

Global time-frequency search for stellar-mass binary black holes in LISA

Diganta Bandopadhyay^{1,2,*}, Christian E. A. Chapman-Bird^{2,†}, and Alberto Vecchio^{2,3,‡}

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

²*Institute for Gravitational Wave Astronomy and School of Physics and Astronomy,
University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom*

³*Center for Computational Astrophysics, Flatiron Institute,
162 5th Avenue, New York, New York 10010, USA*



(Received 22 October 2025; accepted 6 August 2026; published 8 September 2026)

We present a complete pipeline for detecting and identifying gravitational waves (GWs) produced by the inspiral of stellar-mass binary black holes in data from the Laser Interferometer Space Antenna (LISA). The analysis framework relies on an efficient time-frequency implementation of an adaptive semicoherent detection statistic, which we show to be robust against nonstationary noise and the presence of gaps of varying duration and cadence. The search is able to detect signals with coherent signal-to-noise ratios as low as ≈ 11 – 14 over the full parameter space of spin-aligned binaries (using nonspinning templates) with orbital eccentricity ≤ 0.01 when deployed on the 2-year-long LISA Data Challenge *Yorsh*. The search can be run within a day using ≈ 40 GPUs. The techniques presented here have wider applications in GW astronomy, in particular the search for extreme-mass-ratio inspirals in LISA data.

DOI: [10.1103/339z-w985](https://doi.org/10.1103/339z-w985)

I. INTRODUCTION

The challenge of performing a global (i.e., full parameter space) search for gravitational waves (GWs) from binary systems is ubiquitous in GW astronomy. This affects ground-based observations of neutron stars and black holes (see, e.g., Ref. [1] and references therein) as the LIGO [2], Virgo [3] and KAGRA [4] detectors (and more ambitious future generation observatories [5,6]) improve their low-frequency sensitivity and the search domain is extended to subsolar mass objects [7], orbital eccentricity [8] and arbitrary black hole spins [9]. It is also an unsolved problem for the Laser Interferometer Space Antenna (LISA) [10,11] in the context of identifying stellar-origin (~ 1 – $100 M_\odot$) black hole binaries (SOBHBs) ≈ 1 – 100 yr from coalescence, which would provide unique insight into their astrophysical formation pathways [11–15], and extreme-mass-ratio inspirals (EMRIs) of neutron stars and stellar-mass black holes orbiting massive black holes

($\sim 10^5$ – $10^7 M_\odot$), which represent new probes of the formation and evolution of massive black holes at the center of galaxies, their environment [16], and unique laboratories for tests of general relativity [17].

Currently, a computationally viable search strategy is unknown for SOBHBs and EMRIs. The parameter space volume covered by the likelihood function of typical signals is a minuscule fraction of the astrophysical prior, from $\sim \mathcal{O}(10^{-32})$ for SOBHBs to $\mathcal{O}(10^{-46})$ for EMRIs (see Fig. 1 in Ref. [14]). Their complex signals, originating from spinning black holes in binaries with eccentric orbits, have power distributed across (potentially) many modes, $\sim 10^3$, in the case of EMRIs, and must be integrated over $\sim 10^5$ wave cycles to accumulate sufficient signal-to-noise ratio (SNR) to be detectable. These factors make standard match-filtering searches unfeasible [14,18,19]. Alternative methods, including excess-power [18,20–22], machine-learning [23,24] or phenomenological fits [25,26] are being explored, but a viable solution has yet to be found.

The problem of identifying and characterizing SOBHBs in LISA data is under active research [27–34]. These sources offer key insights into both astrophysics and fundamental physics [27,35,36], complementing the late-inspiral observations made by ground-based detectors [37–45]. Their weak-field nature and comparable mass ratios further make them an ideal test-bed for developing techniques that may ultimately be applied to the EMRI search. In addition, SOBHBs are expected to be rare [46],

*Contact author: diganta.bandopadhyay@aei.mpg.de

†Contact author: c.chapman-bird@bham.ac.uk

‡Contact author: a.vecchio@bham.ac.uk

§These authors contributed equally to this work.

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.

TABLE I. Computational costs—timed on NVIDIA A100 GPUs—of the proposed pipeline. The detection statistic $\Upsilon_{N_{\text{seg}}}$ is defined in Eq. (5) and (7). Each search tile is assigned to (and timed on) a single GPU. The search of the *Yorsh* dataset presented here used 4 GPUs. We also provide comparison with the pipeline described in Ref. [31] (BM25). Only the cost of evaluating $\Upsilon_{N_{\text{seg}}}$ is directly comparable; the BM25 total wall time is therefore extrapolated assuming 4 GPUs were used.

	This work	BM25
Cost per Υ evaluation [μs]	1–10	10^3 – 10^4
No. of Υ evaluations per search tile	10^8	
Wall time per tile [h]	0.1–0.5	
Total number of search tiles	2680	
Full search wall time on 4 GPUs [d]	11	~ 3000

which pushes demands on search strategies to be nearly optimal in terms of sensitivity. Furthermore, their radiation leaves the LISA band $\sim$ weeks before the black holes merge, which places severe time-constraints on any global search implementation to allow telescopes to point *in advance* and stare at a merger unfolding [15,47,48], while being also observed in GWs by ground-based detectors.

In this work, we present an end-to-end global search method based on an adaptive semicoherent statistic [30,31]. We demonstrate that this method is able to detect SOBHB signals in year-long LISA (mock) datasets down to SNR ≈ 11 over the full astrophysical parameter space of mildly eccentric, spin-aligned SOBHBs (using nonspinning templates). At the core of our search is a hardware-accelerated time-frequency domain waveform representation—a natural basis for broad-band, slowly chirping signals [28,49]—that reduces computational cost by $\sim 10^3$ with respect to frequency-domain methods. As shown in Table I, this pipeline searches the entirety of the two-year LISA Data Challenge (LDC) *Yorsh* dataset in $\lesssim 12$ days on 4 NVIDIA A100 graphics processing units (GPUs). Due to its computational speed-up of $\sim 250\times$ over the previous state-of-the-art [30,31] and it being massively parallelizable, it can achieve sub-1 day latency on ≈ 40 GPUs. We also demonstrate that our approach naturally deals with (and is robust to) data gaps and nonstationary noise sources, which introduce additional challenges in searching for long-lived signals. It is therefore readily applicable to realistic LISA data.

This work presents a practical solution to the long-standing challenge of SOBHB search, establishes a foundation for tackling the more complex problem of EMRI detection, and several of its underpinning techniques have broader applicability to LIGO-Virgo-KAGRA searches for compact binaries over increasingly wider parameter space. In this study, we use geometric units ($c = G = 1$).

A high level overview of the end-to-end search pipeline described here is illustrated in Fig. 2 of Ref. [30].

II. METHODS

In this section, we introduce the detection statistic, its efficient time-frequency representation and the GPU-optimized acceleration of its implementation that we adopt to build an end-to-end search for SOBHBs.

Our approach exploits a hierarchy of timescales, $T_{\text{obs}} \geq T_{\text{seg}} \geq \Delta T$, that we introduce here: (i) T_{obs} is the *total* duration of the dataset that for LISA is several years (the nominal mission lifetime is 4.5 yr with possible extension to 10 yr [11]); we set the (arbitrary) reference time $t_0 = 0$ at the start of observation, so that the time of coalescence of a SOBHB, t_c , corresponds to their remaining lifetime; (ii) T_{seg} is the duration of the non overlapping *time segments* into which we divide the dataset, and defines the coherence length of our analysis; the total number of segments (which varies as the search progresses) is $N_{\text{seg}} \equiv \tau/T_{\text{seg}} = 100(\tau/10^8 \text{ s})(10^6 \text{ s}/T_{\text{seg}})$ for $\tau = \min(T_{\text{obs}}, t_c)$; (iii) ΔT is the length of the data *tranche* over which an efficient time-frequency approximation of the inner product between the data and the signal’s templates is computed, and is set in advance ensuring that second-order contributions to the Taylor expansion of the signal phase $\approx 0.2(\mathcal{M}_c/30M_\odot)^{10/3}(f/10 \text{ mHz})^{19/3}(\Delta T/10^4 \text{ sec})^3$ can be neglected; here $\mathcal{M}_c$ is the binary’s chirp mass and the total number of tranches in a segment is $M_{\text{tr}} \equiv T_{\text{seg}}/\Delta T$.

A. Time-frequency search statistic

In this section we detail the procedure for constructing a time-frequency formulation of the detection statistic. Crucially, this relies on a sparse representation of the waveform which is derived here. We begin by defining the statistic in the frequency domain and then describe its efficient computation in the time-frequency domain.

Given two real time series $a(t)$ and $b(t)$, with Fourier transform $\tilde{a}(f)$ and $\tilde{b}(f)$, respectively, we define the usual noise-weighted inner product

$$\langle a|b \rangle = 4 \text{Re} \int_0^\infty \frac{\tilde{a}^*(f)\tilde{b}(f)}{S(f)} df, \quad (1)$$

where $S(f)$ is the one-sided noise power spectral density (PSD). We also introduce

$$(a|b) = \int_0^\infty \frac{\tilde{a}^*(f)\tilde{b}(f)}{S(f)} df, \quad (2)$$

which, with slight abuse of terminology, we call the complex inner product, such that $\langle a|b \rangle = 4 \text{Re}\{(a|b)\}$.

We make the standard assumption that the three time-delay interferometry (TDI) channels, labeled by α , are pseudo noise-orthogonal [50] (although this can be relaxed at minor additional cost). From Eq. (1), the coherent inner product between the data, $d = \{d_\alpha\}$, and a GW signal template characterized by parameters θ , $y(\theta) = \{y_\alpha(\theta)\}$, is

$$\begin{aligned}\langle d|y(\theta)\rangle &= \sum_{\alpha=1}^3 \langle d_\alpha|y_\alpha(\theta)\rangle \\ &= \sum_{\alpha=1}^3 4 \operatorname{Re} \left\{ \int_0^\infty \frac{\tilde{d}_\alpha^*(f) \tilde{y}_\alpha(\theta; f)}{S_\alpha(f)} df \right\},\end{aligned}\quad (3)$$

where the noise weighting function for $\langle d_\alpha|y_\alpha(\theta)\rangle$ is the one-sided noise PSD in the relevant TDI channel, $S_\alpha(f)$.

One of the parameters that define the signal in the data is an overall constant phase, ϕ , at some (arbitrary) reference time. Writing $\theta = \{\xi, \phi\}$, such that ξ is the vector of parameters *excluding* an overall constant phase, and substituting Eqs. (1) and (2) into Eq. (3), one can maximize the inner product over this phase as

$$\begin{aligned}\mathcal{O}(\xi) &= \max_{\phi} \langle d|y(\xi) e^{i\phi}\rangle \\ &= \max_{\phi} \sum_{\alpha=1}^3 \langle d_\alpha|y_\alpha(\xi) e^{i\phi}\rangle \\ &= 4 \left| \sum_{\alpha=1}^3 (d_\alpha|y_\alpha(\xi)) \right|.\end{aligned}\quad (4)$$

The quantity $\mathcal{O}(\xi)$ is also known as the overlap of the template with the data, and the last line in Eq. (4) shows that phase maximization can be performed analytically and efficiently in the Fourier domain.

