Appendix:
Discovering State Variables Hidden in Experimental Data

Boyuan Chen[*], Kuang Huang, Sunand Raghupathi,
Ishaan Chandratreya, Qiang Du, Hod Lipson

neural-state-variables.com
Columbia University

[*]To whom correspondence should be addressed; E-mail: bchen@cs.columbia.edu.

Data Collection and Processing

In this section, we describe data collection processes for all datasets. The rigid double pendulum, swing stick, air dancer, fire and lava lamp datasets were collected from real-world physical experiments, while the circular motion, reaction diffusion, single pendulum, and elastic double pendulum datasets were collected from simulations.

Rigid double pendulum.

The physical parameters of our rigid double pendulum system is shown in Fig. 1. The system used in our investigation is a two colored chaotic pendulum from 3D scientific: the first arm is black and the second arm is blue. Using the pivot attachment that came with the pendulum, the pendulum is installed against a brown-beige wall in the laboratory. There are 4 bearings on the pendulum. Three of them are fixed in place and one is left loose to reduce friction. We used an iPhone7 to record videos at 720 p and 240 fps.

Length of the first arm: 20.5 cm	
Length of the second arm: 17.9 cm	
Mass of the first arm: 0.262 kg	
Mass of the second arm: 0.110 kg	
	Double Pendulum
Depth of the arms: 3.8 cm	

Fig. 1: Physical parameters of the rigid double pendulum system

We collected a total of 100 videos, with an approximate length of 15 seconds for each video. We used 80 of these videos for training and validation, and 20 of them for testing. For better video quality, we trimmed each video to 11s in order to avoid the movement at the beginning and

the end of recording caused by humans and small changes in brightness or illumination caused by the camera. Another reason is that the dynamics towards the late part of the recordings are more predictable due to the lack of energy and the loss of momentum. Afterwards, we sub-sampled the video to construct a video dataset with 60 fps to produce sufficient visual difference between subsequent frames in a prediction triplet. To feed the video frames into our visual predictive models, the images are resized to 128×128.

Since we are interested in evaluating the results of prediction from the double pendulum system, we further equalized the background of the pendulum system with a simple color filtering so that our vision algorithms can detect the position and orientation of the pendulum arms with another color filtering during the evaluation process. We performed this additional step only to the double pendulum, for the sake of evaluation alone, while other systems do not involve this extra preprocessing step.

Swing stick.

The physical parameters of the swing sticks can be found in Fig. 2. The system being used is from Geelong Shope, made out of a high-quality base, aluminum sticks, high quality bearings and black rubber feet. We used a GoPro Hero 5 Black camera to capture the motion of the system on a 240 frame per second and 720 pixel setting. The GoPro was mounted on a tripod directly at the height of the swing stick table. For each video sequence, we held the stick to a random position and then applied a force on the arm to cause motions that can last for a longer period of time.

We collected a total of 23 videos of approximately 150 seconds for each video sequence. We used 18 of these videos for training and validation and the remaining 5 for testing. Similar to the rigid double pendulum dataset, we trimmed the videos to obtain high-quality data, resulting in 140 seconds for each video. We further sub-sampled the video to 60 fps and resized the image

3

Mass without batteries: 248 g	
Mass with batteries: 340 g	
Span of the two sticks: 34 cm	
Height of the system: 37 cm	
Width of the base: 8 cm	
Length of the baes: 22.5 cm	
Length of the longer stick: 24 cm	
Length of the shorter stick: 18 cm	
Hinge point for the first arm (from the longer side): 17.5 cm	Swing Stick
Hinge point for the second arm (from longer side): 10.1 cm	

Fig. 2: Physical parameters of the swing stick system

frames to 128×128 before sending the frames to our dynamics predictive model.

Air dancer.

The physics parameters of the air dancer system can be found in Fig. 3. Initially, the air dancer ran on a 9 V capped battery that could directly be placed in the compartment. However, this led to undesirable complications in our study. First, the 9 V voltage was too high for the motor as the rotation of the fan prevented the dancer from showing sustained chaotic behavior. It would often stand upright after a very short duration. Second, the battery life for the system is short. Over the time of recording the dataset, the voltage of the battery quickly dropped.

To resolve these complications, we cut the battery connections from the dancer and plugged

Span of the dancer: 19.2 cm	
Height of the dancer: 30.0 cm	
Overall height: 41.0 cm	
Width of the base: 7.0 cm	Air Dancer
Diameter of the fan: 3.0 cm	

Fig. 3: Physical parameters of the air dancer system

it into a variable dc power supply with 6.70 V, and grounding the negative connection. This produced appropriate airflow through the blower to enforce the repetitive high-low pressure phenomenon. We recorded the video with a GoPro Hero5 Black. In total, we collected 27 videos, each of approximate length 150 seconds. We used 22 videos for training and validation and 5 videos for testing. Following the previous steps, all the videos were trimmed to 140 seconds to obtain the final dataset. The image frames were resized to 128×128 with 60 fps.

