Anomaly detection using convolutional Auto-encoders

4th Einstein Telescope Annual Meeting
12/11/2025

U. Dupletsa, **H. Haigh,** G. Inguglia, K. Vitulova
Marietta Blau Institute

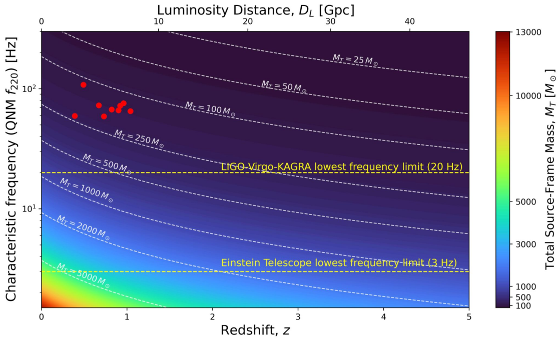
Introduction

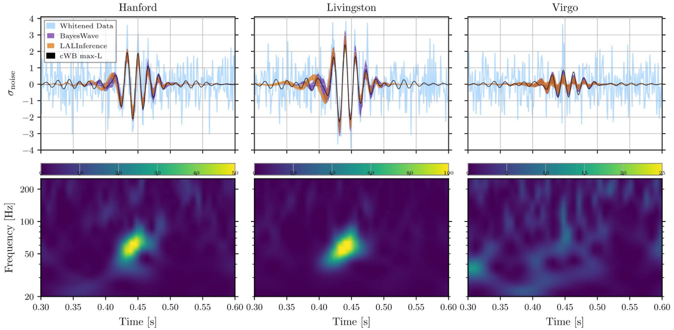
- Einstein Telescope promises vast increase in frequency range -> sensitivity to many more mergers.

- With such reach, matched filtering comes with large **computational overhead.**

- Deep learning offers potential **(semi-) model independent** technique with bulk of computational cost occurring before application.

- **Burst signals** from mergers involving **IMBHs** are short - we study the use of **convolutional autoencoders** to extract from noise.





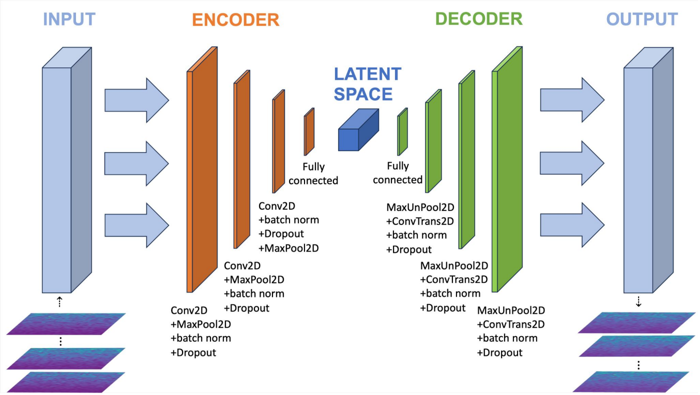
GW190521: [2009.01075]

Convolutional Autoencoders

- Convolutional auto-encoders are deep neural nets designed to compress and then **reconstruct** input image data.

- Structure:
  - **Encoder**: convolutional + pooling layers to pick out important patterns and compress to internal representation
  - **Latent space**: bottle neck that holds learned representation of key features.
  - **Decoder**: reverse of encoder -> reconstructs image from latent space representation.

- **Anomaly detection**: trained on background/noise images, learns to represent features in latent space -> reconstruct noise. **High reconstruction error (MSE) indicates anomalous** input image.



