
Detection of gravitational wave background with machine learning

Hugo Einsle, Marie-Anne Bizouard, Mairi
Sakellariadou, Tania Regimbau

October 7-th, IP2I, Lyon

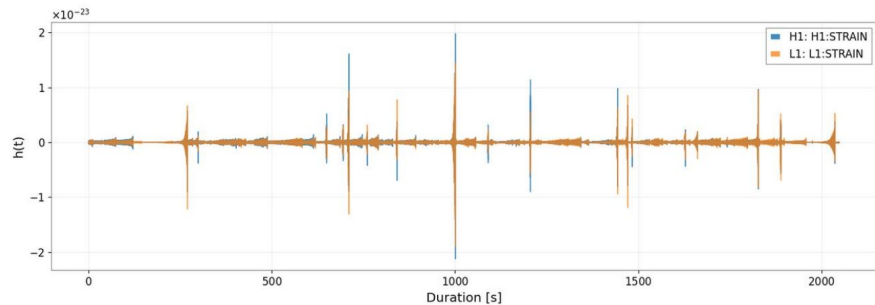
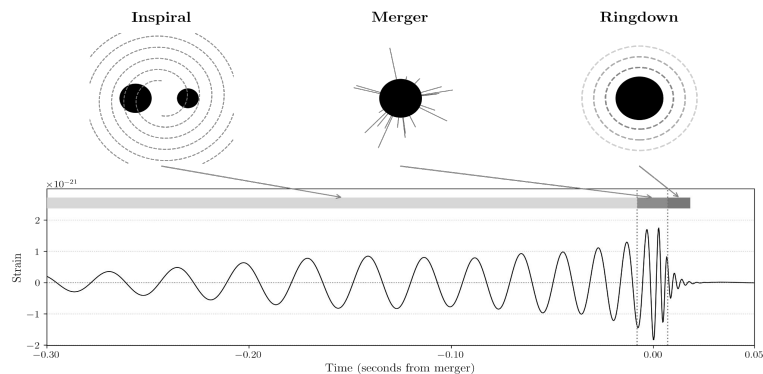


Universiteit
Antwerpen

Gravitational waves and gravitational wave background

Gravitational wave background

- Persistent and incoherent superposition of many sources
- Can be astrophysical (CBCs, magnetars) or cosmological (inflation, cosmic strings)



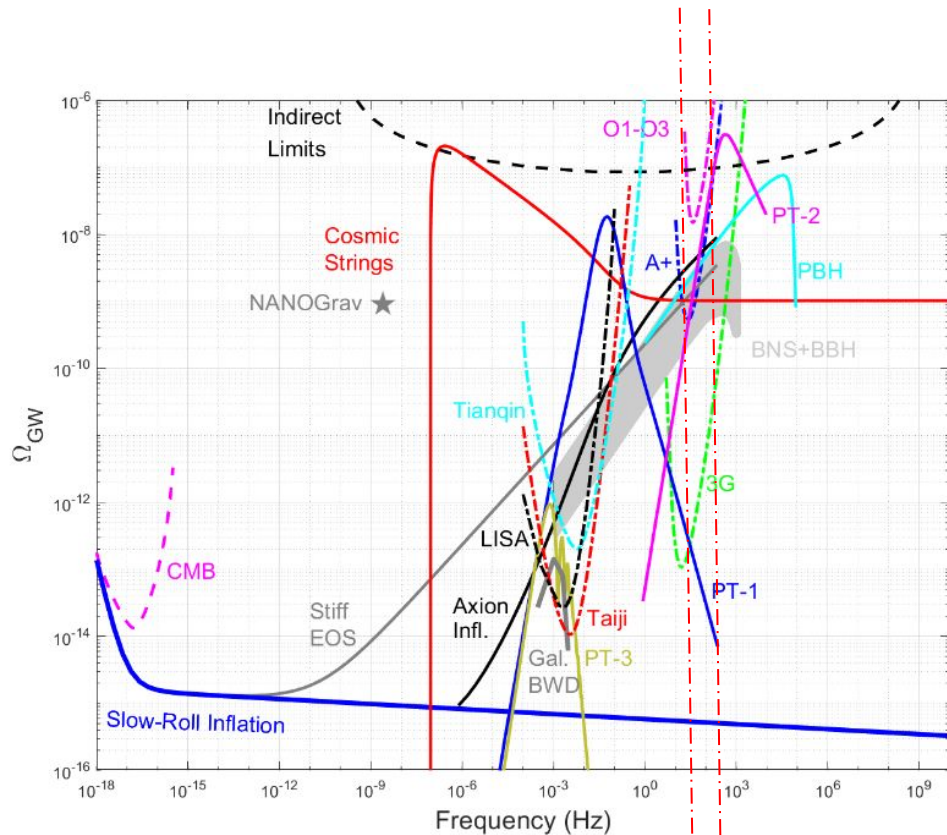
Gravitational wave background

- Ω_{GW} : dimensionless **energy density spectrum** of the stochastic background
- ρ_c : **critical energy density of the Universe**
- Allows to relate strain to **physical energy density**, enabling a statistical search for the SGWB through cross-correlation of detector outputs.

$$\Omega_{\text{GW}}(f) = \frac{1}{\rho_c} \frac{d\rho_{\text{GW}}}{d\ln f}$$

$$\rho_{\text{GW}} = \frac{c^2}{32\pi G} \langle \dot{h}_{ij} \dot{h}^{ij} \rangle$$

$$\rho_c = \frac{3H_0^2}{8\pi G}$$



(Caldwell et al. arXiv:2203.07972)

Problematic and motivations

Optimal method : cross-correlation ([pygwb](#))

- No detection in LVK data so far ([arXiv:2508.20721](#))
- Even once detected, disentangling its astrophysical and other components will remain difficult due to their overlapping spectral shapes, similar statistical properties, contamination from correlated detector noise and weak amplitudes

