Finding gravitational waves with artificial intelligence

Michał Bejger

Ferrara, 19.3.24

INFN FERRARA

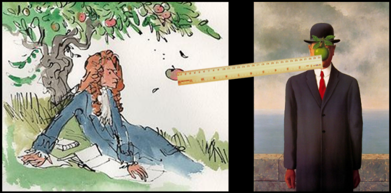
LIGO VIRGO KAGRA

G2NET

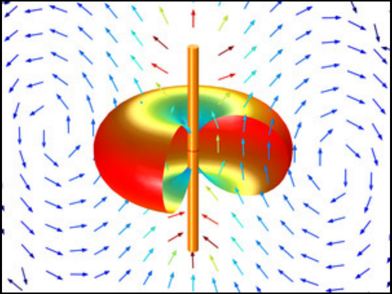
- ⋆ Gravitational-wave detectors and data,
- ⋆ ”Traditional” data analysis techniques,
- ⋆ New approaches: neural networks,
- ⋆ Results and outlook.

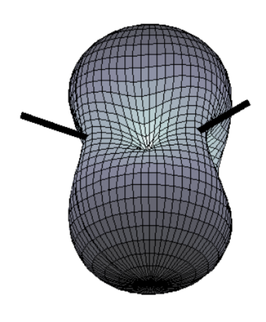
LIGO-Virgo global detector network

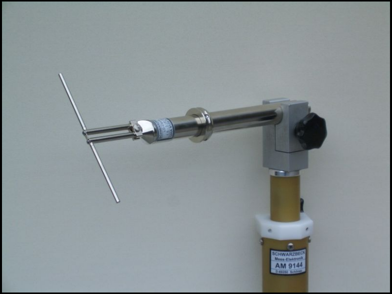
Very precise rulers: measuring distances between free-falling bodies with laser light.



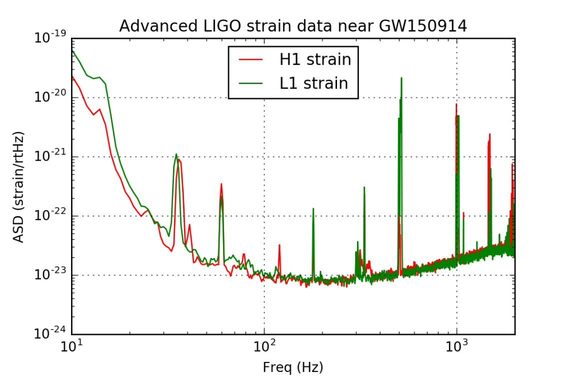
Interferometer = GW antenna





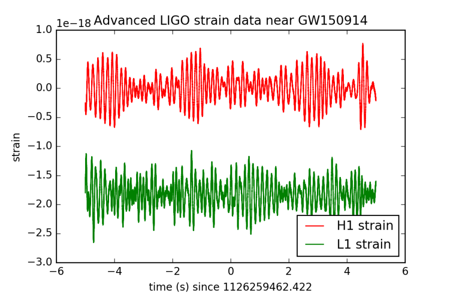


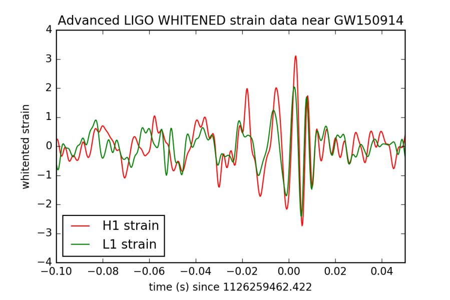
Sensitivity: amplitude spectral density of the noise



- ⋆ *Plot dominated by instrumental noise, lines*: mirror suspension resonances at 500 Hz and harmonics, calibration lines and power lines (60 Hz and harmonics) etc.,
- ⋆ $ASD = \tilde{x}(f) = \frac{1}{\sqrt{T}} \int_0^T x(t) \exp(-2i\pi f t) dt$ (units: $[1/\sqrt{Hz}]$)
- ⋆ One detector produces a stream of data (”main” and auxiliary channels) with 50 MB/s

How the data looks like





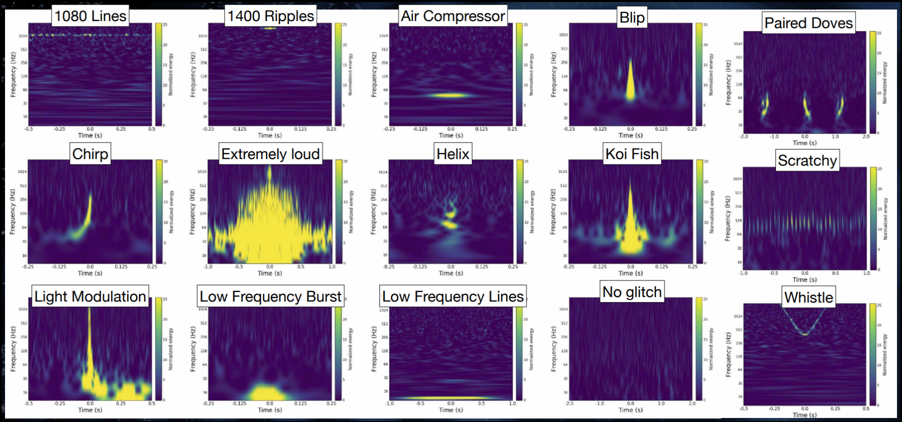
The data are dominated by the **low frequency noise** (L1 offset by -2×10^{-18} due to very low frequency oscillations).

Some usual data treatment:

- ⋆ Whitening (dividing the data by the noise ASD in the Fourier domain),
- ⋆ filtering the frequencies outside the desired band with bandpass filter,
- ⋆ suppressing the instrumental lines.

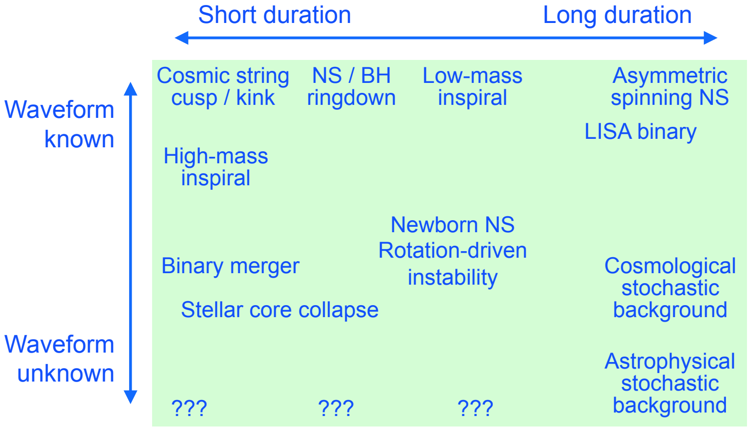
"Glitches zoo": transient instrumental noise

Excess power (glitches) represented as spectrograms - time-frequency maps - suitable for human-eye inspection:



