Alberto Iess

A machine learning approach to Gravitational Wave physics

AI goes MAD – Instituto de Física Teórica, Universidad Autónoma de Madrid - CSC

OUTLINE

- Gravitational Wave
- A Brief Introfuction To Gravitational Wave Detection
- **Machine Learning In Gravitational Wave Physics**

What are Gravitational Waves?

Propagating ripples in the curvature of spacetime, generated by accelerated masses.

(gravitational analogue of electromagnetic waves)

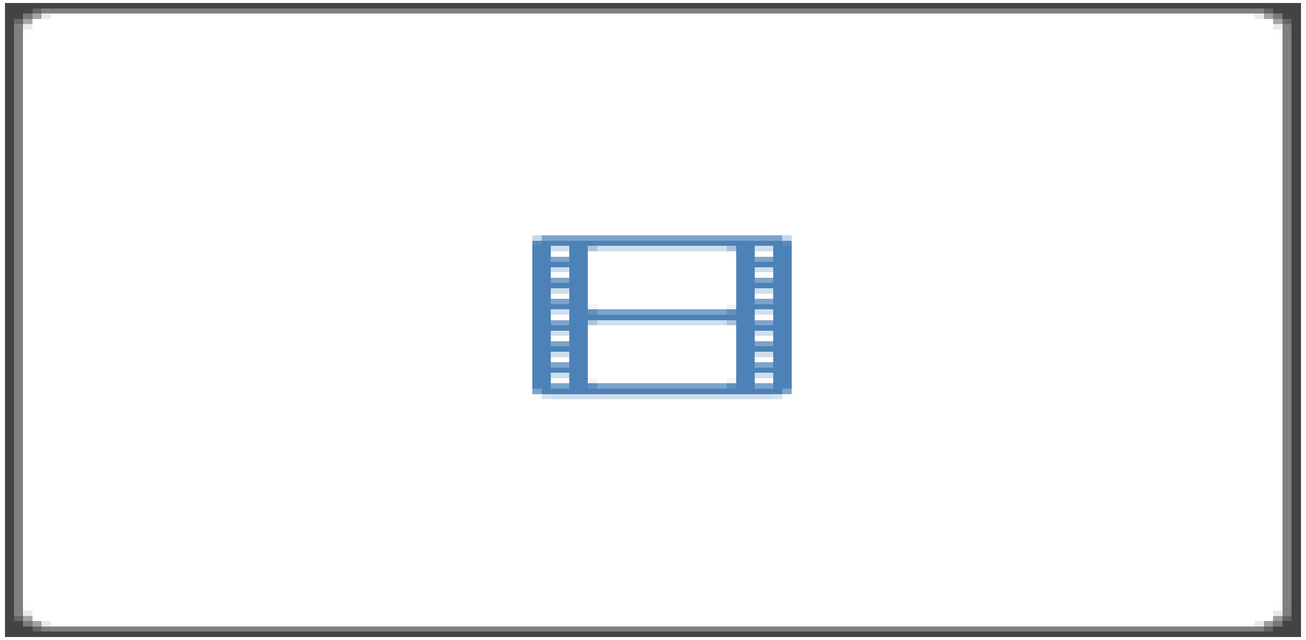


Image credit: NASA Goddard Space Flight Center

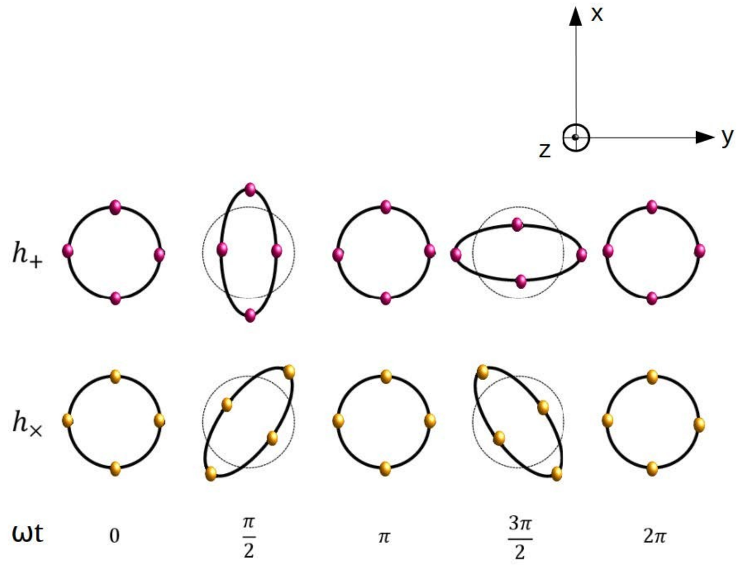
GRAVITATIONAL WAVES

GWs are a solution of the Einstein Field Equation:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}$$

Perturbation of the flat background metric:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad ; \quad |h_{\mu\nu}| \ll 1$$



Belahcene (2019)

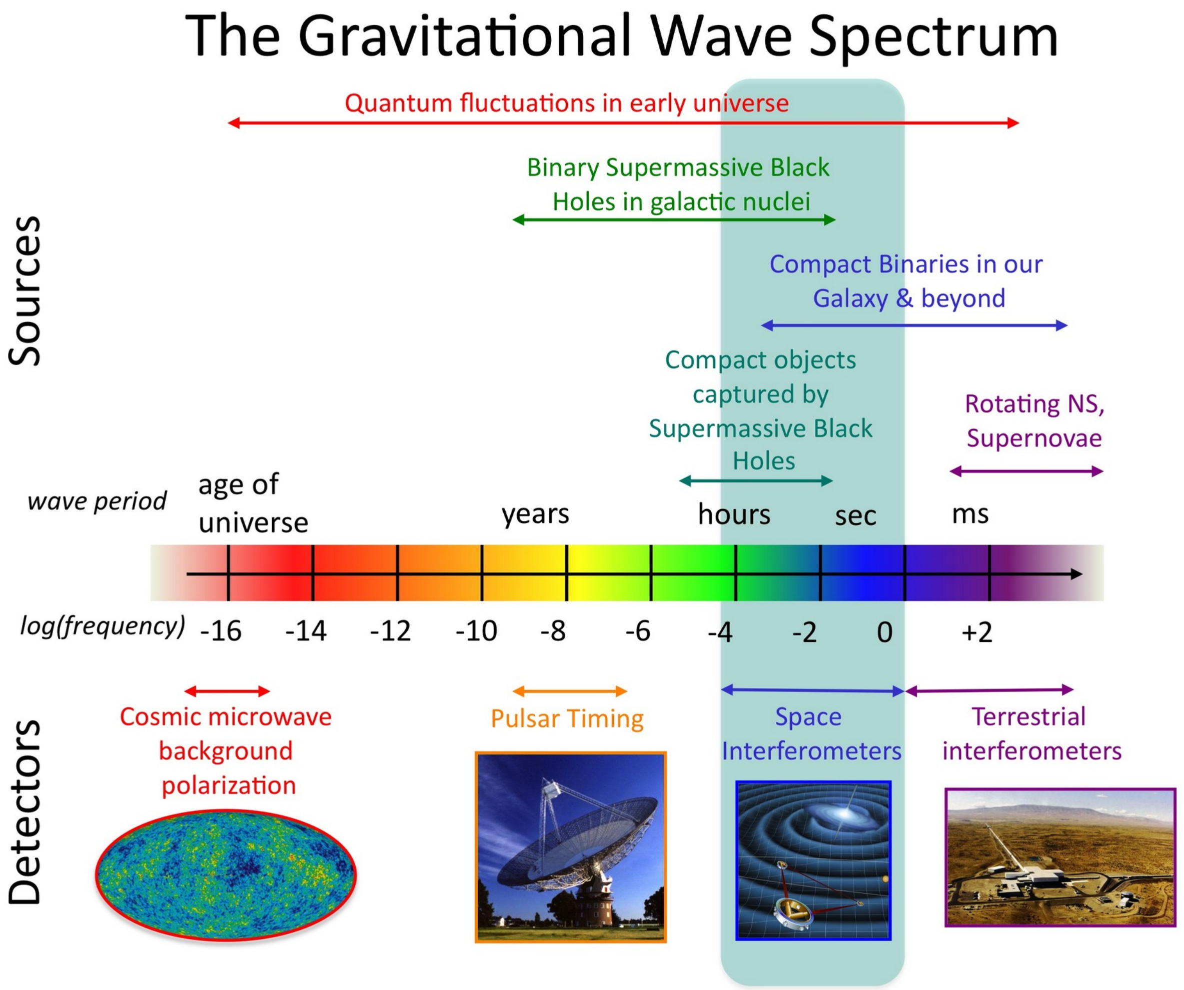


Image credit: NASA Goddard Space Flight Center

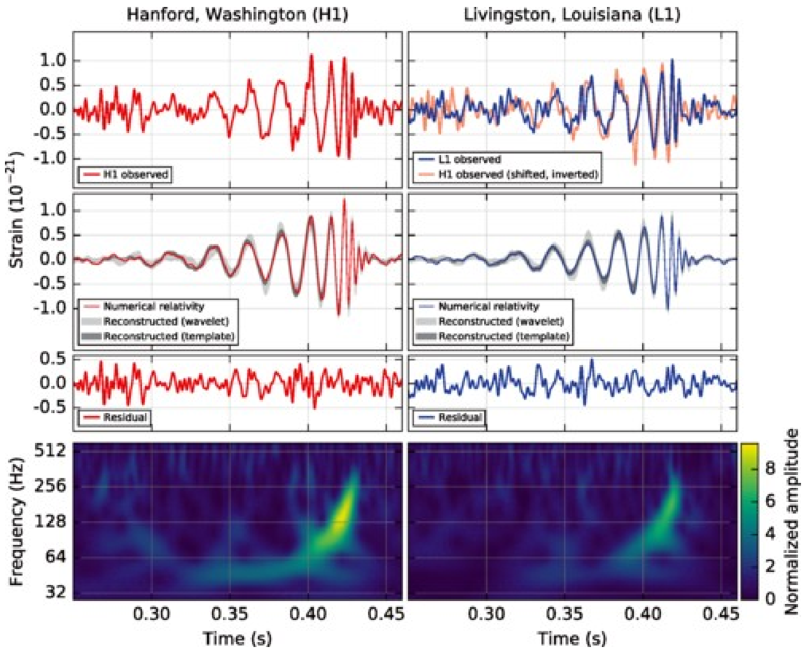
GRAVITATIONAL WAVES

GWs are a solution of the Einstein Field Equation:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}$$

Perturbation of the flat background metric:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad ; \quad |h_{\mu\nu}| \ll 1$$



Abbott et al. 2016

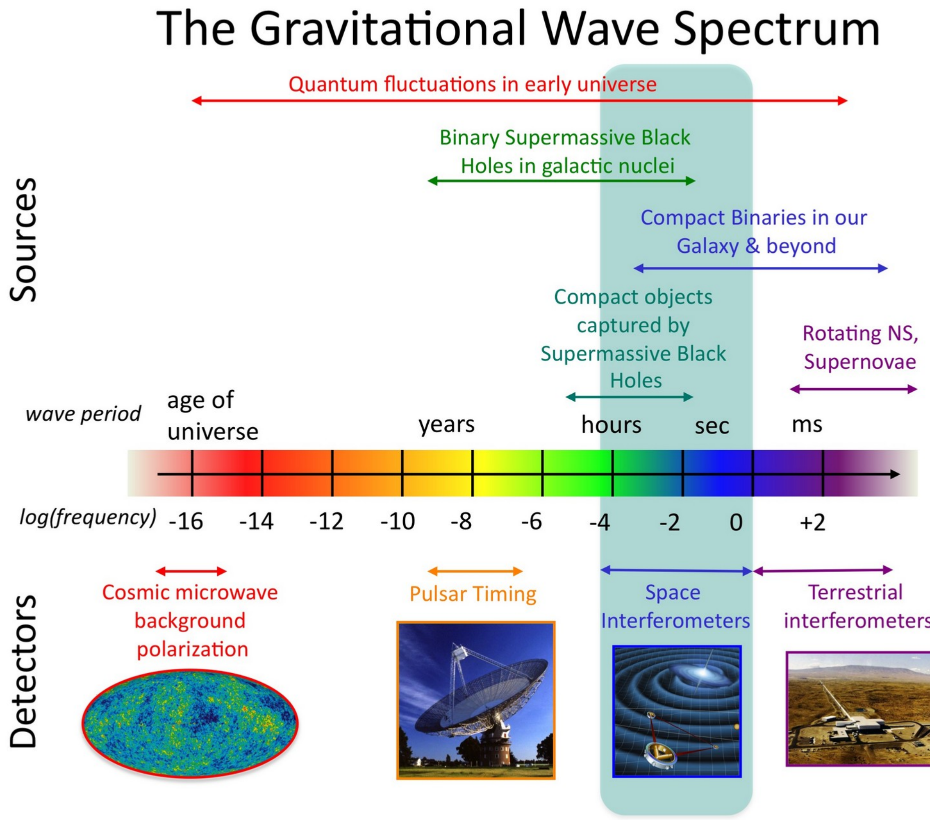
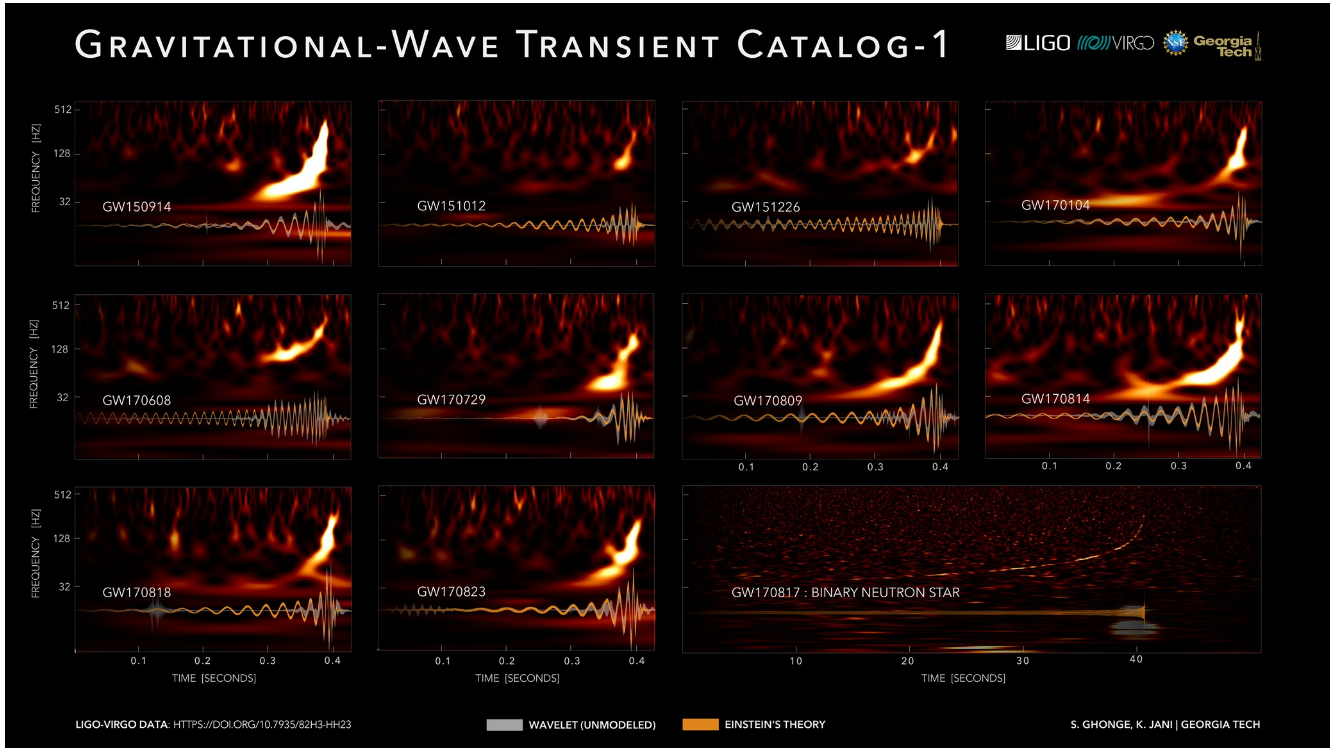


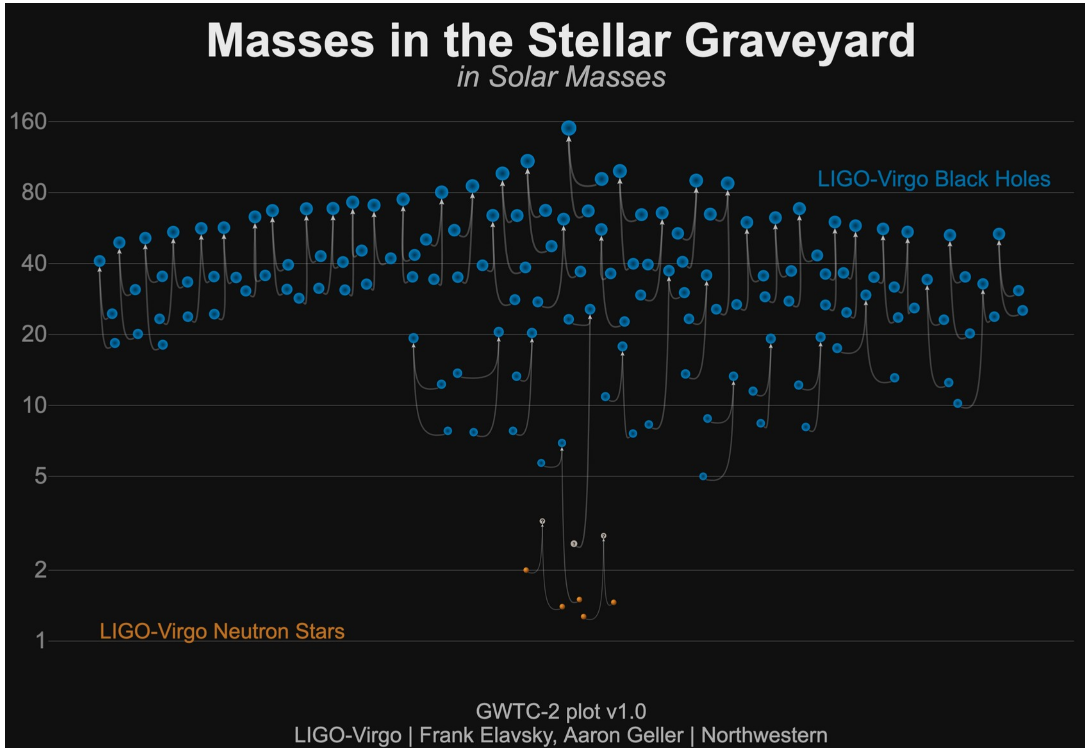
Image credit: NASA Goddard Space Flight Center

- Inspiraling binaries (BBH, BH-NS, BNS)
- Stochastic background
- Rotating asymmetric neutron stars
- Core-collapse supernovae

GRAVITATIONAL WAVES

- O1-O2 events in GWTC-1 (Abbott et al 2019).
- O3a events in GWTC-2 (Abbott et al 2021).





