arXiv:2302.12573

DIFFERENTIATING SIGNAL FROM NOISE: TOWARDS A MULTIVARIATE SPECTRAL ANALYSIS FOR LISA

Quentin Baghi*, Nikos Karnesis, Jean-Baptiste Bayle, Marc Besançon, Henri Inchauspe
*CEA Paris-Saclay

Thursday, July 20th 2023
Data analysis challenges for stochastic gravitational wave backgrounds

1. Problem statement
2. Multivariate time series model
3. Bayesian inference
4. Revisiting null channels

Stochastic GW background generated with Midjourney

- In LISA, **we cannot cross-correlate different detectors** like LIGO-Virgo-KAGRA
  → Instrumental stochastic processes must be very precisely accounted for to allow for a detection

Problem statement

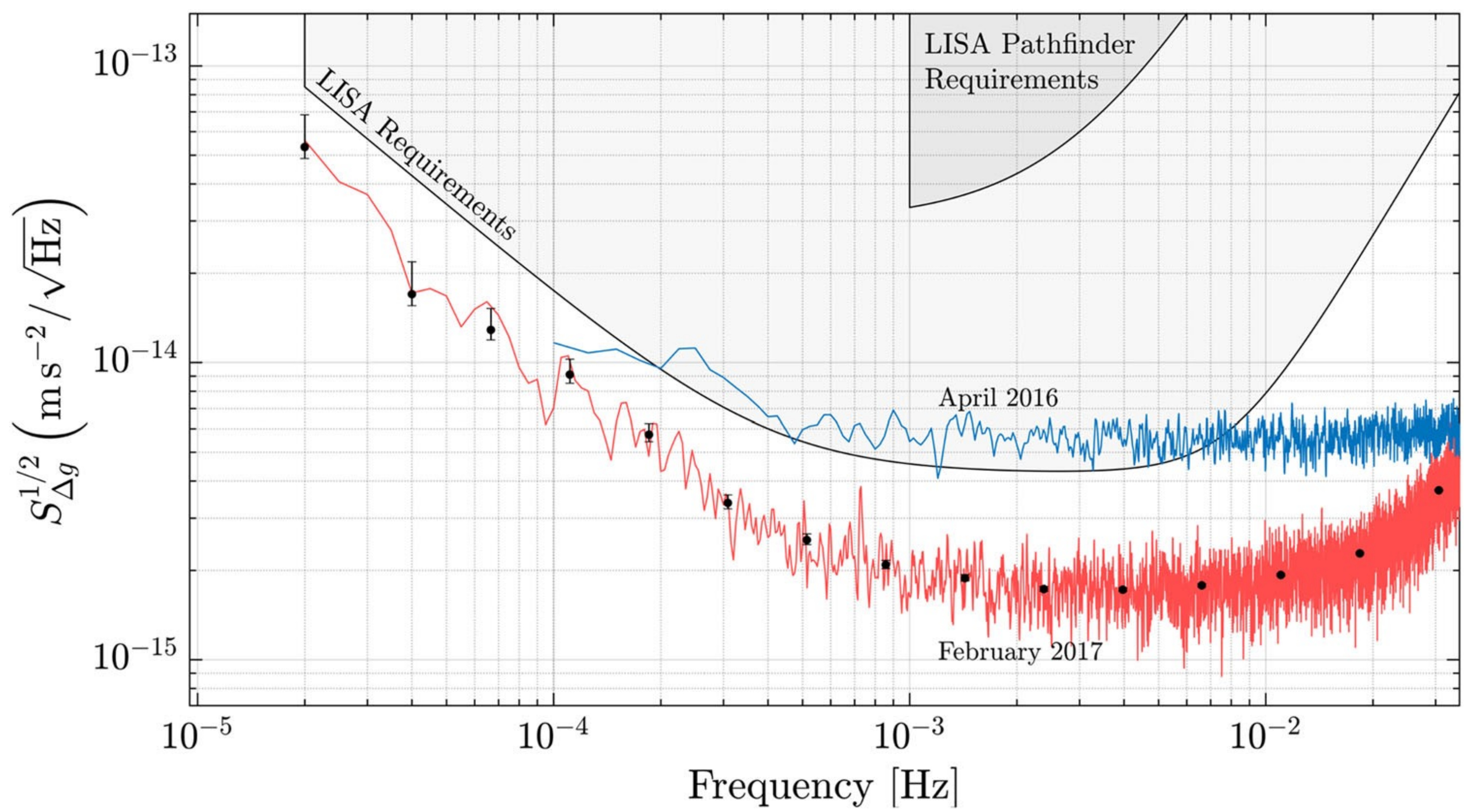
- In LISA, **we cannot cross-correlate different detectors** like LIGO-Virgo-KAGRA
  → Instrumental stochastic processes must be very precisely accounted for to allow for a detection

- **We cannot fully rely on pre-flight instrumental noise models** (Cf. LISA Pathfinder)

- In LISA, **we cannot cross-correlate different detectors** like LIGO-Virgo-KAGRA
  → Instrumental stochastic processes must be very precisely accounted for to allow for a detection

- **We cannot fully rely on pre-flight instrumental noise models** (Cf. LISA Pathfinder)

Overestimated noise level (but great success!)



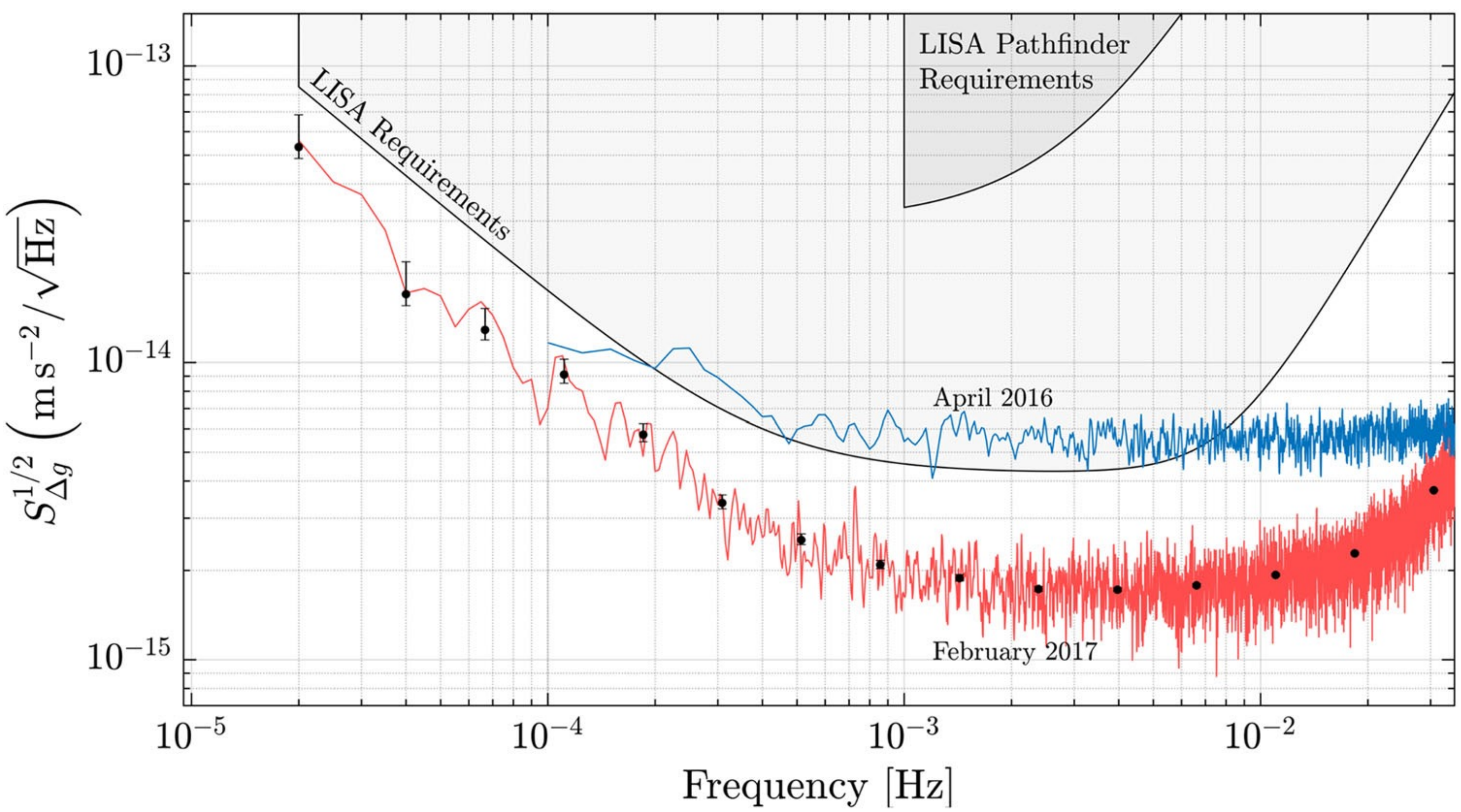
[Armano et al., PRL, 2018]

Problem statement

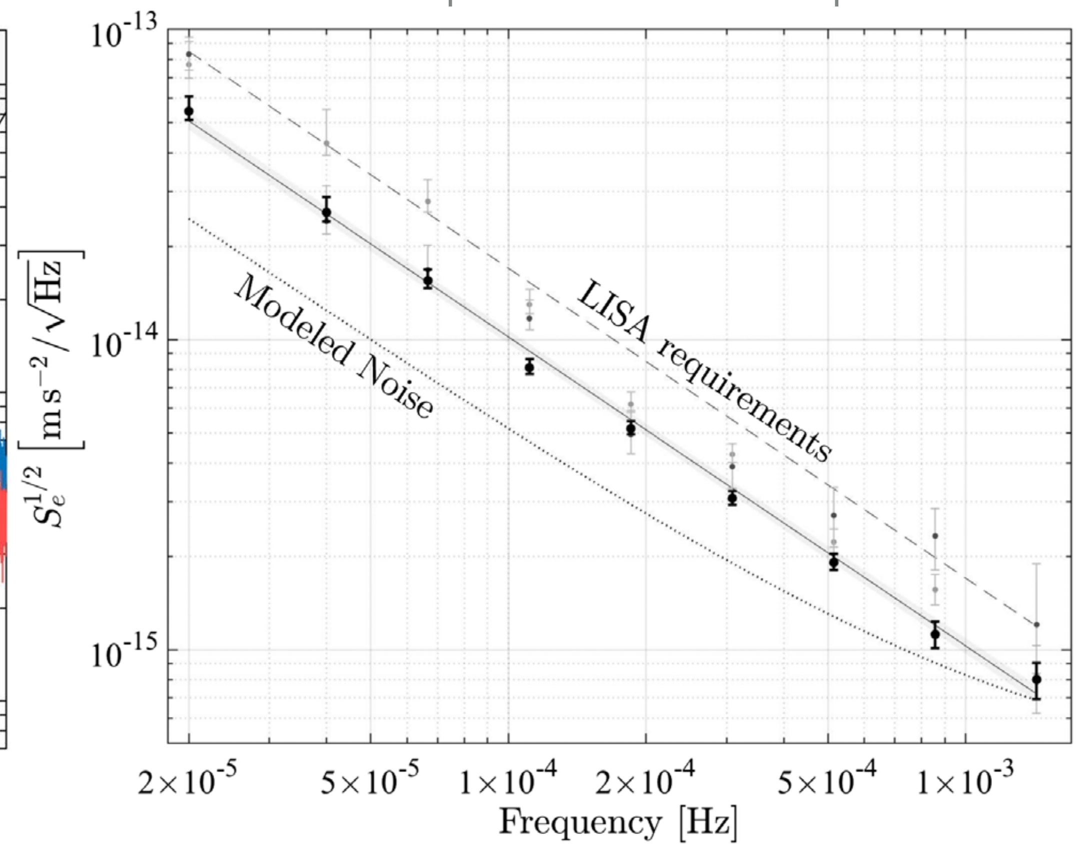
- In LISA, **we cannot cross-correlate different detectors** like LIGO-Virgo-KAGRA
  → Instrumental stochastic processes must be very precisely accounted for to allow for a detection

- **We cannot fully rely on pre-flight instrumental noise models** (Cf. LISA Pathfinder)

Overestimated noise level (but great success!)



