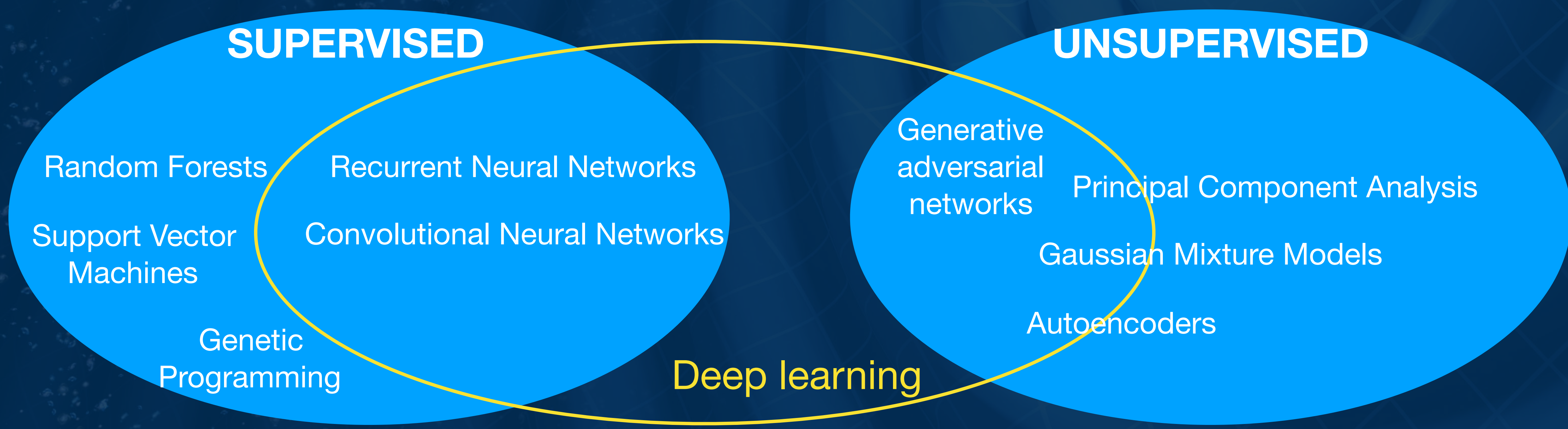
Overview about machine learning in Gravitational Wave astronomy

A. Trovato
CNRS-APC

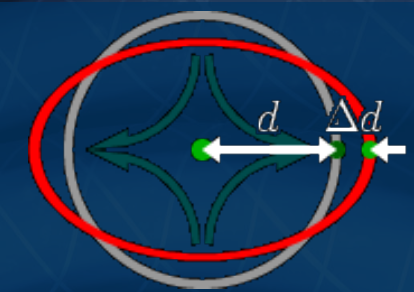
Reminder: Machine Learning

- Machine learning (ML): algorithms that can learn from data and make predictions on it
  - ✓ SUPERVISED: classification and regression by learning from labeled data
  - ✓ UNSUPERVISED: learning from data without labels
- Deep learning: subfield of machine learning
  - ✓ use of raw data (no feature engineering)
  - ✓ typically based on artificial neural networks