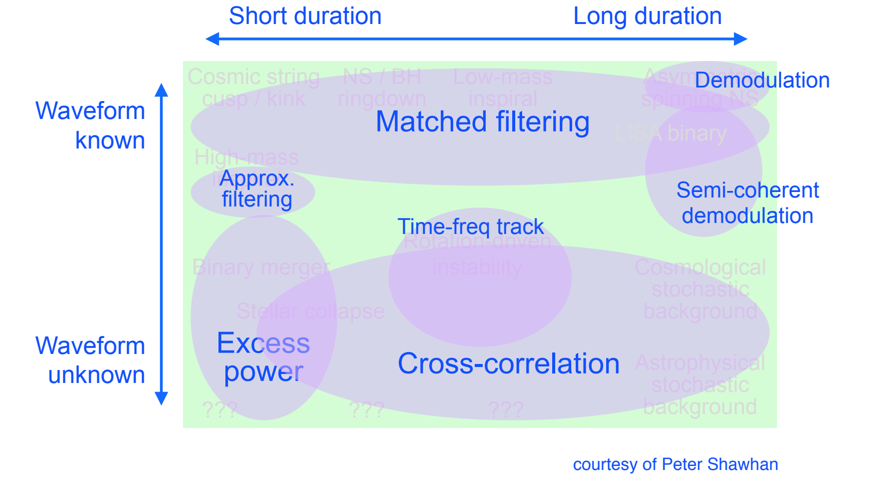
- ⋆ Main problem for the sensitivity of transient searches,
- ⋆ Citizen science: `Gravity Spy`, `Reinforce` (preparation of training data for machine learning).

Taxonomy of signal and search types

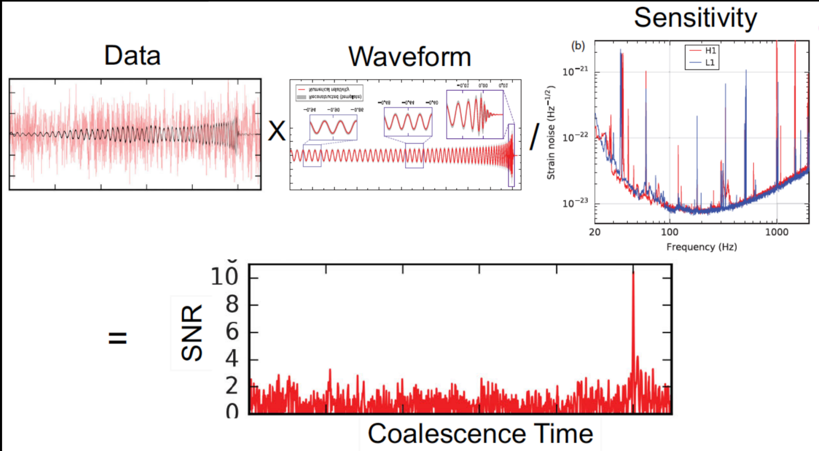


courtesy of Peter Shawhan

Taxonomy of signal and search types



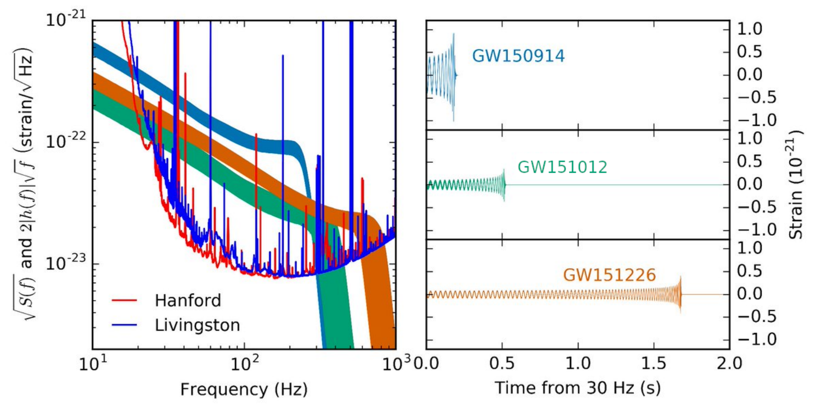
Matched filter in pictures



$$C(t) = (x|h) = 4\int_0^\infty \frac{\tilde{x}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi i f t} df, \quad \text{with } S_n(f) \text{ (noise PSD).}$$

Matched filter SNR: $\rho = (x|h)/\sqrt{(h|h)}$, Optimal SNR: $\rho_{opt} = \sqrt{(h|h)}$

LIGO-Virgo O1 3 events

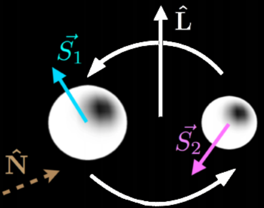


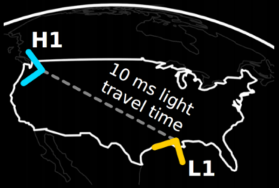
Signal-to-noise : $\rho^2_{opt} = \int_0^\infty \left(\frac{2|\tilde{h}(f)|\sqrt{f}}{\sqrt{S_n(f)}} \right)^2 d\ln(f)$

(GW150914: $\rho \simeq 24$, GW151226: $\rho \simeq 13$, GW151012: $\rho \simeq 10$)

Binary system waveform: 15+ parameters

- Intrinsic:
  - masses
  - spins
  - tidal deformability





Credit: LIGO/Virgo

- Extrinsic:
  - Inclination, distance, polarisation
  - Sky location
  - Time, reference phase

$$f_{\rm GW}^{-8/3}(t) = \frac{(8\pi)^{8/3}}{5}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{5/3}(t_c - t) + \text{higher order corrections}, \qquad \mathcal{M}_c = \frac{(M_1 M_2)^{3/5}}{(M_1+M_2)^{1/5}}.$$

$$h(r) \propto \mathcal{M}_c^{5/3} f_{\rm GW}^{2/3}/r.$$

Astrophysically-interesting parameters

Intrinsic:

- ⋆ Chirp mass $\mathcal{M} = \left(\mu^3 M^2\right)^{1/5} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$,

- ⋆ Mass ratio $q = m_2/m_1$ (at 1PN), alternatively $\nu = m_1 m_2/(m_1 + m_2)^2$,

