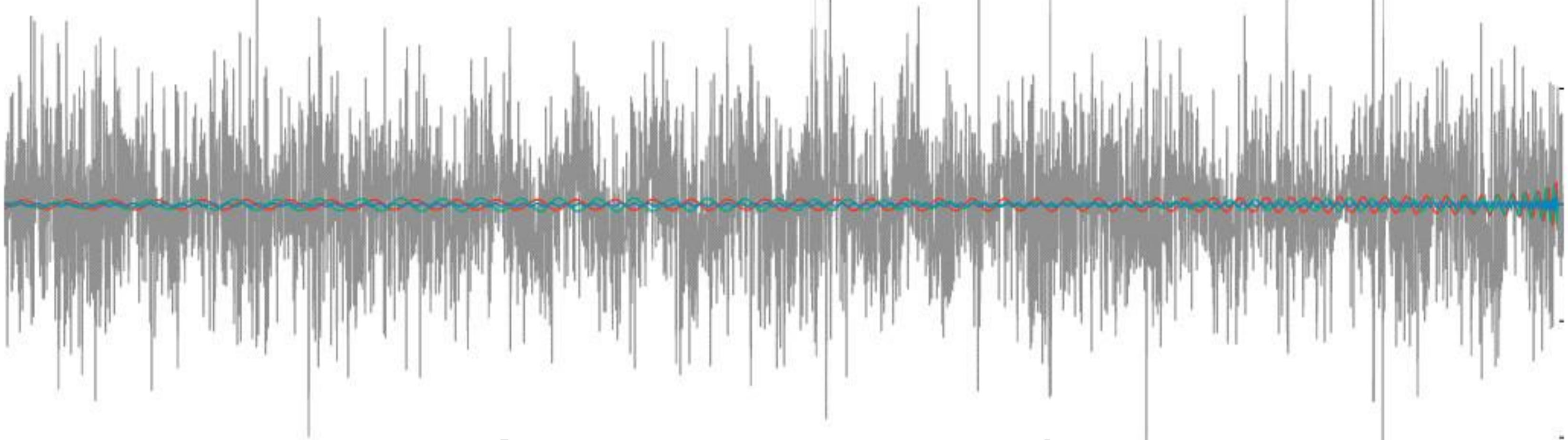
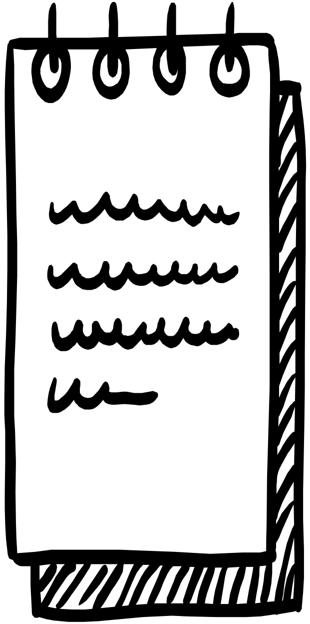
Machine Learning Application for Gravitational Wave science



Elena Cuoco, EGO & SNS & INFN-Pisa

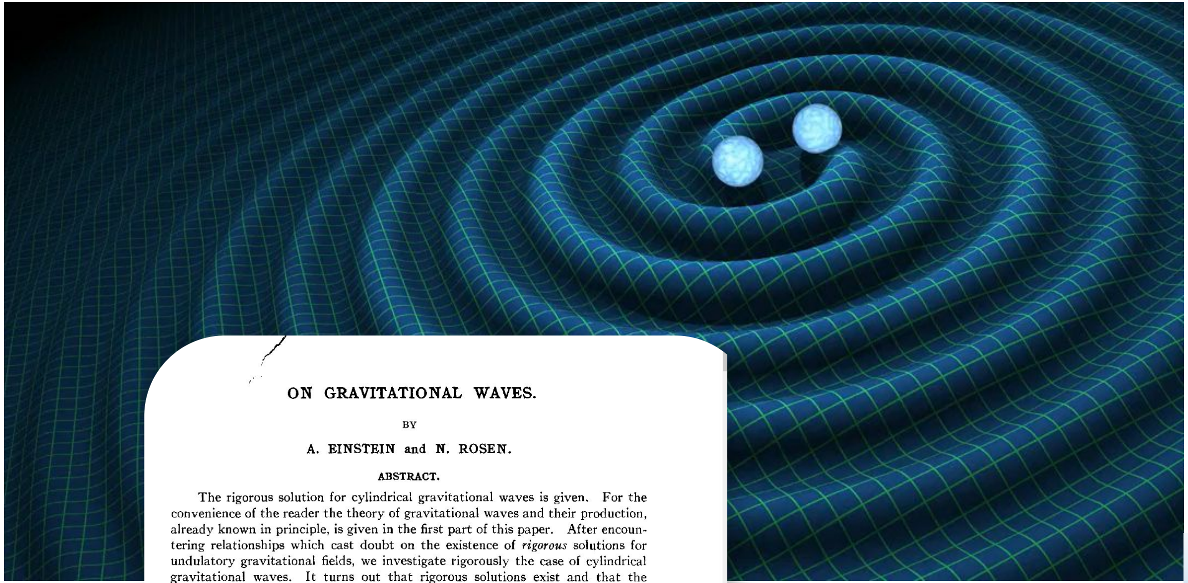
"Big Data within Science and Industry" September 22, 2023 , University of Milano-Bicocca

Outline



- Short overview of GW detector and data
- Data transformation and GW real data
- A roundup of use cases and applications of Machine Learning in Gravitational Wave science

What are gravitational waves



ON GRAVITATIONAL WAVES.

BY

A. EINSTEIN and N. ROSEN.

ABSTRACT.

The rigorous solution for cylindrical gravitational waves is given. For the convenience of the reader the theory of gravitational waves and their production, already known in principle, is given in the first part of this paper. After encountering relationships which cast doubt on the existence of *rigorous* solutions for undulatory gravitational fields, we investigate rigorously the case of cylindrical gravitational waves. It turns out that rigorous solutions exist and that the problem reduces to the usual cylindrical waves in euclidean space.

I. APPROXIMATE SOLUTION OF THE PROBLEM OF PLANE WAVES AND THE PRODUCTION OF GRAVITATIONAL WAVES.

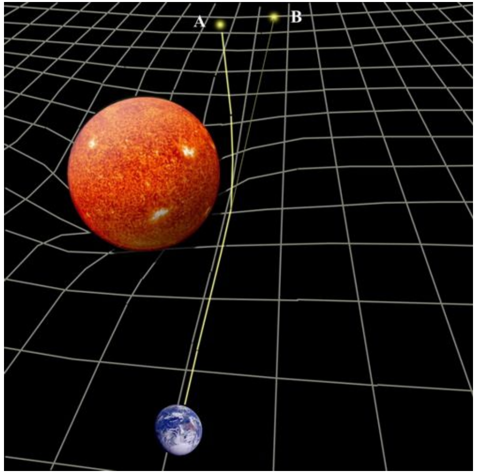
It is well known that the approximate method of integration of the gravitational equations of the general relativity theory leads to the existence of gravitational waves. The method used is as follows: We start with the equations

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R = - T_{\mu\nu}. \qquad (1)$$

We consider that the $g_{\mu\nu}$ are replaced by the expressions

$$g_{\mu\nu} = \delta_{\mu\nu} + \gamma_{\mu\nu}, \qquad (2)$$

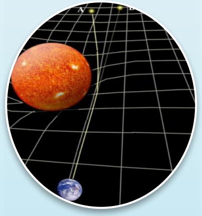
Gravitational Waves (1916)



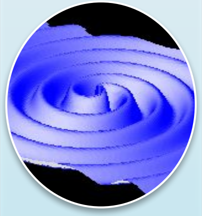
$$\mathbf{G}_{\mu\nu} = \frac{8\pi G}{c^4} \mathbf{T}_{\mu\nu}$$

General Relativity (1915)

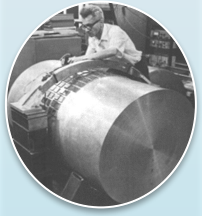
The GW search: a long history…



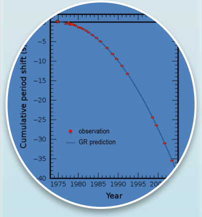
1915
- General Relativity



1916
- Gravitational Waves



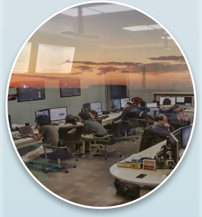
1966
- Weber and Resonant Bars



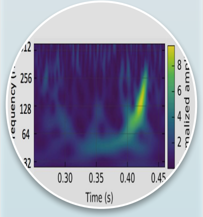
1974
- Hulse-Taylor:Observing the pulsar binary PSR B1913+16

1980-1990
- Cryogenic Resonant Bars

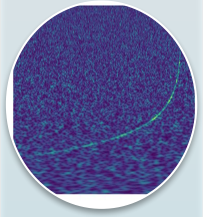
1993
- The approval of Virgo Experiment



1999+
- Data taking from LIGO an Virgo

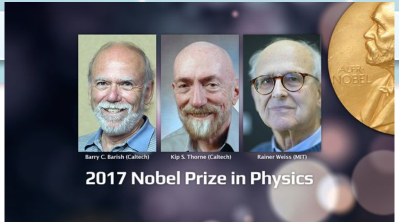


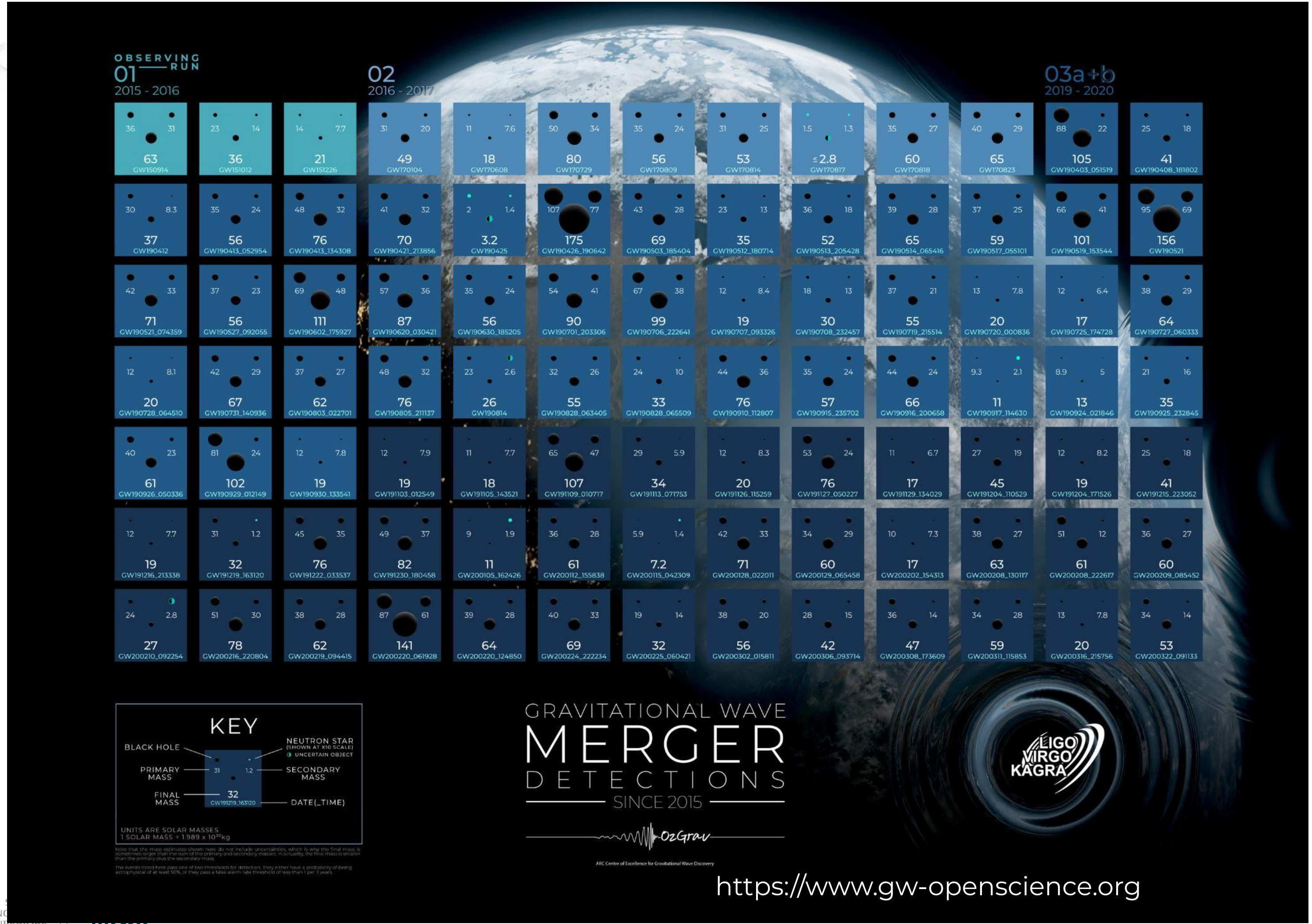
2015
- GW BHBH detection



2017
- BNS detection: Multi-messenger Astronomy

~100 years



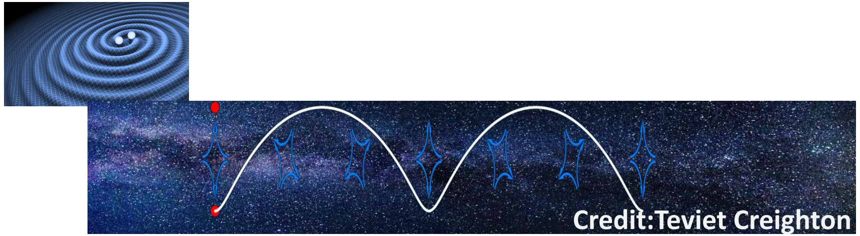


91 events!!!

https://www.gw-openscience.org

Credit: Carl Knox (OzGrav, Swinburne University of Technology).

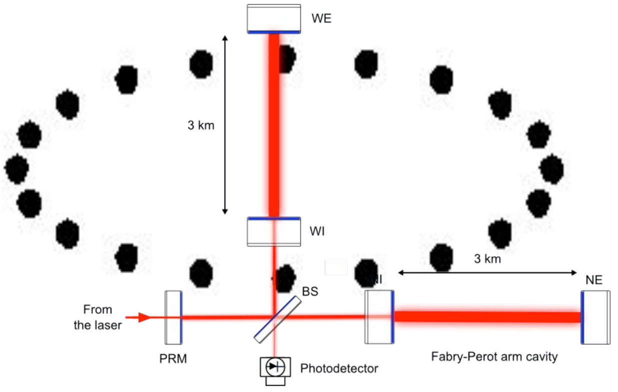
The Gravitational Wave detectors

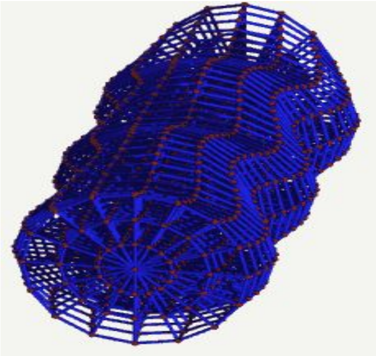


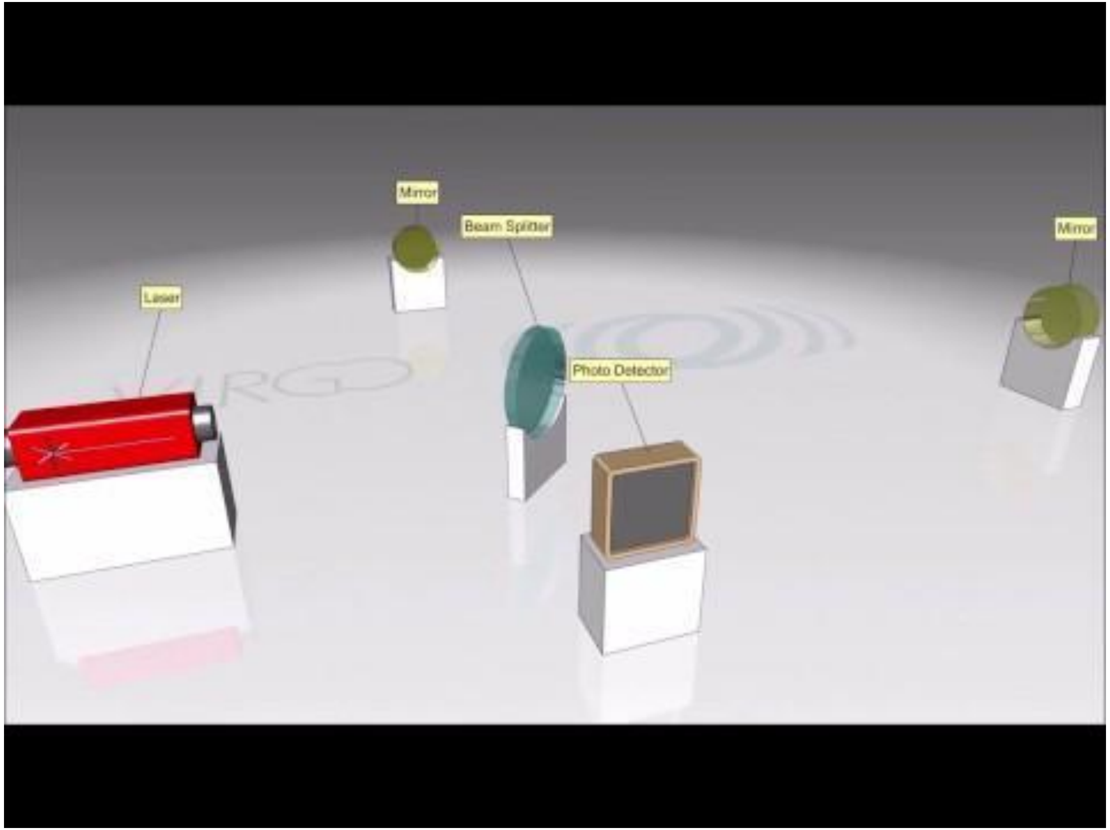
SCUOLA NORMALE SUPERIORE

EGO - Virgo

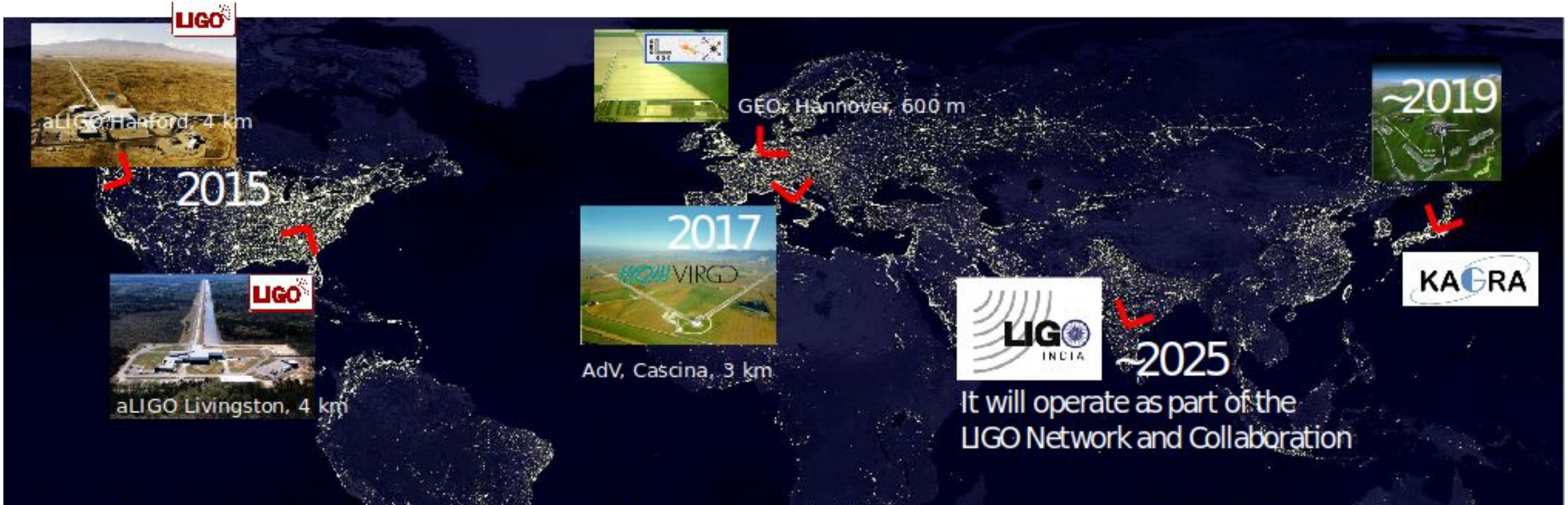
GW detection







International Collaboration: ground based detectors



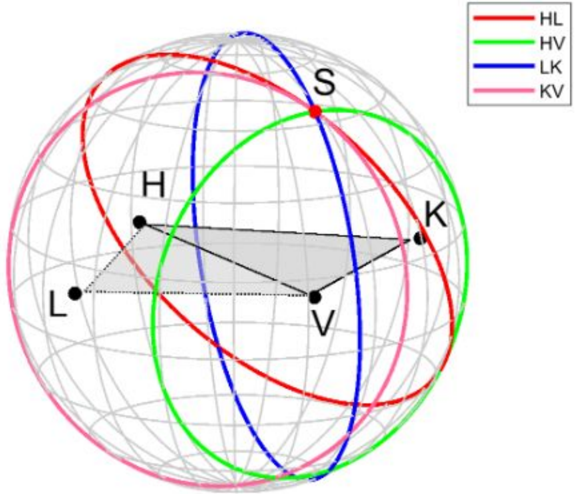
LIGO Scientific Collaboration:
- 1263 collaborators (including GEO)
- 20 countries
- 9 computing centres
- ~1.5 G$ of total investment

Virgo Collaboration:
- 343 collaborators
- 6 countries
- 6 computing centres
- ~0.42 G€ of total investment

KAGRA Collaboration:
- 260 collaborators
- 12 countries
- 5 computing centres
- ~16.4 G¥ of construction costs

Why more than 1 detector?

- *Source localization using only timing for a two-site network yields an annulus on the sky.*
- *For three detectors, the time delays restrict the source to two sky regions which are mirror images with respect to the plane passing through the three sites.*
- *With four or more detectors, timing information alone is sufficient to localize to a single sky region, and the additional baselines help to limit the region to under 10 deg^2 for some signals.*

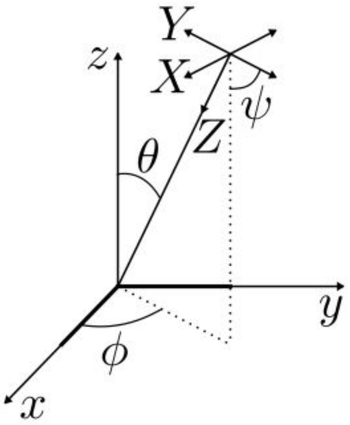


arXiv:1304.0670

- 2 detector → 100 -1000 deg^2
- 3 detector → 10 - 100 deg^2
- 4 detector → $<$ 10 deg^2

Detector response and Antenna Patterns

The detector response can be written



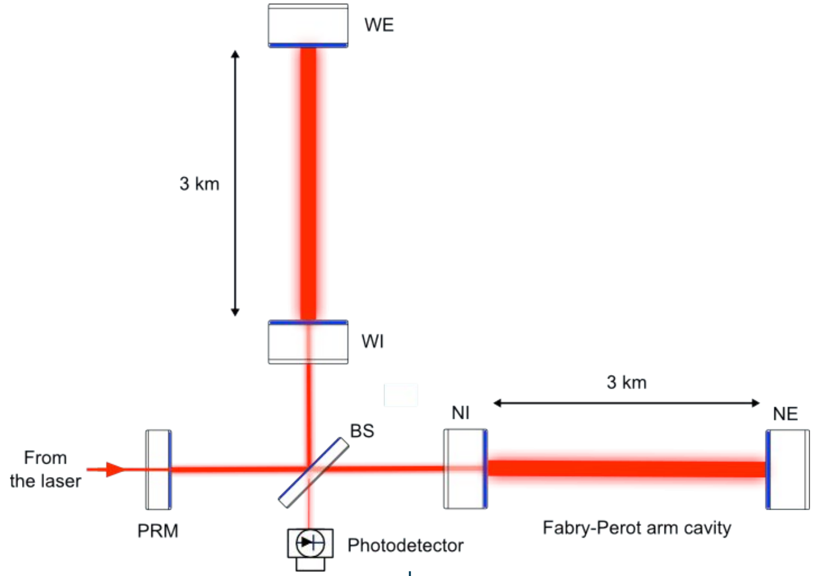
$$F_+ = \frac{1}{2}\left(1+\cos^2\theta\right)\cos 2\phi\cos 2\psi + \cos\theta\sin 2\phi\sin 2\psi$$

$$F_\times = \frac{1}{2}\left(1+\cos^2\theta\right)\cos 2\phi\sin 2\psi + \cos\theta\sin 2\phi\cos 2\psi$$

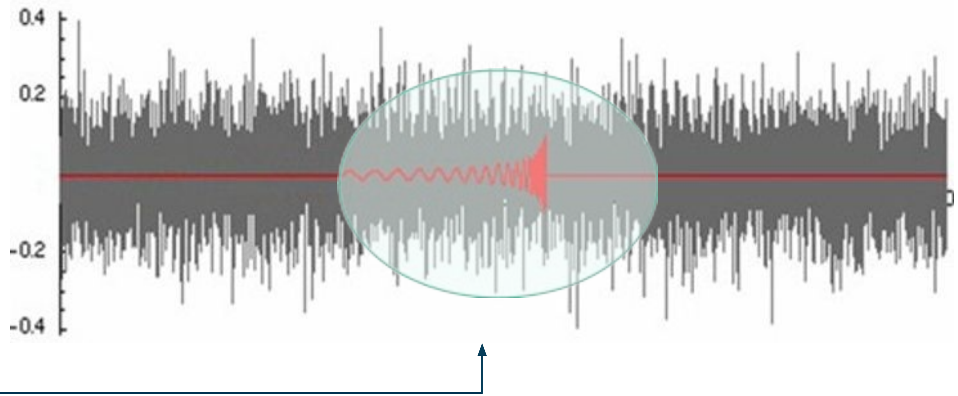
$$h(t) = h_+F_+ + h_\times F_\times$$

This let us have the localization of the source

The GW detector data



- **Time series sequences**... noisy time series with low amplitude GW signal buried in

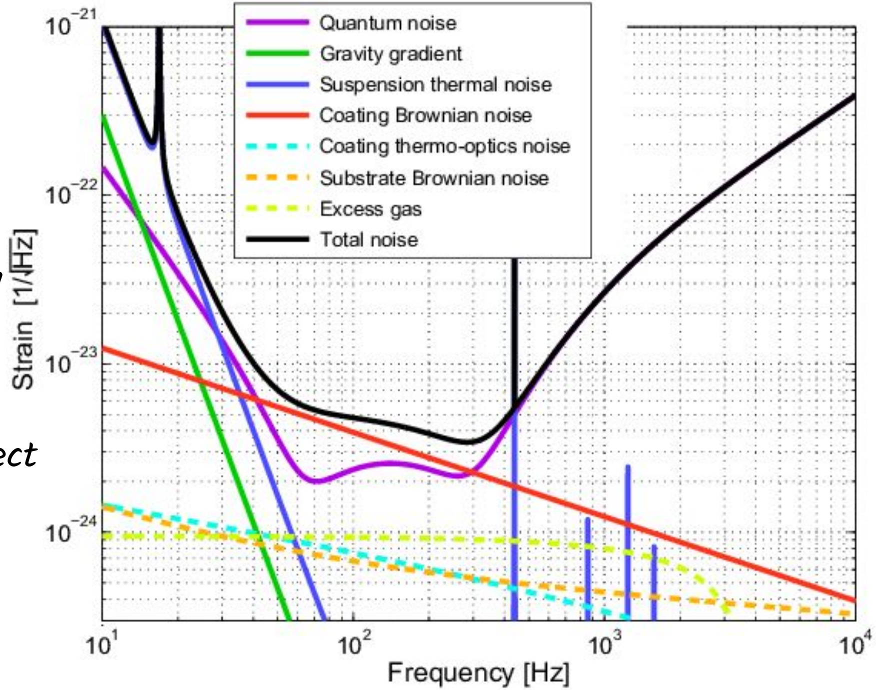


SCUOLA NORMALE SUPERIORE

EGO - Virgo

Theoretical Detector sensitivity

- *Noise at the output of the detector is the superimposition of different contribution*
- *If we consider the noise as a stationary stochastic process we can identify it with Its Covariance and so Power Spectral density*
- *The sensitivity curve define the minimum value of amplitude signal we are able to detect*

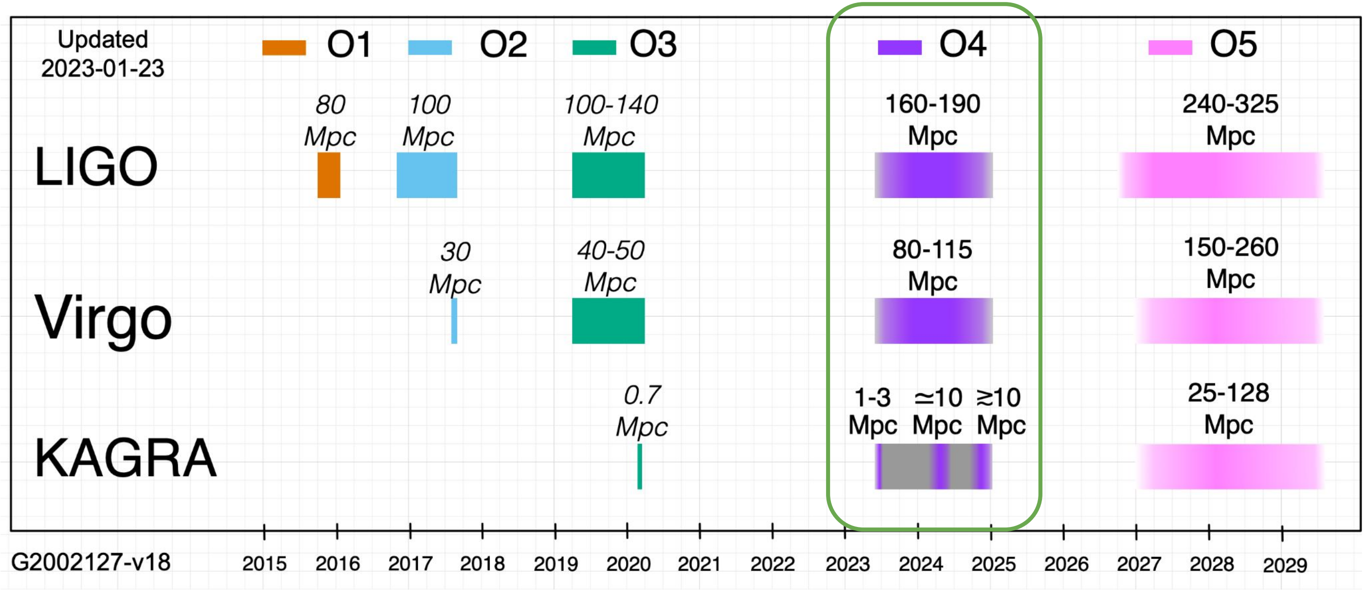


The O-run timeline

The detector **strain sensitivity** is the minimum *detectable* value of the strain produced by an incoming GW:

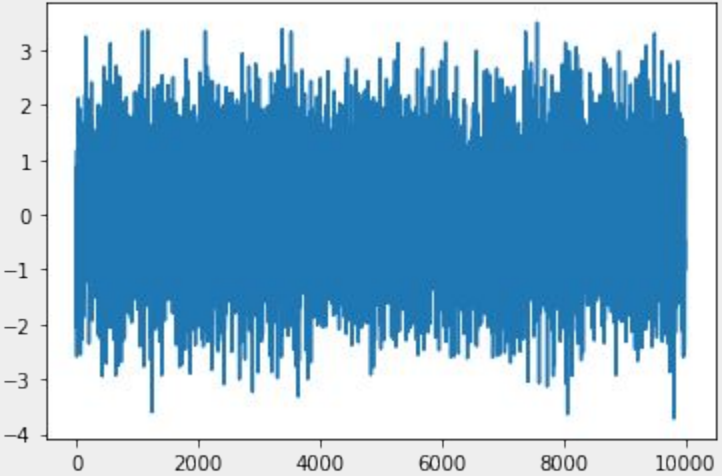
⇒ It is determined by the **detector noise**.

BNS inspiral range: the distance, averaged over GW polarizations and directions in the sky, at which a single detector can observe with matched-filter Signal-to-noise Ratio (SNR) of 8 the inspiral of two neutron stars.



