Introduction to physics informed machine learning

Dr. Suchita Kulkarni

What is physics informed machine learning?

- Machine learning: art of making data driven decisions
  - Works incredibly well in presence of lot of data
  - May inherit data biases, may not extrapolate well
  - Explainability is challenge: feels like a blackbox
- Physics informed machine learning (PIML)
  - Art of injecting known knowledge of physical systems into machine learning
  - Suffers less from data biases, and extrapolates better
  - Contains built-in explainability

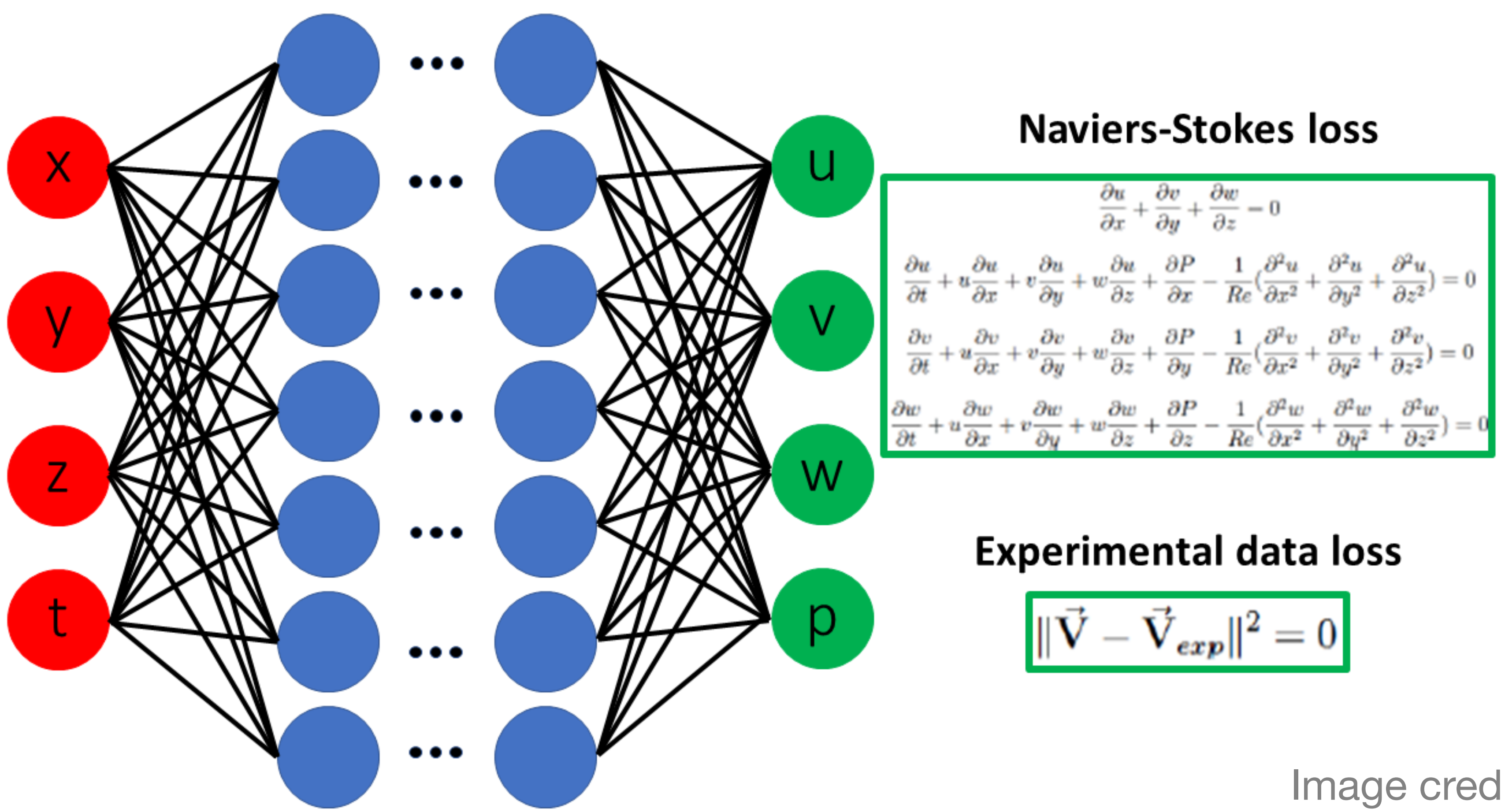


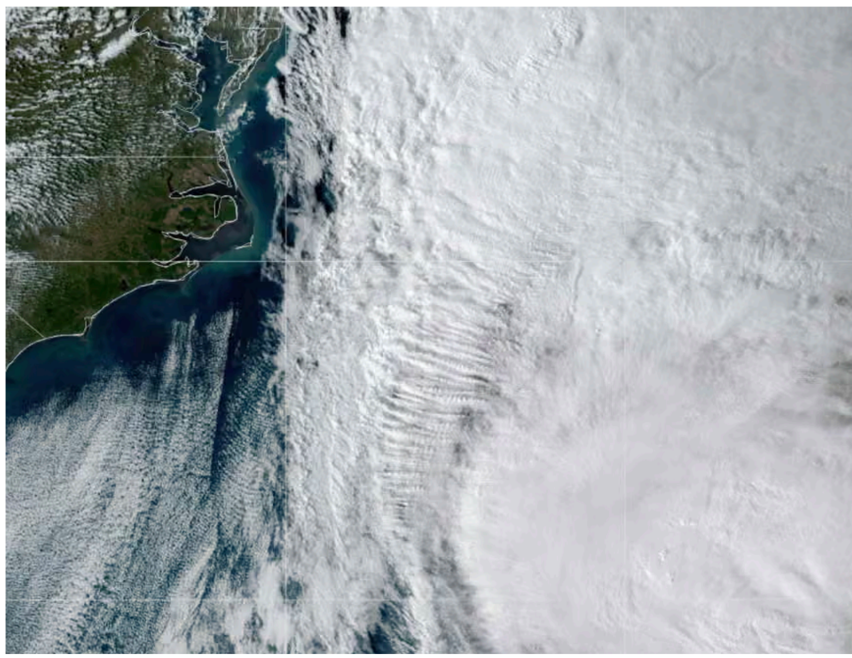
Image credit: Riccardo Munafò

PIML success stories

How Google's DeepMind tool is 'more quickly' forecasting hurricane behavior

'Less expensive and time consuming' model helps with fast and accurate predictions, possibly saving lives and property

Guardian



arXiv:2507.11589

EINSTEIN FIELDS: A NEURAL PERSPECTIVE TO COMPUTATIONAL GENERAL RELATIVITY

Sandeep S. Cranganore[*1] **Andrei Bodnar**[*2] **Arturs Berzins**[*1]
Johannes Brandstetter[1,3]

[*]Equal contribution
[1]LIT AI Lab, Institute for Machine Learning, JKU Linz, Austria
[2]University of Manchester, United Kingdom
[3]Emmi AI GmbH, Linz, Austria
{cranganore, berzins, brandstetter}@ml.jku.at
andrei.bodnar@student.manchester.ac.uk

AURORA: A FOUNDATION MODEL OF THE ATMOSPHERE

Cristian Bodnar[*,1], **Wessel P. Bruinsma**[*,1], **Ana Lucic**[*,1], **Megan Stanley**[*,1], **Johannes Brandstetter**[3,†], **Patrick Garvan**[1], **Maik Riechert**[1], **Jonathan Weyn**[2], **Haiyu Dong**[2], **Anna Vaughan**[4], **Jayesh K. Gupta**[5,†], **Kit Tambiratnam**[2], **Alex Archibald**[4], **Elizabeth Heider**[1], **Max Welling**[6,†], **Richard E. Turner**[1,4], **and Paris Perdikaris**[1]

[1]Microsoft Research AI for Science
[2]Microsoft Corporation [3]JKU Linz [4]University of Cambridge [5]Poly Corporation [6]University of Amsterdam

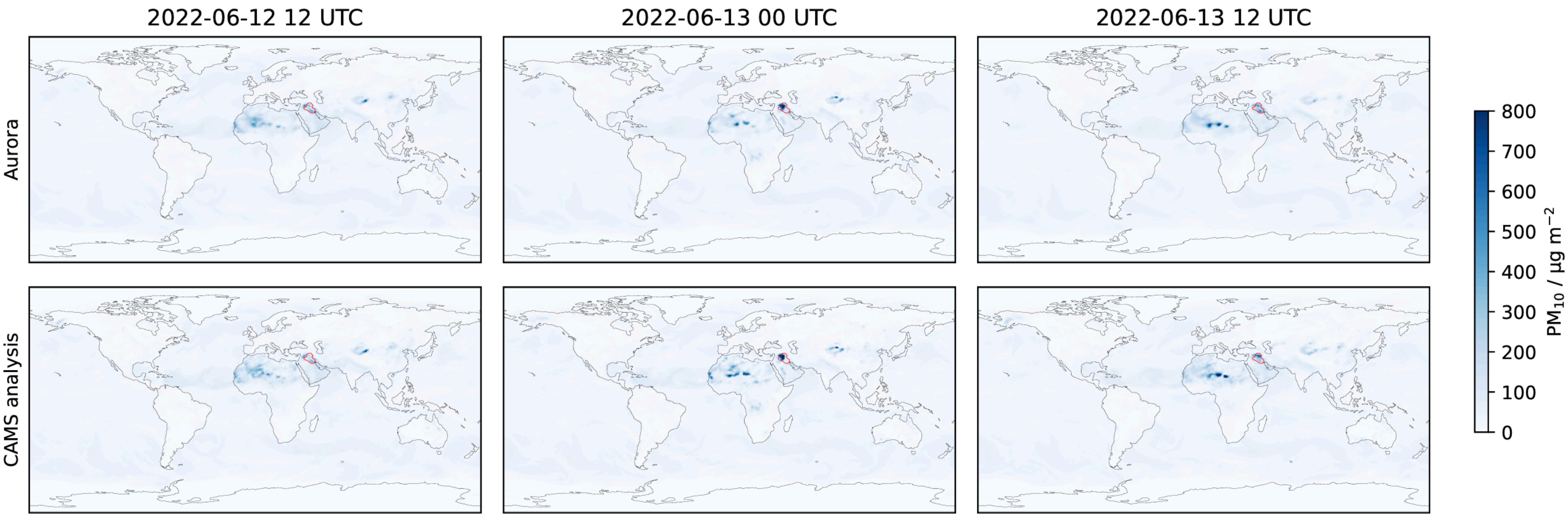
[*]Equal contribution [†]Work done while at Microsoft Research

arXiv:2405.13063

PIML success stories

Aurora see storms brewing a day in advance

arXiv:2405.13063

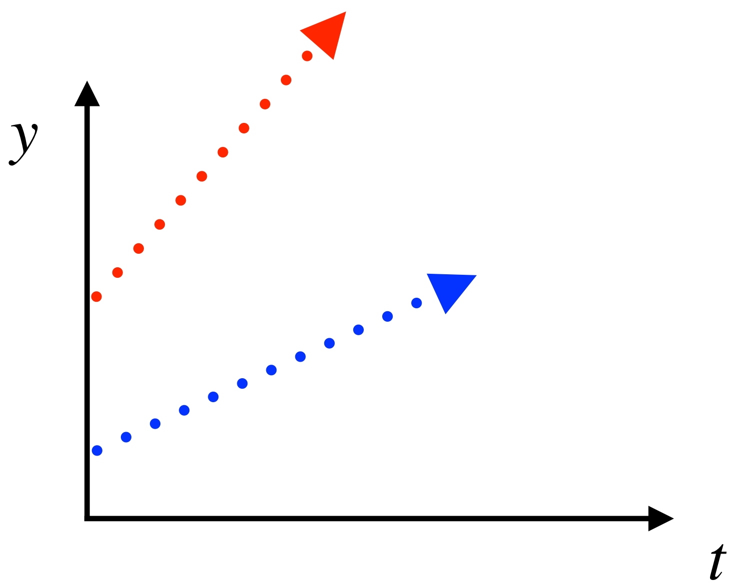


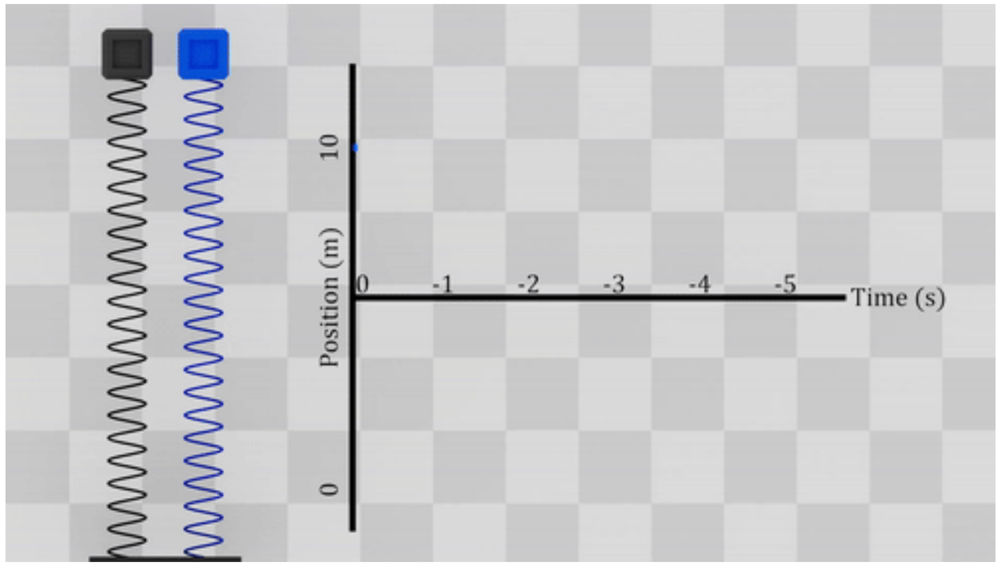
Both Aurora and CAMS initialised at 12 Jun 2022 00 UTC

The physics

- (Partial) differential equations are the backbone of physics
  - Tell us how physical systems change across space and time

$$\frac{dy}{dt} = v \Rightarrow y - y_0 = v \times t \qquad\qquad \frac{d^2y}{dt^2} + \omega^2 y = 0$$





