Reconstruction of GWBs at LISA: SGWBinner and Saqqara

Mauro Pieroni

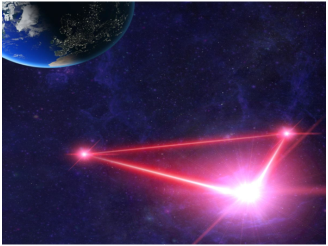
European Organization for Nuclear Research (CERN)
mauro.pieroni@cern.ch

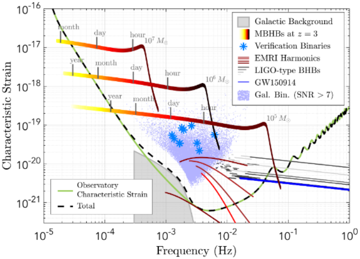
"Towards realistic detection forecasts of primordial GWBs"
IFIC, Valencia

December 10th, 2024

Outline

Laser Interferometer Space Antenna

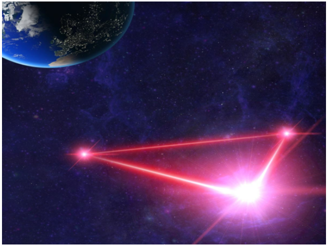


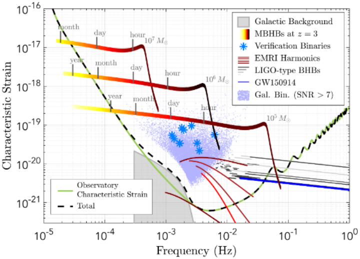


Few details on LISA:

- First GW interferometer in space
- Constellation of three satellites
- 2.5 million km arm lengths
- Peak sensitivity $10^{-2} \div 10^{-3}$Hz
- Three correlated detectors
- Expected launch in mid 2030
- Operating for 4.5 yrs (nominal)

Laser Interferometer Space Antenna





Few details on LISA:

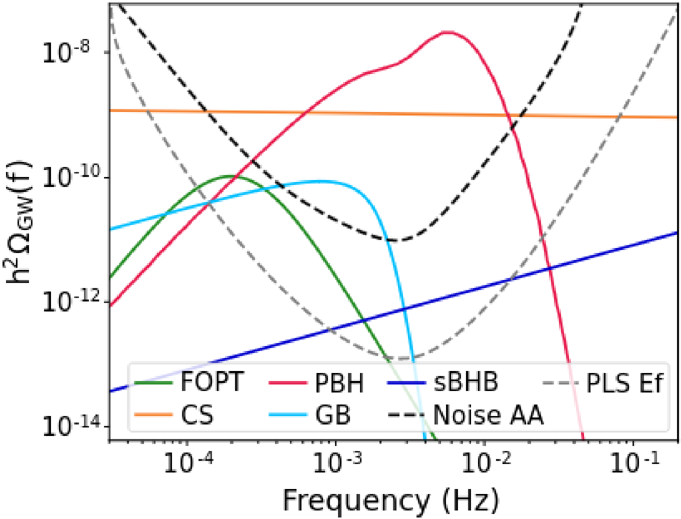
- First GW interferometer in space
- Constellation of three satellites
- 2.5 million km arm lengths
- Peak sensitivity $10^{-2} \div 10^{-3}$Hz
- Three correlated detectors
- Expected launch in mid 2030
- Operating for 4.5 yrs (nominal)

Very interesting for cosmology since we can (among others):

- Measure H_0
- Test modified gravity
- (Hopefully) detect and characterize GWBs!

* Figures from:
https://www.lisamission.org/multimedia/image/lisa-astro2020
LISA Collaboration, P. Amaro-Seoane et al., ArXiv: 1702.00786

Sources of GWBs in the LISA



* Figure from M. Colpi et al., ArXiv:2402.07571

Basics of GW data analysis

Data $\tilde{d}$ (in frequency space) $\longrightarrow$ $\tilde{d} = \tilde{s} + \tilde{n}$

- For individual sources $\langle \tilde{s} \rangle \neq 0$
- For GWBs $\langle \tilde{s} \rangle = 0$
- For noise $\langle \tilde{n} \rangle = 0$

For an isotropic GWB $\longrightarrow$ $\langle h_\lambda(\vec{k})\, h^*_{\lambda'}(\vec{k}')\rangle \propto \delta_{\lambda\lambda'} P_h^\lambda(k) \delta(\vec{k} - \vec{k}')$

Assuming $\langle \tilde{s}\tilde{n} \rangle = 0$ and Gaussian signal and noise

$$\left\langle \tilde{d}^2 \right\rangle = \left\langle \tilde{s}^2 \right\rangle + \left\langle \tilde{n}^2 \right\rangle = \sum_\lambda \mathcal{R}_\lambda\, P_h^\lambda + N \equiv \mathcal{R}\left[P_h + S_n\right]$$

where we have introduced

- The (quadratic) response function of the instrument $\mathcal{R}$
- The (intensity of the) signal power spectrum P_h (in 1/Hz)
- The noise power spectrum N (in 1/Hz)
- The (square of the) Strain sensitivity S_n (in 1/Hz)

Basics of GW data analysis

Data $\tilde{d}$ (in frequency space) $\longrightarrow$ $\tilde{d} = \tilde{s} + \tilde{n}$

- For individual sources $\langle \tilde{s} \rangle \neq 0$
- For GWBs $\langle \tilde{s} \rangle = 0$
- For noise $\langle \tilde{n} \rangle = 0$

For an isotropic GWB $\longrightarrow$ $\langle h_\lambda(\vec{k})\, h^*_{\lambda'}(\vec{k}') \rangle \propto \delta_{\lambda\lambda'} P_h^\lambda(k) \delta(\vec{k} - \vec{k}')$

Assuming $\langle \tilde{s}\tilde{n} \rangle = 0$ and Gaussian signal and noise

$$\left\langle \tilde{d}^2 \right\rangle = \left\langle \tilde{s}^2 \right\rangle + \left\langle \tilde{n}^2 \right\rangle = \sum_\lambda \mathcal{R}_\lambda \, P_h^\lambda + N \equiv \mathcal{R}\left[P_h + S_n\right]$$

where we have introduced

- The (quadratic) response function of the instrument $\mathcal{R}$
- The (intensity of the) signal power spectrum P_h (in 1/Hz)
- The noise power spectrum N (in 1/Hz)
- The (square of the) Strain sensitivity S_n (in 1/Hz)

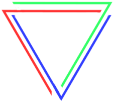
In order to compare with cosmological predictions it's customary to introduce

$$\Omega_{\rm GW} \equiv \frac{1}{3H_0^2 M_p^2} \frac{\partial \rho_{\rm GW}}{\partial \ln f} = \frac{4\pi^2}{3H_0^2} f^3 P_h \qquad \text{and} \qquad \Omega_n(f) = \frac{4\pi^2}{3H_0^2} f^3 S_n(f) \,,$$

where $H_0 \simeq h_0 \times 3.24 \times 10^{-18}\,\mathrm{Hz}$ is the Hubble parameter today.