In this work, we employ the semi-coherent statistic introduced in [29,30,51]. Segmenting the data such that the n th segment covers the time interval $[t_n, t_{n+1}]$, where $t_n = nT_{\text{seg}}$, the detection statistic is

$$\Upsilon_{N_{\text{seg}}}(\xi) = \sum_{n=0}^{N_{\text{seg}}-1} \frac{[\max_{\phi_n} \langle d_n|y_n(\xi) e^{i\phi_n}\rangle]^2}{\langle y_n(\xi)|y_n(\xi)\rangle}, \quad (5)$$

where we have defined the inner product over the segment (for all TDI channels) as

$$\langle d_n|y_n e^{i\phi_n}\rangle = 4 \operatorname{Re} \left\{ e^{i\phi_n} \sum_{\alpha=1}^3 \sum_{k=0}^{N_f-1} \frac{\tilde{d}_{\alpha,n}^*[k] \tilde{y}_{\alpha,n}[k]}{S_{\alpha,n}[k]} \Delta f \right\}, \quad (6)$$

and analogously for $\langle y_n|y_n\rangle$. Here $\tilde{x}$ denotes the Fourier transform of a time series x , $\Delta f = 1/T_{\text{seg}}$, $N_f = T_{\text{seg}}/\Delta t$, Δt is the sampling cadence, and the index k corresponds to the discrete frequency $f_k = k\Delta f$. $S_{\alpha,n}$ is the one-sided noise power spectral density (PSD) in the segment.

In the definition of the detection statistic we have explicitly split the vector of unknown parameters that characterise a signal into $\theta = \{\xi, \phi\}$, where ξ is the vector of parameters *excluding* a constant phase offset ϕ . Note that we are also maximizing over an amplitude parameter implicitly within each segment in the definition of Eq. (5). For a given N_{seg} , the unknown parameters are

$(\xi, \{\phi_n, A_n; n = 1, \dots, N_{\text{seg}}\})$, and therefore the effective dimensionality of the parameter space is $\dim(\xi) + 2N_{\text{seg}}$, with maximization over $2N_{\text{seg}}$ parameters. As the search progresses, N_{seg} is gradually reduced; the final stage, $N_{\text{seg}} = 1$, is fully coherent and the search statistic is $\Upsilon_1 \equiv \max_{\phi}(\rho^2)$, where ρ is the matched filter SNR, this hierarchical process reduces the false-alarm probability of the search [30].

The key to our search is efficiently computing Eq. (6), which we achieve with a hardware-accelerated framework in the time-frequency domain under the assumption of a slowly evolving GW frequency. We divide each segment in tranches, so that the m th tranche covers the time interval $[t_{nm}, t_{nm+1}]$ with $t_{nm} = nT_{\text{seg}} + m\Delta T$. We can now write the inner product Eq. (6) as

$$\begin{aligned}\langle d_n|y_n e^{i\phi_n}\rangle &= 4 \operatorname{Re} \left\{ e^{i\phi_n} \sum_{\alpha=1}^3 \sum_{m=0}^{M_{\text{tr}}-1} \mathcal{T}_{\alpha,nm} \right. \\ &\quad \times \left. \sum_{k=1}^{N_{\mathcal{F}}} \frac{\tilde{d}_{\alpha,nm}^*[k] \tilde{h}_{\alpha,nm}[k]}{S_{\alpha,nm}[k]} \delta f \right\},\end{aligned}\quad (7)$$

with $\delta f = 1/\Delta T$. Equation (7) is significantly more efficient to compute than Eq. (6) because (i) the sum over frequency is compact and can be truncated to $N_{\mathcal{F}} \ll \Delta T/\Delta t$ bins, centred on $f_{nm} \equiv f(t_{nm})$; (ii) the frequency-domain waveform $\tilde{h}_{nm}(f)$ is expressed analytically in terms of Fresnel integrals; (iii) the TDI transfer functions $\mathcal{T}_{\alpha,nm} \equiv \mathcal{T}_\alpha(f_{nm})$ are complex multiplicative factors, computed once per tranche for all frequency bins. As shown in Table I, under assumptions suitable for SOBHBs (and GPU acceleration), this approach is orders of magnitude more efficient than computation of Eq. (6) in the frequency domain, rendering a global SOBHB search feasible.

The detection statistic, Eq. (5), is the sum of (suitably normalized) overlaps over the time segments into which we partition the dataset. We divide the dataset of duration T_{obs} in non overlapping time segments of length T_{seg} , where the total number of segments is $N_{\text{seg}} = T_{\text{obs}}/T_{\text{seg}}$. Using Eq. (4), the detection statistic (5) can be written as

$$\Upsilon_{N_{\text{seg}}}(\xi) = \sum_{n=0}^{N_{\text{seg}}-1} \frac{\mathcal{O}_n^2(\xi)}{\sum_{\alpha=1}^3 \langle y_{\alpha,n}(\xi)|y_{\alpha,n}(\xi)\rangle}, \quad (8)$$

where we have defined the overlap over the segment as

$$\begin{aligned}\mathcal{O}_n(\xi) &= \max_{\phi_n} \sum_{\alpha=1}^3 \langle d_{\alpha,n}|y_{\alpha,n}(\xi) e^{i\phi_n}\rangle, \\ &= 4 \left| \sum_{\alpha=1}^3 (d_{\alpha,n}|y_{\alpha,n}(\xi)) \right|.\end{aligned}\quad (9)$$

In the previous equation, $\langle d_{\alpha,n}|y_{\alpha,n}(\xi)e^{i\phi_n}\rangle$ is given by Eq. (6), and in analogy with Eq. (2) we have defined the segment-level complex inner product

$$\begin{aligned} (d_{\alpha,n}|y_{\alpha,n}(\xi)) &= \int_0^\infty \frac{\tilde{d}_{\alpha,n}^*(f)\tilde{y}_{\alpha,n}(\xi;f)}{S_{\alpha,n}(f)} df, \\ &= \sum_{k=0}^{N_f-1} \frac{\tilde{d}_{\alpha,n}^*[k]\tilde{y}_{\alpha,n}(\xi)[k]}{S_{\alpha,n}[k]} \Delta f, \end{aligned} \quad (10)$$

to emphasise the process by which one maximises over the N_{seg} phase parameters. In Eqs. (8), (9), and (10) we have defined a time series $x_{\alpha,n}$ as the output of the TDI channel α over the time segment n . The data are discretely sampled at times $t_j = j\Delta t$, where Δt is the sampling cadence. We denote the discrete time series as $x_{\alpha,n}[j] = x(nT_{\text{seg}} + j\Delta t)$. The discrete Fourier components of $x_{\alpha,n}$ at frequency $f_k = k\Delta f$ are denoted by $\tilde{x}_{\alpha,n}[k]$, where $\Delta f = 1/T_{\text{seg}}$ and $N_f = T_{\text{seg}}/\Delta t$; $S_{\alpha,n}[k]$ is the one-sided noise spectral density of the data at frequency f_k .

B. Time-frequency waveform representation

It is essential to compute Eq. (10) efficiently, which we do by exploiting the slow phase evolution of GWs from SOBHBs over T_{obs} . Assuming one waveform mode is dominant, the complex GW strain can be written as

$$h(t) = \mathcal{A}(t)e^{i\varphi(t)}. \quad (11)$$

We divide each segment into tranches of duration ΔT . For the n th segment, the m th tranche covers the time interval $[t_{nm}, t_{nm+1}]$, where $t_{nm} = nT_{\text{seg}} + m\Delta T$. Within each tranche, we Taylor-expand the GW phase to second order:

$$\begin{aligned} \varphi(t) &= \varphi_{nm} + 2\pi f_{nm}(t - t_{nm}) + \pi \dot{f}_{nm}(t - t_{nm})^2 \\ &\quad + \mathcal{O}((t - t_{nm})^3), \end{aligned} \quad (12)$$

where

$$\varphi_{nm} \equiv \varphi(t_{nm}), \quad (13)$$

$$f_{nm} \equiv f(t_{nm}) = \frac{1}{2\pi} \left[\frac{d\varphi}{dt} \right]_{t=t_{nm}}, \quad (14)$$

and

$$\dot{f}_{nm} \equiv \frac{1}{2\pi} \left[\frac{d^2\varphi}{dt^2} \right]_{t=t_{nm}}. \quad (15)$$

The Fourier transform of Eq. (11) over the tranche can now be approximated as

$$\tilde{h}_{nm}(f) = \frac{\mathcal{A}_{nm}}{\sqrt{2\dot{f}_{nm}}} e^{i(\varphi_{nm} - \pi \dot{f}_{nm} \zeta_{nm}^2(f))} \mathcal{F}_{nm}(f), \quad (16)$$

where

$$\zeta_{nm}(f) = \frac{f_{nm} - f}{\dot{f}_{nm}}, \quad (17)$$

$$\mathcal{A}_{nm} = \mathcal{A}(t_{nm}). \quad (18)$$

In Eq. (16) we have defined the Fresnel kernel function

$$\mathcal{F}_{nm}(f) = \Delta C_{nm}(f) + i\Delta S_{nm}(f), \quad (19)$$

with

$$\Delta C_{nm}(f) = C[v_{nm}(t_{nm+1}, f)] - C[v_{nm}(t_{nm}, f)], \quad (20)$$

$$\Delta S_{nm}(f) = S[v_{nm}(t_{nm+1}, f)] - S[v_{nm}(t_{nm}, f)], \quad (21)$$

$$v_{nm}(t, f) = \sqrt{2\dot{f}_{nm}}[(t - t_{nm}) + \zeta_{nm}(f)], \quad (22)$$

and $S(z) = \int_0^z \sin(\pi x^2/2) dx$ and $C(z) = \int_0^z \cos(\pi x^2/2) dx$ are the Fresnel integrals [52].

