Prototyping a Global-fit Pipeline for LISA

SENWEN DENG (APC)

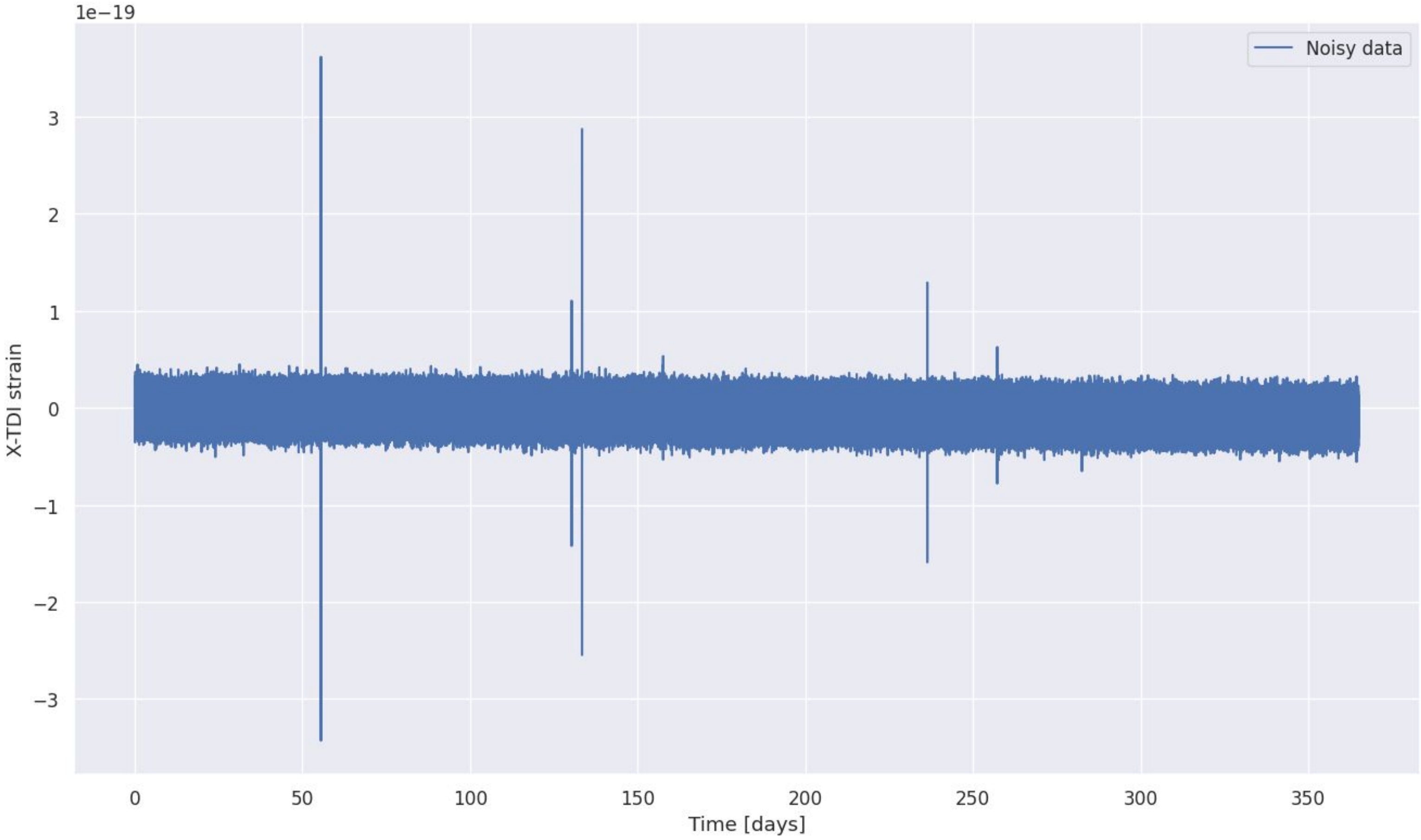
WITH

S. Babak, M. Le Jeune, E. Plagnol, A. Sartirana(APC)
S. Marsat (L2IT)

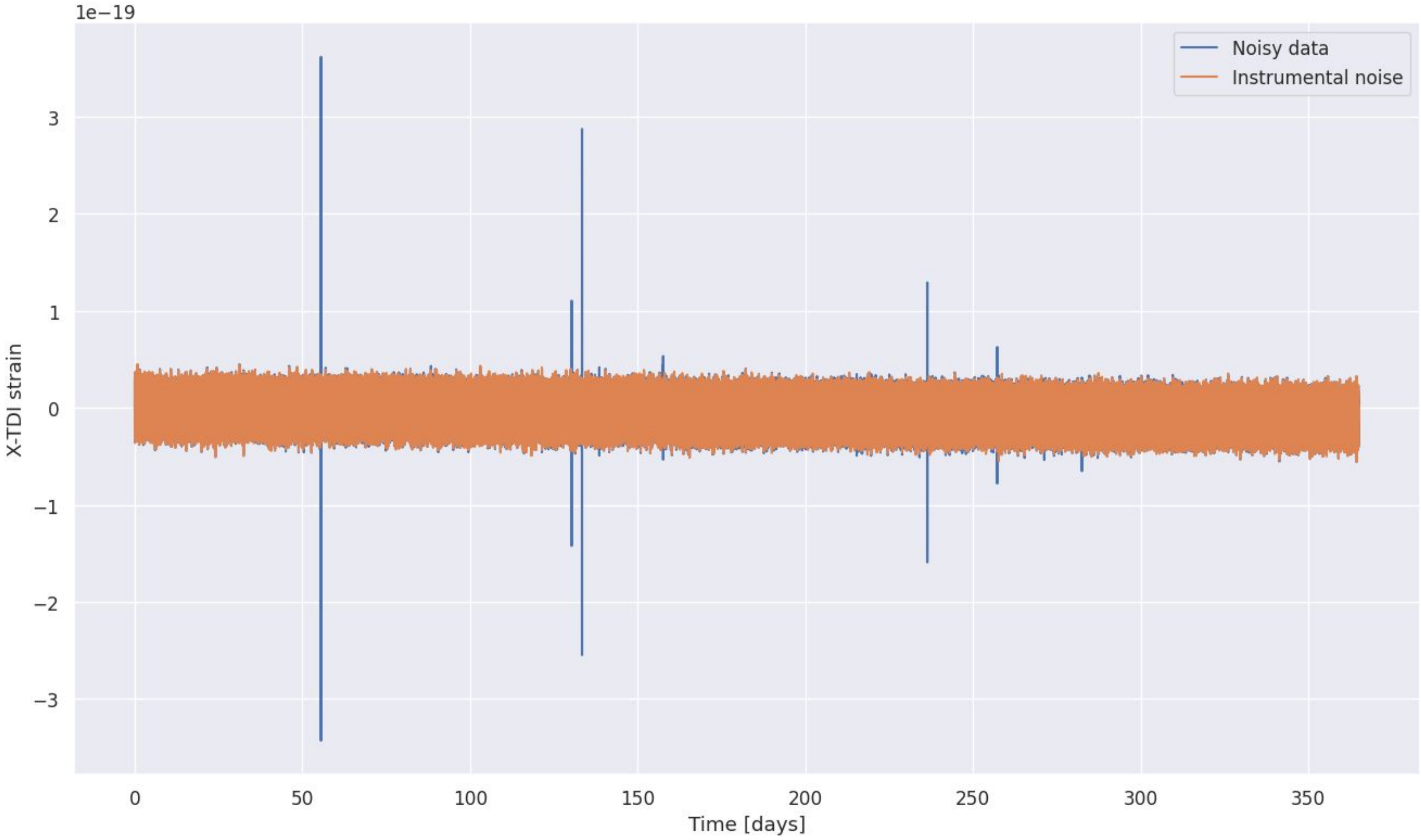
Septième assemblée générale du GdR Ondes Gravitationnelles
16–17 Oct 2023
LUTH, Observatoire de Paris, Meudon

Profile of LISA Data

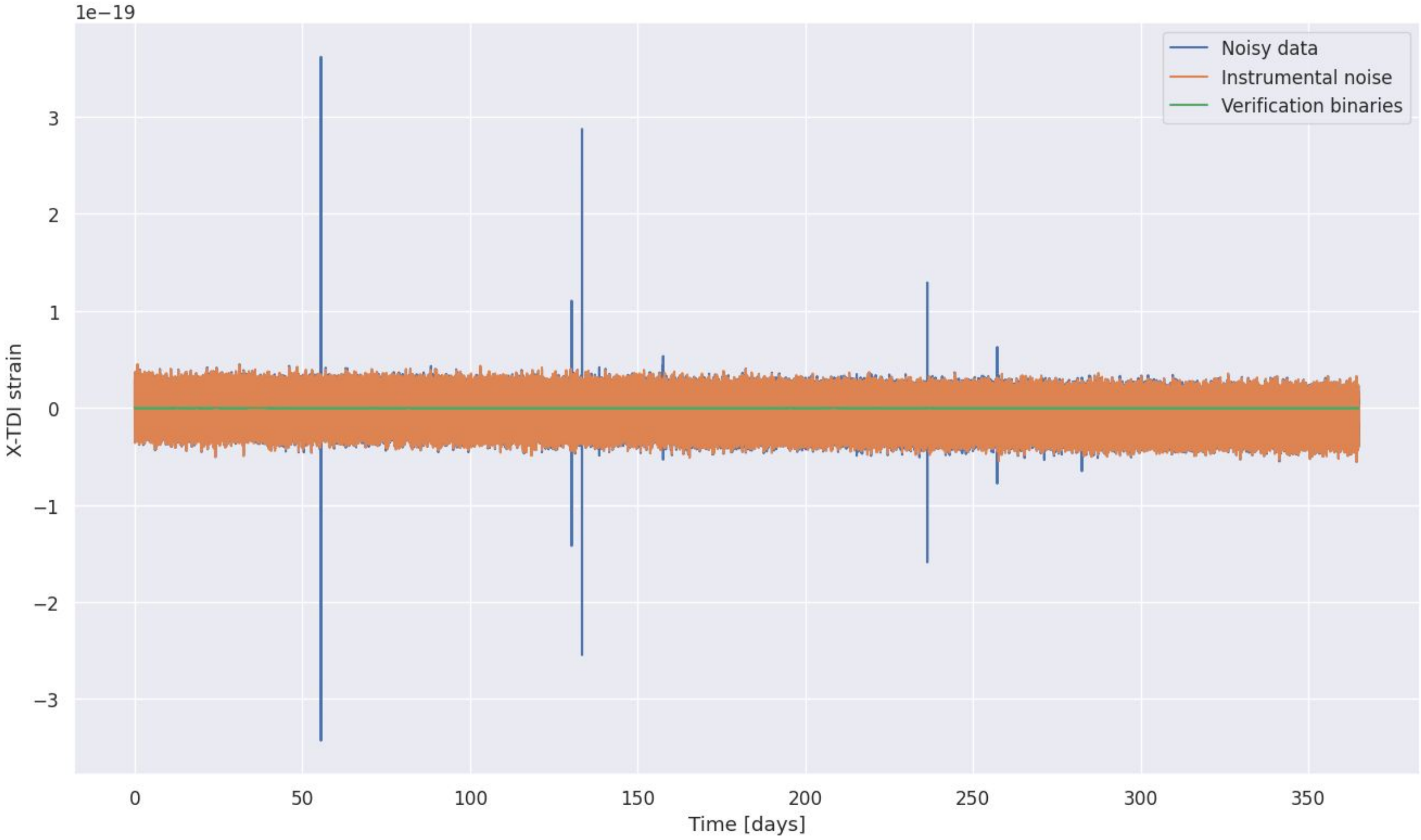
- Dominated by GW signals, all-sky all-time
- Many signals are long-lived (EMRI, GB) and overlapping in F&T



