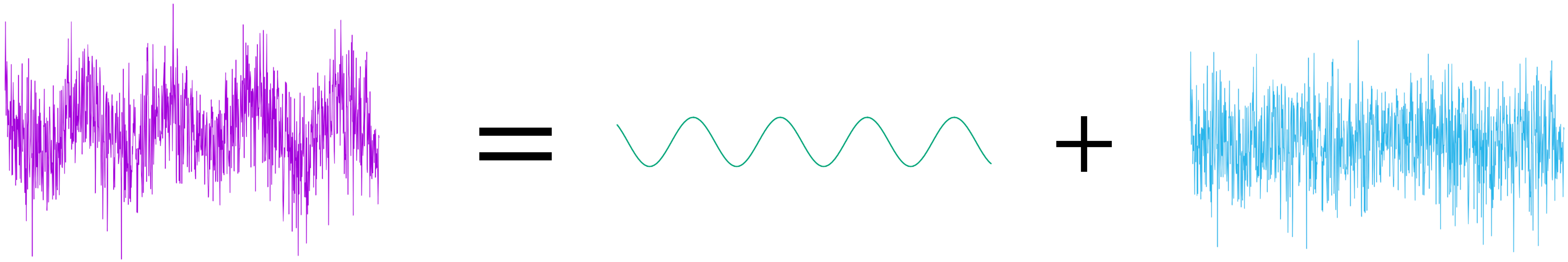
LISA Data Analysis - Global Fit

Neil J. Cornish

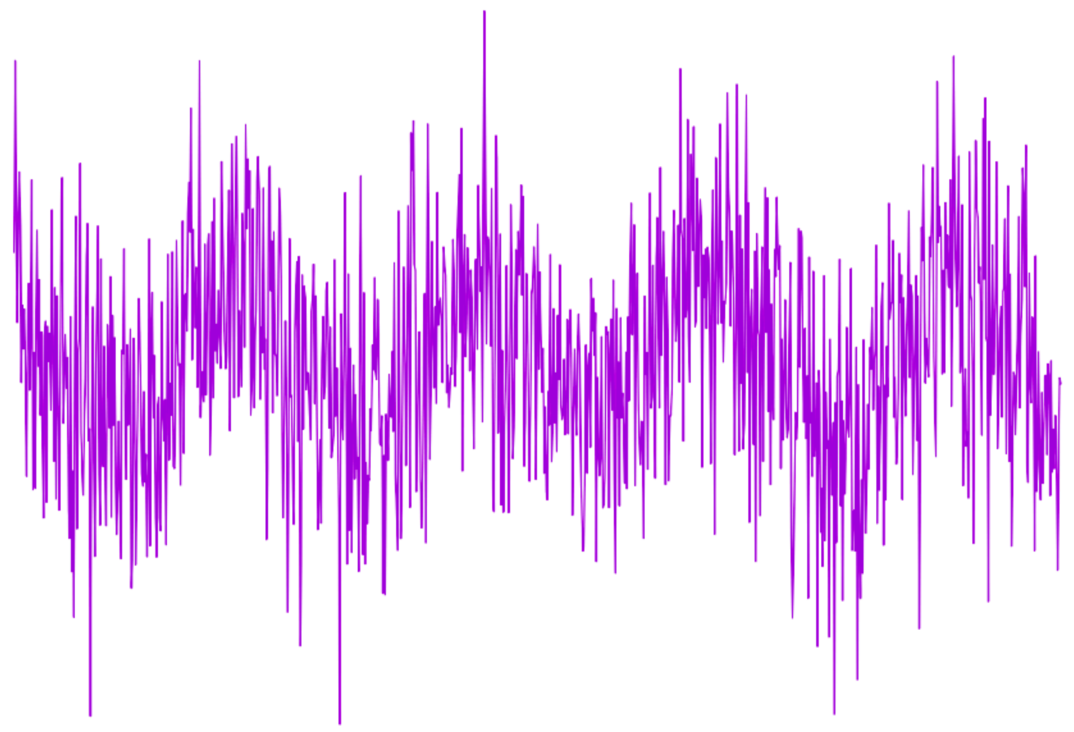
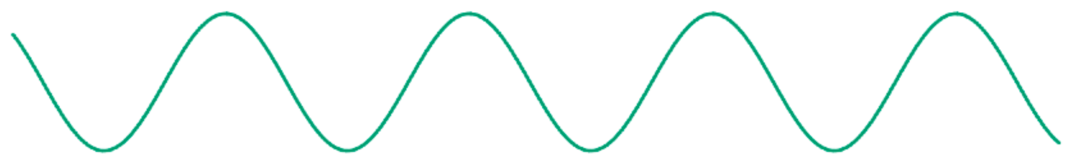
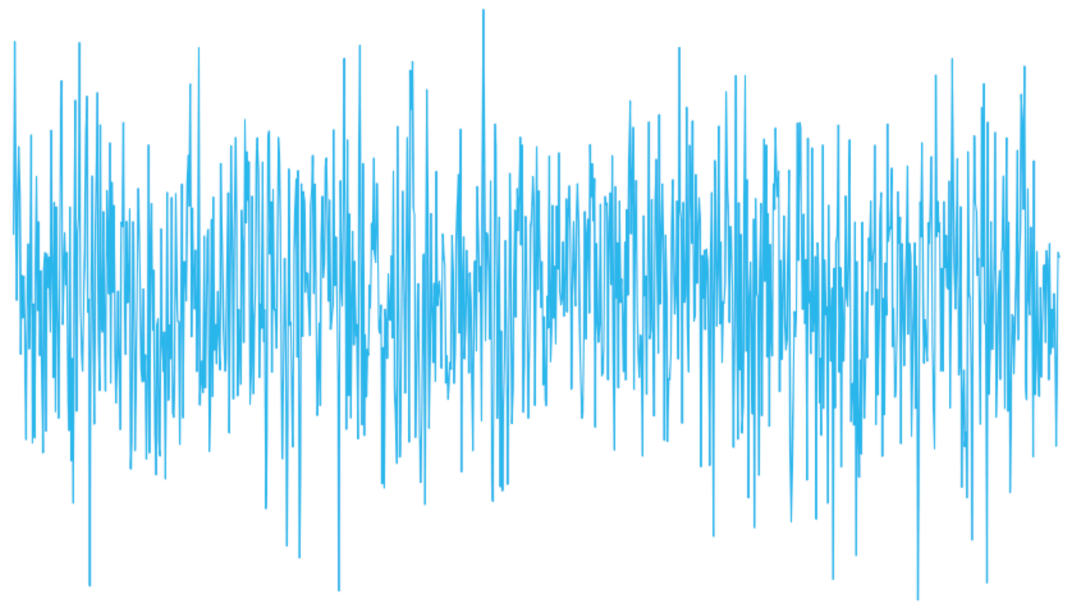
eXtreme Gravity Institute

MONTANA STATE UNIVERSITY

Data analysis 101



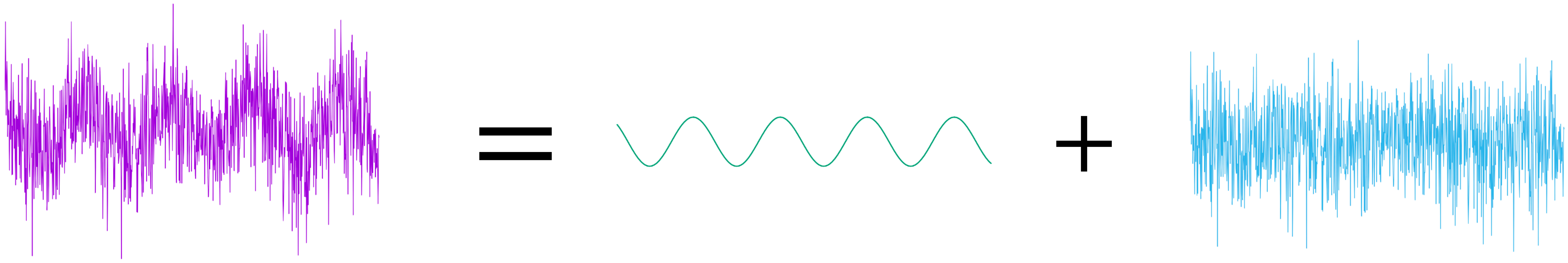
Data analysis 101

 $=$  $+$ 

$$\mathbf{d} = \mathbf{h} + \mathbf{n}$$

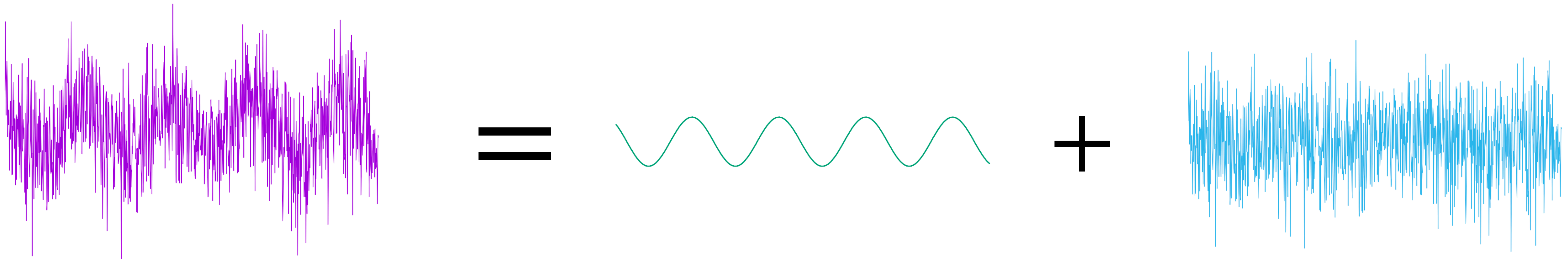
$$\Rightarrow \quad \mathbf{n} = \mathbf{d} - \mathbf{h}$$

Data analysis 101



Noise model $$p(\mathbf{n}) = \frac{1}{\sqrt{\det(2\pi\,\mathbf{C})}} e^{-\frac{1}{2}\mathbf{n}^{\dagger}\cdot\mathbf{C}^{-1}\cdot\mathbf{n}}$$

Data analysis 101



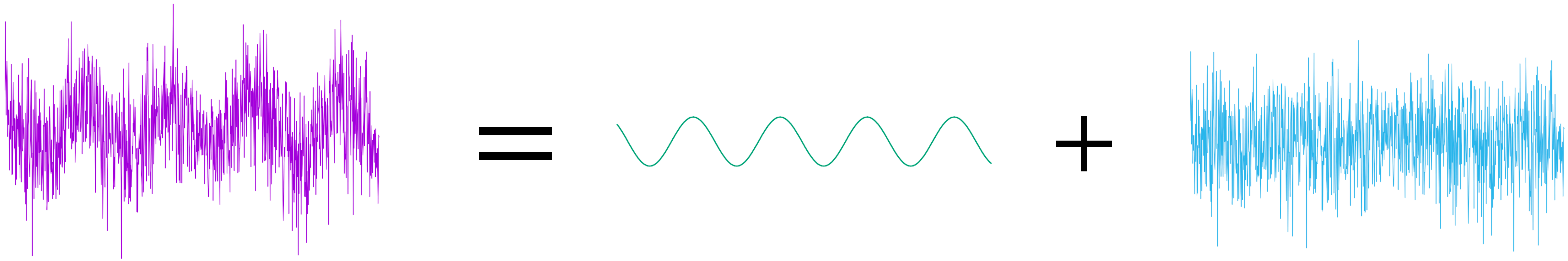
Noise model $$p(\mathbf{n}) = \frac{1}{\sqrt{\det(2\pi\,\mathbf{C})}} e^{-\frac{1}{2}\mathbf{n}^{\dagger}\cdot\mathbf{C}^{-1}\cdot\mathbf{n}}$$

Integrate over noise realizations

$$p(\mathbf{d}\,|\,\mathbf{h}) = \int p(\mathbf{n})\delta(\mathbf{n} - (\mathbf{d} - \mathbf{h}))\,d\mathbf{n} = \frac{1}{\sqrt{\det(2\pi\,\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^{\dagger}\cdot\mathbf{C}^{-1}\cdot(\mathbf{d}-\mathbf{h})}$$

Likelihood

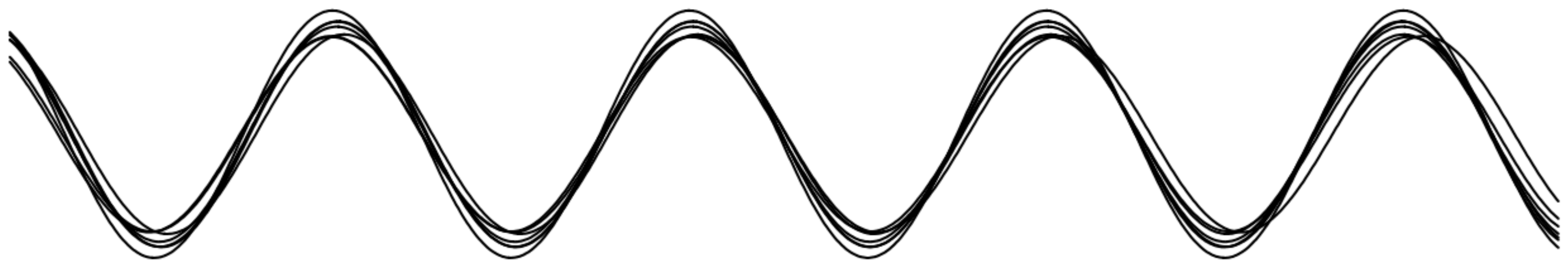
Data analysis 101



Bayes Theorem

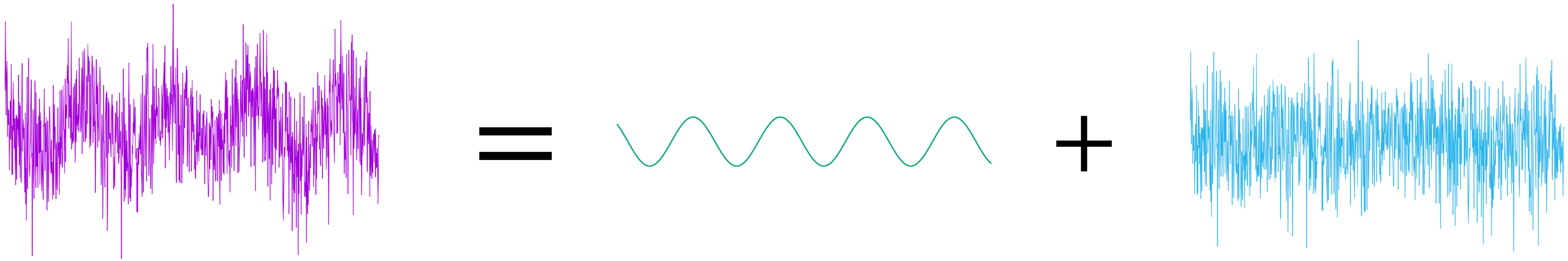
$$p(\mathbf{h}) = \frac{p(\mathbf{d}\,|\,\mathbf{h})\,p(\mathbf{h})}{p(\mathbf{d})}$$

Posterior draws

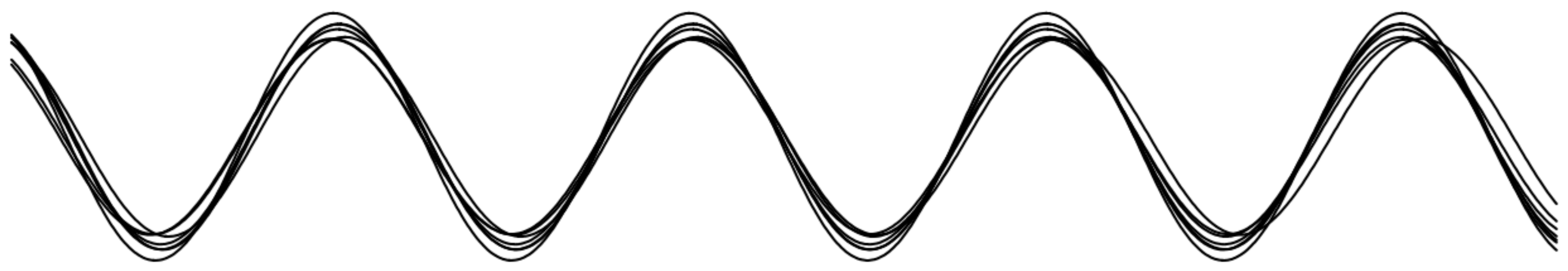


$$\sim \; p(\mathbf{h})$$

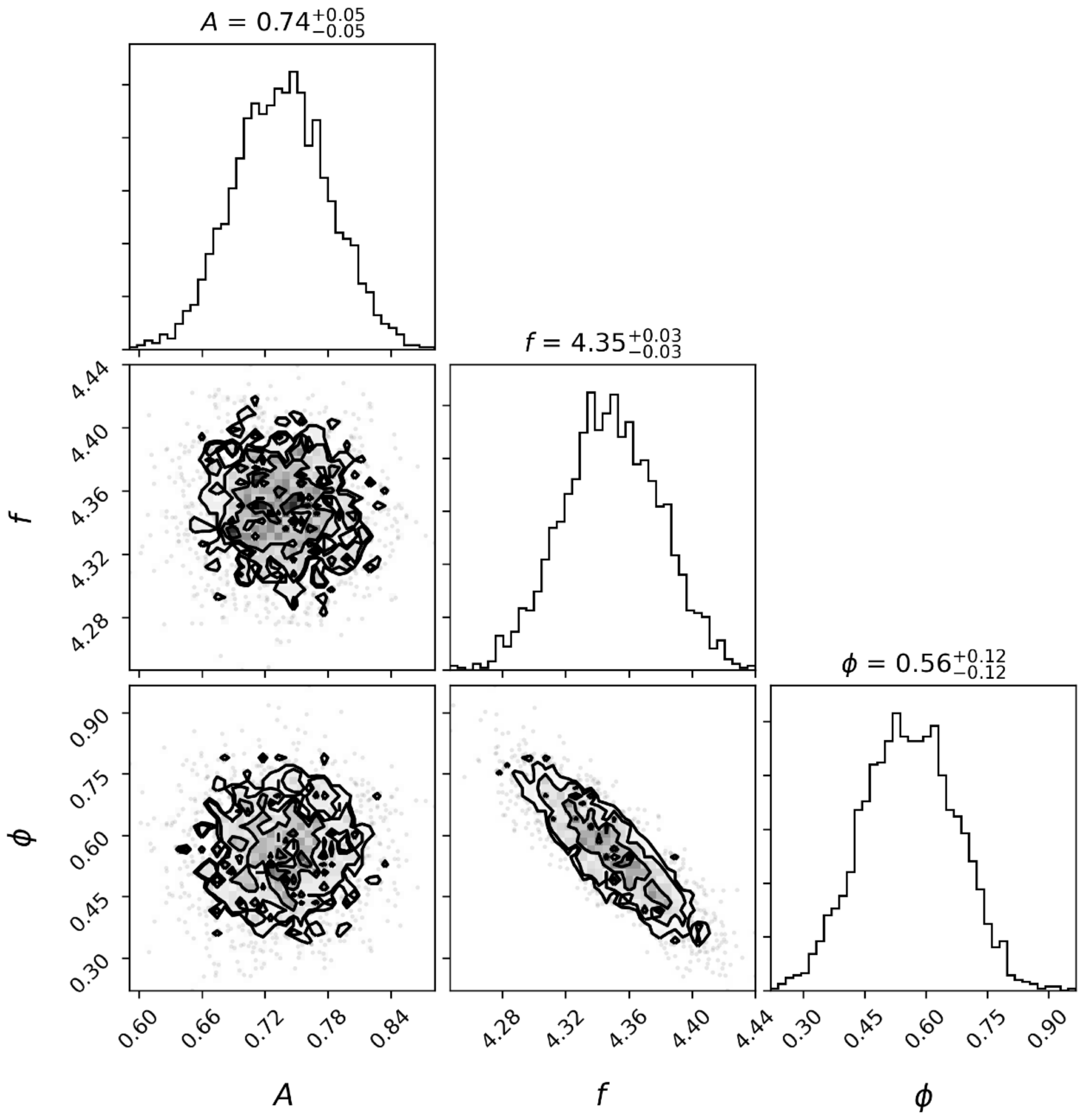
Data analysis 101



$$h(t, \vec{\lambda}) = A\ \cos(2\pi f t + \phi)$$

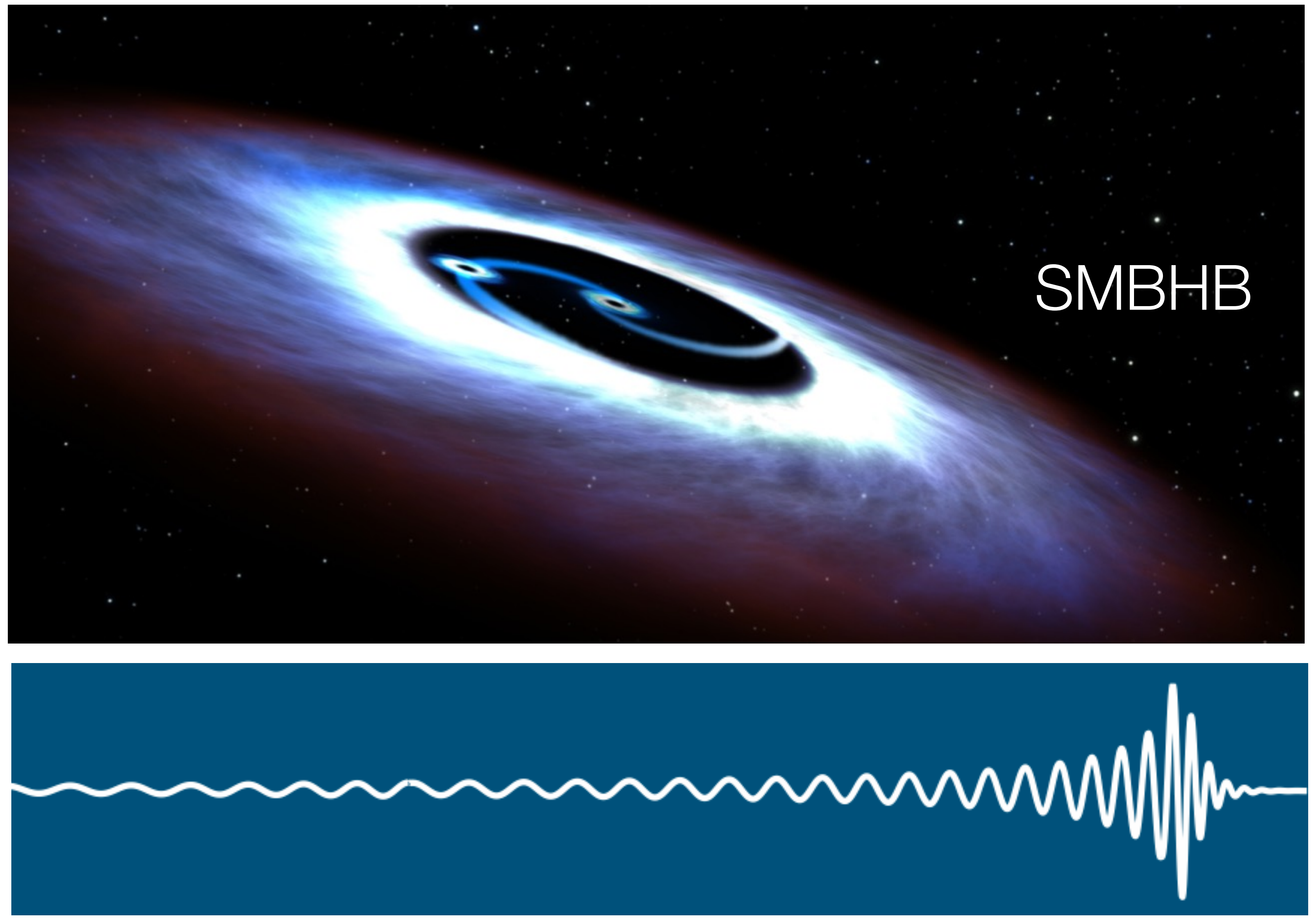


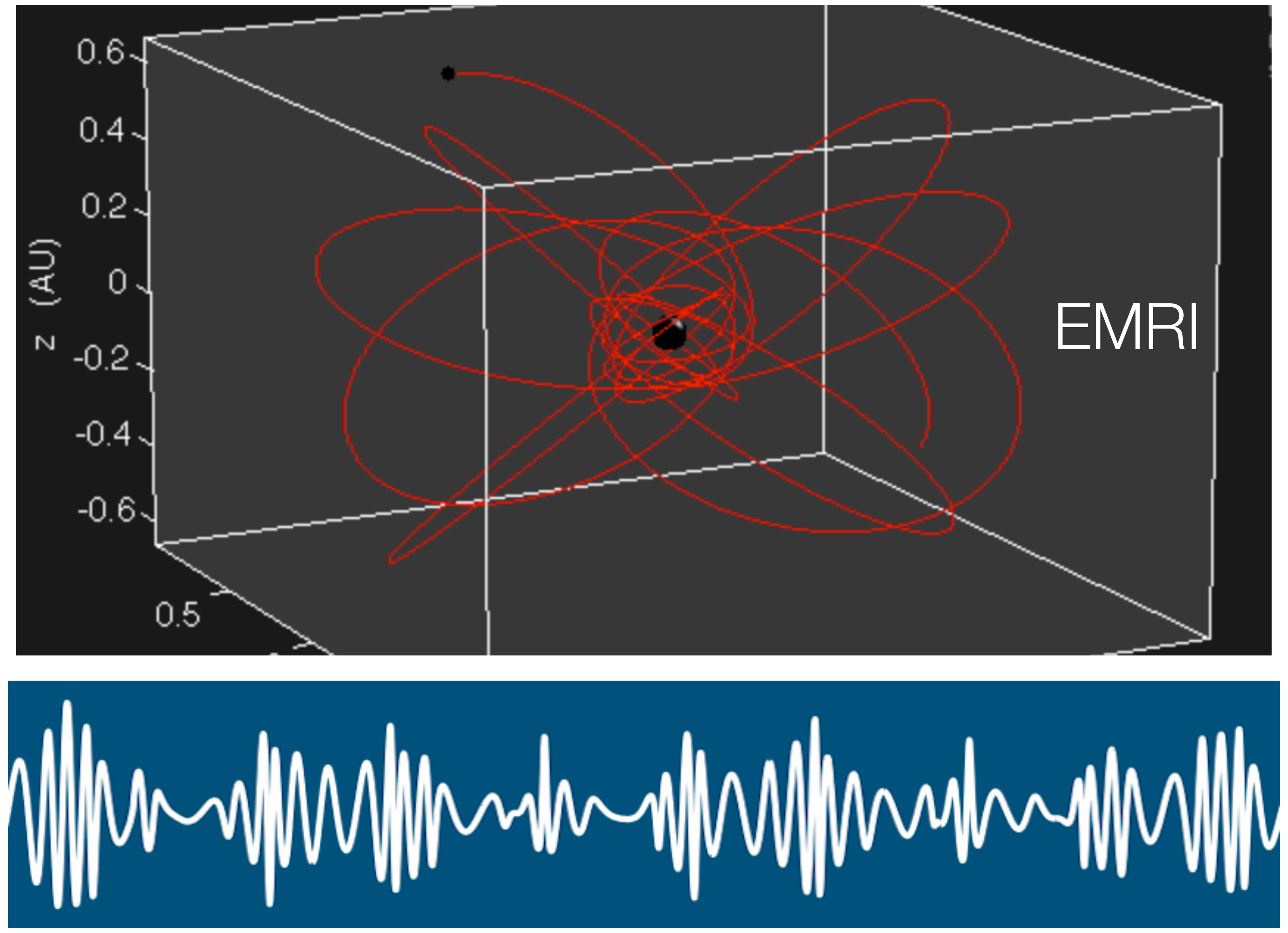
$$p(\mathbf{h}) \rightarrow p(\vec{\lambda})$$

