Application of deep neural networks in physics: black hole quasinormal modes and improving new physics searches

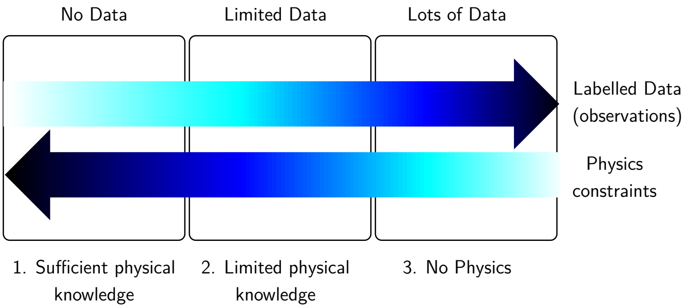
Anele Ncube

Supervised by Prof A. Deandrea, Prof A. S. Cornell & Dr S. R. Herbst

15 April 2025

Outline

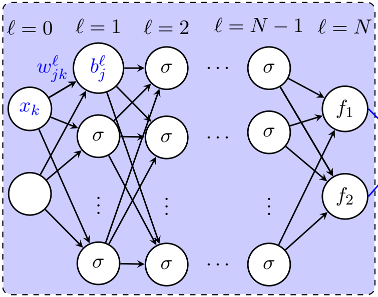
Data & Physics – 3 scenarios



- **Physics-informed neural networks** – PDE solvers
  1. Sturm-Liouville (eigenvalue) problems, BIVPs – data independent.
  2. Inverse problems – partially data driven.
- **Neural networks with data-driven learning**
  1. E.g.: binary classification, regression tasks. – data dependent.

adopted from Karniadakis et al (2020) *Nat Rev. Phys* **3** 422

FNN architecture – generic



$\sigma = \sigma(w^\ell_{jk} x_k + b^\ell_j)$

$\theta = \left\{ w^\ell_{jk}, b^\ell_j \right\}_{1 \le \ell \le L}$

$\mathcal{L}(\theta)$

w^ℓ_{jk}, b^ℓ_j are tunable weights & biases where j and k index neurons of $(\ell - 1)$-th and ℓ-th layers, resp.

- FNN ansatz:

$$f(x) = \sigma^{\ell=L} \left[\sum w^{\ell=L} \sigma^{\ell=L} \left(\dots \sum w^{\ell=2} \sigma^{\ell=1} \left(\sum w^1 x + b^1 \right) + b^2 \dots \right) + b^L \right].$$

- Activation (σ): e.g. sin, tanh,

$$\sigma(z) = \frac{1}{1 + \exp(-z)} \quad \text{(sigmoid)}.$$

- Loss ($\mathcal{L}$): problem-dependent e.g. $\mathcal{L}_{PDE} + \mathcal{L}_{IC/BC}$ for solving EIBVPs.

NN optimisation algorithm – pseudo-code

- Initialise $\theta = \left\{ w^{\ell}_{jk}, b^{\ell}_{j} \right\}_{1 \leq \ell \leq L}$ randomly from e.g. $\mathcal{N}(\mu, \sigma^2)$.
- **For each** epoch **in** training epochs:
  1. Input x_j (training points) + Evaluate $f(x_i)$ and $\mathcal{L}$.
  2. **Backpropagate:** compute $\dfrac{\partial \mathcal{L}}{\partial \theta} = \left\{ \dfrac{\partial \mathcal{L}}{\partial w^{\ell}_{jk}}, \dfrac{\partial \mathcal{L}}{\partial b^{\ell}_{j}} \right\}_{1 \leq \ell \leq L}$ using autodiff.
  3. **Update** θ**:** $w \to w' = w - \dfrac{\partial \mathcal{L}}{\partial w}, \quad b \to b' = b - \dfrac{\partial \mathcal{L}}{\partial b}$ (optimisation algo. dependent).
  4. **Validate:** e.g. for early-stopping.
- **Test** performance of model.

autodiff: exact numerical differentiation.

