

# **cWB+GMM**

**2023 AXIS Lab report**

**2023 HEP workshop | Jiyeon Sun**



# Gravitational Wave

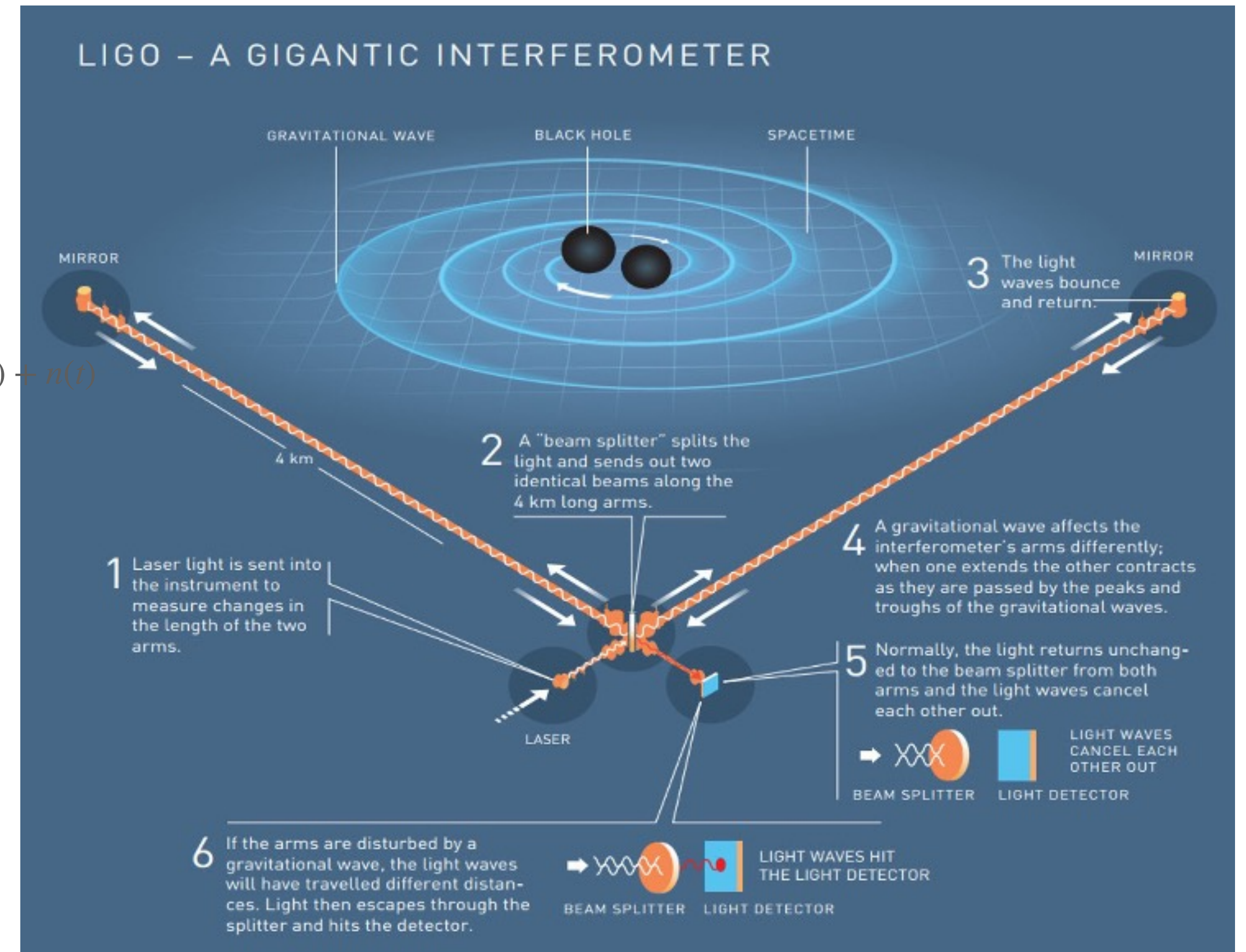
## What is Gravitational Wave and how is it measured?

- Gravitational waves are the perturbations in space-time metric triggered by the movement of massive objects

