Accelerating binary neutron star template generation with machine learning

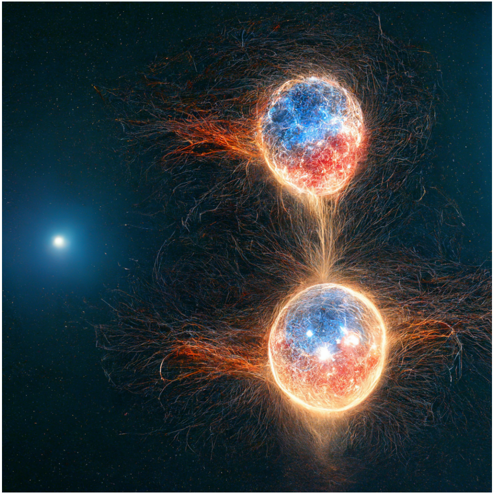
`g2net next challenges`

Jacopo Tissino [1]
with Gregorio Carullo,[2] Matteo Breschi,[2] Rossella Gamba,[2] Stefano Schmidt,[3] Sebastiano Bernuzzi[2]

[1]Gran Sasso Science Institute, [2]Jena University, [3]Utrecht University

EGO, 2022-09-29

A neutron star merger (AI artist's rendition)



The science of neutron star mergers

Measuring the properties of neutron star mergers allows us to explore:

- equations of state of dense, possibly exotic matter;
- strong-field GR, as well as the speed of GWs;
- multi-messenger astronomy: EM follow-up, neutrinos;
- neutron star populations;
- cosmology (BNS mergers are independent standard sirens);
- r-process nucleosynthesis.

GW parameter estimation

The likelihood used in parameter estimation reads:

$$\Lambda(s|\theta) \propto \exp\left((h_\theta|s) - \frac{1}{2}(h_\theta|h_\theta) \right) , \tag{1}$$

where $(a|b)$ is the Wiener product:

$$(a|b) = 4\,\mathrm{Re} \int_0^\infty \frac{\widetilde{a}^*(f)\widetilde{b}(f)}{S_n(f)}\,\mathrm{d}f \ . \tag{2}$$

Typical number of evaluations of the likelihood required: $\gtrsim 10^7$.
Parameter estimation takes days!

Gravitational waveform approximants

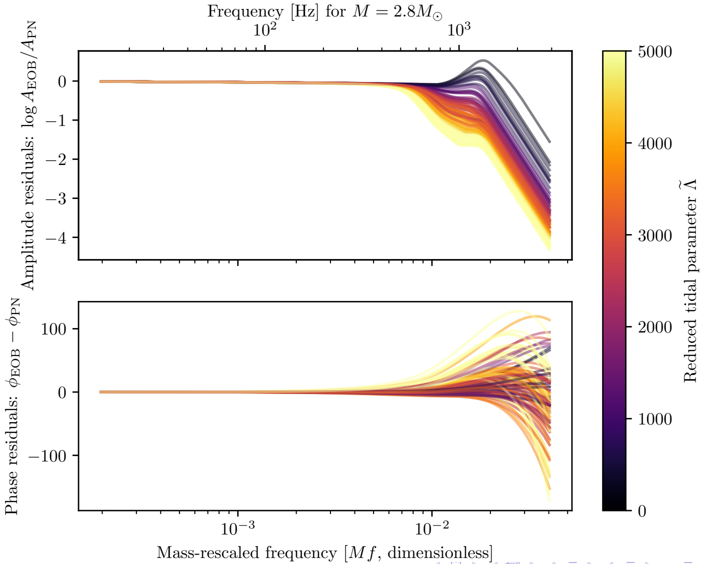
How do we model the emission from a compact binary coalescence?

- **Post Newtonian** (analytic, not accurate up to merger);
- **Effective One Body** (need to solve an ODE, accurate up to merger);
- **Numerical Relativity** (very expensive, our reference point).

