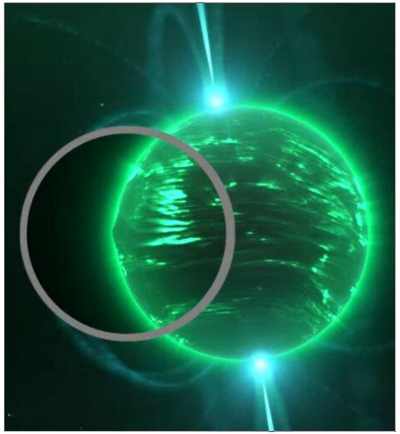
A novel neural-network architecture for continuous gravitational waves

Prasanna Joshi, Reinhard Prix
Date: July 11, 2023

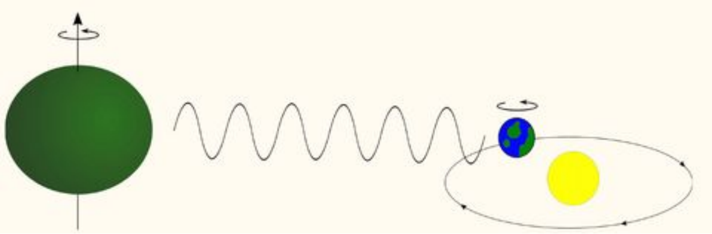
https://arxiv.org/abs/2305.01057

Introduction

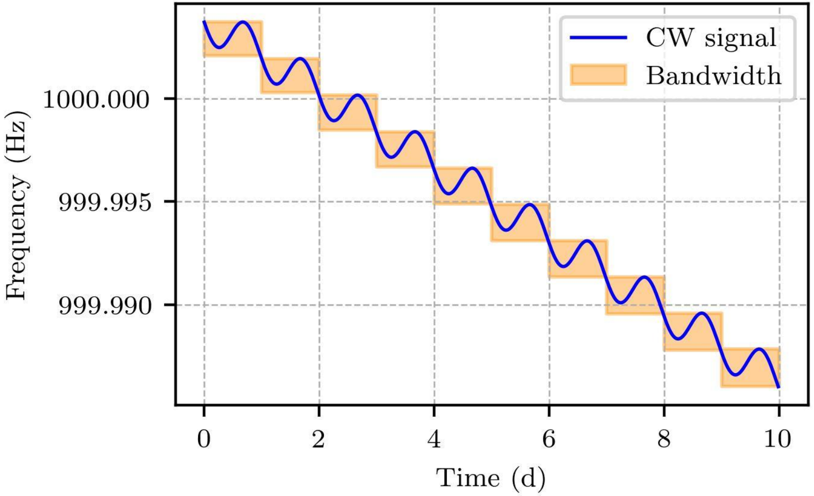


$$h_0 = \frac{16\pi^2 G}{c^4} \frac{I_{zz}\nu^2}{d}\varepsilon\,,$$

$$\mathcal{D} \equiv \frac{\sqrt{S_n}}{h_0}$$



Credit: Greg Ashton



Targeted Search test cases

We trained **10** separate networks, one for each of the **10** targeted search cases.

- Freq (Hz): 20, 100, 200, 500, 1000
- Spindown: -10^{-10} Hz/s
- Skypos: Sky-A (narrow), Sky-B (widest)
- Time-span: 10 days

Sensitivity Depth for 90% detection probability:

$$\mathcal{D}^{90\%}_{\mathrm{Sky-A}} \approx 86.2\,/\sqrt{\mathrm{Hz}},$$
$$\mathcal{D}^{90\%}_{\mathrm{Sky-B}} \approx 81.8\,/\sqrt{\mathrm{Hz}}.$$

Frequency [Hz]	Bandwidth [mHz]	
	Sky-A	Sky-B
20	0.09	0.45
100	0.16	1.94
200	0.38	3.80
500	1.06	9.38
1000	2.19	18.69

