Neural Networks for gravitational wave detection

Ondřej Zelenka

in collaboration with Bernd Brügmann, Frank Ohme, Marlin B. Schäfer and Alexander H. Nitz

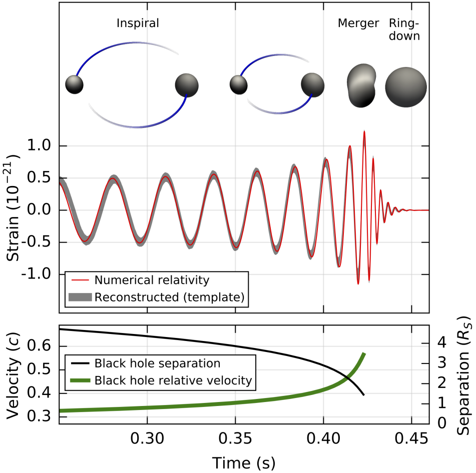
Astronomical Institute of the Czech Academy of Sciences
Theoretical Physics Institute, University of Jena, Germany

4 June 2024

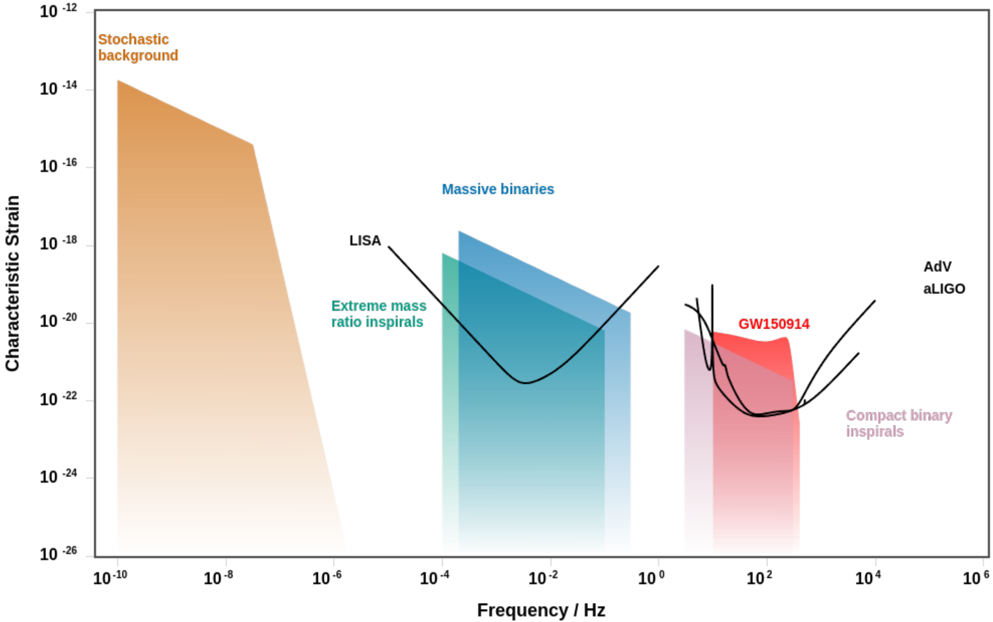
CzechLISA 2024 summer meeting

Gravitational waves

- disturbances in spacetime geometry
- sources: accelerating massive objects
- most frequent: BBH mergers
- crucial for cosmology, SMBH/AGN formation etc.



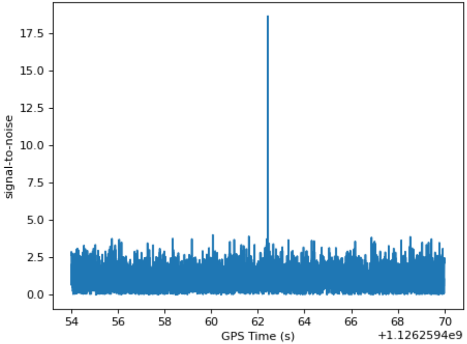
source: [A+16]



