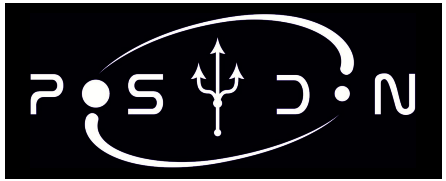
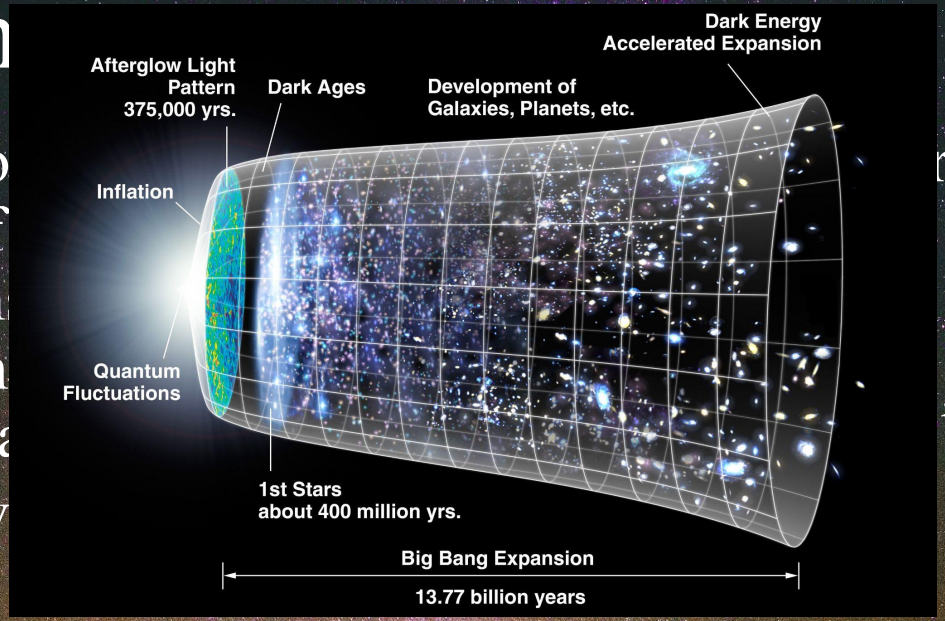
Deep Neural Operators for Detailed Binary Evolution Simulation

Philipp M. Srivastava, Ugur Demir, Patrick Koller, Santiago Lopez, Manuel Ballester, Shamal Lalvani, Aggelos Katsaggelos, and Vicky Kalogera with the POSYDON collaboration



Astronom

- One of the o r fundamental questions of
- Observation visible electromagn
- Now, celesti be studied through grav

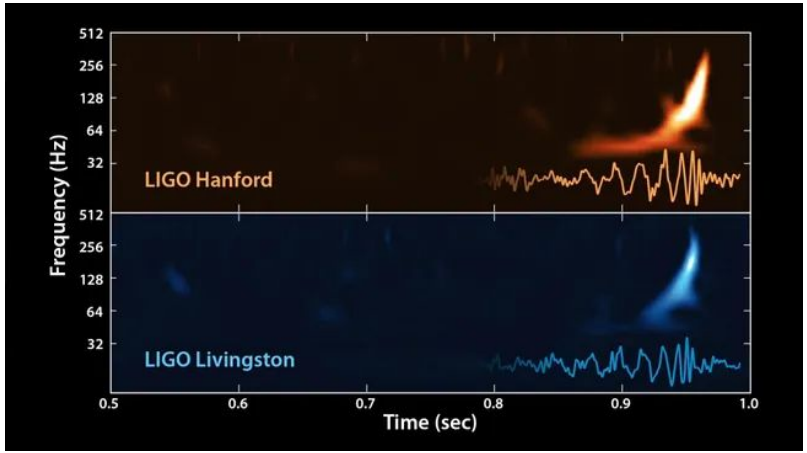


[6]

Photo by Greg Rakozy on Unsplash

Gravitational Waves (GW)

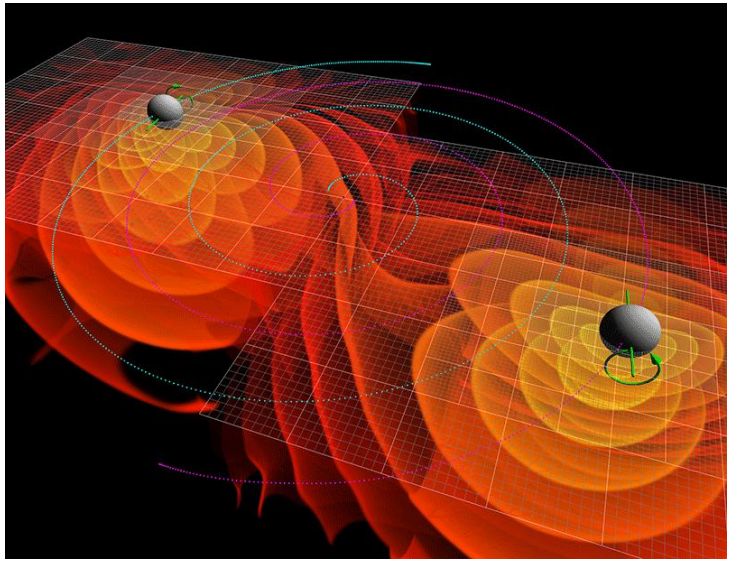
- First detected on September 14, 2015
- Theorized 100 years earlier in Einstein theory of general relativity
- When two compact objects inspiral, they emit GWs
- What kinds of objects create GWs?



