

DENOISING GRAVITATIONAL WAVES WITH ENHANCED DEEP RECURRENT DENOISING AUTO-ENCODERS

Hongyu Shen¹ Daniel George² Eliu. A. Huerta^{2,3} Zhizhen Zhao^{1,3}

University of Illinois at Urbana-Champaign

¹Dept. of Electrical and Computer Engineering, ²Dept. of Astronomy, ³NCSA

ABSTRACT

Denoising of time domain data is a crucial task for many applications such as communication, translation, virtual assistants etc. For this task, a combination of a recurrent neural net (RNNs) with a Denoising Auto-Encoder (DAEs) has shown promising results. However, this combined model is challenged when operating with low signal-to-noise ratio (SNR) data embedded in non-Gaussian and non-stationary noise. To address this issue, we design a novel model, referred to as ‘Enhanced Deep Recurrent Denoising Auto-Encoder’ (EDR-DAE), that incorporates a signal amplifier layer, and applies curriculum learning by first denoising high SNR signals, before gradually decreasing the SNR until the signals become noise dominated. We showcase the performance of EDR-DAE using time-series data that describes gravitational waves embedded in very noisy backgrounds. In addition, we show that EDRDAE can accurately denoise signals whose topology is significantly more complex than those used for training, demonstrating that our model generalizes to new classes of gravitational waves that are beyond the scope of established denoising algorithms.

Index Terms— RNN, denoising algorithm, time series, gravitational waves

1. INTRODUCTION

Denoising time series data in real-time is a timely and fast-paced field of research. The advent of hands-free devices for communication, translation, voice assistants, and hearing aids have triggered the development of new approaches and methods to recognize speech data in realistic, noisy environmental conditions. However, denoising extremely weak time-series signals embedded in non-Gaussian and non-stationary experimental data remains an outstanding challenge.

The exploitation of deep learning algorithms has led to spectacular progress in signal denoising research. Denoising Auto-Encoders (DAEs) stand out among de-noising methods [1, 2]. Denoising auto-encoders (DAEs) follow the architecture of AEs, a type of unsupervised deep net learns an efficient data representation [3, 4, 5], mapping an input x to a hidden representation x' that stores the necessary in-

formation to again reconstruct y of the original input. DAEs are computationally efficient since they utilize convolutional neural networks. Convolutional structures are well-known to be effective in spectrograms. However, some time series require good algorithms in the time domain, for example, gravitational waves (GWs), due to the suboptimality of lossy non-invertible representations of the underlying data introduced by the conversion to spectrograms [6], and to demands on the preserved information in time domain. Even in the time domain, the information of the front part of the data encoded into the model are almost uncorrelated with the terminal part of the data, due to the kernel structure in convolutions. An optimal approach, Deep Recurrent Denoising Auto-Encoder (DRDAE), proposed in [2], to accomplish this consists of feeding the raw time-series data to a model that takes inputs in multiple timesteps. Nevertheless, this approach is still ineffective in denoising low SNR signals, as discussed in [7].

To address this issue, we develop (1) a new RNN DAE architecture, named Enhanced Deep Recurrent De-noising Auto-Encoder (EDRDAE) based on DRDAE, and (2) a new training strategy. We use curriculum learning [8] to ensure that data with SNRs that range from low to very high are accurately recovered. We also describe the rationale to use signal amplifiers (SA) and cross-layer state connections (blue arrows in Fig 1(b)), two critical features in EDRDAE. Specifically, we demonstrate the superior performance of EDRDAE with two gravitational wave datasets, quasi-circular GWs describes two black hole mergers and eccentric GWs. We show that compared with DRDAE, EDRDAE achieves lower mean square error (MSE) and higher overlap [9, 10]. In addition, we show our model can generalize to different types of time series (eccentric GWs) in denoising when only one type of data (quasi-circular GWs) is used for training. Results relative to PCA and dictionary learning is also included in the experiment section.

