Deep Learning Networks & Gravitational Wave Signal Recognization

He Wang (王赫)

[hewang@mail.bnu.edu.cn]

Department of Physics, Beijing Normal University

In collaboration with Zhou-Jian Cao

Aug 23rd, 2019

The 23rd KAGRA face-to-face meeting @Toyama

Introduction

- Problems
  - Current matched filtering techniques are *computationally expensive*.
  - *Non-Gaussian* noise limits the optimality of searches.
  - *Un-modelled* signals?

$\boxed{\text{A trigger generator}}$ $\longrightarrow$ Efficiency+ Completeness + Informative

- Solution:
  - Machine learning (deep learning)
  - ...

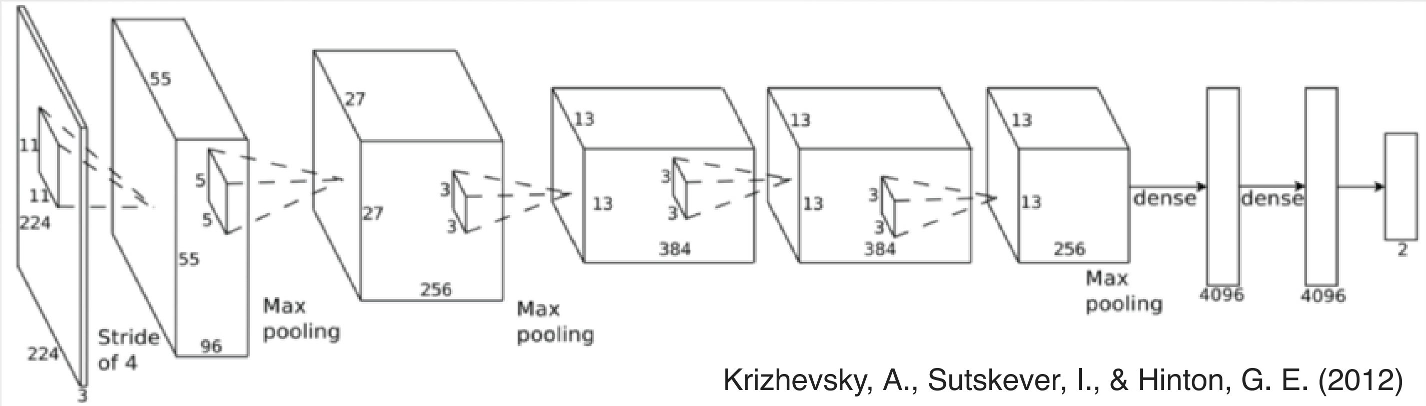
Introduction

- Existing CNN-based approaches:
  - Daniel George & E. A. Huerta (2018)
  - Hunter Gabbard et al. (2018)
  - X. Li et al. (2018)
  - Timothy D. Gebhard et al. (2019)

- Our main contributions:
  - A **brand new** CNN-based architecture (MF-CNN)
  - **Efficient** training process (no bandpass and explicit whitening)
  - **Effective** search methodology (only 4~5 days on O1)
  - Fully **recognized** and predicted **precisely** (<1s) for **all** GW events in O1/O2

MF-ConvNet Model

Convolutional neural network (ConvNet or CNN)

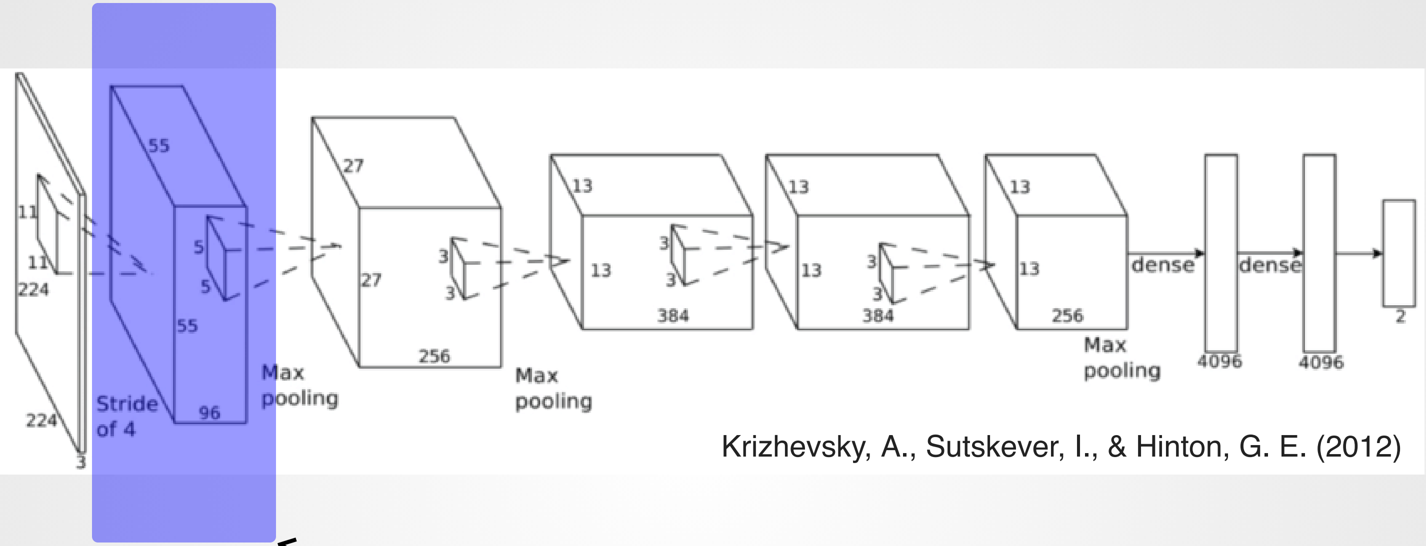


Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012)

"Matched-filtering (cross-correlation with the templates) can be regarded as a convolutional layer with a set of predefined kernels.

