# Sample-efficient non-Gaussian noise reduction in gravitational wave data via learnable wavelets

Arush Pimpalkar,[1,2] Digvijay Wadekar,[1,3] Mark Ho-Yeuk Cheung,[1,4] and Emanuele Berti[1]

[1]*William H. Miller III Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland, 21218, USA*

[2]*Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli 620015, India*

[3]*Weinberg Institute, University of Texas at Austin, Austin, Texas 78712, USA*

[4]*School of Natural Sciences, Institute for Advanced Study, Einstein Drive, Princeton, New Jersey, 08540, USA*

(Received 26 January 2026; accepted 28 May 2026; published 23 June 2026)

We introduce WaveletNet, a wavelet-based neural-network architecture for identifying non-Gaussian noise in gravitational-wave data and reducing its impact on candidate ranking. Traditionally, convolutional neural networks (CNNs) have been widely used as a flexible machine learning method to mitigate non-Gaussian noise. However, training CNNs requires many data samples, especially when the input data segments are long. Glitches that mimic high-mass black hole signals are empirically known to have a waveletlike structure. We exploit this property in WaveletNet by using simple neural networks to learn the best family of wavelets to model glitches in the LIGO-Virgo-KAGRA O3 data. Due to its simplicity, our framework is significantly more sample-efficient than CNNs. As a use case, we build upon the trigger inference using the extended strain representation (TIER) method from [arXiv:2507.08318] and show WaveletNet can improve the performance of search pipelines. We take potential GW candidates from the pipeline, and then down-weight the candidates having noisy strain regions in their vicinity [i.e., within $t_{\mathrm{candidate}} \pm \mathcal{O}(10)$ sec]. We use our framework in a modular way: we provide an output score that can be added to the pipeline's existing detection statistic for the candidates. We test our method using candidates from the IAS-HM search pipeline and find peak improvements of $\sim$15% in the search sensitive volume, concentrated at high total masses and asymmetric mass ratios.

DOI: 10.1103/33k8-331b

## I. INTRODUCTION

Gravitational waves (GWs) are waves on the fabric of spacetime predicted by Einstein's theory of general relativity [1]. They are produced by some of the most violent processes in the Universe, such as the inspiral and merger of binary black holes [2]. The LIGO, Virgo, and KAGRA detectors [3–5] now routinely measure such merger events. Their sensitivity is limited by many noise sources, whose characterization is an essential part of detector operation and data-quality validation [6–10]. Some of these, known as glitches, are non-Gaussian transients that are difficult to remove. Glitches can resemble real GW signals, making it harder to establish the astrophysical origin of some candidates [5,11–13].

Multiple GW search pipelines operate in parallel on interferometer data, using different strategies for signal extraction and background estimation. For example, PyCBC and GstLAL use compact-binary-coalescence template banks to identify coincident triggers across detectors [14], and rank candidates with likelihood-ratio statistics [15–17]. The coherent WaveBurst (cWB) pipeline is an unmodeled search that reconstructs transient signals through coherent excess-power analysis [18–20]. The multiband template analysis (MBTA) pipeline reduces computational cost by splitting matched filtering into multiple frequency bands [21]. The IAS-HM pipeline extends standard matched-filter searches by incorporating higher-order harmonic modes beyond the dominant quadrupole, improving sensitivity to asymmetric and high-mass binaries [22,23]. It performs mode-by-mode filtering and combines the resulting signal-to-noise ratios (SNRs) using physically informed ranking statistics, yielding substantial gains in detection reach [24].

Low-mass compact binary coalescences produce long-duration signals that spend many waveform cycles within the detectors' sensitive band. Although non-Gaussian transients occur across all frequency bands, a single short-duration glitch is unlikely to mimic or mask the full morphology of a long signal. This distinction helps motivate signal-consistency tests that reduce the glitch background for low-mass systems [14,15,25–27]. In contrast, high-mass systems produce short in-band signals whose duration can be comparable to that of a typical glitch. These signals are therefore more susceptible to transient contamination [5,26,28–33], which reduces the discriminating power of standard detection statistics and increases the false-alarm rate. Improved analysis methods are therefore needed to distinguish genuine high-mass signals from transient noise backgrounds.

Flexible machine-learning (ML) architectures such as convolutional neural networks (CNNs) and transformer-based models have been widely applied to the identification of non-Gaussian noise in GW data, see, e.g., [11,34–61]. Closely related applications include supervised and transfer-learning classifiers for glitch morphology [62–74], signal-glitch discrimination and search classification [75–79], and low-latency denoising or noise regression [44,80]. These approaches are highly expressive: they can model complex nonlinear relationships and hierarchical features in the data. However, they typically require large training datasets, particularly for long time-domain segments of tens of seconds or more [38,52]; see Appendix A 1 for a demonstration in our setting. This data demand can be limiting when representative examples of specific glitch morphologies are scarce.

A class of glitches that contaminate searches for high-mass binary black-hole mergers is empirically observed to produce localized, short-duration excess power in the data [11,28,29,81,82], a signature well suited to wavelet-based detection. We therefore use wavelets as a flexible basis for representing these glitch morphologies. The method is best suited to glitches that admit approximately sparse wavelet representations, as is often the case for the short-duration transients most relevant to high-mass binary searches. More complex or irregular glitches may require a superposition of many wavelets; for such cases, other parametrized families, such as chirplets, shapelets, or learned dictionary atoms, could be incorporated into the same framework as replacements for the wavelet basis.

We introduce an explicit inductive bias by modeling glitches with a parametrized wavelet family, the Morlet wavelets, $\psi(t) = e^{-t^2/2\sigma^2} e^{i\omega_0 t}$. Rather than learning arbitrary features from raw data, we use a lightweight multilayer perceptron (MLP) to infer physically interpretable wavelet parameters, such as central frequency and temporal duration. This approach trades architectural flexibility for data efficiency and interpretability, while remaining expressive enough to capture the dominant glitch morphologies relevant to high-mass searches.

In this work, we design a network, WaveletNet, to characterize noisy strain data in the vicinity of gravitational-wave candidates. WaveletNet uses learnable Morlet wavelets as glitchlike templates and performs matched filtering over an extended region surrounding each trigger. By using a wavelet basis to identify localized regions of excess power, the model gains an inductive bias that is more sample efficient than generic CNNs for the nonlocal data considered here; see Appendix A 1. We apply the framework to glitches in data from the LIGO-Virgo-KAGRA O3 observing run. Wavelet-based approaches have also been used in other areas of astrophysics and cosmology to obtain robust inference and improved data efficiency relative to generic deep neural networks (e.g., Ref. [83]).

When integrated modularly into the IAS-HM search pipeline, WaveletNet yields improvements of up to 15% in sensitive spacetime volume, demonstrating its usefulness as a complementary ranking component rather than a replacement for existing pipelines. This builds on the TIER methodology introduced in Ref. [84]. Some instrumental glitches have been empirically observed to “come with friends,” meaning that they cluster in time and occur during periods of elevated noise; scattered-light glitches are a common example [85]. Related behavior has been used manually to reduce contamination in the GWTC-4 catalog; see Fig. 9 of Ref. [86]. This clustering is not universal across all glitch morphologies, but it is relevant for the class of short-duration, broadband transients targeted here, as reflected in the nonlocal summary statistics used in our pipeline; see Sec. V B. Previous search pipelines have evaluated candidate significance using only a narrow ($\sim \pm 0.1$ s) window around the trigger [14,15,26,87], roughly set by the template duration or autocorrelation length. As a result, informative structure in the surrounding strain data is ignored. WaveletNet addresses this limitation by producing an external score from nonlocal data that can be modularly combined with the existing ranking statistic.

The paper is organized as follows. In Sec. II, we describe the architecture of WaveletNet. Section III presents tests of our methodology on a simplified toy dataset. In Sec. IV, we discuss how extended strain information is incorporated into the search ranking statistic. Section V describes the data used to train the ML models and the summary statistics employed to compress the extended strain information. In Sec. VI, we present the resulting sensitivity improvements of the search pipeline, and in Sec. VII, we discuss the implications of our methodology.

### A. Motivation for an inductive-bias-driven architecture

Our methodology is guided by the principle that inference performance can be improved by introducing an explicit inductive bias based on domain knowledge, rather than relying entirely on data-driven feature learning. In many signal detection problems in physics, the signals of interest have partially known structure, while the main challenge comes from noise and limited data.

