GR24-AMALDI16 IOP Publishing
Journal of Physics: Conference Series **3177** (2026) 012112 doi:10.1088/1742-6596/3177/1/012112

A pipeline for searching and fitting instrumental glitches in LISA data

Martina Muratore[1], **Jonathan Gair**[1], **Olaf Hartwig**[2,3], **Michael L. Katz**[4], **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

E-mail: martina.muratore@aei.mpg.de

Abstract. The Laser Interferometer Space Antenna (LISA) relies on robust data analysis pipelines capable of disentangling astrophysical signals from instrumental artefacts. A major challenge will be transient disturbances, or glitches, similar to those observed in LISA Pathfinder. We present a Bayesian pipeline, based on Reversible Jump Markov Chain Monte Carlo and parallel tempering, that simultaneously searches for and characterises multiple glitches while recovering instrumental noise and astrophysical sources such as Massive Black Hole Binaries (MBHBs). We show that the pipeline can fit diverse glitch morphologies, and we apply it to a light version of the 'Spritz' Data Challenge, where gaps and galactic binaries have been removed from the data. Similarly, the MBHB signal in the 'Spritz' data set has been replaced with a model more suitable for our analysis. Our results demonstrate that high-SNR artefacts can bias MBHBs parameter estimation if unmodelled, but accurate recovery is achieved when simultaneously fitting for glitches, noise, and MBHBs signals. This represents a significant step towards a reliable global fit analysis for LISA, providing guidelines for glitch searches and parameter estimation strategies.

1 Introduction

LISA, scheduled for launch in 2035, will observe gravitational waves (GW) from massive black hole binaries (MBHBs), extreme-mass-ratio inspirals, Galactic binaries, and cosmological backgrounds. To achieve these science goals, data analysis pipelines must simultaneously infer astrophysical sources and instrumental noise. The precursor mission LISA Pathfinder (LPF) revealed spurious disturbances of unknown origin, known as glitches. Since LPF observed multiple glitch types [1], and because LISA will share key technologies of the Gravitational Reference Sensor (GRS) with LPF, similar artefacts are likely to appear in LISA. If not modelled, such disturbances could bias astrophysical inference. In this

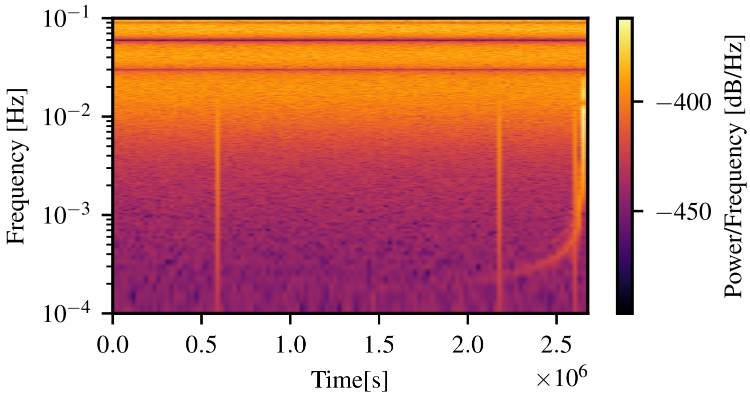

Figure 1: Time-frequency representation of the "Spritz" data set. This plot is taken from Ref. [4].

work, we introduce a Bayesian pipeline that, for the first time, employs the Reversible Jump Markov Chain Monte Carlo (RJ-MCMC) algorithm to search for and fit glitches alongside astrophysical signals and instrumental noise. Specifically, we (i) model noise in the presence of non-Gaussianities, (ii) assess the impact of glitches on MBHB parameter estimation, (iii) develop Bayesian tools for glitch search and parameter estimation using the open-source sampler Eryn [2], (iv) perform joint inference of MBHBs and instrumental noise and finally we (v) apply this algorithm to the "Spritz" data-sets [3]. The "Spritz" dataset comprises three main data sets, of which we focus on the first one, which contains a single loud MBHB, three glitches and instrumental noise as visible in Fig. 1.