Lava lamp.

We downloaded a real world recording of lava lamp system from YouTube. The video lasts about 4 hours with 2.5 fps.

Fire.

We downloaded a real world recording of fire system from YouTube. The video lasts about 3,603 seconds with 24 fps.

Circular Motion.

We simulated a circle moving along a circular path with a constant speed to construct the circular motion dataset. We fixed the center and the radius of the circular path as well as the radius of

the circle. For each sequence, we randomly sampled the circle's initial position and constant speed. In total, we collected 1,100 sequences with 60fps. We used 880 of these sequences for training and validation, and 220 of them for testing.

Reaction diffusion.

We simulated the dynamics of a planar spiral wave to construct our reaction diffusion dataset. The dynamics of the system is driven by the following reaction-diffusion PDEs:

$$u_t = (1 - (u^2 + v^2))u + \beta(u^2 + v^2)v + d(u_{xx} + u_{yy}),$$

$$v_t = -\beta(u^2 + v^2)u + (1 - (u^2 + v^2))v + d(v_{xx} + v_{yy}),$$

with parameters $d = 0.1$ and $\beta = 1$. We ran only one simulation by solving the PDEs, following the original implementations. We then constructed the dataset by rendering the scalar field $u(x, y, t)$ at each time t as a 128×128 image with time t sampled at 5 fps. The long sequence of frames sampled from the simulation was divided into 100 shorter sequences. We used 80 of these sequences for training and validation, and 20 of them for testing.

Single pendulum.

We will present the physics equations of the single pendulum system in the "Physics Equations of Pendulum Systems" section. We set the pendulum mass as $m = 1\text{kg}$ and the pendulum length as $L = 0.5\text{m}$. For each sequence, we randomly sampled the initial position and velocity of the pendulum arm. In total, we collected 1,200 sequences with 60fps. We used 960 of these sequences for training and validation, and 240 of them for testing.

Elastic double pendulum.

We will present the physics equations of the elastic double pendulum system in the "Physics Equations of Pendulum Systems" section. We set the physical parameters of this system to be

the same as the physical parameters of the rigid double pendulum except that the first arm was replaced by a massless spring with a elasticity constant $k = 40\text{kg/s}^2$. For each sequence, we randomly sampled the initial angle and angular velocity of the spring, the initial length and stretch velocity of the spring, and the initial angle and angular velocity of the pendulum arm. In total, we collected 1,200 sequences with 60fps. We used 960 of these sequences for training and validation, and 240 of them for testing.

Physics Equations of Pendulum Systems

In this section, we provide more information on the physical state variables and equations of the three pendulum systems: the single pendulum, the rigid double pendulum, and the elastic double pendulum. See Fig. 4 for a graphical illustration of those systems.

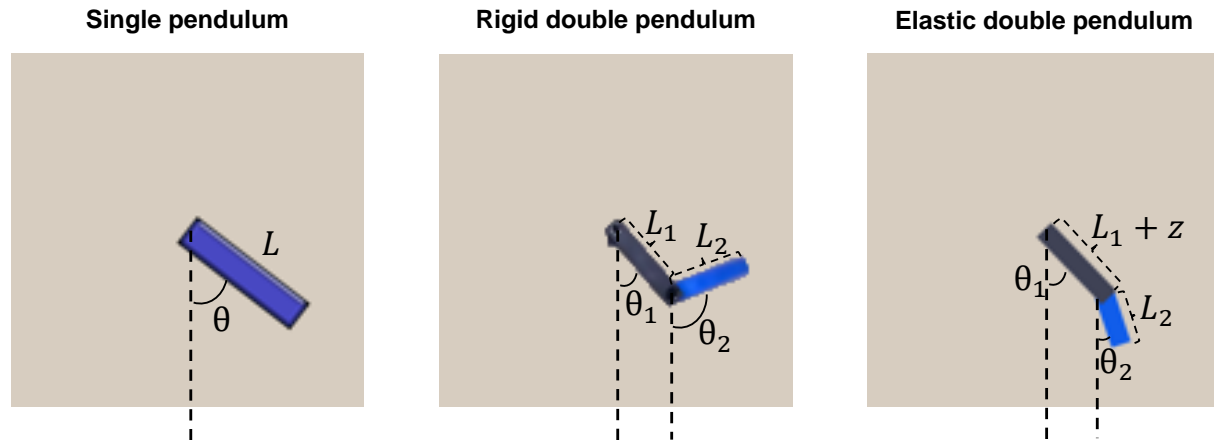Fig. 4: Illustration of our three pendulum systems

