Enhancing transient gravitational wave analyses with machine learning

Ik Siong Heng
with plenty of input from Dixeena Lopez, Gayathri V, Archana Pai, Chris Messenger,…

Big Data in Multi-Messenger Astrophysics
Institute for Pure and Applied Maths,
University of California, Los Angeles
29th November, 2021

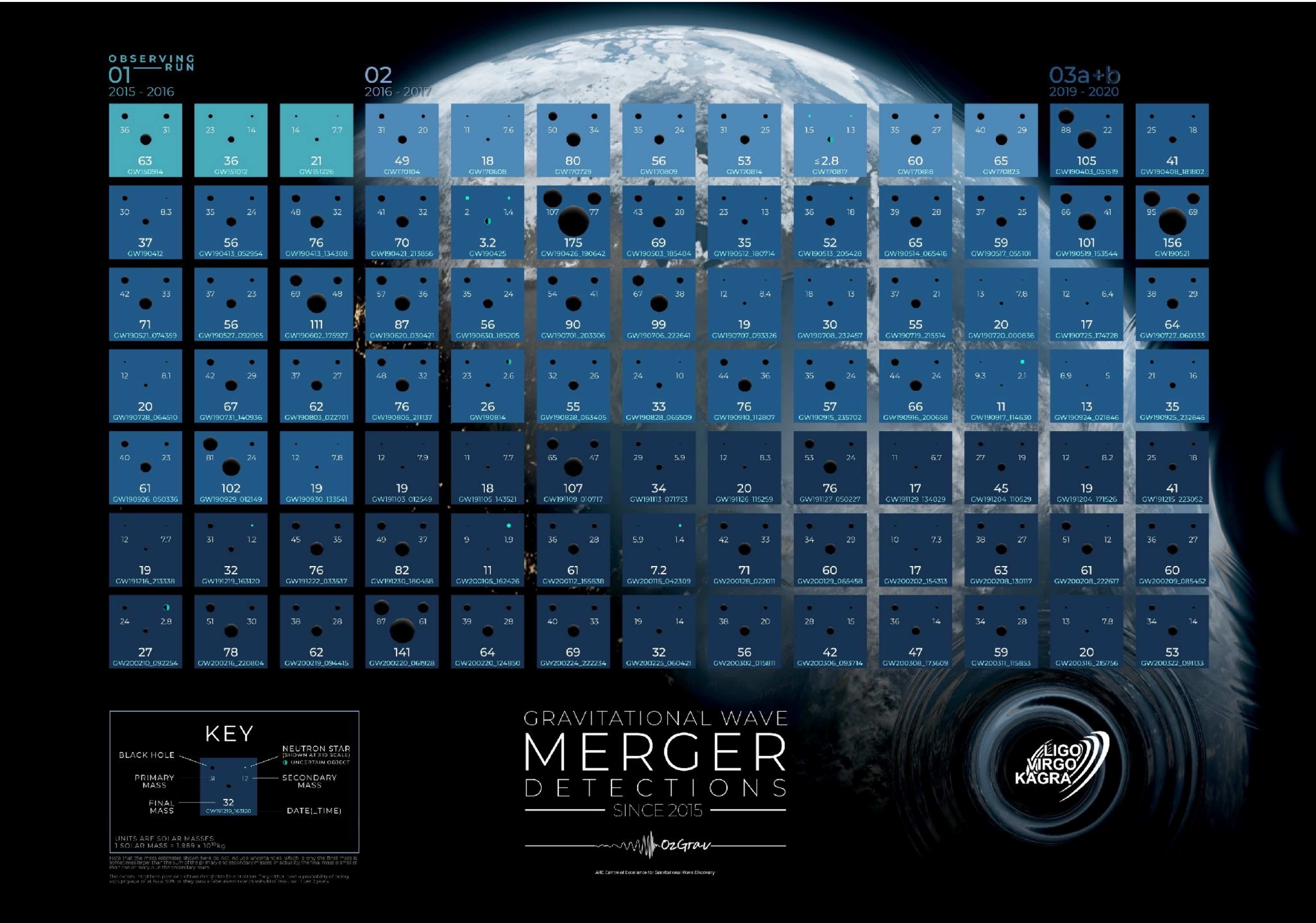
Overview

- Introduction to gravitational wave burst searches
- Coherent Waveburst
- Gaussian Mixture Modelling
- Application to O3a all-sky burst searh
- Summary, discussion & future work
- Bonus (if time permits): Generative adversarial networks for Burst waveform generation

ML4GW@GLA

- Machine learning for gravitational waves at Glasgow:
- Convolutional neural networks (CNNs) for binary black holes and continuous wave searches
  - H. Gabbard *et al*. PRL 2018, arXiv:1712.06041
  - J. Bayley et al., PRD 2020, arXiv:2007.08207
- Rapid inference with conditional variational autoencoders (VItamin)
  - H. Gabbard *et al*., accepted Nature Physics, arXiv:1909.06296
- Speed up of nested sampling with normalising flows (Nessai)
  - M.J. Williams *et al*., PRD 2021, arXiv:2102.11056
- Gaussian process regression for waveform characterisation
  - D. Williams *et al*., PRD 2020, arXiv:1903.09204
- Generative Adversarial Networks (GANs) for generating burst signals
  - J. McGinn *et al*., CQG 2021, arXiv:2103.01641
- Gaussian mixture models for gravitational wave burst searches (this talk)
  - Gayathri V. *et al*., PRD 2020, arXiv:2008.01262

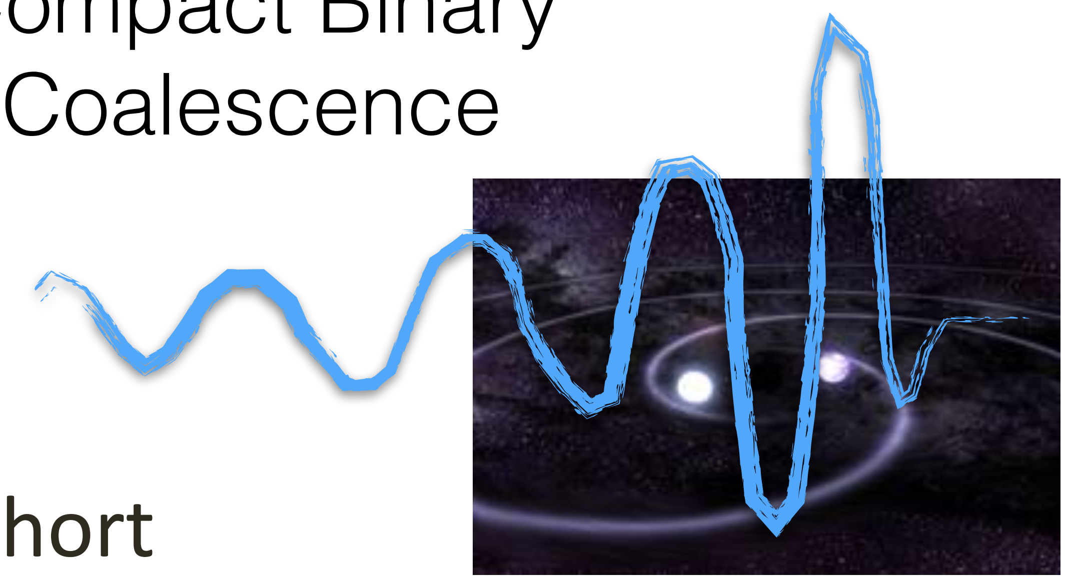
Global gravitational wave network



OBSERVING RUN
01
2015 - 2016
02
2016 - 2017
O3a+b
2019 - 2020
36 31 63 GW150914
23 14 36 GW151012
14 7.7 21 GW151226
31 20 49 GW170104
11 7.6 18 GW170608
50 34 80 GW170729
35 24 56 GW170809
31 25 53 GW170814
1.5 1.3 ≤2.8 GW170817
35 27 60 GW170818
40 29 65 GW170823
88 22 105 GW190403_051519
25 18 41 GW190408_181802
30 8.3 37 GW190412
35 24 56 GW190413_052954
48 32 76 GW190413_134308
41 32 70 GW190421_213856
2 1.4 3.2 GW190425
107 77 175 GW190426_190642
43 28 69 GW190503_185404
23 13 35 GW190512_180714
36 18 52 GW190513_205428
39 28 65 GW190514_065416
37 25 59 GW190517_055101
66 41 101 GW190519_153544
95 69 156 GW190521
42 33 71 GW190521_074359
37 23 56 GW190527_092055
69 48 111 GW190602_175927
57 36 87 GW190620_030421
35 24 56 GW190630_185205
54 41 90 GW190701_203306
67 38 99 GW190706_222641
12 8.4 19 GW190707_093326
18 13 30 GW190708_232457
37 21 55 GW190719_215514
13 7.8 20 GW190720_000836
12 6.4 17 GW190725_174728
38 29 64 GW190727_060333
12 8.1 20 GW190728_064510
42 29 67 GW190731_140936
37 27 62 GW190803_022701
48 32 76 GW190805_211137
23 2.6 26 GW190814
32 26 55 GW190828_063405
24 10 33 GW190828_065509
44 36 76 GW190910_112807
35 24 57 GW190915_235702
44 24 66 GW190916_200658
9.3 2.1 11 GW190917_114630
8.9 5 13 GW190924_021846
21 16 35 GW190925_232845
40 23 61 GW190926_050336
81 24 102 GW190929_012149
12 7.8 19 GW190930_133541
12 7.9 19 GW191103_012549
11 7.7 18 GW191105_143521
65 47 107 GW191109_010717
29 5.9 34 GW191113_071753
12 8.3 20 GW191126_115259
53 24 76 GW191127_050227
11 6.7 17 GW191129_134029
27 19 45 GW191204_110529
12 8.2 19 GW191204_171526
25 18 41 GW191215_223052
12 7.7 19 GW191216_213338
31 1.2 32 GW191219_163120
45 35 76 GW191222_033537
49 37 82 GW191230_180458
9 1.9 11 GW200105_162426
36 28 61 GW200112_155838
5.9 1.4 7.2 GW200115_042309
42 33 71 GW200128_022011
34 29 60 GW200129_065458
10 7.3 17 GW200202_154313
38 27 63 GW200208_130117
51 12 61 GW200208_222617
36 27 60 GW200209_085452
24 2.8 27 GW200210_092254
51 30 78 GW200216_220804
38 28 62 GW200219_094415
87 61 141 GW200220_061928
39 28 64 GW200220_124850
40 33 69 GW200224_222234
19 14 32 GW200225_060421
38 20 56 GW200302_015811
28 15 42 GW200306_093714
36 14 47 GW200308_173609
34 28 59 GW200311_115853
13 7.8 20 GW200316_215756
34 14 53 GW200322_091133
KEY
BLACK HOLE
NEUTRON STAR (SHOWN AT X10 SCALE)
UNCERTAIN OBJECT
PRIMARY MASS
SECONDARY MASS
FINAL MASS
DATE(_TIME)
31 1.2 32 GW191219_163120
UNITS ARE SOLAR MASSES
1 SOLAR MASS = 1.989 x 10^30 kg
GRAVITATIONAL WAVE
MERGER
DETECTIONS
SINCE 2015
OzGrav
LIGO VIRGO KAGRA

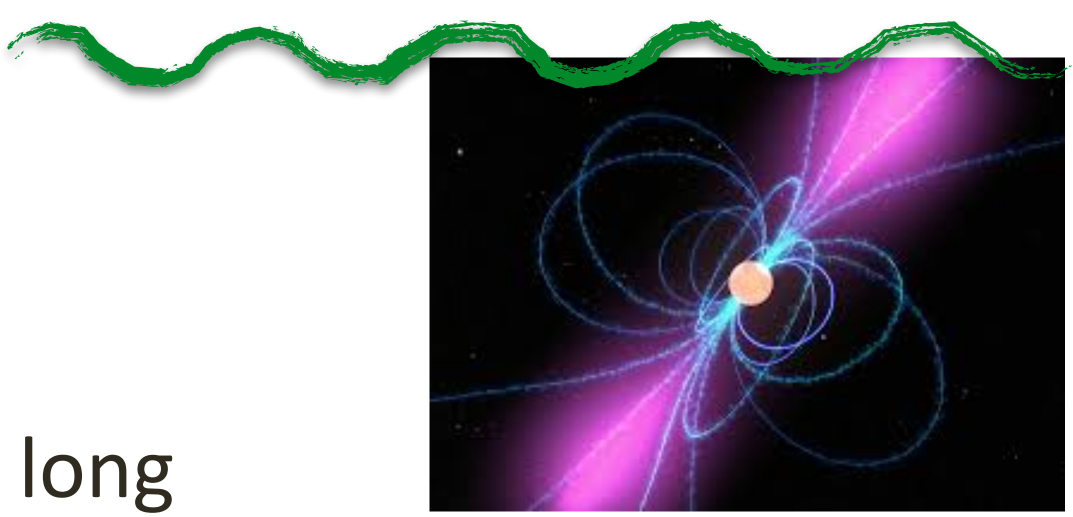
Gravitational wave signal types

modelled	unmodelled
Compact Binary Coalescence	Burst



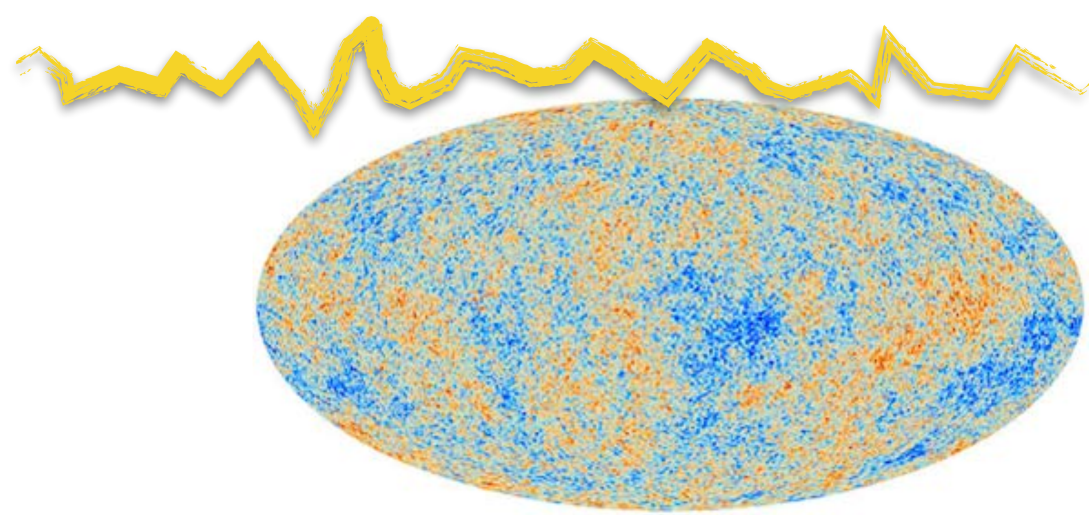
short

Continuous



long

Stochastic



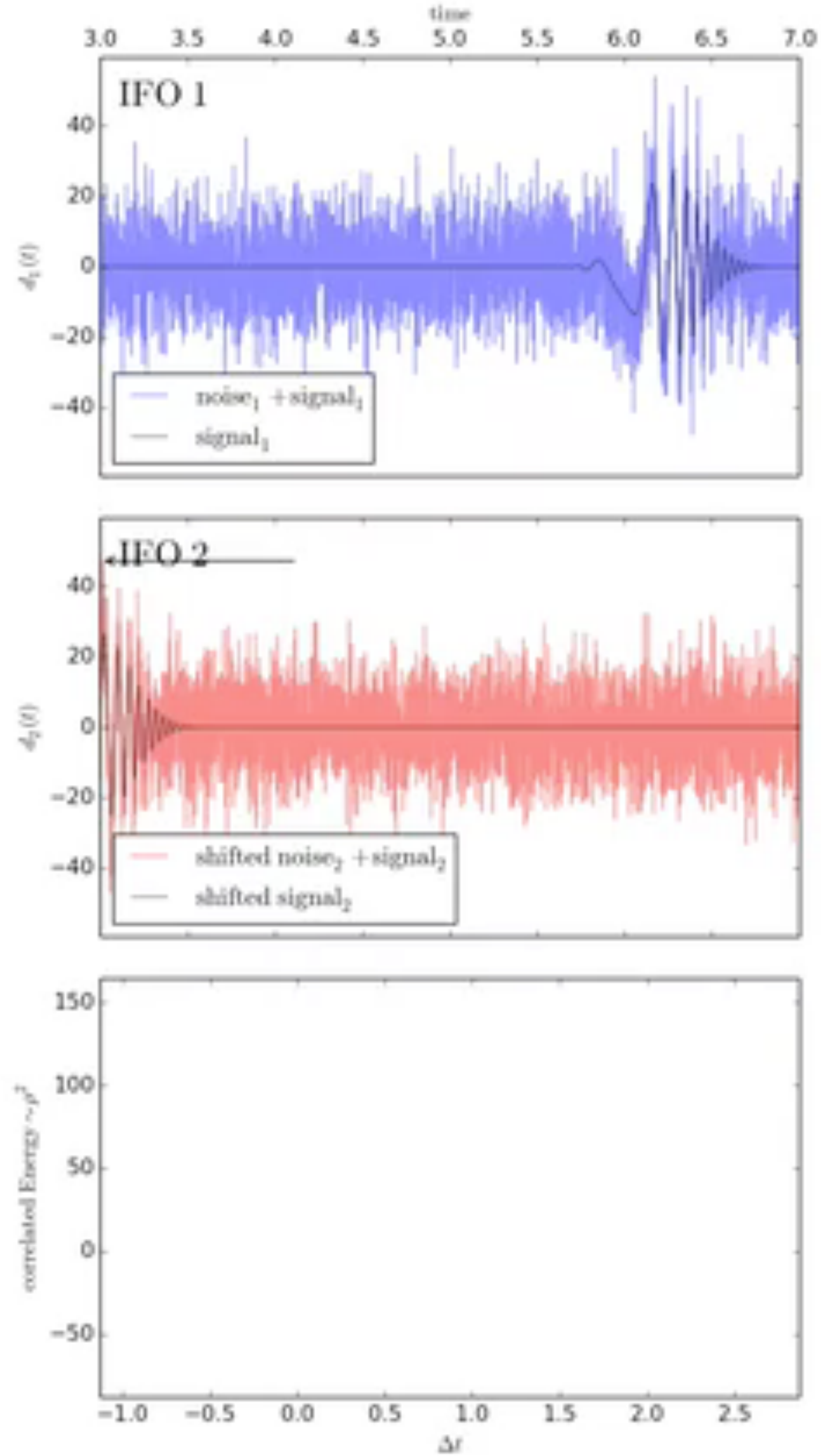
Generic transient (burst) analysis

- Searches for gravitational-wave bursts do not require knowledge about the phase evolution (waveform) of the expected signal
- Burst searches aim to cover a broad parameter space which can overlap with well-modelled signals (eg. binary black holes)
  - Calderon Bustillo et al., PRD (2018) & Romas-Buades et al., PRD (2020) have shown that burst searches can be more sensitive than template-based searches for GWs from high-mass BBH systems, especially if there is significant orbital eccentricity.
  - potential for discovering new sources of gravitational waves
- Steps of a typical generic burst search:
  - weight data by the noise at each frequency (whitening)
  - make time-frequency representation of the data
  - identify correlated excess power in multiple detectors
  - estimate false alarm rate of observations
- Some burst searches target GWs expected from particular sources or are informed by non-GW observations of astrophysical phenomena