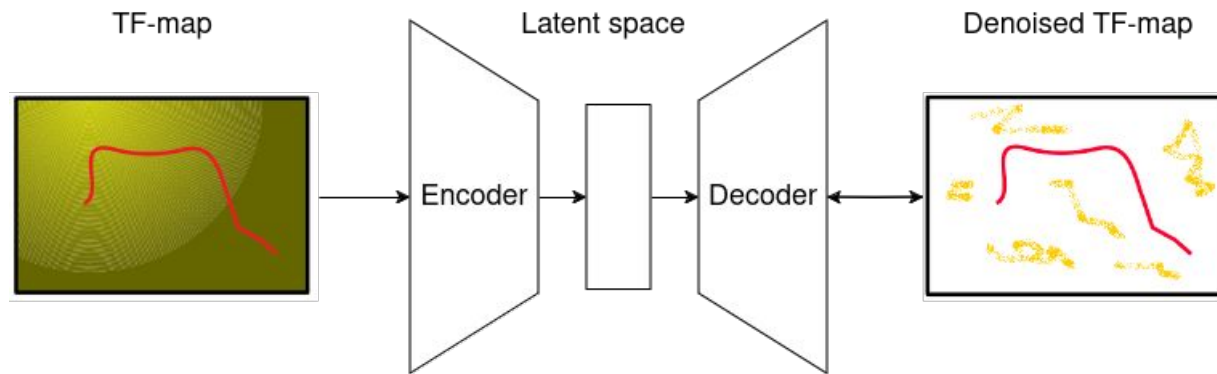
Goal : detecting and disentangling GWB components in LIGO-Virgo data.

Target components: astrophysical (CBC) and flat GWBs.

Proposed solution: hybrid approach combining deep learning and Bayesian inference.

- Custom Multi-Scale Multi-Headed Autoencoder (MSMHAutoencoder) separates noise and GWB.
- Physics-informed training, curriculum learning for low SNR signals.
- MCMC parameter estimation for GWB component inference

Denoising with autoencoders



$$h = f(\tilde{x}) = \sigma_e(W_e \tilde{x} + b_e)$$

$$r = g(h) = \sigma_d(W_d h + b_d)$$

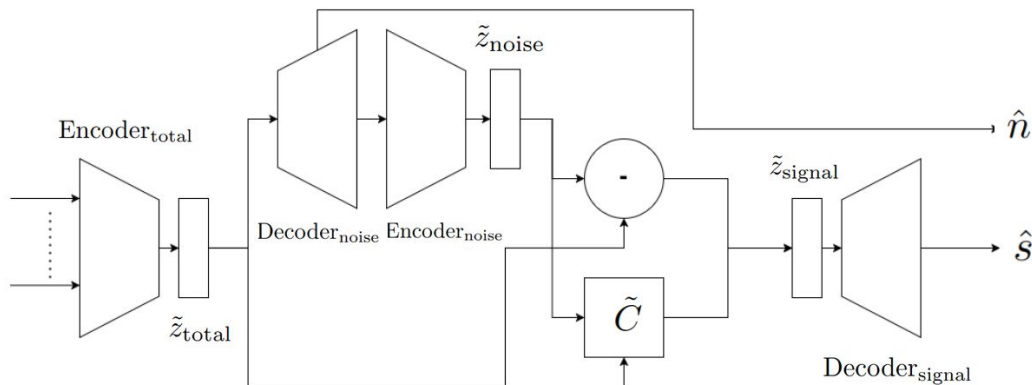
(Schematic representation of the structure of an autoencoder for TF-map denoising)

“Neural Networks for Pattern Recognition”, C. Bishop

“Gradient-Based Learning Applied to Document Recognition”, Y. Lecun, L. Bottou, Y. Bengio, P. Haffner

Gravitational wave background - MSMHAutoencoder

- **Input:** Sequence of M consecutive spectra with N frequency bins each
- **Encoder total:** Compresses the input cross-correlated spectra into a latent representation containing both signal and noise information
- **Noise branch:** The total latent representation is passed through a secondary **Decoder noise** → **Encoder noise** pair to isolate the noise latent representation and reconstructs the estimated noise
- **Latent subtraction:** The estimated noise latent representation is subtracted from the total latent representation
- **Decoder signal :** Reconstructs the cleaned gravitational-wave background from the signal only latent representation



(Einsle et al. Phys. Rev. D **112**, 063056)

Physics-informed loss function

$$\mathcal{L}_{\text{total}} = \lambda_s \mathcal{L}_{\text{spectral}} + \lambda_l \mathcal{L}_{\text{latent}} + \lambda_c \mathcal{L}_{\text{consist}}$$

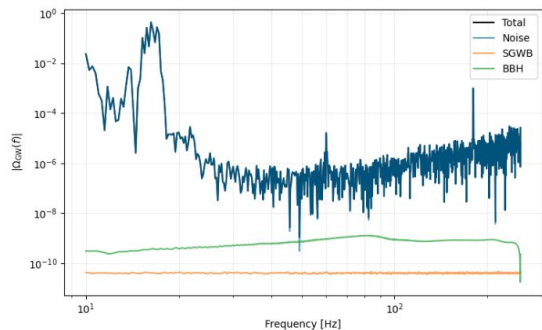
Spectral loss: $\mathcal{L}_{\text{spectral}}$ combines signal and noise reconstruction accuracy with a spectral smoothness penalty

Latent loss: $\mathcal{L}_{\text{latent}}$ ensures internal consistency between predicted and true noise in latent space

Consistency loss: $\mathcal{L}_{\text{consist}}$ imposes self-consistency between reconstructed and original input