Overview of application to physics problems

Use-case	**NN algorithm**
1. Compute QNM frequencies of Kerr BHs (solving ODEs).	PINNs
2. Improving sensitivity of new physics searches (signal-background event classification).	Binary classifiers

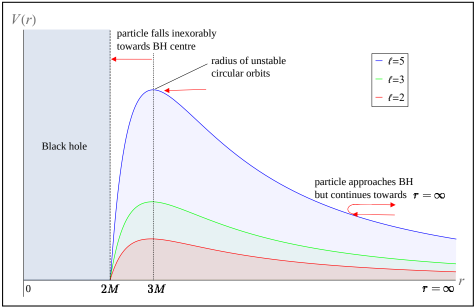
- **Implementation:** Pytorch – an ML open-source library in Python.

BH QNMs – simpler problem: Schwarzschild BH

- Regge-Wheeler-Zerilli eqn.:

$$\frac{d^2\psi(x)}{dx^2} + [\omega^2 - V(r)]\psi(x) = 0,^*$$

tortoise: $-\infty < x < +\infty$,
radial: $2M < r < +\infty$.



- Schwarzschild BH **effective potential**:

$$V(r) = \left(1 - \frac{2M}{r}\right)\left[\frac{\ell(\ell+1)}{r^2} + (1-s^2)\frac{2M}{r^3}\right],$$

M = BH mass, ℓ = multipole number, s = spin of perturbing field.

$^*\omega(\text{SI units}) = \omega(\text{G units}) \times 2\pi(5.142\text{kHz})M_\odot/M$, where G: $G = c = 1$.

Astrophysical constraints & QNM frequencies

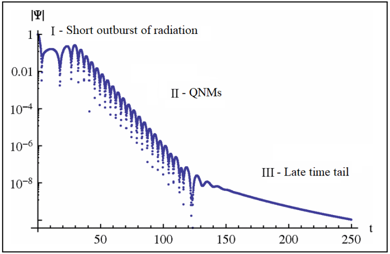
- Asymptotic behaviour of QNMs:

$$\psi(x) = \begin{cases} e^{-i\omega x}, & x \to -\infty \\ e^{+i\omega x}, & x \to +\infty \end{cases},$$

- non-Hermitian eigenvalue problem, i.e for QNMs with $e^{-i\omega t}$:

$$\omega = \omega_{Re} - i\omega_{Im},$$

$$\omega_{Re} = \text{physical oscillation frequency},$$
$$\omega_{Im} \propto \text{damping rate}.$$



simple simulation of GW signal consisting of QNMs

- GW significance of QNMs: “footprints” of BH as frequencies are BH parameter dependent.

Konoplya & Zhidenko (2011) *Rev. Mod. Phys.* **83** 793.

Teukolsky eqn – Kerr BH QNMs

- Radial & angular ODEs:

$$\Delta(r)R''(r) + (s+1)(2r-1)R'(r) + V(r)R(r) = 0,$$

$$(1-u^2)S''(u) - 2uS' + \left[a^2\omega^2u^2 - 2a\omega su + s + A - \frac{(m+su)^2}{1-u^2}\right]S(u) = 0,$$

where $\Delta(r) = r^2 - 2Mr + a^2$, $u = \cos\theta$, a = BH rotation, ℓ, m = multipole indices.

- Eigenfunctions & eigenvalues:

$$R(r) = {}_sR_{\ell m\omega}(r), \quad S(u) = {}_sS_{\ell m\omega},$$
$$\omega = \omega_{Re} - i\omega_{Im},\ A = A_{Re} + iA_{Im}.$$

- Asymptotic behaviour:

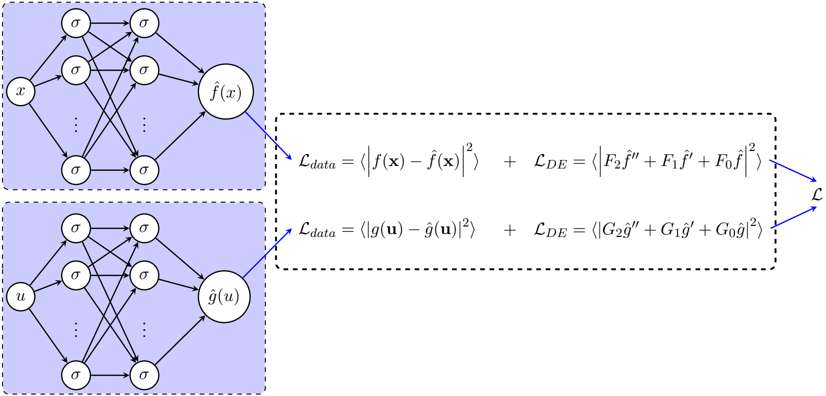
$$R(x) = \begin{cases} e^{-i(\omega - ma/2Mr_+)x}/\Delta^s, & x \to -\infty \\ e^{+i\omega x}/r^{2s+1}, & x \to +\infty \end{cases}.$$

Existing methods

Existing methods for computing QNMs of BHs.