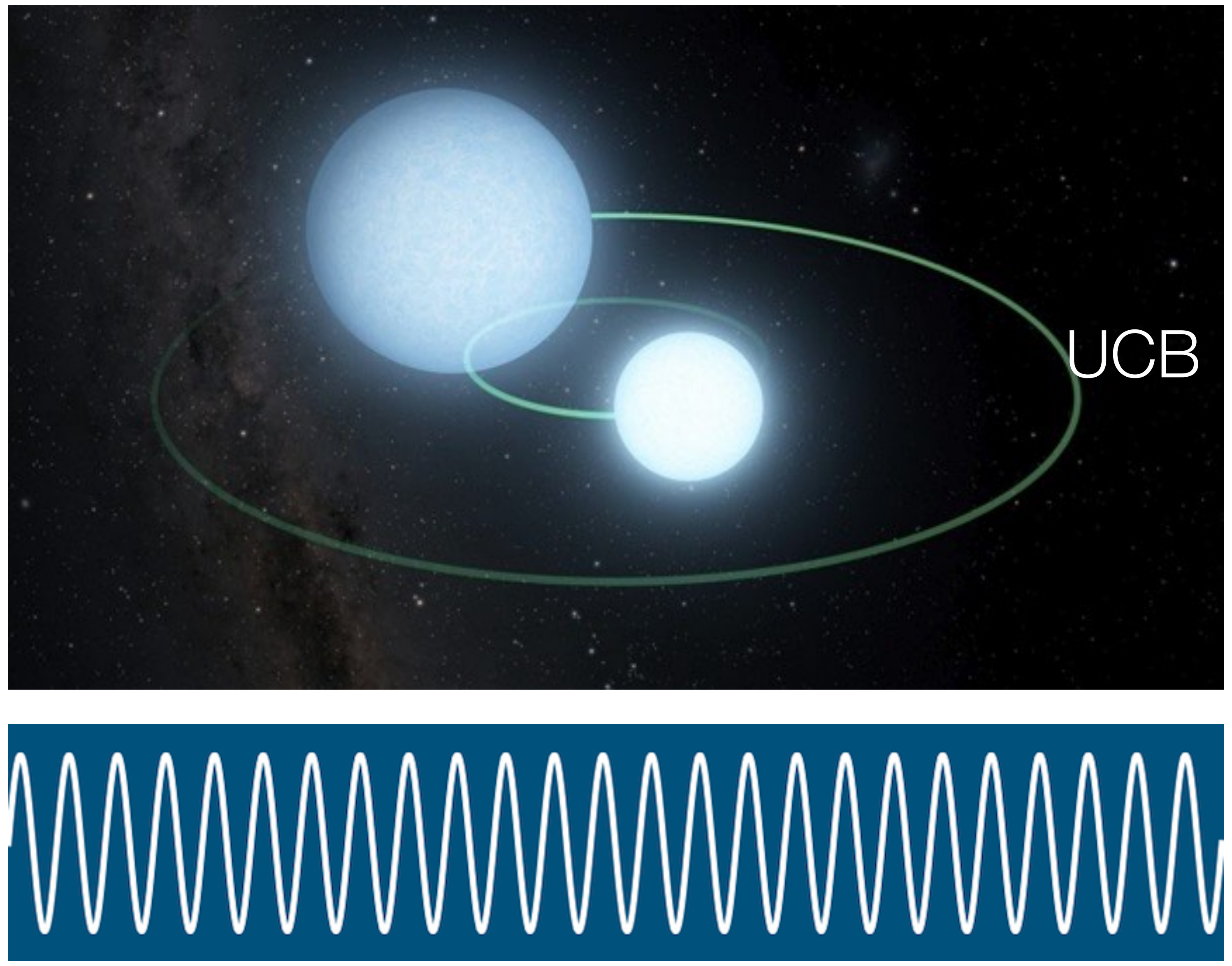


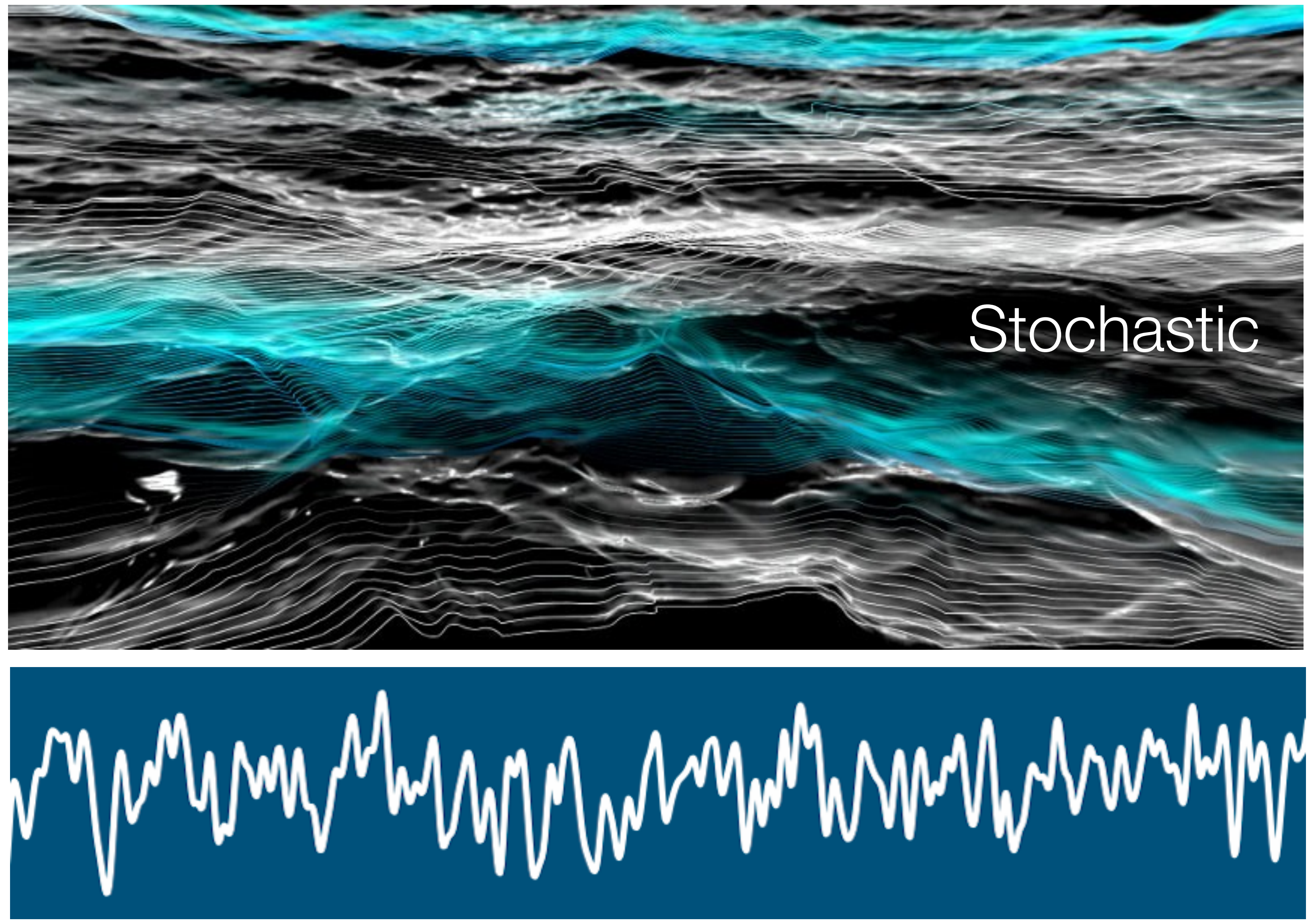
Same idea for LISA, just with hundreds of thousands of parameters

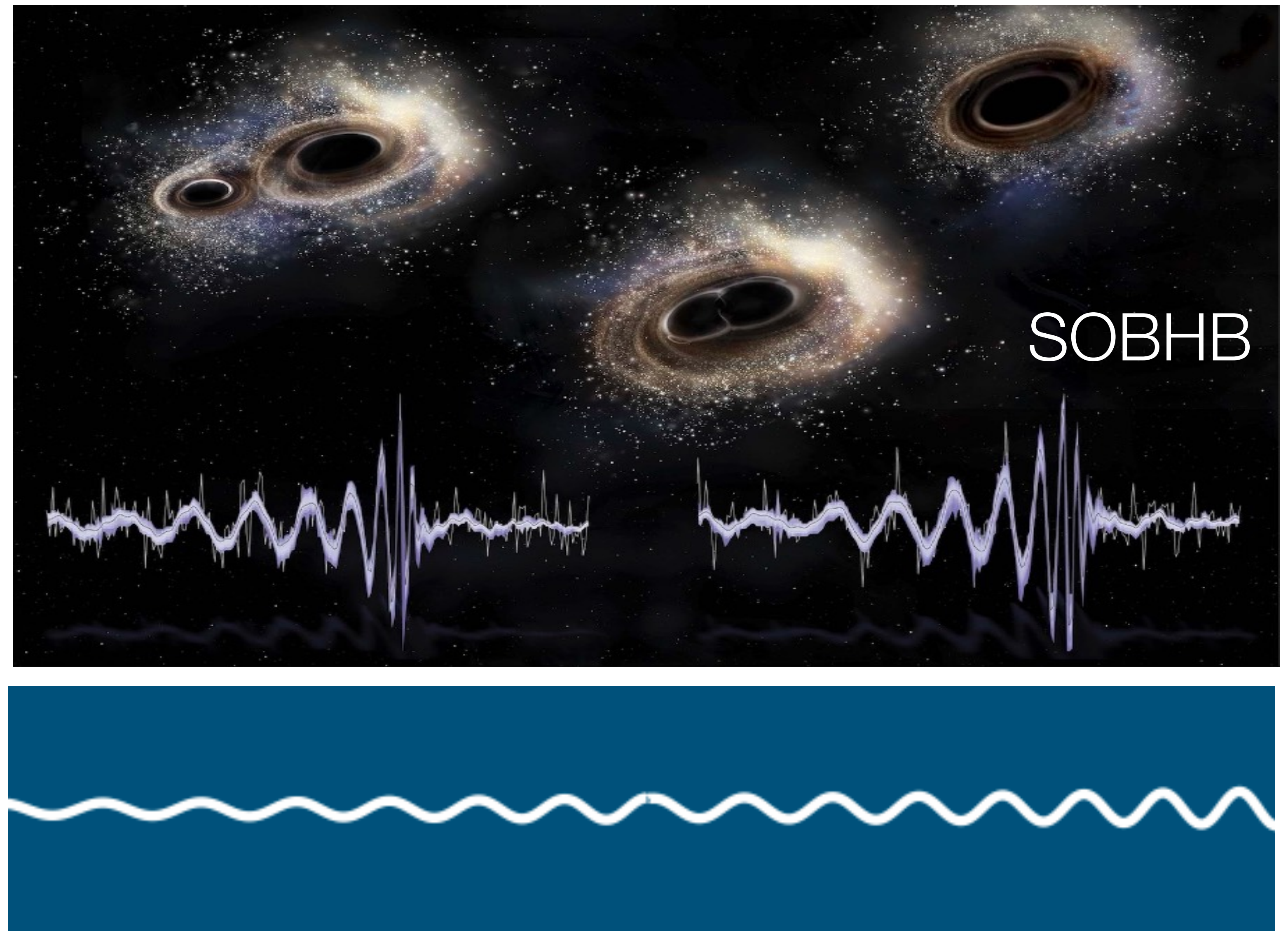
LISA Sources











?

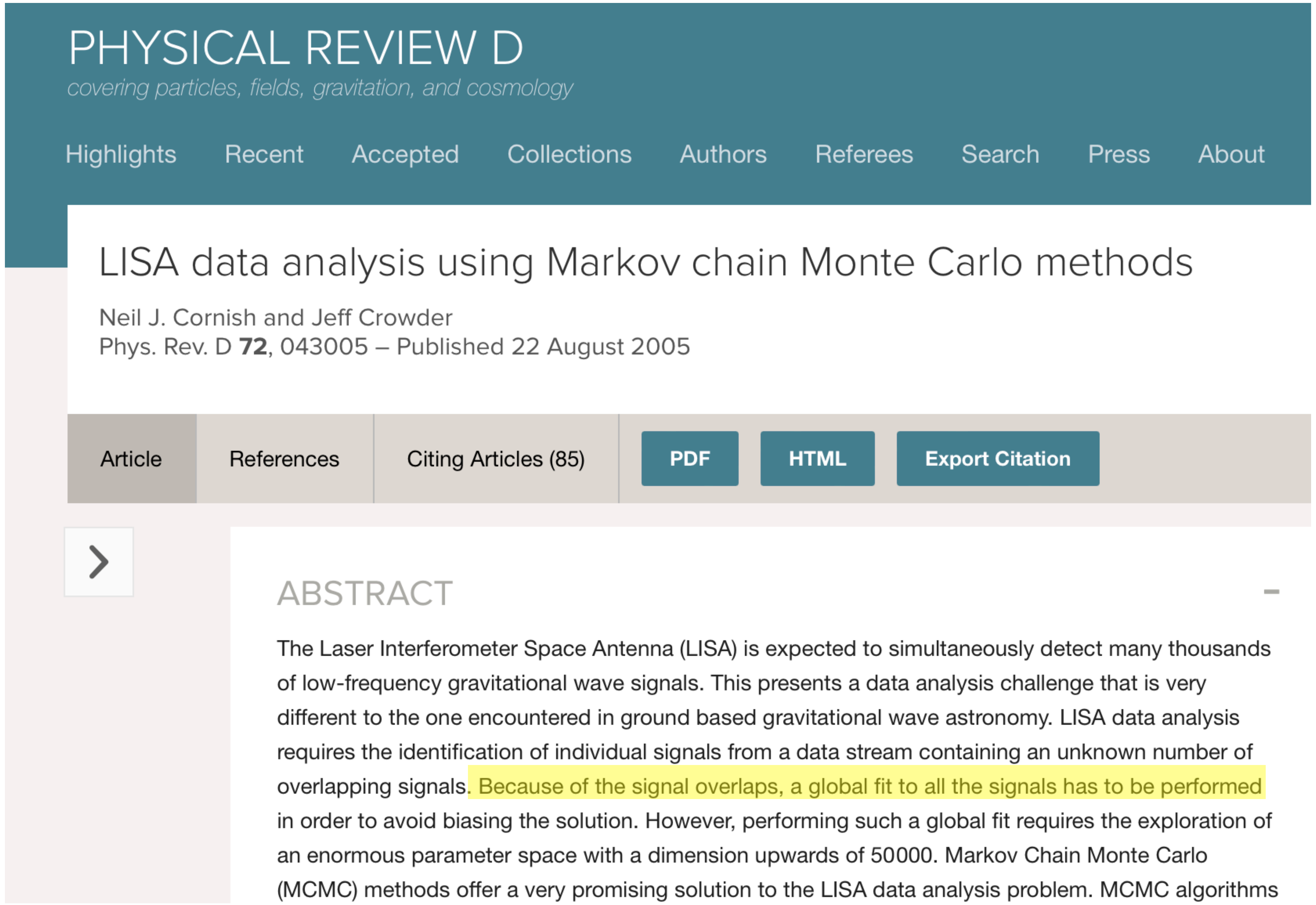
LIGO/Virgo

- Short duration, non overlapping
- Low Latency Search
  - Maximum likelihood inspired
  - Analyze short time segments
  - Grid based search, simple templates
- Longer latency Bayesian follow up
- Also Continuous Wave, Un-modeled and Stochastic searches

LISA

- Millions of overlapping signals
- High dimensional search space
  - Grid based searches impractical
  - Stochastic search methods
- Signal duration often comparable to mission lifetime
- Need a Global Fit: Binaries of all kinds, stochastic signals and un-modeled signals. All together

Because of the signal overlaps, a global fit to all the signals has to be performed


PHYSICAL REVIEW D
covering particles, fields, gravitation, and cosmology

Highlights Recent Accepted Collections Authors Referees Search Press About

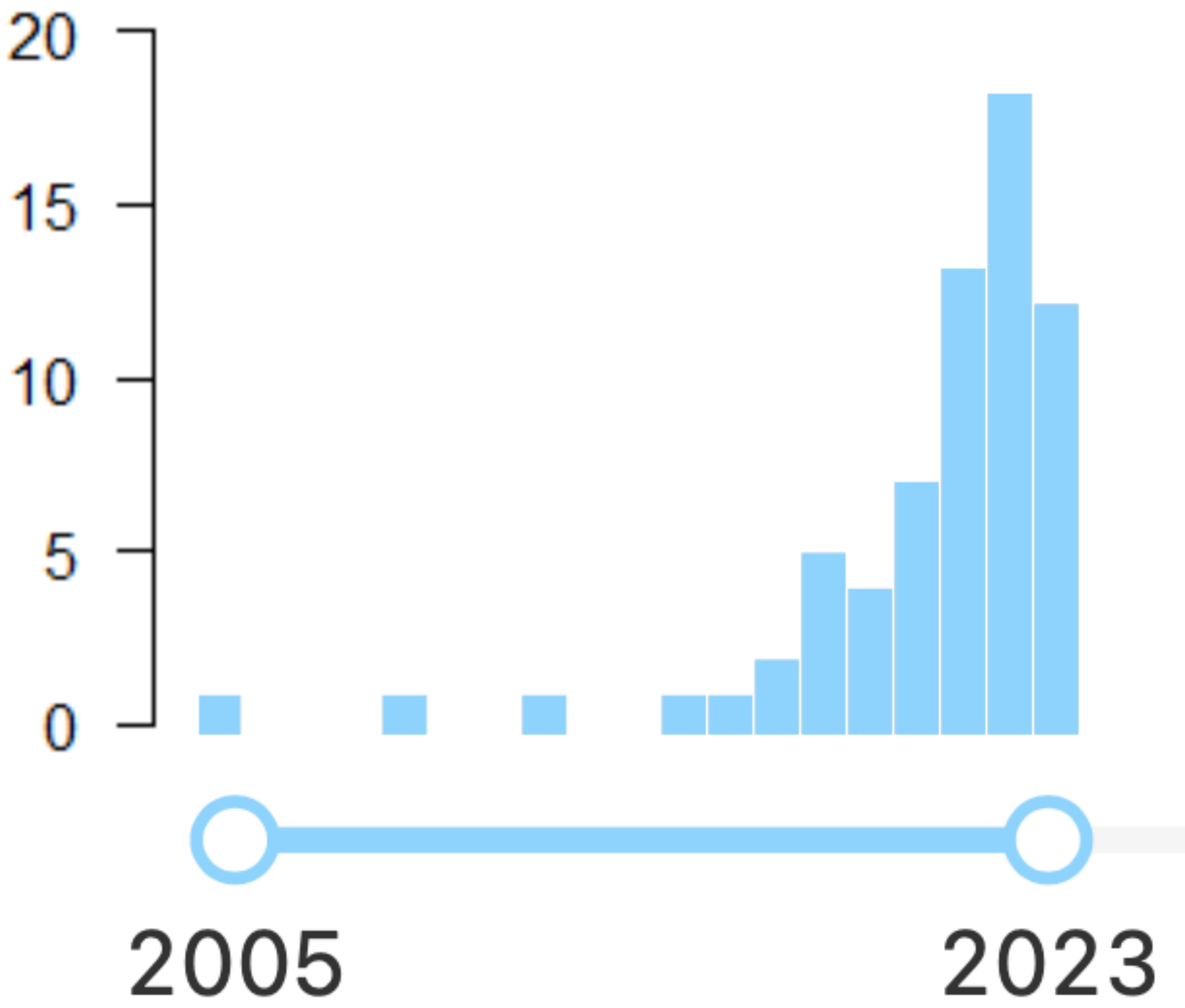
LISA data analysis using Markov chain Monte Carlo methods

Neil J. Cornish and Jeff Crowder
Phys. Rev. D **72**, 043005 – Published 22 August 2005

Article | References | Citing Articles (85) | PDF | HTML | Export Citation

ABSTRACT

The Laser Interferometer Space Antenna (LISA) is expected to simultaneously detect many thousands of low-frequency gravitational wave signals. This presents a data analysis challenge that is very different to the one encountered in ground based gravitational wave astronomy. LISA data analysis requires the identification of individual signals from a data stream containing an unknown number of overlapping signals. Because of the signal overlaps, a global fit to all the signals has to be performed in order to avoid biasing the solution. However, performing such a global fit requires the exploration of an enormous parameter space with a dimension upwards of 50000. Markov Chain Monte Carlo (MCMC) methods offer a very promising solution to the LISA data analysis problem. MCMC algorithms



The Global Solution

Likelihood function for Gaussian noise

$$p(\mathbf{d}|\vec{\lambda}) = \frac{1}{\sqrt{(2\pi)^M \det\mathbf{C}}}\, e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})\cdot\mathbf{C}^{-1}\cdot(\mathbf{d}-\mathbf{h})}$$

$$\mathbf{h} = \sum_{i=1}^{N} \mathbf{h}_i = \text{GW signal model}$$

N unknown, mix of signal types

$$\mathbf{C} = \text{noise correlation matrix}$$

Jointly inferred with signal model . Up to M^3 cost to invert

$$\vec{\lambda} = \text{model parameters}$$

Signal and noise $\mathcal{O}(10^6)$ parameters

The Global Solution

Likelihood function for Gaussian noise

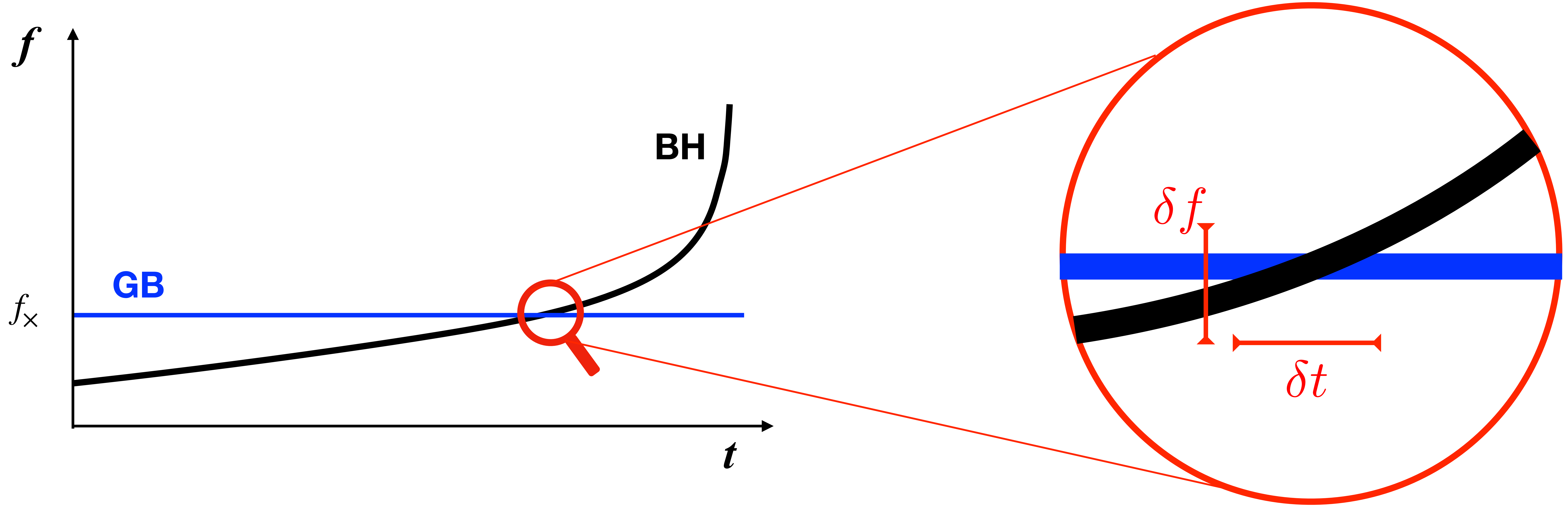
$$p(\mathbf{d}|\vec{\lambda}) = \frac{1}{\sqrt{(2\pi)^M \det \mathbf{C}}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})\cdot\mathbf{C}^{-1}\cdot(\mathbf{d}-\mathbf{h})}$$

$$\log L = (d\,|\,h) - \frac{1}{2}(h\,|\,h) = \sum_{i}^{N} \log L_i - \frac{1}{2}\sum_{i\neq j}(h_i\,|\,h_j)$$

