Nonlinear noise regression with machine learning at LIGO

Tri Nguyen [1] Michael Coughlin [2] Rich Ormiston [3] Rana Adhikari [2]

January 8, 2019

[1] University of Rochester

[2] California Institute of Technology

[3] University of Minnesota – Twin Cities

Table of contents

Objectives

- Strong GWs produce a displacement of 10^{-18} m, about 1000 times smaller than the diameter of a proton.
- Although LIGO has observed GWs from black-hole and neutron-star mergers, many more still lie below the sensitivity limit.

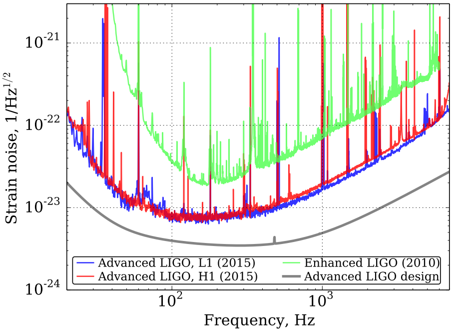


Figure 1: The sensitivity of LIGO Livingston (L1) and LIGO Hanford (H1) during the first observation run O1 [1].

Nonlinear noise regression

- Noise sources are recorded as time series data by physical environmental monitor channels, or **witness channels**.
- Based on the detector's response history, noise regression predicts future response and subtracts the sources.
- Noise sources couple into the output through some physical processes described by a **transfer function**.
- If the transfer function is nonlinear, traditional noise regression can be challenging because the coupling mechanism is often sophisticated.

Goal: Train neural networks to perform linear and nonlinear subtraction on time series data.

Removable v. Non-removable noises

- **Non-Removable noises** define the baseline sensitivity limit. They can only be reduced by improving the detector design.
- Examples: quantum noise, thermal noise, etc.
- **Removable noises** are instrumental and environmental effects. They can be subtracted given there are witness channels monitoring them.
- Examples: seismic noise, magnetic noise, etc.

Goal: Subtract the removable noises while keeping the signals and the non-removable noise intact.

Data

- We perform our analysis on mock and real data.
- Real data are LIGO Hanford (LHO) data during the second observation run on August 14, 2017.
- Mock data are generated by coupling white noises $w(t)$ into the DARM $h(t)$ via the **resonance function**:

$$h(t) = \mathcal{F}^{-1}\left[\frac{\mathcal{F}[w(t)]}{\omega_0^2 - \omega^2 + i\frac{\omega_0 \omega}{Q}}\right]$$

where ω_0 and Q are the angular resonant frequency and the quality factor. $\mathcal{F}$ denotes the Fourier transform.

Visualizing the data

- The input consists of time series from multiple channels.
- Each channel has a duration of 2048 seconds and a sample rate of 512 Hz (total of 1,048,576 samples).

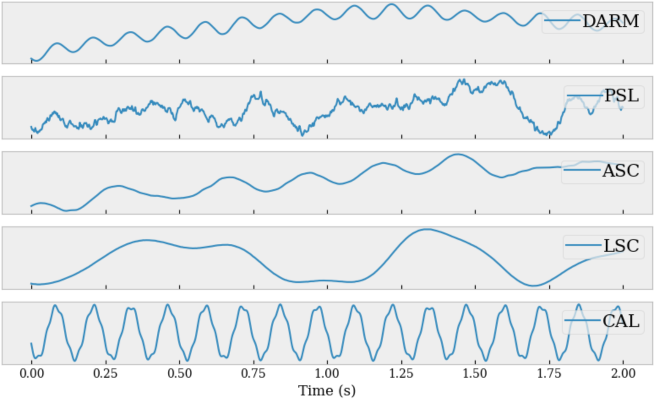


Figure 2: Sample witness channels by subsystems from LHO data on August 14, 2017.