MF-ConvNet Model

Convolutional neural network (ConvNet or CNN)



Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012)

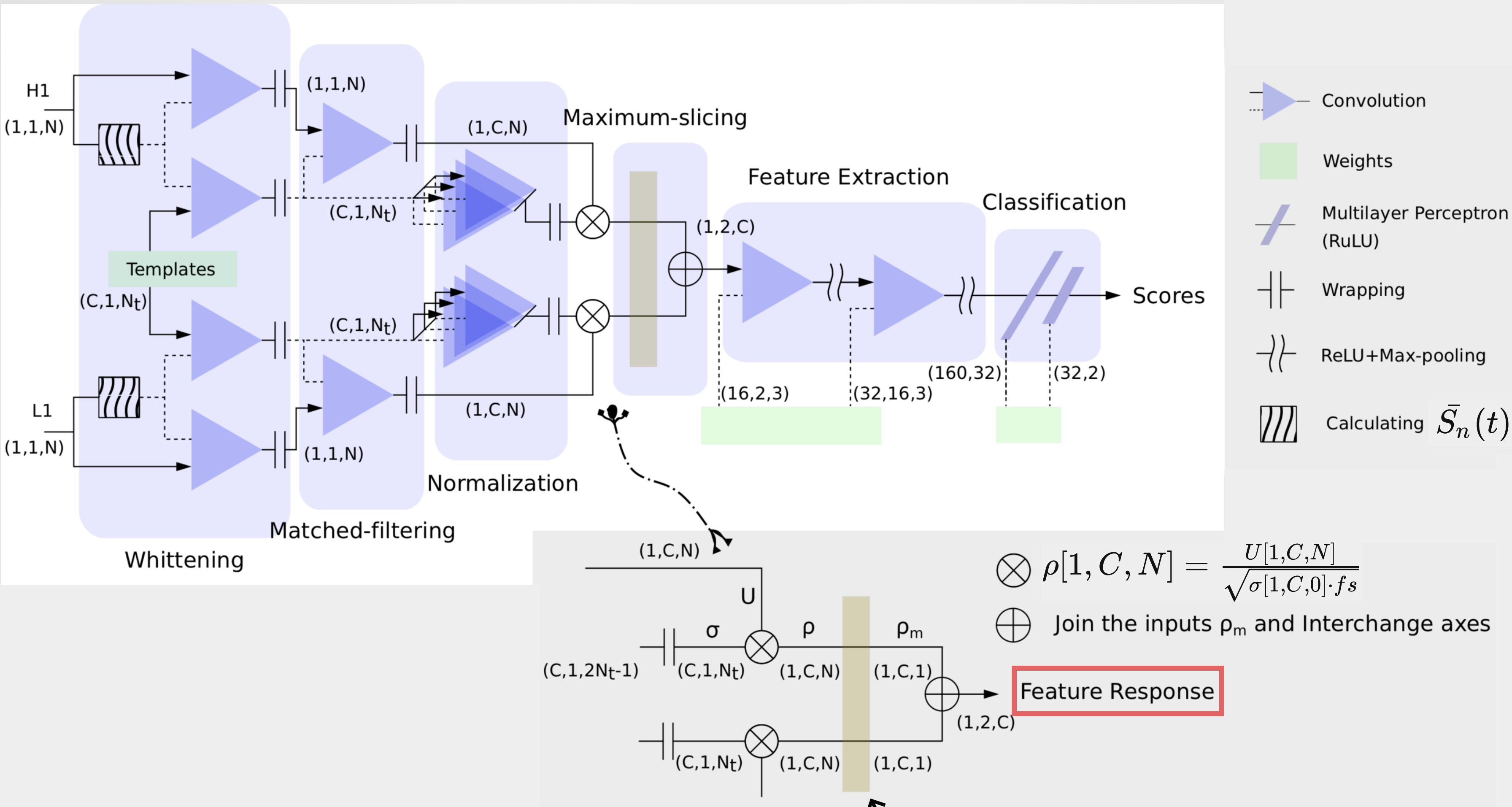
Is it matched-filtering?

"Matched-filtering (cross-correlation with the templates) can be regarded as a convolutional layer with a set of predefined kernels.

- *In practice, we use matched filters as an essential component in the first part of CNN for GW detection.*

