APC, CNRS, FRANCE

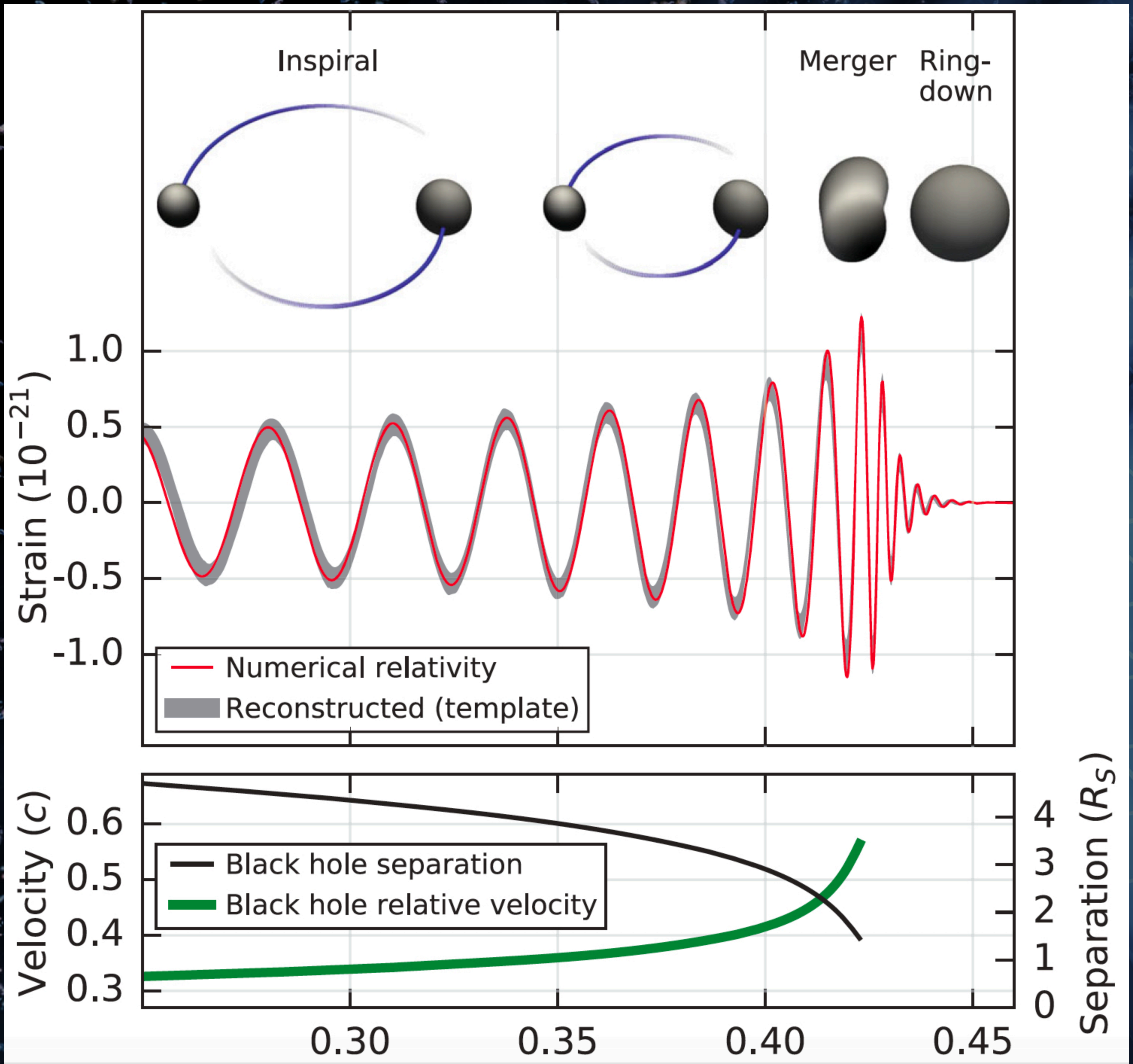
Université de Paris

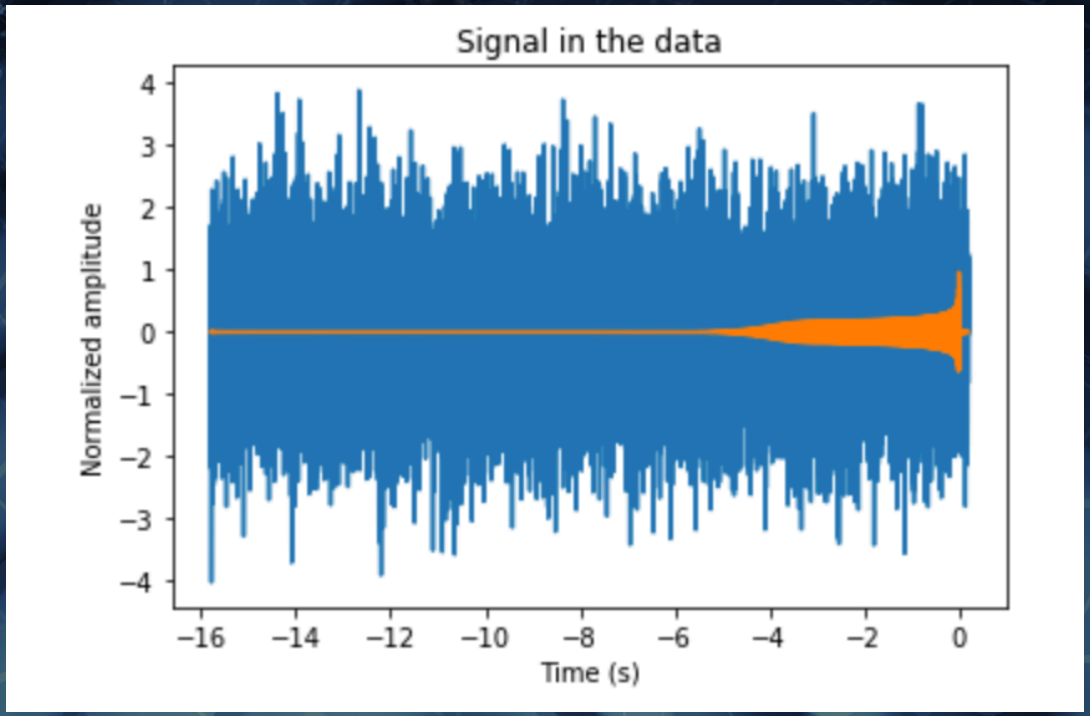
Machine Learning for gravitational waves at APC