[1]

Binary star systems as GW emitters

- Most stars are born as a binary system
- Binary evolution can leave imprints in these GW events
- Binary population synthesis codes allow us to estimate the number of these sources



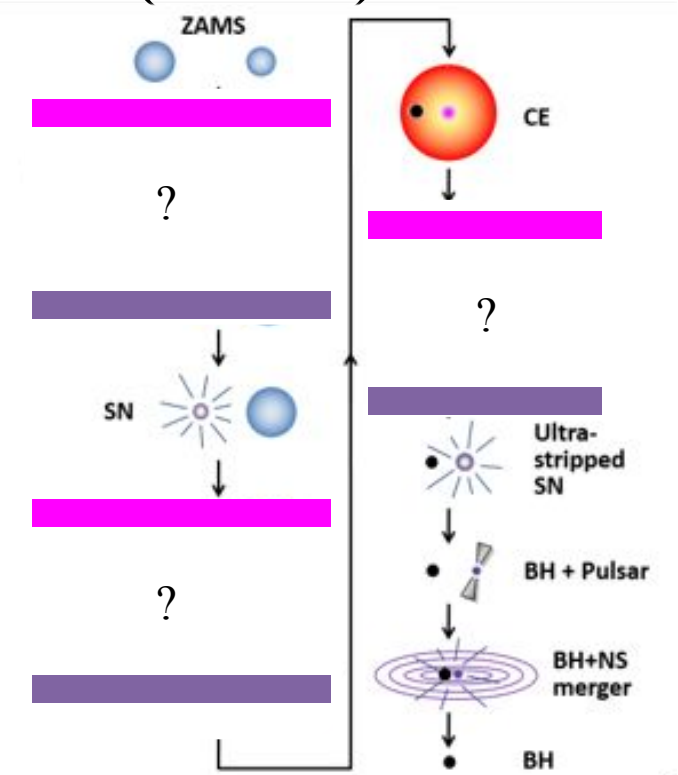
[2]

Outline

1. What is binary population synthesis (BPS)
2. Problem definition
3. Proposed method
4. Results
5. Discussion & future work

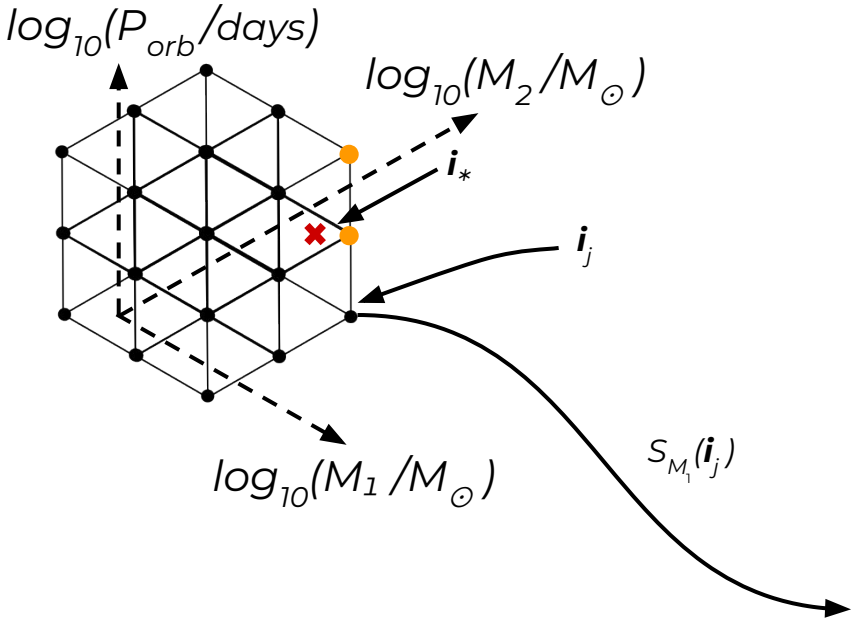
Binary Population Synthesis (BPS)

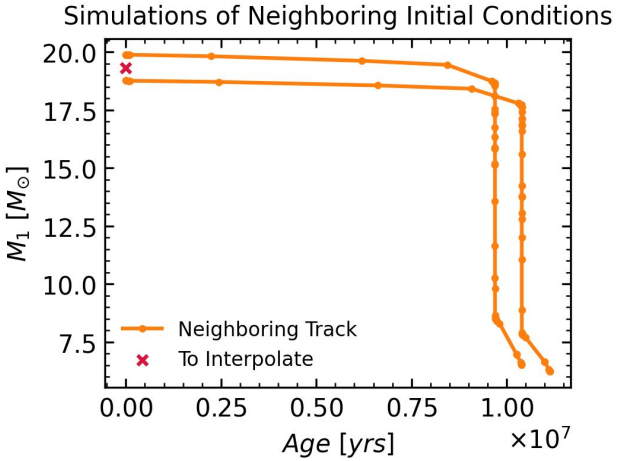
- BPS is the large scale simulation of binary star systems
- Accurate systems omit large parts of binary evolution
- Numerical methods for approximating omitted parts can take up to ~48 hours per binary
- Traditional methods lack sufficient accuracy, don't exploit known physics, and don't account for dependencies between system parameters

