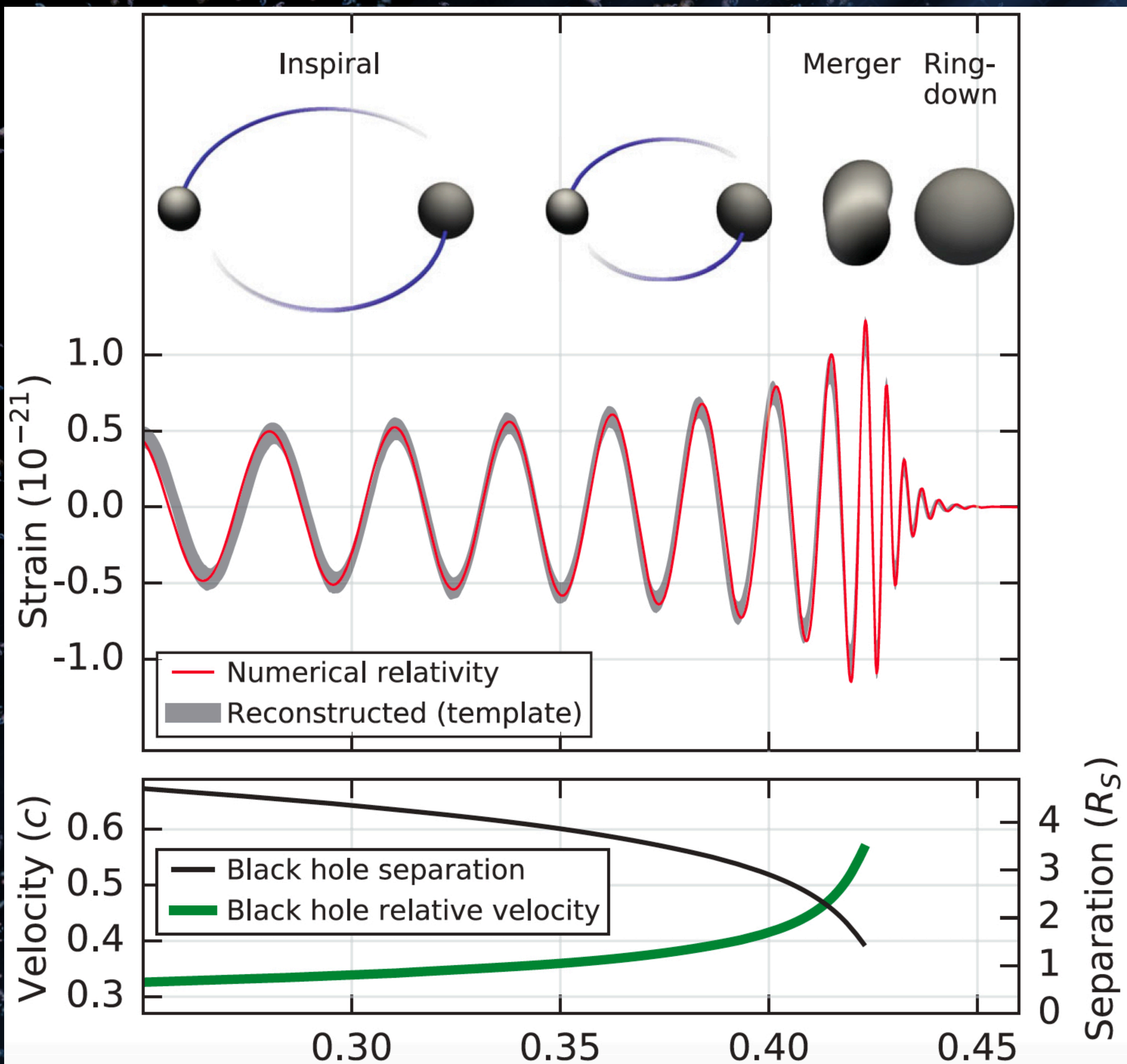
Neural networks for gravitational-wave trigger selection in single-detector periods

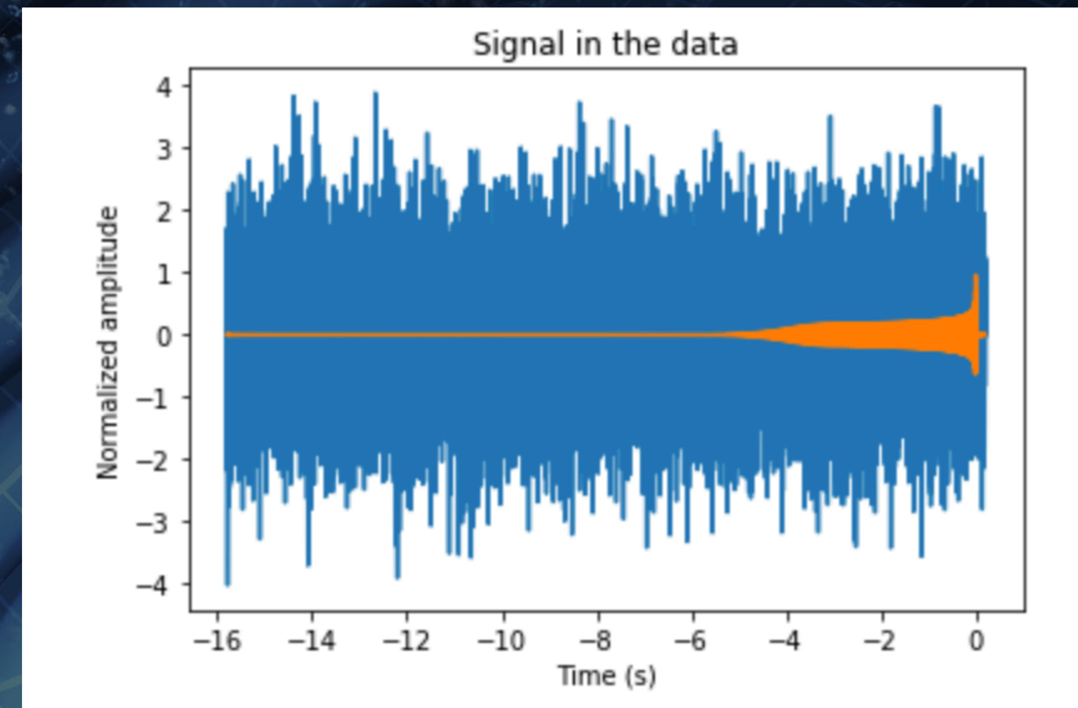
A. Trovato* with M. Bejger and E. Chassande-Mottin,
*APC, CNRS/IN2P3, Université de Paris

