Ameliorating transient noise bursts in gravitational-wave searches for intermediate-mass black holes

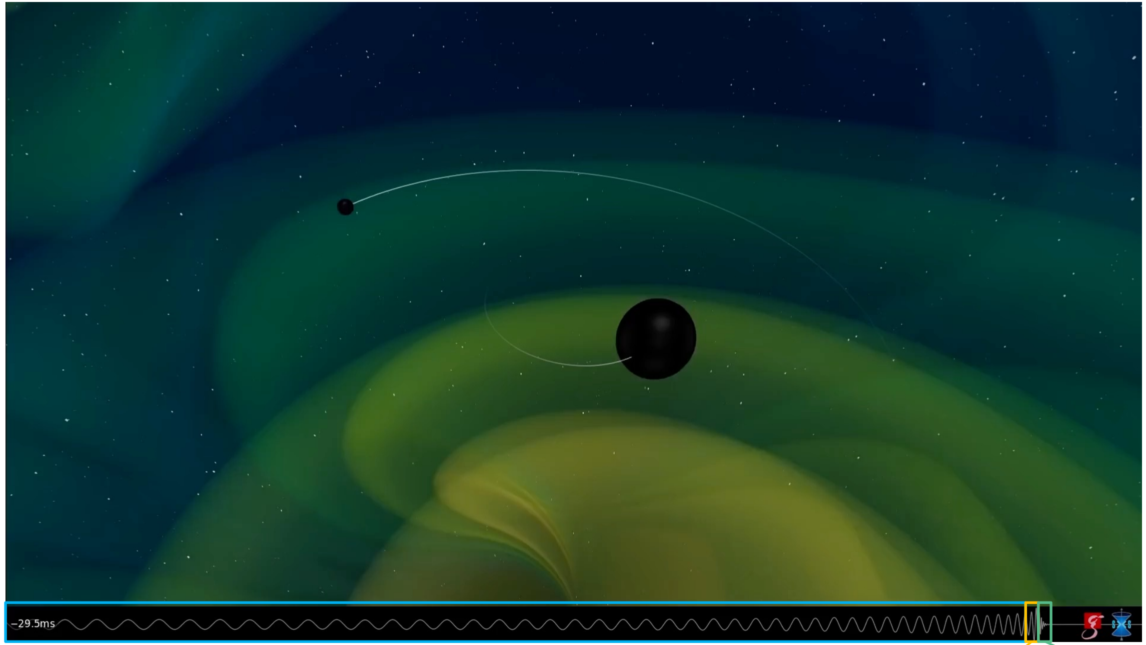
Melissa Lopez, on behalf of the authors

EuCAIFCon 2024

![LIGO Virgo]

What is a compact binary coalescence (CBC)?

Credits: Max Planck Institute



Inspiral

Merger

Ring-down

Binary systems consisting on compact objects in a compact orbit around each other.

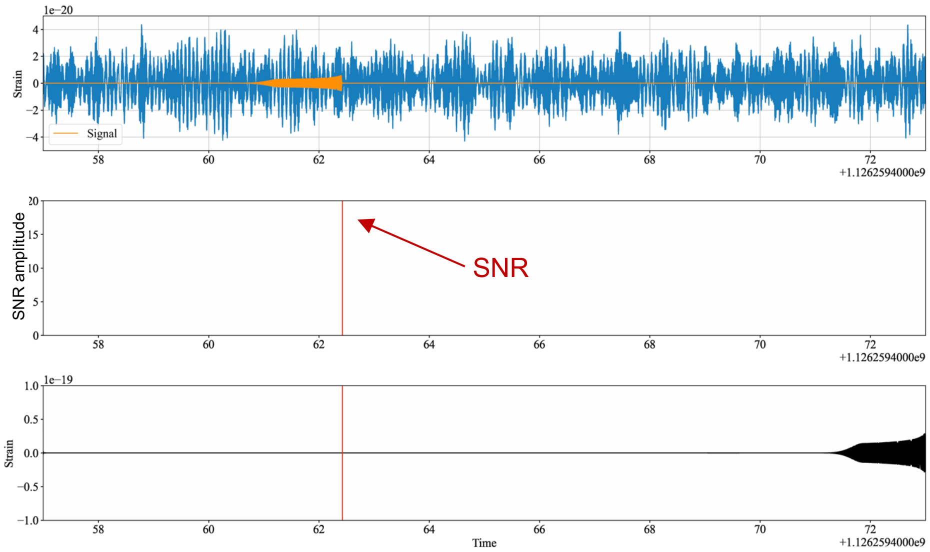
GW simulation parameters:

m_1, m_2: component masses
s_1, s_2: angular momentum or spin

Modelled searches: matched filtering (MF) for CBC

What is matched filtering?