A crucial aspect for the efficient computation of Eq. (10) is the behavior of the kernel function (19) as a function of frequency: for the values of $\dot{f}_{nm}$ typical of SOBHBs in the relevant frequency band ~ 1 –100 mHz, the waveform and thus the Fresnel-kernel in Eq. (19) [and hence Eq. (16)] peak sharply around f_{nm} and fall off rapidly in frequency (see extended discussion on the compactness of this kernel in Appendix). Additionally, as $v_{nm}(t, f) \sim \delta f / \sqrt{\dot{f}}$ is typically large ($\sqrt{\dot{f}} \ll \delta f$ for much of the inspiral), we specifically employ asymptotic forms of the Fresnel integrals to further reduce computational costs [53]. For argument $|z| > 6$, these take the form

$$S(z) = \frac{1}{2} - \frac{1}{\pi z} \cos\left(\frac{\pi z^2}{2}\right) \quad (23)$$

and

$$C(z) = \frac{1}{2} + \frac{1}{\pi z} \sin\left(\frac{\pi z^2}{2}\right), \quad (24)$$

agreeing to better than one part in 10^4 for all $|z| > 6$. Note that $S(-z) = -S(z)$ and $C(-z) = -C(z)$. For $|z| \leq 6$, we use the modified forms

$$S(z) = \frac{1}{2} - f(z) \cos\left(\frac{\pi z^2}{2}\right) - g(z) \sin\left(\frac{\pi z^2}{2}\right) \quad (25)$$

and

$$C(z) = \frac{1}{2} + f(z) \sin\left(\frac{\pi z^2}{2}\right) - g(z) \cos\left(\frac{\pi z^2}{2}\right), \quad (26)$$

with

$$f(z) = \frac{1 + 0.926z}{2 + 1.792z + 3.104z^2} \quad (27)$$

and

$$g(z) = \frac{1}{2 + 4.142z + 3.492z^2 + 6.670z^3}. \quad (28)$$

The polynomial coefficients are sourced from Ref. [54]. This extends the agreement to one part in 10^4 over the remainder of the domain in z .

Finally, as the last step in constructing the signal model, we write the TDI observable as

$$y_{\alpha,nm}(f) = \mathcal{T}_\alpha(f_{nm}) \tilde{h}_{nm}(f). \quad (29)$$

where $\mathcal{T}_\alpha$ are TDI transfer functions, and we have explicitly reintroduced the index α to label the different TDI variables. Note that the transfer function is only computed for the “central” frequency f_{nm} within each tranche and applied to the whole waveform $\tilde{h}(f)$, making use of the relative compactness of $\tilde{h}(f)$ compared to variations in $\mathcal{T}_\alpha(f)$. The waveform is sharply peaked around its central frequency over hour-long timescales, meanwhile the instrument response function varies relatively smoothly and has support at a wide range of frequencies, see for example Fig. 8 in Ref. [55].

Equation. (10) can now be computed by combining the analogous quantities computed at the tranche level:

$$(d_{\alpha,n}|y_{\alpha,n}(\xi)) = \sum_{m=0}^{M_{\text{tr}}-1} (d_{\alpha,nm}|y_{\alpha,nm}(\xi)), \quad (30)$$

where

$$(d_{\alpha,nm}|y_{\alpha,nm}(\xi)) = \mathcal{T}_\alpha(f_{nm}; \xi) \sum_{k=1}^{N_{\mathcal{F}}} \frac{\tilde{d}_{\alpha,nm}^*[k] \tilde{h}_{\alpha,nm}(\xi)[k]}{S_{\alpha,nm}[k]} \delta f, \quad (31)$$

and $M_{\text{tr}} = T_{\text{seg}}/\Delta T$ is the number of tranches in each segment, $\delta f = 1/\Delta T$, and $N_{\mathcal{F}} \neq N_f$ is the number of frequency components over which the summation is performed. Analogously, Eq. (7) can be written in the form:

$$\begin{aligned} \langle d_n|y_n(\xi) e^{i\phi_n} \rangle &= 4 \operatorname{Re} \left\{ e^{i\phi_n} \sum_{\alpha=1}^3 (d_{\alpha,n}|y_{\alpha,n}(\xi)) \right\} \\ &= 4 \operatorname{Re} \left\{ e^{i\phi_n} \times \sum_{\alpha=1}^3 \sum_{m=0}^{M_{\text{tr}}-1} (d_{\alpha,nm}|y_{\alpha,nm}(\xi)) \right\}. \end{aligned} \quad (32)$$

Equation. (30) is significantly more efficient to compute than Eq. (10) for several reasons. The evaluation of Eq. (29) does not require a (relatively) expensive numerical Fourier transform, but it is semianalytical [via Eq. (16)] and the LISA transfer function enters just as a (complex) multiplicative factor. The data are Fourier transformed only once per tranche. Finally, the sum over the frequency components within the tranche inner product (31) is not done over the full frequency grid but on a much smaller number of frequency bins $N_{\mathcal{F}} \ll \Delta T/\Delta t$.

In this work, we find that evaluating Eq. (16) over $N_{\mathcal{F}} = 21$ bins centred on the frequency $\lceil f/f_{nm} \rceil \delta f$ (where $\lceil \cdot \rceil$ is the ceiling operator, i.e., the Fresnel kernel is centred on the nearest frequency bin above f_{nm}), instead of all $2160(\Delta T/6 \text{ hr})(5\text{s}/\Delta t)$ bins, is sufficient to recover $> 99\%$ of the waveform power. This is a hundred-fold reduction in data volume with respect to the full frequency grid, greatly reducing the cost of Eq. (8).

C. Hardware acceleration and computational cost

We can further improve the computational efficiency of Eq. (30), and therefore reduce the overall cost of the evaluation of the detection statistic, by taking advantage of GPU hardware. To do so, we must first identify what stage of the computation should be distributed among each block of GPU threads. The architecture of modern GPUs typically requires specifying a size of 32^n threads per block (where $n \in \mathbb{N}$), so imposing parallelism at the level of the evaluation of Eq. (31) over Fourier bins within a tranche would significantly limit flexibility. One could instead parallelise the computation with respect to the tranche index m , as multiple years of LISA data consists of thousands of these elements. This choice is best when the desired application is to evaluate statistics for $\lesssim 10^2$ sets of waveform parameters in parallel, as it makes full use of the GPU thread-count, but incurs additional memory management and synchronization costs as many thread blocks must collaborate to evaluate the statistic for each set of source parameters.

In our pipeline, we typically evaluate more than 10^5 detection statistics at once to rapidly explore the SOBHB parameter space. This motivates parallelizing Eq. (8) at the highest level possible, devoting one GPU thread to each set of source parameters. Detection statistics can then in principle be computed en masse for the entire batch at each iteration of the optimization algorithm, achieving extremely low wall times per parameter set. However, naïvely attempting to compute 10^5 waveforms in parallel would entail the construction of a matrix of waveforms of

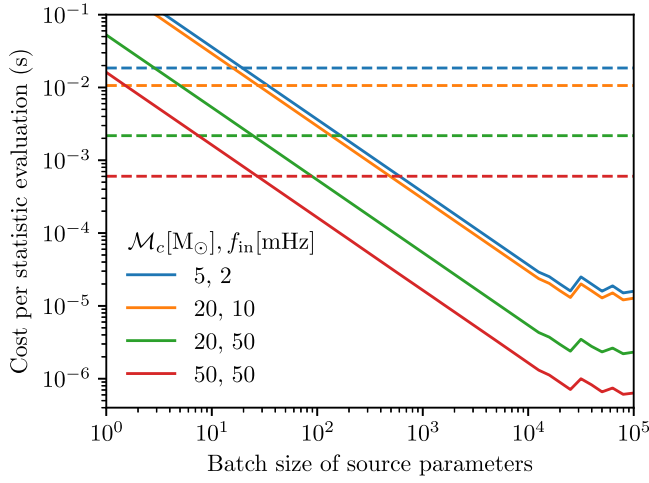


FIG. 1. Wall time per set of source parameters in the evaluation of Eq. (8) as a function of batch size for $N_{\mathcal{F}} = 21$ [see Eq. (30)], $T_{\text{obs}} = 2$ yr and $\Delta t = 5$ s. The implementation on an NVIDIA A100 GPU (solid) outperforms the one on a 2.4 GHz Intel Xeon Platinum single-core CPU (dashed lines) starting from a batch size ≈ 30 (single core). Wall times are averaged uniformly with respect to mass ratio and eccentricity within the bounds targeted by our search. The oscillations in the wall time for batches larger than $\sim 10^4$ are due to saturation of GPU resources.

$(10^5 \times N_{\text{seg}} \times M_{\text{tr}} \times N_{\mathcal{F}}) \approx 10^{10}$ elements, which is expensive and too large for the global memory of most GPUs.

Rather than storing waveforms and then computing statistics as two separate operations, we address this limitation by storing values of $\Upsilon_{N_{\text{seg}}}$ that are updated directly via the inner products $\langle d_{\alpha,n} | y_{\alpha,n} \rangle$ and $\langle y_{\alpha,n} | y_{\alpha,n} \rangle$, tranche by tranche, as the waveform is constructed. In doing so, no waveform storage is necessary, and the large matrix of

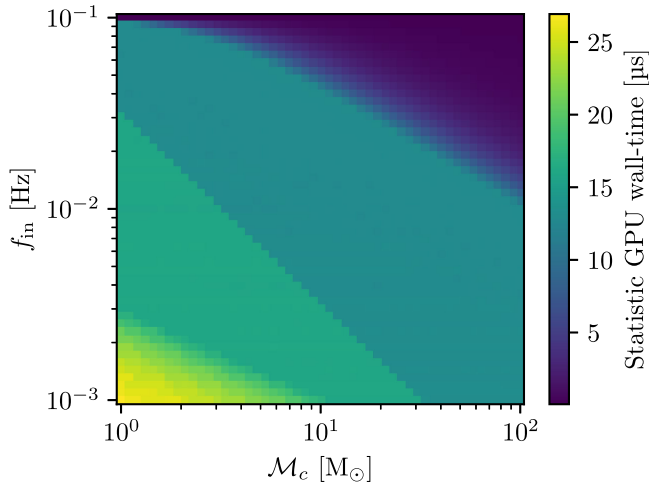


FIG. 2. Computational wall time per evaluation of Eq. (8) in the $(\mathcal{M}_c, f_{\text{in}})$ plane, averaged over the mass ratio and eccentricity ranges probed by our pipeline. The upper-right corner of the plot (where costs are lowest) corresponds to signals that exit the LISA band during the observational window.

waveforms is replaced with a vector of 10^5 statistic values (one per set of source parameters). This level of memory management also necessitates the use of low-level CUDA kernels, as opposed to the use of high-level frameworks such as JAX. Direct evaluation of $\Upsilon_{N_{\text{seg}}}$ is an essential component of our methodology—it minimizes the number of memory allocations required, which ultimately bottlenecks GPU operations, and facilitates sufficiently large batch sizes for Eq. (8) to effectively be computed in microseconds. This is demonstrated in Fig. 1. We observed that speed-ups of a factor $\sim 10^3$ with respect to a single-core CPU are achieved when the batch size is larger than $\sim 10^4$. The resulting algorithm is also extremely lightweight with respect to global memory usage, requiring $\lesssim 1$ GB of memory even for large batch sizes; it is therefore well-suited to deployment on a large number of lower-grade GPUs, for which the memory available can be limited to a few GB.