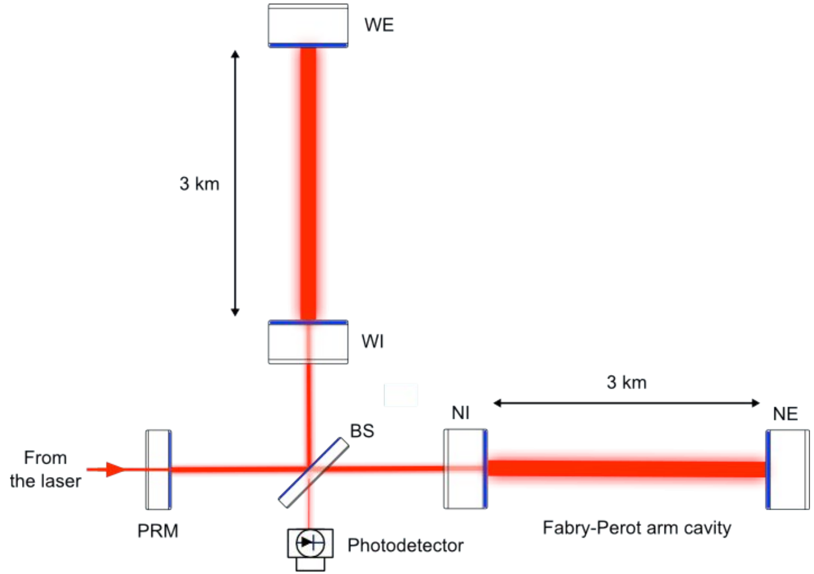
https://observing.docs.ligo.org/plan/

The GW Data

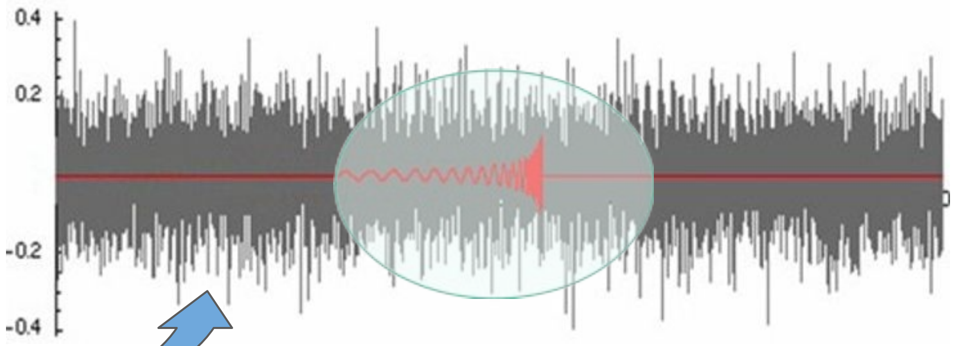


SCUOLA NORMALE SUPERIORE
EGO - Virgo

GW detector data



- Time series sequences... noisy time series with low amplitude GW signal buried in

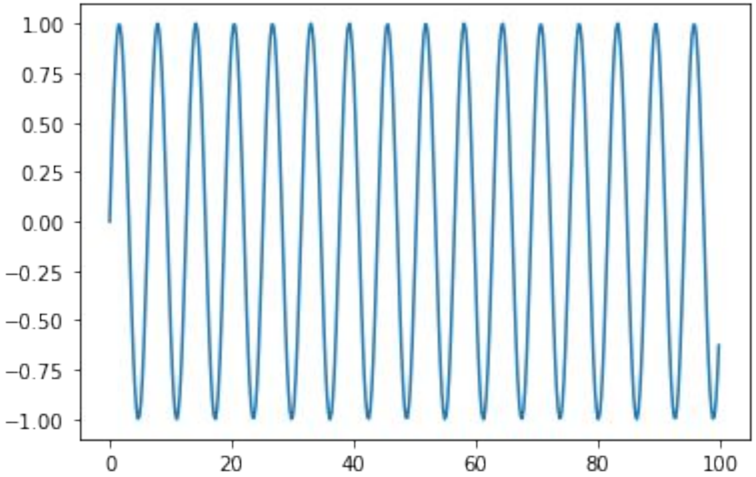


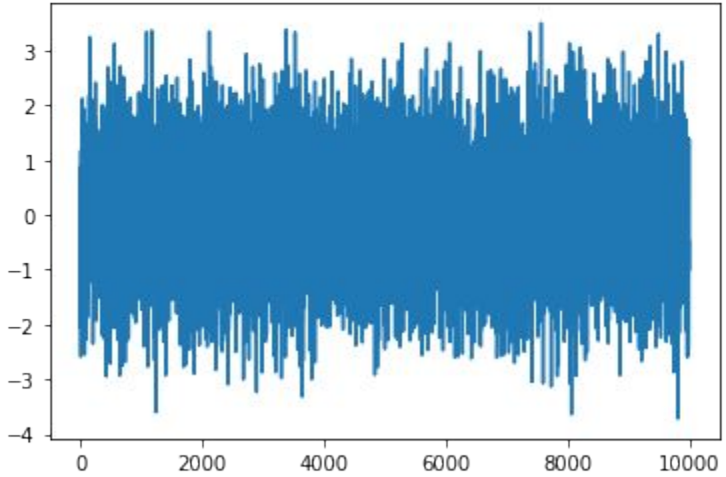
SCUOLA NORMALE SUPERIORE

EGO - Virgo

Time series

- A time series x[n] is a sequence of data points measuring a physical quantity at successive times spaced at uniform time intervals.
- We say that x[n] is a stationary process, if its statistical description does not depend on n.





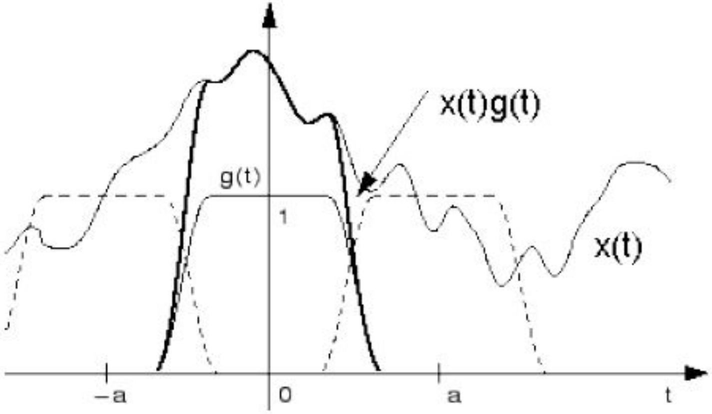
Data mapping, preserving the info

01	**Time-domain**	• Time-series at the output of the detector (be careful with the sampling theorem)
02	**Frequency-domain**	• Fourier transform ○ Useful for stationary data ○ Useful for persistent signals ○ It captures the global frequency information
03	**Time-frequency domain**	• Short Fourier Transform ○ Useful for not stationary data ○ Useful for transient signals
04	**Others**	• Wavelet decomposition (useful for multiresolution analysis) • Q-Transform (useful for transient) • Hough-transform (useful for lines)

Time-Frequency domain: STFT

The short time Fourier transform (STFT) function is simply the Fourier transform operating on a small section of the data. Here a moving window is applied to the signal and the fourier transform is applied to the signal within the window as the window is moved.

$$STFT\{x(t)\} = X(\tau, f) = \int_{-\infty}^{\infty} x(t) g(t-\tau) \exp(-2i\pi f t) dt$$

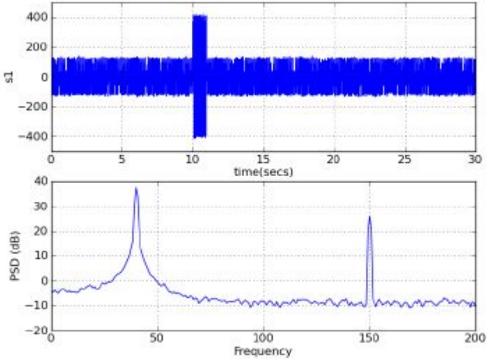


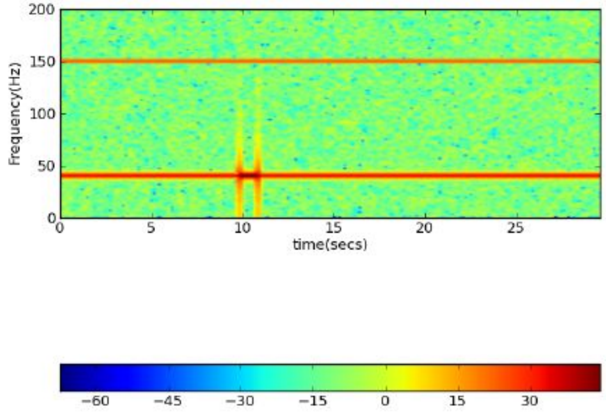
Spectrogram

To have easy access to the information of the STFT we can plot the spectrogram.
It is defined as

$$Spectrogram(\tau, f) = |X(\tau, f)|^2$$

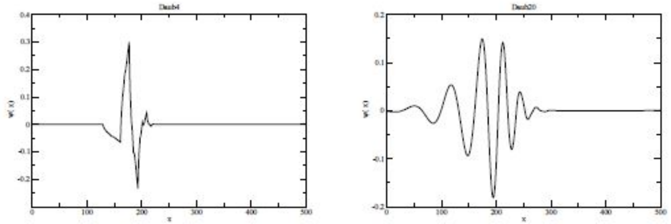
So we will have a bidimensional plot where on x-axis usually is plotted the time, on y-axis the frequency, while the color of the map is the the amplitude of a particular frequency at a particular time.





Wavelet decomposition of time series

The wavelet transform replaces the Fourier transform sinusoidal waves by a family generated by translations and dilations of a window called a wavelet.



$$Wf(a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{b}}\psi^{*}\left(\frac{t-a}{b}\right)\, dt$$

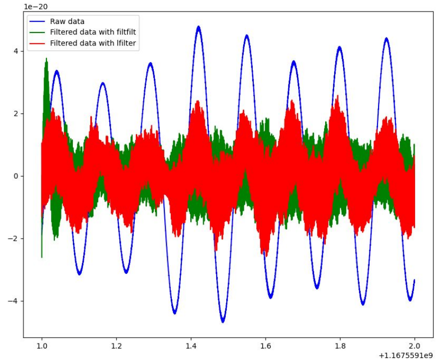
The scale of the wavelet is determined by the parameter **b**.

- When **b** is decreased, the wavelet appears more compressed, allowing it to capture high-frequency information.
- Increasing the value of **b** elongates the wavelet, enabling it to capture low-frequency information.

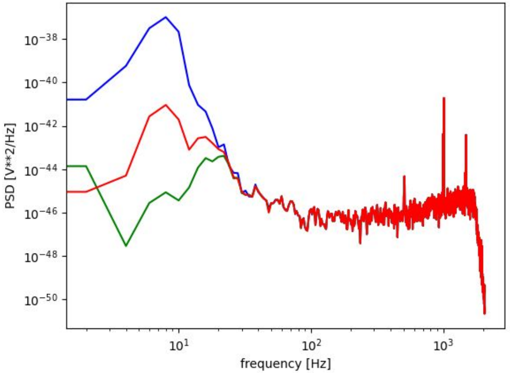
The location of the wavelet is determined by the parameter **a**.

- If we decrease the value of **a**, the wavelet will be shifted to the left, whereas an increase in **a** will shift it to the right.
- Note that the location of the wavelet is crucial because, unlike waves, wavelets are only non-zero within a short interval.

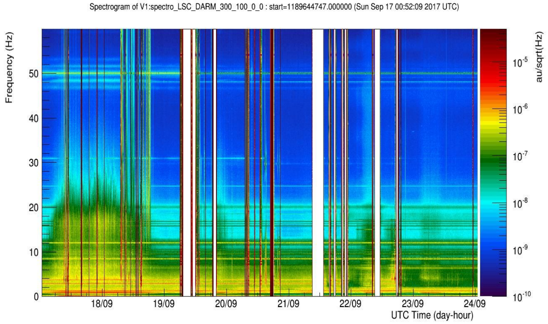
Data representations



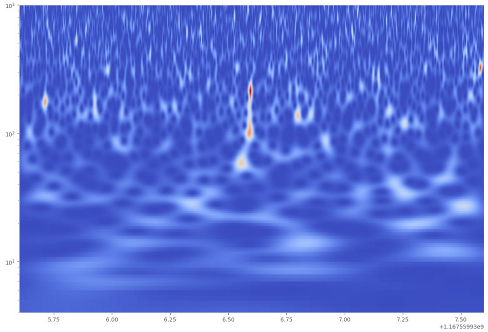
Time-domain



Frequency-domain

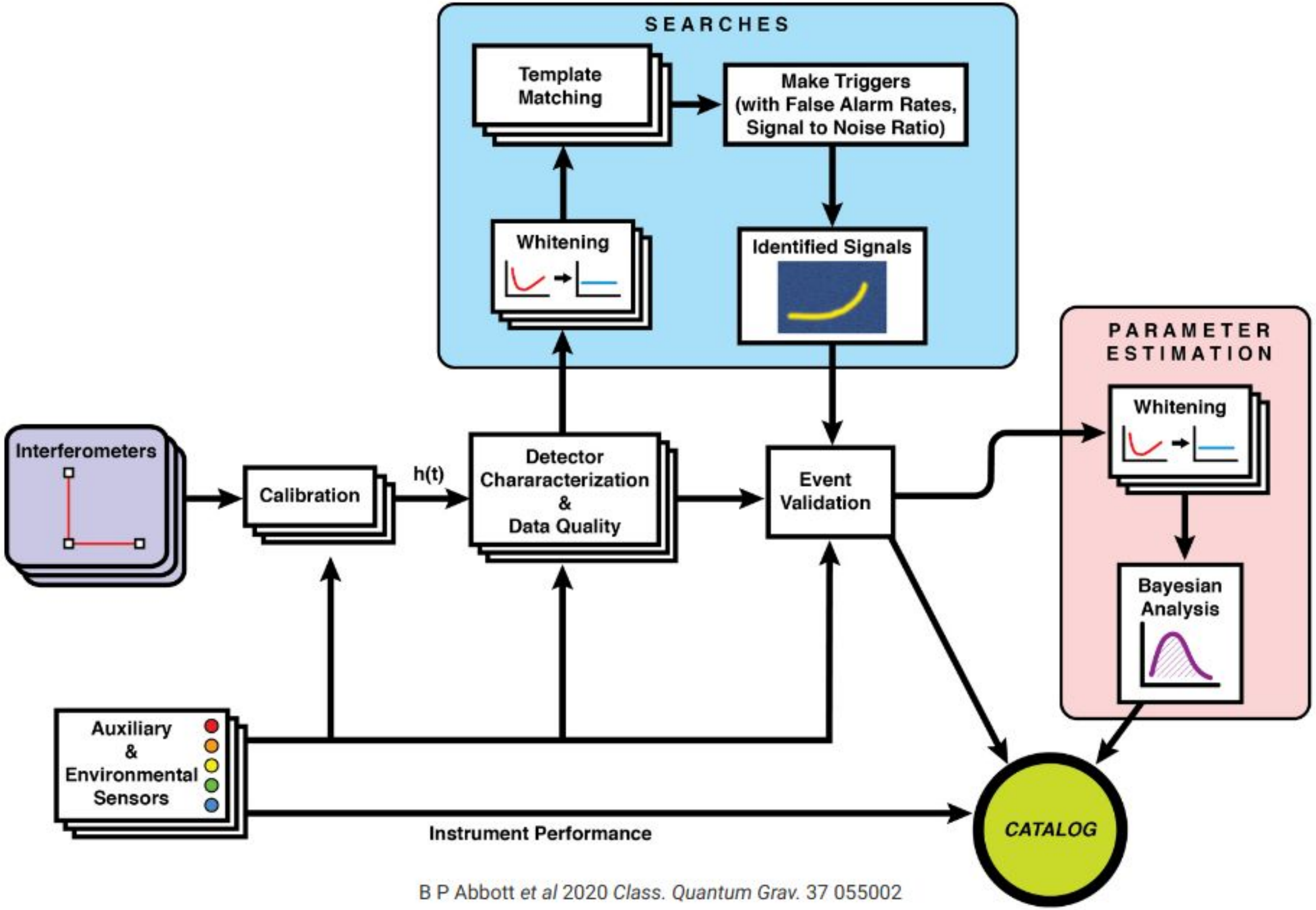


Time-frequency-domain



Wavelet-domain

The data analysis workflow

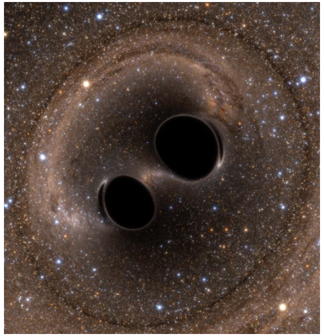


B P Abbott *et al* 2020 *Class. Quantum Grav.* 37 055002

GW astrophysical sources

Short□ long

Known □ unknown form



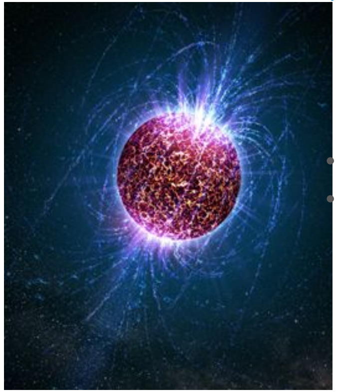
Coalescing Binary Systems CBC

Black hole – black hole (BBH)

Neutron star – neutron star (BNS)

BH-NS

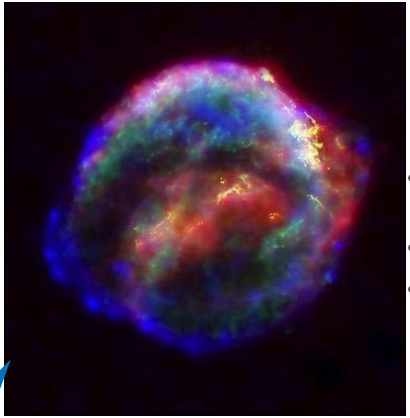
- Analytical waveform



Continuous Sources

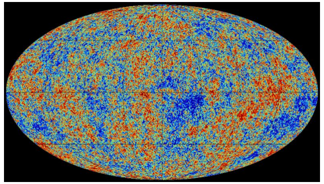
- Spinning neutron stars
- Monotone waveform

What else?



Transient Burst Sources

- Core Collapse Supernovae (CCSN)
- cosmic strings
- unmodeled waveform



Cosmic GW Background

- residue of the Big Bang,
- stochastic, incoherent background

Do we know their Waveforms?

GW Signal Detection and Matched Filter for known waveforms

- Defining the problem
- The Neyman Pearson Criteria
- The Matched Filter

extra pdf slides...:)

Optimal Filter is Matched Filter, if the noise is gaussian distributed

Maximizing the likelihood

Data

Template

$$\rho(t) = 4 \int_0^{\infty} \frac{\tilde{x}(f)\ \tilde{h}^*(f)}{S_n(f)}\ e^{2\pi i f t}\ df$$

Noise power spectral density

Look for maxima of $|\rho(t)|$ above some threshold ➔ trigger

SCUOLA NORMALE SUPERIORE

EGO - Virgo

How we detect transient signals: modeled search

- pyCBC (Usman et al, 2015)
- MBTA (Adams et al. 2015)
- gstlal-SVD (Cannon et al. 2012)

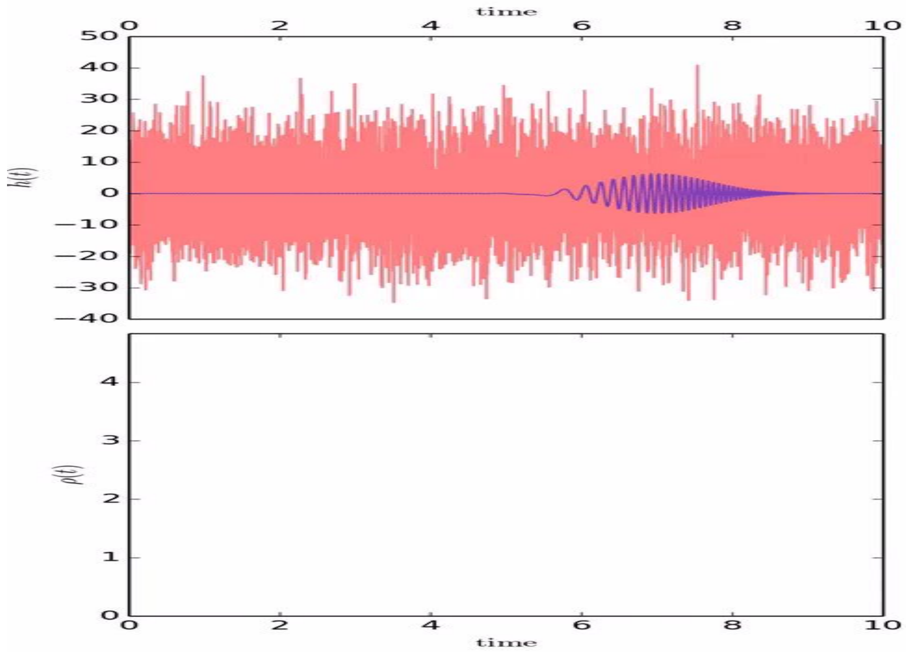
Matched-filter

Data

Template

$$\rho(t) = 4 \int_0^\infty \frac{\tilde{x}(f)\,\tilde{h}^*(f)}{S_n(f)}\, e^{2\pi i f t}\, df$$

Noise power spectral density

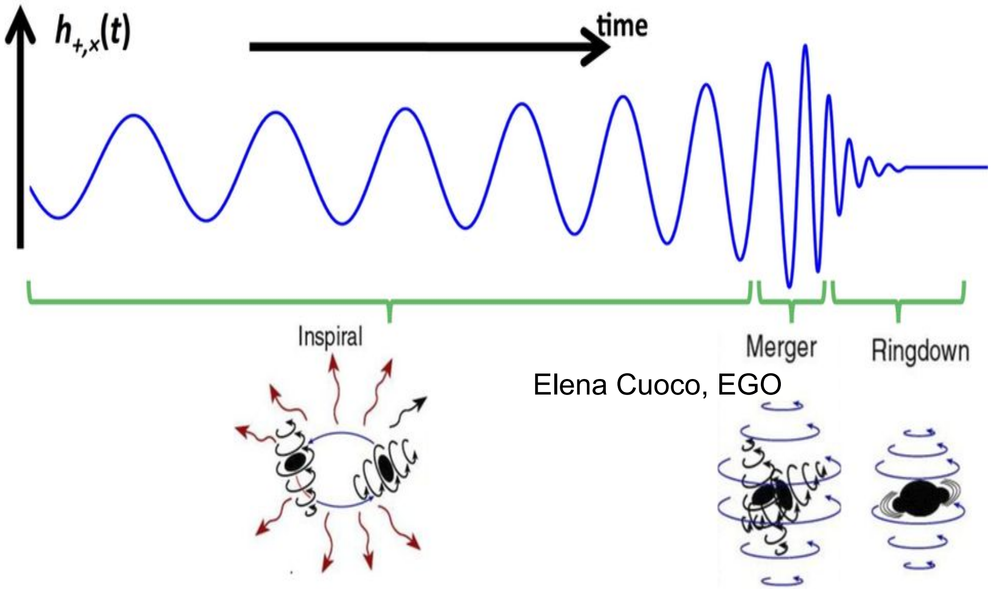


CBC template generation

$$h(f) = A(f)e^{i\Phi(f)}$$

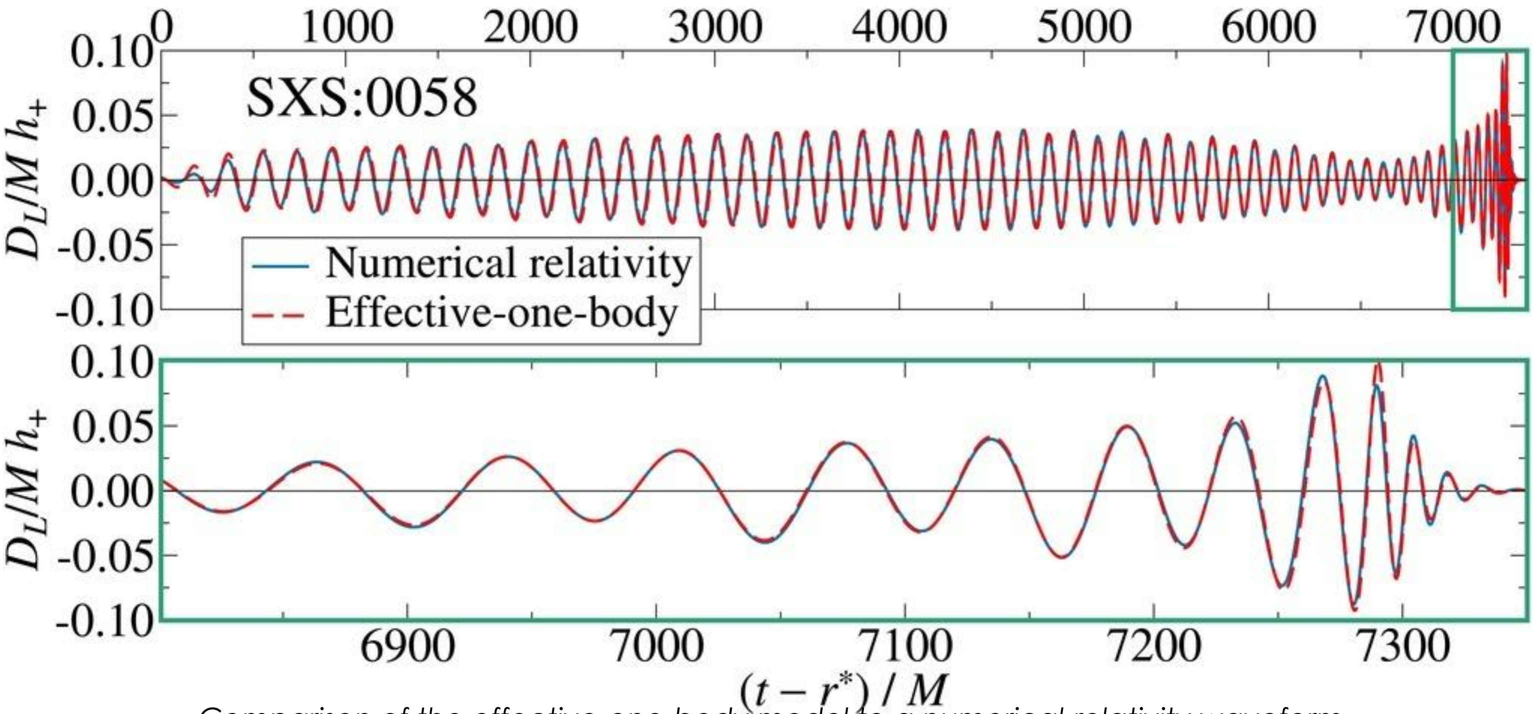
$$\Phi(f) = \sum_{k=1}^{7}(\varphi_k + \varphi_k^l \log(f))f^{(5-k)/3} + \sum_{i\neq k}\varphi_i f^i$$

$$\varphi_j \equiv \varphi_j(m_1, m_2, \vec{s}_1, \vec{s}_2)\ \forall j = k, i$$



Elena Cuoco, EGO

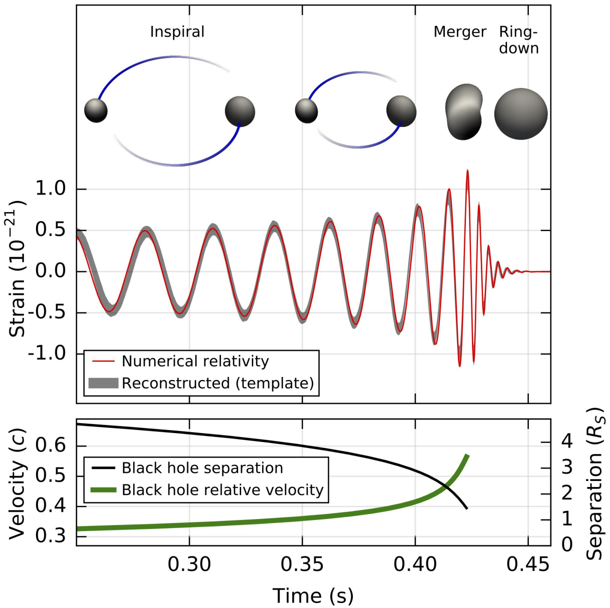
CBC template generation

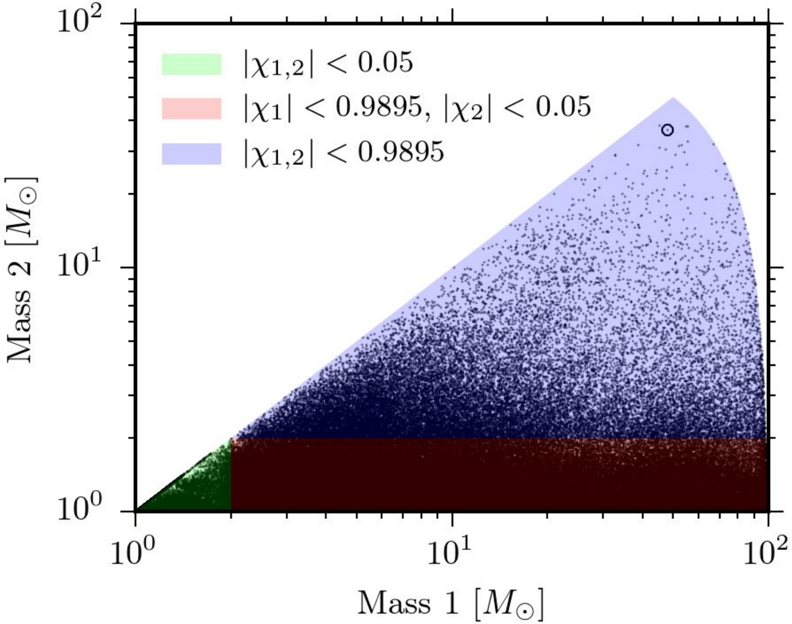


Comparison of the effective-one-body model to a numerical-relativity waveform of a precessing black-hole binary. © A. Taracchini/AEI

How many templates?