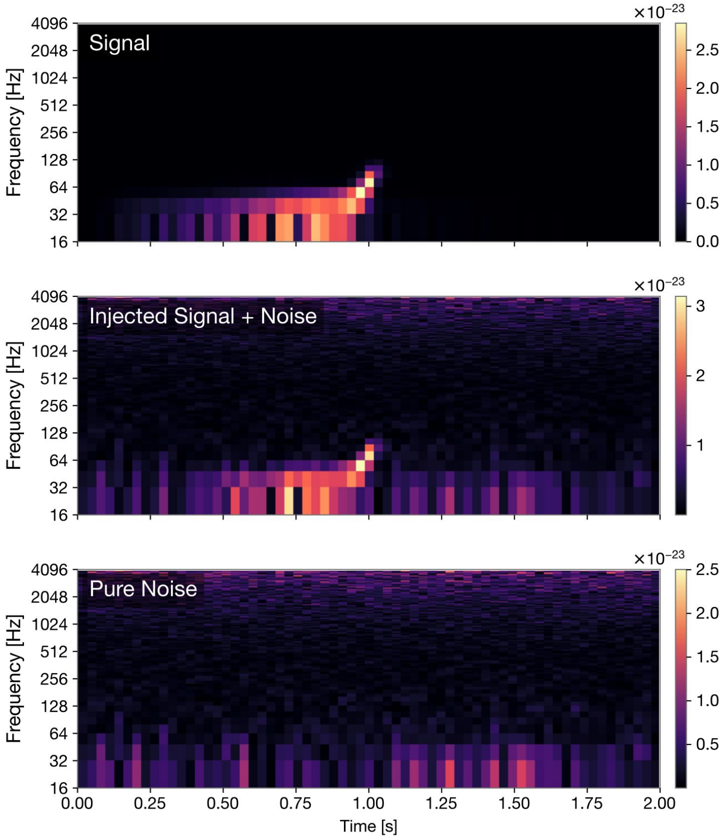
Spectrograms

- Full ET MDC - 1 month of simulated ET noise in 2048s segments sampled at 8192Hz (use only E1 channel).
- Split into **2s chunks**, PSD estimated using FFT, noise whitened.
- Greyscale images of **256 (frequency) x 31 (time) pixels** - normalised to [0,1] before training.
- Waveforms generated using IMRPhenomHM - randomly sampled masses & luminosity distance

Mean squared error loss used in training

$$\mathcal{L}_{\mathrm{MSE}} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2$$

Unsupervised training: model learns only on noise spectrograms



Weak Supervision

- Introduce signal-injected (varied mass and luminosity distance) data set to training, along with slightly altered loss function.
- Hyperparameter, *m*, captures desired/**target level of MSE separation** between noise and anomaly reconstruction.
- When desired separation is achieved, normal **MSE loss of noise-only spectrogram is recovered**.

$\mathcal{L}_{noise}$ - MSE loss of noise spectrograms

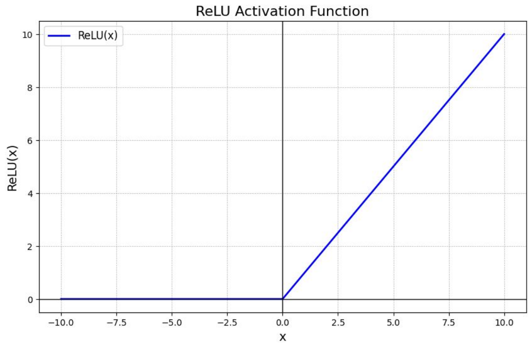
$\mathcal{L}_{anom}$ - MSE loss of anomaly spectrograms

m - Target MSE anom/noise separation

$$\Delta\mathcal{L} = \mathrm{ReLU}(m - (\mathcal{L}_{anom} - \mathcal{L}_{noise}))$$

$$\Delta\mathcal{L} = \begin{cases} m - (\mathcal{L}_{anom} - \mathcal{L}_{noise}), & \text{if } \mathcal{L}_{anom} - \mathcal{L}_{noise} < m, \\ 0, & \text{if } \mathcal{L}_{anom} - \mathcal{L}_{noise} \geq m. \end{cases}$$

$$\boxed{\mathcal{L}_{total} = \mathcal{L}_{MSE}(\hat{x}_{noise}) + \Delta\mathcal{L}}$$



Noise Modelling

- **Overtraining check**: compare MSE distribution of training and testing ET noise spectrograms -> flat ratio implies generalisation to unseen data.