Per-source likelihood

Signal overlaps - why we need a global solution

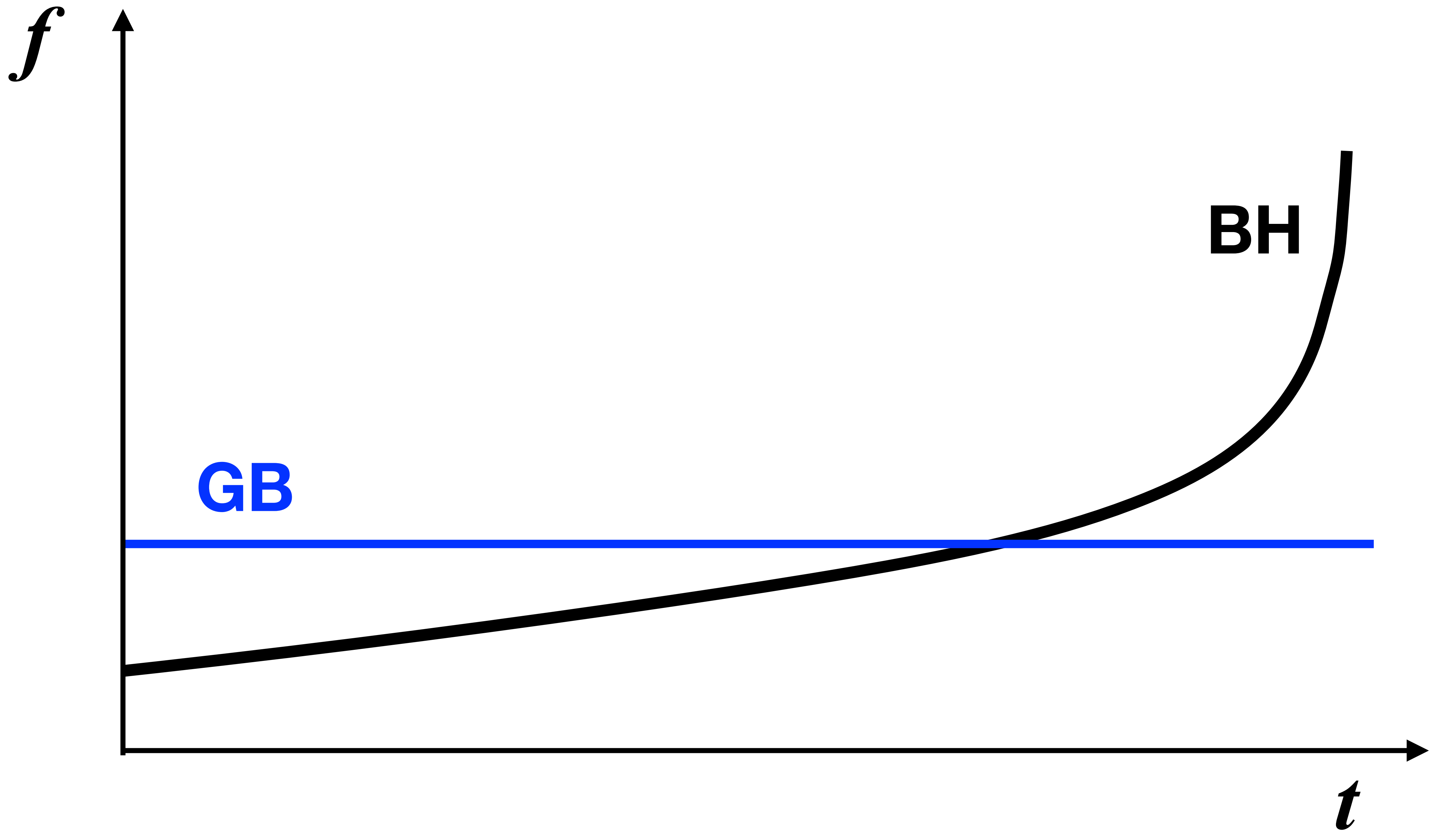
Example: SMBHB-GB Signal Overlaps



$$\delta f = 3 \times 10^{-7}\,\mathrm{mHz} \left(\frac{\mathcal{M}_{\rm GB}}{0.25 M_\odot}\right)^{5/6} \left(\frac{f_\times}{1\,\mathrm{mHz}}\right)^{11/6}$$

$$\delta t = 1.7 \times 10^{3}\,\mathrm{s} \left(\frac{10^6 M_\odot}{\mathcal{M}_{\rm BH}}\right)^{5/6} \left(\frac{1\,\mathrm{mHz}}{f_\times}\right)^{11/6}$$

Example: SMBHB-GB Signal Overlaps



Instantaneous BH SNR

Random phase mismatch

$$(\mathbf{h}_{\rm BH}|\mathbf{h}_{\rm GB}) = \left(\frac{S_{\rm BH}(f_\times)}{S_{\rm n}(f_\times)}\right)^{1/2} \rho_{\rm GB}\, \cos\delta$$

$$\approx 10^{-3}\, \rho_{\rm BH}\, \rho_{\rm GB}\, \cos\delta$$

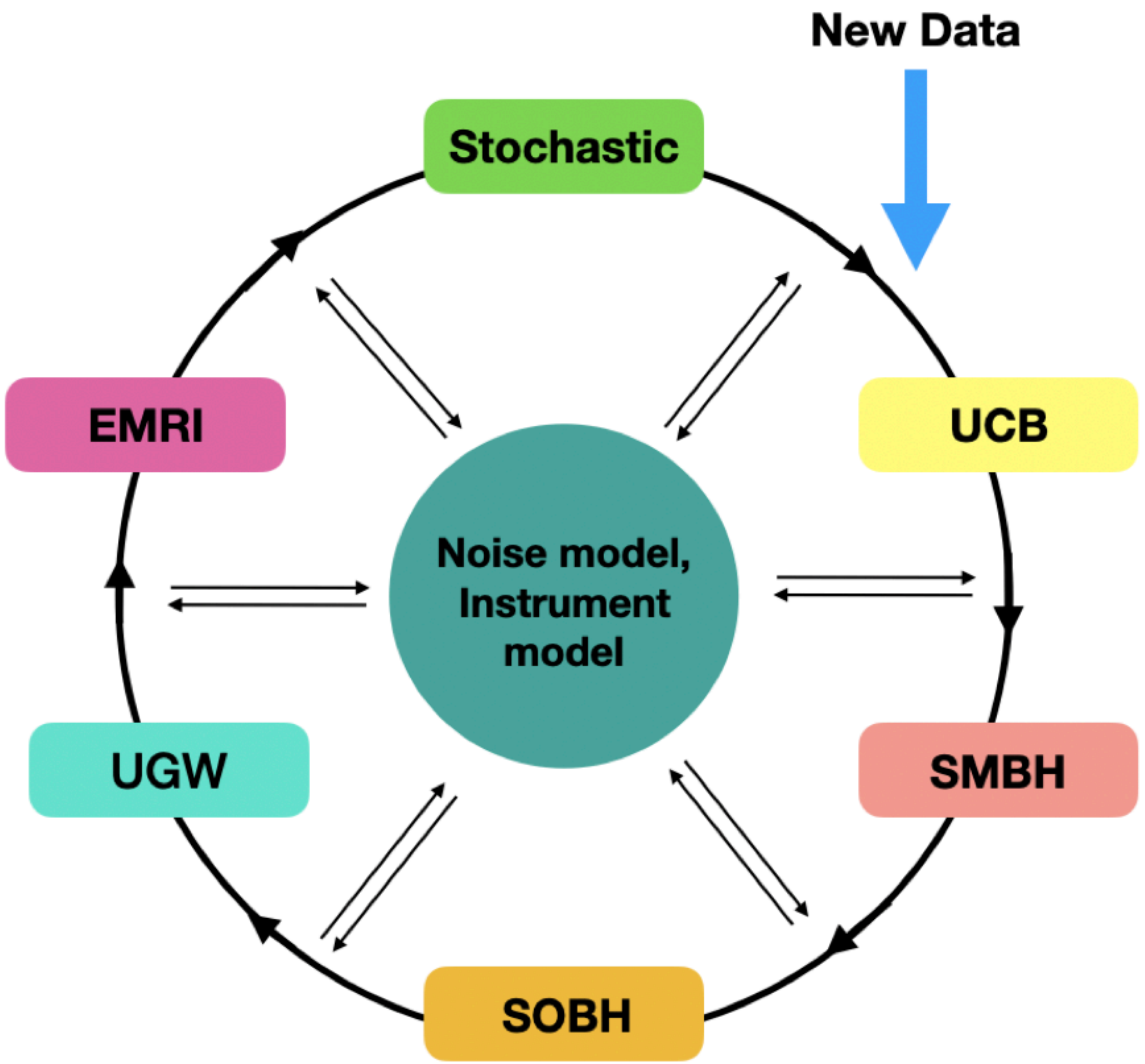
Individual overlaps are small, but there will be tens of thousands of them

$$\sum_{\rm GB} (\mathbf{h}_{\rm BH}|\mathbf{h}_{\rm GB}) \approx \sqrt{N}_{\rm GB}\, 10^{-3}\, \rho_{\rm BH}\, \rho_{\rm GB} \approx 0.2\, \rho_{\rm BH}\, \rho_{\rm GB}$$

Significant bias if not solved for simultaneously

[see Cutler & Harms (2008), Robson & Cornish (2018)]

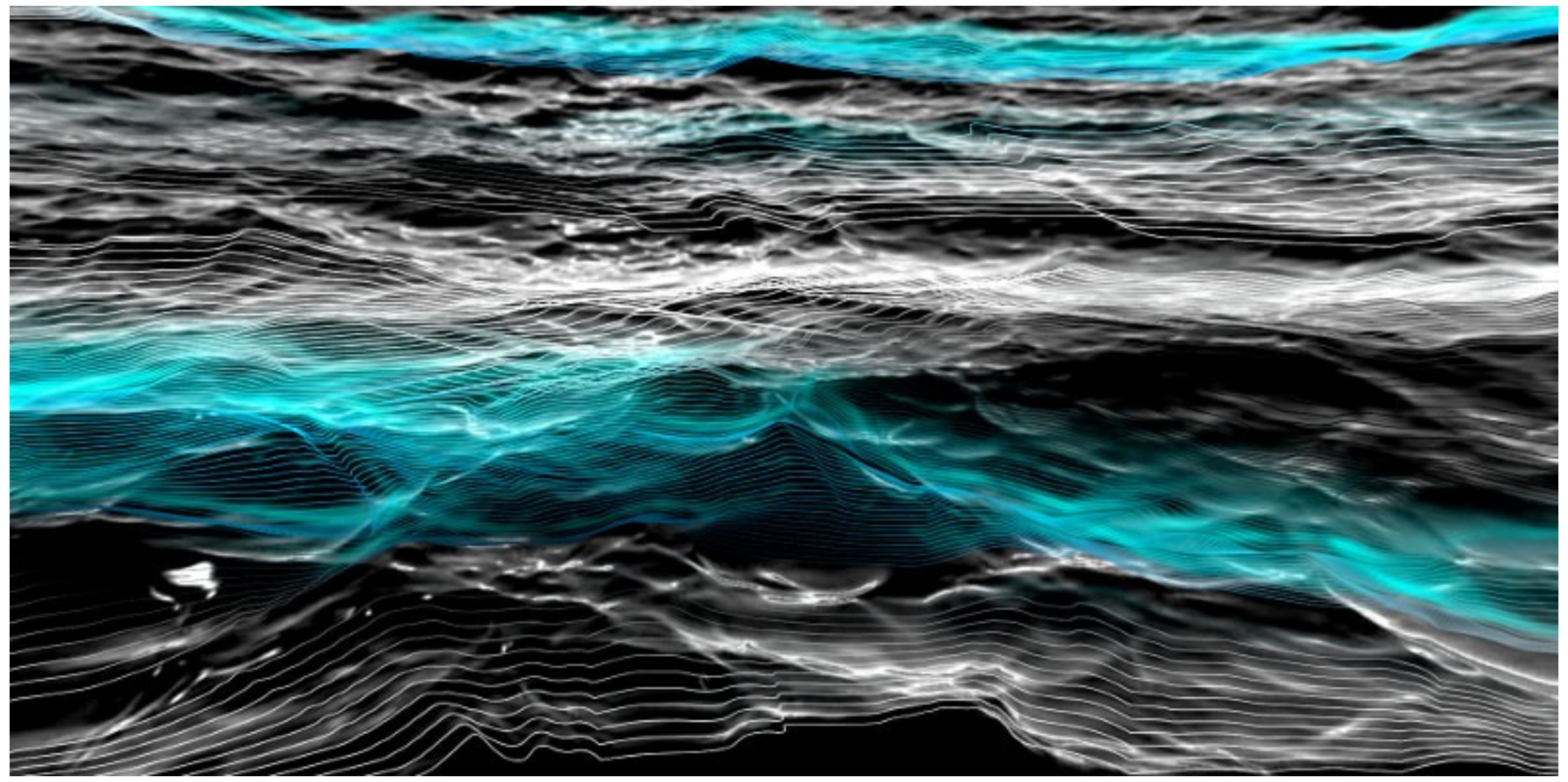
Global Fit via Blocked Sampling



- Transdimensional Markov Chain Monte Carlo (RJMCMC)
- Blocked Metropolis Hastings—update each component of the signal/noise model in circular sweeps
- Only pass residuals - decouples the analysis types
- Update the fit every ~week as new data arrives

How does the stochastic background fit into all this?

$$p(\mathbf{d}|\vec{\lambda}) = \frac{1}{\sqrt{(2\pi)^M \det \mathbf{C}}}\, e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})\cdot\mathbf{C}^{-1}\cdot(\mathbf{d}-\mathbf{h})}$$



$$\mathbf{h} = \sum_{i=1}^{N} \mathbf{h}_i = \text{GW signal model}$$

$$\mathbf{h}_s = \sum_f \tilde{a}_f\, e^{2\pi i f t}$$

Just another template to be included in the signal model

[Cornish & Romano, PRD 87, 122003 (2013)]

[Lentati et al, PRD 87, 104021 (2013)]

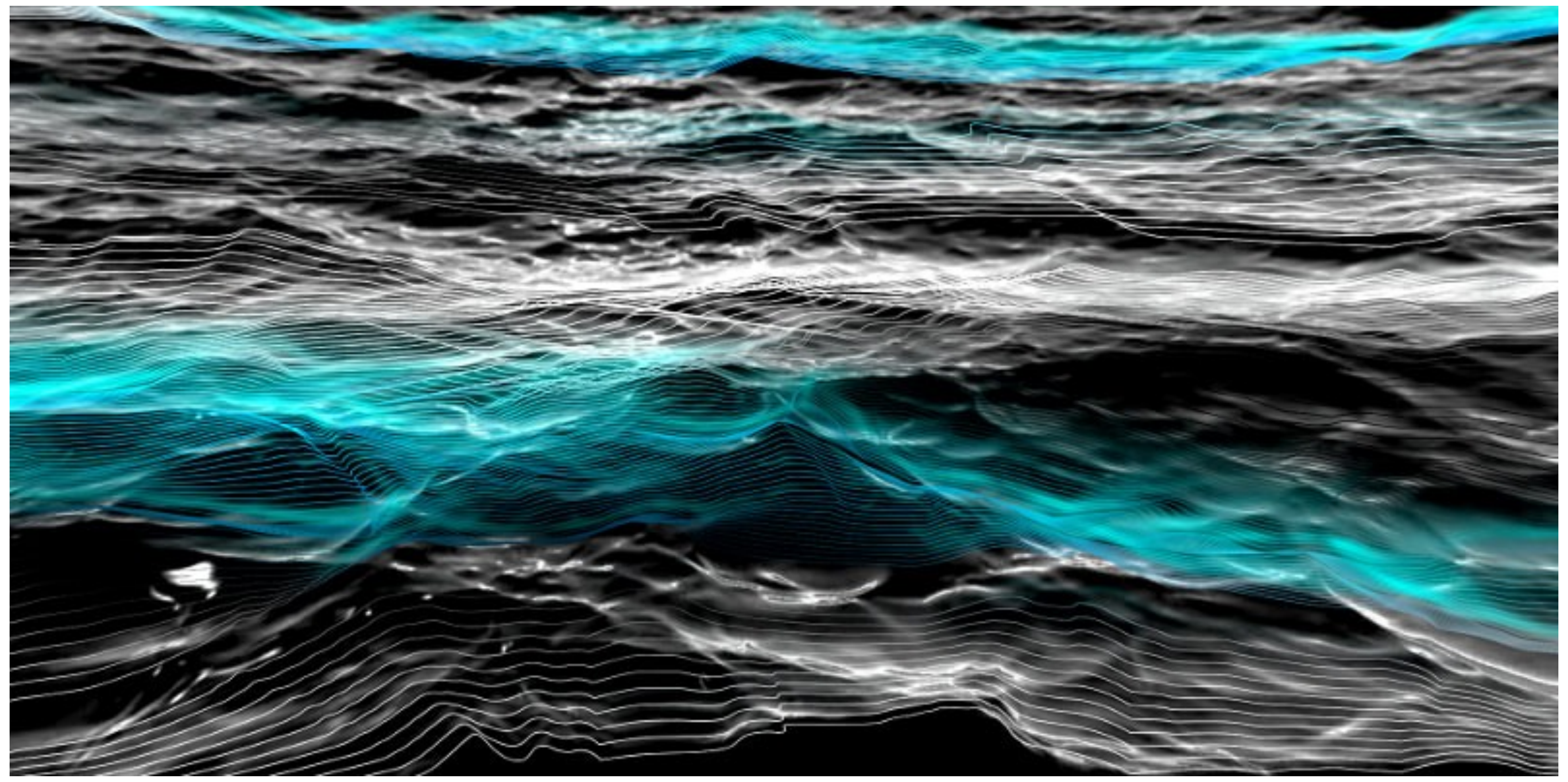
[Romano & Cornish, LRR 20 (2017)]

How does the stochastic background fit into all this?

$$\mathbf{h}_s = \sum_f \tilde{a}_f \, e^{2\pi i f t}$$

Just another template to be included in the signal model

“Stochastic template”. Gaussian prior with spectrum $S_h(f)$



$$p(\tilde{a}_f) = \prod_f \frac{1}{2\pi S_h(f)} \exp\left(\frac{-\tilde{a}_f \, \tilde{a}_f^*}{S_h(f)} \right)$$

[Cornish & Romano, PRD 87, 122003 (2013)] [Lentati et al, PRD 87, 104021 (2013)] [Romano & Cornish, LRR 20 (2017)]

How does the stochastic background fit into all this?

$$\mathbf{h}_s = \tilde{a}_f \, e^{2\pi i f t}$$

$$p(\tilde{a}_f) = \prod_f \frac{1}{2\pi \, S_h(f)} \exp\left(\frac{-\tilde{a}_f \, \tilde{a}_f^*}{S_h(f)} \right)$$

Typically we are not interested in the particular realization of the random amplitudes so we can integrate them out. e.g. for stationary data in the frequency domain we get

$$p(\mathbf{d}) = \frac{1}{\det(2\pi\mathbf{G})} \exp\left[-(\tilde{d}(f) - \tilde{h}(f))_i^* G_{ij}(f)^{-1} (\tilde{d}(f) - \tilde{h}(f))_j \right]$$

Overlap reduction function (LIGO), Hellings-Downs Curve (PTAs)

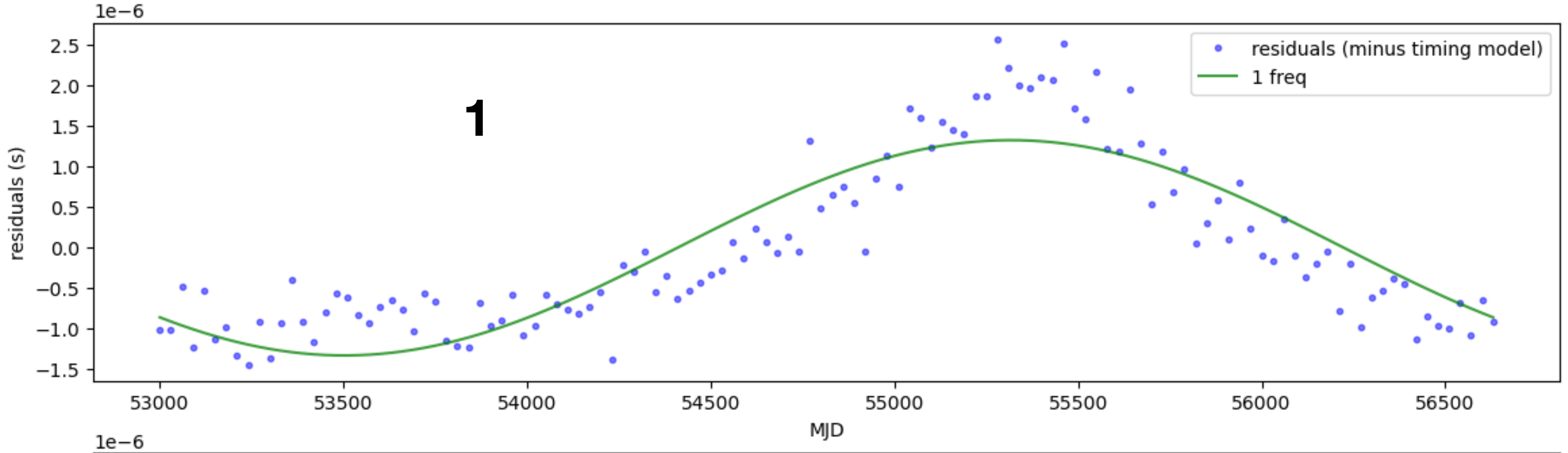
$$G_{ij}(f) = \delta_{ij} \, S_{n,i}(f) + \gamma_{ij}(f) \, S_h(f)$$

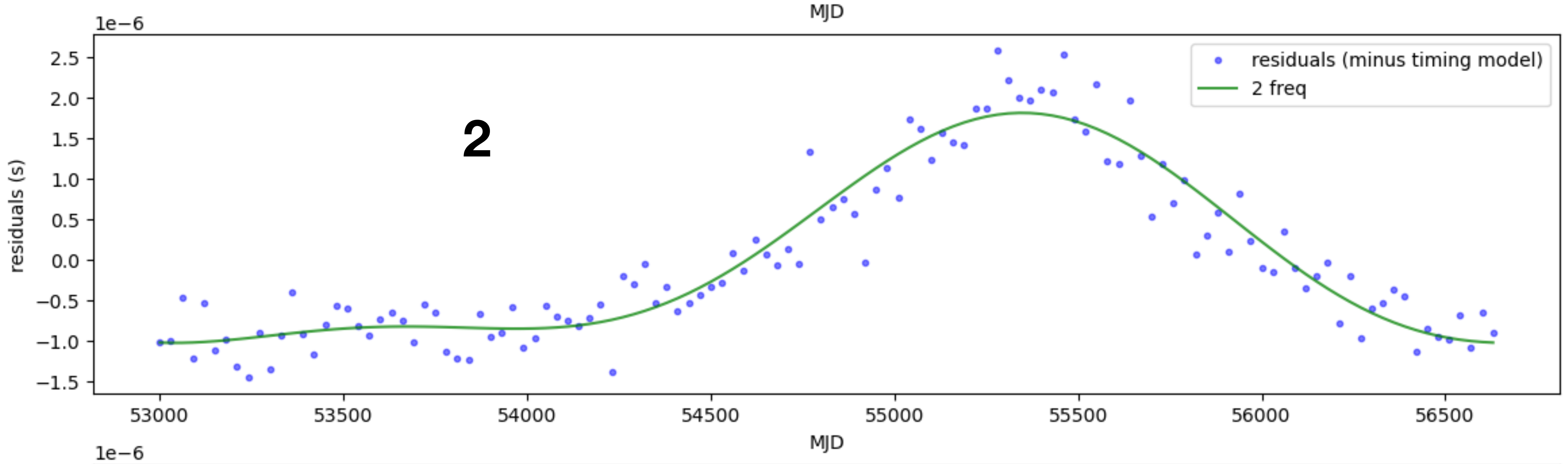
$$\gamma_{ij}(f) = \frac{1}{4\pi} \int (F_i^+(\hat{n}) F_j^+(\hat{n}) + F_i^\times(\hat{n}) F_j^\times(\hat{n})) \, e^{2\pi i f (\vec{x}_i - \vec{x}_j)\cdot\hat{n}} \, d\Omega_{\hat{n}}$$

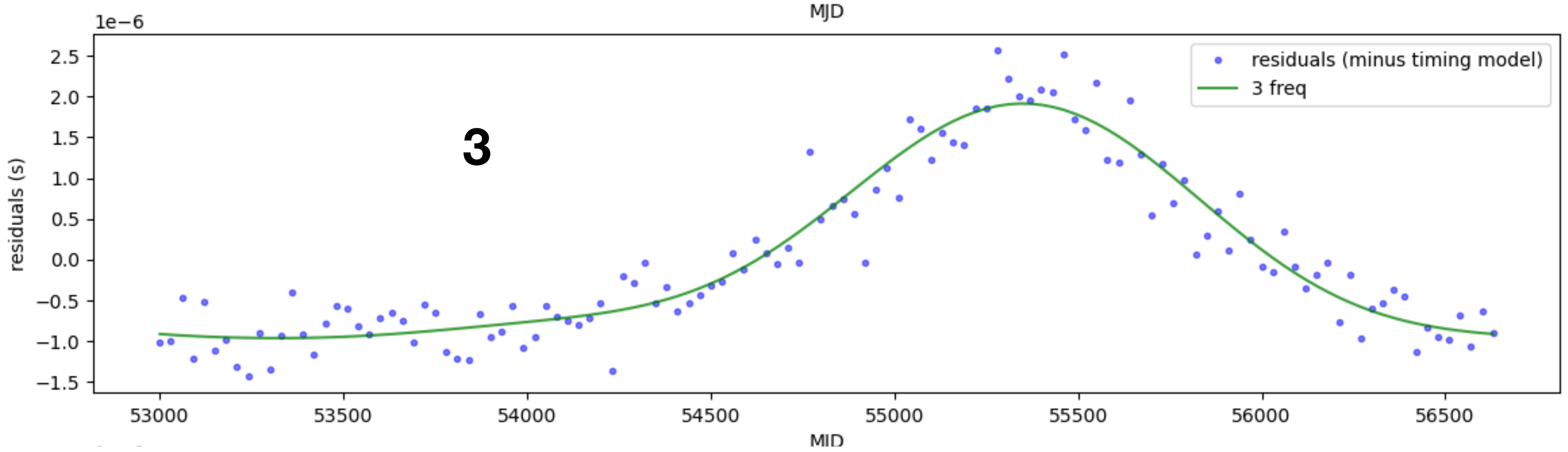
[Cornish & Romano, PRD 87, 122003 (2013)] [Lentati et al, PRD 87, 104021 (2013)] [Romano & Cornish, LRR 20 (2017)]

