass ratio and highly spinning MBHBs with LISA [7,8], and not including higher harmonics could result in severe biases [89]. The GW signal of an MBHB, for the spin-aligned case, depends on 11 parameters $\theta = \{M_t, q, a_1, a_2, d_L, \varphi_{\text{ref}}, i, \lambda_d, \beta_d, \psi, t_{\text{ref}}\}$. They can be split as follows:

- (1) Intrinsic parameters:
 - (a) m_1 is the mass of the first black hole and m_2 of the second, with $m_2 \leq m_1$. We express the masses as dimensionless multiples of the solar mass M_\odot .
 - (b) a_1, a_2 are dimensionless spin magnitudes of the two black holes, aligned or antialigned with the orbital angular momentum.
- (2) Extrinsic parameters:
 - (a) d_L : luminosity distance to the source, scaling the amplitude of the GW signal;
 - (b) φ_{ref} : reference phase at the merger;
 - (c) i : inclination, the angle between the orbital angular momentum and the line of sight, affecting the observed amplitude of higher-order modes;
 - (d) λ_d, β_d : ecliptic longitude and latitude, describing the sky location of the source in the (LISA) detector reference frame;
 - (e) ψ : polarization angle of the waveform, affecting the projection of the signal onto the detector; and
 - (f) t_{ref} : reference time marking a characteristic point in the binary coalescence, such as the merger.

All extrinsic parameters discussed here are defined relative to the LISA constellation reference frame. We also introduce the total mass $M_t = m_1 + m_2$ and the mass ratio $q = m_2/m_1 \leq 1$.

The frequency-domain, source-frame waveform is expressed as [90,91]

$$\tilde{h}_{lm}(f) = A_{lm}(f)e^{-i\phi_{lm}(f)}, \quad (14)$$

where (l, m) represent the harmonic mode indices. The TDI variables are obtained as

$$\tilde{h}_{X,Y,Z}(f) = \sum_{lm} \mathcal{T}_{X,Y,Z}^h(f, t_{lm}(f)) \frac{\tilde{h}_{lm}(f)}{d_L}, \quad (15)$$

where $\mathcal{T}_{X,Y,Z}^h(f, t_{lm}(f))$ are the frequency-domain transfer functions detailed in [90,91]. The transfer functions depend on the extrinsic parameters and vary both in time and frequency due to the LISA constellation's orbital motion. They are evaluated at time $t_{lm}(f)$, which is given by the stationary-phase approximation

$$t_{lm}(f) = t_{\text{ref}} - \frac{1}{2\pi} \frac{d\phi_{lm}(f)}{df}. \quad (16)$$

The phase-meter output (LISA observables) will therefore be the sum of GW signals, $\tilde{h}_{X,Y,Z}(f)$, and glitches, $\tilde{g}_{X,Y,Z}(f)$, propagated through TDI.

E. Noise models and definition of the signal-to-noise ratio

In this work, we assume that primary noise sources such as laser frequency noise, spacecraft longitudinal jitters, clock noises, and tilt-to-length couplings have been corrected and only leave negligible residuals in the postprocessed TDI variables [92–94]. What remains are the so-called secondary noises. These remaining noises [95] fall into two broad categories: residual acceleration noise of each TM and the OMS noise terms limiting the TM position readout. Noises inside these categories can have various physical origins, different levels, and different transfer functions into the final TDI variables [96]. For the analysis that will follow in Sec. V, we just consider two main components modeled as uncorrelated noise on each TM and uncorrelated noise in each interspacecraft optical link, with identical statistical properties for noises of the same kind. We will instead consider a slightly more elaborate noise model for analyzing the light-Spritz dataset, in Sec. VII B.

We characterize these noise sources by their PSD and a corresponding transfer function that describes their propagation in the TDI variables. The TM acceleration noise PSD is

$$S_{\text{TM}}(f) = (N_{\text{TM}})^2 \left(1 + \left(\frac{0.4 \times 10^{-3}}{f} \right)^2 \right) \times \left(1 + \left(\frac{f}{8 \times 10^{-3}} \right)^4 \right), \quad (17)$$

with $N_{\text{TM}} = 2.4 \text{ fm/s}^2/\sqrt{\text{Hz}}$.

The PSD of the displacement noise from the OMS is

$$S_{\text{OMS}}(f) = (N_{\text{OMS}})^2 \left(1 + \left(\frac{2.0 \times 10^{-3}}{f} \right)^4 \right), \quad (18)$$

with $N_{\text{OMS}} = 7.9 \text{ pm}/\sqrt{\text{Hz}}$. This overall OMS noise level is based on the values provided in the LISA Data Challenge (LDC) tools [97].

Note that the noise parameters, $\gamma = \{N_{\text{TM}}, N_{\text{OMS}}\}$, are expressed as equivalent acceleration noise in $\text{m}^2/\text{s}^4/\text{Hz}$ or displacement noise in $\text{m}^2 \text{ Hz}^{-1}$. For our analysis, we instead use units of fractional frequency deviations, to which we will convert them by multiplying the components with $C_{\text{acc}}^{\text{ffd}} = (1/(c2\pi f))^2$ or $C_{\text{disp}}^{\text{ffd}} = (2\pi f/c)^2$, respectively. Here, c is the speed of light in ms^{-1} .

For the first-generation TDI, the noise in the A , E , and T channels is

$$S_A(f) = S_E(f) = 8\sin^2(2\pi Lf) \times [2C_{\text{acc}}^{\text{ffd}} S_{\text{TM}}(3 + 2\cos(2\pi Lf) + \cos(4\pi Lf)) + C_{\text{disp}}^{\text{ffd}} S_{\text{OMS}}(2 + \cos(2\pi Lf))], \quad (19)$$

and

$$S_T(f) = 16C_{\text{disp}}^{\text{ffd}} S_{\text{OMS}}(1 - \cos(2\pi Lf))\sin^2(2\pi Lf) + 128C_{\text{acc}}^{\text{ffd}} S_{\text{TM}}\sin^2(2\pi Lf)\sin^4(\pi Lf). \quad (20)$$

Where, again, L denotes the LISA arm length in seconds.

1. Definition of the signal-to-noise ratio

The SNRs of glitches and MBHB signals are computed based on the three TDI channels using the following formula [Eq. (119) in [98,99]]:

$$\text{SNR} = \left[4\Re \sum_f |\tilde{Q}(f)|^2 \left(\frac{|\tilde{T}_A(f)|^2}{S_A(f)} + \frac{|\tilde{T}_E(f)|^2}{S_E(f)} + \frac{|\tilde{T}_T(f)|^2}{S_T(f)} \right) \Delta f \right]^{1/2}, \quad (21)$$

where Δf is the frequency resolution which is equal to $1/T_{\text{obs}} = 1/(Ndt)$, with T_{obs} the observation time, N the length of the time series, and dt is the inverse of the sampling rate. Note that the above formula is valid in this case as we assume the three TDI channels to be uncorrelated. Here $\tilde{T}_{A/E/T}(f)$ are the transfer functions of a GW signal or a glitch into the combinations A , E , and T , $\tilde{Q}(f)$ is the amplitude of the GW signal [$\tilde{h}_{lm}(f)/d_L$] or glitch [$\Delta\tilde{v}_g(f)/\nu_0$] in Fourier frequency domain, and $S_{A/E/T}(f)$ are the spectral densities of the noises in the A , E , T combinations, respectively. All these quantities have been defined in the previous sections.

IV. BAYESIAN METHODOLOGY AND LIKELIHOOD FORMULATION

The Bayesian framework provides a powerful approach to the problem we aim to solve by treating the detection and parameter estimation of GW signals, instrumental artifacts and noise as an inference problem. In this context, the observed data $d(t)$ is modeled as the sum of contributions from the GW signal $h(t, \theta)$, the glitch $g(t, \phi)$, and the noise $n(t, \gamma)$,

$$d(t, \theta, \phi, \gamma) = h(t, \theta) + g(t, \phi) + n(t, \gamma), \quad (22)$$

where θ , ϕ , and γ represent the parameters of the GW signal, the glitch, and the noise model, respectively. Denoting by $\mu = (\theta, \phi, \gamma)$ the set of all parameters, we write $p(\mu|d, M)$ as the posterior distribution of the parameters. We note M the assumed model, which includes M_1 ,

our chosen shapelet model for the glitches, M_2 , our model for the MBHB signal, and M_3 , our model for the instrumental noise. The posterior distribution $p(\mu|d, M)$ is computed using the Bayes theorem [67,100],

$$p(\mu|d, M) = \frac{p(d|\mu, M)p(\mu|M)}{p(d|M)}, \quad (23)$$

where

- (i) $p(d|\mu, M)$ is the likelihood, quantifying the probability of observing data d under the models M_1 , M_2 , and M_3 with parameters θ , ϕ , and γ ;
- (ii) $p(\mu|M)$ is the prior distribution for the GW, glitch and noise parameters under the models M_1 , M_2 , and M_3 : $p(\theta|M_1)$, $p(\phi|M_2)$, and $p(\gamma|M_3)$; and
- (iii) $p(d|M)$ is the evidence, which acts as a normalization constant for parameter estimation, but can also be used to compare different models M .

For the problem at hand, we will use the Whittle likelihood, for which the log-likelihood in the frequency domain is

$$\log \mathcal{L} = -\frac{1}{2} \sum_f \left[4\Delta f \frac{|\tilde{d}(f) - \tilde{h}(f, \theta) - \tilde{g}(f, \phi)|^2}{S(f, \gamma)} + \log S(f, \gamma) \right], \quad (24)$$

where $\tilde{d}(f)$, $\tilde{h}(f, \theta)$, and $\tilde{g}(f, \phi)$ are the Fourier transforms of the same quantities appearing in Eq. (22); $S(f, \gamma)$ is the one-sided noise PSD, parametrized by γ .

The posterior can be sampled using advanced Bayesian inference techniques such as Markov chain Monte Carlo (MCMC) [101–104]. In this analysis, we wish to fit simultaneously an unknown number of glitches. Therefore, we use a reversible jump MCMC (RJ-MCMC) [105] algorithm, a generalization of the Metropolis-Hastings algorithm that allows us to vary the dimensionality of the parameter space while sampling over the parameters. We use the RJ-MCMC implementation algorithm of the Eryn Sampler [67]. Several walkers are run in parallel and exchange information periodically in order to increase the sampling efficiency. We also use parallel tempering [106–108] to improve the exploration of multimodalities in the posterior distribution. During the run, a new point μ' is proposed from a point μ according to a proposal function $q(\mu'|\mu)$ and is accepted with probability

$$\alpha = \min \left(1, \frac{p(\mu'|d, M)q(\mu|\mu')}{p(\mu|d, M)q(\mu'|\mu)} \right). \quad (25)$$

In our analysis, we use different proposals for the parameters of the MBHB (θ), of the instrumental noise (γ), and of the glitches (ϕ). They are updated successively, using the Gibbs sampling implementation of Eryn [67]. In the Eryn “language,”

branches represent different types of models used to fit the data. Each branch contains leaves, which correspond to individual instances of each model. In our case, leaves represent individual glitches, massive black hole binaries, or noise components, which form different branches [67].

A. Moves

In the RJ-MCMC Eryn sampler, there are two types of moves:

- (i) In-model moves: changes are made to the parameters, keeping the number of leaves in each branch fixed.
- (ii) Out-of-model moves: the dimensionality is varied by adding or removing leaves, in one or several branches.

For instance, in our case, adding or removing glitches. In this work, only the number of glitches is varied, so that the latter type of moves applies only to their parameters. Detailed explanations of the various moves available in Eryn, as well as the most commonly used ones in the literature, can be found in [67,71]. Here, we focus on the in-model moves specifically developed for glitch searches, MBHB parameter estimation, and noise parameter estimation. For the out-of-model move for glitches, we propose new parameters from the prior distribution during the search phase. Instead, during the final parameter estimation phase, as explained in the upcoming section, we propose sampling from a Gaussian distribution, whose covariance matrix is constructed from an initial estimate of the parameters of the detected glitches during the search phase.

1. Glitch search and parameter estimation (in-model moves)

Group stretch move. In the stretch move, new samples are generated by leveraging the distribution of points in the ensemble, using the formula

$$Y = X_j + z(X_k - X_j), \quad (26)$$

where X_k is the current sample, X_j is another point in the ensemble, z is a scaling factor drawn from a predefined distribution, and Y is the proposed new value. In the RJ-MCMC algorithm, however, the variable dimensionality of models complicates the direct use of ensemble-based proposals. To deal with the variable dimensionality of models, the group stretch move is used [67], which builds upon the affine-invariant stretch proposal [109]. This group introduces “friends,” a static subset of reference points that are defined after a first burn of the sampler and as soon as the group move is used. These friends simulate the posterior distribution for a given model and remain fixed throughout multiple iterations. For instance, in our case, we group the estimated glitches into “friend groups” according to their mean time value. The goal is to set X_j properly for each posterior mode such that X_j and X_k are always found on the same mode of the posterior. This approach improves

efficiency by ensuring that proposed points correspond to the posterior distribution of the currently sampled glitch, rather than a different one.

The stretch move is used 50% of the time, and in the implementation of this move, the set of friends associated with each glitch is tracked,³ to improve the efficiency of the RJ-MCMC algorithm to add and remove glitches [67,68].

Selected covariance group move. The “selected covariance group move” replaces the standard Gaussian move. In the context of nested models, the posterior distribution may display distinct modes, corresponding to different glitches in the data. This complexity makes it challenging to efficiently tune covariance matrices for each mode. The core idea behind this move is to group glitches based on their mean injection time values and compute corresponding distributions for glitch parameters using a Gaussian mixture model (GMM) [110]. This approach leverages covariance matrices specific to each glitch group. When glitches are clearly distinguishable, the group proposal becomes highly efficient, aided by the branch supplemental object in `Eryn` [71] to maintain the association between glitches and their respective covariance matrices. Note that we leverage the “fix friends” function from `Eryn` when a proposed birth of a new glitch is accepted. This function assigns a covariance matrix (or friends) to the new glitch from the precomputed friends group.

The computation (update) of the covariance matrices for this move is done as follows:

- (1) Cluster specification: Every 500 iterations, we evaluate the fraction of samples for which the number of active glitch leaves is zero—i.e., samples where no glitch components are present. Physically, this corresponds to samples where the data are explained entirely by the noise model, without requiring any glitch contribution [67]. If this fraction is sufficiently bigger than zero, it indicates that a significant portion of the current samples includes one or more active glitch components—suggesting that the sampler interprets the data as containing non-Gaussian features such as glitches. In this case, a GMM is initialized using the posterior samples, with the number of mixture components set equal to the maximum number of glitch leaves allowed by the model. This enables the sampler to model complex posterior structures as a combination of multiple Gaussian components. Each component in the GMM is assigned a full-rank covariance matrix and is intended to represent a distinct glitch cluster, with the number of components constrained by the dimensionality of the glitch model.

³To best use this move, we leverage the branch supplemental object that has the goal of tracking the friends’ information associated with each glitch. This is particularly needed in the presence of RJ-MCMC and parallel tempering, where we have swaps in temperature or birth/death of new glitches.

- (2) Covariance calculation: Based on the identified number of clusters, a GMM is fitted to the data. The GMM provides the mean and covariance for each cluster, capturing the local structure of the posterior around each glitch.⁴ These parameters are then utilized in the in-model move, enabling it to propose moves that are well aligned with the posterior’s structure. This approach ensures the sampler’s efficiency, particularly when the shape of the posterior varies significantly between glitches, warranting the use of distinct covariance matrices for each glitch. By iteratively estimating the covariance matrices from the samples via a GMM, and adapting the sampler to the posterior structure, this methodology establishes a robust framework suitable for glitch searches. Note that adaptation is applied for the first 50000 iterations—every 500 iterations—and then stopped to preserve detailed balance and ensure the Markovianity of the move. The covariance group move is used 50% of the time.

2. Glitch search and parameter estimation (out-of-model moves)

During the parameter estimation phase, we incorporate a move proposal mechanism based on detected glitches. Specifically, 70% of the time, the glitch parameters are drawn from a mixture of independent Gaussian distributions describing the shapelet parameter τ_0 , $\ln(A_0)$, and β_0 . Each parameter is sampled from a multivariate normal distribution with the corresponding mean and covariance of the previously detected shapelets. We propose from the prior distribution the other 30% of the time.

3. Noise parameter estimation (in-model moves)

We consider both the stretch and Gaussian proposal moves [71]. Initially, the Gaussian move uses a diagonal covariance matrix with small entries (following a suggestion from the `Eryn` tutorial). Then, the covariance matrix is reestimated using the noise samples every 500 iterations,

⁴In a GMM, each Gaussian component does not necessarily correspond to an independent cluster. If a posterior distribution is not very Gaussian in shape, more than one component may be needed to approximate it. Therefore, identifying a GMM component is not the same as identifying a true cluster represented by the posterior of a single glitch. However, for the problem analyzed here, we have verified that the glitch posteriors are sufficiently Gaussian and well separated, so our assumptions remain valid. Dividing a true cluster into more subclusters than are actually present in the data leads to proposal steps that are smaller than necessary. This can result in a high acceptance rate but a low effective sample size, meaning the sampler requires more steps to converge. To improve sampling efficiency, this proposal method should be refined in future work. One development we have already begun testing is using x-means instead of k-means [111] to automatically determine the optimal number of GMM components.

and the adaptation is done for the first 50,000 iterations to preserve detailed balance. The stretch move is used 85% of the time, while the Gaussian move is used 15%.

4. Massive black hole binaries parameter estimation (in-model moves)

It is known that, at low frequencies, MBHB posteriors have an eightfold sky-mode degeneracy. This is reduced to a twofold latitudinal degeneracy at higher frequency. To ensure efficient sampling of the sky-mode posterior, following Ref. [112], we therefore use a sky-mode hopping proposal in addition to the stretch proposal described above. The stretch proposal is used in approximately 80% of proposals. We also add a Gaussian move proposal used 5% of the time with a small covariance matrix. This matrix is then recursively estimated from the MBHB samples with the same cadence used to estimate the noise covariance matrix. We notice that this additional move helps the convergence of the algorithm in possible cases where X_k and X_j [Eq. (26)] are very close in value, as they will spread out slowly otherwise if we only use the stretch move. Instead, the covariance Gaussian move will speed this up.

Regarding the sky-mode proposal, when sampling in the LISA reference frame [91], as we do here, the eight distinct sky modes consist of four longitudinal modes at

$$\{\lambda_d + (0, 1, 2, 3)\pi/2, \psi + (0, 1, 2, 3)\pi/2\}$$

and two latitudinal modes

$$\{\pm\beta_d, \pm\cos\iota, \pm\cos\psi\}.$$

Three varieties of this proposal are used:

- (i) One is to move the walker to a mode drawn randomly with replacement from all eight sky modes. This is used for approximately 4% of proposals.
- (ii) Proposals to change just the longitudinal or latitudinal mode are used approximately 3% and 8% of the time, respectively.

The relative percentages of the sky-mode hopping proposals reflect expectations about their relative importance [112].

V. TESTING THE REVERSIBLE JUMP WITH SIMULATED DATA

To test our pipeline, we generate mock data by injecting glitches—extracted from the LPF catalog [64] and shown in Fig. 1—into a noise-only dataset. For this test, we do not include any astrophysical signals, as our goal is to investigate the ability of the RJ-MCMC algorithm to identify glitches of varying shapes and SNRs and to distinguish them from noise in order to establish an approximate threshold below which such glitches are consistent with instrumental noise.

In Sec. VI, we also quantify the impact of three fiducial glitches—with low, medium, and high SNRs of 21, 72, and 1544, respectively—on the parameter estimation of an MBHB.

A. Data generation

We generate 30 days of LISA data where seven glitches are injected according to the amplitudes and damping times extracted from the LPF catalog [64] and are randomly distributed over these 30 days. In LPF, the glitches were observed to have a Poisson distribution with about one glitch per test mass per day. We use a slightly lower rate here to simplify the computational burden as we do not believe this will affect our results qualitatively. Moreover, we also allow for glitches injected very close in time to show the ability of the methodology to distinguish between such glitches.

Table I reports the parameters of the injected glitches as well as the corresponding number indicating the position of the glitch in the LPF catalog (the run index is in accordance with LPF numbering).

We model the injected glitches using the single exponential model reported in Sec. III Eq. (9), but we use shapelets with $n_0 = 1$ to fit for them. We simulate LISA noise directly in the frequency domain following [113]. Assuming that the real-time series $n(t)$ follows a stationary, zero-mean, Gaussian, and ergodic random process, the Fourier transform of the noise $\tilde{n}(f)$ is normally distributed with zero mean and variance given by the one-sided power spectral density,

$$\langle \tilde{n}^*(f') \tilde{n}(f) \rangle = \frac{1}{2} S_n(f) \delta(f - f'), \quad (27a)$$

$$\langle \tilde{n}(f') \tilde{n}(f) \rangle = 0, \quad (27b)$$

where $\langle \rangle$ denotes the expectation value over the data-generating process. For a finite grid of frequencies, defining $n_k = \tilde{n}(f_k)$, Eq. (27) can be written as

$$\langle \tilde{n}_k^* \tilde{n}_j \rangle = \frac{T}{2} S_n(f_k) \delta_{jk}. \quad (28)$$

TABLE I. Parameters t_0 , $\ln(\Delta v)$, τ , and SNR for each injected glitch. For each glitch, we also indicate the corresponding number from the LPF catalog.