source: gwplotter.com

Gravitational wave searches

- binary black holes
- standard: matched filtering
  - dense template bank
  - slide templates over signal
  - non-optimal
  - computationally demanding
- complicated noise characteristic, glitches



source: PyCBC

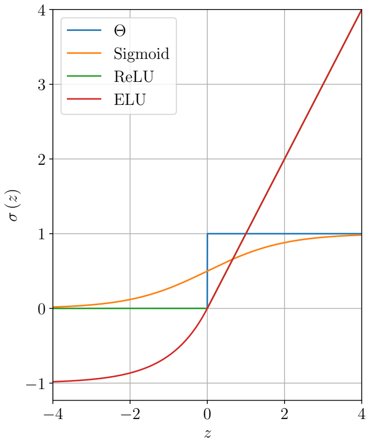
Neural Networks

artificial neurons

$$f(\mathbf{x}) = \sigma\left(\sum_j w_j x_j + b\right)$$

activation function

$$\mathrm{ELU}(z) = \begin{cases} \exp(z) - 1 & \text{if } z < 0 \\ z & \text{if } z \geq 0 \end{cases}$$



Neural Networks

organized in layers:

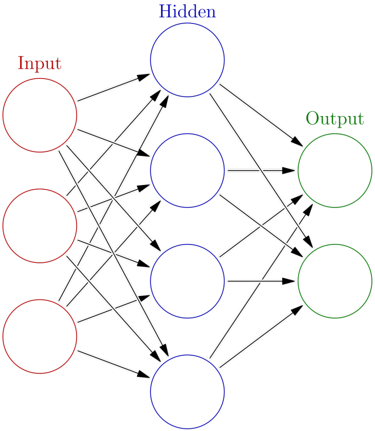
- matrix multiplication
- element-wise non-linear functions

training:

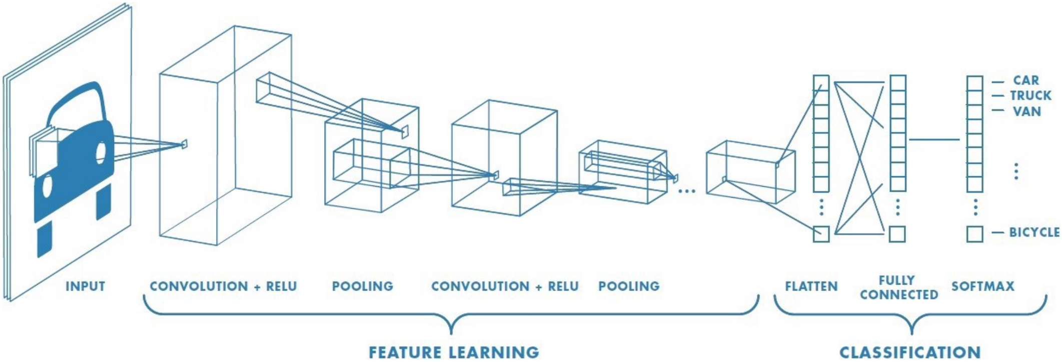
- dataset
- loss function (BCE)

$$\mathrm{BCE}\left(\bar{\mathbf{Y}}, \mathbf{Y}\right) = -\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{2}\log\left(\bar{Y}_{ij}\right) Y_{ij}$$