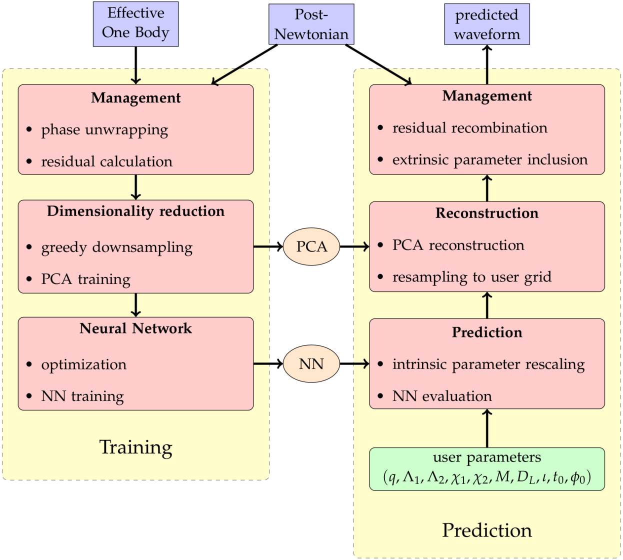
What `mlgw_bns` is

A **surrogate approximant** for BNS,
working in the **frequency domain**,
trained on the EOB approximant `TEOBResumS`.

Residuals from Post-Newtonian waveforms



mlgw_bns structure



Evaluating reconstruction fidelity

Relevant distance measure:

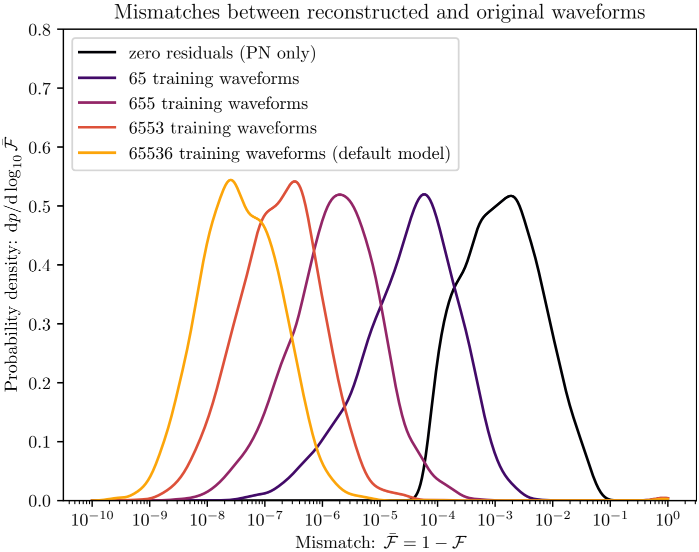
$$\overline{\mathcal{F}}(a, b) = 1 - \mathcal{F}(a, b) = 1 - \frac{\max_{\varphi_0, t_c}(a|b)}{\sqrt{(a|a)(b|b)}}\,. \tag{3}$$

The accuracy requirement depends on the SNR:

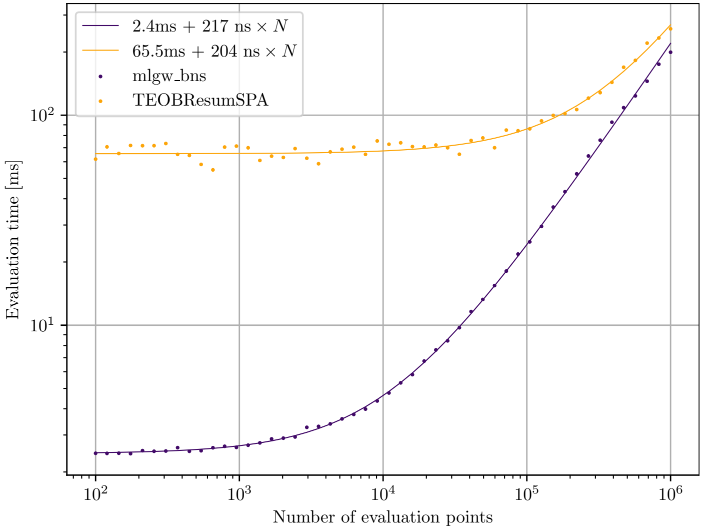
$$\overline{\mathcal{F}} \lesssim \frac{n}{\mathrm{SNR}^2}\,, \tag{4}$$

where n is the number of parameters.