Burst searches

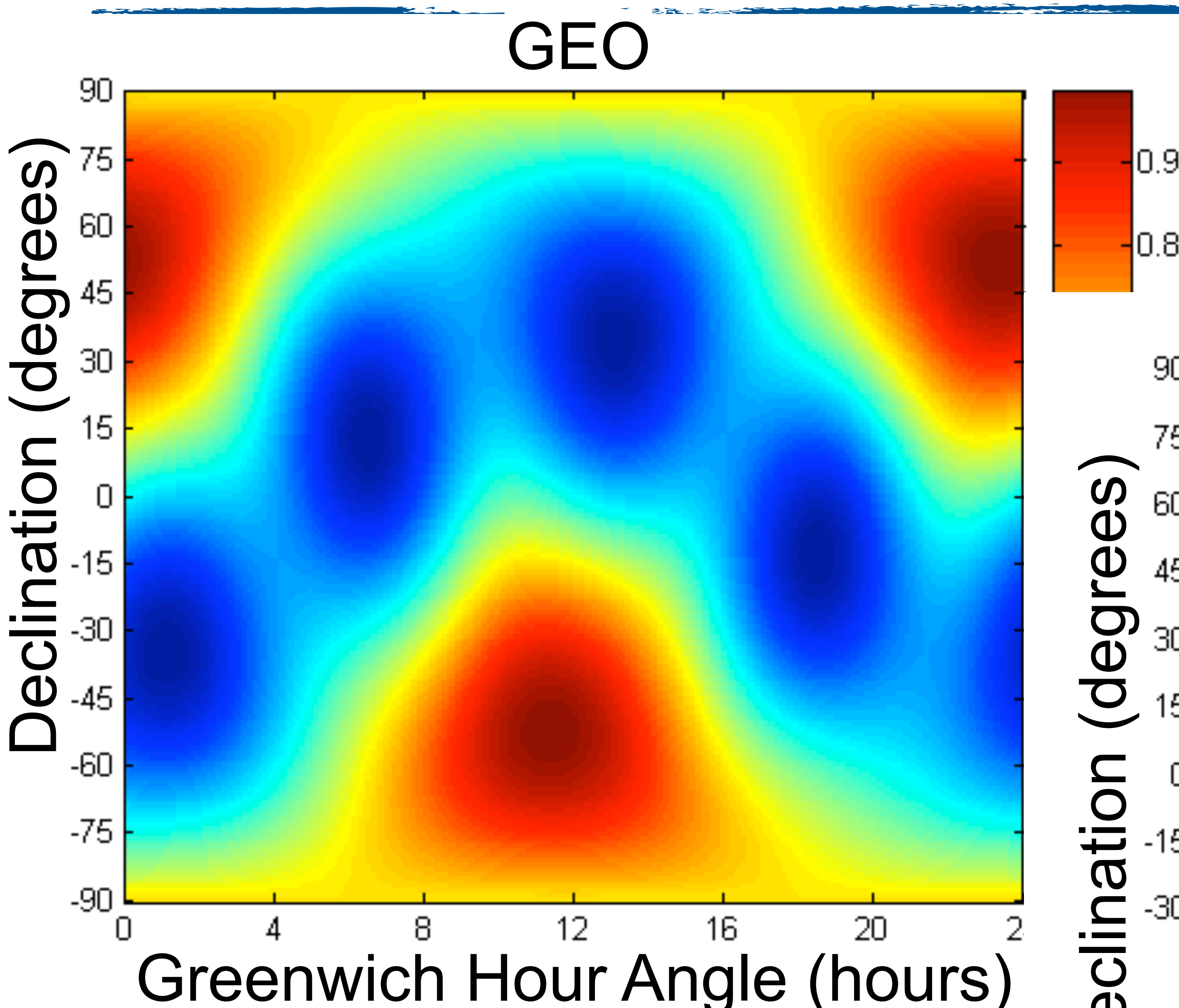
- Correlation between data from multiple detectors should be at a maximum when the data is shifted to correspond to the time delay corresponding to the sky location

- To estimate the background, large, unphysical time-shifts are applied to the data

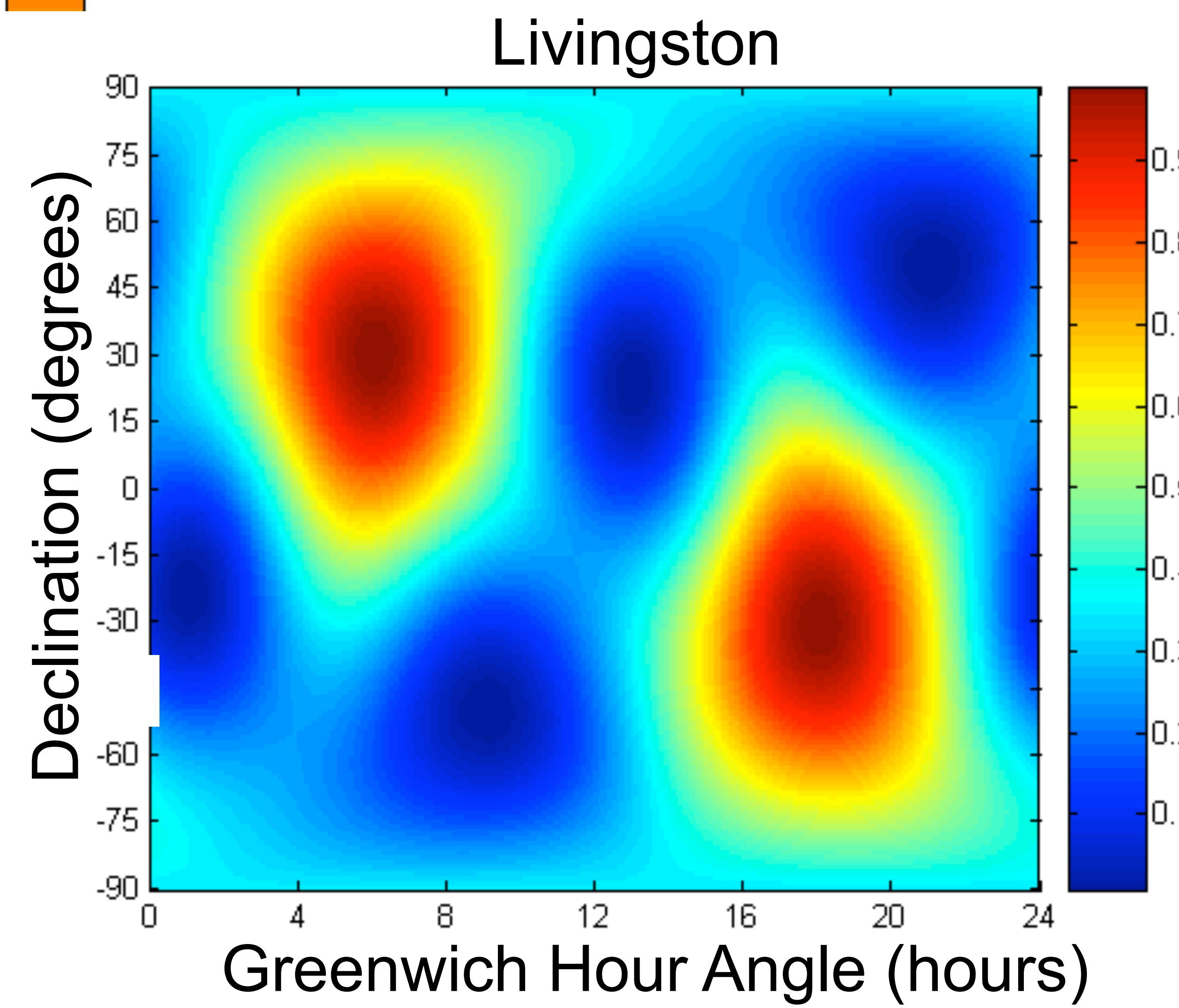


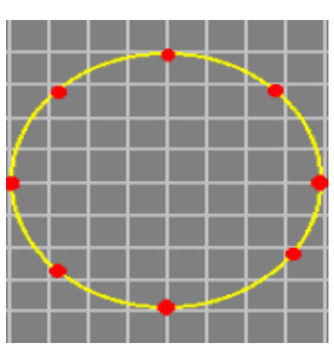
animation: https://github.com/reedessick/pedagogy

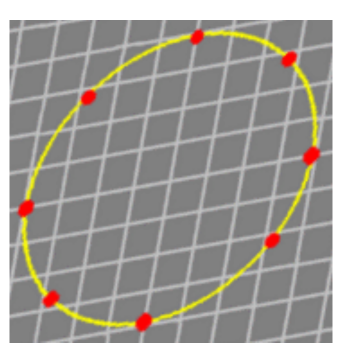
Directional sensitivity



$F_+^2 + F_\times^2$







Burst searches

- Lets formulate our data so that

$$\tilde{\boldsymbol{d}}_{\mathrm{w}} = \boldsymbol{F}_{\mathrm{w}}\tilde{\boldsymbol{h}} + \tilde{\boldsymbol{n}}_{\mathrm{w}},$$

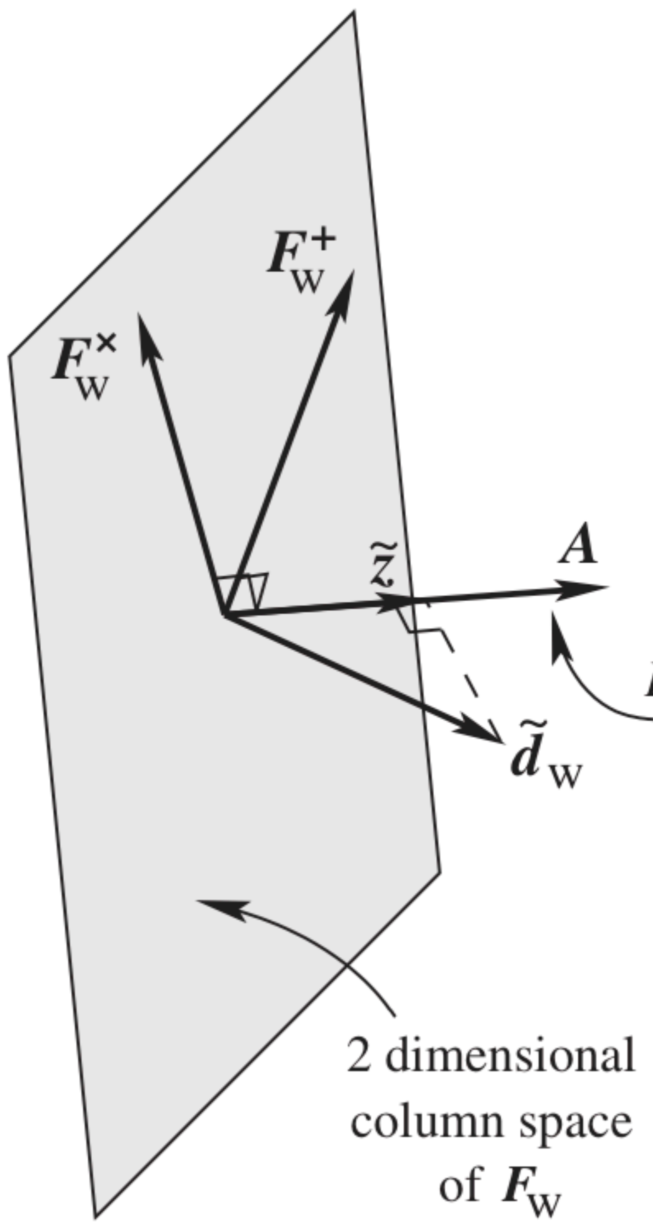
or

$$\overset{\text{data}}{\begin{bmatrix} \tilde{d}_{w1} \\ \tilde{d}_{w2} \\ \vdots \\ \tilde{d}_{wD} \end{bmatrix}} = \overset{\text{antenna pattern}}{\begin{bmatrix} F^{+}_{w1} & F^{\times}_{w1} \\ F^{+}_{w2} & F^{\times}_{w2} \\ \vdots & \vdots \\ F^{+}_{wD} & F^{\times}_{wD} \end{bmatrix}} \overset{\text{signal}}{\begin{bmatrix} \tilde{h}_{+} \\ \tilde{h}_{\times} \end{bmatrix}} + \overset{\text{noise}}{\begin{bmatrix} \tilde{n}_{w1} \\ \tilde{n}_{w2} \\ \vdots \\ \tilde{n}_{wD} \end{bmatrix}},$$

$$\boldsymbol{F}_{\mathrm{w}}(\overset{\text{sky location}}{\hat{\Omega}_s}, \overset{\text{frequency}}{f}) \equiv [\, \boldsymbol{F}^{+}_{\mathrm{w}} \quad \boldsymbol{F}^{\times}_{\mathrm{w}} \,] = \begin{bmatrix} F^{+}_{\mathrm{w}1} & F^{\times}_{\mathrm{w}1} \\ F^{+}_{\mathrm{w}2} & F^{\times}_{\mathrm{w}2} \\ \vdots & \vdots \\ F^{+}_{\mathrm{w}D} & F^{\times}_{\mathrm{w}D} \end{bmatrix}$$

Chatterji et al, PRD 74, 082005 (2006)

Null stream



- **Consider a matrix A whose rows are components of an orthonormal basis**
- **If we construct such that**

$$A\,F_w = 0.$$

- **then, we can use it to construct a null stream**

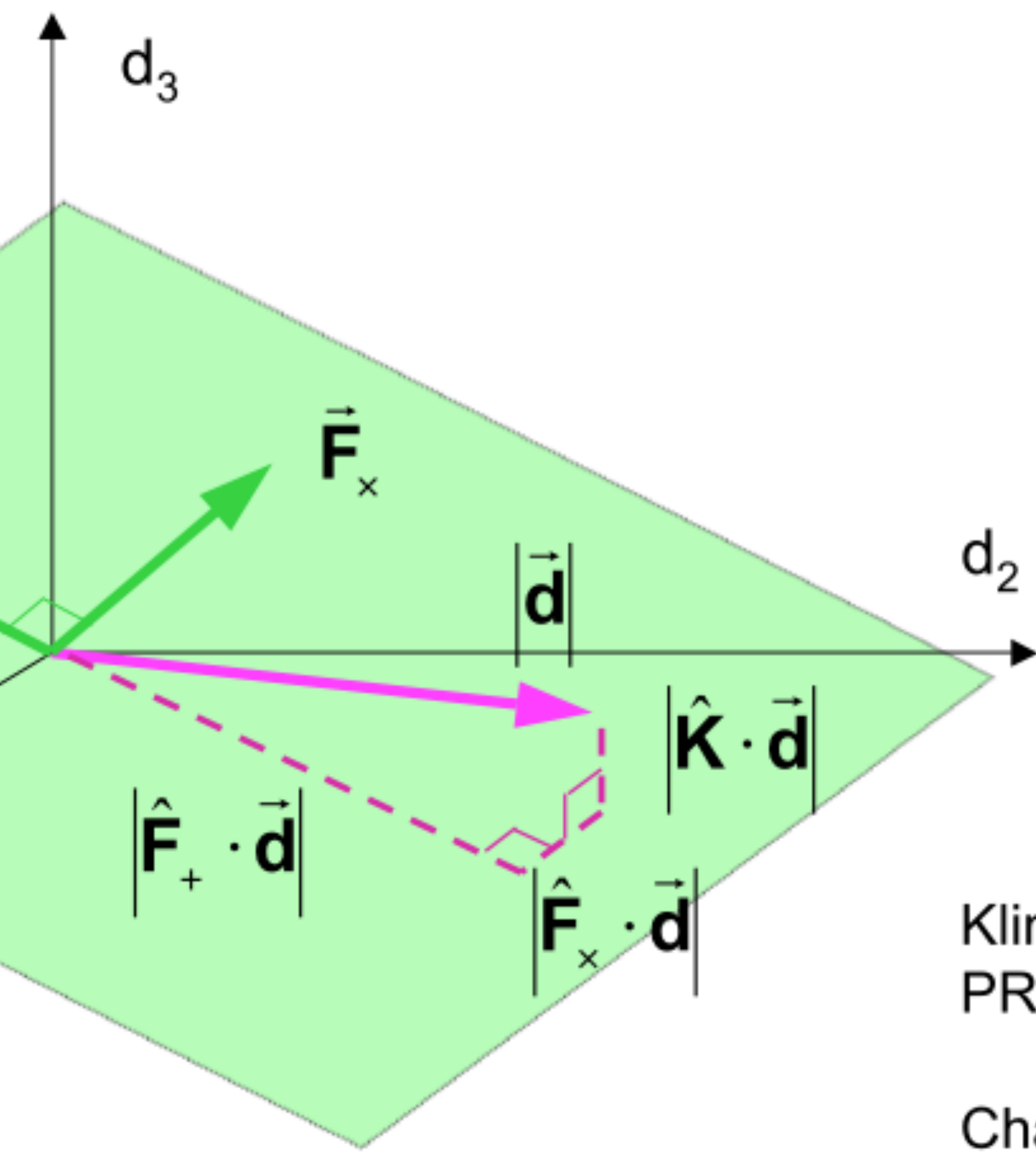
$$\tilde{z} \equiv A\tilde{d}_w = AF_w\tilde{h} + A\tilde{n}_w = A\tilde{n}_w.$$

