LISA Global Fit:
Searching for stochastic signals in such a scheme

Nikolaos Karnesis
Aristotle University of Thessaloniki
17/06/2024

LISA Cosmology WG Workshop 2024

[arXiv:2405.04690] Katz, NK, Korsakova, Gair, Stergioulas

An efficient GPU-accelerated multi-source global fit pipeline for LISA data analysis

Michael L. Katz,[1,2,*] Nikolaos Karnesis,[3] Natalia Korsakova,[4]
Jonathan R. Gair,[2] and Nikolaos Stergioulas[3]

[1] *NASA Marshall Space Flight Center, Huntsville, Alabama 35811, USA*
[2] *Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany*
[3] *Department of Physics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece*
[4] *Astroparticule et Cosmologie, CNRS, Université Paris Cité, F-75013 Paris, France*

(Dated: May 9, 2024)

The large-scale analysis task of deciphering gravitational wave signals in the LISA data stream will be difficult, requiring a large amount of computational resources and extensive development of computational methods. Its high dimensionality, multiple model types, and complicated noise profile require a global fit to all parameters and input models simultaneously. In this work, we detail our global fit algorithm, called "Erebor," designed to accomplish this challenging task. It is capable of analysing current state-of-the-art datasets and then growing into the future as more pieces of the pipeline are completed and added. We describe our pipeline strategy, the algorithmic setup, and the results from our analysis of the LDC2A Sangria dataset, which contains Massive Black Hole Binaries, compact Galactic Binaries, and a parameterized noise spectrum whose parameters are unknown to the user. We recover posterior distributions for all 15 (6) of the injected MBHBs in the LDC2A training (hidden) dataset. We catalog $\sim$ 12000 Galactic Binaries ($\sim$ 8000 as high confidence detections) for both the training and hidden datasets. All of the sources and their posterior distributions are provided in publicly available catalogs.

I. INTRODUCTION

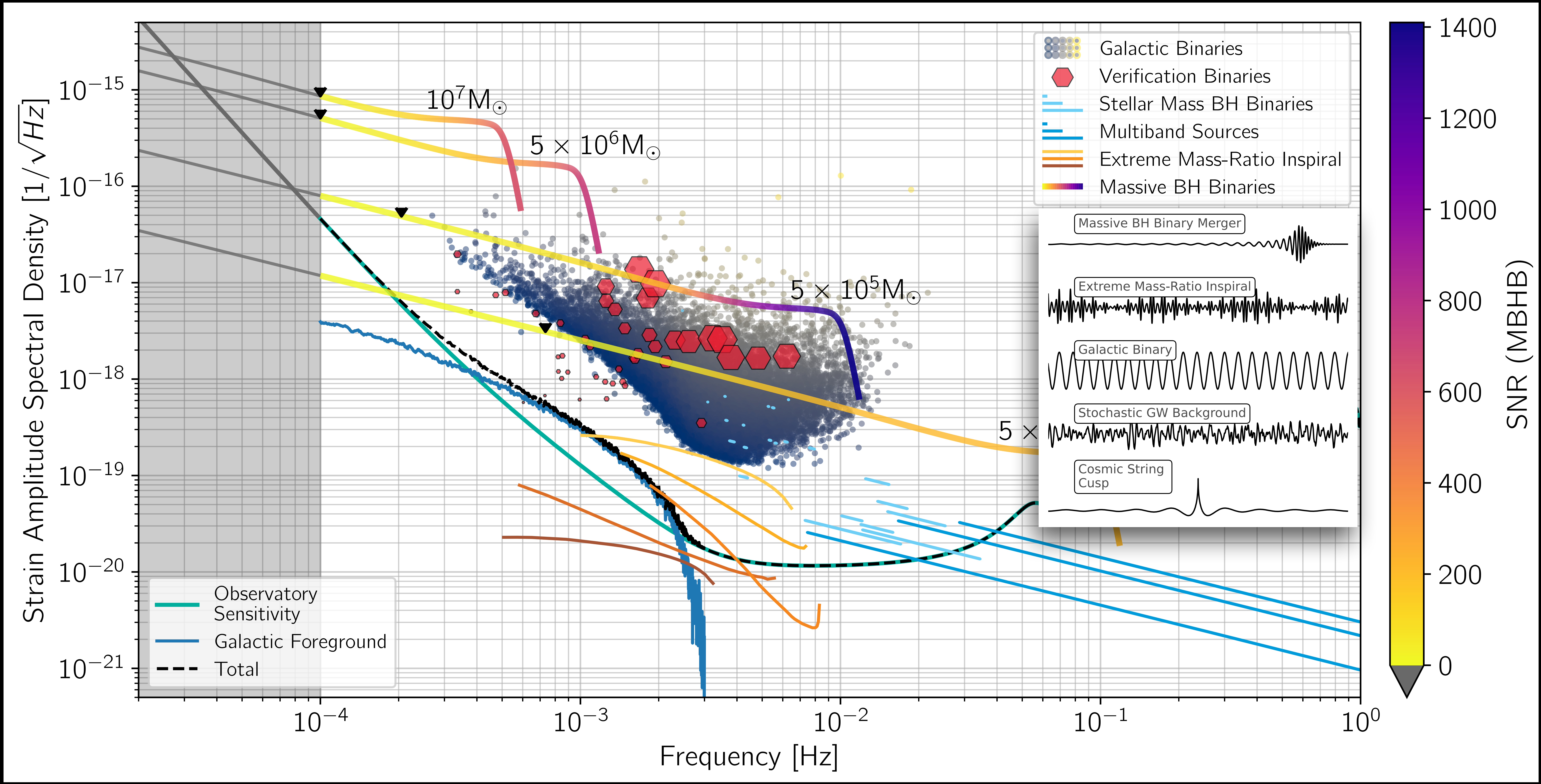
In the mid-2030s, the Laser Interferometer Space Antenna (LISA) will launch into space to measure gravitational waves emanating from a variety of astrophysical and cosmological sources in the millihertz frequency band [1]. LISA will add important information to the gravitational-wave spectrum, following on the discoveries of ground-based observing networks at higher frequencies [2] and pulsar timing arrays at lower frequencies [3]. LISA will detect different astrophysical sources: compact object binaries, also referred to as Galactic Binaries (GB), inside and close to the Milky Way Galaxy in slowly evolving orbits [e.g., 1, 4–7];

glitches [13–17]. Additionally, the orbit and performance of the LISA experiment will need to be included in any global analysis [18–21]. There will also be an initial data reduction pipeline that must be run properly prior to performing the scientific data analysis [22].

