Gravitational Wave Data Analysis with Machine Learning

Jade Powell

OzGrav, Swinburne University of Technology.

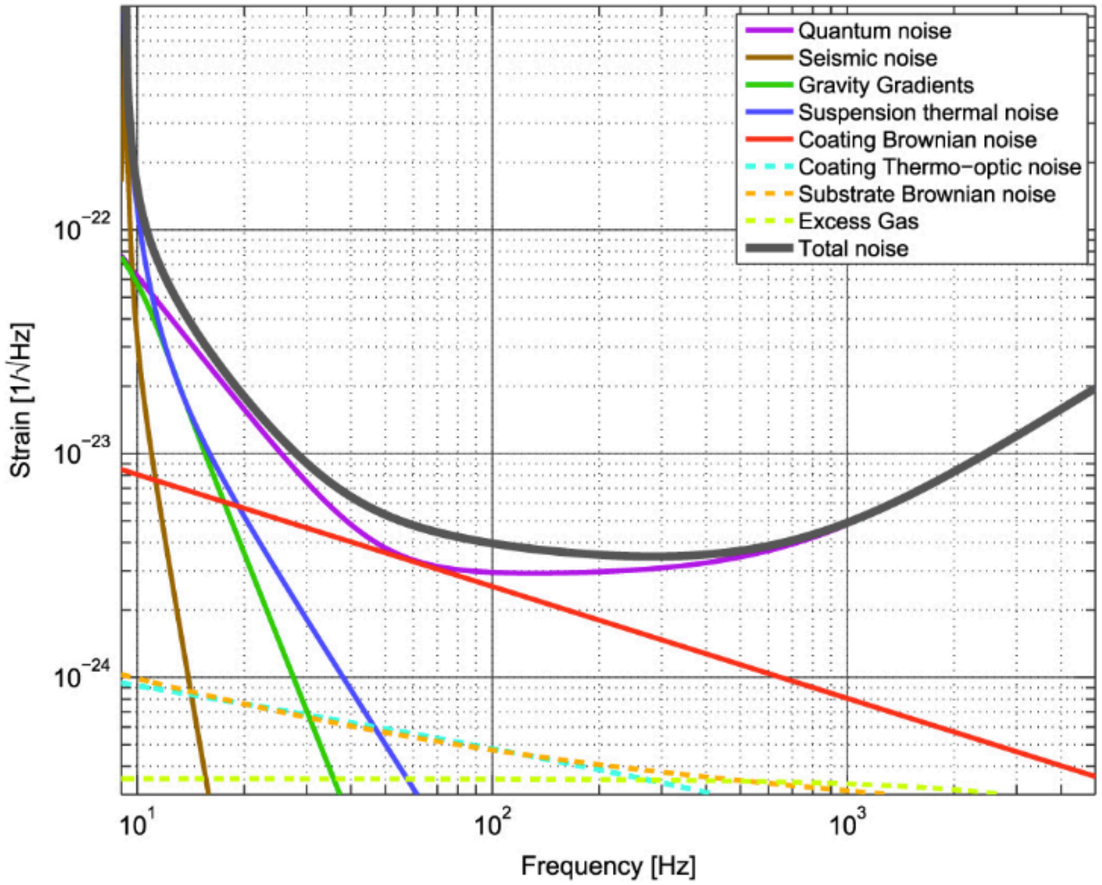
Outline

- I will give an overview of some problems in gravitational wave data analysis and how we are trying to solve them with machine learning.
- I will include improving data quality, searches for binary black holes and unmodelled gravitational wave bursts, and the astrophysics of gravitational wave sources.
- I do not include every study in these areas.

Gravitational Wave Data Quality

Fundamental Noise

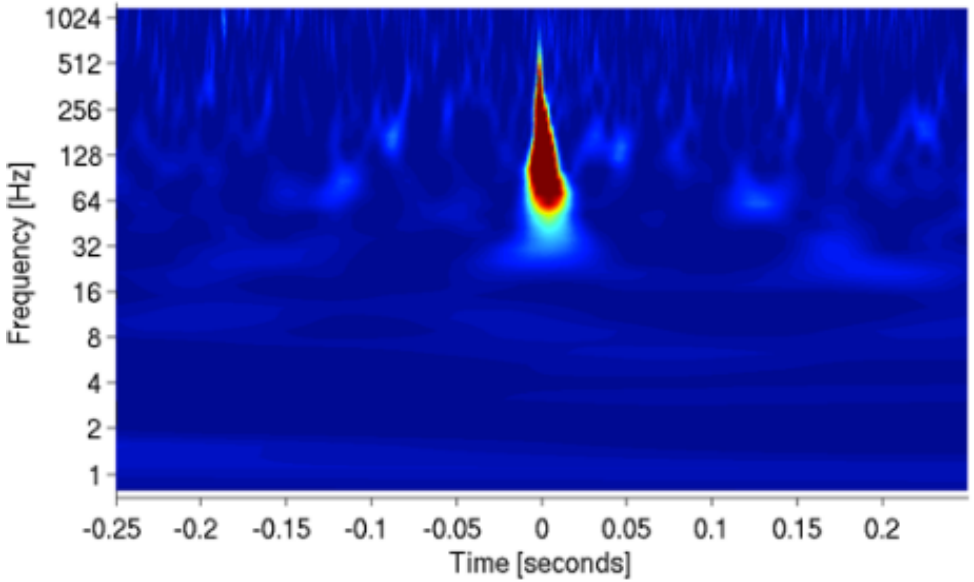
- The sensitivity of a GW detector is limited by fundamental, irreducible noise sources.
- However, the sensitivity of real-world instruments is often limited by noise sources of different origins, related to instrumental and environmental disturbances.
- Often these noise sources are non-stationary, i.e., their statistical properties vary over short or long time scales.

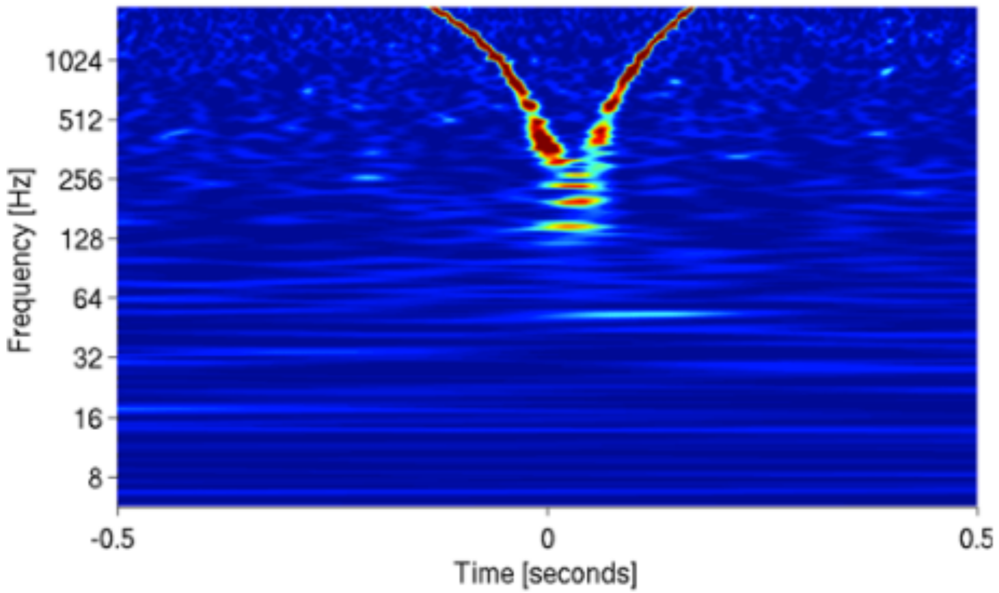


Credit: Advanced LIGO arXiv:1411.4547

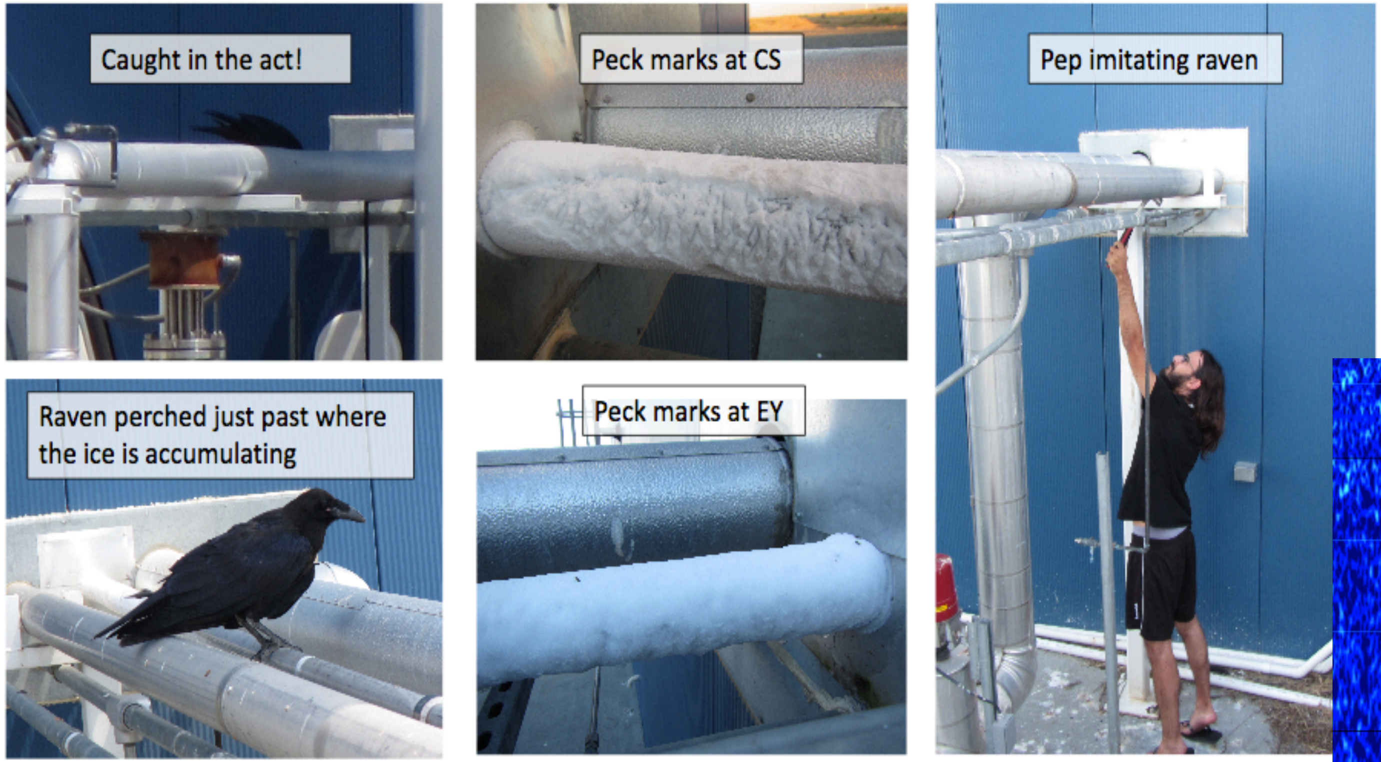
Non-Gaussian Noise

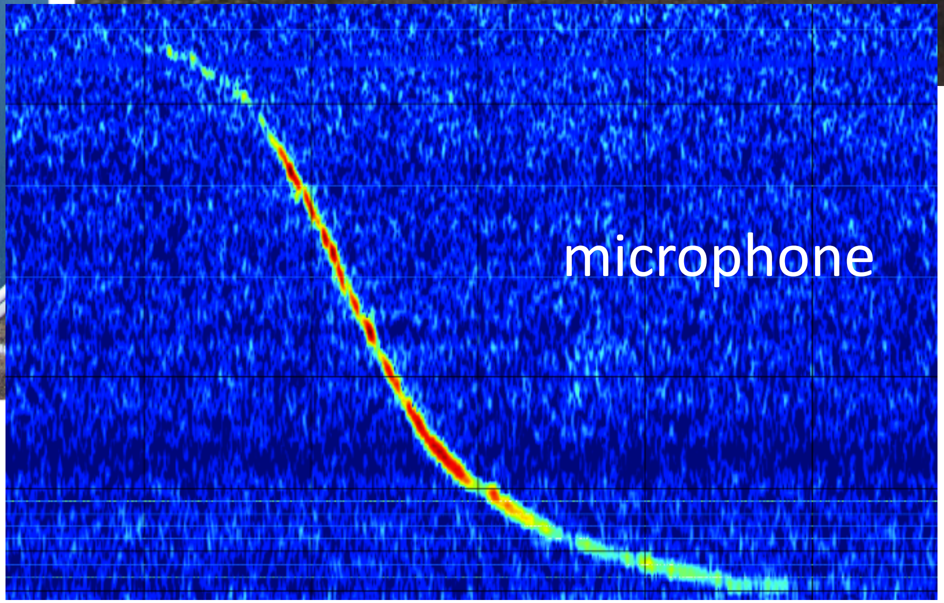
- Glitches are noise transients that produce excess power in the gravitational wave channel.
- They are characterized by their SNR, frequency, and morphology.
- Thousands of glitches can occur everyday.
- They limit the sensitivity of GW searches and can contaminate signals.





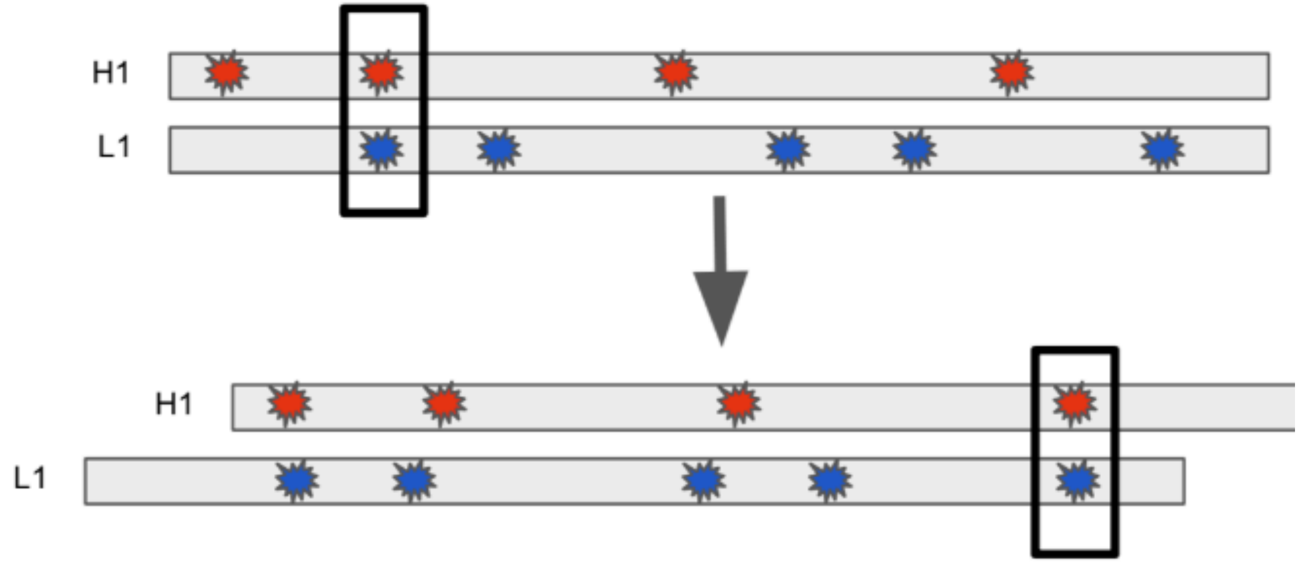
Glitches





Credit: LIGO Scientific Collaboration

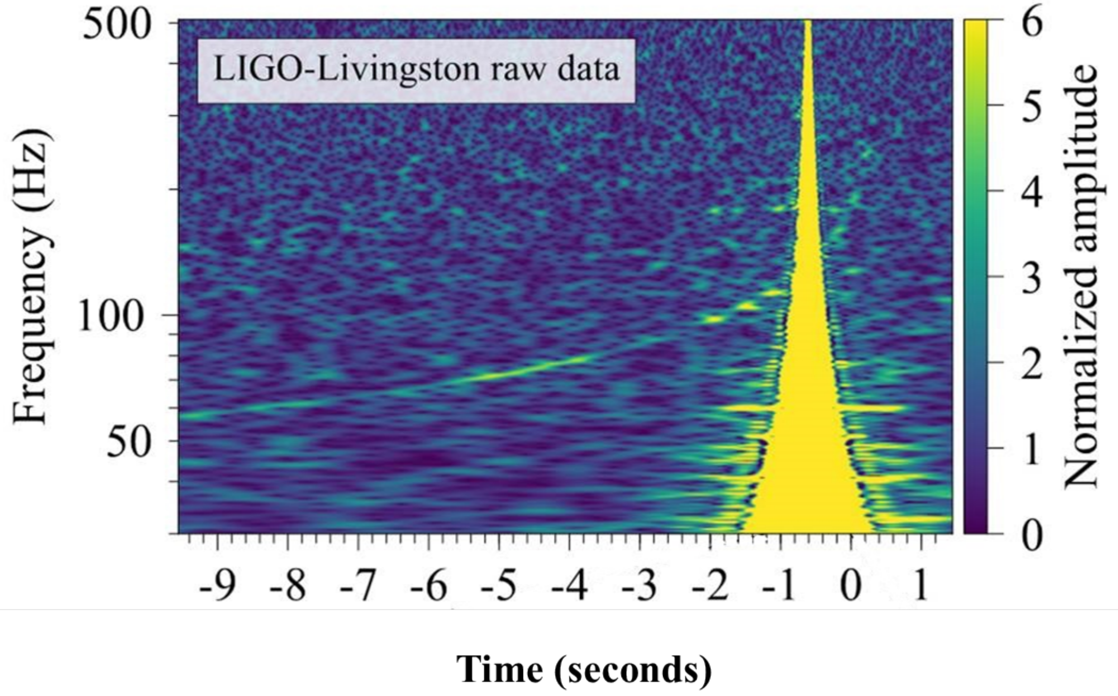
How do glitches limit search sensitivity?