Widely used deep-learning approaches, such as CNNs, treat this as a representation-learning problem in which the model must infer both the signal structure and the decision boundary from raw data. While flexible, this approach typically requires large models and datasets, and offers limited physical interpretability. In contrast, classical signal processing provides established tools, such as matched filtering, that are optimal for improving SNRs when the signal shape and noise distribution are partially or fully

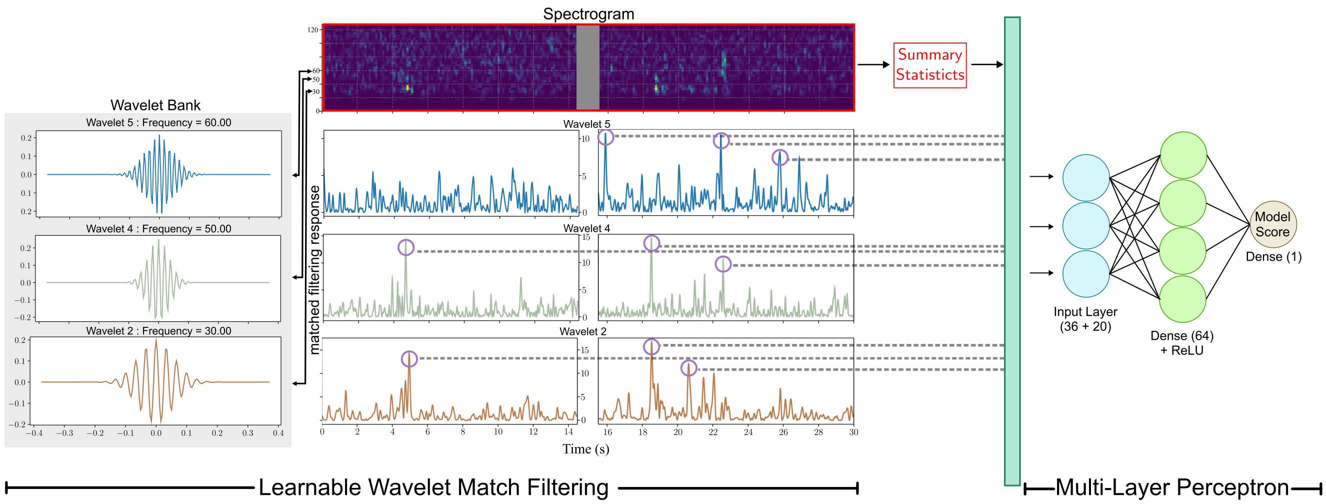


FIG. 1. WaveletNet model architecture for evaluating extended strain information near a GW candidate. Extended strain data in the $\pm\sim 15$ s region surrounding each candidate is filtered through a bank of learnable Morlet wavelets, which act as matched filters and produce time-domain responses sensitive to transient structure in the environmental data. From each wavelet response, we extract the three largest peaks and their time indices, yielding 36 wavelet-derived features across six wavelets. These features are concatenated with 20 summary statistical features (see Sec. V B), forming a 56-dimensional input feature vector. This vector is passed to a multilayer perceptron with one hidden layer of 64 units and a scalar output, which produces the final model score used for candidate ranking.

specified. Our approach combines these ideas by inserting flexible signal-processing structure into the matched-filtering step of the pipeline.

## II. WAVELETNET ARCHITECTURE

The WaveletNet architecture has two main components: a matched-filtering module based on learnable Morlet wavelets, and a MLP that combines the resulting features with summary-statistics inputs to produce a final candidate score. The workflow is shown in Fig. 1. The following sections describe each component.

### A. Learnable wavelet matched filtering module

Glitches that resemble high-mass binary-black-hole signals are empirically observed to exhibit localized time-frequency structure that can be represented by wavelets [81,88–90]. We use this property to learn a compact wavelet basis for identifying glitchy nonlocal data with WaveletNet.

The model performs matched filtering on strain data using the learnable Morlet wavelets. We initialize six Morlet wavelets with center frequencies spanning 15 Hz to 60 Hz, the band most relevant for BBH-mimicking glitches in the O3 data used here. The number of wavelets was chosen empirically: adding more did not improve performance. The lower edge avoids the seismic-noise-dominated regime below 15 Hz, while relatively few glitches in this sample exhibit substantial power above 60 Hz. The Morlet wavelets are defined as

$$\psi(t) = \pi^{-1/4} \cdot e^{i2\pi f t} \cdot e^{-t^2/2}. \tag{1}$$

The wavelet is characterized by its frequency, scale, and duration, with scale and duration intrinsically related. In our model, the scale is the only learnable parameter.[1] The central frequency and duration are held fixed to reduce computational cost, because varying them does not yield measurable performance gains in the O3 data. After scaling, the Morlet wavelet is

$$\psi_s(t) = \frac{1}{\sqrt{s}} \cdot \psi\left(\frac{t}{s}\right). \tag{2}$$

We trained the scale parameters in logarithmic space to prevent convergence to zero. For example, the trained wavelets are visualized in Fig. 2.

Matched filtering of the wavelets against the data is implemented with the CONV1D function from PyTorch. The matched-filter output $y(t)$ is obtained by convolving the input strain data $d_{\mathrm{nonlocal}}(t)$ with the scaled Morlet wavelet $\psi_s(t)$,

$$y(t) = (d_{\mathrm{nonlocal}} * \psi_s^*)(t). \tag{3}$$

To maximize the matched-filter response, we apply analytical phase maximization [87,92], a standard technique in gravitational-wave searches that removes sensitivity to the

[1] The wavelet scale parameter used here is related to the $Q$-transform parametrization of Morlet wavelets discussed in [91], where the quality factor $Q$ determines the trade-off between time and frequency resolution. The two parametrizations are equivalent; we adopt the scale convention as it maps directly onto our implementation.

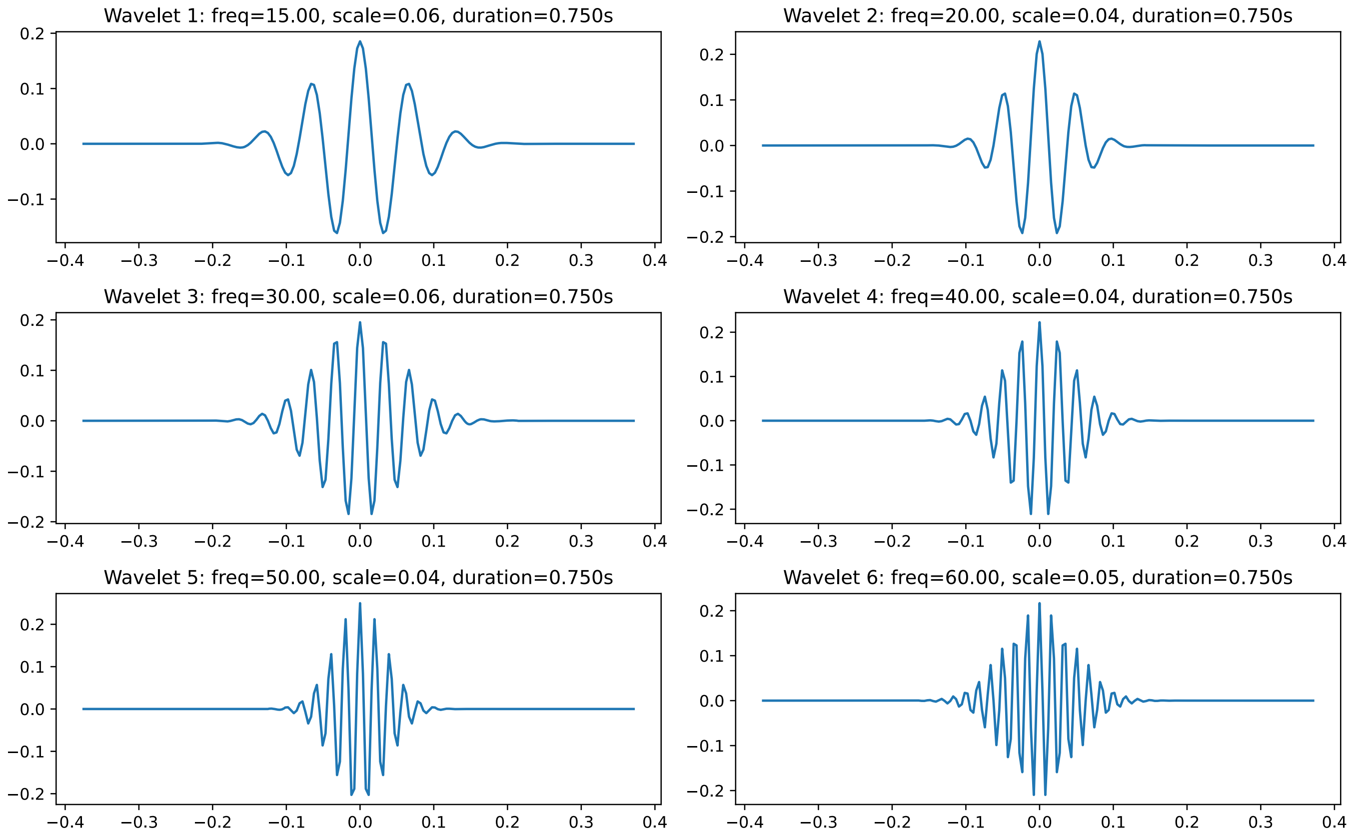


FIG. 2. Learned wavelets corresponding to a high-mass template bank (GW signals with total mass $M_{\mathrm{tot}} \approx 250 M_{\odot}$). Only the scale parameter of each wavelet is trained. The central frequency and duration are held fixed to reduce computational cost, because varying them does not yield measurable performance gains in the O3 data. The resulting learned waveforms effectively function as a bank of glitch templates for this mass range.

arbitrary phase offset between the template and the signal. For a complex-valued filter output $y(t)$, this is equivalent to projecting the complex response onto its optimal phase by rotating $y(t)$ to align along the real axis, yielding the magnitude of the output,

$$\tilde{y}(t) = |y(t)| = \sqrt{\mathrm{Re}(y(t))^2 + \mathrm{Im}(y(t))^2}. \tag{4}$$

This maximizes the response over the unknown coalescence phase without requiring an explicit grid search over it.

Instead of passing the full response to the MLP, we extract the three highest peaks of $\tilde{y}(t)$ and their time indices. Retaining three peaks keeps the feature dimension small and is sufficient empirically, since our data samples rarely contain more than three prominent glitches. These peaks are the loudest glitchlike responses identified by the wavelet filter, providing a compact summary of the strongest transient activity in the data.

Peaks are identified by segmenting the response into 0.1 s intervals and selecting the maximum value within each segment. This representation yields improved classification performance while significantly reducing the training time.

In Fig. 3 we compare the matched-filtering SNR responses of representative wavelets to glitchy and nonglitchy strain data. For the glitchy strain, the wavelet responses show distinct SNR peaks at times corresponding to non-Gaussian transients visible in the spectrograms. In contrast, the responses for the nonglitchy strain remain lower and lack coherent peaks, indicating the absence of significant transient noise. This contrast shows that the learned wavelets are selectively sensitive to glitchlike structure in the nonlocal data.

