Physics-informed Discovery of State Variables in Second-Order and Hamiltonian Systems

Félix Chavelli[1] Khoo Zi-Yu[1] **Dawen Wu**[2,1] Jonathan Sze Choong Low[3]
Stéphane Bressan[1]

[1]National University of Singapore (NUS), Singapore
[2]CNRS@CREATE LTD, Singapore
[3]Agency for Science, Technology and Research (A*STAR), Singapore

April 23, 2025

Outline

Motivation: Modeling Dynamical Systems

- Dynamical systems are ubiquitous (physics, biology, engineering, etc.).
- Understanding, predicting, and controlling them is crucial.
- **Challenge:** Data-driven methods often:
  - Assume relevant state variables are known.
  - Result in overparameterized or non-interpretable state spaces.
- **Goal:** Discover a minimal, non-redundant, and interpretable set of state variables directly from observational data (e.g., videos).

Baseline Approach [1]

- Proposed a nested autoencoder (AE) architecture to identify state variables from videos.
- **Step 1:** Outer AE learns compact representation & predicts future frames.
- **Step 2:** Inner AE compresses this representation further into potential state variables.
- Relies on an external Intrinsic Dimension (ID) estimator to determine the number of state variables (size of inner AE bottleneck).

Limitations of Baseline

- ID estimator is physics-agnostic, yields non-integer values (requires rounding $\Longrightarrow$ potential bias).
- Discovered variables can be correlated, redundant, or entangled.
- No guarantee of physical interpretability.

[1](Chen, B., Huang, K., Raghupathi, S., Chandratreya, I., Du, Q., Lipson, H.: Automated discovery of fundamental variables hidden in experimental data. Nature Computational Science 2, 433–442 (07 2022))

Our Approach: Physics-Informed Discovery

- Leverage physical principles of the systems being modeled.
- Focus on **Second-Order** and **Hamiltonian** systems.
- Incorporate physics knowledge using Physics-Informed Machine Learning (PIML) biases [Karniadakis et al., 2021]:
  - **Observational Bias:** Constraints reflected in data or variable structure (e.g., position/momentum pairs).
  - **Learning Bias:** Constraints via loss functions or training procedure (e.g., VAE's KL divergence for sparsity/independence).
  - **Inductive Bias:** Constraints embedded in the model architecture (e.g., Hamiltonian Neural Network).

Proposed Models: Enhancing the Baseline

We modify the **inner autoencoder** of the baseline model:

1. Physics-Informed AE (PI-AE)

- Adds Observational Bias.
- Enforces 2^{nd}-order structure: Latent variables represent $(\text{position}, \text{momentum})$ pairs.
- Still uses external ID estimator.

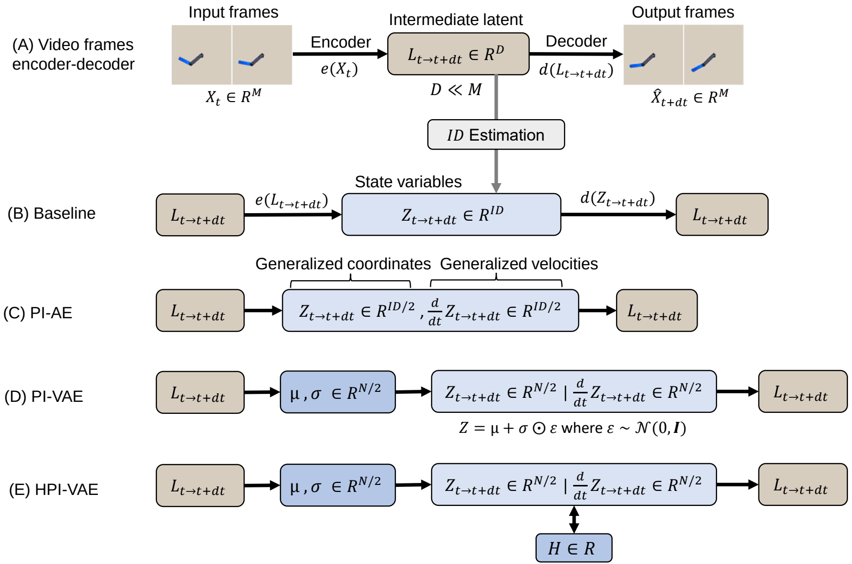
2. Physics-Informed VAE (PI-VAE)

