Application of Gaussian Mixture Modelling to short all-sky burst search

Leigh Smith

Collaborators: Gayathri V., Archana Pai, Ik Siong Heng, Dixeena Lopez, Chris Messenger

G2Net WG1 Meeting Valenica

April 2022

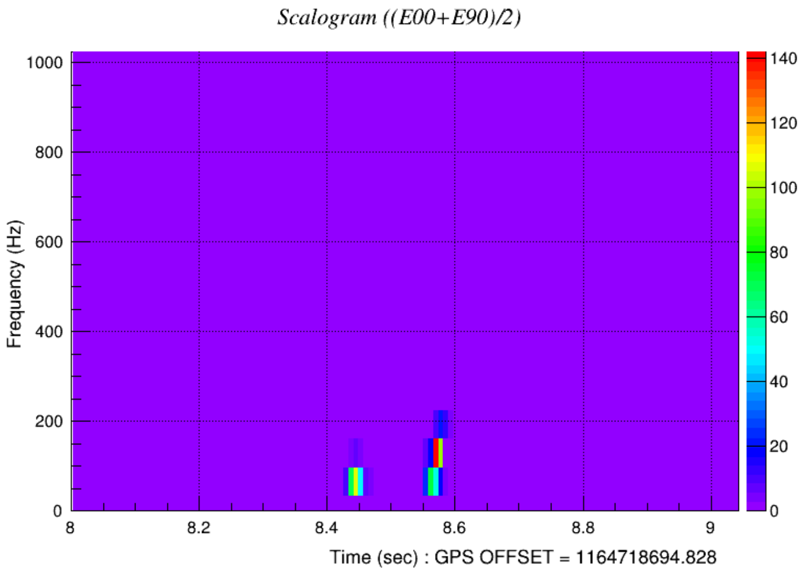
All-sky Short Burst Search

- Burst events: GW transients with generic, un-modelled morphologies
- All-sky short: short duration transients
  - Millisecond to few second, frequency band of 24-4096 Hz
- Searches for generic waveform morphologies & astrophysically motivated waveforms: supernovae and pulsar glitches
- Cannot be detected through usual modelled search algorithms such as matched filtering -> Coherent WaveBurst

Coherent WaveBurst (cWB)

- Does not require a priori knowledge on morphology, time of arrival, sky-direction, polarisation
- Uses excess coherent energy in time-frequency domain
- Combines data from multiple detectors to create a coherent analysis
- Background is estimated by applying an unphysical time shift (greater than light travel time)
- Noise glitches can be difficult to distinguish from short duration signals
    - Short duration, often have low Quality factor Q

See: https://gwburst.gitlab.io/



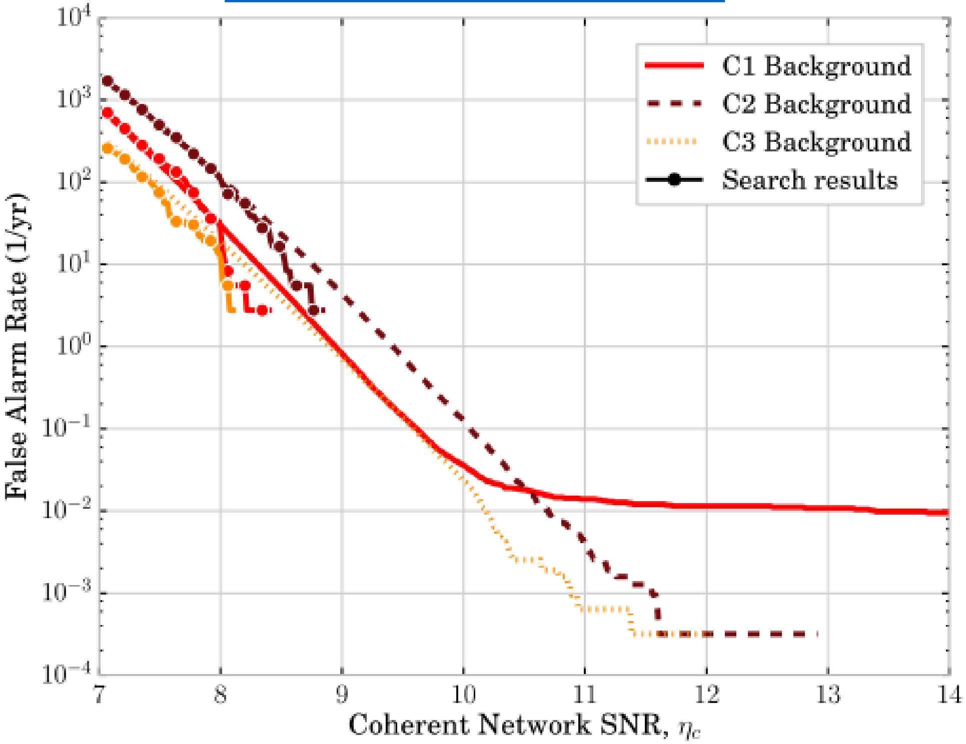
example of waveform in time-frequency plot

Coherent WaveBurst (cWB)

WB coherent WaveBurst
an open source software for gravitational-wave data analysis

- For standard cWB all-sky search:
  - Done in 2 parts: LF (24-1024Hz) & HF (1024-4096Hz)
  - LF has 3 bins in O3 search:
    - LF1 – signal energy confined mostly to one oscillation (& Q ≤ 3)
    - LF2 – Q ≤ 3
    - LF3 – higher Q