Single pendulum

Denote m the mass and L the length of the pendulum. The pendulum's momentum of inertia is $I = \frac{1}{3}mL^2$. We specify the system state by the pendulum's angular position θ and its respective angular velocity $\dot{\theta}$. The system's kinetic energy is

$$T = \frac{1}{2}I\dot{\theta}^2 = \frac{1}{6}mL^2\dot{\theta}^2.$$

Taking the configuration that the pendulum arm is horizontal as the zero point, the system's potential energy is

$$V = -\frac{1}{2}mgL\cos\theta.$$

Therefore, the system's Lagrangian is

$$L = T - V = \frac{1}{6}mL^2\dot{\theta}^2 + \frac{1}{2}mgL\cos\theta,$$

which gives the the equation of motion of the system:

$$\ddot{\theta} = -\frac{3g}{2L}\sin\theta.$$

The total energy of the system is:

$$E = T + V = \frac{1}{6}mL^2\dot{\theta}^2 - \frac{1}{2}mgL\cos\theta.$$

E is also the Hamiltonian of the system.

Rigid double pendulum

Denote m_1 and m_2 the masses of the two arms of the double pendulum, L_1 and W_1 the length and width of the first arm, and L_2 and W_2 the length and width of the second arm. The momenta of inertia of the two arms are:

$$I_1 = \frac{1}{12}(L_1^2 + W_1^2), \quad I_2 = \frac{1}{12}(L_2^2 + W_2^2).$$

We specify the system state by the two arms' angular positions θ_1 and θ_2, and their respective angular velocities $\dot{\theta}_1$ and $\dot{\theta}_2$.

The kinetic energy of the system is the sum of the two arms' translational and rotational kinetic energies, which is given by:

$$T = \frac{1}{2}\left(\frac{1}{4}m_1L_1^2 + m_2L_1^2 + I_1\right)\dot{\theta}_1^2 + \frac{1}{2}\left(\frac{1}{4}m_2L_2^2 + I_2\right)\dot{\theta}_2^2 + \frac{1}{2}m_2L_1L_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2),$$

8

The potential energy of the system is the sum of the two arms' gravitational potential energies. Taking the configuration that both arms are horizontal as the zero point, the potential energy of the system is given by:

$$V = -\left(\frac{1}{2}m_1 + m_2\right)gL_1\cos\theta_1 - \frac{1}{2}m_2gL_2\cos\theta_2,$$

The total energy of the system is the sum of kinetic and potential energies, which is given by:

$$E = \frac{1}{2}\left(\frac{1}{4}m_1L_1^2 + m_2L_1^2 + I_1\right)\dot{\theta}_1^2 + \frac{1}{2}\left(\frac{1}{4}m_2L_2^2 + I_2\right)\dot{\theta}_2^2 + \frac{1}{2}m_2L_1L_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2)$$
$$- \left(\frac{1}{2}m_1 + m_2\right)gL_1\cos\theta_1 - \frac{1}{2}m_2gL_2\cos\theta_2.$$

E is also the Hamiltonian of the system.

Elastic double pendulum

The elastic double pendulum is composed of a massless spring and a rigid pendulum. We denote m the mass of the pendulum. In addition, we denote L_1 the original length and z the stretch of the spring, and L_2 and W_2 the length and width of the pendulum. The pendulum's momentum of inertia is $I = \frac{1}{12}(L_2^2 + W_2^2)$.

We specify the system state by the spring's angular position θ_1, the pendulum's angular position θ_2, the spring stretch z, and their respective velocities $\dot{\theta}_1$, $\dot{\theta}_2$, $\dot{z}$.

The kinetic energy of the system is the sum of the pendulum's translational and rotational kinetic energies, which is given by

$$T = \frac{1}{2}m\left[(L_1 + z)^2\dot{\theta}_1^2 + \frac{1}{4}L_2^2\dot{\theta}_2^2 + \dot{z}^2 + (L_1 + z)L_2\cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + L_2\sin(\theta_1 - \theta_2)\dot{\theta}_2\dot{z}\right]$$
$$+ \frac{1}{2}I\dot{\theta}_2^2.$$

The potential energy of the system is the sum of the gravitational potential energy of the pendulum and the elastic potential energy of the spring. Taking the configuration that both the spring

and the pendulum are horizontal as the zero point of the pendulum's gravitational potential energy, the system's potential energy is given by:

$$V = mg\left[-(L_1 + z)\cos\theta_1 - \frac{1}{2}L_2\cos\theta_2\right] + \frac{1}{2}kz^2.$$

Therefore, the system's Lagrangian $L = T - V$ is given by

$$L = \frac{1}{2}m\left[(L_1 + z)^2\dot{\theta}_1^2 + \frac{1}{4}L_2^2\dot{\theta}_2^2 + \dot{z}^2 + (L_1 + z)L_2\cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + L_2\sin(\theta_1 - \theta_2)\dot{\theta}_2\dot{z}\right]$$

$$+ \frac{1}{2}I\dot{\theta}_2^2 + mg\left[(L_1 + z)\cos\theta_1 + \frac{1}{2}L_2\cos\theta_2\right] - \frac{1}{2}kz^2.$$