In Fig. 2, we show the GPU wall time per statistic computed (for a batch size of 10^4) in the $(\mathcal{M}_c, f_{\text{in}})$ plane explored by our search pipeline. Performance varies from ~ 1 μs for higher $(\mathcal{M}_c, f_{\text{in}})$, to ~ 25 μs for lower $(\mathcal{M}_c, f_{\text{in}})$, over the parameter space. The fuzzy boundary (due to averaging over other source parameters) in Fig. 2 follows the parametric curve $t_c(\mathcal{M}_c, f_{\text{in}}) = T_{\text{obs}}$, with statistics for sources that chirp out-of-band during the observation being less expensive to compute (as the signal occupies fewer tranches). Other features in Fig. 2 are associated with hardware-specific resource limits that lead to differences in performance as a function of parameter space. While significantly varying performance over parameter space can limit the efficiency of large batch sizes (as during synchronizing operations the GPU is forced to wait for the slowest threads to finish), our tile-based approach to the search ensures that all possible sources within a single parameter space tile have similar t_c , so the cost-per-statistic is roughly constant even for large batches of simultaneously evaluated waveforms/search statistics. Our approach is roughly three orders of magnitude faster than the GPU-accelerated frequency-domain implementation of Eq. (8) employed in previous work (as shown in Table I), a sufficient gain in efficiency to enable searches over the full parameter space of mildly eccentric SOBHB signals.

D. Implementation details

Our goal is to detect SOBHBs over multiple years of LISA observations. The dataset consists of the three pseudo-orthogonal TDI time series represented as $d_{\alpha}(t) = n_{\alpha}(t) + y_{\alpha}(t)$, where α labels the second-generation TDI observables A, E and T [55]. Here $n_{\alpha}(t)$ and $y_{\alpha}(t)$ are the total noise and GW signal contributions to the data stream, respectively. In this study, we consider GWs from mildly eccentric (i.e., with eccentricity $e \lesssim 0.01$) SOBHBs for which the two independent polarizations $h_{+,\times}(t)$ are well modeled by the dominant $\ell = |m| = 2$ mode. We will denote with θ the set of parameters that describe a GW

signal at the LISA output. The TDI observable $y_\alpha(t)$ produced by $h_{+,\times}(t)$ can be written as $y_\alpha(t) = \mathfrak{T}_\alpha[t, f, L(t)]h(t)$, where $\mathfrak{T}_\alpha[t, f, L(t)]$ is an operator that projects $h = h_+ - ih_\times$ onto the laser paths connecting the movable optical subassemblies of the LISA constellation (the LISA “arms”) with the appropriate time delays. The complexity of this response function, which depends on time, t , GW frequency, f , and the slowly varying LISA constellation arm-lengths, $L \approx 8.3$ s, further complicates and increases the computational costs of constructing TDI variables, particularly in the frequency band relevant to SOBHBs (and EMRIs) where the induced phase shifts, which scale as $\pi L f \approx 0.33(L/8.3 \text{ sec})(f/10 \text{ mHz})$, cannot be ignored. Following Ref. [31], we approximate TDI-2 variables with rescaled TDI-1.5 variables, which account for the rigid rotation of a constant-arm-length constellation [50,56] while ignoring point-ahead corrections (the accuracy of this approximation is quantified in Appendix E of Ref. [57]). The TDI variables are constructed using single-link observables, the transfer functions for these can be found in Eqs. (62) and (63) in Ref. [58]. Overall, our approach is tailored to take maximum advantage of the hierarchy of timescales, $T_{\text{obs}} \geq T_{\text{seg}} \geq \Delta T$, that we have introduced at the beginning of this section.

To efficiently explore the detection statistic while exploiting the massively parallel nature of the problem, we employ a particle swarm optimiser to identify the global maximum of the detection statistic in each parameter tile. For further details of the implementation and settings of this optimizer, see Refs. [30,59].

III. SEARCHING THE YORSH DATASET

We apply this search pipeline to the most recent LDC release, *Yorsh* 1.b, see Fig. 3. This is a 2 y dataset, sampled

at 5 s cadence, of second-generation TDI variables built under the approximation of constant and unequal constellation arm-lengths.

The instrumental noise is Gaussian and stationary, with an additional cyclo-stationary foreground due to a population of 3×10^7 Galactic ultracompact binaries with signals of $\text{SNR} > 50$ removed to simulate residuals returned by a global fit analysis of these sources [60,61]. The data also include overlapping signals from 8 SOBHBs with $\mathcal{M}_c \approx 7\text{--}40 M_\odot$, mass ratio $q \approx 0.5\text{--}1$, effective spin parameter [62] $|\chi_{\text{eff}}| \approx 0\text{--}0.5$, $f_{\text{in}} \approx 3\text{--}30$ mHz, where $f_{\text{in}} \equiv f(t_0)$, (corresponding to $t_c \approx 1.5\text{--}300$ yr) generated using the spin-aligned and quasicircular IMRPhenomD waveform approximant [63]. Injection optimal SNRs ρ^{inj} range from $\approx 8\text{--}25$, with parameters given in Table II.

In the search, we employ the waveform approximant *TaylorF2Ecc* which allows us to search for spin-aligned and eccentric binaries [64,65], but in the search templates we set $\chi_{\text{eff}} = 0$; as was identified in Ref. [31], aligned spin effects introduce small deviations to the inspiral phase that can be absorbed by a phase maximization and small parameter shift with little to no loss in SNR, and we find the inclusion of χ_{eff} does not impact search performance. This parameter can be readily determined in, e.g., a follow-up Bayesian parameter estimation analysis, which we do not perform in this work, but has been demonstrated previously in Refs. [30,31]. The search parameter vector on which the detection statistic, Eq. (5), depends is $\xi = \{\mathcal{M}_c, q, f_{\text{in}}, e_{\text{in}}, \iota, \psi, \lambda, \beta\}$, where $e_{\text{in}} \equiv e(t_0)$, ψ is the polarization angle, ι the inclination angle of the orbital plane, and (λ, β) identify the source location in the sky through the ecliptic longitude and latitude, respectively. The search covers $\mathcal{M}_c \in [5, 100] M_\odot$, $q \in [0.1, 1]$, $e_{\text{in}} \in [0, 0.01]$ and $f_{\text{in}} \in [1, 100]$ mHz, exploring the entire sky and all inclinations and polarization angles.

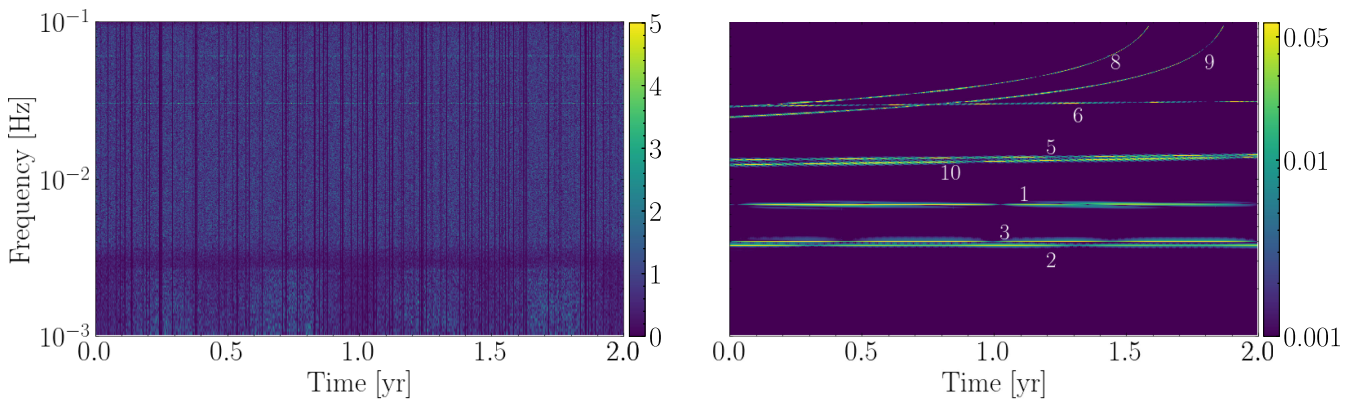


FIG. 3. Time-frequency representations of TDI channel A of the *Yorsh* dataset, containing SOBHBs. In both panels, the data are whitened according to the *Sangria* analytical PSD, which we use in the analysis (see text for more details). Left: The full (GW signals plus noise) *Yorsh* data, with data gaps at 85% duty cycle. Dark vertical lines correspond to the location of the data gaps. Other features visible by eye are described in the text. Right: The GW signal-only contribution to the data (no gaps), showing the time-frequency tracks of the 8 SOBHBs present in the dataset (labels correspond to the *Yorsh* source ID, see II); all signals are completely invisible in the time-frequency data on the left. Both color bars quantify the amplitude of the (dimensionless) whitened strain; the left (right) panel is scaled linearly (logarithmically).

TABLE II. Source parameters of the SOBHBs present in the *Yorsh* LDC dataset. The final column reports the optimal SNR of each injection (ρ^{inj}) computed using the injected waveforms (from the noiseless data streams) and the LDC *Sangria* PSD. We (arbitrarily) set the zero reference time to the beginning of the LISA data, hence the time of coalescence, t_c^{inj} , also corresponds to remaining binary lifetime from the start of LISA observations. All of the injected binaries were on quasicircular orbits (zero eccentricity).

ID	$\mathcal{M}_c^{\text{inj}} [M_\odot]$	$t_c^{\text{inj}} [\text{y}]$	$f_{\text{in}}^{\text{inj}} [\text{mHz}]$	$D_L^{\text{inj}} [\text{Mpc}]$	q^{inj}	$\chi_{\text{eff}}^{\text{inj}}$	ρ^{inj}
1 ^a	29.34741587	65.9176	5.85830665	159.9	0.91	0.50	10.91
2	40.3876144	281.3679	2.78485696	93.06	0.83	-0.23	8.28
3	34.51216704	297.7712	3.00698596	47.0	0.58	0.10	9.88
5 ^a	27.41970433	10.3457	12.24273032	168.3	0.83	-0.55	12.93
6	7.007404972	11.0420	28.02352272	17.3	0.88	-0.17	13.99
8 ^a	22.40969304	1.6501	27.65438527	34.0	0.59	0.002	23.74
9 ^a	26.08583360	1.9185	23.76783772	85.5	0.95	0.10	23.94
10 ^a	39.14942200	7.0604	11.31112717	168.9	0.88	0.03	24.66

^aSources marked with “^a”, are those that are identified by our search pipeline.

