Alberto Iess

12 April 2022

Core-Collapse Supernovae Detection And Classification In Real Noise With CNN And LSTM

Working Group 1 Meeting

Valencia, Spain

SCUOLA NORMALE SUPERIORE

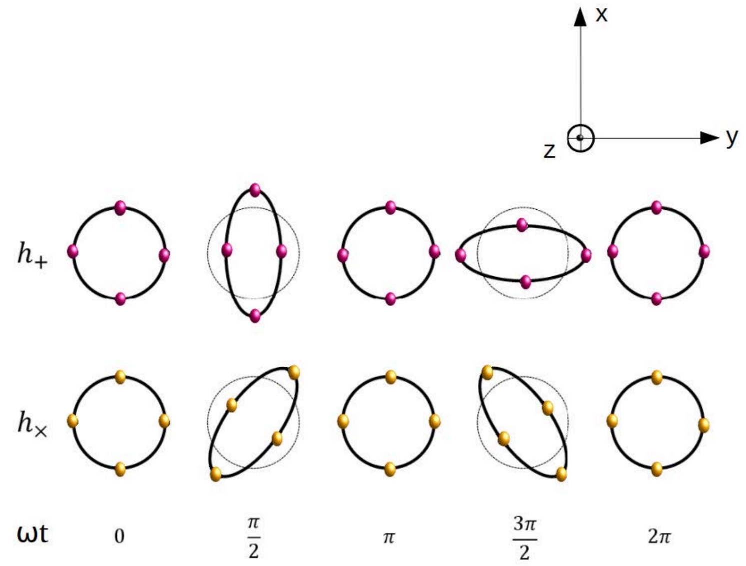
cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

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)

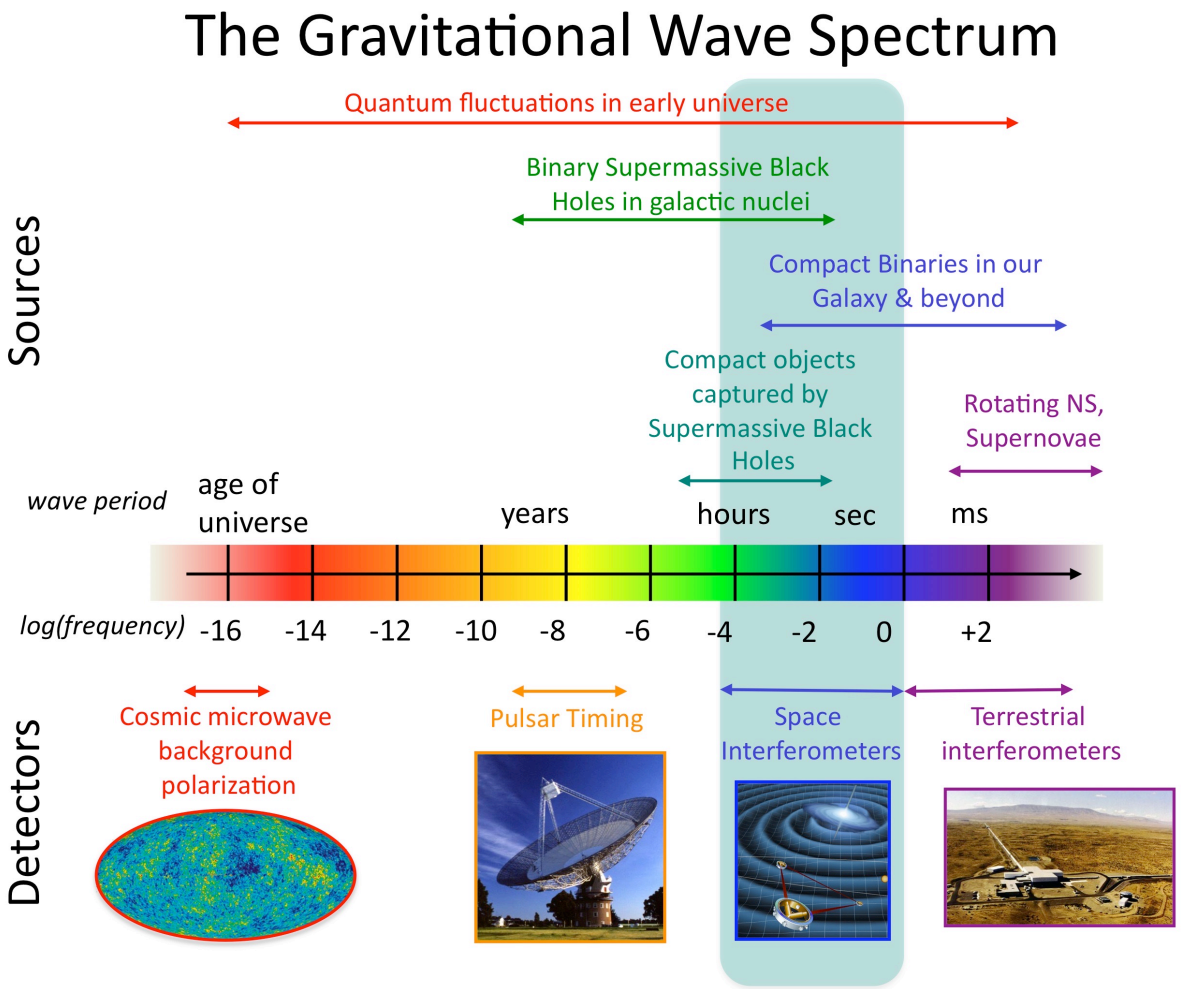
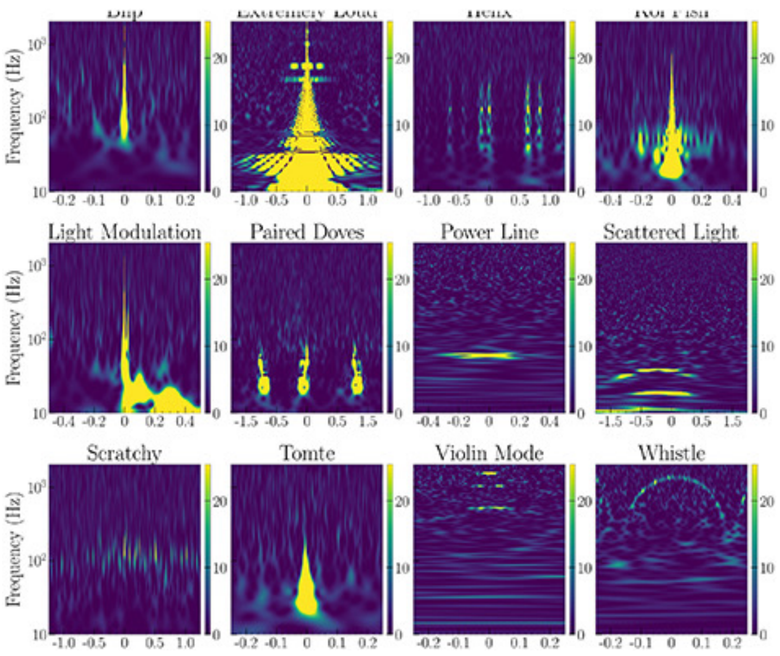


Image credit: NASA Goddard Space Flight Center

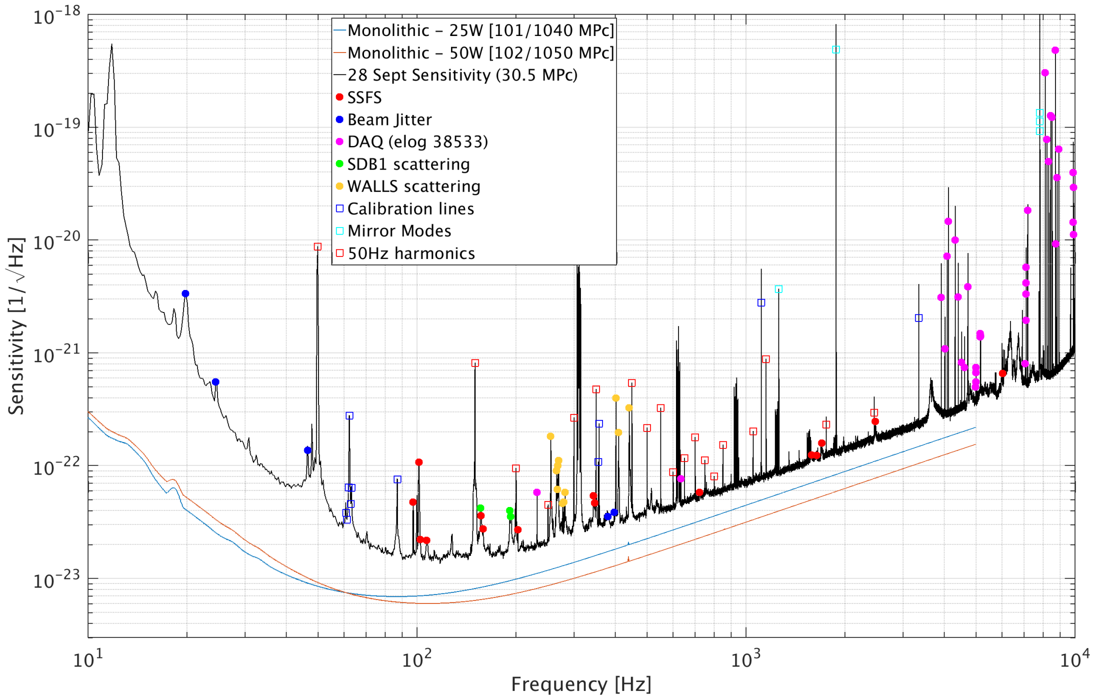
INTERFEROMETER NOISE

TRANSIENT NOISE BURSTS (GLITCHES)