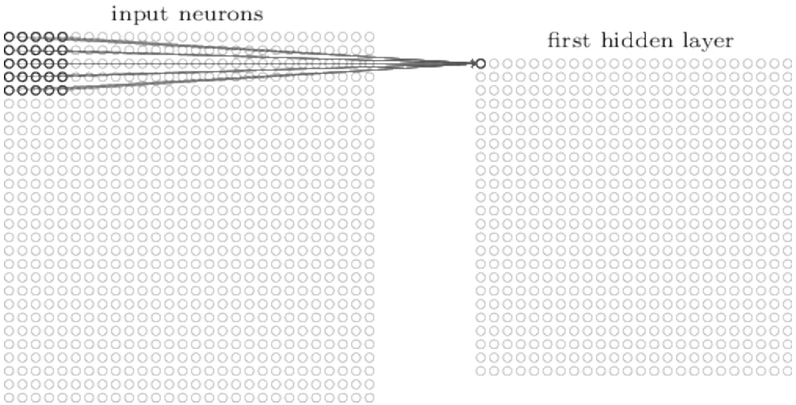
Network Architecture

Inspiration from networks used for Image Recognition

- Convolutional layers
- Small kernel-sizes
- Very deep networks containing ~100 layers

Applied in:

[C. Dreissigacker, et.al, Phys. Rev. D 100, 044009 (2019)](#)
[C. Dreissigacker, et.al, Phys. Rev. D 102, 022005 (2020)](#)



Credit: Neural networks and Deep Learning by Andrew Nielsen

Network Architecture

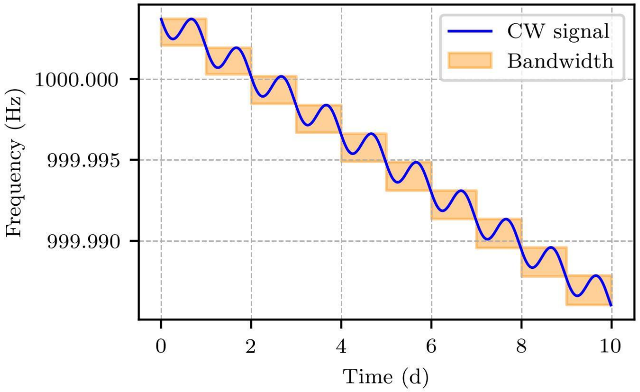
Inspiration from networks used for Image Recognition

- Convolutional layers
- Small kernel-sizes
- Very deep networks containing ~100 layers

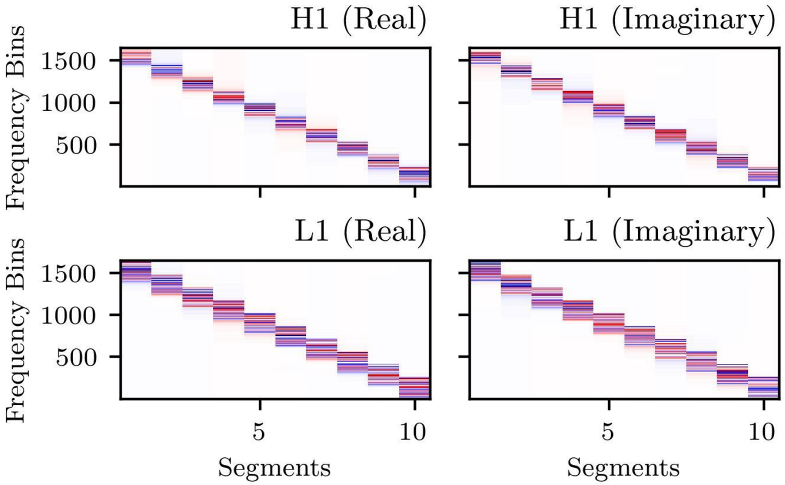
Our network with design features for continuous-wave search

- Convolutional layers
- Large kernel-sizes
- Shallow network containing only 2 convolutional layers and total 5 layers

