

Gravity as a Deep Neural Network

Javier Matulich

Instituto de Física Teórica UAM/CSIC, Madrid

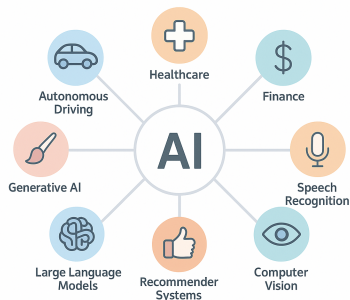


GRASS-SYMBHOL Meeting 2025, Toledo.

In collaboration with Gabriele Barbagallo (IFT).

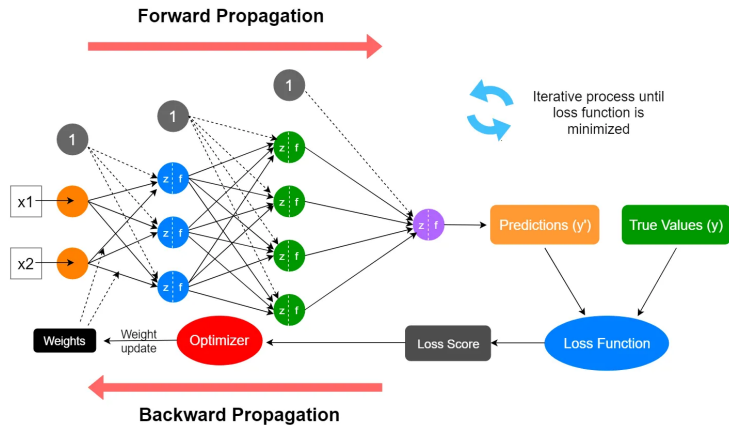
Introduction

- Artificial intelligence (AI) is the simulation of human intelligence in machines.
- At the forefront of recent advances in AI is Machine Learning (ML), with the goal of developing algorithms capable of learning from and adapt to new data, without human assistance.
- The past decade has witnessed the rise of ML impacting countless areas.



Introduction

- The most astonishing advances in ML are based on Neural Networks (NN), powered by Deep Learning (DL) algorithms.



Introduction

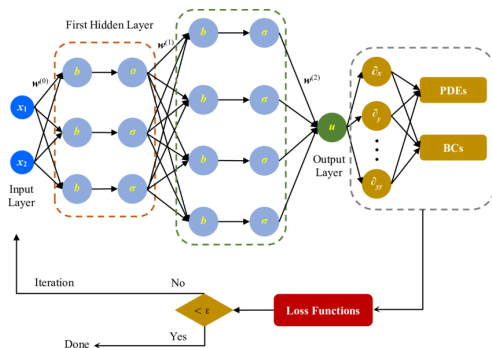
- The scientific community has increasingly become interested in its potential for fundamental research in diverse areas of knowledge.
- In particular, ML have become powerful tools for addressing several problems in physics, such as particle physics, gravitational waves, dark matter, as well as theoretical high-energy physics and mathematical physics.
- ML-based models make connections that have shown, with real-life examples, to be highly counterintuitive, as well as unveiling patterns that can go unnoticed by human intuition. The question is: Could it notice something we have overlooked?
- An uncharted territory to be explored by ML techniques lies at the very foundations of fundamental interactions. Among them, the least understood is the gravitational interaction. Despite its impressive success, General Relativity fails at the microscopic level.

Introduction

- The main goal is to delve into fundamental questions concerning our Universe in the context of gravitational theories, by training a machine to learn how massive objects interact.
- The approach relies in **Physics-informed Deep Learning architectures** which will be used to encode physically meaningful features that could potentially help to delve into some relevant questions regarding gravity.
- If achieved, this could contribute with a new way of approaching fundamental problems in physics. This is what makes the research line ambitious and potentially ground-breaking.