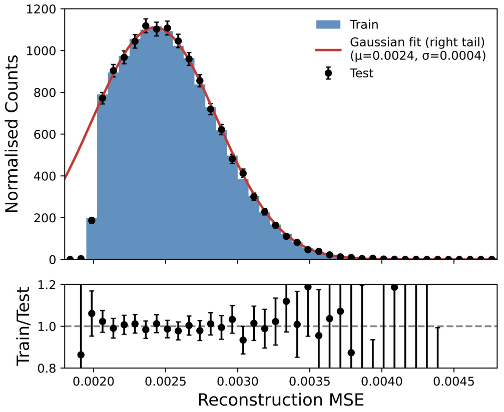
- **Anomaly definition**: chosen threshold value in MSE loss, above which spectrogram is flagged as anomalous.

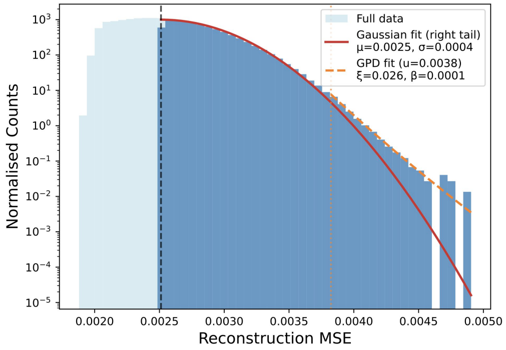
- Define threshold based on statistical **likelihood of noise fluctuation:**

$$\textbf{Threshold} = \boldsymbol{\mu + n\sigma}, \quad (n=3,5)$$

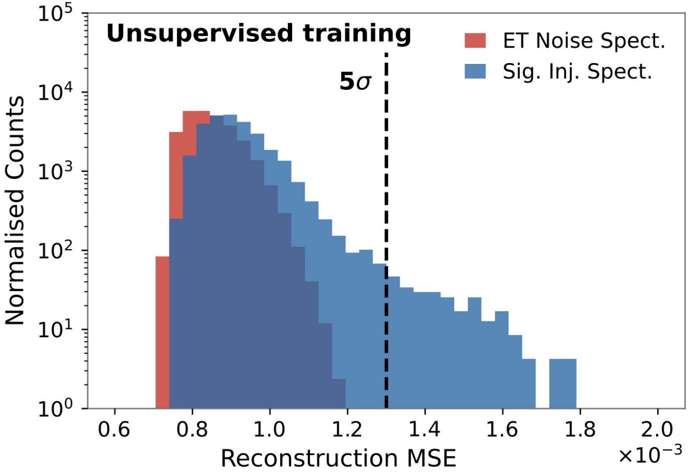
- Assuming 2s spectrograms, 100% duty cycle, 3σ and 5σ equate to FAR:

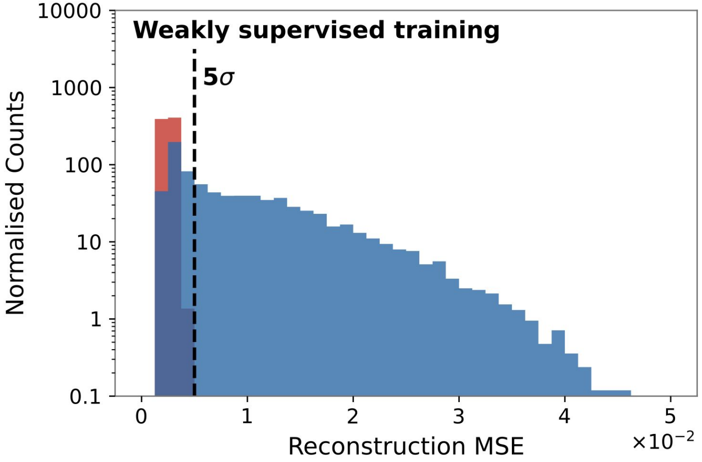
$$\mathbf{3\sigma} \rightarrow p \simeq 1.35\text{x}10^{-3} \rightarrow \text{FAR} \simeq 2.1\text{x}10^{4} \text{ events/year}$$
$$\mathbf{5\sigma} \rightarrow p \simeq 2.87\text{x}10^{-7} \rightarrow \text{FAR} \simeq 4.5 \text{ events/year}$$