Method	**Pros**	**Cons**
Continued fraction method (Leaver, 1985)	High accuracy (standard technique)	No physical intuition
WKB (Schutz & Will, 1985)	Physical intuition	poor for high n i.e. $n > \ell$
Asymptotic Iteration Method (Cho et al, 2012)	Accurate as WKB	Requires n-specific seed values unlike WKB

PINNs approach – Kerr case



- Training minimises $\mathcal{L}(\theta)$:

$$\mathcal{L}(\theta) = \mathcal{L}_{DE+BC} + \mathcal{L}_{data},$$

where

$$F_{0,1,2} = F_{0,1,2}(x, \omega), \quad x \in [0, 1]; \quad G_{0,1,2} = G_{0,1,2}(u, A), \quad u \in [-1, 1].$$

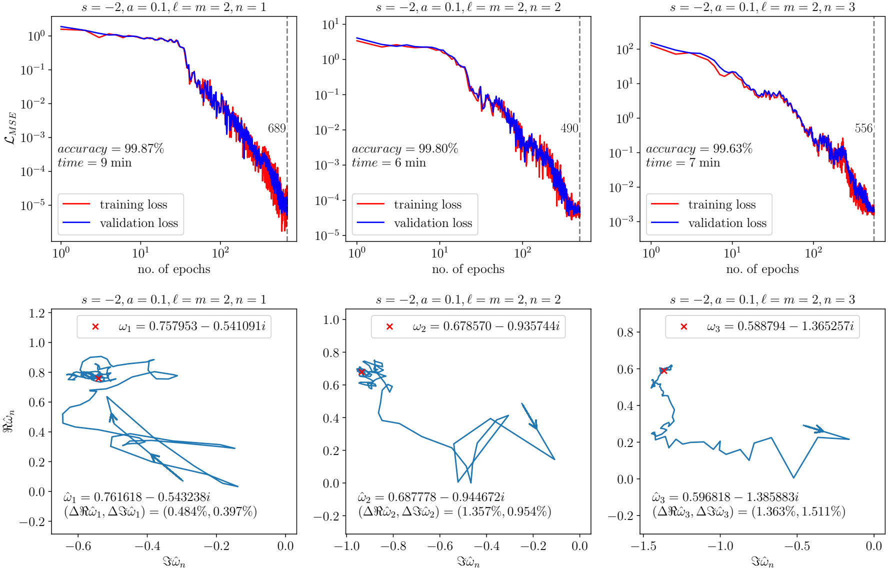
Numerical results

Low-lying QNM frequencies – gravitation perturbations of a Kerr BH
($s = -2, M = 0.5, a = 0.1, \ell = m = 2, n \in \{1, 2, 3\}$).

n	CFM ω_n $\{A_{22}\}$	PINN $\hat{\omega}_n$ $(\Delta\Re\hat{\omega}_n, \Delta\Im\hat{\omega}_n)$	$\hat{A}_{22}$ $(\Delta\Re\hat{A}_{22}, \Delta\Im\hat{A}_{22})$
1.0	0.7580 - 0.5411i {3.7958 + 0.1504i}	0.7616 - 0.5432i (0.484%, 0.397%)	3.7924 + 0.1548i (-0.091%, 0.386%)
2.0	0.6786 - 0.9357i {3.8222 + 0.2589i}	0.6878 - 0.9447i (1.357%, 0.954%)	3.8201 + 0.2614i (-0.054%, 0.196%)
3.0	0.5888 - 1.3653i {3.8543 + 0.3758i}	0.5968 - 1.3859i (1.363%, 1.511%)	3.8413 + 0.3648i (-0.339%, -0.801%)

- **Baseline set-up**: 2 & 3 hidden layers, 20 per layer, 5×10^3 epochs (early-stopping) Adam optimiser, $\sim 10^3$ training pts.

