Enhancing transient gravitational wave analyses with machine learning

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

• Introduction to gravitational wave burst searches

• Coherent Waveburst

• Gaussian Mixture Modelling

• Application to O3a all-sky burst search

• 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
  - 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
  - Gaussian mixture models for gravitational wave burst searches (this talk)
Global gravitational wave network
## Gravitational wave signal types

<table>
<thead>
<tr>
<th>Modelled</th>
<th>Unmodelled</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compact Binary Coalescence</strong></td>
<td><strong>Burst</strong></td>
</tr>
<tr>
<td>short</td>
<td><img src="image" alt="Burst" /></td>
</tr>
<tr>
<td><strong>Continuous</strong></td>
<td><strong>Stochastic</strong></td>
</tr>
<tr>
<td>long</td>
<td><img src="image" alt="Stochastic" /></td>
</tr>
</tbody>
</table>
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

Greenwich Hour Angle (hours)
Declination (degrees)

GEO

$F_+^2 + F_x^2$

Livingston

Declination (degrees)
Greenwich Hour Angle (hours)
Burst searches

- Lets formulate our data so that

\[ \tilde{d}_w = F_w \tilde{h} + \tilde{n}_w, \]

or

\[
\begin{bmatrix}
\tilde{d}_{w_1} \\
\tilde{d}_{w_2} \\
\vdots \\
\tilde{d}_{w_D}
\end{bmatrix} =
\begin{bmatrix}
F^+_{w_1} & F^\times_{w_1} \\
F^+_{w_2} & F^\times_{w_2} \\
\vdots & \vdots \\
F^+_{w_D} & F^\times_{w_D}
\end{bmatrix}
\begin{bmatrix}
\tilde{h}^+ \\
\tilde{h}_\times
\end{bmatrix} +
\begin{bmatrix}
\tilde{n}_{w_1} \\
\tilde{n}_{w_2} \\
\vdots \\
\tilde{n}_{w_D}
\end{bmatrix},
\]

sky location \rightarrow frequency

\[ F_w(\hat{\Omega}_s, f) \equiv \begin{bmatrix} F^+_w & F^\times_w \end{bmatrix} =
\begin{bmatrix}
F^+_{w_1} & F^\times_{w_1} \\
F^+_{w_2} & F^\times_{w_2} \\
\vdots & \vdots \\
F^+_{w_D} & F^\times_{w_D}
\end{bmatrix} \]

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

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

- For 3 detectors, $A$ is constructed by

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

Chatterji et al, PRD 74, 082005 (2006)
Likelihood or Energy Measures

Likelihoods used in this analysis:

Hard Constraint
\[ E_{HC} = |\hat{F}_+ \cdot \vec{d}|^2 \]

Null Energy
\[ E_{NULL} = |\hat{K} \cdot \vec{d}|^2 \]

Incoherent Energy
\[ E_{INC} = \sum_{\alpha} |K_{\alpha} \cdot d_{\alpha}|^2 \]

Klimenko, Mohanty, Rakhmanov, & Mitselmakher,

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

**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 L veto0 L veto1 L veto2 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.


• 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
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 $w_i$, the probability that the data $\mathbf{x}$ belongs to the model is

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

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

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

• The log-likelihood is thus

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

• Maximising the log-likelihood, we find

\[
\hat{\mu}_k = \frac{\sum_{i=1}^{n} r_{ik} x_i}{N_k} \quad \hat{\Sigma}_k = \frac{1}{N_k} \sum_{i=1}^{n} r_{ik} (x_i - \mu_k)(x_i - \mu_k)^T \quad r_{ik} = \frac{w_k \mathcal{N}(x_i|\mu_k, \Sigma_k)}{\sum_{j=1}^{K} w_j \mathcal{N}(x_i|\mu_j, \Sigma_j)}
\]

• 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.
1.3 Using Gaussian mixture models

Figure 1.1: The left panel shows the data points using which the model was constructed. The four colours correspond to the four different Gaussians from which the samples were drawn. The right panel shows a plot for the BIC score for different number of clusters. In this case the BIC score is minimum for the model with four clusters.

Table 1.1: The table shows the means of the Gaussians from which the samples were drawn and the ones obtained after fitting, in the second and third columns respectively.

Figure 1.2: The above plot shows the ellipses corresponding to the Gaussians obtained after fitting.