APC, CNRS, FRANCE

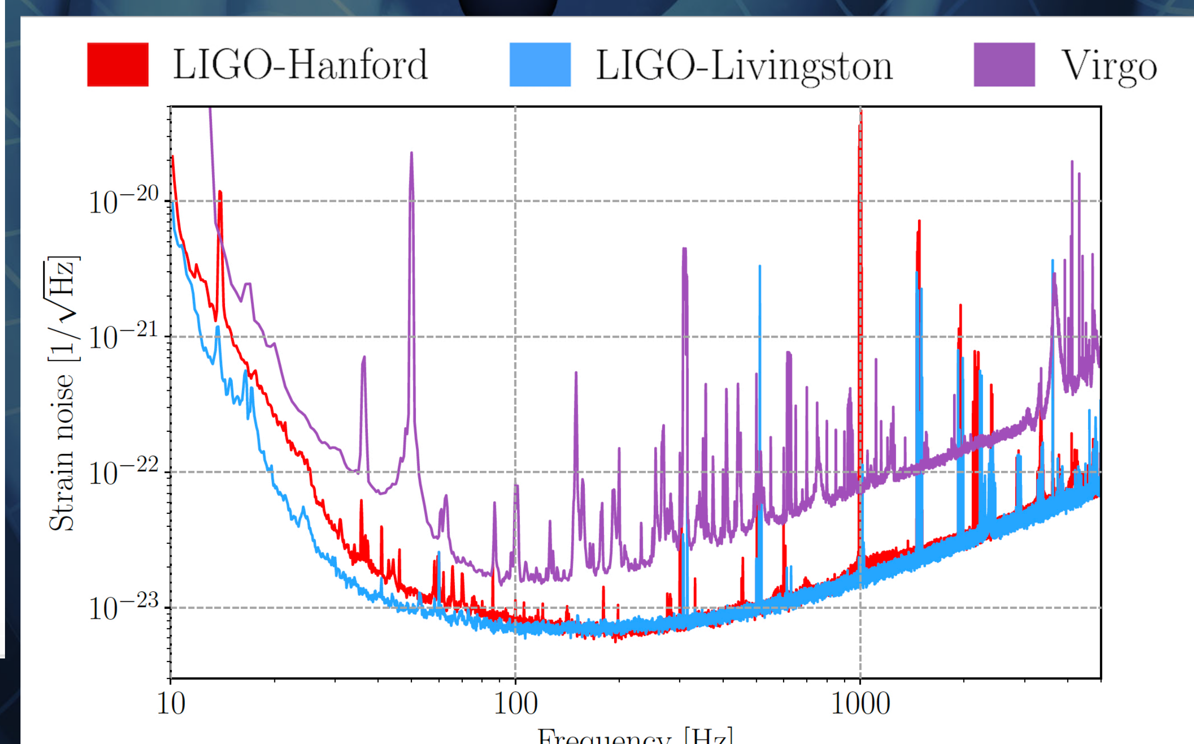
Université de Paris

Gravitational waves detection problem

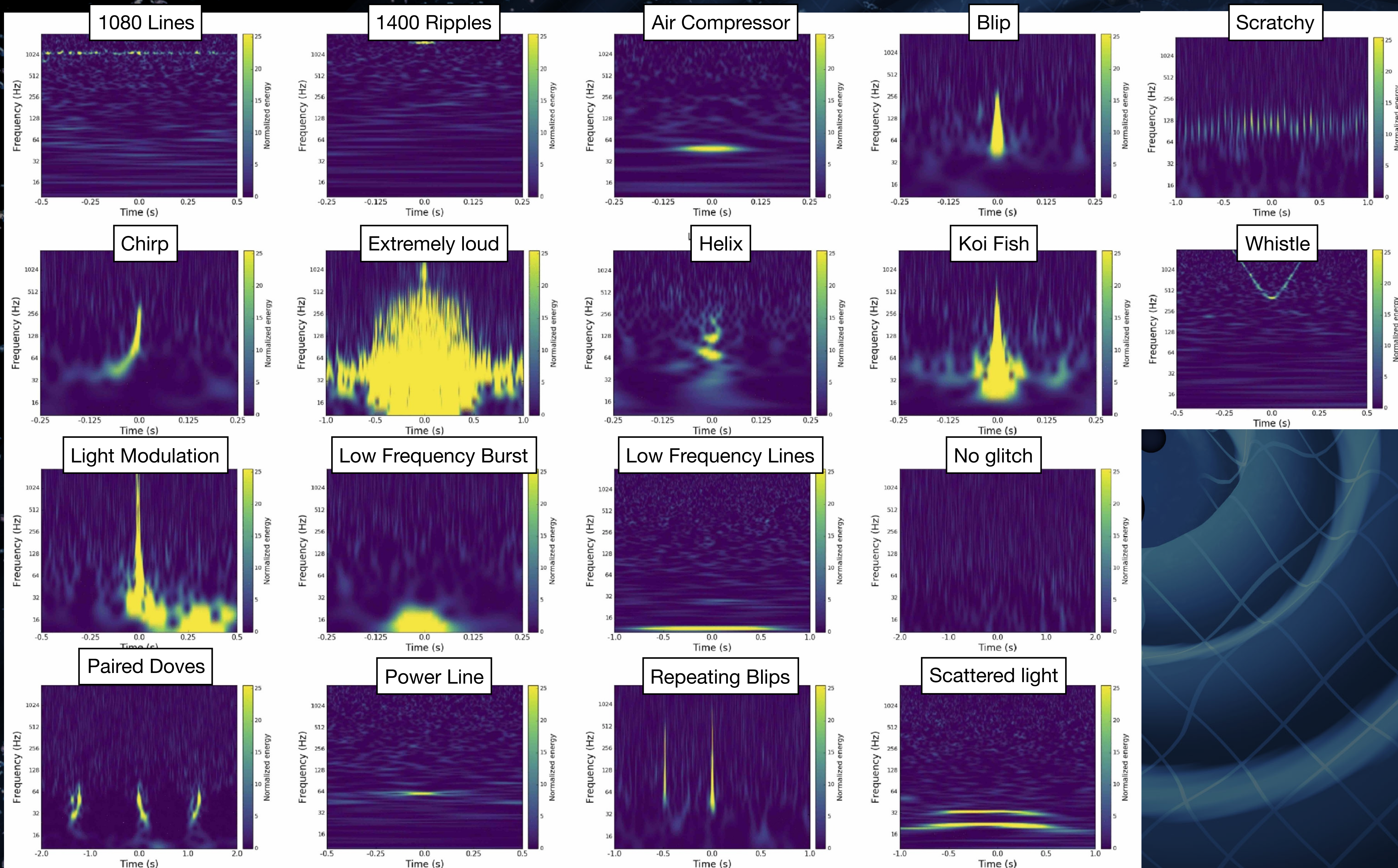




Rare and weak signals in complex background: non-Gaussian non-stationary



Glitches zoo



- ★ **Credits: Gravity Spy dataset**

GW data representation for ML

- Spectrograms representation [e.g. CQG 35 (2018) 095016, Information Sciences 444 (2018) 172]
  - ✓ Deep-learning performs well on images (reuse standard solutions)
  - ✓ Disadvantages:
    - Volume of data (big images)
    - Spectrogram parameters/choice dependent
    - Risk of loosing information due to manipulation

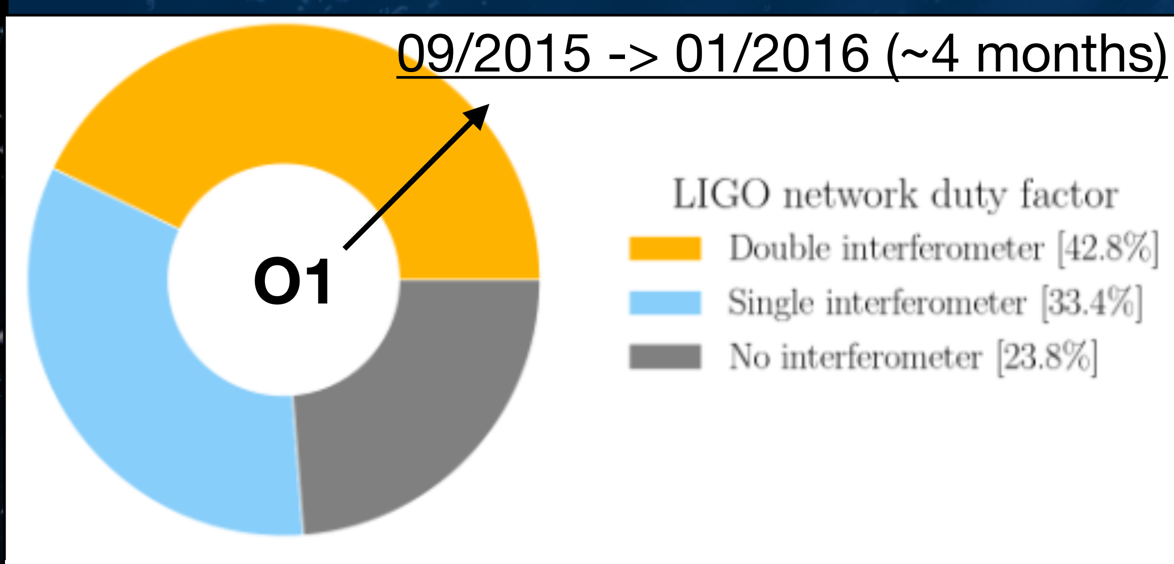
- Time series representation [e.g. Phys. Lett. B 778 (2018) 64, Phys. Rev. D100 (2019) 063015]
  - ✓ full information & reduced volume of data
  - ✓ Multi-detector searches, attempt to make high-confidence detection

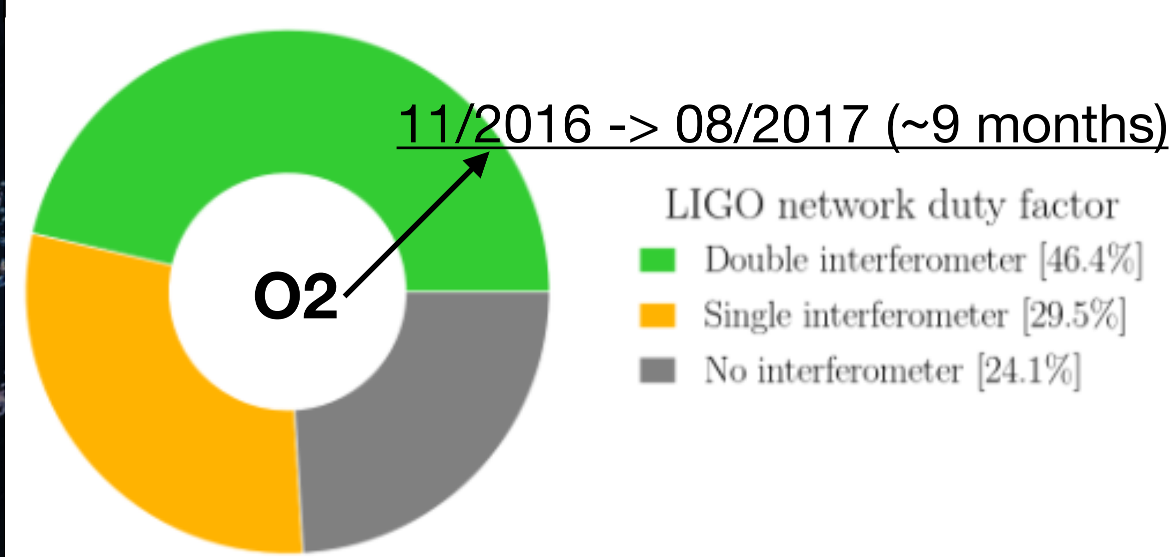
- This work:
  - ✓ time-series representation, single detector, trigger pre-selection