To cover in efficient way the parameters space, we build a templates bank requiring that the signal can be detected with a maximum loss of 3% of its SNR





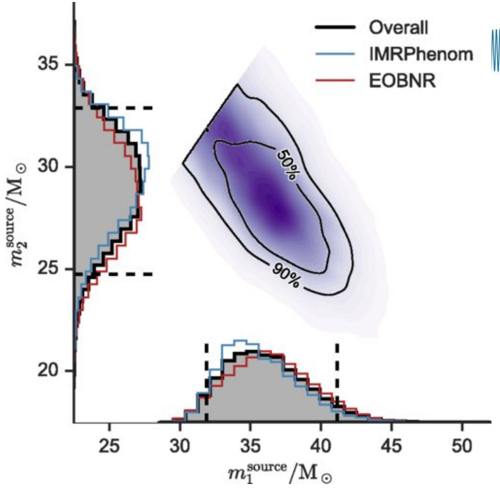
~250000 waveforms used for GW150914

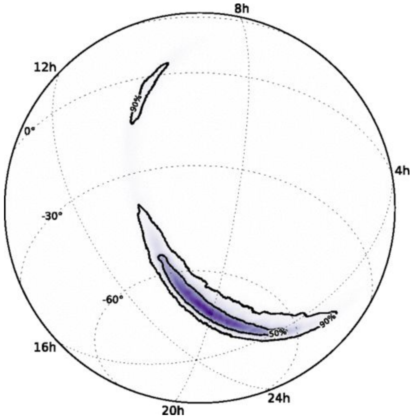
LVC Phys. Rev. X 6 (2016)

Parameter estimation

$$p(\theta|d, H) = \frac{p(\theta|H)p(d|\theta, H)}{p(d|H)}.$$

- *MCMC and Nested Sampling*
  - *MCMC Random steps are taken in parameter space, according to a proposal distribution, and accepted or rejected according to the Metropolis-Hastings algorithm.*
  - *Nested sampling can also compute evidences for model selection.*





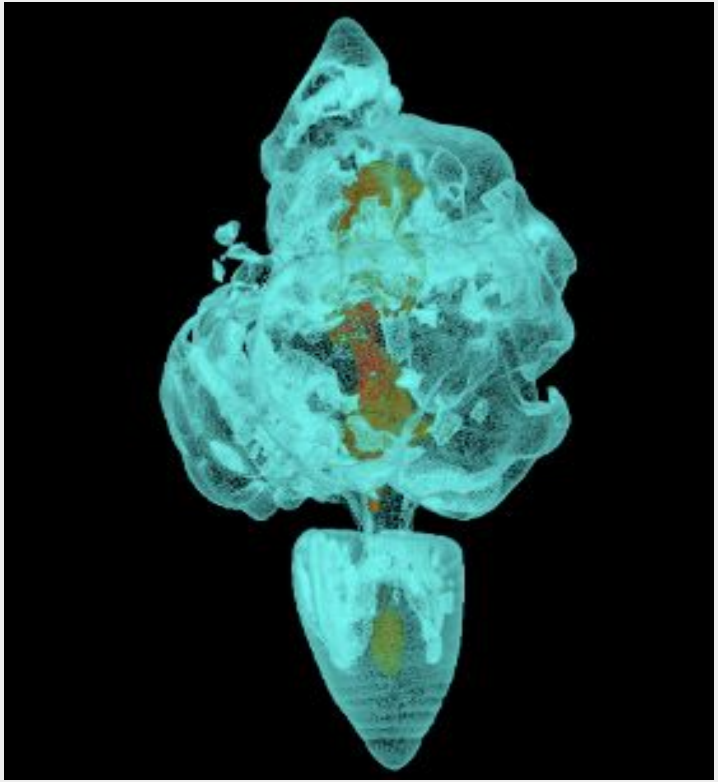
Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library

J. Veitch et al. Phys. Rev. D 91, 042003

LVC (PRL:116, 241102)

Searches for unmodeled signals

What we do for signals with unknown waveforms



Computer simulation of gravitational waves emitted by a supernova. Credit: J Powell / B Mueller

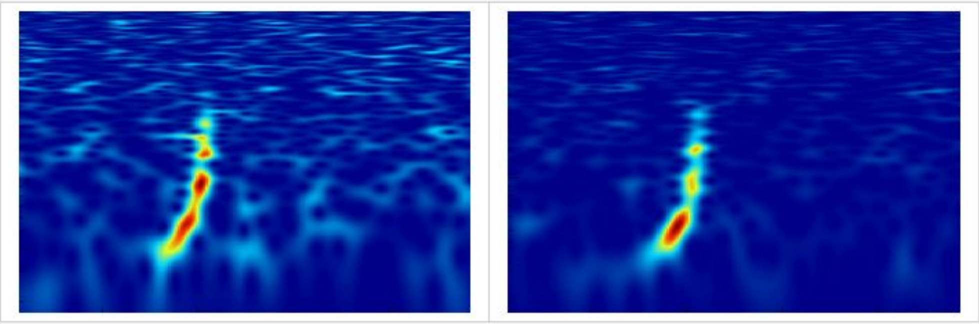
How we detect transient signals: un-modeled search

- *Strategy: look for excess power in single detector or coherent/coincident in network data*

- *Example cWB*

  (https://gwburst.gitlab.io/)
  - *Time-domain data preprocessed*
  - *Wavelet decompostion*
  - *Event reconstruction*

Coherent WaveBurst was used in the first direct detection of gravitational waves (GW150914) by LIGO and is used in the ongoing analyses on LIGO and Virgo data.



Time-Frequency maps of GW150914: Livingston data (left), Hanford data (right)
First screenshot of GW150914 event

Burst search

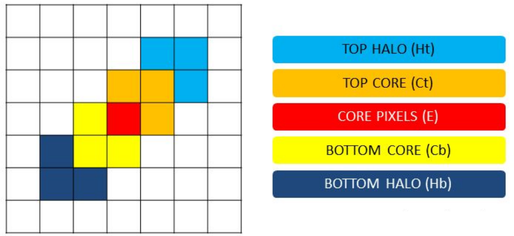
Phys. Rev. D 93, 042004 (2016)
Class.Quant.Grav.25:114029,2008

Coherent WaveBurst

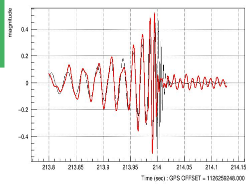
Excess power are selected from a set of wavelet time-frequency maps

Data from both detector are combined together

Triggers are analyzed coherently to estimate signal waveform, wave polarization, source location, using the constrained likelihood method



Selects the best fit waveform which corresponds to the maximum likelihood statistic over a 200000 sky positions



The event are ranked using a variable η_c

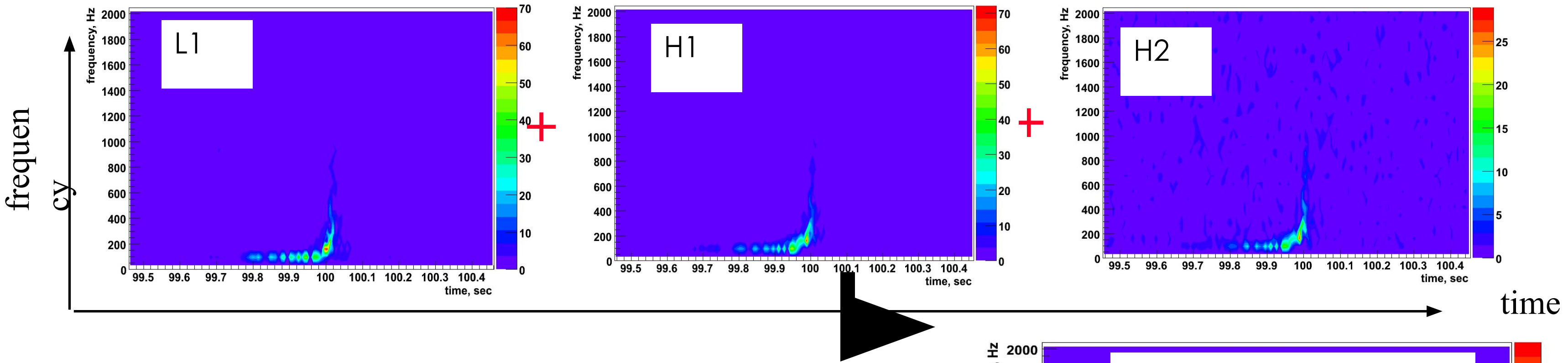
E_c ➔ Normalized coherent energy between the two detectors

E_n ➔ normalized noise energy derived by subtracting the reconstructed signal from the data

$$\eta_c = \sqrt{\frac{2E_c}{(1 + E_n/E_c)}}$$

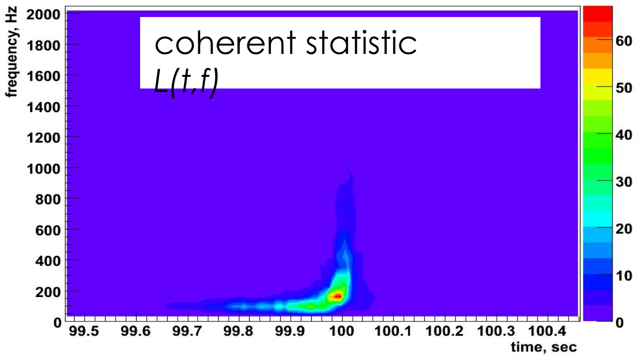
Coherent WaveBurst

- End-to-end multi-detector coherent pipeline
  - construct coherent statistics for detection and rejection of artifacts
  - performs search over the entire sky
  - estimates background with time shifts



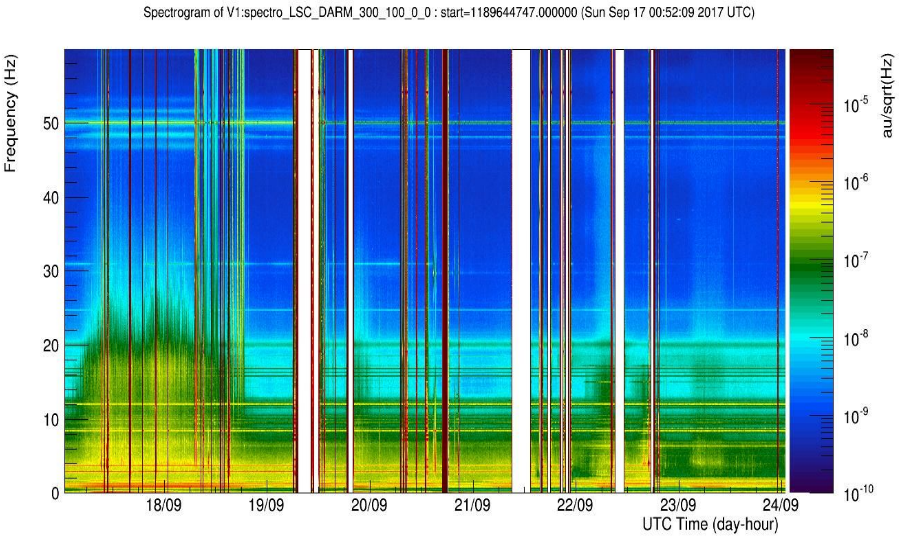
$$\xi_k = h_+ F_{+k} + h_\times F_{\times k}$$

$$L(t,f) = \max_{h_+ h_\times \theta\varphi} \sum_k \frac{x_k^2[t,f] - \left(x_k[t,f] - \xi_k[t,f]\right)^2}{\sigma_k^2(f)}$$



The Real Data

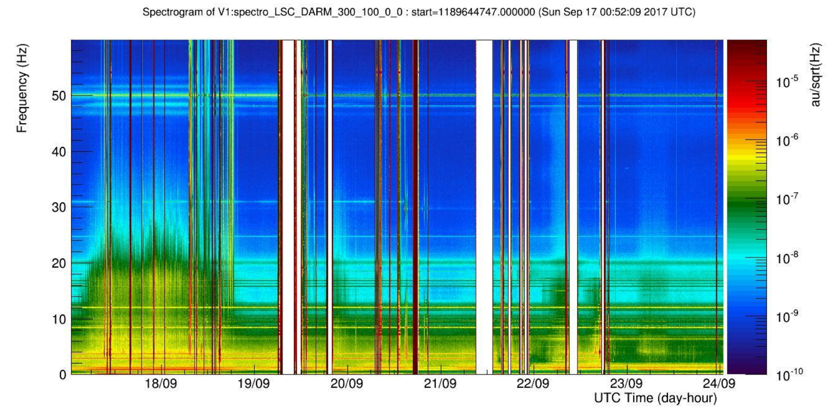
The noise it not at all ideal…

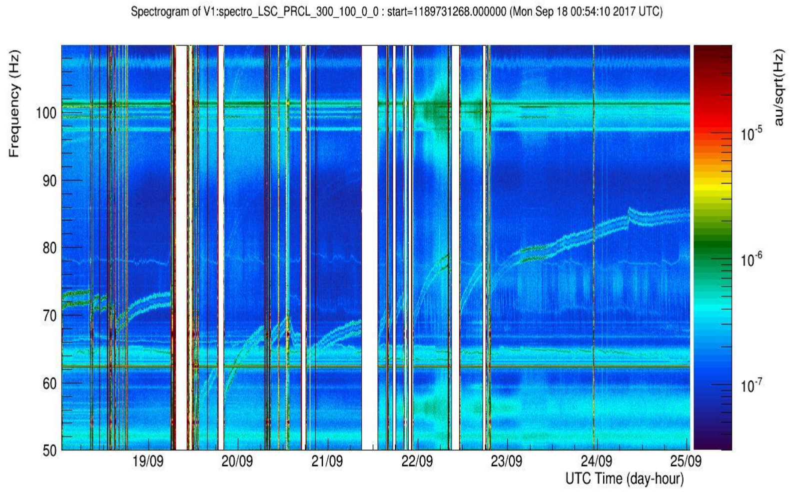


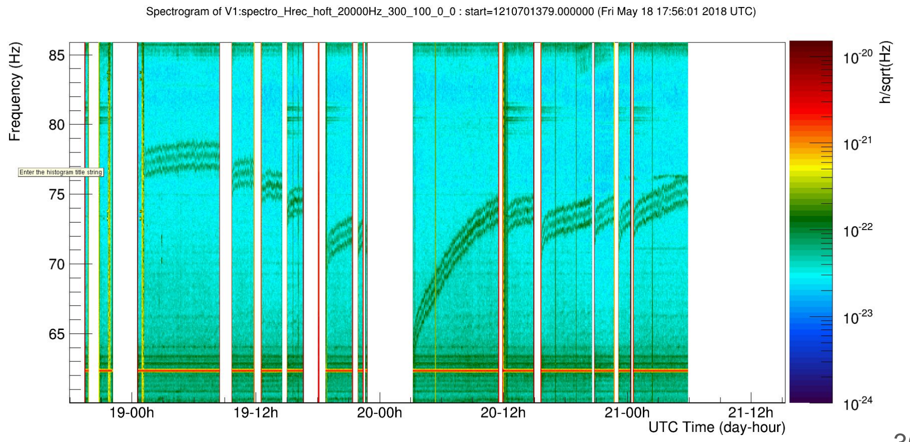
SCUOLA NORMALE SUPERIORE

EGO - Virgo

Not linear and not stationary noise

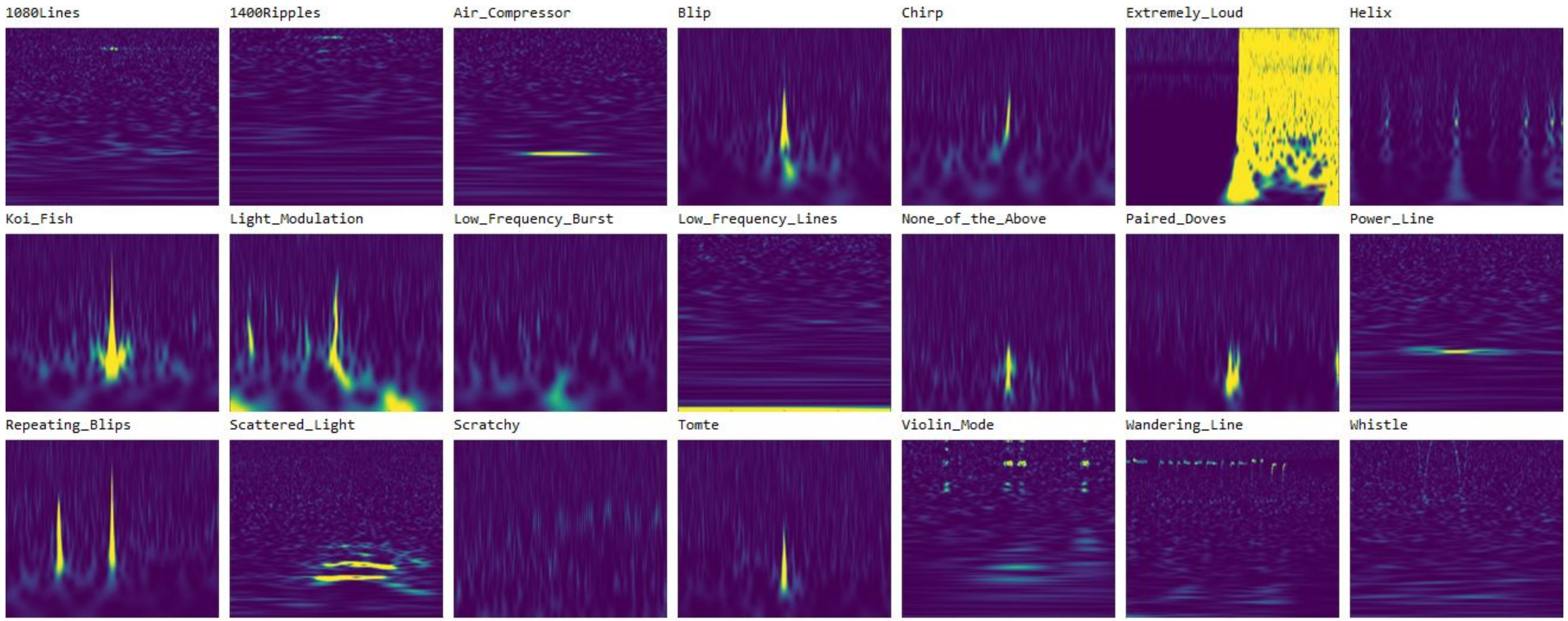






I. Fiori courtesy

Transient noise signals: Glitches



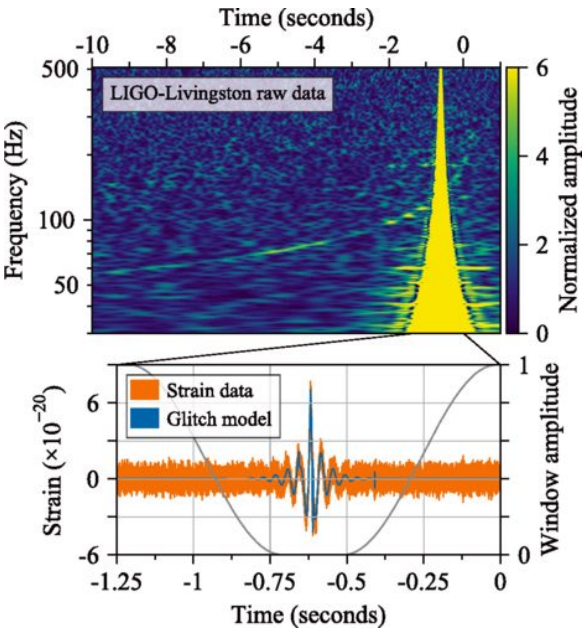
https://www.zooniverse.org/projects/zooniverse/gravity-spy

Gravity Spy, Zevin et al (2017)

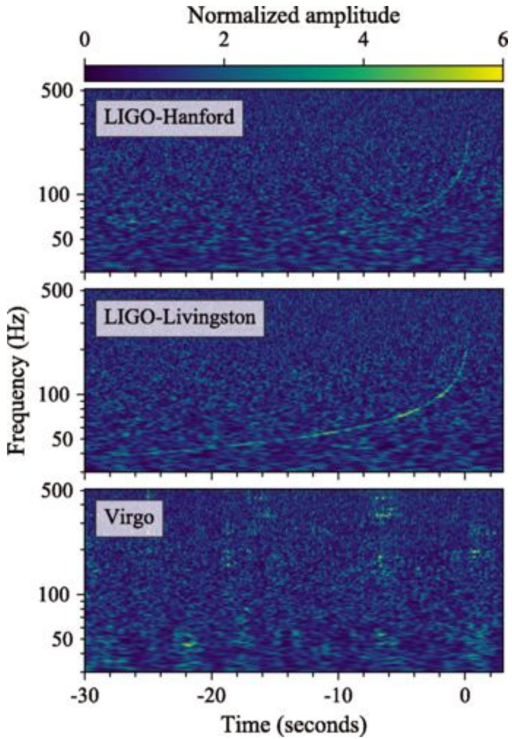
The importance of glitch analysis

Ligo Livingston

GW 170817



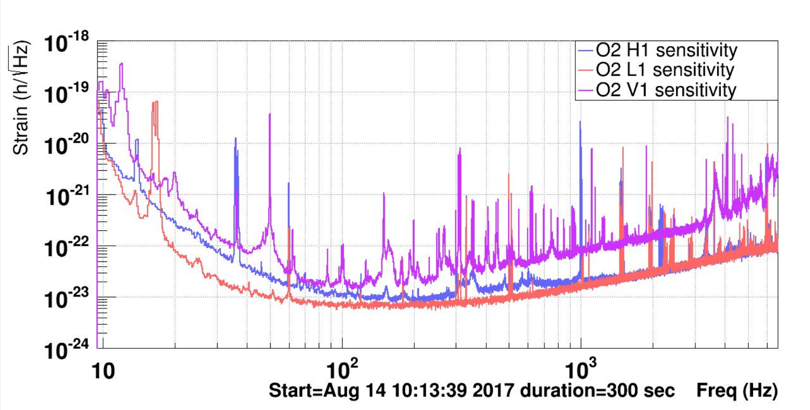
Glitch mitigation

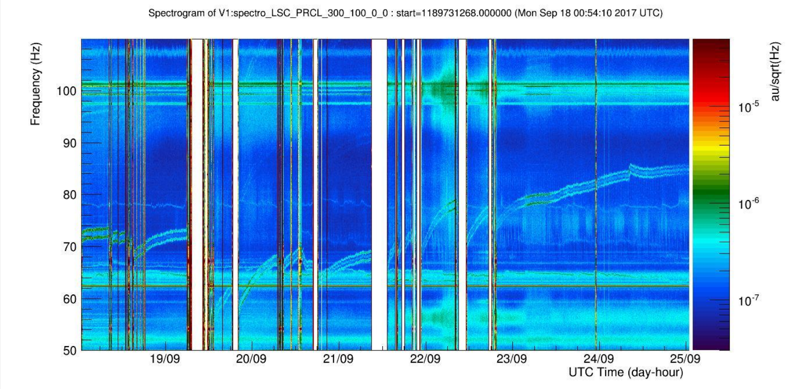


Abbott et al. (2017)

https://arxiv.org/pdf/2002.11668.pdf

Data preprocessing

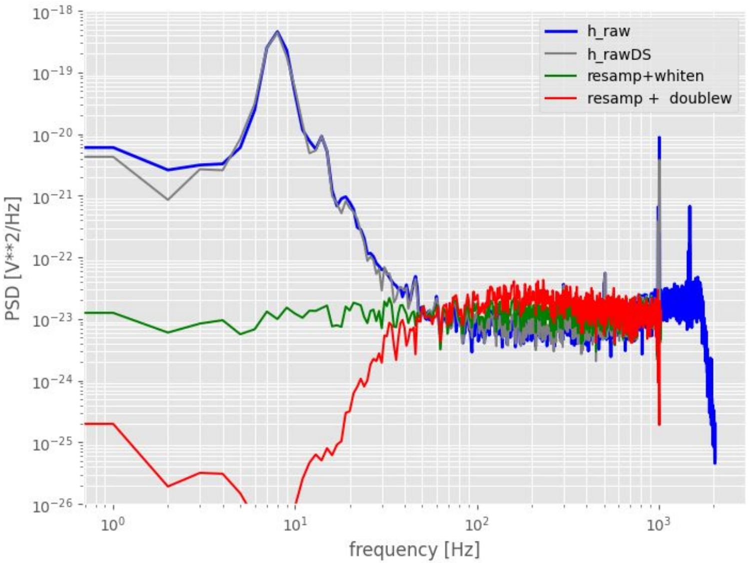




Whitening

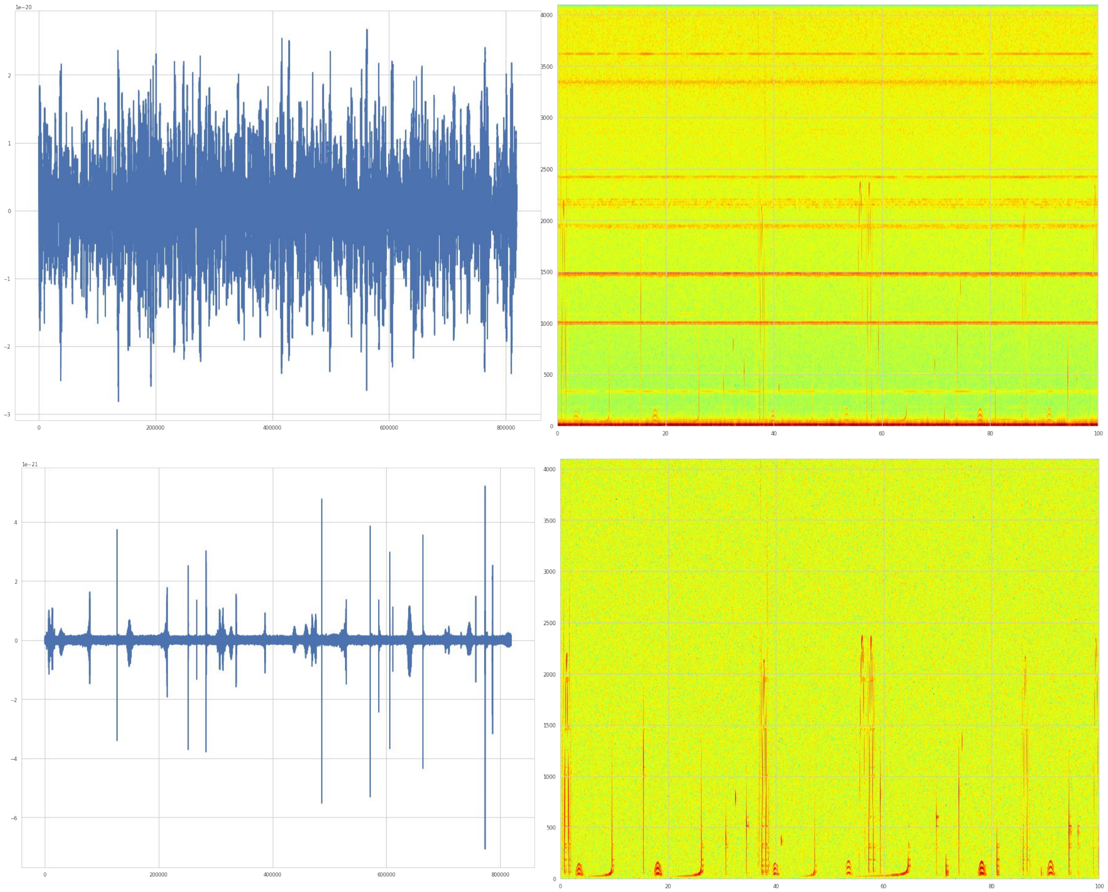
$$\rho(t) = 4\int_0^\infty \frac{\tilde{x}(f)\ \tilde{h}^*(f)}{S_n(f)}\ e^{2\pi i f t}\ df$$

Whitening



We can do in frequency domain estimating the PSD

Signals in whitened data



Not Whitened

Whitened

SCUOLA NORMALE SUPERIORE

EGO - Virgo

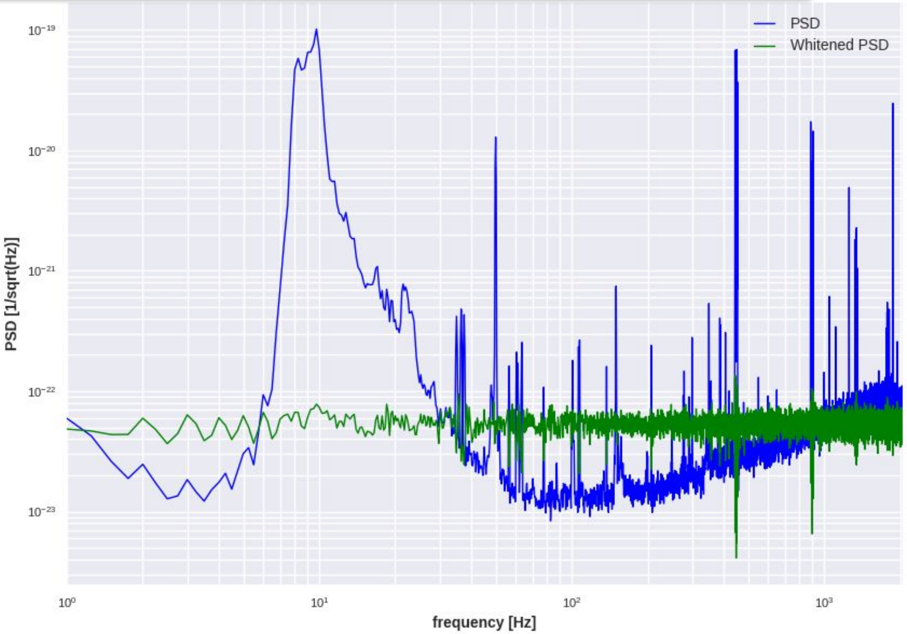
Whitening in time domain

We need parametric modeling

It can be useful for on-line application

It can be implemented for non stationary noise

It can catch the autocorrelation function to larger lags



AR parametric modeling

An AutoRegressive process is governed by this relation

$$x[n] = -\sum_{k=1}^{p} a[k]x[n-k] + w[n],$$

and its PSD for a process of order P is given by

$$P_{AR}(f) = \frac{\sigma^2}{|1 + \sum_{k=1}^{P} a_k \exp(-i2\pi kf)|^2}$$

Kay S 1988 Modern spectral estimation: Theory and Application Prentice Hall Englewood Cliffs

Advantages of AR modeling

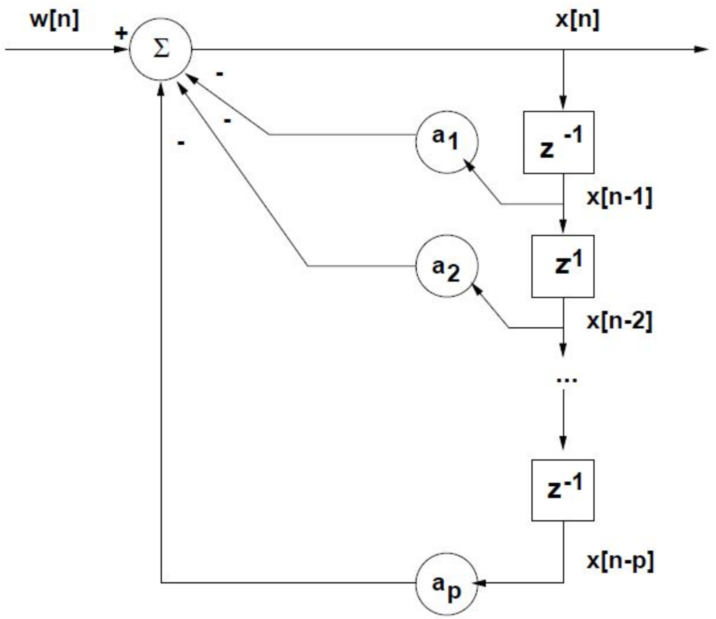
- Stable and causal filter: same solution of **linear predictor filter**

$$\hat{x}[n] = \sum_{k=1}^{P} w_k x[n-k] .$$

$$e[n] = x[n] - \hat{x}[n]$$

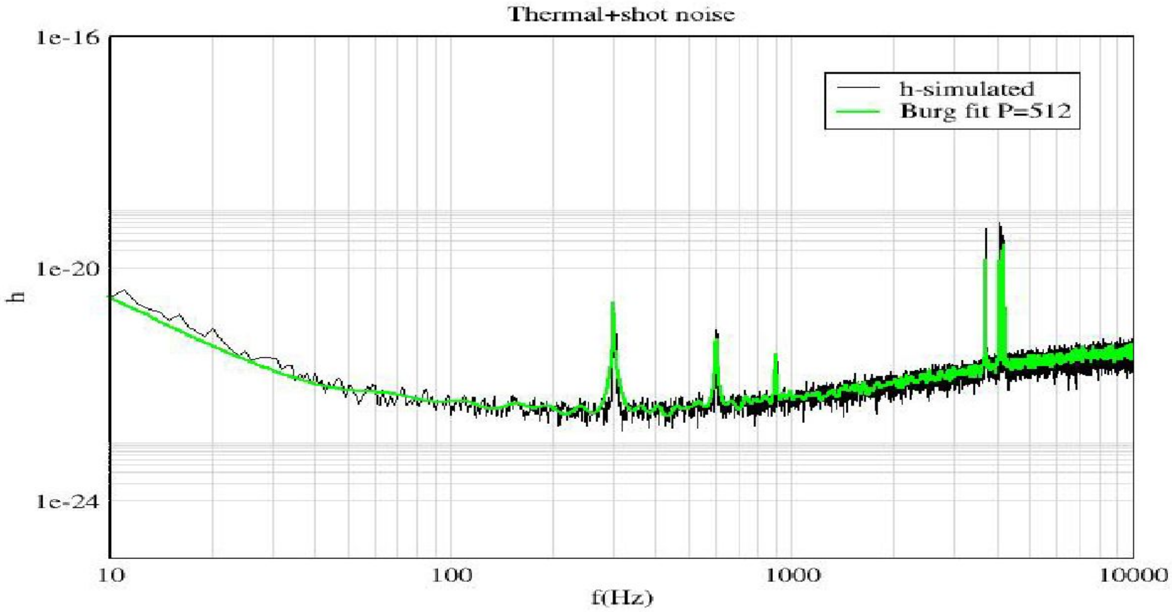
$$\varepsilon_{min} = r_{xx}[0] - \sum_{k=1}^{P} w_k r_{xx}[-k] ,$$

$$\begin{aligned} w_k &= -a_k \\ \varepsilon_{min} &= \sigma^2 \end{aligned}$$



Wiener-Hopf equations

PSD AR(P) Fit



Cuoco et al. Class.Quant.Grav. 18 (2001) 1727-1752 and
Cuoco et al.Phys.Rev.D64:122002,2001

Lattice Filter

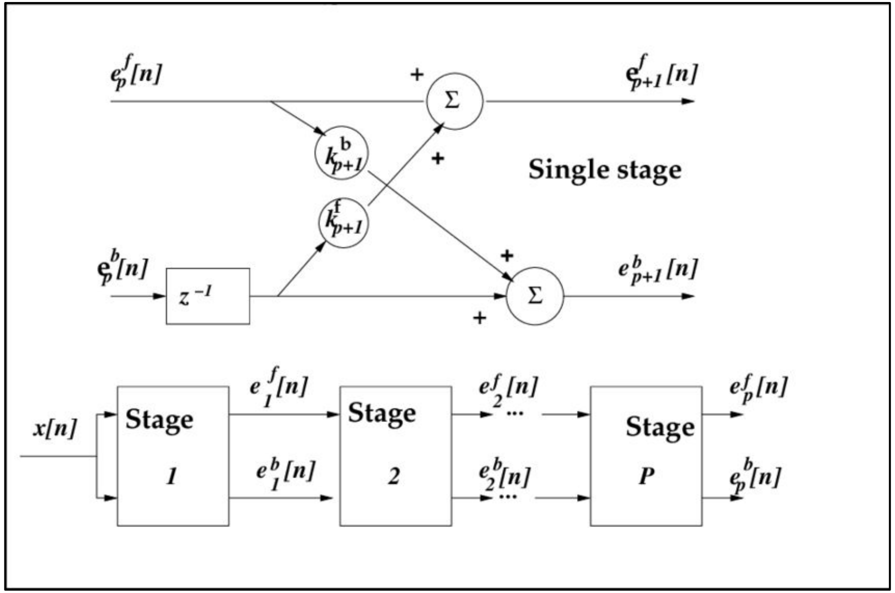
The Least Squares based methods build their cost function using all the information contained in the error function at each step, writing it as the sum of the error at each step up to the iteration n

Forgetting factor

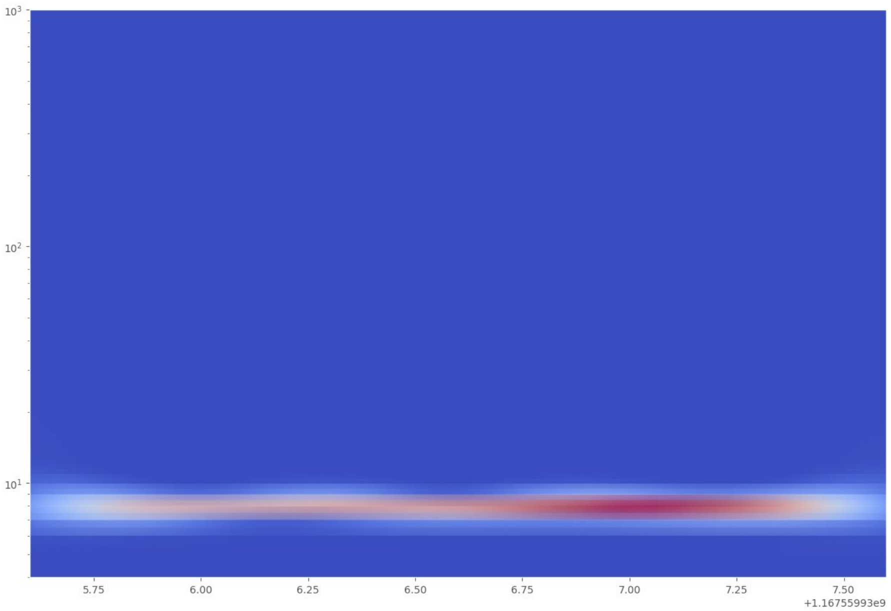
$$\epsilon[n] = \sum_{1}^{n} \lambda^{n-1} e^2(i|n)$$

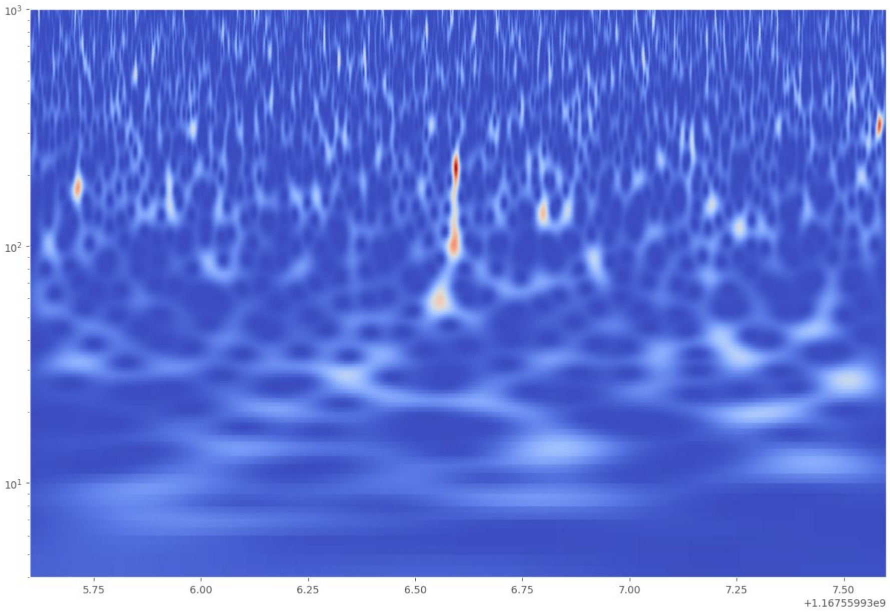
$$e(i|n) = d[i] - \sum_{k=1}^{N} x_{i-k} w_k[n],$$

Desired signal



The effect of whitening





Why artificial Intelligence for GW data?