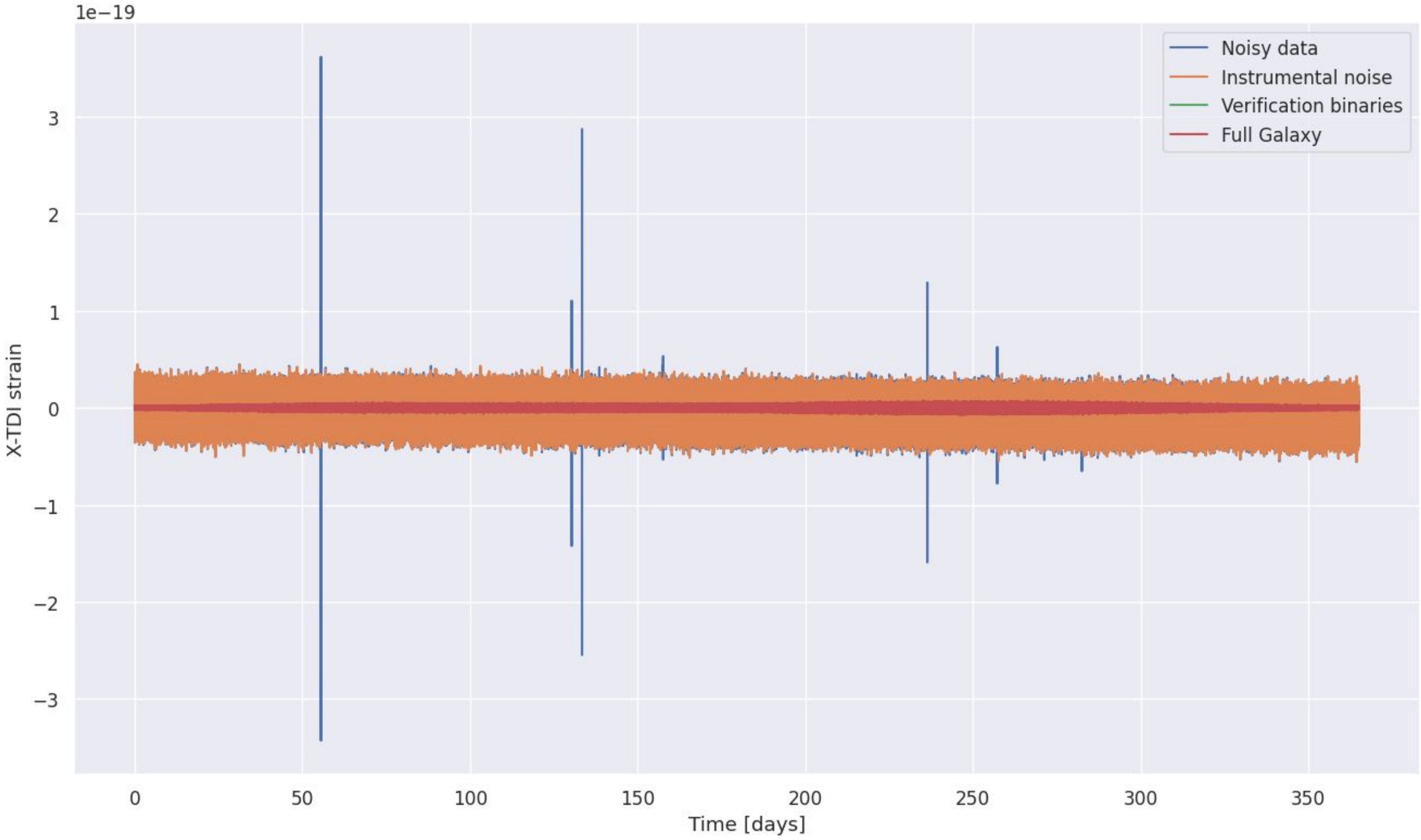
Simulated LISA Data: Sangria



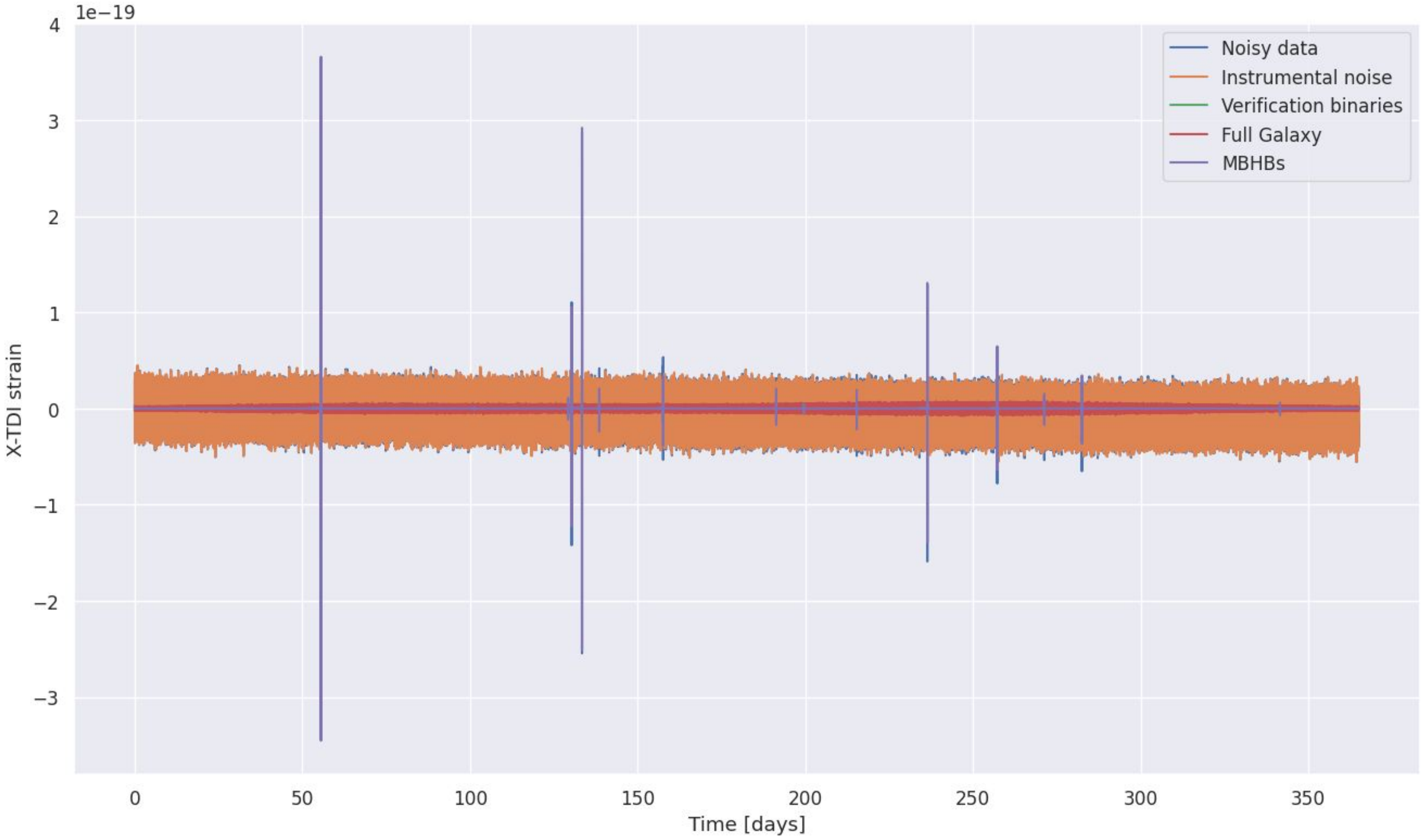
Simulated LISA Data: Sangria



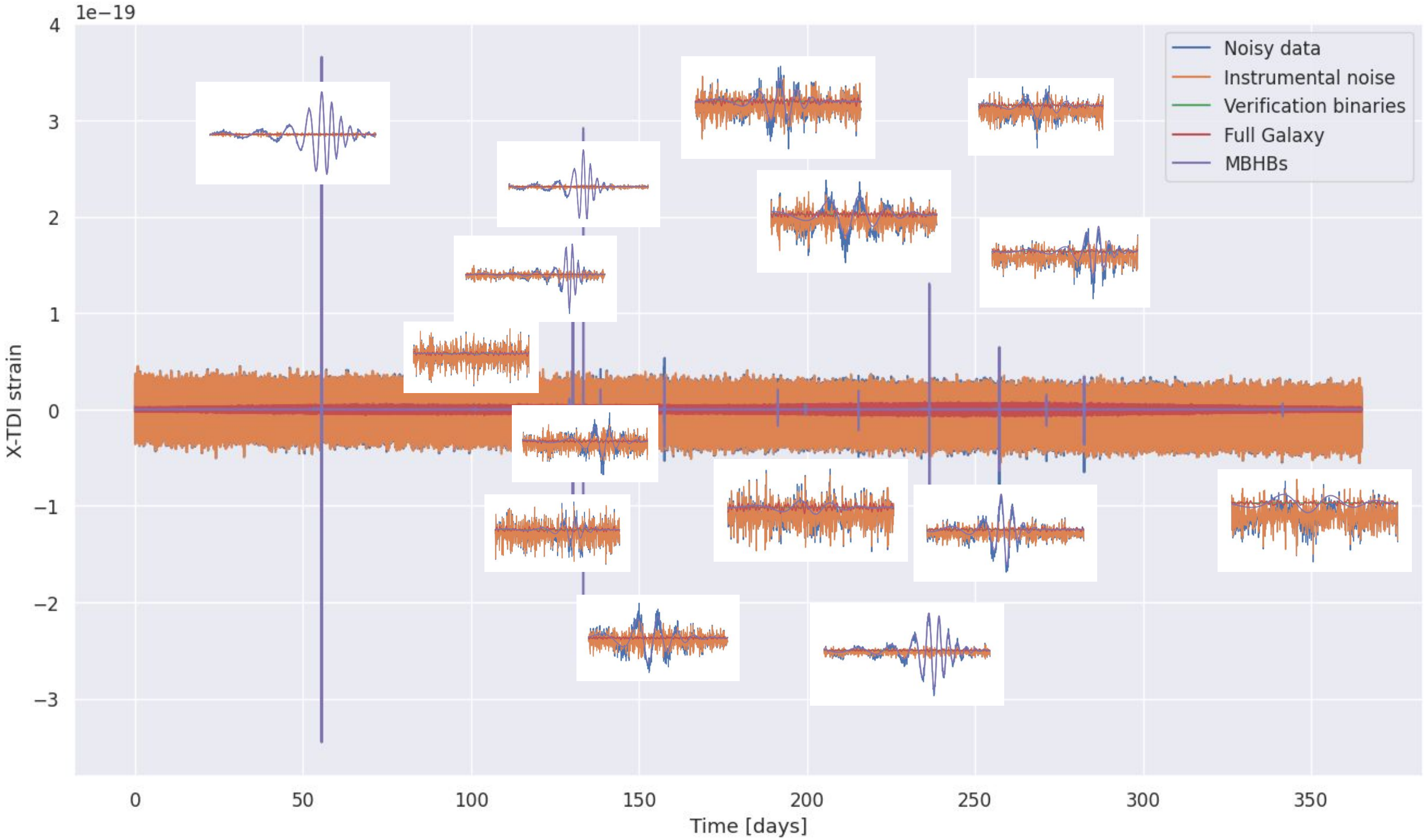
Simulated LISA Data: Sangria



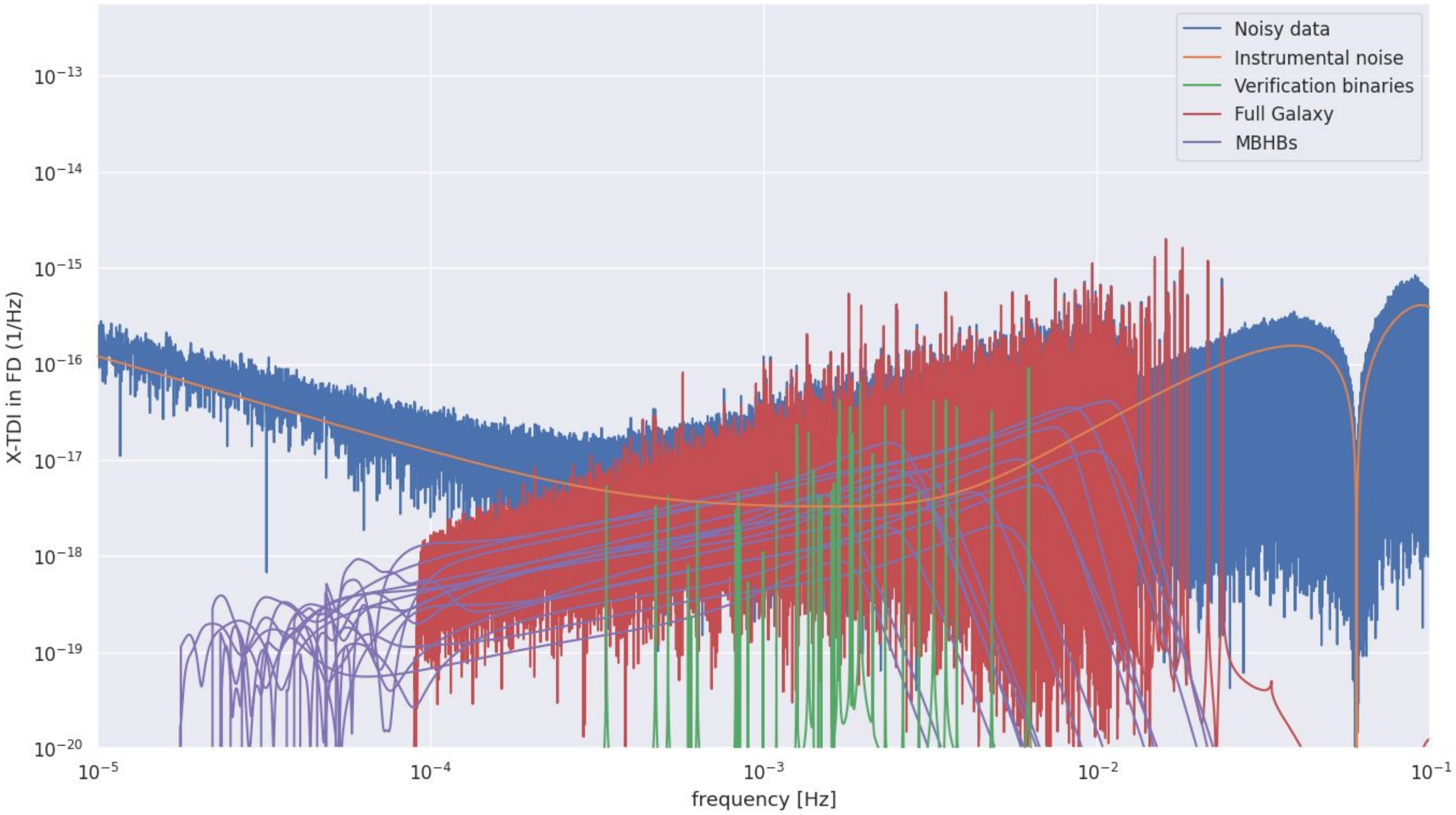
Simulated LISA Data: Sangria



Simulated LISA Data: Sangria



Simulated LISA Data: Sangria



