Using machine learning and the Hough transform to search for gravitational waves from isolated neutron stars

Andrew Miller
Advisers: Pia Astone, Bernard Whiting

9 January 2020

NEWS
NEw WindowS on the universe and technological advancements from trilateral EU-US-Japan collaboration
H2020-MSCA-RISE-2016- Grant Agreement No. 734303

Outline

1. Motivation and background
2. Generalized FrequencyHough
3. The first long-lived postmerger remnant search
4. Machine learning: Convolutional Neural Networks
5. First machine learning search for GW170817

1 Motivation and background

2 Generalized FrequencyHough

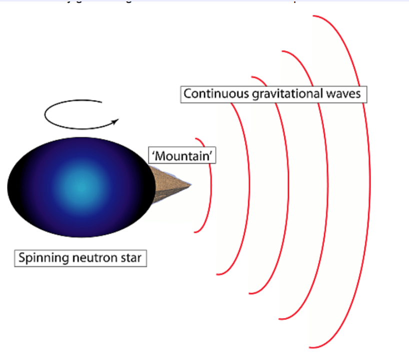
3 The first long-lived postmerger remnant search

4 Machine learning: Convolutional Neural Networks

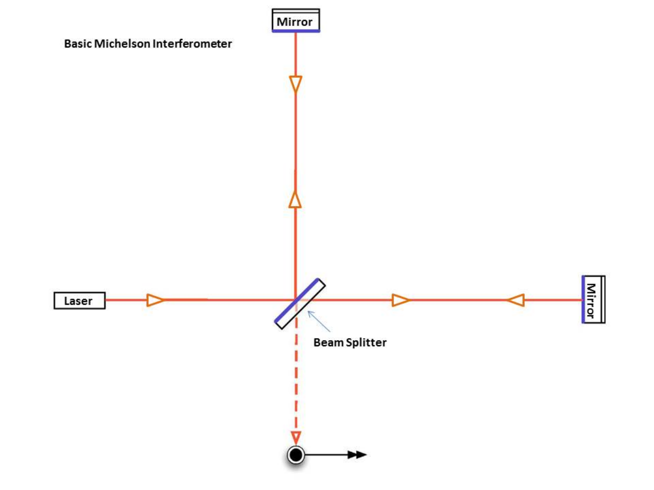
5 First machine learning search for GW170817

Gravitational waves from isolated neutron stars

- Small deformation on the star → gravitational waves (GWs) are radiated [12]
- For older neutron stars, search spindowns/spinups -1×10^{-8} to 2×10^{-9} Hz/s- "continuous waves" [18]
- Model is generally Taylor series expansion of frequency
- For younger neutron stars, $O(10^{-3} - 10^{-1})$ Hz/s, so-called "long duration transients", $O(hours - days)$
- Result of binary neutron star merger or supernova



LIGO/Virgo Design



- We measure a relative displacement of the two arms

The signal model for long duration transients

$$\dot{f} = -kf^n \tag{1}$$

$$f(t) = f_0\left(1 + (n-1)kf_0^{n-1}(t - t_0)\right)^{-1/(n-1)} \tag{2}$$

- $f, \dot{f}$: frequency, spindown
- n: braking index
- k: proportionality constant, some physics is here
- t_0: reference time
- f_0: frequency at t_0
- n indicates emission mechanism [19]:
  - $n = 3 \rightarrow$ rotating magnetic dipole [11]
  - $n = 5 \rightarrow$ GWs due to deformation (ellipticity) [17]
  - $n = 7 \rightarrow$ GWs due to r-modes [16]

1 Motivation and background

2 Generalized FrequencyHough

3 The first long-lived postmerger remnant search

4 Machine learning: Convolutional Neural Networks

5 First machine learning search for GW170817

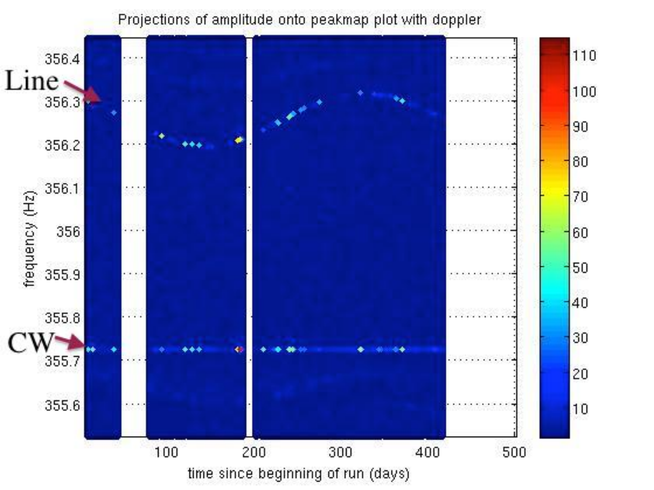
The FrequencyHough

- Maps points in detector's time/frequency plane to lines in source frequency/spindown plane