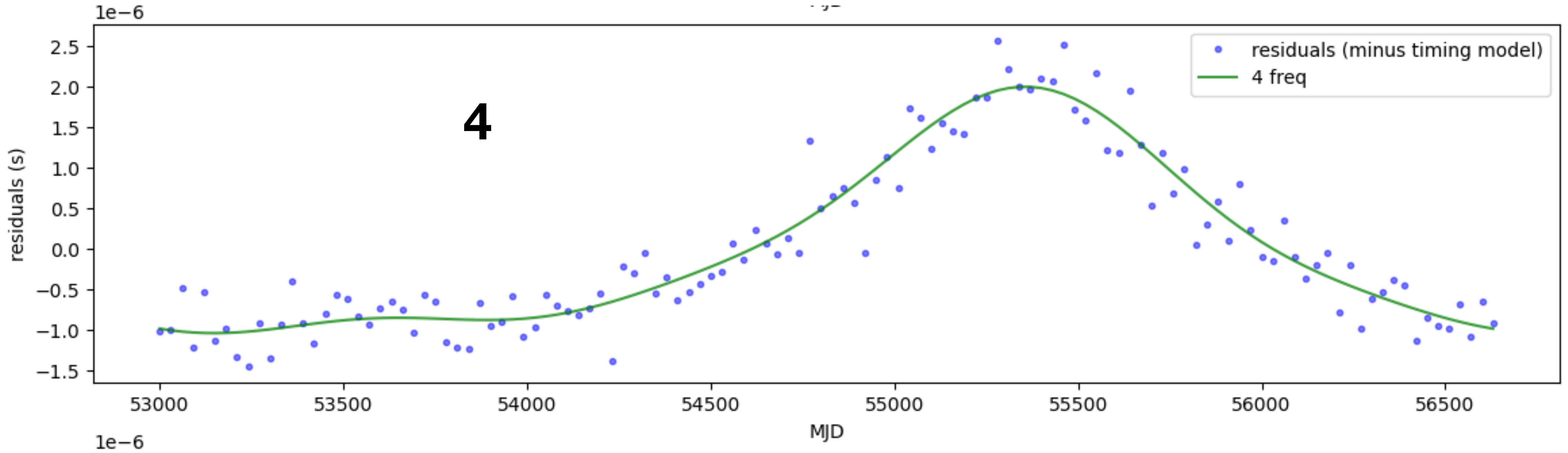
If preferred, can instead subtract the stochastic background signal

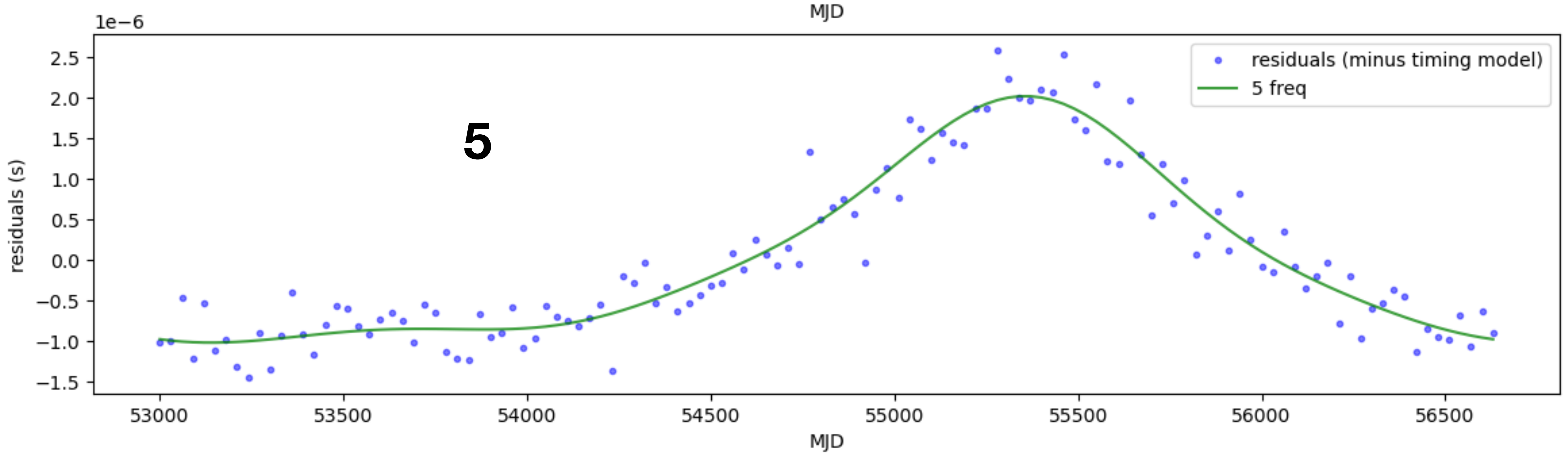
(This is for Germano and Chiara, who at the cosmology working group meeting in Hamburg, October 2016 kept on talking about subtracting the stochastic background)





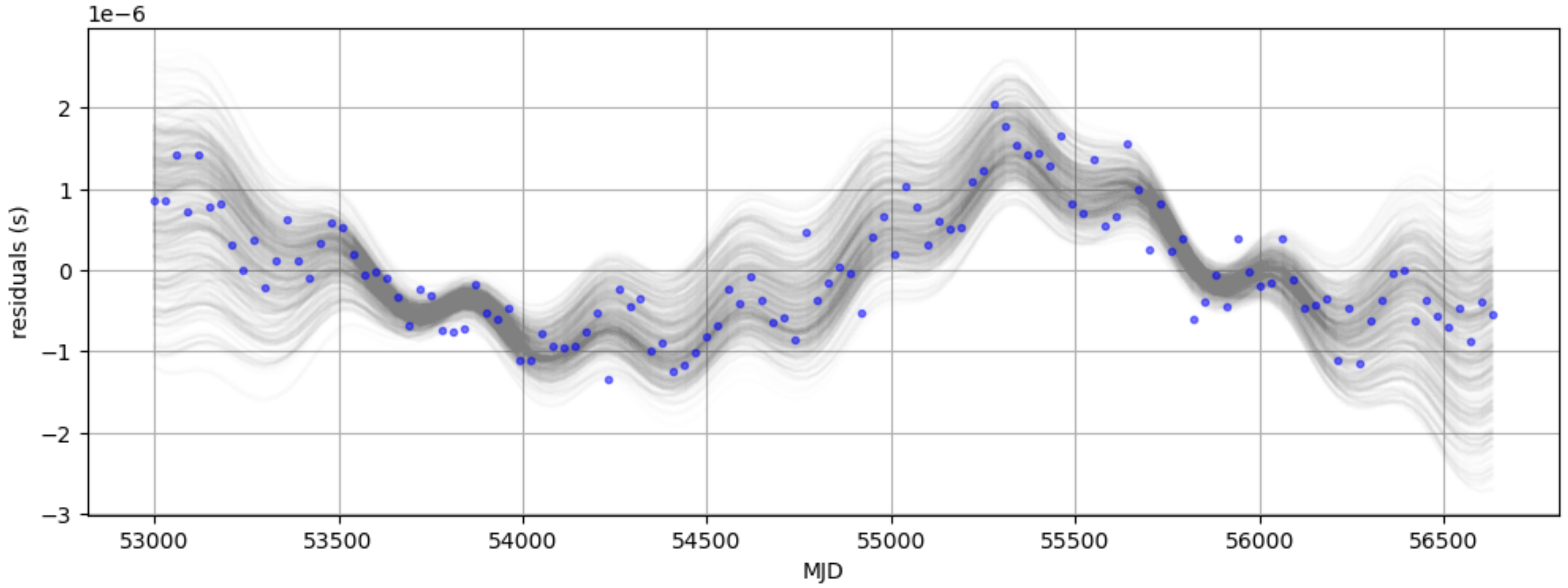


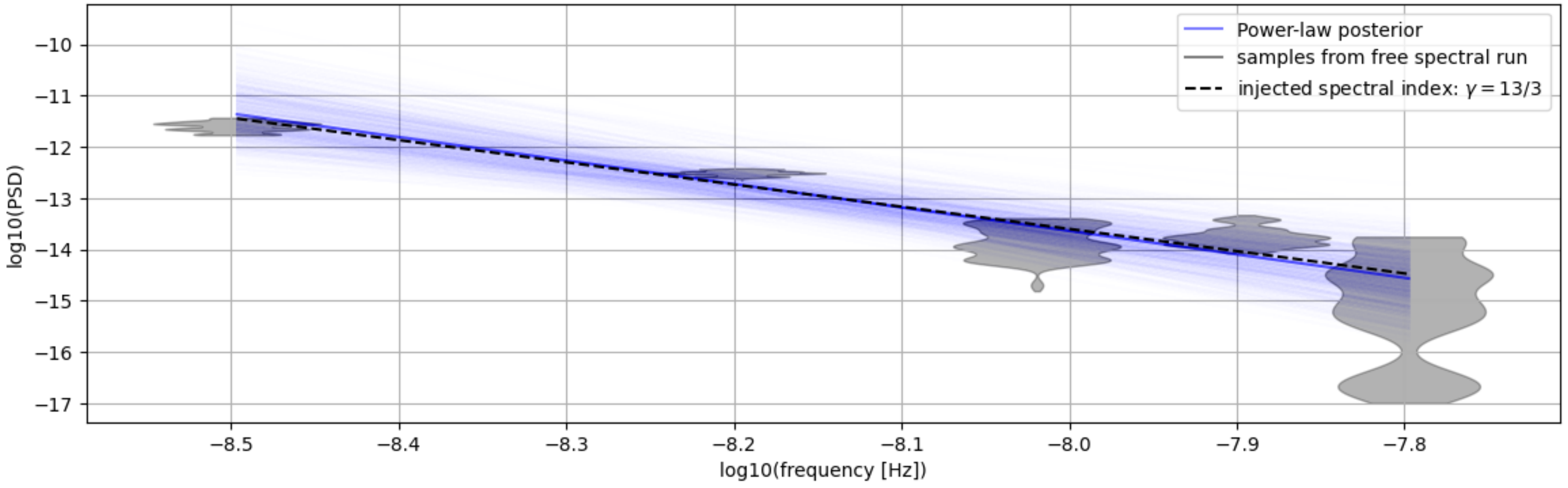


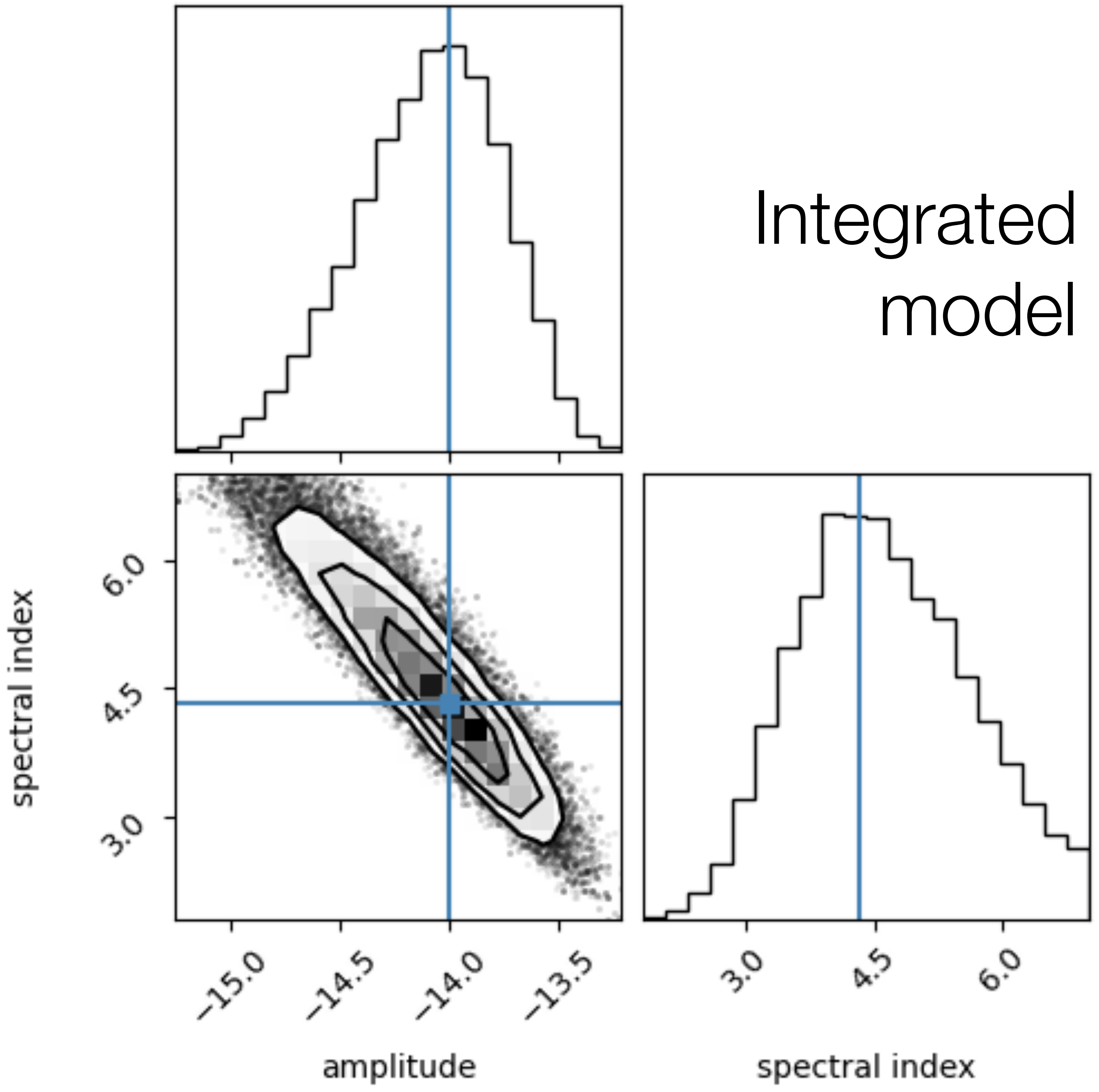


NANOGrav 15 yr Pulsar Timing example.
Just 5 frequency components needed to subtract stochastic signal

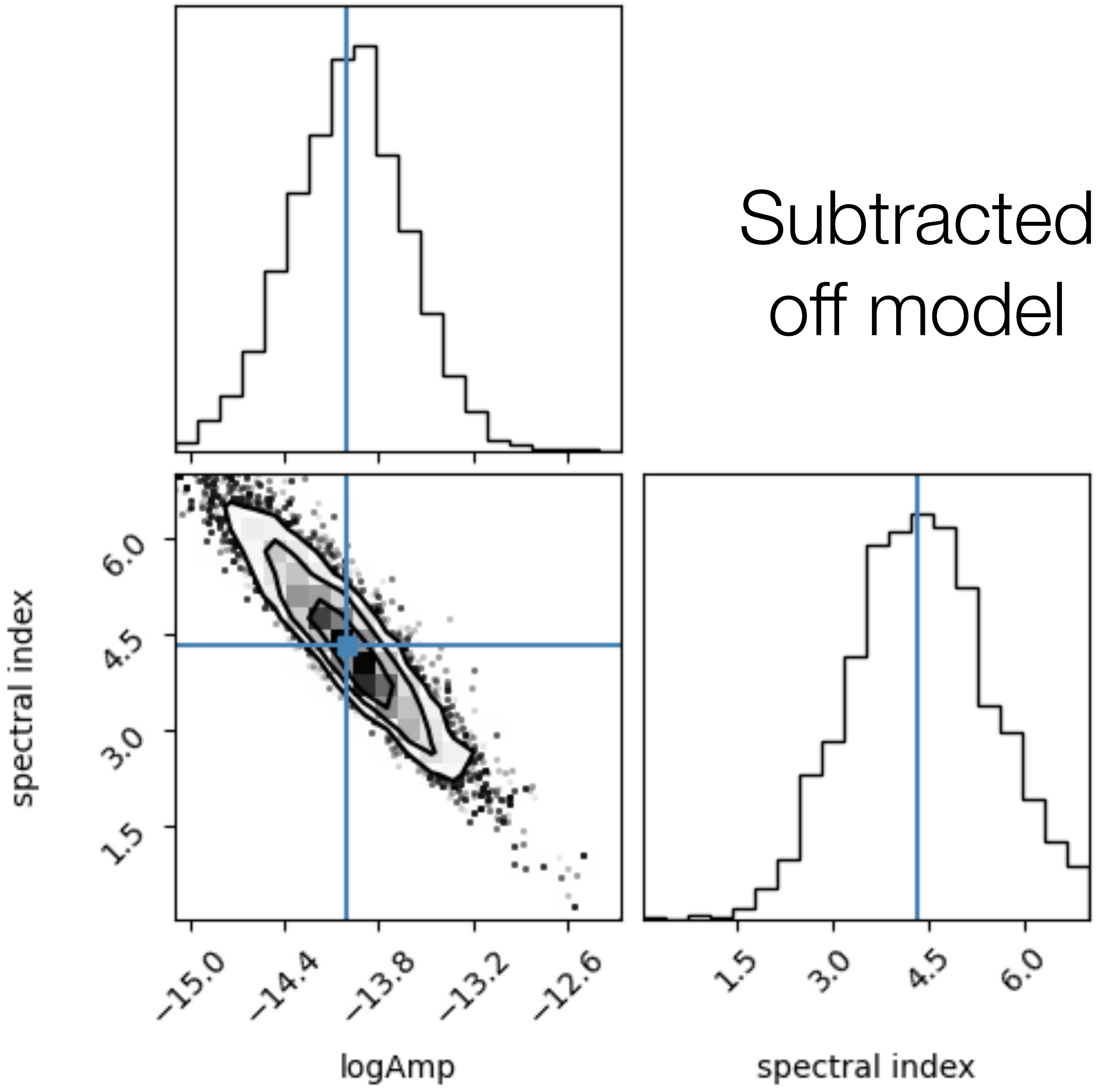
Draws from the MCMC







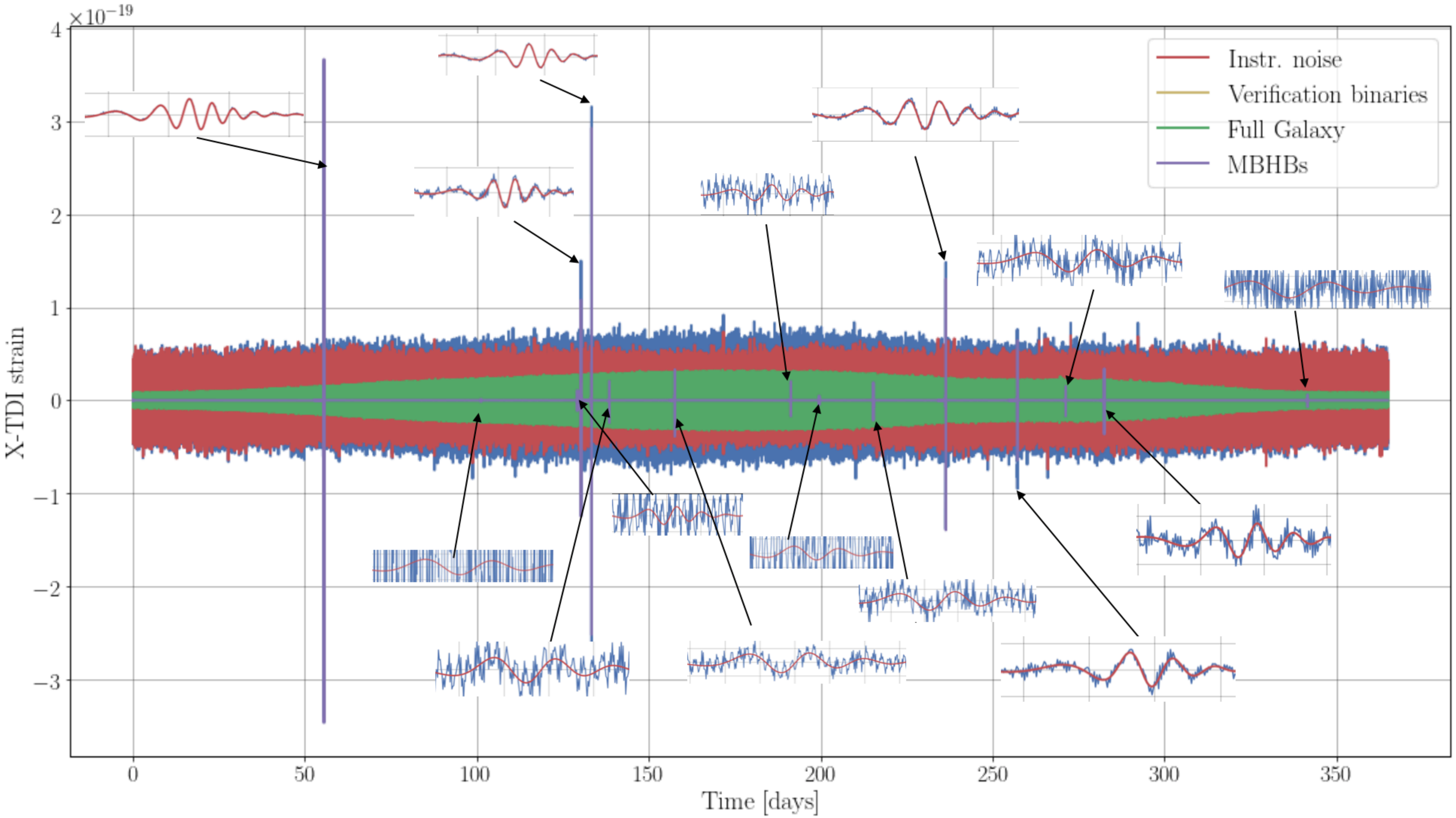
Integrated out model



Subtracted off model

Analysis by Aiden Gundersen

LISA Data Challenge: Sangria Edition

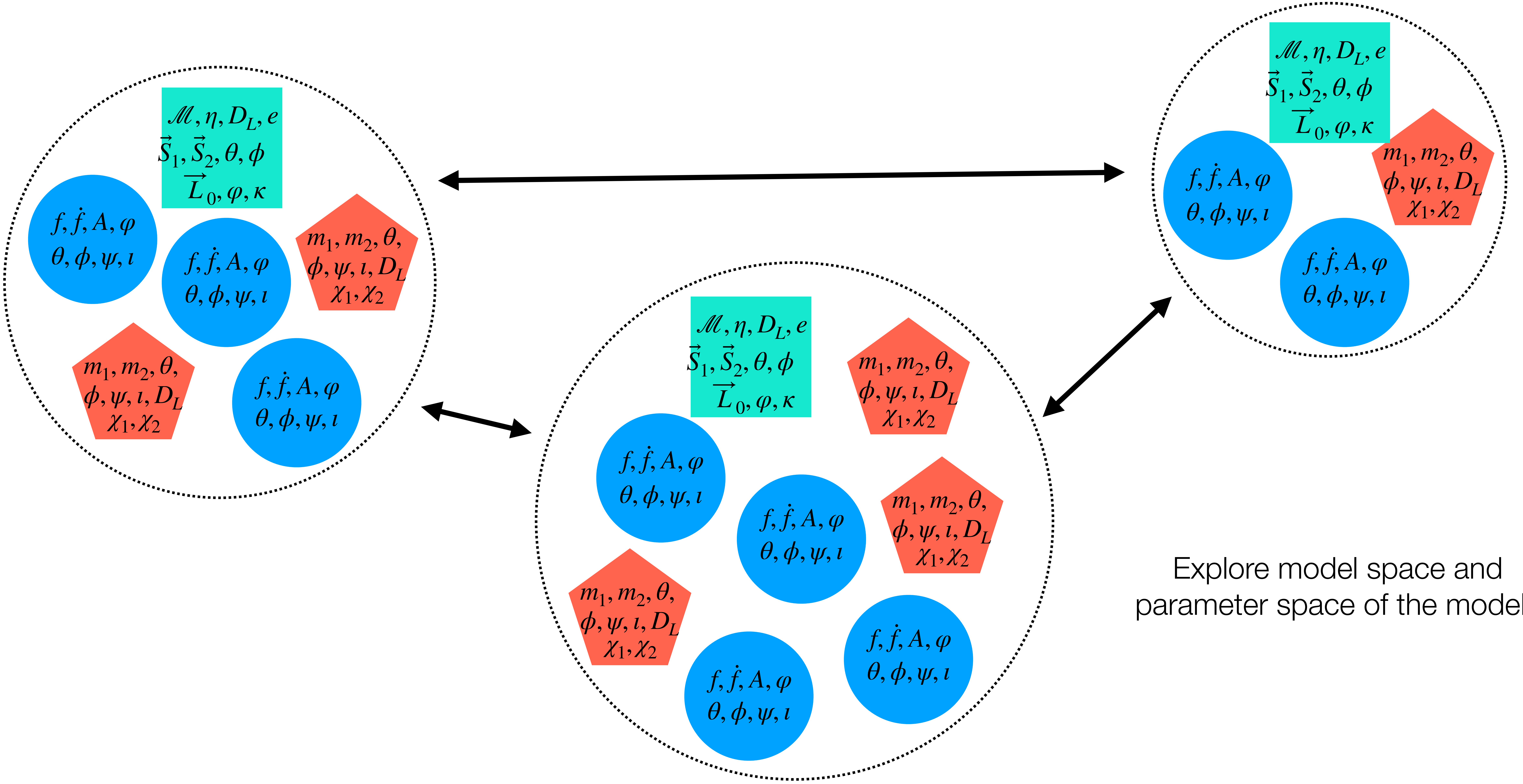


LISA Global Fit (One approach)

