UNIVERSITÀ DEGLI STUDI DI TRIESTE

LIGO VIRGO KAGRA

INFN Istituto Nazionale di Fisica Nucleare

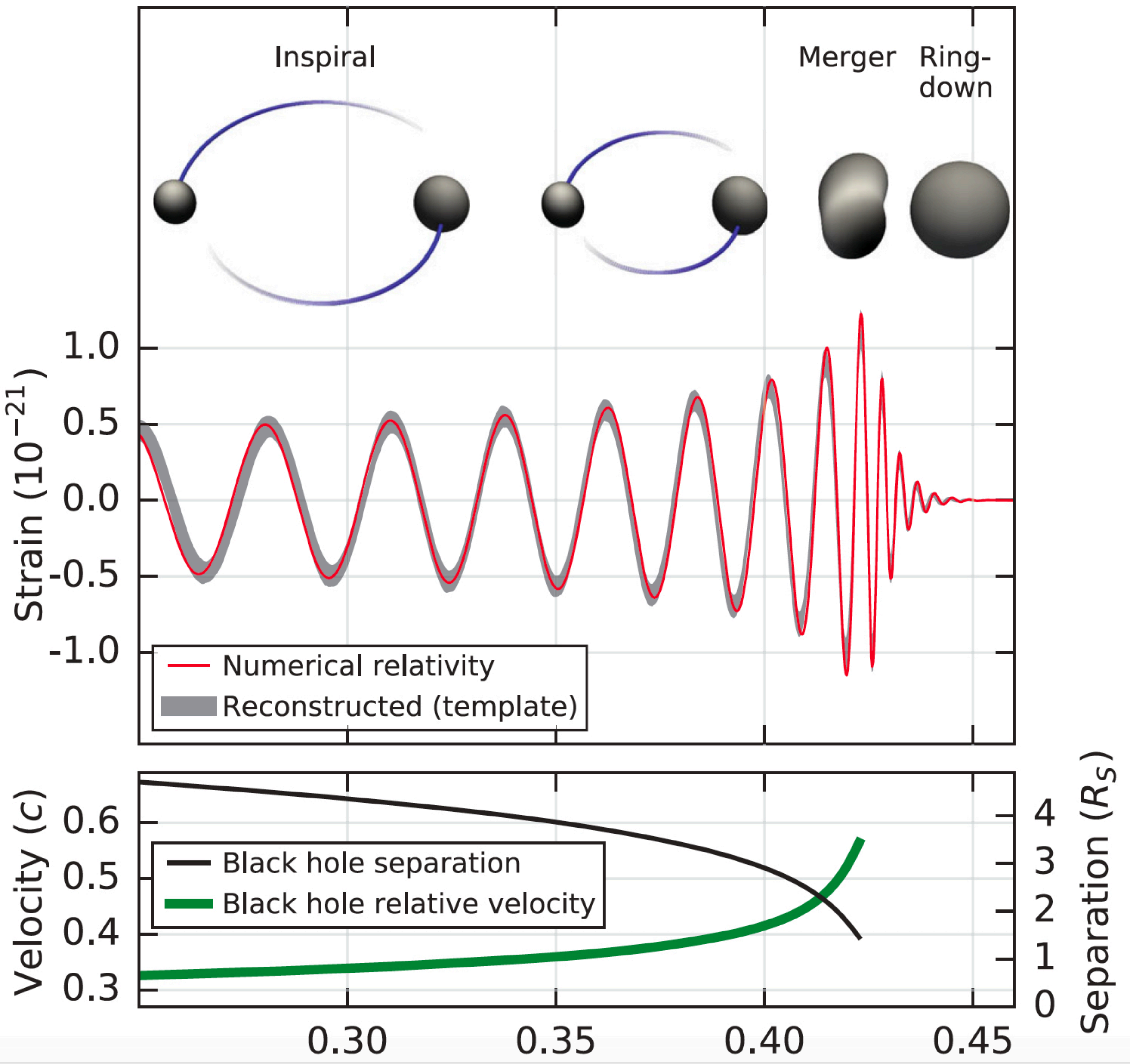
# Gravitational-Wave (Open) Data

A. Trovato on behalf of the LVK Collaboration
Università di Trieste, INFN-Sezione Trieste

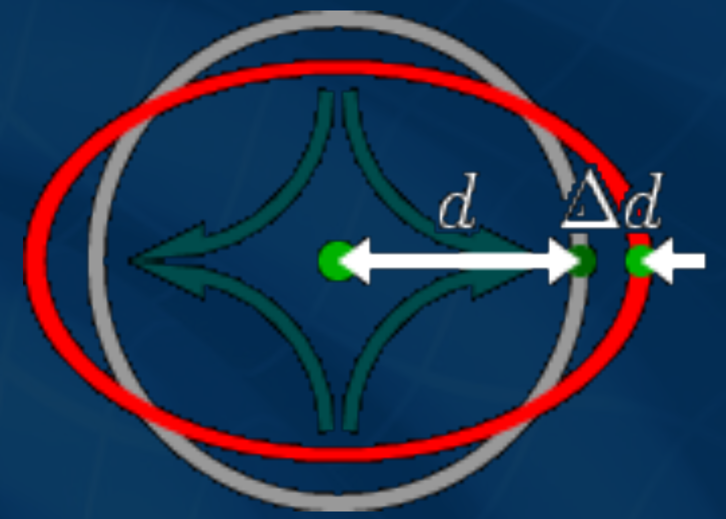
References:

- Gravitational Wave Open Science Center: www.gwosc.org
- Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo: SoftwareX 13 (2021) 100658
- Open data from the third observing run of LIGO, Virgo, KAGRA and GEO: arXiv:2302.03676

# Gravitational waves (GW)

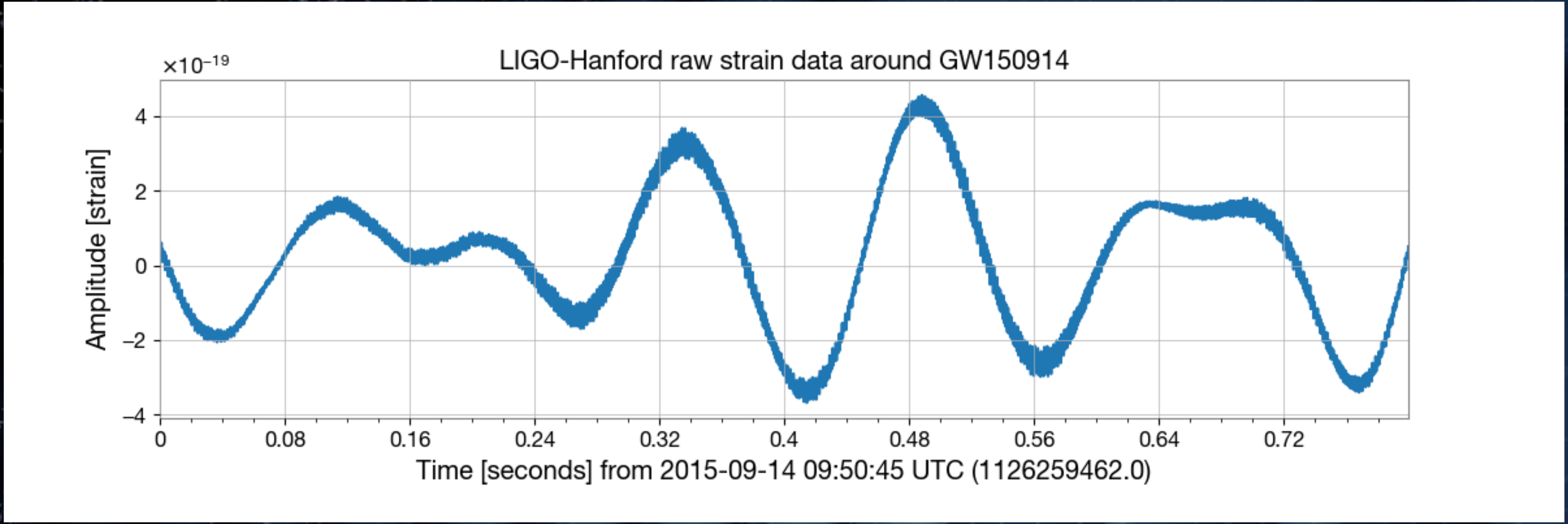



Reminder: strain



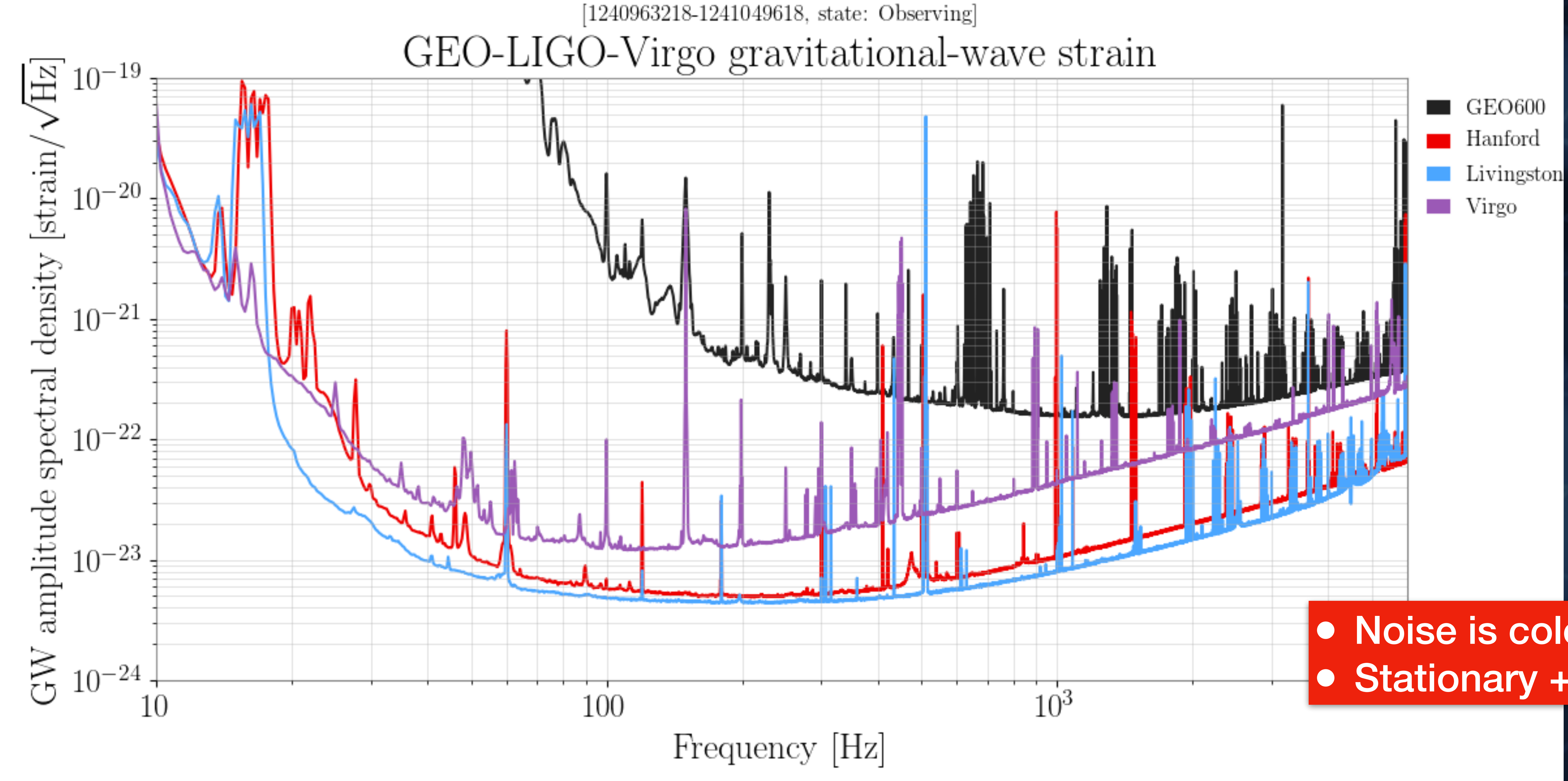

$$h = \frac{\Delta d}{d} = \frac{\textit{change in relative position}}{\textit{separation}}$$

# Recorded data




- Data are time series: stain data h(t) are sampled at 16 kHz (16000 values for each second!) —> main data but not the only type of data recorded
- Great part of data consist just of noise with eventually a very faint GW signal hidden in it
- GW signals are in an audible frequency ( see e.g. Black Hole Hunter)

# Complex noise




- Noise is coloured
- Stationary + transient noise

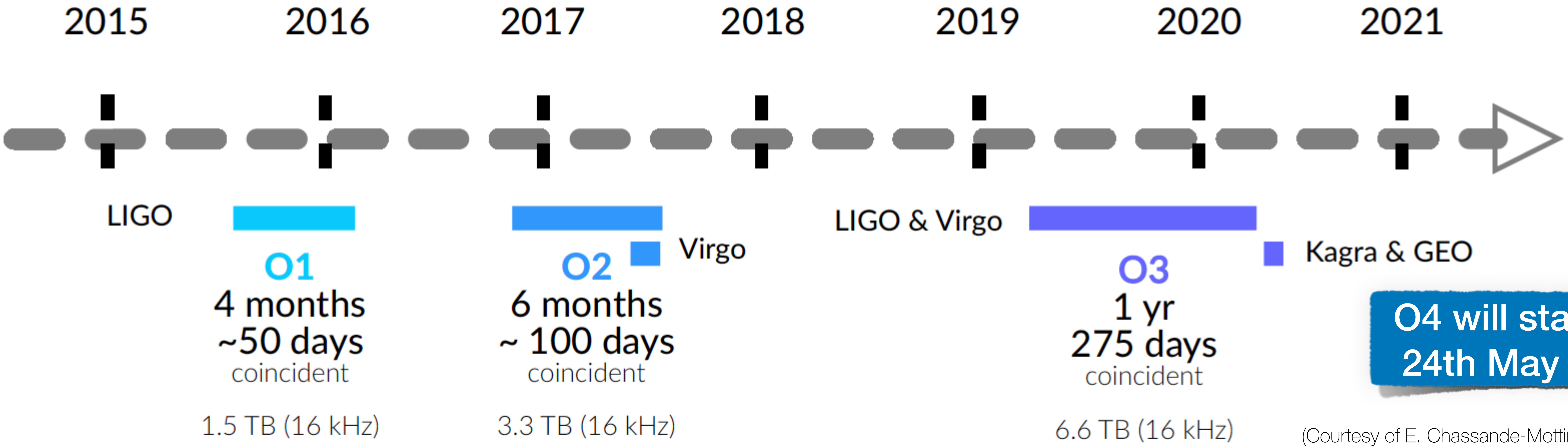
- **GW amplitude spectral density**: which is the amount of noise for each frequency?
- Data needs to be whitened (divided by the ASD) to normalise each frequency contribution

$$d(t) \xrightarrow{\text{FFT}} \tilde{d}(f) \xrightarrow{\text{Whiten}} \tilde{d}_w(f) = \frac{\tilde{d}(f)}{S_n^{1/2}(f)} \xrightarrow{\text{iFFT}} d_w(t)$$

Keep this in mind for the tutorials ;-)

# Observing runs

The detectors alternate periods of data taking and periods of upgrades to the machines

