Deep learning for the Discovery of Parsimonious Physics Models

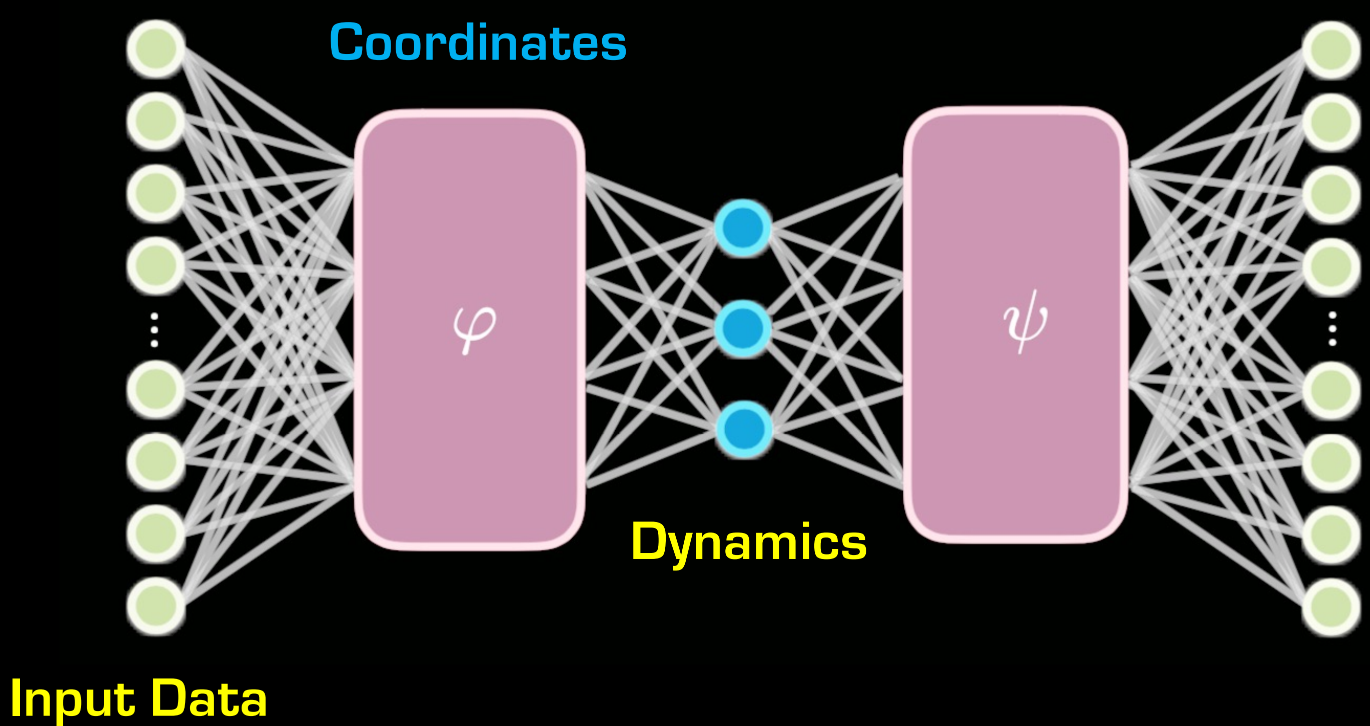
J. Nathan Kutz

Department of Applied Mathematics
University of Washington
Email: **kutz@uw.edu**

Steven L. Brunton

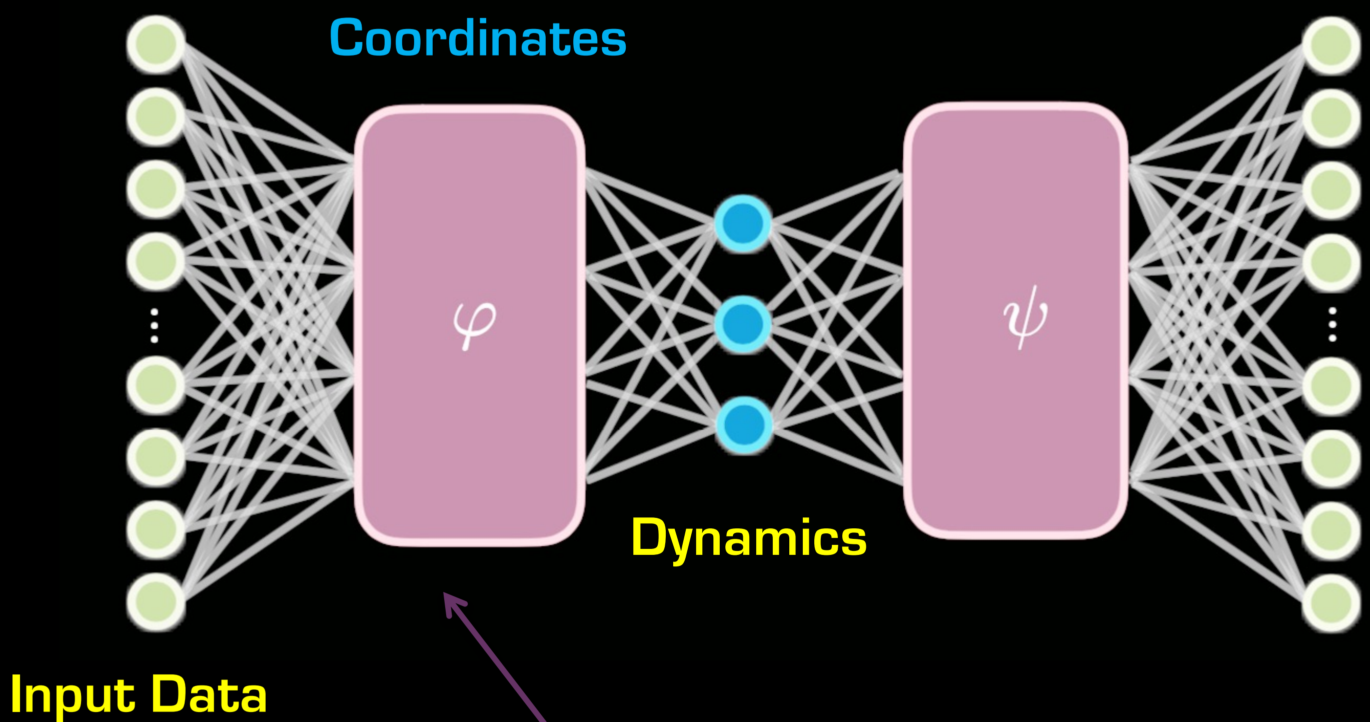
Lisbon – September 16, 2021

Coordinates & Dynamics



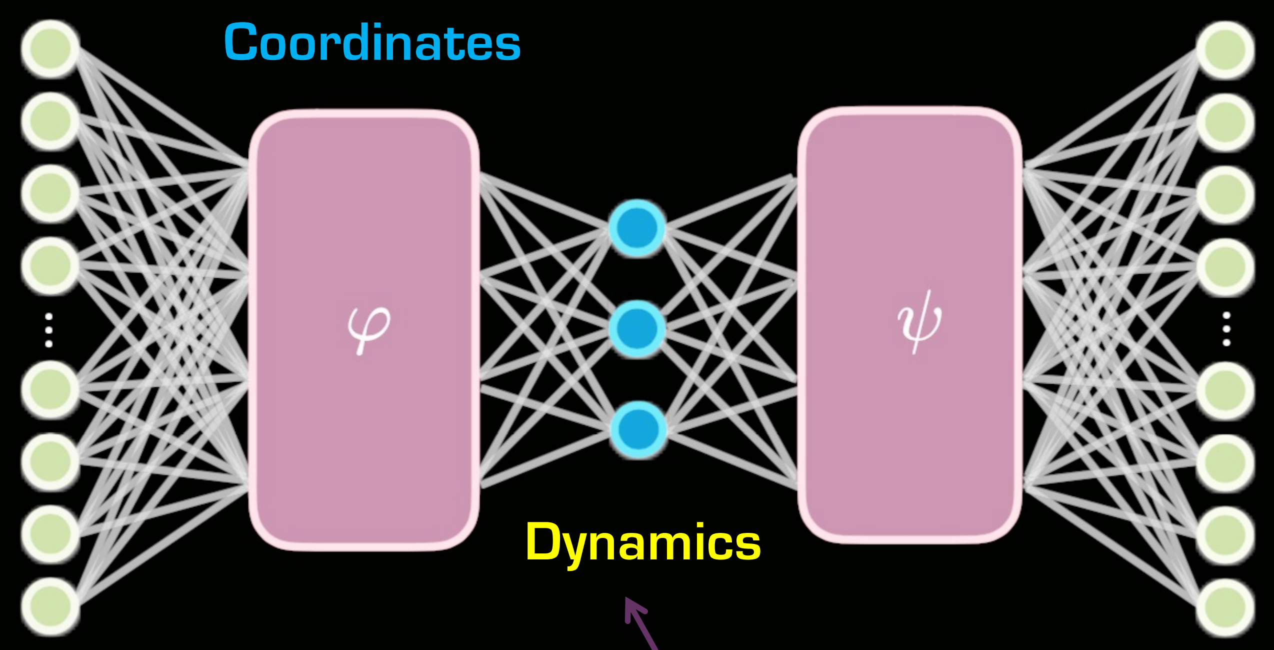
Targeted use of neural networks for discovery coordinate transformations

Coordinates & Dynamics



- Special functions *(Bessel, Hermite, Laguerre,etc)*
- expert knowledge
- SVD-based: POD/PCA/EOF/Hoteling
- Neural Nets

Coordinates & Dynamics



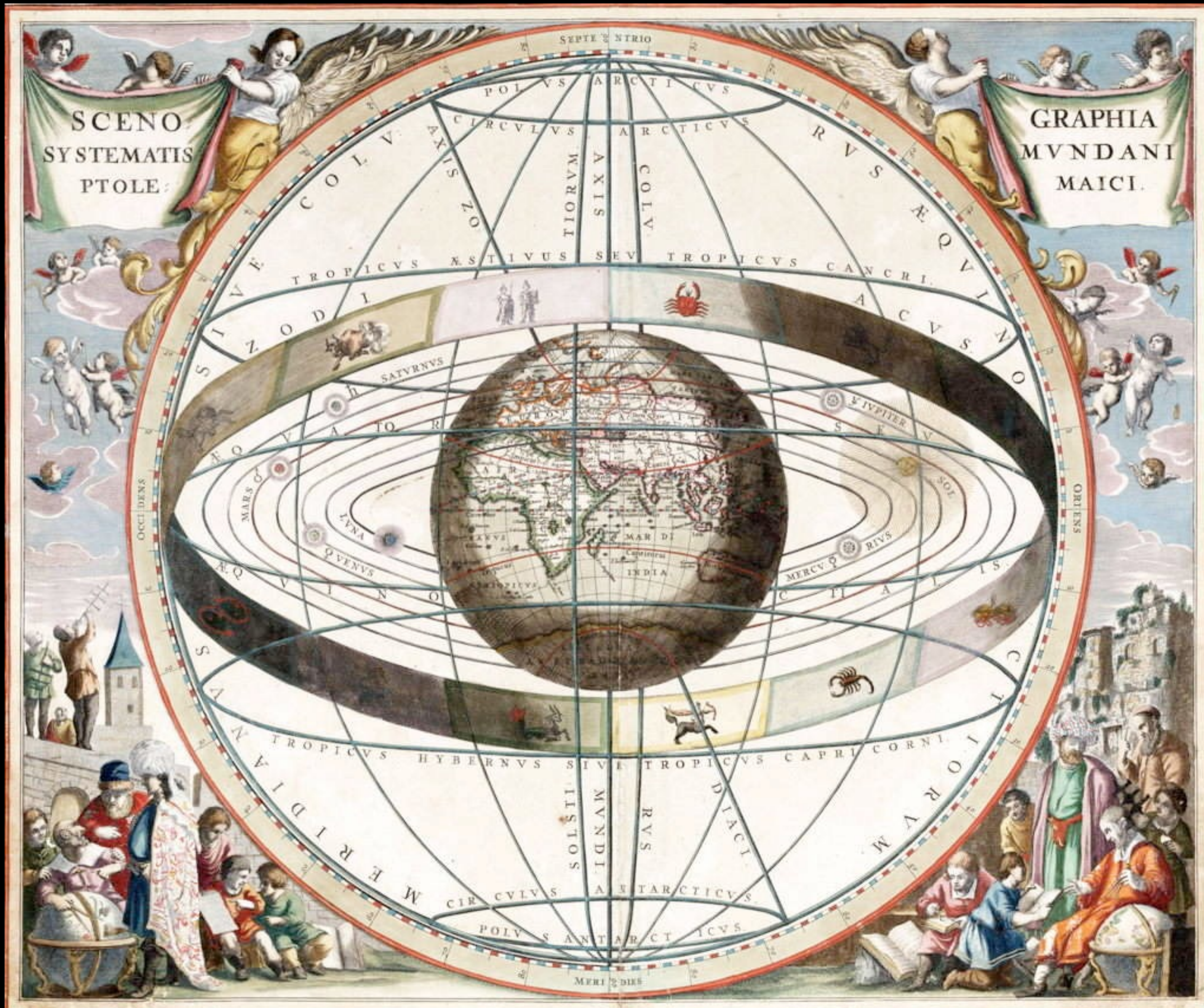
Input data
(measurements)