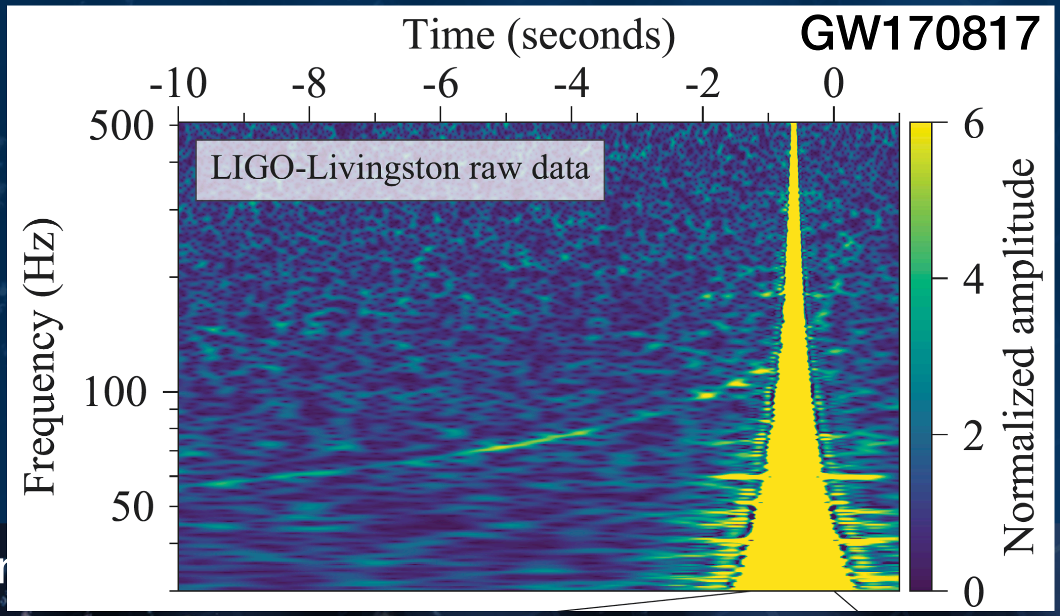


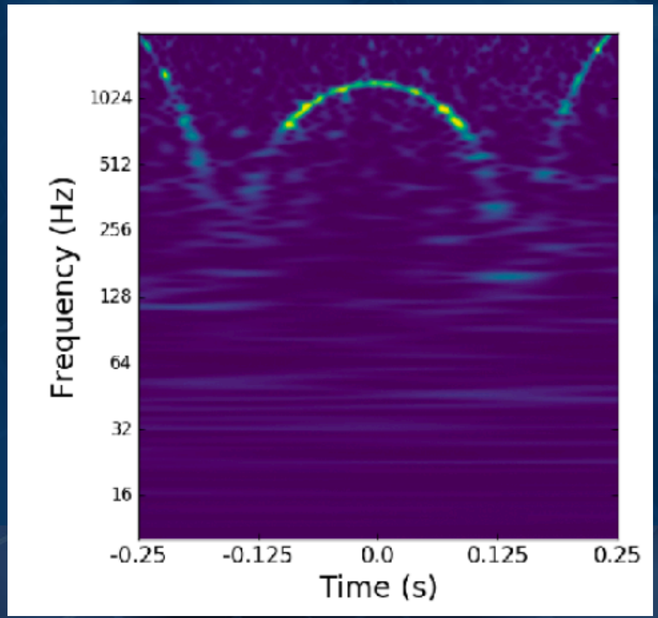
Why ML in GW interferometers (I)?

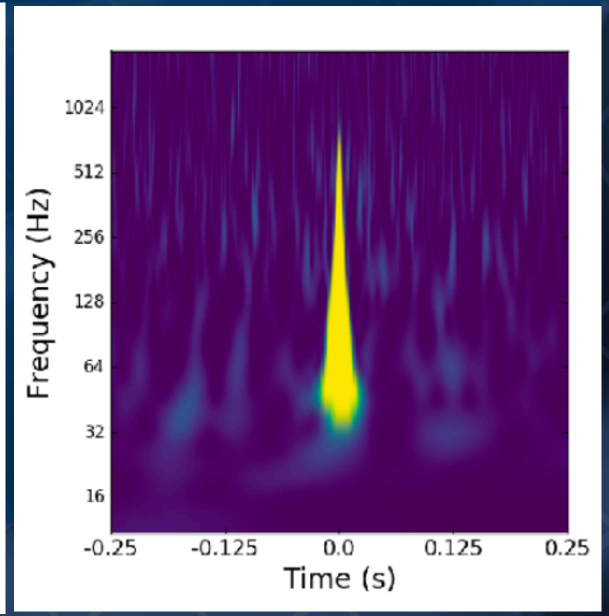
- “Data” in a gravitational wave interferometer:
  - ✓ Strain: h(t)
    - It is a <u>time series</u>
    - It is highly contaminated by noise:
      - <u>Stationary noise</u> (detector sensitivity, detector upgrade and tuning)
      - <u>Transient Noise or glitches</u>: short duration artefacts that can obscure or mimic the gravitational wave signal
        - ➡ Glitches vary widely in duration, frequency range and morphology
        - ➡ No statistical model is able to capture the complexity of the glitch population
        - ➡ <u>Separating the glitches from the astrophysical signal is a challenging task that could be achieved with machine learning algorithms!</u>



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

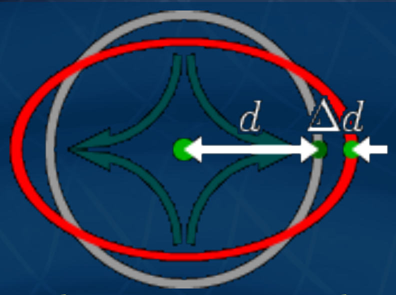






Why ML in GW interferometers (I)?

