SEOBNRv5PHM_NNSur: a fast neural network waveform surrogate for generically precessing binaries

Chris Whittall & Geraint Pratten

UNIVERSITY OF BIRMINGHAM | GRAVITATIONAL WAVE ASTRONOMY

Introduction: SEOBNRv5PHM_NNSur7dq10

- We built a reduced order time-domain surrogate model for SEOBNRv5PHM, using neural networks for parameter space fits.
- SEOBNRv5PHM: state-of-the-art EOB model for quasicircular precessing binaries, including higher order modes. [Ramos-Buades+ (2303.18046),...]
- Covers mass ratios up to 1:10, arbitrary spin magnitudes and directions.
- Covers period 10,000M before merger through to 100M after.

Waveform decomposition

- Work in co-orbital frame:

$$\phi_{\rm orb}(t) := \frac{1}{2}\mathrm{Arg}\left[h_{22}^{\rm CP}(t))\right]$$

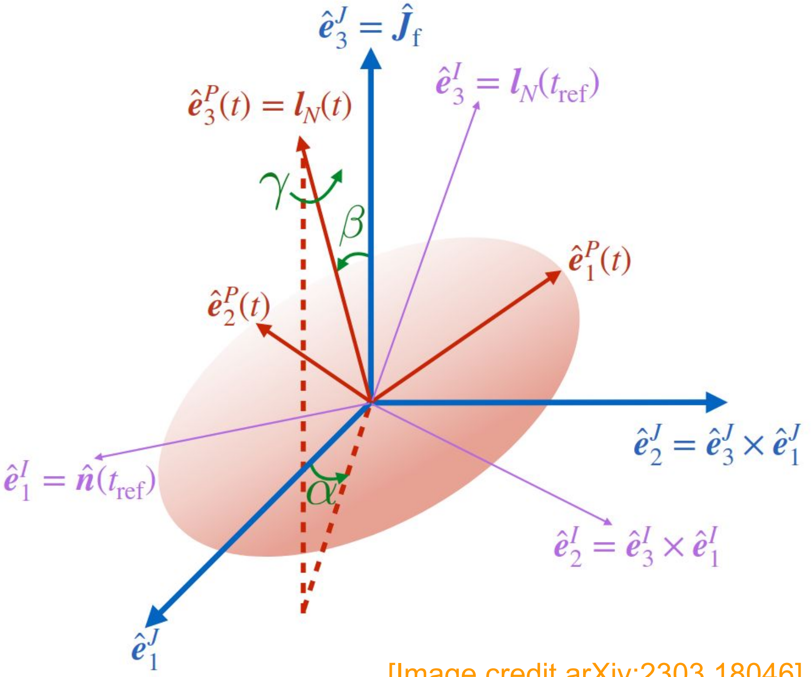
$$h_{\ell m}^{\rm CO}(t) = h_{\ell m}^{\rm CP}(t)\, e^{-im\phi_{\rm orb}(t)}$$

- Rotation from P-frame to inertial frame described by quaternions:

$$\mathbf{q} = \cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right)\hat{\mathbf{n}}$$

- Separate quaternions for P → J_f and J_f → I rotations.

- Not currently including P-frame mode asymmetries.



[Image credit arXiv:2303.18046]

We build separate surrogates for the orbital phase, co-orb frame modes and quaternions q_{J2P} (t) and q_{I2J}

Model structure

For each time-series data piece f(t; λ), we…

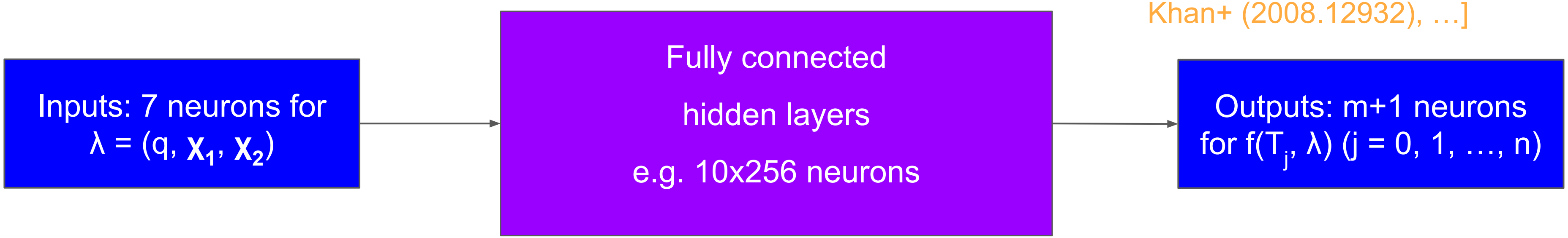
- Use greedy algorithms to find the reduced basis + empirical interpolation matrix $B_i(t)$:

$$f(t,\vec{\lambda}) \approx \mathrm{EI}[f](t,\vec{\lambda}) := \sum_{j=0}^{n} B_j(t) f(T_j,\vec{\lambda}),$$

[Field+ (1308.3565) and many others….]