- ⋆ Spin-orbit and spin-spin coupling (at 2PN and 3PN, resp.) $\rightarrow$

$$\chi_{eff} = (m_1 \chi_{1z} + m_2 \chi_{2z})/(m_1 + m_2)$$

  where χ_{iz} are spin components along system's total angular momentum,

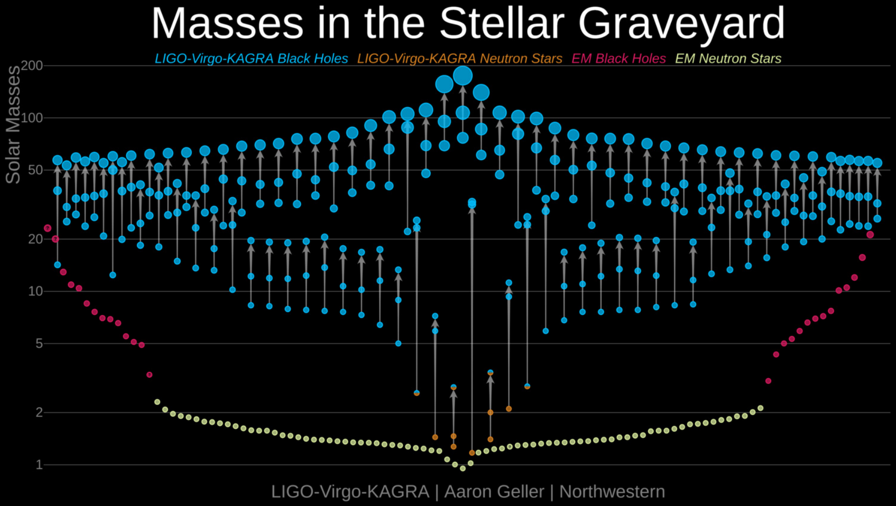
- ⋆ Tidal deformability Λ (at 5PN) $\rightarrow$

$$\tilde{\Lambda} = \frac{16}{13} \frac{(m_1 + 12 m_2) m_1^4 \Lambda_1}{(m_1 + m_2)^5} + (1 \leftrightarrow 2), \qquad \mathcal{R} = 2 \mathcal{M} \tilde{\Lambda}^{1/5}$$

Extrinsic:

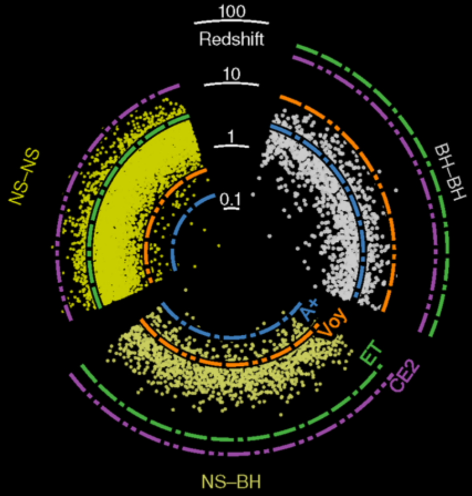
- ⋆ Direct "luminosity" ("loudness") distance: binary systems are "standard sirens".

Compact objects population in GWTC-3 (O1, O2, O3)



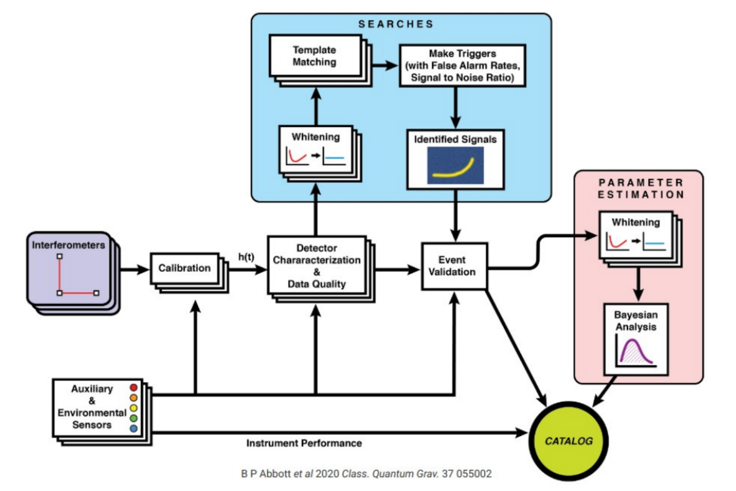
O4 predictions: $\simeq 1.5^3$ more events ($\simeq 300$, one per a few days?)

Einstein Telescope, Cosmic Explorer (2030+)



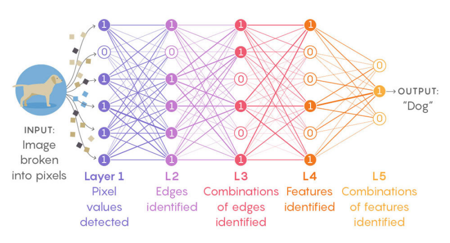
→ *Hundreds* events/day, overlapping signals, new signal types.

Areas for automation and machine learning

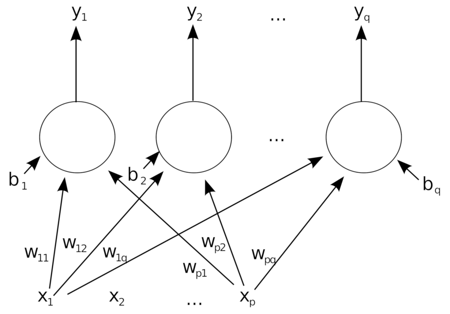


Signal detection and classification ⋆ Parameter estimation ⋆ Data cleaning (e.g., denoising) ⋆ Uncovering relations and patterns in data

Convolutional Neural network for classification



Convolutional Neural Network (CNN)



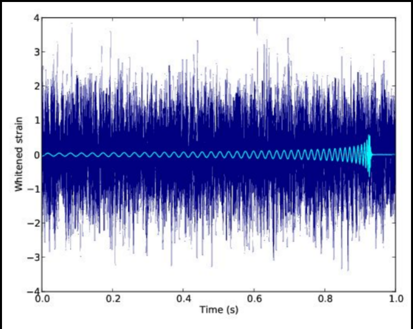
Next layer neuron y_j is a function of previous layer output x_i, with weights and biases (adjusted during training):

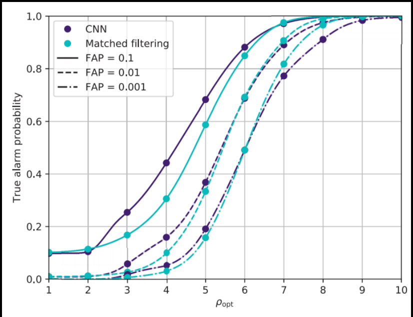
$$y_j = \sum_i x_i \cdot w_{ij} + b_j,$$

neuron activates after crossing threshold ($\rightarrow$ activation function).

- ⋆ Effectively kernel(s) convolution(s) with data on various scales,
- ⋆ Classification of 2D representation (images) or 1D (time series, distributions),
- ⋆ CNN used in many variants: not a complicated architecture, easy to train.

"Deep Filtering": NN detector/classifier



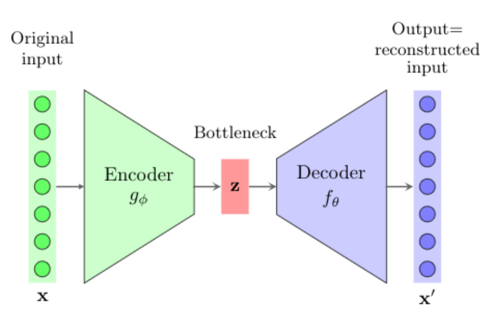


Typical BBH signal (whitened data, $m_1 = 12.06\ M_\odot$, $m_3 = 7.54\ M_\odot$, optimal SNR $\rho_{opt} = 8$.