- "Data" in a gravitational wave interferometer:



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

  - ✓ Strain: h(t)
    - It is a time series
    - It is highly contaminated by noise:
      - Stationary noise (detector sensitivity, detector upgrade and tuning)
      - Transient Noise or glitches: short duration artefacts that can obscure or mimic the gravitational wave signal
        - ➡ Glitches vary widely in duration, frequency range and morphology
        - ➡ No statistical model is able to capture the complexity of the glitch population
        - ➡ Separating the glitches from the astrophysical signal is a challenging task that could be achieved with machine learning algorithms!
  - ✓ Auxiliary channels (Hundreds of thousands auxiliary data streams: monitors status of the detector and of its physical environment)
    - Could provide information about the source of the glitches and coupling with h(t)
    - Detector characterisation
    - Deep-learning algorithms: in principle able to learn and evidence non-linear couplings

Glitches identification and subtraction

- Some ML methods tested:
  - ✓ Convolutional neural networks:
    - Used to classify of glitches —> only strain info
    - Glitches represented usually as spectrograms
  - ✓ Tree-based algorithms (Genetic programming and Random Forests)
    - Determine the origin of glitches though the auxiliary channels
    - Test with two classes of transients with known instrumental origin
    - Use features of the time series
  - ✓ Long Short Term Memory (LSTM)
    - Data denoising using time evolution of the auxiliary channels
    - Test done on calibration lines and 60 Hz mains with few selected witness channels

Why ML in GW interferometers (II)?

- Data analysis: extract small signal buried into a much larger noise

Modelled Searches	Un-modelled Searches
> Matched filtering • Time consuming • Specific waveform families/type of objects > Machine learning • Faster —> could allow to test wider parameter space/more objects	> “Burst” searches • More sensible to glitches > Machine learning • Could provide techniques intermediate between modelled and unmodeled • Better glitch discrimination

- Big-data: ML algorithms could provide a fast way to analyse the increasing volume of data produced by GW detectors