Bins in this plot are from O1 – shows benefit of using bins
arxiv.org/abs/1611.02972



Low Q = wide frequency bandwidth
High Q = narrow frequency bandwidth

Coherent WaveBurst (cWB)

coherent WaveBurst
an open source software for gravitational-wave data analysis

network correlation coefficients: c_{c0}, c_{c2}

Attribute	Definition
f_o	Central frequency
τ	Duration
η_c	Coherent network SNR
c_{c0}	$c_{c0} = E_c/(\vert E_c\vert + E_n)$
c_{c2}	$c_{c2} = (E_c \times c_{c0})/(\vert E_c\vert + E_n)$
N_{ED}	Energy disbalance between detectors
E_c	Coherent energy
N_{norm}	Ratio between reconstructed energy & total energy
χ^2	Residual noise energy measure
Q_{veto0}	Energy distribution of event over different time segments
Q_{veto1}	Quality factor
L_{veto0}	Central frequency of reconstructed signal (identifies narrow band glitches)
L_{veto1}	Root mean square frequency of reconstructed signal
L_{veto2}	Energy ratio between pixel energy and total energy of event

Coherent WaveBurst (cWB)

WB coherent WaveBurst
an open source software for gravitational-wave data analysis

Time-frequency Transform
Wilson-Daubechies-Mayer transform

↓

Data conditioning
Regression & Whitening

↓

Pixel Selection
Selected via energy threshold

↓

Super Cluster
Clusters from multiple resolutions,
Sub-threshold clusters rejected

↓

Likelihood calculation
On selected pixels, computation of detection statistics

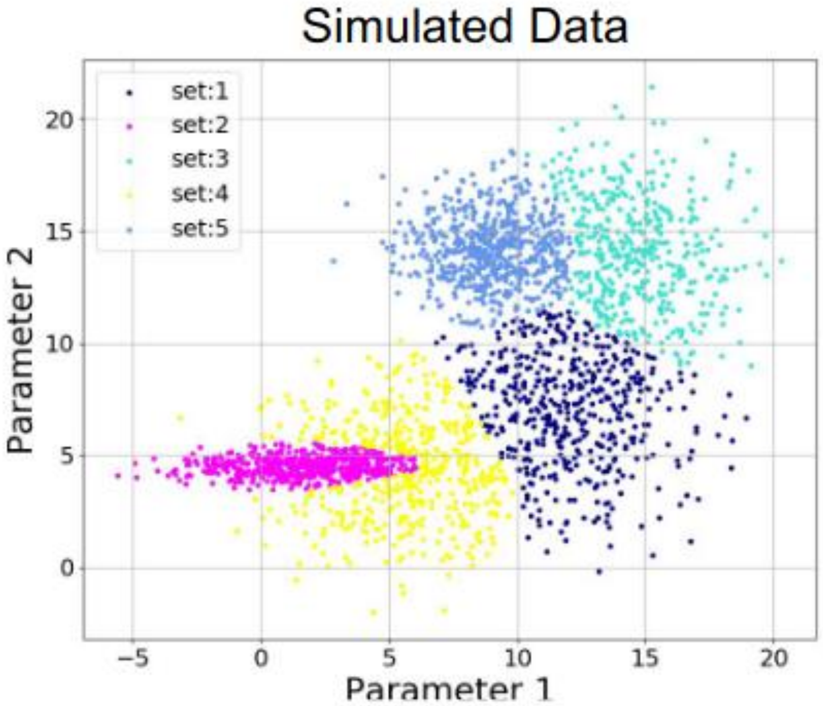
→

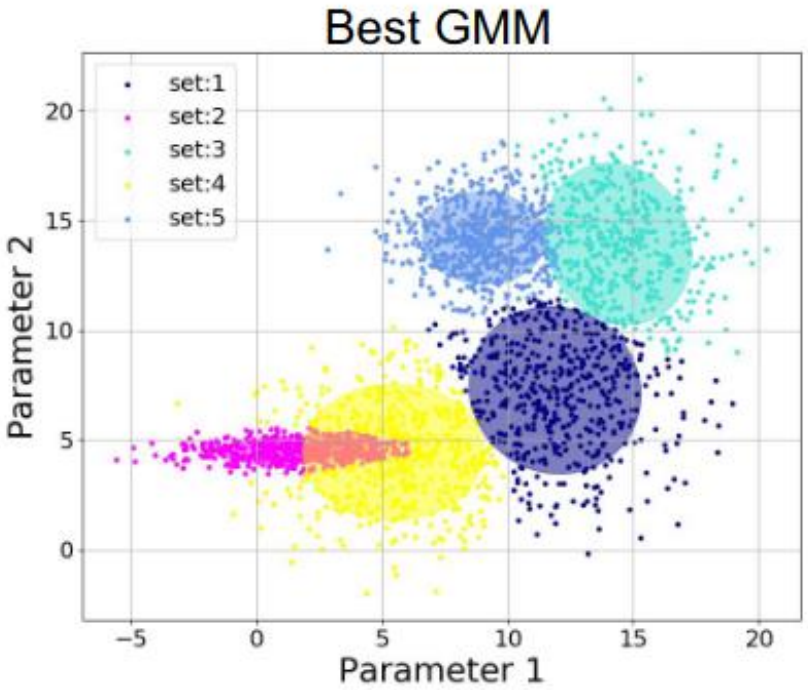
Post-production stage

Gaussian Mixture Modelling applied here

Gaussian Mixture Modelling (GMM)