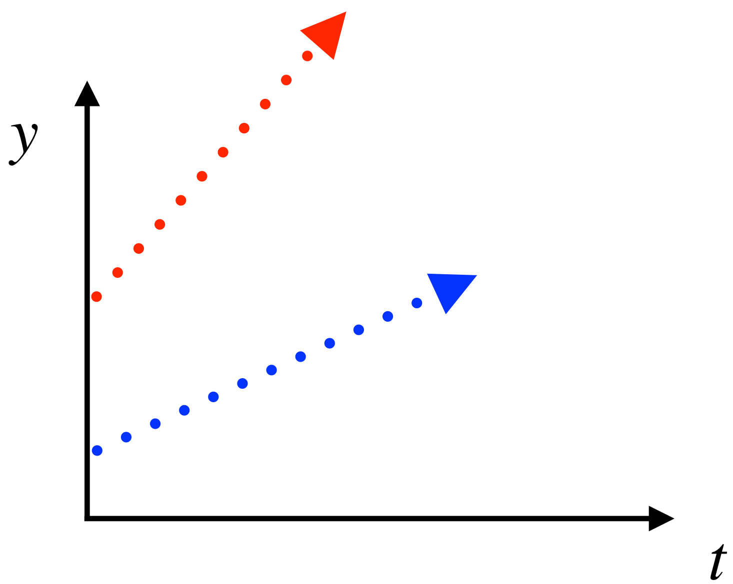
Source wikipedia

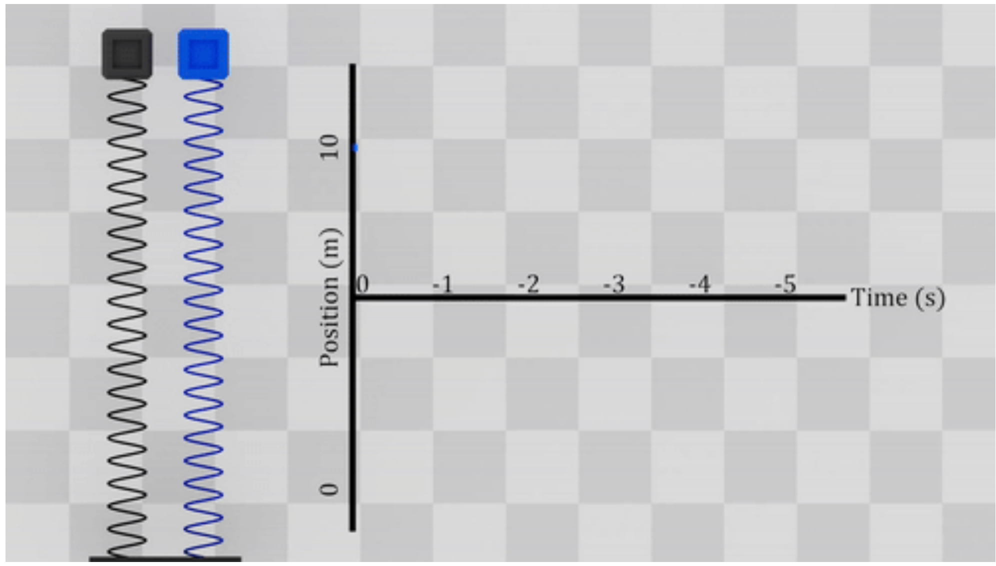
- But apart from partial differential equations, there is a lot of other physics information available
  - Symmetries (not covered today)
  - Conserved quantities (not covered today)

The physics

- (Partial) differential equations are the backbone of physics
  - Tell us how physical systems change across space and time

$$\frac{dy}{dt} = v \Rightarrow y - y_0 = v \times t \qquad\qquad \frac{d^2y}{dt^2} + \omega^2 y = 0$$





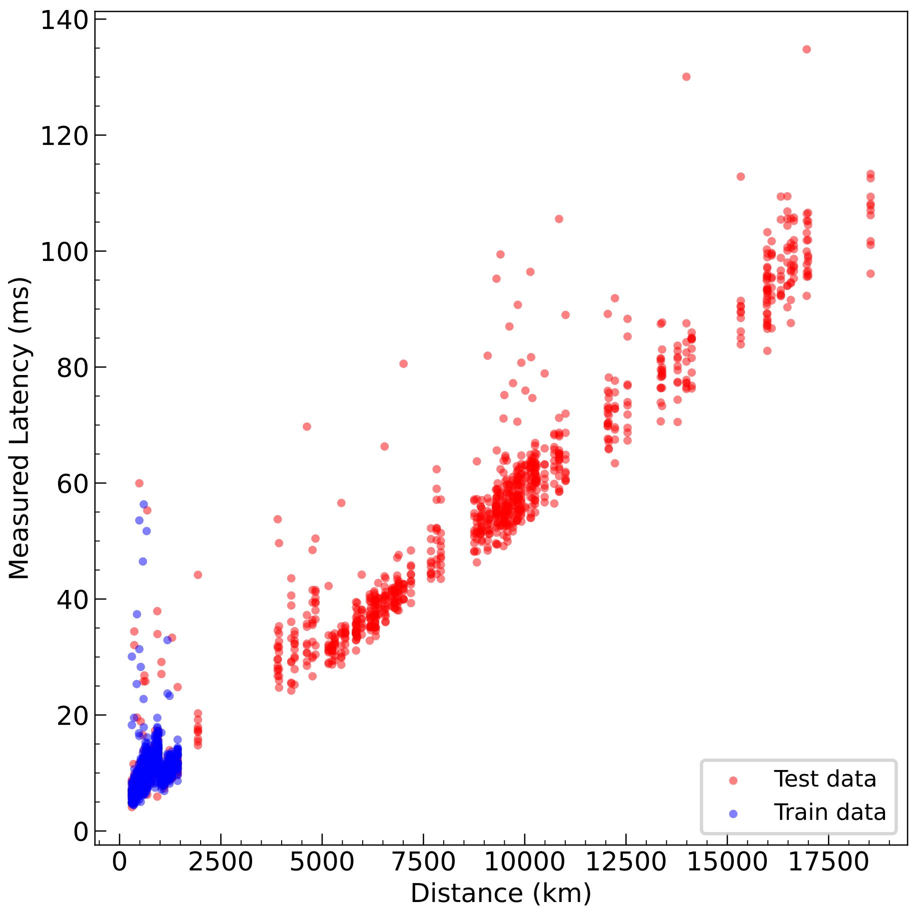
Source wikipedia

- But apart from partial differential equations, there is a lot of other physics information available
  - Symmetries (not covered today)
  - Conserved quantities (not covered today)

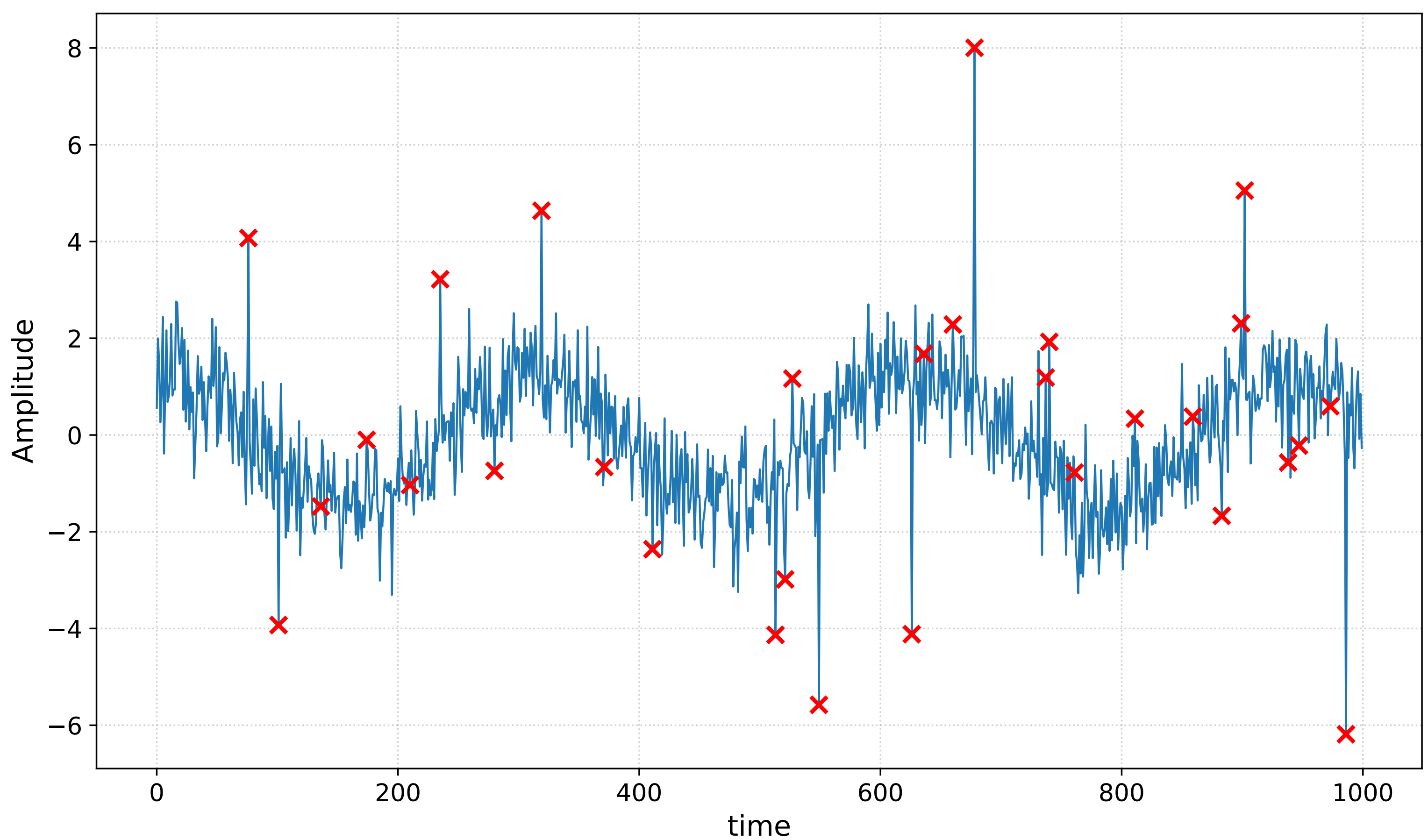
The problem

- Real life data is hardly ever clean
- Real life data is incomplete or sparse or anomalous

Sparse data



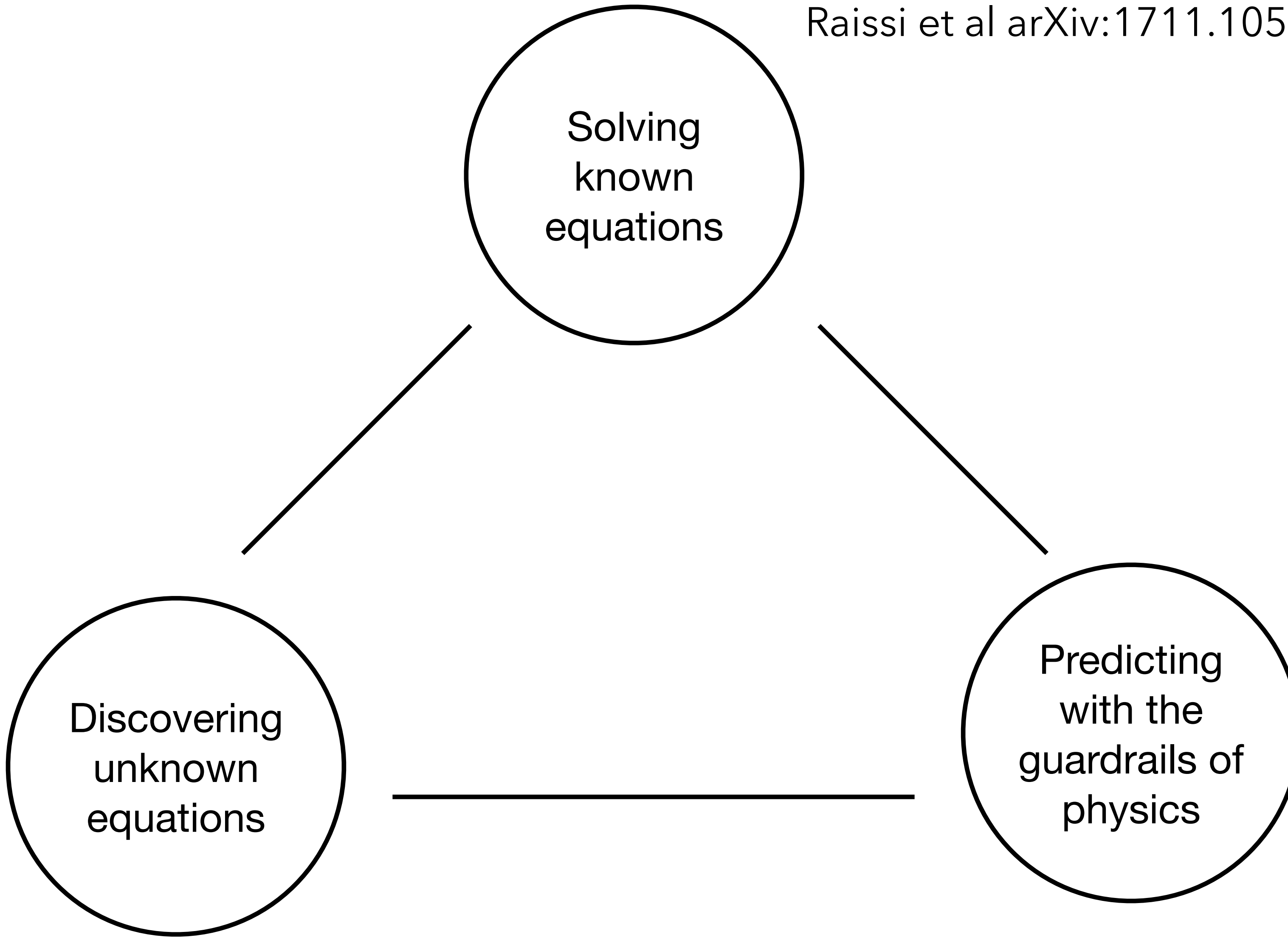
Noisy data



- Often solving differential equations is not sufficient, they need to be aided with some machine learning techniques and this leads to the idea of physics inspired machine learning