Timeslides for measuring the GW background.

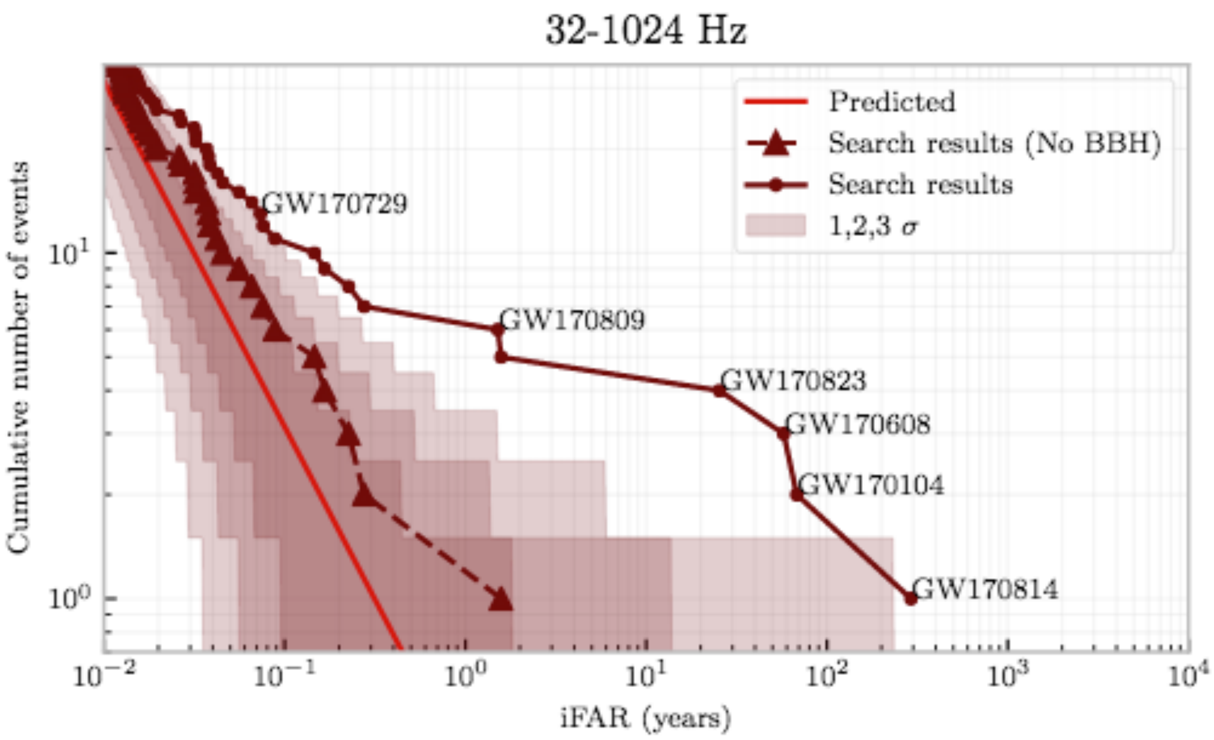
Image Credit: Derek Davis

Image Credit: LIGO Virgo arXiv:1710.05832

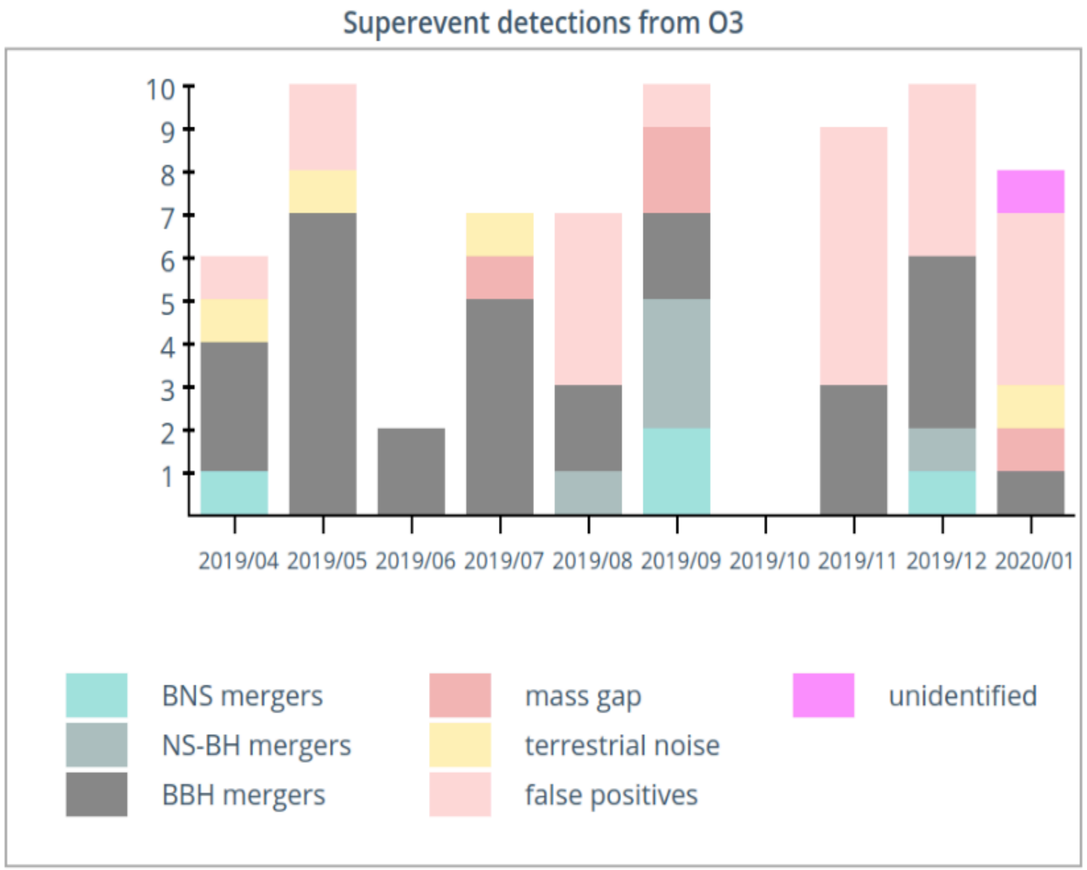


Glitches even occur during signals.

How do glitches limit search sensitivity?



O2 cWB results arXiv:1905.03457

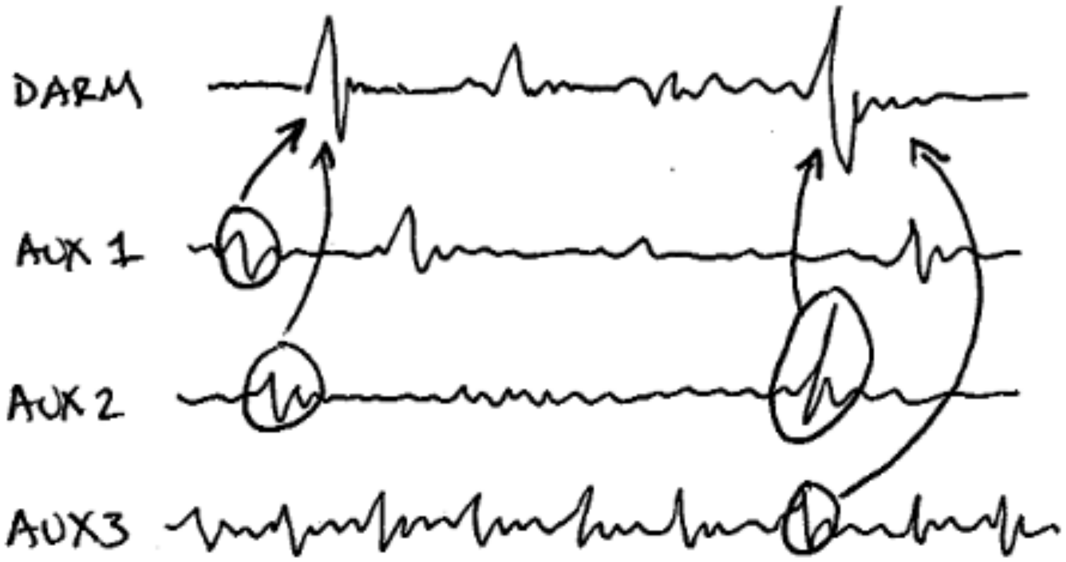


Auxiliary channels

- There are over 200,000 channels which monitor instrument behaviour and environmental conditions.
- These channels witness a broad spectrum of potential coupling mechanisms.
- We look for correlations between data in the gravitational wave channel and these auxiliary channels to identify when glitches occur.
- This can be used to “veto” a gravitational wave detection candidate.

iDQ

- iDQ produces low-latency probabilities of a gravitational wave candidate being a glitch.
- It uses glitches in auxiliary channels to identify glitches in the GW data channel.
- It provides statistical inferences about the probability that a glitch exists within the GW channel, by ingesting triggers from an trigger generator and passing to machine learning algorithms.
- However, not all glitches occur in auxiliary channels.



Credit: Reed Essick

Glitch Classification

- Some glitches occur only in the GW data channel.
- We can try and eliminate them by classifying them into different types to help identify their origin.

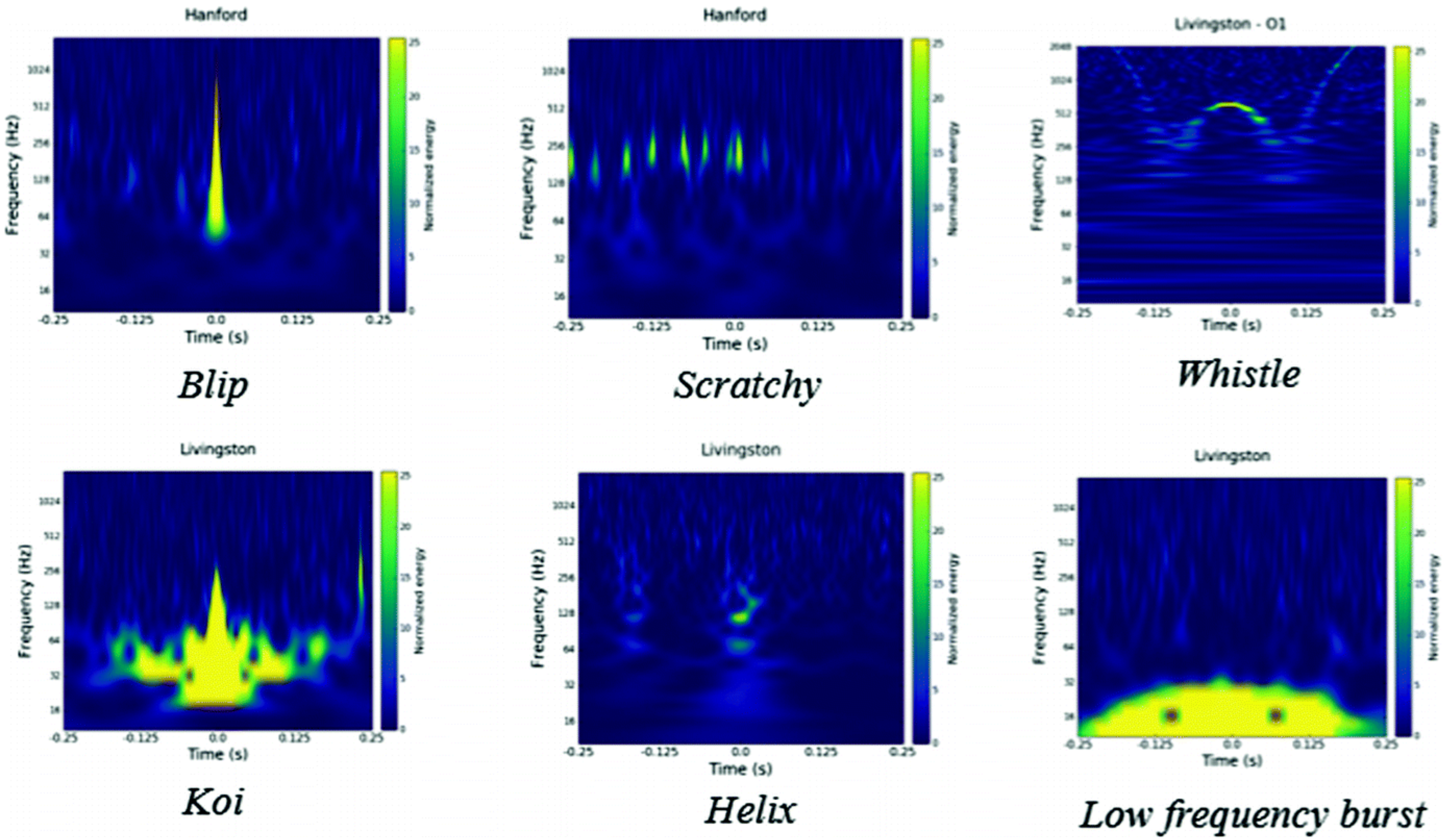
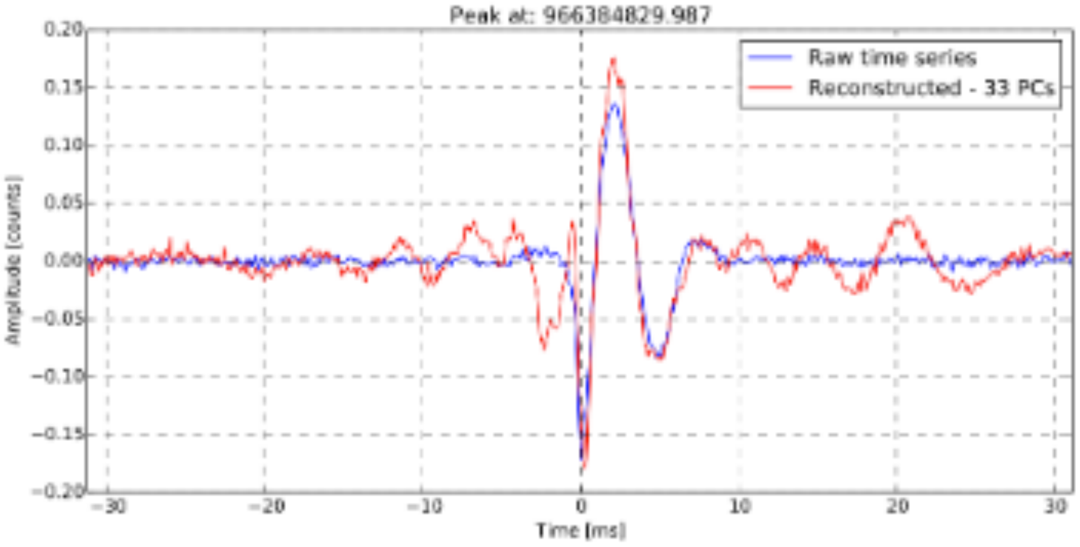


Image Credit: Gravity Spy

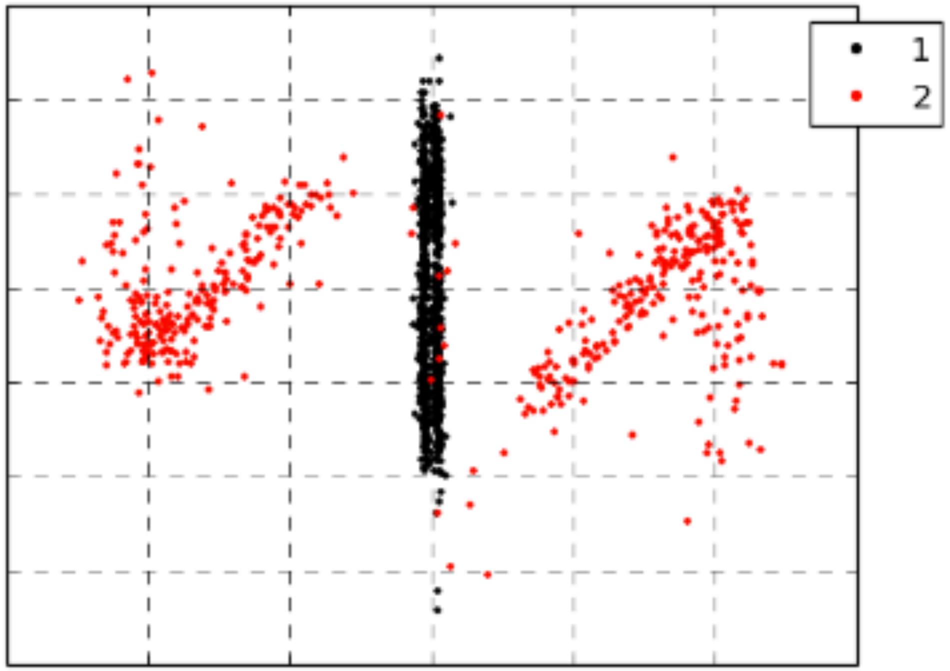
Glitch Classification with PCA

- Early ML studies for glitch classification used Principal Component Analysis (PCA) and Gaussian Mixture Models (GMM).
- (See Powell et al. arXiv:1609.06262 & arXiv:1505.01299)
- A trigger generator finds the glitches.
- The time series of whitened glitches are stored in a matrix D on which PCA is performed.



