The Astrophysical Journal, 919:82 (10pp), 2021 October 1 https://doi.org/10.3847/1538-4357/ac1121

Deep Learning with Quantized Neural Networks for Gravitational-wave Forecasting of Eccentric Compact Binary Coalescence

Wei Wei[1,2,3], E. A. Huerta[1,3,4,5,6], Mengshen Yun[1,2,7], Nicholas Loutrel[8,9], Md Arif Shaikh[10], Prayush Kumar[10,11], Roland Haas[1], and Volodymyr Kindratenko[1,2,7,12]

[1] National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; weiw2@illinois.edu
[2] NCSA Center for Artificial Intelligence Innovation, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[4] Data Science and Learning Division, Argonne National Laboratory, Lemont, IL 60439, USA
[5] Department of Computer Science, University of Chicago, Chicago, IL 60637, USA
[6] Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[7] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[8] Department of Physics, Princeton University, Princeton, NJ 08544, USA
[9] Princeton Gravity Initiative, Princeton University, Princeton, NJ 08544, USA
[10] International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore 560089, India
[11] Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, NY 14853, USA
[12] Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
Received 2020 December 9; revised 2021 June 25; accepted 2021 July 3; published 2021 September 28

Abstract

We present the first application of deep learning forecasting for binary neutron stars, neutron star–black hole systems, and binary black hole mergers that span an eccentricity range $e \leqslant 0.9$. We train neural networks that describe these astrophysical populations, and then test their performance by injecting simulated eccentric signals in advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) noise available at the Gravitational Wave Open Science Center to (1) quantify how fast neural networks identify these signals before the binary components merge; (2) quantify how accurately neural networks estimate the time to merger once gravitational waves are identified; and (3) estimate the time-dependent sky localization of these events from early detection to merger. Our findings show that deep learning can identify eccentric signals from a few seconds (for binary black holes) up to tens of seconds (for binary neutron stars) prior to merger. A quantized version of our neural networks achieves $4\times$ reduction in model size, and up to $2.5\times$ inference speedup. These novel algorithms may be used to facilitate time-sensitive multimessenger astrophysics observations of compact binaries in dense stellar environments.

Unified Astronomy Thesaurus concepts: Gravitational waves (678); Convolutional neural networks (1938); Neutron stars (1108); Eccentricity (441); LIGO (920)

1. Introduction

Multimessenger observations that combine the gravitational and electromagnetic spectra (Abbott et al. 2017a, 2017b, 2017c; Smith 2017; The LIGO Scientific Collaboration et al. 2017; Troja et al. 2017; Mooley et al. 2018) have provided revolutionary insights about the nature of gravity, the engines that power short gamma-ray bursts, cosmology, and fundamental physics (Schutz 1986; Abbott et al. 2017d; Berti et al. 2018; M.S.-S. et al. 2019; Fishbach et al. 2019; Abbott et al. 2019a). These remarkable discoveries provide a glimpse of what multimessenger astrophysics may accomplish once gravitational-wave detectors reach their design sensitivity, and they work in unison with electromagnetic and astroparticle observatories to observe the transient universe with unprecedented precision (Abbott et al. 2019b; Mészáros et al. 2019; Georgescu 2020). To realize these goals, however, there is an urgent need to develop signal-processing tools and computing frameworks that turn computational grand challenges in the big-data era into unique opportunities to enable new modes of data-driven discovery (Huerta et al. 2019; Georgescu 2020; Huerta & Zhao 2021).

In the realm of gravitational-wave observations, recent developments include the production of early-warning systems to forecast the merger of multimessenger sources using template-matching methods in the context of simulated advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) noise (Cannon et al. 2012; Sachdev et al. 2020), and third-generation ground-based detectors (Nitz et al. 2020). Deep learning has emerged as a powerful tool to process data at scale, with a similar sensitivity of traditional algorithms, but at a fraction of their computational cost. Deep learning methods have evolved rapidly in gravitational-wave astrophysics, ranging from the first algorithms that were proposed to enable real-time gravitational-wave detection with advanced LIGO noise (George & Huerta 2018, 2017, 2018), to the production of sophisticated neural networks that span the same signal manifold of traditional low-latency pipelines, process hundreds of hours of advanced LIGO noise faster than real-time with just a handful of GPUs, and identify real events with a minimal number of false positives in real advanced LIGO noise (Huerta et al. 2021; Wei et al. 2021). Deep learning has also been used to forecast the merger of quasi-circular binary neutron stars and black hole–neutron stars systems in real advanced LIGO noise (Wei & Huerta 2021).

To date, neural networks have been developed in the context of quasi-circular gravitational-wave sources, even though some of these models have been used to explore the detection of eccentric binary black hole mergers (Rebei et al. 2019). In this paper we introduce the first class of neural networks that are trained and tested with waveforms that describe binary neutron stars, neutron star–black hole mergers, and binary black hole mergers that cover a broad eccentricity range, $e \leqslant 0.9$. These neural networks target a significantly much more challenging

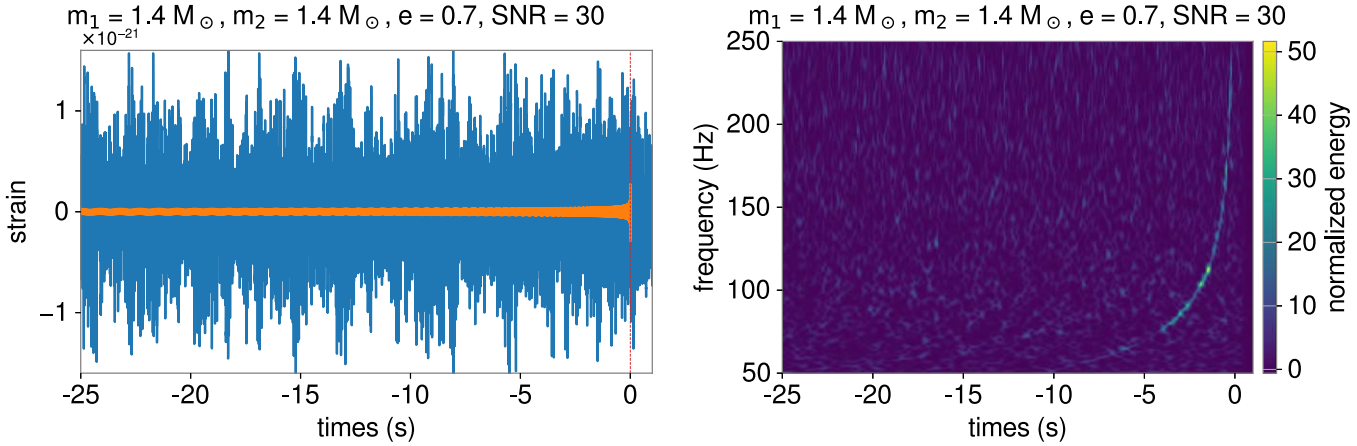

Figure 1. Left panel: gravitational-wave signal injected in advanced LIGO O2 noise that describes a binary neutron star with component masses $m_{\{1,2\}} = 1.4\,M_{\odot}$ and eccentricity $e = 0.7$ measured 150 s before merger. Right panel: spectrum of frequencies of the eccentric waveform signal shown in the left panel.

task than detection, since these models can predict the merger of eccentric signals embedded in advanced LIGO data from a few seconds up to 2 minutes before the merger event.