Time Delay Interferometry

The simplest option is to build three (correlated) Michelson-like data streams

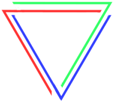


Laser frequency noise is **huge** and must be suppressed with Time Delay Interferometry (TDI)

TDI = Combine measurements at different times to achieve noise reduction!

Time Delay Interferometry

The simplest option is to build three (correlated) Michelson-like data streams

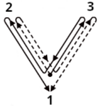


Laser frequency noise is **huge** and must be suppressed with Time Delay Interferometry (TDI)

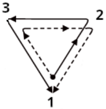
TDI = Combine measurements at different times to achieve noise reduction!

Using η_{ij} measurement in i coming from j and the delay operator D_{ij} we define:

- Michelson variables, dubbed XYZ, defined as (YZ are permutations):
  $X \equiv (1 - D_{13}D_{31})(\eta_{12} + D_{12}\eta_{21}) + (D_{12}D_{21} - 1)(\eta_{13} + D_{13}\eta_{31})$

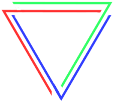


- Sagnac variables, dubbed $\alpha\beta\gamma$, defined as ($\beta\gamma$ permutations):
  $\alpha \equiv \eta_{12} + D_{12}\eta_{23} + D_{12}D_{23}\eta_{31} - (\eta_{13} + D_{13}\eta_{32} + D_{13}D_{32}\eta_{21})$



Time Delay Interferometry

The simplest option is to build three (correlated) Michelson-like data streams

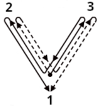


Laser frequency noise is **huge** and must be suppressed with Time Delay Interferometry (TDI)

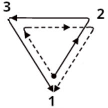
TDI = Combine measurements at different times to achieve noise reduction!

Using η_{ij} measurement in i coming from j and the delay operator D_{ij} we define:

- Michelson variables, dubbed XYZ, defined as (YZ are permutations):
  $X \equiv (1 - D_{13}D_{31})(\eta_{12} + D_{12}\eta_{21}) + (D_{12}D_{21} - 1)(\eta_{13} + D_{13}\eta_{31})$



- Sagnac variables, dubbed $\alpha\beta\gamma$, defined as ($\beta\gamma$ permutations):
  $\alpha \equiv \eta_{12} + D_{12}\eta_{23} + D_{12}D_{23}\eta_{31} - (\eta_{13} + D_{13}\eta_{32} + D_{13}D_{32}\eta_{21})$



In these variables, signal and noise (in different channels) are correlated!

For equal arms, diagonalization via:
(diagonal variables dubbed AET and $\mathcal{AET}$)

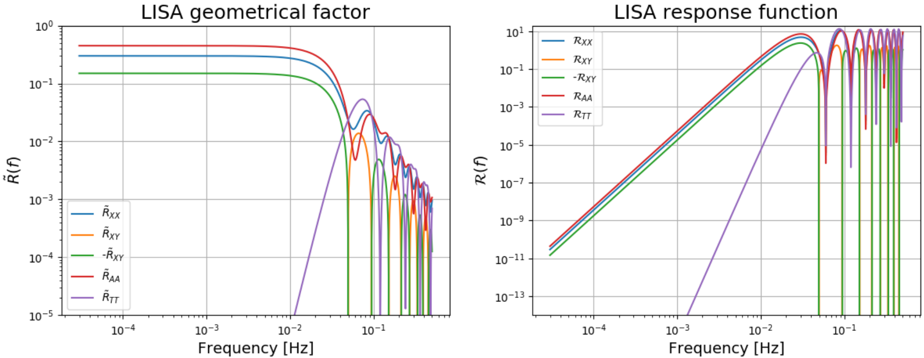
$$\longrightarrow \begin{pmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}$$

The LISA response function

For an isotropic and non-chiral spectrum we get (see, e.g., 2009.11845):

$$\left\langle s_i^{TDI} s_j^{TDI} \right\rangle = \int \mathrm{d}k\, P_h(k) \mathcal{R}_{ij}^{TDI}(k)\,, \qquad \mathcal{R}_{ij}^{TDI}(k) \equiv 4\,(2\pi k L)^2 |W^{TDI}(kL)|^2\, \tilde{R}_{ij}^{TDI}(k)\,.$$

where $\mathcal{R}_{ij}^{TDI}(k)$ is the LISA response and $\tilde{R}_{ij}^{TDI}(k)$ is a geometrical factor.
For XYZ/AET (AET is $\sim$ diagonal) combinations we get:



At low frequencies the TT response is suppressed by a factor f^6!
See https://github.com/Mauropieroni/GW_response

The LISA noise model

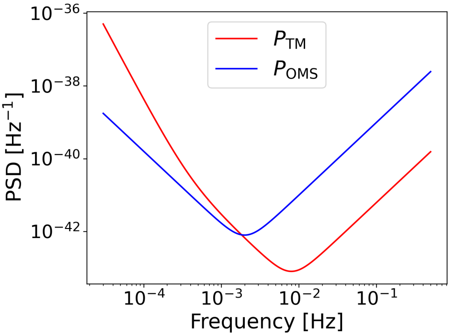
After TDI two main noise components:

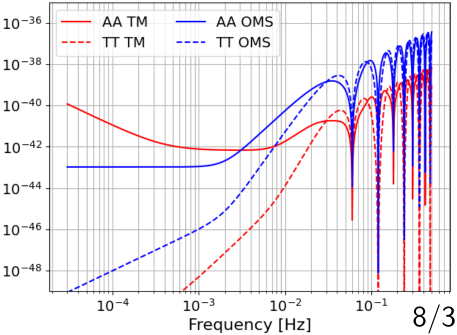
Low frequencies are dominated by Test Mass (TM) noise
large frequencies by Optical Metrology System (OMS) noise:

$$P_{TM}(f, A) = A^2 \times 10^{-30} \times F_{TM}(f) \,,$$

$$P_{OMS}(f, P) = P^2 \times 10^{-24} \times F_{OMS}(f) \,,$$

where $F_{TM}(f)$, $F_{OMS}(f)$ are some functions of frequency.

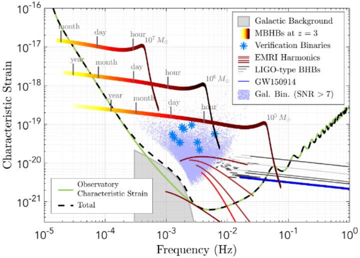


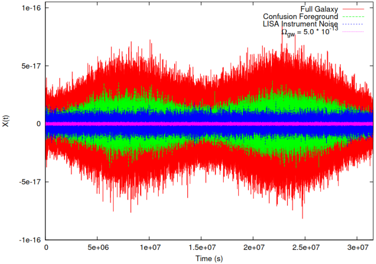


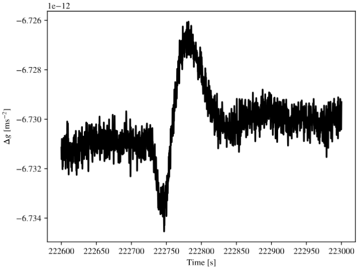
Stationarity won't hold ...