The many aspects of PIML

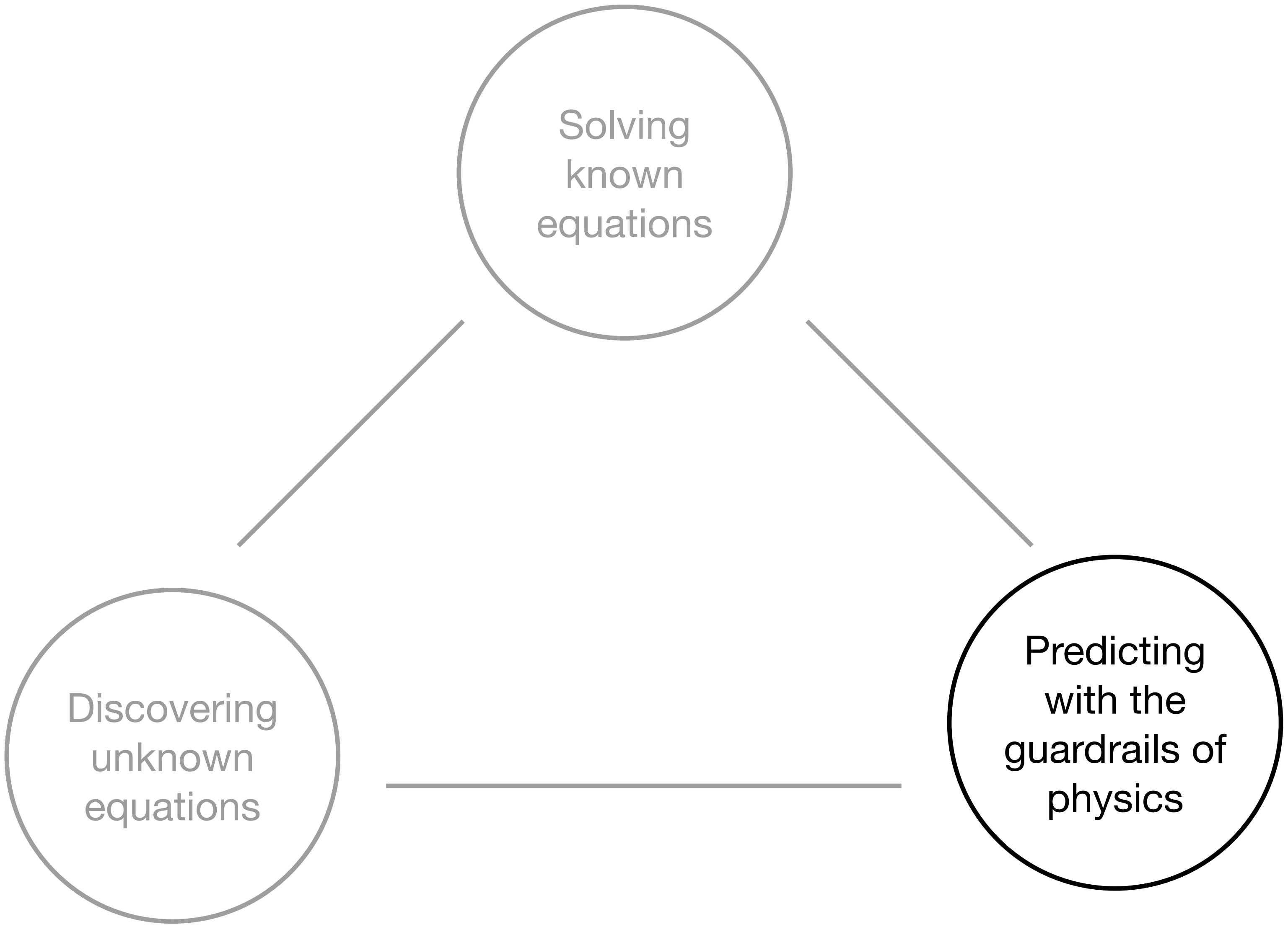
Raissi et al arXiv:1711.10561



Brunton et al, arXiv:1509.03580
Chen et al arXiv:1806.07366

Greydanus et al arXiv:1906.01563
Cranmer et al arXiv: 2003.04630

Plan for today



Use controlled synthetic signals to isolate the effect of the physics constraints. In a real production setting substitute your domain’s signals and keep the same modelling principles.

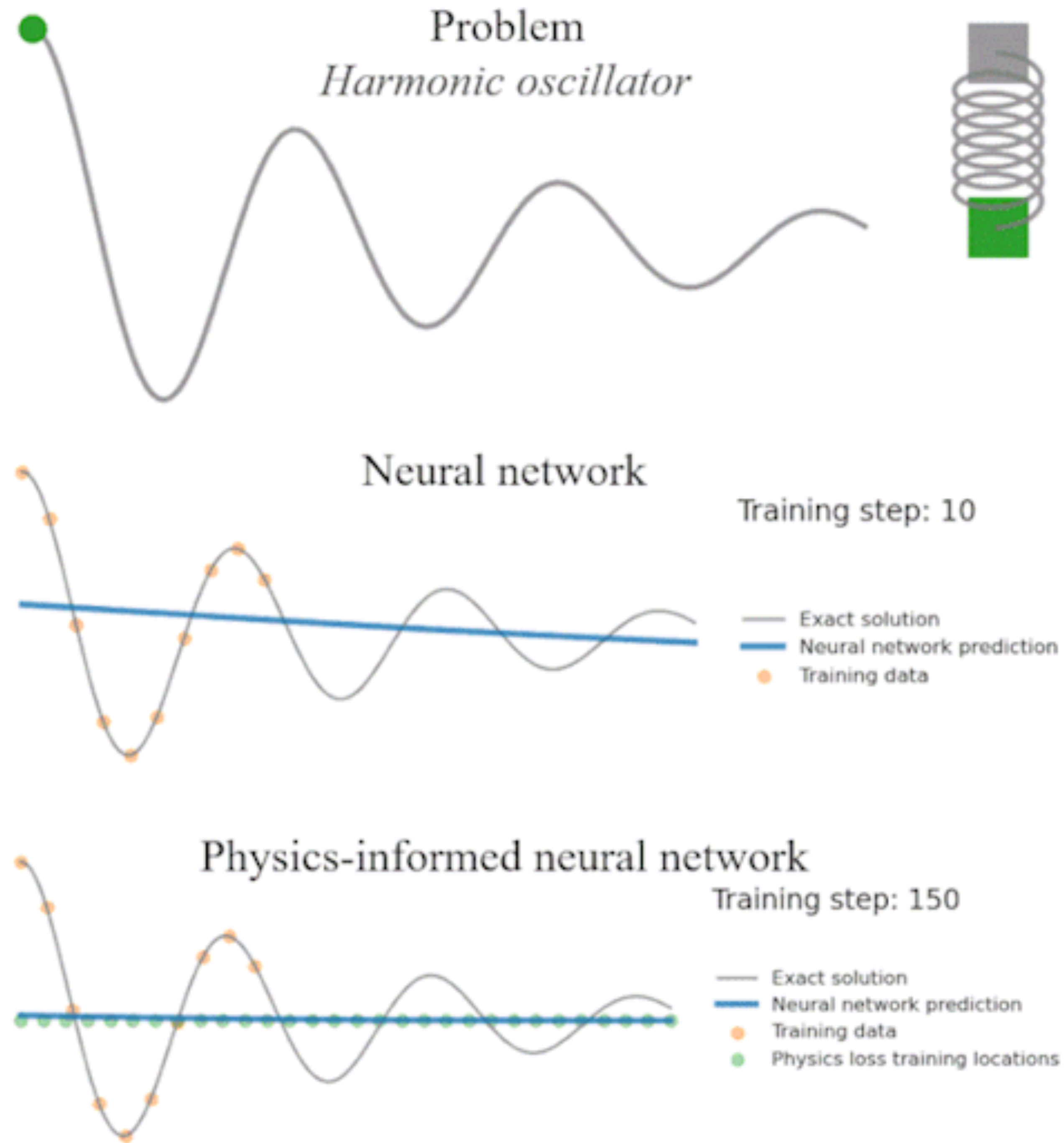
Predictions with physics guardrails

- The central objective of any machine learning project is minimisation of loss function
- For example
  - MSE: mean squared loss, difference between observed and reconstructed data
    - $\mathscr{L}_{\mathrm{MSE}} = \frac{1}{N} || y_{\mathrm{recon}} - y_{\mathrm{obs}} ||$
  - Kullback - Leibler divergence: measure of difference between two probability distributions
    - $\mathscr{L}_{\mathrm{KL}} = \mathscr{D}_{\mathrm{KL}}(p(x) || q(x)) = \int p(x) \log \frac{p(x)}{q(x)} dx$

Linear regression	Neutral network LSTM	Neutral network vAE
$\mathscr{L}_{obj} = \mathscr{L}_{\mathrm{MSE}}$ OR physics $\mathscr{L}_{phy} = t_{\mathrm{obs}} - t_{\mathrm{pred}}$ $= t_{\mathrm{obs}} - x_{\mathrm{obs}}/v_{\mathrm{known}}$	$\mathscr{L}_{obj} = \mathscr{L}_{\mathrm{MSE}}$ + physics $\mathscr{L}_{phy} = \frac{d^2 y_{\mathrm{recon}}}{dt^2} + \omega^2 y_{\mathrm{recon}}$	$\mathscr{L}_{obj} = \mathscr{L}_{\mathrm{MSE}} + \mathscr{L}_{\mathrm{KL}}$ + physics $\mathscr{L}_{phy} = \frac{d^2 y_{\mathrm{recon}}}{dt^2} + \omega^2 y_{\mathrm{recon}}$

Why does physics help?

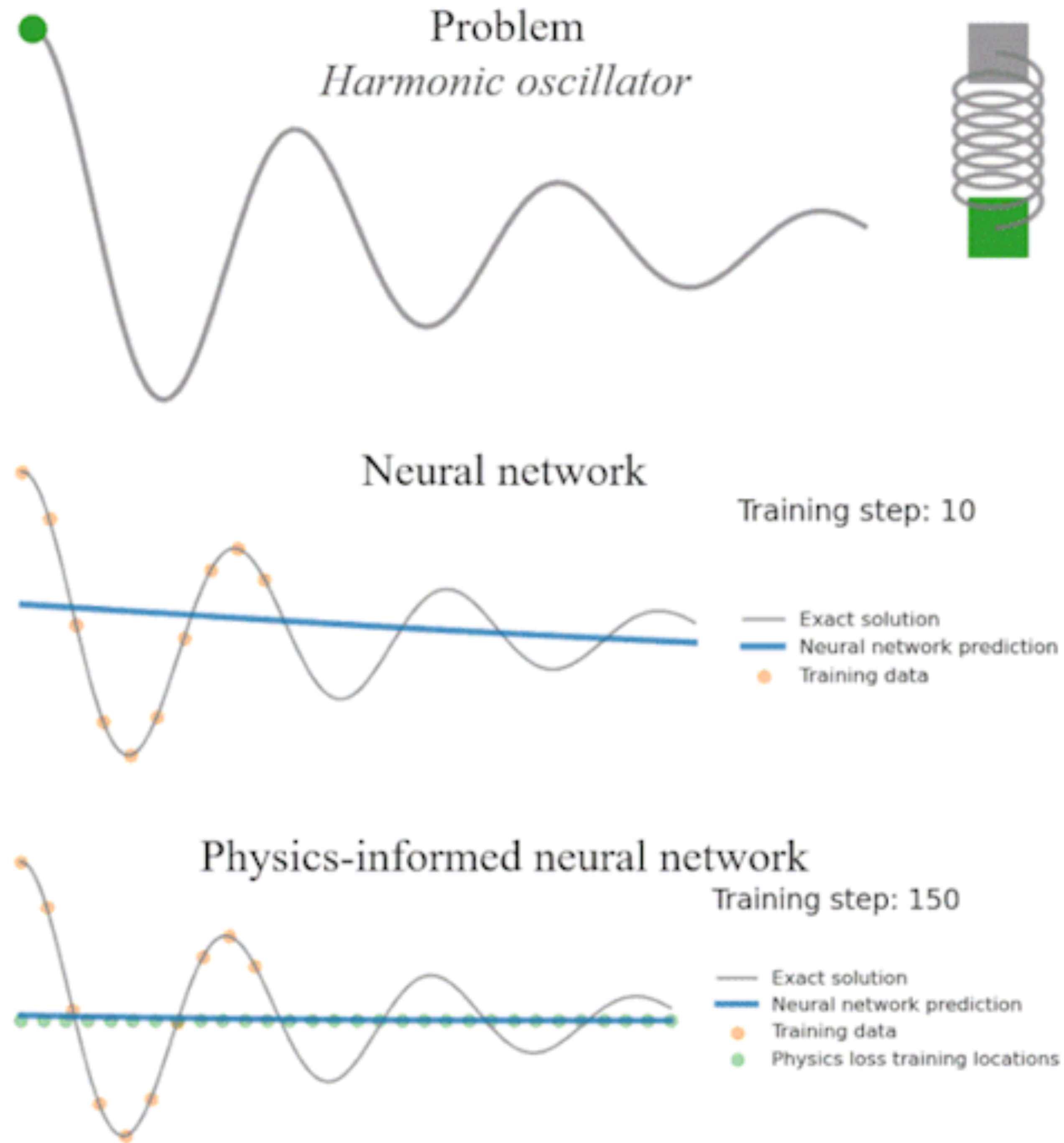
- Regularization: Physics loss acts as a regularizer that restricts the hypothesis space to physically plausible solutions. Without it, the network can fit noise or learn spurious correlations.



Simulation credit: https://github.com/benmoseley/harmonic-oscillator-pinn

Why does physics help?

- Regularization: Physics loss acts as a regularizer that restricts the hypothesis space to physically plausible solutions. Without it, the network can fit noise or learn spurious correlations.