- Train a neural network to predict the values of $\{f(T_j, \lambda): j = 1, 2, \dots, n\}$ at any given value of λ in the training domain:

[Chua+ (1811.05491), Khan+ (2008.12932), …]



- We use 200k training waveforms to build the reduced basis, and train the networks on 1.2M waveforms (2.2M for the orbital phase)

- Single network to predict the time-independent quaternion q_{I2J}

Quaternion downsampling

- Transformation from P-frame to I-frame, e.g.: [Boyle+ (1409.4431)]

$$h_+^I(t;\iota,\phi_0) - i h_\times^I(t;\iota,\phi_0) = \sum_{\ell,m} \sqrt{\frac{2\ell+1}{4\pi}} \mathcal{D}^{\ell}_{m,2}\left(q_f(t;\iota,\phi_0)\right) h^P_{\ell m}(t),$$

- Expensive to evaluate the Wigner D-matrices at every point on the dense time grid: bottleneck!

- Precession timescale longer than orbital period:
  - Evaluate the quaternions on a sparse time grid (piecewise uniform)
  - Evaluate the D-matrices at the sparse time points
  - Interpolate the D-matrices onto the final waveform time grid

- Saves significant amounts of time in some cases.

Model validity: total mass limits

- Waveform polarizations returned on a fixed geometric time grid $[-10^4M, 100M]$ with spacing $\Delta t = 0.5M$.

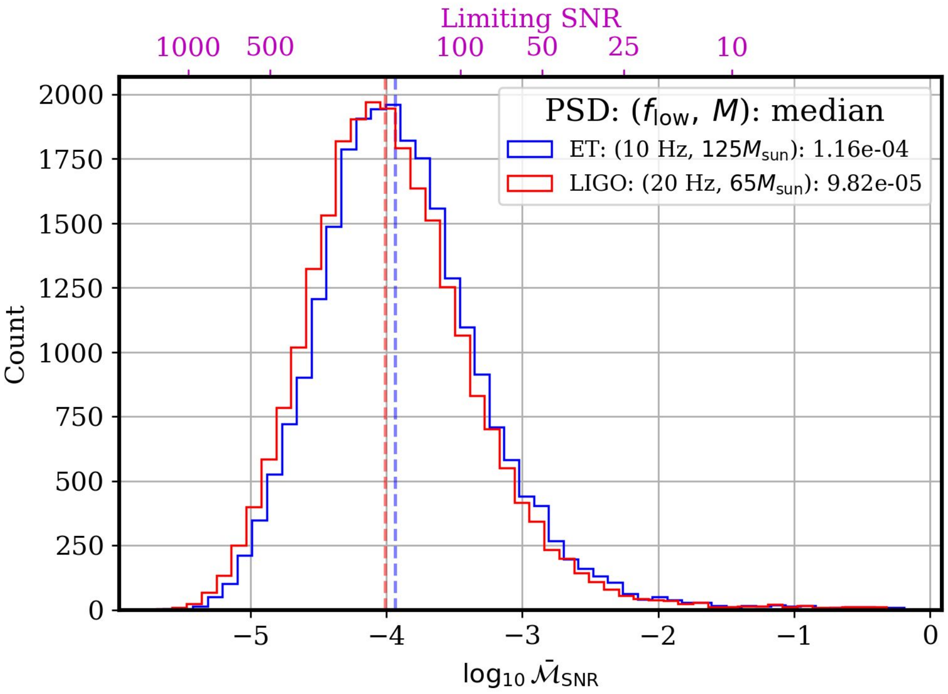
- Starting frequency depends on the total mass.

- Require total mass

$(q \leq 10)$	$(q \leq 4)$	$(q = 1)$
$\mathbf{M > 62.3M_{sun}}$	$\mathbf{M > 48.3M_{sun}}$	$\mathbf{M > 41.1M_{sun}}$

to ensure starting frequency < 20 Hz for all spin configurations.