- GD-based optimizer



Convolutional Neural Networks



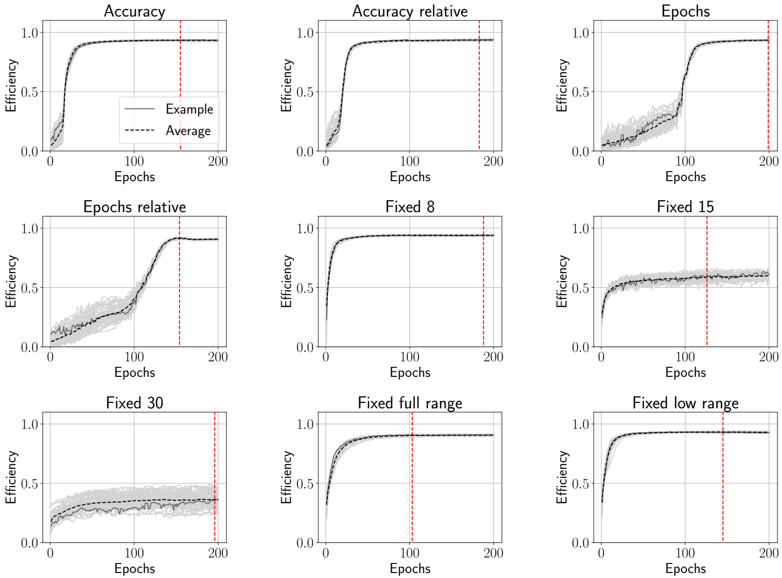
source: MathWorks

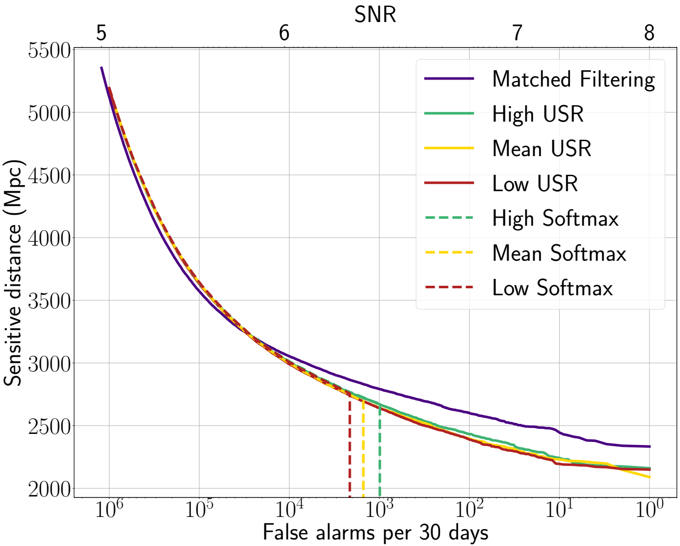
Single-detector searches

- [SZN+22, SZN+21]
- 1 second at 2048 $\mathrm{Hz} \to [0, 1]^2$
- dataset
  - Gaussian noise
  - EOB injections
  - 50% with injection $\mapsto (1, 0)$
  - 50% pure noise $\mapsto (0, 1)$

layer type	output shape
BatchNorm	1×2048
2× Convolution + ELU	32×2034
MaxPool	32×508
2× Convolution + ELU	16×478
MaxPool	16×159
2× Convolution + ELU	16×97
MaxPool	16×48
Flatten	768
Dense + Dropout + ELU	64
Dense + Dropout + ELU	64
Dense + Softmax	2

Training strategies




SNR
5
6
7
8
5500
5000
4500
4000
3500
3000
2500
2000
Sensitive distance (Mpc)
10^6
10^5
10^4
10^3
10^2
10^1
10^0
False alarms per 30 days
Matched Filtering
High USR
Mean USR
Low USR
High Softmax
Mean Softmax
Low Softmax

MLGWSC-1

- mock data challenge [SZN+23, SZ21]
- BBH signals in long segments from 2 detectors
- 4 test datasets, progressive complexity
- 4 ML + 2 traditional contributions