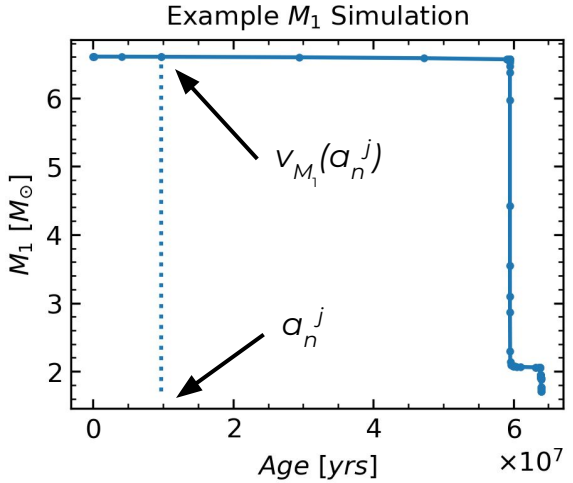


[3, 5]

Problem Definition

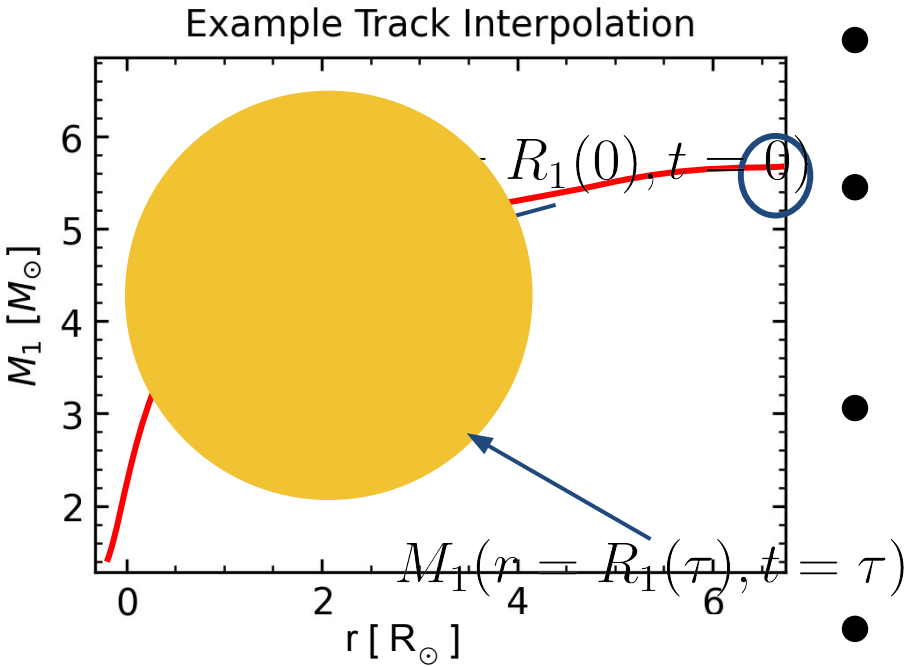






[Srivastava et al. 2025]

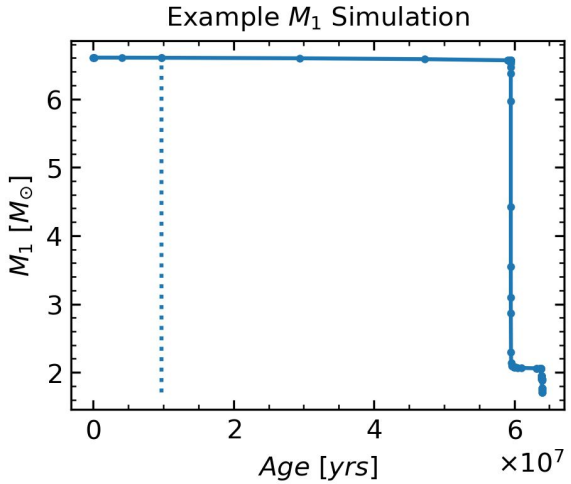
Neural Operator?