This study is motivated by a number of considerations. For instance, it is well known that eccentric waveforms have a rich and complex morphology at early frequencies. Given that neural networks have been particularly successful at time-series processing and pattern identification (LeCun et al. 2015), it is worth exploring whether we can we train neural networks to enable gravitational-wave forecasting by leveraging the rich spectrum of frequencies that characterize eccentric signals in the context of advanced LIGO noise. If this analysis was indeed possible, we would like to quantify how fast deep learning may predict the merger of likely multimessenger sources—neutron star mergers or stellar mass black hole–neutron star systems—or binary black holes that coalesce in dense stellar environments. This is the theme of this article.

We organize this article as follows. Section 2 describes our deep learning algorithms, the modeled waveforms, and advanced LIGO noise used to train and test our models. We present our forecasting results in Section 3. We analyze network quantization results in Section 4. We summarize our findings and discuss future activities in Section 5.

2. Methods

Here we describe the modeled waveforms and advanced LIGO noise used to create our neural networks, and how these models may be used to forecast the merger of compact binary systems in advanced LIGO noise.

2.1. Waveforms and Advanced LIGO Noise

Modeled waveforms. We used the waveform model introduced in East et al. (2013) to describe eccentric compact binary systems. Even though this is an admittedly approximate model, it provides a complete description of the systems' dynamics, from inspiral to ringdown, and more importantly, it enables the modeling of highly eccentric systems. The waveforms used in this study are produced at a sample rate of 16,384 Hz, and describe binary neutron stars, black hole–neutron star systems, and binary black hole mergers. For binary neutron stars, we consider systems with binary components $m_{\{1,2\}} \in [1\,M_{\odot},\ 3\,M_{\odot}]$, whereas for black hole–neutron star systems we considered $m_{\rm BH} \in [3\,M_{\odot},\ 15\,M_{\odot}]$ and $m_{\rm NS} \in [1\,M_{\odot},\ 3\,M_{\odot}]$ for the masses of the black hole and neutron star components, respectively.

The waveforms in the training data set cover the eccentricity range $e \leqslant 0.9$, and are 160 s long. These waveforms are randomly split into training sets (12,423 waveforms for binary neutron stars, 15,593 for black hole–neutron star systems, and 10,677 for binary black holes), and test sets (3075 waveforms for binary neutron stars, 3882 for black hole–neutron star systems, and 2661 for binary black holes).

Advanced LIGO noise. We have used four 4096 s long advanced LIGO noise segments, sampled at 16,384 Hz, from the Hanford and Livingston detectors. These segments have GPS starting times 1186725888, 1187151872, 1187569664, and 1186897920. The first three segments are used for training, while the last is used for testing. All these open-source data were obtained from the Gravitational Wave Open Science Center (Vallisneri et al. 2015).

We rescaled and injected the waveform data sets described above into real advanced LIGO noise, using both Livingston and Hanford data, to simulate eccentric mergers that span a broad range of signal-to-noise ratios (S/Ns). We standardized the data sets for training by normalizing the standard deviation of the advanced LIGO strain data with signal injections.

2.2. Spectrograms of Eccentric Compact Binary Mergers

Eccentric binary neutron stars have long-duration inspiral stages with periodic spikes in time domain, as shown in the right panel of Figure 1. Their spectral decomposition provides a rich spectrum of frequencies, as shown in the right panel of Figure 1. We use the early inspiral-stage presence of the chirp patterns in the spectrograms to provide early warnings for imminent merger events, which may provide early warnings for potential electromagnetic follow-ups.

2.3. Chirp-pattern Recognition and Merger Time Estimation with Deep Learning

Early detection. We use a deep neural network to identify inspiral-stage chirp-like signatures in spectrograms, and provide tens of seconds early warnings for a variety of compact binary systems.

Specifically, we use `ResNet-50` (He et al. 2016) architecture pretrained on `ImageNet` (Deng et al. 2009), implemented

with PyTorch (Paszke et al. 2019) for our pattern-recognition tasks. We produce and stack together spectrograms from 8 s long advanced LIGO strain data to form an image of two channels with the first channel representing Livingston data, and the second channel representing Hanford data. We then pad a third channel with zeros so that the final images have three channels to match the ResNet-50 architecture. The three-channel images made from spectrograms will be used as the input to ResNet-50. We also modify the last layer of the original ResNet-50 design so that the output for input images is a single number from zero to one, indicating the probability for the presence of chirp signals in the input spectrogram image.

To ensure the output for an input image is a number in the range [0, 1], we used the sigmoid function defined as $\sigma(x) = 1/(1 + \exp(-x))$. This function maps real numbers into the range of [0, 1], which can go on be interpreted as the probability for the presence of chirp signals. We choose a threshold of 0.8 for the identification of signals, which means the time steps with output probabilities greater than 0.8 indicate the presence of gravitational-wave signals.

The spectrograms from 8 s long strain data are produced with a blackman window size of 16,384, and a step size of 1024. The generated spectrograms have a size of 8193×113 in the frequency and time domains, respectively. As a part of the preprocessing step, we also take the element-wise log transformation of the spectrograms to accentuate the chirp patterns. All the spectrograms are produced using the spectrogram function provided by SciPy.

Since the neural networks are trained to identify inspiral-stage chirp patterns in the spectrograms, we can truncate the spectrograms above 150 Hz to reduce the size of the input images. We also remove the parts below 20 Hz, which are dominated by low-frequency noise. Therefore, the actual size of the images used as the inputs to the trained neural network is 130×113 with three channels as stated above.

Merger time estimation. The neural network for merger time estimation is the same as the one used for chirp-pattern detection, except that the last layer is a fully connected layer, and the network output corresponds to the estimated merger time.