In constructing the time-frequency representation of $y_a(\xi)$, we set a tranche duration of $\Delta T = 6 \text{ hr}$ (3 hr) for f_{in} below (above) 10 mHz. Shorter tranches are used for higher f_{in} in order to better retain the accuracy of the linear approximation to the frequency evolution. Each tranche is Tukey-windowed with a shape parameter of 0.1 and fast-Fourier-transformed to construct the time-frequency grid over which our analysis is conducted. We model the noise PSD in each tranche according to the *Sangria* LDC model [66], which accounts for both instrumental and Galactic foreground noise sources [67]. This PSD crucially does not model the cyclostationarity of the Galactic foreground, leading to the periodic variations below $\sim 3 \text{ mHz}$ in the left panel of Fig. 3. Note the dip in the whitened spectrogram at $\sim 3 \text{ mHz}$: the foreground contribution to the *Sangria* PSD overestimates confusion noise in this region, which manifests as a feature in our results (Fig. 4) at low frequencies. The spurious horizontal lines in the left panel of Fig. 3 are caused by inconsistencies between the *Yorsh* data and the *Sangria* PSD. Despite these inconsistencies, our pipeline is sufficiently robust to noise modelling errors that the results of our search are not strongly impacted.

To ensure even coverage of the parameter space and enable parallelization over a pool of GPUs, we tile the search volume in $(\mathcal{M}_c, f_{\text{in}})$. Each tile has dimensions of $\Delta \mathcal{M}_c = 5 M_\odot$ and $\Delta f_{\text{in}} = 10^{-4} (10^{-3}) \text{ Hz}$ for f_{in} below (above) 10^{-2} Hz , chosen to approximately balance the per-tile computational cost (the GPU wall time per tile is $\lesssim 0.5 \text{ h}$, see Table I). We exclude a small number of tiles for the largest values of $\mathcal{M}_c$ and f_{in} , as they correspond to signals for which $t_c \ll T_{\text{obs}}$; in this regime the slow GW frequency evolution assumption central to the efficient implementation of the search breaks down. We then search over each of the remaining 2680 tiles with 4 NVIDIA A100 GPUs, which is completed in ~ 11 days. Results are shown in Fig. 4.

A. Search results on the full dataset

In the upper-left panel of Fig. 4, we show the maximum of the detection statistic, $\hat{\Upsilon}_1 \equiv \max_{\xi}(\Upsilon_1)$, returned by the search within each tile. We consider a tile to be statistically significant if $\hat{\Upsilon}_1$ crosses an (arbitrary) threshold of $\Upsilon_1 = 90$, corresponding to a false alarm probability which differs from tile to tile but is $\sim 10^{-4}$. To quantify the detection significance of each source candidate, we construct $\hat{\Upsilon}_1$ background distributions by repeating an identical search over 300 realizations of noise-only data (as described by our noise model) for tiles in which $\Upsilon_1 > 90$, see bottom panel of Fig. 4. The returned $\hat{\Upsilon}_1$ lie far outwith the background distributions in all cases, and we estimate the false alarm probability of these signals to be $\ll 10^{-5}$. Note that the background distributions for the 85% duty cycle dataset (whose analysis we describe below) are statistically equivalent to those of the full *Yorsh* dataset (as they are ultimately only a function of N_{seg} and the dimensions of the search space).

We further verify these detections by examining Υ_1 as a function of sky position conditioned on all other parameters (right panel of Fig. 4), finding that the location of each injected source agrees precisely with the bulk of Υ_1 returned by the search. In Sec. III C, we also show that the values of $(\mathcal{M}_c, f_{\text{in}})$ returned by the search are consistent to better than one part in $\sim 10^4$ with the injection parameters for each of these sources.

The search detects all signals with $\rho^{\text{inj}} \gtrsim 10$, with the exception of source 6. Notably, source 1 ($\rho^{\text{inj}} \approx 11$) is correctly identified, which would not have been possible if our segment ladder started at $N_{\text{seg}} = 100$ [31].

After exploring this result in more detail, we confirmed that this is because source 6 occupies a smaller fraction of the search volume than other signals of lower SNR (likely due to its low chirp mass), see further discussion in Sec. III D. This is an artefact of our relatively simple approach for tiling the $(\mathcal{M}_c, f_{\text{in}})$ space; we expect that

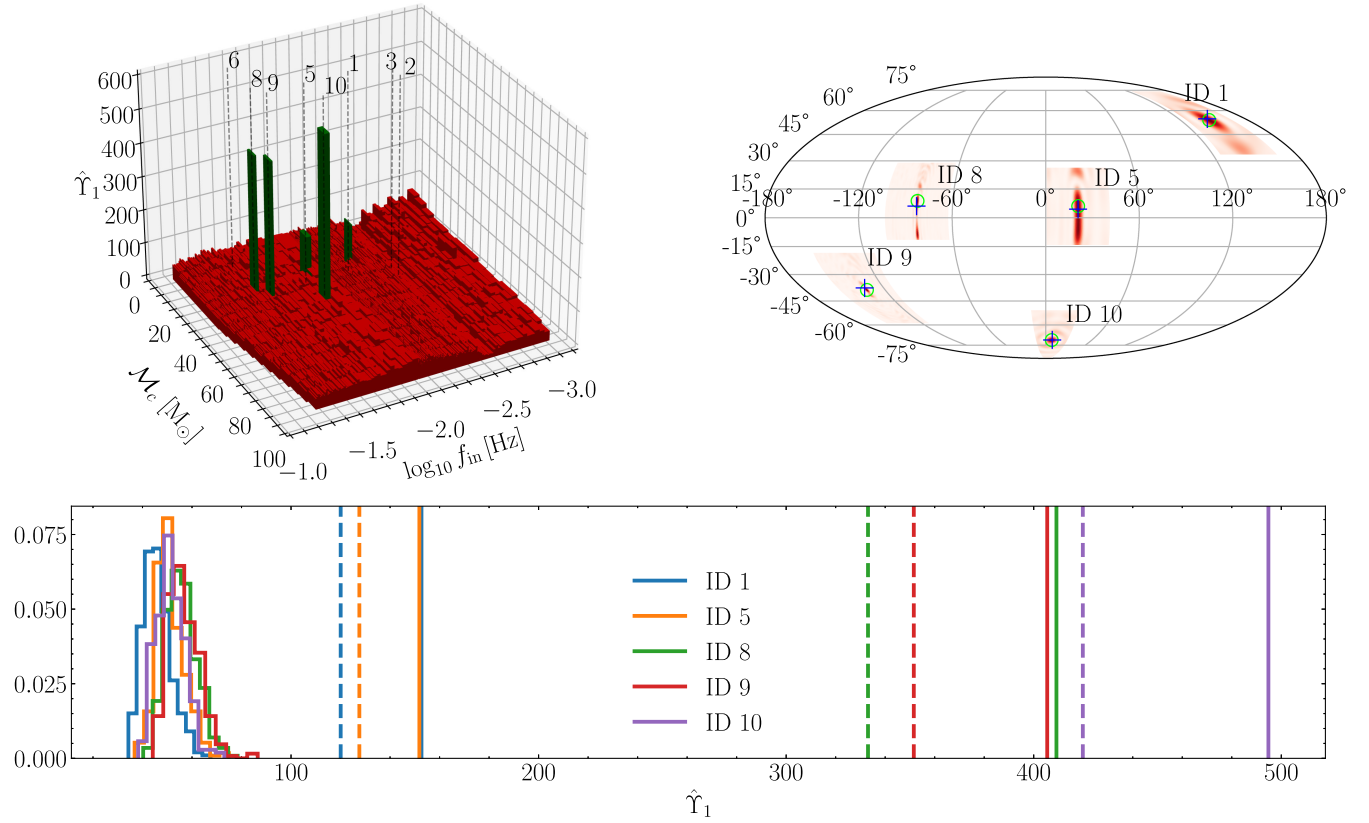


FIG. 4. The results of the search on both 100% and 85% duty cycle datasets. Top left: Maximum detection statistic $\hat{\Upsilon}_1 \equiv \max_{\xi}(\Upsilon_1)$ within each search tile (M_c, f_{in}). A conservative threshold of $\Upsilon_1 \geq 90$ distinguishes significant search triggers (green) from noise triggers (red). Vertical black dotted lines indicate tiles with injected sources. Top right: Heat-map on the sky of $\Upsilon_1(\lambda, \beta)$, conditioned on the search results for each detected source. Blue crosses denote the true location of a source and lime green circles denote the sky position returned by the search. We stress that the density function $\exp(\Upsilon_1(\lambda, \beta)/2)$, which could serve as a proposal for estimating the posterior distribution of the source's location in the sky, is significantly more compact than the region covered by the heat-map (see Table 2 in Ref. [31]). Both top panels refer to the analysis of the 100% duty cycle dataset. Bottom: The distribution of $\hat{\Upsilon}_1$ from background analyses compared to the value of the detection statistic returned by the search for each significant tile, with solid (dashed) lines corresponding to the analysis on the 100% (85%) duty-cycle *Yorsh* dataset. Associated false alarm probability estimates are provided in Table III.

future implementations using tiles with size chosen according to the region of parameter space will be more robust to this behavior.

B. Search extension to data with gaps

The time-frequency framework offers significant flexibility when considering signals of duration longer than the typical timescales over which noise is stationary and data collection suffers interruptions. For LISA, galactic confusion noise (which is expected to dominate instrumental noise below ≈ 3 mHz) is cyclo-stationary and changes on a timescale of 6 months, and the instrumental noise itself is expected to vary on timescales $\ll T_{\text{obs}}$. Moreover, gaps (either due to scheduled instrument adjustments or unforeseen data dropouts or contaminations [68]) will be present; the LISA science data duty cycle requirement is $> 82\%$ [11]. Viable strategies for LISA analysis must be robust to such features of the data stream.

To examine the performance of our approach under such conditions, we zero-out random portions of the *Yorsh* data stream to simulate the presence of gaps and produce an 85% duty-cycle *Yorsh* dataset. Gaps are produced at random times with 6 hour durations, accumulating a total data dropout of 15% over the observation period. The spectrogram of the resulting data is shown in the left panel of Fig. 3. We then reprocess the data with the same pipeline, with the only change being that tranches that overlap with data gaps are not included in the computation of the time-frequency inner product [Eq. (7)], and therefore do not contribute to the detection statistic [Eq. (5)]. For further flexibility, our framework permits a time-frequency PSD that varies between *tranches*, which can be used to model noise spectra which are locally stationary [28,69]; we do not demonstrate this capability in this work, but have verified this performs as expected for our implementation.

The end-to-end data processing remains efficient and is consistent with the results shown in Table I. The results of

this analysis are shown in Fig. 4, and additional results are presented in Sec. III C. The bottom panel of Fig. 4 shows that all of the sources that were identified in the full *Yorsh* dataset are still recovered when gaps are present. As expected, Υ_1 is smaller in all cases, as $\Upsilon_1 \approx \rho^2$, which has decreased due to the loss of data. The fact that the values of Υ_1 obtained are $\approx 15\%$ less than in the search of the full *Yorsh* data (in close agreement with the loss in ρ^2 expected for a perfectly monochromatic signal) verifies that our approach is capable of analysing data containing frequent gaps without incurring additional sensitivity losses or sacrificing computational performance.

