

# Searches for Intermediate Mass Black Holes with Coherent WaveBurst

cWB Lecture 2

Tanmaya Mishra and Marek Szczepanczyk

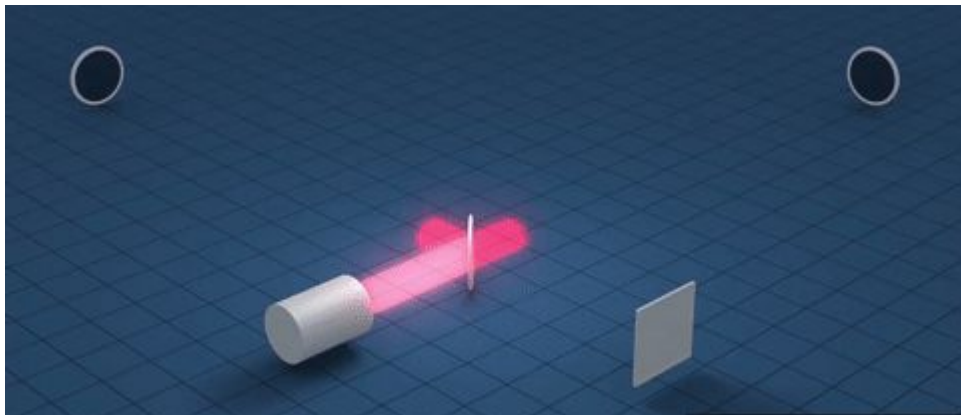


*HEP Software Foundation India - IUCAA Joint Computing Workshop*

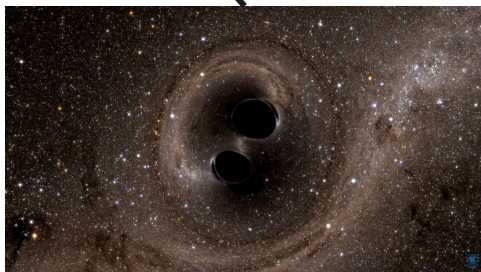
**IUCAA, Pune**

**April 20, 2026**

# Detecting GW Signals from CBC



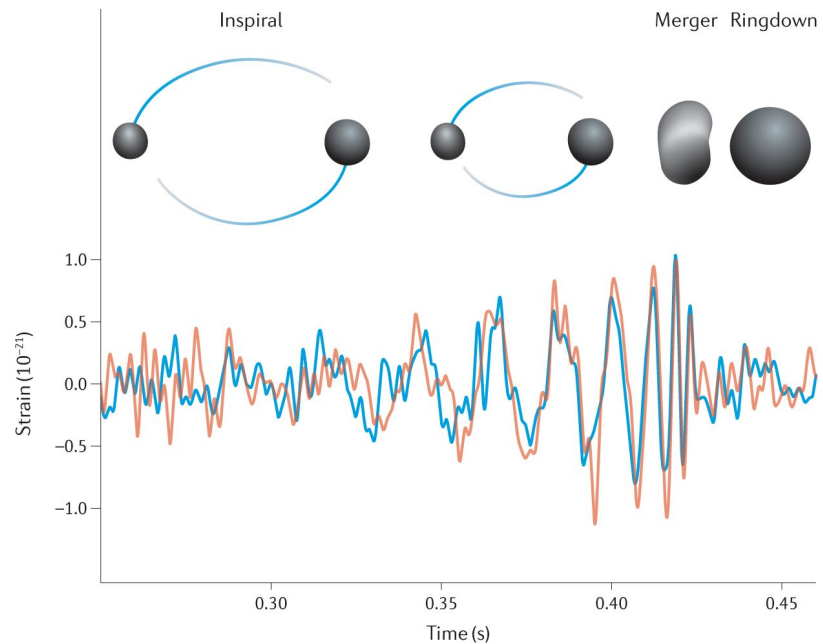
[T. Pyle, Caltech/MIT/LIGO Lab]



[SXS Project]



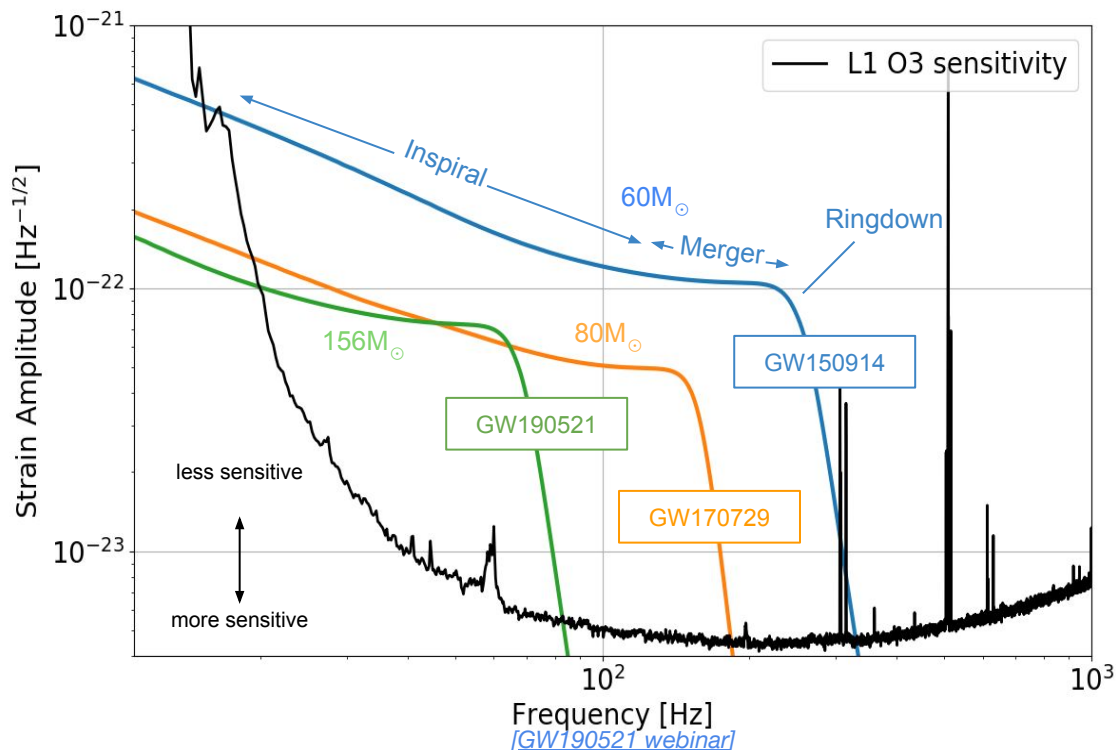
$$h = \frac{\Delta L}{L}$$



[Phys. Rev. Lett. 119, 161101]

# LIGO Detectors Sensitivity

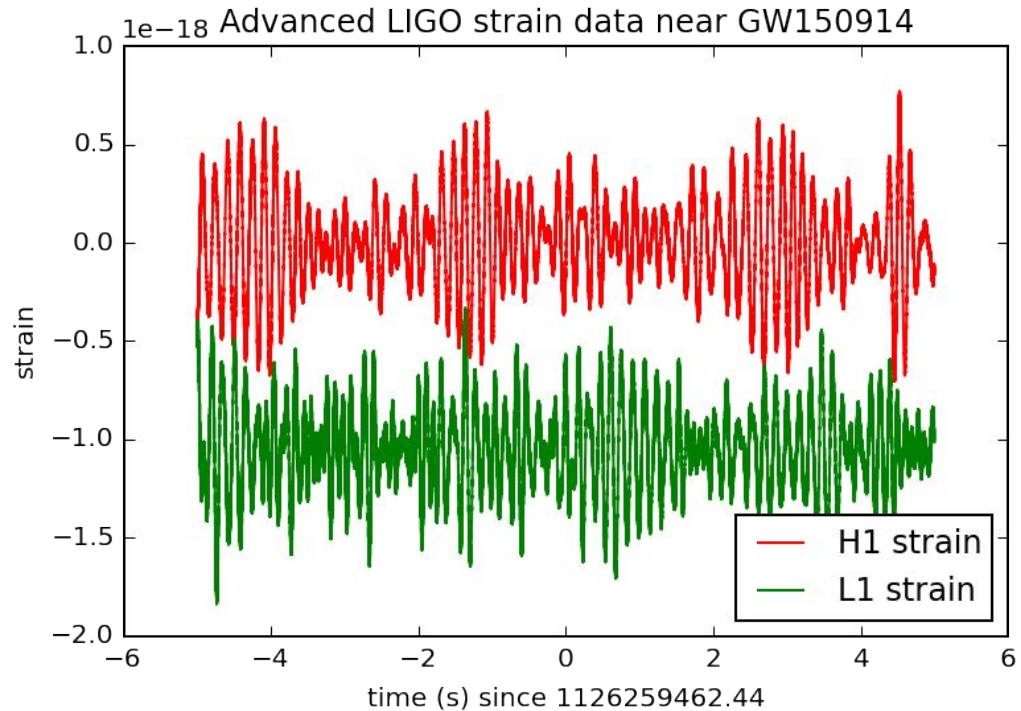
Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors at Hanford (H1) and Livingston (L1).



- Each detector has two 4km long arms, setup like a Michelson interferometer with Fabry-Perot arm cavities.
- Extreme sensitivity attained of the order of  $h = \Delta L/L \sim O(10^{-23})$  which is the expected gravitational wave strain from an astrophysical source.
- LIGO sensitive frequency band 10Hz - 1000Hz.

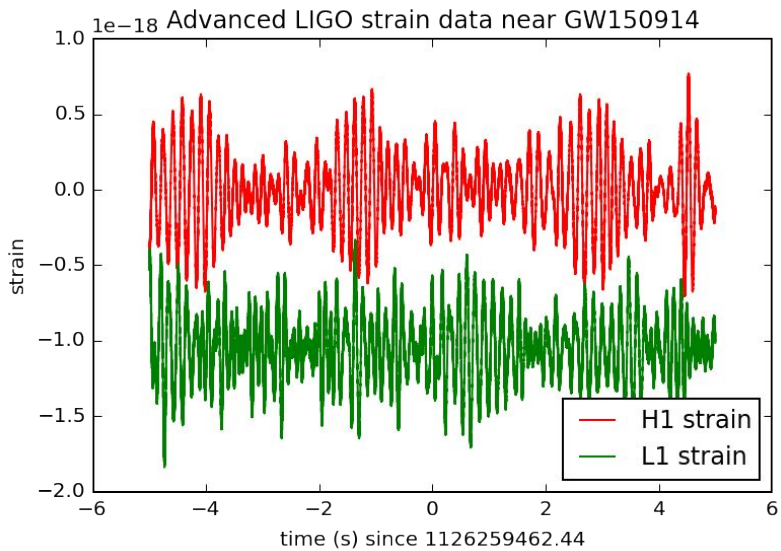
# Strain Data from LIGO

- LIGO captures strain data  $s(t)$  at a sampling rate of **16kHz**
- Downstream analysis, like searches for CBCs use downsampling at **4kHz**
- Strain in the detectors is  $O(10^{-19})$
- Data calibrated in the 10Hz - 5000Hz (limit to Nyquist frequency depending on sampling choice)
- $s(t) = h(t) + n(t)$

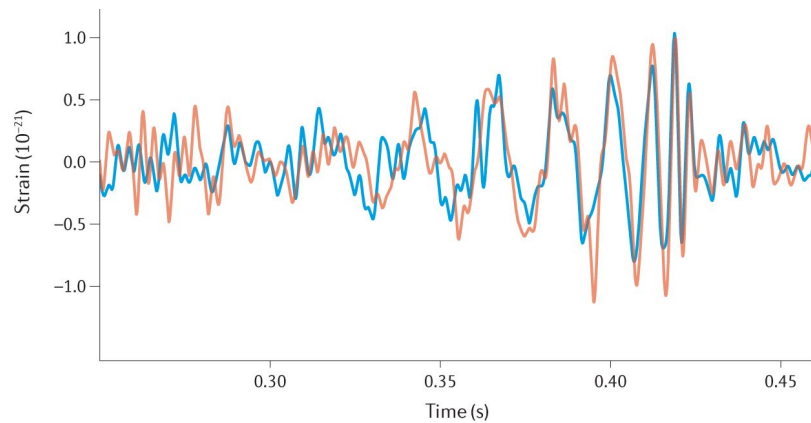


<https://gwosc.org/workshops/>

# Strain Data from LIGO



<https://gwosc.org/workshops/>



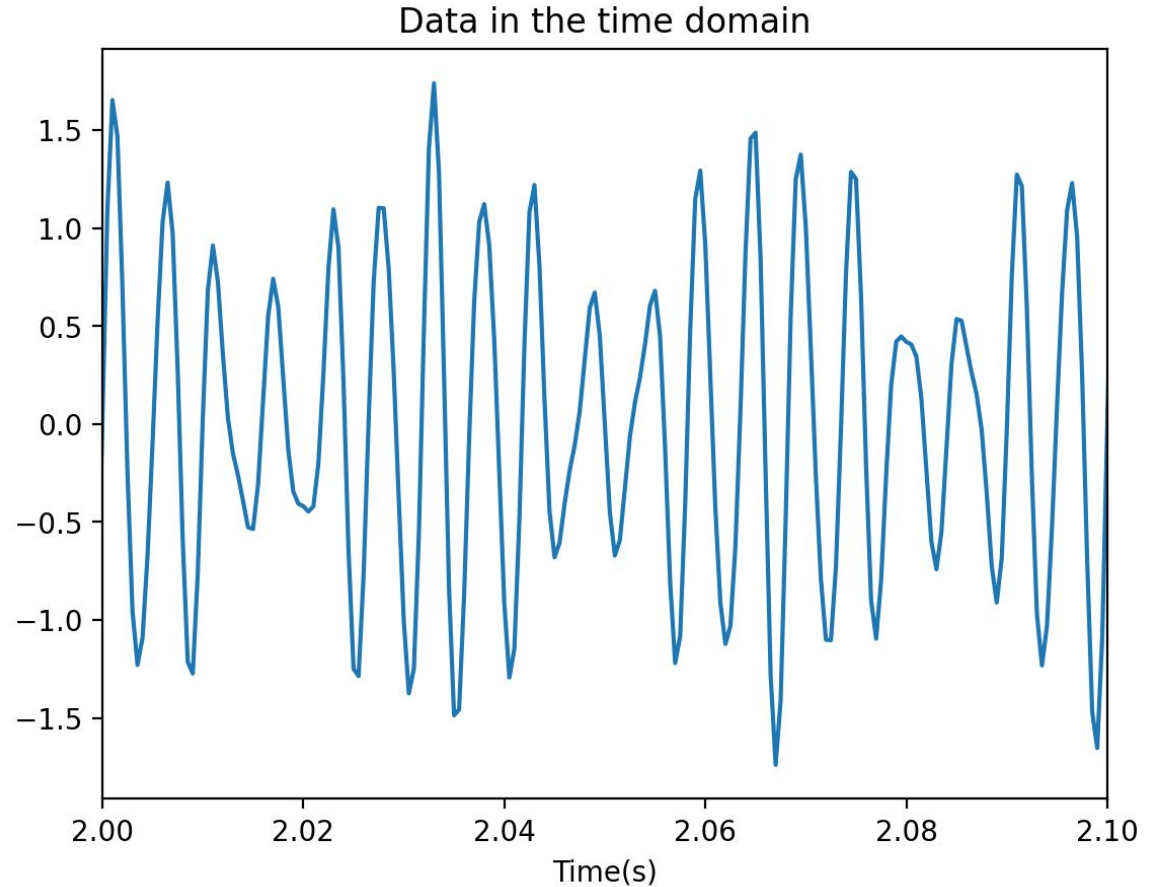
*[Phys. Rev. Lett. 119, 161101]*



# Basic Signal Processing

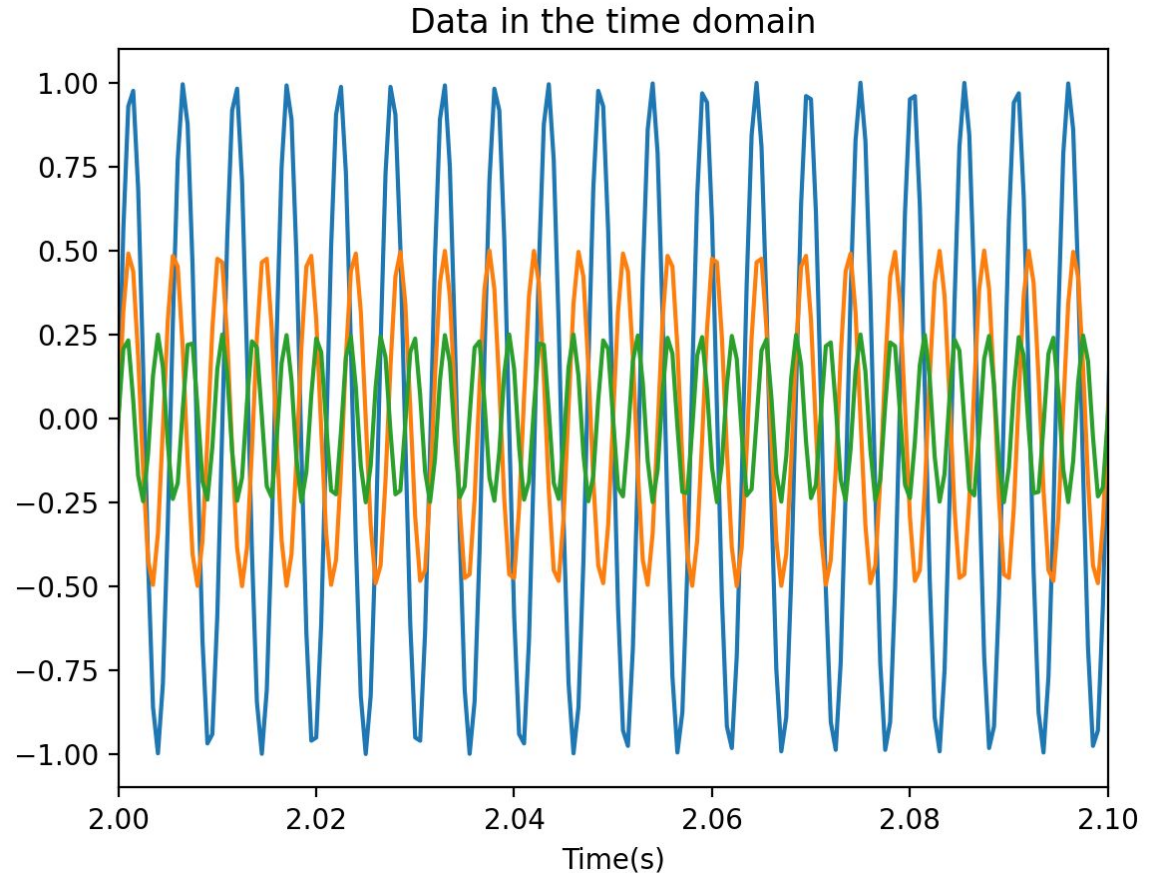
# Fourier transform

- Time-domain data appears noisy and can hide underlying signals



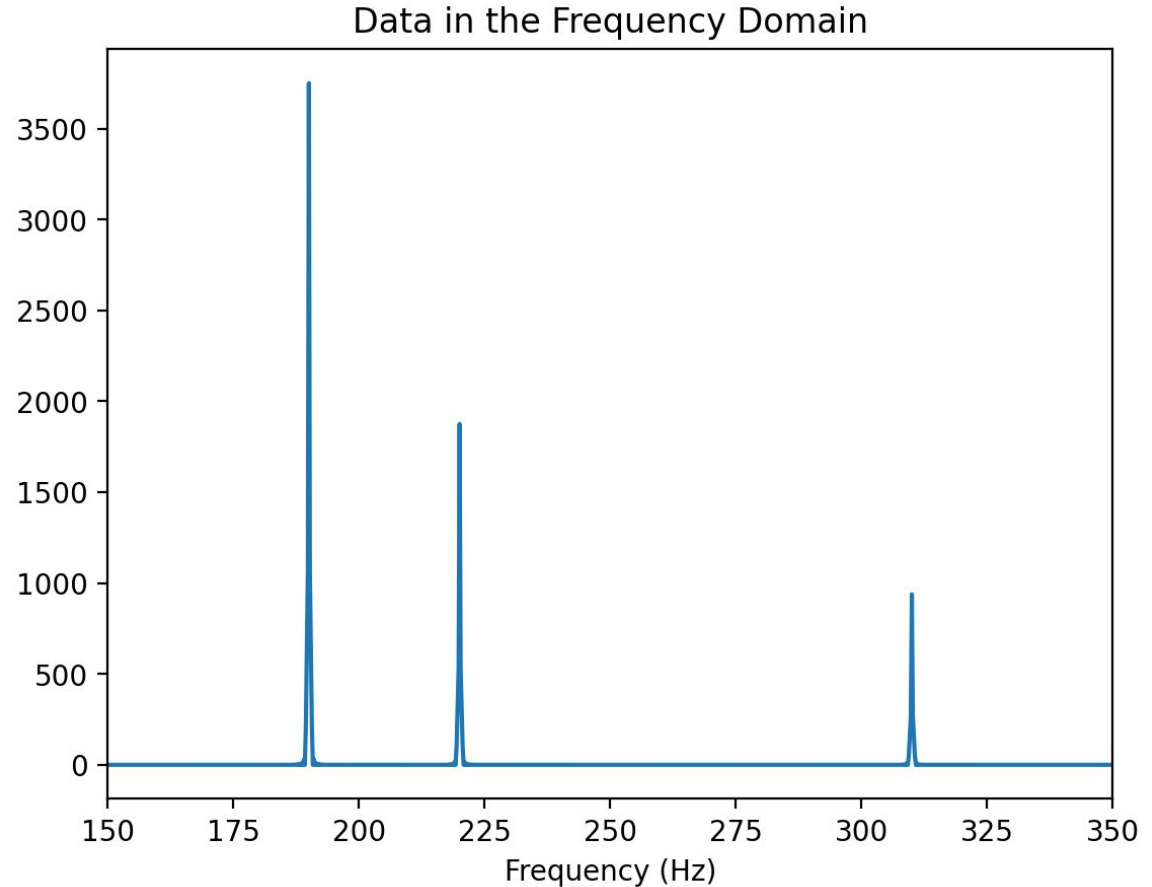
# Fourier transform

- Time-domain data appears noisy and can hide underlying signals
- Access the frequency-domain content of a time-domain signal



