PyGWB

A python-based, user-friendly library for gravitational-wave background (GWB) searches with ground-based interferometers.

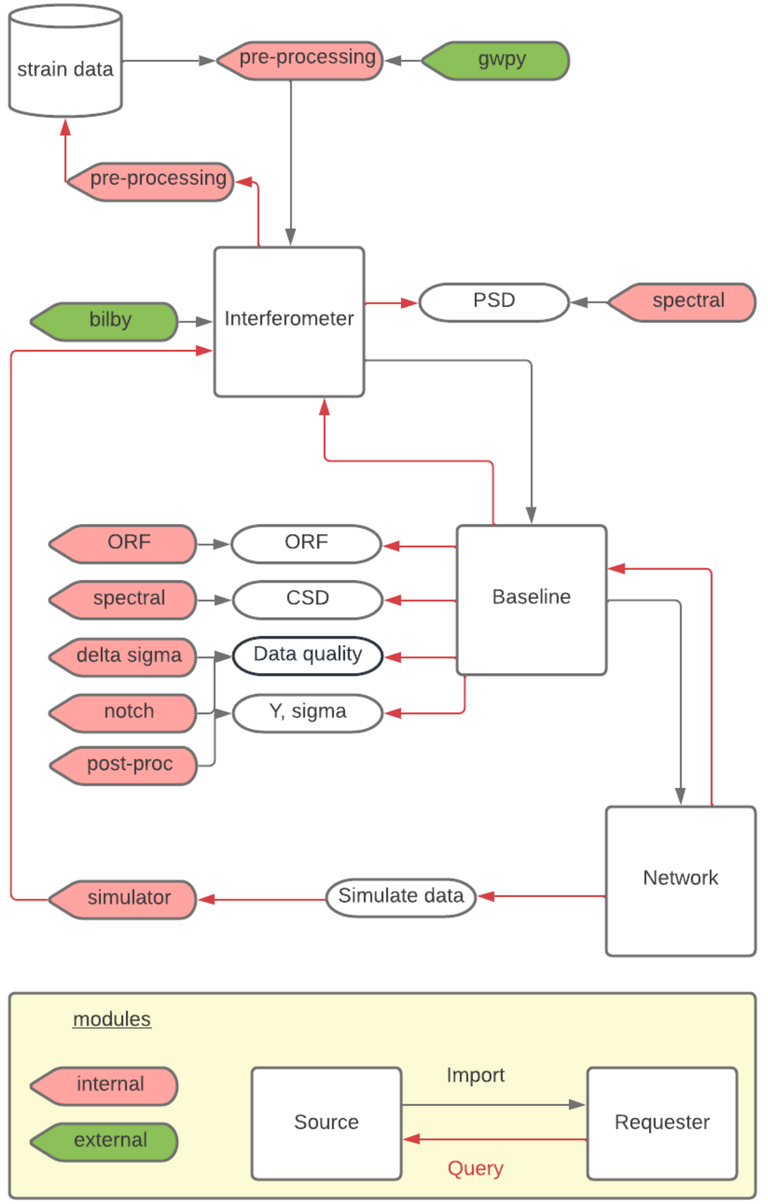
Chia-Hsuan on behave of stochastic isotropic subgroup

Overview of pygwb

- Pipeline created by members of the stochastic group that performs cross correlations between data from the HLVK baselines

  (in the future it can be adapted to more IFOs)

- A bit of history: until O3 we would use stochastic.m + auxiliary codes (mainly in matlab):
  - Stochastic.m → performed the preprocessing of the data and the cross correlation
  - Postprocessing and PE had to be done separately
- Objective of pygwb:
  - Move from matlab to a more friendly and still maintained programming language: Python
  - Self-contained pipeline: no auxiliary codes are needed
  - Open to the general public, not just LVK members and very easy to use
  - We have revisited some functions or parameters and have tried to reuse many functions from gwpy



Pre-processing

Science ready flag work and build with gwpy

Segment information

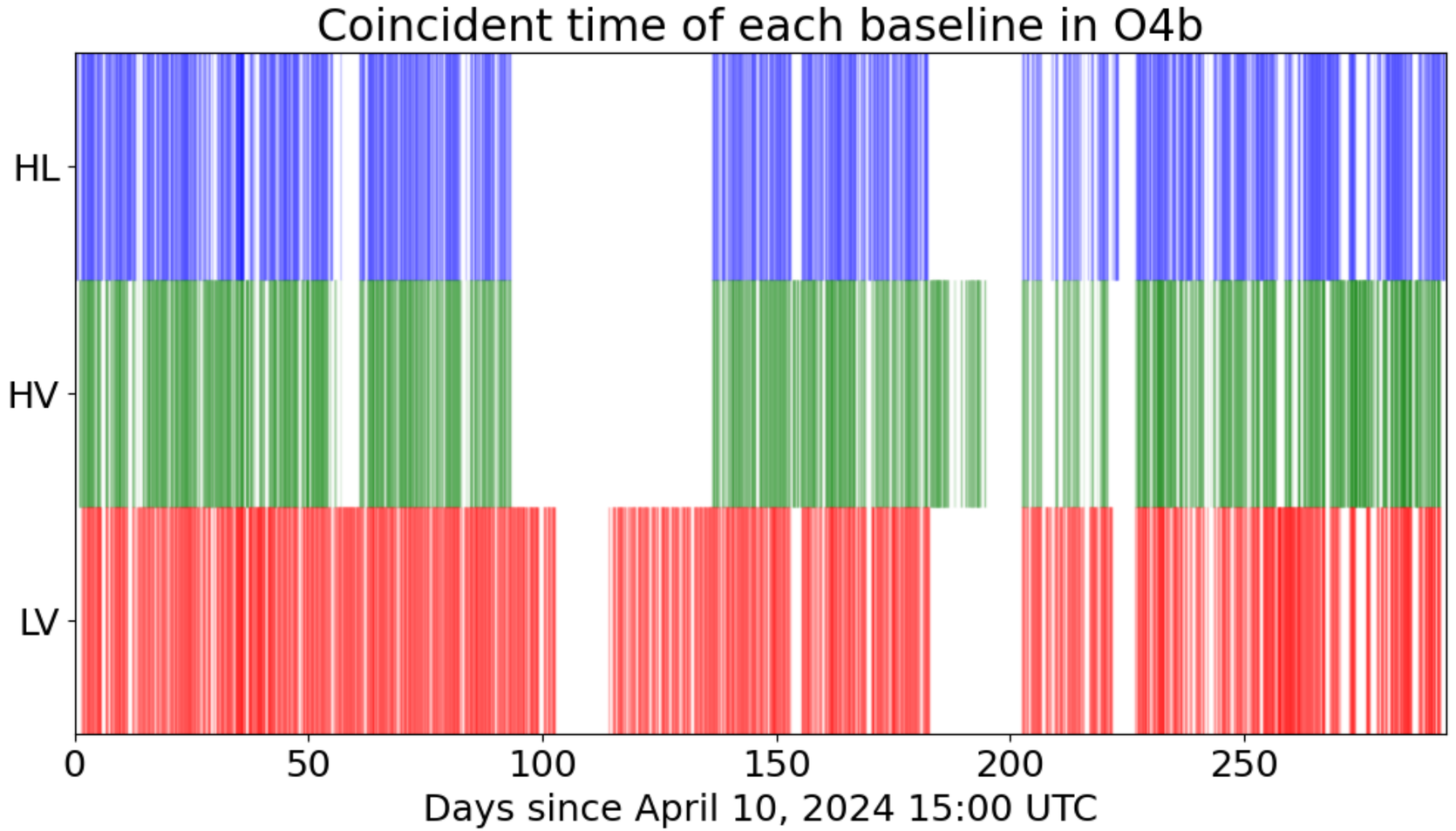
The following flags were used in the above data. This list does not include state information or combinations of flags. Percentages are calculated relative to the total duration of 86400 seconds.

Name	Defined duration [s]	Active duration [s]	Padding	Description
G1:GEO-SCIENCE:1	86400 (100%)	73930 (85.57%)	-	
G1:GEO-UP:1	86400 (100%)	77056 (89.19%)	-	
H1:DMT-ANALYSIS_READY:1	86400 (100%)	68245 (78.99%)	-	H1 Ready for Analysis (Science & Calibrated) from h(t) DQ flags
H1:DMT-GRD_ISC_LOCK_NOMINAL:1	86400 (100%)	68867 (79.71%)	-	Guardian indicates IFO is in nominal lock state
H1:DMT-OBS_READY:1	86400 (100%)	68259 (79%)	-	
L1:DMT-ANALYSIS_READY:1	86400 (100%)	76949 (89.06%)	-	L1 Ready for Analysis (Science & Calibrated) from h(t) DQ flags
L1:DMT-GRD_ISC_LOCK_NOMINAL:1	86400 (100%)	77786 (90.03%)	-	Guardian indicates IFO is in nominal lock state
L1:DMT-OBS_READY:1	86400 (100%)	76963 (89.08%)	-	
V1:ITF_NOMINAL_LOCK:1	86400 (100%)	63979 (74.05%)	-	
V1:ITF_SCIENCE:1	86400 (100%)	61222 (70.86%)	-	Science mode when equal to 1

https://ldas-jobs.ligo.caltech.edu/~detchar/summary/day/20241129/

Pre-processing

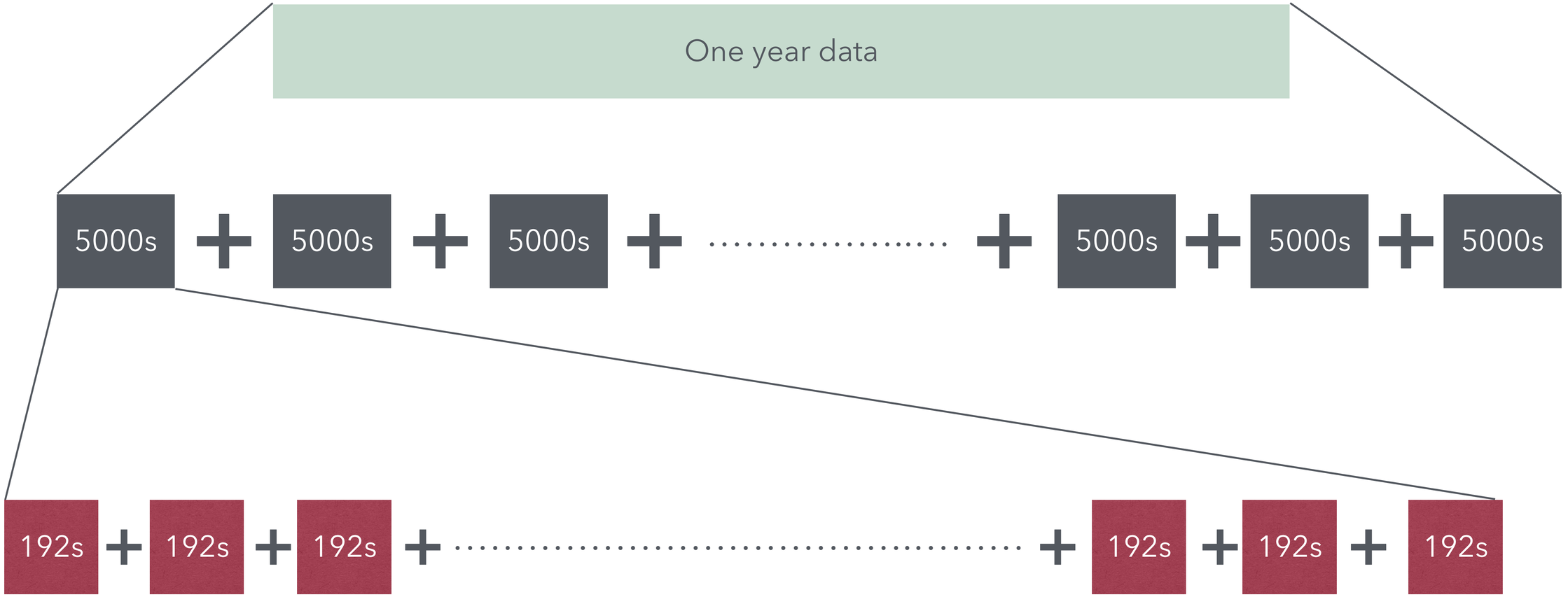
Science ready flag work and build with gwpy




```
1 1370375313 1370376655 1342
1 1370383648 1370388648 5000
1 1370388648 1370393648 5000
1 1370393648 1370398648 5000
1 1370398648 1370400556 1908
1 1370402911 1370407911 5000
1 1370407911 1370412911 5000
1 1370412911 1370415251 2340
1 1370421494 1370426494 5000
1 1370431479 1370436479 5000
1 1370436479 1370441479 5000
1 1370441479 1370446479 5000
1 1370446479 1370451479 5000
1 1370451479 1370456479 5000
1 1370456479 1370461479 5000
1 1370466646 1370471646 5000
1 1370471646 1370475332 3686
1 1370479968 1370480701 733
"jobfile.dat" 2493L, 72213C
```