Excess of power at low frequencies



[Armano et al., PRL, 2018]

- **There is no such thing as a perfect null channel** (Cf. Martina Muratore and Mauro Pieroni's talks)

- **There is no such thing as a perfect null channel** (Cf. Martina Muratore and Mauro Pieroni's talks)
- We need
  - to use **differences in signal and noise features**
  - a noise model that is as robust and flexible as possible
  - realistic instrumental data simulations

Problem statement

- **There is no such thing as a perfect null channel** (Cf. Martina Muratore and Mauro Pieroni's talks)
- We need
  - to use **differences in signal and noise features**
  - a noise model that is as robust and flexible as possible
  - realistic instrumental data simulations
- Nevertheless, we need to make **assumptions**

- **There is no such thing as a perfect null channel** (Cf. Martina Muratore and Mauro Pieroni's talks)

- We need

  - to use **differences in signal and noise features**
  - a noise model that is as robust and flexible as possible
  - realistic instrumental data simulations

- Nevertheless, we need to make **assumptions**

    - Isotropic, stationary SGWB

- **There is no such thing as a perfect null channel** (Cf. Martina Muratore and Mauro Pieroni's talks)
- We need
  - to use **differences in signal and noise features**
  - a noise model that is as robust and flexible as possible
  - realistic instrumental data simulations
- Nevertheless, we need to make **assumptions**
  - Isotropic, stationary SGWB
  - Noise PSD is smooth on mHz scales

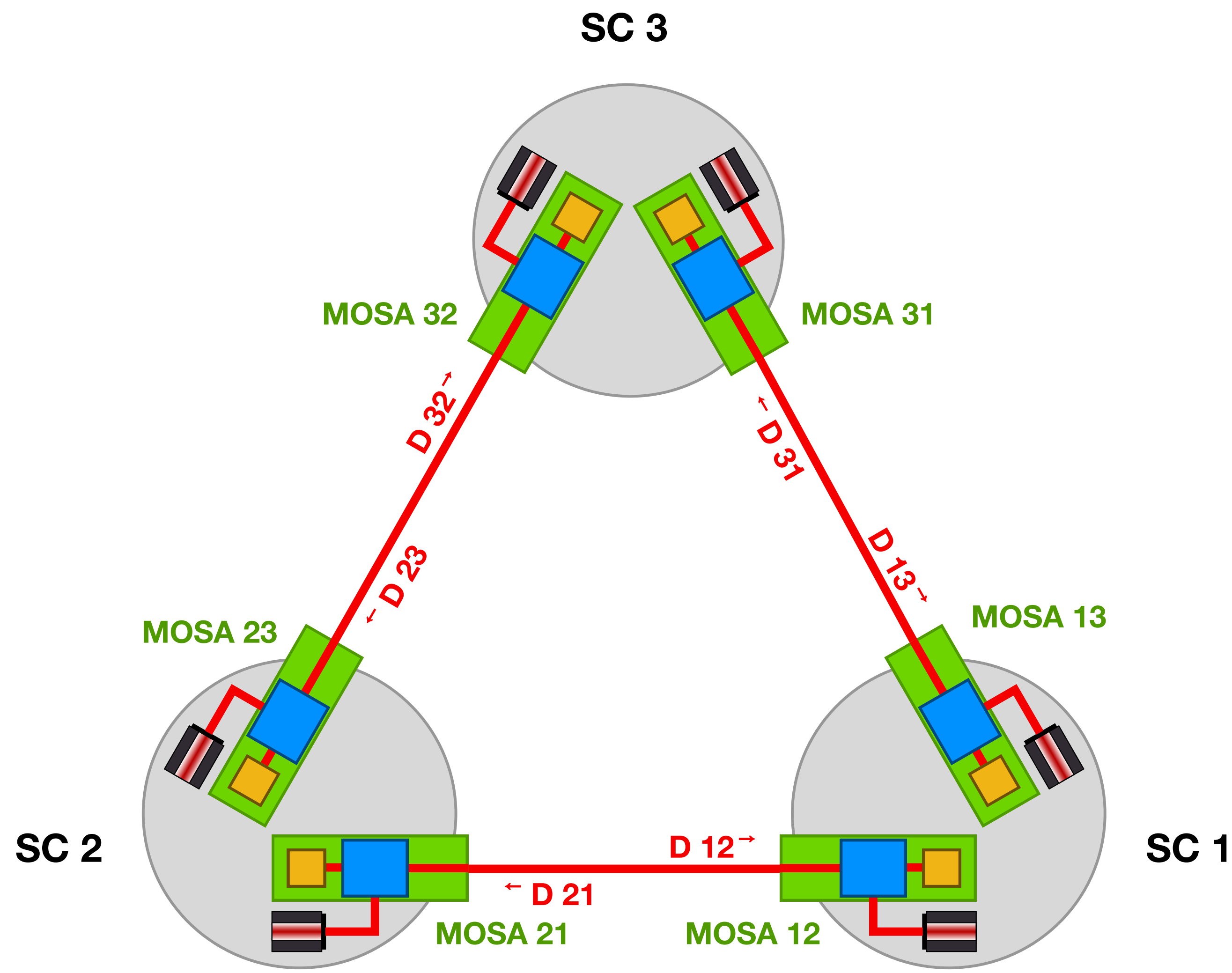
- **There is no such thing as a perfect null channel** (Cf. Martina Muratore and Mauro Pieroni's talks)

- We need
  - to use **differences in signal and noise features**
  - a noise model that is as robust and flexible as possible
  - realistic instrumental data simulations

- Nevertheless, we need to make **assumptions**
  - Isotropic, stationary SGWB
  - Noise PSD is smooth on mHz scales
  - All resolvable GW sources have been removed (!!!!)

- We start from interferometric measurements
- After some combinations to suppress S/C motion and half of the laser noise, we obtain 6 intermediate variables η

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_{12} \\ \eta_{23} \\ \eta_{31} \\ \eta_{13} \\ \eta_{32} \\ \eta_{21} \end{pmatrix}$$



- They are affected by noise and GW signal:

$$\boldsymbol{\eta} = \mathbf{h} + \mathbf{p} + \mathbf{n}$$

GW signal

laser frequency noise

secondary noises

- To cancel laser frequency noise we form TDI variables through a linear operation:

$$\mathbf{d} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \boxed{\mathbf{M}}\boldsymbol{\eta} \quad \rightarrow \text{reduces to 3 effective variables}$$

$$= \mathbf{Mh} + \mathbf{Mn} + \cancel{\mathbf{M}}\mathbf{p}$$

$$= \mathbf{Mh} + \mathbf{Mn}$$

TDI transformation matrix

- To cancel laser frequency noise we form TDI variables through a linear operation:

$$\mathbf{d} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \boxed{\mathbf{M}}\boldsymbol{\eta} \quad \rightarrow \text{reduces to 3 effective variables}$$
$$= \mathbf{Mh} + \mathbf{Mn} + \cancel{\mathbf{M}}\mathbf{p}$$
$$= \mathbf{Mh} + \mathbf{Mn}$$

TDI transformation matrix

- We obtain a 3-dimensional multivariate time series $\mathbf{d}$
- We can take their discrete Fourier transform (DFT): $\tilde{\boldsymbol{d}} = \mathrm{DFT}\left(\boldsymbol{d}\right)$
- Stationarity assumption implies:
  - ✦ DFT components at frequencies f_k approximately uncorrelated
  - ✦ Each frequency bin is characterised by a **spectrum matrix**

$$\tilde{\Sigma}_d(f) \equiv \mathrm{E}\left[\tilde{\mathbf{d}}(f)\tilde{\mathbf{d}}(f)^{\dagger}\right]$$

Multivariate time series model

- We can split the spectrum matrix into 2 independent components: signal and noise

$$\tilde{\mathbf{\Sigma}}_d = \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} + \tilde{\mathbf{\Sigma}}_{\mathrm{GW}}$$

Multivariate time series model

- We can split the spectrum matrix into 2 independent components: signal and noise

$$\tilde{\mathbf{\Sigma}}_d = \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} + \tilde{\mathbf{\Sigma}}_{\mathrm{GW}}$$

- Note that each matrix depends on frequency
- The noise covariance may have a complicated structure
- The signal covariance is easier to model. For a stationary isotropic SGWB, we have

Multivariate time series model

- We can split the spectrum matrix into 2 independent components: signal and noise

$$\tilde{\mathbf{\Sigma}}_d = \tilde{\mathbf{\Sigma}}_{\text{noise}} + \tilde{\mathbf{\Sigma}}_{\text{GW}}$$

- Note that each matrix depends on frequency
- The noise covariance may have a complicated structure
- The signal covariance is easier to model. For a stationary isotropic SGWB, we have

$$\tilde{\mathbf{\Sigma}}_{\text{GW}}(f) = \tilde{\mathbf{R}}_{\text{GW}}(f) S_h(f)$$

- We can split the spectrum matrix into 2 independent components: signal and noise

$$\tilde{\mathbf{\Sigma}}_d = \tilde{\mathbf{\Sigma}}_{\text{noise}} + \tilde{\mathbf{\Sigma}}_{\text{GW}}$$

- Note that each matrix depends on frequency
- The noise covariance may have a complicated structure
- The signal covariance is easier to model. For a stationary isotropic SGWB, we have

$$\tilde{\mathbf{\Sigma}}_{\text{GW}}(f) = \tilde{\mathbf{R}}_{\text{GW}}(f) S_h(f)$$

Sky-averaged TDI response matrix

SGWB PSD (scalar)

Multivariate time series model

- We can split the spectrum matrix into 2 independent components: signal and noise

$$\tilde{\mathbf{\Sigma}}_d = \tilde{\mathbf{\Sigma}}_{\text{noise}} + \tilde{\mathbf{\Sigma}}_{\text{GW}}$$

- Note that each matrix depends on frequency
- The noise covariance may have a complicated structure
- The signal covariance is easier to model. For a stationary isotropic SGWB, we have

$$\underbrace{\tilde{\mathbf{\Sigma}}_{\text{GW}}(f) = \tilde{\mathbf{R}}_{\text{GW}}(f)}_{\text{Sky-averaged TDI response matrix}}\underbrace{S_h(f)}_{\text{SGWB PSD (scalar)}}$$

- What helps distinguishing between noise and signal:

   1. **Differences in correlation structure** between $\tilde{\mathbf{\Sigma}}_{\text{noise}}$ and $\tilde{\mathbf{\Sigma}}_{\text{GW}}$
   2. **Spectral features** of the SGWB template $S_h(f)$
   3. **Priors** on noise and signal parameters

Bayesian inference

- We recently conducted a study with additional assumptions:
  - All interferometric noises have the same transfer function
  - All interferometric noises have the same PSD
  - They are uncorrelated

- We recently conducted a study with additional assumptions:
  - ✦ All interferometric noises have the same transfer function
  - ✦ All interferometric noises have the same PSD
  - ✦ They are uncorrelated

- In our approximation we can model the noise covariance as

$$\underbrace{\tilde{\mathbf{\Sigma}}_{\mathrm{noise}}(f) = \tilde{\mathbf{R}}_{\mathrm{noise}}(f)}_{\text{TDI correlation matrix}}\, \underbrace{S_n(f)}_{\text{Single-link noise PSD (scalar)}}$$

- We recently conducted a study with additional assumptions:
  - ✦ All interferometric noises have the same transfer function
  - ✦ All interferometric noises have the same PSD
  - ✦ They are uncorrelated

- In our approximation we can model the noise covariance as

$$\tilde{\mathbf{\Sigma}}_{\mathrm{noise}}(f) = \tilde{\mathbf{R}}_{\mathrm{noise}}(f) S_n(f)$$

TDI correlation matrix

Single-link noise PSD (scalar)

- We do it with a spline basis:

$$\log S_n(f) = \sum_{j=0}^{K-1} c_j B_j\left(f, \boldsymbol{\xi}\right)$$

- We recently conducted a study with additional assumptions:
  - ✦ All interferometric noises have the same transfer function
  - ✦ All interferometric noises have the same PSD
  - ✦ They are uncorrelated

- In our approximation we can model the noise covariance as

$$\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \tilde{\mathbf{R}}_{\text{noise}}(f) S_n(f)$$

TDI correlation matrix

Single-link noise PSD (scalar)

- We do it with a spline basis:

$$\log S_n(f) = \sum_{j=0}^{K-1} c_j B_j\left(f, \boldsymbol{\xi}\right)$$

Spline amplitudes

- We recently conducted a study with additional assumptions:
  - ✦ All interferometric noises have the same transfer function
  - ✦ All interferometric noises have the same PSD
  - ✦ They are uncorrelated

- In our approximation we can model the noise covariance as

$$\underbrace{\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \tilde{\mathbf{R}}_{\text{noise}}(f)}_{\text{TDI correlation matrix}}\underbrace{S_n(f)}_{\text{Single-link noise PSD (scalar)}}$$

- We do it with a spline basis:

$$\log S_n(f) = \sum_{j=0}^{K-1} \underbrace{c_j}_{\text{Spline amplitudes}} B_j \left(f, \underbrace{\boldsymbol{\xi}}_{\text{Spline locations}}\right)$$

- Power-law template to model the SGWB PSD:

$$S_h(f) = \Omega_{\mathrm{GW}}(f)\frac{3H_0^2}{4\pi^2 f^3} \qquad \Omega_{\mathrm{GW}}(f) = \Omega_{\mathrm{GW},0}\left(\frac{f}{f_0}\right)^n$$

- Power-law template to model the SGWB PSD:

$$S_h(f) = \Omega_{\mathrm{GW}}(f)\frac{3H_0^2}{4\pi^2 f^3} \qquad \Omega_{\mathrm{GW}}(f) = \underbrace{\Omega_{\mathrm{GW},0}}_{\text{Energy density}}\left(\frac{f}{f_0}\right)^n$$

- Power-law template to model the SGWB PSD:

$$S_h(f) = \Omega_{\mathrm{GW}}(f)\frac{3H_0^2}{4\pi^2 f^3} \qquad \Omega_{\mathrm{GW}}(f) = \Omega_{\mathrm{GW},0}\left(\frac{f}{f_0}\right)^n$$

Energy density   Power-law index

- Power-law template to model the SGWB PSD:

$$S_h(f) = \Omega_{\mathrm{GW}}(f) \frac{3H_0^2}{4\pi^2 f^3} \qquad \Omega_{\mathrm{GW}}(f) = \underbrace{\Omega_{\mathrm{GW},0}}_{\text{Energy density}} \left(\frac{f}{f_0}\right)^{\overset{\text{Power-law index}}{n}}$$

- We use Whittle's likelihood (multivariate case):

- Power-law template to model the SGWB PSD:

$$S_h(f) = \Omega_{\mathrm{GW}}(f)\frac{3H_0^2}{4\pi^2 f^3} \qquad \Omega_{\mathrm{GW}}(f) = \underbrace{\Omega_{\mathrm{GW},0}}_{\text{Energy density}}\left(\frac{f}{f_0}\right)^{\overbrace{n}^{\text{Power-law index}}}$$

- We use Whittle's likelihood (multivariate case):

$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) = \Pi_{k=0}^{N-1}\det(\tilde{\Sigma}_d(f_k))^{-1}\exp\left(-\tilde{\mathbf{d}}(f_k)^{\dagger}\tilde{\boldsymbol{\Sigma}}_d^{-1}(f_k)\tilde{\mathbf{d}}(f_k)\right)$$

