Towards a CNN-based Anomaly Detection Pipeline at the Einstein Telescope

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

MBI

Introduction

- Einstein Telescope promises increase in frequency range, **higher sensitivity to many more mergers** -> higher complexity (overlapping signals)

- With such reach, techniques like matched filtering come 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 (<2s) - we study the use of **convolutional autoencoders** to extract from noise.
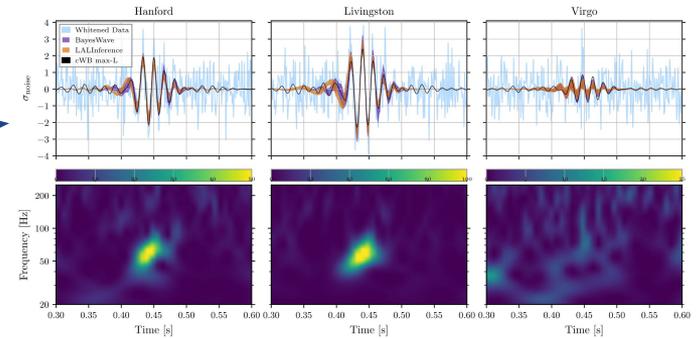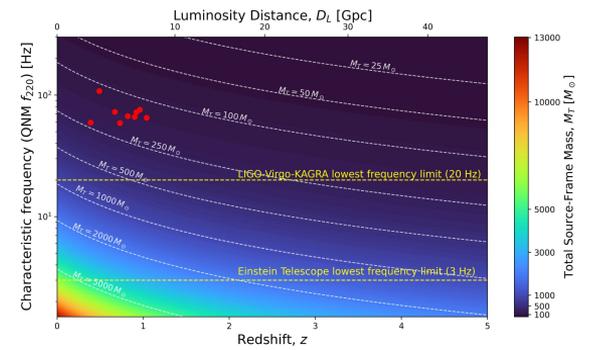
GW190521: **2009.01075**

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.
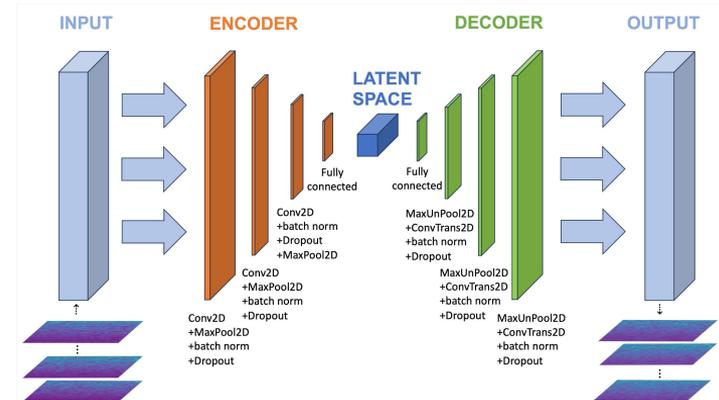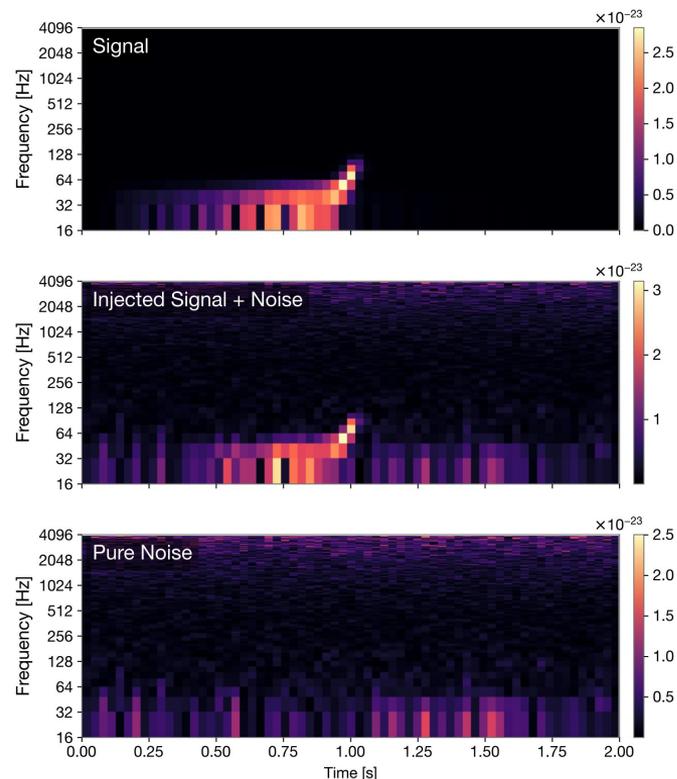
Spectrograms

- 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 STFT

- Greyscale images of **256 (frequency) x 31 (time) pixels** - normalised to [0,1] before training.