The LISA data won't be stationary:

- There will be transient signals
- Measured GWBs change with time
- There will be transient gliches

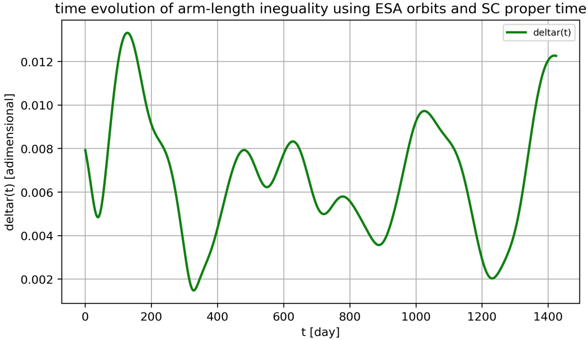






* LISA Collaboration, P. Amaro-Seoane et al., ArXiv: 1702.00786. M. R. Adams, Phys. Rev. D 89 (2014), 022001, ArXiv: 1307.4116. Q. Baghi et al., Phys. Rev. D 105 (2022), 042002, ArXiv: 2112.07490.

... LISA won't have equal arms...



Fluctuations in the arm-lengths of order up to 10^{-2} are expected!

Response functions and noise spectra will be modified $\longrightarrow$

- Orthogonality of TDI variables might be affected
- T is not signal orthogonal anymore
- ...

An accurate description is necessary to avoid biases in the analysis!

... nor the same noise levels in all links!

Let us have a closer look at the problem of noise characterization
(still stick with TM and OMS noise only with known templates)

Each spacecraft contains two test masses and two lasers $\longrightarrow$ 12 (6 Acc +6 OMS) independent noise components are expected!

$$
\begin{pmatrix} A & 0 \\ 0 & P \end{pmatrix} \longrightarrow \begin{pmatrix} \begin{matrix} A_{12} & 0 & 0 & D_{12}A_{21} & 0 & 0 \\ 0 & A_{23} & 0 & 0 & D_{23}A_{32} & 0 \\ 0 & 0 & A_{31} & 0 & 0 & D_{31}A_{13} \\ D_{12}A_{12} & 0 & 0 & A_{21} & 0 & 0 \\ 0 & D_{23}A_{23} & 0 & 0 & A_{32} & 0 \\ 0 & 0 & D_{31}A_{31} & 0 & 0 & A_{13} \end{matrix} & 0 \\ 0 & \begin{matrix} P_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & P_{23} & 0 & 0 & 0 & 0 \\ 0 & 0 & P_{31} & 0 & 0 & 0 \\ 0 & 0 & 0 & P_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & P_{32} & 0 \\ 0 & 0 & 0 & 0 & 0 & P_{13} \end{matrix} \end{pmatrix}
$$

Several complications are added in the problem:

- Noise components propagate differently in different TDI variables
- Higher dimensionality of the parameter space
- Correlations between the noise parameters

Again, this requires care!

Data generation (in frequency domain)

Assuming signal and noise to be Gaussian, stationary, and isotropic

independent realizations for each data segment and frequency are drawn:

$$\tilde{s}_c(f_i) = \frac{\mathcal{N}(0, \sqrt{\Omega_{\rm GW}(f_i)}) + i\, \mathcal{N}(0, \sqrt{\Omega_{\rm GW}(f_i)})}{\sqrt{2}}$$

$$\tilde{n}_c(f_i) = \frac{\mathcal{N}(0, \sqrt{\Omega_n(f_i)}) + i\, \mathcal{N}(0, \sqrt{\Omega_n(f_i)})}{\sqrt{2}}$$

so that the spectra ($\Omega_{\rm GW}$ and Ω_n) quantify the variance of fluctuations

Data generation (in frequency domain)

Assuming signal and noise to be Gaussian, stationary, and isotropic

independent realizations for each data segment and frequency are drawn:

$$\tilde{s}_c(f_i) = \frac{\mathcal{N}(0, \sqrt{\Omega_{\rm GW}(f_i)}) + i\,\mathcal{N}(0, \sqrt{\Omega_{\rm GW}(f_i)})}{\sqrt{2}}$$

$$\tilde{n}_c(f_i) = \frac{\mathcal{N}(0, \sqrt{\Omega_n(f_i)}) + i\,\mathcal{N}(0, \sqrt{\Omega_n(f_i)})}{\sqrt{2}}$$

so that the spectra ($\Omega_{\rm GW}$ and Ω_n) quantify the variance of fluctuations

Given that:

- LISA will be operating for 4yrs (and assume also 75% efficiency)
- We choose data segments of roughly 12 days

in practice we have:

- Roughly 95 independent measurements at each frequency.
- A frequency resolution of around 10^{-6}Hz

This implies: $\sim 5 \times 10^5 \times 95$ points in total (per channel), which is a lot!

Data pre-processing

Is it possible to get something similar but faster??

Data pre-processing

Is it possible to get something similar but faster??

Let us start by defining $D_c(f_i)$ (our new data), as:

$$D_c(f_i) \equiv \langle \tilde{d}_c^2(f_i) \rangle = \langle (\tilde{s}_c(f_i) + \tilde{n}_c(f_i))^2 \rangle = \langle \tilde{s}_c^2(f_i) \rangle + \langle \tilde{n}_c^2(f_i) \rangle \ .$$

We can reduce the complexity of the problem by performing two operations:

Data pre-processing

Is it possible to get something similar but faster??

Let us start by defining $D_c(f_i)$ (our new data), as:

$$D_c(f_i) \equiv \langle \tilde{d}_c^2(f_i) \rangle = \langle (\tilde{s}_c(f_i) + \tilde{n}_c(f_i))^2 \rangle = \langle \tilde{s}_c^2(f_i) \rangle + \langle \tilde{n}_c^2(f_i) \rangle \ .$$

We can reduce the complexity of the problem by performing two operations:

- We average over the (95) data segments:
  This leaves us with some $D(f_i)$ (the averaged data) and
  an estimate of the error $\sigma(f_i)$ (the standard deviation or the data).

Data pre-processing

Is it possible to get something similar but faster??