MF-ConvNet Model

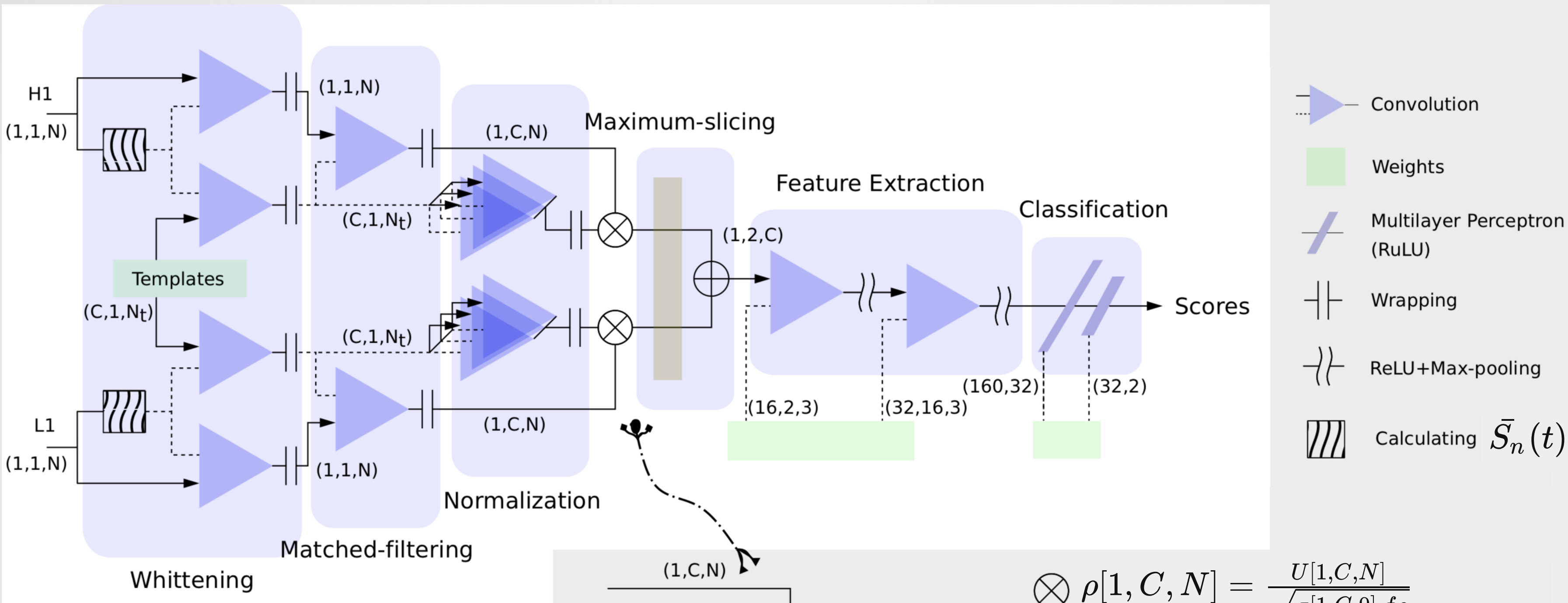
Architechture

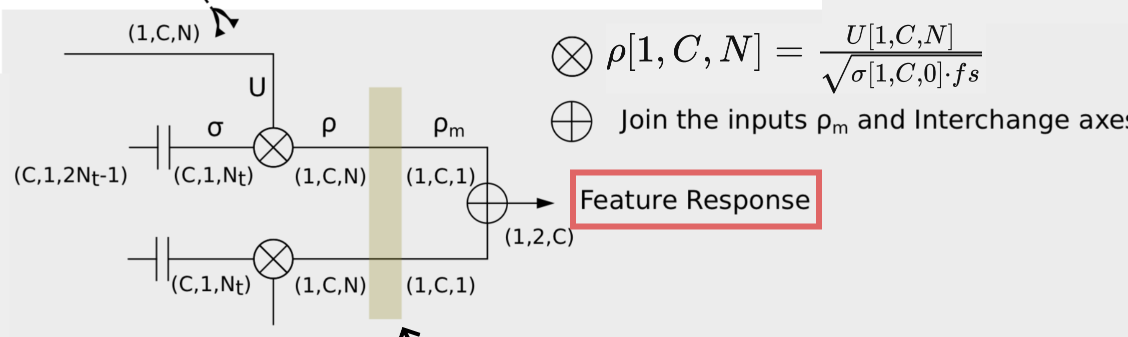


$$\rho_m[1, C, 1] = \max_N \rho[1, C, N]$$

MF-ConvNet Model

Architechture





$$\boxed{\begin{aligned} C_0 &= \arg\max_C \rho[1,C,N]\,, \\ N_0 &= \arg\max_N U[1,C_0,N] \end{aligned}}$$

$$\rho_m[1,C,1] = \max_N \rho[1,C,N]$$

In the meanwhile, we can obtain the **optimal time** N_0 (relative to the input) of feature response of matching by recording the location of the maxima value corresponding to the **optimal template** C_0

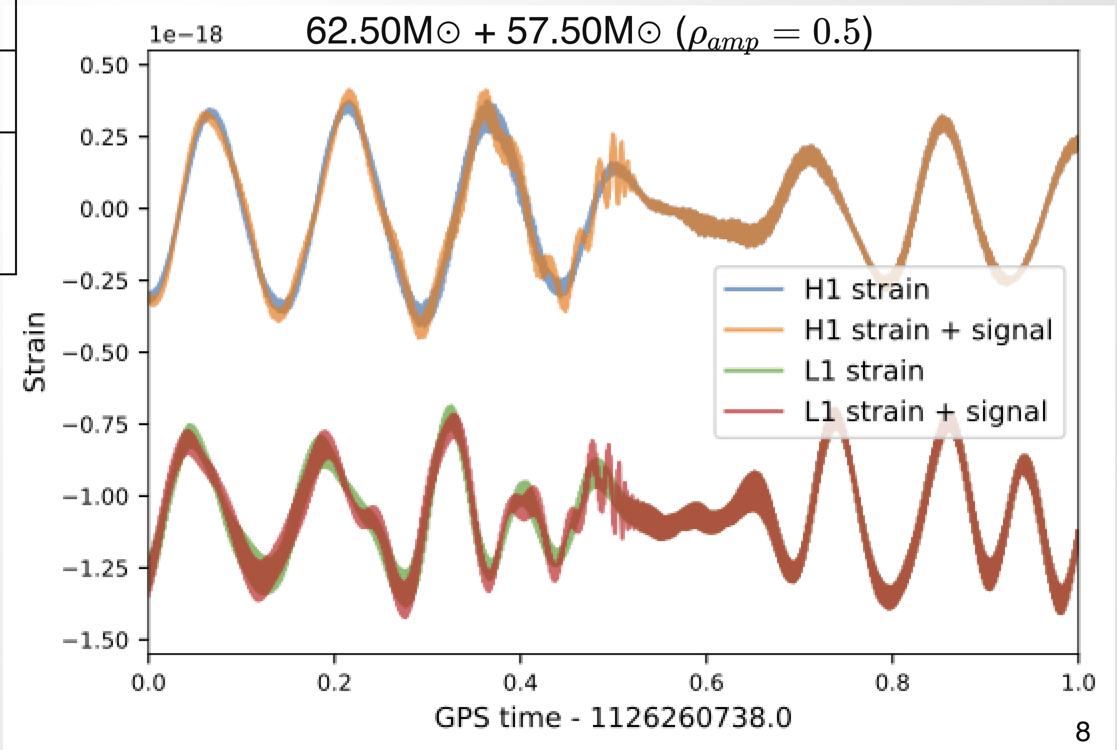
Experiments & Results

(In preprint)

- We use **SEOBNRE** model [Cao et al. (2017)] to generate waveform, we only consider **circular**, **spinless** binary black holes.

	template	**waveform (train/test)**
Number	35	1610
Length (s)	1	5
	equal mass	

- A sample as input



- The background noises for training/testing are sampled from a **closed** set (33x4096s) in the first observation run (O1) in the absence of the segments (4096s) containing the first 3 GW events.