- Waveforms generated using IMRPhenomPV2 - 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

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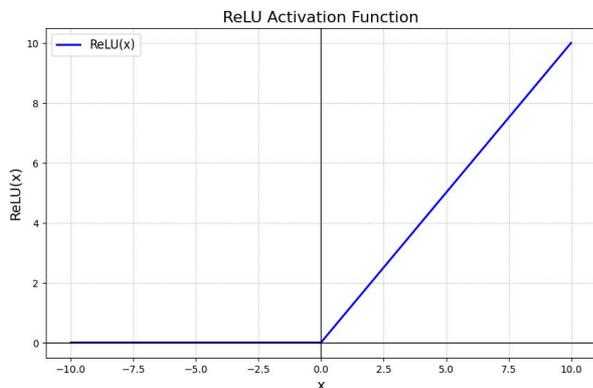
$$\Delta\mathcal{L} = \begin{cases} m - (\mathcal{L}_{anom} - \mathcal{L}_{noise}), & \text{if } \mathcal{L}_{anom} - \mathcal{L}_{noise} < m, \\ 0, & \text{if } \mathcal{L}_{anom} - \mathcal{L}_{noise} \geq m. \end{cases}$$

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


ReLU Activation Function

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 error, above which spectrogram is flagged as anomalous.

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

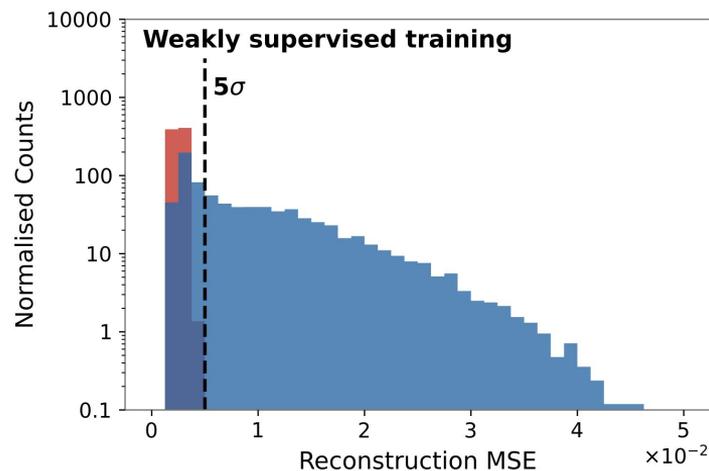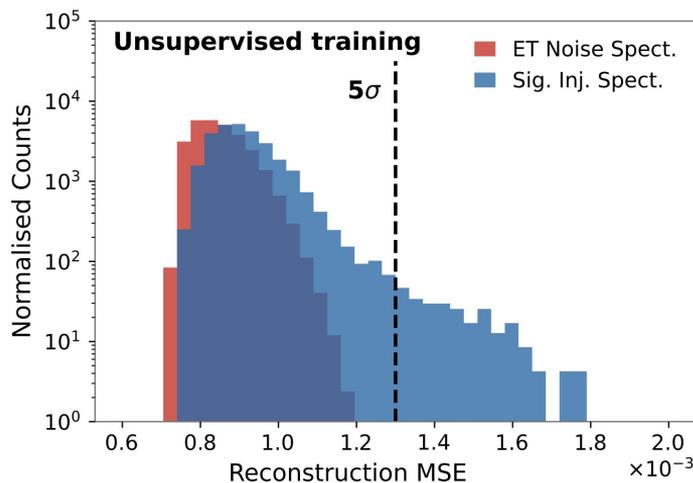$$\text{Threshold} = \mu + n\sigma, \quad (n=3,5)$$