ML for gravitational wave detectors

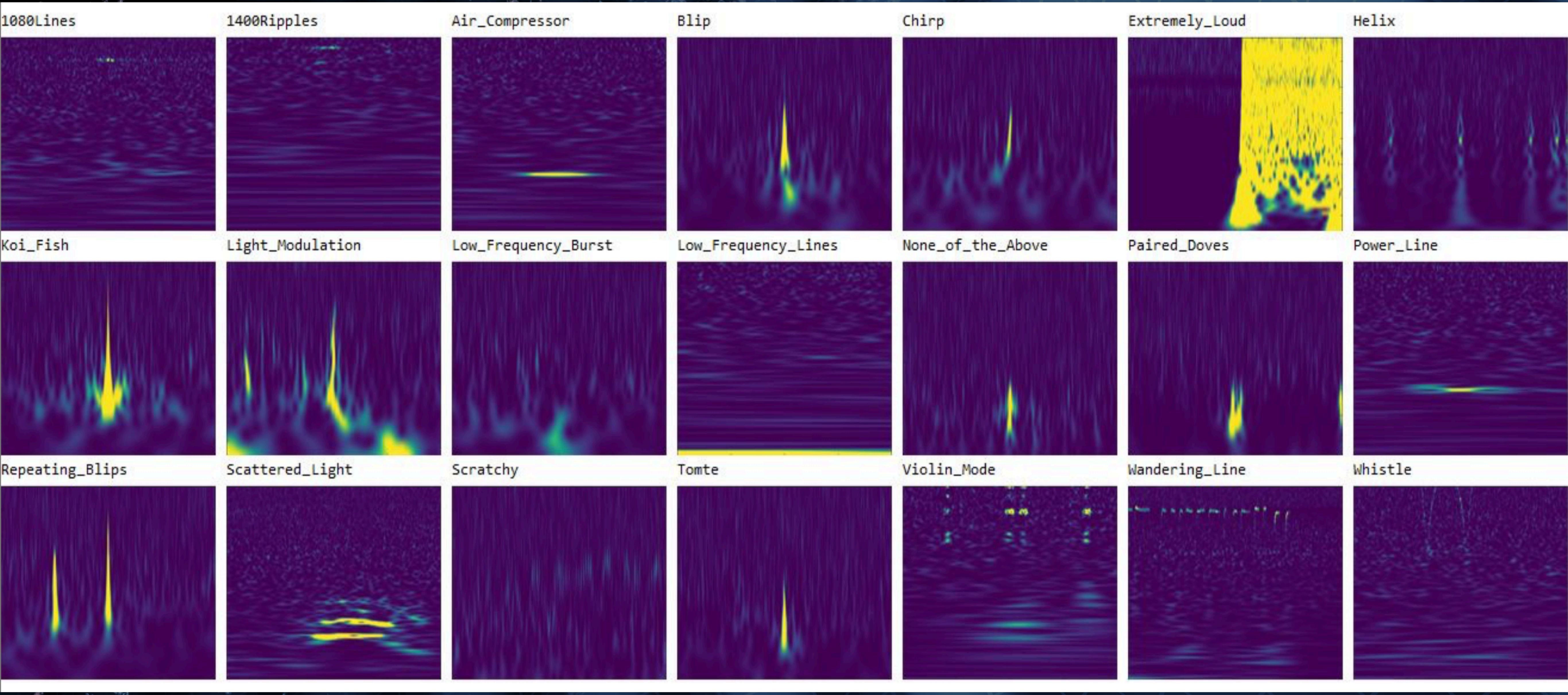
- Main applications:
  - ✓ Transient Noise classification and subtraction
    - citizen science (Gravity Spy, Class. Quantum Gravity 34 (2017) 064003)
    - Class. Quantum Grav. 32 (2015) 215012; Class. Quantum Grav. 34 (2017) 034002; Class. Quantum Grav. 35 (2018) 095016; Phys. Rev. D 95 (2017) 104059; Phys. Rev. D 97 (2018) 101501(R)
  - ✓ Astrophysical signal searches
    - Detecting Compact Binary Coalescence (Phys. Rev. Lett. 120 (2018) 141103; Physics Letters B 778 (2018) 64; Phys. Rev. D 91 (2015) 062004; Phys. Rev. D 96 (2017) 104015)
    - Detecting Bursts (Class. Quantum Grav. 32 (2015) 245002;)
    - Supernova searches (2 papers in preparation)
    - Continuous waves searches (papers in prearation)
  - ✓ Parameter estimation (Mon. Not. Roy. Astron. Soc. 421 (2012) 169; Phys. Rev. D 97 (2018) 044039;)
  - ✓ System Control
    - Lock acquisition, beam spot position control
- Data
  - ✓ strain h(t) and/or auxiliary channels
  - ✓ Data representation: spectrograms or time series

Gravity Spy

- Citizen Science and Glitch Classification
- LIGO looks for variation 1000 times smaller than the diameter of a proton.
  - ✓ Susceptible to a great deal of instrumental and environmental noise.
  - ✓ Transient noise (glitches) can mimic astrophysical signals.
  - ✓ Machine Learning is incredibly helpful, but does not alone resolve all noise.
- Human intuition remains a valuable asset in the effort to characterise glitches
- Web + glitch classification + Deep Learning = citizen science done right!
- Two way path: human classification provides labeled classes as training data for ML + LIGO glitches classified by CNN with lowest “confidence score” back to the citizen scientists for further analysis = final label.