$$f(t) = f_0 + \dot{f}(t - t_0) + \ldots \quad (3)$$

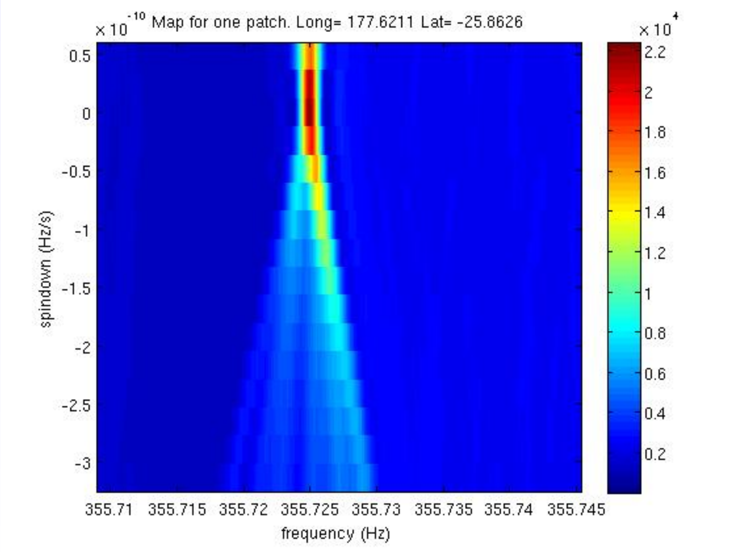
- Slope of line is $-1/(t - t_0)$ in $(f, \dot{f})$ plane
- Line drawn for each spindown in the grid at each time
- Used for continuous wave searches [5]

Doppler-corrected peakmap for continuous wave signal



- Whiten the noise, then select time/frequency points ("peaks") with values above 2.5 and local maxes

Hough map for continuous wave signal



Generalizing the FrequencyHough

Start with [13]:

$$f(t) = f_0 \left(1 + (n-1) k f_0^{n-1} (t - t_0)\right)^{-1/(n-1)}$$

Change coordinates:

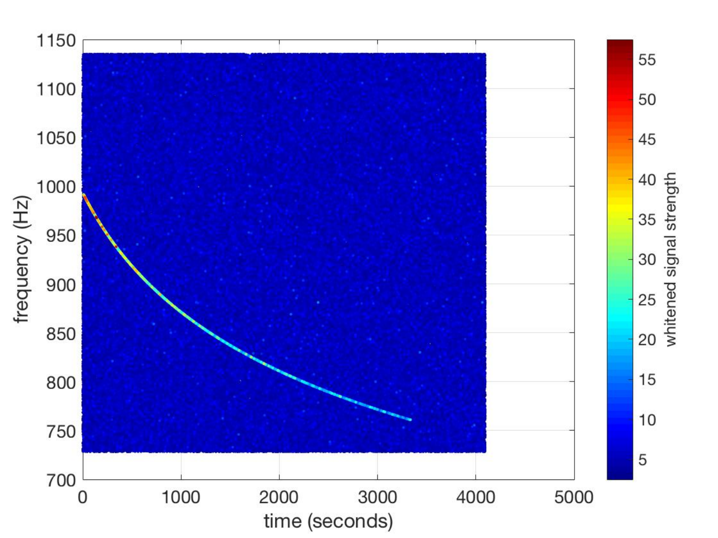
$$x_0 = \frac{1}{f_0^{n-1}}$$

$$x = \frac{1}{f^{n-1}}$$

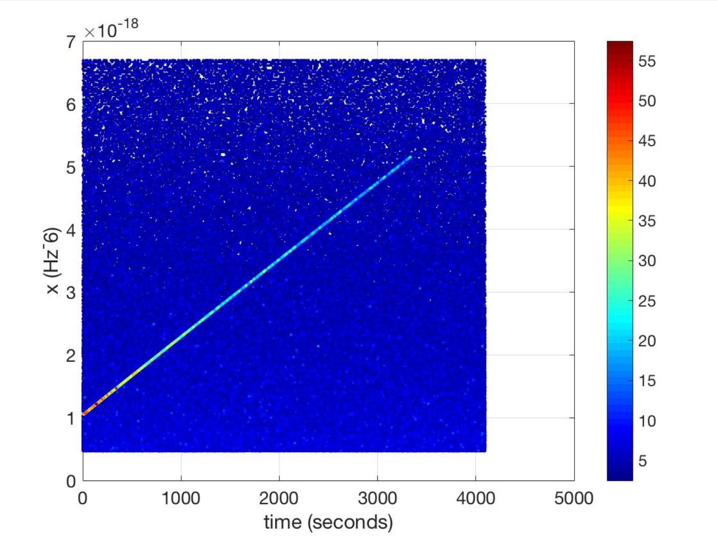
Do some algebra:

$$x = x_0 + (n-1) k (t - t_0)$$

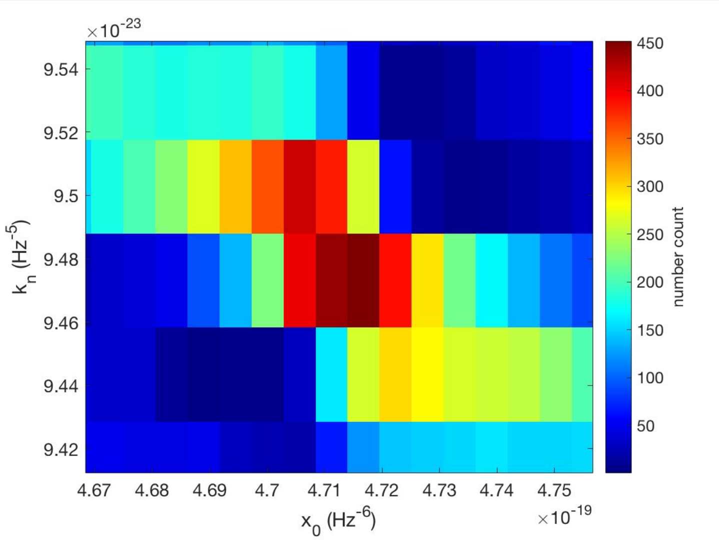
Peakmap before transformation (r-mode, white noise)