MDC - MSE



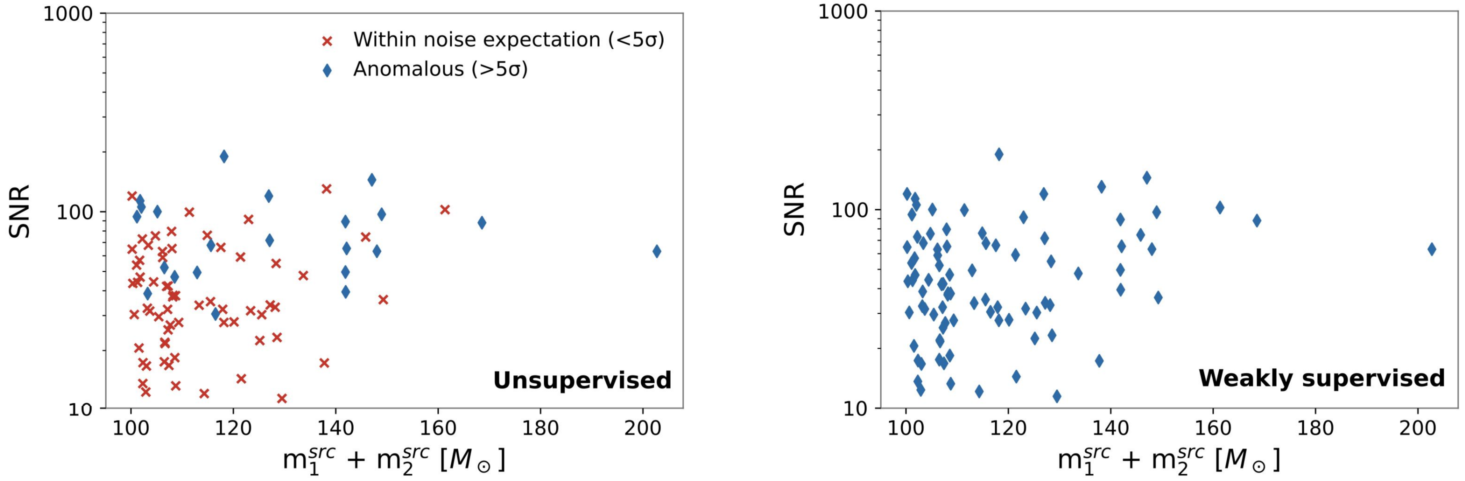


- Weak supervision shows large improvement in separating MSE between noise and injected spectrograms.

- Effect of hyperparameter *m* (=5x10^{-2}) quite clear in distribution of MSE after weakly supervised training.

