Extreme mass-ratio inspirals for Mojito Dataset: Simulating and Validating a Relevant Catalog of EMRIs for LISA

Bert Depoorter (KU Leuven | Royal Observatory of Belgium)

Supervisors: Dr. Henri Inchauspé (KU Leuven), Dr. Ollie Burke (UGlasgow), Manuel Piarulli (L2IT Toulouse)

EMRIs and LISA

- LISA: space-based GW observatory; sensitive to 10^{-4}–10^{-1} Hz GWs.
- EMRIs: stellar-mass BH (m) orbiting SMBH (M), $\varepsilon = \frac{m}{M} \in [10^{-8}, 10^{-4}]$

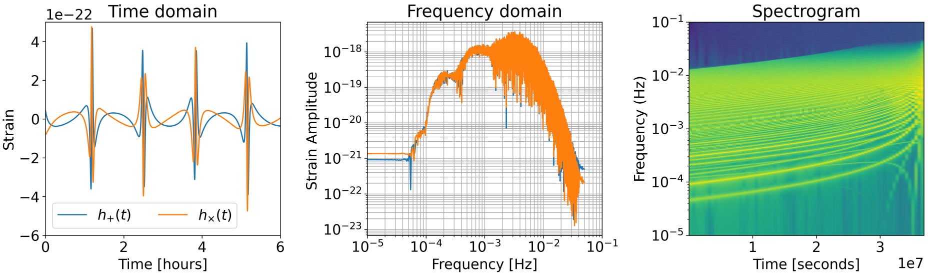


Figure: Example EMRI waveform in time, frequency and time-frequency domain. Created with `FastEMRIWaveforms` [1].

Global-fit challenge

Real LISA observatory:

- Simultaneous, overlapping signals from $> 10^4$ sources.
- Non-stationary instrumental noise.
- Reconstructing GW measurements from raw telemetry via TDI.

Data processing and source inference infrastructure needed.
To prepare: **realistic simulation needed!** LISA common Dataset 1: **Mojito**

- Galactic binaries
- Massive black hole binaries
- EMRIs
- Stellar origin black hole binaries
- Instrumental noise

Light version: planned for release end of 2025.

EMRIs: opportunities and challenges

Characteristics:

- 10^3–10^4 orbital cycles $\Rightarrow$ high-precision parameter estimation.
- Rich phenomenology: high spins, eccentric orbits, strong-field effects.
- Complex waveforms: many harmonics, sensitive to intrinsic parameters.

Waveform model: `KerrEccentricEquatorial` in `FastEMRIwaveforms` [1].

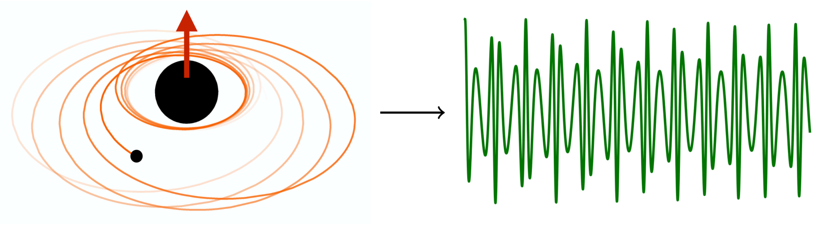


Figure credit: L. Speri [2].

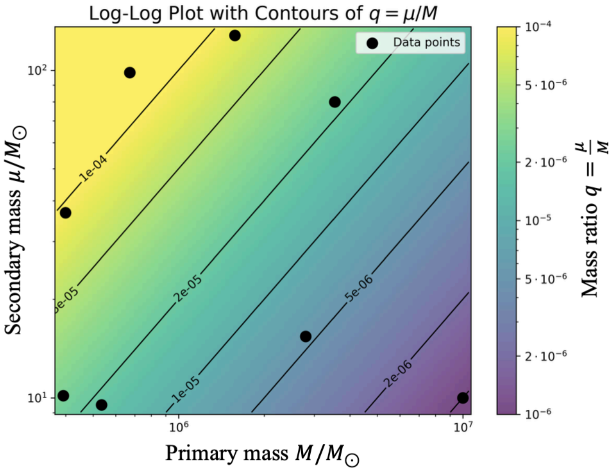
Selection of source parameters

Wishlist:

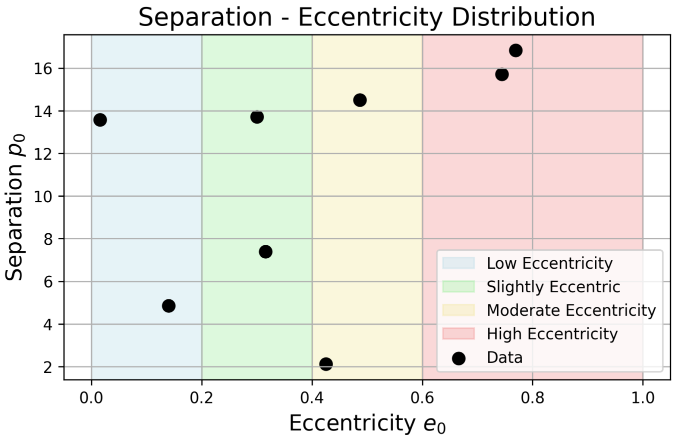
- 8 EMRIs covering range of masses / mass-ratios.
- Both pro-/retrograde spins.
- Continuous and plunging sources.
- Eccentricities up to $e \sim 0.9$.
- Uniform sky distribution.
- SNR threshold > 30.

Starting point: Catalog from *Babak et al.*, [3].
Selection based on SNR and intrinsic parameters.
Tweaked parameters to meet requirements.

Mass-ratio distribution



Eccentricity - separation distribution



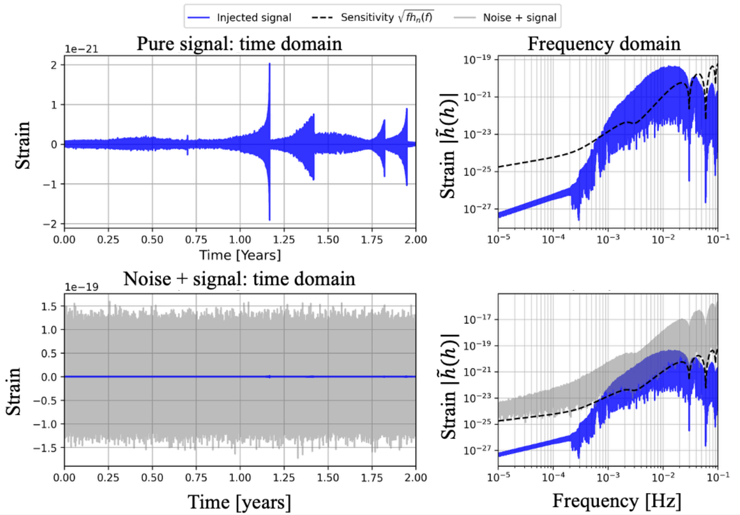
Instrumental simulation - implementation

- Generate detector-frame strain
- Project on LISA response: returns single link time series
- Post-process single link responses to TDI variables XYZ (2nd generation)

Characteristics:

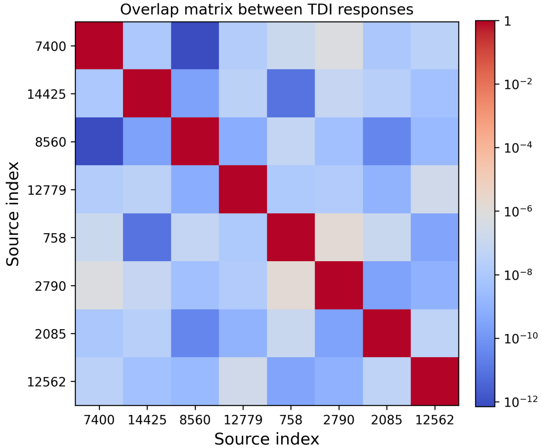
- $T_{\mathrm{obs}} = 2$ years, cadence $\Delta t = 5$ s.
- Noise: gaussian & stationary, TM and OMS contributions [4].
- Spacecraft orbits: equal and static arm length orbits.

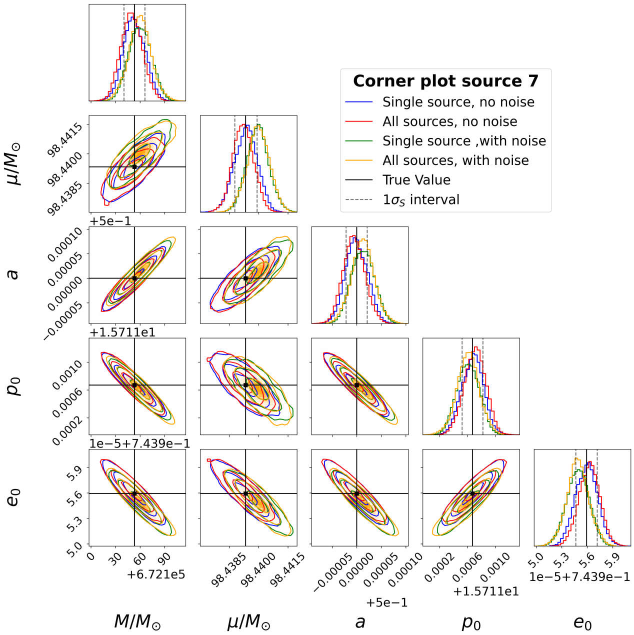
Clean vs. Noisy Signal



Results: Overlap matrix

Computed overlaps $\mathcal{O}(h_i, h_j) = \frac{<h_i|h_j>}{\sqrt{<h_i|h_i><h_j|h_j>}}$ across the 8-source catalog.