- **For 3 detectors, A is constructed by**

$$A = \frac{F_w^+ \times F_w^\times}{|F_w^+ \times F_w^\times|}$$

Chatterji et al, PRD 74, 082005 (2006)

Likelihood or Energy Measures



Likelihoods used in this analysis:

Hard Constraint $\mathbf{E}_{\mathbf{HC}} = \left|\hat{\mathbf{F}}_+ \cdot \vec{\mathbf{d}}\right|^2$

Null Energy $\mathbf{E}_{\mathbf{NULL}} = \left|\hat{\mathbf{K}} \cdot \vec{\mathbf{d}}\right|^2$

Incoherent Energy $\mathbf{E}_{\mathbf{INC}} = \sum_\alpha \left|\mathbf{K}_\alpha \cdot \mathbf{d}_\alpha\right|^2$

Klimenko, Mohanty, Rakhmanov,& Mitselmakher,
PRD **72** 122002 (2005); J. Phys. Conf. Ser. **32** 12 (2006)

Chatterji, Lazzarini, Stein, Sutton, Searle, & Tinto,
PRD **74** 082005 (2006)

Coherent Waveburst

- Coherent Waveburst (cWB) is an algorithm for detecting generic gravitational-wave transients
  - made the first detection of gravitational waves from binary black holes (GW150914)
- The cWB algorithm identifies coherent excess power in multiple detectors
- Excess power must be consistent with detector response (amplitude, time-delay,…) for a gravitational wave signal originating from somewhere in the sky
  - excess noise from environmental sources are not likely to have consistent signal features, time delay,…
- To determine the background, cWB is run on data with an unphysical time shift
- When a significant cluster of excess power is identified, it is stored as a trigger.
- Each trigger is characterised by a large range of attributes in an effort to capture the various properties of the identified excess power

<u>cWB trigger attributes (!!)</u>

```
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
```

Coherent Waveburst

- For each search, coherent Waveburst will generate a list of background triggers from the time-shifted data and a list of “zero-lag” triggers (which may contain gravitational wave signals).
- Coherent Waveburst is used for a range of searches, including searches for binary black holes, searches for gravitational waves associated with supernovae, searches for long-duration bursts,…
- The “standard” coherent Waveburst procedure is to characterise triggers by binning and thresholding in an effort to distinguish gravitational wave signals from spurious transients. (*post processing*)
- This process is typically optimised manually, often by looking at scatter plots of various attributes, and Receiver Operator Characteristic (ROC) curves.
- In the absence of well-modelled signal morphologies, the search sensitivities are characterised by ad hoc waveforms (more later)

Gaussian Mixture Models

- We wanted to develop an approach that minimises the need to binning and thresholding.
- We adopted a Gaussian Mixture Model (GMM) approach.
- Gaussian Mixture Models uses a combination of Gaussian distributions to model the parameter space covered by a set of data points
- We construct one GMM to model the attribute space covered by simulated signals and another GMM to model the time-shifted background triggers.
- Once the models are constructed, we can calculate the likelihood that a trigger belongs to the signal or noise model.
- Note that the GMM step is only applied at the post processing stage.
- Publish demonstration here: Gayathri *et al*. PRD (2020) arXiv:2008.01262
- There has also been work using boosted decision-tree approach (XGBoost) for better discrimination of binary black hole signals from background: T. Mishra *et al.* PRD (2021) arXiv:2105.04739

IGR University of Glasgow

Gaussian Mixture Models

- Gaussian Mixture Models (GMM) uses a combination of Gaussian distributions to model the parameter space covered by a set of data points
- The process of selecting the properties of the Gaussian distributions is iterative and unsupervised
- For a model with K Gaussian distributions, each of weight $\boldsymbol{w_i}$, the probability that the data $\vec{x}$ belongs to the model is

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

- Now, let $\vec{x}$ be the trigger attributes for 1 event. So, a trigger list of n events is $\mathbf{X} = \{\vec{x}_1, \vec{x}_2, \ldots \vec{x}_n\}$ and the corresponding likelihood is

$$\mathcal{L} = p(\mathbf{X}|\Theta) = \prod_{i=1}^{n} p(\vec{x}_i|\Theta) \qquad \Theta = \phi_j, \mu_j, \Sigma_j, \{j = 1...K\}$$

Gaussian Mixture Models

- The log-likelihood is thus

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

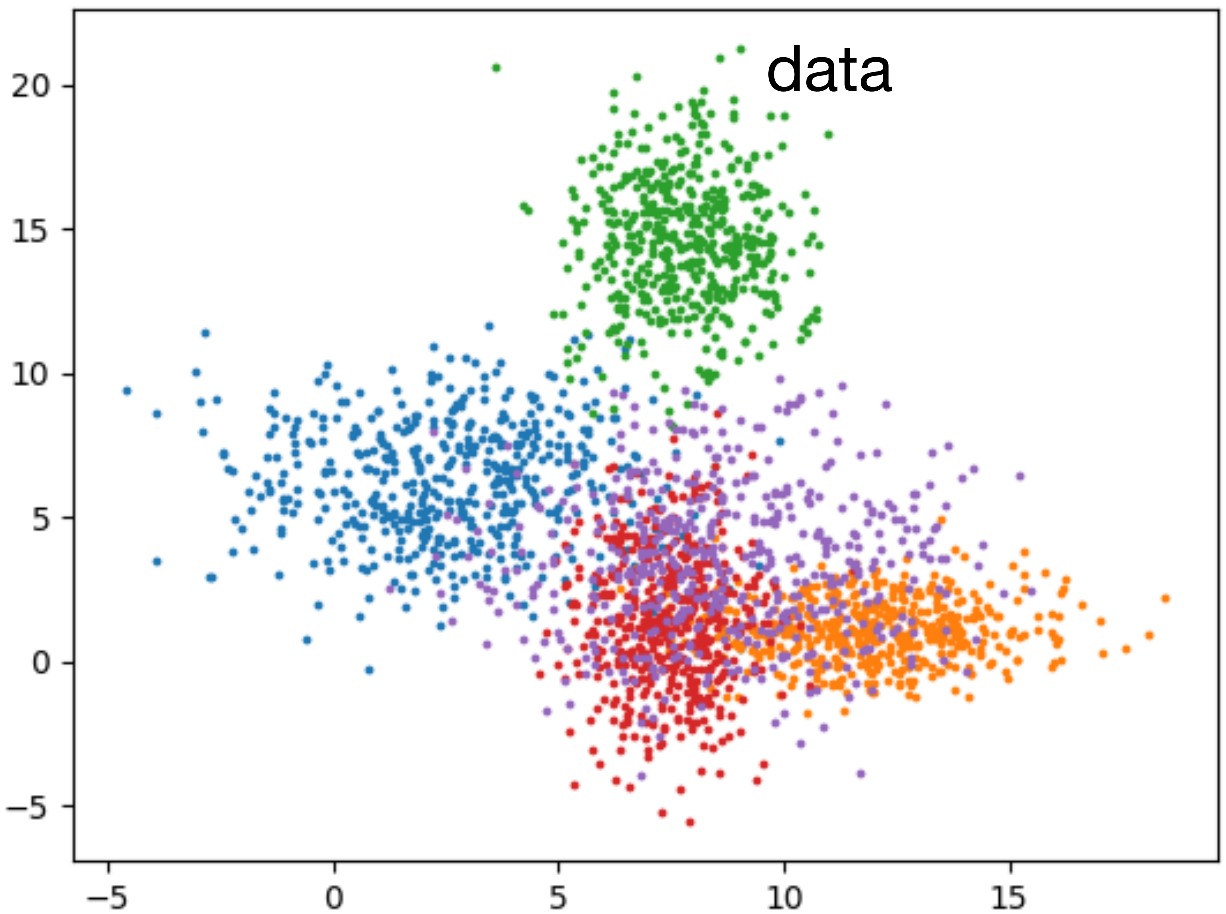
- Maximising the log-likelihood, we find

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{k}} = \frac{\sum_{i=1}^{n} r_{ik}\mathbf{x}_i}{N_k} \qquad \hat{\Sigma}_k = \frac{1}{N_k}\sum_{i=1}^{n} r_{ik}\left(\mathbf{x}_i - \boldsymbol{\mu}_{\boldsymbol{k}}\right)\left(\mathbf{x}_i - \boldsymbol{\mu}_{\boldsymbol{k}}\right)^T \qquad r_{ik} = \frac{w_k \mathcal{N}\left(\mathbf{x}_i|\boldsymbol{\mu}_{\boldsymbol{k}}, \Sigma_k\right)}{\sum_{j=1} w_j \mathcal{N}\left(\mathbf{x}_i|\boldsymbol{\mu}_{\boldsymbol{j}}, \Sigma_j\right)}$$

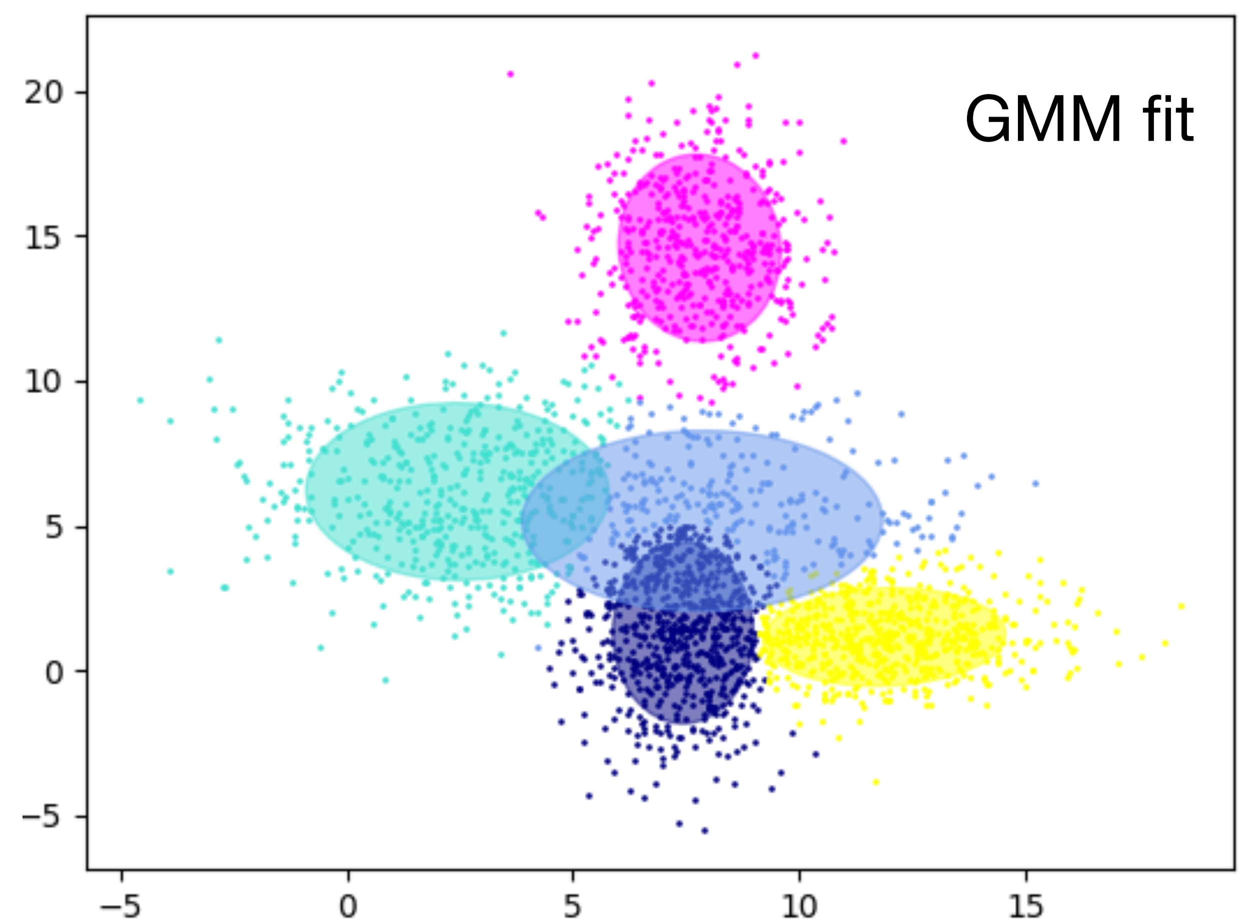
- In practice, Expectation Maximization solve for the maximum likelihood.
- This gives us the locations $\hat{\mu}_k$ and widths $\hat{\Sigma}_k$ of the Gaussians which we use to fit the data.

- Reminder: we create one GMM model for the signal space and another for the background space, then calculate the likelihood that triggers belong to either of these models.

Gaussian Mixture Model example



from scikit-learn examples

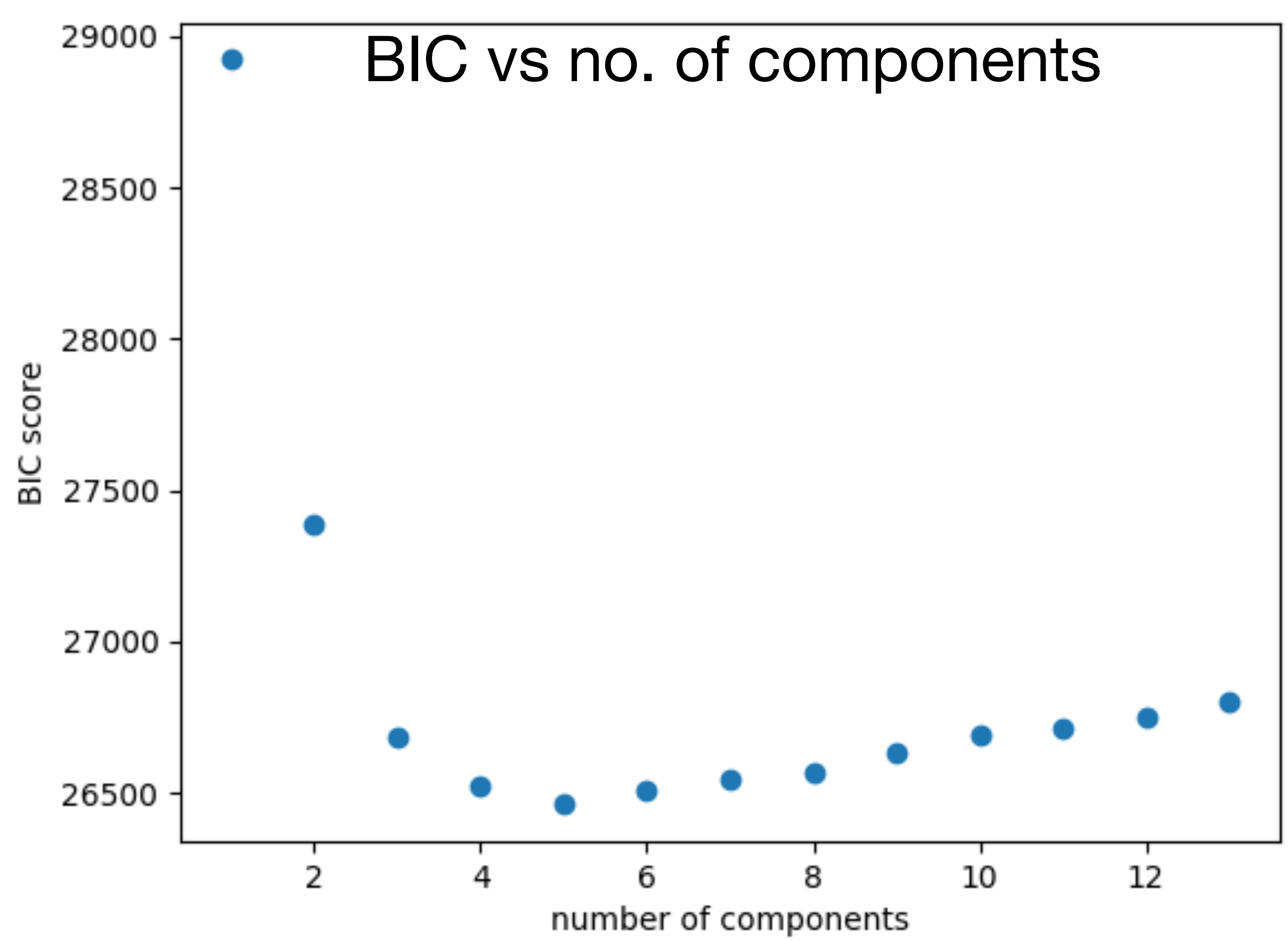


Gaussian Mixture Models

- The Bayesian Information Criterion (BIC) is used to determine the optimal number of Gaussian distributions to fit the data set

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

$\hat{L}$ is the maximized value of the likelihood function, *k* is the number of parameters estimated by the model and *n* is the number of data points in the sample



from scikit-learn examples

GMM for O3a data

- Perform cWB+GMM search over cWB all-sky short duration low frequency burst analysis over O3a with HL network (R. Abbott et al., arXiv: 2107.03701v1) and report confident events.
- Consider cWB all-sky short duration low frequency burst analysis over O3a with HL network (select events with rho0 >5.5 and netcc0 >0.5).
- We choose 12 attributes to characterise each trigger: rho0, netcc0, netcc2, neted0, norm, penalty, Qveto0, Qveto1, ecor, Lveto0, Lveto1, Lveto2.
- We excluded frequency and duration because it is a function of the injected signal population and the Burst MDCs are currently not astrophysically motivated.
- We reparametrize the cWB attributes such that it can fit with optimum Gaussians using functions like log and inverse sigmoid.
  - Combine two attributes to produce Lratio: Lveto1/Lveto0.
  - Reparametrized the attributes netcc0, netcc2 and Lveto2 and Lratio with inverse sigmoid function and penalty, rho0, neted, ecor and Qvetos with logarithmic function.