P. Bacon, M. Bejger and E. Chassande-Mottin, A. Trovato*

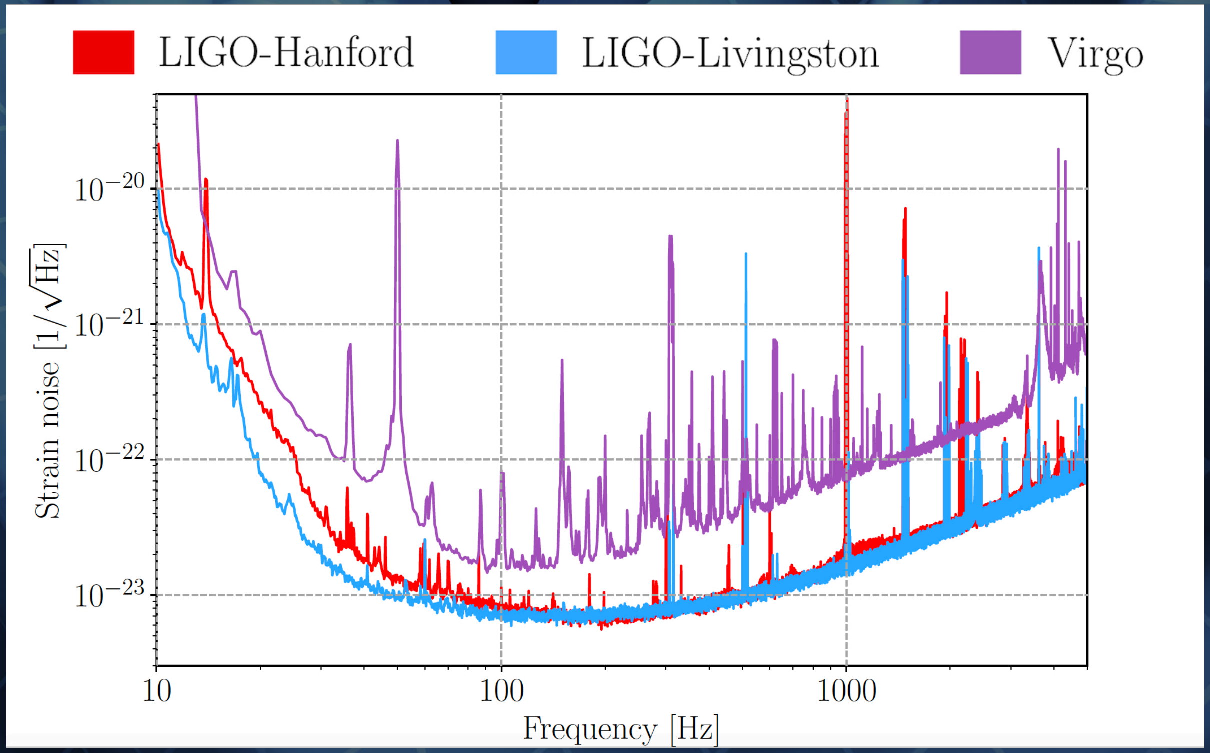
*APC, CNRS/IN2P3, Univ. Paris Diderot

Gravitational waves detection problem



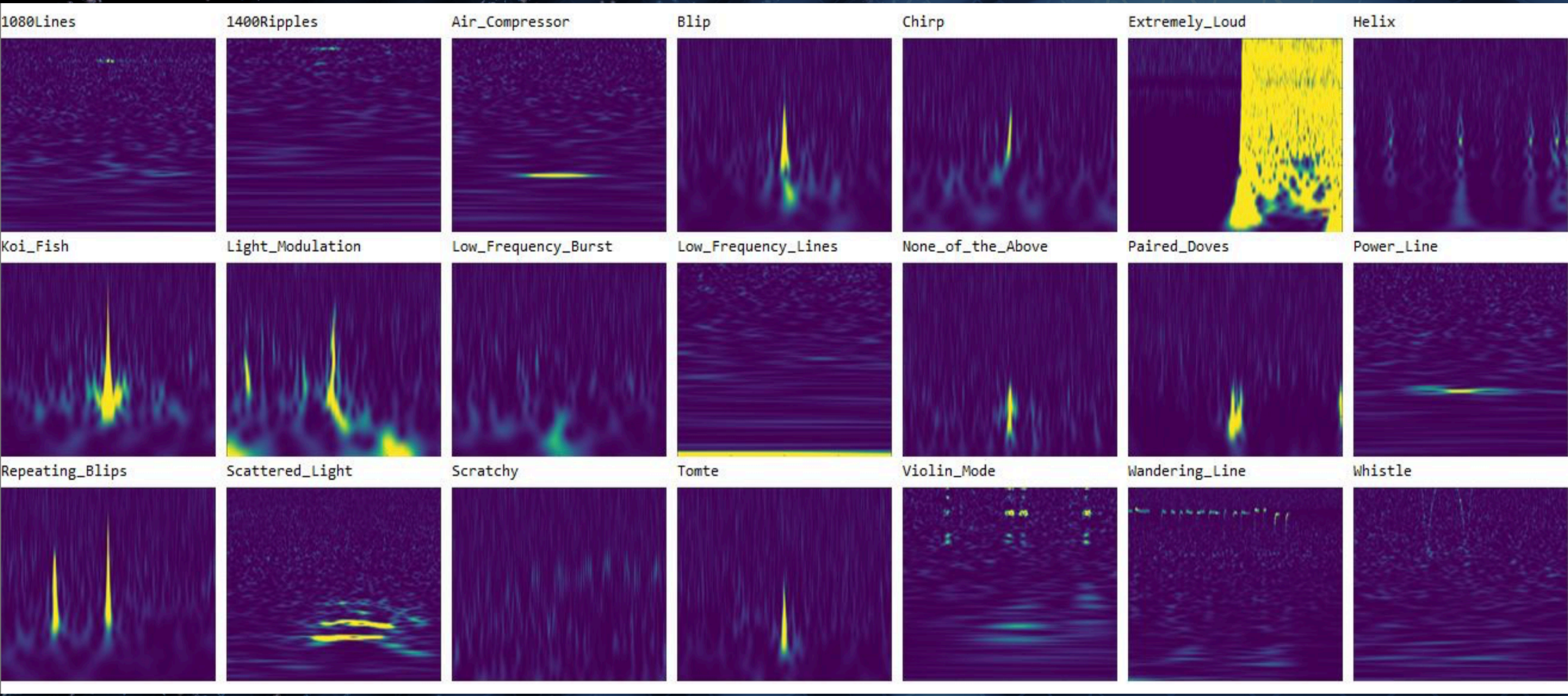


Complex background: non-Gaussian non-stationary



Glitches zoo

- Gravity Spy dataset —> glitches represented as spectrograms
- Only high SNR (SNR>7.5)



Projects presented

1. A classifier based on a Convolutional Neural Network to distinguish signals vs glitches vs gaussian background
   - Single detector application
   - Input: time-series / Output: label

2. A convolutional autoencoder used to denoise the gravitational wave signals
   - Input: time-series / output: time-series

Project 1

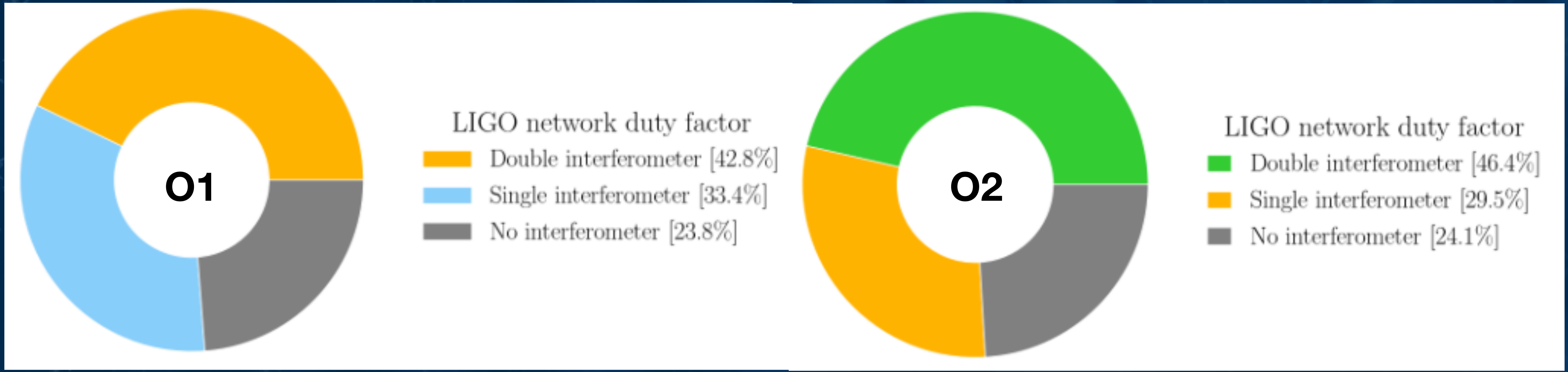
General ideas

- Study, identify and reduce the transient noise present in the gravitational wave detectors through deep learning techniques
  - ✓ Huge amount of noisy data
  - ✓ Impact data quality
  - ✓ mimic the gravitational wave signal
  - ✓ Complex population —> No statistical model
  - ✓ Task for machine learning algorithms!
  - ✓ Interesting topic: other projects in LIGO/Virgo deal with it

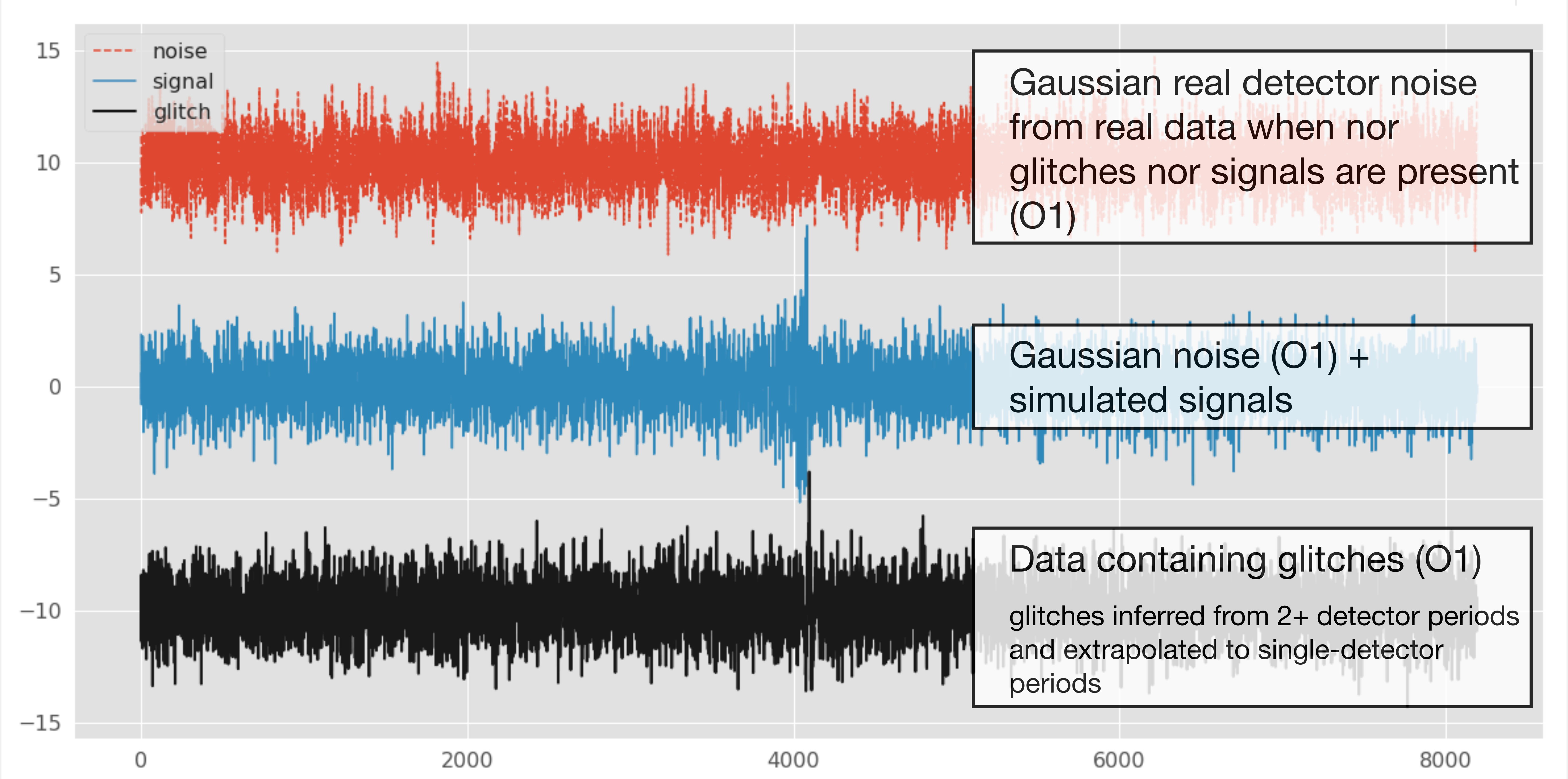
- Final goal: analyse single-detector data
  - ✓ Potentially interesting detector time: since O2 gstlal provides triggers for a single-detector case —> only Binary Neutron Stars (BNS) until now