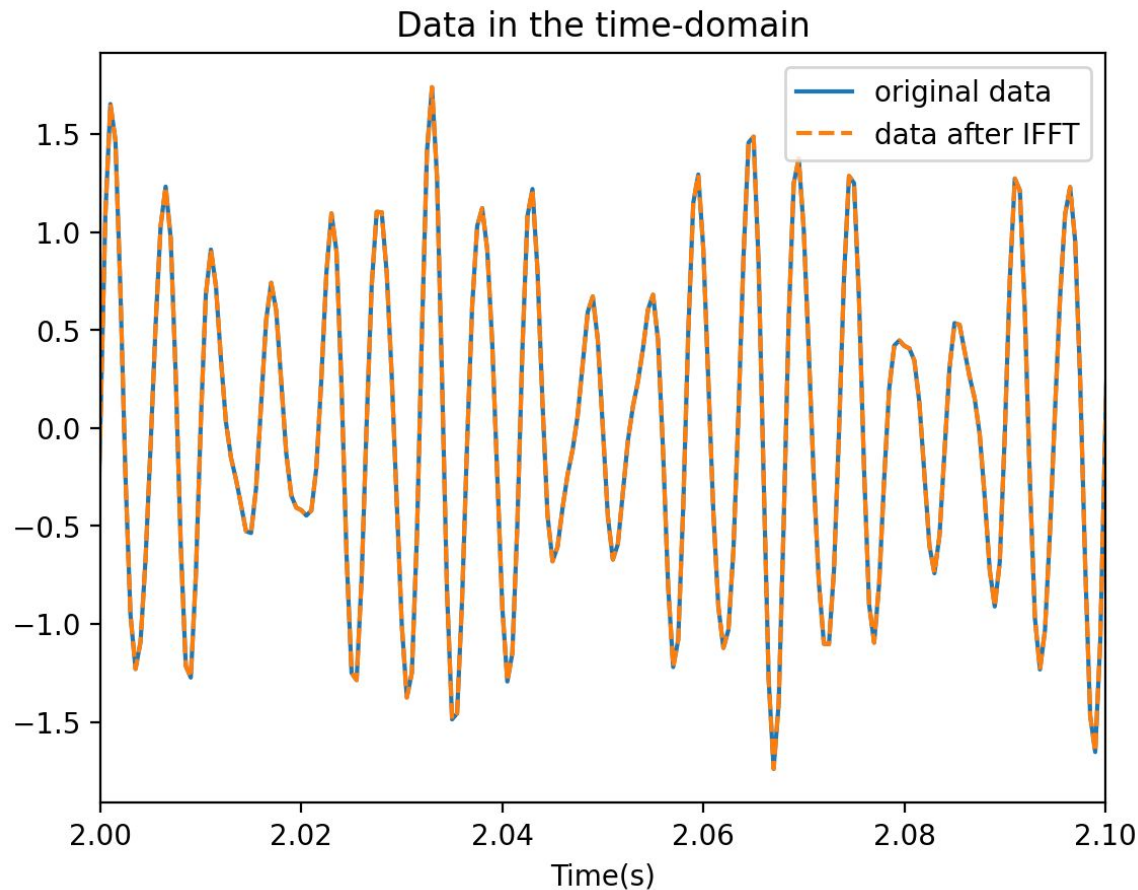
# Fourier transform

- Access the frequency-domain content of a time-domain signal
- Applying Fast Fourier Transform
- $n(t) \rightarrow n(f)$



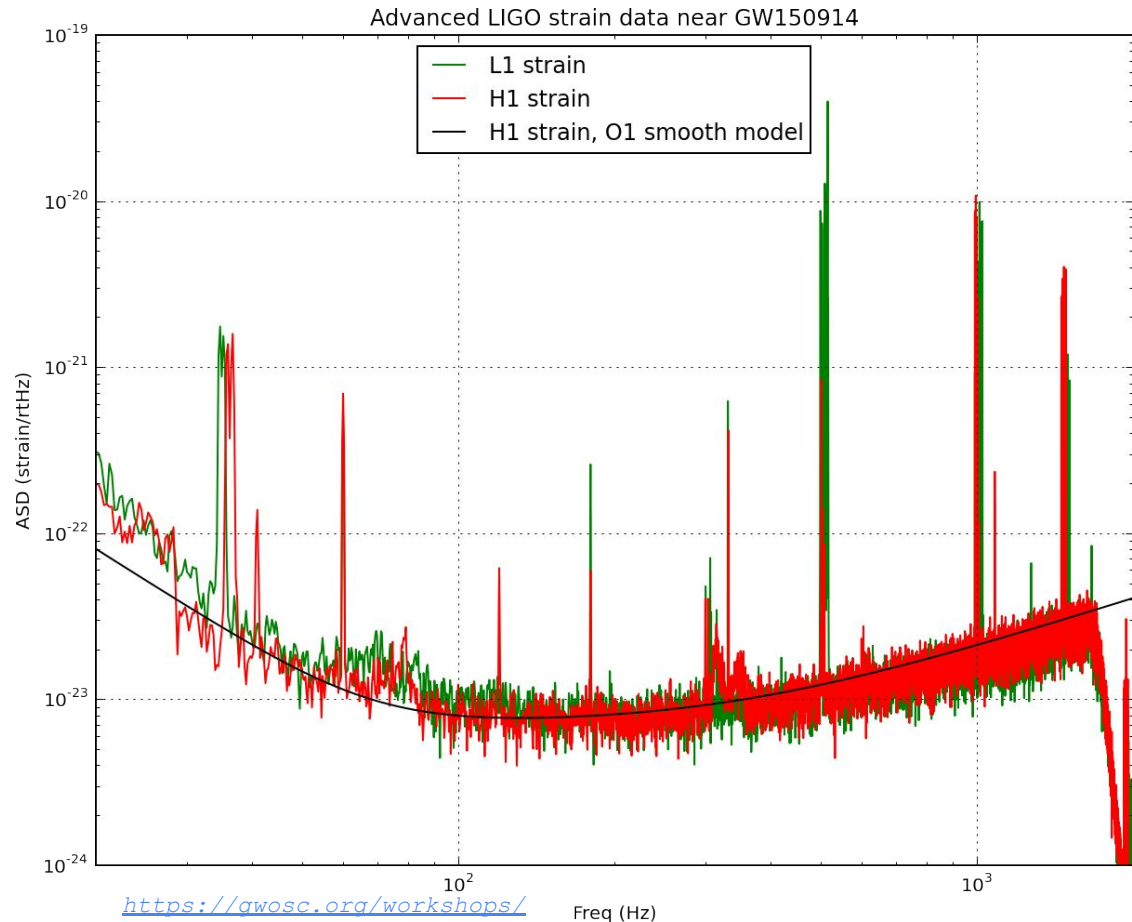
# Fourier transform

- Access the frequency-domain content of a time-domain signal
- Applying Fast Fourier Transform
- $n(t) \rightarrow n(f)$
- Inverse FFT to get back time-domain signal



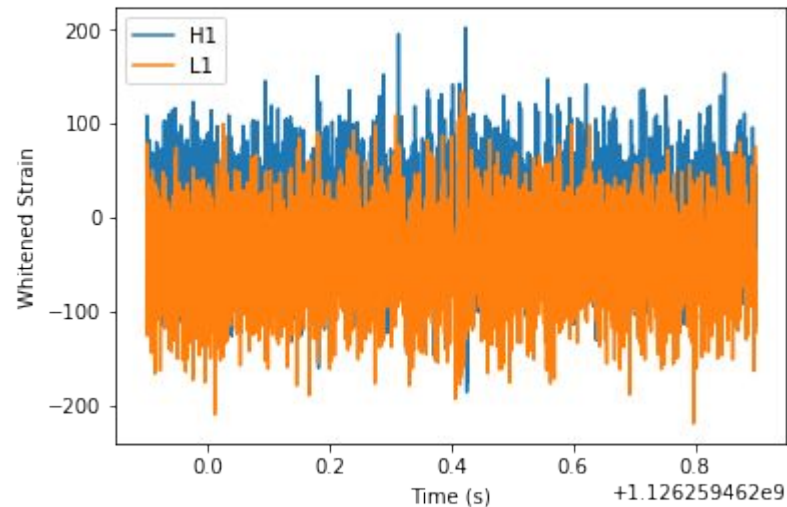
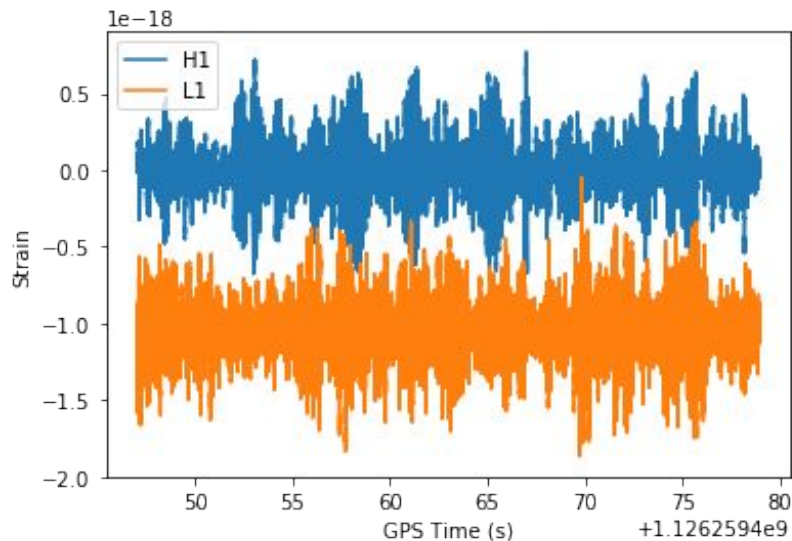
# Power Spectral Density

- The noise power spectral density is given by the mean square energy at each frequency
- $S_n(f) = |n(f)|^2$
- $ASD = \text{sqrt}(S_n(f))$
- Averaging over  $T$  ( $\sim 32\text{s}$ ) gets rid of transient glitches and signals and makes the ASD cleaner



# Whitening

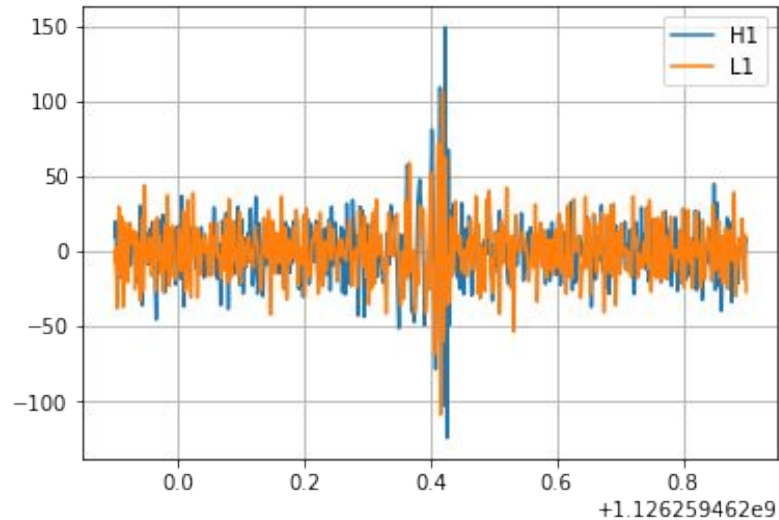
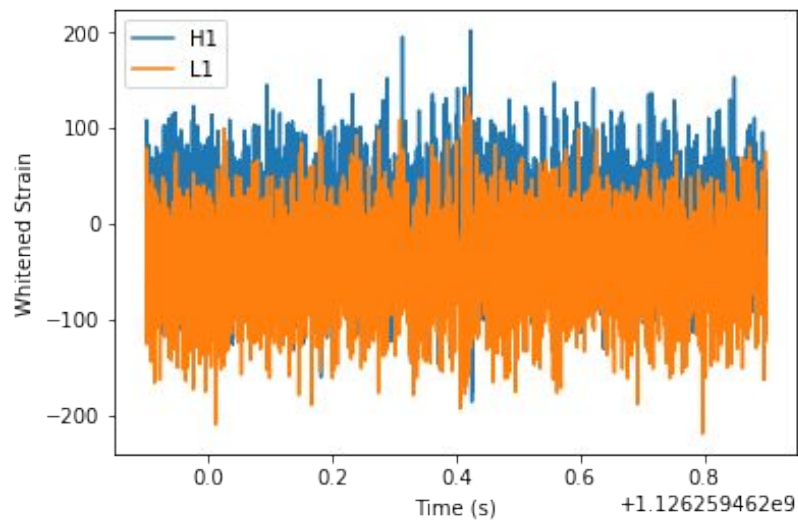
- $s(t) \rightarrow s(f)$ , divide by the ASD:  $s(f)/\sqrt{S_n(f)/2}$ , convert back to time-domain



<https://gwosc.org/workshops/>

# Band passing

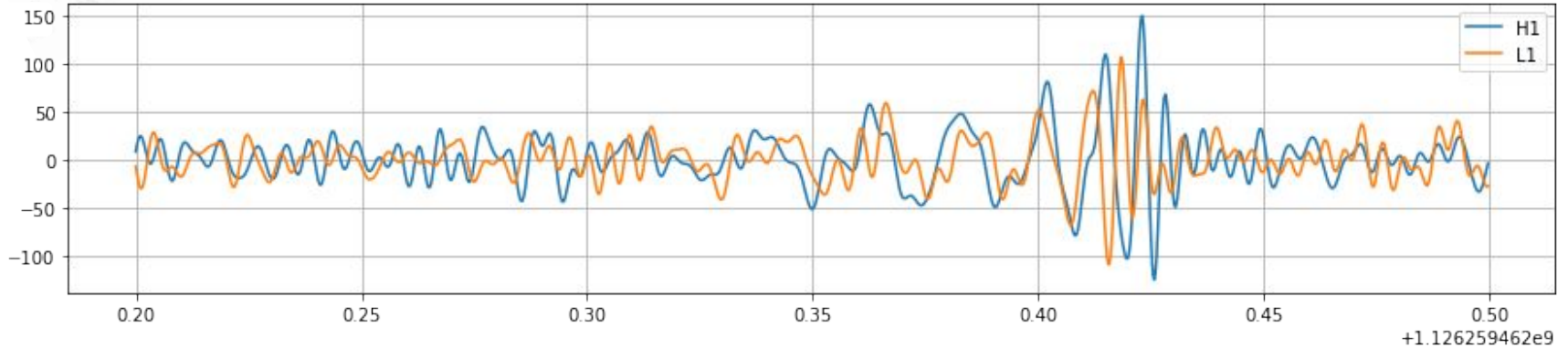
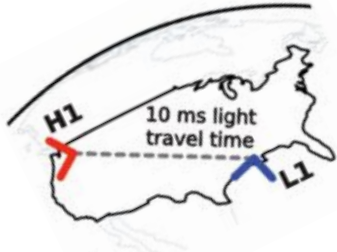
- Remove frequency where you do not expect any signals (keep 30 Hz - 250 Hz)



<https://gwosc.org/workshops/>

# Adjust for time of flight

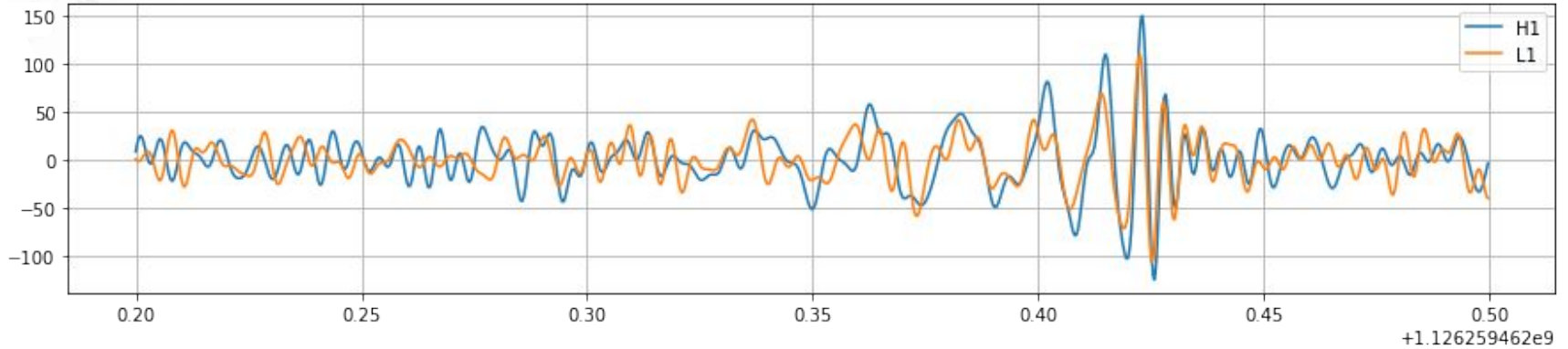
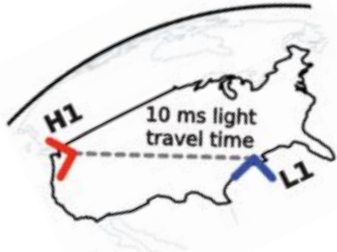
- For GW150914, the time of flight between H and L detectors was 7ms