2.4. Training Strategy

Early detection. We separately trained three ResNet-50s on eccentric binary neutron star, black hole–neutron star, and binary black hole data sets, as described in Section 2.1. We followed the same training strategy for all cases. As mentioned above, we first injected clean waveforms into advanced LIGO noise data to simulate noisy signals with different S/Ns. Then we produced spectrograms from those stimulated gravitational-wave events to generate input images for ResNet-50. Finally, we trained ResNet-50 using a batch size of 256, and a learning rate of 10^{-4} with the ADAM (Kingma & Ba 2014) optimizer. The training and testing was done using four NVIDIA V100 GPUs. For robust performance, we exposed the neural network to a variety of scenarios during the training. Specifically, 50% of the input spectrogram images contain no gravitational waves, while 25% have simulated signals only in either Livingston or Hanford data, while the remaining 25% have waveform signals in both Livingston and Hanford data. Upon thoroughly testing the performance of our three separate neural networks, we found that the network trained on binary neutron star waveforms provided the best performance for early detection for all systems under consideration. We may understand this finding if we consider that forecasting depends critically on information the neural network extracts from the inspiral phase, and in the case of black hole–neutron star systems and binary black holes most of the power is concentrated in the vicinity of the merger. Thus, in what follows we present forecasting results using the neural network trained with binary neutron star waveforms for the three classes of binaries under consideration.

Merger time estimation. For merger time estimation, the training process follows the same approach described above, except that the neural network is now trained to predict the merger time of the binary systems, and the input spectrogram images contain signals that merge at different times. We separately trained three neural networks on eccentric binary neutron star, neutron star–black hole, and binary black hole data sets, as described in Section 2.1.

2.5. Sky Localization

In addition to forecasting and quantifying the time to merger with deep learning, we have adapted an algorithm to rapidly estimate the sky localization area as a function of time, namely, from the time our neural networks identify a given signal until the merger event. The localization sky area is estimated via triangulation using a Fisher-matrix-based method, as described in Fairhurst (2009, 2011). The sky area may be estimated from the separation between the detectors, their individual effective bandwidth, and the S/Ns. The effective bandwidth, σ_f, of a detector is computed from the frequency moments as

$$\sigma_f^2 = \bar{f}^2 - \bar{f}^2, \quad \text{where}$$
$$\bar{f}^n = 4\int_0^\infty df \frac{f^n |\tilde{h}(f)|^2}{S_n(f)}, \tag{1}$$

where S_n is the noise spectral density, and $\tilde{h}$ is the frequency-domain waveform. The effective bandwidth and the S/N, $\rho = \sqrt{\bar{f}^0}$, may be combined to compute the timing accuracy, σ_t, through the relation

$$\sigma_t = \frac{1}{2\pi\rho\,\sigma_f}. \tag{2}$$

The timing accuracy may be used to construct a posterior distribution of the source location (Fairhurst 2009, 2011). In this study, we truncate our time-domain signal at different times before merger and Fourier transform the truncated signal to get the frequency-domain signal to compute the frequency moments that are then used to estimate the sky localization area as a function of time from early detection to merger. The results we present below assume a three-detector network that encompasses the Hanford and Livingston LIGO detectors and the Virgo detector. We have computed power spectral densities (PSDs) for each of these interferometers using the O2 noise segments described above.

2.6. Quantized Neural Networks

We have explored the use of quantized neural networks to enable forecasting at the edge in view of their compact size and power efficiency. To quantize our fully trained ResNet-50

models, post-training static quantization is used to convert the weights and activations of the models from 32- to 8-bit representation. We utilize quantization tools provided by `PyTorch` to perform static quantization. First we define a `ResNet-50` model and insert a quantization layer at the beginning and a dequantization layer at the end for handling the input and output tensors during inference. The weights of our trained `FP32` model are then loaded into the new model definition. Layer fusions are performed to fuse `Conv2D`, `BatchNorm`, and `ReLU` modules when possible to obtain better performance.

We prepared our networks for inference by collecting statistics for each layer input, running calibration for the quantized model, and quantizing the trained weights into `INT8`. Asymmetric linear quantization is used here to scale and offset the values in activation tensors. The calibration step adjusts the scales and offsets to minimize accuracy loss due to quantization. The spectrogram images used for this step are randomly selected from the testing data set, and only 30 images are required to fully calibrate the quantized parameters. After the networks have been quantized, we use `Intel Xeon E5-2620 CPU` with `AVX2` support as the backend for running the quantized networks. As we show below, our quantized networks have the same forecasting performance of regular neural networks, but are 4× more compact and 2.5× faster. These features promote them as ideal tools to enable gravitational-wave forecasting at the edge.

3. Results

We present results for three types of sources, binary neutron stars, black hole–neutron star systems, and binary black hole mergers. As mentioned above, our neural networks are used to search for patterns in spectrograms that characterize eccentric compact binary mergers. We used a sliding window of 8 s, with a step size of 1 s, that is applied to the spectrograms generated from strain data that are up to 160 s long. The data within the sliding window are fed into the neural networks, and the neural networks output the probability for the existence of a gravitational-wave signals in advanced LIGO noise.

3.1. Eccentric Binary Neutron Stars

Our first set of results comprise binary neutron stars with component masses $1\,M_\odot \leqslant m_{\{1,2\}} \leqslant 2.1\,M_\odot$, and eccentricities $e \leqslant 0.9$. To test the performance of our neural networks, we prepared injections that sampled a broad range of inclinations, sky locations, and S/Ns.

In Figure 2 we present three sets of results for binary neutron stars with component masses $(m_1 = m_2 = 1.4\,M_\odot)$ and $(m_1 = 2.1\,M_\odot,\ m_2 = 1.4\,M_\odot)$. In both cases we consider binaries with $\mathrm{S/N} = 30$. The top panels in Figure 2 show that our neural networks identify these signals up to 15 s before their binary components coalesce. The middle panels show that our neural networks may provide a reliable estimate for the time to merger about 10 s before the binary components collide. Finally, the bottom panels in Figure 2 show that the sky localization is rather sensitive to the eccentricity of the binary. We notice that sky localization improves by nearly two orders of magnitude from early detection up to merger for the most eccentric systems, and by three orders of magnitude for the least eccentric systems.

An extensive body of research in the literature argues strongly for the modeling of eccentric binary neutron stars, and the new physics that may be learned by detecting these sources (East et al. 2012; East & Pretorius 2012; Gold et al. 2012; Lehner & Pretorius 2014; Paschalidis et al. 2015; East et al. 2016; Chaurasia et al. 2018, 2018; Yang et al. 2018; Vick & Lai 2019, 2018; Yang 2019). This new tool provides the means to enable such observations, and to even forecast when such objects may coalesce. If flybys during the inspiral evolution produce tidal disruptions with electromagnetic counterparts (Tsang 2013), then deep learning forecasting may help associate these electromagnetic observations with the physics of eccentric neutron star systems.