Pre-processing

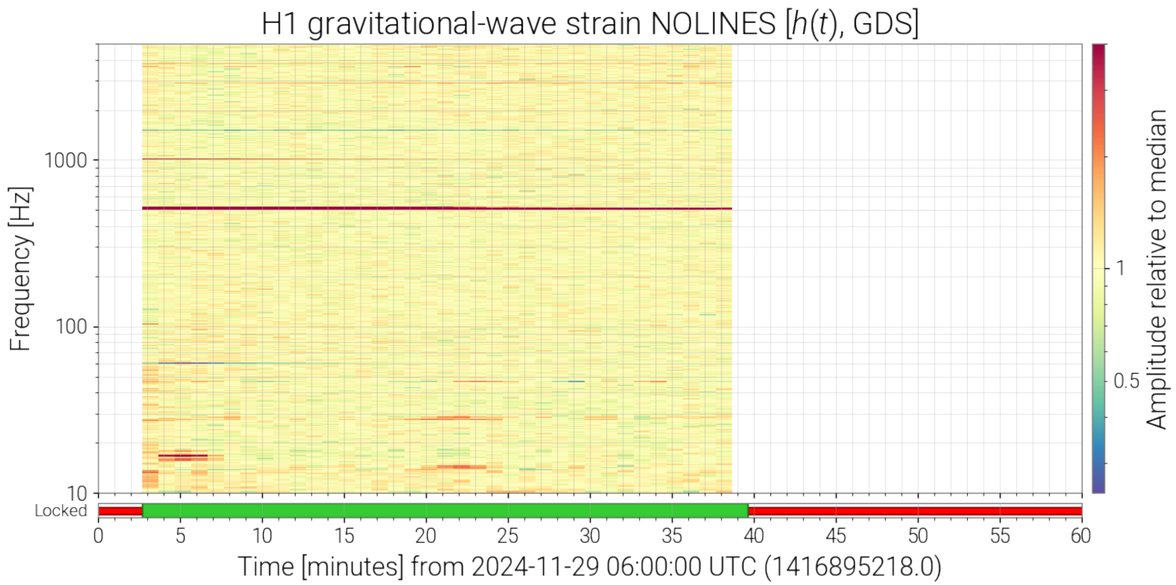
Science ready flag work and build with gwpy



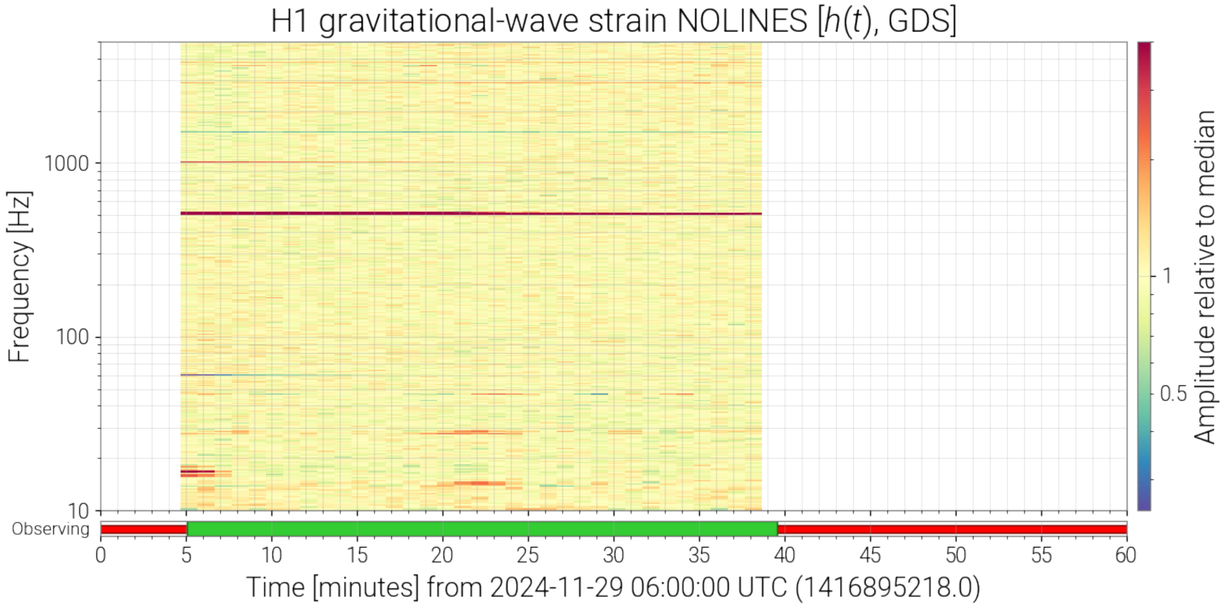
FFT and also consider the change of ORF

Pre-processing

Locked



Observing



Pre-processing

1. Loading data according to the job file with correct channel

```
[data_specs]
t0 = 1377106176.0
tf = 1377106920.0
interferometer_list = ['H1', 'L1']
data_type = private
channel = GDS-CALIB_STRAIN_CLEAN_AR
frametype = H1:H1_HOFT_C00_AR,L1:L1_HOFT_C00_AR
time_shift = H1:0,L1:1
random_time_shift = False
```

2. Down sampling and Applying highpass filter (11Hz)

3. Time shifting

(if necessary)

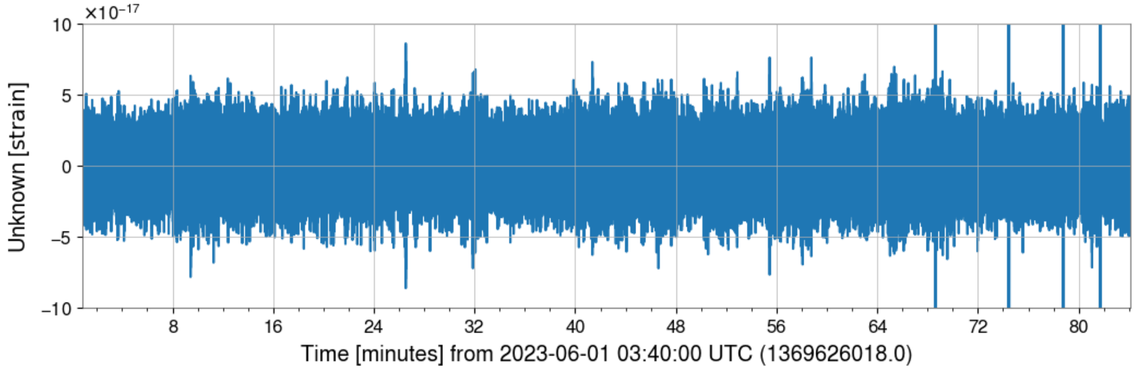
```
[preprocessing]
new_sample_rate = 4096
cutoff_frequency = 11.0
segment_duration = 192
number_cropped_seconds = 2
window_downsampling = hamming
```

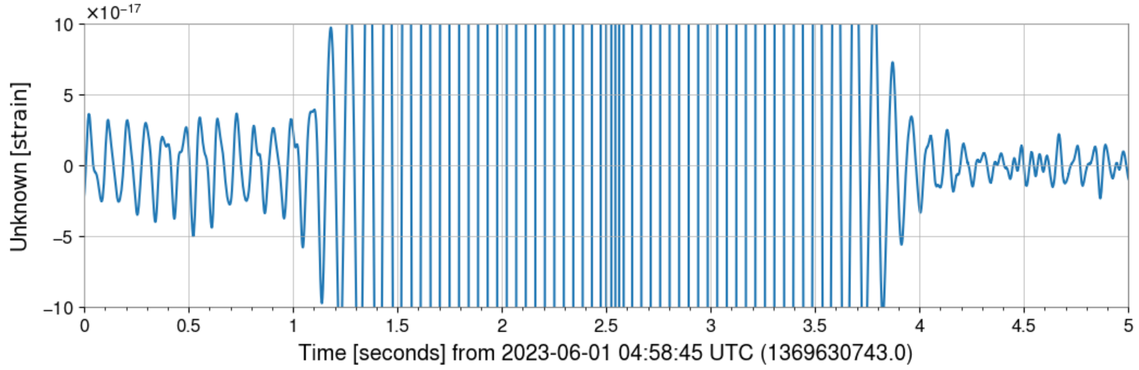
4. Gating

```
gate_data = True
gate_whiten = True
gate_tzero = 1.0
gate_tpad = 0.5
gate_threshold = 50.0
cluster_window = 0.5
```