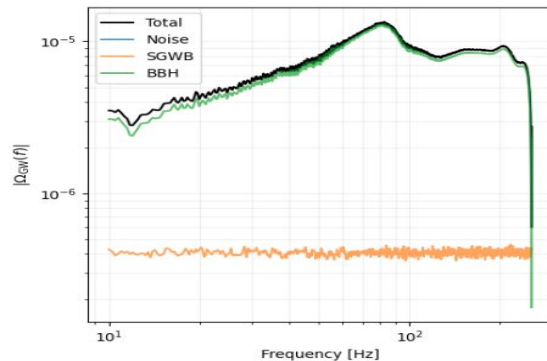
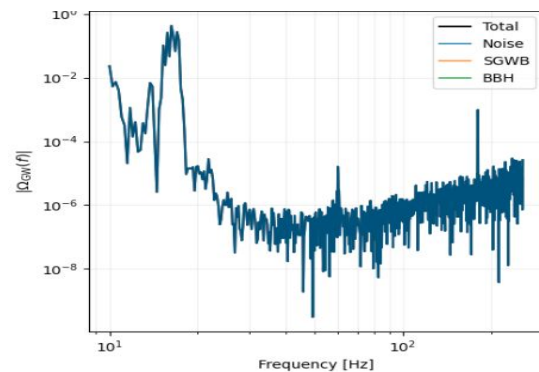
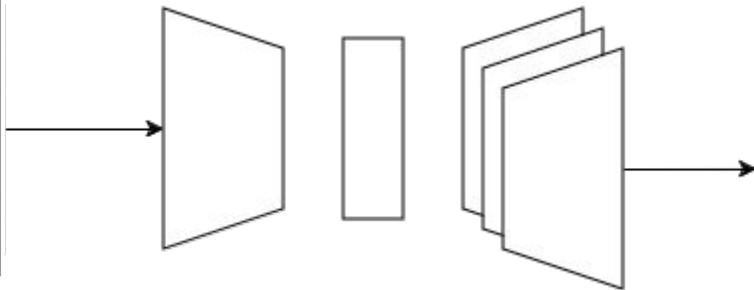
Loss weights: λ_s λ_l λ_c

Gravitational wave background - Denoising



1 input :

- Full spectrum



2 outputs :

- Detector noise
- What remains (signals)

Bayesian inference

Power-law models :

$$\log_{10}(\Omega_{\text{model}}(f|\theta)) = \log_{10}(10^{\log_{10}(\Omega_{\alpha})} \left(\frac{f}{f_{\text{ref}}}\right)^{2/3} + \epsilon)$$
$$\log_{10}(\Omega_{\text{model}}(f|\theta)) = \log_{10}(10^{\log_{10}(\Omega_0)} + 10^{\log_{10}(\Omega_{\alpha})} \left(\frac{f}{f_{\text{ref}}}\right)^{2/3} + \epsilon)$$

Likelihood :

$$\mathcal{L}(\theta) = -\frac{1}{2} \sum_f \left[\frac{(\log_{10}(\hat{s}(f)) - \log_{10}(\Omega_{\text{model}}(f|\theta)))^2}{\sigma_{\log}^2(f)} + \ln(2\pi\sigma_{\log}^2(f)) \right]$$

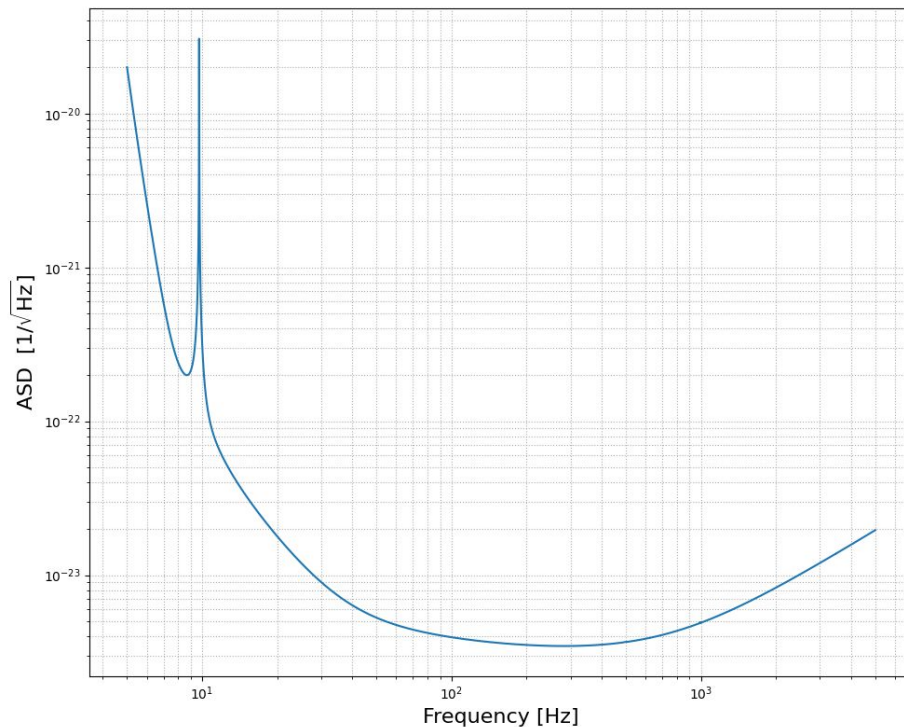
Bayes factor for best model evidence :

$$\mathcal{B} = \mathcal{Z}_{\text{CBC}} / \mathcal{Z}_{\text{Noise}}$$
$$\mathcal{B}_{\text{cosmo}} = \mathcal{Z}_{\text{CBC}+\text{cosmo}} / \mathcal{Z}_{\text{CBC}}$$