Gravity Spy dataset

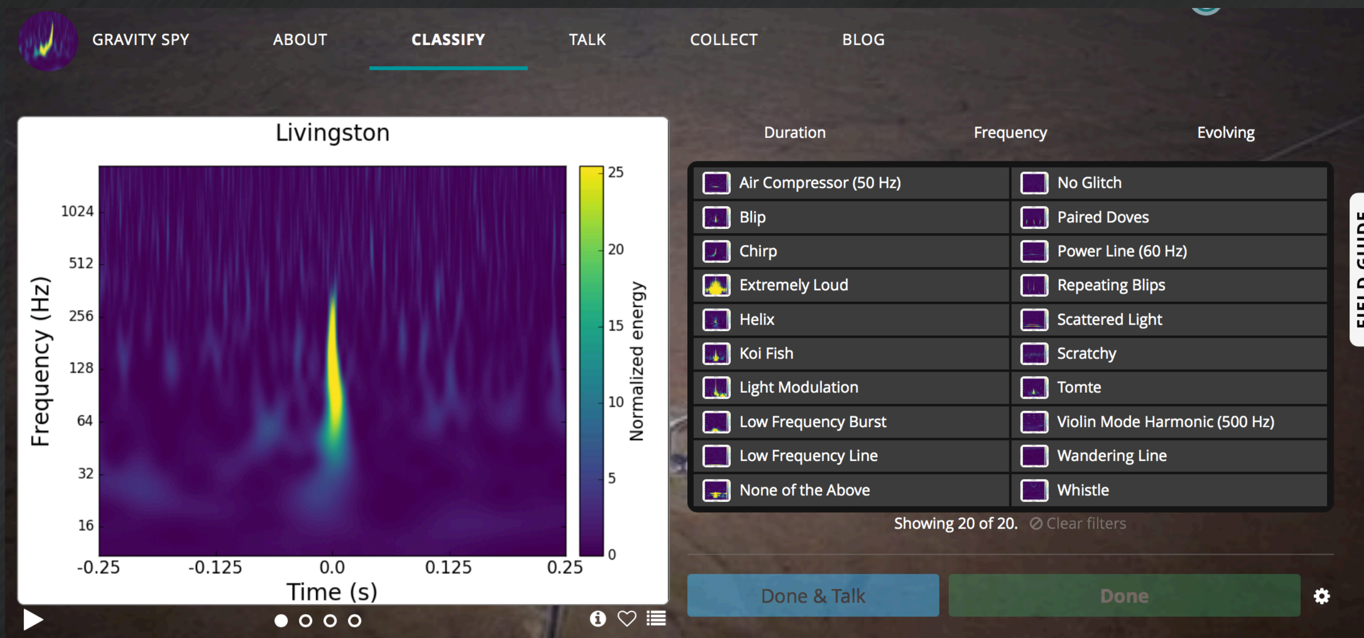
- Glitches represented as spectrograms
- Only high SNR (SNR>7.5)