Peakmap after transformation



Hough map (zoomed)



1 Motivation and background

2 Generalized FrequencyHough

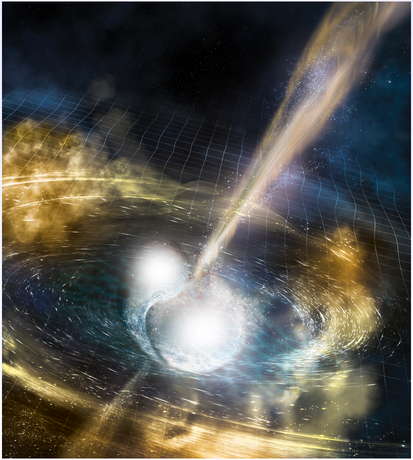
3 The first long-lived postmerger remnant search

4 Machine learning: Convolutional Neural Networks

5 First machine learning search for GW170817

Why search for a remnant?

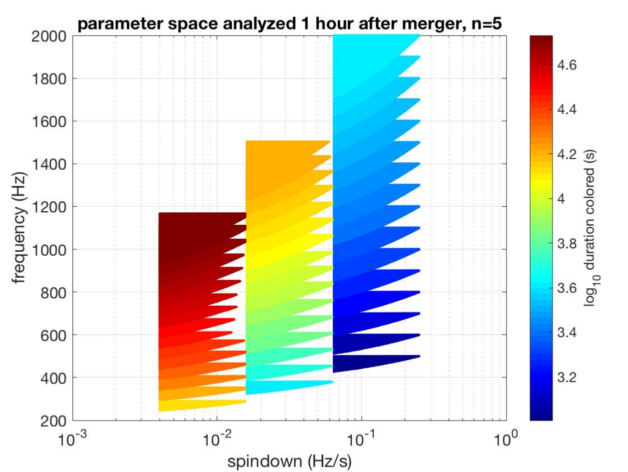
- Kilanova model (r-process) cannot fully explain the spectra: hybrid models considered [20]
- Search for $O(s)$-$O(mins)$ signals done already [1].
- Parameter space explored for long-lived remnant- could be produced with stiff equation of state (EoS) [3]
- Constrain pre/post merger EoS [7]



Search parameters

- f_0: [500, 2000] Hz
- n: [2.5, 7]
- $\dot{f}$: $[-\frac{1}{2^2}, -\frac{1}{16^2}]$ Hz/s
- Began the search about 1 hour after the merger
  - Signal immediately after merger is very complicated
  - Less frequency variation later because spindown is smaller
  - Improves sensitivity because we use longer Fast Fourier Transforms
- Looking for sources lasting at most 1 day
- Grids in x_0, k and n are constructed
- Search run in 3 configurations, one for FFT length (2,4,8 s)

Parameter space explored (1 hour after merger)

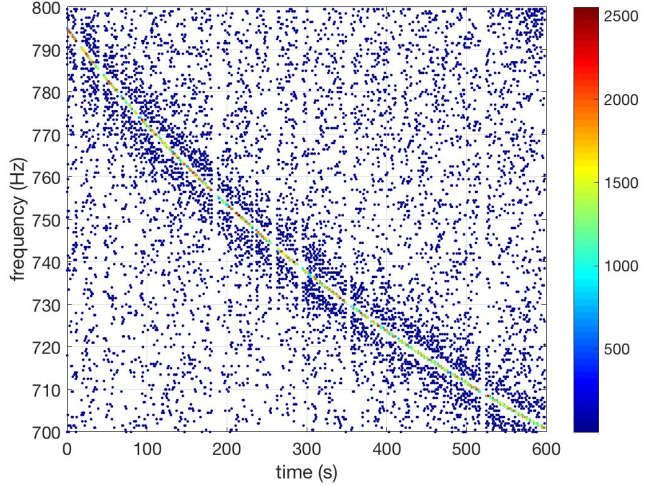


- "Holes": still sensitive to, just not as much
- Done to reduce computation cost and maximize sensitivity