Visualizing the data

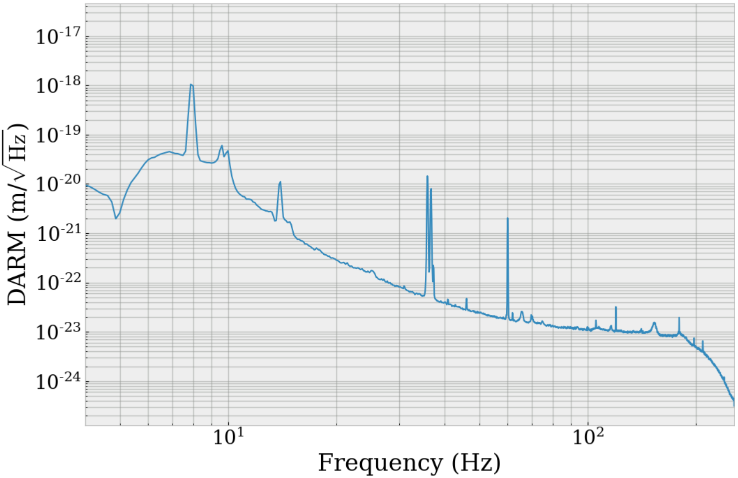


Figure 3: The power spectrum density of LHO data on August 14, 2017.

Methods

Neural networks

- **Neural networks** are a set of non-parametric, supervised machine learning algorithms.
- Neural networks can learn and perform tasks such as data classification and regression.
- A neural network may consist of thousands of interconnected nodes. Each node has parameters to characterize the data.

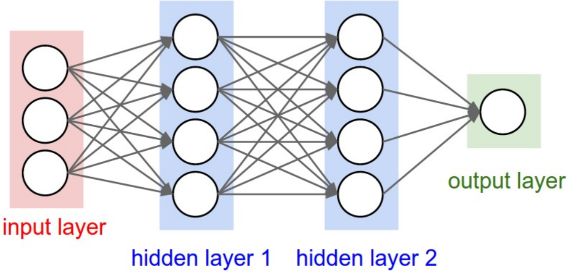


Figure 4: Neural networks block diagram [2].

- **Long Short-Term Memory** (LSTM) networks are special neural networks to process sequential inputs, like time series data [3].
- A major strength of LSTMs is the ability to store and use information from past inputs.

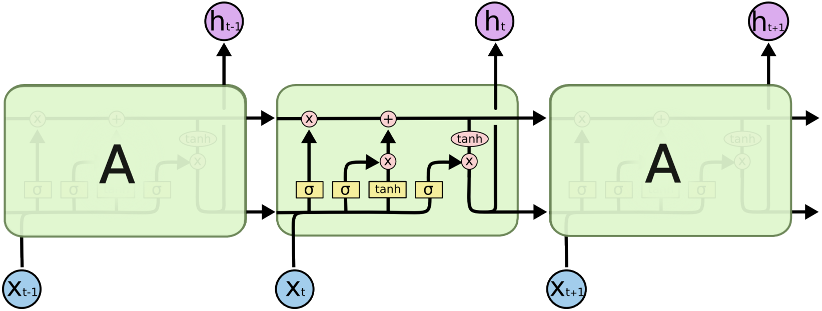
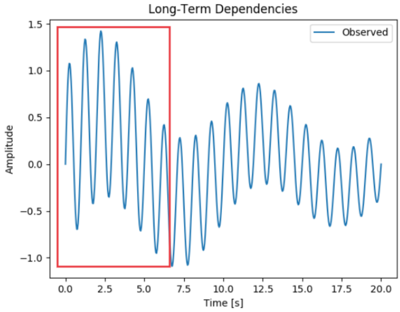


Figure 5: A typical LSTM layer contains four interacting sub-layers [4].

Why do we need LSTMs?

- LIGO noise budget includes slowly migrating signals.
- Example: seismic waves at test masses take several seconds to get to the witness channels.
- Unlike standard neural networks, LSTMs can capture the long-term dependencies in the data.



$$\text{Observed} = e^{-0.1t}\left(\frac{3}{5}\sin(10t) + \sin(t)\right)$$

Figure 6: Seeing only the data inside the red box, an LSTM can still predict the time series (Credit: Rich Ormiston).

Data preprocessing

- Sample-rate conversion: Up-sample or down-sample so all channels have the same sample rate of 512 Hz.
- Bandpass filter: Attenuate frequencies outside a certain range.
- Standard scaling: Normalize the mean and standard deviation of each channel to 0 and 1.
- Lookback window: Divide each channel into short and overlapping series. There are 1,048,546 mini series, each with 16 data points (or 0.03125 seconds).

