AI and extreme scale computing to learn and infer the physics of higher order gravitational wave modes of quasi-circular, spinning, non-precessing binary black hole mergers

Asad Khan\textsuperscript{a,b,c}, E. A. Huerta\textsuperscript{a,b,d}

\textsuperscript{a}Data Science and Learning Division, Argonne National Laboratory, Lemont, Illinois 60439, USA
\textsuperscript{b}Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
\textsuperscript{c}National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
\textsuperscript{d}Department of Computer Science, University of Chicago, Chicago, Illinois 60637, USA

Abstract

We use artificial intelligence (AI) to learn and infer the physics of higher order gravitational wave modes of quasi-circular, spinning, non-precessing binary black hole mergers. We trained AI models using 14 million waveforms, produced with the surrogate model \texttt{NRHbSur3dq8}, that include modes up to $\ell \leq 4$ and $(5, 5)$, except for $(4, 0)$ and $(4, 1)$, that describe binaries with mass-ratios $q \leq 8$ and individual spins $s_{z1,2} \in [-0.8, 0.8]$. We use our AI models to obtain deterministic and probabilistic estimates of the mass-ratio, individual spins, effective spin, and inclination angle of numerical relativity waveforms that describe such signal manifold. Our studies indicate that AI provides informative estimates for these physical parameters. This work marks the first time AI is capable of characterizing this high-dimensional signal manifold. Our AI models were trained within 3.4 hours using distributed training on 256 nodes (1,536 NVIDIA V100 GPUs) in the Summit supercomputer.

Keywords: AI, High Performance Computing, Higher Order Modes, Black Hole Mergers

1. Introduction

The development of rigorous, reproducible, statistically and domain informed artificial intelligence (AI) is leading to remarkable breakthroughs in science and engineering \cite{1}, and powering human intelligence to enable significant advances in pure mathematics \cite{2}. AI applications in gravitational wave astrophysics are evolving at a breakneck speed. It is only four years ago that we proposed that AI may provide a scalable, computationally efficient method to search for and
find gravitational waves [3, 4]. This approach has been embraced and further developed by an international community of researchers [5–7]. To this date, AI has been explored for a variety of signal processing tasks, including detection [8–24], gravitational wave denoising and data cleaning [25–27], parameter estimation [28–34], rapid waveform production [35, 36], waveform forecasting [37, 38], and early warning systems for multi-messenger sources [39–41], among others.

Production scale AI frameworks for gravitational wave detection that harness high performance computing (HPC) and scientific data infrastructure have also been developed [42], furnishing evidence for the scalability, reproducibility and computational efficiency of AI-driven methodologies. Figure 1 provides a glimpse of the rapid convergence of AI and extreme scale computing to study astrophysical scenarios that require waveforms with richer and more complex morphology. In view of these developments, it is time to further push the frontiers of AI applications to quantify their suitability to describe high-dimensional signal manifolds that describe waveforms whose morphology is significantly richer and much more complex than what has already been explored in the literature. This article represents a step in that direction. The driver we have selected for this study consists of characterizing higher order gravitational wave modes that describe quasi-circular, spinning, non-precessing binary black hole mergers. We densely sample a parameter space that consists of binaries with mass-ratios \( q \leq 8 \) and individual spins \( s_{\{1,2\}} \in [-0.8, 0.8] \), using a training dataset of 14M waveforms. The sheer size of this training dataset requires the combination of AI and HPC, and thus we have used the Summit supercomputer to reduce time-to-insight through distributed training.

While this article showcases the convergence of AI and HPC for a computational grand challenge, the main goal of this analysis is to explore what new insights we may obtain by conducting AI-driven studies of the signal manifold of higher order gravitational wave modes. In particular, in this article we aim to address a few issues that we discussed in [43], namely:

1. Is it possible for AI to learn and accurately characterize the physics of high-dimensional gravitational wave signal manifolds?
2. Is it possible to exploit the computational efficiency and scalability of AI to train models with tens of millions of modeled waveforms and conduct fast data-driven analyses with fully trained AI models?
3. Is it true that the convergence and performance of AI models improves as we consider large volume training datasets that describe signal manifolds with higher degrees of freedom?
4. What insights do we gain when we characterize gravitational wave signal manifolds with waveforms with complex morphology?