- Neural operators are typically used to solve initial value problems
- In the context of partial differential equations (PDEs) this is an initial value function
- We can define the initial value function as $M_1(r, t = 0)$ which we sample at $M_1(r = R_1(0), t = 0)$
- The task becomes to approximate $M_1(r = R_1(\tau), t = \tau)$ at arbitrary τ

Simply put

$$f(\mathbf{i}_*, t) \rightarrow \hat{\mathbf{z}}^t_{\mathbf{i}_*}$$



Initial value
$\mathbf{i}_* \in \mathbb{R}^3$

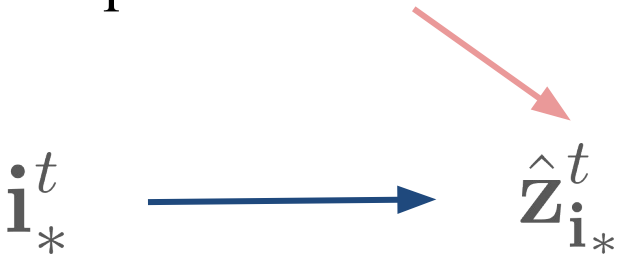
Arbitrary time
$t \in \mathbb{R}$

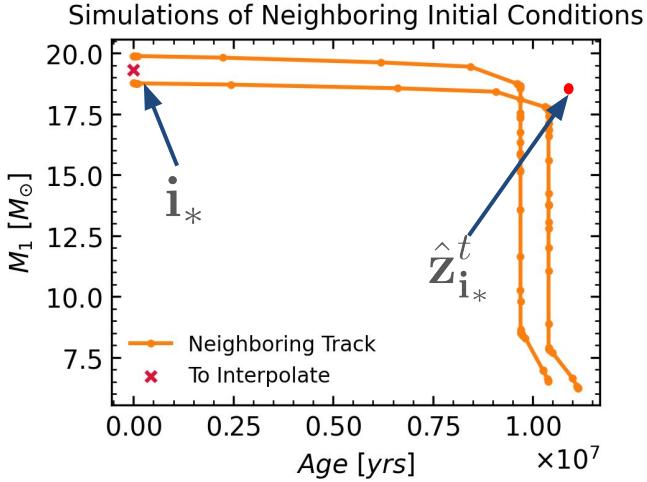
Initial value with time
$\mathbf{i}^t_* = (\mathbf{i}_*, t) \in \mathbb{R}^4$

A different perspective

Let's suppose this is an approximation of our desired output

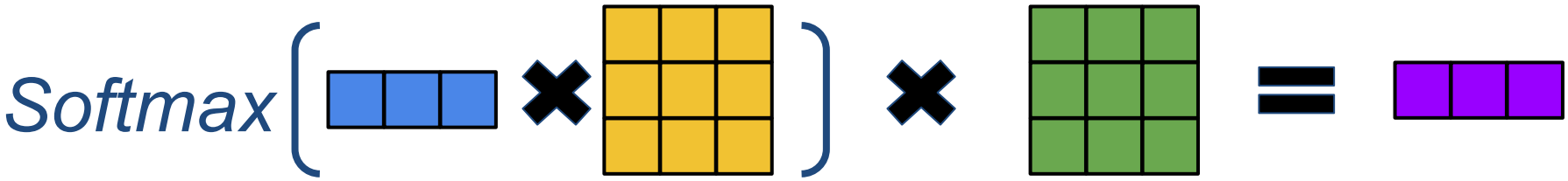
N





Let's see which parts of the neighbors our prediction most resembles and use that