Public DCC images, LIGO-Virgo Collaboration

GRAVITATIONAL WAVES

LIGO Livingston

LIGO Hanford

Virgo

KAGRA

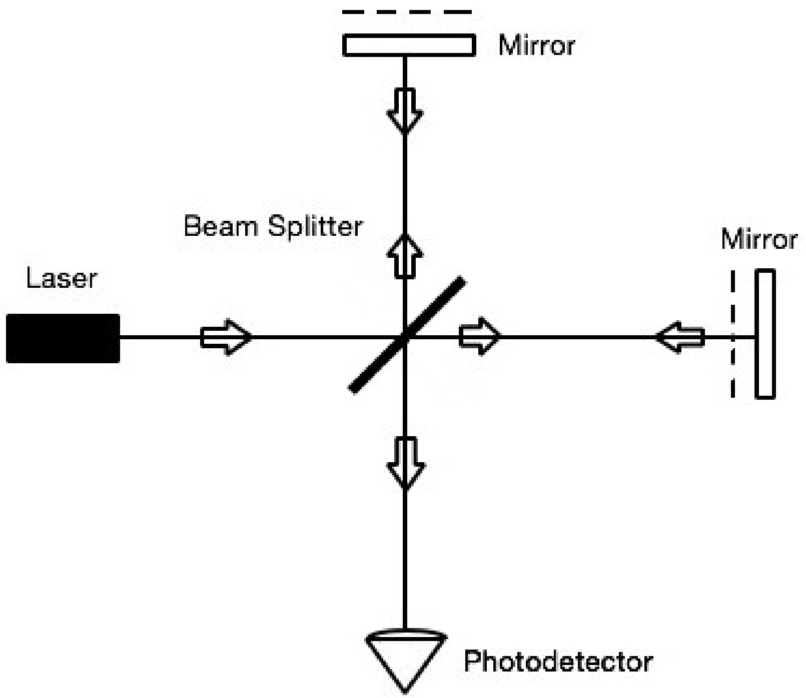
Current and future ground and space based GW detectors

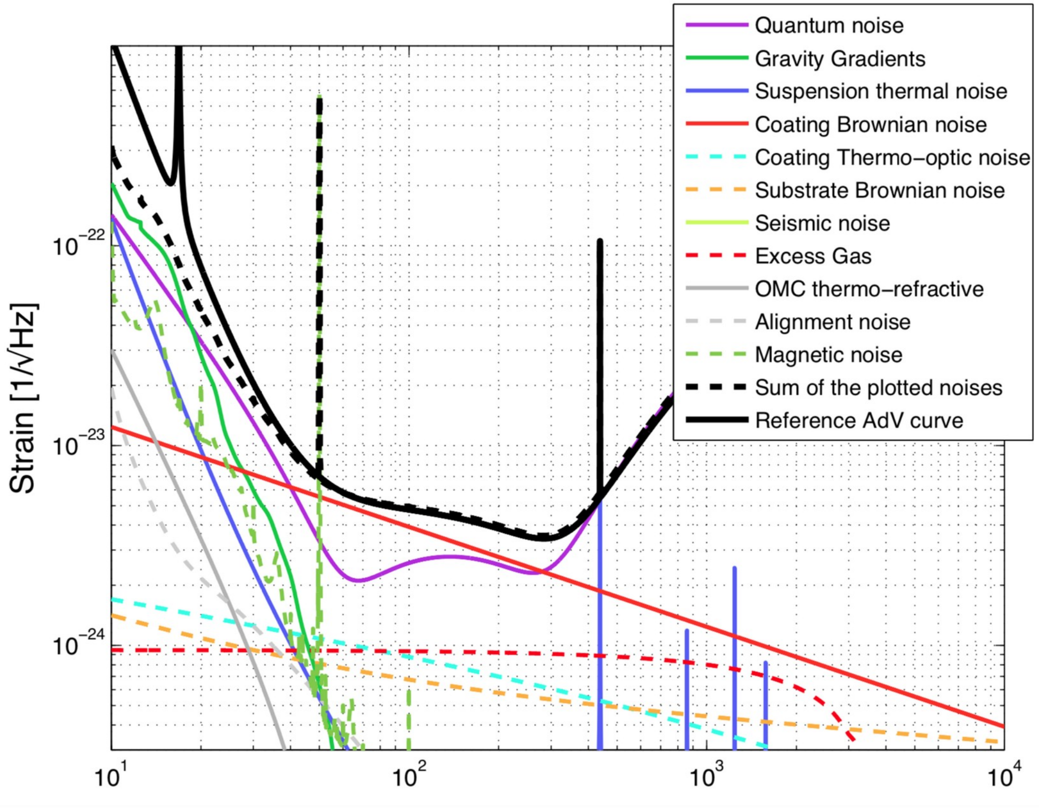
LISA

Cosmic Explorer

GRAVITATIONAL WAVE INTERFEROMETERS

- Ground-based interferometers are complex instruments, with strict sensitivity requirements.
- Many noise sources can affect measurements.

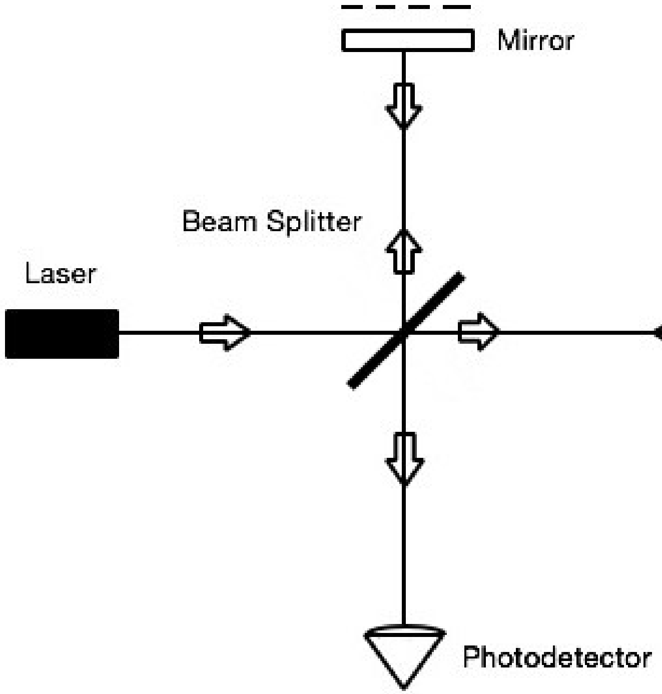




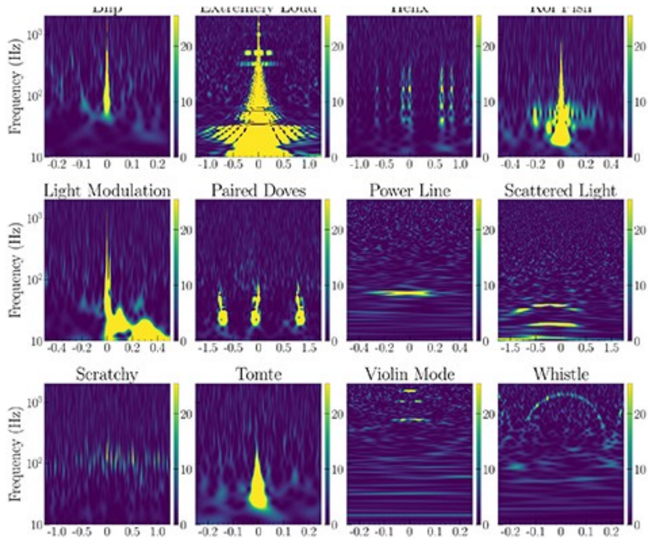
Acernese et al. (2015)

GRAVITATIONAL WAVE INTERFEROMETERS

- Ground-based interferometers are complex instruments, with <u>strict sensitivity requirements</u>.
- <u>Many noise sources</u> can affect measurements.

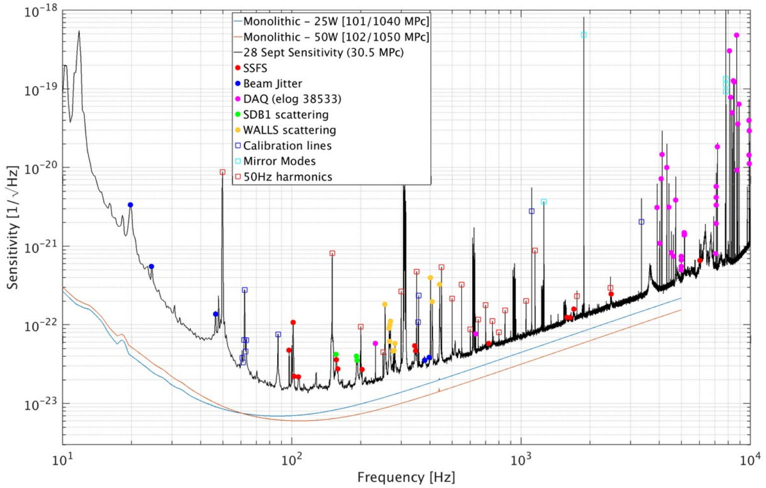


TRANSIENT NOISE BURSTS(GLITCHES)



Bahaadini et al. (2018)

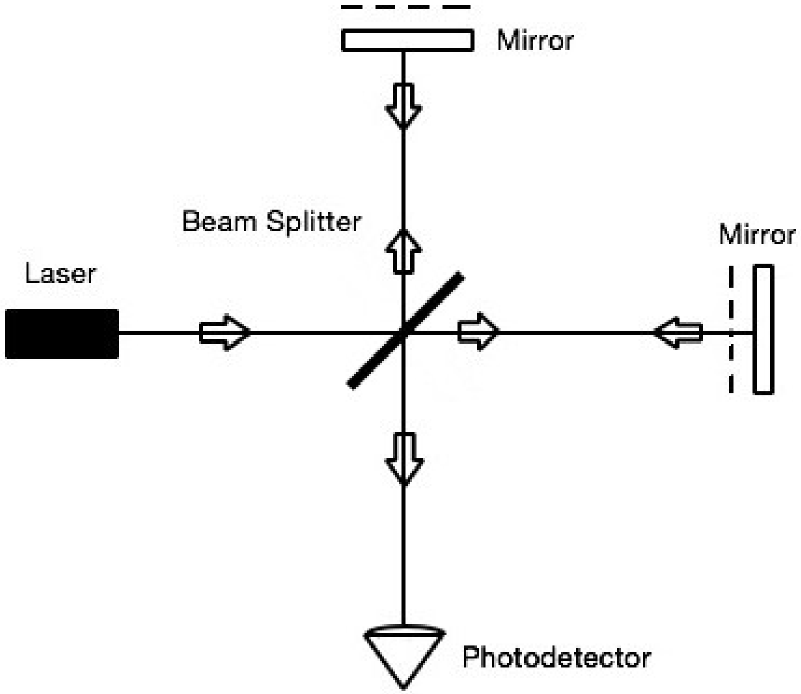
DISTURBANCES, NOISE LINES



Virgo logbook entry elog40306

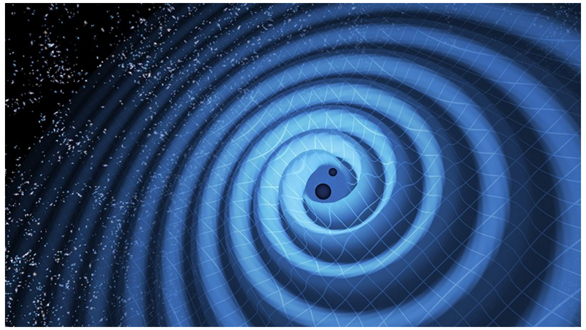
GRAVITATIONAL WAVE INTERFEROMETERS

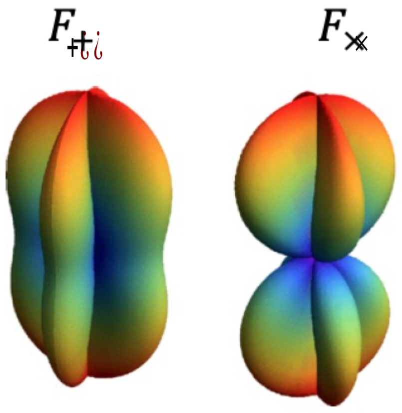
- Ground-based interferometers are complex instruments, with strict sensitivity requirements.
- Many noise sources can affect measurements.