Results: faithfulness to SEOBNRv5PHM



- Mismatches with SNR-weighted average over inclination, phase and polarisation.

$$\bar{\mathcal{M}}_{\rm SNR} := 1 - \left(\frac{\sum_i \mathcal{M}_i^3 \rho_i^3}{\sum_i \rho_i^3} \right)^{1/3}$$

- Orbital phase model is greatest source of error, followed by the quaternion models.

Results: computational cost

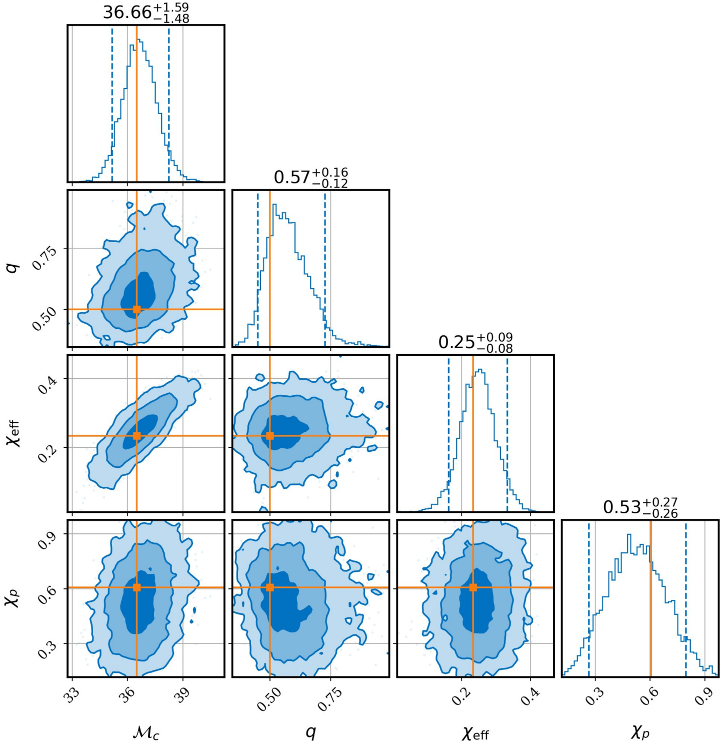
Mean cost per waveform (wall time/batch size)

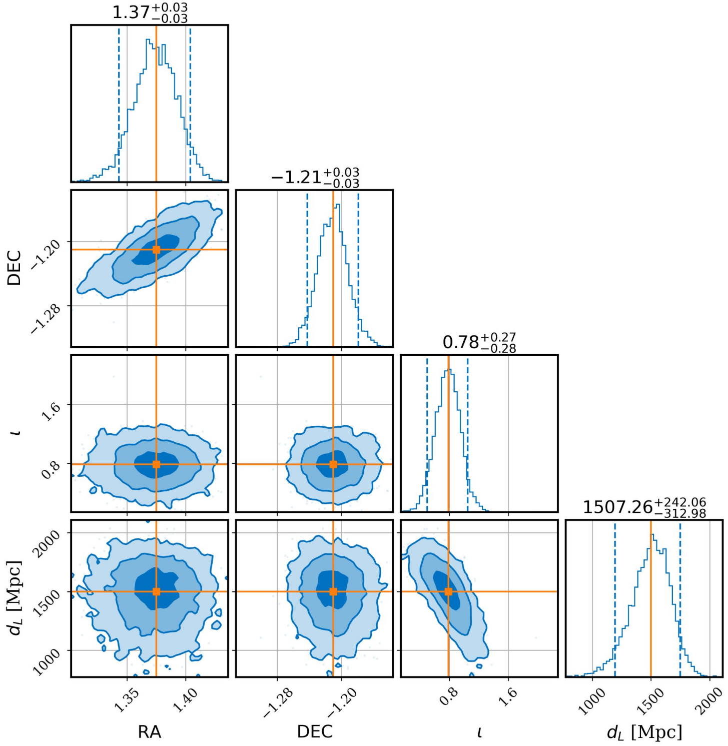
	With DS	Without DS	Best-case speedup[3]
Laptop CPU[1] [single wf]	**12.52ms**	15.47ms	5.2x
Laptop GPU[2] [single wf]	13.79ms	**13.38ms**	4.9x
Laptop GPU[2] [250x batch]	**0.93ms**	1.82ms	71x
Nvidia A100 [1500 batch]	**0.16ms**	0.25ms	410x

[1] Single-thread on Intel Core Ultra 7 155H
[2] Nvidia RTX 1000 Ada
[3] Average SEOB cost = 65.6ms per waveform