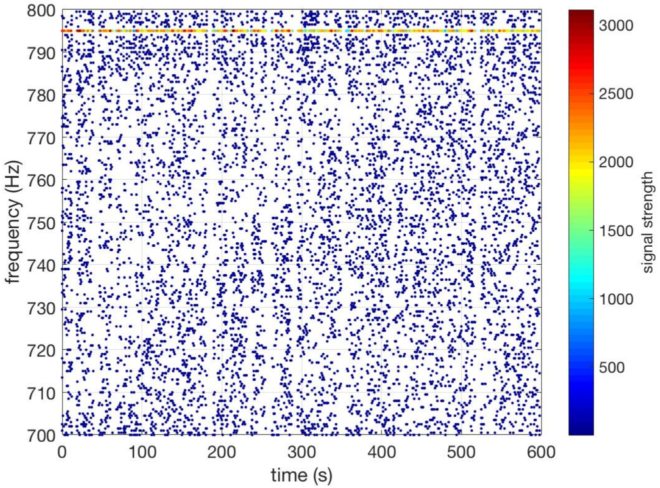
Candidate selection and follow-up

- Candidate = x_0, k, t_0, lon, lat, $\rightarrow f_0$, $\dot{f}_0$
- Nonuniform noise $\rightarrow$ select 1 candidate in each "square" of the Hough map, uniformly in f_0/k
- 25 – 50 candidates per map
- Coincidences between detectors' candidates and each configuration
- Correct for phase evolution of signal

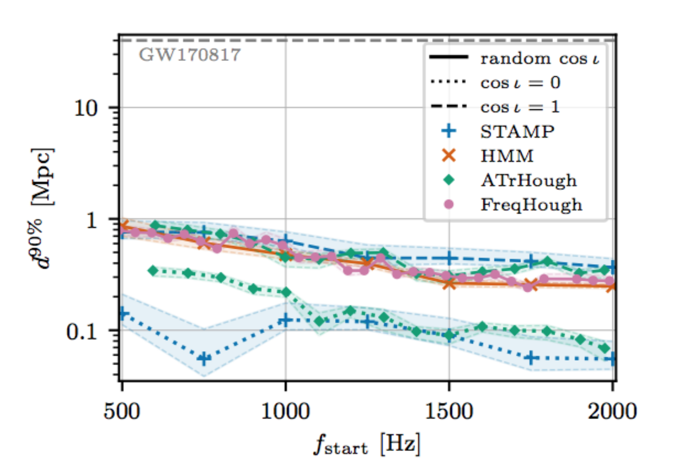
Peakmap before follow-up



Peakmap after follow-up

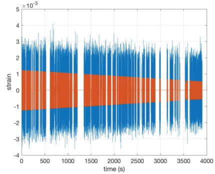


Upper limits

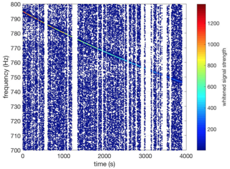


- 4 pipelines (2 modelled and 2 unmodelled) searched for a remnant [2]

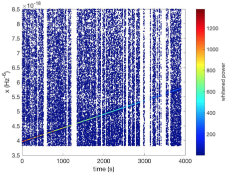
Method summary



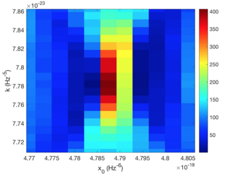
(a) Time series strain O2 Livingston data



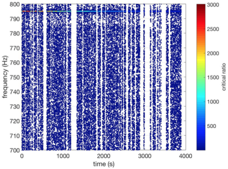
(b) Time/frequency map made with $T_{FFT} = 8$ s.



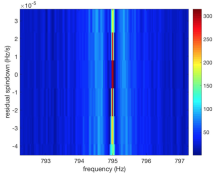
(c) Transformation of 9-13(b) using equation 9–1



(d) Hough map showing injected signal's parameters



(e) Phase-corrected time/frequency map using 9-13(d)



(f) Hough map, phase-corrected signal's parameters

1. Motivation and background
2. Generalized FrequencyHough
3. The first long-lived postmerger remnant search
4. Machine learning: Convolutional Neural Networks
5. First machine learning search for GW170817

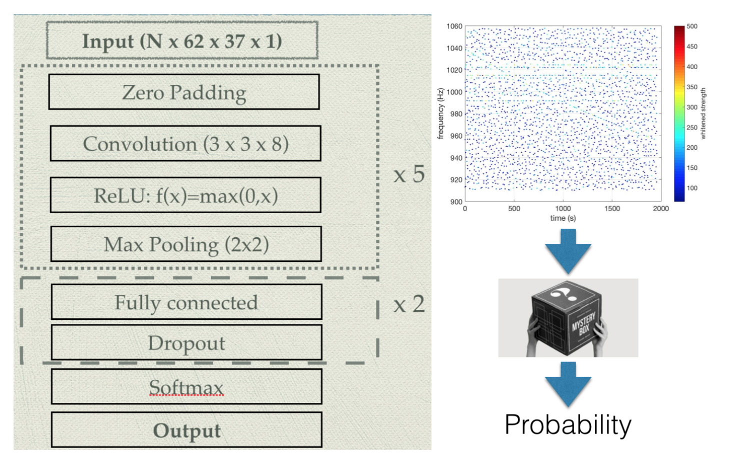
Why machine learning?