Evolution of PINN approximations



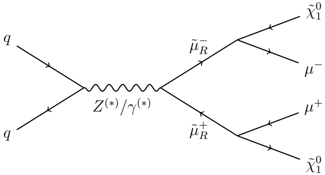
Loss (top row); $\hat{\omega}_{PINN}$ (bottom row).

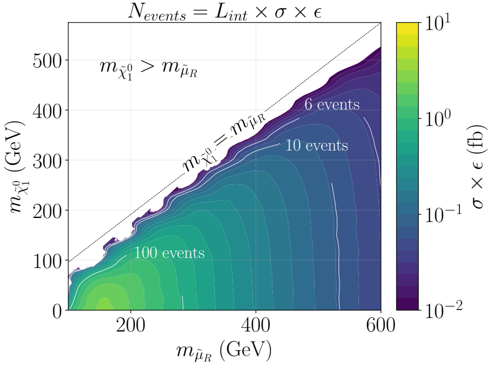
Improving $\tilde{\mu}_R^{\pm}$ searches in LHC data

- Simulations on **MG5** with Run 2 setup (i.e. $\sqrt{s} = 13$ TeV; $L_{int} = 139$ fb^{-1}).
- $\tilde{\chi}_1^0$ stable LSP $\therefore$ DM candidate.
- Backgrounds (SM QCD):

$$\begin{aligned} p\bar{p} &\to t\bar{t}, \\ p\bar{p} &\to \ell^+\ell^-\nu\bar{\nu}, \\ p\bar{p} &\to \ell\ell\ell\nu. \end{aligned}$$

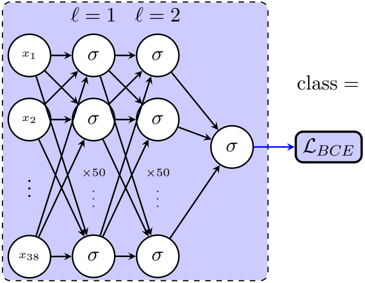
- $s/b \ll 1$
  e.g. $b \sim 750$, $s \sim 10$ events for $(m_{\tilde{\mu}_R}, m_{\tilde{\chi}_1^0}) = (500, 0)$ GeV.





Concerning signals

NN binary classifier

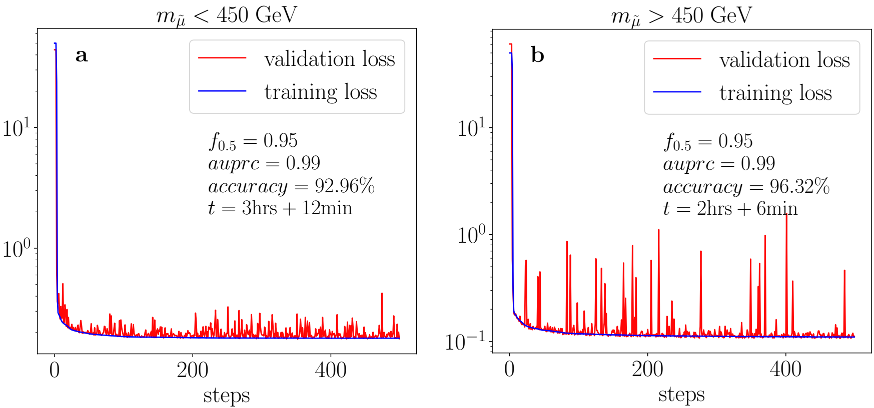


$$\text{class} = \begin{cases} 0 \text{ (bkgs)} & \text{for } \sigma^{\ell=L} \leq 0.7, \\ 1 \text{ (sigs)} & \text{for } \sigma^{\ell=L} > 0.7. \end{cases}$$

x_i = kinematic variables; σ = sigmoid; $\mathcal{L}_{BCE}$ = binary cross-entropy

- SGD optimisation algorithm,
- standardised inputs; train:validation:test ratio: 8:1:1; balanced train set.
- Gain in sensitivity compared to existing approaches (i.e. jet veto & BDTs)???