This complicated analysis requires a global fit over all parameters characterising the models for the signals and the noise. The design of this type of pipeline has been, and will continue to be, a large developmental project requiring many participating groups and expertise. LISA global fit pipelines are likely to be built on the concept of the global fit "wheel": separate modules specifically designed for different source or instrumental analyses will run in parallel and communicate with each other

LISA Global Fit

And why we do it?



LISA Science Definition Document

LISA Global Fit

And why we do it?

- Correlations between sources become important for that many signals
- Computational reasons: sequential fits are inefficient
- Grid searches are almost impossible
- Imperfect source subtraction yields imperfect residuals
- Uncertainties propagation
- Not fixed dimensions!

Also see talk by R. Buscicchio

Erebor

An efficient GPU-accelerated multi-source global fit pipeline for LISA data analysis

[arXiv:2405.04690] Katz, NK, Korsakova, Gair, Stergioulas

Erebor

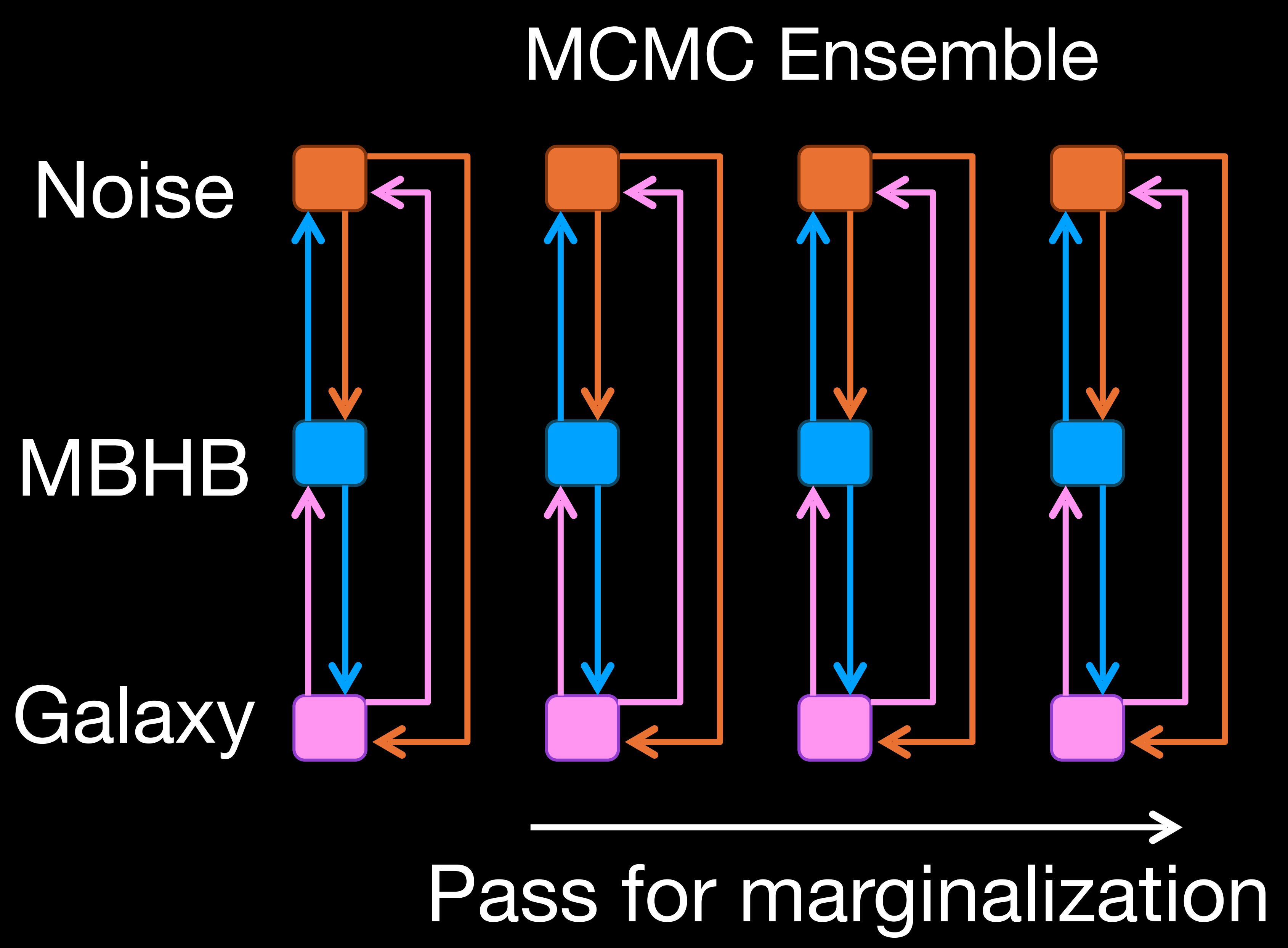
An efficient GPU-accelerated multi-source global fit pipeline for LISA data analysis

[arXiv:2405.04690] Katz, NK, Korsakova, Gair, Stergioulas

- Accomplished LDC2A analysis
- Publicly available code given in Data/Code availability statement in paper.
- Published output catalogs readable by `lisacattools` (all datasets)
- Clean up code, document, tutorials, full open-source code.