- Unmodeled approach to detecting GWs
- Modeled searches are slow, computationally expensive, and not ideal if you don't fully trust your model
- Can see signals with time-varying braking indices
- Lots of applications already in GW physics
  - Neural networks, support vector machines [15]
  - Convolutional Neural Networks (CNNs) for binary black hole detection [9, 10]
  - Determining if neutron star remnant remains after a merger [8]
- Here we are extending the work in [15] to CNNs

Key questions not well answered in GW machine learning

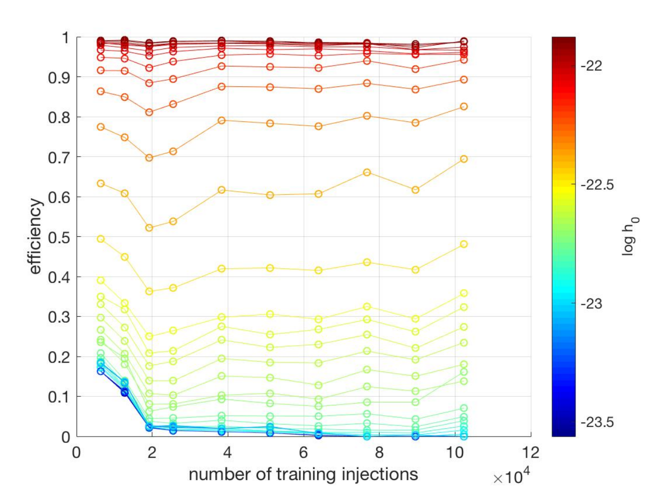
- How much to train?
- How to train (types of architectures)?
- How robust, effects on false alarm probability
- Effects of unpredictable noise/ finite data
- How can we use them in a search?

Convolutional neural network architecture



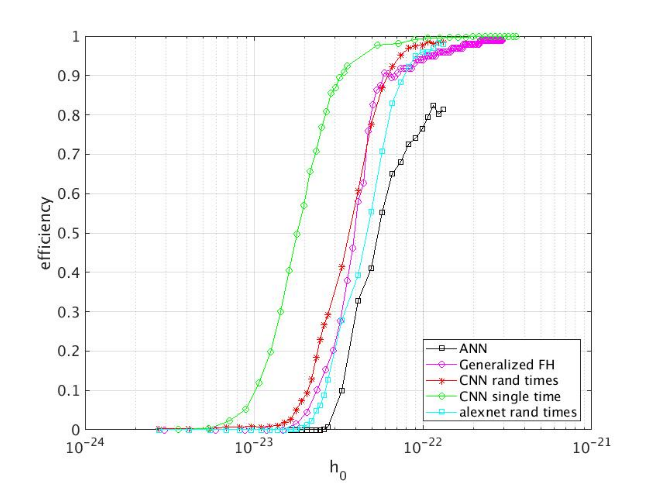
- Input: time/frequency map
- Output: probability of signal p_{out}, apply threshold $p_{thr} = 0.9$ to control false alarm probability (FAP)
- Architecture used in [4, 14]

How much to train?



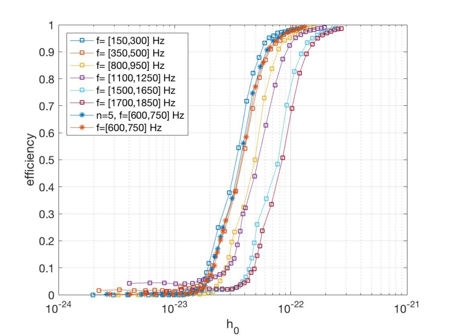
- Not many injections are needed

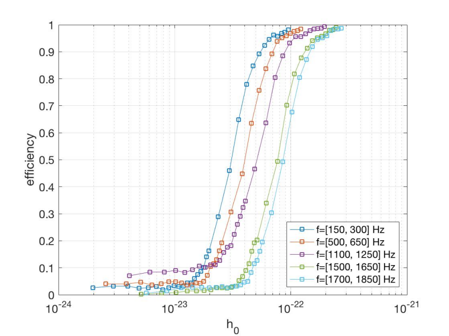
A comparison of the networks with the Hough



- The efficiencies are similar but the false alarm probabilities are different: $O(1\%)$ vs. $O(0.01\%)$
- Green = no noise

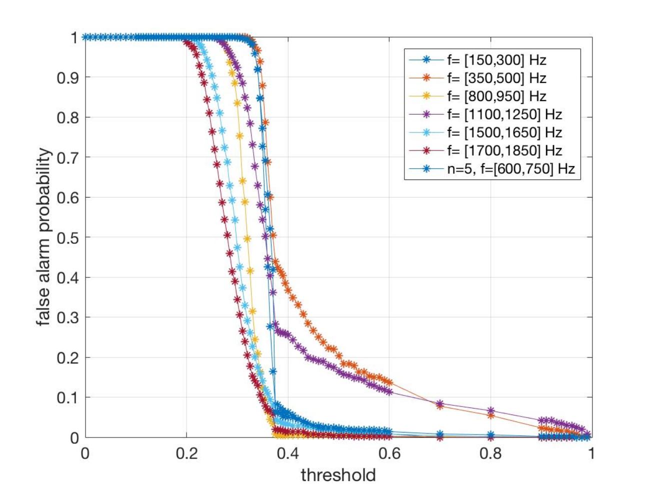
Are the networks robust?