<https://gwosc.org/workshops/>

# Adjust for time of flight

- For GW150914, the time of flight between H and L detectors was 7ms

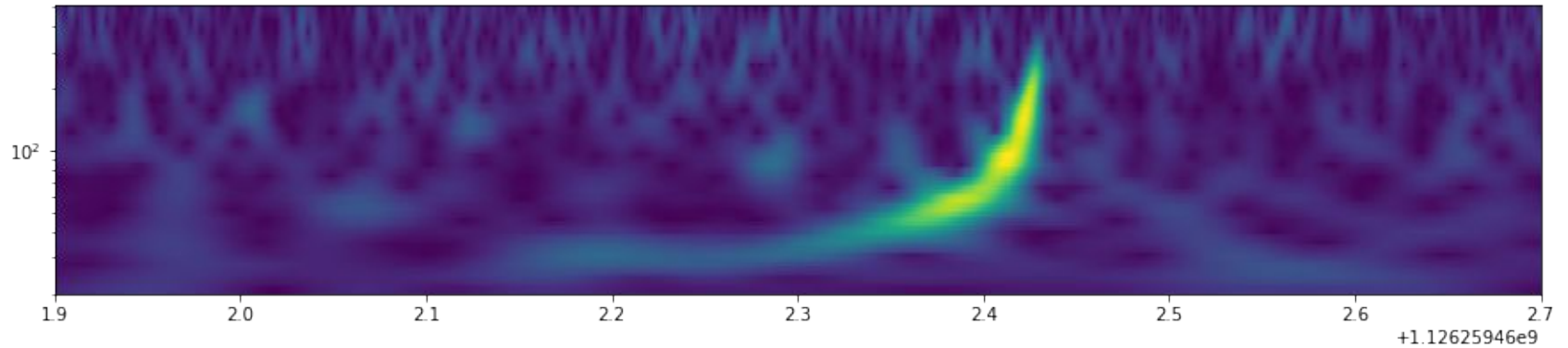


<https://gwosc.org/workshops/>

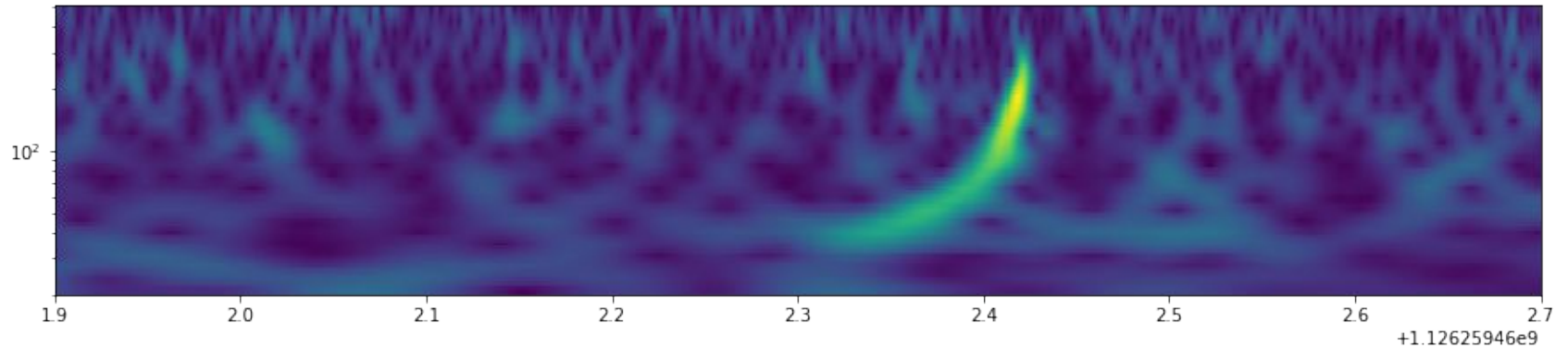
# Time-Frequency Spectrograms



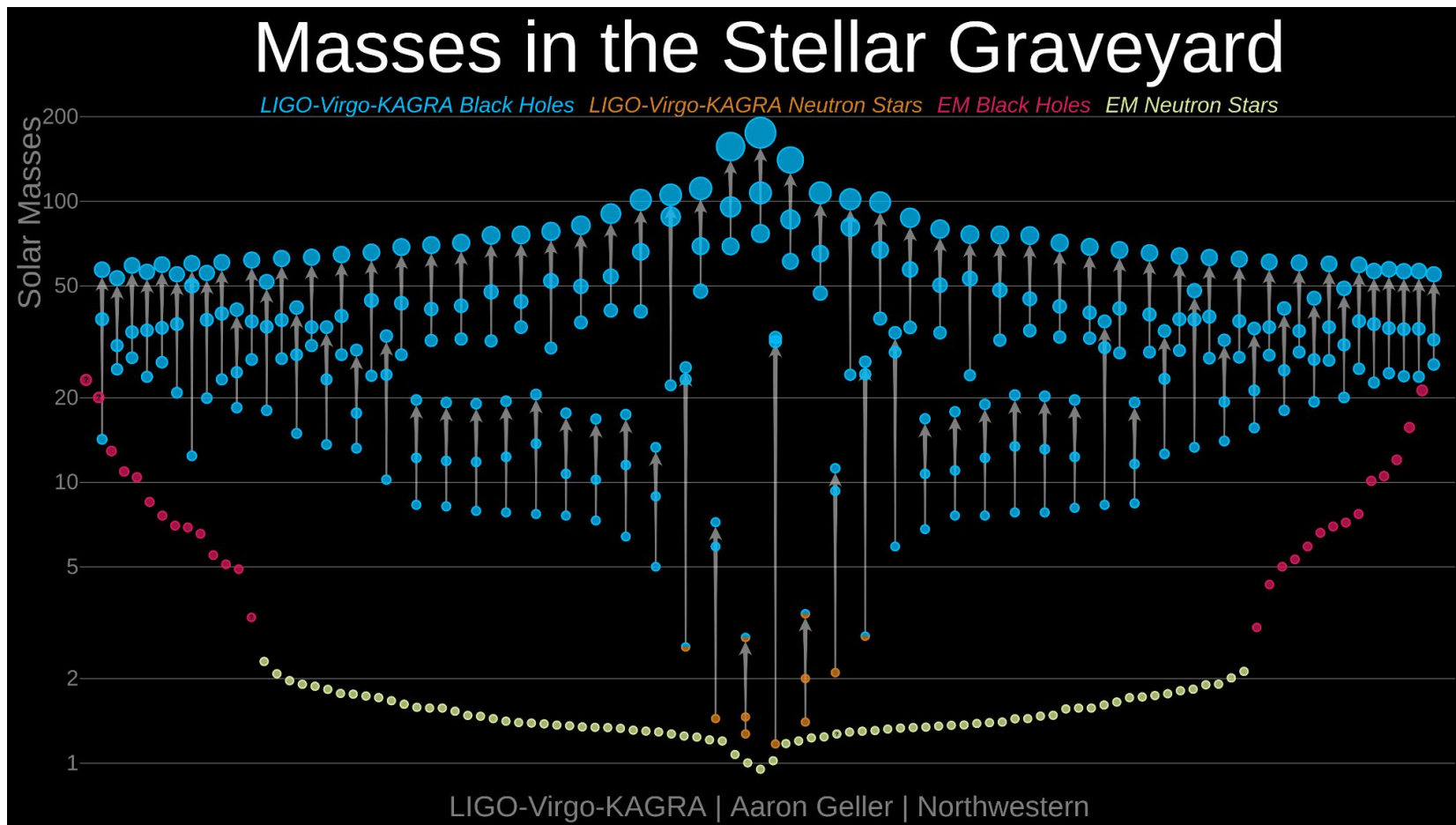
H1



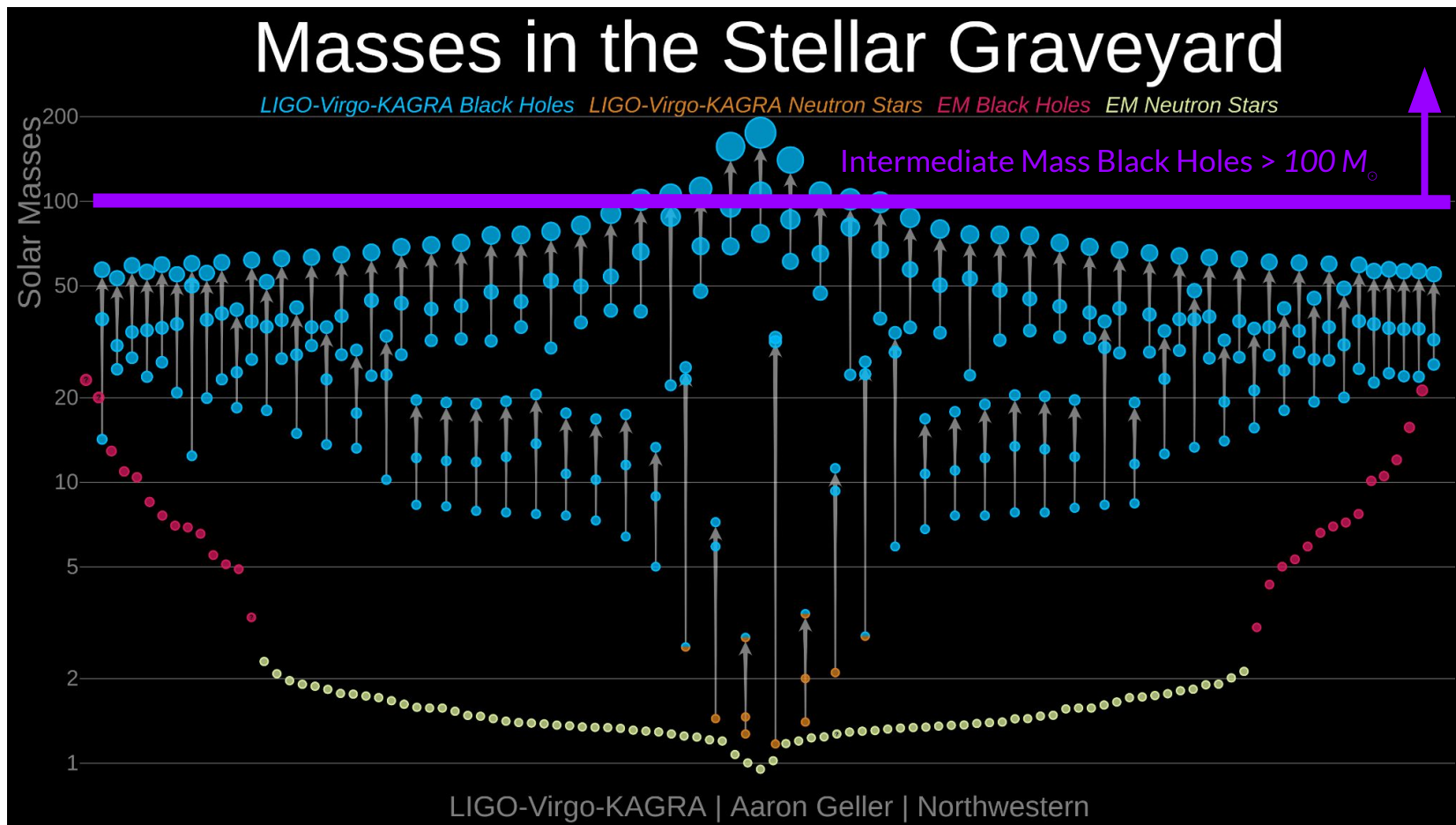
L1



# GW Detections from the First Three Observing Runs



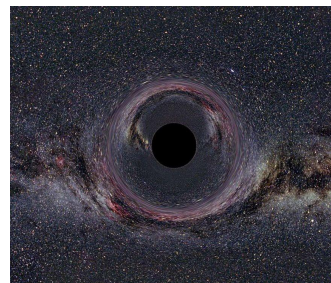
# GW Detections from the First Three Observing Runs



# Intermediate Mass Black Holes (IMBHs)

- Black holes (BH) with masses in the range  $100 M_{\odot}$  -  $10^5 M_{\odot}$ .
- Missing link between stellar mass ( $< 100 M_{\odot}$ ) and supermassive ( $> 10^5 M_{\odot}$ ) BH.
- Various IMBH candidates proposed from EM observations, empirical mass-scaling relations, hyper-luminous X-ray sources, but aren't conclusive.
- Most massive BH observed via GW in the third observing run (O3) of LIGO - **GW190521** event.
  - First conclusive evidence of IMBH  $< 10^3 M_{\odot}$ , with its total mass  $142 M_{\odot}$  (component masses  $85 M_{\odot}$  and  $66 M_{\odot}$ ).
- Primary component of GW190521 - interesting candidate with far-reaching astrophysical implications.

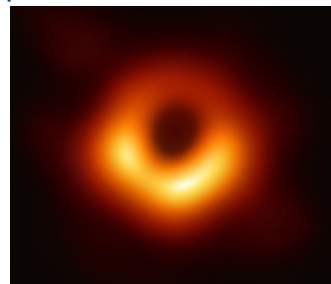
Stellar mass BH  $< 100 M_{\odot}$



[Ute Kraus/Wikipedia, CC BY-SA]

$100 M_{\odot} < \text{IMBH} < 10^5 M_{\odot}$

Supermassive BH  $> 10^5 M_{\odot}$



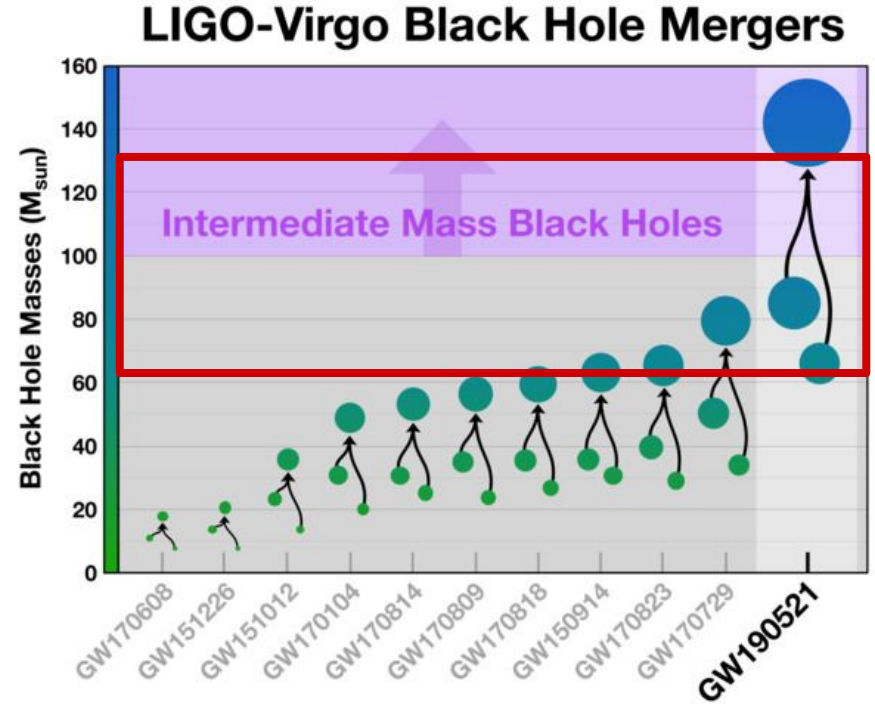
[Event Horizon Telescope]

# Pair Instability Mass Gap

Sufficiently heavy He core mass can provide suitable conditions for abundant pair production (electron-positron) which causes the star to violently implode.

- Stars with He core mass in ( $32 M_{\odot}$ ,  $64 M_{\odot}$ ) - pulsational pair instability supernova (PPSN) - remnant BH mass  $< 64 M_{\odot}$ .
- Stars with He core mass in ( $64 M_{\odot}$ ,  $135 M_{\odot}$ ) - pair instability supernova (PISN) - no compact remnant (PISN mass gap).
- Stars with He core mass  $> 135 M_{\odot}$  directly collapse to form an IMBH.

Primary component BH of **GW190521** lies in the **PISN mass gap** and can't be a SN remnant - raises questions on the possible formation channels.



Masses of the GW detections from first and second observing runs (O1 and O2) of LIGO + GW190521 event.  
*[LIGO/Caltech/MIT/R. Hurt (IPAC)].*

# Possible Formation Channels

- **Hierarchical mergers:**

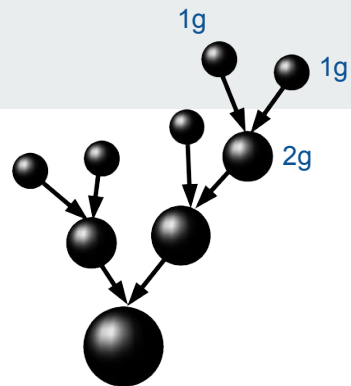
- Mergers involving one or two second generation (2g) BHs.
- First generation (1g) merger of BHs can produce a 2g BH in the PISN mass gap.
- Triple systems or dynamic capture in a dense cluster.

- **Active Galactic Nucleus (AGN) disk:**

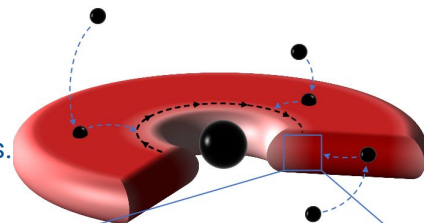
- Dense and hot environment hosting  $O(10,000)$  BHs.
- BBH mergers are expected to stay in the AGN disk and acquire another companion BH.

- **Stellar merger scenario:**

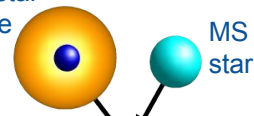
- Star with over-sized Hydrogen envelope.
- Direct collapse into BH in the PISN mass gap without encountering PISN/PPSN.
- Merger with companion main sequence star required to get over-sized H envelope.



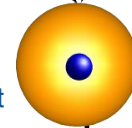
Different merger scenarios.  
[[GW190521 webinar](#)].



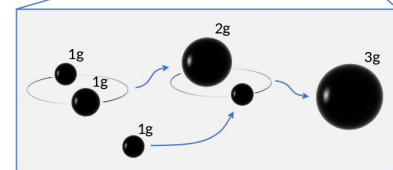
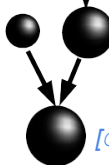
Giant star with He core  
MS star



Stellar merger product



Black hole in PI gap



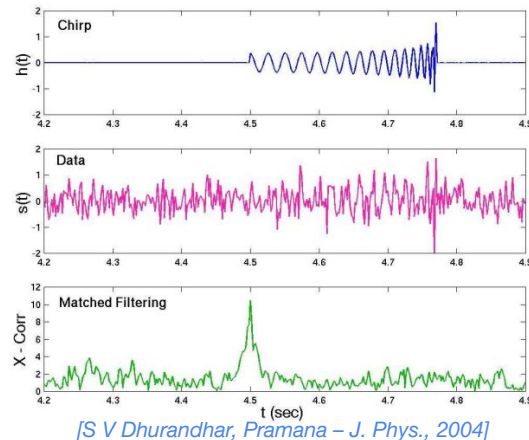
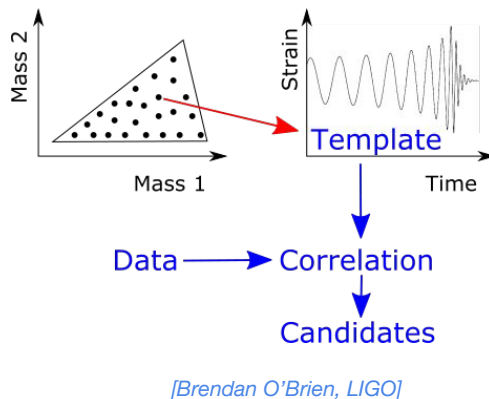
[Credit: Imre Bartos]

[Credit: Ugo N. Di Carlo]

# Search Types

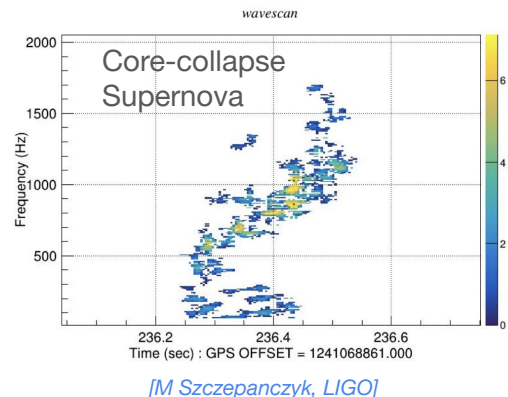
## ● Modeled Search:

- Looks for sources that are well modeled.
- Accurate waveforms are used to construct template banks e.g. CBC sources.
- Uses Matched Filtering, example - PyCBC, GstLAL.



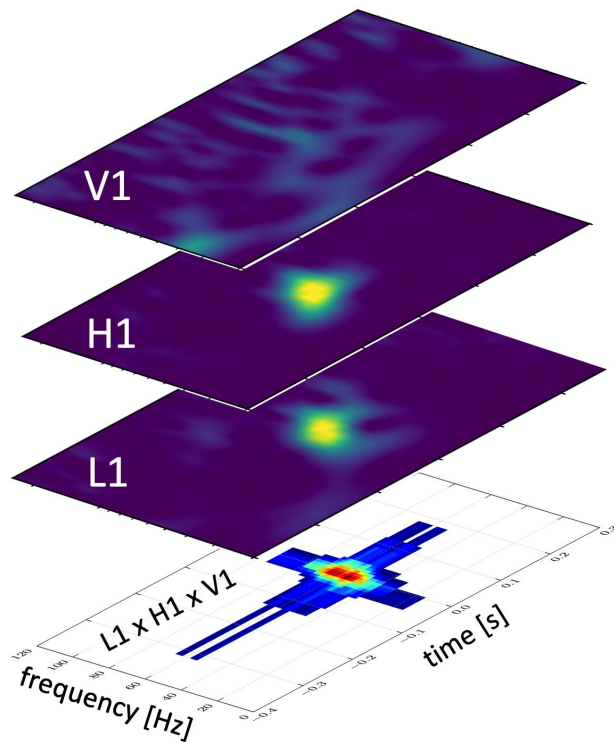
## ● Unmodeled Search:

- Looks for astrophysical sources not covered by template banks.
- Can **detect unmodeled sources** e.g. CCSN and poorly modeled CBC sources.
- Searches for excess power in the time-frequency domain, example - **Coherent WaveBurst (cWB)**.



# Coherent WaveBurst

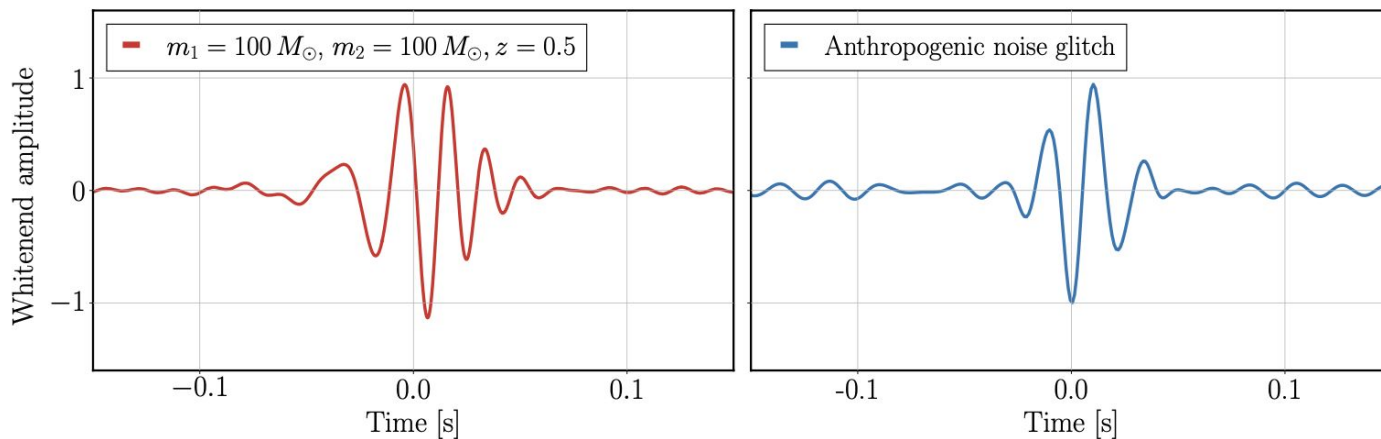
- cWB is a model-agnostic search algorithm [[S. Klimenko+ 2008](#), [S. Klimenko+ 2016](#)].
- cWB maps generic properties of GW events into summary statistics.
- GW signals can be mimicked by short duration detector glitches.
- The standard veto method uses a priori defined thresholds on the cWB summary statistics to distinguish between a GW signal and noise.