Let us start by defining $D_c(f_i)$ (our new data), as:

$$D_c(f_i) \equiv \langle \tilde{d}_c^2(f_i)\rangle = \langle(\tilde{s}_c(f_i) + \tilde{n}_c(f_i))^2\rangle = \langle \tilde{s}_c^2(f_i)\rangle + \langle \tilde{n}_c^2(f_i)\rangle \ .$$

We can reduce the complexity of the problem by performing two operations:

- We average over the (95) data segments:
  This leaves us with some $D(f_i)$ (the averaged data) and
  an estimate of the error $\sigma(f_i)$ (the standard deviation or the data).
- We coarse grain (i.e. bin) the data (in frequency):
  i.e. from the initial linear 10^{-6}Hz spacing ($\sim 5 \times 10^5$ points)
  $\longrightarrow$ we go to some final (and less dense) set of frequencies f_i
  This leaves us with the final dataset f_i, D_i, with errors σ_i and weights w_i.

Data pre-processing

Is it possible to get something similar but faster??

Let us start by defining $D_c(f_i)$ (our new data), as:

$$D_c(f_i) \equiv \langle \tilde{d}_c^2(f_i) \rangle = \langle (\tilde{s}_c(f_i) + \tilde{n}_c(f_i))^2 \rangle = \langle \tilde{s}_c^2(f_i) \rangle + \langle \tilde{n}_c^2(f_i) \rangle \ .$$

We can reduce the complexity of the problem by performing two operations:

- We average over the (95) data segments:
  This leaves us with some $D(f_i)$ (the averaged data) and
  an estimate of the error $\sigma(f_i)$ (the standard deviation or the data).
- We coarse grain (i.e. bin) the data (in frequency):
  i.e. from the initial linear 10^{-6}Hz spacing ($\sim 5 \times 10^5$ points)
  $\longrightarrow$ we go to some final (and less dense) set of frequencies f_i
  This leaves us with the final dataset f_i, D_i, with errors σ_i and weights w_i.

To gain a factor $\sim \mathcal{O}(100 \times 100)$ in computation time!

Data modeling and likelihood

To perform the analysis (in a Bayesian framework) we need to specify

$$p(\theta|\tilde{d}) = \mathcal{L}(\tilde{d}|\theta)\pi(\theta)/\pi(\tilde{d}) \,,$$

meaning, we have to choose a likelihood (and the priors).

Data modeling and likelihood

To perform the analysis (in a Bayesian framework) we need to specify

$$p(\theta|\tilde{d}) = \mathcal{L}(\tilde{d}|\theta)\pi(\theta)/\pi(\tilde{d}) \, ,$$

meaning, we have to choose a likelihood (and the priors).

Whittle likelihood should work (see any of the existing reviews):

$$-2 \ln \mathcal{L}(\tilde{d}|\theta) \propto \sum_f \tilde{d}_i C_{ij}^{-1}(f|\theta)\tilde{d}_j^* + \ln\left[\det C_{ij}(f|\theta)\right] \, ,$$

where C_{ij} is the covariance matrix for $\langle \tilde{d}_i \tilde{d}_j^* \rangle$ given by:

$$C_{ij}(f,\theta) = S_{ij}\left(f, \vec{\theta}_s\right) + N_{ij}\left(f, \vec{\theta}_n\right) \, .$$

Data modeling and likelihood

To perform the analysis (in a Bayesian framework) we need to specify

$$p(\theta|\tilde{d}) = \mathcal{L}(\tilde{d}|\theta)\pi(\theta)/\pi(\tilde{d}) \,,$$

meaning, we have to choose a likelihood (and the priors).

Whittle likelihood should work (see any of the existing reviews):

$$-2\ln\mathcal{L}(\tilde{d}|\theta) \propto \sum_f \tilde{d}_i C_{ij}^{-1}(f|\theta)\tilde{d}_j^* + \ln\left[\det C_{ij}(f|\theta)\right] \,,$$

where C_{ij} is the covariance matrix for $\langle \tilde{d}_i \tilde{d}_j^* \rangle$ given by:

$$C_{ij}(f,\theta) = S_{ij}\left(f,\,\vec{\theta}_s\right) + N_{ij}\left(f,\vec{\theta}_n\right) \,.$$

For fast estimates of the uncertainties on parameter reconstruction we can use the Fisher Information Matrix (FIM):

$$F_{ij} \equiv -\left.\frac{\partial^2 \ln\mathcal{L}(\tilde{d}|\theta)}{\partial\theta_i\partial\theta_j}\right|_{\theta_0} = \sum_f \mathrm{Tr}\left[C^{-1}\frac{\partial C}{\partial\theta_i}C^{-1}\frac{\partial C}{\partial\theta_j}\right] \,,$$

NB $\longrightarrow$

- FIM only works if the posterior is Gaussian
- FIM won't tell us anything about possible biases

An accurate likelihood for the compressed data

What about the compressed data??

A Gaussian likelihood would give a systematic low bias!

(astro-ph/9808264, astro-ph/0205387, astro-ph/0302218, 0801.0554)

Consider the Gaussian likelihood:

$$\ln \mathcal{L}_G\left(\vec{\theta}, \vec{n}\right) \propto -\frac{N_{chunks}}{2} \sum_{i,j} \sum_k w_{ij}^{(k)} \left(\frac{D_{ij}^{(k)} - h^2\Omega_{\rm GW}\left(f_{ij}^{(k)}, \vec{\theta}\right) - h^2\Omega_{n,ij}\left(f_{ij}^{(k)}, \vec{n}\right)}{h^2\Omega_{\rm GW}\left(f_{ij}^{(k)}, \vec{\theta}\right) + h^2\Omega_{n,ij}\left(f_{ij}^{(k)}, \vec{n}\right)} \right)^2$$

and the Lognormal likelihood:

$$\ln \mathcal{L}_{LN}\left(\vec{\theta}, \vec{n}\right) \propto -\frac{N_{chunks}}{2} \sum_{i,j} \sum_k w_{ij}^{(k)} \ln^2 \left(\frac{h^2\Omega_{\rm GW}\left(f_{ij}^{(k)}, \vec{\theta}\right) + h^2\Omega_{n,ij}\left(f_{ij}^{(k)}, \vec{n}\right)}{D_{ij}^{(k)}} \right)$$

Then we define our likelihood as (astro-ph/0302218, 2009.11845)

$$\ln \mathcal{L} = \frac{1}{3} \ln \mathcal{L}_G + \frac{2}{3} \ln \mathcal{L}_{LN}$$

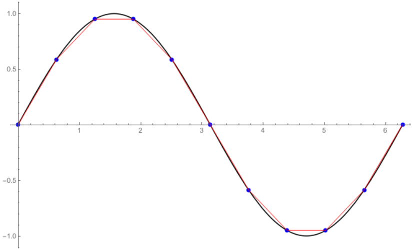
which includes the skewness contributions and thus is more accurate.

SGWBinner algorithm

We look for best approximation of the signal with a multi-PL

$$h^2\Omega_{\rm GW}\left(f,\,\vec{\theta}\right) = \sum_i 10^{\alpha_i} \left(\frac{f}{\sqrt{f_{\min,i}\, f_{\max,i}}}\right)^{p_i} \Theta\left(f - f_{\min,i}\right) \Theta\left(f_{\max,i} - f\right) .$$

where Θ is the Heaviside step function.