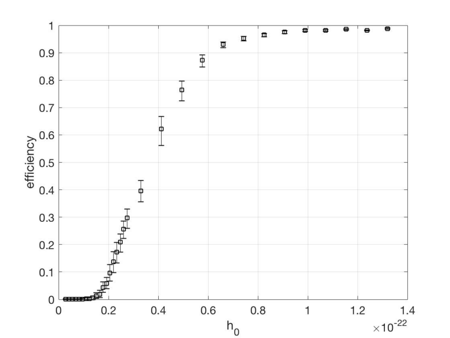
- The network is robust towards signals with different behaviors than those on which it was trained.
- Even variations in *n* can be seen with the network.
- Tails represent the false alarm probability

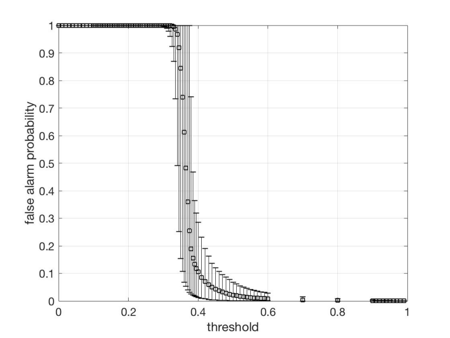
How wrong can the networks be?



- We choose the threshold to be 0.9

How bad is the error?



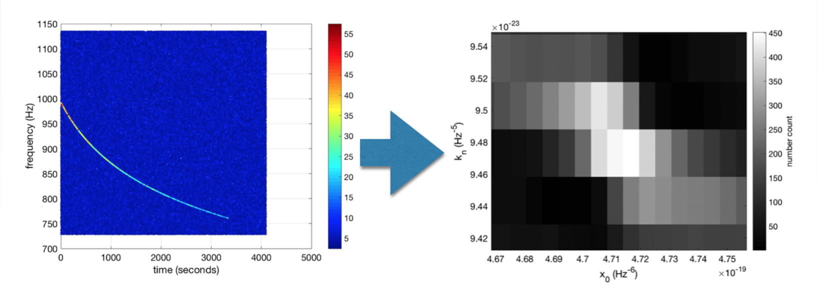


- The false alarm probability and false dismissal probability must be computed empirically
- Choosing threshold of 0.9 allows repeatable experiments

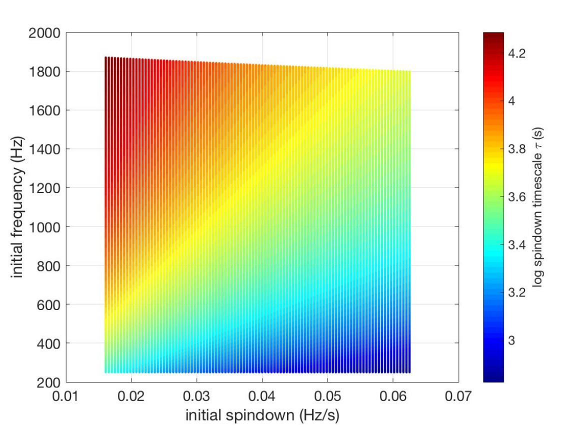
1. Motivation and background
2. Generalized FrequencyHough
3. The first long-lived postmerger remnant search
4. Machine learning: Convolutional Neural Networks
5. First machine learning search for GW170817

Search design

- Start with Short Fast Fourier Transform Database (SFDB) [6]
- Choose T_{FFT}, construct 2000 s x 150 Hz time/frequency maps, give to CNNs
- Look for coincident maps in H/L when $p_{out} > p_{thr}$
- For triggered maps, perform follow-up using Generalized FrequencyHough Transform [13] to estimate parameters

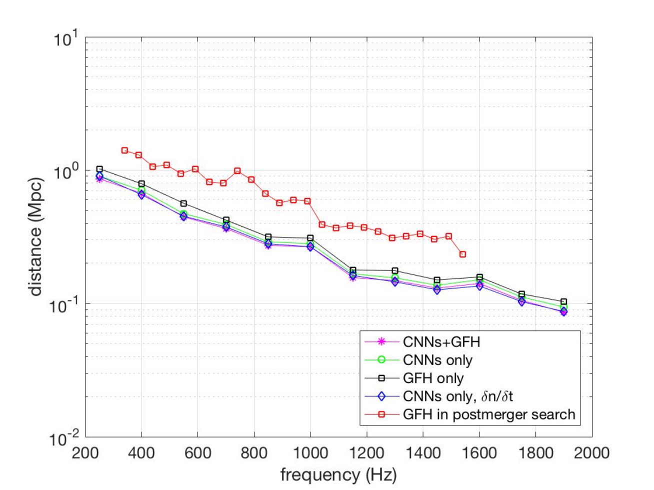


Parameter space explored and search design



- Searched the 1 week of data after GW170817 in 2000 s pieces
- Made a network for each detector, then performed coincidences in time between maps with output > 0.9

Upper limits at 50% confidence

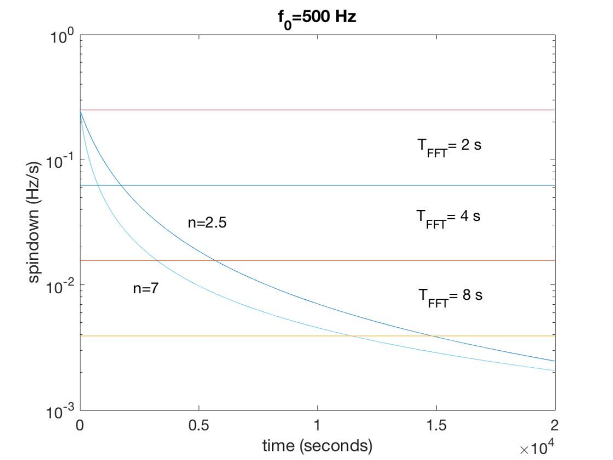


Future work

- Expanding classification of neural networks to include separate categories for glitches and time-varying braking indices
- Parameter estimation of GW signal using machine learning
- In the event of detection, develop ways to extract meaningful signal parameters
    - Fitting techniques at time-domain level, weighting of different power laws, etc.