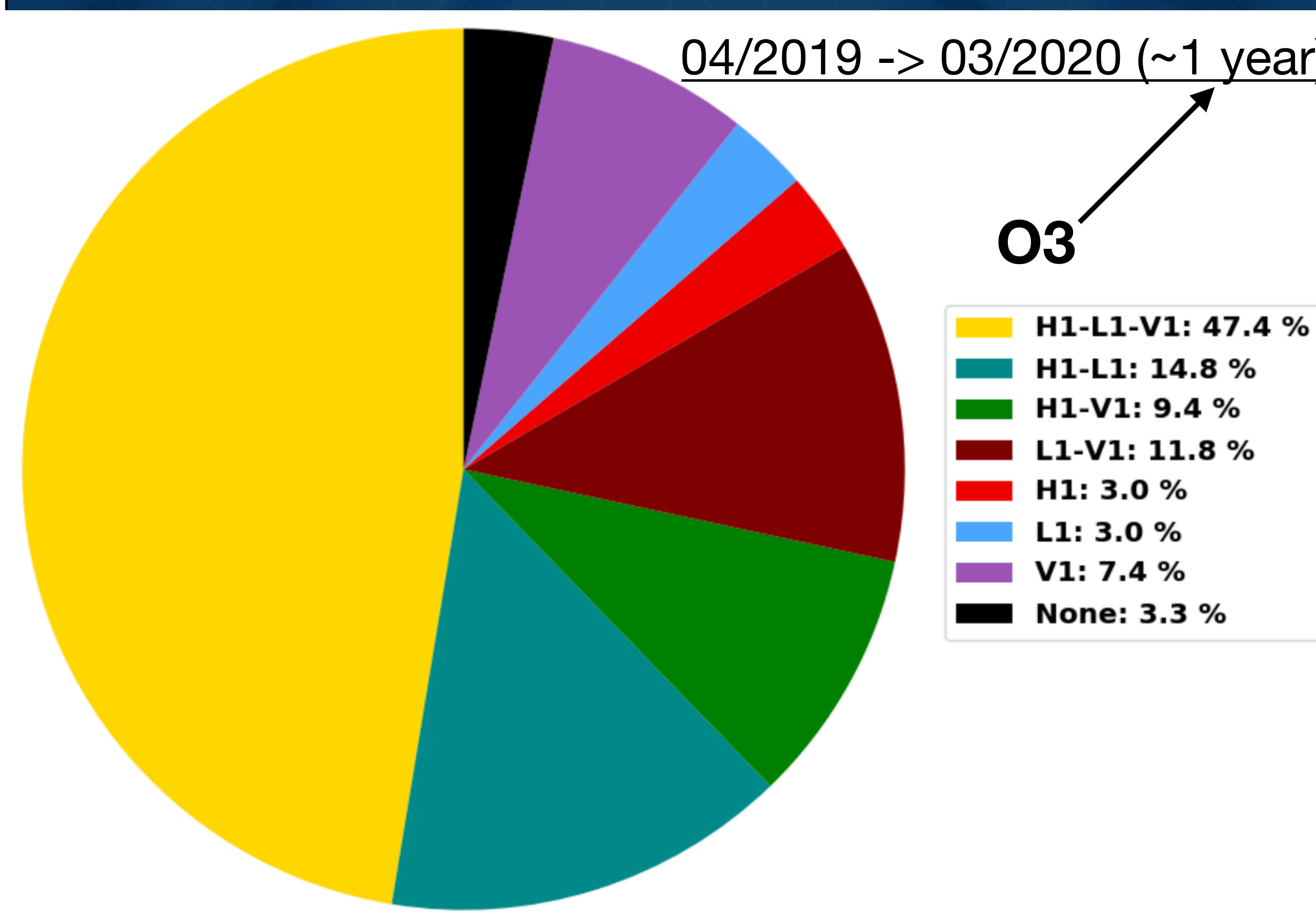
Single-detector time

- Glitch impact on sensitivity is larger during single-detector periods as coincidence with additional detector is impossible. Can machine learning help?

- Single-detector time:
  - ✓ 2.7 months in O1+O2; 1.6 month in O3

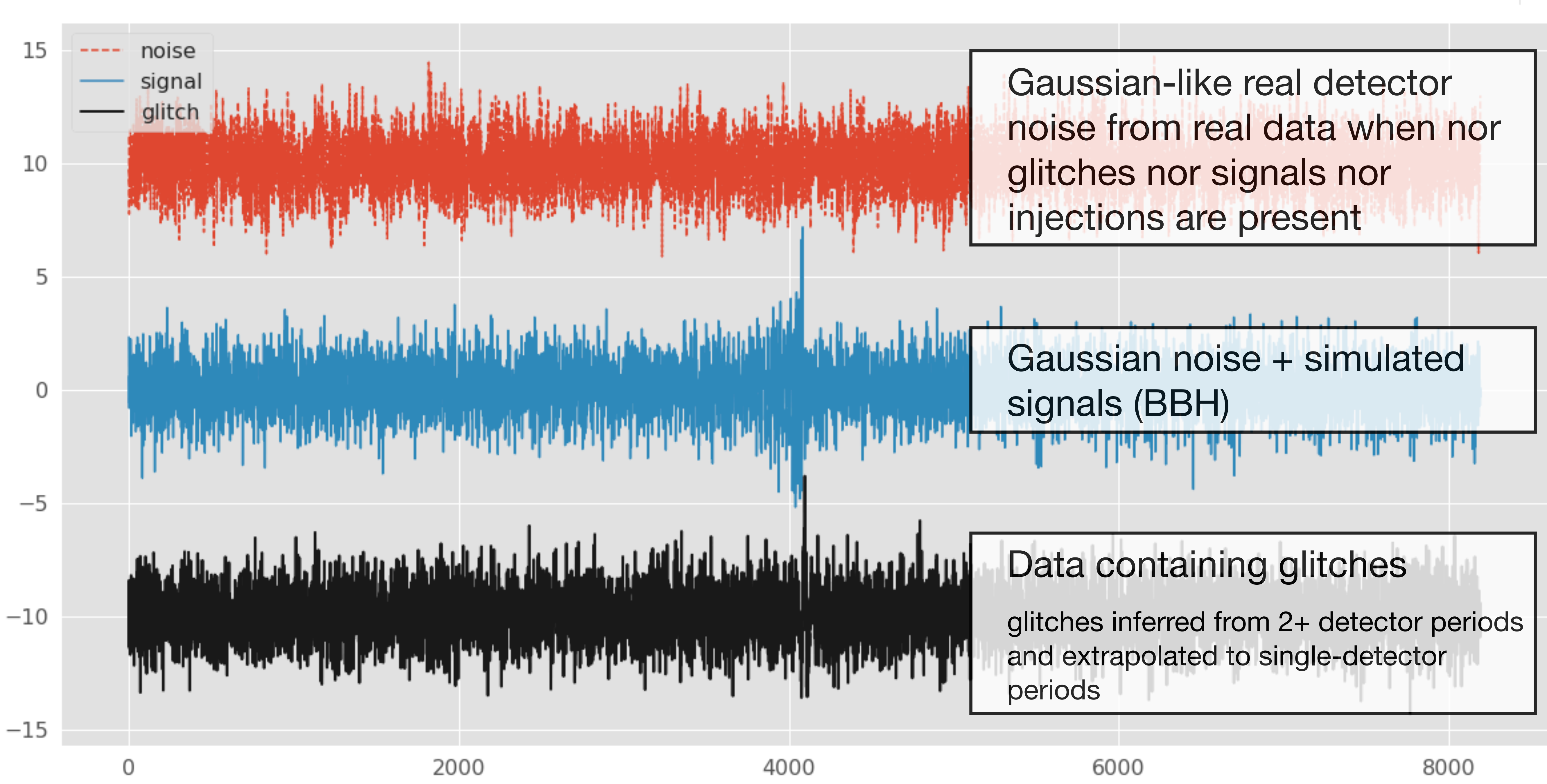






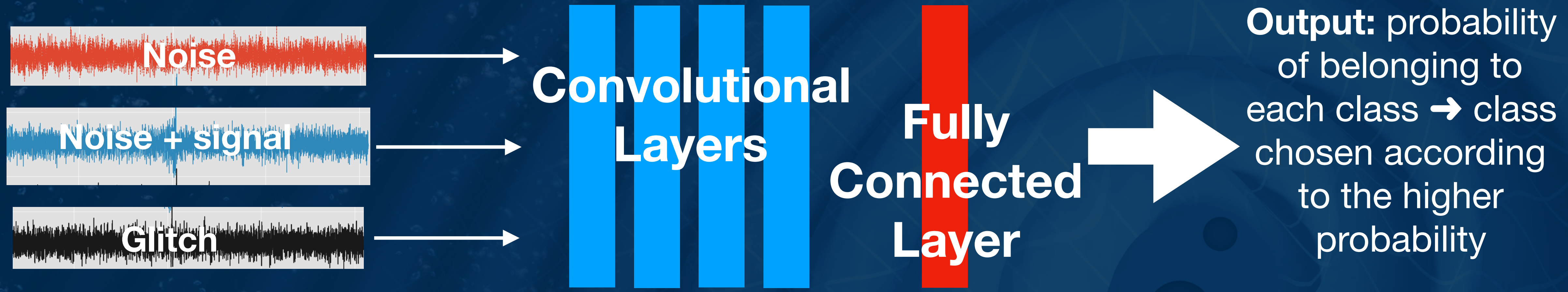
Training data: 3 classes

Segments of glitches and “clean” noise data samples from the one month of LIGO O1 run (downsampled to 2048 Hz), whitened by the amplitude spectral density of the noise.