Simulated LISA Data: Sangria

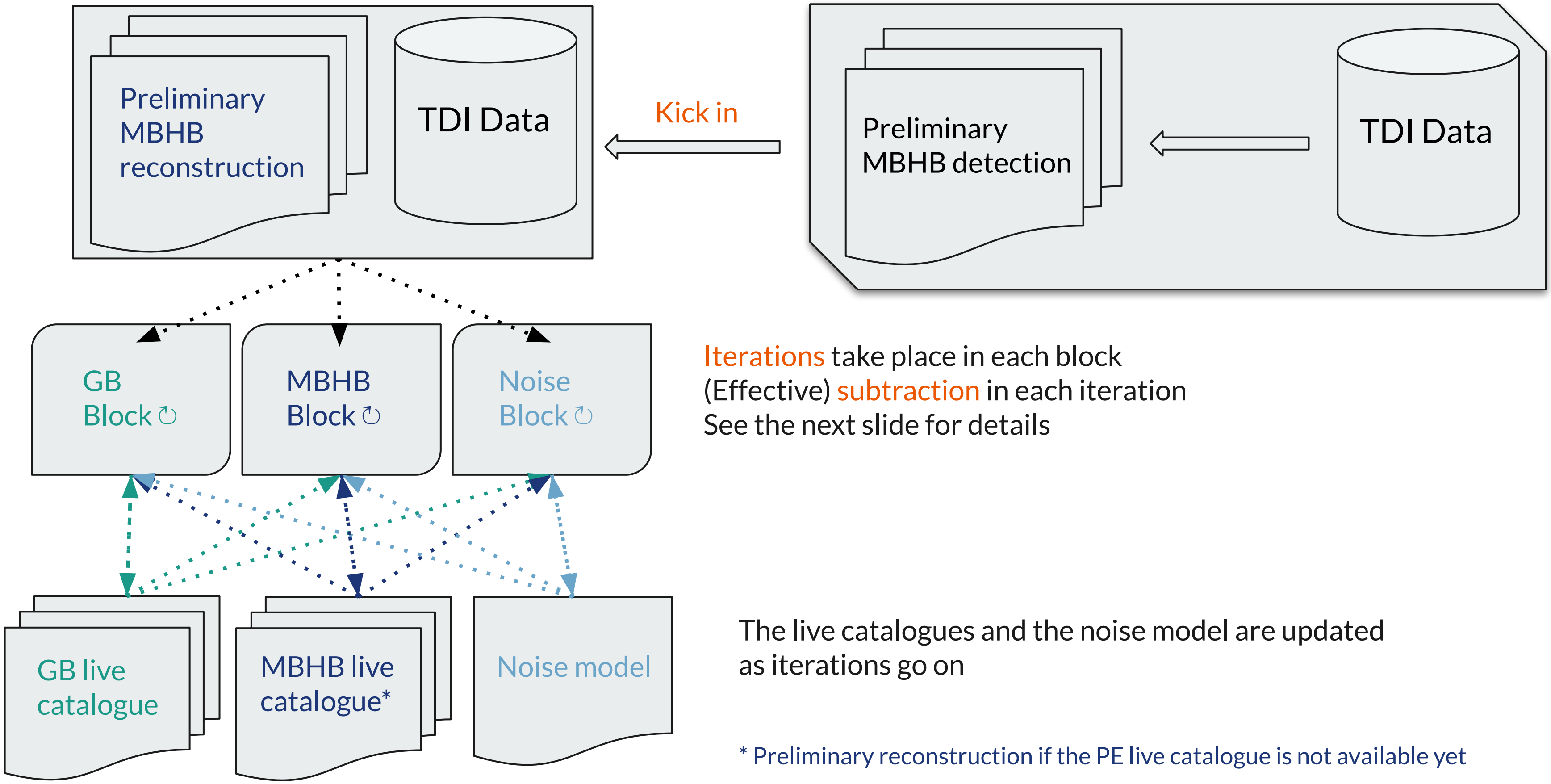
Profile of LISA Data

- Dominated by GW signals, all-sky all-time
- Many signals are long-lived (EMRI, GB) and overlapping in F&T
- Unresolved GW signals contribute to the noise budget
- Non-stationary noise: gaps, glitches
- Pioneer's problem: unknown event rate, unknown parameter distribution

Challenging!