3.2. Eccentric Neutron Star–Black Hole Binaries

As described above, we model black hole–neutron star binaries assuming systems with component masses $m_{\rm BH} \in [3\,M_\odot,\ 15\,M_\odot]$ and $m_{\rm NS} \in [1\,M_\odot,\ 3\,M_\odot]$, and eccentricities $e \leqslant 0.9$. As in the case of binary neutron stars, we prepared injections that sampled a broad range of inclinations, sky locations, and S/Ns.

Figure 3 summarizes our findings for injections that describe binaries with component masses $(m_1 = 5\,M_\odot,\ m_2 = 1.4\,M_\odot)$ and $(m_1 = 10\,M_\odot,\ m_2 = 1.4\,M_\odot)$ and $\mathrm{S/N} = 30$. The top panels in this figure show that forecasting is weakly dependent on eccentricity, and that systems with lower total mass may be identified up to 12 s before merger, whereas heavier systems are identified 10 s before merger. This behavior is expected due to several factors. First, forecasting results are optimal for low-mass black hole–neutron star systems because, as shown in the left panel of Figure 1, the time evolution of the whitened waveform amplitude undergoes a gradual increase as it nears merger. On the other hand, this evolution becomes more asymmetrical, characterized by a sharp amplitude growth near merger, as we consider heavier black holes. As a result, the spectrograms used to identify the existence of waveforms in advanced LIGO noise contain a wealth of information near merger where the power is concentrated. In turn, neural nets become more confident of the existence of these systems closer to merger.

The middle panels in Figure 3 show that neural networks may provide reliable information about the merger time about 6 s before merger. The bottom panels show that, as in the case of binary neutron stars, sky localization depends strongly on orbital eccentricity, and that the most eccentric systems may be better localized before merger. From the time our deep learning algorithms identify these injections, their sky localization is reduced from $\sim O(10^3)$ deg^2 for the most eccentric systems and $\sim O(10^4)$ deg^2 for the least eccentric systems, to only $\sim O(10)$ deg^2 at merger. It is worth noting that this improvement takes place within 10 s.

3.3. Eccentric Binary Black Hole Mergers

The third set of systems we considered are binary black hole mergers with component masses $m_{\rm BH} \in [3\,M_\odot,\ 15\,M_\odot]$ and eccentricities $e \leqslant 0.9$. As in the two previous cases, we used injections that sample a wide range of inclination angles, sky locations, and S/Ns.

Our findings are summarized in Figure 4 for two sample systems with component masses $(m_1 = 10\,M_\odot,\ m_2 = 5\,M_\odot)$ and $(m_1 = 12\,M_\odot,\ m_2 = 8\,M_\odot)$ and $\mathrm{S/N} = 30$.

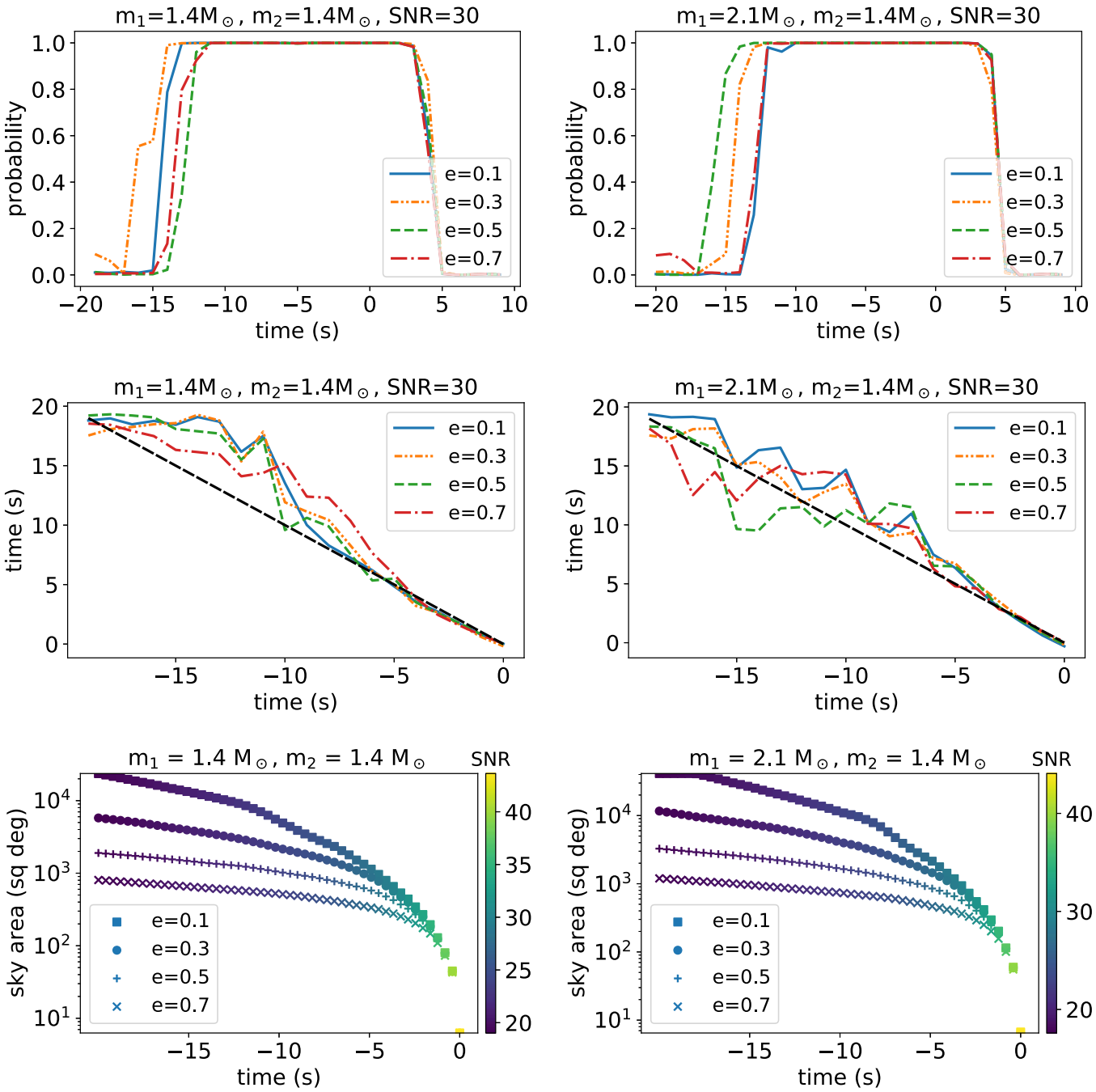

Figure 2. Top panels: neural networks identify injections of modeled binary neutron star waveforms in O2 LIGO data up to 15 s before merger. Middle panels: neural networks provide reliable estimates of time to merger 10 s ahead of the actual merger event. Bottom panels: time-dependent sky localization of injected binary neutron star waveforms in O2 LIGO noise.