Modelling the spectrum matrix

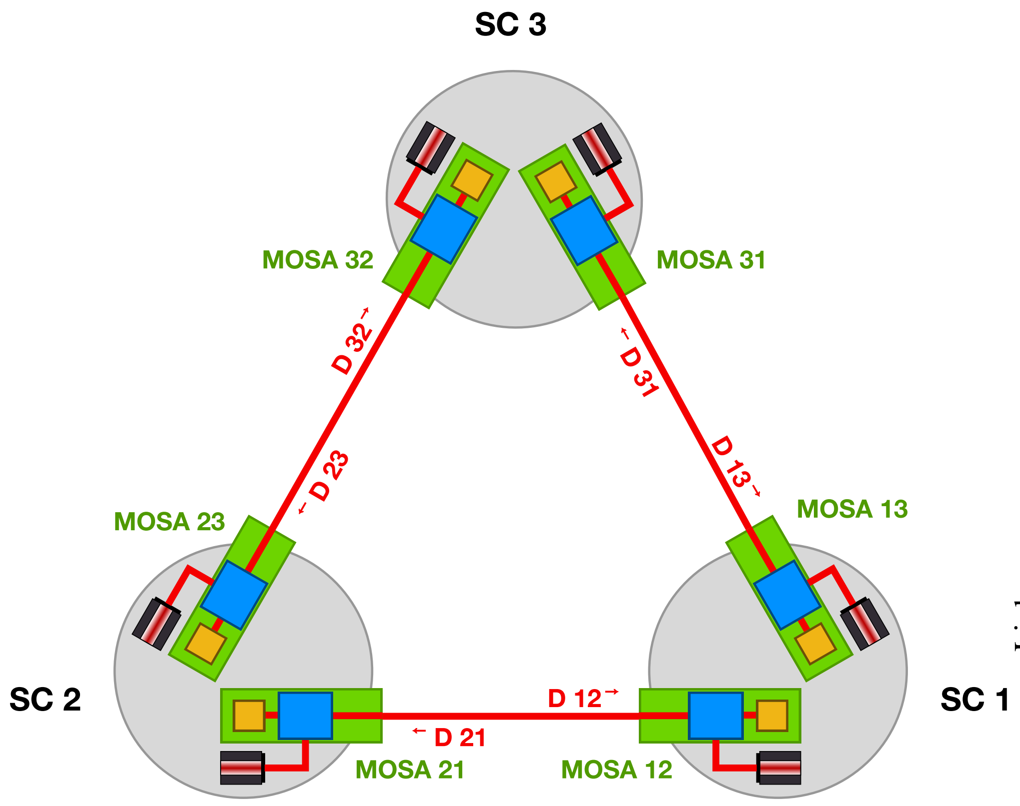
- Power-law template to model the SGWB PSD:

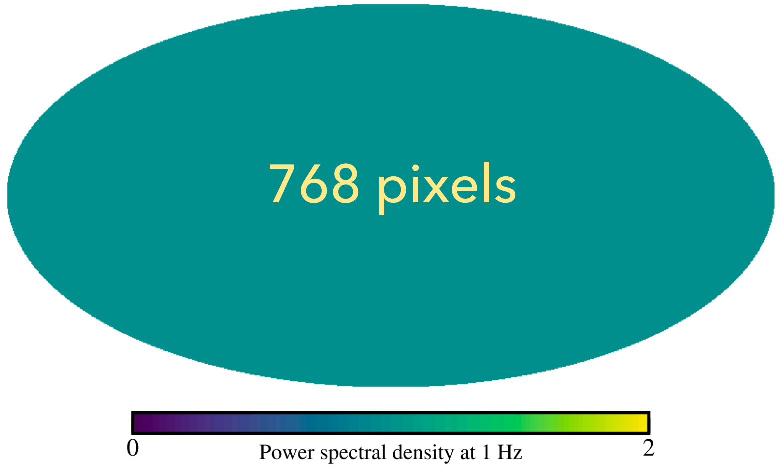
$$S_h(f) = \Omega_{\mathrm{GW}}(f)\frac{3H_0^2}{4\pi^2 f^3} \qquad \Omega_{\mathrm{GW}}(f) = \underbrace{\Omega_{\mathrm{GW},0}}_{\text{Energy density}}\left(\frac{f}{f_0}\right)^{\overbrace{n}^{\text{Power-law index}}}$$

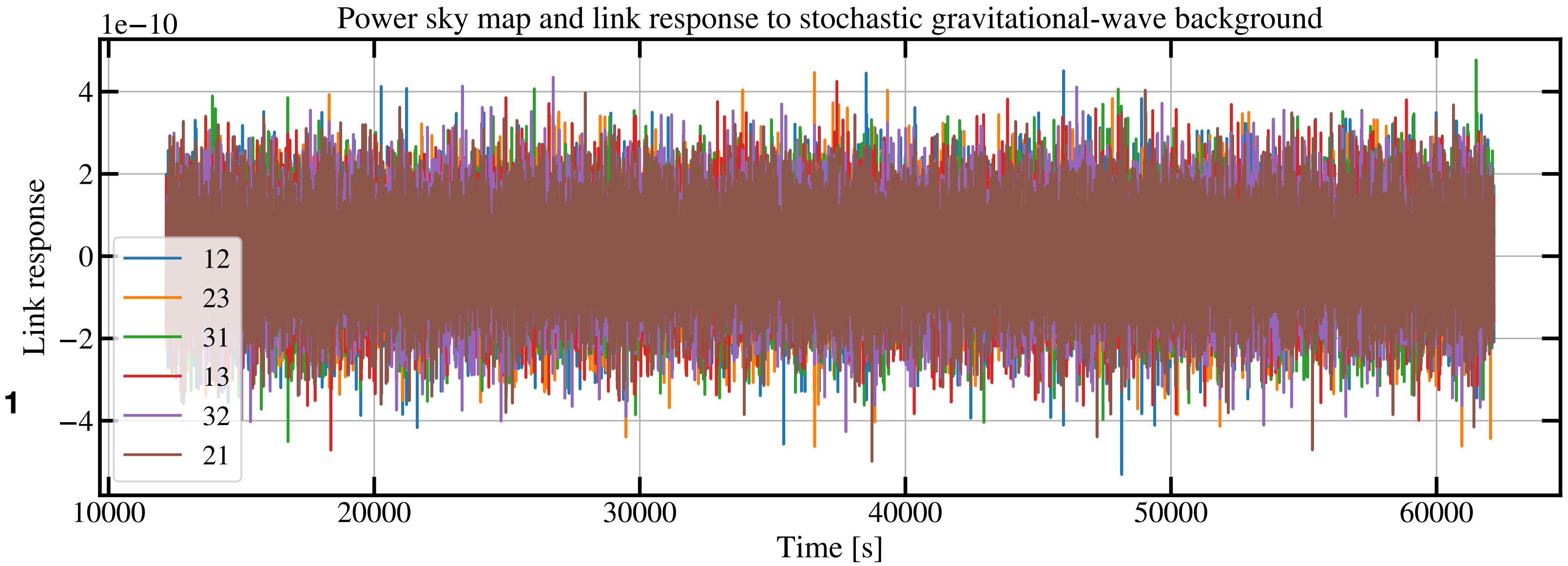
- We use Whittle's likelihood (multivariate case):

$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) = \Pi_{k=0}^{N-1}\det(\tilde{\Sigma}_d(f_k))^{-1}\exp\left(-\underbrace{\tilde{\mathbf{d}}(f_k)^\dagger}_{\text{Fourier-transformed TDI data}}\underbrace{\tilde{\boldsymbol{\Sigma}}_d^{-1}(f_k)}_{\text{Full TDI covariance matrix}}\tilde{\mathbf{d}}(f_k)\right)$$

- We generate single-link measurements using LISA GW Response [Bayle, Baghi, Renzini, Le Jeune 2022]
- Obtain 6 science interferometer time series
- Add instrumental noise from prescribed PSD







- Process data through time-delay interferometry (TDI) using the pyTDI code [Staab, Bayle, Hartwig 2022]
- Time-varying arm lengths → second-generation TDI

Bayesian inference

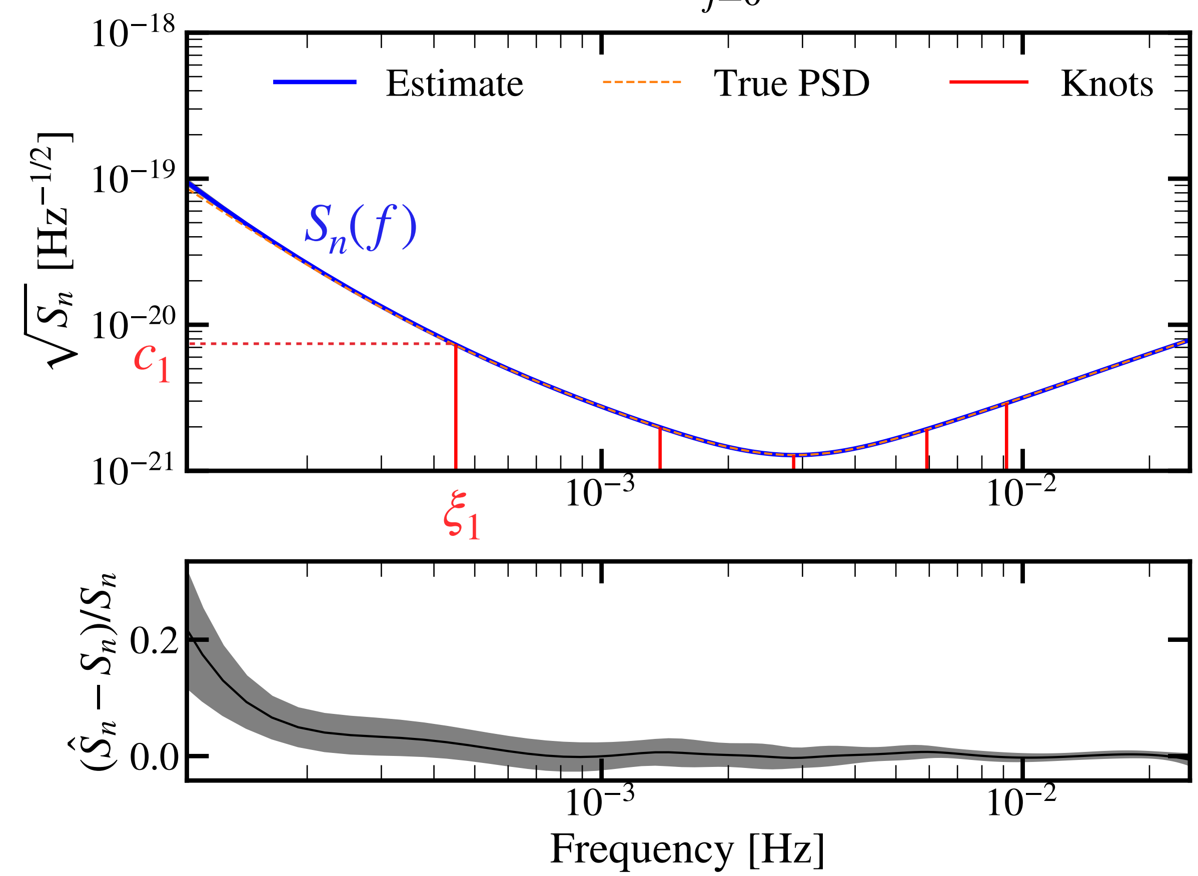
- Bayesian data analysis :
  - Sampling posterior distributions with parallel tempered Markov chain Monte Carlo (MCMC)
  - Uniform priors on GW parameters $\Omega_{m0} \in [10^{-16}, 10^{-14}]$ and $n \in [-5, 7]$
  - Uniform prior on noise PSD level: 1 order of magnitude deviation allowed

- Bayesian data analysis :
  - ✦ Sampling posterior distributions with parallel tempered Markov chain Monte Carlo (MCMC)
  - ✦ Uniform priors on GW parameters $\Omega_{m0} \in [10^{-16}, 10^{-14}]$ and $n \in [-5, 7]$
  - ✦ Uniform prior on noise PSD level: 1 order of magnitude deviation allowed

Bayesian inference

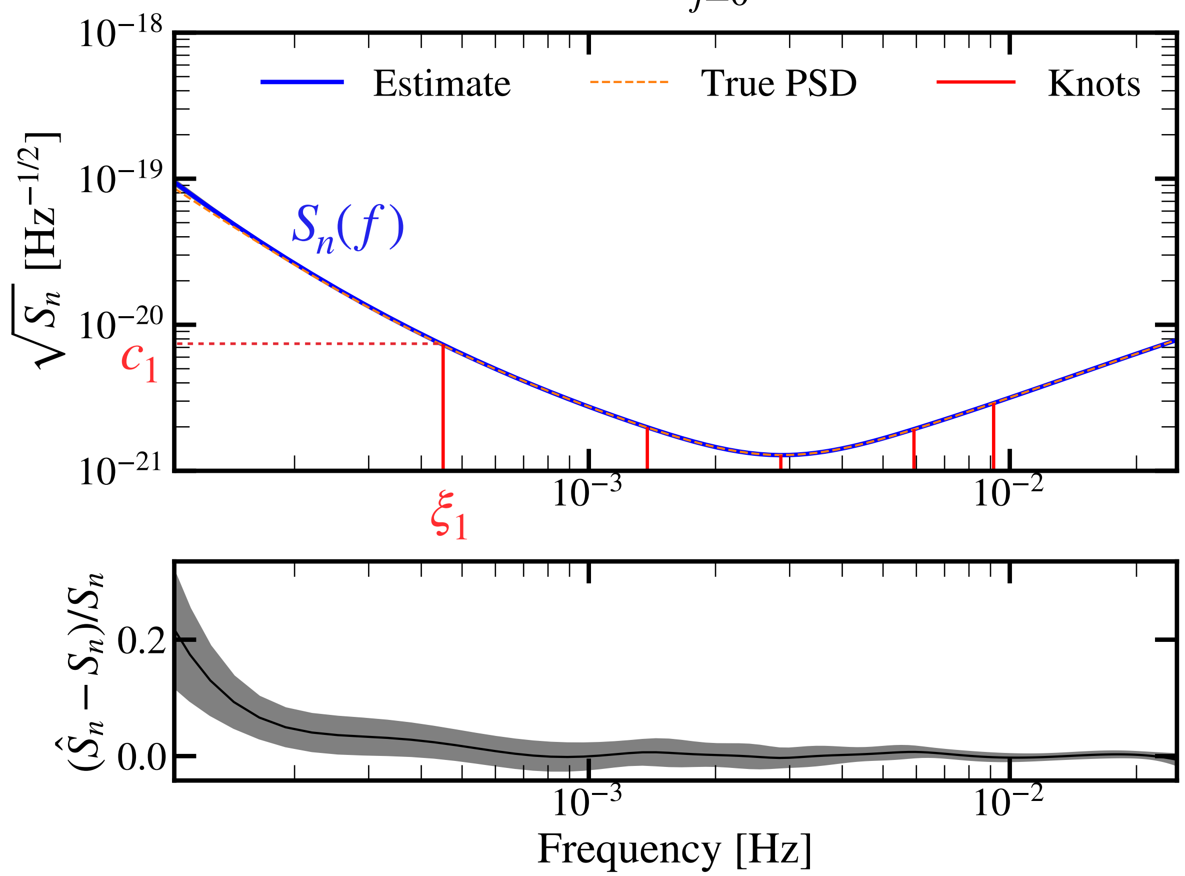
- Bayesian data analysis :
  - ✦ Sampling posterior distributions with parallel tempered Markov chain Monte Carlo (MCMC)
  - ✦ Uniform priors on GW parameters $\Omega_{m0} \in [10^{-16}, 10^{-14}]$ and $n \in [-5, 7]$
  - ✦ Uniform prior on noise PSD level: 1 order of magnitude deviation allowed