Single-detector time

- Current pipelines: signal has to appear in coincidence in two or more detectors
  - ✓ distinguish true astrophysical signals from the transient noise
  - ✓ highly reduces the number of false positives allowing to detect gravitational waves with very high statistical confidence.
- Single-detector time could be exploited better
  - ✓ 2.7 months in O1+O2 => could contain 3 events
  - ✓ In O3 about 16% at the moment

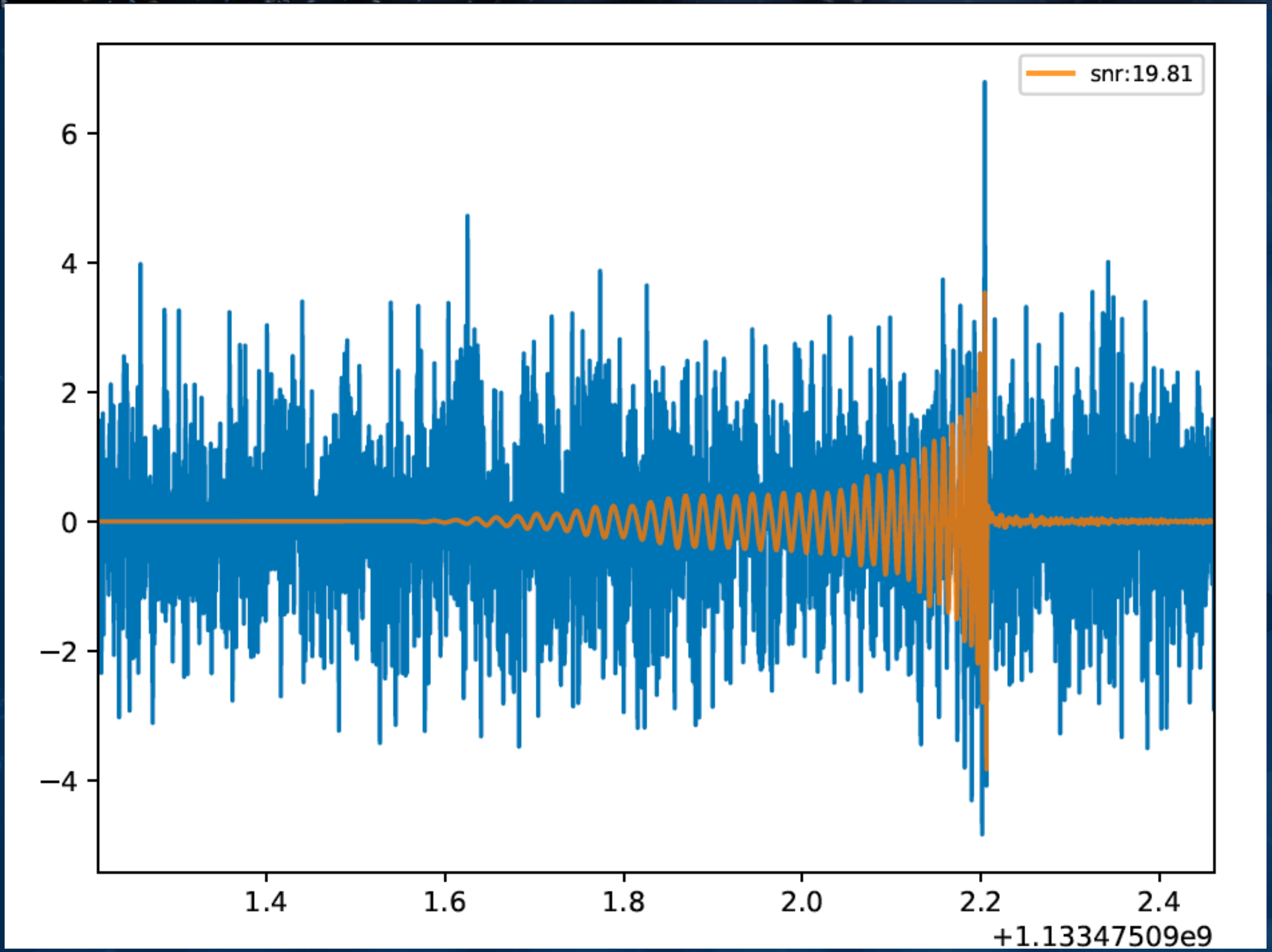


Training data: 3 classes



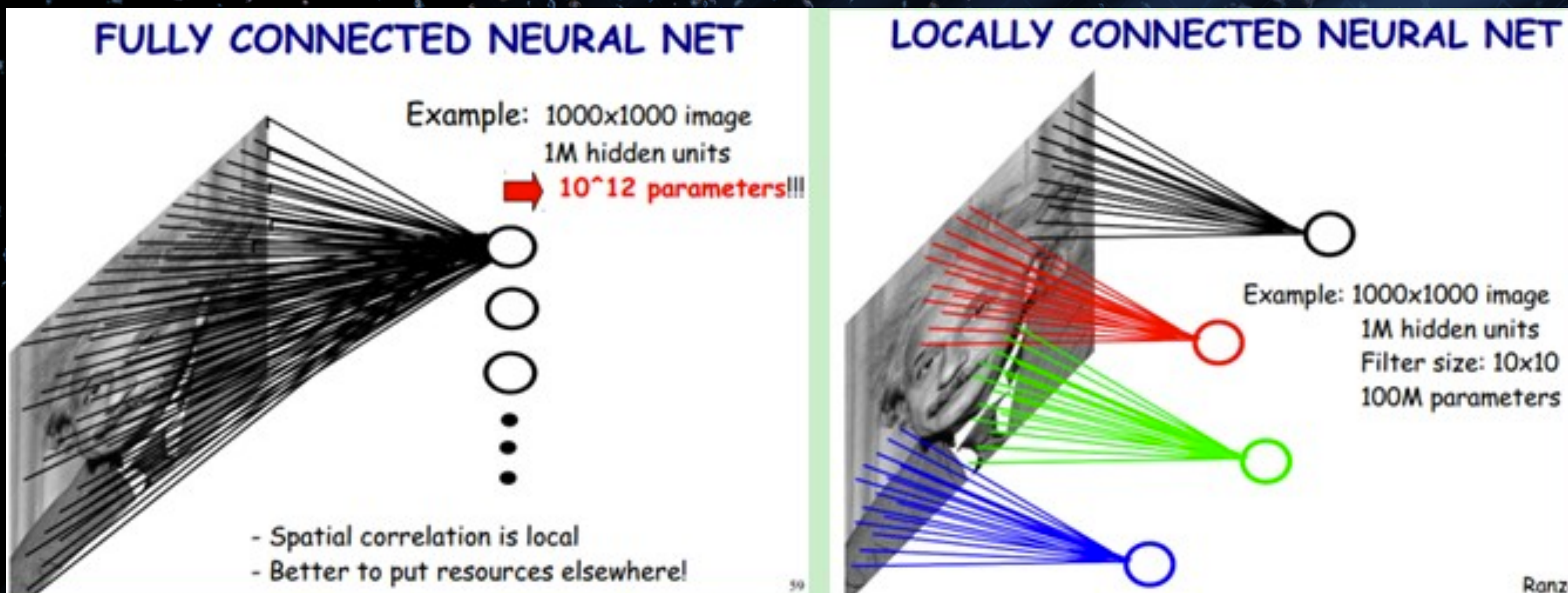
Glitches and "clean" noise data samples from the one month of LIGO O1 run (downsampled to 2048 Hz, duration: 4s —> 8192 points), whitened by the amplitude spectral density of the noise.

Simulated signals



- Randomly selected binary black holes' system merger waveforms: m_1, $m_2 \in (8, 16)\ M_\odot$, signal-to-noise $\in (15, 45)$, added to ”clean” noise samples, whitened.