2 Methodology

The analysis of LPF glitches has shown that they can be phenomenologically modelled with exponential decays characterised by specific timescales and amplitudes [5]. More recently, shapelet functions have been proposed as a flexible basis for fitting LPF glitches [6], and we adopt this approach here. Glitches are propagated through LISA's Time-Delay Interferometry (TDI) combinations [7], while astrophysical signals are modelled using the PhenomHM waveform for spin-aligned MBHBs [8], as implemented in the BBHx software package [9]. This frequency-domain model incorporates higher harmonics beyond the dominant $(2, 2)$ mode, mitigating biases in asymmetric systems. The noise spectra include contributions from the reference and test-mass interferometer backlink, test-mass acceleration, and optical metrology noises [10]. We adopt a Whittle likelihood, for which the frequency-domain log-likelihood is given by

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

where $\tilde{d}(f)$, $\tilde{h}(f, \theta)$, and $\tilde{g}(f, \phi)$ are the Fourier transforms of the data $d(t)$, the GW signal $h(t; \theta)$, and the glitch $g(t; \phi)$, respectively. The parameter sets θ, ϕ, and γ describe the GW signal, glitch, and noise model. $S(f, \gamma)$ denotes the one-sided noise power spectral density, parameterized by γ, Δf is the frequency resolution and Δt is the sampling interval. The posteriors are sampled with the open-source Eryn sampler [2]. RJ-MCMC extends standard MCMC by allowing the number of model components—in this case, glitches—to vary during sampling. Parallel tempering runs chains at different temperatures, enabling exploration of multi-modal posteriors and improving the robustness of glitch recovery [11].

3 Results

We first test the pipeline on 30 days of simulated LISA data containing seven glitches extracted from the LISA Pathfinder catalogue [1] with SNRs 21, 72, 230, 479, 637, 1172 and 1544, respectively. The left panel of Fig. 2 shows that five injected glitches are successfully recovered, with the exception of the one with the lowest SNR ($\sim$21) and the two glitches (SNRs $\sim$1172 and 1544) that occur close in time. The latters overlap in their posteriors, yielding in the 1D distribution a single broader peak that resembles a fifth, larger Gaussian at the injection time τ_0. The inability to recover the low-SNR glitch is most likely explained by the search failing to identify it, since proposed samples seldom land in the relevant region of parameter space. This limitation seems to stem from employing a uniform prior on β_0—combined with the glitch's relatively small amplitude—which is too wide compared to the τ_0 values the injected glitches. Nevertheless, this choice is still in line with the broader LPF decay time distribution [4]. This undetected event has a negligible impact on MBHB parameter recovery [4]. The right panel of Fig.2 demonstrates that a high-SNR glitch ($\sim$1544) induces significant biases in the recovered MBHB parameters—the total mass M_t, mass ratio q, and component spins a_1 and a_2—when not included in the model. Incorporating glitches in a joint fit with MBHBs removes these biases and restores accurate parameter recovery[4].