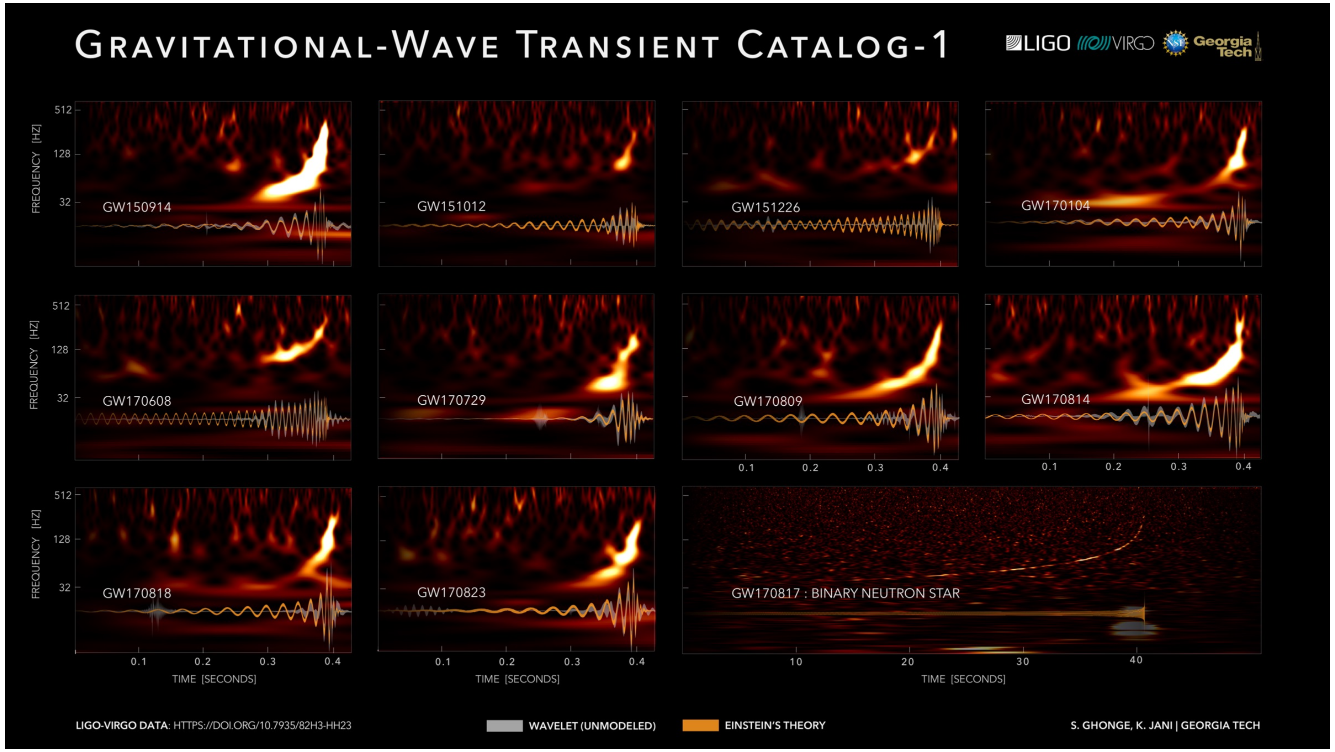
Bahaadini et al. (2018)

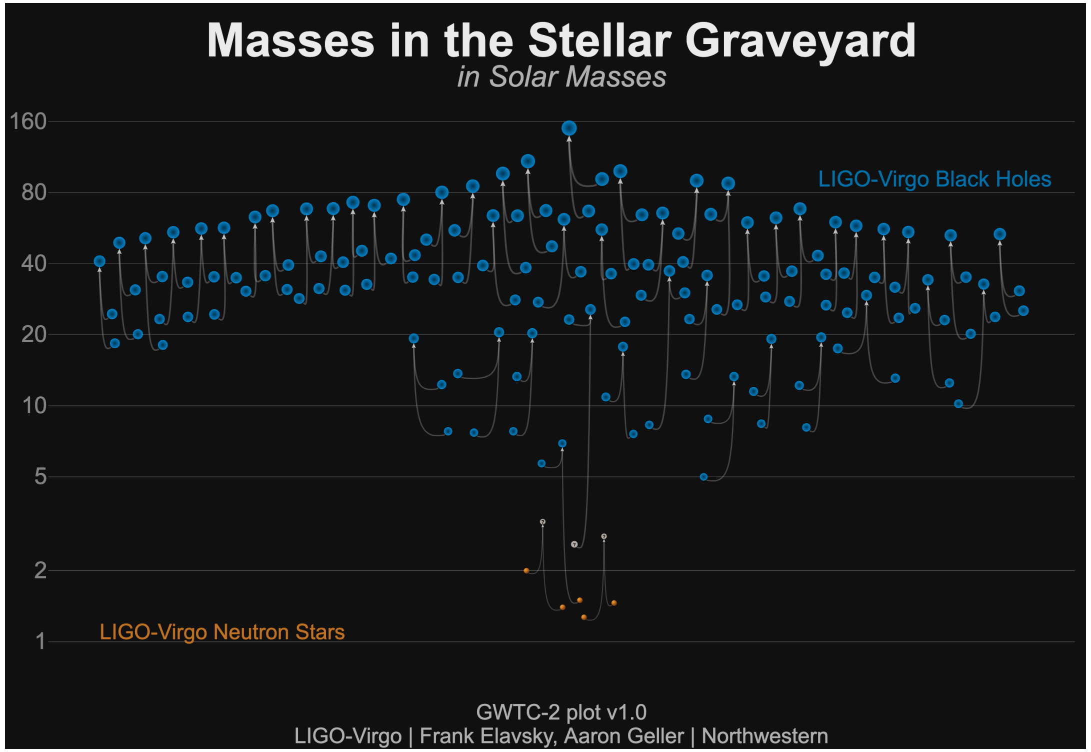
LONG LIVED DISTURBANCES, NOISE LINES



Virgo logbook entry elog40306

- 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

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

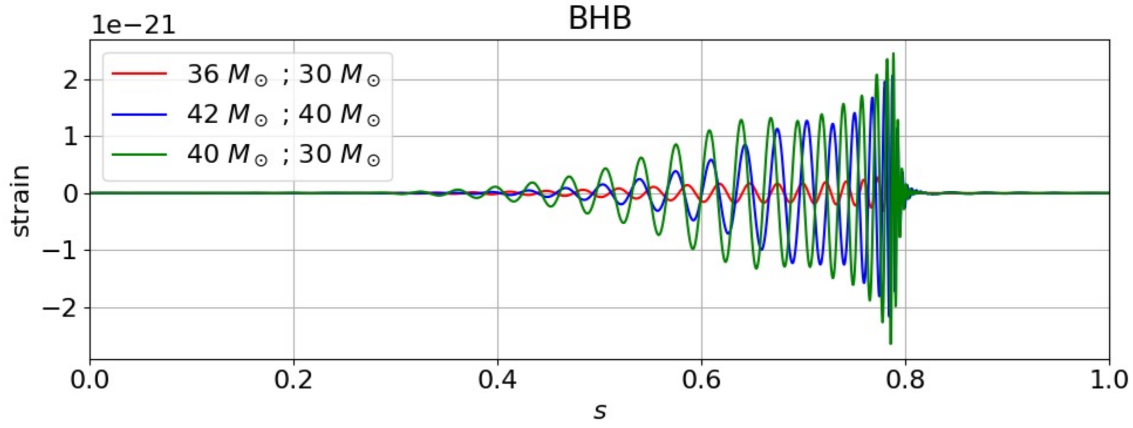
$$\langle s|h\rangle(t) = 4\int_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi i f t}\,\mathrm{d}f$$

Matched filter correlation

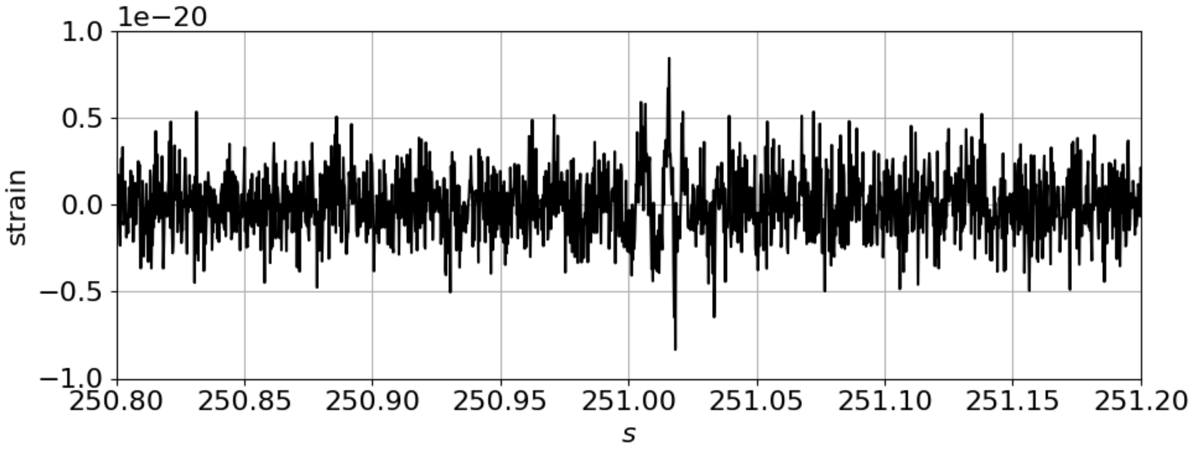
$$\langle h|h\rangle = 4\int_0^\infty \frac{\tilde{h}(f)\tilde{h}^*(f)}{S_n(f)}\,\mathrm{d}f$$

Template Normalization

- Large template bank to cover the parameter space (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.

