Preparing for the real thing:

A joint effort to reconstruct SGWB signals from realistic synthetic data

Nikolaos Karnesis on behalf of
Baghi, Bayle, Buscicchio, Caprini, Dam, Hartwig, Inchauspé, Muratore, Nardini, Pieroni, Pozzoli, Santini

Towards a realistic forecast detection of Primordial Gravitational Wave Backgrounds

Valencia – 10 December, 2024

A lot of new people, methods, codes, agencies, institutes, teams…

Recent updates

- **Consortium restructuring.**
  - New version of the working groups.
- **Distribute Data Processing Center (DDPC) is officially launched!**
  - Countries and institutes are jumping in.
  - Official pipelines are being discussed.
  - Data engineers are in to help.
  - Data products are being discussed.
- **The LISA Science Team (LST) started working!**
  - Science goals, catalogues, thresholds, etc, are being defined.
- **New players:** new groups, institutes, methods, codes are joining the effort**.**
- **Therefore:** A bit of coordination on projected capabilities (concerning SGWB detection) is needed. **That's one of the goals of this collaborative effort**.

Purpose and status of joint project

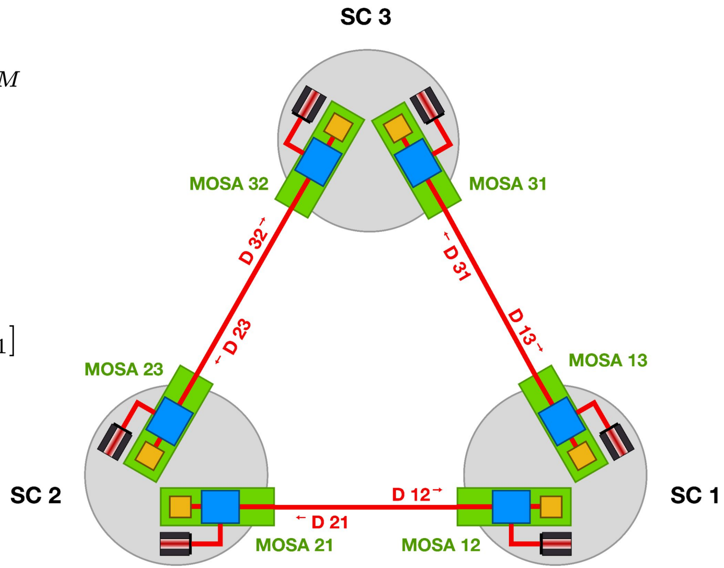
- **Assess capabilities of LISA to detect and characterize SGWB signals.**
    - This is the final goal, but there are assumptions involved.
- **Introduce realism!**
    - From simpler analytic orbits, to realistic orbits simulated by ESA.
    - Noise and/or signal stationarity is explored.
    - Abandon the “ideal” explicit spectral model of LISA.
    - Channel correlations!
- **Compare the performance of different methods in a controlled setup.**
    - Test methods based on different statistical approaches/assumptions.
    - Test the effect of noise and signal model types on the detection of SGWB.
- **Try different injection signals:** Explore parameter spaces for different cosmological models.
- Associated with the Science Investigation Work Package, it was first discussed and planned during the workshop in Geneva.

Instrumental setup

- Time Delay Interferometry:
  - Simulator starts with 0.5 Hz sampling frequency. Generates data at the ISI, RFI, TMI.
  - Build the eta variables: $\eta_{12} = N^{\mathrm{oms}}_{\mathrm{isi}_{12}} - \nu_0 \dot{H}_{12} + N^{TM}_{12} + \mathbf{D}_{12} N^{TM}_{21}$ where $\mathbf{D}_{ij} x(t) = x(t - L_{ij}(t))$
  - And then the TDI variables as (1st generation):

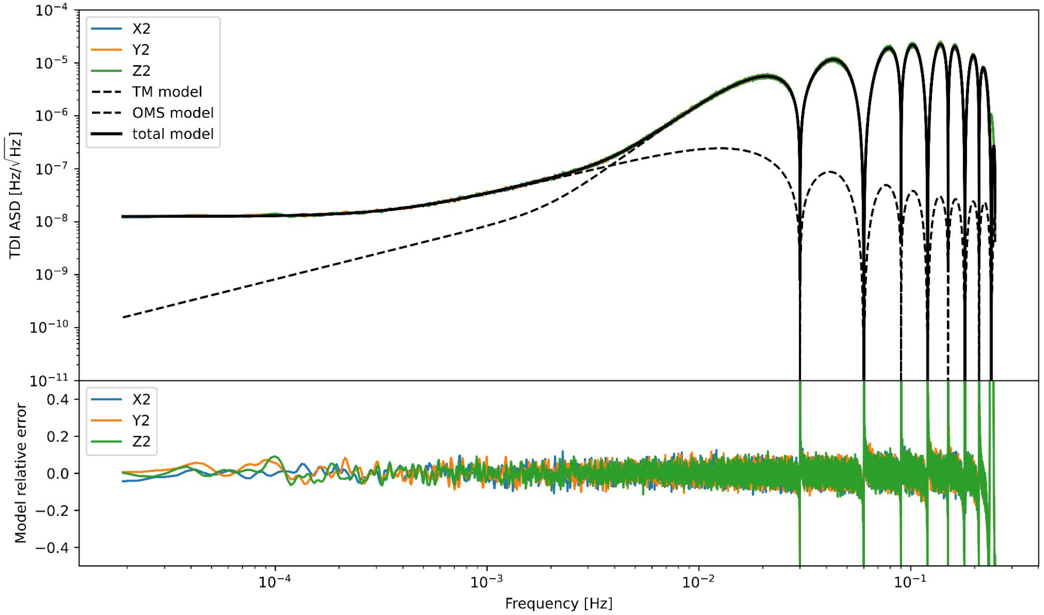
$$
\begin{aligned}
X_1 = &\ \eta_{13} + \mathbf{D}_{13}\eta_{31} + \mathbf{D}_{131}\eta_{12} + \mathbf{D}_{1312}\eta_{21} \\
&- [\eta_{12} + \mathbf{D}_{12}\eta_{21} + \mathbf{D}_{121}\eta_{13} + \mathbf{D}_{1213}\eta_{31}]
\end{aligned}
$$

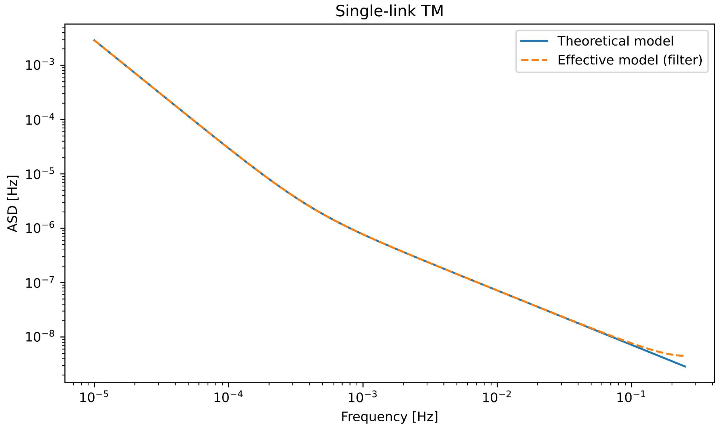
- Noises and signals:
  - The usual test-masses (TM) and optical metrology system (OMS).
  - An extra term that comes from the filters is added.
  - For now we have begun our investigations with equal noises across distant S/C.
  - Signal is produced by time-domain filtering white noise.