- ⋆ CNN on time series to classify data with and without signal
- ⋆ Comparison with matched template method (template banks of $\sim$1000 templates).

Several implementations: George & Huerta, `arXiv:1701.00008`, Gabbard et al., `arXiv:1712.06041`

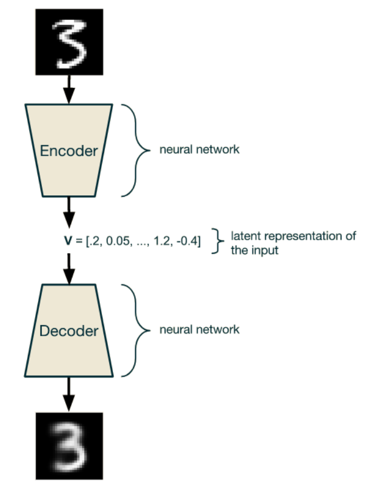
AutoEncoder architecture



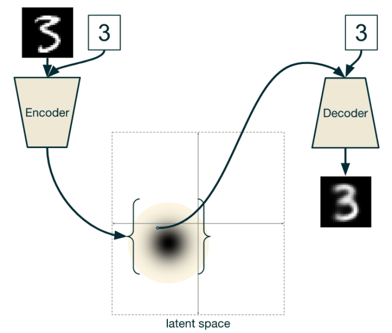
- ⋆ Identity function: compresses the representation of input, to later decompress it, in an unsupervised way (i.e., representation learning),
- ⋆ AEs are composed of two networks: an encoder g_ϕ, and a decoder f_θ,
- ⋆ Latent space representation **z** (the 'bottleneck'),
- ⋆ Training by minimizing a loss function, e.g.

$$L_{AE}(\theta, \phi) = \sum_{i=1}^{N} (x_i - f_\theta(g_\phi(x_i)))^2$$

(Conditional) Variational AutoEncoder

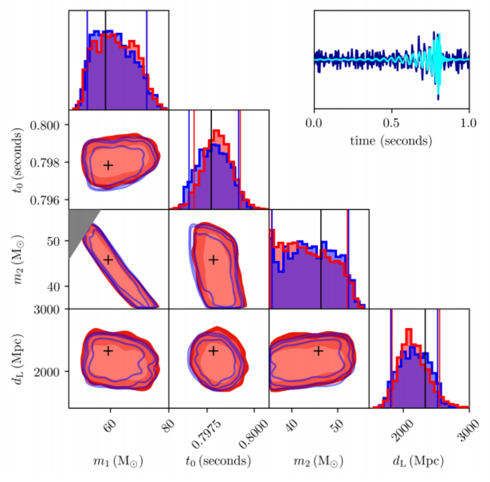


- ⋆ Latent space: convenient way of data reduction.



- ⋆ VAE produces a probability distribution in the latent variables space (for e.g. error in parameter estimation).
- ⋆ Conditional training: data + parameters (e.g. physical values generating the GW waveform)

Conditional VAE (CVAE) for GW parameter estimation



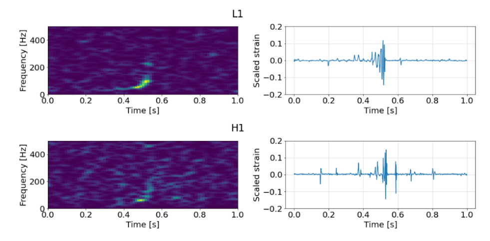
Simulates the Bayesian approach to inference (usually done with Markov Chain Monte Carlo parameter search) $\rightarrow$ very efficient in obtaining the posterior distributions of parameters

$\leftarrow$ comparison of the trained CVAE with one of the MCMC samplers (`Bilby`) used by the LIGO-Virgo Collaboration.

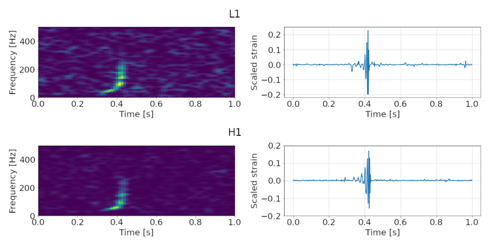
H. Gabbard et al., `arXiv:1909.06296`

Detecting GWs as data anomalies

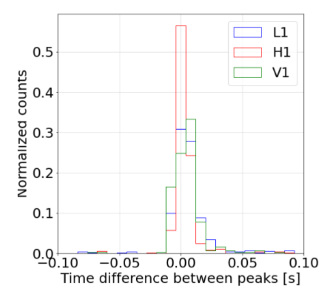
AE for noise reconstruction; anomaly = reconstructed output - input



Real data: GW170814



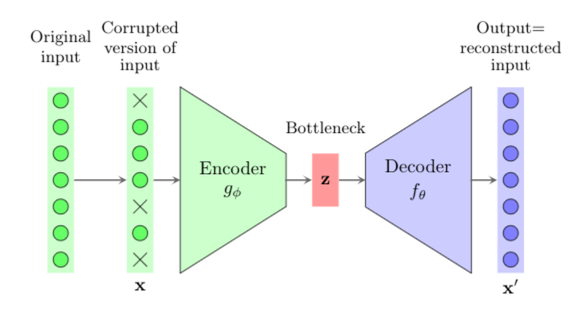
Real data: GW150914



Anomaly localisation in time based on the difference in peak positions between the reconstructed and injected signals. F. Morawski, MB, E. Couco, L. Petre (Mach. Learn.: Sci. Technol. 2 2021 045014)

Denoising gravitational waveforms

Denosing Convolutional AutoEncoder: noisy time series at the input, requesting clean time series at the output.

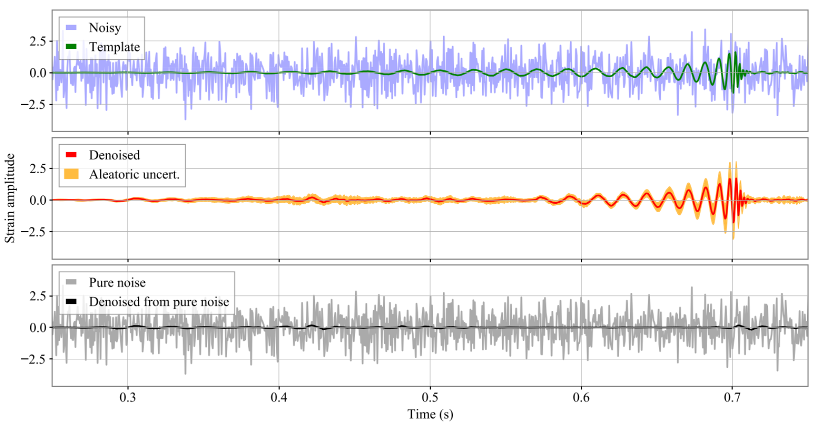


DAE loss function - the mean square error between the corrupted version of the ground truth $\mathbf{X}$ and the reconstructed output $\mathbf{x}' = f_\theta(g_\phi(\mathbf{x}))$:

$$L_{DAE}(\theta, \phi) = \sum_{i=1}^{N} (X_i - f_\theta(g_\phi(x_i)))^2$$

P. Bacon, A. Trovato, MB, Mach. Learn.: Sci. Technol. 4 035024 (2023)
arXiv:2205.13513

Training DAE on 1 month of O1 Livingston data



$M_1 = 22.74\ M_\odot$, $M_2 = 17.34\ M_\odot$, distance $r = 479.27$ Mpc, corresponding to the optimal matched filter signal-to-noise ratio $\rho_{\text{opt}} = 9.3$ The denoised $\rho_{\text{opt},d}$ is 10.3.