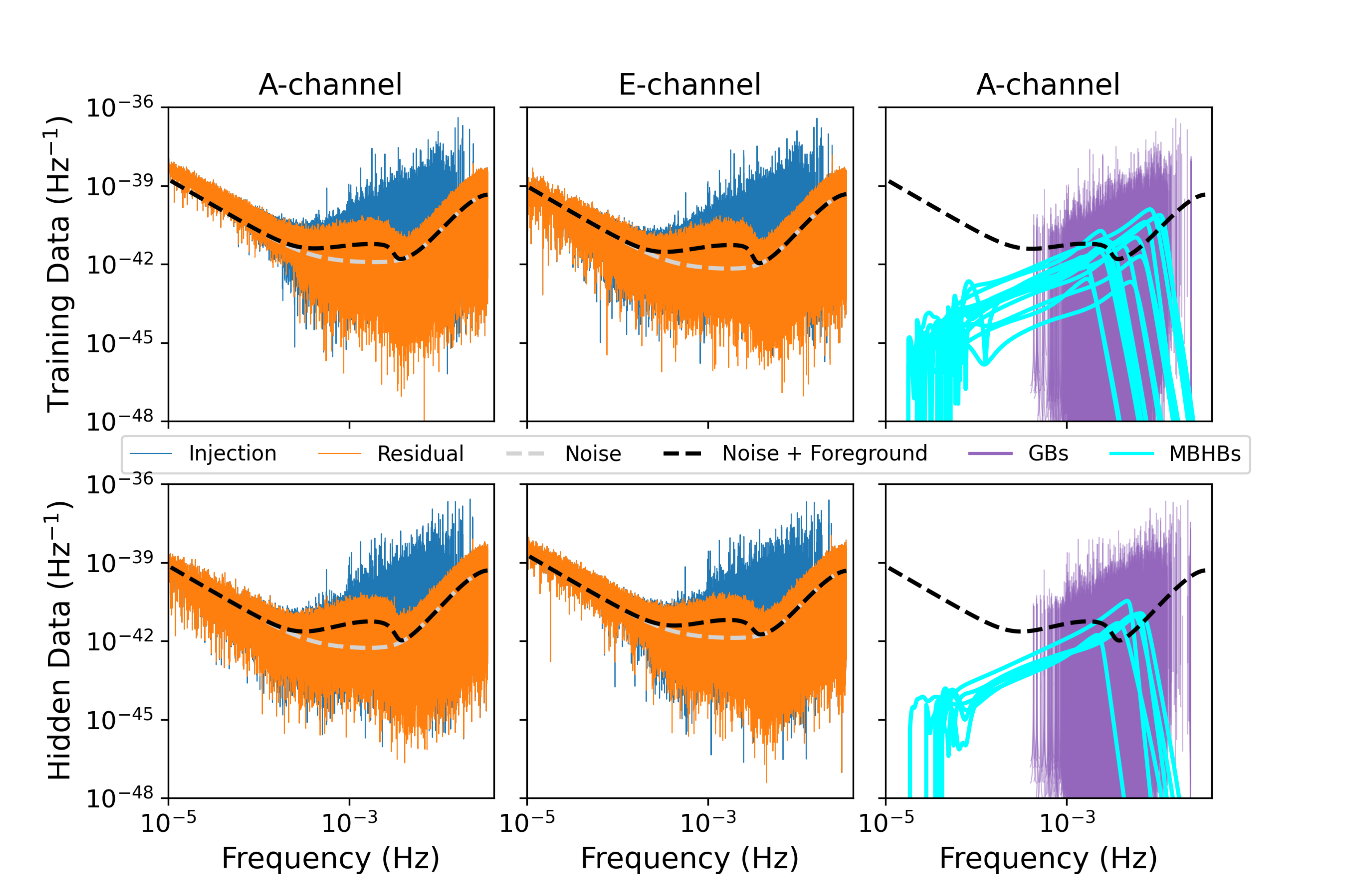
Main differences of Erebor

compared to efforts so far

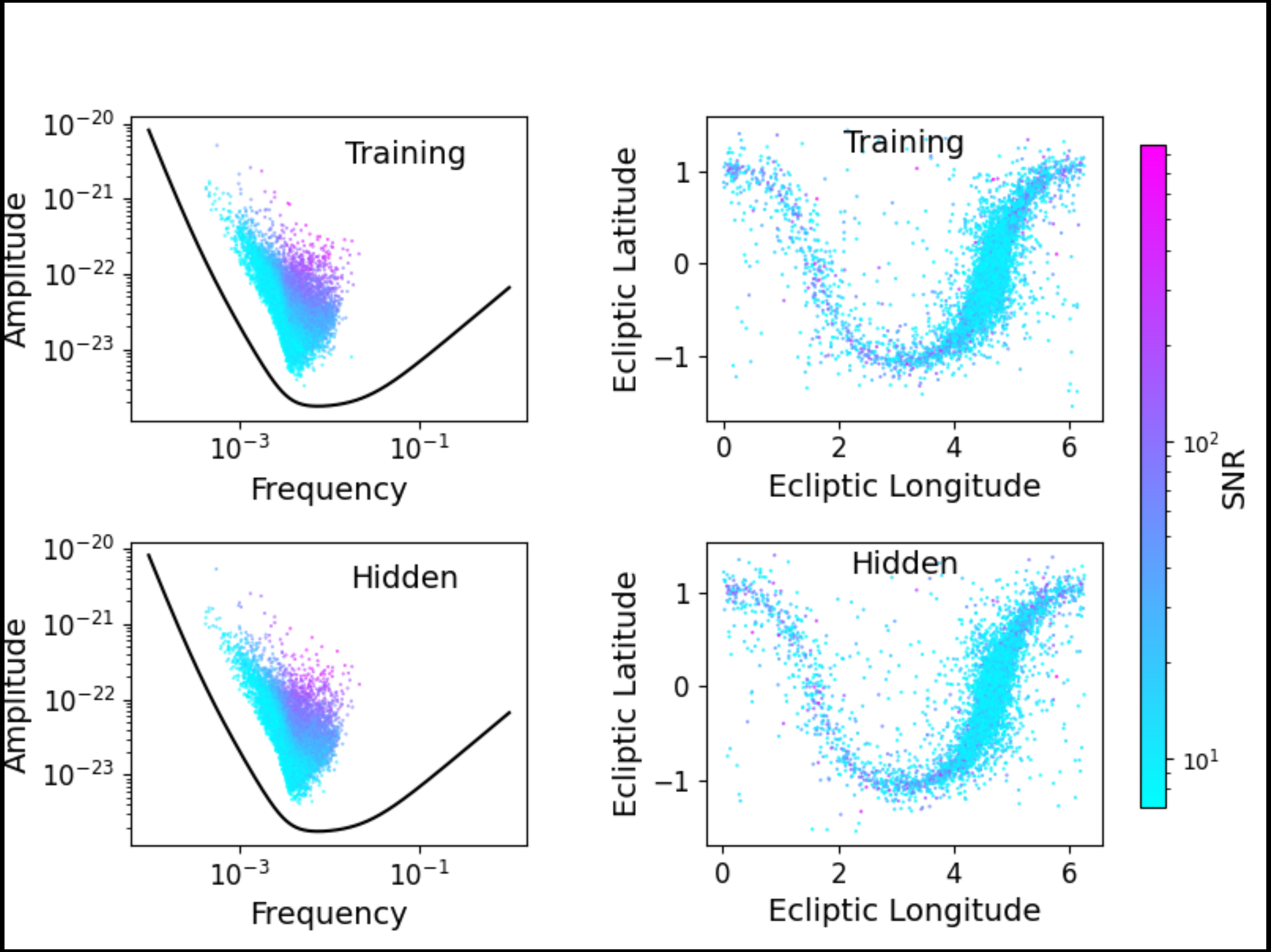


- GPUs (~5x cost improvement [?])
- Ensemble sampling
- Online Reversible Jump Proposal updates. [GMM on past samples]