A sketch of the algorithm:

1. Build a robust noise prior
2. Bin the frequency range and reconstruct the signal
3. Merge as many bins as possible (to avoid overfitting)
4. Define a procedure to compute the error on the reconstruction
5. Final MCMC run with common noise parameters

Criterion for merging

Two motivations for merging:

- Larger N means smaller bins which implies larger errors
- For large values of N, the reconstruction with N bins (*i.e.* $2N$ parameters) may overfit the signal

Typically reducing the number N of bins may improve the analysis!

Criterion for merging

Two motivations for merging:

- Larger N means smaller bins which implies larger errors
- For large values of N, the reconstruction with N bins ($i.e.$ $2N$ parameters) may overfit the signal

Typically reducing the number N of bins may improve the analysis!

For a couple of consecutive bins $(i, i+1)$ we can compute

$$\begin{aligned}\Delta\,\mathrm{AIC} &= \mathrm{AIC}_{\mathrm{after\ \ merging}} - \mathrm{AIC}_{\mathrm{before\ \ merging}} \\ &= \chi^2_{\mathrm{after\ \ merging}} - \chi^2_{\mathrm{before\ \ merging}} - 2\,k_{\mathrm{1\text{-}bin}}\end{aligned}$$

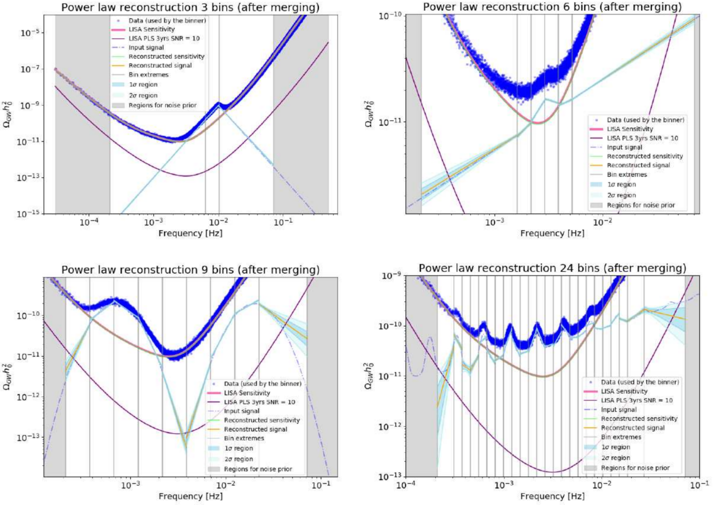
where the Akaike Information Criterion (AIC) is $\mathrm{AIC} \equiv \chi^2 + 2k$.

In practice:

$\Delta\mathrm{AIC} < 0 \quad \longrightarrow \quad$ It is convenient to merge the two bins

Frequency shape reconstruction

Some examples of reconstructions with the SGWBinner code



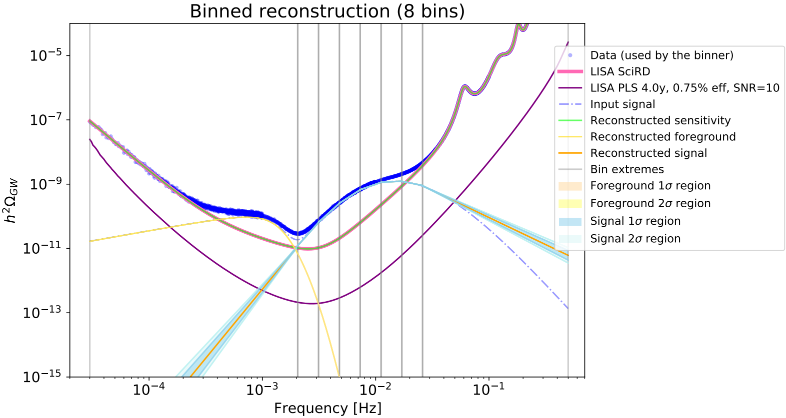
C. Caprini et al. JCAP 11 (2019) 017, ArXiv: 1906.09244.

Joint treatment of SGWB + noise + foreground

The code can work with several components in the data!

Consider for example the signal due to Galactic binaries:

$$h^2\Omega_{\rm GW} = 10^{\alpha_{FG}} f^{2/3} {\rm e}^{-a_1 f + a_2 f \sin(a_3 f)} \left\{1 + \tanh\left[a_4(f_k - f)\right]\right\} \ .$$



R. Flauger et al. JCAP 01 (2021) 059, ArXiv:2009.11845.

Forecasting LISA constraints I (FIM)

Choose a template

↓

Get forecasts (e.g., using FIM) on the template parameters

↓

Convert in constraints on model parameter

↓

Forecast constrains on fundamental physics!

Forecasting LISA constraints I (FIM)

Choose a template

↓

Get forecasts (e.g., using FIM) on the template parameters

↓

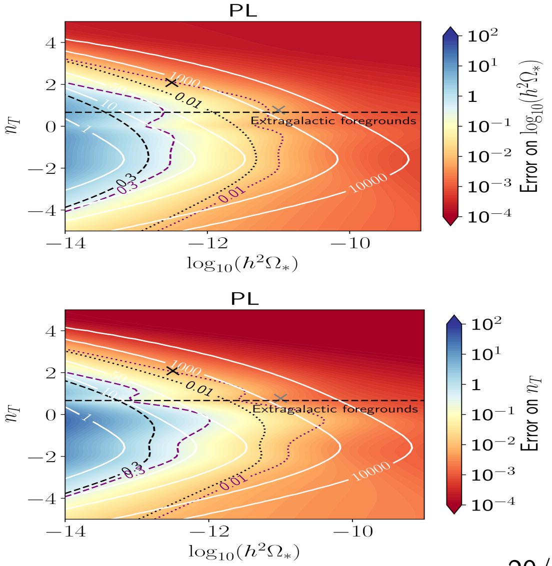
Convert in constraints on model parameter

↓

Forecast constrains on fundamental physics!