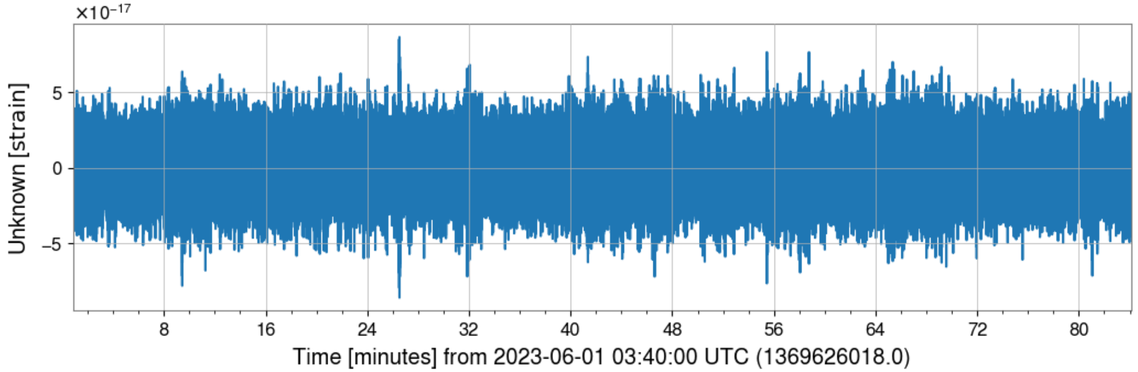
Pre-processing

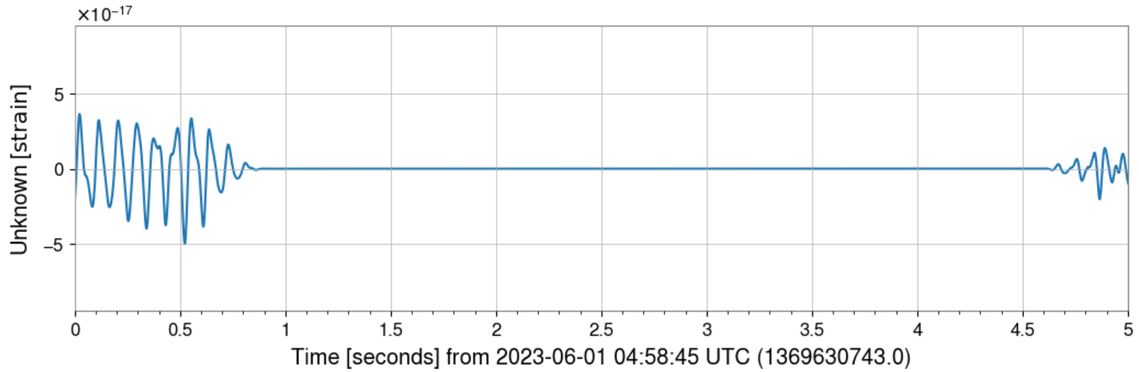
Before gating



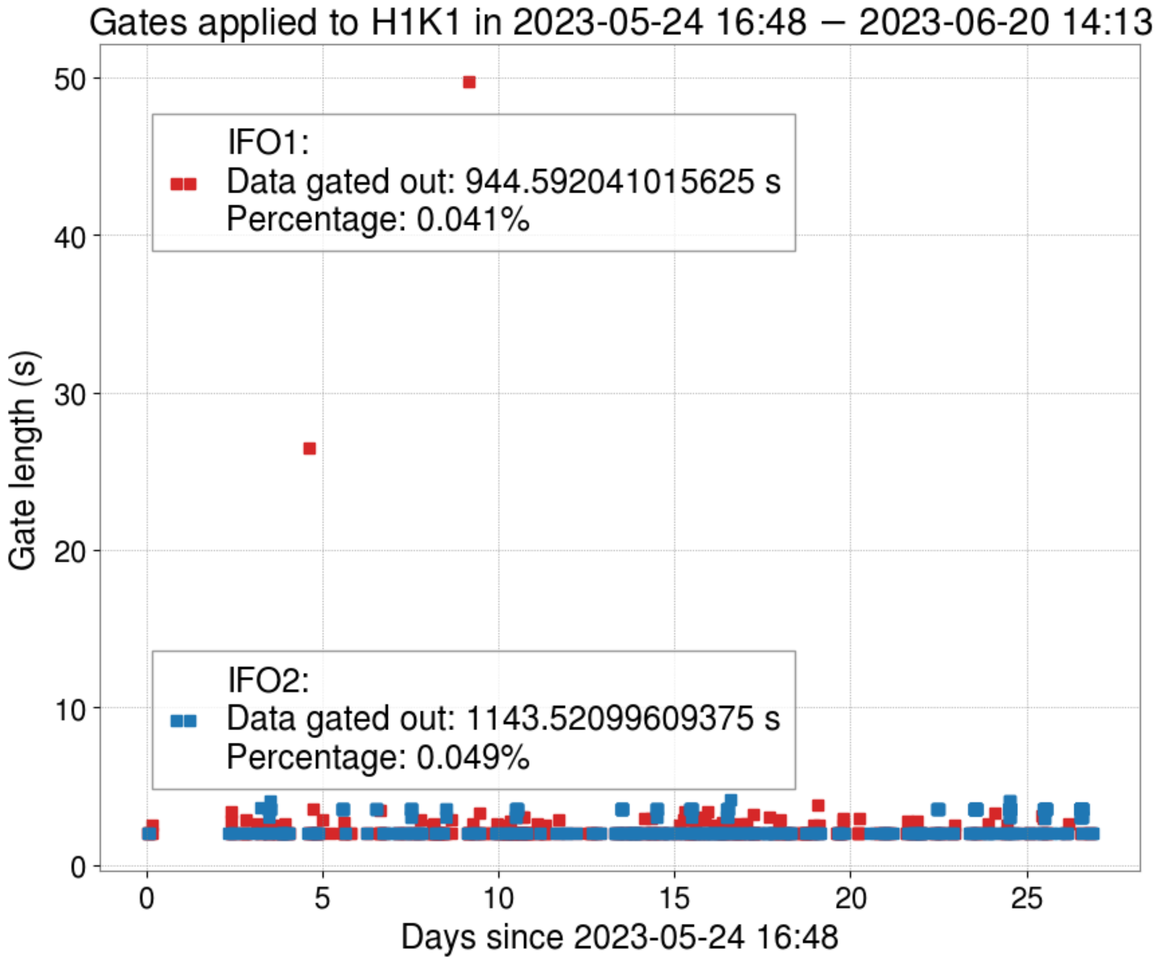


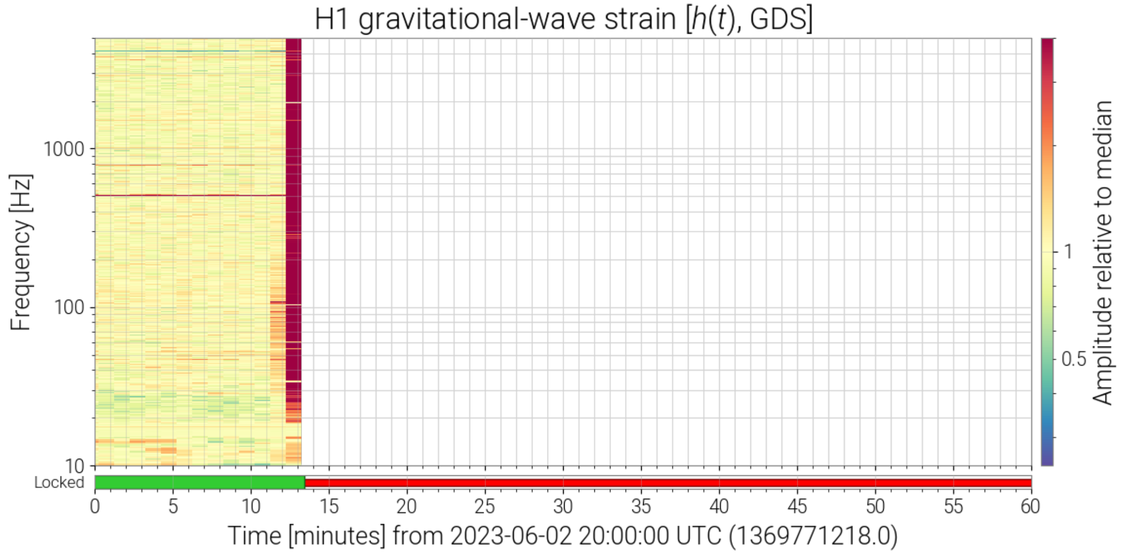
After gating

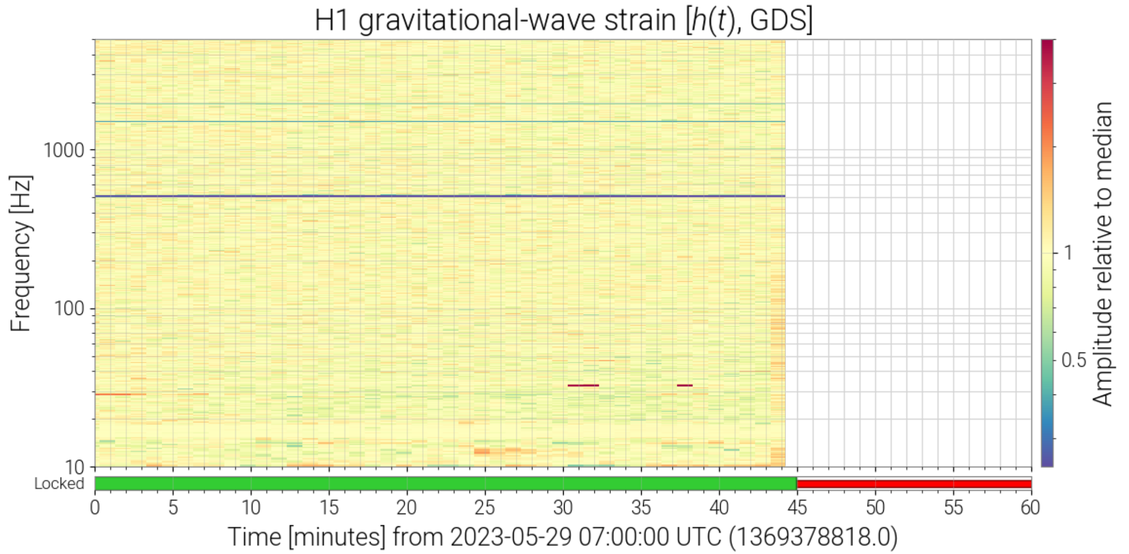




Pre-processing

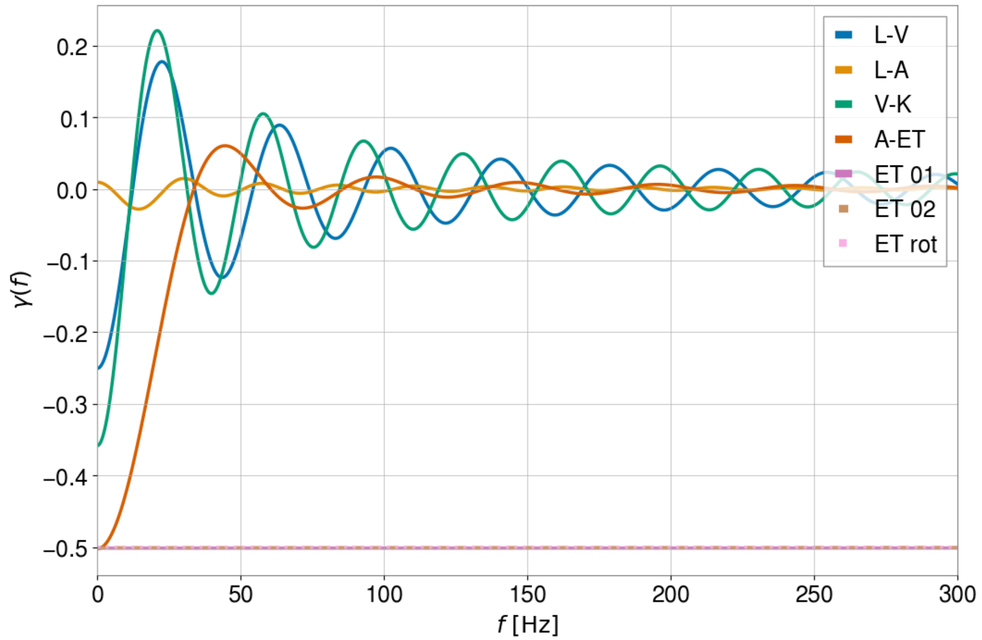






ORF

```
for ii,baseline in enumerate([L1V1, L1A1, V1K1, A1ET, ET01, ET02, ET_test120]):
    if baseline.name == 'ET 02':
        ls='--'
    elif baseline.name == 'ET rot':
        ls=':'
    else:
        ls='-'
    plt.plot(freqs, baseline.overlap_reduction_function, label=baseline.name, ls=ls, lw=2.5,
c=my_palette[ii])
```

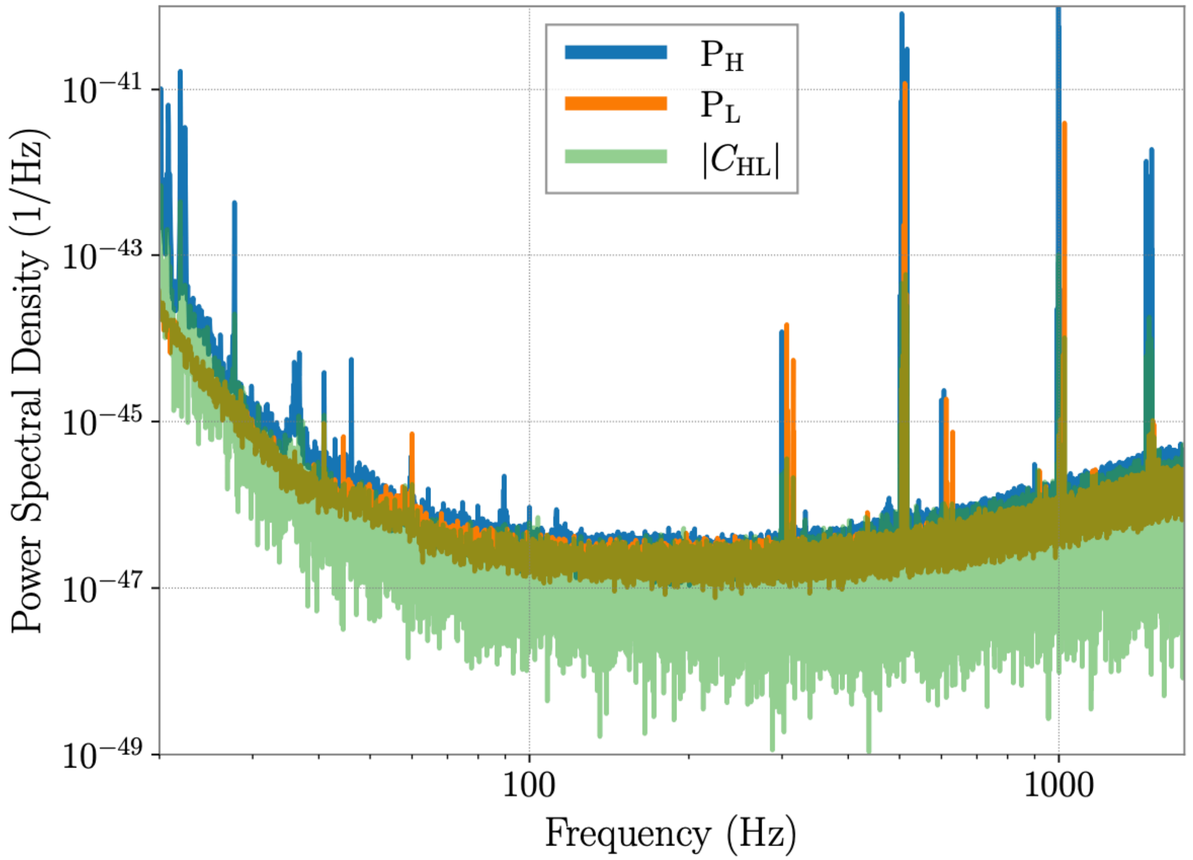


Spectral

```
logger.info(f"Setting PSDs and CSDs of the baseline...")
base_12.set_cross_and_power_spectral_density(params.frequency_resolution)
base_12.set_average_power_spectral_densities()
base_12.set_average_cross_spectral_density()
```