DEEP CLEAN

Basic workflow

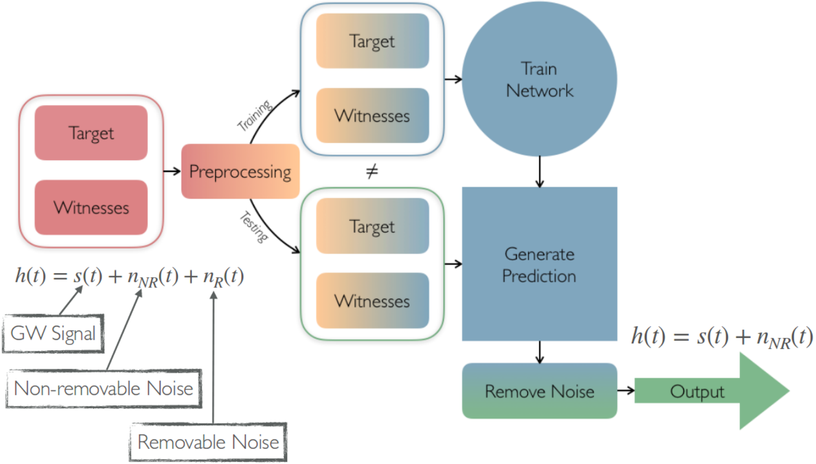


Figure 7: The noise regression pipeline of DeepClean (Credit: Rich Ormiston).

Inside each network

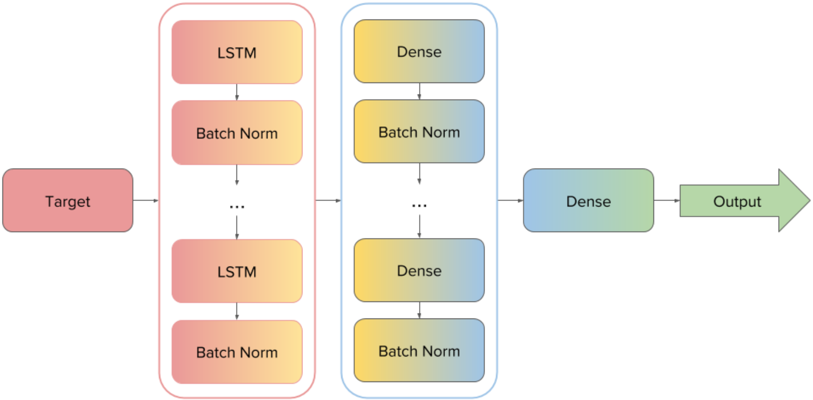


Figure 8: The architecture of each DeepClean network.

Visualizing the output

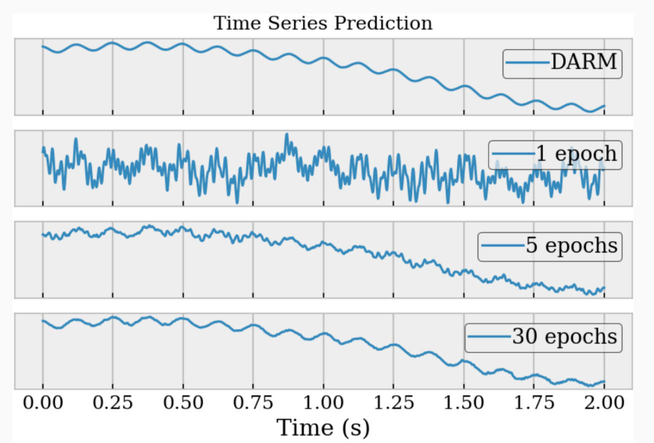


Figure 9: DeepClean LSTM prediction output (Credit: Rich Ormiston).

Results

Mock data

DeepClean successfully subtracted the resonant noise ($\omega_0 = 13.15$ rad/s, $Q = 1000$) and left out the 60-Hz AC power line and the rest of the background (bucket noise).