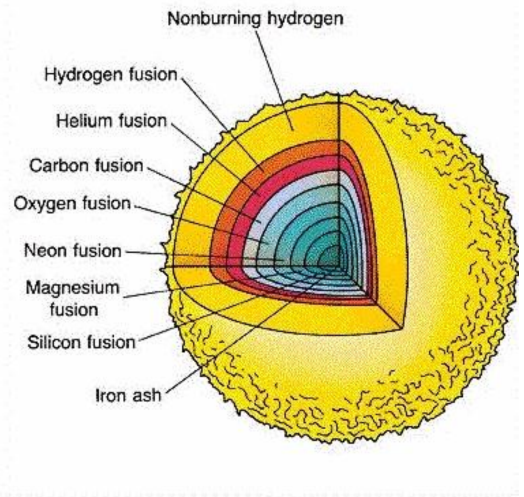


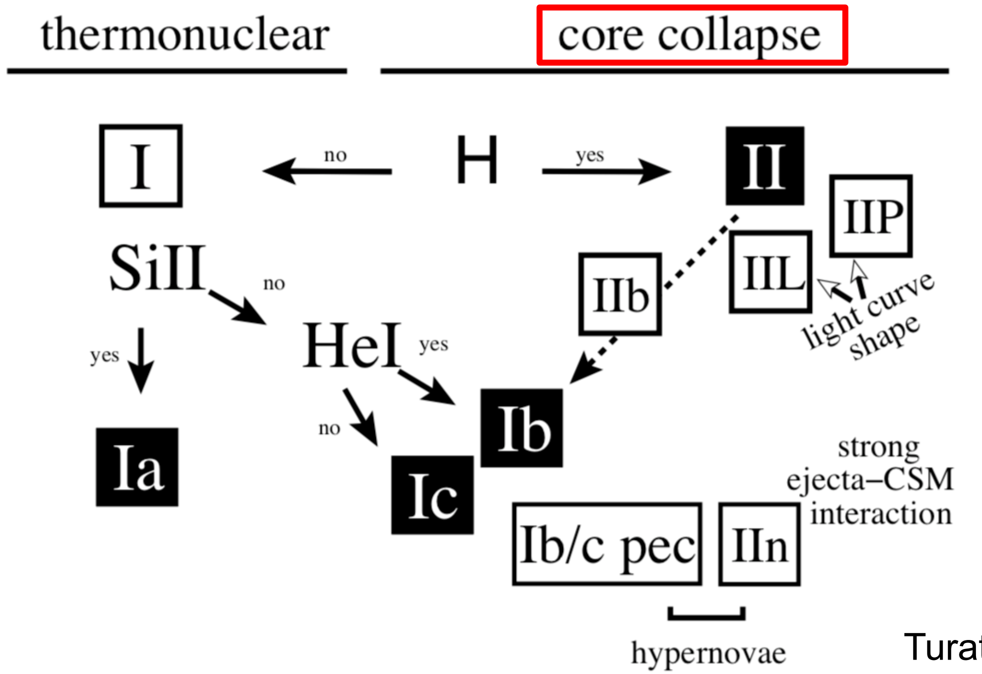
Waveform template



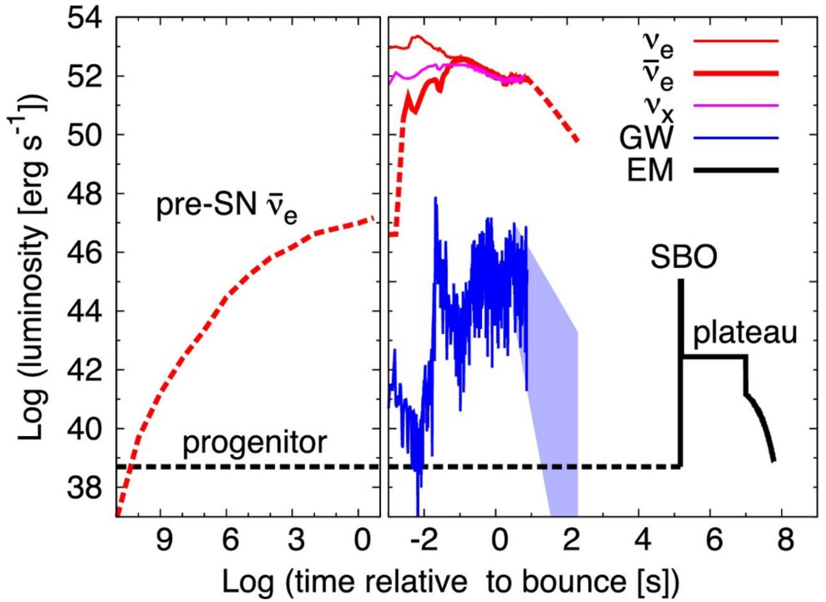
Whitened interferometer data

- 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)



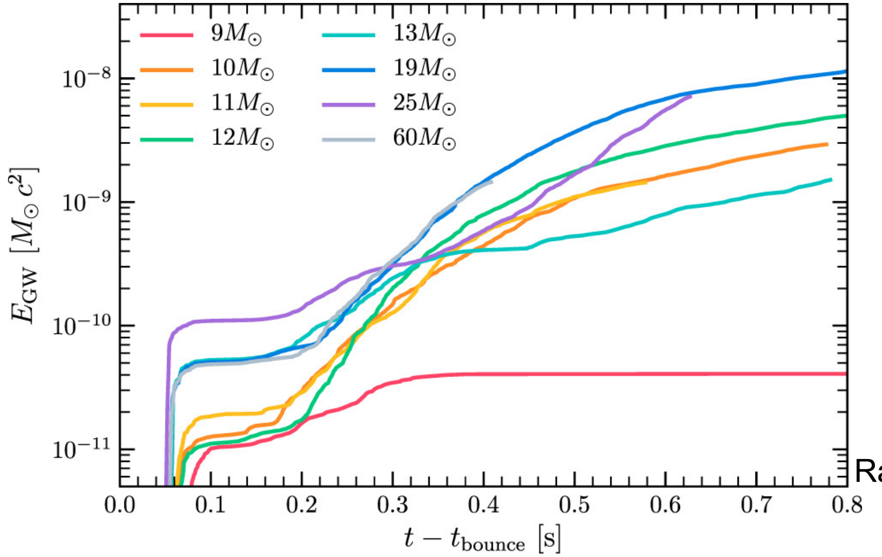
Nakamura et al. (2016)

GWs FROM CORE-COLLAPSE SUPERNOVAE

- Waveform depends on progenitor star.
- Different emission mechanisms.
- Largely Stochastic.
- Broadband emission.
- Best waveform models from computationally expensive 3D simulations.

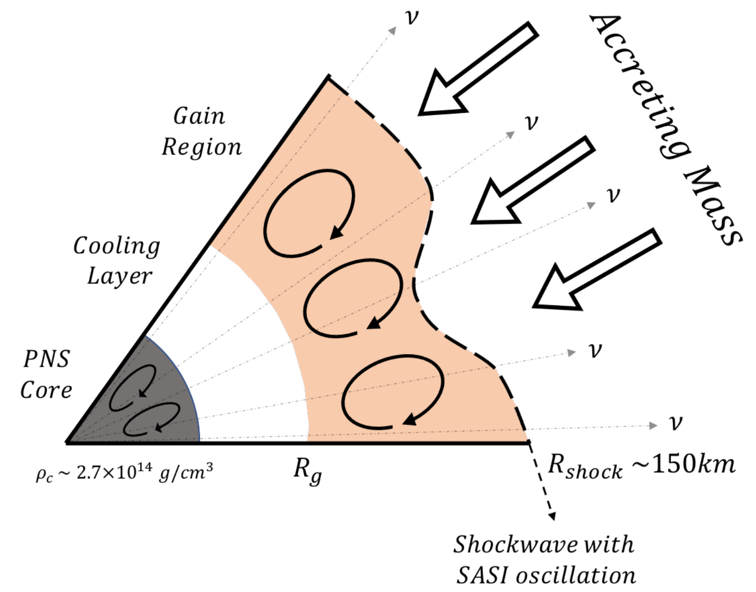
↓

Matched filter not feasible!



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

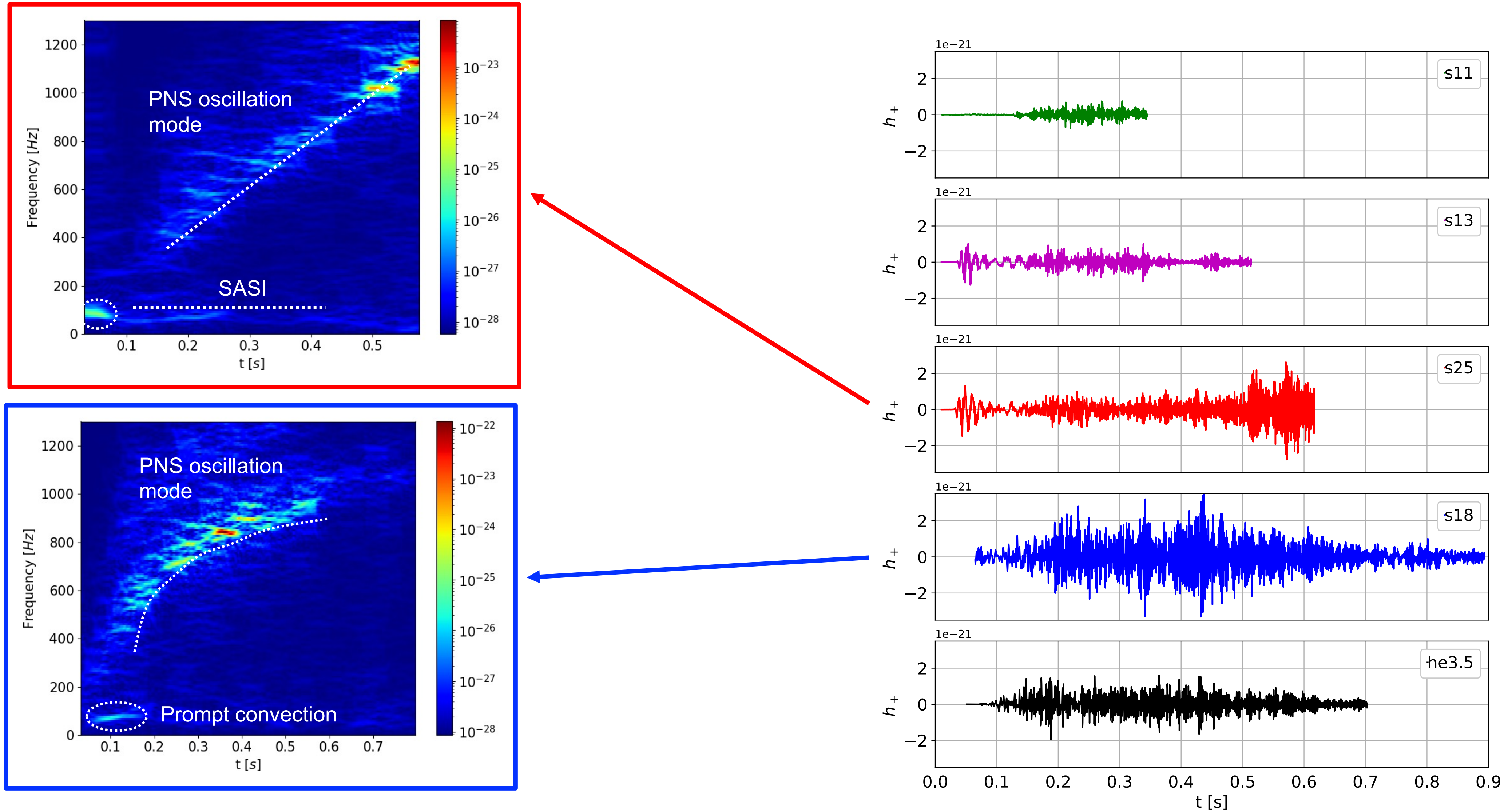
Radice et al. (2017)