$$\log S_n(f) = \sum_{j=0}^{K-1} c_j B_j\left(f, \boldsymbol{\xi}\right)$$

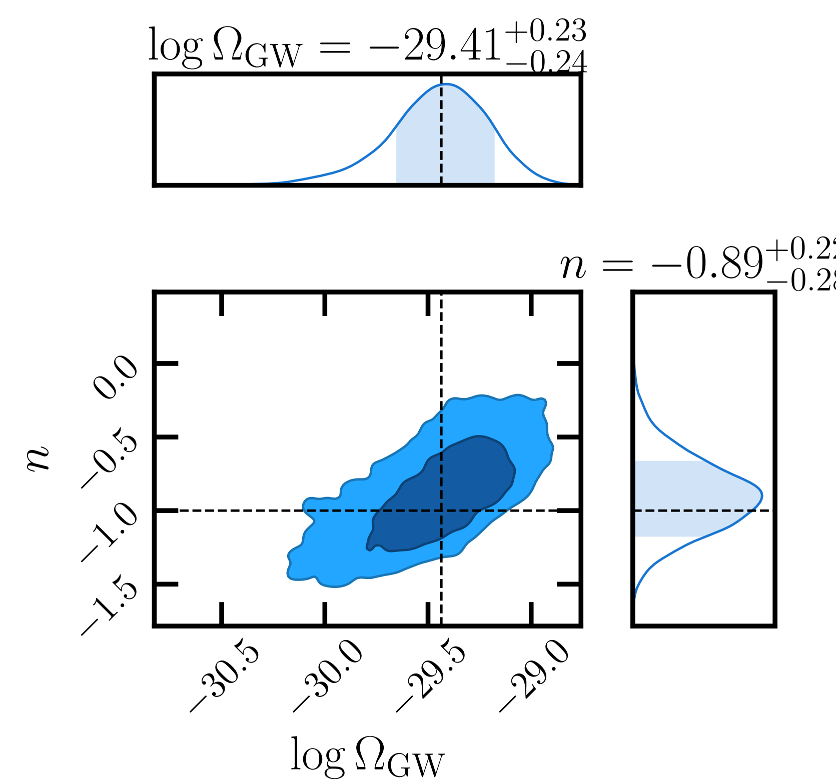


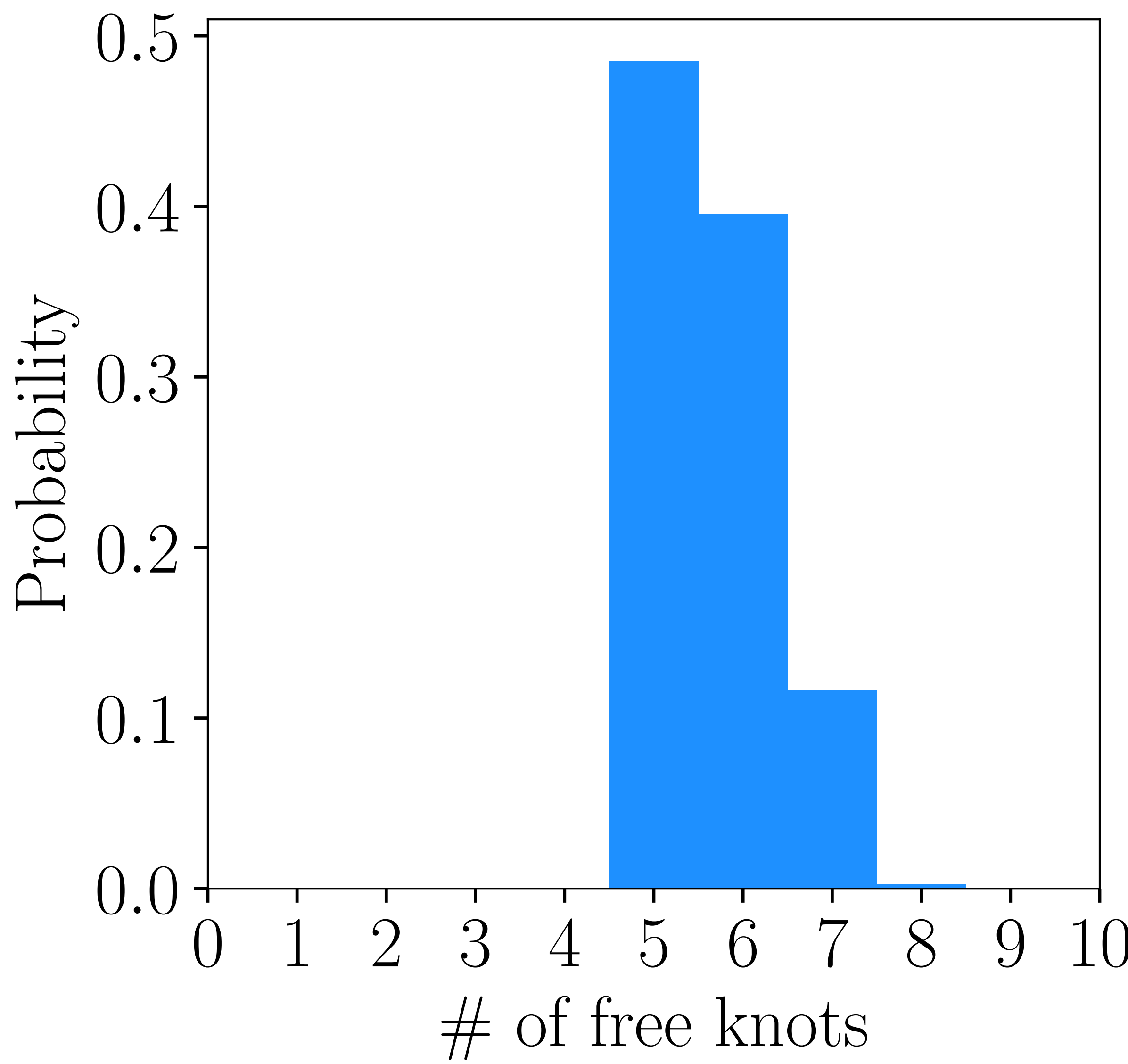
- Bayesian data analysis :
  - ✦ Sampling posterior distributions with parallel tempered Markov chain Monte Carlo (MCMC)
  - ✦ Uniform priors on GW parameters $\Omega_{m0} \in [10^{-16}, 10^{-14}]$ and $n \in [-5, 7]$
  - ✦ Uniform prior on noise PSD level: 1 order of magnitude deviation allowed

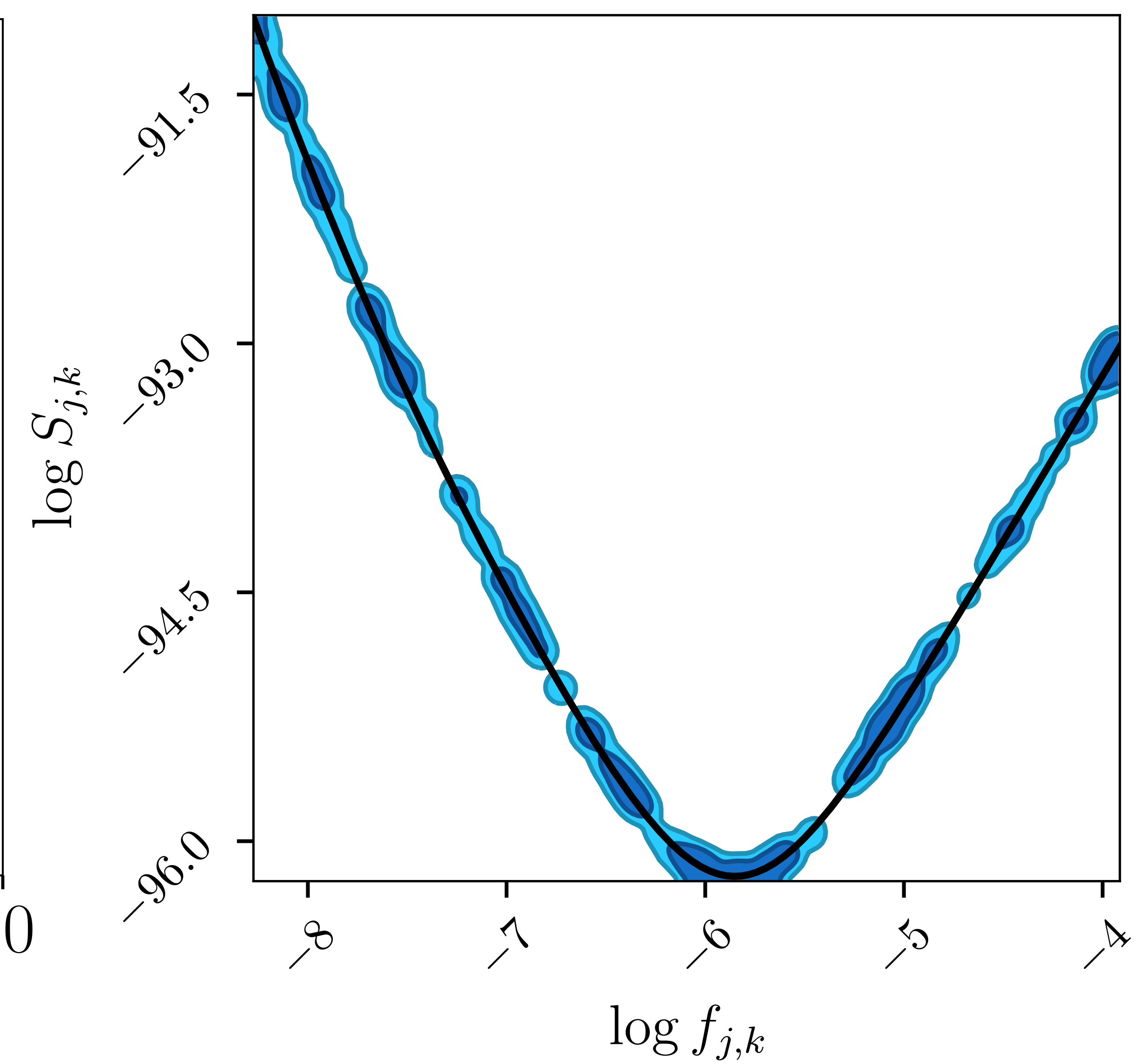
$$\log S_n(f) = \sum_{j=0}^{K-1} c_j B_j\left(f, \boldsymbol{\xi}\right)$$



$$\Omega_{\mathrm{GW}}(f) = \Omega_{\mathrm{GW},0}\left(\frac{f}{f_0}\right)^n$$







[Done with Eryn: Karnesis et al., 2023]

Bayesian inference

- Detection using Bayesian model comparison
  - Hypothesis H_0: only noise in the data $\quad \tilde{d} = M\tilde{n}$
  - Hypothesis H_1: presence of a SGWB $\quad \tilde{d} = M\left(\tilde{h} + \tilde{n}\right)$

$$Z_i = \int_\Theta p(d \,|\, \theta, H_i)\, p(\theta)\, d\theta$$

$$\mathscr{B}_{10} = \frac{Z_1}{Z_0}$$

- Detection using Bayesian model comparison
  - ✦ Hypothesis H_0: only noise in the data $\tilde{d} = M\tilde{n}$
  - ✦ Hypothesis H_1: presence of a SGWB $\tilde{d} = M\left(\tilde{h} + \tilde{n}\right)$

$$Z_i = \int_{\Theta} p(d \,|\, \theta, H_i)\, p(\theta)\, d\theta$$

- Aim: compute the Bayes factors for a range of configurations (Ω_{m0}, n)

$$\mathscr{B}_{10} = \frac{Z_1}{Z_0}$$

- Detection using Bayesian model comparison

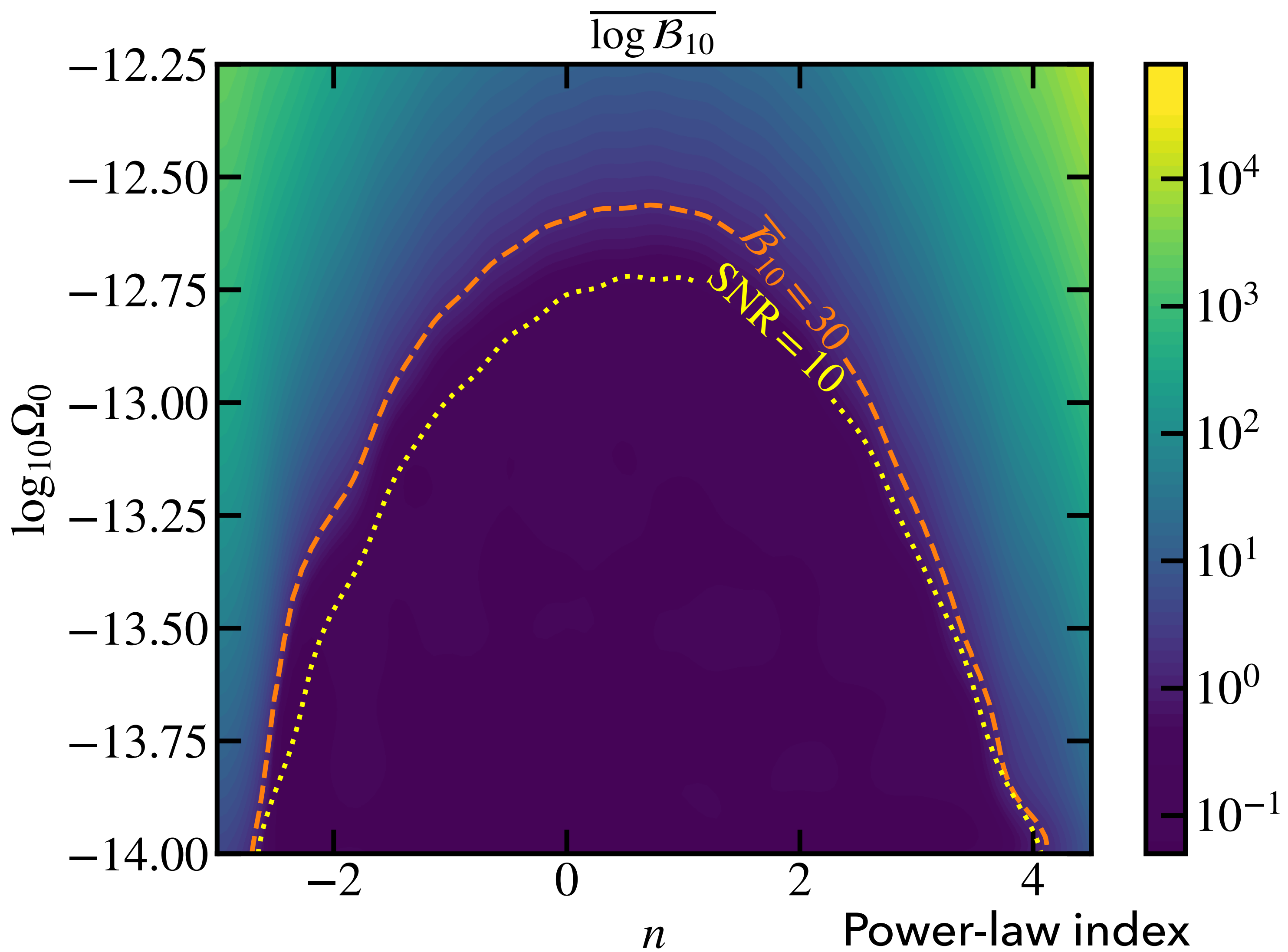
✦ Hypothesis H_0: only noise in the data $\tilde{d} = M\tilde{n}$