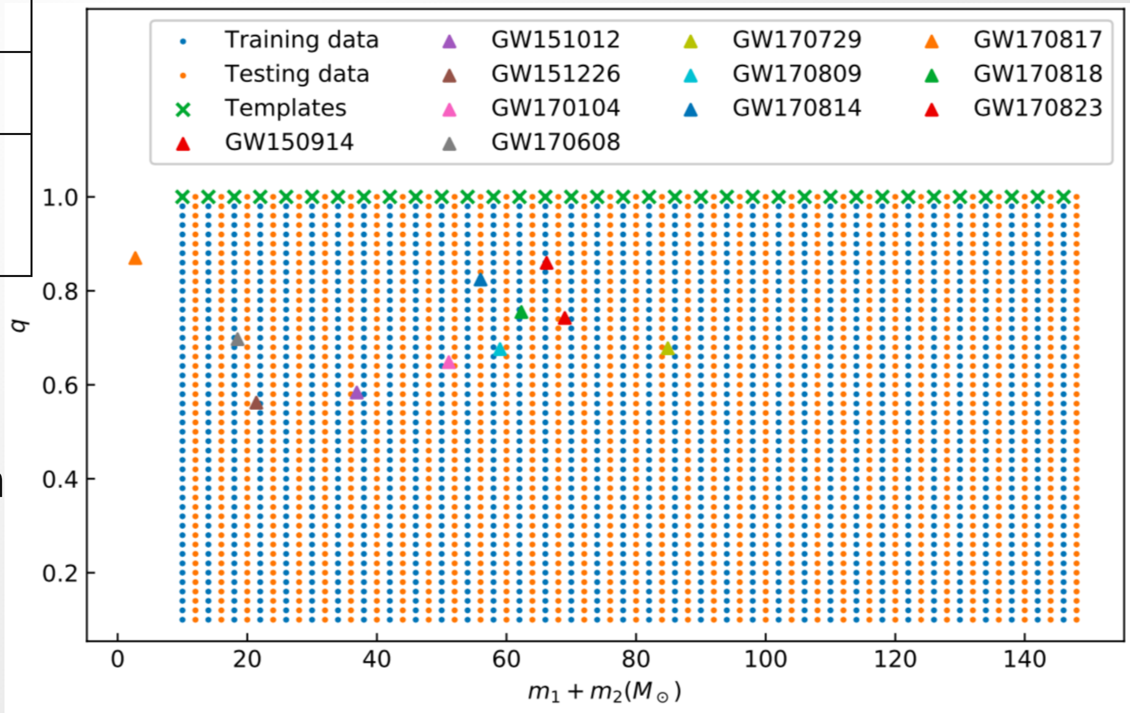
Experiments & Results

(In preprint)

- We use **SEOBNRE** model [Cao et al. (2017)] to generate waveform, we only consider **circular**, **spinless** binary black holes.

	template	**waveform (train/test)**
Number	35	1610
Length (s)	1	5
	equal mass	

- Mass distribution of dataset / templates / events

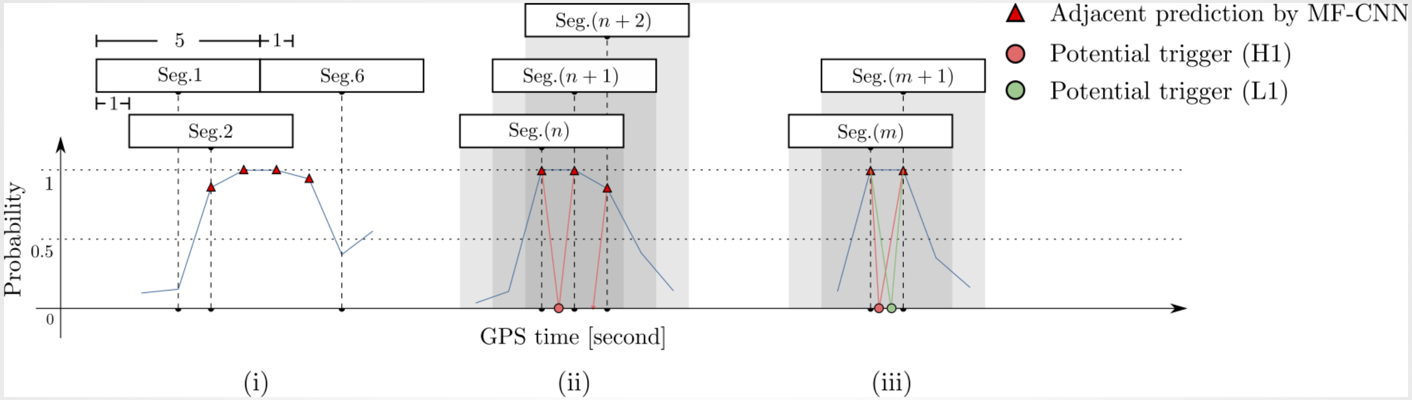


- The background noises for training/testing are sampled from a **closed** set (33x4096s) in the first observation run (O1) in the absence of the segments (4096s) containing the first 3 GW events.