Backup Slides

Choice of Fast Fourier Transform length



- Time/frequency bands chosen for a spindown window and n

Grid in x_0

$$x_0 = \frac{1}{f_0^{n-1}} \tag{4}$$

$$dx_0 = \frac{1-n}{f_0^n} df_0 \tag{5}$$

- Changes as a function of frequency, but we over-resolve the grid by taking $f_0 = f_{0,max}$ so the grid is uniform
- Computationally, doesn't increase burden of analysis

Grid in k

- The idea: in the $log_{10}\ f$-$log_{10}\ \dot{f}$ space, the power law equation forms lines:

$$\begin{aligned} \dot{f} &= -kf^n \\ log\ |\dot{f}| &= n\ log\ f + log\ k \\ y &= nx + b \end{aligned}$$

- Consider $f \rightarrow f + df$ and $k \rightarrow k + dk$
- Find dk so that spindown remains constant when shifting one frequency bin df for a fixed n

$$dk = k\left[\left(1 + \frac{df}{f}\right)^{-n} - 1\right] \tag{6}$$

$$dk \approx -nk\frac{df}{f} \tag{7}$$

Grid on n

- Calculated numerically with the criteria: stepping from n to $n + dn$ will result in a signal's frequency varying by less than $df = 1/T_{FFT}$
- For each point in this grid, we do a Hough

Coincidences

- Define a distance in the parameter space as:

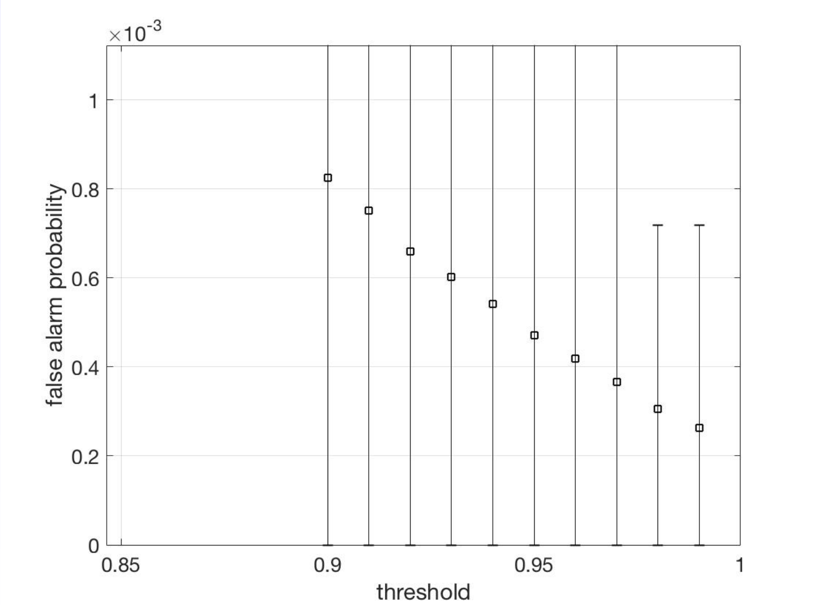
$$d = \sqrt{\left(\frac{x_2 - x_1}{\delta x}\right)^2 + \left(\frac{k_2 - k_1}{\delta k}\right)^2} \tag{8}$$

- If 2 candidates within a certain distance, a coincidence has occurred

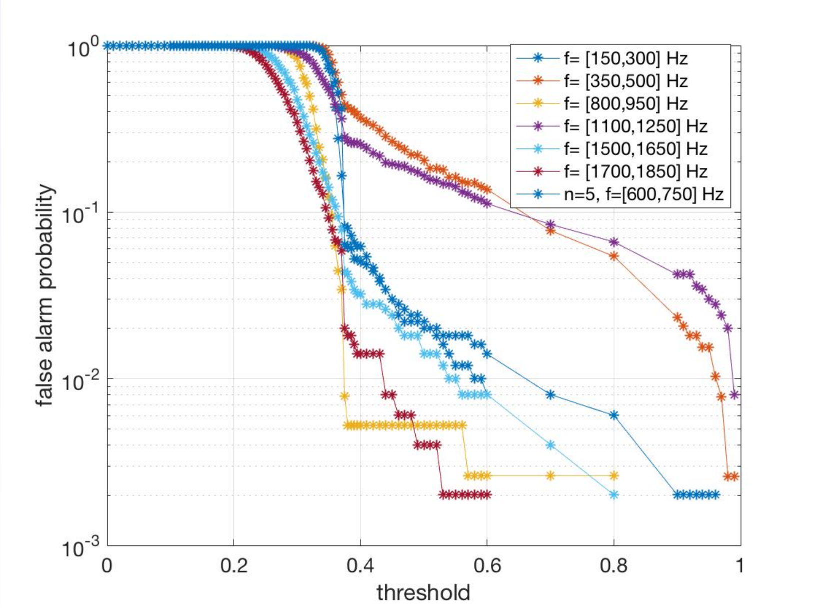
Doing the follow-up

- Look for coincidences among candidates in each configuration run, 3 bins
- Each candidate has $n, f_0, \dot{f}_0, t_0, \Delta t \rightarrow$ we know $f(t)$
- Correct the phase of the signal in time domain
- Increase T_{FFT} depending on frequency band and what kind of grid we consider around candidate
- Ideally, we would get a horizontal line in the time/frequency plane (monochromatic signal) at f_0
- In practice, frequency will spin up/ down near true f_0
- Do original Hough