- Supervised machine learning method
- Probabilistic model which uses uni-modal Gaussian distributions to represent a multi-modal data set
- Allows for sub-populations to be identified in data and modelled as a superposition of Gaussians





example with [sklearn package]()

Gaussian Mixture Modelling (GMM)

- Data x has d attributes. Modelled by GMM as a superposition of K Gaussians:

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

  - Gaussian distribution $\mathcal{N}(\mathbf{x}_i|\boldsymbol{\mu}_j, \Sigma_j)$ has d-dimensional mean $\boldsymbol{\mu}_j$, d x d covariance matrix Σ_j , a weight on each Gaussian of w_j

- Individual log-likelihood: $ln(\mathcal{L}) = \sum_{i=1}^{n} ln(p(\mathbf{x}|\Theta)) = \sum_{i=1}^{n} ln\left\{ \sum_{j=1}^{K} w_j \mathcal{N}(\mathbf{x}_i|\boldsymbol{\mu}_j, \Sigma_j) \right\}$

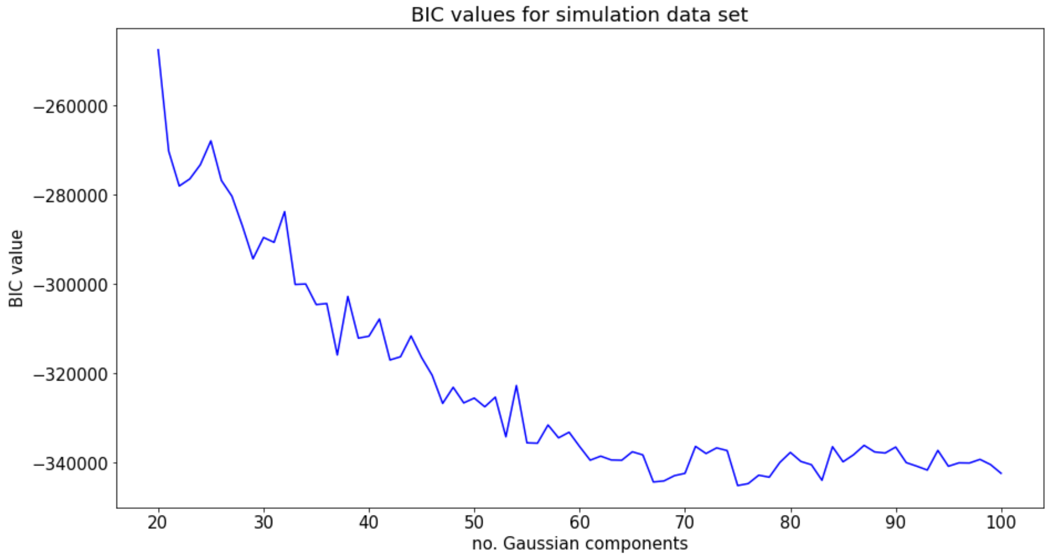
$\Theta := w_j, \boldsymbol{\mu}_j, \Sigma_j, \{j = 1, ..., K\}$

Gayathri V. et al. 2020

Gaussian Mixture Modelling (GMM)

- Model parameters estimated through Expectation maximisation (EM) technique
  - Cannot predict optimal no. Gaussians
- Account for overfitting of data through Bayesian Information Criterion (BIC):
  - For *n* no. of attributes and given no. of Gaussians *K*, and max likelihood $\hat{\mathcal{L}}$:

$$BIC = Kln(n) - 2ln(\hat{\mathcal{L}})$$



  - Lowest BIC score gives optimal number of Gaussian components

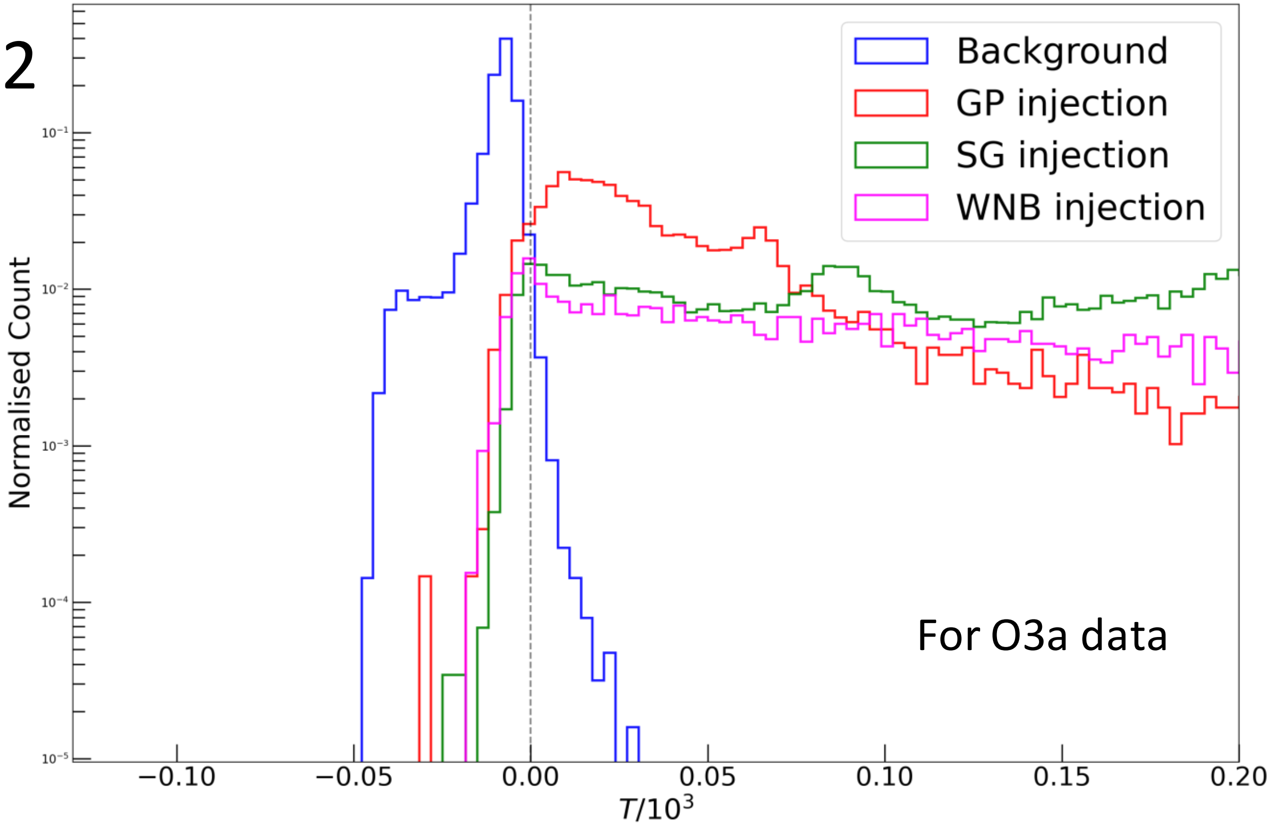
Gaussian Mixture Modelling (GMM)