As we describe below, the answer to questions 1-3 above is a resounding YES. In terms of new insights, we find that our deterministic and probabilistic AI models provide informative constraints for the mass-ratio, individual spins and inclination angle of higher order wave modes. These are important results, since our previous work [43] showed that, when we only consider \( \ell = |m| = 2 \) modes, it
was difficult to constrain the individual spins of comparable mass-ratio systems, as well as the spin of the secondary for asymmetric mass-ratio systems. This study shows that the inclusion of higher order modes alleviates these problems, and thus provide an informed description of the ability of AI to characterize this high dimensional signal manifold in the absence of noise. Throughout this paper we use geometric units in which $G = c = 1$.

This paper is organized as follows. Section 2 described the approach used to create our AI models. We present and discuss our findings in Section 3. Future directions of work are outlined in Section 4.

2. Methods

Here we describe the datasets, AI architectures and training methods followed to create our AI models.

Datasets We use the surrogate model NRHybSur3dq8 [41] to generate time-series datasets that include both the plus, $h_+$, and cross, $h_\times$, polarizations. These may be represented as a complex time-series $h = h_+ - i h_\times$. $h$ may also be expressed as a sum of spin-weighted spherical harmonic modes, $h_{\ell m}$, on the sphere $[45]$

$$h(t, \theta, \phi_0) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} h_{\ell m}(t) -2Y_{\ell m}(\theta, \phi_0),$$  

(1)
where \(-2Y_{\ell m}\) are the spin-weight \(-2\) spherical harmonics, \(\theta\) is the inclination angle between the orbital angular momentum of the binary and line of sight to the detector, and \(\phi_0\) is the initial binary phase, that we set to zero in this study. Our waveforms include higher order modes \(\ell \leq 4\) and \((5,5)\), excluding the \((4,0)\) and \((4,1)\) modes; cover the time span \(t \in [-10,000 \, M, 130 \, M]\) with the merger peak occurring at \(t = 0 \, M\); and are sampled with a time step \(\Delta t = 1 \, M\).

In Figure 2 we showcase the importance of including higher order modes to accurately model binary black hole mergers. It is apparent that the inclusion of higher order modes produces significant modifications in the amplitude and phase evolution of waveform signals.

**Figure 2: Impact of higher order modes.** For a binary black hole merger with parameters \(\{q, s_1^z, s_2^z\} = \{7.05, 0.77, 0.77\}\), we show waveforms with \((l = |m| = 2; \theta = 0)\) (top panel), and waveforms including higher order modes up to \(\ell \leq 4\) and \((5,5)\), excluding the \((4,0)\) and \((4,1)\) modes, for \(\theta = \{0, \pi/4, \pi/2\}\) (bottom panels).

**Training dataset** It consists of \(\sim 14\) million waveforms that cover a 4-D parameter space that encompasses mass-ratio, individual spins and inclination angle \(\{q, s_1^z, s_2^z, \theta\}\), respectively. We generate it by sampling the mass-ratio \(q \in [1, 8]\) in steps of \(\Delta q = 0.1\); individual spins \(s_1^z \in [-0.8, 0.8]\) in steps of \(\Delta s_1^z = 0.02\); and the inclination angle \(\theta \in [0, \pi]\) in steps of \(\Delta \theta = 0.1\).

**Validation and test datasets** Each of these sets consist of \(\sim 800,000\) waveforms, and are generated by alternately sampling values that are inbetween the training set values.

**AI architecture** We use a slightly modified WaveNet [46] neural network architecture for our model. WaveNet’s main features that are relevant for this work include dilated causal convolutions, gated activation units, and the usage of residual and skip connections. These features help capture long range correlations in the input time-series, and facilitate the training of deeper neural networks. Furthermore, since we are interested in regression analyses, we turn off the causal padding in the convolutional layers. We use a filter size of 2 in all convolutional layers and stack 3 residual blocks each consisting of 14 dilated convolutions. For a more in-depth discussion of the architecture we refer the reader to [43, 46]. The output from the WaveNet is then fed into three separate branches of fully-connected layers. Each branch is trained to predict the
mass ratio $q$, the effective spin parameters $S_{\text{eff}}$, $\sigma_{\text{eff}}$, and the inclination angle $\theta$ respectively, where

$$S_{\text{eff}} = \sigma_1 s_1^z + \sigma_2 s_2^z,$$

with $\sigma_1 \equiv 1 + \frac{3}{4q}$ and $\sigma_2 \equiv 1 + \frac{3q}{4}$, \hspace{1cm} (2)

$$\sigma_{\text{eff}} = \frac{m_1 s_1^z + m_2 s_2^z}{m_1 + m_2} = \frac{q s_1^z + s_2^z}{1 + q}. \hspace{1cm} (3)$$

Since our goal is to predict $\{q, s_1^z, s_2^z, \theta\}$, we solve Eqs. (2) and (3) in conjunction with the predicted $q$ values in order to extract the individual spins $s_i^z$.