### B. MLP classification head

The MLP takes as input the concatenation of wavelet-derived features and nonlocal summary statistics, as discussed in Sec. V B. From each wavelet response, the three largest peaks and their corresponding time indices are extracted, yielding a total of 36 wavelet-based features across the full wavelet bank. These features are concatenated with 20 summary statistical features, forming a 56-dimensional input vector. The input is processed by a fully connected layer with rectified linear unit (ReLU) activation, followed by a final linear layer that produces a scalar output score.

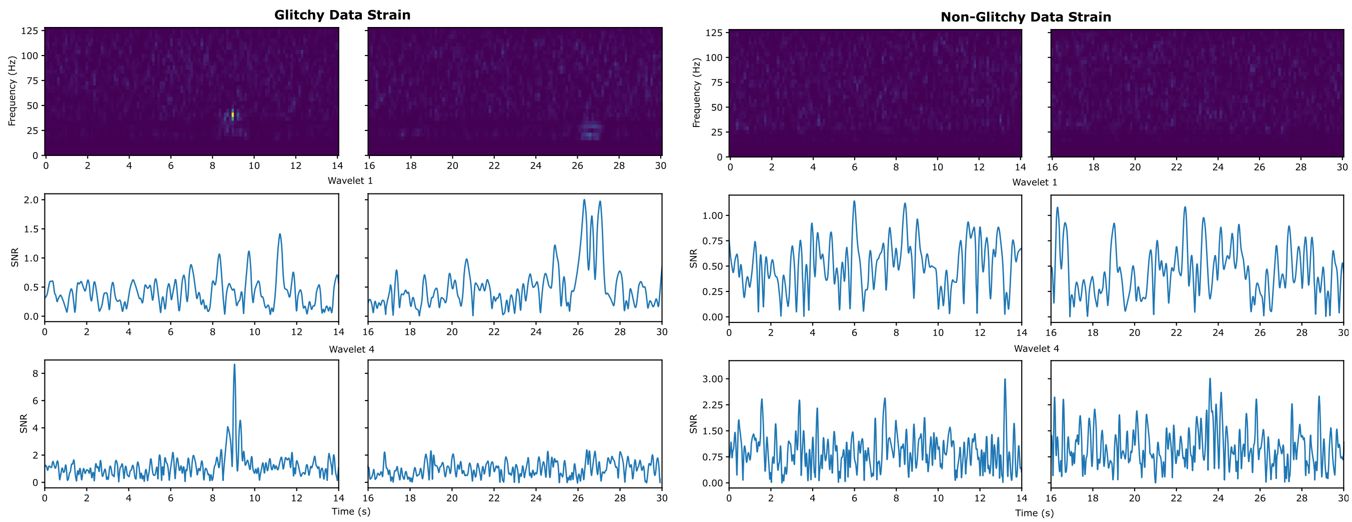


FIG. 3. Examples of SNR responses of the wavelet with a glitchy data strain (left panels) and a nonglitchy data strain (right panels). The glitchy strain, which contains stronger non-Gaussian noise, produces higher wavelet-response SNR peaks than the nonglitchy strain. The panels use the same SNR scale to facilitate direct comparison.

Training is performed using the Adam optimizer with a learning rate of $10^{-4}$ and binary cross-entropy loss with logits. We monitor the validation loss for early stopping and adaptive learning-rate reduction using a ReduceLROnPlateau schedule, retaining the model state with the lowest validation loss.

The ML component is designed to augment the traditional matched-filtering search pipeline, rather than replace it. The glitches used for training are noise transients that produce triggers in the existing GW template banks, and therefore span a range of morphologies and characteristic frequencies. Capturing this diversity with a single wavelet model is challenging. We therefore train separate model instances for triggers from different GW template banks; the banks are approximately partitioned by total mass, as described in Ref. [93].

Since each template bank corresponds to a characteristic range of total masses and signal durations, glitches that trigger a given bank tend to share similar time-frequency morphologies. We therefore train a separate WaveletNet model for each bank, using only glitches identified by templates from that bank. During inference, each candidate is associated with the template that produced it, and the WaveletNet model trained on the corresponding bank is used to compute the probability that the candidate is a glitch.

## III. PERFORMANCE ON A TOY DATASET

To compare CNNs with WaveletNet, we perform a toy experiment using a generic CNN and a simplified version of the wavelet-based model described in Sec. II. The dataset is a synthetic time-series dataset designed to study binary signal detection under controlled conditions. Each sample consists of a 10 s, one-dimensional time series sampled at 2048 Hz. There are two classes: pure Gaussian noise, and Gaussian noise containing a weak, localized signal.

The signal is a Haar wavelet modulated by a Gaussian envelope and randomly positioned within the time series to test temporal invariance. Although both Haar and Morlet wavelets are compactly supported, they are morphologically distinct: Haar wavelets are piecewise-constant step functions, whereas Morlet wavelets are smooth sinusoidal oscillations. The model's performance advantage is therefore not a trivial consequence of template-signal identity. The signal amplitude is chosen so that the prefiltering SNR is 0.6, placing the signal well below the noise floor and making it indistinguishable by eye. This challenging regime tests whether each model can exploit time-frequency structure rather than raw amplitude. The matched-filtering step in WaveletNet provides a substantial SNR boost, enabling effective detection despite the low prefiltering SNR. All datasets are class balanced and generated deterministically, with multiple dataset sizes used to assess scaling with training data.

We first apply matched filtering with learnable wavelets to emphasize relevant time-frequency structure. From the resulting matched-filtering responses, we extract a small set of summary statistics, including the mean, standard deviation, range, and peak prominence, to characterize signal strength and distinctiveness. This structured representation reduces model complexity, stabilizes training, and preserves interpretability while retaining the benefits of ML-based optimization.

Figure 4 shows the validation performance of the two models as a function of training-set size and training epoch. The wavelet-based model achieves higher validation accuracy with fewer training samples and exhibits more stable convergence. In contrast, the CNN improves more slowly

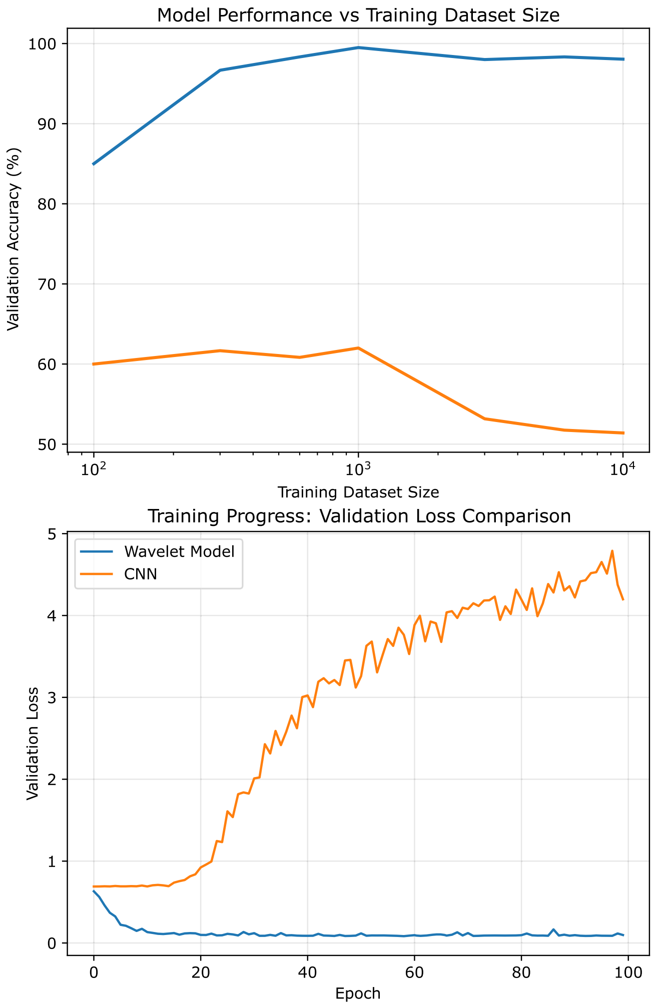


FIG. 4. Comparison of a wavelet-based neural network and a CNN on a toy-model dataset (see Sec. III for details). This figure illustrates the effect of the inductive bias used in WaveletNet on data efficiency and training stability. Top panel: validation accuracy as a function of training-set size for a wavelet-based model and a CNN. The wavelet-based model achieves higher accuracy with fewer training samples. Bottom panel: validation loss as a function of training epoch for the same models, obtained using a fixed training set of 400 examples per class. The wavelet-based model exhibits stable convergence, while the CNN shows larger fluctuations. These results qualitatively demonstrate the advantage of incorporating signal-processing structure in limited-data regimes.

with dataset size and shows larger fluctuations in validation loss, consistent with higher data requirements and greater sensitivity to overfitting. These trends qualitatively demonstrate how incorporating problem-specific structure can improve data efficiency and training robustness, even in simplified settings. At very small training-set sizes, stochasticity in the training procedure can be significant; this likely explains the change in relative behavior in the left panel of Fig. 4. A Jupyter notebook implementing the toy-model experiments is publicly available at [94].

More broadly, synergies between wavelet-based representations and flexible deep learning architectures have been explored in the machine learning literature, demonstrating that combining explicit time-frequency structure with learnable models can improve robustness and generalization in limited-data regimes [83,95–98].