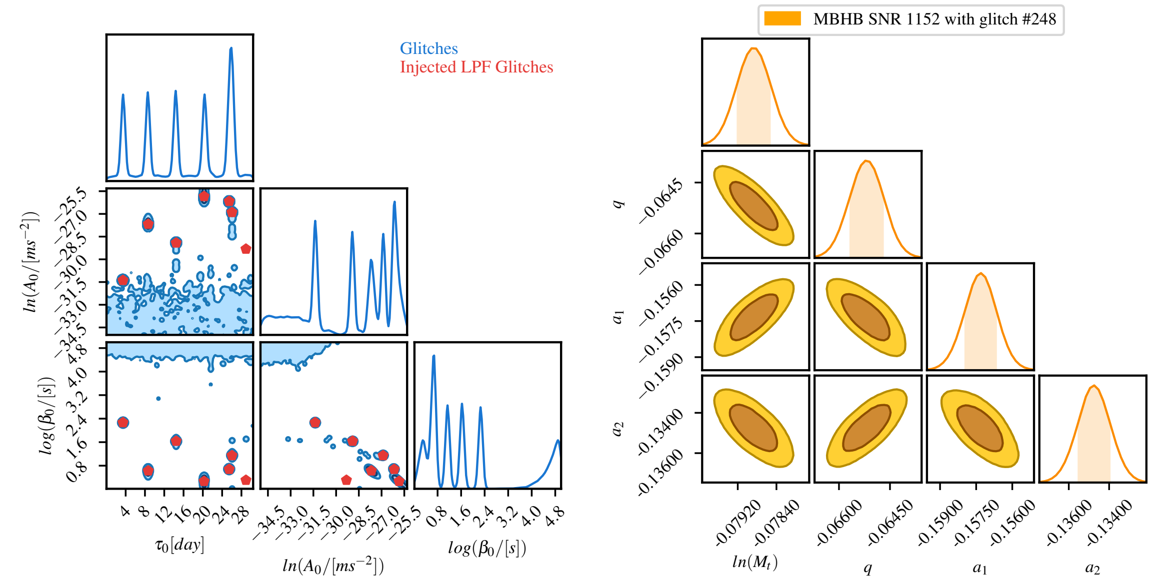

Figure 2: Left: posterior distributions of the shapelet parameters for seven injected LPF-like glitches, fitted using first-order shapelets. Multiple injected glitches are recovered simultaneously, while low-SNR components remain consistent with noise. Right: Posterior distributions of the intrinsic parameters of an MBHB with SNR 1152 in the presence of a glitch with SNR $\sim$1544, where we have shifted the parameters relative to the injected values. The fact that the posteriors do not peak at 0 illustrates the bias introduced when the glitch is not modelled. These plots are taken from Ref.[4].

We further validated the pipeline on a simplified version of the "Spritz" dataset [3], with Galactic binaries and data gaps removed. During the search phase (Fig. 3), the dataset is divided into one-day segments. Broad priors are adopted for the glitch parameters, extending beyond the ranges observed in LPF, to remain as agnostic as possible. In each segment, we first fit for noise and then test for the presence of a GW or instrumental signal, using the in-model and out-of-model moves described in REF [4]. If the majority of RJ-MCMC samples favour the no-signal hypothesis, the segment is discarded; otherwise, it is flagged for further analysis, and the Bayes factors between the MBHB and glitch hypotheses are computed. The parameter estimation phase then refines the glitch, MBHB, and noise parameters by performing joint estimation of instrumental noise and astrophysical signal.
This search and parameter estimation pipeline successfully detects and models the three glitches present in the data (Fig. 1) while recovering both the instrumental noise and the MBHB parameters without bias. The methods for glitch search, identification, and parameter estimation developed in this work will form the foundation of the glitch-handling module in Erebor [12], complementing existing modules for MBHBs, Galactic binaries, and noise/foreground fitting. These tools will be refined through improved sampling strategies and updated noise models, with future extensions to address instrumental non-stationarities such as data gaps and the Galactic foreground, enabling the analysis of realistic data sets like "Spritz." While our study is based on LPF-derived glitch distributions and does not encompass all MBHB populations, the shapelet approach remains flexible enough to capture a broad range of glitch morphologies. Given the technological continuity between LPF and LISA, especially in the GRS, the results presented here provide a realistic benchmark for the expected performance of glitch mitigation in LISA. All code used in this work is open-source and can be accessed at https://github.comMartyMuratore/artifacts.

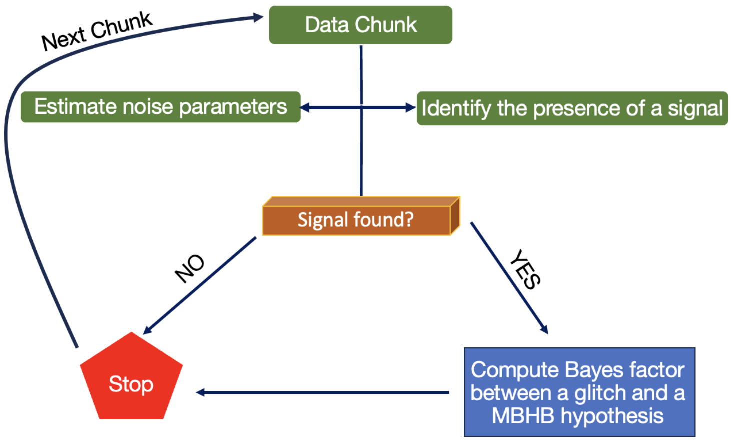

Figure 3: Schematic view of the search estimation algorithm.