Then we can derive the system's equations of motion from the above Lagrangian, resulting in a 3×3 system of ODEs:

$$A\begin{bmatrix}\ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{z}\end{bmatrix} = b,$$

where

$$A = \begin{bmatrix} (L_1 + z)^2 & \frac{1}{2}(L_1 + z)L_2\cos(\theta_1 - \theta_2) & 0 \\ \frac{1}{2}(L_1 + z)L_2\cos(\theta_1 - \theta_2) & \frac{1}{4}L_2^2 + \frac{I}{m} & \frac{1}{2}L_2\sin(\theta_1 - \theta_2) \\ 0 & \frac{1}{2}L_2\sin(\theta_1 - \theta_2) & 1 \end{bmatrix},$$

and

$$b = \begin{bmatrix} -\frac{1}{2}(L_1 + z)L_2\sin(\theta_1 - \theta_2)(\dot{\theta}_2^2) - 2(L_1 + z)\dot{\theta}_1\dot{z} - g(L_1 + z)\sin\theta_1 \\ \frac{1}{2}(L_1 + z)L_2\sin(\theta_1 - \theta_2)(\dot{\theta}_1^2) - L_2\cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{z} - \frac{1}{2}gL_2\sin\theta_2 \\ (L_1 + z)\dot{\theta}_1^2 + \frac{1}{2}L_2\cos(\theta_1 - \theta_2)\dot{\theta}_2^2 + g\cos\theta_1 - \frac{k}{m}z \end{bmatrix}.$$

It is straightforward to verify that the matrix A is symmetric and positive definite. Therefore the ODE system is always solvable.

The system's total energy is the sum of its kinetic and potential energies, which is given by:

$$E = \frac{1}{2}m\left[(L_1 + z)^2\dot{\theta}_1^2 + \frac{1}{4}L_2^2\dot{\theta}_2^2 + \dot{z}^2 + (L_1 + z)L_2\cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + L_2\sin(\theta_1 - \theta_2)\dot{\theta}_2\dot{z}\right]$$

$$+ \frac{1}{2}I\dot{\theta}_2^2 - mg\left[(L_1 + z)\cos\theta_1 + \frac{1}{2}L_2\cos\theta_2\right] + \frac{1}{2}kz^2.$$

E is also the Hamiltonian of the system.

Model Details

In this section, we provide details about our models for dynamics prediction, latent reconstruction, and latent dynamics prediction with Neural State Variables.

Dynamics predictive model architecture

The dynamics predictive model is an auto-encoder with specific parameters listed in Fig. 5. All convolutional or transposed convolutional layers are accompanied with a batch normalization layer and a specified activation function. For the encoder network, after each "Conv" layer as shown in Fig. 5, we attach another convolutional layer with the same number of filters as the current convolutional layer but with 3×3 kernel and 1 as stride. For the decoder network, along with each "Deconv" layer as shown in Fig. 5 except for the last one, the input is also passed through a transposed convolutional layer with kernel size 4×4, 2 as stride, and a Sigmoid activation function. The output of this branch will then be concatenated with each "Deconv" layer along the feature dimension as the input of the next "Deconv" layer.

Latent reconstruction model architecture

The latent reconstruction model is also an auto-encoder with specific parameters listed in Fig. 6. Each layer is a linear layer accompanied with a sine activation function, and ID refers to the system's intrinsic dimension. The intermediate latent vectors whose dimension is ID are identified as Neural State Variables.

Neural latent dynamics model architecture

The neural latent dynamics model is a simple six-layer MLP, the widths of first five layers are (32, 64, 64, 64, 32, ID) where ID is the dimension of Neural State Variables, and each layer is accompanied with the ReLU activation function.

Layer	Kernel Size	#Filters	Stride	Padding	Activation
Conv1	4×4	32	2	1	ReLU
Conv2	4×4	32	2	1	ReLU
Conv3	4×4	64	2	1	ReLU
Conv4	4×4	128	2	1	ReLU
Conv5	3×4	128	(1,2)	1	ReLU
Deconv5	3×4	64	(1,2)	1	ReLU
Deconv4	4×4	64	2	1	ReLU
Deconv3	4×4	32	2	1	ReLU
Deconv2	4×4	16	2	1	ReLU
Deconv1	4×4	3	2	1	Sigmoid

Fig. 5: Dynamics predictive model architecture

Encoder Layer	#Filters	Activation	Decoder Layer	#Filters	Activation
Layer1	128	Sine	Layer5	32	Sine
Layer2	64	Sine	Layer6	64	Sine
Layer3	32	Sine	Layer7	128	Sine
Layer4	ID	Sine	Layer8	64	Sine

Fig. 6: Latent reconstruction model architecture

More Physics Evaluation Results for Dynamics Predictive Model

In this section, we show physics evaluation results for the high-dimensional dynamics predictive model on the single pendulum and rigid double pendulum systems. As shown in Fig. 7 and Fig. 8, our dynamics predictive model outperforms both the copy data and linear extrapolation baselines on both systems.