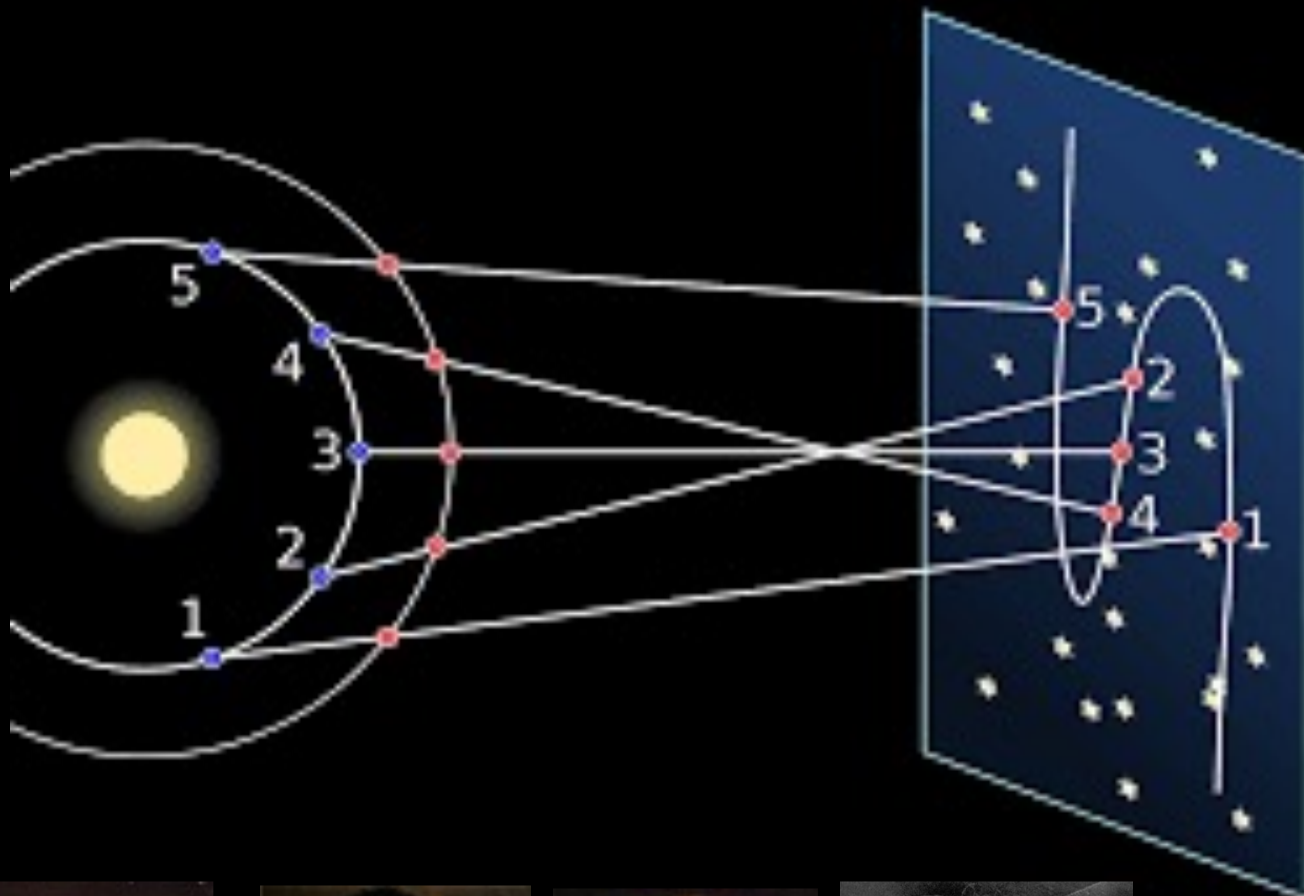
- ARIMA
- DMD & Koopman
- SINDy: Discover ODEs or PDEs
- Normal Forms
- LSTM, GRU, etc