- Builds on PI-AE.
- Adds Learning Bias (Variational Autoencoder).
- KL divergence promotes sparsity

3. Hamiltonian PI-VAE (HPI-VAE)

- Builds on PI-VAE.
- Adds Inductive Bias.
- Incorporates a Hamiltonian Neural Network (HNN) layer acting on latent variables.
- Loss includes term for Hamilton's equations $(\dot{\tau} = \partial H/\partial \rho, \dot{\rho} = -\partial H/\partial \tau)$.
- Aims for canonical, energy-conserving variables.
- Also bypasses external ID estimator.

Model Architectures Overview



Experimental Setup

- **Data:** Video frames (real & simulated) from Chen et al. (2022).
  - Systems: Reaction-Diffusion, Single Pendulum, Double Pendulum, Swing Stick, Elastic Pendulum.
  - 100k frames per system (100 frames x 1000 trajectories).
  - 80
- **Models:**
  - Outer AE: CNN-based, latent dim = 64 (fixed).
  - Inner AE: Fully Connected layers.
  - Baseline/PI-AE bottleneck = ID (from estimator).
  - PI-VAE/HPI-VAE bottleneck = 10 (fixed, larger than any true DoF).
- **Evaluation:**
  - Estimated Degrees of Freedom (DoF).
  - Interpretability of latent variables (comparison with physical variables).
  - Reconstruction loss (check: models must still be predictive).

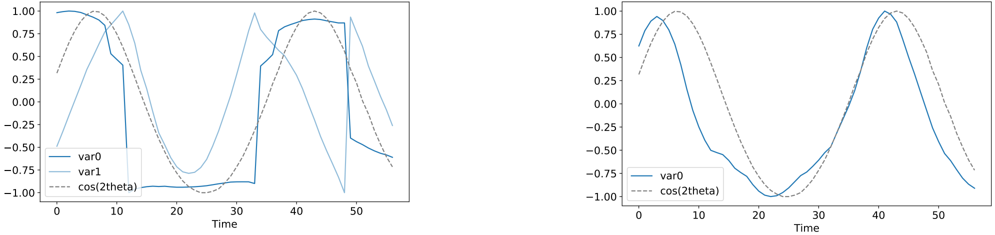
Results: Degrees of Freedom Estimation

Table: Estimated Latent Space Dimension (Degrees of Freedom)

System	Ground Truth	Baseline (Chen et.al 2022)	PI-AE	PI-VAE	HPI-VAE
Reaction-diffusion	2	$2.16 \approx 2$	$2.16 \approx 2$	2	2
Single pendulum	2	$2.05 \approx 2$	$2.05 \approx 2$	2	2
Double pendulum	4	$4.71 \approx 4$	$4.71 \approx 4$	4	4
Swingstick system	4	$4.89 \approx 4$	$4.89 \approx 4$	4	4
Elastic pendulum	6	$5.34 \approx 6$	$5.34 \approx 6$	6	6

- Baseline & PI-AE rely on external estimator + rounding (prone to bias).
- PI-VAE & HPI-VAE automatically discover the correct DoF via VAE latent sparsity (using variance threshold > 0.01).