$$\rho_{\text{opt}} = \sqrt{(h|h)} = \sqrt{4\int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)}\mathrm{d}f},$$

(1 s segments, 2048 Hz sampling rate, $M_i \in (10, 30)\ M_\odot$, zero spins)

Properties of the DAE output

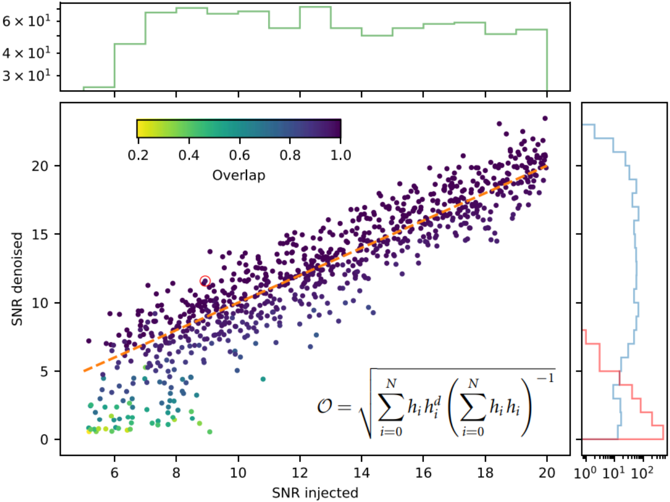


Figure 4. Denoised SNR (calculated from the DAE output analogously to equation (7), but with denoised output waveform h^d instead of the originally-injected signal h, vertical axis) as a function of the injected SNR (horizontal axis) for a testing dataset of 1000 data instances with added astrophysical GW waveforms. Points are colored by their corresponding overlap values. Orange dashed line denotes the denoised SNR equal to the injected SNR. Example waveform presented in figure 2 is denoted by a red circle. Side histograms (in logarithmic scale) show the distribution of the injected SNR (upper plot), and SNRs denoised from samples containing added GW waveforms (blue histogram), and—for comparison—not containing GW signals (i.e. pure noise, red histogram), respectively.

DAE output on glitches

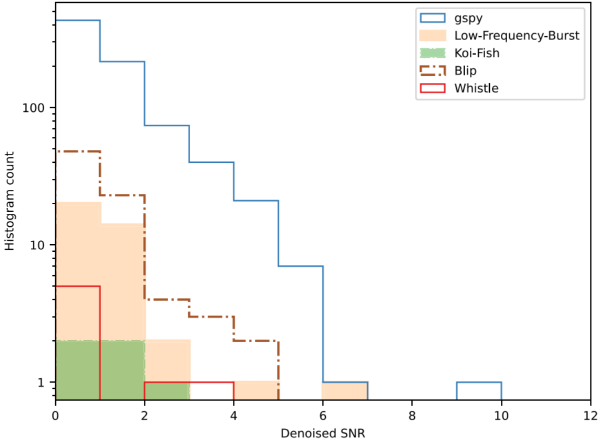


Figure 5. Evaluation of the DAE on instrumental glitches. The logarithmic vertical scale plot shows histograms of denoised output SNR for a selection of 38 Low Frequency Burst glitches, 5 Koi Fish type glitches, 80 Blips and 7 Whistle glitches. The blue line marks the evaluation of 792 assorted various types of glitches. All the glitches data are obtained from the `Gravity Spy` database [55]. The glitches have their estimated intrinsic SNR > 10.

DAE output on real data from O2: GW170104@H1

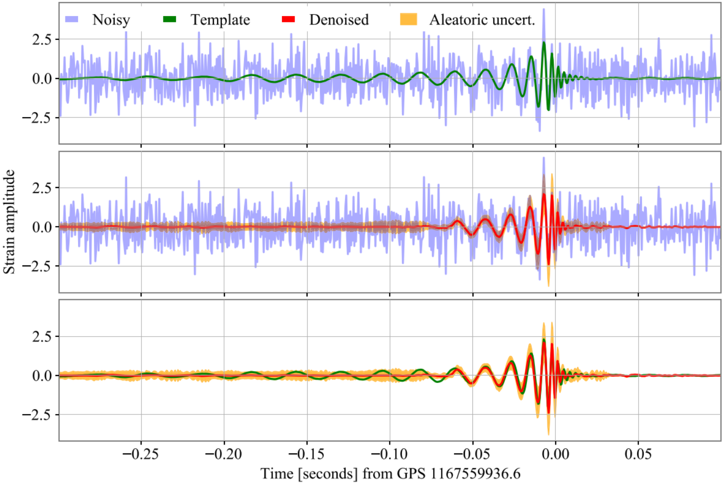


Figure B1. Denoising applied to the O2 data GW170104 event for the L1 detector (3 top panels) and the H1 detector (3 bottom panels). The component masses are $30.8^{+7.3}_{-5.6}\ M_\odot$ and $20.0^{+4.9}_{-4.6}\ M_\odot$, and the single-detector optimal SNRs are $9.9^{+1.5}_{-1.3}$ for L1 and $9.5^{+1.3}_{-1.6}$ for H1.