In this paper, we use SNR_{peak} to quantify the noise level, defined as $\text{SNR}_{\text{peak}} = \frac{\text{Maximum value of a given signal}}{\text{Standard deviation of noise}}$. To facilitate direct comparisons with typical conventions in the GW community, a conversion between SNR_{peak} and matched-filtering SNR, SNR_{MF} , is also provided, which is approximately $\text{SNR}_{\text{MF}} = 13 \times \text{SNR}_{\text{peak}}$.

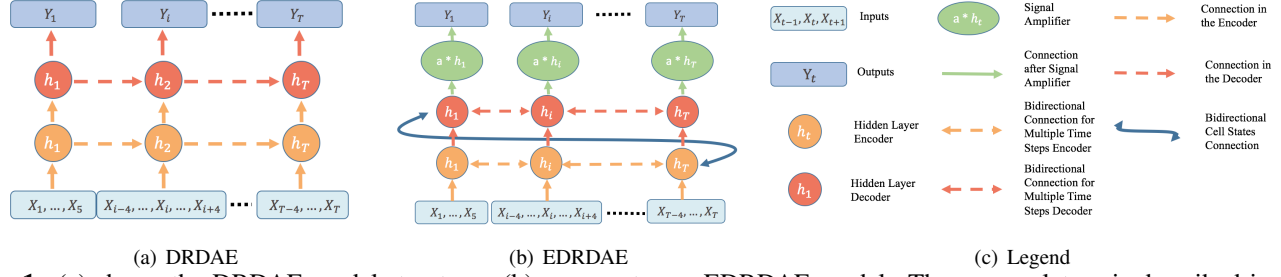


Fig. 1. (a) shows the DRDAE model structure. (b) represents our EDRDAE model. The nomenclature is described in (c). Orange circles and red circles indicate encoder and decoder part of the two models. Double arrows indicate the bidirectional LSTM connections, whereas one-directional arrows refer to conventional LSTM structure. Both DRDAE and EDRDAE has multiple timestep inputs. However, for EDRDAE, we add a cross-layer connection (blue arrows (BA)) and signal amplifier (SA) in the green ellipse. Notice that the diagrams in (a) and (b) only display the basic structure of the models, the number of layers in the actual experiments differ and are provided in the experiment section.

2. METHODS

2.1. Model Architecture

The architecture of EDRDAE (Fig 1(b)) empowers this new model to better denoise time series embedded in non-Gaussian and non-stationary noise datasets. There are three major structural differences compared to DRDAE (Fig 1(a)). First of all, besides the multiple time step inputs, we apply bidirectional LSTMs [11], rather than conventional one-way LSTMs, since the bidirectional structure will pass the information of the input data through time in two directions, forward and backward. Especially for multiple timestep inputs, where we use several neighboring time points to predict the central timestep, we found bidirectional layers to help pass information from neighboring timesteps (before and after the central timestep) to adjacent layers. Intuitively, this boosts denoising performance significantly.

The second difference is the introduction of the Signal Amplifier (SA) (see Fig 1(b)). It is beneficial in denoising signals when the amplitude of the signal is lower than that of the background noise. This new structure is inspired by speech data, which is nearly symmetric concerning the mean-value axis (a horizontal line). SA right before the output layer assists the network in learning more evident patterns and magnifying the reconstructed values to reach the true values of the clean references.

Finally, to make the training faster, we apply a cross-layer connection, which is in spirit similar to the connection in [12]. It passes through both forward and backward cells in bidirectional layers. The cell states from the output timestep of the encoder are passed to the cell states of the first timestep of the decoder layer, as illustrated in Fig 1(b) using blue arrows. Empirically, we observe this structure helps reconstruction in noisy environments and achieves higher accuracy compared to models that have different cell states for different layers.

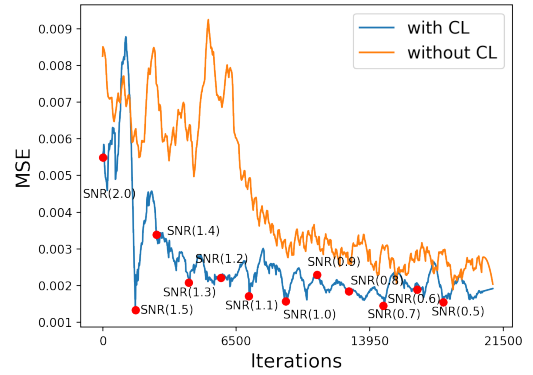


Fig. 2. Validation error comparison between training models with and without curriculum learning. The model with curriculum learning achieves lower MSE compared to the model without curriculum learning. The red markers indicate the change of SNRs with curriculum learning.