✦ Hypothesis H_1: presence of a SGWB $\tilde{d} = M\left(\tilde{h} + \tilde{n}\right)$

$$Z_i = \int_\Theta p(d\,|\,\theta, H_i)\,p(\theta)\,d\theta$$

- Aim: compute the Bayes factors for a range of configurations (Ω_{m0}, n)

$$\mathscr{B}_{10} = \frac{Z_1}{Z_0}$$



Log-energy density

Power-law index

- Detection using Bayesian model comparison

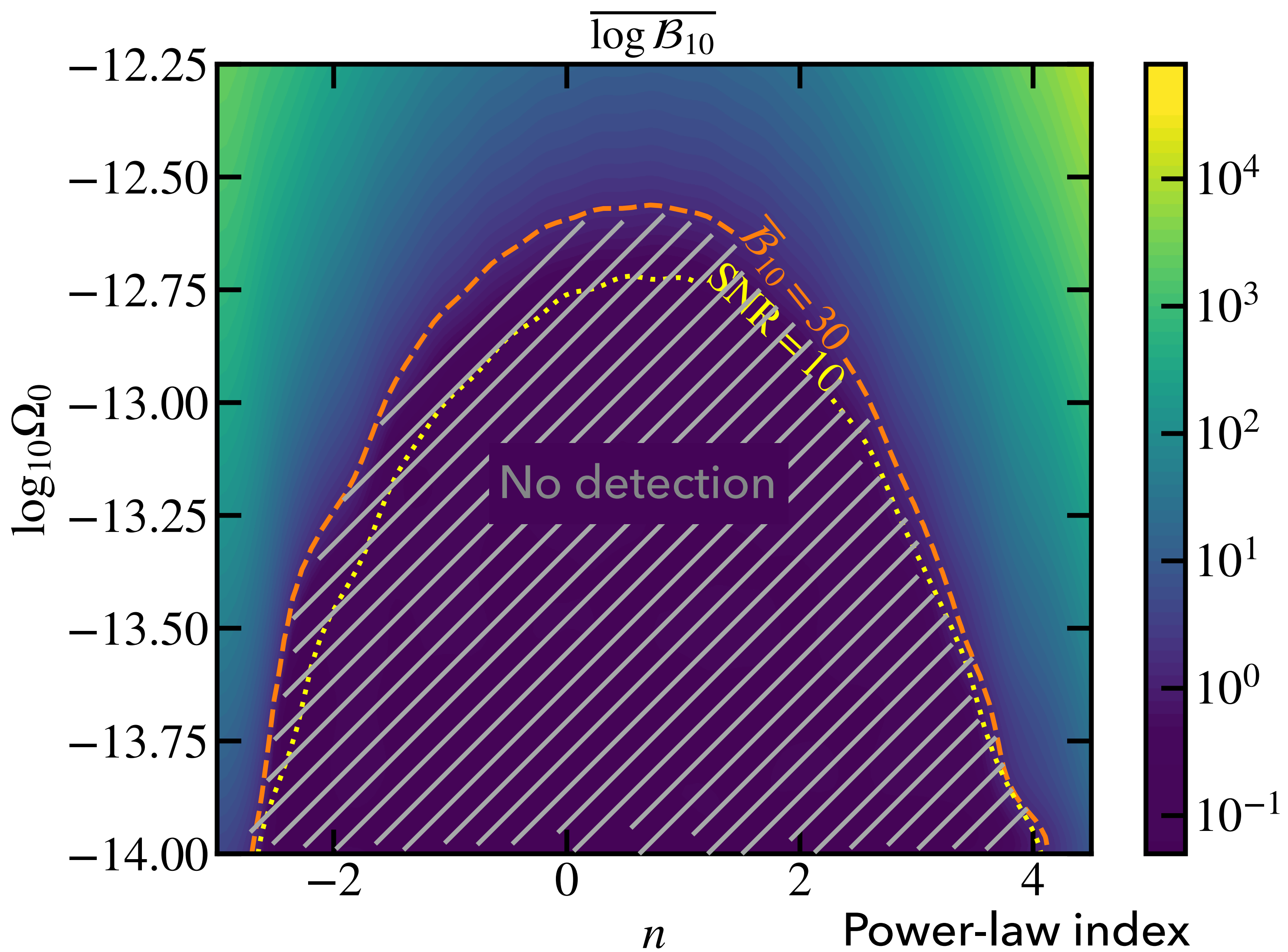
  - Hypothesis H_0: only noise in the data $\tilde{d} = M\tilde{n}$
  - Hypothesis H_1: presence of a SGWB $\tilde{d} = M\left(\tilde{h} + \tilde{n}\right)$

$$Z_i = \int_\Theta p(d \,|\, \theta, H_i)\, p(\theta)\, d\theta$$

- Aim: compute the Bayes factors for a range of configurations (Ω_{m0}, n)

$$\mathscr{B}_{10} = \frac{Z_1}{Z_0}$$

Log-energy density



Power-law index

- Detection using Bayesian model comparison

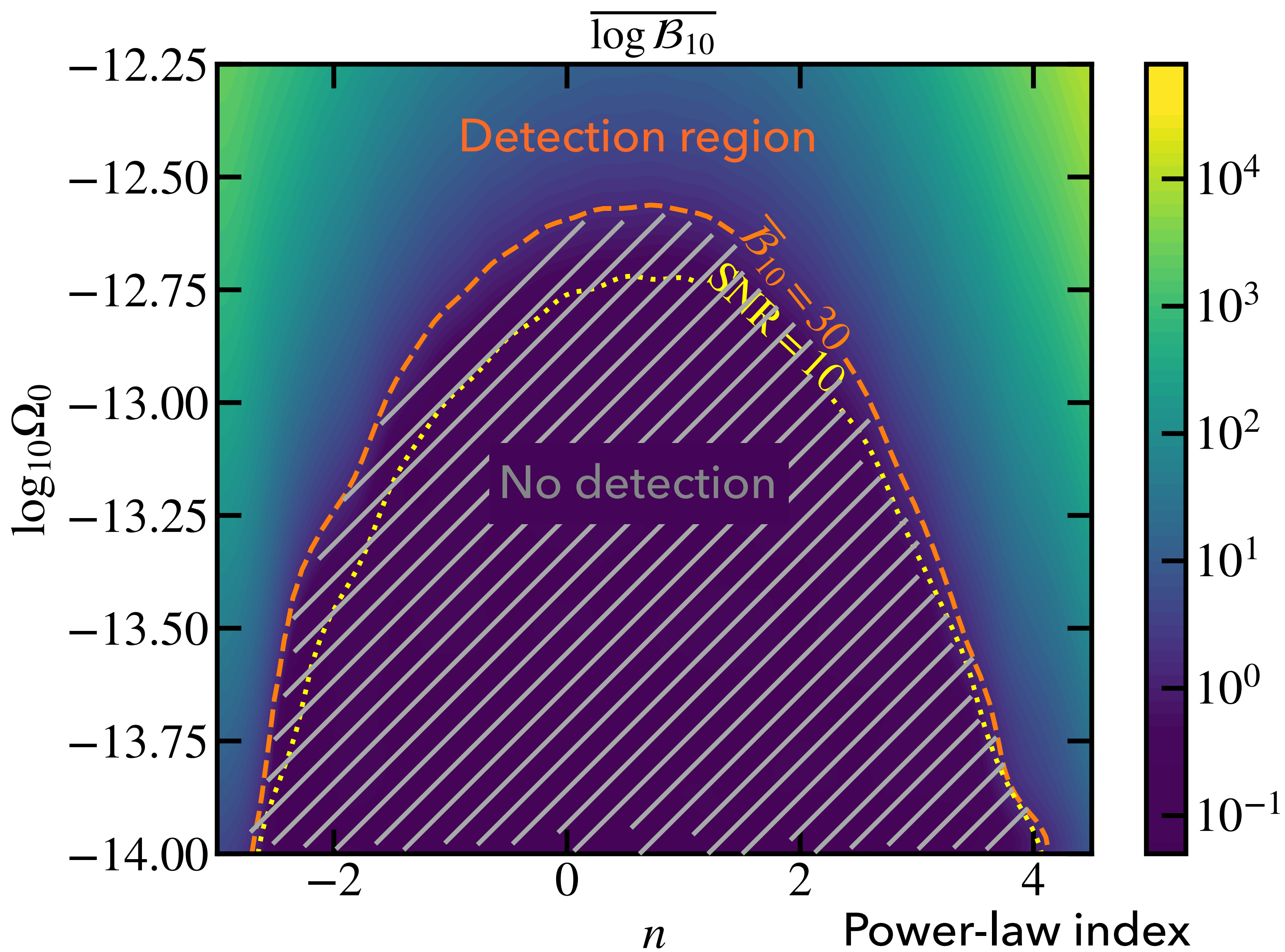
✦ Hypothesis H_0: only noise in the data $\tilde{d} = M\tilde{n}$

✦ Hypothesis H_1: presence of a SGWB $\tilde{d} = M\left(\tilde{h} + \tilde{n}\right)$

$$Z_i = \int_\Theta p(d \mid \theta, H_i)\, p(\theta)\, d\theta$$

- Aim: compute the Bayes factors for a range of configurations (Ω_{m0}, n)

$$\mathscr{B}_{10} = \frac{Z_1}{Z_0}$$



Log-energy density

Power-law index

- In reality, we will have different noises, with different transfer functions

$$\tilde{\mathbf{\Sigma}}_{\mathrm{noise}}(f) = \sum_{i} \tilde{\mathbf{R}}_{\mathrm{noise},i}(f) S_{n,i}(f)$$

→ One spline for each PSD

- In reality, we will have different noises, with different transfer functions

$$\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \sum_i \tilde{\mathbf{R}}_{\text{noise},i}(f) S_{n,i}(f)$$

→ One spline for each PSD

- Even worse: the single-link measurements η_{ij} will have different PSDs!

$$\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \sum_i \tilde{\mathbf{M}}_{\text{noise},i}(f) \mathbf{S}_{n,i}(f) \tilde{\mathbf{M}}_{\text{noise},i}(f)^{\dagger}$$

- In reality, we will have different noises, with different transfer functions

$$\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \sum_i \tilde{\mathbf{R}}_{\text{noise},i}(f) S_{n,i}(f)$$

→ One spline for each PSD

- Even worse: the single-link measurements η_{ij} will have different PSDs!