Glitch no.	t_0 (s)	$\ln(\Delta v / [\text{ms}^{-2}])$	τ (s)	SNR
3	2.97×10^5	-25.5	184.47	72
118	7.50×10^5	-25.5	4.23	230
90	1.25×10^6	-24.4	43.21	479
72	1.75×10^6	-24.5	1.86	637
179	2.20×10^6	-23.9	4.77	1172
248	2.25×10^6	-23.6	13.96	1544
41	2.50×10^6	-28.3	5.57	21

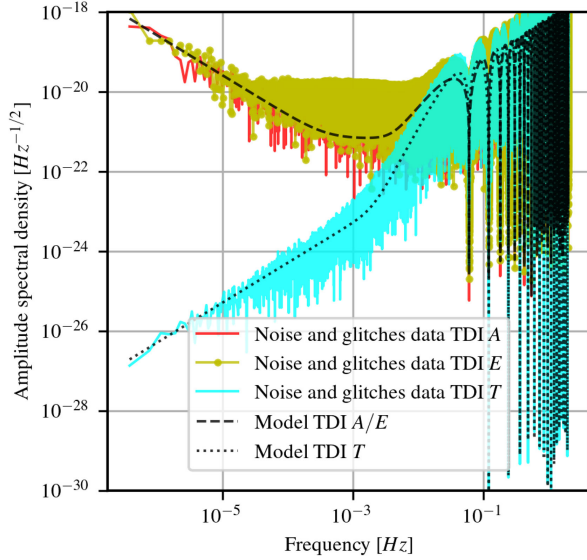


FIG. 3. Amplitude spectral density of the TDI channels A , E , and T when seven glitches are injected on TM_{12} at different times.

Therefore the variance of the real and imaginary parts of the noise is given by (where the real and imaginary part are uncorrelated)

$$\langle \Re[\tilde{n}_k]^2 \rangle = \langle \Im[\tilde{n}_k]^2 \rangle = \frac{T}{4} S_n(f_k) = \sigma_k^2. \quad (29)$$

The noise in the Fourier domain is then generated as

$$\tilde{n}_k = \sigma_k \cdot (X_{\text{real}} + iX_{\text{imag}}), \quad (30)$$

where $X_{\text{real}} \sim \mathcal{N}(0, 1)$ and $X_{\text{imag}} \sim \mathcal{N}(0, 1)$ are independent normally distributed random variables with mean 0 and variance 1.

For this example, the noise PSD is given by the sum of the TM acceleration and OMS noises which follow the models described in Sec. III E. The assumed value for the TM acceleration noise is $N_{\text{TM}}^2 = (2.4 \times 10^{-15} \text{ ms}^{-2} \text{ Hz}^{-1/2})^2$ and for the OMS noise $N_{\text{OMS}}^2 = (7.9 \text{ pm}/\sqrt{\text{Hz}})^2$.

Figure 3 shows the amplitude spectral density (ASD) in the TDI channels resulting from the injected glitches. The channels A and E show an excess of power with respect to the reference noise model in nominal conditions, shown in black. This excess of power is mild in the null channel T , as anticipated in [62]. The reason is that for null channels (acceleration) glitches have a mild effect due to the OMS and TM acceleration noise transfer functions of these channels. In particular, the latter shows a suppression at low frequencies with respect to the OMS such that the OMS noise dominates for these channels at all frequencies [54].

TABLE II. Table of priors used for shapelets and noise parameters.

Parameter	Lower bound	Upper bound
τ_0 (s)	0	2,592,000
$\ln(A_0/[\text{ms}^{-2}])$	-35	-20
β_0 (s)	1	1×10^5
N_{TM} ($\text{ms}^{-2} \text{ Hz}^{-1/2}$)	(7×10^{-12})	(8×10^{-12})
N_{OMS} ($\text{mHz}^{-1/2}$)	(2×10^{-15})	(3×10^{-15})

B. Data analysis

We analyze a complete 30-day dataset in a single run. Using the shapelet model with $n_0 = 1$, we fit for glitches, allowing up to 10 glitches in the data, also including the possibility of no glitches. The noise shape is considered fixed, while the two amplitude parameters of the OMS and TM noise, N_{OMS} and N_{TM} , vary within the uniform prior range specified in Table II. We also use uniform priors for the time of injection (τ_0) and the decay time (β_0) of the shapelet, whereas for the amplitude (A_0) of the shapelet, we consider a uniform prior in log amplitude. In this way, we avoid sampling parameters with vastly different numerical scales—such as amplitudes on the order of 10^{-12} and injection times around 10^4 – 10^6 . This choice tends to give more weight to low-SNR glitches. The goal here is to ensure that the sampler spends enough time in this region of parameter space, where most glitches occur. Figure 1 illustrates this relationship by showing the impulse, duration, and SNRs of the glitches in the LPF catalog.

The analysis takes about 2 days using 70 walkers and 10 temperatures, running on a single A100 GPU. We expect that this timing can be improved in the future. Concerning convergence of the algorithm, it was assessed by monitoring the stability of the posterior distributions over successive blocks of 500 iterations. Specifically, we updated the covariance group move (see Sec. IVA 1) every 500 iterations and checked that the posteriors did not change significantly between consecutive blocks. In practice, convergence was often achieved earlier. We plan to investigate alternative stopping criteria in future work. Figure 4 reports the posterior for the parameters of the shapelets used to fit the injected glitches. The parameters of the injected glitches (Table I)—converted to shapelet parameters according to $\tau_0 = t_0$, $A_0 = \frac{\Delta v}{2\beta_0}$, and $\beta_0 = \tau$ —are indicated within the figure.

It can be observed in Fig. 4 that the parameters of the glitch with the lowest SNR (21) cannot be recovered, and that the posteriors of the two glitches that are close in time overlap with each other such that in the 1D plot we observe a fifth bigger Gaussian for the time of injection τ_0 . The failure to recover the low-SNR glitch is likely due to the fact that the search did not detect it, possibly because proposed points rarely fall within the correct region of parameter space. This issue is probably linked to the choice

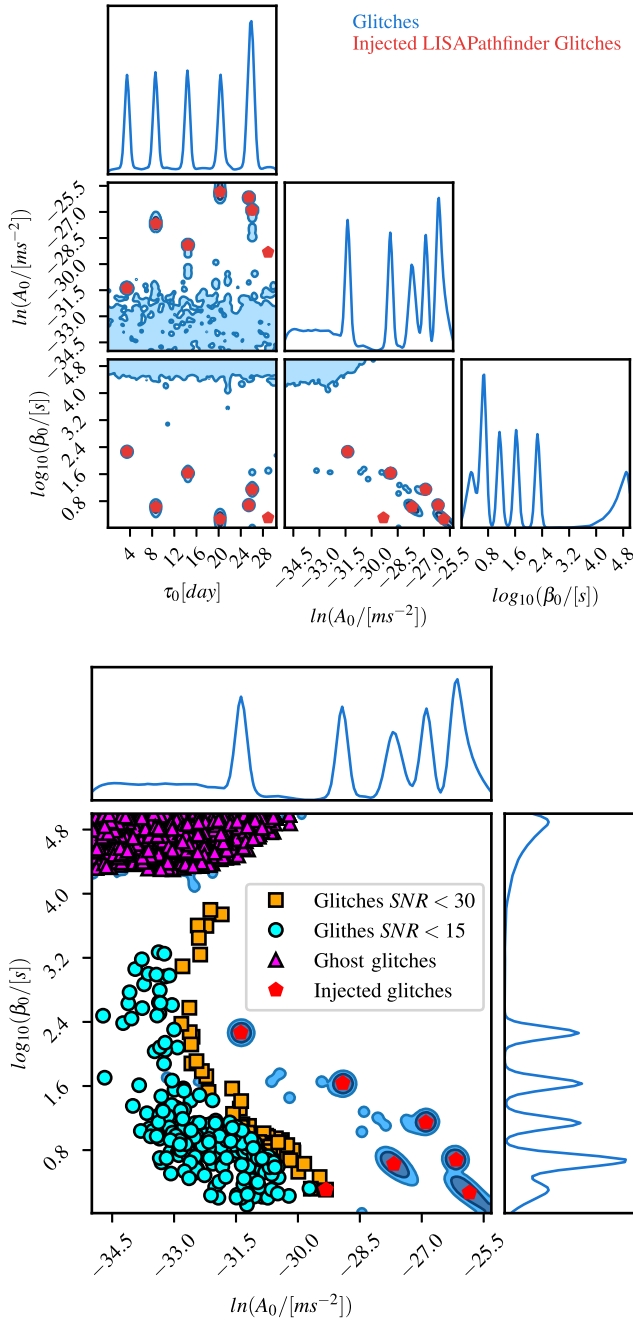


FIG. 4. Upper plot: posterior distributions of the shapelet parameters τ_0 , $\ln(A_0)$, and β_0 [plotted as $\log_{10}(\beta)$ for better visualization] for the seven injected glitches, fitted using first-order shapelets. Lower plot: enlargement of the 2D posteriors of $\ln(A_0)$ and β_0 [plotted as $\log_{10}(\beta_0)$], overlaid with SNR contours of the injected glitches.

of a uniform prior on β_0 (along with its lower amplitude), which turns out to be too broad relative to the τ_0 values for some of the injected glitches—despite this choice being consistent with the general LPF decay time distribution shown in Fig. 1. To verify this hypothesis, we repeated the analysis using a logarithmic prior on β_0 and we observed

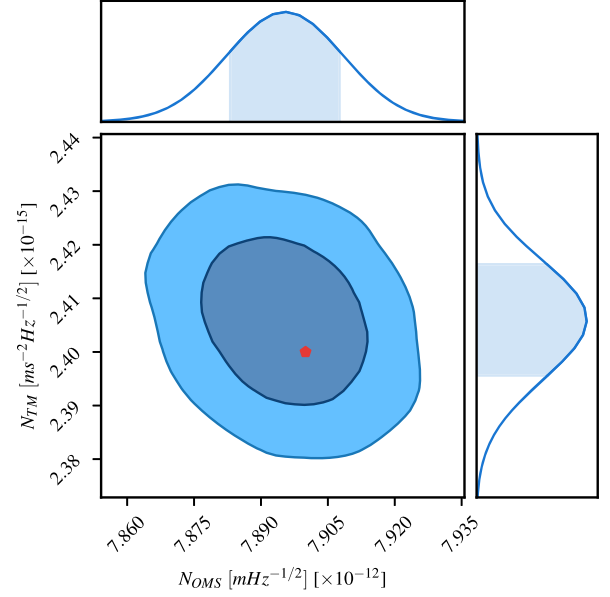


FIG. 5. Posterior distribution of the amplitudes, N_{TM} and N_{OMS} , for the test-mass acceleration noise and optical metrology system noise.

that the low-SNR glitch is better constrained.⁵ However, it is important to underline that the recovery of the two noise parameters is nonetheless unbiased, as shown in Fig. 5. Therefore, the recovered PSD is well constrained (Fig. 6) and closely aligns with the PSD model used to simulate the data, as illustrated in Fig. 6. In the lower plot of Fig. 4, we also observe a larger number of fictitious low-amplitude glitches (“ghost glitches”) arising from random fluctuations in the stationary component of the noise. These glitches have SNRs $\lesssim 24$. We also plot the glitches from LPF with $\text{SNR} < 15$ and $15 < \text{SNR} < 30$. Notably, the low-SNR glitch with $\text{SNR} = 21$, which was not fitted, lies within the region of the latter. This suggests that glitches with parameters in the same region as those are consistent with noise, given our choice of prior.

Finally, Fig. 7 shows the number of estimated glitches in the data at zero temperature. Even though seven glitches were injected, the sampler prefers to have eight glitches, which means adding two ghost glitches, since one of the injected ones is not recovered. The key difference between the ghost glitches and the missing one is that the amplitudes of the former are so low that the associated sampled value of β_0 and τ_0 have no impact on the likelihood, and so, if

⁵Additionally, to exclude the possibility that the low-SNR glitch is consistent with noise, we have performed another test in which we fixed the number of glitches to seven, initialized the sampler at the true parameters, and allowed the sampler to adjust slightly. Once the parameters stabilized, we restarted the run with RJ turned on, allowing the sampler to remove glitches. The sampler did not remove any of the glitches once they were found, indicating that—at least under the specific settings used in this example—this low-SNR glitch is not consistent with noise.

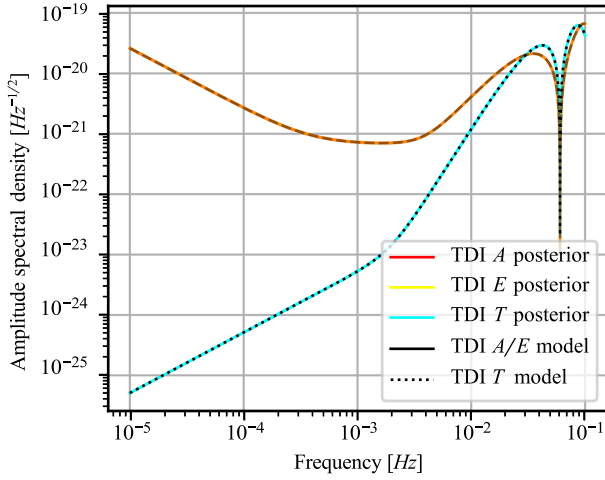


FIG. 6. Amplitude spectral density of the TDI A , E , and T channels constructed from 1000 random draws from the posterior distribution of the amplitude noise parameters of test-mass acceleration noise N_{TM} and optical metrology noise N_{OMS} , after glitch fitting is performed. The spread of the ASD across the posterior is very small, and the recovered ASDs are completely consistent with those used to generate the data.

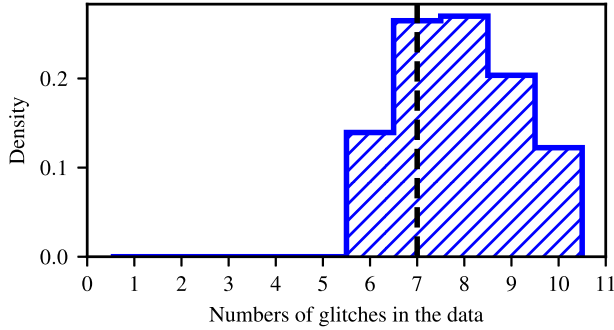


FIG. 7. Distribution of the number of glitches at zero temperature estimated by the dynamical parameter estimation procedure. The true value is marked with a dashed black line.

such signals have significant prior support, they will readily be added and removed by the RJ step of the sampler. The amplitude of the missing glitch is small enough to evade recovery, but not so low that the associated τ_0 and β_0 have no impact on the likelihood. Therefore, only values of τ_0 and β_0 in the right portion of the parameter space give high enough likelihood, and glitches with amplitude similar to that of the missing glitch are less likely to be accepted. The addition of ghost glitches is the result of allowing for low-amplitude glitches in the prior, which can be used to fit random excursions in the instrumental noise. As these fictitious glitches do not absorb much of the power in the data, this does not impact noise estimation, but nonetheless, there are techniques to mitigate this effect. One approach, used for example in [114] in the context of including white dwarf binaries in the LISA global fit, is to use a stricter

prior on the SNR of the white dwarf binaries that are proposed to be added.

To estimate the lowest SNR glitch that can be distinguished from noise, we considered different scenarios in which the no. 41 glitch, the low-SNR one that was not recovered in our analysis, was replaced by other glitches of increasing SNRs starting from SNR 20 to 65. We concluded that, for the noise model used in this study, glitches with an SNR of a few tens cannot be distinguished from noise. In the next section we show that such glitches do not lead to significant biases in MBHB parameter estimation. Biases start to appear from SNRs of ~ 60 . This observation could also be used to define a prior for glitches based on their SNR to help mitigate the confusion noise from the low-SNR sources discussed in the preceding paragraph.

VI. EFFECT OF GLITCHES ON PARAMETER ESTIMATION OF A MASSIVE BLACK HOLE BINARY

We consider now three different cases where we inject three glitches with SNRs of 21, 72, and 1544, respectively. The first glitch corresponds to no. 41, the second to no. 3, and the third to no. 248 of the LPF catalog. The parameters of these glitches are reported in Table I. We inject one glitch at a time, substituting it with another in each subsequent example. We simulate for all three cases a fiducial MBHB signal at redshift 4, with SNR 1152, in noisy data. The choice of the MBHB is driven by Fig. 3.5 of [2], which shows the typical SNRs and redshift of MBHB sources expected in LISA, as well as Fig. 3 in [1].

The MBHB priors used for this analysis are given in Table III. Uniform distributions are used on parameters $\{\log_{10}(M_t), q, a_1, a_2, d_L, \varphi_{\text{ref}}, \lambda_d, \psi, t_{\text{ref}}\}$. The prior on total mass is log-uniform, and we express the masses as dimensionless multiples of the solar mass M_\odot . For the inclination, we use a uniform prior in $\cos i$ between -1 and 1 , while the prior on the ecliptic latitude is uniform in $\sin \beta_d$.

TABLE III. True values of the MBHB parameters together with the prior range used for parameter estimations. We assume a flat prior on the specified range. Values of the masses are given in the detector frame.

Parameter	True	Lower bound	Upper bound
$\log_{10}(M_t)$	6.29	5	8
q	0.4628	0.01	0.999
a_1	0.7474	-1	1
a_2	0.8388	-1	1
d_L (Gpc)	37	0.01	1000
φ_{ref} (rad)	1.194	0	2π
$\cos i$	0.5	-1	1
λ_d (rad)	1.288	0	2π
$\sin \beta_d$	-0.298	-1	1
ψ (rad)	$\pi/6$	0	π
t_{ref} (s)	2.628×10^6	$t_{\text{ref}} - 100$	$t_{\text{ref}} + 100$

We choose the injection time for each of the three glitches to maximize the impact of the glitch on the MBHB parameter estimation, following the procedure described in [115]. The idea is to maximize the overlap between the glitch and the MBHB signal. The noise-weighted inner product can be defined as

$$(h_1|h_2) = 4\Re \int_0^\infty h_1^\dagger(f) \mathbf{C}^{-1}(f) h_2(f) df. \quad (31)$$

Here, $h_1(f)$ and $h_2(f)$, respectively, stand for the vectors of MBHB signals and glitches as they appear in the N_{TDI} TDI combinations considered in our analysis. $\mathbf{C}(f)$ is the cross-spectral density matrix of the noise in those channels and \dagger denotes the Hermitian transpose. Using this inner product, the normalized overlap is given by

$$\mathcal{O}(h_1, h_2) = \frac{(h_1|h_2)}{\sqrt{(h_1|h_1)(h_2|h_2)}}. \quad (32)$$

When considering multiple independent channels, for example, TDI *A* and *E*, the inner product reduces to a sum over the channels,

$$(h_1|h_2) = 4\Re \int_0^\infty \sum_{j=1}^{N_{\text{TDI}}} \frac{h_1^{j*}(f) h_2^j(f)}{S_n(f)} df, \quad (33)$$

where we consider the *A* and *E* channels to not only be uncorrelated but also to have equal noise PSD, $S_n(f)$.

Defining Δt as the difference between the MBHB merger time and the glitch injection time, we found that the best match for the glitch with SNR 21 is at $|\Delta t| = 6.80361$ min before the merger, for the glitch with SNR 72 it is at $|\Delta t| = 25.2338$ min before the merger, and for the glitch with SNR 1544 at $|\Delta t| = 11.5665$ min before the merger. Figure 8 shows the variation of the overlap versus time for the three glitches, where we have fixed the amplitudes and decay time of the glitches while varying the injection time. The dotted vertical lines indicate the injection time, with respect to the merger, at which the overlap is maximized for each of the three glitches. We show in Fig. 9 the three glitches injected at the time that maximizes the overlap between the glitch and the fiducial MBHB in TDI channel *A*. Note that the frequency of the MBHB evolves over time, and the maximum impact of a glitch is expected when its frequency overlaps with the instantaneous frequency of the MBHB signal. Acceleration glitches typically occur at frequencies on the order of 10^{-3} Hz, which aligns more closely with the early inspiral phase. However, if the MBHB merger occurred at lower frequencies—as in the case studied in [115]—the glitch would instead have the strongest overlap during the merger phase.

The posteriors shown in Figs. 10 and 11 illustrate the effect of the presence of these different glitches if we

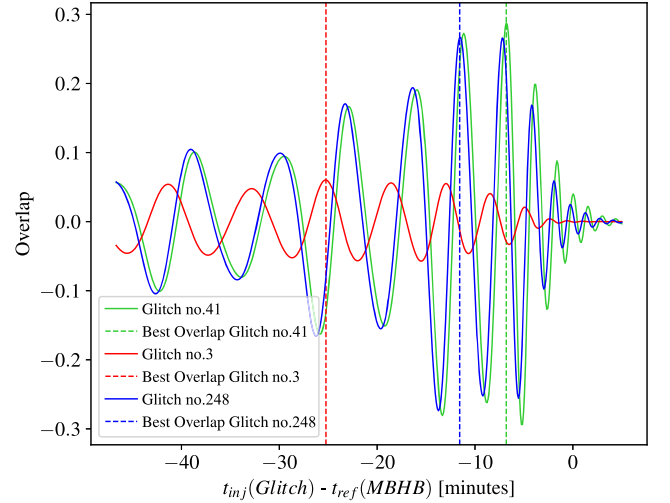


FIG. 8. Overlap between the MBHB and glitch signals as a function of the glitch injection time relative to the merger, where the decay time and amplitude of the glitches are kept fixed. The red, blue, and green solid curves are for the three glitches with SNRs of 21, 72, and 1544. The GW signal is fixed to that of our fiducial MBHB. Dashed vertical lines with matching colors denote the times that maximize the overlap.

attempt to recover the parameters of this MBHB without simultaneously fitting for both the artifact and the GW source. For comparison, the posteriors of the MBHB parameters obtained from data containing no glitches are reported in both plots, and the true values of the parameters are indicated by a dashed black line.