$$s(t) = \xi(\theta, \phi)h(t) + n(t)$$

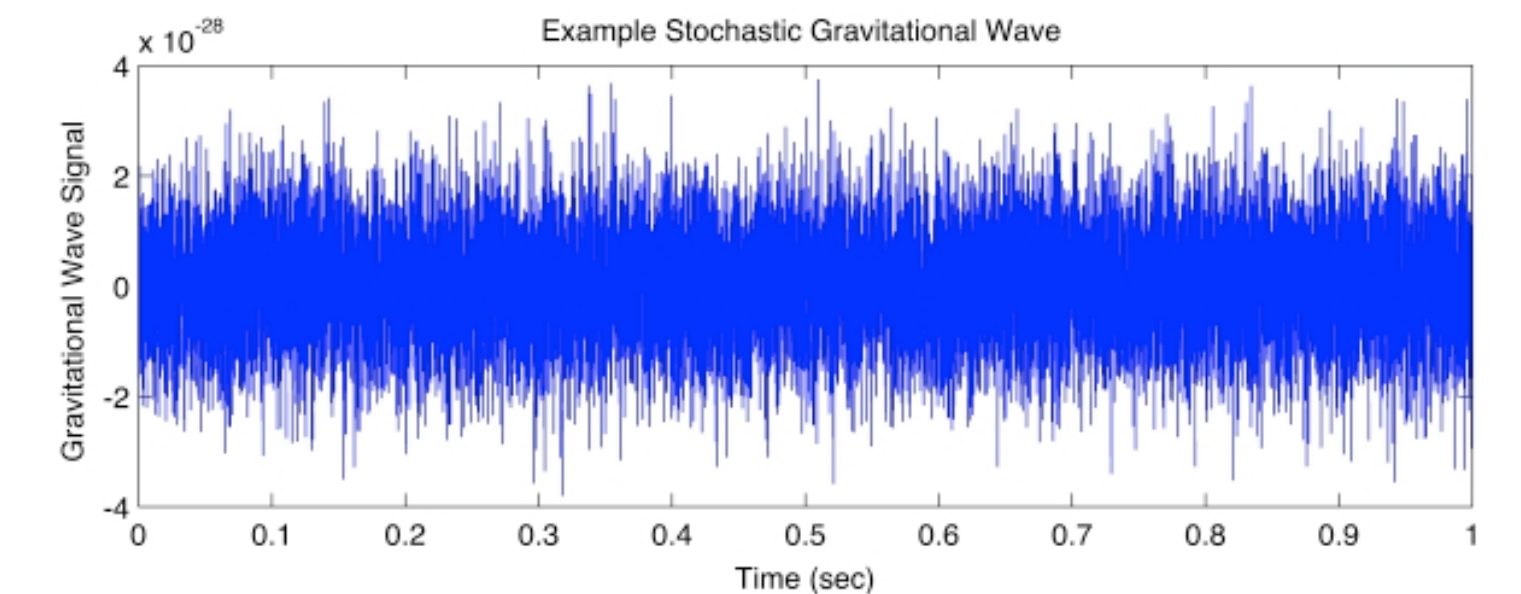
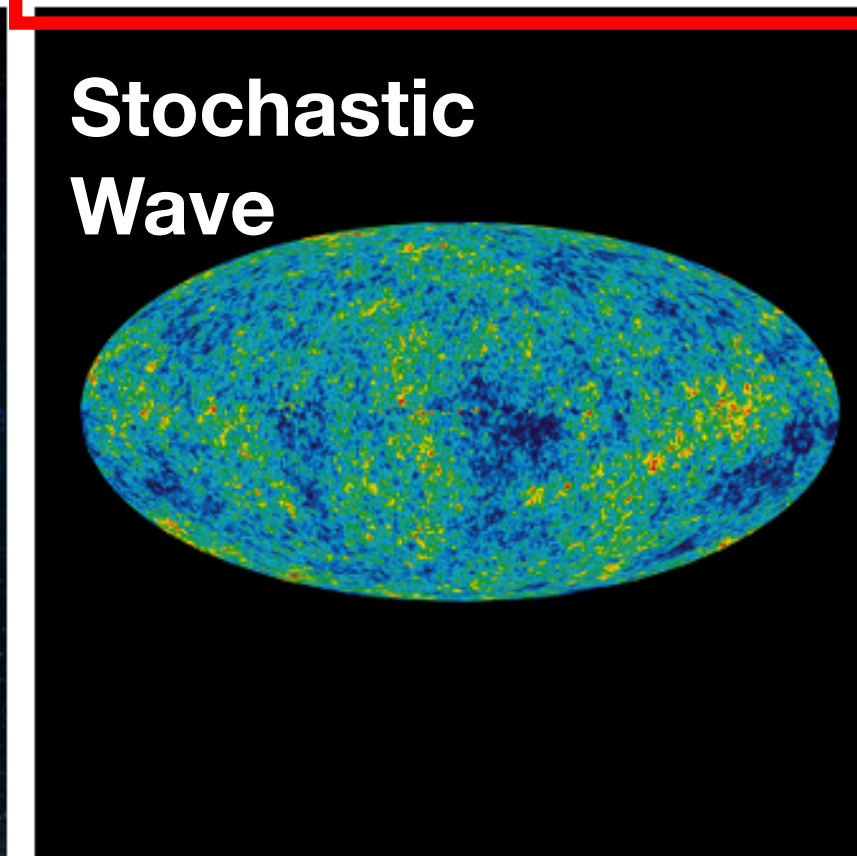
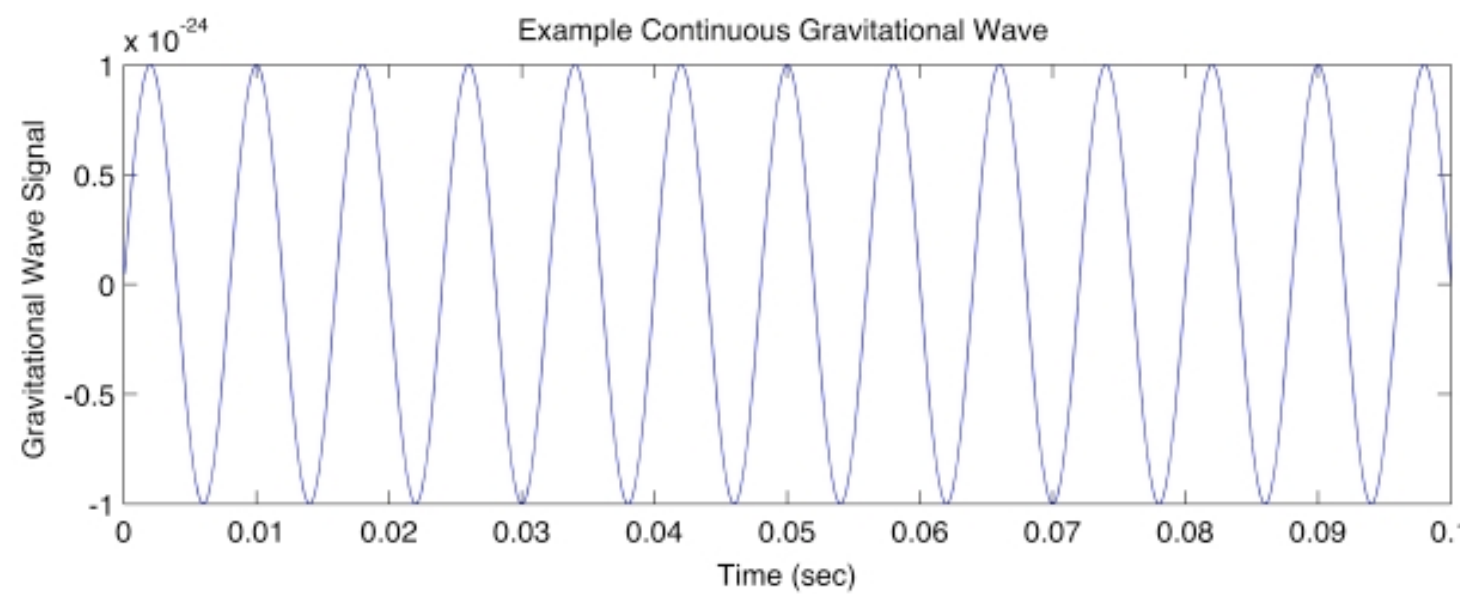
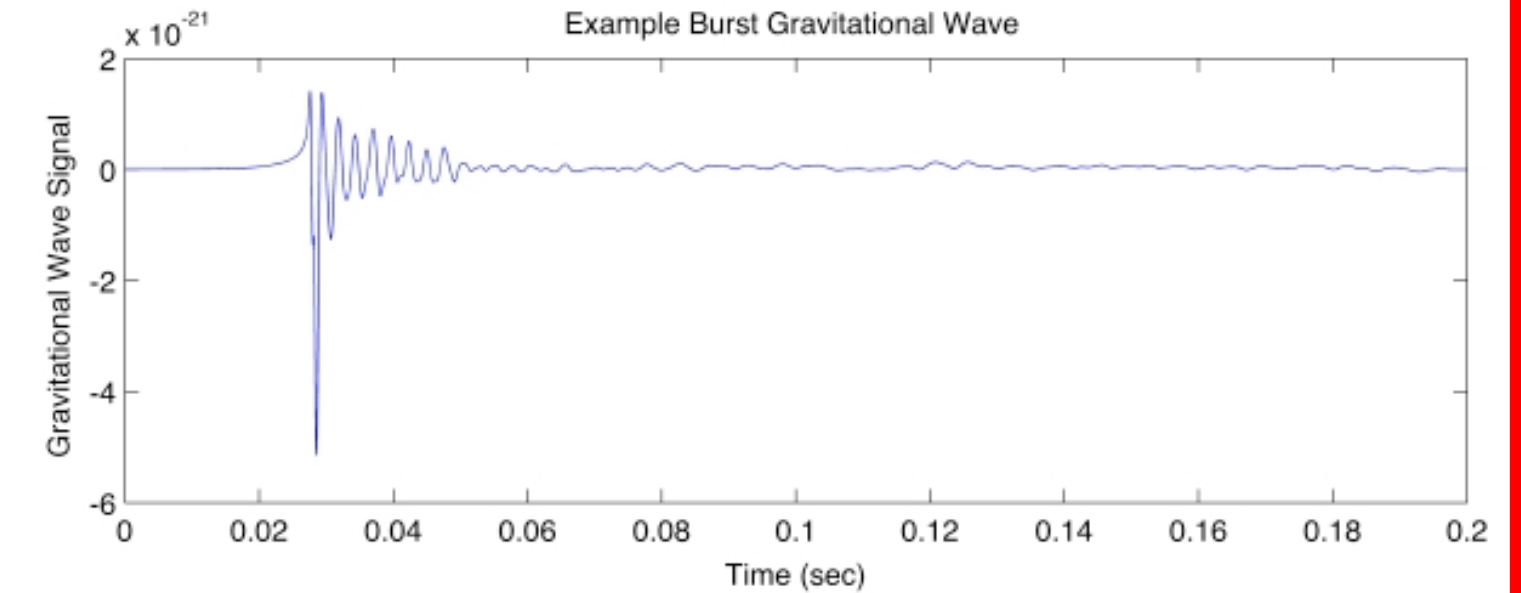