- Trans-dimensional modeling using Reversible Jump MCMC
- Build highly informative jump proposal from initial search stage
- Searches typically use stochastic hill climbing, approximate likelihoods, phenomenological models, likelihood maximization etc (anything goes)
- Time annealing - build up solution as new data arrives. Posteriors become proposals for next stage

[Litttenberg & Cornish, arXiv: 2301.03673]

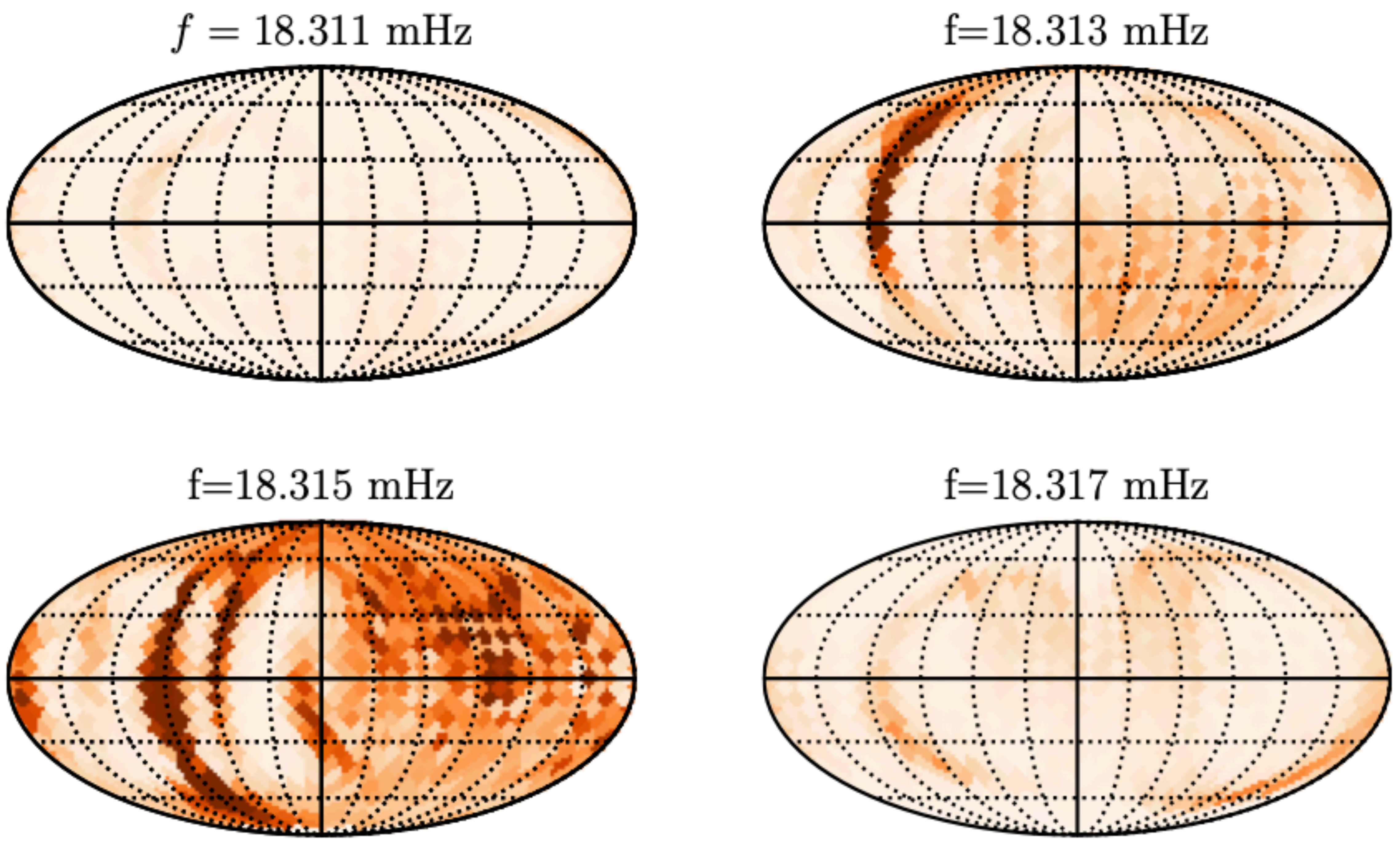
Trans-dimensional Inference



Explore model space and
parameter space of the models

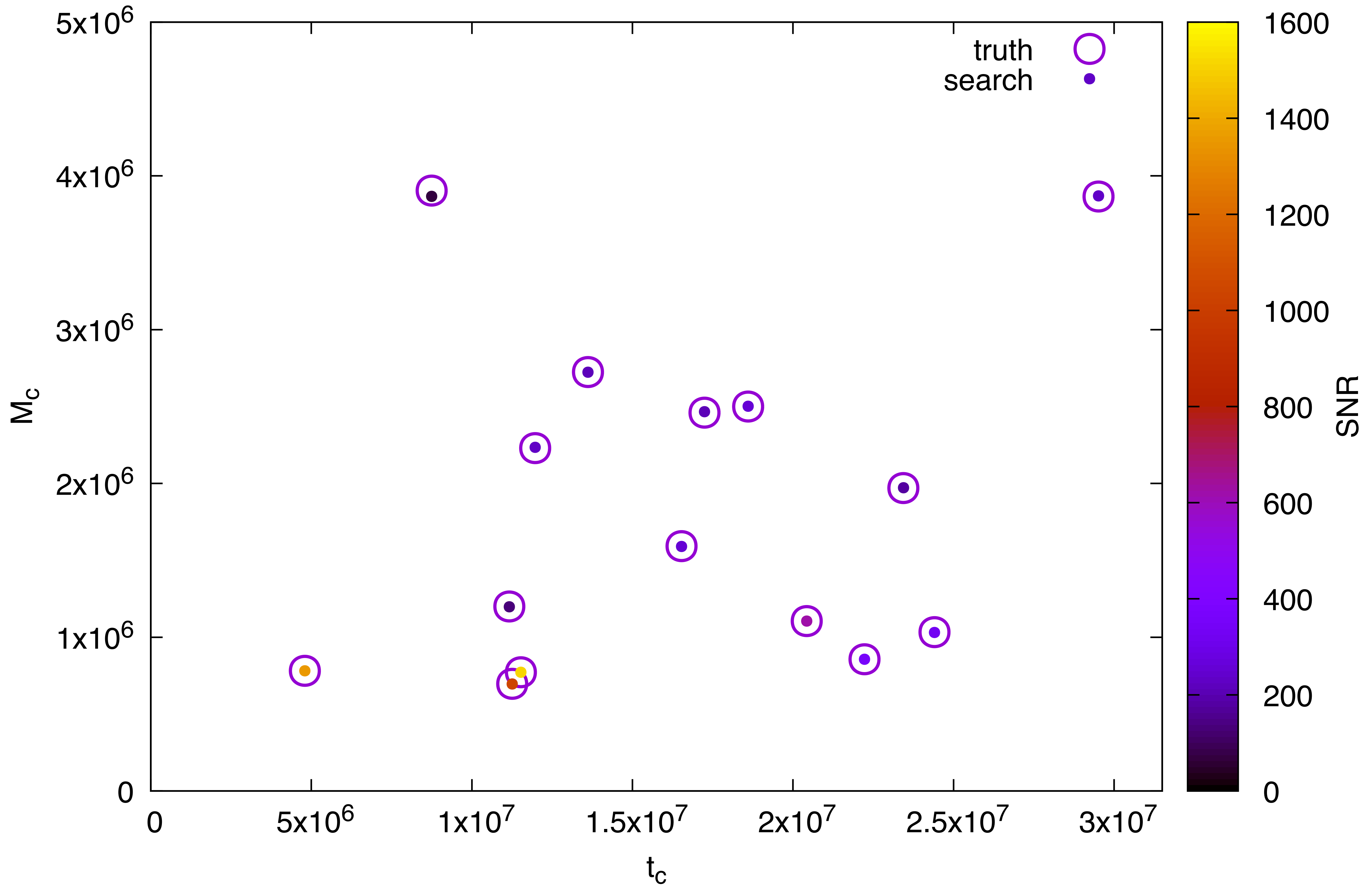
Low latency single-source search results used as proposals in global fit

F-statistic maps for GBs



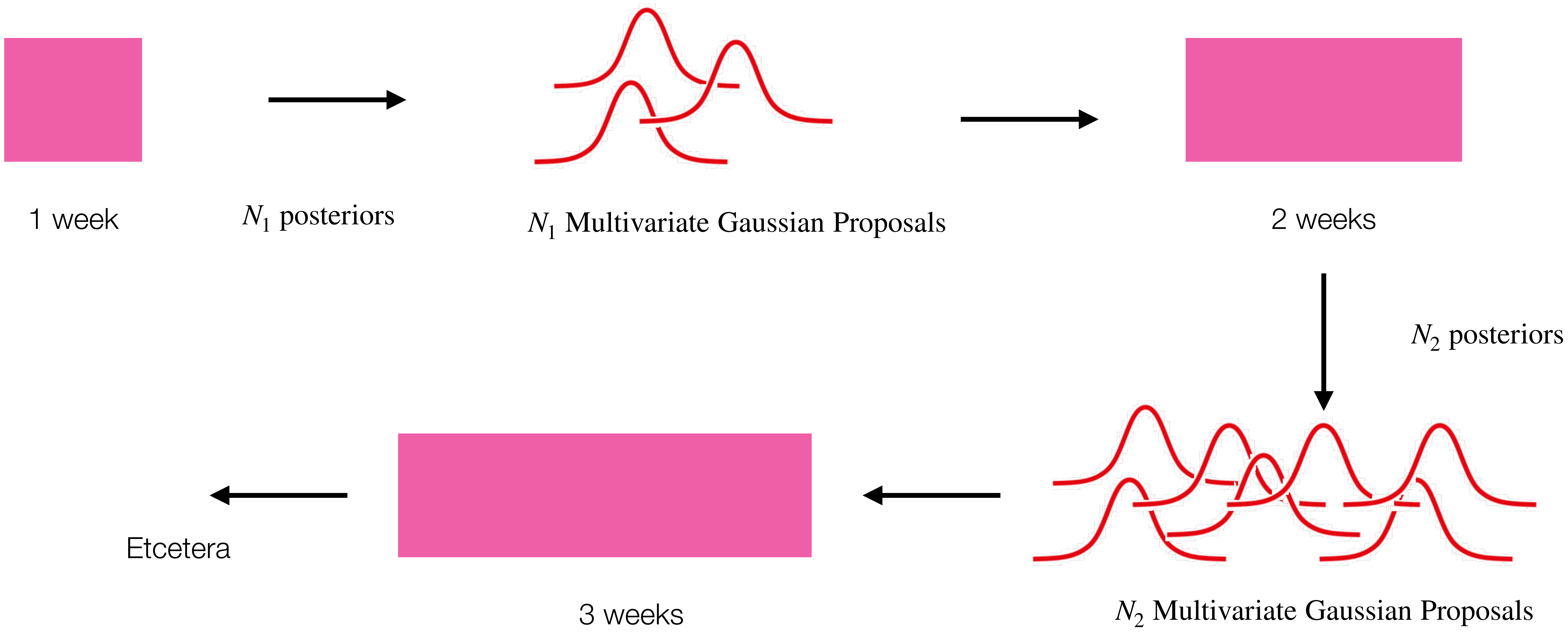
[Littenber, Cornish, Lackeos & Robson, arXiv:2004.08464]

Low latency BH search



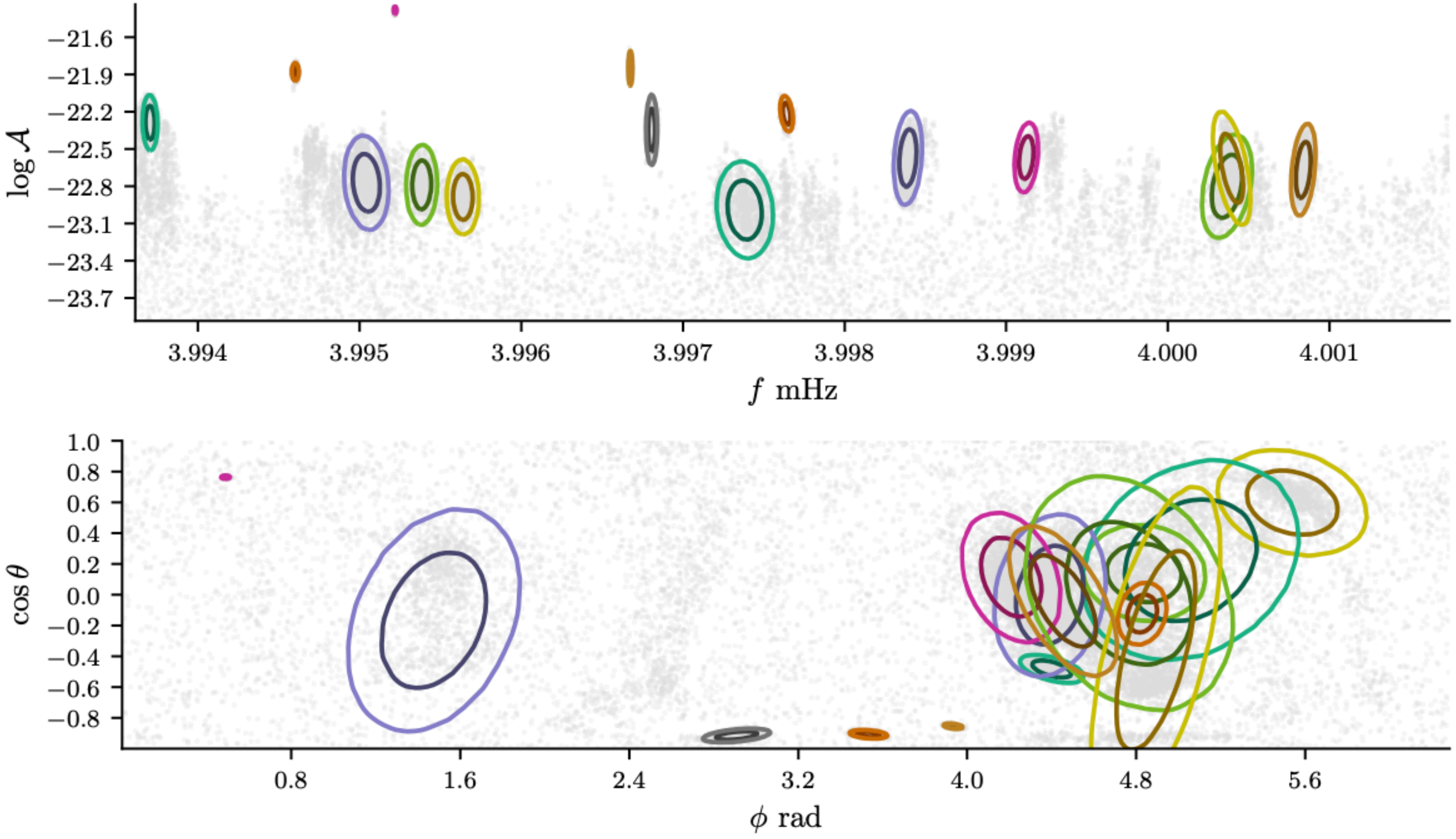
[Cornish, arXiv: 2110.06238]

Building up the solution - "time annealing"



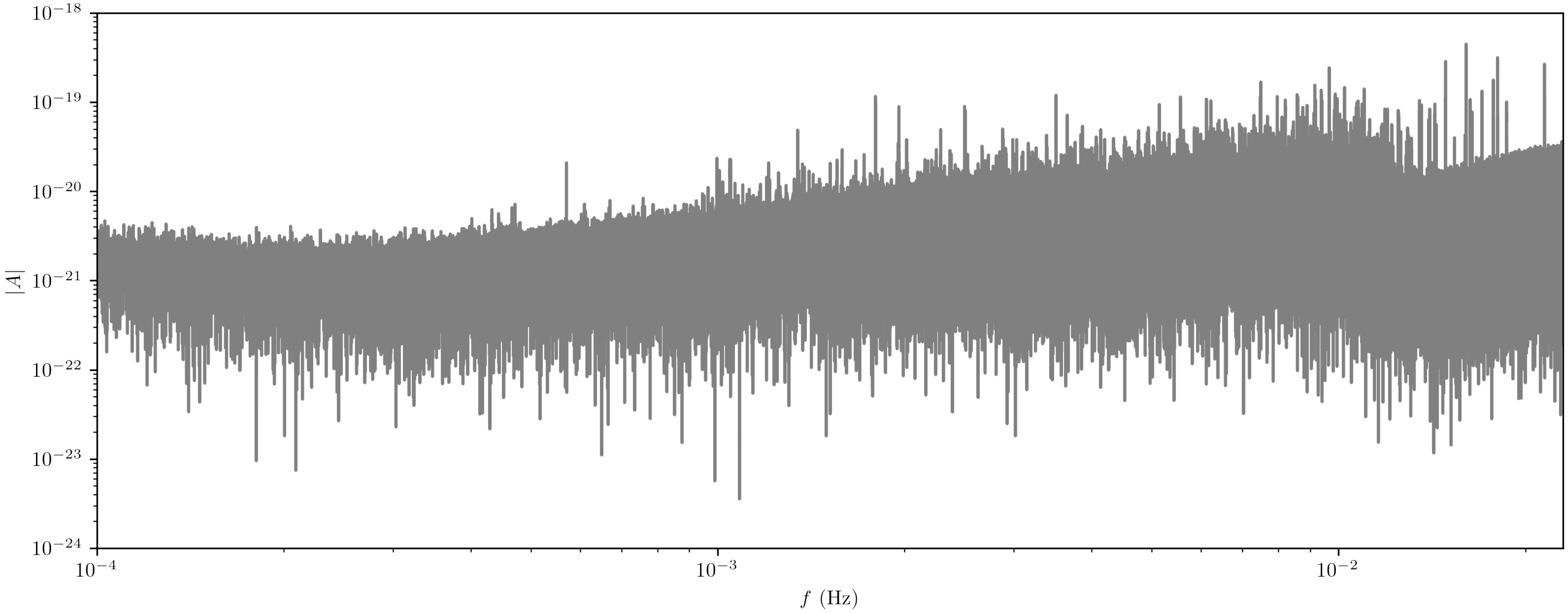
We used 1 month -> 3 months -> 6 months -> 12 months

Building up the solution - “time annealing”



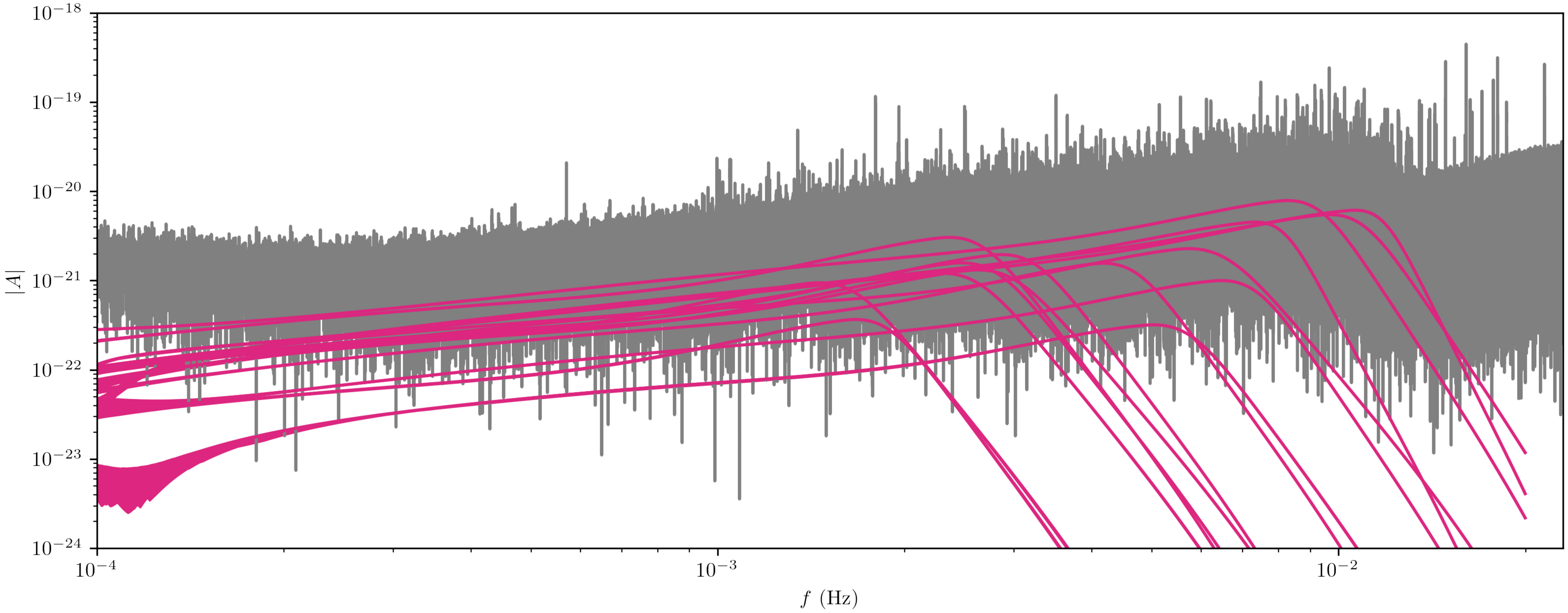
We used 1 month -> 3 months -> 6 months -> 12 months

12 months of Sangria data - A TDI channel

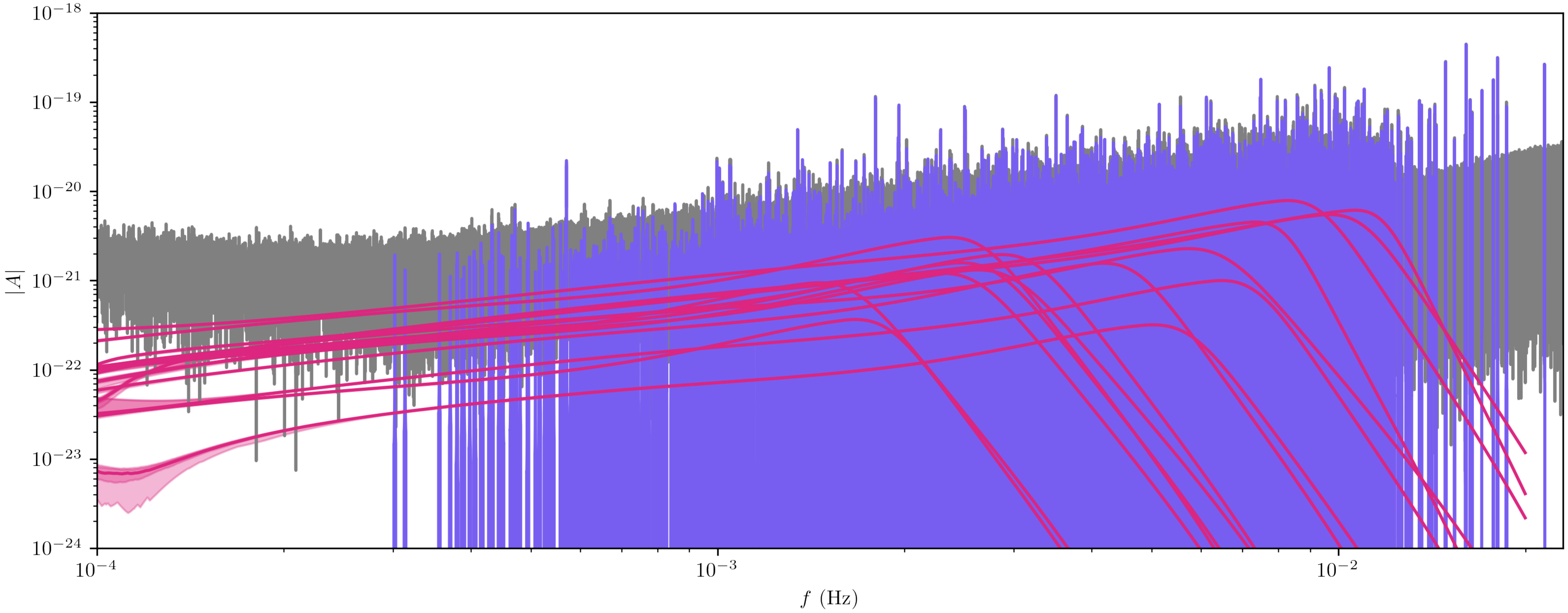


[Litttenberg & Cornish, arXiv: 2301.03673]

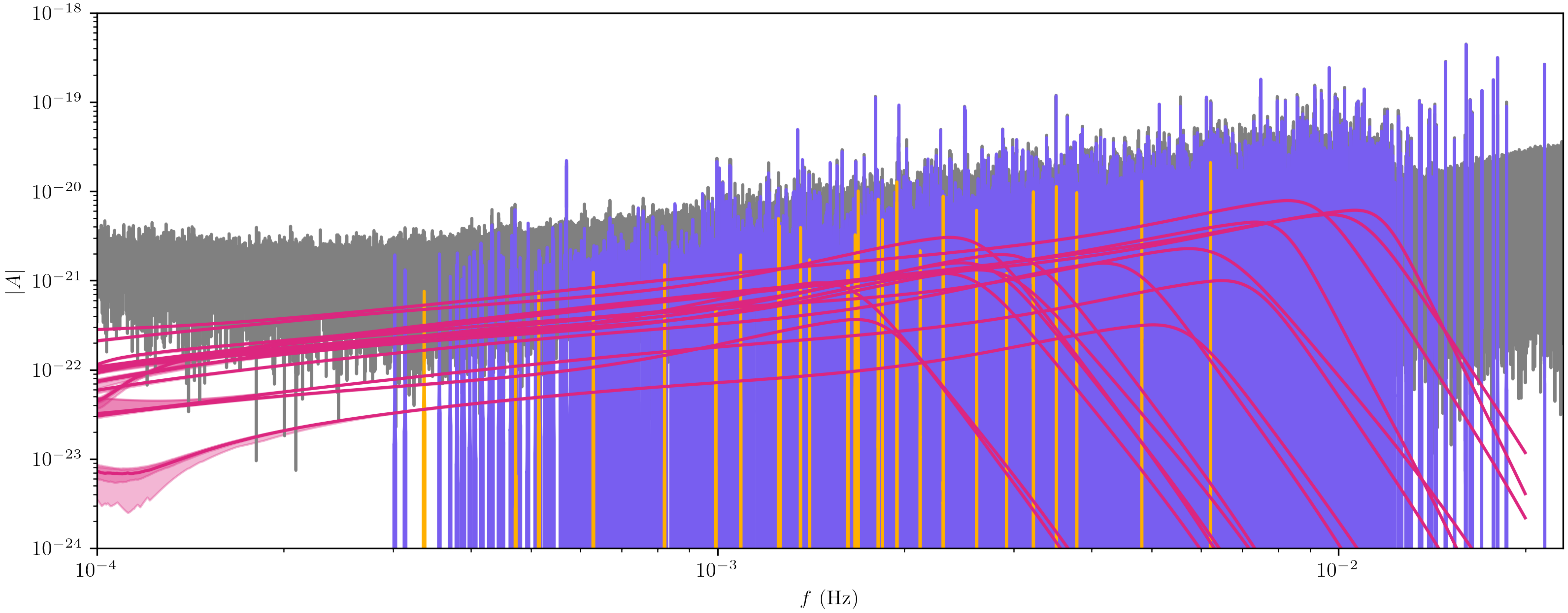
12 months of Sangria data - A TDI channel