In the case of the SNR 21 glitch, shown in Fig. 10, the glitch has less impact in the intrinsic parameters with respect to the impact we see in the extrinsic ones, as the shift in the posterior is less than the width of the posterior.

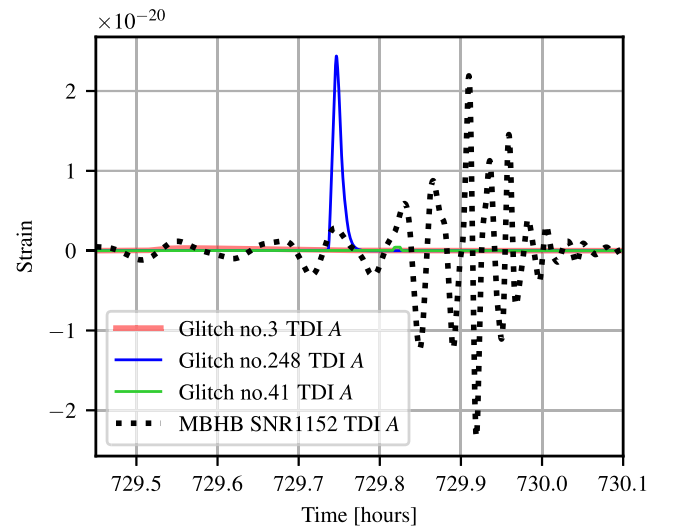


FIG. 9. Fiducial MBHB waveform along with the glitches injected at their maximum overlap time in TDI channel *A*.

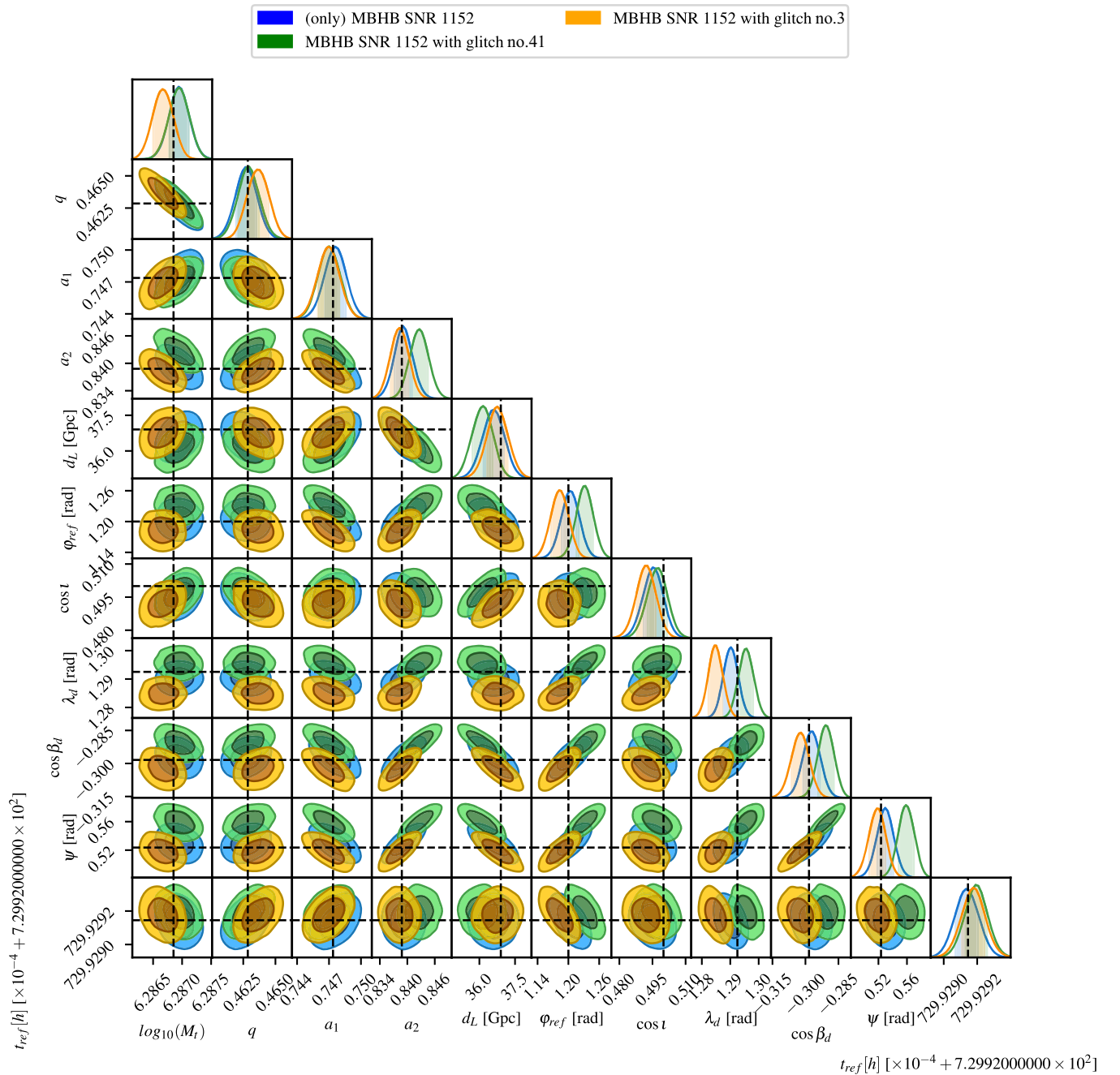


FIG. 10. Posteriors of an MBHB with SNR 1152 compared to two cases: a glitch of SNR 21 is injected at the maximum match but not fitted; a glitch of SNR 72 is injected at the maximum match but not fitted.

The bias starts to be significant for the glitch with SNR 72 for several of the parameters, e.g., the total mass, the reference phase at merger, and the ecliptic longitude and latitude. For the SNR 1544 glitch shown in Fig. 11, the two posteriors are completely disjoint, with all parameters significantly biased. To help the reader visualize the effect of not fitting such a glitch, we report only the posteriors for the intrinsic parameters. It is important to specify that the noise model has been fixed for this analysis, and we expect

that variations in the noise model would allow to partially absorb the impact of the glitch making the MBHB posterior more consistent with the injected values.

Let us recall that the injection times of the glitches were chosen specifically to maximize their impact on the parameter estimation of the MBHB. Their effect is expected to be smaller if they occur during the early inspiral. This point is further discussed in Appendix B, where we present posteriors illustrating how the presence of

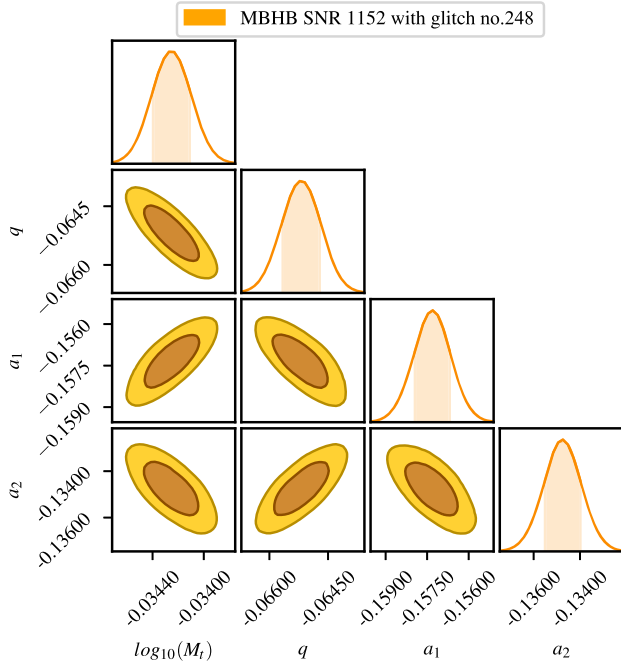


FIG. 11. Posteriors of the intrinsic parameters of an MBHB with SNR 1152 in the case where a glitch of SNR 1544 is injected—and not fitted—at the time maximizing its match with the injected signal. We have shifted the parameters relative to the injected values. The fact that the posteriors do not peak at 0 illustrates the bias caused by the glitch. For the sake of clarity, we do show the corresponding posteriors for the only MBHB case, as the statistical error is much smaller than the systematic bias observed in the figure.

the SNR 1544 glitch affects parameter estimation for the same MBHB, depending on its injection time. However, since the coalescence is the loudest part of MBHB signals and largely drives parameter estimation, failing to capture glitches occurring near to it would severely jeopardize the science objectives of LISA. For instance, biases in sky localization and distance could prevent the detection of an electromagnetic counterpart and lead to misidentification of the host galaxy, thereby compromising cosmological measurements based on bright or dark siren methods [26–28,116]. This issue also extends to applications in fundamental physics, which rely on detecting extremely small deviations from general relativity. In this context, even small biases can lead to false detections and must be mitigated as much as possible [117,118]. Our analysis shows that, to avoid such pitfalls, glitches with SNRs of several tens or higher will need to be included in the model fitted to the data. This threshold also corresponds to the point at which glitches first become identifiable in the data, as discussed in the previous paragraph.

In Fig. 12, we now show the posterior on the parameters of the MBHB signal obtained when simultaneously fitting for the glitch and compare this to the posterior obtained in the absence of the glitch. We observe that by fitting the

glitch, the MBHB posteriors are no longer biased, and the posteriors for the (only) MBHB case overlap with those where the glitch has been accounted for, indicating that the best strategy is to fit for both. For completeness, we also show posteriors for the fitted glitch parameters in Fig. 13.

We investigated the effect of injecting glitches (nos. 3 and 41) into MBHB with an SNR of 320, in order to reproduce the example studied in Ref. [115]. The posteriors plots are reported in Appendix C. For the short-duration glitch with SNR 21 (no. 41), we find results consistent with those reported there: the inferred parameters remain largely unbiased. In contrast, glitch no. 3, with an SNR of 72, though also short in duration, introduces a more significant bias.

VII. (LIGHT-)SPRITZ DATASET

In this section, we describe the Spritz dataset and the strategy we have used to analyze it. We describe the mock data in Sec. VII A and the noise model used for this analysis in Sec. VII B, since it differs from the noise that was used to generate and analyze the data in the previous section, Sec. VI. Finally, in Sec. VIII, we present the methodology we use for doing the initial search and parameter estimation of glitches, noise, and MBHB and for the subsequent analysis of the full dataset to refine the posterior estimates of the noise, GW signal, and artifacts.

A. Introduction

Full details of the Spritz datasets can be found in [119]. The Spritz data release comprised three main datasets, of which we focus on the first one, which contains a single loud MBHB and three glitches as visible in Fig. 14, which shows the dataset in the time-frequency domain, and Fig. 15 which shows it in the frequency domain. The glitches are so-called double decay exponential glitches [74]. All the datasets are realized with TDI X , Y , and Z second generation assuming Keplerian orbits. The dataset also features more realistic instrumental noise (see Sec. VII B), a background of Galactic binaries and data gaps. Auxiliary data included in the dataset allow individual components to be perfectly subtracted, allowing analysts to focus on selected aspects of the dataset. Since we focus on the detection of glitches and their impact on MBHB parameter estimation, we removed the gaps and Galactic binaries, and for this reason, we renamed the modified dataset as light-Spritz. Furthermore, the version of `BBHx` we are using for this paper (version 0.0.0) does not support the response model used to generate the MBHB in Spritz. The Spritz data were generated using the `PhenomD` waveform model [120,121], with the LISA response evaluated for Keplerian orbits (time-varying arm lengths), but our version of `BBHx` does not include time-varying arm-length response. Therefore, to rule out any systematic errors from this waveform mismatch [122], we replace

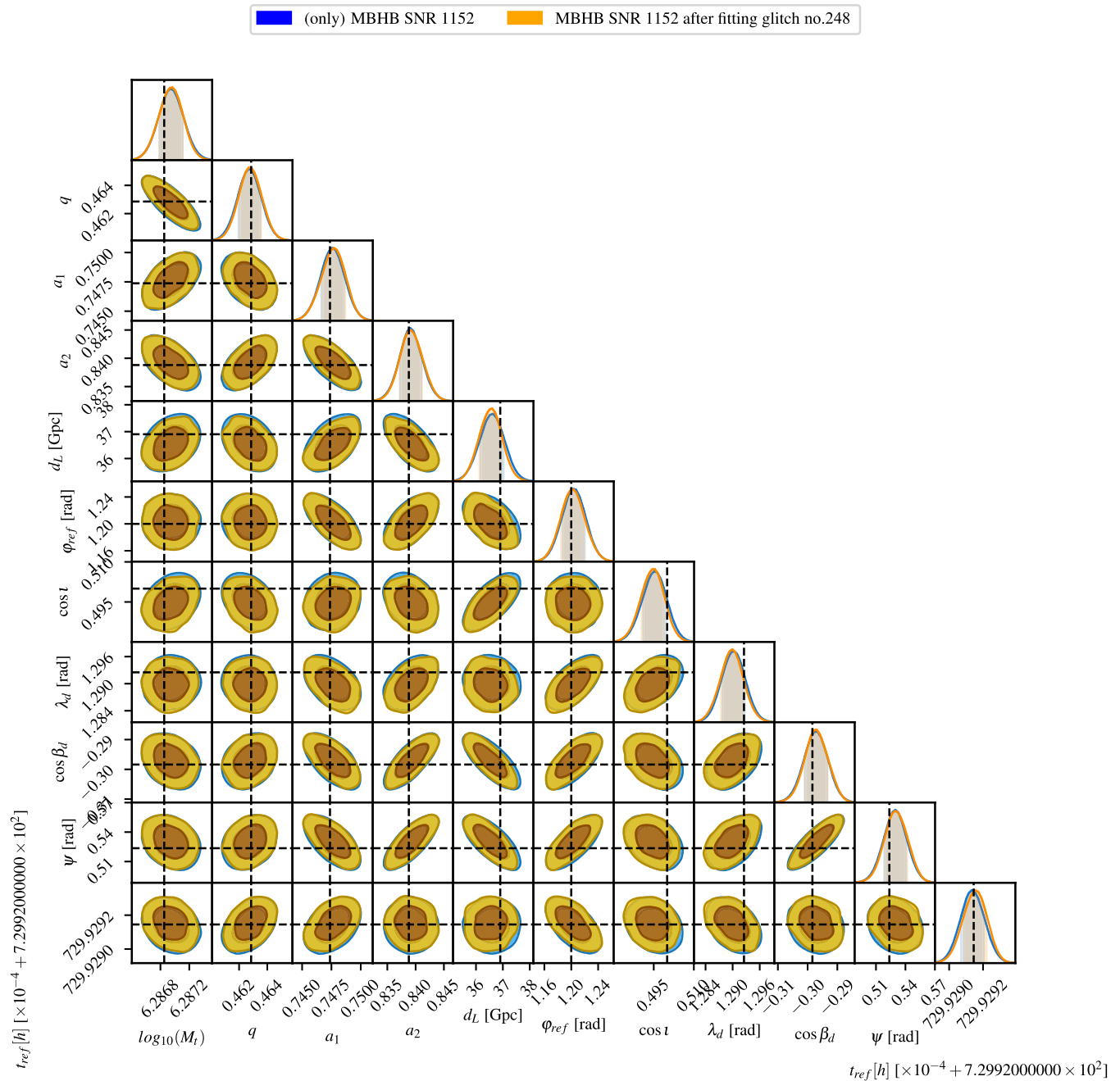


FIG. 12. Posteriors for the parameters of an MBHB with SNR 1152 obtained from data containing only the MBHB, compared to those obtained from data containing a glitch of SNR 1544, when the latter is also fitted.

the original Spritz injected signal with one with the same parameters (see Table III in Sec. VI), but generated using the PhenomHM waveform (instead of PhenomD) with an equal-arm response. We plan to include a more realistic treatment of the LISA constellation, to compute the response, in a follow-up study that will analyze the full dataset.

Regarding glitch modeling, Ref. [62] showed that the mismatch between equal and unequal/time-varying-arm models does not significantly impact glitch parameter

reconstruction. For the noise, however, using equal, unequal, or time-varying arm-length models affects the high-frequency range, where the zeros of the TDI appear. To avoid related issues, we restrict our analysis to frequencies up to 2×10^{-2} Hz [123], which does not alter the outcome of our analysis. Additionally, the null channel T cannot be well approximated using the equal-arm-length model [124,125]. For these reasons, we perform the analysis using only the two sensitive channels, A and E ,

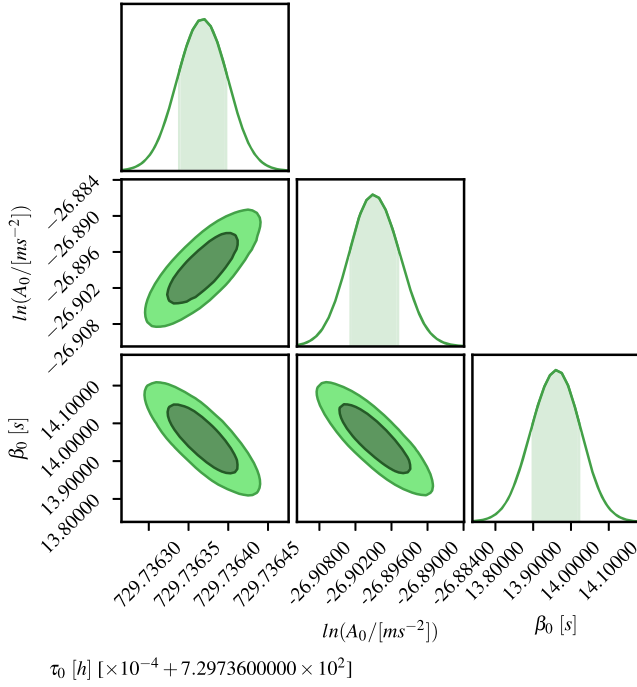


FIG. 13. Posteriors on the parameters of the first-order shapelet used to fit the injected glitch (no. 248) with SNR 1544.

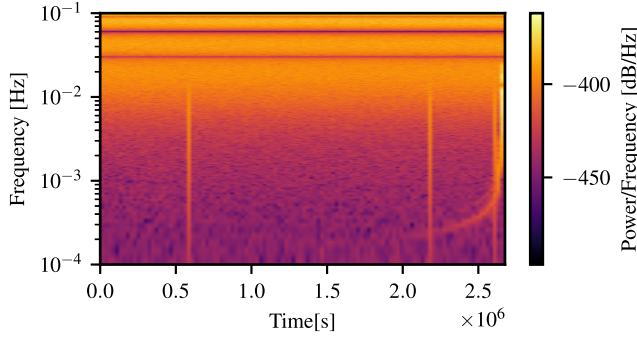


FIG. 14. Time-frequency representation of the light-Spritz dataset.

employing second-generation TDI under the assumption of equal-arm lengths, for which sensitivity to discrepancies in arm-length modeling is negligible. For consistency, the MBHB signal is also computed using second-generation TDI variables [see Eq. (3)], assuming constant and equal-arm lengths.