- Maximum log-likelihood: $W = ln(\hat{\mathcal{L}})|_{\hat{K}}$

- Noise and signals are distinguished as 2 separate models
- Detection statistic for each trigger:

$$T = W_s - W_n$$

- Clear separation between the signal and background distributions



Application to cWB

- Applied at post-production stage of cWB
- No need to separate into bins
  - removes need to use trial factor
- Use of 11 cWB defined attributes
- Re-parametrisation of attributes gives optimal results
  - New parameter introduced: $L_{ratio} = L_{veto1}/L_{veto0}$

Original attribute set	Re-parametrized attribute set
η_c	$log_{10}(\eta_c)$
c_{c0}	$logit(c_{c0})$
c_{c2}	$logit(c_{c2})$
N_{ED}	$log_{10}(N_{ED} + 10^3)$
E_c	$log_{10}(E_c)$
N_{norm}	N_{norm}
χ^2	χ^2
Q_{veto0}	$log_{10}(Q_{\text{veto0}} + 1)$
Q_{veto1}	$log_{10}(Q_{\text{veto1}})$
L_{ratio}	$logit(L_{\text{ratio}})$
L_{veto2}	$logit(L_{\text{veto2}} \times 0.99)$

GMM on O3a all sky data

- Consider triggers with η_c (SNR) > 5.5 & c_c (cross-correlation) > 0.5 in HL network
- Simulated signals used for training
  - Consists of Sine-Gaussian (SQ), Gaussian Pulse (GP) and White Noise Burst (WNB) events

Sine-Gaussian Burst (SGW)			
No.	f_0 (Hz)	Q	-
1	70	3	-
2	70	9	-
3	70	100	-
4	100	9	-
5	153	9	-
6	235	3	-
7	235	9	-
8	235	100	-
9	361	9	-
10	554	9	-
11	849	3	-
12	849	9	-
13	849	100	-

White-Noise Burst (WNB)			
	$f_{\rm low}$ (Hz)	Δf (Hz)	τ (s)
14	150	100	0.1
15	300	100	0.1
16	750	100	0.1
Gaussian Pulse (GP)			
	-	-	τ (s)
17	-	-	0.1
18	-	-	1
19	-	-	2.5
20	-	-	4

Paper: D. Lopez et al. 2021

GMM on O3a all sky data

- Data randomly split into 3 sets
  - 10% for validation
  - 70% for training
  - 20% for testing