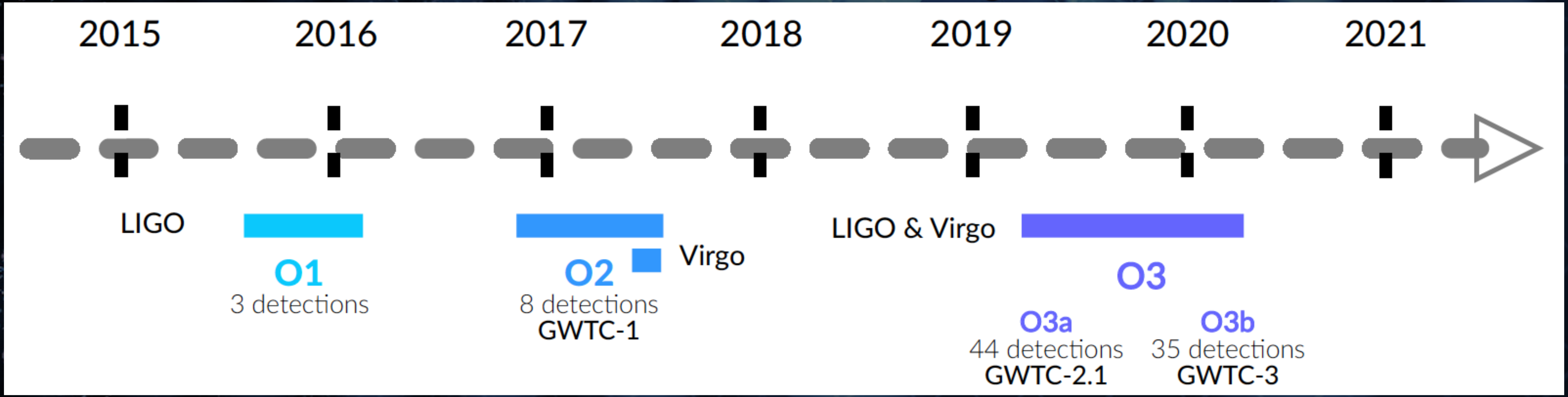




O4 will start the 24th May 2023

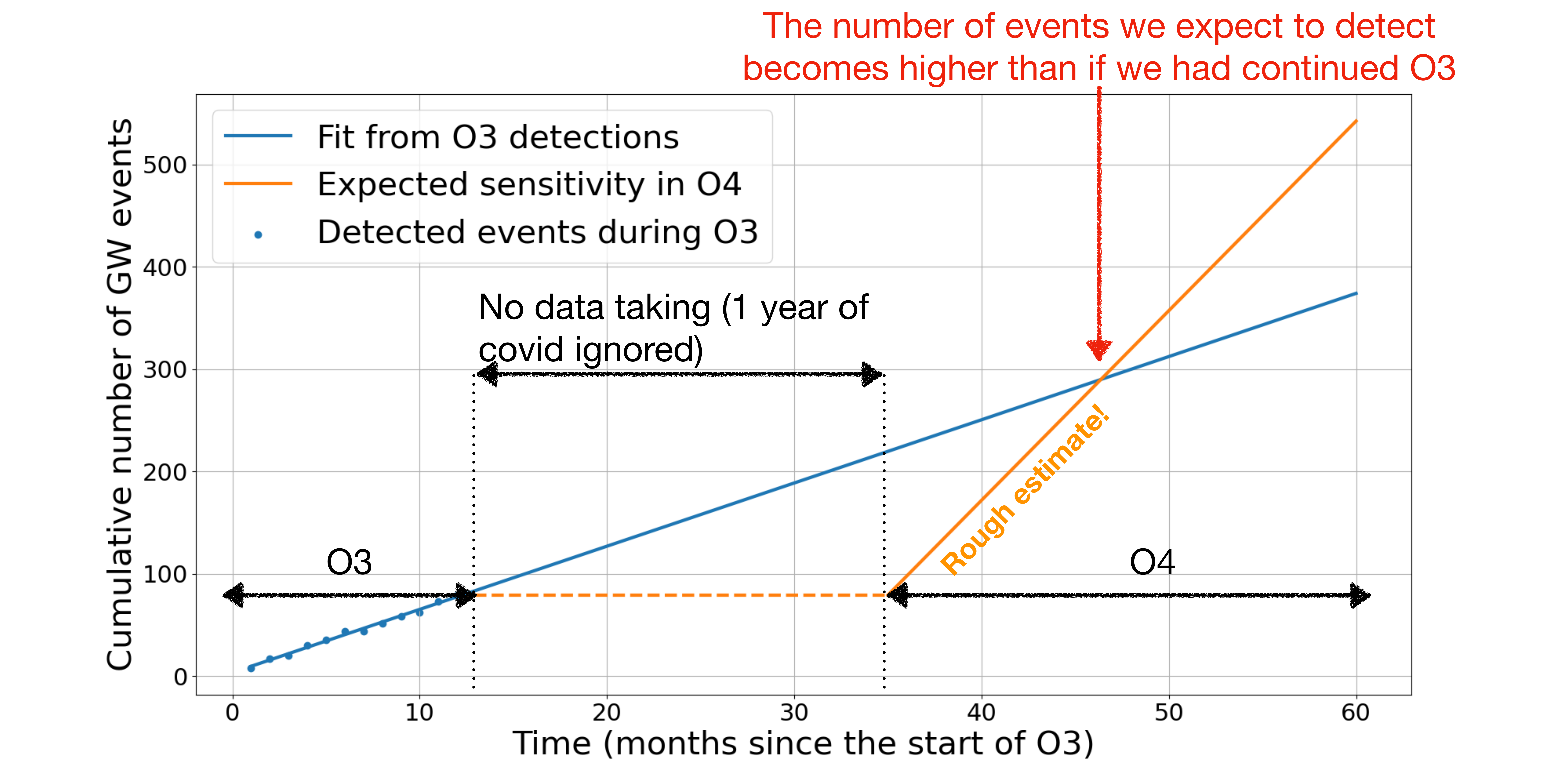
(Courtesy of E. Chassande-Mottin)

# GW detections and catalogs

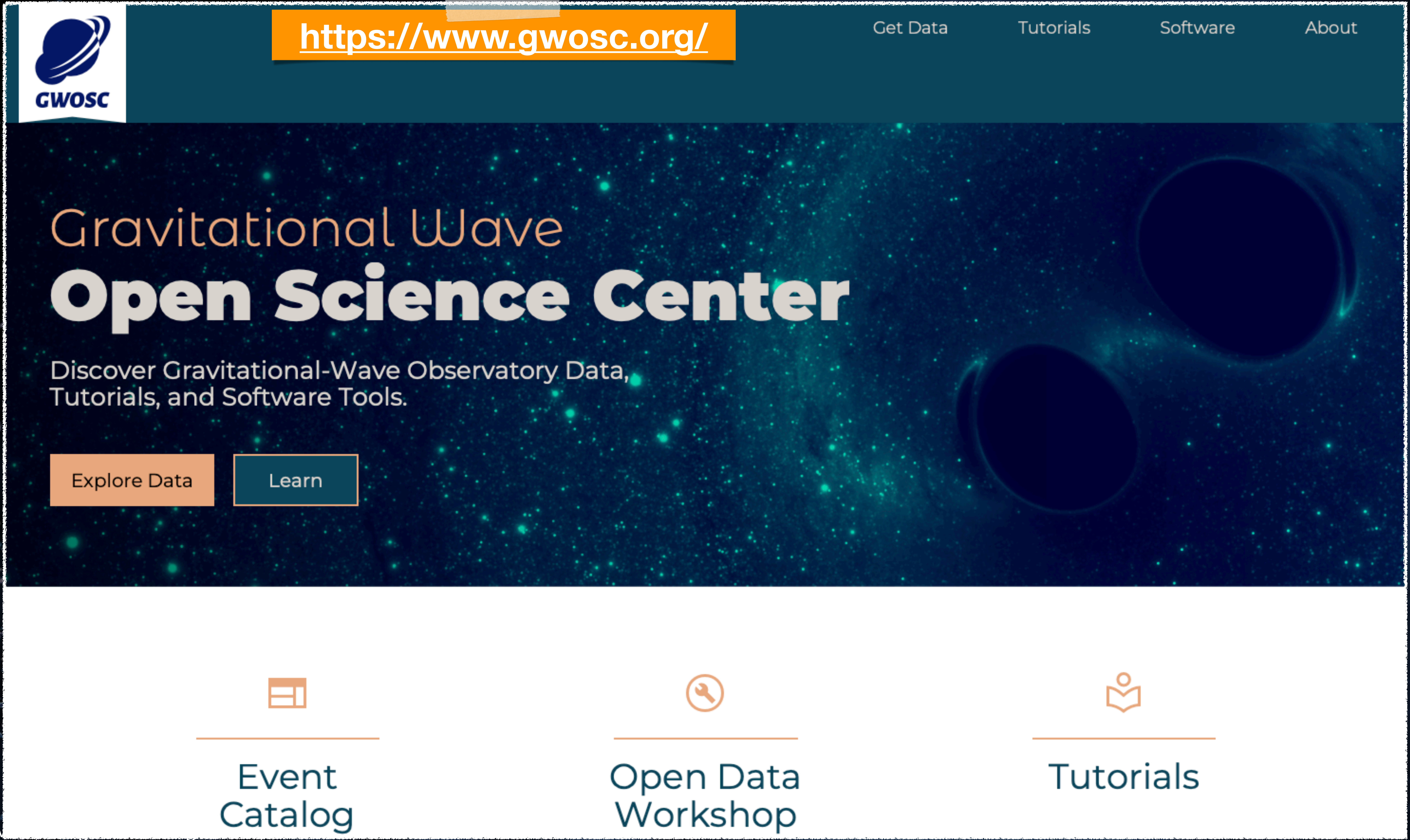
- A total of 90 GW events have been collected until now
- Events are named after the date of their discovery
- Events are grouped in catalogs named “Gravitational Wave Transient Catalog” (GWTC)



# Why on/off data taking?

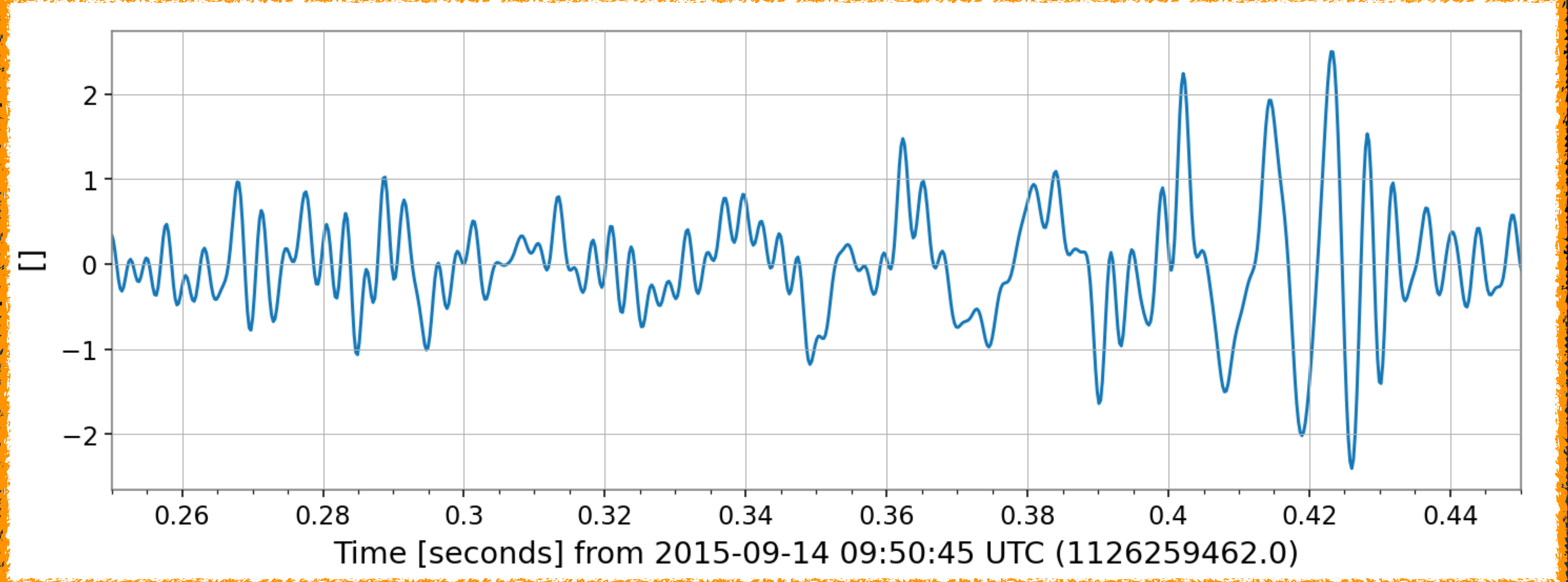


https://www.gwosc.org/



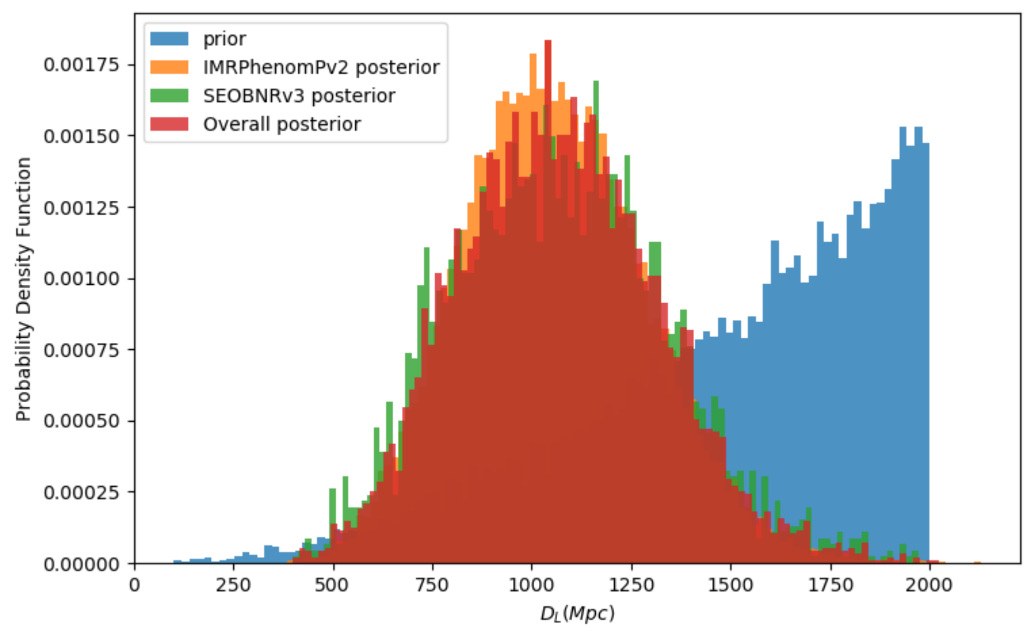
# Available data products

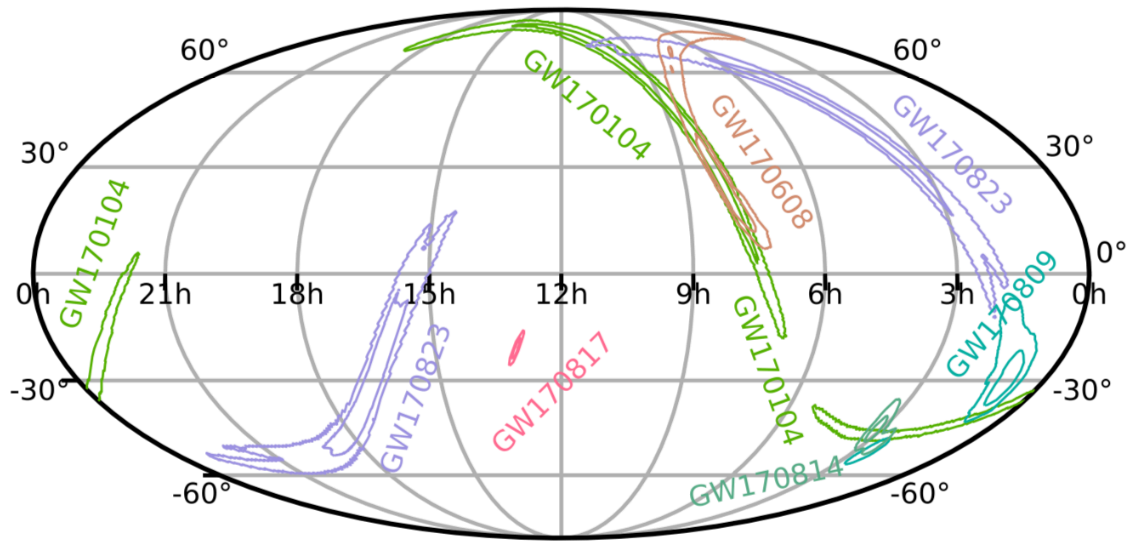
## Strain Data




## Analysis Results

(mainly related to events or event catalogs)




## Segments (Timelines)

| Start | Stop | Duration |
|---|---|---|
| 1164559440 | 1164559654 | 214 |
| 1164560599 | 1164561392 | 793 |
| 1164562093 | 1164569775 | 7682 |

# GWOSC releases

Two different types of strain data releases (data page)