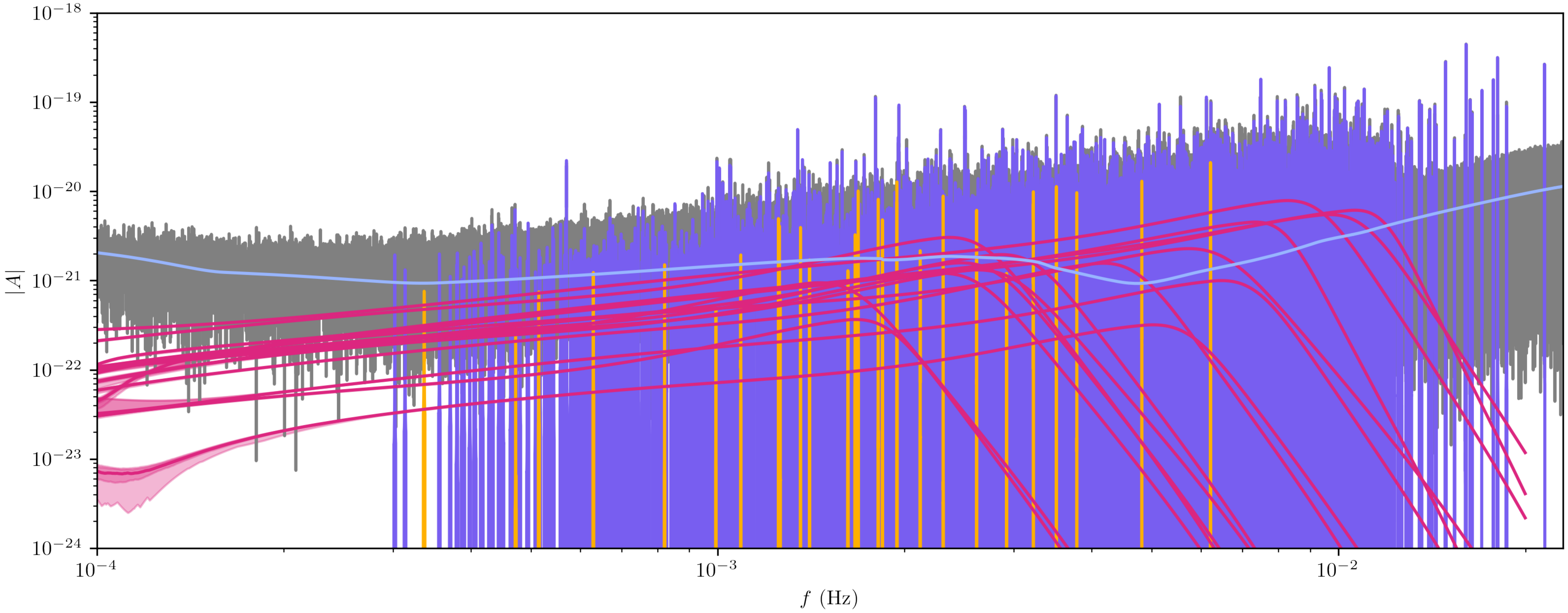
12 months of Sangria data - A TDI channel



12 months of Sangria data - A TDI channel

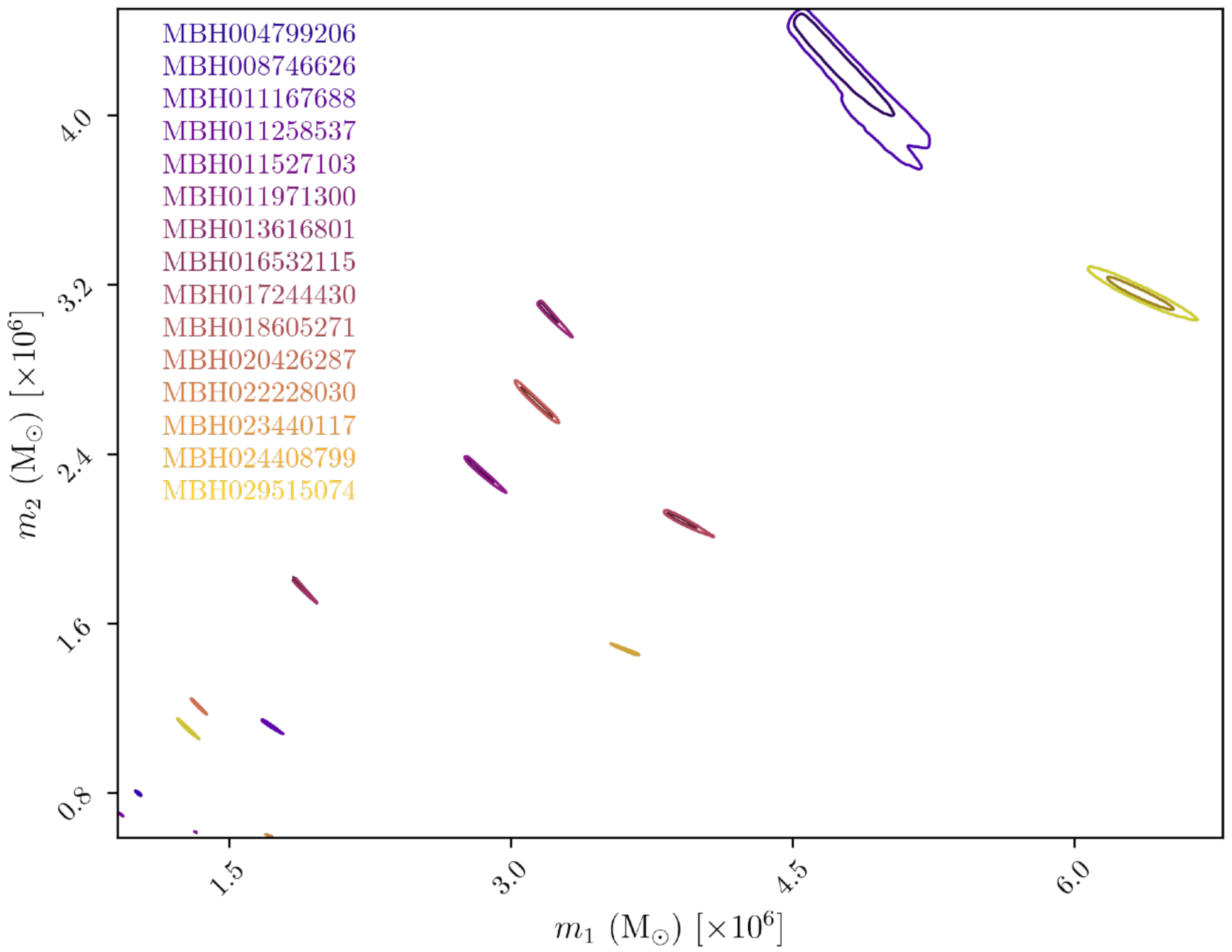


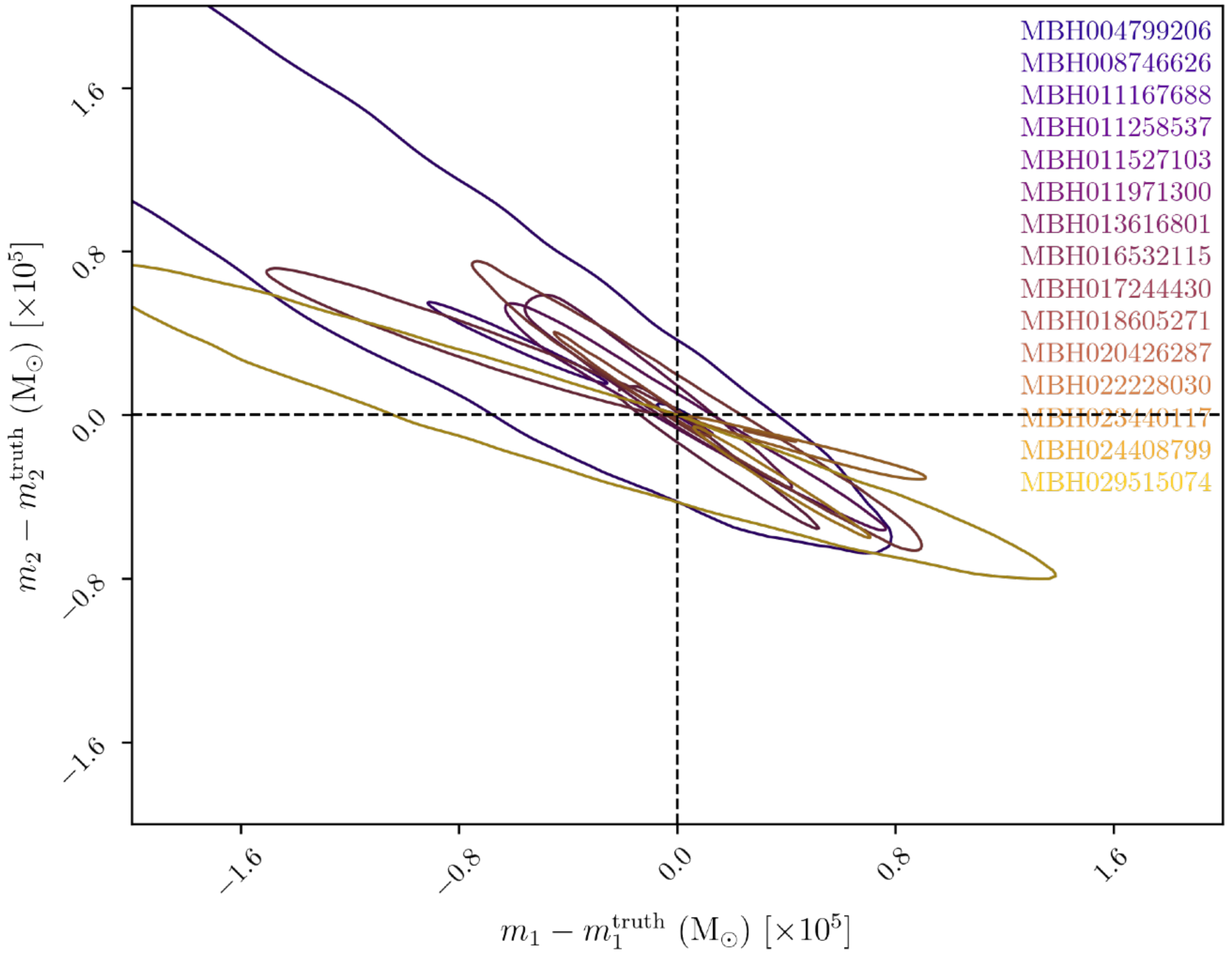
12 months of Sangria data - A TDI channel