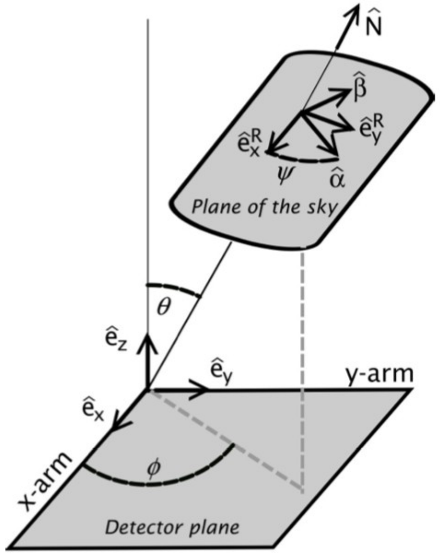
INTERFEROMETER ANTENNA PATTERN FUNCTIONS

$$h(t) = F_{+}(\alpha,\delta,\lambda,\beta,\chi,\eta)\,h_{+}(t) + F_{\times}(\alpha,\delta,\lambda,\beta,\chi,\eta)\,h_{\times}(t)$$





Adhikari (2014)



Schutz (2011)

The usual approach in GW data analysis for extraction of a signal from a signal+noise data-stream $s(t) = n(t) + h(t)$ is to implement **matched filtering** (see Allen et al 2012, Abbott et al. 2016):

$$\rho^2(t) \equiv \frac{1}{\langle h|h\rangle} |\langle s|h\rangle(t)|^2$$

Matched filter SNR time-series

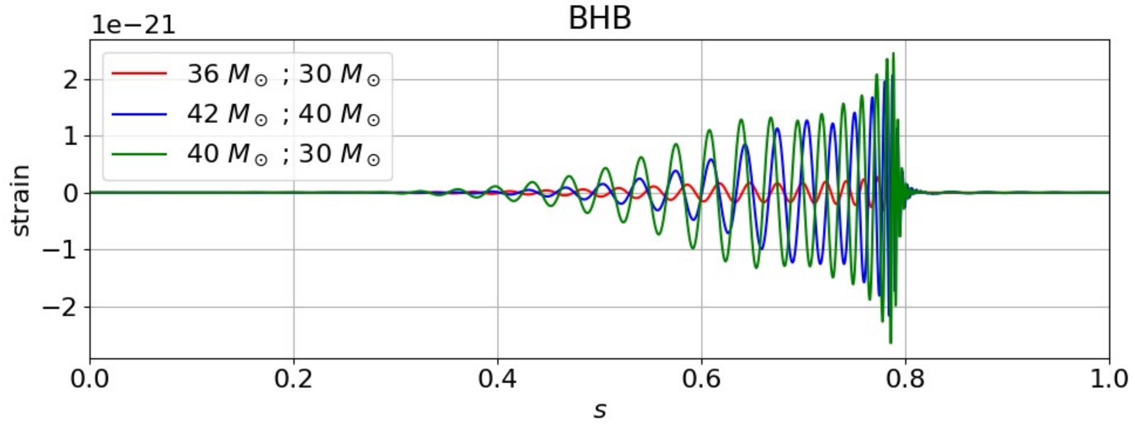
$$\langle s|h\rangle(t) = 4\int_0^{\infty} \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi i f t}\, \mathrm{d}f$$

Matched filter correlation

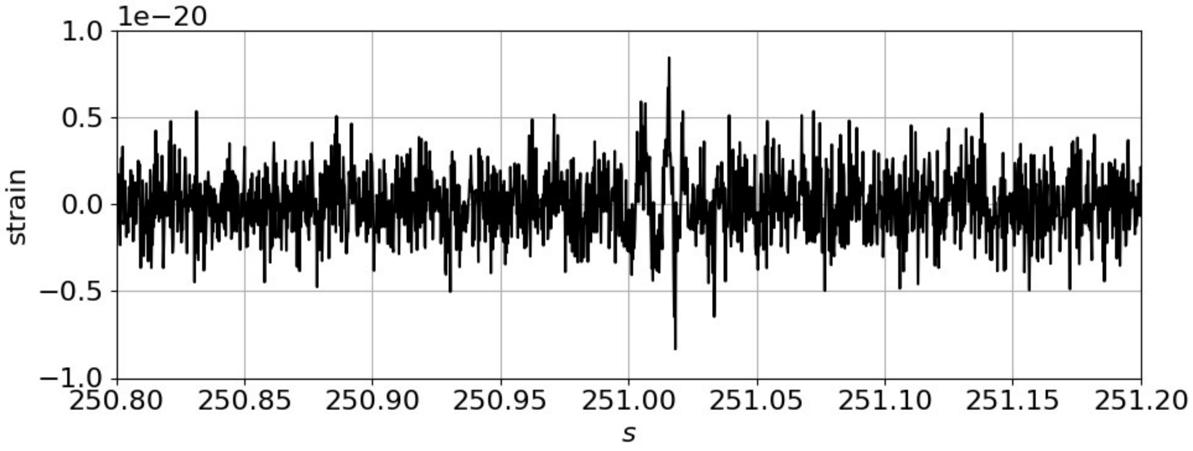
$$\langle h|h\rangle = 4\int_0^{\infty} \frac{\tilde{h}(f)\tilde{h}^*(f)}{S_n(f)}\, \mathrm{d}f$$

Template Normalization

- Large template bank to cover the parameter space ($\mathcal{M}_c$, ϕ_c , t_c , ι , D , θ , φ).
- Requires perfect modeling of phase evolution of the signal.
- Optimal for stationary gaussian background.
- Computationally expensive.

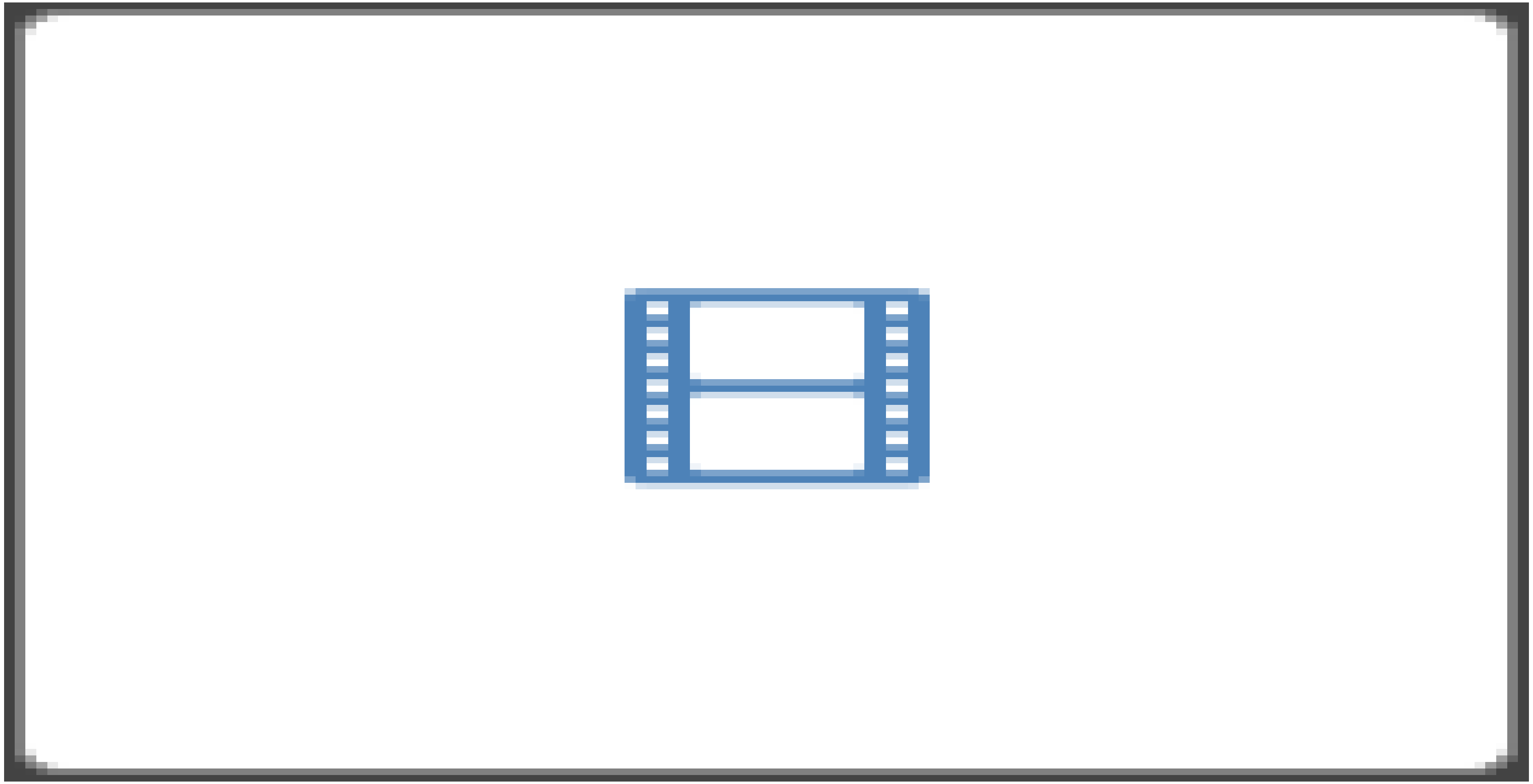


Waveform template



Whitened interferometer data

GRAVITATIONAL WAVES



Why Machine Learning In Gravitational Wave Astronomy?

The science case for the use of ML in GW astronomy is well justified:

- Increased interferometer sensitivity → more events to be processed
- Large datastream
- Low latency analysis for multimessenger → faster sky localization needed
- Joint multimessenger analyses
- Interferometer monitoring, control and noise subtraction

⇩

RECENT PROLIFERATION OF ML RELATED PROJECTS
(see Cuoco et al. 2021, *Enhancing gravitational-wave science with machine learning*)

GW170817 → Multi-messenger astrophysics (Abbott et al. 2017 and refs. therein)

- Coincident short GRBs detected in gamma rays
- Host galaxy identification (NGC 4993)
- Optical/infrared/UV counterpart (AT2017gfo) has been detected
- First spectroscopic identification of a kilonova
- X-ray and a radio counterparts have been identified

GRB
Fermi GBM, INTEGRAL, Astrosat, IPN, Insight-HXMT, Swift, AGILE, CALET, H.E.S.S., HAWC, Konus-Wind

X-RAY
Swift, MAXI/GSC, NuSTAR, Chandra, Integral

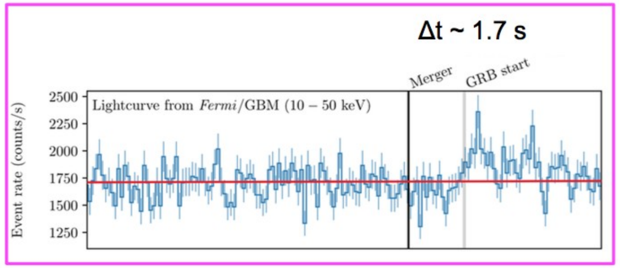
UV
Swift, HST

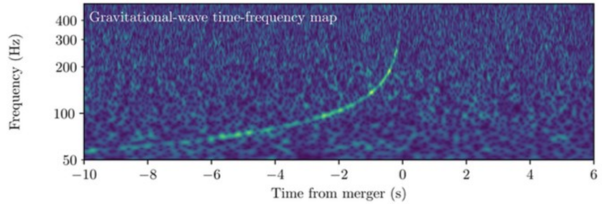
RADIO
ATCA, VLA, ASKAP, VLBA, GMRT, MWA, LOFAR, LWA, ALMA, OVRO, EVN, e-MERLIN, MeerKAT, Parkes, SRT, Effelsberg

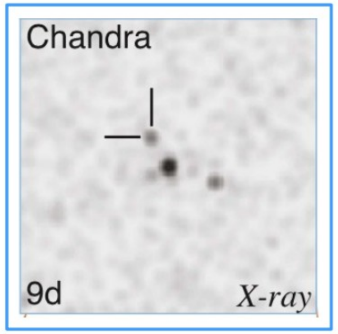
IR
REM-ROS2, VISTA, Gemini-South, 2MASS, SPITZER, NTT, GROND, SOAR, NOT, ESO-VLT, Kanata Telescope, HST

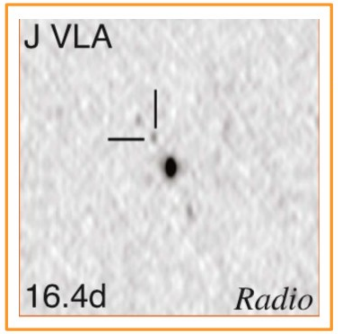
OPTICAL
Swope, DECam, DLT40, MASTER, VISTA, ESO-VLT + others

And possibly **NEUTRINOS** (IceCube, ANTARES, Pierre Auger Observatory)

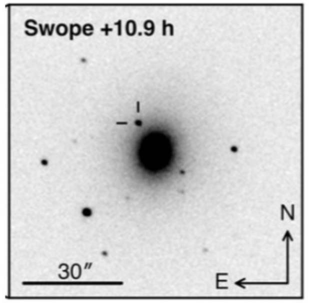


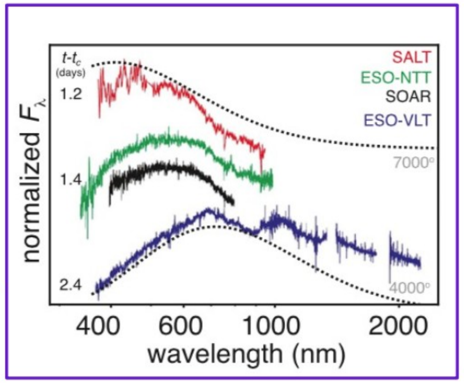


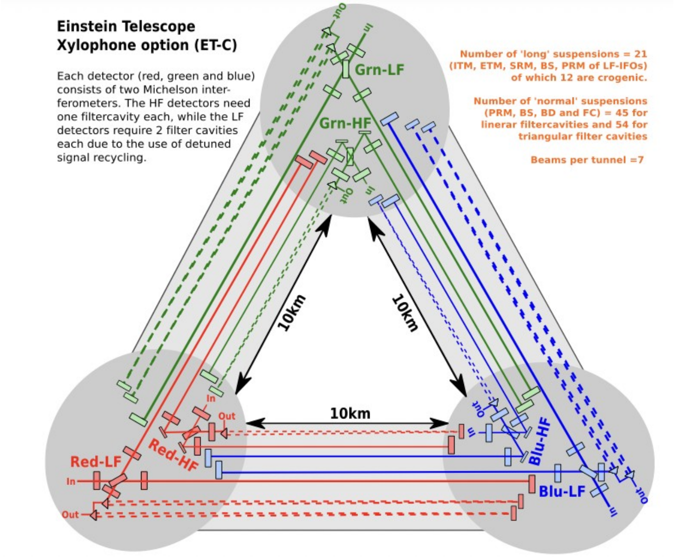




NGC 4993



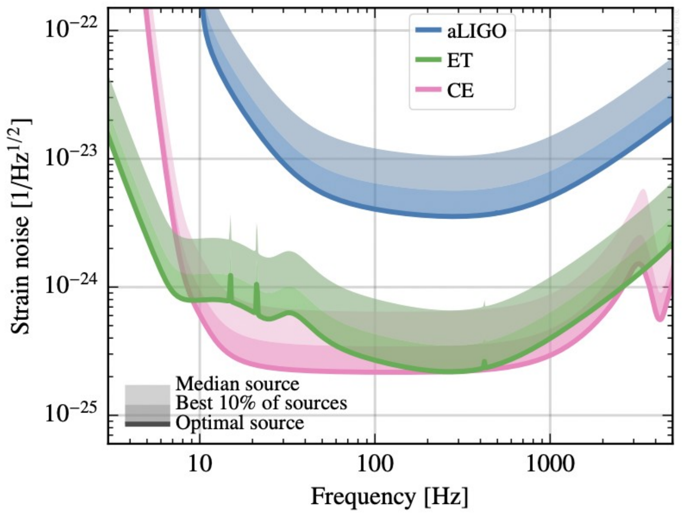


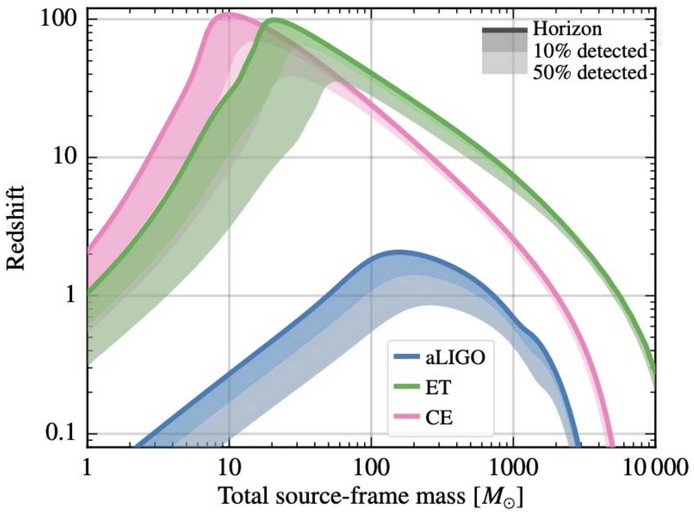


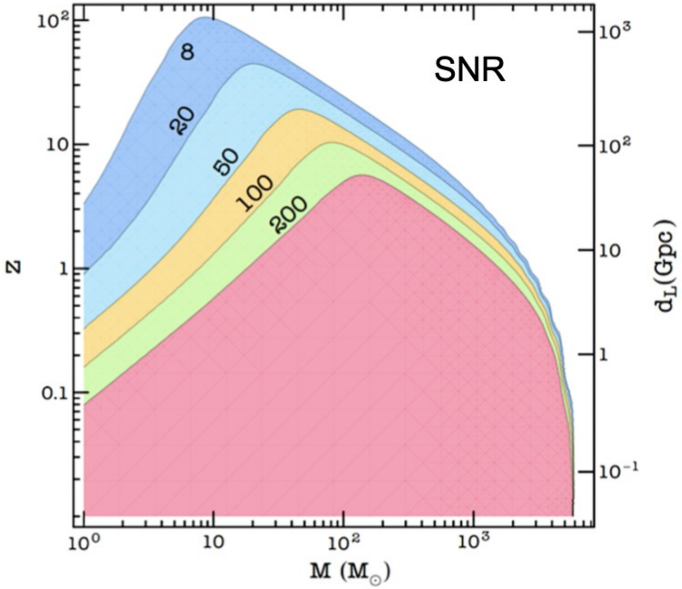
Einstein Telescope (ET) is a 3rd generation ground based interferometer planned for the early 2030s.