A+ design datasets

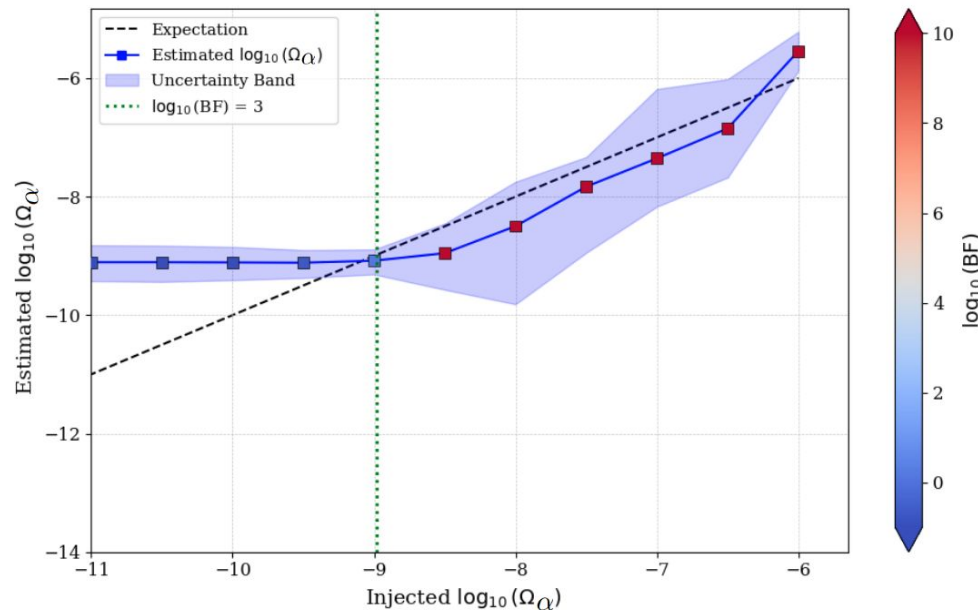
Datasets:

- Segments of 2048 s sampled at 512 Hz.
- Training and validation sets: 2000 segments equivalent to 47.4 days of data ; mixture of pure noise, astrophysical and flat GWB injections.
- Testing dataset: 1000 segments equivalent to 23.7 days of data; mixture of pure noise, astrophysical and flat GWB injections.



Bayesian inference : only CBC GWB

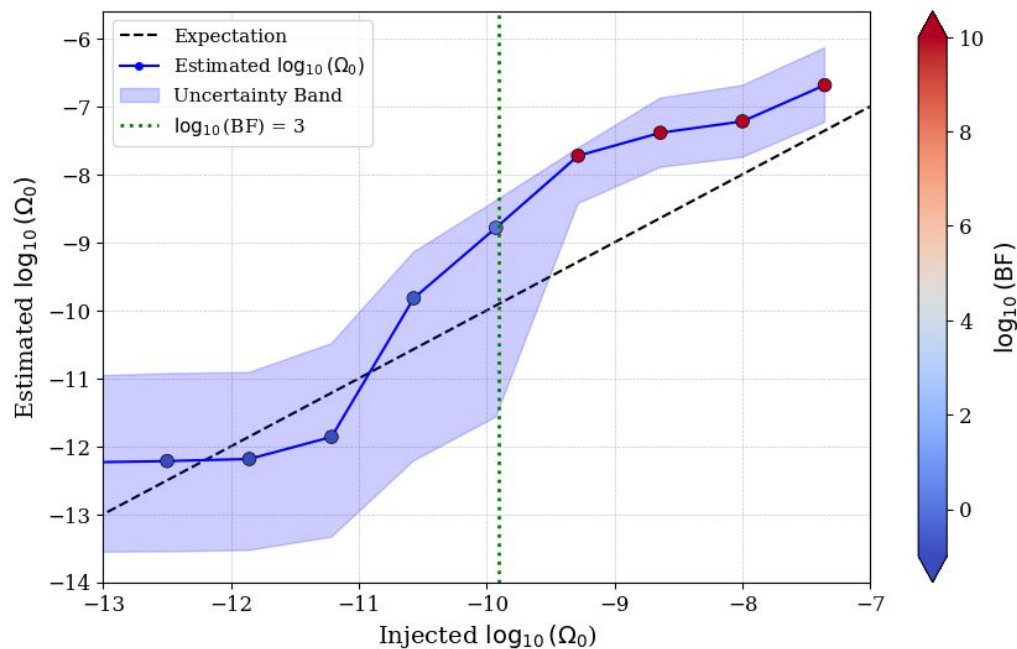
- We inject and retrieve a BBH-only GWB with varying amplitudes
- **Detection regime:** The estimated $\log_{10}(\Omega_\alpha)$ closely follows the injected values down to $\Omega_\alpha \approx 10^{-9}$
- **Evidence threshold:** Below this level the Bayes factor drops sharply, indicating the model cannot distinguish signal from noise.
- **Noise-limited regime:** For lower amplitudes, the estimated Ω_α plateaus with small uncertainty, reflecting the model fitting residual noise rather than true signal content.