$$C_{IJ,f} = \frac{2}{T}\tilde{s}^*_{I,f}\tilde{s}_{J,f}\,, \qquad P_{I,f} = \frac{2}{T}|\tilde{s}_{I,f}|^2$$

$$\tilde{s}_f \equiv \sum_{t_k=0}^{T-\delta t} s(t_k)\, e^{-i2\pi m t_k/T}\,,$$

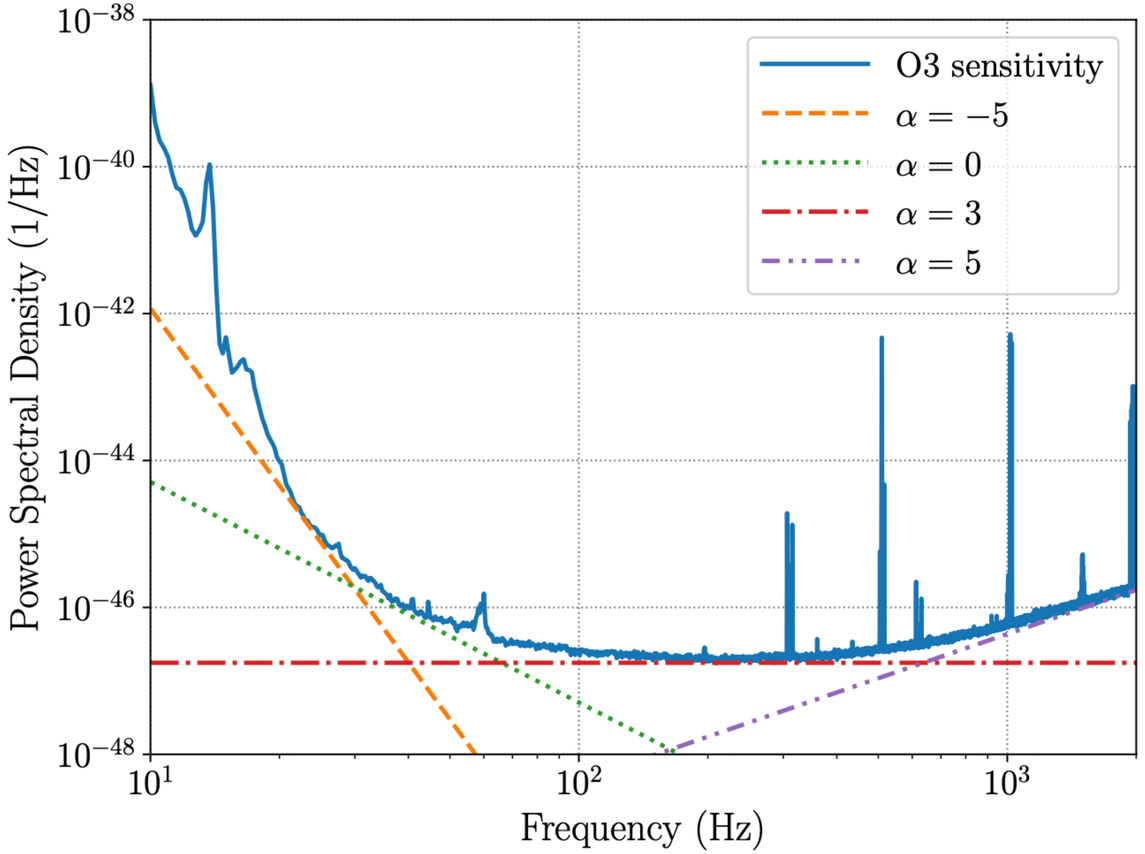


Delta sigma cut

$$\frac{|\sigma_i - \sigma_{i+1}| + |\sigma_i - \sigma_{i-1}|}{2\sigma_i} > \text{threshold}$$

```
alphas_delta_sigma_cut = ['-5', '0', '3', '5']
delta_sigma_cut = 0.2
```

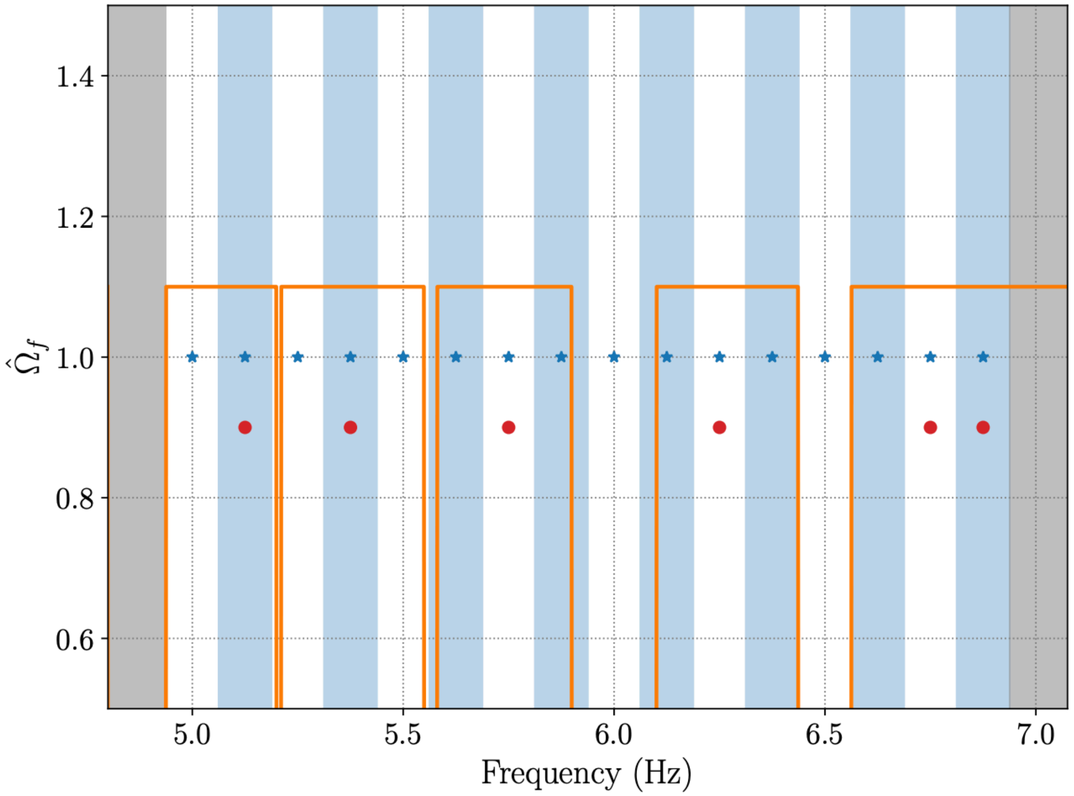
$$\frac{|\bar{\sigma}_{t,\alpha} b_{\text{avg}} - \sigma_{t,\alpha} b_{\text{nav}}|}{\bar{\sigma}_{t,\alpha} b_{\text{avg}}} > \text{threshold}$$



Notch

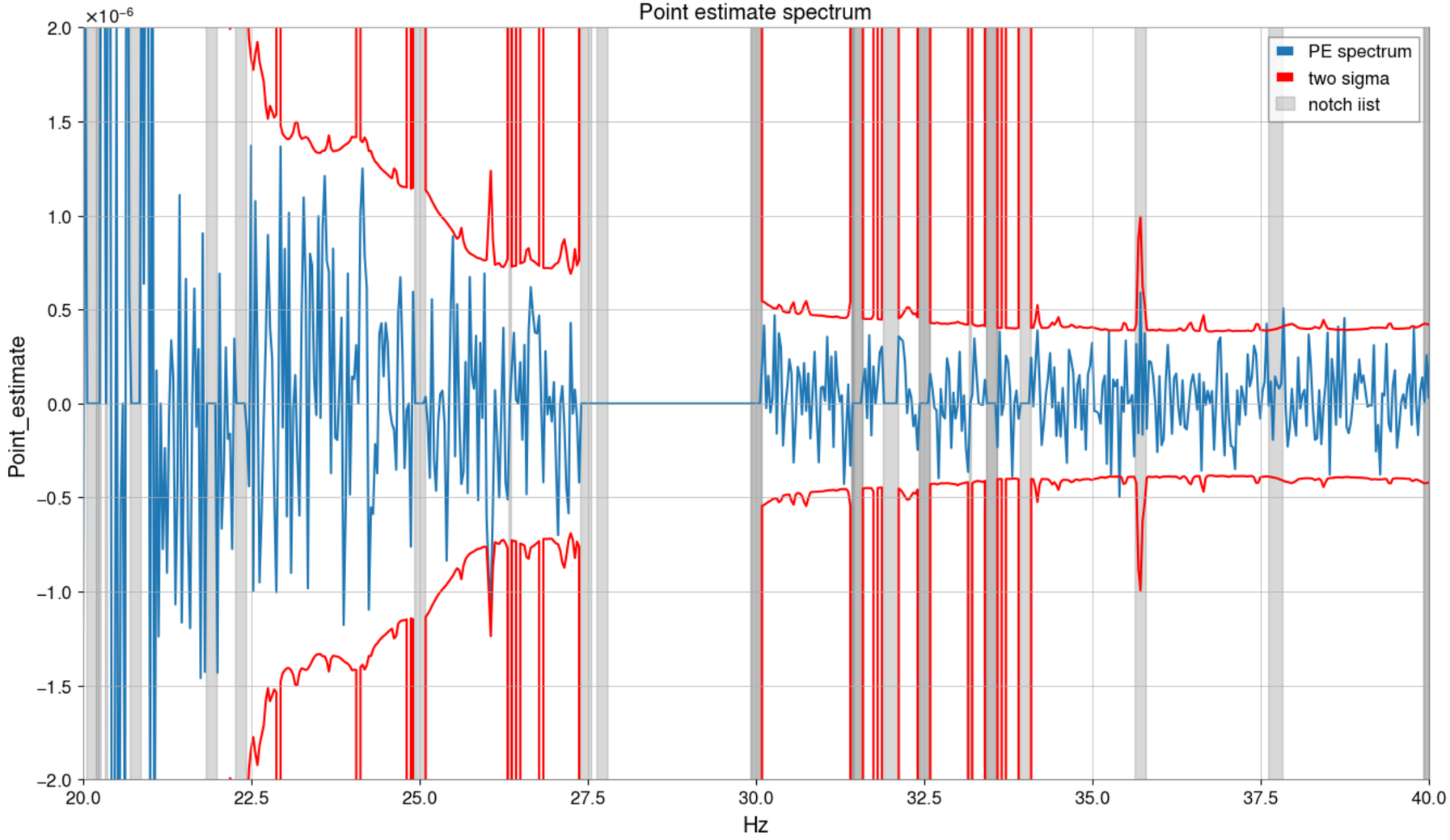
To remove the contribution from known noise

Bin width = δf by default ($\frac{\delta f}{2}$ for both left and right)



Notch

/home/chia-hsuan.hsiung/notchlist.ipynb



Short summary — Parameter setting

```
t0 = 1374206066.0
tf = 1374206957.0
interferometer_list = ['H1', 'L1']
data_type = private
channel = GDS-CALIB_STRAIN_CLEAN_AR
frametype = H1:H1_HOFT_C00_AR,L1:L1_HOFT_C00_AR
time_shift = H1:0,L1:1
random_time_shift = False
```

data_type = {Public, Private, Local}

`FrChannels (path/to/data)` command will help

Short summary — Parameter setting

```
[preprocessing]
new_sample_rate = 4096
cutoff_frequency = 11.0
segment_duration = 192
number_cropped_seconds = 2
window_downsampling = hamming
ftype = fir

[gating]
path_gate_data =
gate_data = True
gate_whiten = True
gate_tzero = 1.0
gate_tpad = 0.5
gate_threshold = 50.0
cluster_window = 0.5

[window_fft_specs]
window_fftgram = hann

[window_fft_welch_specs]
window_fftgram = hann
```

```
[density_estimation]
frequency_resolution = 0.03125
N_average_segments_psd = 2
coarse_grain_psd = False
coarse_grain_csd = True
overlap_factor_welch = 0.5
overlap_factor = 0.5

[postprocessing]
polarization = tensor
alpha = 0.0
fref = 25.0
flow = 20.0
fhigh = 500.0

[data_quality]
notch_list_path = /home/chia-hsuan.hsiung/pygwb_dir/pygwb/tut
calibration_epsilon = 0.0
alphas_delta_sigma_cut = ['-5', '0', '3', '5']
delta_sigma_cut = 0.2
return_naive_and_averaged_sigmas = False

[local_data]
local_data_path = H1://home/chia-hsuan.hsiung/practice_data/m
```

Additional parameter : custom_job_file=/path/to/custom_file

Pygwb_pipe

Starting point of the analysis.