Re-parameterization

Model	No. of optimum Gaussians in GMM	
	Original attribute set	Re-parametrized attribute set
Signal	113	90
Noise	115	82

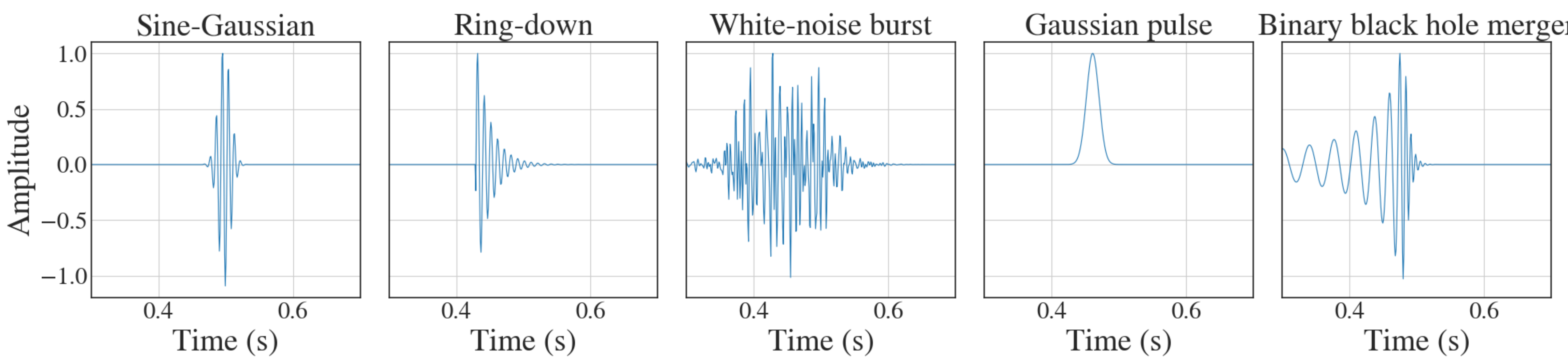
Original attribute set	Re-parametrized attribute set
η_c	$log_{10}(\eta_c)$
c_{c0}	$logit(c_{c0})$
c_{c2}	$logit(c_{c2})$
N_{ED}	$log_{10}(N_{ED} + 10^3)$
E_c	$log_{10}(E_c)$
N_{norm}	N_{norm}
χ^2	χ^2
Q_{veto0}	$log_{10}(Q_{veto0} + 1)$
Q_{veto1}	$log_{10}(Q_{veto1})$
L_{ratio}	$logit(L_{ratio})$
L_{veto2}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

Training and training

- Around 1000 years of background are generated by time-shifting the HL network zero lag livetime of 102.56 days data.
- O3a background and simulation data divided into 3 datasets
- Tuning data: (10% of the full data) - Used to determine optimum attribute subset, by minimising BIC value on Signal Injections.
- Training data: (70% of the full data) - Training data used to obtain means, variances and weights of the GMM. (Training the model parameters).
- Testing data: (20% of the full data) - We are unable to estimate event significance estimate to beyond 1 per 200 years (0.2 x 1000 years)

Modelling the signal space

- We consider the same set of simulated signals as the O3a all-sky burst search (described in R. Abbott *et al*., arXiv:2107.03701)
    - A set of ad hoc waveforms sine-Gaussian wavelets (SG), Gaussian pulses (GA), and band-limited white-noise bursts (WNB).
    - The simulation tuning/training/testing data is split in the same proportions as the background data



- Waveforms from core-collapse supernovae distributed uniform-in-distance.

Constructing a detection statistic

- For given model parameters and number of Gaussian, the maximum log likelihood statistics for each trigger given by -

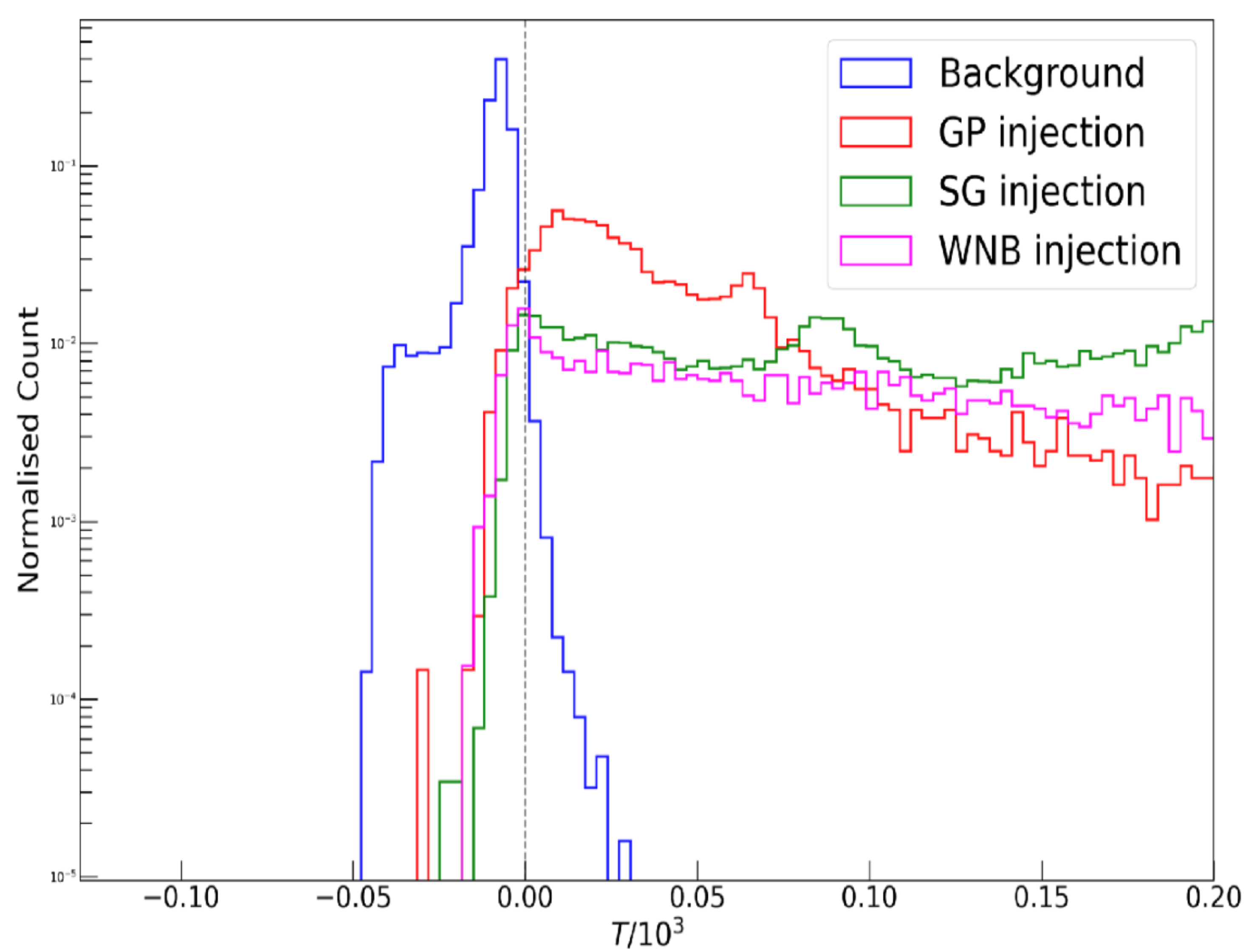
$$W = \ln(\hat{\mathcal{L}})|_{\hat{K}}$$

- The GMM detection statistic for each trigger as follows -

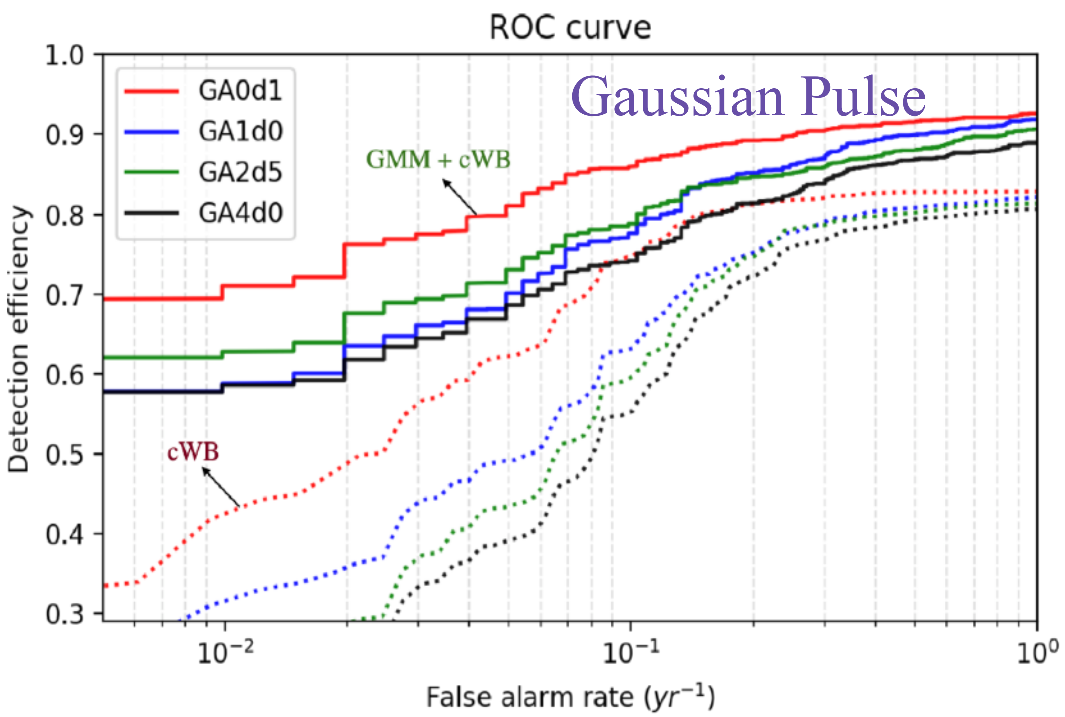
$$T = W_{\mathrm{s}} - W_{\mathrm{g}}$$

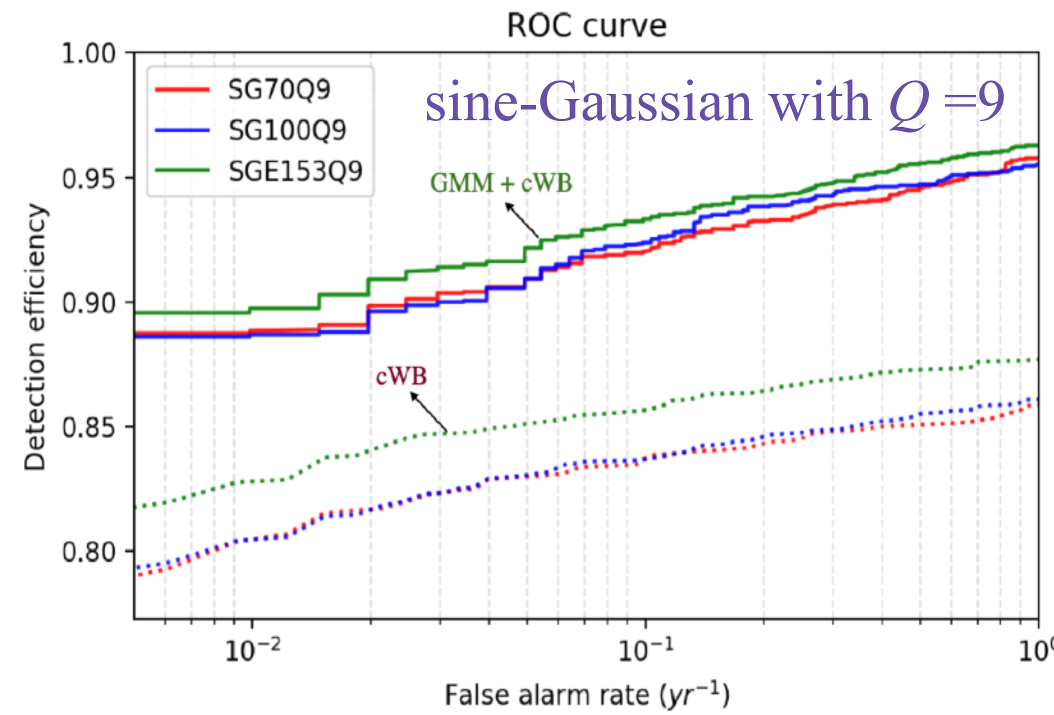
For data consist of two distinct classes, signals (s) and noisy background glitches (g).

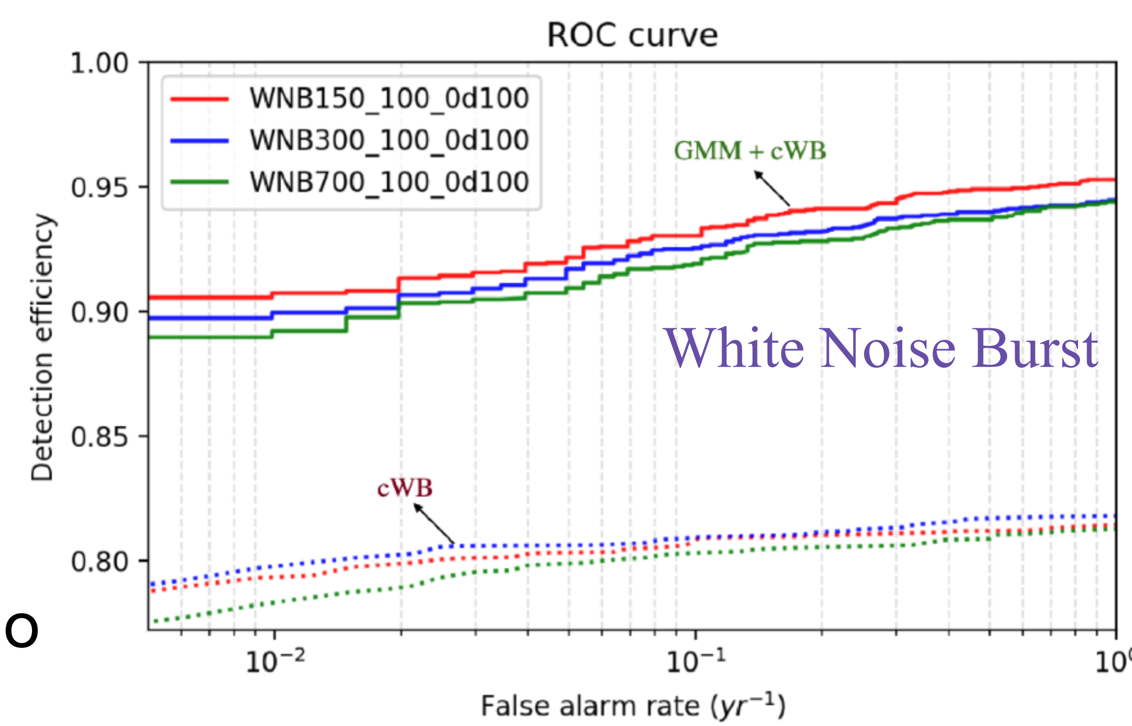
GMM detection statistic distribution for O3a short gravitational-wave bursts