Excess power in time-frequency domain  
[[GW190521 webinar](#)]

# Detection Confidence

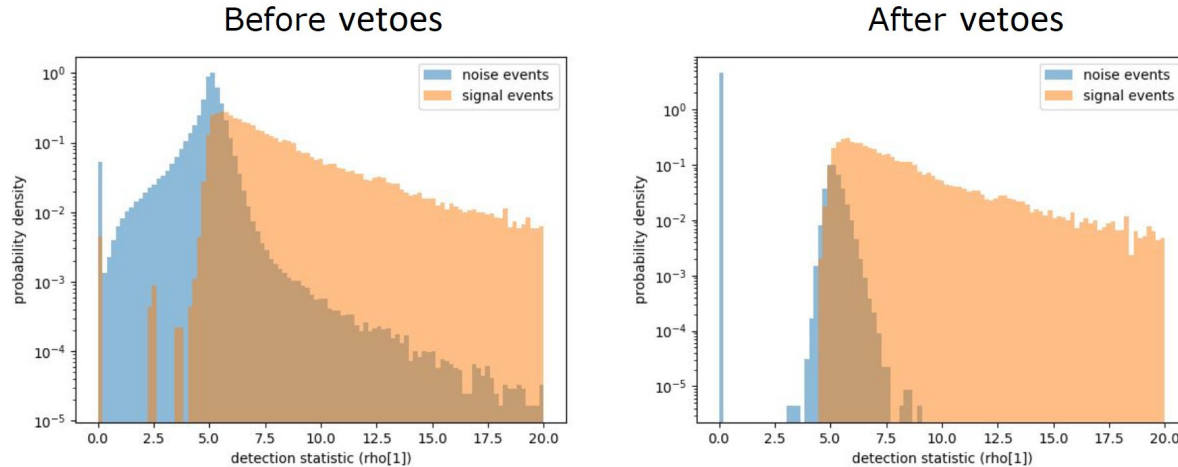
- The detection confidence can be compromised by the presence of non-Gaussian noise transients and instrumental noise known as **glitches**.



High mass BBH and anthropogenic noise glitch comparison [\[Brendan O'Brien, PhD Thesis\]](#).

- High mass BBH signals are of shorter duration and exhibit mostly the merger-ringdown parts in the LIGO sensitive frequency band.
- Lack of inspiral makes the signal prone to be mimicked by instrumental and other non-Gaussian glitches.

# Separation of Signal from Glitches



Standard veto method illustration [Brendan O'Brien, UF LIGO].

- All events below the a priori defined thresholds are discarded as noise. Example of a standard veto:  $\mathbf{c}_c > 0.7$  where  $\mathbf{c}_c$  is the network correlation coefficient, a summary statistic estimated by cWB.
- Designing the vetoes in the multidimensional summary statistics space is time consuming and complex.
- The vetoes need redefinition for each observing run and detector network.



# Enhancing cWB with Machine Learning

[T. Mishra+ 2021](#) (Methodology, O1+O2 observing run results)

[T. Mishra+ 2022](#) (Search for binary mergers in O3 observing run)

[M. Szczepanczyk+ 2023](#) (Search for generic GW bursts in O3)

**Documentation:**

<https://gwburst.gitlab.io/documentation/latest/html/xgboost.html>

# Enhancing cWB with Machine Learning

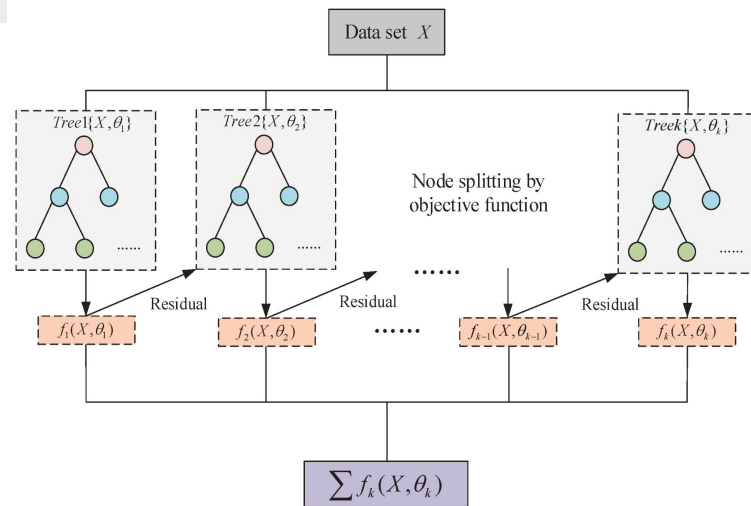
- Standard veto method has been replaced by an automated ML method that provides improved separation between GW signals and glitches.
- A subset of summary statistics is used as input features into the ML algorithm **XGBoost** - a boosted decision-tree based ensemble learning classifier algorithm.
- The XGBoost hyper-parameters are optimized to avoid over-fitting. An  $\eta$ -dependent sample weight is introduced to reduce class imbalance.
- Detection statistic used:

- Production - 
$$\eta_0 = \sqrt{\frac{E_c}{1 + \chi^2(\max(1, \chi^2) - 1)}}$$

- Post Production - 
$$\eta_r = \eta_0 \cdot W_{\text{XGB}}$$

;  $E_c$  = coherent energy,  $\chi^2$  = reduced chi-sq statistic  
( $E_{\text{res}}/N_{\text{d.o.f.}}$ )

;  $W_{\text{XGB}}$  is the penalty factor provided by the ML algorithm [[T. Mishra+ 2021](#), [T. Mishra+ 2022](#)].




XGBoost flow chart for building an ensemble of trees. [[Rui Guo+ 2020](#)]

# XGBoost Update: New summary statistics

```
ML_list = ['netNoise', 'norm', 'chi2', 'netCC', 'corrED', 'pixden', 'superXP', 'XP/Xcorr', 'max_snr_ratio',  
          'anet', 'gnet', 'netcc1', 'TFvolumelikelihood', 'size1', 'frequency0', 'bandwidth0', 'duration0',  
          'sqrt_ecor/likelihood', 'chirp6', 'chirp1', 'chirp3', 'rhoG_20d0', 'Qa', 'Qp']
```

- netNoise : normalized network noise
- norm : wavescan oversampling factor
- chi2 : residual noise / Gaussian noise
- netCC : network correlation (netcc[4])
- corrED : correlation energy disbalance (netcc[5])
- pixden : event pixel density (netcc[6])
- superXP : event super-coherence (netcc[7])
- XP/Xcorr : ratio of crosspower and crosscorrelation ( $\rho[6]/\rho[1]$ )
- max\_snr\_ratio : network snr asymmetry ( $(nifo * \max(\text{snr}[0], \text{snr}[1]) / (\text{snr}[0] + \text{snr}[1])) - 1$ )
- TFvolumelikelihood : size in TF map ( $(\text{bandwidth}[1] * \text{duration}[1]) / (\text{volume}[0] / \text{likelihood})$ )

# XGBoost Regression: Intrinsic hyperparameters



XGBoost hyper-parameter	entry
objective	reg:pseudohubererror
tree_method	hist
grow_policy	lossguide
n_estimators	20,000†
max_depth	<b>8</b>
learning_rate	<b>0.03</b>
min_child_weight	<b>10.0</b>
colsample_bytree	<b>1.0</b>
subsample	<b>0.8</b>
gamma	<b>5.0</b>
base_score	$\overline{M}_{inj}$

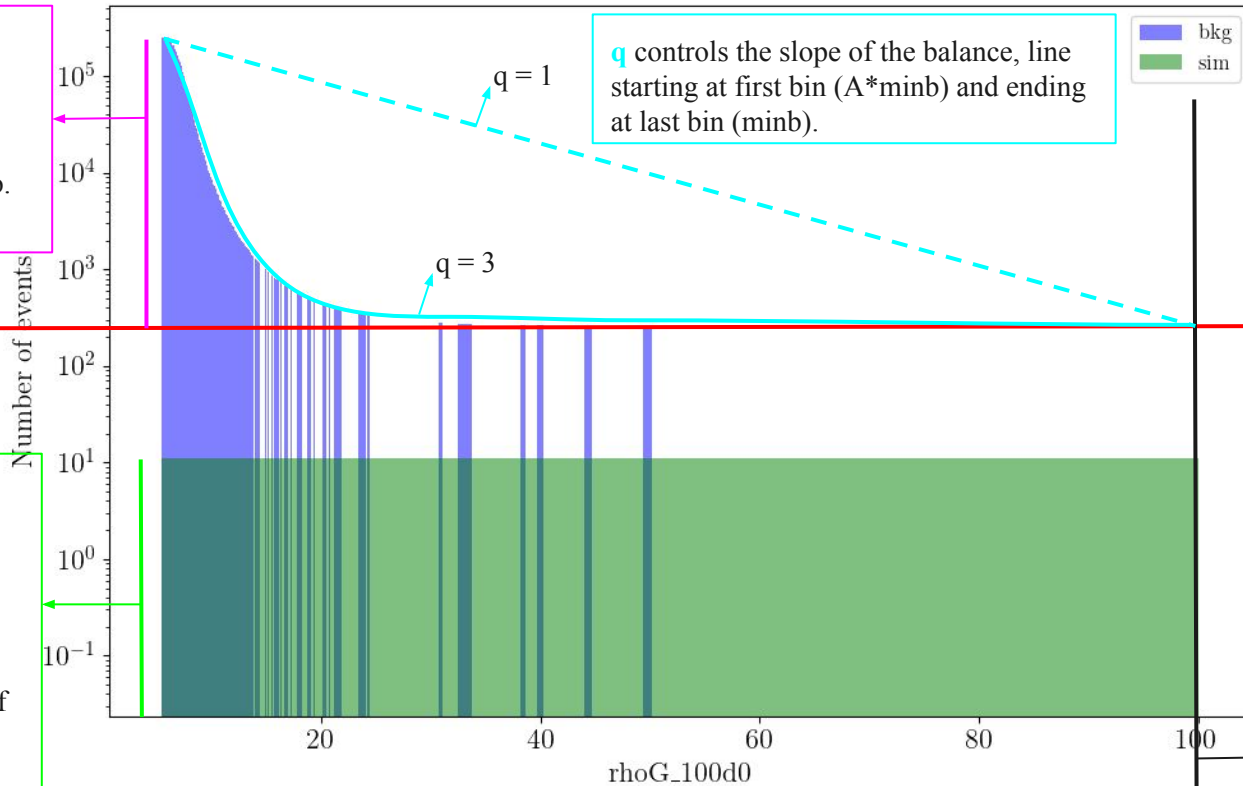
# XGB Training: User-defined custom hyperparameters

## Background/Simulation Balanced Distributions

**A** controls the class balance in first bin w.r.t. minb value .. final weighted BKG hist starts at  $A \cdot \text{minb}$ .  
[A = 1000]

**q** controls the slope of the balance, line starting at first bin ( $A \cdot \text{minb}$ ) and ending at last bin ( $\text{minb}$ ).

bkg  
sim

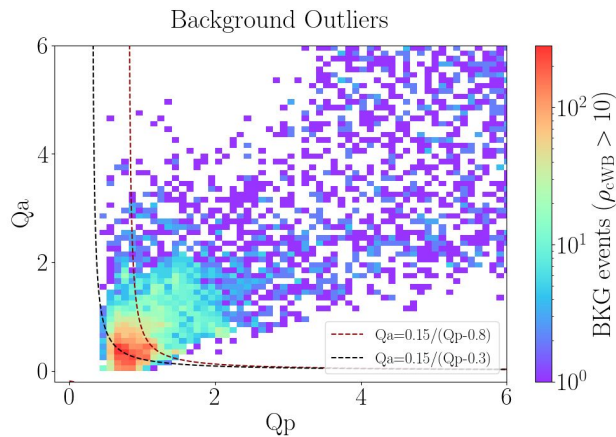
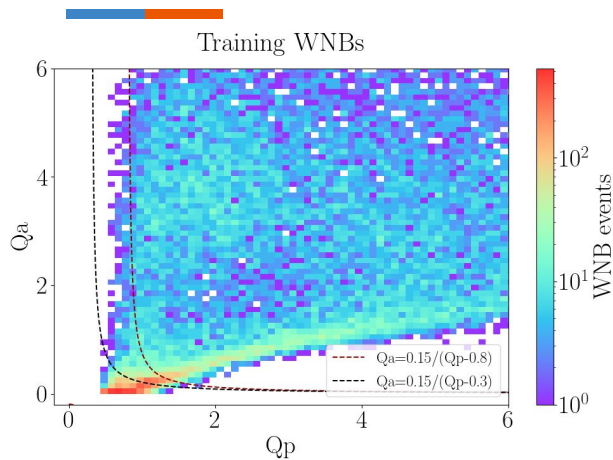


final weighted BKG hist saturates at the value set by **minb**.  
[minb = 250]

sets the limit on the input rho parameter in ML\_list **rho\_cap**.  
[rho\_cap = 100]

percentile binning is done based on input SIM events. **spw** assigns an overall weight to each SIM event.. final weighted SIM hist has a value of  $\text{spw} \cdot \text{initial}$ .  
[spw = 0.05]

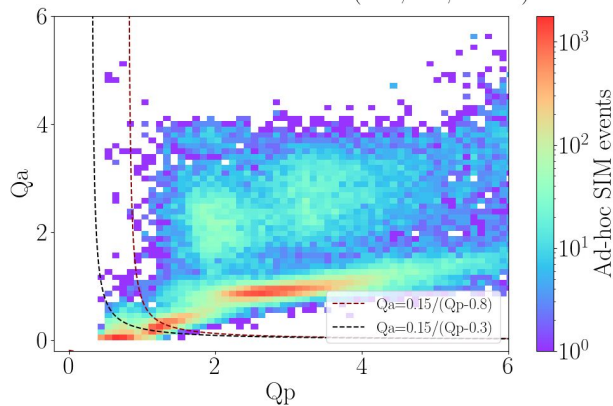
# XGB Training: Qa-Qp (plots from AllSky search)



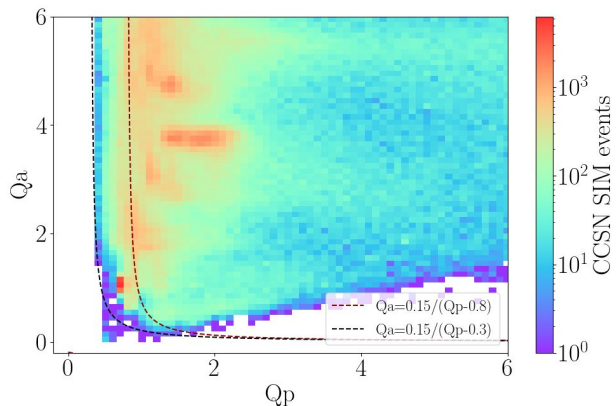
## Qa-Qp correction

$$P_{XGB} = \begin{cases} P_{XGB} - (0.15 - Q_a(Q_p - 0.8)), & \text{if } Q_a(Q_p - 0.8) \leq 0.15 \\ & \text{(under the curve)} \\ P_{XGB}, & \text{if } Q_a(Q_p - 0.8) > 0.15 \\ & \text{(above the curve).} \end{cases}$$

Ad-hoc Standard Simulations (GA, SG, WNB)



CCSN Simulations

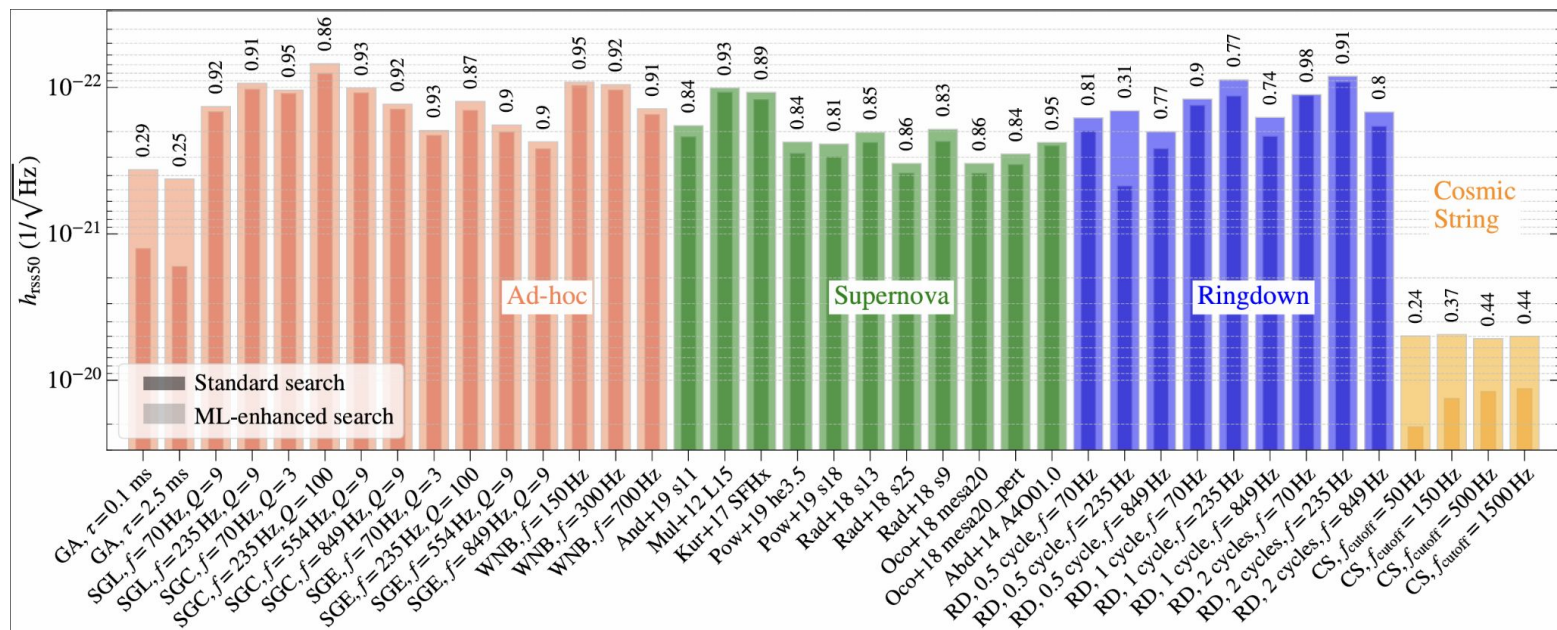


## Monotonic transformation

$$W_{XGB}(P_{XGB}) = \frac{-\log(1.0 - 0.995\sqrt{P_{XGB}})}{5.3}$$

# Generic GW bursts search with ML-enhanced cWB

- Not rely on astrophysical waveform models to train XGBoost:
  - Trained using ad-hoc band-limited White-Noise-Bursts (WNB) to maintain the model independent nature of cWB.



$h_{rss50}$  at IFAR > 100 yr for simulations representing GW bursts.

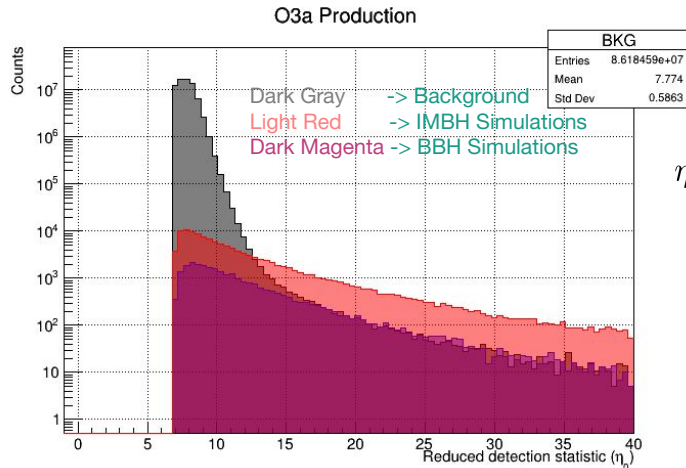
# Targeted XGBoost model for BBH/IMBH mergers

- XGBoost model trained for BBH and IMBH searches.
- Training data used per chunk:
  - BKG (Background - 100 yr accumulated by time-shifting detector data)
  - SIM (Simulations - 1000 events from stellar-mass: BBH search OR Intermediate-mass injections: IMBH search)
- False Alarm Rate (FAR) is assigned based on reranked BKG data.