As discussed in Sec. IV, we analyze the data in the frequency domain. However, converting the time-domain data provided in light-Spritz to the frequency domain by using a fast Fourier transform (FFT) without any preprocessing causes significant artifacts in the form of spectral leakage [86], linked to significant high-frequency content that leaks into the low-frequency part of the light-Spritz dataset. We mitigate this effect by applying a low-pass filter to the data before taking the FFT. Another approach would be

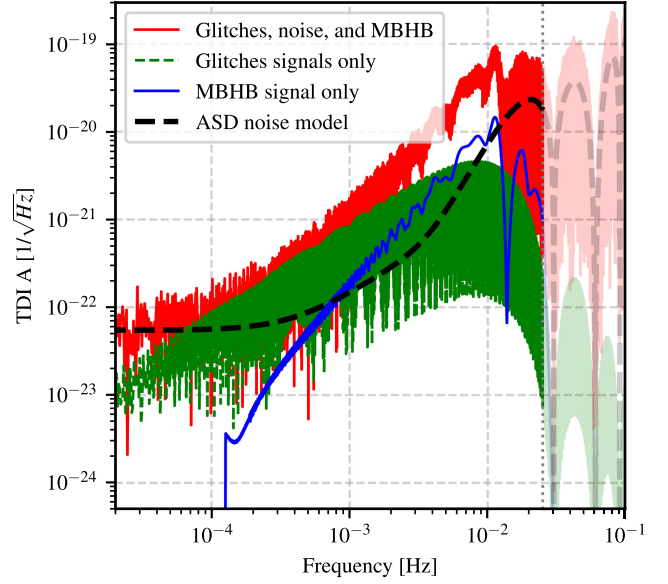


FIG. 15. Fourier domain representation of the light-Spritz dataset.

to apply a window function to the data before taking the FFT. This has the downside of introducing correlations between adjacent frequency bins, potentially biasing the parameter estimation using our diagonal likelihood model. Instead, by filtering the data, we obtain a diagonal likelihood (see Appendix E). Concretely, we use a first-order Butterworth filter with a cutoff frequency of 1×10^{-3} Hz using the `SciPy.signal.butter` function and then use the `SciPy.signal.filtfilt` function to apply the filter twice in opposite time directions, such that it has zero impact on the phase of the underlying data (i.e., it does not cause any time shift of the data). Moreover, in order not to bias our analysis, we apply a frequency-domain model of the filter response to the noise, glitches, and GW signal templates used in the likelihood as well. Due to the double application of the filter, the Fourier transforms appearing in the likelihood are scaled by $|H(f)|^2$, whereas PSDs are scaled by $|H(f)|^4$, where $H(f)$ is the complex filter transfer function,

$$\begin{aligned} \log \mathcal{L} &= -\frac{1}{2} \sum_f \left[4\Delta f \frac{|\tilde{d}(f) - \tilde{h}(f, \theta)|^2 |H(f)|^2 - \tilde{g}(f, \phi) |H(f)|^2}{S(f, \gamma) |H(f)|^4} \right. \\ &\quad \left. + \log \{S(f, \gamma) |H(f)|^4\} \right]. \end{aligned} \quad (34)$$

B. Noise model

To study—and model—just the noise in the light-Spritz dataset, we further removed the glitch and MBHB components described in Sec. VII A. We then estimated the spectra using the `SciPy.signal.welch` method with the following parameters: `window=(“kaiser,” 10)`,

$n_{\text{perseg}}=10$, 713 , $n_{\text{overlap}}=5356$, and $\text{dettrend}=\text{"linear"}$. The LDC manual for Spritz [126] suggests that laser, backlink, acceleration, and readout noises were enabled in the simulation. However, we find that the corresponding noise model provided as part of the LDC software suite ([97], v.1.2.4) does not perfectly describe the Spritz simulated data.

To accurately fit the noise of the dataset, we cross-checked against the default values for the noise sources included in the instrument simulation at the time [127] and consequently included the following noise components, expressed as PSDs:

- (i) Interspacecraft interferometer OMS noise, $S_{\text{OMS}}^{\text{ISI}} = (N_{\text{OMS}}^{\text{ISI}})^2$ (white), with $N_{\text{OMS}}^{\text{ISI}} = 6.35 \text{ pm}/\sqrt{\text{Hz}}$.
- (ii) Reference interferometer OMS noise, $S_{\text{OMS}}^{\text{RSI}} = (N_{\text{OMS}}^{\text{RSI}})^2$ (white), with $N_{\text{OMS}}^{\text{RSI}} = 3.32 \text{ pm}/\sqrt{\text{Hz}}$.
- (iii) Test-mass interferometer OMS noise, $S_{\text{OMS}}^{\text{TMI}} = (N_{\text{OMS}}^{\text{TMI}})^2$ (white), with $N_{\text{OMS}}^{\text{TMI}} = 1.42 \text{ pm}/\sqrt{\text{Hz}}$.
- (iv) TM acceleration noise, $S_{\text{TM}} = (N_{\text{TM}})^2 \cdot (1 + (\frac{0.4 \times 10^{-3}}{f})^2)$, with $N_{\text{TM}} = 2.4 \text{ fm}/\text{s}^2/\sqrt{\text{Hz}}$.
- (v) Reference Interferometer (RFI) and Test Mass Interferometer (TMI) backlink noise, $S_{\text{BL}}^{\text{RFI/TMI}} = (N_{\text{BL}}^{\text{RFI/TMI}})^2 \cdot (1 + (\frac{2 \times 10^{-3}}{f})^4)$, with $N_{\text{BL}}^{\text{RFI/TMI}} = 3 \text{ pm}/\sqrt{\text{Hz}}$.

For the analysis, we define the following combinations of the noise parameters to estimate the effective noise levels:

$$\gamma = \begin{cases} N_1 \equiv [(N_{\text{OMS}}^{\text{ISI}})^2 + (N_{\text{OMS}}^{\text{RFI}})^2 + (N_{\text{BL}}^{\text{RFI/TMI}})^2]^{1/2}, \\ N_2 \equiv [(N_{\text{OMS}}^{\text{TMI}})^2 + (N_{\text{BL}}^{\text{RFI/TMI}})^2]^{1/2}, \\ N_{\text{TM}}, \end{cases} \quad (35)$$

with $N_1 = 3.32 \times 10^{-12}$, $N_2 = 7.78 \times 10^{-12}$, and $N_{\text{TM}} = 2.4 \times 10^{-15}$. As in Sec. III E, we first express the noises as equivalent displacement noise in $\text{m}^2 \text{ Hz}^{-1}$ or acceleration noise in $\text{m}^2/\text{s}^4/\text{Hz}$. We then convert them to units of fractional frequency fluctuations as used in the LDC by applying the factors $C_{\text{disp}}^{\text{ffd}} = (2\pi f/c)^2$ and $C_{\text{acc}}^{\text{ffd}} = (c2\pi f)^{-2}$, where c is again the speed of light.

Figure 16 shows the ratio between the estimated spectra (the PSD computed from the Spritz only noise data), the model included with the LDC tools, and the model we used. We see artifacts around the zeros of the TDI transfer function for both plots, owing to the use of an equal-arm approximation in the model we are using to fit the noise. For our model, we estimated the arm length to be used in the equal-arm model as an average over all six time-varying arm lengths, evaluating to $L/c \approx 8.323 \text{ s}$. Whereas we used the default value set in the LDC tools for the LDC model. This causes slightly increased artifacts for the LDC model around the zeros. We also observe that the LDC model slightly underestimates the noise in the upper half of the frequency band.

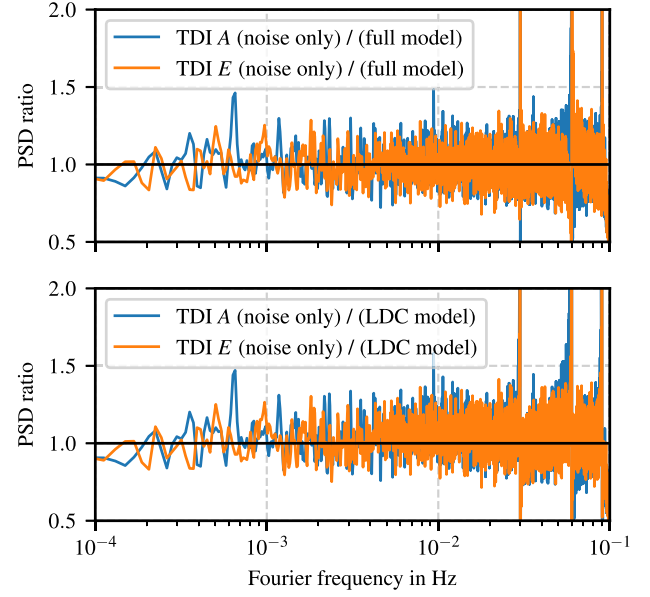


FIG. 16. Comparison between the power spectral density of the Spritz noise dataset with the noise model reported in the LDC manual.

To conclude, some of these noises have different TDI transfer functions, which are derived in [96] for TDI X , Y , Z as well as A , E , T . For our work, we only need the transfer functions for A and E , which turn out to be identical under our assumptions. We summarize them here for convenience. They include a common TDI factor for second-generation TDI,

$$C_{xx} = 16 \sin^2(2\pi fL) \sin^2(4\pi fL). \quad (36)$$

Using this, we have

$$|H_{\text{ISI/RFI(OMS)}}^{\text{AA}}|^2 = 2C_{xx}(2 + \cos(2\pi fL)) \quad (37)$$

$$|H_{\text{TMI(OMS)}}^{\text{AA}}|^2 = C_{xx}(3 + 2\cos(2\pi fL) + \cos(4\pi fL)) \quad (38)$$

$$|H_{\text{TM}}^{\text{AA}}|^2 = 4 \cdot |H_{\text{TMI(OMS)}}^{\text{AA}}|^2, \quad (39)$$

such that the total noise can be described by

$$\begin{aligned} S_{\text{AA}}^{\text{total}} = S_{\text{EE}}^{\text{total}} = & C_{\text{disp}}^{\text{ffd}} [|H_{\text{TMI(OMS)}}^{\text{AA}}|^2 (S_{\text{OMS}}^{\text{TMI}} + S_{\text{BL}}^{\text{TMI}}) \\ & + |H_{\text{ISI/RFI(OMS)}}^{\text{AA}}|^2 (S_{\text{OMS}}^{\text{ISI}} + S_{\text{OMS}}^{\text{RFI}} + S_{\text{BL}}^{\text{RFI}}) \\ & + C_{\text{acc}}^{\text{ffd}} |H_{\text{TM}}^{\text{AA}}|^2 S_{\text{TM}}. \end{aligned} \quad (40)$$

VIII. ANALYSIS OF THE DATASET

The search-phase algorithm is visualized in Fig. 17 and explained in Sec. VIII A, while the parameter estimation one is visualized in Fig. 18 and explained in Sec. VIII B.

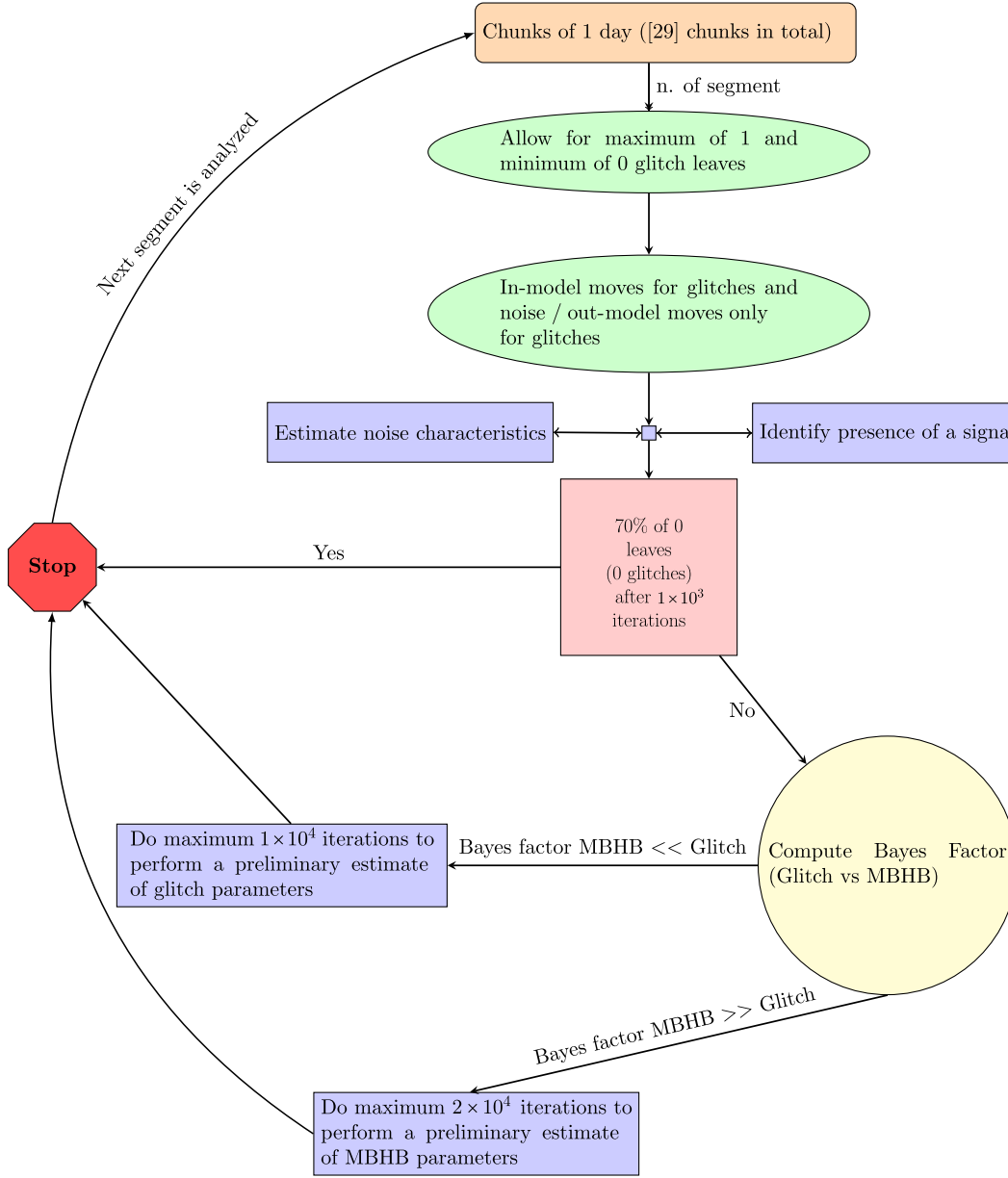


FIG. 17. Schematic view of the search algorithm.

A. Search phase

We divide the time series into segments of approximately one day, with a 20% overlap between adjacent segments. We therefore segment the light-Spritz dataset into 29 stretches of data and we regenerate a time grid spanning from $t_0 = 0$ to $t_{\text{end}} = 2,675,890$ s by subtracting the initial time of the Spritz dataset $t_{\text{in}} = 8,899,200$ s. We consider broad prior ranges for the glitch parameters, allowing the distribution of A_0 and β_0 to extend beyond what is observed in LPF (Fig. 1), aiming to remain as agnostic as possible.

The priors and the corresponding true values on the glitch and noise parameters for the search phase are reported in Table IV. The true values of the amplitude

and decay times for the glitch parameters are taken from the Spritz LDC manual [69]. In particular the amplitude Δv is 2.20 pm s^{-1} , while the two decay times are $\tau_1 = 10$ and $\tau_2 = 11$ s. The true value of the glitch injection times τ_{inj} are provided in the Spritz dataset in days $\tau_{\text{inj}_{gl_1, gl_2, gl_3}} = \{110, 128, 133\}$. We thus convert each of them to the timescale $\tau_{0_{gl_1, gl_2, gl_3}}$, used in our analysis, as

$$\tau_{0_{gl_1, gl_2, gl_3}} = \frac{\tau_{\text{inj}_{gl_1, gl_2, gl_3}} \times 86,400 - t_{\text{in}} - n_{\text{skip}} \times dt}{3600},$$

where $n_{\text{skip}} = 1000$ is the number of samples skipped to account for the burn-in period of the filter.

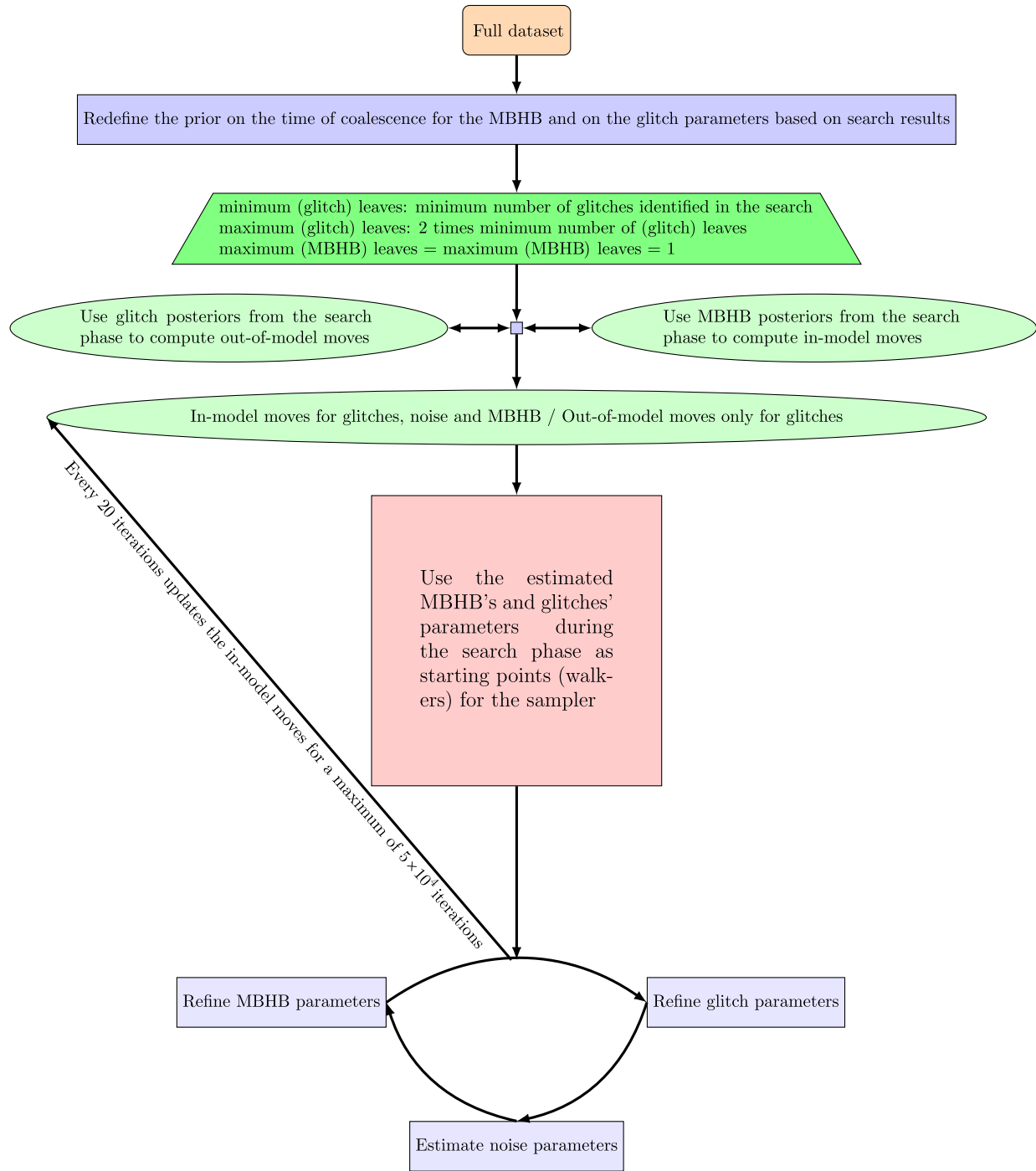


FIG. 18. Schematic view of the parameter estimation algorithm.

Our search is split in two phases: the first phase and second phase.

1. First phase

First, we fit only for both noise and glitches in each segment separately. We use the in-model and out-of-model moves described in Sec. IV A, to explore the parameter space. The number of glitches in each segment is estimated between

0 and 1 (i.e., the number of leaves in the glitch branch is 0 or 1). The search includes a stopping criterion such that if, after 1000 iterations, more than 70% of the samples have no glitch leaves, we conclude that no glitches are present in the segment and proceed to analyze the next one. Otherwise, the segment is flagged as containing a possible glitch signal.

Table V reports the results of the first phase of the search pipeline, showing in which stretches of data glitches have been found and their SNRs. Note that the SNR given in the

TABLE IV. Table of priors for the glitch and noise parameters along with the true values. The variables “t start” and “t end” change depending on the stretch considered. Note that the following relationships hold between a double decay exponential glitch and the shapelet model used to fit it: $A_0 = \Delta v \frac{1}{1+\tau_2/\tau_1}$ and $\beta_0 = \frac{\tau_1+\tau_2}{2}$.

Parameter	True value	Lower bound	Upper bound
$\tau_{0_{gl_1}}$ (s)	166.61	t start	t end
$\tau_{0_{gl_2}}$ (s)	598.61	t start	t end
$\tau_{0_{gl_3}}$ (s)	718.61	t start	t end
$\ln(A_0/[\text{ms}^{-2}])$	-27.58	-35	-20
β_0 (s)	10.5	1	1×10^5
N_1 (m/ $\sqrt{\text{Hz}}$)	7.77×10^{-12}	7×10^{-12}	8×10^{-12}
N_2 (m/ $\sqrt{\text{Hz}}$)	3.32×10^{-12}	2.5×10^{-12}	3.5×10^{-12}
N_{TM} (m/s ² / $\sqrt{\text{Hz}}$)	2.4×10^{-15}	2×10^{-15}	3×10^{-15}

table is estimated from the simultaneous fitting of glitches and noise. Specifically, to calculate the SNR, we compute the mean of the posterior distribution of SNR values for each data stretch during the search phase. This mean is obtained by drawing 3000 random samples from the posterior distribution of the glitch and noise parameters. However, the instrumental noise is not sufficiently constrained; therefore, these SNRs have to be considered as a rough estimate. Indeed, we observed that only a day of data is not enough to estimate and constrain all the noise parameters properly. This is particularly true in stretches of data where the late inspiral part of the MBHB is present (N26–N28). Therefore, we need to properly constrain the noise when processing the full month of data. This is further discussed in Appendix D.

To perform this initial search, we used 25 walkers and 10 temperatures, and the run for each segment takes about 20 min.