Physics-Informed Neural Networks (PINNs)

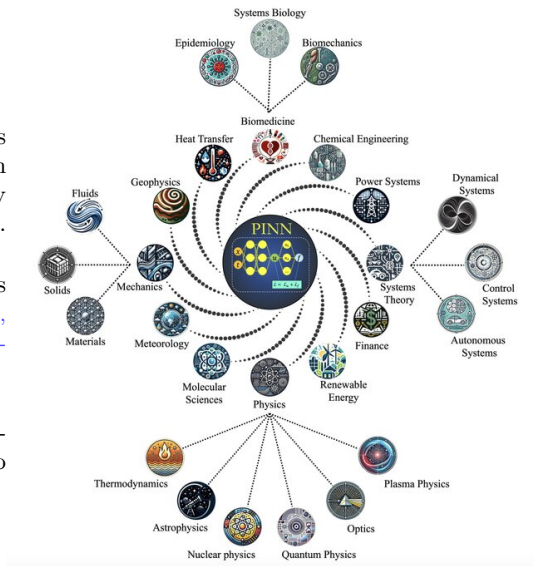
- In a nutshell, **PINNs** are DNN trained to solve (un)supervised learning tasks while respecting any symmetries, invariances, or conservation principles that can be modeled by general time-dependent and nonlinear partial differential equations.



- The differential equations are not solved in the usual way by time evolution; instead, the neural network provides a solution across the entire domain at once.

Applications of PINNs

- PINNs algorithm has been extensively used in the last 6 years in many different physics problems.
- The original article has 18329 citations [M. Raissi, P. Perdikaris, G.E. Karniadakis (2019)].
- There are very few examples of its applications to gravity.



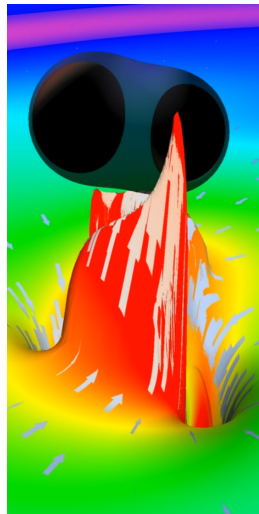
Credit: José Garcia Ventura

Applications of PINNs

- It is then part of this work to deepen into the scope of physics-informed Neural, looking for a more efficient and scalable way of simulating dynamical systems in General Relativity.
- The expected advantage of the method is that PINNs trained with a simpler data set, will allow to encode a more complex dynamical evolution. For instance, in the case of binary black hole coalescence, the PINNs can be trained (or "informed") by the data from numerical methods, effective one-body (EOB) or from Post-Newtonian (PN) approximations, so that the entire evolution, including merger, can be ultimately obtained from the PINNs.
- First we need to control some of the state-of-the-art approaches.

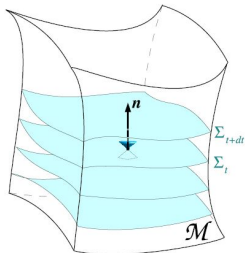
Numerical Relativity in a Nutshell

- **3+1 Decomposition:** Split spacetime into spatial hypersurfaces evolving in time (ADM / BSSN).
- **Gauge Choices:** Choose lapse function and shift vector to control coordinate evolution.
- **Initial Data:** Solve constraint equations to obtain consistent starting conditions.
- **Numerical Methods:**
 - Grid-based:** Discretize spacetime on a mesh; robust for binaries.
 - Spectral:** Represent fields as global basis expansions; high precision for smooth fields.
- **Stability and Boundaries:** Proper outer boundary conditions.
- **Diagnostics:** Extract gravitational waves, apparent horizons, and conserved quantities.



Gravity in its Hamiltonian form for numerical relativity

- Arnowitt–Deser–Misner (ADM) decomposition: For globally hyperbolic spacetimes $(\mathcal{M}, g)$: $\mathcal{M} \cong \mathbb{R} \times \Sigma$. They can be foliated $\mathcal{M} = \bigcup_{t \in \mathbb{R}} \Sigma_t$



$$g_{\mu\nu} = \begin{pmatrix} -\alpha^2 + \beta_k \beta^k & \beta_i \\ \beta_i & \gamma_{ij} \end{pmatrix}$$

$$g^{\mu\nu} = \begin{pmatrix} -\frac{1}{\alpha^2} & \frac{\beta^i}{\alpha^2} \\ \frac{\beta^i}{\alpha^2} & \gamma^{ij} - \frac{\beta^i \beta^j}{\alpha^2} \end{pmatrix}$$

Where $n = -\alpha \nabla t = \frac{1}{\alpha}(1, -\beta^i)$ is the unit vector to the hypersurface.

- ADM leads to Baumgarte, Shapiro, Shibata, Nakamura (BSSN) formalism by introducing conformally related variables

$$\tilde{\gamma}_{ij} = e^{-4\phi} \gamma_{ij}, \quad \tilde{A}_{ij} = e^{-4\phi} \left(K_{ij} - \frac{1}{3} \gamma_{ij} K \right), \quad \tilde{\Gamma}^i = -\partial_j \tilde{\gamma}^{ij}$$

where ϕ the conformal factor satisfying $\det(\tilde{\gamma}_{ij}) = 1$.