Improved sensitivity with GMM





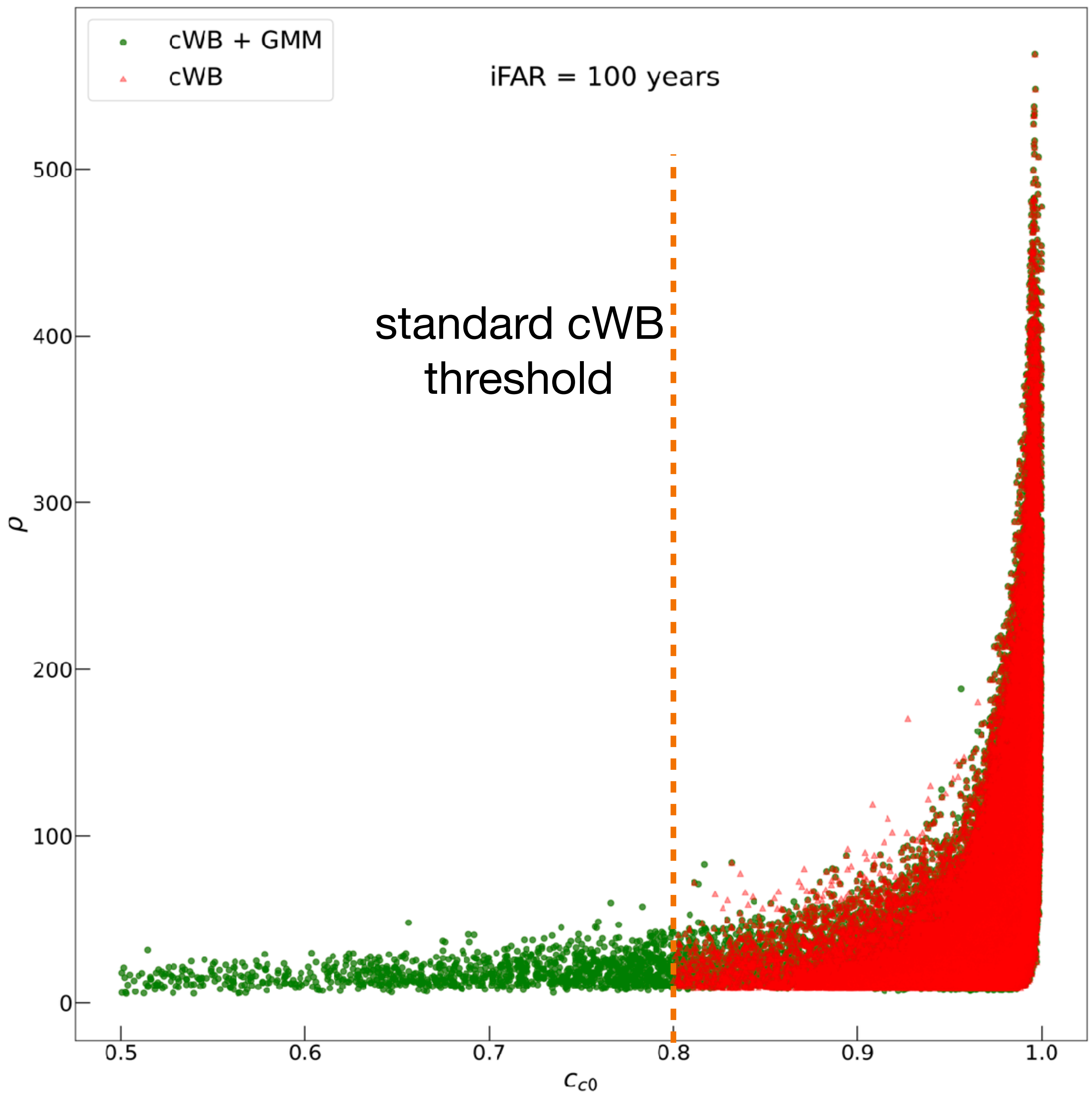


GMM method mitigates "blip-type" glitches during O3 run.

cWB+GMM: vary T
standard cWB: vary rho

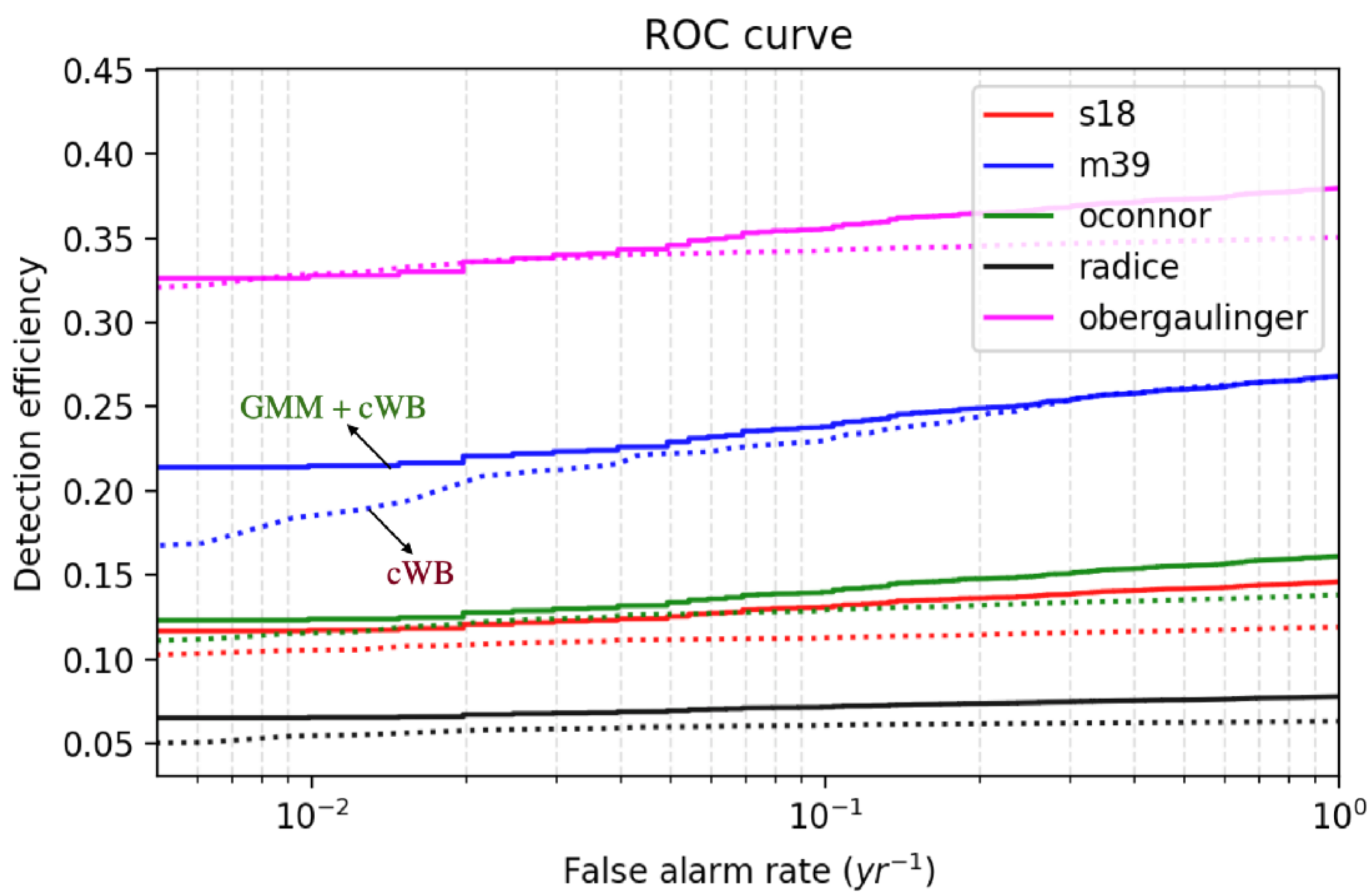
https://wiki.ligo.org/Bursts/O3-Cwb-LF

Improved sensitivity with GMM



Robustness test

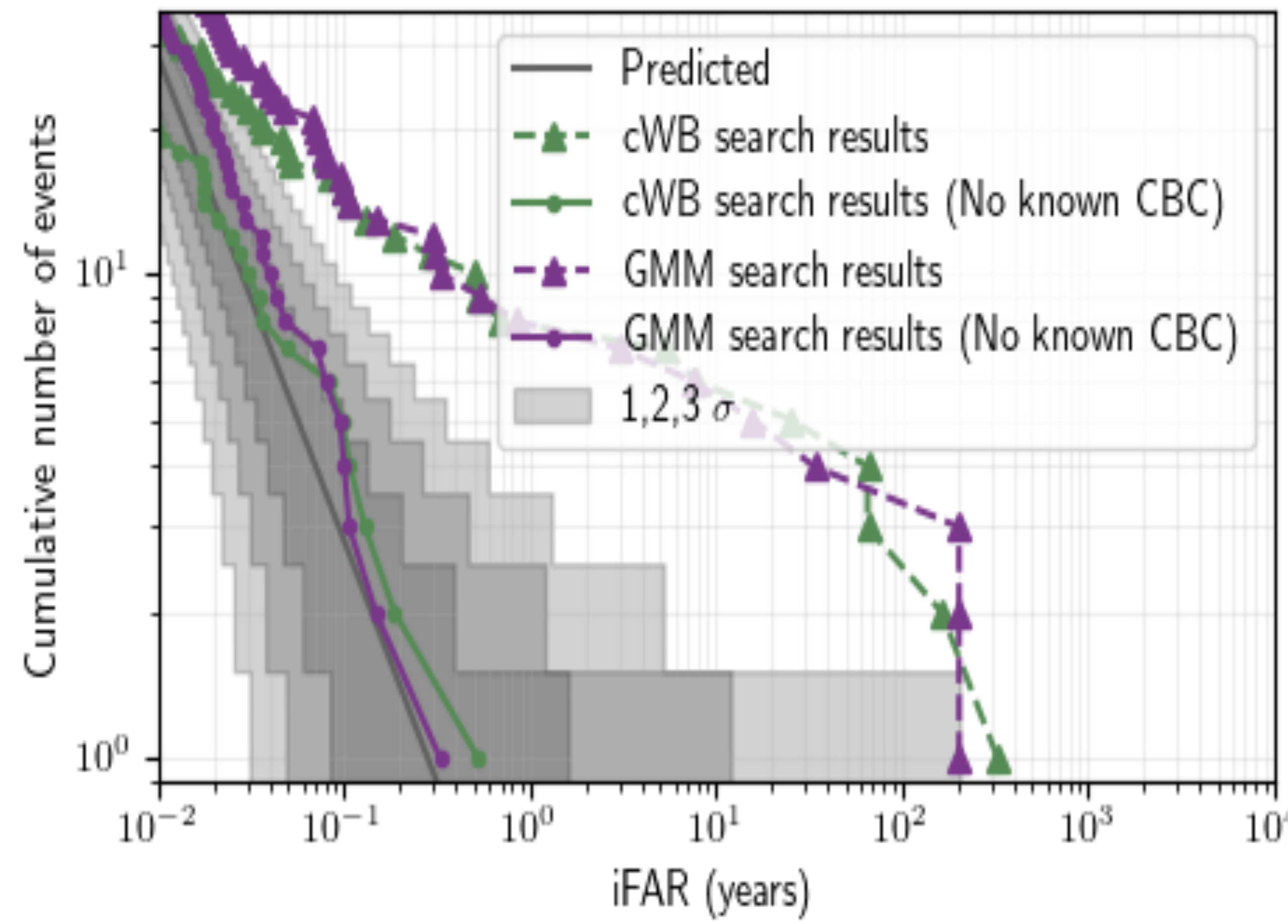
- Tested on CCSNe injections, which not included in the training data set to proves the robustness against the different morphologies of waveforms and distribution.
- Training data: Set of waveforms (Gaussian Pulse, sine-Gaussian wavelets and White Noise Burst).



O3a BBH observations

- Analyzed the low frequency [16,1024] Hz region of HL network.

- Found 15 known BBHs (same as cWB all sky search).
- The loudest event excluding known CBC are at UTC 2019-09-30 23:46:52 has an iFAR of 0.33 years (0.008 years in cWB) and at UTC 2019-05-11 04:12:15 with and iFAR of 0.15 years (0.002 years in cWB).



- The loudest events excluding the known CBC in cWB search are occurred at UTC 2019-09-28 02:11:45 and UTC 2019-08-04 08:35:43 has an iFAR of 0.53 years and 0.19 years respectively. Those two events in cWB plus GMM search shows an iFAR of 0.006 and 0.05 respectively.

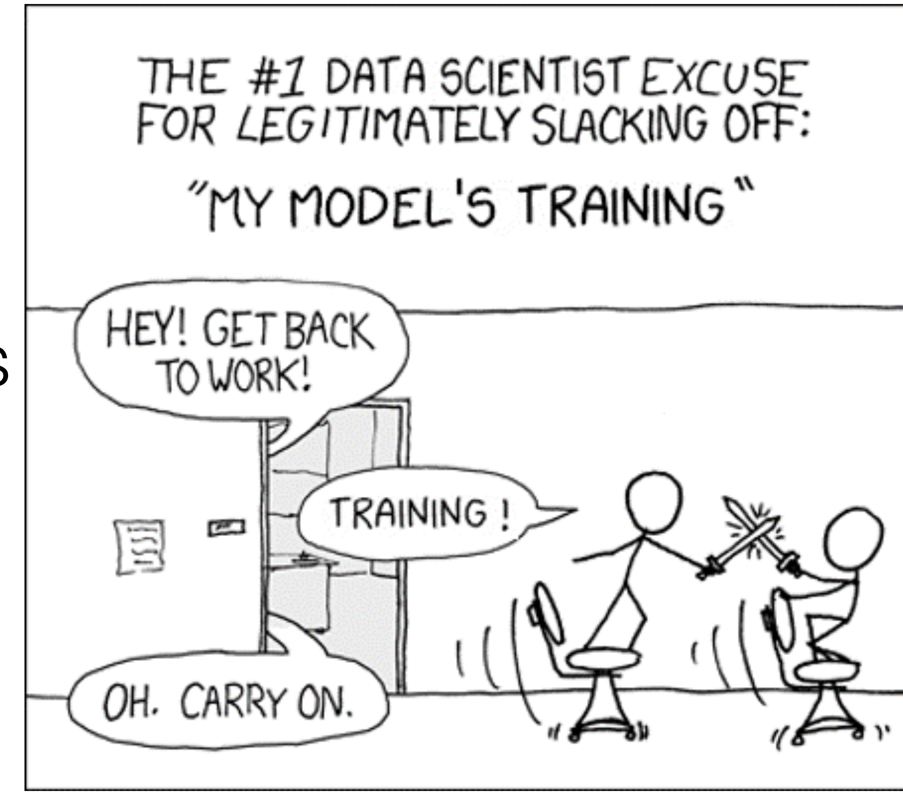
O3a BBH observations

- We consider only 200 years of background, since we used 20% of O3a background as test data.
- GMM method is more sensitive to BBH Events with total source mass > 60 solar mass when compared to the cWB all sky O3a search,
- GW190412m has significantly asymmetric component masses and is an exception.

Event	η_c	Q_{veto}	c_{c0}	χ^2	N_{norm}	T	$M(M_\odot)$	iFAR in years: cWB+GMM	iFAR in years: cWB
GW190408_181802	8.59	0.92	0.96	0.13	5.09	-0.41	$43.0^{+4.2}_{-3.0}$	0.30	25.14
GW190412	11.69	4.16	0.95	0.06	5.4	13.21	$38.4^{+3.8}_{-3.7}$	15.62	14.86
GW190421_213856	6.46	0.31	0.97	-0.07	4.41	-0.38	$72.9^{+13.4}_{-9.2}$	0.30	0.04
GW190426_190642	5.52	0.45	0.88	0.08	4.07	-4.85	$184.4^{+41.7}_{-36.6}$	0.02	0.01
GW190503_185404	7.34	0.34	0.93	-0.02	4.76	1.65	$71.7^{+9.4}_{-8.3}$	0.84	0.70
GW190513_205428	7.05	1.67	0.86	0.15	3.77	-2.99	$53.9^{+8.6}_{-5.9}$	0.07	0.28
GW190517_055101	6.08	0.19	0.88	-0.15	3.05	-2.79	$63.5^{+9.6}_{-9.6}$	0.08	0.01
GW190519_153544	10.13	0.53	0.89	0.01	7.63	18.04	$106.6^{+13.5}_{-14.8}$	33.83	7.78
GW190521	9.24	0.60	0.92	-0.16	10.53	32.45	$163.9^{+39.2}_{-23.5}$	> 200	65.38
GW190521_074359	14.19	0.56	0.96	-0.08	8.44	72.77	$74.7^{+7.0}_{-4.8}$	> 200	326.88
GW190602_175927	7.25	0.43	0.95	-0.13	6.5	0.73	$116.3^{+19.0}_{-15.6}$	0.54	0.51
GW190706_222641	9.29	0.79	0.83	-0.10	7.36	24.93	$104.1^{+20.2}_{-13.9}$	> 200	65.38
GW190727_060333	5.86	0.35	0.96	0.17	4.96	-2.94	$67.1^{+11.7}_{-8.0}$	0.07	0.006
GW190728_064510	6.50	3.94	0.87	-0.13	2.55	-4.93	$20.6^{+4.5}_{-1.3}$	0.02	0.051
GW190828_063405	10.27	0.84	0.82	0.10	5.01	8.78	$58.0^{+7.7}_{-4.8}$	7.52	163.44
GW190915_235702	8.07	0.42	0.95	0.06	4.29	5.29	$59.9^{+7.5}_{-6.4}$	3.07	5.36
GW190929_012149	5.97	0.22	0.85	0.103	3.44	-6.20	$104.3^{+34.9}_{-25.2}$	0.01	0.009

Some thoughts + future work

- By using GMM to model our trigger attribute space, we were able to increase the sensitivity of the cWB all-sky burst search.
- We are now implementing this for the O4 analysis
- It is possible to use the GMM approach for specific signal classes.
- The properties of the simulated signals used to train the GMM (and burst searches in general) should be carefully considered.