## IV. INCLUDING ENVIRONMENTAL INFORMATION IN THE SEARCH RANKING STATISTIC

When searching for GW signals, each search pipeline finds a number of signal candidates. To determine the significance of the candidates, a ranking score/statistic is assigned to each candidate. The optimal ranking score is given by the Neyman-Pearson lemma, and corresponds to the ratio of the likelihood of the candidate being either a signal ($\mathcal{S}$) or noise ($\mathcal{N}$),

$$\mathcal{L} \equiv \frac{P(d|\mathcal{S})}{P(d|\mathcal{N})} \equiv \frac{P(d_{\rm local}, d_{\rm nonlocal}|\mathcal{S})}{P(d_{\rm local}, d_{\rm nonlocal}|\mathcal{N})}, \tag{5}$$

where $d_{\rm local}$ is the region within $\pm 0.1$ s of the candidate, and $d_{\rm nonlocal}$ extends to $\pm 15$ s around the candidate, as described below. Expanding Eq. (5), we obtain

$$\mathcal{L} \equiv \frac{P(d_{\rm local}|\mathcal{S})}{P(d_{\rm local}|\mathcal{N})} \frac{P(d_{\rm nonlocal}|d_{\rm local}, \mathcal{S})}{P(d_{\rm nonlocal}|d_{\rm local}, \mathcal{N})}. \tag{6}$$

Calculating $P(d_{\rm nonlocal}|d_{\rm local}, \mathcal{N})$ is challenging because of its non-Gaussian and nonstationary nature. Therefore, search pipelines often use the approximation

$$\mathcal{L}_{\rm pipeline} \approx \frac{P(d_{\rm local}|\mathcal{S})}{P(d_{\rm local}|\mathcal{N})}, \tag{7}$$

which effectively assumes the second fraction to be unity. This corresponds to treating the noise far from the candidate event as statistically independent of the local data $d_{\rm local}$.

However, as mentioned above, glitches can "come with friends": if $d_{\rm local}$ contains a glitch (i.e., the true hypothesis is $\mathcal{N}$), the surrounding data $d_{\rm nonlocal}$ can also exhibit elevated noise levels. Consequently, the standard approximation is not strictly valid, since the statistical properties of $d_{\rm nonlocal}$ need not resemble those of a typical Gaussian background.

It is worth clarifying the role of $d_{\rm nonlocal}$ in this framework. The nonlocal environment is not used to refine parameter estimation of the candidate itself, but to inform candidate ranking. When a candidate occurs during elevated glitch activity, it is more likely to be a noise transient mimicking a signal and is therefore down-ranked. This increases the separation in the ranking statistic between genuine astrophysical events and glitch-induced triggers, improving overall pipeline sensitivity. A nonlocal window may also contain a glitch unrelated to the candidate. In such cases a true signal could be mildly down-weighted, but the cost in sensitivity is small because such coincidences are rare compared with the gain from correctly down-weighting glitch-induced triggers. This asymmetry between signals and glitches, together with the known nonstationarity and nonlinear coupling of detector noise on timescales from seconds to months [12,99], motivates the inclusion of $d_{\mathrm{nonlocal}}$ in the ranking statistic, and is empirically supported by the gain in sensitive spacetime volume reported in Sec. VI.

We use our ML method to address this issue; we learn the distribution of non-Gaussian noise from the detector data and down-weight non-Gaussian transients. As shown schematically in Fig. 5, the resulting ML score can be modularly combined with the traditional pipeline statistic to further calibrate the ranking score,

$$\mathcal{L}_{\mathrm{new}} \simeq \mathcal{L}_{\mathrm{pipeline}} \times \prod_{i\in\mathrm{det}} \frac{p_i^{\mathrm{ext}}}{1-p_i^{\mathrm{ext}}}. \tag{8}$$

Here $p_i^{\mathrm{ext}}$ is the classifier output trained to estimate $p_i^{\mathrm{ext}} = p(d|\mathcal{S})/[p(d|\mathcal{S}) + p(d|\mathcal{N})]$ in detector $i$. The ratio $p_i^{\mathrm{ext}}/(1-p_i^{\mathrm{ext}})$ therefore yields $p(d|\mathcal{S})/p(d|\mathcal{N})$, providing a direct estimate of the likelihood ratio contribution from the nonlocal data that can be consistently combined with

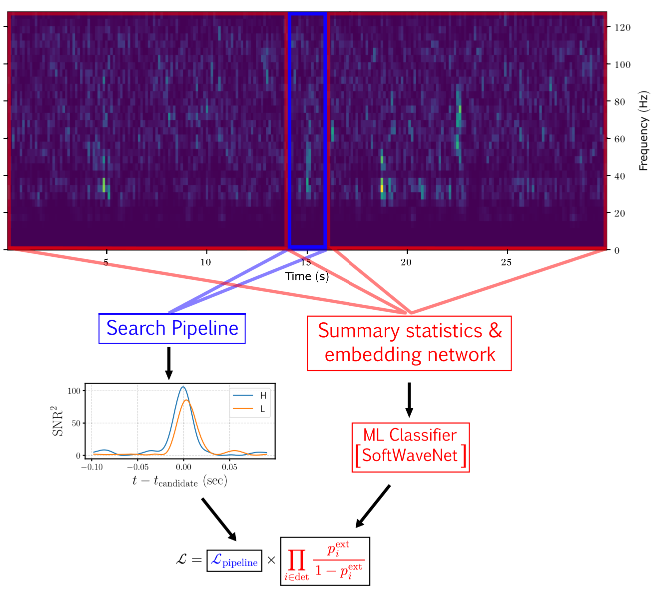


FIG. 5. Ranking statistic of a typical search pipeline ($\mathcal{L}_{\mathrm{pipeline}}$) uses local SNR from matched filtering to sort or rank a candidate event by its likelihood of being an astrophysical signal rather than noise. The TIER framework can be used to augment the ranking statistic of a pipeline by including complementary information about the extended strain data next to the candidate time [see Eq. (5)]. We thus obtain the probability that the candidate is an astrophysical signal ($p^{\mathrm{ext}}$) based on the extended strain data and include it in the candidate ranking statistic.

$\mathcal{L}_{\mathrm{pipeline}}$. The new likelihood $\mathcal{L}_{\mathrm{new}}$ should approximate the true likelihood better than $\mathcal{L}_{\mathrm{pipeline}}$, as calculated in Eq. (7).

## V. TRAINING DATA

In this section, we describe the input data streams used to train WaveletNet and obtain $p_i^{\mathrm{ext}}$. WaveletNet uses two inputs: a whitened strain time series from the nonlocal region surrounding each candidate, and a vector of statistical features summarizing the candidate's data environment. Together, these inputs provide complementary contextual and summary information. Separate models are trained for each of the 17 binary-black-hole template banks, whose construction and boundaries are described in Sec. II of Ref. [100]. For each bank, the models are trained using approximately 15,000 injection examples and 23,000 glitch examples.

### A. Whitened time series

We extract 30 s time-series segments from the nonlocal region surrounding each candidate event. These segments exclude the immediate vicinity of the trigger, ensuring that the model captures background noise characteristics rather than features of the candidate itself. By providing raw contextual information about the surrounding noise environment, the time-series input enables WaveletNet's learnable wavelet filters to characterize noise properties relevant for classification.

### B. ADDITIONAL NONLOCAL STATISTICAL FEATURES OF THE GW CANDIDATES

In addition to the time series input described above, we incorporate a set of summary statistical features derived from the nonlocal data surrounding each candidate. These features, provided by the IAS-HM search pipeline, offer a compact description of the broader data environment and are intended to complement the information extracted directly from the strain time series.

We present below the statistical features computed from the nonlocal data. These features were developed empirically to capture aspects of the surrounding data that may indicate elevated noise levels or nonstationarity. Similar summary statistics are used in the random-forest models described in Ref. [84]. More informative summary statistics, or ML-based compression of the environmental data, may further improve classifier performance.

(1) $\Delta t_{\mathrm{triggers}}$: This quantity denotes the time interval between the candidate and the other neighboring triggers, in the range of $0.5\ \mathrm{s} < |t - t_{\mathrm{candidate}}| < 15\ \mathrm{s}$, in each detector's catalog picked up in the matched filtering stage of the pipeline. We only include binary black holes with individual masses $>10M_{\odot}$, which safely have duration shorter than 0.5 s, to avoid signal overlap.

(2) SNR of nearby triggers: We include information about the SNRs of neighboring triggers observed in each detector. We use full-harmonic SNRs calculated as $|\rho^2_{\rm HM}| \equiv |\rho_{22}|^2 + |\rho^{\perp}_{33}|^2 + |\rho^{\perp}_{44}|^2$, where $\rho^{\perp}_{33}$ and $\rho^{\perp}_{44}$ correspond to the SNR of the orthogonalized higher-harmonic templates. We rescale these values to reduce the numerical range of the inputs for the ML models.

(3) Template ID of the neighboring triggers: In the IAS pipeline, the template ID corresponds to singular vector components ($c_0$, $c_1$) [101,102]. Glitches are caught more frequently by certain templates (short-duration waveforms) and providing this information to the ML model could be useful.

(4) $\int \frac{|h_{\rm norm}(f)|^2}{{\rm PSD}_{\rm local}(f)} df$: This quantity measures the sensitivity of a given local strain data segment to gravitational waveforms from a particular template bank. Essentially, this quantity penalizes data segments with high power spectral density (PSD) values, often dominant in the segments with gravitational strains.

(5) $\Delta t_{\rm hole}$: Distance to the nearest "hole." These are the regions excised by the IAS-HM search pipeline. They correspond to very loud noise transients, and are expected to have SNR $\gtrsim 20$. We include this parameter separately from the nearby triggers, as these data regions do not have an accurate SNR estimate in our pipeline (as we inpaint such loud noise data segments before estimating the PSD).

(6) $N_{\rm noisy\,bands}$: These are the number of noisy frequency bands/channels overlapping with the triggers. We use results from the `Band Eraser` tool from Ref. [100] to estimate this quantity. The `Band Eraser` tool removes individual noisy segments with dimensions 64 s × 2 Hz in the spectrogram of the strain data.

