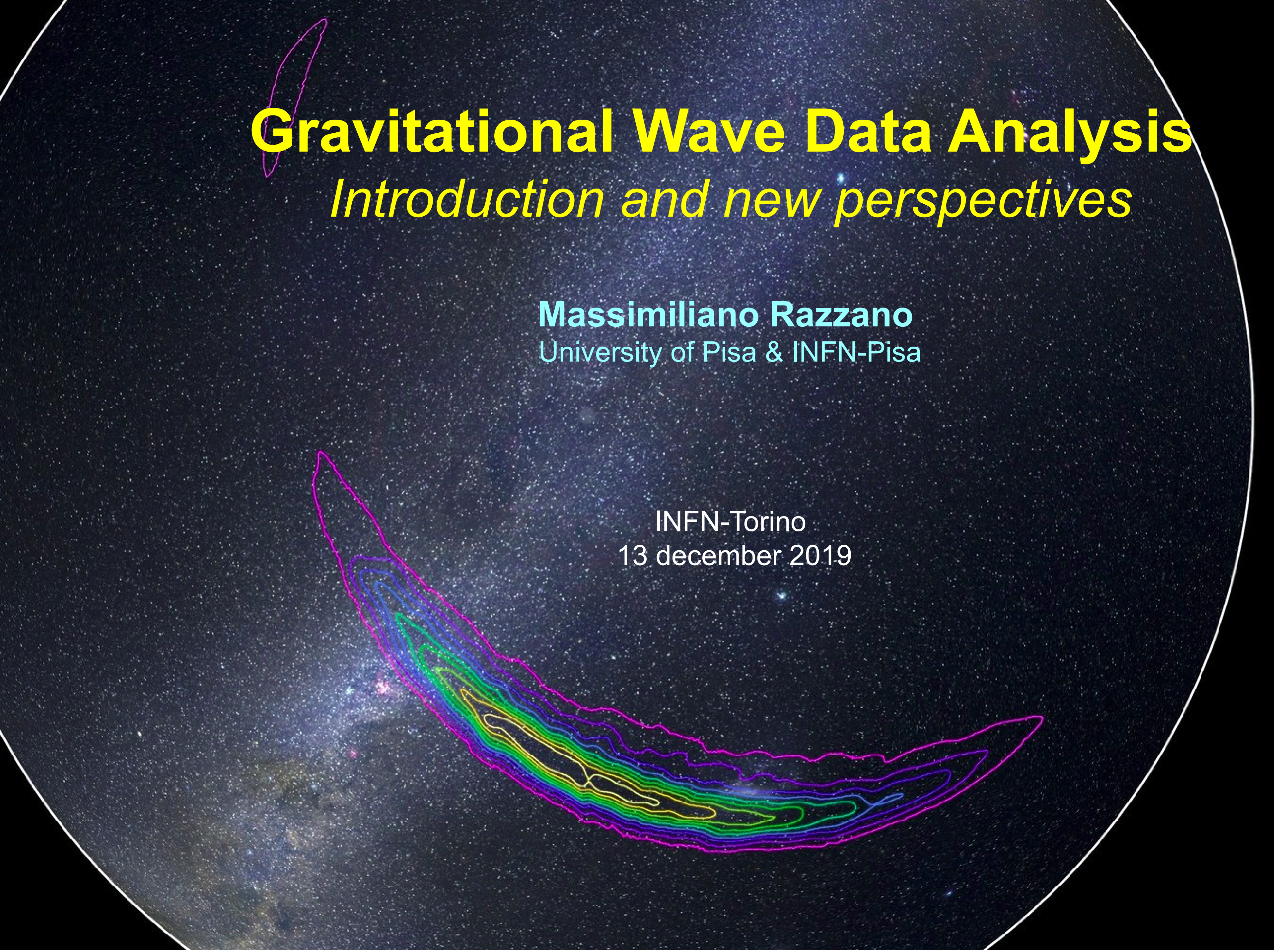
Gravitational Wave Data Analysis

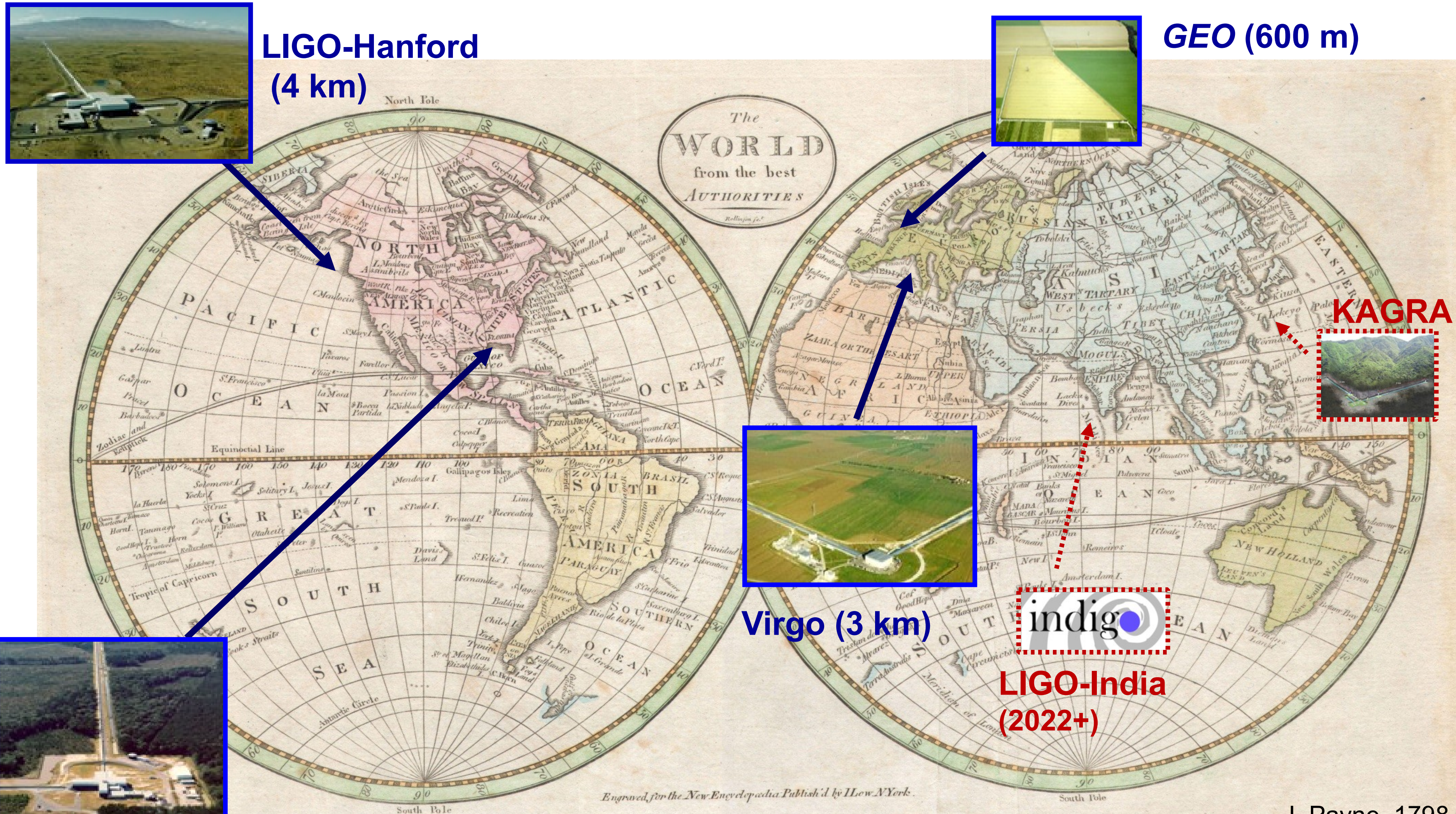
Introduction and new perspectives

Massimiliano Razzano
University of Pisa & INFN-Pisa

INFN-Torino
13 december 2019



The era of advanced GW detectors



M. Razzano

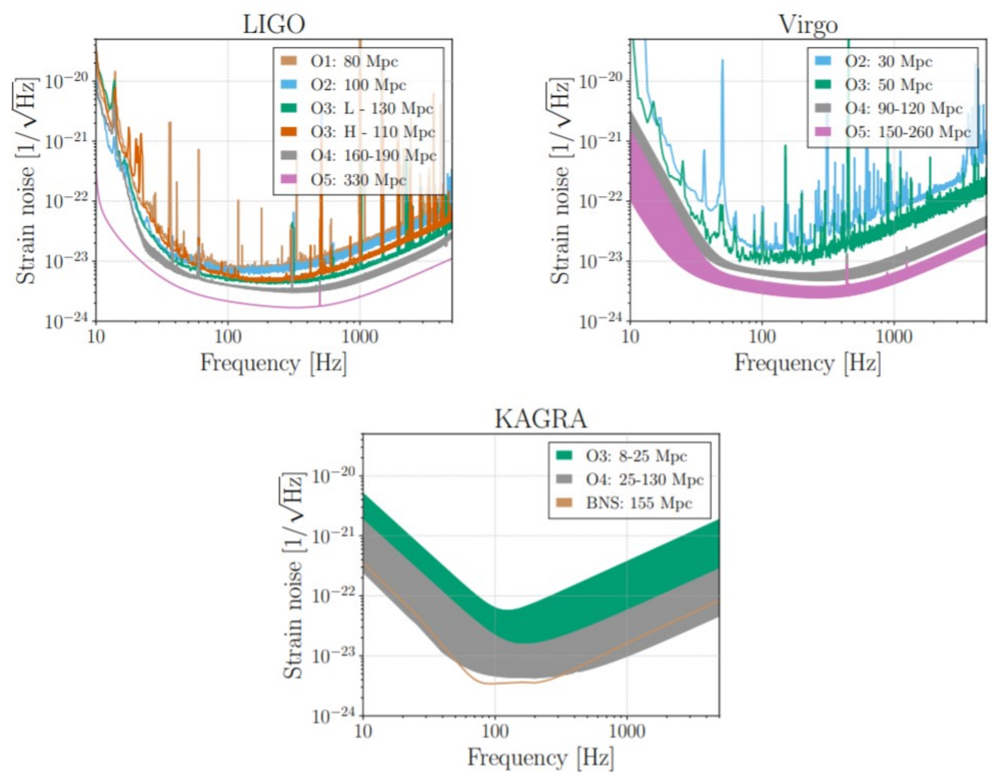
Why go to the Advanced detectors?

- Abbott et al. 2019 “Observing scenario” paper, LRR

Significant upgrade to increase sensivitity by 10x
Wrt “previous generation” LIGO and Virgo (2010s’)

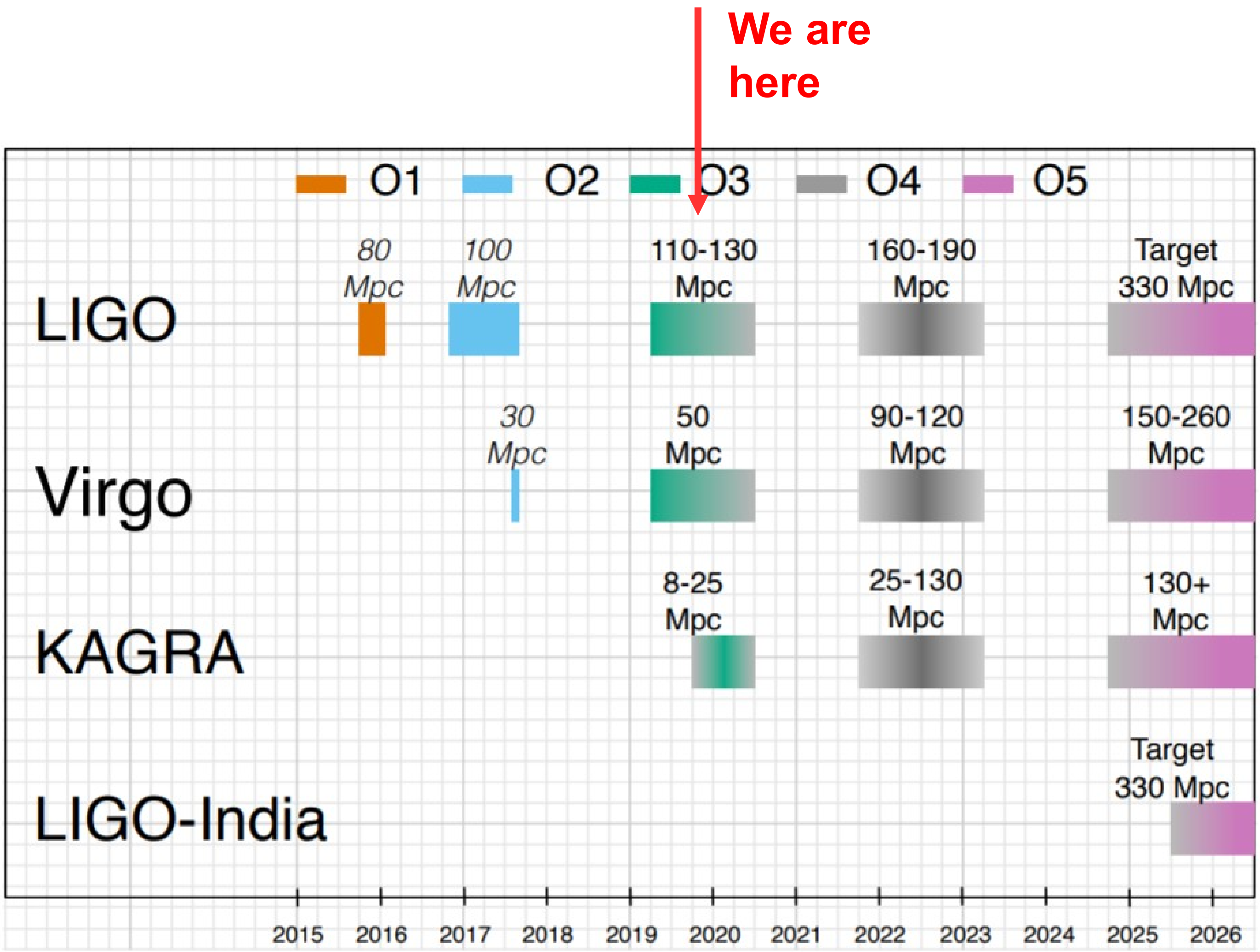
This means 10x distance reach
→ 10^3x larger volume
→ 10^3 number of events

Extremely tiny signals
Arm deformation ~10^{-18} m

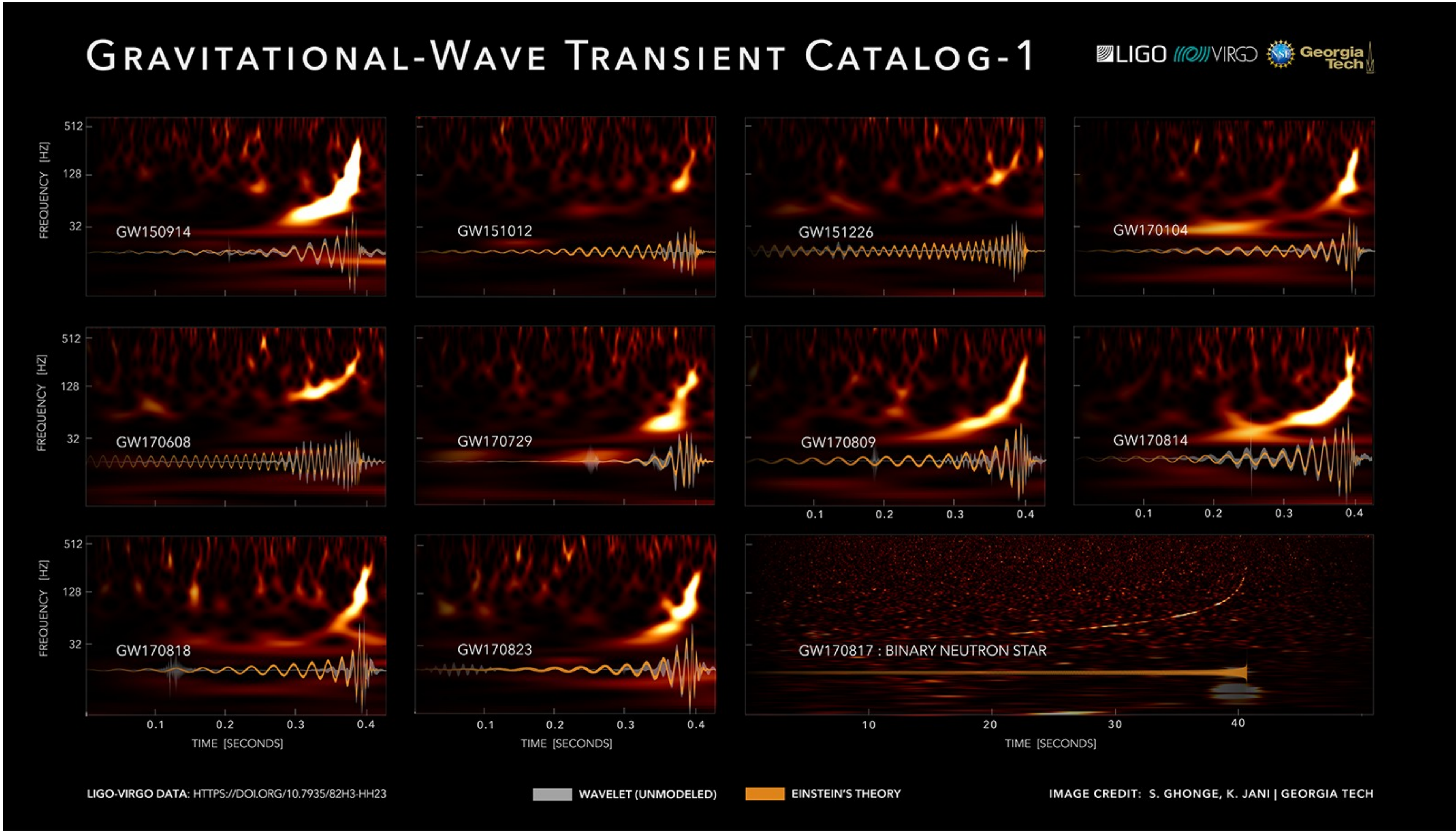


The era of Advanced GW detectors

- Abbott et al. 2019 “Observing scenario” paper, LRR



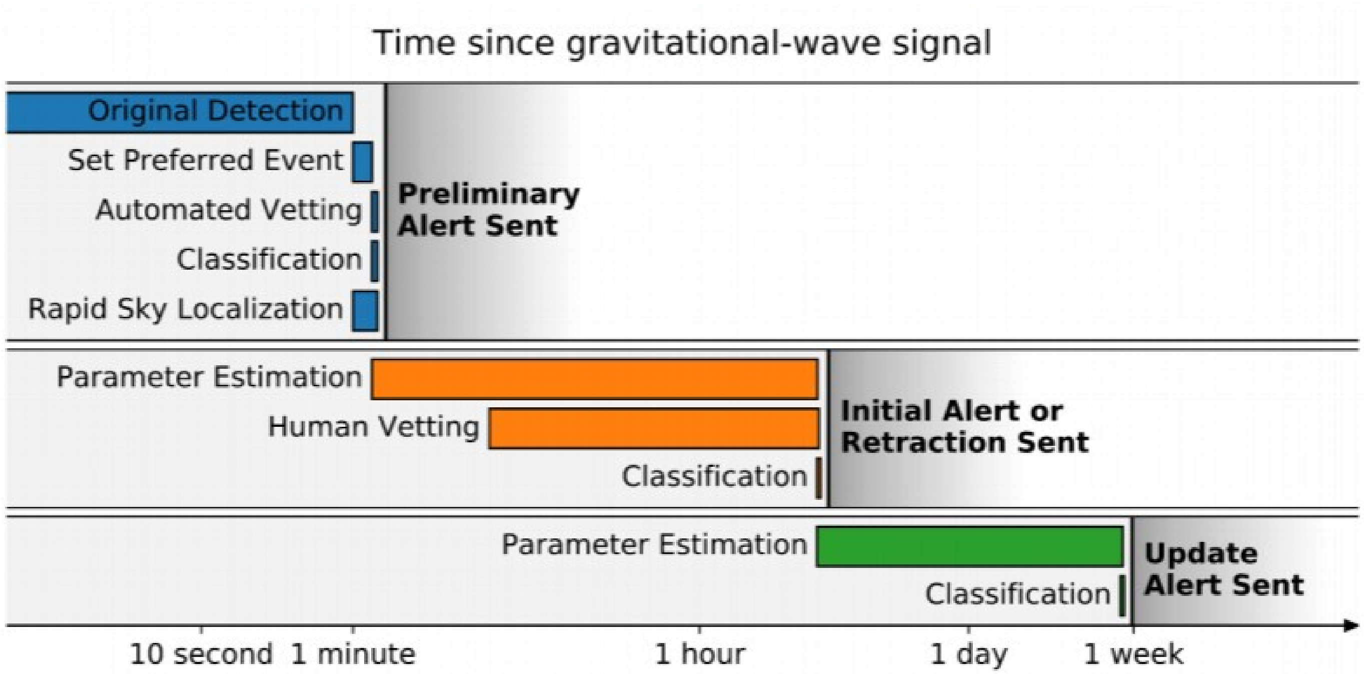
The first GW catalog of transients



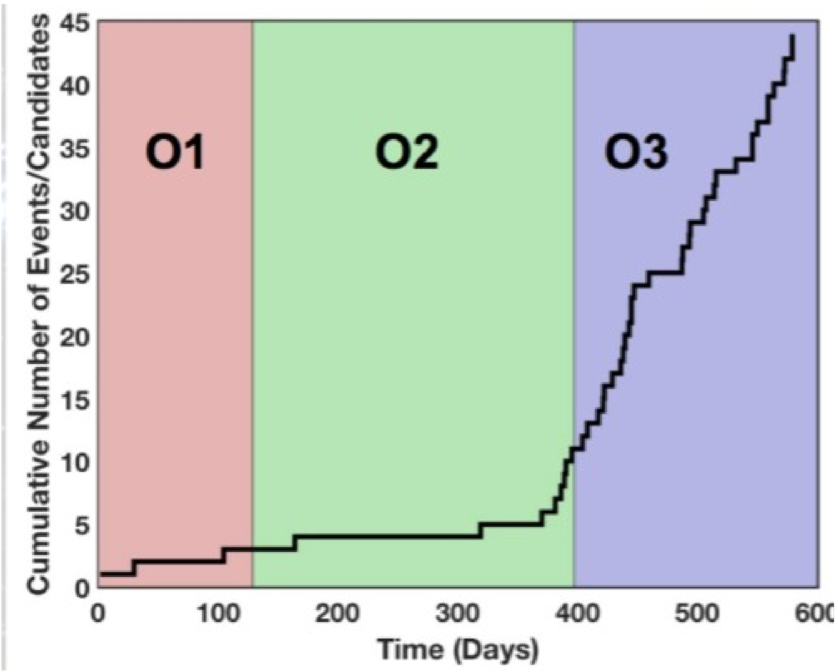
GWTC-1
Abbott et al 2019, arXiv:1811.12907

Open Public Alerts

- EM alerts are public now
- Public alerts user guide
  - https://emfollow.docs.ligo.org/userguide/
- Gravitational Wave Candidate Event Database (GraceDB)
  - https://gracedb.ligo.org/



Latest from O3



~40 alerts, ~10 retractions

- Mostly Binary Black holes,
- 7 candidates binary neuron stars or neutron star-black hole binary

Back to Einstein's field equations...

Uyuni Train Cemetery
(Bolivia)

Einstein's field equations

Set of 10 equations

$$\boxed{R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}} = \boxed{\frac{8\pi G}{c^4} T_{\mu\nu}}$$

Geometric part
(aka Einstein's tensor $G_{\mu\nu}$)
=
Geometry of spacetime

Stress-Enery part
(aka momentum-energy tensor)
=
Matter distribution

Linearized equations

We can linearize the field equation and choose a proper gauge (Lorenz Gauge) + orientation (TT)
The Einstein's equations become:

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}$$

where $\Box = -(1/c^2)\partial_t^2 + \nabla^2$

In vacuum (outside a source) $T_{\mu\nu} = 0$
In this case, solutions are plane waves

→ Gravitational waves!

Gravitational waves

In this gauge we have just 2 components.

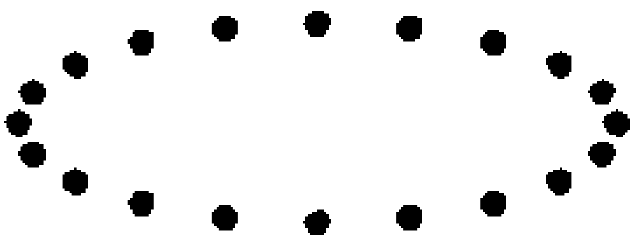
The solutions reads as:

$$h_{\mu\nu}^{TT}(t,z) = \begin{pmatrix} h_+ & h_x & 0 \\ h_x & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix}_{\mu\nu} \cos[\omega(t - z/c)]$$

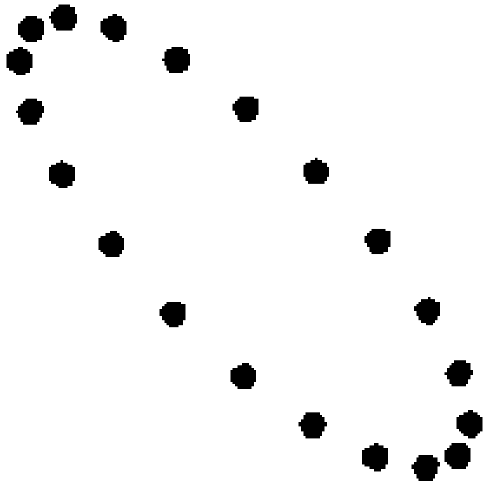
h_+ and h_x are the gravitational wave polarizations

GWs propagate at the speed of light

Gravitational waves Polarizations



Plus (+) polarization



Cross (x) polarization

Sources of Gravitational waves

Far from the source (i.e. the "far" zone), we can connect the GWs generated from the source, with the mass distribution of the source

In particular, under condition of slow motion and weak field, when we have a non-vanishing quadrupole momentum of the mass distribution, we have GW emission

$$h_{ij}^{TT}(t,z) \simeq \frac{2G}{c^4 r} \ddot{I}_{ij}^{TT}(t - r/c)$$

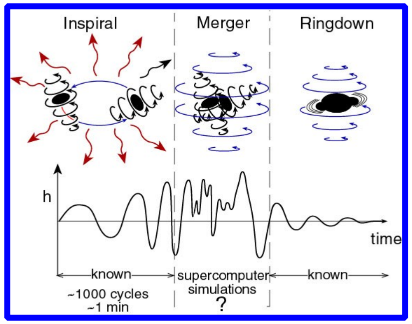
$G/c^4 \sim 10^{-49}\ s^2 g^{-2} cm^{-1}$

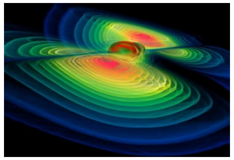
1/r dependence

Expected sources detectable by LIGO/Virgo

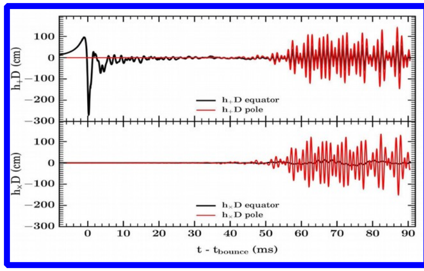
Transients

- **Coalescence of compact binary systems (NSs and/or BHs)**
  - Known waveforms (template banks)
  - $E_{gw} \sim 10^{-2}\ Mc^2$





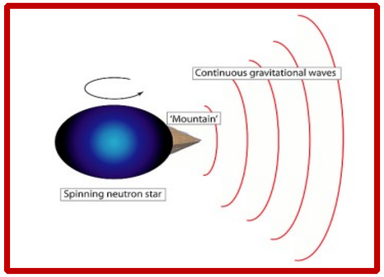
- **Core-collapse of massive stars**
  - Uncertain waveforms
  - $E_{gw} \sim 10^{-8} – 10^{-4}\ Mc^2$



Ott, C. 2009

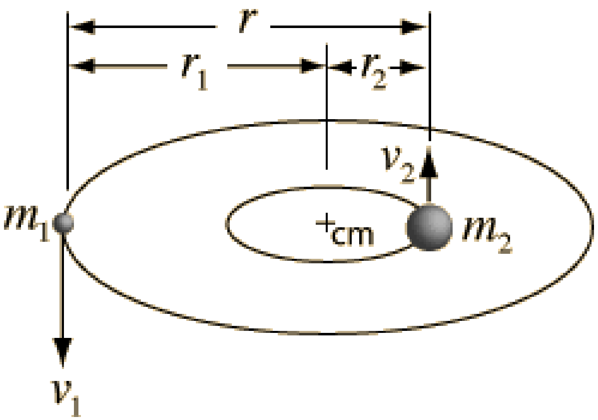
Non transients

- **Rotating neutron stars**
  - Quadrupole emission from star's asymmetry
  - Continuous and Periodic



- **Stochastic background**
  - Superposition of many signals (mergers, cosmological, etc)
  - Low frequency

The case of binary systems



Working out calculations (a = major semiaxis, μ = reduced mass):

$$\ddot{I}_{11} = -2\mu a^2 \omega^2 \cos[2(\omega t + \phi_0)]$$

and the corresponding wave component:

$$h_{+}^{TT} = h_{11}^{TT} = -\frac{4G\mu a^2\omega^2}{c^4 r} \cos[2(\omega t + \phi_0)]$$

Twice the frequency of the orbit

System loses energy by GWs → orbital contraction!