Strategy of Global-fit

Keywords: kick-in, subtraction, iteration

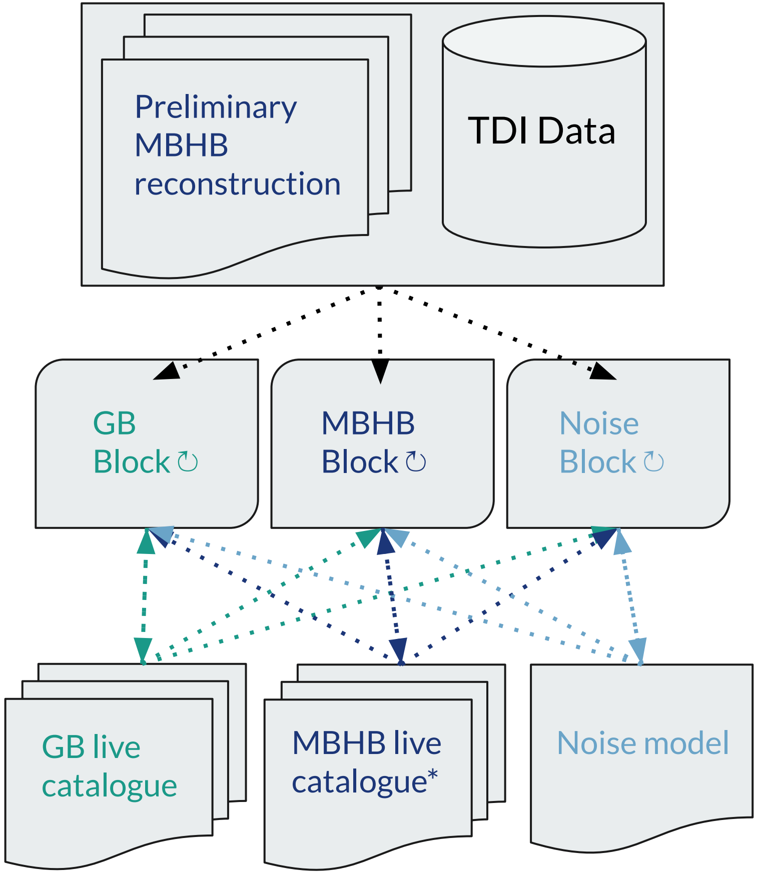


Iterations take place in each block
(Effective) subtraction in each iteration
See the next slide for details

The live catalogues and the noise model are updated as iterations go on

* Preliminary reconstruction if the PE live catalogue is not available yet

Strategy of Global-fit: prototype architecture



Each block iteration

1. subtracts unattended live catalogue signals from data
2. refines (MCMC)
3. updates the live catalogue/model

Specifics:

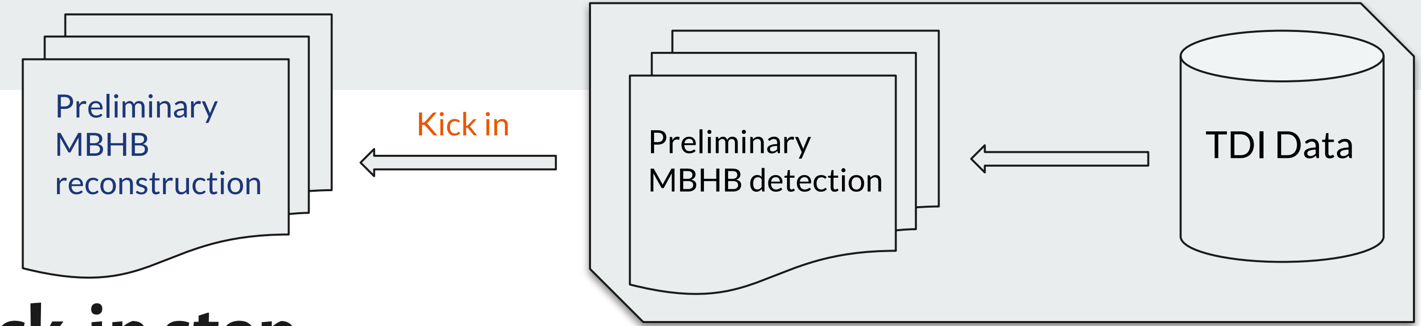
- There are up to thousands of jobs on an HTC cluster.
- The blocks can be async.
- MCMC Chains, plots, logs and debug information are stored at each iteration.

More blocks for other sources: SMBH, EMRI, …

The live catalogues and the noise model are updated as iterations go on

* Preliminary reconstruction if the PE live catalogue is not available yet

Strategy of Global-fit: prototype architecture



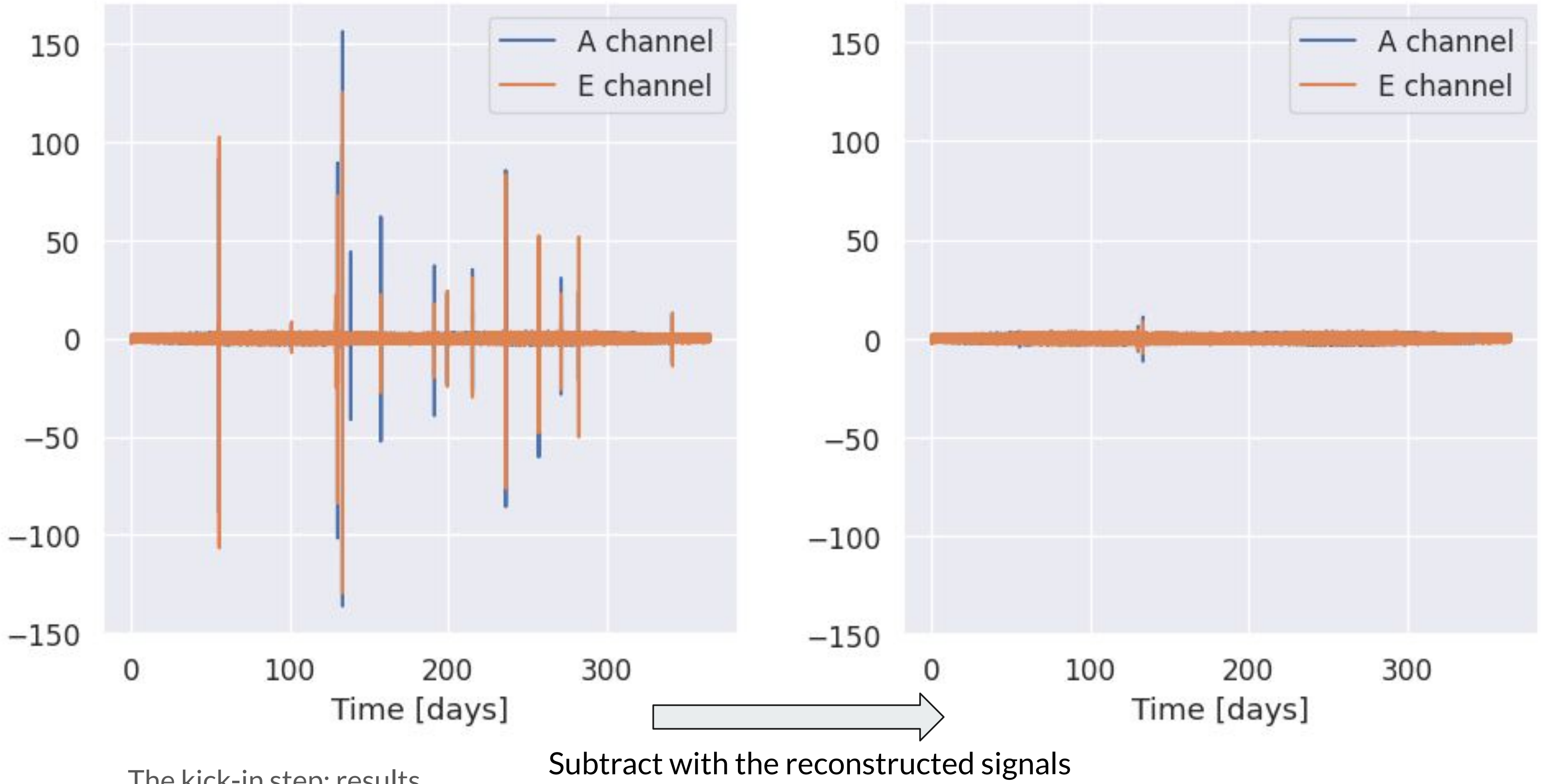
The kick-in step

We neglect the LISA motion and we assume the long-wavelength regime.

- Detect and reconstruct with a time-sliding F-statistic: log-likelihood ratio maximised over time of arrival (merger), distance, inclination, sky, initial phase, polarisation

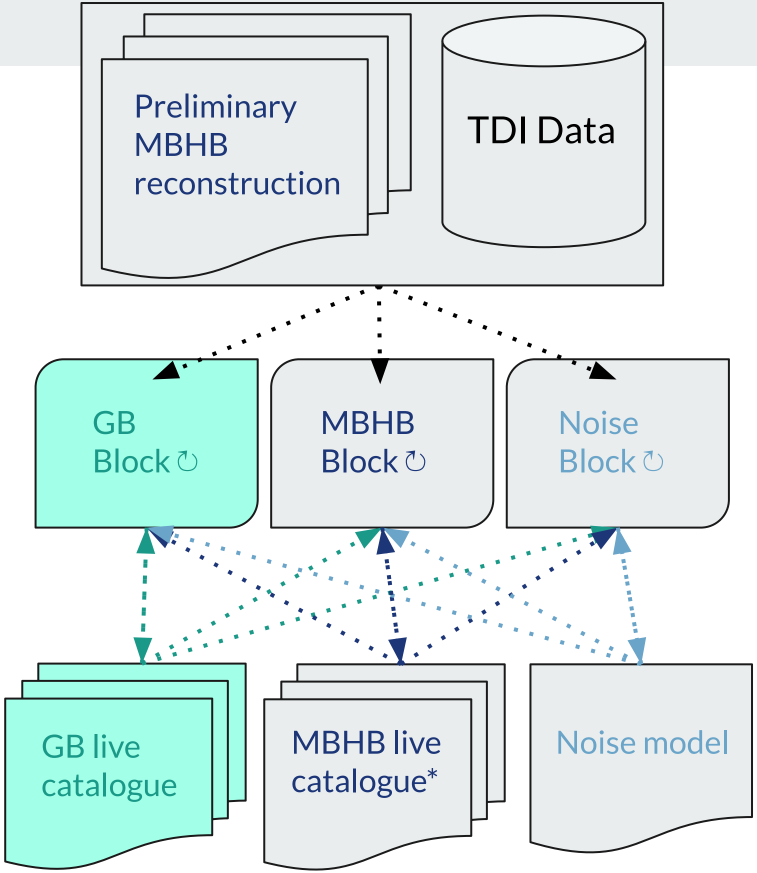
How to find the maximum fast? One possible way: mesh-refinement driven by Vegas