- 10^5 BNS detections per year
- 10^5 BBH detections per year
- Order of magnitude gain on sensitivity
- Access lower frequencies (few Hz)

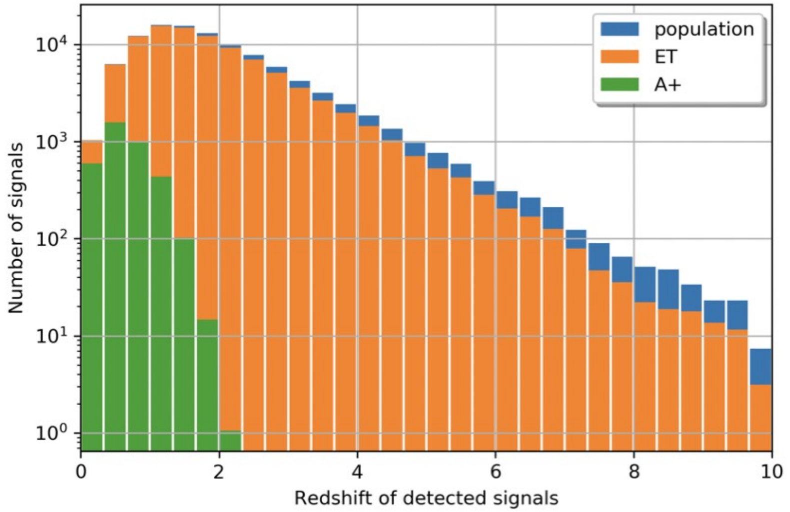
ET Observational Science Board Kick-off meeting



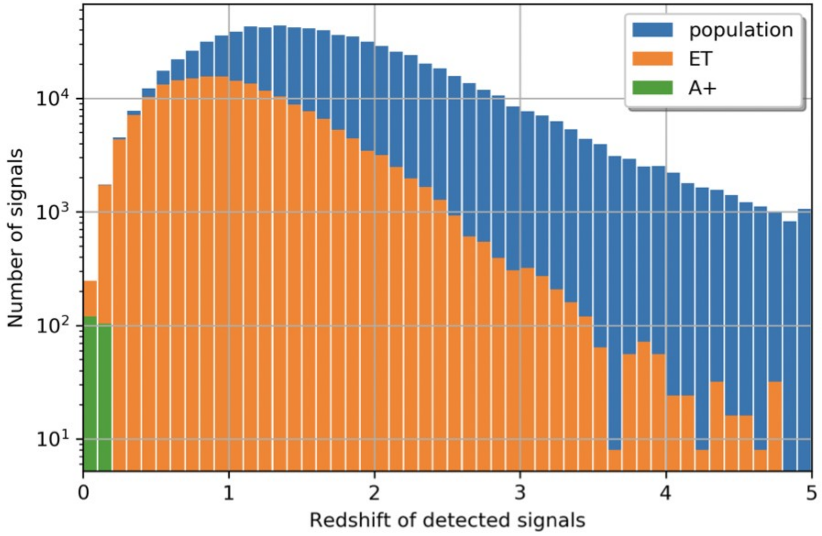




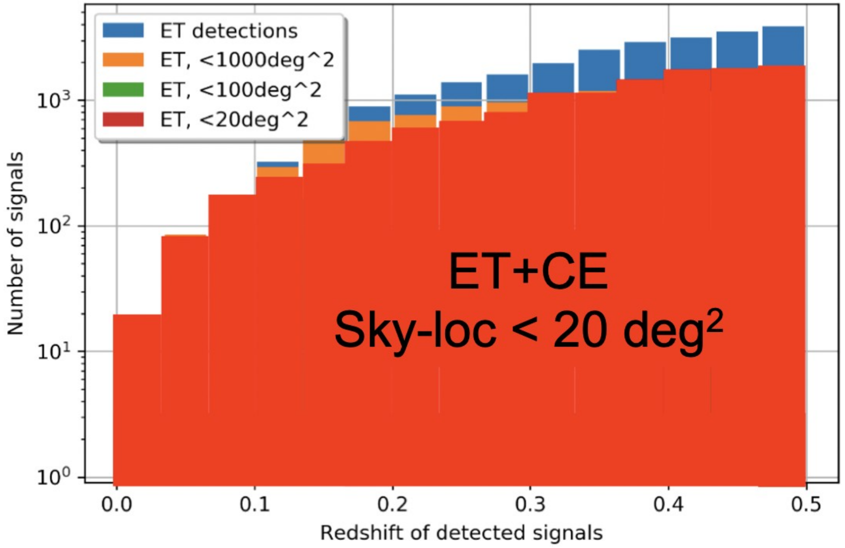
BINARY BLACK-HOLE MERGERS



BINARY NEUTRON-STAR MERGERS



We will reach higher redshifts
And improve sky localization!



MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

GLITCH CLASSIFICATION	NOISE SUBTRACTION	PARAMETER ESTIMATION	SKY LOCALIZATION
WAVEFORM MODELING	GRAVITATIONAL WAVE SEARCHES	INTERFEROMETER CONTROL	And much more…

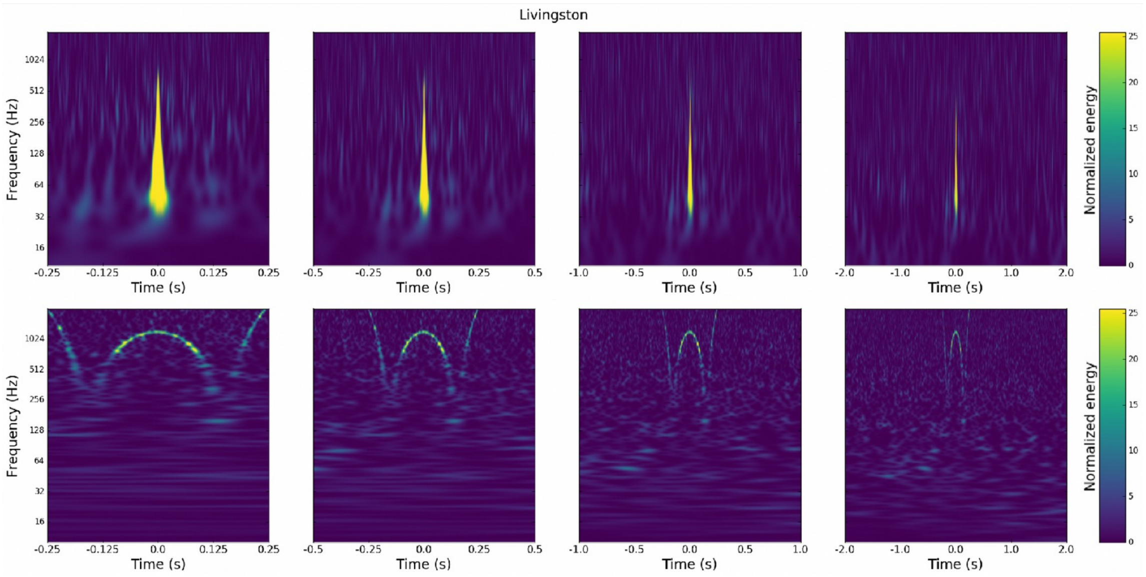
MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

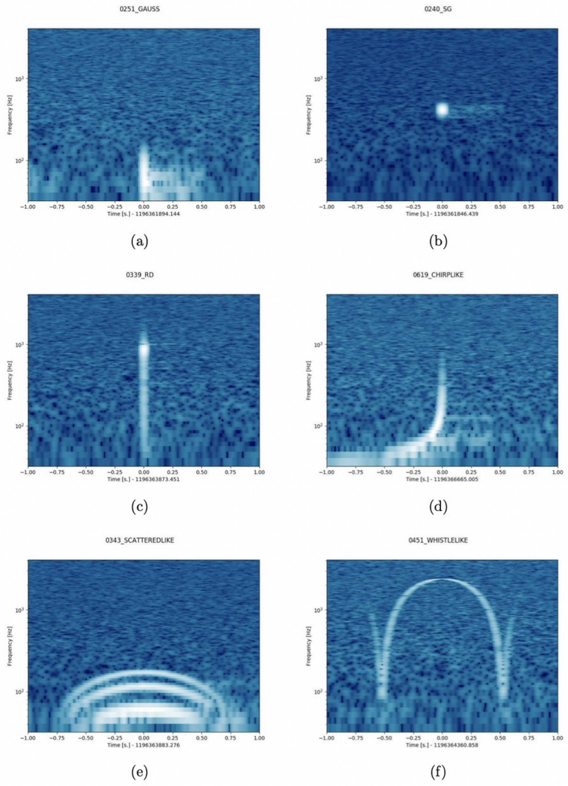
GLITCH CLASSIFICATION	NOISE SUBTRACTION	PARAMETER ESTIMATION	SKY LOCALIZATION
WAVEFORM MODELING	GRAVITATIONAL WAVE SEARCHES	INTERFEROMETER CONTROL	And much more…

GLITCH CLASSIFICATION

Glitches hamper matched filter searches increasing the FAR and triggering vetoes on pipelines. ML Can be used to:

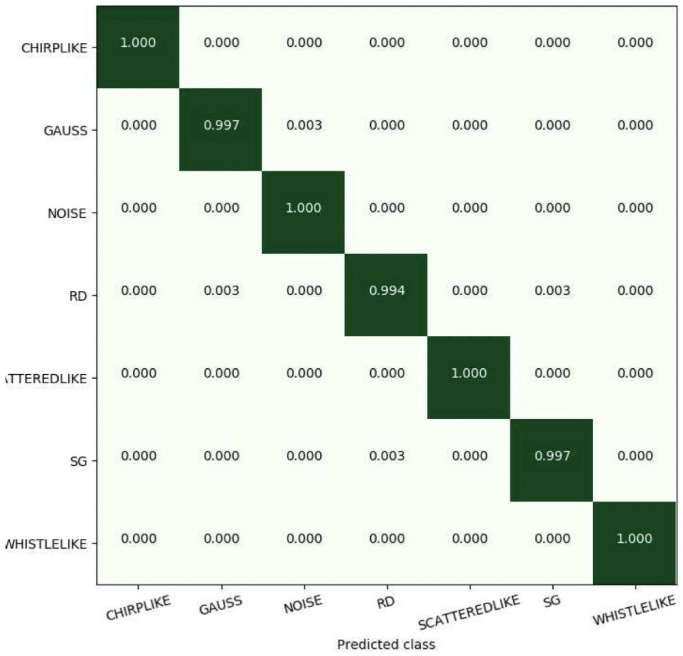
- Recognize noise burst transients
- Subtract glitches
- Classify into categories depending on origin
- Discover new classes





GRAVITY SPY
(citizen science project)

Zevin et al. (2017)



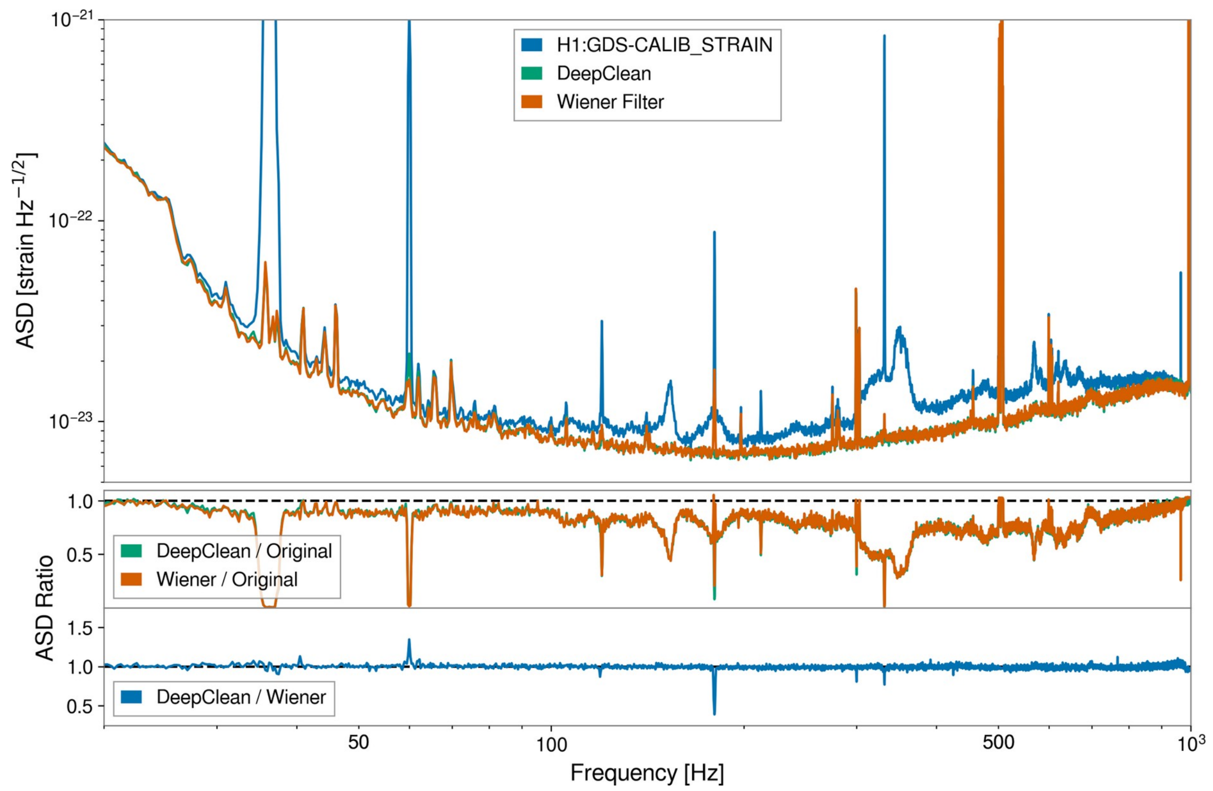
Cuoco, Razzano (2018)

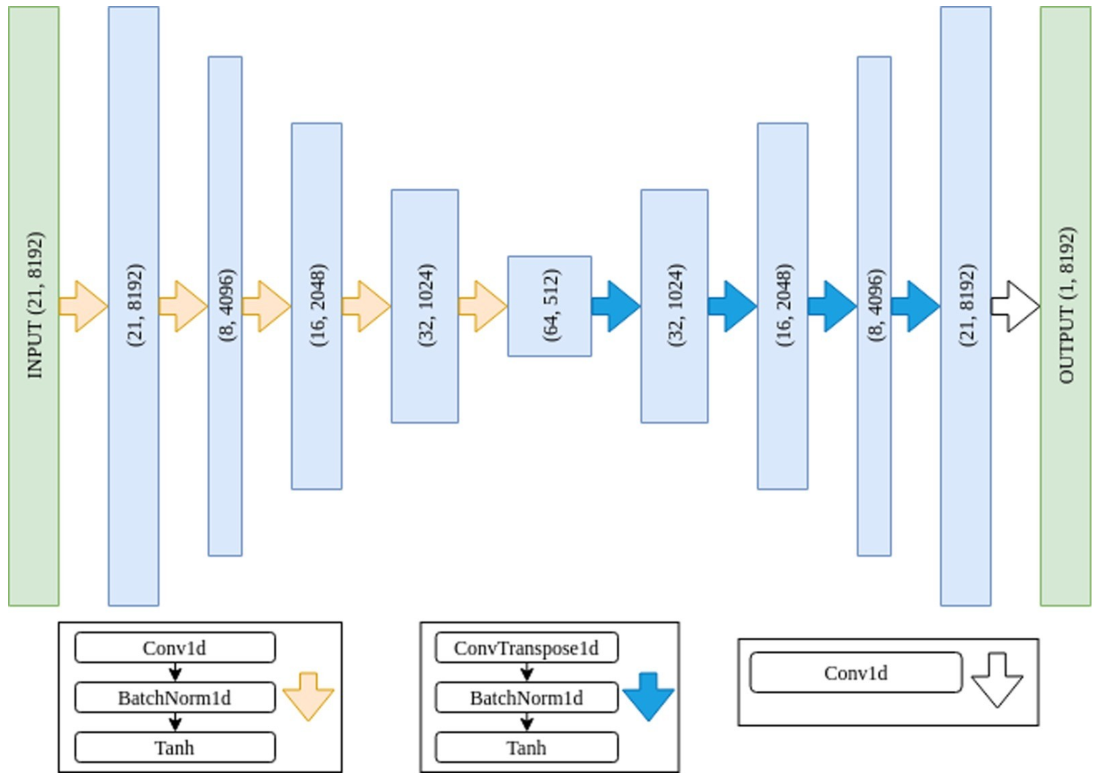
And many more…

NOISE SUBTRACTION

Use witness channels $w_i(t)$ for environmental noise subtraction.