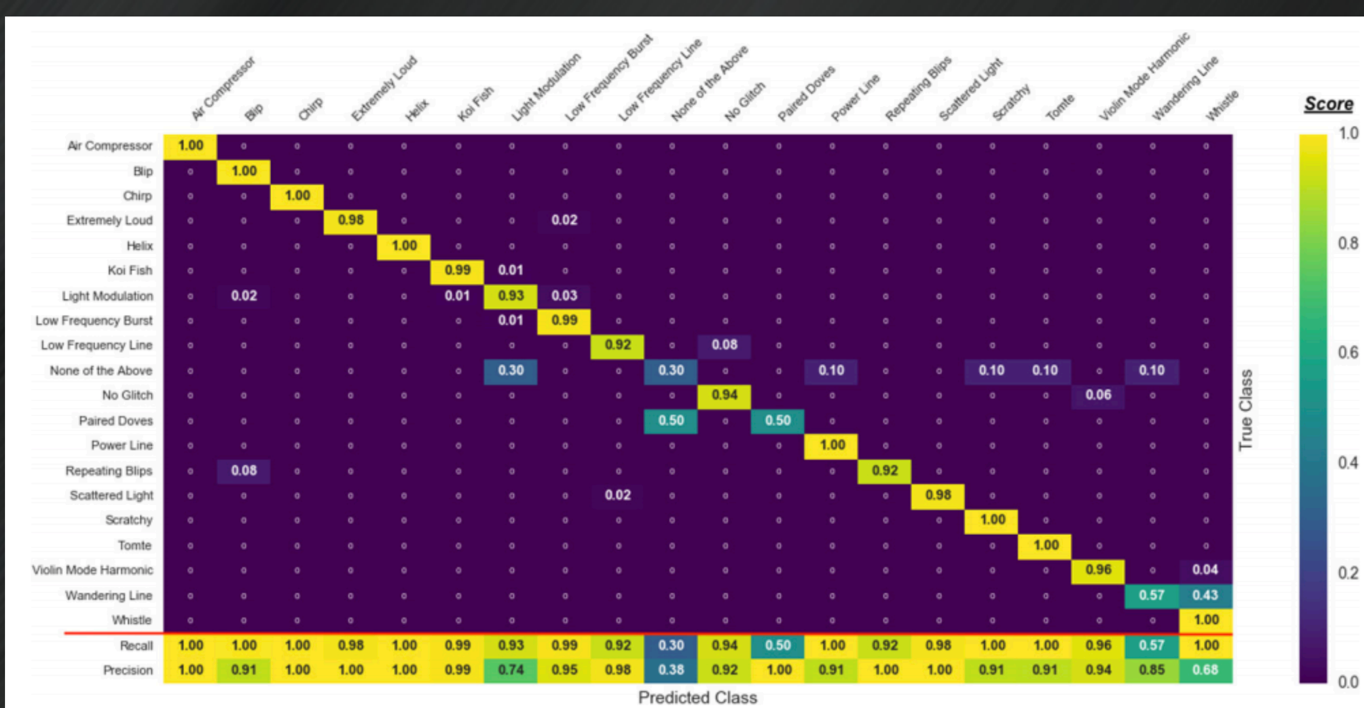
GRAVITY SPY

CALIFORNIA STATE UNIVERSITY FULLERTON · LSC LIGO Scientific Collaboration · NSF · Northwestern CIERA · SYRACUSE UNIVERSITY School of Information Studies

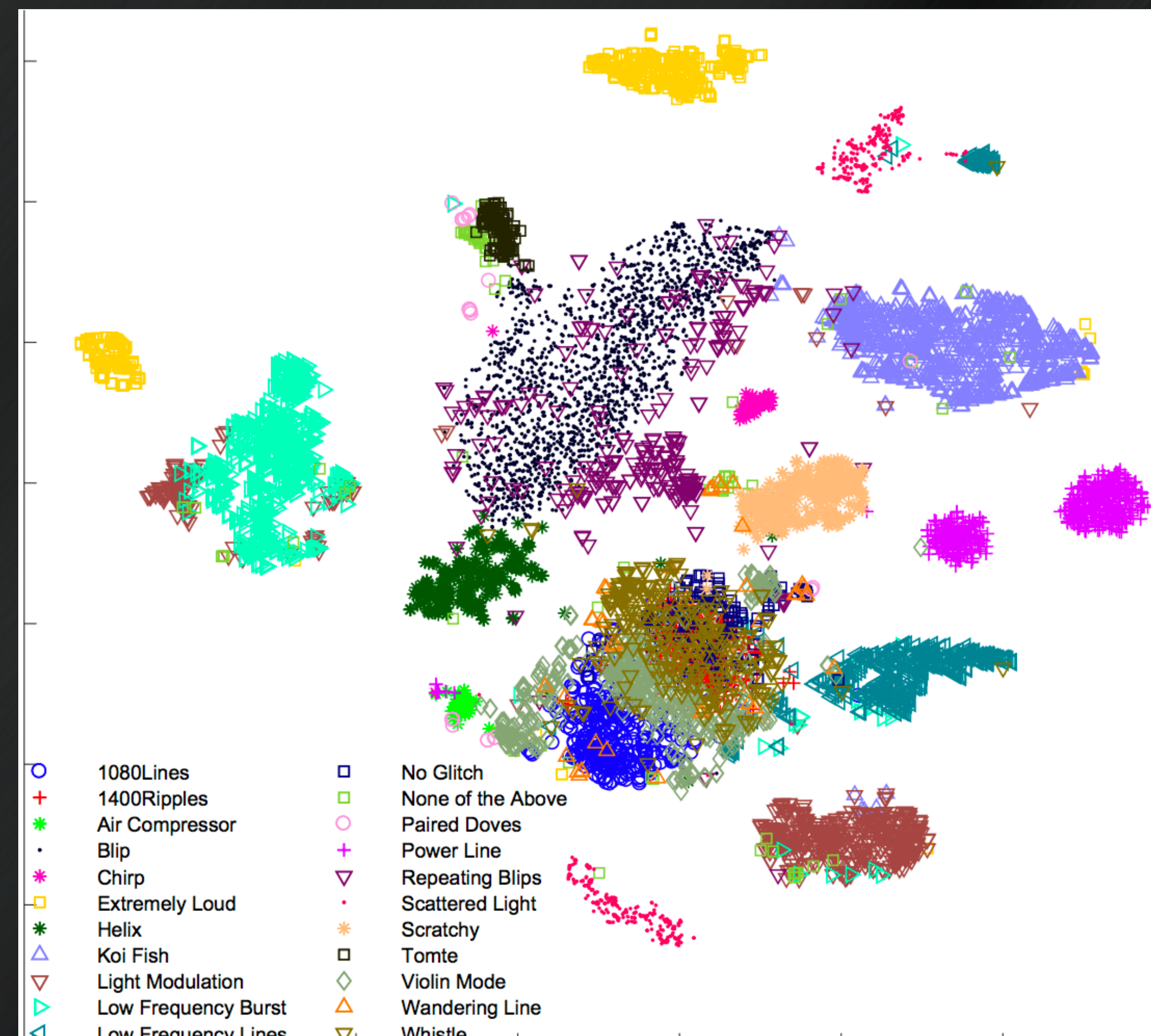
LVC Members: Scotty Coughlin, Mike Zevin, Josh Smith, Andy Lundgren, Duncan MacLeod, Vicky Kalogera