TABLE V. Table showing the number of leaves in each model and the glitch SNR for each stretch of data analyzed. N labels the segment considered. The “% of 0 or 1 leaf” columns indicate the fraction of samples with 0 or 1 leaf, respectively. The bold fonts indicate stretches of data where glitches were identified, the italic fonts mark regions where the inspiral phase of the MBHB becomes significant, and the bolditalic fonts stretches where the merger of the MBHB is present.

N	Time (s)	% of 0 leaves	% of 1 leaves	Glitch SNR
0	0.0–107,995	86	13	0.60
1	72,000–197,995	66	36	1.27
2	162,000–287,995	87	13	0.39
3	252,000–377,995	86	14	0.47
4	342,000–467,995	85	15	0.50
5	432,000–557,995	83	16	0.54
6	522,000–647,995	10	90	294
7	612,000–737,995	58	42	1.28
8	702,000–827,995	87	13	0.42
9	792,000–917,995	87	13	0.43
10	882,000–1,007,995	56	44	1.42
11	972,000–1,097,995	79	21	1.46
12	1,062,000–1,187,995	85	15	0.51
13	1,152,000–1,277,995	84	15	0.52
14	1,242,000–1,367,995	86	14	0.45
15	1,332,000–1,457,995	87	12	0.40
16	1,422,000–1,547,995	87	13	0.45
17	1,512,000–1,637,995	80	20	0.65
18	1,602,000–1,727,995	84	16	0.65
19	1,692,000–1,817,995	86	14	0.48
20	1,782,000–1,907,995	86	14	0.47
21	1,872,000–1,997,995	85	15	3.34
22	1,962,000–2,087,995	85	15	0.44
23	2,052,000–2,177,995	3	97	269.14
24	2,142,000–2,267,995	6	93	265.68
25	2,232,000–2,357,995	81	18	0.77
26	2,322,000–2,447,995	8	92	3.74
27	2,412,000–2,537,995	3	97	5.85
28	2,502,000–2,627,995	1	98	188.99
29	2,592,000–2,675,890	1	99	520.26

We observe five stretches with high SNR signals. Of these, four have been identified as glitches (bold font) and one as an MBHB (bold italic font). All these stretches have at least 90% of samples with only one leaf. Additionally, two other stretches (italic font) have more than 92% of samples with one leaf, although their estimated signal SNRs are much lower. This occurs because the inspiral phase of the MBHB waveform starts to contribute a non-negligible SNR, triggering the algorithm with evidence of a low-SNR signal. In the remaining stretches, the SNR values result from random noise fluctuation.

In the first phase of the search, the glitch model serves as a trigger to detect the presence of a signal in each stretch, whether it corresponds to a glitch or an MBHB signal. This approach is effective because the MBHB in light-Spritz has a high SNR, leading to behavior similar to matched filtering, where the overlap between our template bank and the GW signal increases with the SNR (the characteristic frequency of the glitch and the instantaneous GW frequency of the MBHB). However, in cases with lower SNR MBHBs, the search may not be sensitive enough to detect such signals. To resolve ambiguities between the presence of a glitch or an MBHB signal in each stretch, the second phase of the search evaluates the Bayes factor comparing the glitch and MBHB hypotheses.

2. Second phase

The first phase of our search aims to identify the presence of a signal in each stretch of data. In the second phase, we determine whether the signal is better described as a glitch or an MBHB. To this end, we perform a dedicated MBHB search to assess the presence of a GW signal in stretches where a signal was detected or where the estimated SNR is sufficiently high (N6, N23, N24, N26, N27, N28, and N29). We then compute the Bayes factor comparing the MBHB and glitch hypotheses.⁶ For this, we employ both the stepping-stone method [128] and thermodynamic integration [67]. This information is subsequently used to initialize the walkers for the parameter estimation phase. We then compare the Bayes factors between the glitch and MBHB hypotheses for each of these stretches. To this end, we use 30 temperatures and run the sampler with a fixed-dimensional parameter space. We fit only for the signal, assuming the noise to be at a fiducial level, following the strategy outlined in [57]. This approach is robust given the high SNR of both glitches and MBHBs, and no ambiguous Bayes factor values were observed that would suggest the simultaneous presence of both

⁶Note that this analysis could alternatively be performed using the RJ-MCMC framework, which would allow for automatic computation of Bayes factors between glitches and MBHBs by leveraging the EREBOR global fit infrastructure [57]. However, incorporating this additional complexity is not necessary for the current analysis.

TABLE VI. Log Bayes factors comparing the MBHB and glitch models for datasets N6, N24, and N29 using the stepping-stone method. $d\log_{10}\mathcal{B}_{\text{MBHB}/\text{glitch}}$ gives an estimate on the error on $\log_{10}\mathcal{B}_{\text{MBHB}/\text{glitch}}$.

	N6	N24	N29
$\log_{10}\mathcal{B}_{\text{MBHB}/\text{glitch}}$	$\sim -6.1 \times 10^4$	$\sim -3.9 \times 10^4$	$\sim 1.5 \times 10^7$
$d\log_{10}\mathcal{B}_{\text{MBHB}/\text{glitch}}$	~ 0.61	~ 0.22	~ 0.44

components. In potential edge cases, the existing methodology in the EREBOR global fit framework [57] offers a viable and complementary strategy for MBHB detection.

We show in Table VI three representative examples of the log Bayes factor for stretches N6, N24, and N29, where a signal (either a glitch or an MBHB) was identified. For brevity, we report only the results obtained using the stepping-stone method. In Appendix F are reported the results obtained using the thermodynamic integration for completeness. The Bayes factors across all considered stretches were found to be highly decisive, strongly favoring one hypothesis over the other, either the presence of a glitch or that of an MBHB. For this reason, we present only a subset of representative results here, as the conclusions were clear-cut in all cases. Specifically, for stretches N6 and N24, the presence of a glitch is strongly favored, while for stretch N29, the MBHB hypothesis is preferred.

During the search phase, along with computing the Bayes factor, we also obtain a rough estimate of the parameters of the MBHB. However, doing parameter estimation for the MBHB using only a day of data highlights several key considerations regarding frequency-domain analysis and the use of the discrete Fourier transform. One primary concern is spectral leakage, particularly evident when the premerger/inspiral phase of the signal is analyzed, which necessitates the use of windowing techniques to mitigate its effects [129]. Indeed, when analyzing frequency-domain waveforms over short time segments, particular caution is required to ensure accuracy. However, in the context of this research, we did not window the data as it was not strictly necessary since we were not interested in characterizing the premerger part of the signals.⁷ These issues reflect a fundamental challenge in performing low-latency analyses of MBHB signals [130], but once again, we do not need to be so careful in the search phase, as the final characterization of the signals uses the full dataset.

To conclude on a more technical side, to be able to perform the analysis of the MBHB with BBH_x but restricting it to a single day, we need to account for the time window considered when generating the MBHB template with BBH_x . This is achieved by shifting the phase with respect to the starting time of our window and by defining a start and an end time of the observation to cut the waveform,

⁷As illustrated in Appendix E, filtering the data favors no correlation between adjacent frequency bins.

TABLE VII. Table of priors for the glitch and noise parameters. We consider a single broad prior on $\tau_{0_{gl_1, gl_2, gl_3}}$ that encompasses all three glitches' injection times.

Parameter	Lower bound	Upper bound
$\tau_{0_{gl_1, gl_2, gl_3}}$ (s)	432,000	2,627,995
$\ln(A_0/[\text{ms}^{-2}])$	-32	-20
β_0 (s)	1	1×10^4
N_1 (m/ $\sqrt{\text{Hz}}$)	7×10^{-12}	8×10^{-12}
N_2 (m/ $\sqrt{\text{Hz}}$)	2.5×10^{-12}	3.5×10^{-12}
N_{TM} (m/s ² / $\sqrt{\text{Hz}}$)	2.2×10^{-15}	2.7×10^{-15}

such that any frequency components of the signal that are emitted outside of this range are excluded.

B. Parameter estimation phase

For the final parameter estimation phase we consider the full month of data. We include an MBHB template in our likelihood and consider the same prior as the one reported in Table III, but restrict the prior on the time of coalescence to $t_{\text{ref}} - 24 \text{ s} < t_{\text{ref}} < t_{\text{ref}} + 24 \text{ s}$ according to the MBHB posteriors found during the search phase. We also consider a slightly restricted prior range for the glitch parameters based on the findings from the search phase, as shown in Table VII. In particular, we restrict the times of injection

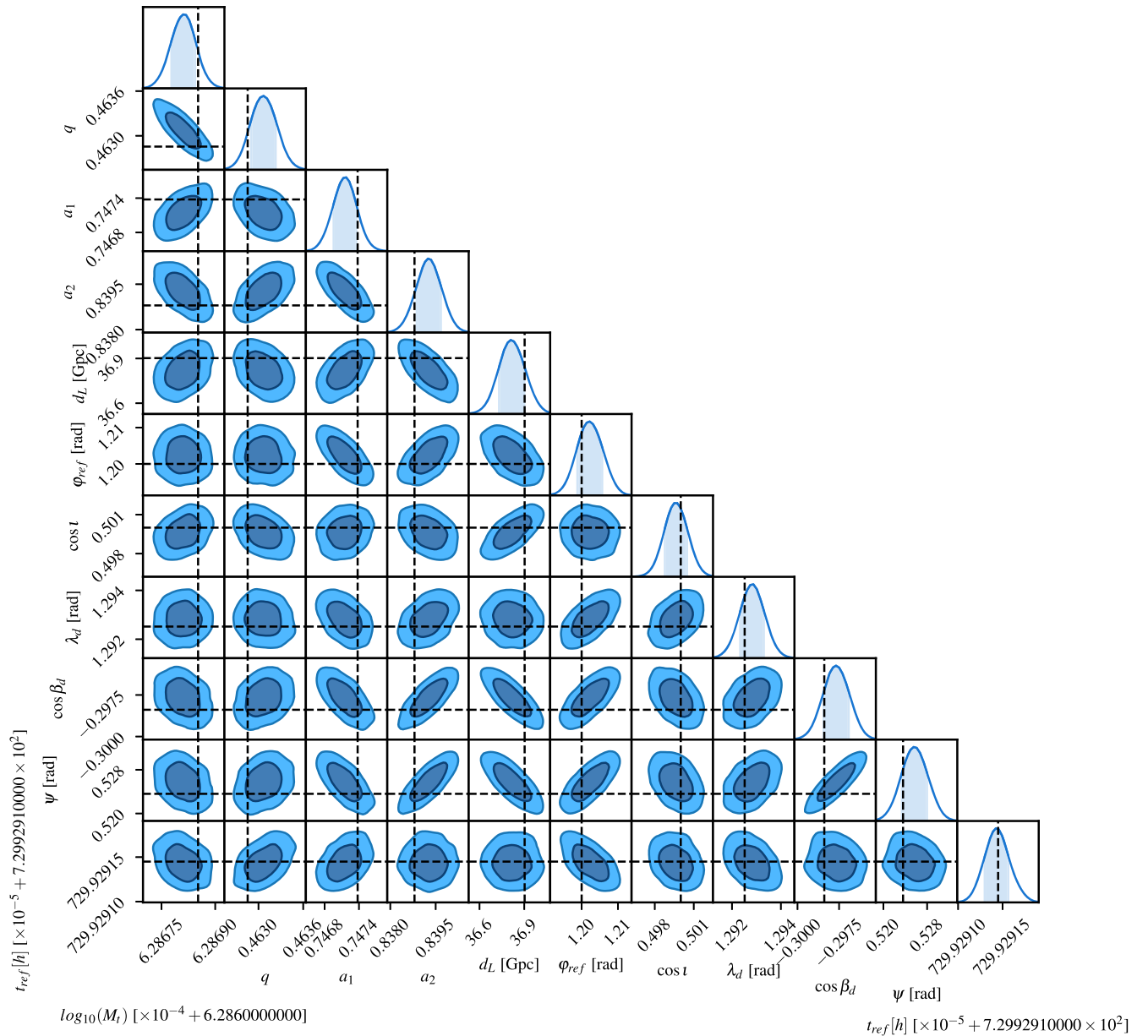


FIG. 19. Posteriors of the MBHB in the light-Spritz dataset.

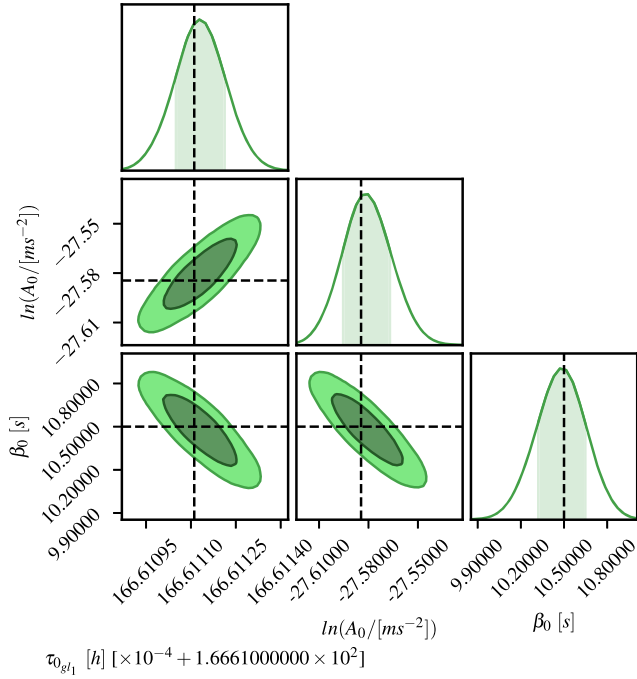


FIG. 20. Posteriors of the glitch injected around 166 h in the light-Spritz dataset.

and exclude combinations of β_0 and A_0 that do not reflect the glitch distribution found in LPF. If such glitches were present, they would have already been identified during the search phase, given the high SNR of those glitches.

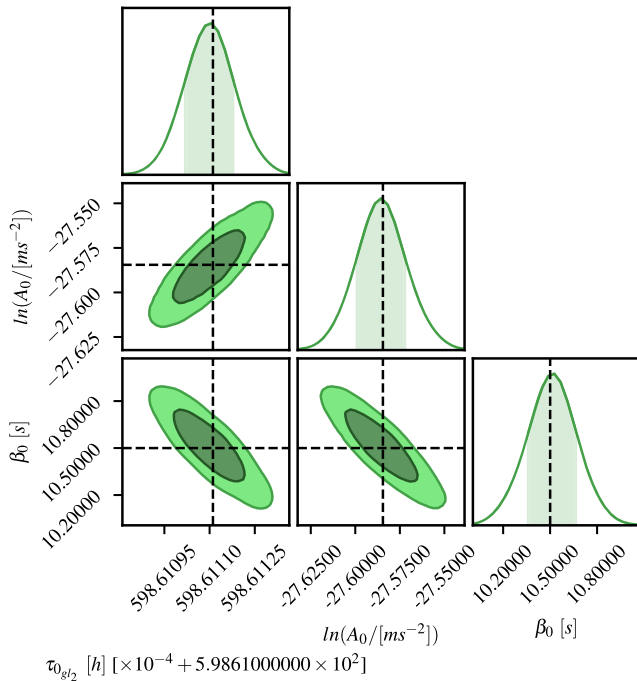


FIG. 21. Posteriors of the glitch injected around 598 h in the light-Spritz dataset.

We then set the minimum number of glitches for the reversible jump to match the number of high-SNR glitches identified in the search phase, and we set the maximum number of glitches to be double the minimum to allow for glitches that could have been missed during the search.

We consider the same in-model and out-of-model moves—that were used during the first search phase—for the glitches. The in-model moves get updated every 20 iterations up to a total of 50,000 iterations. We also incorporate the covariance matrices computed from the three estimated glitches—during the search phase—to define an additional reversible jump move proposal (see Sec. IV A 2) which proposes the birth or death of a glitch based on those identified during the search. To be on the safe side, we still allow 40% of proposed glitch parameters to be drawn from the full prior range, ensuring that we do not miss any potential glitches.

Instead, for the in-model Gaussian move proposal of the MBHB we estimate the covariance matrix using samples from the MBHB parameter posteriors obtained during the search phase.

Finally, we initialize the MCMC walkers using samples found during the search phase, and we estimate the parameters of the noise, signal, and glitches simultaneously following a Gibbs sampling approach [67].

The posteriors for the MBHB are shown in Fig. 19, for the three glitches are shown in Figs. 20–22, and for the noise parameters in Fig. 23. The MBHB has an SNR of 1522, while the glitches have SNRs of around 300. For all the posteriors, we only plot the last 500 samples of each walker, as we consider the previous samples, where we

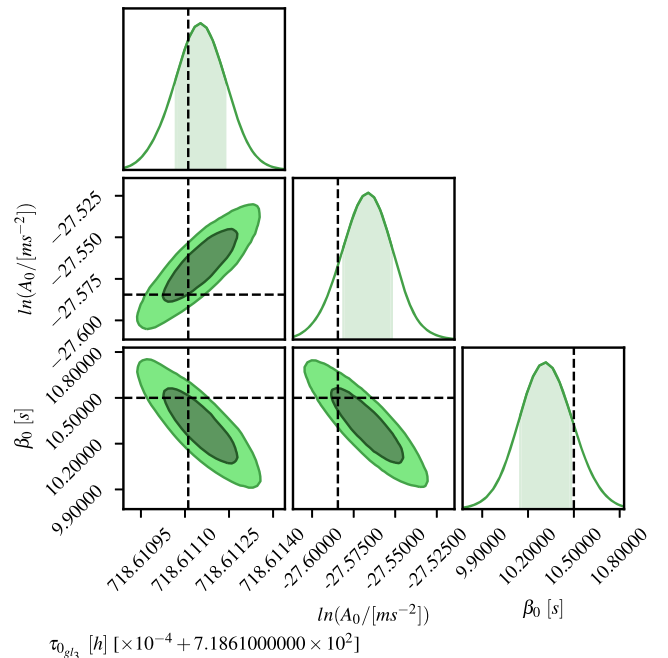


FIG. 22. Posteriors of the glitch injected around 718 h in the light-Spritz dataset.

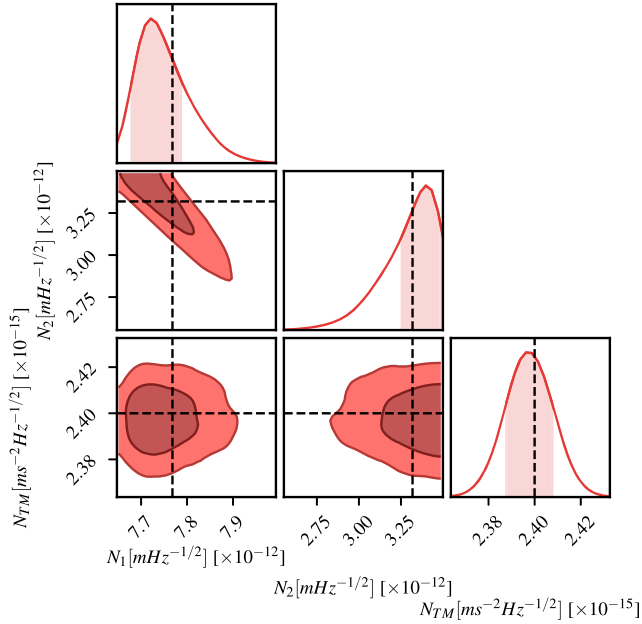


FIG. 23. Posteriors of noise parameters estimated in the light-Spritz dataset.

perform recursive updates of the in-model moves, as a burn-in period.

IX. CONCLUSION AND FUTURE PERSPECTIVE

In this work, we have demonstrated a potential global fit approach for LISA to estimate MBHB parameters in the presence of noise artifacts by implementing a search pipeline to find and fit for glitches along with the noise properties of the instrument and the parameters of an MBHB, using realistic simulated LISA data. Our methodology provides a general framework for detecting and modeling glitches, ensuring that such instrumental artifacts do not compromise LISA’s scientific output.

We have developed and validated our pipeline in two steps. First, we generated synthetic data where multiple glitches of varying SNRs, including those closely spaced in time, were injected to test the effectiveness of our reversible jump approach in handling diverse glitch morphologies. Our injected glitches were randomly drawn from the LPF glitch distribution to provide a realistic test scenario. We established an approximate threshold for distinguishing glitches from instrumental noise and investigated the impact of low-, medium-, and high-SNR glitches on the parameter estimation of a fiducial MBHB with an SNR of 1152. We found that unfitted glitches with $\text{SNR} \gtrsim 60$ introduce non-negligible biases in MBHB parameter recovery. For instance, for an unfitted glitch with SNR 21 we observed a slight shift in the parameter posterior relative to the injected values, but this shift was small compared to the statistical uncertainties, whereas for a glitch with SNR 72,

the true values were excluded at high confidence from the posterior support for several parameters, such as the total mass or the sky location. For the extreme case of an SNR 1544 glitch, biases were very large, and the posterior was completely disjointed from the case where no glitch is present. Moreover, we noticed that the relative SNR between the glitch and the MBHB signal, as well as the glitch’s temporal proximity to the merger, are critical factors in determining the extent of bias. However, these are not the sole contributors. For instance, when comparing the posterior distributions of MBHB systems with SNRs 1152 and 320, glitch number no. 41 in the LPF catalog appears to cause a more pronounced bias in the recovery of parameters for the low-SNR system, hinting that also the relationship between the glitch frequency and the instantaneous frequency of the MBHB plays a role. When fitting for the glitch(es) (single glitch as well as multiple glitches), we corrected for these biases and estimated the parameters of the MBHB correctly, with almost the same accuracy when no glitch was present. This study stresses the importance of accounting for glitches to fully exploit the tremendous scientific potential of MBHB observations with LISA—whether for fundamental physics [14–22], astrophysics [3–13], or cosmology [26–28].