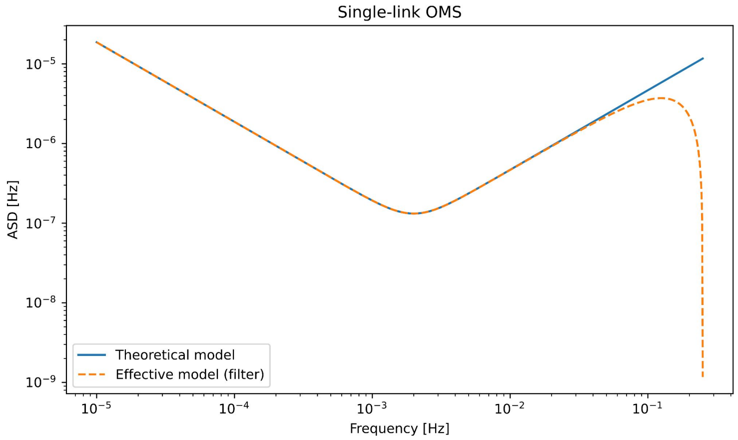


Instrumental setup

Here is how the noise looks like. Single links and the final TDI synthesization.







Plots by JB Bayle

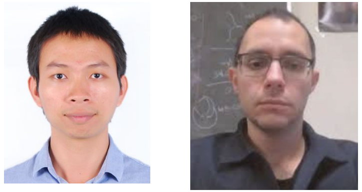
The noise data sets

Data set	Realizations	Description
noise-1	`10 (a-j)`	LDC Spritz noise curves, equal-armlength orbits, TDI2.
noise-2	`2 (a-b)`	LDC Spritz noise curves, ESA trailing orbits, TDI2.
noise-3	`10 (a-j)`	LDC Spritz noise curves. Constant *unequal* light travel times, TDI2.
noise-4	`10 (a-j)`	LDC Spritz noise curves. Constant *equal* light travel times (`8.3 s`), TDI2.

The signal data sets

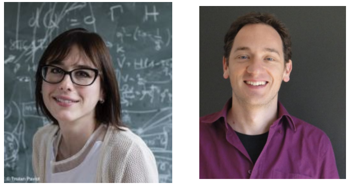
Data set	Realizations	Description
cs-1	2 (a-b)	**Cosmic Strings** just beyond the reach of SKA. Equal-armlength orbits, TDI2.
cs-4	4 (a-d)	**Cosmic Strings** as above. Constant equal light travel times (of 8.3 s), TDI2.
fopt-1	2 (a-b)	**First-Order Phase Transitions**. Using equal-armlength orbits, TDI2.
gb-1	1 (a)	**Galactic Binaries** (LDC Spritz). Equal-armlength orbits, TDI2.
sobh-1	2 (a-b)	**Stellar-mass black hole binaries** (merger rate from LVK). Equal-armlength orbits, TDI2.
sobh-2	2 (a-b)	**Stellar-mass black hole binaries** (merger rate from LVK), ESA trailing orbits, TDI2.

Meet the group, the “plumbers” and their “pipelines”



The Binner: Dam, Pieroni

The NQJBHM: Bayle, NK, Baghi, Inchauspé, Besancon



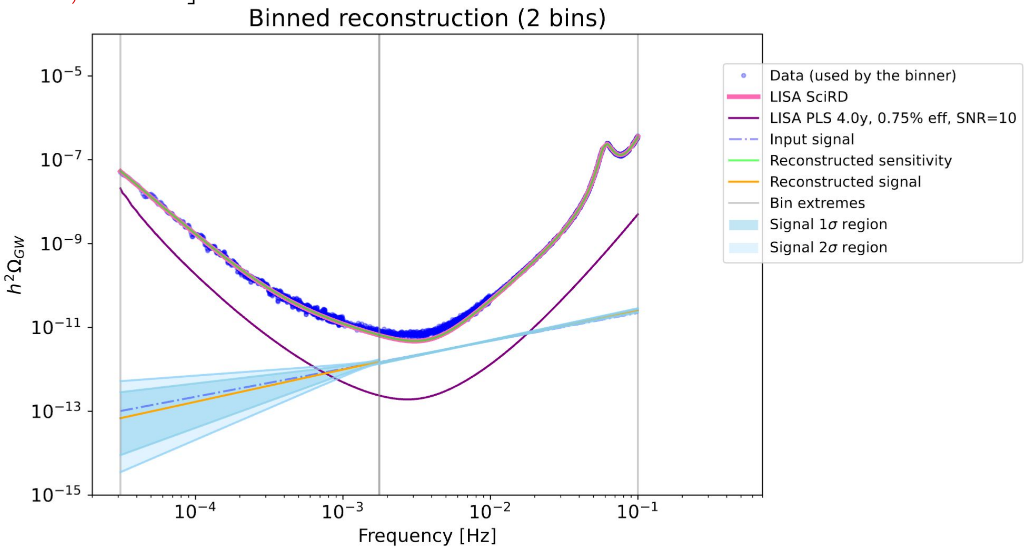
Coordination: Caprini, Nardini

The AEI: Santini, Muratore, Hartwig

The Balrog: Pozzoli, Buscicchio

The Binner

- Based in the SGWBinner software [Flauger *et al* JCAP, 01, 059, 2021]
  - Developed *by and for* the Cosmology WG.
  - Agnostic reconstruction method for the signal.
  - Based on “binning” the data and fitting them with sets of power-law models.
  - Number and width of bins are decided with model selection criteria.
  - Coarse-graining the data, using a mixture of Normal-Lognormal likelihood.
  - Check [Caprini *et al* JCAP, 11, 017, 2019].
  - Talk by M. Pieroni earlier!



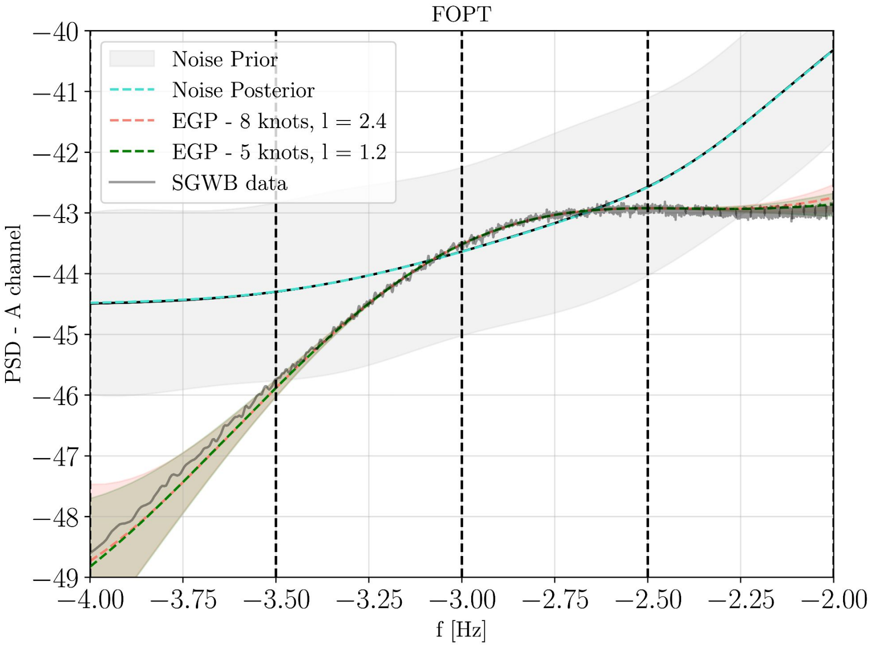
The Balrog

- Based in the Balrog software suite
  - Developed as a LISA simulation and DA tools, in general.
  - Aimed to be developed into a full Global Fit pipeline.
  - Based on Gaussian Processes methods, fitting both signal and noise PSDs (at the TDI level) with weakly parametric models.