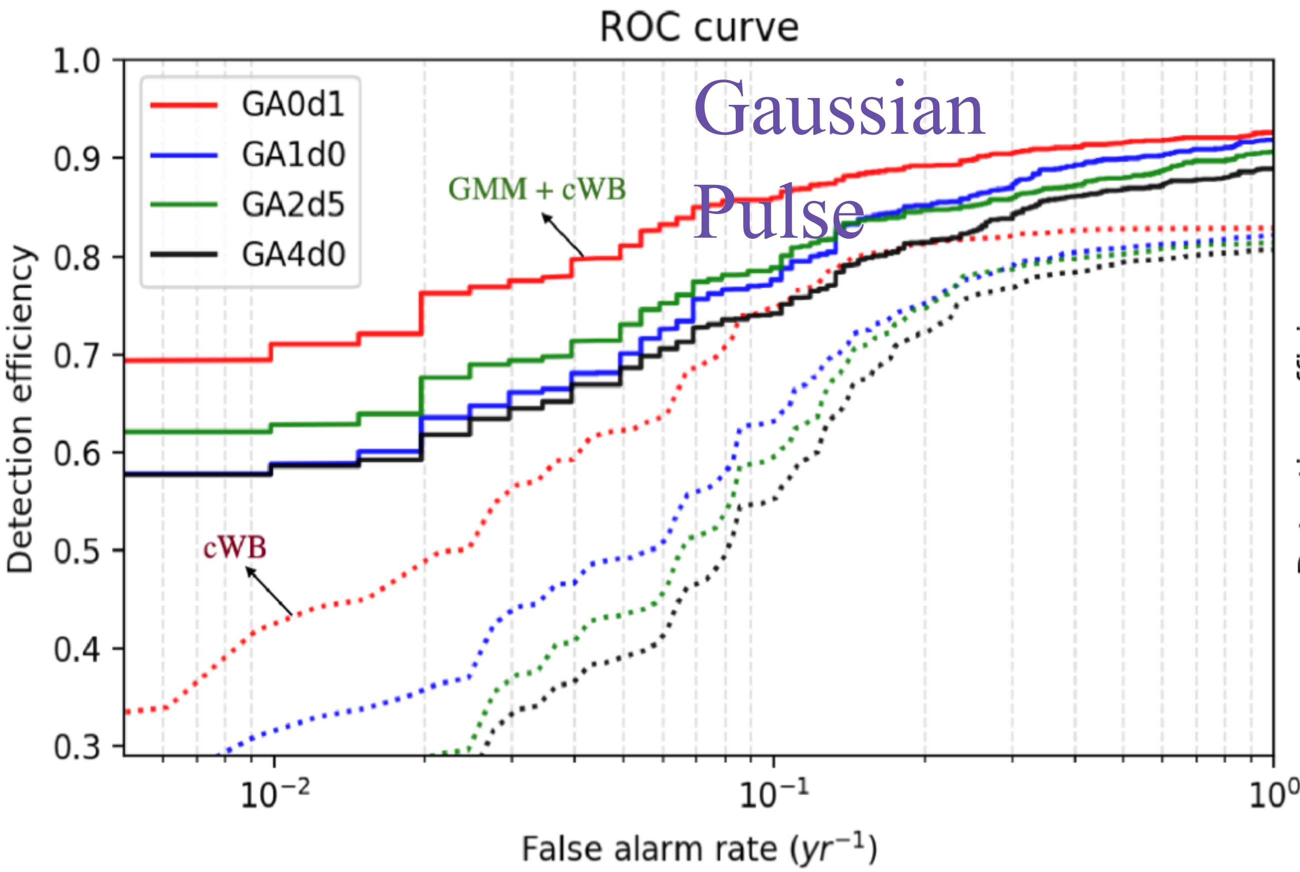


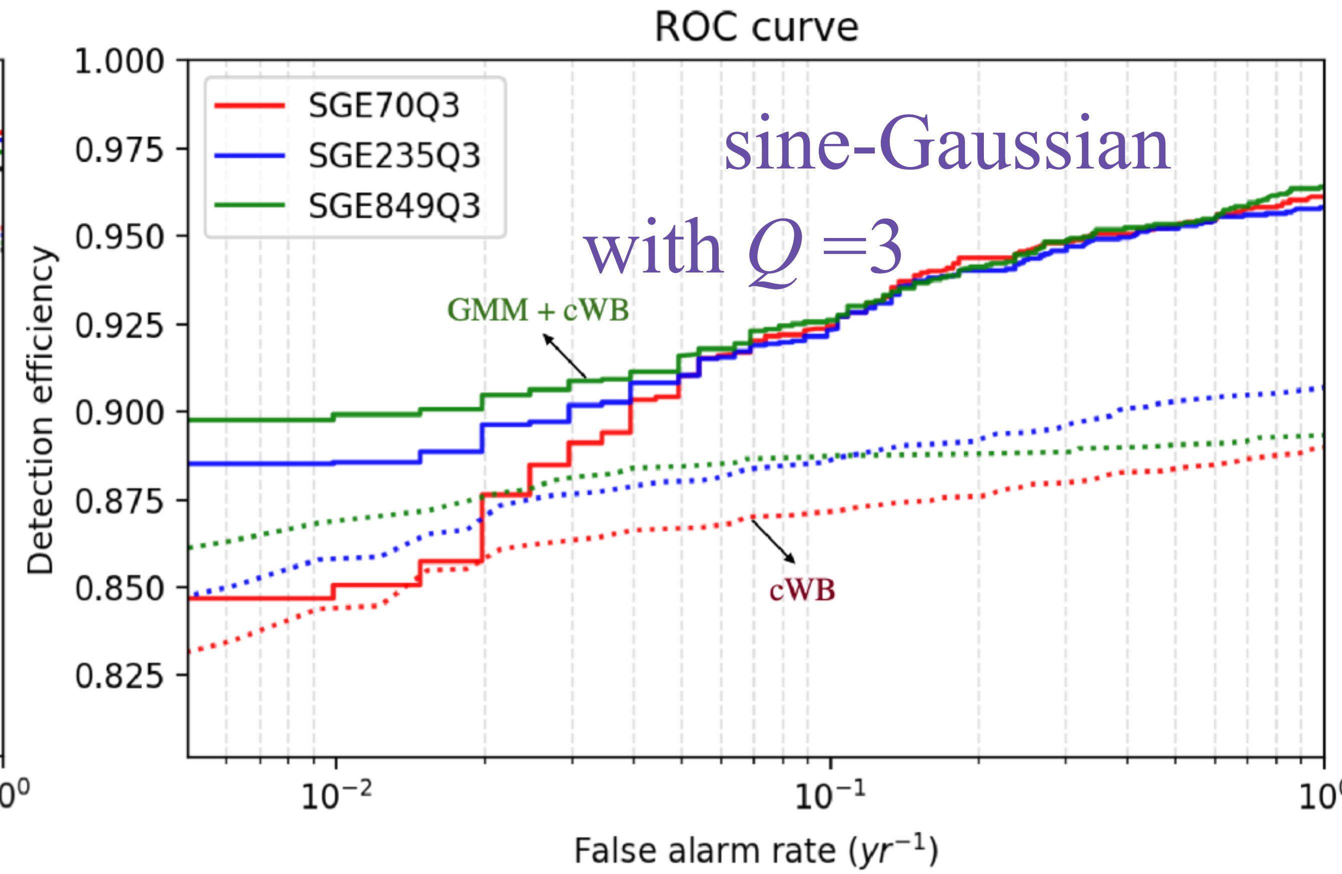
https://xkcd-excuse.com

- Exploring the use of a neural network to classify signal/background triggers.
- In our approach, cWB encodes our data into a set of attributes. What if we use different encoders; could we come up with a more optimal encoding of the data for burst searches?
- Autoencoders for burst searches (anomaly detection): F. Morawski et al., MLST (2021) arXiv:2103.07688

Extra slides

Results



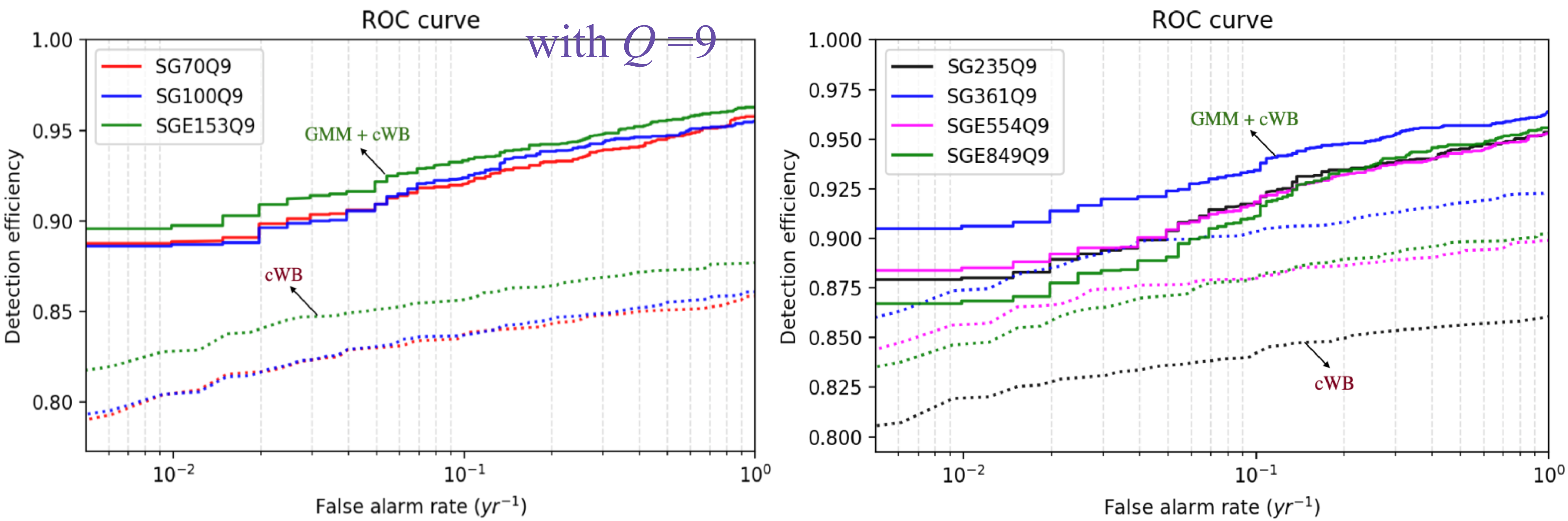


- cWB + GMM improves in detection efficiency for GA which falls in the dirty bin containing a population of very short and very loud (blip-type) glitches.
- GMM method mitigates "blip-type" glitches during O3 run.

https://wiki.ligo.org/Bursts/O3-Cwb-LF

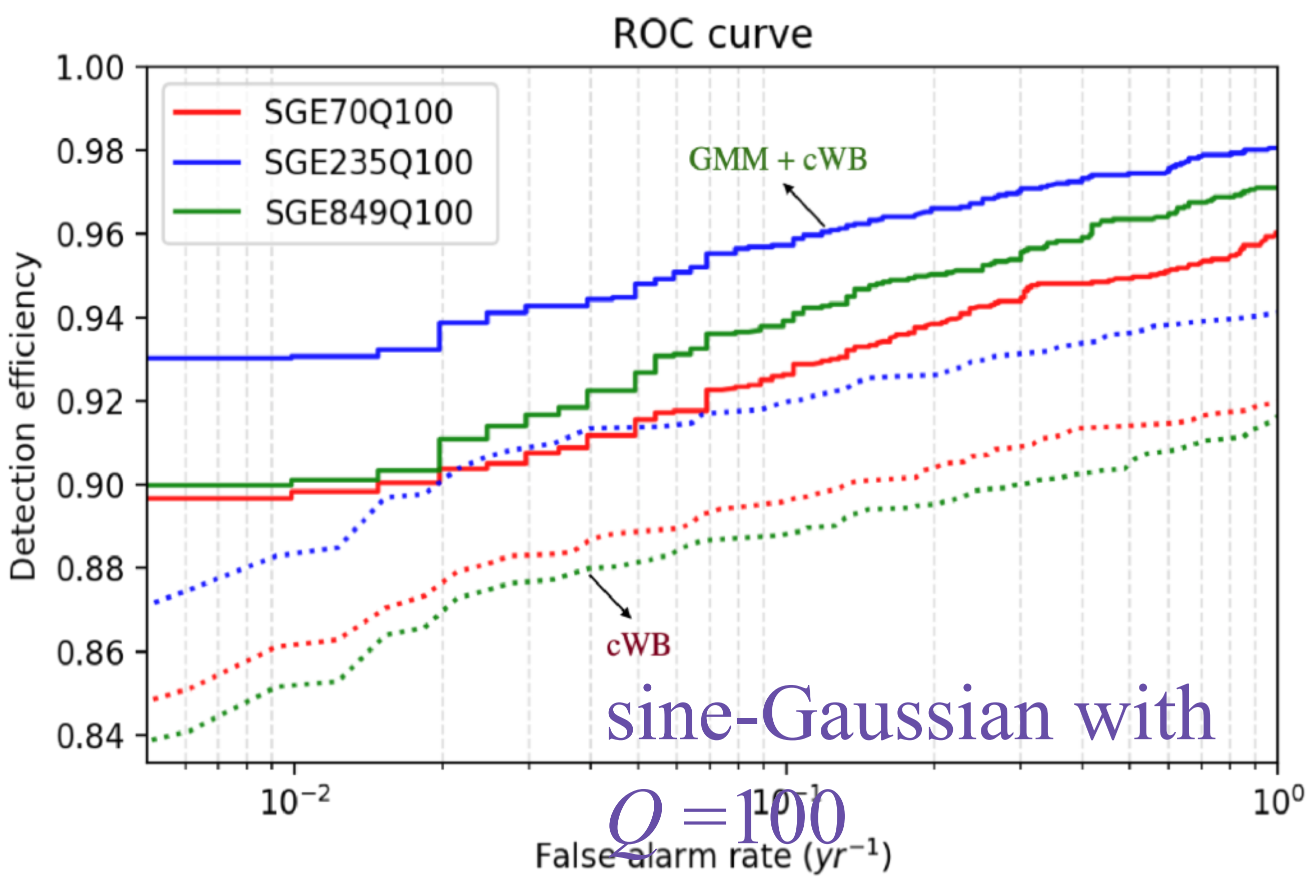
Results

sine-Gaussian with Q =9

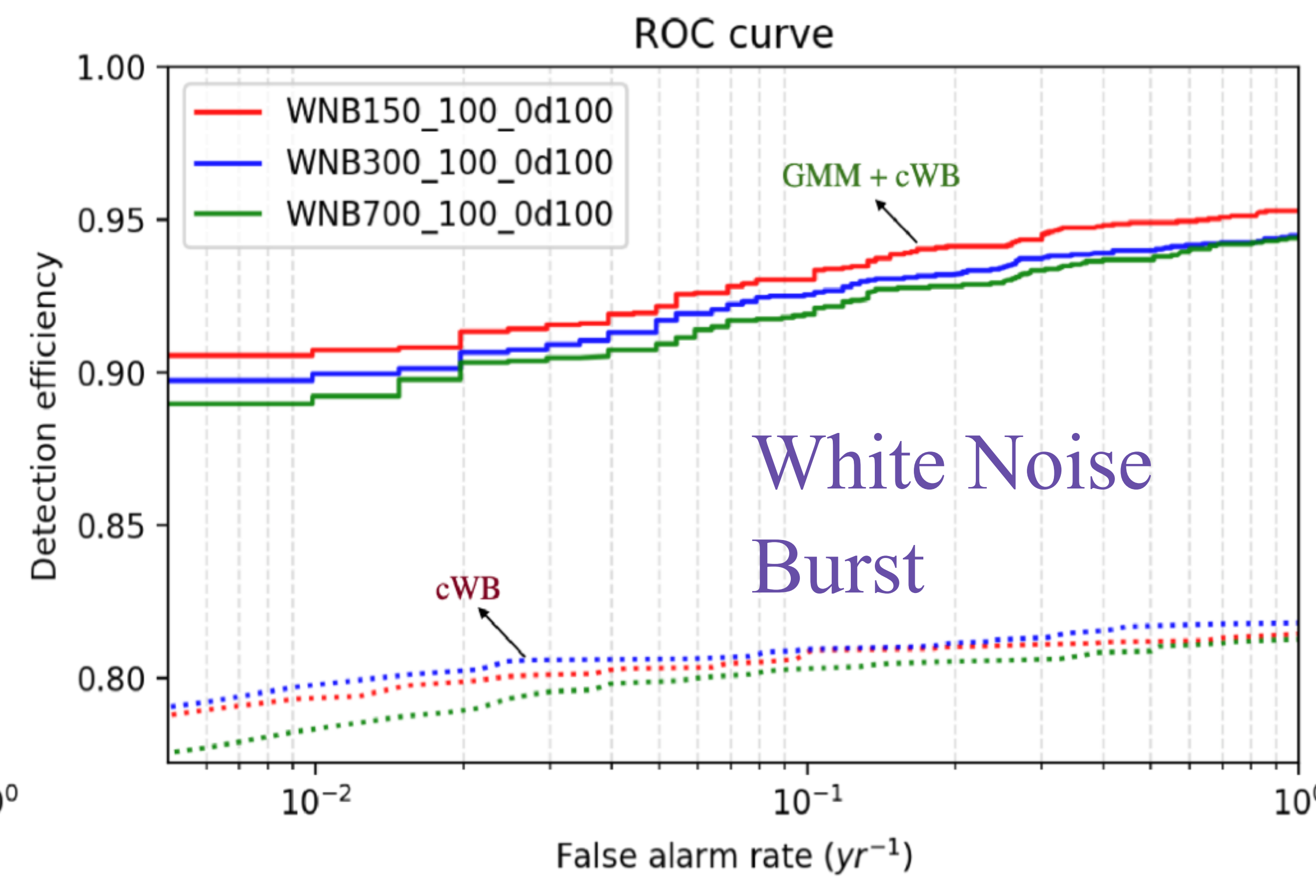


IGR University of Glasgow

Results



sine-Gaussian with $Q = 100$



White Noise Burst

Generative Adversarial Networks

- Generative Adversarial Networks (GANs) pit two neural networks against each other.
- The generator network (**G**) tries to generate data that re (desired) training data.
- The discriminator (**D**) is a classifier network that labels i from the generator (“fake”) or from the training data (“re