DAE output on real data from O2: GW170608@H1, $f_{low} = 50$ Hz

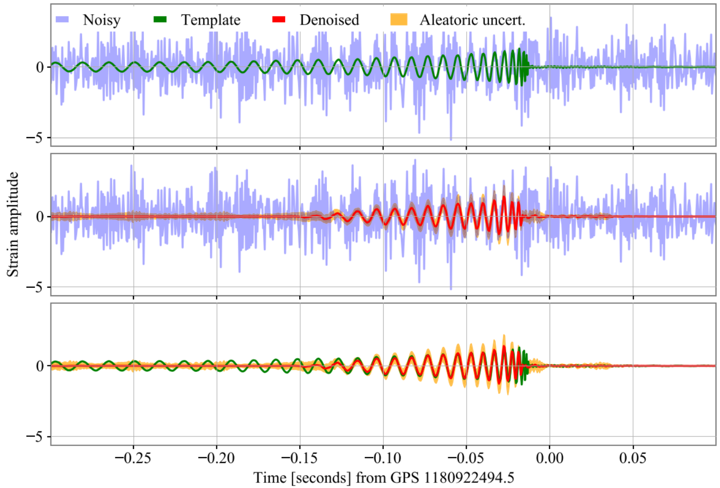
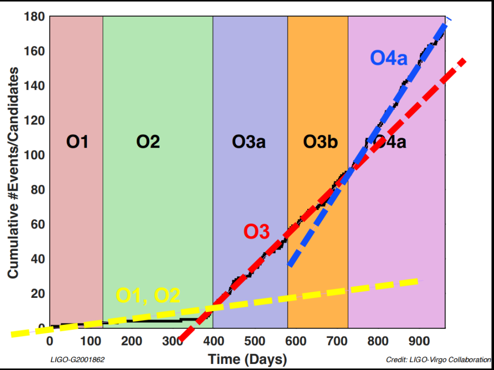
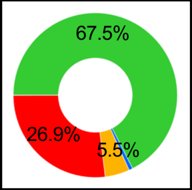


Figure B4. Denoising applied to the O2 data GW170608 event for the L1 detector (3 top panels) and the H1 detector (3 bottom panels). While for the other plots the high-pass filter was set to 30 Hz (as in the training set), in this case we apply a high pass at 50 Hz to the original data before the denoising.

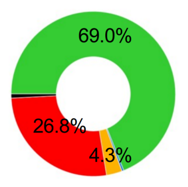
O4a (May 24, 2023 - January 16, 2024)



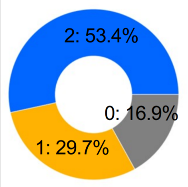
H1 Performance



L1 Performance

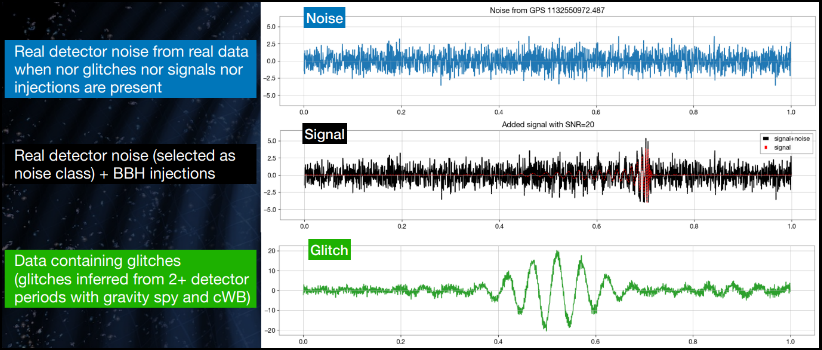


Network Performance



Single detector GW signal classifier: data

Data: 1 month of L1 data without know GW detections (between Nov 25, 2015 and Dec 25, 2015) + known glitches from the `Gravity Spy` database; 1s duration time-domain input, 2048 Hz sampling date



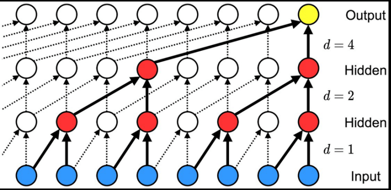
A. Trovato, É. Chassande-Mottin, MB, R. Flamary, N. Courty, CQG in review, `arXiv:2307.09268`

GW classifier: ML architectures

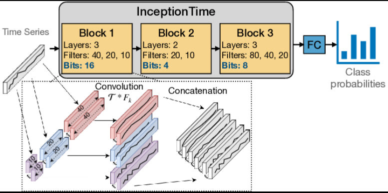
Standard scheme: input data → NN → classification probability P_s

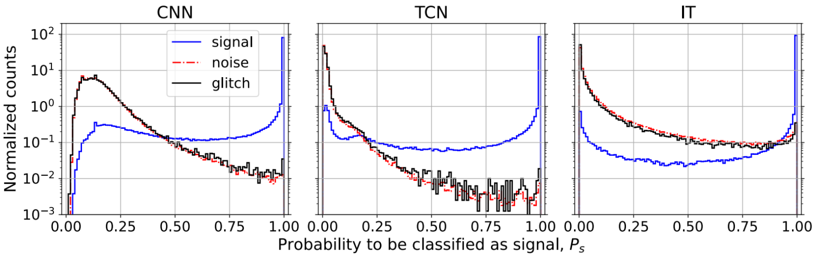
In addition to ”vanilla” 1D CNN:

Temporal CNN (TCN)



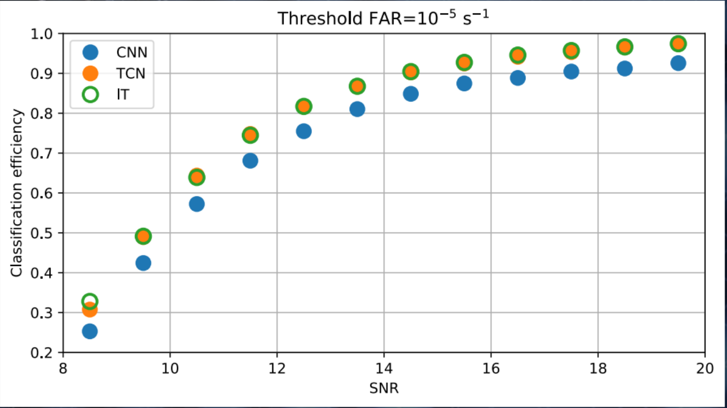
Inception Time (IT)





(effectively, 2-class problem: signal vs glitch+noise)

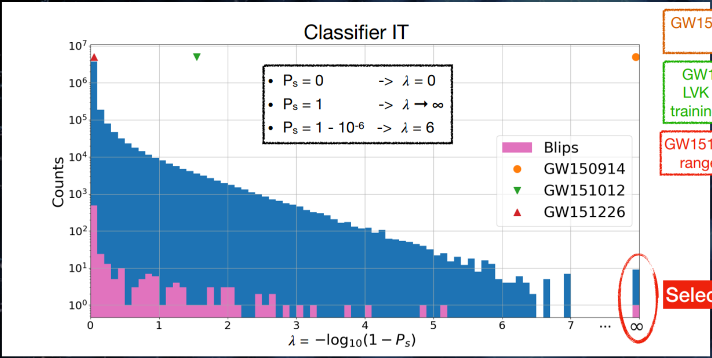
Classification efficiency vs SNR for fixed False Alarm Rate



- TCN and IT perform similarly and outperform CNN
- Efficiency better than 0.5 for SNR>9 at this level of FAR
  - (1 alarm per 10^5 s = 0.864 alarms per day)

Application to remaining 3 months of O1

- ⋆ Analysis of remaining 3 months of O1 L1 data, excluding the 1 month period already used for training and testing,
- ⋆ known GW detections (3 in O1) have been examined separately.