Pre-processing the input



Detector time-series

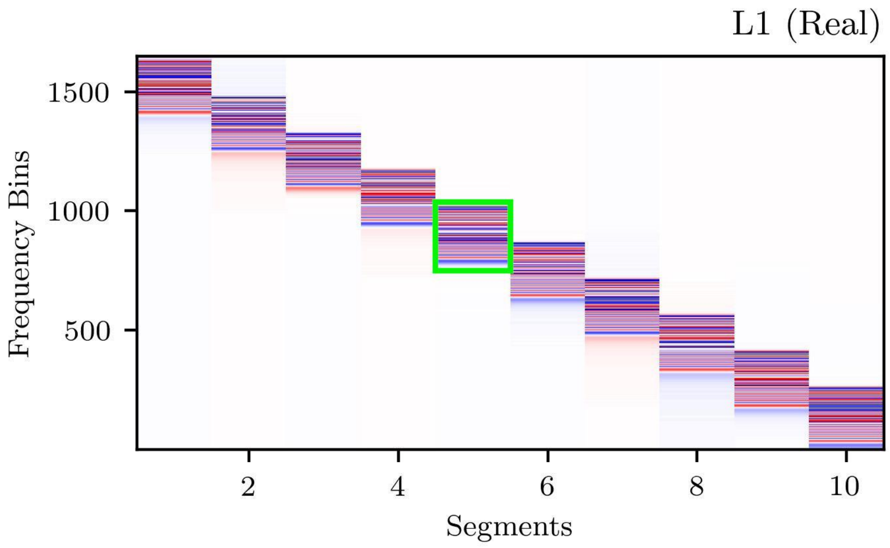


Spectrogram over ten one-day segments

Network Architecture

First layer: 1D convolution

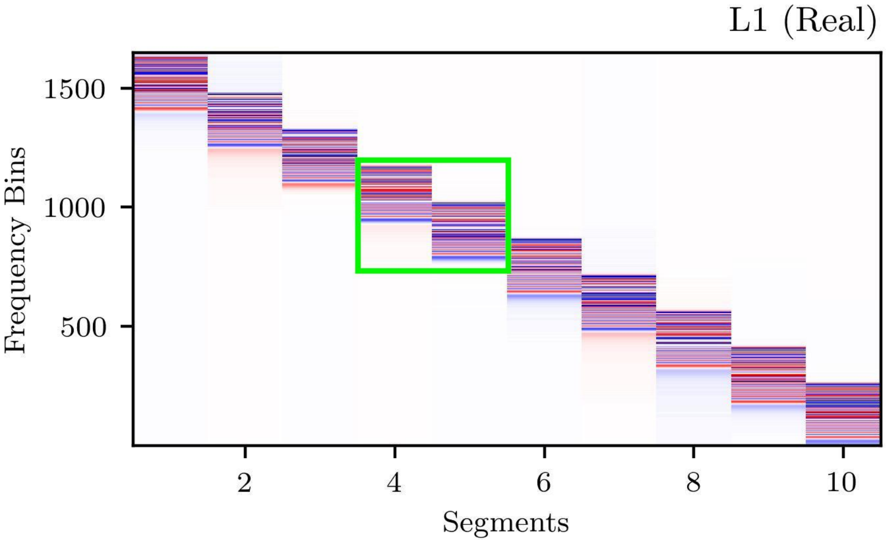
- Convolution along the frequency-axis
- Kernel-size = 1x313 bins
- Kernel completely overlaps the widest signal-bandwidth in the one-day segments (3.4 mHz)
- Number of kernels = 64



Network Architecture

Second layer: 2D convolution

- Kernel-size = 2x40 bins
- “visual field” of kernel overlaps the widest output “track” width over the two-day span
- Number of kernels = 64