Cross-Attention

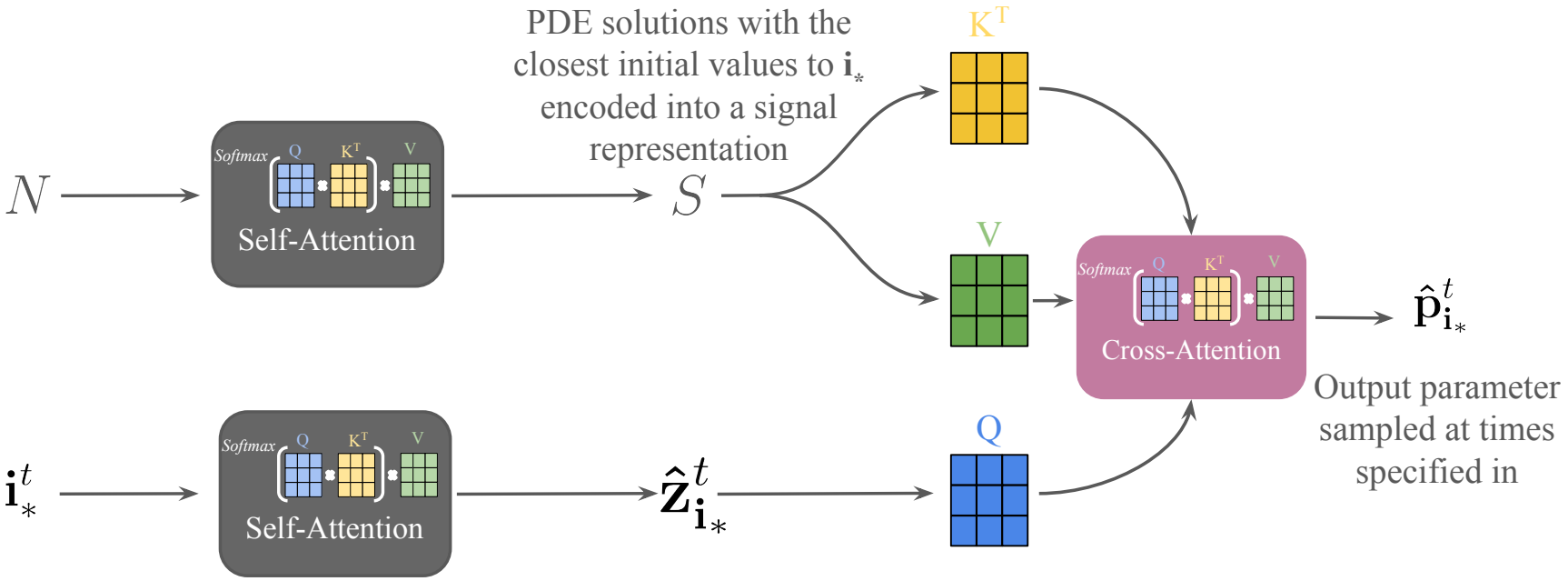


$$Q = \hat{\mathbf{z}}_{\mathbf{i}_*}^t W_Q \quad K = NW_K \quad V = NW_V \qquad \hat{\mathbf{p}}_{\mathbf{i}_*}^t$$

$$W_Q, W_K, W_V \in \mathbb{R}^{(d \times d')}$$

*Followed by fully connected layers

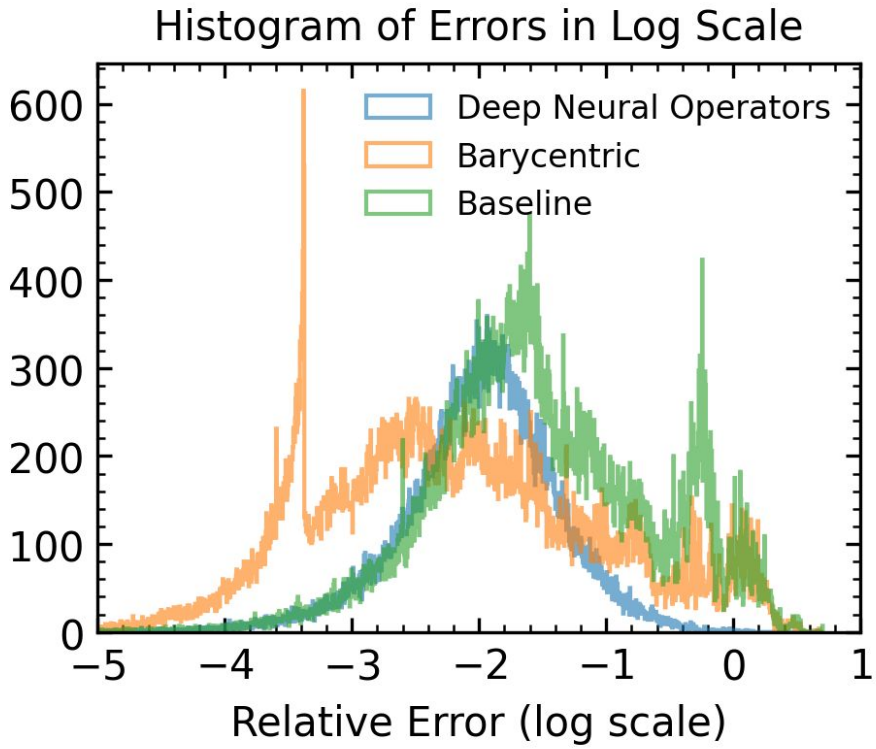
System Design



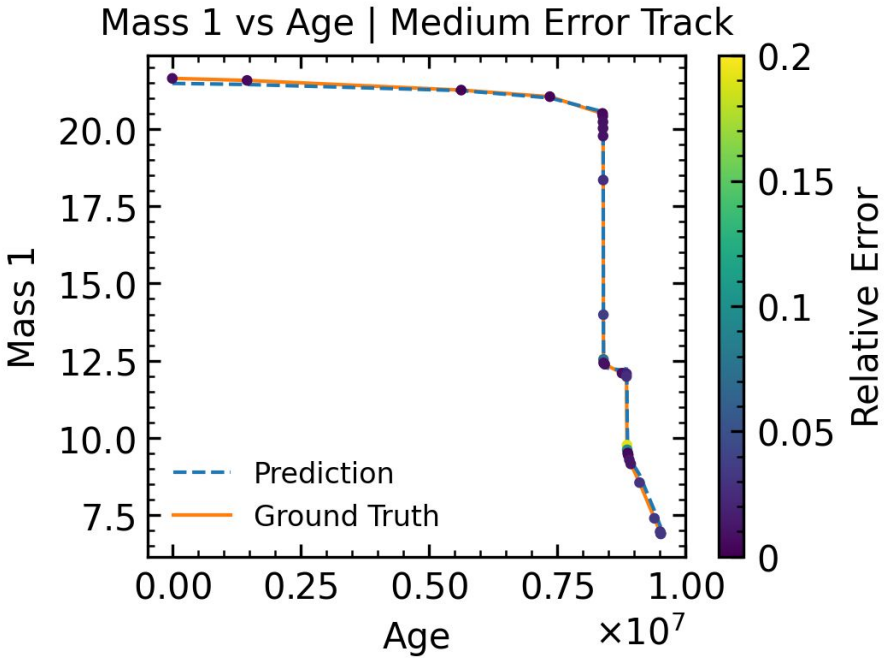
...supported by the Simons Foundation (PD fellowship grant awards GBMF8477 and GBMF12343) and the ... Simons AI Institute for the Sky (SkAI) via grants NSF AST-2421845 and Schmidt Sciences MPS-AI-00010513.