Data taken during a whole observation run

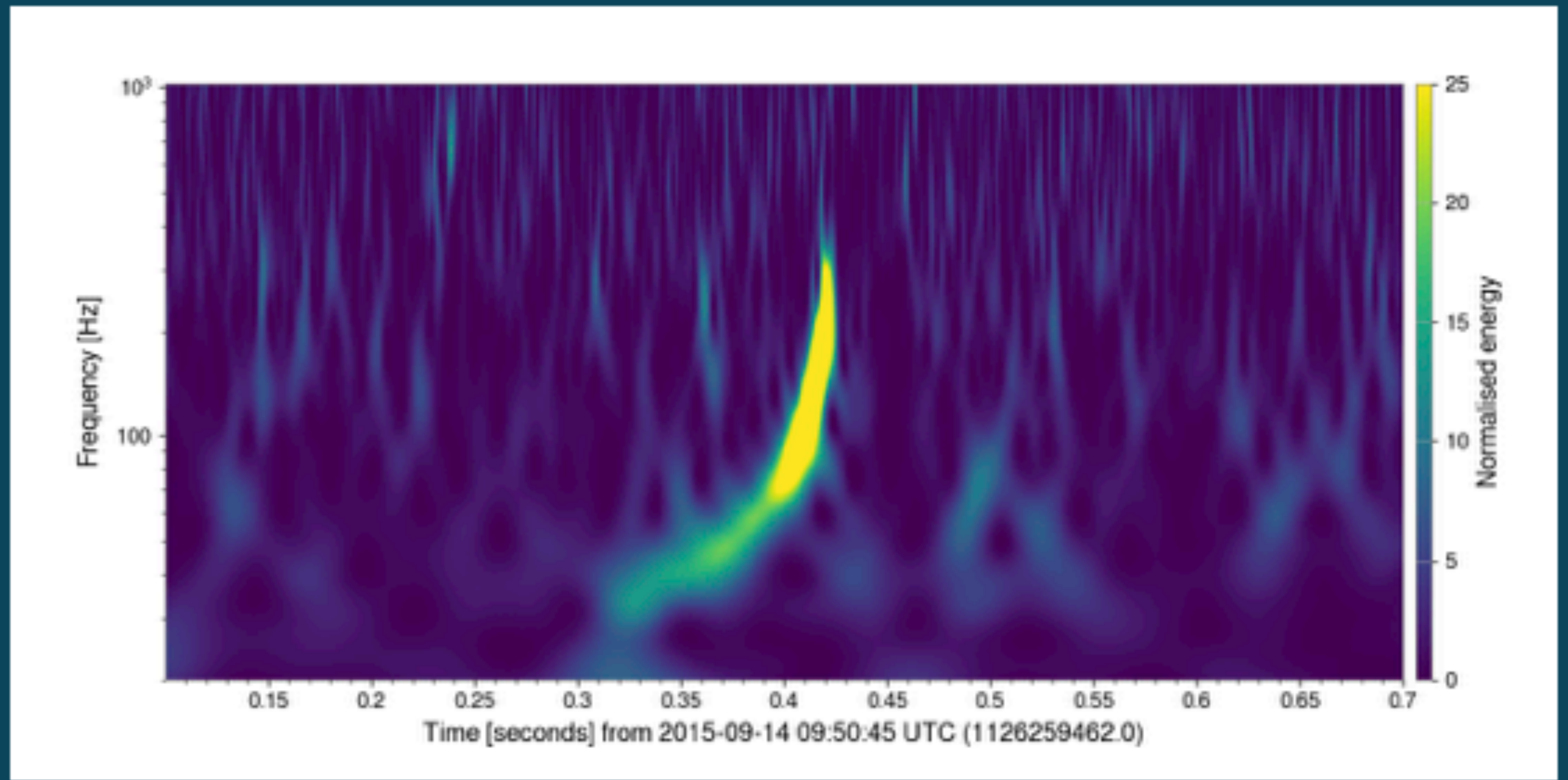
Gravitational wave data surrounding discoveries

O3b Data Release

4 kHz Data | 16 kHz Data | Documents | Timeline

**O3b Time Range**: November 1, 2019 through March 27, 2020

**Detectors**: H1, L1 and V1




| | | | |
|---|---|---|---|
| 32sec · 16KHz: | GWF | HDF | TXT |
| 32sec · 4KHz: | GWF | HDF | TXT |
| 4096sec · 16KHz: | GWF | HDF | TXT |
| 4096sec · 4KHz: | GWF | HDF | TXT |

For details on LVK open data see:

- O1+O2: SoftwareX 13 (2021) 100658
- O3: arXiv:2302.03676

# Example of a run page

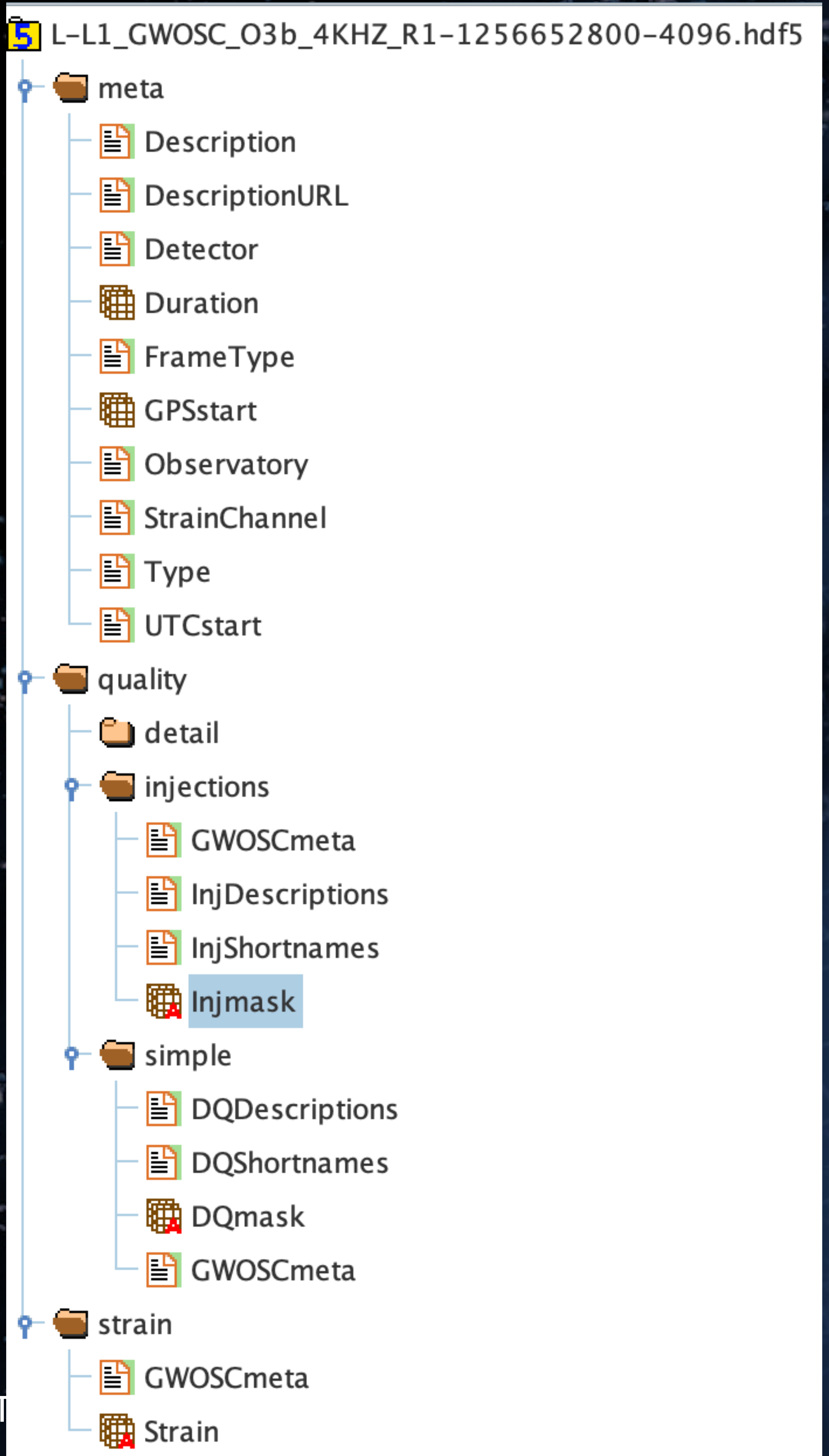
This page lists:

Link to this page

- Files that are available from the **O3b_4KHZ_R1** dataset
- Times from 1256655618 to 1269363618
- **L1** detector.
- Each file covers a 4096-second period, with strain data at either 16kHz or downsampled to 4 kHz.
- File sizes are shown in megabytes.
- The time of the beginning of the file is shown as 'GPS start time', and is linked to a timeline showing which parts of the tile have science-mode data.
- The last column of the table shows the percentage of each file that has data.
- For instructions on downloading many files, see the Automatic Download Tutorial.

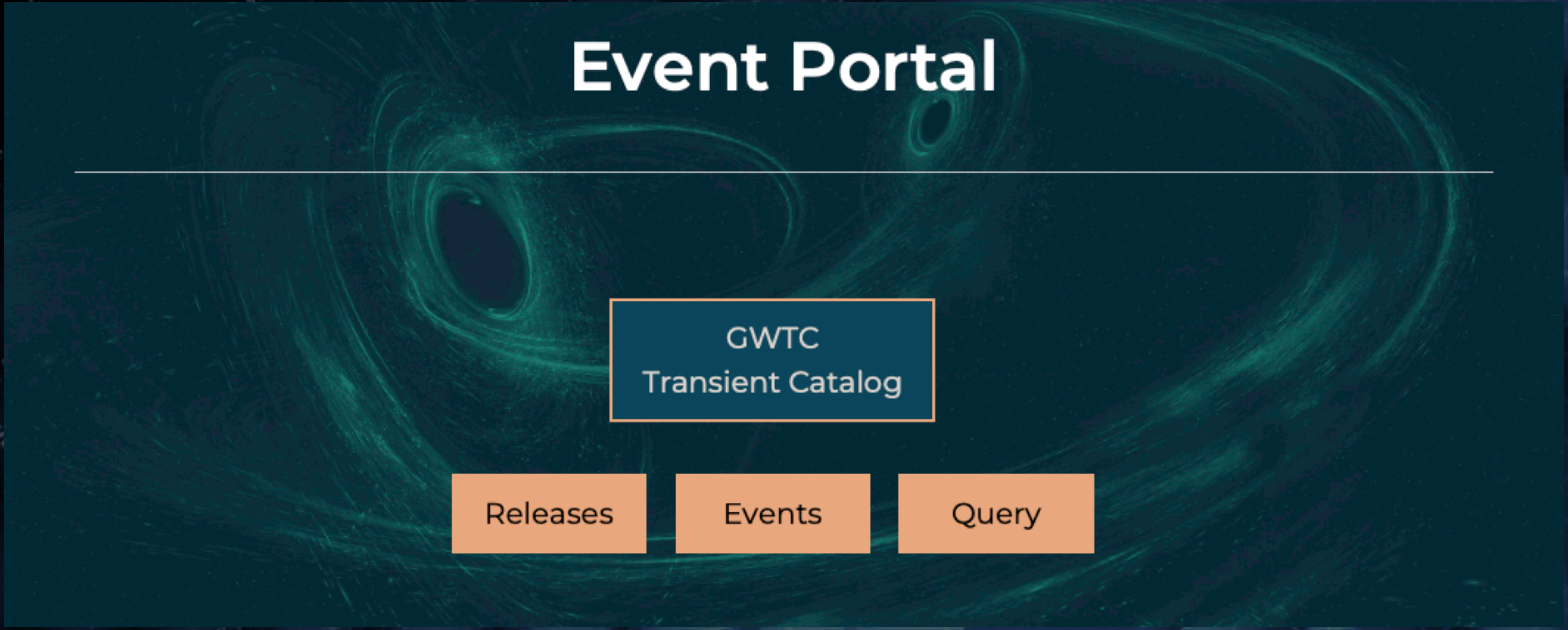
| Timeline | UTC | Mbytes | HDF5 | Frame | Percent |
|---|---|---|---|---|---|
| 1256652800 | 2019-11-01T14:13:02 | 40 MB | HDF5 | Frame | 30 |
| 1256656896 | 2019-11-01T15:21:18 | 27 MB | HDF5 | Frame | 20 |
| 1256660992 | 2019-11-01T16:29:34 | 36 MB | HDF5 | Frame | 27 |

# Content of a strain data file




- Metadata of the file
- Informations about the data quality and injections (sampled at 1 Hz)
- Strain data (sampled at 16 kHz or 4 kHz)

| Bit | Short Name | Description |
|---|---|---|
| Data Quality Bits | | |
| 0 | DATA | data present |
| 1 | CBC_CAT1 | passes the cbc CAT1 test |
| 2 | CBC_CAT2 | passes cbc CAT2 test |
| 3 | CBC_CAT3 | passes cbc CAT3 test |
| 4 | BURST_CAT1 | passes burst CAT1 test |
| 5 | BURST_CAT2 | passes burst CAT2 test |
| 6 | BURST_CAT3 | passes burst CAT3 test |
| Injection Bits | | |
| 0 | NO_CBC_HW_INJ | no cbc injection |
| 1 | NO_BURST_HW_INJ | no burst injections |
| 2 | NO_DETCHAR_HW_INJ | no detchar injections |
| 3 | NO_CW_HW_INJ | no continuous wave injections |
| 4 | NO_STOCH_HW_INJ | no stoch injections |


Event Portal
GWTC
Transient Catalog
Releases
Events
Query

# Event Portal

GWTC
Transient Catalog

Releases Events Query

## Release List

| Release Name | Description |
|---|---|
| GWTC | The Gravitational-wave Transient Catalog (GWTC) is a cumulative set of gravitational wave transients maintained by the LIGO/Virgo/KAGRA collaboration. The online GWTC contains confidently-detected events from multiple data releases. |
| GWTC-1-confident | Confident detections from "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs." Additional data products, including PE samples and skymaps, are linked from the documentation at https://doi.org/10.7935/82H3-HH23 |
| GWTC-1-marginal | Marginal candidates from "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs." Additional data products are linked from the documentation at https://doi.org/10.7935/82H3-HH23 |
| GWTC-2 | Events from the O3A observation run of LIGO and Virgo, as described in the GWTC-2 catalog paper. These events are also included in a cumulative list of all GWTC events published to date. Details and additional data products are linked from the documentation page. |
| GWTC-2.1-auxiliary | This release/list/table contains candidates from GWTC-2 which, based on the [updated] analysis presented in the GWTC-2.1 catalog paper |

A

# Event Portal

GWTC
Transient Catalog

Releases | Events | Query

| Name | Version | Release | GPS | Mass 1 ($M_\odot$) | Mass 2 ($M_\odot$) | Network SNR | Distance (Mpc) | $\chi_{eff}$ | Total Mass ($M_\odot$) |
|---|---|---|---|---|---|---|---|---|---|
| GW200322_091133 | v1 | GWTC-3-confident | 1268903511.3 | $34^{+48}_{-18}$ | $14.0^{+16.8}_{-8.7}$ | $6.0^{+1.7}_{-1.2}$ | $3600^{+7000}_{-2000}$ | $0.24^{+0.45}_{-0.51}$ | $55^{+37}_{-27}$ |
| GW200316_215756 | v1 | GWTC-3-confident | 1268431094.1 | $13.1^{+10.2}_{-2.9}$ | $7.8^{+1.9}_{-2.9}$ | $10.3^{+0.4}_{-0.7}$ | $1120^{+470}_{-440}$ | $0.13^{+0.27}_{-0.10}$ | $21.2^{+7.2}_{-2.0}$ |
| GW200311_115853 | v1 | GWTC-3-confident | 1267963151.3 | $34.2^{+6.4}_{-3.8}$ | $27.7^{+4.1}_{-5.9}$ | $17.8^{+0.2}_{-0.2}$ | $1170^{+280}_{-400}$ | $-0.02^{+0.16}_{-0.20}$ | $61.9^{+5.3}_{-4.2}$ |
| GW200311_103121 | v1 | GWTC-3-marginal | 1267957899.7 | -- | -- | 9.2 | -- | -- | -- |
| GW200308_173609 | v1 | GWTC-3-confident | 1267724187.7 | $36.4^{+11.2}_{-9.6}$ | $13.8^{+7.2}_{-3.3}$ | $7.1^{+0.5}_{-0.5}$ | $5400^{+2700}_{-2600}$ | $0.65^{+0.17}_{-0.21}$ | $50.6^{+10.9}_{-8.5}$ |
| GW200306_093714 | v1 | GWTC-3-confident | 1267522652.1 | $28.3^{+17.1}_{-7.7}$ | $14.8^{+6.5}_{-6.4}$ | $7.8^{+0.4}_{-0.6}$ | $2100^{+1700}_{-1100}$ | $0.32^{+0.28}_{-0.46}$ | $43.9^{+11.8}_{-7.5}$ |
| GW200302_015811 | v1 | GWTC-3-confident | 1267149509.5 | $37.8^{+8.7}_{-8.5}$ | $20.0^{+8.1}_{-5.7}$ | $10.8^{+0.3}_{-0.4}$ | $1480^{+1020}_{-700}$ | $0.01^{+0.25}_{-0.26}$ | $57.8^{+9.6}_{-6.9}$ |
| GW200225_060421 | v1 | GWTC-3-confident | 1266645879.3 | $19.3^{+5.0}_{-3.0}$ | $14.0^{+2.8}_{-3.5}$ | $12.5^{+0.3}_{-0.4}$ | $1150^{+510}_{-530}$ | $-0.12^{+0.17}_{-0.28}$ | $33.5^{+3.6}_{-3.0}$ |
| GW200224_222234 | v1 | GWTC-3-confident | 1266618172.4 | $40.0^{+6.9}_{-4.5}$ | $32.5^{+5.0}_{-7.2}$ | $20.0^{+0.2}_{-0.2}$ | $1710^{+490}_{-640}$ | $0.10^{+0.15}_{-0.15}$ | $72.2^{+7.2}_{-5.1}$ |
| GW200220_124850 | v1 | GWTC-3-confident | 1266238148.1 | $38.9^{+14.1}_{-8.6}$ | $27.9^{+9.2}_{-9.0}$ | $8.5^{+0.3}_{-0.5}$ | $4000^{+2800}_{-2200}$ | $-0.07^{+0.27}_{-0.33}$ | $67^{+17}_{-12}$ |
| GW200220_061928 | v1 | GWTC-3-confident | 1266214786.7 | $87^{+40}_{-23}$ | $61^{+26}_{-25}$ | $7.2^{+0.4}_{-0.7}$ | $6000^{+4800}_{-3100}$ | $0.06^{+0.40}_{-0.38}$ | $148^{+55}_{-33}$ |