## VI. SEARCH PIPELINE SENSITIVITY IMPROVEMENT

We apply the trained WaveletNet models to candidates produced by the IAS-HM search pipeline, computing the external probability $p^{\rm ext}$ for both background and coincident events. These values are combined with the traditional ranking statistic using Eq. (8) to form a new ranking statistic. This reranking modifies the false alarm rate (FAR) and the inferred astrophysical probability, $p_{\rm astro}$, assigned to each candidate.

As a first diagnostic of the trained models, we examine the learned wavelet filters. In Fig. 2 we show the trained wavelets for one representative template bank, BBH-14, corresponding to total masses of $M_{\rm tot} \approx 250 M_\odot$. These learned wavelets can be interpreted as data-driven templates for the short-duration, clustered transients targeted by our method in the detector data.

To translate this wavelet-level sensitivity into an impact on search results, we compute an external probability score, $p^{\rm ext}$, for each coincident candidate identified by the IAS-HM pipeline. For a given candidate, the learned wavelet responses and nonlocal summary statistics are processed by WaveletNet to produce $p_i^{\rm ext}$ for each participating detector. These scores are then combined with the traditional pipeline ranking statistic $\mathcal{L}_{\rm pipeline}$ according to Eq. (8) to form a reranked statistic $\mathcal{L}_{\rm new}$, yielding an updated ordering of candidates.

In Fig. 6 we illustrate the resulting changes in inverse false-alarm rate (IFAR) for candidates in the BBH-14 template bank obtained using this method. Here, the old IFAR refers to the original ranking produced by the IAS-HM pipeline without WaveletNet, and we plot the ratio of the updated IFAR to the original value after incorporating $p^{\rm ext}$. For many candidates, particularly those with

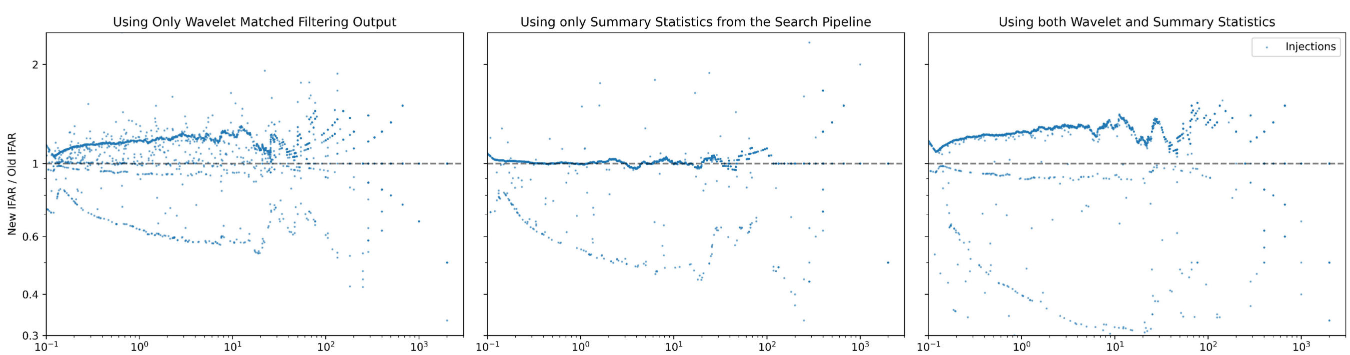


FIG. 6. Two classes of inputs to WaveletNet are the SNR responses obtained from matched filtering with the learned wavelets, and the additional summary statistics provided by the search pipeline (see Fig. 1). We illustrate the relative importance of these input classes by quantifying their effect on the IFAR improvement for injections in the data. The left panel shows the performance obtained using only the wavelet matched-filtering outputs. The center panel shows the IFAR improvement obtained using only the summary statistics from the search pipeline (see Sec. V B). The right panel shows the full WaveletNet model, which combines the wavelet-based features with the pipeline summary statistics. The wavelet-based features provide a stronger contribution to the IFAR improvement than the summary statistics alone.

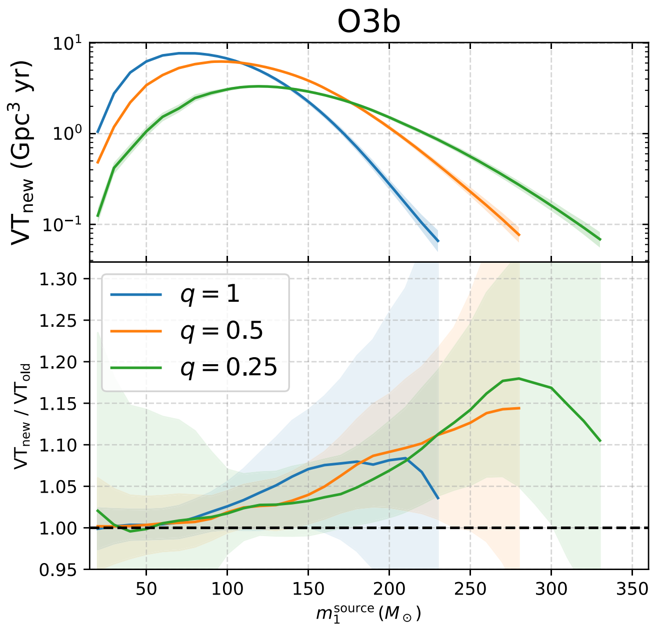


FIG. 7. Improvement in the sensitive spacetime volume ($VT$) of the IAS-HM search obtained by incorporating WaveletNet. The curves labeled "old" correspond to the $VT$ reported in Ref. [24], which does not include WaveletNet. We use the injection catalog released by the LVK Collaboration on Zenodo [104]. The shaded bands denote the 90% confidence intervals obtained via bootstrap resampling of the injection set. The observed sensitivity gains are primarily concentrated at high total masses and asymmetric mass ratios, corresponding to short-duration signals that are most susceptible to contamination from transient noise. We also observe an increase in the significance of several near-threshold GW candidates identified by the IAS-HM search.

modest IFAR in the original ranking, the inclusion of $p^{\rm ext}$ leads to a noticeable suppression of events occurring in noisy nonlocal environments, while leaving cleaner candidates largely unaffected or modestly enhanced. This behavior is consistent with the intended role of $p^{\rm ext}$ as a noise-aware correction to the pipeline ranking statistic.

Using these reranked candidate lists, we then compute updated false-alarm rates and astrophysical probabilities for simulated signal injections. The sensitive spacetime volume, $VT$, quantifies the four-dimensional volume over which a search is sensitive to astrophysical signals and is the primary figure of merit for comparing detection efficiency at fixed false-alarm rate; improvements in $VT$ correspond directly to an increased expected detection rate. In its simplest form, $VT$ is often estimated as

$$VT \approx \frac{N_{\rm det}}{N_{\rm inj}} V_{\rm inj} T_{\rm inj}, \tag{9}$$

where $N_{\rm det}$ is the number of detected injections above a given threshold, $N_{\rm inj}$ is the total number of injections, and $V_{\rm inj}T_{\rm inj}$ is the spacetime volume over which injections are distributed.

Our computation of the volume-time, $\overline{VT}$, employs a more accurate weighted Monte Carlo integral over the injections recovered by the pipeline with IFAR > 1 yr, following [103]:

$$\overline{VT} \simeq \frac{T_{\rm obs}}{N_{\rm draw}} \sum_{j=1}^{N_{\rm found}} \frac{f(z_j)\,\frac{1}{1+z_j}\frac{dV_c(z_j)}{dz}\,p(\theta_j)}{\pi_{\rm draw}(\theta_j, z_j)}, \tag{10}$$

where $T_{\rm obs}$ is the observation time of the analyzed dataset, $p(\theta)$ is the astrophysical probability distribution over the intrinsic source parameters $\theta = (m_1^{\rm s}, m_2^{\rm s}, \vec{s}_1, \vec{s}_2)$, $V_c$ is the comoving volume, $z$ is the redshift, $N_{\rm draw}$ is the total number of injections generated from the sampling distribution, and $\pi_{\rm draw}(\theta, z)$ is the corresponding injection sampling probability elaborated in Ref. [24].

In Fig. 7 we show the resulting improvement in $VT$ as a function of inverse false-alarm rate (IFAR) when incorporating $p^{\rm ext}$ into the ranking statistic. The shaded bands denote the 90% confidence intervals obtained via bootstrap resampling of the injection set. The error bars are wider in regions with fewer recovered injections. A larger injection campaign at high mass and asymmetric mass ratio would reduce these uncertainties. Nevertheless, across a wide range of IFAR thresholds, we observe a systematic increase in $VT$, indicating enhanced detection efficiency at fixed false-alarm rate.

## VII. DISCUSSION

The primary goal of this work is to introduce and evaluate a wavelet-based approach for characterizing nonlocal noise environments surrounding GW candidates, and to assess its impact when integrated into existing ranking pipelines. In this section, we discuss how WaveletNet compares to TIER, introduced in Ref. [84], and highlight the regimes in which the wavelet-based framework provides complementary or potentially improved performance.

A key distinction between WaveletNet and TIER lies in how nonlocal noise is modeled. The TIER method relies on matched filtering with template banks (originally constructed to identify compact binary coalescence signals). While this approach is effective at capturing certain forms of excess power, these templates are not optimized for instrumental glitches, which often exhibit time-frequency morphologies that can differ substantially from those of astrophysical signals. In contrast, WaveletNet learns a set of data-driven wavelet templates directly from environmental data, enabling it to adapt to the characteristic time-frequency structure of glitches. This design choice allows the model to assess the noise environment surrounding a candidate in a manner more closely aligned with the underlying glitch phenomenology.