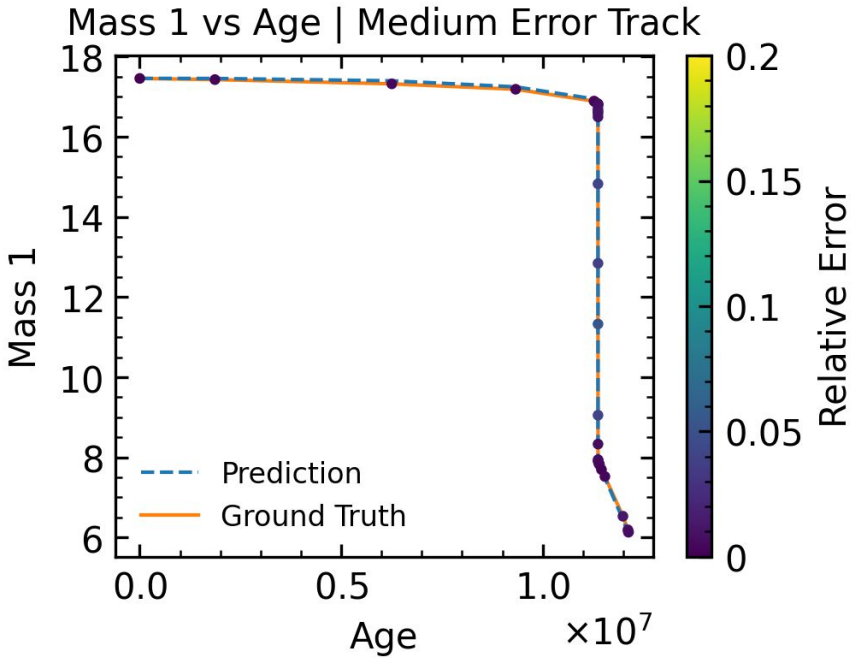
With predicted classes results

	Median	Mean
Barycentric	**0.53%**	10.82%
Neural Ops	1.1%	**2.43%**

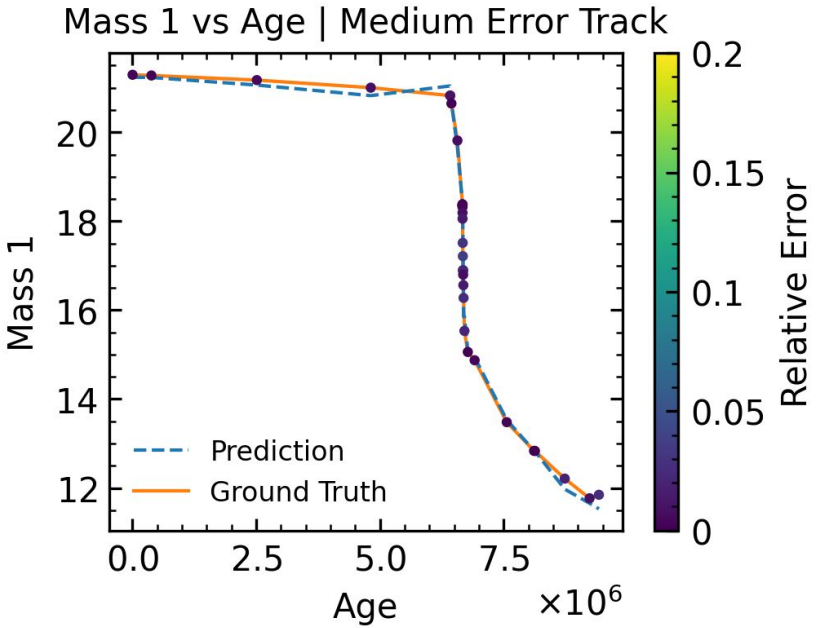


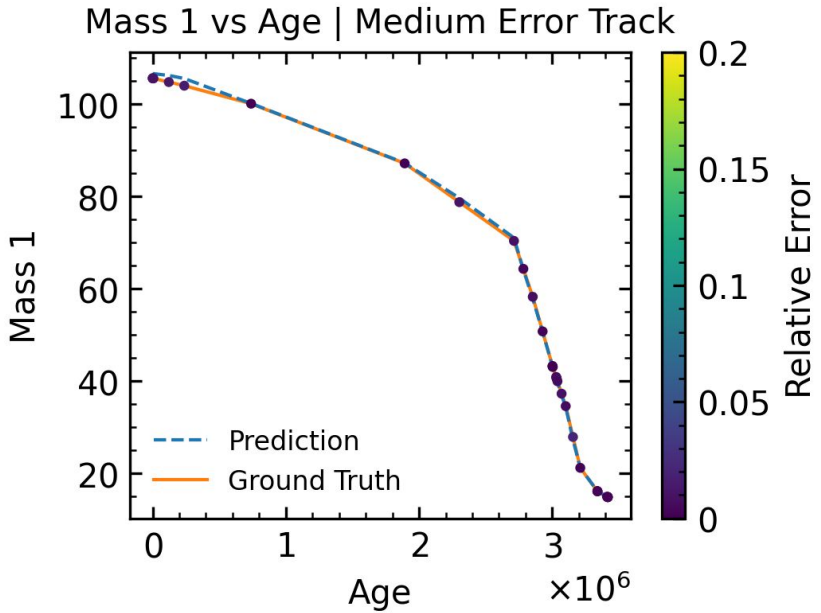
Some tracks


Mass 1 vs Age | Medium Error Track
Mass 1
Age
×10^7
Prediction
Ground Truth
Relative Error
20.0
17.5
15.0
12.5
10.0
7.5
0.00
0.25
0.50
0.75
1.00
0.2
0.15
0.1
0.05
0