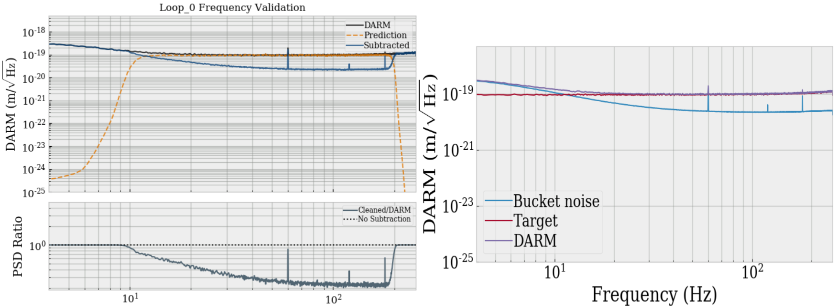


Figure 10: Left: The subtraction in the 10-250 Hz band. Right: The expected outcome.

DeepClean successfully subtracted the 7-Hz calibration line and 60-Hz AC power line.

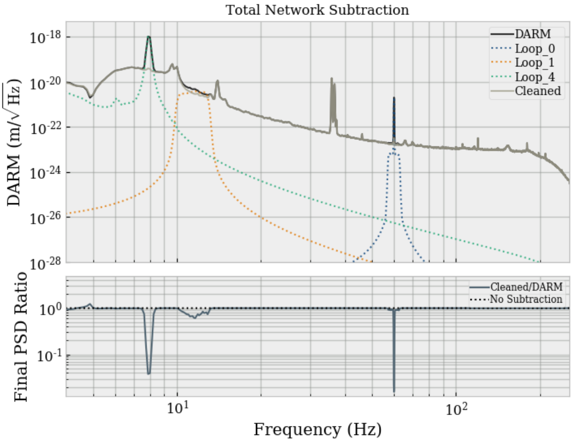


Figure 11: Total subtraction.

LHO data

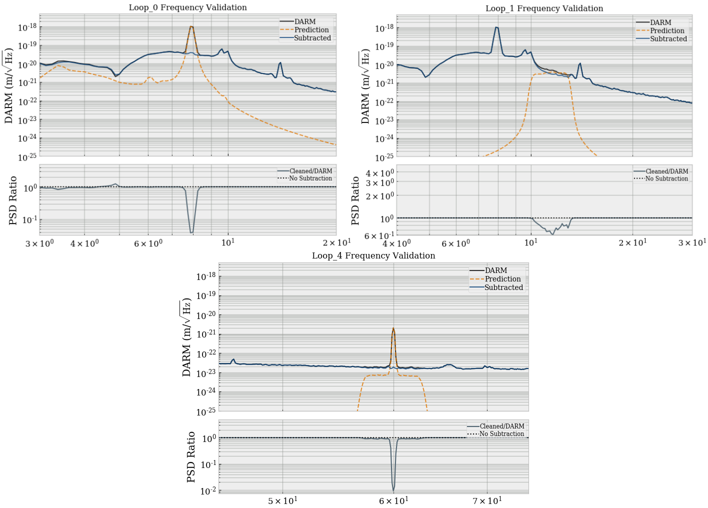


Figure 12: Subtraction on the 3-9 Hz (top left), 10-13 Hz (top right), and 57-63 Hz (bottom) bands.

Conclusion

Future work

- Obtain better subtraction in the 20-100 Hz band $\Rightarrow$ Discover more compact binaries and uncover additional information on past observations.
- Use different mock data sets with different coupling mechanisms.
- Conduct an ablation and/or correlation study to understand which witness channels are important for the subtraction.
- Optimize for real-time noise regression (i.e. analyze an hour of data in less than an hour).

References i

Benjamin P. Abbott et al.
Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy.
Phys. Rev., D93(11):112004, 2016.
[Addendum: Phys. Rev.D97,no.5,059901(2018)].

Andrej Karpathy.
Neural networks part 1: Setting up the architecture, November 2017.

Andrej Karpathy.
The Unreasonable Effectiveness of Recurrent Neural Networks, 2015.

Christopher Olah.
Understanding LSTM Networks, 2015.

- The LIGO Scientific Collaboration.
  Advanced LIGO.
  Classical and Quantum Gravity, 32(7):074001, 2015.
- R. Adhikari.
  Sensitivity and noise analysis of 4 km laser interferometric gravitational wave antennae.
  PhD thesis, Massachusetts Institute of Technology, Massachusetts, USA, 2004.
- Vaibhav Tiwari et al.
  Regression of Environmental Noise in LIGO Data.
  Class. Quant. Grav., 32(16):165014, 2015.

Thank You!