Method	θ (deg)	$\dot{\theta}$ (deg/s)	Energy (J)
Copy data	10.09 (±0.04)	37.44 (±0.14)	0.10 (±0.00)
Linear extrapolation	0.67 (±0.00)	37.44 (±0.14)	0.28 (±0.00)
Our model	0.22 (±0.00)	18.97 (±0.12)	0.15 (±0.00)

Fig. 7: Physics evaluation results on the single pendulum system

Method	θ_1 (deg)	θ_2 (deg)	$\dot{\theta}_1$ (deg/s)	$\dot{\theta}_2$ (deg/s)	Energy (J)
Copy data	8.29 (±0.04)	17.96 (±0.10)	111.64 (±0.78)	170.75 (±0.96)	0.06 (±0.00)
Linear extrapolation	2.39 (±0.02)	3.12 (±0.02)	111.64 (±0.78)	170.75 (±0.96)	0.08 (±0.00)
Our model	0.89 (±0.01)	1.55 (±0.01)	66.67 (±0.53)	113.94 (±0.84)	0.06 (±0.00)

Fig. 8: Physics evaluation results on the rigid double pendulum system

More Results from Intrinsic Dimension Estimation

There are several intrinsic dimension estimation algorithms in addition to the Levina-Bickel's algorithm. In this section we provide the intrinsic dimension estimation results from another four widely used algorithms: MiND_ML, MiND_KL, Hein, and CD using the same latent vectors. Those results together with results from the Levina-Bickel's algorithm and ground truth values are shown in Fig. 9. As shown in the results, Levina-Bickel's algorithm provides robust and accurate estimations across all six systems.

System	Ground Truth	Levina-Bickel	MiND_ML	MiND_KL	Hein	CD
Circular motion	2	2.19 (±0.05)	2.66 (±0.07)	3.00 (±0.00)	6.33 (±0.47)	5.23 (±0.41)
Reaction diffusion	2	2.16 (±0.14)	2.34 (±0.08)	3.00 (±0.00)	3.00 (±1.41)	3.08 (±1.04)
Single pendulum	2	2.05 (±0.02)	2.05 (±0.01)	2.00 (±0.00)	2.00 (±0.00)	2.03 (±0.04)
Rigid double pendulum	4	4.71 (±0.03)	4.69 (±0.03)	5.00 (±0.00)	3.33 (±0.47)	3.63 (±0.10)
Swing stick	4	4.89 (±0.33)	4.38 (±0.10)	4.80 (±0.40)	3.33 (±0.47)	3.15 (±0.10)
Elastic double pendulum	6	5.34 (±0.20)	5.15 (±0.13)	5.13 (±0.34)	4.00 (±0.00)	4.26 (±0.12)

Fig. 9: Intrinsic Dimension Estimation Results from Different Algorithms.

More Physics Evaluation Results for Hybrid Schemes

In this section, we provide physics evaluation results testing the hybrid schemes for long term predictions on the rigid double pendulum system. The hybrid scheme follows a $N + 1$ pattern where every N steps performed with the high-dimensional latent vectors are followed by a one step prediction using Neural State Variables. We implemented the hybrid $N + 1$ scheme with $N = 3, 4, 5, 6$ on the double pendulum system and compared physics evaluation results from the generated long term predictions. As shown in Fig. 10, the hybrid scheme always produces reasonable long-term predictions, and performance is not sensitive to the choice of N.