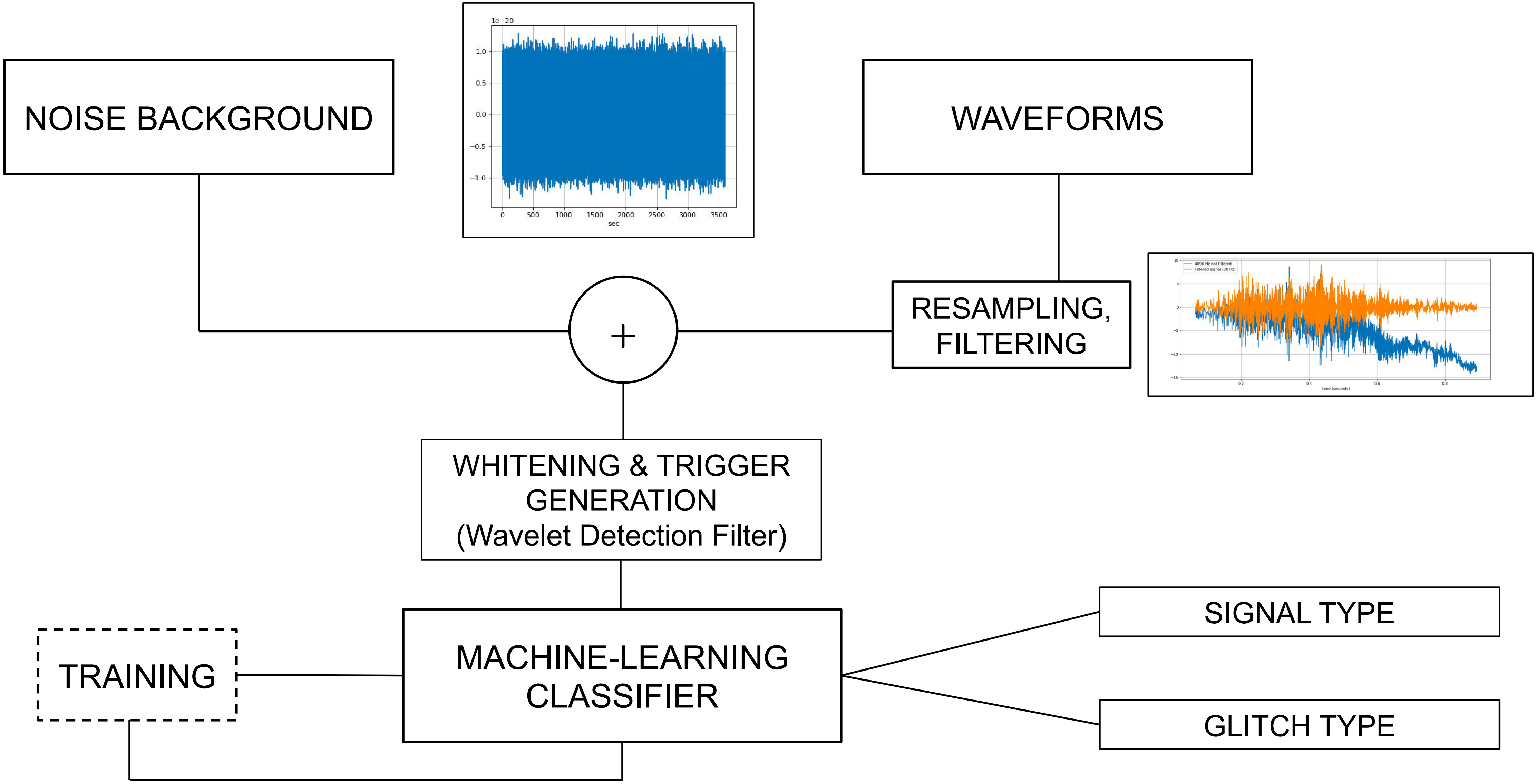
GW emission Process	Potential explosion mechanism		
	MHD mechanism (rapid rotation)	Neutrino mechanism (slow/no rotation)	Acoustic mechanism (slow/no rotation)
Rotating collapse and Bounce	Strong	None/weak	None/weak
3D rotational instabilities	Strong	None	None
Convection & SASI	None/weak	Weak	Weak
PNS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

Ott et al. (2017)

Core-Collapse Supernovae Detection And Classification In Real Noise With CNN And LSTM

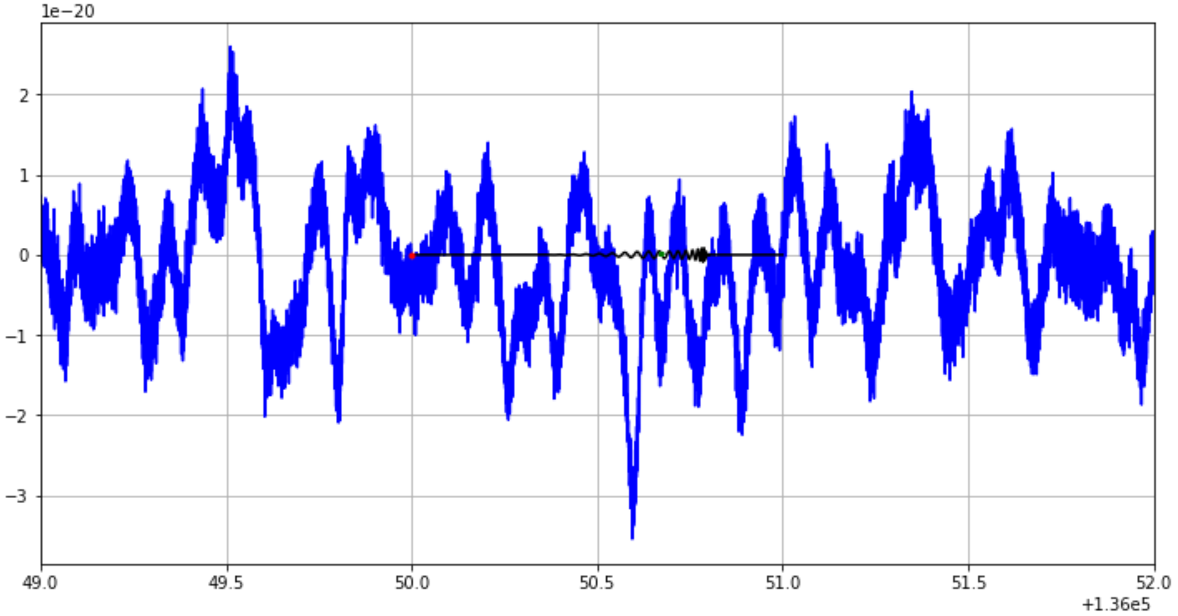


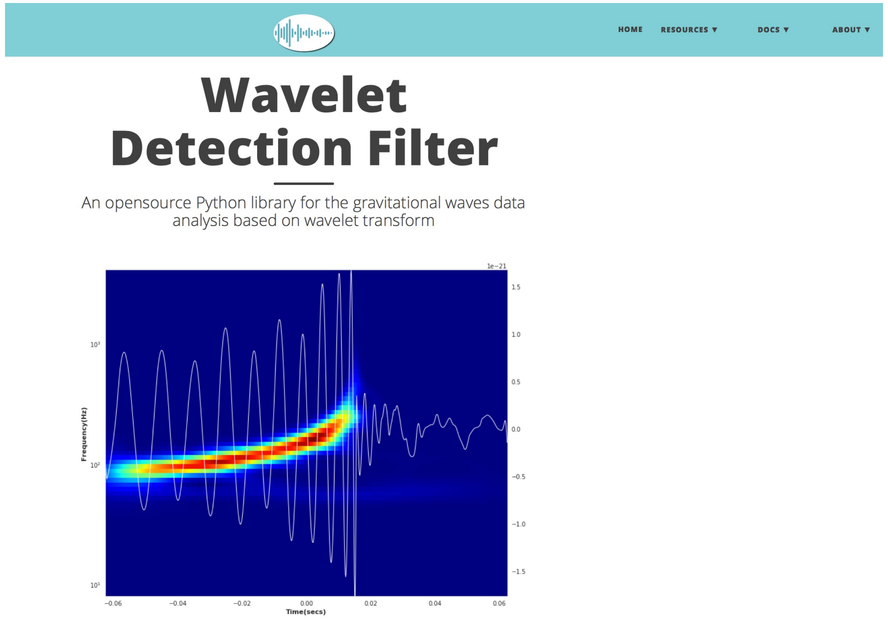
Iess, Cuoco, Morawski, Powell (2020)