- Similar results to Wiener filter for linear noise
- Can learn non linear noise



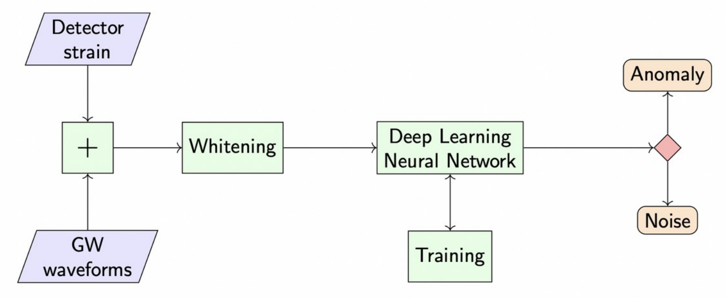


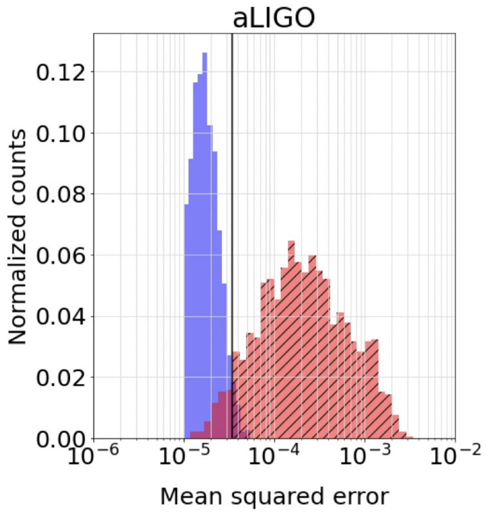
Ormiston et al. (2020)

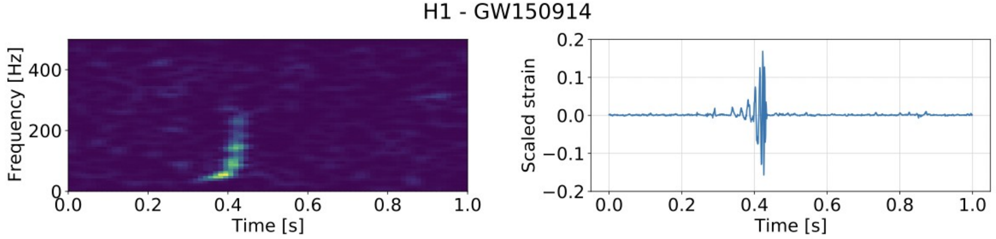
- Different sampling for channels
- 1-D time-series inputs

GRAVITATIONAL WAVE SEARCHES

A number of studies use as inputs whitened time-series to search for gravitational wave signals from compact binary mergers

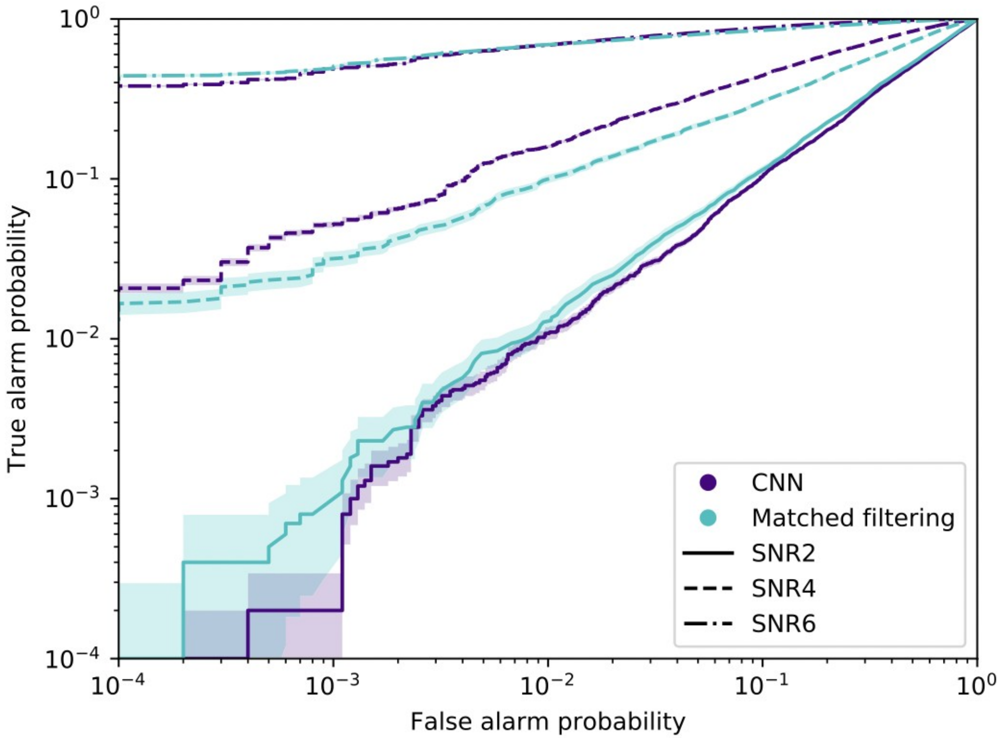






Morawski et al. (2021)

Anomaly Detection

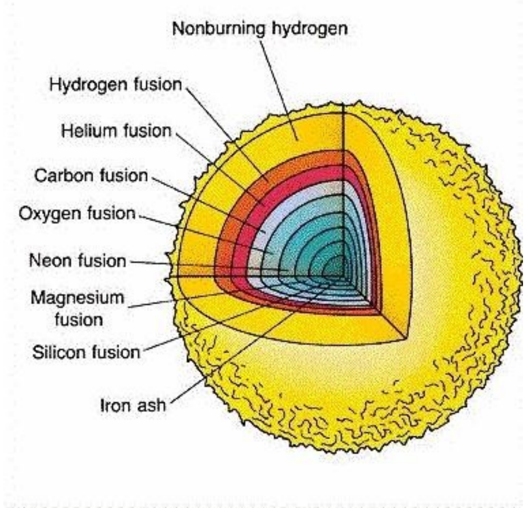


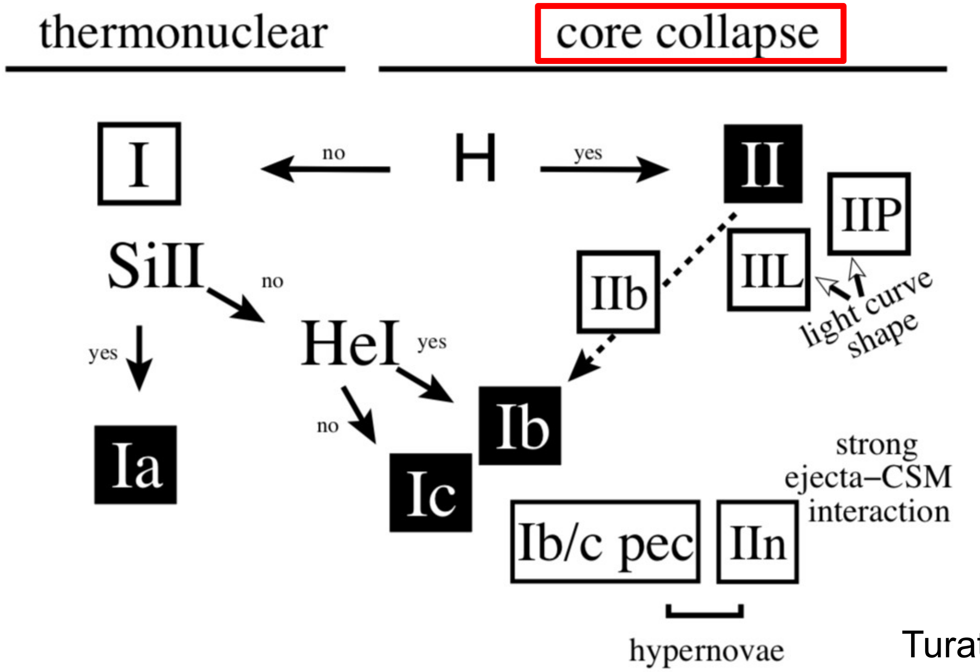
Gabbard et al. (2020)

Convolutional Neural Networks

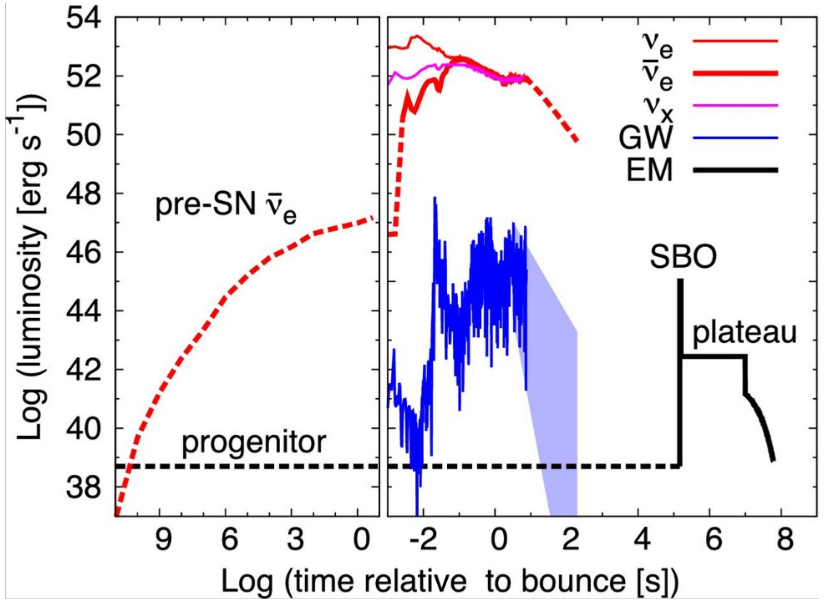
DEEP LEARNING FOR BURST SIGNALS: CORE-COLLAPSE SUPERNOVAE

- **Final evolutionary stages of massive stars**, ($\gtrsim 8-9\,M_\odot$), after nuclear burning accumulates iron core mass greater than Chandrasekhar mass limit.
- **Multi-messenger emission, 99% energy in neutrinos**.
- **Rare, optimistically** (~50-100 yrs in Milky Way).
- **Prompt neutrino emission at ~10MeV**.
- Weak prompt GW signals, expected less than 0.0001% supernova energy, h.
- Late EM emission.





Turatto (2003)

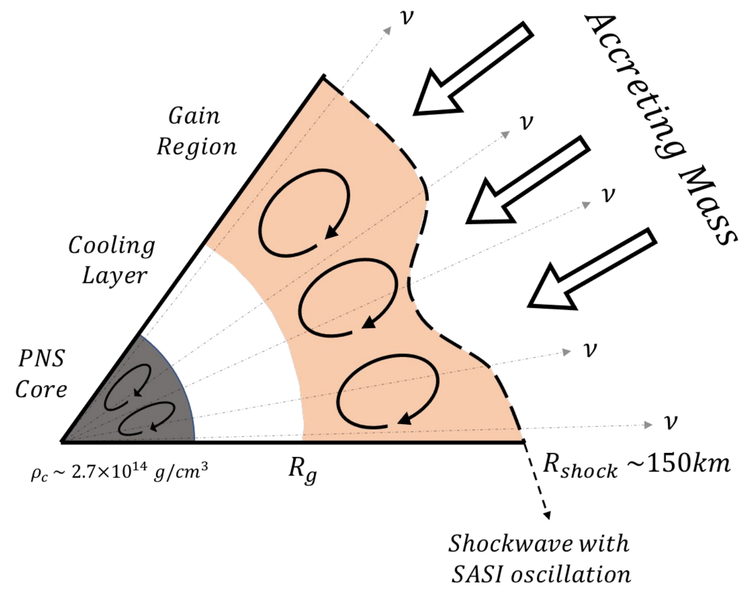


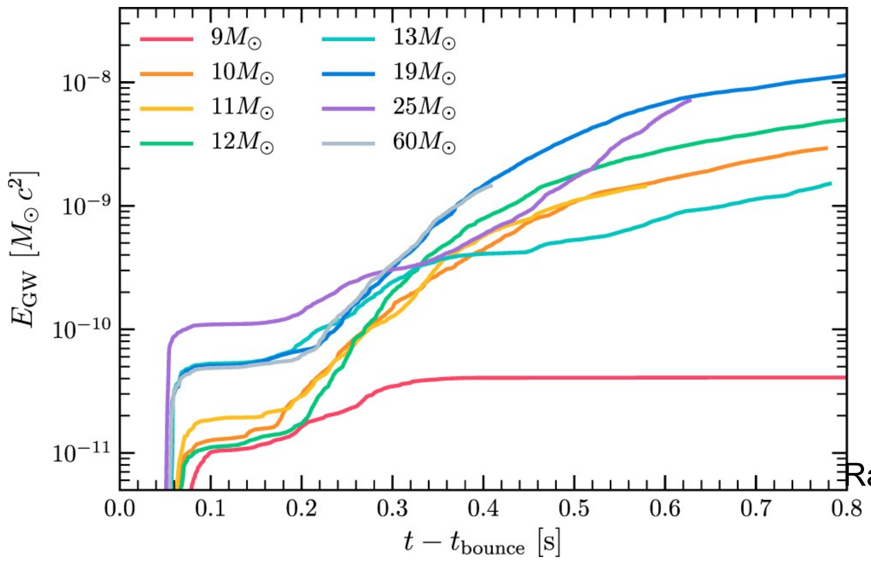
Nakamura et al. (2016)

GWs FROM CORE-COLLAPSE SUPERNOVAE

- Waveform depends on progenitor star
- Different possible emission mechanisms
- Large degree of stochasticity
- Broadband emission
- Best waveform models from computationally expensive 3D simulations

Matched filter not feasible!





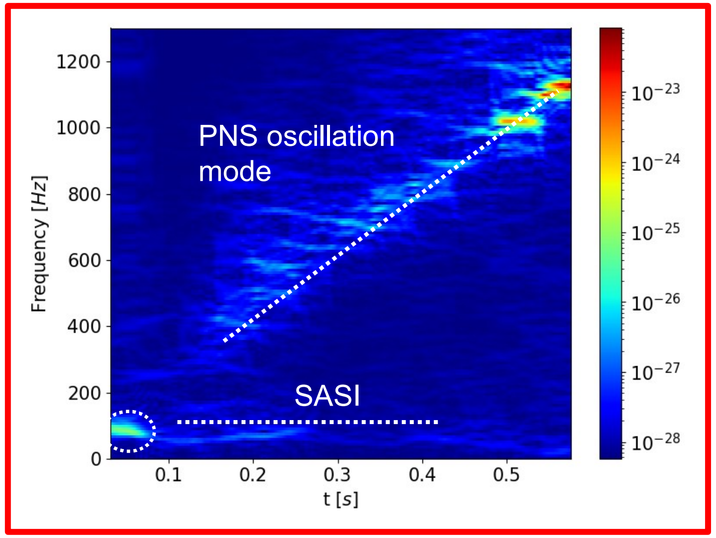
Radice et al. (2017)

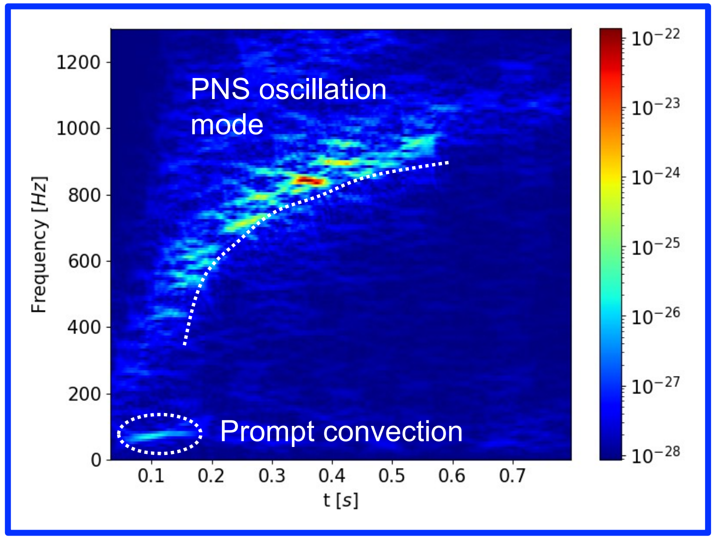
$E_{gw} \sim 10^{44}$ erg 10^{47} erg
$E_{SN} \sim 10^{50}$ erg 10^{53} erg