Binary coealescence

Orbital decay → Closer orbit → Faster spin → More GW emission →Larger orbital decay

→Runaway process

$$f_{dot,GW} = k\, f_{GW}^{11/3}$$

In some time, the 2 object will coalesce and merge (e.g. PSR B1916+13 in 300 Myr)

GW from chirp

We define phase as

$$\Phi(t) = \int_{t_0}^{t} \omega_{GW}(t) dt$$

Then the h(t) becomes (as usual, i is the inclination):

$$h_+(t) = \frac{1}{r}\left(\frac{GM_c}{c^2}\right)^{5/4}\left(\frac{5}{c\tau}\right)^{1/4}\left(\frac{1+\cos^2 i}{2}\right)\cos[\Phi(\tau)]$$

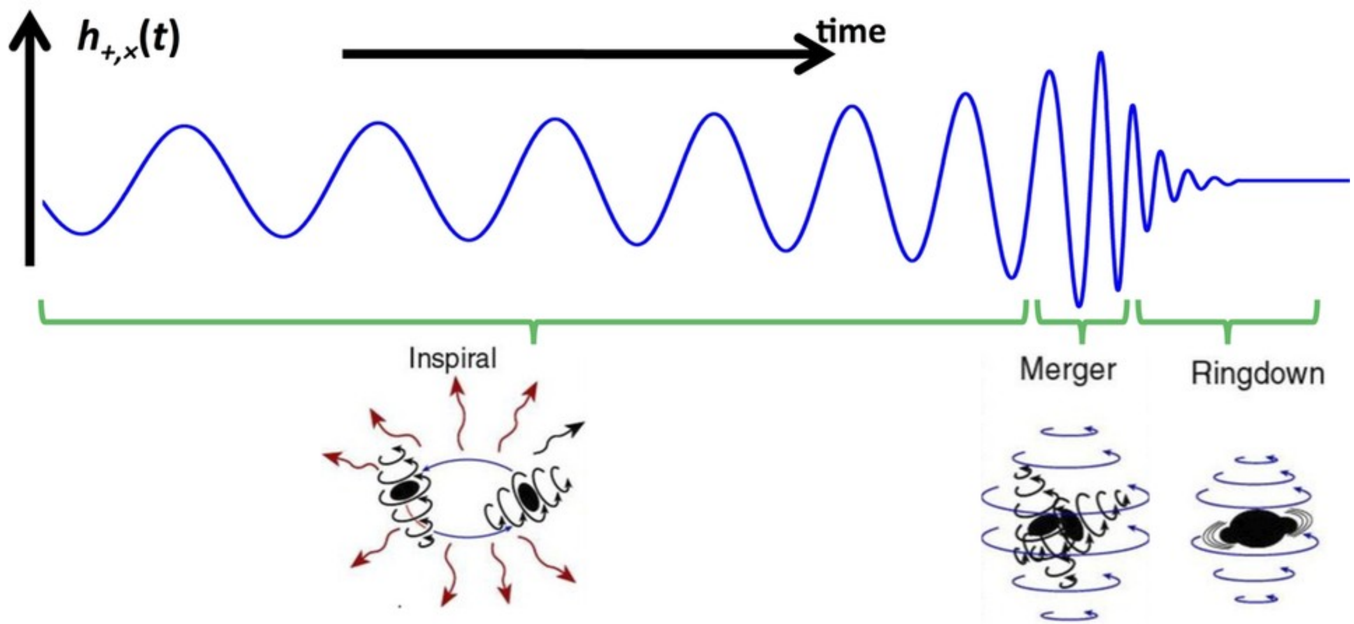
$$h_X(t) = \frac{1}{r}\left(\frac{GM_c}{c^2}\right)^{5/4}\left(\frac{5}{c\tau}\right)^{1/4}\cos i \sin[\Phi(\tau)]$$

Binary coealescence

Closer to the coalescence, GW amplitude and frequency increase → the so-called *chirp*

- *No longer easy to model as point sources*
- *After the merger, the final object (e.g. a BH) will undergo a ringdown phase*

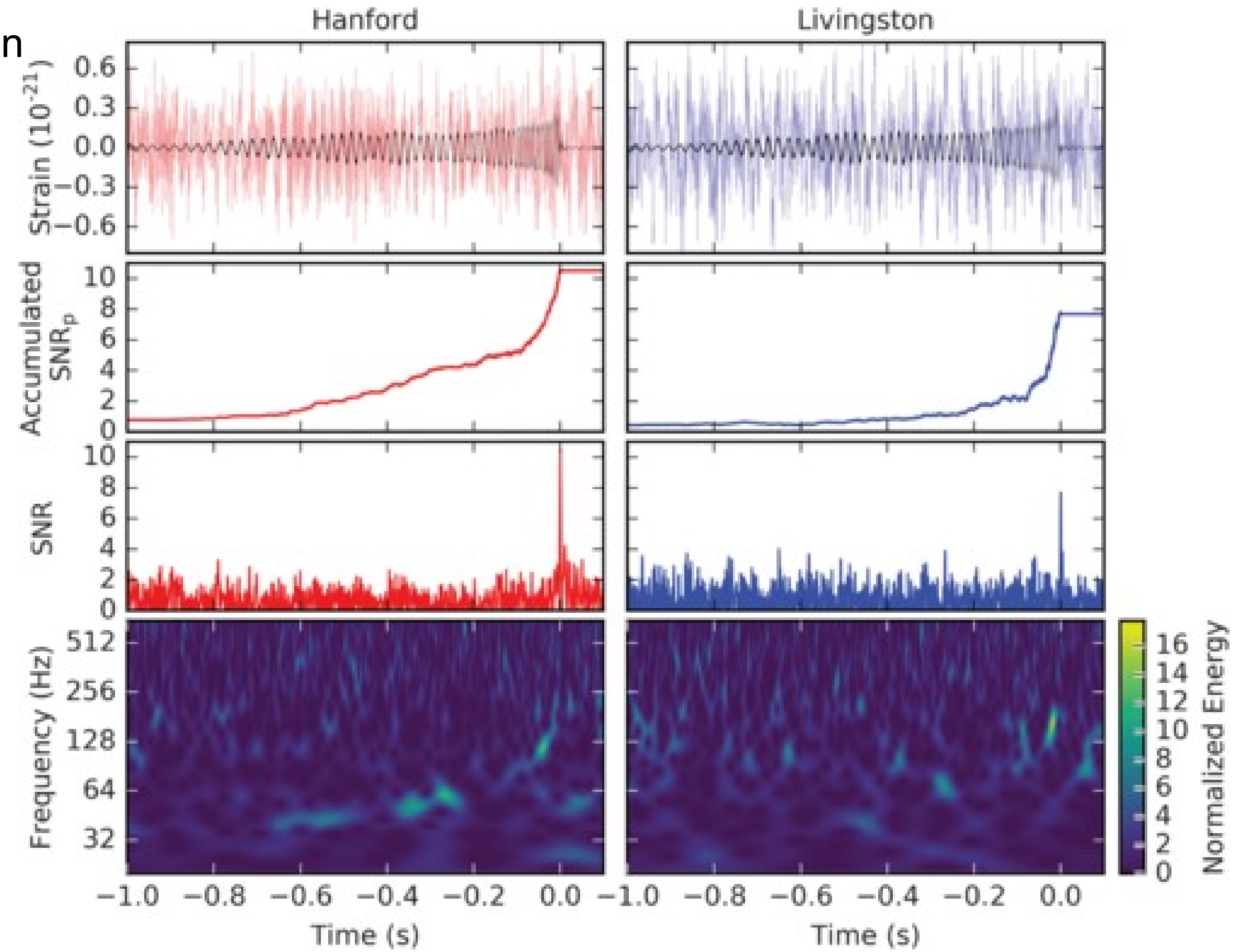
Can we use
This information?



M. Favata/SXS/K. Thorne

Signals buried in noise

Low signals + high noise
→ Detection
→ Characterization



GW151226

*

Random processes

Sequence of random variables
Instrumental noise is an example of random process timeseries x(t)

If we know the probability density function p(x), we can evaluate the expectation value:

$$\langle x \rangle := \int x \, p_x(x) dx \, .$$

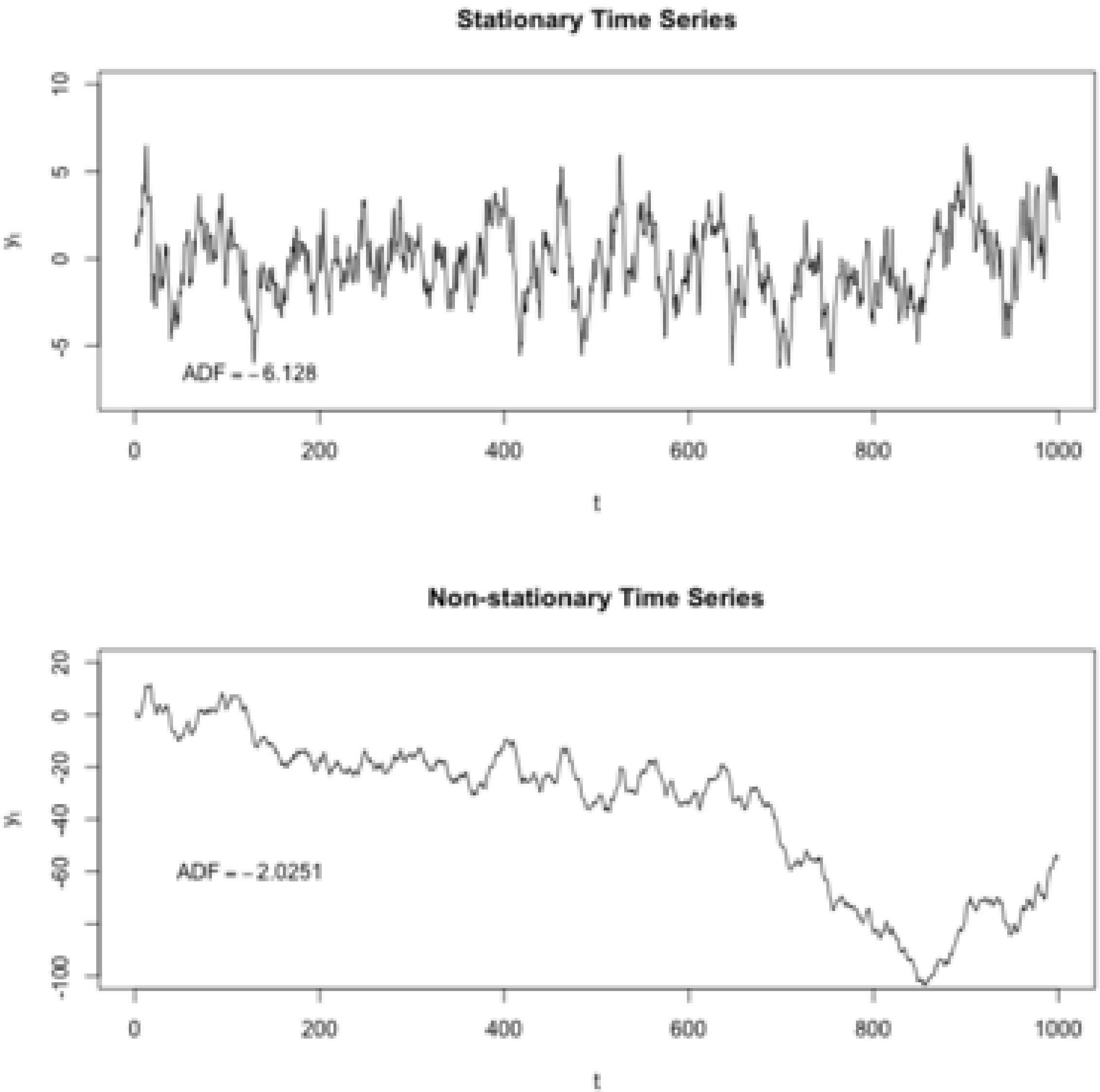
If the statistical properties of the signal do not change, we say it is *stationary*, and

$$\langle x \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt \, .$$

Statistical ensemble average
=
Long time average

*

Random processes



*

Power Spectrum

We can evaluate the average of $x^2(t)$

$$\langle x^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)\,dt\ .$$

Windowed signal

$$x_T(t) = \begin{cases} x(t) & -T/2 < t < T/2 \\ 0 & \text{otherwise.} \end{cases}$$

$$\langle x^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} x_T^2(t)\,dt$$

Parseval's Theorem

$$= \lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} |\tilde{x}_T^2(f)|^2\,df$$

x(t) is real → symmetry:

$$\tilde{x}_T(-f) = \tilde{x}_T^*(f)$$

$$= \lim_{T\to\infty} \frac{2}{T} \int_{0}^{\infty} |\tilde{x}_T^2(f)|^2\,df$$

$$= \int_{0}^{\infty} \boxed{S_x(f)}\,df\ ,$$

Power Spectral Density (PSD)

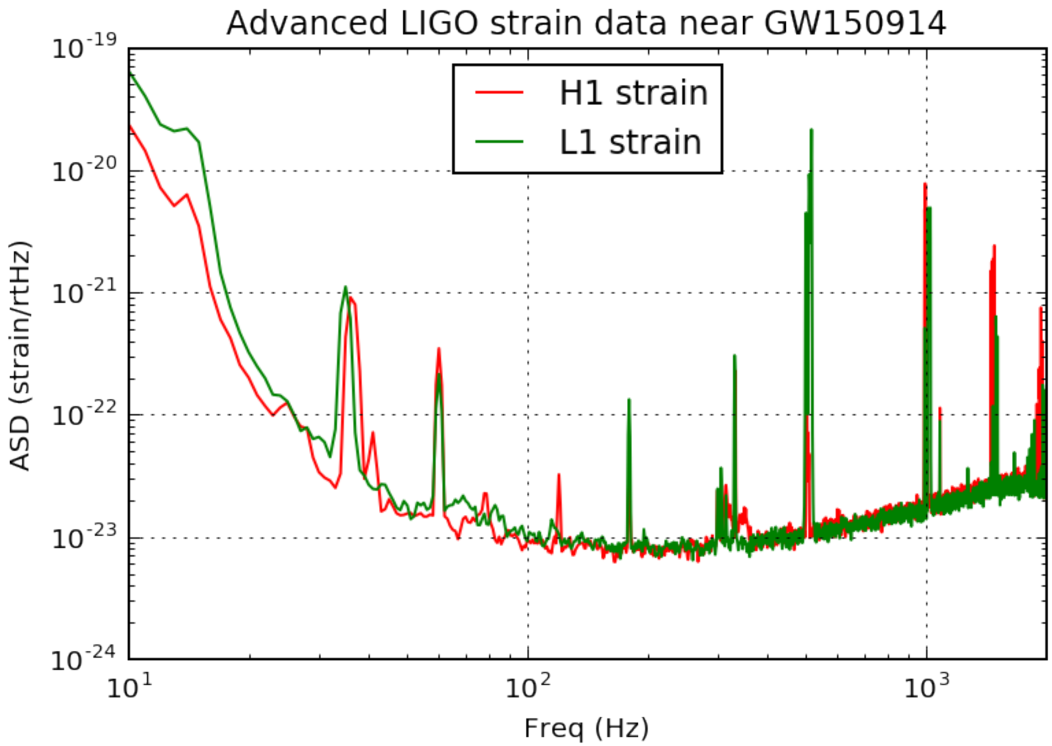
*

Power Spectrum

The definition of S(f) becomes:

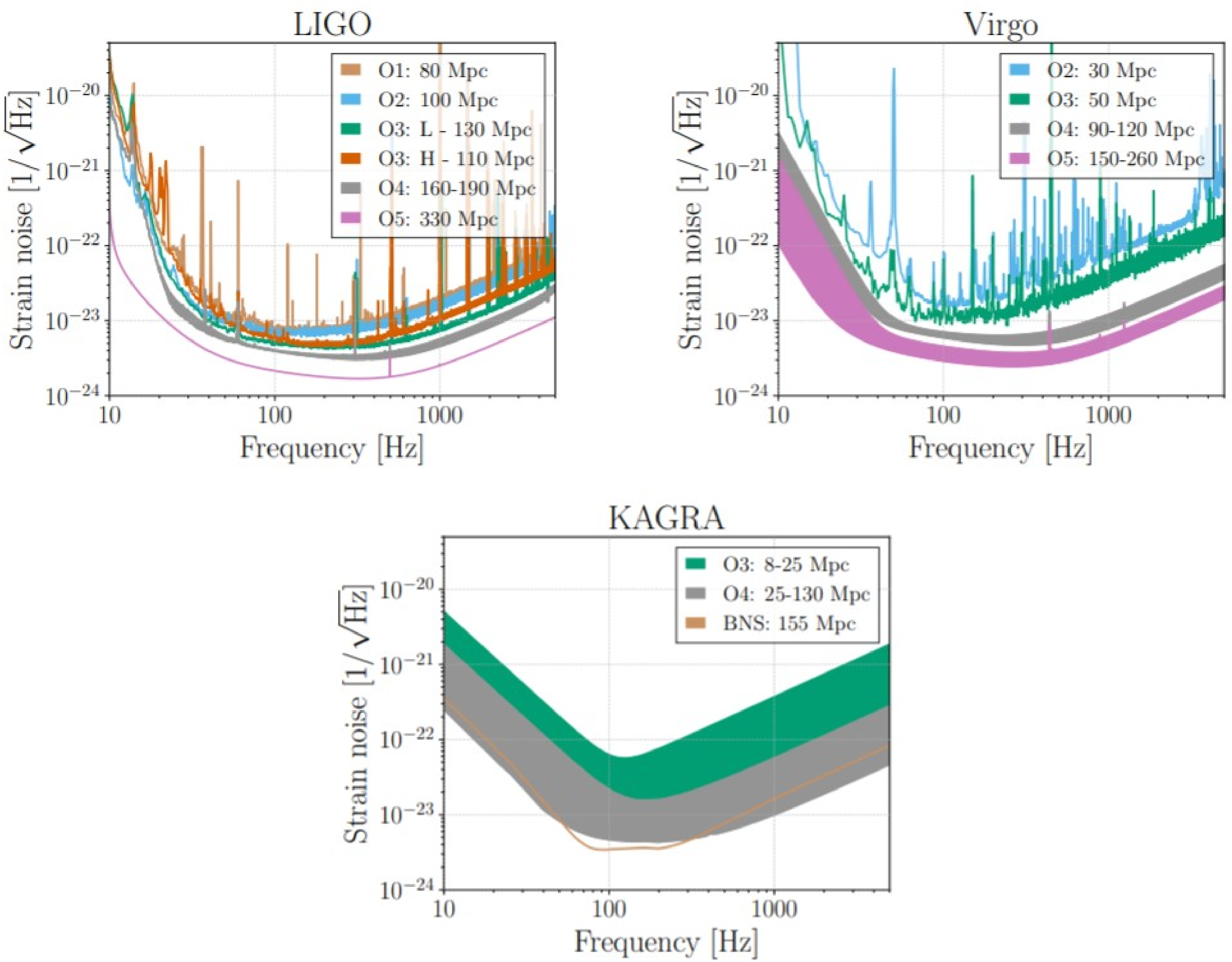
$$S_x(f) := \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} x(t) e^{-2\pi i f t} dt \right|^2$$

The amplitude spectral density (ASD) is the sqrt of PSD



*

Power Spectrum



*

Optimal detection filter

We want to test against null hypothesis (H0=No signal)
Using Bayes theorem one can calculate the likelihood ratio under assumption
Of gaussian noise.

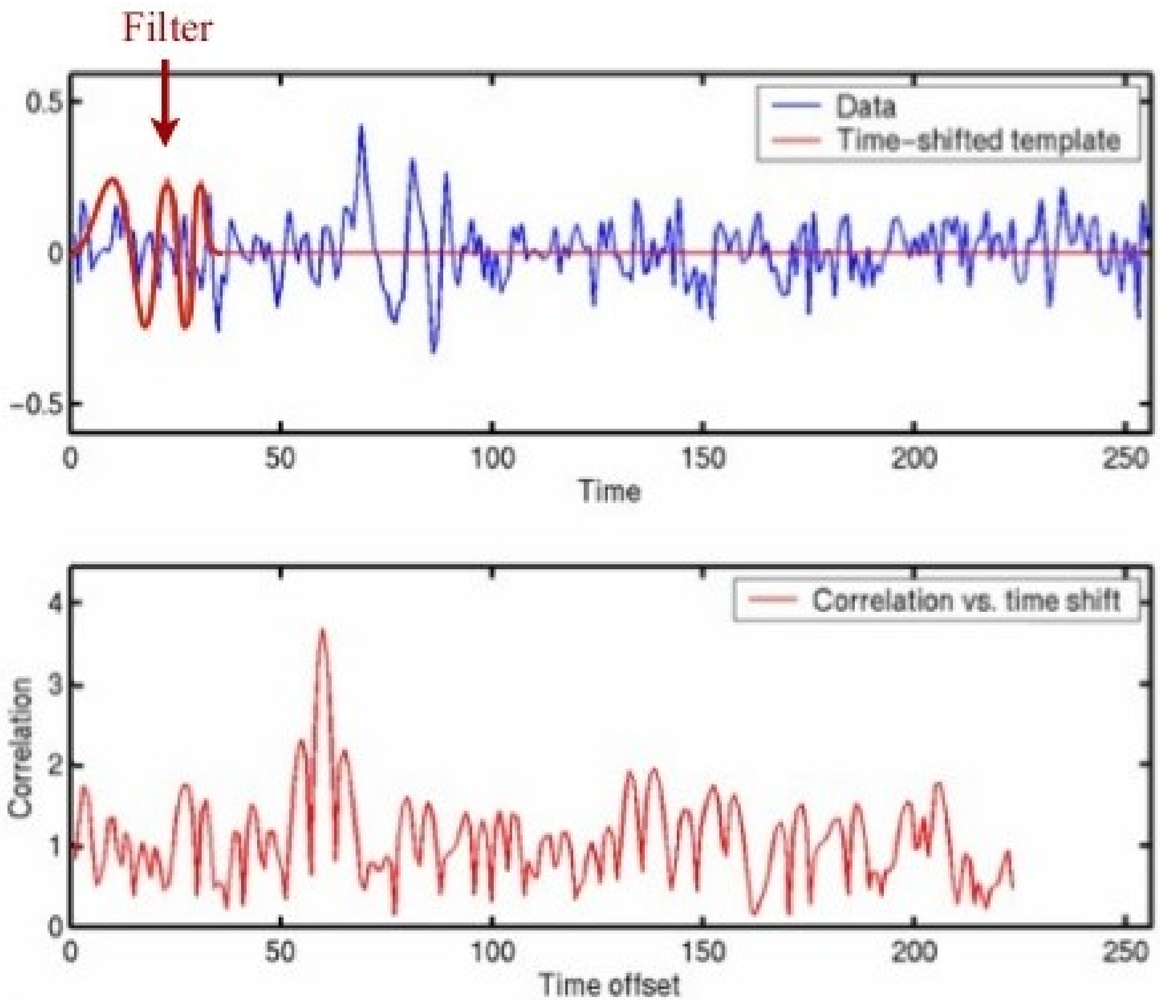
Likelihood depends monotonically on (s,h), therefore we call it optimal statistic

$$(s,h) = 4\,\mathrm{Re}\int_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)}\,df$$

Also called this **matched filter**
(a noise-weighted correlation of anticipated signal with data)

*

An example of matched filter



slide courtesy of Damir Buskulic

*

GW open data on the web

Hosted at the Gravitational Wave Open Science Center (GWOSC)
https://www.gw-openscience.org/

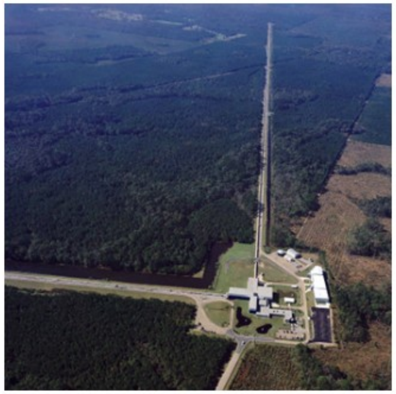
Gravitational Wave Open Science Center

Data▾ Software▾ Online Status▾ About GWOSC▾

The Gravitational Wave Open Science Center provides data from gravitational-wave observatories, along with access to tutorials and software tools.

LIGO Hanford Observatory, Washington

(Credits: C. Gray)



LIGO Livingston Observatory, Louisiana

(Credits: J. Giaime)

Virgo detector, Italy

(Credits: Virgo Collaboration)

Get started!