1D Convolutional Neural network

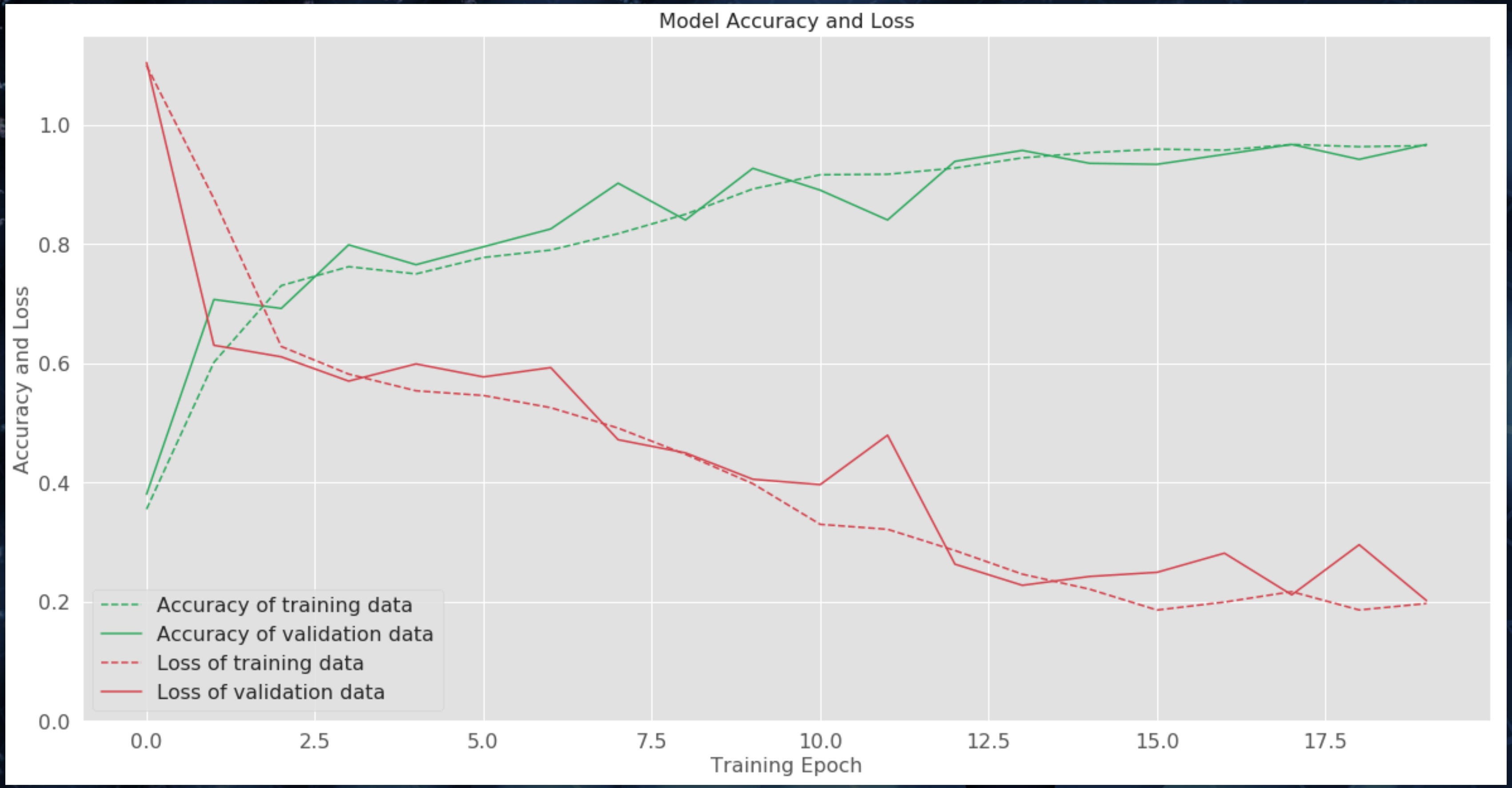


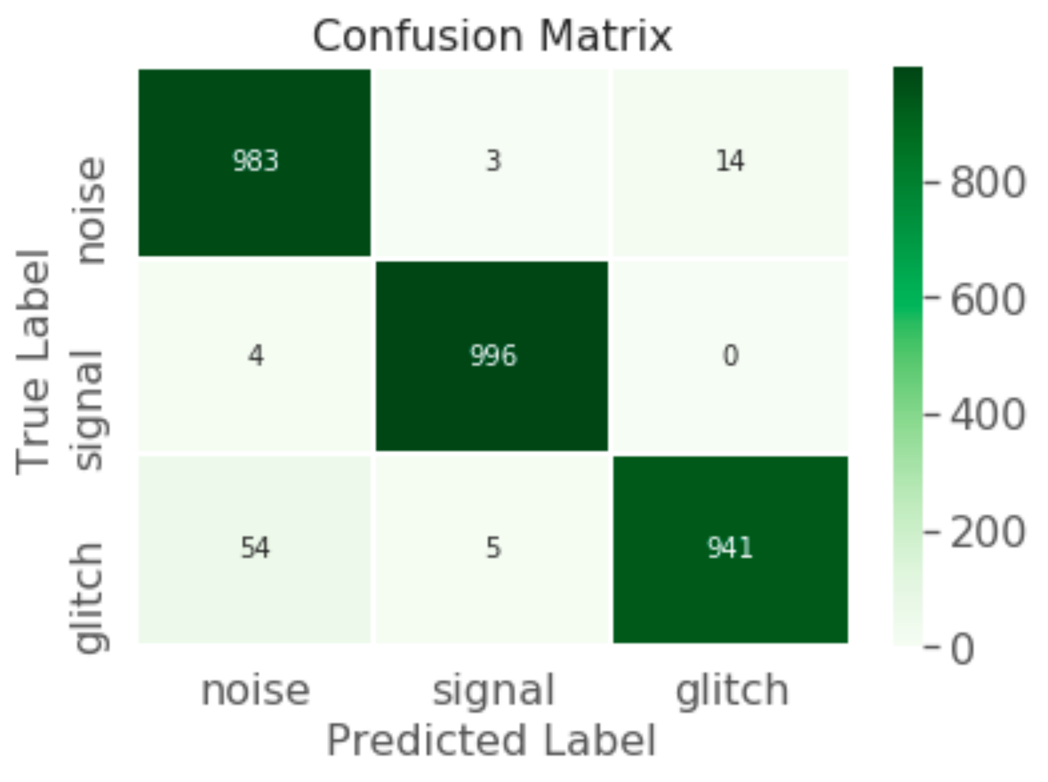
- ✓ Suited for data with a known grid-like topology.
- ✓ In a convolutional layer neurons receive input from only a restricted subarea of the previous layer

Architecture of the network used

```
Layer (type)                 Output Shape              Param #
=================================================================
reshape_1 (Reshape)          (None, 8192, 1)           0
_________________________________________________________________
conv1d_1 (Conv1D)            (None, 8188, 500)         3000
_________________________________________________________________
max_pooling1d_1 (MaxPooling1 (None, 2729, 500)         0
_________________________________________________________________
conv1d_2 (Conv1D)            (None, 2725, 250)         625250
_________________________________________________________________
conv1d_3 (Conv1D)            (None, 2721, 250)         312750
_________________________________________________________________
max_pooling1d_2 (MaxPooling1 (None, 907, 250)          0
_________________________________________________________________
conv1d_4 (Conv1D)            (None, 903, 150)          187650
_________________________________________________________________
global_average_pooling1d_1 ( (None, 150)               0
_________________________________________________________________
dropout_1 (Dropout)          (None, 150)               0
_________________________________________________________________
dense_1 (Dense)              (None, 3)                 453
=================================================================
Total params: 1,129,103
Trainable params: 1,129,103
Non-trainable params: 0
```

Classification results

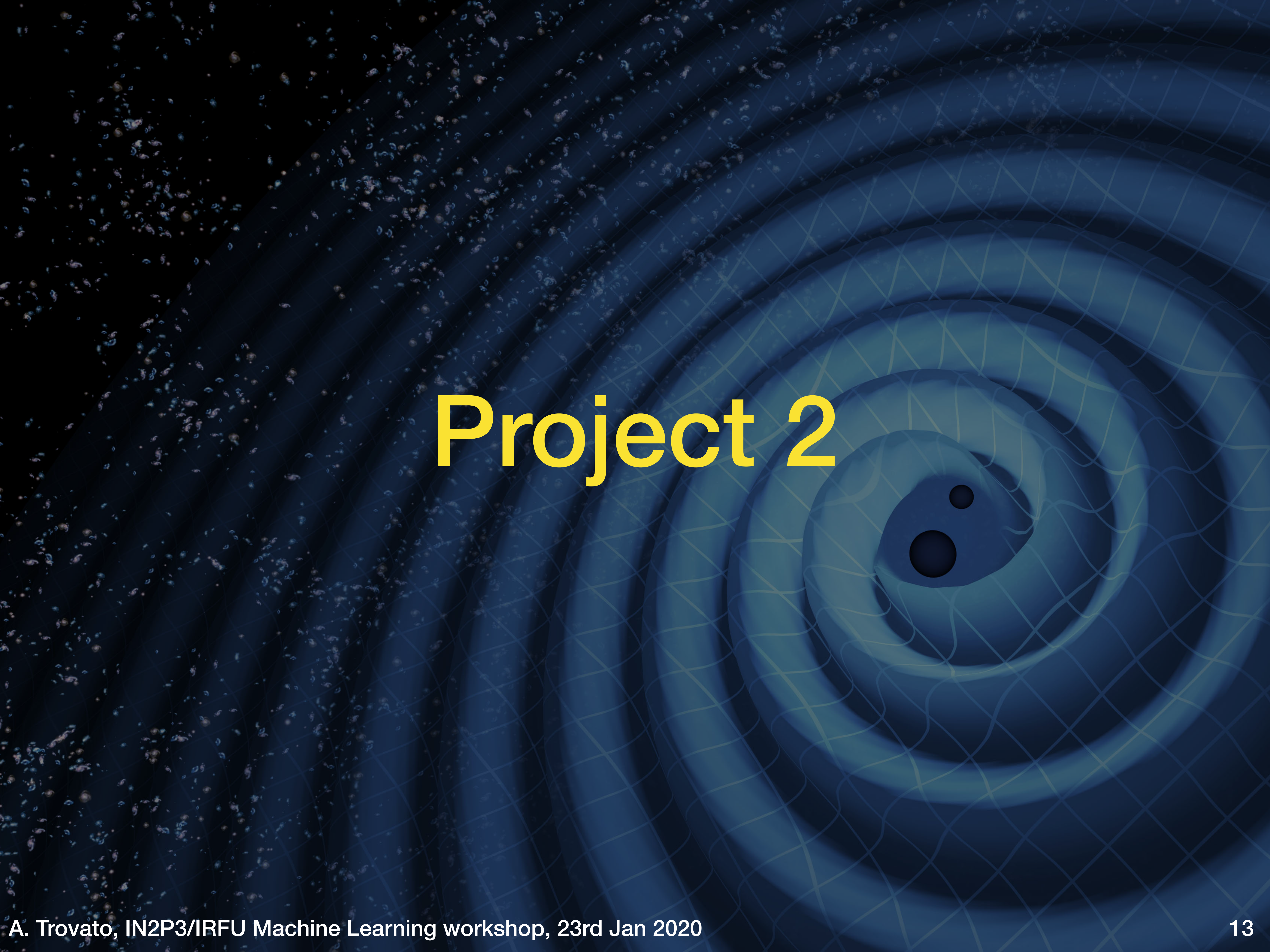




- Training data: 1000 instances, 3 classes
- Training time: '10 minutes for 20 epochs @Nvidia Tesla K40XL
- Accuracy on test data: 0.97