NOISE BACKGROUND
1e−20
sec
WAVEFORMS
RESAMPLING, FILTERING
time (seconds)
+
WHITENING & TRIGGER GENERATION (Wavelet Detection Filter)
TRAINING
MACHINE-LEARNING CLASSIFIER
SIGNAL TYPE
GLITCH TYPE

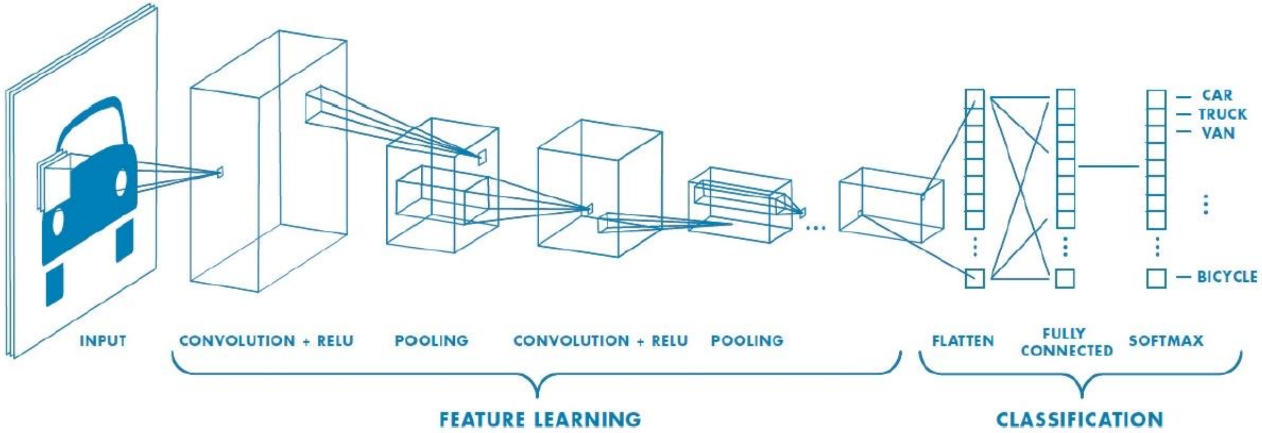
WAVELET DETECTION FILTER (WDF) (Cuoco et al. 2001, 2018)

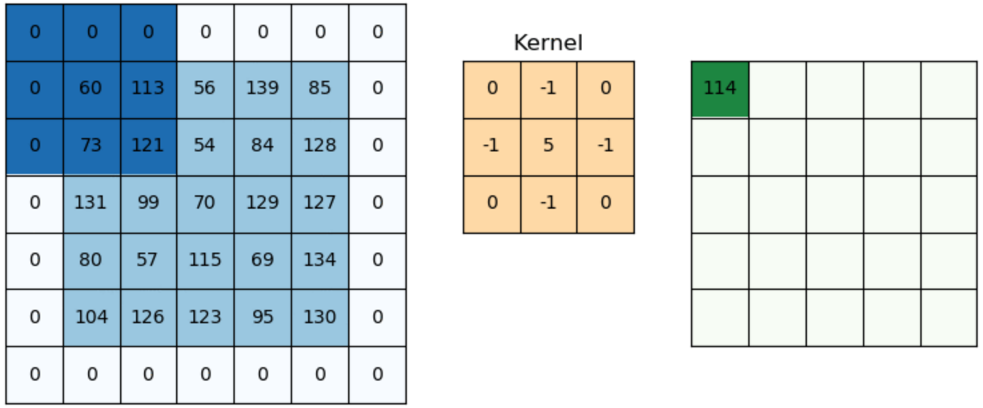
- Python Library for Signal Processing and Transient Detection based on wavelet transform.
- Dockers implementation for lightweight use on any platform, including HPC clusters.
- Website: https://wdfpipe.gitlab.io
- ReadTheDocs: https://wdf.readthedocs.io/en/latest





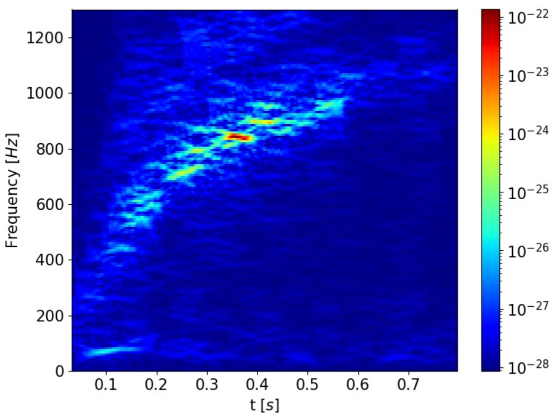
CONVOLUTIONAL NEURAL NETWORKS





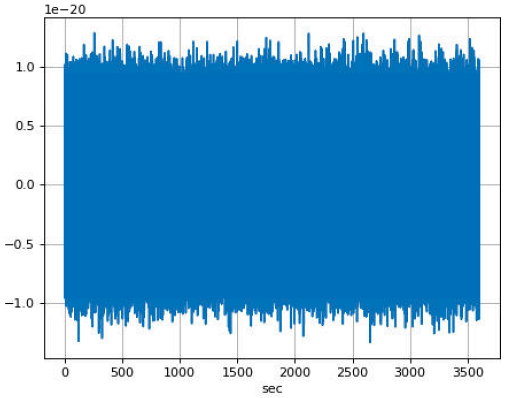
2-D CNN

Spectrogram images



1-D CNN

Whitened time series



SCUOLA NORMALE SUPERIORE

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

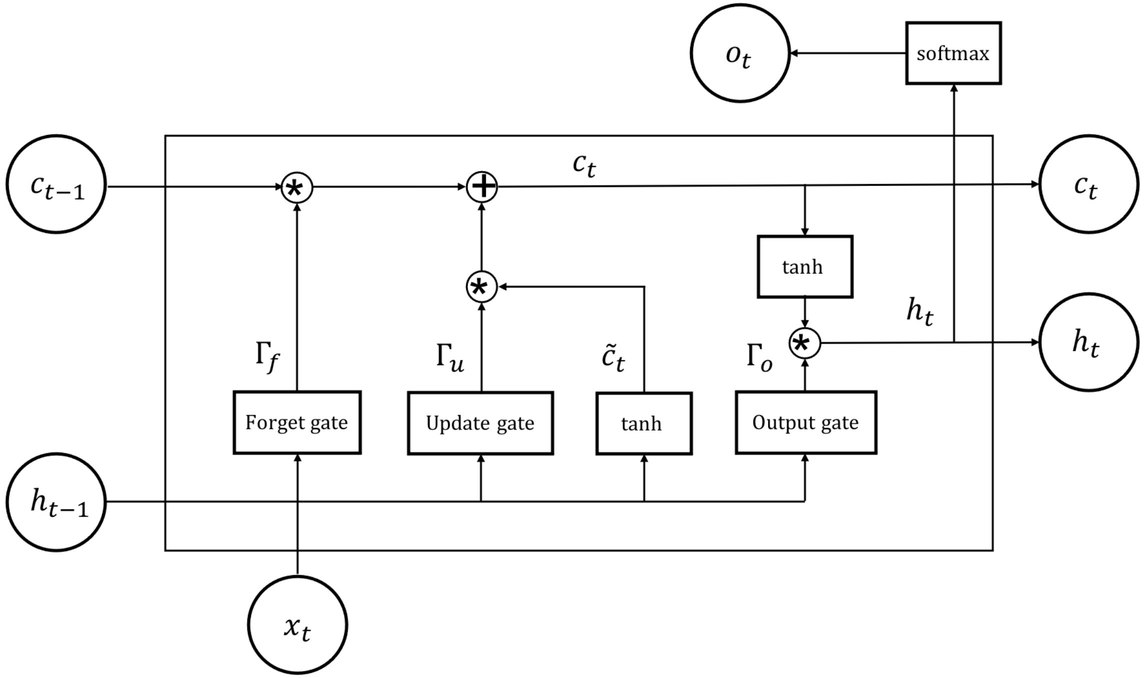
LONG SHORT TERM MEMORY NETWORKS

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

- Long training times.
- Hyperparameter tuning can be challenging.
- Decreased performances for sequences above O(1000).