Test data

- 1-month strain in H1 and L1
- `IMRPhenomXPHM` injections

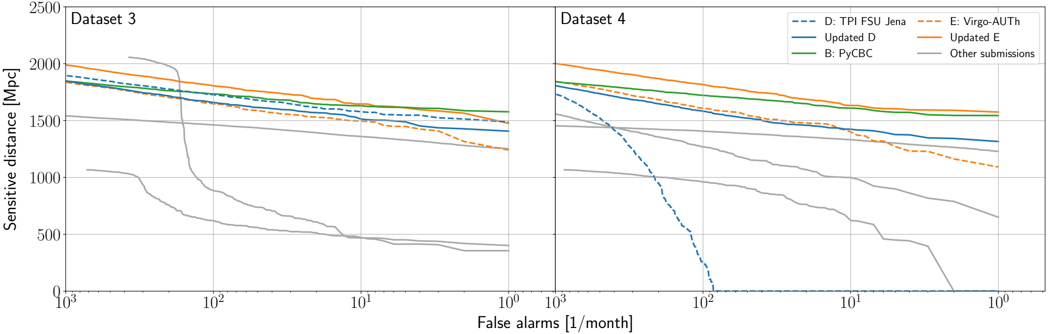
DS	injections			noise	
	spins	modes	masses	noise	PSD
1	zero	$l = 2,\ m = \pm 2$	$10 - 50 M_\odot$	Gaussian	fixed
2	aligned	$l = 2,\ m = \pm 2$	$7 - 50 M_\odot$	Gaussian	O3a
3	generic	all implemented[1]	$7 - 50 M_\odot$	Gaussian	O3a
4	generic	all implemented	$7 - 50 M_\odot$	real O3a noise	

[1] (2, 2), (2, -2), (2, 1), (2, -1), (3, 3), (3, -3), (3, 2), (3, -2), (4, 4), (4, -4)

TPI FSU Jena submission

- [ZBO24, SZN+22, Zel23]
- 1 second segments at 2048 Hz
- real noise from O3a
- waveforms of DS 3 and 4

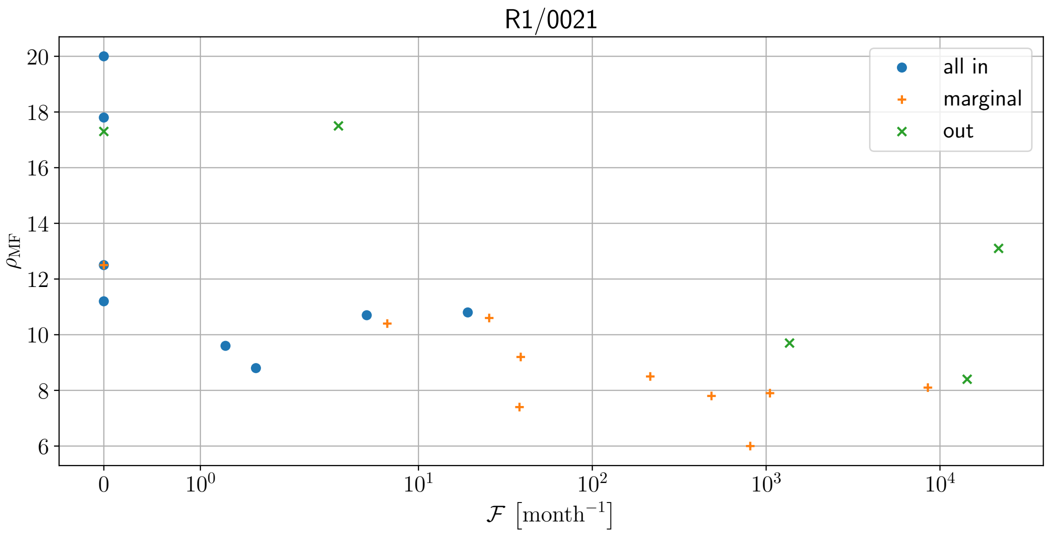
layer type	output shape
Input	2×2048
$4\times$ (Convolution + ELU)	16×1954
MaxPool	16×488
$4\times$ (Convolution + ELU)	32×442
MaxPool	32×147
$4\times$ (Convolution + ELU)	64×117
MaxPool	64×58
Flatten	3712
Dense + Dropout + ELU	128
Dense + Dropout + ELU	128
Dense + (Softmax)	2



Dataset 3
Dataset 4
D: TPI FSU Jena
Updated D
B: PyCBC
E: Virgo-AUTh
Updated E
Other submissions
Sensitive distance [Mpc]
False alarms [1/month]
2500
2000
1500
1000
500
0
10^3
10^2
10^1
10^0