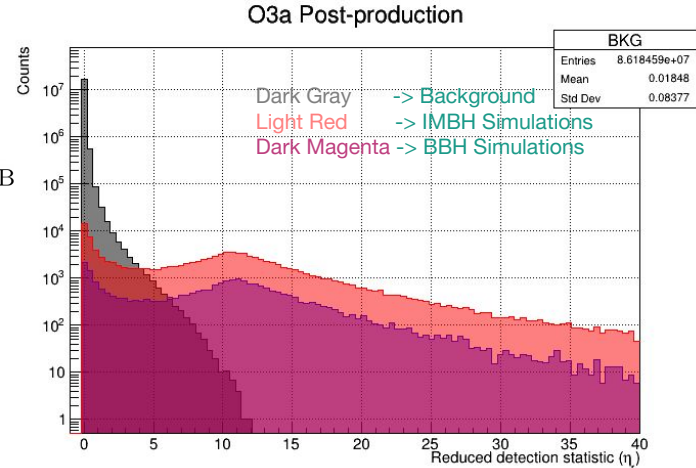
FAR (False Alarm Rate) w.r.t. test statistic  $\eta$  :

$$\text{FAR}(\eta) = \frac{N_{\text{bkg}}(\eta)}{T_{\text{bkg}}}$$

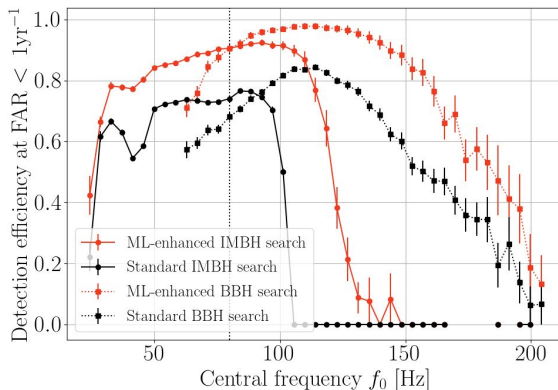
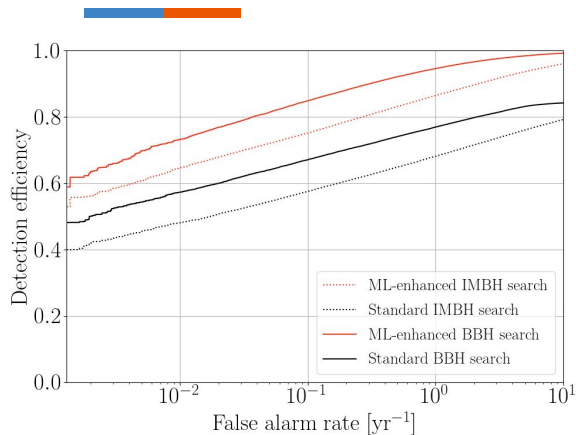
where  $N_{\text{bkg}}(\eta)$  is the number of reconstructed background events with test statistic greater than the threshold  $\eta$ , and  $T_{\text{bkg}}$  is the total time (in years) of background data.



$$\eta_r = \eta_0 \cdot W_{\text{XGB}}$$

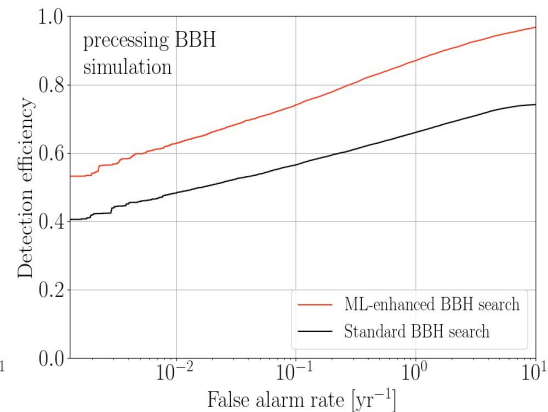
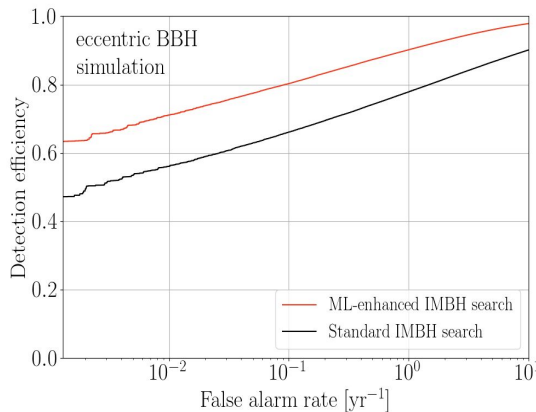


# ML-enhanced cWB search results



- ML-enhanced cWB searches improve detection efficiency of both BBH and IMBH simulations across the frequency band.

- XGBoost models are **robust** in detecting BBH signals originating beyond the training simulation set.





# cWB upgrade for the Fourth Observing Run of LIGO-Virgo-KAGRA

[T. Mishra+ 2025](#) (Updated cWB search, O3 reanalysis results)

Collaboration paper contributions in the ongoing O4 observing run:

*GWTC-4* (Catalog of GW events - data products + simulation studies)

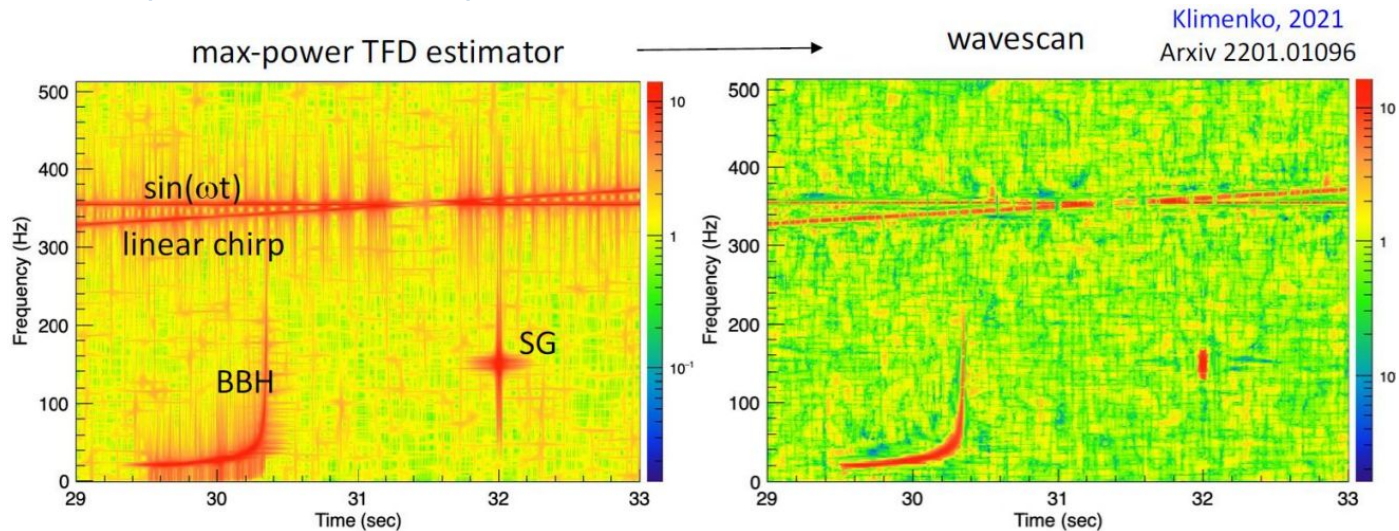
*O4a All-Sky search* (Generic GW bursts search - end-to-end cWB analysis+results)

*SN 2023ixf paper* [arXiv:2410.16565](https://arxiv.org/abs/2410.16565) (Targeted CCSN search - developed and ran the search)

*Exceptional event paper* (Follow-up of exceptional event detected by cWB)

# cWB upgrade for O4 - WaveScan

- New time-frequency (TF) transform called wavescan [[S Klimenko 2022](#)].
  - High-resolution TF maps with suppressed temporal and spectral leakage.
- Both cross-power and excess-power statistics used for efficient selection of transient events.

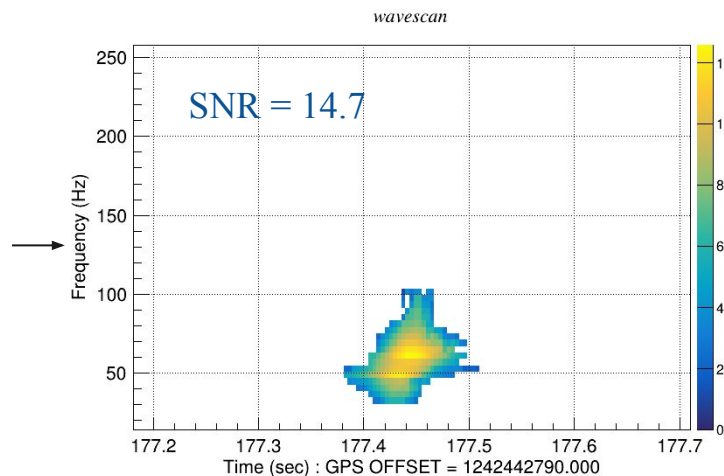
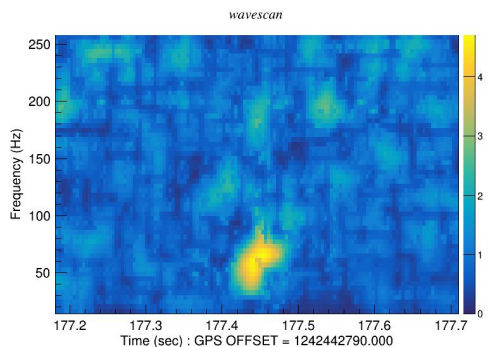
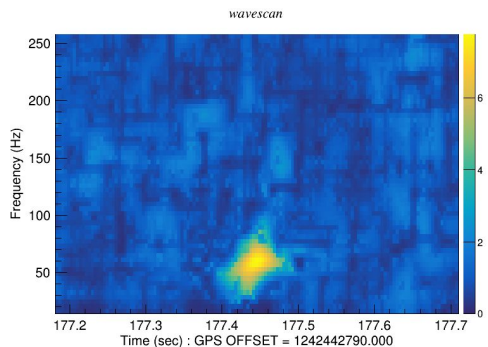


- Single XGB model is trained to identify both stellar mass and intermediate mass binary black hole (BBH) mergers: Combined BBH-IMBH search.

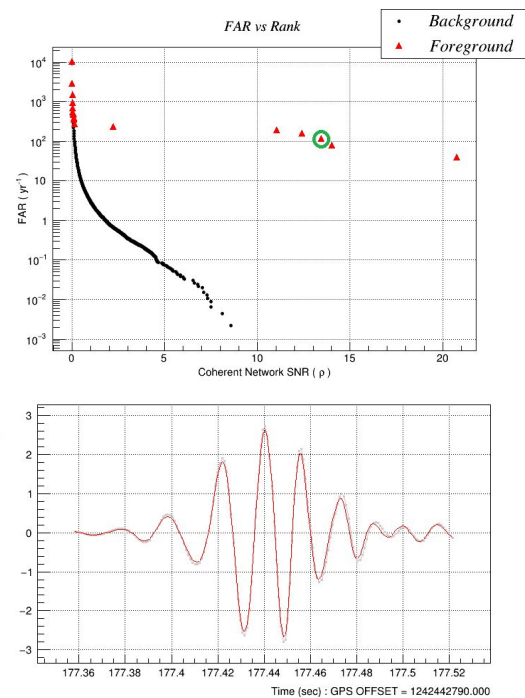
# GW190521 IMBH event

- First conclusive evidence of IMBH.
  - Detected by **cWB** with **highest significance** as compared to template-based searches.

L1

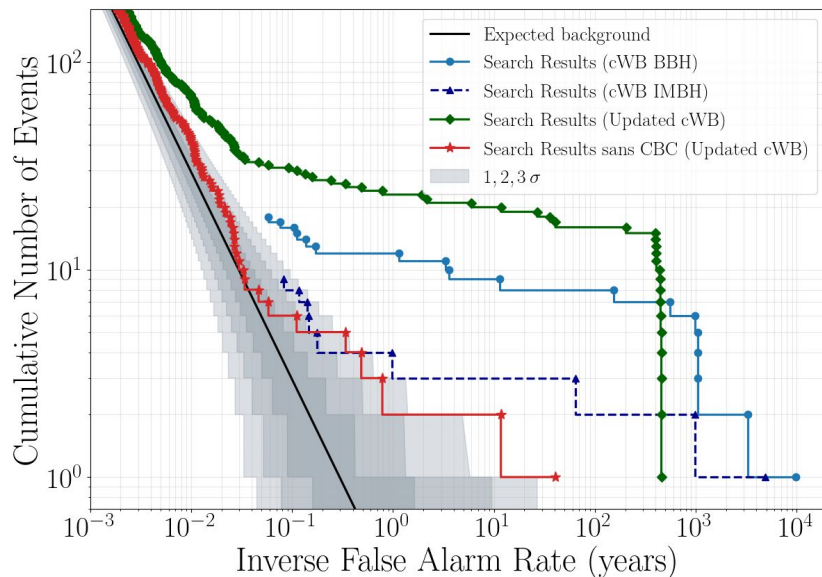
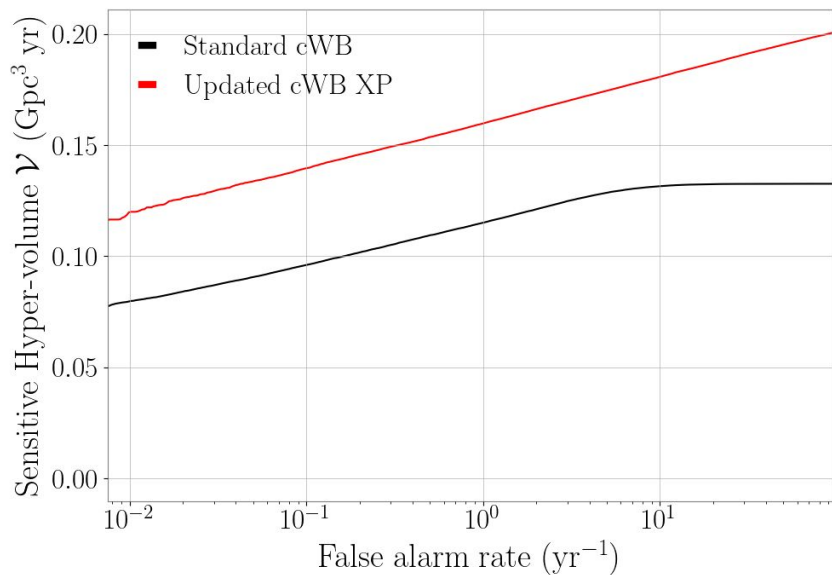


H1



# Search Results on O3

*T. Mishra+ 2025*  
(Phys. Rev. D 111, 023054)



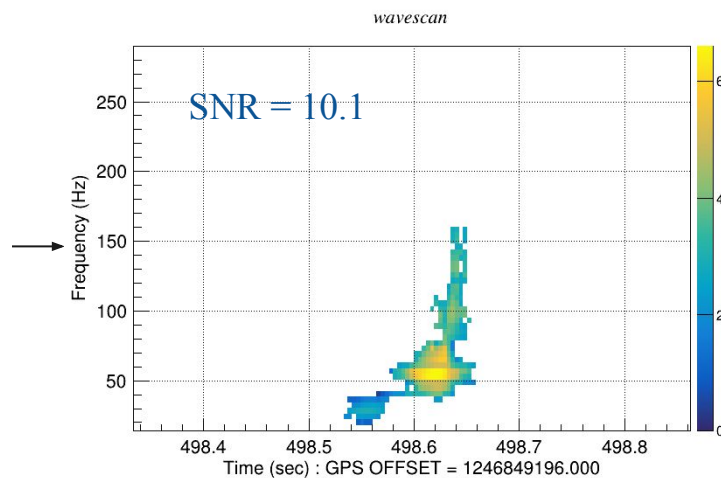
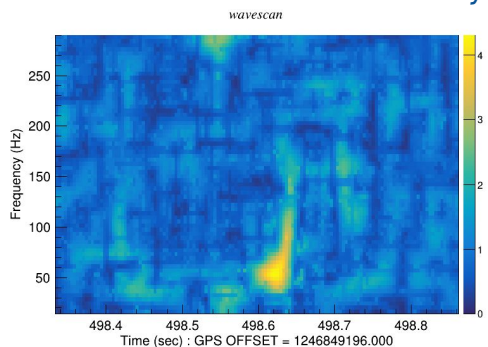
- Improved sensitivity by approximately **~40% at IFAR > 1yr** for all BBH simulations as compared to the standard cWB search.
- The updated cWB search detects 33 events present in GWTC-3 with equivalent or higher significance; 9 of these GW events were previously missed by the standard cWB search.
- **3 new cWB-only events** detected with IFAR > 1yr with **combined significance of 3.6  $\sigma$** .

# cWB-only event: 190711

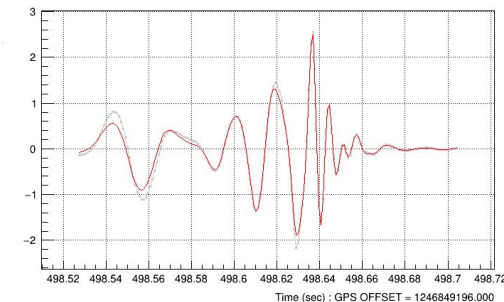
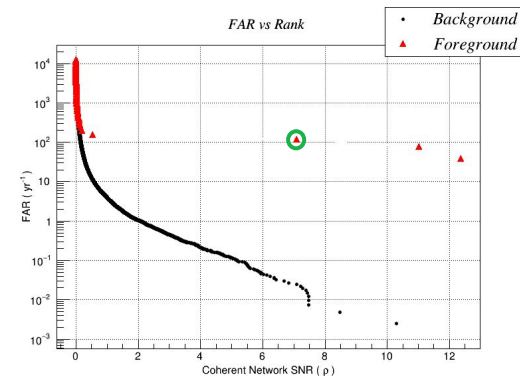
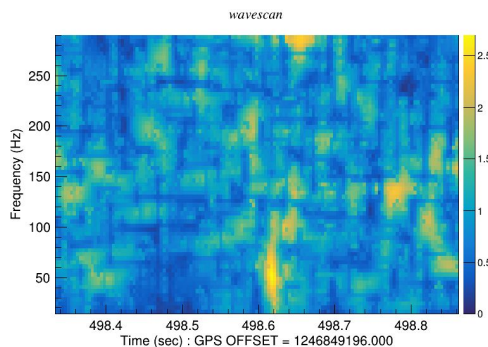
T. Mishra+ 2025  
(Phys. Rev. D 111, 023054)

- **IFAR = 40.7 yr** [*pastro* = 0.992] by the O3a Offline cWB XP search. GPS = 1246849694.62
  - Event not present in GWTC-3.
  - Found in the low latency search by GstLAL with FAR  $\sim 2/\text{yr}$  - [S190711ai](#)

L1



H1



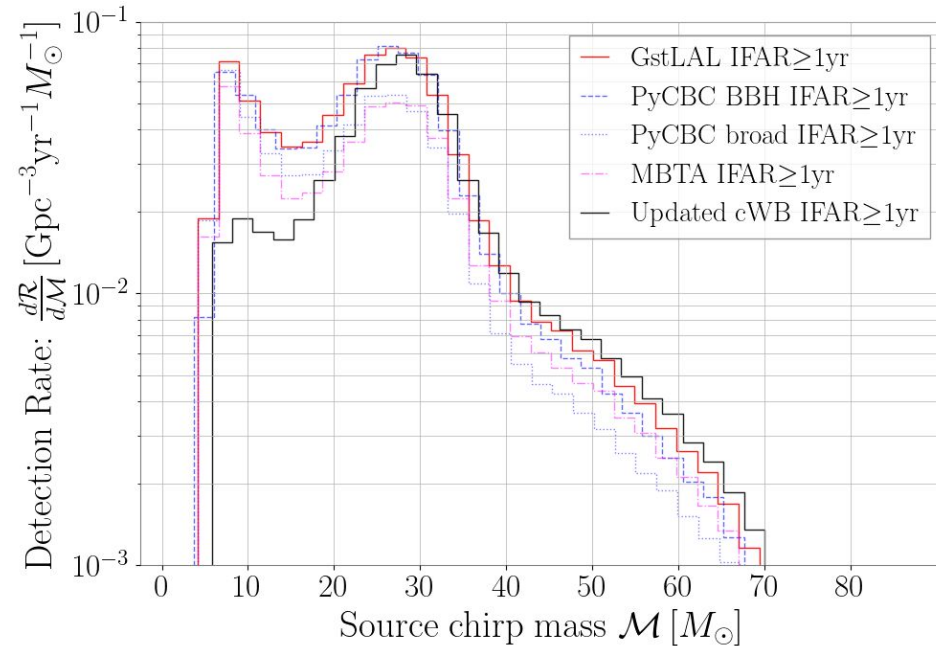
$m_1 \sim 70 M_\odot$  : suggests a BH in the PISN mass gap

$q \sim 0.3$  : potentially dynamic/AGN assisted formation scenario

# Explanation of cWB-only events

*T. Mishra+ 2025*  
(Phys. Rev. D 111, 023054)

- Consider reweighted simulations following the **Power-Law Peak (PLP)** pop model from GWTC-3.
- O3 CBC pipelines considered in this simulation study:
  - GstLAL, PyCBC (BBH, broad), MBTA
- CBC searches perform best for the low chirp mass events, while the updated cWB search start to surpass the CBC detection efficiency for  $\mathcal{M} > 30 M_{\odot}$ .
- **cWB detects 7.4%** of the stellar-mass BBH mergers missed by other CBC searches.
- O3 detections from GWTC-3:
  - 65 detections with IFAR>1yr.
- Expected cWB-only detections **4.8 +/- 2.1**



We observe 3 cWB-only events in O3 - consistent with the expectation.