Experiments & Results

Search methodology

(In preprint)

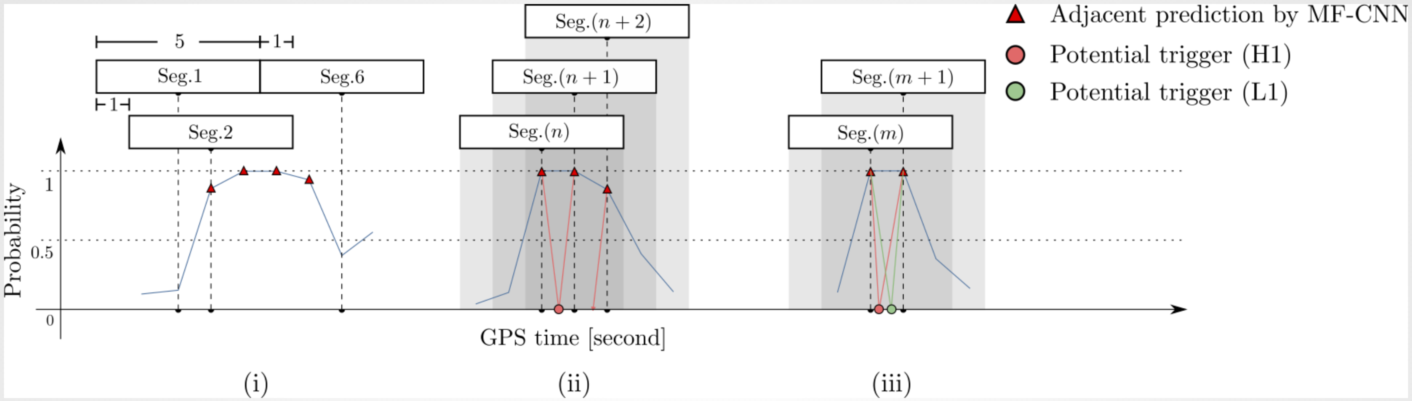


- Every 5 seconds segment as input of our MF-CNN with a step size of 1 second.
- The model can scan the whole range of the input segment and output a probability score.
- In the ideal case, with a GW signal hiding in somewhere, there should be **5** adjacent predictions for it with respect to a threshold.

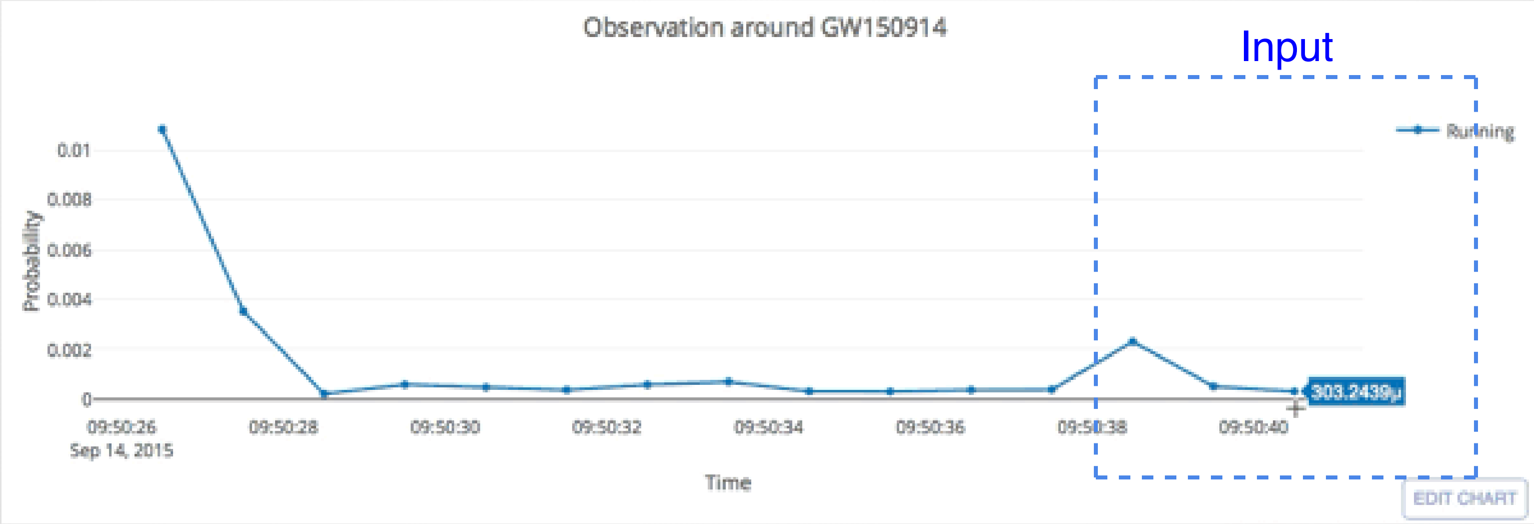
Experiments & Results

Search methodology

(In preprint)



- Every 5 seconds segment as input of our MF-CNN with a step size of 1 second.
- The model can scan the whole range of the input segment and output a probability score.
- In the ideal case, with a GW signal hiding in somewhere, there should be **5** adjacent predictions for it with respect to a threshold.