Example: a power-law

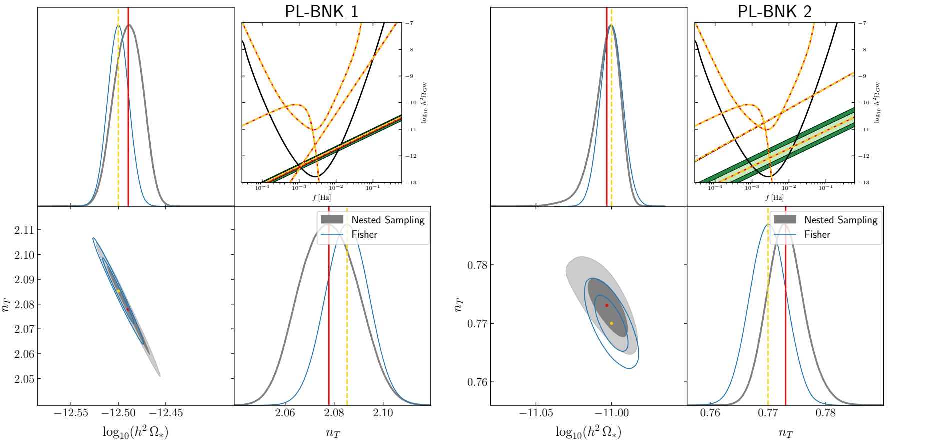
$$\Omega_{\rm GW} h^2 = 10^{\log_{10}(h^2\Omega_*)} \left(\frac{f}{f_*}\right)^{n_T}$$



LISA Cosmology Working Group, JCAP 11 (2024) 032, ArXiv:2405.03740.

Forecasting LISA constraints II (Full analysis)

Validate FIM results generating data and running some data analysis pipeline



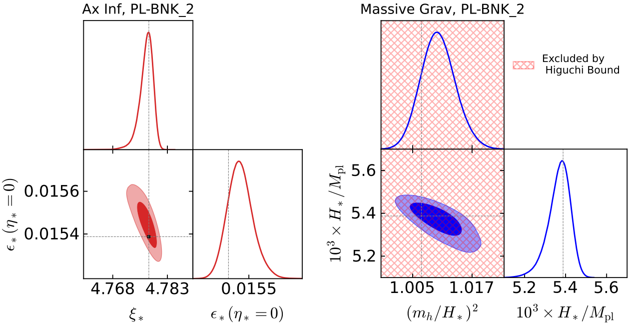
LISA Cosmology Working Group, JCAP 11 (2024) 032, ArXiv:2405.03740.

Forecasting LISA constraints III (model parameters)

Again PL
(from inflation)

LISA Cosmology Working Group, JCAP 11 (2024) 032, ArXiv:2405.03740.

→

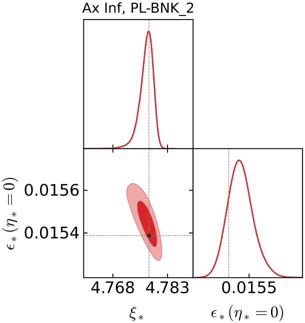


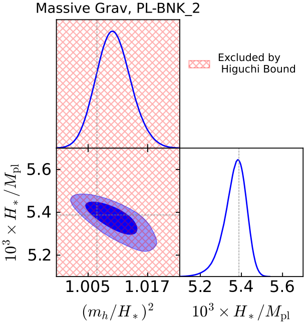
Forecasting LISA constraints III (model parameters)

Again PL
(from inflation)

LISA Cosmology Working Group, JCAP 11 (2024) 032, ArXiv:2405.03740.

→

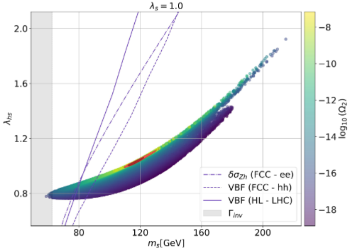


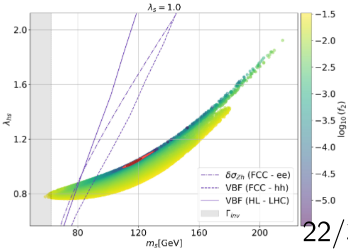


Or a BPL
(from FOPTs)

LISA Cosmology Working Group, JCAP 10 (2024) 020, ArXiv:2403.03723.

→





Also more complex noise ...

Let's consider a more complex noise model

- 6 TM parameters
- 6 OMS parameters

Does that impact the signal reconstruction?

Also more complex noise ...

Let's consider a more complex noise model

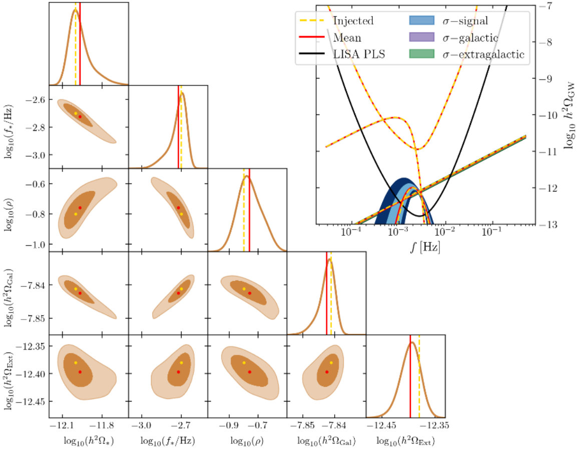
- 6 TM parameters
- 6 OMS parameters

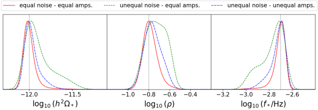
Does that impact the signal reconstruction?

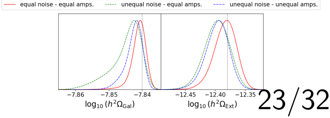
Tested for:

- Power Law
  Hartwig et al. Phys.Rev.D 107 (2023), 123531, ArXiv: 2303.15929.
- Lognormal bump
  Kume et al. ArXiv: 2410.10342.

Seems like not much!

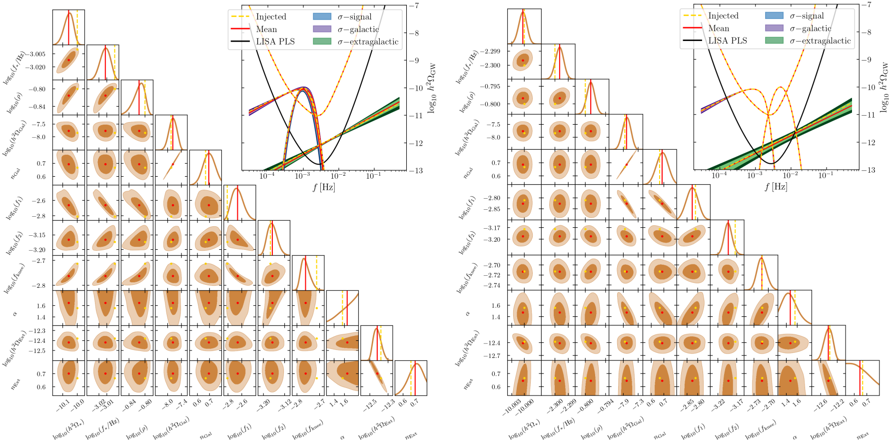






… and foreground models

What about relaxing the assumptions on the foreground?



Depends on how degenerate the components are…

Kume et al. ArXiv: 2410.10342.