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 Sangër, Alessandro Santini, Lorenzo Speri, Alice Spadaro, Mauro Pieroni and Sebastian Völkel. 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 the 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.

References

[1] Armano M, Audley H, Baird J, Binetruy P, Born M, Bortoluzzi D, Castelli E, Cavalleri A, Cesarini A, Chiavegato V, Cruise A M, Dal Bosco D, Danzmann K, De Deus Silva M, Diepholz I, Dixon G, Dolesi R, Ferraioli L, Ferroni V, Fitzsimons E D, Freschi M, Gesa L, Giardini D, Gibert F, Giusteri R, Grimani C, Grzymisch J, Harrison I, Hartig M S, Heinzel G, Hewitson M, Hollington D, Hoyland D, Hueller M, Inchauspé H, Jennrich O, Jetzer P, Johlander B, Karnesis N, Kaune B, Korsakova N, Killow C J, Lobo J A, López-Zaragoza J P, Maarschalkerweerd R, Mance D, Martín V, Martin-Polo L, Martin-Porqueras F, Martino J, McNamara P W, Mendes J, Mendes L, Meshksar N, Nofrarias M, Paczkowski S, Perreur-Lloyd M, Petiteau A, Plagnol E, Ramos-Castro J, Reiche J, Rivas F, Robertson D I, Russano G, Sala L, Sarra P, Slutsky J, Sopuerta C F, Sumner T, Texier D, Thorpe J I, Vetrugno D, Vitale S, Wanner G, Ward H, Wass P, Weber W J, Wissel L, Wittchen A, Zanoni C and Zweifel P (LISA Pathfinder Collaboration) 2022 *Phys. Rev. D* **106**(6) 062001 URL `https://link.aps.org/doi/10.1103/PhysRevD.106.062001`

[2] Karnesis N, Katz M L, Korsakova N, Gair J R and Stergioulas N 2023 *Monthly Notices of the Royal Astronomical Society* **526** 4814–4830 ISSN 1365-2966 URL `http://dx.doi.org/10.1093/mnras/stad2939`

[3] Le Jeune M, Babak S, Baghi Q, Bayle J B, Castelli E and Korsakova N 2022 Lisa data challenge: Spritz lISA-LCST-XXX-TN-001

[4] Muratore M, Gair J, Hartwig O, Katz M L and Toubiana A 2025 *Phys. Rev. D* **112**(6) 063041 URL `https://link.aps.org/doi/10.1103/1sj2-219n`

[5] Sala L 2023 *Residual test mass acceleration in LISA Pathfinder: in-depth statistical analysis and physical sources* Ph.D. thesis Trento U.

[6] Baghi Q, Korsakova N, Slutsky J, Castelli E, Karnesis N and Bayle J B 2022 *Phys. Rev. D* **105** 042002 (*Preprint* 2112.07490)

Journal of Physics: Conference Series **3177** (2026) 012112 doi:10.1088/1742-6596/3177/1/012112

[7] Tinto M and Dhurandhar S V 2014 *Living Reviews in Relativity* **17** 6 ISSN 1433-8351 URL `https://doi.org/10.12942/lrr-2014-6`

[8] London L, Khan S, Fauchon-Jones E, García C, Hannam M, Husa S, Jiménez-Forteza X, Kalaghatgi C, Ohme F and Pannarale F 2018 *Phys. Rev. Lett.* **120** 161102 (*Preprint* `1708.00404`)

[9] Katz M, Karnesis N and Korsakova N 2023 mikekatz04/eryn: first full release URL `https://doi.org/10.5281/zenodo.7705496`

[10] Bayle J B and Hartwig O 2023 *Physical Review D* **107** ISSN 2470-0029 URL `http://dx.doi.org/10.1103/PhysRevD.107.083019`

[11] Vousden W D, Farr W M and Mandel I 2015 *Monthly Notices of the Royal Astronomical Society* **455** 1919–1937 ISSN 0035-8711 URL `https://doi.org/10.1093/mnras/stv2422`

[12] Katz M L, Karnesis N, Korsakova N, Gair J R and Stergioulas N 2024 An efficient gpu-accelerated multi-source global fit pipeline for lisa data analysis (*Preprint* `2405.04690`) URL `https://arxiv.org/abs/2405.04690`