NEW **LIGO/Virgo alerts began April 2, 2019**

Released data

GWOSC provides two main types of data

- GW related to events (e.g. Binary Black Holes, etc)
  - About 1-hour window centered on the event(s)
  - Released with the publication of the event(s)
  - GW Strain data, size ~Gb

- GW "bulk" data
  - Bulk datasets of each observing run (size ~Tb)
  - Releases after 18 months from the end of the run
  - Data blocks of 6 months, released every 6 months
  - First chunck of O3 will come in April 2021

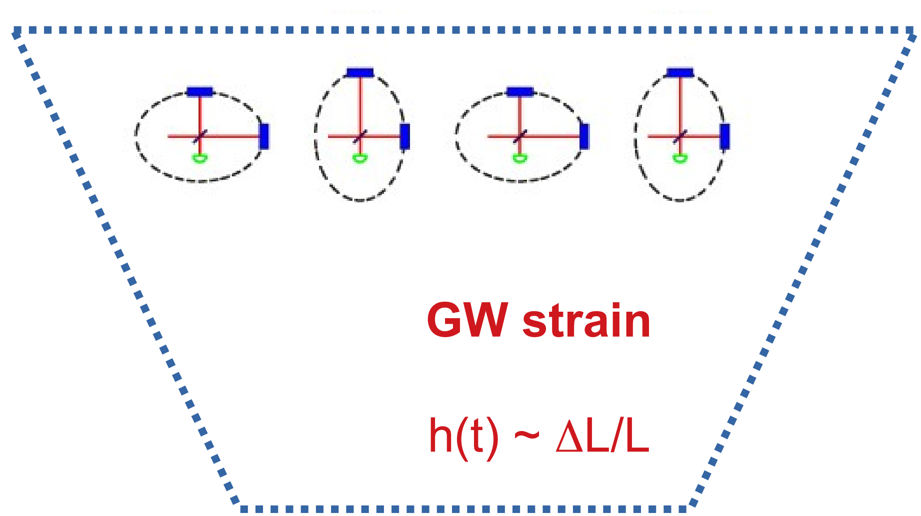
- Supporting documentation and tools
  - Help the external community in using data
  - Lots of tutorials
  - Materials from periodic Open Data Workshops (Last one this April in Paris)

GW data products

- Releases include GW strain, data quality and injections
  - Timeseries
  - Various formats, including standard "frame" files (GWF) and HDF5

Available Releases

- LIGO
  - S5 (2005 - 2007)
  - S6 (2009 – 2010)
- Advanced LIGO
  - O1 (2015 – 2016)
  - O2 (2016 – 2017)
- Advanced Virgo
  - O2 (2016 – 2017)



Timelines

- Information on data availability over time

Gravitational Wave Open Science Center

Data ▾ Software ▾ Online Status ▾ About GWOSC ▾

Strain Data
Catalogs
Timelines

Timeline

Timeline technology provides instant access to time-based metadata on scales from seconds to centuries. **Timeline** is developed by the Gravitational Wave Open Science Center, which is supported by the U.S. National Science Foundation.

Welcome to **Timeline**. This page provides information on when LIGO was collecting science mode data, as well as data quality classifications and injection times.

Timeline Queries

- Use the **All Timeline Query Form** to request any of the Timeline or Segment Lists
- Use the **Run Timeline Query Form** to request any of the Run Timeline or Segment Lists.
- Use the **Pre-Catalog Event Timeline Query Form** to request any of the Pre-Catalog Event Timeline or Segment Lists.
- Use the **Catalog Timeline Query Form** to request any of the Catalog Timeline or Segment Lists.
- Use the **Marginal CBC Trigger Timeline Query Form** to request any of the Marginal CBC Trigger Timeline or Segment Lists.

Timeline Quick Links

Some common example plots are linked below:

Science Mode Timelines

- Five detectors since 2005

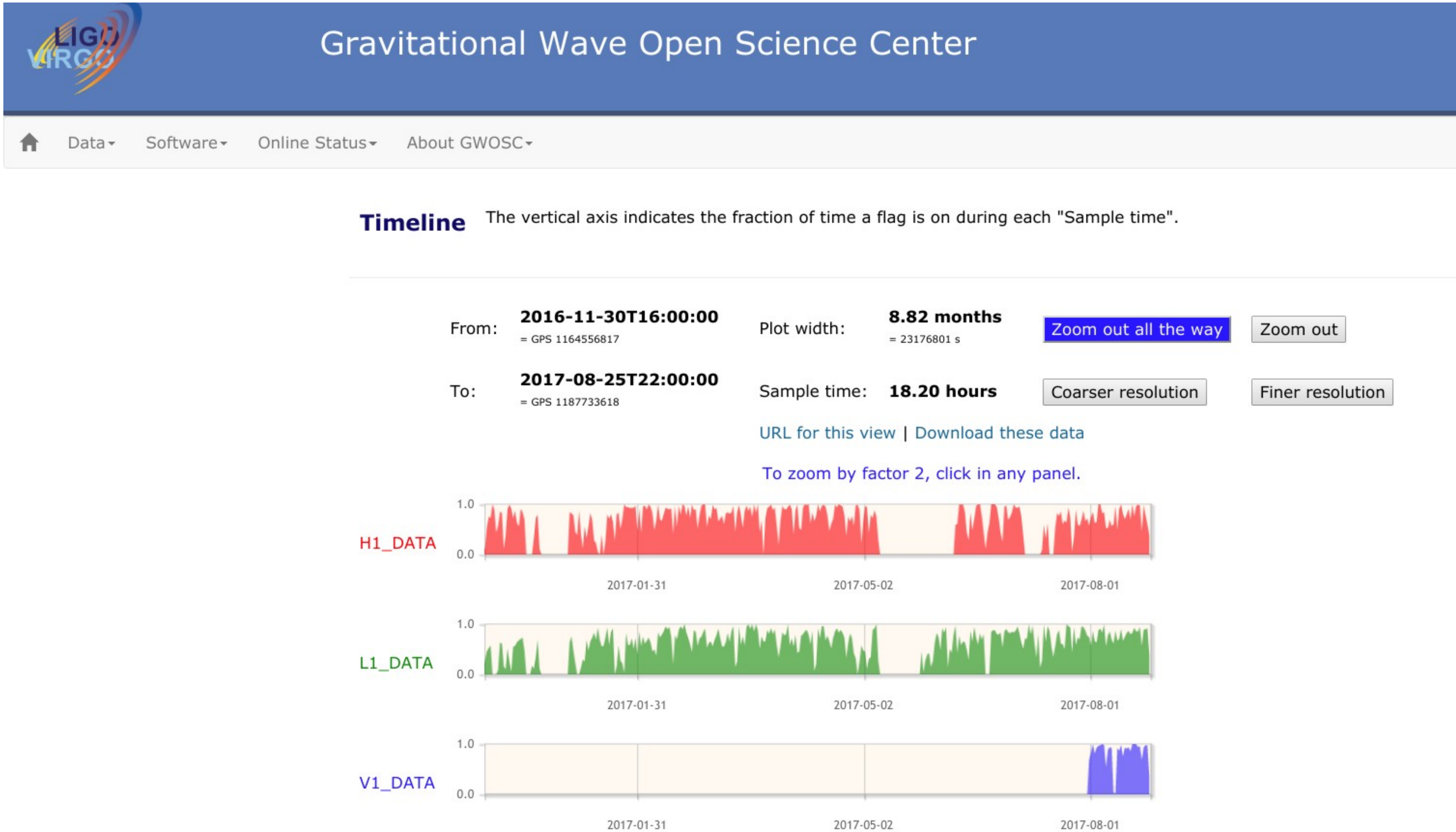
Timelines from the O2 run, 2016-2017

- Three detectors over the O2 run
- Passes O2 Burst checks for H1, L1, V1
- Passes O2 CBC checks for H1, L1, V1
- Times with no Continuous-Wave injections

Timelines from the O1 run, 2015-2016

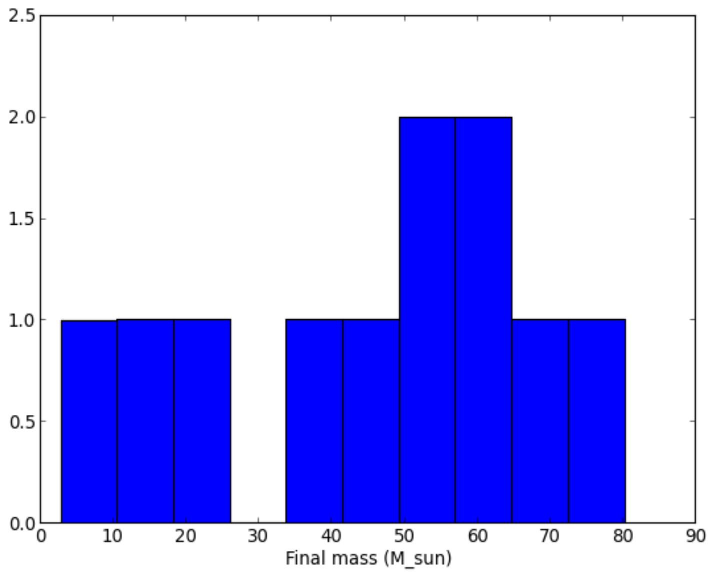
- Two detectors over the O1 run
- Passes O1 Burst checks for H1, L1
- Passes O1 CBC checks for H1, L1

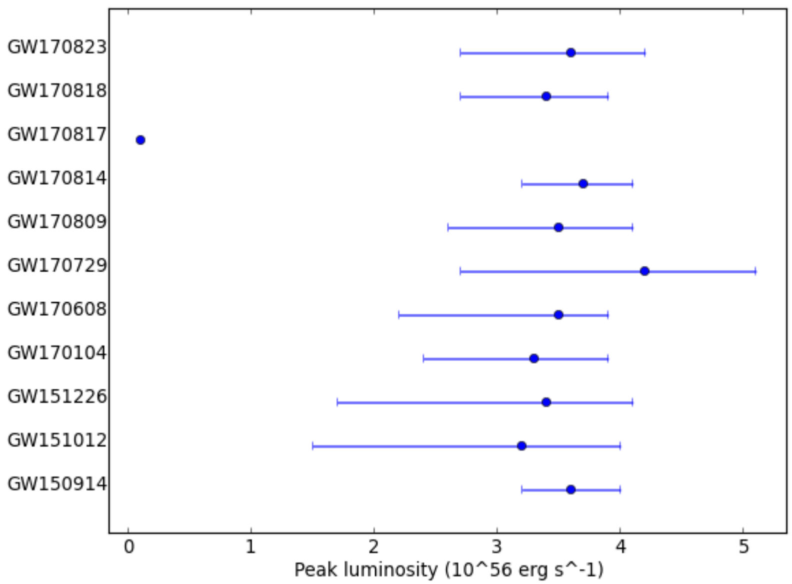
Timelines – example from O2



Strain data – Catalogs

- Previously, each event published separately, now included in catalogs
  - https://www.gw-openscience.org/catalog/
- GWTC-1
  - Released in Dec 2018 (arXiv:1811.12907)
  - 11 confident detections + 14 marginal triggers
  - Strain data + skymaps, etc...





Strain data – Catalogs

Event	Primary mass (M_sun)	Secondary mass (M_sun)	Effective inspiral spin	chirp mass (M_sun)	Final spin	Final mass (M_sun)	Luminosity distance (Mpc)	GPS time (s)	Radiated energy (M_sun X c^2)	Peak luminosity (10^56 erg s^-1)	FAR cWB (yr^-1)	FAR gstLAL (yr^-1)	FAR PyCBC (yr^-1)	Source redshift	Sky localization (deg^2)	Network SNR cWB	Network SNR gstLAL	Network SNR PyCBC	UTCtime
GW150914	$35.6^{+4.8}_{-3.0}$	$30.6^{+3.0}_{-4.4}$	$-0.01^{+0.12}_{-0.13}$	$28.6^{+1.6}_{-1.5}$	$0.69^{+0.05}_{-0.04}$	$63.1^{+3.3}_{-3.0}$	430^{+150}_{-170}	1126259462.4	$3.1^{+0.4}_{-0.4}$	$3.6^{+0.4}_{-0.4}$	$<$ 1.63e-04	$<$ 1.00e-07	$<$ 1.53e-05	$0.09^{+0.03}_{-0.03}$	179	25.2	24.4	23.6	09:50:45.4
GW151012	$23.3^{+14.0}_{-5.5}$	$13.6^{+4.1}_{-4.8}$	$0.04^{+0.28}_{-0.19}$	$15.2^{+2.0}_{-1.1}$	$0.67^{+0.13}_{-0.11}$	$35.7^{+9.9}_{-3.8}$	1060^{+540}_{-480}	1128678900.4	$1.5^{+0.5}_{-0.5}$	$3.2^{+0.8}_{-1.7}$	NA	7.92e-03	0.17	$0.21^{+0.09}_{-0.09}$	1555	NA	10.0	9.5	09:54:43.4
GW151226	$13.7^{+8.8}_{-3.2}$	$7.7^{+2.2}_{-2.6}$	$0.18^{+0.20}_{-0.12}$	$8.9^{+0.3}_{-0.3}$	$0.74^{+0.07}_{-0.05}$	$20.5^{+6.4}_{-1.5}$	440^{+180}_{-190}	1135136350.6	$1.0^{+0.1}_{-0.2}$	$3.4^{+0.7}_{-1.7}$	0.02	$<$ 1.00e-07	$<$ 1.69e-05	$0.09^{+0.04}_{-0.04}$	1033	11.9	13.1	13.1	03:38:53.6
GW170104	$31.0^{+7.2}_{-5.6}$	$20.1^{+4.9}_{-4.5}$	$-0.04^{+0.17}_{-0.20}$	$21.5^{+2.1}_{-1.7}$	$0.66^{+0.08}_{-0.10}$	$49.1^{+5.2}_{-3.9}$	960^{+430}_{-410}	1167559936.6	$2.2^{+0.5}_{-0.5}$	$3.3^{+0.6}_{-0.9}$	2.91e-04	$<$ 1.00e-07	$<$ 1.37e-05	$0.19^{+0.07}_{-0.08}$	924	13.0	13.0	13.0	10:11:58.6
GW170608	$10.9^{+5.3}_{-1.7}$	$7.6^{+1.3}_{-2.1}$	$0.03^{+0.19}_{-0.07}$	$7.9^{+0.2}_{-0.2}$	$0.69^{+0.04}_{-0.04}$	$17.8^{+3.2}_{-0.7}$	320^{+120}_{-110}	1180922494.5	$0.9^{+0.0}_{-0.1}$	$3.5^{+0.4}_{-1.3}$	1.44e-04	$<$ 1.00e-07	$<$ 3.09e-04	$0.07^{+0.02}_{-0.02}$	396	14.1	14.9	15.4	02:01:16.5
GW170729	$50.6^{+16.6}_{-10.2}$	$34.3^{+9.1}_{-10.1}$	$0.36^{+0.21}_{-0.25}$	$35.7^{+6.5}_{-4.7}$	$0.81^{+0.07}_{-0.13}$	$80.3^{+14.6}_{-10.2}$	2750^{+1350}_{-1320}	1185389807.3	$4.8^{+1.7}_{-1.7}$	$4.2^{+0.9}_{-1.5}$	0.02	0.18	1.36	$0.48^{+0.19}_{-0.20}$	1033	10.2	10.8	9.8	18:56:29.3
GW170809	$35.2^{+8.3}_{-6.0}$	$23.8^{+5.2}_{-5.1}$	$0.07^{+0.16}_{-0.16}$	$25.0^{+2.1}_{-1.6}$	$0.70^{+0.08}_{-0.09}$	$56.4^{+5.2}_{-3.7}$	990^{+320}_{-380}	1186302519.8	$2.7^{+0.6}_{-0.6}$	$3.5^{+0.6}_{-0.9}$	NA	$<$ 1.00e-07	1.45e-04	$0.20^{+0.05}_{-0.07}$	340	NA	12.4	12.2	08:28:21.8
GW170814	$30.7^{+5.7}_{-3.0}$	$25.3^{+2.9}_{-4.1}$	$0.07^{+0.12}_{-0.11}$	$24.2^{+1.4}_{-1.1}$	$0.72^{+0.07}_{-0.05}$	$53.4^{+3.2}_{-2.4}$	580^{+160}_{-210}	1186741861.5	$2.7^{+0.4}_{-0.3}$	$3.7^{+0.4}_{-0.5}$	$<$ 2.08e-04	$<$ 1.00e-07	$<$ 1.25e-05	$0.12^{+0.03}_{-0.04}$	87	17.2	15.9	16.3	10:30:43.5
GW170817	$1.46^{+0.12}_{-0.10}$	$1.27^{+0.09}_{-0.09}$	$0.00^{+0.02}_{-0.01}$	$1.186^{+0.001}_{-0.001}$	$\leq$ 0.89	$\leq$ 2.8	40^{+10}_{-10}	1187008882.4	$\geq$ 0.04	$\geq$ 0.1	NA	$<$ 1.00e-07	$<$ 1.25e-05	$0.01^{+0.00}_{-0.00}$	16	NA	33.0	30.9	12:41:04.4
GW170818	$35.5^{+7.5}_{-4.7}$	$26.8^{+4.3}_{-5.2}$	$-0.09^{+0.18}_{-0.21}$	$26.7^{+2.1}_{-1.7}$	$0.67^{+0.07}_{-0.08}$	$59.8^{+4.8}_{-3.8}$	1020^{+430}_{-360}	1187058327.1	$2.7^{+0.5}_{-0.5}$	$3.4^{+0.5}_{-0.7}$	NA	4.20e-05	NA	$0.20^{+0.07}_{-0.07}$	39	NA	11.3	NA	02:25:09.1
GW170823	$39.6^{+10.0}_{-6.6}$	$29.4^{+6.3}_{-7.1}$	$0.08^{+0.20}_{-0.22}$	$29.3^{+4.2}_{-3.2}$	$0.71^{+0.08}_{-0.10}$	$65.6^{+9.4}_{-6.6}$	1850^{+840}_{-840}	1187529256.5	$3.3^{+0.9}_{-0.8}$	$3.6^{+0.6}_{-0.9}$	2.14e-03	$<$ 1.00e-07	$<$ 3.29e-05	$0.34^{+0.13}_{-0.14}$	1651	10.8	11.5	11.1	13:13:58.5

Strain data – Single events

GWTC-1-confident

GW170814

Effective inspiral spin,	**$0.07^{+0.12}_{-0.11}$**
Luminosity distance, *Mpc*	**580^{+160}_{-210}**
Final spin,	**$0.72^{+0.07}_{-0.05}$**
Primary mass, *M_sun*	**$30.7^{+5.7}_{-3.0}$**
FAR gstLAL, *yr^-1*	**< 1.00e-07**
FAR PyCBC, *yr^-1*	**< 1.25e-05**
Secondary mass, *M_sun*	**$25.3^{+2.9}_{-4.1}$**
chirp mass, *M_sun*	**$24.2^{+1.4}_{-1.1}$**
Radiated energy, *M_sun X c^2*	**$2.7^{+0.4}_{-0.3}$**
Network SNR gstLAL,	**15.9**
Source redshift,	**$0.12^{+0.03}_{-0.04}$**
FAR cWB, *yr^-1*	**< 2.08e-04**
UTCtime,	**10:30:43.5**
Peak luminosity, *10^56 erg s^-1*	**$3.7^{+0.4}_{-0.5}$**
Sky localization, *deg^2*	**87**
Final mass, *M_sun*	**$53.4^{+3.2}_{-2.4}$**
GPS time (s),	**1186741861.5**
Network SNR PyCBC,	**16.3**
Network SNR cWB,	**17.2**

Download Data Files:

V1 4096sec 4KHz	V-V1_GWOSC_4KHZ_R1-1186739814-4096.hdf5 V-V1_GWOSC_4KHZ_R1-1186739814-4096.gwf V-V1_GWOSC_4KHZ_R1-1186739814-4096.txt.gz
V1 4096sec 16KHz	V-V1_GWOSC_16KHZ_R1-1186739814-4096.hdf5 V-V1_GWOSC_16KHZ_R1-1186739814-4096.gwf V-V1_GWOSC_16KHZ_R1-1186739814-4096.txt.gz
V1 32sec 4KHz	V-V1_GWOSC_4KHZ_R1-1186741846-32.hdf5 V-V1_GWOSC_4KHZ_R1-1186741846-32.gwf V-V1_GWOSC_4KHZ_R1-1186741846-32.txt.gz
V1 32sec 16KHz	V-V1_GWOSC_16KHZ_R1-1186741846-32.hdf5 V-V1_GWOSC_16KHZ_R1-1186741846-32.gwf V-V1_GWOSC_16KHZ_R1-1186741846-32.txt.gz
H1 4096sec 4KHz	H-H1_GWOSC_4KHZ_R1-1186739814-4096.hdf5 H-H1_GWOSC_4KHZ_R1-1186739814-4096.gwf H-H1_GWOSC_4KHZ_R1-1186739814-4096.txt.gz
H1 4096sec 16KHz	H-H1_GWOSC_16KHZ_R1-1186739814-4096.hdf5 H-H1_GWOSC_16KHZ_R1-1186739814-4096.gwf H-H1_GWOSC_16KHZ_R1-1186739814-4096.txt.gz
H1 32sec 4KHz	H-H1_GWOSC_4KHZ_R1-1186741846-32.hdf5 H-H1_GWOSC_4KHZ_R1-1186741846-32.gwf H-H1_GWOSC_4KHZ_R1-1186741846-32.txt.gz
H1 32sec 16KHz	H-H1_GWOSC_16KHZ_R1-1186741846-32.hdf5 H-H1_GWOSC_16KHZ_R1-1186741846-32.gwf H-H1_GWOSC_16KHZ_R1-1186741846-32.txt.gz

each file:

- Metadata
- GW strain
- Data quality (1 Hz rate)

Bulk data

Available in 2 ways

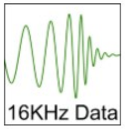
Large Data Sets for High Performance Computing

For users of computing clusters, CernVM-FS is the preferred method to access large data sets:

CernVM FS
i.e. mount a network disk on your PC

O2 Data Release

O2 Time Range: November 30, 2016 through August 25, 2017
Detectors: H1, L1 and V1



O1 Data Release

O1 Time Range: September 12, 2015 through January 19, 2016
Detectors: H1 and L1

Search archive

Detector status

GWOSC Calendar Today Yesterday Observing Run 1 Summary Observing Run 2 Summary

Gravitational-Wave Observatory Status

Please select a day from the calendar above to see archived or current status.

Information is available for dates after November 30, 2016. The Advanced LIGO and Advanced Virgo detectors are currently in the third observing run, known as O3, which began April 1, 2019. Summaries of previous observing runs are available in the menu above. For overviews of LIGO and Virgo observing runs, see the O3 schedule or arXiv:1304.0670.

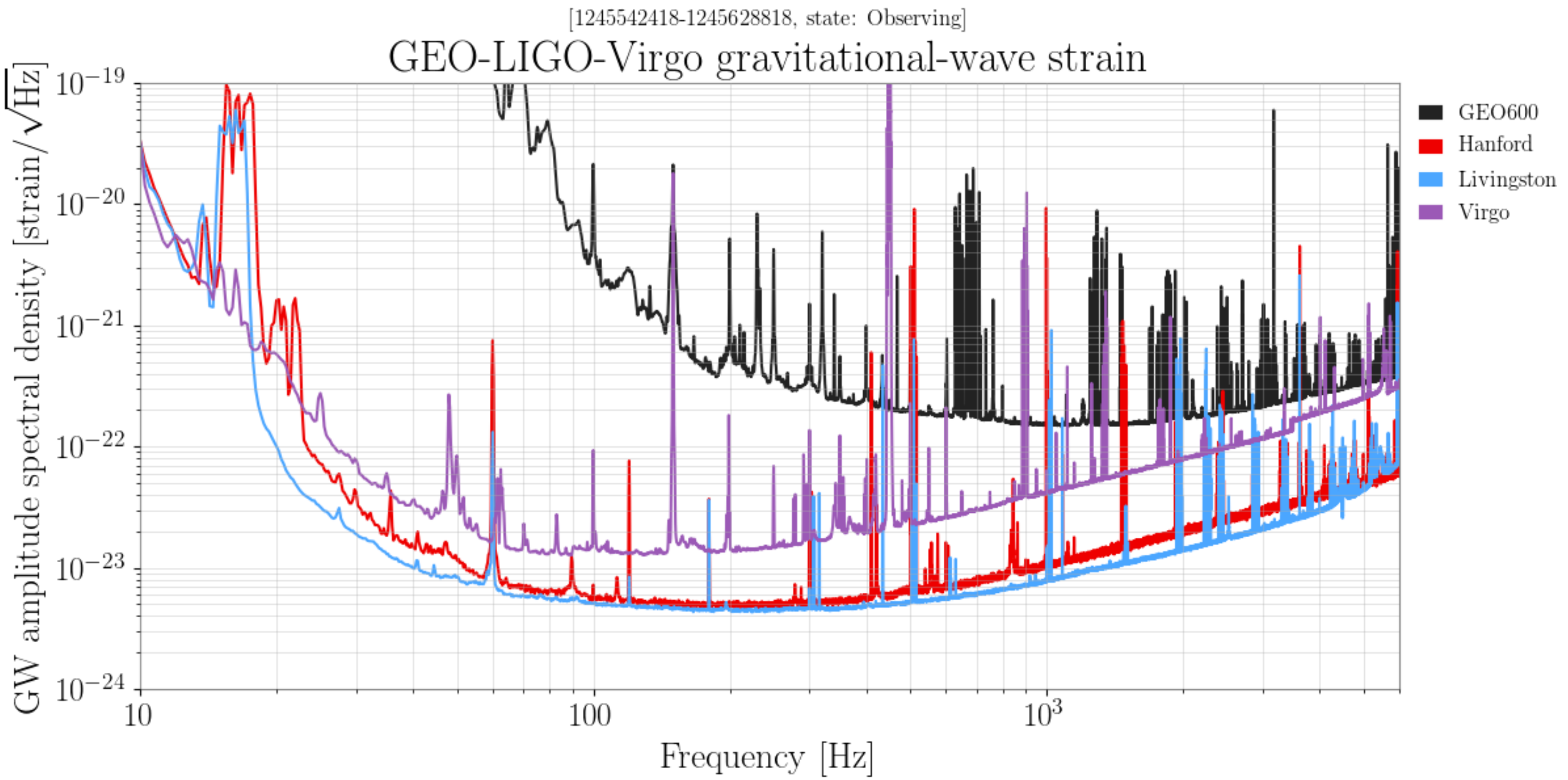
- Today's Summary Page
- Virgo Status Page
- Current Status (GWISTAT)
- LIGO/Virgo Alerts (GraceDB)

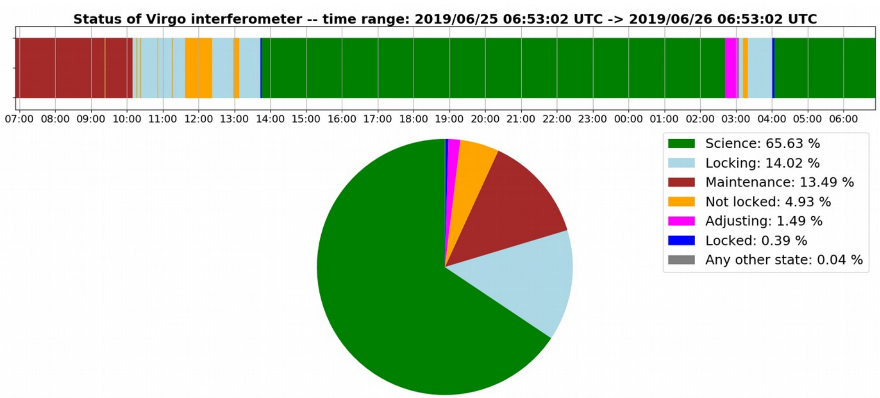
LIGO Hanford

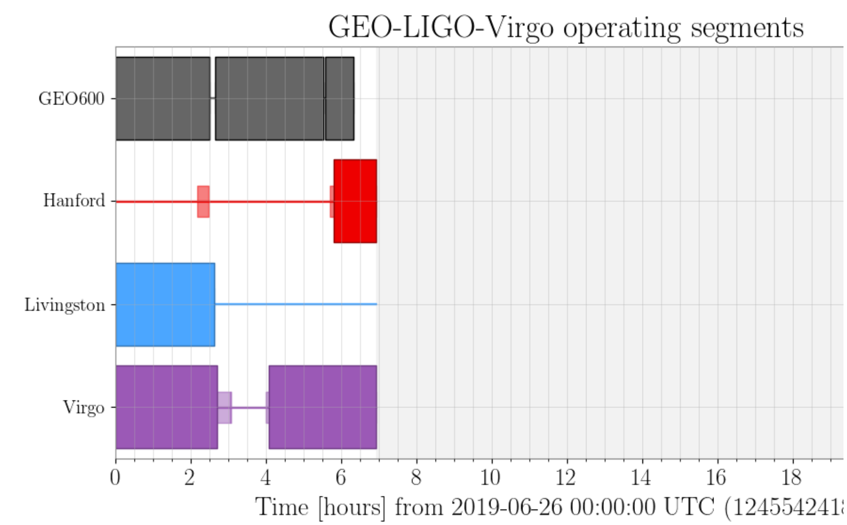
LIGO Livingston

Virgo

GEO600


[1245542418-1245628818, state: Observing]
GEO-LIGO-Virgo gravitational-wave strain
GW amplitude spectral density [strain/√Hz]
10^−19
10^−20
10^−21
10^−22
10^−23
10^−24
10
100
10^3
Frequency [Hz]
GEO600
Hanford
Livingston
Virgo


Status of Virgo interferometer -- time range: 2019/06/25 06:53:02 UTC -> 2019/06/26 06:53:02 UTC
07:00 08:00 09:00 10:00 11:00 12:00 13:00 14:00 15:00 16:00 17:00 18:00 19:00 20:00 21:00 22:00 23:00 00:00 01:00 02:00 03:00 04:00 05:00 06:00
Science: 65.63 %
Locking: 14.02 %
Maintenance: 13.49 %
Not locked: 4.93 %
Adjusting: 1.49 %
Locked: 0.39 %
Any other state: 0.04 %


GEO-LIGO-Virgo operating segments
GEO600
Hanford
Livingston
Virgo
0 2 4 6 8 10 12 14 16 18
Time [hours] from 2019-06-26 00:00:00 UTC (1245542418

LIGO-Virgo alerts

Gravitational Wave Open Science Center

Data▾ Software▾ Online Status▾ About GWOSC▾

LIGO/Virgo Public Alerts

Announcement

From GCN Circular 24045:

Our third observing run ("O3") began as scheduled on 2019 April 1 at 15:00 UTC. At that time the LIGO Hanford, LIGO Livingston, and Virgo Observatories transitioned from engineering and commissioning to observing. All three detectors are operating at good sensitivity and stability. We are analyzing data in low latency and processing candidate transient events automatically.