(Einsle et al. Phys. Rev. D **112**, 063056)

Bayesian inference : CBC and cosmological GWB

- We inject and retrieve a flat GWB with varying amplitudes in the presence of a BBH GWB such that $\Omega_{\text{BBH}} = 10^{-9}$
- **Detection regime:** The model successfully recovers the injected flat background amplitude down to $\Omega_0 \approx 1.3 \times 10^{-10}$
- **Consistency with injections :** For higher amplitudes the estimated values of $\log_{10}(\Omega_0)$ align well with the injected ones
- **Noise-limited regime:** Below $\Omega_0 \approx 10^{-10}$ the credible intervals widen by up to two orders of magnitude and the estimated $\log_{10}(\Omega_0)$ plateaus

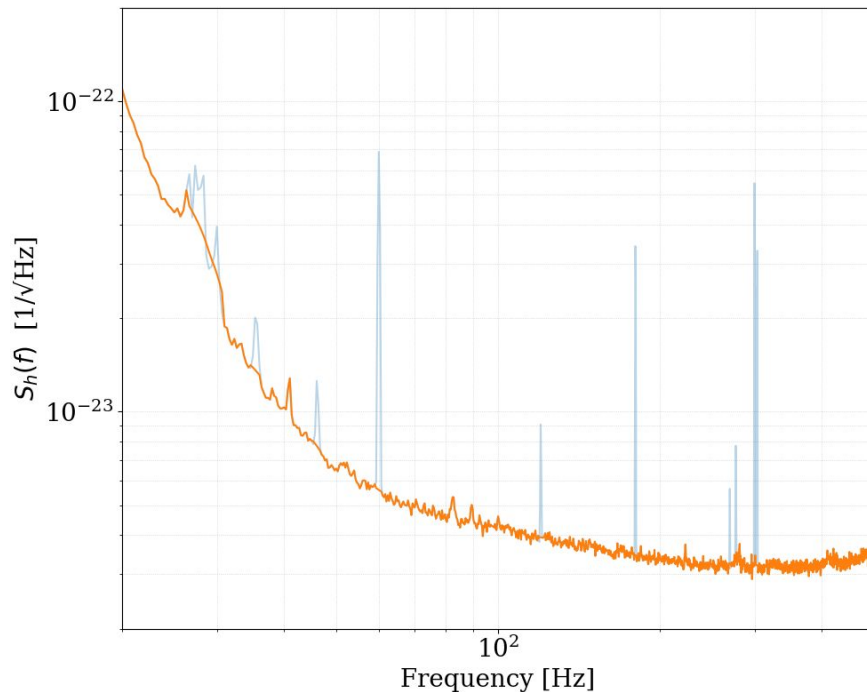


(Einsle et al. Phys. Rev. D **112**, 063056)

O4a-like datasets

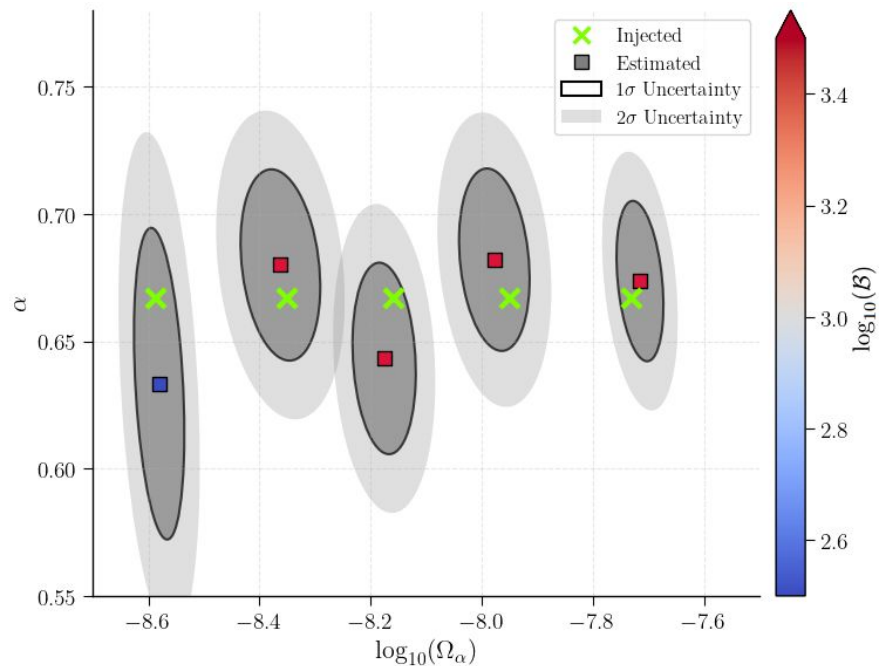
Datasets:

- Segments of 2048 s sampled at 512 Hz.
- One dataset with spectral lines removed and one dataset where they remain
- Training and validation sets: 108 days of data ; mixture of pure noise, astrophysical and flat GWB injections.
- Testing dataset: 108 days of data; mixture of pure noise, astrophysical and flat GWB injections.