Second, we applied our search algorithm to a simplified version of the Spritz dataset, which was obtained by removing the Galactic binaries and gaps from the original dataset. Our results show that we can successfully implement a search and a parameter estimation pipeline to detect and model glitches present in the data while recovering both the instrumental noise and the MBHB parameters without bias. Additionally, we provided a detailed noise model for this dataset after identifying inconsistencies between the official LDC Spritz documentation [119] and the actual procedure used to generate data for this mock dataset. This information will be valuable for future researchers analyzing the same dataset.

A key aspect that warrants further investigation is the accuracy and efficiency of the GMM clustering approach used to build proposal distributions for glitches in the parameter estimation phase, particularly in determining the number of components. This is particularly important in view of including the search and parameter estimation for glitches, developed in this work, in the framework of the EREBOR global fit [57], in which multiple MBHBs and other types of GW sources are present. In the current implementation, the number of clusters is fixed to match the maximum number of leaves (components) in the model, without attempting to infer the true number of underlying glitch populations. While this approach offers flexibility and has demonstrated to work very well in our framework, it may introduce inefficiencies when the number of GMM components significantly exceeds the actual number of distinct glitches. More specifically, in scenarios where the

GMM overestimates the number of glitch clusters, the model may fragment individual glitch posteriors into subposteriors, which can hinder efficient sampling. As a result, MCMC chains may struggle to traverse the full posterior landscape, especially across different glitch modes, even if local acceptance rates remain high. We have begun exploring alternative strategies to dynamically estimate the appropriate number of clusters, potentially improving the exchange of samples across posterior modes and enhancing the overall sampling efficiency. This might play a more significant role when more GW sources or more complicated noise components, with respect to the light-Spritz dataset, are considered.

Nonetheless, the glitch search, determination, and parameter estimation components developed in this work will form the foundation of the glitch handling module within EREBOR. From EREBOR, we will incorporate the already developed modules for MBHB searches and parameter estimation, as well as the search and parameter estimation for Galactic binaries, noise, and foreground fitting. Over time, we expect to refine both the glitch and MBHB modules based on the insights gained here, particularly in transdimensional sampling strategies. We will also update the noise fitting by leveraging the more complete noise models reported in this work.

Future work will also extend this approach to account for instrumental nonstationarities, such as data gaps and the presence of the Galactic foreground, so that data with the more realistic properties of the Spritz dataset can be handled. While our study serves as a guideline for glitch mitigation in LISA, it does not encompass all possible MBHB populations, and our reference glitch distribution is based on LPF data. However, the shapelet-based approach should readily adapt to different glitch morphologies, with respect to the signal exponential ones addressed in this paper, present in the data, in particular by leveraging higher order shapelets, but its efficacy needs to be properly quantified. However, since LISA will use similar technology to LPF, at least for the GRS, the acceleration glitches are unlikely to differ substantially, so the results obtained here should represent the expected performance on actual LISA data.

To conclude, in this work, we have adopted simplifying assumptions in constructing the orthogonal TDI combinations A , E , and T . Specifically, we assumed equal and constant arm lengths and uncorrelated noise across different channels. These assumptions simplify the likelihood by yielding a diagonal covariance matrix and lead to identical TDI transfer functions for the A and E channels. These assumptions are appropriate for our analysis of the Spritz data, which we restricted to utilize the A and E channels, where the impact of using the equal-arm approximation is less severe than for the null-channel T . Moreover, we excluded in the analysis the high-frequency region near the

nulls of the TDI transfer functions, where A and E would also be heavily affected; this should not meaningfully affect our results, as all signals considered here lie at lower frequencies. While such idealizations are appropriate for the scope of this study, we acknowledge that, in the actual LISA mission, the arm lengths will be unequal and time varying, and noise correlations between channels might not be negligible. This is especially true for the high-frequency regime, and ongoing studies are investigating how to properly model these effects and incorporate nonstationarities arising from orbital dynamics. However, these effects are more relevant for sources such as stellar-origin black hole binaries inspiraling in the LISA band at high frequencies. For instance, preliminary investigations in [123] suggest that alternative TDI combinations with fewer delays than the standard Michelson channels X , Y , and Z could be preferable, as the nulls in the TDI transfer function occur at higher frequencies with respect to the Michelson channels. For the low-frequency regime, the main improvement by moving away from our simplifying assumptions lies in being able to exploit a null channel, like T , which might allow us to put stronger constraints on the noise during the search phase or to better characterize it during the parameter estimation phase.

All codes used in this work are open-source and can be accessed on GitHub [131].

ACKNOWLEDGMENTS

The authors thank the following people (alphabetically ordered) for useful discussion in improving the work: Quentin Baghi, Jean-Baptiste Bayle, Ollie Burke, Eleonora Castelli, Raffi Enficiaud, Hector Hestelles, Niklas Houba, Nikos Karnesis, Lorenzo Pompili, Lorenzo Sala, Elise Sang er, Alessandro Santini, Lorenzo Speri, Alice Spadaro, Mauro Pieroni, and Sebastian V lkel. M.M. and O.H. gratefully acknowledge the support of the German Space Agency, DLR. The work is supported by the Federal Ministry for Economic Affairs and Climate Action based on a resolution of the German Bundestag (Project Ref. No. FKZ 50 OQ 2301). A. T. is supported by MUR Young Researchers Grant No. SOE2024-0000125, ERC Starting Grant No. 945155–GWmining, Cariplo Foundation Grant No. 2021-0555, MUR PRIN Grant No. 2022-Z9X4XS, MUR Grant “Progetto Dipartimenti di Eccellenza 2023–2027” (BiCoQ), and the ICSC National Research Centre funded by NextGenerationEU.

DATA AVAILABILITY

All the codes and data that I have been used for the paper are available in my github page [131] and the data for Spritz analysis can be found here [69,119].

APPENDIX A: GLITCH MODEL

1. Derivation of shapelet model in time and frequency domains

Shapelets can be defined in the time domain as follows:

$$\psi_n(t) = 2 \frac{t}{n} e^{-\frac{t}{n}} L_{n-1}^1 \left(2 \frac{t}{n} \right) \Theta(t), \quad (\text{A1})$$

with L_n^α denoting a generalized Laguerre polynomial. A glitch perturbation is then modeled by a linear combination of shapelets,

$$g(t) = \sum_{j=0}^P A_j \psi_{n_j} \left(\frac{t - \tau_j}{\beta_j} \right), \quad (\text{A2})$$

with $P + 1$ being the number of shapelets. Substituting Eq. (A1) into Eq. (A2) we have

$$g(t) = \sum_{j=0}^P 2A_j \frac{1}{n_j} \frac{t - \tau_j}{\beta_j} e^{-\frac{t - \tau_j}{\beta_j}} \times L_{n_j-1}^1 \left(2 \frac{t - \tau_j}{\beta_j} \right) \Theta \left(\frac{t - \tau_j}{\beta_j} \right). \quad (\text{A3})$$

Note that Eq. (A3) refers to a glitch in acceleration, but we need to derive the corresponding fractional frequency shift, which means integrating $g(t)$ and dividing by the velocity of light c .

To perform the integral, we can consider the generating function

$$\sum_{n=0}^{\infty} v^n L_n^\alpha(x) = \frac{1}{(1-v)^{\alpha+1}} e^{-\frac{vx}{1-v}}. \quad (\text{A4})$$

To align it with Eq. (A1), we adjust the index of the series as follows:

$$\sum_{n=1}^{\infty} v^n L_{n-1}^\alpha(x) = \frac{v}{(1-v)^{\alpha+1}} e^{-\frac{vx}{1-v}} \quad (\text{A5})$$

and then define

$$\begin{aligned} G(t, \tau, n, \beta, A, v) &= \frac{2A}{c} \int_0^t \left(\frac{t' - \tau}{n\beta} \right) \exp \left[-\frac{t' - \tau}{n\beta} \right] \\ &\times \frac{v}{(1-v)^{\alpha+1}} \exp \left[-2 \frac{t' - \tau}{n\beta} \frac{v}{(1-v)} \right] \\ &\times \Theta \left(\frac{t' - \tau}{\beta} \right) dt'. \end{aligned} \quad (\text{A6})$$

This expression enables us to obtain different shapelet orders by rewriting the relationship between the left and right components of Eq. (A6) as follows:

$$\begin{aligned} \frac{\Delta v_g}{\nu_0}(t, \vec{\tau}, \vec{n}, \vec{\beta}, \vec{A}) &\equiv \int_0^t \frac{g(t')}{c} dt' \\ &= \sum_{i=0}^P \frac{1}{n_i!} \left[\frac{d^{n_i}}{dv^{n_i}} G(t, \tau_i, n_i, \beta_i, A_i, v) \right]_{|v=0}. \end{aligned} \quad (\text{A7})$$

Equation (A6) can be explicitly written as

$$\begin{aligned} G(t, \tau, n, \beta, A, v) &= \frac{2vA}{c(1-v)^{\alpha+1}} n\beta \int_{-\frac{\tau}{n\beta}}^{\frac{t-\tau}{n\beta}} u e^{-u} e^{-\frac{2vu}{1-v}} \Theta(nu) du \\ &= -\frac{2vnA\beta(1-v)^{-\alpha}}{c(1+v)^2} \times \Theta \left(\frac{t-\tau}{\beta} \right) \\ &\times \left(-1 + v + \frac{e^{\frac{(1+v)(t-\tau)}{n(-1+v)\beta}} (t + tv + n\beta - nv\beta - (1+v)\tau)}{n\beta} \right). \end{aligned} \quad (\text{A8})$$

In this way, the generating function is used to perform the integral. Equation (A8) needs to be computed for each value of A_j , τ_j , and β_j that is needed such that the sum over the shapelets can be performed.

We can now derive the Fourier transform of $\frac{\Delta v_g}{\nu_0}(t, \vec{\tau}, \vec{n}, \vec{\beta}, \vec{A})$. This reads

$$\frac{\Delta \tilde{v}_g}{\nu_0}(f, \vec{\tau}, \vec{n}, \vec{\beta}, \vec{A}) = \sum_{j=0}^P \frac{1}{n_j!} \left[\frac{d^{n_j}}{dv^{n_j}} \tilde{G}(f, \tau_j, n_j, \beta_j, A_j, v) \right]_{|v=0}. \quad (\text{A9})$$

Considering Eq. (A8), we obtain

$$\tilde{G}(f, \tau, n, \beta, A, v) = \frac{iAn(1-v)^{1-\alpha} v \beta e^{-2if\pi\tau}}{cf\pi(i(1+v) + 2fn\pi(-1+v)\beta)^2}, \quad (\text{A10})$$

and therefore,

$$\frac{\Delta \tilde{v}_g}{\nu_0}(f, \vec{\tau}, \vec{n}, \vec{\beta}, \vec{A}) = \quad (\text{A11})$$

$$\sum_{j=0}^P \frac{1}{n_j!} \left[\frac{d^{n_j}}{dv^{n_j}} \left(\frac{iAn(1-v)^{1-\alpha} v \beta e^{-2if\pi\tau}}{cf\pi(i(1+v) + 2fn\pi(-1+v)\beta)^2} \right) \right]_{|v=0}. \quad (\text{A12})$$

The first-order shapelet ($j = 0$) in the frequency domain is

$$\frac{\Delta \tilde{v}_g}{\nu_0}(f, \tau_0, \beta_0, A_0) = \frac{iA_0\beta_0 e^{2if\pi\tau_0}}{cf\pi(i - 2f\pi\beta_0)^2}. \quad (\text{A13})$$

2. Glitch model in time and frequency domain and respective TDI transfer functions

The effect of a typical LPF glitch is evaluated starting from a model for its differential acceleration, $\Delta g(t)$ [62,74],

$$h_g(t) = \frac{\Delta v}{2} t' e^{-\frac{t'}{\tau}} \Theta(t'), \quad t' = t - t_0 \quad (\text{A14})$$

where t_0 is the occurrence time, $\Theta(t)$ is the Heaviside step function, Δv is the total transferred impulse per unit mass, and the τ parameter determines the duration of the event. We can also express the fractional frequency shift from the glitch as

$$\frac{\Delta \nu_g(t)}{\nu_0} = \frac{\Delta v}{c} \left(1 - \frac{e^{-\frac{t-T}{\tau}}(t-T+\tau)}{\tau} \right) \Theta(t-t_0). \quad (\text{A15})$$

From a frequency-domain perspective, Δv is also the zero-frequency limit of the Fourier transform of the $h_g(t)$ template. This is relevant, since it implies that it has a strong low-frequency component. The Fourier transform of the template is

$$\frac{\Delta \nu_g(f)}{\nu_0} = -\frac{\Delta v e^{-i2\pi f t_0}}{c i 2\pi f (-i + 2\pi f \tau)^2}, \quad (\text{A16})$$

where we need to divide by $c\omega$ to express it as fractional frequency shift. The expression is similar to the shapelet when $n_0 = 1$ [Eq. (10)] with $\Delta v = 2A\beta$ and $t_0 = \tau$ and $\tau = \beta$.

Considering now Eq. (2) in Sec. III C, we can model the TDI response to a glitch as a product between the glitch expression in the frequency domain [Eq. (A16)] and the transfer function $\mathcal{T}_{X,Y,Z}^g(f)$ which takes into account TDI (see Sec. III C). We provide below the expressions for the first-generation TDI transfer functions, noting $\mathcal{F}D_{pq} = e^{-i(2\pi f)L}$:

In case of a glitch happening on TM_{21} we have

$$\mathcal{T}_{X_1}^g(f) = -2e^{-3iL(2\pi f)}(-1 + e^{2iL(2\pi f)}) \quad (\text{A17a})$$

$$\mathcal{T}_{Y_1}^g(f) = 1 - e^{-4iL(2\pi f)} \quad (\text{A17b})$$

$$\mathcal{T}_{Z_1}^g(f) = 0. \quad (\text{A17c})$$

In case the same glitch is happening on TM_{13} ,

$$\mathcal{T}_{X_1}^g(f) = 1 - e^{-4iL(2\pi f)} \quad (\text{A18a})$$

$$\mathcal{T}_{Y_1}^g(f) = 0 \quad (\text{A18b})$$

$$\mathcal{T}_{Z_1}^g(f) = -2e^{-3iL(2\pi f)}(-1 + e^{2iL(2\pi f)}) \quad (\text{A18c})$$

or on TM_{31}

$$\mathcal{T}_{X_1}^g(f) = 4ie^{-2iL(2\pi f)} \sin(L(2\pi f)) \quad (\text{A19a})$$

$$\mathcal{T}_{Y_1}^g(f) = 0 \quad (\text{A19b})$$

$$\mathcal{T}_{Z_1}^g(f) = -1 + e^{-4iL(2\pi f)} \quad (\text{A19c})$$

or on TM_{23}

$$\mathcal{T}_{X_1}^g(f) = 0 \quad (\text{A20a})$$

$$\mathcal{T}_{Y_1}^g(f) = -1 + e^{-4iL(2\pi f)} \quad (\text{A20b})$$

$$\mathcal{T}_{Z_1}^g(f) = 4ie^{-2iL(2\pi f)} \sin(L(2\pi f)) \quad (\text{A20c})$$

and finally in case of TM_{23} ,

$$\mathcal{T}_{X_1}^g(f) = 0 \quad (\text{A21a})$$

$$\mathcal{T}_{Y_1}^g(f) = -2e^{-3iL(2\pi f)}(-1 + e^{2iL(2\pi f)}) \quad (\text{A21b})$$

$$\mathcal{T}_{Z_1}^g(f) = 1 - e^{-4iL(2\pi f)}. \quad (\text{A21c})$$

For the second-generation TDI, a similar computation can be done using Eq. (3) in Sec. III C.

APPENDIX B: POSTERiors OF SNR 1152 MBHB WITH HIGH SNR GLITCH OF 1544 AT DIFFERENT INJECTION TIMES WITH RESPECT TO ITS MERGER TIME

The impact of a glitch in the recovery of an MBHB is maximized when the glitch occurs closer to the merger; this, in practice, means that the same type of glitch would not have as great an impact if it happened at a different time. To show this, we also report below the impact on parameter estimation of the fiducial MBHB with SNR 1152 and redshift 4—considered also in the analysis of the light-Spritz dataset—with the 1544 SNR glitch injected at different times with respect to maximum overlap computed in Sec. VI: five hours before the merger and five minutes after the merger (Fig. 24). In all cases, noisy data have been used, which introduce some random statistical fluctuations that are visible in the posteriors.

Looking at Fig. 24, where the glitch has been injected five hours before the merger time, we can see that not fitting for the glitch causes less bias with respect to the glitch occurring five minutes after the merger. We expect that glitches happening earlier would have an even smaller effect. If the glitch occurs closer to or at the merger, not fitting for it clearly biases the parameters of the MBHB, such as the masses, spins, or sky positions. This behavior stems from the fact that the merger is the loudest part of the signal and drives the parameter estimation of the MBHB, so glitches that corrupt that part of the signal have a stronger effect.

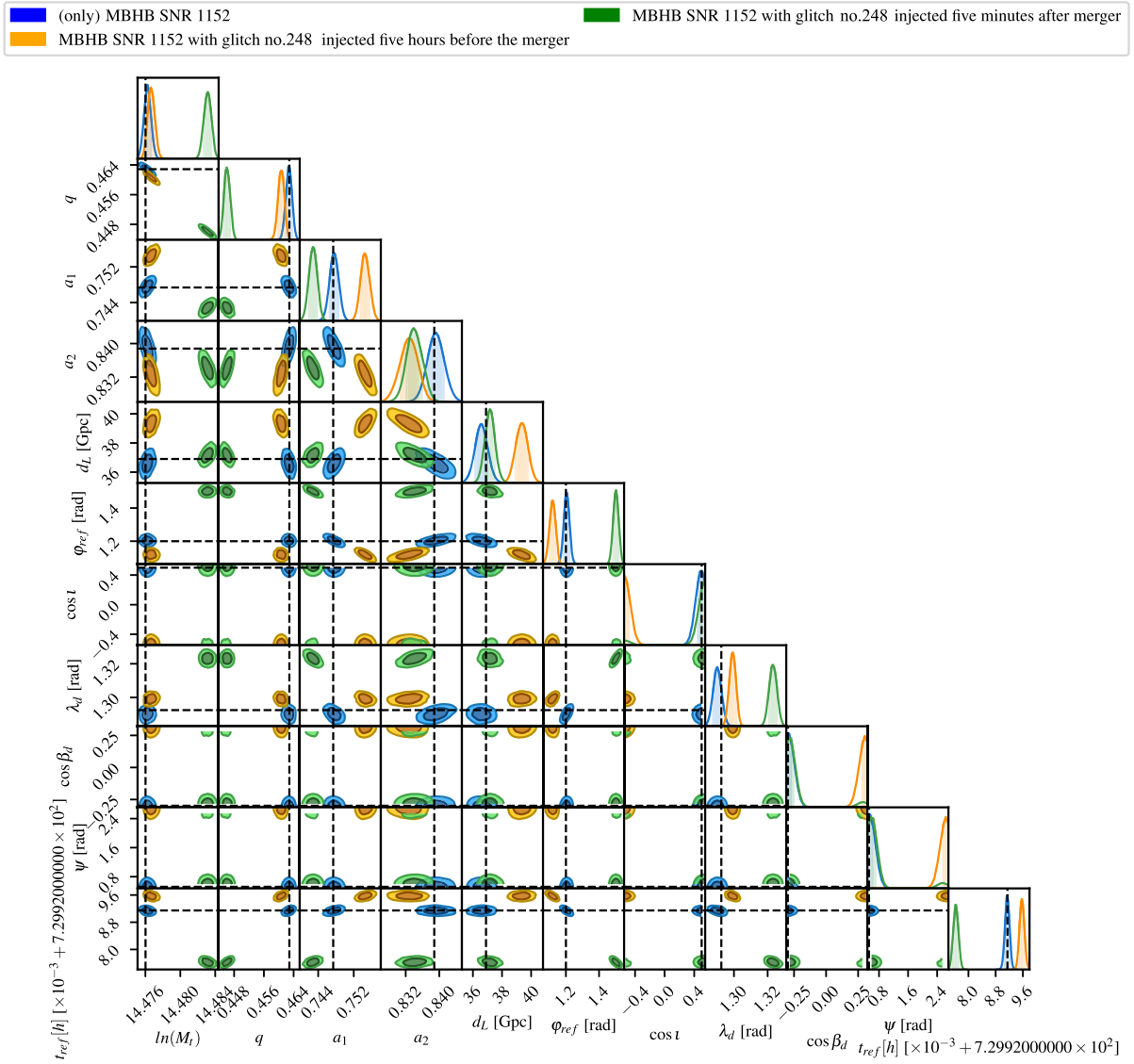


FIG. 24. Posteriors of MBHB 1152 SNR along with posterior of the same MBHB with a 1544 SNR glitch (no. 248) injected five minutes after the merger time and injected five hours before the merger time.