2.2. Curriculum learning with Decreasing SNR

Different from the conventional image and audio denoising setups [13, 14], for which the noise level is typically too low to obscure image background or audio utterances, the data we focus on, however, are always embedded within extreme noise. For example, raw gravitational waves. As a result, it is difficult to learn the original signal structure and remove the noise from raw data when training directly starts with very low SNRs. We found that gradually reducing SNRs during training, an idea taken from curriculum learning literature [8], provides regularization, which allows the network to distill more accurate information of the underlying signals with larger SNRs to signals with lower SNRs. This guarantees that the information of the true signal is maintained even with extreme noise. Moreover, a decreasing SNR strategy, in combination with the previously mentioned signal amplifier (SA), ensures the reconstruction accuracy and smoothness of the data. With this approach, the model accurately denoises GWs under extremely small SNRs in testing, and it

Table 1. CL training scheme for GW denoising. This table illustrates the change of SNR_{peak} happens along training for DRDAE and EDRDAE models. We start with SNR_{peak} 2.0 for the first 2000 iterations and then decrease SNR_{peak} with specified iterations in the table to the final $\text{SNR}_{\text{peak}} = 0.5$. The number of iterations shown in the table where a decrease of SNR is performed is based on the loss function.

ITER. (K)	2	3	4.5	6	7.5	9	10.5	13	14.5	16	17.5
SNR_{peak}	1.5	1.4	1.3	1.2	1.1	1.0	0.9	0.8	0.7	0.6	0.5

reaches lower MSE with fewer iterations compared with models trained without curriculum learning (Fig 2).

3. EXPERIMENT

3.1. Gravitational Wave Datasets

We designed experiments to illustrate the performance of DRDAE and EDRDAE on GW datasets. We use simulated GWs describe binary black hole (BBH) mergers, generated with the waveform model introduced in [15], which is available in LIGO’s Algorithm Library [16]. We consider BBH systems with mass-ratios $q \leq 10$ in steps of 0.1, and with total mass $M \in [5M_{\odot}, 75M_{\odot}]$, in steps of $1M_{\odot}$. The waveforms are generated with a sampling rate of 8192 Hz and whitened with the design sensitivity of LIGO [17]. We consider the late inspiral, merger and ringdown evolution of BBHs, since it is representative of the BBH GW signals reported by ground-based GW detectors [18, 19]. Specifically, we have 9861 (75%) samples for training and 2459 (25%) for testing. There are additional 41 eccentric waveforms with eccentricities up to 0.2 ten cycles before the black holes collide for testing the generalizability of our model. For simulating non-Gaussian and non-stationary noise cases, we take real non-Gaussian noise, 4096s from the LIGO Open Science Center (LOSC) around the LVT151012 event.

3.2. Data Preprocessing

The SNR of astrophysical GW sources in the LIGO detectors cannot be known prior to detection. Therefore, standardization of the datasets is a crucial step to guarantee that our model can detect GWs with a wide range of astrophysically motivated SNRs. As a result, we add additive white Gaussian noise to the clean data for each iteration during training with a predetermined SNR_{peak} and normalize our data to have variance 1.0. Apart from the timestep augmentation [2], we also add random shifts to the training data, to make the model more resilient to variations in the location of the signal. For every input signal, we randomly generate an integer between 0 and 200 as shift length for left and right shifts. The

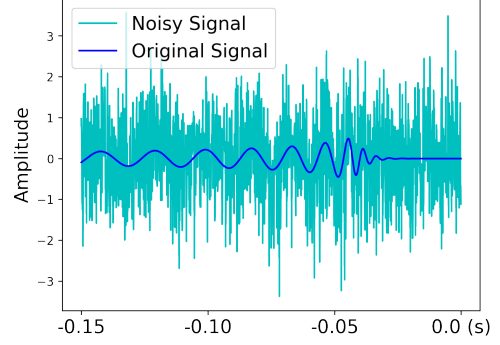


Fig. 3. Comparison of the noisy gravitational wave at $\text{SNR}_{\text{peak}} = 0.5$ and the clean signal.