**Training methodology** We employ mean-squared error (MSE) between the predicted and the ground-truth values as the loss function. During training, we monitor the loss on the validation set to dynamically reduce learning rate as well as to stop training before over-fitting. We reduce the learning rate by a factor of 2 whenever the validation loss does not decrease for 3 consecutive epochs, and stop the training when the validation loss does not decrease for 5 consecutive epochs. Training the model on 256 nodes, equivalent to 1,536 NVIDIA V100 GPUs, in the Summit Supercomputer then takes about 71 epochs, using LAMB optimizer with initial learning rate set to 0.001.

**Normalizing Flow:** In addition to the above analysis for point estimates, we also train a normalizing flow model to estimate the posterior distribution. To do that, instead of extracting the parameters directly, we use the WaveNet model as a feature extractor and then condition a normalizing flow model on the extracted features to estimate the posterior distribution. This method was first delineated in [48], and then used in other studies [28, 33]. We follow the same procedure, making use of the nflows library [49]. Normalizing flow is an example of a “likelihood-free” inference method, and is made up of a composition of invertible maps to transform a simple base probability distribution (e.g., a Multivariate Gaussian) into a desired posterior distribution which could be very complicated. The transformed distribution is then given by the change of variable formula:

$$p_X(x) = p_Z(z) |\det[J_f(z)]|^{-1}, \hspace{1cm} (4)$$

where $Z$ is the random variable for the base distribution, $X$ is the random variable for the transformed distribution, and $f$ is the normalizing flow (i.e., the invertible transformation), such that $X = f(Z)$. In our case, the goal is to model the conditional posterior distribution $p(y|h)$ for the parameters $y$ corresponding to the waveform strain $h$. We do this in two steps (as illustrated in Figure 3): first we pass the strain $h$ through the WaveNet model to extract a feature vector $\hat{h}$. We then use a conditional version of Normalizing Flow $f_{\hat{h}, \vartheta}$, specifically a Neural Spline Flow [50], to transform the base Standard Multivariate Gaussian $\mathcal{N}(\mu = 0, \Sigma = I)$ to the predicted posterior distribution $q(y|h)$. The function $f_{\hat{h}, \vartheta}$ therefore depends on the input waveform $h$ (through the feature vector $\hat{h}$), and is parametrized by learnable weights $\vartheta$. The normalizing flow model is then trained by updating the parameter $\vartheta$ so that predicted distribution $q(y|h)$ matches the true posterior distribution $p(y|h)$. This is achieved by minimizing
the negative log-likelihood, i.e., given a batch of \( N \) ground-truth parameters \( y_i \) and their corresponding waveforms \( h_i \), the objective for the Normalizing Flow model is to minimize:

\[
L = -\frac{1}{N} \sum_{i=1}^{N} \log q(y_i|h_i) \tag{5}
\]

where, according to equation \( \text{[4]} \)

\[
q(y_i|h_i) = p_N(z) \left| \det[J_{f_{h,\theta}}(z)] \right|^{-1} . \tag{6}
\]

3. Results

We present results using deterministic and probabilistic AI models, which we described in Section \( \text{[2]} \). As described above, these AI models have been designed to take in time-series waveform signals that includes both polarizations \( (h_+, h_\times) \), and then output the most likely values for \( \{q, s_1^z, s_2^z, \theta\} \) that best reproduce the input signal.
3.1. Deterministic AI models

In Figures 4 and 5 we provide a sample of results of the predictive capabilities of our deterministic AI models for a variety of input waveform signals. Note that ground truth waveforms are shown in blue, whereas waveforms whose parameters, \( \{q, s^1_z, s^2_z, \theta\} \), are predicted by AI are shown in dotted red. We quantify the accuracy of our AI models by computing the overlap, \( \mathcal{O}(h_t| h_p) \), between ground-truth, \( h_t \), and AI-predicted, \( h_p \), waveforms using

\[
\mathcal{O}(h_t| h_p) = \max_{[t_c, \phi_c]} \left( \hat{h}_t|\hat{h}_p[t_c, \phi_c] \right), \quad \text{with} \quad \hat{h}_t = h_t| h_t)^{-1/2},
\]

where \([t_c, \phi_c]\) indicates that the normalized waveform \(\hat{h}_p\) has been time- and phase-shifted.