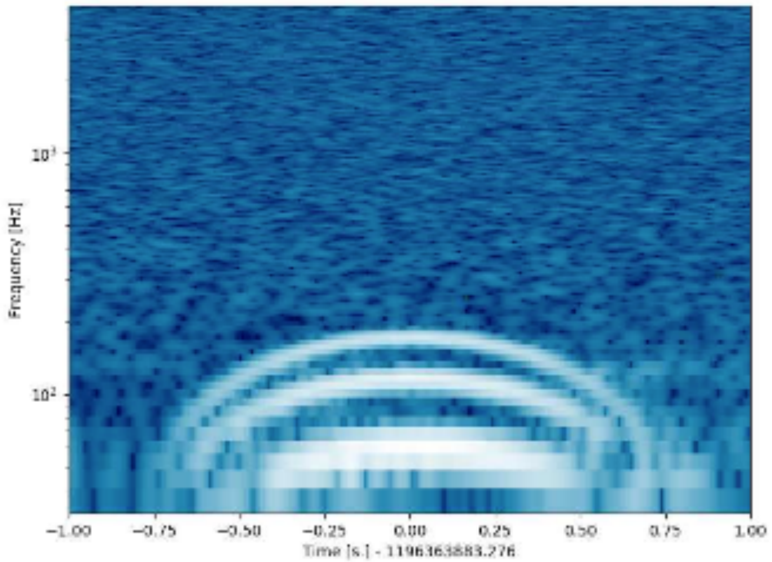
Glitch Classification with PCA

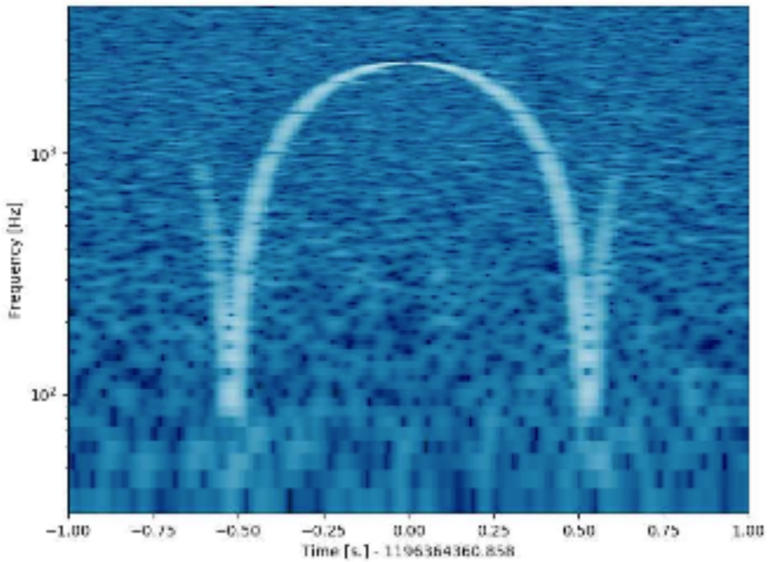
- PCA is an orthogonal linear transformation that transforms a set of correlated variables into another set of linearly uncorrelated variables, called Principal Components (PCs).
- The matrix D is factored so that D= $U\Sigma V^T$ where V=A^TA, Σ contains eigenvalues, and U is the PCs.
- PC coefficients are calculated by taking the dot product of the PCs and the whitened glitch.
- Then GMM clustering is applied to the PC coefficients.



Glitch Classification with NN

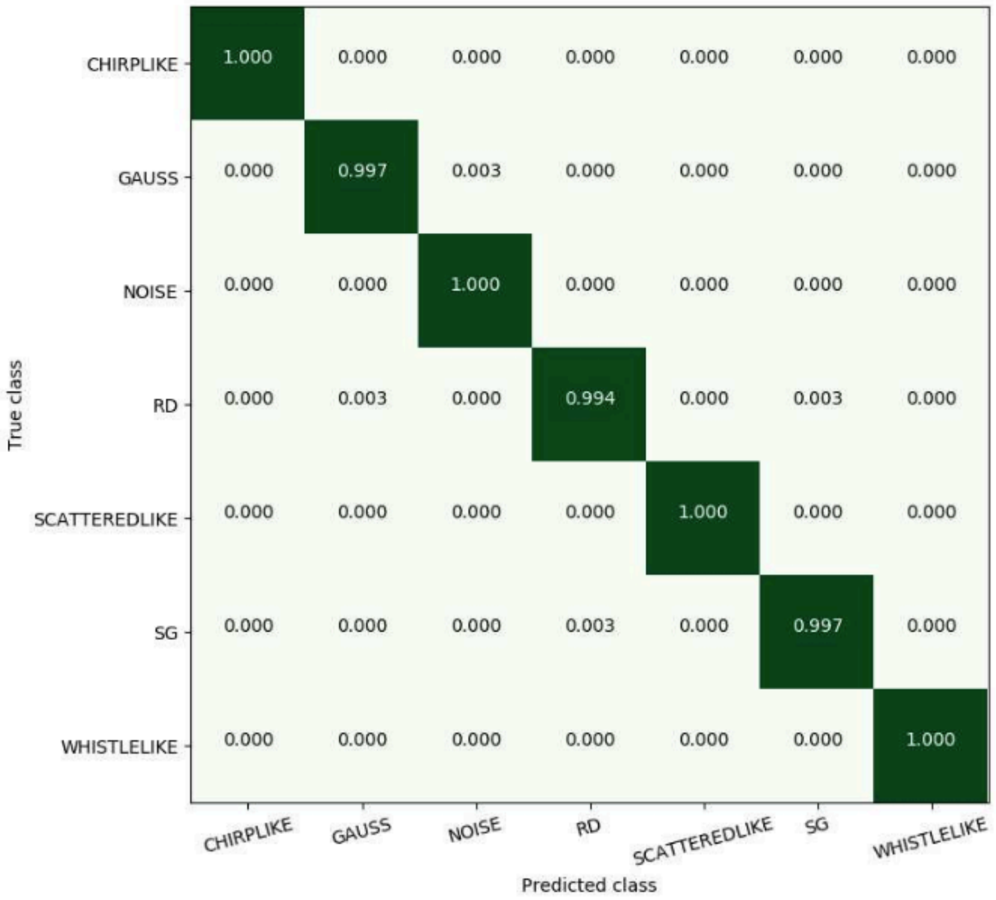
- These studies were then improved with the use of Neural Networks.
- Razzano & Cuoco 2018 (arxiv:1803.09933) apply a CNN to simulated glitches.





Glitch Classification with NN

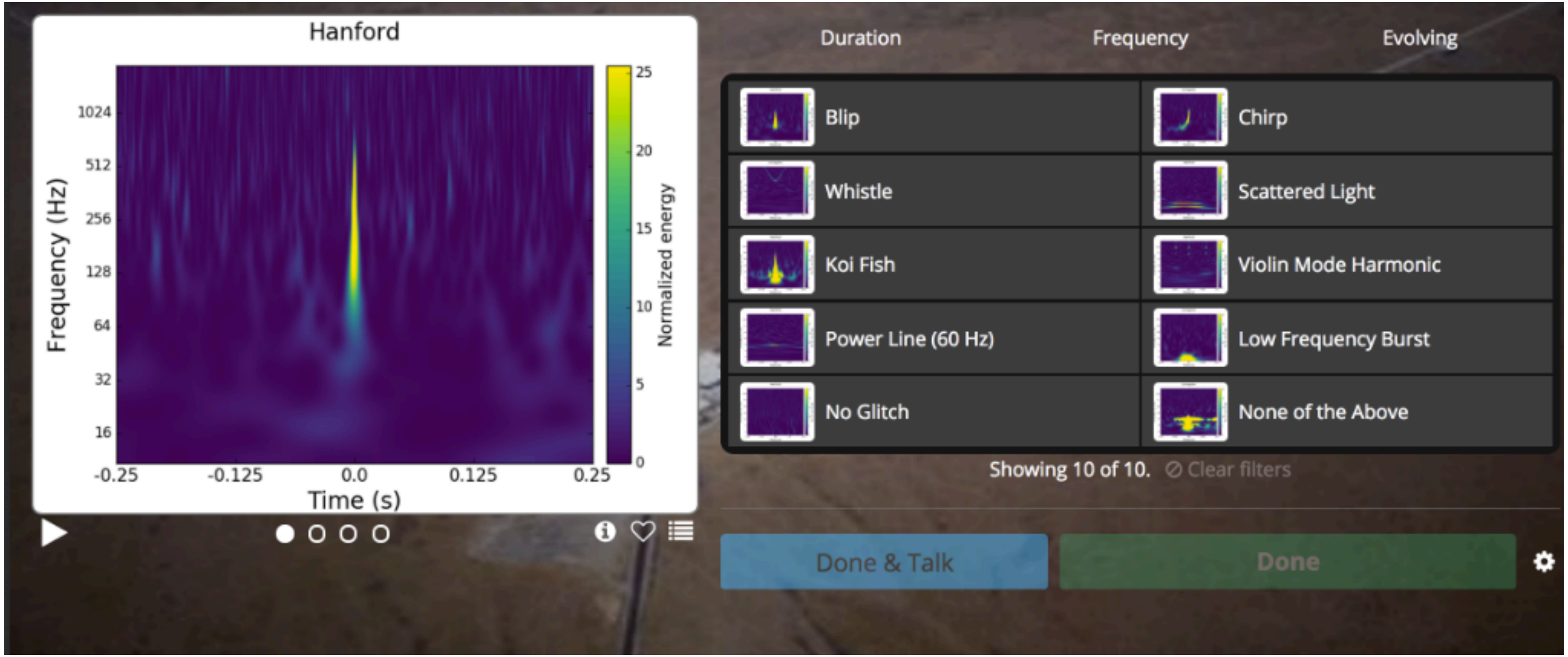
- They build images that cover 2 seconds around each glitch from the whitened time series.
- Simulated six families of signals.
- Training, validation, and test set with ratio 70:15:15.



Gravity Spy

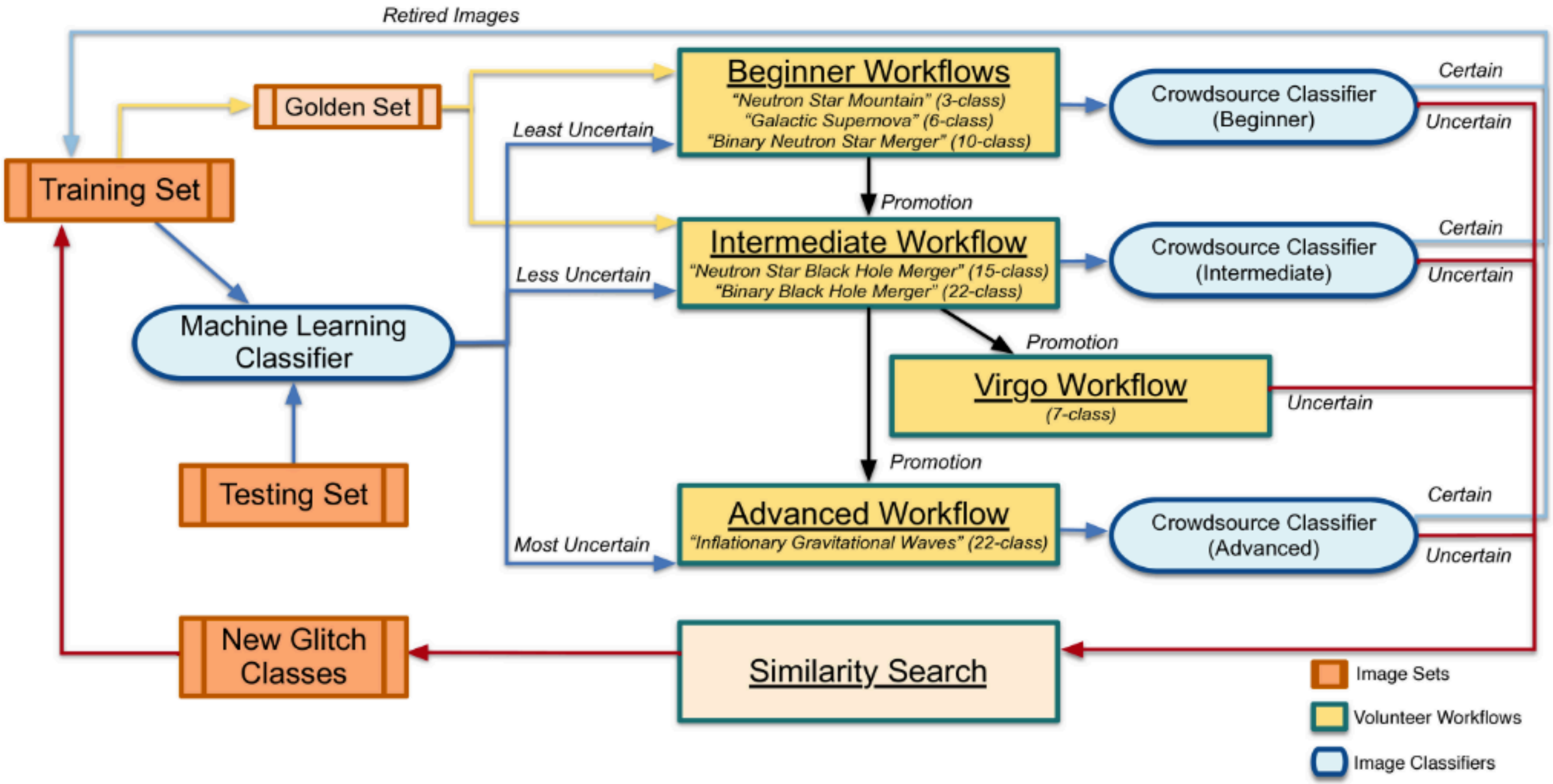
gravityspy.org

- GravitySpy uses citizen scientists to produce training sets for machine learning glitch classification.



Gravity Spy

gravityspy.org



Coughlin *et al* arXiv:1903.04058

How do I try it myself?

- Log into gravityspy.org to try classifying glitches
- Download already labelled LIGO glitches for training your algorithm from zenodo
- https://zenodo.org/record/1476156
- https://zenodo.org/record/1476551

How do I get more LIGO & Virgo data?

- The Gravitational Wave Open Science Center provides data from gravitational-wave observatories, along with access to tutorials and software tools.
- You can download the data from previous observing runs.
- https://www.gw-openscience.org/about/

Compact Binaries Mergers (CBC)

Waveform Modelling

- Signal models are needed for matched filtering and parameter estimation.
- Solutions of the Einstein equations can be obtained with numerical relativity simulations - High computational cost!
- LIGO and Virgo rely on approximate solutions obtained through phenomenological modelling
- Gaussian process regression has been used to produce new waveforms by providing a direct interpolation between numerical simulations.
- For Example, Z. Doctor arXiv:1706.05408 (2017)