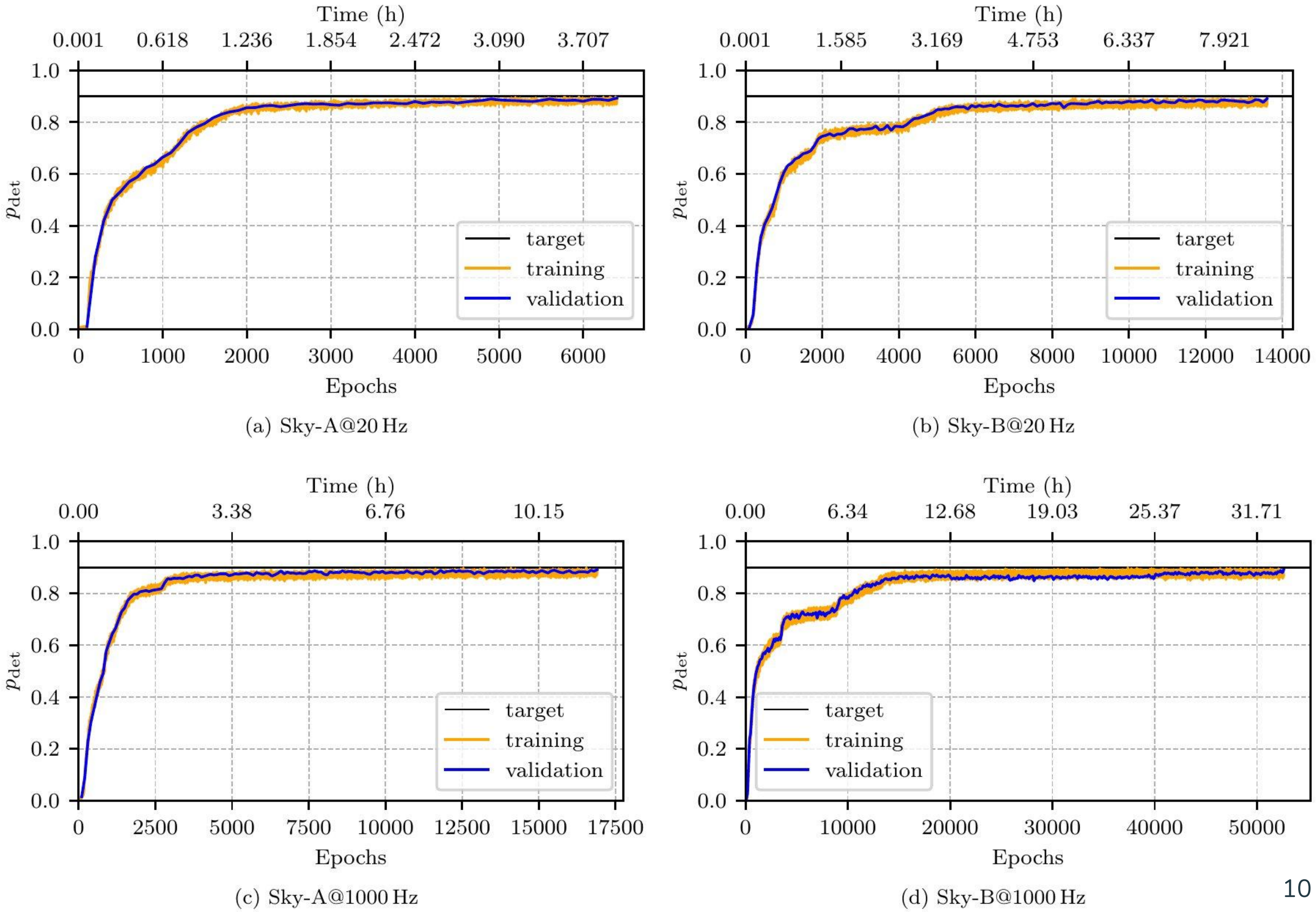
Network Architecture

- Each layer (except final) has ReLU activation function
- Final layer has the softmax activation function
- Inference based on whether network output value crosses a 1% false-alarm threshold

Layer	Output shape (T, F, C)
Input	(10, 1647, 4)
Conv1D Kernel - (1, 313, 2) Stride - 16	(10, 103, 64)
Conv2D Kernel - (2, 40, 64) Stride - (1, 4)	(10, 26, 64)
Flatten	(16640)
Dense	(32)
Output	(1)

Results

Detection probability (at 1% false-alarm probability) during the training for four of our benchmark test cases