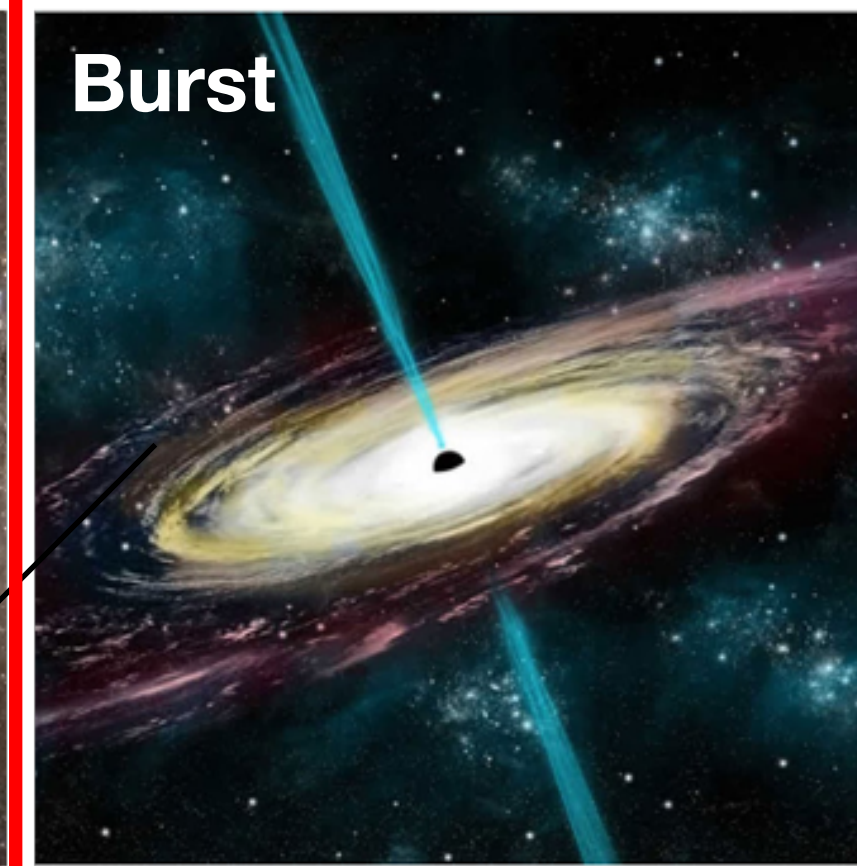
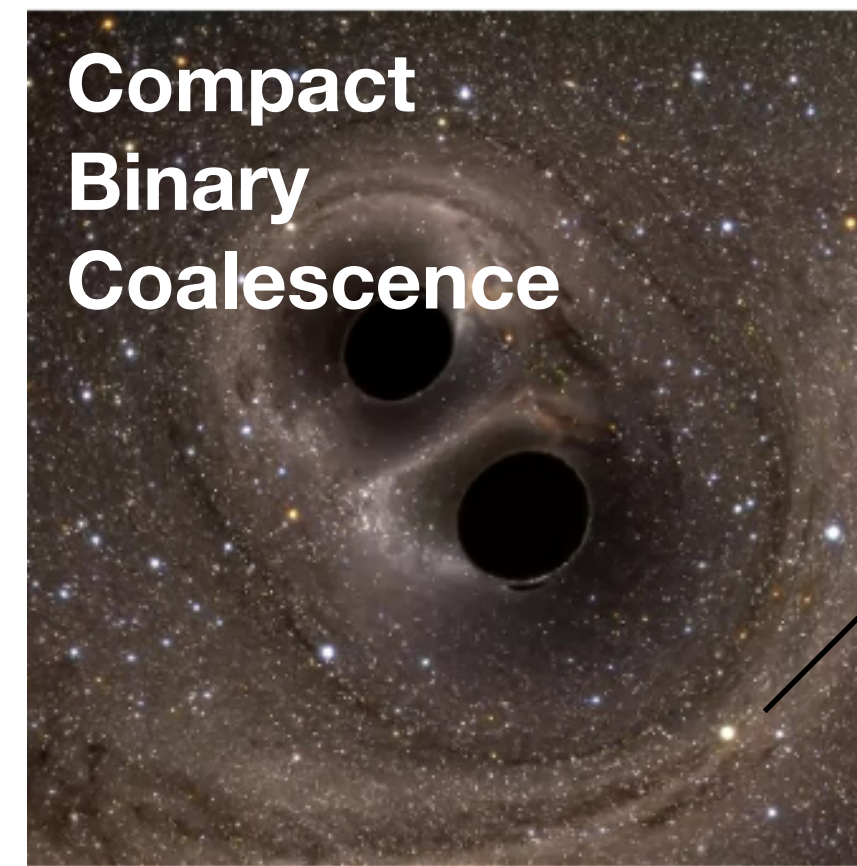
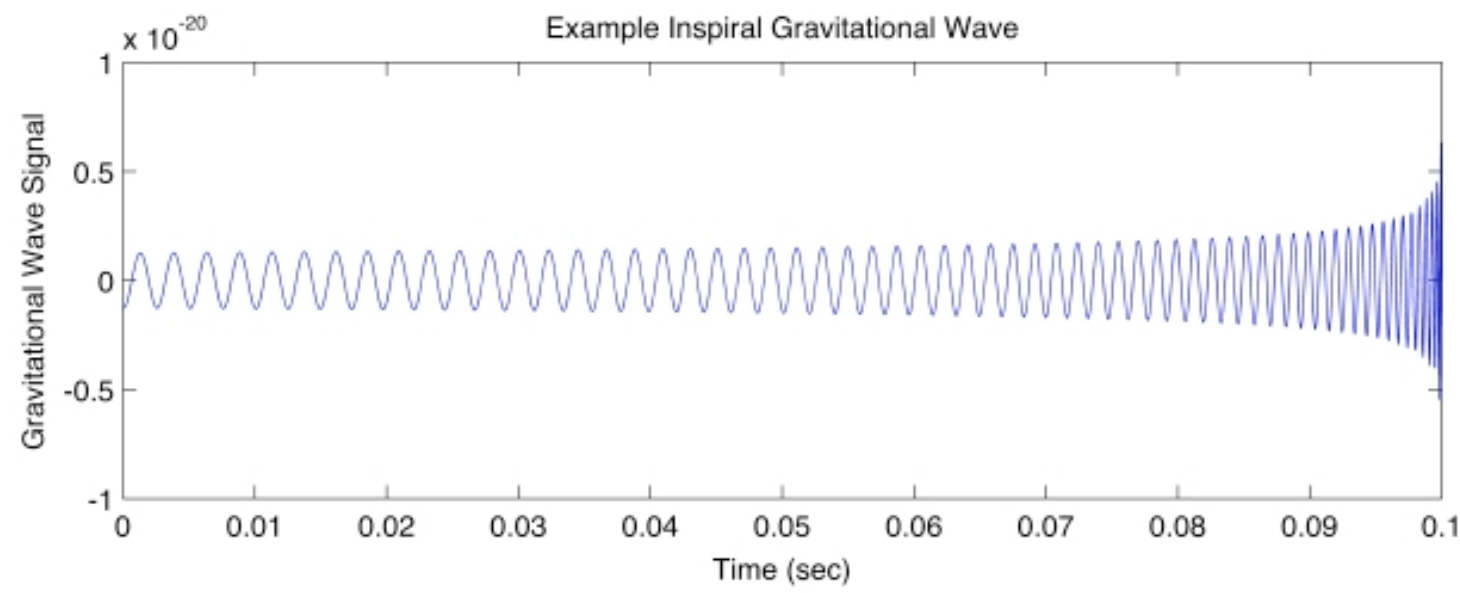
- Strain is measured by gigantic interferometers