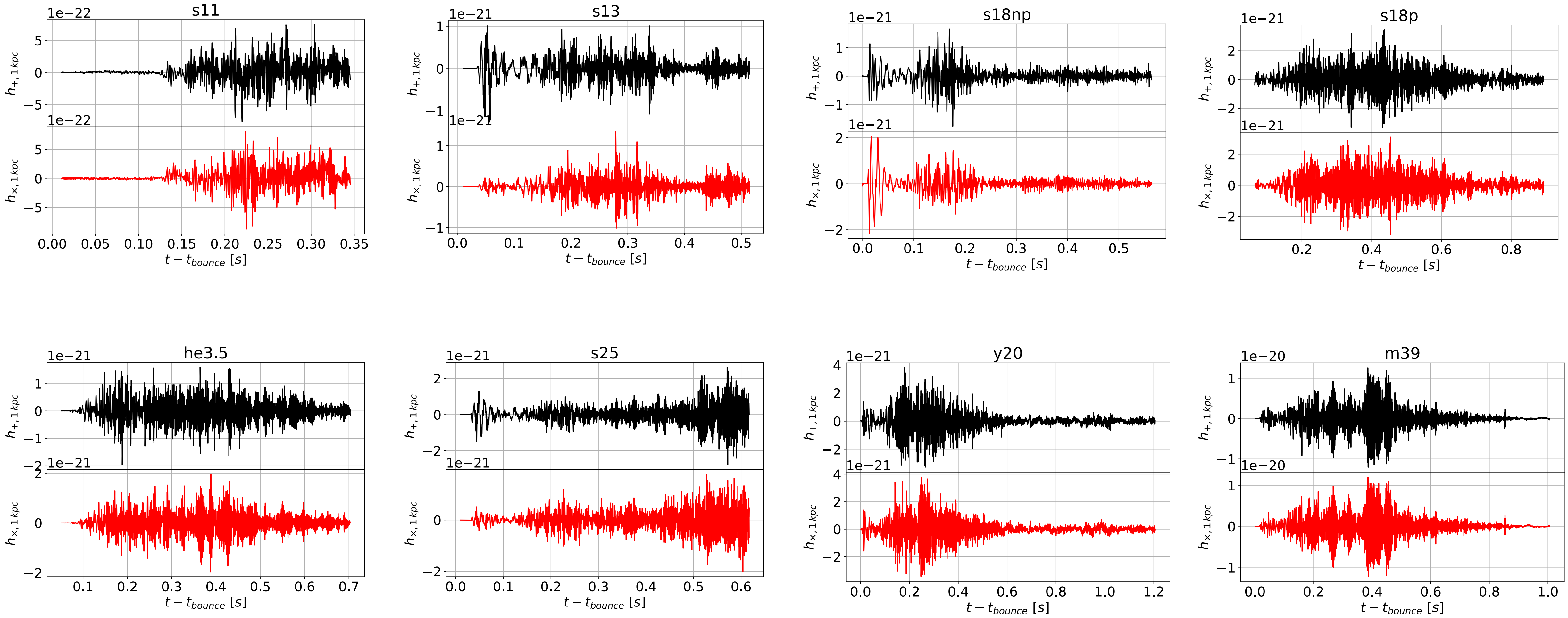


LSTM EQUATIONS

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

CORE-COLLAPSE SUPERNOVAE DATASET

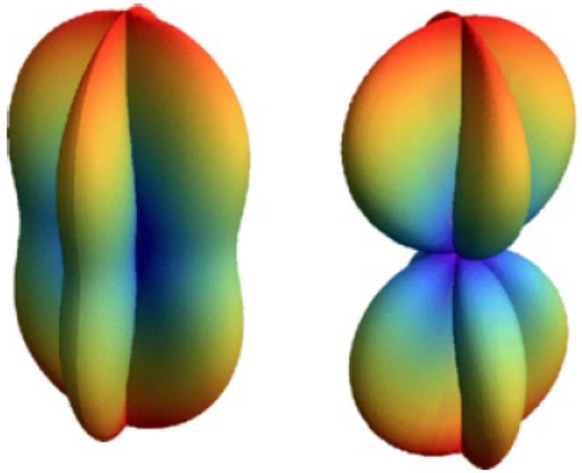
(neutrino-driven explosion mechanism)



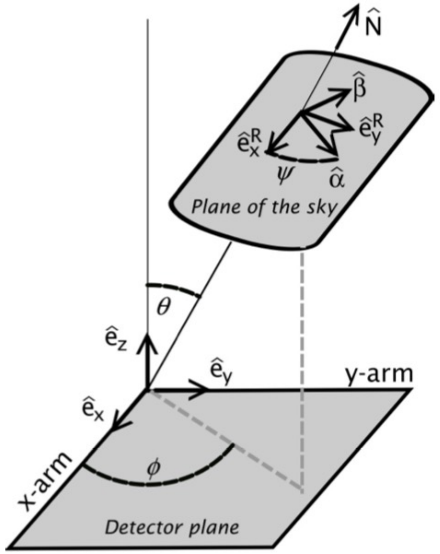
GW INJECTIONS

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

F_+ $F_\times$



Adhikari 2014



Schutz 2011

LIGO Hanford

LIGO Livingston

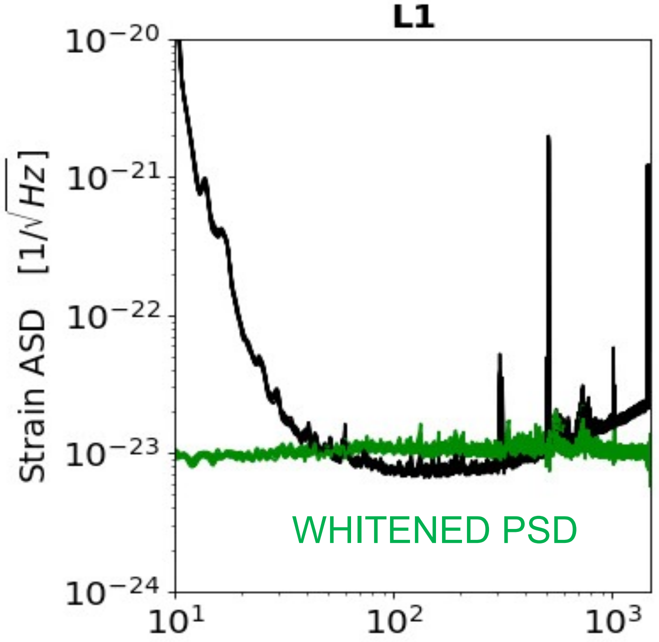
Virgo

DATASET CHARACTERISTICS

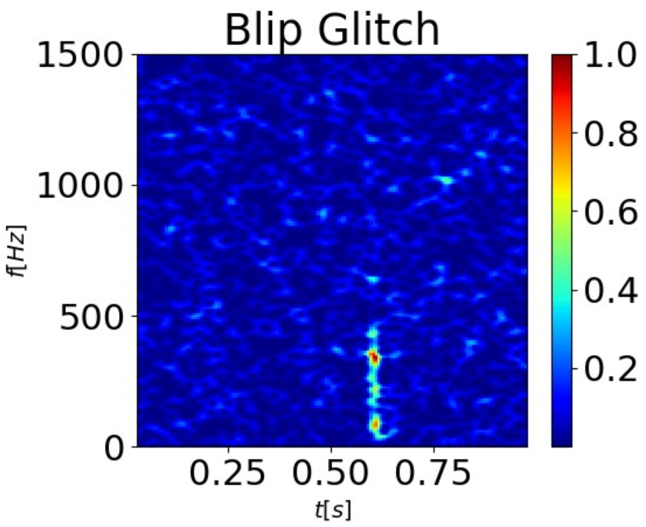
- 44 segments (4096s per segment) from O2 science run
- CCSN Dataset: ~15000 samples, **Fixed distance of 1 kpc**.
- Multiple Glitch Families.

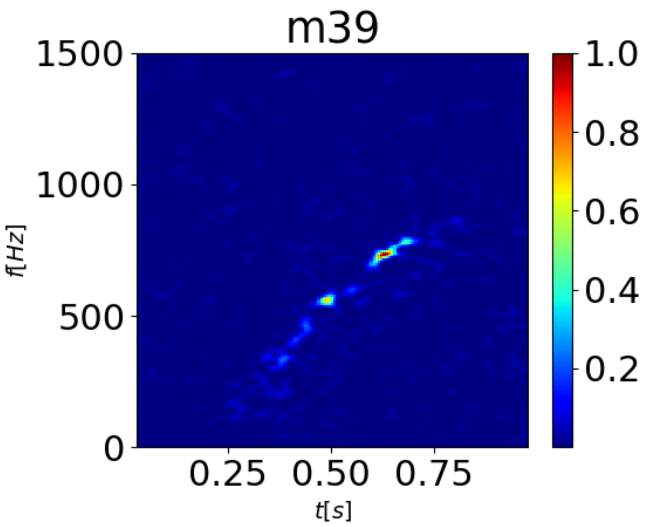
- SNR distribution is affected by ITF antenna pattern (and sensitivity).
- Detector noise PSD is non stationary.
- **Imbalanced Dataset** due to different model amplitudes.