./pygwb_pipe ––param_file parameter_file

```
igwn) [chia-hsuan.hsiung@ldas-pcdev7 0-5000]$ pygwb_pipe --param_file parameters_0-5000_final.ini
026-04-21 17:50:18.388 | INFO     | __main__:main:81 - Successfully imported parameters from paramfile and input.
026-04-21 17:50:18.403 | INFO     | __main__:main:123 - Saved final param file at ./parameters_0-5000_final_0-5000_final.ini.
eading (<built-in function format>): |                                                                                | 1/1 (100%) ETA 00:00
/cvmfs/software.igwn.org/conda/envs/igwn-py311-20260331/lib/python3.11/site-packages/gwpy/timeseries/timeseries.py:930: UserWarning: re
ample() rate matches current sample_rate (1024.0 Hz), returning input data unmodified; please double-check your parameters
  warnings.warn(
eading (<built-in function format>): |                                                                                | 1/1 (100%) ETA 00:00
026-04-21 17:50:20.941 | INFO     | __main__:main:129 - Loaded up interferometers with selected parameters.
026-04-21 17:50:20.941 | INFO     | __main__:main:150 - Applying autogating procedure.
026-04-21 17:50:22.917 | INFO     | __main__:main:160 - Gates for IFO H1:[]
026-04-21 17:50:24.878 | INFO     | __main__:main:160 - Gates for IFO L1:[]
026-04-21 17:50:24.878 | INFO     | __main__:main:163 - Baseline with interferometers H1, L1 initialised.
026-04-21 17:50:24.878 | INFO     | __main__:main:167 - Setting PSDs and CSDs of the baseline...
026-04-21 17:50:27.874 | INFO     | __main__:main:172 - ... done.
026-04-21 17:50:27.877 | DEBUG    | pygwb.baseline:calculate_delta_sigma_cut:1385 - loading notches for delta sigma cut from /home/chi
-hsuan.hsiung/pygwb_dir/pygwb/tutorials/output/Toy_example_HL_notchlist.txt
026-04-21 17:50:27.877 | DEBUG    | pygwb.baseline:set_frequency_mask:269 - loading notches from /home/chia-hsuan.hsiung/pygwb_dir/pyg
wb/tutorials/output/Toy_example_HL_notchlist.txt
026-04-21 17:50:28.231 | INFO     | pygwb.delta_sigma_cut:run_dsc:168 - Running delta sigma cut
026-04-21 17:50:31.188 | INFO     | __main__:main:190 - times flagged by the delta sigma cut as badGPStimes:[]
026-04-21 17:50:31.188 | INFO     | __main__:main:201 - calculating the point estimate and sigma...
026-04-21 17:50:31.188 | INFO     | pygwb.baseline:set_point_estimate_sigma:1268 - Point estimate and sigma spectra have not been set
before. Setting it now...
026-04-21 17:50:31.188 | DEBUG    | pygwb.baseline:set_point_estimate_sigma:1271 - No weighting supplied in setting of spectrum. Suppl
ed when combining for final sigma
026-04-21 17:50:31.188 | INFO     | pygwb.baseline:set_point_estimate_sigma_spectrum:1134 - Point estimate and sigma spectrograms are
ot set yet. setting now...
026-04-21 17:50:31.227 | INFO     | pygwb.baseline:set_point_estimate_sigma_spectrum:1149 - 0 bad segments removed.
```

Short summary —pygwb_pipe

What you will see when it’s finished

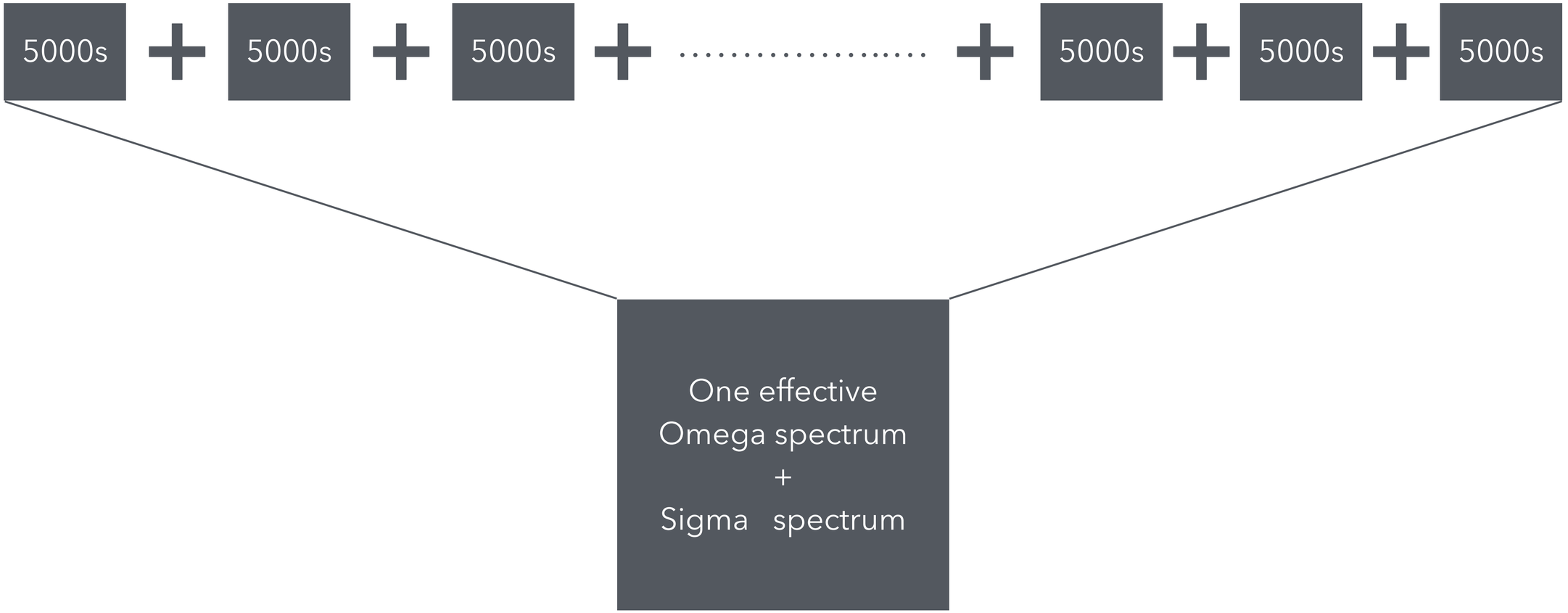
```
2026-04-21 17:50:31.227 | INFO     | pygwb.baseline:set_point_estimate_sigma_spectrum:1149 - 0 bad segments removed.
2026-04-21 17:50:31.227 | DEBUG    | pygwb.baseline:set_frequency_mask:269 - loading notches from /home/chia-hsuan.hsiung/pygwb_dir/pyg
wb/tutorials/output/Toy_example_HL_notchlist.txt
2026-04-21 17:50:31.330 | DEBUG    | pygwb.postprocessing:postprocess_Y_sigma:181 - Bias factor: 1.0466610426846534
2026-04-21 17:50:31.331 | DEBUG    | pygwb.baseline:set_frequency_mask:269 - loading notches from /home/chia-hsuan.hsiung/pygwb_dir/pyg
wb/tutorials/output/Toy_example_HL_notchlist.txt
2026-04-21 17:50:31.709 | SUCCESS  | __main__:main:211 - Ran stochastic search over times 0-5000
2026-04-21 17:50:31.709 | SUCCESS  | __main__:main:214 -        POINT ESTIMATE: -8.665371e-08
2026-04-21 17:50:31.709 | SUCCESS  | __main__:main:215 -        SIGMA: 8.266330e-08
2026-04-21 17:50:31.710 | INFO     | __main__:main:219 - Saving point_estimate and sigma spectrograms, spectra, and final values to fil
e.
2026-04-21 17:50:31.710 | INFO     | __main__:main:222 - Saving average psds and csd to file.
```

With two outputs

```
parameters_0-5000_final.ini  point_estimate_sigma_0-5000.npz  psds_csds_0-5000.npz
```

Pygwb_combine

Combining all thunks of results into Omega spectrum and sigma spectrum.