C. Injected and recovered signals

Here we provide additional details of the signals injected in the LDC dataset *Yorsh*, as well as results supporting the accuracy and significance of the detections reported by our search pipeline.

Table II shows the details of the binaries whose GW signal was included in the *Yorsh* LDC dataset. The *Yorsh* catalog does not provide t_c ; the values provided in Table II are inferred from the provided parameters using the TaylorF2Ecc model (with $e = 0$). We also report the optimal signal-to-noise ratio of the injections, computed with the LDC *Sangria* PSD model used in the search. Our source naming convention follows that of the *Yorsh* technical document [61]; note that source IDs 4 and 7 are not included in the table because these signals are not present in the dataset.

Assessing statistical significance of candidates will generally be nontrivial for LISA data due to the challenge of generating (many statistically identical) signal-free datasets. Here, we calculate the false alarm probabilities for source candidates by comparing against a background distribution computed using *simulated* noise-only data; this quantity corresponds to the probability that the search will find a source candidate with $\Upsilon_1 \geq \Upsilon_{1|\text{candidate}}$. This involves generating several statistically independent realizations of noise-only data and running the search on each realization, empirically building the background distribution of $p_0(\Upsilon_1)$. Following Ref. [30], we extrapolate this

TABLE III. False alarm probabilities extrapolated from Υ_1 background distributions in search tiles passing the detection threshold. The power-law and Gumbel fits provide conservative and optimistic estimates respectively (see text for details).

Source		Power law		Gumbel	
ID	ρ^{inj}	100%	85%	100%	85%
1	10.91	5.6×10^{-7}	8.7×10^{-6}	9.2×10^{-11}	9.0×10^{-8}
5	12.93	3.2×10^{-8}	4.9×10^{-7}	8.8×10^{-10}	1.2×10^{-7}
8	23.74	5.2×10^{-10}	4.4×10^{-9}	$\sim 10^{-28}$	$\sim 10^{-22}$
9	23.94	3.5×10^{-13}	2.7×10^{-12}	$\sim 10^{-28}$	$\sim 10^{-24}$
10	24.66	8.3×10^{-18}	1.5×10^{-16}	$\sim 10^{-39}$	$\sim 10^{-32}$

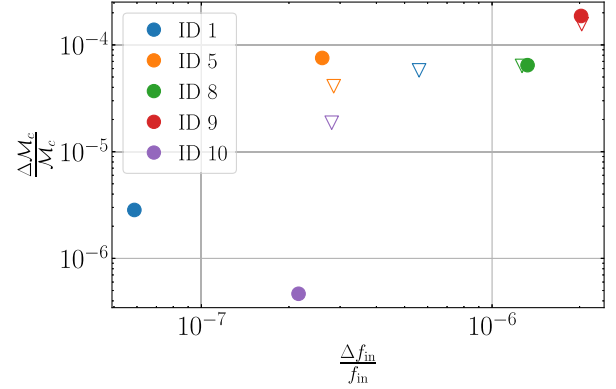


FIG. 5. Relative error in the recovery of $\mathcal{M}_c$ and f_{in} (with respect to injected values) for detected signals at either 100% (filled circles) or 85% (triangles) duty cycle. As we search with a stochastic optimiser, some variability in the maximum-statistic parameters is expected, which can cause relative error to decrease despite the loss in SNR when gaps are present.

distribution using a power-law tail (conservative) or Gumbel (optimistic); these extrapolated fits are used to compute the upper and lower bounds for the false-alarm-probabilities for each of the candidate detections. Note that the background distributions in Fig. 4 are computed individually for each search tile (where there is a candidate detection); the variation in the background distributions is caused by the difference in the prior-to-posterior ratio across parameter space; using a metric-based approach (such as that employed in Ref. [70]) to place search tiles would reduce the differences between these distributions.

In Table III, we report false alarm probabilities with which the signals are recovered in the search, for both the full (100%) and 85% duty cycle *Yorsh* datasets. In the case of a power-law extrapolation, we fit δ for $p_0(\Upsilon_1) \propto \Upsilon_1^\delta$; for the Gumbel distribution $p_0(\Upsilon_1) \propto G(\mu; \beta)$, we fit the free parameters μ and β (location and scale, respectively).

To demonstrate that our search pipeline successfully identifies SOBHBs, as opposed to merely detecting them, we also compute relative errors in the recovered $\mathcal{M}_c$ and f_{in} for each source relative to their injected values (Table II), which complements the results concerning the sky location that are reported in Fig. 4. The results are shown in Fig. 5. In all cases, the injected and recovered parameters are in close agreement, with relative errors on scales typical of the measurement precision of these parameters, see, e.g., Ref. [27,31]. The results of the analysis can then be used to seed a dedicated follow-up parameter estimation stage, which can efficiently return posterior distribution functions on the source parameters.

D. Analysis of source ID 6

Source 6 is the only injected signal with $\rho > 10$ that was not found by the search.

Our search can fail to identify a signal for one of two reasons: (i) ρ is too low for the maximum segment count of

$N_{\text{seg}} = 50$ (i.e., the signal’s statistic value is not significant with respect to noise fluctuations), or (ii) the peak in the search surface is too narrow at $N_{\text{seg}} = 50$ to be reliably encountered by the stochastic optimizer. While sources 2 and 3 are not identified due to the former reason (as their SNRs are lower, $\rho^{\text{inj}} \approx 8\text{--}9$), source 6 with $\rho^{\text{inj}} \approx 14$ is not found due to the latter reason. Indeed, we have tested that the signal is reliably detected by our pipeline once the search prior is narrowed by a factor of ~ 4 .

To understand this behavior, we performed targeted exploration of the primary mode of the posterior distributions $p(\xi) \propto \exp(\Upsilon_{50}/2)$ for each source via nested sampling, as shown in Fig. 6. This confirms that source 6 occupies a smaller fraction of the prior space than source 5 (the most comparable source in both SNR, $\rho^{\text{inj}} \approx 13$, and f_{in}), which explains why this source was identified despite having a lower SNR. While the 50% confidence intervals of the posterior distributions for sources 8 and 9 (which were both successfully identified) occupy a smaller region of parameter space than source 6, the tails of these distributions extend much further across parameter space before sinking into the noise floor due to their larger SNR. This means that the particle swarms are far more likely to encounter (and then climb) the deformations in the statistic manifold corresponding to these sources. The volume occupied by the primary mode (for constant SNR) is an intrinsic property of the waveform model (and hence the

search statistic) as a function of the $(\mathcal{M}_c, f_{\text{in}})$ parameter space. An adaptive tiling strategy that accounts for this behavior (similarly to what is employed in the construction of template banks for ground-based searches [71,72]) and redistributes search efforts will improve the reliability of the search at little additional cost. It is our expectation that such a tiling strategy will allow us to find source 6. We will revisit strategies for parameter-space tiling in future work.

IV. OUTLOOK

The time-frequency GPU-accelerated search approach for long-duration, broad-band GW signals that we have successfully demonstrated here provides a solution to the long-standing problem of searching globally for SOBHBs in LISA. In the most naïve version in which each tile is assigned to a GPU, it meets the requirement of completing a global search every few weeks as the mission progresses—the SNR accumulates slowly, approximately $\propto T_{\text{obs}}^{1/2}$ —while achieving $\lesssim 1$ hr latency in specific tiles, should detection candidates, and the need for alerts, start to emerge. Data processing in $\approx$ hr—long time segments seamlessly accounts for data-gaps and the complexity of noise nonstationarities over the (much longer) observation timescale [75,76].

There are several extensions and improvements to the approach presented here that we intend to pursue in the

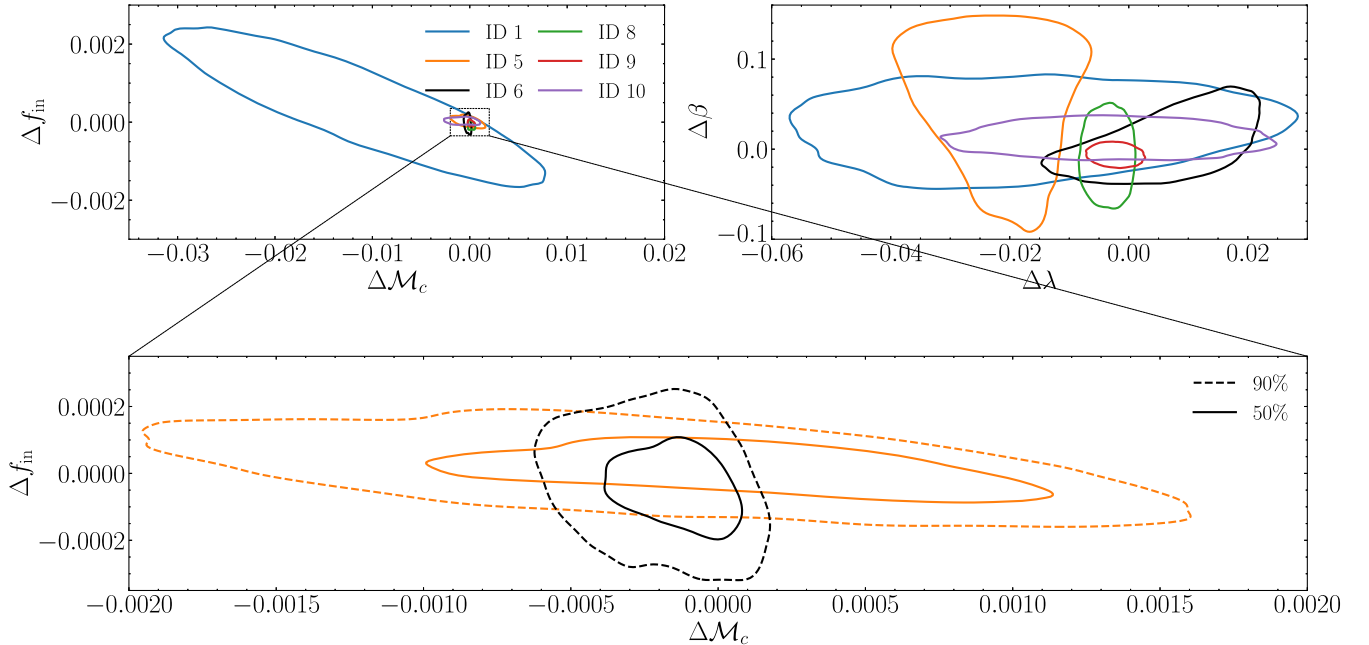


FIG. 6. Comparison of marginal posterior distributions for source 6 (not detected by our pipeline) with successfully identified sources in the *Yorsh* dataset (full duty cycle). Top panel: Marginal posterior distributions of $p(\xi) \propto \exp(\Upsilon_{50}/2)$ in $(\mathcal{M}_c, f_{\text{in}})$ and (λ, β) , presented in units of search space dimensions, for the *Yorsh* injections (Table II). Bottom panel: Closer view of the (50%, 90%) intervals for sources 5 and 6, which are similar in SNR and f_{in} (and thus are the most comparable). The larger extent of the distribution for source 5 (particularly with respect to $\mathcal{M}_c$) explains why this source was identified while source 6 was not. We emphasize that these are *not* coherent marginal posterior distributions. The nested sampling implementation within `nessai` was used to stochastically sample these distributions [73,74].