# Fourth Observing Run (O4)

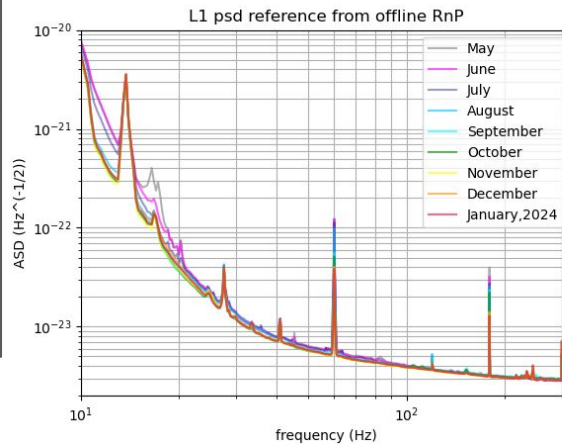
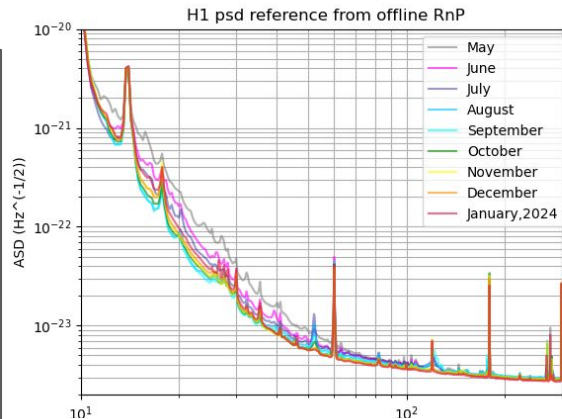


- The updated cWB was used in the low latency searches during the first half of the fourth observing run (O4a) [May 2023 - Jan 2024]. In charge of the cWB BBH-IMBH search in O4.
  - Out of the 81 significant detection candidates in O4a [[Public Alerts GraceDB](#)], the Online cWB search detected 57 (70% of all significant detections).
- Offline analysis for O4a is complete. O4b data release soon. O4c underway.

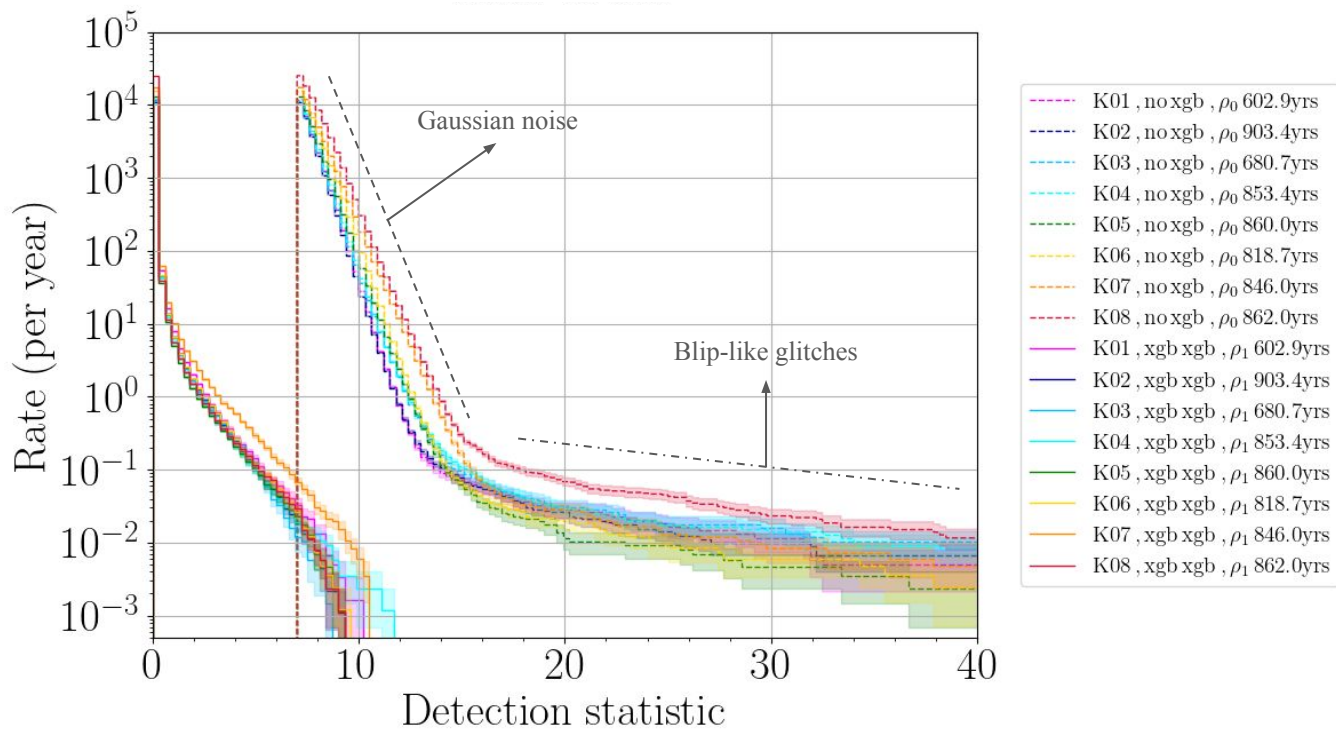
# Chunks in O4a



		May
12.78 days	Chunk K01	June
17.47 days	Chunk K02	July
13.88 days	Chunk K03	August
18.32 days	Chunk K04	September
18.43 days	Chunk K05	October
14.52 days	Chunk K06	November
12.72 days	Chunk K07	December
15.36 days	Chunk K08	January (2024)



# Offline BBH-IMBH search analysis



# Sensitive Volume Time

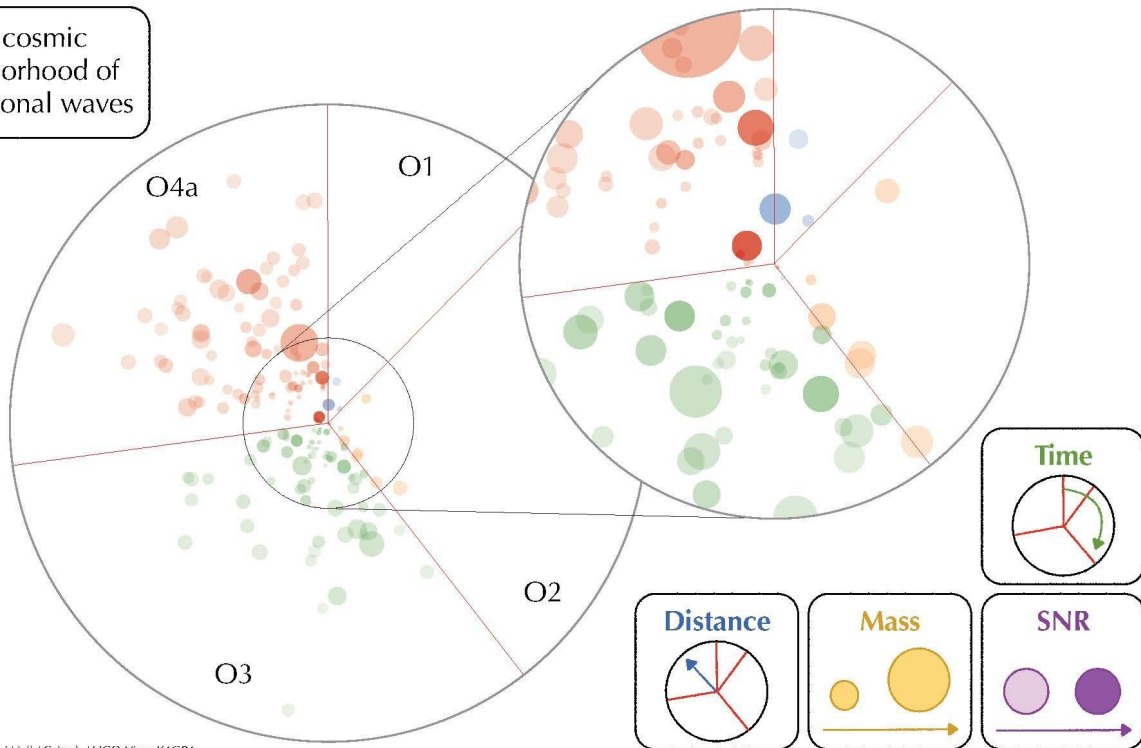
- To estimate the rate of detected mergers, we estimate the sensitive hypervolume spanned by the search

$$\langle VT \rangle_{sen} \sim \frac{N_{rec}}{N_{tot}} \langle VT \rangle_{tot}$$

$$V_{sen} = \frac{\langle VT \rangle_{sen}}{T_a}$$

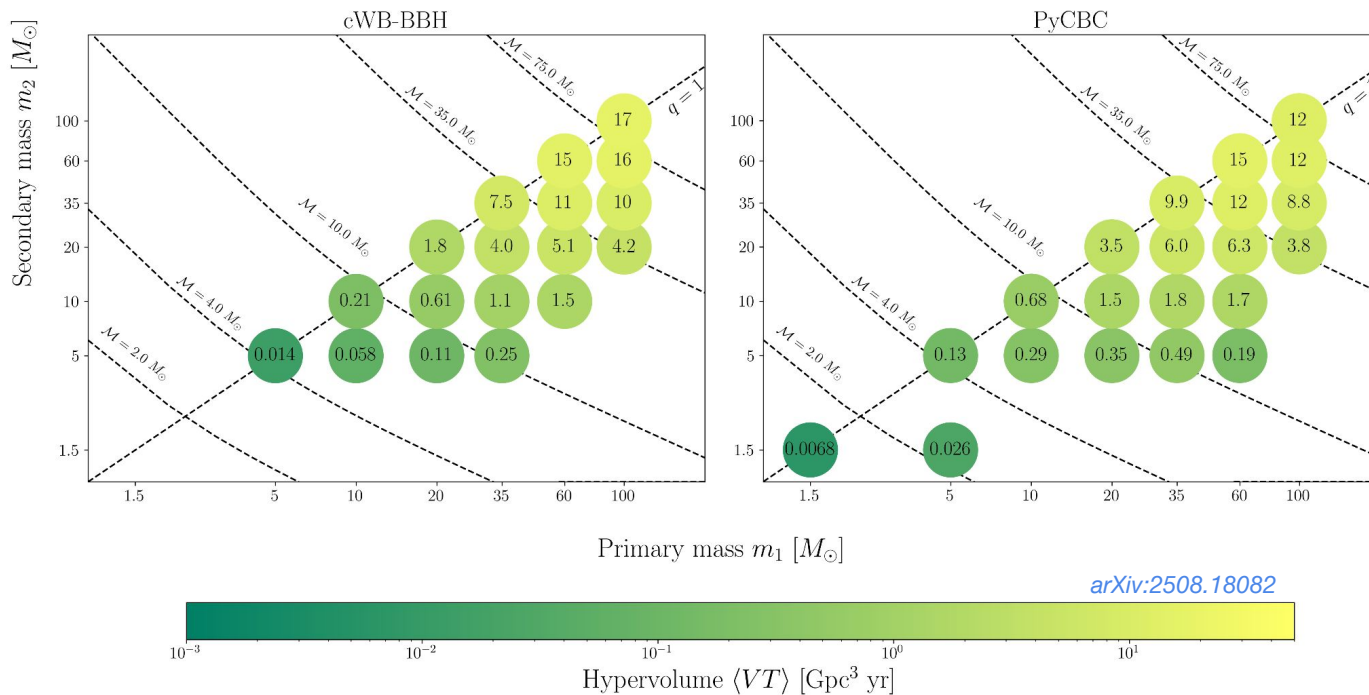
$$D_{\langle VT \rangle_{sen}} = \left( \frac{3V_{sen}}{4\pi} \right)^{\frac{1}{3}}$$

The cosmic neighborhood of gravitational waves



# Fourth Observing Run (O4a)

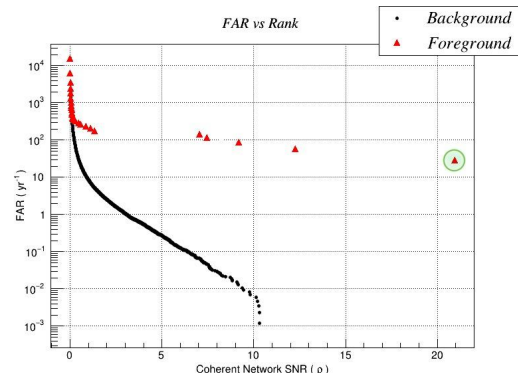
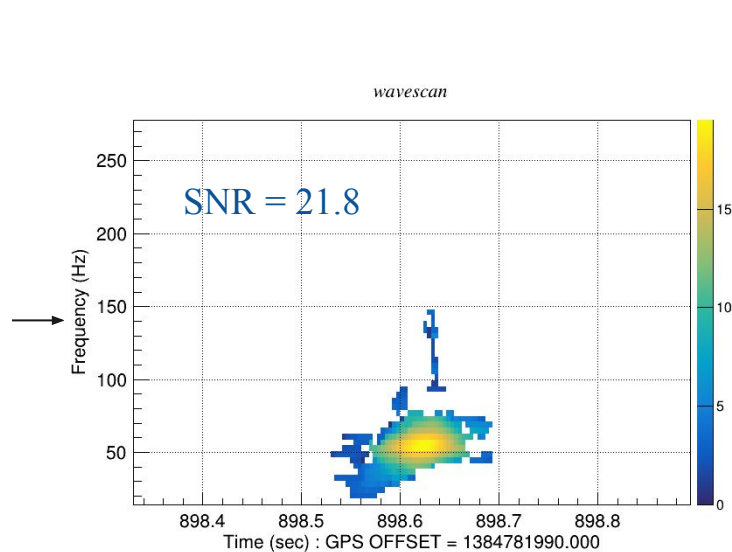
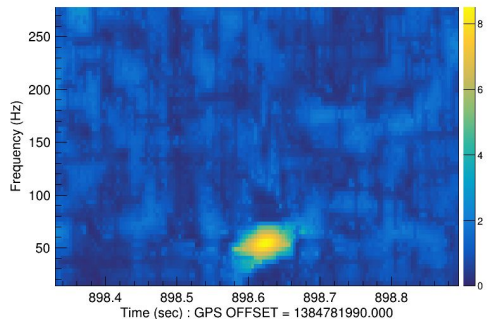
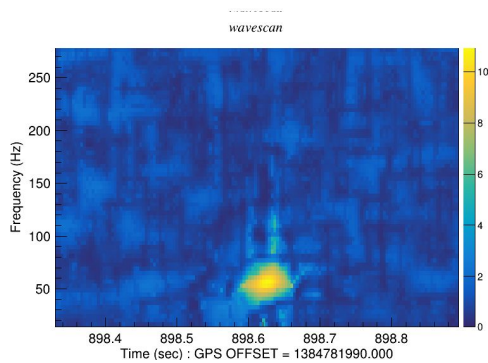
- The updated cWB was used in the low latency searches during the first half of O4 [May 2023 - Jan 2024].
  - Out of the 81 significant detection candidates in O4a [[Public Alerts GraceDB](#)], the Online cWB search detected 57 (70% of all significant detections).



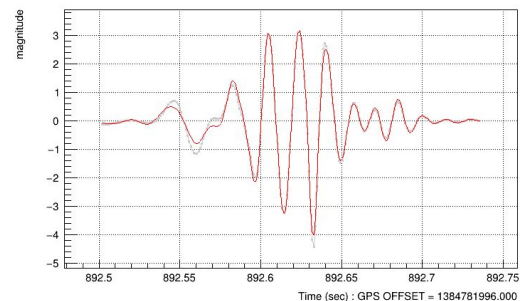
# GW231123 event

- Heaviest black hole merger detected by LVK so far.
  - Both mass components lie in the PISN mass gap. **cWB IFAR~9700 years**, **PyCBC IFAR~160 years**.

L1



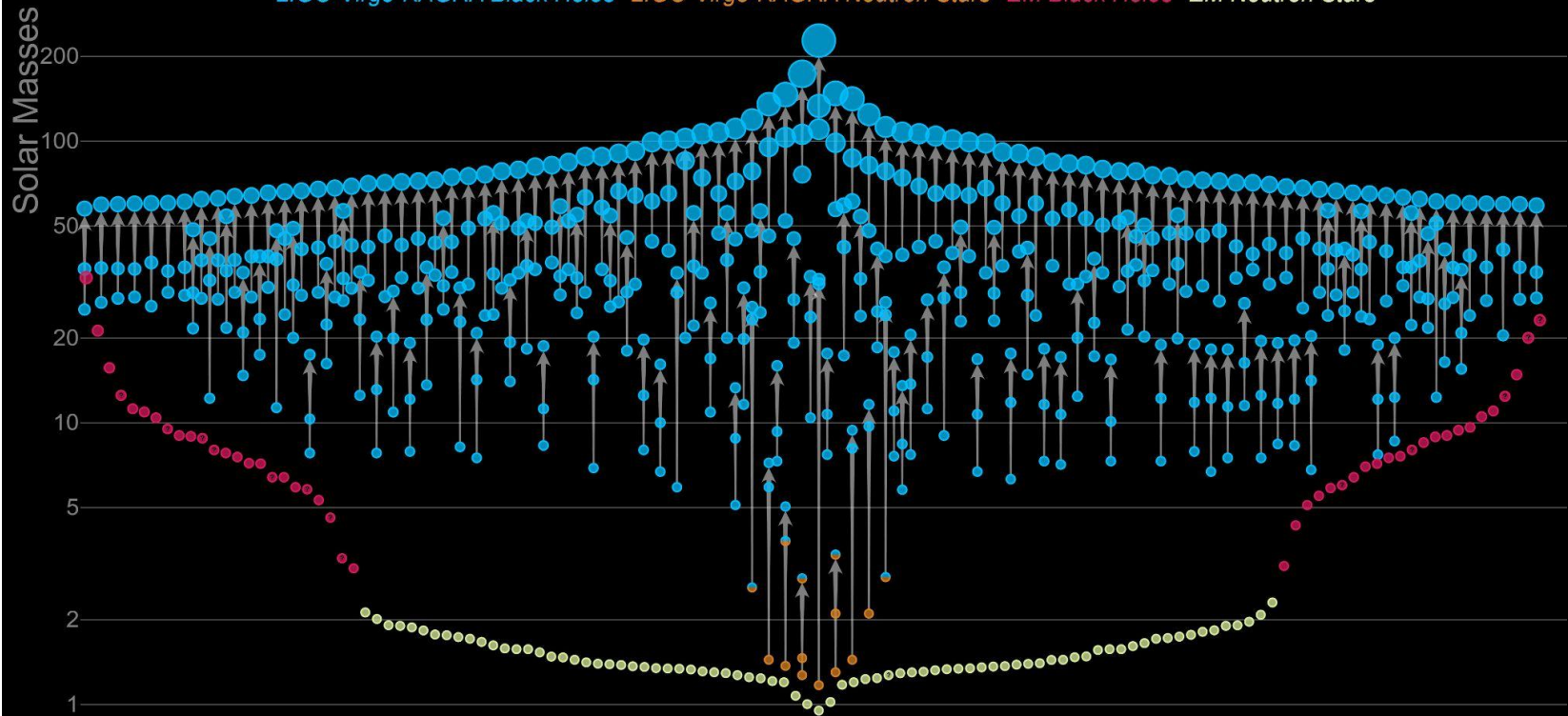
H1



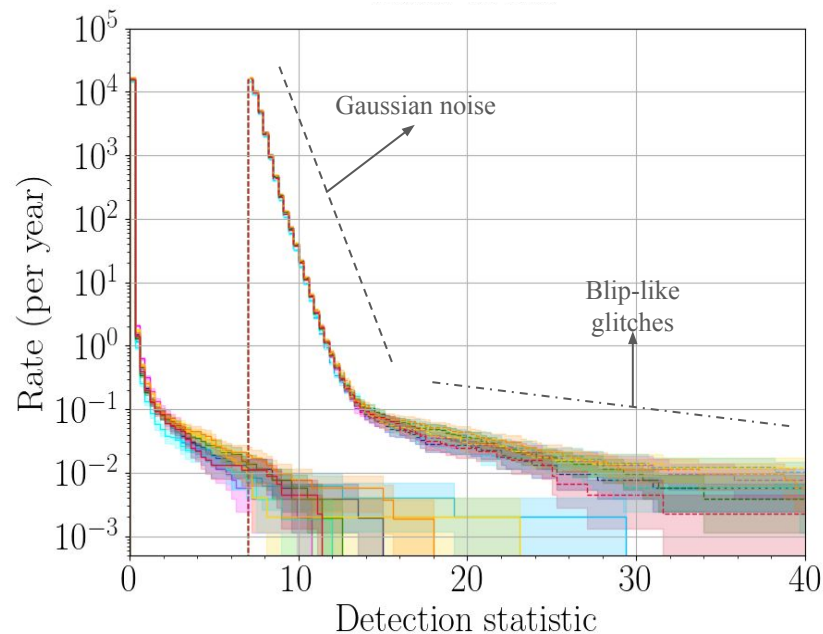
# Black hole mass spectrum - O4a

## Masses in the Stellar Graveyard

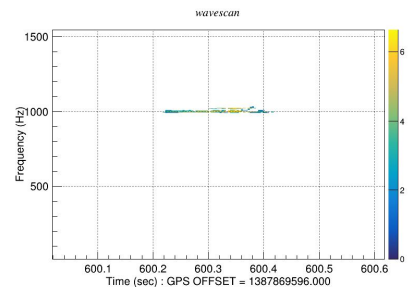
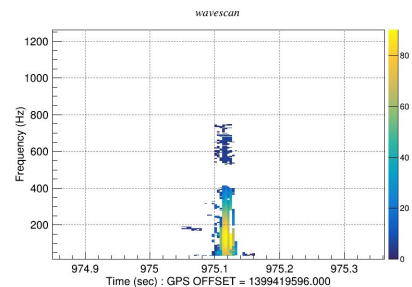
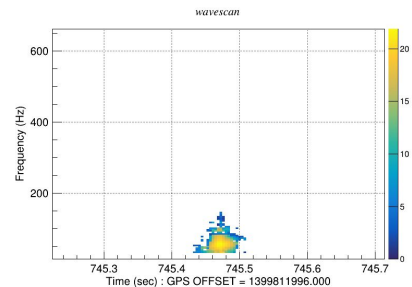
LIGO-Virgo-KAGRA Black Holes LIGO-Virgo-KAGRA Neutron Stars EM Black Holes EM Neutron Stars