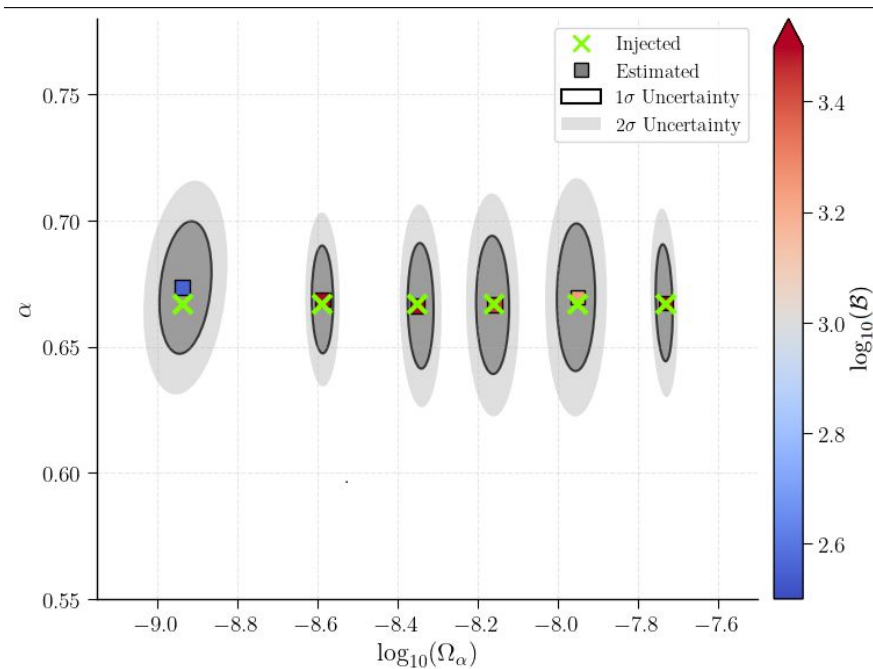


Bayesian inference : only CBC GWB

With spectral lines

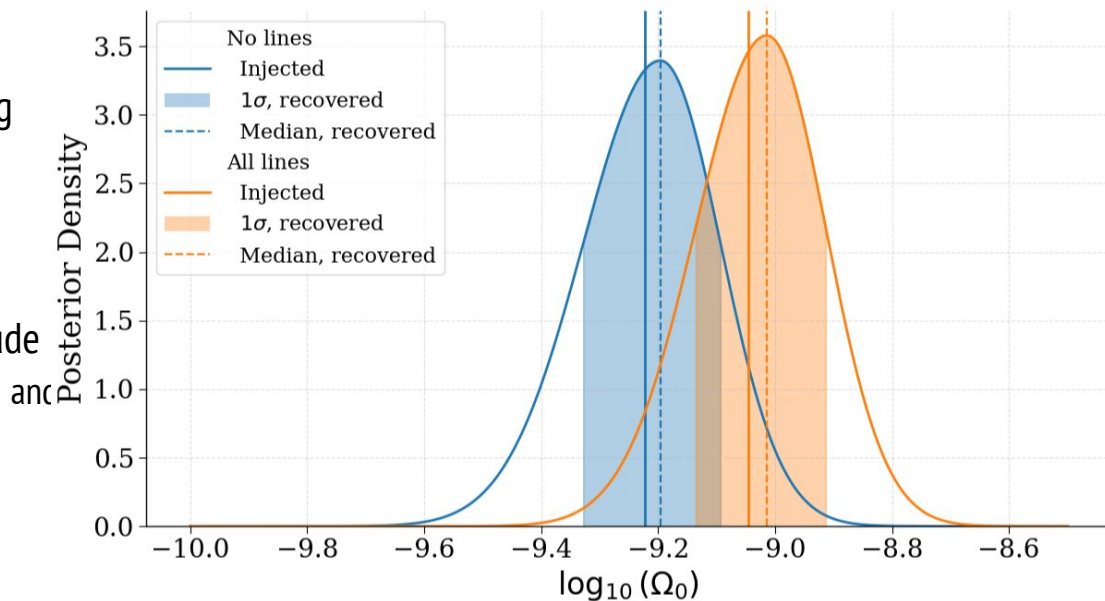


Without spectral lines



Bayesian inference : CBC and cosmological GWB

- We inject and retrieve a flat GWB with varying amplitudes in the presence of a CBC GWB at respective detection threshold
- **Detection regime:** The model successfully recovers the injected flat background amplitude down to $\Omega_0 \approx 6.4 \times 10^{-10}$ without spectral lines and $\approx 9.7 \Omega_0^{-10}$ with spectral lines



Blind test comparison with pygwb

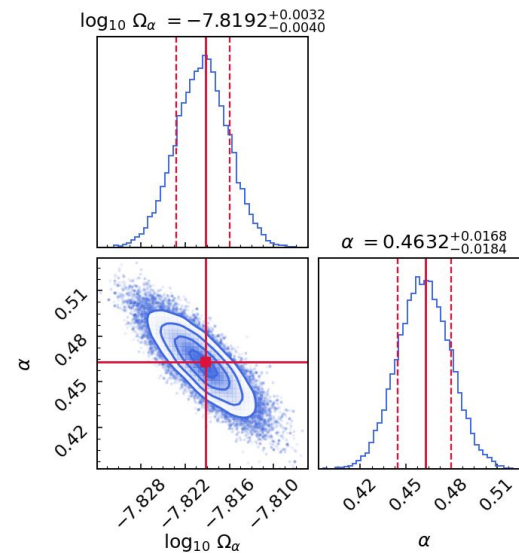
Dataset	Injected		
	$\Omega_{\text{CBC}}^{\text{inj}}$	$\Omega_{\text{cosmo}}^{\text{inj}}$	$\Omega_{\text{tot}}^{\text{inj}}$
1	10.0	4.0	14.0
2	3.0	0.8	3.8
3	3.0	1.0	4.0
4	1.0	0.0	1.0
5	10.0	0.8	10.8

Each dataset is generated with either a CBC GWB with a $\frac{2}{3}$ spectral index, a cosmological GWB with a 0 spectral index, or a combination of both. Each dataset represents 108 of simulated data representative of O4a data after data quality cuts. All amplitudes are quoted in units of $1\text{e-}9$ at $\text{fref} = 25 \text{ Hz}$.

Dataset 1 - DeepGWB

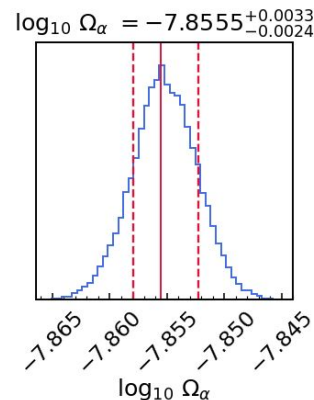
CBC (alpha free) search :

$\log_{10}(\text{BF}) = 31$
 $\log_{10}(\Omega_{\alpha}) = -7.82 (+0.003/-0.004)$
 $\Omega_{\alpha} = 1.52\text{e-}08 (+1.13\text{e-}10/-1.38\text{e-}10)$
 $\alpha = 0.46 (+0.02/-0.02)$



CBC (alpha fixed at 2/3) search :