future. While in this study analytic orbits are used to model the spacecraft [77], this can be generalized, at no additional computational costs to numerical and fully generic orbits with varying arm-length and point-ahead corrections [78,79]. A more rigorous metric-based approach to the $(\mathcal{M}_c, f_{\text{in}})$ tiling would ensure that the ratio between the search volume vs the posterior volume is constant across the parameter space, solving the issue that we have encountered with source 6 in *Yorsh* [71,72]. We plan to further expand the search parameter space to cover black holes with arbitrary spin orientation (which induces precession of the orbital plane and modulations of the emitted GWs), orbits with large eccentricity and radiation's multiple modes, all of which will be particularly pertinent when tackling the even more challenging problem of EMRI identification. The time-frequency representation of the waveform can be extended to the multiharmonic case, simply by computing the waveform quantities $\{A, f, \dot{f}, \phi\}$ for each harmonic, and subsequently applying the Fresnel expressions to each harmonic.

The increased computational burden incurred by considering waveforms which include additional physical effects discussed above can be mitigated by using a parameter space tiling scheme [e.g., as done here over $(f_0, \mathcal{M}_c)$] which is *adaptive*. Such a tiling scheme would have greater resolution (i.e., more tiles) in parts of parameter space where the posterior is expected to be narrower. The search is carried out independently in each tile, and as a result the cost of the search can be crudely estimated to scale as $\mathcal{O}(N_{\text{tiles}}/N_{\text{GPUs}})$. Additionally, improvements to GPU hardware, particularly those which allow use of higher numbers of concurrent threads, will also significantly speed up the methods presented here.

More broadly, our approach is applicable to a wide variety of GW searches, both for LISA and the current ground-based detector network, providing a viable path to expanding the parameter space that can be covered.

ACKNOWLEDGMENTS

D. B. is supported by MPG and acknowledges support from UK Space Agency Grant No. UKRI971. C. E. A. C-B. and A. V. are supported by UKSA Space Agency Grant No. UKRI971. A. V. acknowledges the support of the Royal Society and Wolfson Foundation. A. V. thanks the Flatiron Institute for hospitality while a portion of this research was carried out. The Flatiron Institute is a division of the Simons Foundation. The computations for this paper were performed using the University of Birmingham's BlueBEAR High Performance Computing facility. D. B. is supported by the Max Planck Institute for Gravitational Physics (Albert Einstein Institute).

DATA AVAILABILITY

Data supporting the figures, tables and conclusions of this work are openly available [80].

APPENDIX: COMPACTNESS OF THE FRESNEL-KERNEL FUNCTION

The compactness of the Fresnel-kernel function, Eq. (19), i.e., the fact that its evaluation over 21 frequency bins centered on f_{nm} for each tranche provides a faithful approximation of the Fourier transform of a waveform via Eq. (16), is one of the key elements that enables the orders of magnitude reduction in computational cost demonstrated in this papers. Here we discuss the variation of this compactness across the parameter space of the search that we have presented. The Fresnel-kernel function remains compact over the vast majority of the $\mathcal{M}_c \in [5, 100]M_\odot$ and $f_{\text{in}} \in [1, 100]$ mHz parameter space, with the exception of a small number of tiles at the edge of the search space ($f_{\text{in}} \gtrsim 0.05$ Hz, $\mathcal{M}_c \gtrsim 60M_\odot$). However, the systems in this region are in band for $\lesssim 1$ month, and as a result are unlikely to accumulate enough SNR to be detectable in LISA as they would need to be unreasonably close. Additionally, the loss of compactness is limited to the tranches containing the portion of a signal that is about to exit the LISA frequency band, at which point the noise PSD is rising steeply and thus the contribution to the inner product is strongly suppressed from this tranche. The parameter space dependence of this compactness is illustrated in Fig. 7, the impact of this behavior on the evaluation of the Fourier transform of a signal using Eq. (16) is shown for selected values of f_{in} and $\mathcal{M}_c$ in Fig. 8.

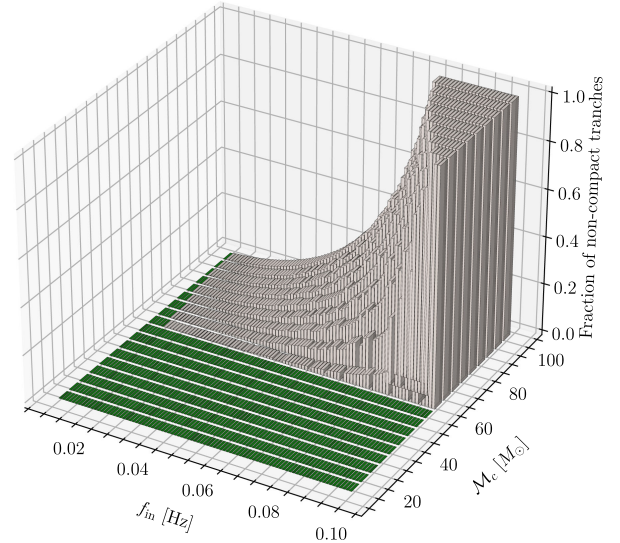


FIG. 7. Fraction of SFT tranches which violate the “compactness condition,” for simplicity we consider the signal to be compact within a tranche if it satisfies the condition $\dot{f}\Delta T \leq 10\delta f$, i.e., it crosses less than 11 Fourier bins within ΔT . A signal violating the compactness condition does not necessarily imply that the Fresnel integral is no longer valid within that tranche, as shown in the bottom row of Fig. 8.

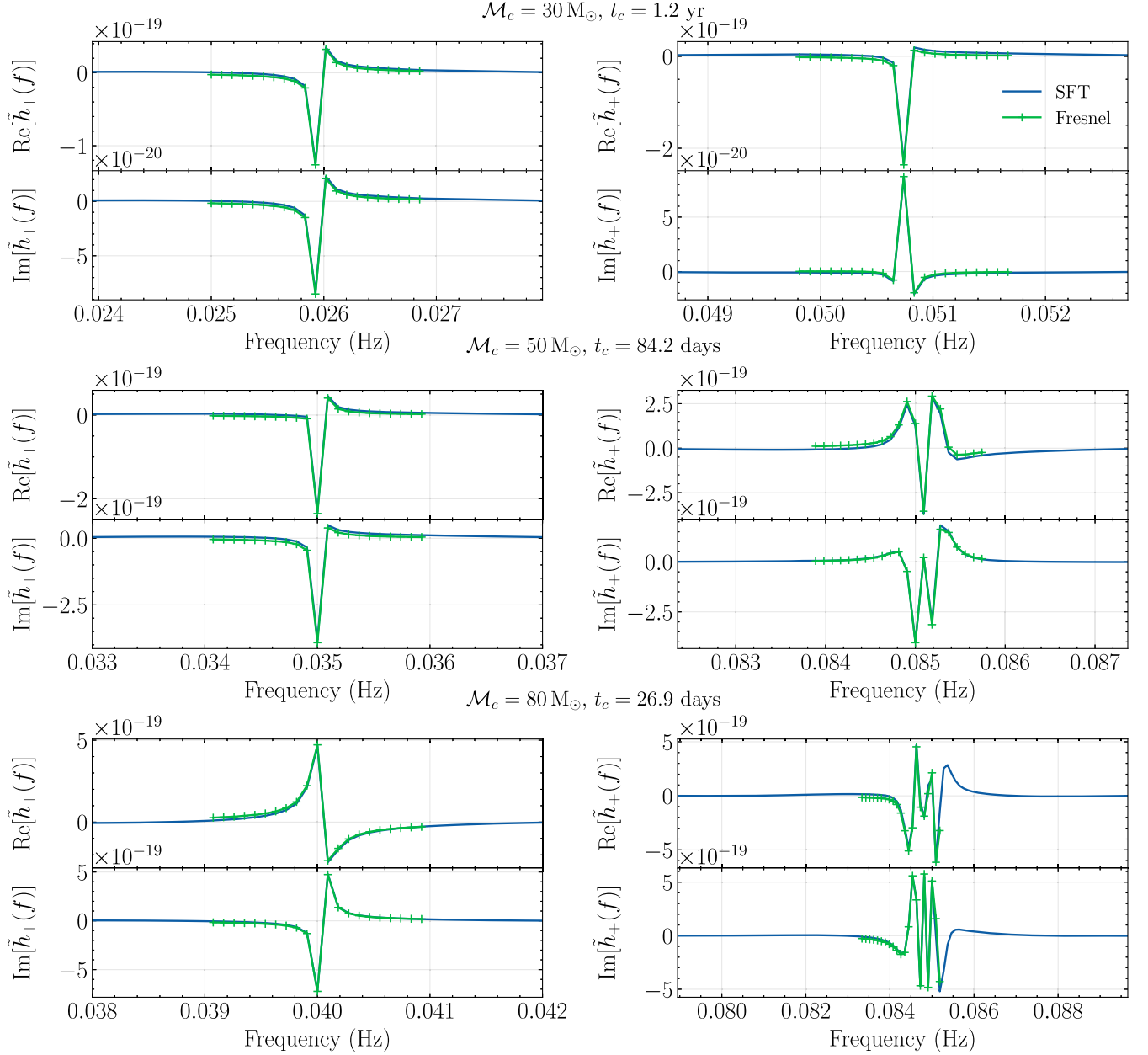


FIG. 8. Fresnel-kernel approximation of the Fourier transform of a waveform. Waveforms generated using the Fresnel-kernel function over a tranche (green lines) are compared to the short-Fourier transformed (SFT) time domain waveform (blue lines) in the first (left panels) and last (right panels) SFT time-interval. Top panel: a mildly chirping source ($\mathcal{M}_c = 30M_\odot$) that remains in the LISA frequency band over the whole observation period (in this figure $T_{\text{obs}} = 1$ yr). Middle panel: a moderately chirping source ($\mathcal{M}_c = 50M_\odot$) that exits the LISA band after 80 days. Bottom panel: a highly chirping source ($\mathcal{M}_c = 80M_\odot$) that exits the LISA frequency band within less than a month. Waveforms computed using the Fresnel-kernel function are generated at 20 points around f_{nm} , which is sufficient for the full evolution of the mildly and moderately chirping sources. However, in the last tranche of the highly chirping source, the SFT of the waveforms is not captured fully by the approximation (bottom right plot). This illustrates that the compactness of the Fresnel-kernel breaks down in the late inspiral for highly chirping systems. Note that in this case extending the number of frequency points over which the Fresnel-kernel function is computed would recover the SFT waveform, but add to the computational burden. For simplicity, in this figure we have computed waveforms retaining only the zeroth order post-Newtonian term of the waveform phase, which is fully described by the chirp mass and time-to-merger.