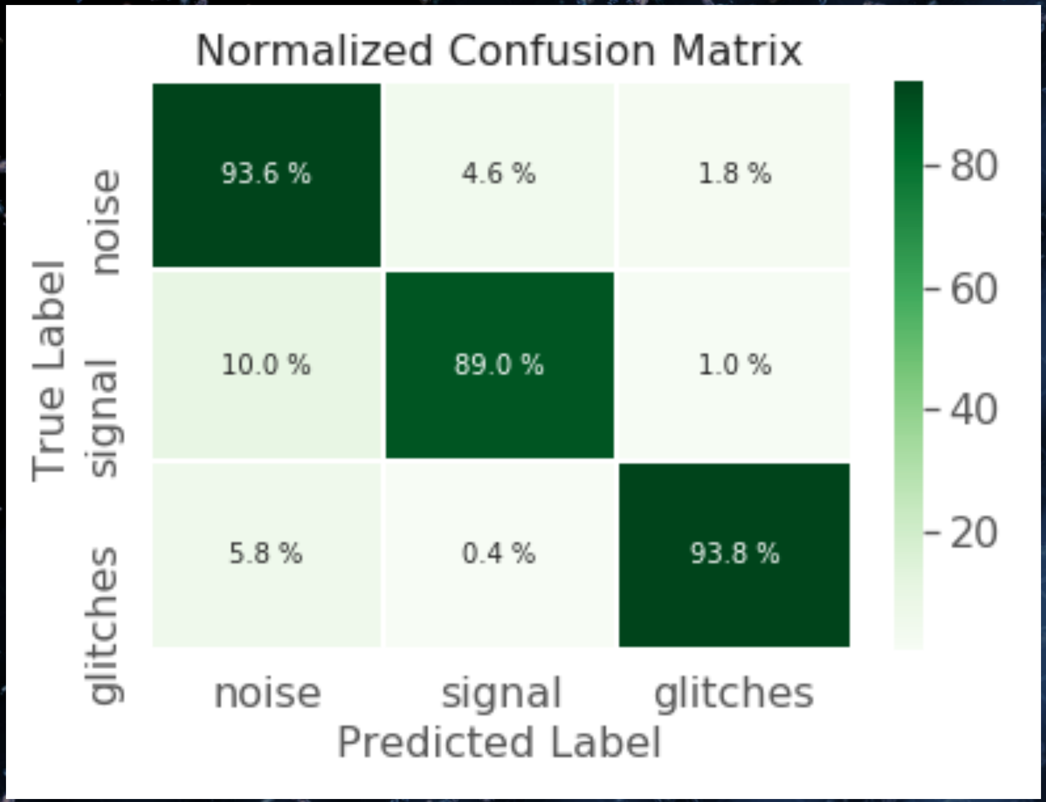
Network used

- CNN used: small network with 4 convolution layers (with dropouts and pooling) used as classifier to distinguish the 3 classes: noise, noise+signal, glitches



Layer #	1	2	3	4	5
Type	Conv	Conv	Conv	Conv	Dense
Filters	64	32	16	8	-
Kernel Size	16	8	8	4	-
Strides	4	2	2	1	-
Activation	relu	relu	relu	relu	softmax
Dropout	0.5	0.5	0.25	0.25	-
Max Pool	4	2	2	2	-

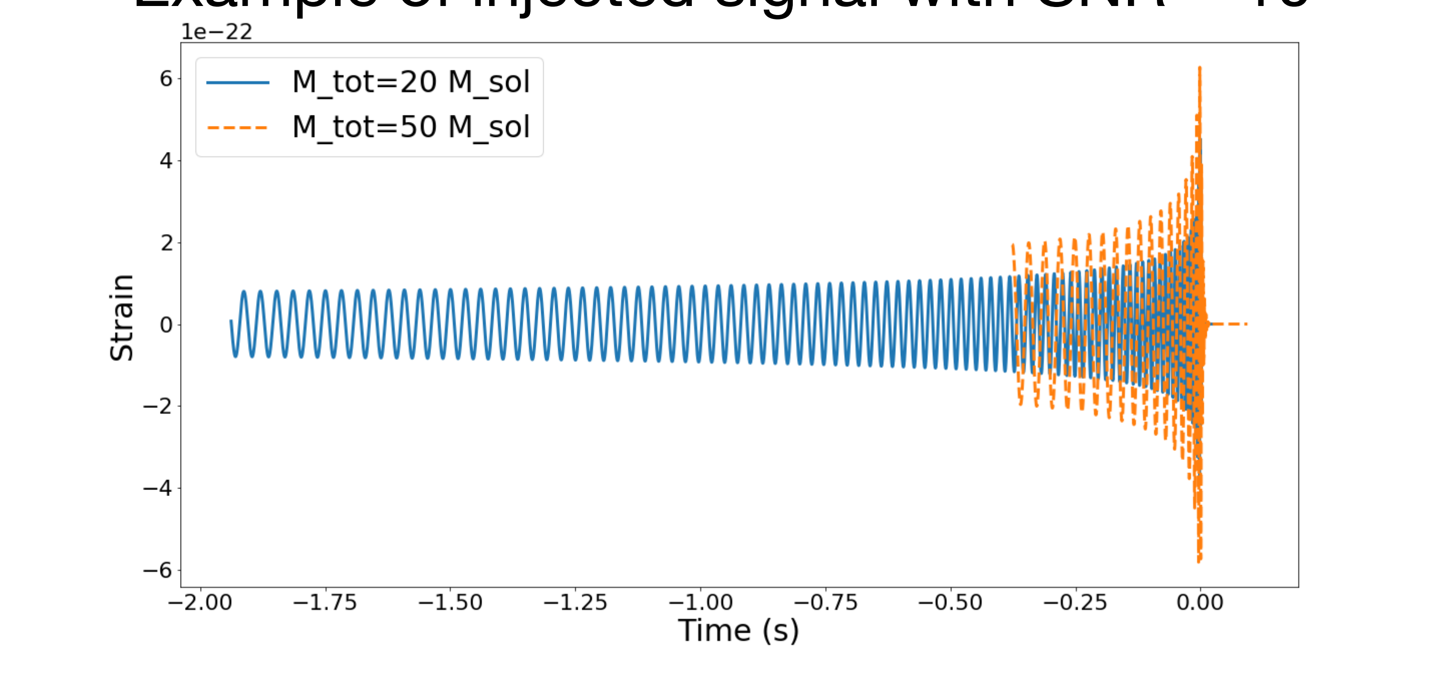
Confusion matrix and dataset details



Additional dataset details

- **Segments length: 1 second**
- **Injected signals (BBH)**
  - $m_1+m_2 \in (33,60)\ M_\odot$
  - SNR $\in$ (8,20)
- **Selected glitches**
  - SNR > 10