$$g(f) \sim \mathcal{GP}(m(f), k(f, f'))$$

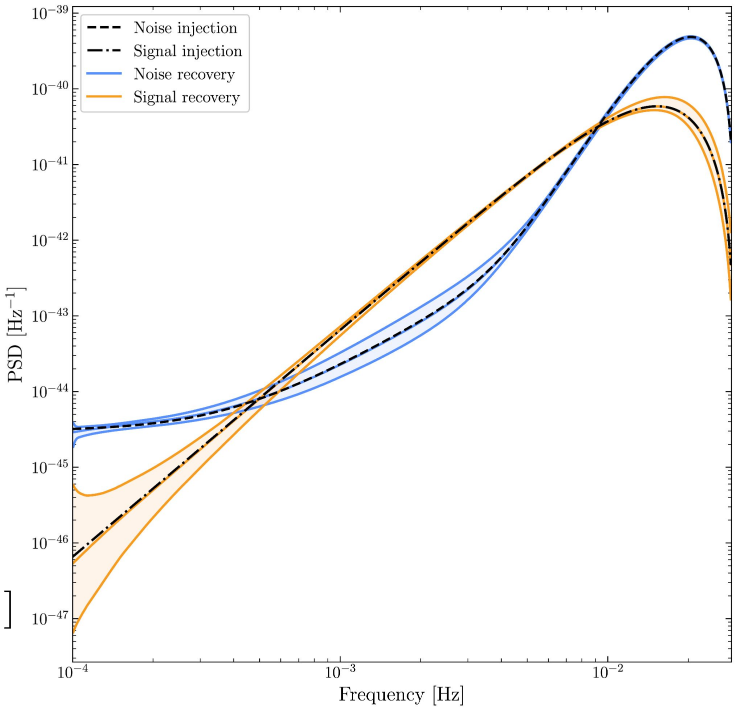
$$A(f) - m(f) = \Sigma(f \mid f_{\text{knots}})\Sigma^{-1}(f_{\text{knots}}, f_{\text{knots}})(g(f) - m(f))$$

  - [Pozzoli *et al*, PhysRevD.109.083029]



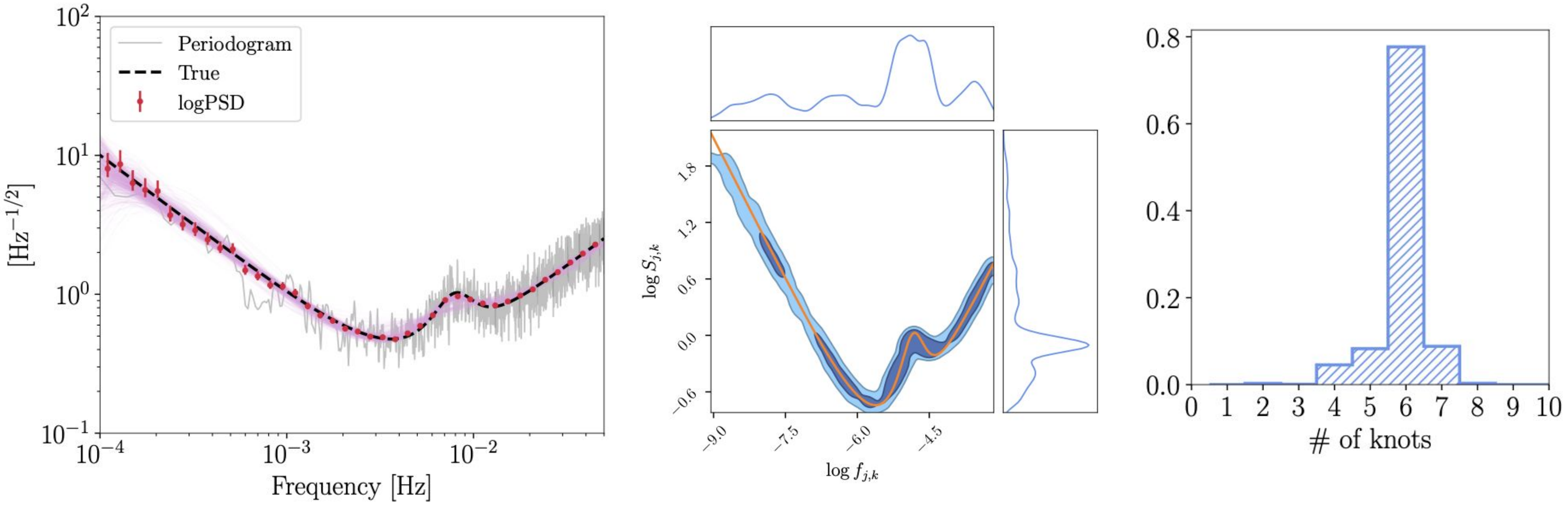
The AEI

- Based in the Bayesian trans-dimensional methods and spline interpolation methods
  - Adopt spline models as flexible PSD models (Akima interpolation).
  - Model *deviations* from the true injected PSD data.
  - Model the data at the TDI level, fit the noise-orthogonal AET channels.
  - Fit smoothed PSD of the data, use the complex Wishart likelihood [More by Quentin Baghi later].
  - Use of trans-dimensional methods to decide on number of degrees of freedom [again, more by Quentin Baghi later].
  - `Eryn` sampler: [NK *et al* MNRAS, 526, 4, 2023]



Parenthesis on the trans-dimensional algorithms

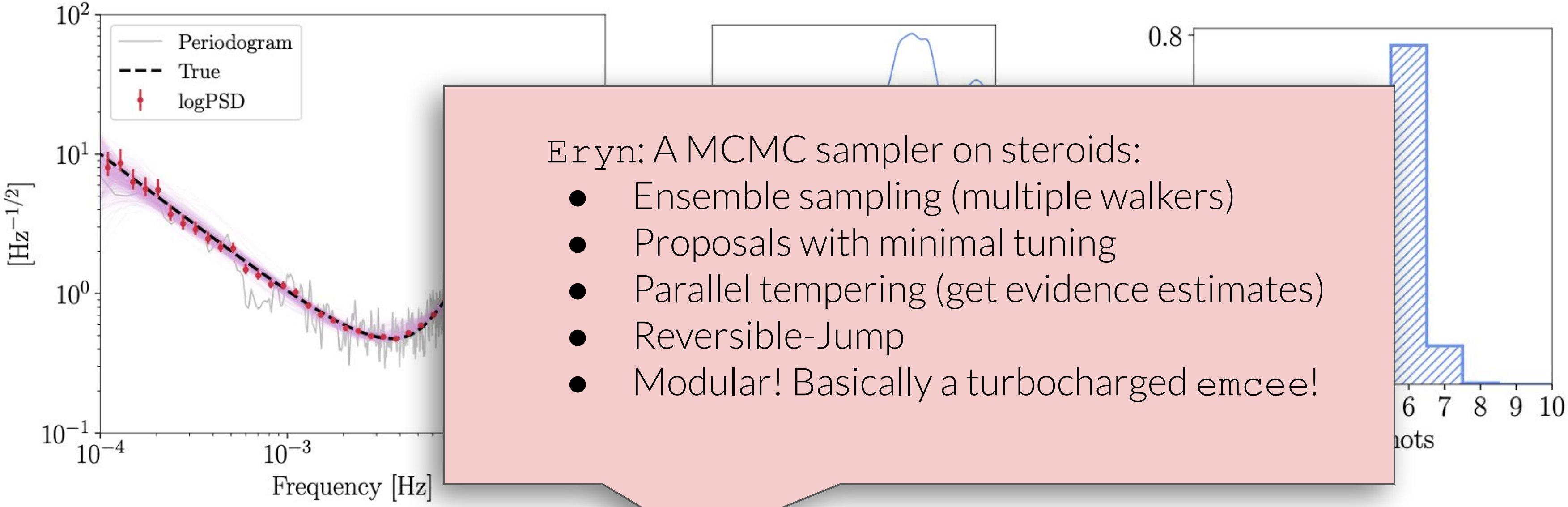
- Reversible Jump MCMC
  - Basically a generalization of the Metropolis-Hastings algorithm.
  - Propose the extension or reduction of the order of the model at each step.
  - In our case it's the number of spline knots!
  - See the `Eryn` sampler [NK *et al* MNRAS, 526, 4, 2023] for examples!



Grab the source here: https://github.com/mikekatz04/Eryn , or just **`pip install eryn`**

Parenthesis on the trans-dimensional algorithms

- Reversible Jump MCMC
  - Basically a generalization of the Metropolis-Hastings algorithm.
  - Propose the extension or reduction of the order of the model at each step.
  - In our case it's the number of spline knots!
  - See the `Eryn` sampler [NK *et al* MNRAS, 526, 4, 2023] for examples!