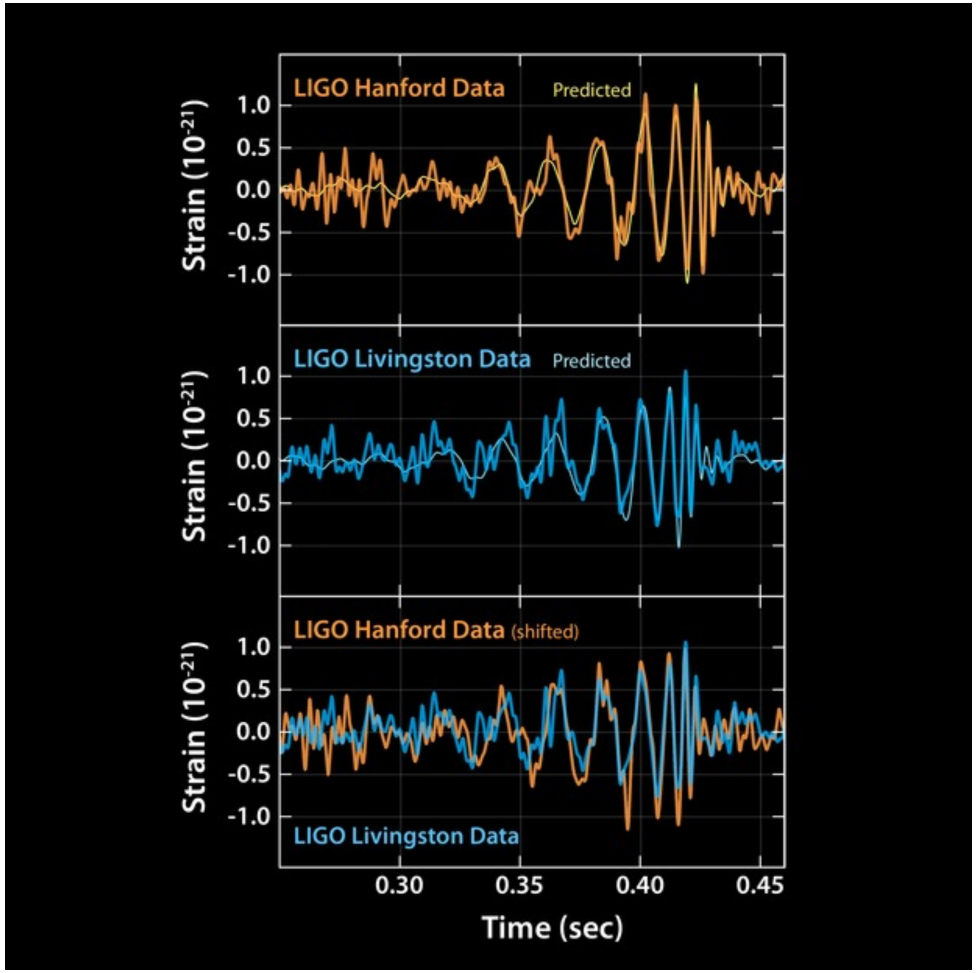


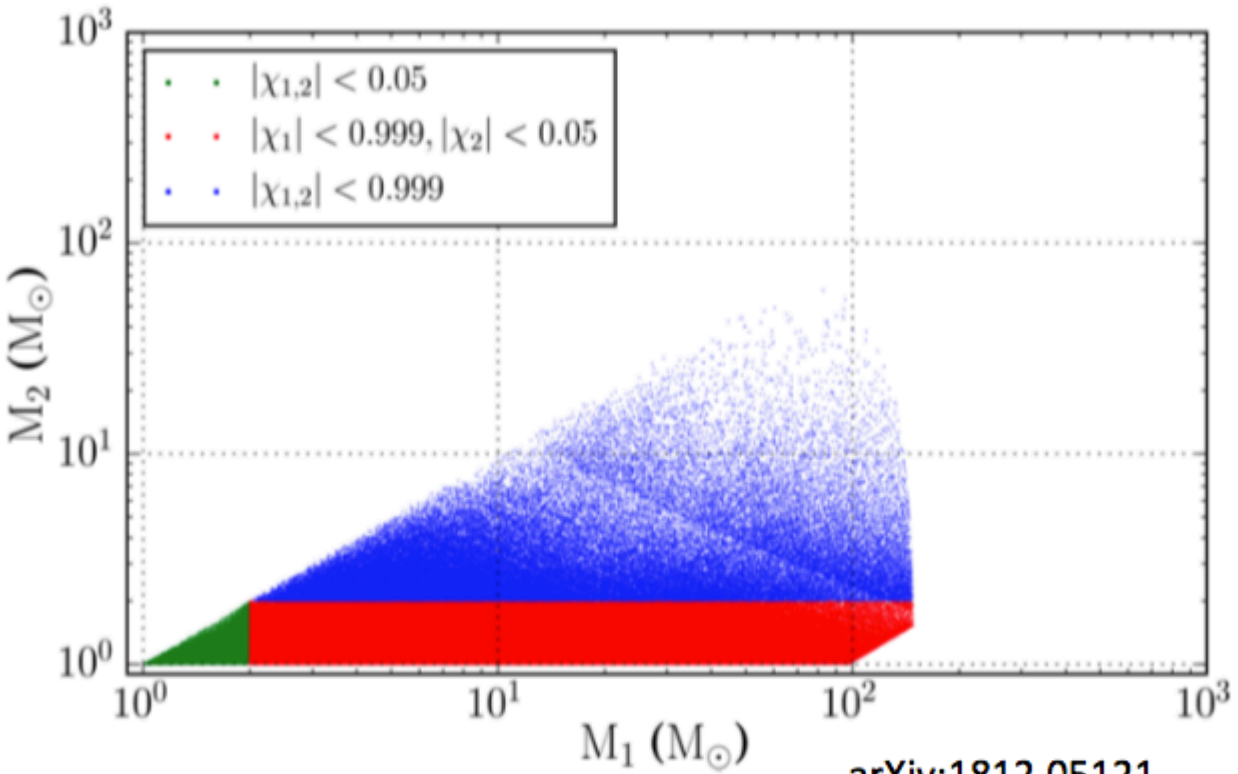
Image Credit: LIGO and Virgo

Matched filter searches

- Searches for gravitational wave signals from compact binaries use matched filtering.
- GW detector noise is non-Stationary and non-Gaussian.
- Inner product: $(\mathbf{a}|\mathbf{b}) = 2\int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f) + \tilde{a}^*(f)\tilde{b}(f)}{S_n(f)}\,df$
- Matched filter: $(\mathbf{d}|\mathbf{h}(\boldsymbol{\theta}))$
  - **d** – data
  - **h** – waveform model
  - $\boldsymbol{\theta}$ - binary parameters

Matched filter searches

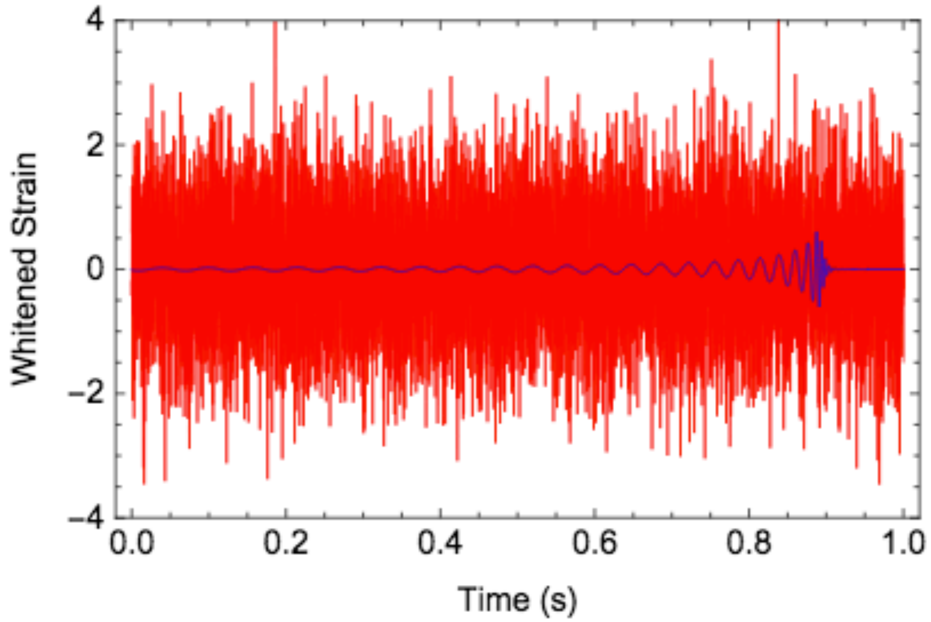
- Discrete template bank is built to cover the mass-spin parameter space for potential sources.
- Density of the bank is determined by the minimum overlap requirement of 0.97 so only 3% of SNR is lost.
- Assume spins are aligned to the orbital angular momentum.
- Do not account for tidal deformability of neutron stars.

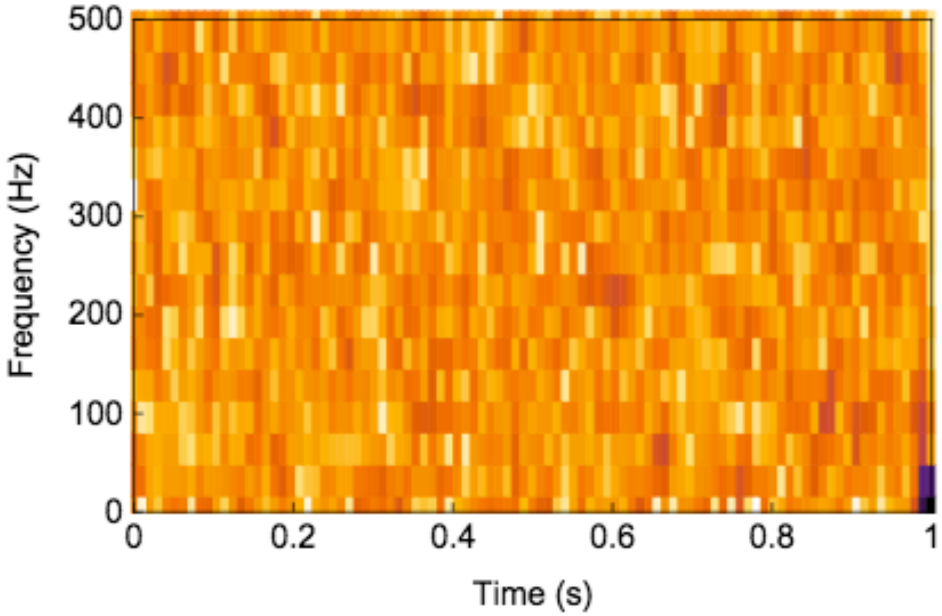


arXiv:1812.05121

Machine learning CBC searches

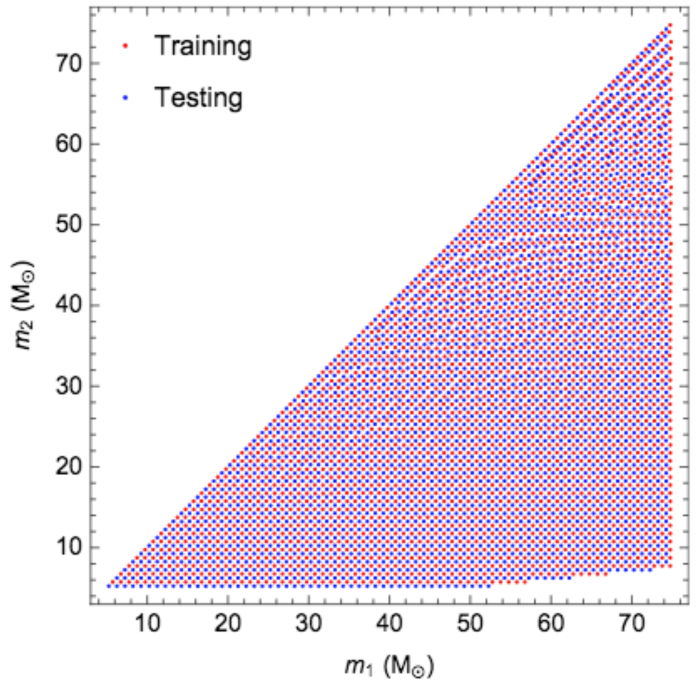
- George & Huerta (arXiv:1701.00008) use a system of two deep convolutional neural networks to rapidly detect CBC signals.
- They use time series as input so that they can find signals too small for image recognition.

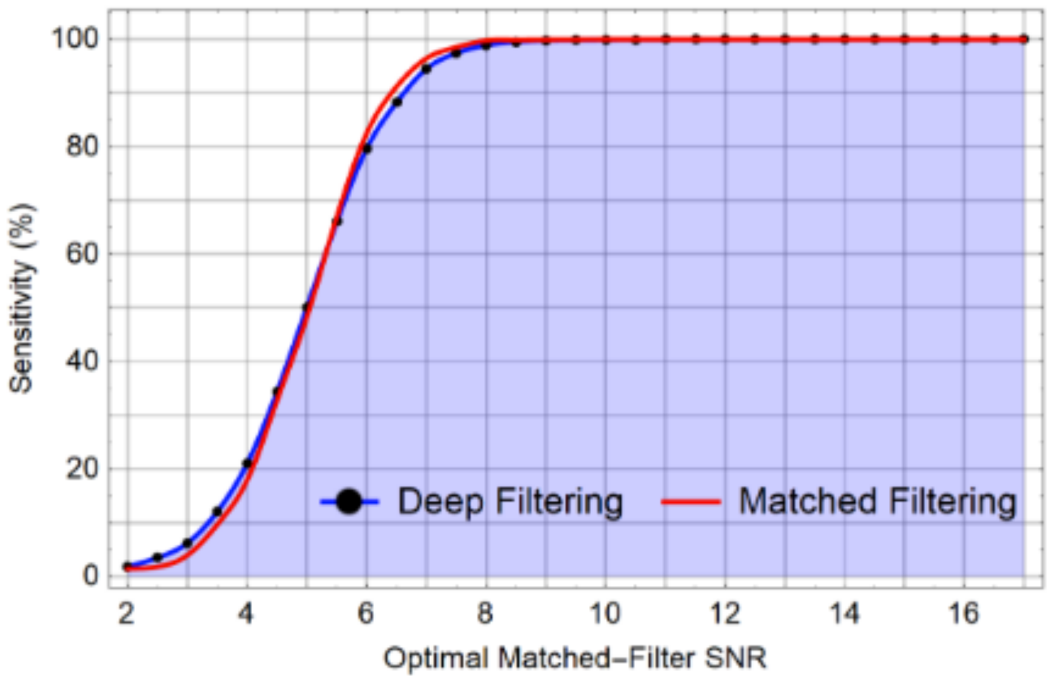




Machine learning CBC searches

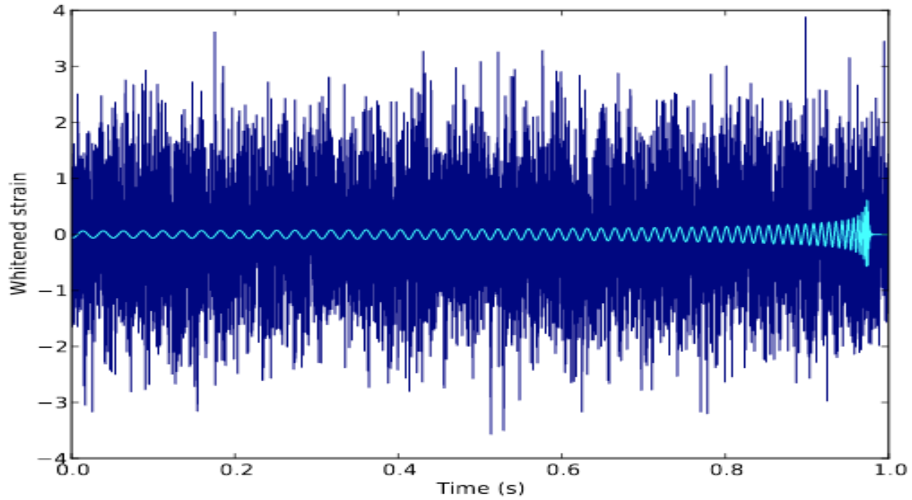
- George & Huerta (arXiv:1701.00008)
- They find their method significantly outperforms conventional machine learning techniques, achieves similar performance compared to matched-filtering while being several orders of magnitude faster.

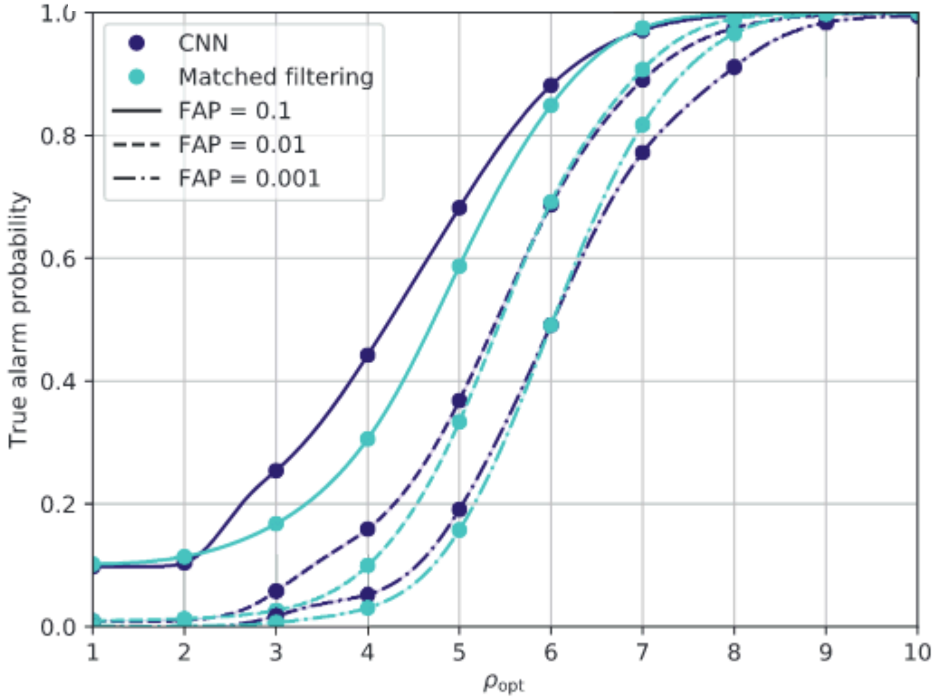




Machine learning CBC searches

- Gabbard et al. (arXiv:1712.06041) also perform a CBC search with a deep convolutional neural network.
- They use whitened time series of measured gravitational-wave strain as an input.
- Train and test on simulated binary black hole signals in synthetic Gaussian LIGO noise.
- They find they can reproduce the sensitivity of a matched filter search.