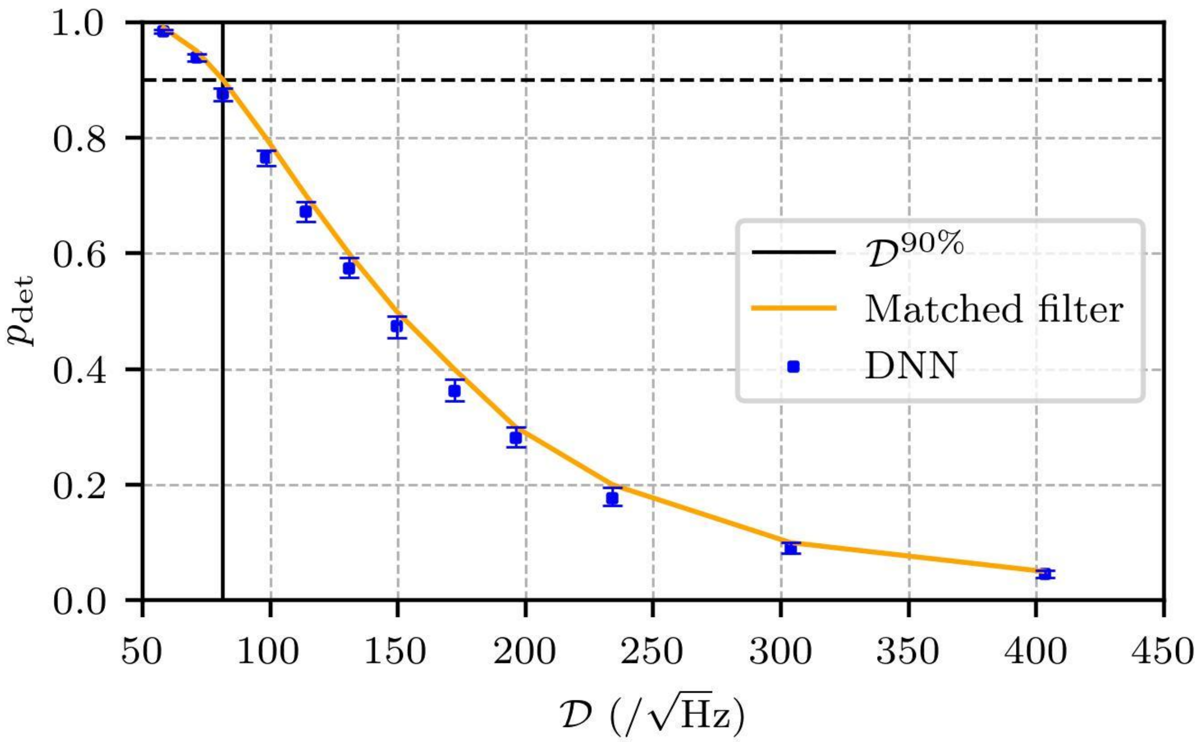
Results

Detection Probability (at 1% false-alarm probability) evaluated on a test set

Frequency	Sky-A	Sky-B
20	$89.0^{+0.8}_{-1.2}$	$88.5^{+0.8}_{-1.0}$
100	$87.8^{+0.8}_{-1.1}$	$87.4^{+1.0}_{-1.0}$
200	$89.0^{+0.8}_{-1.0}$	$89.0^{+0.9}_{-1.0}$
500	$88.4^{+0.7}_{-1.0}$	$88.8^{+1.0}_{-0.9}$
1000	$87.6^{+0.8}_{-1.1}$	$87.6^{+1.0}_{-1.2}$

Results

Detection probability (at 1% false-alarm probability) as a function of sensitivity depth of the injected signal



Future Work

- Train deep neural networks to efficiently search for CW signals over a longer time span (upto 1 year)
- Develop wide-parameter space searches for CW signals using deep neural networks
- Train deep neural networks to estimate parameters of continuous-wave signals

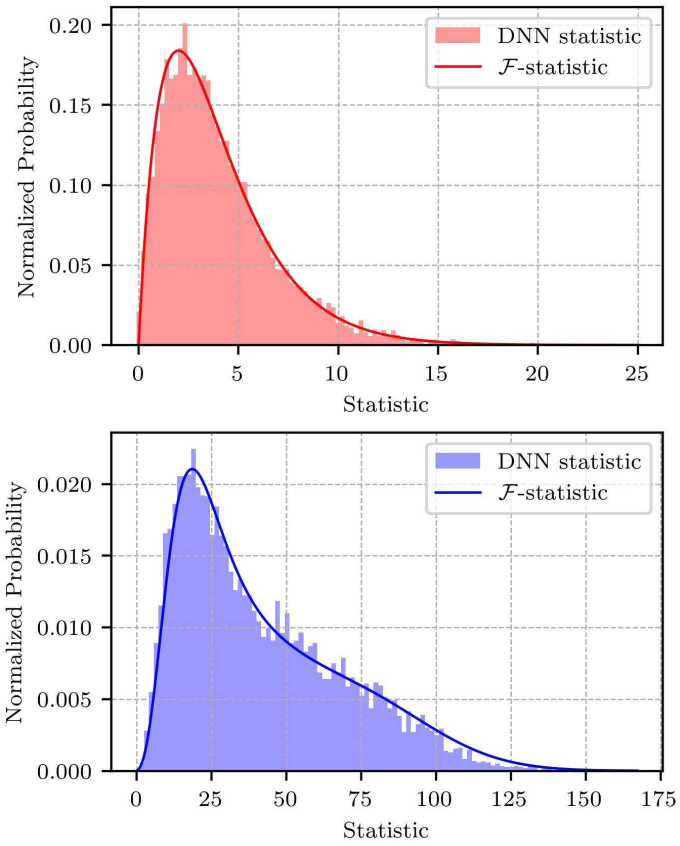
Thank you!

Extra Slides

DNN Training

- No. of signals for training = 8192
- No. of signals for validation = 8192
- Loss function: Binary Cross-entropy
- Optimizer: Adam

Results



DNN statistic
$\mathcal{F}$-statistic
Normalized Probability
Statistic
0.20
0.15
0.10
0.05
0.00
0
5
10
15
20
25
DNN statistic
$\mathcal{F}$-statistic
Normalized Probability
Statistic
0.020
0.015
0.010
0.005
0.000
0
25
50
75
100
125
150
175