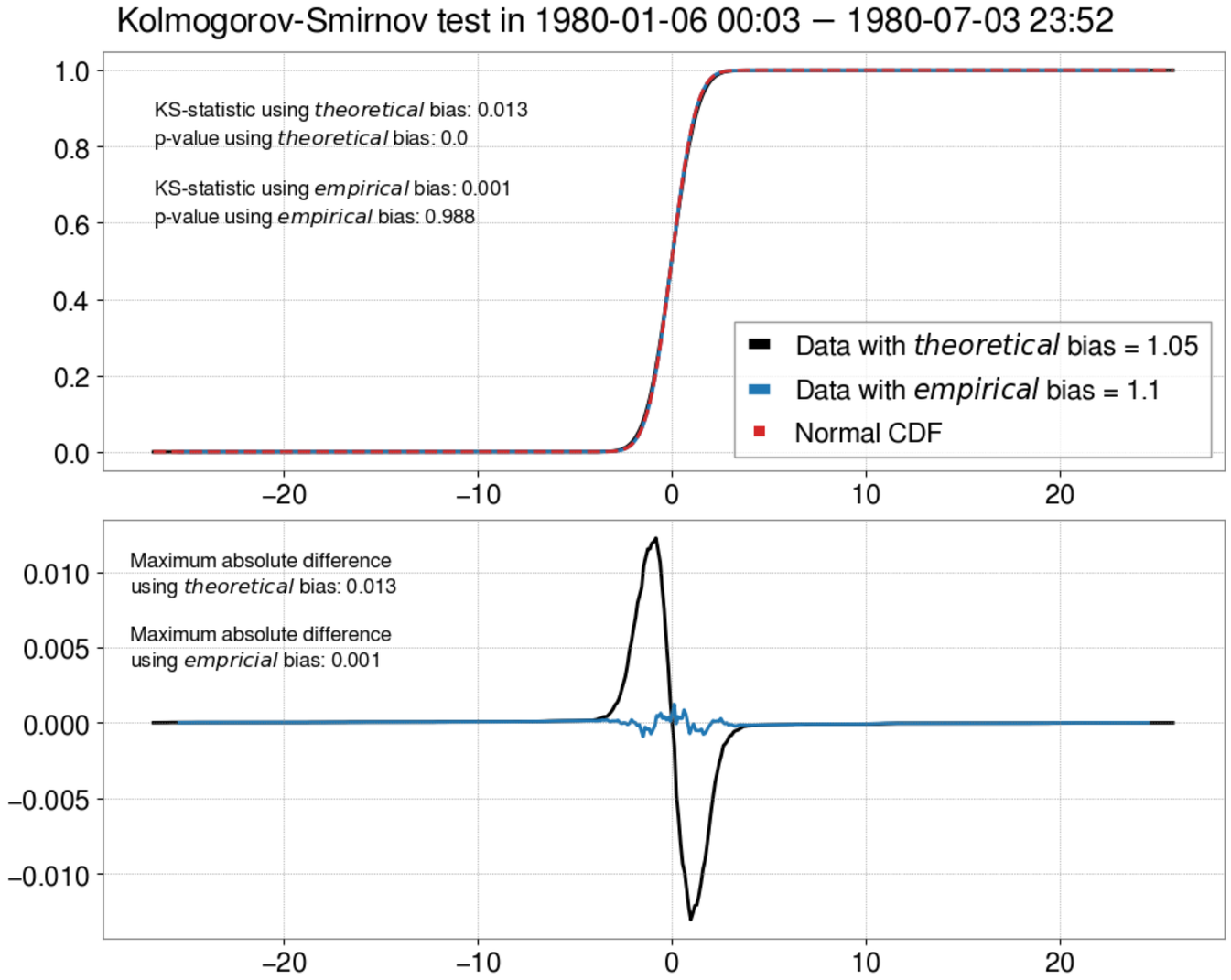
Pygwb_combine

```
2026-04-21 16:56:20.652 | INFO     | __main__:main:199 - Unpacking files...
^M  0%|          | 0/4 [00:00<?, ?it/s]^M 25%|██        | 1/4 [00:00<00:02,  1.10it/s]^M100%|██████████| 4/4 [00:00<00:00,  4.17it/s]
2026-04-21 16:56:21.623 | INFO     | __main__:main:245 - saved file with unweighted point estimate and sigma values for all times in ru
n:
 /home/chia-hsuan.hsiung/runuing_test/isotropic/running_file/test1/output/combined_results/point_estimate_sigma_0-110000_UNWEIGHTED.npz
.
2026-04-21 16:56:21.662 | INFO     | __main__:main:284 - saved file with all sigma information related to the delta sigma cut for all t
imes in run:
 /home/chia-hsuan.hsiung/runuing_test/isotropic/running_file/test1/output/combined_results/delta_sigma_cut_0-110000.npz.
2026-04-21 16:56:21.698 | INFO     | __main__:main:325 - Saved file with combined point estimate and sigma OmegaSpectrum objects for th
is run. These are weighted with alpha=0.0
2026-04-21 16:56:21.699 | INFO     | __main__:main:339 - Final point estimate re-weighted with alpha=0.00 at reference frequency fref=2
5 with h0=0.6766:
 [-1.8393811914036025e-08 +/- 1.7645669225191066e-08]
2026-04-21 16:56:22.291 | INFO     | __main__:main:393 - Saved file with re-weighted point estimate and sigma values and spectra:
 /home/chia-hsuan.hsiung/runuing_test/isotropic/running_file/test1/output/combined_results/point_estimate_sigma_spectra_alpha_0.0_fref_
25_0-110000.npz.
2026-04-21 16:56:22.320 | INFO     | __main__:main:408 - Combining coherences over all files...
^M  0%|          | 0/4 [00:00<?, ?it/s]^M100%|██████████| 4/4 [00:00<00:00, 76.97it/s]
2026-04-21 16:56:22.388 | INFO     | __main__:main:447 - Saved file with coherence spectrum:
 /home/chia-hsuan.hsiung/runuing_test/isotropic/running_file/test1/output/combined_results/coherence_spectrum_0-110000.npz.
```

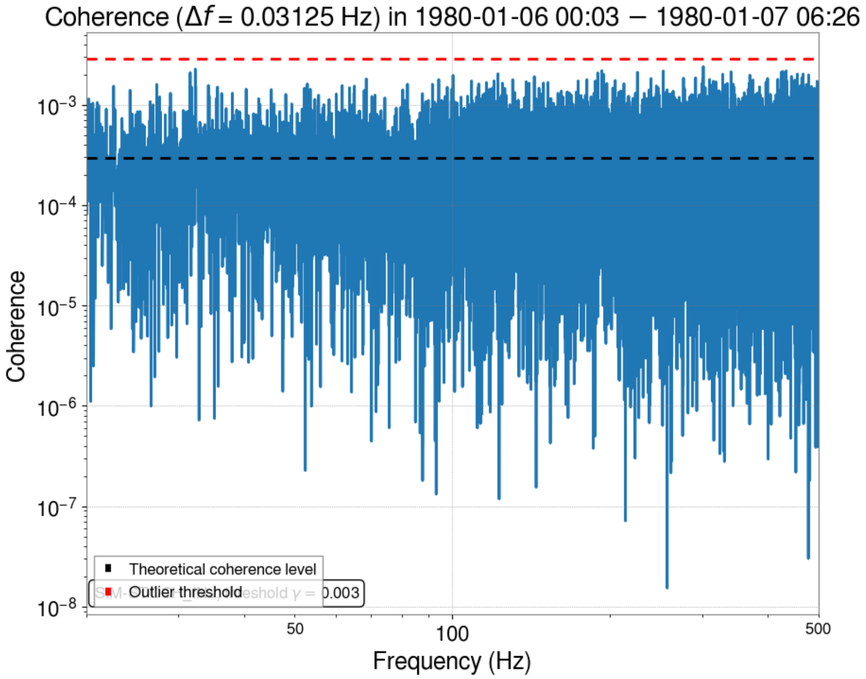
Pygwb_pe

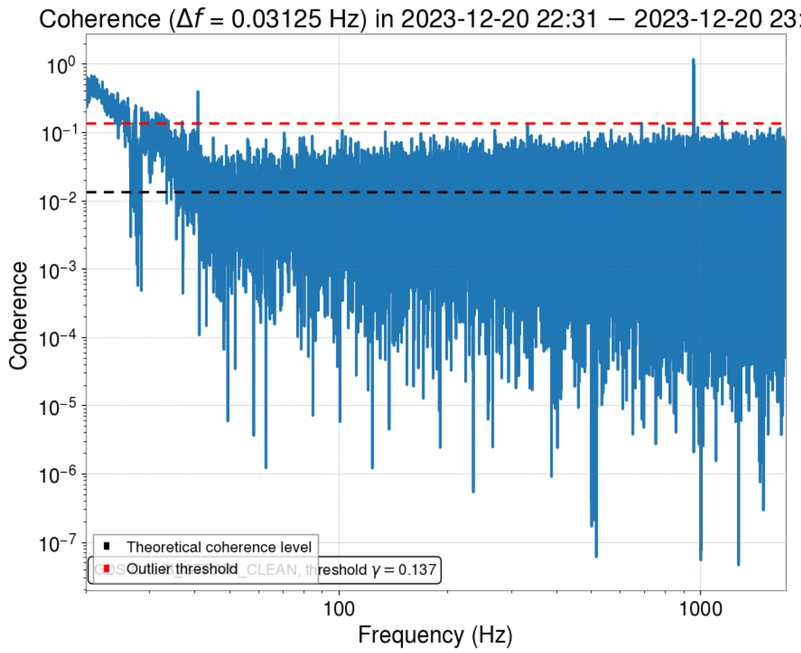
/home/chia-hsuan.hsiung/pygwb_dir/pygwb/tutorials/run_pe.ipynb

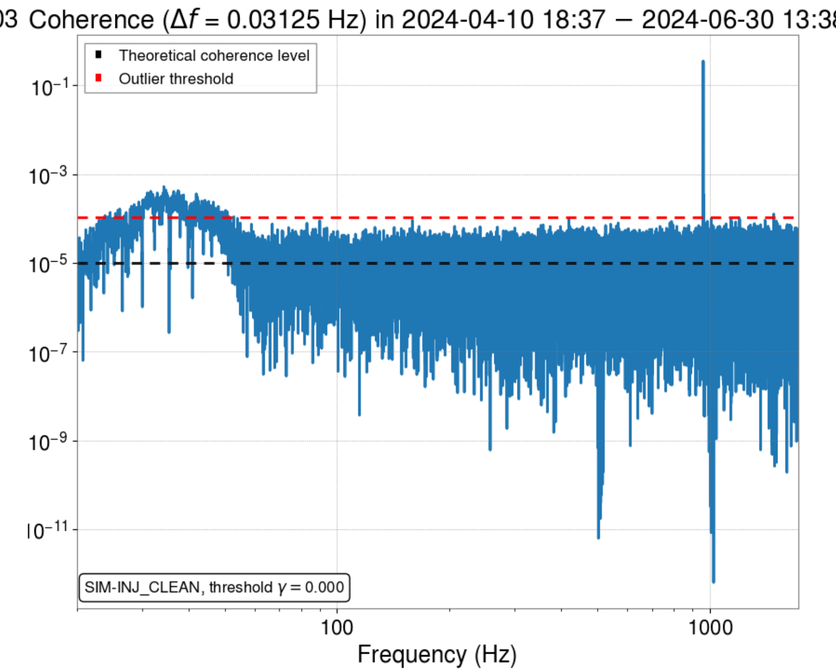
Pygwb_stat


Kolmogorov-Smirnov test in 1980-01-06 00:03 – 1980-07-03 23:52
KS-statistic using theoretical bias: 0.013
p-value using theoretical bias: 0.0
KS-statistic using empirical bias: 0.001
p-value using empirical bias: 0.988
Data with theoretical bias = 1.05
Data with empirical bias = 1.1
Normal CDF
1.0
0.8
0.6
0.4
0.2
0.0
–20
–10
0
10
20
Maximum absolute difference
using theoretical bias: 0.013
Maximum absolute difference
using empricial bias: 0.001
0.010
0.005
0.000
–0.005
–0.010
–20
–10
0
10
20