Specifically, TIER uses the same template banks employed by the IAS-HM pipeline to generate matched filtering triggers for glitch identification. This strategy is computationally efficient, but it represents noise transients solely through their projections onto astrophysical GW templates. As a result, it can be suboptimal when instrumental glitches exhibit morphologies that differ significantly from compact binary coalescence signals.

We demonstrate that WaveletNet can improve the sensitivity of search pipelines targeting short-duration, clustered transients, achieving improvements in sensitive spacetime volume comparable to existing approaches while offering a complementary, data-driven alternative to fixed signal-based template banks. Because the model learns glitch-sensitive wavelet features directly from the data, it can adapt to changes in detector behavior. We therefore expect this approach to be useful in future observing runs, such as O4 and beyond, where glitch populations and morphologies may differ from those seen in O3. As long as the relevant glitches admit a reasonably sparse wavelet representation, the method can be retrained and recalibrated without relying solely on fixed signal-based template banks.

In this paper, we do not attempt to classify or characterize the morphology of individual glitches. Rather, we use wavelet triggers solely to flag time windows containing anomalous power, agnostic to the specific glitch type. The learned wavelet parameters associated with glitchy data may also be useful as priors in parameter-estimation or data-cleaning methods such as BayesWave [90].

Unlike the approach of Ref. [84], which characterizes nonlocal regions through hand-crafted summary statistics such as trigger time separations, neighboring SNRs, and spectral-noise metrics fed into a random-forest classifier, WaveletNet performs matched filtering directly against learned wavelet templates on the raw strain time series. This provides a flexible measure of transient noise activity: the matched-filter response quantifies the similarity between the nonlocal data and the characteristic time-frequency structure of glitches, rather than relying only on precomputed scalar summaries that may not fully capture local structure [5,12,56,105,106]. This inductive bias improves sample efficiency relative to more generic CNNs, which typically require larger training datasets to learn comparable representations. The scalar output score produced by WaveletNet is also directly interpretable as a ranking correction and can be incorporated into existing detection pipelines.

We also explored several alternative ML approaches based on CNNs, including architectures augmented with attention mechanisms. These models offer greater flexibility and, in principle, can learn a broader class of time-frequency patterns than wavelet-based models. In practice, however, CNN-based approaches required significantly more training data to achieve stable performance and were more prone to overfitting in the limited-data regime relevant for environmental glitch characterization. For a fair comparison, the CNNs were trained using the same dataset sizes as WaveletNet, namely 15,000 injection examples and 23,000 glitch examples per bank. While CNNs may outperform wavelet-based methods given sufficiently large and diverse training sets, WaveletNet provides a more data-efficient solution for the problem considered here. A detailed discussion of these alternative models and their performance is provided in Appendix A.

In Appendix B, we compare WaveletNet with the random-forest regression model introduced in Ref. [84], as well as with variants of WaveletNet that differ in architectural details. As described in Sec. V, our model integrates two input streams: wavelet-based features derived from nonlocal strain data and a set of summary statistical features. The additional comparisons presented in Appendix B and Fig. 6 show that the wavelet-based features provide most of the observed performance improvement, while the summary statistics add complementary contextual information. This supports the conclusion that learned wavelet features are a key contributor to the effectiveness of the proposed approach.

## VIII. CONCLUSIONS

Using a simple wavelet-based model architecture, we demonstrate that it is possible to account for certain characteristics of the nonlocal noise environment, yielding meaningful improvements in search-pipeline sensitivity. Instead of relying on large, data-hungry CNNs, we introduce a lightweight wavelet-based model that incorporates prior knowledge about glitch structure and can be trained with fewer examples. This architecture is easier to train, more stable in limited-data regimes, and still captures features that improve candidate ranking.

Gravitational-wave search pipelines typically assess the origin of a candidate using data localized to the immediate vicinity of the event. For stationary Gaussian noise, this is sufficient because data whitening already accounts for the broader noise environment. However, instrumental glitches are neither stationary nor Gaussian: some cluster in time and arise from common underlying disturbances, so elevated noise activity in the nonlocal environment can be correlated with glitches near the candidate. In this non-Gaussian regime, the broader data environment provides useful context for distinguishing astrophysical signals from instrumental noise transients. In this work, we develop machine-learning models that analyze the nonlocal environment surrounding a candidate by matched filtering learnable wavelets against the strain data, complemented by sparse summary statistical features.

The scalar output of these models can be incorporated into existing ranking statistics, enabling candidate reranking without disrupting established detection workflows. Applying this approach to the IAS-HM search, we observe peak improvements of $\sim 15\%$ in sensitive spacetime volume, concentrated at high total masses and asymmetric mass ratios ($q = 0.25$), where short-duration signals are most susceptible to contamination by transient noise (see Fig. 7). This value is the maximum gain across the mass and mass-ratio grid explored, not an average over all sources, demonstrating enhanced detection efficiency at fixed false-alarm rate in the regime most affected by non-Gaussian noise.

The proposed framework can be integrated with gravitational-wave search pipelines that produce candidate triggers and associated environmental data. Because the wavelet templates are learned directly from the data, the method can adapt to changes in detector noise characteristics. Future work will extend this analysis to O4 and subsequent observing runs, where evolving glitch populations and increased detector sensitivity may further benefit from data-driven, adaptive noise characterization.

## ACKNOWLEDGMENTS

A. P., D. W., M. H. Y. C., and E. B. are supported by NSF Grants No. AST-2307146, No. PHY-2513337, No. PHY-090003, and No. PHY-20043, by NASA Grant No. 21-ATP21-0010, by the John Templeton Foundation Grant No. 62840, by the Simons Foundation (MPS-SIP-00001698, E. B.), by the Simons Foundation International (SFI-MPS-BH-00012593-02), and by the Italian Ministry of Foreign Affairs and International Cooperation Grant No. PGR01167. M. H. Y. C. is a Croucher Fellow supported by the Croucher Foundation. M. H. Y. C. is supported by the Jonathan M. Nelson Center for Collaborative Research at the Institute for Advanced Study. This work was carried out at the Advanced Research Computing at Hopkins (ARCH) core facility [107] which is supported by the NSF Grant No. OAC-1920103.

## DATA AVAILABILITY

The data that support the findings of this article are openly available.

## APPENDIX A: OTHER APPROACHES

In this section, we present alternate ML model architectures that we tried, but that did not yield usable results. While the reasons for the failures are often complicated, we present some plausible explanations.

### 1. Convolutional neural network (CNN)

We experimented with directly training CNNs on the GW strain examples described in Sec. V. We also tested an attention layer designed to focus the model on glitchlike structure. We describe these attempts below.

#### a. Training on spectrograms of the nonlocal region

We generated spectrograms from the time-series data using the `specgram` function from the MATPLOTLIB library. The spectrograms were computed with a sampling frequency `Fs = 256`, an FFT window size `NFFT = 64`, and an overlap of `noverlap = 32` between successive windows. The spectrogram was divided into the left and right region of the candidate and passed separately to the CNN (see Fig. 5).

One model used multiple stacked layers of 2D convolutions, batch normalization, ReLU activations, and max pooling. The goal was to learn spectral patterns associated with glitches directly from the input spectrograms. We found that this CNN struggled to distinguish the two classes effectively.

We also tested a model with a SpectralAttention mechanism after several convolutional blocks, inspired by Ref. [108]. Each convolution block consisted of a 2D convolutional layer with progressively larger kernel sizes, $(5 \times 5)$, $(11 \times 11)$, and $(21 \times 21)$, followed by batch normalization, LeakyReLU activation, and max pooling. After reshaping the output, we applied a custom attention layer that projected the features into query and value representations and used soft attention weights to emphasize salient regions. The motivation was to help the model focus on glitch-specific patterns, especially in noisy or ambiguous cases. However, the attention-enhanced model did not yield a significant improvement. A possible explanation is that the spectrograms from the glitch and injection classes often appeared quite similar. This suggests that the CNN may have failed to isolate features that were truly distinctive of glitchy backgrounds, potentially because of subtle underlying spectral differences.

WaveletNet performs better in this limited-data regime because its wavelet filters act as convolution kernels optimized to detect non-Gaussian transient structure. By imposing an inductive bias tailored to glitch morphology, the model focuses on a restricted and physically motivated feature space rather than learning arbitrary spectral patterns. As a result, WaveletNet achieves improved data efficiency and can be trained with fewer samples than more flexible CNN-based architectures.

#### b. Training on raw time series

Another approach was to train one-dimensional CNNs on raw time series without preprocessing. The goal was to retain the raw amplitude and phase information without imposing a time-frequency trade-off or depending on spectrogram preprocessing choices. Additionally, we experimented with two types of attention mechanisms within our model structure—TransformerSelfAttention and SpectralAttention.

CNNs can capture short-range temporal features with shallow layers, but identifying long-range structure generally requires deeper architectures. TransformerSelfAttention creates direct connections between all elements of the time series, allowing the model to act as a global pattern detector. In principle, this could help connect noisy nonlocal regions with nearby glitches.

## APPENDIX B: MORE FIGURES AND COMPARISONS

In our previous work [84], we trained a random-forest (RF) model only on summary statistics and achieved an almost 20% improvement in $VT$. WaveletNet provides a different way to evaluate the nonlocal environment of candidates. In Fig. 8, we compare the RF and WaveletNet outputs for glitchy and nonglitchy strain data. We also test a variant of WaveletNet in which the classification-head MLP is replaced by an RF model; the corresponding $VT$ increase is shown in Fig. 9. In Figs. 10 and 11, we show that WaveletNet achieves more stable validation performance than CNNs, particularly in the limited-data regime.