# Event Portal

GWTC
Transient Catalog

Releases | Events | Query

Query Events

| | | | | | |
|---|---|---|---|---|---|
| Event Name: | | | | | |
| Release: | GWTC-1-marginal<br>GWTC-1-confident<br>O1_O2-Preliminary<br>O3_Discovery_Papers | | | | |
| Mass 1 Range: | 0 | ∞ | Mass 2 Range: | 0 | ∞ |
| Total Mass Range: | 0 | ∞ | Final Mass Range: | 0 | ∞ |
| Chirp Mass Range: | 0 | ∞ | Detector Frame Chirp Mass Range: | 0 | ∞ |
| Distance (Mpc) Range: | 0 | ∞ | Redshift Range: | 0 | ∞ |
| Network SNR Range: | 0 | ∞ | $\chi_{eff}$ Range: | -1 | 1 |
| False Alarm Rate Range: | 0 | ∞ | $P_{astro}$ Range: | 0 | 1 |
| UTC Time Range: | | | | | |

# Example of an event page

## GWTC-2.1 PE for GW190412 (update)

Waveform Family: C01:Mixed

Date added: May 13, 2022

show / hide parameters

| | |
|---|---|
| chi_eff | $0.21^{+0.12}_{-0.13}$ |
| chirp_mass_source (M_sun) | $13.3^{+0.5}_{-0.5}$ |
| final_mass_source (M_sun) | $35.6^{+4.8}_{-4.5}$ |
| final_spin | $0.66^{+0.05}_{-0.04}$ |
| luminosity_distance (Mpc) | $720^{+240}_{-220}$ |
| mass_1_source (M_sun) | $27.7^{+6.0}_{-6.0}$ |
| mass_2_source (M_sun) | $9.0^{+2.0}_{-1.4}$ |
| network_matched_filter_snr | $19.8^{+0.2}_{-0.3}$ |
| redshift | $0.15^{+0.04}_{-0.04}$ |
| sky_area (deg^2) | 240 |
| total_mass_source (M_sun) | $36.8^{+4.7}_{-4.4}$ |

Source File

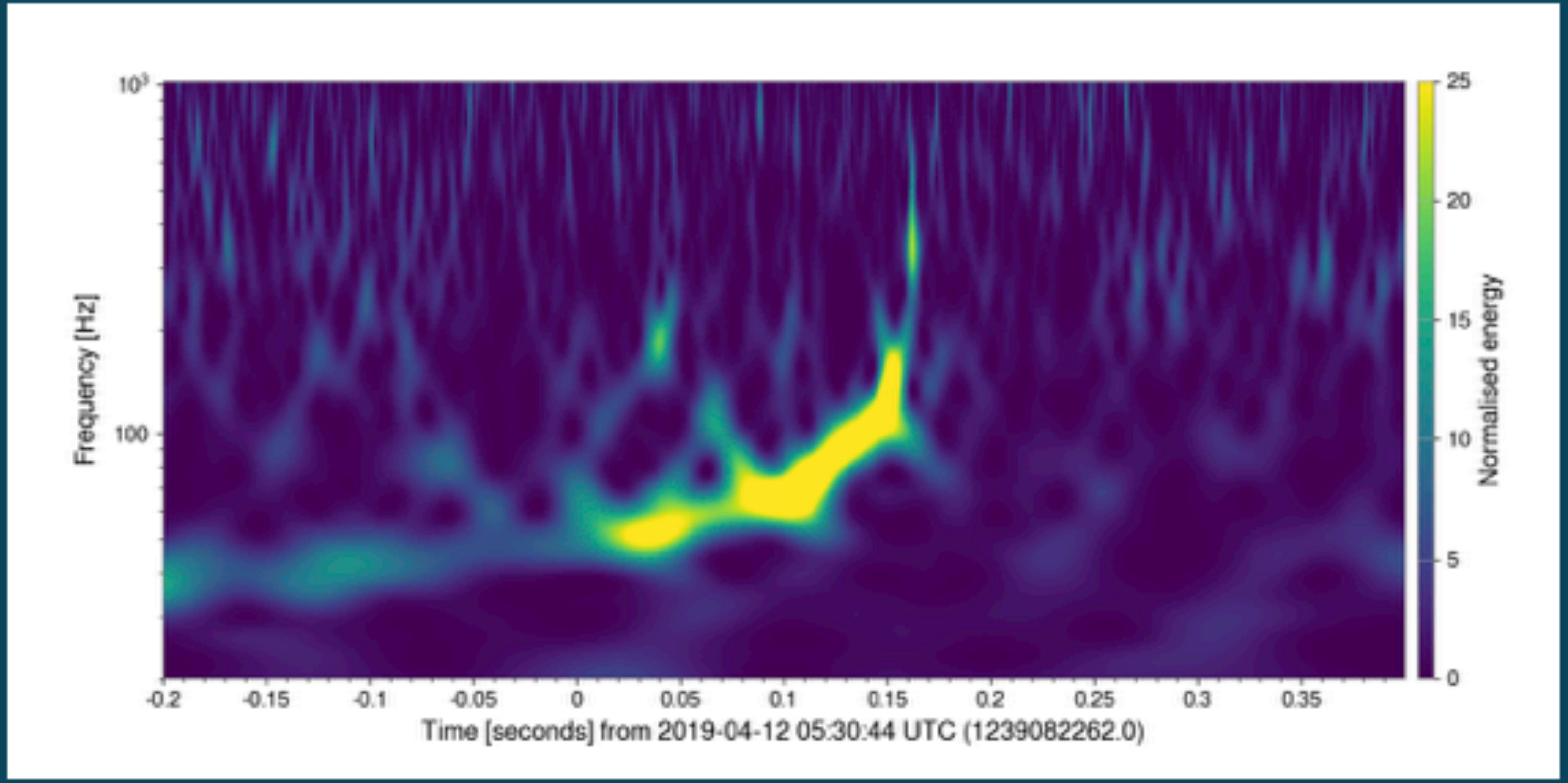
Posterior Samples Zenodo Entry

Skymap for GW190412

Default PE

Link to this page

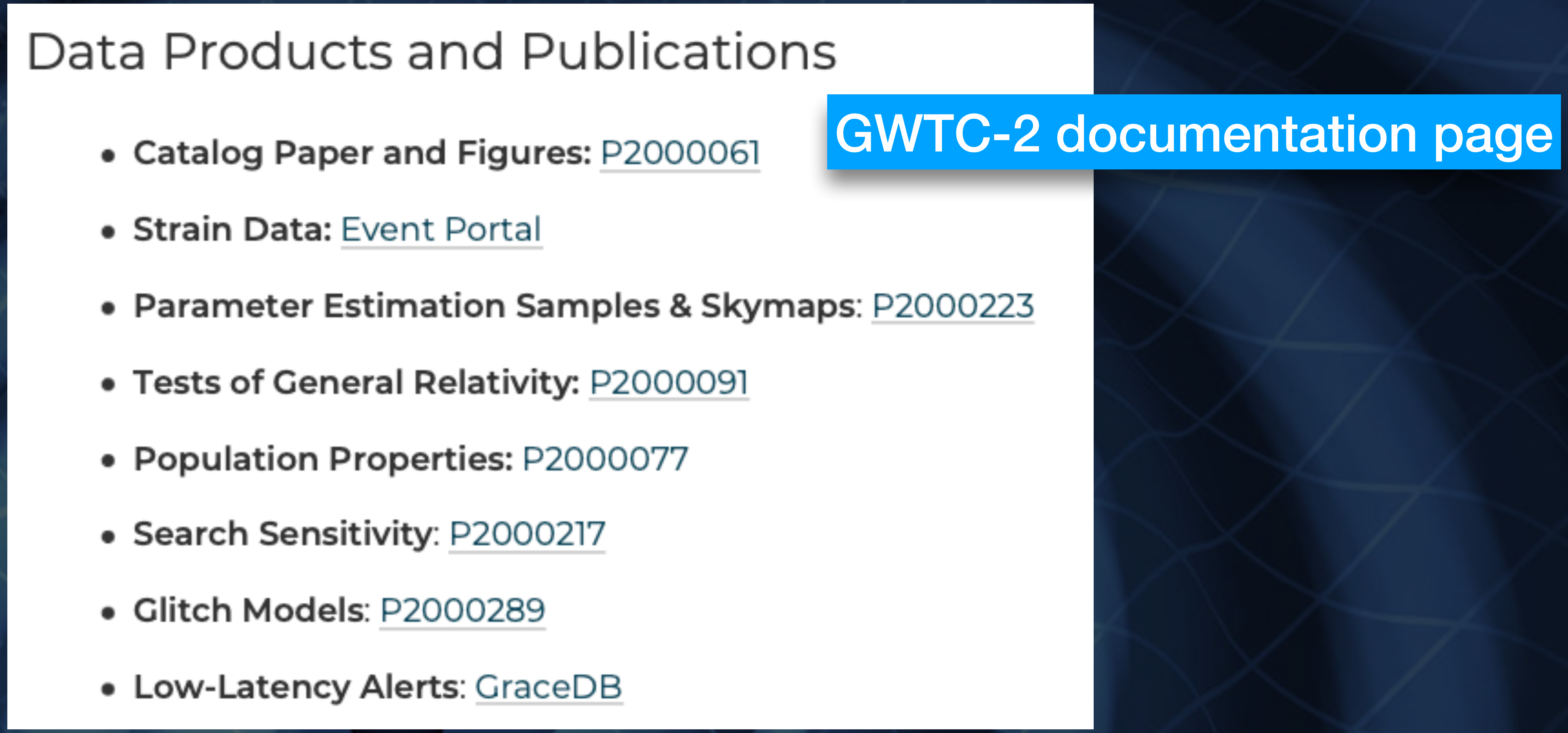
## L1 strain

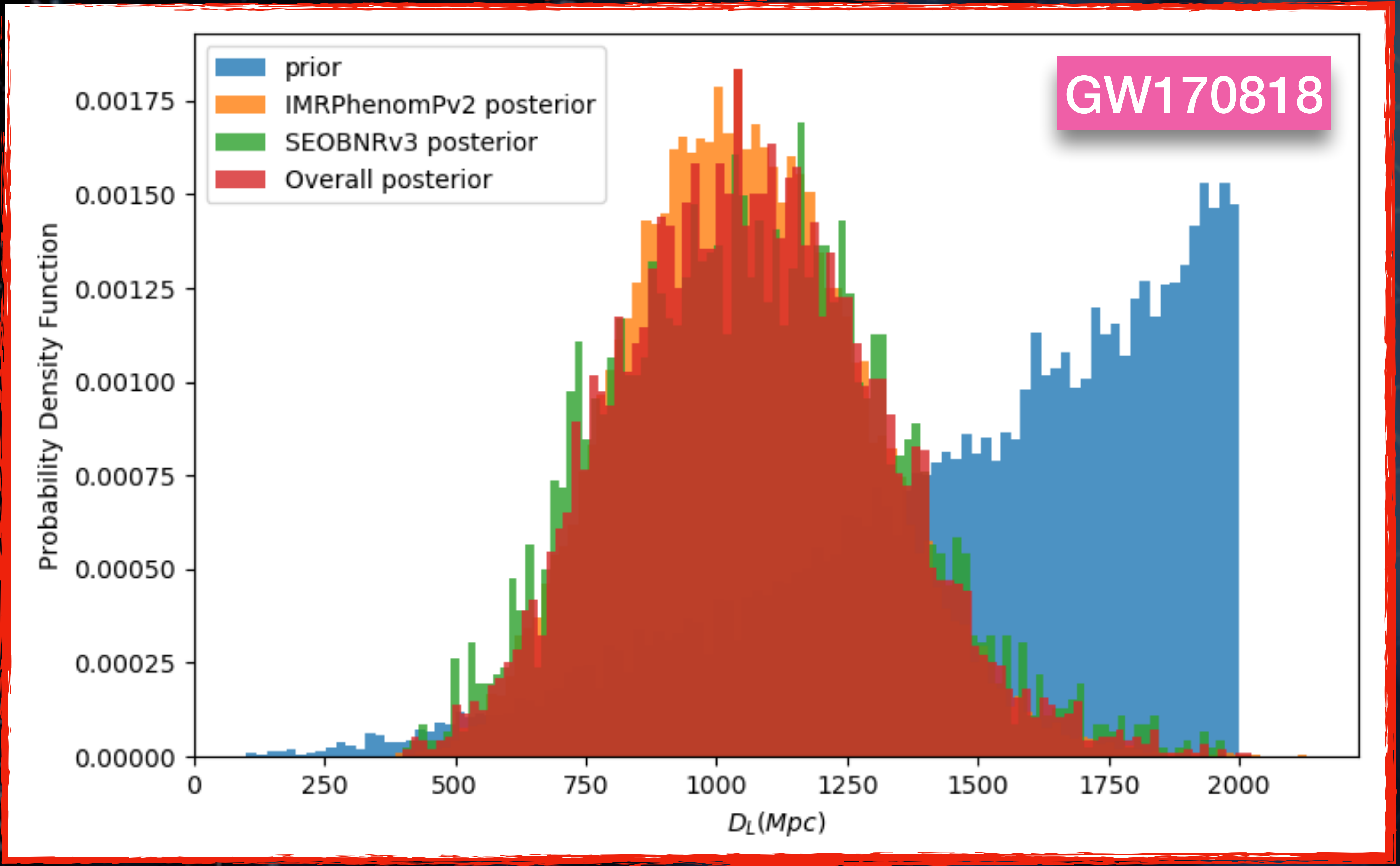


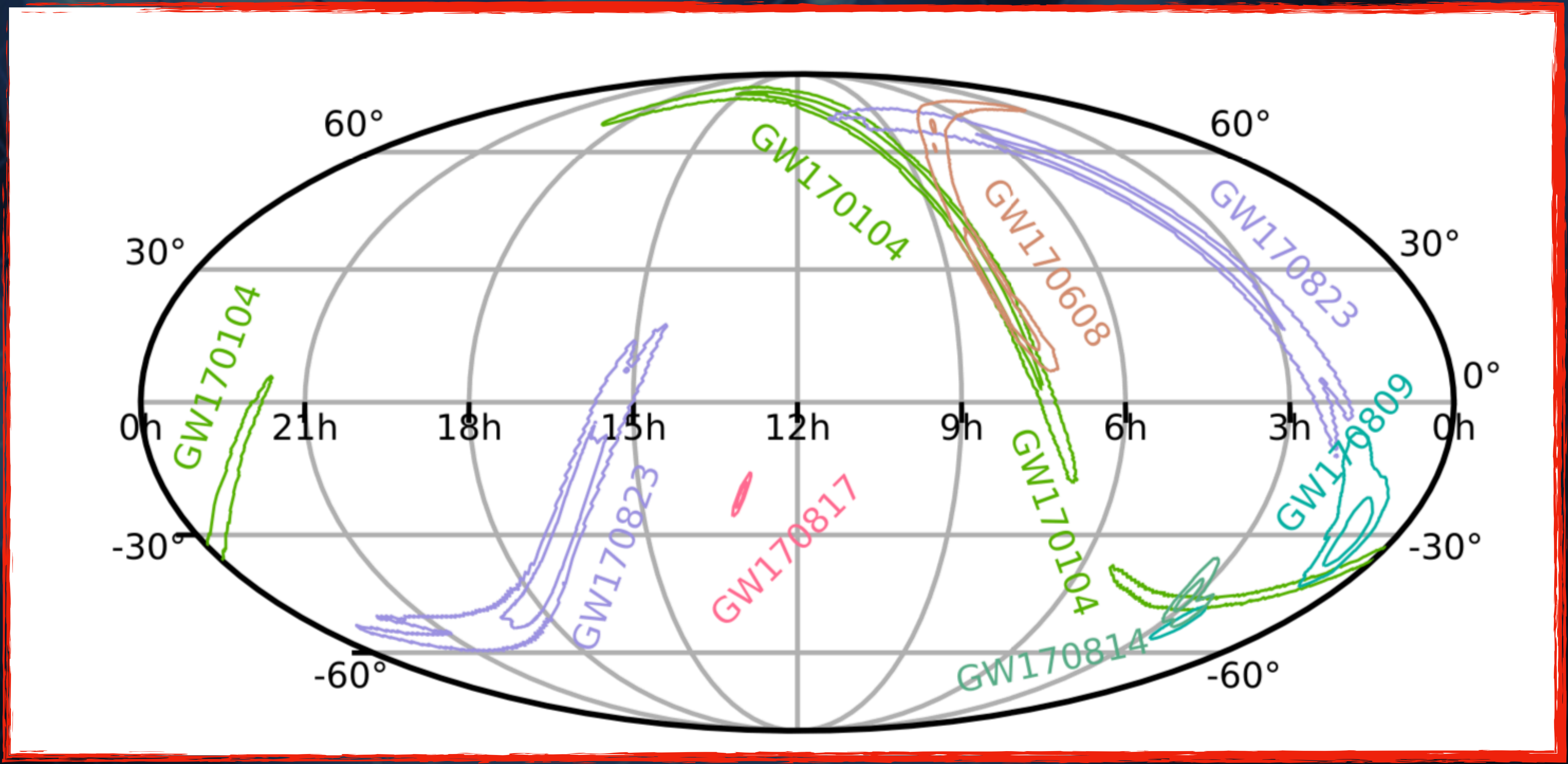

| | | | |
|---|---|---|---|
| 32sec • 16KHz: | GWF | HDF | TXT |
| 32sec • 4KHz: | GWF | HDF | TXT |
| 4096sec • 16KHz: | GWF | HDF | TXT |
| 4096sec • 4KHz: | GWF | HDF | TXT |