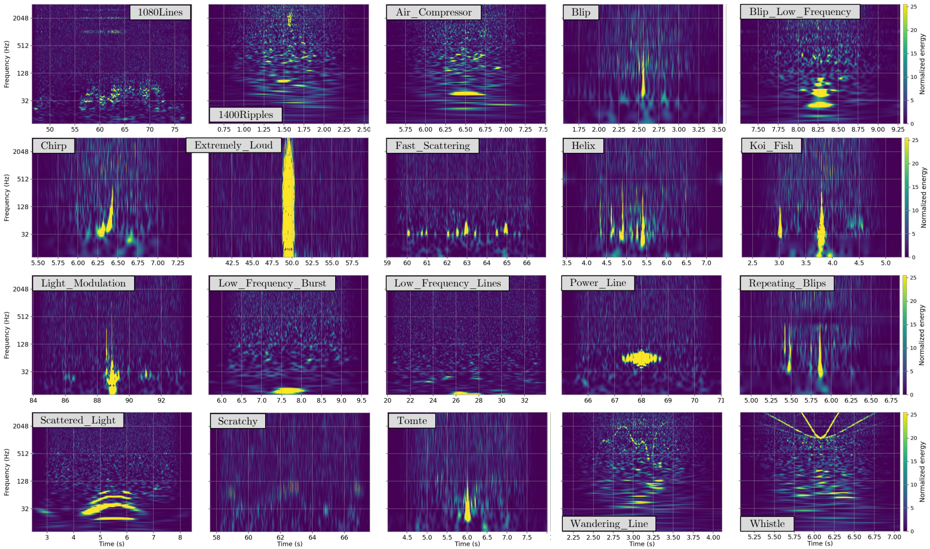
Unknown signal * Template Bank (simulated GW) → Triggers

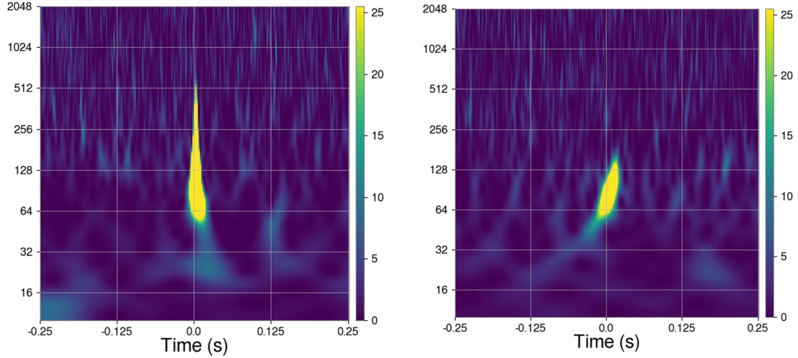


Idea: unknown signals generate *multiple* triggers. Can we find *patterns* with Machine Learning?

Transient noise burst (glitches)

- Caused by instruments or environment (known or unknown)
- Diminish scientific data available
- Hinder GW detection (mask and/or mimic)



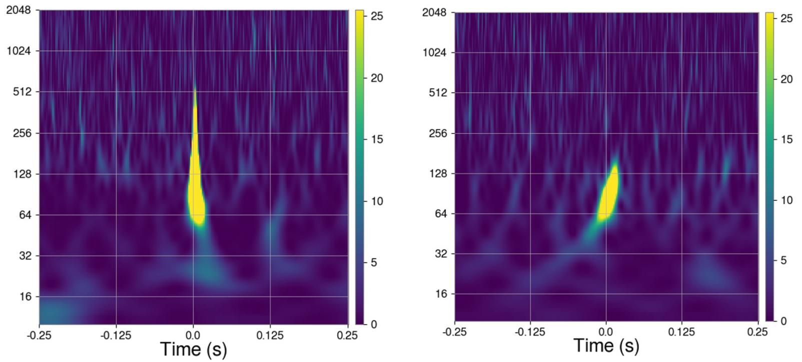


Example of a blip glitch (left) and a intermediate-mass black hole (right)

Motivation

Context: intermediate-mass black holes (IMBH) are the missing link between stellar black holes and supermassive black holes, but they are hard to detect!