- Equal distribution of injected waveforms

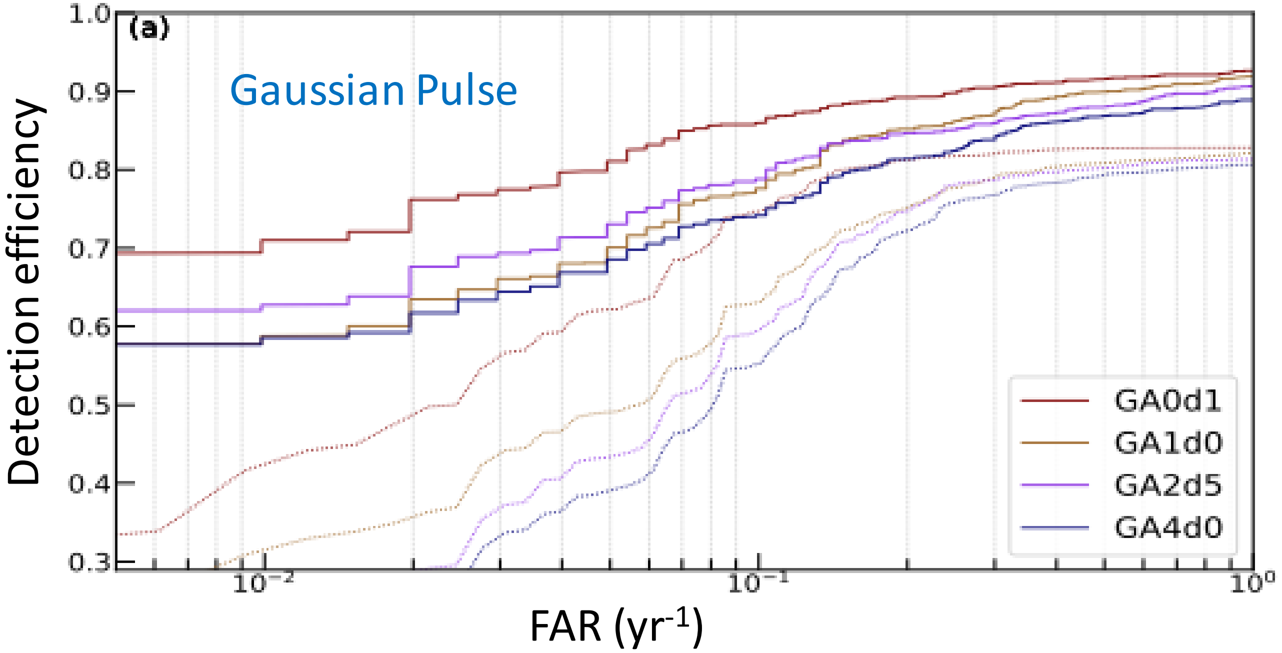
- Approx. 1000 years of background data
  - Only tested on 200 years due to other data being used for training

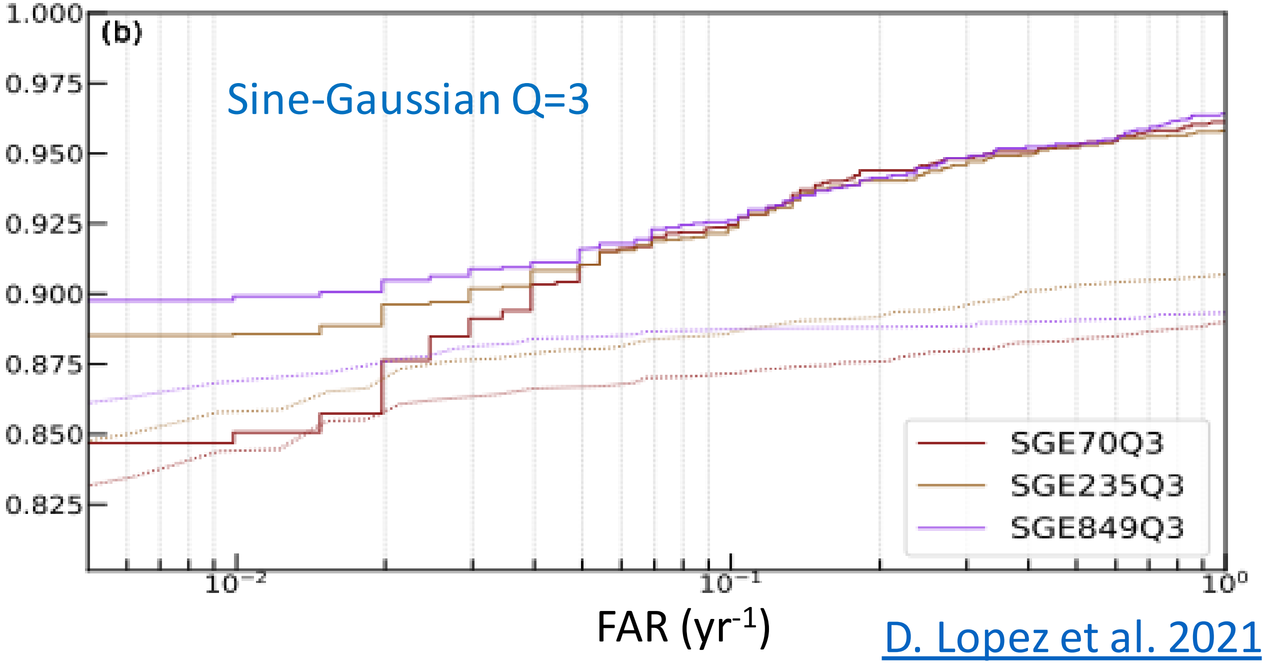
Paper: D. Lopez et al. 2021

Results – injected waveforms

- Efficiency calculated based on T
- Best improvement on detection efficiency for GP
  - often fall into bin containing a population of very short and very loud blip-type glitches
- GP has improvement of ~ x125 in sensitive volume