Pygwb_stat





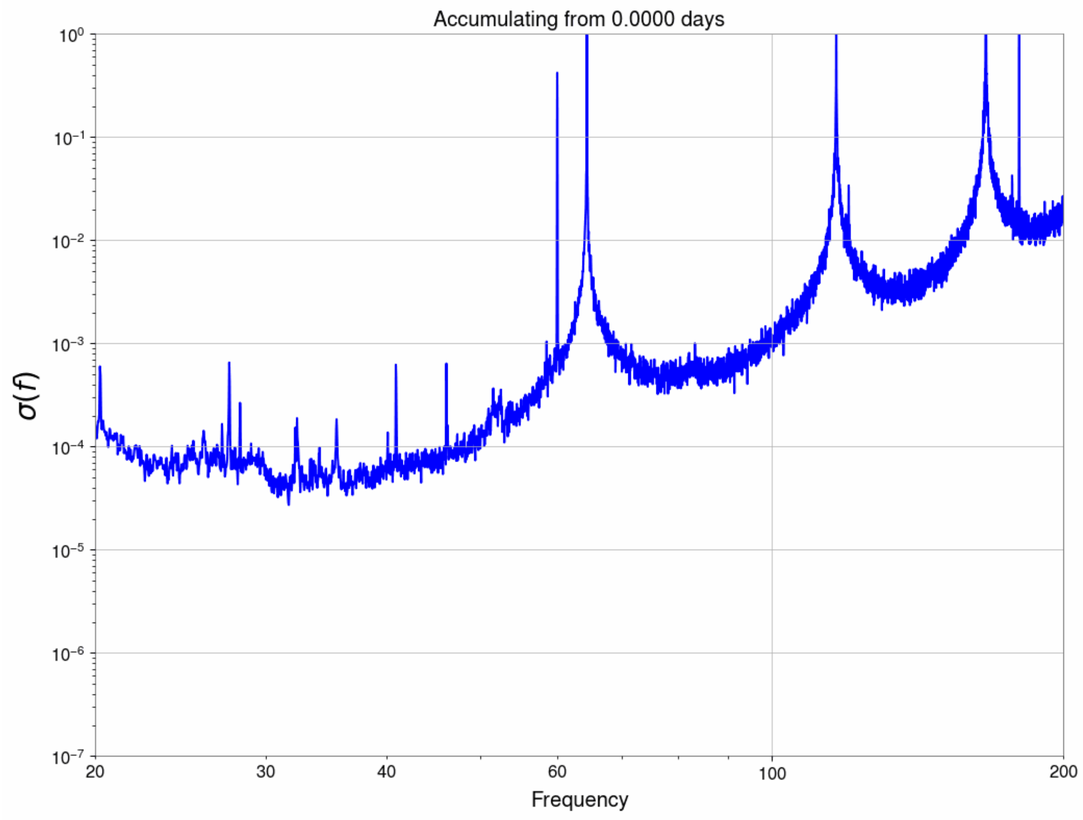


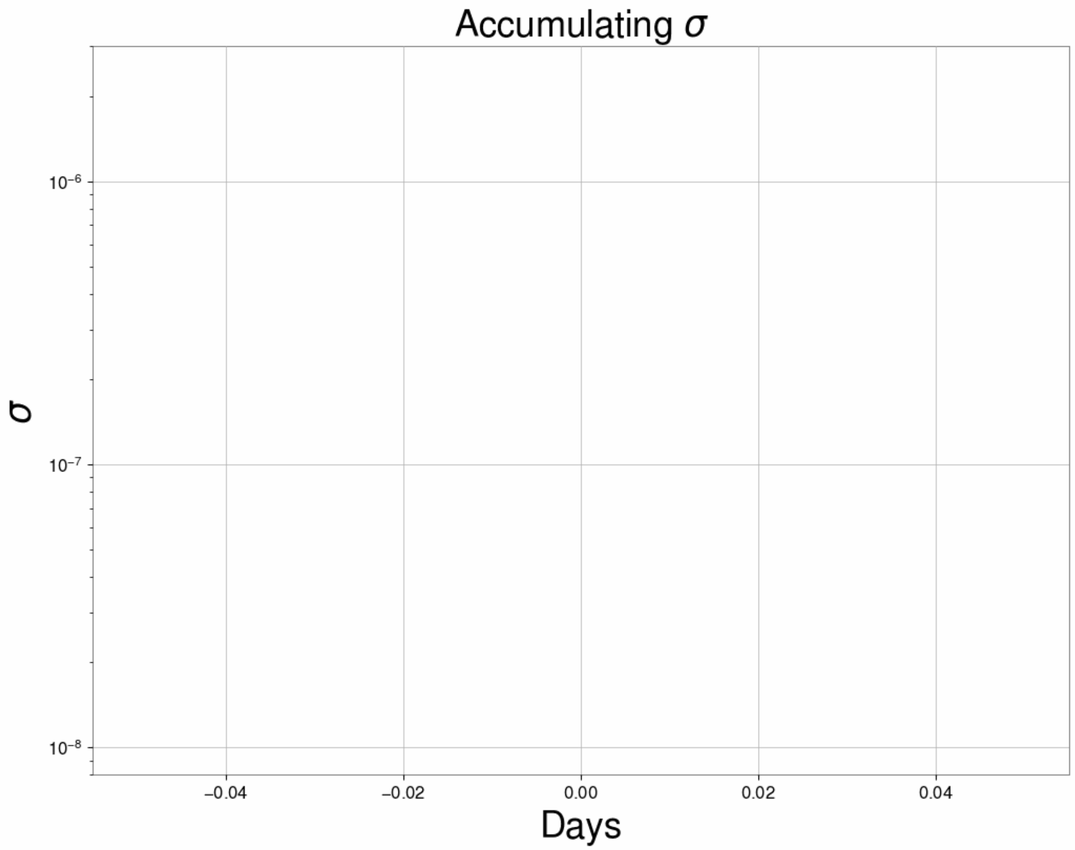
Pure noise

Hardware injection

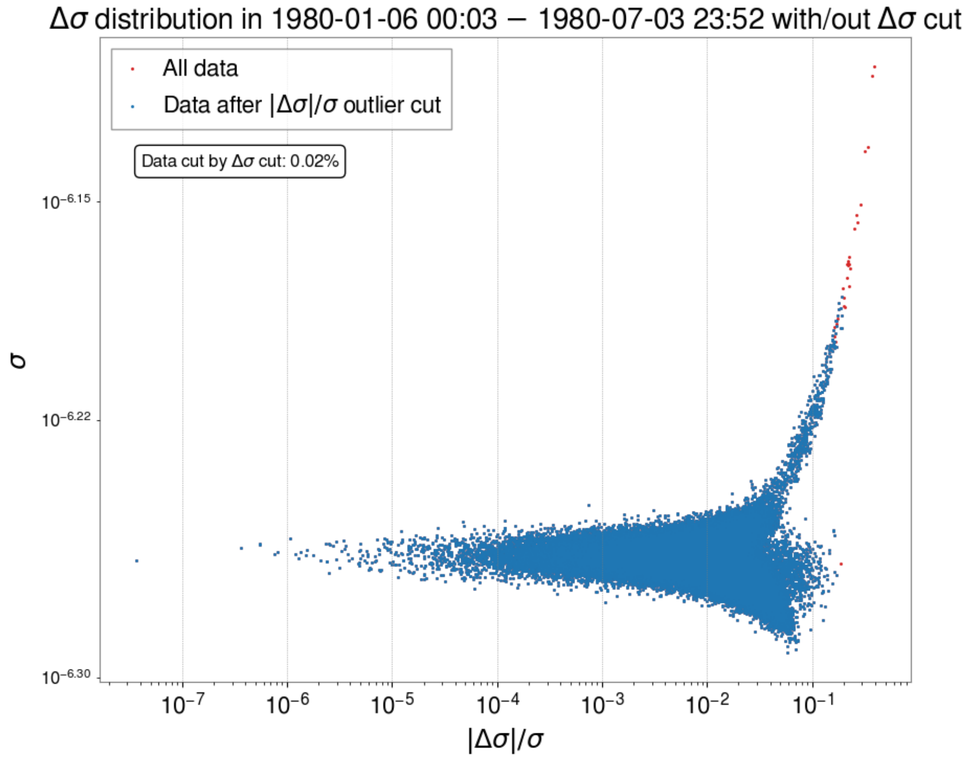
Software injection

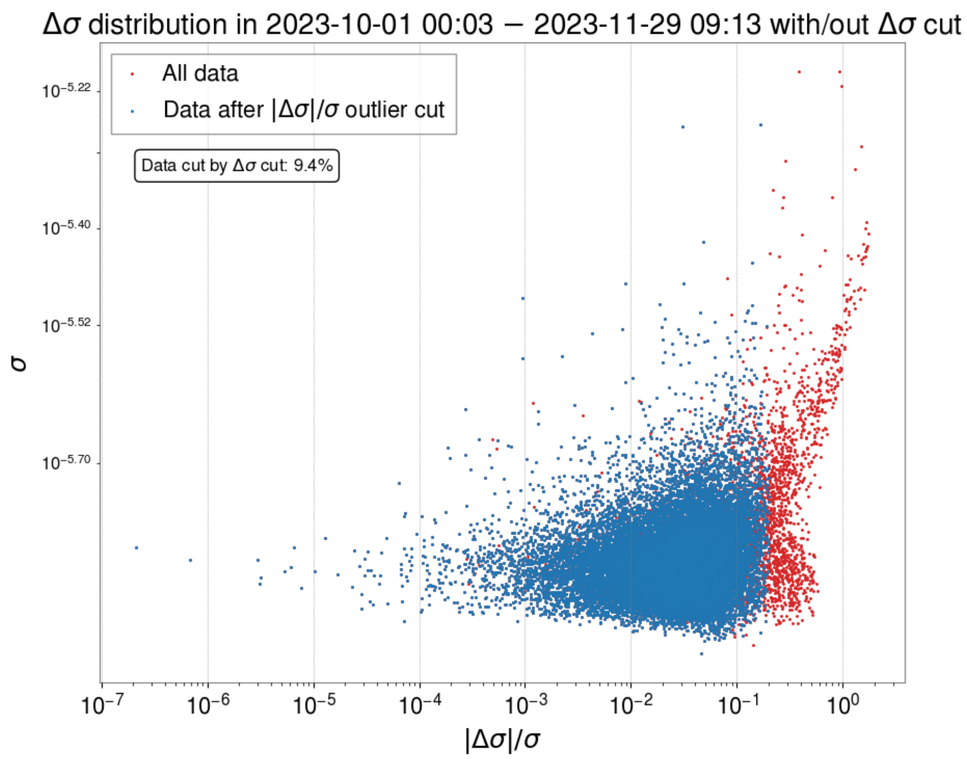
Pygwb_stat


Accumulating from 0.0000 days
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
σ(f)
20
30
40
60
100
200
Frequency


Accumulating σ
10^{-6}
10^{-7}
10^{-8}
σ
−0.04
−0.02
0.00
0.02
0.04
Days

Pygwb_stat

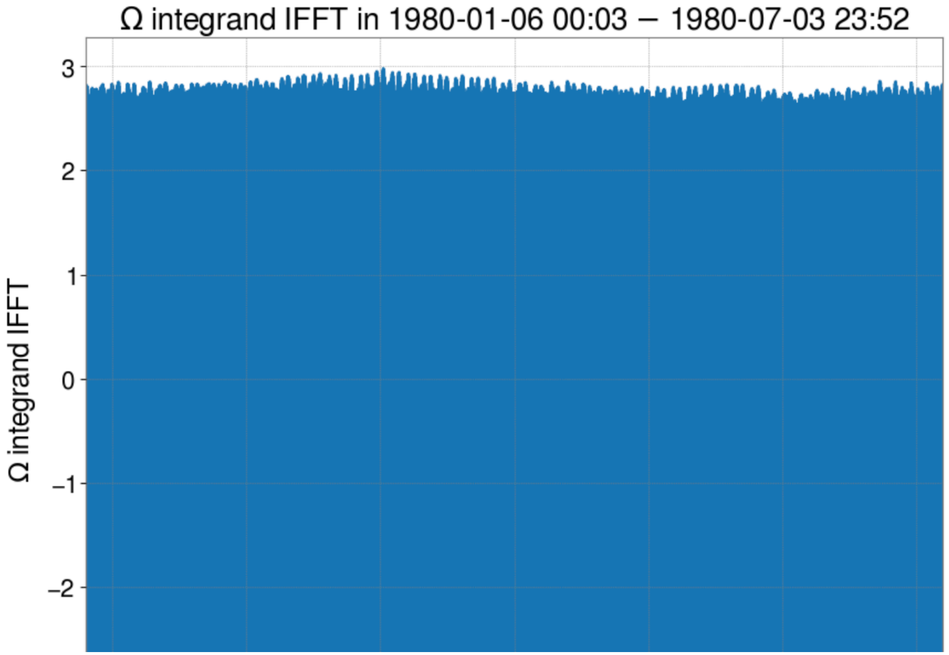

Δσ distribution in 1980-01-06 00:03 – 1980-07-03 23:52 with/out Δσ cut
All data
Data after |Δσ|/σ outlier cut
Data cut by Δσ cut: 0.02%
σ
10^-6.15
10^-6.22
10^-6.30
10^-7
10^-6
10^-5
10^-4
10^-3
10^-2
10^-1
|Δσ|/σ

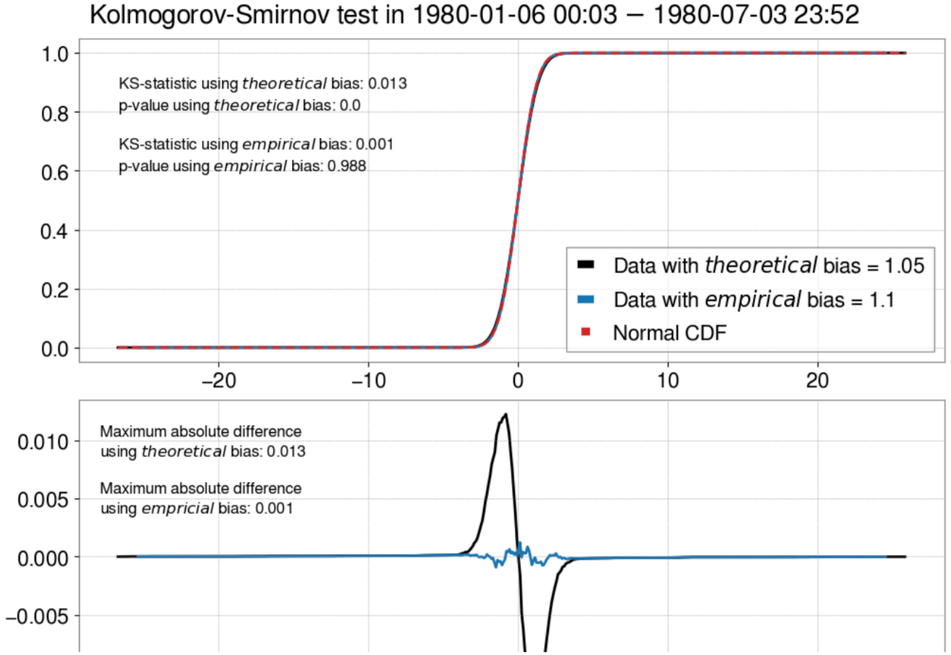

Δσ distribution in 2023-10-01 00:03 – 2023-11-29 09:13 with/out Δσ cut
All data
Data after |Δσ|/σ outlier cut
Data cut by Δσ cut: 9.4%
σ
10^-5.22
10^-5.40
10^-5.52
10^-5.70
10^-7
10^-6
10^-5
10^-4
10^-3
10^-2
10^-1
10^0
|Δσ|/σ