- Single-source MCMC used for Reversible Jump proposal distribution.
- No time build up yet, directly run on the full year-long data.

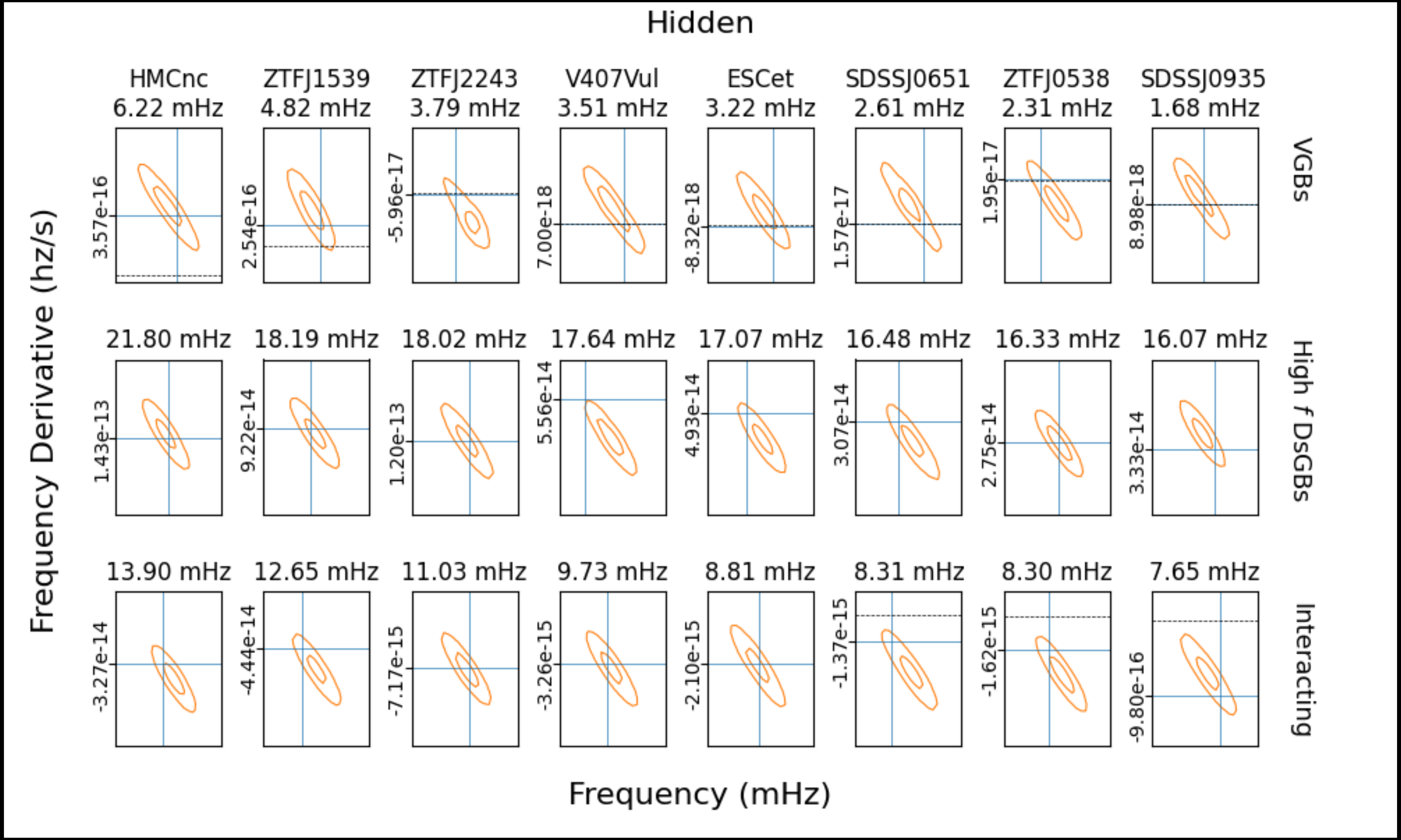
Some results [General overview]