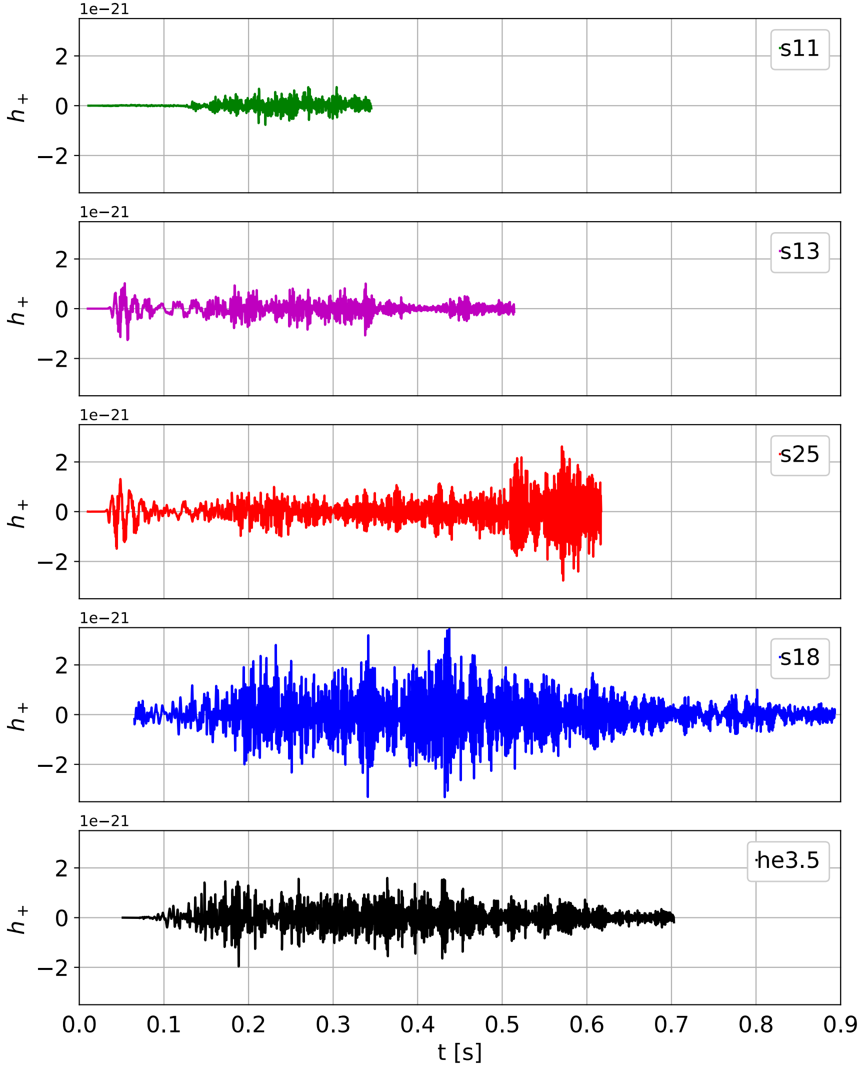
	Potential explosion mechanism		
GW emission Process	MHD mechanism (rapid rotation)	Neutrino mechanism (slow/no rotation)	Acoustic mechanism (slow/no rotation)
Rotating collapse and Bounce	Strong	None/weak	None/weak
3D rotational instabilities	Strong	None	None
Convection & SASI	None/weak	Weak	Weak
PNS g-modes	None/weak	None/weak	Strong

Ott et al. (2017)

DEEP LEARNING FOR BURST SIGNAL CLASSIFICATION



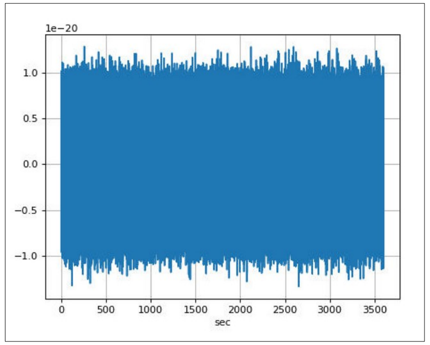




Iess, Cuoco, Morawski, Powell (2020)

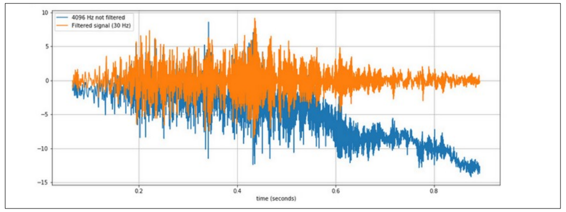
DEEP LEARNING FOR BURST SIGNAL CLASSIFICATION

NOISE BACKGROUND
(*simulated gaussian or real*)

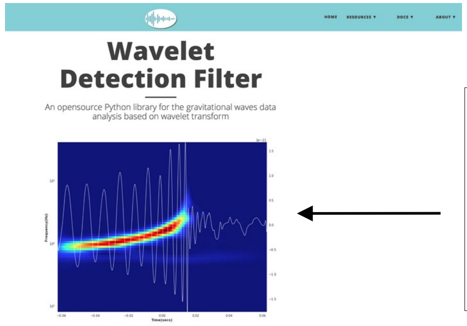


WAVEFORMS

RESAMPLING, FILTERING AND ANTENNA PATTERN



+



WHITENING & TRIGGER GENERATION
(Wavelet Detection Filter https://wdfpipe.gitlab.io)

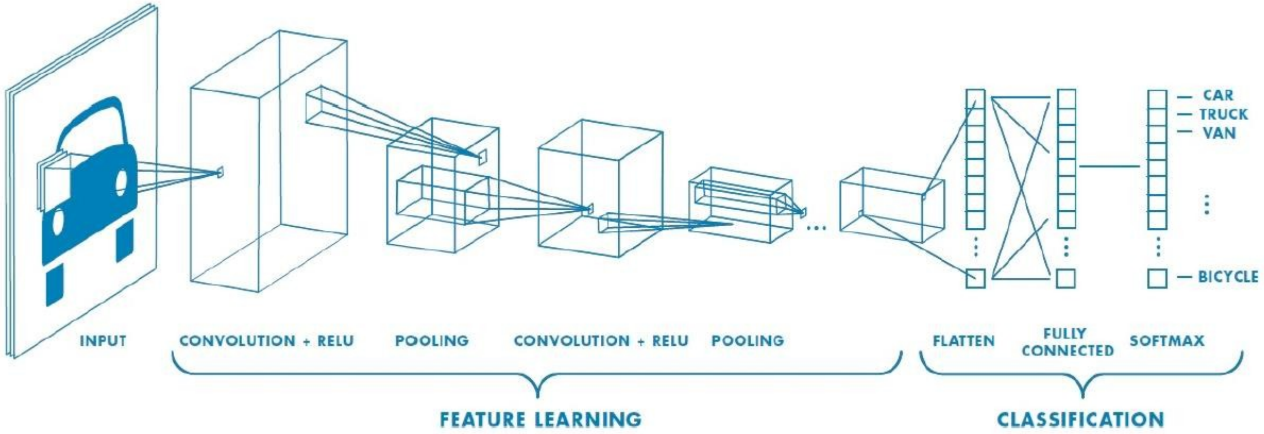
TRAINING

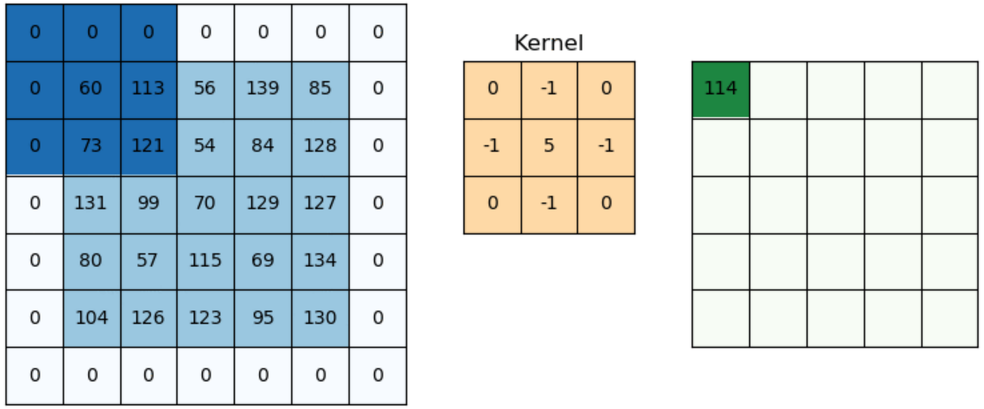
MACHINE-LEARNING CLASSIFIER

SIGNAL TYPE

GLITCH TYPE

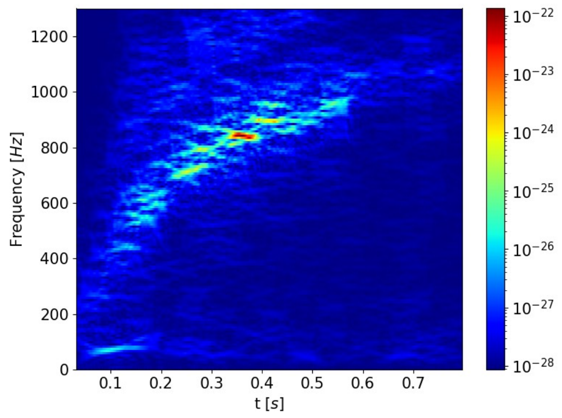
CONVOLUTIONAL NEURAL NETWORKS





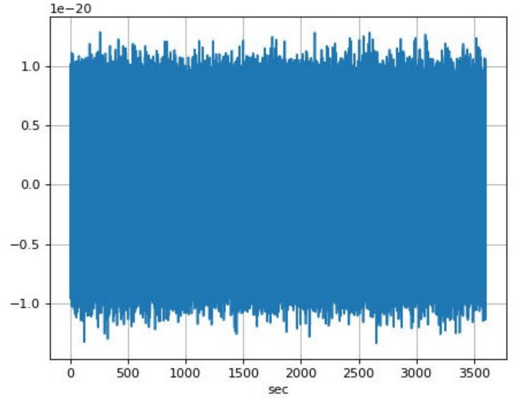
2-D CNN

Spectrogram images



1-D CNN

Whitened time series

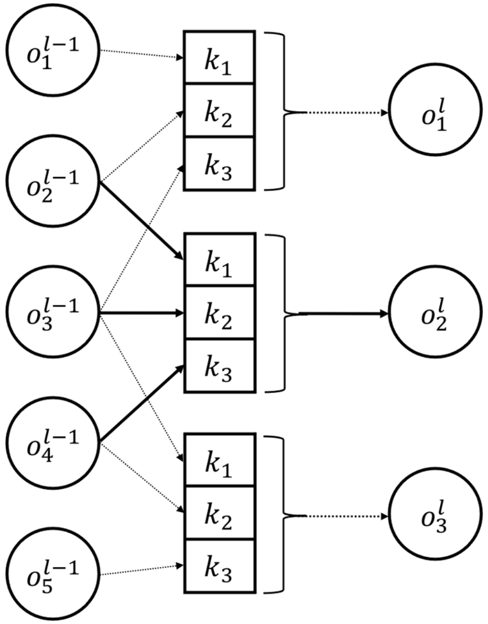


CONVOLUTIONAL NEURAL NETWORKS

CNN (Hubel and Wiesel 1962, LeCun 1998, Fukushima 1980**)**

- Recognizes patterns in data by building feature maps.
- Easy to implement, fast to train.
- Translation-invariant.

<u>NN with a strong prior on internal weights</u>: for each hidden unit all weights are zero but those that describe the kernel, shared among the different neurons.