Figure 4: Comparison between ground truth and AI-predicted waveforms

Ground truth waveforms (in blue) and AI predicted waveforms (dotted red). Top panels: \( \{q, s^1_z, s^2_z\} = \{2.05, -0.79, 0.77\} \). Bottom panels: \( \{q, s^1_z, s^2_z\} = \{3.05, -0.79, 0.77\} \). We show waveforms, from left to right, with inclination angles \( \theta = \{0, \pi/4, \pi/2\} \). In all these cases, the overlap between ground truth and predicted waveforms is \( \mathcal{O} \geq 0.99 \).

Figure 6 summarizes the accuracy with which deterministic AI models estimate the parameters of higher order modes for all mass-ratios and spins for a sample of inclination angles. There we notice that predictions degrade in accuracy for edge-on systems, i.e., \( \theta = \pi/2 \). We provide a more comprehensive analysis of these results in Figure 7 where we show overlap calculations between ground truth and predicted signals in terms of symmetric mass-ratio and effective spin \((\eta, \sigma_{\text{eff}})\) for all mass-ratios and spins under consideration for a sample of inclination angles. Again, here we find that our results are optimal for all angles except for edge-on binaries. We can understand this if we recall that for \( \theta = \pi/2 \) we lose half of the information we feed into our AI models since \( h_\times(t, \theta = \pi/2) \to 0 \).

These results demonstrate the ability of AI to search across the signal manifold of higher order modes, and pinpoint a set of parameters, \( \{q, s^1_z, s^2_z, \theta\} \), that best describes the amplitude and phase evolution of higher order modes. While this is informative, we also want to know the uncertainty associated with such AI-predicted tetrad of values. To extract such information, we estimate posterior distributions using a combination of WaveNet with normalizing flow, as described in Section 2.
3.2. Probabilistic AI models

We have selected a number of binary black hole systems to quantify the ability of AI to reconstruct the astrophysical parameters of systems that are known to be hard to characterize. In particular, we consider comparable mass-ratio binaries, for which it is difficult to tell apart individual spins, as well as asymmetric mass-ratio systems, for which it is difficult to accurately constrain the spin of the secondary. These results are presented in Figures 8-11. In all these figures, blue lines represent ground truth values and red dotted lines are point-parameter
Figure 7: Overlap between ground truth signals and those predicted by deterministic AI models

Overlap between ground truth and AI-predicted signals in terms of symmetric mass-ratio, $\eta$, and effective spin, $\sigma_{\text{eff}}$, for a sample of inclination angles.

estimates provided by our deterministic WaveNet model. The histograms are produced by drawing 10,000 samples from the multivariate Gaussian base distribution, and then applying the normalizing flow conditioned on the input waveform. This is therefore equivalent to drawing 10,000 samples from the posterior distribution.

The results we present below encompass the following scenarios:

- **Comparable mass-ratio systems** Figure 8 shows that AI provides informative constraints for the individual spins ($s_{z1}$, $s_{z2}$), the mass-ratio, $q$ and the inclination angle, $\theta$. We also reconstructed the effective spin parameters ($S_{\text{eff}}$, $\sigma_{\text{eff}}$), given by Equations (2) and (3). These results indicate that AI is capable of providing reliable constraints for ($q$, $s_{z1}$, $s_{z2}$, $\theta$, $S_{\text{eff}}$, $\sigma_{\text{eff}}$) of comparable mass-ratio systems.