As of April 2 20:00 UTC, we have configured our low-latency analysis pipeline to send public alerts for significant gravitational-wave transient candidates that are detected in coincidence across two or more gravitational-wave detectors.

Automated Preliminary GCN Notices will be sent immediately without any human intervention. Shortly afterward, they will be vetted by an LSC/Virgo rapid response team and either confirmed with an Initial GCN Notice and Circular, or withdrawn with a Retraction.

Retraction notices may be issued more frequently over the next few weeks as our understanding of the instrumental background improves.

For further information about vetting procedures, analysis methodology, and the contents of LIGO/Virgo public alerts, refer to the LIGO/Virgo Public Alerts User Guide: https://emfollow.docs.ligo.org/userguide/

This marks the beginning of the era of public alerts for the field of gravitational-wave astronomy.

Resources

- LIGO/Virgo Alerts User Guide
- Gravitational Wave Candidate Event Database (GraceDB)
- GCN: The Gamma-ray Coordinates Network
- GW Events iPhone app
- Press release on start of O3

LIGO-Virgo alerts

GraceDB — Gravitational Wave Candidate Event Database

HOME | SEARCH | LATEST | DOCUMENTATION | LOGIN

Latest — as of 26 June 2019 07:01:58 UTC

Test and MDC events and superevents are not included in the search results by default; see the query help for information on how to search for events and superevents in those categories.

Query:

Search for: Superevent

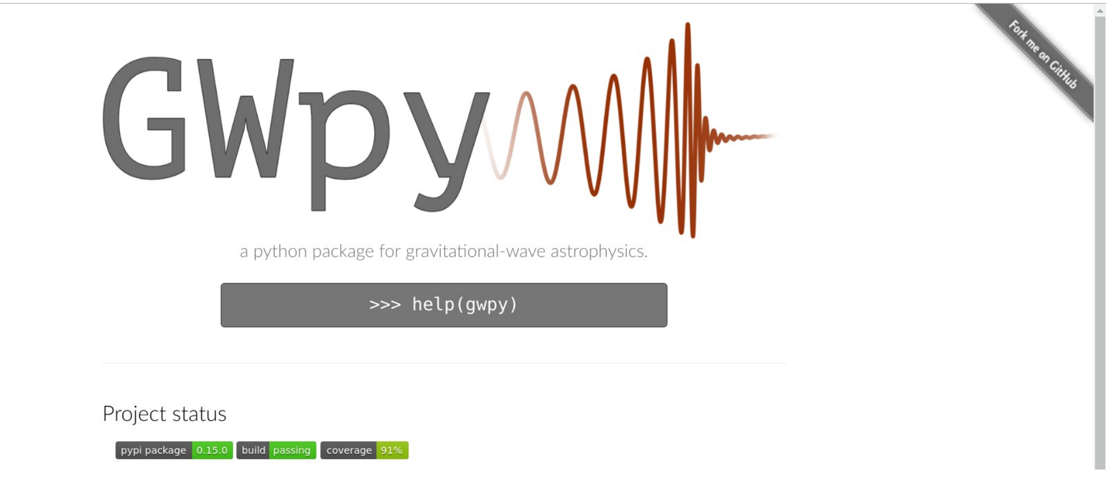
Search

UID	Labels	t_start	t_0	t_end	FAR (Hz)	UTC Created
S190602aq	PE_READY ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1243533584.081266	1243533585.089355	1243533586.346191	1.901e-09	2019-06-02 17:59:51 UTC
S190524q	ADVNO SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1242708743.678669	1242708744.678669	1242708746.133301	6.971e-09	2019-05-24 04:52:30 UTC
S190521r	PE_READY ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1242459856.453418	1242459857.460739	1242459858.642090	3.168e-10	2019-05-21 07:44:22 UTC
S190521g	PE_READY ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1242442966.447266	1242442967.606934	1242442968.888184	3.801e-09	2019-05-21 03:02:49 UTC
S190519bj	PE_READY ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1242315361.378873	1242315362.655762	1242315363.676270	5.702e-09	2019-05-19 15:36:04 UTC
S190518bb	ADVNO SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1242242376.474609	1242242377.474609	1242242380.922655	1.004e-08	2019-05-18 19:19:39 UTC
S190517h	PE_READY ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1242107478.819517	1242107479.994141	1242107480.994141	2.373e-09	2019-05-17 05:51:23 UTC
S190513bm	ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1241816085.736106	1241816086.869141	1241816087.869141	3.734e-13	2019-05-13 20:54:48 UTC
S190512at	PE_READY ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1241719651.411441	1241719652.416286	1241719653.518066	1.901e-09	2019-05-12 18:07:42 UTC
S190510g	ADVOK SKYMAP_READY EMBRIGHT_READY PASTRO_READY DQOK GCN_PRELIM_SENT	1241492396.291636	1241492397.291636	1241492398.293185	8.834e-	2019-05-10 03:00:03

Www.gracedb.org

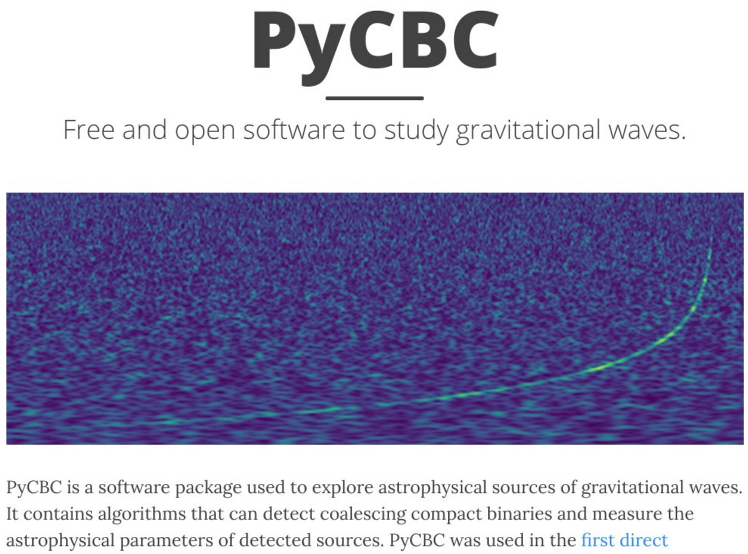
How to run analysis on your own?

Standard routines have been published in Python-based software packges, some of them publicly available



www.gwpy.org

www.pycbc.org



*

The gwpy Package

Strain data h(t)
4KHz

Read the strain data and segments

```
from gwpy.timeseries import TimeSeries
from gwpy.segments import DataQualityFlag

# Define the interval
dt_win=256
ev_t0_min = ev_gps-dt_win
ev_t0_max = ev_gps+dt_win

print("Get data for %s (%s) GPS: %.2f - %.2f" % (ev_name,ev_ifo,ev_t0_min,ev_t0_max))

#fetch the data. Use cache=True to keep the data in the cache memory (to speed things up)
data = TimeSeries.fetch_open_data(ev_ifo,ev_t0_min,ev_t0_max,cache=True)

#get the segments in a larger time window (just to have a bigger time span to look over)
segments = DataQualityFlag.fetch_open_data(ev_ifo+"_DATA",ev_t0_min-3600,ev_t0_max+3600)
print("Done")
```

```
Get data for GW150914 (H1) GPS: 1126259206.40 - 1126259718.40
```

```
/opt/anaconda/envs/gw-env-py27/lib/python2.7/site-packages/gwpy/types/series.py:921: UserWarning: TimeSeries.crop g
iven end larger than current end, crop will end when the Series actually ends.
  % type(self).__name__)
```

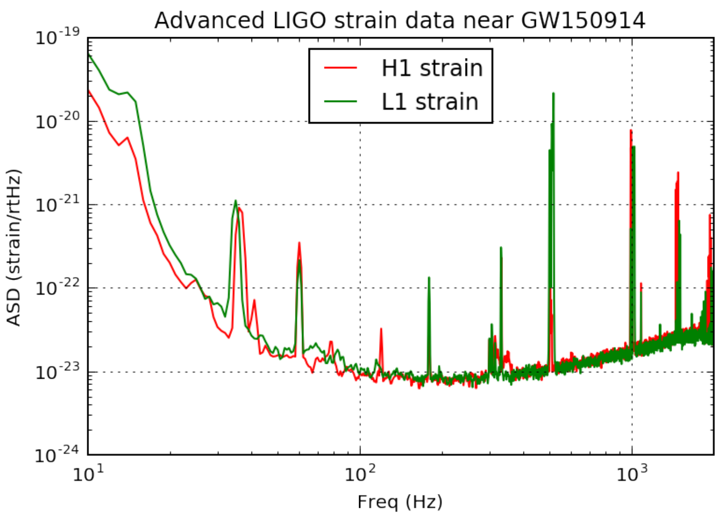
```
Done
```

Data Quality Segments

1 Hz

*

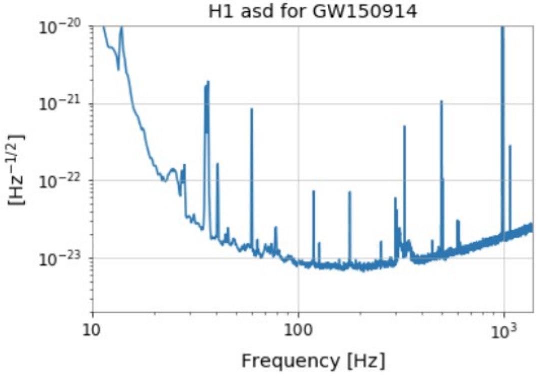
Building the ASD



In [11]:
```
#Now we can plot the ASD as well, using the Welch method with a overlapping window of 4 seconds

asd = data.asd(fftlength=4, method="median")
plot = asd.plot()

ax = plot.gca()
ax.set_xlim(10, 1400)
ax.set_ylim(2e-24, 1e-20)
ax.set_title(ev_ifo+ " asd for "+ev_name)
plot.refresh()
```



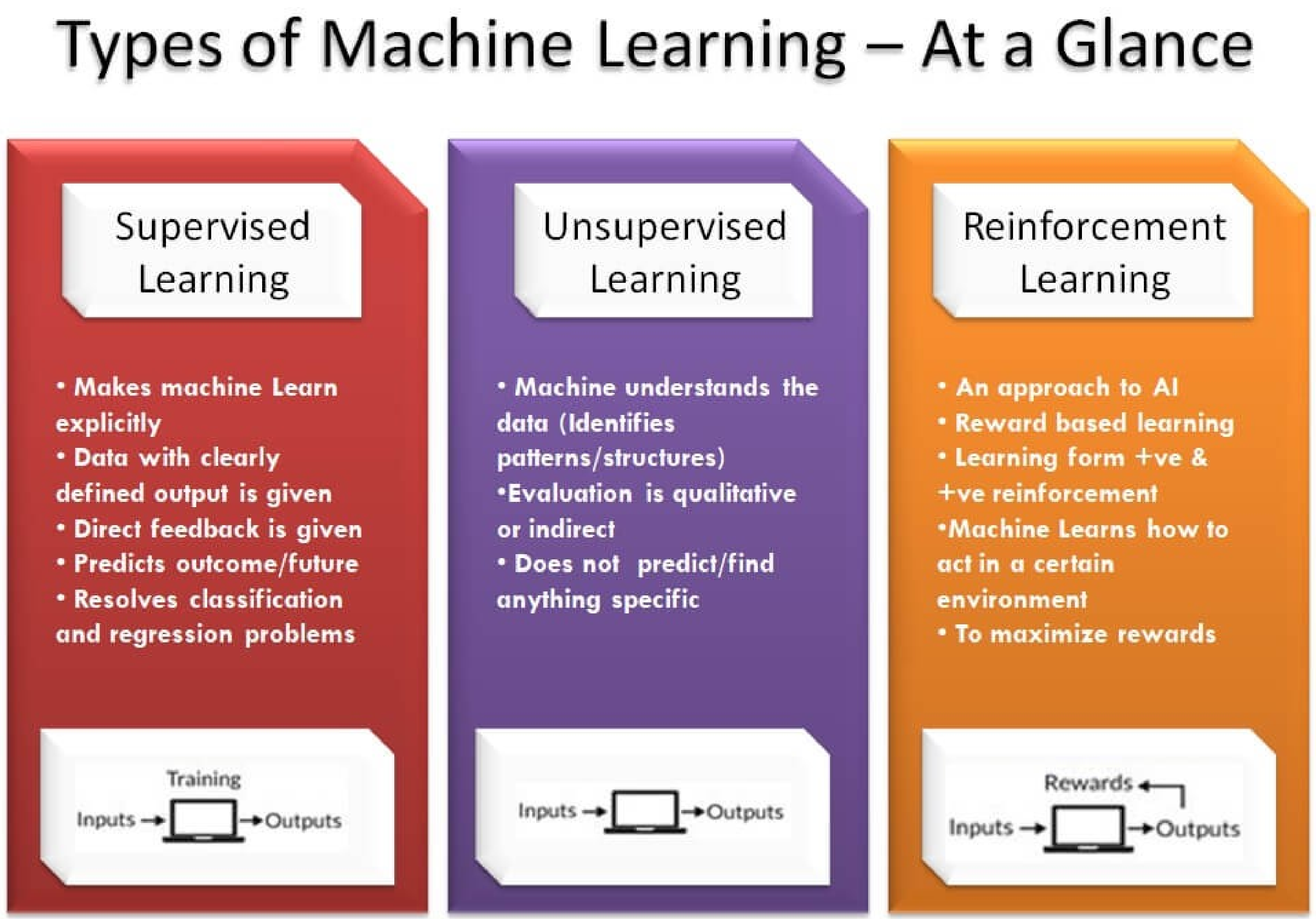
Gravitational Waves and Big Data

- Interferometers are producing lots of data everyday
- Order of ~50 MB/s → about 0.5 TB/day from ~10^3-10^5 channels
- Signals are buried in a high noise

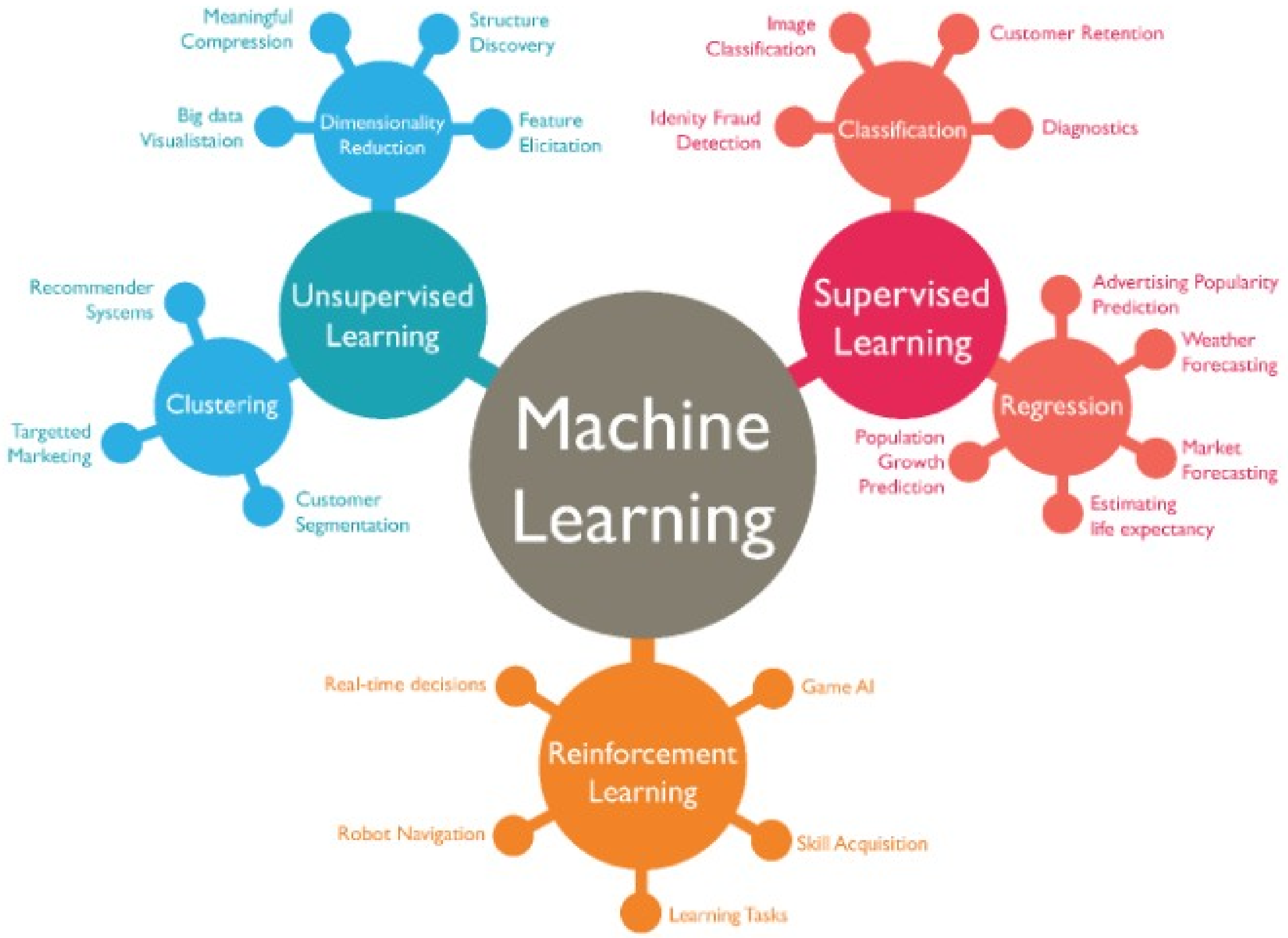
- Big data stimulates use of machine learning methods for 2 main reasons:
  - Shorter timescales
    - Detector characterization
    - Detection and quick localization
    - Low-latency analysis for quick EM alert
  - Longer timescales
    - Search for new sources (not just CBC but also CW etc)

Intro to machine learning

- Developed to combine pattern recognition + artificial intelligence
- Recent Reinassance due to more computing power



Intro to machine learning

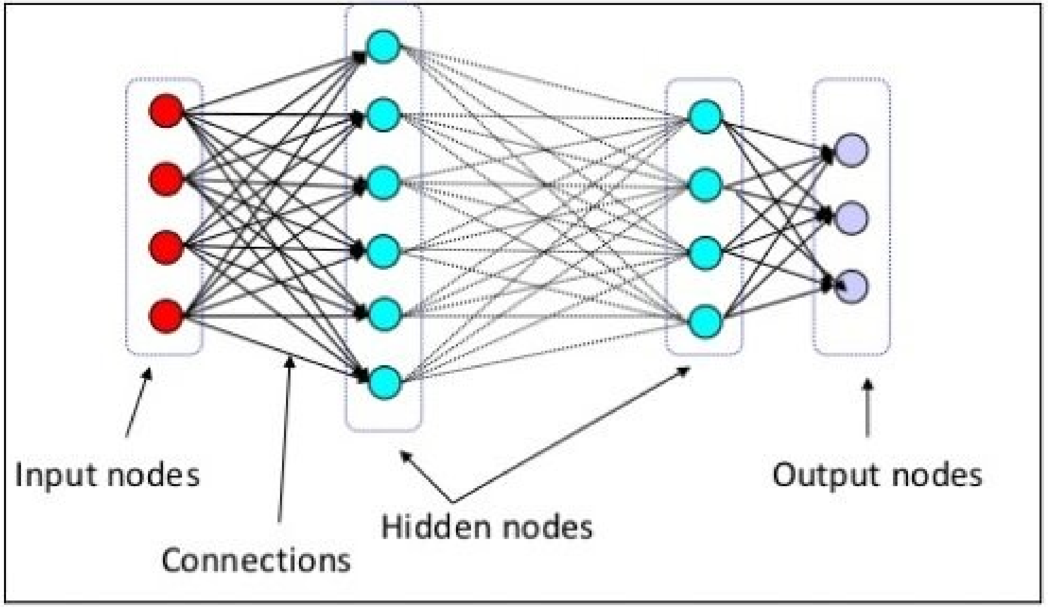


Why Deep Learning?

- Deep Learning (DL) is at the frontier of machine learning studies
  - Born from works on neural networks and artificial intelligence
- It combines neural network architecture & power of machine learning
- The building block is the artificial neuron (perceptron), acting as a nonlinear processing unit
- A single perceptron → multilayer network of perceptrons (i.e. “deep”)

- In principle, a deep network can approximate any continuous function (universal approximation)

- Various projects in progress in LIGO/Virgo to apply machine learning to grawitational waves. Many at this workshop!



Deep learning for Gravitational waves

- Looks like a very promising approach due to data space complexity
- Can be applied to various aspects, from detection of new sources to characterization of the detector

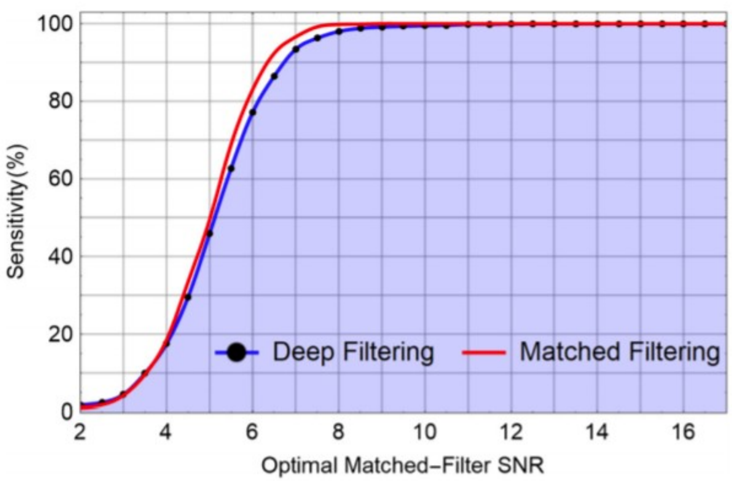


Fig. 4. Sensitivity of detection with real LIGO noise. The figure shows the sensitivity of detecting GW signals injected in real LIGO noise (from LOSC) from our test set using `Deep Filtering` compared with matched filtering with the same template bank used for training the CNN. Note that the SNR is on average proportional to 10 ± 1.5 times the ratio of the amplitude of the signal to the standard deviation of the noise for our test set. This implies that we are capable of detecting signals significantly weaker than the background noise. While matched-filtering has the advantage of being optimized with the PSD of the LIGO noise in the test set, `Deep Filtering` was only trained on noise from other events, therefore our results demonstrate the ability of the CNNs to automatically generalize to non-stationary LIGO noise having different PSDs without retraining.

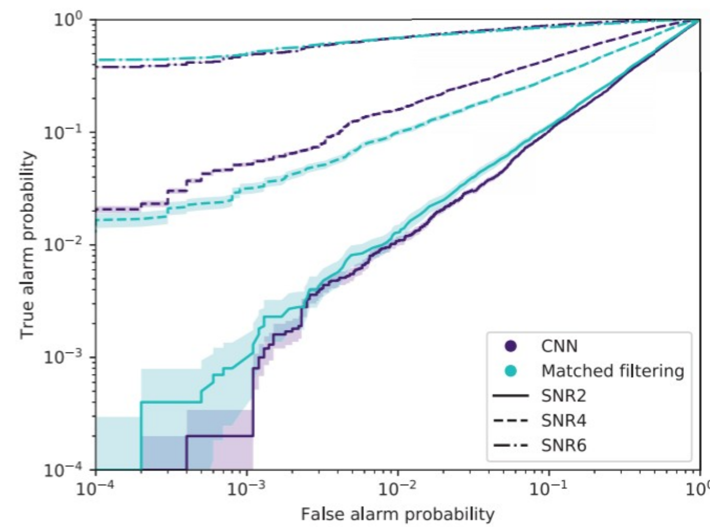


FIG. 2. The ROC curves for test datasets containing signals with optimal SNR, $\rho_{\rm opt} = 2, 4, 6$. We plot the true alarm probability versus the false alarm probability estimated from the output of the CNN (purple) and matched-filtering (cyan) approaches. Uncertainties in the true alarm probability correspond to 1-σ bounds assuming a binomial distribution.