	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



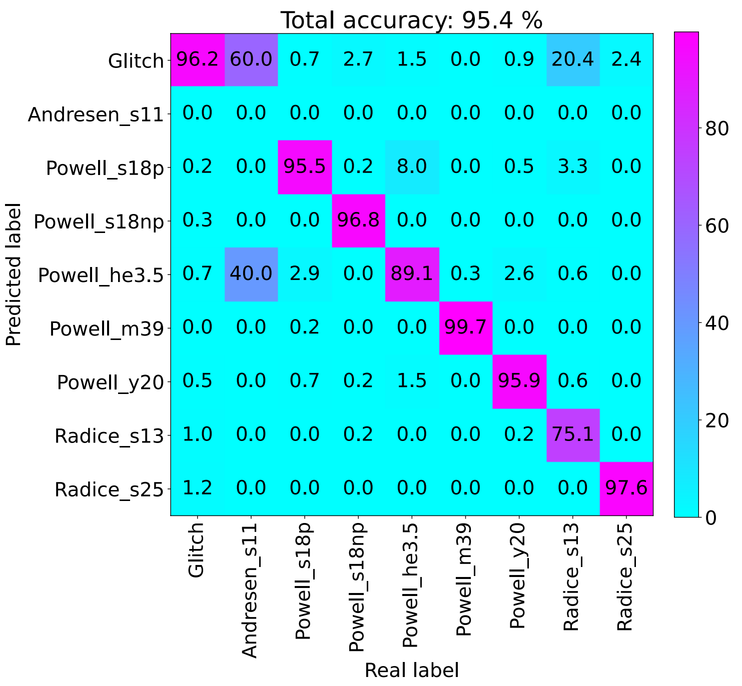
… AFTER WDF WHITENING



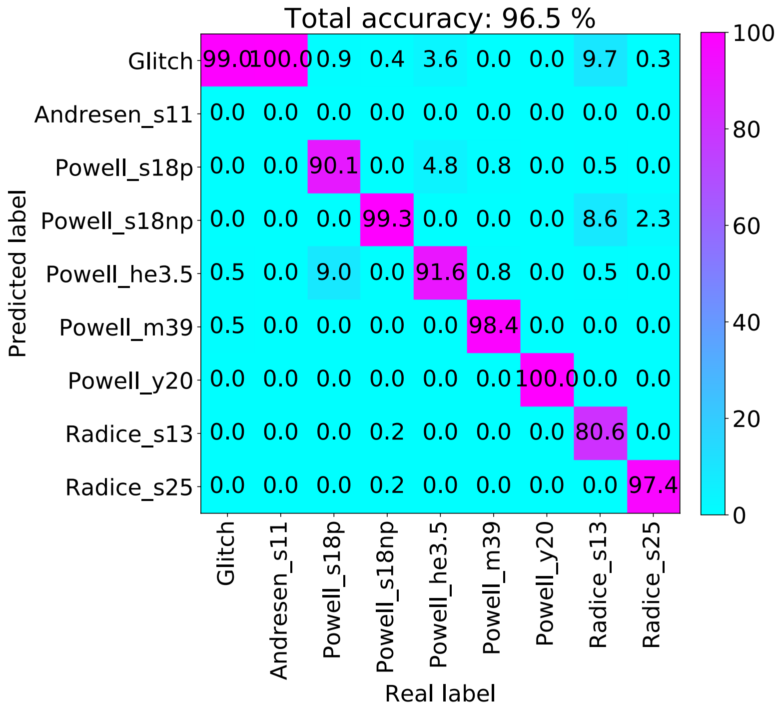


COMPARISON OF ML MODEL ACCURACY
(Single interferometer, LIGO H1)

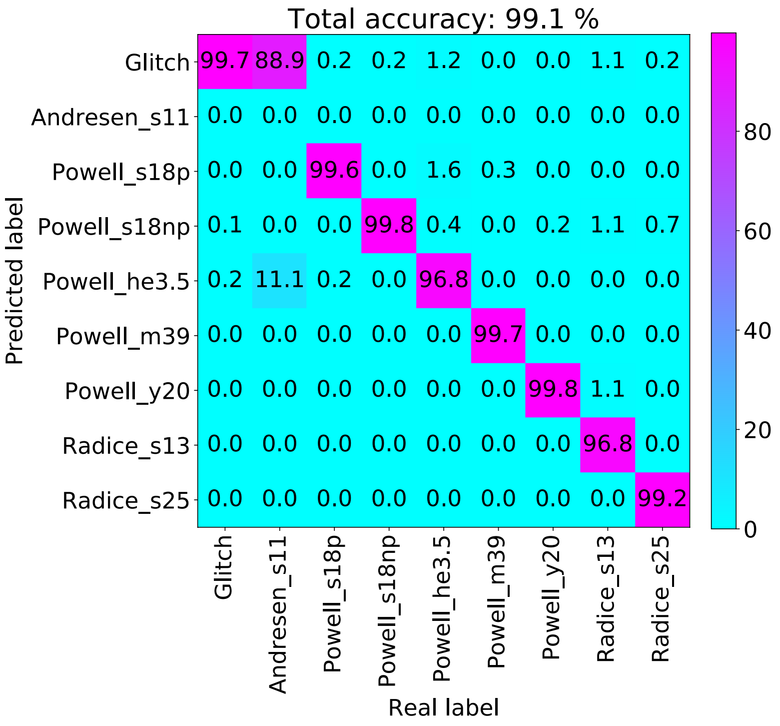
- **Bi-LSTM**, 2 recurrent layers
- ~10 ms/sample
- Best weights over 100 epochs



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



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



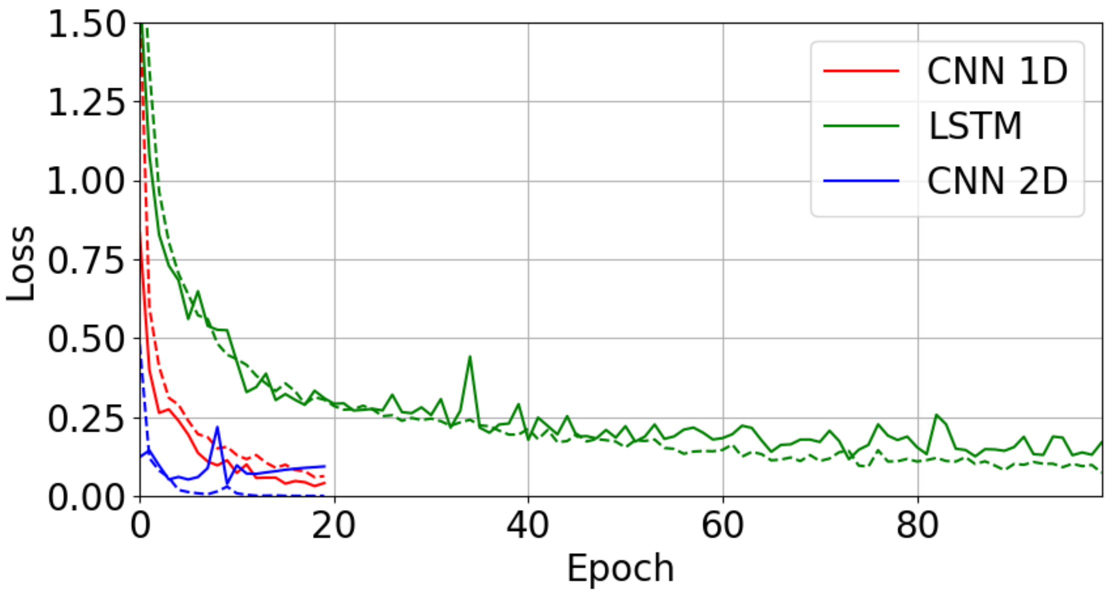
COMPARISON OF ML MODEL ACCURACY

(Single interferometer, V1,H1,L1)

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

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

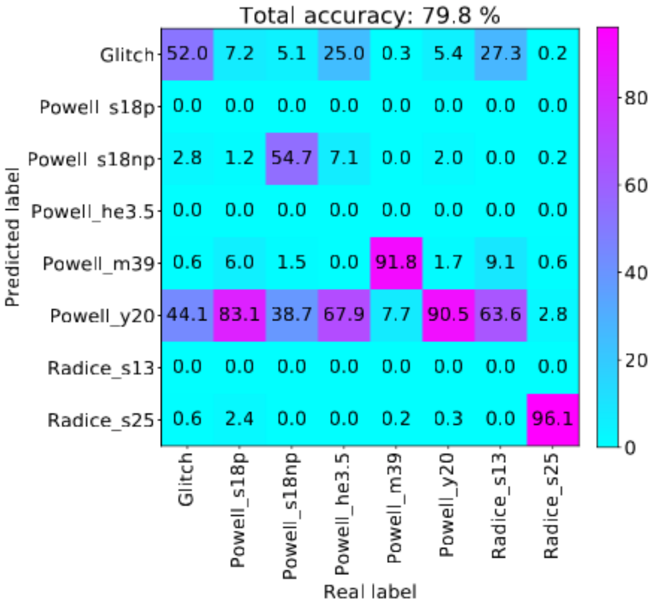
- **2D-CNN**, 4 convolutional layers
- ~3 ms/sample
- Best weights over 20 epochs



		Waveform Model									
ITF	**ML Model**	*Noise*	*s11*	*s18p*	*s18np*	*He3.5*	*m39*	*y20*	*s13*	*s25*	**Total**
	LSTM	49.2	*	3.6	58.4	0.0	89.5	69.9	0.0	89.8	73.7
V1	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
	LSTM	90.1	0.0	98.2	92.8	85.4	98.7	96.0	87.1	94.8	93.6
L1	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
	LSTM	96.2	0.0	95.5	96.8	89.1	99.7	95.9	75.1	97.6	95.4
H1	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