# Other analysis products

- The event portal contains links to:
  - ✓ Posterior samples
  - ✓ Confidence intervals
  - ✓ Skymaps

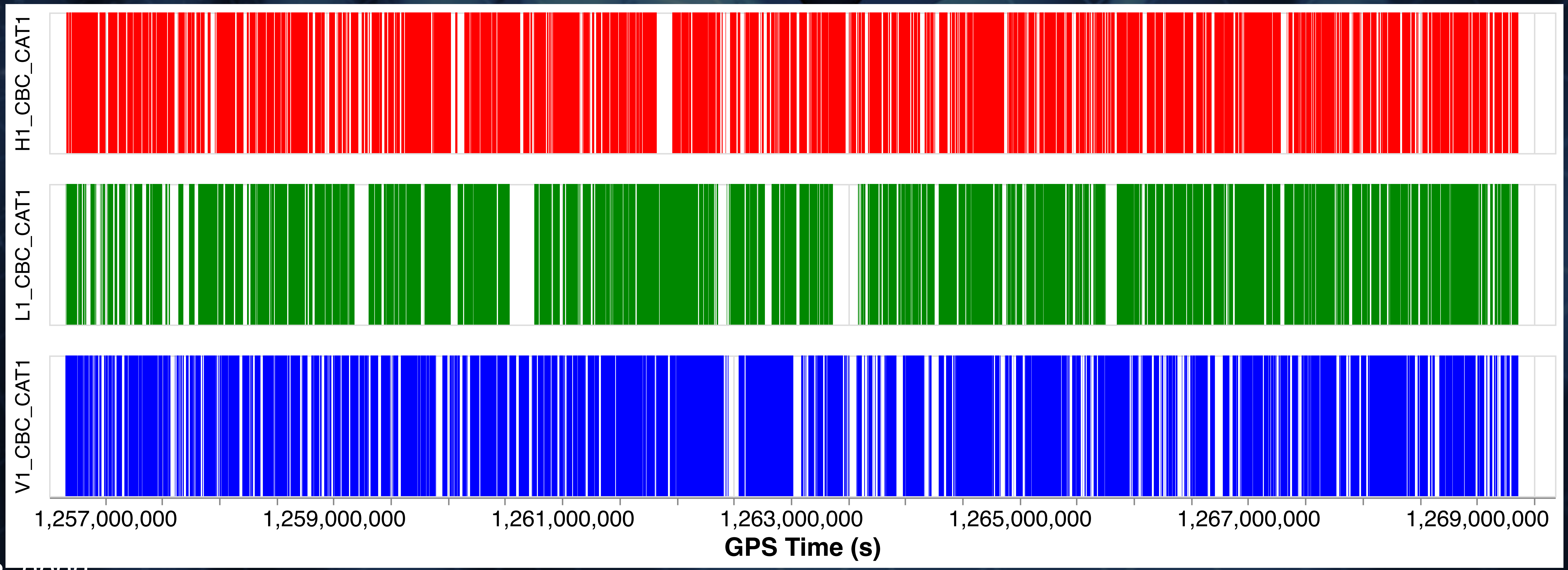


# Segment Lists

- LIGO/Virgo detectors are not always 'on' (in science mode)
- Data quality may not meet basic requirements
- Consequence : GW data analysis is applied to data segments (different in each detector)
- The “Timeline” tool of the GWOSC website allows you to select segments that fulfill specific data quality checks in a time period that can be selected specifying GPS time and duration




Timelines for O3b

# Software for GW data

- Software for working with Gravitational Wave Data available to the public: https://www.gwosc.org/software/
- API on the GWOSC web site + python client (called “gwosc”)
- Part of the software developed by LIGO/Virgo and open-source

GWpy

a python package for gravitational-wave astrophysics.

PyCBC

Free and open software to study gravitational waves.

Used in the tutorials just after this lecture

**ligo.skymap**

The ligo.skymap package provides tools for reading, writing, generating, and visualizing gravitational-wave probability maps from LIGO and Virgo. It includes the rapid sky localization code BAYESTAR, tools for making sky maps from MCMC samples, observation planning utilities, and tools for making beautiful astronomical maps.

**Bilby**

Bilby: a user-friendly Bayesian inference library.

**GstLAL**

gstlal provides a suite of GStreamer elements that expose gravitational-wave data analysis tools from the LALSuite library for use in GStreamer signal-processing pipelines.

**LALSuite**

The LSC Algorithm Library Suite (LALSuite) is a collection of component packages, each of which is tagged, packaged, and released separately.

# GW Open Data Workshops

https://www.gwosc.org/workshops

- Organised every year since 2018
- Last year we had about 1000 registered
- Lectures + hands-on sessions
- Workshops become online courses
- Register to the next one if you are interested!
- Hybrid format (in person study hubs)

## GW Open Data Workshops

Gravitational Wave Open Data Workshop #6 (2023)

May 15-17, 2023

Workshop

Gravitational Wave Open Data Workshop #5 (20

May 23-25, 2022

Workshop

Gravitational Wave Open Data Workshop #4 (20

Remote workshop, May 10-14, 2021

Online Course Workshop

# New program for students

https://www.gwosc.org/path

Web Apps with plotting tools without coding!

## Learning Path

**Teachers & Students**

Objectives

- Outreach oriented
- Material for undergrads or younger students

**Step 1: Watch Introductory Video (1)**

*7 Minutes*

- How are gravitational waves created?
- How are gravitational waves measured?


Footnote: Construction of the KAGRA detector in Japan has been completed since the production of this video.

**Select Data Time and Detector**

How do you want to find data?

By event name

Select Event

GW150914

Detector

H1

**Set Plot Parameters**

Time Range (seconds)

1.00

0.10 8.00

Whitened and band-passed data

Whiten?

Band-pass frequency range (Hz)

30 400

10 2000

Q-tranform plot

Colorbar Max Energy

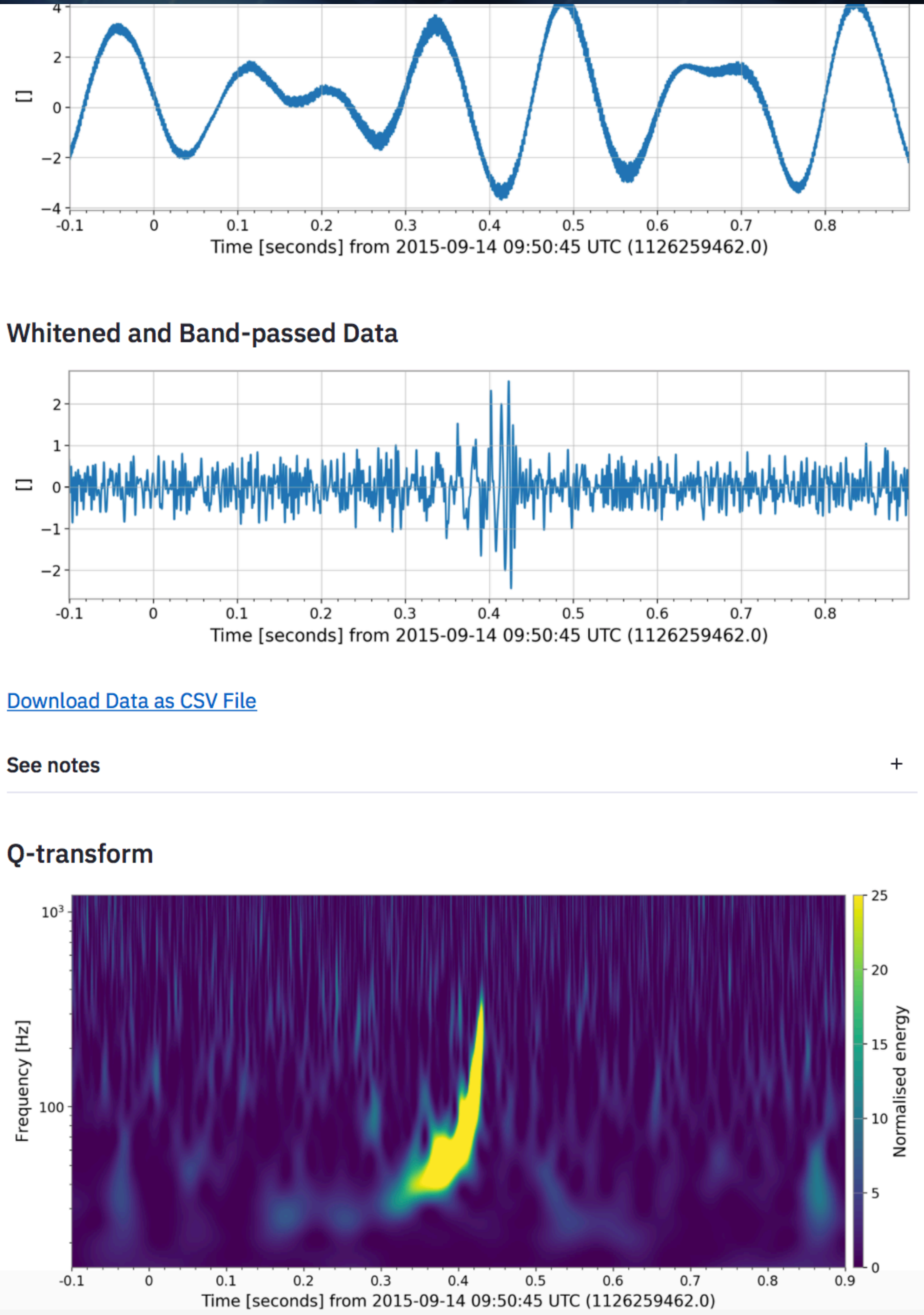
25

10 500

Q-value

5

5 120

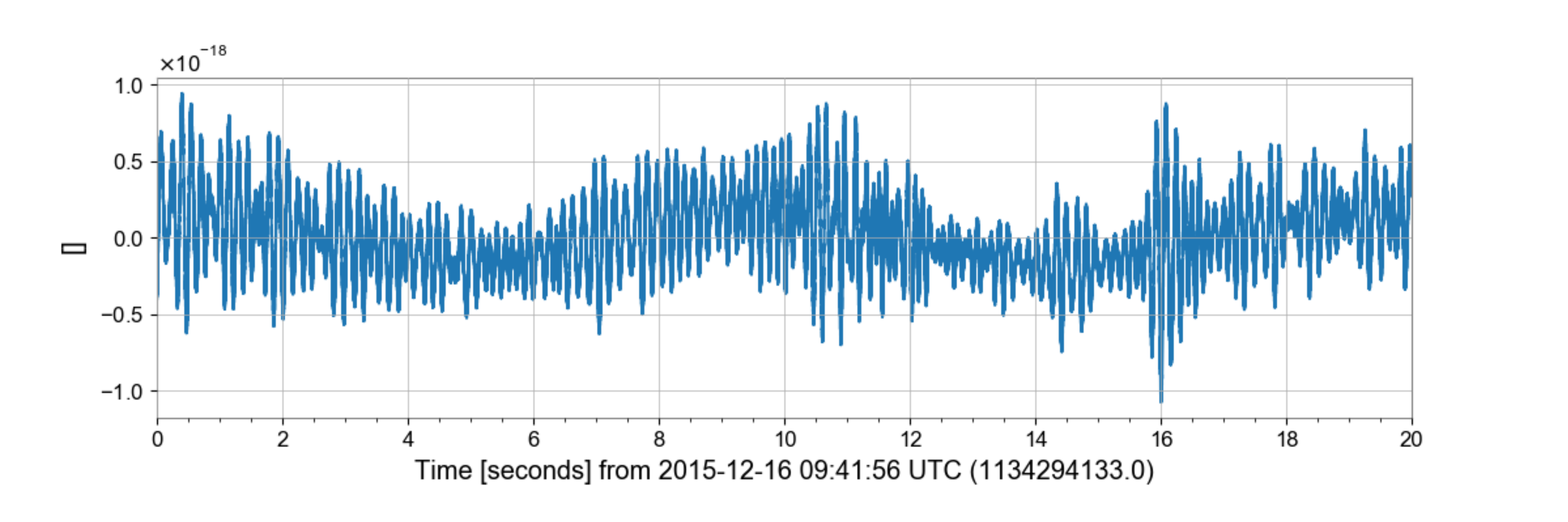


# Conclusion

- Gravitational-wave data are accessible through the GWOSC website!
- GWOSC is a successful open science project
  - ✓ A large and growing community of users
  - ✓ Data used in a good fraction of the publications related to GW astronomy
- Reproducibility of LVK results (see e. g. arXiv:2010.07244)
- Next observing run will start in about 2 month!!!
- A wealth of science to come and to share!
- A post-doc position opening in Trieste at the end of this year (contact agata.trovato@units.it if you are interested)
- For further questions write to gwosc@igwn.org or use ask.igwn.org

# Part II: Hands-on session

# Intro to gwpy




| | |
|---|---|
| Install it: | pip install gwpy |
| Import the class you need: | from gwpy.timeseries import TimeSeries |
| Download the data: | data = TimeSeries.fetch_open_data('H1', 1134294133,1134294153) |
| Plot the data: | plot = data.plot() |
| Show the plot: | plot.show() |

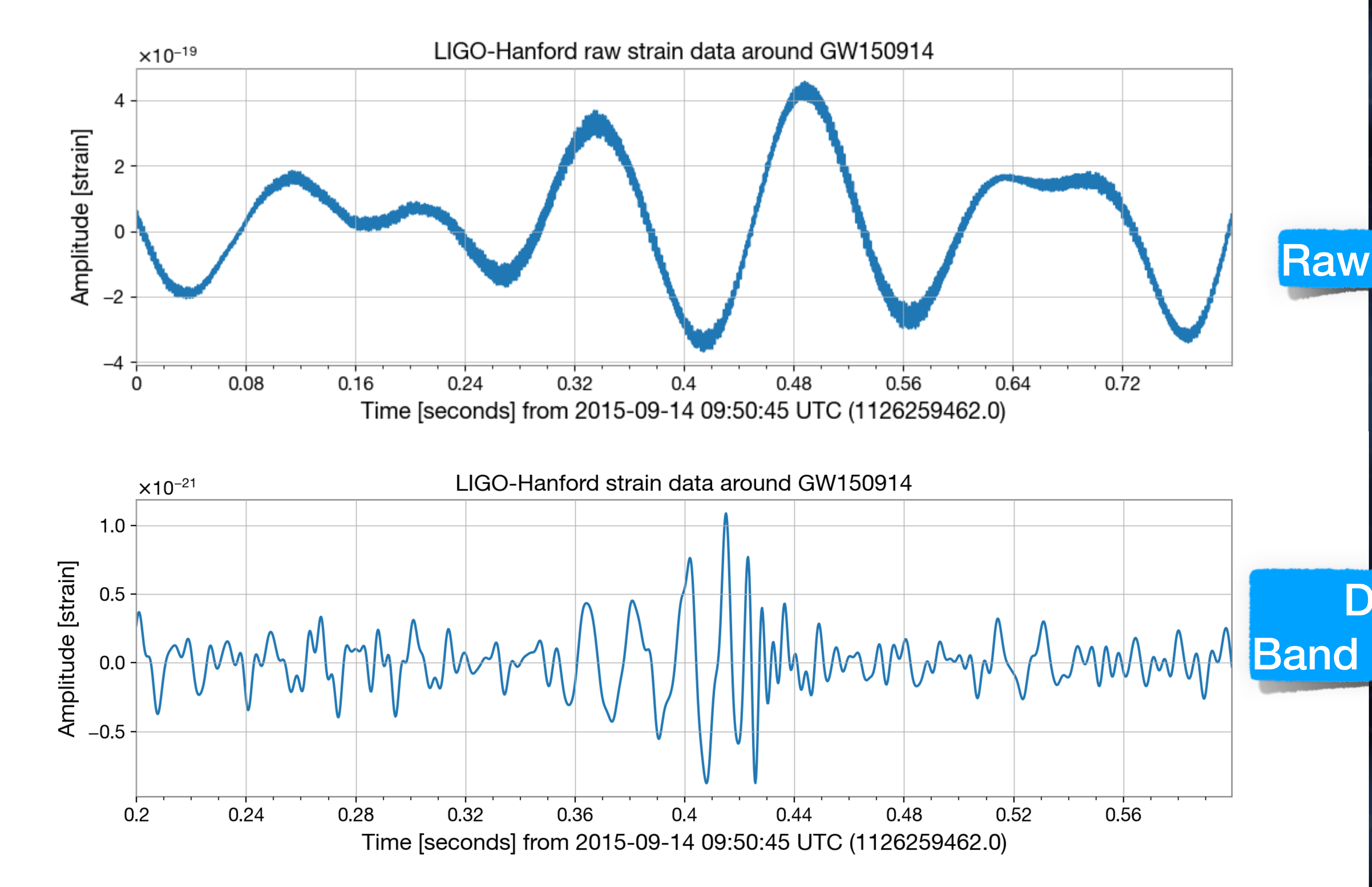
# Example to plot GW150914 data

```
from gwpy.timeseries import TimeSeries
from gwpy.signal import filter_design

t0 = 1126259462.4
hdata = TimeSeries.fetch_open_data('H1', t0-0.4,t0+0.4)

bp = filter_design.bandpass(50, 250, hdata.sample_rate)
notches = [filter_design.notch(line, hdata.sample_rate) for line in (60, 120, 180)]
zpk = filter_design.concatenate_zpks(bp, *notches)
hfilt = hdata.filter(zpk, filtfilt=True)
hfilt = hfilt.crop(t0-0.2,t0+0.2)
plot = hfilt.plot()
ax = plot.gca()
ax.set_title('LIGO-Hanford strain data around GW150914')
ax.set_ylabel('Amplitude [strain]')
plot.show()
```