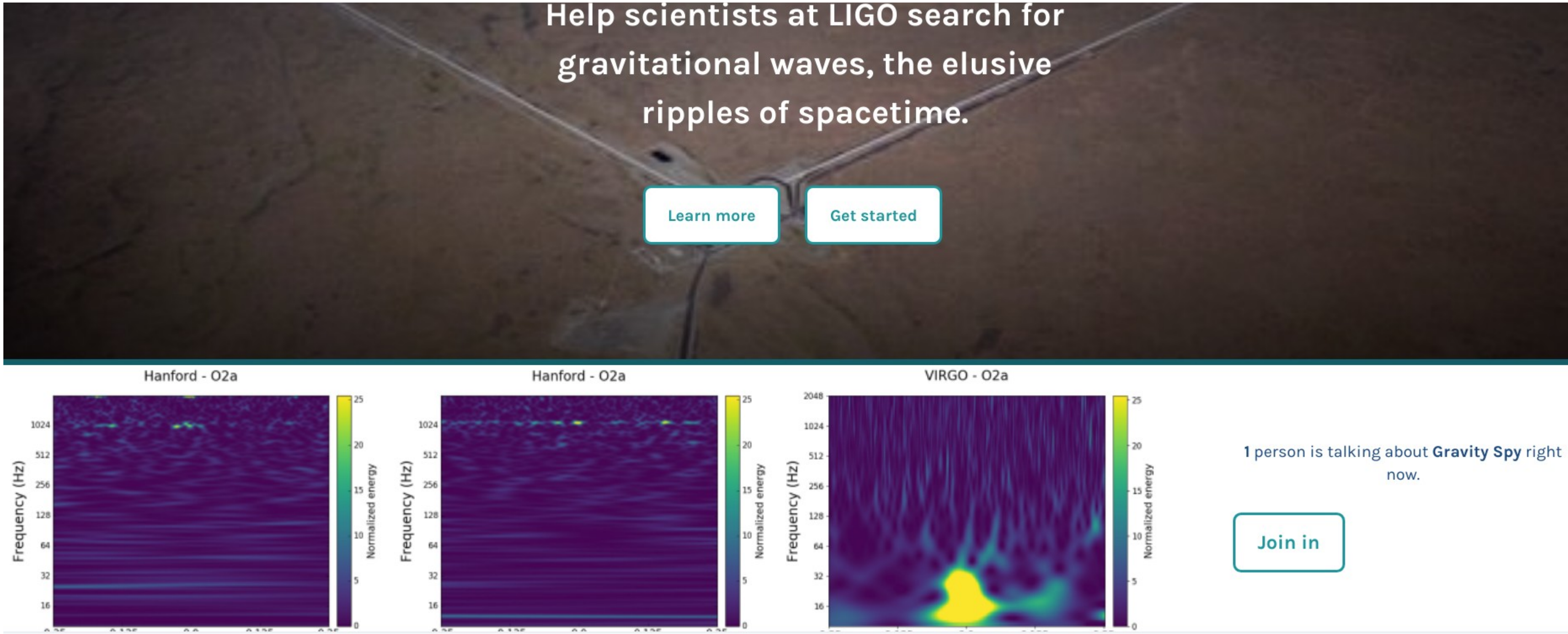
George & Huerta, 2018

Gabbard et al 2019

Deep learning for glitch characterization & classification

- Interferometers are limited by stationary and nonstationary noise
- Transient noise events (glitches) can impact data quality and mimic real astrophysical signals
- Detect and classify glitches is one of the most important tasks for detector characterization and data analysis (e.g. low-latency & detector optimization)
- Glitches can have complex time-frequency signatures → difficult to classify manually
- Automatic methods have been tested (e.g. Powell+15, CQG,32,215021, Mukund+17,PRD,95,104059)
- Many groups working on this in the LVC, already various publications!

Glitch & Citizen science: GravitySpy

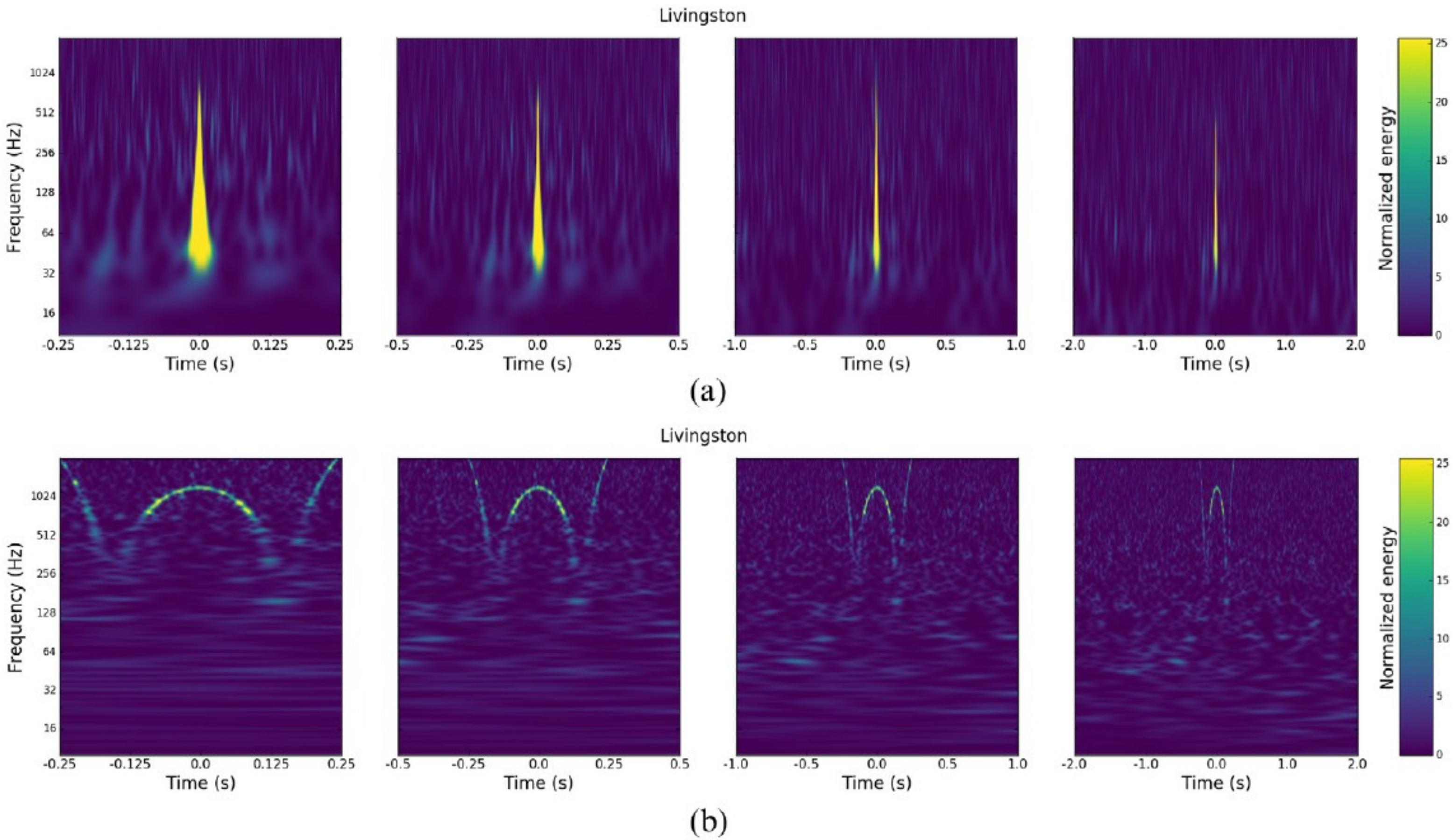


[Www.gravityspy.org](http://www.gravityspy.org)

Needs for a training set
Citizen scientists contribute to classify glitches

More details in Zevin+17

Sample glitch gallery

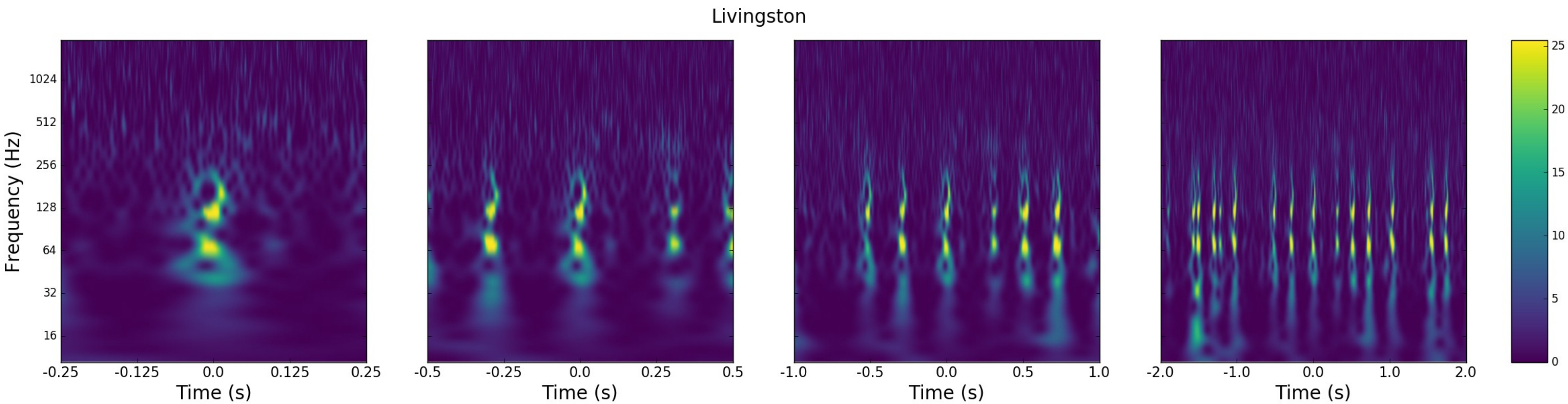


Blip glitches

Whistle glitches

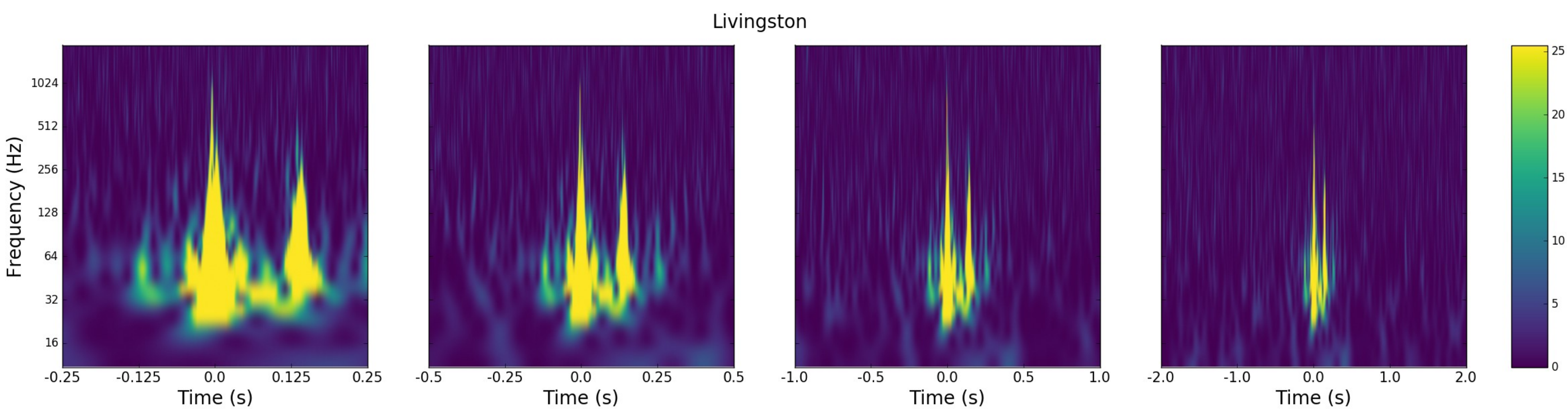
Examples of time-frequency glitch morphology (Zevin+17)

Sample glitch gallery



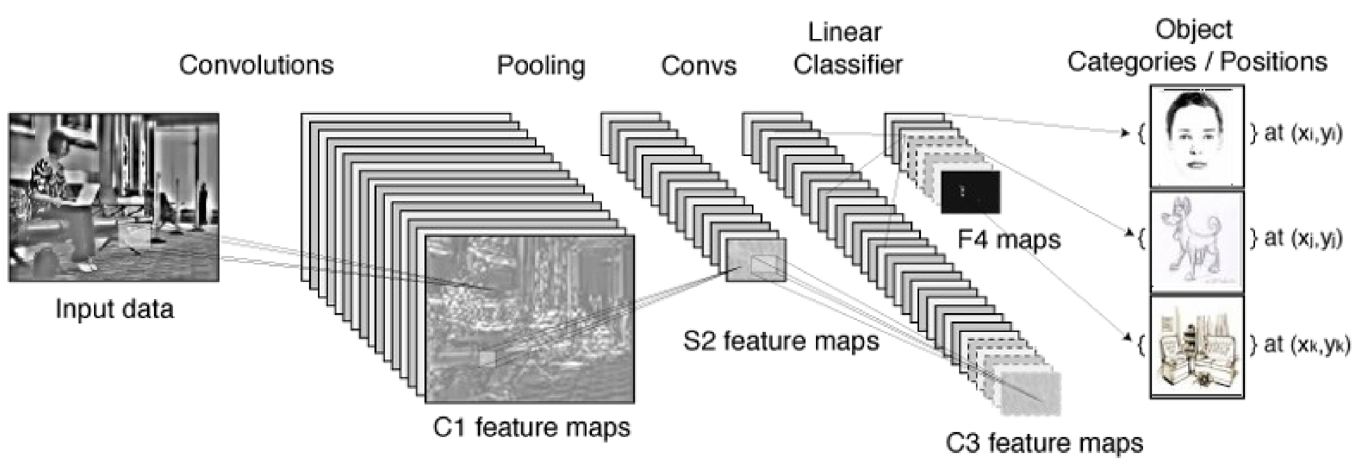
Helix glitches

Koi fish glitches



Deep Learning & glitches

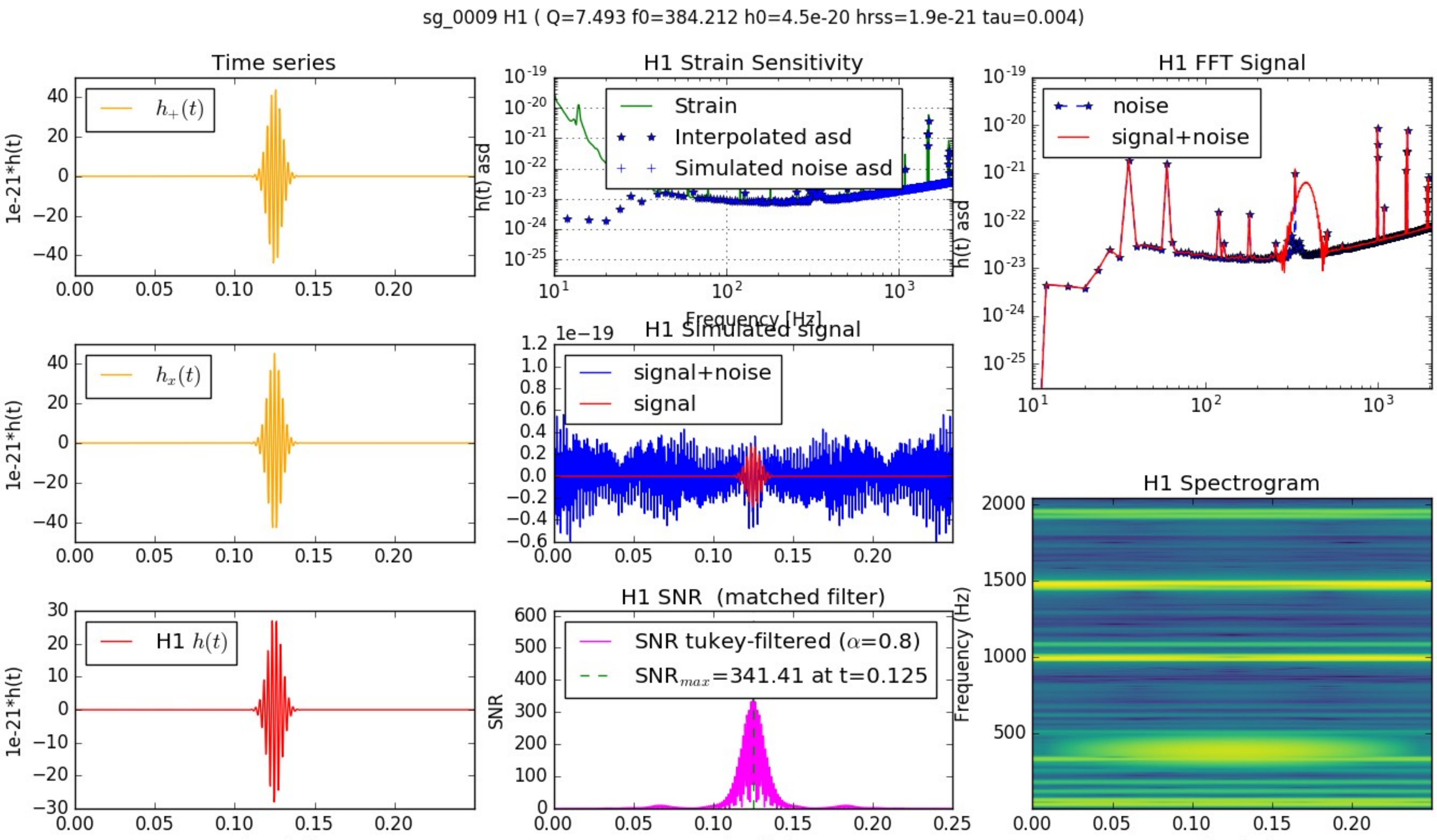
- Promising tool to classify complex patterns
- Deep network to approximate a classification function
- In our case, the function is:
  - GW data → glitch class
- We focus on images
  - Easy to spot signal “types” (training)
  - Compress long data stream (time-frequency)
  - Preprocessing stage → Image recognition techniques
- Fully connected deep neural networks are CPU expensive
- We exploit Convolutional deep Neural Networks (CNNs)
  - Designed to extract features
  - Optimized for image classification



Work in collaboration with E. Cuoco

Tests on simulations (I)

- Ad hoc simulations for tests (e.g. Powell+2015)
- Simulate colored noise using public sensitivity curve
- 6 classes of glitch shapes (+ NOISE one to check detection)



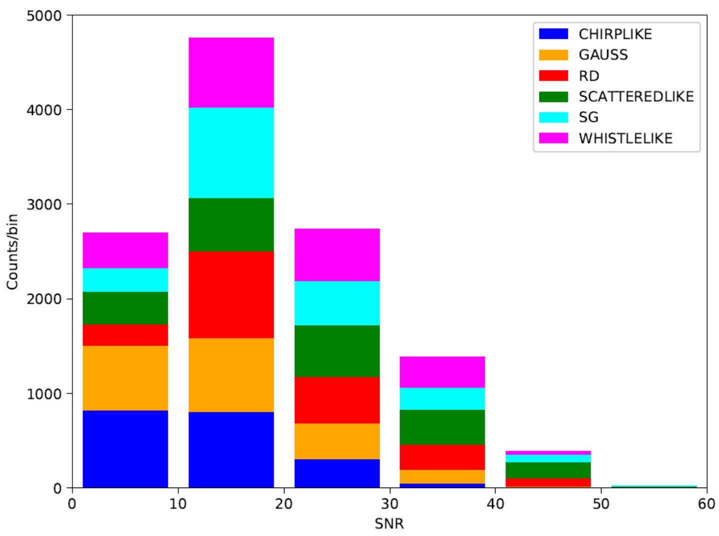
Example of H1 simulation

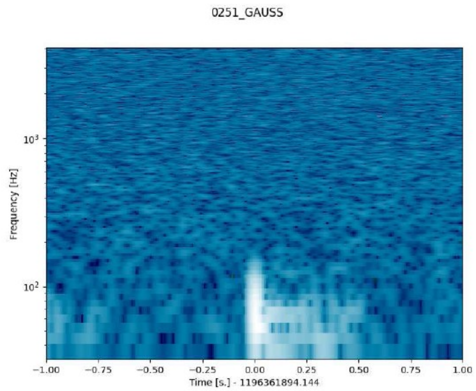
Building the images

- 2k glitches per family
- Class for signals
- Spectrogram for each image
- 2-seconds time window

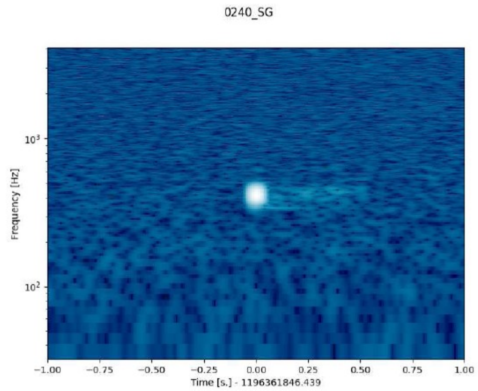
to highlight fatures in long glitches
(if needed, can use multi-window)

- Data is whitened
- Optional contrast stretch

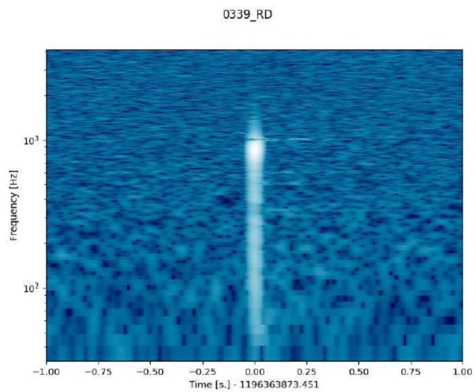




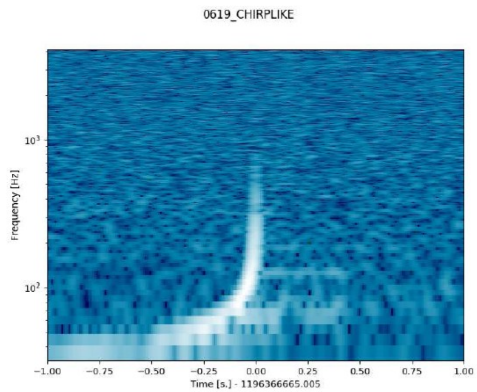
(a)



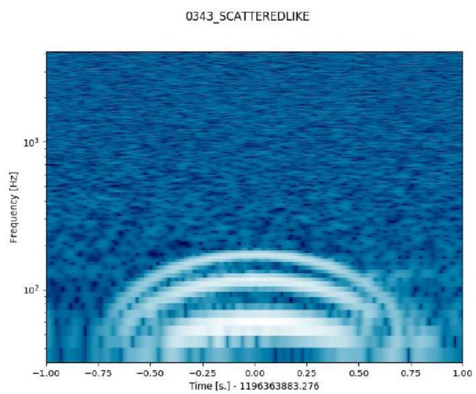
(b)



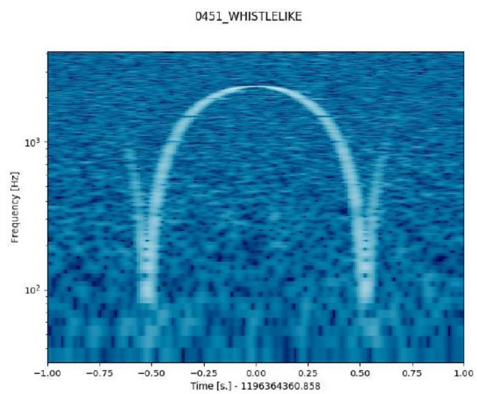
(c)



(d)



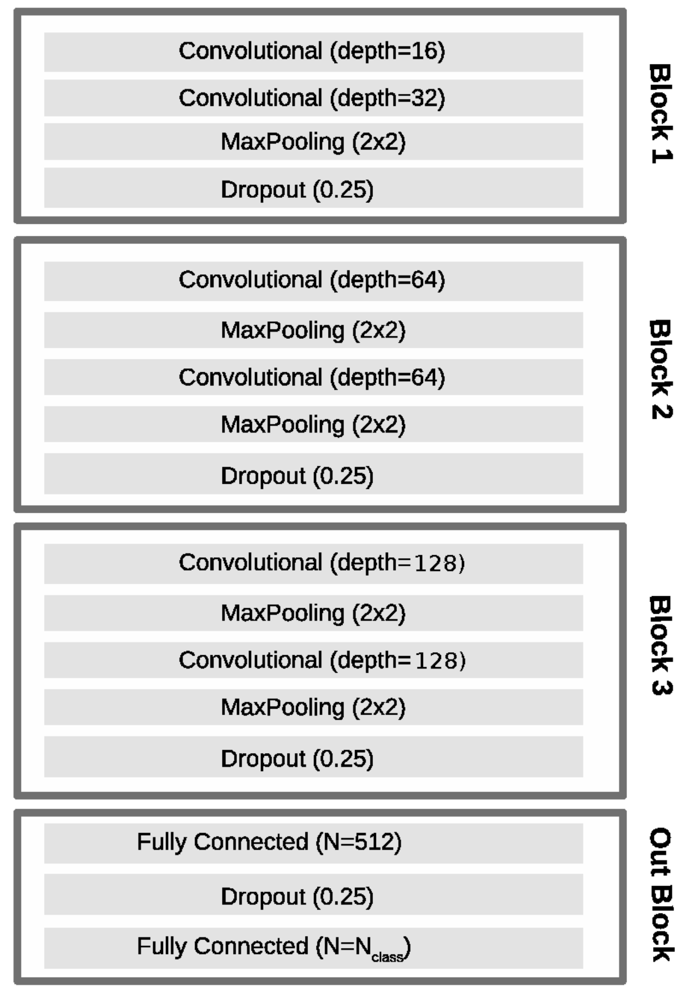
(e)



(f)

Developing the CNN architecture

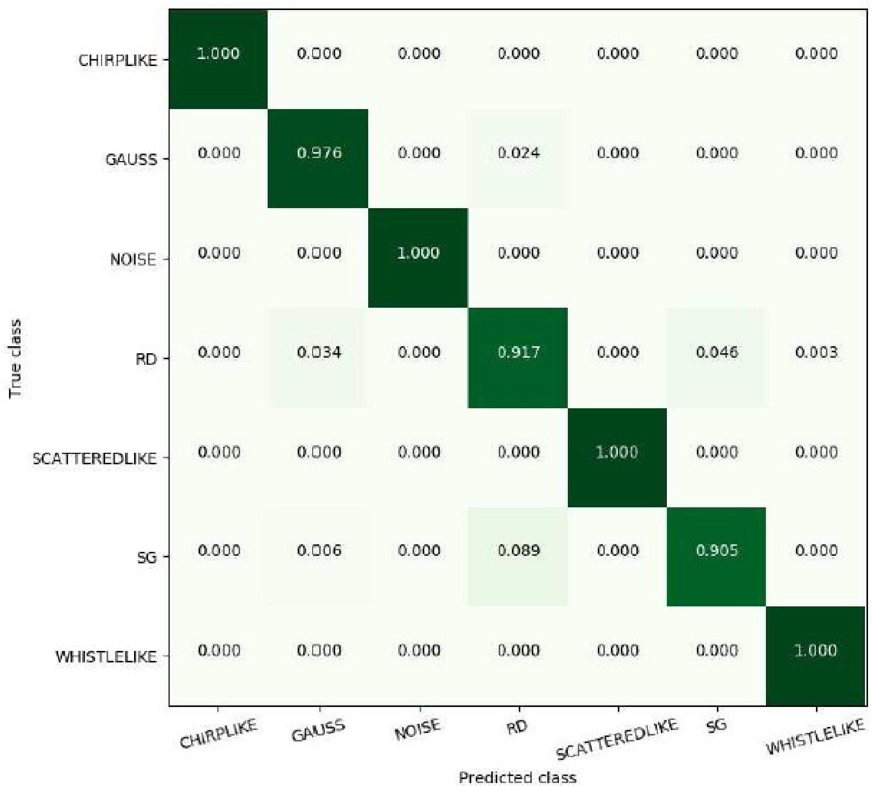
- Input GW data
  - Timeseries whitening (See Elena Cuoco’s talk)
  - Image processing
  - Image creation from time series (FFT or Q transform)
  - (if needed) Image equalization & contrast enhancement
- Classification
  - A probability for each class, take the max
  - Add a NOISE (i.e. no glitch) class to check detectio
- Network layout
  - Tested various networks:
    - Shallow
    - 1 CNN block
    - 3 CNN blocks
- Run on GPU Nvidia GeForce (GTX 780/Titan Xp)
  - Python + CUDA-optimized libraries
    - Keras+Tensorflow+etc