Simulation credit: https://github.com/benmoseley/harmonic-oscillator-pinn

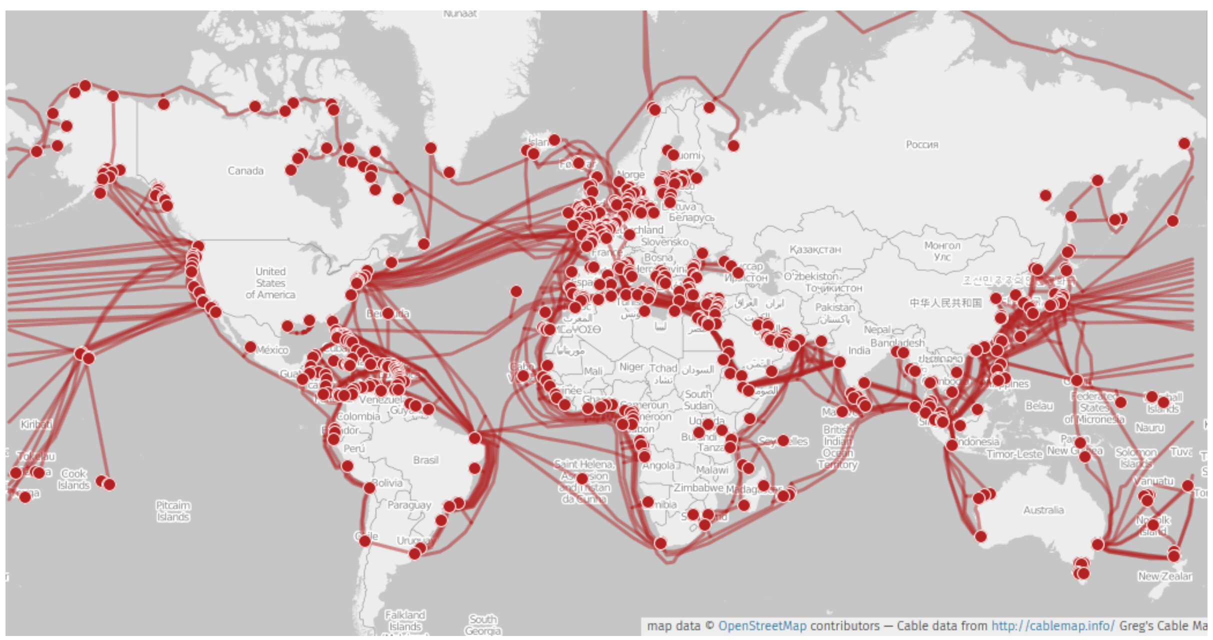
Part - I

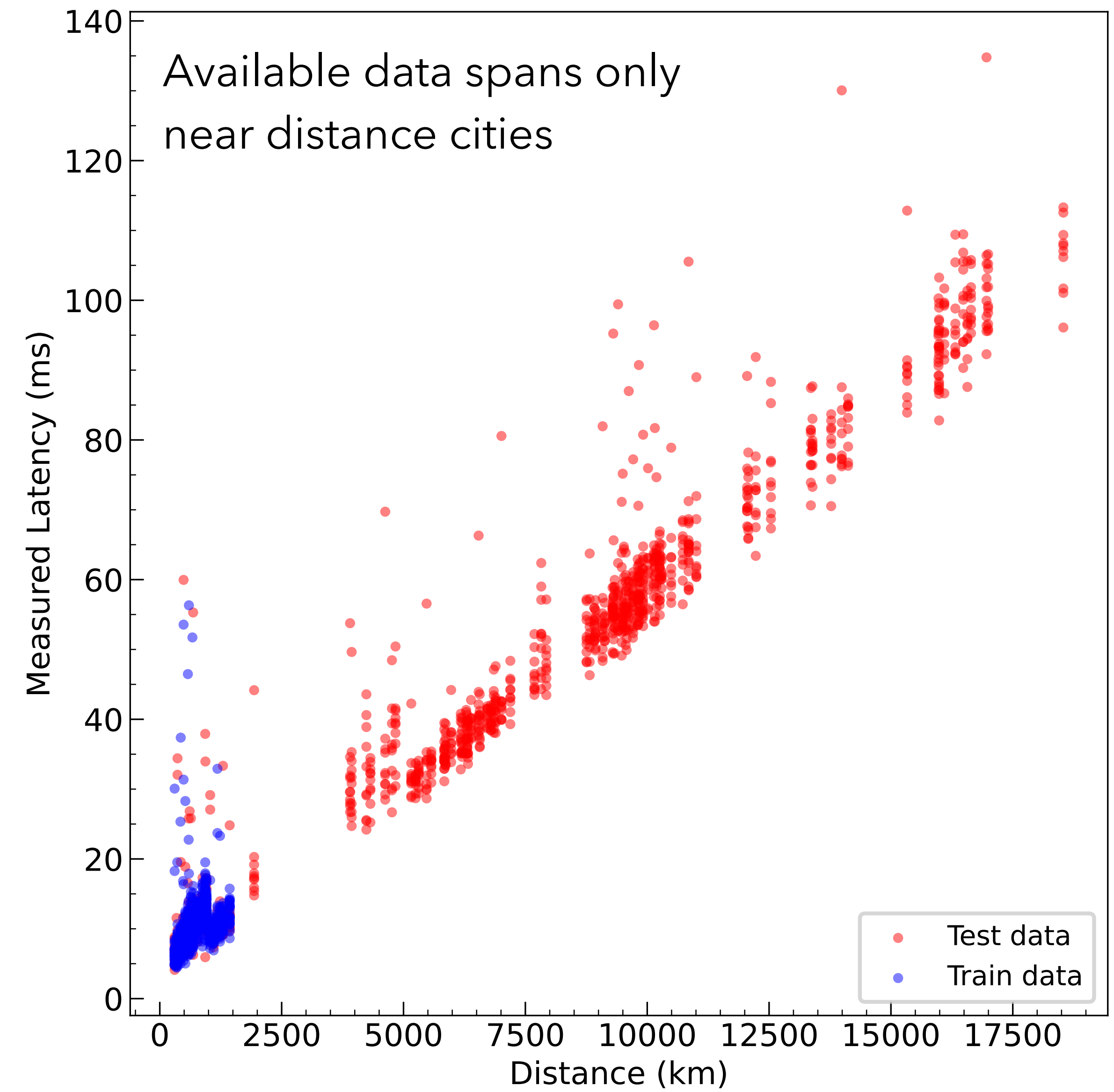
Physics informed latency prediction

Contrasting a learned model with a pure physics baseline

Physics informed latency prediction

- Case: We've been given data of time taken for internet signals to reach from point A to point B. We know the geographical distance between the two cities but not the length of the fibre. We are supposed to flag anomalous network communication.





Physics informed latency prediction

- Case: We've been given data of time taken for internet signals to reach from point A to point B. We know the geographical distance between the two cities but not the length of the fibre. We are supposed to flag anomalous network communication.

Physics model (no ML)

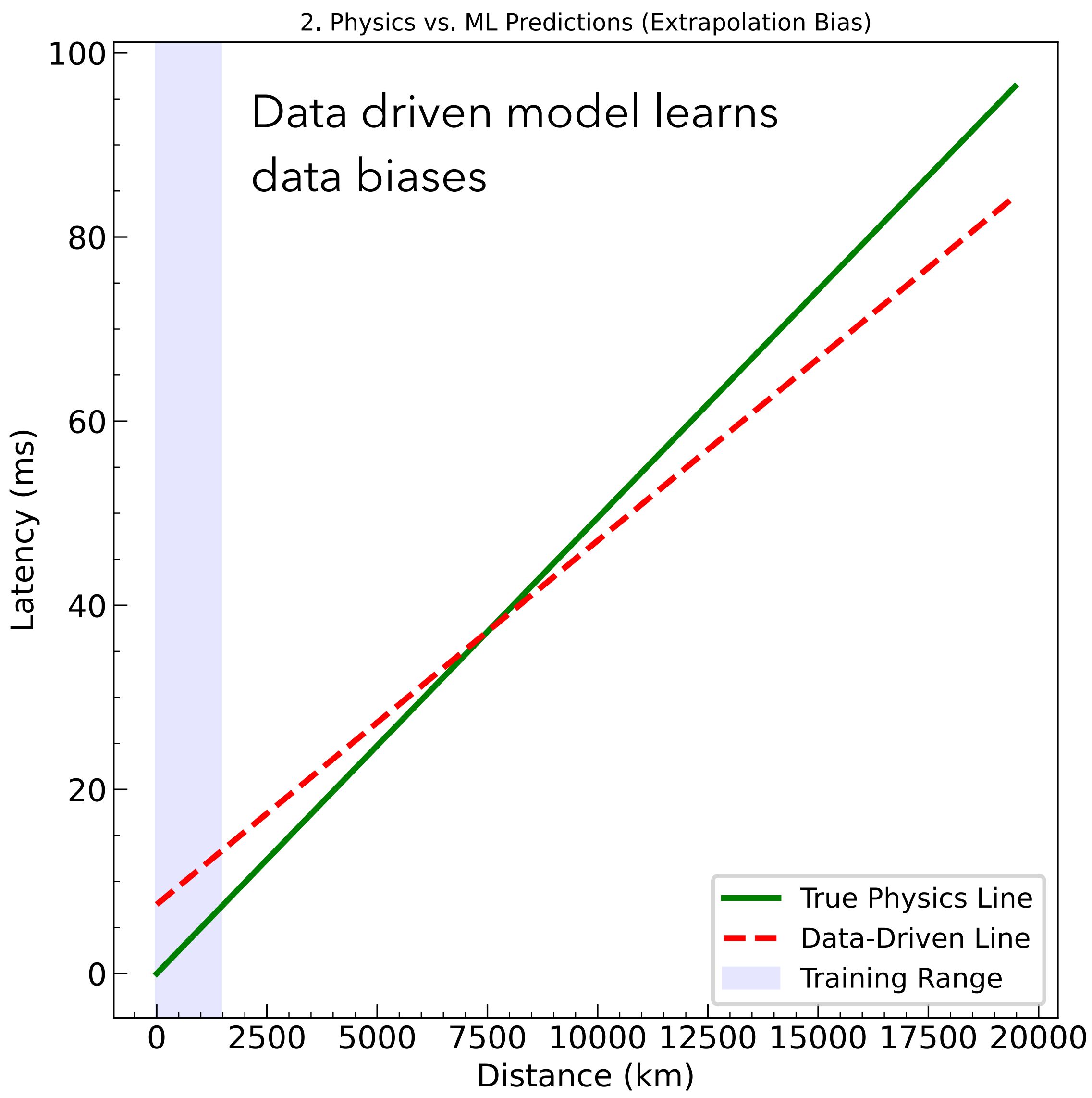
$$\frac{dx}{dt} = v \Rightarrow \text{latency}_{\text{pred}} = \frac{x_{\text{city}_1} - x_{\text{city}_2}}{\text{speed of light in fibre}}$$

If $\text{latency}_{\text{pred}} < \text{latency}_{\text{measured}}$ flag anomaly

Find quantitive estimate of this inequality via threshold tuning of precision and recall

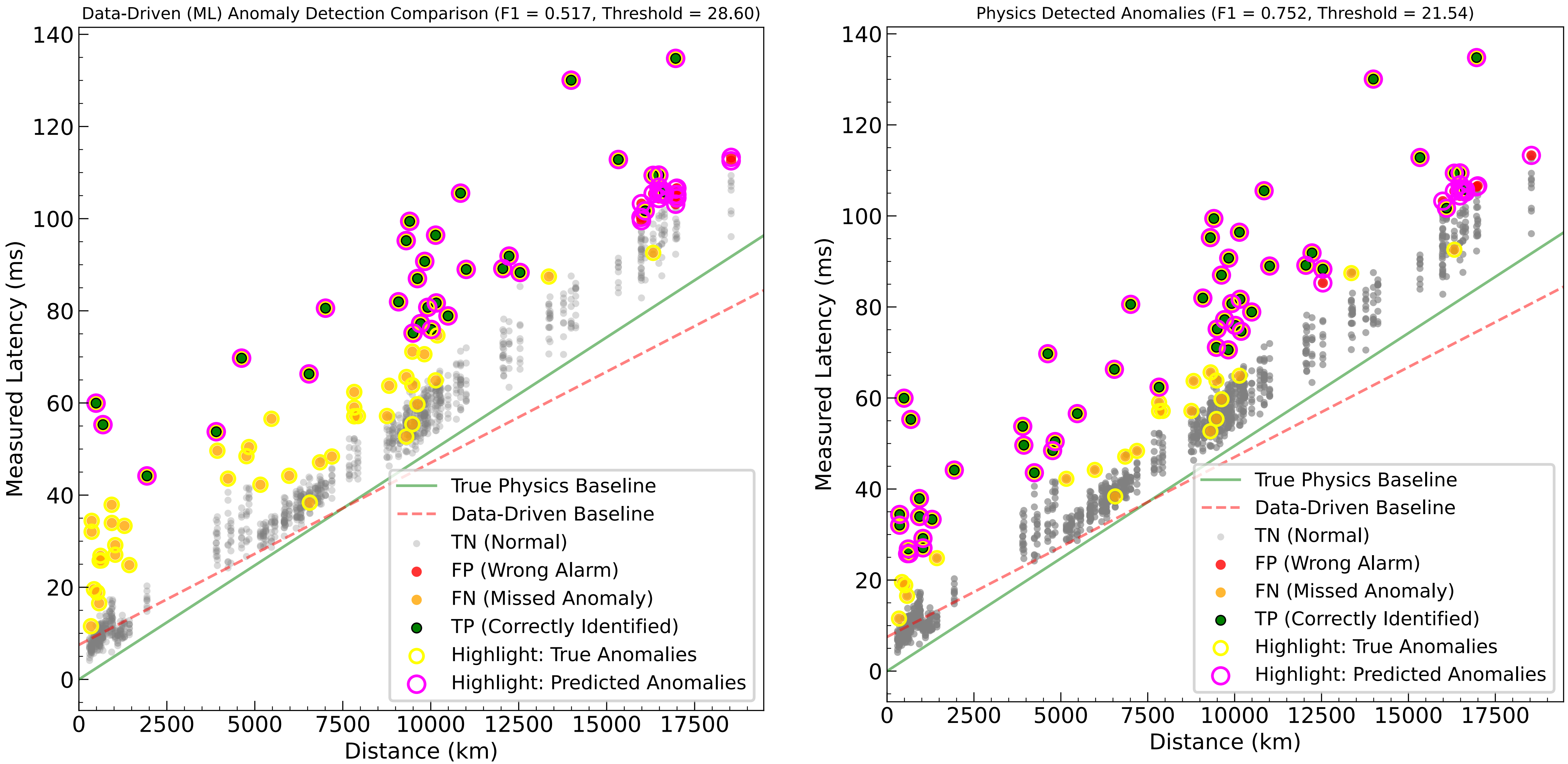
ML model

Fit linear regression, find optimal anomaly detection threshold by tuning of precision and recall