- **Moderately asymmetric mass-ratio systems** Figures 9 and 10 present parameter estimation results for binaries with mass-ratio $q = 4$. Here we considered two spin configurations to explore whether AI can still provide informative predictions when the spin of the secondary is high, $s_{z2} \sim 0.8$, Figure 9, and moderate, $s_{z2} \sim -0.5$, Figure 10. We notice that in both cases, AI provides reliable deterministic and probabilistic estimates for ($q$, $s_{z1}$, $s_{z2}$, $\theta$, $S_{\text{eff}}$, $\sigma_{\text{eff}}$).

- Figures 11 and 12 present parameter estimation results for mass-ratio $q = 7$ binaries. The cases presented therein shed light on whether AI can provide informative constraints when the spin of the secondary is high $s_{z2} \sim 0.8$, Figure 11, or negligible $s_{z2} \sim 0$, Figure 12. In principle, one may expect that given the mass-ratio of the system, it may be difficult to constrain the spin...
of the secondary regardless of its spin. However, these results indicate that AI provides informative constraints not only for the individual spins, but the entire set \( (q, s_{1z}, s_{2z}, \theta, S_{\text{eff}}, \sigma_{\text{eff}}) \).

In summary, both our deterministic and probabilistic AI models provide informative constraints on the astrophysical properties of gravitational waves that include higher order modes to accurately describe the physics of quasi-circular, spinning, non-precessing binary black holes.

Another important result of this paper is that we have designed a methodology to train AI models that adequately handle training datasets that include tens of millions of modeled waveforms, thereby paving the way to extend this analysis for the case in which these types of signals are contaminated with simulated and advanced LIGO noise. The methods introduced in this paper will enable us to quantify the biases introduced by noise in parameter estimation analyses, and how to handle them to extract informative AI-driven parameter estimation results using higher order gravitational wave modes.
Figure 8: **Deterministic and probabilistic AI results** AI posterior distributions for the mass-ratio, individual spins, effective spins and inclination angle, \((q, s_1^z, s_2^z, \sigma_{\text{eff}}, S_{\text{eff}}),\) respectively, of an equal mass binary black hole merger. Blue lines indicate ground truth values, red dotted lines show point-wise predictions from our deterministic AI model. The histograms show the distribution of 10,000 samples drawn from the posterior.
Figure 9: **Deterministic and probabilistic AI results** As Figure 8 but now for a binary with mass-ratio $q = 4$ whose binary components are rapidly spinning.
Figure 10: **Deterministic and probabilistic AI results** As Figure 9 but now the secondary is slowly rotating in the $-z$ direction.
Figure 11: Deterministic and probabilistic AI results As Figure 12 but now both binary components are rapidly rotating.
Figure 12: **Deterministic and probabilistic AI results** As Figure 8 but now for a $q = 7$ binary with a rapidly rotating primary and a nearly non-spinning secondary.
4. Conclusions

We have developed scalable and computationally efficient methods to design AI models that are capable of characterizing the signal manifold of higher order wave modes of quasi-circular, spinning, non-precessing binary black hole mergers. Our approach enabled us to train several AI models using a dataset of over 14M waveforms within 3.4 hours with 256 nodes, equivalent to 1,536 NVIDIA V100 GPUs, achieving optimal convergence and state-of-the-art regression results.

We have demonstrated that AI can abstract knowledge from time-series data that helps constrain the physical parameters that determine the dynamical evolution of higher order modes of black hole mergers. In particular, we have provided evidence that AI provides deterministic and probabilistic predictions that tightly constrain the mass-ratio, individual spins, inclination angle, and effective spin parameters for a variety of astrophysical scenarios. We also found that deterministic and probabilistic predictions are consistent with each other, and in good accord with ground truth physical parameters.

The results we have introduced in this article provide benchmarks for the expected performance of AI to estimate the astrophysical parameters of binary black hole mergers in the absence of noise. In future work, we will present studies for the impact of simulated and advanced LIGO noise to conduct informative AI-driven inference for high dimensional signal manifolds.