APPENDIX C: POSTERIOR OF SNR 320 MBHB WITH GLITCHES N41 AND N3 INJECTED AT THEIR MAXIMUM OVERLAP

In Fig. 10 are reported the posteriors of an SNR 320 MBHB when the two glitches corresponding to glitch nos. 41 and 3 of the LPF catalog are injected. The parameters of the MBHB are reported in Table VIII. The priors are the same as Table III, but the prior on the total mass M_t has been converted to Log_e , i.e., (\ln).

The glitches were injected at their maximum overlap accordingly to Eq. (33), that is, 0.19 h for glitch no. 41 and 0.11 h for glitch no. 3, as can be seen from Fig. 25. In Fig. 25, we show the two glitches injected at the time that maximizes the overlap between the glitch and the fiducial MBHB in the TDI channel A. It is possible to observe that

TABLE VIII. Massive black hole binary true values. Values of the masses are given in the detector frame. We express the masses as dimensionless multiples of the solar mass M_\odot .

Parameter	True value
$\ln(M_t)$	17.91
q	0.33
a_1	0.3
a_2	0.4
d_L (Gpc)	47.6
φ_{ref} (rad)	1.2
$\cos i$	0.29
λ (rad)	2.0
$\sin \beta$	0.82
ψ (rad)	1.6
t_{ref} (h)	30

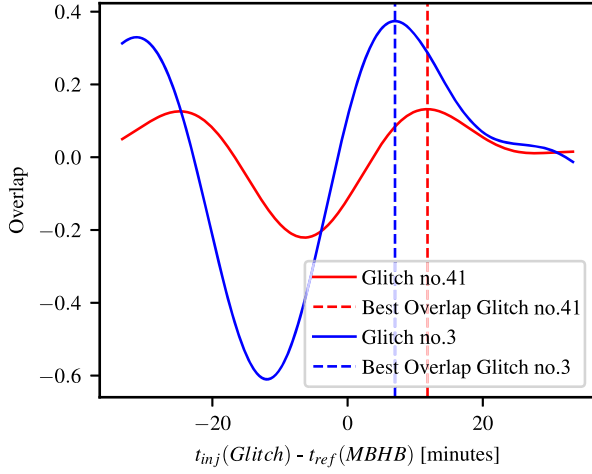


FIG. 25. Overlap between the MBHB and glitch signals as a function of the glitch injection time relative to the merger, where the decay time and amplitude of the glitches are kept fixed. The red and blue solid curves are for the two glitches with SNRs of 21 and 72. The GW signal is fixed to that of the parameters in Table VIII. Dashed vertical lines with matching colors denote the times that maximize the overlap.

sky parameters have multimodality. Indeed, there are up to eight modes in the sky that are exactly degenerate for a fixed LISA configuration and nonevolving sources. However, this degeneracy is broken by LISA's motion and changes in frequency. High-mass and short-duration sources, like this case, are more likely to exhibit such degeneracies. The degeneracy primarily involves sky location and polarization (β_d , ψ , λ_d , and $\cos \iota$) [91,112], as can

also be seen in the posteriors plot showing the extrinsic parameters (right plot Fig. 26) for the case of (only) MBHB posteriors or when a low-SNR 21 glitch is injected but not fitted. The left plot in Fig. 26 shows the posteriors for the intrinsic parameters.

APPENDIX D: DISCUSSION ON POSTERIORES FOR NOISE ESTIMATED DURING THE SEARCH PHASE

Figure 27 shows the posterior distributions for the three noise parameters (N_1 , N_2 , and N_{TM}), estimated during the search for segments N0, N2, N5, N8, N10, N13, N15, N20, and N26, along with the true values reported in Sec. VII B. The TM acceleration noise is not well characterized, showing a bias away from the true value. This bias increases for data chunks that include the late inspiral phase of the MBHB (segments N26–N28), as this occurs in the frequency range around 10^{-3} Hz where TM noise dominates, as also visible in Fig. 28.

The N_2 component spans the entire prior range, indicating poor constraint. In contrast, N_1 is better estimated, as it dominates at higher frequencies. Even with just about a day of data, its value can be meaningfully constrained. On the other hand, the TMI noise remains poorly characterized, as its contribution lies in a frequency range where it is overshadowed by the stronger OMS noise, as discussed in Sec. VII B, or in other words the OMS noise that enters in N_2 is much smaller than the OMS noise that contributes to N_1 . Longer data segments are needed to improve the inference of TMI noise, as demonstrated in Sec. VI, where the fit is performed on the full light-Spritz dataset.

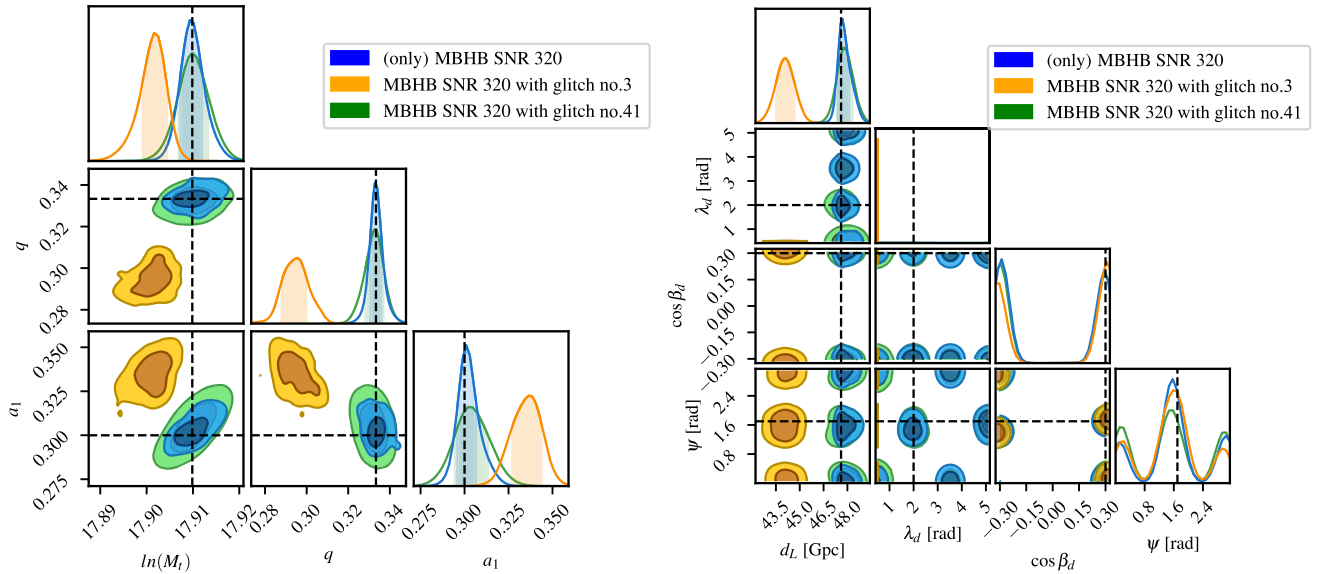


FIG. 26. Posteriors of a subset of intrinsic parameters (left) and a subset of extrinsic parameters (right) for the parameters of an MBHB with SNR 320 obtained from data containing only the MBHB, compared to those obtained from data containing a glitch of SNR 21 and of SNR 72, when the latter are not fitted.

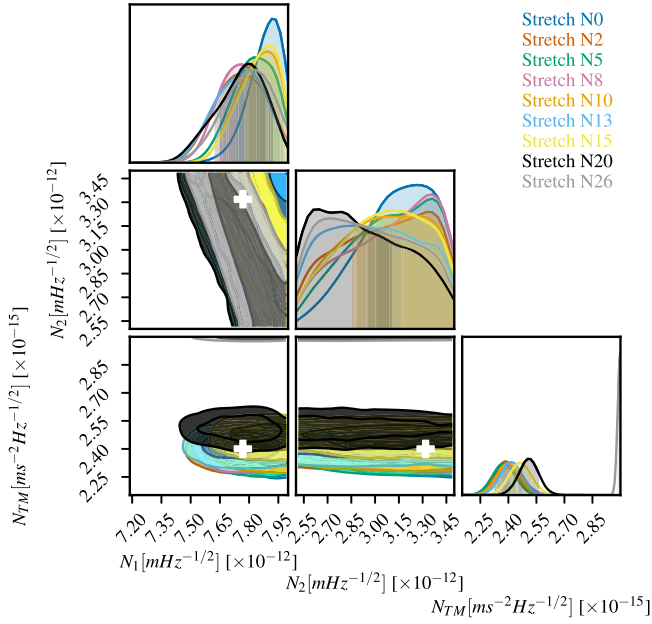


FIG. 27. Estimation of amplitude for the components of the test-mass acceleration noise, test-mass interferometer, interspacecraft interferometer, and reference interferometer.

Nonetheless, this initial estimate is useful as an indicator of a reasonable prior range to be used during the parameter estimation phase, where the full dataset is analyzed.

APPENDIX E: ORTHOGONALITY OF FAST FOURIER TRANSFORM OF SPRITZ-LIKE DATA

In Sec. VII A, we propose to low-pass filter the light-Spritz data before taking the FFT to better approximate a diagonal likelihood. To verify this approach, we produce a

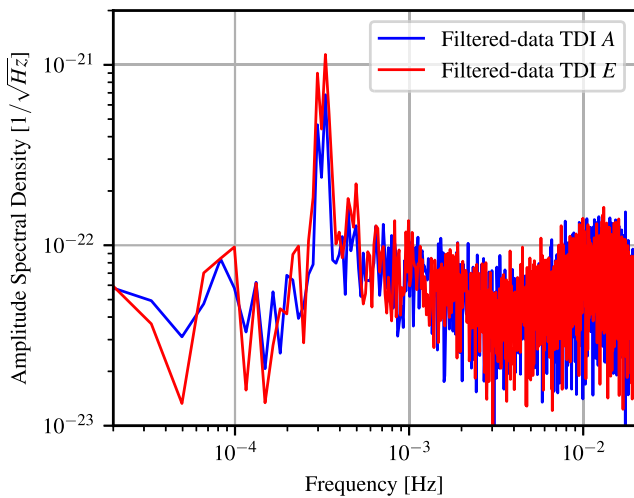


FIG. 28. Amplitude spectral density of the TDI A and E estimated during the search phase for the stretch N26, where the inspiral phase of the MBHB is present.

toy model with simplified Spritz-like noise in the time domain and investigate the covariance matrix of the FFT. We use the same sampling rate as in Spritz, $f_s = 0.2$ Hz, and produce 500 realizations of a noise data stream of 4000 samples. The noise in this toy model is the sum of the two components,

$$n_{\text{toy}} = n_{\text{TM}} + n_{\text{OMS}}, \quad (\text{E1})$$

where the noise PSDs of n_{TM} and n_{OMS} are given by $C_{\text{acc}}^{\text{ffd}} S_{\text{TM}}$ and $C_{\text{disp}}^{\text{ffd}} S_{\text{OMS}}^{\text{ISI}}$, as described in Sec. VII B.

We approximate these noise shapes by adding white noise time series, shaped using repeated applications of `cumsum` and `gradient` operations. We then inject this noise into the variable η_{12} of the second-generation TDI Michelson combination [see Eq. (3)], evaluated with equal arms and computed using the open-source software PYTDI [132]. Expressed in terms of delay operators, this corresponds to

$$n_{\text{toy}}^{\text{TDI}} = (1 - D^2)(1 - D^4)n_{\text{toy}}. \quad (\text{E2})$$

We then consider three scenarios:

- (1) Directly compute the FFT without any preprocessing.
- (2) Compute the FFT after applying a window function (hann window).
- (3) Compute the FFT after applying a low-pass filter, as described in Sec. VII A.

After computing the FFT series, we estimate the covariance matrix by computing

$$C_{jk} = E[\text{FFT}_j \text{FFT}_k^*], \quad (\text{E3})$$

where FFT_j is the j th frequency bin of the FFT of $n_{\text{toy}}^{\text{TDI}}$, considering only positive frequencies. The expectation value is approximated by averaging over our 500 realizations. The square root of the diagonal elements of this

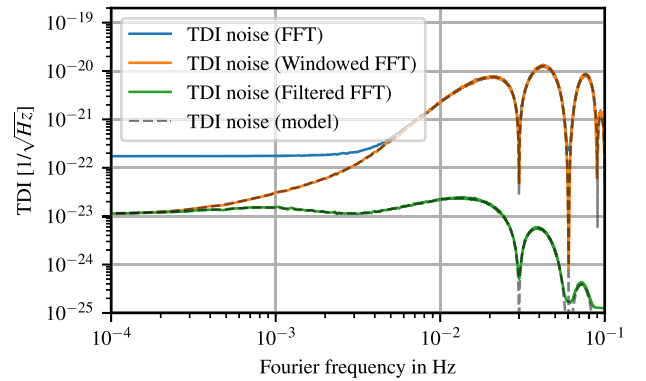


FIG. 29. Diagonal of FFT covariance matrix, scaled to amplitude spectral density for a toy model equal and constant arm-length TDI X_2 .

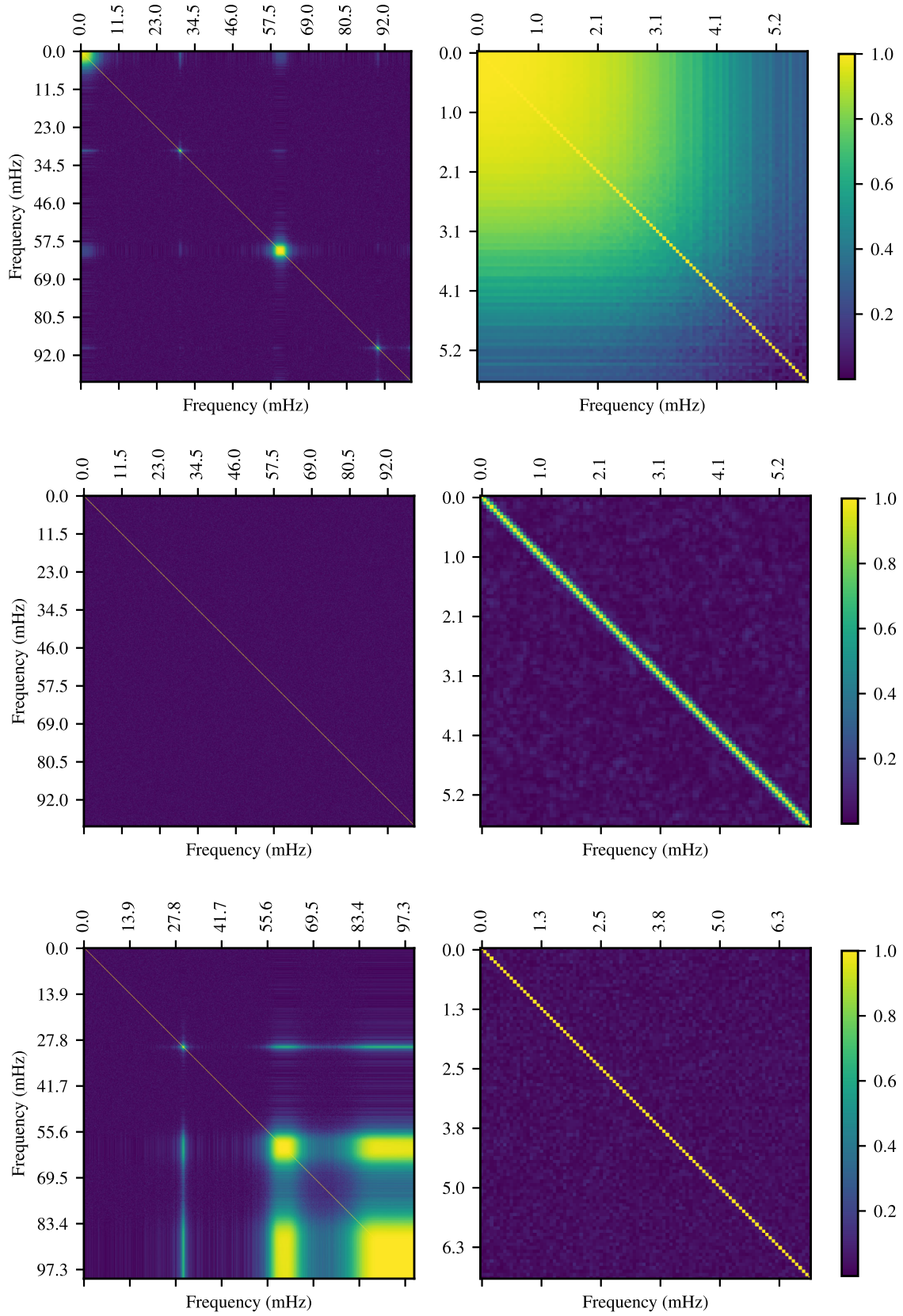


FIG. 30. Numerically estimated covariance matrix of FFT. Top: no preprocessing; middle: window function; bottom: low-pass filter.

matrix are displayed in Fig. 29, rescaled by $\sqrt{\frac{2}{f_s \sum w[j]^2}}$ to approximate an ASD [133]. Here, $w[j]$ are the window function coefficients applied to the j th sample.

The direct FFT is strongly affected by spectral leakage and deviates from the expected ASD at low frequencies, whereas the windowed FFT and filtered FFT approximate their respective model well. Note that the low-pass filter is accounted for in the modeled noise shape for the latter.

To study correlations between frequency bins with the three methods, we normalize C_{jk} by the values on the diagonal and plot

$$\tilde{C}_{jk} = |C_{jk} / \sqrt{C_{jj}C_{kk}}|. \quad (\text{E4})$$

The results are displayed in Fig. 30. For all three scenarios, we display the matrix for the full frequency range on the left as well as an enlarged version of just the first 100 frequency bins on the right. The top row of Fig. 30 is the direct FFT. It shows strong long-range correlations around the frequencies corresponding to the zeros of the TDI transfer function as well as at dc, but otherwise a narrow diagonal. The middle row is the windowed FFT. The window strongly suppresses the long-range correlations, but introduces strong correlations

in a narrow band along the diagonal. Finally, the filtered FFT shows even stronger correlations at high frequencies, but is perfectly diagonal in the band up to 20 mHz used for our analysis. This validates the use of a diagonal likelihood model in Sec. VII A.

APPENDIX F: BAYES FACTOR COMPUTATION

In Fig. 31, we report three examples of the stretches N6, N24, and N29 as an illustration of the methodology to compute the Bayes factor with thermodynamic integration. We use 70 temperatures to perform the thermodynamic integration method, where $\beta_k = \frac{1}{T_k}$ and T_k represents the temperatures. As expected [67], the stepping-stone algorithm produces more accurate estimates of the marginal likelihood than methods that use thermodynamic integration [134]; this is why in the computation shown here, the errors on the Bayes factor are higher. The y axis shows the average log-likelihood ($\log L$) at each temperature for the MBHB (glitch) model, and the x axis shows the inverse of the temperature. This curve is integrated to calculate the log evidence, $\log_{10} Z$. The difference between the log evidence of the MBHB model versus the glitch model gives the log Bayes factor, $\log_{10} \mathcal{B}_{\text{MBHB}/\text{glitch}}$.