Table 2. Comparisons of MSE and overlap across different approaches. \cdot/\cdot refers to ”metric for quasi-circular” / ”metric for eccentric”. Here ”DL” refers to dictionary learning. ”WT” refers to wavelet thresholding.

MODEL	PCA	DL	WT	DRDAE	EDRDAE
MSE	.033/.034	.108/.107	.017/.018	.008/.004	.001/.005
OVERLAP	.644/.639	.466/.476	.684/.671	.906/.955	.994/.985

length is 0% to 15% proportional to the total signal length. Zero padding is performed when necessary.

3.3. Training and Evaluation

All trainings are performed on NVIDIA Tesla V100 GPUs using TensorFlow [20]. We train the models with batch size 30. To ensure a fair model comparison, each encoder and decoder has four layers for DRDAE, with a one-way LSTM cell that has 64 channels for each layer. For EDRDAE, we have one encoder layer followed by two decoder layers with bidirectional LSTM cells. With the defined structures, DRDAE has 250,177 parameters, and EDRDAE has 235,650 parameters. We tuned the optimizers (ADAM, RMSprop, and SGD), learning rate (0.0001 to 0.001), batch size (10 to 50), number of layers (2 to 4) and the size of input multiple time steps (5 to 20). Only results of the best model (Adam optimizer, 0.0001 learning rate, batch size 30, 4-layer model 9 timesteps for input) displayed in Fig 4. We use the default ADAM setup described in [21]. To denoise signals with extremely low SNR, for example, $\text{SNR}_{\text{peak}} = 0.5$ (Fig 3), our training starts with $\text{SNR}_{\text{peak}} = 2.0$. We gradually decrease the SNR at every subsequent training step until we reach a $\text{SNR}_{\text{peak}} = 0.5$ (Shown in Table 1 and Fig 2). We only input BBH GWs and additive Gaussian noise for training. Real LIGO noise and both BBH GWs and eccentric GWs are used for testing. In general, the models become stable after training for about 20,000 iterations. We summarize results after the fi-