More Theoretical Analysis on Long-Term Prediction Stability

In this section, we provide more theoretical analysis on long term prediction stability. We will fix the notations that $\mathcal{S} \subset \mathbb{R}^D$ is the system's state space where $\mathbb{R}^D$ is the high-dimensional image space, $\mathrm{ID} = \dim \mathcal{S}$ is the system's intrinsic dimension, and LD is the latent space dimension of a given dynamics predictive model. The dimensions are assumed to satisfy $\mathrm{ID} \ll \mathrm{LD} \ll D$. The ground truth dynamics is given by $\boldsymbol{X}_{t+1} = F(\boldsymbol{X}_t)$ where $F : \mathcal{S} \to \mathcal{S}$ is the system evolution mapping.
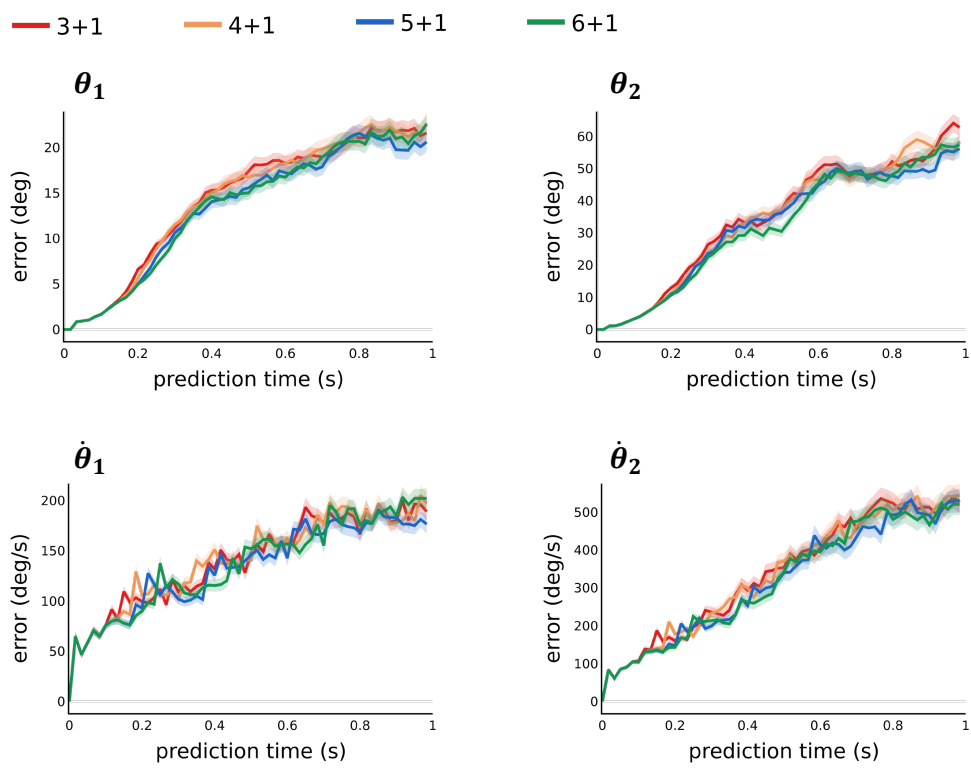
Fig. 10: Physics evaluation with hybrid schemes on the double pendulum system

For the theoretical analysis purpose, let us define:

$$M_{\mathcal{S}}(\hat{\boldsymbol{X}}) \triangleq \mathrm{dist}(\hat{\boldsymbol{X}}, \mathcal{S}) = \inf_{\boldsymbol{X} \in \mathcal{S}} \left\| \hat{\boldsymbol{X}} - \boldsymbol{X} \right\|,$$

as the metric measuring the deviation from any predicted state $\hat{\boldsymbol{X}}$ to the state space $\mathcal{S}$, where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^D$.

Now let us consider the long term predictions generated from model rollouts through Neural State Variables. Starting from any initial state $\boldsymbol{X}_0 \in \mathcal{S}$, the model rollouts produce:

$$\boldsymbol{X}_0 = \hat{\boldsymbol{X}}_0 \to \hat{\boldsymbol{V}}_0 \to \hat{\boldsymbol{X}}_{dt} \to \hat{\boldsymbol{V}}_{dt} \to \hat{\boldsymbol{X}}_{2dt} \to \hat{\boldsymbol{V}}_{2dt} \to \hat{\boldsymbol{X}}_{3dt} \to \cdots,$$

where $\hat{\boldsymbol{X}}_t \in \mathbb{R}^D$, $t = 0, dt, 2dt, \cdots$ are the predicted frames, and $\hat{\boldsymbol{V}}_t \in \mathbb{R}^{\mathrm{ID}}$, $t = 0, dt, 2dt, \cdots$ are the corresponding Neural State Variables. We denote $\phi_{\mathrm{E}} = h_{\mathrm{E}} \circ g_{\mathrm{E}}$ that maps every $\hat{\boldsymbol{X}}_t$ to $\hat{\boldsymbol{V}}_t$ and $\phi_{\mathrm{D}} = g_{\mathrm{D}} \circ h_{\mathrm{D}}$ that maps every $\hat{\boldsymbol{V}}_t$ to $\hat{\boldsymbol{X}}_{t+dt}$ for $t = 0, dt, 2dt, \cdots$. Both ϕ_{E} and ϕ_{D} are compositions of neural networks that are trained to minimize the one-step prediction error. Therefore, $\hat{F} = \phi_{\mathrm{D}} \circ \phi_{\mathrm{E}}$ provides an approximation of the ground truth evolution mapping F through Neural State Variables. We will make the following assumption regarding the one-step dynamics approximation error.

Assumption 1. $M_{\mathcal{S}}(\hat{\boldsymbol{X}}) \leq \varepsilon \quad \forall \hat{\boldsymbol{X}} = \hat{F}(\boldsymbol{X}), \ \boldsymbol{X} \in \mathcal{S}$ for some $\varepsilon > 0$.