5. Acknowledgements

A.K. and E.A.H. gratefully acknowledge National Science Foundation (NSF) awards OAC-1931561 and OAC-1934757, and the Innovative and Novel Computational Impact on Theory and Experiment project ‘Multi-Messenger Astrophysics at Extreme Scale in Summit’. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under contract no. DE-AC05-00OR22725. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. This work utilized resources supported by the NSF’s Major Research Instrumentation program, the HAL cluster (grant no. OAC-1725729), as well as the University of Illinois at Urbana-Champaign. We thank NVIDIA for their continued support.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[2] A. Davies, P. Veličković, L. Buesing, S. Blackwell, D. Zheng, N. Tomašev, R. Tanburn, P. Battaglia, C. Blundell, A. Juhász, M. Lackenby, G. Williamson, D. Hassabis, and P. Kohli, “Advancing mathematics by guiding human intuition with ai,” Nature, vol. 600, pp. 70–74, 12 2021.
[3] D. George and E. A. Huerta, “Deep neural networks to enable real-time multimessenger astrophysics,” *Phys. Rev. D*, vol. 97, p. 044039, Feb 2018.

[4] D. George and E. Huerta, “Deep learning for real-time gravitational wave detection and parameter estimation: Results with advanced ligo data,” *Physics Letters B*, vol. 778, pp. 64 – 70, 2018.

[5] E. A. Huerta, G. Allen, I. Andreoni, J. M. Antelis, E. Bachelet, G. B. Berriman, F. B. Bianco, R. Biswas, M. Carrasco Kind, K. Chard, M. Cho, P. S. Cowperthwaite, Z. B. Etienne, M. Fishbach, F. Forster, D. George, T. Gibbs, M. Graham, W. Gropp, R. Gruendl, A. Gupta, R. Haas, S. Habib, E. Jennings, M. W. G. Johnson, E. Katsonidou, D. S. Katz, A. Khan, V. Kindratenko, W. T. C. Kramer, X. Liu, A. Mahabal, Z. Marka, K. McHenry, J. M. Miller, C. Moreno, M. S. Neubauer, S. Oberlin, A. R. Olivas, D. Petravick, A. Rebei, S. Rosofsky, M. Ruiz, A. Saxton, B. F. Schutz, A. Schwing, E. Seidel, S. L. Shapiro, H. Shen, Y. Shen, L. P. Singer, B. M. Sipocz, L. Sun, J. Towns, A. Tsokaros, W. Wei, J. Wells, T. J. Williams, J. Xiong, and Z. Zhao, “Enabling real-time multi-messenger astrophysics discoveries with deep learning,” *Nature Reviews Physics*, vol. 1, pp. 600–608, Oct. 2019.

[6] E. A. Huerta and Z. Zhao, *Advances in Machine and Deep Learning for Modeling and Real-Time Detection of Multi-messenger Sources*, pp. 1–27. Singapore: Springer Singapore, 2020.

[7] E. Cuoco, J. Powell, M. Cavaglià, K. Ackley, M. Bejger, C. Chatterjee, M. Coughlin, S. Coughlin, P. Easter, R. Essick, H. Gabbard, T. Gebhard, S. Ghosh, L. Haegel, A. Iess, D. Keitel, Z. Marka, S. Marka, F. Morawski, T. Nguyen, R. Ormiston, M. Puerer, M. Razzano, K. Staats, G. Vajente, and D. Williams, “Enhancing Gravitational-Wave Science with Machine Learning,” *Mach. Learn. Sci. Tech.*, vol. 2, no. 1, p. 011002, 2021.

[8] H. Gabbard, M. Williams, F. Hayes, and C. Messenger, “Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy,” *Physical Review Letters*, vol. 120, p. 141103, Apr. 2018.

[9] V. Skliris, M. R. K. Norman, and P. J. Sutton, “Real-Time Detection of Unmodeled Gravitational-Wave Transients Using Convolutional Neural Networks,” *arXiv e-prints*, p. arXiv:2009.14611, Sept. 2020.

[10] Y.-C. Lin and J.-H. P. Wu, “Detection of gravitational waves using bayesian neural networks,” *Phys. Rev. D*, vol. 103, p. 063034, Mar 2021.

[11] H. Wang, S. Wu, Z. Cao, X. Liu, and J.-Y. Zhu, “Gravitational-wave signal recognition of LIGO data by deep learning,” *Phys. Rev. D*, vol. 101, no. 10, p. 104003, 2020.
[12] X. Fan, J. Li, X. Li, Y. Zhong, and J. Cao, “Applying deep neural networks to the detection and space parameter estimation of compact binary coalescence with a network of gravitational wave detectors,” Sci. China Phys. Mech. Astron., vol. 62, no. 6, p. 969512, 2019.