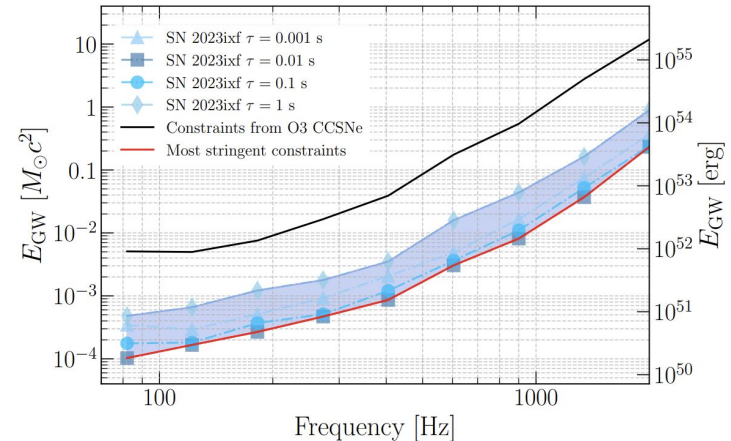
# Offline All-Sky search analysis



- K01 , no xgb ,  $\rho_0$  517.0yrs
- K02 , no xgb ,  $\rho_0$  519.0yrs
- K03 , no xgb ,  $\rho_0$  487.0yrs
- K04 , no xgb ,  $\rho_0$  521.0yrs
- K05 , no xgb ,  $\rho_0$  514.0yrs
- K06 , no xgb ,  $\rho_0$  500.0yrs
- K07 , no xgb ,  $\rho_0$  524.0yrs
- K08 , no xgb ,  $\rho_0$  446.0yrs
- K01 , xgb ,  $\rho_1$  517.0yrs
- K02 , xgb ,  $\rho_1$  519.0yrs
- K03 , xgb ,  $\rho_1$  487.0yrs
- K04 , xgb ,  $\rho_1$  521.0yrs
- K05 , xgb ,  $\rho_1$  514.0yrs
- K06 , xgb ,  $\rho_1$  500.0yrs
- K07 , xgb ,  $\rho_1$  524.0yrs
- K08 , xgb ,  $\rho_1$  446.0yrs



- Follow-up GW search performed for CCSN detected by EM observations within nearby galaxies.
  - SN2023ixf -> Type II supernova in M101 host galaxy 6.7 Mpc away.
- Conservative on-source-window of 5 days -> HL network live time of 0.68 days. No GW counterpart detected.
- Developed and performed the targeted cWB search:
  - Training WNBs injected around the SN2023ixf sky location.
  - XGBoost input features include sky location for reconstructed triggers.
- Most stringent constraints on GW energy emission from CCSN to date.



GW energy as a function of frequency  
[\[SN2023ixf arXiv:2410.16565\]](https://arxiv.org/abs/2410.16565).

# Summary



- Model-independent searches like Coherent WaveBurst (cWB) are crucial in detecting GW signals from **uncommon sources like IMBH and eBBH mergers**.
- cWB was recently upgraded with **WaveScan**, and now employs an **ML method**, resulting in a **~40% improvement** in the detection efficiency for BBH signals.
  - Updated cWB search uncovers **3 new BBH events** not detected by other LVK searches.
- O4 run underway: **Detected 70%** of 81 significant detection candidates in O4a. Detected heaviest merger GW231123 with an IFAR of 9700 yr.
- Provided most stringent constraints on GW energy emission from CCSN by following up SN2023ixf.
- Leading search for identifying generic GW burst sources!
- With next-generation ground-based GW detectors (Cosmic Explorer, Einstein Telescope, LIGO India) joining in future, we expect to detect GWs from unprecedented and exotic sources like eccentric binary mergers, hyperbolic encounters, unequal mass ratio mergers, Core-Collapse SN, and so on!



**Thank You.**



## Hands-on session notes

Relevant README on Sarathi: </home/tanmaya.mishra/README.cWB>

## Interactive Job

Prior to starting production, it is important to check if the jobs run without errors. To do this, run a single job interactively from a CIT headnode:

```
cwb_inet JobID
```

The `JobID` identifies the job segment (1200s) that will be analyzed and is unique for a given chunk.

If the job finishes without running into any errors, you will find a wave root file in the data directory (e.g. `data/wave_....root`). **!!**

The root file contains all reconstructed events with the estimated summary statistics values. An optional check is to print a few summary statistics by opening and scanning the root file:

```
root data/wave_...root
ScanWAVE("rho[0]:rho[1]:time[0]:netcc[2]", "netcc[0]>0.8, "colsize=14")
.q
```

## Producing CEDs

The Coherent Event Display (CED) pages helps one investigate the time-frequency WaveScans and additional details for a reconstructed event and can be produced by running the following command:

```
cwb_inet JobID GPS ced factor lag
```

- For loudest background outliers, run the following to produce the corresponding CED:
  - `cwb_inet JobID 0 ced 0 lag`

Here `JobID` and `lag` are printed on the background report pages for loudest outliers.

- For zero-lag events run:
  - `cwb_inet JobID 0 ced 0 0`

One can access the `JobID` by running the `cwb_dump job` (or `sjob` in case of all possible segments/extended segments). This produces a file containing all the job segments and stores it in the `report/dump/____.job` (`sjob`) file. The `JobID` is given by the line number corresponding to the required job segment in the `____.job` file. **!!**

## Production

Once the working directories are finalized, the cWB XP production is done by submitting jobs to the CIT cluster through `condor`. The condor job system allows us to analyze multiple jobs simultaneously on different nodes across the CIT cluster. The information about the different jobs in a given working directory are created and stored in dag file in the condor folder.

```
cwb_condor create
```

Running the above command creates the following files:

- `condor/____.dag`

Here, the `.sub` file enables condor to access information related to the scripts and config necessary to run the job (explained in the following bullet point). The `PID` denotes the `JobID`, the `CWB_HOME_WAT` points to the installed library that will be used to run the job on a condor specified execute node (it sources the `cit_watenv_cmake.sh` so make sure to correctly setup the config) **!!**, the `RETRY ___ 5` denotes the number of times the job is automatically submitted to find a node.

- `condor/____.sub`

Stores the common information used to run all the jobs defined in the `.dag` file.

The condor job scheduler uses the `request_memory` and `request_disk` provided in the `.sub` file to find the suitable execute nodes that meet the job requirements.

The `output`, `error`, `log` files should be used to debug any failing jobs on condor. **!!**

# cWB on HTC Condor

## Submitting condor jobs

Use the `condor_submit_dag` option when submitting more than 1000 jobs at once to ensure the job scheduler does not push your priority to lower **!!!**. For less than 1000 jobs you can directly submit, from the working directory, by running the `cwb_condor submit` command.

```
cd condor
condor_submit_dag -maxjobs 1000 _____.dag
```

The `-maxjobs` option constraints the total number of `RUN + HOLD + IDLE` to 1000.

To check the status of condor jobs, run the following

```
condor_q OR condor_watch_q
```

If you find that there are jobs on HOLD, you can check all the jobs and the reason for going to hold by running the following command

```
condor_q --held
```

This will print the condor DAGMANjobID or the process ID (something like this `37348619.0` ), and you can follow up on this job by either searching for the process ID in the `condor/____.Log` file, or by running the following

# cWB on HTC Condor

In case a job is going to `IDLE` for too long, you can check if nodes with the requested specification exist

```
condor_q process_ID --better-analyze
```

This tells you how many nodes on CIT can run your job.

In case a job goes to `HOLD` because it is going over the requested memory/disk, then you can increase the job's requirements on the fly by running the following:

```
condor_qedit --jobids process_ID 'RequestMemory=10000'  
condor_qedit --jobids process_ID 'RequestDisk=10000000'
```

In both cases shown above, we are increasing the requested memory and disk to 10 GB (or 10000 MB). Note that for the request disk update, the disk size is specified in KB whereas memory is requested in MB. **!!**

# cWB Hands-on Tutorial on Sarathi

Follow README on Sarathi:

```
/home/tanmaya.mishra/README.cWB
```

Working Directories:

- Looking at GW231123 event - Produce CED:

```
/home/tanmaya.mishra/O4/SEARCHES/OFFLINE/BBH/LH/BKG/O4_K07_C00_LH_BBH_BKG_tst2
```

- Running BBH injections using HTCondor:

```
/home/tanmaya.mishra/O4/SEARCHES/OFFLINE/BBH/LH/SIM/O4_K07_LH_BBH_SIM_PLP_Aug2024_IUCAA_tst2
```

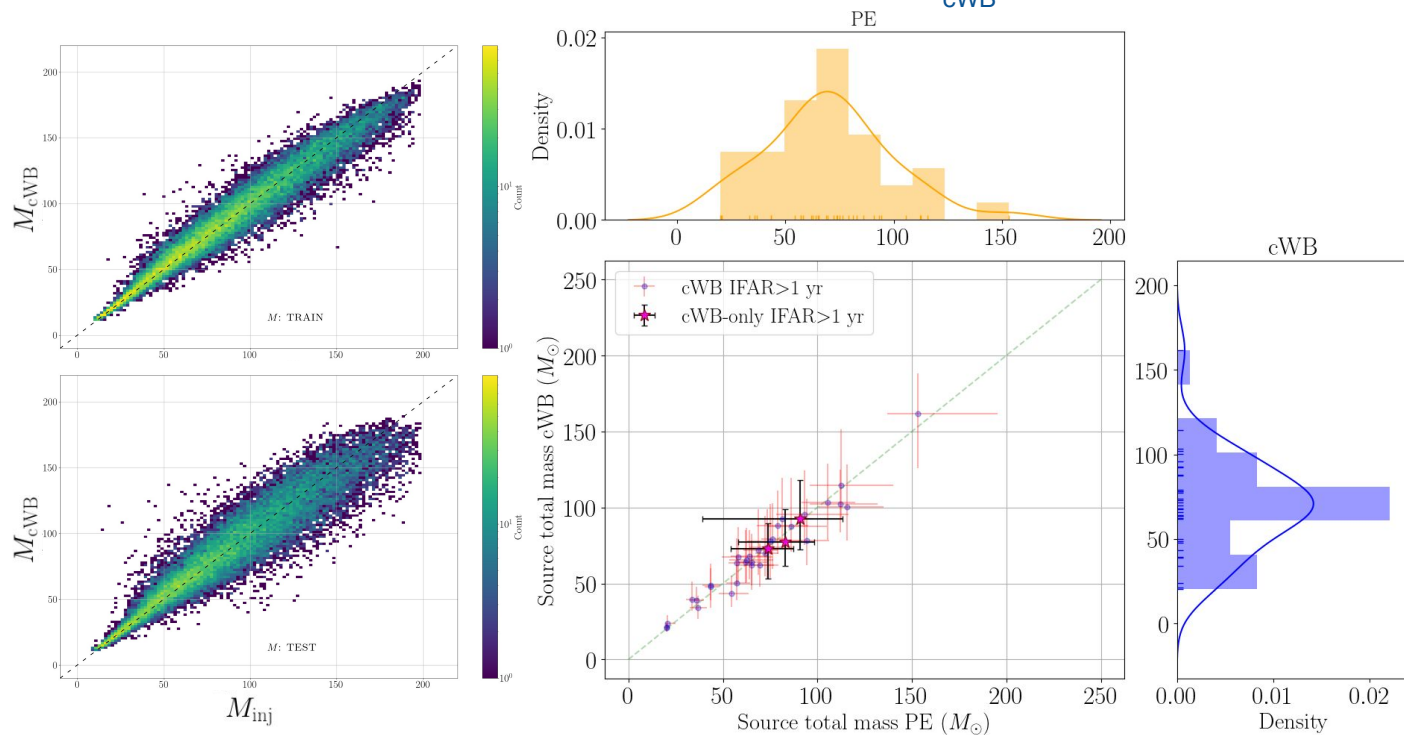
- Produce merged root files, apply trained XGB model, assign significance, assess sensitivity by making ROC/detection efficiency plot

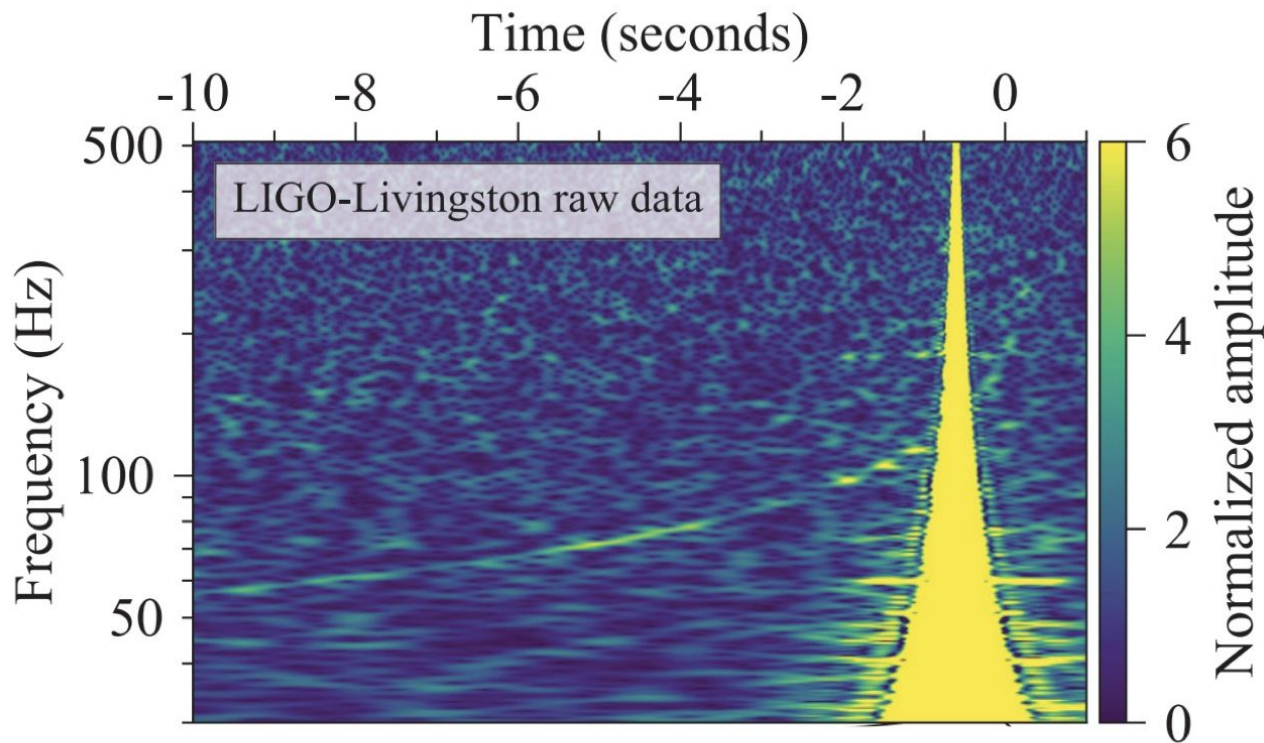


## Revisiting session 1 QnA ↓

# Rapid Parameter estimation (PE) with cWB-ML

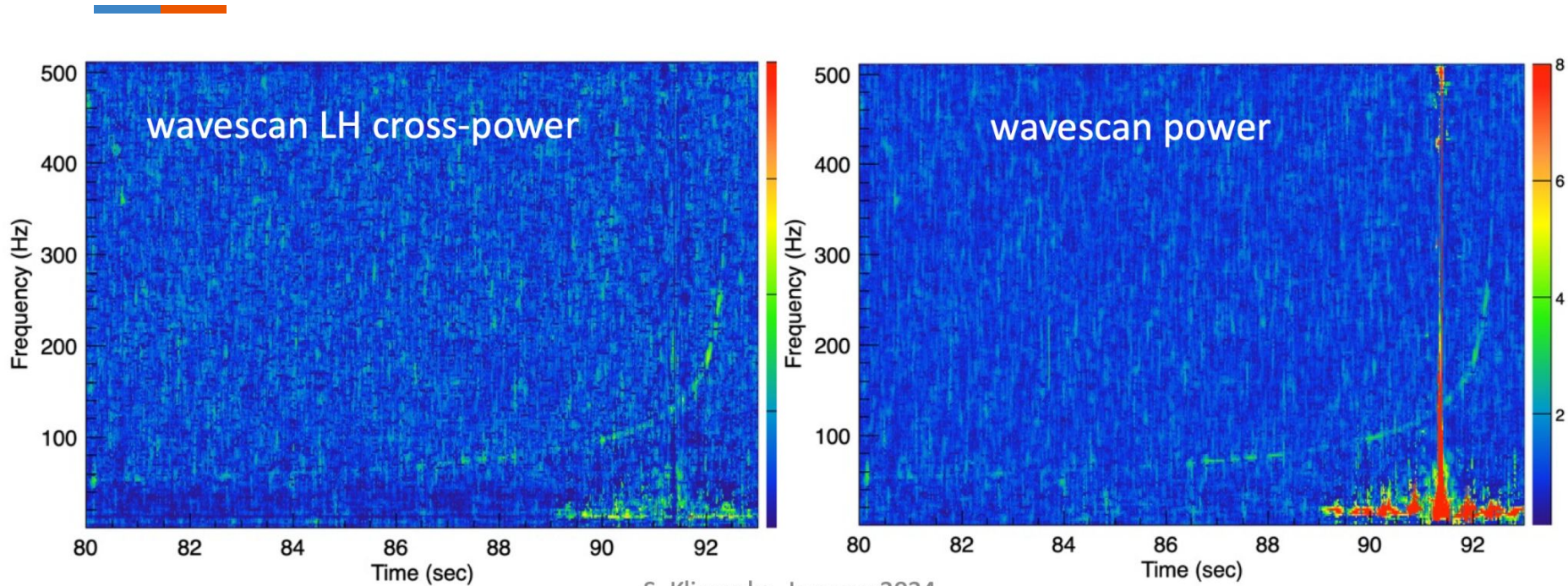
- We build a simple XGBoost regression model using a selected set of **summary statistics** for recovered injections with cWB.
  - The model is used to estimate the source total mass  $M_{\text{cWB}}$ .





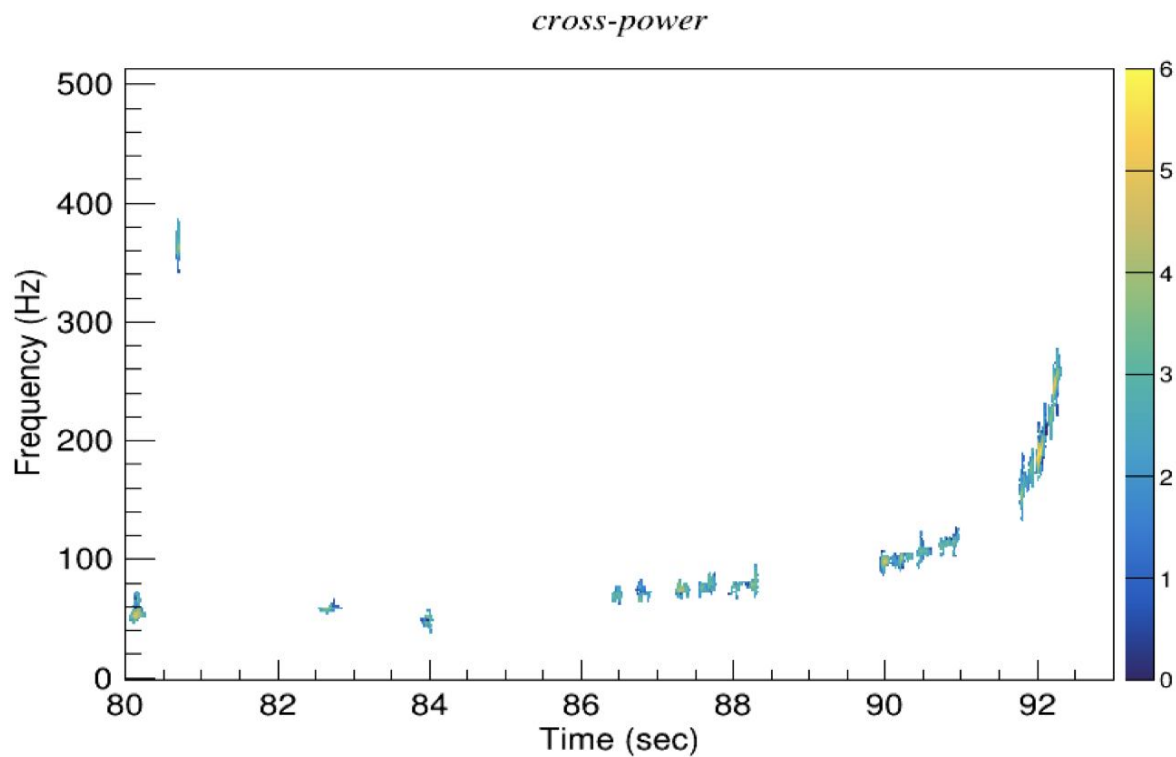
[Phys. Rev. Lett. \*\*119\*\*, 161101 \(2017\)](#)

# GW170817



S. Klimenko, January 2024

# GW170817





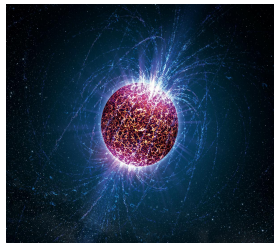
**Extra Slides ↓**

# Sources of GW for ground-based detectors

Long  
Duration

## Continuous

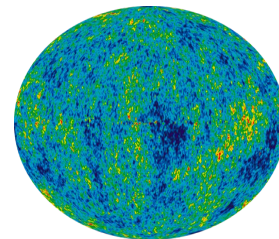
Rotating Neutron Star (NS)  
with surface  
imperfections.