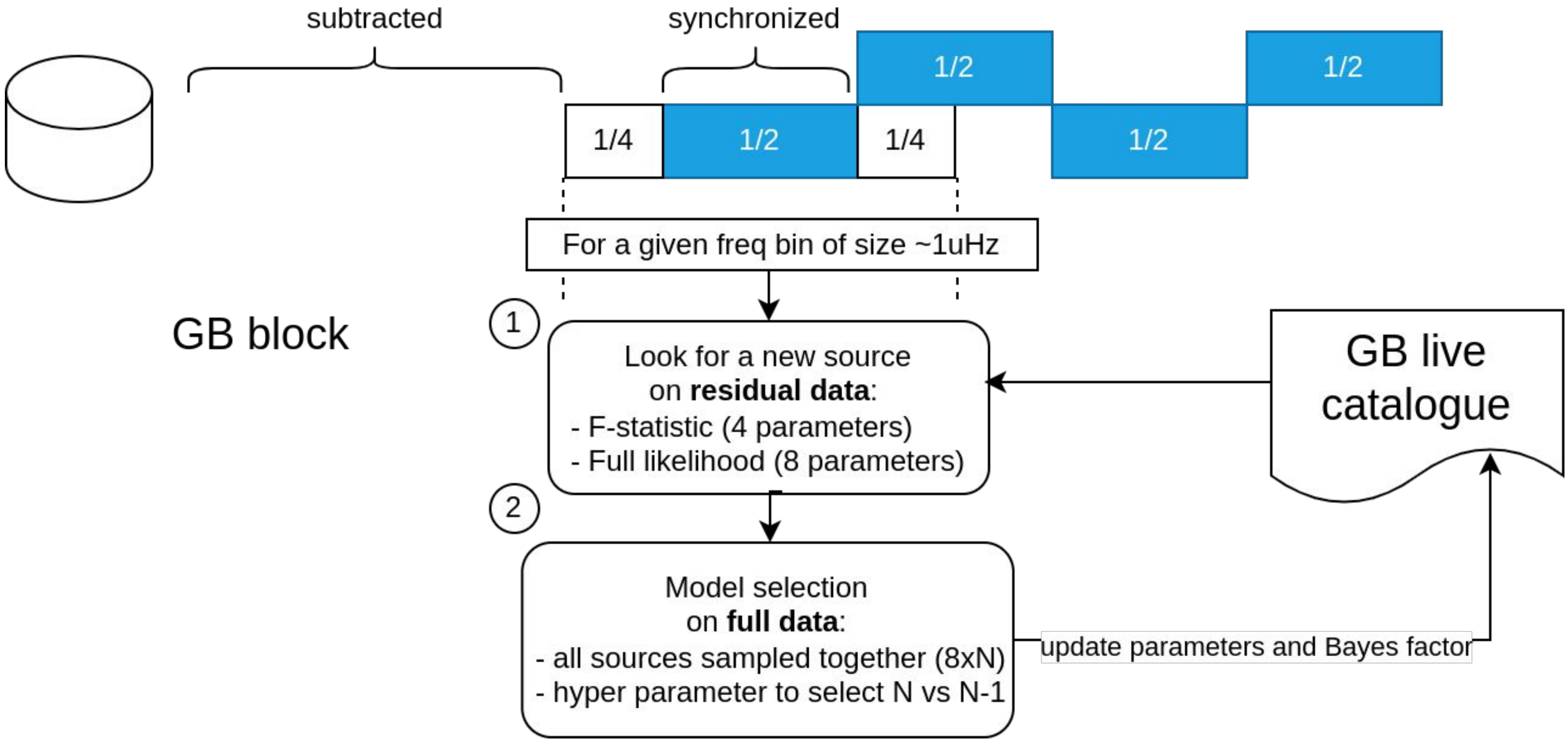
- Adapt the meshgrid by doing Monte-Carlo integrations
- Embarrassingly parallelizable



The kick-in step: results

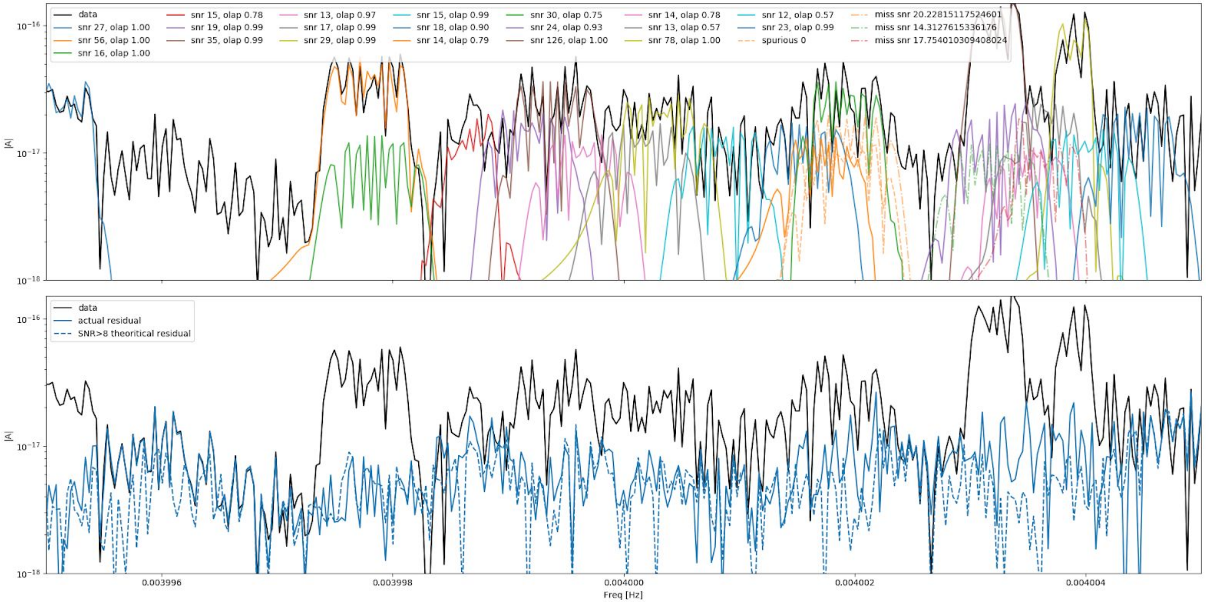
Dealing with the Galactic binaries



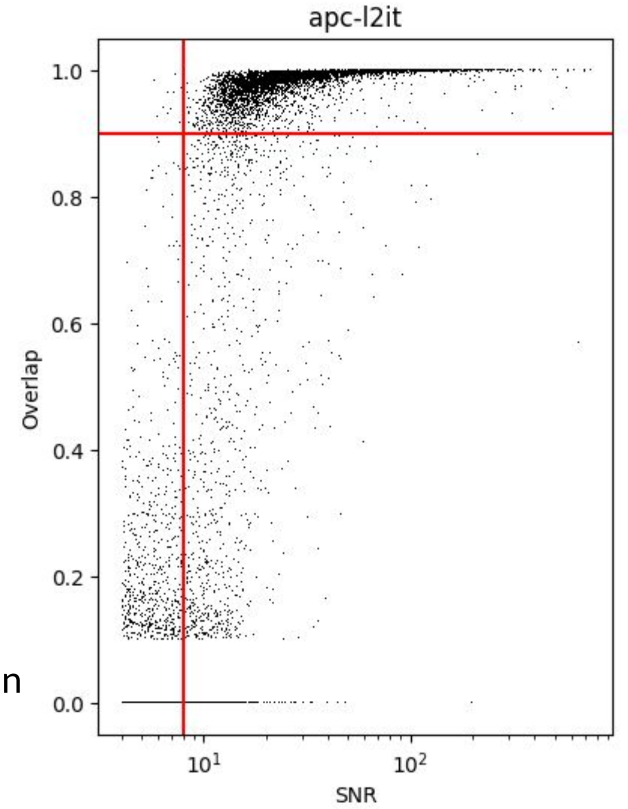


Dealing with the Galactic binaries

Preliminary results @4mHz



Submission



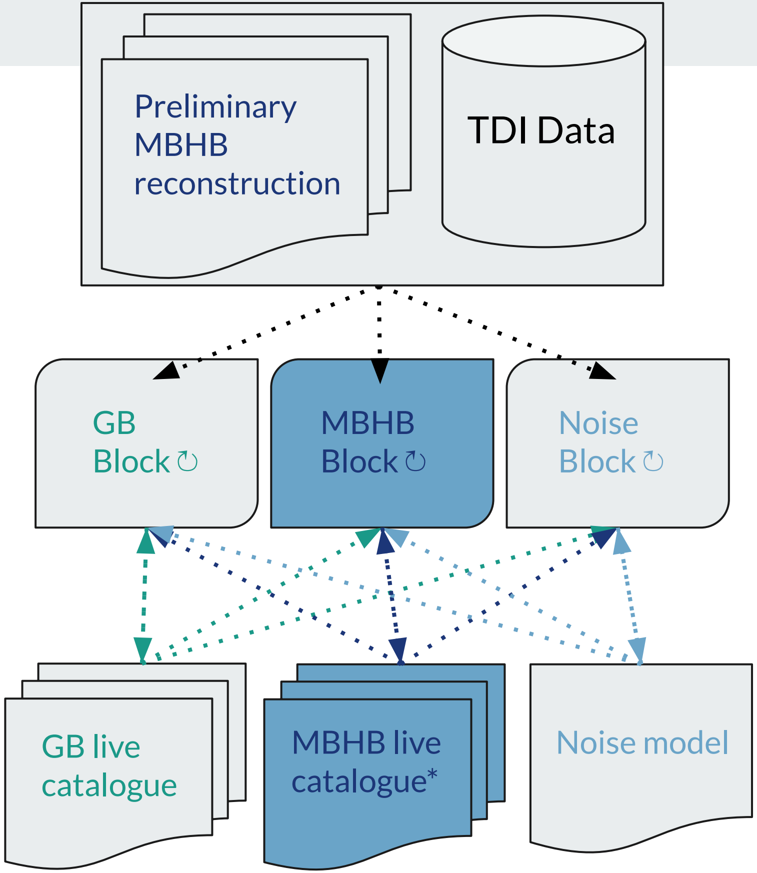
Dealing with the Galactic binaries: results

Refine MBHB PE

We start from the preliminary results of the kick-in step

- Reconstructed signal for heterodyning (Cornish & Shuman 2020)
- Initial points for the sampling

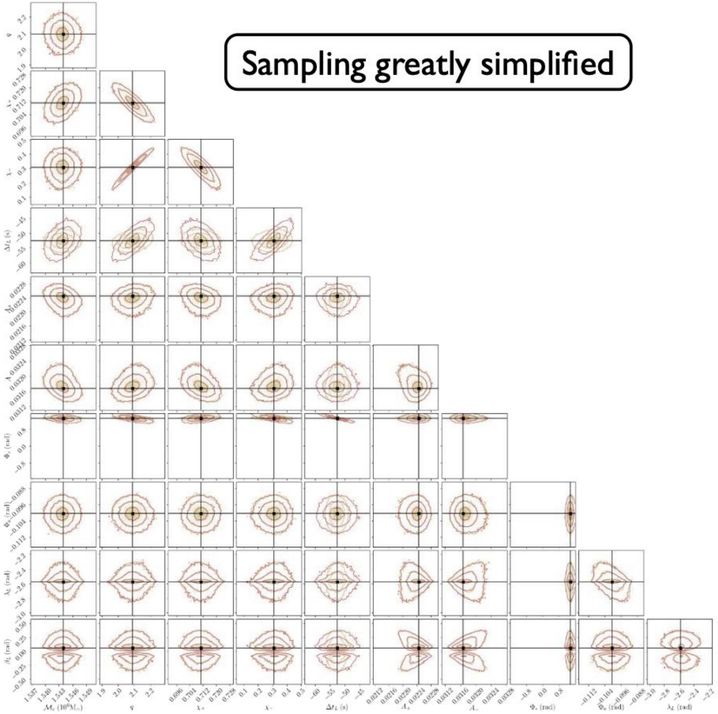
Parameter mapping is helpful (Marsat et al. 2021)