Physics informed latency prediction

- When generalised, physics driven model captures anomalies better



- Learn more at: https://physics-informed-latency-pred.streamlit.app/

Part - II

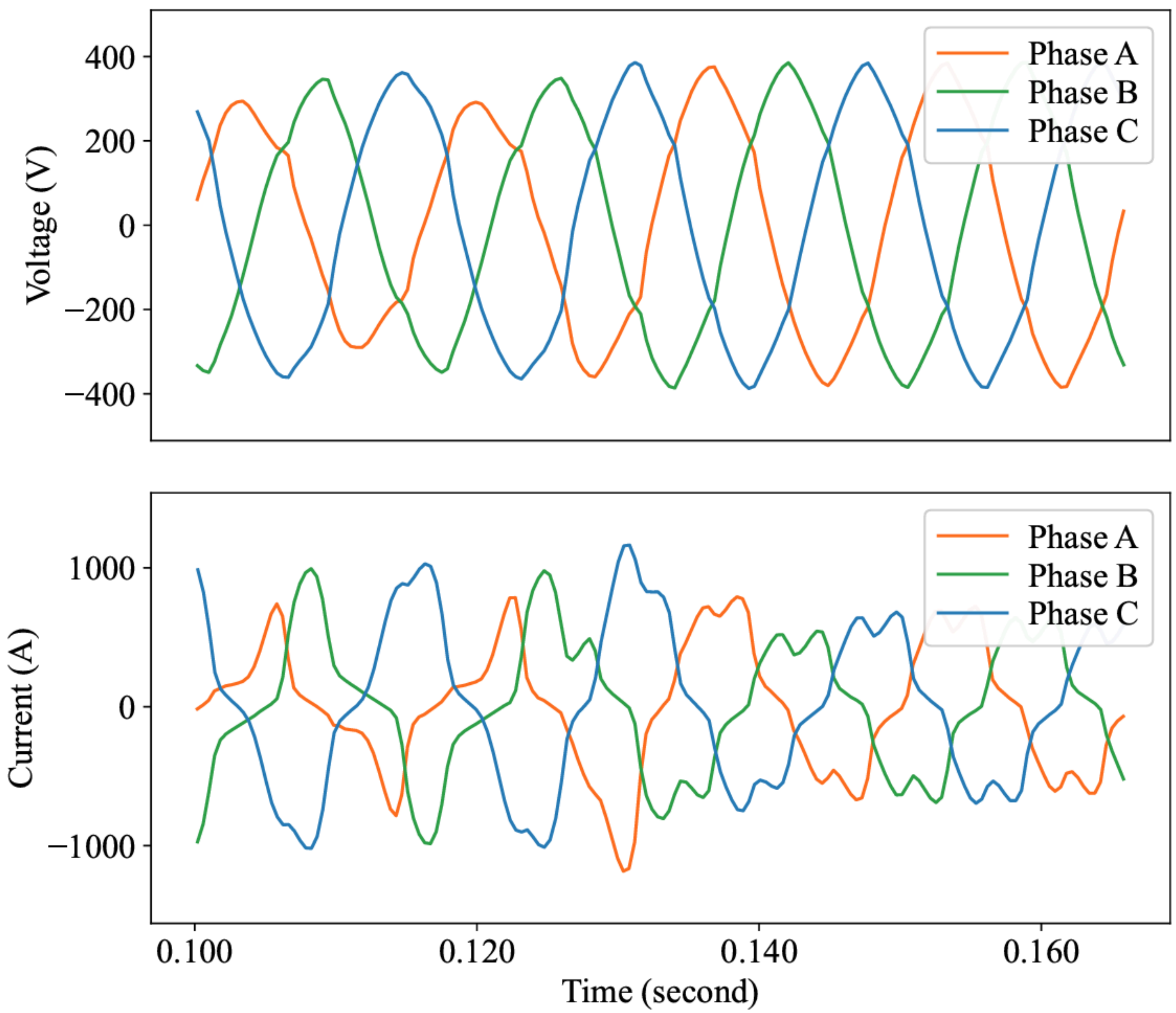
Anomaly detection in time series data

LSTM behaviour under simple harmonic dynamics

Anomaly detection in simple harmonic oscillator

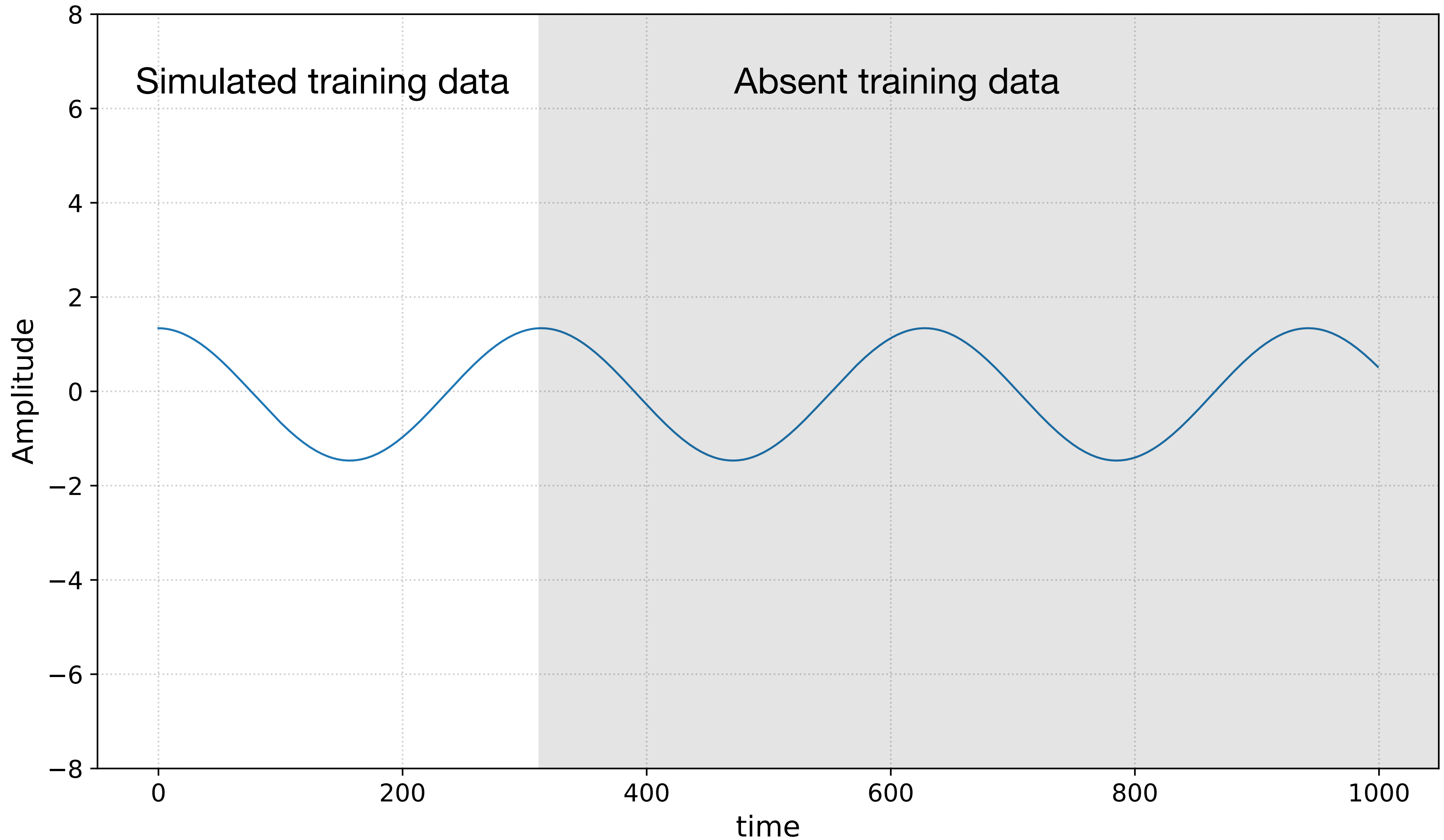
- Case: Many systems have oscillatory output
- Example: voltage and currents in power-grid

arXiv:2504.06588



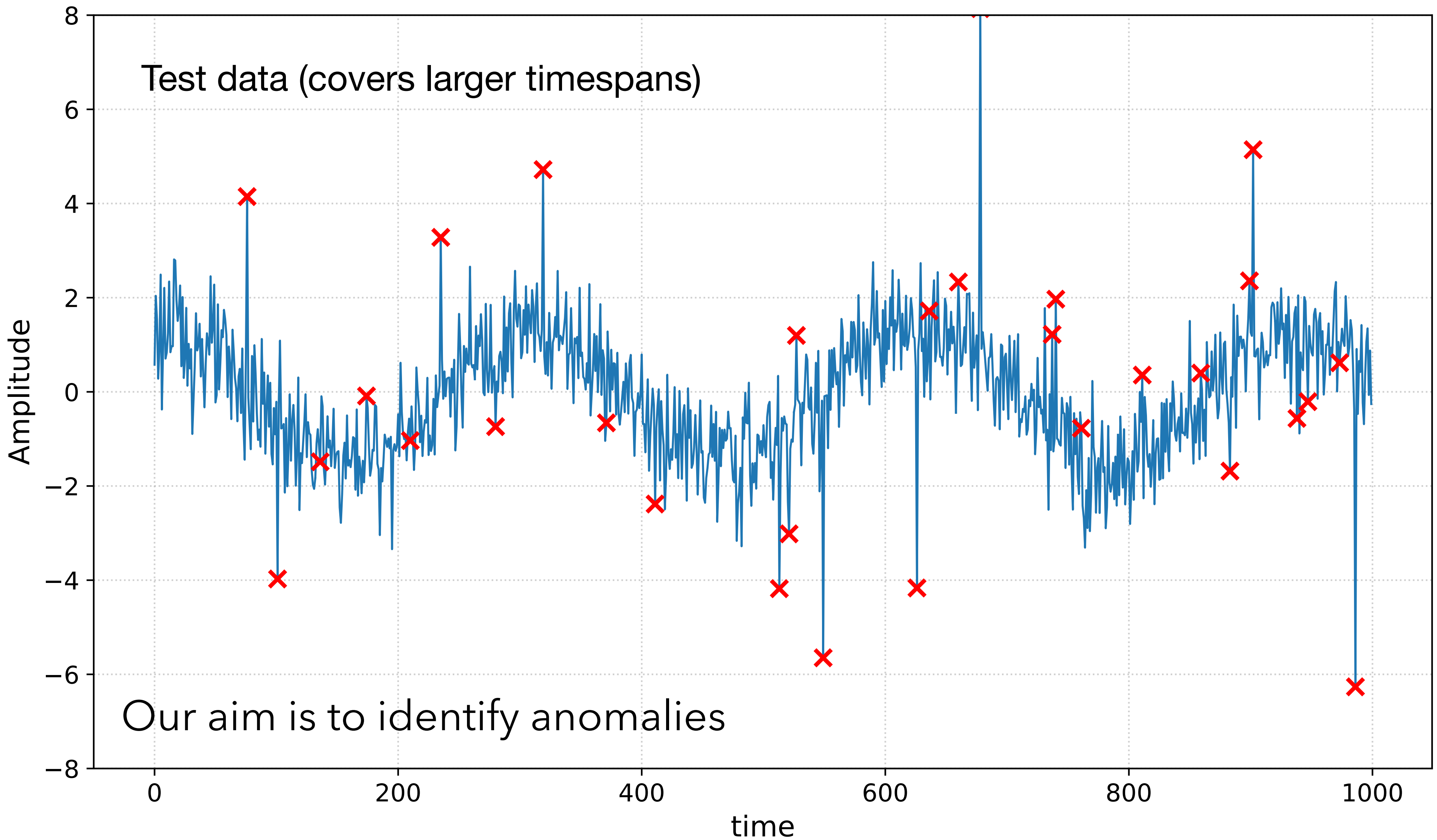
- Real-world synchrophasor dataset i.e. high resolution time synchronised power system measurements
- Three-phase voltage and current injection waveform measurement
- Shows phase imbalance, harmonic distortion, non-unitary power factor, and a transient event
- Also presents an important example of oscillatory systems

Anomaly detection in simple harmonic oscillator



- Given signal contains a fixed frequency and a number of anomalies of variable strength
    - Real life situation will contain noise, drift, unknown frequency etc.
    - These problems can be mitigated using additional processing, e.g., smoothing within windows, making frequency a learnable parameter or using fast-Fourier transform to estimate frequency etc. These will not be discussed here.

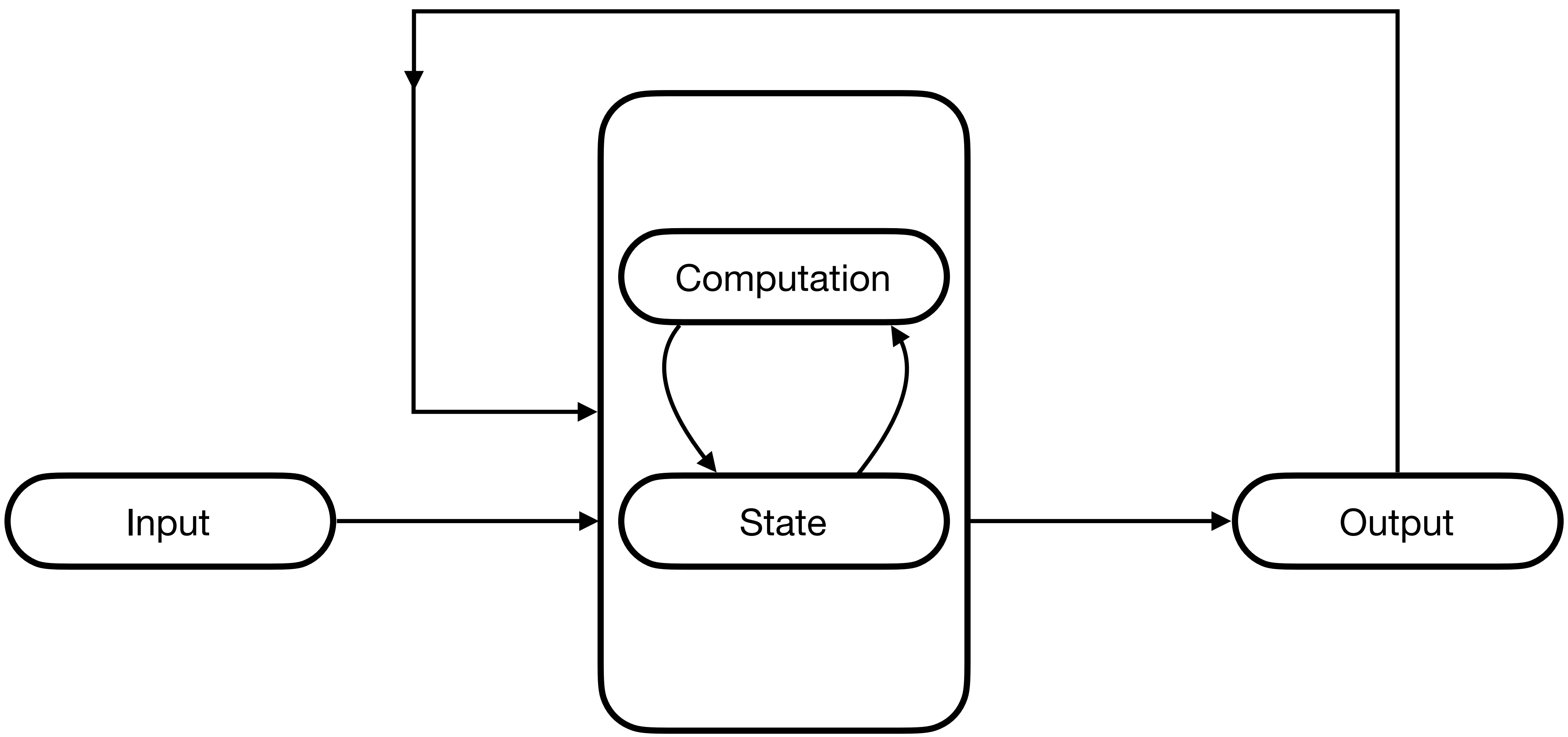
Anomaly detection in simple harmonic oscillator



- Given signal contains a fixed frequency and a number of anomalies of variable strength
  - Real life situation will contain noise, drift, unknown frequency etc.
  - These problems can be mitigated using additional processing, e.g., smoothing within windows, making frequency a learnable parameter or using fast-Fourier transform to estimate frequency etc. These will not be discussed here.