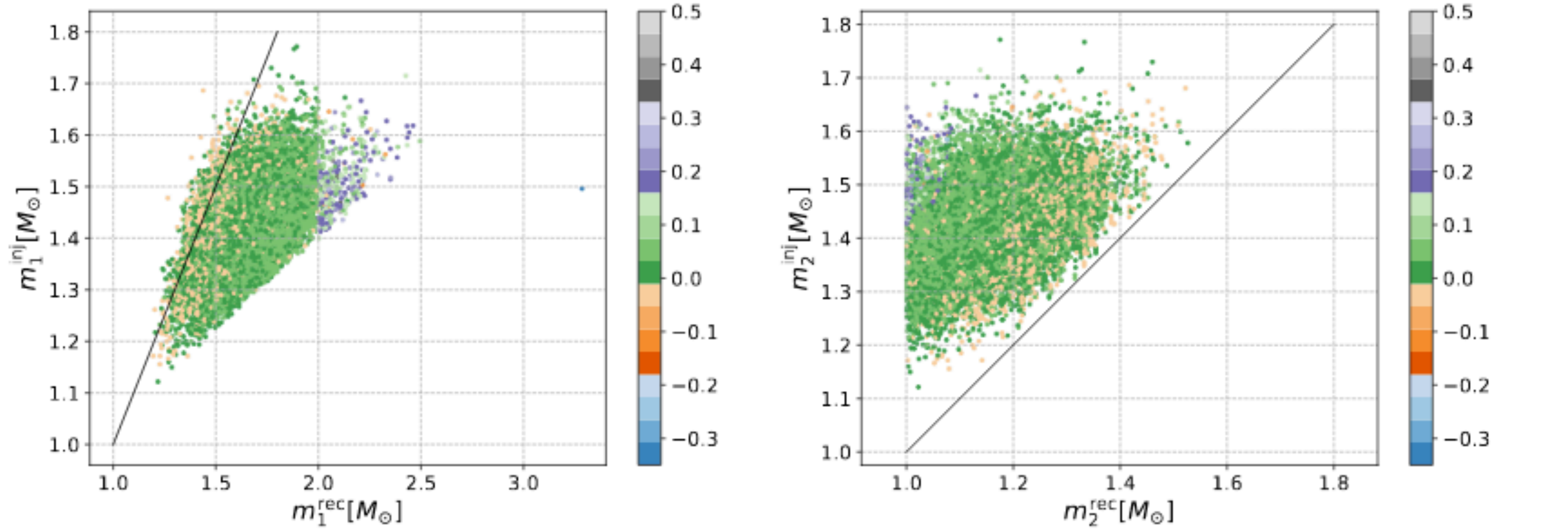


Low-latency source-properties (EM-bright)

- LIGO & Virgo provide two probabilities in low-latency. (See arXiv:1911.00116)
- The probability that there is a neutron star in the CBC system, P(HasNS)
- The probability that there exists tidally disrupted matter outside the final coalesced object after the merger, P(HasRemnant).
- Matched filter searches give point estimates of mass and spin but they have large errors!
- To solve this a machine learning classification is used. (scikit learn K nearest neighbours, also tried random forest)

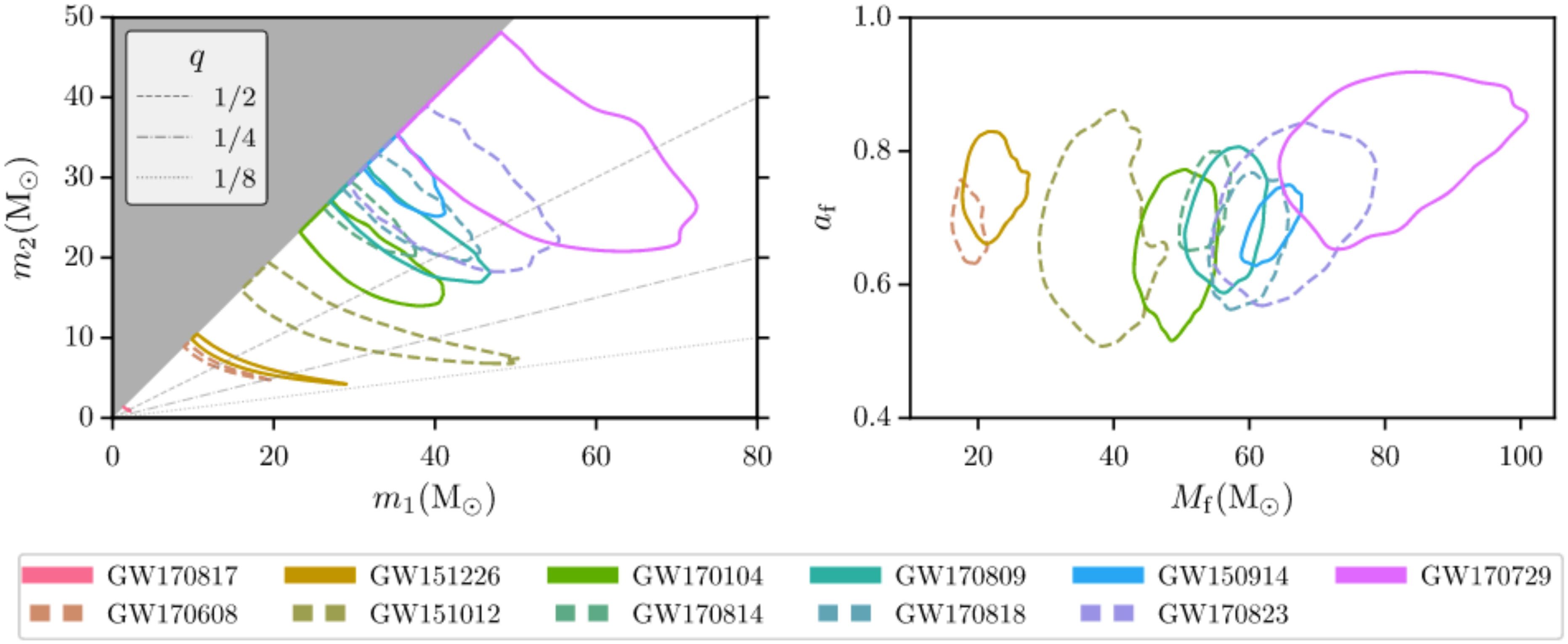
Low-latency source-properties (EM-bright)

- A training set is created by injecting fake signals into gravitational wave data and performing a search.
- This then produces a map between true values and matched filter search point estimates which is learnt by the classifier.



CBC Parameter Estimation

- Characterized by 15 parameters. Masses, spins, distance, inclination, sky position, polarization.



Credit: LIGO & Virgo GWTC-1 arXiv:1811.12907

CBC Parameter Estimation

Bayes' Theorem:

$$\underbrace{p(\theta|d,M)}_{\text{posterior}} = \frac{\overbrace{p(\theta|M)}^{\text{prior}}\overbrace{p(d|\theta,M)}^{\text{likelihood}}}{\underbrace{p(d|M)}_{\text{evidence}}}$$

Notation:

$p(x)$ = Probability distribution of x

$p(x|y)$ = Conditional probability of x given y

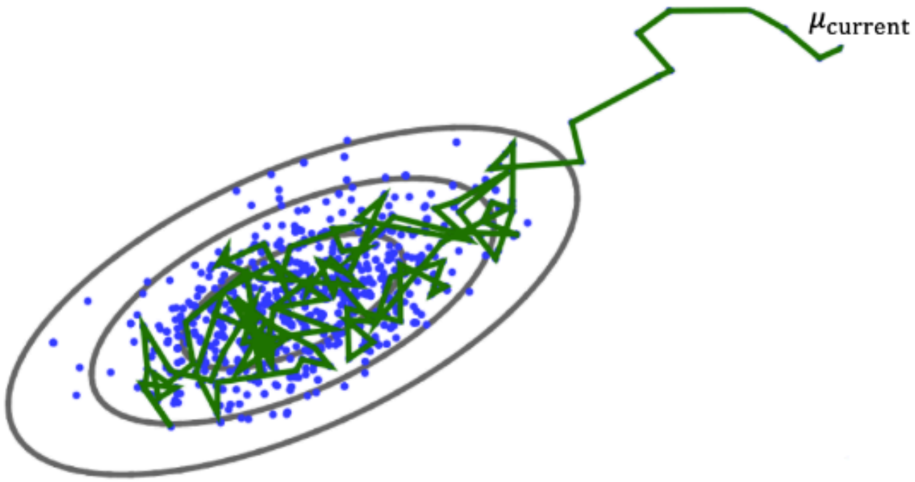
Data: d = noise + GW strain ,

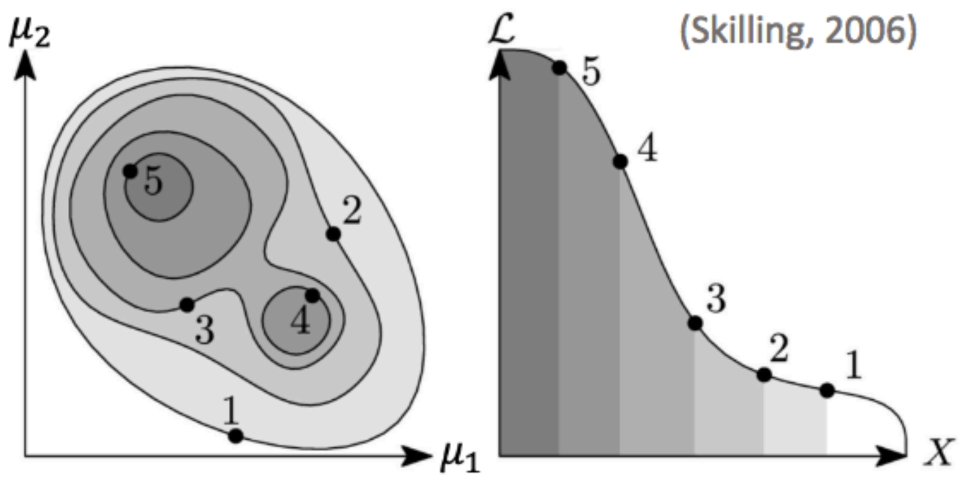
Model: $M(f)$,

Model parameters: $\theta = \{m_1, m_2, d_L, ..\}$

MCMC and Nested Sampling

- We have two main PE codes LALInference and Bilby
- MCMC Random steps are taken in parameter space, according to a proposal distribution, and accepted or rejected according to the Metropolis-Hastings algorithm.
- Nested sampling can also compute evidences for model selection.





Machine Learning Parameter Estimation

- BAMBI: blind accelerated multimodal Bayesian inference combines the benefits of nested sampling and artificial neural networks. (arXiv: 1110.2997)
- An artificial neural network learns the likelihood function to increase significantly the speed of the analysis.

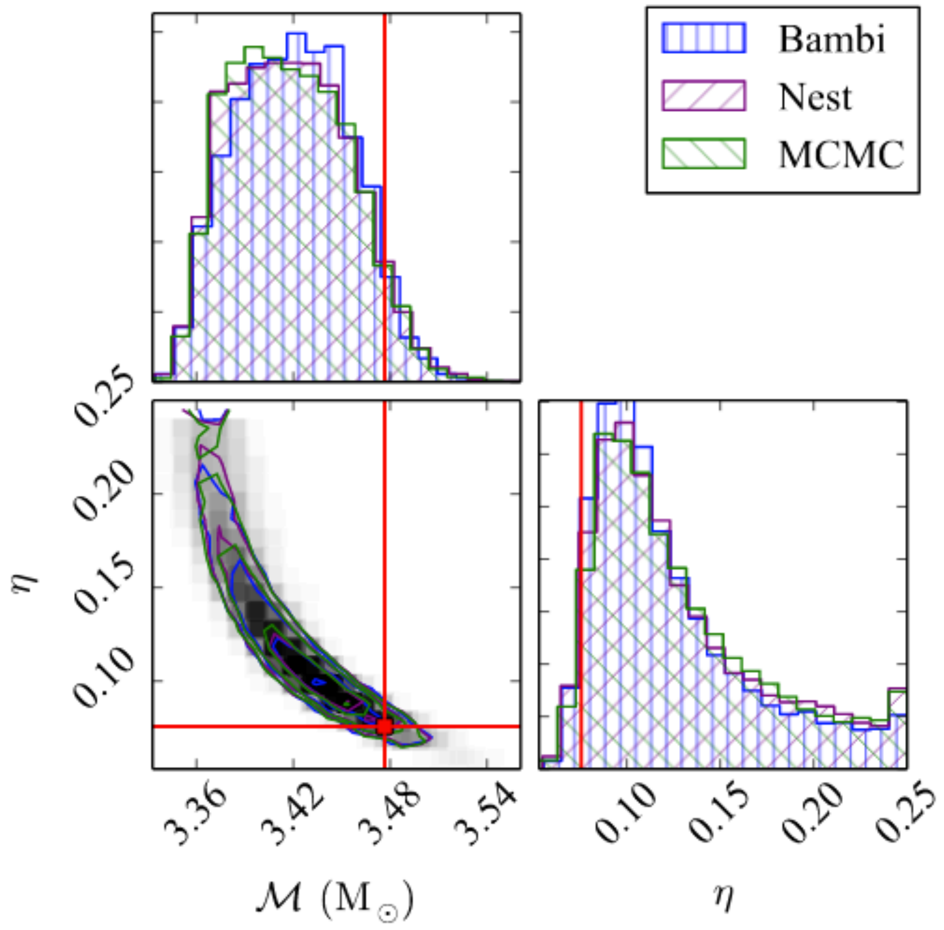
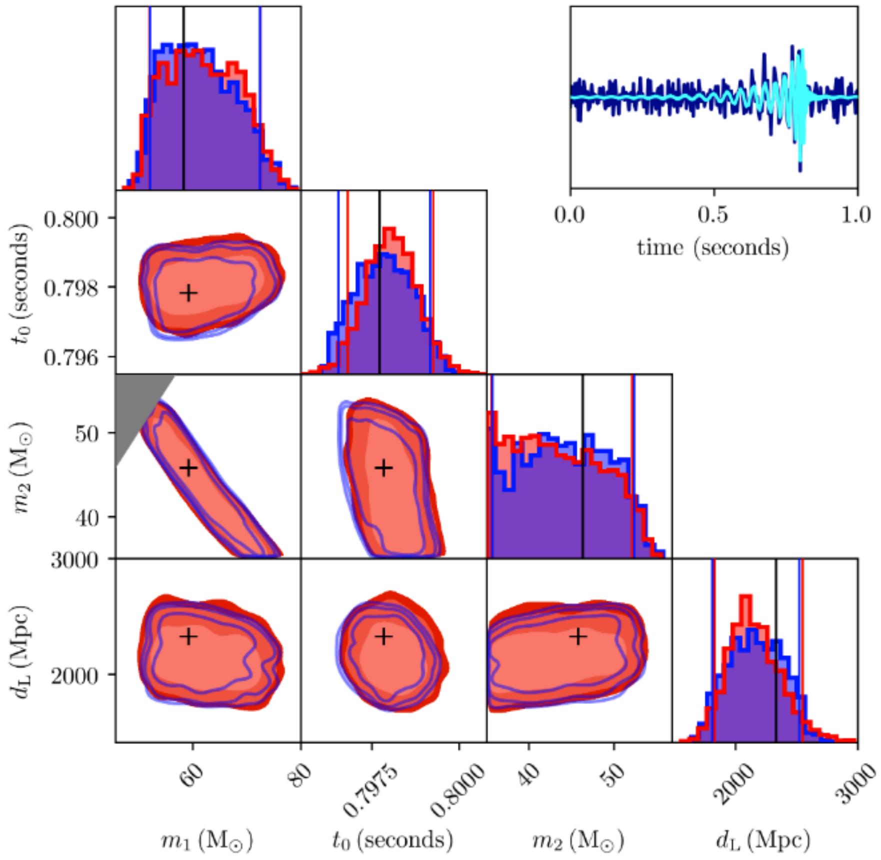


Image Credit: Vietch et al. arXiv: 1409.7215

Machine Learning Parameter Estimation

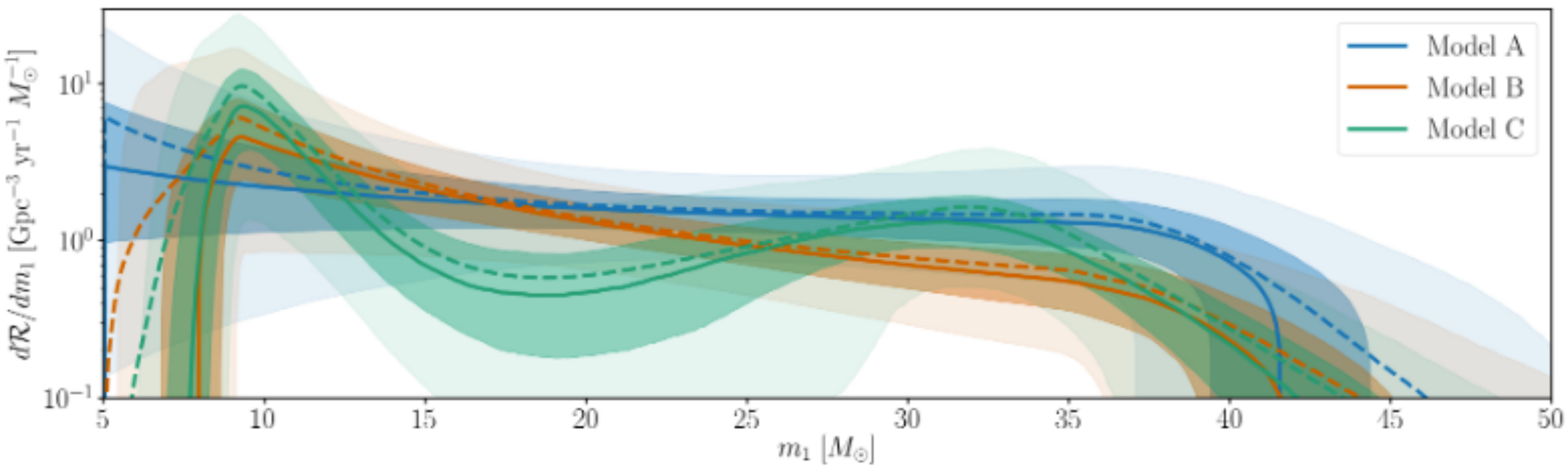
- Chua et al. (arXiv:1909.05966) produce Bayesian posteriors using neural networks.
- Gabbard et al. (arXiv:1909.06296) use a conditional variational autoencoder pre-trained on binary black hole signals.
- ~6 orders of magnitude faster than existing techniques.