Fidelity results



Evaluation time



Reduced Order Quadratures

GW170817 uniform grid: $N_{\text{points}} = (f_{\text{max}} - f_{\text{min}})\, T_{\text{signal}} \approx 2 \times 10^5$.
With ROQ:

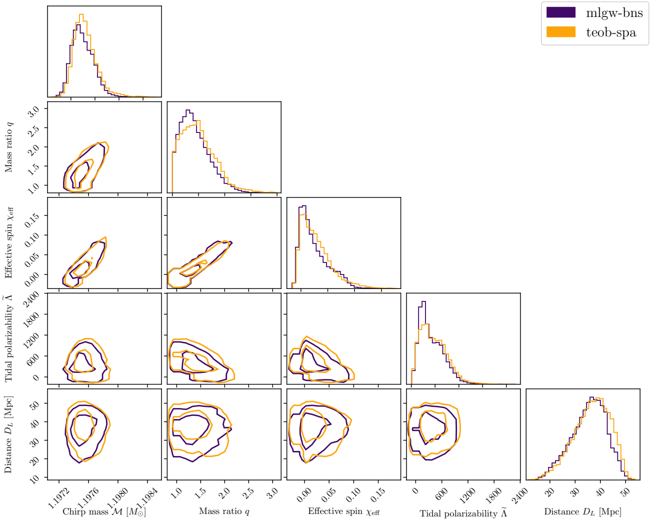
$$(d, h_\theta)_{\text{ROQ}} = \sum_{i=1}^{N_L} \omega_j(t_c) h_\theta(f_i; \theta) \,, \tag{5}$$

and similarly for the (h_θ, h_θ) term with N_Q points.
For GW170817, $N_Q + N_L \approx 300$.

Full parameter estimation timings

- `TEOBResumS` + No ROQ $\sim$ 57h 30m;
- `TEOBResumS` + ROQ $\gtrsim$ 40h (not measured);
- `mlgw_bns` + No ROQ: $\sim$ 49h 20m;
- `mlgw_bns` + ROQ: $\sim$ 4h 20m.

A posterior distribution: GW170817



Get mlgw-bns

`pip install mlgw-bns`

Documentation is available at `mlgw-bns.readthedocs.io`.

Backup slides

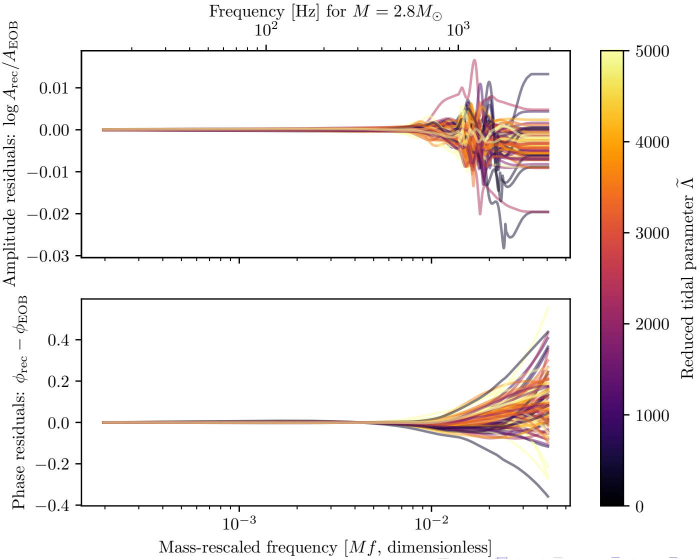
Backup slides!

Technologies

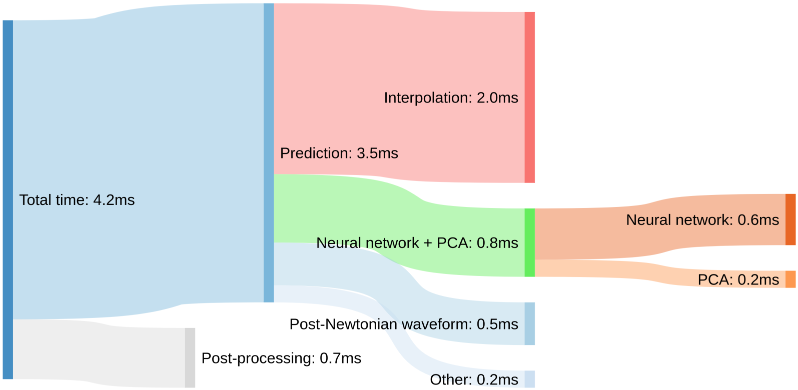
mlgw_bns is implemented as a python package, and it makes use of

- scikit-learn for the neural network (upgrading to pytorch);
- optuna for the hyperparameter optimization;
- pytest and tox for automated testing;
- numba for just-in-time compilation and acceleration.