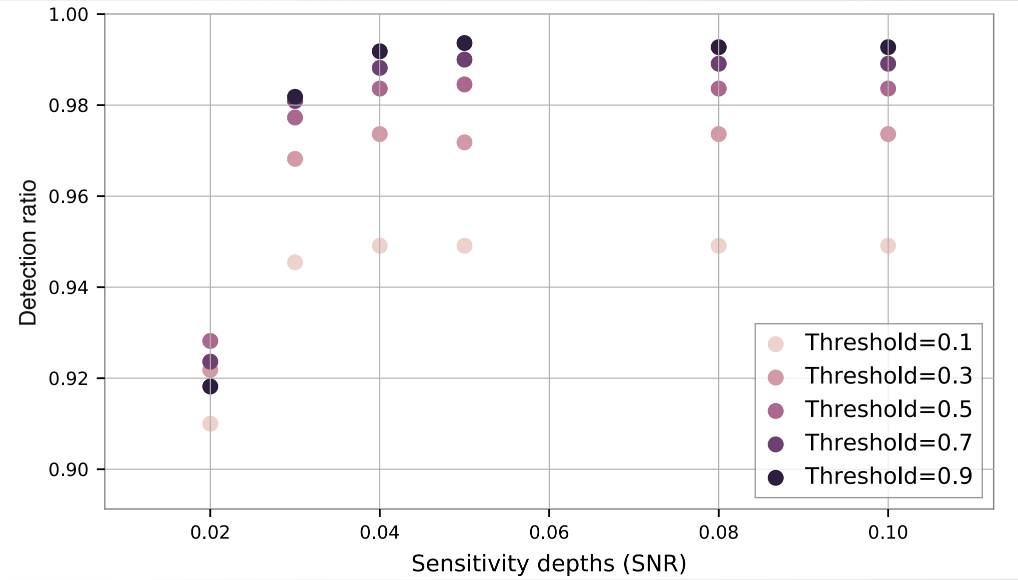


Experiments & Results

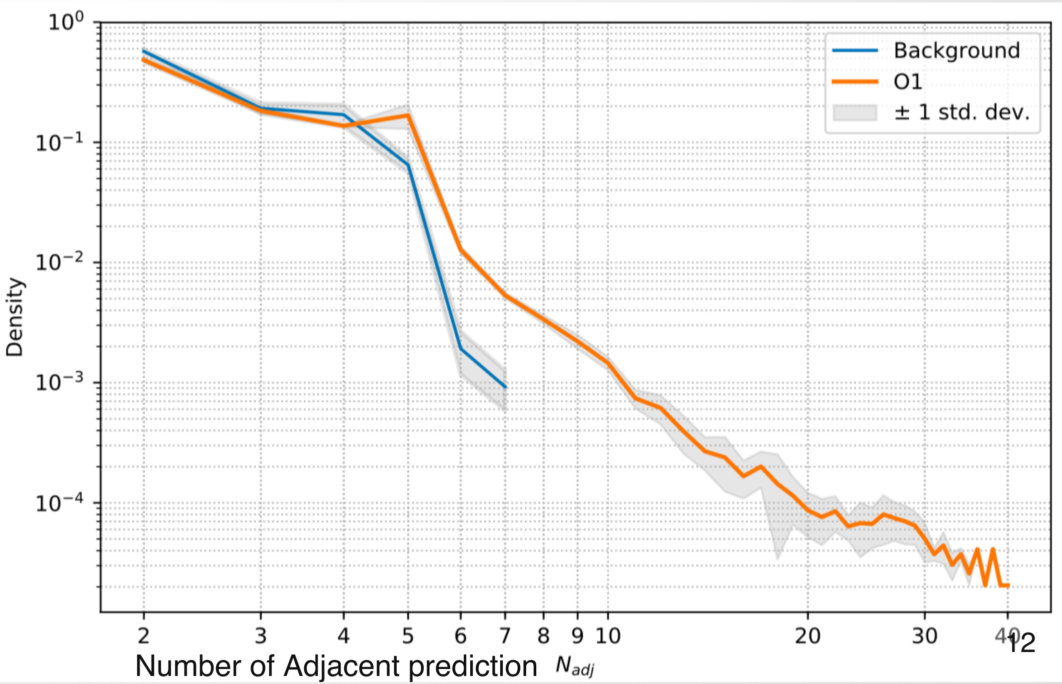
Population property on O1

(In progress)

- Sensitivity estimation
  - **Background**: using time-shifting on the closed set from real LIGO recordings in O1
  - **Injection**: random simulated waveforms



- Statistical significance on O1
  - Count a group of adjacent predictions as one "**trigger block**".
  - For pure background (non-Gaussian), monotone trend should be observed.
  - In the **ideal** case, with a GW signal hiding in somewhere, there should be **5 adjacent predictions** for it with respect to a threshold.