Idea: use triggers from matched filtering (free information) from detection algorithms to learn the background (glitches) and foreground (GW signals) with ML



Example of a blip glitch (left) and a IMBH (right)

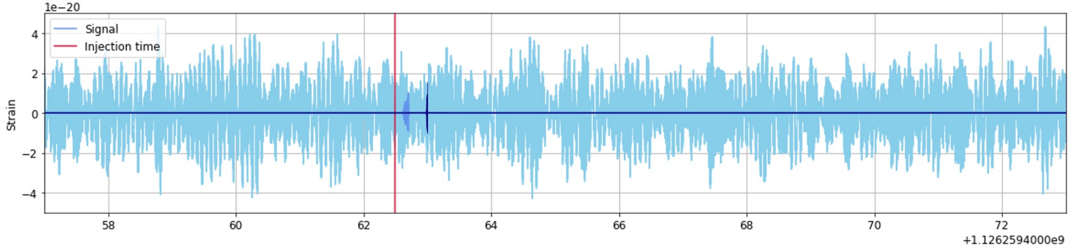
- MF searches use *strict* conditions for detection.
- Can we *relax* the search with the interpolation ability of ML?

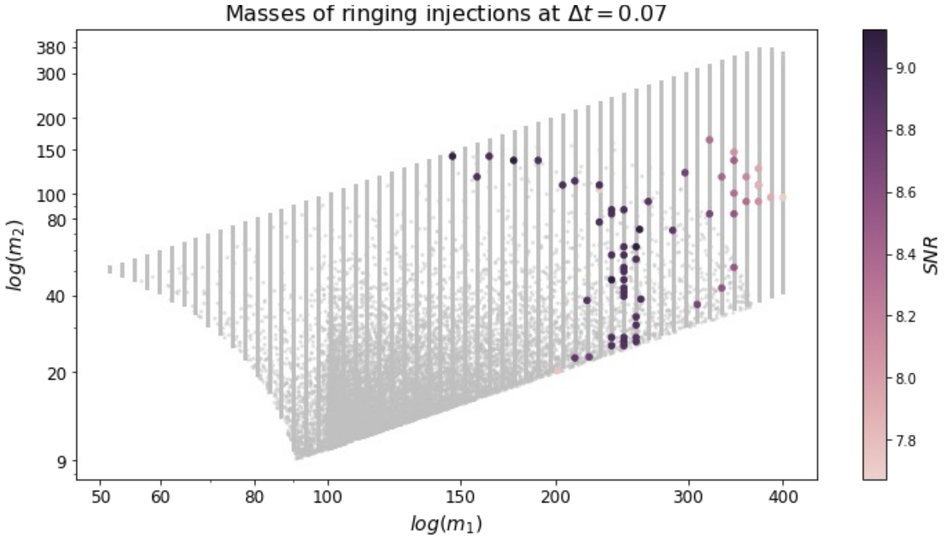
Similar ideas with cWB:

- Gayathri et al. 2020 (XGBoost)
- Lopez et al. 2021 (GMM)

A simulated GW through a detection pipeline

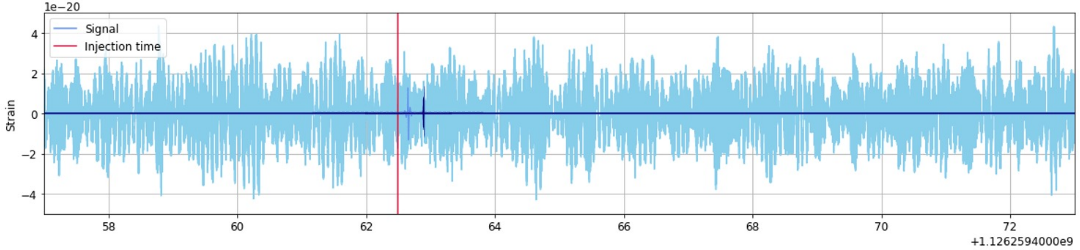
Δt: time when trigger happened – time when GW signal was added to the noise