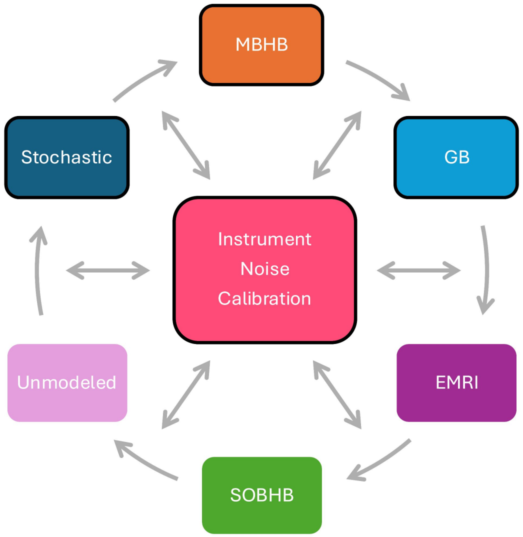
`Eryn`: A MCMC sampler on steroids:

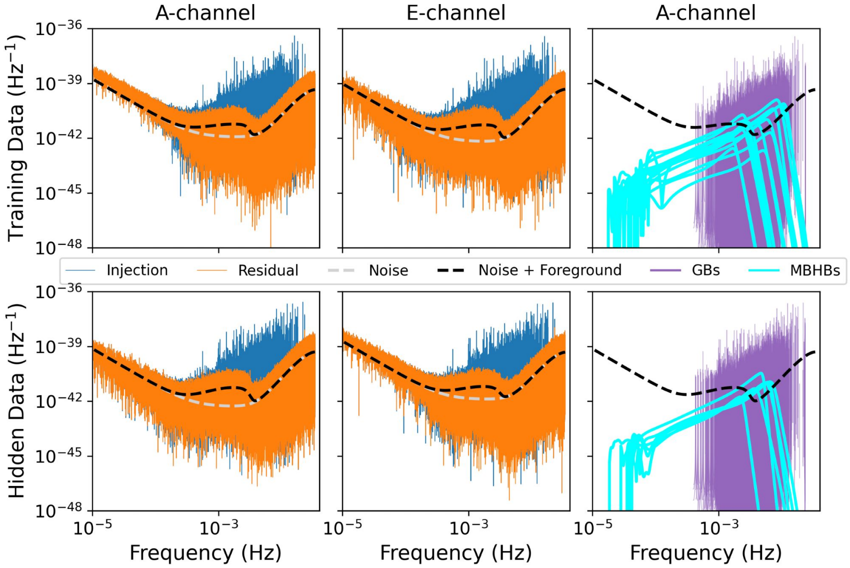
- Ensemble sampling (multiple walkers)
- Proposals with minimal tuning
- Parallel tempering (get evidence estimates)
- Reversible-Jump
- Modular! Basically a turbocharged `emcee`!

Grab the source here: https://github.com/mikekatz04/Eryn , or just **`pip install eryn`**

Parenthesis on the trans-dimensional algorithms

- Reversible Jump MCMC
  - See the `Eryn` sampler [NK *et al* MNRAS, 526, 4, 2023] for examples!
  - Used as a building block of the *Erebor* Global Fit pipeline!
  - Take a look at [Katz *et al* arXiv:2405.04690, accepted to PRD]

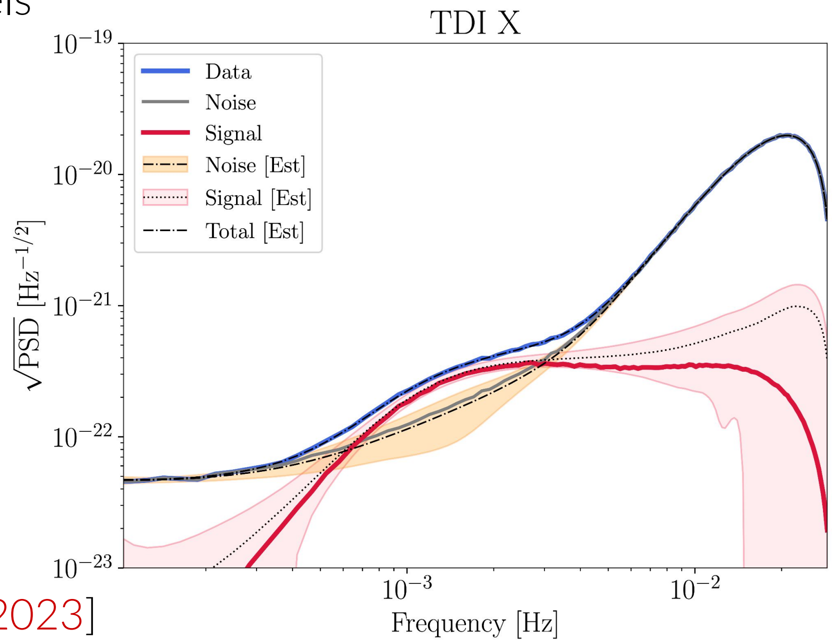




Grab the source here: https://github.com/mikekatz04/Eryn , or just **`pip install eryn`**

The NQJBHM [a cooler name needed here]

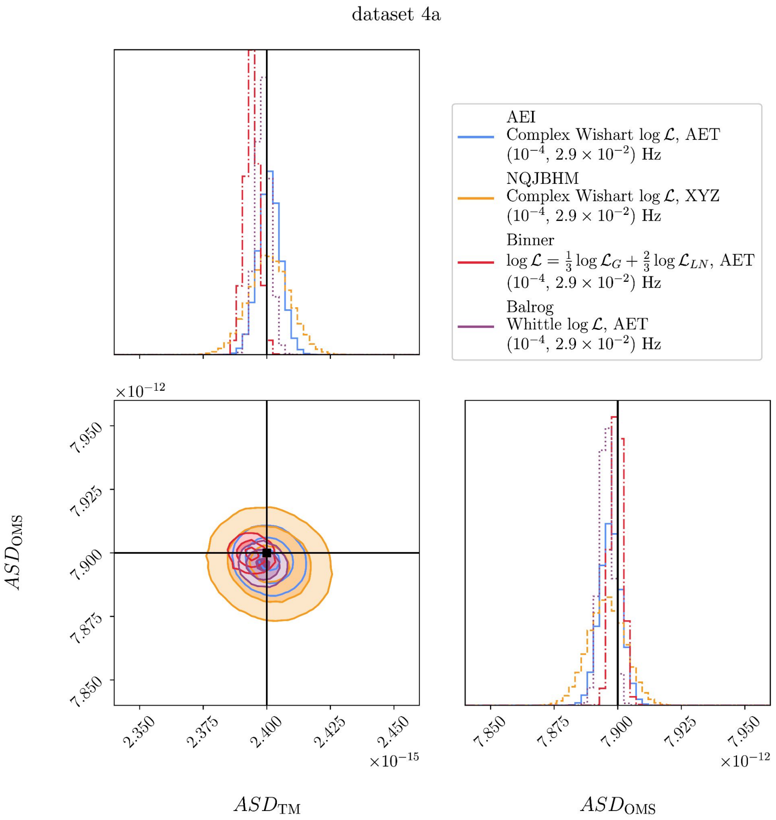
- Based in the Bayesian trans-dimensional methods and spline interpolation methods
  - Adopt spline models as flexible PSD models (Akima interpolation).
  - Model *deviations* from the true injected PSD data.
  - Model the data at the link level, fit the full covariance matrix XYZ.
  - Fit smoothed PSD of the data, use the complex Wishart likelihood [More by Quentin Baghi later].
  - Use of trans-dimensional methods to decide on number of degrees of freedom [again, more by Quentin Baghi later].
  - `Eryn` sampler: [NK *et al* MNRAS, 526, 4, 2023]
  - Based on the work of [Baghi *et al* JCAP, 04, 066, 2023]