Evolution Equations

$$\begin{aligned}
 \partial_t \phi &= -\frac{1}{6} \alpha K + \beta^k \partial_k \phi + \frac{1}{6} \partial_k \beta^k, \\
 \partial_t \tilde{\gamma}_{ij} &= -2\alpha \tilde{A}_{ij} + \tilde{\gamma}_{ik} \partial_j \beta^k + \tilde{\gamma}_{jk} \partial_i \beta^k \\
 &\quad - \frac{2}{3} \tilde{\gamma}_{ij} \partial_k \beta^k + \beta^k \partial_k \tilde{\gamma}_{ij}, \\
 \partial_t K &= -D^i D_i \alpha + \beta^i \partial_i K \\
 &\quad + \alpha \left(\tilde{A}_{ij} \tilde{A}^{ij} + \frac{1}{3} K^2 \right), \\
 \partial_t \tilde{A}_{ij} &= e^{-4\phi} \left[-D_i D_j \alpha + \alpha (R_{ij} - 8\pi S_{ij}) \right] \\
 &\quad + \alpha (K \tilde{A}_{ij} - 2 \tilde{A}_{ik} \tilde{A}^k{}_j) + \beta^k \partial_k \tilde{A}_{ij} \\
 &\quad + \tilde{A}_{ik} \partial_j \beta^k + \tilde{A}_{jk} \partial_i \beta^k, \\
 \partial_t \tilde{\Gamma}^i &= -2 \tilde{A}^{ij} \partial_j \alpha + \beta^j \partial_j \tilde{\Gamma}^i - \tilde{\Gamma}^j \partial_j \beta^i \\
 &\quad + 2\alpha \left(\tilde{\Gamma}^i{}_{jk} \tilde{A}^{jk} - \frac{2}{3} \tilde{\gamma}^{ij} \partial_j K \right) \\
 &\quad + \frac{2}{3} \tilde{\Gamma}^i \partial_j \beta^j + \tilde{\gamma}^{jk} \partial_j \partial_k \beta^i.
 \end{aligned}$$

Constraint Equations

$$\begin{aligned}
 \mathcal{H} &\equiv R + K^2 - K_{ij} K^{ij} - 16\pi \rho = 0 \\
 \mathcal{M}^i &\equiv D_j (K^{ij} - \gamma^{ij} K) - 8\pi S^i = 0
 \end{aligned}$$

Typical Gauge Choices

- *1+log slicing (lapse):*

$$\partial_t \alpha = -2\alpha K + \beta^i \partial_i \alpha.$$

- *Gamma-driver shift condition:*

$$\begin{aligned}
 \partial_t \beta^i &= \frac{3}{4} B^i, \\
 \partial_t B^i &= \partial_t \tilde{\Gamma}^i - \eta B^i,
 \end{aligned}$$

where η is a damping parameter.

- Since every step involves differential equations, PINNs offer a natural alternative to numerical relativity.
- Despite the impressive advances in waveforms generation through PN, EOB and NR, challenges such as mismatches and accuracy problems are still present. These issues could become critical as the next generations of GW detectors achieve unprecedented precision. It is in this context that exploring alternative approaches will become essential and therefore, worth to be developed.
- The most natural problem to address is the relativistic 2-body dynamics.
- We begin by simplifying the problem for exploratory purposes, considering the post-Newtonian approximation to general relativity.

Post-Newtonian expansion

- In General Relativity, the exact equations of motion are too complex to solve for most systems.
- For systems where the gravitational field is weak and velocities are small compared to the speed of light ($v \ll c$), an expansion around the Newtonian limit provides excellent approximations.
- Expand in powers of (v/c) such that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ for $h_{\mu\nu} = \mathcal{O}\left(\frac{v^2}{c^2}\right)$.