Application to O3b data

- 1 November 2019 - 27 March 2020: $\sim$ 147 days
- high data quality in both L1 and H1: $\sim$ 95 days
- cropped GWTC-3 [A^+21]: 31 confident events
- 90% intervals: $m_i \in [10M_\odot, 50M_\odot]$?
  - all in: found 8/9
  - marginal: found 10/11
  - out: found 5/11


R1/0021
all in
marginal
out
ρ_{MF}
20
18
16
14
12
10
8
6
0
10^0
10^1
10^2
10^3
10^4
$\mathcal{F}$ [month^{-1}]

Conclusions

- ML-based searches competitive on Gaussian noise
- handling real noise not fully understood
- speed

Outlook for Earth-based searches

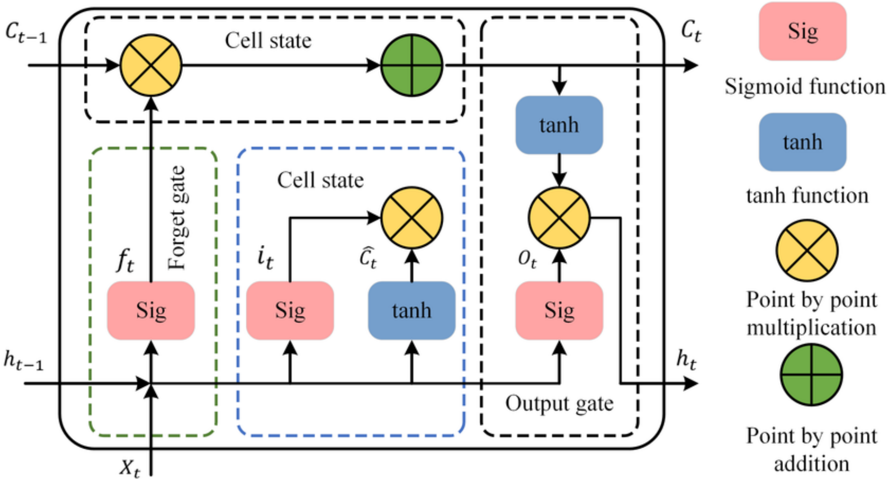
- hybrid ML/MF pipelines
- curriculum learning
- longer signals (BNS...)

Extension to LISA data

Problems:

- longer signals
    - computational complexity
    - motion of detector $\rightarrow$ larger effective parameter space
- overlap
    - need for full separation of signals

Recurrent architecture: LSTM



source: [RXH+24]

References I

B. P. Abbott et al.
Observation of Gravitational Waves from a Binary Black Hole Merger.
Phys. Rev. Lett., 116(6):061102, 2016.

R. Abbott et al.
GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run.
11 2021.

Amir Rehman, Huanlai Xing, Mehboob Hussain, Nighat Gulzar, Muhammad Adnan Khan, et al.
Hcdp-delm: Heterogeneous chronic disease prediction with temporal perspective enabled deep extreme learning machine.
Knowledge-Based Systems, 284:111316, 2024.

Marlin Schäfer and Ondřej Zelenka.
MLGWSC-1 Github repository, 2021.

Marlin B. Schäfer, Ondřej Zelenka, Alexander H. Nitz, Frank Ohme, and Bernd Brügmann.
Training Strategies for Deep Learning Gravitational-Wave Searches, 2021.

Marlin B. Schäfer, Ondřej Zelenka, Alexander H. Nitz, Frank Ohme, and Bernd Brügmann.
Training strategies for deep learning gravitational-wave searches.
Phys. Rev. D, 105:043002, February 2022.

References II

Marlin B. Schäfer, Ondřej Zelenka, Alexander H. Nitz, He Wang, Shichao Wu, et al.
First machine learning gravitational-wave search mock data challenge.
Phys. Rev. D, 107:023021, January 2023.

Ondřej Zelenka, Bernd Brügmann, and Frank Ohme.
Convolutional Neural Networks for signal detection in real LIGO data.
2 2024.
accepted in Phys. Rev. D.

Ondřej Zelenka.
ml-gw-search github repository release v1.1, 2023.