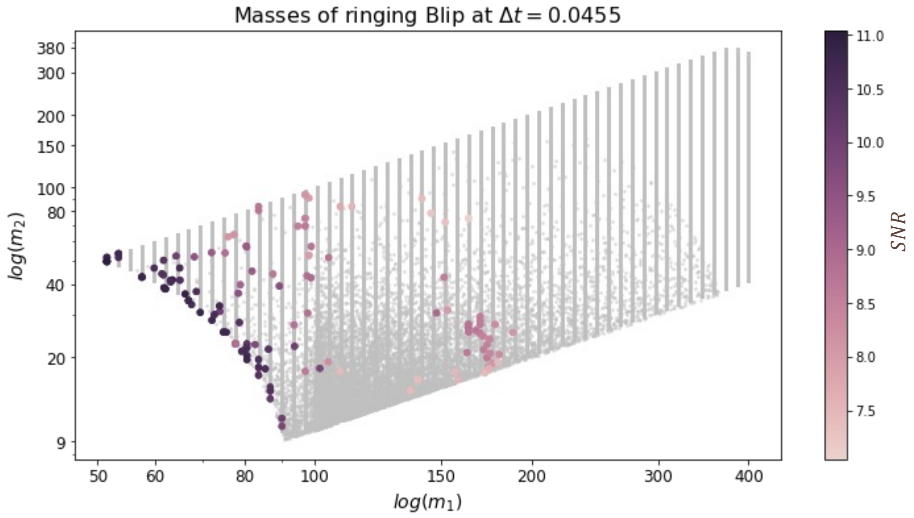




A glitch through a detection pipeline

Δt: time when trigger happened – time when glitch happened





Methodology

Task: Distinguish IMBH from different glitch classes in <u>single detector</u> → we have 3 detectors!

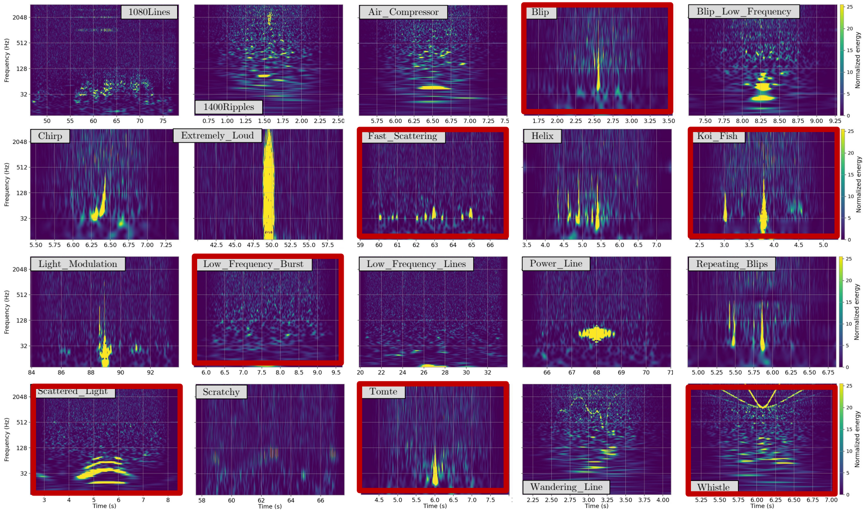
Data: reproduce IMBH search GstLAL O3 but truncate it → only matched filtering

Algorithm: Multi-layer perceptron (MLP)

Input: Adding time is hard, so let's simplify the problem. Each template is defined by $m_1, m_2, s_{1z}, s_{2z}, \chi^2, SNR$. We weight average by SNR to get the feature vector

$$\mu(m_1, m_2, s_{1z}, s_{2z}, \chi^2, SNR)$$

Output: class probability



Idea: MLP differentiates **6 classes**: 5 different types of background (glitches) and single foreground (GW signals). It uses only **6 parameters** in **single detector**

Methodology

1. Accounting for imbalanced data (boostrapping with replacement)

Injections (GR GW)