$$s(t) = \xi(\theta, \phi)h(t) + n(t)$$



# Burst

## Unmodelled and Transient GW signal



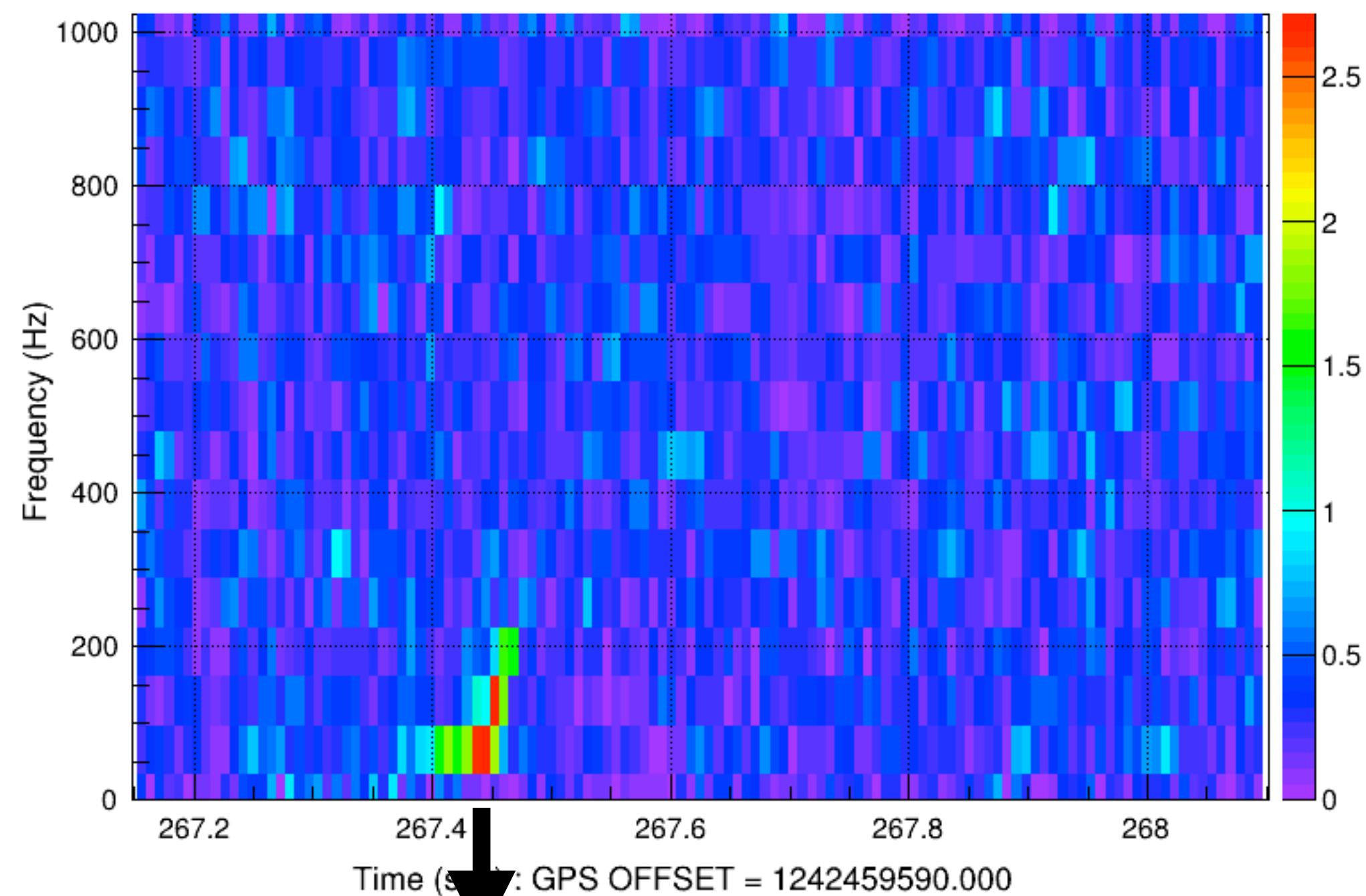
# coherent WaveBurst

## Algorithm for detecting generic transient signal



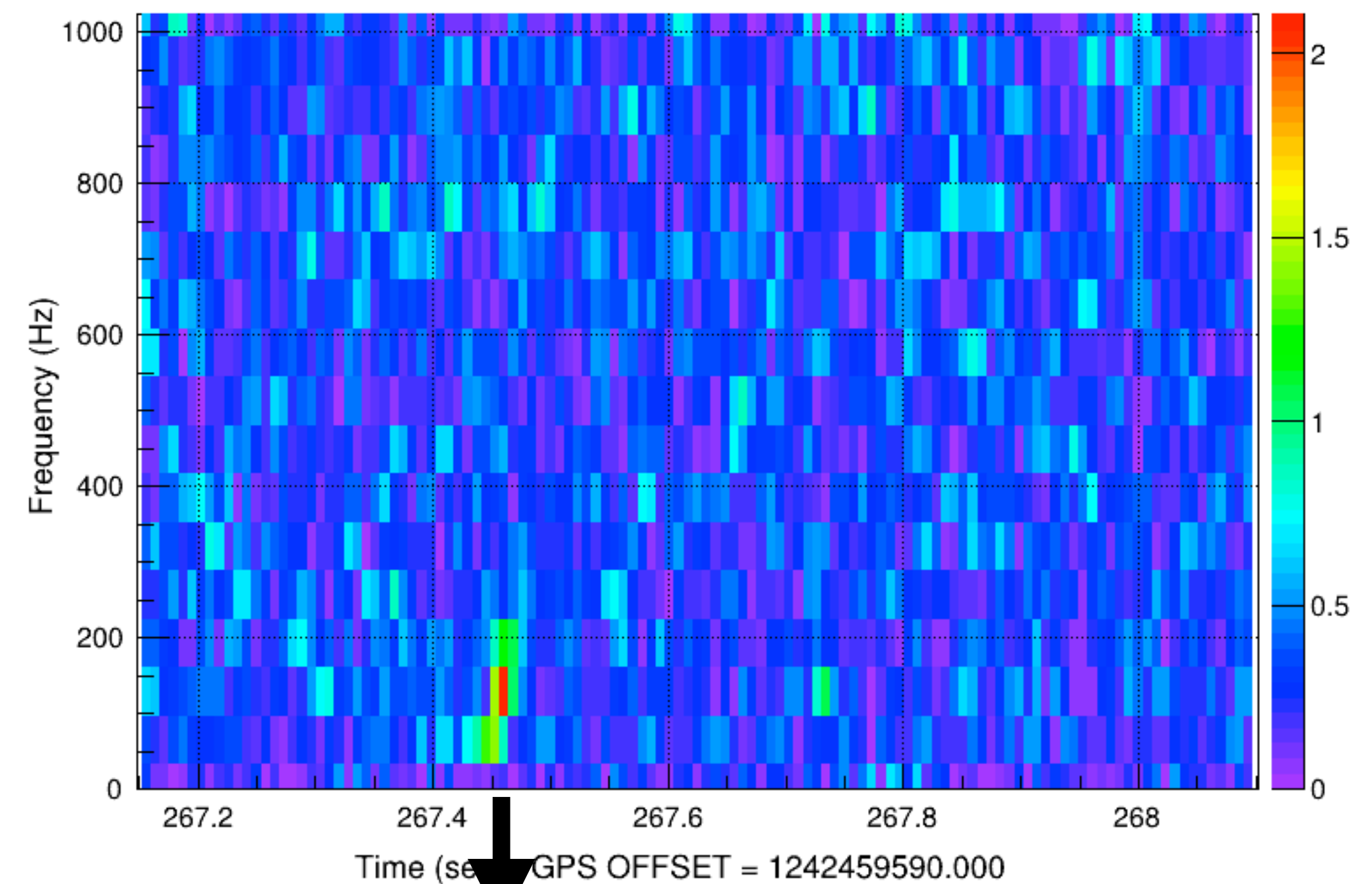
- cWB identifies the coherent excess power in multiple detectors

Scalogram ( $\sqrt{(E_{00}+E_{90})/2}$ )



1242459857.4385 from LIGO Livingston

Scalogram ( $\sqrt{(E_{00}+E_{90})/2}$ )