Pygwb_html

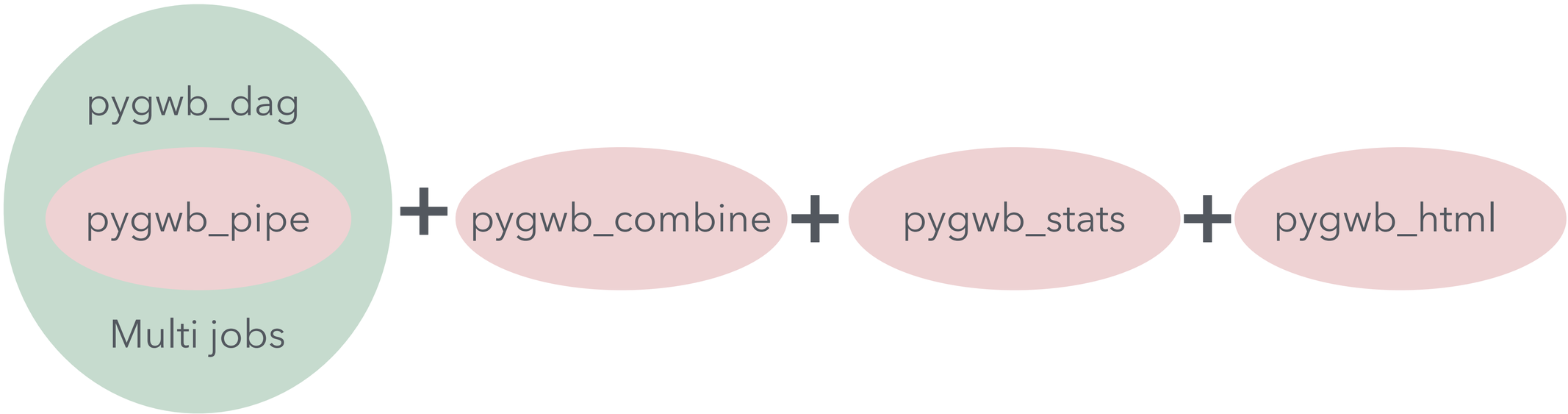
pygwb combined results

pygwb combined results

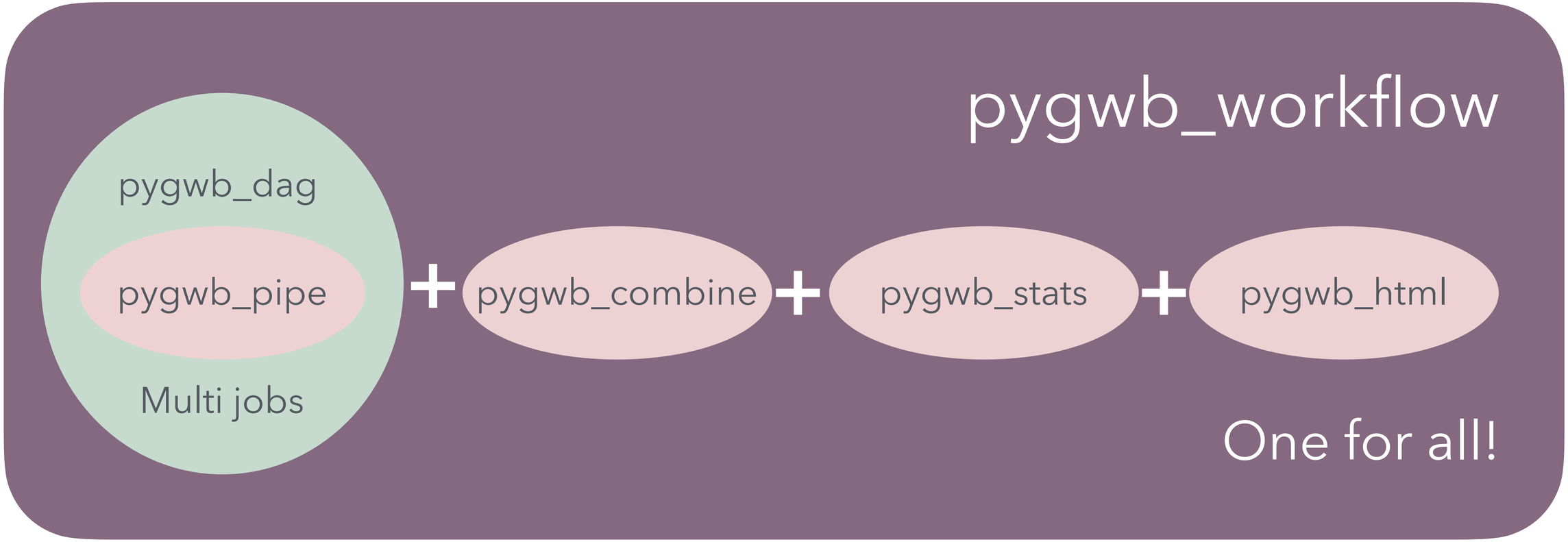




Pygwb_dag


pygwb_dag
pygwb_pipe
Multi jobs
+
pygwb_combine
+
pygwb_stats
+
pygwb_html

Pygwb_workflow



Pygwb_workflow

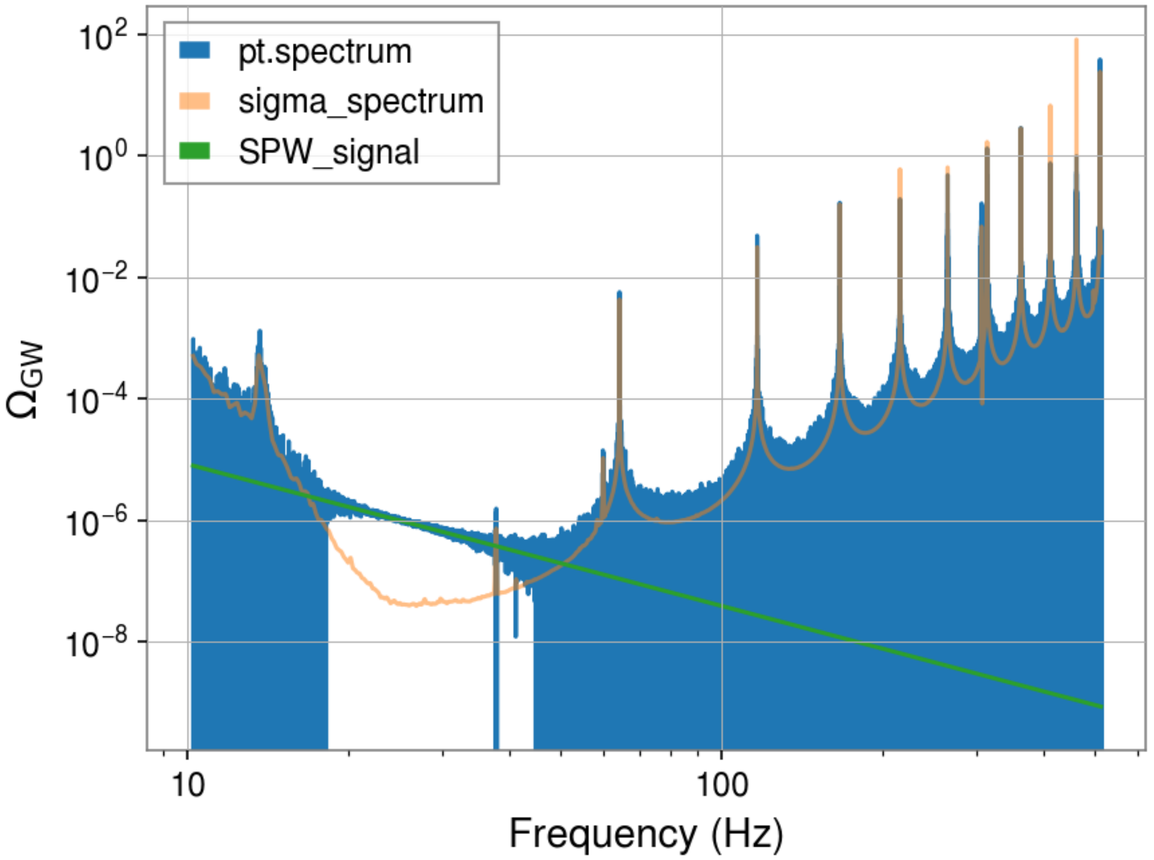
```
[general]
accounting_group = ligo.dev.o4.sgwb.isotropic.stochastic
ifos = H1 L1
plot_segment_results = False
max_job_dur = 5000
min_job_dur = 600
t0 = 1368975618
tf = 1389456018
combine_factor = 12

[executables]
pygwb_pipe = pygwb_pipe
pygwb_combine = pygwb_combine
pygwb_stats = pygwb_stats
pygwb_html = pygwb_html

[data_quality]
science_segment = DMT-ANALYSIS_READY
veto_definer = /home/hsiung.chiahsuan/public_html/isotropic-o4-bimonthly/H1L1-HOFT_C00_O4_CBC.xml
```

Job file is created automatically even considering the data flag and VDF.

Simulation



Simple power law

Jupyter-notebook

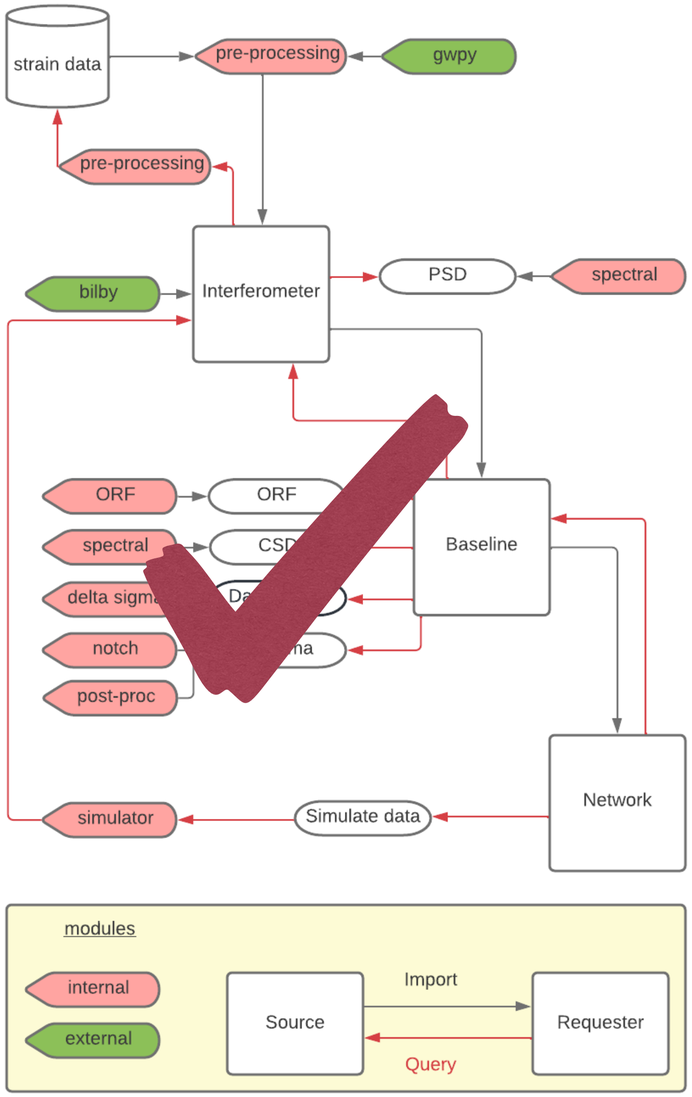
`/home/chia-hsuan.hsiung/simulating_gaussian_background.ipynb`

HTCondor

`/home/chia-hsuan.hsiung/script-dir/1m1e-7/make_data.sh`

`/home/chia-hsuan.hsiung/script-dir/1m1e-7/hello.submit`

Summary



You are here now