NN Loss



2 NNs trained on data from 2 signal regions: low and high $m_{\tilde{\mu}_R}$.

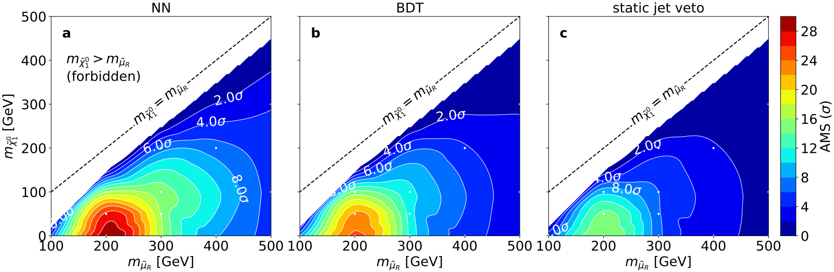
Sig acceptance & bkg rejection – NN vs. BDT vs. jet veto cuts

Signal acceptance: BDT > NN > jet veto.

$m_{\tilde{\mu}_R}$ (GeV)	$m_{\tilde{\chi}_1^0}$ (GeV)	s (TPs) NN		BDT		static jet veto	
low $m_{\tilde{\mu}_R}$ signal region							
100	0	19.5	(21.7%)	53.3	(59.5%)	37.2	(41.6%)
200	0	172.1	(72.3%)	214.5	(90.1%)	129.6	(54.4%)
300	0	74.3	(87.6%)	82.1	(96.8%)	45.2	(53.3%)
400	0	29.3	(91.4%)	31.6	(98.5%)	16.6	(51.7%)
high $m_{\tilde{\mu}_R}$ signal region							
500	0	12.0	(90.5%)	13.2	(99.3%)	13.2	(99.3%)

- Background rejection (b/ FPs): low $m_{\tilde{\mu}_R}$ NN 5.9 (0.8%), high $m_{\tilde{\mu}_R}$ NN 2.4 (0.3%), BDT 28.4 (4.0%) & 33.8 (4.5%).

Statistical analysis (CL_s method) – NN vs. BDT vs. jet veto cuts



Projected exclusion limit plots (NN > BDT > jet veto).

- AMS:

$$Z = \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right)-s\right)}, \quad b = \text{FP},\ s = \text{TP}.$$

- $Z = 2\sigma$ $(= 95\%\text{CL}) \Rightarrow$ exclusion.

Summary

PINNs for Kerr QNMs; NN classifiers for $\tilde{\mu}_R$ searches.

Topic	**Kerr BH QNMs**	$\tilde{\mu}_R$ **searches**
ML approach	supervised PINNs	NN classifier in 2 signal regions
Results	frequencies <5% error for low-lying overtones ($n \leq 3$) & slow rotation ($a \leq 0.1$)	Greater sensitivity for all sig events at Run 2 (more decisive exclusion limits).
Limitations & improvements	Importance of higher overtones & improved PINN design	Limited success in compressed region (i.e. $m_{\tilde{\chi}_1^0}/m_{\tilde{\mu}_R} \leq 1.5$), ongoing research.

- **Articles:**
- "Solving the Regge-Wheeler and Teukolsky equations: supervised versus unsupervised physics-informed neural networks" (arXiv:2402.11343 [gr-qc]).
- "Improving smuon searches with Neural Networks" (arXiv:2411.04526 [hep-ph]).

PINN application to eigenvalues in nilmanifolds

Laplacian eigenmodes in twisted periodic topologies for new physics models

Ncube A., Deandrea A., Cornell A., Herbst R.

Abstract: Laplacian eigenmodes in non-trivial topologies (e.g. having twisted periodicity) are important in constructing a complete picture of the physics at play within models that incorporate compact extradimensional spaces. Determining them analytically is generally unwieldy, and the existing standard numerical methods have limited ability as spatial dimensions increase and when computing higher-index eigenmodes is required. To determine the feasibility of using physics-informed neural networks to compute Laplacian eigenmodes, we apply them to three primitive test cases: the Möbius strip, the real projective plane and the 3-torus in Cartesian coordinates. The neural networks approach's potential performance beyond solving the simpler cases is estimated in terms of the approximation errors obtained by comparing with known analytical solutions.

---

for SAIP2025 conference on 7 - 11 July, 2025.

Thank you