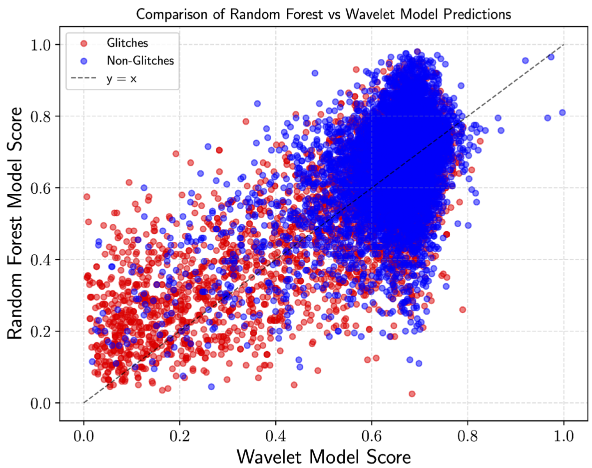


FIG. 8. Comparison of output scores from the RF model trained in Ref. [84] and the WaveletNet model. The scatter away from the $y = x$ line shows that WaveletNet learns features that differ from those used by the RF model.

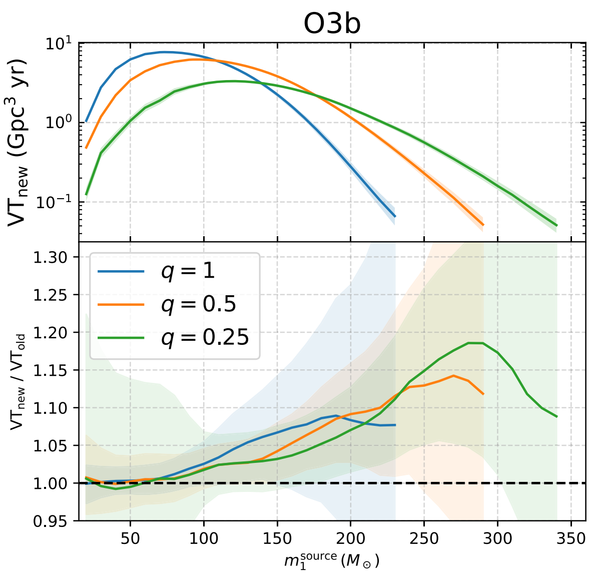


FIG. 9. Similar to Fig. 7, but using an RF regression model instead of the multilayer perceptron as the final classification module (see Fig. 1). This comparison tests the choice of final classifier. We find that the resulting increase in $VT$ is comparable to that obtained with the MLP-based WaveletNet.

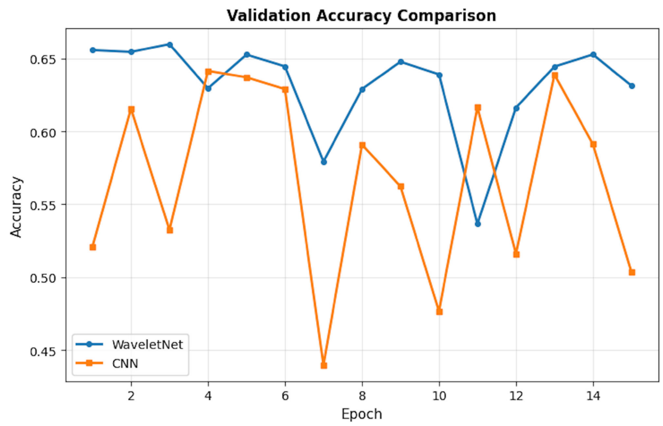


FIG. 10. Validation accuracy over training epochs for WaveletNet and the CNN. WaveletNet achieves higher and more stable performance than the CNN. Training CNNs requires large datasets, especially when the input segments are long.

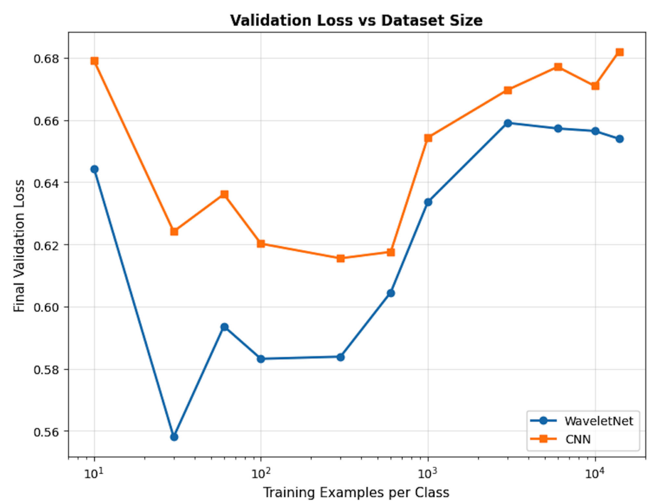


FIG. 11. Final validation loss as a function of training set size for WaveletNet and CNN. WaveletNet maintains lower and more stable loss across dataset sizes, while CNN performance degrades at larger scales, reflecting its higher data requirements.

[1] A. Einstein, Ann. Phys. (Berlin) **49**, 769 (1916).
[2] B. P. Abbott *et al.*, Phys. Rev. Lett. **116**, 061102 (2016).
[3] J. Aasi *et al.* (LIGO Scientific Collaboration), Classical Quantum Gravity **32**, 074001 (2015).
[4] T. Akutsu *et al.* (KAGRA Collaboration), Prog. Theor. Exp. Phys. **2021**, 05A101 (2021).
[5] M. Cabero *et al.*, Classical Quantum Gravity **36**, 15 (2019).
[6] D. Davis *et al.* (LIGO Collaboration), Classical Quantum Gravity **38**, 135014 (2021).
[7] F. Acernese *et al.* (Virgo Collaboration), Classical Quantum Gravity **40**, 185006 (2023).
[8] S. Soni *et al.* (LIGO Collaboration), Classical Quantum Gravity **42**, 085016 (2025).
[9] F. Badaracco, C. De Rossi, I. Fiori, J. Harms, K. Miyo, F. Paoletti, T. Tanaka, T. Washimi, and T. Yokozawa, Phys. Rev. D **104**, 042006 (2021).
[10] J. Buchli *et al.* (The LIGO Instrument Team), Science **389**, 1012 (2025).
[11] M. Zevin *et al.*, Classical Quantum Gravity **34**, 064003 (2017).
[12] S. Soni *et al.* (LIGO Collaboration), Classical Quantum Gravity **38**, 025016 (2020).
[13] M. H.-Y. Cheung, D. Wadekar, A. K. Mehta, T. Islam, J. Roulet, E. Berti, T. Venumadhav, B. Zackay, and M. Zaldarriaga, Phys. Rev. D **113**, 023003 (2026).
[14] S. A. Usman *et al.*, Classical Quantum Gravity **33**, 215004 (2016).
[15] C. Messick *et al.*, Phys. Rev. D **95**, 042001 (2017).
[16] B. Ewing *et al.*, Phys. Rev. D **109**, 042008 (2024).
[17] P. Godwin *et al.*, arXiv:2010.15282.
[18] S. Klimenko *et al.*, Phys. Rev. D **93**, 042004 (2016).
[19] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. D **93**, 122004 (2016); **94**, 069903(A) (2016).
[20] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. D **95**, 042003 (2017).
[21] F. Aubin *et al.*, Classical Quantum Gravity **38**, 095004 (2021).
[22] D. Wadekar, T. Venumadhav, J. Roulet, A. K. Mehta, B. Zackay, J. Mushkin, and M. Zaldarriaga, Phys. Rev. D **110**, 044063 (2024).
[23] D. Wadekar, J. Roulet, T. Venumadhav, A. K. Mehta, B. Zackay, J. Mushkin, S. Olsen, and M. Zaldarriaga, arXiv:2312.06631.
[24] A. K. Mehta, D. Wadekar, J. Roulet, I. Anantpurkar, T. Venumadhav, J. Mushkin, B. Zackay, M. Zaldarriaga, and T. Islam, arXiv:2501.17939.
[25] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, Phys. Rev. D **100**, 023011 (2019).
[26] A. H. Nitz, Classical Quantum Gravity **35**, 035016 (2018).
[27] B. Allen, Phys. Rev. D **71**, 062001 (2005).
[28] P. Joshi, R. Dhurkunde, S. Dhurandhar, and S. Bose, Phys. Rev. D **103**, 044035 (2021).
[29] S. Choudhary, S. Bose, S. Dhurandhar, and P. Joshi, Phys. Rev. D **110**, 044051 (2024).
[30] T. Ghosh, S. Bose, S. Dhurandhar, and S. Choudhary, Classical Quantum Gravity **43**, 045016 (2026).
[31] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **119**, 161101 (2017).
[32] R. Abbott *et al.* (LIGO Scientific, VIRGO, and KAGRA Collaborations), Astron. Astrophys. **659**, A84 (2022).
[33] K. Chandra, J. Calderón Bustillo, A. Pai, and I. Harry, Phys. Rev. D **106**, 123003 (2022).
[34] A. E. Koloniari, E. C. Koursoumpa, P. Nousi, P. Lampropoulos, N. Passalis, A. Tefas, and N. Stergioulas, Mach. Learn. Sci. Tech. **6**, 015054 (2025).
[35] T. Zhao, R. Shi, Y. Zhou, Z. Cao, and Z. Ren, Front. Phys. (Beijing) **20**, 45301 (2025).
[36] T. S. Yamamoto, A. L. Miller, M. Sieniawska, and T. Tanaka, Phys. Rev. D **106**, 024025 (2022).
[37] M. Zevin *et al.*, Eur. Phys. J. Plus **139**, 100 (2024).
[38] T. D. Gebhard, N. Kilbertus, I. Harry, and B. Schölkopf, Phys. Rev. D **100**, 063015 (2019).
[39] X. Li, W. Yu, X. Fan, and G. Jogesh Babu, Front. Phys. **15**, 54501 (2020).
[40] R. Biswas *et al.*, Phys. Rev. D **88**, 062003 (2013).
[41] E. Cuoco, M. Cavaglià, I. S. Heng, D. Keitel, and C. Messenger, Living Rev. Relativity **28**, 2 (2025).
[42] J. Fernandes, A. Pai, and K. Chandra, Phys. Rev. D **112**, 083043 (2025).
[43] K. Sharma, K. Chandra, and A. Pai, arXiv:2208.02545.
[44] M. Saleem *et al.*, Classical Quantum Gravity **41**, 195024 (2024).
[45] C. Reissel, S. Soni, M. Saleem, M. Coughlin, P. Harris, and E. Katsavounidis, arXiv:2501.04883.
[46] S. Alvarez-Lopez, A. Liyanage, J. Ding, R. Ng, and J. McIver, Classical Quantum Gravity **41**, 085007 (2024).
[47] M. Lopez, G. C. Santoro, A. Martins, S. Schmidt, J. Schoppink, W. van Straalen, C. Capano, and S. Caudill, Phys. Rev. D **111**, 103020 (2025).
[48] A.-K. Malz and J. Veitch, Phys. Rev. D **112**, 024071 (2025).
[49] R. Legin, M. Isi, K. W. K. Wong, Y. Hezaveh, and L. Perreault-Levasseur, Astrophys. J. Lett. **985**, L46 (2025).
[50] N. Shah, A. M. Knee, J. McIver, and D. C. Stenning, Classical Quantum Gravity **40**, 235008 (2023).
[51] E. Marx *et al.*, Phys. Rev. D **111**, 042010 (2025).
[52] M. B. Schäfer *et al.*, Phys. Rev. D **107**, 023021 (2023).
[53] A. McLeod, D. Beveridge, L. Wen, and A. Wicenec, Phys. Rev. D **111**, 024035 (2025).
[54] S. J. Kapadia, T. Dent, and T. Dal Canton, Phys. Rev. D **96**, 104015 (2017).
[55] S. Bini, G. Vedovato, M. Drago, F. Salemi, and G. A. Prodi, Classical Quantum Gravity **40**, 135008 (2023).
[56] R. Essick, P. Godwin, C. Hanna, L. Blackburn, and E. Katsavounidis, Mach. Learn. **2**, 015004 (2021).
[57] N. Nagarajan and C. Messenger, Phys. Rev. D **112**, 103002 (2025).
[58] C. Chatterjee, A. Petulante, K. Jani, J. Spencer-Smith, Y. Hu, R. Lau, H. Fu, T. Hoang, S. C. Zhao, and S. Deshmukh, arXiv:2412.20789.
[59] C. Chatterjee and K. Jani, Astrophys. J. **973**, 112 (2024).
[60] C. Chatterjee and K. Jani, Astrophys. J. **982**, 102 (2025).
[61] C. McIsaac and I. Harry, Phys. Rev. D **105**, 104056 (2022).
[62] J. Powell, D. Trifirò, E. Cuoco, I. S. Heng, and M. Cavaglià, Classical Quantum Gravity **32**, 215012 (2015).
[63] J. Powell, A. Torres-Forné, R. Lynch, D. Trifirò, E. Cuoco, M. Cavaglià, I. S. Heng, and J. A. Font, Classical Quantum Gravity **34**, 034002 (2017).