# Example to plot GW150914 data

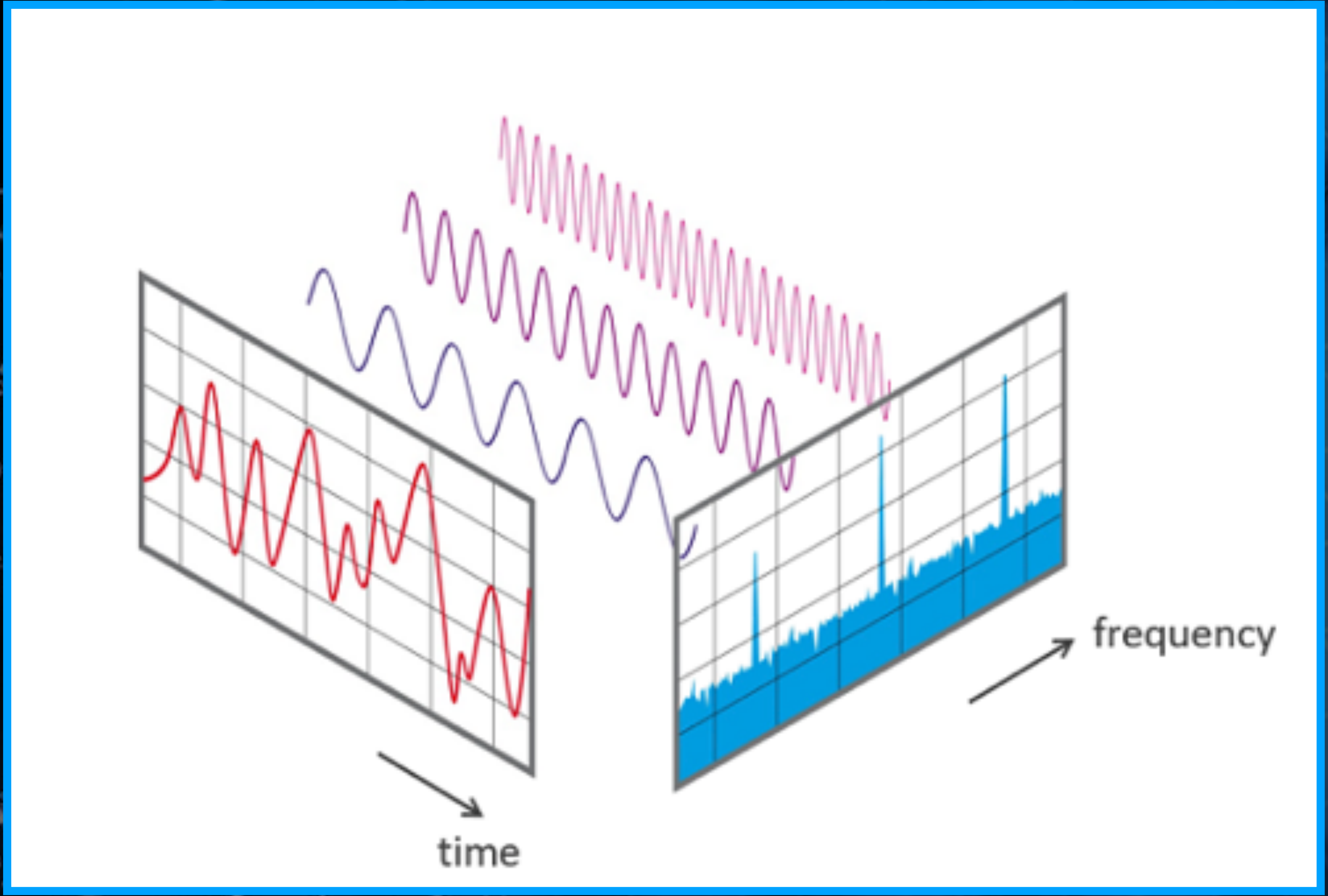



Raw data

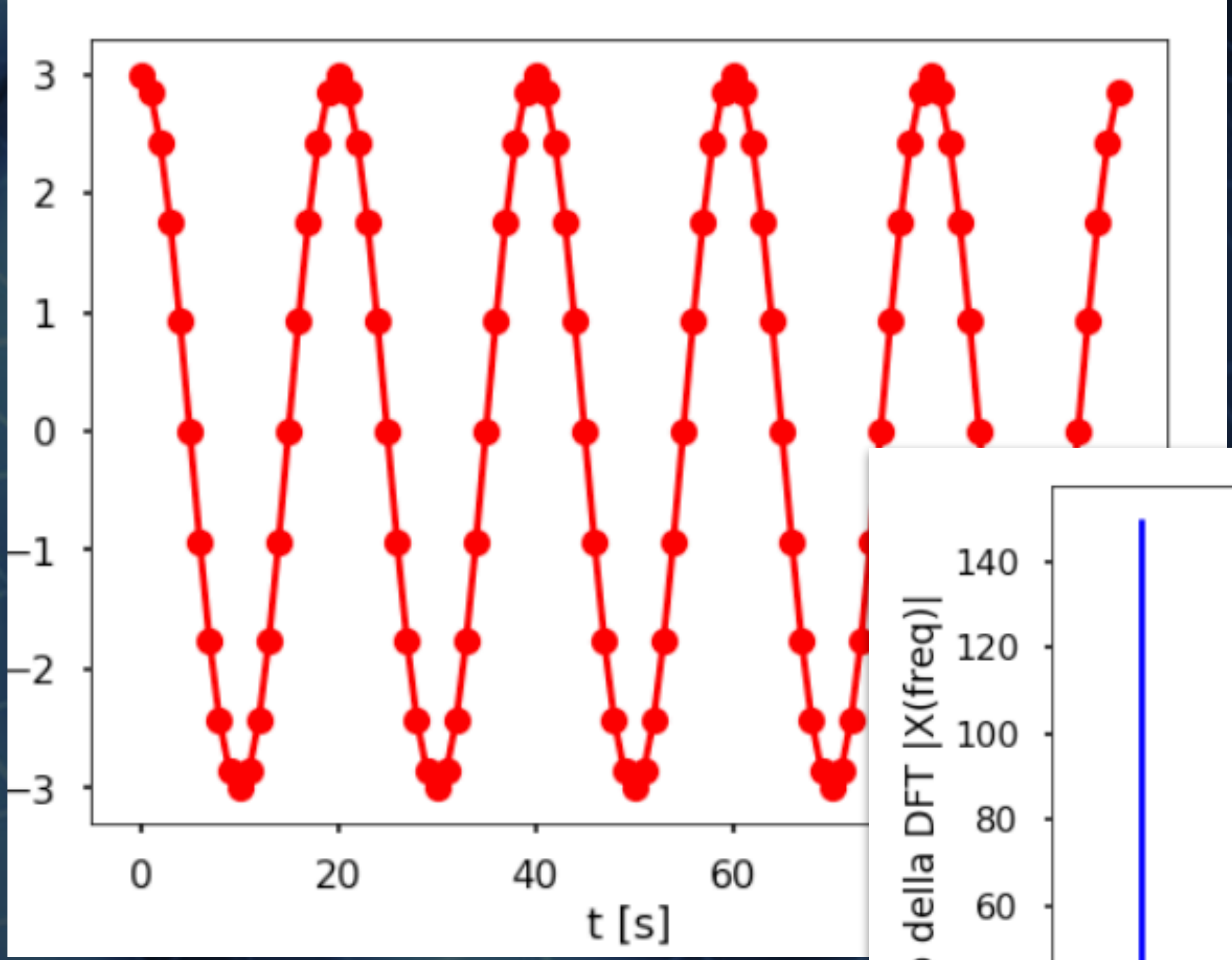
Data after
Band pass + notch

# Fourier Transforms and FFT

Fourier transforms decompose a *continuous* signal into its constituent frequencies

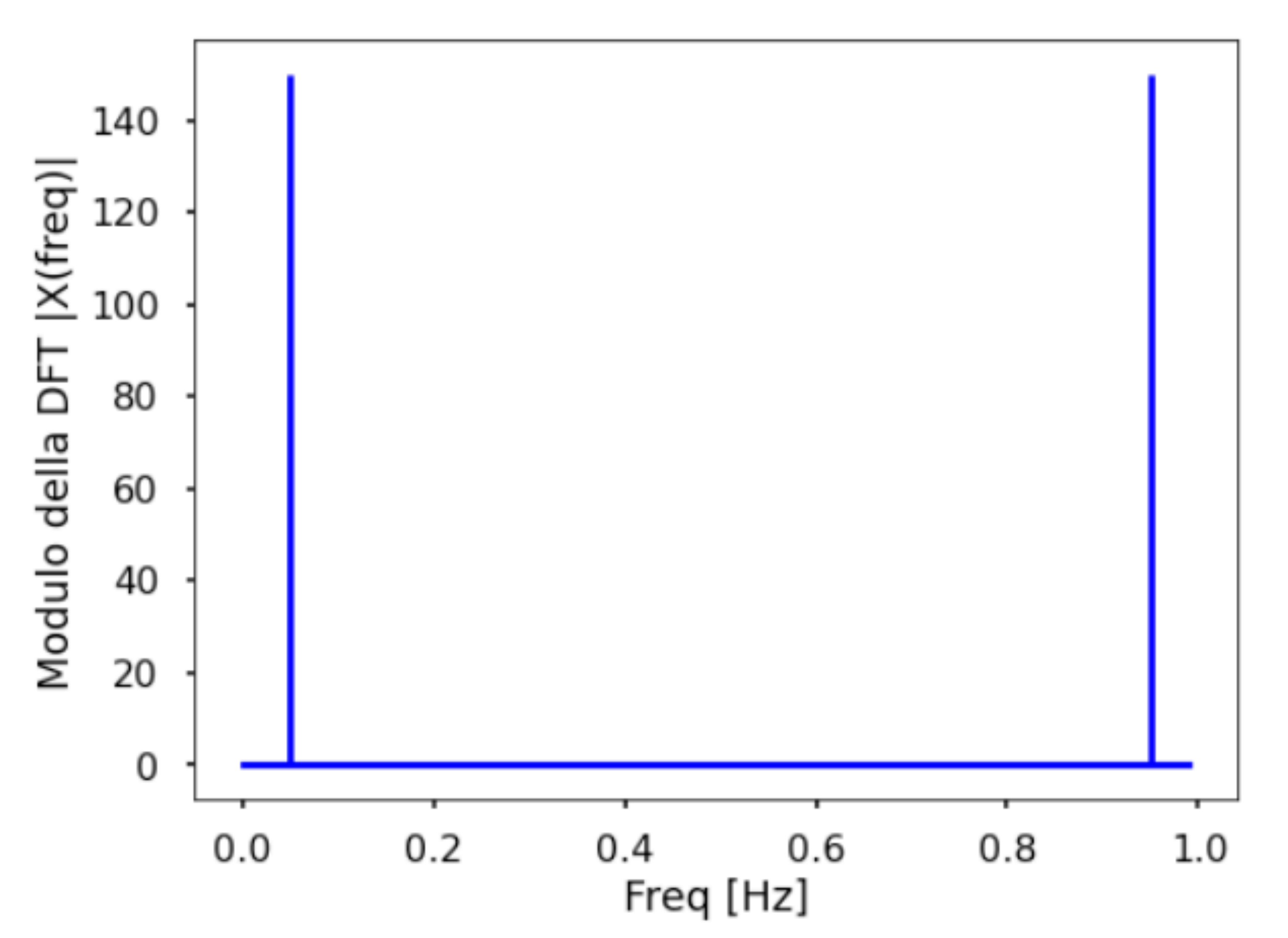



Fast Fourier Transformation (FFT) is an optimized mathematical algorithm to find the frequency components of a *discrete* signal