We note that all possible ground truth and predicted states fall into the set

$$\mathcal{Z} \triangleq \bigcup_{n=0}^{\infty} \hat{F}^{(n)}(\mathcal{S}),$$

where $\hat{F}^{(0)}(\mathcal{S}) = \mathcal{S}$ and

$$\hat{F}^{(n)}(\mathcal{S}) = \{ \hat{\boldsymbol{X}} = \underbrace{\hat{F} \circ \hat{F} \cdots \circ \hat{F}}_{n \text{ times}}(\boldsymbol{X}) : \ \boldsymbol{X} \in \mathcal{S} \}, \quad n = 1, 2, \cdots$$

By definition $\mathcal{S} \subset \mathcal{Z} \subset \mathbb{R}^D$, which yields $\phi_{\mathrm{E}}(\mathcal{S}) \subset \phi_{\mathrm{E}}(\mathcal{Z}) \subset \mathbb{R}^{\mathrm{ID}}$. Here we make the second assumption saying that $\phi_{\mathrm{E}}(\mathcal{Z})$ is no larger than $\phi_{\mathrm{E}}(\mathcal{S})$ even if $\mathcal{Z}$ might be strictly larger than $\mathcal{S}$.

Assumption 2. $\phi_E(\mathcal{S}) = \phi_E(\mathcal{Z})$.

This assumption essentially relies on the dimensionality constraint. The state space $\mathcal{S}$ has dimension ID, while $\mathcal{Z}$ may have a higher dimension than ID. By projecting to the space of Neural State Variables whose dimension is ID, the extra dimensions of $\mathcal{Z}$ will be eliminated.

By Assumption 2 and the definition of $\mathcal{Z}$, we have:

$$\hat{F}(\mathcal{S}) = \phi_D \circ \phi_E(\mathcal{S}) = \phi_D \circ \phi_E(\mathcal{Z}) = \hat{F}(\mathcal{Z}) = \mathcal{Z}.$$

It together with Assumption 1 gives the following conclusion.

Proposition 1. Under Assumptions 1 and 2, we have $M_{\mathcal{S}}(\hat{\boldsymbol{X}}) \leq \varepsilon \quad \forall \hat{\boldsymbol{X}} \in \mathcal{Z}$.

As a direct corollary, for any long term prediction sequence $\{\hat{\boldsymbol{X}}_0, \hat{\boldsymbol{X}}_{dt}, \hat{\boldsymbol{X}}_{2dt}, \cdots\}$ from any initial state $\hat{\boldsymbol{X}}_0 = \boldsymbol{X}_0 \in \mathcal{S}$, we have $M_{\mathcal{S}}(\hat{\boldsymbol{X}}_t) \leq \varepsilon$ for all $t = 0, dt, 2dt, \cdots$. That is, the prediction sequence always stays in a fixed neighborhood of the ground truth state space $\mathcal{S}$ and the growth of $M_{\mathcal{S}}(\hat{\boldsymbol{X}}_t)$ is well-controlled, which guarantees the stability of long term predictions generated with Neural State Variables. For the model rollouts through latent vectors whose dimension LD $\gg$ ID, similar arguments yield that $g_E(\mathcal{S}) \subset g_E(\mathcal{Z}) \subset \mathbb{R}^{LD}$ and $g_E(\mathcal{Z})$ could have a dimension higher than ID. Therefore, the long term predictions could escape from the state space $\mathcal{S}$ through the extra dimensions. By shrinking the latent dimension to the system's intrinsic dimension, such instability formation can be avoided.

More Results on Neural State Variables for Dynamic Stability Indicators

In this section, we show long term stability evaluation results using dynamic stability indicators with Neural State Variables. In Fig. 11, the stability plots for six systems are shown. See the supplementary video for the corresponding visual predictions.
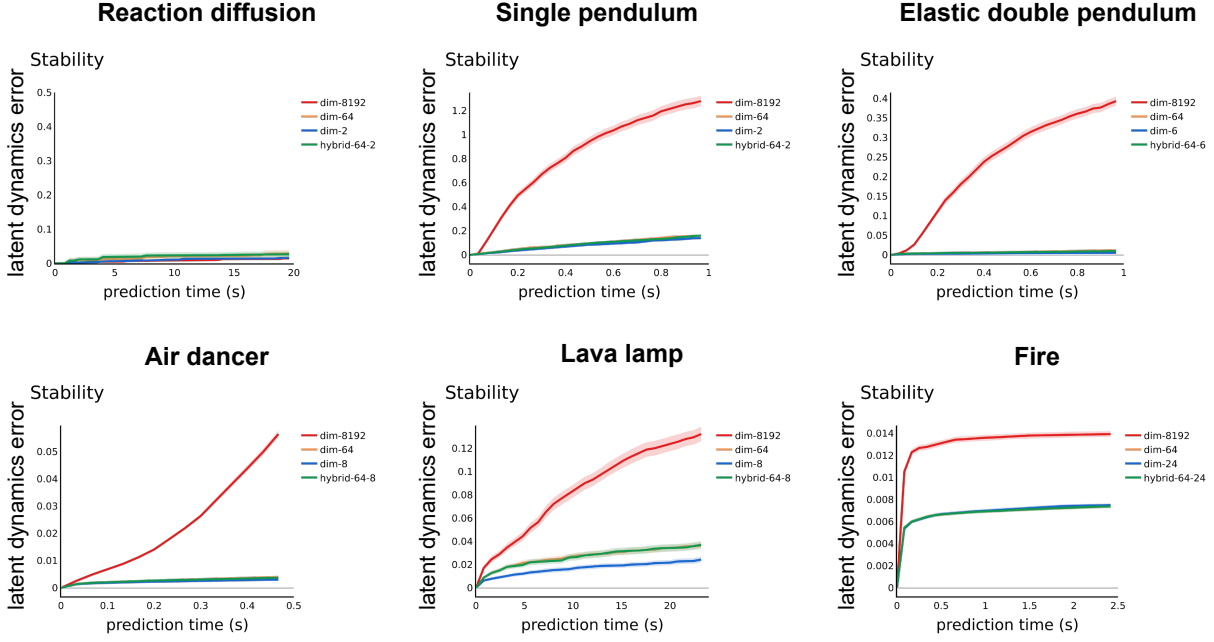
Fig. 11: Stability evaluation with Neural State Variables