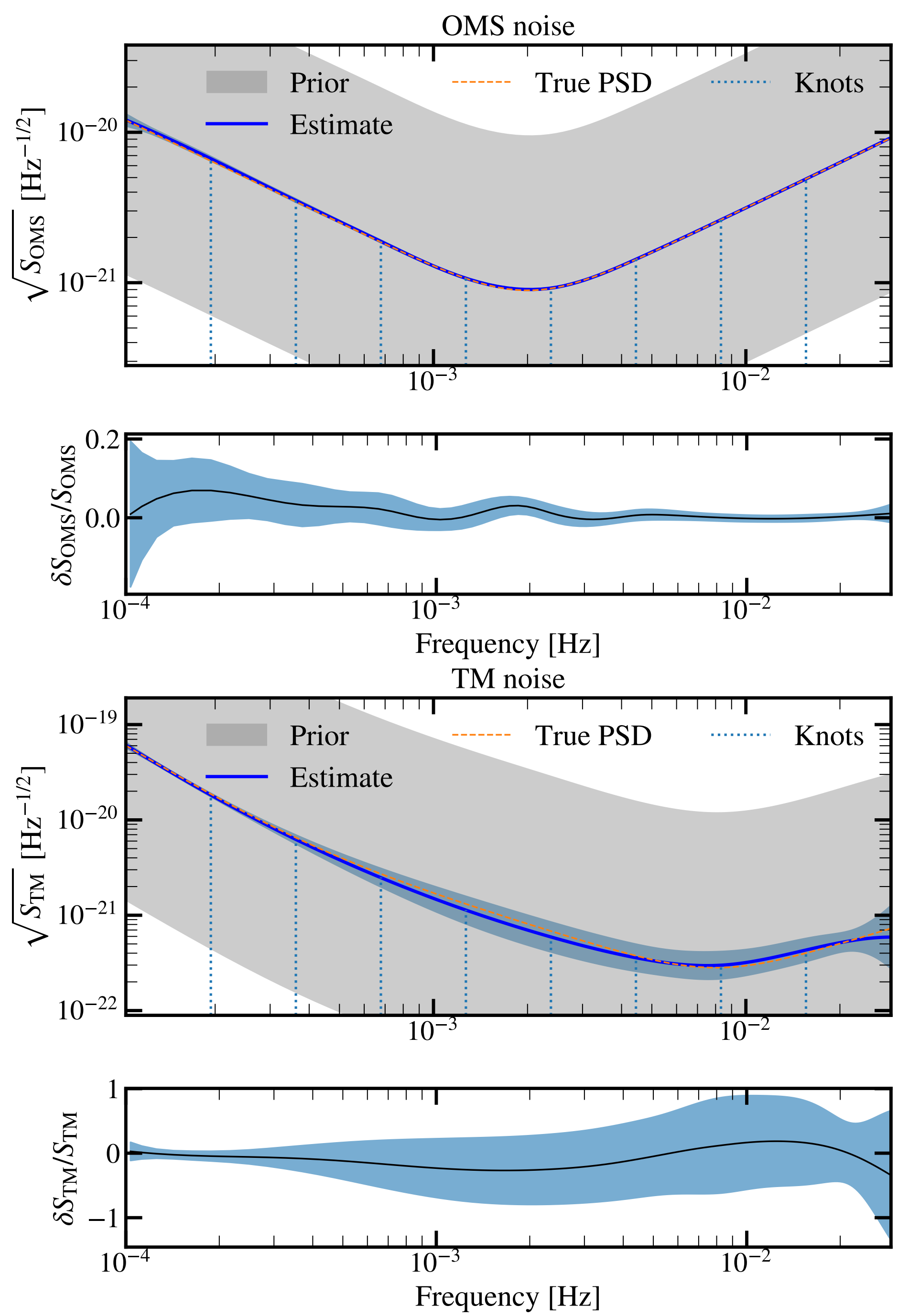
$$\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \sum_i \tilde{\mathbf{M}}_{\text{noise},i}(f) \mathbf{S}_{n,i}(f) \tilde{\mathbf{M}}_{\text{noise},i}(f)^{\dagger}$$

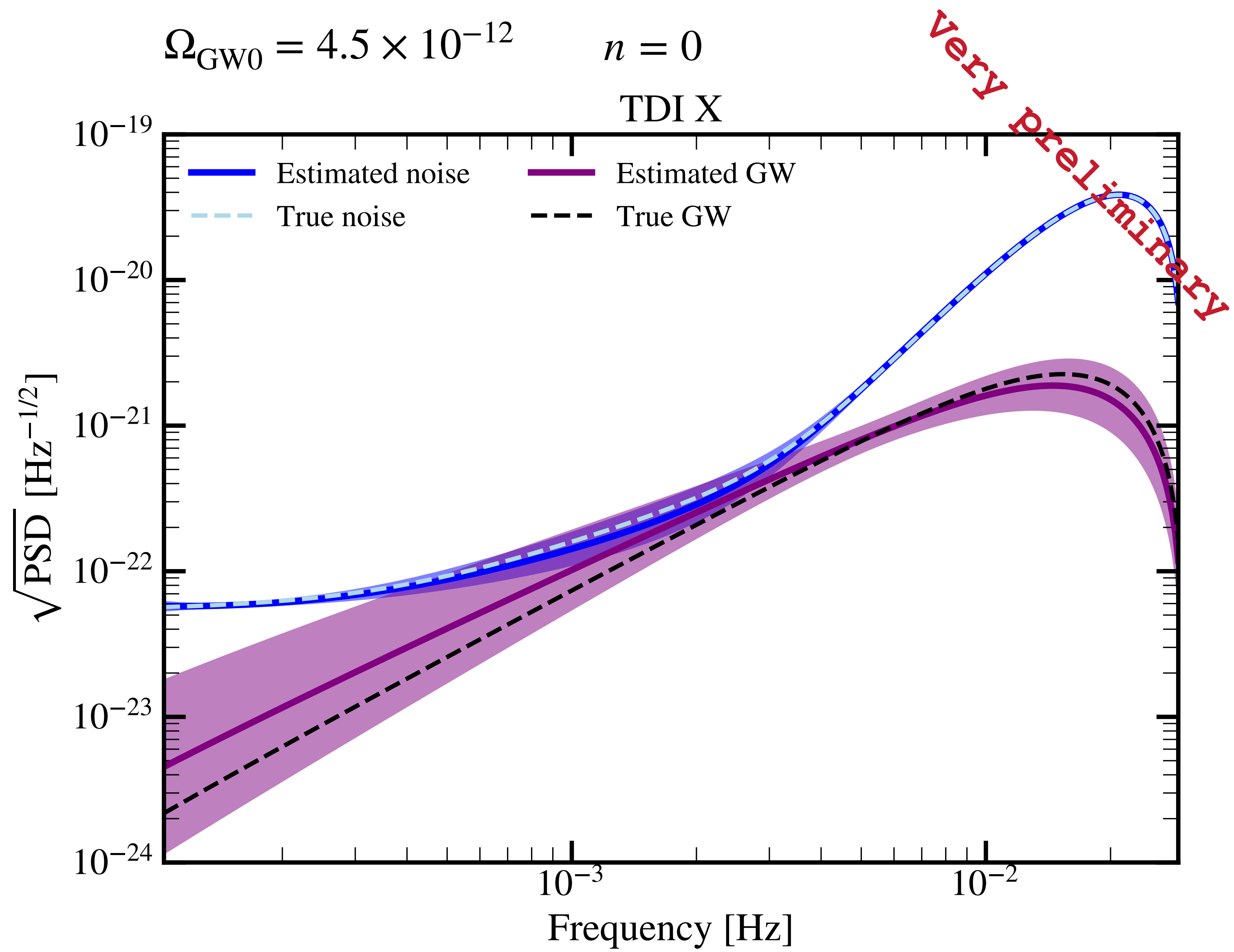
- But we only have 9 observable degrees of freedom, so we might rather directly fit for them…?

$$\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \begin{pmatrix} S_{XX}(f) & S_{XY}(f) & S_{XZ}(f) \\ S_{YX}(f) & S_{YY}(f) & S_{YZ}(f) \\ S_{ZX}(f) & S_{ZY}(f) & S_{ZZ}(f) \end{pmatrix}$$

- Attempt with 2 noise components: optical metrology system (OMS) and test-mass (TM) noises

$$\tilde{\mathbf{\Sigma}}_{\mathrm{noise}}(f) = \tilde{\mathbf{R}}_{\mathrm{OMS}}(f)S_{\mathrm{OMS}}(f) + \tilde{\mathbf{R}}_{\mathrm{TM}}(f)S_{\mathrm{TM}}(f)$$





- Going beyond: can the concept of null channel help?

Revisiting null channels

- **Classic definition following Prince et al.** (2002). We can look for a combination e of TDI variables that maximises signal-to-noise ratio (SNR) of a deterministic signal $\mathbf{s}$:

$$e = a_1 X + a_2 Y + a_3 Z \qquad \mathrm{SNR} = \int_{f_l}^{f_u} \frac{\mathbf{a}^\dagger \tilde{\boldsymbol{A}}_{\mathrm{GW}} \mathbf{a}}{\mathbf{a}^\dagger \tilde{\boldsymbol{\Sigma}}_{\mathrm{noise}} \mathbf{a}} df \qquad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Where $\tilde{\boldsymbol{A}}_{\mathrm{GW}}$ is the matrix of cross-products of the signal vector $\tilde{\mathbf{s}}$: $\tilde{\boldsymbol{A}}_{\mathrm{GW}} = \tilde{\mathbf{s}}\tilde{\mathbf{s}}^\dagger$

- **Classic definition following Prince et al.** (2002). We can look for a combination e of TDI variables that maximises signal-to-noise ratio (SNR) of a deterministic signal $\mathbf{s}$:

$$e = a_1 X + a_2 Y + a_3 Z \qquad \mathrm{SNR} = \int_{f_l}^{f_u} \frac{\mathbf{a}^\dagger \tilde{A}_{\mathrm{GW}} \mathbf{a}}{\mathbf{a}^\dagger \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} \mathbf{a}} df \qquad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Where $\tilde{A}_{\mathrm{GW}}$ is the matrix of cross-products of the signal vector $\tilde{\mathbf{s}}$: $\tilde{A}_{\mathrm{GW}} = \tilde{\mathbf{s}}\tilde{\mathbf{s}}^\dagger$

- Which from Rayleigh's principle is equivalent to the generalised eigenvalue problem [Borloz & Xerri 2005]:

$$\tilde{A}_{\mathrm{GW}} \mathbf{a} = \lambda \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} \mathbf{a}$$

- **Classic definition following Prince et al.** (2002). We can look for a combination e of TDI variables that maximises signal-to-noise ratio (SNR) of a deterministic signal $\mathbf{s}$:

$$e = a_1 X + a_2 Y + a_3 Z \qquad \mathrm{SNR} = \int_{f_l}^{f_u} \frac{\mathbf{a}^\dagger \tilde{\boldsymbol{A}}_{\mathrm{GW}} \mathbf{a}}{\mathbf{a}^\dagger \tilde{\boldsymbol{\Sigma}}_{\mathrm{noise}} \mathbf{a}} df \qquad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Where $\tilde{\boldsymbol{A}}_{\mathrm{GW}}$ is the matrix of cross-products of the signal vector $\tilde{\mathbf{s}}$: $\tilde{\boldsymbol{A}}_{\mathrm{GW}} = \tilde{\mathbf{s}}\tilde{\mathbf{s}}^\dagger$

- Which from Rayleigh's principle is equivalent to the generalised eigenvalue problem [Borloz & Xerri 2005]:

$$\tilde{\boldsymbol{A}}_{\mathrm{GW}} \mathbf{a} = \lambda \tilde{\boldsymbol{\Sigma}}_{\mathrm{noise}} \mathbf{a}$$

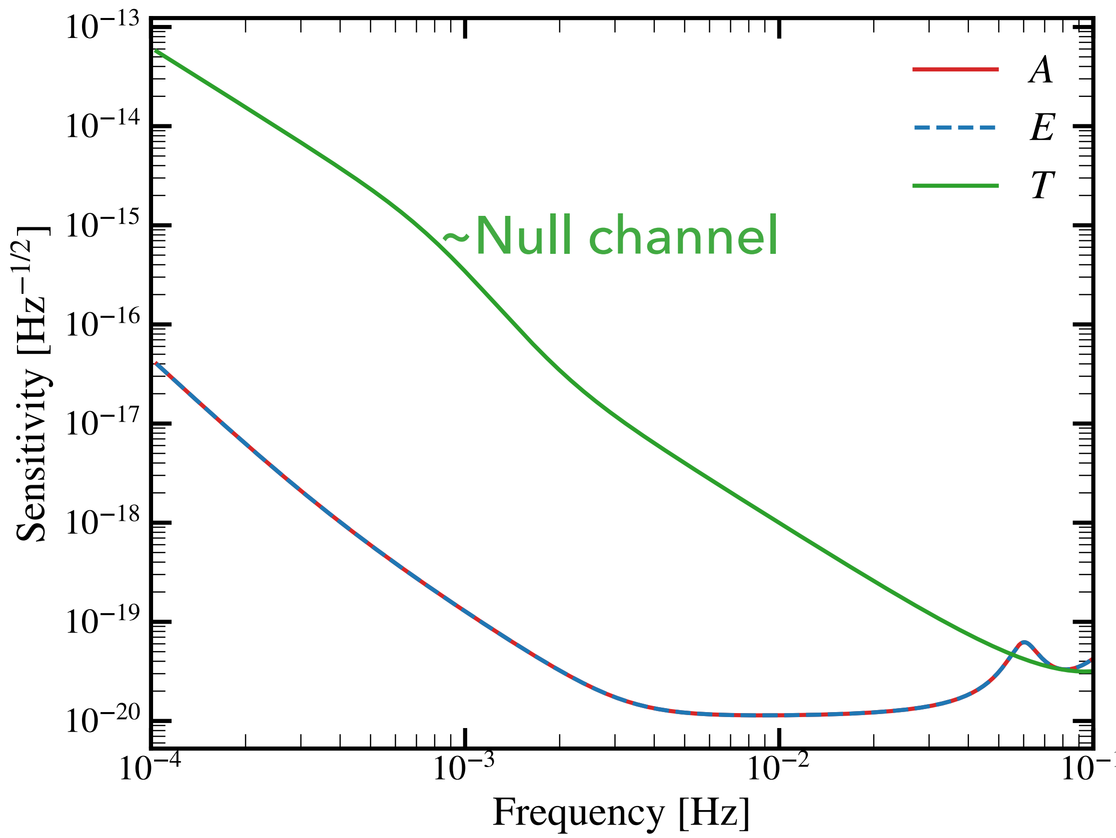
- For deterministic signals $\tilde{\boldsymbol{A}}_{\mathrm{GW}}$ is of rank 1 so it is equivalent to performing the **eigendecomposition of the noise covariance matrix $\tilde{\boldsymbol{\Sigma}}_{\mathrm{noise}}$**

$$\tilde{\boldsymbol{\Sigma}}_{\mathrm{noise}}(f) = \begin{pmatrix} S_{XX}(f) & S_{XY}(f) & S_{XZ}(f) \\ S_{YX}(f) & S_{YY}(f) & S_{YZ}(f) \\ S_{ZX}(f) & S_{ZY}(f) & S_{ZZ}(f) \end{pmatrix}$$

- That means we compute $\tilde{\mathbf{\Sigma}}_{\text{noise}}(f) = \mathbf{V}(f)\mathbf{\Lambda}(f)\mathbf{V}(f)^{\dagger}$ so that the transformation $\tilde{\mathbf{e}} = \mathbf{V}^{\dagger}\tilde{\mathbf{d}}$ forms a set of orthogonal channels
- **The null channel corresponds to the eigenvector with smallest SNR**
- Under specific assumptions
  - ✦ Fixed arm lengths → first-generation TDI
  - ✦ Equal armlengths
  - ✦ Equal interferometric noises → all η_{ij} have the same PSDs
- The diagonalisation is independent of frequency and noise levels → A, E, T

$$\mathbf{V} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

$\propto \mathrm{SNR}^{-1}$



Both noise and signal are orthogonal

But unrealistic assumptions!