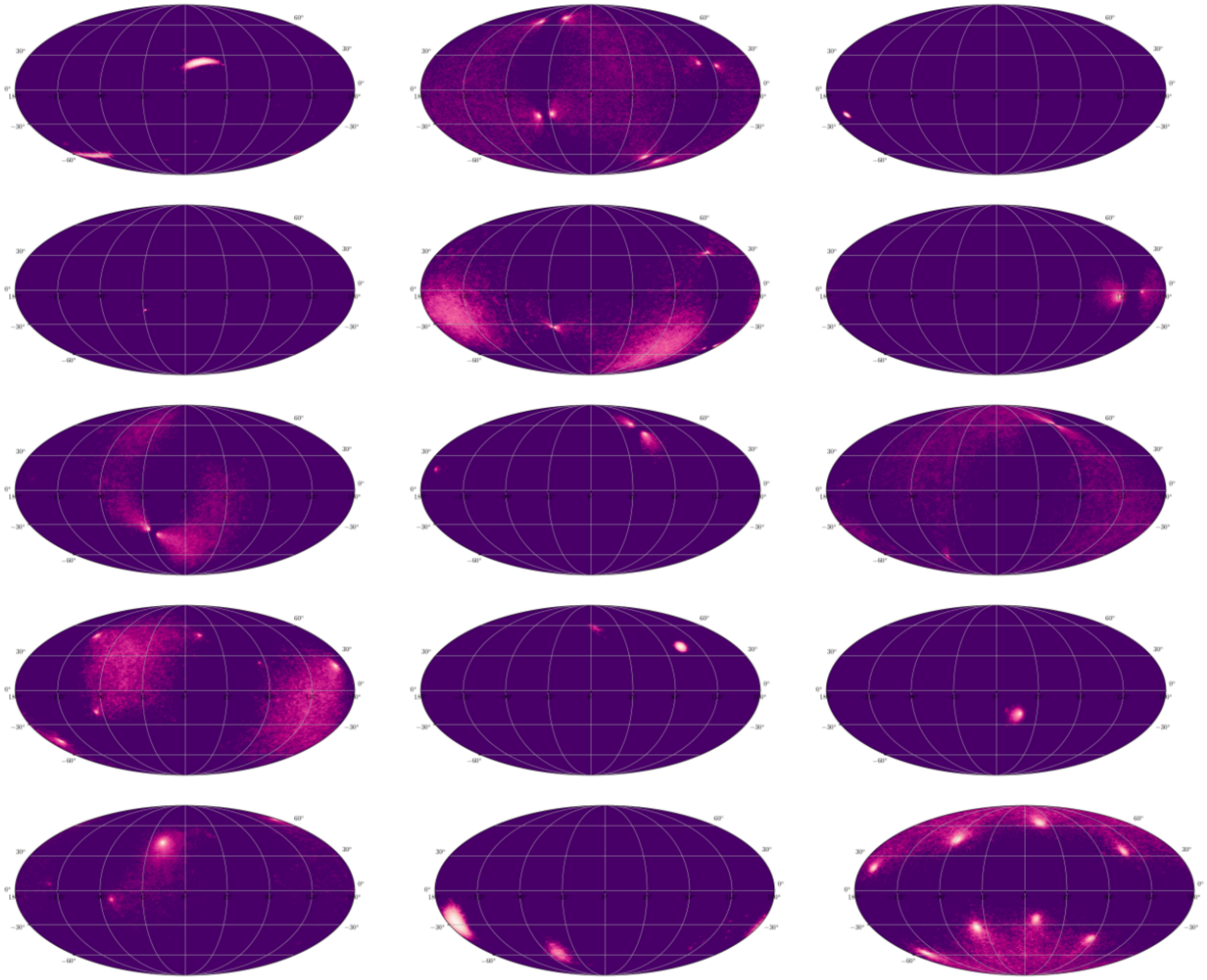


Run time ~ 2 days on AWS, $12K

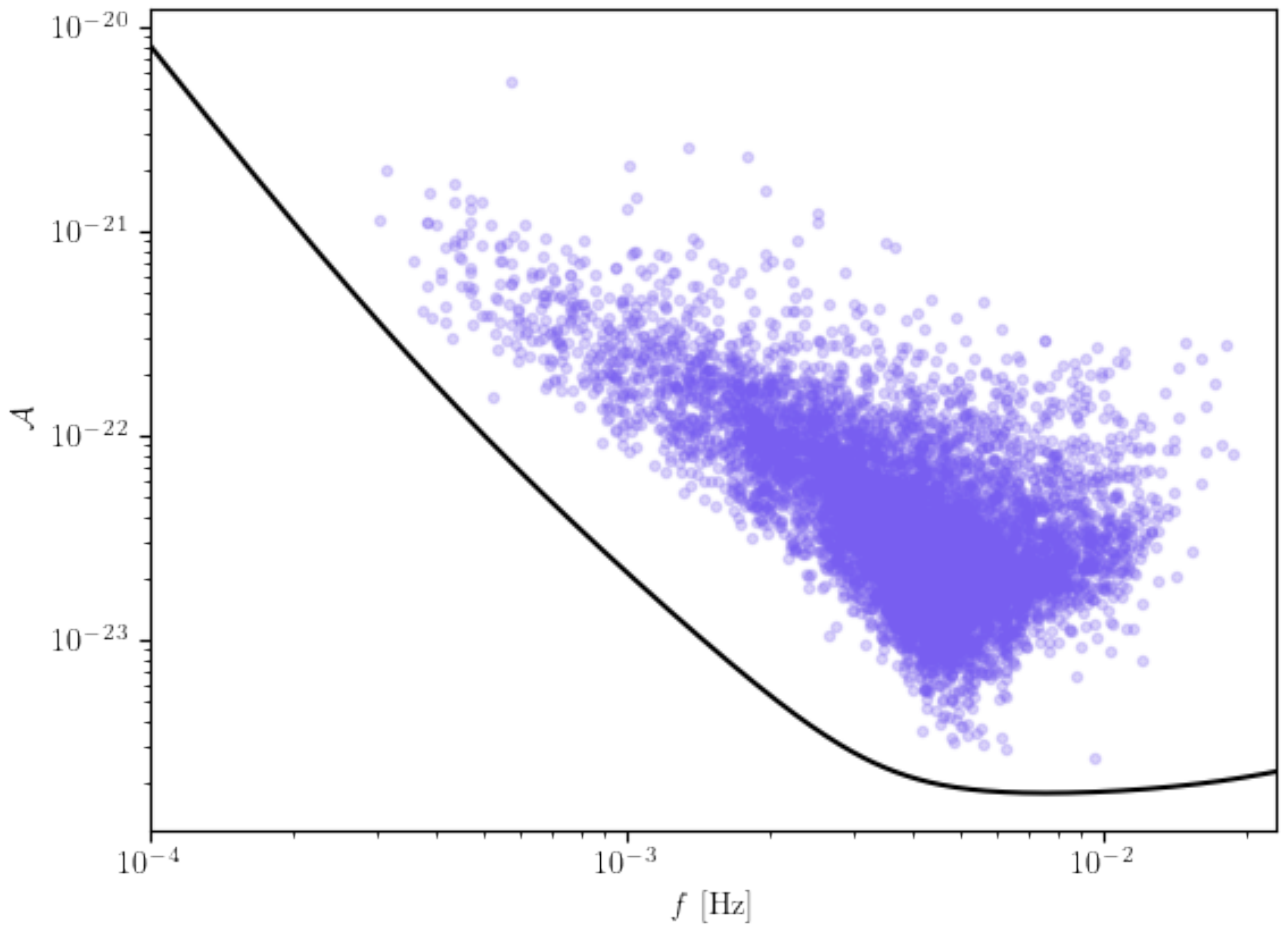
12 months of Sangria data - MBHBs



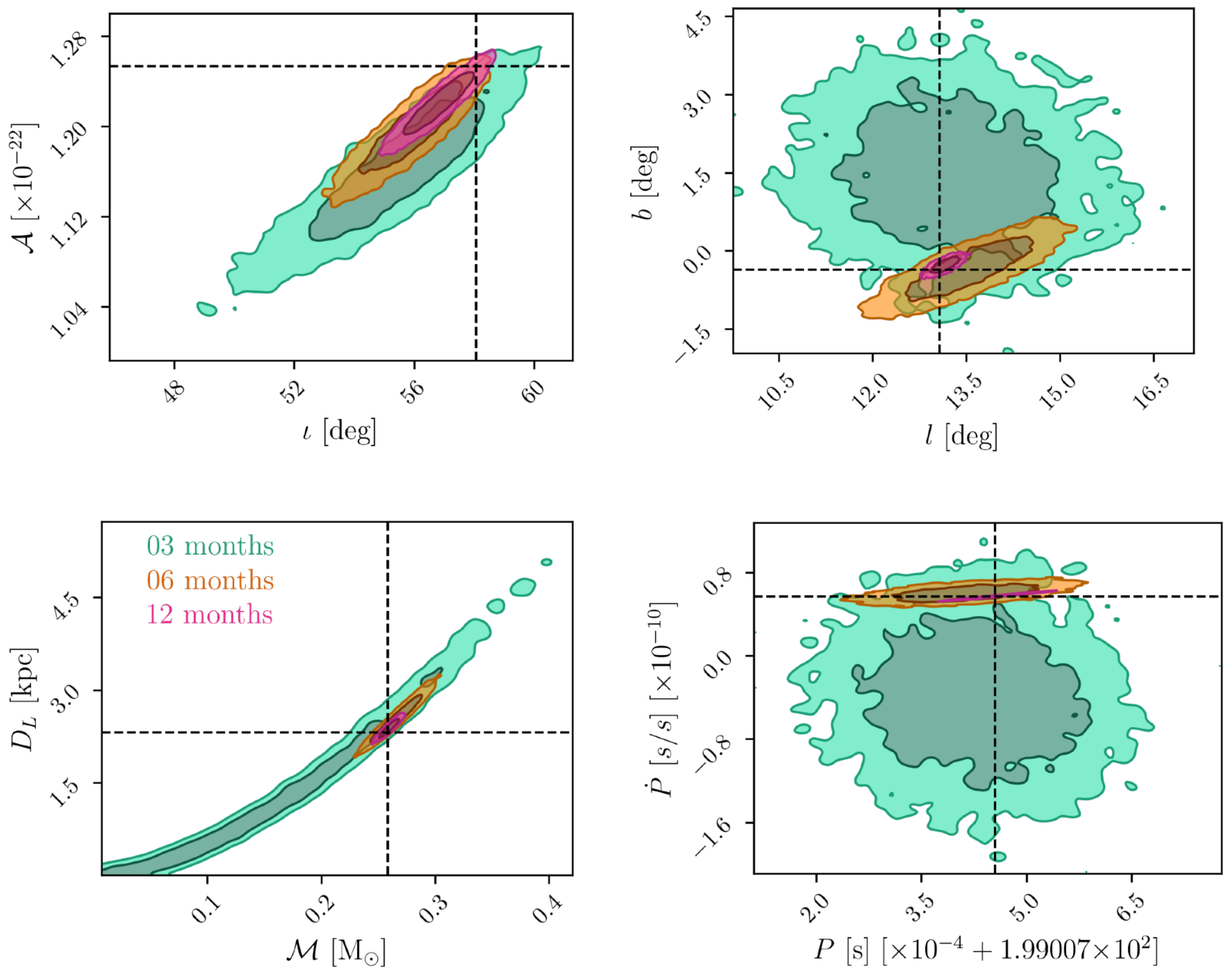




Sangria data - Galactic Binaries

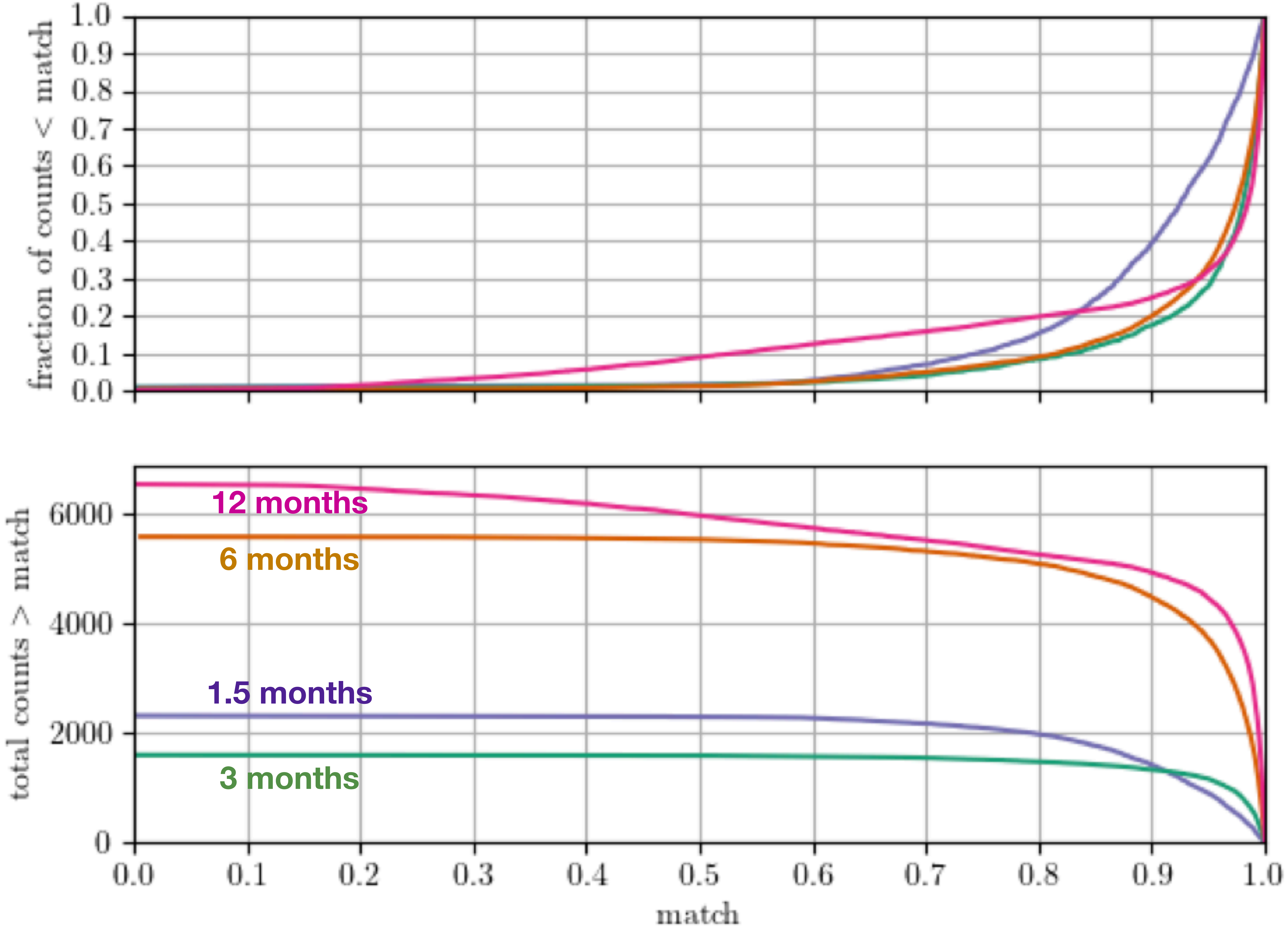


All candidate sources at 12 months



Example of how a source resolves with time

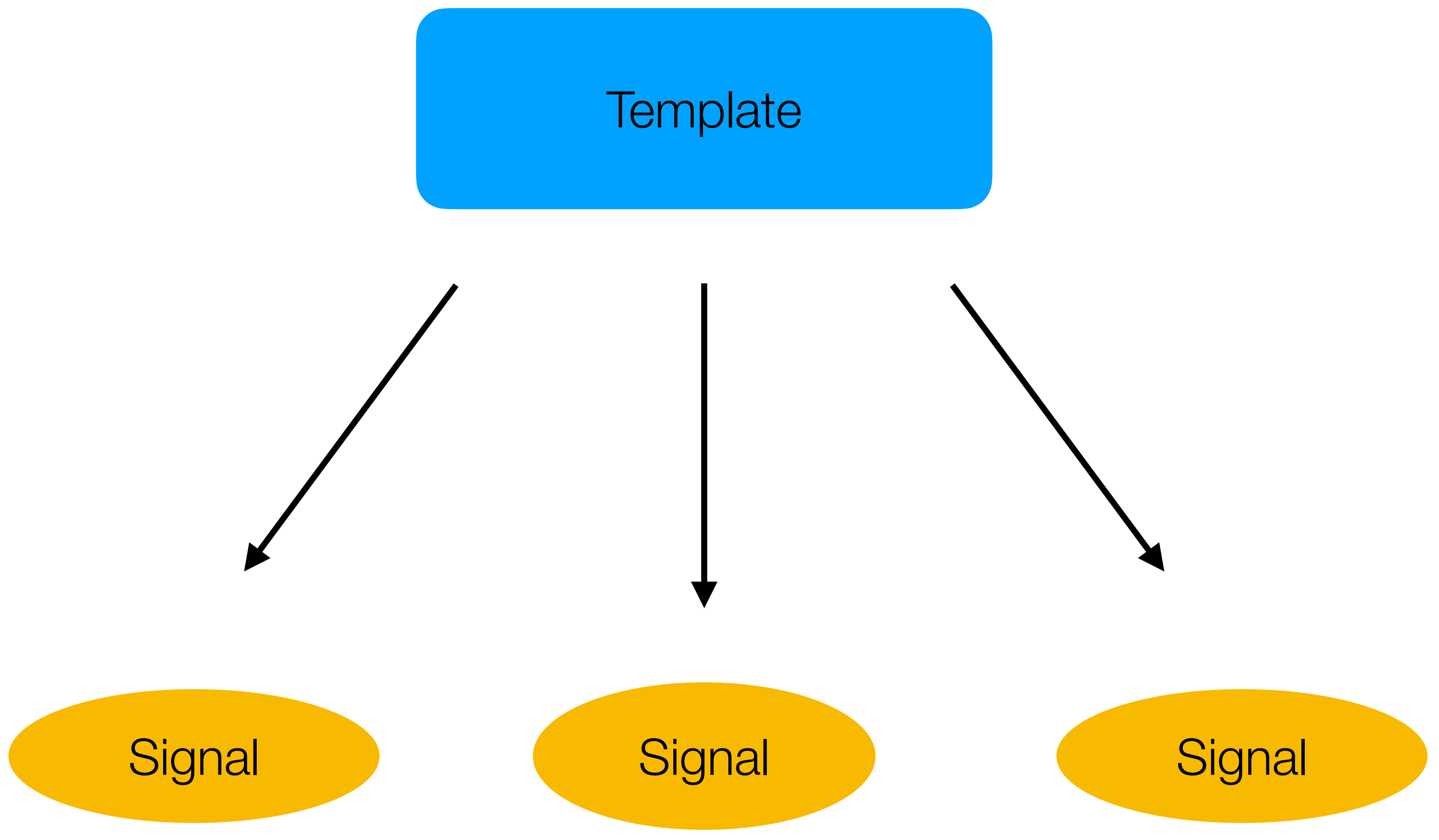
GB matches over time for 90+% confident detections



Theory: $M \approx 1 - \frac{D}{2\,\mathrm{SNR}^2}$

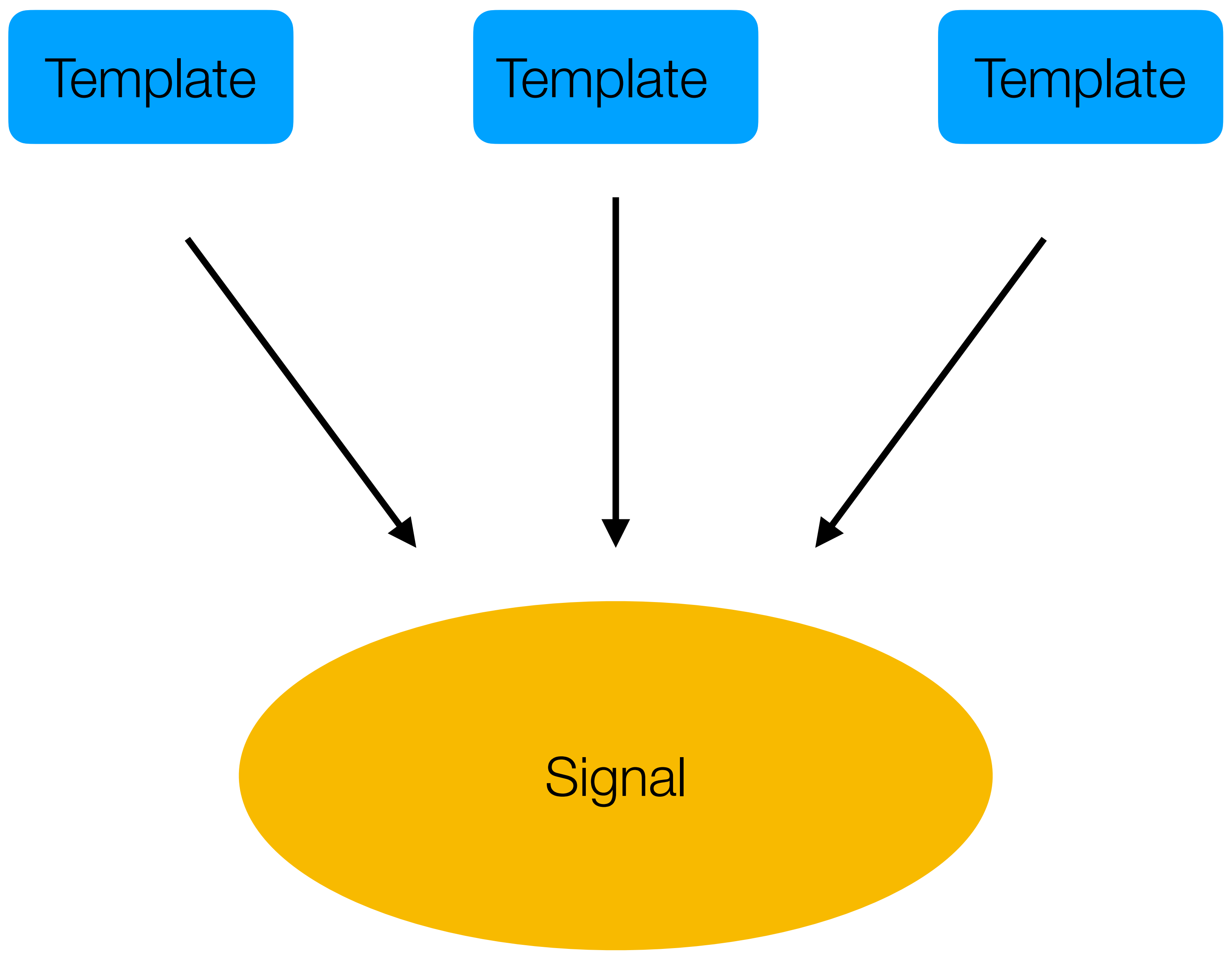
Galactic Binaries - what went wrong at 12 months?

One to Many



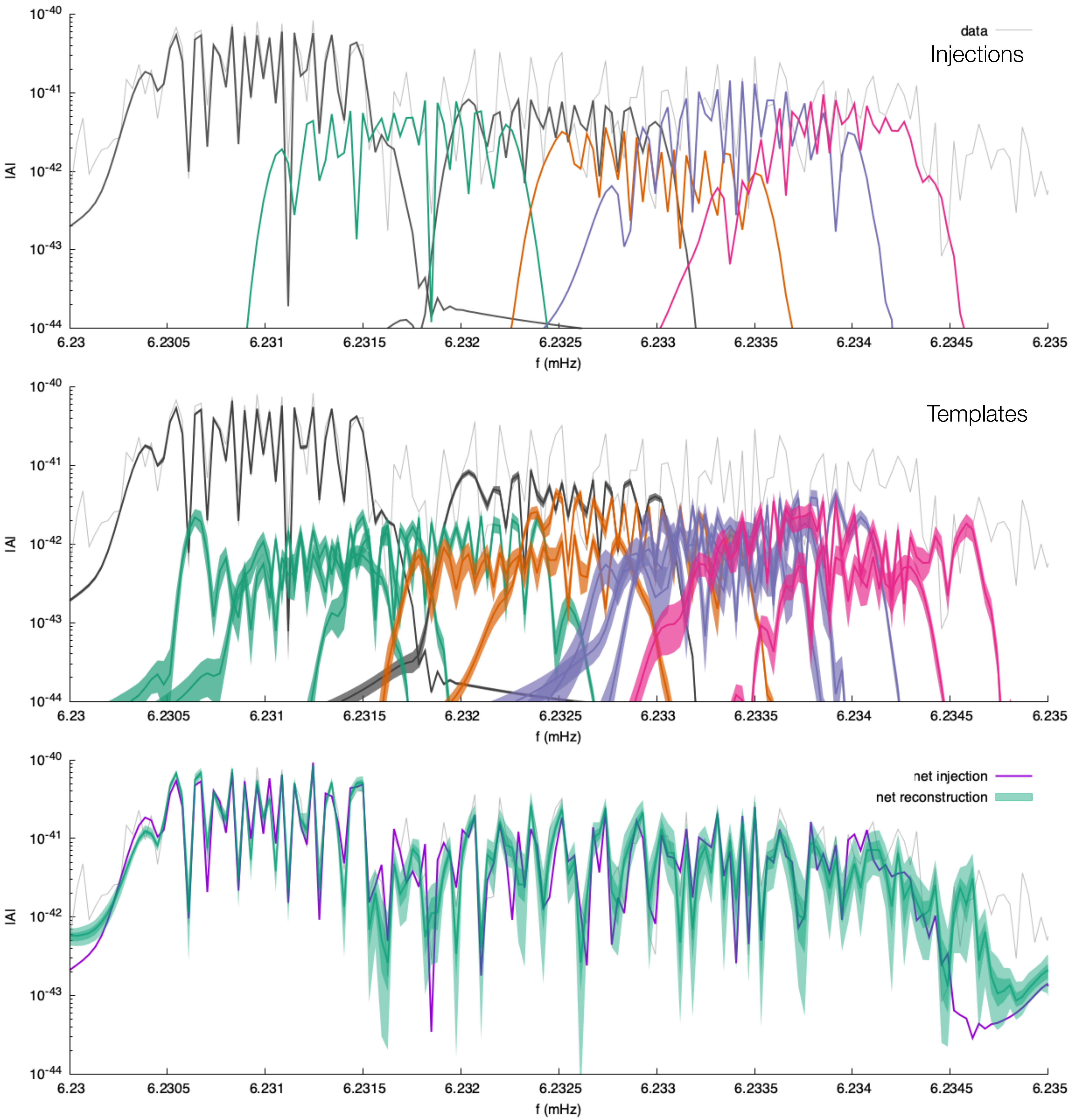
Can be the right answer in a Bayesian sense

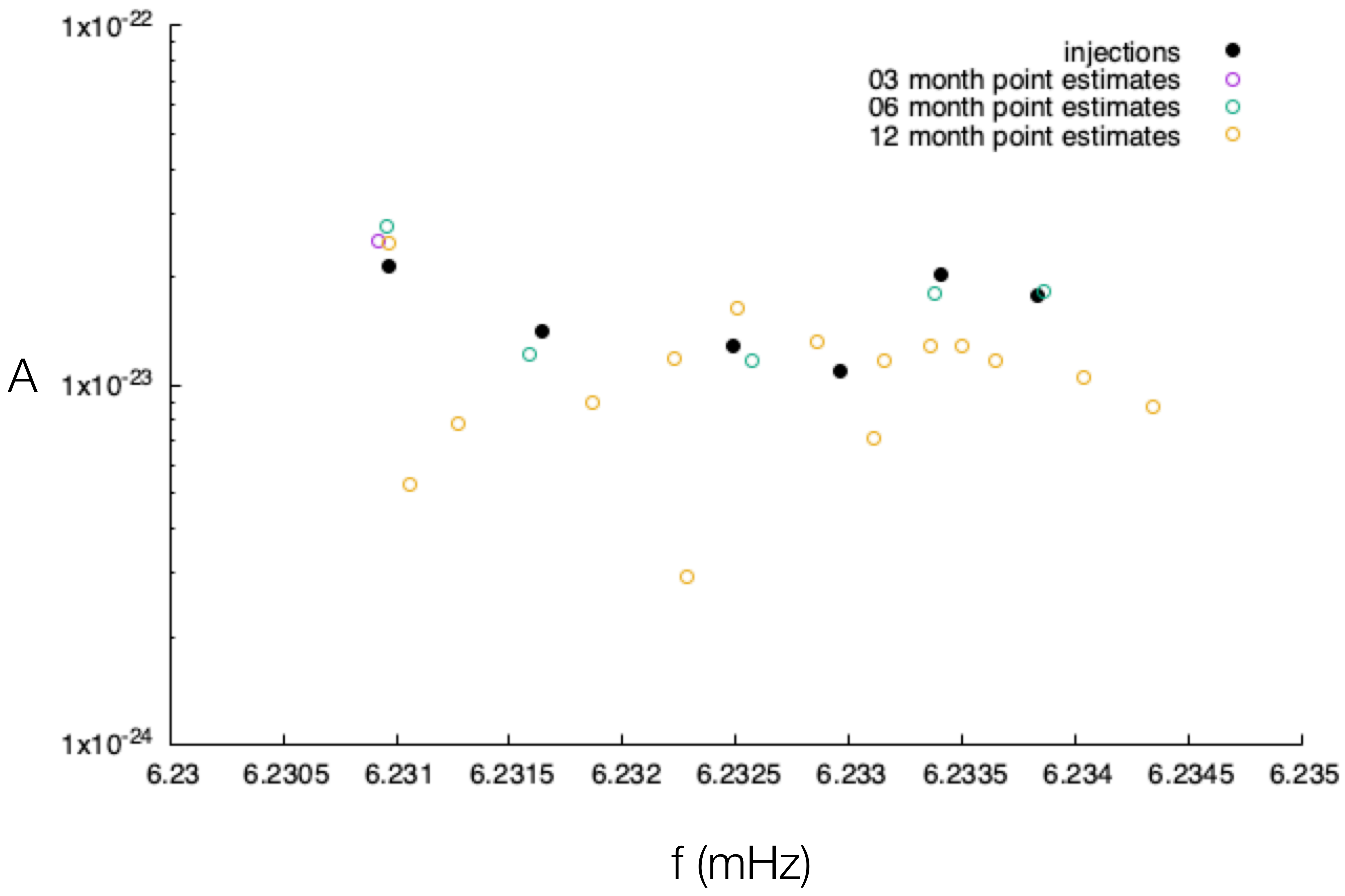
Many to One



Never the right answer - poor sampling

Galactic Binaries - what went wrong at 12 months?





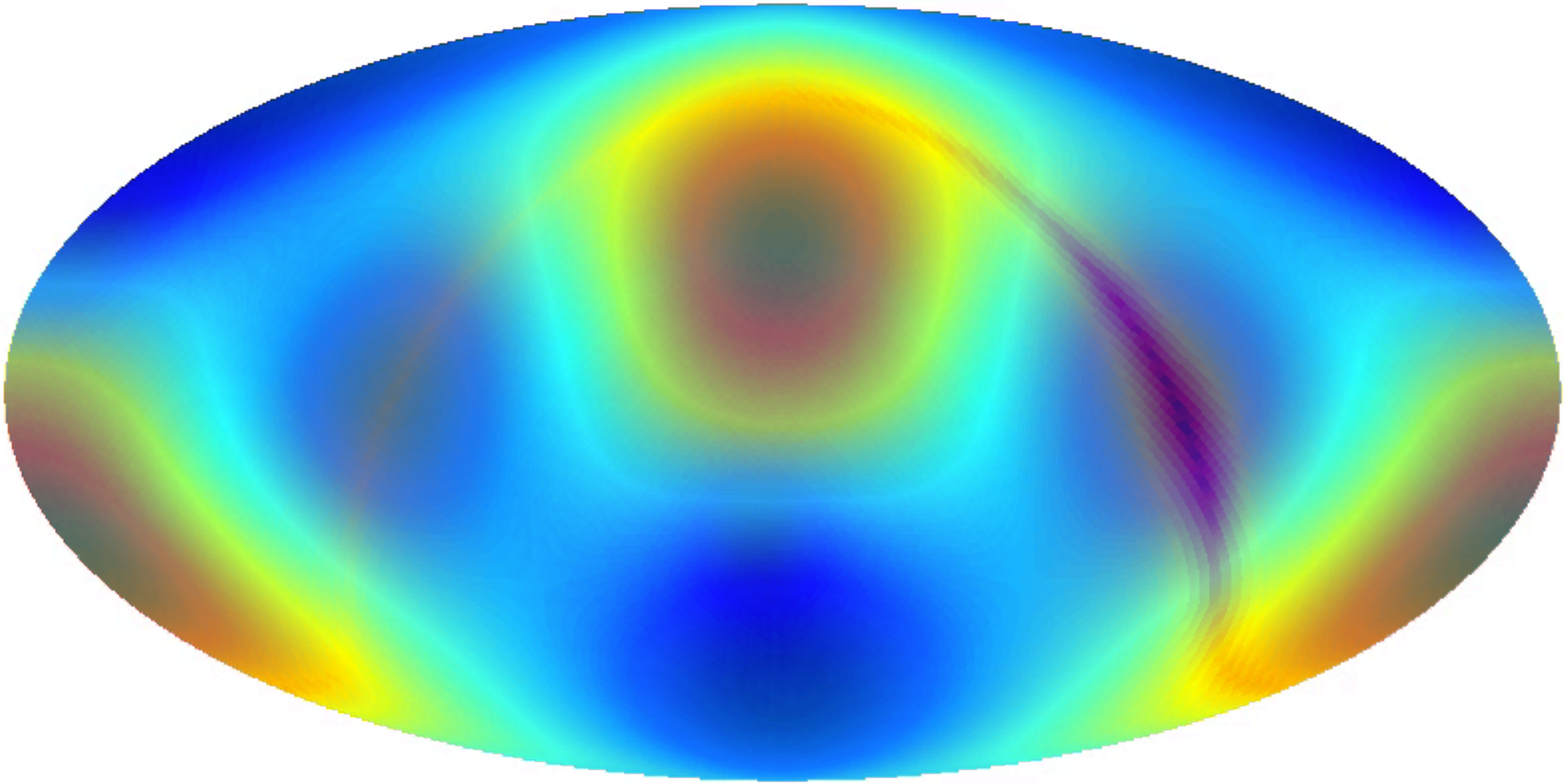
How to do better?

- Better proposals - easiest fix is to increment by smaller amounts in time
- Time-frequency modeling of signals and noise
- Include all three data channels, A, E & T
- Treat the unresolved binaries as a stochastic background (signal), and model the noise component by component
- Include a galaxy shape prior with hyper-parameters for bulge radius, disk radius, disk height etc

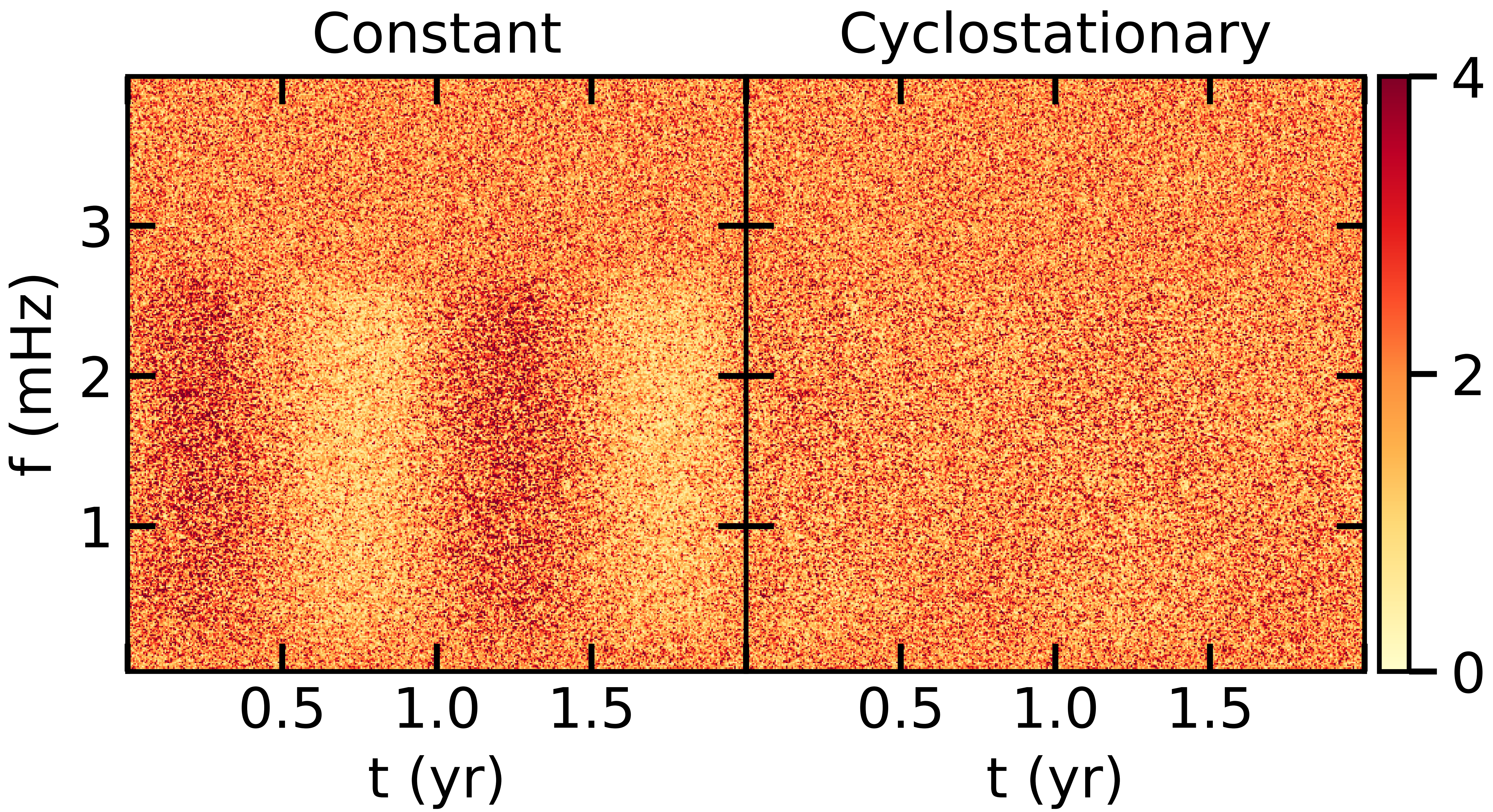
How to do better?