Example of injected signal with SNR = 10

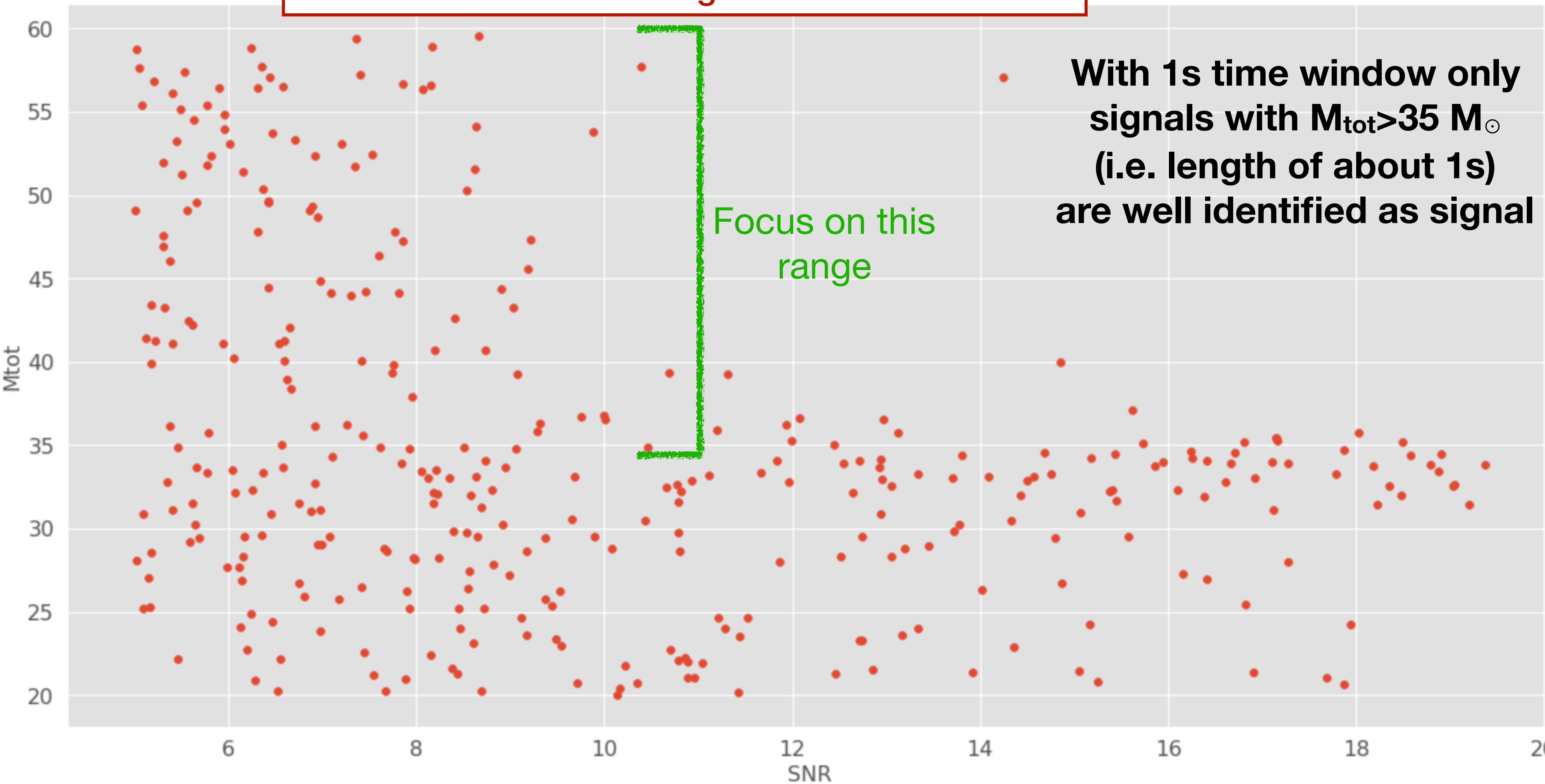


Detectability across the parameter space

1 s time window

Signal with $m_{tot} \in (20, 60)\ M_\odot$

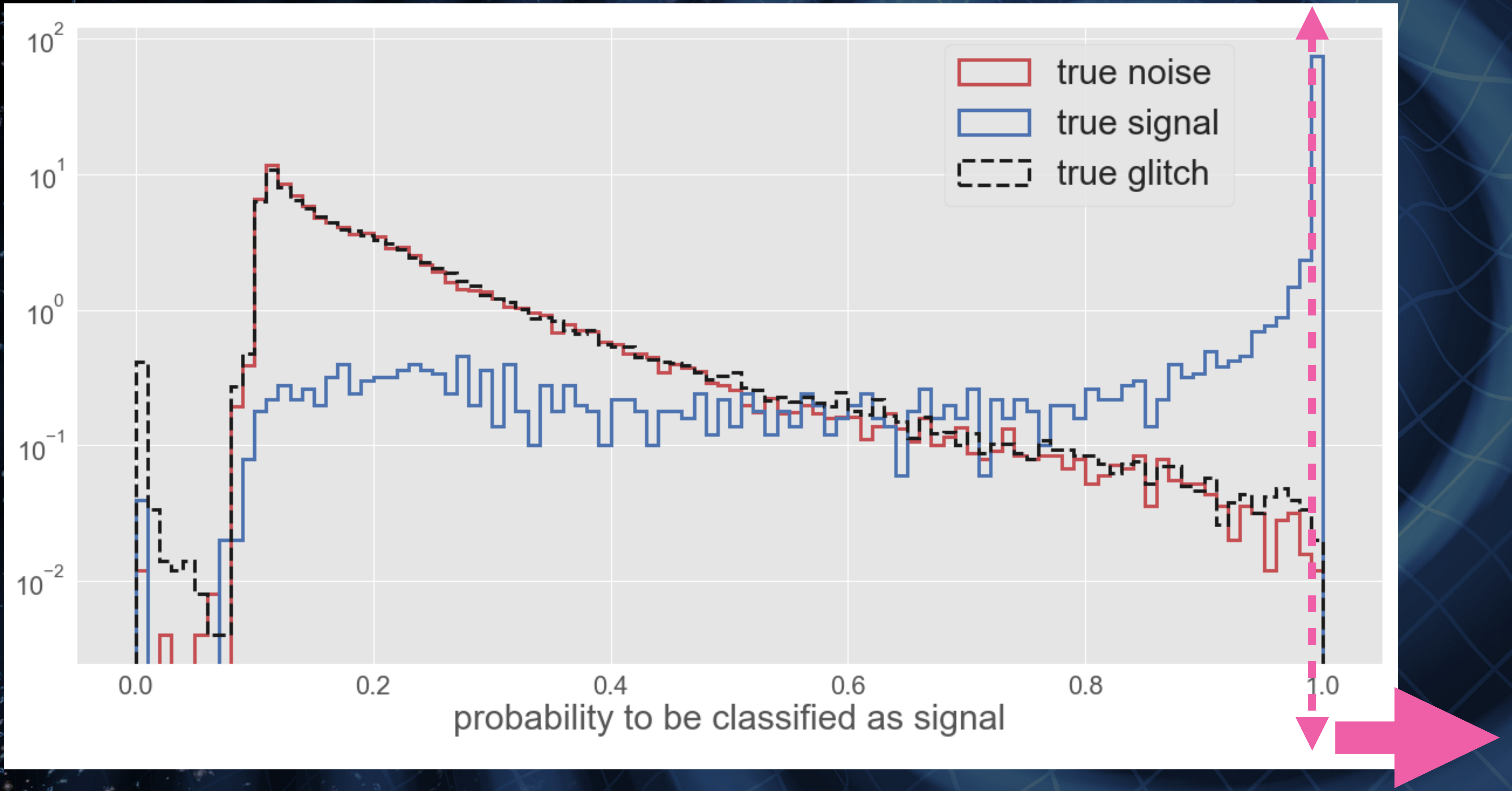
Missed detections: true signals classified as noise



With 1s time window only signals with M_{tot}>35 $M_\odot$ (i.e. length of about 1s) are well identified as signal

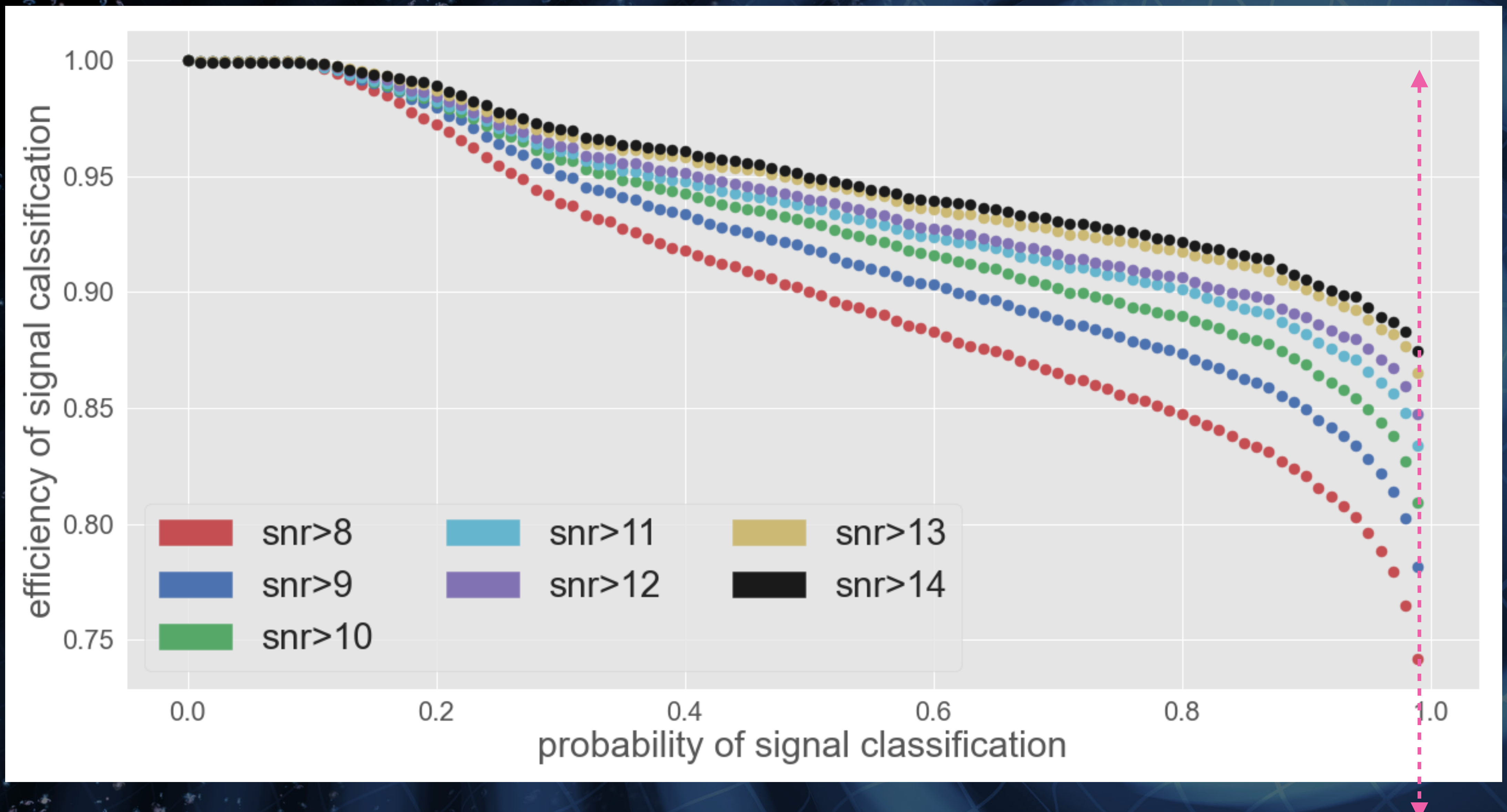
Probabilities of classification

Test: use the probability of the signal classification as statistic to distinguish signal vs noise+glitches