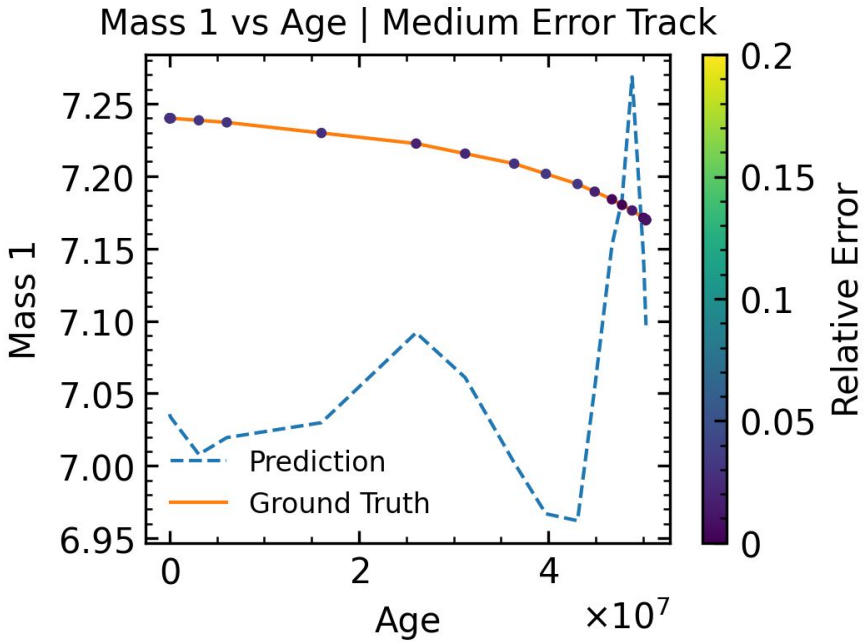

Mass 1 vs Age | Medium Error Track
Mass 1
Age
×10^7
Prediction
Ground Truth
Relative Error
18
16
14
12
10
8
6
0.0
0.5
1.0
0.2
0.15
0.1
0.05
0

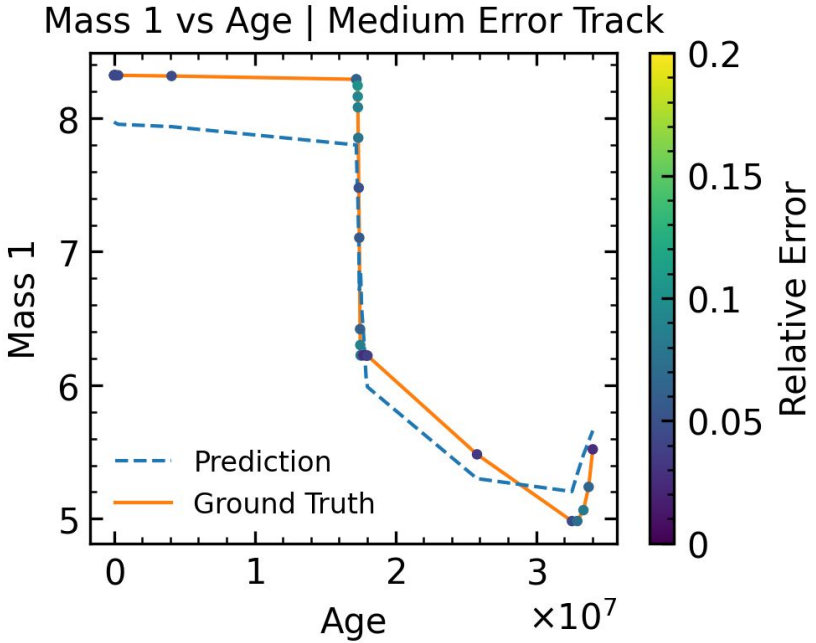
Some tracks


Mass 1 vs Age | Medium Error Track
Mass 1
Age
×10^6
Prediction
Ground Truth
Relative Error


Mass 1 vs Age | Medium Error Track
Mass 1
Age
×10^6
Prediction
Ground Truth
Relative Error

Problematic tracks





Discussion & future work

- Want to predict more parameters simultaneously to exploit dependencies between them
- Using known Physics to train the system is an enticing prospect
- Efficacy needs to be verified through usage by astrophysical community

References

[1] B. P. Abbott *et al.*, “Observation of Gravitational Waves from a Binary Black Hole Merger,” *Physical Review Letters*, vol. 116, no. 6, Feb. 2016, doi: 10.1103/physrevlett.116.061102.