- Our data: a lot of noise and few GW signals (soon will be many)
- Low SNR signals (overlapping signals)
- Many transient noise disturbances (glitches)
- Not stationary/not linear noise (strange noise coupling)
- Many monitoring auxiliary channels ("big" data)
- Computational and timing efficiency (Fast alert system)

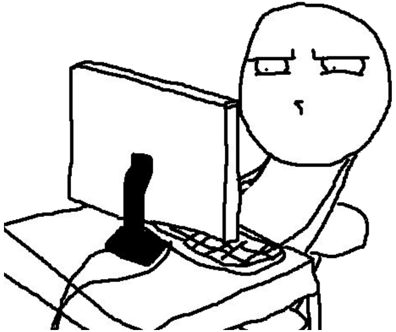
How Machine Learning can help

Data conditioning

- Identify Non linear noise coupling
- Use Deep Learning to remove noise
- Extract useful features to clean data

Signal Detection/Classification/PE

- A lot of fake signals due to noise
- Fast alert system
- Manage parameter estimation



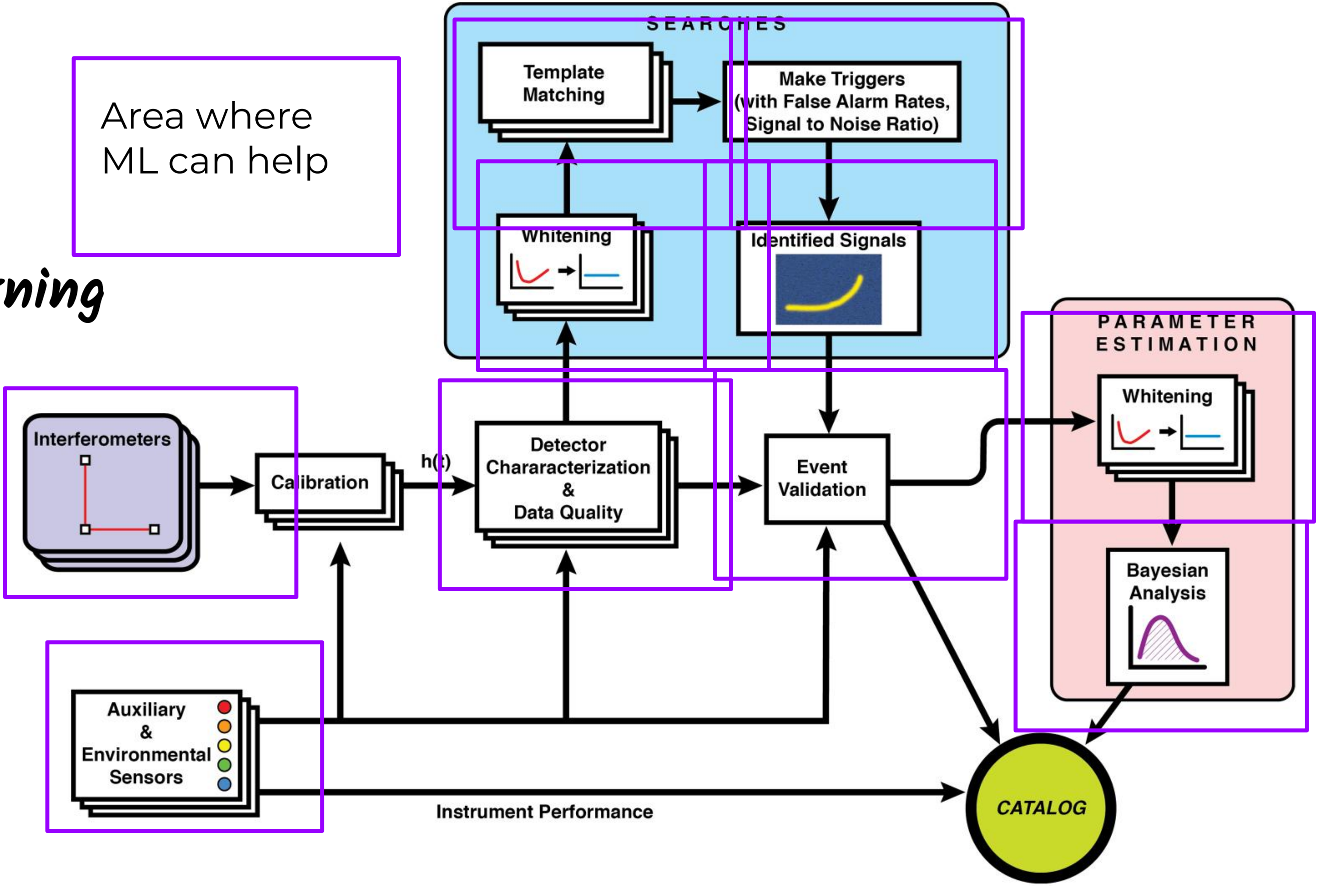
SCUOLA NORMALE SUPERIORE

EGO - Virgo

The data analysis workflow and ML

Area where ML can help

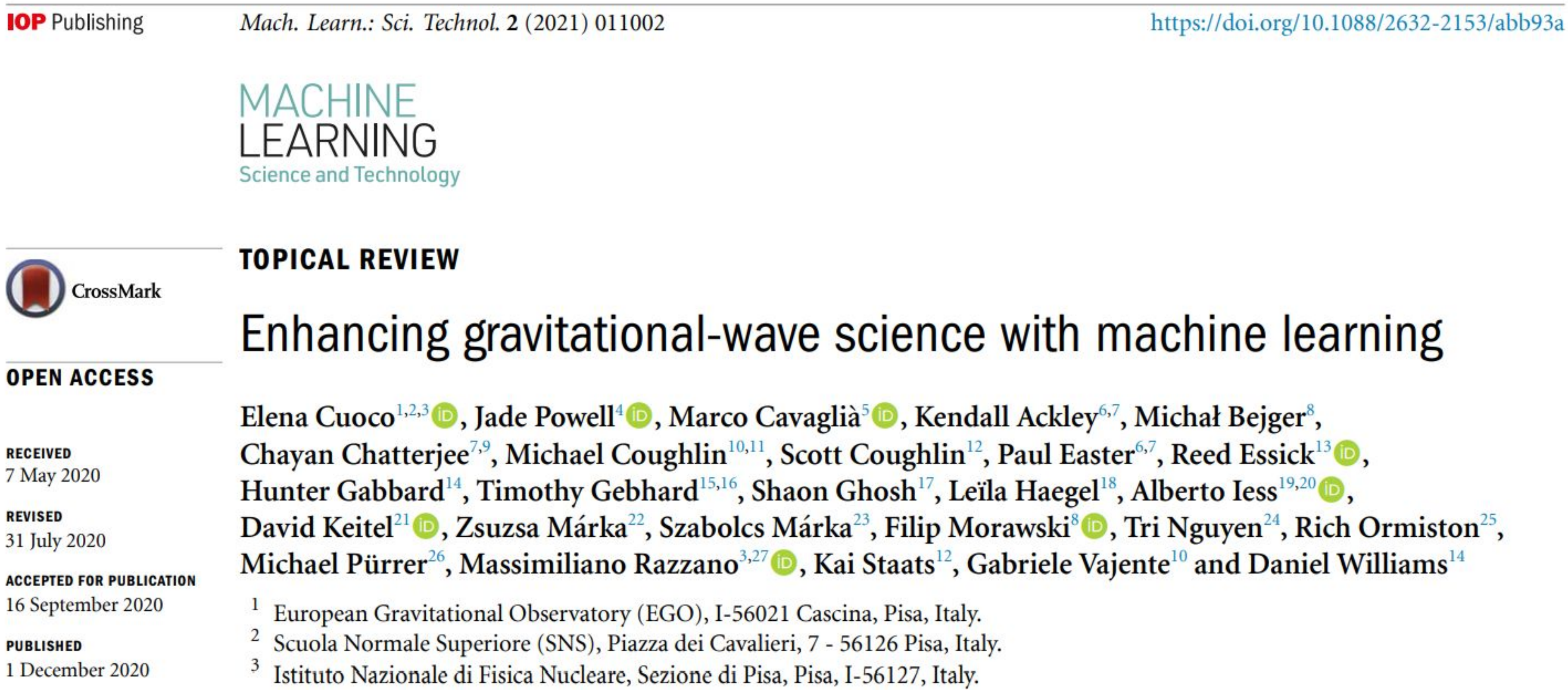
Machine Learning everywhere



SCUOLA NORMALE SUPERIORE

EGO - Virgo

An outdated overview


IOP Publishing Mach. Learn.: Sci. Technol. 2 (2021) 011002 https://doi.org/10.1088/2632-2153/abb93a

MACHINE LEARNING
Science and Technology

CrossMark

OPEN ACCESS

RECEIVED
7 May 2020

REVISED
31 July 2020

ACCEPTED FOR PUBLICATION
16 September 2020

PUBLISHED
1 December 2020

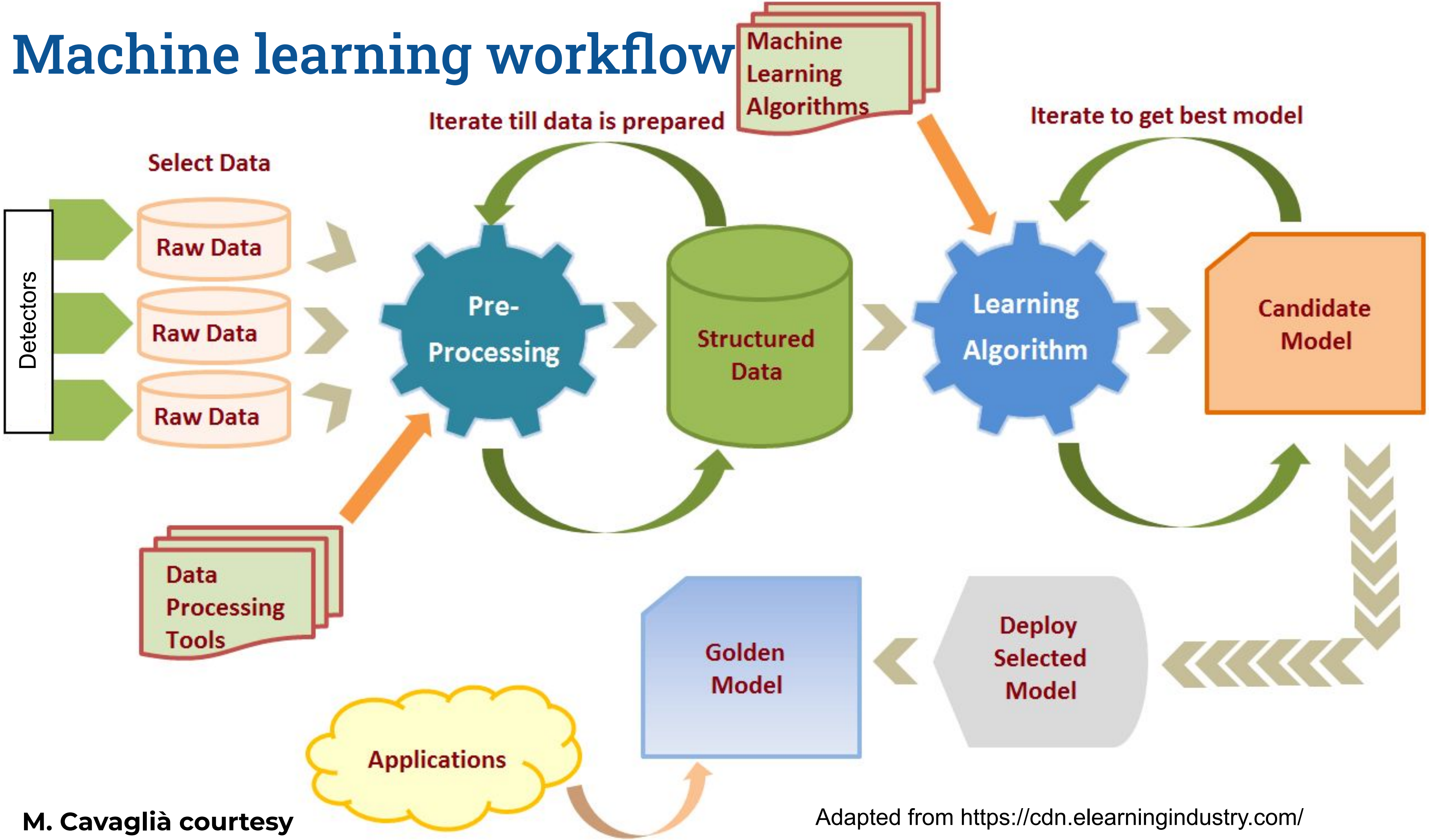
TOPICAL REVIEW

Enhancing gravitational-wave science with machine learning

Elena Cuoco[1,2,3], Jade Powell[4], Marco Cavaglià[5], Kendall Ackley[6,7], Michał Bejger[8], Chayan Chatterjee[7,9], Michael Coughlin[10,11], Scott Coughlin[12], Paul Easter[6,7], Reed Essick[13], Hunter Gabbard[14], Timothy Gebhard[15,16], Shaon Ghosh[17], Leïla Haegel[18], Alberto Iess[19,20], David Keitel[21], Zsuzsa Márka[22], Szabolcs Márka[23], Filip Morawski[8], Tri Nguyen[24], Rich Ormiston[25], Michael Pürrer[26], Massimiliano Razzano[3,27], Kai Staats[12], Gabriele Vajente[10] and Daniel Williams[14]

[1] European Gravitational Observatory (EGO), I-56021 Cascina, Pisa, Italy.
[2] Scuola Normale Superiore (SNS), Piazza dei Cavalieri, 7 - 56126 Pisa, Italy.
[3] Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Pisa, I-56127, Italy.

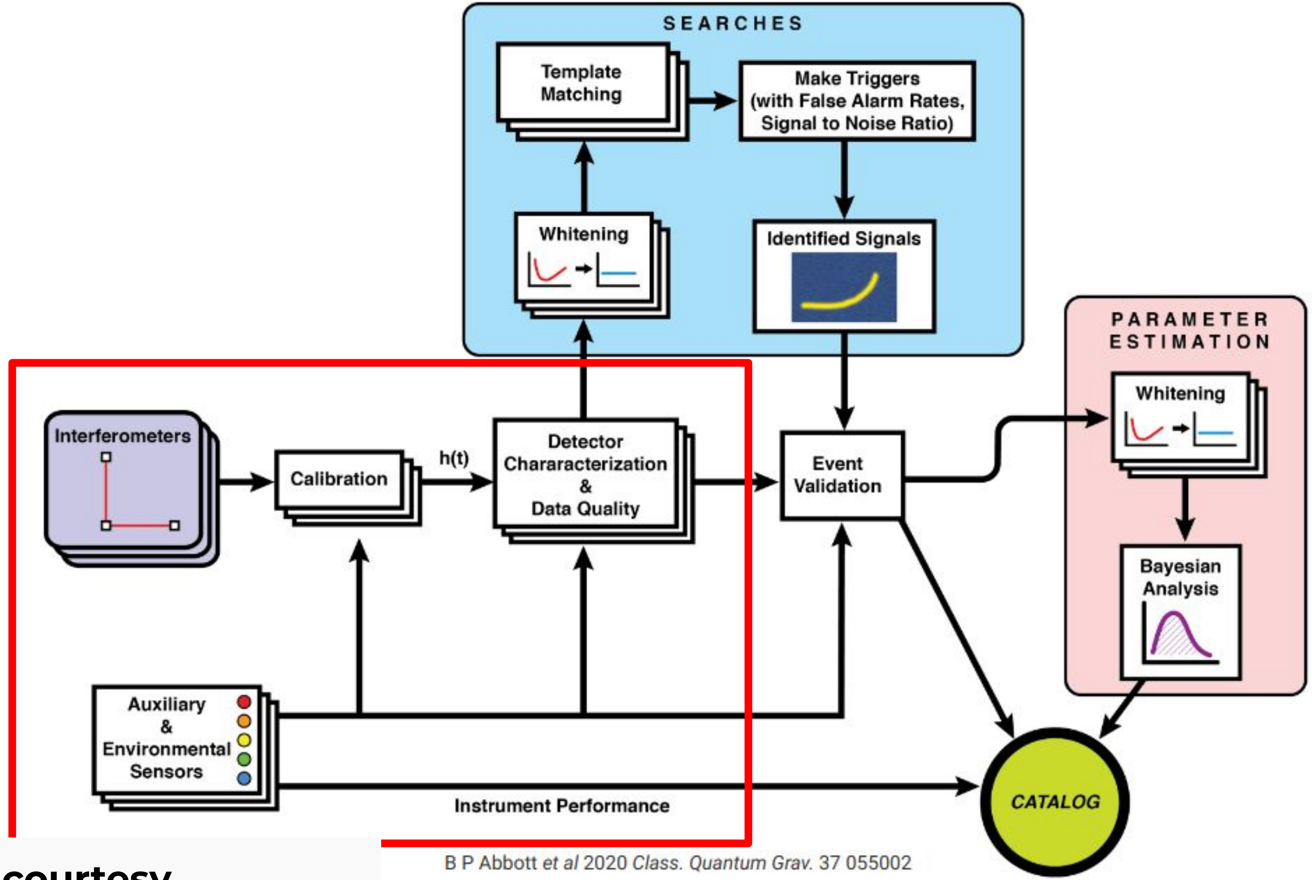
Machine learning workflow



M. Cavaglià courtesy

Adapted from https://cdn.elearningindustry.com/

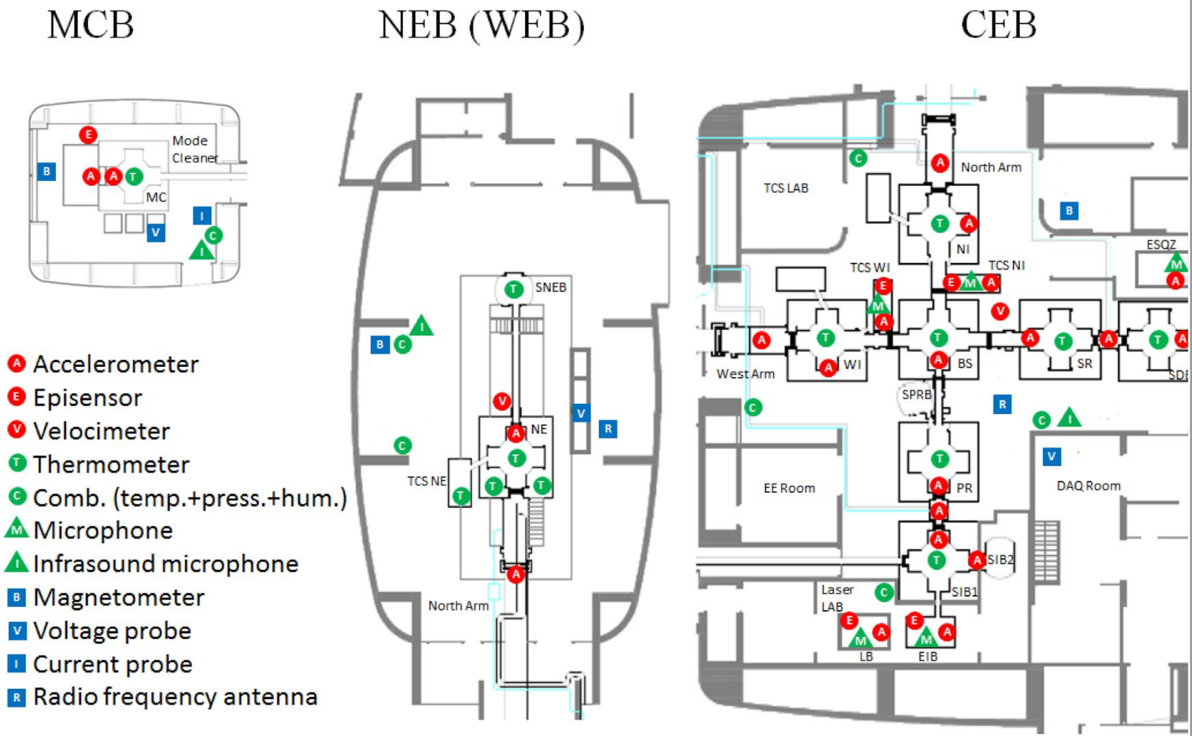
ML for GW data quality



B P Abbott *et al* 2020 *Class. Quantum Grav.* 37 055002

M. Cavaglià courtesy

Algorithms for GW Data Quality improvement



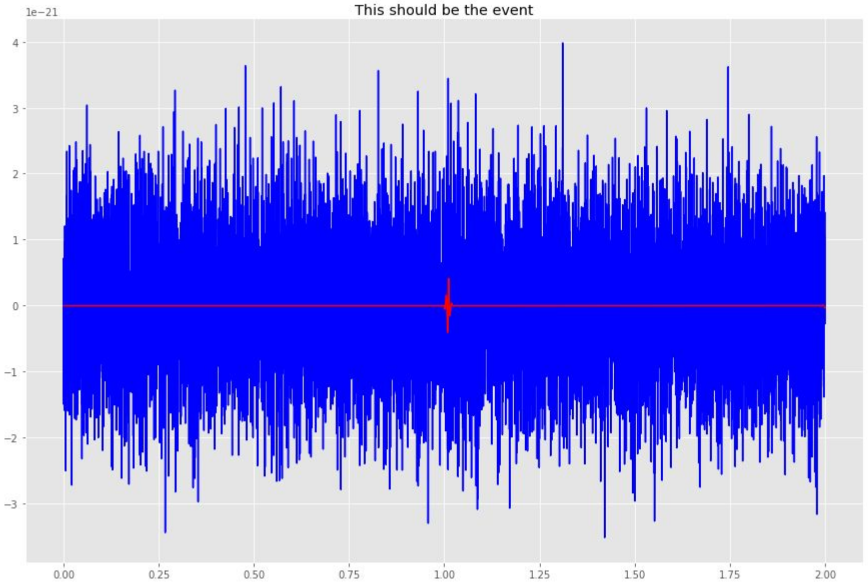
- Sensors continuously monitor the behavior of the detectors and their environment.
- Sensor data are used to characterize noise that may negatively impact searches and signal estimation.
- Information is in the form of time series.
- Invalid data due to detector malfunctions, calibration errors, data acquisition are removed from analyses.
- Flags are created according to different levels of data quality.

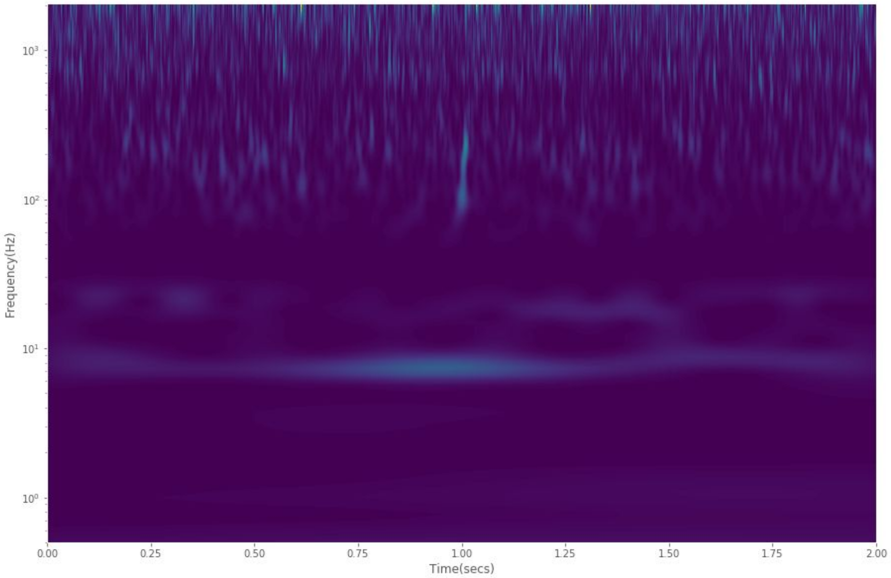
Galaxies **2020**, *8*(4), 82; **https://doi.org/10.3390/galaxies8040082**

M. Cavaglià courtesy

How to deal with data: Time series or Images?