Zoomed-in CNN errorbars plot



Log-scale CNN false alarm probability plot



[1] B. P. Abbott et al.
Search for post-merger gravitational waves from the remnant of the binary neutron star merger GW170817.
The Astrophysical Journal, 851(1):L16, dec 2017.

[2] B. P. Abbott et al.
Search for gravitational waves from a long-lived remnant of the binary neutron star merger GW170817.
The Astrophysical Journal, 875(2):160, apr 2019.

[3] S. Ai, H. Gao, Z.-G. Dai, X.-F. Wu, A. Li, B. Zhang, and M.-Z. Li.
The allowed parameter space of a long-lived neutron star as the merger remnant of gw170817.
The Astrophysical Journal, 860(1):57, 2018.

[4] P. Astone, P. Cerdá-Durán, I. Di Palma, M. Drago, F. Muciaccia, C. Palomba, and F. Ricci.
New method to observe gravitational waves emitted by core collapse supernovae.
Physical Review D, 98(12):122002, 2018.

[5] P. Astone, A. Colla, S. D'Antonio, S. Frasca, and C. Palomba.
Method for all-sky searches of continuous gravitational wave signals using the frequency-hough transform.
Physical Review D, 90(4):042002, 2014.

[6] P. Astone, S. Frasca, and C. Palomba.
The short fft database and the peak map for the hierarchical search of periodic sources.
Classical and Quantum Gravity, 22(18):S1197, 2005.

[7] A. Bauswein, O. Just, H.-T. Janka, and N. Stergioulas.
Neutron-star radius constraints from gw170817 and future detections.
arXiv preprint arXiv:1710.06843, 2017.

[8] D. Chatterjee, S. Ghosh, P. R. Brady, S. J. Kapadia, A. L. Miller, S. Nissanke, and F. Pannarale.
A machine learning based source property inference for compact binary mergers.
arXiv preprint arXiv:1911.00116, 2019.

[9] H. Gabbard, M. Williams, F. Hayes, and C. Messenger.
Matching matched filtering with deep networks for gravitational-wave astronomy.
Physical review letters, 120(14):141103, 2018.

[10] D. George and E. Huerta.
Deep neural networks to enable real-time multimessenger astrophysics.
Physical Review D, 97(4):044039, 2018.

[11] P. D. Lasky, C. Leris, A. Rowlinson, and K. Glampedakis.
The braking index of millisecond magnetars.
The Astrophysical Journal Letters, 843(1):L1, 2017.

[12] M. Maggiore.
Gravitational Waves: Volume 1: Theory and Experiments, volume 1.
Oxford university press, 2008.

[13] A. Miller, P. Astone, S. D'Antonio, S. Frasca, G. Intini, I. La Rosa, P. Leaci, S. Mastrogiovanni, F. Muciaccia, C. Palomba, O. J. Piccinni, A. Singhal, and B. F. Whiting. Method to search for long duration gravitational wave transients from isolated neutron stars using the generalized frequency-hough transform. *Phys. Rev. D*, 98:102004, Nov 2018.

[14] A. L. Miller, P. Astone, S. D'Antonio, S. Frasca, G. Intini, I. La Rosa, P. Leaci, S. Mastrogiovanni, F. Muciaccia, A. Mitidis, et al. How effective is machine learning to detect long transient gravitational waves from neutron stars in a real search? *Physical Review D*, 100(6):062005, 2019.

[15] A. Mytidis, A. A. Panagopoulos, O. P. Panagopoulos, and B. Whiting.
Sensitivity study using machine learning algorithms on simulated r-mode gravitational wave signals from newborn neutron stars.
arXiv preprint arXiv:1508.02064, 2015.

[16] B. J. Owen, L. Lindblom, C. Cutler, B. F. Schutz, A. Vecchio, and N. Andersson.
Gravitational waves from hot young rapidly rotating neutron stars.
Physical Review D, 58(8):084020, 1998.

[17] C. Palomba.
Pulsars ellipticity revised.
Astronomy and Astrophysics, 354:163–168, 2000.

[18] A. Pisarski, L. S. Collaboration, V. Collaboration, et al.
All-sky search for continuous gravitational waves from isolated neutron stars using advanced ligo o2 data.
arXiv preprint arXiv:1903.01901, 2019.

[19] L. Stella, S. Dall'Osso, G. Israel, and A. Vecchio.
Gravitational radiation from newborn magnetars in the virgo cluster.
The Astrophysical Journal Letters, 634(2):L165, 2005.

[20] Y.-W. Yu, L.-D. Liu, and Z.-G. Dai.
A long-lived remnant neutron star after gw170817 inferred from its associated kilonova.
The Astrophysical Journal, 861(2):114, 2018.