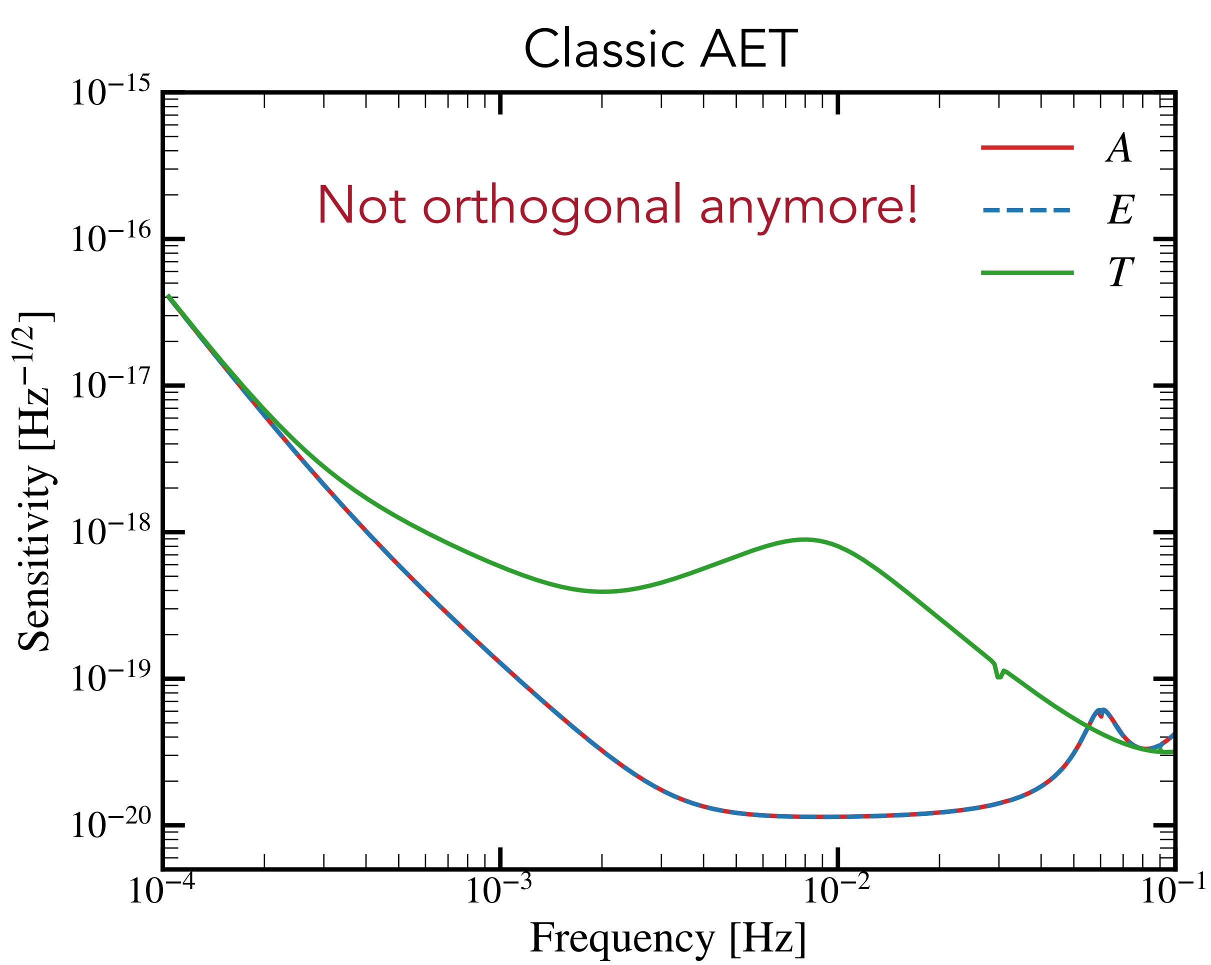
- Now assume

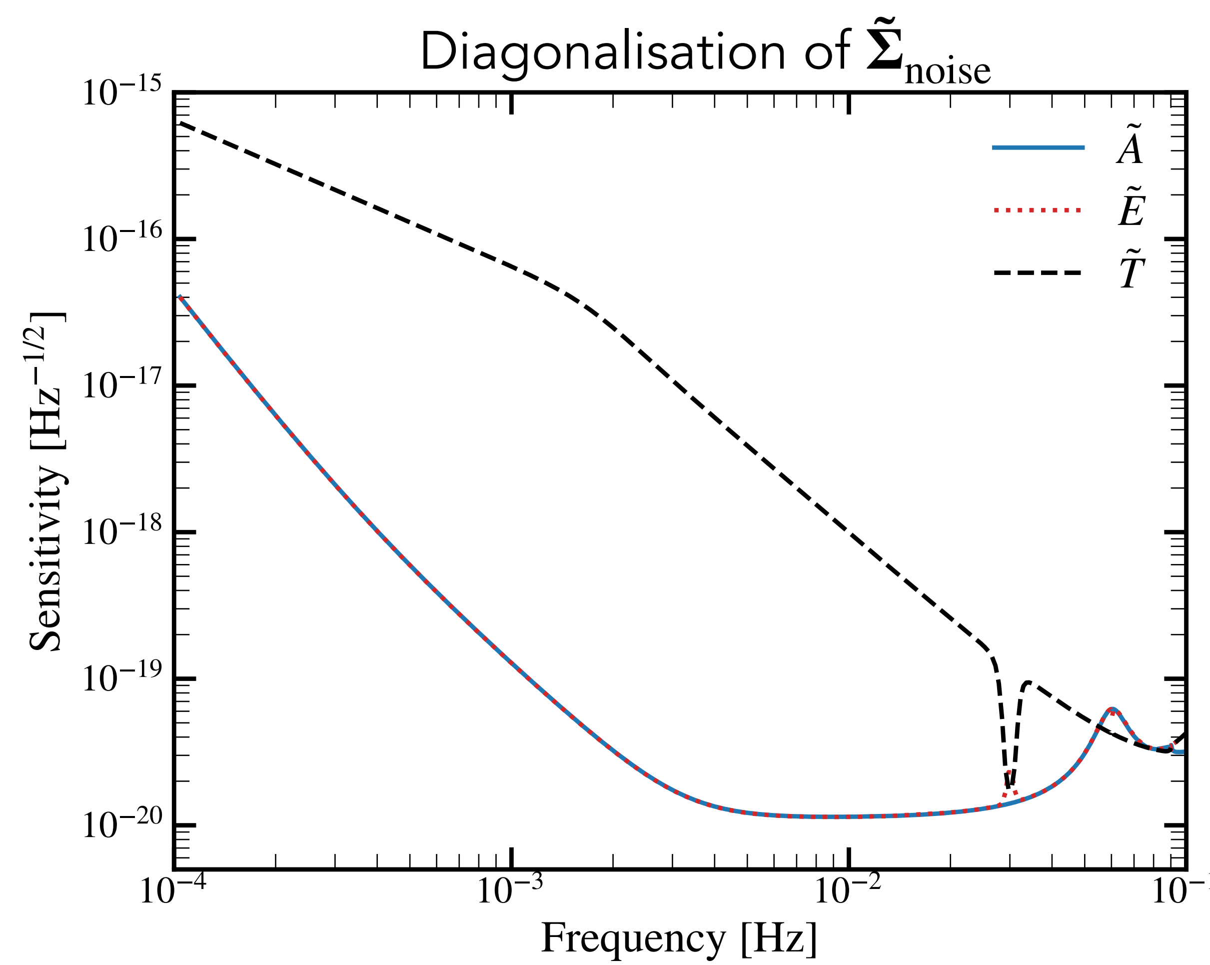
    ✦ **Flexing armlenghts** → second-generation TDI

    ✦ **Unequal armlengths**

    ✦ Equal interferometer noises

- The eigenvectors $\mathbf{V}(f)$ of $\tilde{\mathbf{\Sigma}}_{\text{noise}}(f)$ are now frequency dependent!





- But we are looking for stochastic signals. What is a null channel in this case?
- The definition of the optimal SNR needs to account for SGWB correlations:

$$\mathrm{SNR} = \int_{f_l}^{f_u} \frac{\mathbf{a}^{\dagger}\tilde{\mathbf{\Sigma}}_{\mathrm{GW}}\mathbf{a}}{\mathbf{a}^{\dagger}\tilde{\mathbf{\Sigma}}_{\mathrm{noise}}\mathbf{a}} df \qquad \tilde{\mathbf{\Sigma}}_{\mathrm{GW}} = \mathrm{E}\left[\mathbf{s}\mathbf{s}^{\dagger}\right]$$

- But we are looking for stochastic signals. What is a null channel in this case?
- The definition of the optimal SNR needs to account for SGWB correlations:

$$\mathrm{SNR} = \int_{f_l}^{f_u} \frac{\mathbf{a}^\dagger \tilde{\boldsymbol{\Sigma}}_{\mathrm{GW}} \mathbf{a}}{\mathbf{a}^\dagger \tilde{\boldsymbol{\Sigma}}_{\mathrm{noise}} \mathbf{a}} df \qquad \tilde{\boldsymbol{\Sigma}}_{\mathrm{GW}} = \mathrm{E}\left[\mathbf{s}\mathbf{s}^\dagger\right]$$

- The generalised eigenvalue problem we need to solve is now

$$\tilde{\boldsymbol{\Sigma}}_{\mathrm{GW}} \mathbf{a} = \lambda \tilde{\boldsymbol{\Sigma}}_{\mathrm{noise}} \mathbf{a} \qquad \tilde{\boldsymbol{\Sigma}}_{\mathrm{GW}}(f) = \tilde{\mathbf{R}}_{\mathrm{GW}}(f) S_h(f)$$

- But we are looking for stochastic signals. What is a null channel in this case?
- The definition of the optimal SNR needs to account for SGWB correlations:

$$\mathrm{SNR} = \int_{f_l}^{f_u} \frac{\mathbf{a}^\dagger \tilde{\mathbf{\Sigma}}_{\mathrm{GW}} \mathbf{a}}{\mathbf{a}^\dagger \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} \mathbf{a}} df \qquad \tilde{\mathbf{\Sigma}}_{\mathrm{GW}} = \mathrm{E}\left[\mathbf{s}\mathbf{s}^\dagger\right]$$

- The generalised eigenvalue problem we need to solve is now

$$\tilde{\mathbf{\Sigma}}_{\mathrm{GW}} \mathbf{a} = \lambda \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} \mathbf{a} \qquad \tilde{\mathbf{\Sigma}}_{\mathrm{GW}}(f) = \tilde{\mathbf{R}}_{\mathrm{GW}}(f) S_h(f)$$

- **Involves both noise and signal orthogonalisation**! Equivalent to solving the eigenvalue problem:

- But we are looking for stochastic signals. What is a null channel in this case?
- The definition of the optimal SNR needs to account for SGWB correlations:

$$\mathrm{SNR} = \int_{f_l}^{f_u} \frac{\mathbf{a}^\dagger \tilde{\mathbf{\Sigma}}_{\mathrm{GW}} \mathbf{a}}{\mathbf{a}^\dagger \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} \mathbf{a}} df \qquad \tilde{\mathbf{\Sigma}}_{\mathrm{GW}} = \mathrm{E}\left[\mathbf{s}\mathbf{s}^\dagger\right]$$

- The generalised eigenvalue problem we need to solve is now

$$\tilde{\mathbf{\Sigma}}_{\mathrm{GW}} \mathbf{a} = \lambda \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} \mathbf{a} \qquad \tilde{\mathbf{\Sigma}}_{\mathrm{GW}}(f) = \tilde{\mathbf{R}}_{\mathrm{GW}}(f) S_h(f)$$

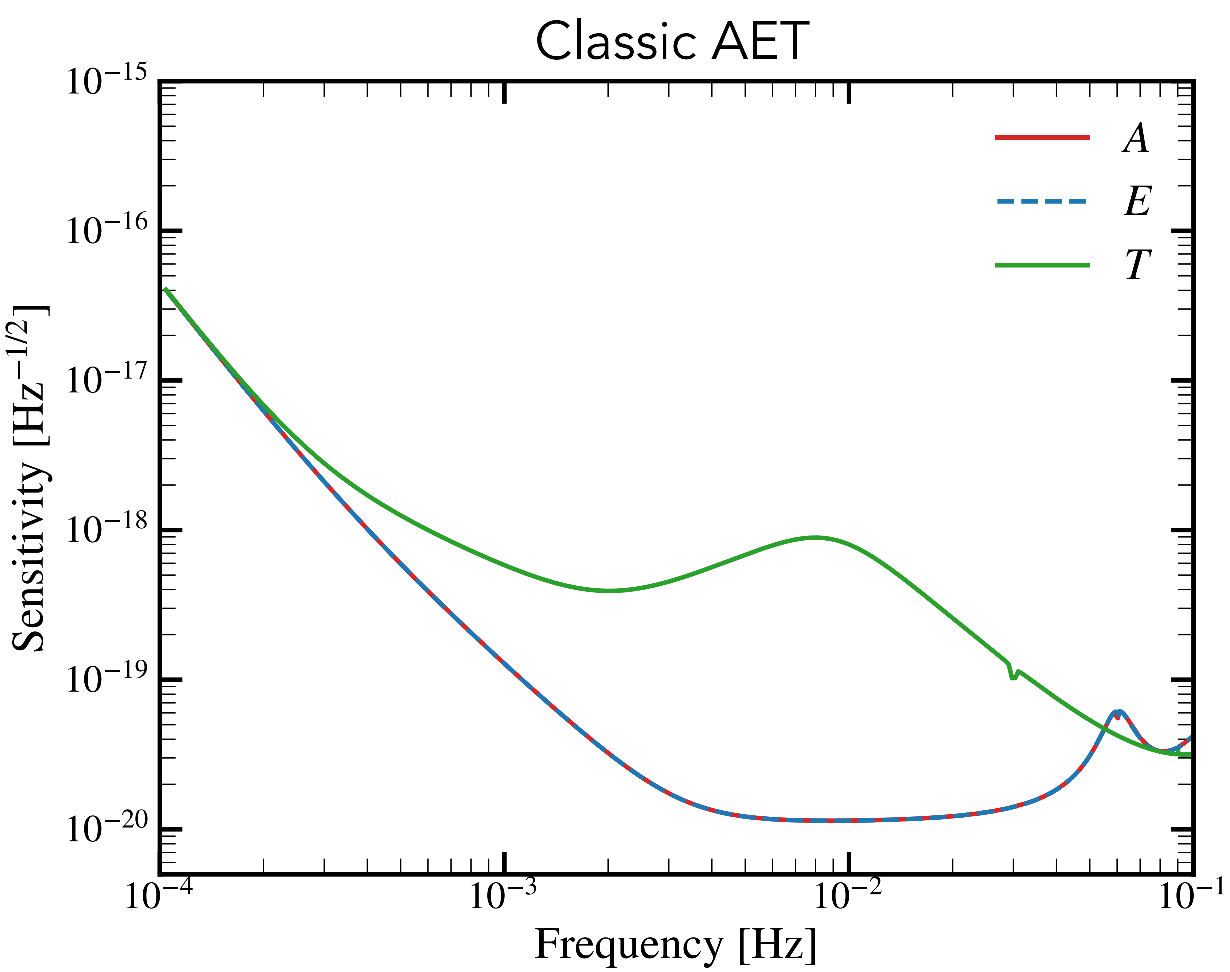
- **Involves both noise and signal orthogonalisation**! Equivalent to solving the eigenvalue problem:

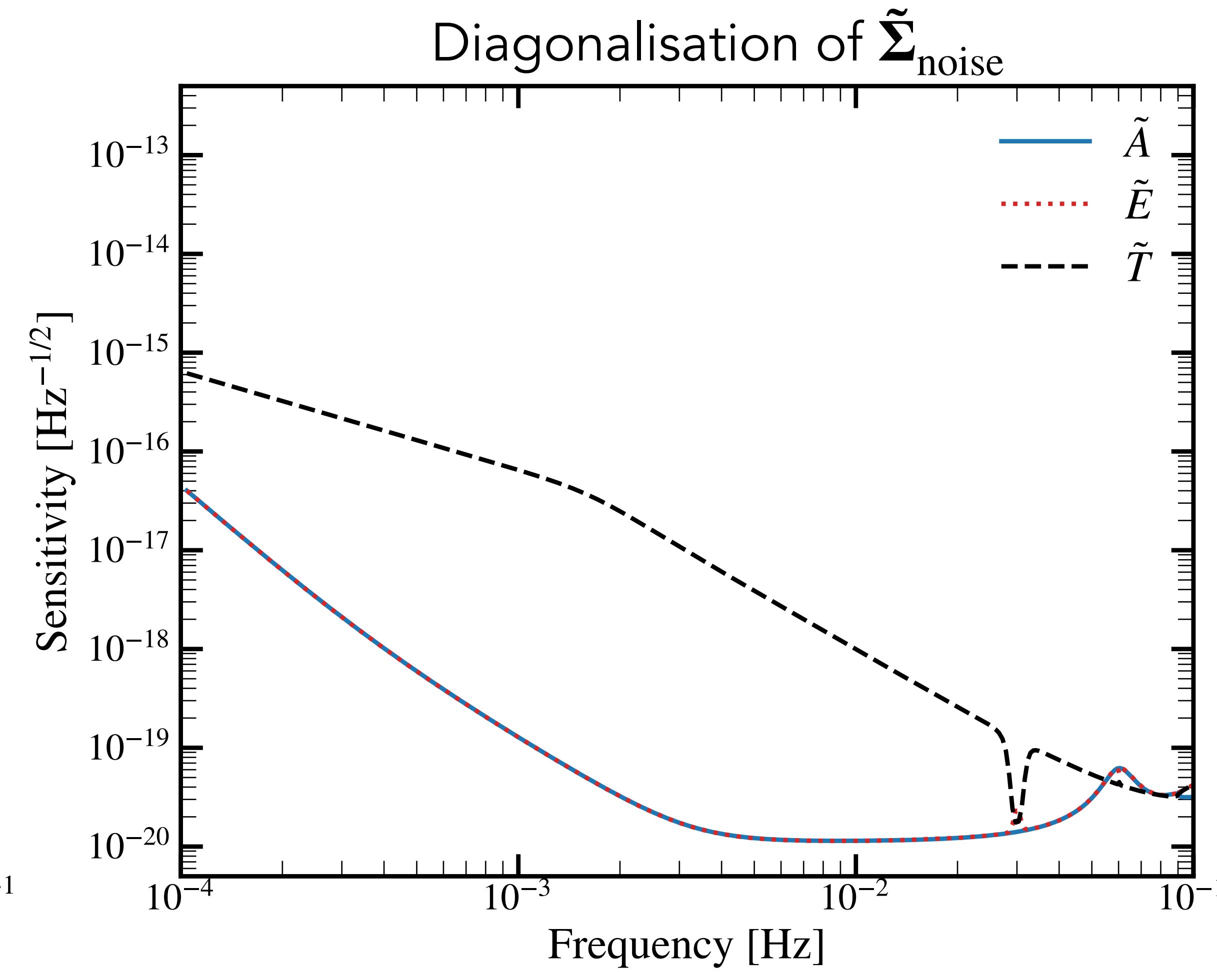
$$\mathbf{B}\mathbf{z} = \tilde{\lambda}\mathbf{z}$$

Where we set $$\mathbf{B} = \mathbf{L}^{-1} \tilde{\mathbf{\Sigma}}_{\mathrm{noise}} \mathbf{L}^{\dagger -1} \qquad \mathbf{z} = \mathbf{L}^\dagger \mathbf{a}$$

With $\mathbf{L}$ the Cholesky decomposition of the SGWB response matrix: $\tilde{\mathbf{R}}_{\mathrm{GW}}(f) = \mathbf{L}\mathbf{L}^\dagger$

- Let us revisit the null channel with this new definition of the optimal SNR:

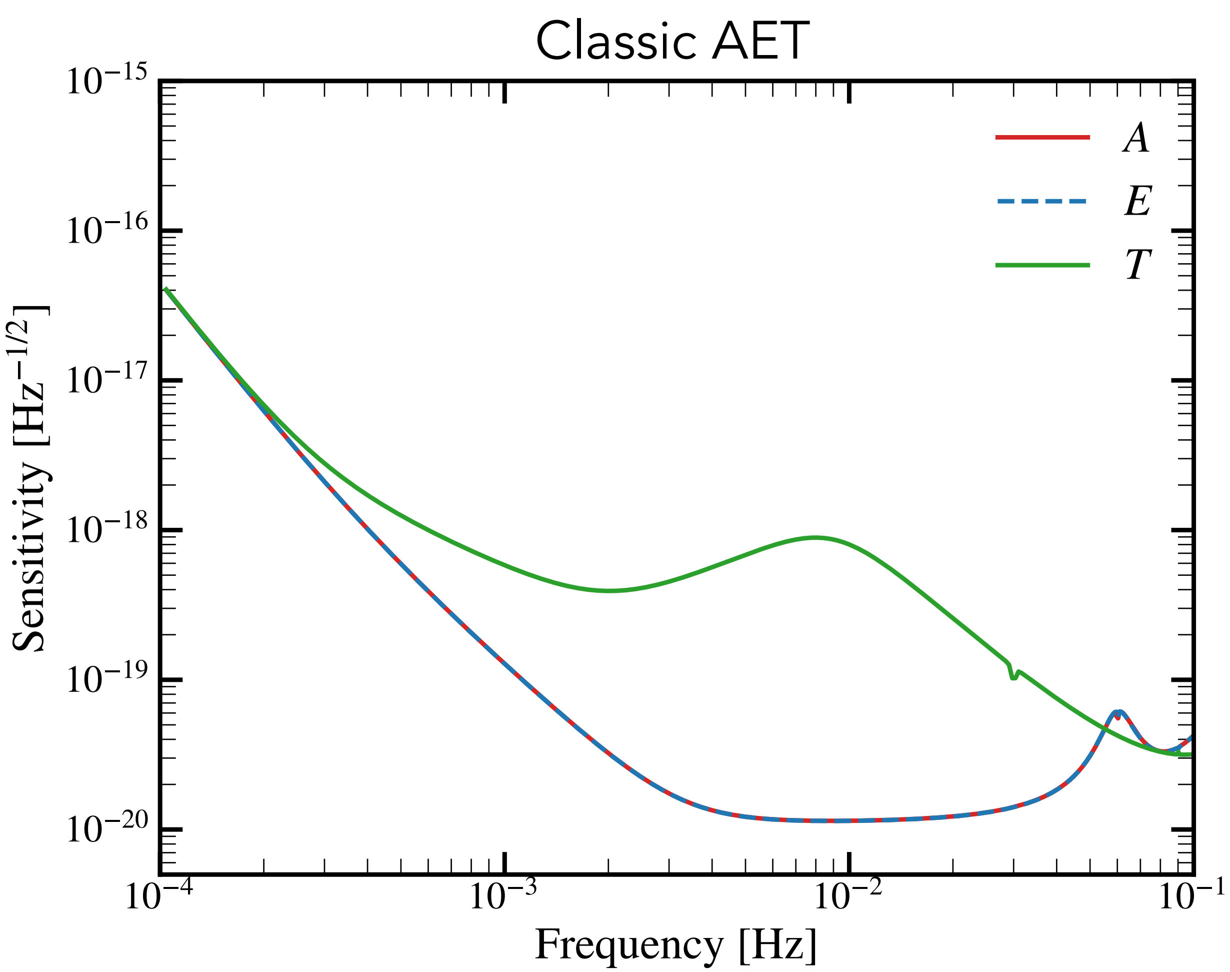


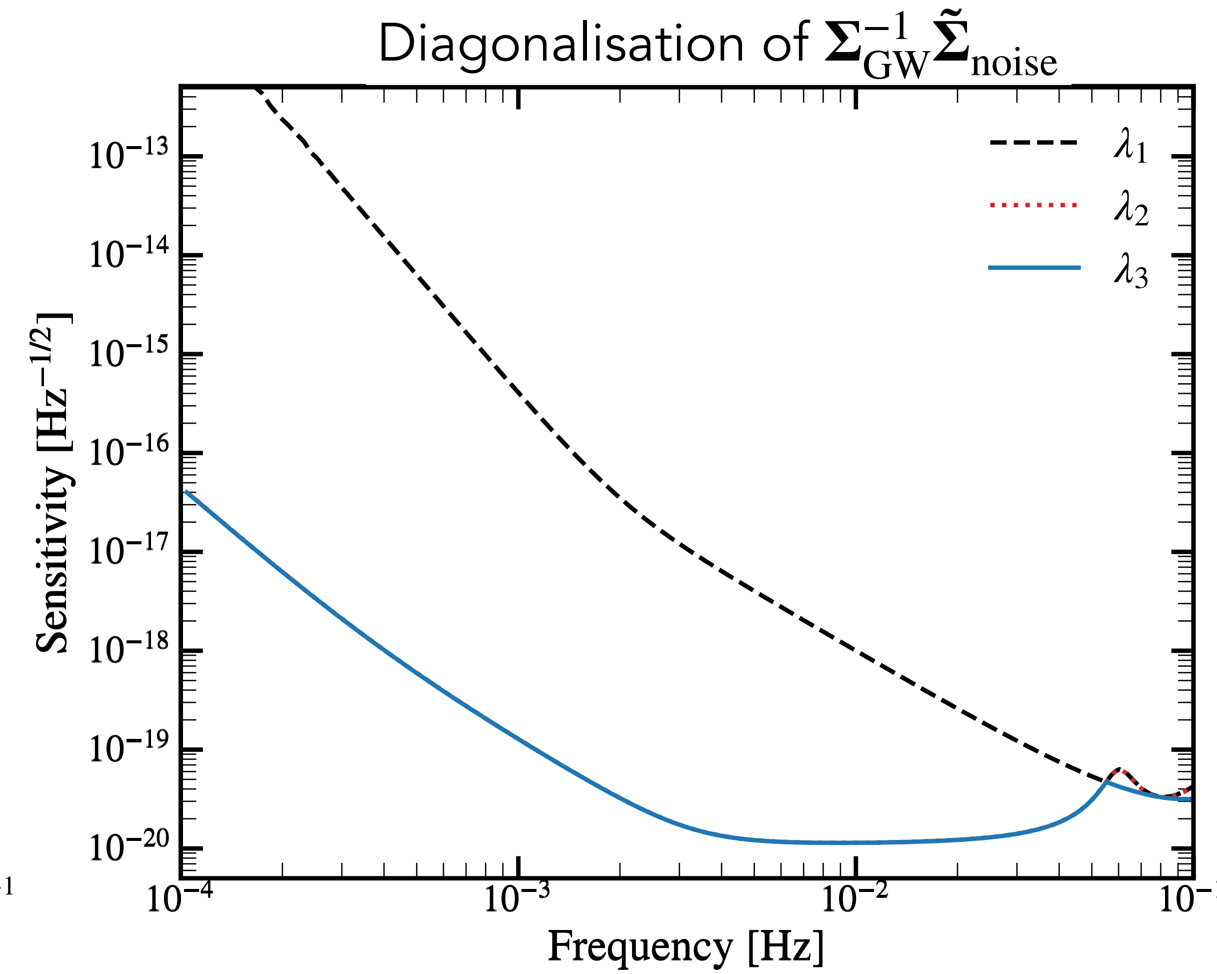


Very similar to ζ channel, see Martina's talk!

But now completely orthogonal channels!

- Let us revisit the null channel with this new definition of the optimal SNR:





Very similar to ζ channel, see Martina's talk!

But now completely orthogonal channels!

Prospects

- **We need a non-parametric spectral estimation for the full TDI** covariance matrix
- There is a **trade-off** to find **between** inserting **instrumental knowledge** (breaking down noise in different components with known transfer functions) **and model complexity** (number of parameters)
- We need to **drop assumptions** one by one and see if SGWB is still distinguishable
- We **need to go to time-frequency domain** to use the time information as a discriminant

Prospects

- **We need a non-parametric spectral estimation for the full TDI** covariance matrix
- There is a **trade-off** to find **between** inserting **instrumental knowledge** (breaking down noise in different components with known transfer functions) **and model complexity** (number of parameters)
- We need to **drop assumptions** one by one and see if SGWB is still distinguishable
- We **need to go to time-frequency domain** to use the time information as a discriminant

Thank you for your attention !