- Pre-processing analysis (whitening, band pass filtering)
- Change of domain space: Time-Frequency projections





SCUOLA NORMALE SUPERIORE
EGO - Virgo

Noise transient classification, time series based

OPEN ACCESS
IOP Publishing
Classical and Quantum Gravity
Class. Quantum Grav. 32 (2015) 215012 (20pp)
doi:10.1088/0264-9381/32/21/215012

Classification methods for noise transients in advanced gravitational-wave detectors

Jade Powell[1], Daniele Trifirò[2,3], Elena Cuoco[4,5],
Ik Siong Heng[1] and Marco Cavaglià[3]

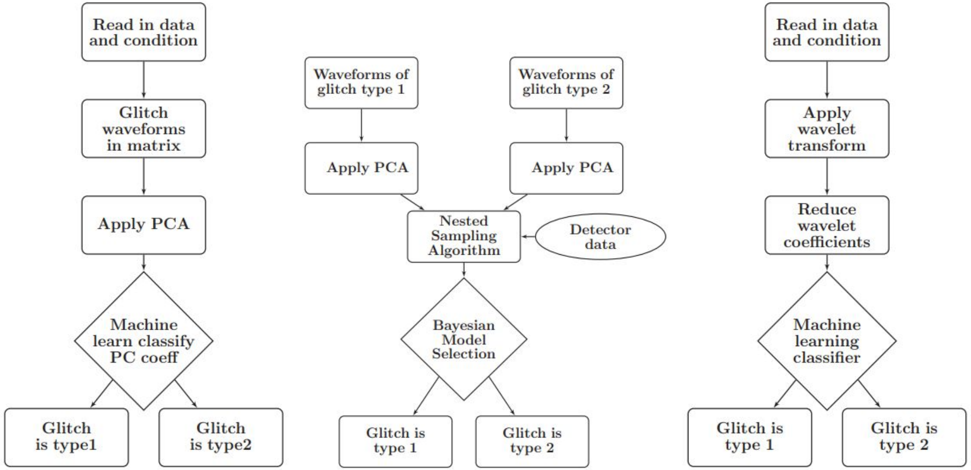
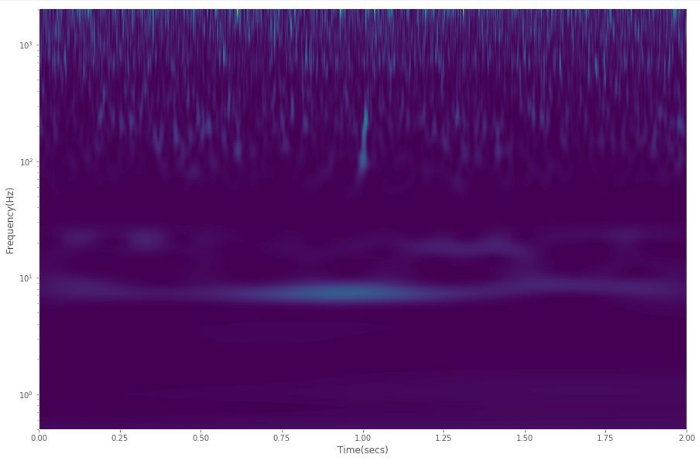


Image-based classification

Simulated and real data

- Transient Noise classification and Images as input data

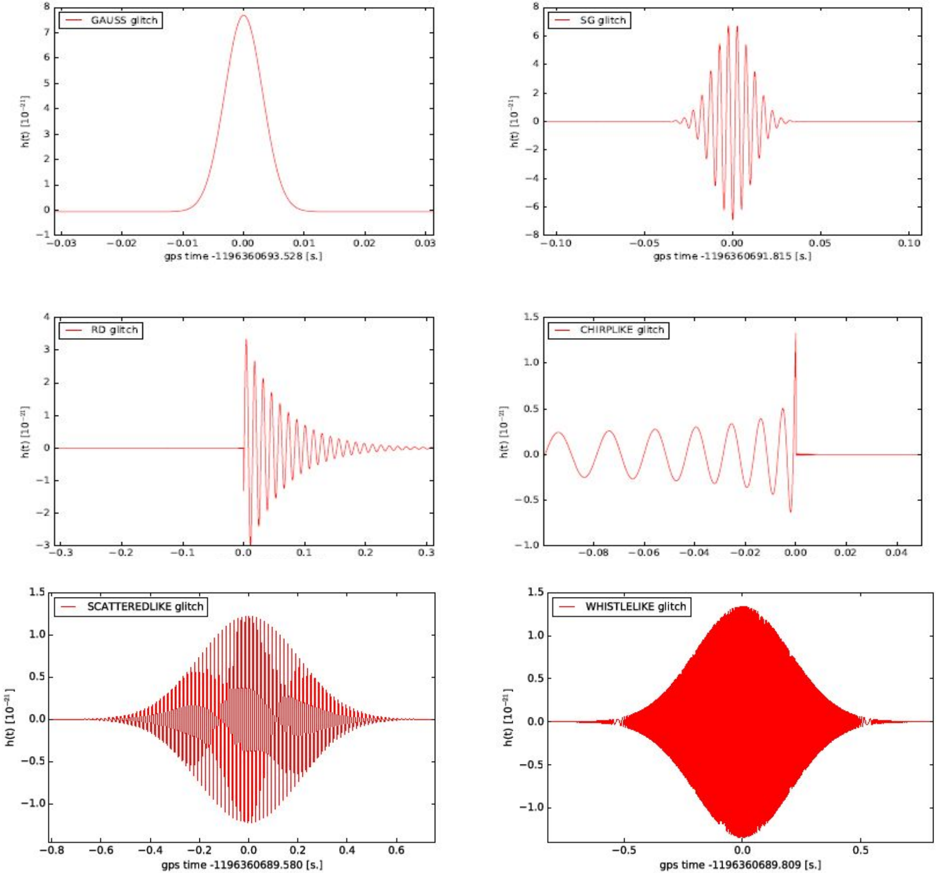


2018

SCUOLA NORMALE SUPERIORE

EGO - Virgo

How we started: Data simulation (transient signal families + Detector colored Noise)



Waveform
Gaussian
Sine-Gaussian
Ring-Down
Chirp-like
Scattered-like
Whistle-like
NOISE (random)

To show the glitch time-series here we don't show the noise contribution

Razzano M., Cuoco E. CQG-104381.R3

Building the images

Spectrogram for each image

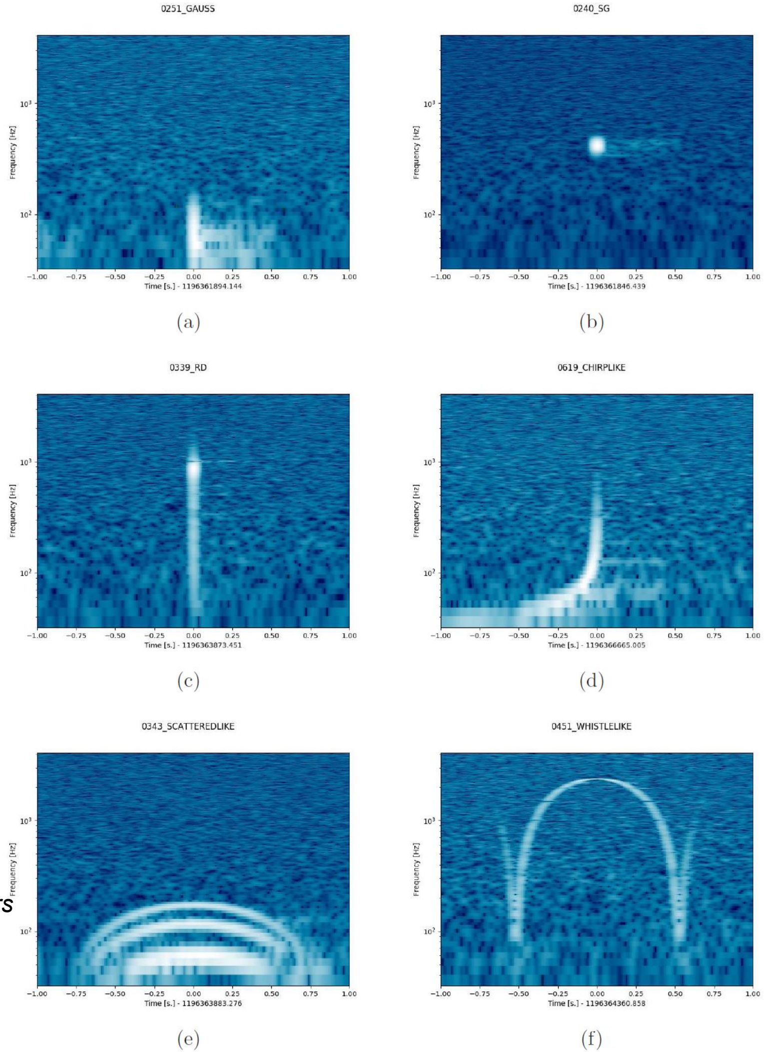
2-seconds time window to highlight features in long glitches

Data is whitened

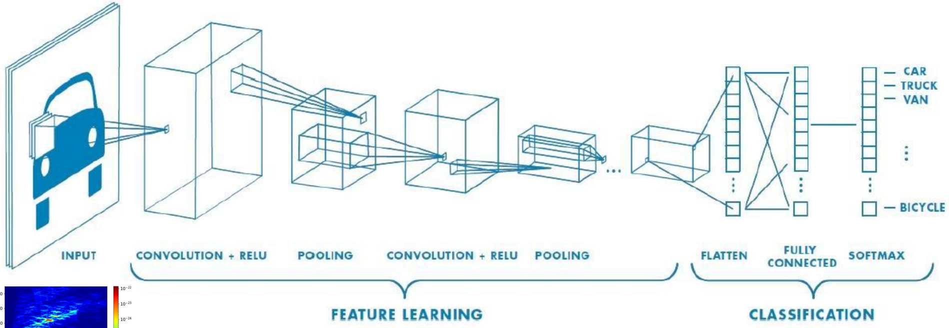
Optional contrast stretch

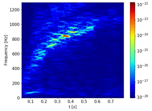
Simulations now available on FigShare

Razzano, Massimiliano; Cuoco, Elena (2018): Simulated image data for testing machine learning classification of noise transients in gravitational wave detectors (Razzano & Cuoco 2018). figshare. Collection. https://doi.org/10.6084/m9.figshare.c.4254017.v1

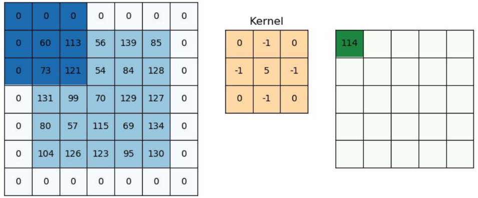


Deep learning. Convolutional Neural Network





Spectrogram images

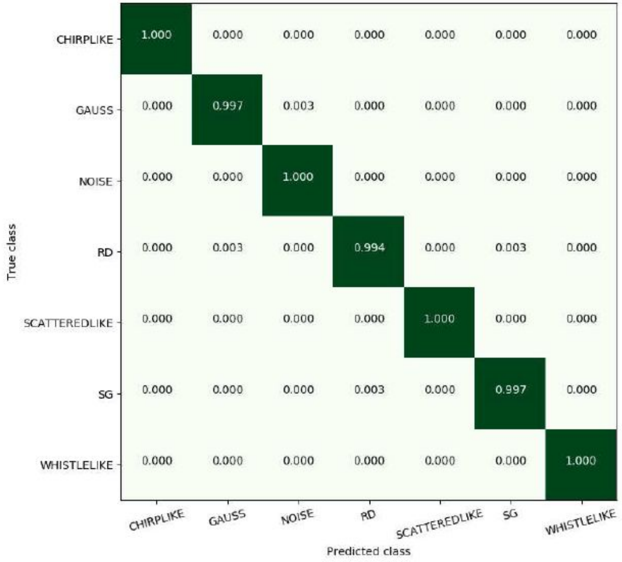


Block 1
- Convolutional (depth=16)
- Convolutional (depth=32)
- MaxPooling (2x2)
- Dropout (0.25)

Block 2
- Convolutional (depth=64)
- MaxPooling (2x2)
- Convolutional (depth=64)
- MaxPooling (2x2)
- Dropout (0.25)

Block 3
- Convolutional (depth=128)
- MaxPooling (2x2)
- Convolutional (depth=128)
- MaxPooling (2x2)
- Dropout (0.25)

Out Block
- Fully Connected (N=512)
- Dropout (0.25)
- Fully Connected ($N=N_{class}$)

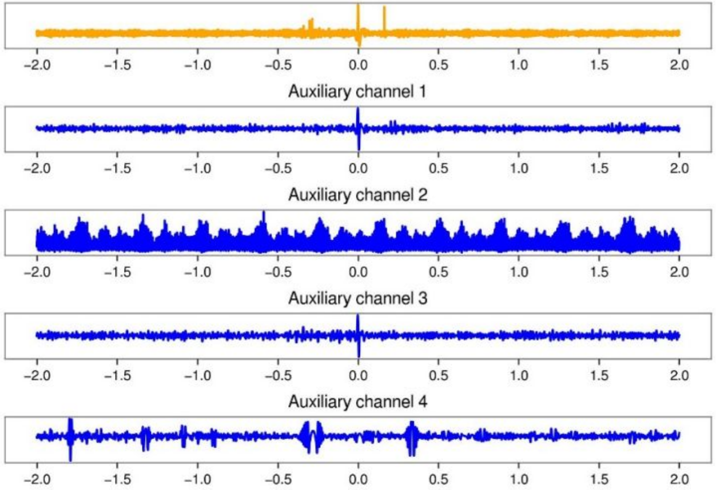


Normalized Confusion Matrix

Razzano M., Cuoco E.
CQG-104381.R3

SCUOLA NORMALE SUPERIORE
EGO - Virgo

Noise transient classification on auxiliary channels

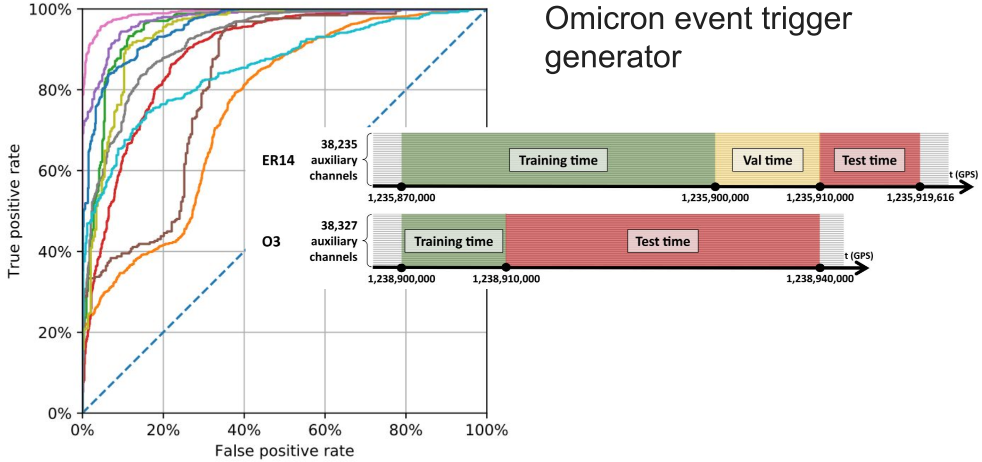


PHYSICAL REVIEW D **88,** 062003 (2013)

Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data

Rahul Biswas,[1] Lindy Blackburn,[2] Junwei Cao,[3] Reed Essick,[4] Kari Alison Hodge,[5] Erotokritos Katsavounidis,[4] Kyungmin Kim,[6,7] Young-Min Kim,[8,7] Eric-Olivier Le Bigot,[3] Chang-Hwan Lee,[8] John J. Oh,[7] Sang Hoon Oh,[7] Edwin J. Son,[7] Ye Tao,[9] Ruslan Vaulin,[4,*] and Xiaoge Wang[9]

Logistic regression and Omicron event trigger generator



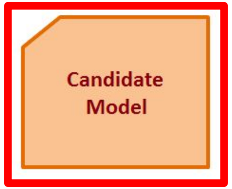
PHYSICAL REVIEW D **101,** 102003 (2020)

Efficient gravitational-wave glitch identification from environmental data through machine learning

Robert E. Colgan,[1,2] K. Rainer Corley,[3,4] Yenson Lau,[2,5] Imre Bartos,[6] John N. Wright,[2,5] Zsuzsa Márka,[4] and Szabolcs Márka[3]

M. Cavaglià courtesy

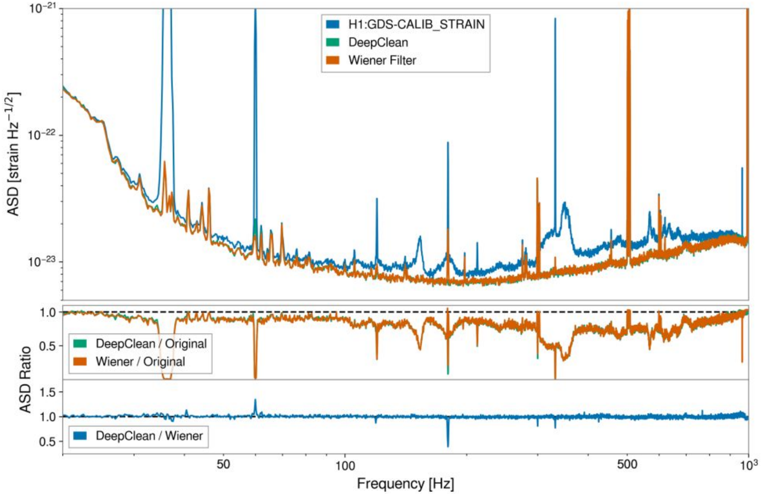
Deep Clean



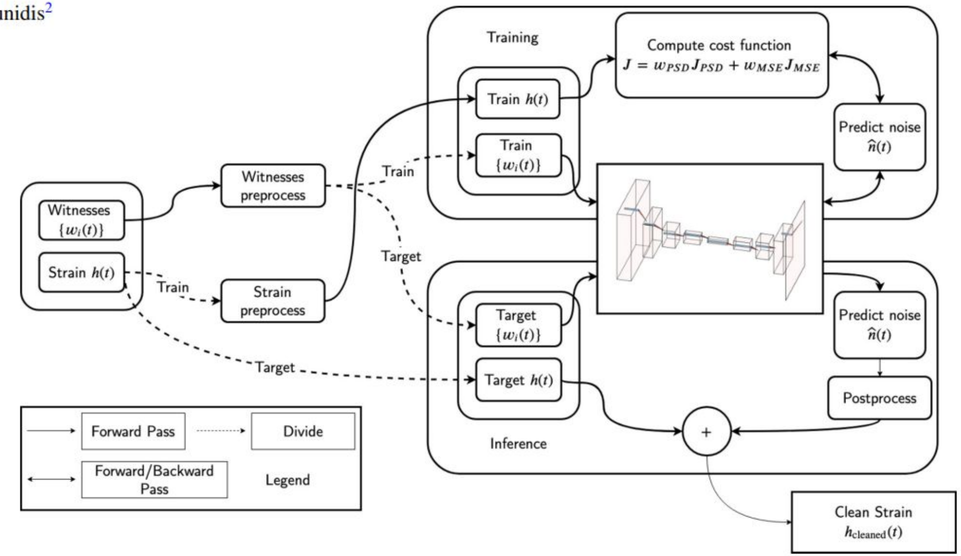
PHYSICAL REVIEW RESEARCH **2**, 033066 (2020)

Noise reduction in gravitational-wave data via deep learning

Rich Ormiston,[1] Tri Nguyen,[2] Michael Coughlin,[1,3] Rana X. Adhikari,[3] and Erik Katsavounidis[2]

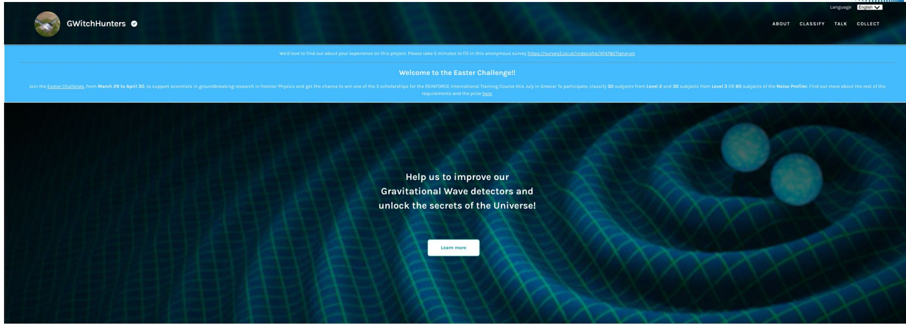


One-dimensional Convolutional Neural Network which takes a specified set of witness channels and subsequently outputs the predicted noise in strain.

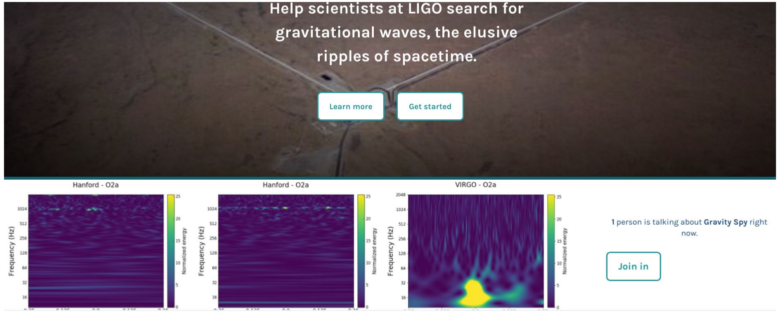


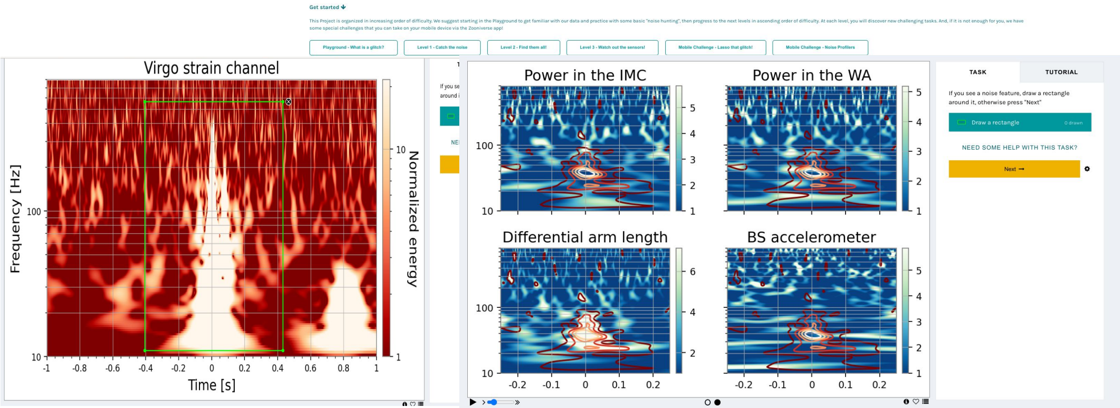
Citizen science for GW-AI

GWitchHunters



Gravity Spy





Citizen scientists contribute to classify glitches

More details in Zevin+17 [10.1088/1361-6382/aa5cea](10.1088/1361-6382/aa5cea)

https://doi.org/10.1016/j.ins.2018.02.068

REINFORCE
REsearch INfrastructures FOR Citizens in Europe

- *Team: M. Razzano, F. Di Renzo, F. Fidecaro (@Unipi), G. Hemming, S. Katsanevas (@EGO)*
- *Launched @ Nov 2019 – REINFORCE Project H2020-SWAFS (2019-2022)*

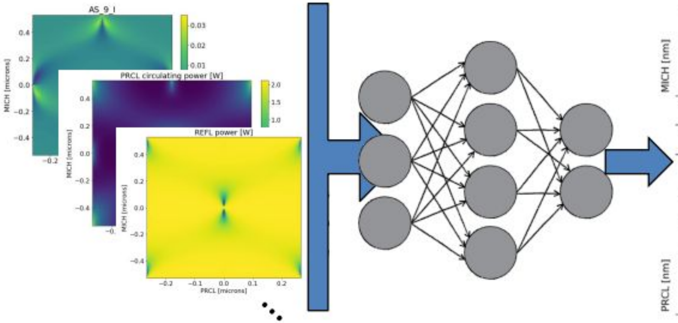
https://www.zooniverse.org/projects/reinforce/gwitchhunters

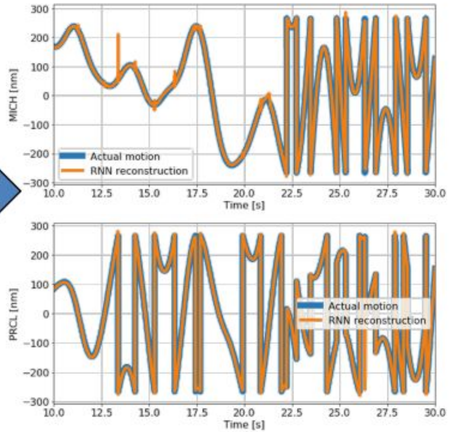
Machine learning for control system

Machine Learning for Lock Acquisition

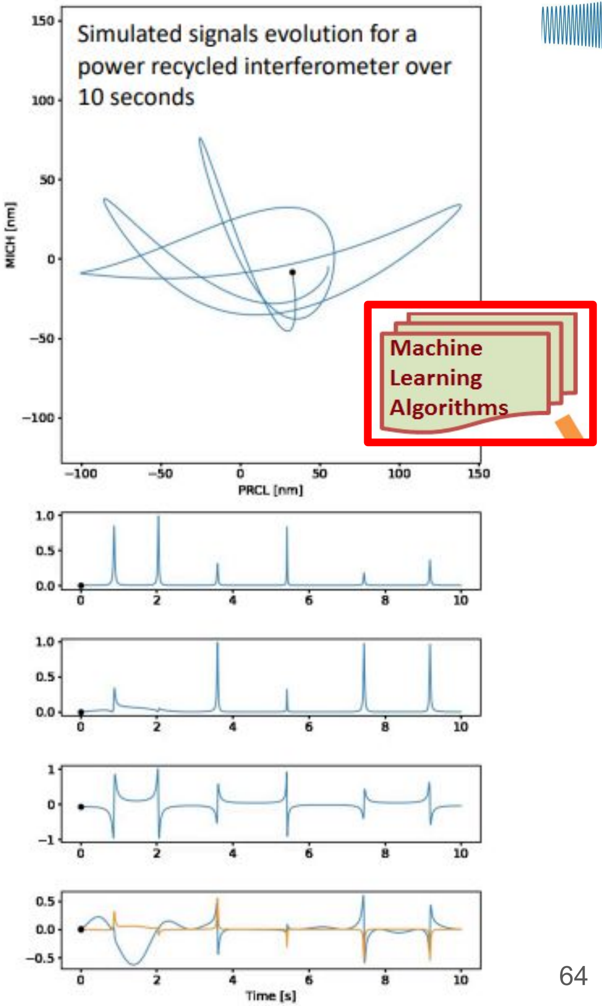
Non-linear control problem: drive the system into a narrow region of the 5d phase space, where linear control is possible

- Construct a **non-linear state estimator**: use all available signals as input, build an estimate of the degree of freedom positions that works (almost) everywhere





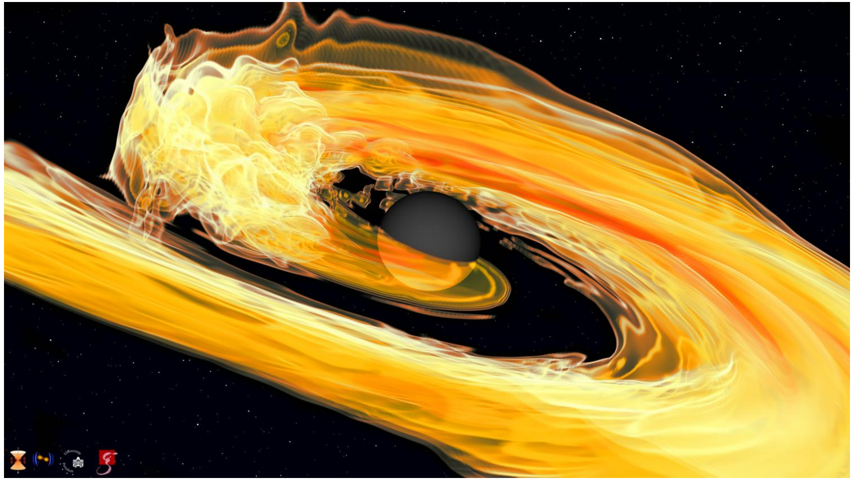
In simulation we have both **input signals** and **target coordinates**: supervised learning



Astrophysical GW source and ML

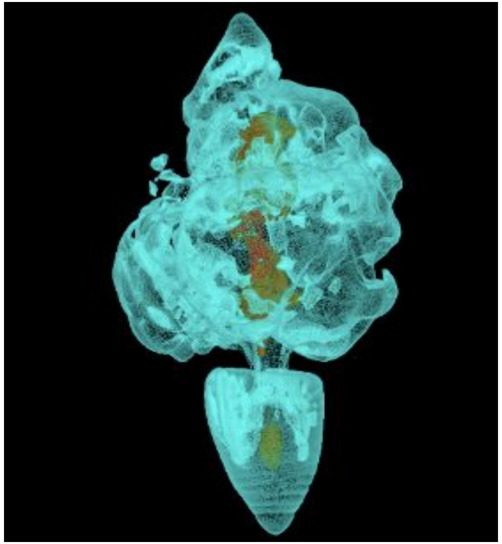
Compact Binary Coalescences

Matched filter modeled searches



Core Collapse Supernovae

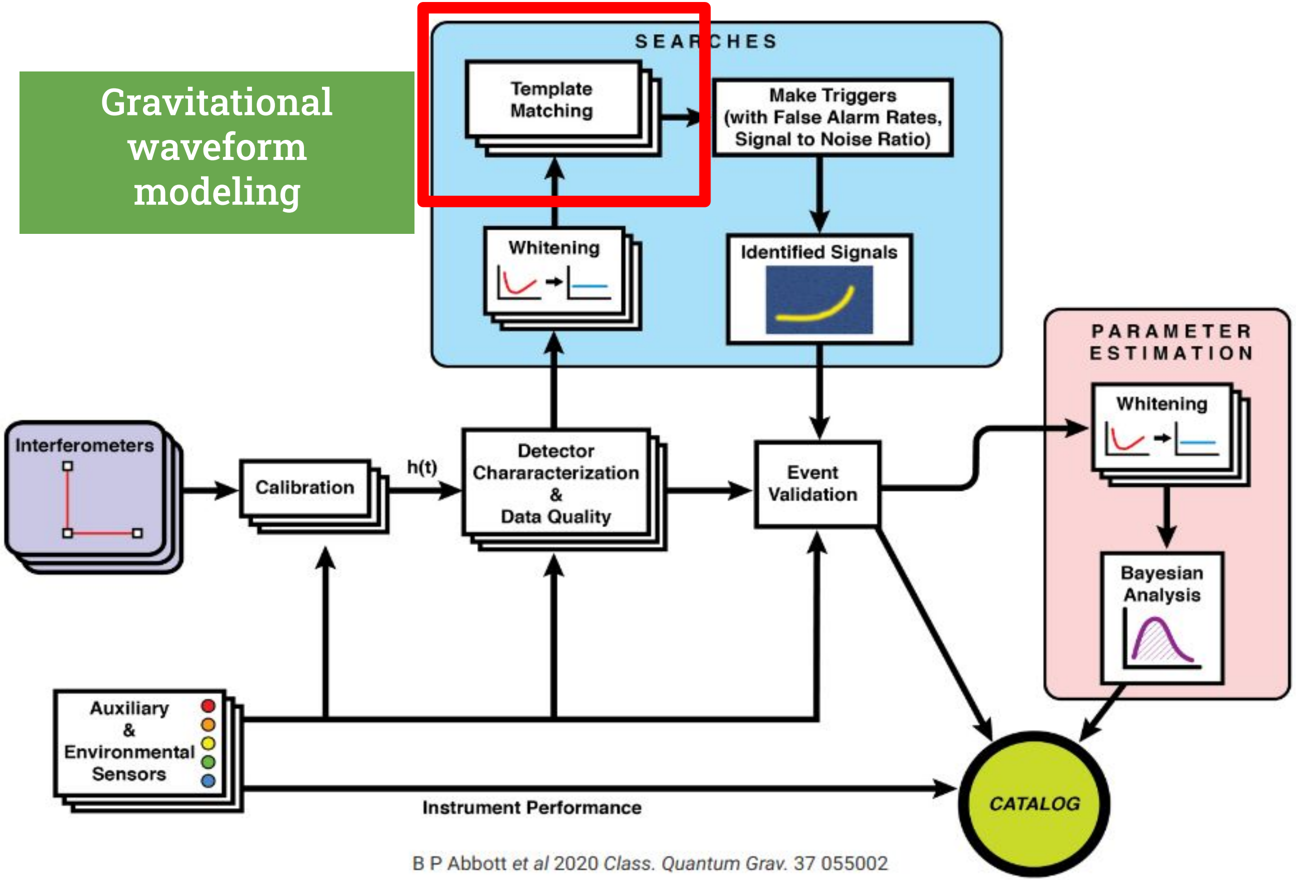
Unmodeled searches



Computer simulation of gravitational waves emitted by a supernova. Credit: J Powell / B Mueller

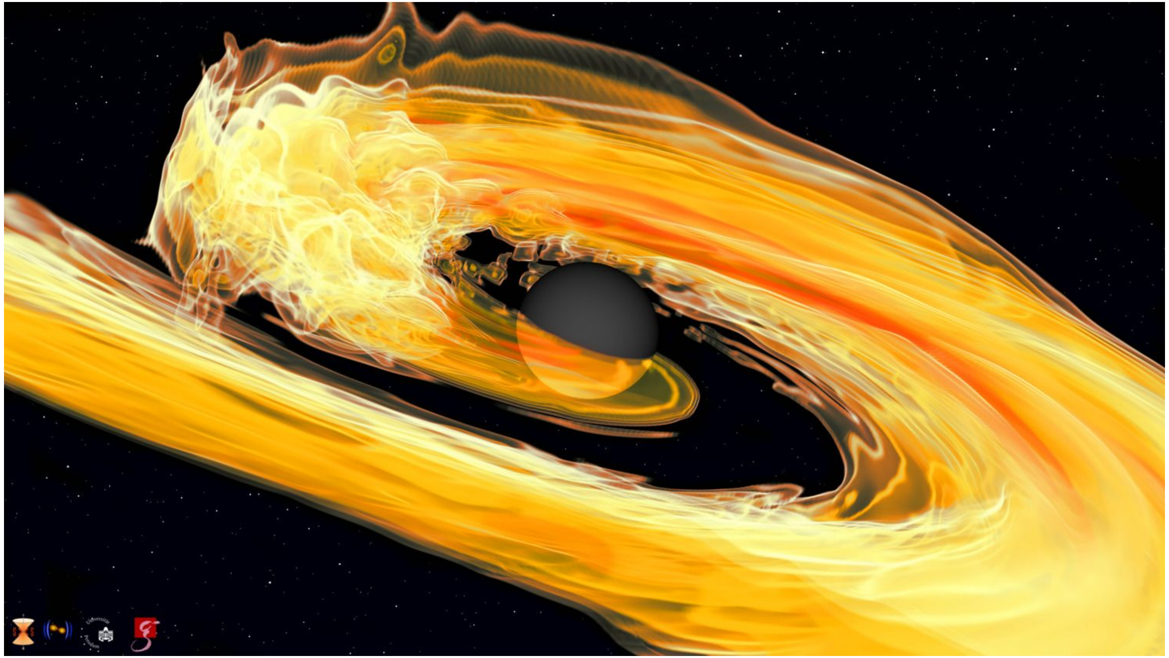
Template matching

Gravitational waveform modeling



B P Abbott *et al* 2020 *Class. Quantum Grav.* 37 055002

Gravitational wave modeling



- GW detection of binary systems relies on matched-filter analysis. Template accuracy is crucial!
- Accurate solutions of the Einstein equations for binary sources can be obtained with Numerical Relativity (NR) simulations.
- High computational cost prevent the production of NR waveforms catalogs spanning the full parameter space.
- LIGO and Virgo rely on approximate solutions that are traditionally obtained through the effective-one-body or phenomenological modeling approaches.
- How can machine learning help?

Waveform Building

Machine Learning Algorithms

Gaussian process regression to compute the waveform at points of the parameter space not covered by numerical relativity.

GPR has been used to build surrogate models of both non-precessing and precessing BBH systems.