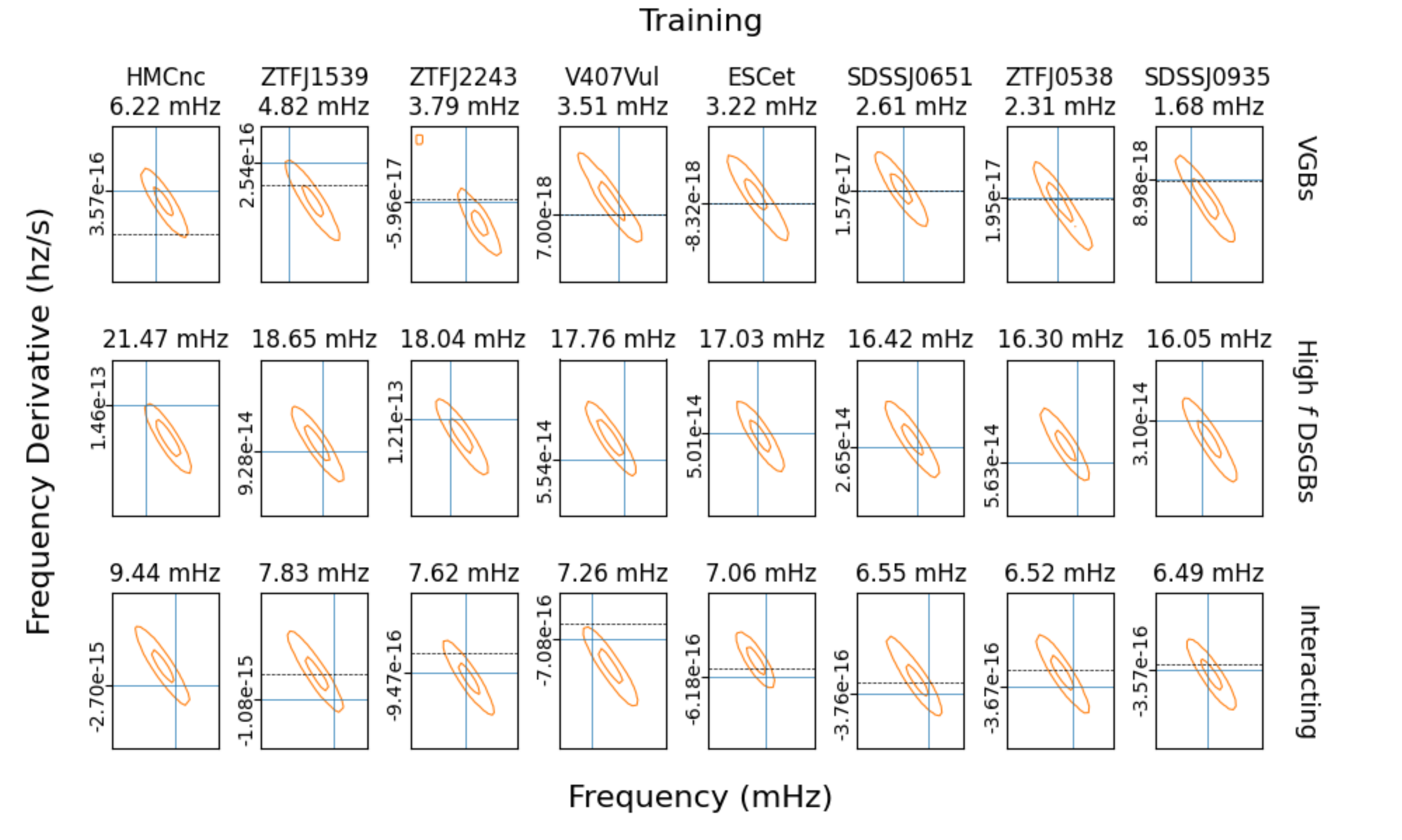
Some results [DWDs]



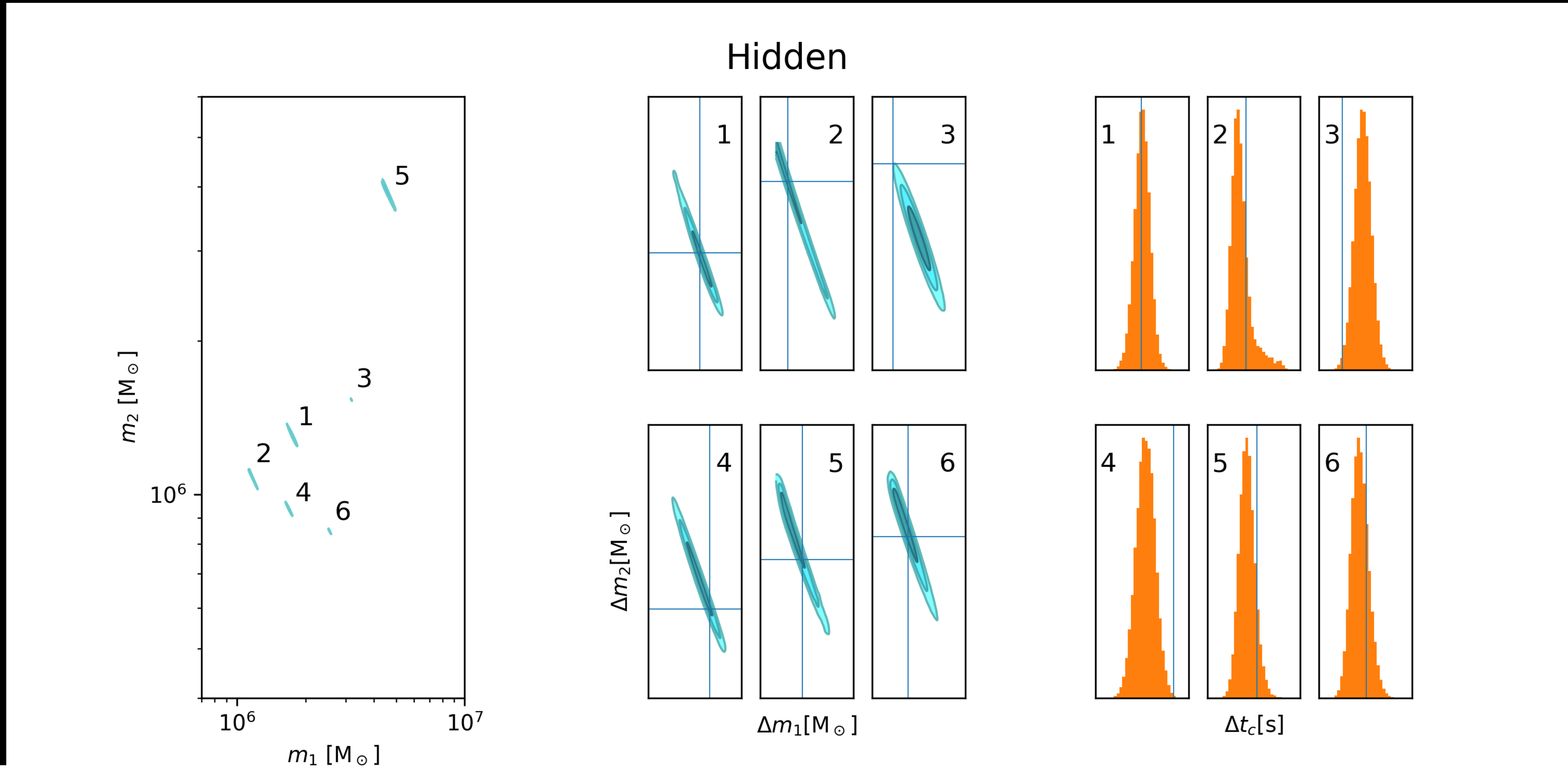
Some results [DWDs]