Blips

Blips

…

Whistles

Whistles

Whistles

Whistles

N=100 → Shuffled data

2. Improving model prediction (k-fold cross-validation): train & validate

K=1

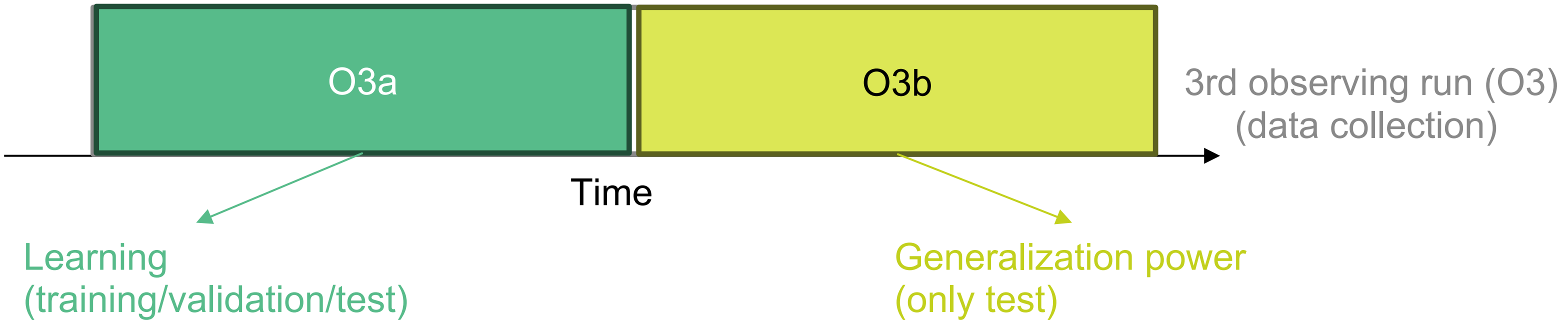
K=2

…

K=9

We avoid unbalanced data and overfitting

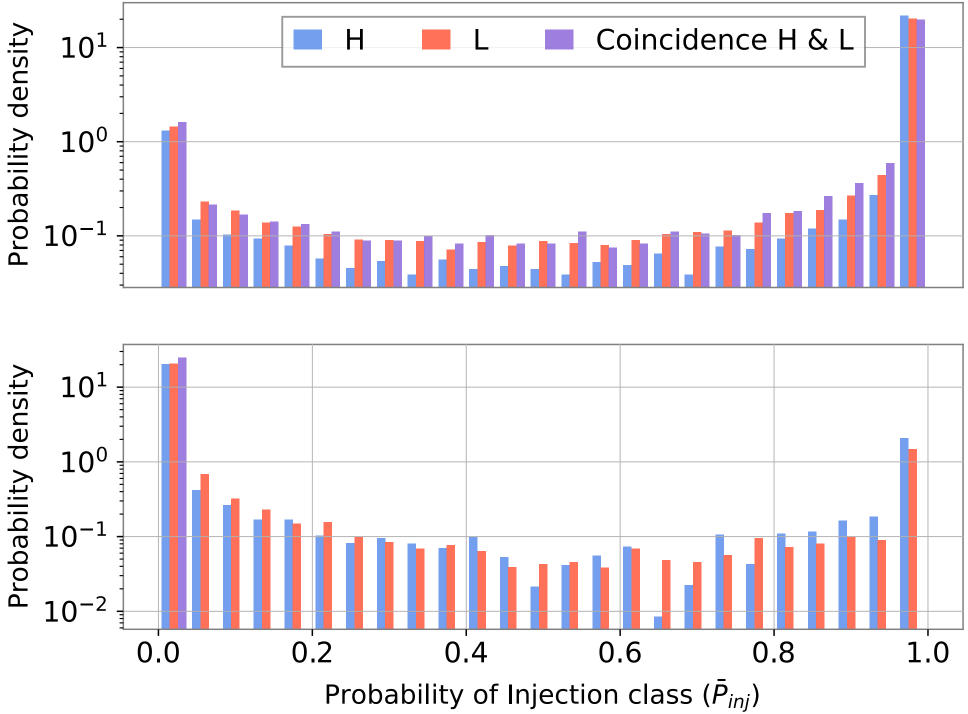
Results: diving into the known (controlled data set)



Detector	Accuracy O3a (%)	Accuracy O3b (%)
LIGO Hanford (H1)	98.75	73.01
LIGO Livingston (L1)	95.81	67.45
Virgo	99.27	75.91

Accuracy drops due to loud and confussing background.
Idea: *time coincidence* among detectors since *most* glitches do not overlap

Results: diving into the known (time coincidence)