GW150914 identified with $P_s = 1$ by all networks

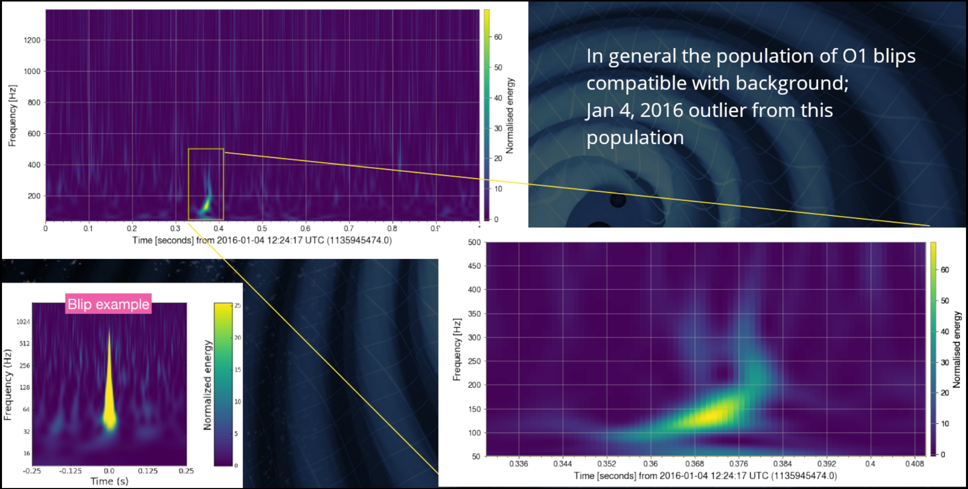
GW151012 was detected by LVK in L1 with a SNR~6 (our training set has a minimum of 8)

GW151226 has masses not in the range used in our training set

Selected triggers

(in pink: blip glitches classified by `Gravity Spy`)

$P_s = 1$ glitch at GPS=1135945474.373 (Jan 04, 2016)



Previous single-detector O1 analyses

Turns out, this O1 outlier was analyzed before:

THE ASTROPHYSICAL JOURNAL, 897:169 (9pp), 2020 July 10 https://doi.org/10.3847/1538-4357/ab96c7

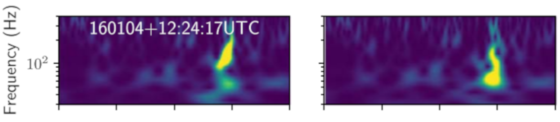
CrossMark

A Search for Gravitational Waves from Binary Mergers with a Single Observatory

Alexander H. Nitz[1,2], Thomas Dent[3], Gareth S. Davies[3], and Ian Harry[4,5]
[1] Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), D-30167 Hannover, Germany; alex.nitz@aei.mpg.de
[2] Leibniz Universität Hannover, D-30167 Hannover, Germany
[3] Instituto Galego de Física de Altas Enerxías, Universidade de Santiago de Compostela, Santiago de Compostela, Galicia, Spain
[4] University of Portsmouth, Portsmouth, PO1 3FX, UK
[5] Kavli Institute of Theoretical Physics, UC Santa Barbara, CA, USA
Received 2020 April 22; revised 2020 May 20; accepted 2020 May 25; published 2020 July 15

				LIGO-Livingston						
Date Designation	GPS Time	Obs.	Known	$\tilde{\Lambda}_s$	$p_{\mathrm{astro},s}$	ln $\mathcal{B}_{c/s}$	$p_{\mathrm{astro},c}$	$\mathcal{O}_c$	$\mathcal{M}^{\mathrm{src}}$	χ_{eff}
160104+12:24:17UTC	1135945474.38	L	⋯	12.21	0.47	⋯	0.47	0.90	$32.7^{+6.6}_{-8.2}$	$0.2^{+0.3}_{-0.5}$

For the top candidates, we perform an additional diagnostic to confirm that the signal morphology is consistent with a gravitational-wave source. We take the best-fit parameters for a GR gravitational-wave signal and subtract them from the data. The result for selected candidates is shown in Figure 4. For some cases, this diagnostic disfavors a candidate due to missing or excess power, such as the case of 160104+12:24:17UTC, which otherwise would have been the most significant candidate. We see that subtracting off the best-fit estimate of

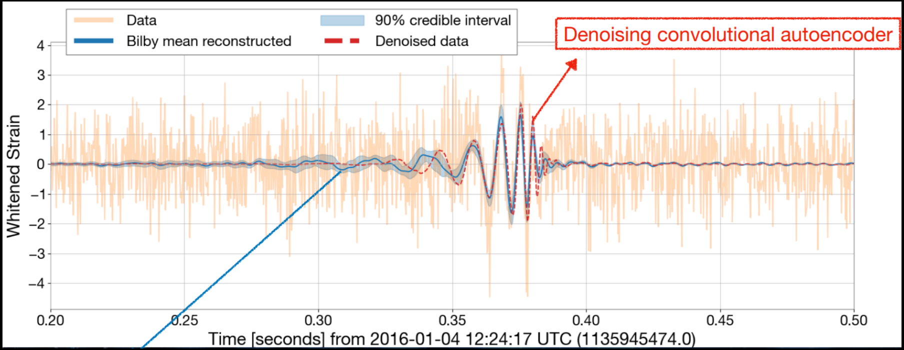


event spectrogram residual

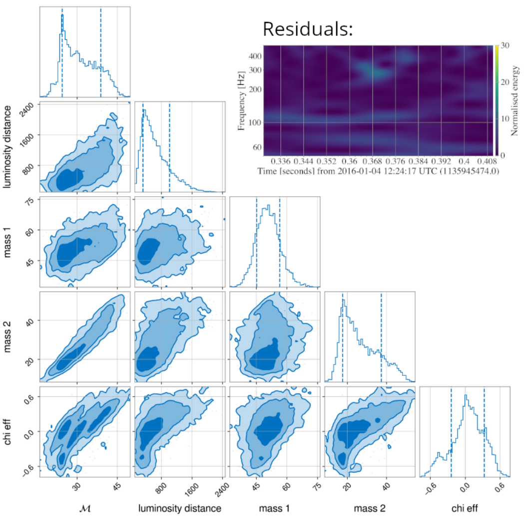
Astrophysical origin of GPS=1135945474.373 (Jan 04, 2016) event?

We have checked, if it's compatible with GW waveform ("chirp"), using:

- ⋆ standard-choice parameter estimation & MCMC sampler library `bilby`,
- ⋆ 1D CNN DAE (Bacon et al., 2023).