Efficiency

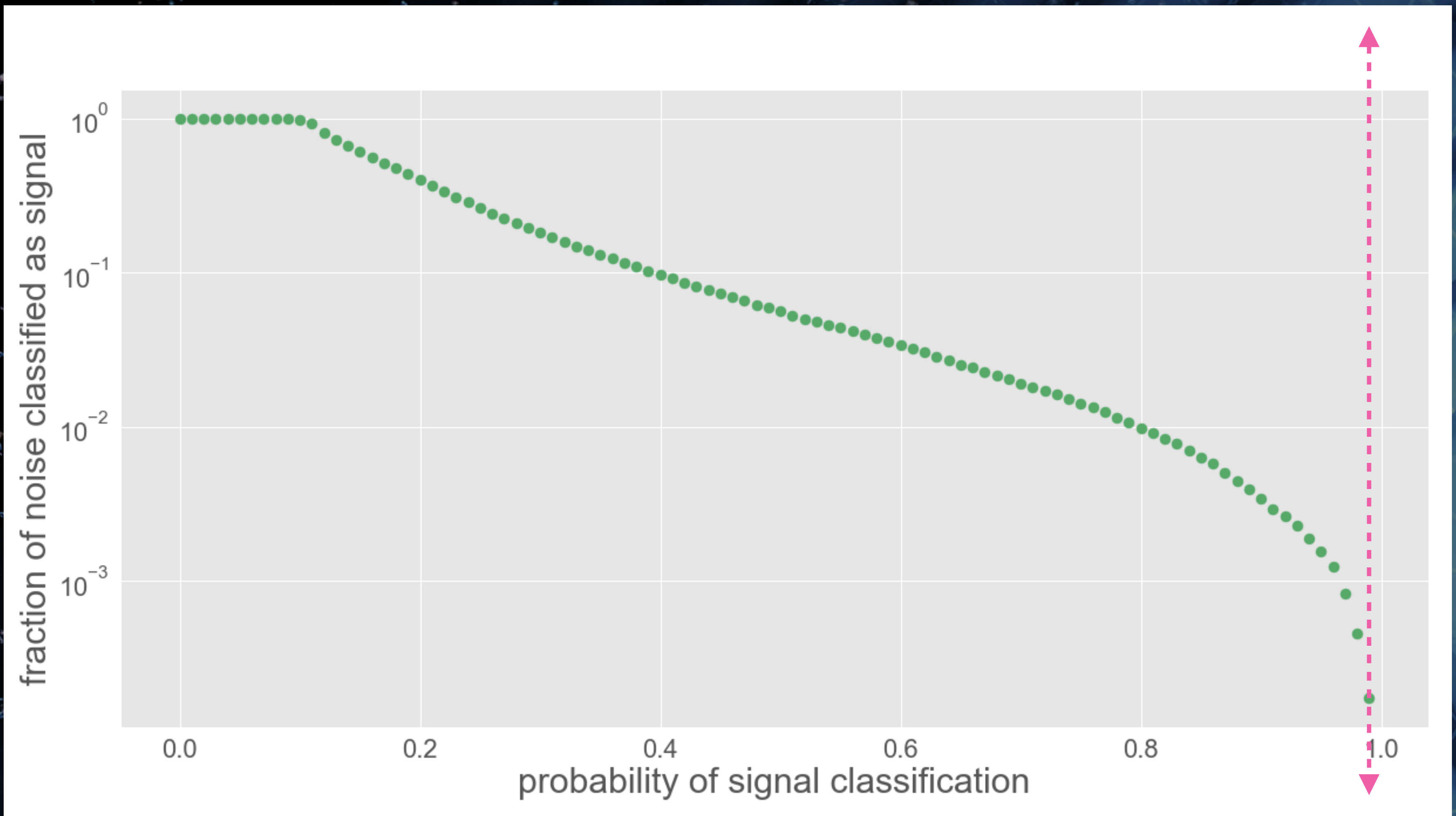
Efficiency = Fraction of signal well classified w.r.t. all the signals present in the dataset



With a stringent cut on probab_signal>0.99, reasonable efficiency around 80-90% for signals with SNR>10

False Alarm Rate

FAR = Fraction of noise+glitches classified as signals w.r.t. all the noise+glitches present in the dataset



With a stringent cut on probab_signal>0.99, FAR ≈ 1/83 min ➜ this means about 2000 false alarms in O1!

Noise rejection is too limited

Trigger pre-selection (rather than high-confidence detection)

Results similar to other works on the subject

e.g. arXiv:1904.08693*, 1701.00008**, arXiv:1711.03121

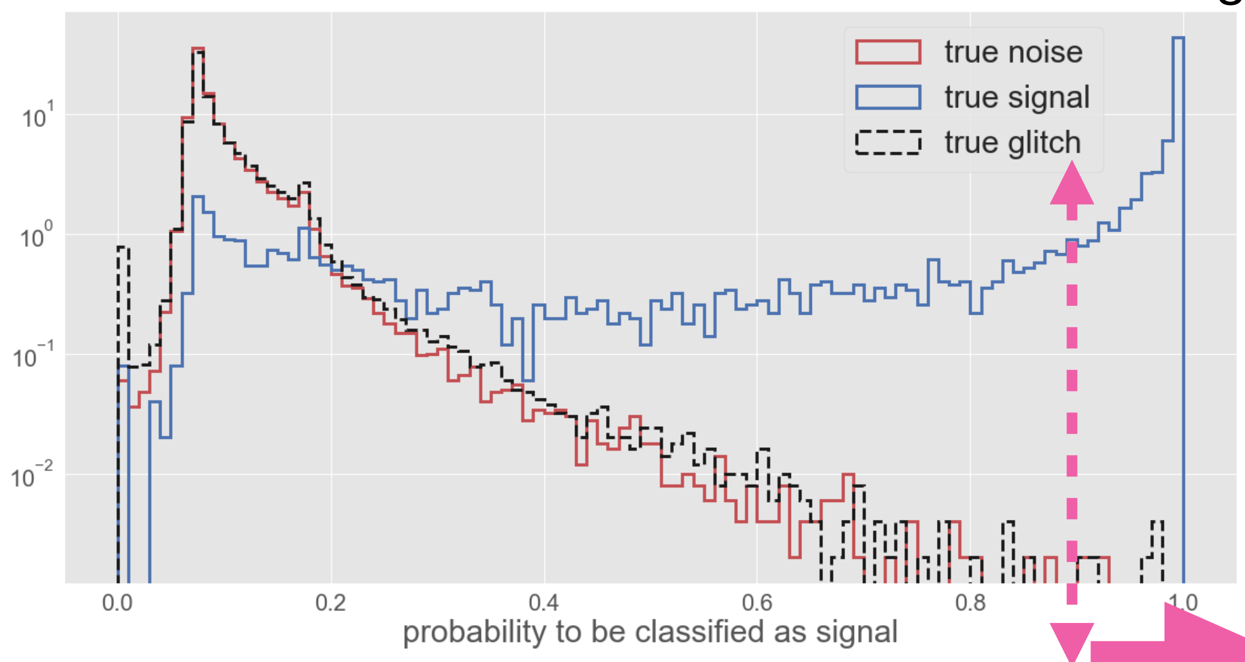
*FAR of 1/40 minutes with detection ratio of 86%

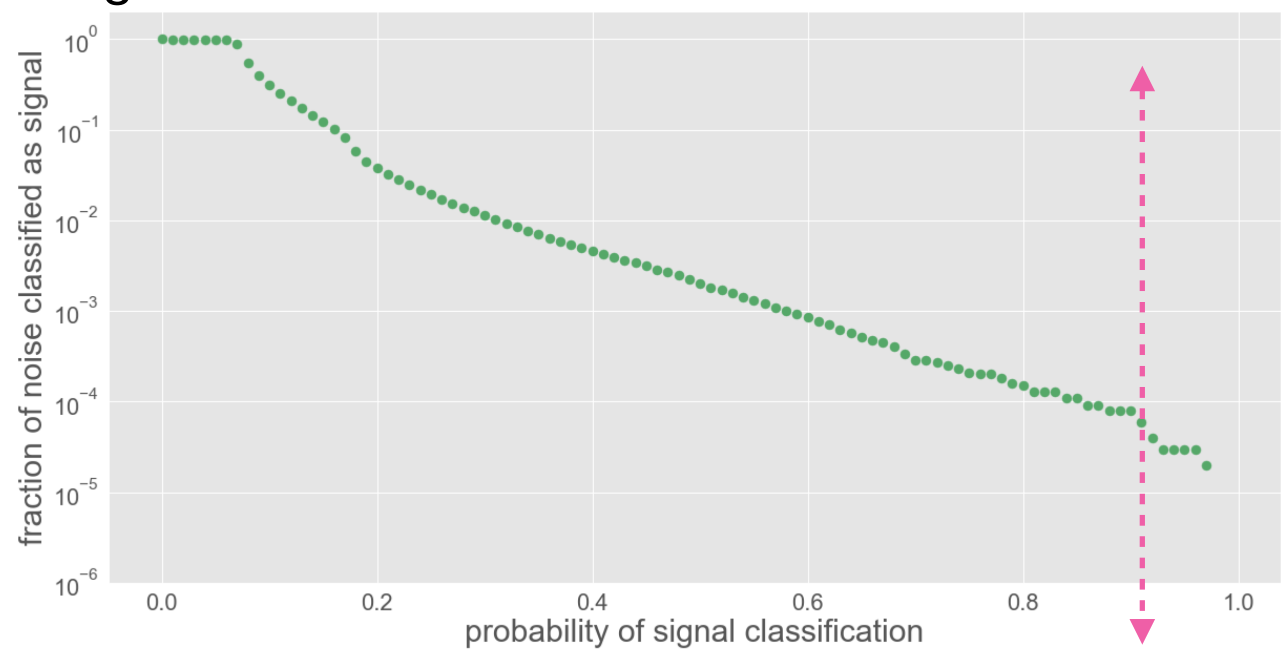
**FAR of 0.6% and 100% sensitivity for SNR>10