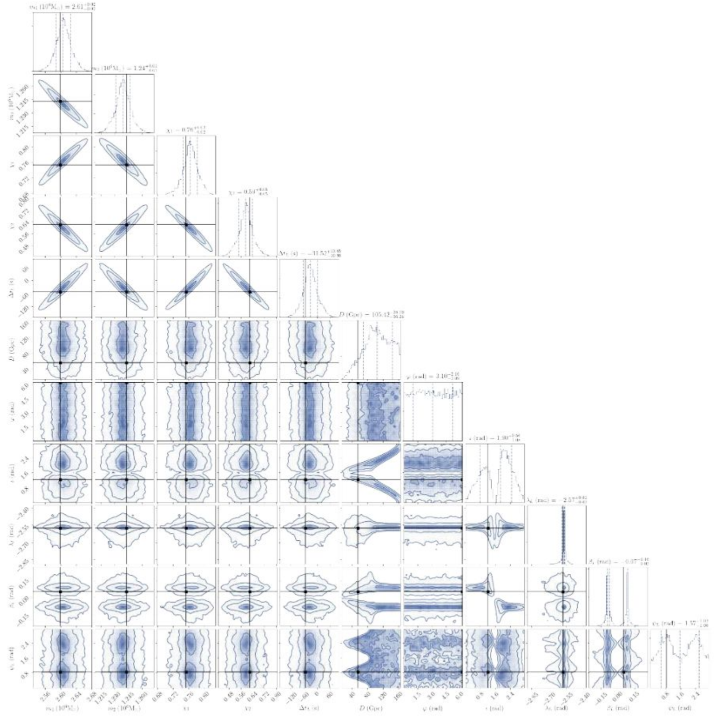
MBHB sampling with degeneracies: parameter map

Transformed parameters

Sampling greatly simplified



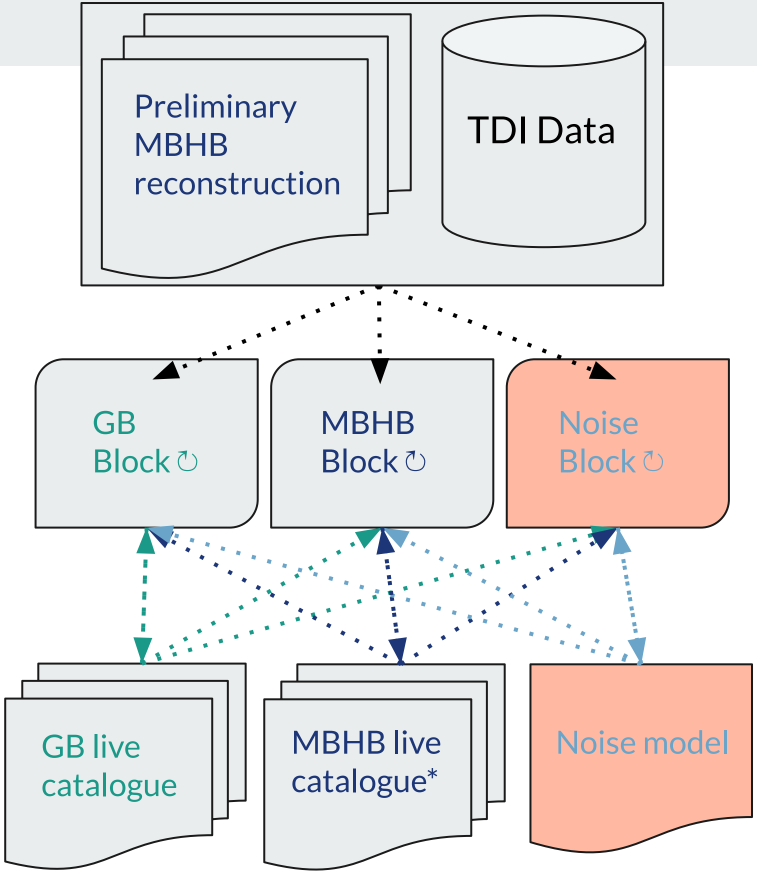
Original (physical) parameters

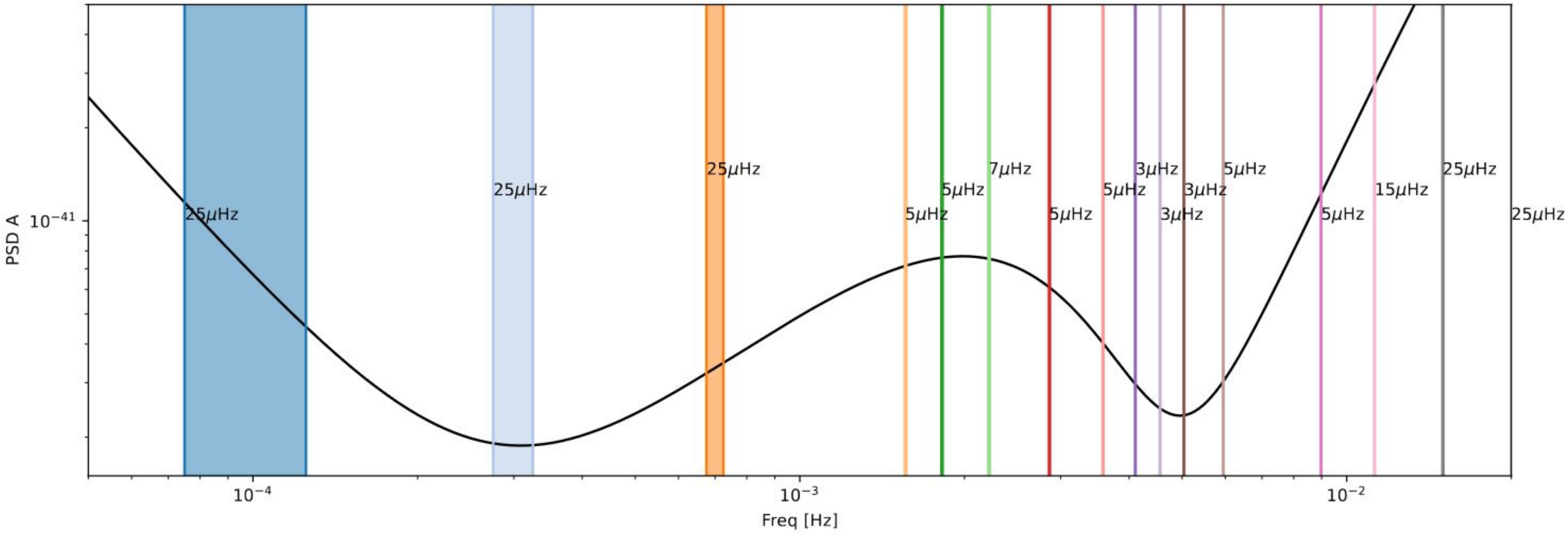


Refine MBHB PE

Get the noise level

- Starting from the PSD estimation
- Plug in a parametric model and fit





Simultaneously fit the parametric model using a dozen of bins

Get the noise model

Summary

With the Sangria analysis, we have demonstrated all the four components of the global-fit prototype

- Fast and preliminary detection/removal of MBHBs
- Search for Galactic binaries in small overlapped frequency bands
- Fast PE for MBHBs
- Noise model fitting

We built a modular architecture combining the components in concert

Next steps

Short term

- Robust stopping criterion for new source discovery (GB Block)
- Time iteration: Data accumulates with time. Each type of source has its own good cadence for data analysis.
- Dealing with gaps and glitches

Long term

- Add the modules (blocks) for other sources (SMBH, EMRI, ...)

Backup slides

$$f_{l1} = 4 \times 10^{-4}\text{Hz} \qquad f_{u1} = 2 \times 10^{-3}\text{Hz}$$
$$f_{l2} = 8 \times 10^{-3}\text{Hz} \qquad L_{\text{arm}} = 8.3391023799538\text{s}$$

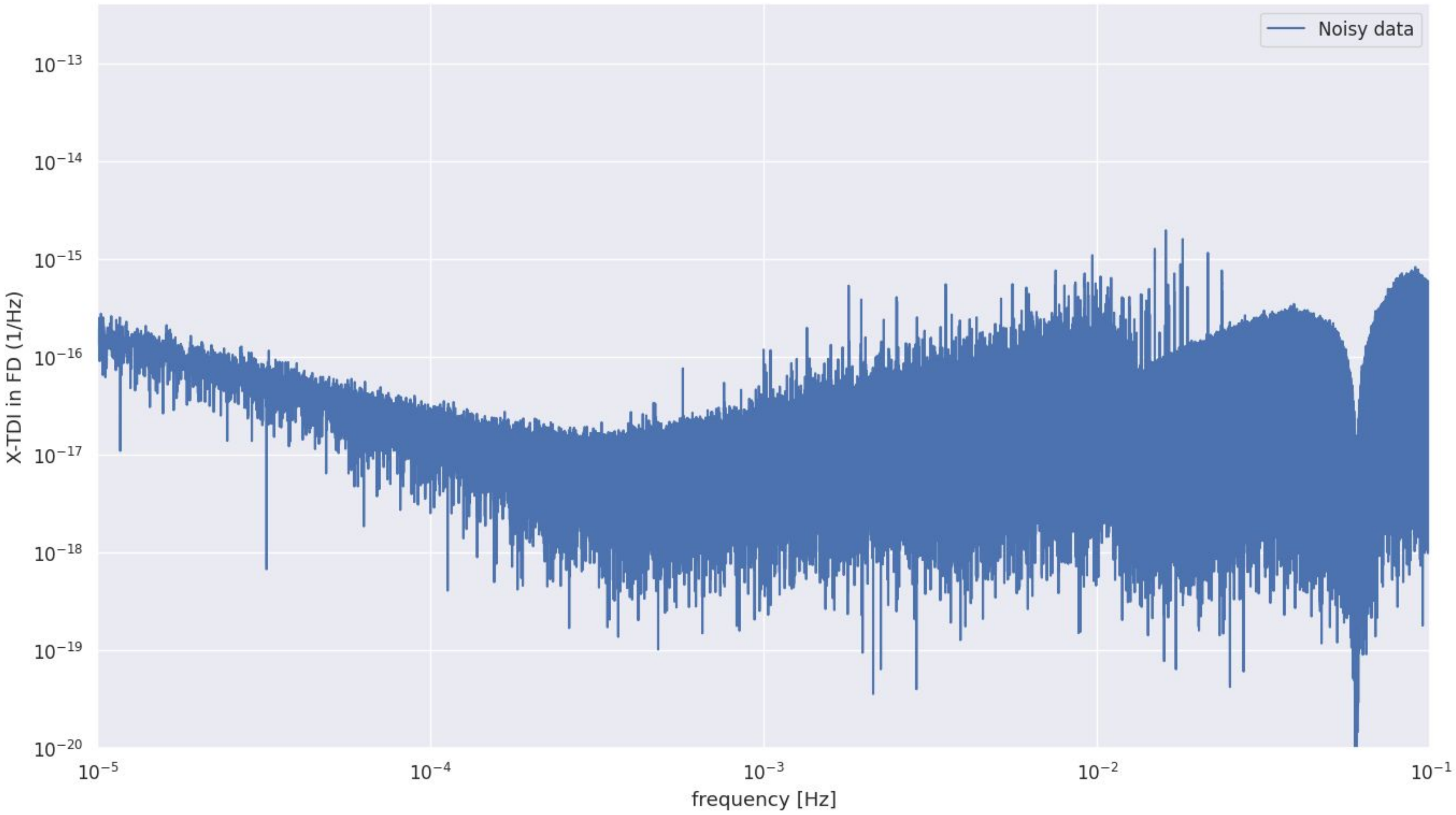
$$S_{\text{pm}} = S_{\text{acc}}(1 + (f_{l1}/f)^2(1 + (f/f_{l2})^4)/(2c\pi f)^2$$
$$S_{\text{op}} = S_{\text{oms}}(1 + (f_{u1}/f)^4/(2c\pi f)^2$$
$$x = 2c\pi f L_{\text{arm}}$$