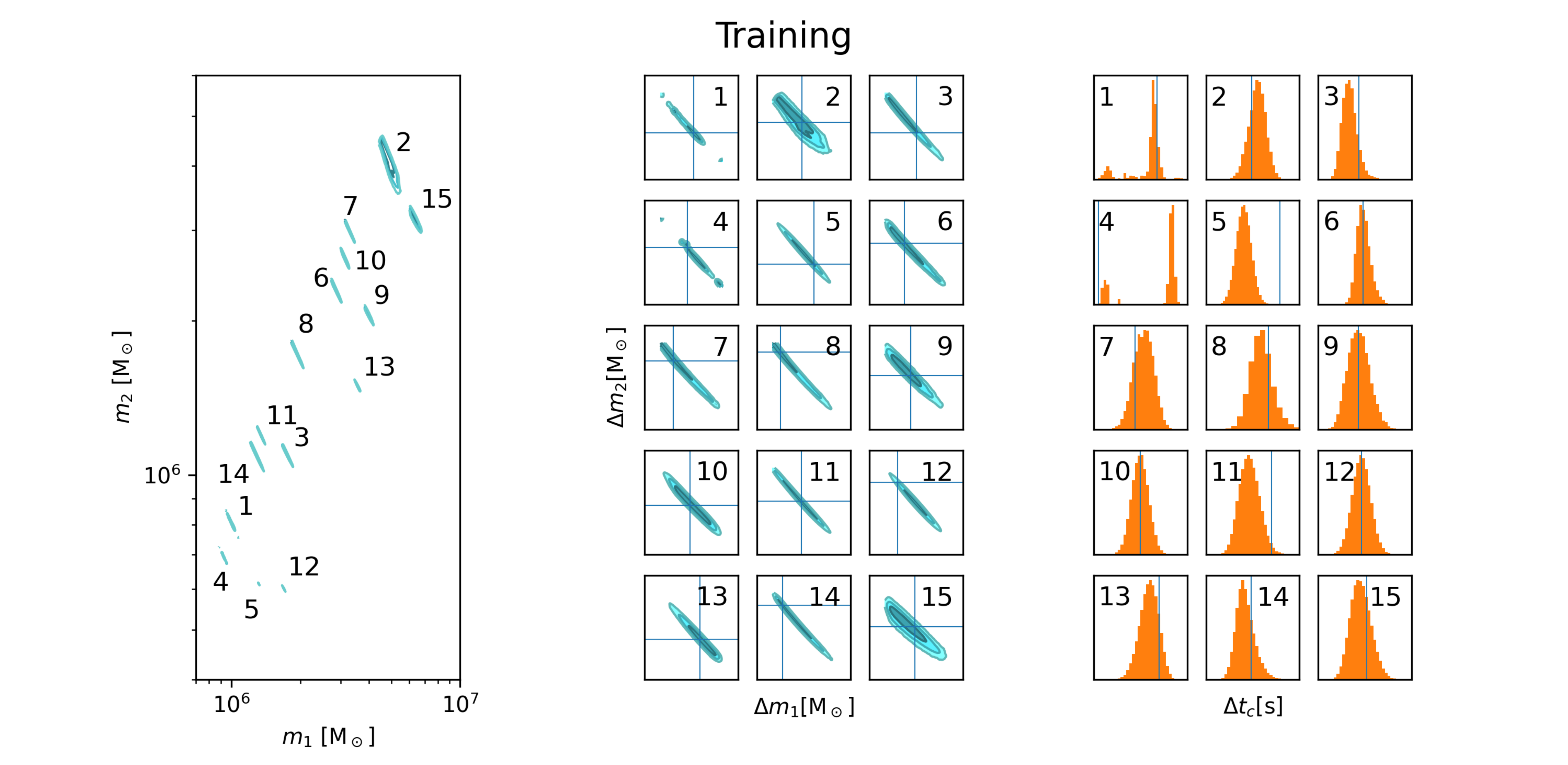


Some results [DWDs]