SBI for GWB data analysis

Traditional methods (MCMC, nested sampling, whatever) are quite efficient and guaranteed to converge (in some cases)

but

scale poorly with number of parameters and require explicit likelihoods

Can alternative approaches perform better in some cases?

SBI for GWB data analysis

Traditional methods (MCMC, nested sampling, whatever) are quite efficient and guaranteed to converge (in some cases)

but

scale poorly with number of parameters and require explicit likelihoods

Can alternative approaches perform better in some cases?

Normally, with Bayesian inference, we try to study the posterior probability:

$$p(\theta|d) = \frac{p(d|\theta)\,\pi(\theta)}{p(d)} \equiv r(d,\theta)\,\pi(\theta)\,,$$

where we have introduced:

$$r(d,\theta) \equiv \frac{p(d|\theta)}{p(d)} = \frac{p(\theta|d)}{\pi(\theta)} = \frac{p(\theta,d)}{p(d)\,\pi(\theta)}\,,$$

i.e., $r(d,\theta)$ is the ratio between joint probability and marginal probability.

Given a pair (θ, d), $r(d,\theta)$ can be used to assess whether θ can generate d!

SBI for GWB data analysis

Traditional methods (MCMC, nested sampling, whatever) are quite efficient and guaranteed to converge (in some cases)

but

scale poorly with number of parameters and require explicit likelihoods

Can alternative approaches perform better in some cases?

Normally, with Bayesian inference, we try to study the posterior probability:

$$p(\theta|d) = \frac{p(d|\theta)\,\pi(\theta)}{p(d)} \equiv r(d,\theta)\,\pi(\theta)\,,$$

where we have introduced:

$$r(d,\theta) \equiv \frac{p(d|\theta)}{p(d)} = \frac{p(\theta|d)}{\pi(\theta)} = \frac{p(\theta,d)}{p(d)\,\pi(\theta)}\,,$$

i.e., $r(d,\theta)$ is the ratio between joint probability and marginal probability.

Given a pair (θ, d), $r(d,\theta)$ can be used to assess whether θ can generate d!

This can be cast in a minimization problem that can be solved with ML
the approach is typically referred to as Neural Ratio Estimation (NRE)
(basically build a classifier to say whether θ, d are joint or marginal...).

A very simple example

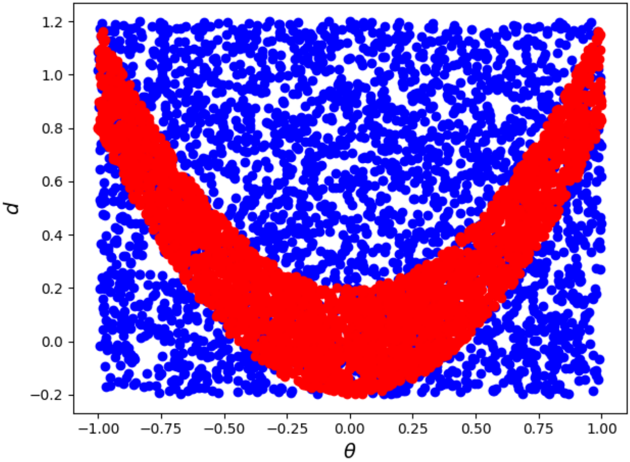
Let's consider:

$$d = \theta^2 + \epsilon$$

with
$\theta \sim U[-1, 1]$
$\epsilon \sim U[-.2, .2]$

Marginal samples
$d, \theta \sim p(d)p(\theta)$

Joint samples
$d, \theta \sim p(d, \theta)$



A very simple example

Let's consider:

$$d = \theta^2 + \epsilon$$

with

$\theta \sim U[-1, 1]$

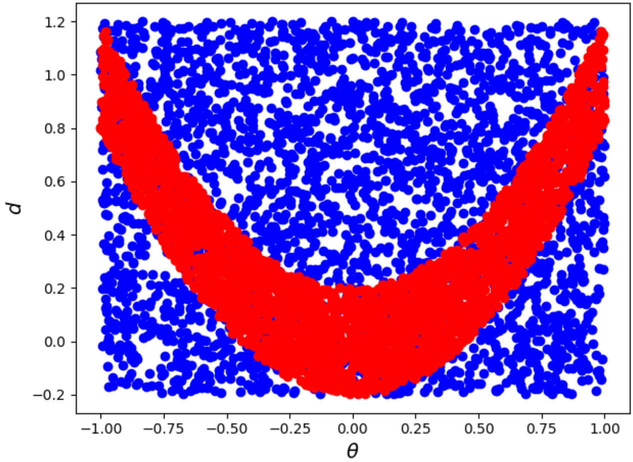
$\epsilon \sim U[-.2, .2]$

Marginal samples

$d, \theta \sim p(d)p(\theta)$

Joint samples

$d, \theta \sim p(d, \theta)$



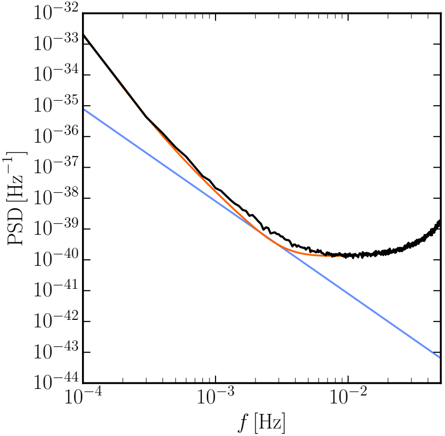
In very large parameter spaces you can target a subset of the parameters!

↓

Automatically marginalize over the other parameters (here ϵ)!

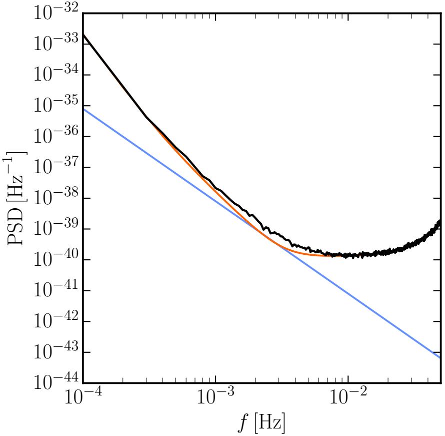
Recover previous results ...

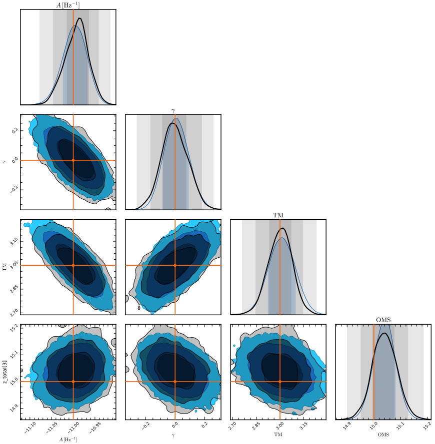
Assume we inject a power law signal:
Can we recover it with the same level of accuracy?