Conclusion 1

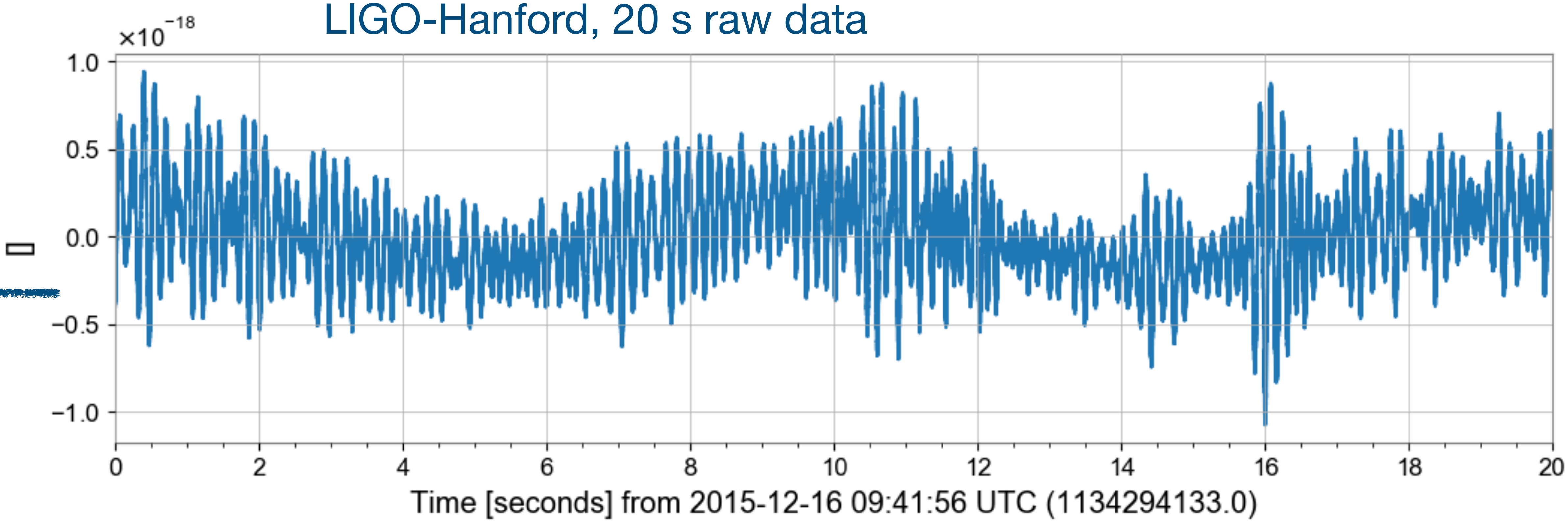
- Proof-of-concept single-detector low-latency classifier implemented (gaussian noise vs gaussian noise+glitch vs gaussian noise+signals)
  - ✓ Paper in preparation
- Extension of the training data set:
  - ✓ environmental channels besides time-series
  - ✓ specific classification for glitches (e.g. using labeled data from Gravity Spy)
- Different types of networks:
  - ✓ Recurrent Neural Networks (RNN), Long-Short Term Memory (LSTM) for classification
  - ✓ Addition if dilated convolutions to the CNN
  - ✓ Explore bayesian neural networks



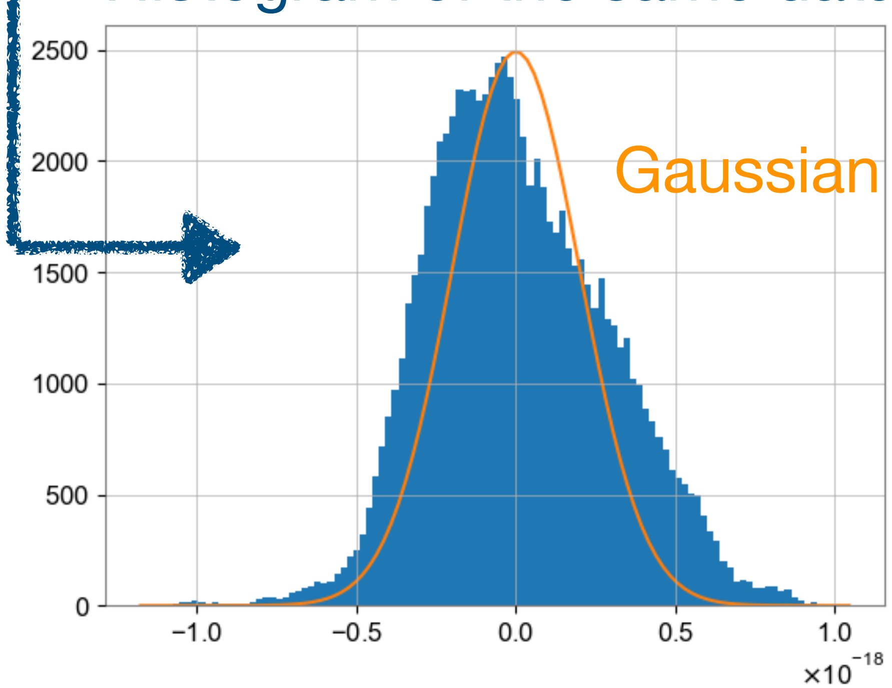
Project 2

Non-Gaussian data

LIGO-Hanford, 20 s raw data



Histogram of the same data



The data are far from being Gaussian and stationary:

- Standard match-filter approach assume Gaussian data

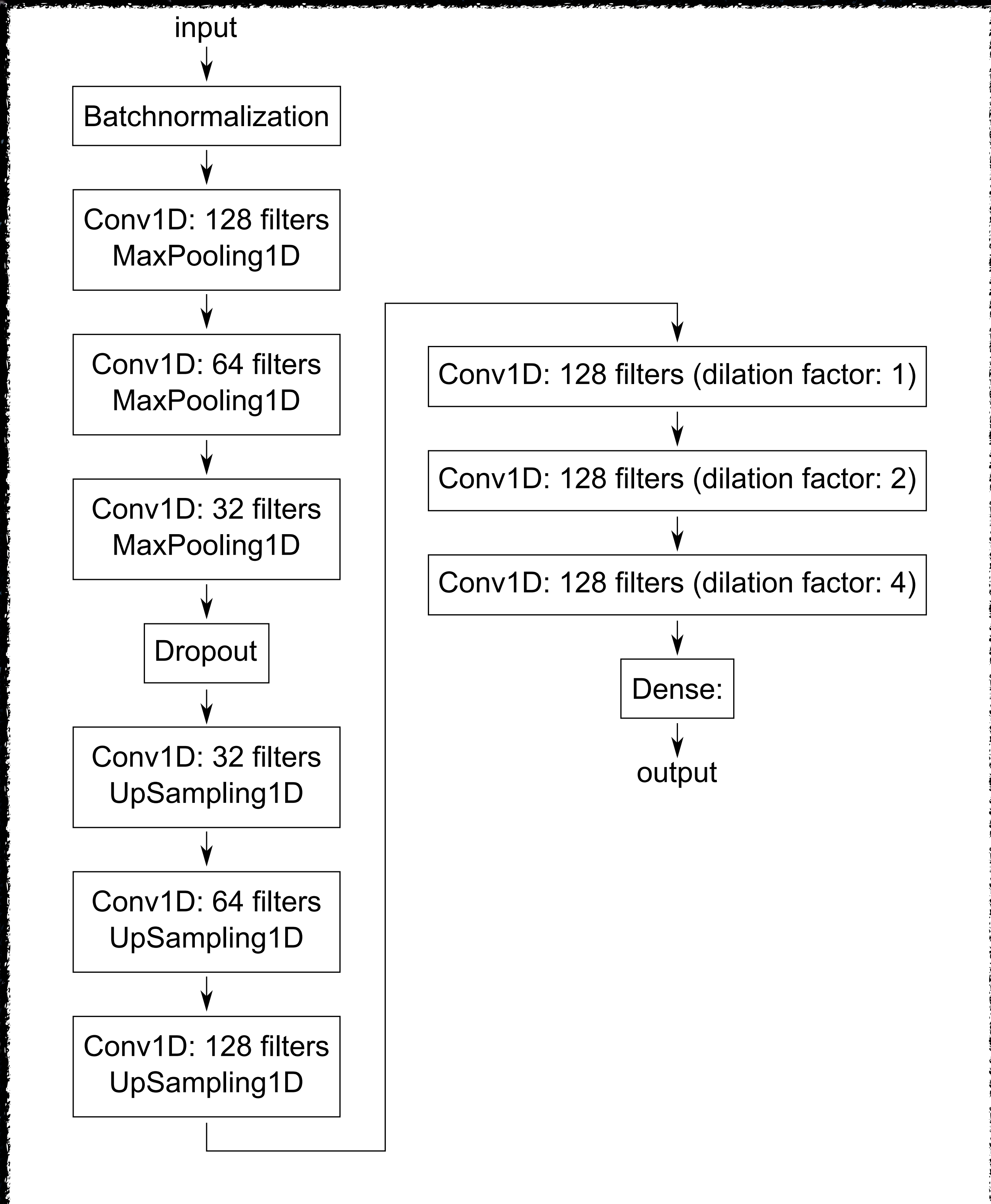
Denoising autoencoder based on CNN

- **Denoising**: model that take noisy signals and return clean signals

- **Autoencoder**: learn a representation of input data in an unsupervised way
  - ✓ bottleneck-shaped -> encoder + decoder
  - ✓ Sparsity -> primordial when dealing with noise