Some [preliminary] findings

Some [preliminary] **findings**

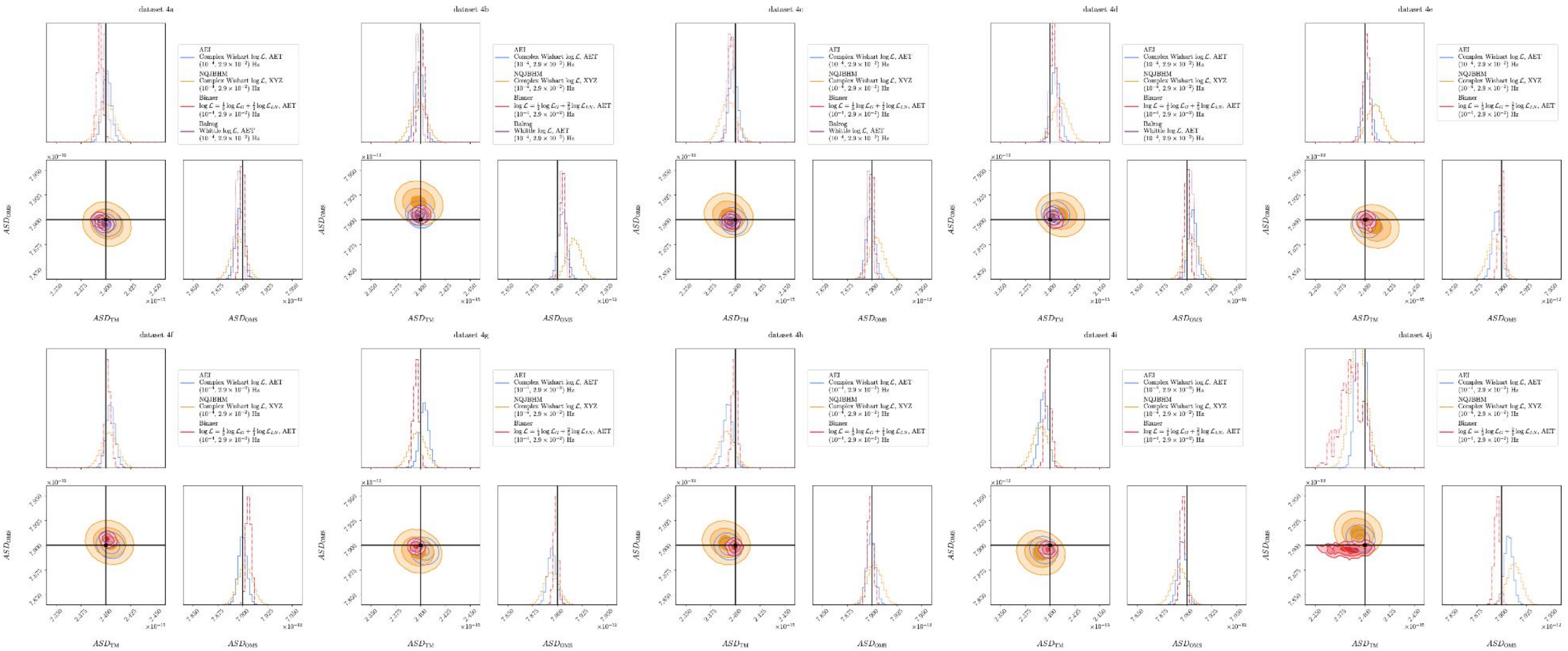
Starting point: Test statistical models and TDI-level modelling.



- Use the noise-4 data, which correspond to a rigid, equal armlength LISA configuration.
- No signal injection.
- Use analytic models only for the noises.
- Run all pipelines, and compare the posteriors for the noise parameters.

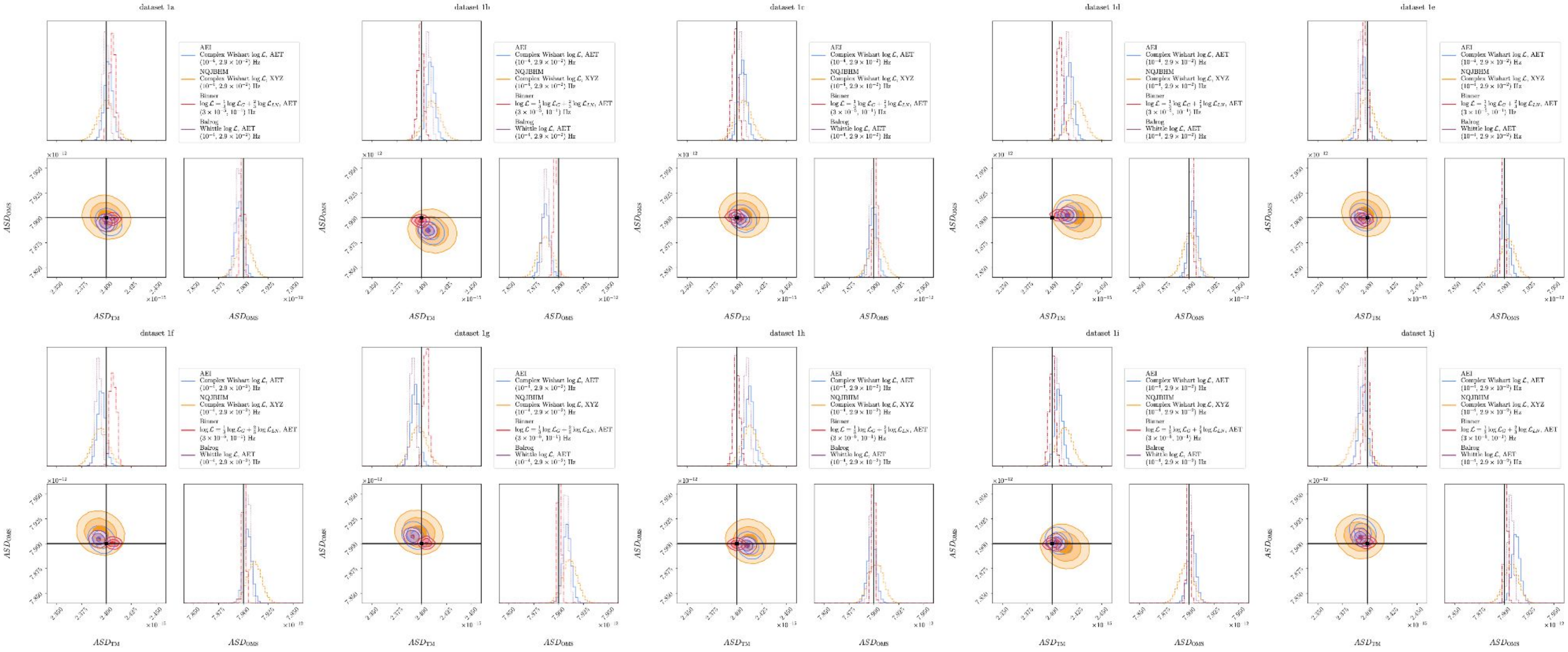
Some [preliminary] **findings**

Starting point: Test statistical models and TDI-level modelling by fitting the analytic models. This is for the noise-4, with the rigid, equal-arm LISA.



Some [preliminary] **findings**

Then: Test statistical models and TDI-level modelling by fitting the analytic models. This is for the noise-1, with the slowly varying, almost equal-arm LISA.