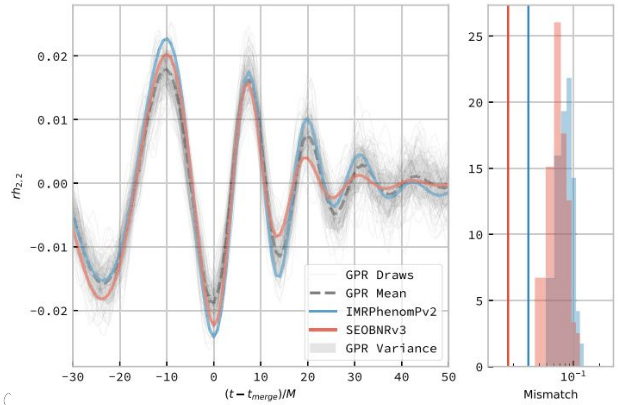
PHYSICAL REVIEW D **101,** 063011 (2020)

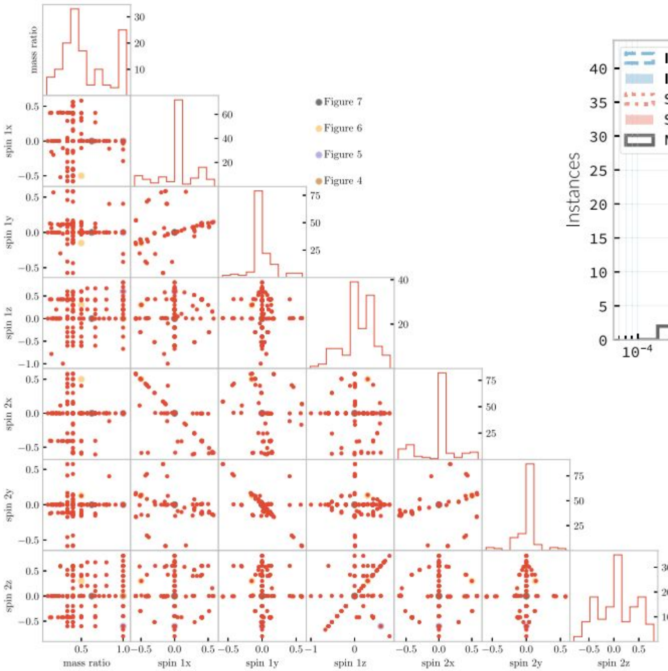
Precessing numerical relativity waveform surrogate model for binary black holes: A Gaussian process regression approach

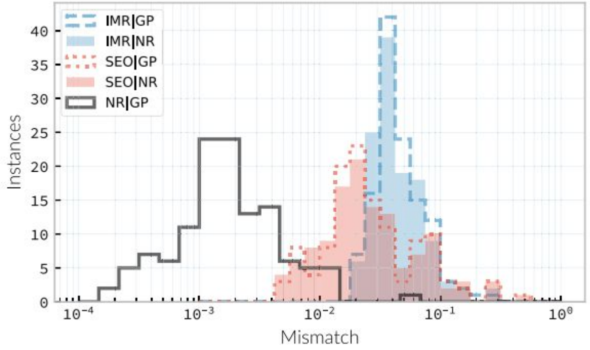
D. Williams* and I. S. Heng
SUPA, University of Glasgow, Glasgow G12 8QQ, United Kingdom

J. Gair
Max Planck Institute for Gravitational Physics, Potsdam Science Park, Am Mühlenberg 1, D-14476 Potsdam, Germany

J. A. Clark and B. Khamesra
Center for Relativistic Astrophysics and School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

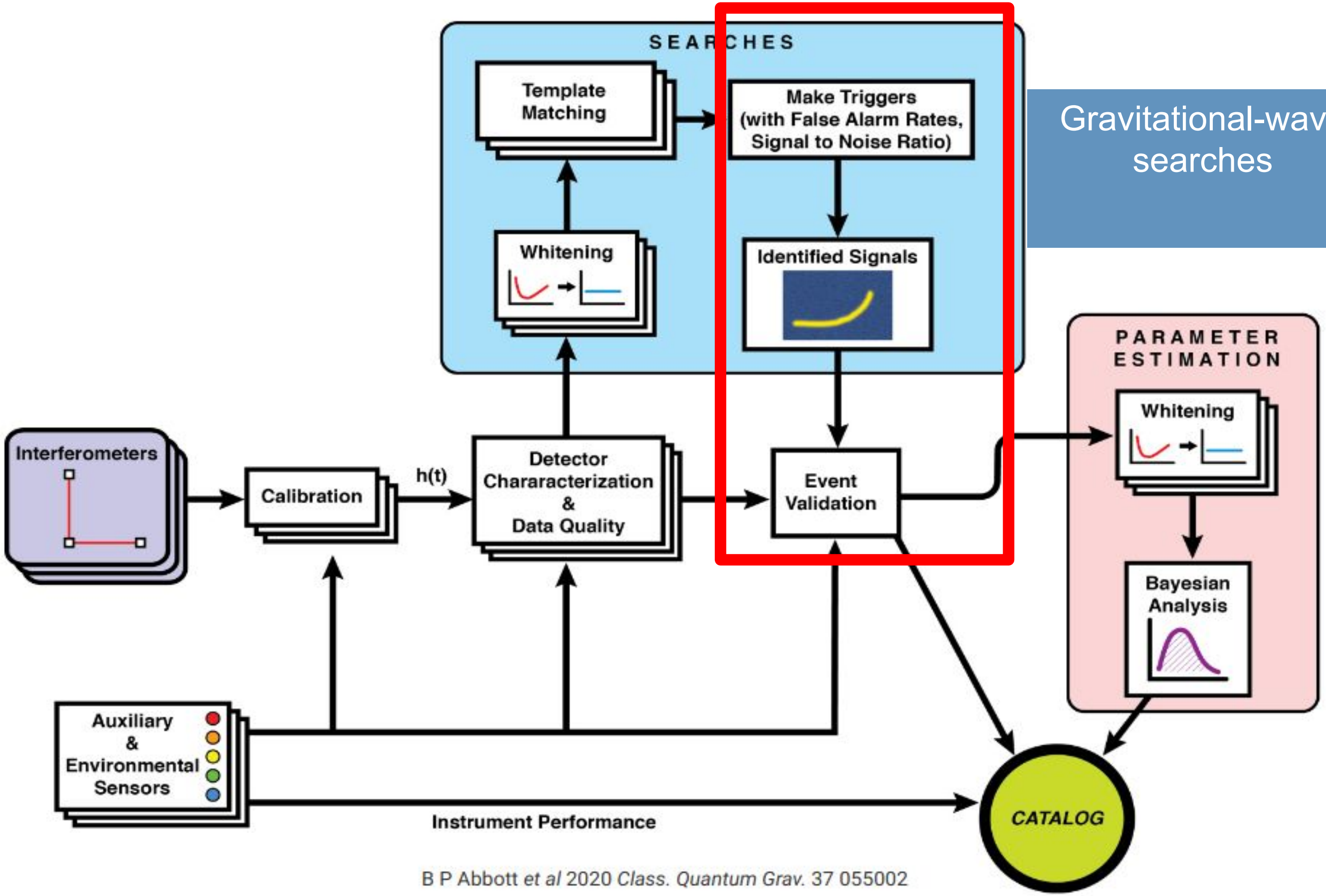






See also:
Z. Doctor et al, "Statistical gravitational waveform models: What to simulate next?"
Phys. Rev. D 96, 123011 (2017)

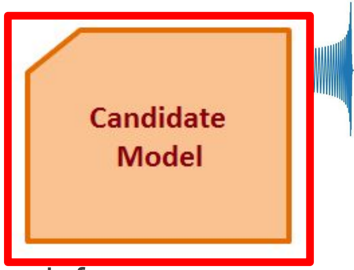
Gravitational wave signal detection



Gravitational-wave searches

B P Abbott *et al* 2020 *Class. Quantum Grav.* 37 055002

CBC detection



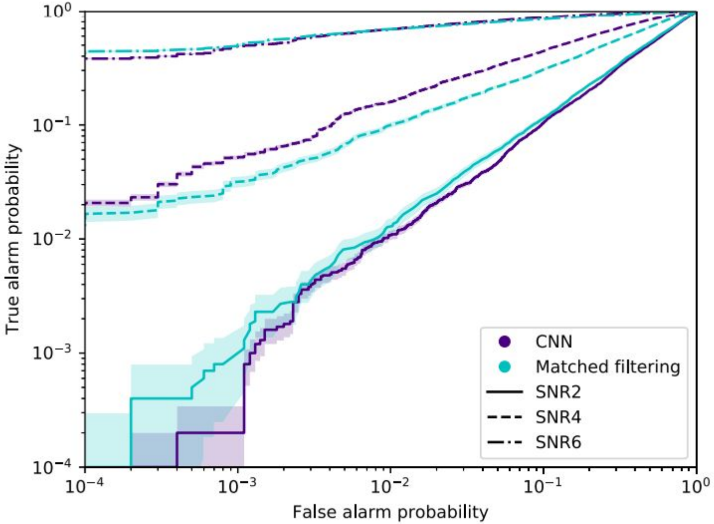
- Deep convolutional neural network to search for binary black hole gravitational-wave signals.
- Input is the whitened time series of measured gravitational-wave strain in Gaussian noise.
- Sensitivity comparable to match filtering.

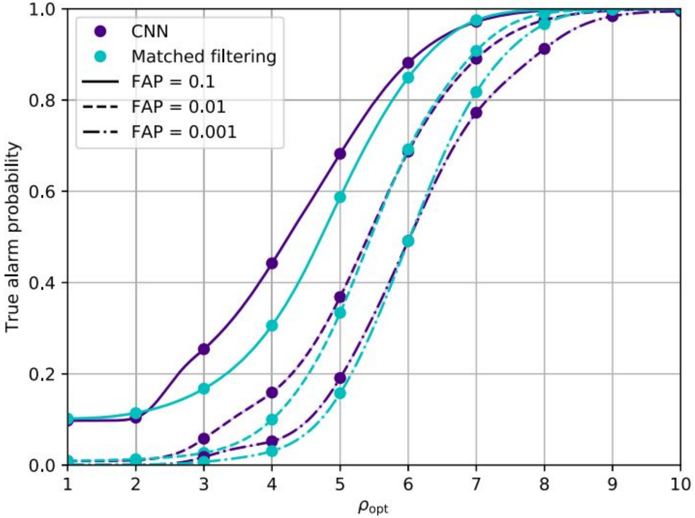
PHYSICAL REVIEW LETTERS **120**, 141103 (2018)

Editors' Suggestion

Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy

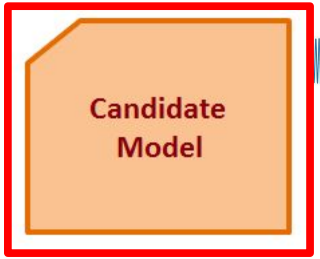
Hunter Gabbard,* Michael Williams, Fergus Hayes, and Chris Messenger
SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom





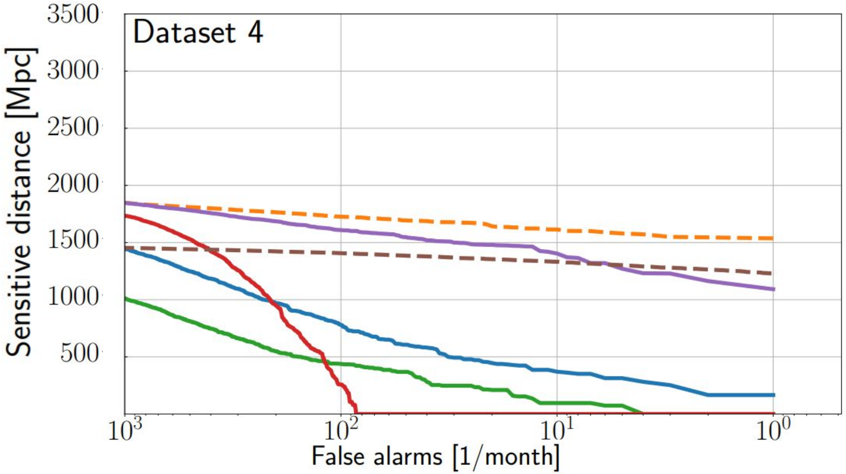
See also:
D. George and E.A. Huerta
Phys. Lett. B 778 64–70 (2018)

CBC detection



MLGWSC-1: The first Machine Learning Gravitational-Wave Search Mock Data Challenge

Marlin B. Schäfer[1,2], Ondřej Zelenka[3,4], Alexander H. Nitz[1,2], He Wang[5], Shichao Wu[1,2], Zong-Kuan Guo[5], Zhoujian Cao[6], Zhixiang Ren[7], Paraskevi Nousi[8], Nikolaos Stergioulas[9], Panagiotis Iosif[10,9], Alexandra E. Koloniari[9], Anastasios Tefas[8], Nikolaos Passalis[8], Francesco Salemi[11,12], Gabriele Vedovato[13], Sergey Klimenko[14], Tanmaya Mishra[14], Bernd Brügmann[3,4], Elena Cuoco[15,16,17], E. A. Huerta[18,19], Chris Messenger[20], Frank Ohme[1,2]



- Comparison of 6 algorithms for binary black hole searches.
- Four different data sets of different complexity (from Gaussian noise to varying real detector PSD)
- Benchmark data set for algorithm testing.

A few excerpts from the paper conclusions:

- Machine learning search algorithms are competitive in sensitivity compared to state-of-the-art searches on simulated data and the limited parameter space explored in this challenge.
- Most of the tested machine learning algorithms struggle to effectively handle real noise, which is contaminated with non-Gaussian noise artifacts.
- Traditional search algorithms are capable of detecting signals at lower FARs, thus making detections more confident.
- The tested machine learning searches struggle to identify long duration signals.

Anomaly Detection in Gravitational Waves data using Convolutional AutoEncoders for CBC signals

Filip Morawski, Michał Bejger, Elena Cuoco, Luigia Petre, https://doi.org/10.1088/2632-2153/abf3d0

SCUOLA NORMALE SUPERIORE

EGO - Virgo

CAMK PAN

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

G2NET

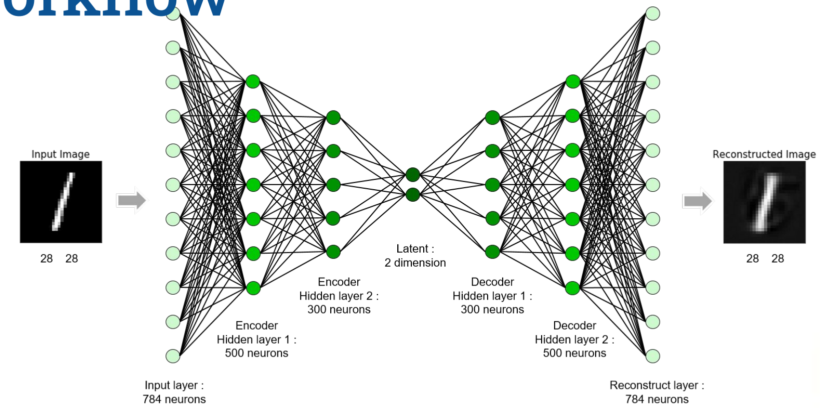
Example for detection/classification for CBC signals

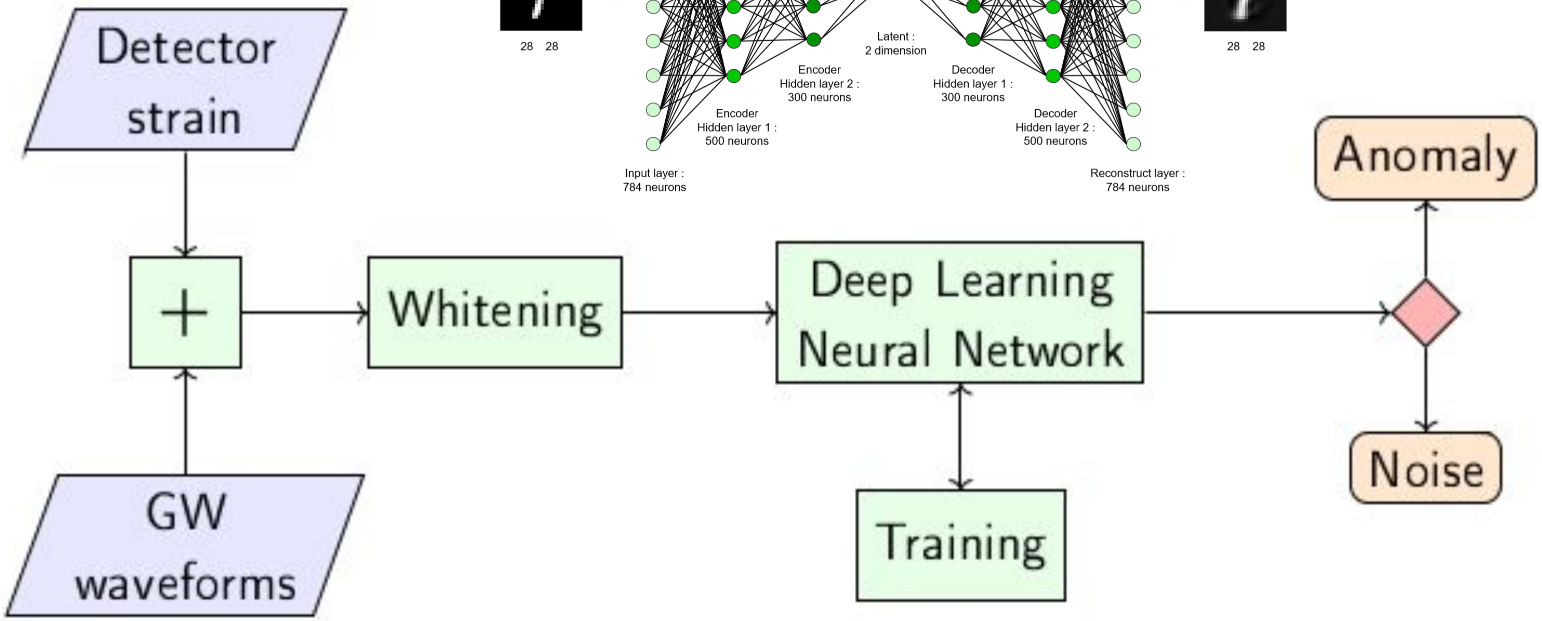
Create a deep learning pipeline allowing detection of anomalies defined in terms of **transient signals**: gravitational waves as well as glitches.

Additionally: Consider **reconstruction of the signal** for the found anomalies.

Filip Morawski, *Michał Bejger, Elena Cuoco, Luigia Petre, https://iopscience.iop.org/article/10.1088/2632-2153/abf3d0*

Auto-encoder workflow



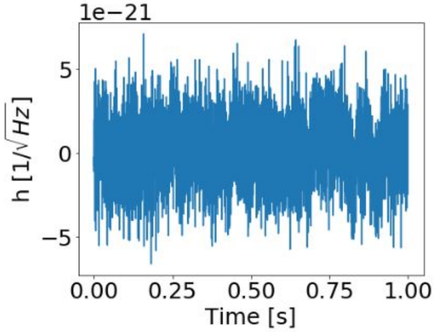


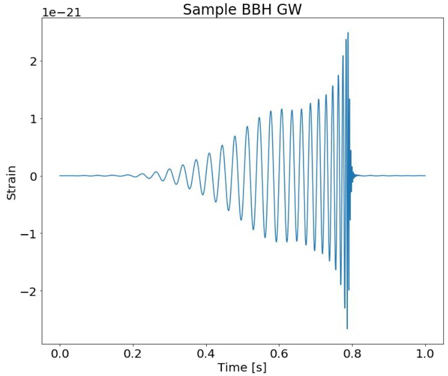
Filip Morawski, *Michał Bejger, Elena Cuoco, Luigia Petre, https://iopscience.iop.org/article/10.1088/2632-2153/abf3d0*

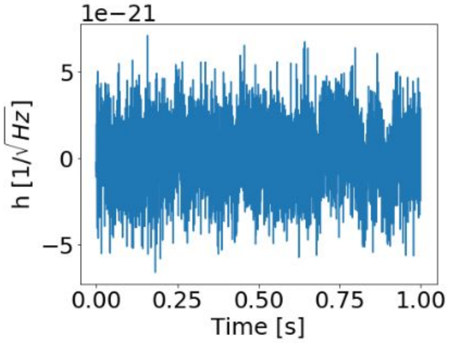
Auto-encoder workflow

Model input

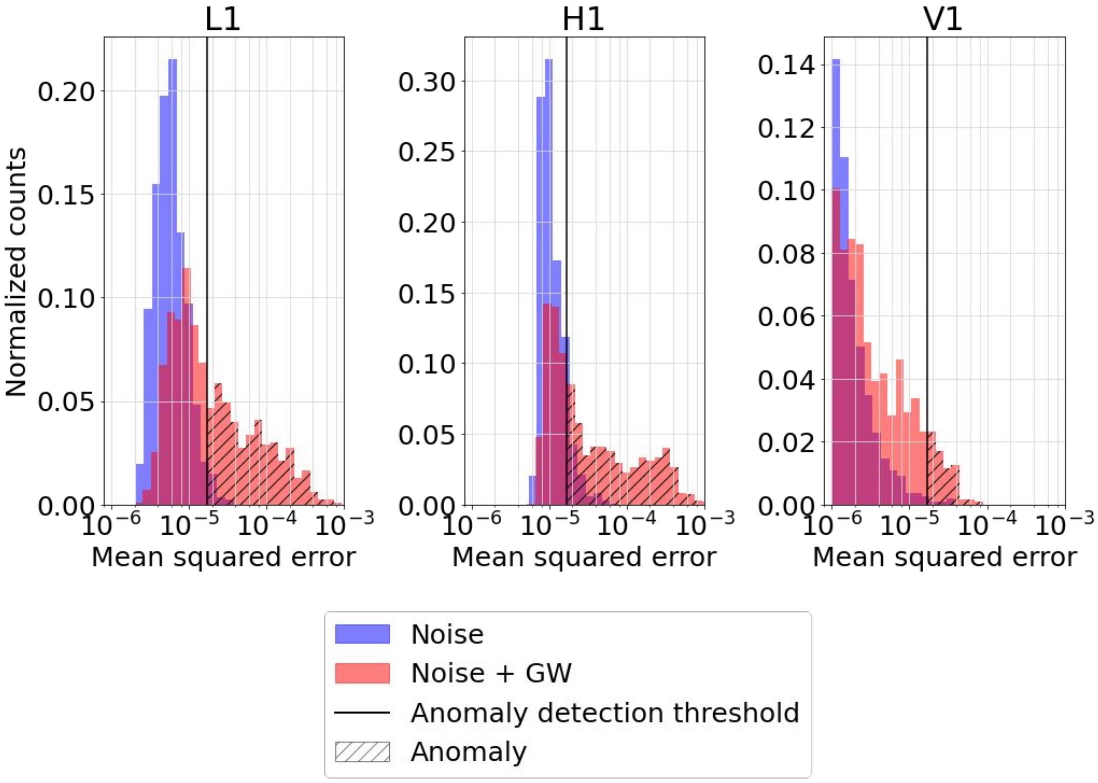
Model prediction





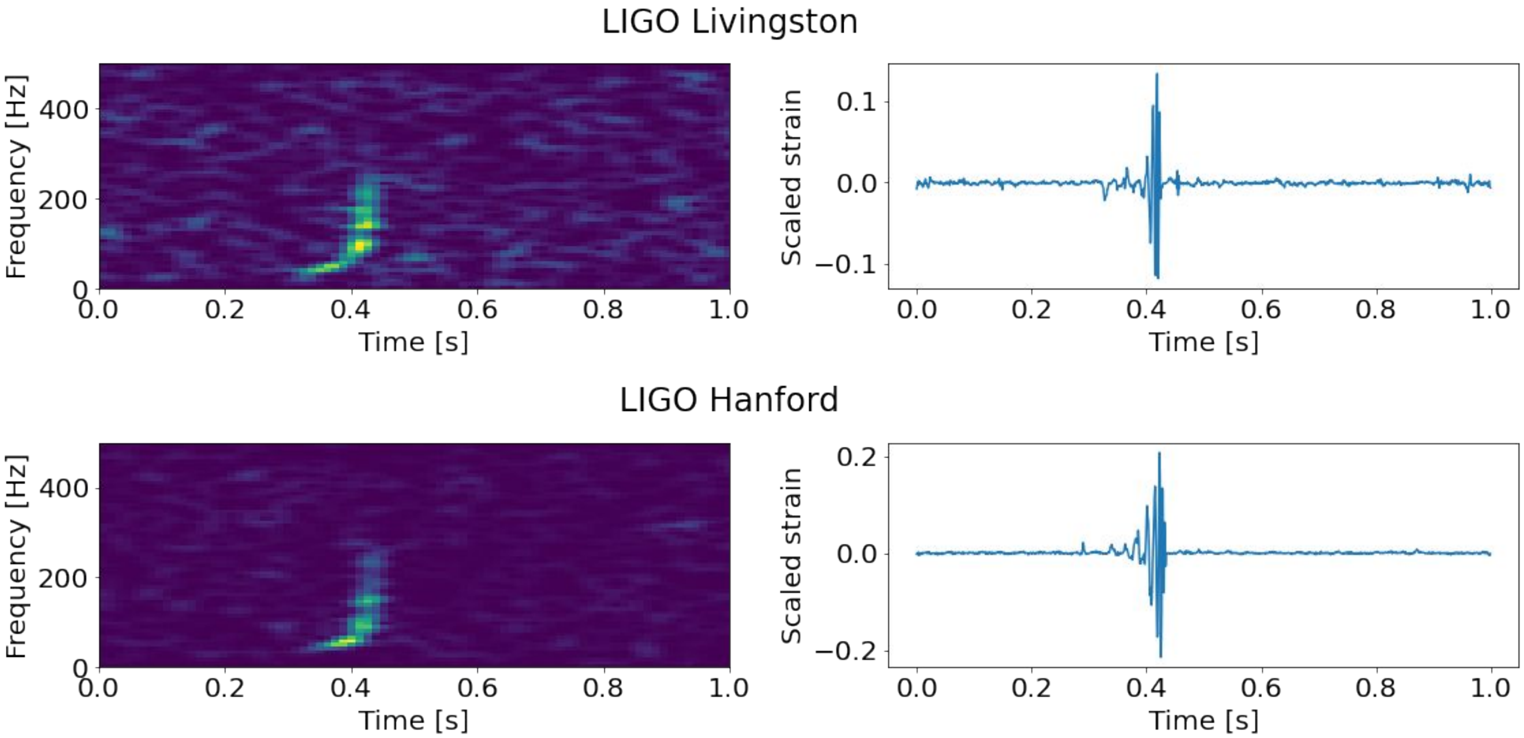


O2 data - MSE Distributions



SCUOLA NORMALE SUPERIORE
EGO - Virgo

GW150914



Filip Morawski, *Michał Bejger, Elena Cuoco, Luigia Petre,*
https://iopscience.iop.org/article/10.1088/2632-2153/abf3d0

Rapid localization of sources

Machine Learning Algorithms

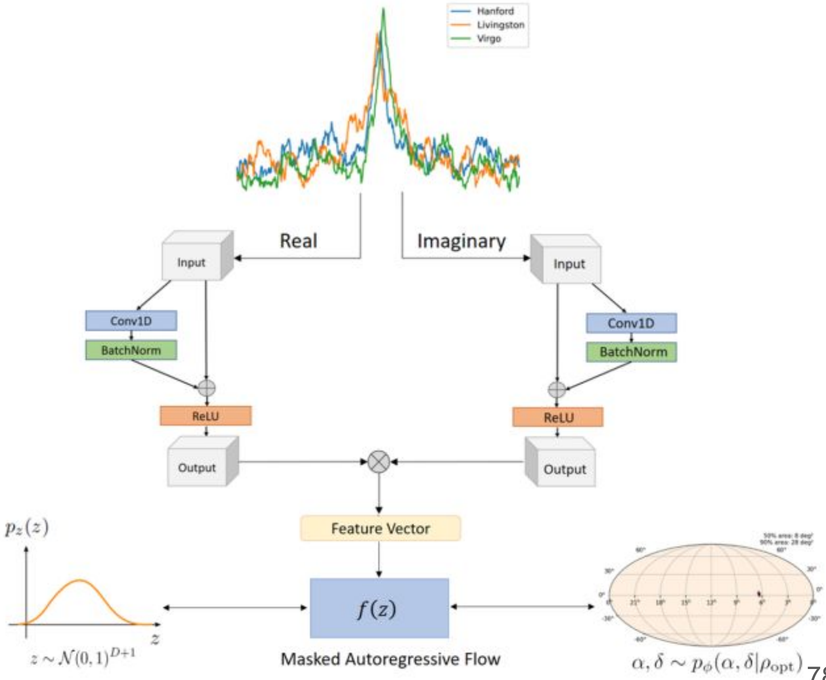
Rapid localization of gravitational wave sources from compact binary coalescences using deep learning

Chayan Chatterjee,* Linqing Wen,† and Damon Beveridge‡
Department of Physics, OzGrav-UWA, The University of Western Australia, 35 Stirling Hwy, Crawley, Western Australia 6009, Australia

Foivos Diakogiannis§
The Commonwealth Scientific and Industrial Research Organisation 7 Conlon St, Waterford, WA, Australia

Kevin Vinsen¶
International Centre for Radio Astronomy Research, The University of Western Australia M468, 35 Stirling Hwy, Crawley, WA, Australia
(Dated: August 1, 2022)

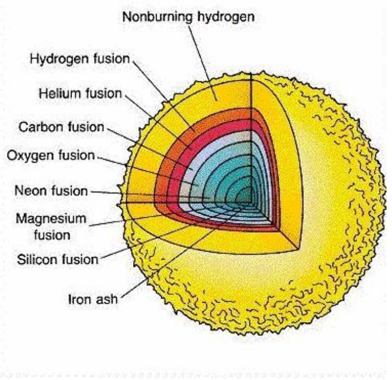
- Deep learning-based approach for sky localization of binary coalescences
- Train and test a normalizing flow model on matched-filtering output from GW searches.
- Fast sky localizations.