[13] X.-R. Li, G. Babu, W.-L. Yu, and X.-L. Fan, “Some optimizations on detecting gravitational wave using convolutional neural network,” Front. Phys. (Beijing), vol. 15, no. 5, p. 54501, 2020.

[14] D. S. Deighan, S. E. Field, C. D. Capano, and G. Khanna, “Genetic-algorithm-optimized neural networks for gravitational wave classification,” arXiv e-prints, p. arXiv:2010.04340, Oct. 2020.

[15] A. L. Miller et al., “How effective is machine learning to detect long transient gravitational waves from neutron stars in a real search?,” Phys. Rev. D, vol. 100, no. 6, p. 062005, 2019.

[16] P. G. Krastev, “Real-Time Detection of Gravitational Waves from Binary Neutron Stars using Artificial Neural Networks,” Phys. Lett. B, vol. 803, p. 135330, 2020.

[17] M. B. Schäfer, F. Ohme, and A. H. Nitz, “Detection of gravitational-wave signals from binary neutron star mergers using machine learning,” Phys. Rev. D, vol. 102, p. 063015, Sept. 2020.

[18] C. Dreissigacker and R. Prix, “Deep-Learning Continuous Gravitational Waves: Multiple detectors and realistic noise,” Phys. Rev. D, vol. 102, no. 2, p. 022005, 2020.

[19] A. Rebei, E. A. Huerta, S. Wang, S. Habib, R. Haas, D. Johnson, and D. George, “Fusing numerical relativity and deep learning to detect higher-order multipole waveforms from eccentric binary black hole mergers,” Phys. Rev. D, vol. 100, p. 044025, Aug 2019.

[20] C. Dreissigacker, R. Sharma, C. Messenger, R. Zhao, and R. Prix, “Deep-Learning Continuous Gravitational Waves,” Phys. Rev. D, vol. 100, no. 4, p. 044009, 2019.

[21] B. Beheshtipour and M. A. Papa, “Deep learning for clustering of continuous gravitational wave candidates,” Phys. Rev. D, vol. 101, p. 064009, Mar. 2020.

[22] M. B. Schäfer and A. H. Nitz, “From One to Many: A Deep Learning Coincident Gravitational-Wave Search,” arXiv e-prints, p. arXiv:2108.10715, Aug. 2021.

[23] M. B. Schäfer, O. Zelenka, A. H. Nitz, F. Ohme, and B. Brügmann, “Training Strategies for Deep Learning Gravitational-Wave Searches,” arXiv e-prints, p. arXiv:2106.03741, June 2021.
[24] A. Gunny, D. Rankin, J. Krupa, M. Saleem, T. Nguyen, M. Coughlin, P. Harris, E. Katsavounidis, S. Timm, and B. Holzman, “Hardware-accelerated Inference for Real-Time Gravitational-Wave Astronomy,” arXiv e-prints, p. arXiv:2108.12430, Aug. 2021.

[25] H. Shen, D. George, E. A. Huerta, and Z. Zhao, “Denoising gravitational waves with enhanced deep recurrent denoising auto-encoders,” in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3237–3241, IEEE, 2019.

[26] W. Wei and E. A. Huerta, “Gravitational Wave Denoising of Binary Black Hole Mergers with Deep Learning,” Phys. Lett., vol. B800, p. 135081, 2020.

[27] R. Ormiston, T. Nguyen, M. Coughlin, R. X. Adhikari, and E. Katsavounidis, “Noise reduction in gravitational-wave data via deep learning,” Phys. Rev. Research, vol. 2, p. 033066, Jul 2020.

[28] H. Shen, E. A. Huerta, E. O’Shea, P. Kumar, and Z. Zhao, “Statistically-informed deep learning for gravitational wave parameter estimation,” Machine Learning: Science and Technology, vol. 3, p. 015007, Nov 2021.

[29] H. Gabbard, C. Messenger, I. S. Heng, F. Tonolini, and R. Murray-Smith, “Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy,” arXiv preprint arXiv:1909.06296, 2019.

[30] A. J. Chua and M. Vallisneri, “Learning Bayesian posteriors with neural networks for gravitational-wave inference,” Phys. Rev. Lett., vol. 124, no. 4, p. 041102, 2020.

[31] S. R. Green, C. Simpson, and J. Gair, “Gravitational-wave parameter estimation with autoregressive neural network flows,” Phys. Rev. D, vol. 102, p. 104057, Nov 2020.