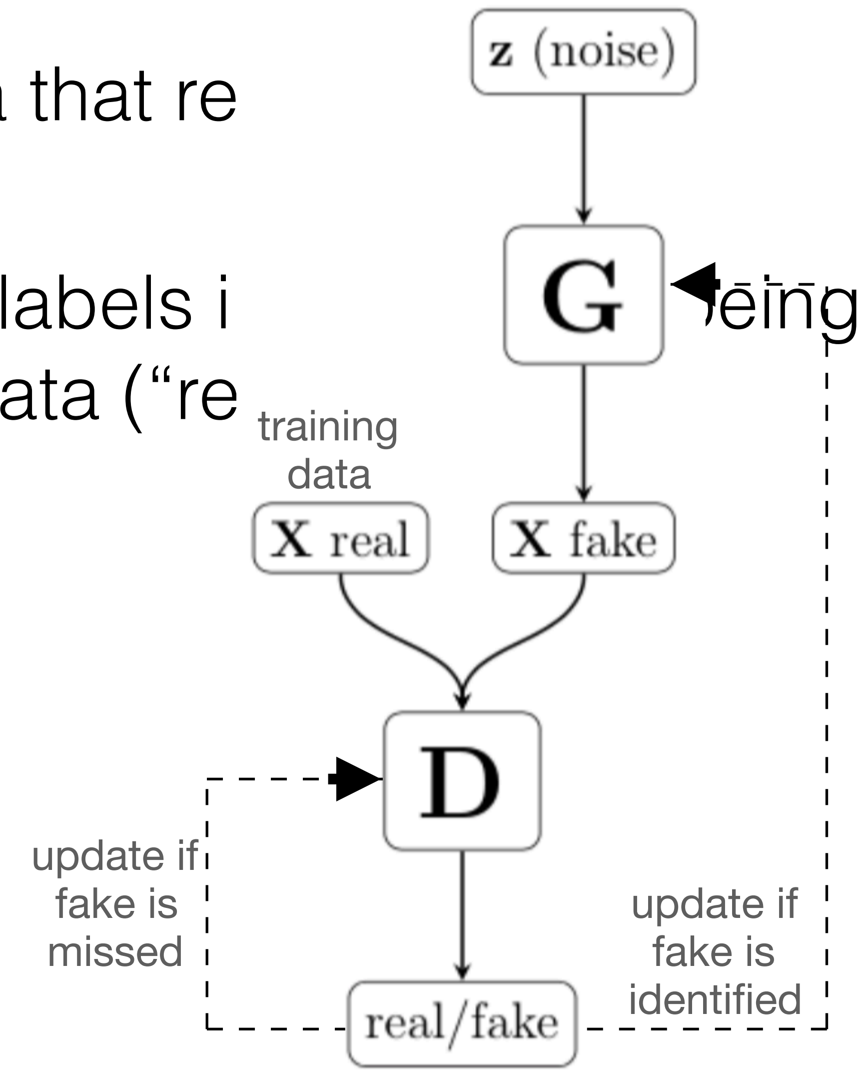


Image generation

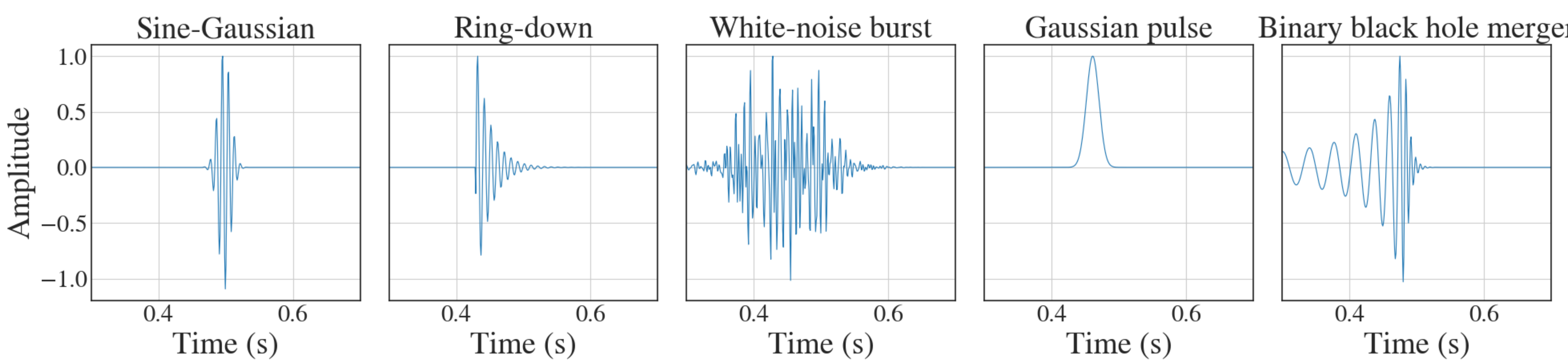
- Recent works have been incredibly

- With conditioning, it is also possibl
  from each class.
  - eg. If GAN is trained on images of ca

https://thispersondoesnotexist.com

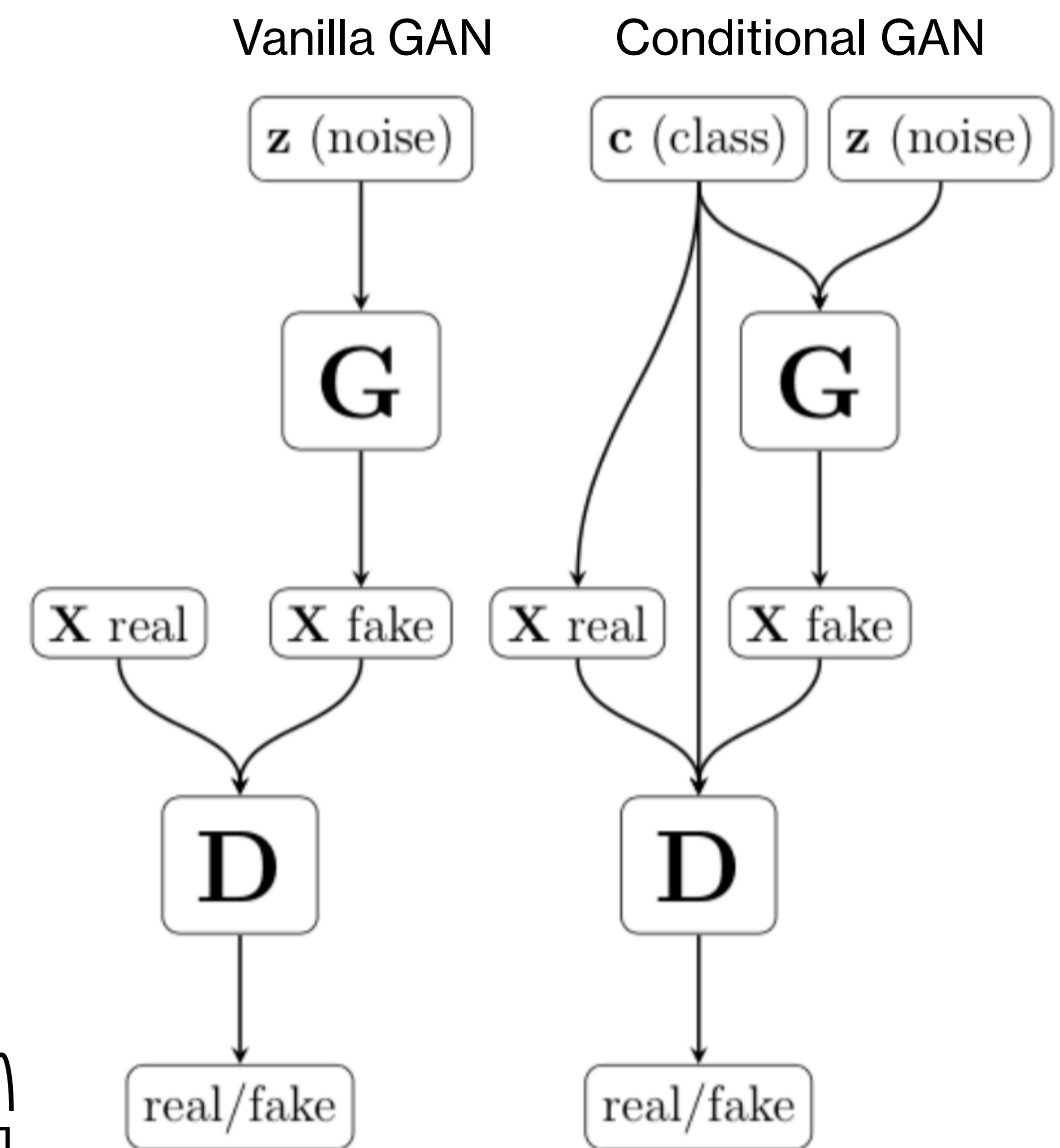
Gravitational-wave bursts

- For Burst (generic transient) searches, we typically do not have accurately modelled signal predictions.
- We use ad hoc waveforms to characterise the sensitivity of Burst searches.
- A well-trained Burst search should be able to span the parameter space between the defined ad hoc waveforms.
- We use GANs to explore the waveform morphologies of “mixed” Burst waveforms.



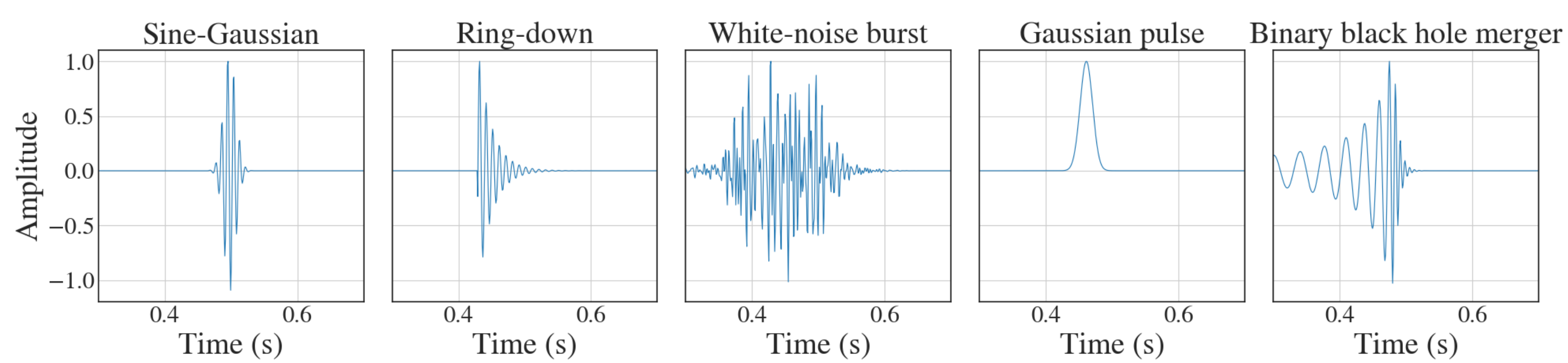
Conditional GANs

- We use a conditional GAN (cGAN) where ea assigned a label.
- “One hot encoding” is used for each wavefc
- There is also a latent space within each clas to vary



Labels

[1,0,0,0,0] [0,1,0,0,0] [0,0,1,0,0] [0,0,0,1,0] [0,0,0,01]



https://arxiv.org/abs/2103.01641

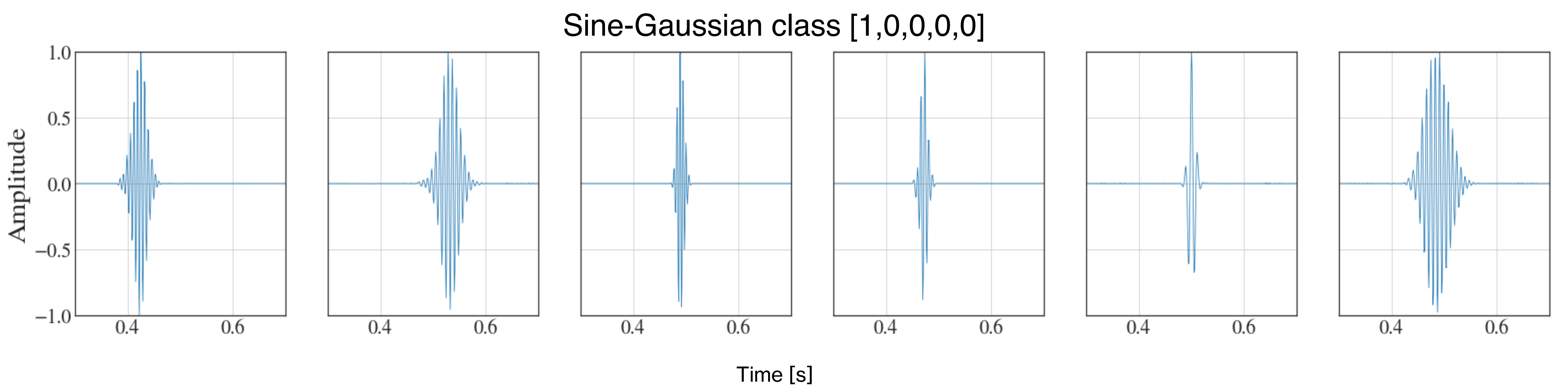
Training data

- Binary Black Hole parameters - masses: power law [30, 70] M_{sol}, d_L: fixed, zero spin, other parameters: usual astrophysical priors.
- All waveforms sampled at 1024 Hz.

Signal type	Duration	Frequency (Hz)	Decay (s)	Central time epoch (s)
Sine-Gaussian	1s	70 - 250	0.004 - 0.03	0.4 - 0.6
White-noise burst	1s	70 - 250	0.004 - 0.03	0.4 - 0.6
Gaussian Pulse	1s	-	0.004 - 0.03	0.4 - 0.6
Ring-down	1s	70 - 250	0.004 - 0.03	0.4 - 0.6
BBH*	1s	-	-	0.4 - 0.6

Multiple generations from one class

- The latent variable (*z*) for each waveform class is a 100-number array
- By fixing the class vector but changing the latent variable, we can produce signals from the same class with different physical properties.



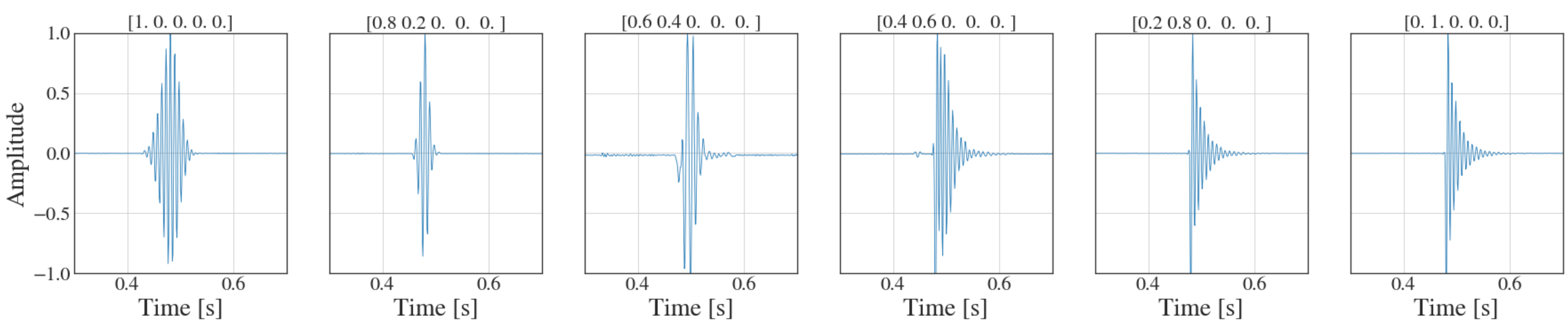
https://arxiv.org/abs/2103.01641

Interpolating between classes

- Assuming that the cGAN has a 'smooth' space between all five classes, we explore the signal morphologies by interpolating between class labels:
- Sine gaussian [1, 0] → Ringdown [0, 1]

[1, 0] → [0.8, 0.2] → [0.6, 0.4] → [0.4, 0.6] → [0.2, 0.8] → [0, 1]