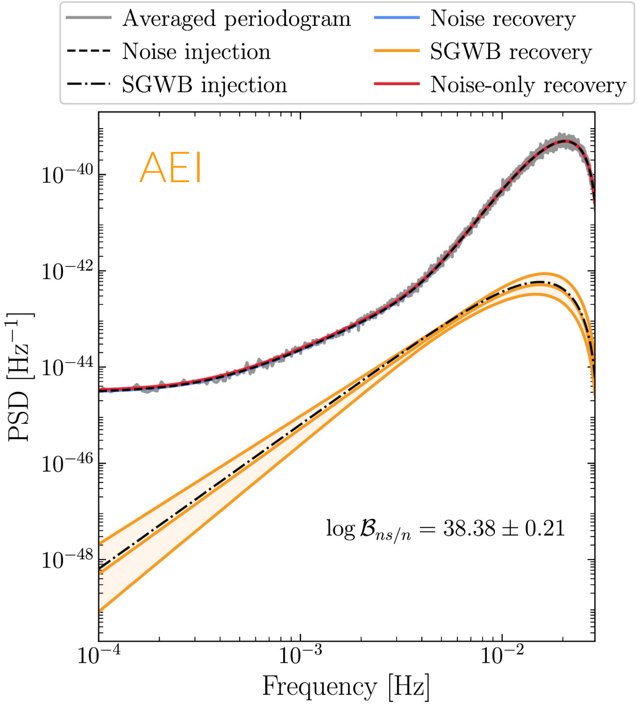


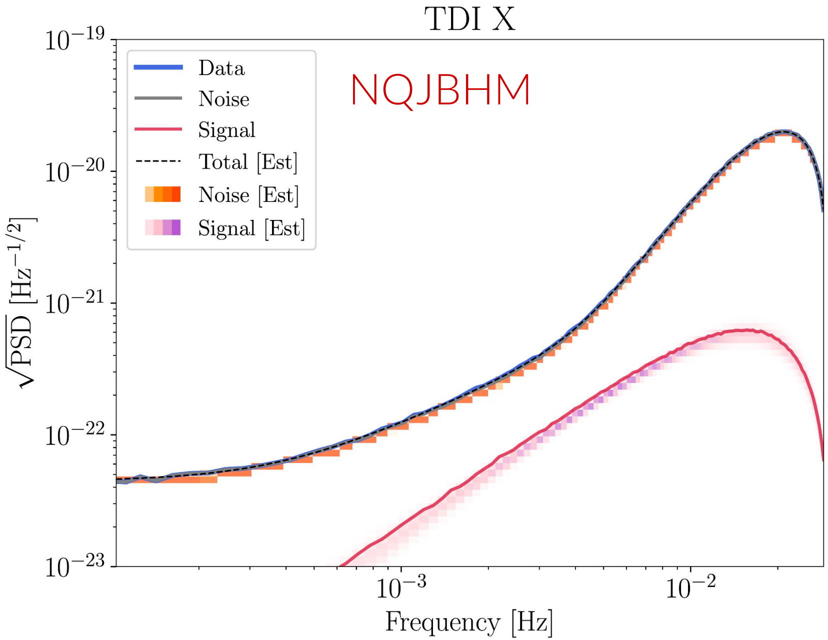
Some [preliminary] findings

Test detectability: Compute Bayes Factors in order to pinpoint differences in all approaches. We use the Cosmic Strings (`*0.1`) data set, and full analytic models. Also for the rigid LISA case.

Some [preliminary] findings

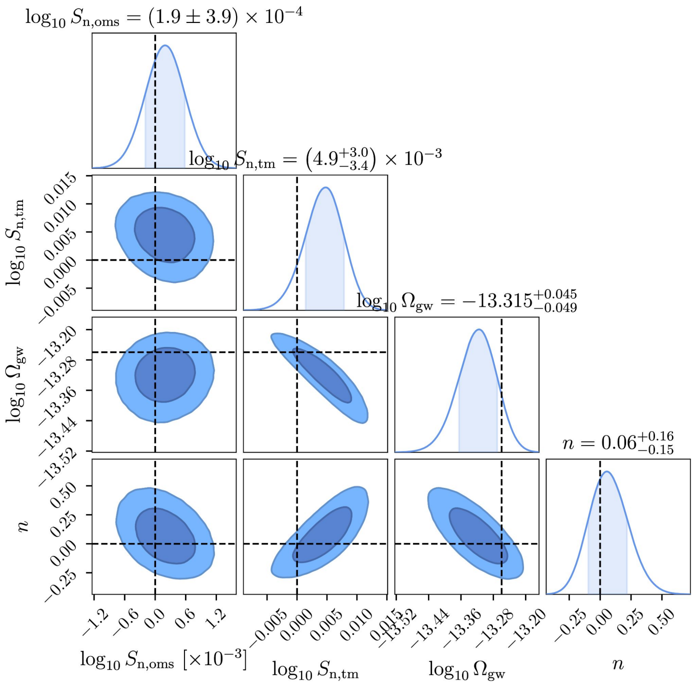
Test detectability: Compute Bayes Factors in order to pinpoint differences in all approaches. We use the Cosmic Strings (`*0.1`) data set, and full analytic models. Also for the rigid LISA case.

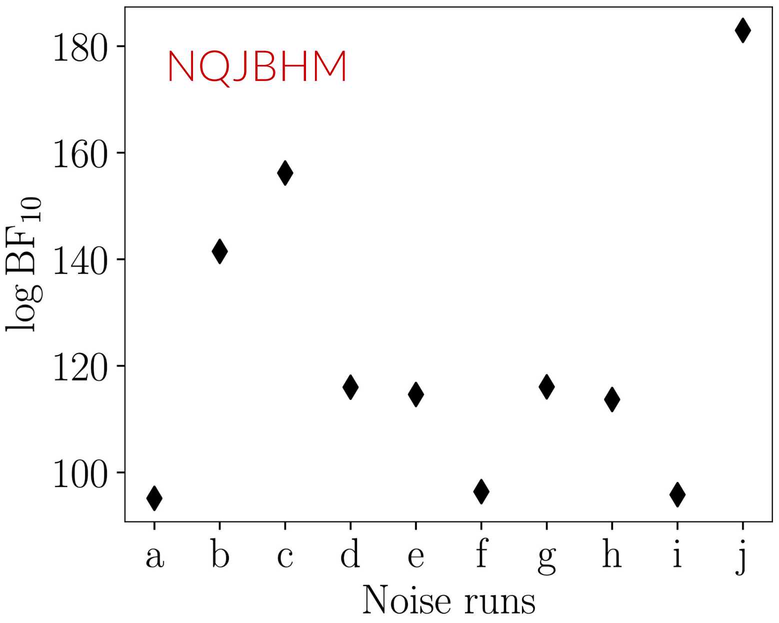




Some [preliminary] **findings**

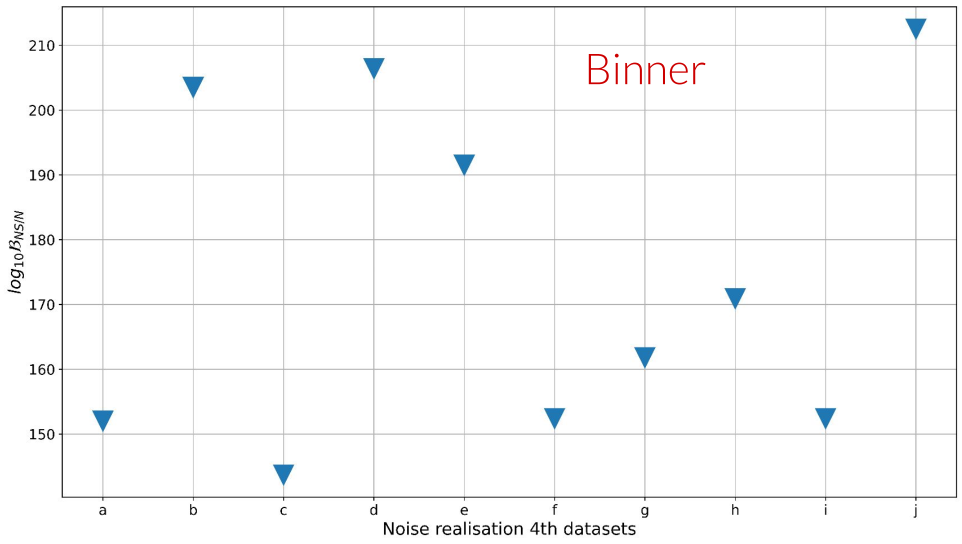
Test detectability: Compute Bayes Factors in order to pinpoint differences in all approaches. We use the Cosmic Strings (`*0.1`) data set, and full analytic models. Also for the rigid LISA case.