MDC - IMBH mergers

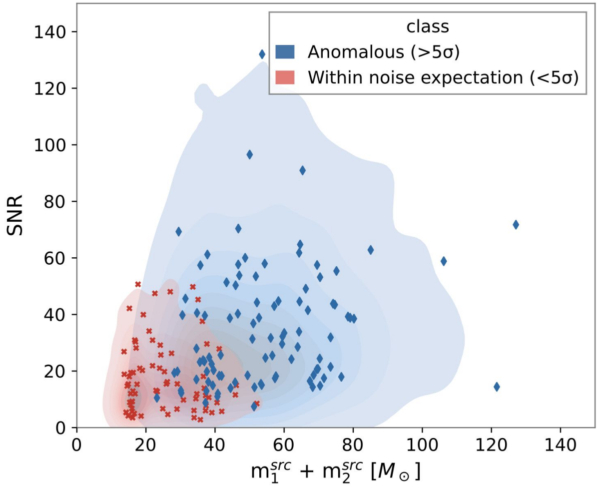
MDC mergers of M ≥ 100

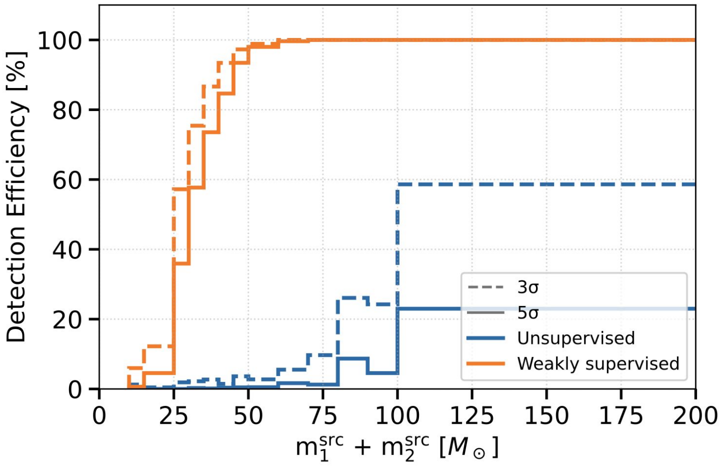


- Events falling below 5σ shown as red crosses, above with blue diamond.
- Weak supervision increases efficiency from 23% to 100% of mergers which involve or form IMBH

MDC - other masses

Subsample of full MDC1 merger injections





- Introduction of weak supervision provides huge increase in achievable efficiency for target burst signals

- Model shows strong ability to generalise to other merger signals outside of initial target

Conclusion

- Large frequency range and sensitivity of Einstein Telescope presents substantial increase in computational overheads - ML could offer a solution with **computational work done in training stage**.

- **CNN autoencoder** for anomaly detection shows ability to recover (up to 100%) **IMBH burst signals** from ET noise.

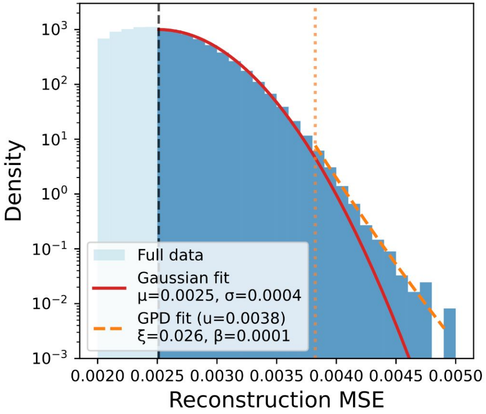
- Inclusion of weak supervision to training greatly improves efficiency and **strong ability to generalise** to signals outside initial target range.

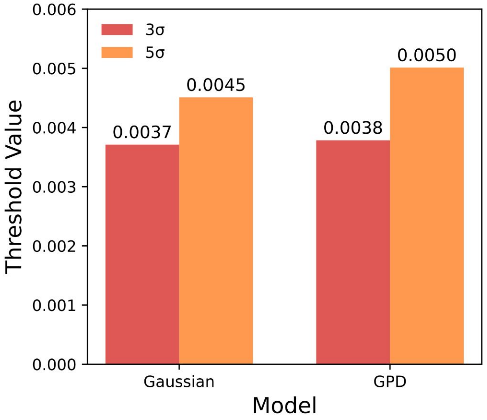
- **Next steps:**
  - Inclusion of detector response.
  - Injected glitches alongside mergers
  - Study possibility of analysing multiple channels

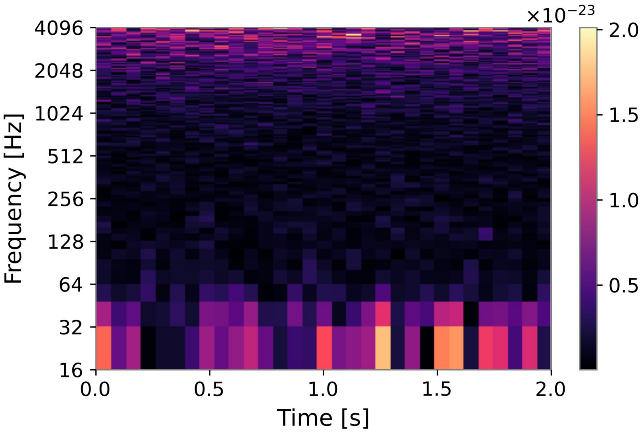
Work already started!