More Analysis on Neural State Variables

In this section, we show more results to demonstrate the rich physics information contained in our learned Neural State Variables. First, we give the quantitative latent regression results on the elastic double pendulum system in Fig. 12. It can be observed that the learned Neural State Variables capture much richer information about the system dynamics than the variables obtained through PCA from high dimensional latent embedding vectors.

We also show visualizations of the Neural State Variables after applying PCA on them. See Fig. 13.

Latent Variables	θ_1 (deg)	θ_2 (deg)	z (m)	$\dot{\theta}_1$ (deg/s)	$\dot{\theta}_2$ (deg/s)	$\dot{z}$ (m/s)	Total energy (J)
dim-6 PCA of dim-8192 Latents	23.86 (±1.14)	43.03 (±0.72)	0.02 (±0.00)	149.36 (±3.20)	397.20 (±9.50)	0.52 (±0.01)	0.08 (±0.00)
dim-6 PCA of dim-64 Latents	14.10 (±0.88)	27.96 (±0.85)	0.01 (±0.00)	170.68 (±6.37)	396.70 (±2.90)	0.59 (±0.01)	0.07 (±0.00)
dim-6 Latents	7.85 (±0.54)	14.11 (±0.50)	0.01 (±0.00)	70.03 (±1.41)	163.13 (±4.10)	0.27 (±0.01)	0.05 (±0.00)

Fig. 12: Latent regression results on the elastic double pendulum system
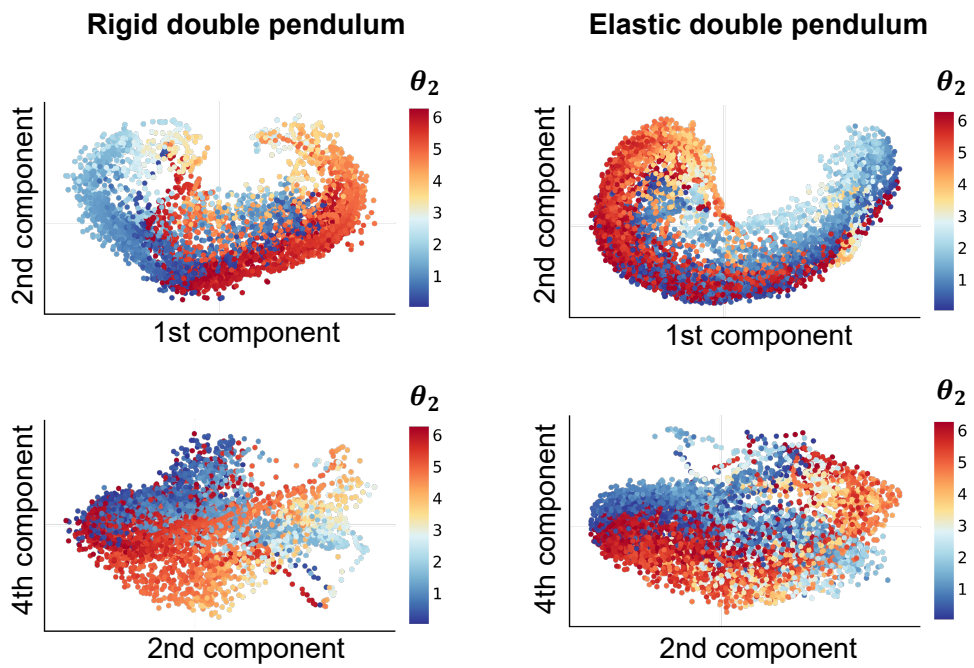
Rigid double pendulum   **Elastic double pendulum**

Fig. 13: Visualization of Neural State Variables