Solid - GMM + cWB
Dashed - cWB





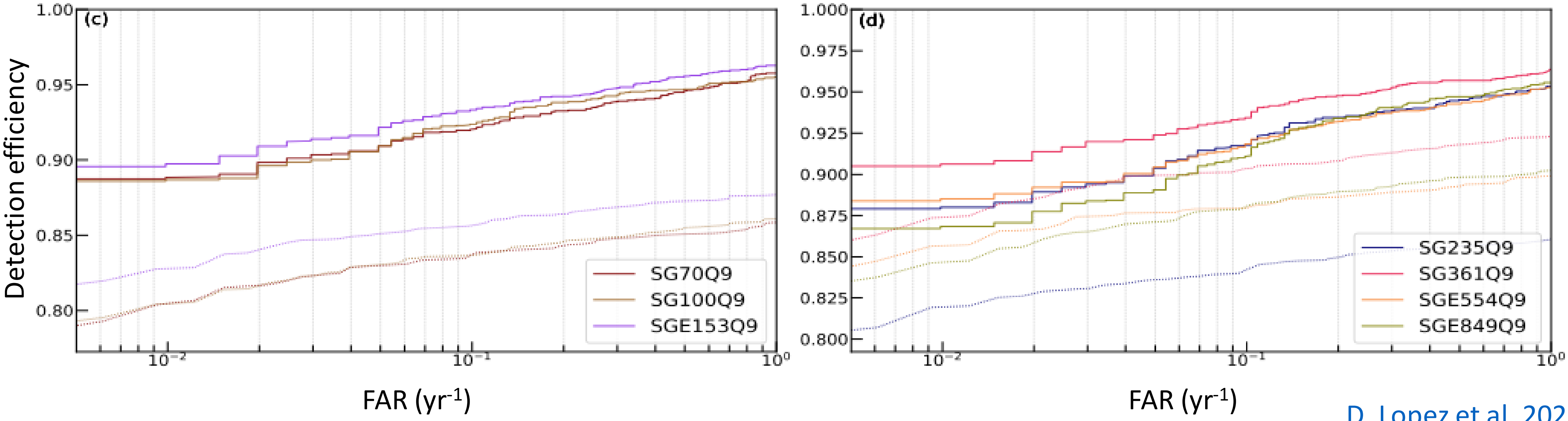
D. Lopez et al. 2021

Results – injected waveforms

- Improves detection efficiency for a given FAR over all waveforms
- Typical improvement of ~ x1.33 in sensitive volume for SG & WNB