- **Convolutional Neural Networks** are used as encoder and decoder
  - ✓ Less parameters to train than more complex networks (e.g. Recurrent Neural Networks)

Model & dataset



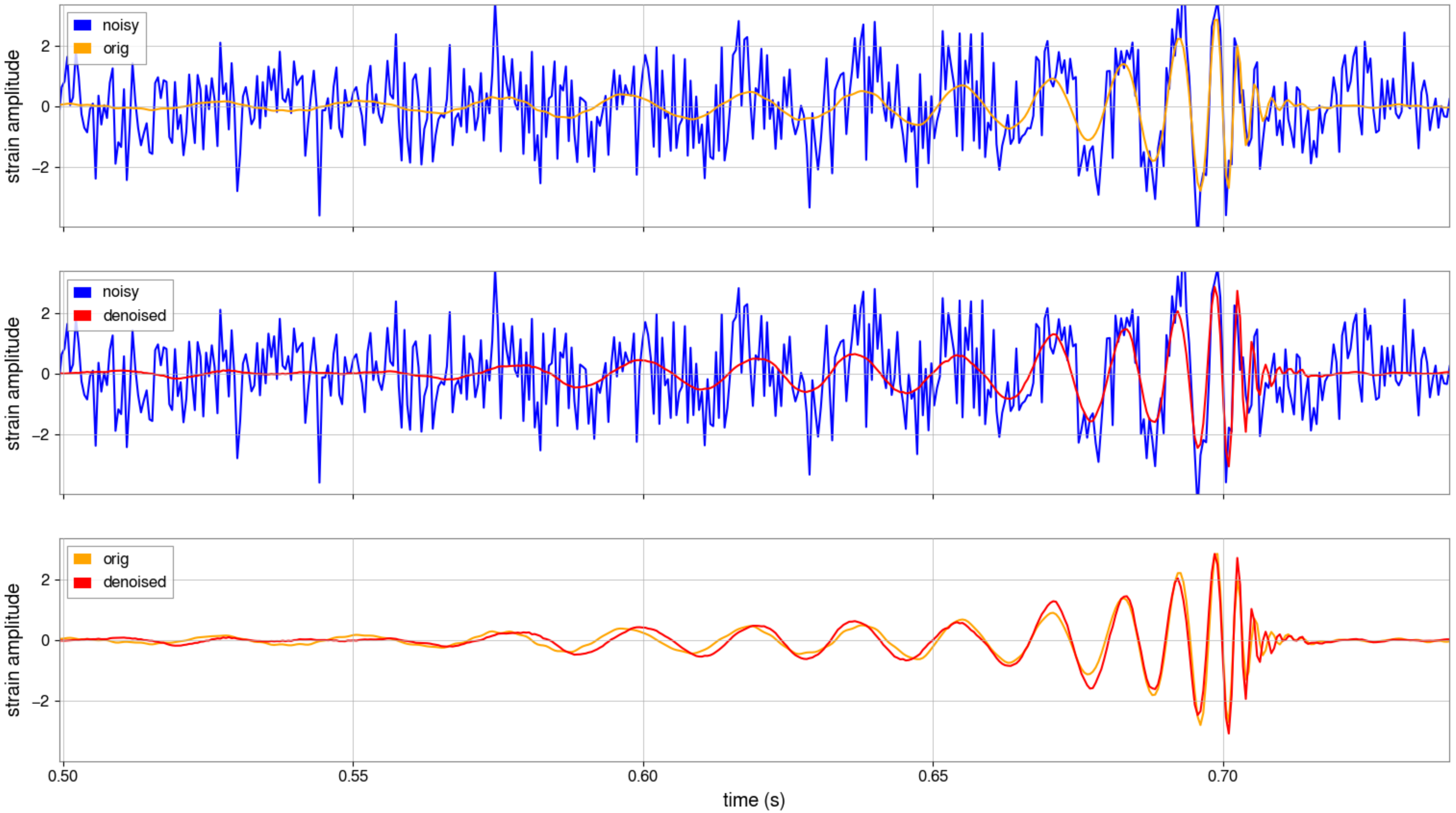
- loss function: mean squared error (MSE)
- activation function (every layer): ReLU
- optimizer: Adam

Dataset

- sampling rate: 2kHz
- <u>injections</u>: SEOBNRv4 GW signals with m1, m2 ∈ [7, 20] $M_{\odot}$, f_{low} = 30 Hz, signal-to-noise ratio (SNR) in [5, 50].
- <u>Input</u>: GW injected signals + real O1 data (away from known glitches and GW signals)
- <u>Expected output</u>: GW injected signals