GWs from Core Collapse Supernovae

- Waveform depends on progenitor star
- Different emission mechanisms (Proto-neutron star oscillation, Standing Accretion Shock Instability (SASI),..)
- Largely Stochastic
- Best waveform models from computationally expensive 3D simulations
- Different simulation models
- Rare (~100 yrs in Milky Way)

Need an alternative to matched filter approach

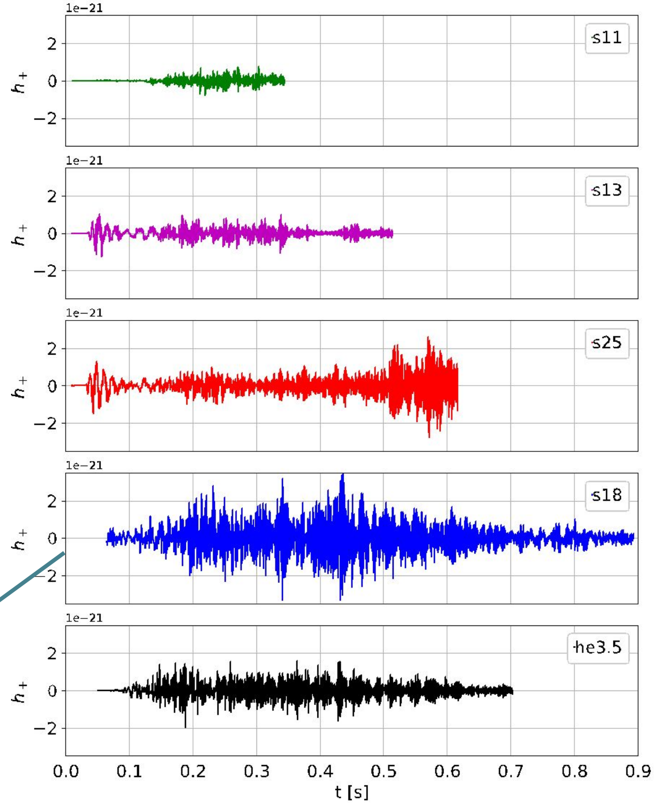


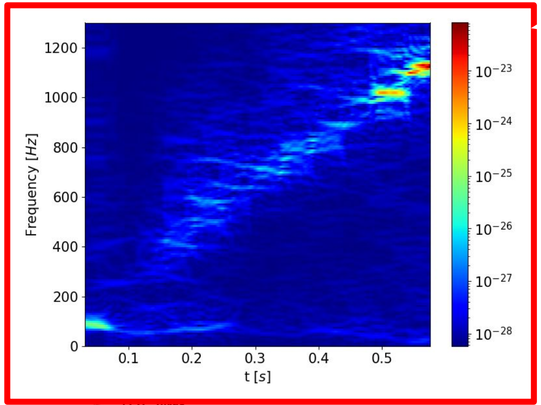
GW emission Process	Potential explosion mechanism: MHD mechanism (rapid rotation)	Neutrino mechanism (slow/no rotation)	Acoustic mechanism (slow/no rotation)
Rotating collapse and Bounce	Strong	None/weak	None/weak
3D rotational instabilities	Strong	None	None
Convection & SASI	None/weak	Weak	Weak
PNS g-modes	None/weak	None/weak	Strong

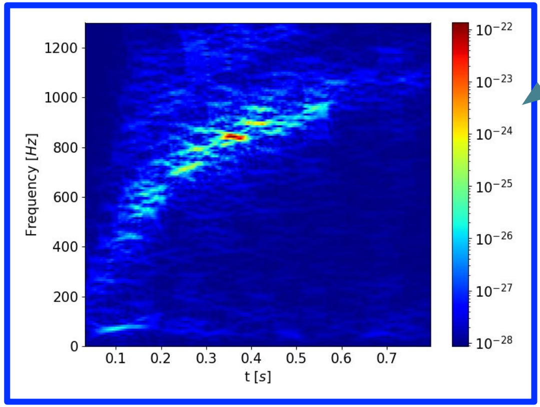
Ott et al. (2017)

Core-Collapse Supernovae models

- *Andresen s11: Low amplitude, non-exploding, peak emission at lower frequencies*
- *Radice s13: Non-exploding, lower amplitudes*
- *Radice s25: Late explosion time, standing accretion shock instability (SASI), high peak frequency*
- *Powell s18: High peak frequency, exploding model*
- *Powell He3.5: ultra-stripped helium star, high peak frequency, exploding model*





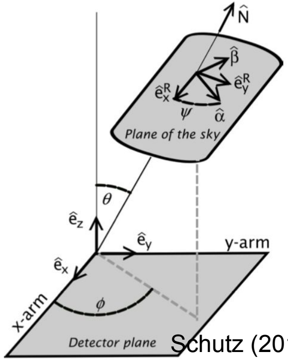


Iess, Cuoco, Morawski, Powell, https://doi.org/10.1088/2632-2153/ab7d31

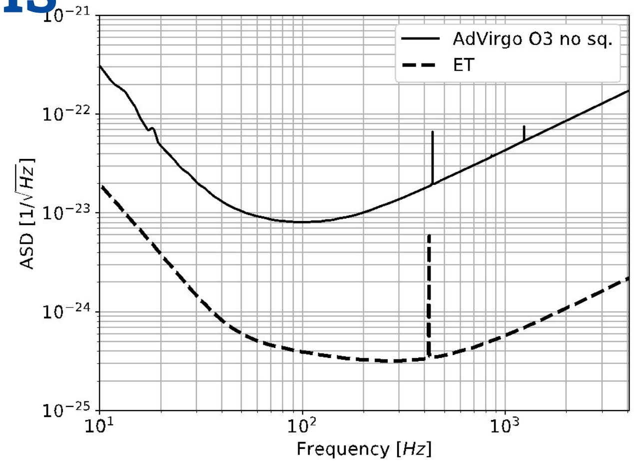
MDC and CCSN GW simulations

$$h(t) = F_+ h_+(t) + F_\times h_\times(t)$$

- Distances:
  VO3 0.01 kpc to 10 kpc
  ET 0.1 kpc to 1000 kpc
- Random sky localization
- Large SNR range



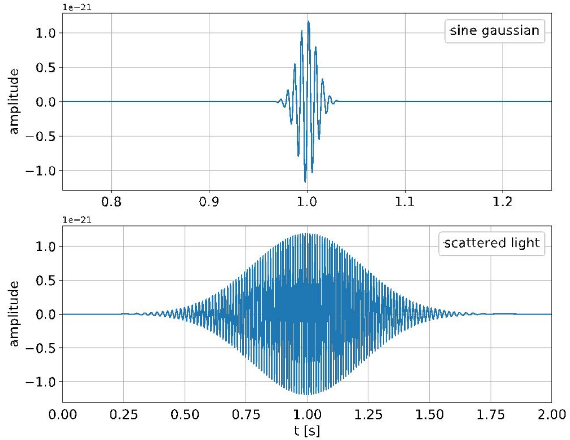
Schutz (2011)



SINE GAUSSIAN & SCATTERED LIGHT GLITCHES

$$h_{SG}(t) = h_0 \sin(2\pi f_0 (t - t_0)) e^{-\frac{(t-t_0)^2}{2\tau^2}}$$

$$h_{SL}(t) = h_0 \sin(\phi_{SL}) e^{-\frac{(t-t_0)^2}{2\tau}} \quad \phi_{SL} = 2\pi f_0 (t - t_0)[1 - K(t - t_0)^2]$$

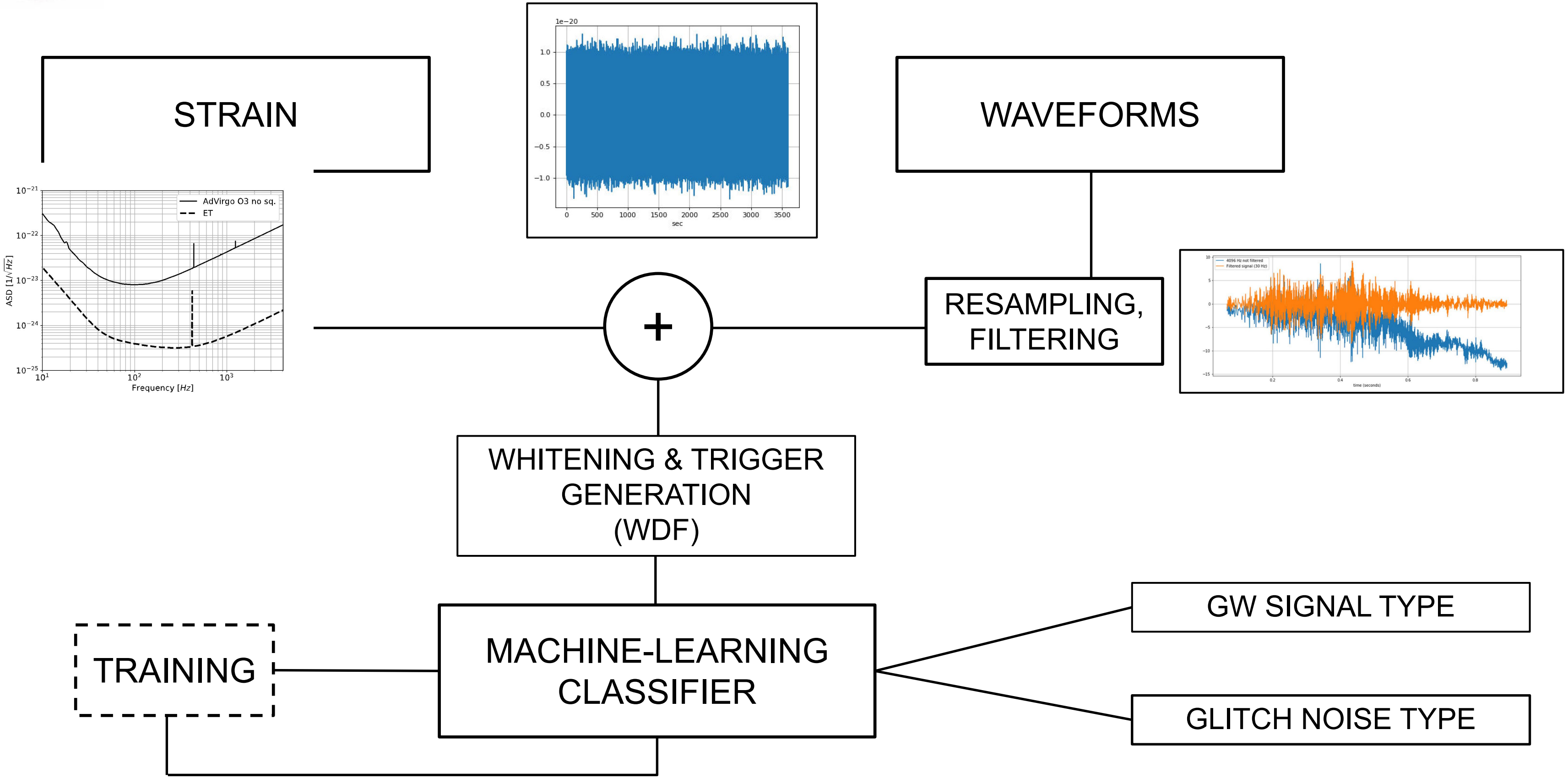


BACKGROUND STRAIN : simulated data sampled at 4096 Hz built from VO3 and ET projected sensitivities

SCUOLA NORMALE SUPERIORE

EGO - Virgo

Pipeline Workflow

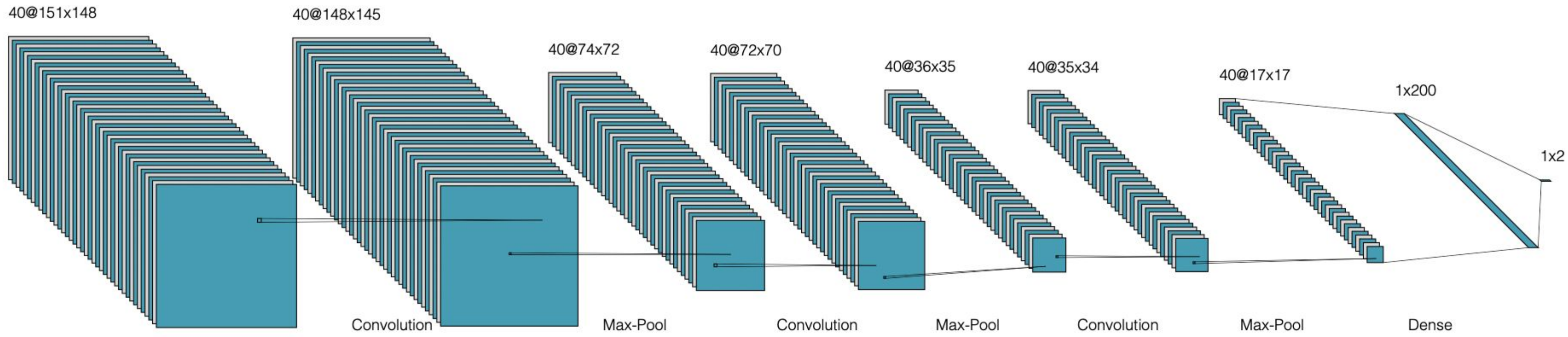


Neural Network architecture

- ***Train, Validation, Test sets: 60%, 10%, 30%***
- 3 or 4 Convolutional layers
- Activation function *f*: ReLU
- Adam optimizer, learning rate α = 0.001, decay rate of 0.066667
- Early stopping
- Batch Size: 64 or 128
- Loss function: Categorical-cross entropy

Dataset: chunks of 3 hr data with 1000 injections/h

GPU: Tesla k40



Binary Classification

- Train on **<u>all</u>** CCSNe waveforms and glitches.
- Test on **<u>all</u>**.

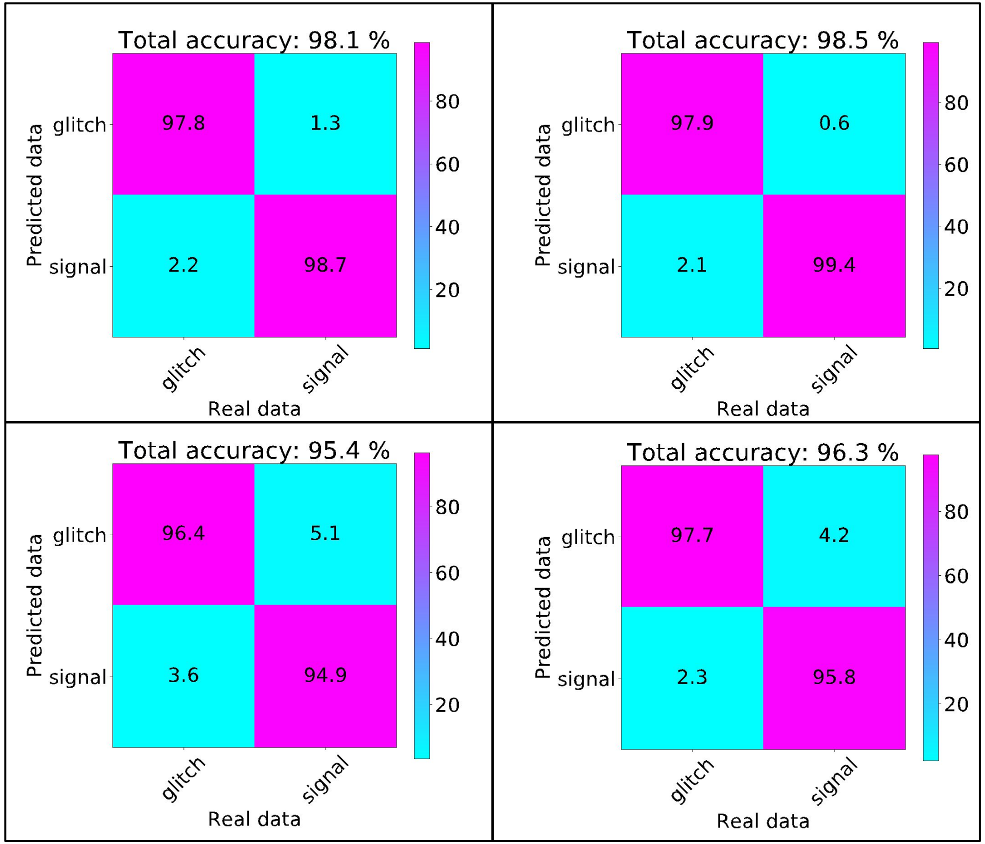
Test samples → TRAINED CNN MODEL → SIGNAL / GLITCH

- Training time: ~ 30 min

1D CNN | 2D CNN

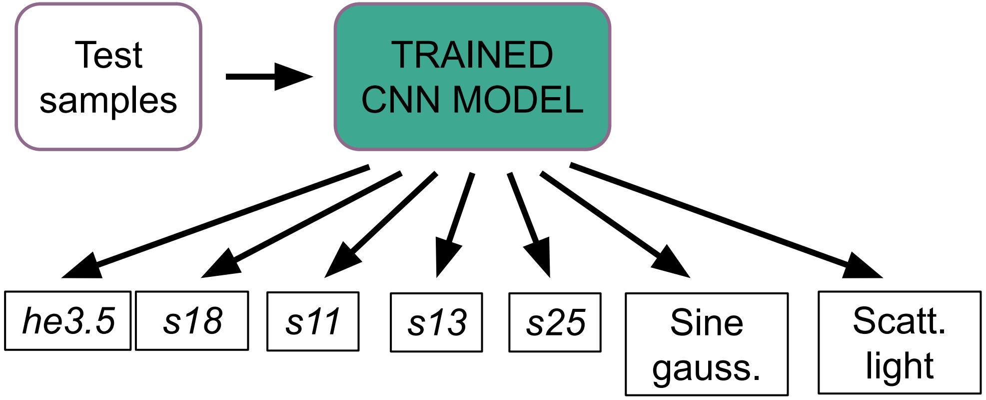
ET

VO3



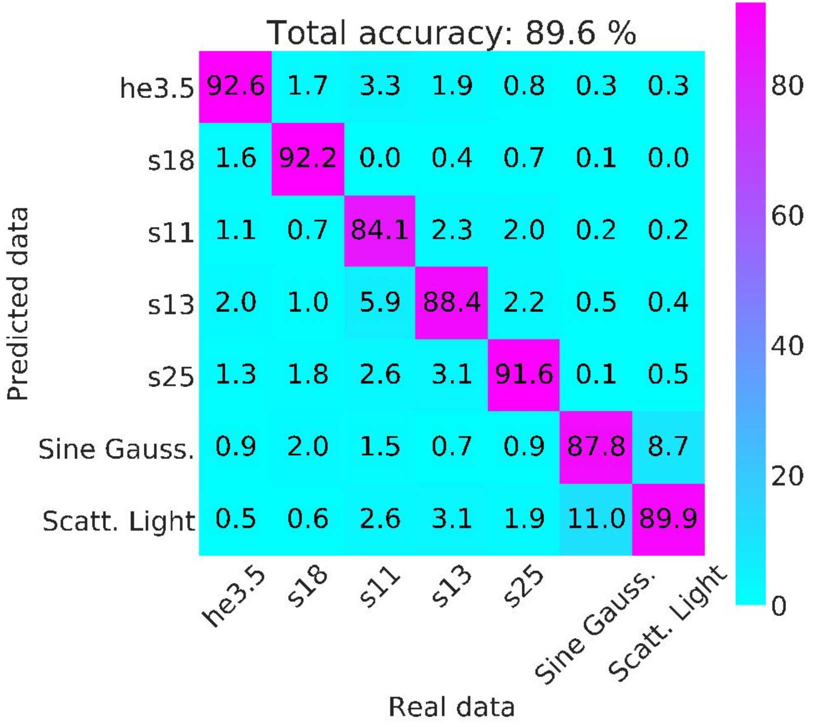
MultiLabel classification

- Train on **<u>all</u>** (4 CCSNe waveform models + glitches).
- Test on **<u>all.</u>**



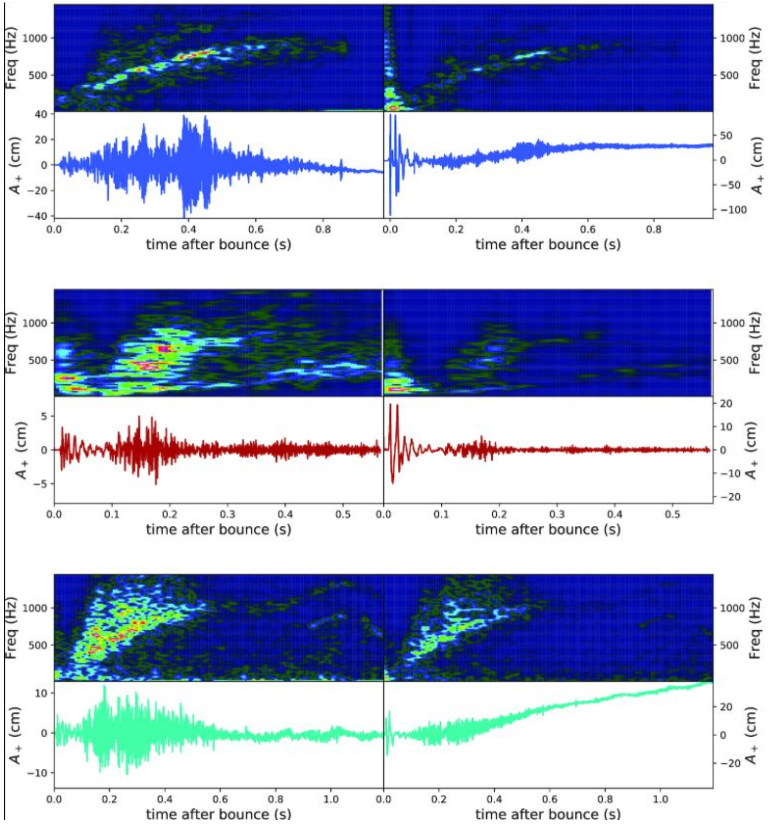
COMPLEX TASK → LONGER TRAINING (> 1 hr)

ET, MERGED 1D & 2D CNN



Test on O2 real Data

- 44 segments (4096s per segment) from O2 science run.
- Added m39, y20, s18np models (Powell, Mueller 2020).
- **<u>Fixed distance of 1 kpc</u>**.
- Added LSTM Networks, suited for timeseries data.
- **Added Three ITF classification.**



- *Powell s18np*: differs from s18 since simulation does not include perturbations from the convective oxygen shell. As a result, this model develops strong SASI after collapse.

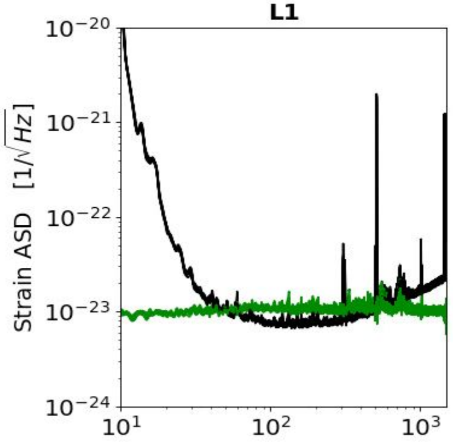
- *Powell y20*: non-rotating, 20 solar mass Wolf-Rayet star with solar metallicity.

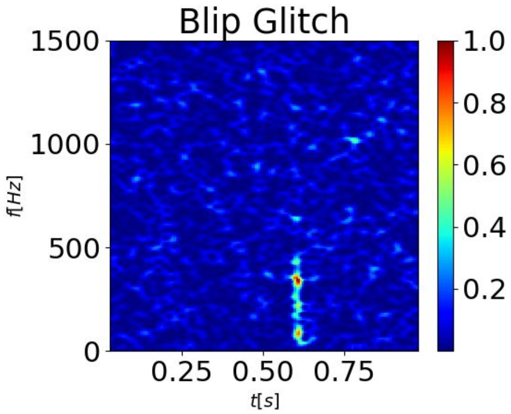
- *Powell m39*: rapidly rotating Wolf-Rayet star with an initial helium star mass of 39 solar masses

Real noise from O2 science run

- Noise PSD is non stationary.
- Multiple Glitch Families.
- SNR distribution is affected by ITF antenna pattern.
- Dataset: ~15000 samples.
- Imbalanced Dataset due to different model amplitudes.

	Triggers		
Detector	*Signal*	*Noise*	*Total*
Virgo V1	9273	47901	57174
Ligo L1	10480	3810	14290
Ligo H1	10984	4103	15087
L1, H1, V1	5647	675	6322





CCSN Classification on Simulated and Real O2 Data with CNNs and LSTMs
***A. Iess**, E. Cuoco, F. Morawski, C. Nicolaou, O. Lahav, accepted for A&A*

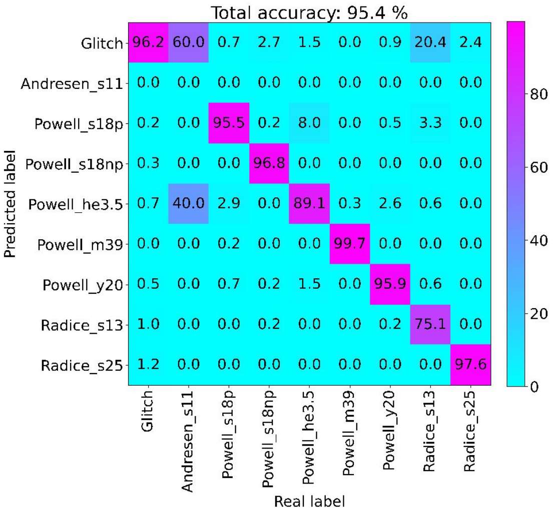
Multi-label task

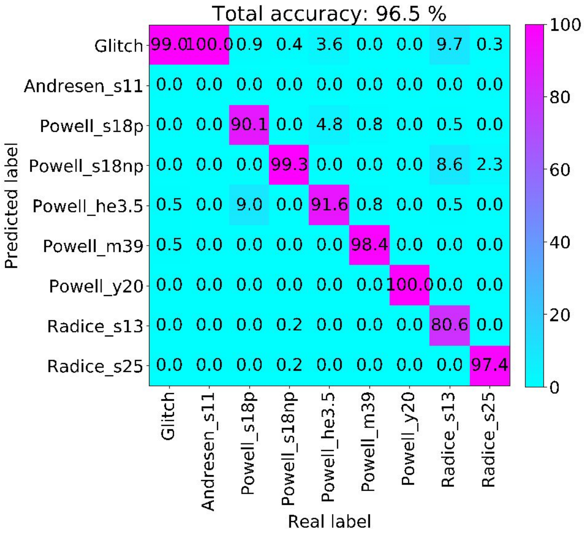
MULTILABEL CLASSIFICATION ON REAL O2 NOISE (SINGLE ITF, LIGO H1, DIFFERENT MODELS)

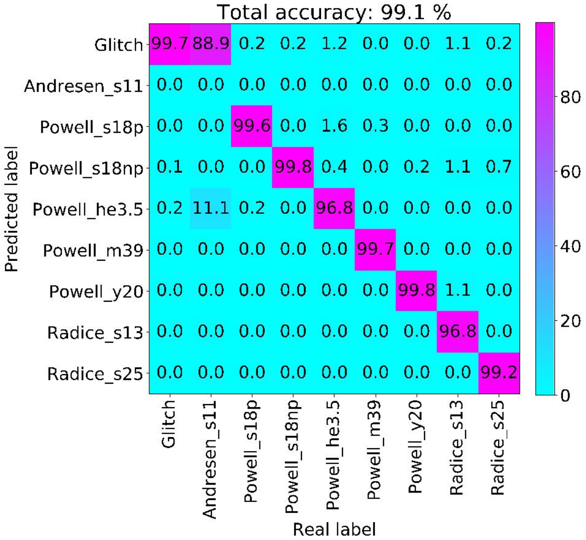
- **<u>Bi-LSTM</u>**, 2 recurrent layers
- ~10 ms/sample
- Best weights over 100 epochs

- **<u>1D-CNN</u>**, 4 convolutional layers
- ~2 ms/sample
- Best weights over 20 epochs

- **<u>2D-CNN</u>**, 4 convolutional layers
- ~3 ms/sample
- Best weights over 20 epochs



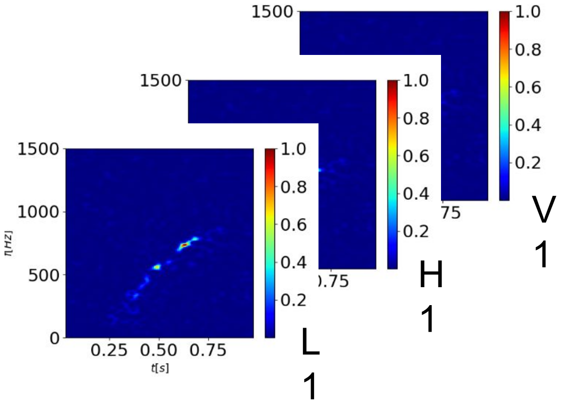


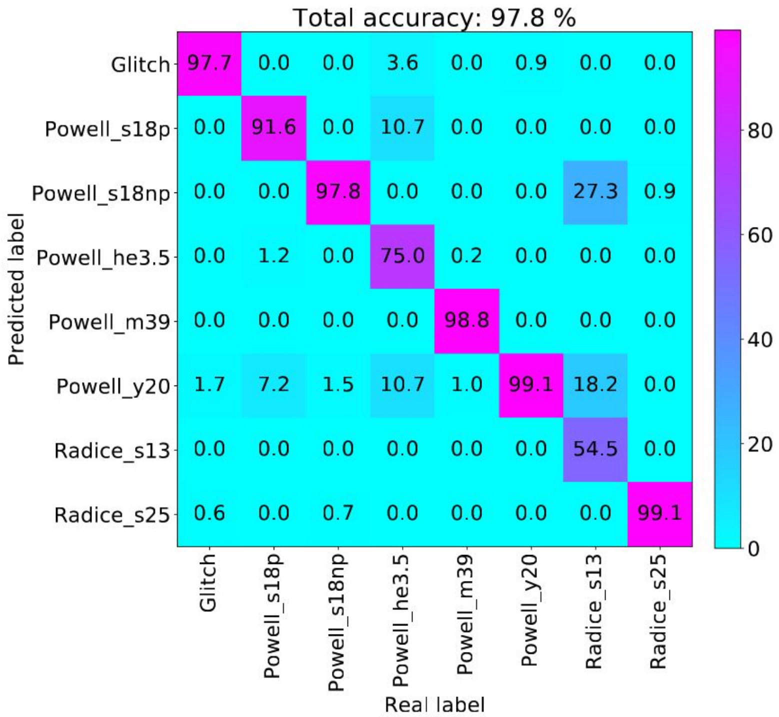


__A. Iess__, E. Cuoco, F. Morawski, C. Nicolaou, O. Lahav, 2022

Analysis on 3 detectors and merged models on O2 data

- Dataset breakdown:
  675 noise, 329 s18p, 491 s18np, 115 he3.5, 1940 m39, 1139 y20, 76 s13, 1557 s25.
- Input to NNs have additional dimension (ITF)

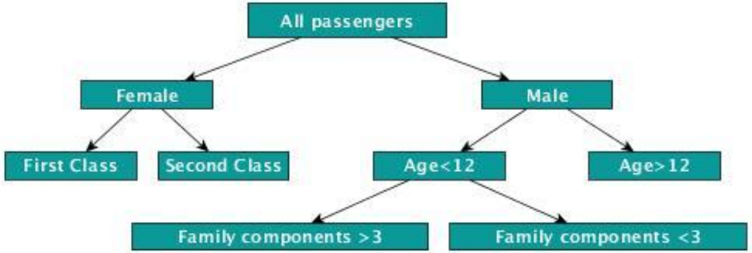




__A. Iess__, E. Cuoco, F. Morawski, C. Nicolaou, O. Lahav, accepted in A&A

Supervised Classification: eXtreme Gradient Boosting

- https://github.com/dmlc/xgboost
- Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In 22nd SIGKDD Conference on Knowledge Discovery and Data Mining, 2016
- XGBoost originates from research project at University of Washington, see also the Project Page at UW.