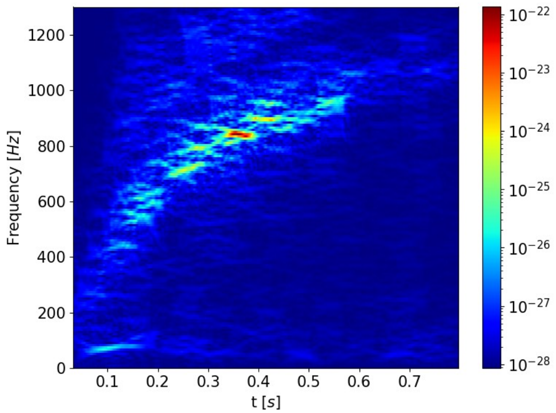


CNN EQUATION

$$o_j^l = f\left(\sum_{m=0}^{L-1} o_{j+m}^{l-1} K_m\right)$$

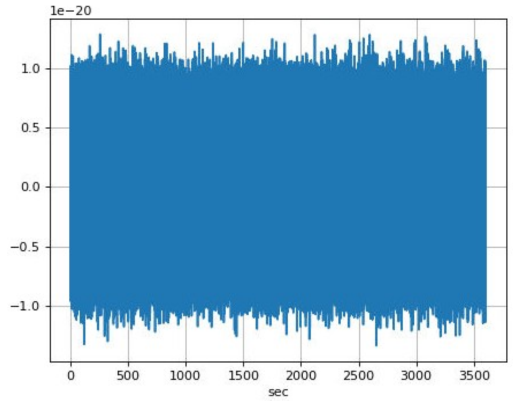
2-D CNN

Spectrogram images

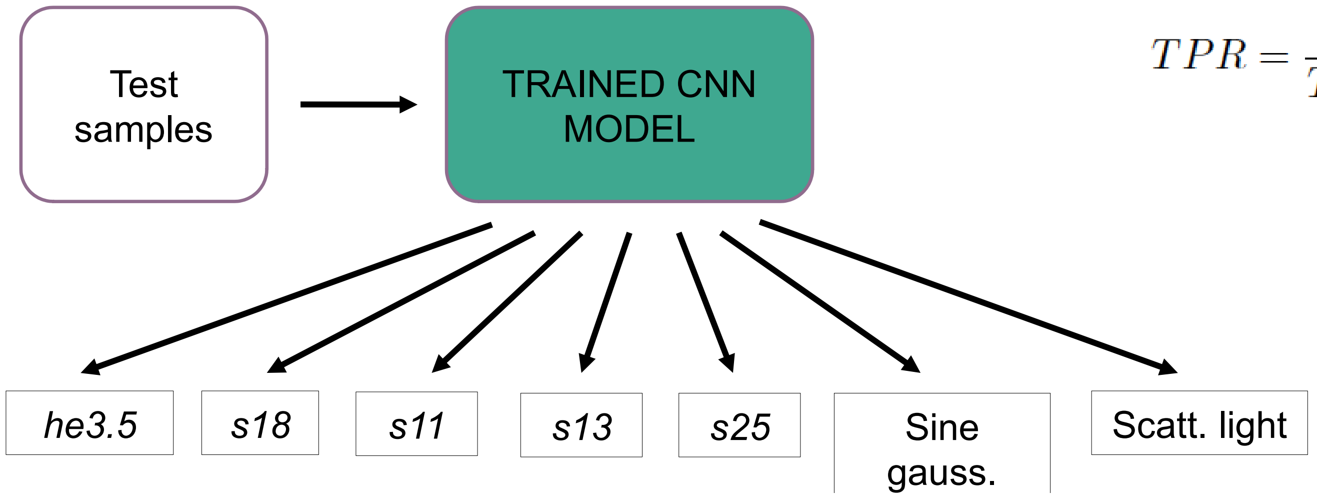


1-D CNN

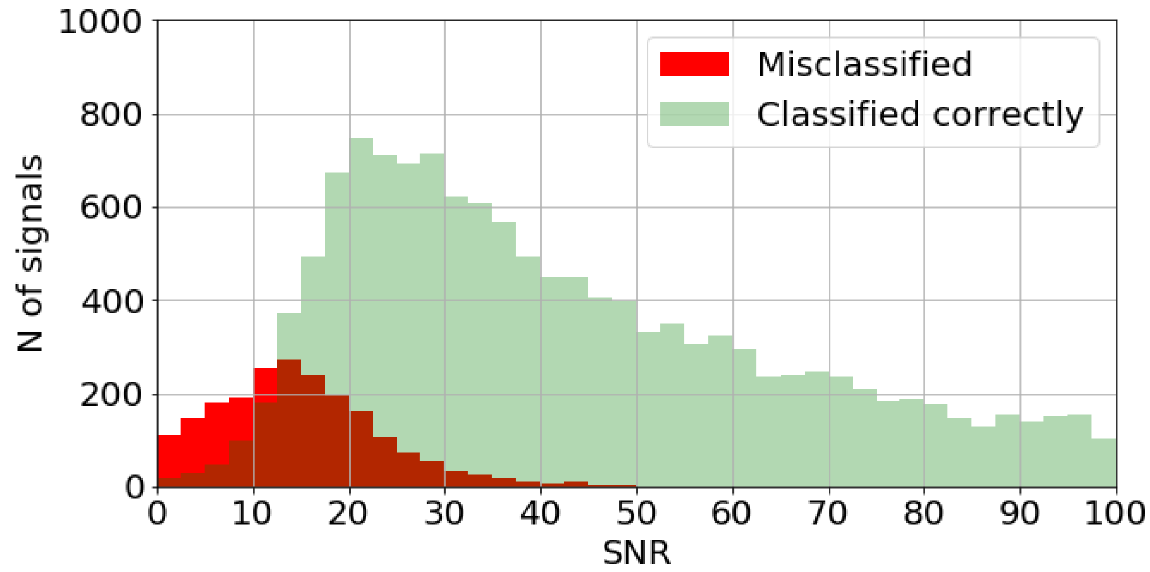
Whitened time series



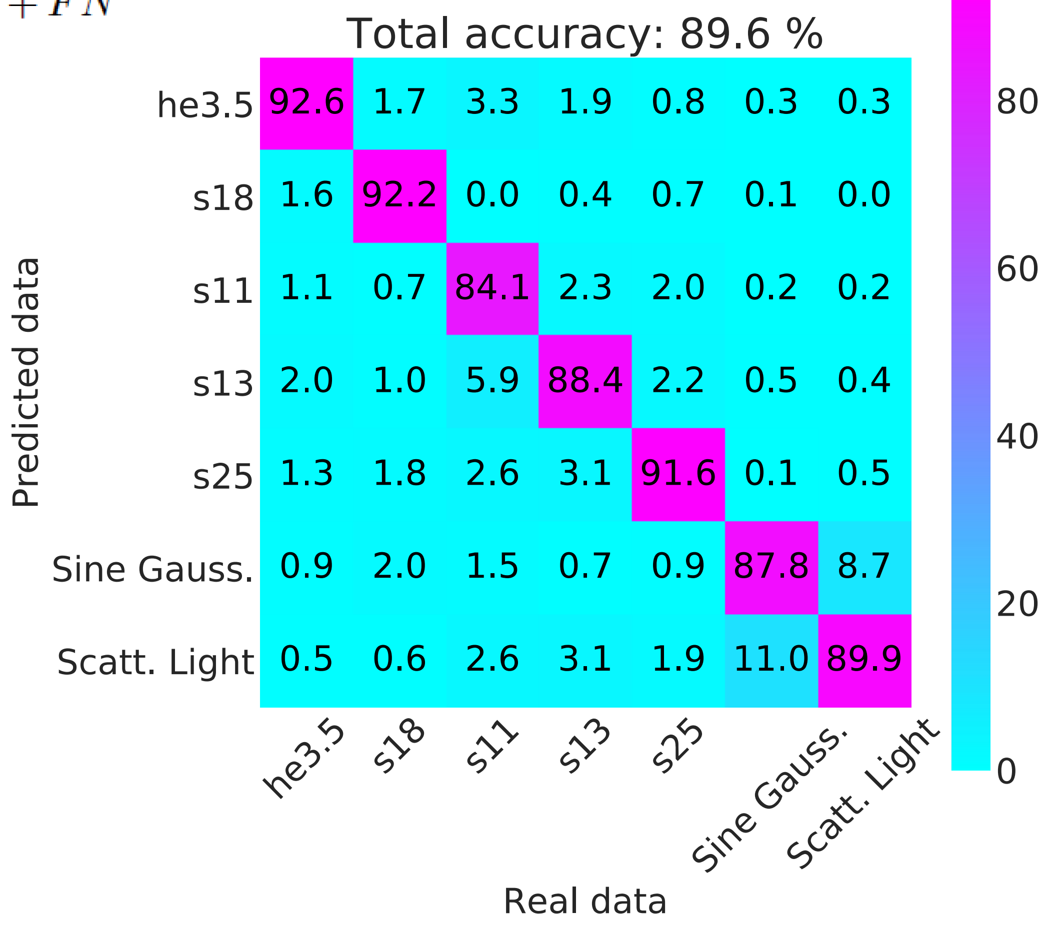
DEEP LEARNING FOR BURST SIGNAL CLASSIFICATION



$$TPR = \frac{TP}{TP + FN}$$



ET, MERGED 1D & 2D CNN



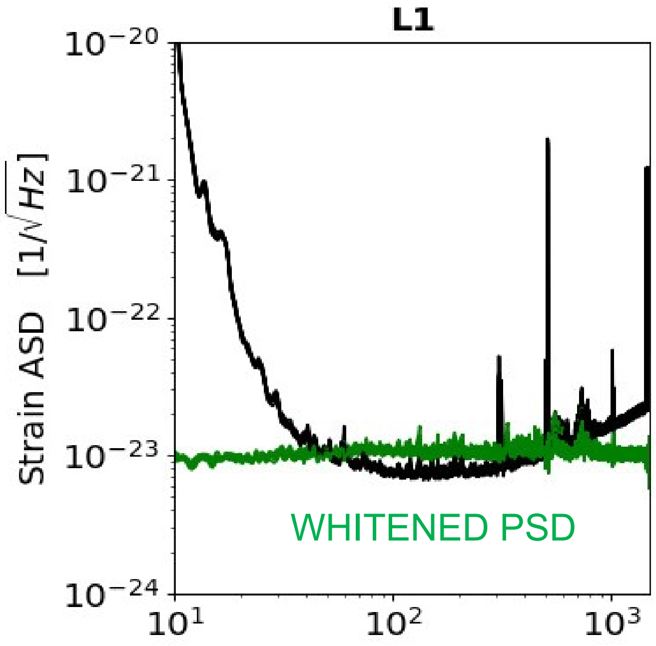
Iess, Cuoco, Morawski, Powell (2020)

What happens with real interferometer noise?

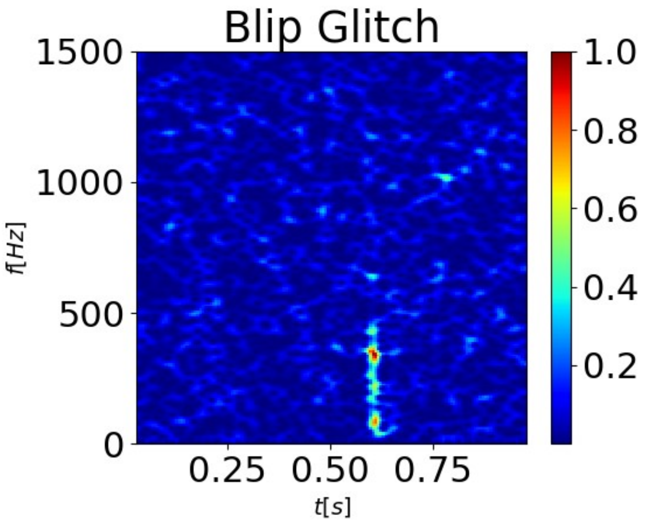
DATASET CHARACTERISTICS

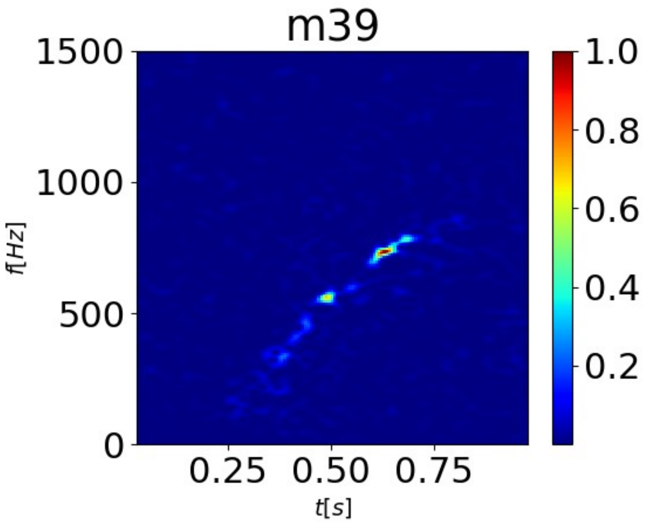
- Detector noise PSD is non stationary.
- Multiple Glitch Families.
- CCSN Dataset (at 1 kpc): ~15000 samples.
- Imbalanced Dataset due to different model amplitudes.

	Triggers		
Detector	*Signal*	*Noise*	*Total*
Virgo V1	9273	47901	57174
Ligo L1	10480	3810	14290
Ligo H1	10984	4103	15087
L1, H1, V1	5647	675	6322



… AFTER WDF WHITENING





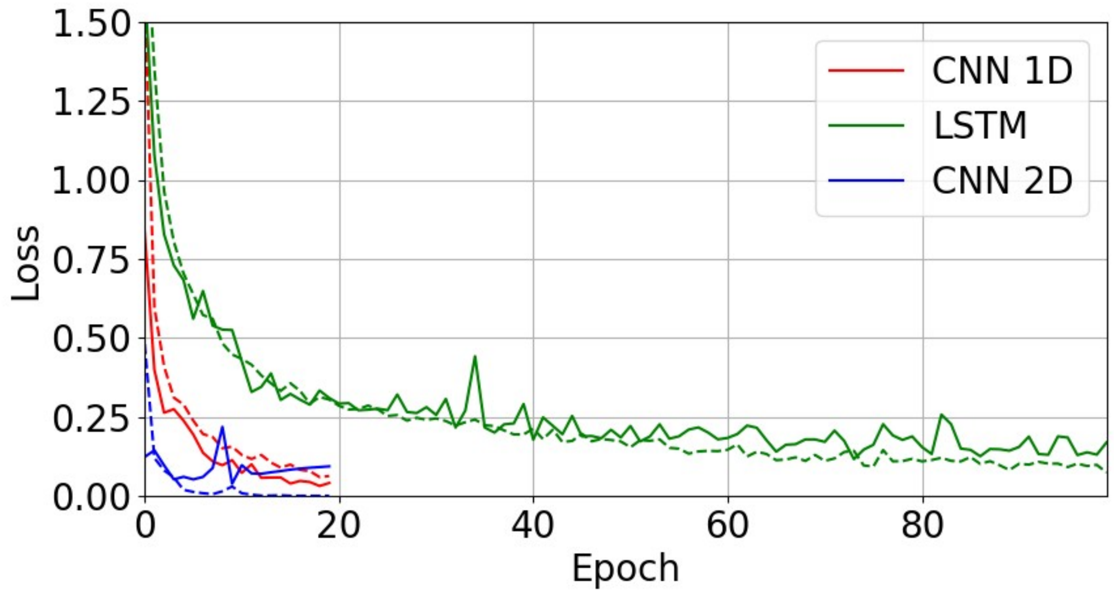
COMPARISON OF ML MODEL ACCURACY

(Single interferometer, V1,H1,L1)

- **<u>Bi-LSTM</u>**, 2 recurrent layers
- ~10 ms/sample
- Best weights over 100 epochs

- **<u>1D-CNN</u>**, 4 convolutional layers
- ~2 ms/sample
- Best weights over 20 epochs

- **<u>2D-CNN</u>**, 4 convolutional layers
- ~3 ms/sample
- Best weights over 20 epochs



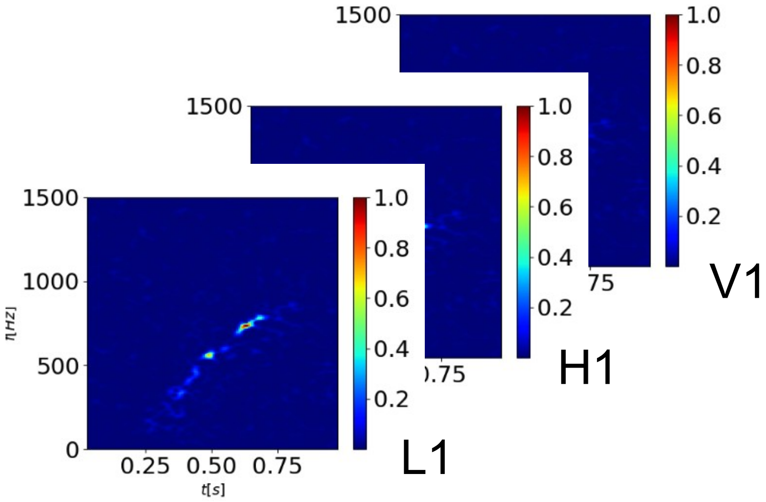
ITF	ML Model	Waveform Model: *Noise*	*s11*	*s18p*	*s18np*	*He3.5*	*m39*	*y20*	*s13*	*s25*	**Total**
V1	LSTM	49.2	*	3.6	58.4	0.0	89.5	69.9	0.0	89.8	73.7
	CNN 1-D	44.6	*	8.4	10.9	0.0	84.3	73.1	0.0	87.4	68.3
	CNN 2-D	48.6	*	9.6	39.4	3.6	92.3	72.5	0.0	94.6	75.2
L1	LSTM	90.1	0.0	98.2	92.8	85.4	98.7	96.0	87.1	94.8	93.6
	CNN 1-D	99.4	0.0	89.5	95.3	82.2	99.2	98.2	75.5	98.8	95.9
	CNN 2-D	99.8	0.0	99.1	99.3	97.4	100.0	99.7	91.6	99.8	99.3
H1	LSTM	96.2	0.0	95.5	96.8	89.1	99.7	95.9	75.1	97.6	95.4
	CNN 1-D	99.0	0.0	90.1	99.3	91.6	98.4	100.0	80.6	97.4	96.5
	CNN 2-D	99.7	0.0	99.6	99.8	96.8	99.7	99.8	96.8	99.2	99.1

TRUE POSITIVE RATE

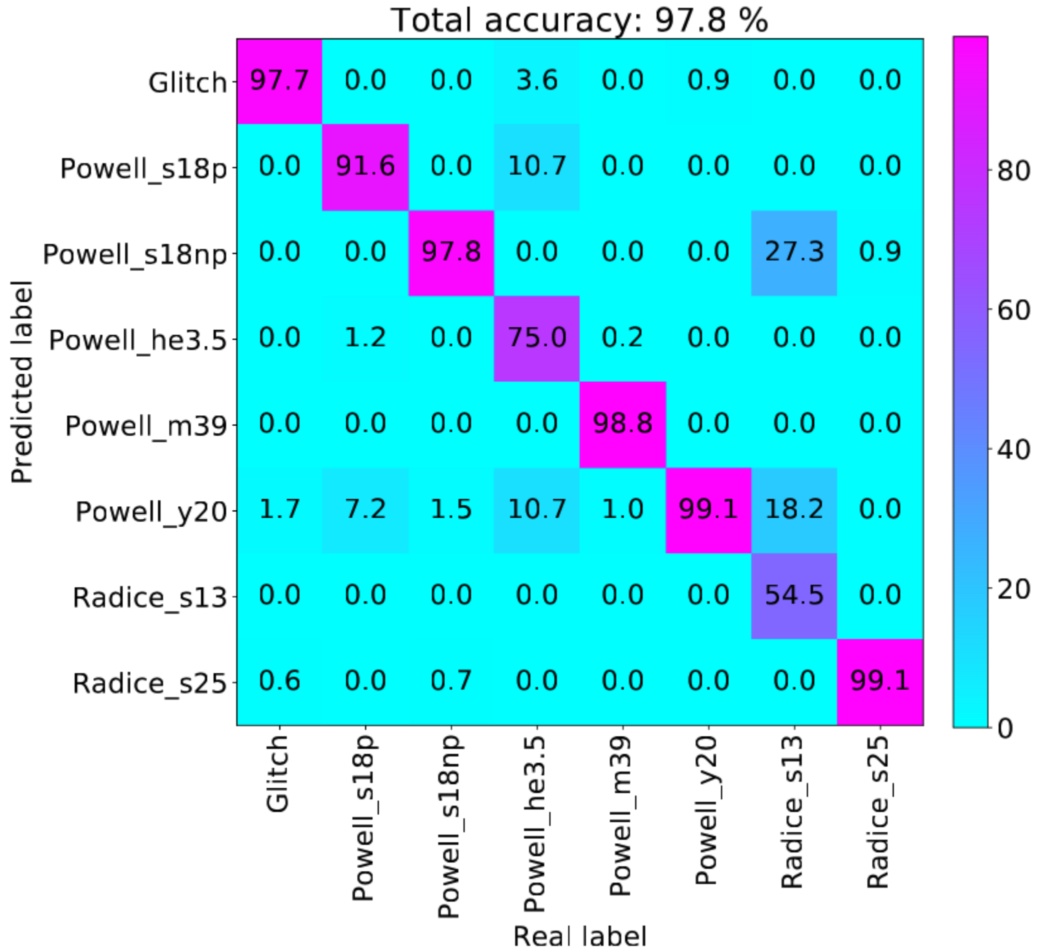
$$TPR = \frac{TP}{TP + FN}$$

COINCIDENT MERGED MODEL APPROACH

- Take only triple coincident triggers
- Input to NNs have additional dimension (ITF)
- Merge information from time-series and time-frequency representation.