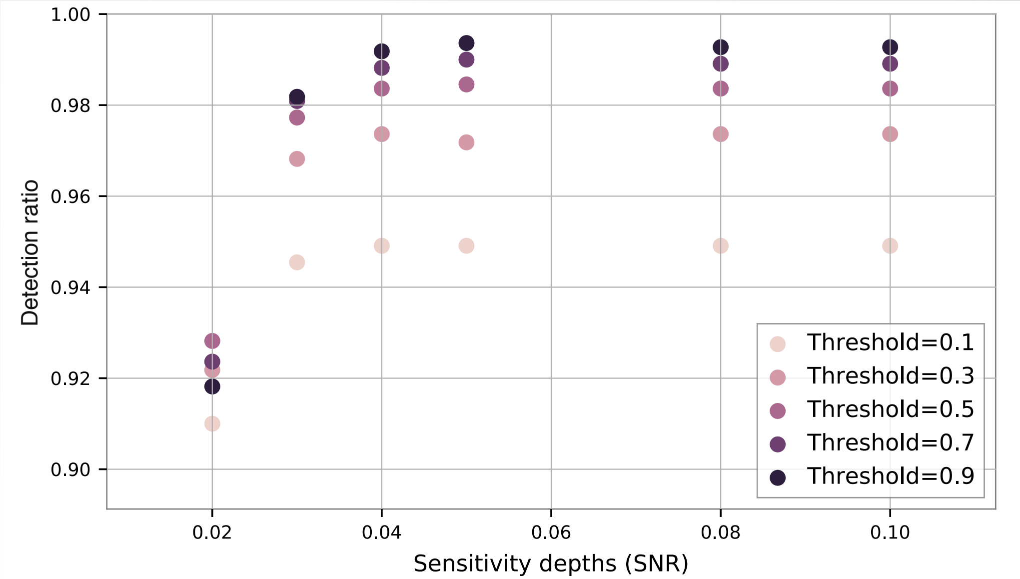


Experiments & Results

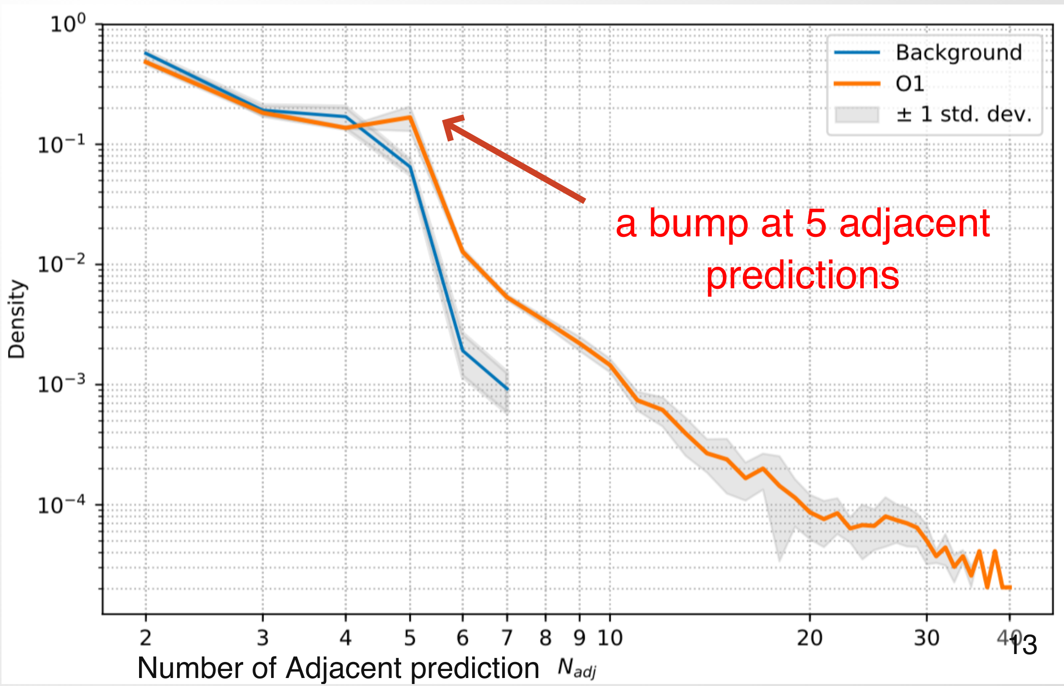
Population property on O1

(In progress)

- Sensitivity estimation
  - **Background**: using time-shifting on the closed set from real LIGO recordings in O1
  - **Injection**: random simulated waveforms



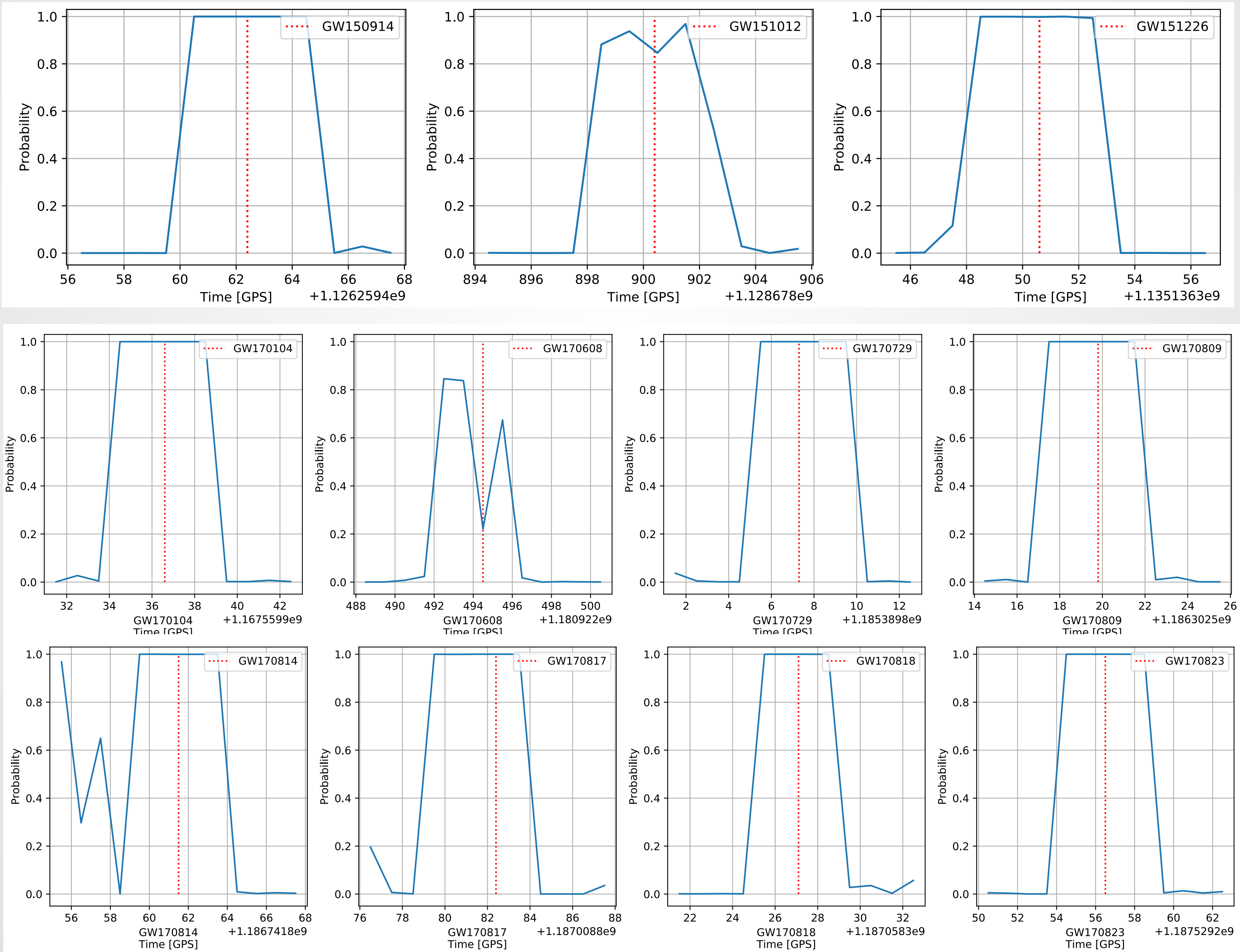
- Statistical significance on O1
  - Count a group of adjacent predictions as one "**trigger block**".
  - For pure background (non-Gaussian), monotone trend should be observed.
  - In the **ideal** case, with a GW signal hiding in somewhere, there should be **5 adjacent predictions** for it with respect to a threshold.