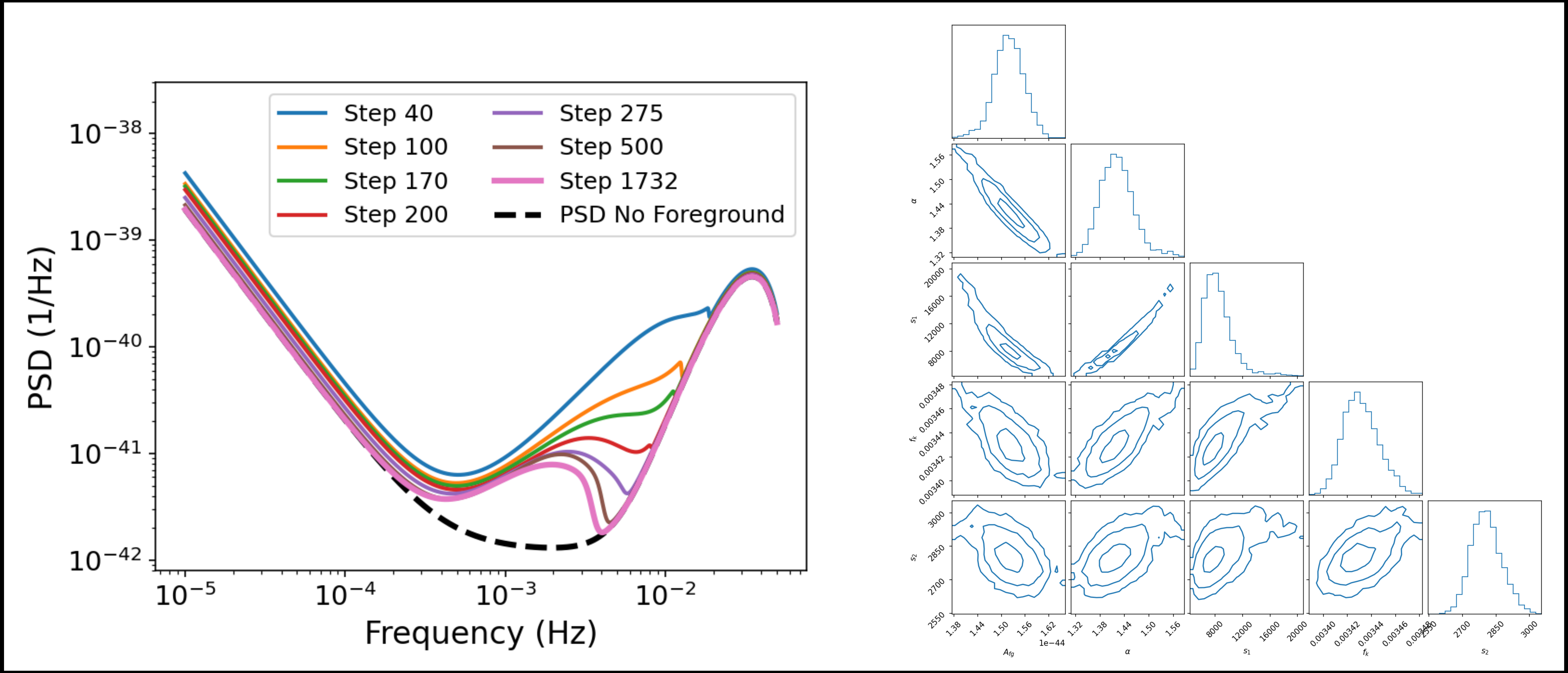
Some results [MBHBs]





Some results [Confusion]

$$S_{\text{gal}} = \frac{A}{2} e^{-(\mathrm{f}/\mathrm{f}_1)^{\alpha}} \mathrm{f}^{-7/3} \left(1 + \tanh\left(\frac{\mathrm{f}_{\text{knee}} - \mathrm{f}}{\mathrm{f}_2}\right)\right)$$



Discussion on the stochastic part

Lessons learned and future work

Disclaimer

- The purpose of this work was to focus on the transient and monochromatic source separation problem.

  [which is the first one to solve]

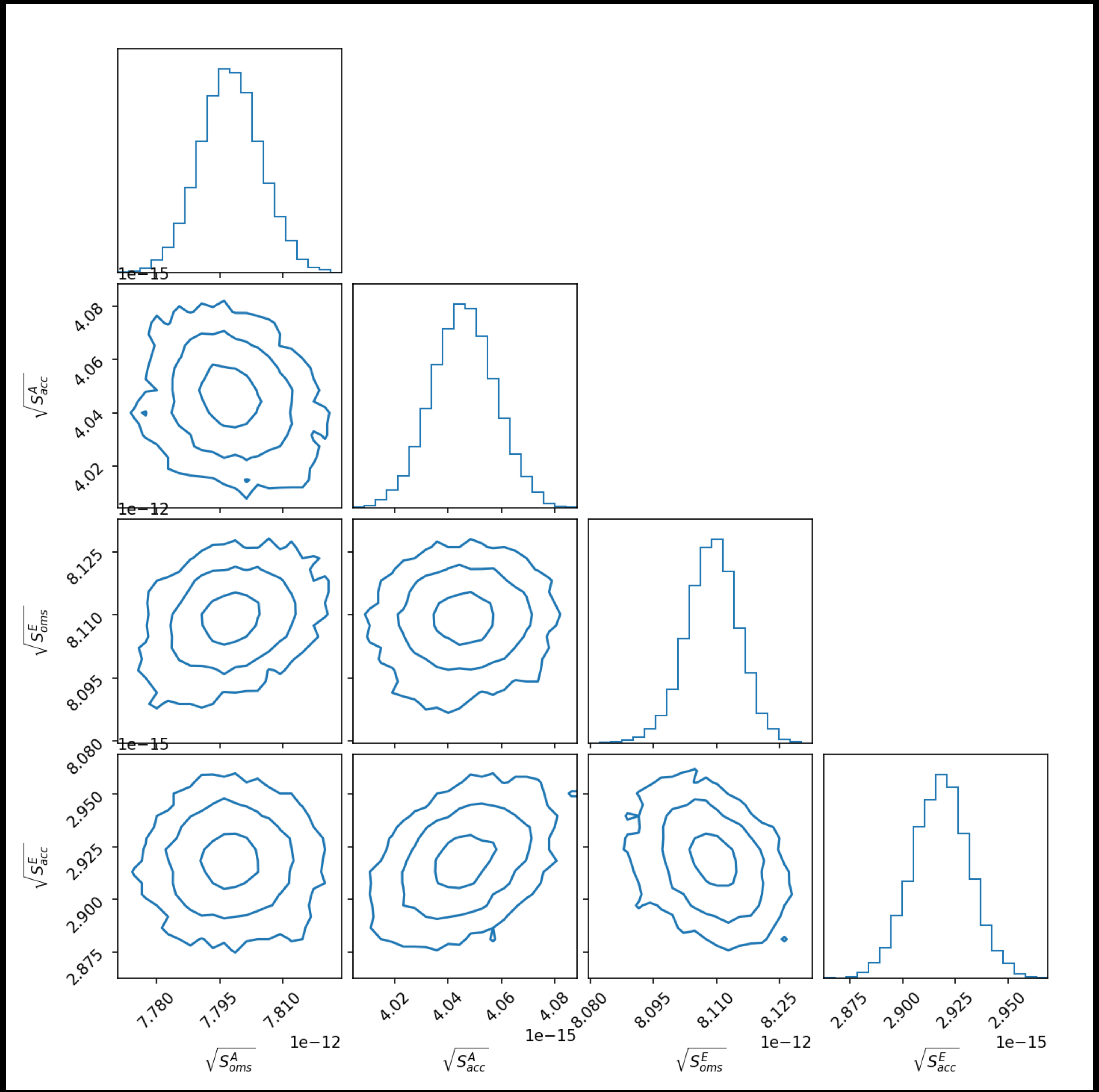
- Enhancements, improvements, and updates will follow.