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

$$3\sigma \rightarrow p \approx 1.35 \times 10^{-3} \rightarrow FAR \approx 2.1 \times 10^{4} \text{ events/year}$$
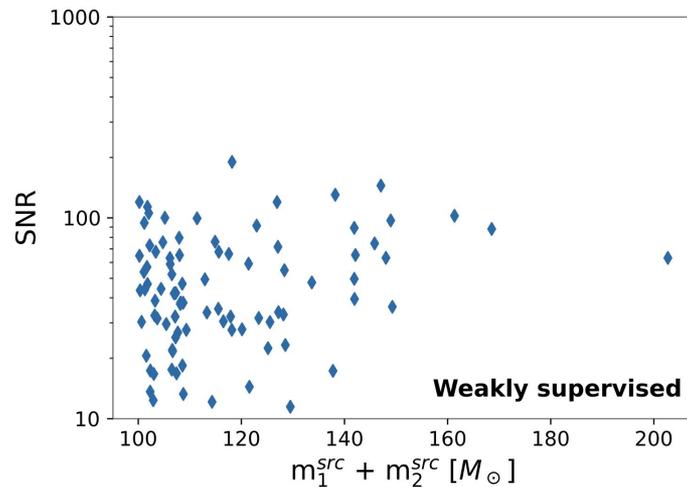$$5\sigma \rightarrow p \approx 2.87 \times 10^{-7} \rightarrow FAR \approx 4.5 \text{ events/year}$$
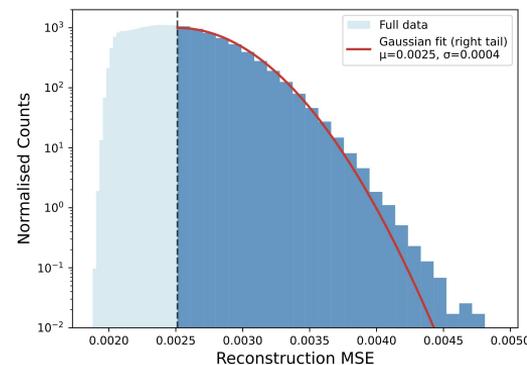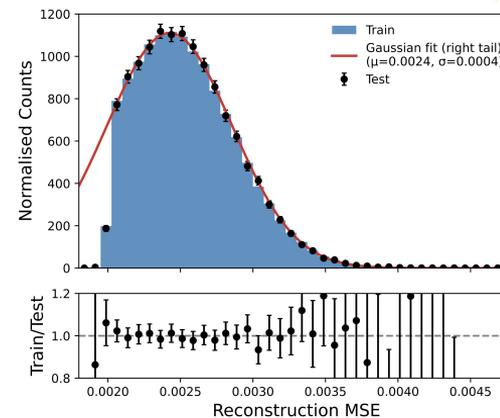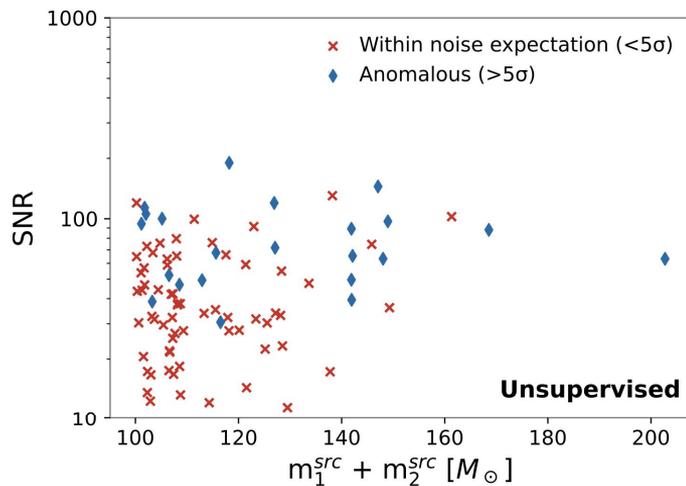
2506.12237 - MDC contains 6119 BBH merger samples

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

- Effect of hyperparameter m (=5×10^{-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

Subsample of full MDC1 merger injections
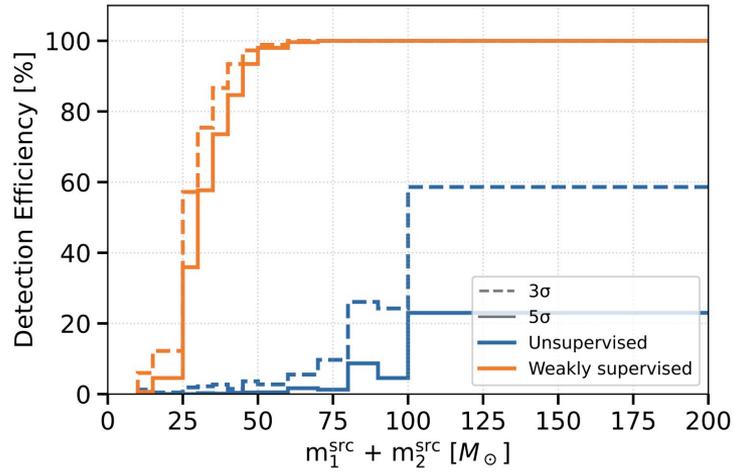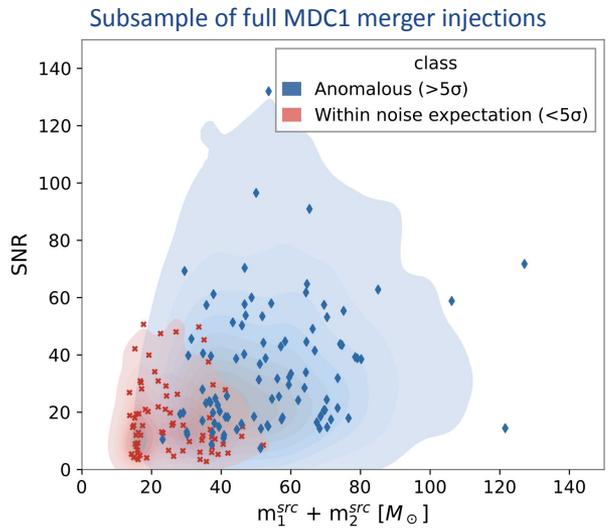
- Introduction of weak supervision provides huge increase in achievable efficiency for target burst signals