Anomaly detection in simple harmonic oscillator

- Our aim is to identify anomalies i.e. how many times does the data deviate from the ideal trajectory
  - Use LSTM: type of recurrent neural network that allows long term memory relevant for time series data



Loss functions

- Loss functions get extended by physics informed loss

Neutral network

LSTM

$$\mathcal{L}_{obj} = \mathcal{L}_{\mathrm{MSE}}$$

+ physics

$$\mathcal{L}_{phy} = \frac{d^2 y_{\mathrm{recon}}}{dt^2} + \omega^2 y_{\mathrm{recon}}$$

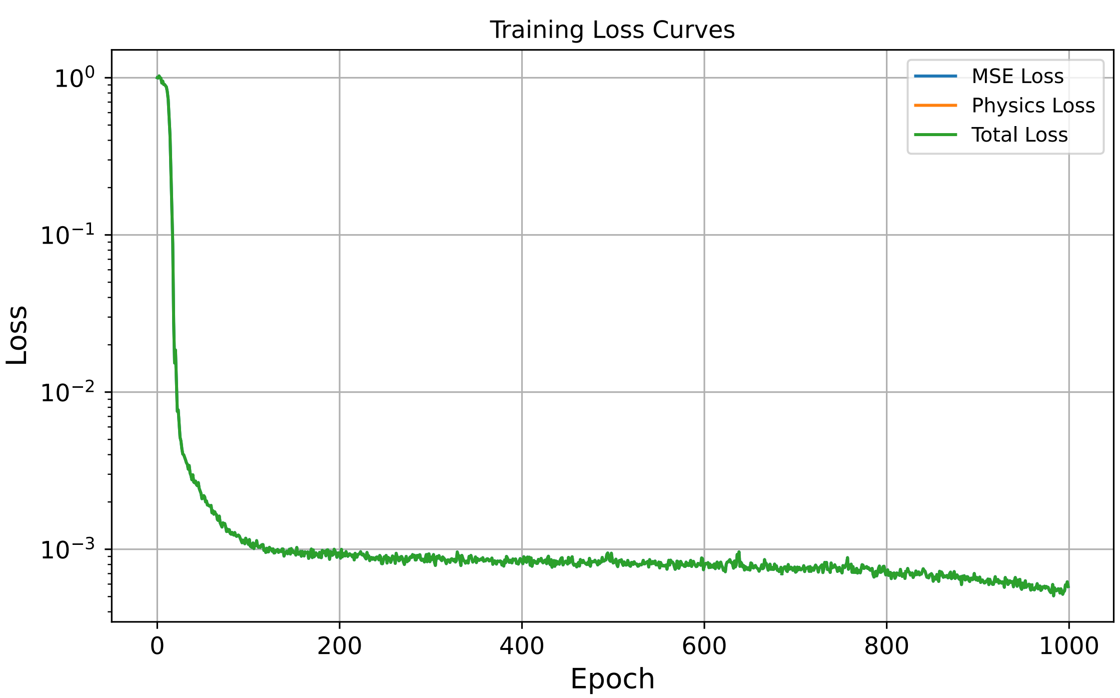
$\Rightarrow$

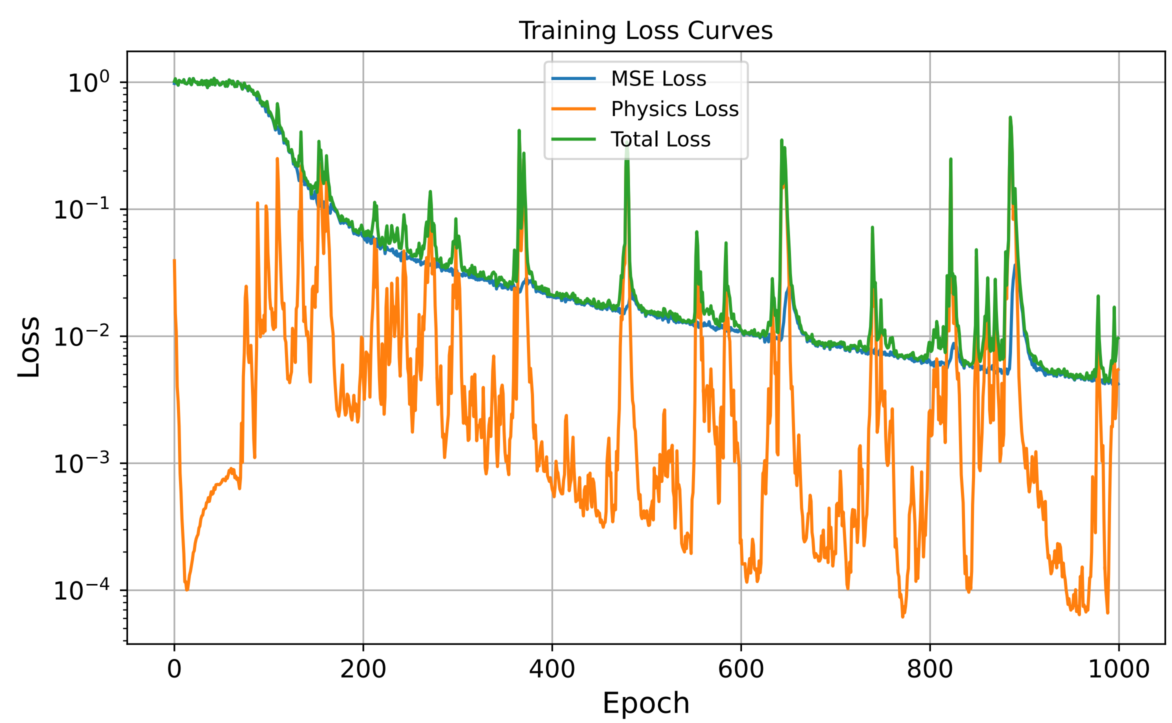
$$\mathcal{L}_{total} = \mathcal{L}_{obj} + \beta \times \mathcal{L}_{physics}$$

β: additional hyper-parameter, needs tuning
(for today, manual tuning)

Large β: physics loss dominates, trivial solutions

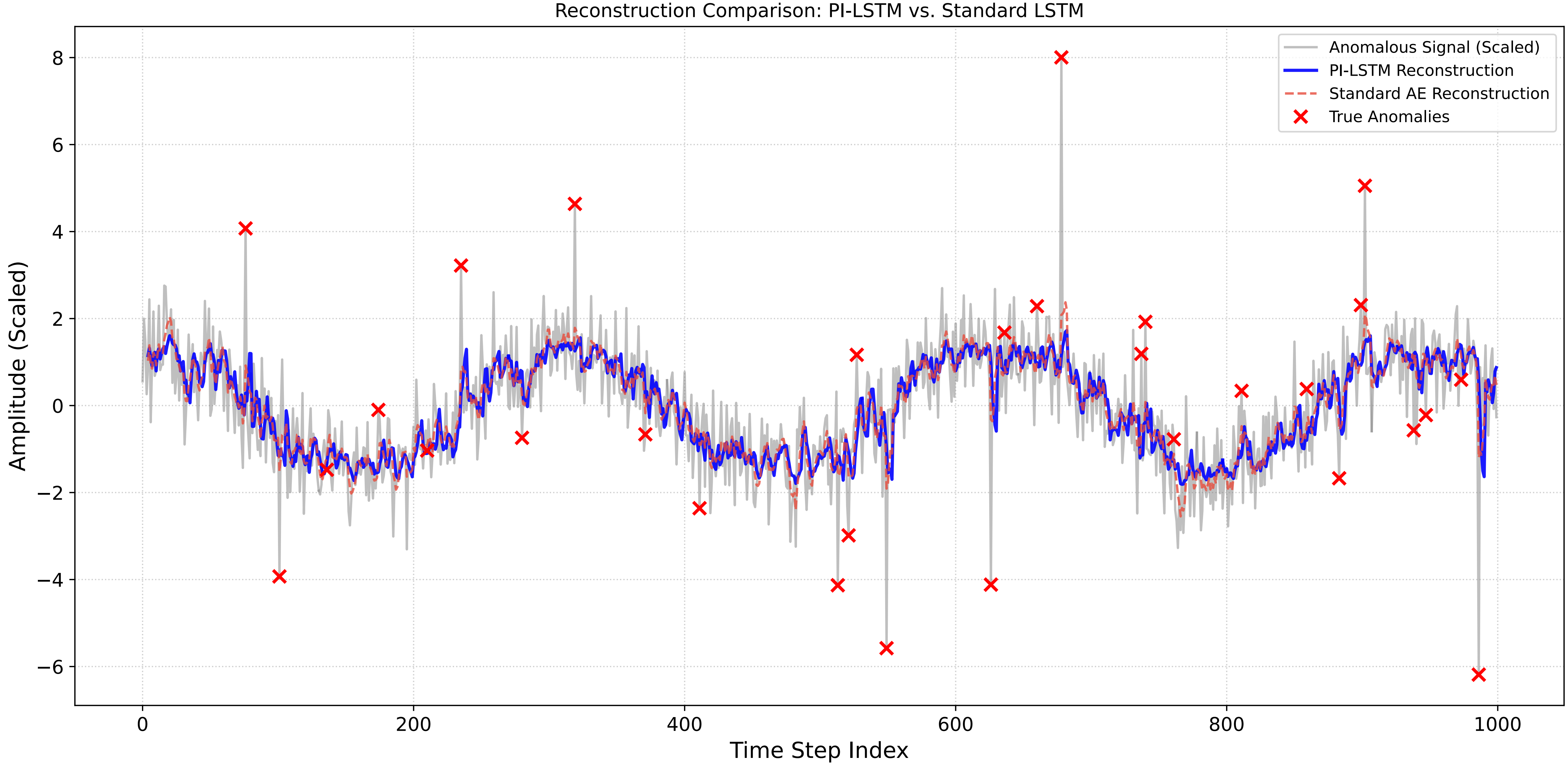
Small β too little physics importance





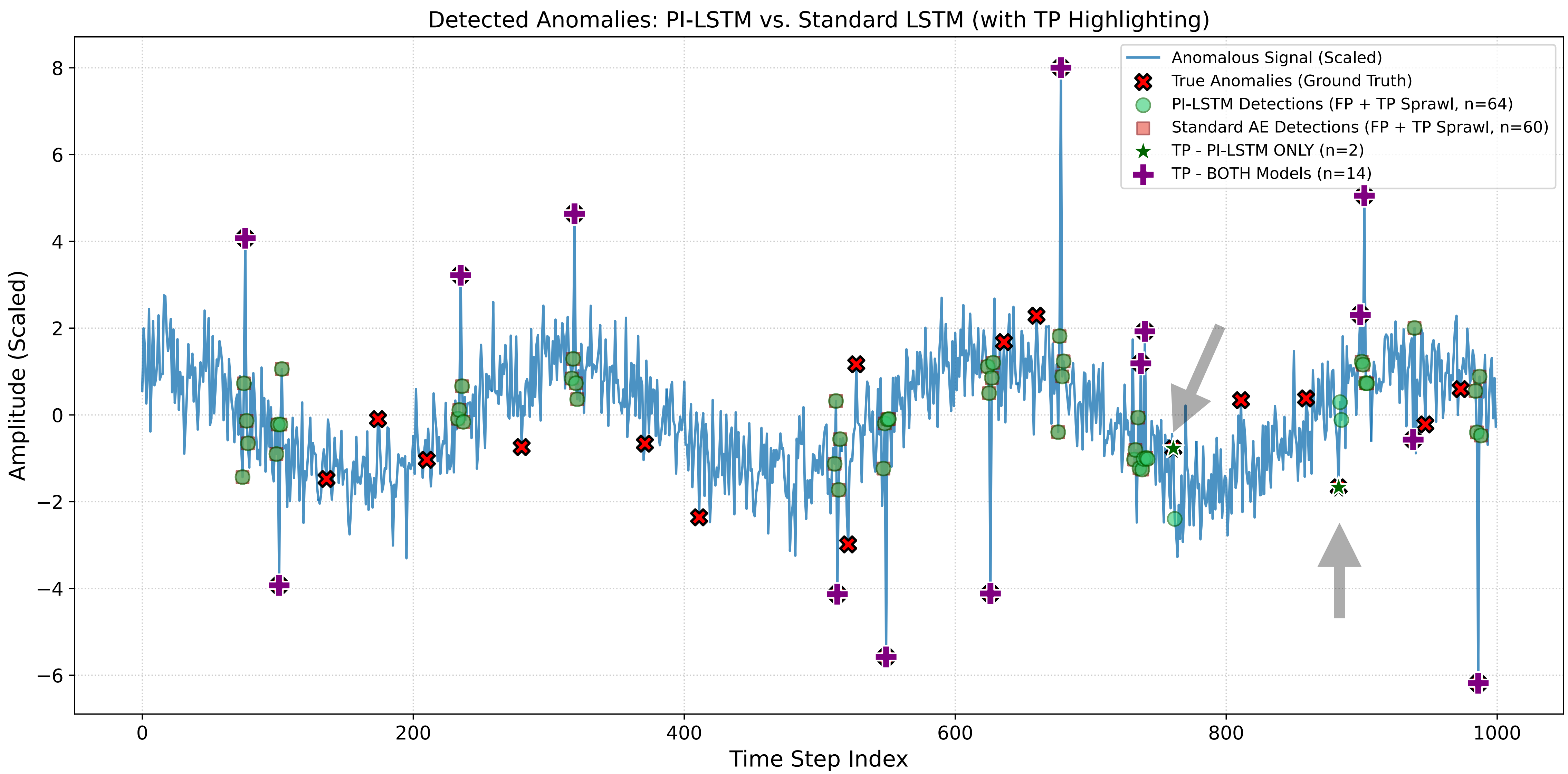
Anomaly detection in simple harmonic oscillator