[32] S. R. Green and J. Gair, “Complete parameter inference for gw150914 using deep learning,” Machine Learning: Science and Technology, vol. 2, p. 03LT01, June 2020.

[33] M. Dax, S. R. Green, J. Gair, J. H. Macke, A. Buonanno, and B. Schölkopf, “Real-time gravitational-wave science with neural posterior estimation,” arXiv e-prints, p. arXiv:2106.12594, June 2021.

[34] M. Dax, S. R. Green, J. Gair, M. Deistler, B. Schölkopf, and J. H. Macke, “Group equivariant neural posterior estimation,” arXiv e-prints, p. arXiv:2111.13139, Nov. 2021.

[35] S. Khan and R. Green, “Gravitational-wave surrogate models powered by artificial neural networks,” Phys. Rev. D, vol. 103, no. 6, p. 064015, 2021.

[36] A. J. K. Chua, C. R. Galley, and M. Vallisneri, “Reduced-order modeling with artificial neurons for gravitational-wave inference,” Phys. Rev. Lett., vol. 122, p. 211101, May 2019.
[37] J. Lee, S. H. Oh, K. Kim, G. Cho, J. J. Oh, E. J. Son, and H. M. Lee, “Deep learning model on gravitational waveforms in merging and ringdown phases of binary black hole coalescences,” Phys. Rev. D, vol. 103, p. 123023, June 2021.

[38] A. Khan, E. A. Huerta, and H. Zheng, “Interpretable AI forecasting for numerical relativity waveforms of quasi-circular, spinning, non-precessing binary black hole mergers,” arXiv e-prints, p. arXiv:2110.06968, Oct. 2021.

[39] W. Wei, E. A. Huerta, M. Yun, N. Loutrel, M. A. Shaikh, P. Kumar, R. Haas, and V. Kindratenko, “Deep Learning with Quantized Neural Networks for Gravitational-wave Forecasting of Eccentric Compact Binary Coalescence,” Astrophys. J., vol. 919, p. 82, Oct. 2021.

[40] W. Wei and E. A. Huerta, “Deep learning for gravitational wave forecasting of neutron star mergers,” Phys. Lett. B, vol. 816, p. 136185, 2021.

[41] H. Yu, R. X. Adhikari, R. Magee, S. Sachdev, and Y. Chen, “Early warning of coalescing neutron-star and neutron-star-black-hole binaries from the nonstationary noise background using neural networks,” Phys. Rev. D, vol. 104, p. 062004, Sept. 2021.

[42] E. A. Huerta, A. Khan, X. Huang, M. Tian, M. Levental, R. Chard, W. Wei, M. Heflin, D. S. Katz, V. Kindratenko, D. Mu, B. Blaiszik, and I. Foster, “Accelerated, scalable and reproducible AI-driven gravitational wave detection,” Nature Astronomy, vol. 5, pp. 1062–1068, July 2021.

[43] A. Khan, E. A. Huerta, and A. Das, “Physics-inspired deep learning to characterize the signal manifold of quasi-circular, spinning, non-precessing binary black hole mergers,” Physics Letters B, vol. 808, p. 135628, Sept. 2020.

[44] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, “Surrogate model of hybridized numerical relativity binary black hole waveforms,” Phys. Rev. D, vol. 99, p. 064045, Mar 2019.

[45] E. T. Newman and R. Penrose, “Note on the bondi-metzner-sachs group,” Journal of Mathematical Physics, vol. 7, no. 5, pp. 863–870, 1966.

[46] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, “Wavenet: A generative model for raw audio,” 2016.

[47] Y. You, J. Li, S. Reddi, J. Hseu, S. Kumar, S. Bhojanapalli, X. Song, J. Demmel, K. Keutzer, and C.-J. Hsieh, “Large batch optimization for deep learning: Training bert in 76 minutes,” 2019.

[48] S. R. Green and J. Gair, “Complete parameter inference for gw150914 using deep learning,” Machine Learning: Science and Technology, vol. 2, no. 3, p. 03LT01, 2021.
[49] C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios, “nflows: normalizing flows in PyTorch,” Nov. 2020.

[50] C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios, “Neural spline flows,” Advances in Neural Information Processing Systems, vol. 32, pp. 7511–7522, 2019.