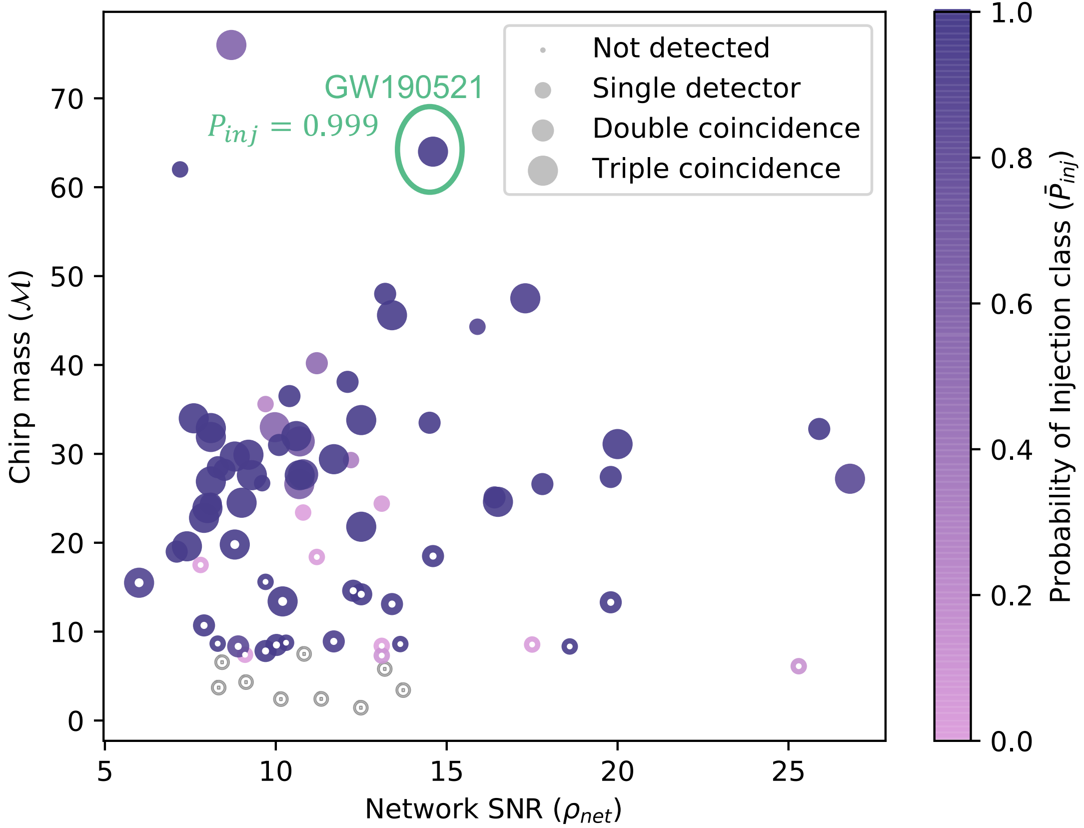


Known foreground (simulated IMBH signals)

Known background (all glitches)

GW simulation = Injection

Results: diving into the (un)known – GWOSC catalogue

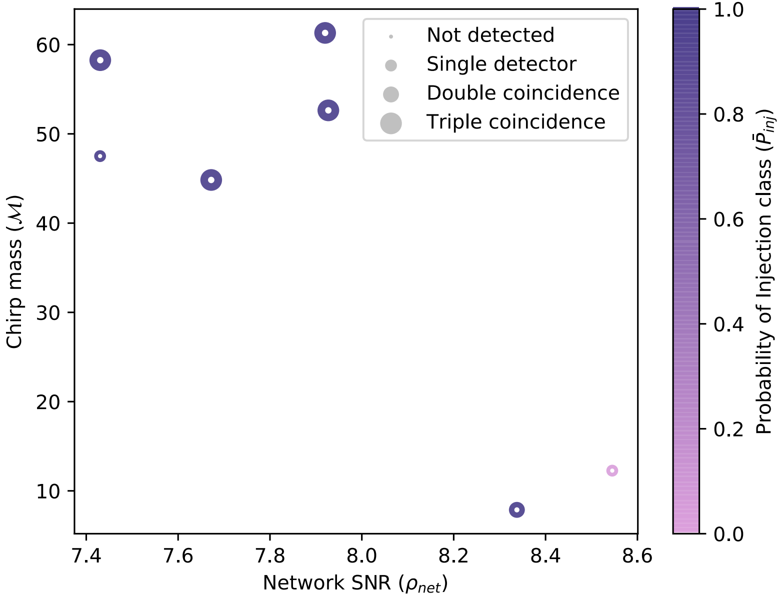


It has learnt to identify GW simulations. Can we see real GW signals?