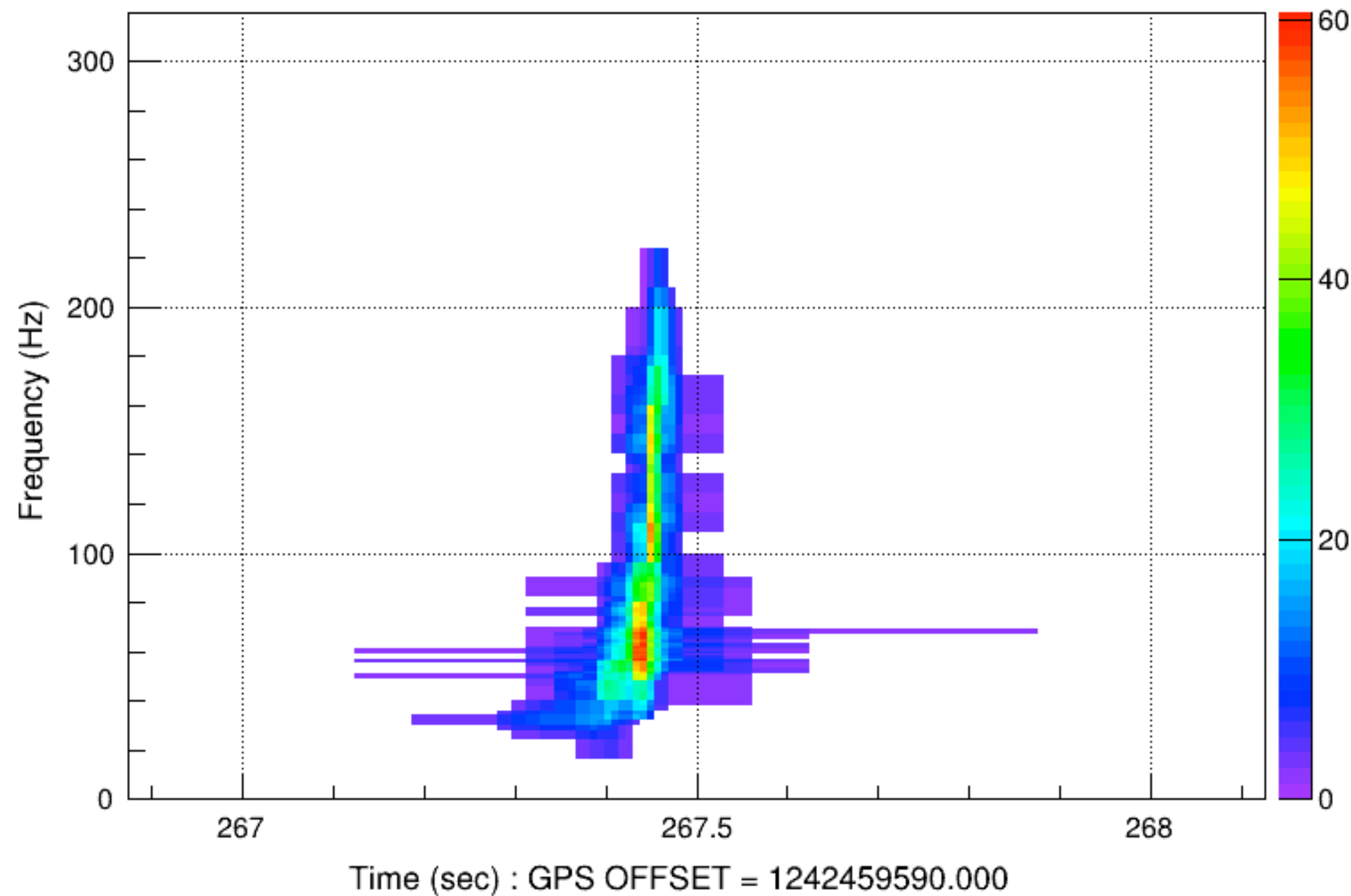


1242459857.4438 from LIGO Hanford

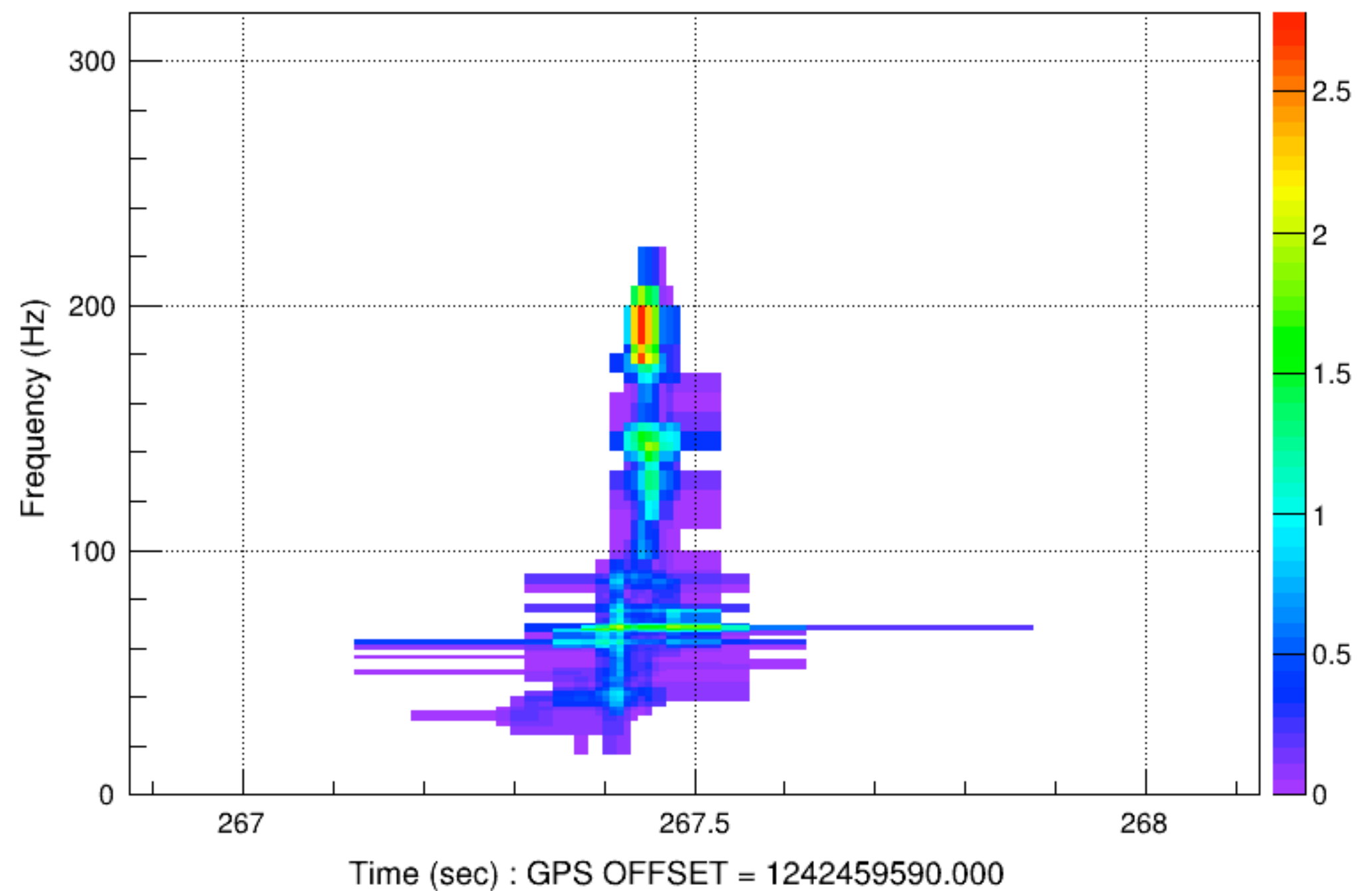
# coherent WaveBurst

## Attributes to capture the properties of the identified triggers

Likelihood 584 - dt(ms) [7.8125:250] - df(hz) [2:64] - npix 134



Null 21 - dt(ms) [7.8125:250] - df(hz) [2:64] - npix 130



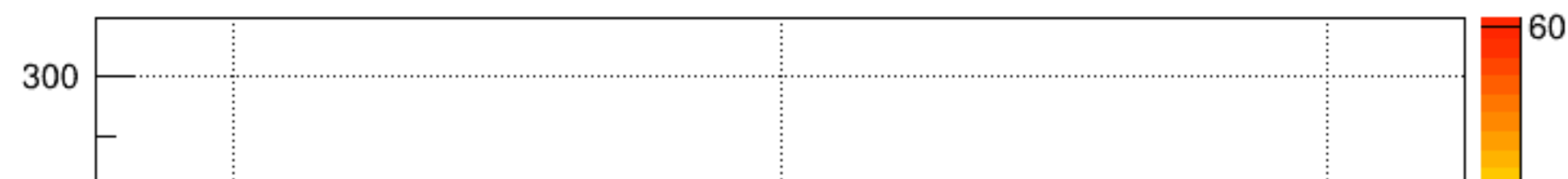
- Likelihood  
= Coherent Energy + Incoherent Energy  
>> GW signal component of the total E

- Null  
>> Background noise component  
of the total E

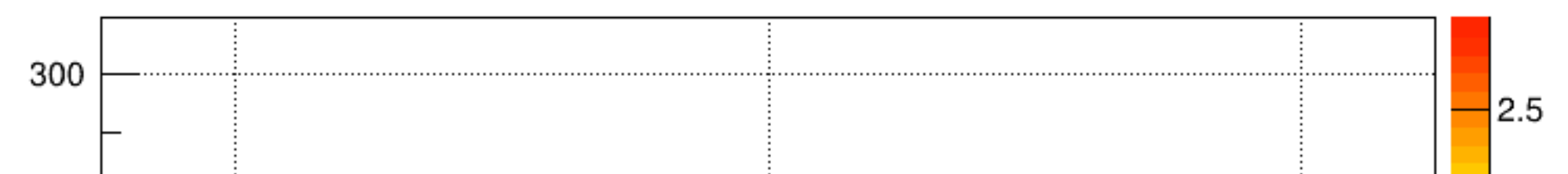
# coherent WaveBurst

## Attributes to capture the properties of the identified triggers

Likelihood 584 - dt(ms) [7.8125:250] - df(hz) [2:64] - npix 134

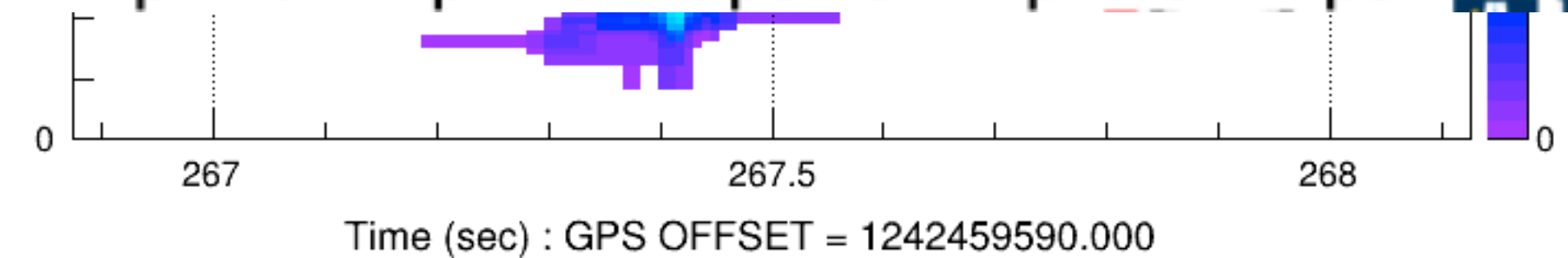
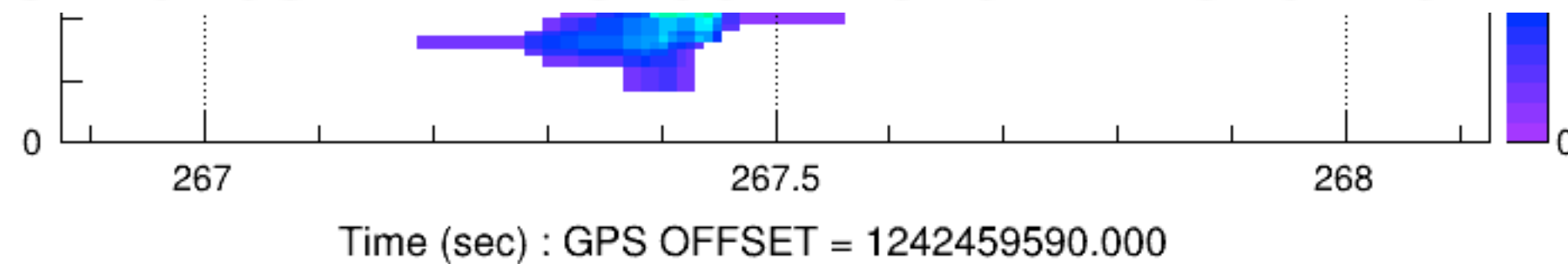


Null 21 - dt(ms) [7.8125:250] - df(hz) [2:64] - npix 130



### cWB trigger attributes (!!)

```
mass0 mass1 spin0 spin1 spin2 spin3 spin4 spin5 time0 lag0 lag1 lag2 slag0 slag1 slag2 rho0  
rho1 gnet anet netcc0 netcc1 netcc2 netcc3 neted0 neted1 neted2 neted3 neted4 likelihood  
norm_penalty Ecor factor Qveto0 Qveto1 frequency0 frequency1 dtL dtH reconstructed_snr  
null0 null1 strain0 strain1 hrss0 hrss1 noise0 noise1 duration0 duration1 volume0 volume1  
size0 size1 ecor bandwidth0 bandwidth1 snr0 snr1 xSNR0 xSNR1 sSNR0 sSNR1 iSNR0 iSNR1 ioSNR0  
ioSNR1 oSNR0 oSNR1 Lveto0 Lveto1 Lveto2 chirp0 chirp1 chirp2 chirp3 chirp4 chirp5
```