Data filtering

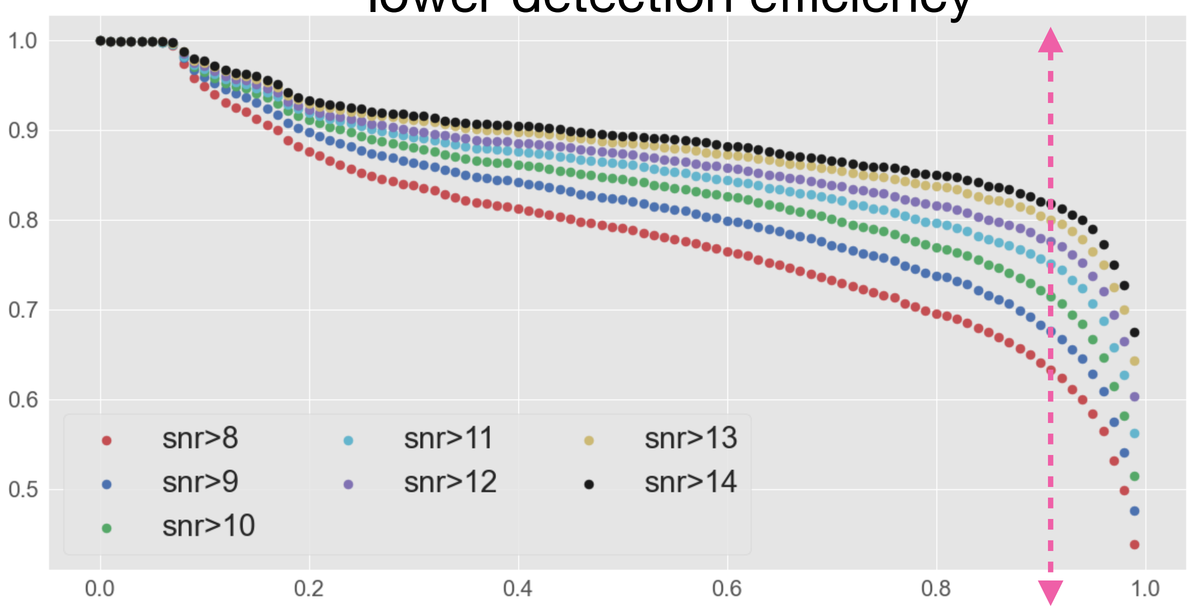
Band pass filter to all data [20,1e3] Hz

reduced noise/glitch background tails!





lower detection efficiency



In this case with a cut on probab_signal>0.90, FAR ≈ 1/3.5 h ➜ ~ 900 false alarms in O1!

Conclusion and perspectives

- GW signal classifier from single-detector time-series
  - ✓ Able to reach correct classification to the percent scale
  - ✓ However, not sufficient for high-confidence detection (too many false alarms)
  - ✓ Due to large class imbalance in the observations (signal very rare, noise very common)
  - ✓ Hint of improvement with a band pass filter -> more statistic needed

- Can noise rejection be improved? Can we optimize the CNN with this objective specifically?
  - ✓ Focus on the imbalance between classes -> explore different loss functions
    - f1 loss tested, Neyman-Pearson under study
  - ✓ Extend the data set
  - ✓ Consider different architectures and hypermarameters
  - ✓ Suggestions are welcome


Backup slides

Precision and recall

Choose a relevant class: e.g. signal

$$\text{Precision} = \frac{tp}{tp + fp}$$

Fraction of signal well classified w.r.t. those classified as signal

$$\text{Recall} = \frac{tp}{tp + fn}$$

Fraction of signal well classified w.r.t. all the signals present in the sample

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn}$$

can be a misleading metric for imbalanced data sets

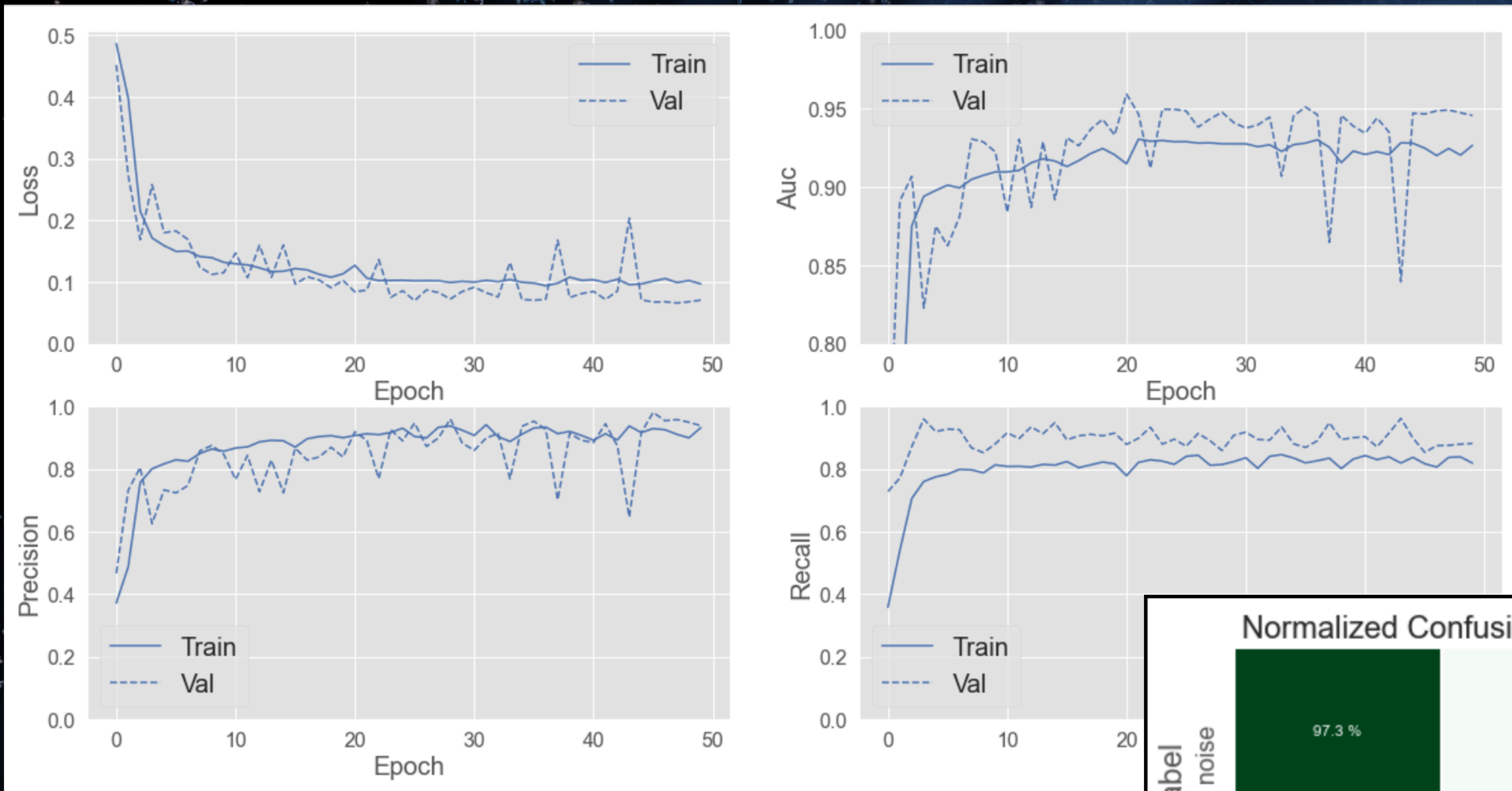
$$F_1 = \left(\frac{2}{\text{recall}^{-1} + \text{precision}^{-1}}\right) = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}.$$

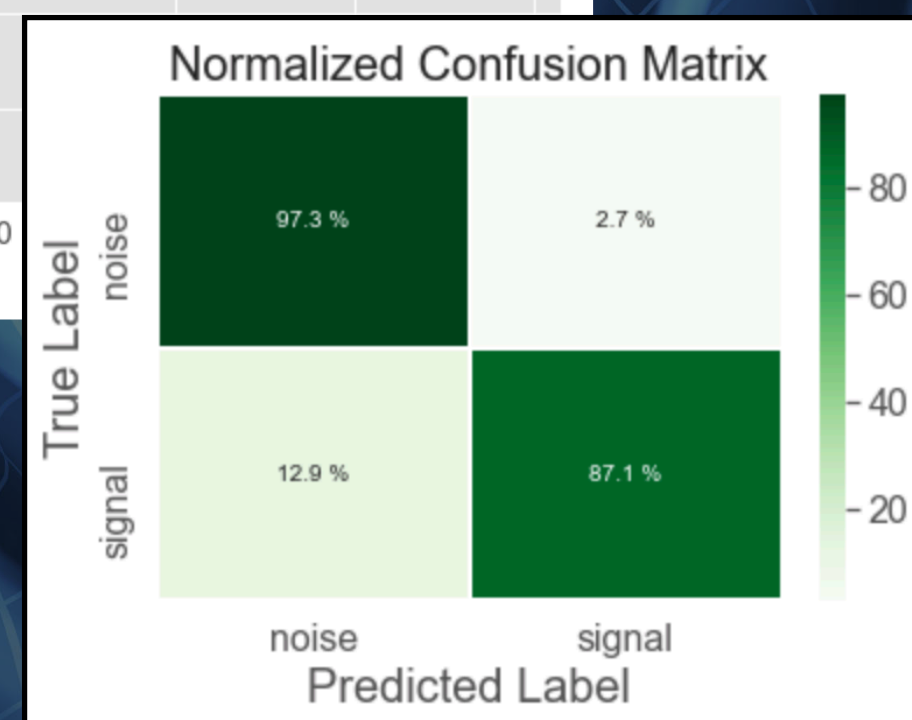
Harmon mean of precision and recall

https://arxiv.org/pdf/1402.1892.pdf

if the classifier is completely uninformative, then the optimal behavior is to classify all examples as positive.