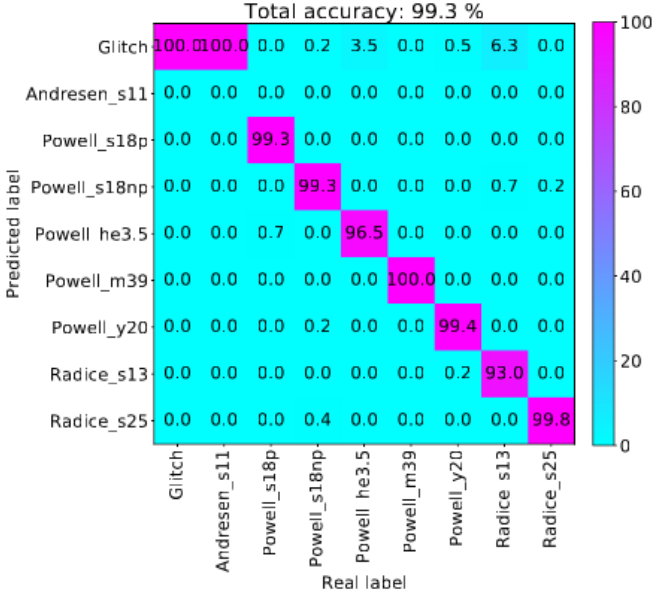
TRUE POSITIVE RATE

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



V1

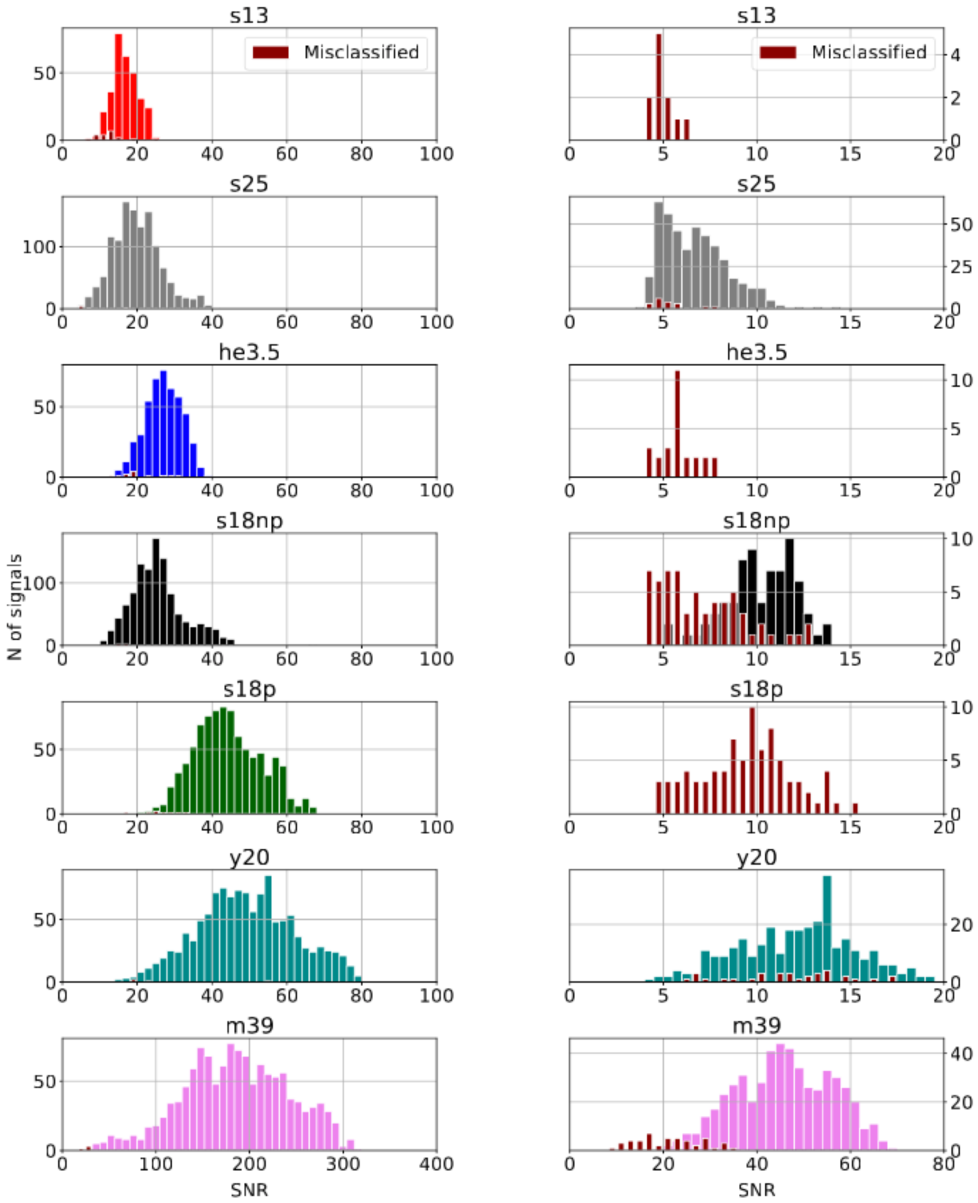
675 → noise
0 → s11
329 → s18p
491 → s18np
115 → he3.5
1940 → m39
1139 → y20
76 → s13
1557 → s25



L1

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

L1+H1

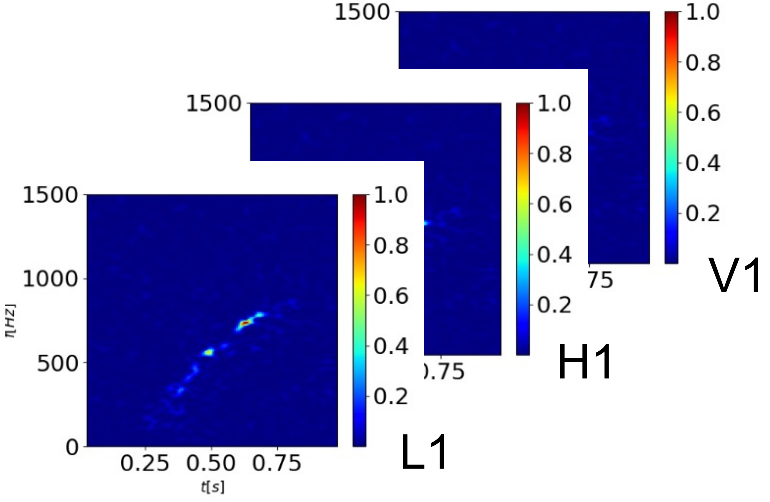


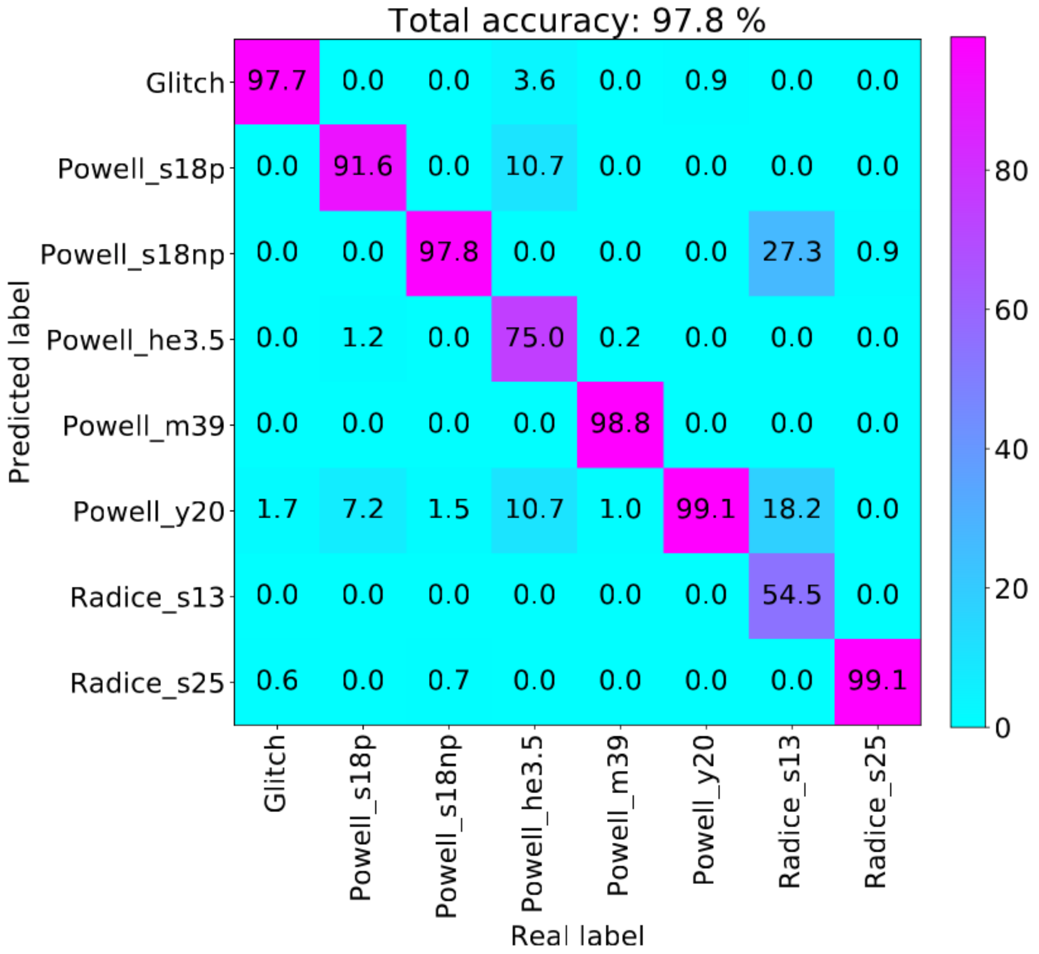
V1

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

CLASSIFICATION WITH 3 ITF INPUT

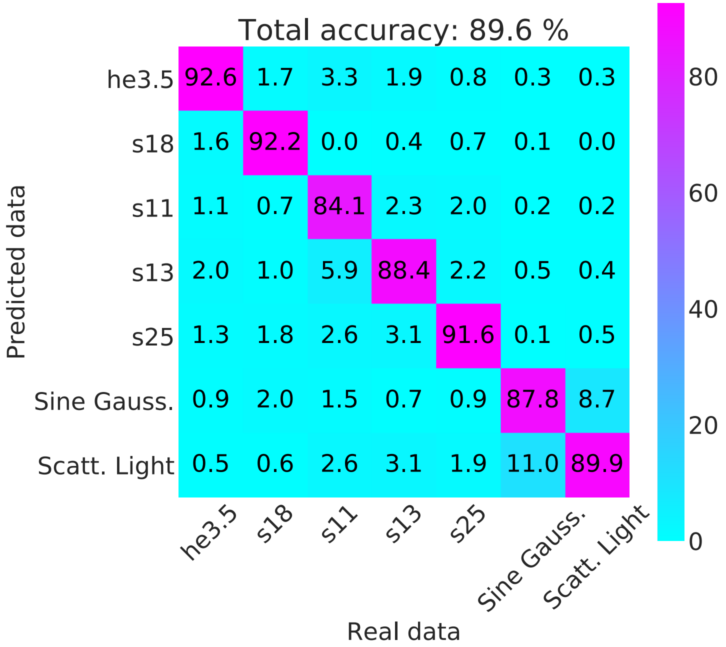
- Dataset breakdown (ONLY triple coincident triggers): 675 noise, 329 s18p, 491 s18np, 115 he3.5, 1940 m39, 1139 y20, 76 s13, 1557 s25.
- Input to NNs have additional dimension (ITF)





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

SIMULATED NOISE, SINGLE ITF *(Iess et al. 2020)*



Only low Q values