Challenges in Producing Final Products and L3 Data for Galactic Binaries in the LISA Mission.

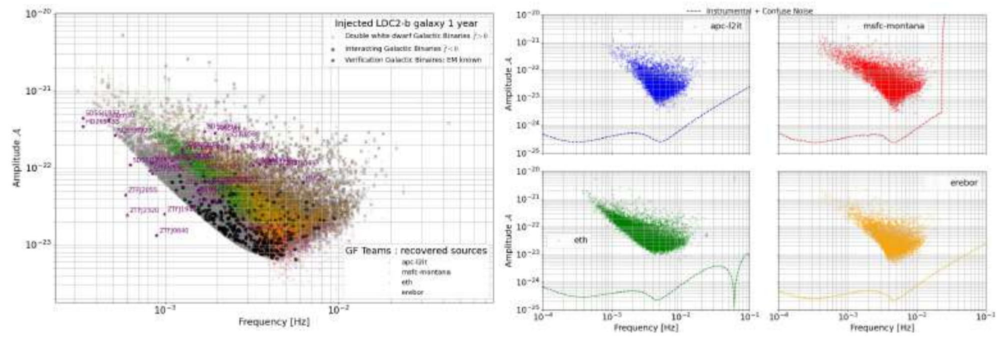
Guillaume Boileau, Astrid Lamberts, Martin Vannier, Christophe Ordenovic, Mathias Schultheis, Philippe Berio

guillaume.boileau@oca.eu Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, 06304 Nice, France

1. Introduction

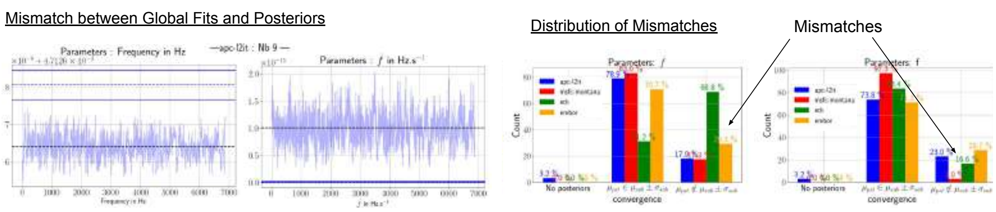
For the LISA mission, the creation of final products is a crucial step in the data processing process. Multiple global fit (L2) pipelines will identify and fit the same sources in different ways, and the goal is to merge these into a single comprehensive catalog. This involves comparing data from global fits, consolidating the data, and statistically validating it. Several global fit algorithms have distinct properties, resulting in varied submissions results to the LISA Data Challenge. The challenge lies in producing final products that encompass all observed sources while establishing connections between Global fits and providing all necessary informatin for scientific interpretation in an easily accessible way. Furthermore, algorithms capable of assessing the quality of adjustments and ensuring convergence have been developed as part of this study. To achieve this, we propose a preliminary protocol that outlines the essential steps needed to produce the final products. We have analyzed Sangria [1] data challenge for only Galactic binaries. Additionally, we propose an algorithm to merge L2 Global Fits [2, 3, 4, 5] and validate them against injected data.

2. Injection and Submissions



The inputs are from the LISA Data Challenge Sangria V2 [1], which includes both fully specified and blind datasets with simulated waveforms and Gaussian noise from millions of Galactic white dwarf binaries and merging massive black-hole binaries. The data also incorporates LISA noise produced using LISACode to generate "TDI-1.5" observables X, Y, Z. The submissions are the products of the global fit L2 [2, 3, 4, 5], estimating individual signals from the dataset and submitting a table with the parameters of each observed source and the posteriors for each of the sources.

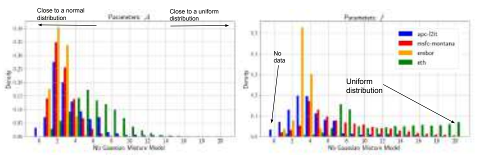
3. Convergence check| L3 preprocessing



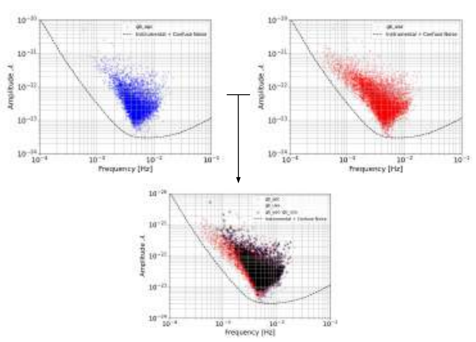
Before proceeding to the L3 pipeline, it is crucial to ensure the proper convergence of the global fits. To achieve this, we examine the consistency between the submitted L2 catalog and the posteriors. On the left, mismatches and overlap issues between Global Fits and posteriors are observed. The black and blue dotted lines shows respectively the trace mean and the quoted value of the parameter, and solid lines 5σ interval, which do not correspond to the trace of the GF. On the right, the distribution of the various Global Fits under study is depicted. A significant portion of L2 entries exhibit mismatches for at least one parameter.