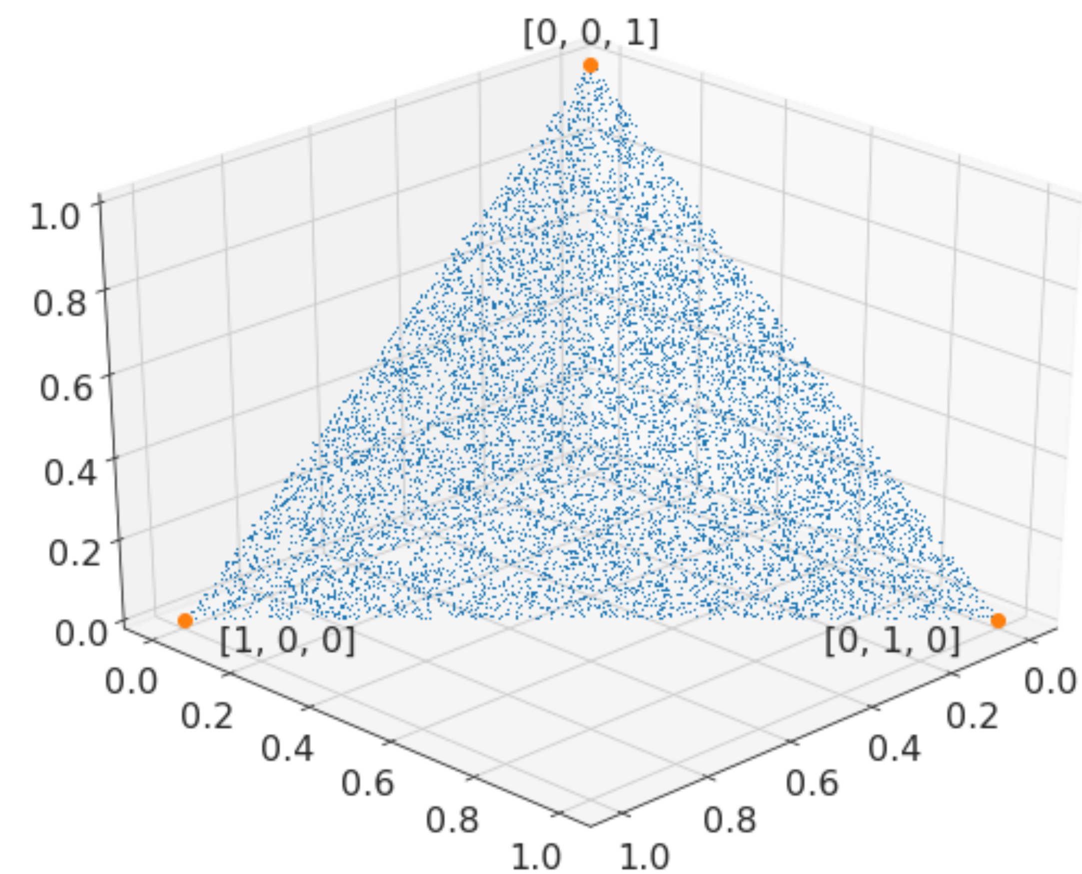
Modelled signal 'Unmodelled signal' Modelled signal



https://arxiv.org/abs/2103.01641

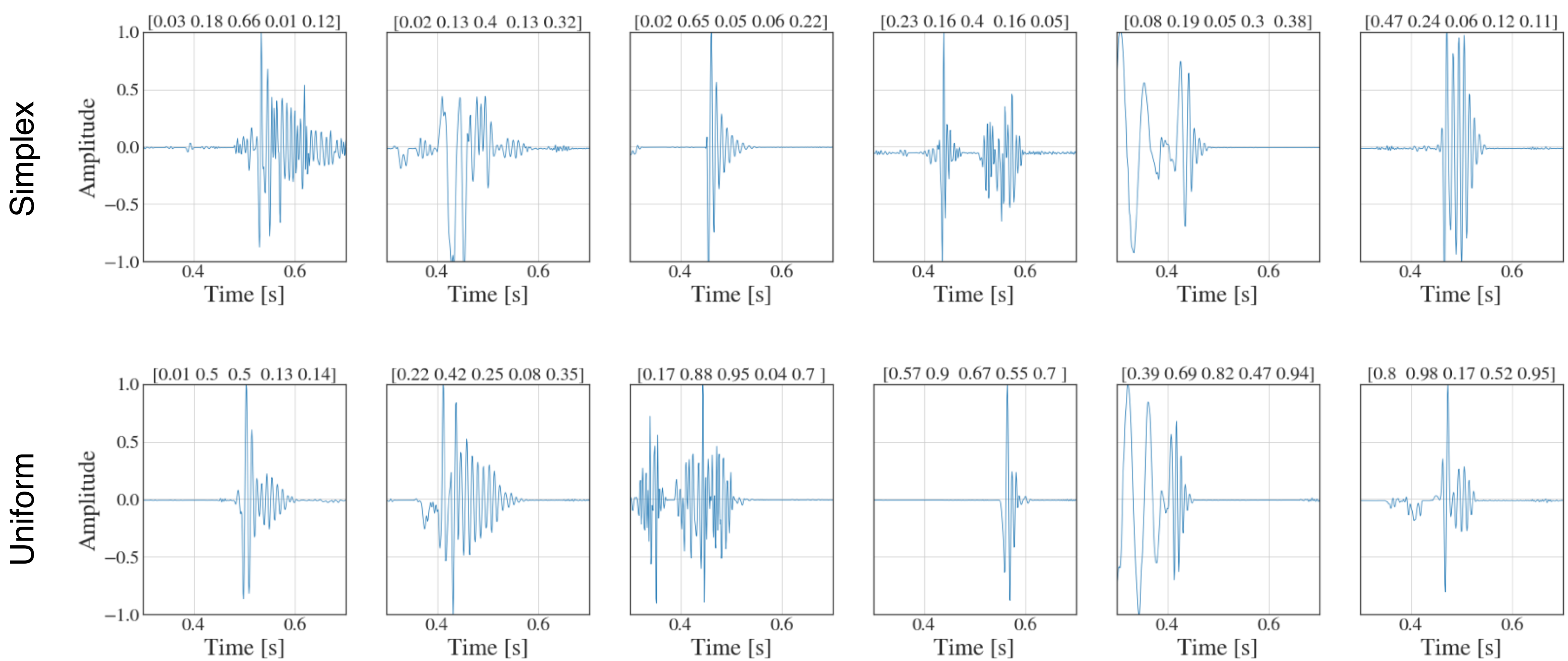
Interpolating between classes

- We can consider random mixtures of class
- **Vertex** - Points that lie at the corners of th Closest to training data.
- **Simplex** – Sampled points on a simplex, 5 links all vertices. It is a subspace of Unifor
- **Uniform** - sampling uniformly within the 5- hyper-cube.



https://arxiv.org/abs/2103.01641

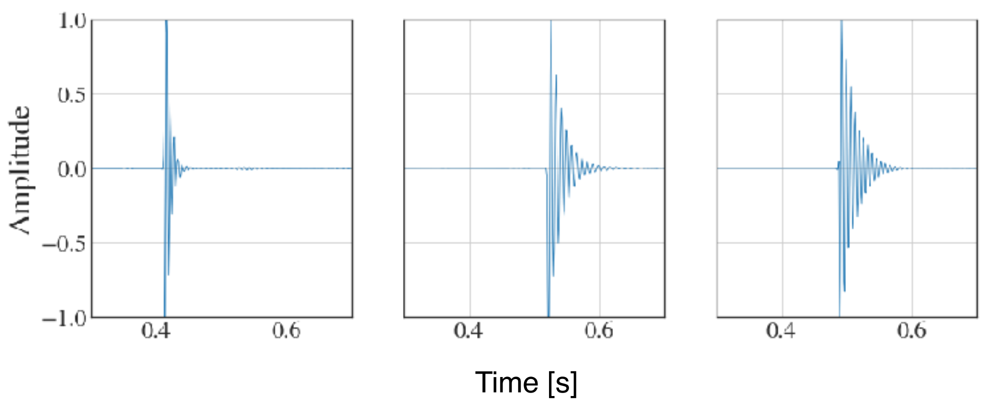
Interpolating between classes



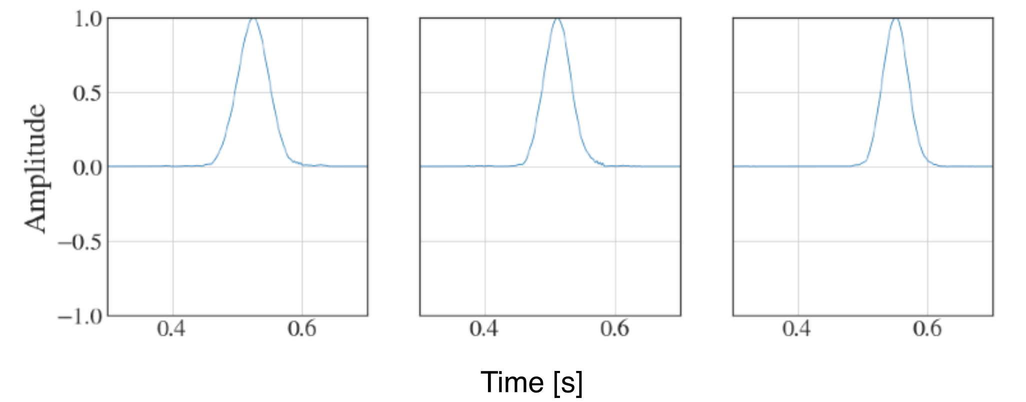
https://arxiv.org/abs/2103.01641

GAN generations

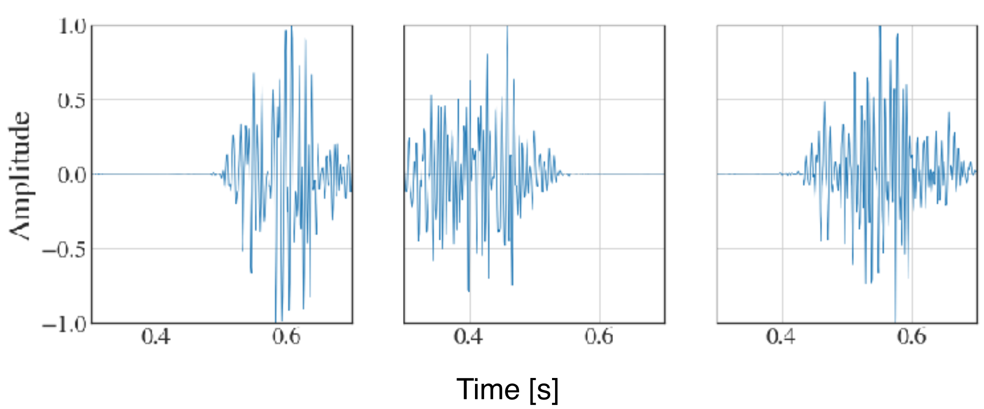
Ring-down



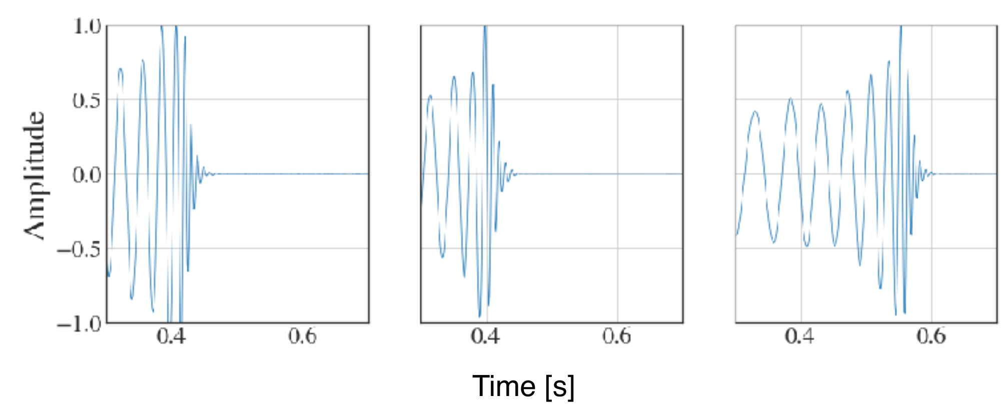
Gaussian blip



White-noise burst



Binary black hole



Characterising waveform generations

- A basic search pipeline using a CNN in order to compare the sensitivity of such a search using different GAN generated waveforms in Gaussian noise
  - 3 different CNN networks; one trained on Vertex generations, another on Simplex generations and the last on Uniform generations
- We are interested in the relative sensitivity as a function of the types of waveforms used for training the network.
- Set a threshold corresponding to a false alarm probability of 10^{-3}

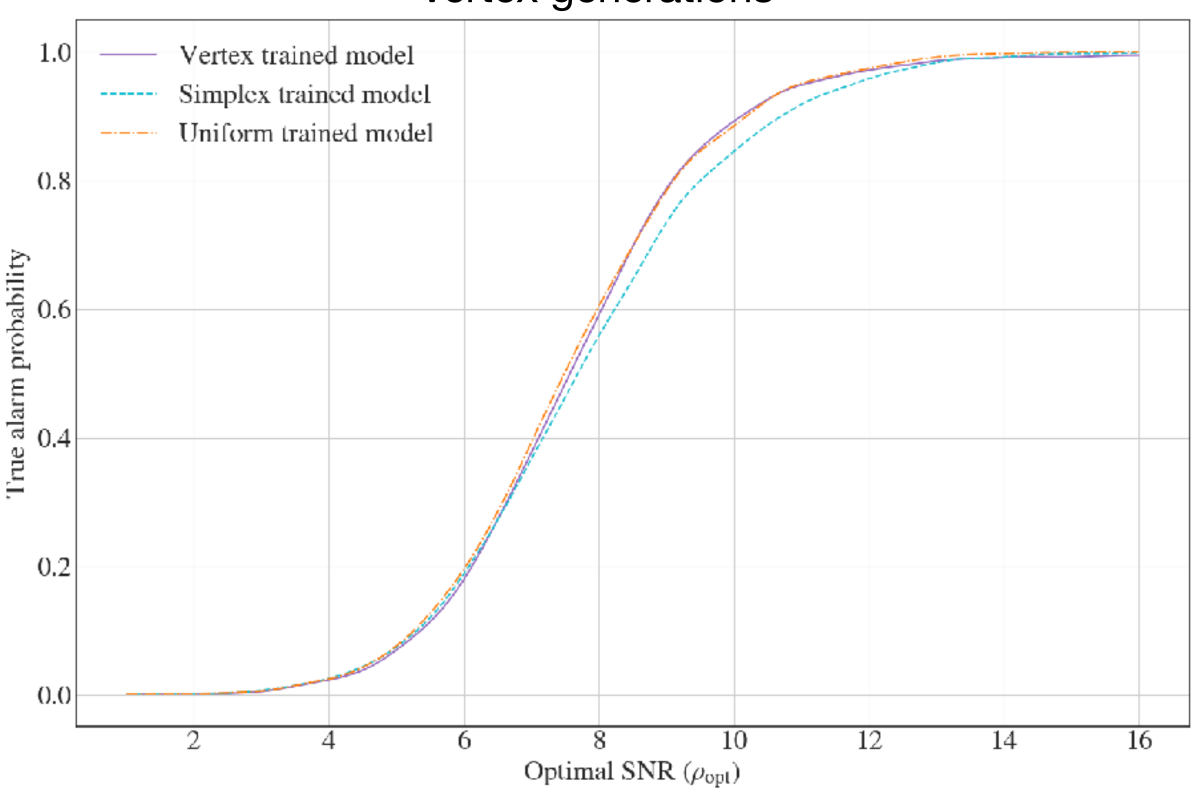
- Reminder: Vertex generations correspond to the standard set of waveforms used in Burst searches

https://arxiv.org/abs/2103.01641

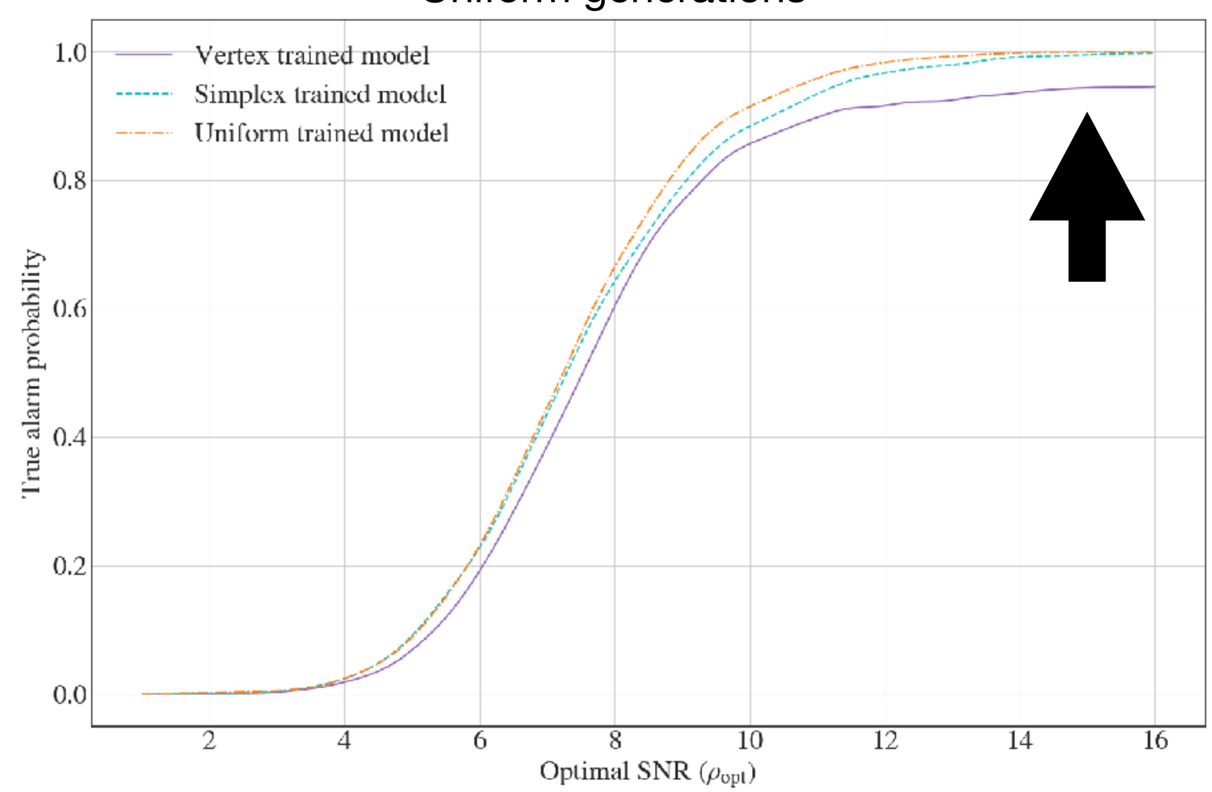
Characterising waveform generations

- Vertex model only manages full detection when tested on vertex data and misses even the strongest signals from Uniform generations
- Of the two methods of generating unmodelled signals, the Uniform generation produces more general morphologies that do not negatively effect the performance on the modelled signals.
- CNN-based burst searches will be more sensitive to a wider parameter space if trained on signals generated from entire cGAN burst waveform space.

Vertex generations



Uniform generations



https://arxiv.org/abs/2103.01641