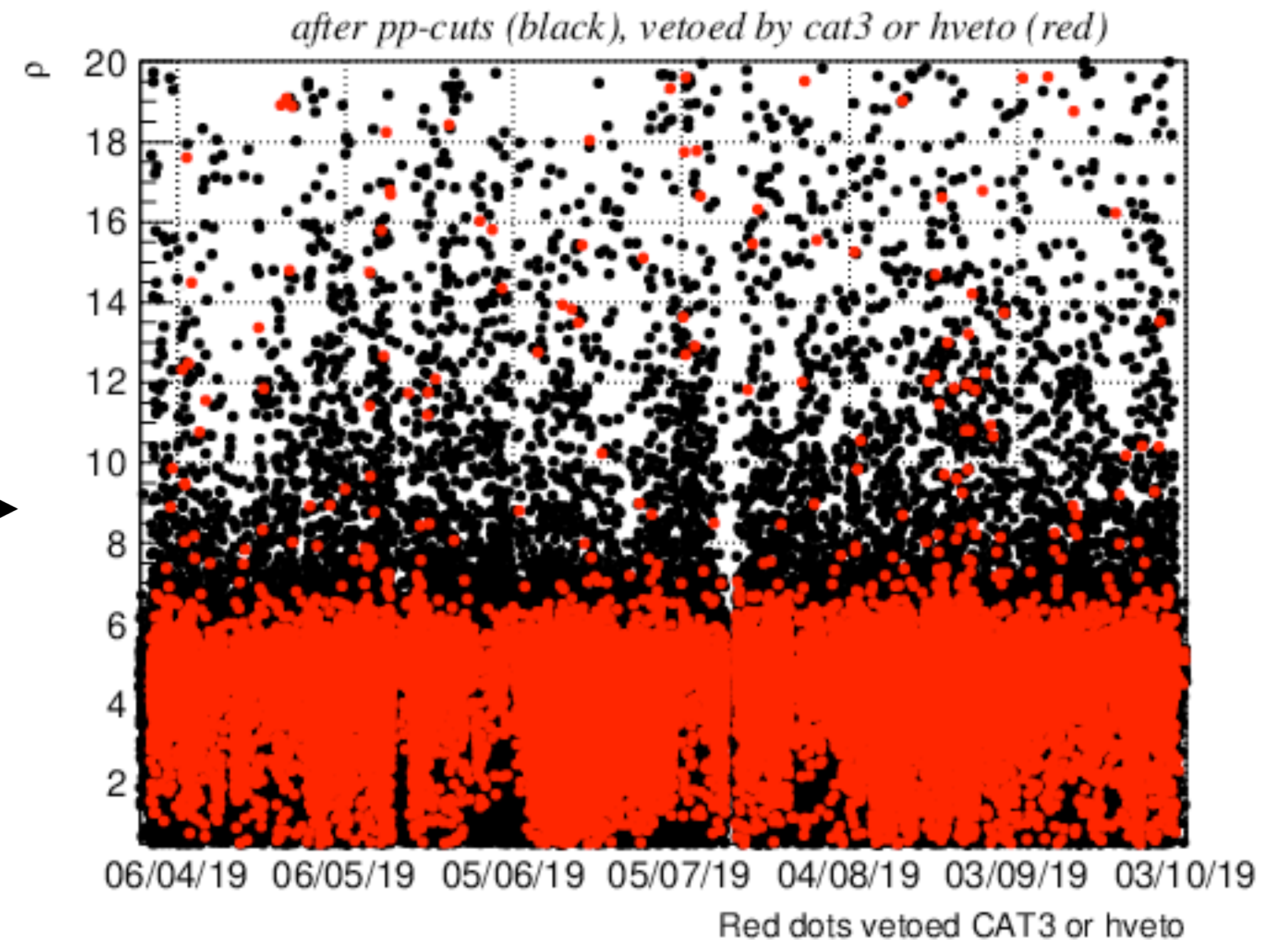
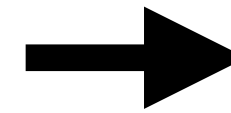
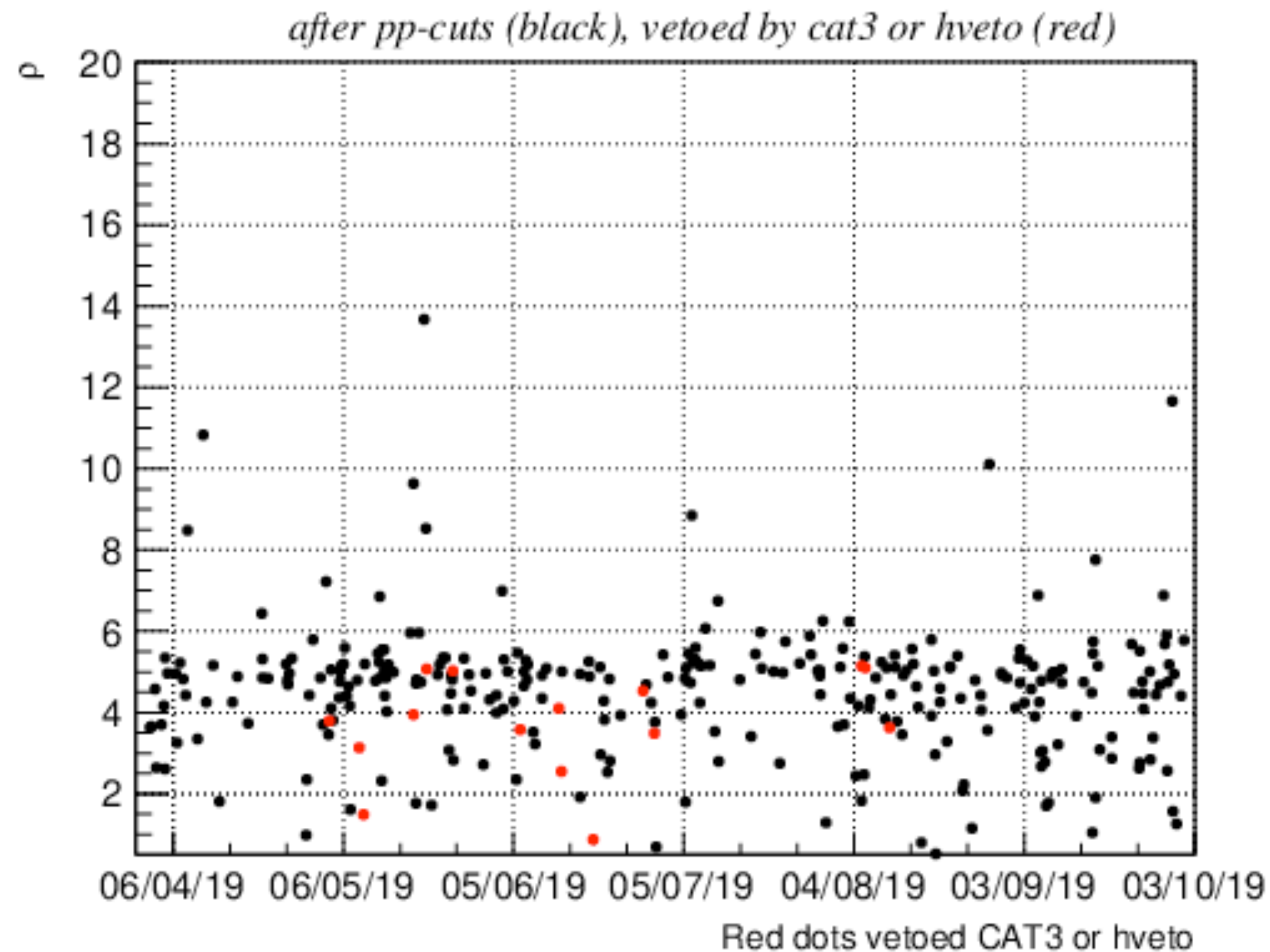


- Likelihood  
= Coherent Energy + Incoherent Energy  
>> GW signal component of the total E