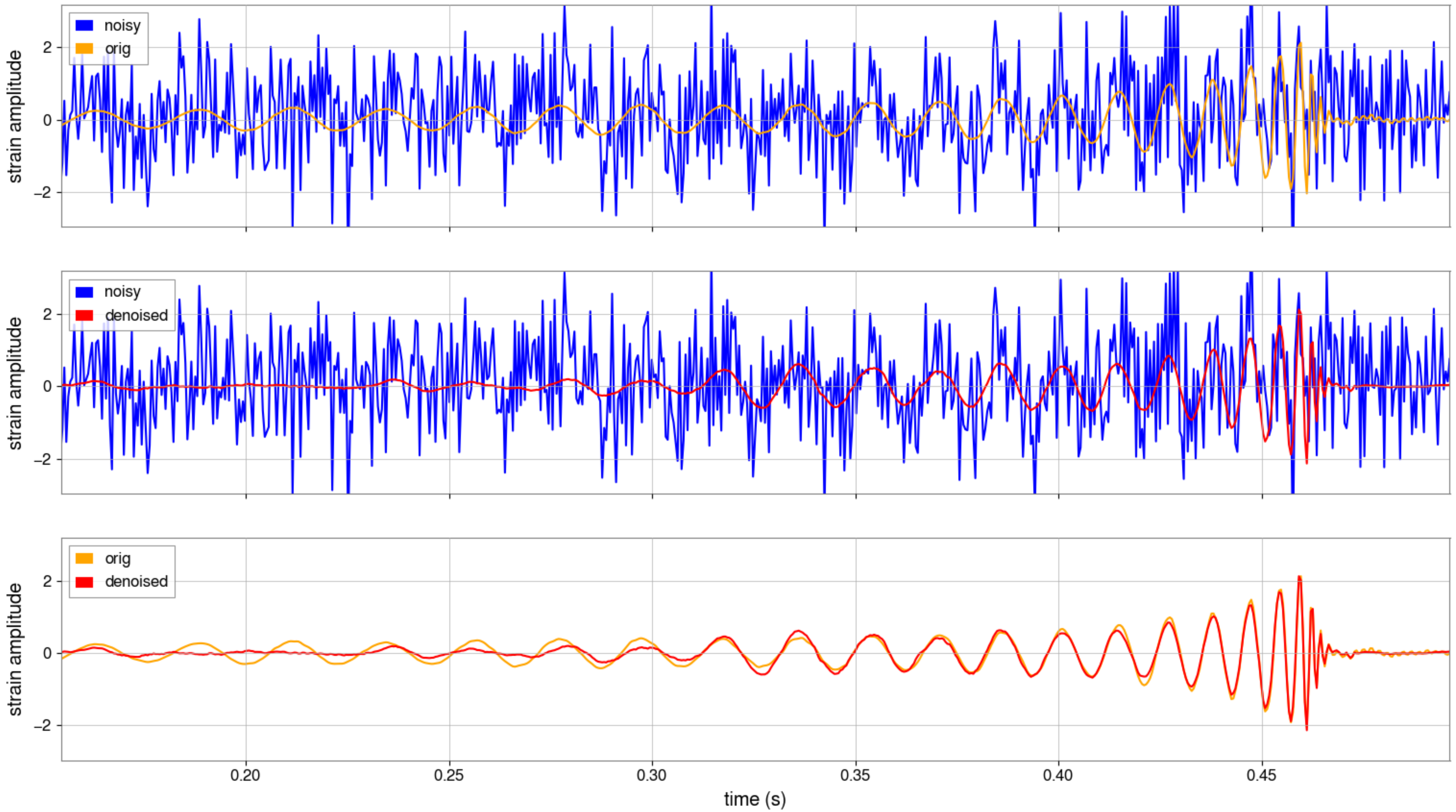
Result on a GW150914-like event

network SNR=25 / Overlap=0.91



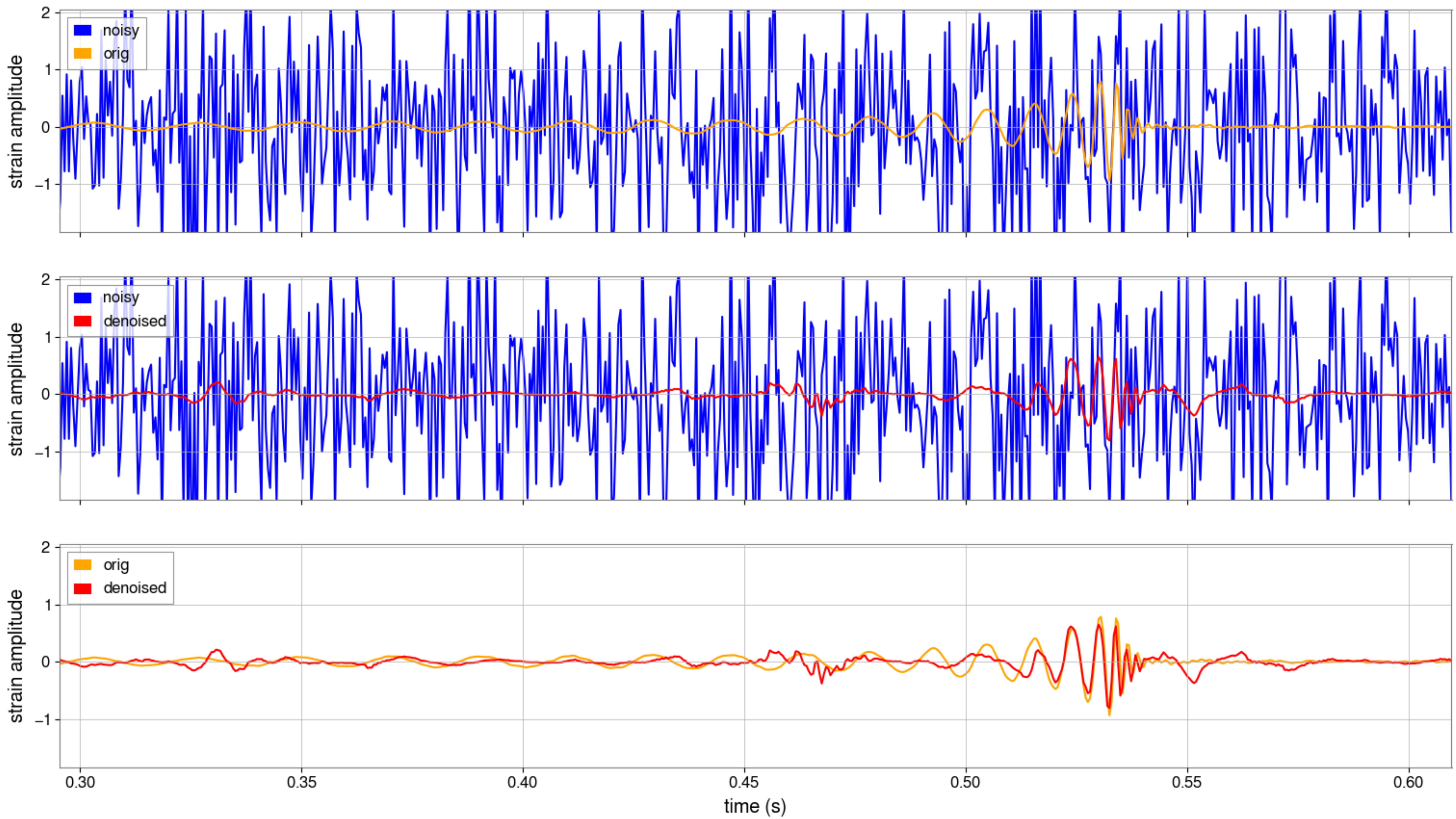
Result on a GW151226-like event

network SNR=11 / Overlap=0.98

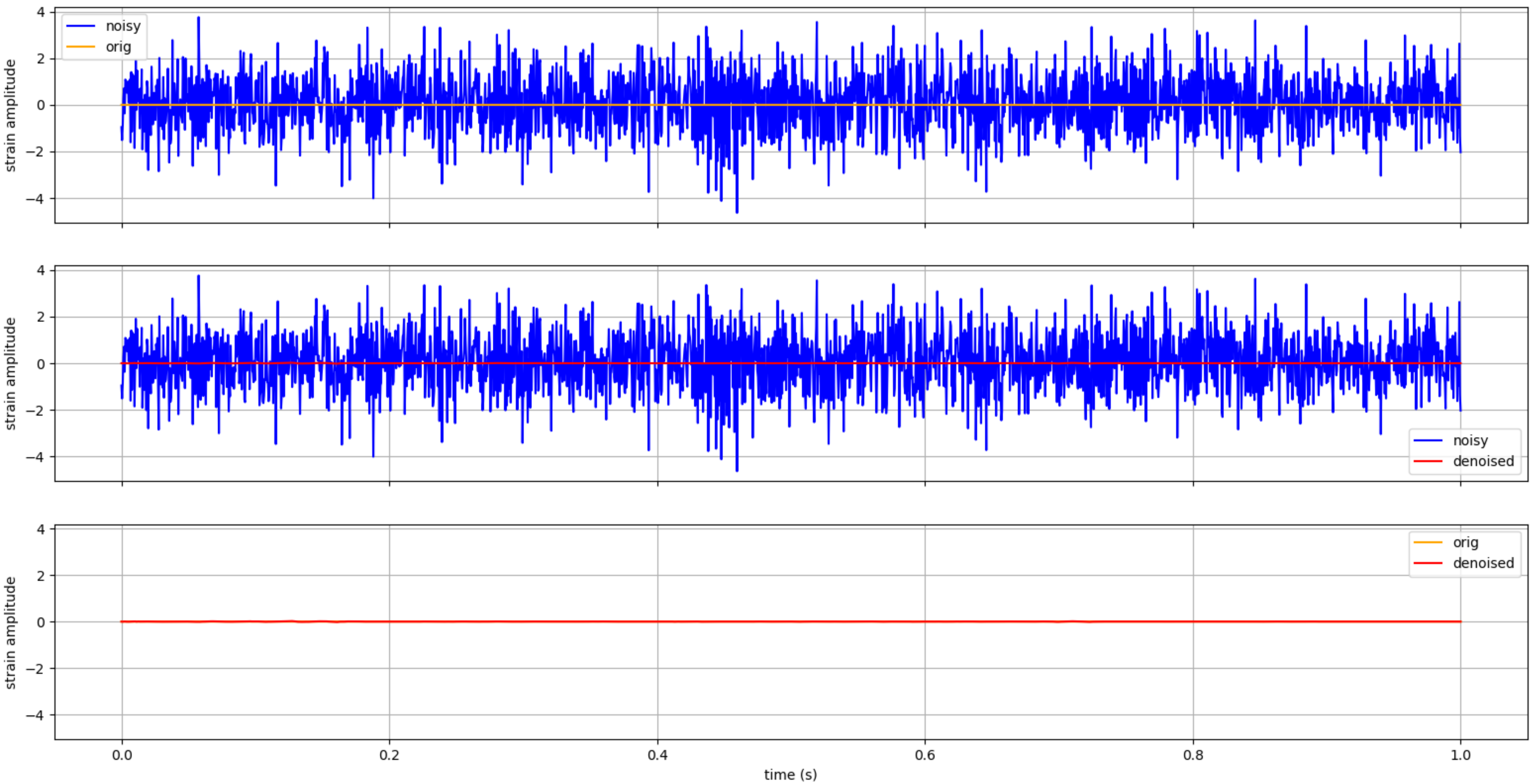


Result on a faint signal

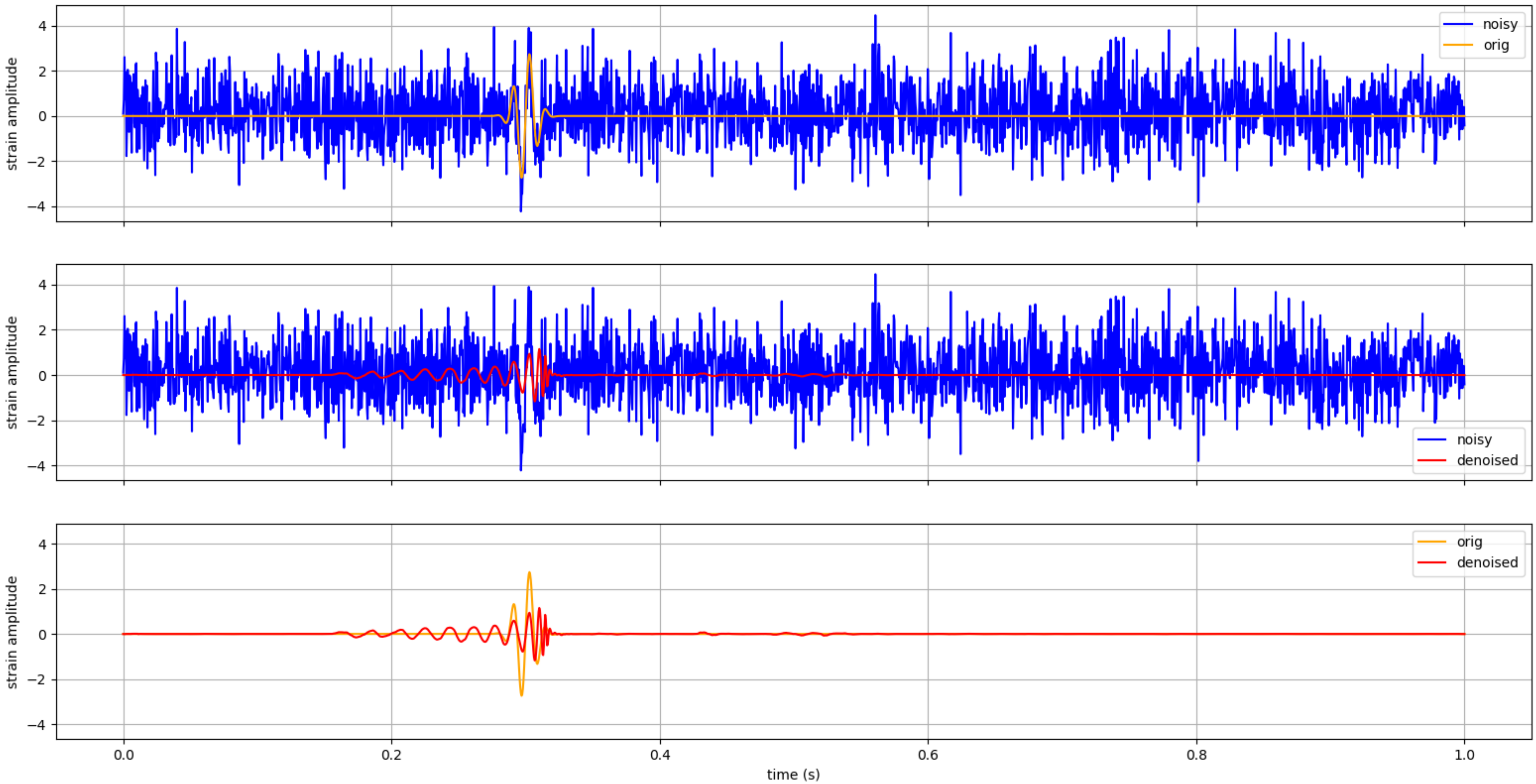
network SNR=5 / Overlap=0.79



Result without injection



Result with a synthetic glitch



Fake detections occur with glitches (sine-gaussian or mexican hat wavelets). Ongoing work.