4. Gaussianity Test

The Gaussian Mixture Model (GMM) is used to determine the number of Gaussian components for each chain from global fits to assess the convergence of each Global Fit. This test is crucial because we do not expect the data to be Gaussian. If many Gaussian components are found, it indicates a more uniform distribution, suggesting that the fit is poor and the model may not be suitable. GMM is a probabilistic model that represents data as a mixture of K Gaussian distributions, each with a mean vector $\boldsymbol{\mu}_k$, a covariance matrix $\boldsymbol{\Sigma}_k$, and a mixing coefficient π_k. The Probability Density Function (PDF) of a GMM is given by $p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$



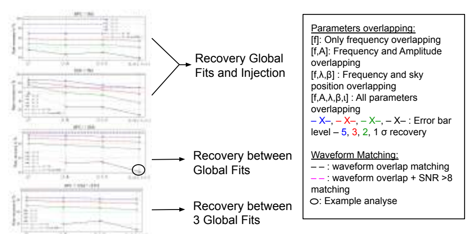
5. Proposal of fusion L3 Algorithm

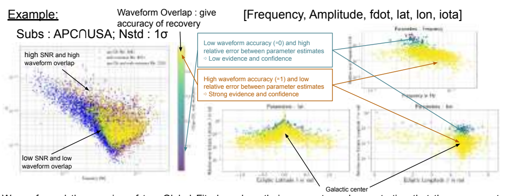


The Global Fits merging method integrates parameter information from two Global Fits, assessing overlap based on specified standard deviation levels. It aims to merge these Global Fits into a unified catalog, ensuring each candidate from one catalog overlaps optimally with candidates from the other. This overlap, which represents the correlation between waveforms, also serves to prioritize candidates when multiple entries coincide. This facilitates a streamlined selection process within overlapping regions.

6. Results : A first comparison for GB

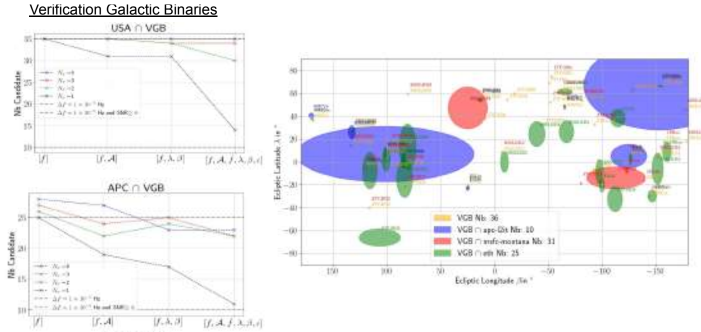
Result of recovery between APC and USA / recovery between Injection:





We performed the merging of two Global Fits based on their parameters, demonstrating that these parameters significantly impact the merging outcome. Additionally, the quality level (– X– Error bar level) also provides different levels of recovery, which have a significant impact. This level allows the construction of catalogs with an emphasis on either purity or completeness. This illustrates the current recovery capability enabled by this tool, allowing direct interaction with L2 to accurately assess the performance of various Global Fits.

Verification Galactic Binaries



Steps for an Frist transition L2 /L3:

- Convergence criteria :
  - Check Mismatch posterior/Global Fits
  - Check Priors
- Statistical test of Gaussianity
  - Global Quality check of the convergence
  - Comparison of normality distributions.
- Overlap and relative error between Global Fits
  - Comprehensive view of the recovery.
  - Quality assessment criteria.

7. Conclusions/Next step

1. Continuous integration with new Global Fits
2. Fusion catalogue with quality user variation
   - Purity vs completeness.
   - Adaptability to different LISA sources.
3. Developp Visualization tools in collaboration
4. Implement other sources

8. References

[1] Stas Babak et al.. Lisa data challenge 2a: Sangria. Technical report, 2020.

[2] Stefan H. Strub et al.. *arXiv e-prints*, page arXiv:2403.15318, March 2024.

[3] Tyson B. Littenberg and Neil J. Cornish. *PRD*, 107(6):063004, March 2023.

[4] Michael L. Katz et al.. *arXiv e-prints*, page arXiv:2405.04690, May 2024.

[5] APC. In preparation. 2024.

GB thanks the Centre national d'études spatiales (CNES) for support for this research and the LISA Data Challenge (LDC) for providing input data and continuous feedback on this work.