[64] M. Razzano and E. Cuoco, Classical Quantum Gravity **35**, 095016 (2018).
[65] S. Bahaadini, V. Noroozi, N. Rohani, S. Coughlin, M. Zevin, J. R. Smith, V. Kalogera, and A. Katsaggelos, Inf. Sci. (NY) **444**, 172 (2018).
[66] S. Soni *et al.*, Classical Quantum Gravity **38**, 195016 (2021).
[67] J. Glanzer *et al.*, Classical Quantum Gravity **40**, 065004 (2023).
[68] D. George, H. Shen, and E. A. Huerta, Phys. Rev. D **97**, 101501(R) (2018).
[69] D. George, H. Shen, and E. A. Huerta, Phys. Rev. D **97**, 101501 (2018).
[70] M. Llorens-Monteagudo, A. Torres-Forné, J. A. Font, and A. Marquina, Classical Quantum Gravity **36**, 075005 (2019).
[71] O. T. Bişkin, İ. Kırbaş, and A. Çelik, Astron. Comput. **42**, 100683 (2023).
[72] N. Koyama, Y. Sakai, S. Sasaoka, D. Dominguez, K. Somiya, Y. Omae, Y. Terada, M. Meyer-Conde, and H. Takahashi, Mach. Learn. Sci. Tech. **5**, 035028 (2024).
[73] A.-K. Malz, G. Ashton, and N. Colombo, Phys. Rev. D **111**, 084078 (2025).
[74] Y. Wu, M. Zevin, C. P. L. Berry, K. Crowston, C. Østerlund, Z. Doctor, S. Banagiri, C. B. Jackson, V. Kalogera, and A. K. Katsaggelos, Classical Quantum Gravity **42**, 165015 (2025).
[75] D. George and E. A. Huerta, Phys. Lett. B **778**, 64 (2018).
[76] S. Choudhary, A. More, S. Suyamprakasam, and S. Bose, Phys. Rev. D **107**, 024030 (2023).
[77] A. Trovato, É. Chassande-Mottin, M. Bejger, R. Flamary, and N. Courty, Classical Quantum Gravity **41**, 125003 (2024).
[78] Q. Meijer, M. Lopez, D. Tsuna, and S. Caudill, Phys. Rev. D **109**, 022006 (2024).
[79] Z. Yarbrough, A. Guimaraes, P. Joshi, G. González, A. Valentini, and U. Shah, Classical Quantum Gravity **42**, 165014 (2025).
[80] A. Gunny, D. Rankin, J. Krupa, M. Saleem, T. Nguyen, M. Coughlin, P. Harris, E. Katsavounidis, S. Timm, and B. Holzman, Nat. Astron. **6**, 529 (2022).
[81] K. Chatziioannou, N. Cornish, M. Wijngaarden, and T. B. Littenberg, Phys. Rev. D **103**, 044013 (2021).
[82] L. Vazsonyi and D. Davis, Classical Quantum Gravity **40**, 035008 (2023).
[83] S. Cheng, Y.-S. Ting, B. Ménard, and J. Bruna, Mon. Not. R. Astron. Soc. **499**, 5902 (2020).
[84] D. Wadekar *et al.*, arXiv:2507.08318.
[85] D. Nandi, A. Effler, S. Soni, T. A. Ferreira, R. M. Schofield, H. Pham, T. Ohanlon, V. V. Frolov, and G. González, arXiv:2605.14143.
[86] A. G. Abac *et al.* (LIGO Scientific, VIRGO, and KAGRA Collaborations), arXiv:2508.18082.
[87] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, Phys. Rev. D **85**, 122006 (2012).
[88] S. Hourihane, K. Chatziioannou, M. Wijngaarden, D. Davis, T. Littenberg, and N. Cornish, Phys. Rev. D **106**, 042006 (2022).
[89] N. J. Cornish, arXiv:2511.10632.
[90] N. J. Cornish and T. B. Littenberg, Classical Quantum Gravity **32**, 135012 (2015).
[91] S. K. Chatterji, The search for gravitational wave bursts in data from the second LIGO science run, Ph.D. thesis, Massachusetts Institute of Technology, 2005.
[92] C. Cutler and E. E. Flanagan, Phys. Rev. D **49**, 2658 (1994).
[93] D. Wadekar, T. Venumadhav, A. K. Mehta, J. Roulet, S. Olsen, J. Mushkin, B. Zackay, and M. Zaldarriaga, arXiv:2310.15233.
[94] https://github.com/Arush-Pimpalkar/WaveletNet.
[95] E. Oyallon, S. Zagoruyko, G. Huang, N. Komodakis, S. Lacoste-Julien, M. Blaschko, and E. Belilovsky, IEEE Trans. Pattern Anal. Mach. Intell. **41**, 2208 (2019).
[96] G. Michau, G. Frusque, and O. Fink, Proc. Natl. Acad. Sci. U.S.A. **119**, e2106598119 (2022).
[97] V. Khemani, M. Azarian, and M. Pecht, Electronics **11**, 451 (2022).
[98] J. Bruna and S. Mallat, arXiv:1203.1513.
[99] B. Hall, S. Suyamprakasam, N. Mazumder, A. More, and S. Bose, Classical Quantum Gravity **41**, 245016 (2024).
[100] D. Wadekar, T. Venumadhav, J. Roulet, A. K. Mehta, B. Zackay, J. Mushkin, and M. Zaldarriaga, Phys. Rev. D **110**, 044063 (2024).
[101] J. Roulet, L. Dai, T. Venumadhav, B. Zackay, and M. Zaldarriaga, Phys. Rev. D **99**, 123022 (2019).
[102] D. Wadekar, T. Venumadhav, A. K. Mehta, J. Roulet, S. Olsen, J. Mushkin, B. Zackay, and M. Zaldarriaga, Phys. Rev. D **110**, 084035 (2024).
[103] V. Tiwari, Classical Quantum Gravity **35**, 145009 (2018).
[104] LIGO-Virgo-KAGRA Collaboration, 10.5281/zenodo.5546675 (2021).
[105] D. Davis, L. V. White, and P. R. Saulson, Classical Quantum Gravity **37**, 145001 (2020).
[106] R. Huxford, R. George, M. Trevor, Z. Yarbrough, and P. Godwin, arXiv:2412.04638.
[107] https://www.arch.jhu.edu/.
[108] P. Melchior, Y. Liang, C. Hahn, and A. Goulding, Astron. J. **166**, 74 (2023).