PN Order	Approximation	Meaning	Typical Terms / Effects
0PN	Newtonian	Classical gravity	$\Phi = -\frac{GM}{r}$
1PN	$(v/c)^2$ corrections	Special relativistic and curved-space effects	Mercury's perihelion precession, gravitational time dilation
1.5PN	$(v/c)^3$	Gravitational radiation reaction appears	Energy loss due to gravitational waves
2PN-3PN	Higher orders	Precise corrections for compact binaries	Accurate waveform models for inspiraling black holes and neutron stars

Post-Newtonian expansion

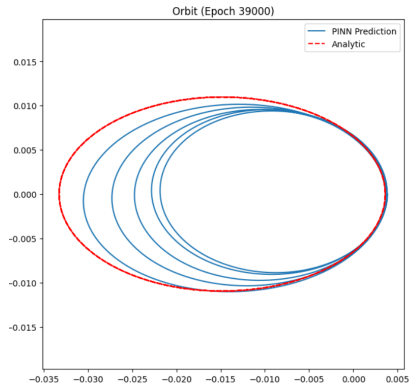
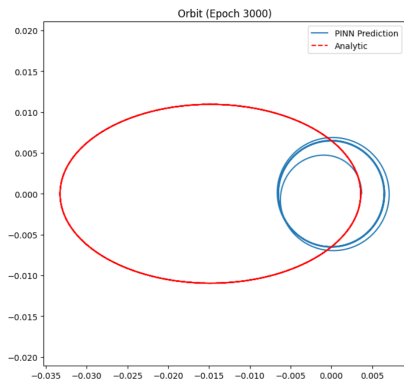
- For instance, the eom up to 2PN is (center-of-mass frame)

$$\begin{aligned} \frac{d^2 \mathbf{r}}{dt^2} = & -\frac{GM}{r^2} \hat{\mathbf{r}} \left\{ 1 - \frac{1}{c^2} \left[\frac{4GM}{r} - v^2 + 4(\mathbf{v} \cdot \hat{\mathbf{r}})^2 \right] \right. \\ & \left. + \frac{1}{c^4} \left[\frac{15}{4} \left(\frac{GM}{r} \right)^2 + \frac{9}{2} \frac{GM}{r} v^2 - \frac{15}{2} \frac{GM}{r} (\mathbf{v} \cdot \hat{\mathbf{r}})^2 + \frac{3}{8} v^4 \right] + \mathcal{O}\left(\frac{1}{c^6}\right) \right\}. \end{aligned}$$

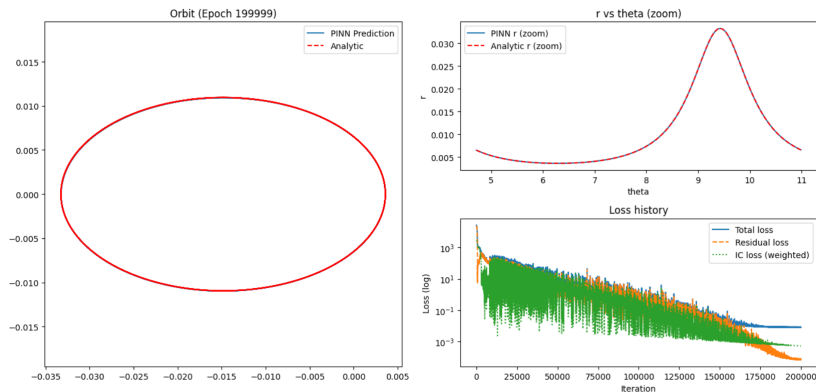
- Applications:
 - Binary pulsar dynamics and gravitational-wave templates.
 - Modeling of coalescing black holes and neutron stars.
- Conservative dynamics; radiation reaction enters at 1.5PN.
- The approach starts to fail at 5PN for the merger phase, so the use of NR becomes necessary.

Gravity-informed Neural Networks

- As a warm up, we can train the neural network to provide Mercury's orbit.
- $M_{sun} = 1,9 \times 10^{30}$, $M_{mercury} = 3,2 \times 10^{23}$, $a = 0.387$, $e = 0.21$.
- 1PN correction (perihelion precession) needs over 1000 orbits to become obvious in the simulation.



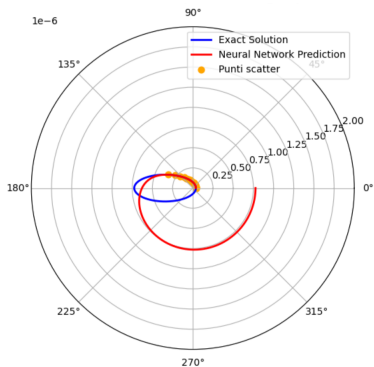
Gravity-informed Neural Networks



- To have into consideration: weighted loss functions, architecture of the NN, initial conditions, ...