LIGO Citizen Science: >2.9 million classifications



Combining Machine Learning & Crowdsourcing

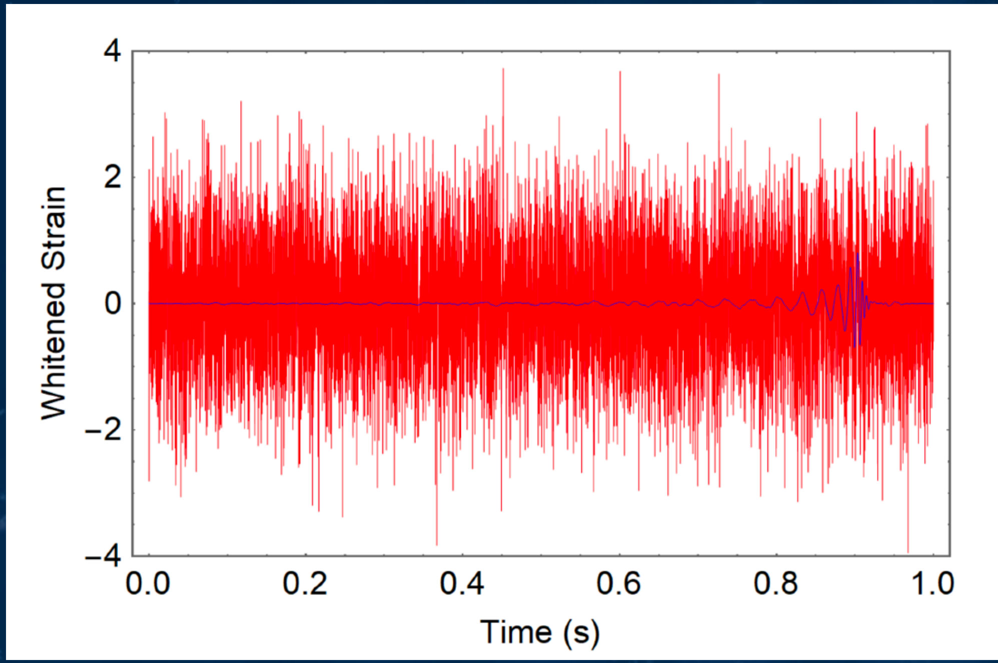


Clustering algorithms / identification of novel glitch classes

arXiv: 1611.04596   www.gravityspy.org   @gravityspyzoo

Deep Filtering

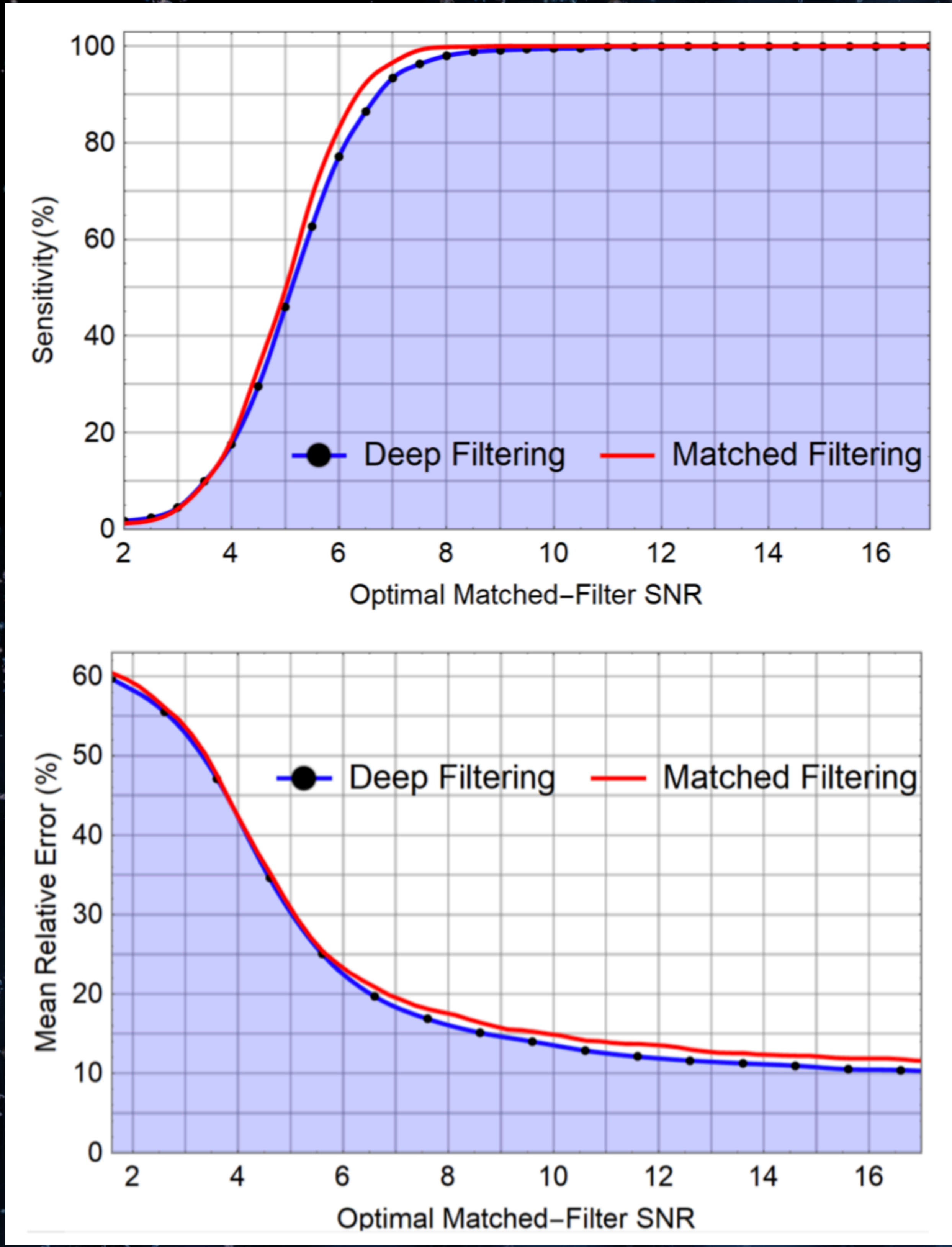
- Autors: Daniel George & E. A. Huerta Physics Letters B 778 (2018) 64, Phys. Rev. D 97 (2018) 044039;
- Data: raw time series (no spectrograms)



Example of LIGO whitened data (noise) + injected signal (SNR=7.5)

- System of two deep convolutional neural networks:
  1. Find a signal in a highly noisy time-series data stream
  2. Estimate parameters of signals
- Training with injections of GW templates originating from quasi-circular, non-spinning, stellar mass BBH systems

Deep Filtering (plots)



- 100% sensitivity when SNR is greater than 10 (false alarm rate tuned to be less than 1%, i.e., 1 per 100 seconds of noise in the test set was classified as signals

- Errors follow a Gaussian distribution for each region of the parameter space for SNR greater than 10.
- Deep Filtering error < 5% for SNR>50
- Matched-Filtering error with same template bank is always > 11%

Deep Filtering (conclusion)

- Similar performance compared to matched-filtering while being several orders of magnitude faster
- Demo of real-time detection of GW150914 https://www.youtube.com/watch?v=87zEll_hkBE&feature=youtu.be
- Not integrated in the current pipelines but promising results
- Possible extension for online searches
- Tests ongoing with other classes of signals

Conclusion

- ML in GW astronomy looks very promising!
- It will be fundamental for the next future and to handle in a fast way the high volume of data we expect

BUT:

– None of the ML methods have been integrated in the current detection pipelines/production

– Tests with a limited volume of data

– Frequency-time representations usually used: not necessarily optimal

– Auxiliary channels mainly used for detector characterisation/ denoising —>no application for signal detection