Good Coordinates

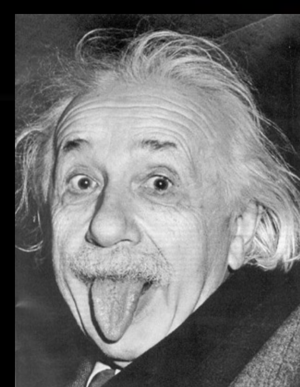

Saturn & Mars

Doctrine of the Perfect Circle

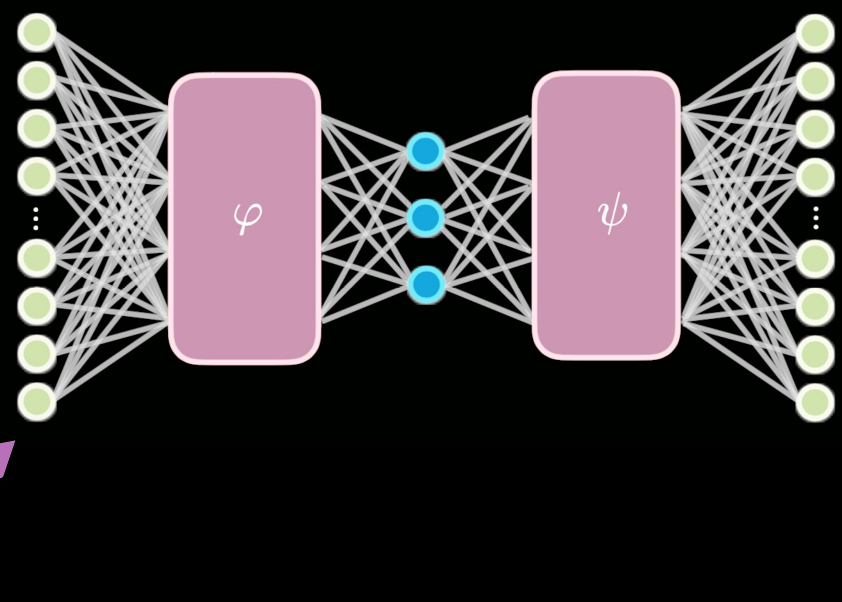


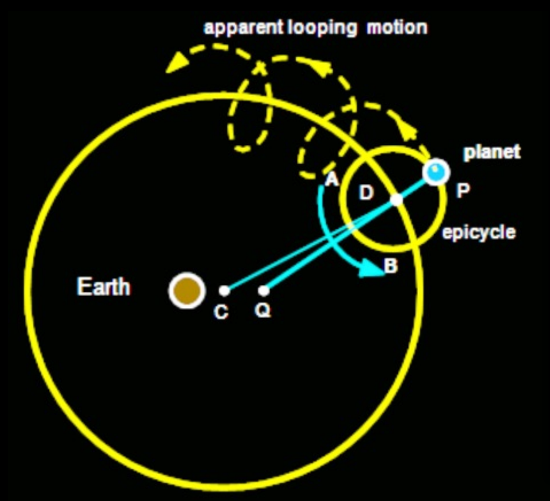

5
4
3
2
1
5
2
3
4
1



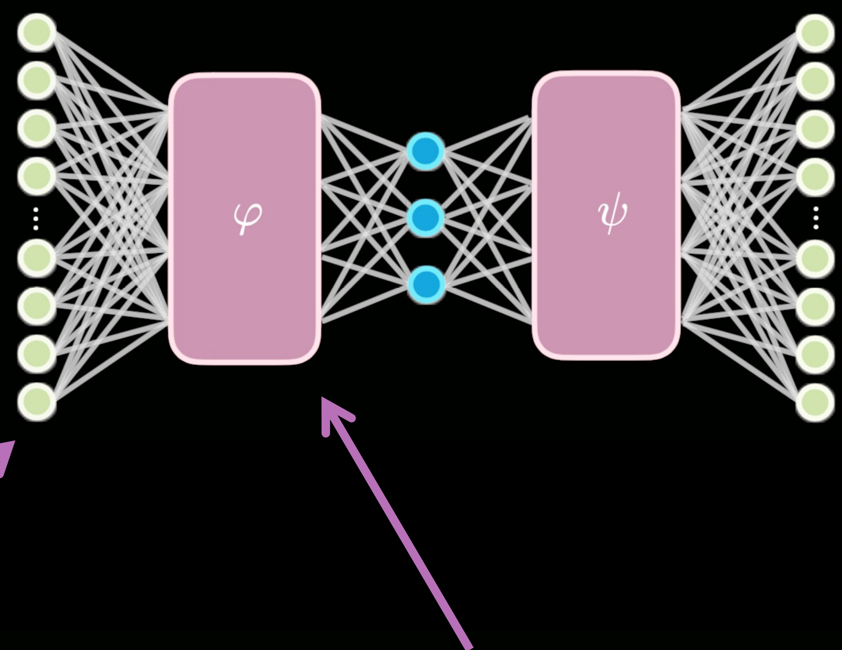



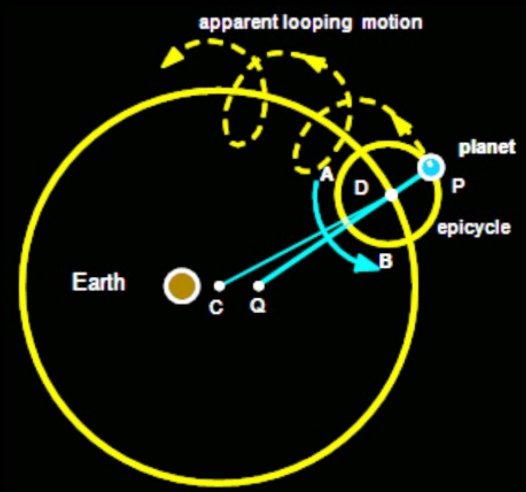
Discovery Paradigm

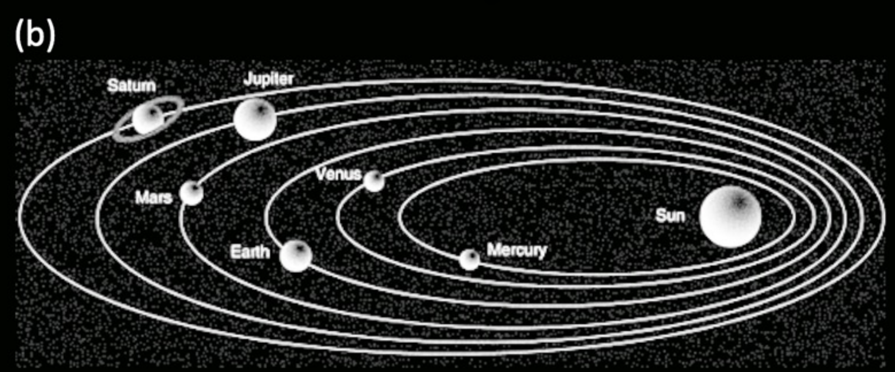




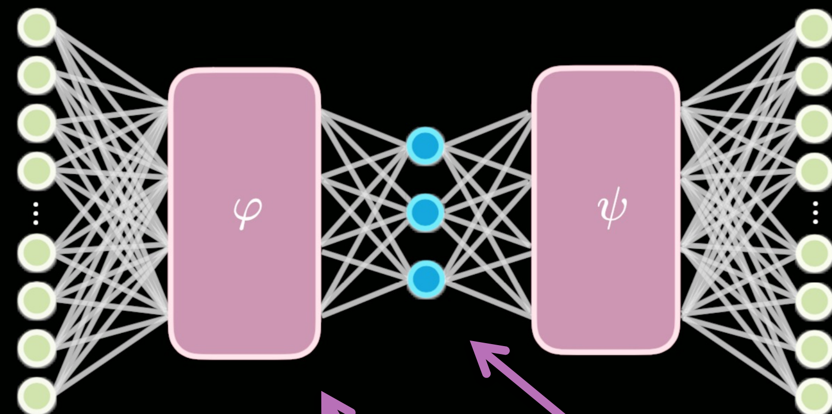
Discovery Paradigm

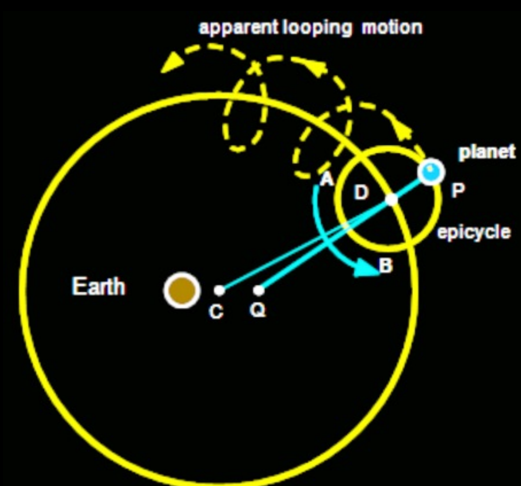


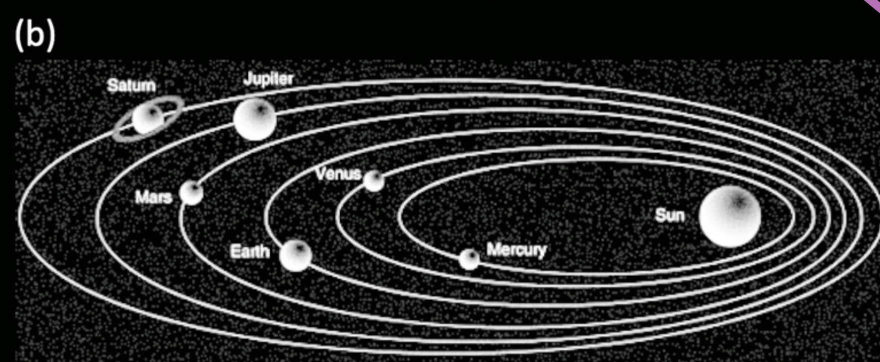




Discovery Paradigm







$$F = G\frac{m_1 m_2}{r^2}$$

Kepler vs Newton

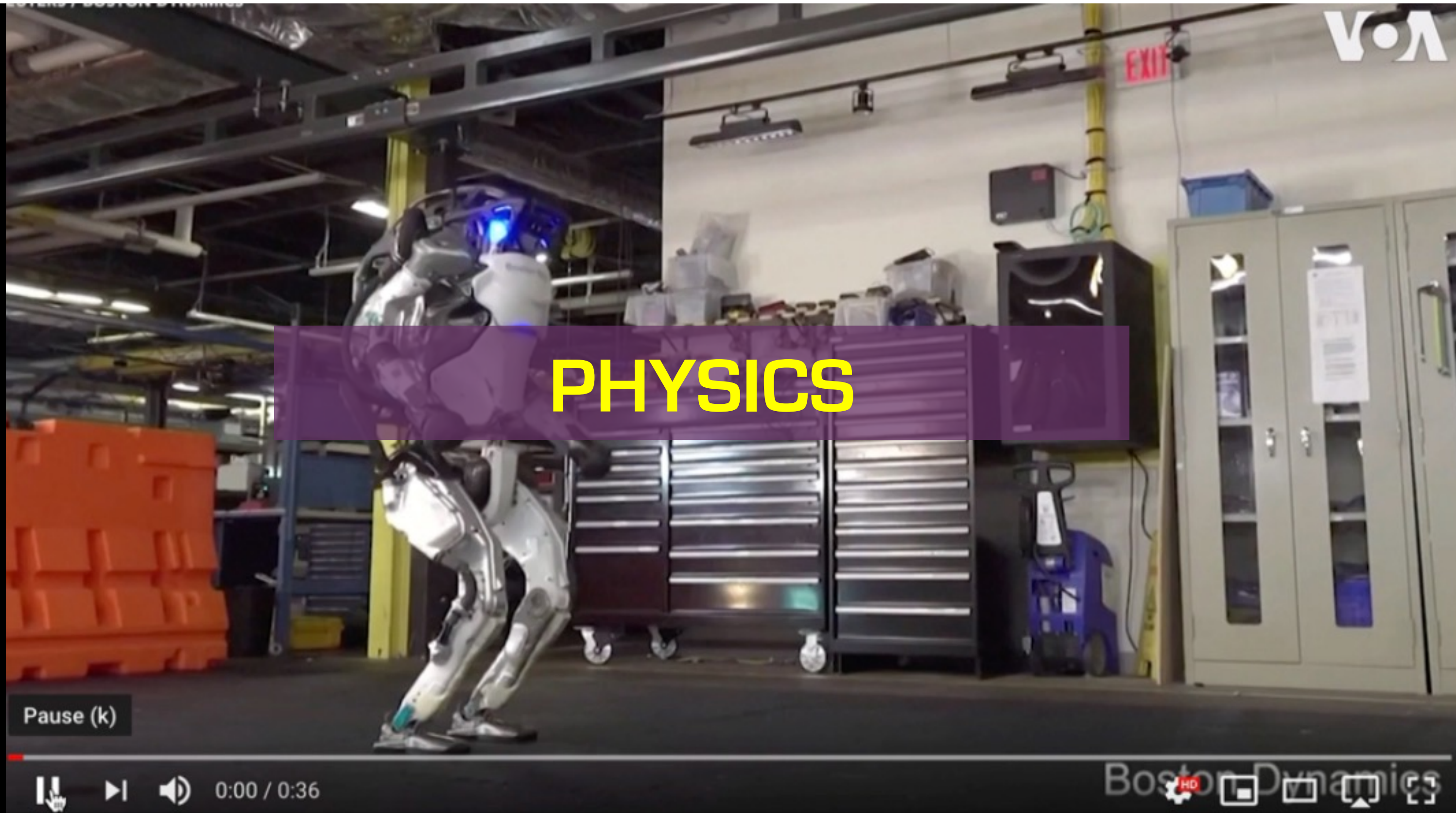
function approximation (ellipses)

F=ma (ellipses)

Newton

Kepler

Breiman – Two Cultures


VOA
EXIT
PHYSICS
Pause (k)
0:00 / 0:36
Boston Dynamics


NVIDIA
NO PHYSICS
Pause (k)
It began its driving lessons in a parking lot

Mathematical Formulation

Mathematical Framework

Dynamics

State-space

Parameters

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, t, \Theta, \Omega)$$

Dynamics

Stochastic effects

Measurement

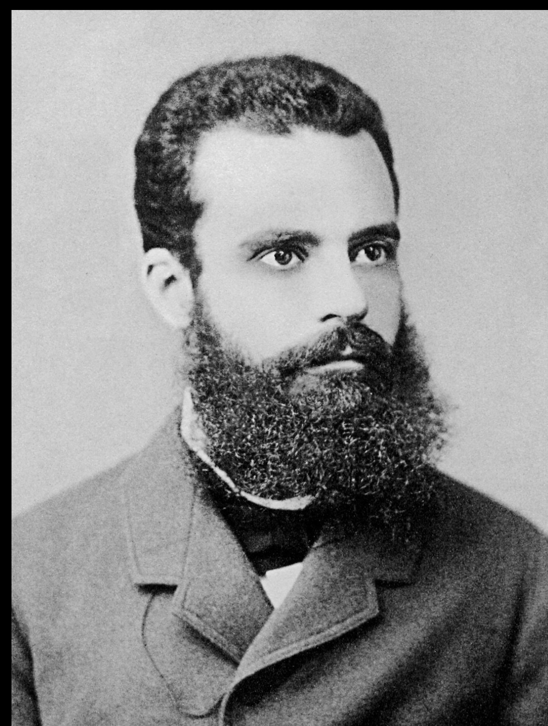
$$\mathbf{y}(t_k) = h(t_k, \mathbf{x}(t_k), \Xi)$$

Measurement model

Measurement noise

Interpretability & Parsimony

William of Occam



Vilfredo Pareto

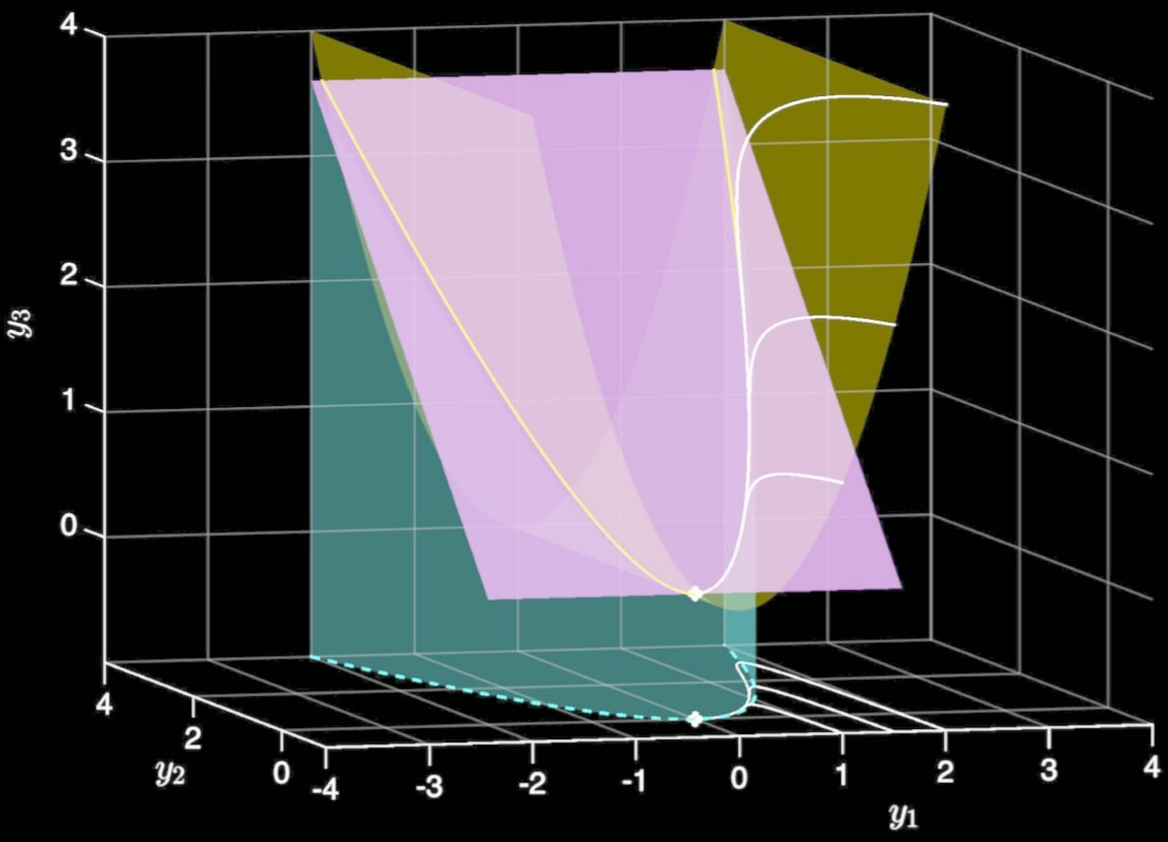
The Ultimate Physics
Regularization

# of terms
# of dimensions

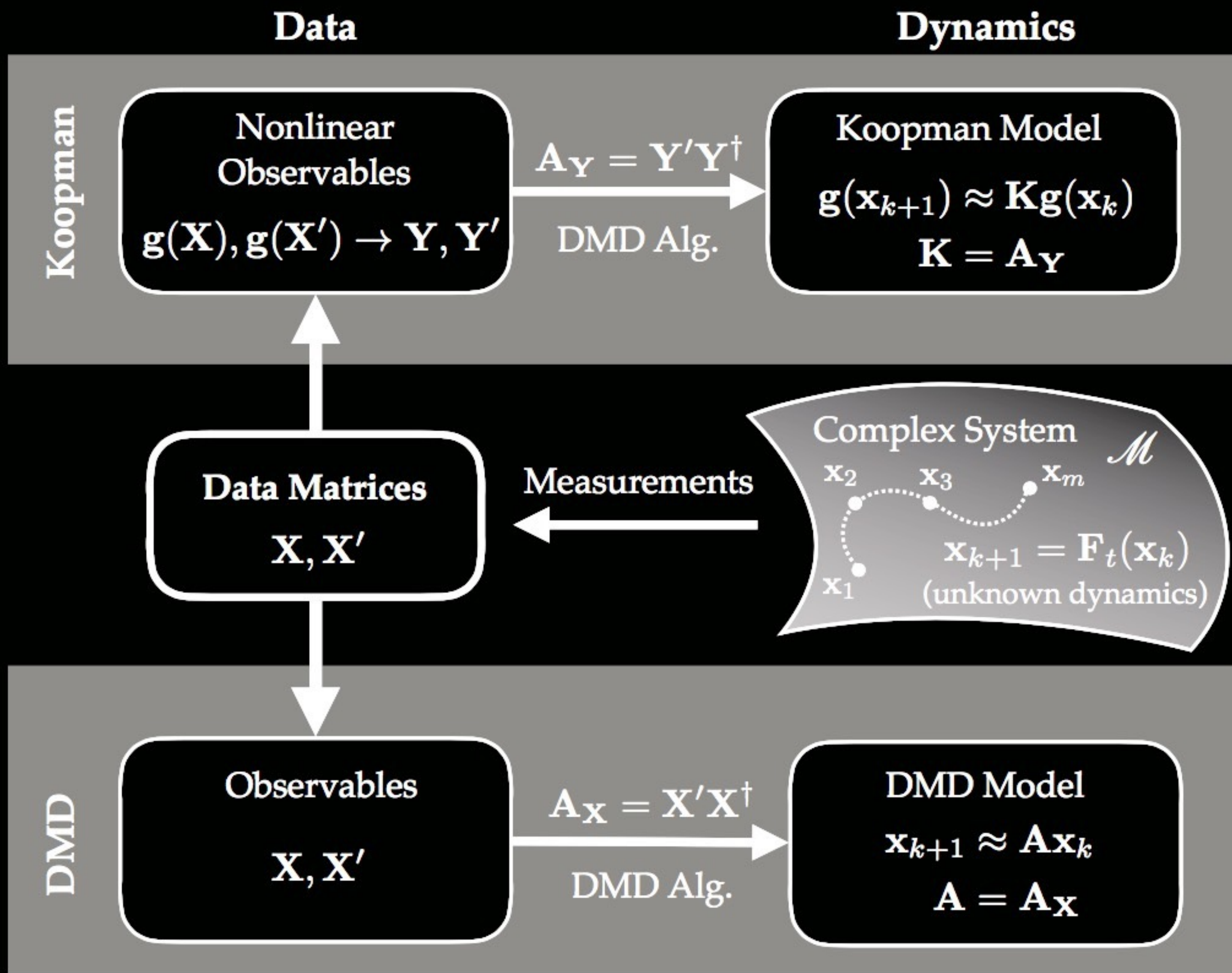
Linear Models

Koopman Invariant Subspaces

$$\left.\begin{aligned}\dot{x}_1 &= \mu x_1 \\ \dot{x}_2 &= \lambda(x_2 - x_1^2)\end{aligned}\right\} \implies \frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \mu & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & 2\mu \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad \text{for} \quad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \end{bmatrix}$$