Red is Gabbard et al. method and blue is Bilby nested sampling.

CBC Population Studies

- Now that we have started to detect a population of black hole signals we can try to do population studies to try and understand signals formation mechanisms.
- Population properties paper from O1+02 (arXiv:1811.12940).
- Uses phenomenological models (like power laws) combined with Bayesian hierarchical modelling.



CBC Population Studies

- Bayesian hierarchical modelling involves some assumptions of populations mass and spin distributions.
- Does not scale well for high dimensional models and a large number of GW detections.
- We can use unmodelling clustering!

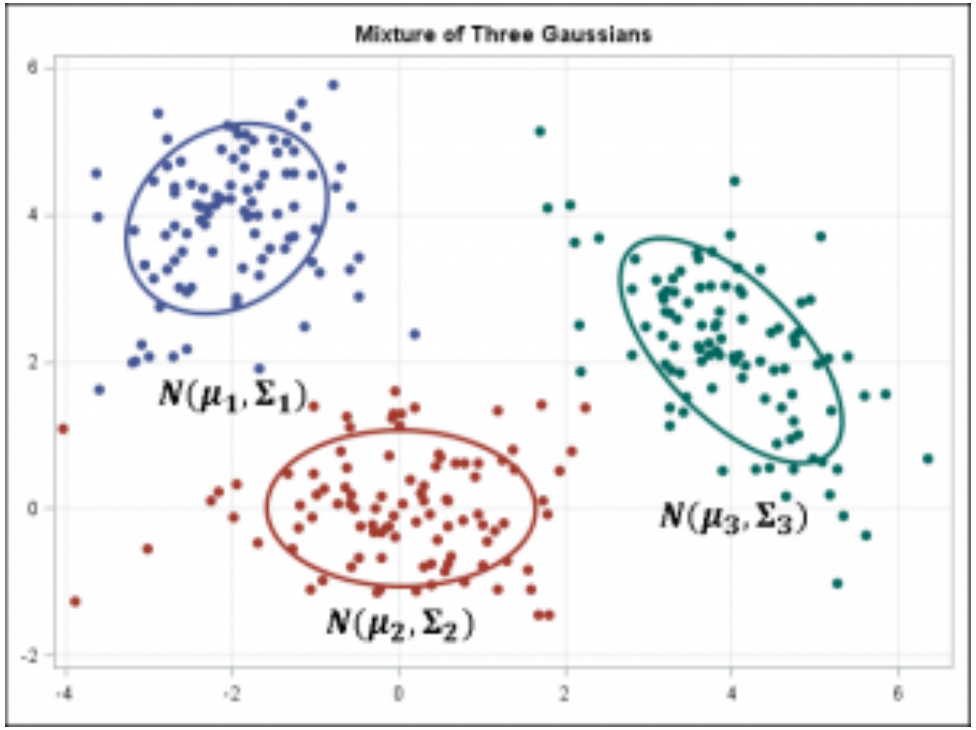
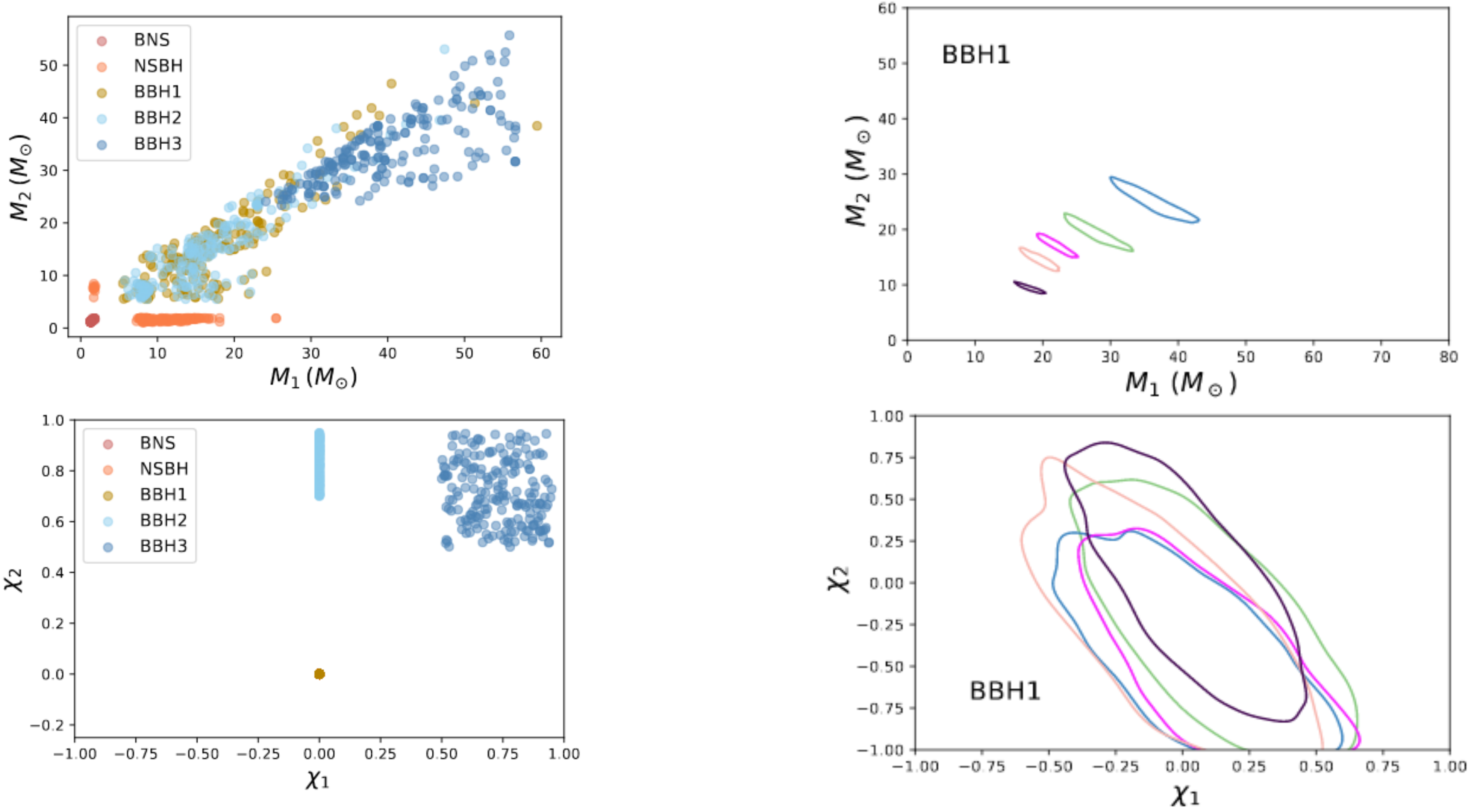


Image Credit: Alison Bolen, The SAS Data Science Blog

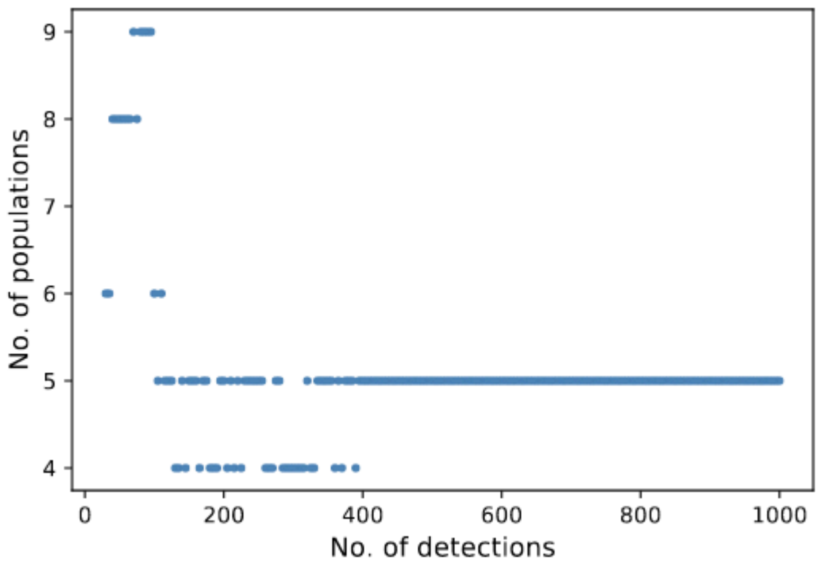
CBC Population Studies

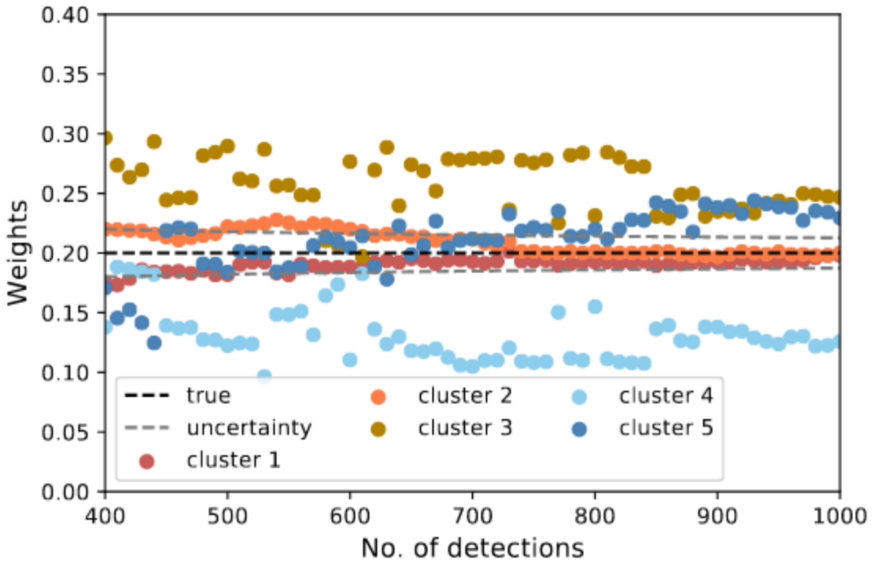
- In Powell et al. (arXiv:1905.04825) we apply unmodelled clustering to masses and spins.



CBC population studies

- Two of the populations have identical mass distributions and different spin.
- This is difficult because spin is poorly measured.
- Determine the number of populations and the number of CBC signals in each population.





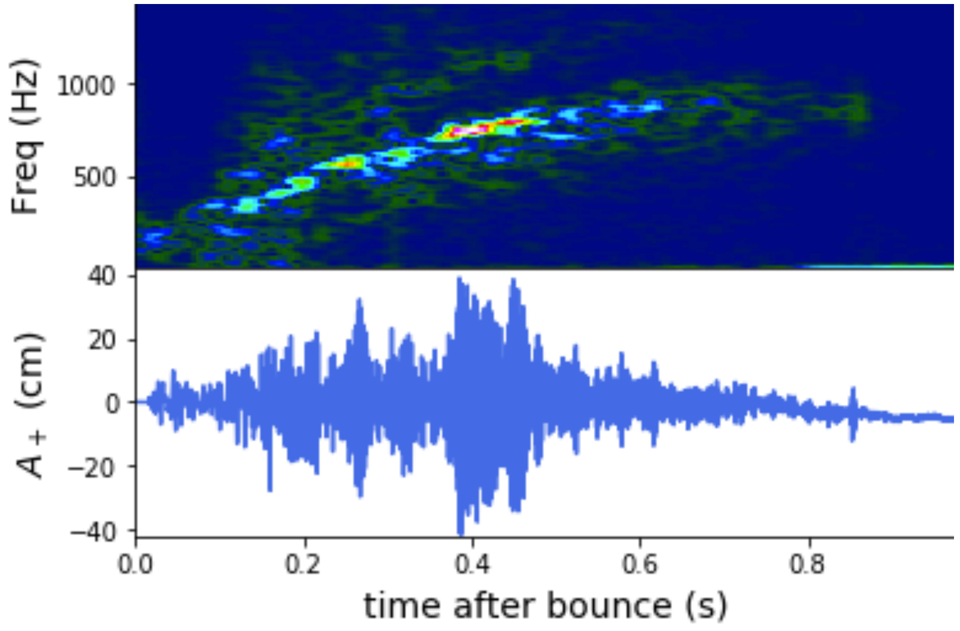
How do I try it myself?

- The Gravitational Wave Open Science Center has the data, parameter estimates, and matched filtering tutorials that you can download.
- You can get code to produce synthetic parameter estimates for compact binaries https://git.ligo.org/daniel.wysocki/synthetic-PE-posteriors

Gravitational Wave Bursts

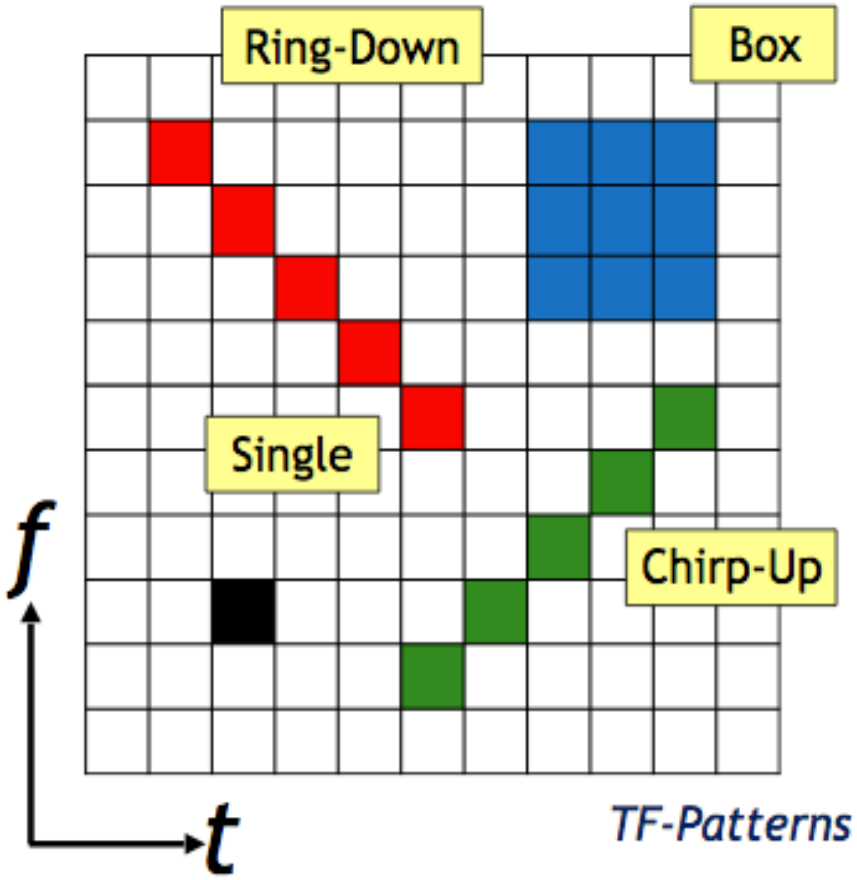
What are gravitational wave bursts?

- A burst is a gravitational wave signal where the waveform morphology is partially or completely unknown.
- The source could be an unknown unknown, a supernova, cosmic string, fast radio burst, compact binaries and others.
- The main burst search is called coherent Wave Burst (cWB).