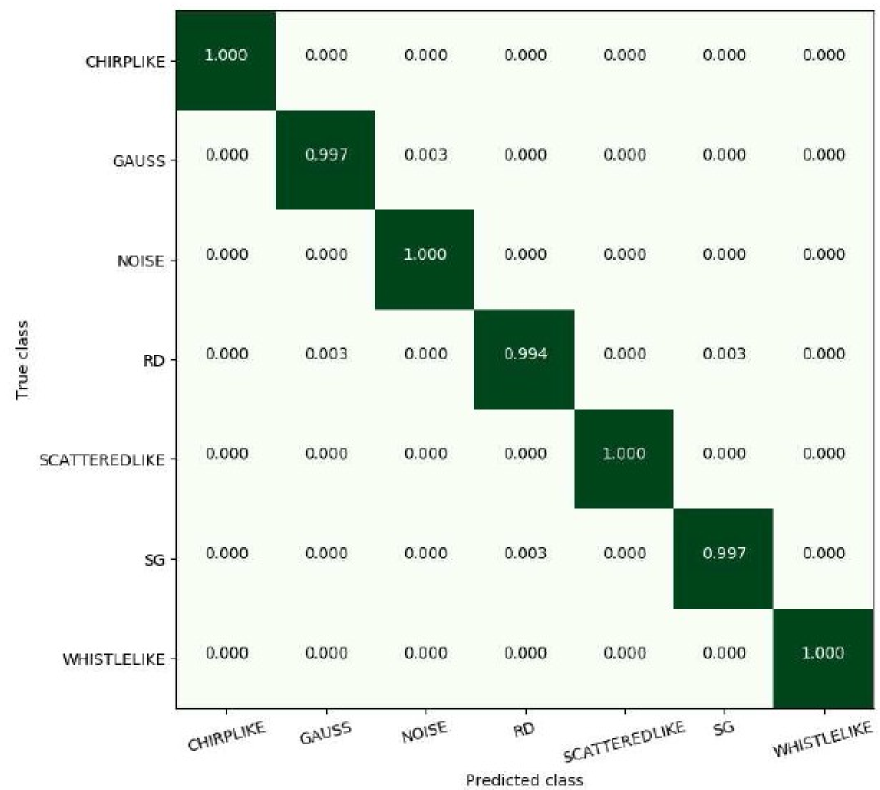
Confusion matrices

Normalized Confusion Matrix

SVM



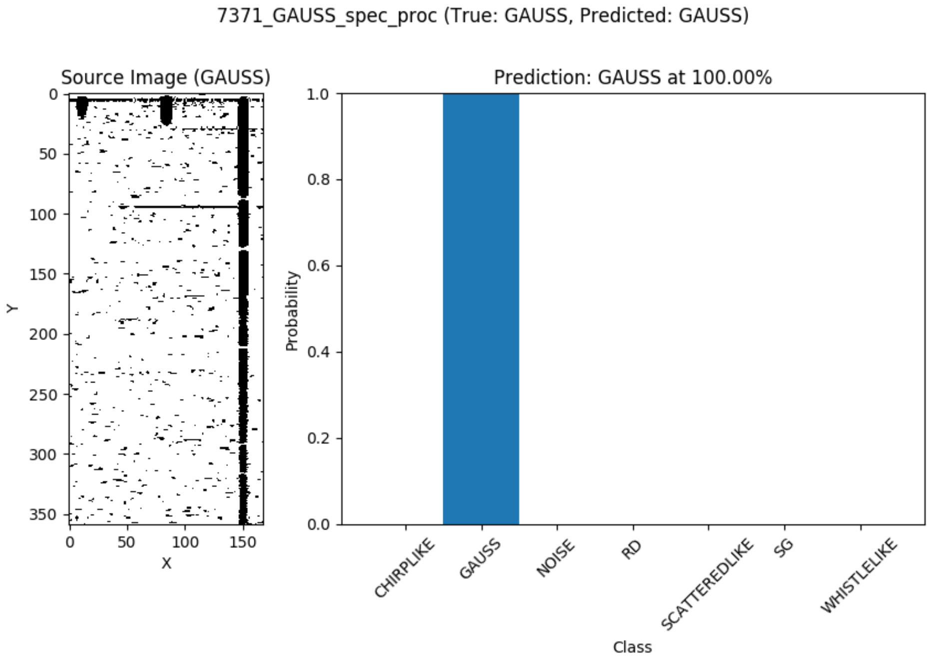
Deep CNN



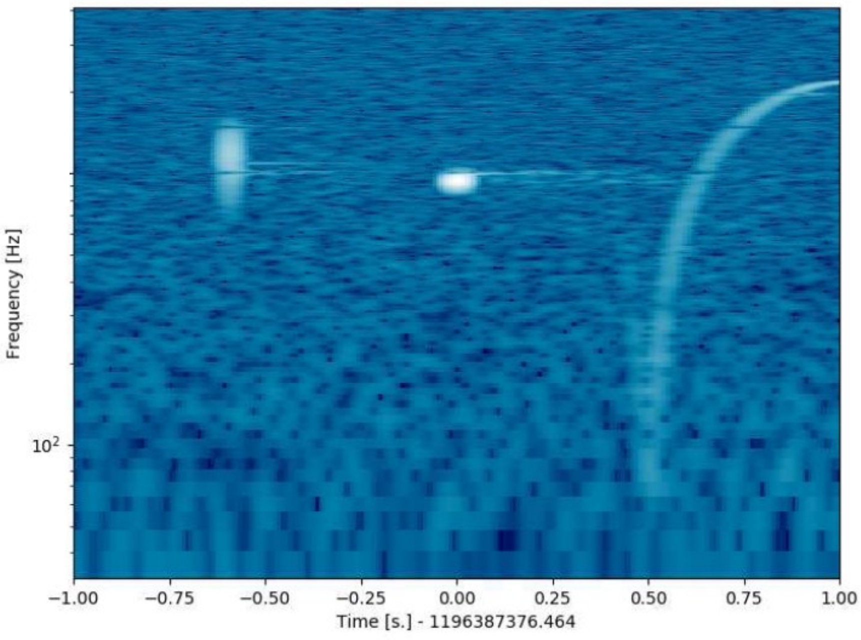
Deep CNN better at distinguishing similar morphologies

Picking & classifying glitches

Some cases of more glitches in the time window, always identify the right class



100% Sin-Gauss



More details in
Razzano & Cuoco 2018, CQG,35,9

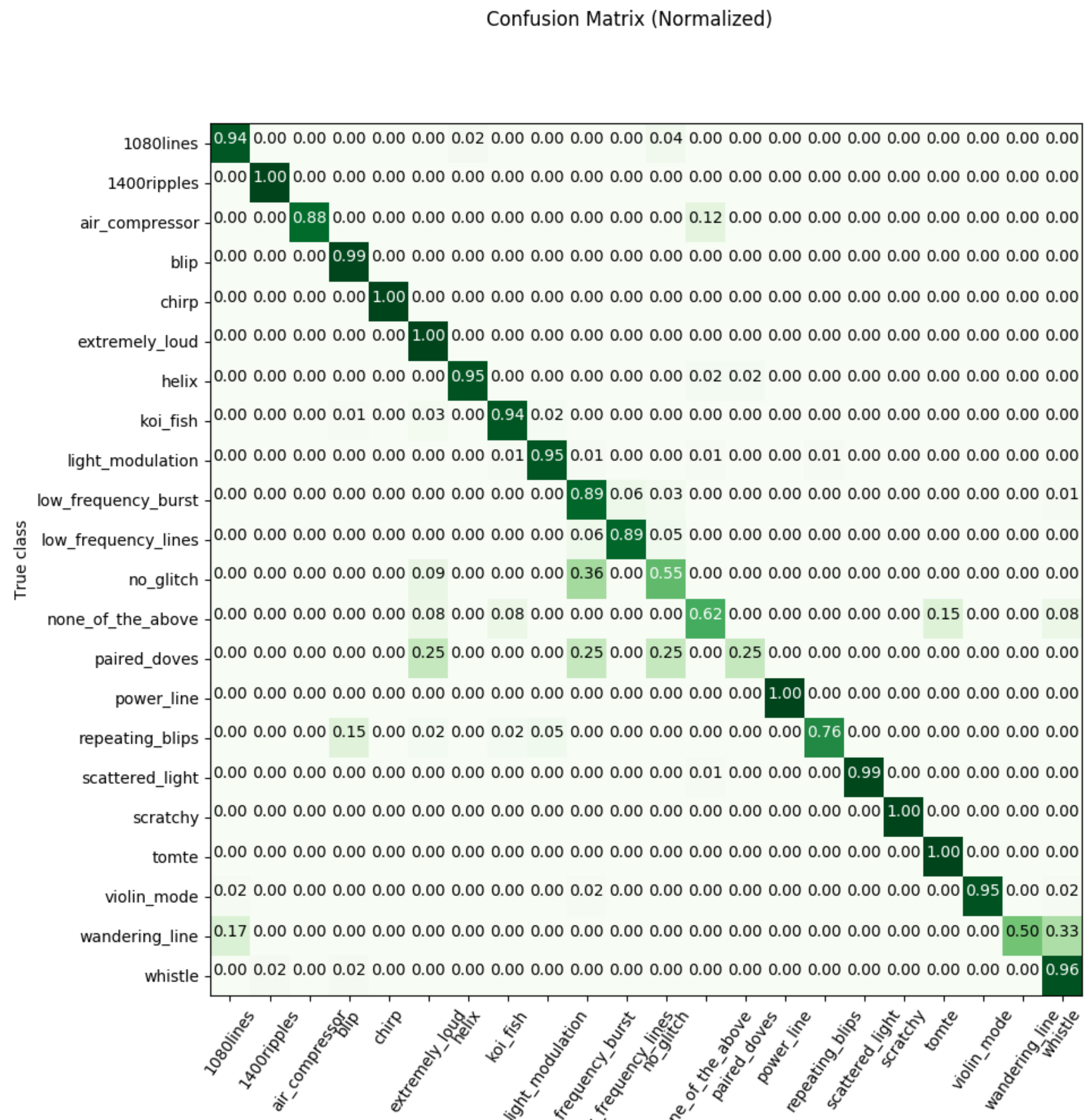
Move to the real data: O1 run

Dataset from GravitySpy images

Glitch name	# in H1	# in L1
Air compressor	55	3
Blip	1495	374
Chirp	34	32
Extremely Loud	266	188
Helix	3	276
Koi fish	580	250
Light Modulation	568	5
Low_frequency_burst	184	473
Low_frequency_lines	82	371
No_Glitch	117	64
None_of_the_above	57	31
Paired doves	27	-
Power_line	274	179
Repeating blips	249	36
Scattered_light	393	66
Scratchy	95	259
Tomte	70	46
Violin_mode	179	-
Wandering_line	44	-
Whistle	2	303

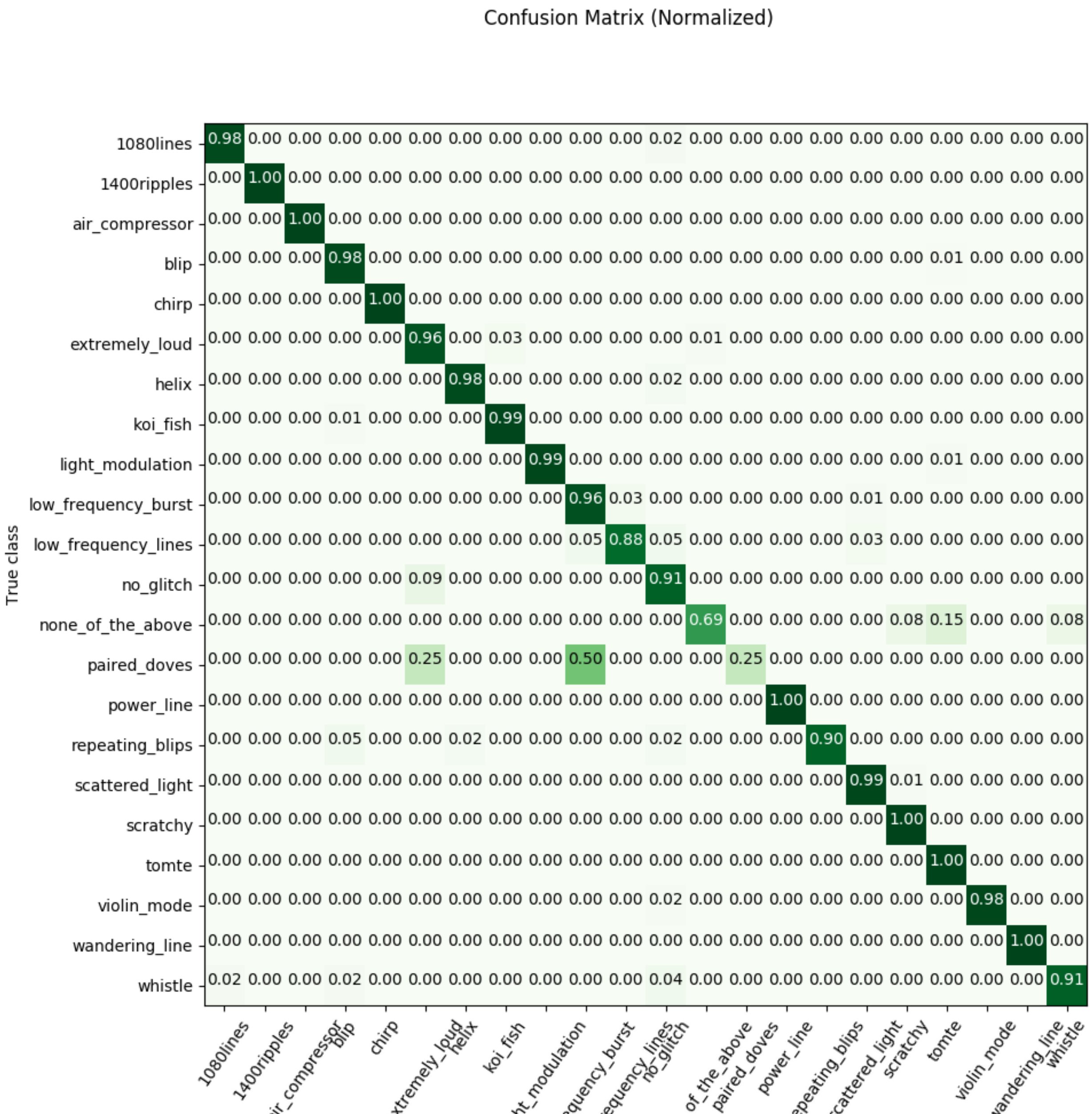
Classification results

Confusion
(or confusing?)
matrix



1 CNN block

Classification results



Full CNN stack

Consistent with Zevin+2017

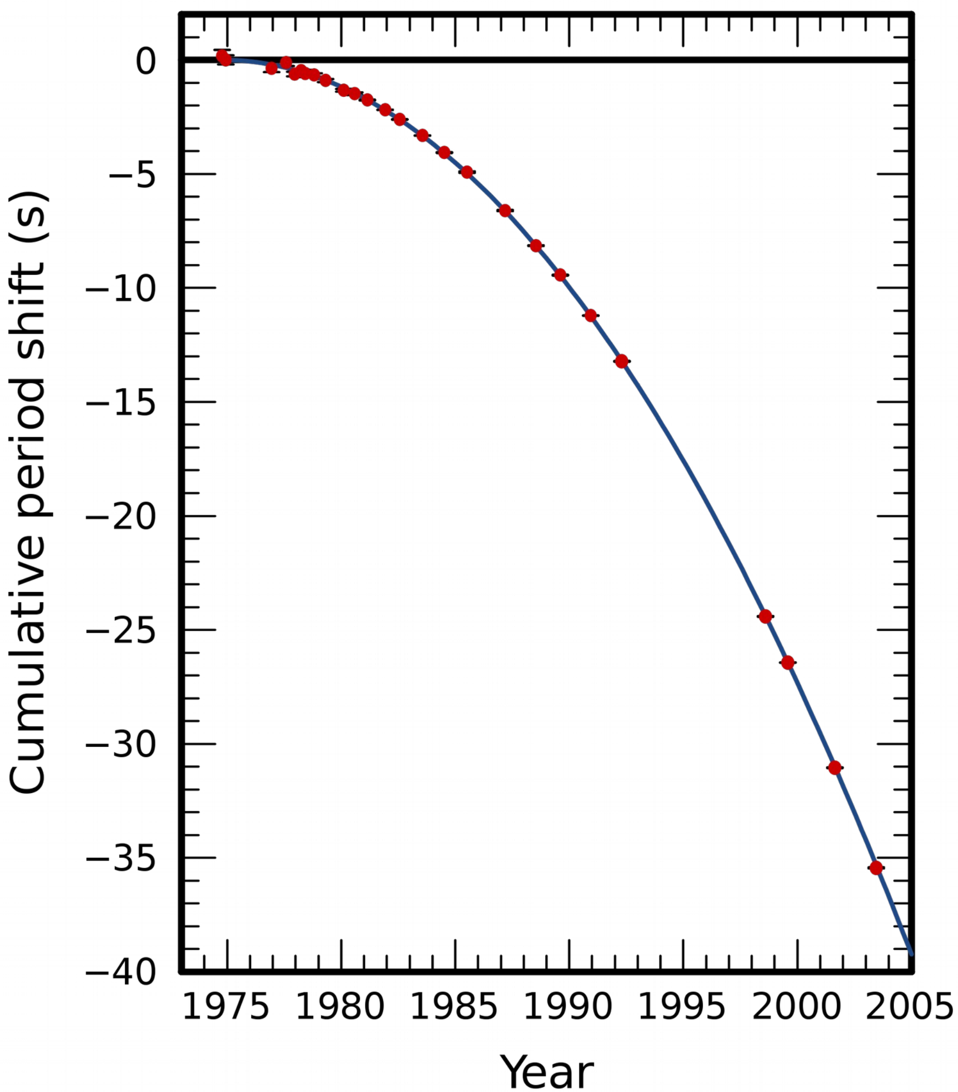
Conclusions and next steps

- Machine and deep learning methods are growing fast in GW community
- We have tested and developed image-based deep learning for classification of noise glitches
- Time-frequency images as input data
- Tested on simulations, real data and output of other pipelines
- Starting from image mapping, we developed a multichannel approach
- All methods and pipelines under testing in order to be ready for O3

Backup

The case of binary system PSR1916+13

- Discovered in 1978
- Distance 6.4 kpc
- P=7.7 h
- P pulsar 59 ms
- Orbital decay of 3.5 m/yr, or 76.5 us/year
-
- Nobel Prize on Physics 1993: R. A. Hulse e J. H. Taylor



First indirect proof of Gravitational Waves !

The picnic problem

Picnic planned, but the sky is cloudy

- 50% of rainy days start cloudy
- Cloudy mornings are common (40%)
- This is a dry month (usually 10% rain)

What is the chance of rain?

A=cloud
B=rain

$P_{rain} = 0.1$

P^{clouds} given rain $= 0.5$

$$P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{B})\,P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A})}$$

$P_{cloud} = 0.4$

$\rightarrow P = 0.12$

*

Completeness relationship

$$P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{B})\,P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A}|\mathcal{B})\,P(\mathcal{B}) + P(\mathcal{A}|\neg\mathcal{B})\,P(\neg\mathcal{B})} = \frac{\Lambda(\mathcal{B}|\mathcal{A})}{\Lambda(\mathcal{B}|\mathcal{A}) + P(\neg\mathcal{B})/P(\mathcal{B})}$$

$P(\neg\mathcal{B}) = 1 - P(\mathcal{B})$ P of not B

$\Lambda(\mathcal{B}|\mathcal{A}) := \dfrac{P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A}|\neg\mathcal{B})}$ Likelihood ratio

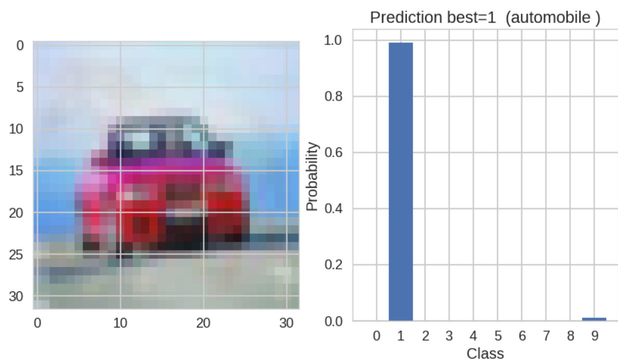
Using the Odd Ratios:

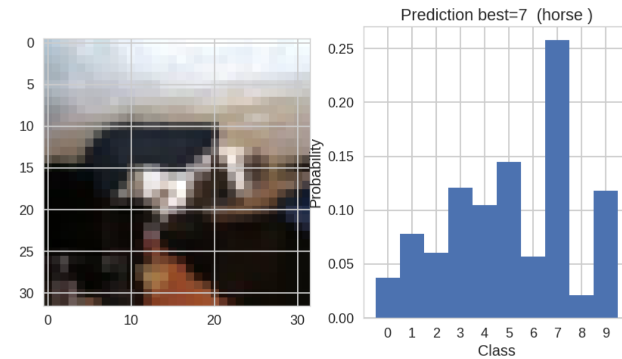
$$O(\mathcal{B}|\mathcal{A}) = O(\mathcal{B})\,\Lambda(\mathcal{B}|\mathcal{A})$$

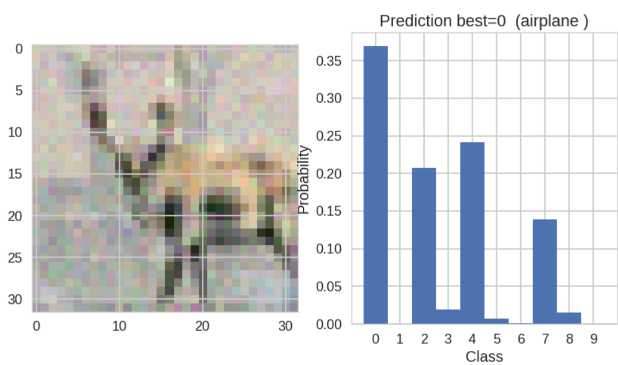
Where O(B) = P(B)/P(--B)

*

Testing CNN on standard images







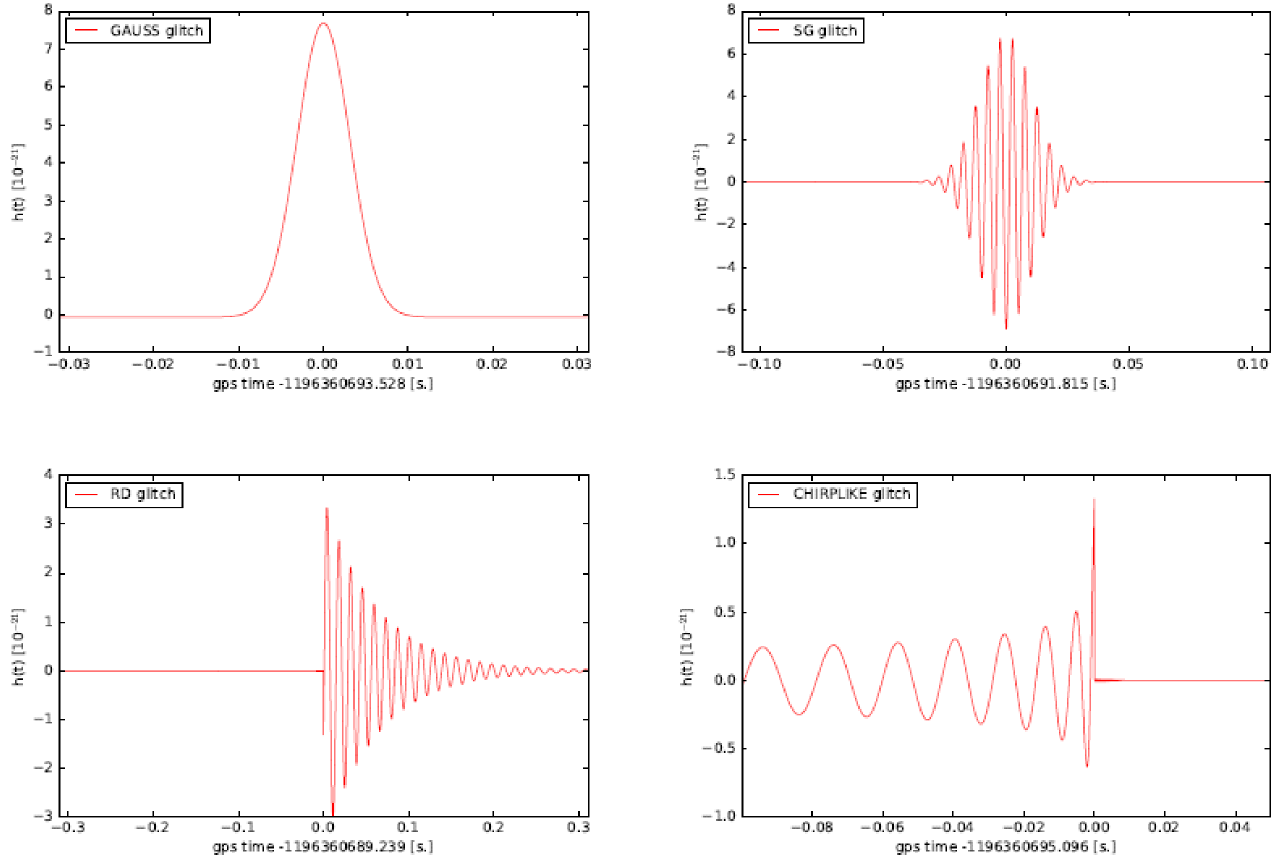
Deer (4)

Dataset:
- CIFAR-10
- 60k 32x32 images
- 10 classes

CNN implementation
- Python(Keras+TensorFlow)
- Run on GPU

Tests on simulations (II)