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

\[
BIC = k \times \ln(n) - 2\ln(\hat{L})
\]

\(\hat{L}\) 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

![BIC vs no. of components](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

<table>
<thead>
<tr>
<th>Model</th>
<th>No. of optimum Gaussians in GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original attribute set</td>
</tr>
<tr>
<td>Signal</td>
<td>113</td>
</tr>
<tr>
<td>Noise</td>
<td>115</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Original attribute set</th>
<th>Re-parametrized attribute set</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_c$</td>
<td>$\log_{10}(\eta_c)$</td>
</tr>
<tr>
<td>$c_{c0}$</td>
<td>$\logit(c_{c0})$</td>
</tr>
<tr>
<td>$c_{c2}$</td>
<td>$\logit(c_{c2})$</td>
</tr>
<tr>
<td>$N_{ED}$</td>
<td>$\log_{10}(N_{ED} + 10^3)$</td>
</tr>
<tr>
<td>$E_c$</td>
<td>$\log_{10}(E_c)$</td>
</tr>
<tr>
<td>$N_{norm}$</td>
<td>$N_{norm}$</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>$\chi^2$</td>
</tr>
<tr>
<td>$Q_{veto0}$</td>
<td>$\log_{10}(Q_{veto0} + 1)$</td>
</tr>
<tr>
<td>$Q_{veto1}$</td>
<td>$\log_{10}(Q_{veto1})$</td>
</tr>
<tr>
<td>$L_{ratio}$</td>
<td>$\logit(L_{ratio})$</td>
</tr>
<tr>
<td>$L_{veto2}$</td>
<td>$\logit(L_{veto2} \times 0.99)$</td>
</tr>
</tbody>
</table>
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:

\[ W = \ln(\hat{L})|_{\hat{K}} \]

• The GMM detection statistic for each trigger as follows:

\[ T = W_s - W_g \]

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

standard cWB threshold

iFAR = 100 years
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 an 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.

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

• 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 $Q=9$
Results

sine-Gaussian with $Q = 100$

White Noise Burst

ROC curve

Detection efficiency vs. False alarm rate ($yr^{-1}$)
Generative Adversarial Networks

• Generative Adversarial Networks (GANs) pit two neural networks against each other.

• The generator network \((G)\) tries to generate data that resembles the training data.

• The discriminator \((D)\) is a classifier network that labels input data as being from the generator (“fake”) or from the training data (“real”).
Image generation

• Recent works have been incredibly successful in image generation.

• With conditioning, it is also possible to control the combination of features from each class.
  - eg. If GAN is trained on images of cats and pizza, it can create a pizza-cat.

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 each waveform morphology is assigned a label.
• “One hot encoding” is used for each waveform.
• There is also a latent space within each class to vary.

Training data

- All waveforms sampled at 1024 Hz.

<table>
<thead>
<tr>
<th>Signal type</th>
<th>Duration</th>
<th>Frequency (Hz)</th>
<th>Decay (s)</th>
<th>Central time epoch (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sine-Gaussian</td>
<td>1s</td>
<td>70 - 250</td>
<td>0.004 - 0.03</td>
<td>0.4 - 0.6</td>
</tr>
<tr>
<td>White-noise burst</td>
<td>1s</td>
<td>70 - 250</td>
<td>0.004 - 0.03</td>
<td>0.4 - 0.6</td>
</tr>
<tr>
<td>Gaussian Pulse</td>
<td>1s</td>
<td>-</td>
<td>0.004 - 0.03</td>
<td>0.4 - 0.6</td>
</tr>
<tr>
<td>Ring-down</td>
<td>1s</td>
<td>70 - 250</td>
<td>0.004 - 0.03</td>
<td>0.4 - 0.6</td>
</tr>
<tr>
<td>BBH*</td>
<td>1s</td>
<td>-</td>
<td>-</td>
<td>0.4 - 0.6</td>
</tr>
</tbody>
</table>
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.

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 \([1, 0]\) \rightarrow \text{Ringdown} \([0, 1]\)

\[
\begin{align*}
[1, 0] & \rightarrow [0.8, 0.2] \rightarrow [0.6, 0.4] \rightarrow [0.4, 0.6] \rightarrow [0.2, 0.8] \rightarrow [0, 1]
\end{align*}
\]

Modelled signal ‘Unmodelled signal’ Modelled signal

Interpolating between classes

- We can consider random mixtures of classes.
- **Vertex** - Points that lie at the corners of the 5-dimensional class space, closest to training data.
- **Simplex** – Sampled points on a simplex, 5-dimensional hyperplane which links all vertices. It is a subspace of Uniform.
- **Uniform** - Sampling uniformly within the 5-dimensional one-hot encoding hyper-cube.

Interpolating between classes

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

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.