Koopman and DMD

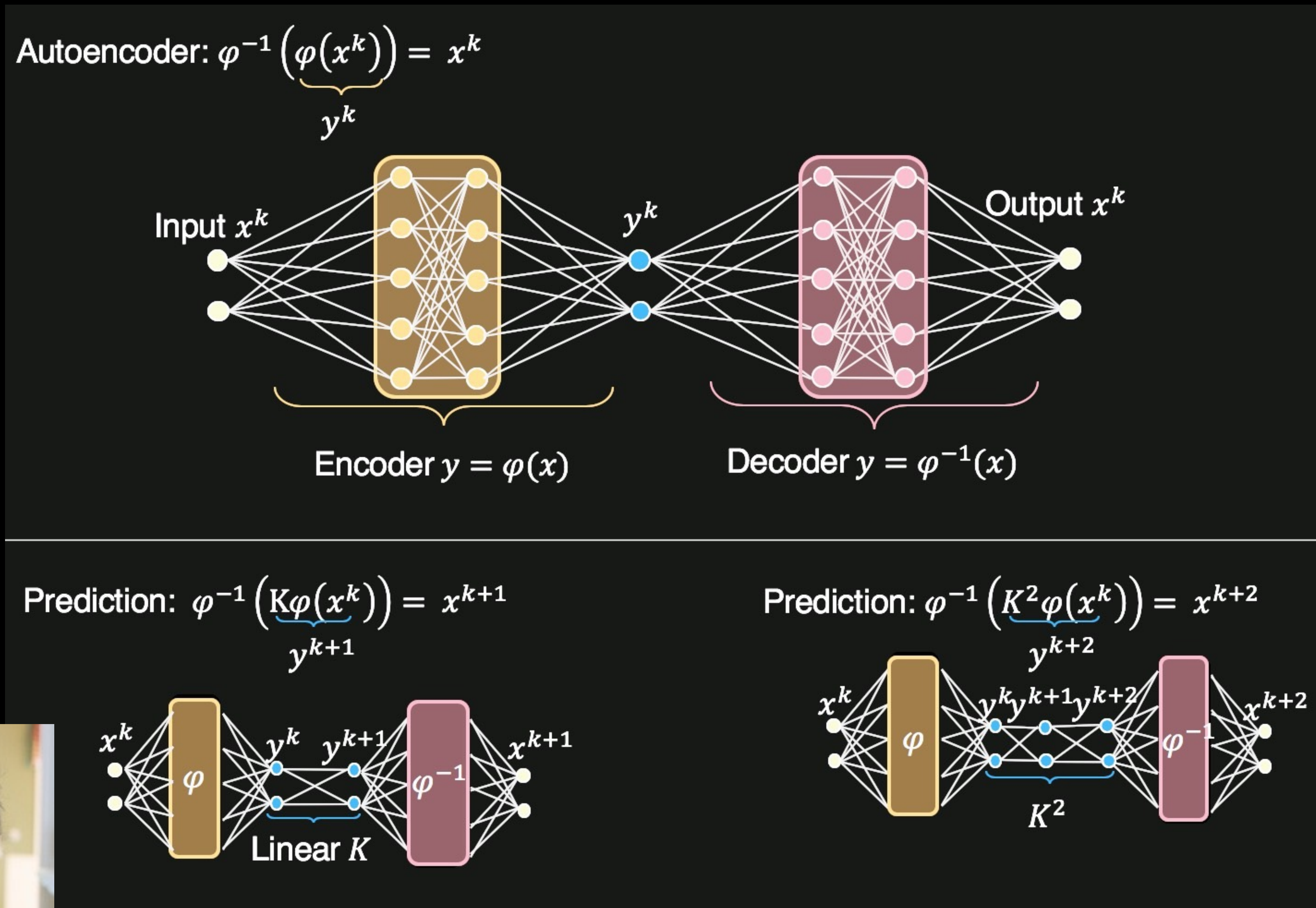


Neural Nets

"Supervised learning is a high-dimensional interpolation problem."

S. Mallat, PRSA (2016)

NNs for Koopman Embedding



Bethany Lusch

Lusch et al. Nat. Comm (2018)

Failure!

(obviously)

Duffing Oscillator

Poincare-Lindstedt Expansion: let $\tau = \omega t$ so that

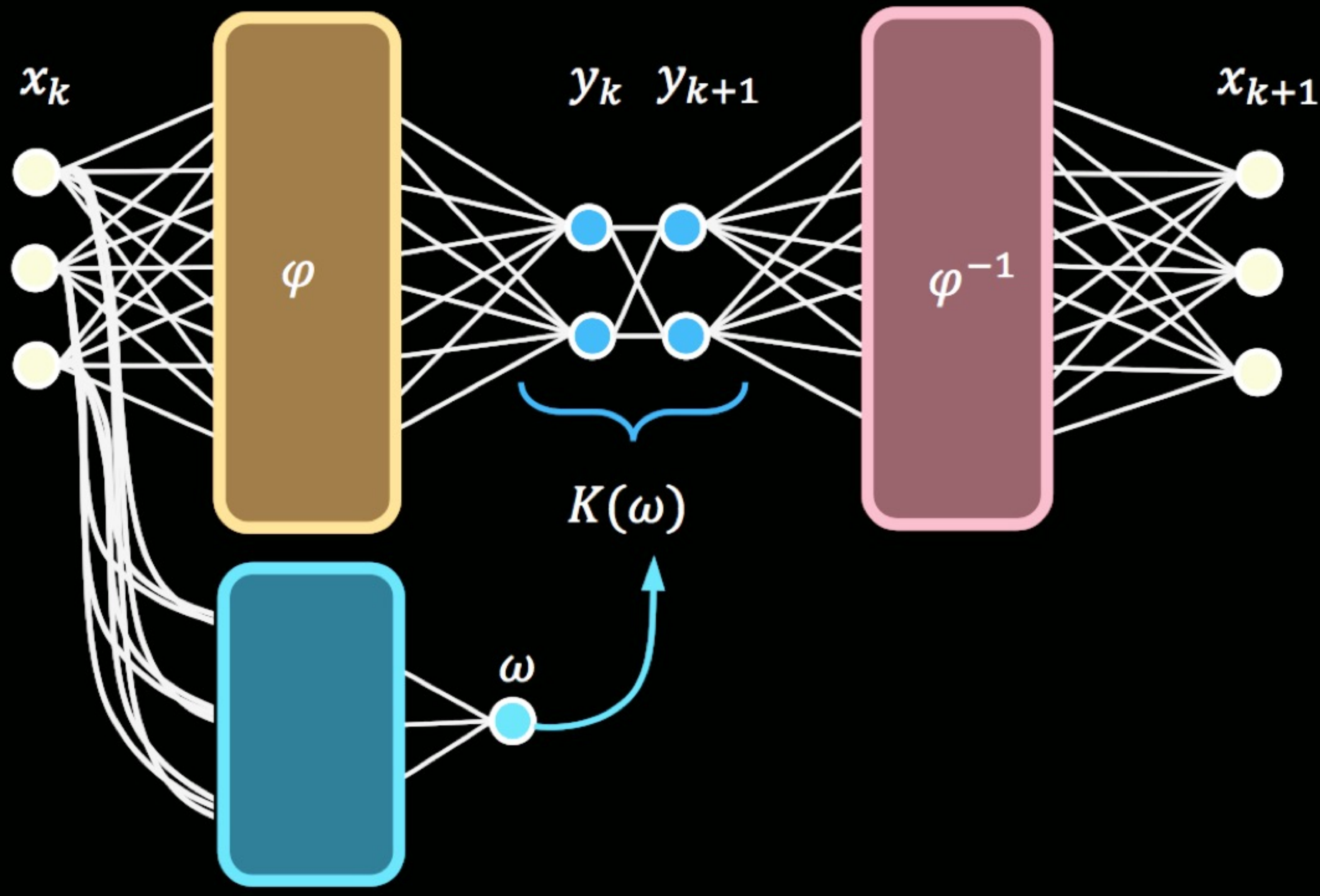
$$y_{tt} + y + \epsilon y^3 = 0 \Rightarrow \omega^2 y_{\tau\tau} + y + \epsilon y^3 = 0$$

Nonlinearity: Shifts Frequencies + Generates Harmonics

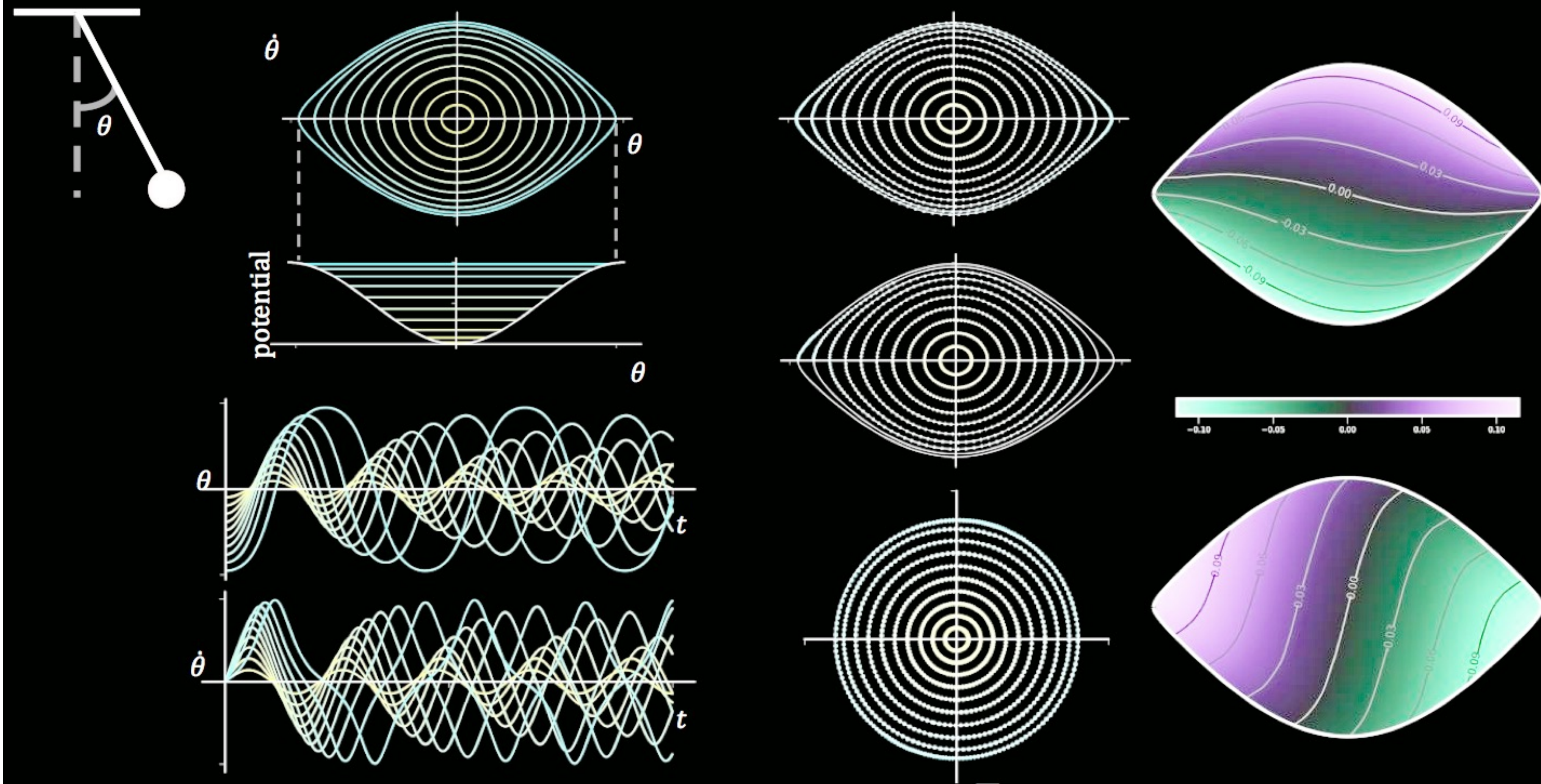
$$y = A \sin[(1+\epsilon 3A^2/8)t] + \epsilon \left\{ \frac{3A^3}{32} \sin\left[\left(1+\epsilon \frac{3A^2}{8}\right)t\right] - \frac{A^3}{32} \sin\left[3\left(1+\epsilon \frac{3A^2}{8}\right)t\right] \right\}$$