$\log_{10}(\text{BF}) = 33$
 $\log_{10}(\Omega_{\alpha}) = -7.86 (+0.003/-0.002)$
 $\Omega_{\alpha} = 1.40\text{e-}08 (+1.05\text{e-}10/-7.72\text{e-}11)$



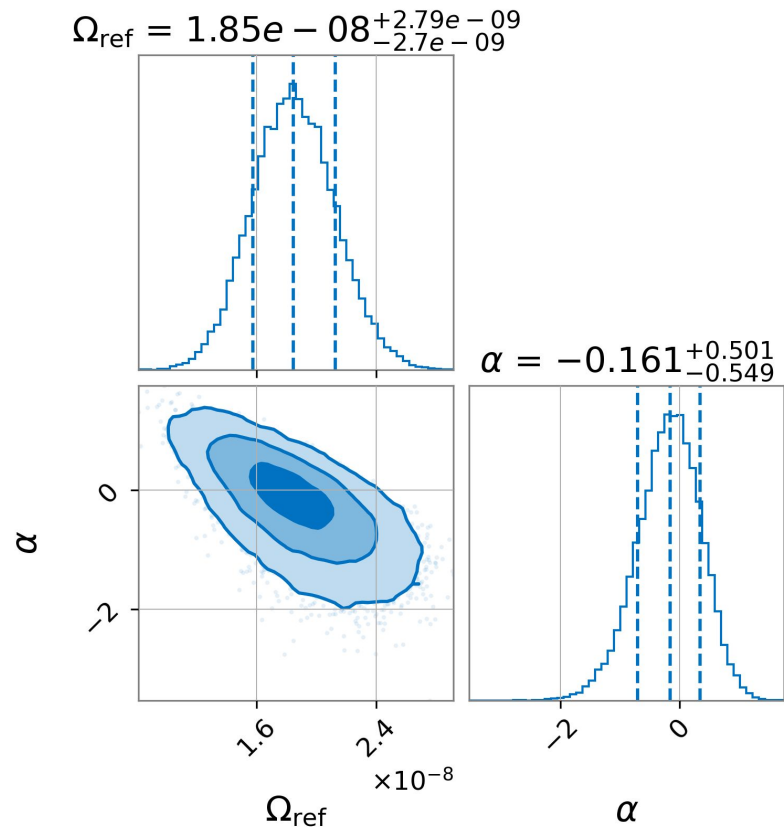
Dataset 1 - PyGWB

CBC (alpha free) search :

log10(BF) = 39.5
Omega_alpha = 1.85e-08 (+2.79e-9/-2.7e-9)
alpha = -0.161 (+0.501/-0.549)

CBC (alpha fixed at $\frac{2}{3}$) search :

SNR = 9.40
Omega_alpha = 1.52e-08 (+1.62e-9/-1.62e-9)

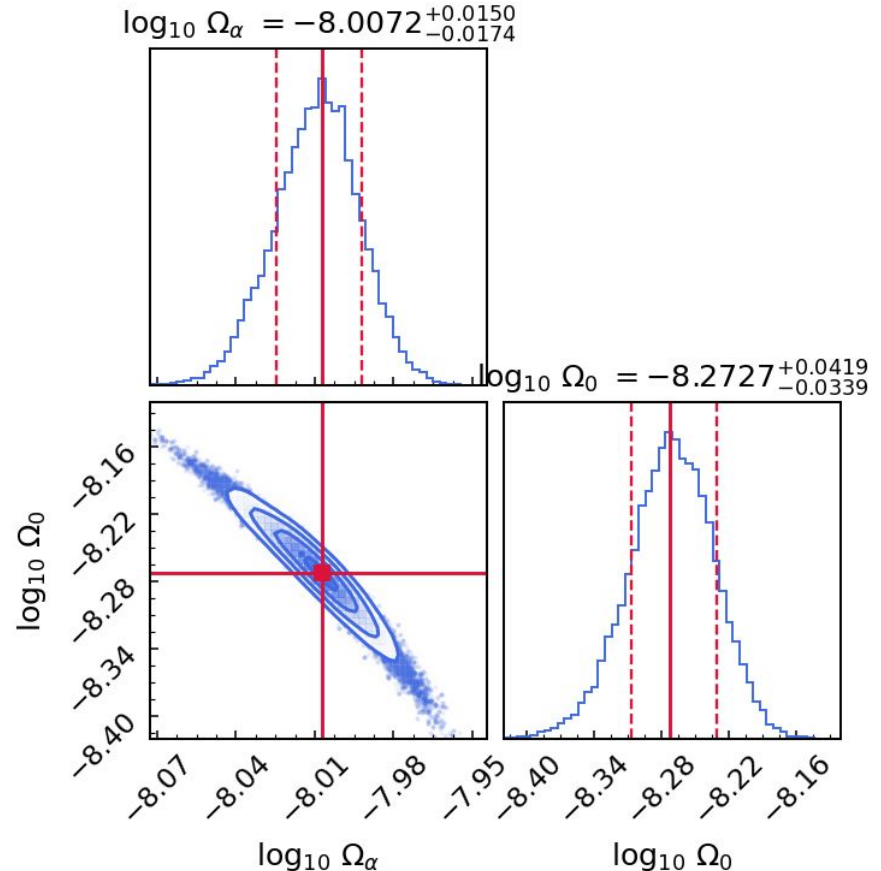


Dataset 1 - DeepGWB

CBC (alpha = 2/3)+Cosmo (alpha =0) search:

$\log_{10}(\text{BF}_{\text{cosmo}})$ = 7
 $\log_{10}(\Omega_{\alpha})$ = -8.0 (+0.02/-0.02)
 Ω_{α} = $9.84\text{e-}09$ (+3.45e-10/-3.85e-10)
 $\log_{10}(\Omega_0)$ = -8.27 (+0.04/-0.03)
 Ω_0 = $5.34\text{e-}09$ (+5.40e-10/-4.01e-10)

DeepGWB is able to disentangle the components of the background !