[C. Reed, Penn State]

## Stochastic

Cosmic GW  
background from the  
early universe.



[NASA/WMAP Science Team]

Short  
Duration

## Compact Binary Coalescence (CBC)

- Binary Black Hole (BBH)
- Binary Neutron Star (BNS)
- Neutron Star - Black Hole (NSBH)



[SXS Project]

## Burst

Core-Collapse  
Supernova (CCSN).



[NASA, ESA, J. Hester, ASU]

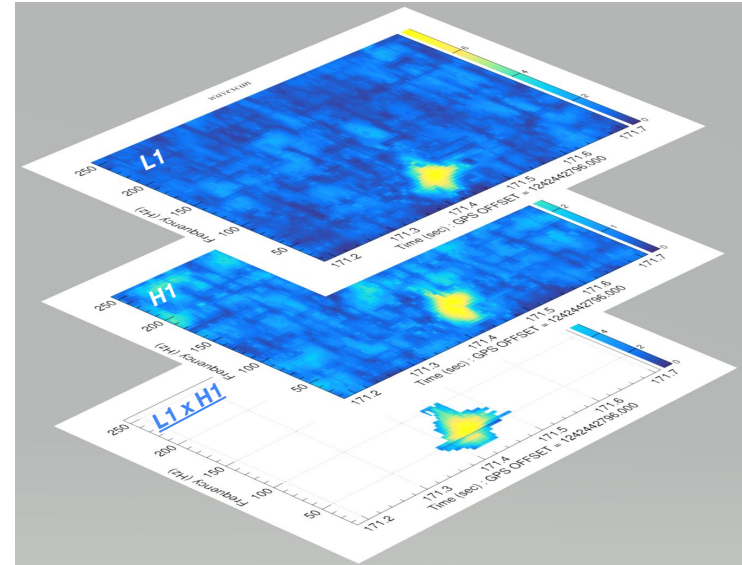
90 CBC signals detected by ground-based detectors!

Modeled

Unmodeled

# Coherent WaveBurst

- Recently, cWB was upgraded for O4:
  - New time-frequency transform called **wavescan** [[S Klimenko 2022](#)]
  - Machine Learning (ML) method employed for separating GW signals and noise [[T. Mishra+ 2021](#), [T. Mishra+ 2022](#)]



GW190521 detection with updated cWB

- The updated cWB search detected **3 new GW event candidates** not found by the LVK matched-filter searches:
- The three new candidates together have a **combined significance of  $3.6 \sigma$** , representing a statistical excess of events detected only by the burst search.
- PE results for the events performed with the IMRPhenomXPHM waveform in the Bilby framework:

Event ID	IFAR [yr]	SNR	$P_{\text{astro}}$	GPS time	$\mathcal{M}_c[M_\odot]$	$M[M_\odot]$	$m_1[M_\odot]$	$m_2[M_\odot]$	$q$	$\chi_p$
190711	<b>40.7</b>	10.1	0.99	1246849694.6	$31.3^{+8.5}_{-6.0}$	$91.0^{+51.9}_{-22.2}$	$69.9^{+56.8}_{-26.3}$	$20.1^{+12.5}_{-8.7}$	$0.29^{+0.38}_{-0.19}$	$0.29^{+0.36}_{-0.21}$
190607	<b>11.7</b>	9.4	0.99	1243931925.9	$31.4^{+8.7}_{-6.0}$	$74.0^{+19.9}_{-13.2}$	$43.1^{+13.7}_{-9.6}$	$30.9^{+11.2}_{-9.7}$	$0.74^{+0.23}_{-0.32}$	$0.49^{+0.40}_{-0.37}$
200318	<b>4.5</b>	8.4	0.94	1268594035.1	$35.0^{+10.6}_{-7.2}$	$83.0^{+25.0}_{-15.5}$	$49.9^{+17.9}_{-12.2}$	$33.5^{+13.5}_{-11.6}$	$0.69^{+0.27}_{-0.31}$	$0.49^{+0.40}_{-0.36}$

- Notably, 190711 with  $m_1 \sim 70 M_\odot$  suggests a black hole in the **PISN mass gap**, and shows highly asymmetric masses ( $q \sim 0.3$ ) indicative of a potential dynamical or AGN-assisted formation scenario.

# Gravitational wave network



LIGO Hanford



GEO600



KAGRA

Underground detector.

LIGO Livingston

Virgo

LIGO India

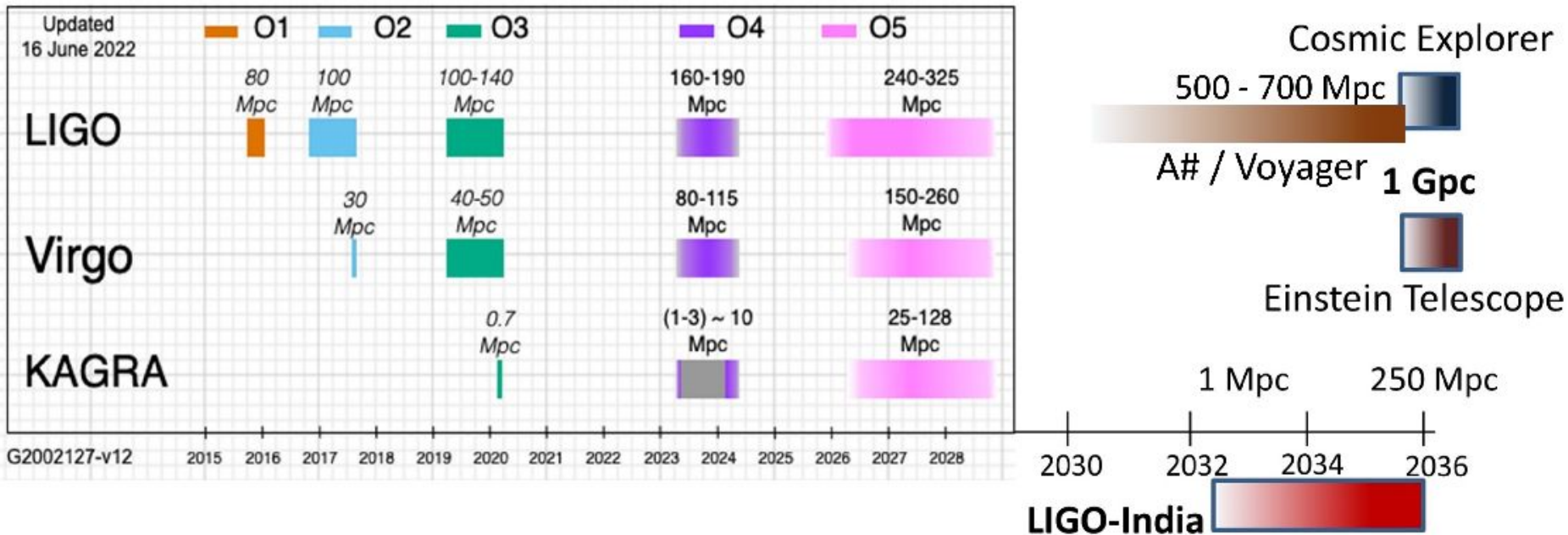


CAD drawing of the proposed LIGO India observatory

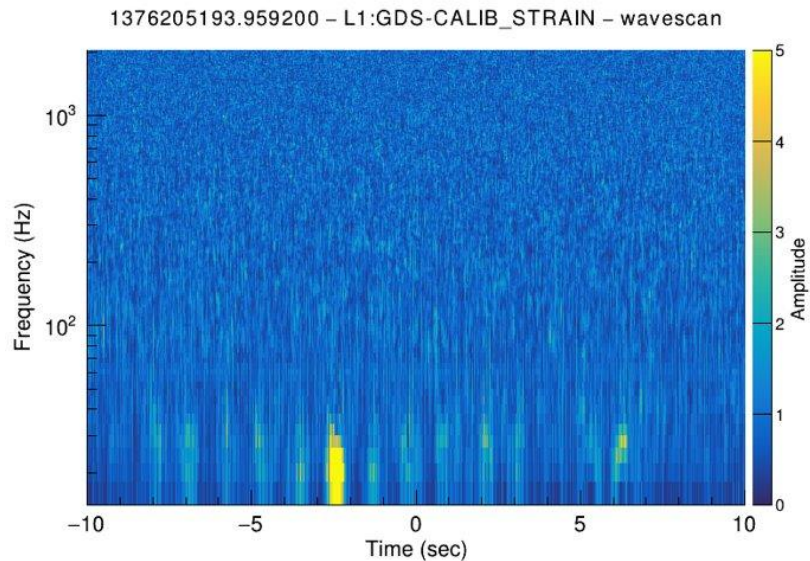
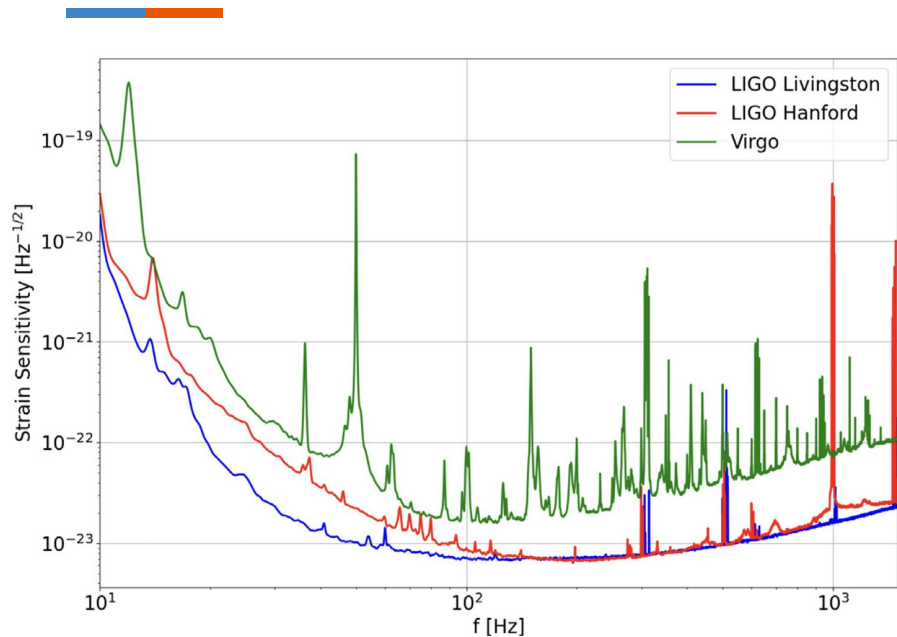
## Gravitational Wave Observatories

A third LIGO detector in India (2032/2036)

# GW sensitivity across observing runs

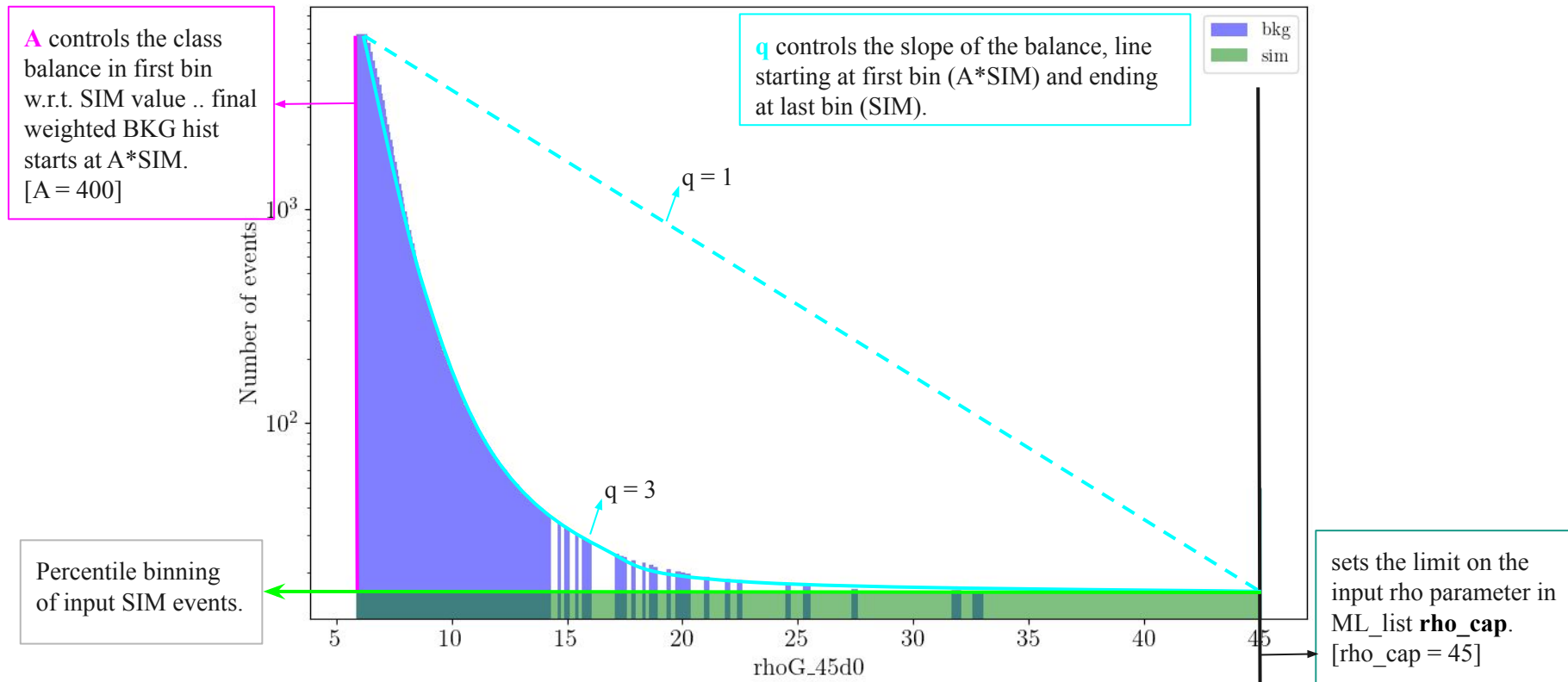


# Detector sensitivity and Scatter noise example



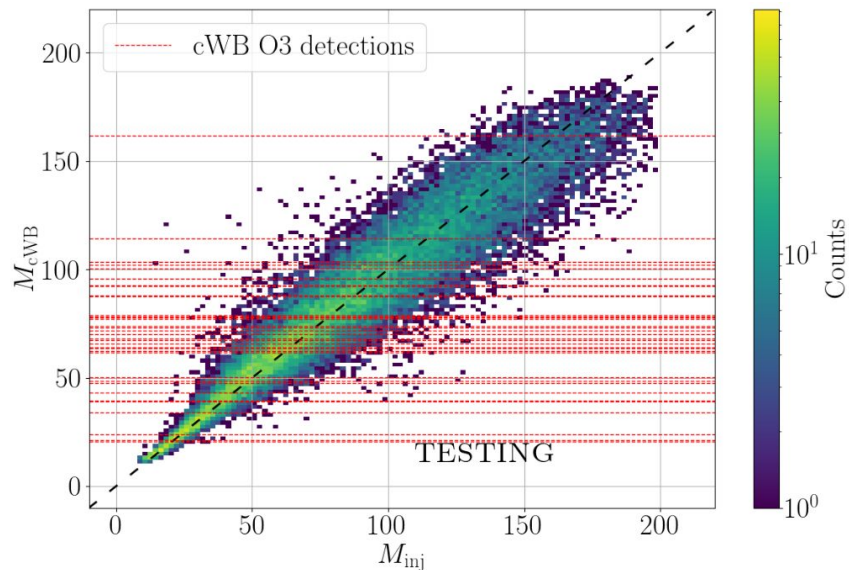
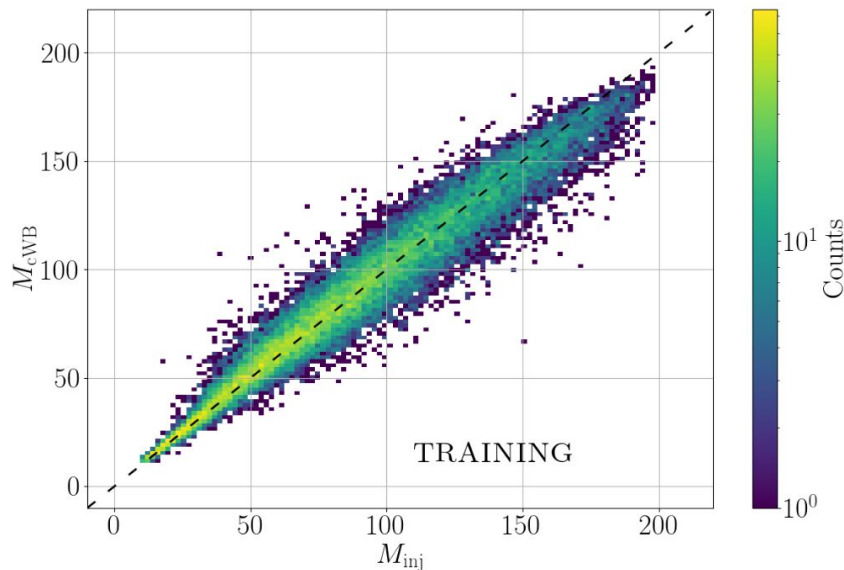
# XGB Training: Degrees of freedom

## Background/Simulation Balanced Distributions



# Rapid Parameter estimation with cWB-ML: Errors

- We build a simple XGBoost regression model using a selected set of summary statistics for recovered injections with cWB.
  - The model is used to estimate the source total mass  $M_{\text{cWB}}$ .



# 3G Detectors



SENSITIVITY OF ET AND CE COMPARED TO ADVANCED LIGO & THE REACH FOR 3G OBSERVATORIES

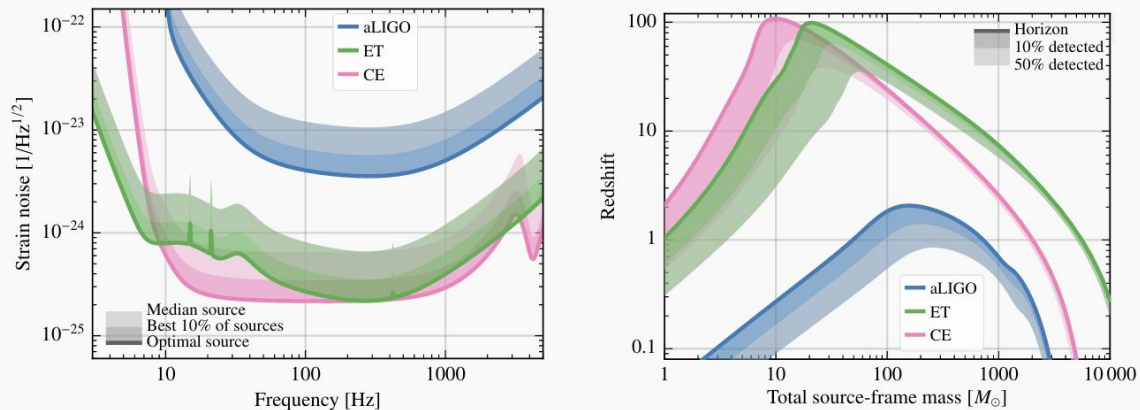


Figure 1.1: GW strain noise for current and future detectors (left) and astrophysical reach for equal-mass, nonspinning binaries distributed isotropically in sky and inclination (right).

VISIBILITY OF BLACK HOLE BINARIES IN 3G

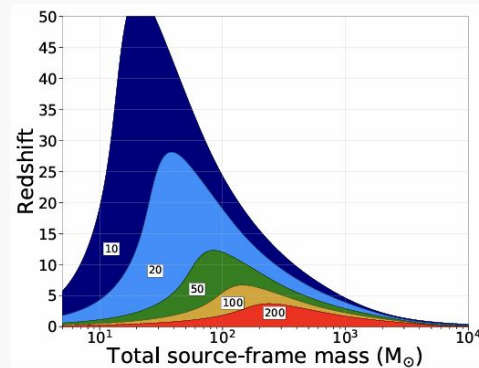


Figure 1.2: Signal-to-noise ratio contours as a function of binary's total mass and its redshift for equal mass binaries averaged over sky position and orientation in the 3G network.