- Trained on clean simple harmonic oscillator signal, tested on noisy signal containing anomalies



- Standard LSTM introduces distortions in order to recreate test signal

Anomaly detection in simple harmonic oscillator



Model	Unique Anomalies Detected (TP)	Total Windows Flagged	F1-Threshold
PINN (β = 0.01)	16	64	1.41
Standard AE (β = 0.00)	14	60	1.23

14% more anomalies and 15% higher detection threshold with PINN

- Learn more at: https://pinnlearning.streamlit.app/

Part - III

Generative modelling with structure

VAEs that learn to produce physically valid signals

Signal generation capabilities

- Case: Can we combine generative AI with physics information to create more accurate and controlled simulation?
- Aim: Demonstrate with a simple harmonic oscillator
- Use VAE: provides an interpretable latent space, that allows generative capabilities and physics integration

Neutral network

vAE

$$\mathscr{L}_{obj} = \mathscr{L}_{\mathrm{MSE}} + \mathscr{L}_{\mathrm{KL}}$$

vAE architecture

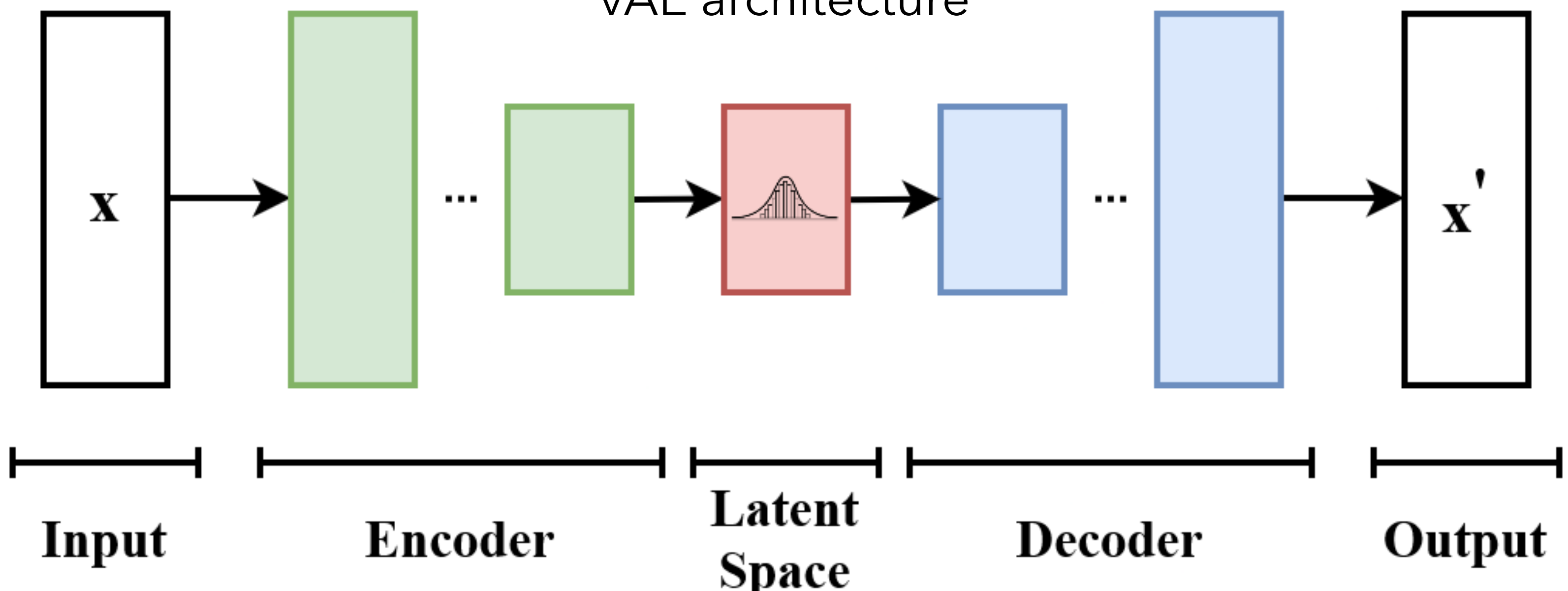


Image source: wiki

β: additional hyper-parameter, needs tuning

Large β: physics loss dominates, trivial solutions

Small β too little physics importance

Signal generation capabilities

- Case: Can we combine generative AI with physics information to create more accurate and controlled simulation?
- Aim: Demonstrate with a simple harmonic oscillator
- Use VAE: provides an interpretable latent space, that allows generative capabilities and physics integration

Neutral network

vAE

$$\mathscr{L}_{obj} = \mathscr{L}_{\mathrm{MSE}} + \mathscr{L}_{\mathrm{KL}}$$

vAE architecture

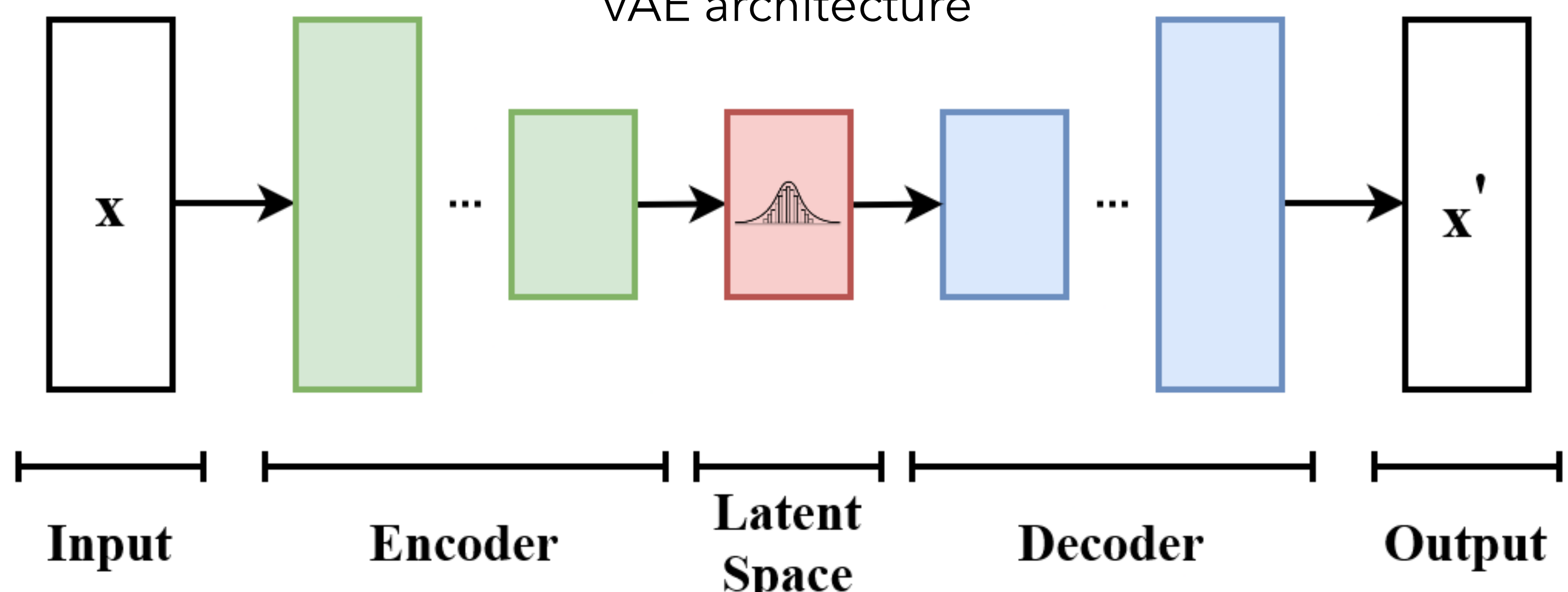


Image source: wiki

$$\Rightarrow \qquad \mathscr{L}_{total} = \mathscr{L}_{obj} + \beta \times \mathscr{L}_{physics}$$

β: additional hyper-parameter, needs tuning

Large β: physics loss dominates, trivial solutions

Small β too little physics importance

Signal generation capabilities

- Case: Can we combine generative AI with physics information to create more accurate and controlled simulation?
- Aim: Demonstrate with a simple harmonic oscillator
- Use VAE: provides an interpretable latent space, that allows generative capabilities and physics integration

Neutral network

vAE

$$\mathscr{L}_{obj} = \mathscr{L}_{\mathrm{MSE}} + \mathscr{L}_{\mathrm{KL}}$$

+ physics

$$\mathscr{L}_{phy} = \frac{d^2 y_{\mathrm{recon}}}{dt^2} + \omega^2 y_{\mathrm{recon}}$$

vAE architecture

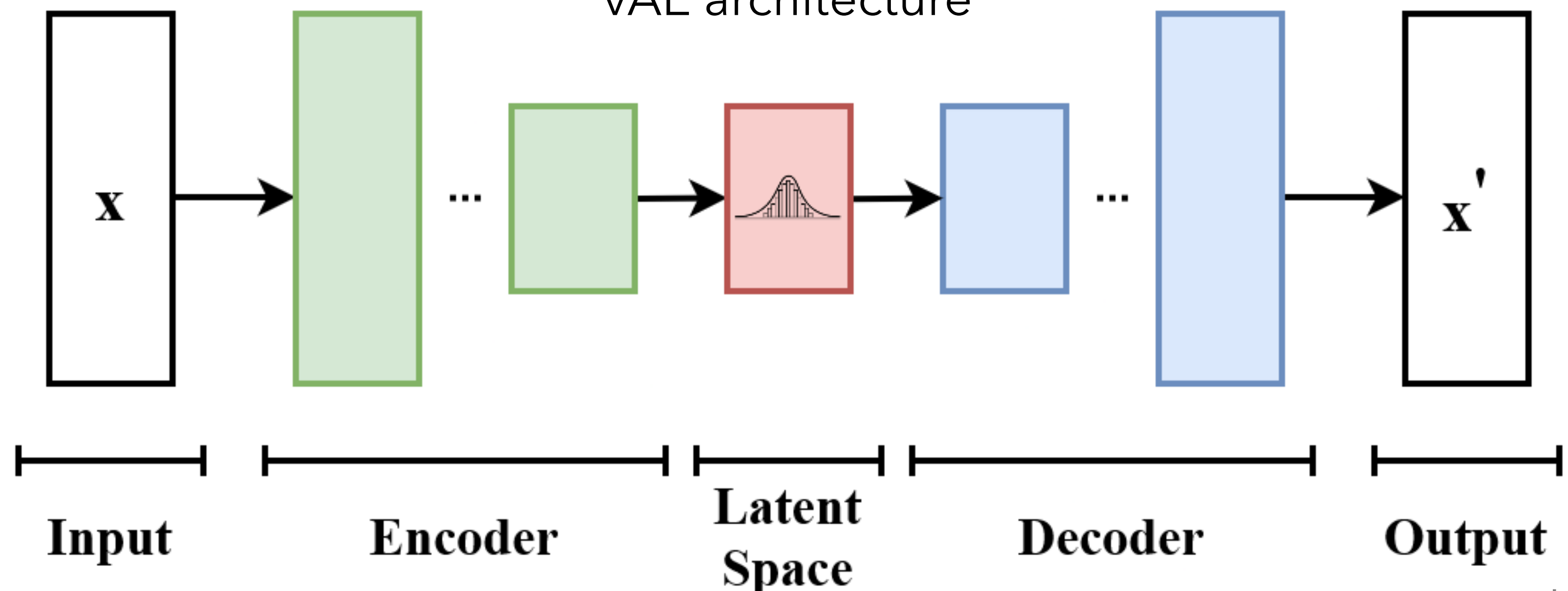


Image source: wiki

$\Rightarrow$ $$\mathscr{L}_{total} = \mathscr{L}_{obj} + \beta \times \mathscr{L}_{physics}$$