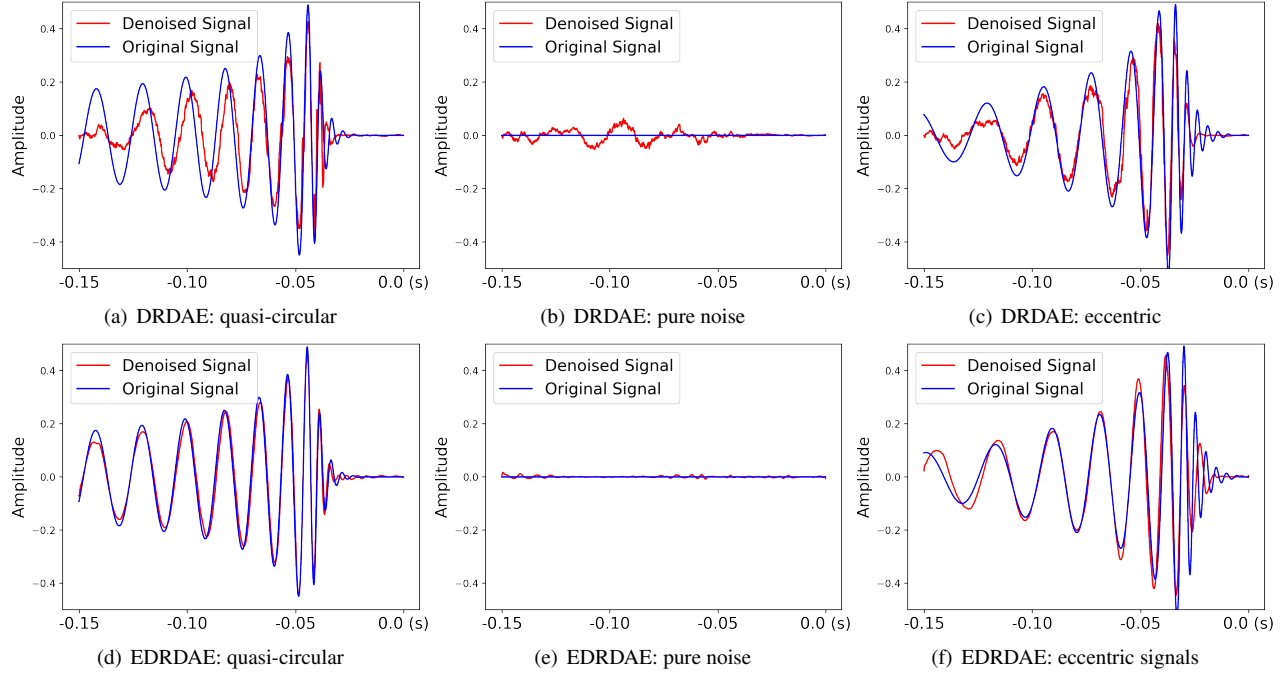


Fig. 4. Performance on GW signals contaminated by real LIGO detector noise with $\text{SNR}_{\text{peak}} = 0.5$. The plots include reconstructed outputs of DRDAE and EDRDAE on quasi-circular signals, pure noise input, and eccentric gravitational waves.

Table 3. Ablation study for major parts in EDRDAE

MISSING PARTS	W/O SA	W/O BA	W/O CL
MSE	.017/.002	.007/.003	.010/.001
OVERLAP	.769/.979	.941/.974	.941/.960

nal iteration of training in Fig 4. Two metrics are used for model evaluation for GWs. Mean Squared Error (MSE) is a conventional metric for GW evaluation measuring the reconstruction quality. We also evaluate a metric known as overlap, which measures how similar any two GWs are. Overlap values equal to unity indicate perfect agreement between any two given signals. For GW detection, LIGO typically requires a minimum overlap value 0.97 [9, 10]. Measurements are concluded in Table 2 with benchmarks that include PCA with Wiener filtering and dictionary learning optimized by coordinate descent, as well as wavelet universal thresholding with 10 levels [22]. An ablation study for the SA, cross-layer connections and decreasing SNRs is also provided in Table 3.

Results on DRDAE and EDRDAE with real LIGO noise are in Fig 4. We compare the denoising results for a quasi-circular GW in Fig 4(a) and Fig 4(d). To prevent spurious detection, which is crucial in GW detection and denoising, we also test the model when the input is only pure noise. Notice that our model is also able to prevent fake wave generation from pure noise (see Fig. 4(e)), whereas the output from DRDAE still contains wave-like structure with noticeable amplitude (see Fig. 4(b)). Additionally, we also test the

resilience of the model to denoise signals that are not used during training, i.e., eccentric GWs. Currently, there is no signal-processing algorithm adequate for the detection or denoising of eccentric GWs due to their complex topology [23]. We chose eccentric GWs to showcase the robustness of our algorithm to denoise signals that are significantly different from the training data. The sample waveforms are presented in Fig 4(c) and Fig 4(f). It is clear our model achieves better reconstruction in both quasi-circular GWs and eccentric GWs with lower MSEs and higher overlaps (Table 2).

4. CONCLUSION

In this paper, we proposed a new deep recurrent denoising auto-encoder to denoise gravitational wave signals contaminated by an extremely high level of noise often encountered in realistic detection scenarios. By introducing additional structures to the model (cross-layer connection, signal amplifier), and by adopting a training approach that gradually reduces the SNRs of the training samples, we show that our model outperforms DRDAE and other tested popular denoising algorithms (PCA, dictionary learning and wavelet thresholding) for GW denoising. It is also noteworthy that although our denoising auto-encoder was only trained with quasi-circular GWs contaminated with additive white Gaussian noise, it is able to handle both quasi-circular GWs with different mass ratios and eccentric GWs embedded in real LIGO noise. Therefore the proposed method has great generalization performance.