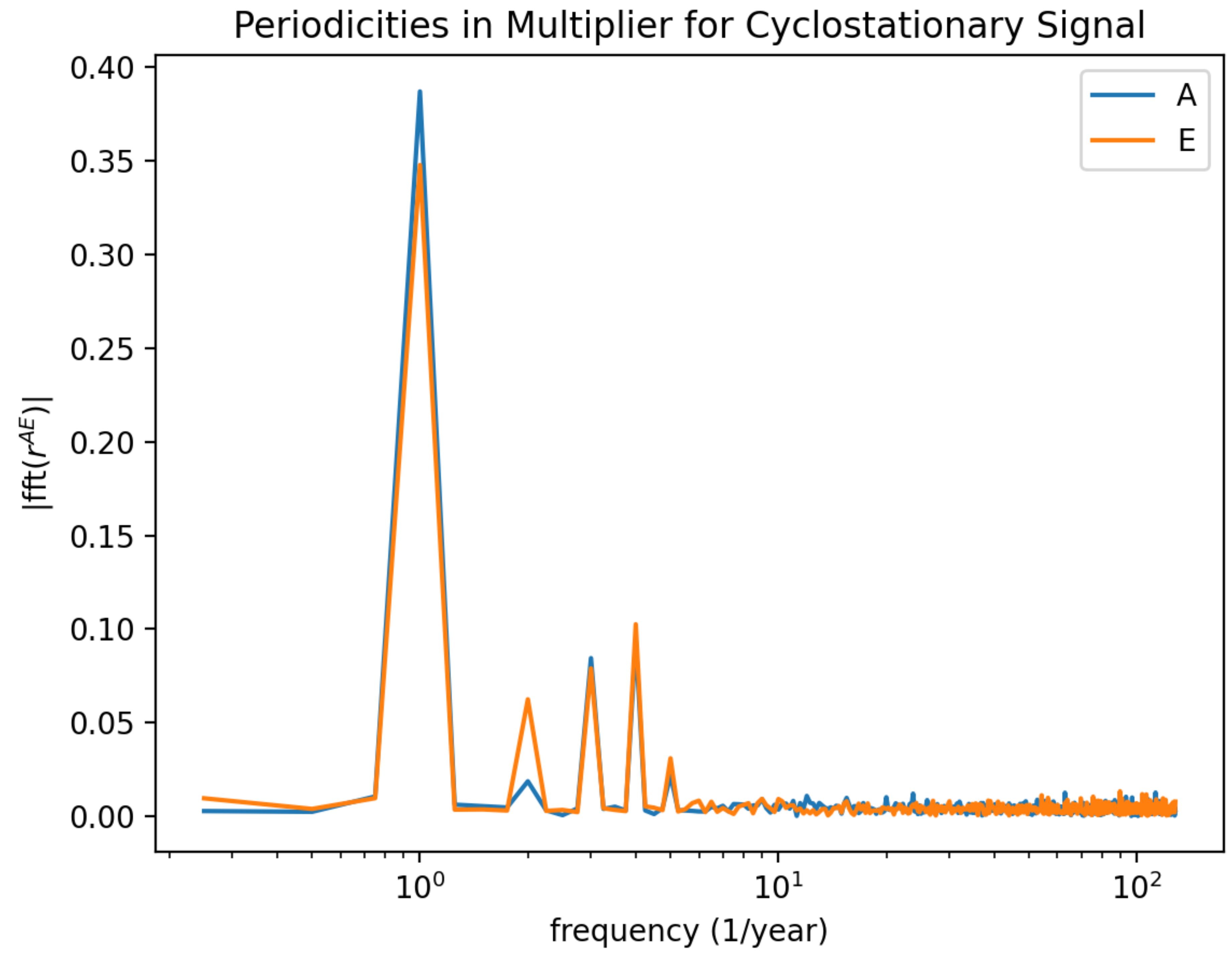
- Better proposals - easiest fix is to increment by smaller amounts in time
- Time-frequency modeling of signals and noise
- Include all three data channels, A, E & T
- Treat the unresolved binaries as a stochastic background (signal), and model the noise component by component
- Include a galaxy shape prior with hyper-parameters for bulge radius, disk radius, disk height etc

Non-Stationary Data



Cyclostationary Noise

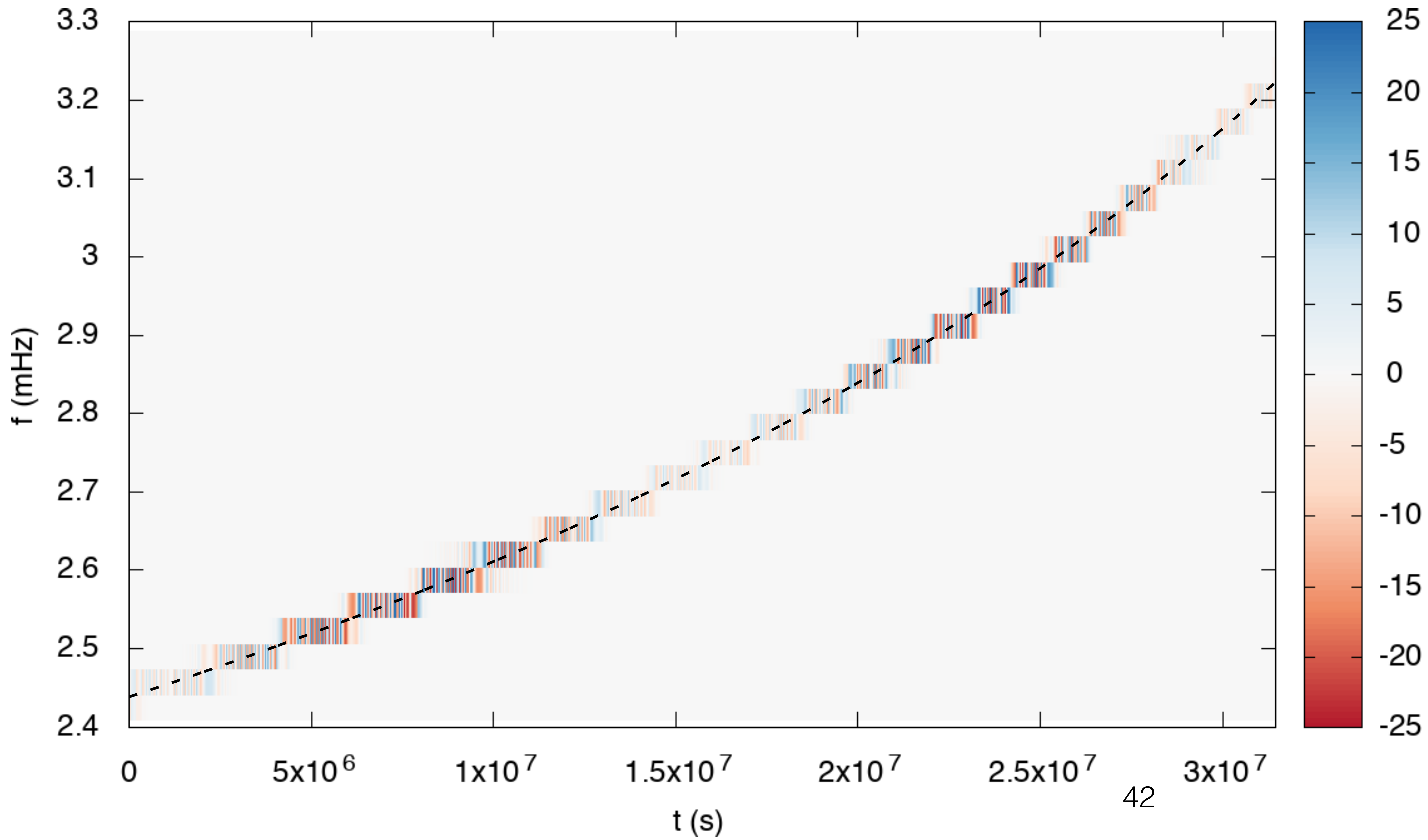




Whitening using constant PSD and dynamic PSD

Wavelet domain waveforms

$$p(\mathbf{d}|\mathbf{h}) = \frac{1}{\sqrt{\det(2\pi\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^{\dagger}\mathbf{C}^{-1}(\mathbf{d}-\mathbf{h})}$$



Fast wavelet transforms of the signals for computational efficiency

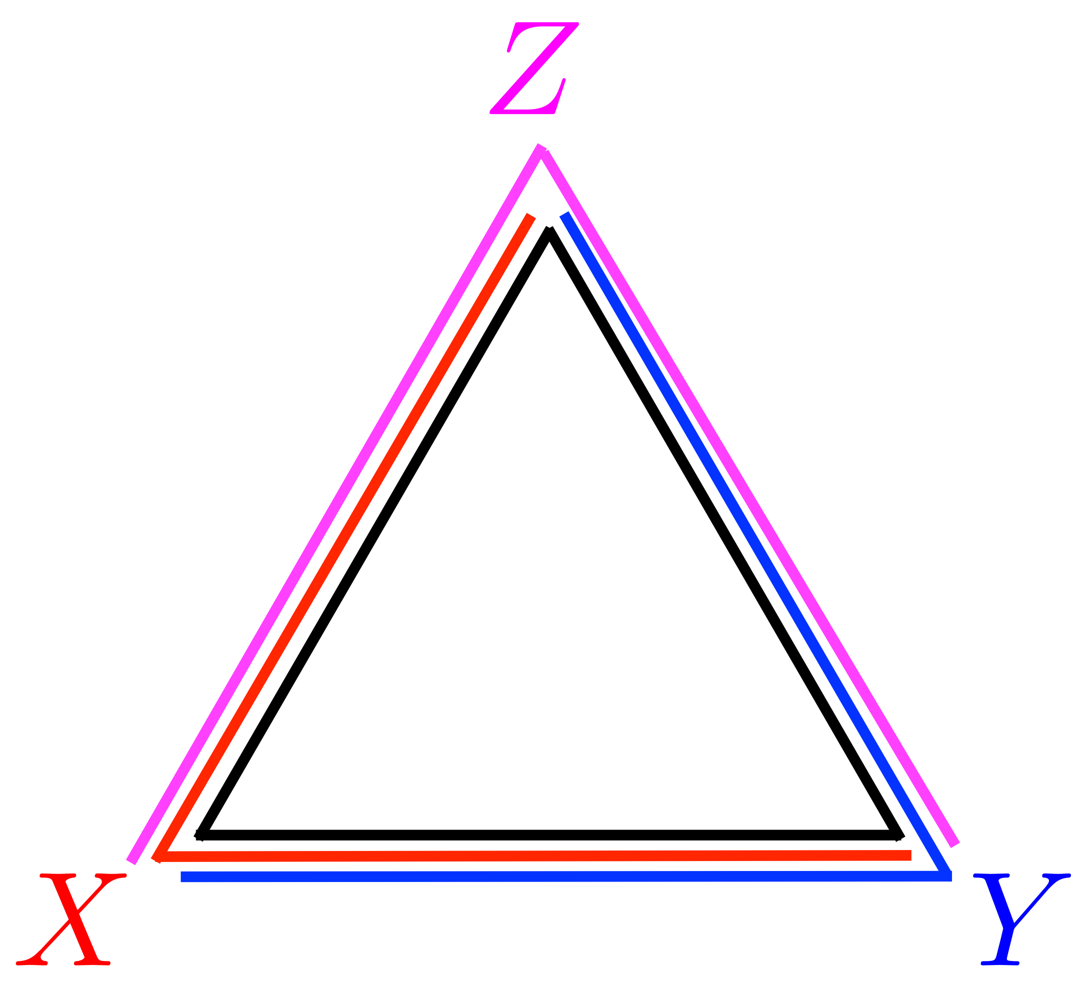
Faster than frequency domain, $\sqrt{N}$ scaling

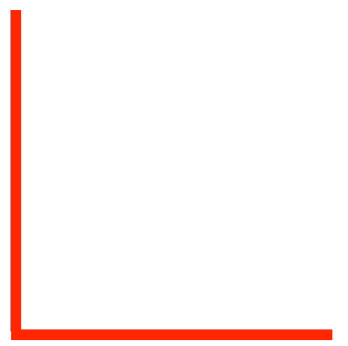
[Cornish, Phys Rev D **102,** 124038 (2020)]

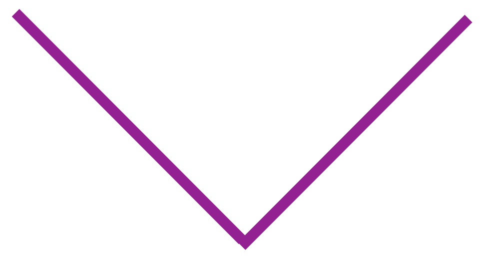
How to do better?

- Better proposals - easiest fix is to increment by smaller amounts in time
- Time-frequency modeling of signals and noise
- Include all three data channels, A, E & T
- Treat the unresolved binaries as a stochastic background (signal), and model the noise component by component
- Include a galaxy shape prior with hyper-parameters for bulge radius, disk radius, disk height etc

Multiple Data LISA Channels



A $S_+ = \frac{\sqrt{3}}{2} X \Rightarrow$ 

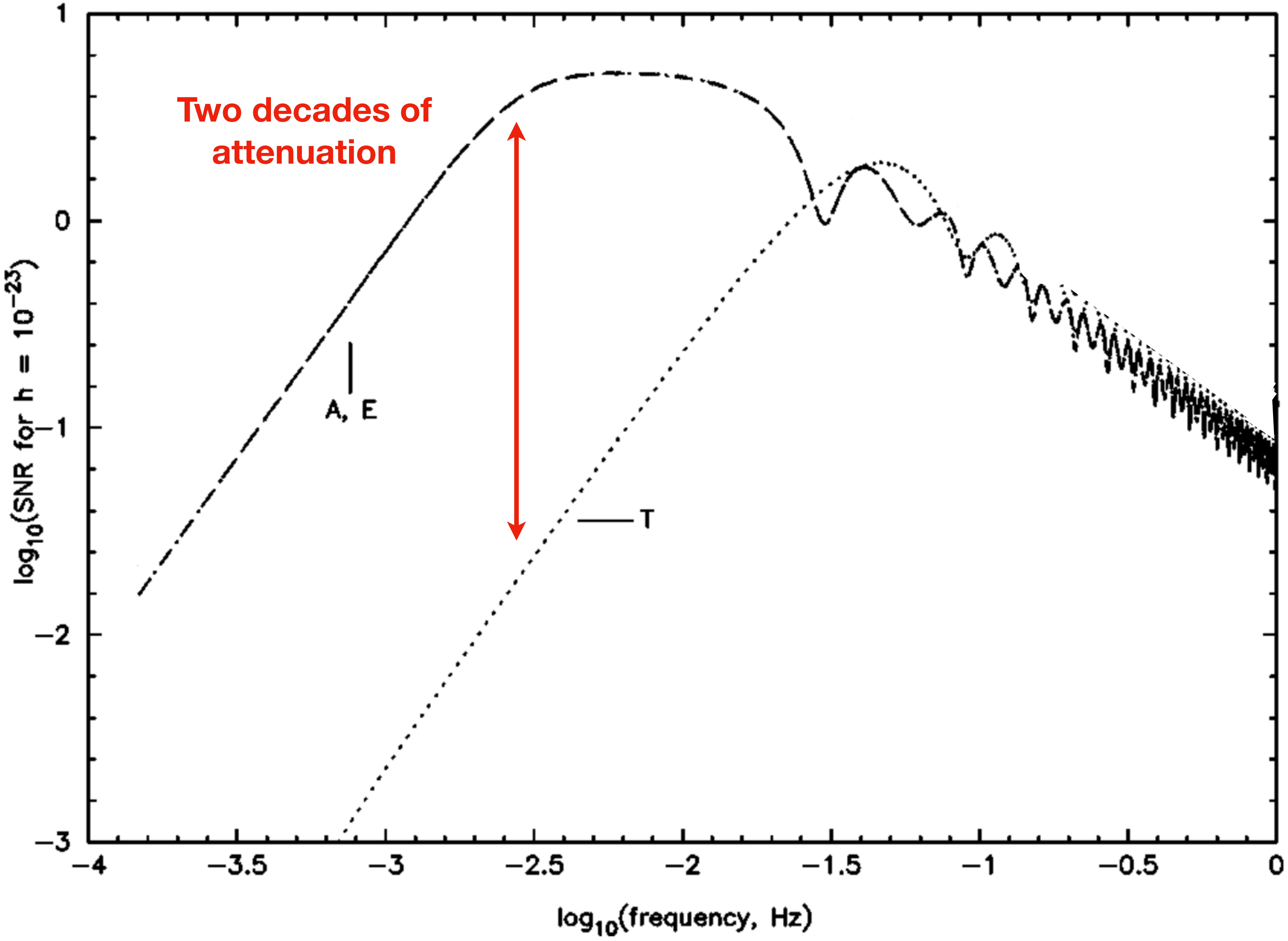
E $S_\times = \frac{1}{2}(X + 2Y) \Rightarrow$ 

} Sensitive to GWs

T $S_\odot = \frac{1}{3}(X + Y + Z) \Rightarrow$

} Insensitive to GWs

LISA Sensitivity - T channel as noise monitor



Key to detecting a stochastic background

[Prince et al, Phys. Rev. D66 (2002)]

[Tinto, Armstrong & Estabrook, Phys.Rev.D63 (2001)]

[Hogan & Bender, Phys. Rev. D64 (2001)]

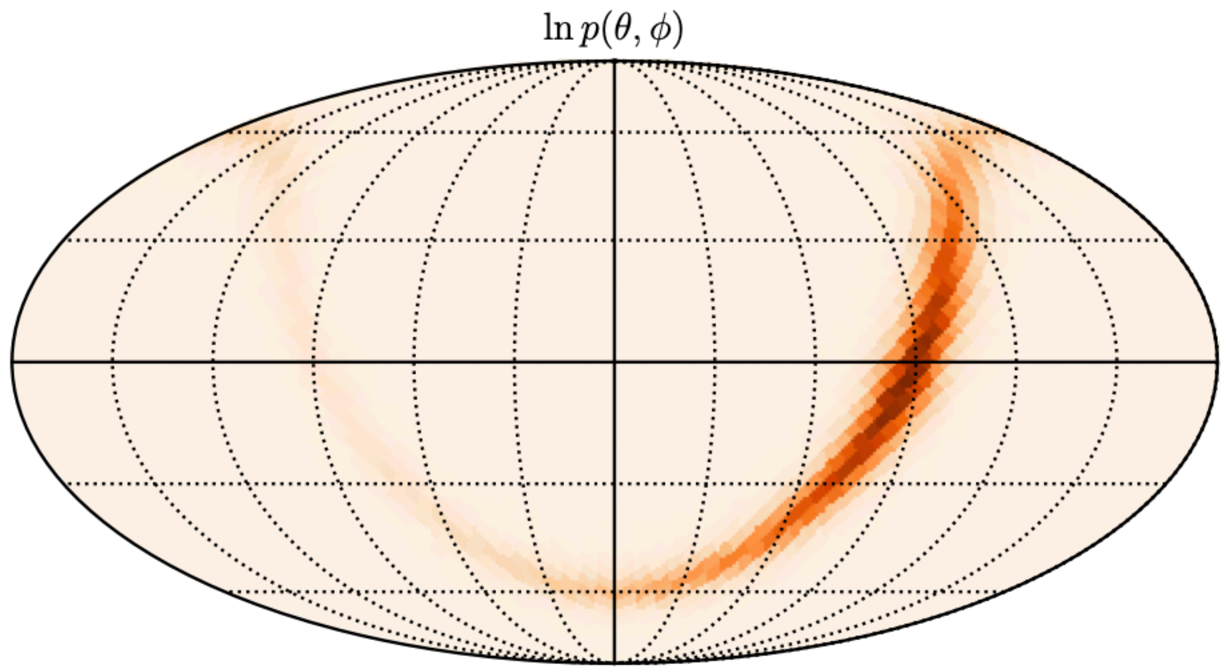
[Adams & Cornish, Phys. Rev. D86 (2010)]

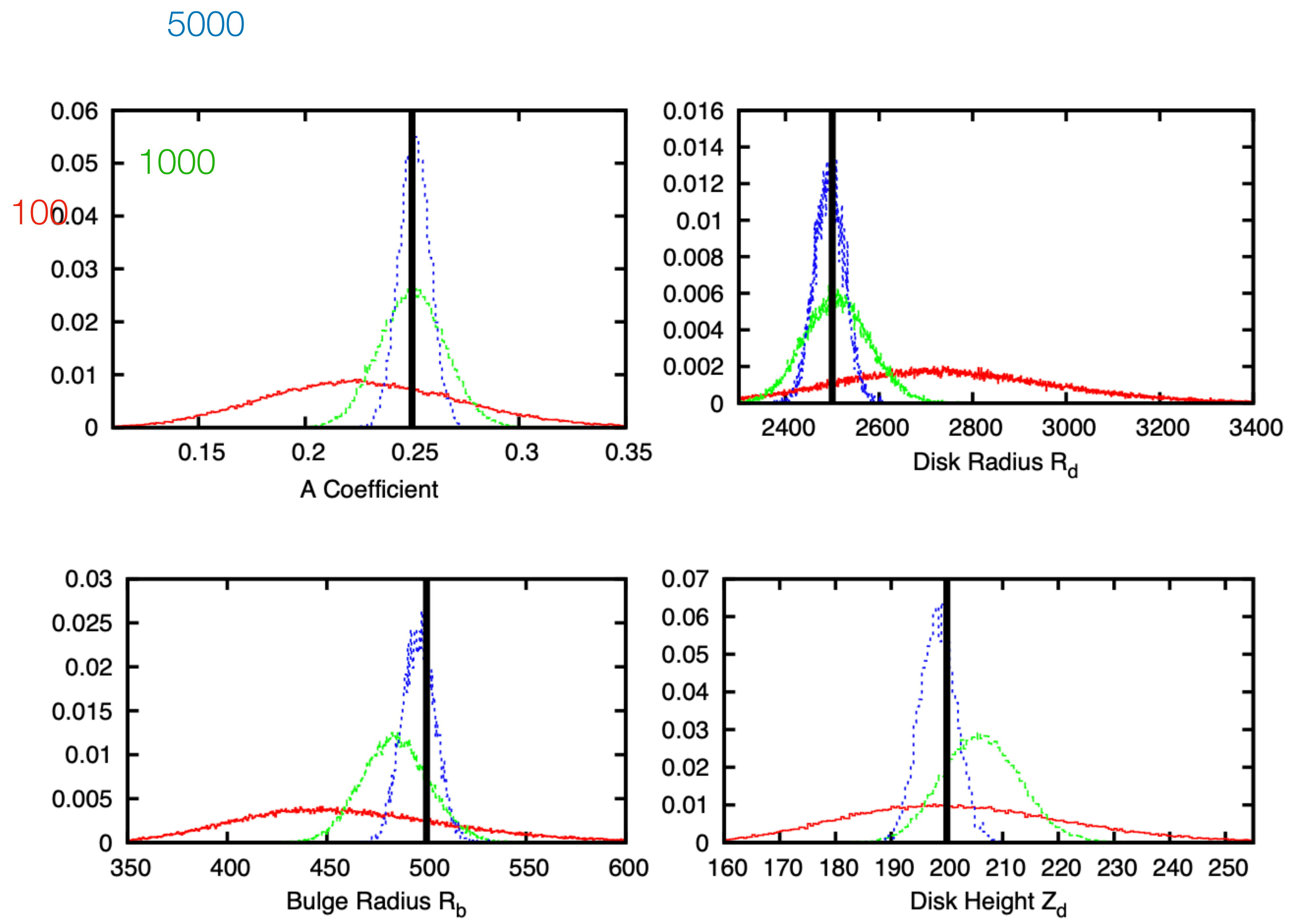
How to do better?

- Better proposals - easiest fix is to increment by smaller amounts in time
- Time-frequency modeling of signals and noise
- Include all three data channels, A, E & T
- Treat the unresolved binaries as a stochastic background (signal), and model the noise component by component
- Include a galaxy shape prior with hyper-parameters for bulge radius, disk radius, disk height etc

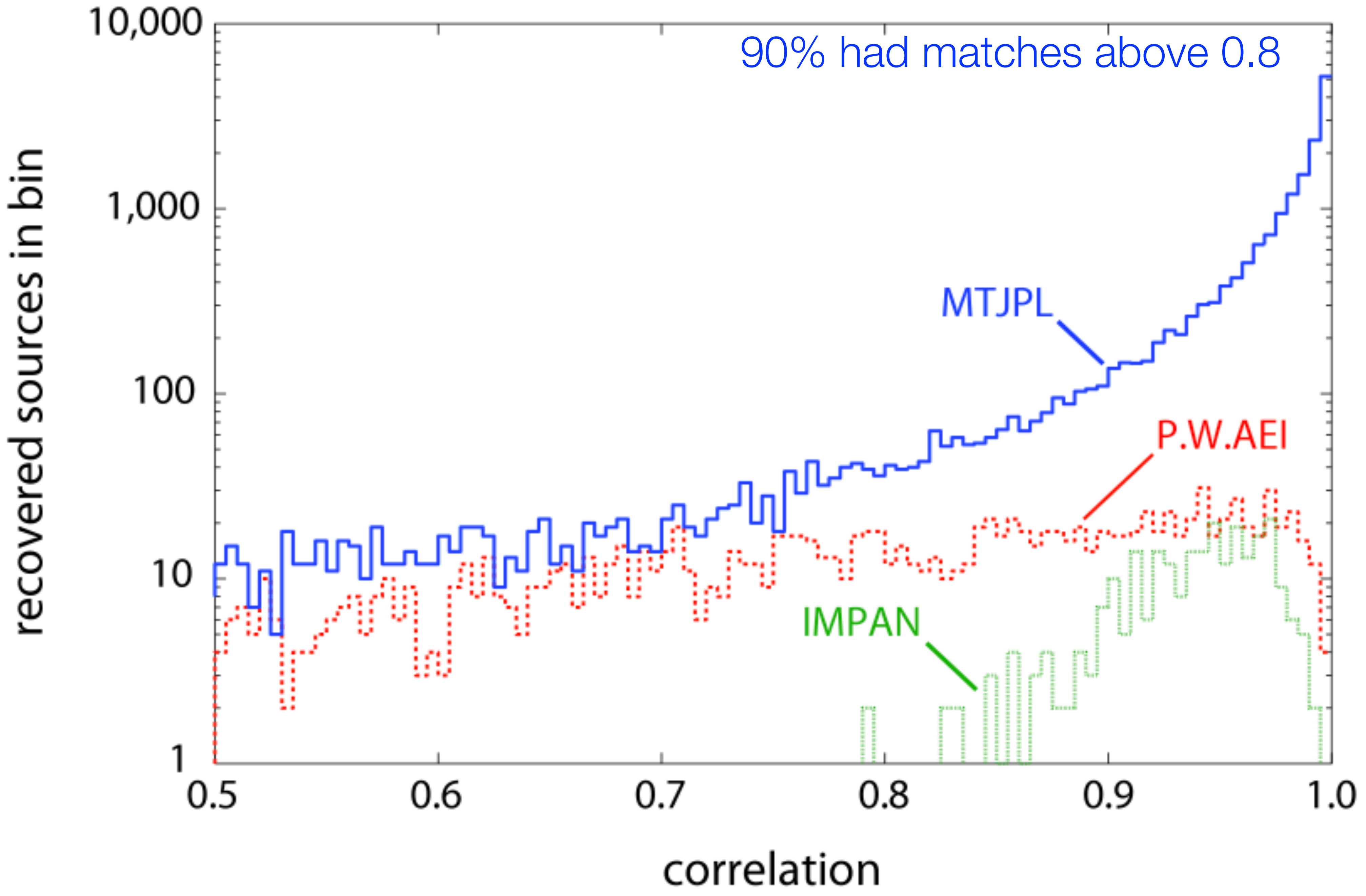
Galaxy shape prior with hyper-parameters

$$\rho(x, y, z) = \rho_0 \left(A_b e^{-r^2/R_b^2} + (1 - A_b) e^{-\sqrt{x^2 + y^2}/R_d} \operatorname{sech}^2(z/Z_d) \right)$$





[Adams, Cornish & Littenberg, PRD 86, 124032 (2012)]



Crowder & Cornish (2007)