[2] E. Berti, “The first sounds of merging black holes,” *Physics*, Feb. 11, 2016. https://physics.aps.org/articles/v9/17

[3] Tauris, T. M. and van den Heuvel, E. P. J., Physics of Binary Star Evolution. From Stars to X-ray Binaries and Gravitational Wave Sources. 2023. doi:10.48550/arXiv.2305.09388.

[4] Srivastava, P. M., “Irregularly Sampled Time Series Interpolation for Detailed Binary Evolution Simulations”, The Astrophysical Journal, vol. 984, no. 2, Art. no. 154, IOP, 2025. doi:10.3847/1538-4357/adbe6b.

[5] Fragos, T., “POSYDON: A General-purpose Population Synthesis Code with Detailed Binary-evolution Simulations”, The Astrophysical Journal Supplement Series, vol. 264, no. 2, Art. no. 45, IOP, 2023. doi:10.3847/1538-4365/ac90c1.

[6] “Timeline of the Universe image.” https://map.gsfc.nasa.gov/media/060915/index.html

Questions?

Please contact us at philipp.msrivastava@northwestern.edu

Definitions

Initial value

$\mathbf{i}_* \in \mathbb{R}^3$

Arbitrary time

$t \in \mathbb{R}$

Initial value

$\boldsymbol{i}'_* = (\boldsymbol{i}_*, t) \in \mathbb{R}^4$

Neighbors defined as

$\mathbf{n}_1, \mathbf{n}_2 \in \mathbb{R}^{(32 \times d)}$

Concatenated neighbors

$\mathbf{N} \in \mathbb{R}^{([2 \cdot 32] \times d)}$

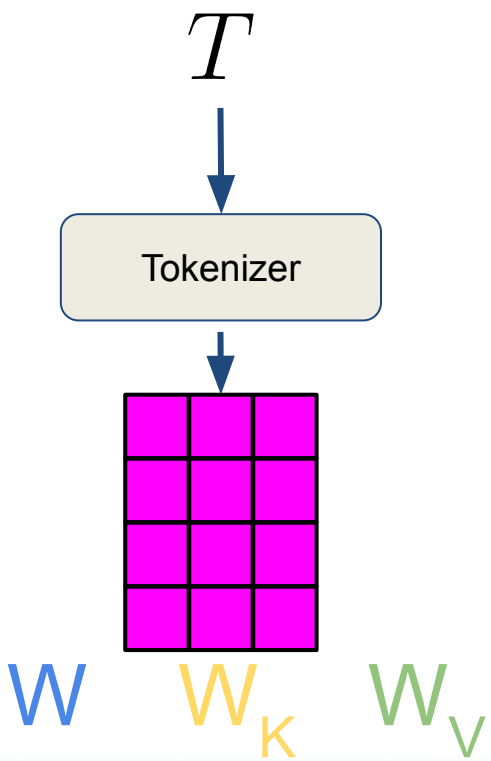
Arbitrary track

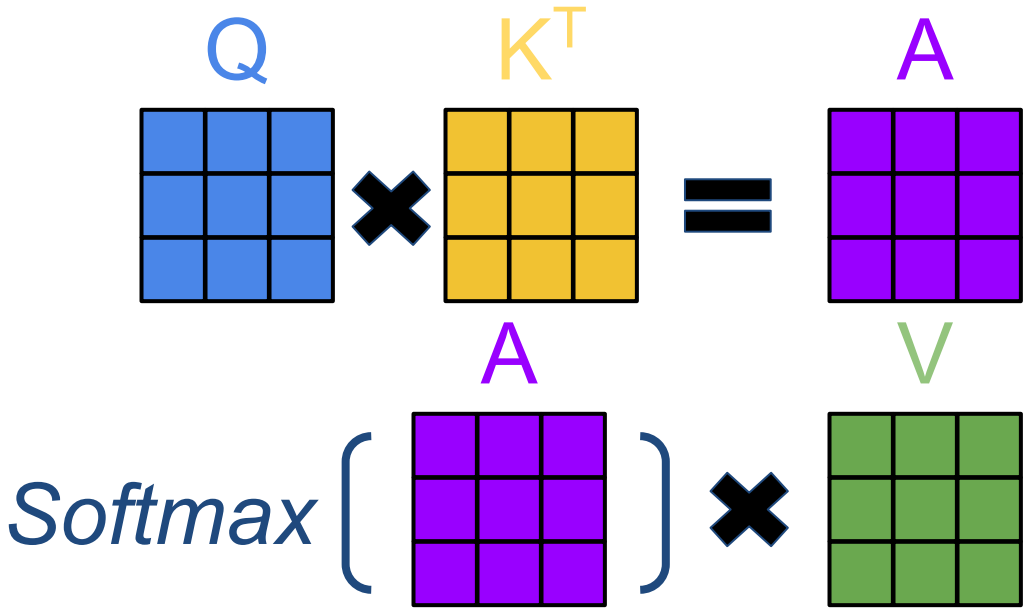
$T \in \mathbb{R}^{(32 \times d)}$

Embedded vector

$\mathbf{z}^{(\cdot)}_{(\cdot)}$

Self-Attention





*Followed by fully connected layers