- [1] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), *Astrophys. J. Lett.* **1004**, L21 (2026).
- [2] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
- [3] F. Acernese *et al.* (VIRGO Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [4] T. Akutsu *et al.* (KAGRA Collaboration), *Prog. Theor. Exp. Phys.* **2021**, 05A101 (2021).
- [5] M. Punturo *et al.*, *Classical Quantum Gravity* **27**, 194002 (2010).
- [6] M. Evans *et al.*, [arXiv:2109.09882](https://arxiv.org/abs/2109.09882).
- [7] LVK Collaboration, *Mon. Not. R. Astron. Soc.* **524**, 5984 (2023).
- [8] A. G. Abac *et al.*, *Astrophys. J.* **973**, 132 (2024).
- [9] I. Harry, S. Privitera, A. Bohé, and A. Buonanno, *Phys. Rev. D* **94**, 024012 (2016).
- [10] P. Amaro-Seoane *et al.*, [arXiv:1702.00786](https://arxiv.org/abs/1702.00786).
- [11] M. Colpi *et al.*, [arXiv:2402.07571](https://arxiv.org/abs/2402.07571).
- [12] A. Sesana, *Phys. Rev. Lett.* **116**, 231102 (2016).
- [13] D. Gerosa, S. Ma, K. W. K. Wong, E. Berti, R. O’Shaughnessy, Y. Chen, and K. Belczynski, *Phys. Rev. D* **99**, 103004 (2019).
- [14] C. J. Moore, D. Gerosa, and A. Klein, *Mon. Not. R. Astron. Soc.* **488**, L94 (2019).
- [15] A. Klein, G. Pratten, R. Buscicchio, P. Schmidt, C. J. Moore, E. Finch, A. Bonino, L. M. Thomas, N. Williams, D. Gerosa, S. McGee, M. Nicholl, and A. Vecchio, [arXiv:2204.03423](https://arxiv.org/abs/2204.03423).
- [16] P. A. Seoane *et al.* (LISA Collaboration), *Living Rev. Relativity* **26**, 2 (2023).
- [17] K. G. Arun *et al.* (LISA Collaboration), *Living Rev. Relativity* **25**, 4 (2022).
- [18] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, E. S. Phinney, and M. Vallisneri, *Classical Quantum Gravity* **21**, S1595 (2004).
- [19] A. J. K. Chua and C. J. Cutler, *Phys. Rev. D* **106**, 124046 (2022).
- [20] S. H. Strub, L. Speri, and D. Giardini, *Phys. Rev. D* **113**, 084048 (2026).
- [21] J. R. Gair and G. Jones, *Classical Quantum Gravity* **24**, 1145 (2007).
- [22] L.-q. Wen and J. R. Gair, *Classical Quantum Gravity* **22**, S445 (2005).
- [23] X.-T. Zhang, C. Messenger, N. Korsakova, M. L. Chan, Y.-M. Hu, and J.-d. Zhang, *Phys. Rev. D* **105**, 123027 (2022).
- [24] P. S. Cole, J. Alvey, L. Speri, C. Weniger, U. Bhardwaj, D. Gerosa, and G. Bertone, *Phys. Rev. D* **113**, 063030 (2026).
- [25] C.-Q. Ye, H.-M. Fan, A. Torres-Orjuela, J.-d. Zhang, and Y.-M. Hu, *Phys. Rev. D* **109**, 124034 (2024).
- [26] Y. Wang, Y. Shang, S. Babak, Y. Shang, and S. Babak, *Phys. Rev. D* **86**, 104050 (2012).
- [27] R. Buscicchio, A. Klein, E. Roebber, C. J. Moore, D. Gerosa, E. Finch, and A. Vecchio, *Phys. Rev. D* **104**, 044065 (2021).
- [28] M. C. Digman and N. J. Cornish, *Phys. Rev. D* **108**, 023022 (2023).
- [29] D. Bandopadhyay and C. J. Moore, *Phys. Rev. D* **108**, 084014 (2023).
- [30] D. Bandopadhyay and C. J. Moore, *Phys. Rev. D* **110**, 103026 (2024).
- [31] D. Bandopadhyay and C. J. Moore, *Phys. Rev. D* **111**, 044039 (2025).
- [32] H. Wang, M. J. Williams, I. Harry, and Y.-M. Hu, *Phys. Rev. D* **113**, 063040 (2026).
- [33] H. Wang, I. Harry, A. Nitz, and Y.-M. Hu, *Phys. Rev. D* **109**, 063029 (2024).
- [34] S. Wu, A. H. Nitz, I. Harry, S. Babak, M. J. Williams, C. Capano, and C. Weaving, [arXiv:2506.01898](https://arxiv.org/abs/2506.01898).
- [35] A. Toubiana, S. Marsat, E. Barausse, S. Babak, and J. Baker, *Phys. Rev. D* **101**, 104038 (2020).
- [36] A. Gupta, S. Datta, S. Kastha, S. Borhanian, K. G. Arun, and B. S. Sathyaprakash, *Phys. Rev. Lett.* **125**, 201101 (2020).
- [37] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **116**, 061102 (2016).
- [38] B. P. Abbott *et al.*, *Phys. Rev. X* **6**, 041015 (2016).
- [39] B. P. Abbott *et al.*, *Phys. Rev. X* **9**, 031040 (2019).
- [40] R. Abbott *et al.*, *Phys. Rev. Lett.* **125**, 101102 (2020).
- [41] R. Abbott *et al.*, *Phys. Rev. X* **11**, 021053 (2021).
- [42] R. Abbott *et al.*, *Phys. Rev. X* **13**, 041039 (2023).
- [43] A. G. Abac *et al.*, *Astrophys. J. Lett.* **993**, L25 (2025).
- [44] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), *Astrophys. J. Lett.* **1004**, L22 (2026).
- [45] A. G. Abac *et al.* (KAGRA Collaboration, Virgo Collaboration, and LIGO Scientific Collaboration), *Phys. Rev. Lett.* **135**, 111403 (2025).
- [46] R. Buscicchio, J. Torrado, C. Caprini, G. Nardini, N. Karnesis, M. Pieroni, and A. Sesana, *J. Cosmol. Astropart. Phys.* **01** (2025) 084.
- [47] S. E. de Mink and A. King, *Astrophys. J. Lett.* **839**, L7 (2017).
- [48] S.-X. Yi and K. S. Cheng, *Astrophys. J. Lett.* **884**, L12 (2019).
- [49] R. Tenorio and D. Gerosa, *Phys. Rev. D* **111**, 104044 (2025).
- [50] M. Tinto and S. V. Dhurandhar, *Living Rev. Relativity* **24**, 1 (2021).
- [51] A. J. K. Chua, C. J. Moore, and J. R. Gair, *Phys. Rev. D* **96**, 044005 (2017).
- [52] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover Publications Inc., New York, 1964), p. 300.
- [53] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, DLMF, NIST Digital Library of Mathematical Functions, <https://dlmf.nist.gov/>, Release 1.2.4 of 2025-03-15, equations 7.5.3 and 7.5.4;
- [54] M. E. McCormick and D. R. Kraemer, *Coastal Eng.* **44**, 261 (2002).
- [55] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, *Phys. Rev. D* **66**, 122002 (2002).
- [56] N. J. Cornish and R. W. Hellings, *Classical Quantum Gravity* **20**, 4851 (2003).
- [57] J. Valencia and S. Husa, *Phys. Rev. D* **112**, 044020 (2025).
- [58] S. Marsat and J. G. Baker, [arXiv:1806.10734](https://arxiv.org/abs/1806.10734).
- [59] D. Bandopadhyay and C. J. Moore, Pyso, <https://github.com/dig07/PySO>.

- [60] P. Bacon and G. Servignat, [10.5281/zenodo.17544375](https://zenodo.org/record/17544375) (2025).
- [61] G. Servignat, Yorsh stellar black hole binaries dataset (2025), IDC Technical note.
- [62] P. Schmidt, M. Hannam, and S. Husa, *Phys. Rev. D* **86**, 104063 (2012).
- [63] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044007 (2016).
- [64] B. Moore, M. Favata, K. G. Arun, and C. K. Mishra, *Phys. Rev. D* **93**, 124061 (2016).
- [65] M. Favata, C. Kim, K. G. Arun, J. Kim, and H. W. Lee, *Phys. Rev. D* **105**, 023003 (2022).
- [66] (L. D. C. Working Group), Lisa data challenge software (2022).
- [67] S. Babak, M. L. Jeune, A. Petiteau, and M. Vallisneri, Lisa data challenge: Sangria (2020), IDC technical note.
- [68] Q. Baghi, N. Korsakova, J. Slutsky, E. Castelli, N. Karnesis, and J.-B. Bayle, *Phys. Rev. D* **105**, 042002 (2022).
- [69] M. C. Digman and N. J. Cornish, *Astrophys. J.* **940**, 10 (2022).
- [70] I. W. Harry, B. Allen, and B. S. Sathyaprakash, *Phys. Rev. D* **80**, 104014 (2009).
- [71] B. J. Owen, *Phys. Rev. D* **53**, 6749 (1996).
- [72] S. Sakon *et al.*, *Phys. Rev. D* **109**, 044066 (2024).
- [73] M. J. Williams, nessai: Nested sampling with artificial intelligence (2021).
- [74] M. J. Williams, J. Veitch, and C. Messenger, *Phys. Rev. D* **103**, 103006 (2021).
- [75] O. Burke, M. Muratore, and G. Woan, *Phys. Rev. Appl.* **25**, 034041 (2026).
- [76] O. Burke, S. Marsat, J. R. Gair, and M. L. Katz, *Phys. Rev. D* **111**, 124053 (2025).
- [77] N. J. Cornish and L. J. Rubbo, *Phys. Rev. D* **67**, 022001 (2003); **67**, 029905(E) (2003).
- [78] W. Martens and E. JoffrejX, *Journal of the Astronautical Sciences* **68**, 402 (2021).
- [79] M. L. Katz, J.-B. Bayle, A. J. K. Chua, and M. Vallisneri, *Phys. Rev. D* **106**, 103001 (2022).
- [80] C. Chapman-Bird and D. Bandopadhyay, dig07/yorsh-end-to-end-search: Resubmission copy for prd (2026).