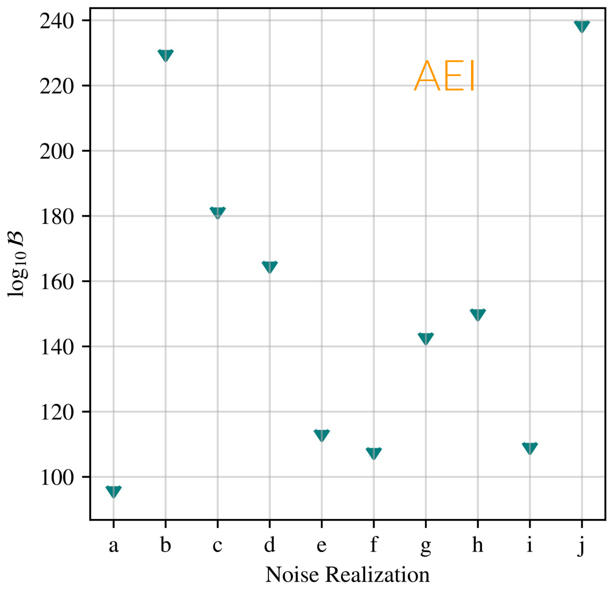




Some [preliminary] findings

Test detectability: Compute Bayes Factors in order to pinpoint differences in all approaches. We use the Cosmic Strings (`*0.1`) data set, and full analytic models. Also for the rigid LISA case.

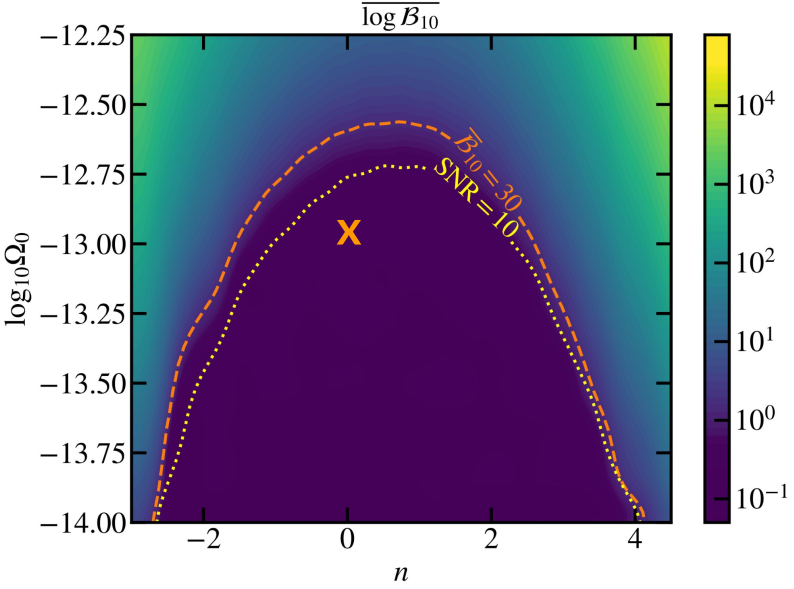


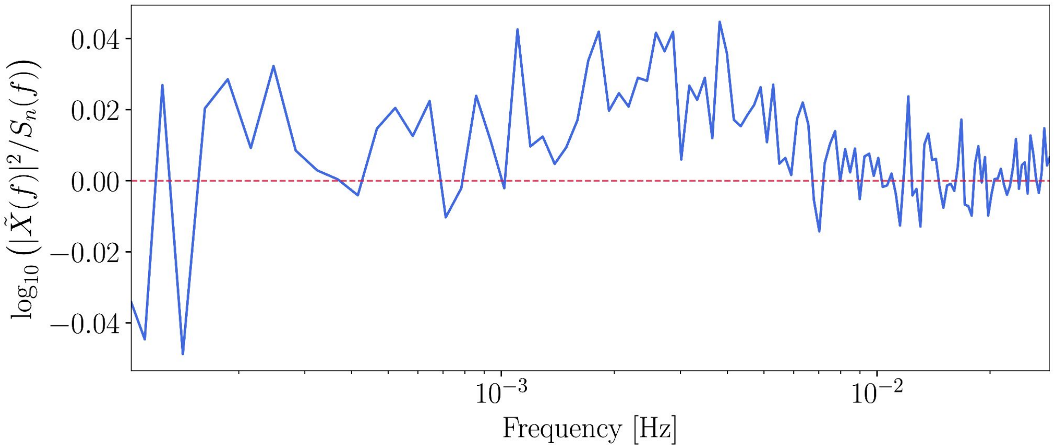


Some [preliminary] findings

This shows something important: Note that with flexible models, this signal is **NOT** detectable!

[Baghi *et al*, JCAP04, 066, 2023]

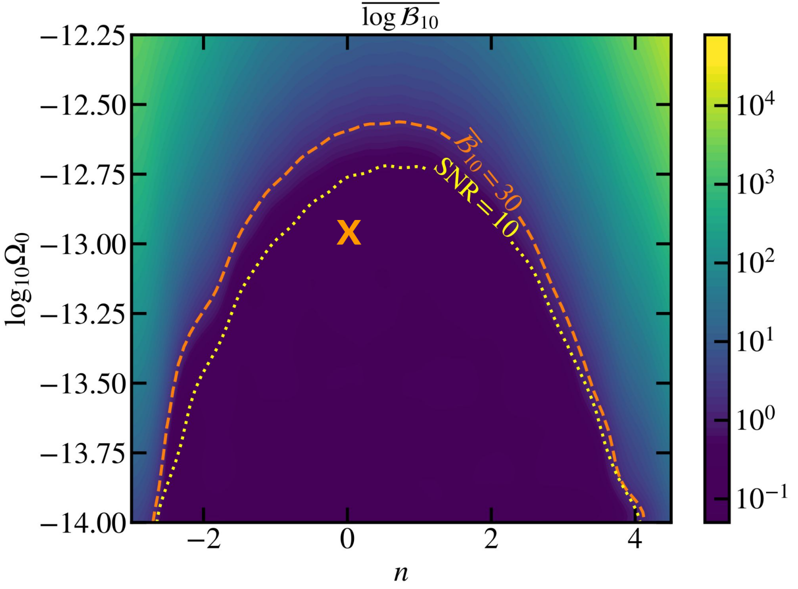


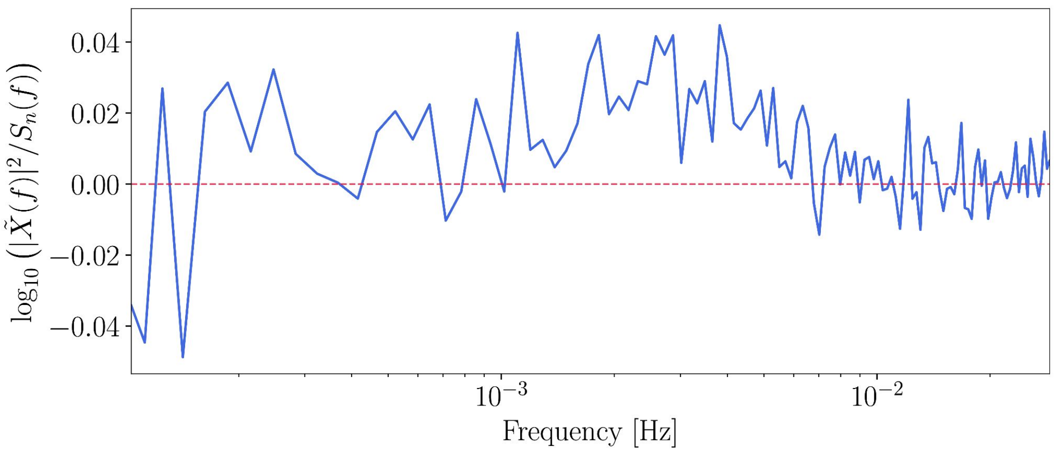


Some [preliminary] findings

This shows something important: Note that with flexible models, this signal is NOT detectable!