Tree Ensemble

$$y_n = \sum_{k=1}^{K} f_k(x_n)$$

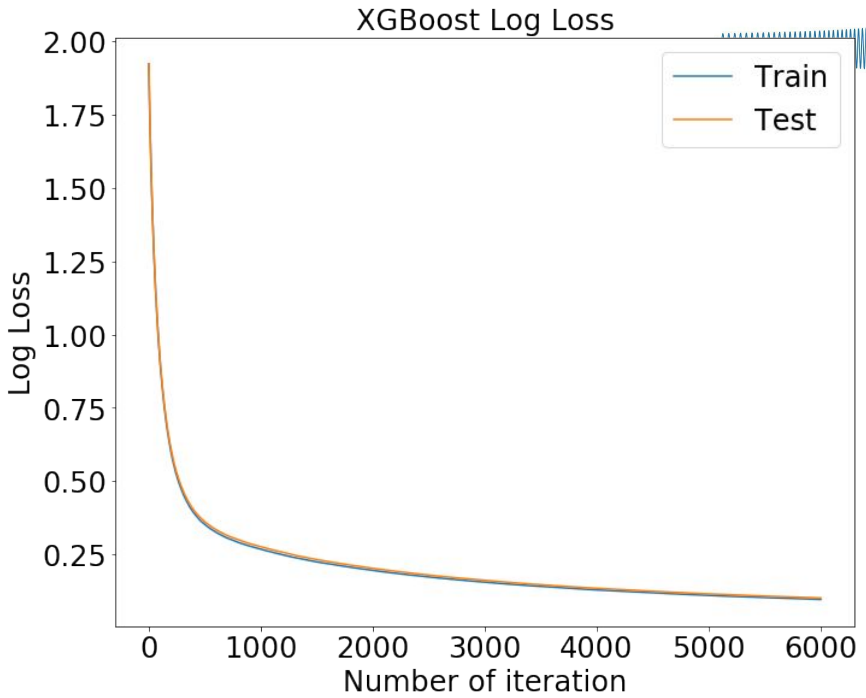
dmlc XGBoost

Xgboost performance

$$L = -\frac{1}{N}\sum_{1}^{N}((y_i \log(p_i) + (1 - y_i)(log(1 - p_i)) + \Omega$$

Cost function

Train/validation/test set: 70/15/15

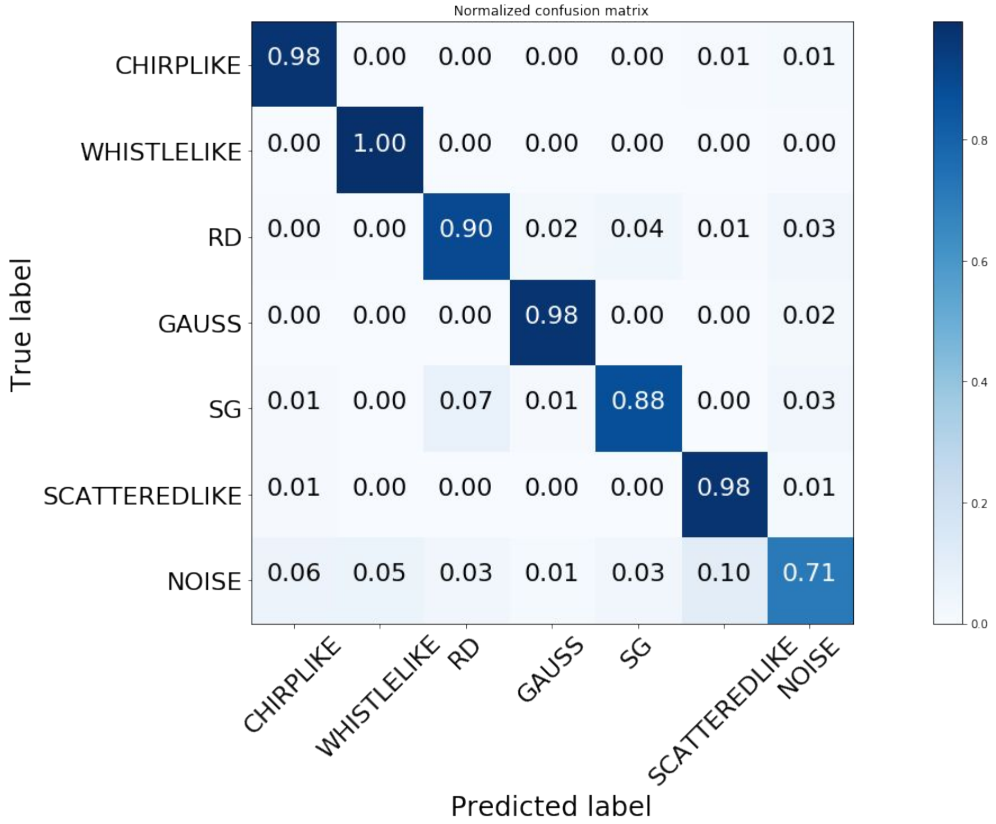


Data set Features: Wavelet coefficients

task	Classes	Learning-rate	Max_depth	estimators
Binary	2	0.01	7	5000
Multi-label	7	0.01	10	6000

WDFX Results: Multi-Label Classification

Overall accuracy
>93%



Cuoco et al.
10.23919/EUSIPCO.2018.8553393
2018 26th European Signal Processing Conference (EUSIPCO)

SCUOLA NORMALE SUPERIORE
EGO - Virgo

WDFX: Binary Classification Results

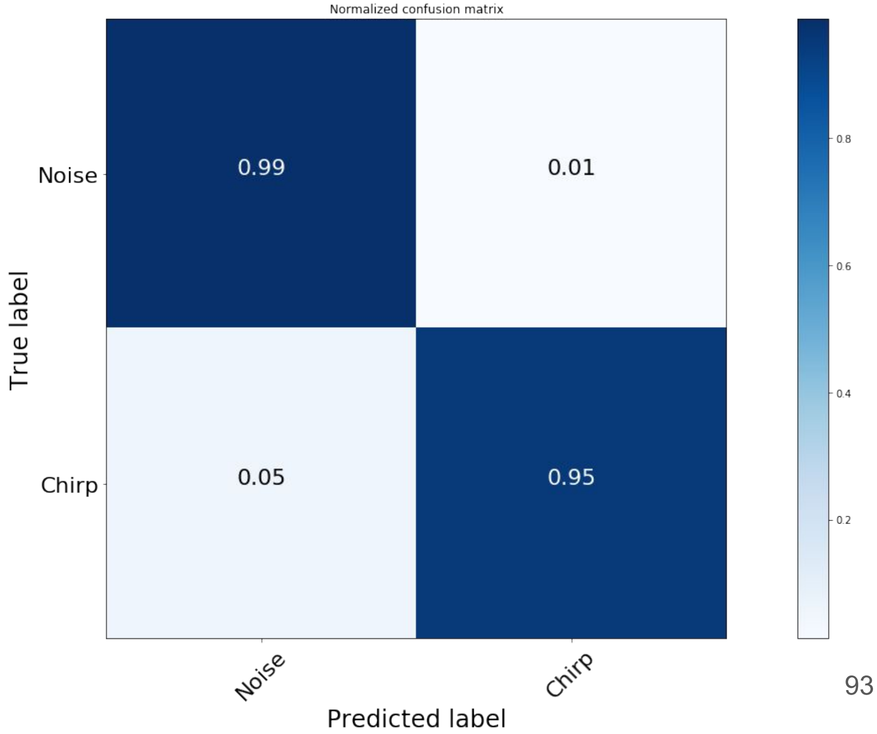
Chirp-like signals

OR

Noise

Cuoco et al.
10.23919/EUSIPCO.2018.8553393
2018 26th European Signal Processing Conference (EUSIPCO)

Overall accuracy >98%



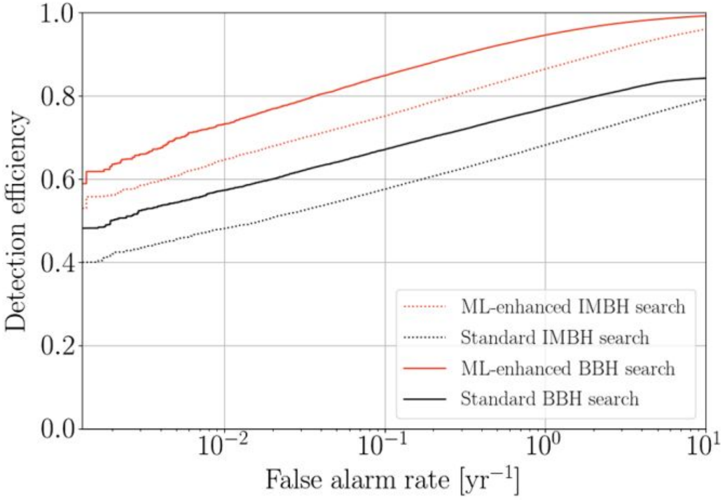
Burst search

- Decision tree-based machine learning algorithm (eXtreme-Gradient Boost) to automate signal vs. noise classification in coherent WaveBurst searches for binary black hole mergers.
- Post-processing application replacing standard veto techniques.

PHYSICAL REVIEW D **105,** 083018 (2022)

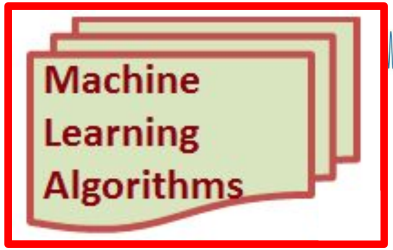
Search for binary black hole mergers in the third observing run of Advanced LIGO-Virgo using coherent WaveBurst enhanced with machine learning

T. Mishra,[1] B. O'Brien,[1] M. Szczepańczyk,[1] G. Vedovato,[2] S. Bhaumik,[1] V. Gayathri,[1] G. Prodi,[3,4] F. Salemi,[5,4] E. Milotti,[6,7] I. Bartos,[1] and S. Klimenko[1]



	Standard cWB	ML-enhanced cWB		
Event	FAR (yr^{-1})	FAR (yr^{-1})	SNR	P_{astro}
GW190408_181802	$<9.5 \times 10^{-4}$	$<1.0 \times 10^{-3}$	14.8	0.999
GW190412	$<9.5 \times 10^{-4}$	$<1.0 \times 10^{-3}$	19.7	0.999
GW190421_213856	3.0×10^{-1}	1.8×10^{-2}	9.3	0.997
GW190503_185404	1.8×10^{-3}	$<9.9 \times 10^{-4}$	11.5	0.999
GW190512_180714	3.0×10^{-1}	1.8×10^{-1}	10.7	0.941
GW190513_205428	...	$1.0 \times 10^{+0}$	11.5	0.703
GW190517_055101	6.5×10^{-3}	6.2×10^{-4}	10.7	0.999
GW190519_153544	3.1×10^{-4}	$<1.0 \times 10^{-4}$	14.0	1.000
GW190521	2.0×10^{-4}	$<1.0 \times 10^{-4}$	14.4	1.000
GW190521_074359	$<1.0 \times 10^{-4}$	$<1.0 \times 10^{-4}$	24.7	0.999
GW190602_175927	1.5×10^{-2}	$<8.8 \times 10^{-4}$	11.1	0.999
GW190701_203306	5.5×10^{-1}	1.1×10^{-2}	10.2	0.997
GW190706_222641	$<1.0 \times 10^{-3}$	$<1.1 \times 10^{-3}$	12.7	0.999
GW190707_093326	...	1.1×10^{-1}	11.2	0.976
GW190727_060333	8.8×10^{-2}	3.4×10^{-3}	11.4	0.998
GW190728_064510	...	2.6×10^{-2}	10.5	0.993
GW190828_063405	$<9.6 \times 10^{-4}$	$<1.1 \times 10^{-3}$	16.6	0.999

Supernova searches



- Genetic evolutionary algorithm to perform single-interferometer supernova searches.
- Post-processing method on top of cWB.

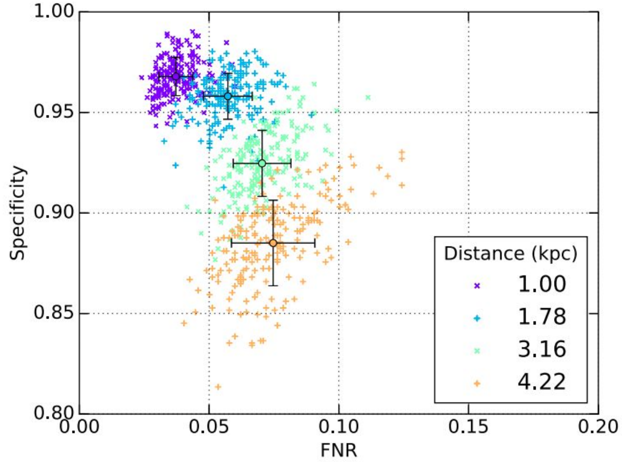
Mach. Learn.: Sci. Technol. **1** (2020) 015005 https://doi.org/10.1088/2632-2153/ab527d

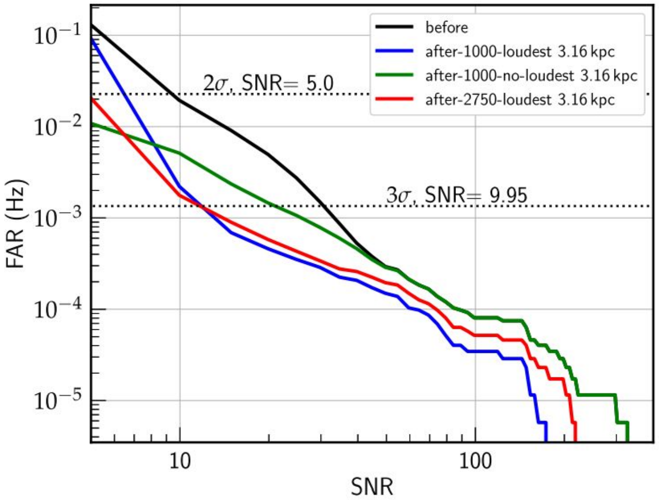
MACHINE LEARNING Science and Technology

PAPER

Improving the background of gravitational-wave searches for core collapse supernovae: a machine learning approach

M Cavaglià[1,5], S Gaudio[2], T Hansen[2], K Staats[3], M Szczepańczyk[4] and M Zanolin[2]

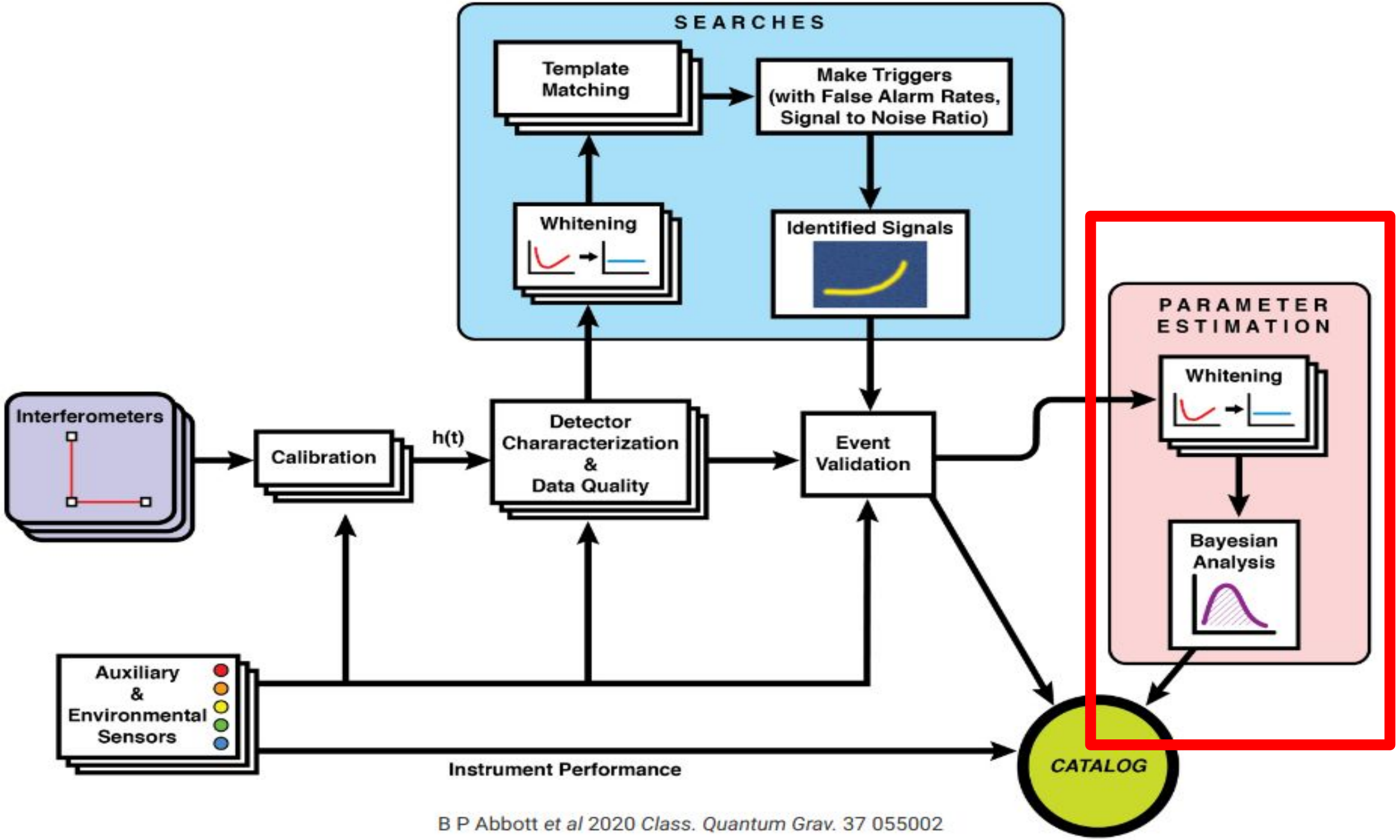




Trigger time	n_+	$P(s\|n_+)$	Actual
1137221362.849899	0	0.03	BKG
1137221296.450439	12	0.35	BKG
1137221270.478584	7	0.26	BKG
1137221270.315765	0	0.03	BKG
1137221256.461151	0	0.03	BKG
1137221254.992889	0	0.03	BKG
1137221206.790939	0	0.03	BKG
1137221187.891924	0	0.03	BKG
1137088411.819580	0	0.03	BKG
1137088400.326843	91	0.50	BKG
1137123606.447540	146	0.50	SIG (Yak, 3.16 kpc)
1137234559.739685	188	0.65	SIG (Yak, 3.16 kpc)
1137250081.748009	167	0.52	SIG (Yak, 3.16 kpc)
1137215815.308205	188	0.65	SIG (Yak, 1 kpc)
1137240747.519287	188	0.65	SIG (Yak, 1 kpc)
1137251495.131439	188	0.65	SIG(Yak, 1 kpc)
1137232392.167053	188	0.65	SIG (Yak, 1 kpc)
1137237558.365189	186	0.62	SIG (Yak, 1 kpc)

Currently extending to multi-detector and applying to O3 data

Astrophysical interpretation of GW sources



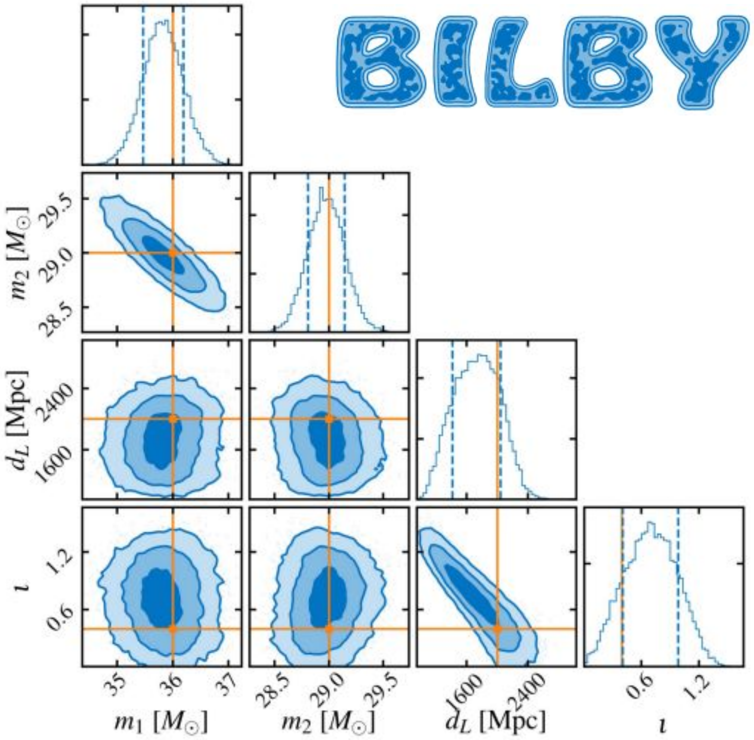
B P Abbott *et al* 2020 *Class. Quantum Grav.* 37 055002

Astrophysical interpretation of GW sources

The Astrophysical Journal Supplement Series, 241:27 (13pp), 2019 April https://doi.org/10.3847/1538-4365/ab06fc

Bilby: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy

Gregory Ashton[1,2], Moritz Hübner[1,2], Paul D. Lasky[1,2], Colm Talbot[1,2], Kendall Ackley[1,2], Sylvia Biscoveanu[1,2,3], Qi Chu[4,5], Atul Divakarla[1,2,6], Paul J. Easter[1,2], Boris Goncharov[1,2], Francisco Hernandez Vivanco[1,2], Jan Harms[7,8], Marcus E. Lower[1,9,10], Grant D. Meadors[1,2], Denyz Melchor[1,2,11], Ethan Payne[1,2], Matthew D. Pitkin[12], Jade Powell[9,10], Nikhil Sarin[1,2], Rory J. E. Smith[1,2], and Eric Thrane[1,2]



PHYSICAL REVIEW D 91, 042003 (2015)

Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library

J. Veitch,[1,2,*] V. Raymond,[3] B. Farr,[4,5] W. Farr,[1] P. Graff,[6] S. Vitale,[7] B. Aylott,[1] K. Blackburn,[3] N. Christensen,[8] M. Coughlin,[9] W. Del Pozzo,[1] F. Feroz,[10] J. Gair,[11] C.-J. Haster,[1] V. Kalogera,[5] T. Littenberg,[5] I. Mandel,[1] R. O'Shaughnessy,[12,13] M. Pitkin,[14] C. Rodriguez,[5] C. Röver,[15,16] T. Sidery,[1] R. Smith,[3] M. Van Der Sluys,[17] A. Vecchio,[1] W. Vousden,[1] and L. Wade[12]

Publications of the Astronomical Society of the Pacific, 131:024503 (16pp), 2019 February https://doi.org/10.1088/1538-3873/aaef0b

PyCBC Inference: A Python-based Parameter Estimation Toolkit for Compact Binary Coalescence Signals

C. M. Biwer[1,2], Collin D. Capano[3], Soumi De[2], Miriam Cabero[3], Duncan A. Brown[2], Alexander H. Nitz[3], and V. Raymond[4,5]

Rapid and accurate parameter inference for coalescing, precessing compact binaries

J. Lange,[1] R. O'Shaughnessy,[1] and M. Rizzo[1]
[1]Center for Computational Relativity and Gravitation, Rochester Institute of Technology, Rochester, New York 14623, USA

- Current parameter estimation techniques for compact binary coalesce signals rely on Bayesian analysis (posteriors + evidence).
- Computationally costly!
- Need to dramatically speed up the process!
- How can machine learning help?

Rapid inference of source parameters

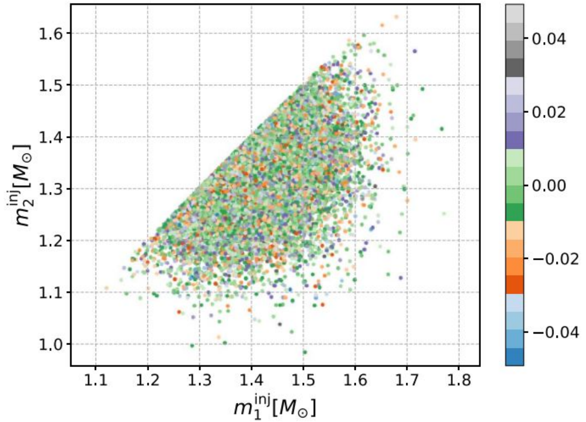
- Classifiers (Kneighbors, genetic, random forests) for HasNS and HasRemnant properties of sources in low-latency
- Train and test on LIGO-Virgo online MDC
- Integrate in the LVK low-latency infrastructure and run in O4

The Astrophysical Journal, 896:54 (10pp), 2020 June 10
https://doi.org/10.3847/1538-4357/ab8dbe

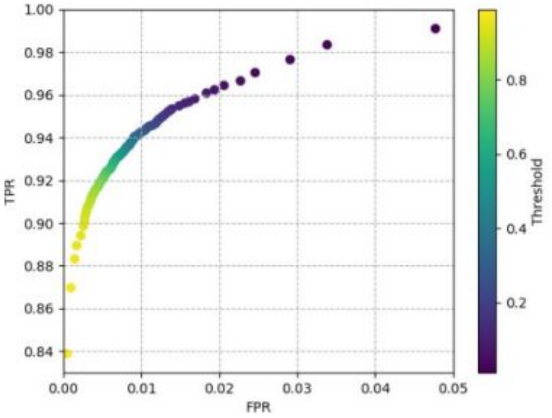
A Machine Learning-based Source Property Inference for Compact Binary Mergers

Deep Chatterjee[1], Shaon Ghosh[1,2], Patrick R. Brady[1], Shasvath J. Kapadia[1,3], Andrew L. Miller[4], Samaya Nissanke[5], and Francesco Pannarale[6,7]
[1] Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, WI 53211, USA
[2] Department of Physics and Astronomy, Montclair State University, 1 Normal Avenue, Montclair, NJ 07043, USA
[3] International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore 560012, India
[4] Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, Chemin du Cyclotron, 2 B-1348 Louvain-la-Neuve, Belgium
[5] GRAPPA, Anton Pannekoek Institute for Astronomy and Institute of High-Energy Physics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[6] Dipartimento di Fisica, Università di Roma "Sapienza," Piazzale A. Moro 5, I-00185 Roma, Italy
[7] INFN Sezione di Roma, Piazzale A. Moro 5, I-00185 Roma, Italy
Received 2019 October 31; revised 2020 April 24; accepted 2020 April 26; published 2020 June 12

Deploy Selected Model



O3 event	p(HasNS)	p(HasRemnant)
GW190425	0.999	0.9959
GW190426	0.9676	0.0029
GW190421	0.0057	0.0012
GW190915	0.0057	0.0012
GW200115	0.967	0.0029
GW20012	0.0057	0.0012



See also:
S. Sharma Chaudhary, MC, D. Chatterjee, S. Ghosh, in preparation

Parameter estimation

Classical and Quantum Gravity

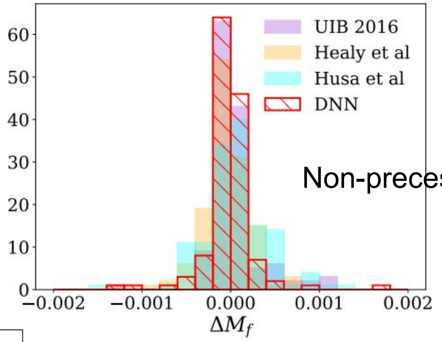
PAPER

Predicting the properties of black-hole merger remnants with deep neural networks

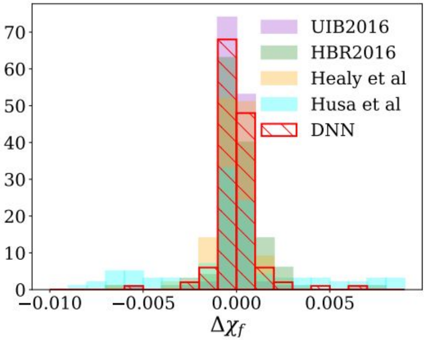
L Haegel[3,1,2] and S Husa[1]

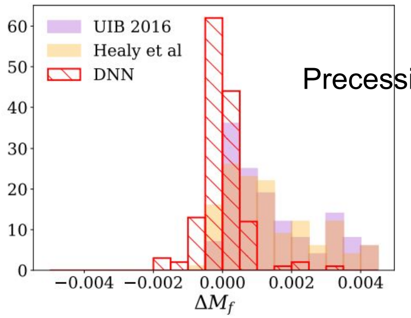
Published 10 June 2020 • © 2020 IOP Publishing Ltd
Classical and Quantum Gravity, Volume 37, Number 13
Citation L Haegel and S Husa 2020 *Class. Quantum Grav.* **37** 135005

- Deep neural networks to infer the relationship between the initial BBHs parameters and the remnant final mass and fina spin of a binary black hole merger.
- Trained with publicly available NR catalogs.

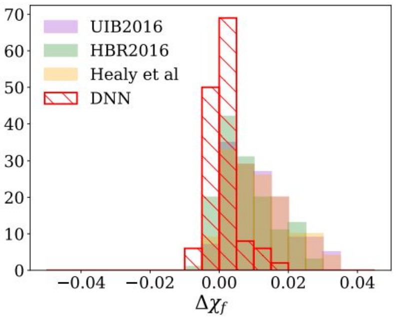


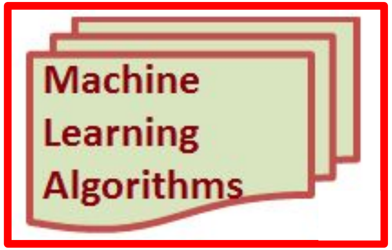
Non-precessing





Precessing





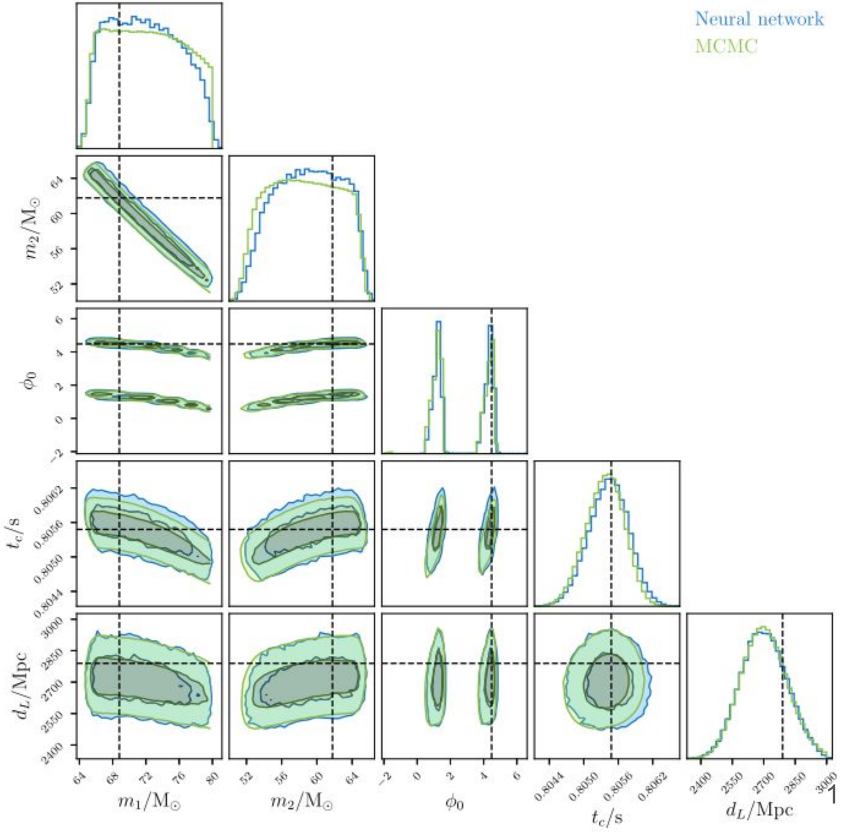
Parameter estimation

PHYSICAL REVIEW D **102,** 104057 (2020)

Gravitational-wave parameter estimation with autoregressive neural network flows

Stephen R. Green,[1,*] Christine Simpson,[2,†] and Jonathan Gair[1,‡]

- Autoregressive normalizing flows for rapid likelihood-free inference of binary black hole system parameters.
- Maps a **multivariate standard normal distribution into the posterior distribution of system parameters.**
- Performance comparable to Markov chain Monte Carlo.



Parameter estimation

ARTICLES
https://doi.org/10.1038/s41567-021-01425-7
nature physics

Check for updates

Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy

Hunter Gabbard[1], Chris Messenger[1], Ik Siong Heng[1], Francesco Tonolini[2] and Roderick Murray-Smith[2]

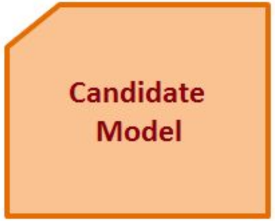
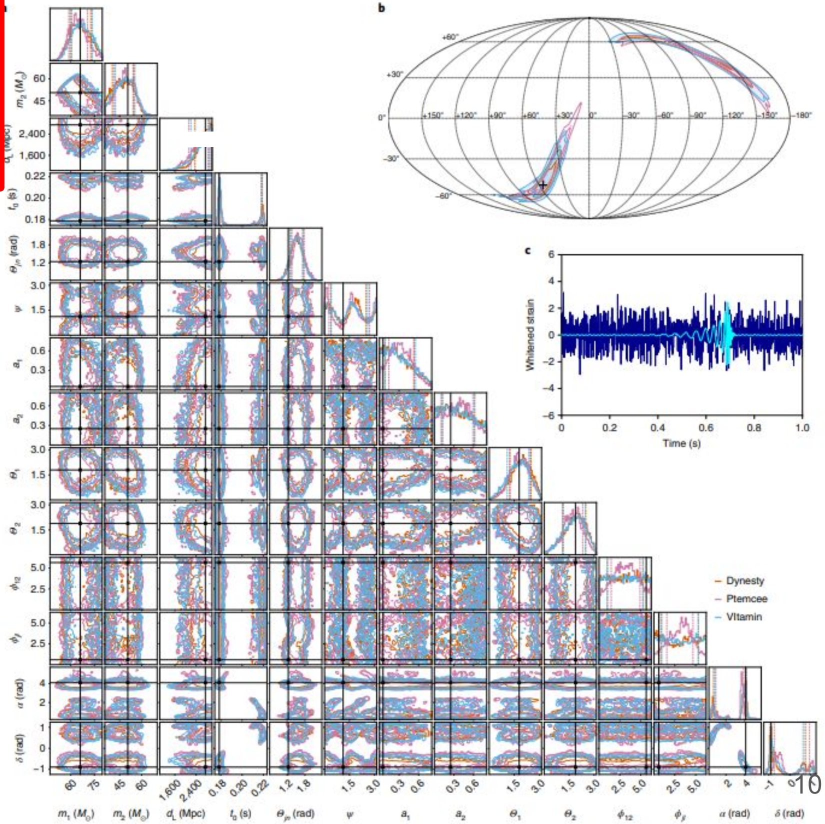


Table 2 | Durations required to produce samples from each of the sampling approaches

Sampler	Run time (s)			Ratio $\frac{\tau_{\text{VItamin}}}{\tau_X}$
	Min.	Max.	Median	
Dynesty[a7]	21,564	261,268	45,607[b]	2.2×10^{-6}
emcee[8]	16,712	39,930	19,821	5.1×10^{-6}
ptemcee[9]	2,392	501,632	41,151.0	2.4×10^{-6}
CPNest[6]	10,309	437,008	83,807	1.2×10^{-6}
VItamin[c]	**1×10^{-1}**			1

- Pre-trained conditional variational autoencoder
- Standard advanced detector power spectral density.
- Full-parameter estimation ~ 1 s.



SCUOLA NORMALE SUPERIORE
EGO - Virgo

Machine learning applications in LVK: a long list

Glitches classification	GW signal detection	Parameter estimation	Sky localization
Easy access information	Data quality	Waveform modelling	…

Review paper: Enhancing gravitational-wave science with machine learning Elena Cuoco *et al* 2021 *Mach. Learn.: Sci. Technol.* 2 011002…We are preparing the updated version

Hands-on session

- GW Data reading and plotting
- Whitening procedure
- CNN-1D classification for CCSN signals
- CNN-2D classification for CCSN signals
- Code repo: https://github.com/elenacuoco/GWML-masterclass
  - please download the code and try to run CCN-1D on your pc, to setup the data sets.

Thank you

twitter: @elenacuoco
elena.cuoco@ego-gw.it