Spectrogram

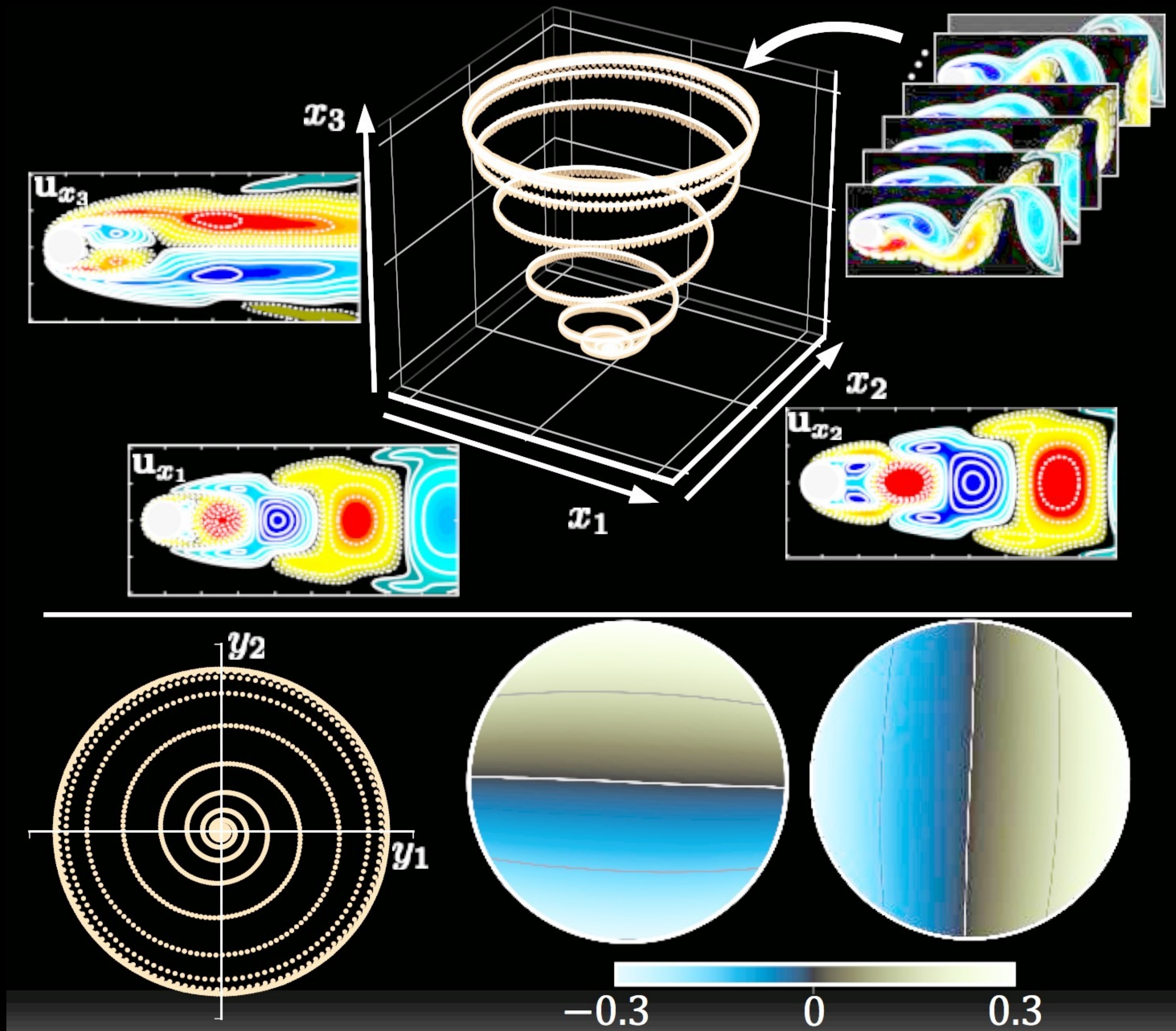
Handling the Continuous Spectra



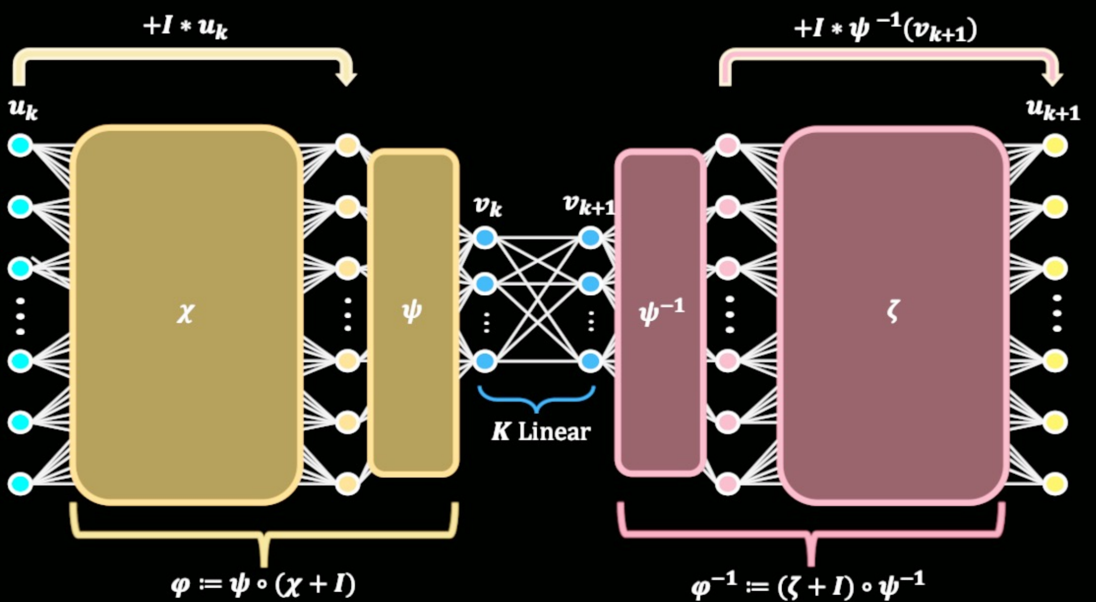
The Pendulum

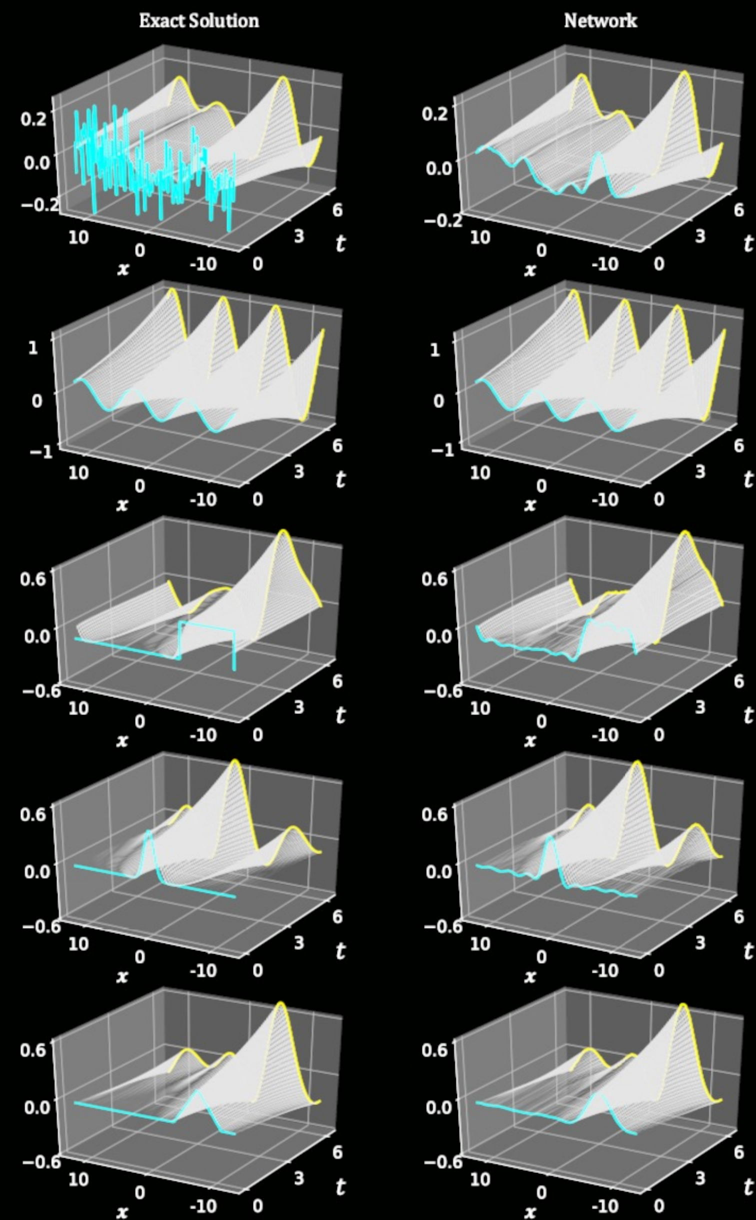


Flow Around a Cylinder



NNs for PDE Koopman Embedding





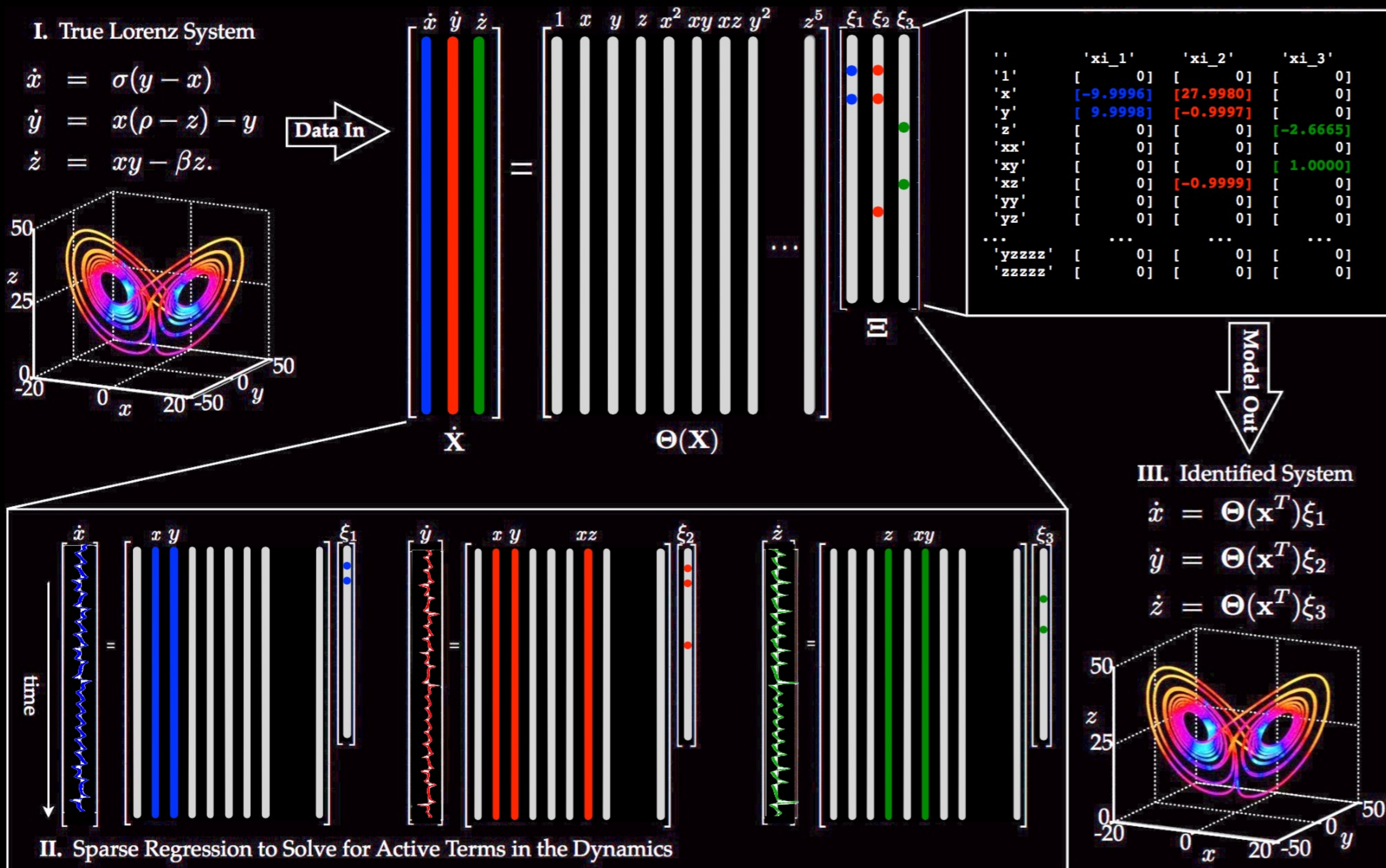
Like IST: Linearize Kuramoto-Sivashinsky

Craig Gin

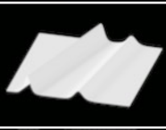
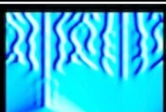
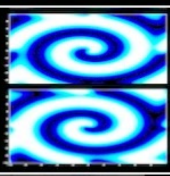
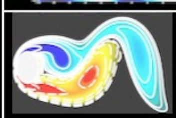
Gin et al. arxiv (2019)

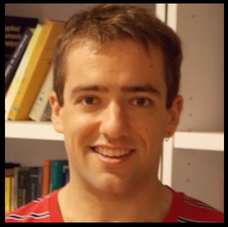
Governing Equations

Sparse Identification of Nonlinear Dynamics (SINDy)



Brunton, Proctor & Kutz, PNAS (2016)