For TDI A channel,

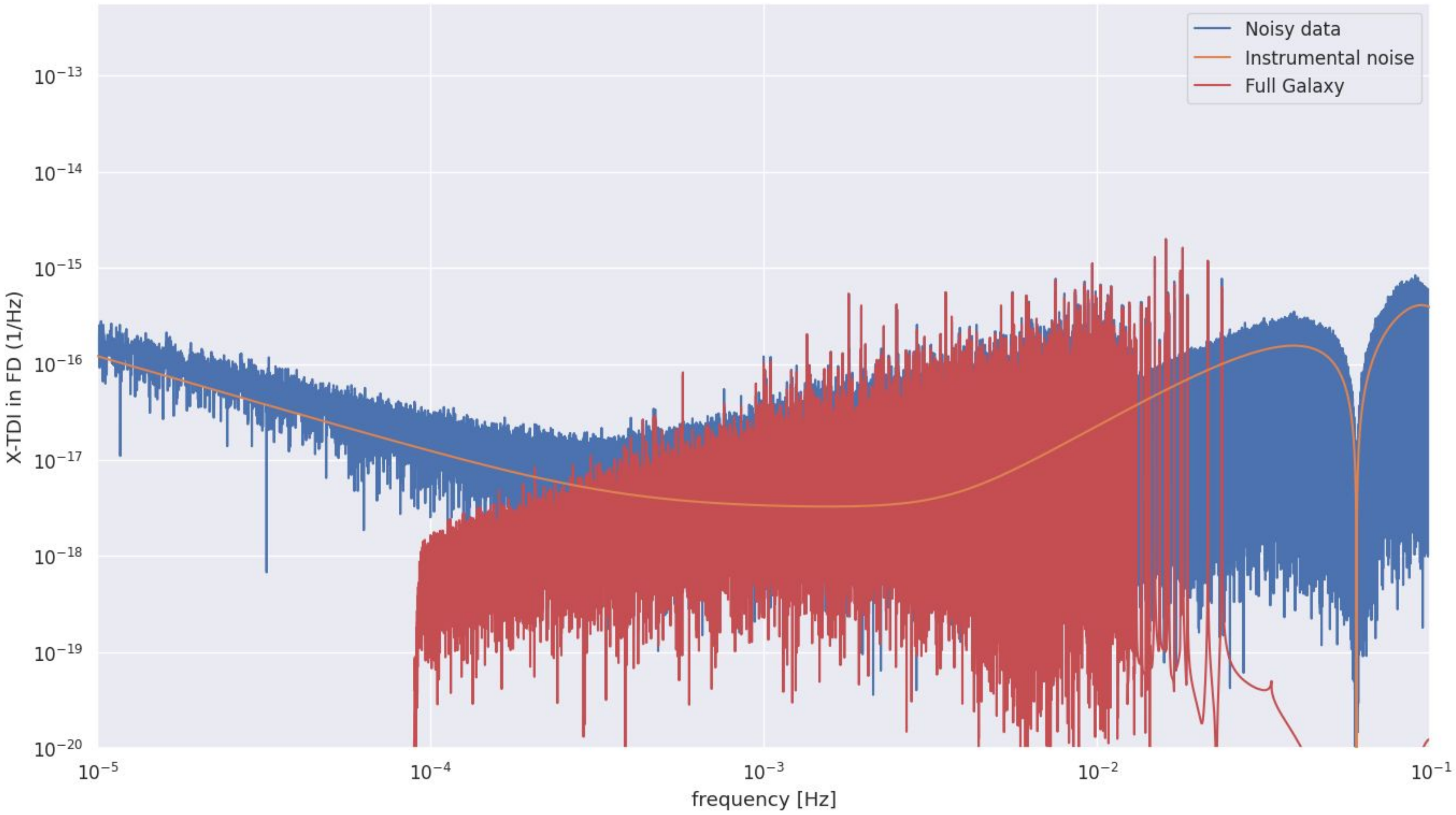
$$S_{\text{instr}} = 8\sin^2(x)(2S_{\text{pm}}(3 + 2\cos(x) + \cos(2x)) + S_{\text{op}}(2 + \cos(x)))$$
$$S_{\text{gal}} = 6(x\sin(x))^2 A \cdot f^{-\frac{7}{3}} \cdot \exp\left(-\left(\frac{f}{f1}\right)^{\alpha}\right) \cdot \frac{1}{2}\left(1 + \tanh\left(-\frac{f - f_{\text{knee}}}{f2}\right)\right)$$
$$S = S_{\text{instr}} + S_{\text{gal}}$$

The parameters are $S_{\text{acc}}, S_{\text{oms}}, A, f1, f2, \alpha, f_{\text{knee}}$