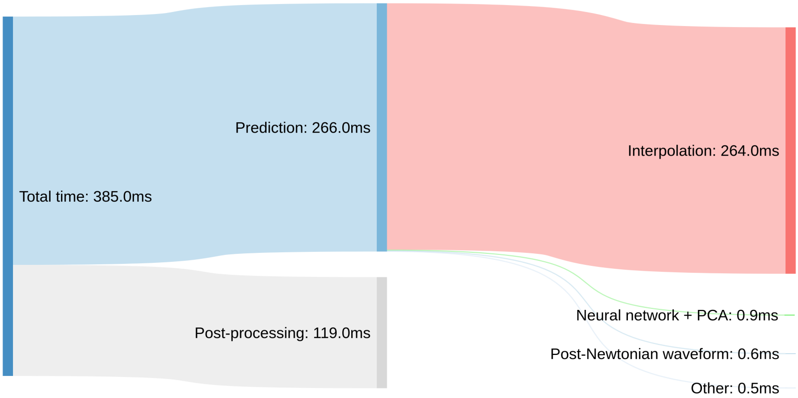
Reconstruction residuals



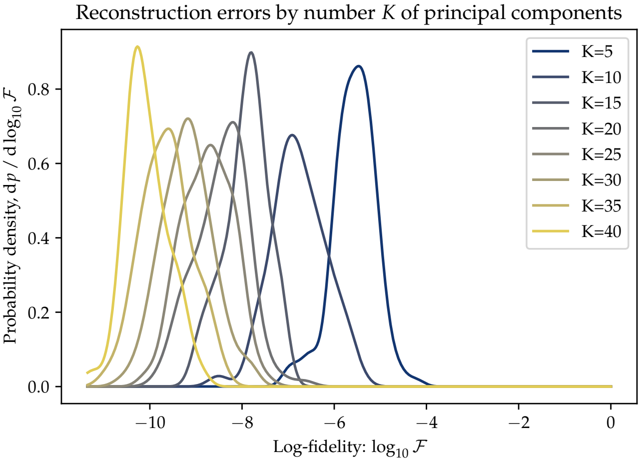
Profiling the evaluation: 8×10^3 interpolation points



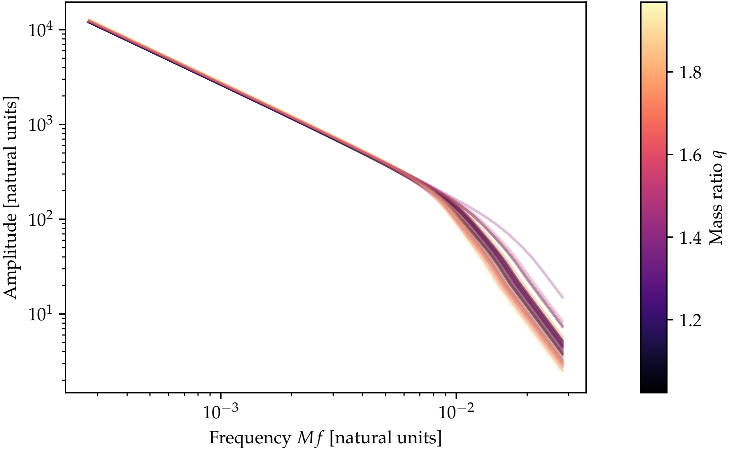
Profiling the evaluation: 2×10^6 interpolation points



PCA-only reconstruction error

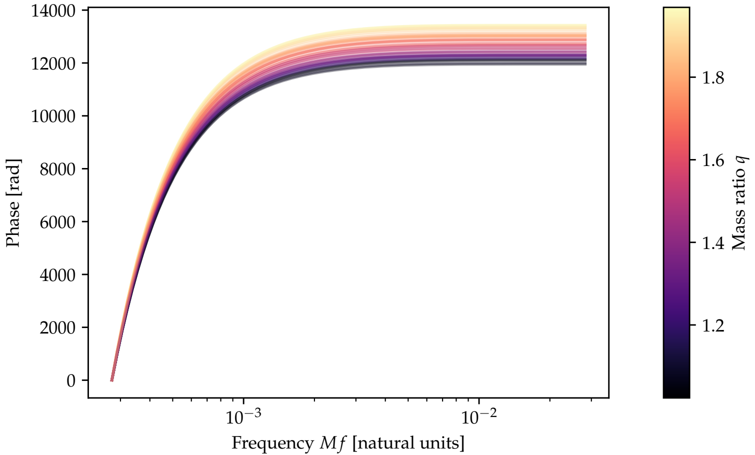


Amplitudes



10^4
10^3
10^2
10^1
Amplitude [natural units]
10^{-3}
10^{-2}
Frequency Mf [natural units]
1.8
1.6
1.4
1.2
Mass ratio q