- Null  
>> Background noise component  
of the total E

# coherent WaveBurst

## Background analysis using unphysical Time shift



- Background analysis gives glitches that are identified as event triggers

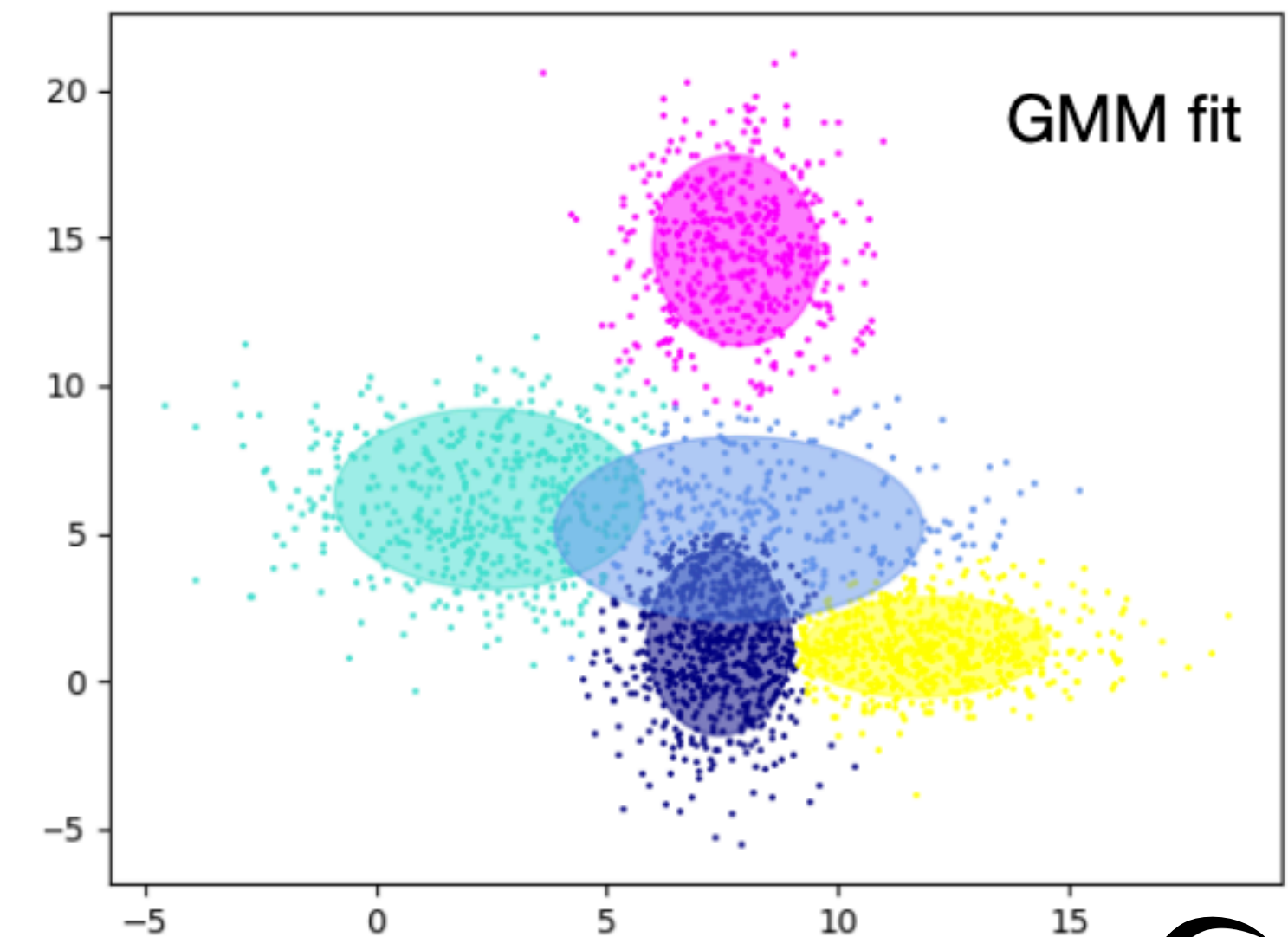
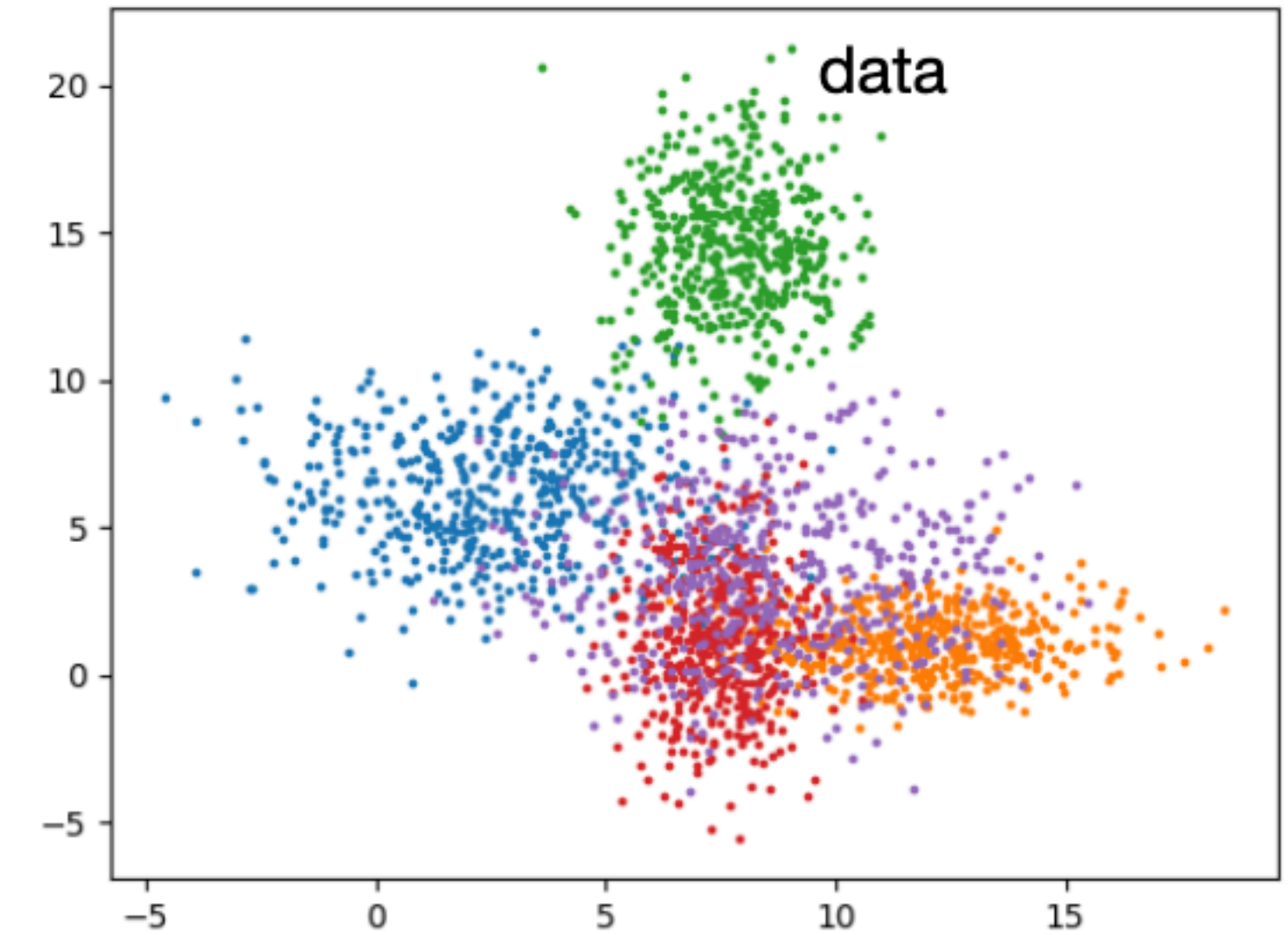
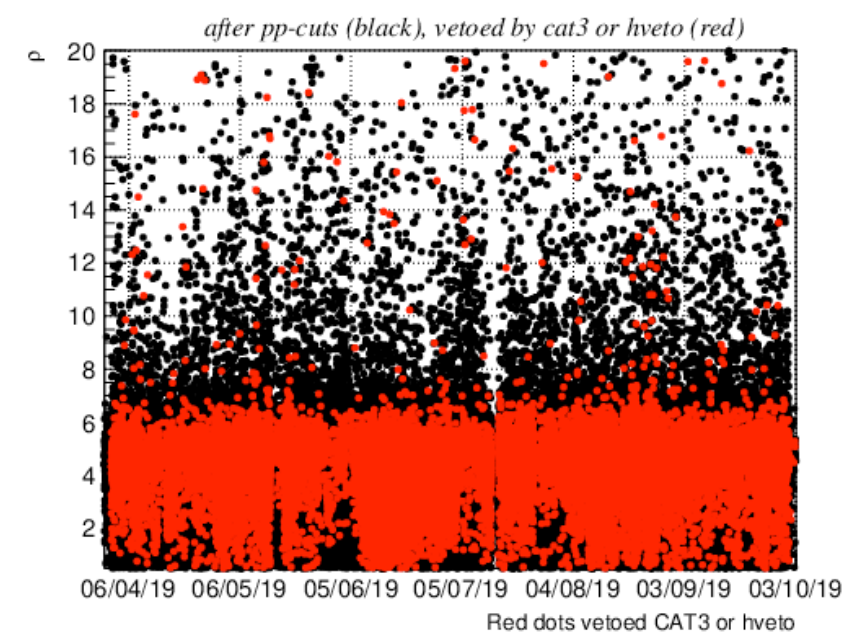
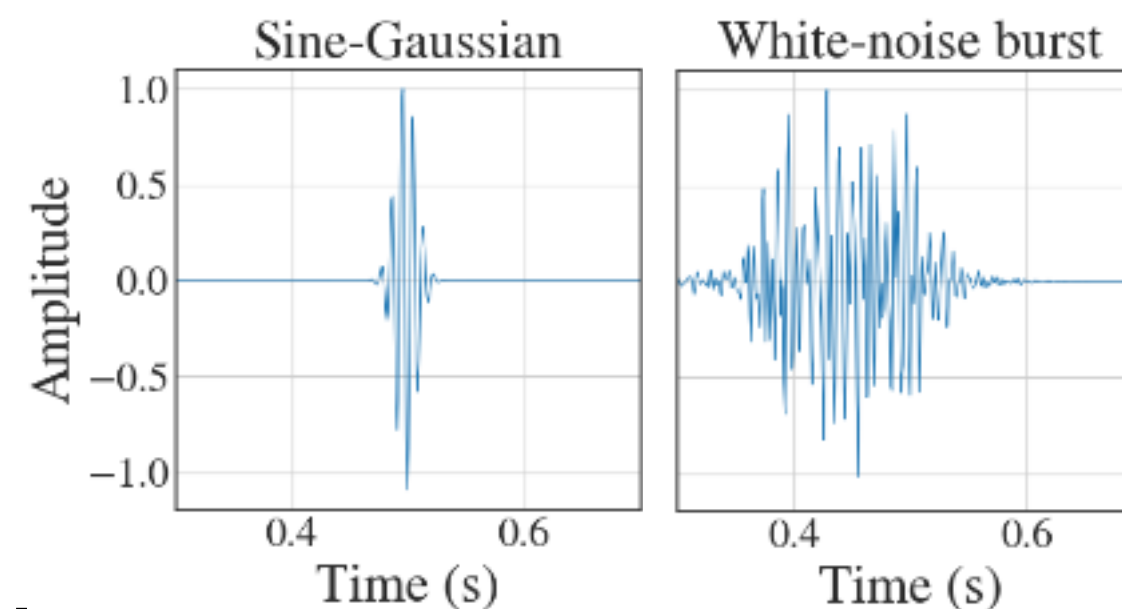
# Gaussian Mixture Modelling