Corner plot source 7
Single source, no noise
All sources, no noise
Single source ,with noise
All sources, with noise
True Value
$1\sigma_S$ interval
$M/M_\odot$
$\mu/M_\odot$
a
p_0
e_0
+6.721e5
+5e−1
+1.5711e1
1e−5+7.439e−1

Estimated precision of different MCMC runs

Parameter	Parameter Precision σ/μ
$M/M_\odot$	$1.9919111e-05$
$\mu/M_\odot$	$5.6867135e-06$
a	$4.4838883e-05$
p_0	$9.9302374e-06$
e_0	$1.7070927e-06$
D_s/Gpc	$1.2862370e-02$
θ_S	$9.1194543e-03$
ϕ_S	$1.5574695e-03$
θ_K	$1.5409376e-02$
ϕ_K	$4.3824116e-03$
Φ_{ϕ_0}	$1.5778561e-02$
Φ_{r_0}	$1.3118491e-02$

Table: Parameter precision estimated from MCMC runs.

Recovered SNR from recovered parameters

- Signal-to-noise ratio (SNR): metric for detectability
- From posterior distribution: recover parameters
- Residual SNR: difference of signal left in data

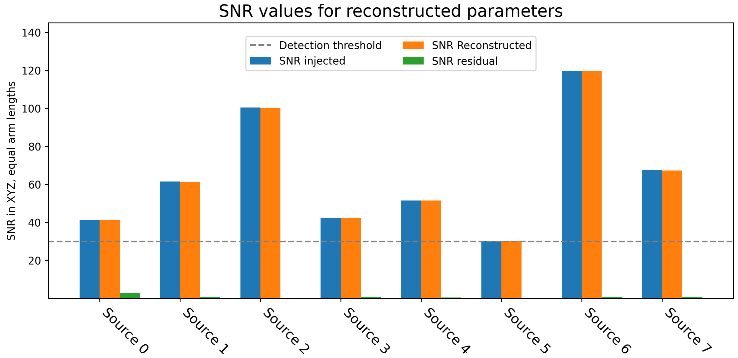


Figure: Injected vs. recovered SNR.

Last step: validation for final dataset production

Goal: validate numerical results and robustness of tools.

- Test behaviour and data processing over different sampling rates.
- Search for possible artifacts in data processing pipeline.
- Compare source selection metrics (SNR, overlaps) match actual simulation outputs.
- Consistency between fastlisaresponse and gw-response implementation: check for biases.
- Final check: rerun MCMC run with fast response model on actual dataset.

Conclusion

- Constructed a targeted EMRI mock dataset (part of CD1 Mojito) enabling realistic LISA-style tests.
- Demonstrated: parameter inference is possible provided
  - sufficient SNR.
  - well-behaved and known noise model.
  - restricted prior around true values to speed up MCMC convergence.
- Confirmed high precision expected from EMRI observation:
  - intrinsic parameters: $\sim 10^{-5}$ relative precision.
  - extrinsic parameters: $\sim 10^{-2}$ relative precision.
- Demonstrated: most Mojito EMRI sources are separable, orthogonal signals

Final step: Validate results with actual data pipeline.

References

[1] C. E. A. Chapman-Bird et al., *The Fast and the Frame-Dragging: Efficient waveforms for asymmetric-mass eccentric equatorial inspirals into rapidly-spinning black holes*, (June 11, 2025) arXiv:2506.09470 [gr-qc], http://arxiv.org/abs/2506.09470 (visited on 06/13/2025), pre-published.

[2] *EMRI Waveform Modelling Schedule session LISA Symposium 2024*, https://virtual.oxfordabstracts.com/event/5189/program?session=108076&s=2095 (visited on 06/22/2025).

[3] S. Babak et al., *Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals*, (Apr. 11, 2018) arXiv:1703.09722, http://arxiv.org/abs/1703.09722 (visited on 11/20/2024), pre-published.

[4] S. Babak et al., *LISA Sensitivity and SNR Calculations*, (Aug. 2, 2021) arXiv:2108.01167 [astro-ph], http://arxiv.org/abs/2108.01167 (visited on 01/27/2025), pre-published.