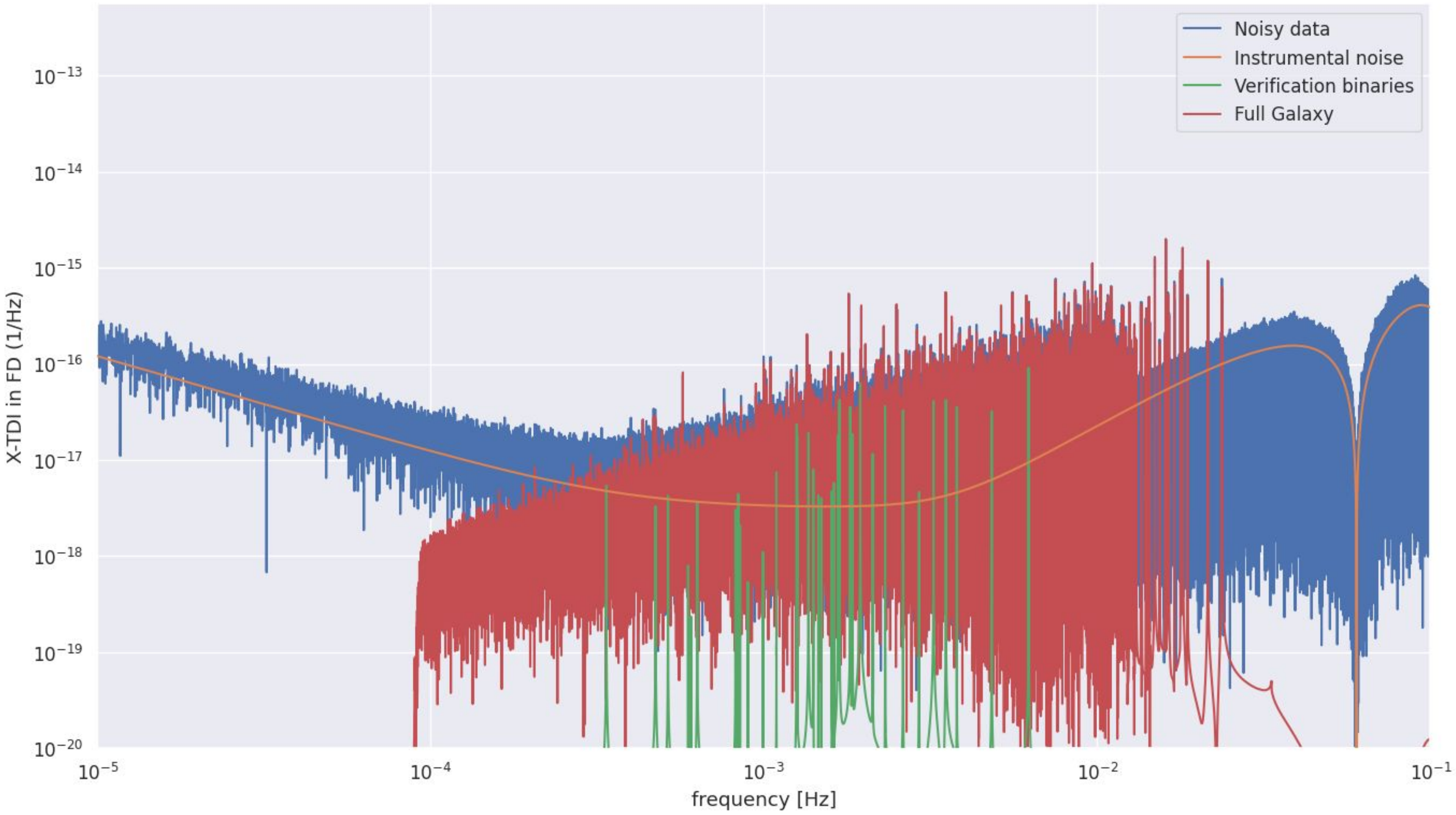
The parameteric noise model



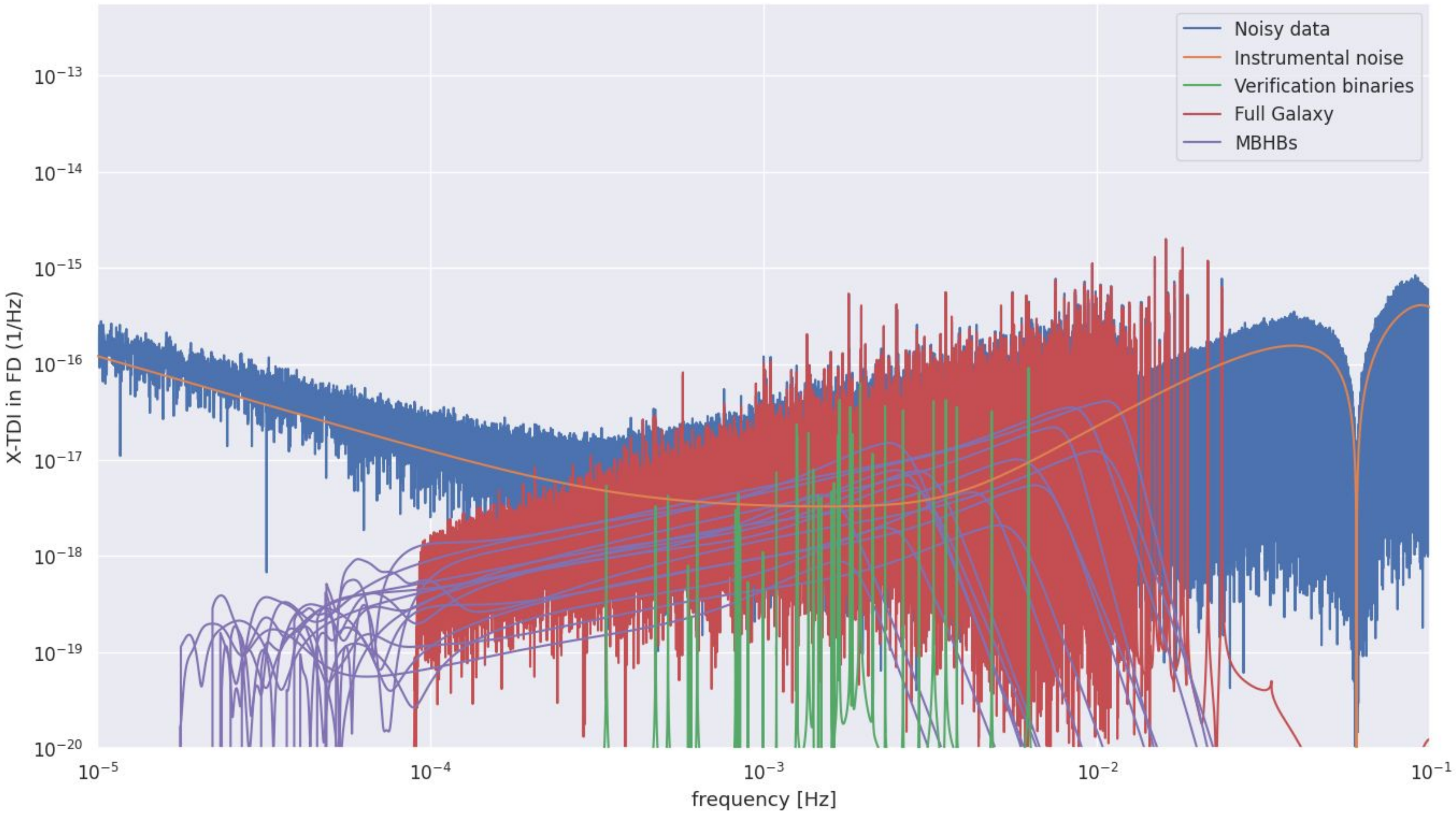
Sangria in FD



Sangria in FD

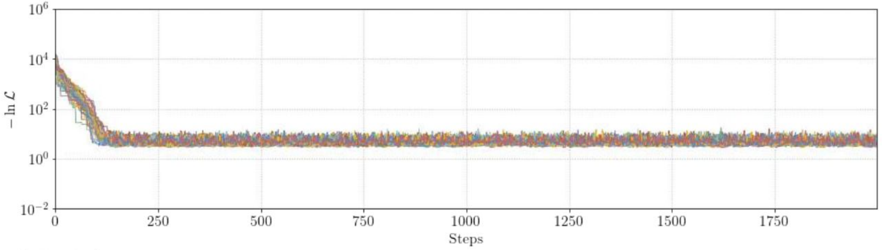


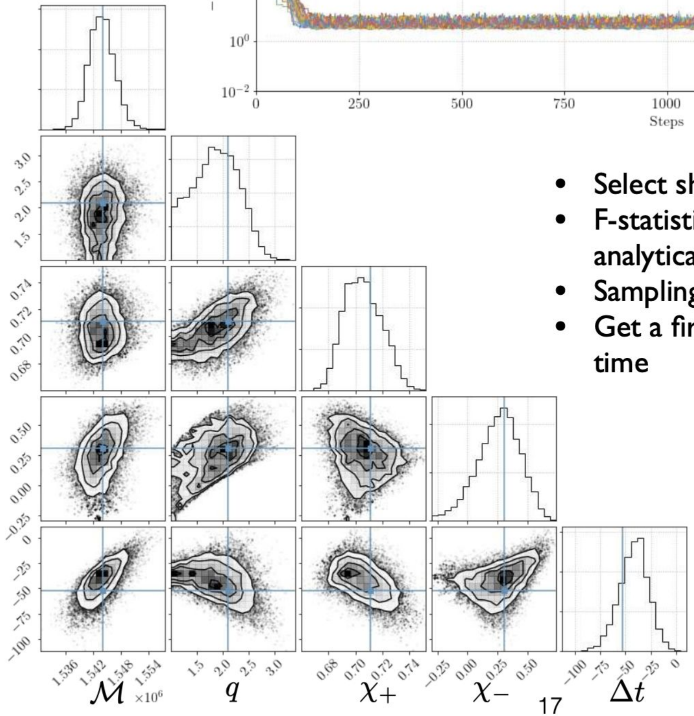
Sangria in FD



Sangria in FD

MBHB initial search: F-statistic on small data segments

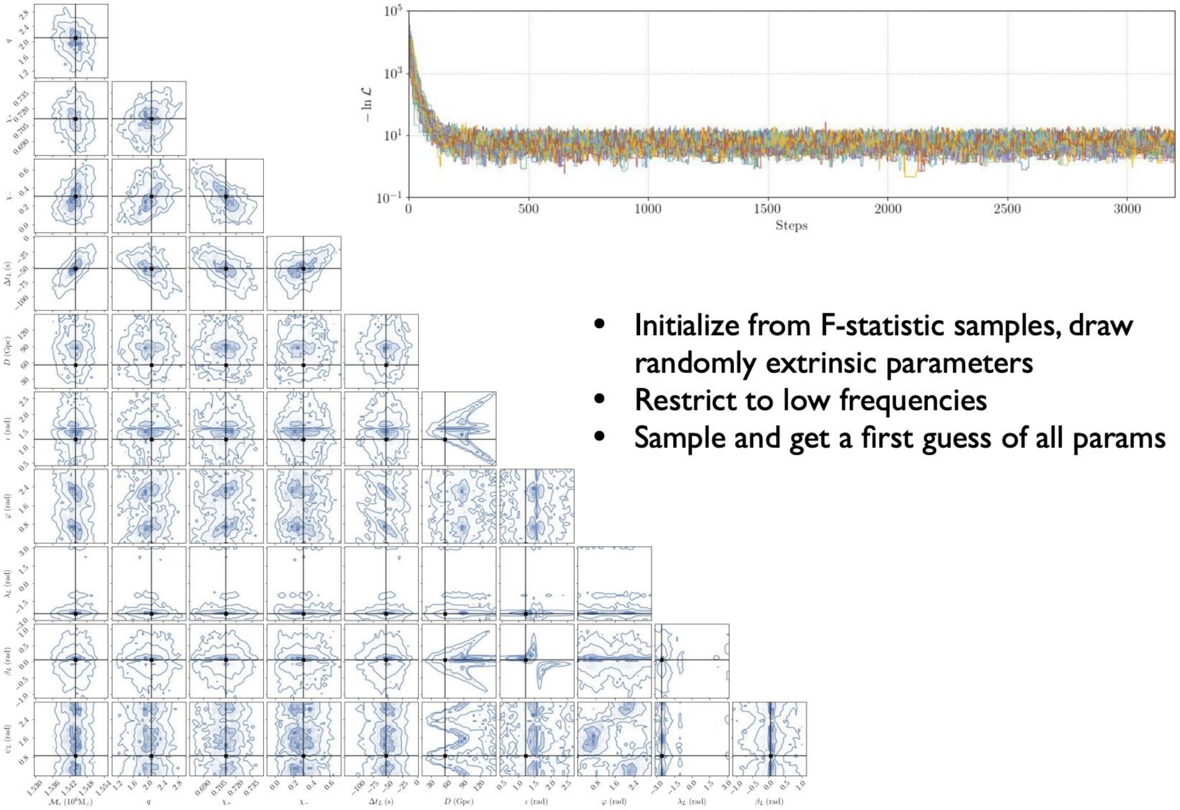




- Select short data segment
- F-statistic: approximate response, optimize analytically over extrinsic parameters
- Sampling easier for a lower dimensionality
- Get a first guess of intrinsic parameters + time

Low frequency sampling of MBHB signals

MBHB initial PE: sampling with low frequencies



- Initialize from F-statistic samples, draw randomly extrinsic parameters
- Restrict to low frequencies
- Sample and get a first guess of all params

Low frequency sampling of MBHB signals

MBHB signal: heterodyned likelihood

Decomposing the likelihood:

$$\ln \mathcal{L} = -\frac{1}{2}(s-d|s-d)$$
$$= -\frac{1}{2}(s-s_0|s-s_0) + (s-s_0|d-s_0) - \frac{1}{2}(s_0-d|s_0-d)$$

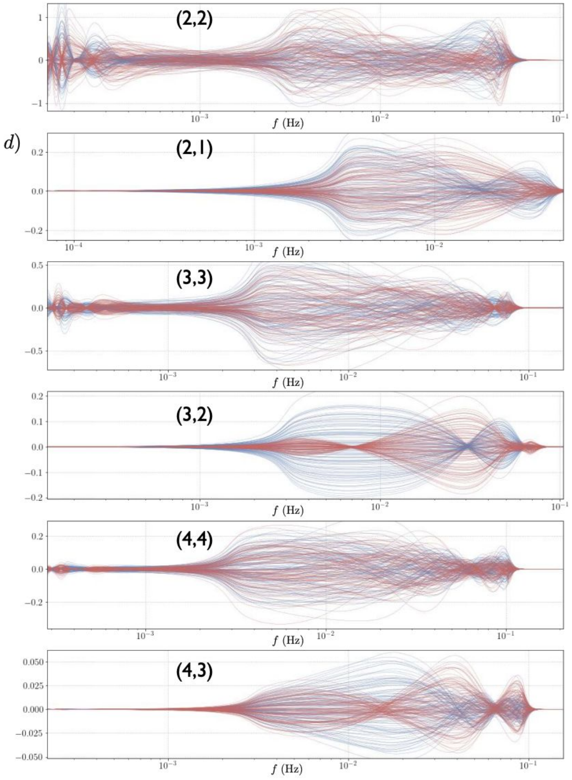
Residuals from reference waveform:

$$s_{\ell m} = r_{\ell m} e^{i\Phi^0_{\ell m}}$$

Implementation:

$$(s-s_0|s-s_0) = \sum_{\ell m}\sum_{\ell' m'} (r_{\ell m} r^*_{\ell' m'} | e^{i(\Phi^0_{\ell' m'} - \Phi^0_{\ell m})})$$
$$(s-s_0|d-s_0) = \sum_{\ell m} (r_{\ell m} | e^{-i\Phi^0_{\ell m}} (d-s_0))$$

- Fix a sparse frequency grid (~128)
- Linear interpolation of the residuals, mode-by-mode
- Precompute 0-th and 1st polynomial inner products against phase and data terms, with a fine resolution



Low frequency sampling of MBHB signals