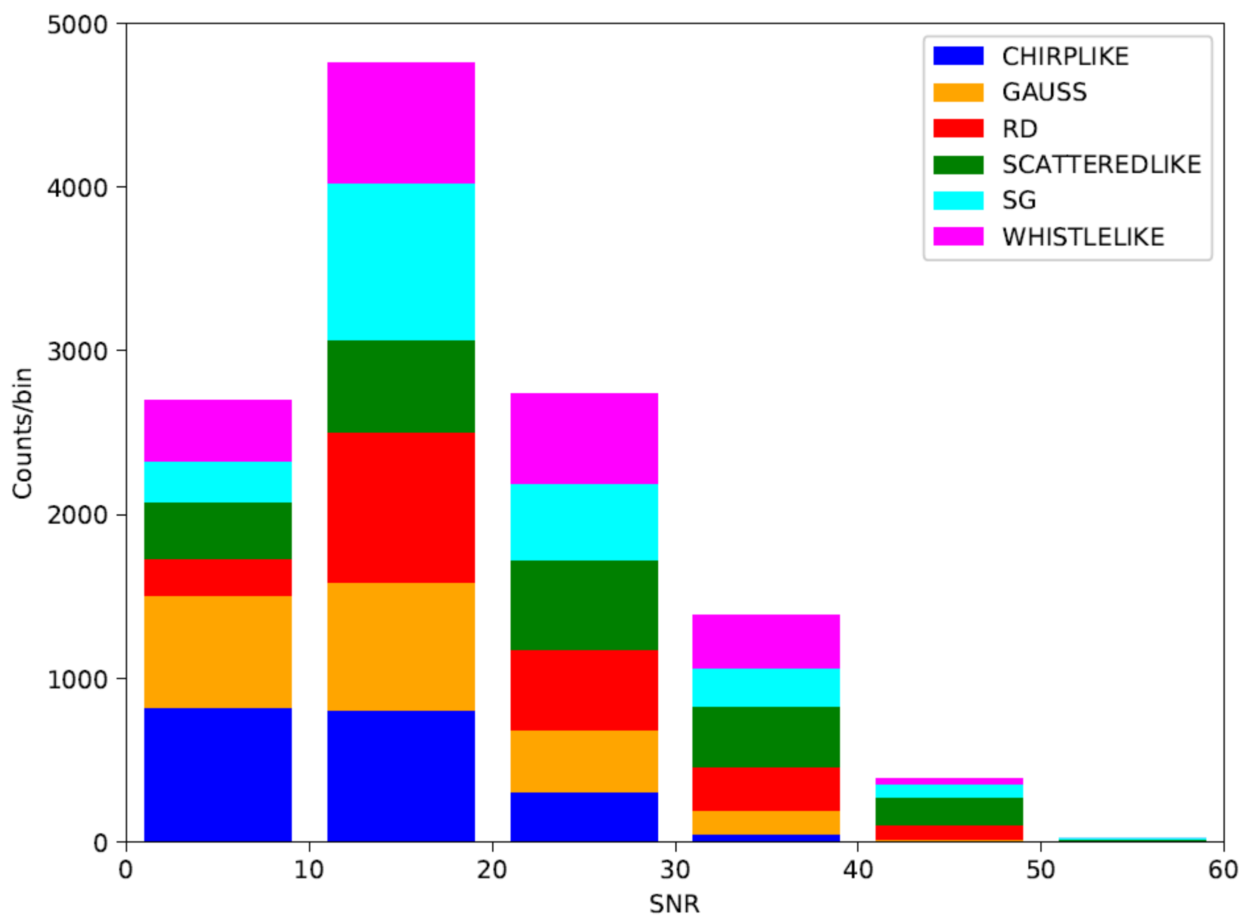
Examples of simulated timeseries



To show the glitch timeseries here we don’t show the noise contribution

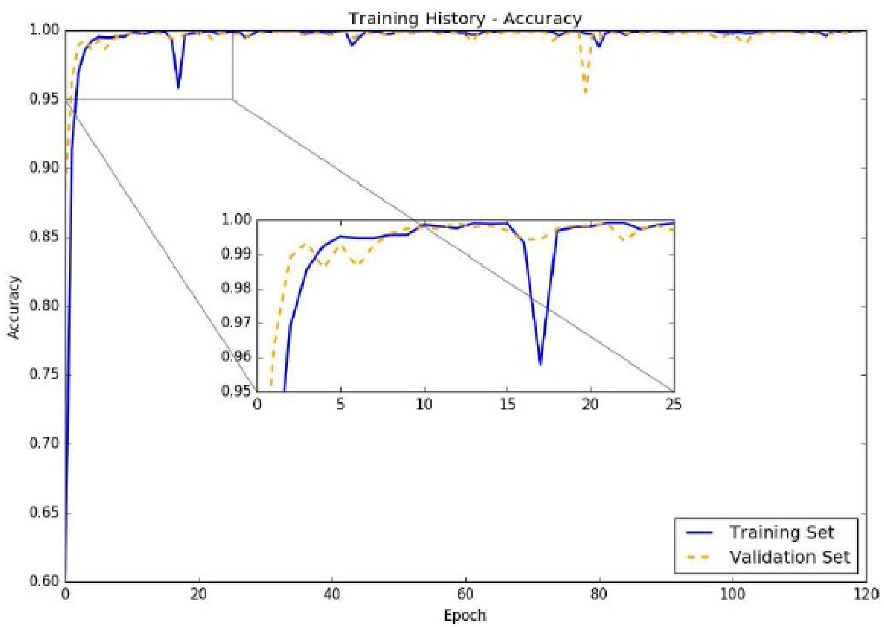
Tests on simulations (II)

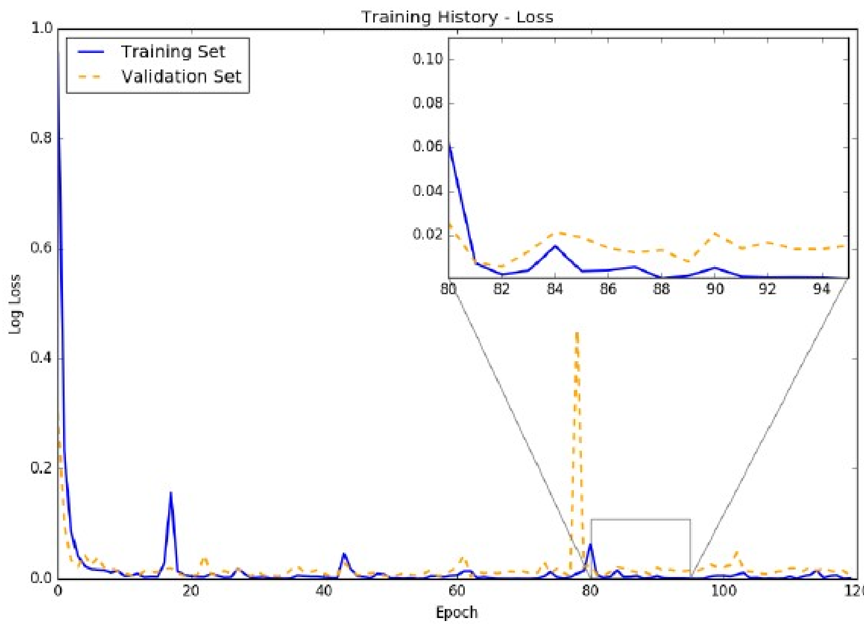
- Simulated time series with 8kHz sampling rate
- Glitches distributed according to Poisson statistics
- 2000 glitches per each family
- Glitch parameters random to achieve various shapes and Signal-To-Noise ratio



Training the CNN

- Datasets of 14k images
- Training:validation:test → 75:15:15
- Image size 241 x 513 pixels
- Depending on memory constraints, image size are reduced
- Perform grid search to tune hyperparameters
- Training time ~ few hrs for ~100 epochs
- Classification on the fly (~1 ms/image)





Classification results – metrics

- We checked performances of different architectures (including a baseline one)

Metric	Accuracy	Precision	Recall	F1 score	Log loss
SVM	0.971	0.972	0.971	0.971	0.08
Shallow CNN	0.986	0.986	0.986	0.986	0.04
1 CNN block	0.991	0.991	0.991	0.991	0.02
3 CNN blocks	0.998	0.998	0.998	0.998	0.008

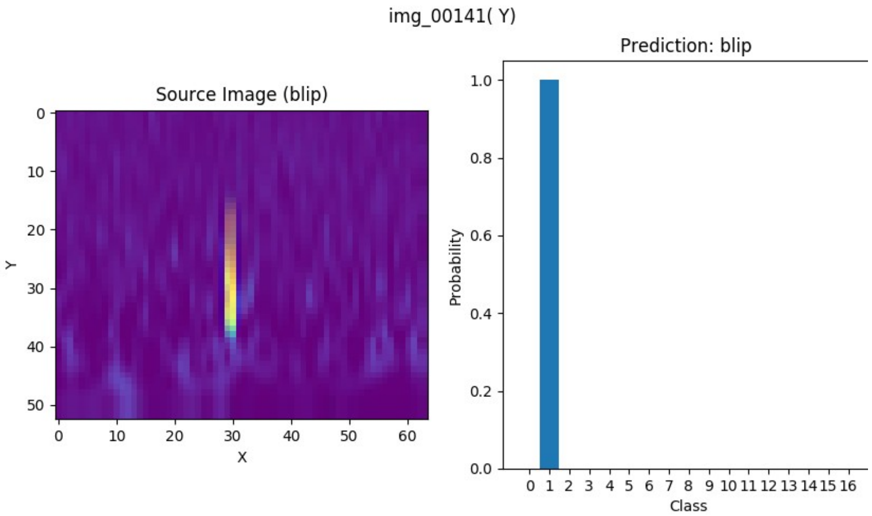
Linear Support Vector Machine

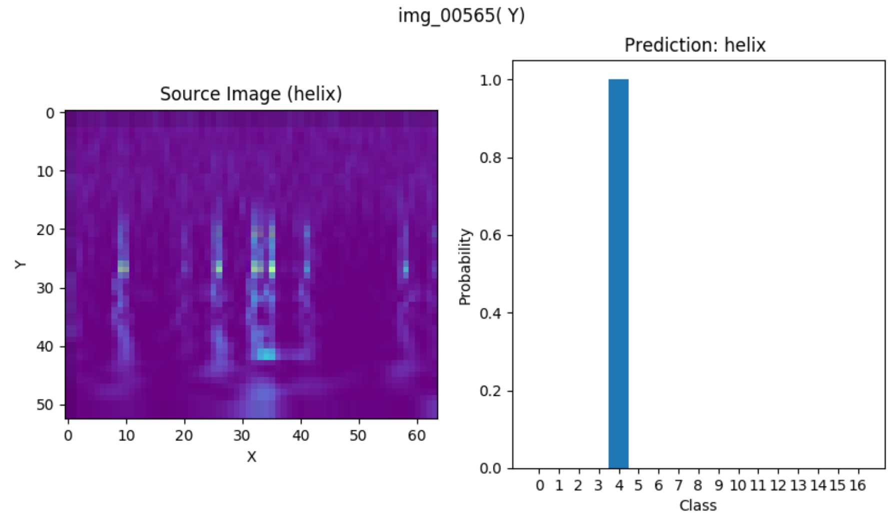
CNN with 1 hidden layer

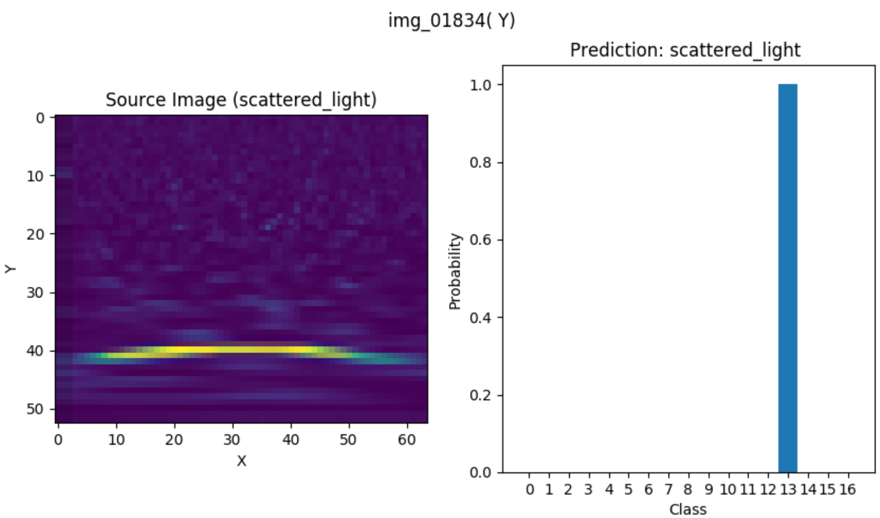
CNN with one block
(2 CNNs+Pooling&Dropout)

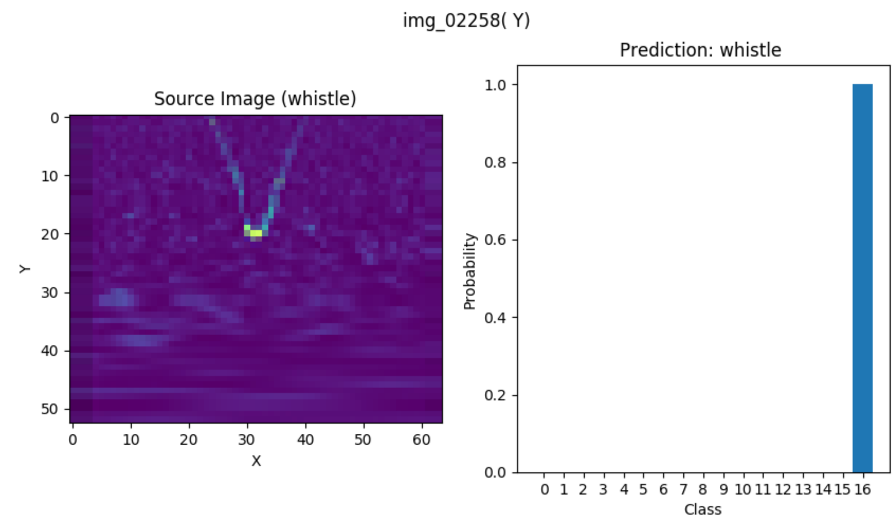
Deep 4-blocks CNNs

Sample classification results







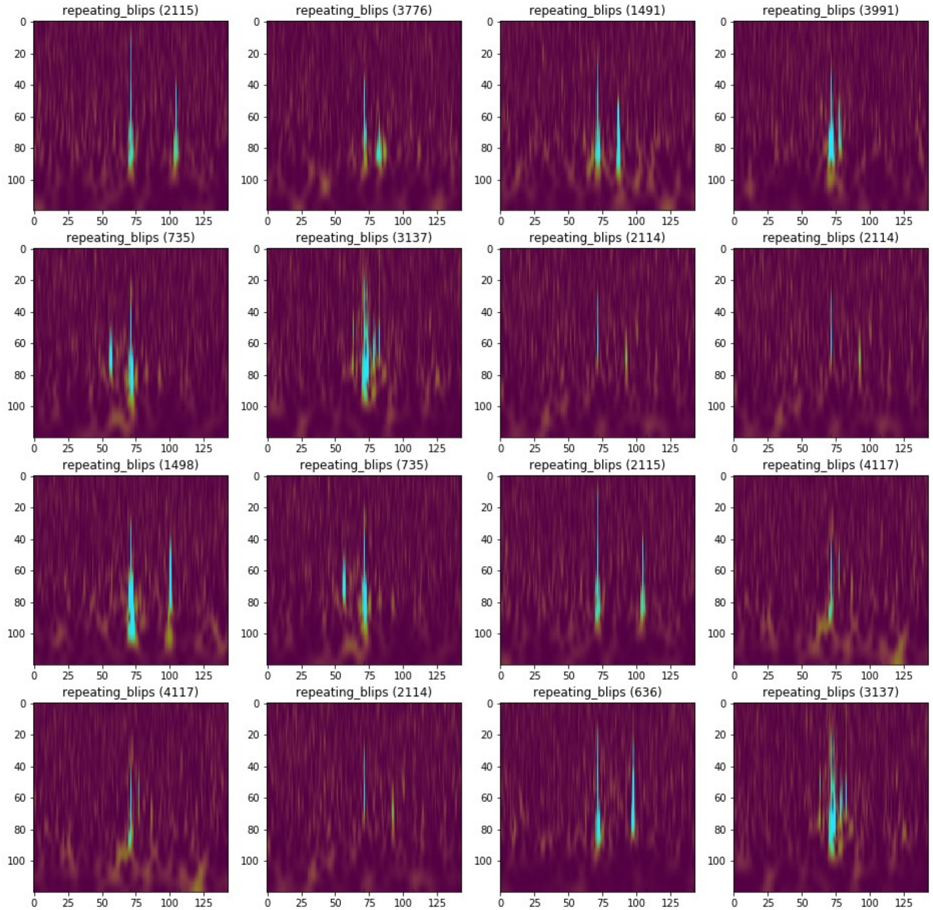


Preparing for O3

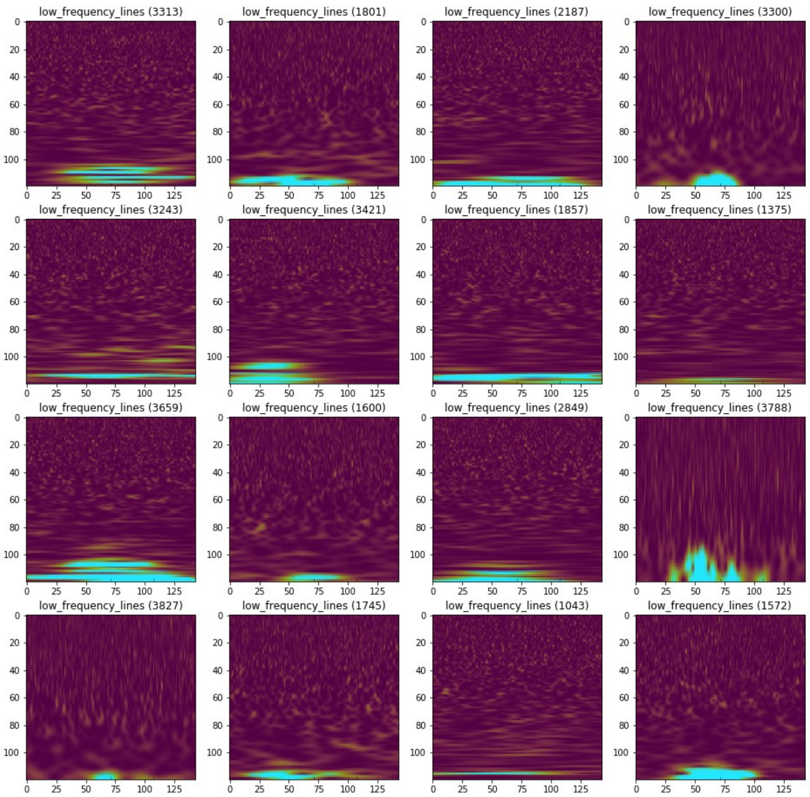
Pipeline tested and CNN trained

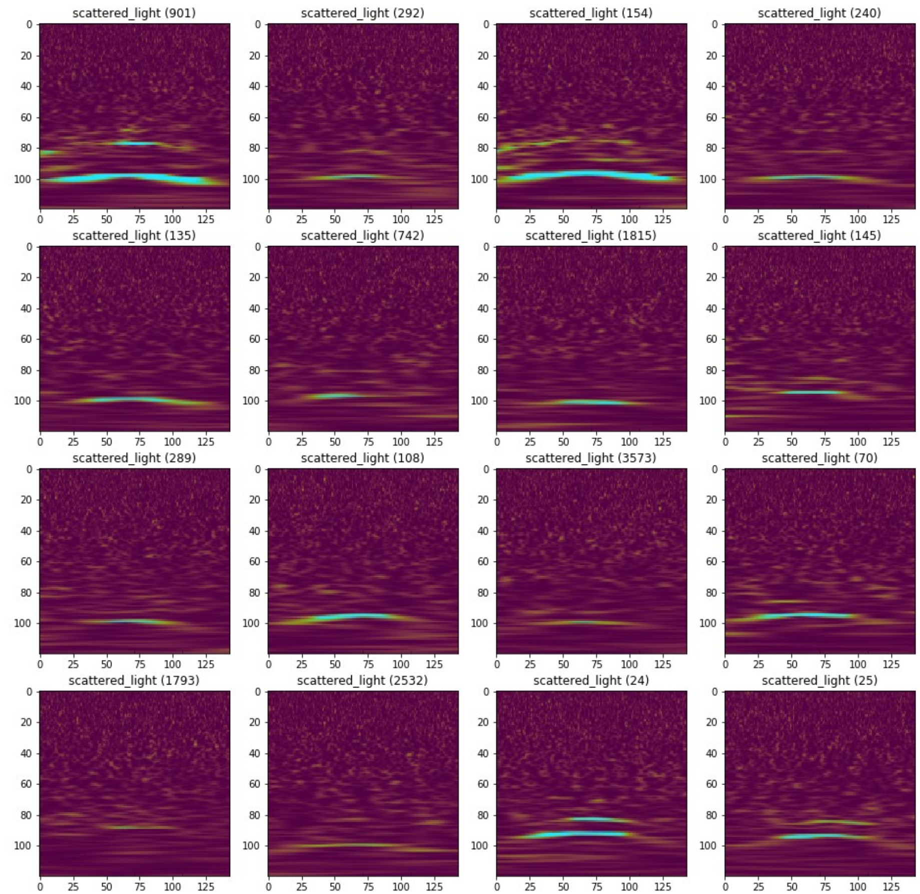
- Prepare to run continuosly on data

- Test on stretch of O1 public data
  - Used 1 week from GPS 1132704017. SNR>10
  - ~4500 glitches
- Pick the glitches and classify them
  - Check against LIGO-DB classes



Preparing for O3

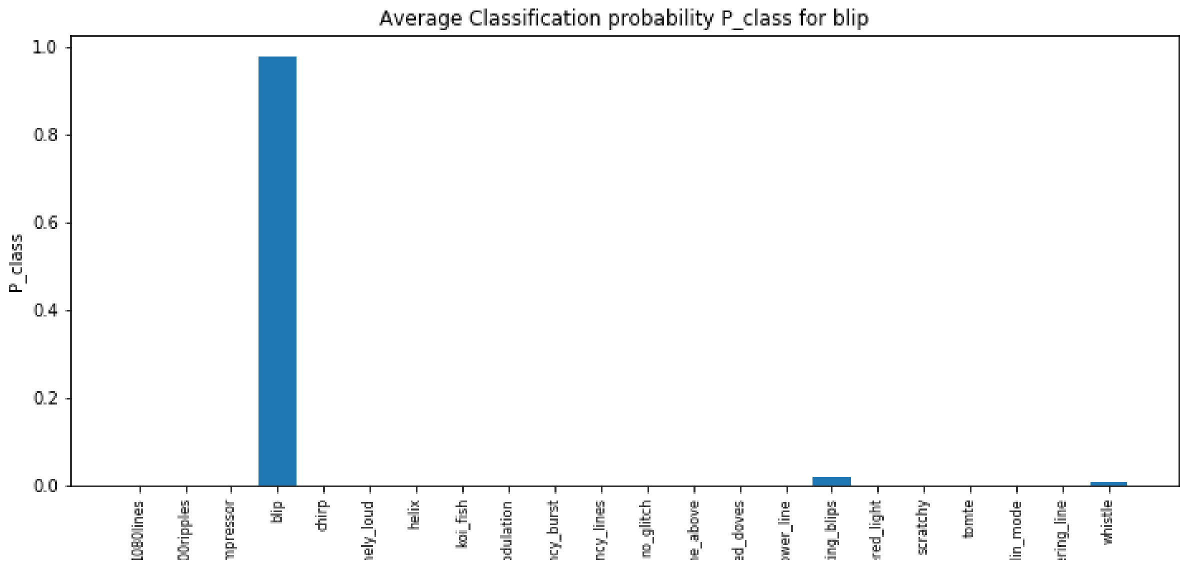


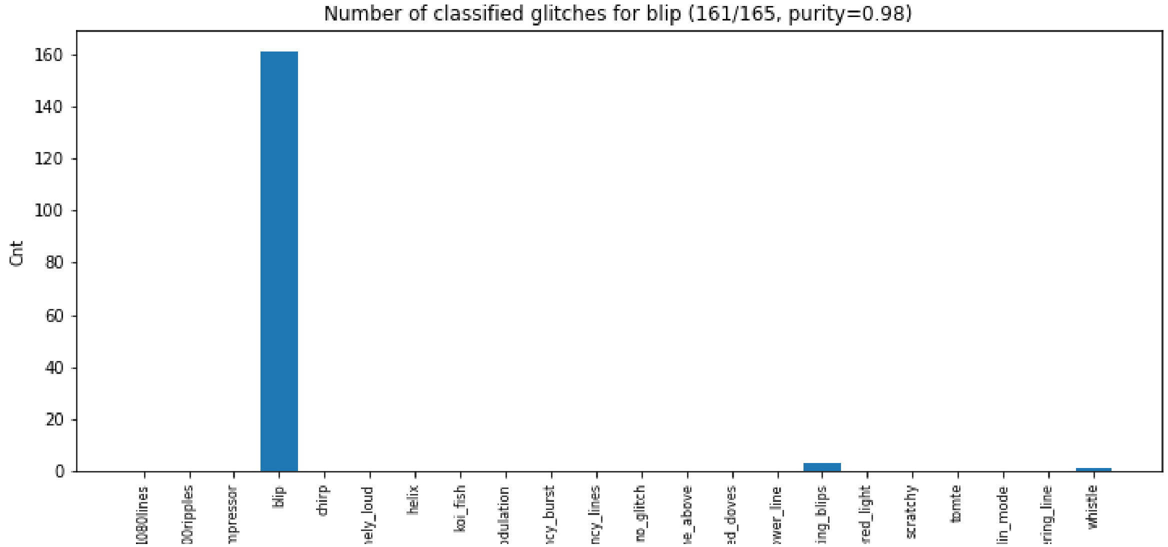


Other random examples

Preparing for O3

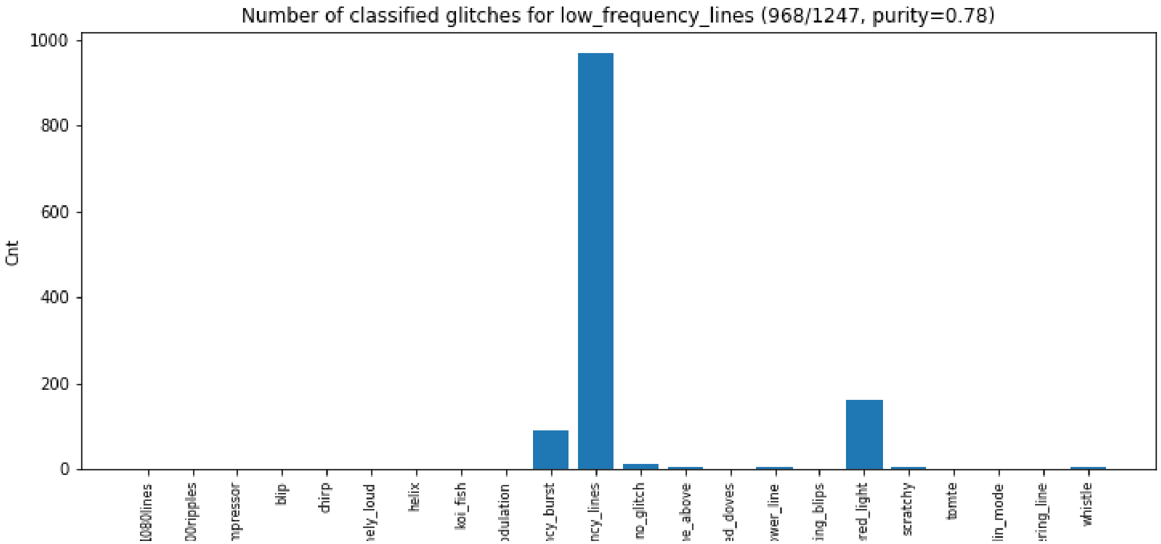
Compare with the classes in the DB



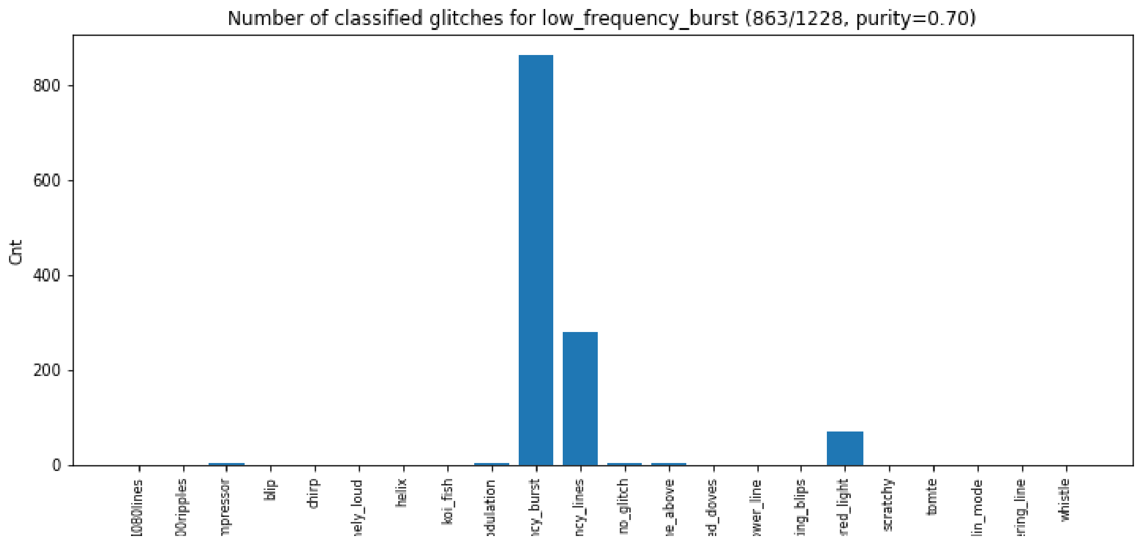


Preparing for O3

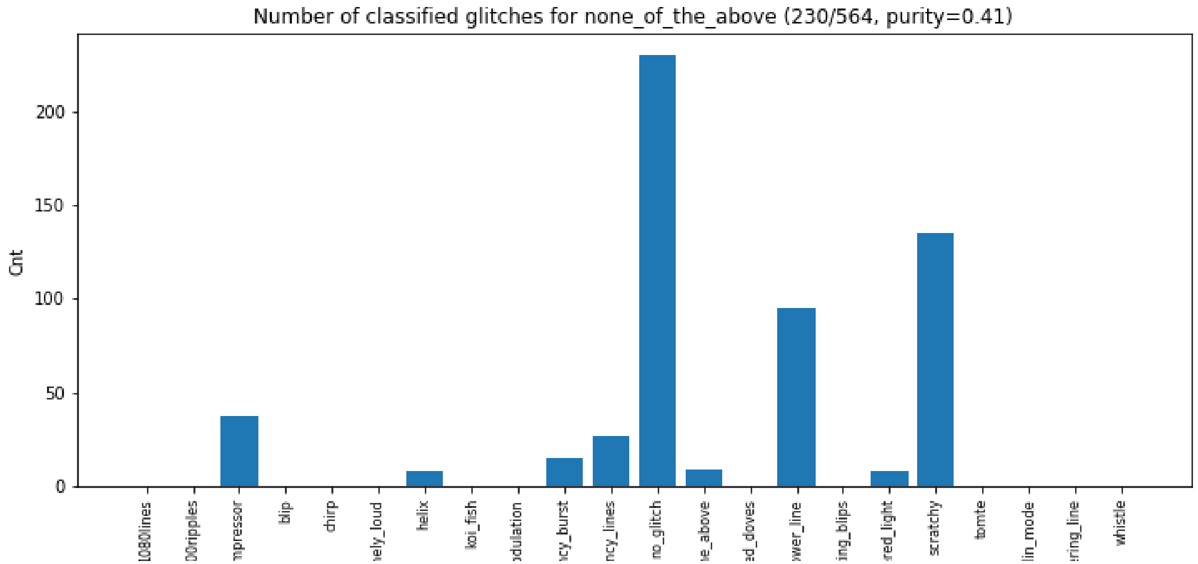
Sometimes a slight mismatch (as found in simple display)