Recover previous results ...

Assume we inject a power law signal:
Can we recover it with the same level of accuracy?





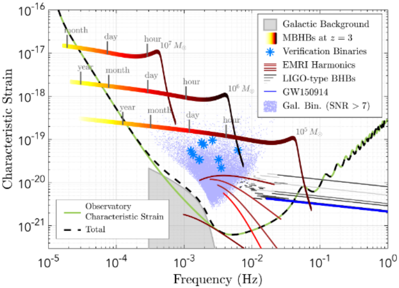
Good news!

James Alvey et al., Phys. Rev. D 109 (2024) 083008, ArXiv:2309.07954.

Code available at `https://github.com/PEREGRINE-GW/saqqara/`

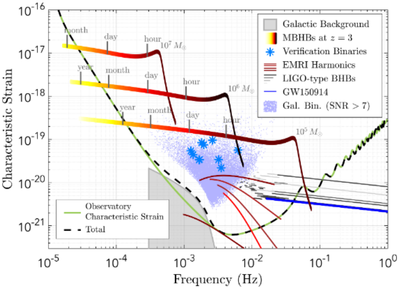
... plus something completely new!

What if there's something else beyond GWB and noise?



… plus something completely new!

What if there's something else beyond GWB and noise?

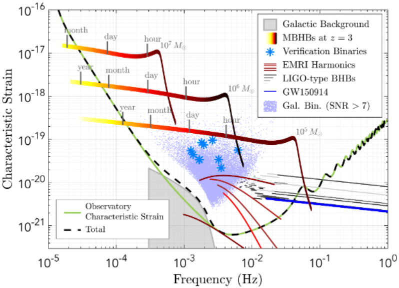


A minimal setup for this: randomly inject transients slightly below the threshold for detection

Would this still work??

... plus something completely new!

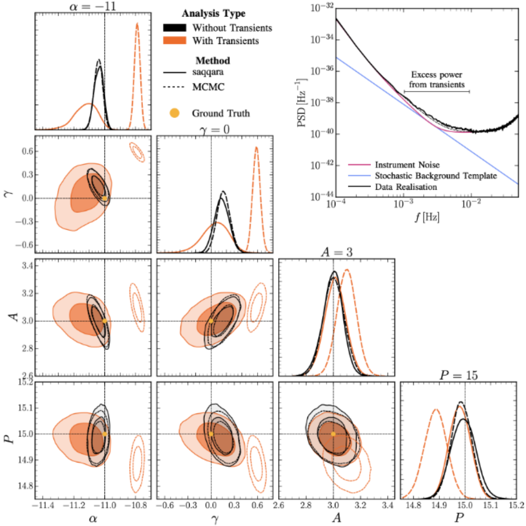
What if there's something else beyond GWB and noise?



A minimal setup for this: randomly inject transients slightly below the threshold for detection

Would this still work??

Yes!!



James Alvey et al., Phys. Rev. D 109 (2024) 083008, ArXiv:2309.07954.
Code available at `https://github.com/PEREGRINE-GW/saqqara/`

What about noise non-stationarities? (FIM)

The noise won't be stationary for the whole mission duration ...

How does this impact the signal parameters reconstruction?

What about noise non-stationarities? (FIM)

The noise won't be stationary for the whole mission duration ...

How does this impact the signal parameters reconstruction?

A strategy to answer this question:

1. Cut the data into segments (where stationarity holds)
2. Analyze segment-by-segment (FIM or training a network)
3. Combine the results (assuming each segment is an independent draw)

What about noise non-stationarities? (FIM)

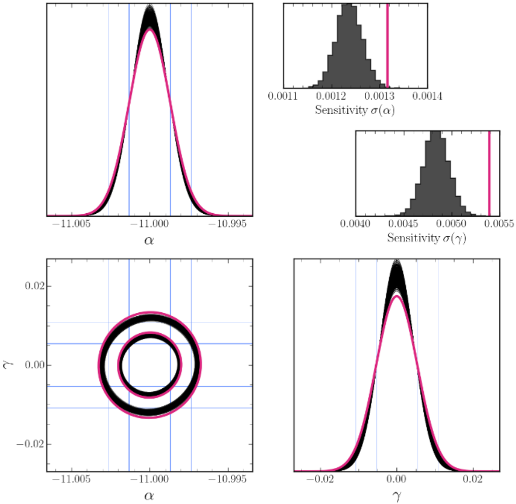
The noise won't be stationary for the whole mission duration ...

How does this impact the signal parameters reconstruction?

A strategy to answer this question:

1. Cut the data into segments (where stationarity holds)
2. Analyze segment-by-segment (FIM or training a network)
3. Combine the results (assuming each segment is an independent draw)

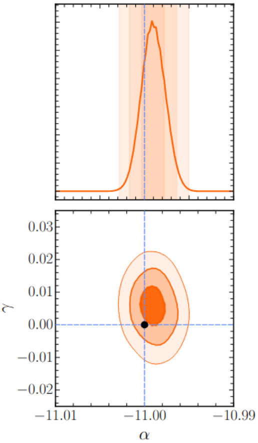
First test with FIM:
Looks like you actually do better!

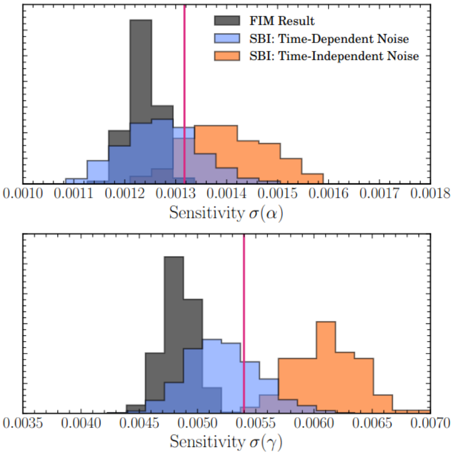


James Alvey et al., ArXiv:2408.00832.
Code available at `https://github.com/PEREGRINE-GW/saqqara/`
See also `https://github.com/Mauropieroni/GW_response`

What about noise non-stationarities? (Full analysis)

Validate FIM results with a full analysis





Moreover, training segment-by-segment is 100 times faster!

James Alvey et al., ArXiv:2408.00832.
Code available at `https://github.com/PEREGRINE-GW/saqqara/`
See also `https://github.com/Mauropieroni/GW_response`

Conclusions and outlook

Conclusions

- Analyzing LISA data will be quite difficult
- Agnostic (or at least not very signal dependent) methods are useful
- Traditional methods work (but they might be expensive)
- SBI (and ML in general) looks quite promising ...
- ... it works with transients and non-stationarities in the noise!

Future perspectives

- Keep improving on detector modeling
- More realistic noise model and data generation procedure
- Include anisotropies in the signal
- New ideas/techniques?
- ...

Last slide

Thank you for your attention!