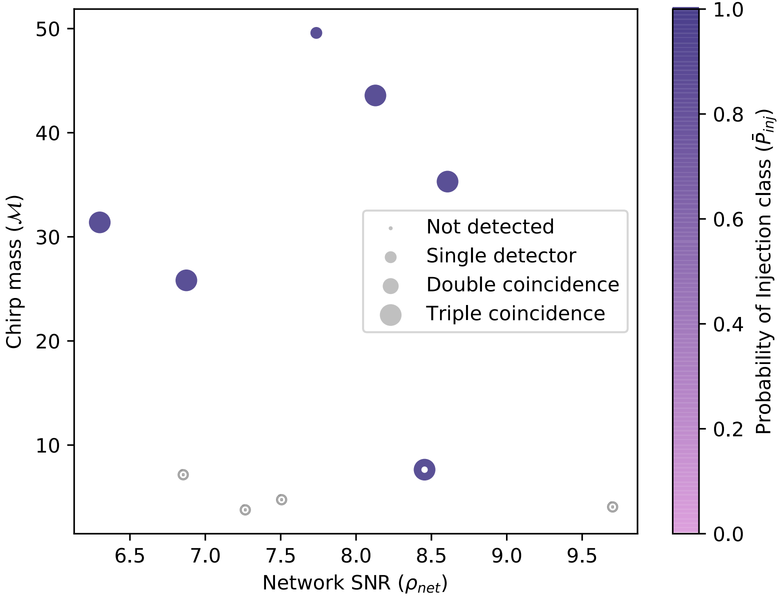
WARNING: this is *not* a detection, i.e. **no** background estimation

Results: diving into the (un)known – Other catalogues

WARNING: this is *not* a detection, i.e. **no** background estimation



4-OGC, Nitz et al. 2023 *ApJ* **946** 59



Olsen et al. 2023, Phys. Rev. D **106**, 043009

Conclusions & future work

- ✓ We can differentiate signals from glitches with MF triggers & neural networks
- ✓ We learned O3a and explored generalization power in O3b
- ✓ Identified GW190521 and other GW signals
- ✓ Paper out soon. Stay tuned!

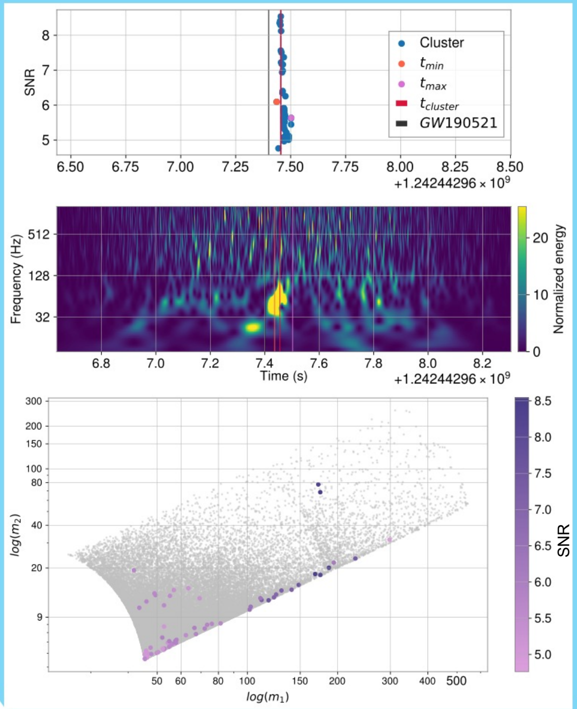
- ➢ Background estimation
- ➢ Include time component: ordering might be relevant
- ➢ Extend to other template banks and/or pipelines

Thank you for listening! Questions?