PDE	Form	Error (no noise, noise)	Discretization
 KdV	$u_t + 6uu_x + u_{xxx} = 0$	1%±0.2%, 7%±5%	$x \in [-30, 30], n=512, t \in [0, 20], m=201$
 Burgers	$u_t + uu_x - \epsilon u_{xx} = 0$	0.15%±0.06%, 0.8%±0.6%	$x \in [-8, 8], n=256, t \in [0, 10], m=101$
 Schrodinger	$iu_t + \frac{1}{2}u_{xx} - \frac{x^2}{2}u = 0$	0.25%±0.01%, 10%±7%	$x \in [-7.5, 7.5], n=512, t \in [0, 10], m=401$
 NLS	$iu_t + \frac{1}{2}u_{xx} + \|u\|^2 u = 0$	0.05%±0.01%, 3%±1%	$x \in [-5, 5], n=512, t \in [0, \pi], m=501$
 KS	$u_t + uu_x + u_{xx} + u_{xxxx} = 0$	1.3%±1.3%, 70%±27%	$x \in [0, 100], n=1024, t \in [0, 100], m=251$
 R-D	$u_t = 0.1\nabla^2 u + \lambda(A)u - \omega(A)v$ $v_t = 0.1\nabla^2 v + \omega(A)u + \lambda(A)v$ $A = u^2 + v^2, \omega = -\beta A^2, \lambda = 1 - A^2$	0.02% ± 0.01%, 3.8% ± 2.4%	$x, y \in [-10, 10], n=256, t \in [0, 10], m=201$ subsample $3 \cdot 10^5$
 Navier Stokes	$\omega_t + (\mathbf{u} \cdot \nabla)\omega = \frac{1}{Re}\nabla^2 \omega$	1% ± 0.2% , 7% ± 6%	$x \in [0, 9], n_x=449, y \in [0, 4], n_y=199,$ $t \in [0, 30], m=151,$ subsample $3 \cdot 10^5$



Sam Rudy

Rudy, Brunton, Proctor & Kutz, Sci. Adv (2017)
Schaeffer, PRSA (2017)

Discrepancy Modeling

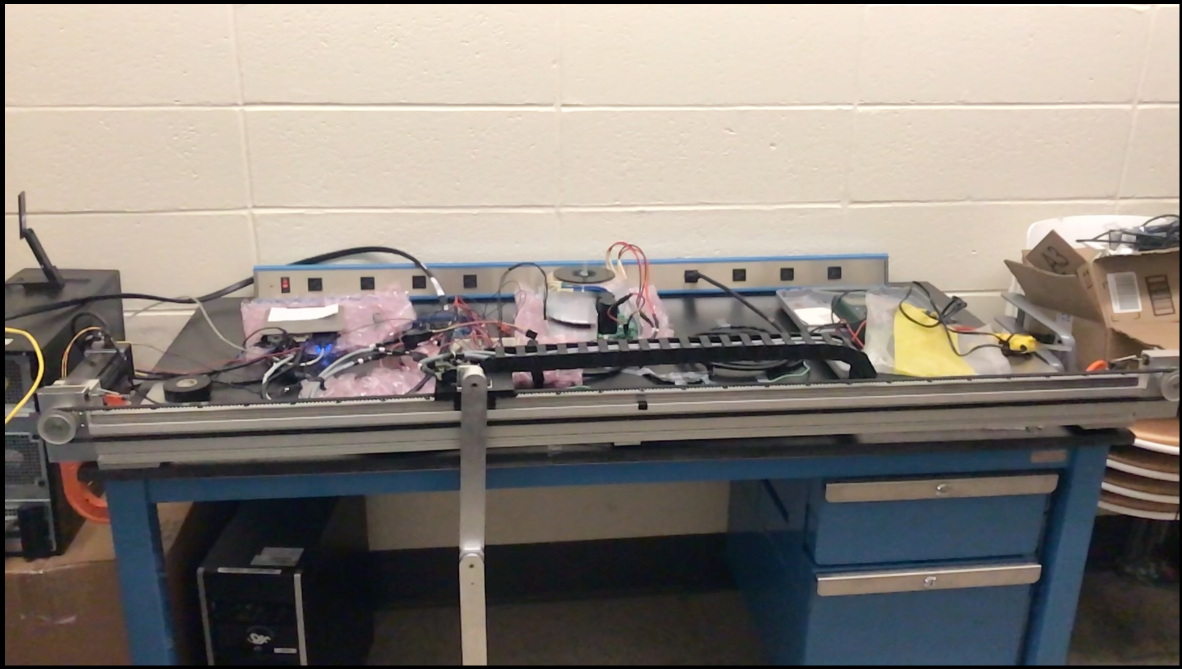
Instead of model discovery from scratch…

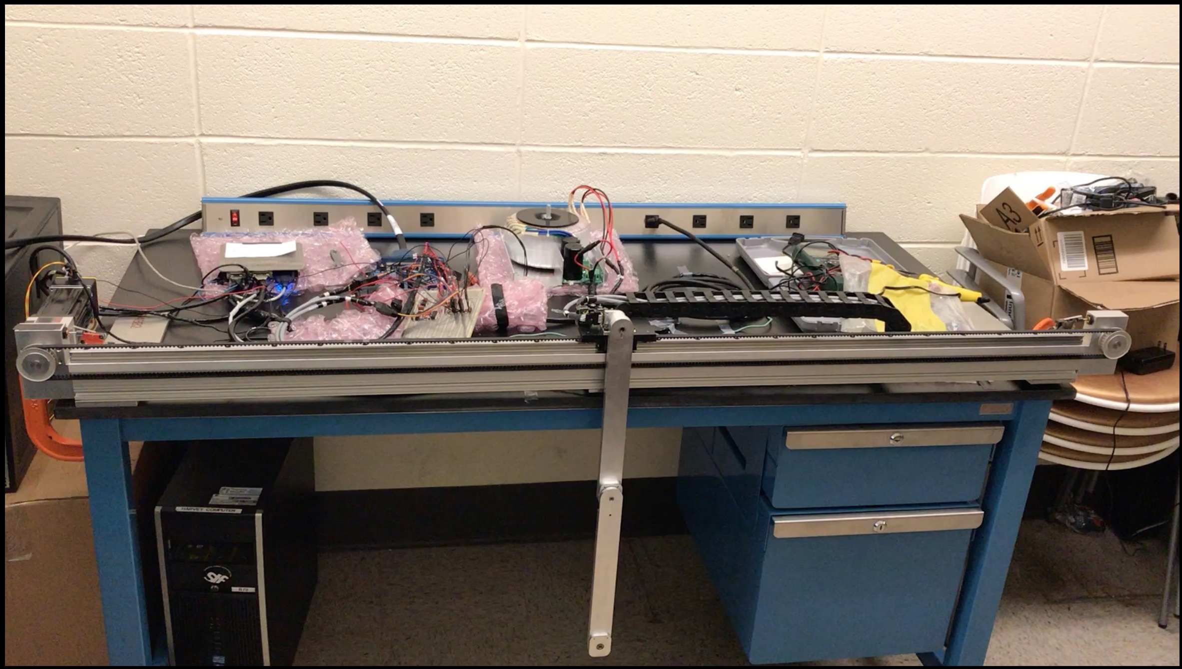
…we often start with partial knowledge of the physics

- Idealized Hamiltonian or Lagrangian system
- Knowledge of constraints, conservation laws, symmetries

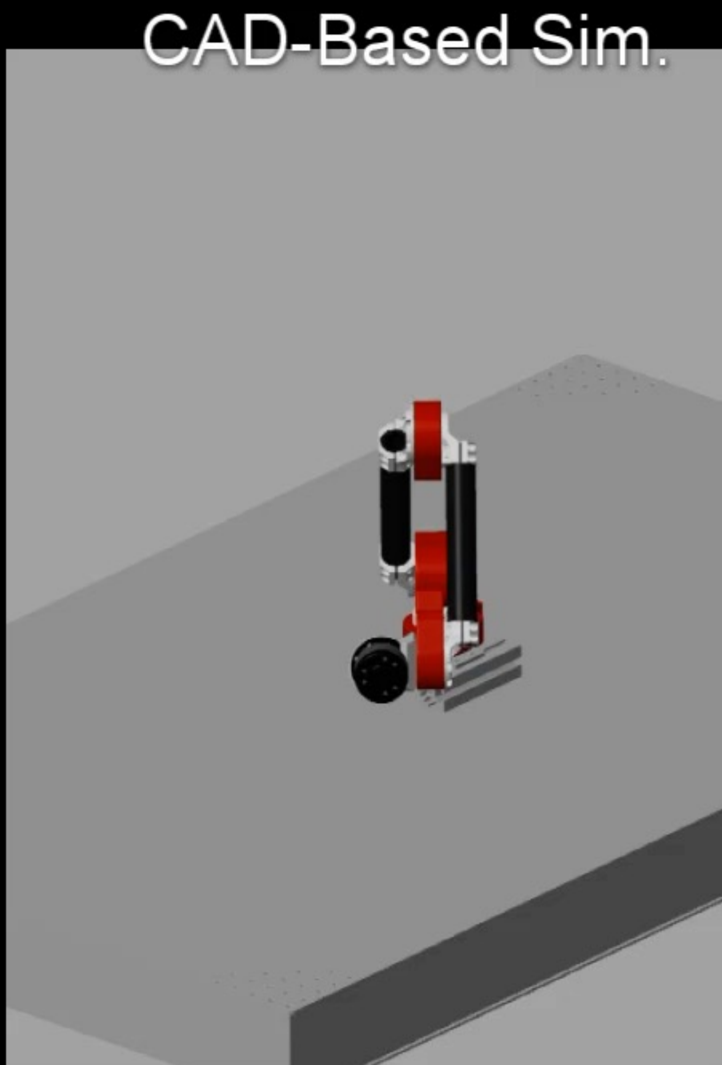
$$\frac{d}{dt}\mathbf{x} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\delta}\mathbf{g}(\mathbf{x})$$

Imperfect model — Discrepancy



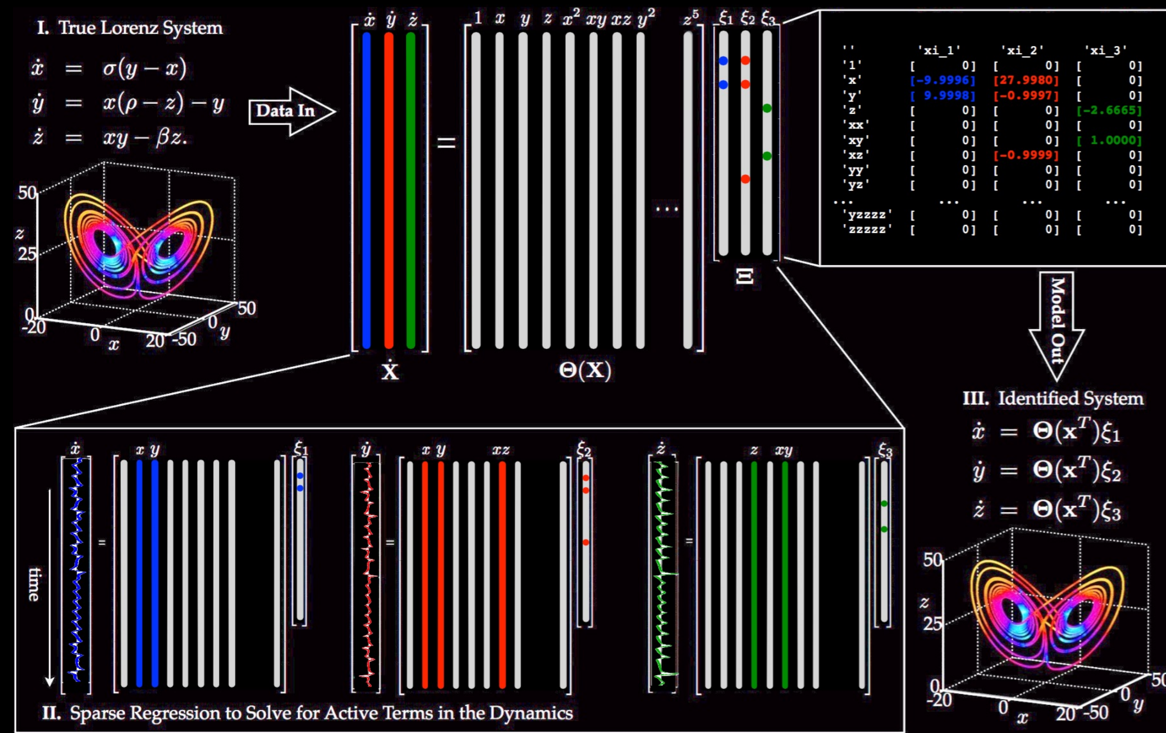


Digital Twins



Sparse Identification of Nonlinear Dynamics (SINDy)