Experiments & Results

(In preprint)

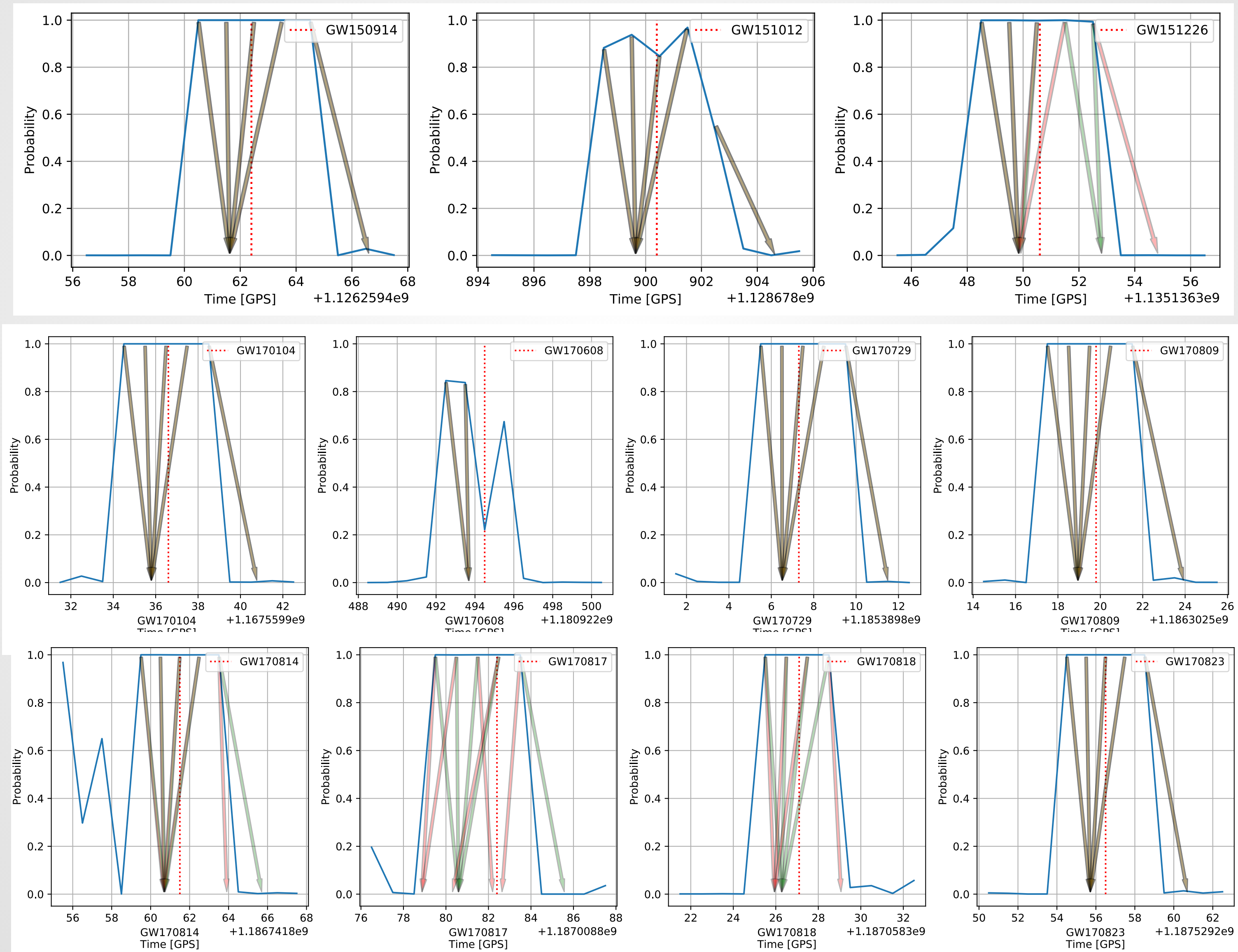
- Recovering all GW events in both O1 and O2



Experiments & Results

(In preprint)

- Recovering all GW events in both O1 and O2



Summary

- Some benefits from MF-CNN architechure
  - Simple configuration for GW data generation
  - Almost no data pre-processing
  - Works on **non-stationary** background
  - Easy parallel deployments, multiple detectors can be benefit a lot from this design
  - More templates / smaller steps for searching can improve further
- Main understanding of the algorithms:
  - GW templates are used as **likely features** for matching
  - **Generalization** of both matched-filtering and neural networks
  - Matched-filtering can be **rewritten** as convolutional neural layers

Summary

- Some benefits from MF-CNN architechure
  - Simple configuration for GW data generation
  - Almost no data pre-processing
  - Works on **non-stationary** background
  - Easy parallel deployments, multiple detectors can be benefit a lot from this design
  - More templates / smaller steps for searching can improve further
- Main understanding of the algorithms:
  - GW templates are used as **likely features** for matching
  - **Generalization** of both matched-filtering and neural networks
  - Matched-filtering can be **rewritten** as convolutional neural layers

`Thank you for your attention!`