The top panels in Figure 4 show that deep learning may forecast the merger of low-mass binary black hole mergers about 2 s before merger. The middle panels indicate that deep learning may indicate the time to merger about 1 s before the binary components collide. In other words, our deep learning algorithms are actually working as real-time gravitational-wave classifiers. The bottom panels show how accurately we may constrain the sky location of binary black hole mergers. These results show that we may localize these sources within $\sim O(10^2)$ deg^2 in the vicinity of merger merger, and down to $\sim O(10)$ deg^2 at merger.

4. Quantized Neural Networks for Rapid, Energy-efficient Forecasting

We compare the results of inference speed and error rate for unquantized and quantized networks trained with spectrogram images. Prior to quantization, the networks need to be calibrated using a set of testing images. We randomly select 30 images from the testing data set to yield the optimal scales and zero-points for activation tensors. We selected 1000 spectrogram images from the testing set to benchmark the performance. After quantization, the inference latency of the quantized model is 8 ms per prediction, while that of the

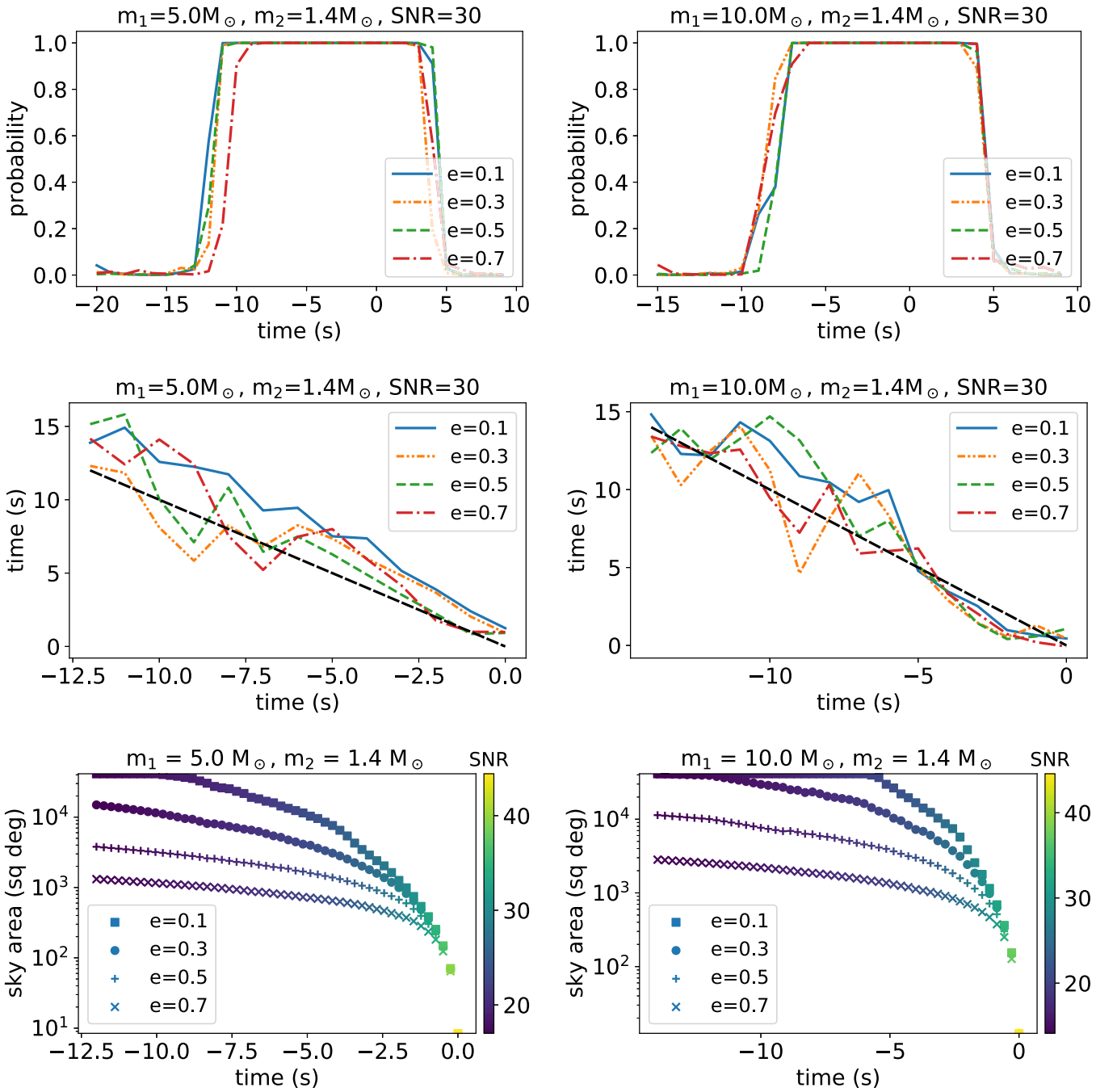

Figure 3. Top panels: neural networks identify injections of neutron star–black hole waveforms in O2 LIGO data up to 12 s before merger. Middle panels: neural networks provide reliable estimates of time to merger 5 s ahead of the actual merger event. Bottom panels: time-dependent sky localization of injected neutron star–black hole waveforms in O2 LIGO noise.

unquantized model is 20 ms per prediction, showing a 2.5× speedup. Here we used a batch number of 16, which means the network predicted 16 images at a time to exploit the parallelism of the multicore processor. By converting the model parameters from floating-point to fixed-point representation, we can reduce the overall model size by 4×, from 92 MB to 23 MB. Our results show that quantization does not negatively affect inference performance, and, in our case, it is able to decrease the top-1 error rates by 10%.

Figure 5 presents three forecasting results produced by our unquantized and quantized networks. We select one scenario from each of these three sources: binary neutron stars, black hole–neutron star systems, and binary black hole mergers. As illustrated, the forecasting results produced by the quantized network are very similar to those produced by the unquantized network, but faster. Furthermore, our quantized networks alleviate the high demand for computational and memory resources, and, as a result, the power efficiency is increased for gravitational-wave forecasting, which is critical for edge devices.