Results - Power-law and CBC models searches

	Injection			DEEPGWB		pygwb		DEEPGWB			pygwb		
	$\Omega_{\alpha=2/3}^{\text{inj}}$	Ω_0^{inj}	$\Omega_{\text{tot}}^{\text{inj}}$	$\Omega_{\alpha=2/3}$	$\log_{10}(\mathcal{B})$	$\Omega_{\alpha=2/3}$	$\log_{10}(\mathcal{B})$	Ω_{α}	α	$\log_{10}(\mathcal{B})$	Ω_{α}	α	$\log_{10}(\mathcal{B})$
1	10.0	4.0	14.0	$14.0^{+0.01}_{-0.01}$	33	$15.05^{+1.62}_{-1.62}$	17.49	$15.20^{+0.11}_{-0.14}$	$0.46^{+0.02}_{-0.02}$	31	$19^{+2.8}_{-2.7}$	$-0.16^{+0.50}_{-0.55}$	7.46
2	3.0	0.8	3.8	$3.83^{+0.04}_{-0.04}$	14	$3.93^{+1.85}_{-2.39}$	0.84	$4.17^{+0.07}_{-0.07}$	$0.45^{+0.04}_{-0.03}$	12	$6.1^{+3.2}_{-3.9}$	$-1.7^{+1.90}_{-1.92}$	0.37
3	3.0	1.0	4.0	$4.00^{+0.04}_{-0.04}$	15	$4.20^{+1.81}_{-2.20}$	0.96	$4.39^{+0.08}_{-0.06}$	$0.43^{+0.03}_{-0.03}$	13	$6.5^{+3.2}_{-3.6}$	$-1.6^{+1.85}_{-1.92}$	0.42
4	1.0	-	1.0	$1.01^{+0.02}_{-0.02}$	4.3	$0.05^{+1.27}_{-0.05}$	-0.05	$1.08^{+0.04}_{-0.04}$	$0.49^{+0.06}_{-0.08}$	3.1	$0.05^{+241}_{-0.05}$	$-1.1^{+3.49}_{-2.71}$	-0.02
5	10.0	0.8	10.8	$11.3^{+0.01}_{-0.01}$	73	$12.03^{+1.63}_{-1.63}$	10.80	$11.70^{+0.11}_{-0.11}$	$0.57^{+0.02}_{-0.02}$	71	$15^{+2.8}_{-2.8}$	$-0.19^{+0.63}_{-0.70}$	4.52

(All amplitudes are quoted in units of $1\text{e-}9$ at $\text{fref} = 25\text{ Hz.}$)

DeepGWB outperforms PyGWB giving a more precise estimate of both the amplitude and the spectral index while being sensitive in regimes where PyGWB does not achieve a significant detection

Results - CBC + flat cosmological model search

	Injection			DEEPGWB						
	$\Omega_{\alpha=2/3}^{\text{inj}}$	Ω_0^{inj}	$\Omega_{\text{tot}}^{\text{inj}}$	$\Omega_{\alpha=2/3}$	Ω_0	$\log_{10}(\mathcal{B}_{\text{cosmo}})$	Ω_{α}	α	Ω_0	$\log_{10}(\mathcal{B}_{\text{cosmo}})$
1	10.0	4.0	14.0	$9.84^{+0.35}_{-0.39}$	$5.34^{+0.54}_{-0.40}$	7.0	$9.78^{+0.45}_{-0.38}$	$0.66^{+0.01}_{-0.01}$	$5.40^{+0.54}_{-0.50}$	6.0
2	3.0	0.8	3.8	$2.61^{+0.24}_{-0.15}$	$1.58^{+0.25}_{-0.26}$	6.2	$2.66^{+0.22}_{-0.19}$	$0.66^{+0.01}_{-0.01}$	$1.51^{+0.31}_{-0.23}$	5.7
3	3.0	1.0	4.0	$2.62^{+0.22}_{-0.18}$	$1.81^{+0.27}_{-0.27}$	5.3	$2.67^{+0.22}_{-0.20}$	$0.66^{+0.01}_{-0.01}$	$1.72^{+0.31}_{-0.22}$	4.7
4	1.0	-	1.0	$0.76^{+0.06}_{-0.05}$	$0.33^{+0.07}_{-0.06}$	2.9	$0.78^{+0.06}_{-0.05}$	$0.64^{+0.04}_{-0.05}$	$0.32^{+0.08}_{-0.06}$	2.8
5	10.0	0.8	10.8	$9.84^{+0.24}_{-0.24}$	$1.84^{+0.35}_{-0.26}$	8.2	$10.00^{+0.39}_{-0.32}$	$0.65^{+0.03}_{-0.03}$	$1.68^{+0.41}_{-0.32}$	7.3

(All amplitudes are quoted in units of $1\text{e-}9$ at $\text{fref} = 25\text{ Hz.}$)

DeepGWB is able to separate the components from the injected GWB. However the CBC background is systematically underestimated while the cosmological background is overestimated

Conclusion

- First hybrid framework combining **deep learning** with **Bayesian inference** to detect and disentangle the GWB components
- While cross-correlation remains optimal asymptotically, the MSMHAutoencoder **reaches equivalent sensitivity much faster**, providing an efficient alternative for near-term searches
- **Complement, not replacement**: ML/DL doesn't replace physical models but rather tackles problems that have become too large, complex, or computationally expensive for standard methods alone.
- **Detector context**: increasing sensitivity → more data, more complex noise, weaker signals to search for
- **Accessing information**: ML/DL helps extract information that would otherwise be hard to reach
- **Concrete benefits**: faster searches, better denoising, detection of weak or poorly modelled signals beyond the reach of classical approaches

Thank you for listening !