Coherent Wave Burst (cWB)

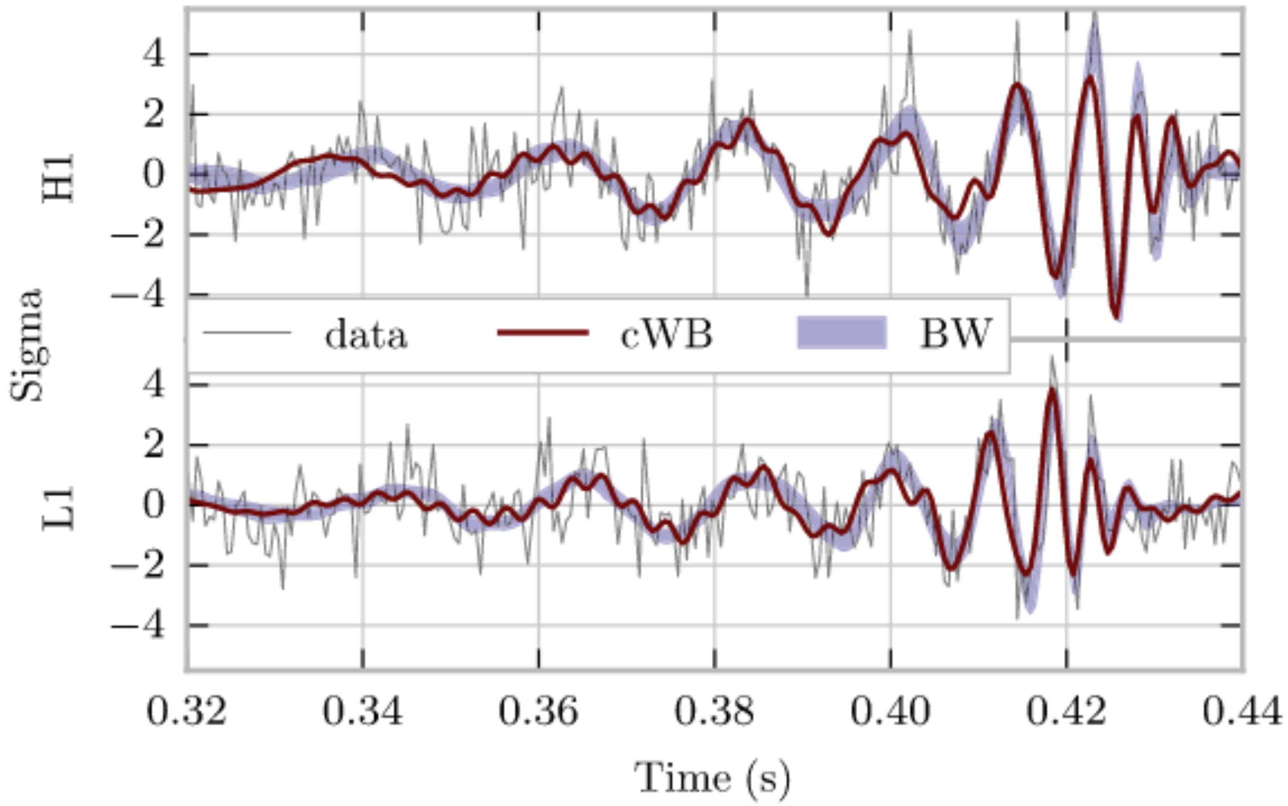
- cWB relies upon the excess coherent power in a network of detectors.
- The data is transformed into time-frequency domain and the clusters of time-frequency pixels above certain energy threshold are identified for each detector.
- Time frequency map of the single detectors is then combined using the maximisation of the likelihood over all possible sky locations and the events are then ranked according to this likelihood.
- We can also inform our un-modelled search about the morphology of the expected signal



Coherent Wave Burst (cWB)

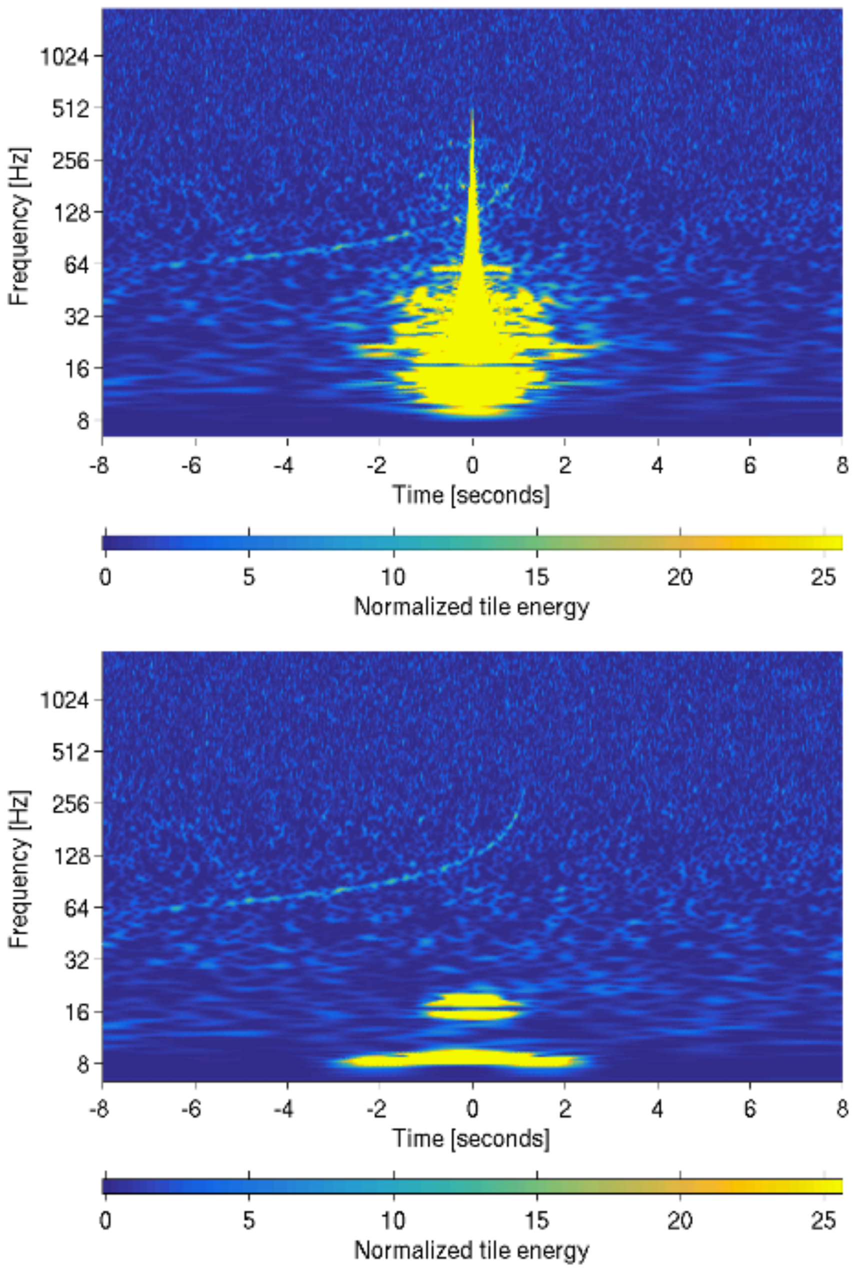
- cWB produces reconstructions of gravitational wave signals.
- It can detect CBC signals as well as bursts.
- https://gwburst.gitlab.io/

GW150914



BayesWave

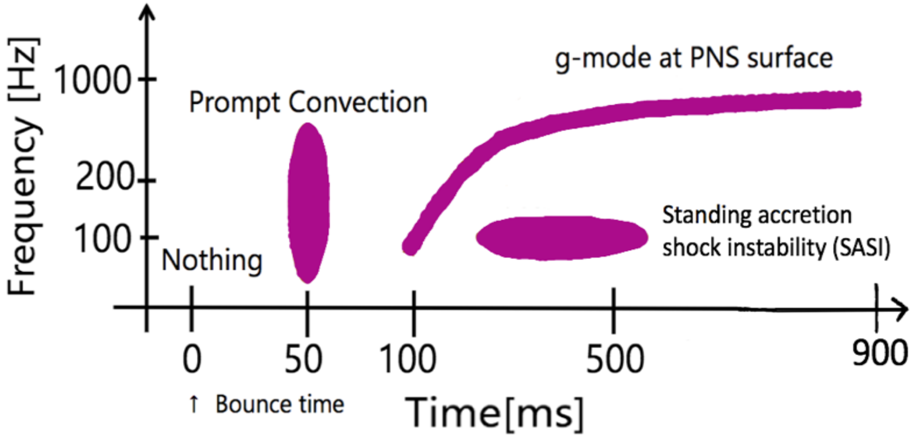
- BayesWave is another standard burst tool. (Cornish, Littenberg, arXiv:1410.3835)
- Models signals as a variable number of sine-Gaussian wavelets with power coherent across detectors.
- It produces unmodelled waveform reconstructions and can remove glitches that occur during signals.

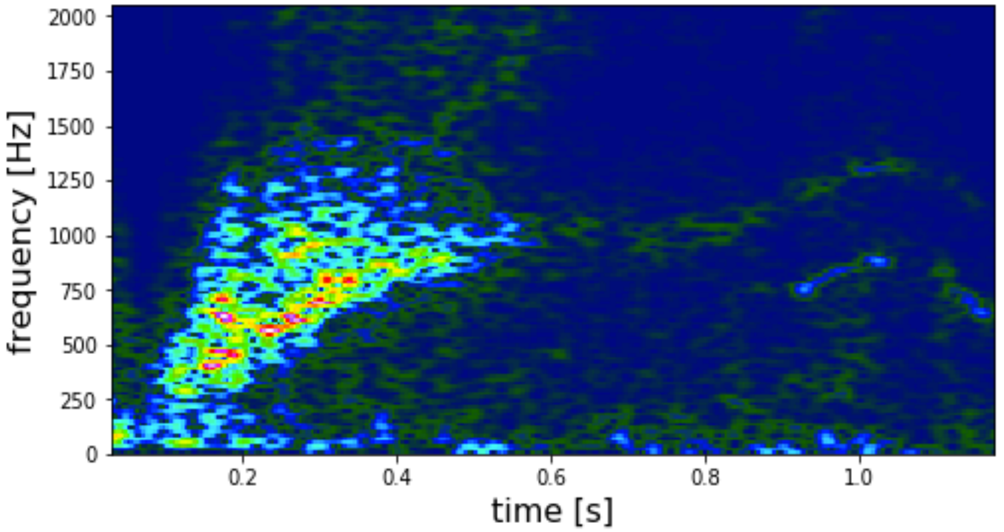


Credit: Pankow arXiv:1808.03619

Supernova Search

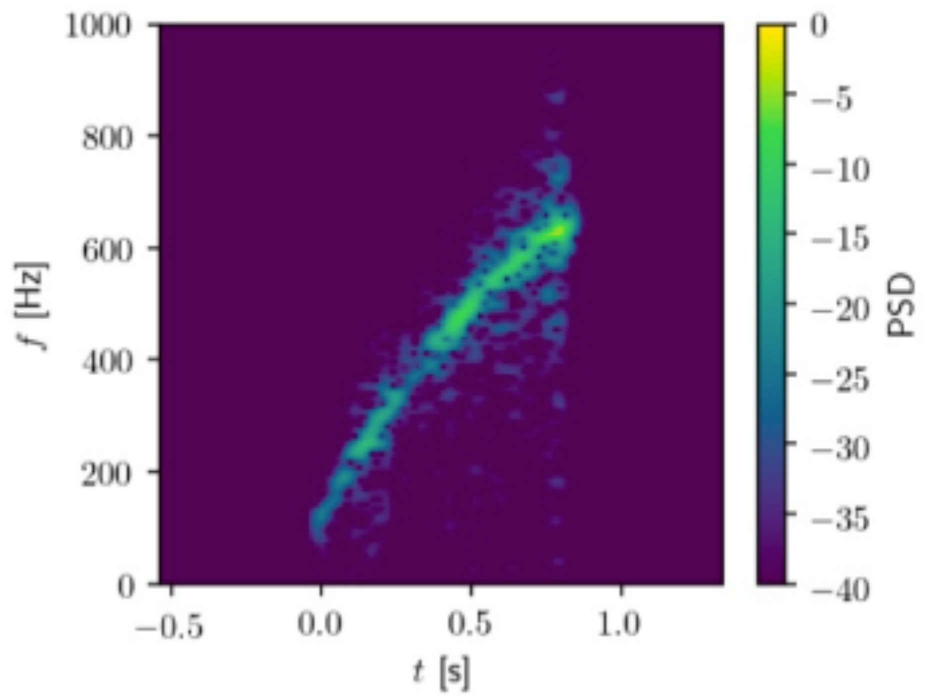
- Some burst searches are for targeted sources like supernovae.
- There is not enough supernova waveforms to match filter search but some supernova waveform features are known.
- The known features from supernova simulations can be incorporated into supernova searches using machine learning.





Supernova Search

- Astone et al. (arXiv:1812.05363) enhance the efficiency of cWB using a neural network.
- The network is trained on phenomenological waveforms that represent the g-mode emission in supernova waveforms.
- They use cWB to prepare images of the data.



Supernova Search Astone et al.

- They use colours to determine which detectors find the signal.

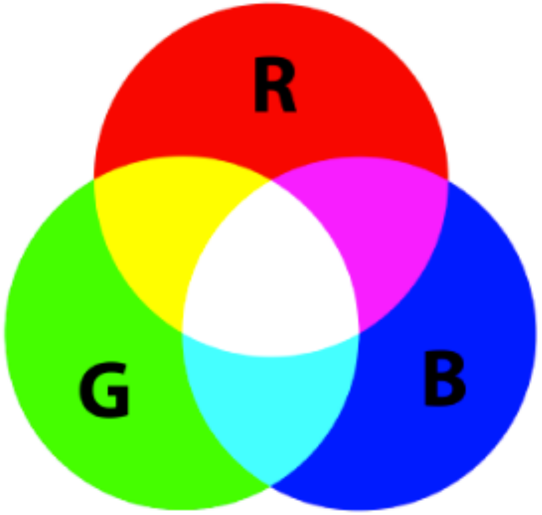


Figure 2: *The mechanism of additive color synthesis. LIGO Hanford is assigned to red, LIGO Livingston to green and Virgo to blue. A triple coincidence will appear in white, while a double coincidence in yellow, magenta or cyan, depending on which couple of detectors is involved.*

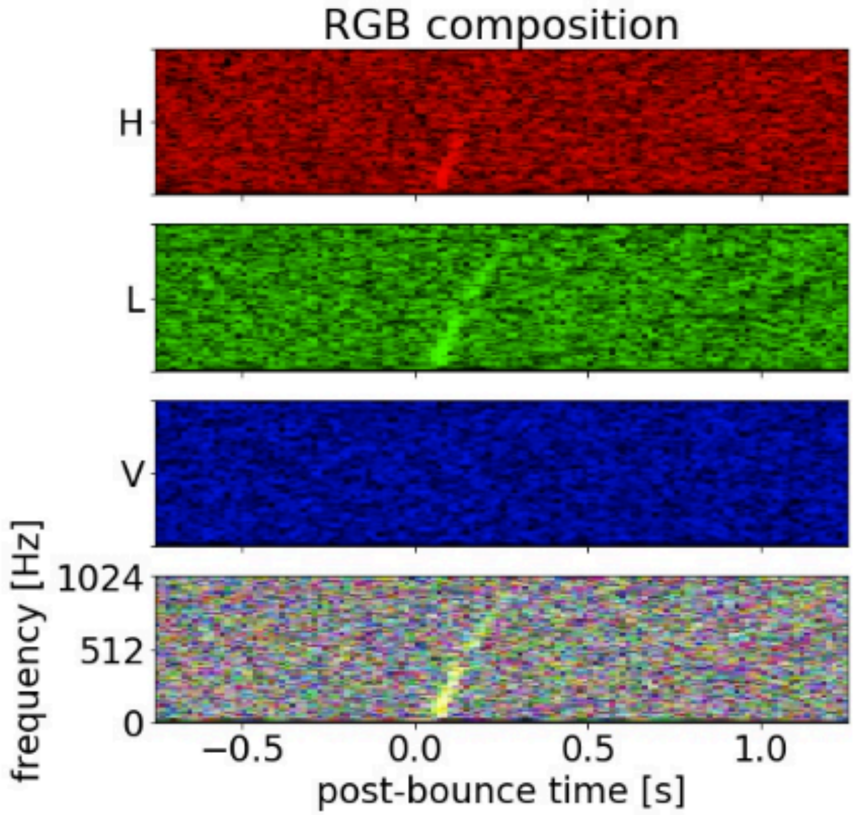
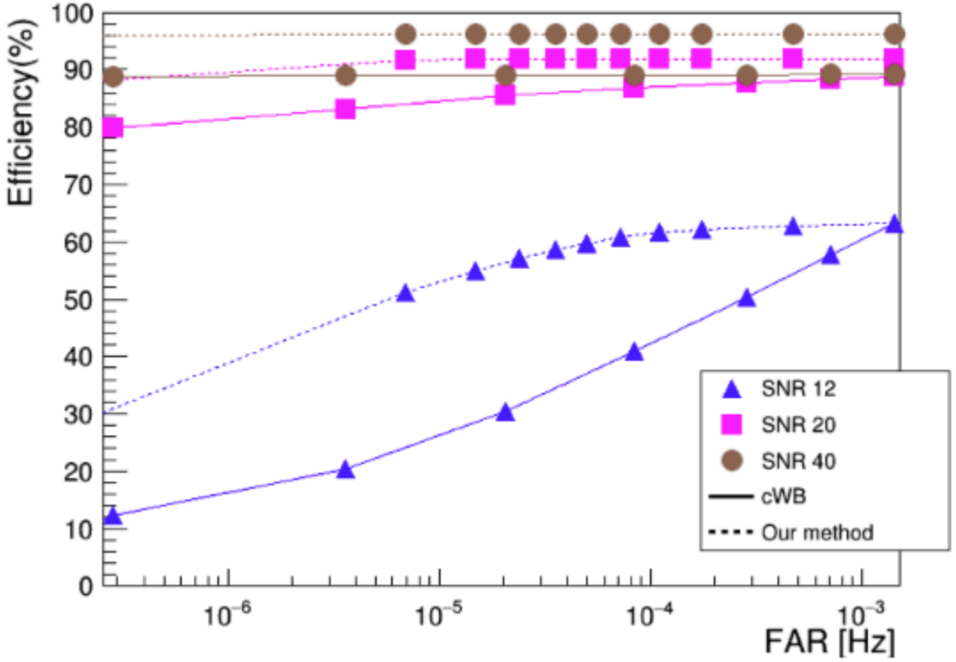


Figure 3: *From the top; the spectrogram of LIGO Hanford is red, then that of LIGO Livingston is green and Virgo is blue. At the bottom: the RGB image obtained by stacking the previous three spectrograms. In this case, the signal is present just in Hanford and Livingston so that the combined signal at the bottom is in yellow.*

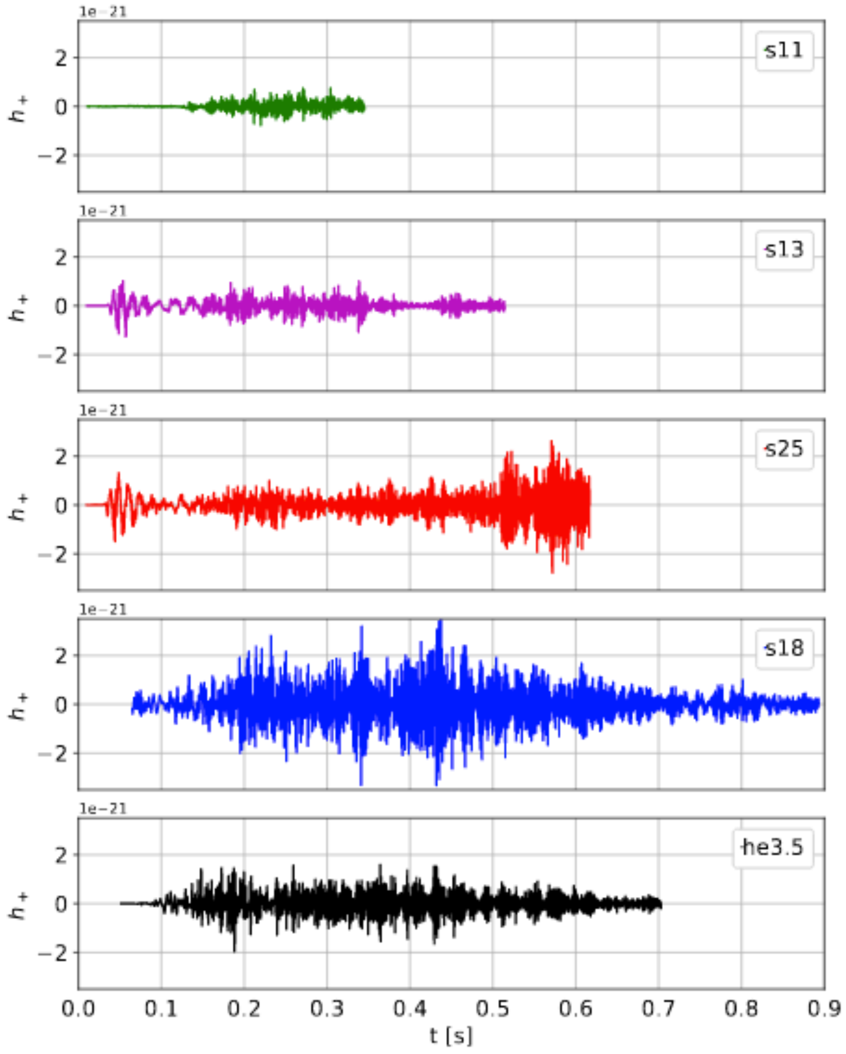
Supernova Search Astone et al.