5. Conclusions

We have introduced the first application of deep learning to forecast the merger of eccentric binary systems in advanced LIGO noise. The neural networks introduced in this study can

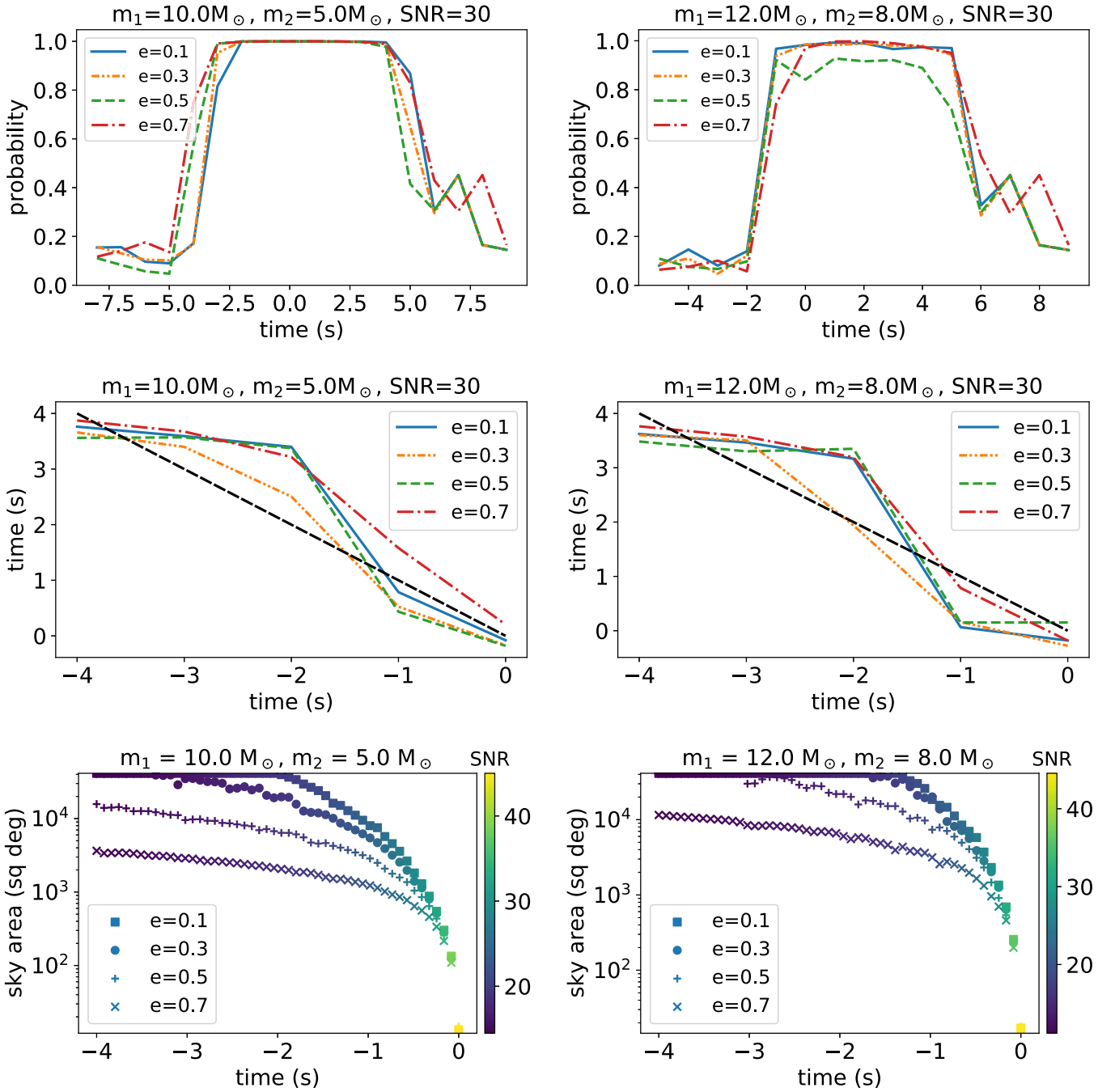

Figure 4. Top panels: neural networks identify injections of binary black hole waveforms in O2 LIGO data a few seconds before merger. Forecasting results transition into real-time alerts for these black hole mergers. Middle panels: neural networks provide a reliable estimate of the merger once the coalescence is imminent. Bottom panels: time-dependent sky localization of injected binary black holes in O2 LIGO noise.

readily identify eccentric waveforms injected in O2 noise up to 15 s before merger, and estimate the time within which the binary components will coalesce. We have adapted a sky localization method to estimate the time-dependent sky area within which we may observe these signals, from the time they are identified by our forecasting algorithms up to merger, with a three-detector network encompassing the twin LIGO detectors (Hanford and Livingston) and the Virgo detector. These results were obtained using open-source LIGO and Virgo O2 noise. Our findings indicate that the performance of our forecasting neural networks is fairly independent of the eccentricity of the binary system under consideration. However, sky localization improves significantly for larger values of eccentricity.

We quantized our neural networks and found that they improve the latency and power efficiency of gravitational-wave forecasting, making it suitable for edge computing. These algorithms may then enable a broader cross section of the community to readily use these algorithms for time-sensitive multimessenger astrophysics discovery campaigns that target compact binary systems that may reside in dense stellar environments, but that due to their complex morphology are difficult to capture with other signal-processing tools.