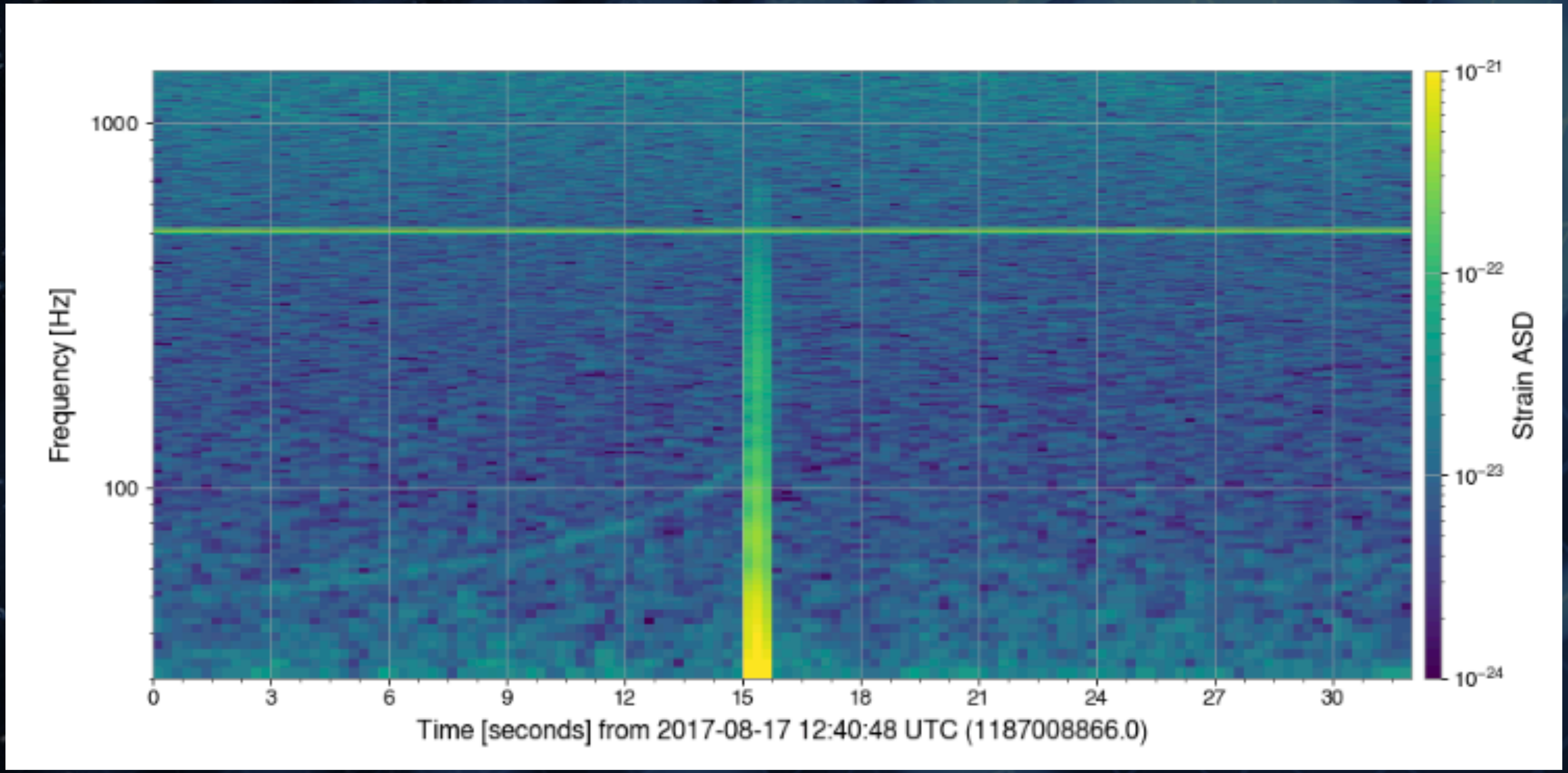

Time domain

Frequency domain



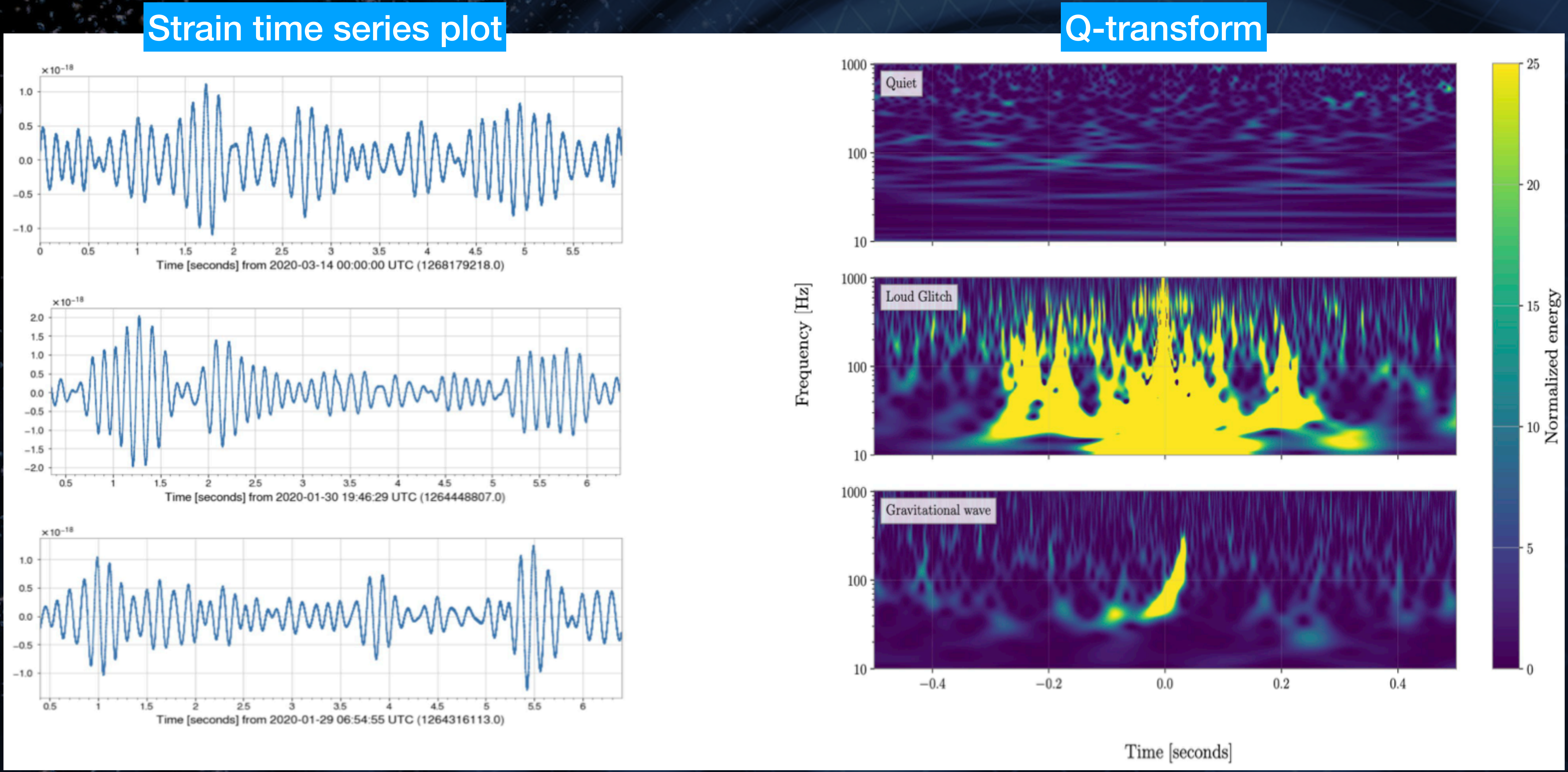
# Spectrograms

- Spectrograms: frequencies vs time
- The signal is broken into smaller time windows and the FFT is calculated for each window
- Time windows are usually partially overlapping




Plot taken from gwpy documentation

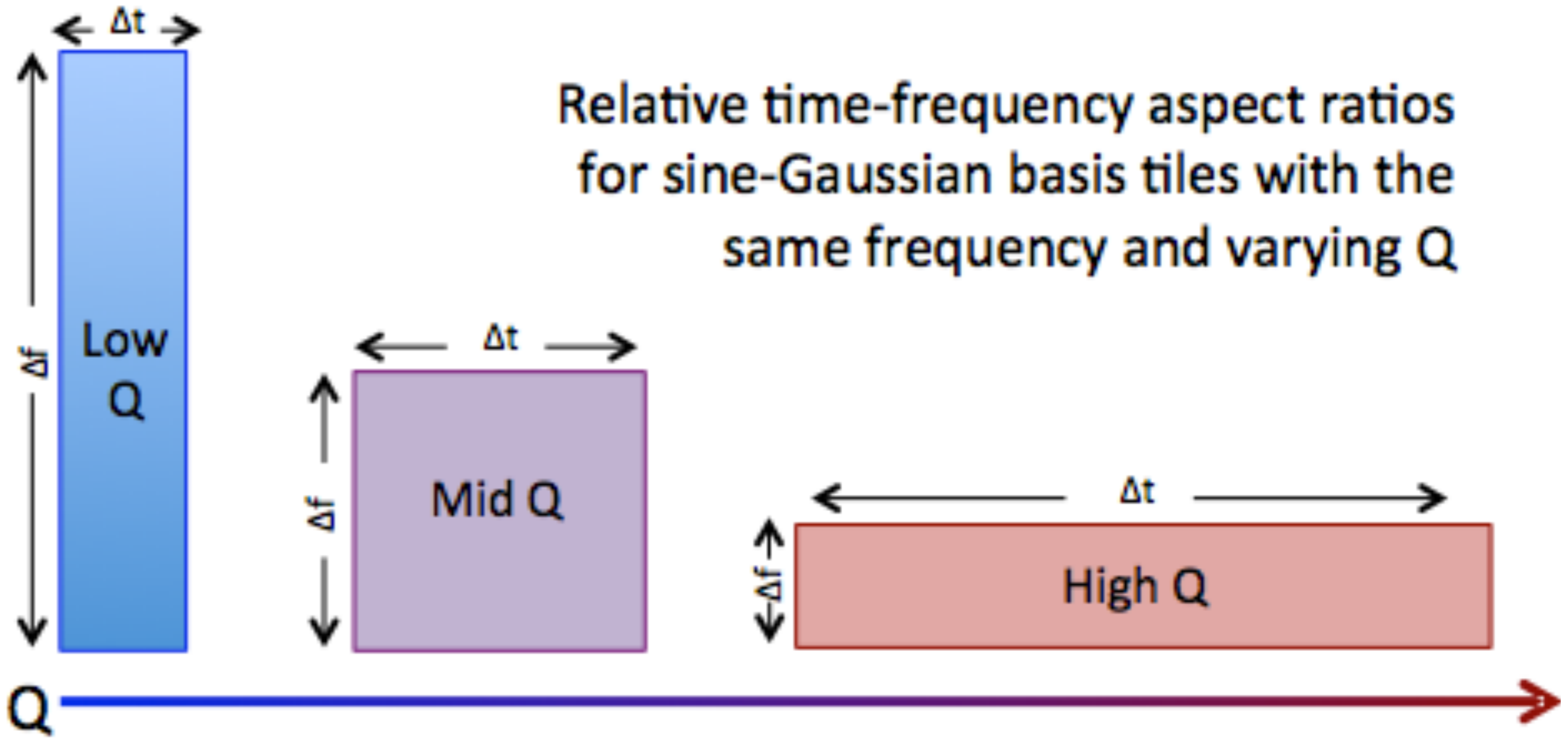
# Intro to Q-transform - I

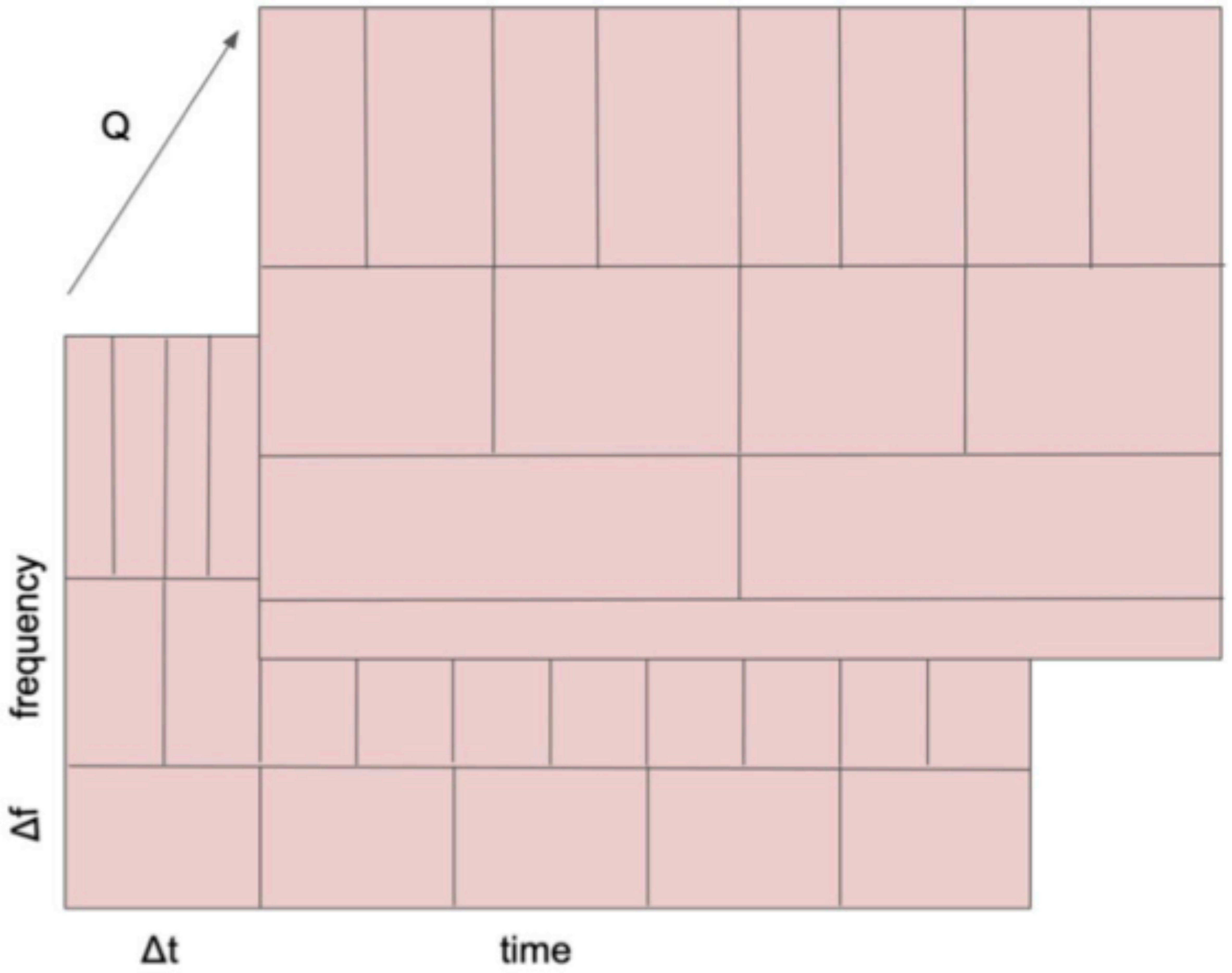


# Intro to Q-transform - II

- Q transform allows us to visualize the data in time-frequency space
- Multiple spectrograms with constant ratio of duration to bandwidth (Q factor)
- Optimize over time-freq tile energy to pick a plane