[Baghi *et al*, JCAP04, 066, 2023]





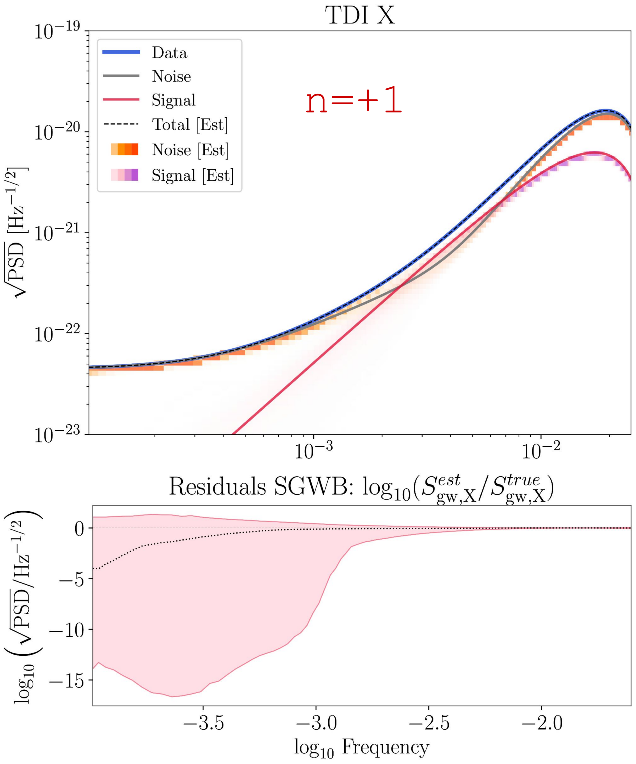
Assuming a spectral shape ~*is*~ prior knowledge!

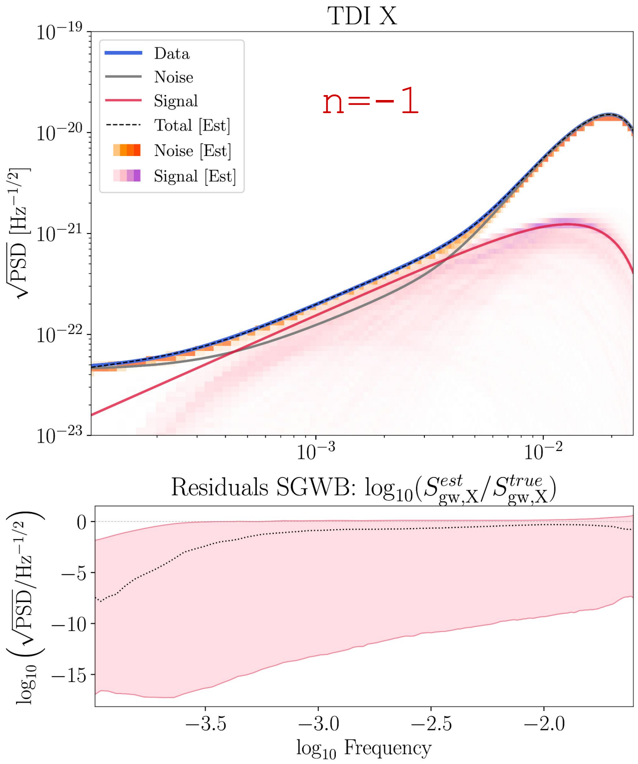
Some [preliminary**] findings**

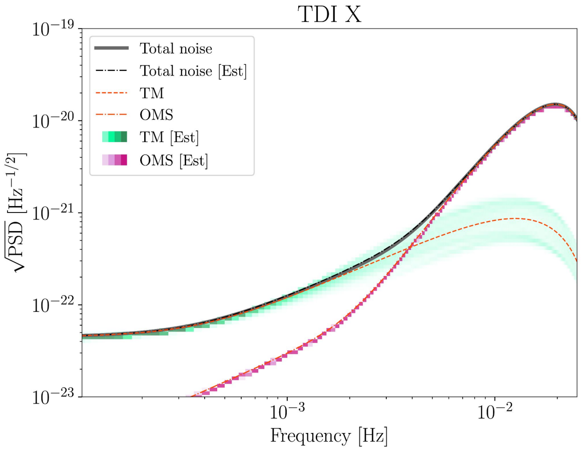
Another example on the importance of our level of knowledge on the noise and spectral shape priors!

Some [preliminary] findings

Another example on the importance of our level of knowledge on the noise and spectral shape priors!







See Quentins talk later!

Some [preliminary] **findings**

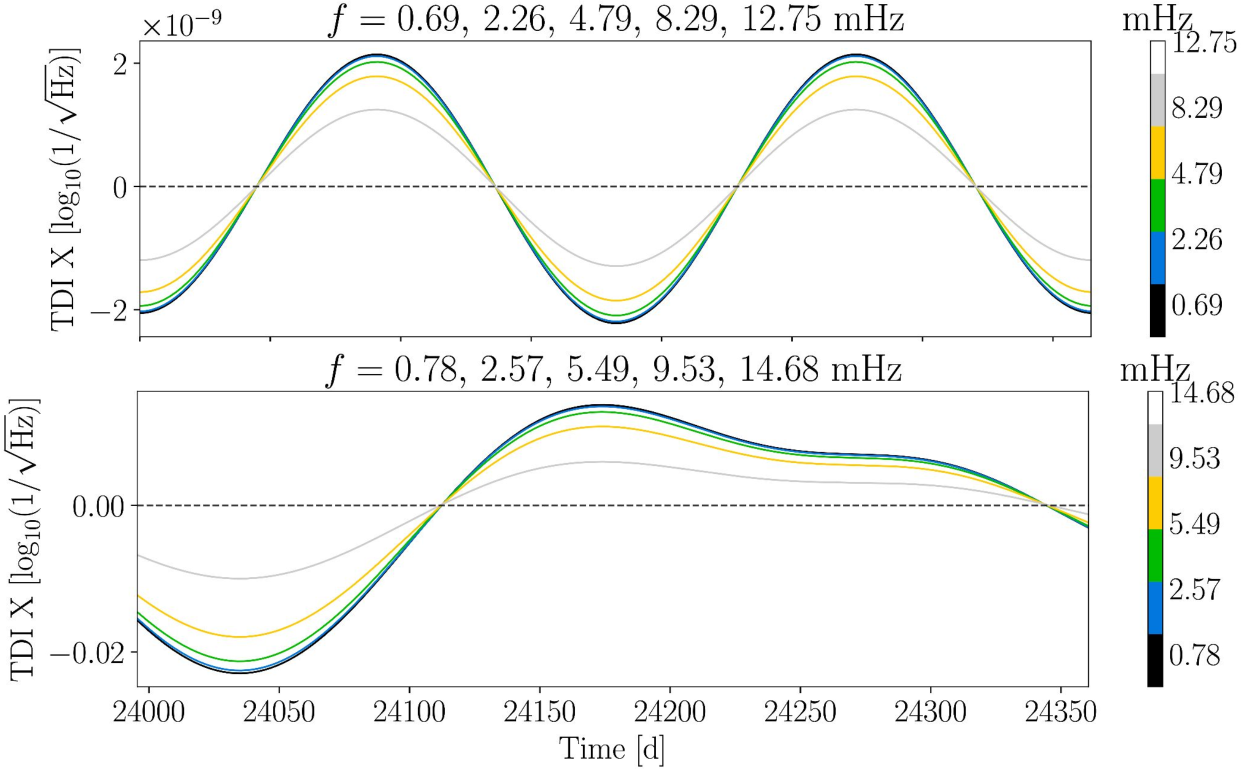
An example on the stationarity of the noise!

Some [preliminary] **findings**

Another example on the stationarity of the noise! Deviations of PSD value from t0 at given f.

Analytic orbits

ESA orbits



See Quentins talk later!

Summary

- Effort to compare the different methods **[Ongoing]**
- Assess impact of modelling (flexible, type, analytic) to the detectability of SGWB signals **[Ongoing]**
- All the above in a realistic, but controlled setup **[Ongoing]**

Questions to answer

- Although not part of this study, we need to include astrophysical components.
    - Time variations become *very* relevant. Non-stationarity is our next step.
- How do priors impact the analysis? Stay tuned to find out!
- Understand statistical assumptions about the data!
- Understand the performance of pipelines that adopt flexible spectral modelling techniques.

To do list

- Finalize the analysis of all datasets and record strengths and limitations of all approaches.
- Basically figure out *which* method is best for *which* task!