Discussion on the stochastic part

Lessons learned and future work

A. The instrumental noise

- Analytic model: Simple two-parameter model, fitted separately for the two channels.
- Limited spectral shape flexibility.
- The different parameters for the two channels.
- No time dependence.
  [OK though, for this particular dataset]
- Some minor "*features*" are being treated now, updated results to be released.



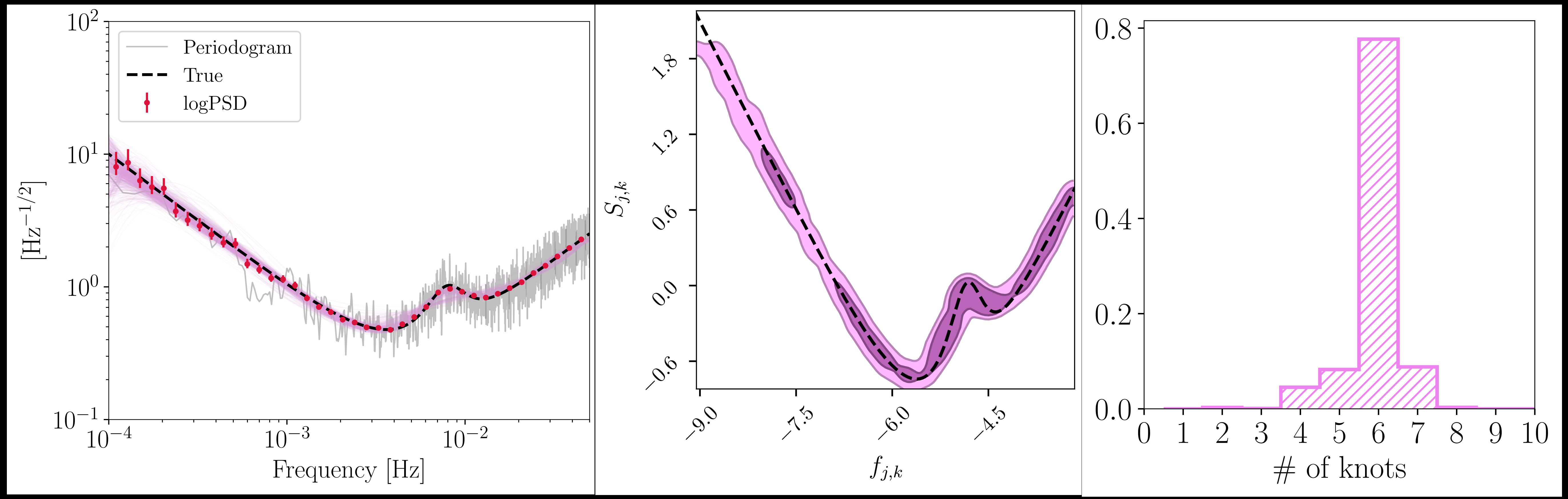
[Phys. Rev. D 107, 063004, 2023] Littenberg, Cornish

Discussion on the stochastic part

Lessons learned and future work

B. The astrophysical confusion noise

- Analytic model: While flexible enough for the “bump”, not suitable for data with other features. For example we could follow [Phys. Rev. D 107, 063004, 2023] by Littenberg, Cornish.



[MNRAS, 526, 4, 2023] NK, Katz+

- The same model was used for both channels.
- No time-variability assumed.
- Due to correlations, other stochastic signals searches *might need* to enter the loop.

Discussion on the stochastic part

Lessons learned and future work

- Collaborative project: that involves (in random order) **Pieroni**, **Muratore**, **Hartwig**, **Baghi**, **NK**, **Bayle**, **Caprini**, **Nardini**, **Dam Quang**, **Santini**, **Pozzoli**, **Buscicchio**, [...].
- Increase realism, take advantage of orbits, transfer functions, time variations, flexible models [...]

See talk by F. Pozzoli

Also see talk by R. Buscicchio

LISA global fitting stochastic signals

Summary & a step forward:

- Another block can be added in the blocked Gibbs loop!
- For example add searches of spectral shape-agnostic models.
- Or use global fit residuals for a hierarchical analysis [was discussed already here].
- Add time dependence: use orbits, or *fit* for it, make use of transfer functions […].
- Concerning the Erebor effort we need to add more source-blocks, improve existing modules, clean-up code and write examples.
- Again: code is intended to be public.

See talk by F. Pozzoli | See talk by H. Inchauspè | See talk by L. Speri | See talk by J. Kume

Extras

Some results [DWDs]

This work