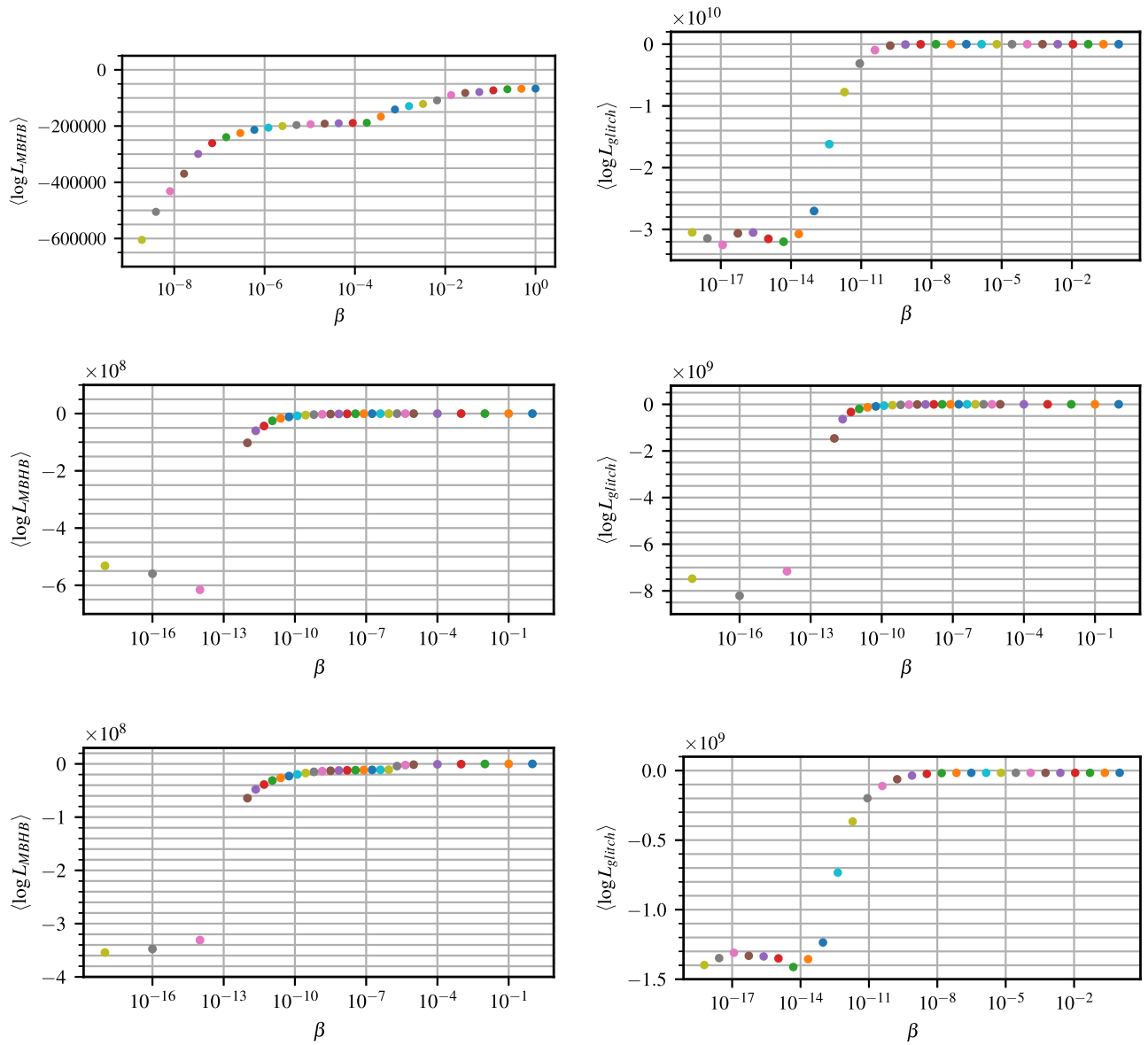


FIG. 31. The y axis shows the average log-likelihood ($\log L$) at each temperature for the MBHB (glitch) model, and the x axis shows the inverse of the temperature. For the top row: $\log_{10} \mathcal{B}_{MBHB/glitch} \sim -5.57 \times 10^4$ with $d \log_{10} \mathcal{B}_{MBHB/glitch} \sim 11843.7$ for stretch N6; for the middle row: $\log_{10} \mathcal{B}_{MBHB/glitch} \sim -2.93 \times 10^4$ with $d \log_{10} \mathcal{B}_{MBHB/glitch} \sim 55796.5$ for stretch N24; and for the bottom row: $\log_{10} \mathcal{B}_{MBHB/glitch} \sim -1.56 \times 10^7$ with $d \log_{10} \mathcal{B}_{MBHB/glitch} \sim 28048.2$ for stretch N29. $d \log_{10} \mathcal{B}_{MBHB/glitch}$ gives an estimate on the error on $\log_{10} \mathcal{B}_{MBHB/glitch}$. Note that we have omitted the base 10 in the plots.

[1] P. Amaro-Seoane *et al.*, arXiv:1702.00786.
 [2] M. Colpi *et al.*, arXiv:2402.07571.
 [3] J. R. Gair, A. Sesana, E. Berti, and M. Volonteri, *Classical Quantum Gravity* **28**, 094018 (2011).
 [4] A. Sesana, J. Gair, E. Berti, and M. Volonteri, *Phys. Rev. D* **83**, 044036 (2011).

[5] A. Klein *et al.*, *Phys. Rev. D* **93**, 024003 (2016).
 [6] P. Dayal, E. M. Rossi, B. Shiralilou, O. Piana, T. R. Choudhury, and M. Volonteri, *Mon. Not. R. Astron. Soc.* **486**, 2336 (2019).
 [7] E. Barausse, I. Dvorkin, M. Tremmel, M. Volonteri, and M. Bonetti, *Astrophys. J.* **904**, 16 (2020).

- [8] A. Toubiana, K. W. K. Wong, S. Babak, E. Barausse, E. Berti, J. R. Gair, S. Marsat, and S. R. Taylor, *Phys. Rev. D* **104**, 083027 (2021).
- [9] H.-Y. Chen, A. Ricarte, and F. Pacucci, [arXiv:2202.04764](https://arxiv.org/abs/2202.04764).
- [10] Y. Fang and H. Yang, *Mon. Not. R. Astron. Soc.* **523**, 5120 (2023).
- [11] A. Spadaro, R. Buscicchio, D. Izquierdo-Villalba, D. Gerosa, A. Klein, and G. Pratten, *Phys. Rev. D* **111**, 023004 (2025).
- [12] V. Langen, N. Tamanini, S. Marsat, and E. Bortolas, *Mon. Not. R. Astron. Soc.* **536**, 3366 (2025).
- [13] A. Toubiana, L. Sberna, M. Volonteri, E. Barausse, S. Babak, R. Enficiaud, D. I. Villalba, J. R. Gair, J. E. Greene, and H. Quelquejay Leclere, [arXiv:2410.17916](https://arxiv.org/abs/2410.17916).
- [14] E. Berti, V. Cardoso, and C. M. Will, *Phys. Rev. D* **73**, 064030 (2006).
- [15] E. Berti, A. Sesana, E. Barausse, V. Cardoso, and K. Belczynski, *Phys. Rev. Lett.* **117**, 101102 (2016).
- [16] K. Chamberlain and N. Yunes, *Phys. Rev. D* **96**, 084039 (2017).
- [17] E. Barausse *et al.*, *Gen. Relativ. Gravit.* **52**, 81 (2020).
- [18] E. Maggio, L. Buoninfante, A. Mazumdar, and P. Pani, *Phys. Rev. D* **102**, 064053 (2020).
- [19] S. Bhagwat, C. Pacilio, E. Barausse, and P. Pani, *Phys. Rev. D* **105**, 124063 (2022).
- [20] M. Corman, A. Ghosh, C. Escamilla-Rivera, M. A. Hendry, S. Marsat, and N. Tamanini, *Phys. Rev. D* **105**, 064061 (2022).
- [21] A. Toubiana, L. Pompili, A. Buonanno, J. R. Gair, and M. L. Katz, *Phys. Rev. D* **109**, 104019 (2024).
- [22] C. Pitte, Q. Baghi, M. Besançon, and A. Petiteau, *Phys. Rev. D* **110**, 104003 (2024).
- [23] A. Mangiagli, C. Caprini, M. Volonteri, S. Marsat, S. Vergani, N. Tamanini, and H. Inchauspé, *Phys. Rev. D* **106**, 103017 (2022).
- [24] C. A. Dong-Páez, M. Volonteri, R. S. Beckmann, Y. Dubois, A. Mangiagli, M. Trebitsch, S. Vergani, and N. Webb, *Astron. Astrophys.* **676**, A2 (2023).
- [25] D. Izquierdo-Villalba, A. Sesana, M. Colpi, D. Spinoso, M. Bonetti, S. Bonoli, and R. Valiante, *Astron. Astrophys.* **686**, A183 (2024).
- [26] C. Caprini and N. Tamanini, *J. Cosmol. Astropart. Phys.* **10** (2016) 006.
- [27] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, *J. Cosmol. Astropart. Phys.* **04** (2016) 002.
- [28] A. Mangiagli, C. Caprini, S. Marsat, L. Speri, R. R. Caldwell, and N. Tamanini, *Phys. Rev. D* **111**, 083043 (2025).
- [29] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
- [30] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [31] K. Somiya (KAGRA Collaboration), *Classical Quantum Gravity* **29**, 124007 (2012).
- [32] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto (KAGRA Collaboration), *Phys. Rev. D* **88**, 043007 (2013).
- [33] G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart, *Astron. Astrophys.* **375**, 890 (2001).
- [34] A. J. Ruiter, K. Belczynski, M. Benacquista, S. L. Larson, and G. Williams, *Astrophys. J.* **717**, 1006 (2010).
- [35] V. Korol, E. M. Rossi, P. J. Groot, G. Nelemans, S. Toonen, and A. G. A. Brown, *Mon. Not. R. Astron. Soc.* **470**, 1894 (2017).
- [36] A. Lamberts, S. Blunt, T. B. Littenberg, S. Garrison-Kimmel, T. Kupfer, and R. E. Sanderson, *Mon. Not. R. Astron. Soc.* **490**, 5888 (2019).
- [37] V. Korol, N. Hallakoun, S. Toonen, and N. Karnesis, *Mon. Not. R. Astron. Soc.* **511**, 5936 (2022).
- [38] A. Toubiana, N. Karnesis, A. Lamberts, and M. C. Miller, *Astron. Astrophys.* **692**, A165 (2024).
- [39] J. R. Gair, C. Tang, and M. Volonteri, *Phys. Rev. D* **81**, 104014 (2010).
- [40] P. Amaro-Seoane, *Living Rev. Relativity* **21**, 4 (2018).
- [41] 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, *Phys. Rev. D* **95**, 103012 (2017).
- [42] Z. Pan and H. Yang, *Phys. Rev. D* **103**, 103018 (2021).
- [43] P. A. Seoane, Y. Lin, and K. Tzanavaris, *Phys. Rev. D* **110**, 064011 (2024).
- [44] M. Bonetti and A. Sesana, *Phys. Rev. D* **102**, 103023 (2020).
- [45] F. Pozzoli, S. Babak, A. Sesana, M. Bonetti, and N. Karnesis, *Phys. Rev. D* **108**, 103039 (2023); **110**, 049903(E) (2024).
- [46] A. Sesana, *Phys. Rev. Lett.* **116**, 231102 (2016).
- [47] K. W. K. Wong, E. D. Kovetz, C. Cutler, and E. Berti, *Phys. Rev. Lett.* **121**, 251102 (2018).
- [48] D. Gerosa, S. Ma, K. W. K. Wong, E. Berti, R. O’Shaughnessy, Y. Chen, and K. Belczynski, *Phys. Rev. D* **99**, 103004 (2019).
- [49] A. Toubiana, S. Marsat, S. Babak, J. Baker, and T. Dal Canton, *Phys. Rev. D* **102**, 124037 (2020).
- [50] R. Buscicchio, A. Klein, E. Roebber, C. J. Moore, D. Gerosa, E. Finch, and A. Vecchio, *Phys. Rev. D* **104**, 044065 (2021).
- [51] A. Toubiana, S. Babak, S. Marsat, and S. Ossokine, *Phys. Rev. D* **106**, 104034 (2022).
- [52] R. Buscicchio, J. Torrado, C. Caprini, G. Nardini, N. Karnesis, M. Pieroni, and A. Sesana, *J. Cosmol. Astropart. Phys.* **01** (2025) 084.
- [53] P. Auclair *et al.* (LISA Cosmology Working Group), *Living Rev. Relativity* **26**, 5 (2023).
- [54] M. Muratore, O. Hartwig, D. Vetrugno, S. Vitale, and W. J. Weber, *Phys. Rev. D* **107**, 082004 (2023).
- [55] N. J. Cornish and J. Crowder, *Phys. Rev. D* **72**, 043005 (2005).
- [56] T. B. Littenberg and N. J. Cornish, *Phys. Rev. D* **107**, 063004 (2023).
- [57] M. L. Katz, N. Karnesis, N. Korsakova, J. R. Gair, and N. Stergioulas, *Phys. Rev. D* **111**, 024060 (2025).
- [58] S. H. Strub, L. Ferraioli, C. Schmelzbach, S. C. Stähler, and D. Giardini, *Phys. Rev. D* **110**, 024005 (2024).
- [59] S. Deng, S. Babak, M. Le Jeune, S. Marsat, E. Plagnol, and A. Sartirana, *Phys. Rev. D* **111**, 103014 (2025).
- [60] M. Armano *et al.* (LISA Pathfinder Collaboration), *Phys. Rev. D* **110**, 042004 (2024).
- [61] M. Armano *et al.*, *Phys. Rev. D* **106**, 062001 (2022).
- [62] M. Muratore, Time delay interferometry for LISA science and instrument characterization, Ph.D. thesis, Trento University, 2021.
- [63] M. Zevin *et al.*, *Classical Quantum Gravity* **34**, 064003 (2017).

- [64] M. Armano *et al.* (LISA Pathfinder Collaboration), *Phys. Rev. D* **106**, 062001 (2022).
- [65] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **119**, 161101 (2017).
- [66] C. Pankow *et al.*, *Phys. Rev. D* **98**, 084016 (2018).
- [67] N. Karnesis, M. L. Katz, N. Korsakova, J. R. Gair, and N. Stergioulas, *Mon. Not. R. Astron. Soc.* **526**, 4814 (2023).
- [68] M. Katz and J. Roberts, MIKEKATZ04/BBHx: New release! (2023).
- [69] LISA LDC Team, LISA data challenges: Challenge 2b (accessed: January 22, 2025), <https://zenodo.org/records/7436568>.
- [70] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forzeza, C. Kalaghatgi, F. Ohme, and F. Pannarale, *Phys. Rev. Lett.* **120**, 161102 (2018).
- [71] M. Katz, N. Karnesis, and N. Korsakova, MIKEKATZ04/Eryn: First full release (2023).
- [72] M. Armano *et al.*, *Phys. Rev. Lett.* **120**, 061101 (2018).
- [73] M. Armano *et al.*, *Phys. Rev. D* **110**, 042004 (2024).
- [74] L. Sala, Residual test mass acceleration in LISA Pathfinder: In-depth statistical analysis and physical sources, Ph.D. thesis, Trento University, 2023.
- [75] E. Castelli, LISA Pathfinder noise performance results: Disturbances in the sub-mHz frequency band and projection to LISA, Ph.D. thesis, Trento University, 2020.
- [76] Q. Baghi, N. Korsakova, J. Slutsky, E. Castelli, N. Karnesis, and J.-B. Bayle, *Phys. Rev. D* **105**, 042002 (2022).
- [77] D. A. Shaddock, B. Ware, R. E. Spero, and M. Vallisneri, *Phys. Rev. D* **70**, 081101 (2004).
- [78] J. W. Armstrong, F. B. Estabrook, and M. Tinto, *Astrophys. J.* **527**, 814 (1999).
- [79] M. Muratore, D. Vetrugno, and S. Vitale, *Classical Quantum Gravity* **37**, 185019 (2020).
- [80] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, *Phys. Rev. D* **66**, 122002 (2002).
- [81] D. A. Shaddock, M. Tinto, F. B. Estabrook, and J. W. Armstrong, *Phys. Rev. D* **68**, 061303 (2003).
- [82] M. Vallisneri, *Phys. Rev. D* **72**, 042003 (2005).
- [83] O. Hartwig and M. Muratore, *Phys. Rev. D* **105**, 062006 (2022).
- [84] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, *Phys. Rev. D* **66**, 122002 (2002).
- [85] G. Heinzl, J. J. Esteban, S. Barke, M. Otto, Y. Wang, A. F. Garcia, and K. Danzmann, *Classical Quantum Gravity* **28**, 094008 (2011).
- [86] O. Hartwig, Instrumental modelling and noise reduction algorithms for the Laser Interferometer Space Antenna, Ph.D. thesis, Leibniz University, Hannover, 2021.
- [87] J. Bergé, R. Massey, Q. Baghi, and P. Touboul, *Mon. Not. R. Astron. Soc.* **486**, 544 (2019).
- [88] Wikipedia Contributors, Laguerre polynomials (accessed: January 22, 2025), https://en.wikipedia.org/wiki/Laguerre_polynomials.
- [89] C. Pitte, Q. Baghi, S. Marsat, M. Besançon, and A. Petiteau, *Phys. Rev. D* **108**, 044053 (2023).
- [90] S. Marsat and J. G. Baker, [arXiv:1806.10734](https://arxiv.org/abs/1806.10734).
- [91] S. Marsat, J. G. Baker, and T. D. Canton, *Phys. Rev. D* **103**, 083011 (2021).
- [92] O. Hartwig and J.-B. Bayle, *Phys. Rev. D* **103**, 123027 (2021).
- [93] M.-S. Hartig, S. Schuster, and G. Wanner, *J. Opt.* **24**, 065601 (2022).
- [94] M. Chwalla, K. Danzmann, M. D. Álvarez, J. E. Delgado, G. Fernández Barranco, E. Fitzsimons, O. Gerberding, G. Heinzl, C. Killow, M. Lieser, M. Perreur-Lloyd, D. Robertson, J. Rohr, S. Schuster, T. Schwarze, M. Tröbs, G. Wanner, and H. Ward, *Phys. Rev. Appl.* **14**, 014030 (2020).
- [95] D. Quang Nam, J. Martino, Y. Lemièrre, A. Petiteau, J.-B. Bayle, O. Hartwig, and M. Staab, *Phys. Rev. D* **108**, 082004 (2023).
- [96] D. Quang Nam, J. Martino, Y. Lemièrre, A. Petiteau, J.-B. Bayle, O. Hartwig, and M. Staab, *Phys. Rev. D* **108**, 082004 (2023).
- [97] LISA Data Challenge Working Group, LISA data challenge software, Ver. 1.2.4 (2022), <https://zenodo.org/records/7332221>.
- [98] M. Tinto and S. Dhurandhar, *Living Rev. Relativity* **24**, 1 (2021).
- [99] S. Babak, M. Hewitson, and A. Petiteau, [arXiv:2108.01167](https://arxiv.org/abs/2108.01167).
- [100] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, 2nd ed. (Oxford University Press, New York, 2006).
- [101] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [102] W. K. Hastings, *Biometrika* **57**, 97 (1970).
- [103] C. Hitchcock, *Contemporary Debates in Philosophy of Science* (Blackwell Publishing, 2003).
- [104] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (Chapman & Hall/CRC, London, 1995).
- [105] P. J. Green, *Biometrika* **82**, 711 (1995).
- [106] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **57**, 2607 (1986).
- [107] K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604 (1996).
- [108] W. D. Vousden, W. M. Farr, and I. Mandel, *Mon. Not. R. Astron. Soc.* **455**, 1919 (2015).
- [109] J. Goodman and J. Weare, *Commun. Appl. Math. Comput. Sci.* **5**, 65 (2010).
- [110] <https://scikit-learn.org/stable/modules/mixture.html>
- [111] https://scikit-learn.org/stable/modules/generated/sklearn_cluster.KMeans.html
- [112] M. L. Katz, *Phys. Rev. D* **105**, 044055 (2022).
- [113] M. Muratore, J. Gair, and L. Speri, *Phys. Rev. D* **109**, 042001 (2024).
- [114] T. B. Littenberg, N. J. Cornish, K. Lackeos, and T. Robson, *Phys. Rev. D* **101**, 123021 (2020).
- [115] A. Spadaro, R. Busicchio, D. Vetrugno, A. Klein, D. Gerosa, S. Vitale, R. Dolesi, W. J. Weber, and M. Colpi, *Phys. Rev. D* **108**, 123029 (2023).
- [116] J. R. Gair *et al.*, *Astron. J.* **166**, 22 (2023).
- [117] J. Y. L. Kwok, R. K. L. Lo, A. J. Weinstein, and T. G. F. Li, *Phys. Rev. D* **105**, 024066 (2022).
- [118] A. Gupta *et al.*, *SciPost Phys. Community Rep.* **5** (2025).
- [119] LISA-LCST-XXX-TN-001, [10.5281/zenodo.7436568](https://zenodo.org/records/10.5281/zenodo.7436568) (2022).
- [120] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forzeza, and A. Bohé, *Phys. Rev. D* **93**, 044006 (2016).

- [121] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044007 (2016).
- [122] A. Dhani, S. Völkel, A. Buonanno, H. Estelles, J. Gair, H. P. Pfeiffer, L. Pompili, and A. Toubiana, [arXiv:2404.05811](https://arxiv.org/abs/2404.05811).
- [123] O. Hartwig and M. Muratore, *Phys. Rev. D* **105**, 062006 (2022).
- [124] M. Muratore, D. Vetrugno, S. Vitale, and O. Hartwig, *Phys. Rev. D* **105**, 023009 (2022).
- [125] M. R. Adams and N. J. Cornish, *Phys. Rev. D* **82**, 022002 (2010).
- [126] LISA-LCST-SGS-MAN-001, [10.5281/zenodo.7132178](https://zenodo.org/records/7132178) (2020).
- [127] J.-B. Bayle, O. Hartwig, A. Petiteau, and M. Lilley, Lisanode (2022), <https://zenodo.org/records/6461078>.
- [128] P. Maturana-Russel, R. Meyer, J. Veitch, and N. Christensen, *Phys. Rev. D* **99**, 084006 (2019).
- [129] G. C. Davies, I. Harry, M. J. Williams, D. Bandopadhyay, L. Barack, J.-B. Bayle, C. Hoy, A. Klein, H. Middleton, C. J. Moore, L. Nuttall, G. Pratten, A. Vecchio, and G. Woan, *Phys. Rev. D* **111**, 043045 (2025).
- [130] N. J. Cornish, *Phys. Rev. D* **105**, 044007 (2022).
- [131] <https://github.com/martinaAEI/artifacts>.
- [132] M. Staab, J.-B. Bayle, and O. Hartwig, PYTDI (2023).
- [133] G. Heinzel, A. Rüdiger, and R. Schilling, Spectrum and spectral density estimation by the Discrete Fourier transform (DFT), including a comprehensive list of window functions and some new at-top windows (Max Plank Institute, 2002), Vol. **12**, <https://hdl.handle.net/11858/00-001M-0000-0013-557A-5>.
- [134] W. Xie, P. O. Lewis, Y. Fan, L. Kuo, and M.-H. Chen, *Syst. Biol.* **60**, 150 (2011).