- They find their method increases the sensitivity of traditional cWB



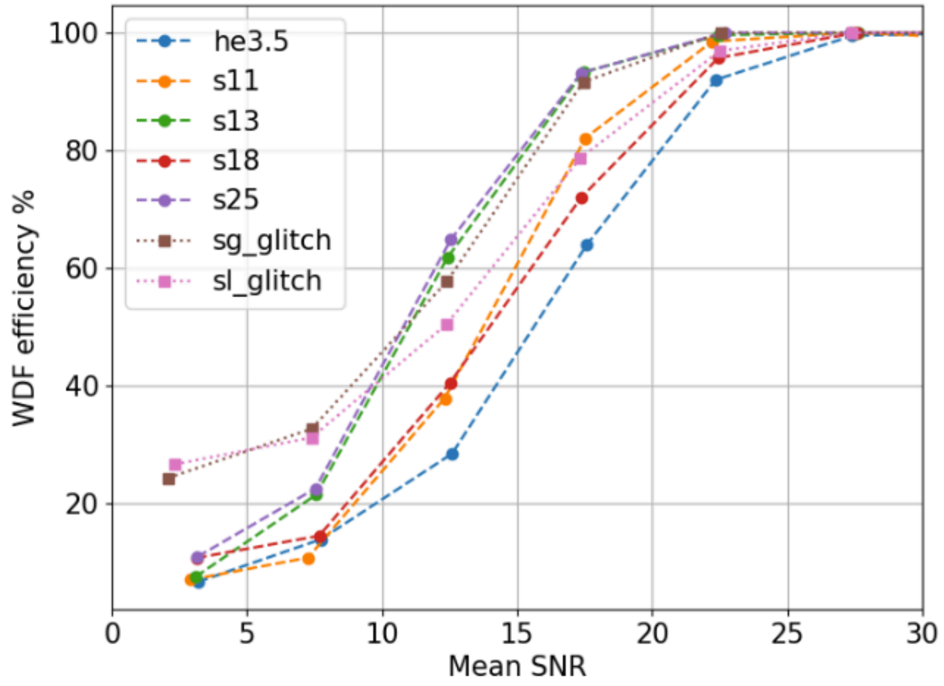
Supernova Search

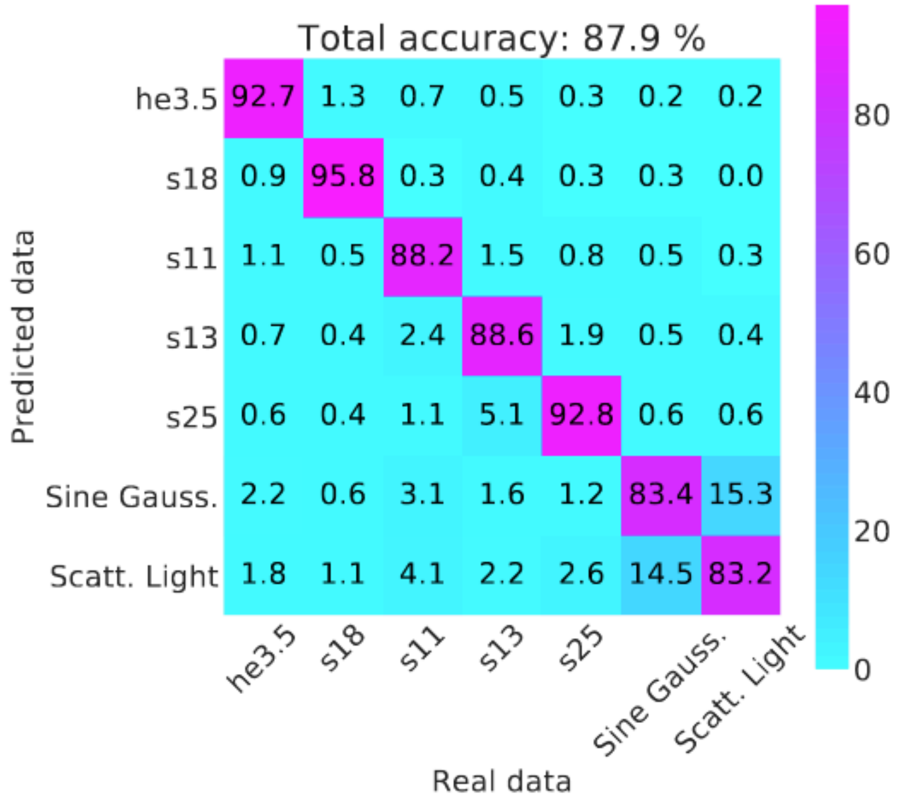
- Iess et al. (arXiv:2001.00279) have a different approach that does not involve cWB.
- They use a trigger generator called WDF to find excess power in the detector.
- Then they do a neural network classification to decide if the trigger is a signal or noise.
- They train directly on supernova waveforms.



Supernova Search – Iess et al.

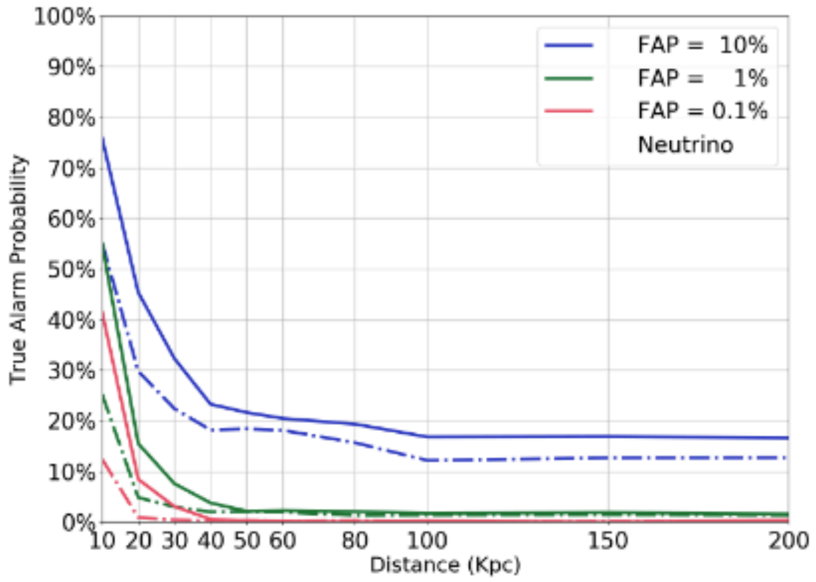
- They use both time series and images of data.
- They obtain high accuracies with both methods and include glitches.



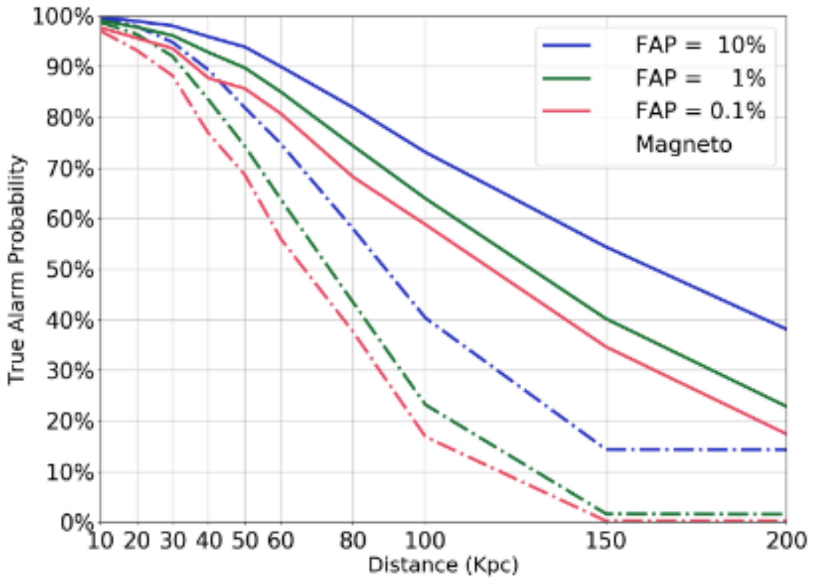


Supernova Search

- Chan et al. (arXiv:1912.13517) also train directly on supernova waveforms.
- They use only the time series waveforms from different explosion mechanisms.



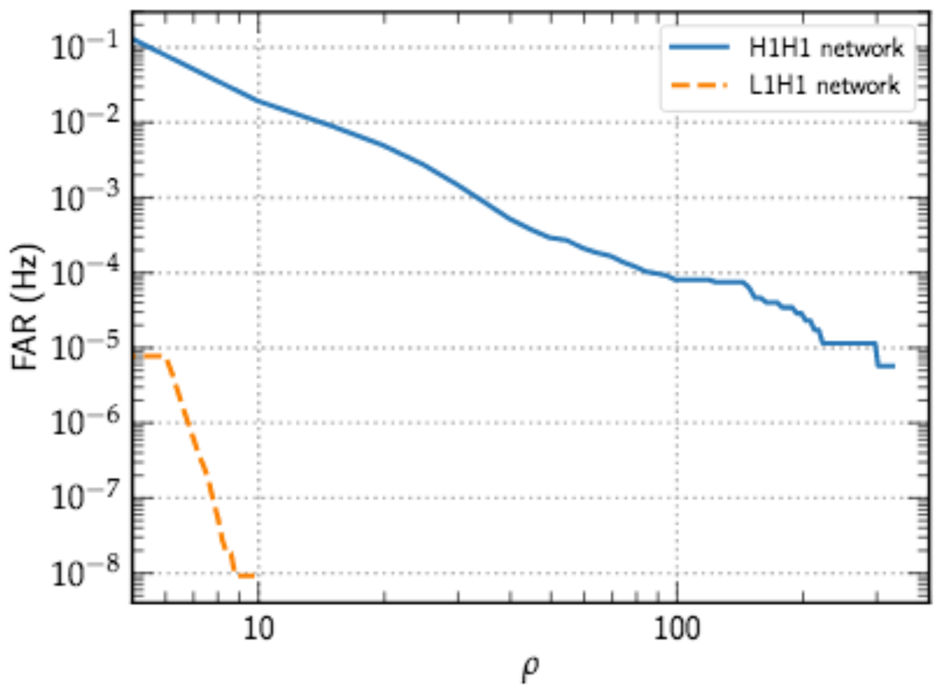
(a)



(b)

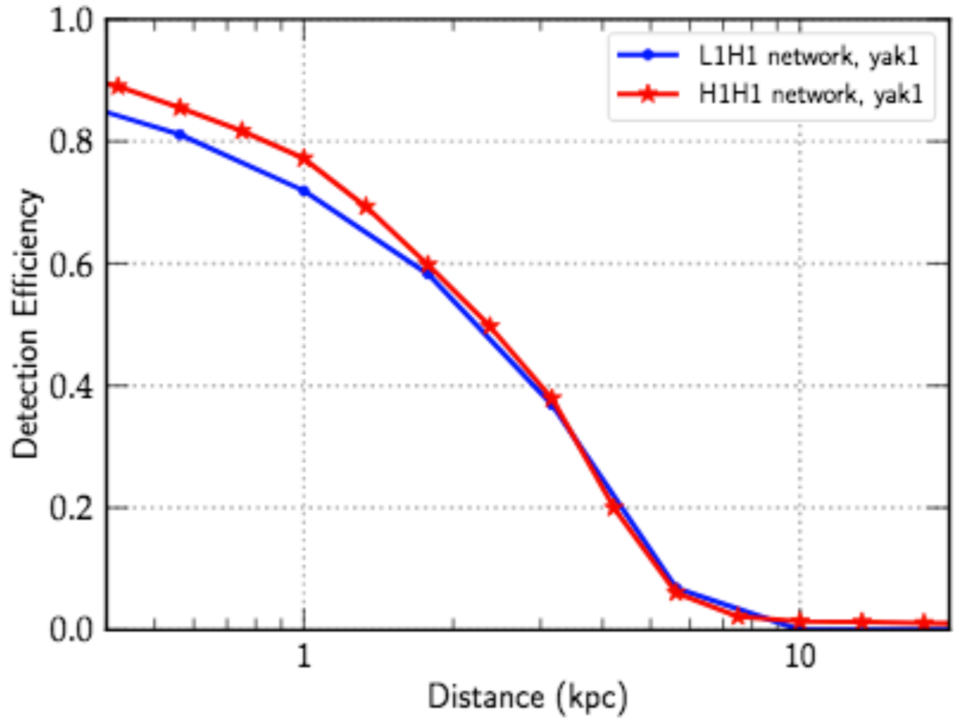
Single Detector Search

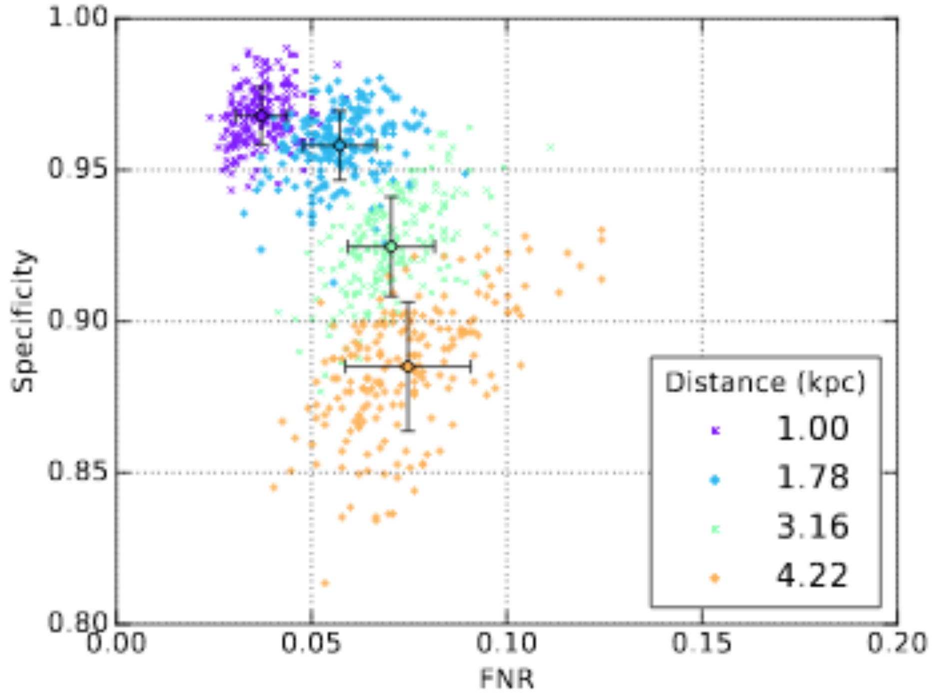
- 30% of gravitational wave data is collected when only 1 detector is in observing mode.
- Can’t do time slides to measure the background if there is only 1 detector.
- Cavaglia et al. (arXiv:2002:04591) use machine learning combined with cWB to perform a single detector search for supernovae.
- They train a genetic programming algorithm on the output parameters of cWB.



Single Detector Search – Cavaglia et al.

- FNR is false negative rate
- Specificity is true negative rate





How can I try it myself?

- You can download some supernova gravitational wave signals here http://www.phys.utk.edu/smc/data.html
- You can get KarooGP here http://kstaats.github.io/karoo_gp/
- You can get Coherent WaveBurst here https://gwburst.gitlab.io/

Conclusions

- Machine learning can improve many areas of gravitational wave science.
- There is still plenty of work left to do!
- You can download the data and try it yourself.