Solid - GMM + cWB
Dashed - cWB

Sine-Gaussian Q=9

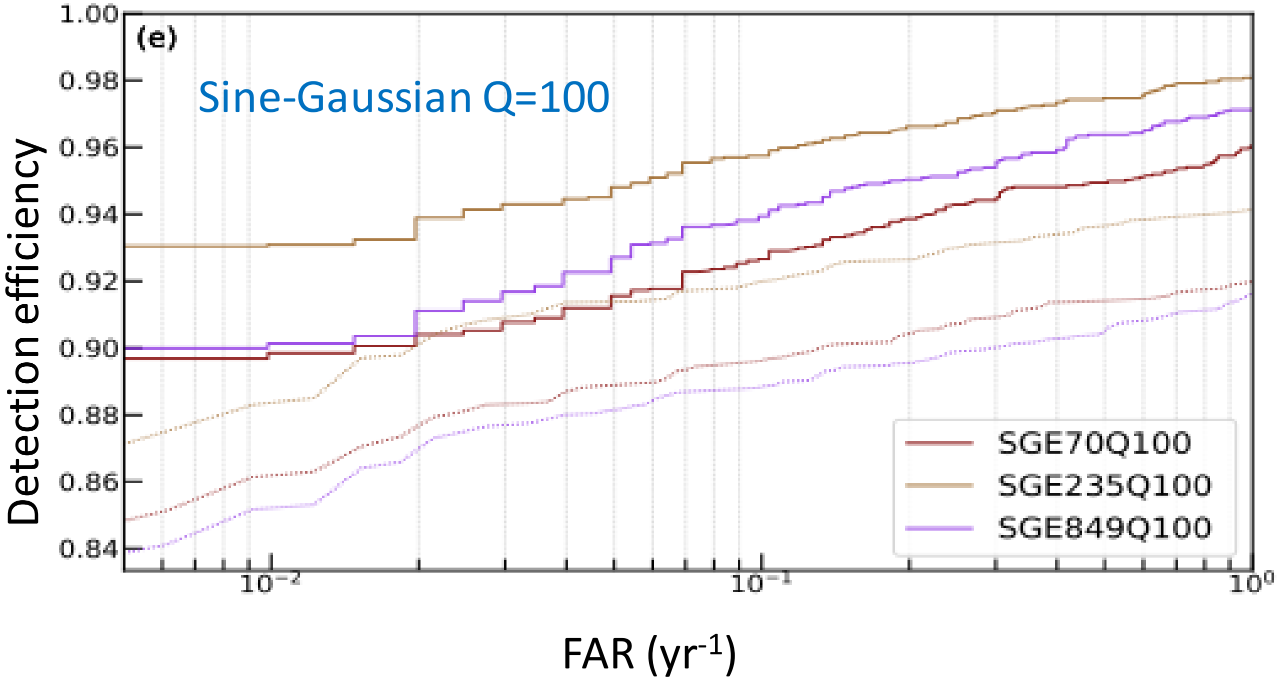


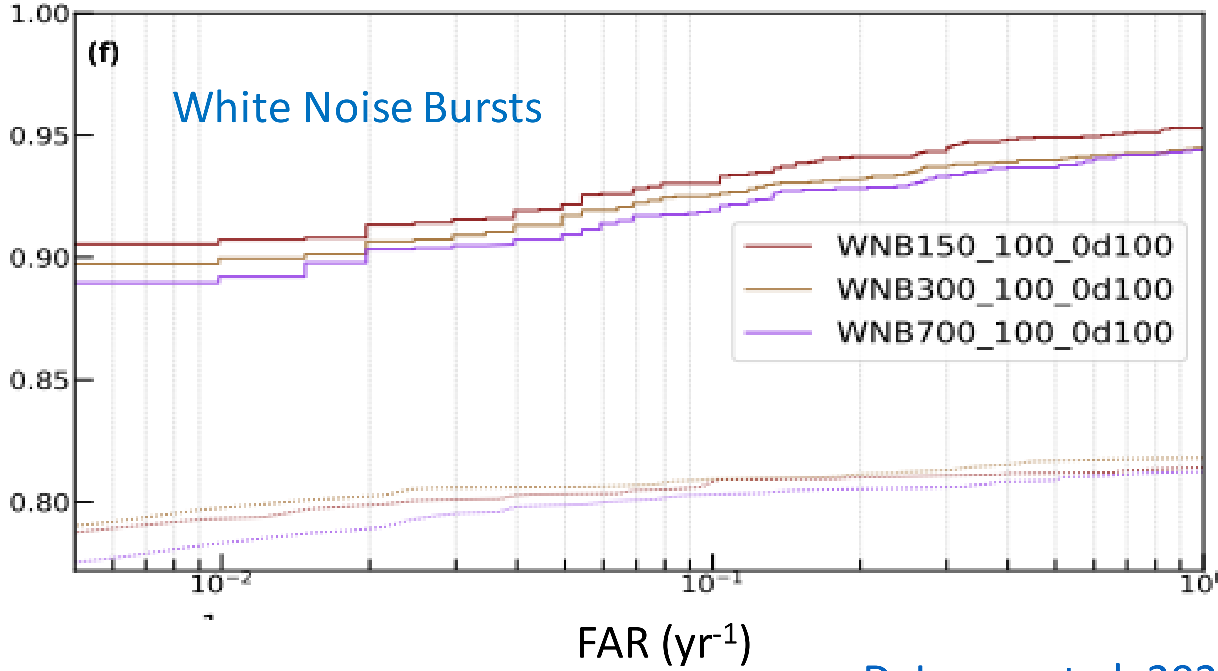
D. Lopez et al. 2021

Results – injected waveforms

- Improves detection efficiency for a given FAR over all waveforms
- Typical improvement of ~ x1.33 in sensitive volume for SG & WNB