Phases



14000
12000
10000
8000
6000
4000
2000
0
Phase [rad]
10^{-3}
10^{-2}
Frequency Mf [natural units]
1.8
1.6
1.4
1.2
Mass ratio q

Signal parameters

We do (currently) consider:

- intrinsic parameters: (source-frame) total mass $M = m_1 + m_2$, mass ratio $q = m_1 / m_2$, aligned spins χ_{1z} and χ_{2z}, tidal polarizabilities Λ_1 and Λ_2;
- extrinsic parameters: luminosity distance D_L, inclination ι, initial phase ϕ_0, coalescence time t_c;
- considered but not in the waveform: sky position (ra and dec), polarization angle ψ, redshift z.

We do not consider: precessing spins, eccentricity (both decent approximations for BNS).

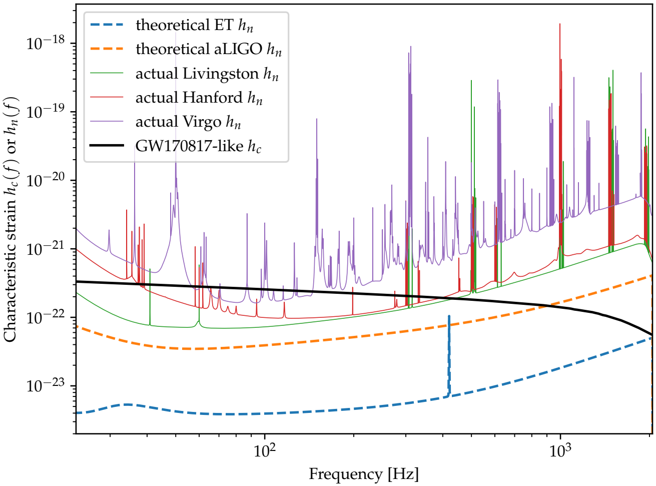
Hyperparameters optimized

- PC exponent α: the network reconstructs $\mathrm{PC}_i \lambda_i^\alpha$, where λ_i is the eingenvalue and PC_i is its eigenvector;
- layer sizes: 2 to 4 layers between 10 and 200 nodes;
- activation function and L2 regularization parameter;
- training parameters: initial learning rate, tolerance, batch size, stopping criterion;
- number of waveforms available for training.

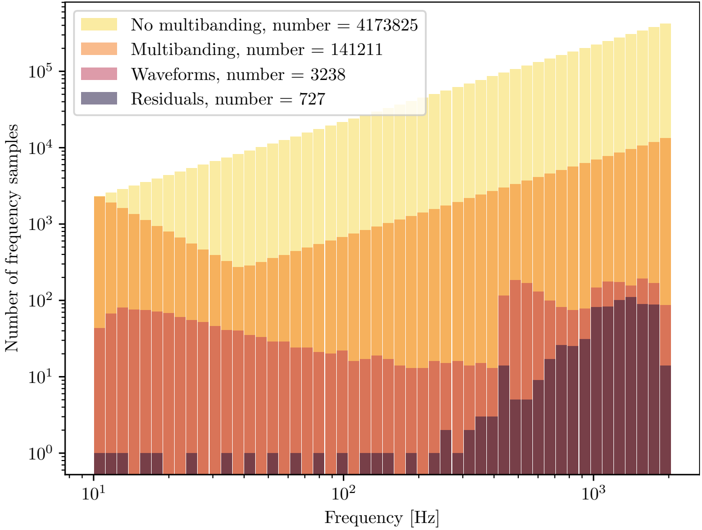
are optimized by a MOTPE (Multi-objective tree-structured Parzen Estimator) of:

- (L2) reconstruction error;
- network training time.

Power Spectral densities and GW170817



Frequency grid histograms



Hyperparameter optimization