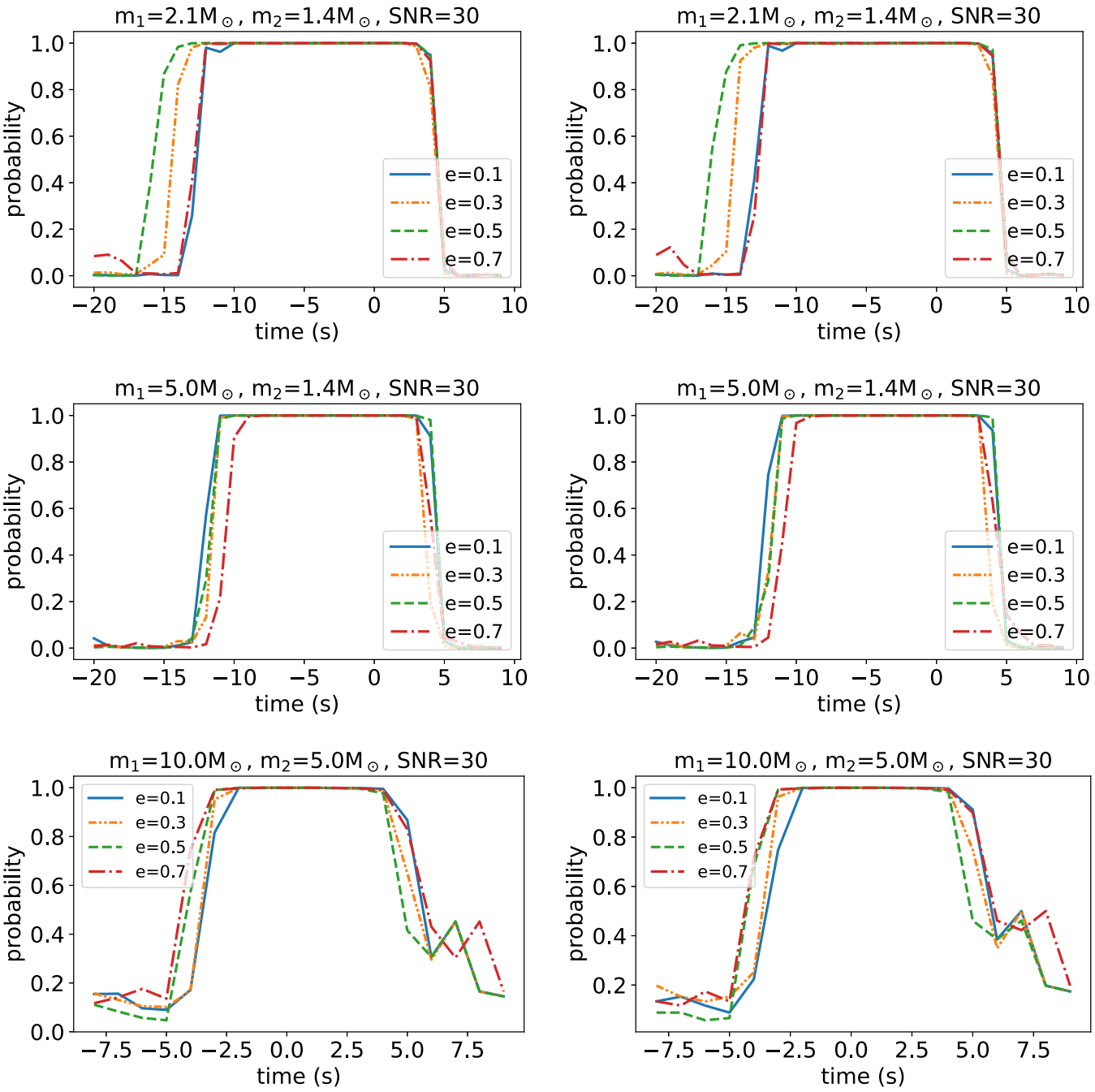

Figure 5. Deep learning forecasting for binary neutron stars, black hole–neutron star systems, and binary black hole mergers in advanced LIGO data, from top to bottom, respectively. The left panels show results produced by unquantized networks, and the right panels show results produced by quantized networks.

We thank Frans Pretorius for insightful conversations that led to the conceptualization of this project. We gratefully acknowledge National Science Foundation (NSF) awards OAC-1931561 and OAC-1934757. We thank NVIDIA for their continued support. This work utilized resources supported by the NSF's Major Research Instrumentation program, the Hardware–Learning Accelerated (HAL) cluster, grant OAC-1725729, as well as the University of Illinois at Urbana-Champaign. N.L. acknowledges support from NSF grant PHY-1912171, the Simons Foundation, and the Canadian Institute for Advanced Research (CIFAR).

Appendix A
Waveforms Used for Neural Network Training

We have mentioned in previous studies that deep learning enables the generalization to new types of signals, beyond the waveform set used for training. We have explored this assertion by comparing forecasting predictions for the binary neutron star system GW170817 using neural networks trained with the eccentric waveform model used in this analysis (East et al. 2013), and the `IMRPhenomD_NRTidal` (Dietrich et al. 2019) waveform model.

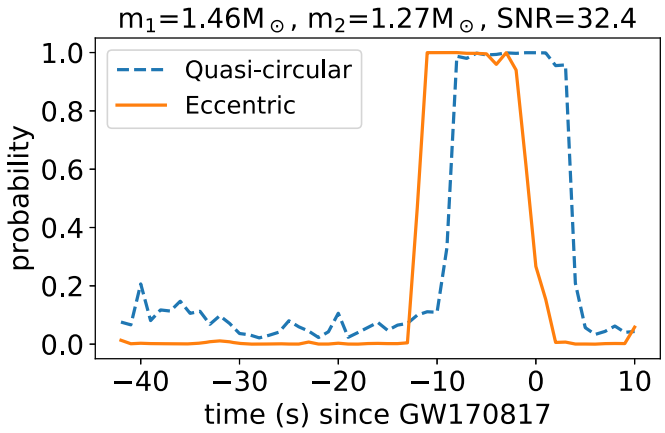

Figure A1. Forecasting of the binary neutron star GW170817 with a neural network trained with `IMRPhenomD_NRTidal` waveforms (Quasi-circular) and with eccentric waveforms produced by the model introduced in East et al. (2013) (Eccentric).

Figure A1 indicates that both neural network models have similar forecasting capabilities (Wei & Huerta 2021). These results indicate that while we should continue to use the best waveform models available to train deep learning algorithms, neural networks may also be used to enable data-driven discovery by guiding them toward the right answer with approximate models that describe complex physical processes.

Appendix B
Gravitational-wave Forecasting for the Gravitational-wave Events GW190814 and GW190412

Figure B2 presents forecasting results for the events GW190814 and GW190412. These results are consistent with short early-warning or real-time detection alerts that we discussed in Sections 3.2 and 3.3. The key point here is that the larger the total mass of the systems under consideration the closer to the merger event our forecasting algorithms identify gravitational-wave signals.