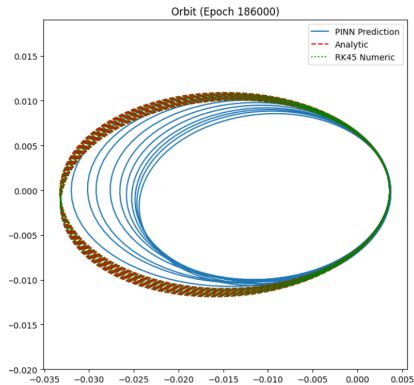
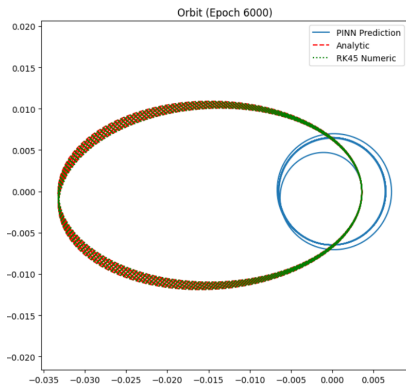
Gravity-informed Neural Networks

- What happens if we consider a vanilla NN?

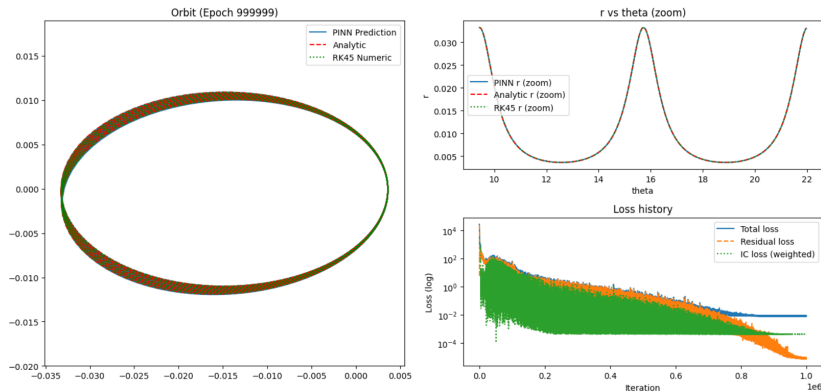


Gravity-informed Neural Networks

- Considering the same setup, we can artificially magnify the relativistic 1PN effects by decreasing the value of the speed of light.



Gravity-informed Neural Networks



- This proposal is difficult to generalize to the full Einstein equations.

PN expansion and PINNs: Symplectic Hamiltonian

- Symplectic Hamiltonian Neural Networks (SHNN) enforce the *Hamiltonian structure* $H(\mathbf{q}, \mathbf{p})$, ensuring long-term stability and physical consistency.
- Compute dynamics via the symplectic form:

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}.$$

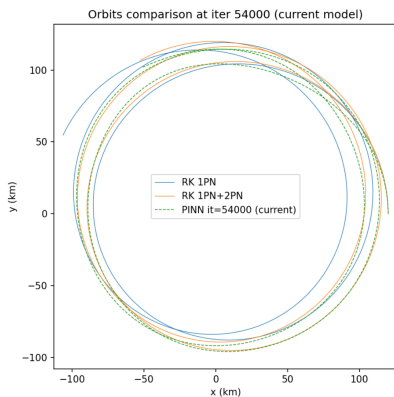
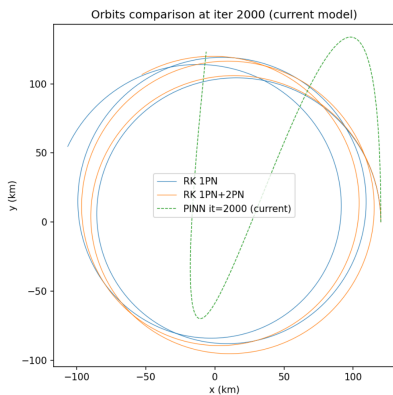
- So that, the loss function is given by

$$\mathcal{L} = \left\| \dot{\mathbf{q}} - \frac{\partial H}{\partial \mathbf{p}} \right\|^2 + \left\| \dot{\mathbf{p}} + \frac{\partial H}{\partial \mathbf{q}} \right\|^2$$

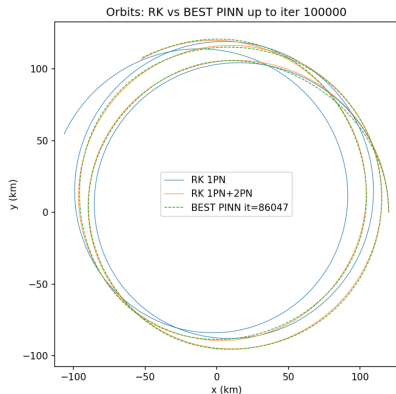
- **Advantages:** Preserves symplectic structure and energy over long trajectories, provides physically consistent extrapolation, ideal for learning conservative dynamical systems.
- Easily generalized via Hamiltonian corrections; compatible with EOB implementation.