Paper draft submitted - feedback very much appreciated!

Backup

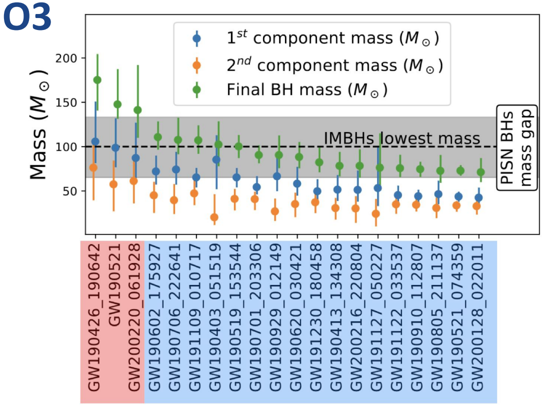


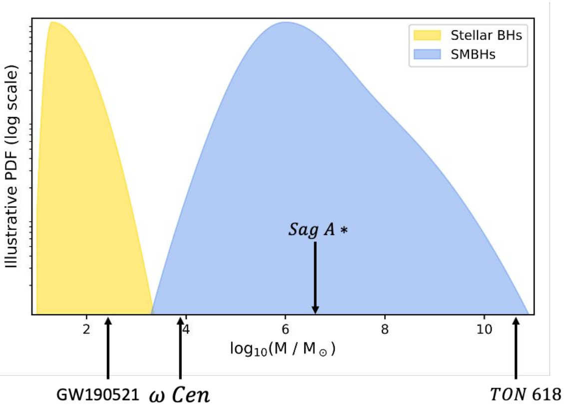


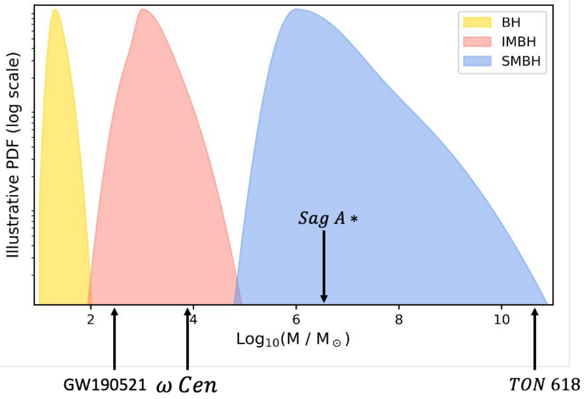


ET noise with high reconstruction MSE

O3


1st component mass ($M_\odot$)
2nd component mass ($M_\odot$)
Final BH mass ($M_\odot$)
Mass ($M_\odot$)
IMBHs lowest mass
PISN BHs mass gap
GW190426_190642
GW190521
GW200220_061928
GW190602_175927
GW190706_222641
GW191109_010717
GW190403_051519
GW190519_153544
GW190701_203306
GW190929_012149
GW190620_030421
GW191230_180458
GW190413_134308
GW200216_220804
GW191127_050227
GW191122_033537
GW190910_112807
GW190805_211137
GW190521_074359
GW200128_022011


Stellar BHs
SMBHs
Illustrative PDF (log scale)
$Sag\ A*$
$\log_{10}(M / M_\odot)$
GW190521
$\omega\ Cen$
$TON\ 618$


BH
IMBH
SMBH
Illustrative PDF (log scale)
$Sag\ A*$
$\mathrm{Log}_{10}(M / M_\odot)$
GW190521
$\omega\ Cen$
$TON\ 618$