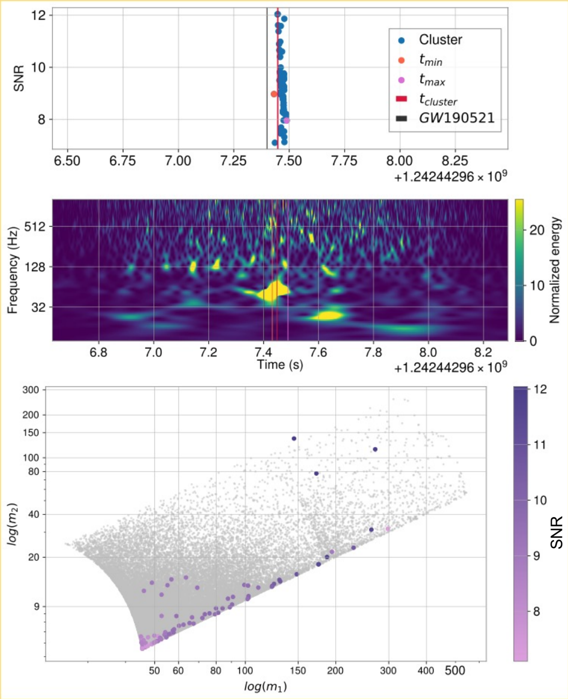
Extra slides

Results: diving into the (un)known - GW190521

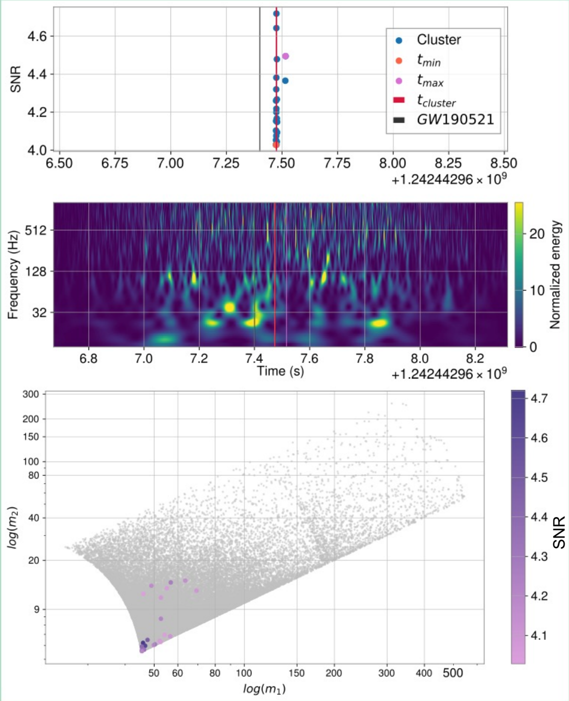
LIGO Hanford (H1)



LIGO Livingston (L1)

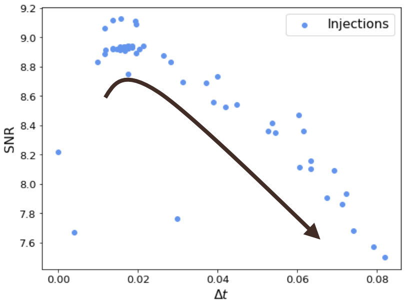


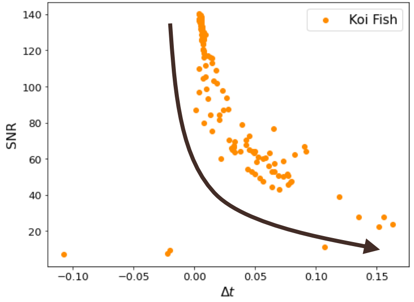
Virgo (V1)

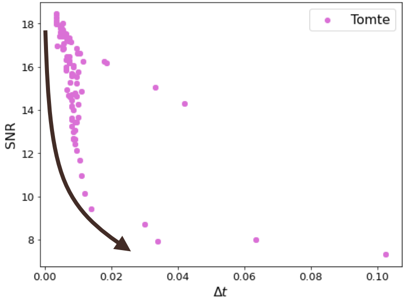


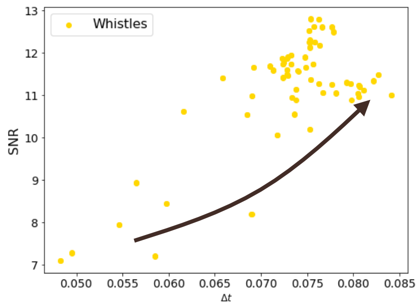
What does the dynamic SNR pattern look like?

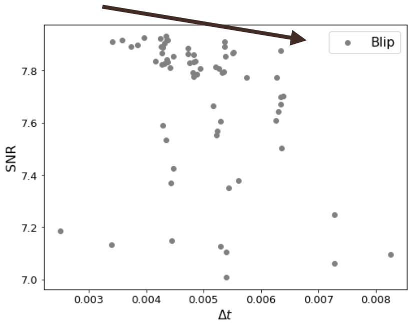
Taking time interval: -1s < event time < 1s











Not all glitches produce many injections/triggers