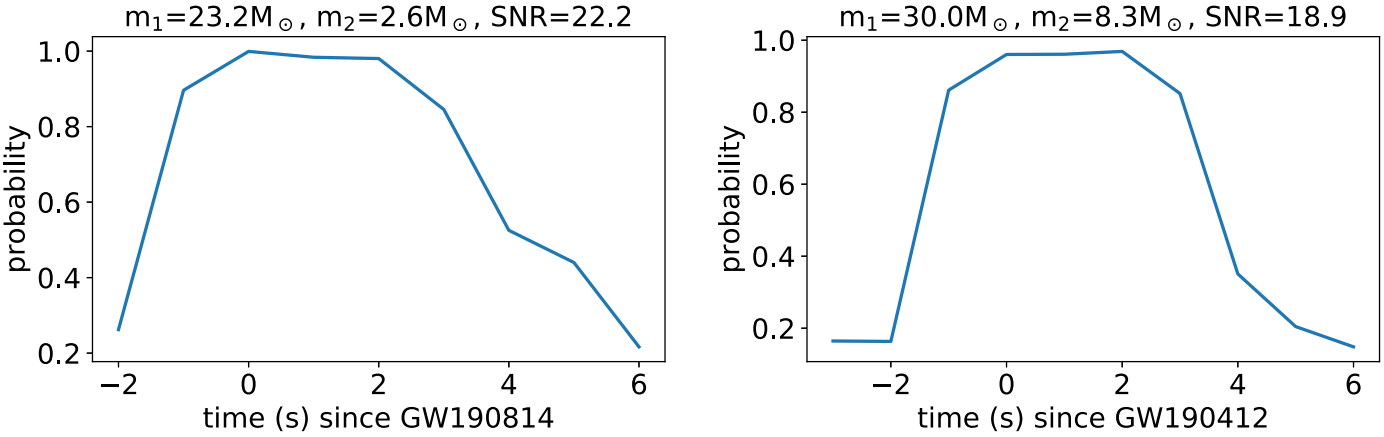

Figure B2. Application of our neural networks to forecast the merger of the events GW190814 and GW190412.

ORCID iDs

Wei Wei https://orcid.org/0000-0003-1187-2253
E. A. Huerta https://orcid.org/0000-0002-9682-3604
Mengshen Yun https://orcid.org/0000-0001-5461-636X
Nicholas Loutrel https://orcid.org/0000-0002-1597-3281
Md Arif Shaikh https://orcid.org/0000-0003-0826-6164
Prayush Kumar https://orcid.org/0000-0001-5523-4603
Roland Haas https://orcid.org/0000-0003-1424-6178
Volodymyr Kindratenko https://orcid.org/0000-0002-9336-4756

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, PhRvL, 119, 161101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, ApJL, 848, L12
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, ApJL, 848, L13
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017d, Natur., 551, 85
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019a, PhRvD, 100, 104036
Abbott, B., Abbott, R., Abbott, T. D., et al. 2019b, ApJ, 875, 161
Berti, E., Yagi, K., & Yunes, N. 2018, GReGr, 50, 46
Cannon, K., Cariou, R., Chapman, A., et al. 2012, ApJ, 748, 136
Chaurasia, S. V., Dietrich, T., Johnson-McDaniel, N. K., et al. 2018, PhRvD, 98, 104005
Chaurasia, S. V., Dietrich, T., Johnson-McDaniel, N. K., et al. 2018, PhRvD, 98, 104005
Deng, J., Dong, W., Socher, R., et al. 2009, in 2009 IEEE Conference on Computer Vision and Pattern Recognition (Piscataway, NJ: IEEE)
Dietrich, T., Khan, S., Dudi, R., et al. 2019, PhRvD, 99, 024029
East, W. E., McWilliams, S. T., Levin, J., & Pretorius, F. 2013, PhRvD, 87, 043004
East, W. E., Paschalidis, V., Pretorius, F., & Shapiro, S. L. 2016, PhRvD, 93, 024011
East, W. E., & Pretorius, F. 2012, ApJL, 760, L4
East, W. E., Pretorius, F., & Stephens, B. C. 2012, PhRvD, 85, 124009
M.S.-S. et al. 2019, ApJ, 876, L7
Fairhurst, S. 2009, NJPh, 11, 123006
Fairhurst, S. 2011, CQGra, 28, 105021
Fishbach, M., Gray, R., Magaña Hernandez, I., et al. 2019, ApJL, 871, L13
George, D., & Huerta, E. A. 2017, arXiv:1711.07966
George, D., & Huerta, E. A. 2018, PhLB, 778, 64
George, D., & Huerta, E. A. 2018, PhRvD, 97, 044039
Georgescu, I. 2020, NatRP, https://go.nature.com/2YY1NLn
Gold, R., Bernuzzi, S., Thierfelder, M., Brügmann, B., & Pretorius, F. 2012, PhRvD, 86, 121501
He, K., Zhang, X., Ren, S., & Sun, J. 2016, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE), 770
Huerta, E. A., Allen, G., Andreoni, I., et al. 2019, NatRP, 1, 600
Huerta, E. A., Khan, A., Huang, X., et al. 2021, NatAs, at press
Huerta, E. A., & Zhao, Z. 2021, arXiv:2105.06479
Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980
LeCun, Y., Bengio, Y., & Hinton, G. 2015, Natur, 521, 436
Lehner, L., & Pretorius, F. 2014, ARA&A, 52, 661
Mészáros, P., Fox, D. B., Hanna, C., & Murase, K. 2019, NatRP, 1, 585
Mooley, K., Nakar, E., Hotokezaka, K., et al. 2018, Natur, 554, 207
Nitz, A. H., Schäfer, M., & Dal Canton, T. 2020, ApJL, 902, L29
Paschalidis, V., East, W. E., Pretorius, F., & Shapiro, S. L. 2015, PhRvD, 92, 121502
Paszke, A., Gross, S., Massa, F., et al. 2019, in Advances in Neural Information Processing Systems 32, ed. H. Wallach (Red Hook, NY: Curran Associates, Inc.), 8024
Rebei, A., Huerta, E. A., Wang, S., et al. 2019, PhRvD, 100, 044025
Sachdev, S., Magee, R., Hanna, C., et al. 2020, ApJL, 905, L25
Schutz, B. F. 1986, Natur, 323, 310
Smith, K. T. 2017, Sci, 358, 1551
The LIGO Scientific Collaboration, the Virgo Collaboration et al. 2017, ApJ, 850, L39
Troja, E., Piro, L., van Eerten, H., et al. 2017, Natur, 551, 71
Tsang, D. 2013, ApJ, 777, 103
Vallisneri, M., Kanner, J., Williams, R., Weinstein, A., & Stephens, B. 2015, JPhCS, 610, 012021
Vick, M., & Lai, D. 2018, MNRAS, 476, 482
Vick, M., & Lai, D. 2019, PhRvD, 100, 063001
Wei, W., & Huerta, E. A. 2021, PhLB, 816, 136185
Wei, W., Khan, A., Huerta, E. A., Huang, X., & Tian, M. 2021, PhLB, 812, 136029
Yang, H. 2019, PhRvD, 100, 064023
Yang, H., East, W. E., Paschalidis, V., Pretorius, F., & Mendes, R. F. P. 2018, PhRvD, 98, 044007