Symplectic Hamiltonian: Results

- The setup is more interesting: binary system with $m = 1.4 M_{sun}$ each, $d = 120$ Km.
- Better control on the initial conditions, conservation of energy and angular momentum in the loss functions (closer to NR).



Symplectic Hamiltonian: Results



- To have into consideration: weighted loss functions, architecture of the NN, initial conditions, ...

Symplectic Hamiltonian: Transfer Learning

- To facilitate training on the full Einstein equations, we propose an incremental approach: pretrain the network on simplified data and refine it through [Transfer Learning](#).

- **Motivation:**

- Solving PDEs from scratch is costly.
- Many PDEs share similar structures (e.g., parameters, domains, boundary conditions).
- Use a previously trained PINN as a starting point.

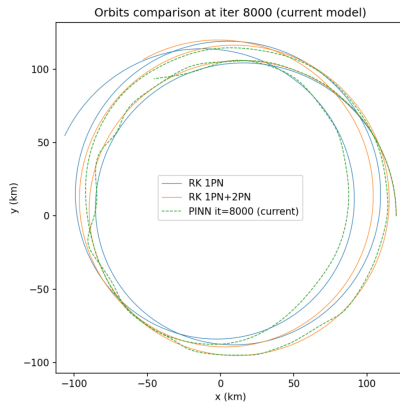
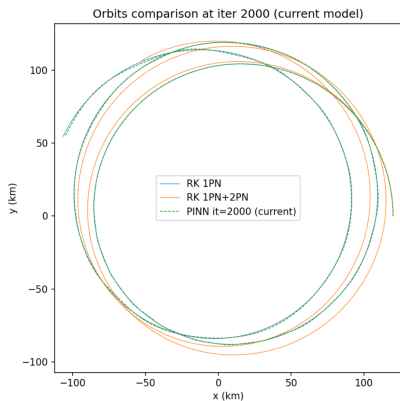
- **Procedure:**

- Train a PINNs on a source problem by minimizing $\mathcal{L}_{\text{pre-trained}}^{(S)}$.
- Then transfer this knowledge by minimizing $(1 - \mathcal{N}_T)\mathcal{L}_{\text{physics}}^{(T)} + \mathcal{L}_{\text{data}}^{(T)}$, where $\mathcal{N}_T$ goes from 0 to 1 in a controlled way.

- **Benefit:** Faster convergence and better generalization to new but related physics.

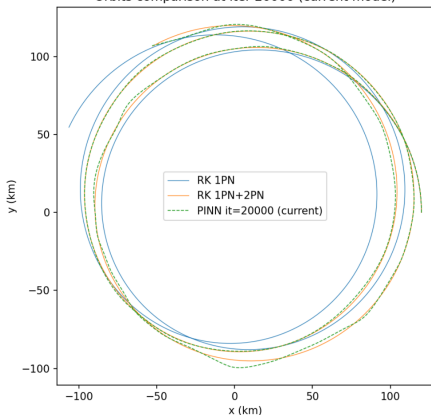
Symplectic Hamiltonian: Transfer Learning results

- Binary Black Hole system.

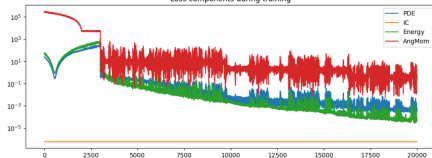


Symplectic Hamiltonian: Transfer Learning results

Orbits comparison at iter 20000 (current model)



Loss components during training



- 17 minutes vs 2 hours.
- To emulate progressively richer gravitational physics, we initially freeze selected biases and weights and then selectively unfreeze them as higher-order interactions are introduced.

Further directions

- Full Einstein's equations.
- Realistic astrophysical scenarios involve the gravitational interaction with matter. Notably, PINNs have already been successfully applied to problems in fluid dynamics.
- Many body problem for structure formation. There are interesting examples.
- Alternative approach: Hamiltonian NN.