$$Q = \frac{f_0}{\Delta f}$$

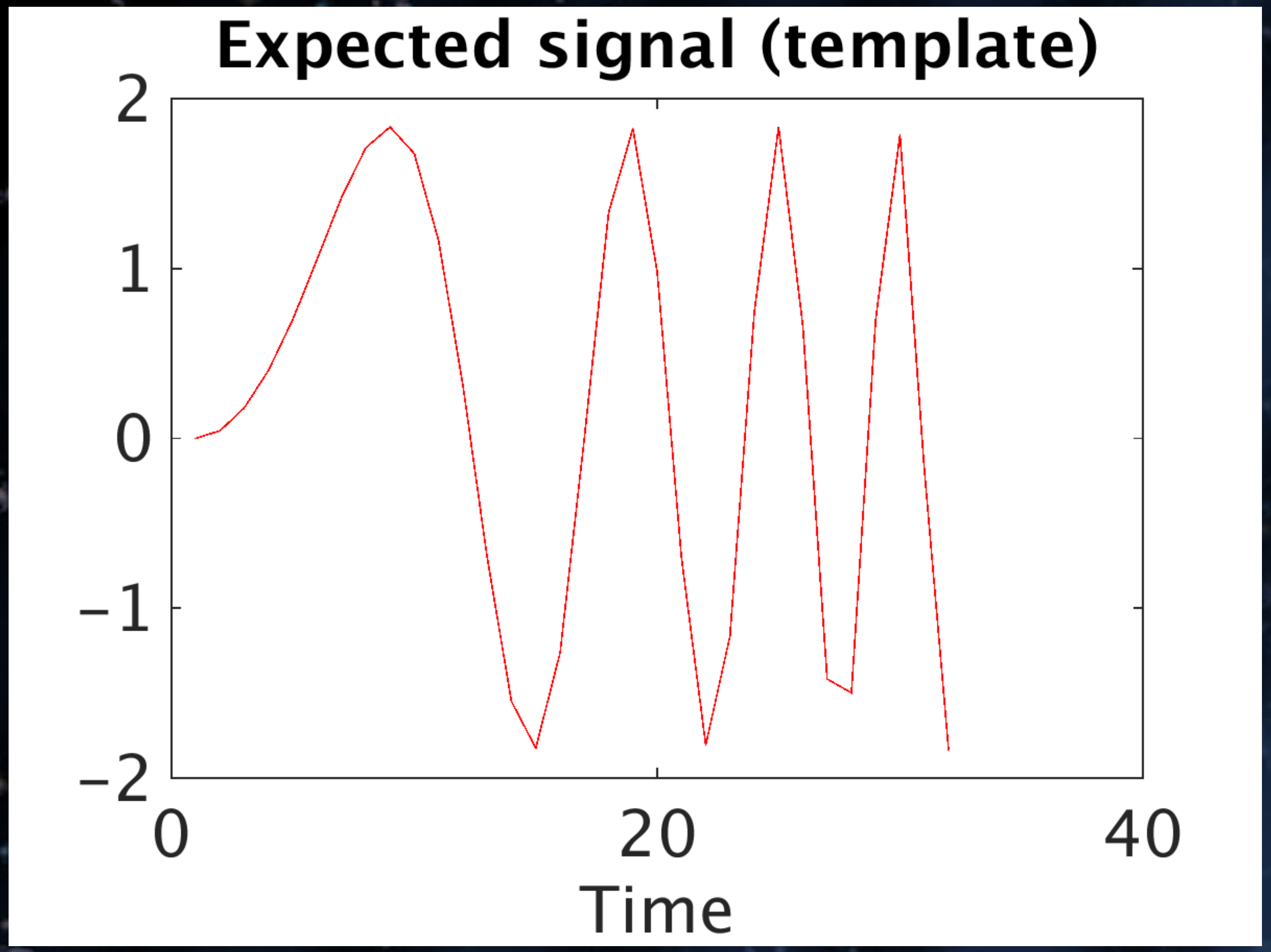


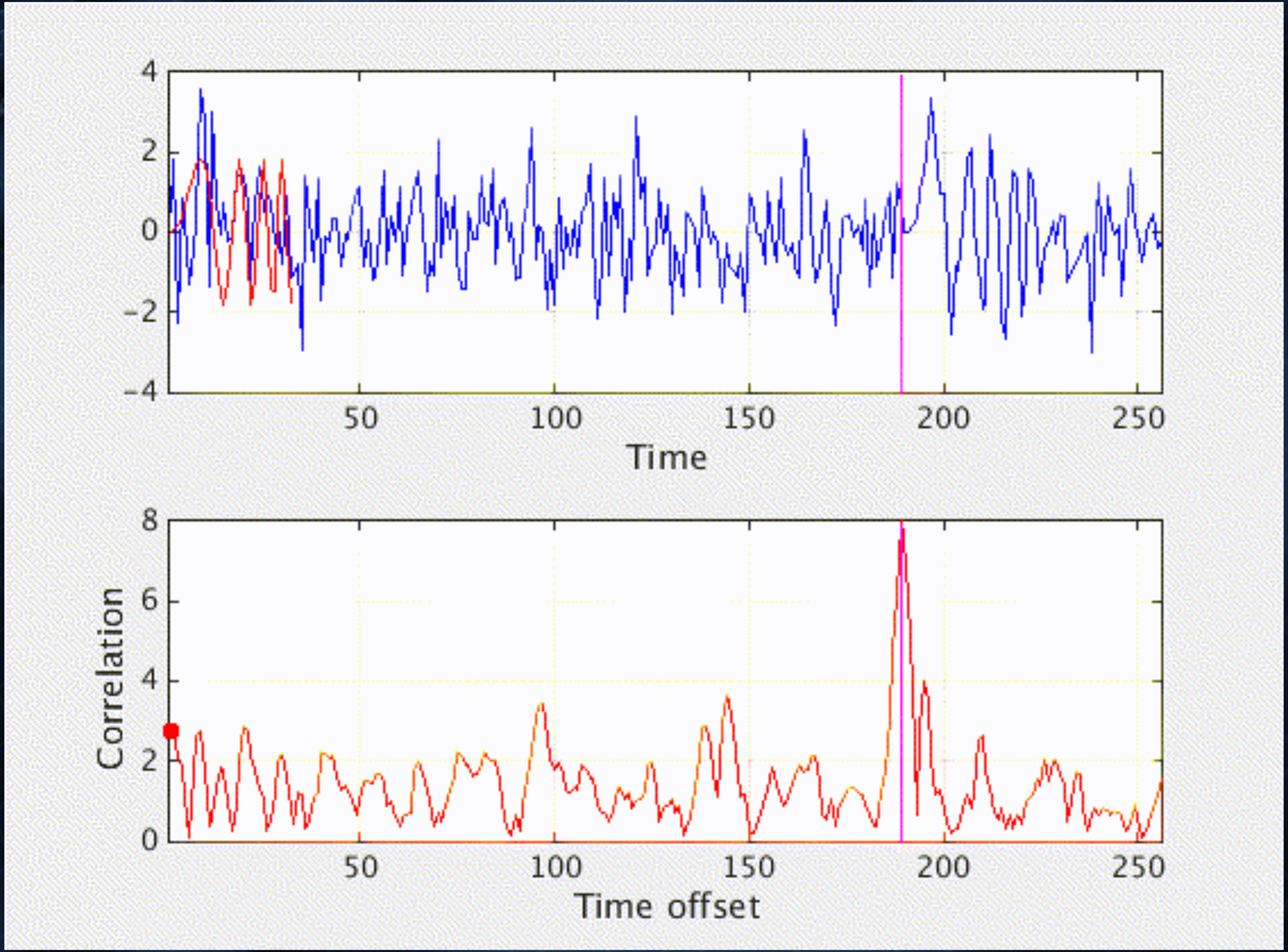

S. Chatterji et al., *Multiresolution techniques for the detection of gravitational-wave bursts*, CQG **21** S1809 (2004)

# Matched filtering (I)

Simple example of a template

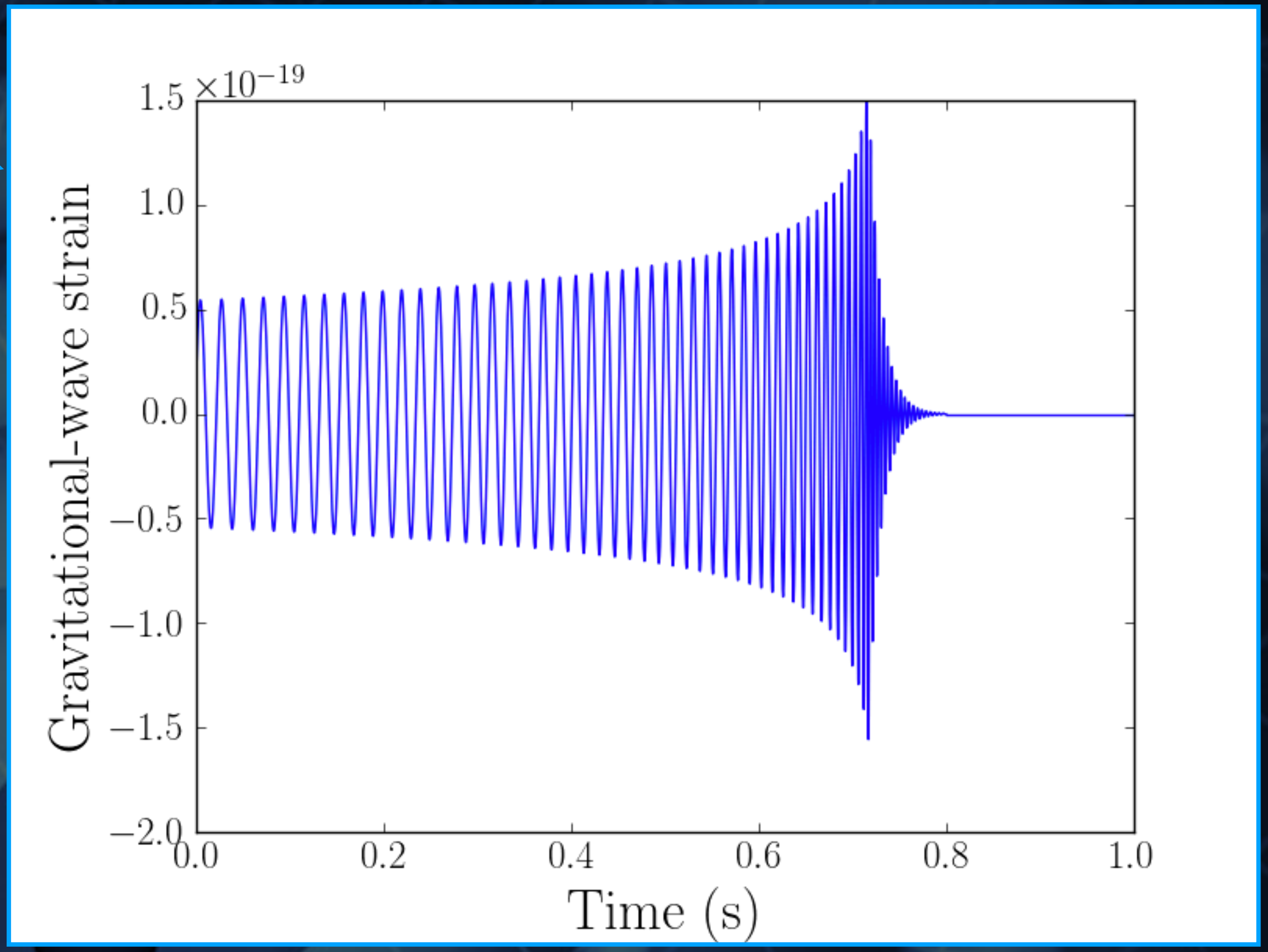


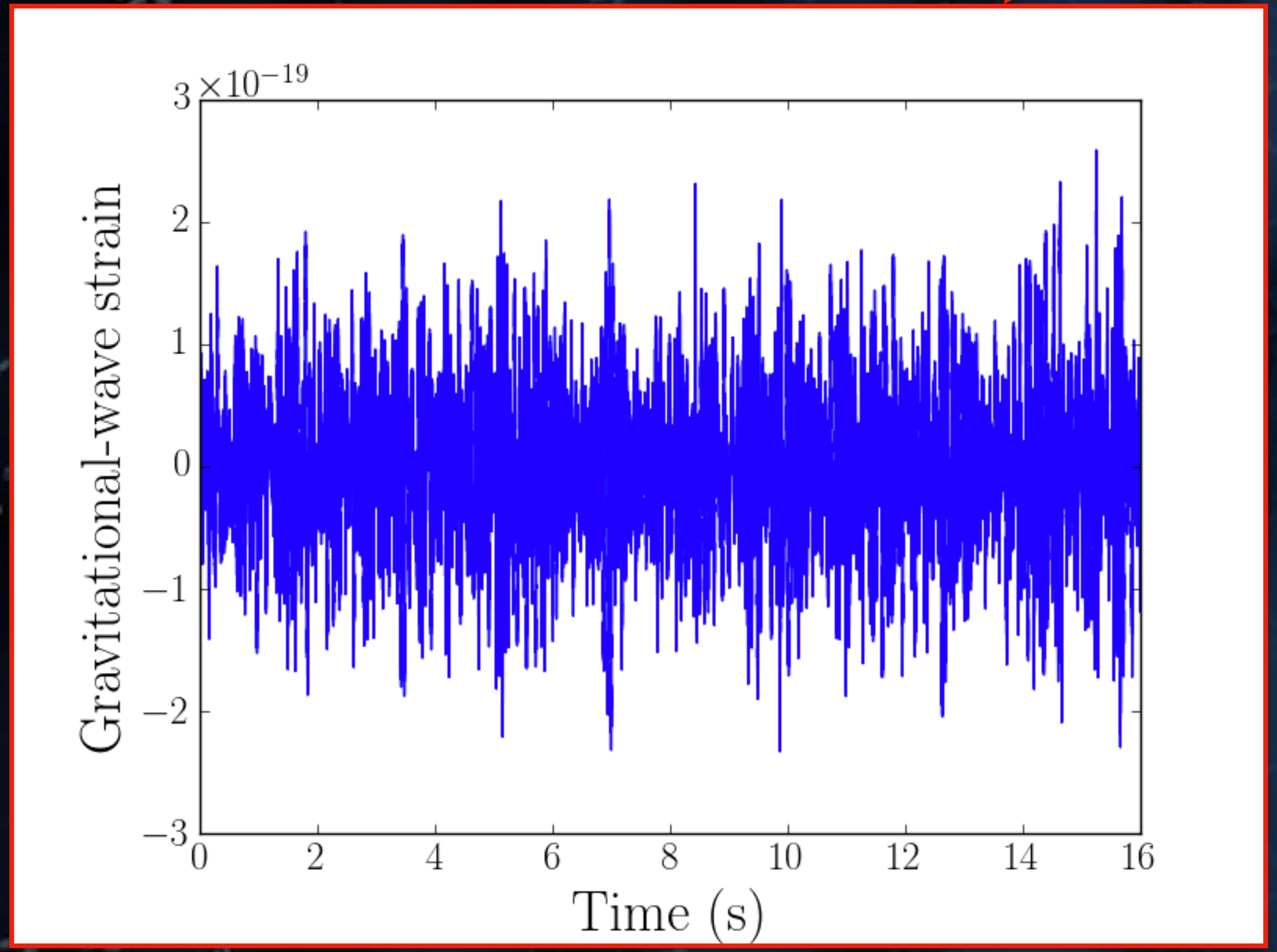

Filter: correlation between the template and the data

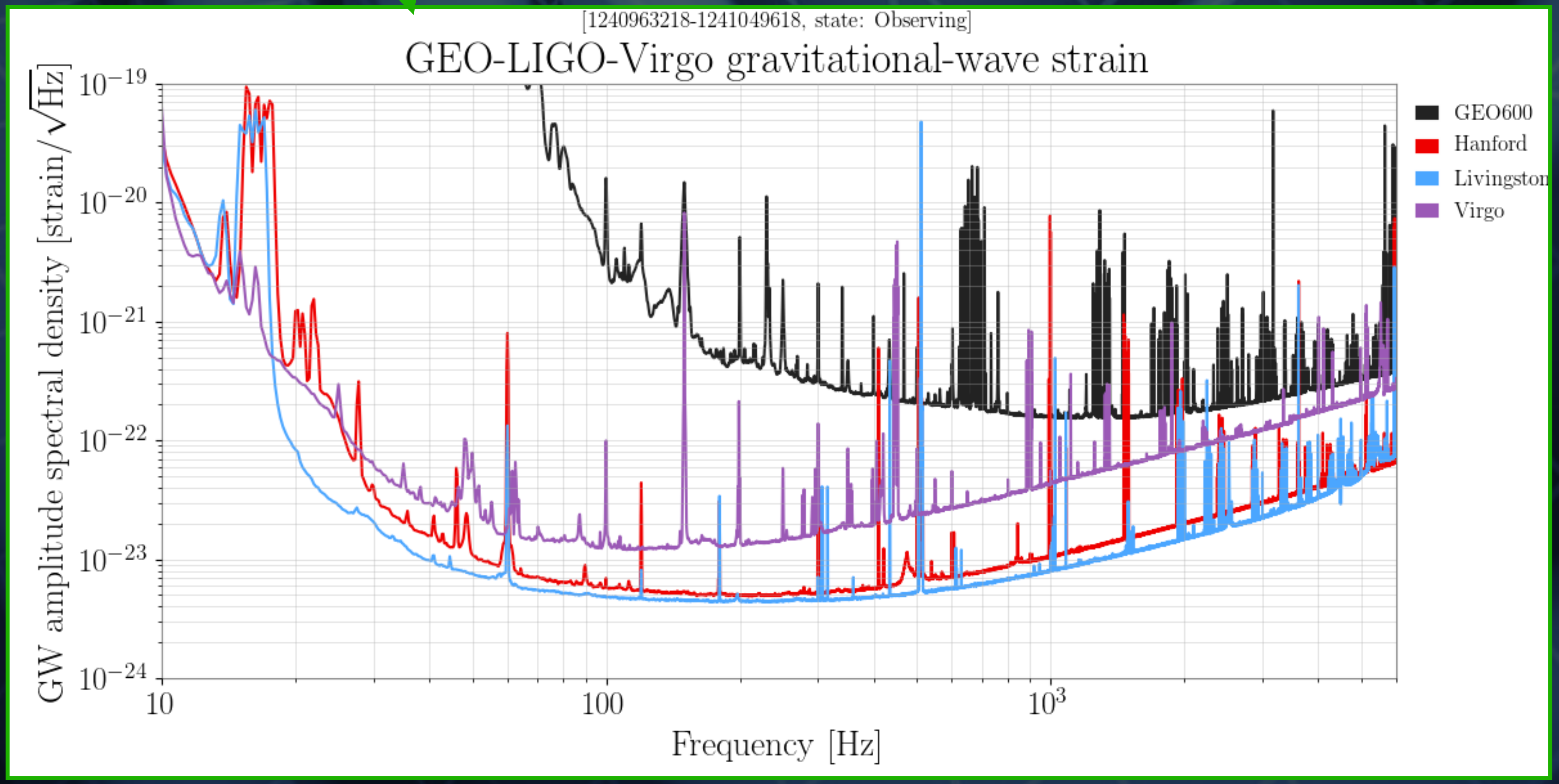


# Matched filtering (II)

$$(s|h) = 4\Re \int_0^\infty \frac{\tilde{s}(f)\tilde{h}^\star(f)}{S_h(f)}df$$




$$\tilde{a}(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} a(t) \exp^{-i2\pi f t} dt$$

# Acknowledgements

This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. The authors also gratefully acknowledge the support of the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Netherlands Organization for Scientific Research (NWO), for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors also gratefully acknowledge research support from these agencies as well as by the Council of Scientific and Industrial Research of India, the Department of Science and Technology, India, the Science & Engineering Research Board (SERB), India, the Ministry of Human Resource Development, India, the Spanish Agencia Estatal de Investigación, the Vicepresidència i Conselleria d'Innovació, Recerca i Turisme and the Conselleria d'Educació i Universitat del Govern de les Illes Balears, the Conselleria d'Innovació, Universitats, Ciència i Societat Digital de la Generalitat Valenciana and the CERCA Programme Generalitat de Catalunya, Spain, the National Science Centre of Poland and the European Union – European Regional Development Fund; Foundation for Polish Science (FNP), the Swiss National Science Foundation (SNSF), the Russian Foundation for Basic Research, the Russian Science Foundation, the European Commission, the European Regional Development Funds (ERDF), the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, the Hungarian Scientific Research Fund (OTKA), the French Lyon Institute of Origins (LIO), the Belgian Fonds de la Recherche Scientifique (FRS-FNRS), Actions de Recherche Concertées (ARC) and Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO), Belgium, the Paris Ile-de-France Region, the National Research, Development and Innovation Office Hungary (NKFIH), the National Research Foundation of Korea, the Natural Science and Engineering Research Council Canada, Canadian Foundation for Innovation (CFI), the Brazilian Ministry of Science, Technology, and Innovations, the International Center for Theoretical Physics South American Institute for Fundamental Research (ICTP-SAIFR), the Research Grants Council of Hong Kong, the National Natural Science Foundation of China (NSFC), the Leverhulme Trust, the Research Corporation, the Ministry of Science and Technology (MOST), Taiwan, the United States Department of Energy, and the Kavli Foundation. The authors gratefully acknowledge the support of the NSF, STFC, INFN and CNRS for provision of computational resources.

# GWOSC Impact

- Examples of projects using GWOSC data: gw-openscience.org/projects/
  - ✓ Professional research, student projects, classroom activities, text books, art projects, workshops, training
  - ✓ Around 3000 visitors each month (unique IP) and thousands of strain file downloads

**About 250 published papers acknowledge use of GWOSC (INSPIRE-HEP)**