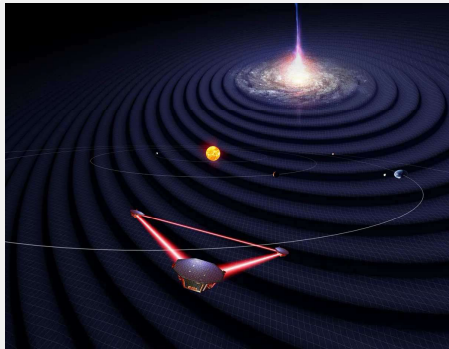# Disentangling and characterizing astrophysical GW signals with LISA
Modular global-fit pipeline for LISA data analysis

Senwen Deng*, Stanislav Babak, Maude Le Jeune, Sylvain Marsat, Éric Plagnol and Andrea Sartirana
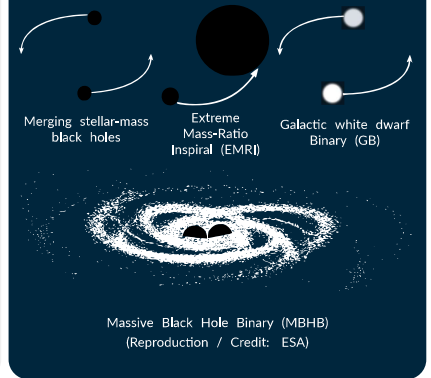*deng@apc.in2p3.fr

LISA constellation in the solar system
(Credit: University of Florida / Simon Barke / CC BY-SA 4.0)

What is LISA?

LISA is the acronym for Laser Interferometer Space Antenna, a space-based gravitational wave observatory.

- Collaboration between ESA, its member states and NASA.
- Constellation of three spacecrafts, exchanging laser beams, forming an equilateral triangle with 2.5 million km sides.
- Detect gravitational waves in the millihertz frequency band.

Astrophysical sources

Merging stellar-mass black holes

Extreme Mass-Ratio Inspiral (EMRI)

Galactic white dwarf Binary (GB)

Massive Black Hole Binary (MBHB)
(Reproduction / Credit: ESA)

Why is LISA data analysis challenging?

- LISA data is dominated by signal, making the noise level estimation difficult.
- Signals are long-lasting, and they overlap in time and frequency domains.
- All the Galactic binary signals cannot be resolved individually, and the unresolved ones form a confusion foreground.
- High dimensionality due to the large number of sources.
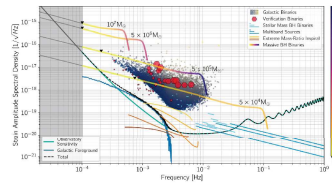
Primary LISA sources in the frequency-amplitude plane
(Credit: ESA / LISA Definition Study Report - ESA-SCI-DIR-RP-002)
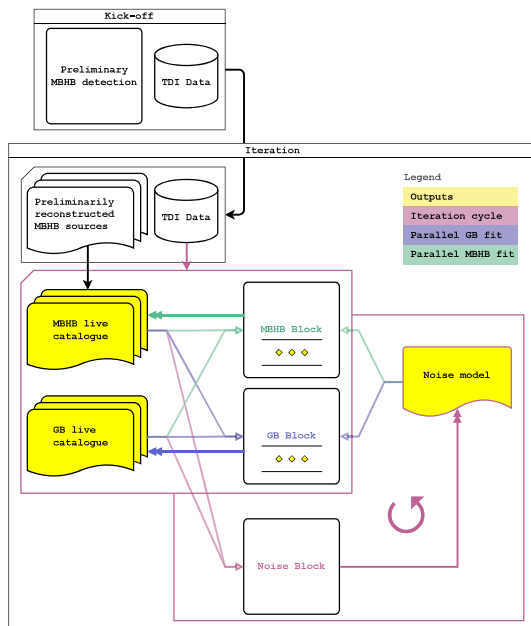
"Sangria" dataset
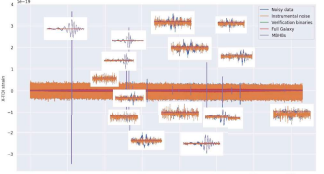
LISA Data Challenge 2 (aka "Sangria") training dataset (simulation of 15 MBHBs + 30 million GBs + stationary instrumental noise)
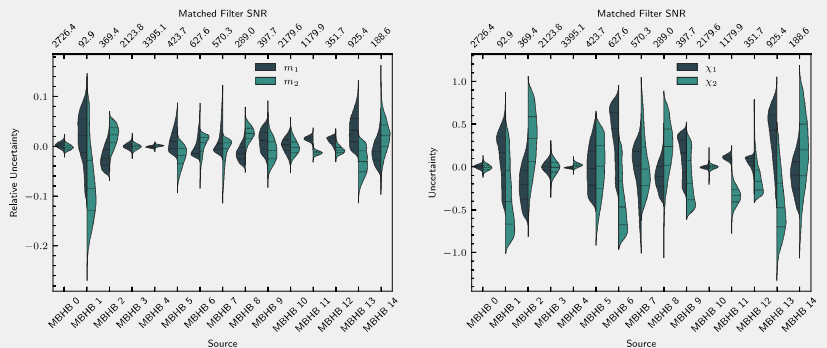
How to tackle the challenge?

We designed a modular framework applied to analyse the "Sangria" dataset. The framework relies on:

- The transitional nature of broadband loud MBHB signals
- Parallel tempering MCMC algorithm
- Block parallel Gibbs sampling technique
- Refinement through iterations
- Product space model selection
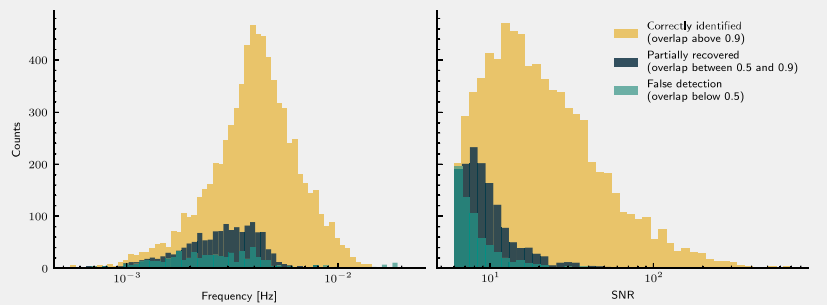- Accelerated likelihood computation with heterodyning

Schema of the pipeline

Our results

Posterior distribution for MBHB masses m_1, m_2 and spins χ_1, χ_2

- Correctly identified (overlap above 0.9)
- Partially recovered (overlap between 0.5 and 0.9)
- False detection (overlap below 0.5)

Detected GB counts as a function of frequency and SNR

Conclusion

- We developed a global-fit pipeline for LISA data analysis.
- We applied it to the "Sangria" training dataset and obtained promising results.
- Our pipeline is modular and can be adapted to future datasets.