Conclusions 2

- Worth exploring higher sampling rates:
  - ✓ do we catch enough signal variability ?
- No pre-processing up to know (except whitening):
  - ✓ investigate band pass filter.
- Working on further improving noise robustness.
- Try to apply the method to yet uncovered regions of the parameter space (eccentricity ?)
- Caveat with using current NN architectures in physics: no proper measurement of the uncertainty/degree of belief.

COST action g2net

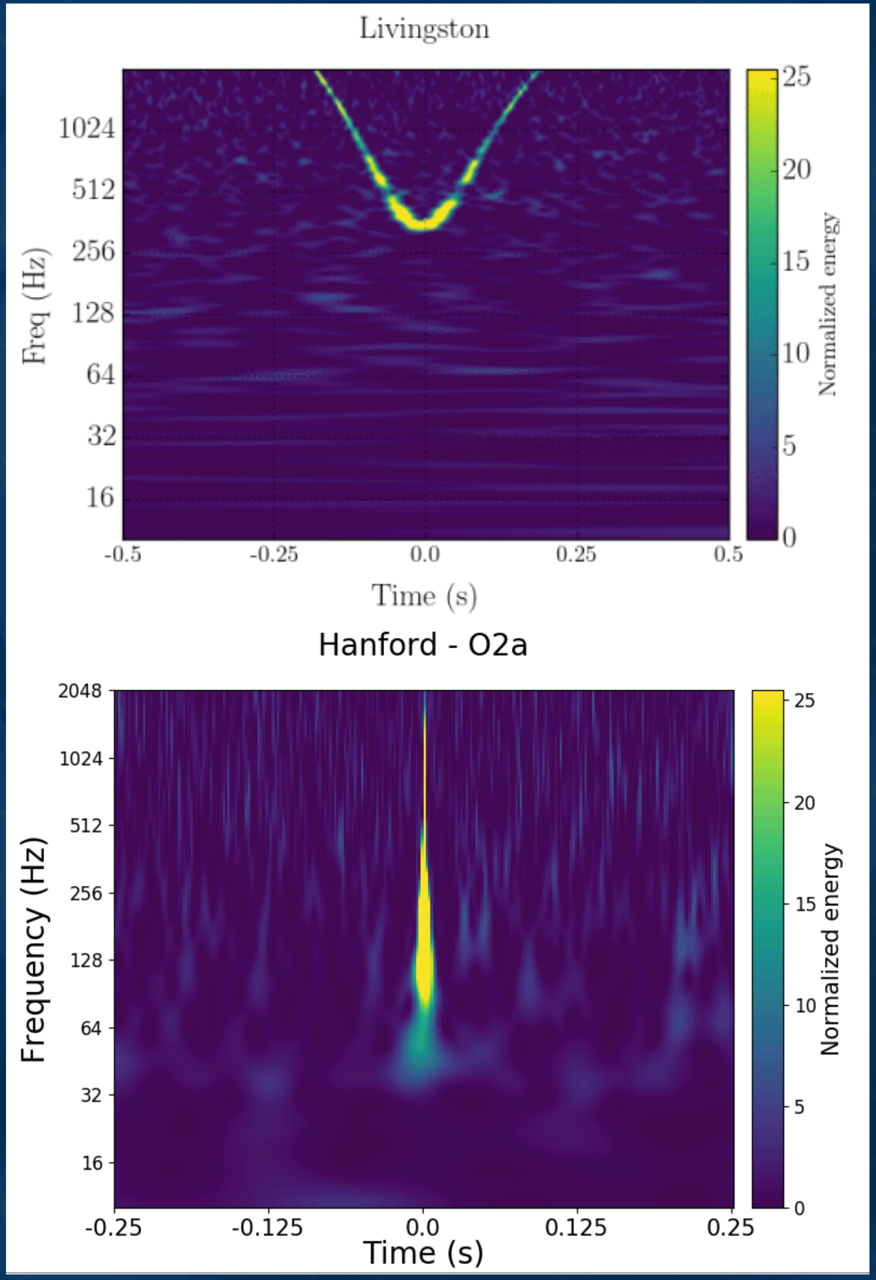
www.g2net.eu

- **WG1**: Machine Learning for Gravitational Wave astronomy
- **WG2**: Machine Learning for low-frequency seismic measurement
- **WG3**: Machine Learning for Advanced Control techniques

Join the action if you are interested!

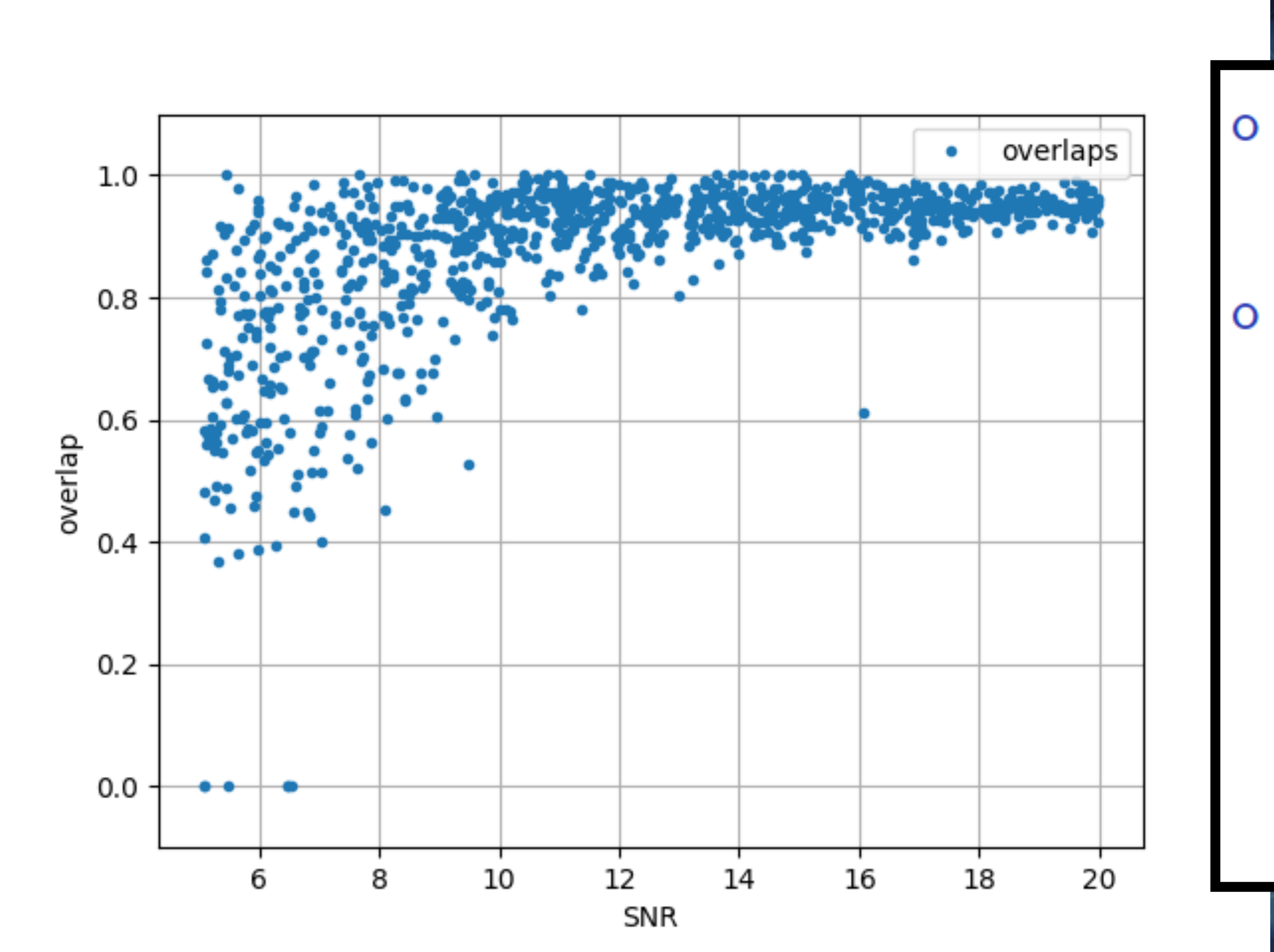
Glitches representation

- Spectrograms representation
  - ✓ Deep-learning performs well on images
  - ✓ Disadvantages:
    - Volume of data (big images)
    - Spectrogram parameters/choice dependent
    - Risk of loosing information due to manipulation
    - Deep learning algorithms learn on raw data



- Time series representation
  - ✓ full information
  - ✓ Reduced volume of data

Overlap



- SNR: standard matched filter SNR
- Overlap:

$$\mathcal{O}(h_c, h_d)^2 = \frac{\sum_i h_c[i] h_d[i]}{\sum_i h_c[i] h_c[i]}$$

with

h_c: clean signal
h_d: denoised signal.