5. REFERENCES

- [1] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol, “Extracting and Composing Robust Features with Denoising Autoencoders,” New York, NY, USA, 2008, ICML 2008, pp. 1096–1103, ACM.
- [2] A. L. Maas, Q. V. Le, T. M. O’Neil, O. Vinyals, P. Nguyen, and A. Y. Ng, “Recurrent neural networks for noise reduction in robust ASR,” in *INTER-SPEECH, 2012*, 2012, pp. 22–25.
- [3] G. E. Hinton, S. Osindero, and Y. Teh, “A fast learning algorithm for deep belief nets,” *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [4] G. Hinton and R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504 – 507, 2006.
- [5] Y. Bengio and Y. LeCun, “Scaling learning algorithms towards AI,” 2007, vol. 34, pp. 1–41.
- [6] Daniel George and E. A. Huerta, “Deep Neural Networks to Enable Real-time Multimessenger Astrophysics,” *Phys. Rev.*, vol. D97, no. 4, pp. 044039, 2018.
- [7] N. Kumar, M. Van Segbroeck, K. Audhkhasi, P. Drotár, and S. S. Narayanan, “Fusion of diverse denoising systems for robust automatic speech recognition,” in *ICASSP*, 2014, pp. 5557–5561.
- [8] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, “Curriculum learning,” in *ICML*, 2009, pp. 41–48.
- [9] T. D. Canton, A. J. Miller, A. P. Lundgren, D. A. Brown, and others., “Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors,” *Phys. Rev.*, vol. D90, no. 8, pp. 082004, 2014.
- [10] S. A. Usman, A. H. Nitz, I. W. Harry, C. M. Biwer, and others., “The PyCBC search for gravitational waves from compact binary coalescence,” *Class. Quant. Grav.*, vol. 33, no. 21, pp. 215004, 2016.
- [11] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional lstm and other neural network architectures,” *Neural Networks*, vol. 18, no. 5-6, pp. 602–610, 2005.
- [12] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le, “Sequence to sequence learning with neural networks,” 2014, NIPS’14, pp. 3104–3112.
- [13] Z. Xia and A. Chakrabarti, “Identifying recurring patterns with deep neural networks for natural image denoising,” *arXiv preprint arXiv:1806.05229*, 2018.
- [14] D. Liu, B. Wen, Y. Fan, C. C. Loy, and T. S. Huang, “Non-local recurrent network for image restoration,” 2018.
- [15] A. Bohé, V. Raymond, L. E. Kidder, M. Pürrer, and others., “Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors,” *Phys. Rev. D*, vol. 95, no. 4, pp. 044028, 2017.
- [16] LSC, “LSC Algorithm Library software packages LAL, LALWRAPPER, and LALAPPS,” .
- [17] David Shoemaker, “Advanced LIGO anticipated sensitivity curves,” 2010.
- [18] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, and others., “Observation of Gravitational Waves from a Binary Black Hole Merger,” *Phys. Rev. Lett.*, vol. 116, no. 6, pp. 061102, 2016.
- [19] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, and others., “GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence,” *Phys. Rev. Lett.*, vol. 116, no. 24, pp. 241103, 2016.
- [20] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, and others., “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems,” 2016.
- [21] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [22] David L. Donoho and Iain M. Johnstone, “Ideal denoising in an orthonormal basis chosen from a library of bases,” *Comptes Rendus Acad. Sci., Ser. I*, vol. 319, pp. 1317–1322, 1994.
- [23] E. A. Huerta, Prayush Kumar, Bhanu Agarwal, Daniel George, Hsi-Yu Schive, Harald P. Pfeiffer, Roland Haas, Wei Ren, Tony Chu, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Mark A. Scheel, and Bela Szilagyi, “Complete waveform model for compact binaries on eccentric orbits,” *Phys. Rev. D*, vol. 95, pp. 024038, Jan. 2017.