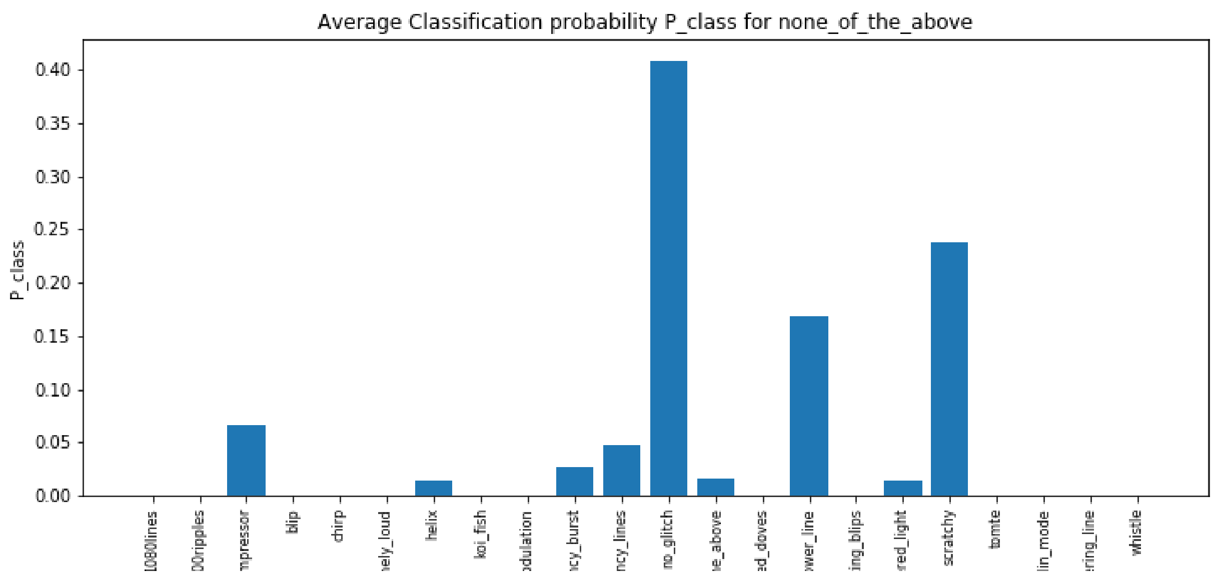
Average classification probability ~0.7-0.8



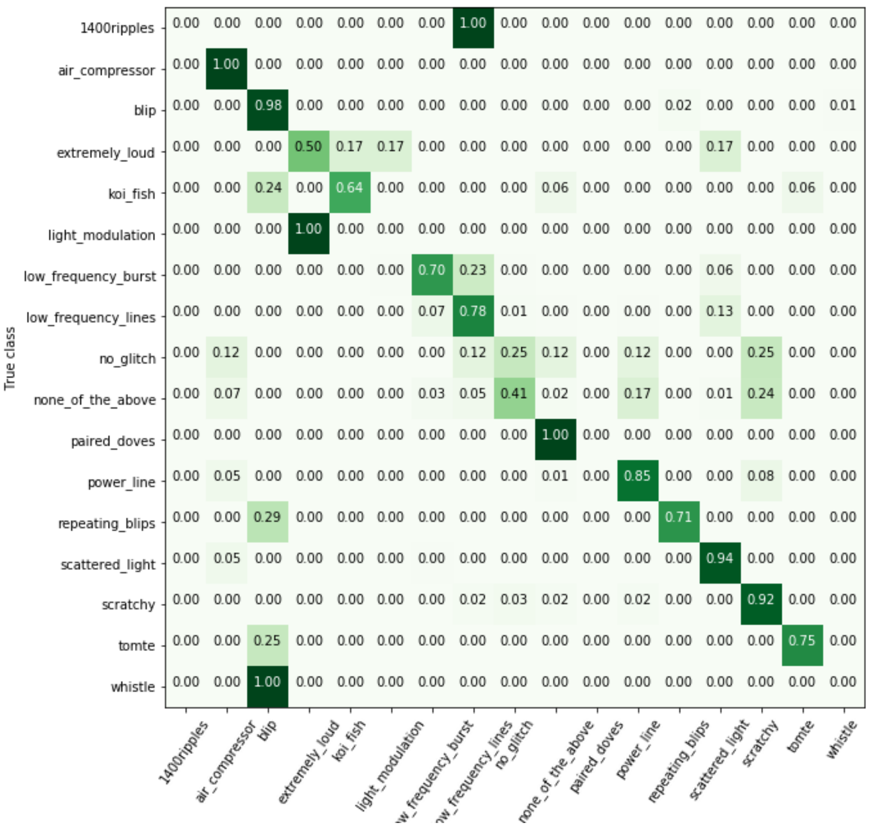
Exploring "None of the above"

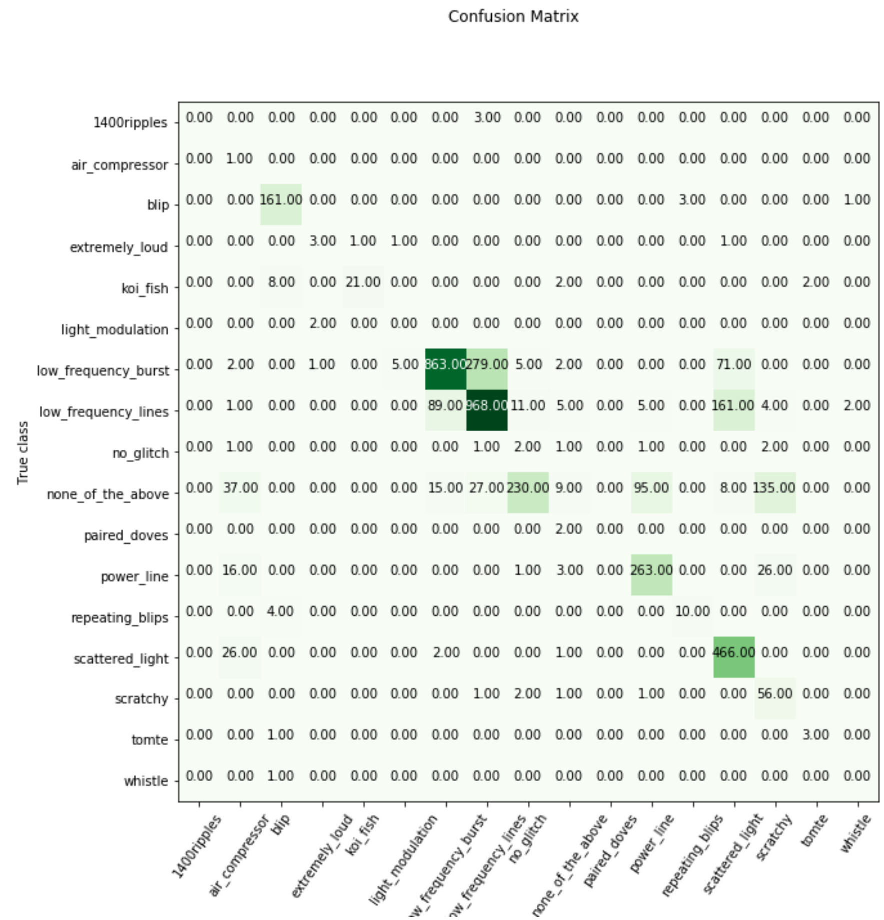


Low classification probabilities, as expected



Confusion matrix

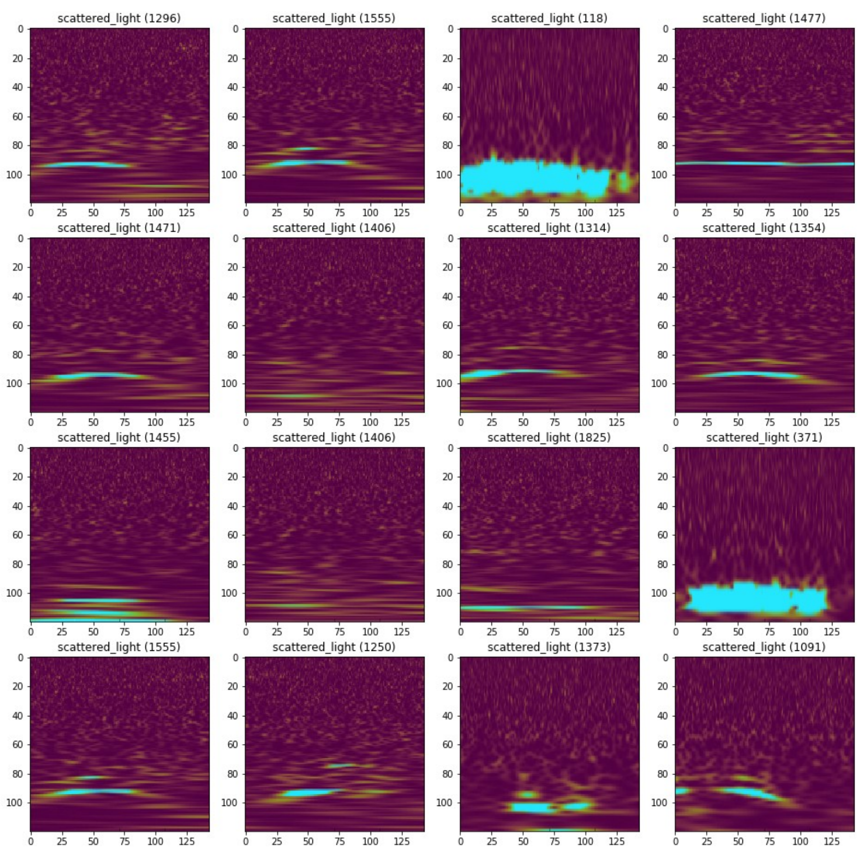


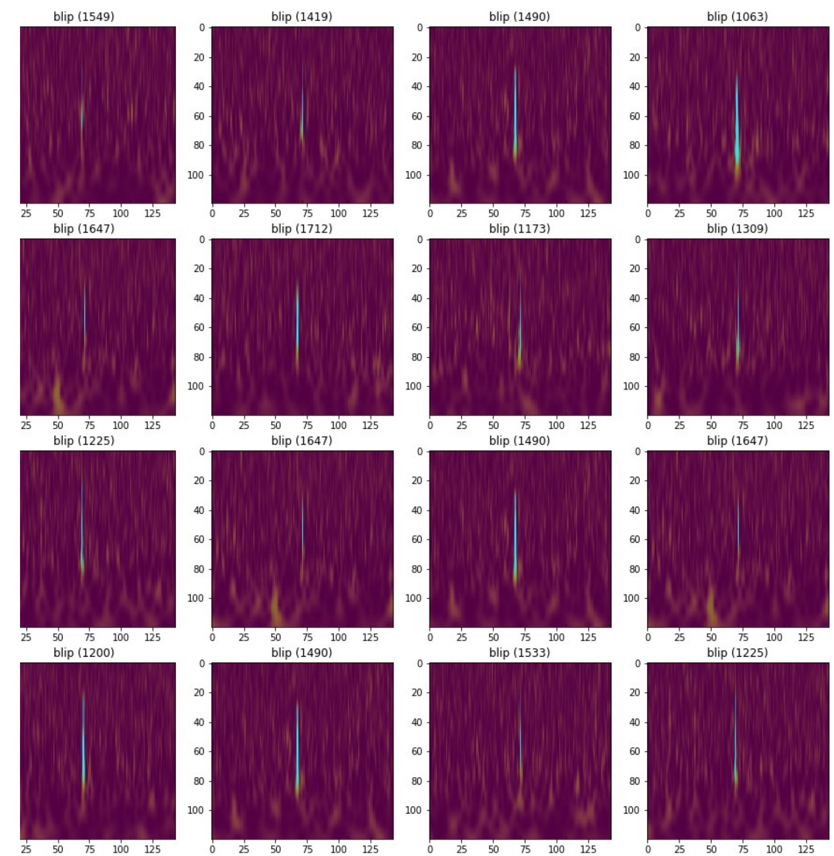


Repeat with more statistic

Interfacing with WDF pipeline

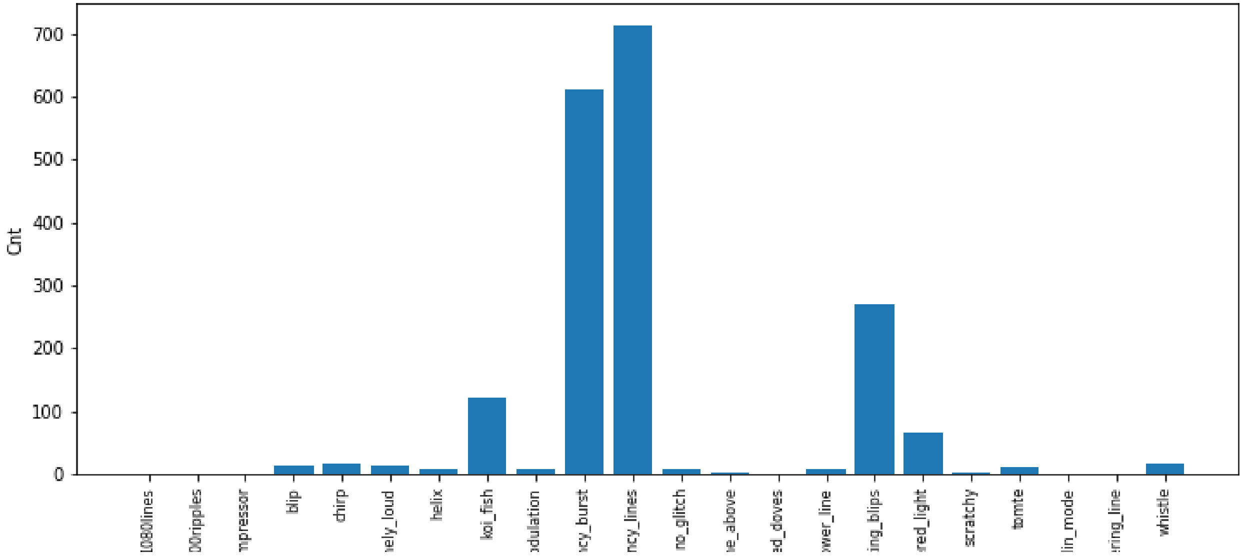
Apply the learned model to the glitches found with an external pipeline,e.g. the Wavelet Detection Filter pipeline (see Elena Cuoco's talk)
Same GPS time, sample of ~1800 glitches



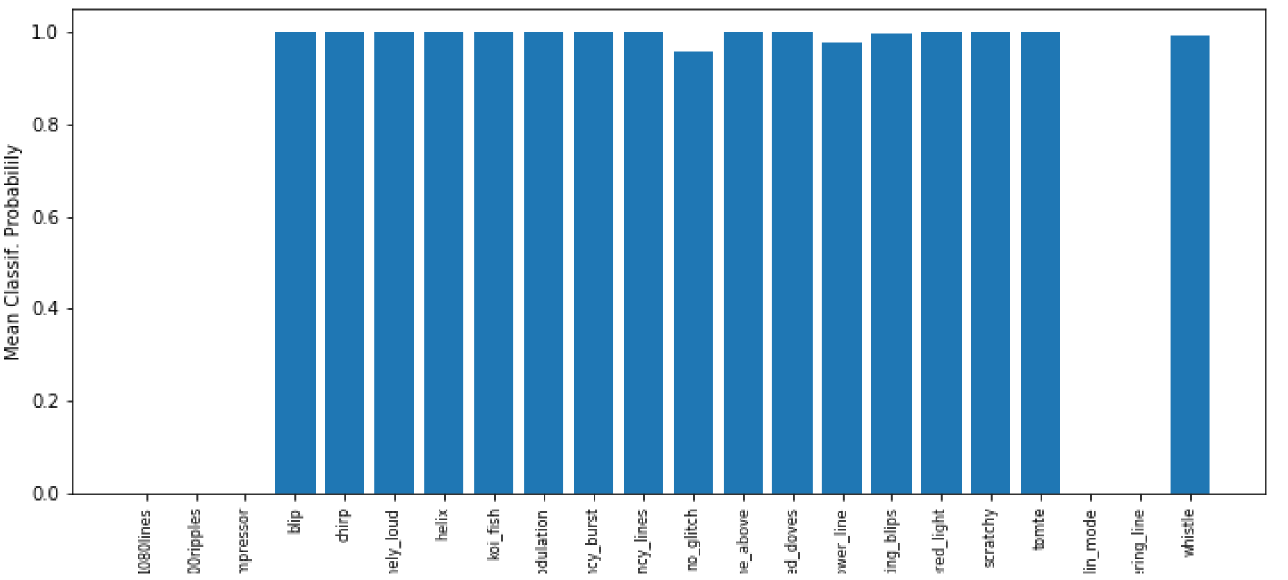


Interfacing with WDF pipeline

Study the distribution and mean classification probability across classes

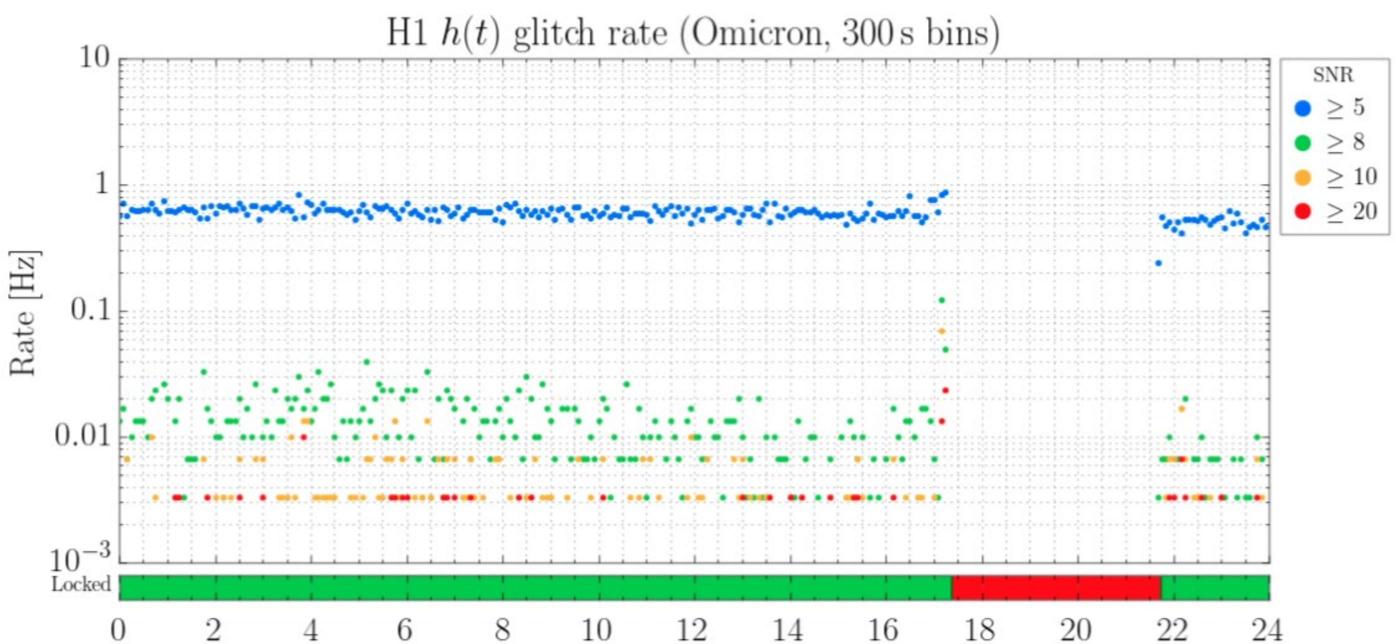


Classifiction looks good



Glitches and multichannel images

- For some glitch classes origin is unknown (e.g. blips)
  - Goal: estimate how important is a channel in the formation of a glitch
- Method
  - Identify a list of glitches (e.g. omicron, etc)
  - Select list of auxiliary channels and get timeseries
  - Condition timeseries
- Rank them according to a “importance” figure
- We started from representing multichannel FFT as images..



Work by K. Mogushi, MR, M. Cavaglià, G. Nicolini, G. Cella

Gaussian Noise

Noise time series of N points (duration T, sampling dt) with every x_i is random Gaussian (e.g. mean = 0,variance=1)

We call it Gaussian noise.

$$p_x(\{x_j\}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1} x_j^2\right\}$$

The limit to the continuum

$$\lim_{\Delta t \to 0} \sum_{j=0}^{N-1} x_j^2 \Delta t = \int_0^T x^2(t)\,dt \approx \int_{-\infty}^{\infty} |\tilde{x}(f)|^2\,df$$

$$S_x(f) = 2\int_{-\infty}^{\infty} R_x(\tau)\,e^{-2\pi i f\tau}\,d\tau = \lim_{\Delta t \to 0} 2\sigma^2 \Delta t$$

*

Optimal detection filter

We test the H_1 hypothesis (signal) vs the null Hypothesis H_0 (no signal)
→ H1 = not H0

So, if B = H_1 and A = s(t), we can compute the odds ratio $O(H_1|s)$

We can use $O(H_1|s) = O(H_1)\Lambda(H_1|s)$ Likelihood ratio $\Lambda(\mathcal{H}_1|s) = \frac{p(s|\mathcal{H}_1)}{p(s|\mathcal{H}_0)}$.

No dependence on the data

For H_0 :

$$p(s|\mathcal{H}_0) = p_n[s(t)] \propto e^{-(s,s)/2}$$

For H_1 :

$$p(s|\mathcal{H}_1) = p_n[s(t) - h(t)] \propto e^{-(s-h,s-h)/2}$$

(If noise n(t) = s(t)-h(t) is Gaussian)

*

Optimal detection

We need a way to evaluate the presence of a signal in the data.
$n(t)$: noise
$h(t)$: signal

$s(t) = n(t)+h(t)$: recorded signal

We test these hypotheses

H0 (Null Hypothesis) = “no signal”, i.e. $s(t) = n(t)$
H1 (Alternative) = “signal”, i.e. $s(t) = n(t) + h(t)$

We can compute the ratios between P(H1) and P(H0), given the recorded data $s(t)$

This is something that we can address with Bayes Theorem

*

Bayes Theorem (I)

P that A is true given B

Joint probability (A and B)

$$P(\mathcal{A}|\mathcal{B}) := \frac{P(\mathcal{A},\mathcal{B})}{P(\mathcal{B})},$$

As well as: $$P(\mathcal{B}|\mathcal{A}) := \frac{P(\mathcal{A},\mathcal{B})}{P(\mathcal{A})}$$

This leads to Bayes theorem:

Prior

P of A given B

Posterior

$$P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{B})\,P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A})}$$

Evidence

*

Optimal detection filter

We test the H_1 hypothesis (signal) vs the null Hypothesis H_0 (no signal)
→ H1 = not H0

So, if B = H_1 and A = s(t), we can compute the likelihood ratio

Likelihood ratio $$\Lambda(\mathcal{H}_1|s) = \frac{p(s|\mathcal{H}_1)}{p(s|\mathcal{H}_0)} .$$

For H_0 :

$$p(s|\mathcal{H}_0) = p_n[s(t)] \propto e^{-(s,s)/2}$$

For H_1 :

$$p(s|\mathcal{H}_1) = p_n[s(t) - h(t)] \propto e^{-(s-h,s-h)/2}$$

(If noise n(t) = s(t)-h(t) is
Gaussian)

*

Optimal detection filter

The likelihood becomes

$$\Lambda(\mathcal{H}_1|s) = \frac{e^{-(s-h,s-h)/2}}{e^{-(s,s)/2}} = e^{(s,h)}e^{-(h,h)/2}$$

(s,h) Only dependence from data

Likelihood depends monotonically on (s,h), therefore we call it optimal statistic

$$(s,h) = 4\,\mathrm{Re}\int_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)}\,df$$

Also called this **matched filter**

(a noise-weighted correlation of anticipated signal with data)

*