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

Next Steps: Glitches

- So far Signal vs background only - glitches pose clear problem especially for burst signals.

- Exploring various approaches:

    - **Multi-step classifier** - CNN to detect anomalies -> further NN model for classification.

    - **Variational Autoencoder** - class-structured latent space representation for noise, glitch, signal classification.
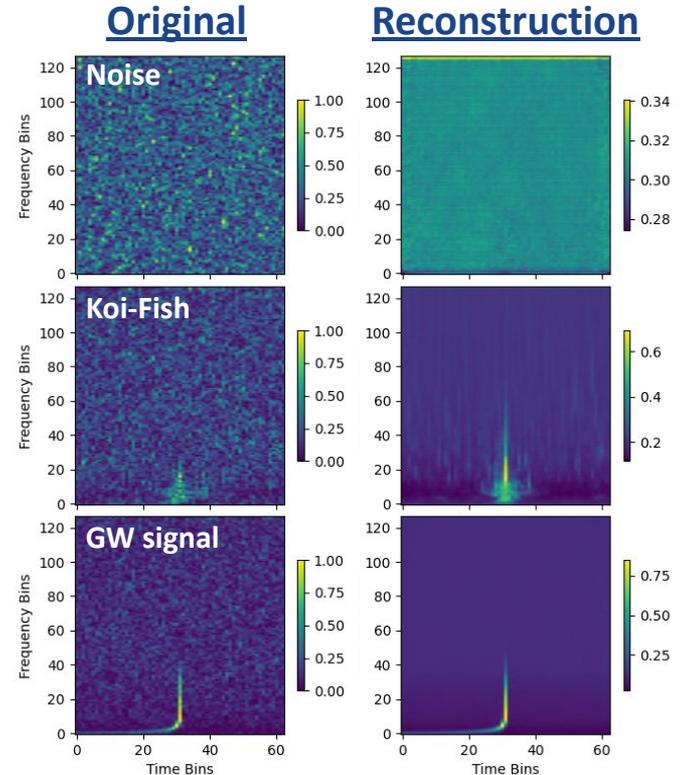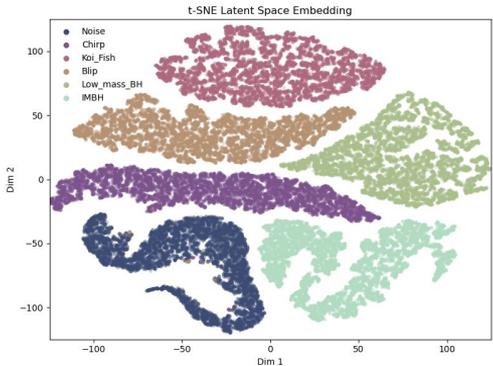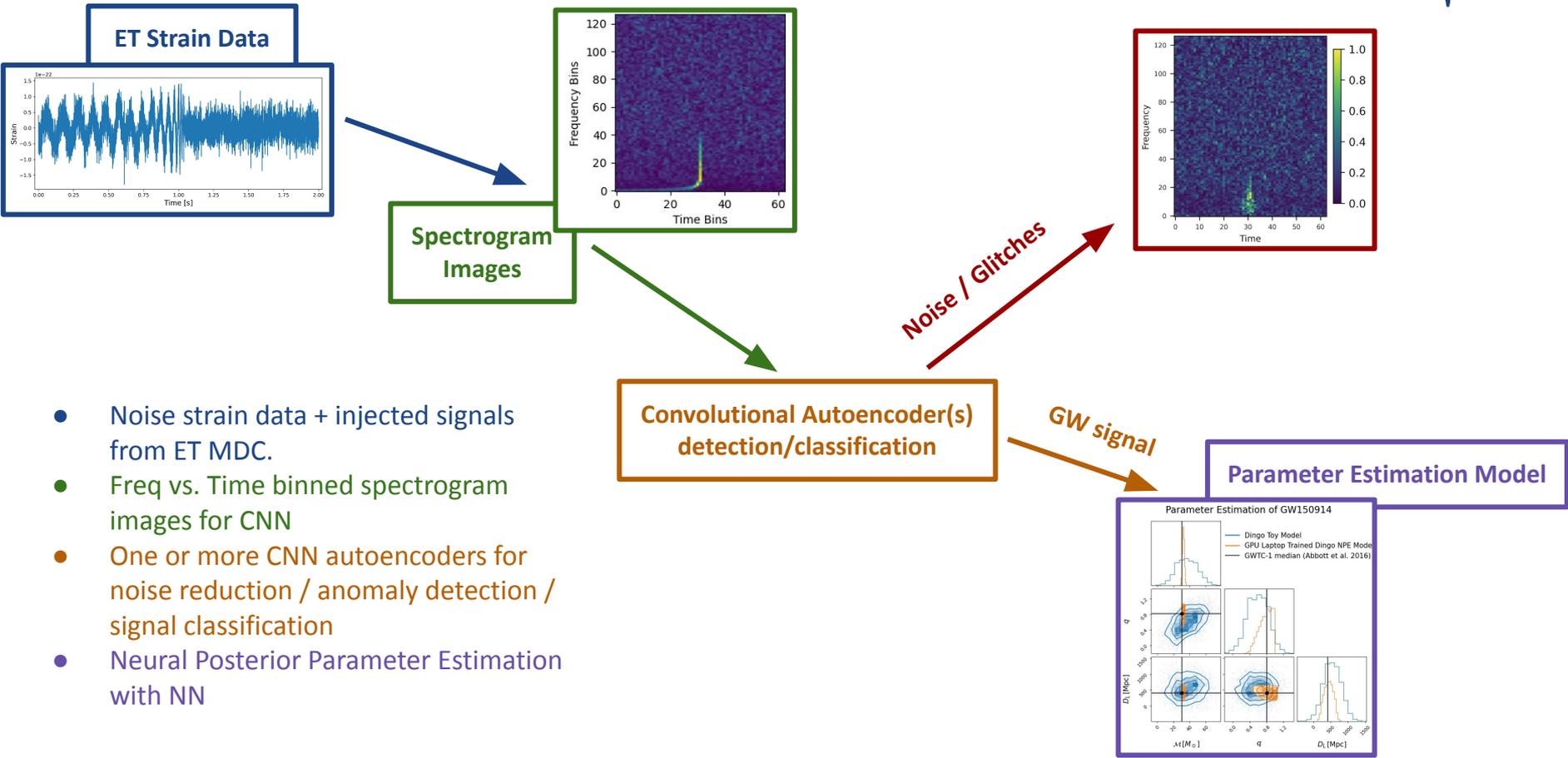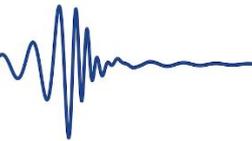

Preliminary work using samples generated with GlitchGen (GAN based approach) - **2207.00207**

Conclusion

- Einstein Telescope presents substantial increase event rates - deep learning offers a solution for **fast detection/classification 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 - weak supervision shows ability to **generalise to signals outside initial target range.**

- Initial anomaly detection study **published in Physics Letters B 2511.13154**

- Current/upcoming work:

  - Including glitches - Explore methods of **signal/glitch classification** (variational AE, multi-step classifiers)

  - **Parameter estimation** - Neural posterior estimation network integrated into detection pipeline for rapid parameter estimation of signals detected by CNN.

  - Expand models to **multiple input streams** - 2 or 3 depending on ET design.

  - Longer term - **Validation of methods with LVK (O3) data**

Backup

Pipeline Overview

ET Strain Data

Spectrogram Images

Noise / Glitches

Convolutional Autoencoder(s) detection/classification

GW signal

Parameter Estimation Model

- Noise strain data + injected signals from ET MDC.
- Freq vs. Time binned spectrogram images for CNN
- One or more CNN autoencoders for noise reduction / anomaly detection / signal classification
- Neural Posterior Parameter Estimation with NN