Results: State Variable Interpretability (Single Pendulum)

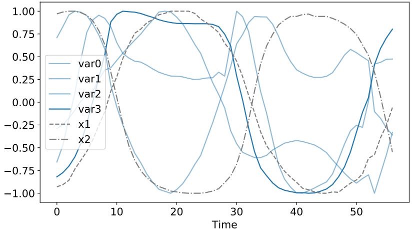


(a) Baseline (b) HPI-VAE

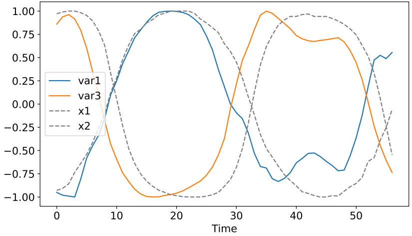
Figure: Latent variables vs. time for one trajectory. Black dashed line: $\cos(2\theta(t))$.

- **Baseline:** Variables appear entangled/redundant, somewhat discontinuous.
- **HPI-VAE:** Learns smoother variables. One variable clearly correlates with $\cos(2\theta)$ (position-related), the other (not shown) is its derivative (momentum-related, due to 2^{nd} order constraint). VAE promotes continuity.

Results: State Variable Interpretability (Double Pendulum)



(a) Baseline

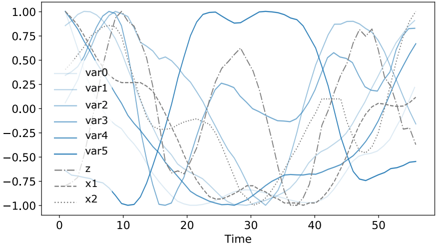


(b) HPI-VAE

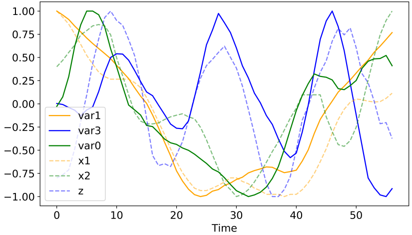
Figure: Latent variables vs. time. Black dashed lines: horizontal positions $x_1(t), x_2(t)$.

- **Baseline:** Variables entangled. Some correlation with x_1, x_2 exists but is mixed.
- **HPI-VAE:** Clearer separation. Two variables (shown) correlate well with x_1, x_2. Other two (not shown) are derivatives. VAE helps disentangle.

Results: State Variable Interpretability (Elastic Pendulum)



(a) Baseline



(b) HPI-VAE

Figure: Latent variables vs. time. Black dashed lines: $x_1(t), x_2(t)$, arm length $z(t)$.

- **Baseline:** Difficult to interpret; only one variable seems weakly correlated to x_1.
- **HPI-VAE:** Captures dynamics better. Variables (3 shown) show correlation with x_1, x_2, z. Hamiltonian structure encourages mapping to canonical coordinates. Match isn't perfect but significantly improved.

Conclusion

- Integrating physics-informed ML (specifically VAEs and Hamiltonian constraints) improves upon baseline data-driven discovery of state variables.
- **Key Advantages:**
  - Automatic DoF Discovery: PI-VAE/HPI-VAE bypass the need for external, physics-agnostic ID estimators and potential rounding biases.
  - Improved Interpretability: Discovered variables are less redundant/entangled, smoother, and correlate better with physically meaningful quantities.
  - Minimal Representation: VAE sparsity naturally finds a minimal set of variables.
  - Faithfully captures system characteristics (e.g., 2^{nd} order dynamics).
- This approach offers a better balance between data-driven flexibility and physical consistency.
- Potential for wide application in science and engineering.

Future Work

- Further improve interpretation of latent variables:
  - Develop methods to automatically map latent variables to known physical coordinates.
  - Explore symbolic regression on the discovered latent space.
- Scale to more complex, higher-dimensional systems.
- Incorporate other physics constraints (e.g., symmetries, other conservation laws).
- Investigate adaptive penalization of errors based on spatial/temporal scales for better variable identification.

Acknowledgements

- Collaborators: Félix Chavelli, Khoo Zi-Yu, Jonathan Sze Choong Low, Stéphane Bressan.
- Institutions: National University of Singapore, CNRS@CREATE, A*STAR.

- Thank you for your attention!

Thank You and Questions?