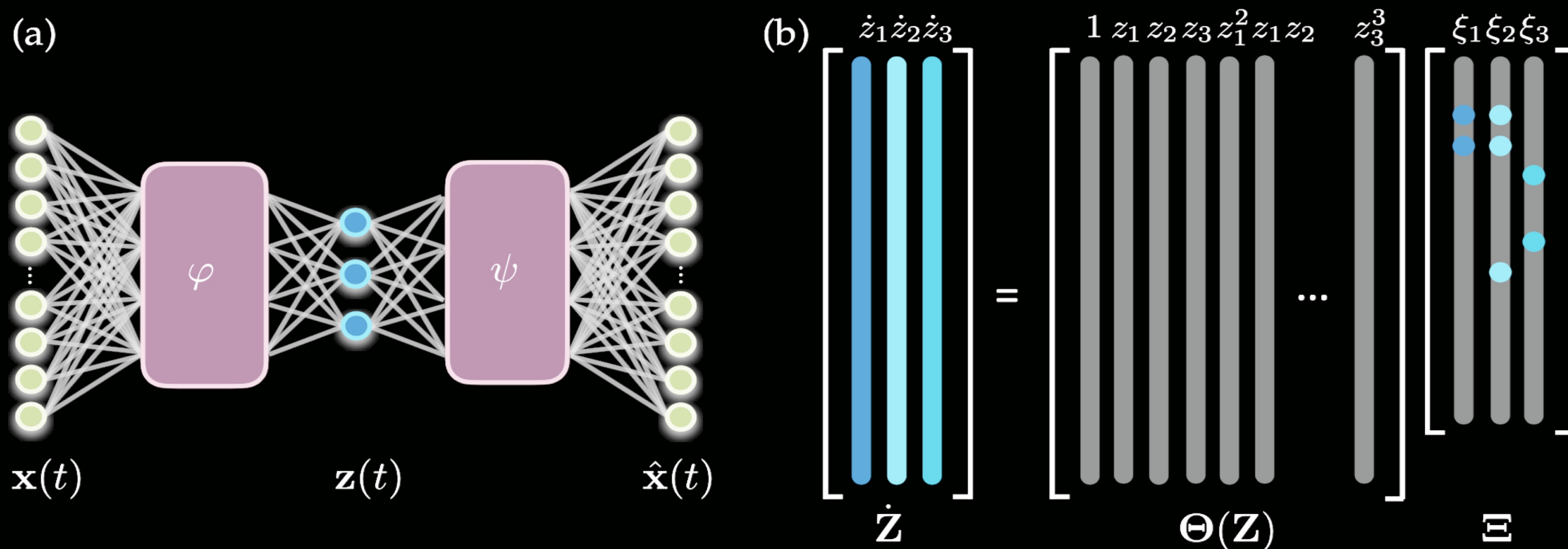
Modular, flexible and adaptive



- PDEs (Rudy et al 2017, Schaeffer et al 2017)
- Parametric ODEs/PDEs (Rudy et al 2018)
- Weak (integral) formulation (Schaeffer et al 2018, Bortz et al 2020)
- Multiscale physics (Champion et al 2019)
- Nonlinear Control (Kaheman et al 2020)
- Implicit dynamical systems (Mangan et al 2018, Lin et al 2019, Kaheman et al 2020)
- Hybrid systems (Mangan et al 2019)
- Low-data limit (Kaiser et al 2018, Xiu et al 2019)
- Course-graining SINDy (Owens et al 2020)
- Boundary value problems (Shea et al 2020)
- Stochastic systems (Clementi et al 2018)
- Dynamics with constraints (Loiseau et al 2018)
- Poincare & Flow maps & Floquet theory (Bramburger et al 2019)

Dynamics & Coordinates

Coordinates + Dynamics

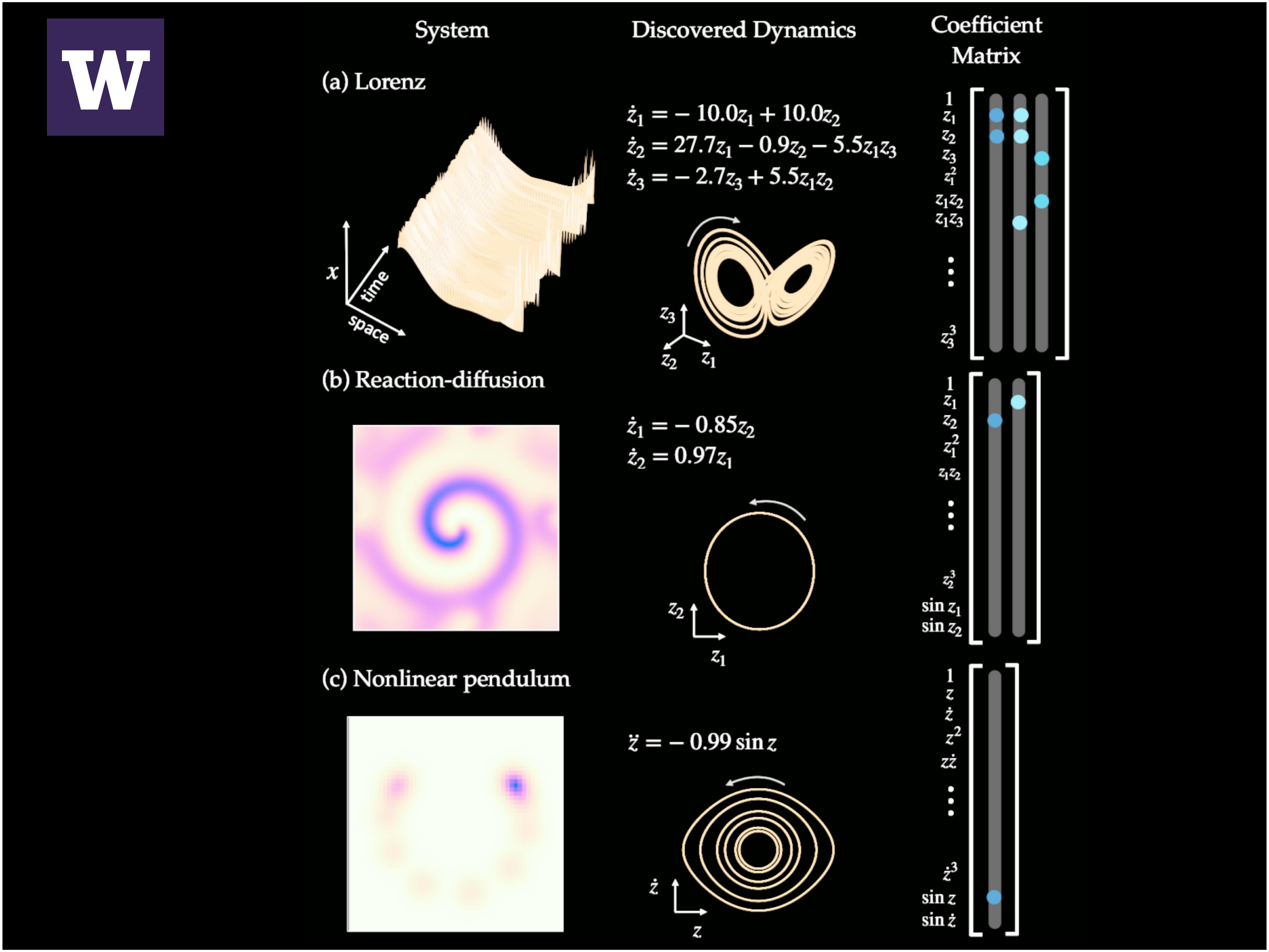


$\dot{\mathbf{z}}_i = \nabla_{\mathbf{x}}\varphi(\mathbf{x}_i)\dot{\mathbf{x}}_i \qquad \boldsymbol{\Theta}(\mathbf{z}_i^T) = \boldsymbol{\Theta}(\varphi(\mathbf{x}_i)^T)$

$$\underbrace{\|\mathbf{x} - \psi(\mathbf{z})\|_2^2}_{\text{reconstruction loss}} + \underbrace{\lambda_1 \left\| \dot{\mathbf{x}} - (\nabla_{\mathbf{z}}\psi(\mathbf{z}))\left(\boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right)\right\|_2^2}_{\text{SINDy loss in } \dot{\mathbf{x}}} + \underbrace{\lambda_2 \left\| (\nabla_{\mathbf{x}}\mathbf{z})\,\dot{\mathbf{x}} - \boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right\|_2^2}_{\text{SINDy loss in } \dot{\mathbf{z}}} + \underbrace{\lambda_3 \|\boldsymbol{\Xi}\|_1}_{\text{SINDy regularization}}$$

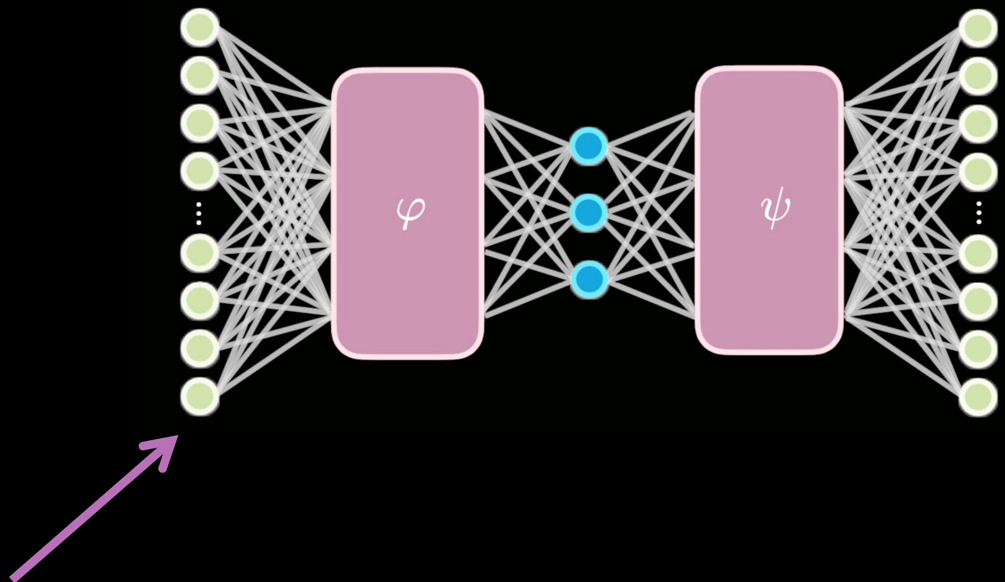
Kathleen Champion

Champion, Lusch, Kutz, Brunton, PNAS (2019)
Zheng et al, SR3 – IEEE Access (2019)

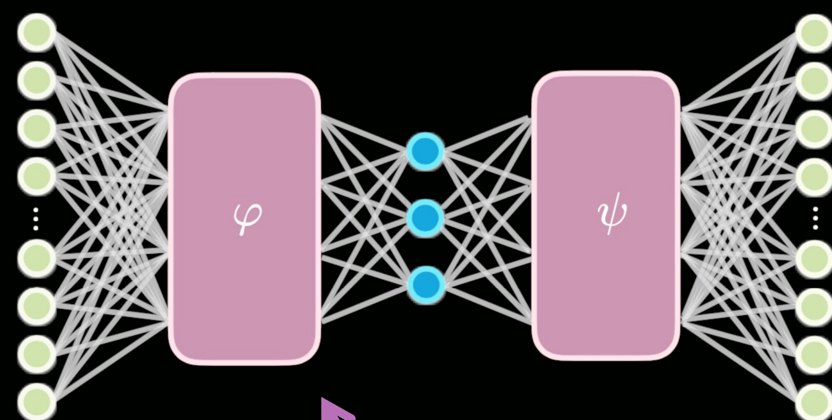


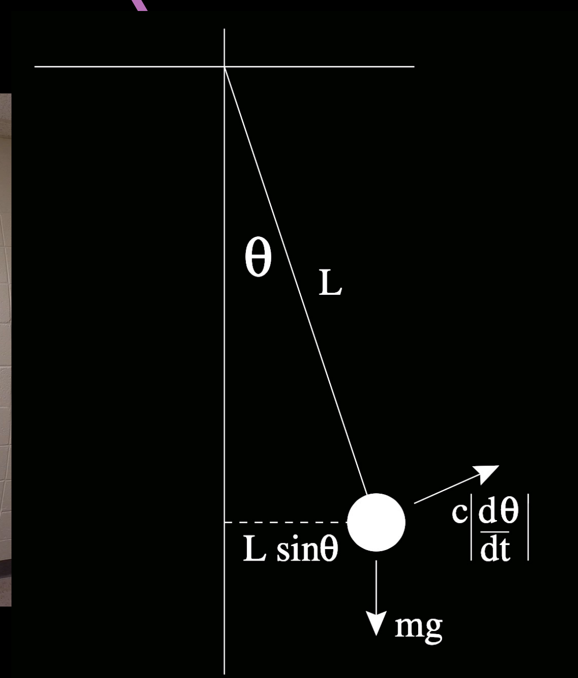
System
Discovered Dynamics
Coefficient Matrix
(a) Lorenz
x
time
space
$\dot{z}_1 = -10.0z_1 + 10.0z_2$
$\dot{z}_2 = 27.7z_1 - 0.9z_2 - 5.5z_1z_3$
$\dot{z}_3 = -2.7z_3 + 5.5z_1z_2$
z_3
z_2
z_1
1
z_1
z_2
z_3
z_1^2
z_1z_2
z_1z_3
z_3^3
(b) Reaction-diffusion
$\dot{z}_1 = -0.85z_2$
$\dot{z}_2 = 0.97z_1$
z_2
z_1
1
z_1
z_2
z_1^2
z_1z_2
z_2^3
$\sin z_1$
$\sin z_2$
(c) Nonlinear pendulum
$\ddot{z} = -0.99 \sin z$
$\dot{z}$
z
1
z
$\dot{z}$
z^2
$z\dot{z}$
$\dot{z}^3$
$\sin z$
$\sin \dot{z}$