Results: injection-recovery

Inject SEOBNRv5PHM into zero noise, recover with the surrogate

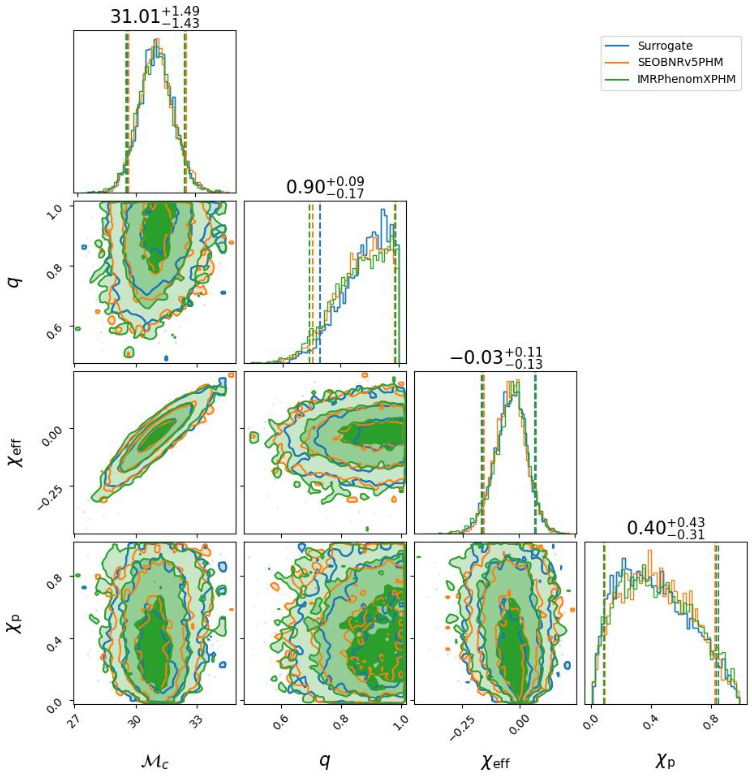


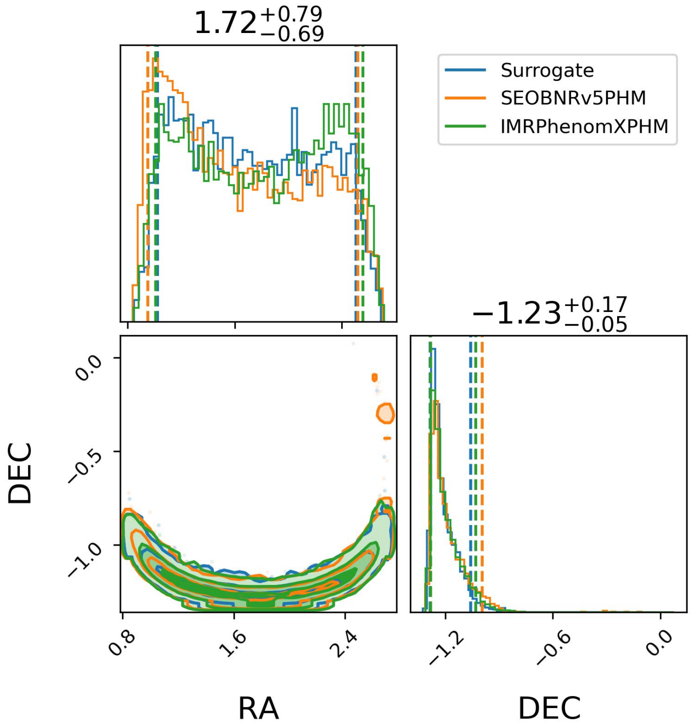


	Injected
$\mathbf{m_1}$	60.0
$\mathbf{m_2}$	30.0
$\mathbf{a_1}$	0.7
$\boldsymbol{\theta_1}$	π/3
$\mathbf{a_2}$	0.0

Network (HLV) optimal SNR ~ 25.7

Results: GW150914





	Sur	SEOB	XPHM
Wall time	22hr 56m	63hr 49m	9hr 10m

8s segment, nlive=1000, npool=32

~ 2.8x speedup compared to SEOBNRv5PHM

Conclusions

- Demonstrated neural network surrogate approach for 7d precessing problem.
- Median model mismatches ~ 1e-4 against base SEOBNRv5PHM.
- Speedup ~5x for single waveform on a laptop CPU or GPU
- Surrogate achieves best (average) performance when evaluating large waveform batches on the GPU.
- Demonstrated practical application of surrogate to PE, with significant speedup over SEOBNRv5PHM.

Thank you for listening