Solid - GMM + cWB
Dashed - cWB

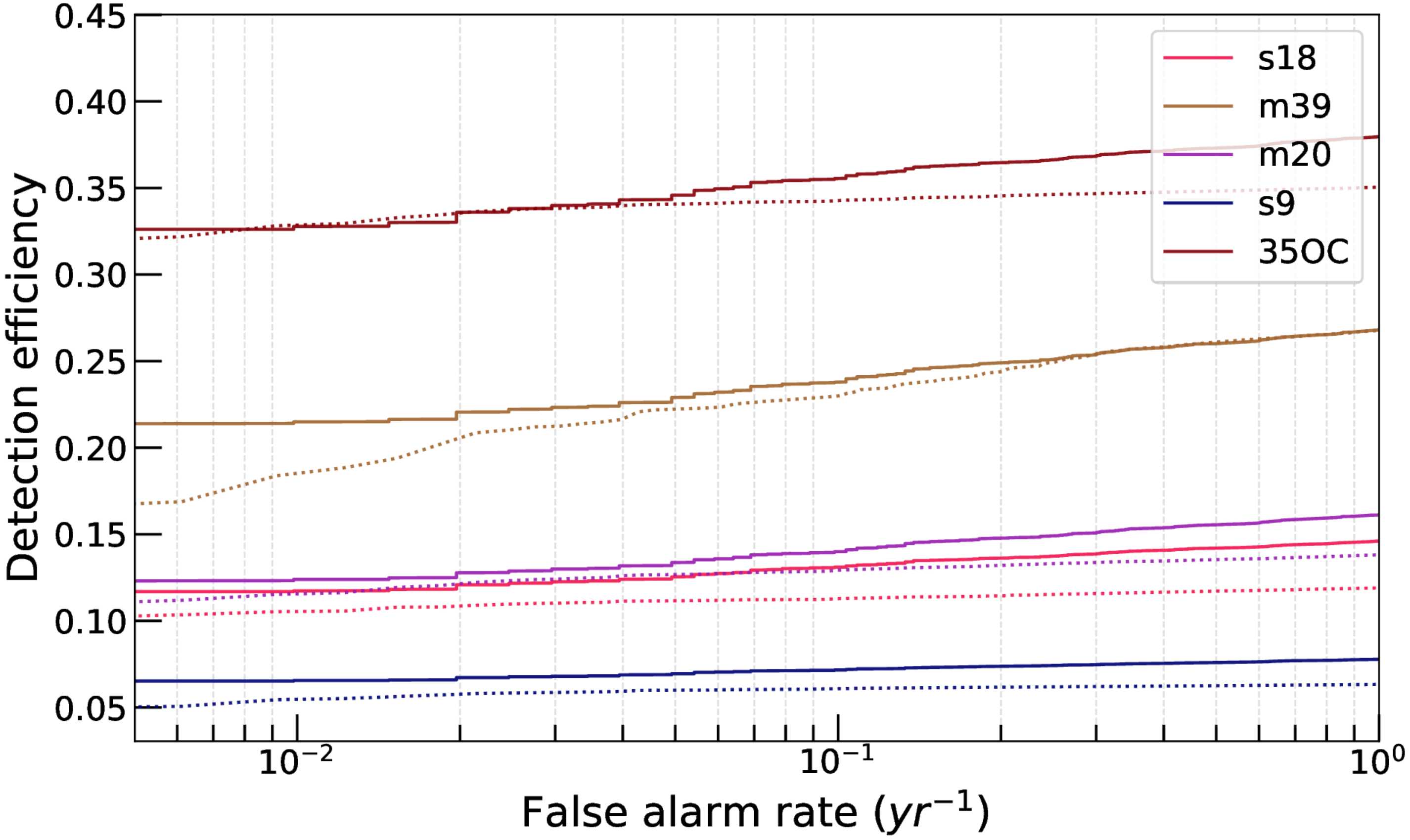




D. Lopez et al. 2021

Results – CCSN injections

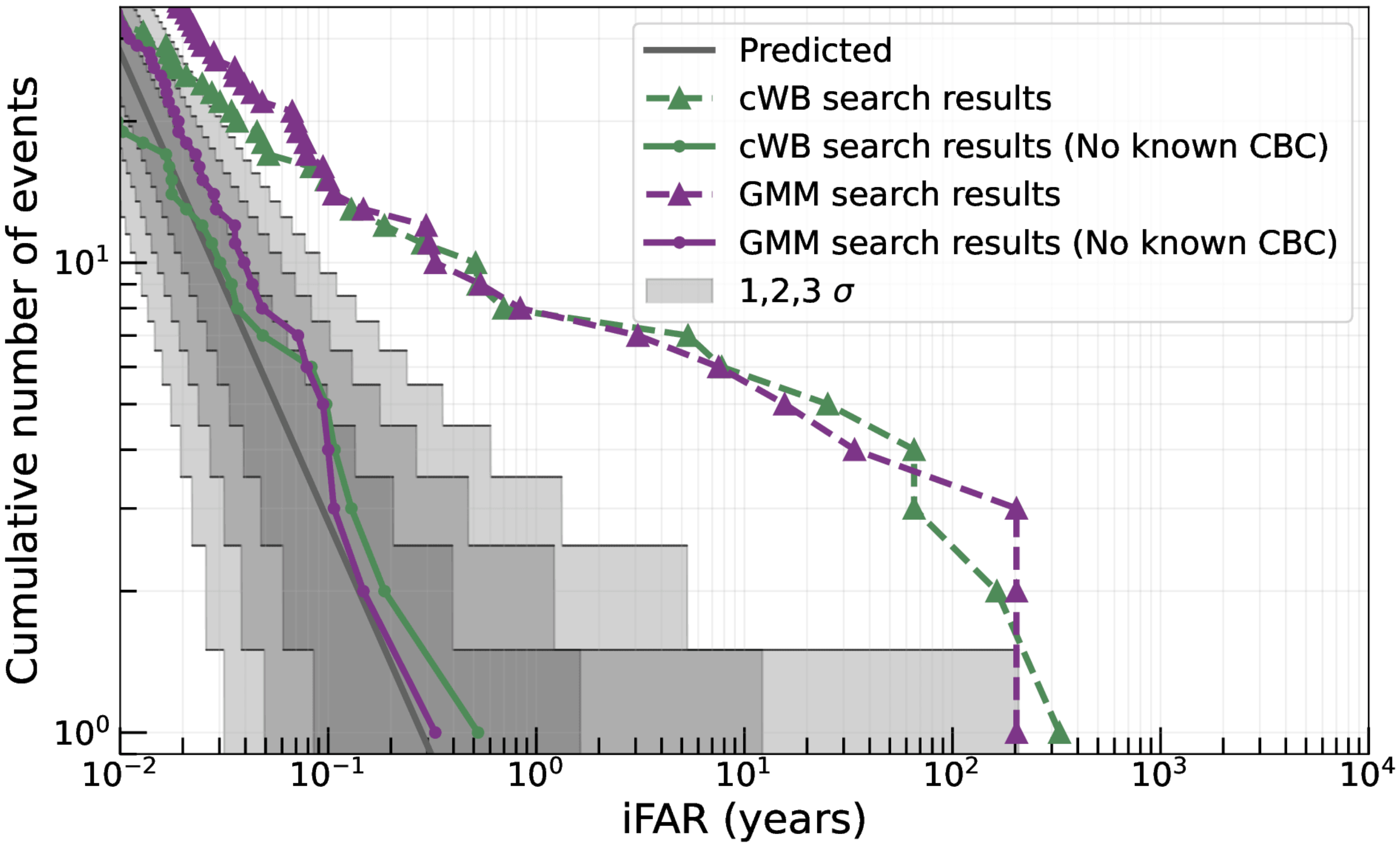
- CCSN waveforms were not included in training set
  - Tests robustness over variety of morphologies
- Largest improvement seen for m39 type SN
- Comparative efficiency for other models, however still has improvement



D. Lopez et al. 2021

Results

- Recovers all CBC events identified by targeted cWB search
- Higher significance for some events than found with standard cWB
- Less sensitive to low-mass BBH systems
  - Potentially due to injected signals used



D. Lopez et al. 2021

Summary

- cWB plus GMM enhances the detection performance for generic morphologies
  - improved detection probability at given FAR across simulated waveforms
  - Typical improvement of 1.33 in sensitive volume, 125 for GP
- Assists with classification of blip glitches
- Improvement in CCSNe injections demonstrates robustness
- Work currently being done to make available for O4 run

Papers: method: arxiv:2008.01262 , application to O3a: arxiv:2112.06608

For O3a simulations:

- SG & GP injected over grid of max strain values given by $h_{rss} = (\sqrt{3})^{N} 5$ X10^{-23} Hz$^{-1/2}$
- WNB injected uniform in square of signal distance