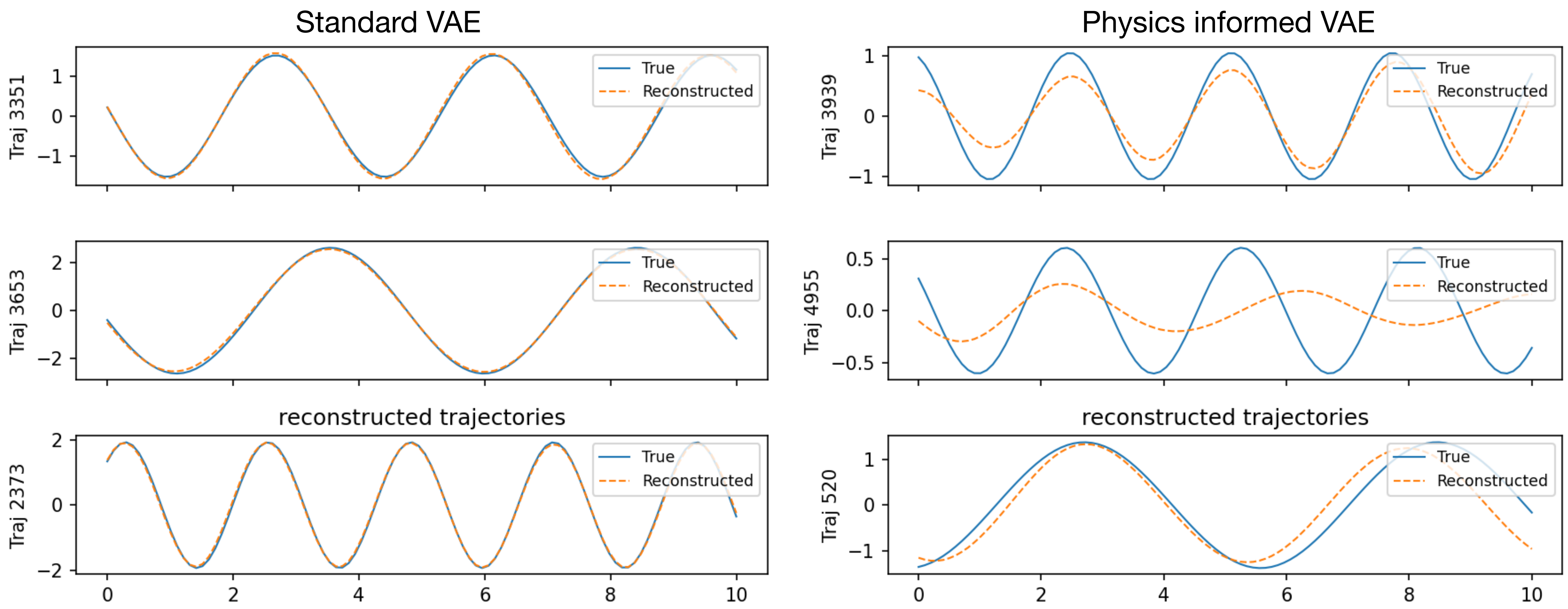
β: additional hyper-parameter, needs tuning

Large β: physics loss dominates, trivial solutions

Small β too little physics importance

Training a physics informed VAE

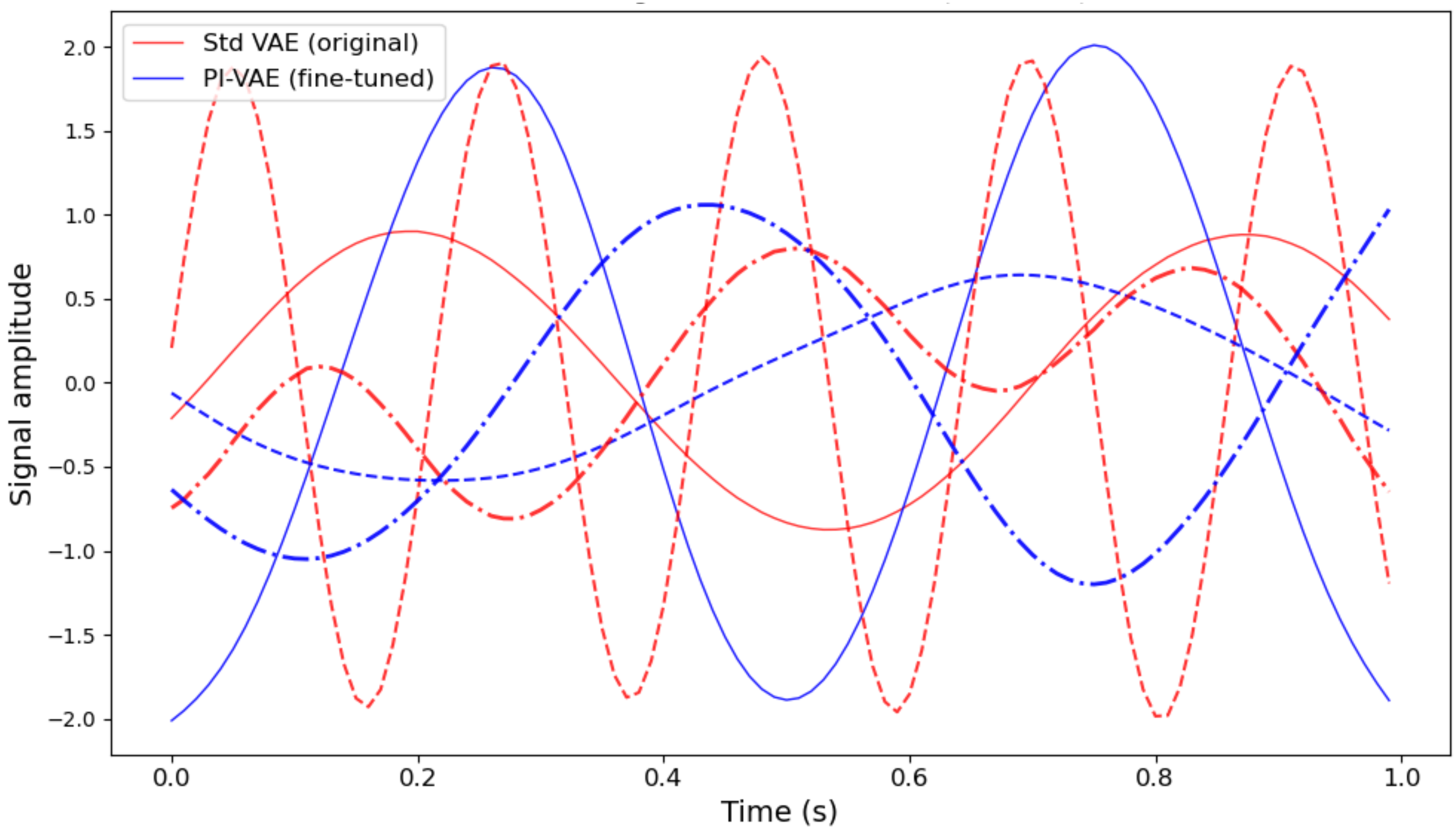
- Train on a clean simple harmonic oscillator signal of variable frequency, amplitude and phase
- Aim: use trained model to generate new simple harmonic oscillator signals



- The physics informed VAE appears to have worse reconstruction performance
- Sacrifice on reconstruction during training for physics consistency, can be improved by fine-tuning hyper-parameter β, here no fine-tuned β : chosen value 'good enough' not ' the best'

Signal generation

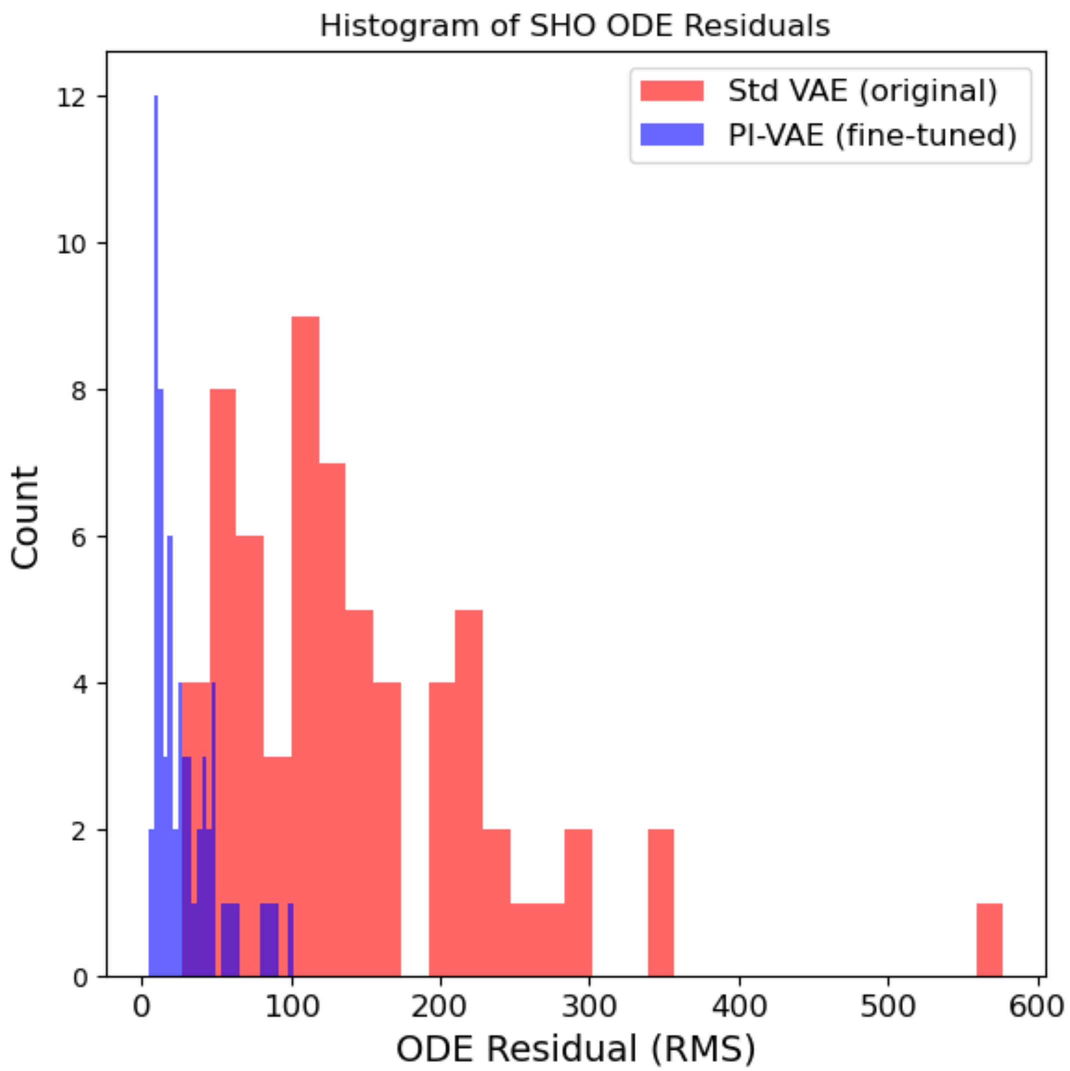
- Small fine-tuning (minimise physics loss only) of PI-VAE post training as β was not optimised
- Procedure: Sample $z \sim \mathcal{N}(0, I)$ $\rightarrow$ Signal $x =$ Decoder(z)
- Standard VAE doesn't do very good with signal generation

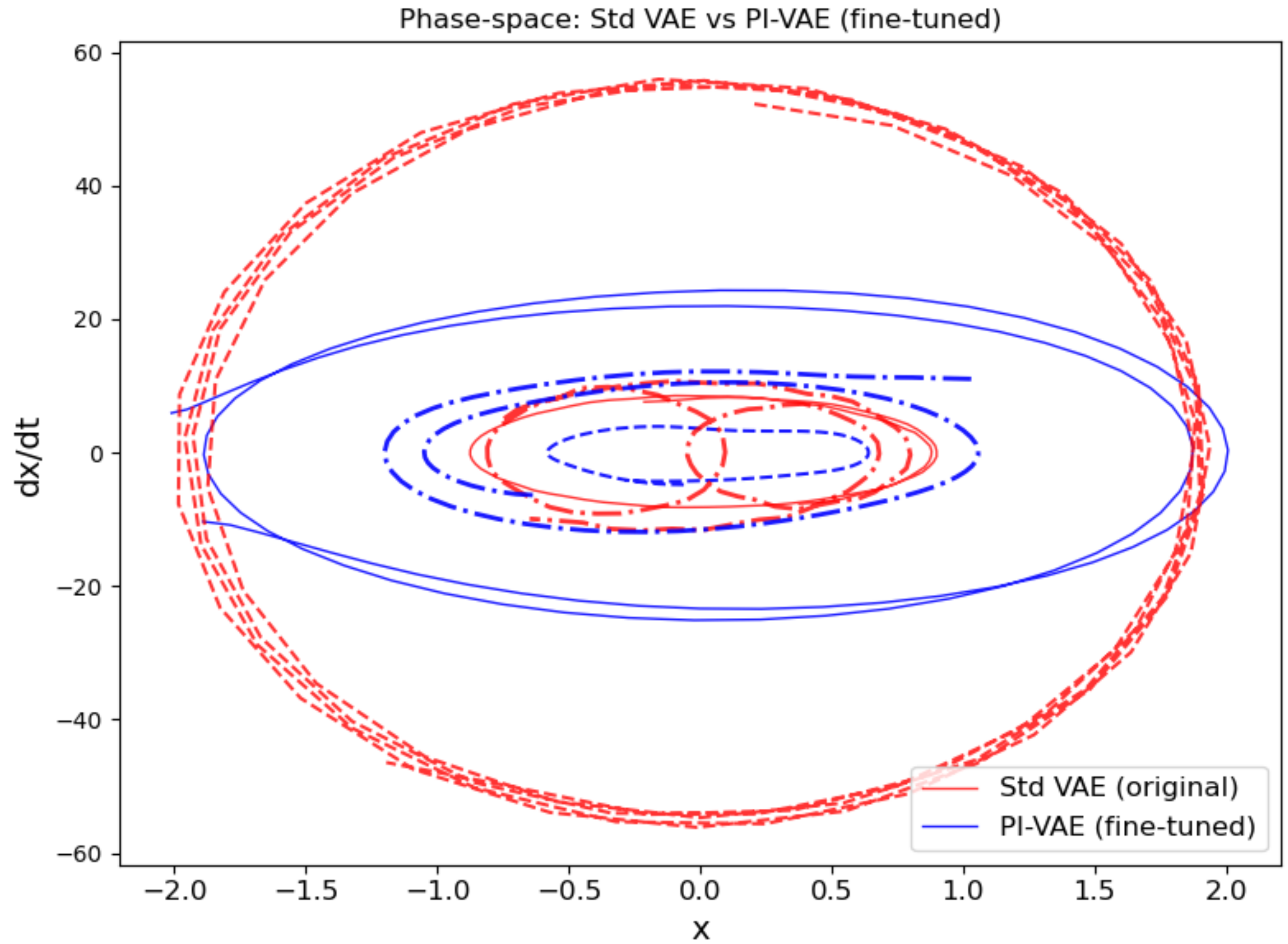


- PI-VAE has a more stable performance in signal generation

Signal generation

- Physics informed signal generation adheres to physically meaningful trajectories
  - Average ODE residual (standard VAE): 145.47
  - Average ODE residual (PI-VAE): 61.41
  - Average ODE residual after fine-tuning (PI-VAE): 30.08





Conclusions

- Physics constraints = better extrapolation and data efficiency when equations are known
- Trade-off: beta tuning matters, wrong physics hurts more than helps
- PIML works best for: sparse data, known governing equations, need for reliability

https://qr.link/4khP9z

Resources and slids

- Happy to discuss and connect