GPS=1135945474.373 (Jan 04, 2016) astrophysical origin? `bilby` parameter estimation



$$GPS = 1135945474.373^{+0.076}_{-0.07}$$

$$SNR = 11.34^{+1.8}_{-1.6}$$

$$\mathcal{M} = 30.18^{+12.3}_{-7.3} M_\odot$$

$$m_1 = 50.7^{+10.4}_{-8.9}\ M_\odot$$

$$m_2 = 24.4^{+20.2}_{-9.3}\ M_\odot$$

$$\chi_{\rm eff} = 0.06^{+0.4}_{-0.5}$$

$$d_L = 564^{+812}_{-338}\ \mathrm{Mpc}$$

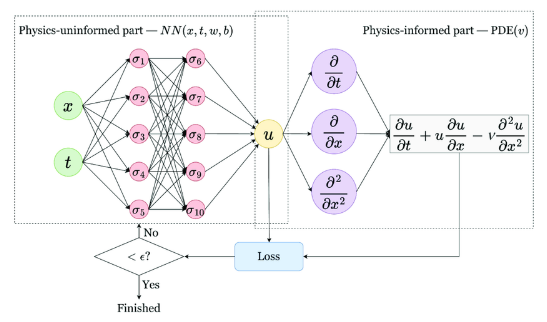
Consistent with BBH population observed so far 🥳 🤖

Physically Inspired Neural Networks (PINNs)

Model discovery with PINNs: incorporating physical principles into ML (solution to obey certain equations) to *infer* physical parameters from signal events discovered in *unmodeled* GW searches, i.e.

(collaboration with Matteo Scialpi)

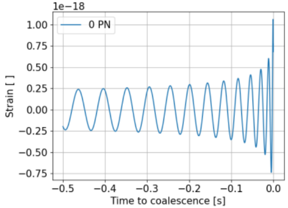
$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{96}{5}\pi^{8/3}\left(\frac{G\mathcal{M}}{c^3}\right)^{5/3} f^{11/3}$$

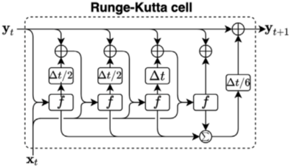


PINNs for waveform parameter estimation: Newtonian case

Master thesis version:

- **Data**: a clean waveform, made by $\{(t_k, h_k)\}_{k=1}^{N}$ data points.
- We can impose (1), where $\mathcal{F}$ is a known functional and $\vec{\theta}$ are the parameters to be inferred.
- **We solve (1) thanks to a Recurrent Neural Network** (RNN) with a Runge-Kutta integrator at 4th order implemented inside.





$$\frac{df}{dt} = \mathcal{F}[f, \vec{\theta}]$$

$$h(t) = \mathcal{H}[f, \vec{\theta}]$$

$$\mathcal{L} = \mathcal{L}\left[\sum_{k=1}^{N} (h_k - h(t_k, \vec{\theta}))^{\beta}\right]$$

PINNs: 1.5 post-Newtonian case

$$M_{tot} = m_1 + m_2, \quad \eta = \frac{m_1 m_2}{M_{tot}^2},$$

$$\varepsilon = \frac{GM_\odot}{c^3}\left(\frac{M_{tot}}{M_\odot}\right)\pi f,$$

$$\frac{df}{dt} = \frac{96}{5}\pi^{8/3}\left(\frac{GM_\odot}{c^3}\right)^{5/3}\left(\frac{M_{tot}}{M_\odot}\right)^{5/3}\eta f^{11/3}\times$$

$$\times\left[1-\left(\frac{743}{336}+\frac{11}{4}\eta\right)\varepsilon^{2/3}+4\pi\varepsilon\right]$$

$$t(f) = t_c - 5(8\pi f)^{-8/3}\left(\frac{GM_\odot}{c^3}\frac{M_{tot}}{M_\odot}\right)^{-5/3}\eta^{-1}\times$$

$$\times\left[1+\frac{4}{3}\left(\frac{743}{336}+\frac{11}{4}\eta\right)\varepsilon^{2/3}+\frac{32}{5}\pi\varepsilon\right]$$

$$\Psi(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \frac{3}{4}\eta\left(8\pi\frac{GM_\odot}{c^3}\frac{M_{tot}}{M_\odot}f\right)^{-5/3}\times$$

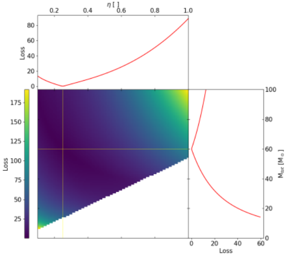
$$\times\left[1+\frac{20}{9}\left(\frac{743}{336}+\frac{11}{4}\eta\right)\varepsilon^{2/3}-16\pi\varepsilon\right]$$

$$\tilde{h}(f) = \frac{\chi c}{D}\sqrt{\frac{5}{24}}\pi^{-2/3}\left(\frac{GM_\odot}{c^3}\right)^{1/3}\left(\frac{M_{tot}}{M_\odot}\right)^{5/6}\eta^{1/2}f^{-7/6}e^{i\Psi(f)}$$

$$\mathcal{L} = \frac{\beta_f}{N}\sum_{k=1}^{N}|f_k - f(t_k)| +$$

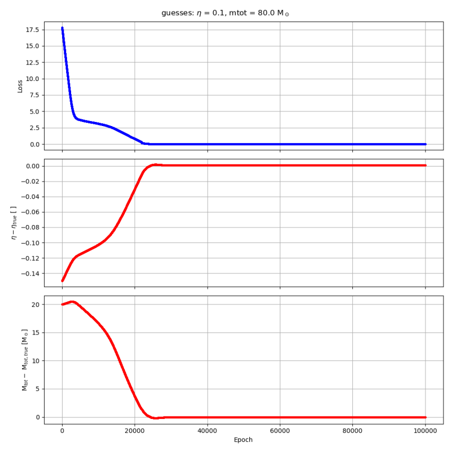
$$+\frac{\beta_t}{N}\sum_{k=1}^{N}|t_k - t(f(t_k))| +$$

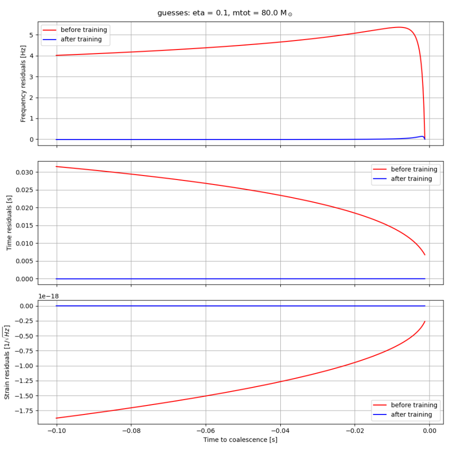
$$+\frac{\beta_h}{N}\sum_{k=1}^{N}|h_k - h(f(t_k))|$$



PINNs: 1.5 post-Newtonian case

- ⋆ "Ground truth" values $m_1 = m_2 = 30\,M_\odot$ ($\eta = m_1 m_2 / M_{tot}^2 = 0.25$),
- ⋆ initial values $M_{tot} = 80\,M_\odot$, $\eta = 0.1$.





Summary & outlook

- $\star$ There are ML implementations that may even be useful practically,
- $\star$ Now working on bringing physics laws into the game
- $\rightarrow$ need to introduce complications: non-stationary noise etc.,
- $\star$ Other interesting areas of GW-related astrophysics: NS interiors, equation of state, ...