Enhance the detection efficiency using machine learning

- GMM uses a superposition of Gaussian functions to characterize the parameter space covered by a set of data points.

$$p(\mathbf{x}) = \sum_{j=1}^K w_j \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$$

- One GMM to model the attribute space by simulated signals
- Another GMM to model the attribute space by time-shifted background triggers





# Gaussian Mixture Modelling

## Signal-Glitch classification

- Use trained GMMs to calculate the log-likelihood

$$\ln \mathcal{L} = \sum_{i=1}^n \ln p(\vec{x}_i | \Theta) = \sum_{i=1}^n \ln \left\{ \sum_{j=1}^K w_j N(\vec{x}_i | \mu_j, \Sigma_j) \right\}$$

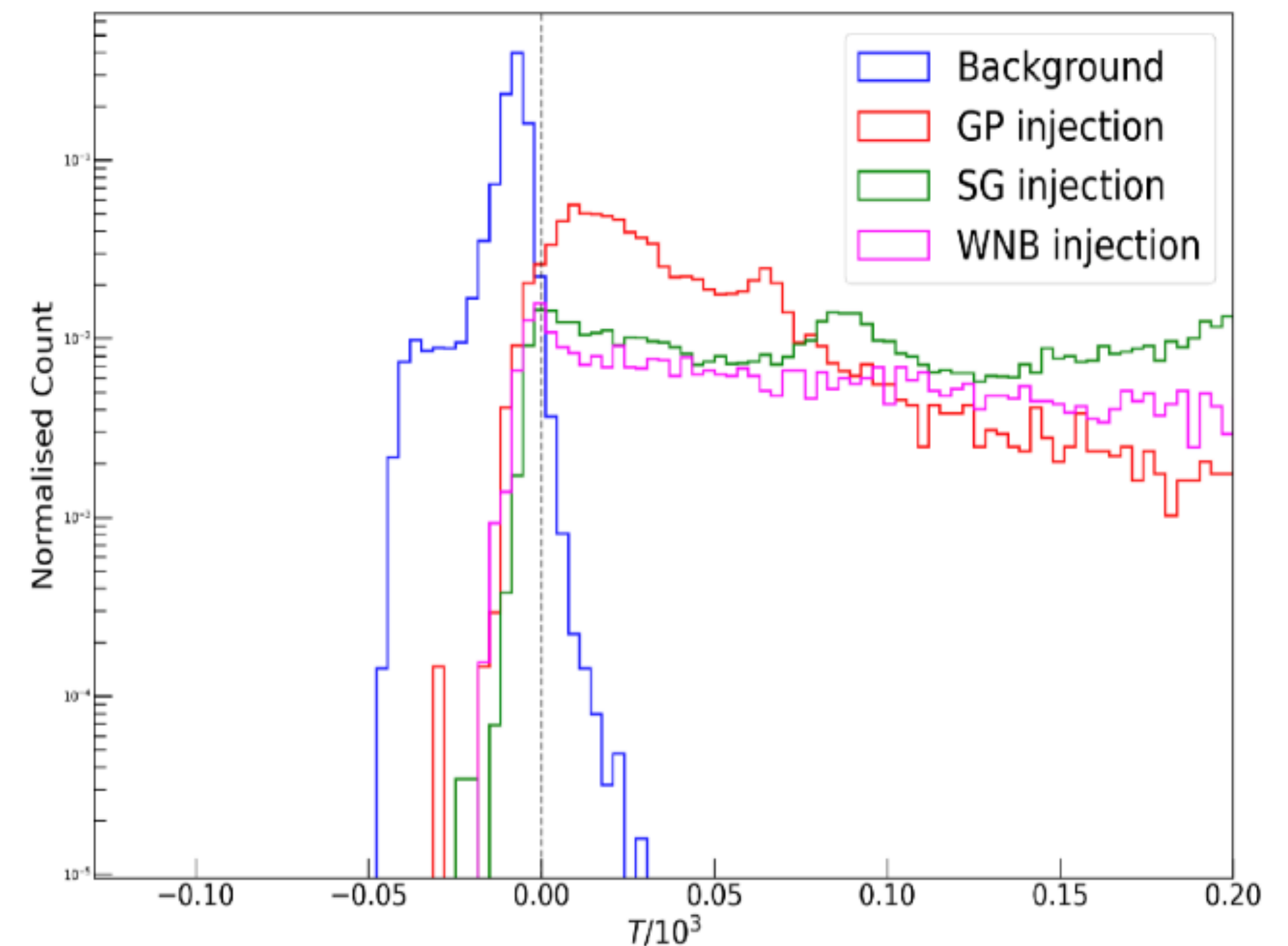
$$\Theta = \phi_j, \mu_j, \Sigma_j, \{j = 1 \dots K\}$$

- With 2 models representing signal and glitch, the GMM detection statistic for each trigger is calculated

$$W = \ln(\hat{\mathcal{L}}) |_{\hat{K}}$$

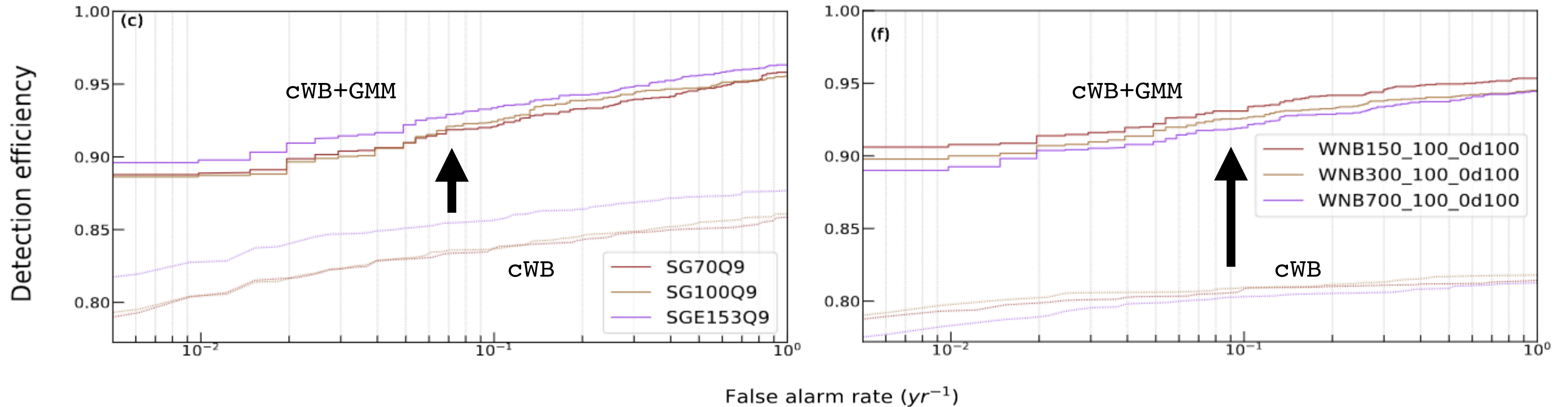
$$T = W_s - W_g$$

GMM detection statistic distribution for O3a short gravitational-wave bursts



# Gaussian Mixture Modelling

Result : cWB vs cWB+GMM and updates ...



- GMM enhances the detection performance of standard cWB

# cWB+GMM

## Summary

- The coherent WaveBurst covers a broad range of parameter space without requiring model waveforms
- GMM method enhances the detection efficiency of cWB algorithm
- Updates for O4 run are being prepared
  - New method to determine the optimal number of Gaussians
  - Test on parallelizing the training process to use more training data which was not possible due to high computational cost