Loss = f1-score

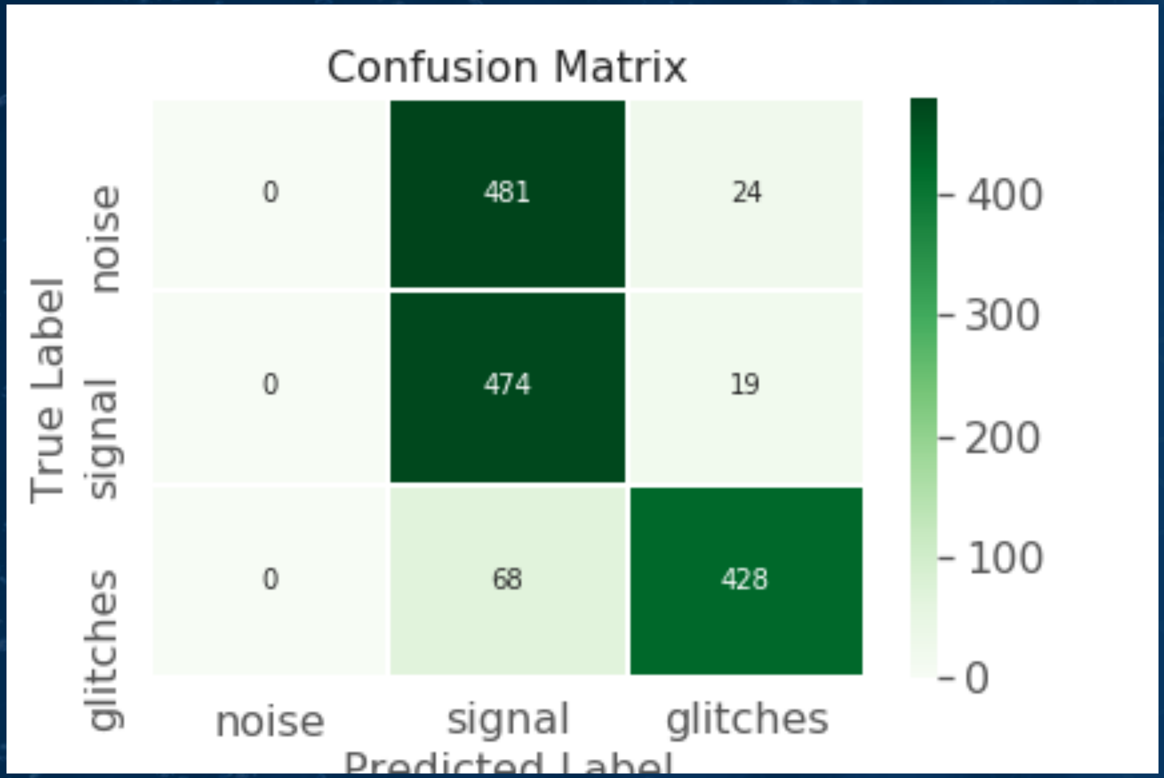




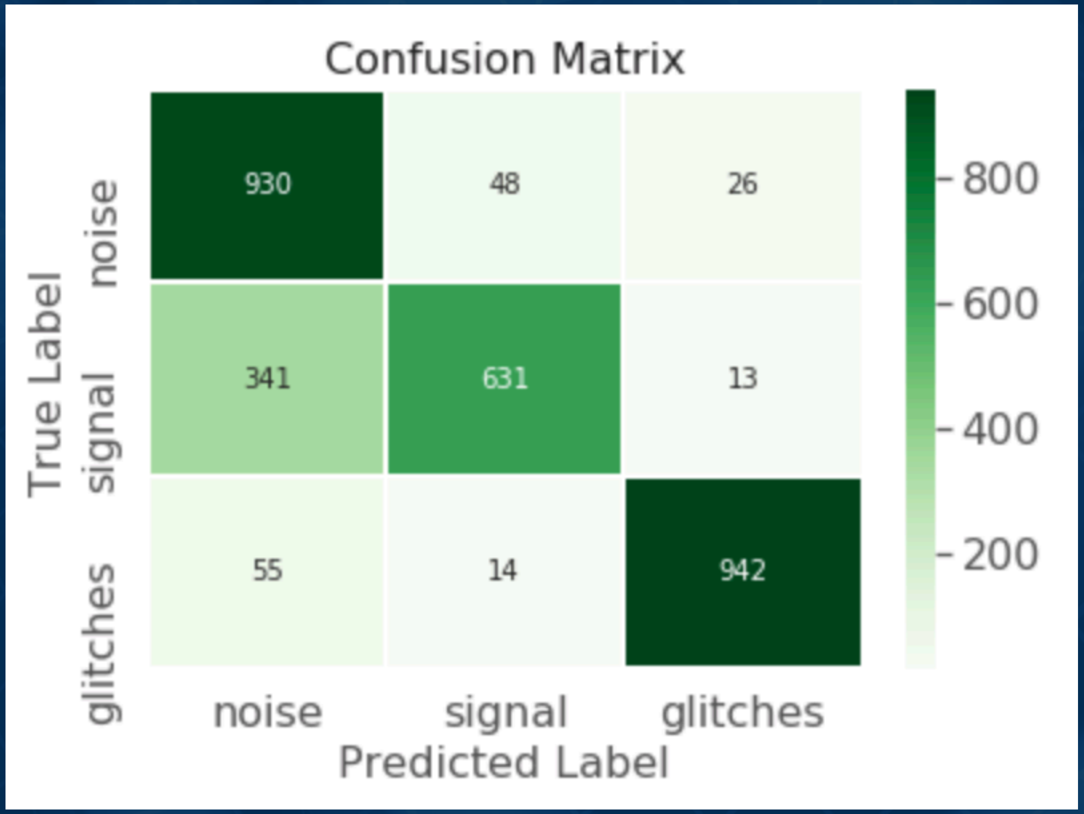
Length of the time window (I)

- Length of the time window (= size of the input data segment) coupled with the masses of the simulated signals
  - ✓ Signals with m_1, $m_2 \in (10, 30)\ M_{\odot}$, 5<SNR<20 ; glitches with SNR>10

4 s time window



1 s time window



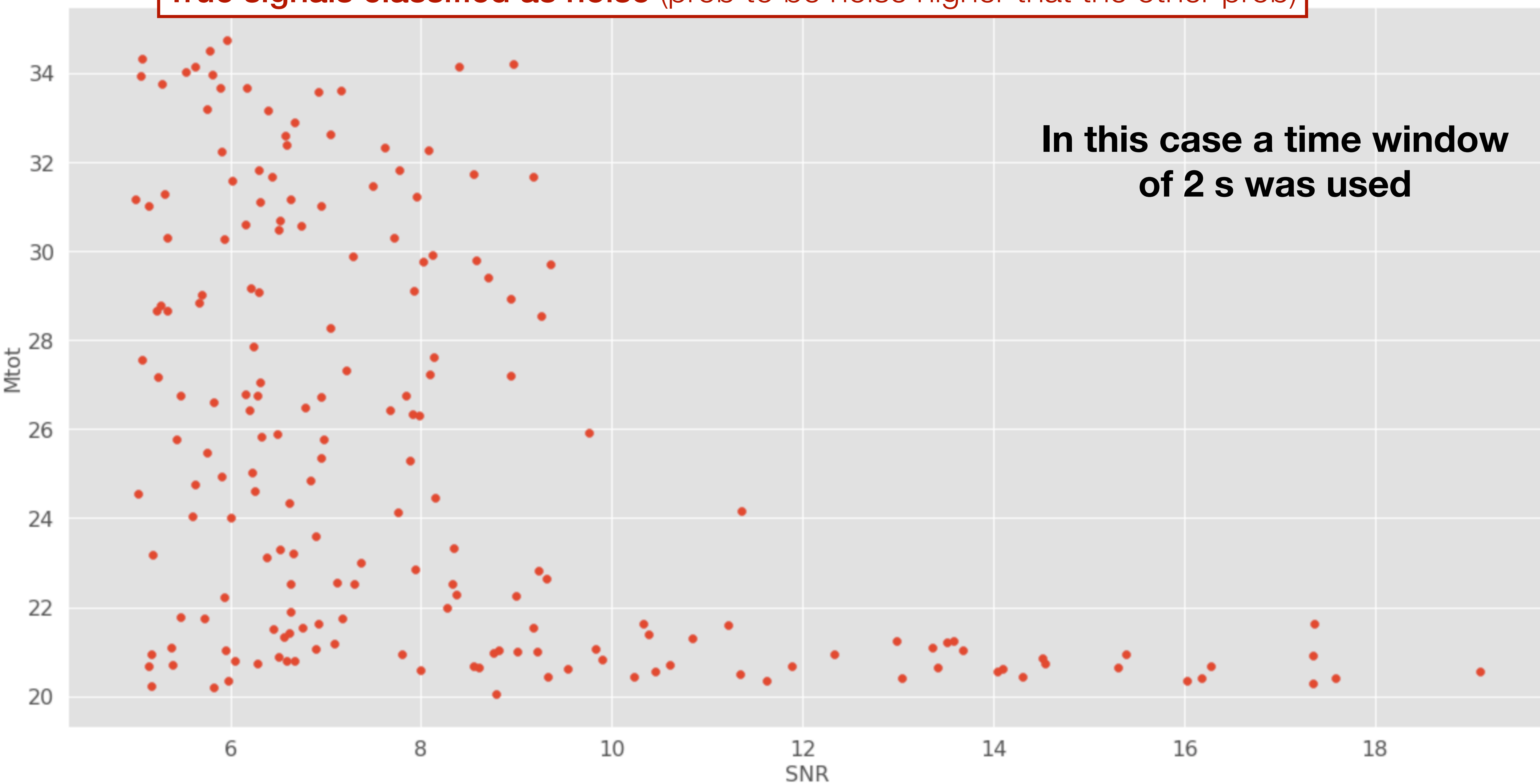
CNN:

Conv1D (500, 5)	MaxPooling1D (3)	Conv1D (250, 5)	Conv1D (500, 5)	MaxPooling1D (3)	Conv1D (150, 5)	MaxPooling1D (3)	Dropout (0.5)

Length of the time window (II)

A network working with windows of 1s could be combined with another one with 2 s windows, each optimised for different ranges in masses

True signals classified as noise (prob to be noise higher that the other prob)



In this case a time window of 2 s was used