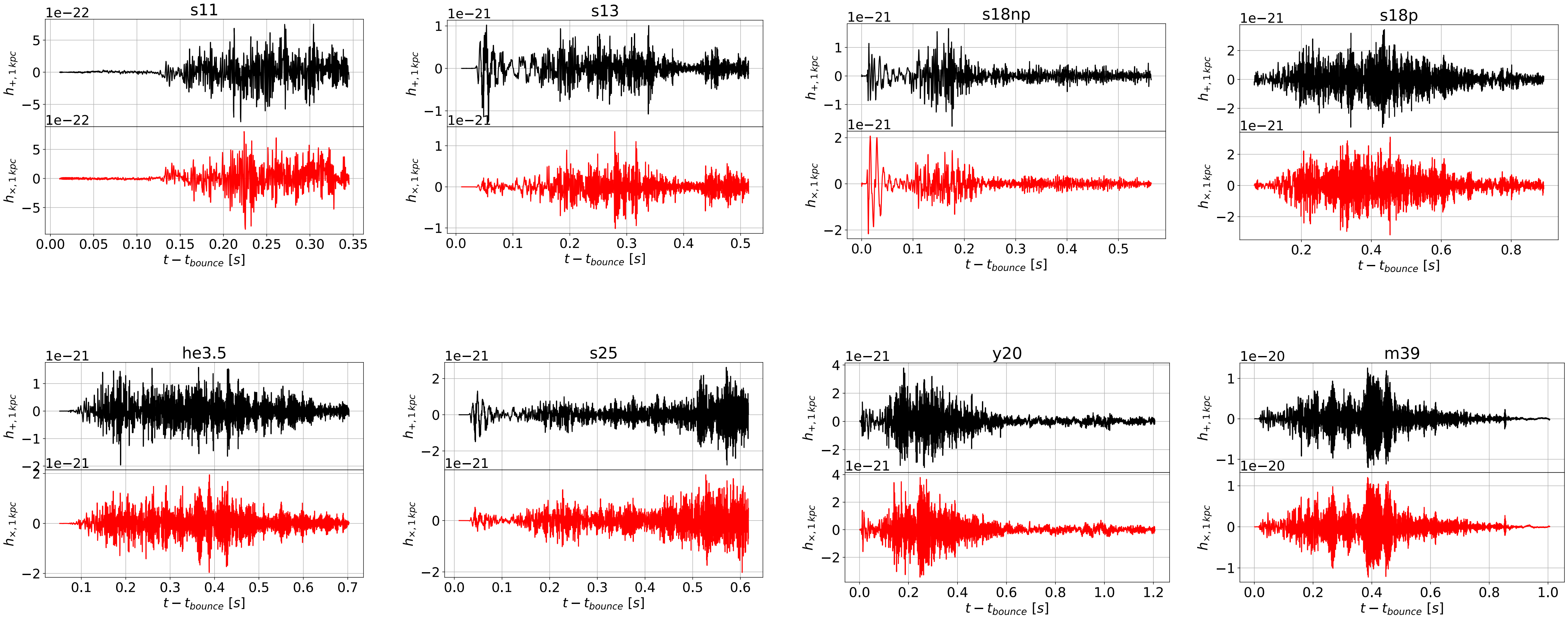
1. ACCUMULATE SNR
2. DECREASE FAR THROUGH COINCIDENCE REQUIREMENT
3. REDUCED TRAINING SET



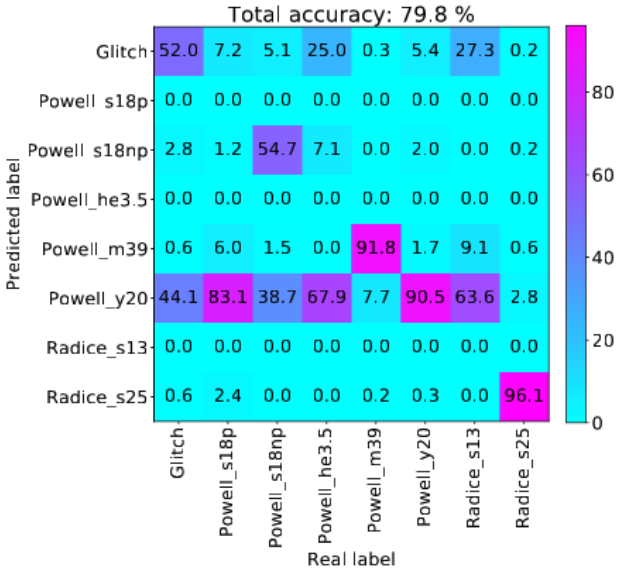
Iess, Cuoco, Morawski, Nicolaou, Lahav 2022 (accepted)

EXTRA SLIDES

CORE-COLLAPSE SUPERNOVAE DATASET
(neutrino-driven explosion mechanism)

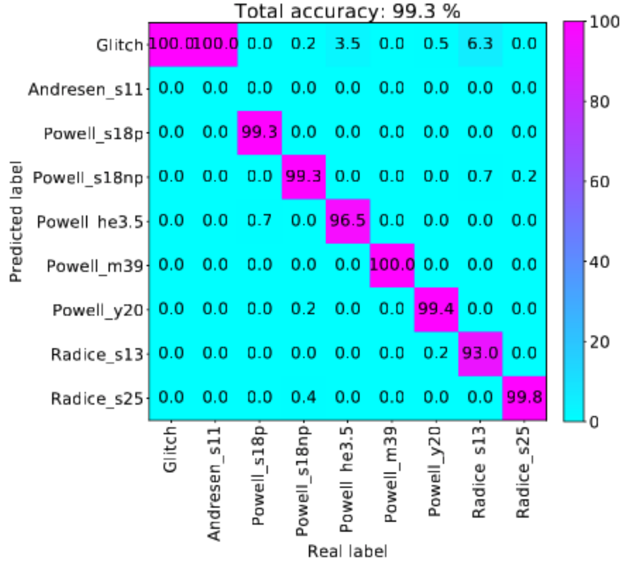


EXTRA SLIDES



V1

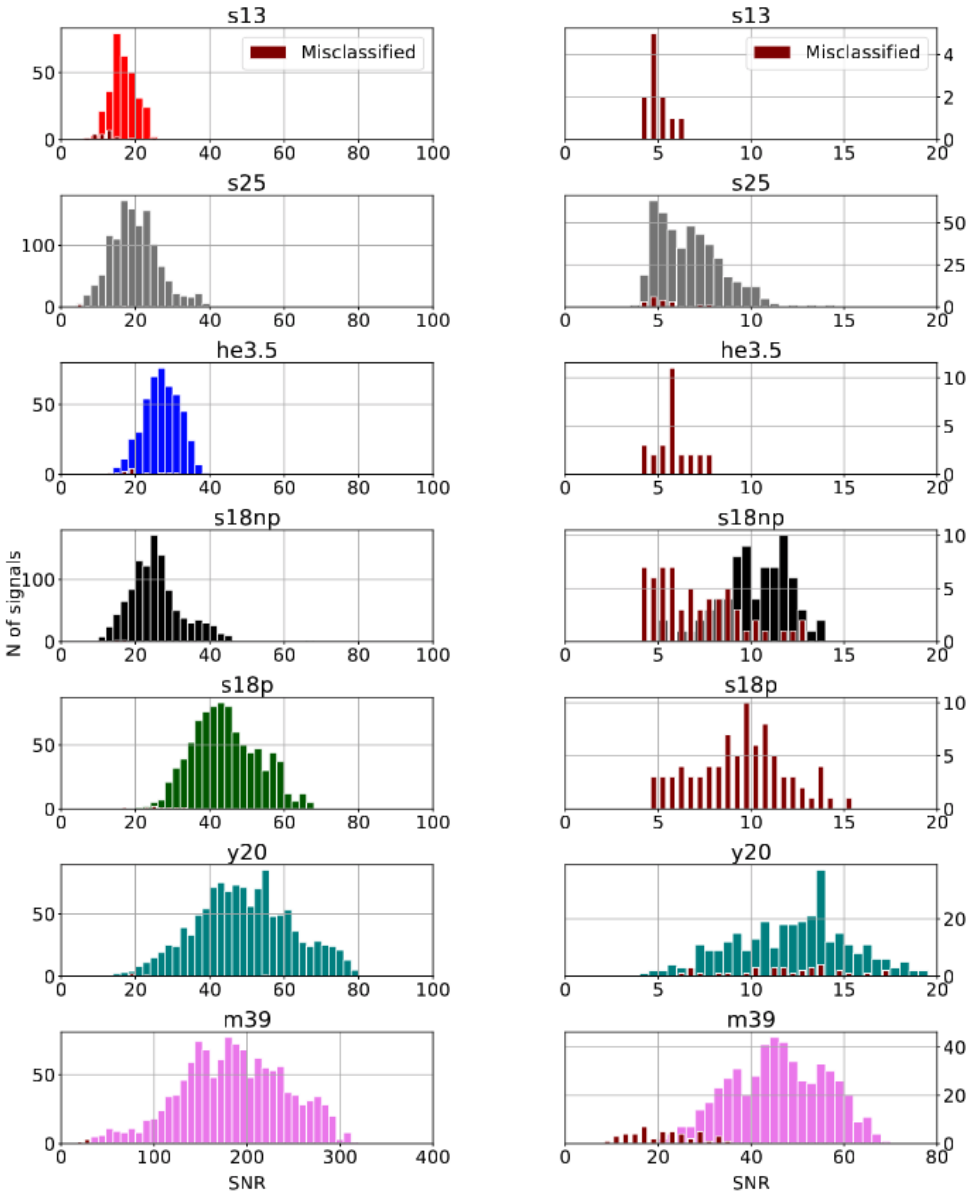
675 → noise
0 → s11
329 → s18p
491 → s18np
115 → he3.5
1940 → m39
1139 → y20
76 → s13
1557 → s25



L1

3810 → noise
12 → s11
1438 → s18p
1782 → s18np
704 → he 3.5
2052 → m39
1969 → y20
476 → s13
2047 → s25

L1+H1



V1

Iess, Cuoco, Morawski, Nicolaou, Lahav 2021 (submitted)

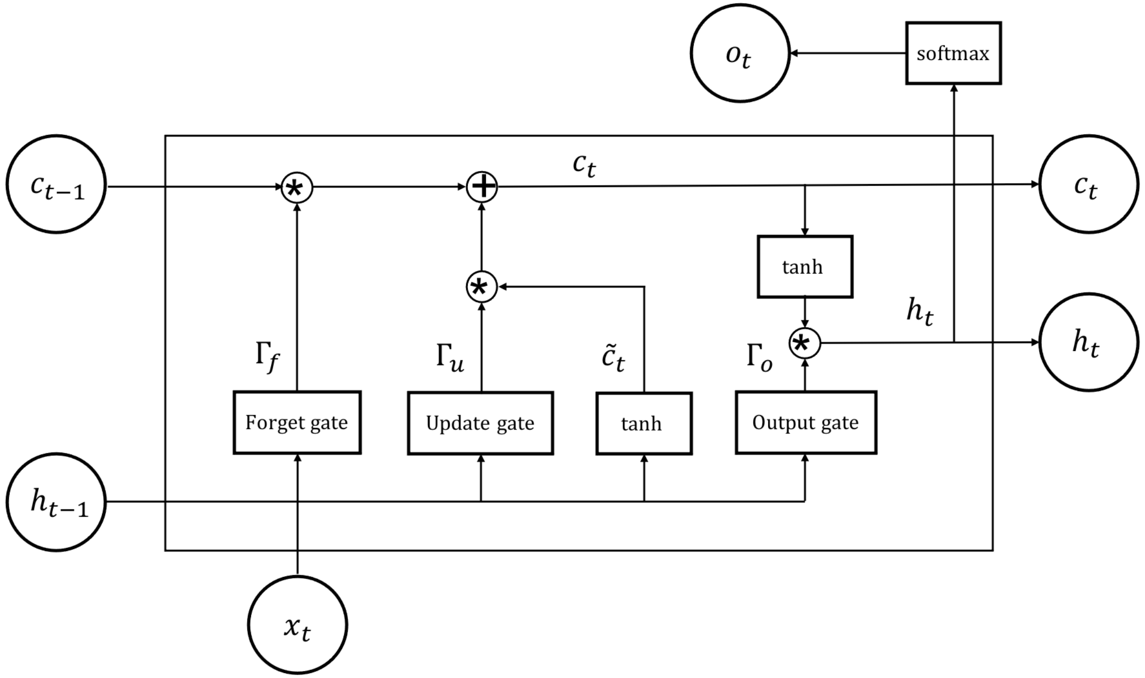
EXTRA SLIDES

PROS

- Keeps track of dependencies in time-series with internal loop updating a “state” cell (Hochreiter and Schmidhuber 1997).
- Avoids the *Vanishing Gradient* problem.

CONS

- Many parameters to train, long training times.
- Hyperparameter tuning can be challenging.
- Decreased performances for sequences above O(1000).

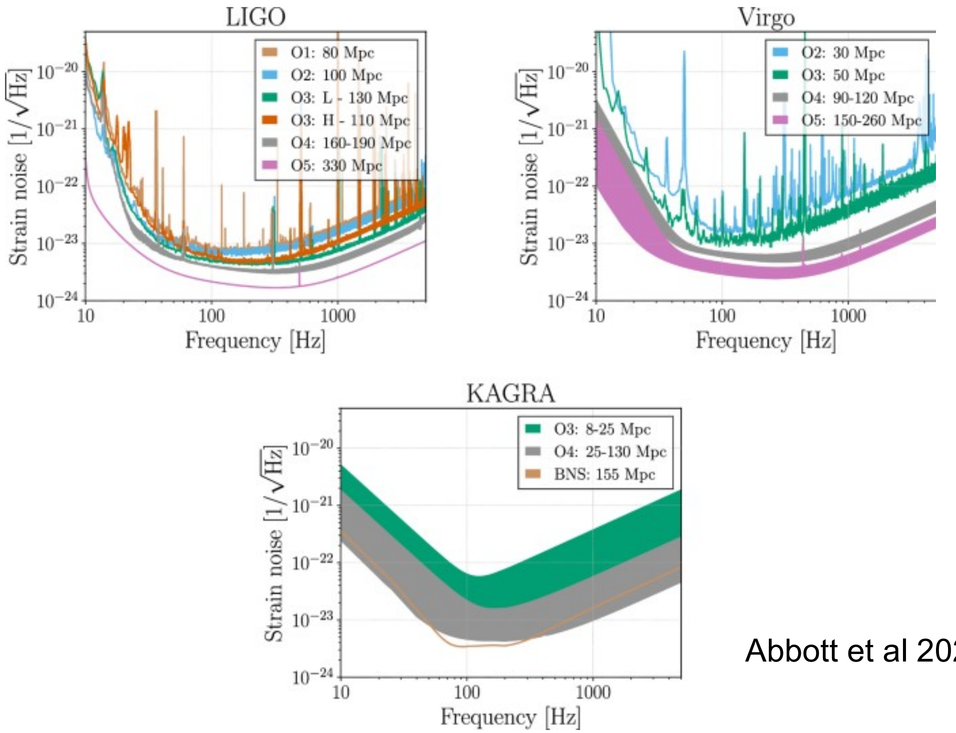


LSTM EQUATIONS

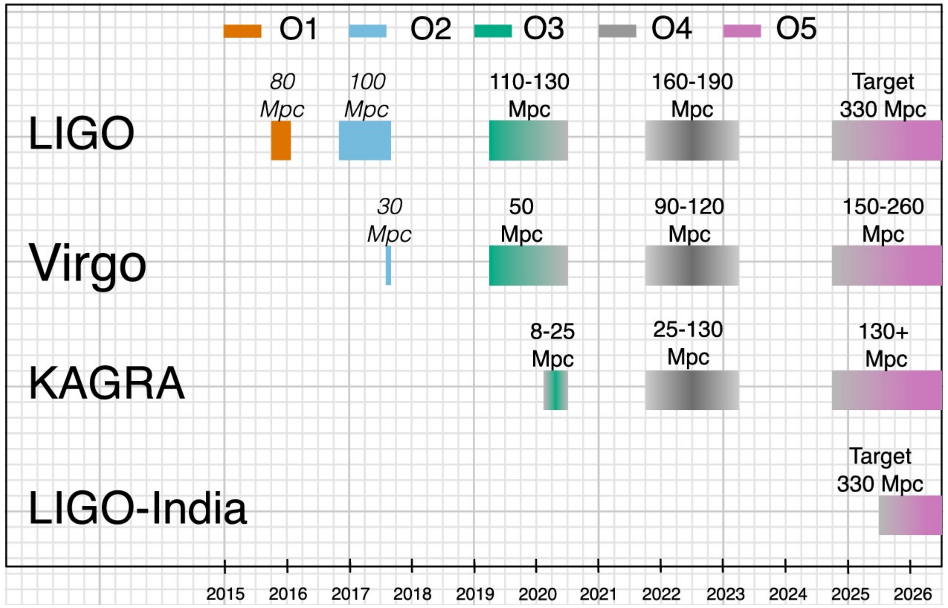
$$\tilde{c}_t = \tanh\left(W_{ch}h_{t-1} + W_{cx}x_t + b_c\right)$$
$$\Gamma_f = \sigma\left(W_{fh}h_{t-1} + W_{fx}x_t + b_f\right)$$
$$\Gamma_u = \sigma\left(W_{uh}h_{t-1} + W_{ux}x_t + b_u\right)$$
$$\Gamma_o = \sigma\left(W_{oh}h_{t-1} + W_{ox}x_t + b_o\right)$$
$$c_t = \Gamma_u * \tilde{c}_t + \Gamma_f * c_{t-1}$$
$$h_t = \Gamma_o * \tanh\left(c_t\right)$$

EXTRA SLIDES

- Deep Learning: a fast tool to implement in search pipelines.
- Can reduce the FAR by providing glitch vetoing.
- Can be adapted to multiple sources, in particular when matched filtering cannot be applied.
- Machine Learning models can analyze simultaneously different types of data (EM,GW, Neutrino..) and therefore constitute a promising framework for multi-messenger data.
- Possibility to detect Core-Collapse supernovae GW for events in Milky Way for neutrino-driven models with current generation of detectors, in Milky Way neighbourhood (Mpcs) with 3rd generation ET and CE.



Abbott et al 2020



EXTRA SLIDES

DeepClean