Discovery Paradigm

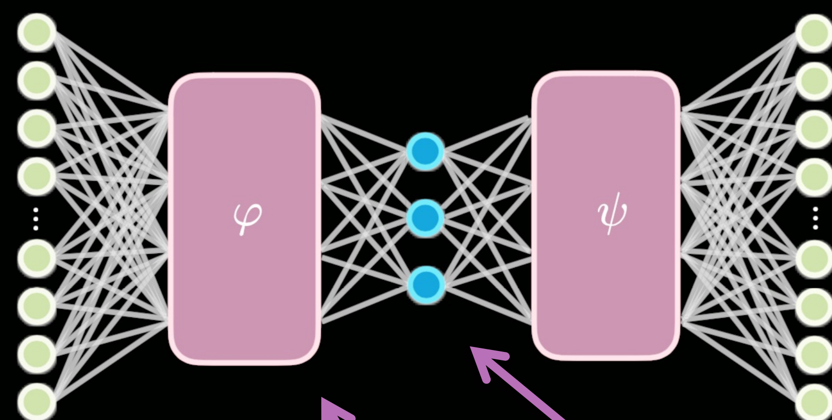


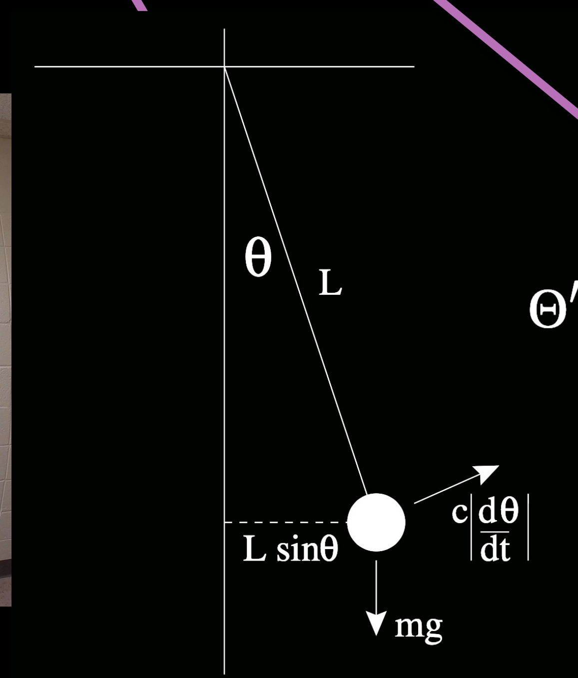
Discovery Paradigm





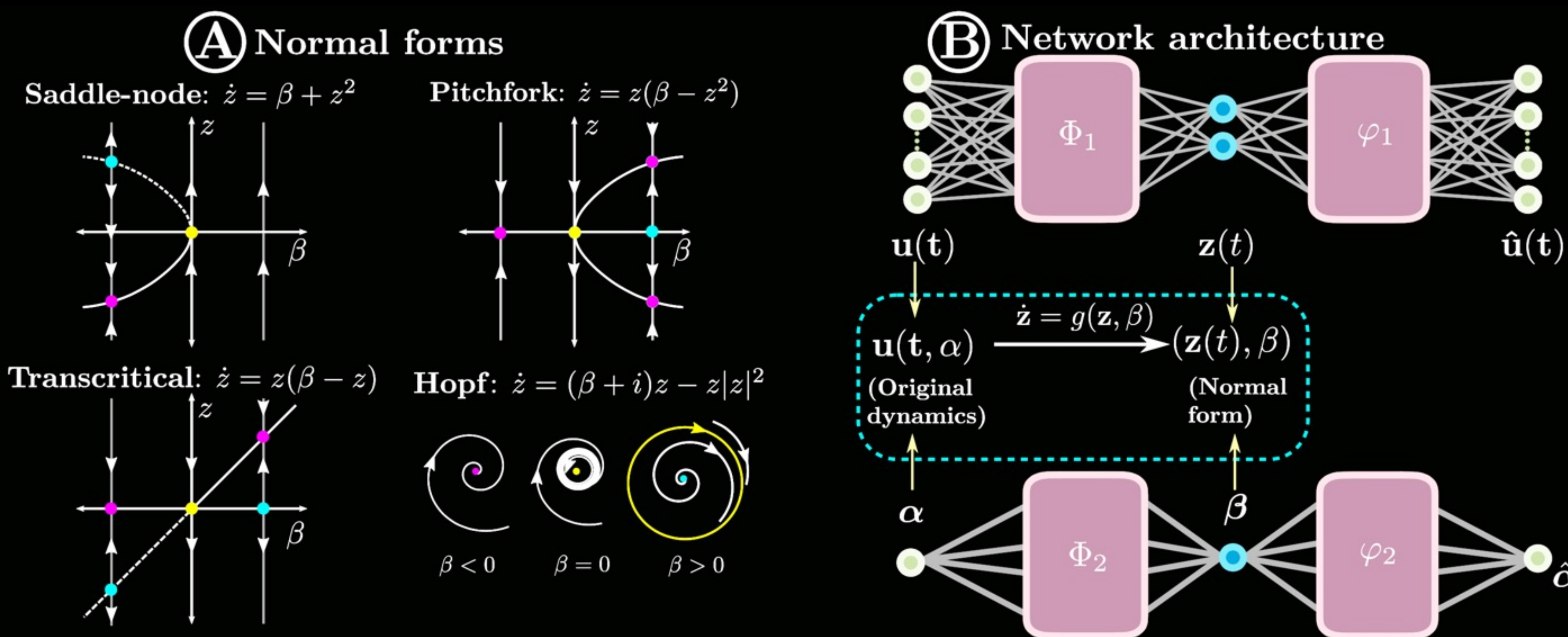
Discovery Paradigm





$$\Theta'' + \gamma\Theta' + \omega^2 \sin\Theta = 0$$

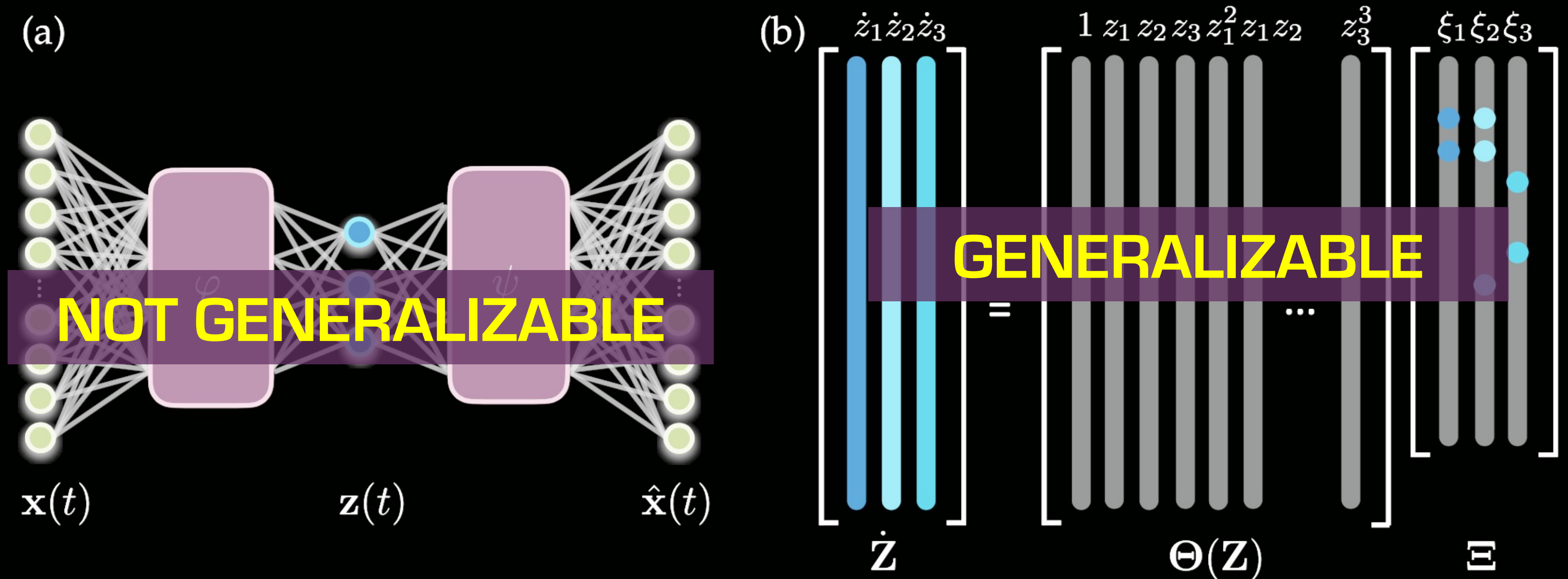
Learn Normal Forms



Manu Kalia

Kalia et al (arxiv 2021)

Coordinates + Dynamics



$$\dot{\mathbf{z}}_i = \nabla_{\mathbf{x}}\varphi(\mathbf{x}_i)\dot{\mathbf{x}}_i \qquad \boldsymbol{\Theta}(\mathbf{z}_i^T) = \boldsymbol{\Theta}(\varphi(\mathbf{x}_i)^T)$$

$$\underbrace{\|\mathbf{x}-\psi(\mathbf{z})\|_2^2}_{\text{reconstruction loss}} + \lambda_1 \underbrace{\left\|\dot{\mathbf{x}} - (\nabla_{\mathbf{z}}\psi(\mathbf{z}))\left(\boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right)\right\|_2^2}_{\text{SINDy loss in } \dot{\mathbf{x}}} + \lambda_2 \underbrace{\left\|(\nabla_{\mathbf{x}}\mathbf{z})\,\dot{\mathbf{x}} - \boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right\|_2^2}_{\text{SINDy loss in } \dot{\mathbf{z}}} + \lambda_3 \underbrace{\|\boldsymbol{\Xi}\|_1}_{\text{SINDy regularization}}$$

Time Coordinates

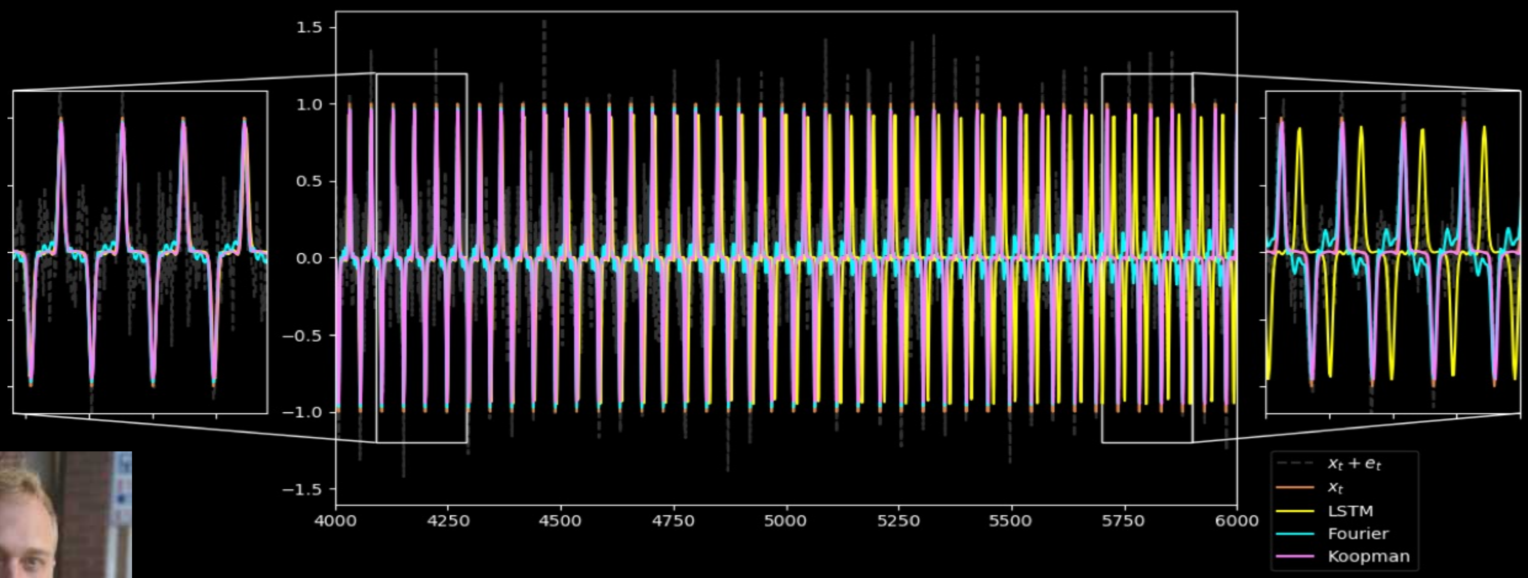
Fourier & Koopman Forecasting

Learn NN to make things sinusoidal

Koopman: $$E(\Theta, \omega) = \sum_{t=1}^{T} \left(\mathbf{x}_t - f_\Theta(\Omega(\omega t))\right)^2$$

Fourier: $$E(A, \omega) = \sum_{t=1}^{T} \left(\mathbf{x}_t - \mathbf{A}\Omega(\omega t)\right)^2$$

$$\Omega(\omega t) = \begin{bmatrix} \sin(\omega_1 t) \\ \vdots \\ \sin(\omega_N t) \\ \cos(\omega_1 t) \\ \vdots \\ \cos(\omega_N t) \end{bmatrix}$$



Henning Lange

Lange et al, arxiv (2019)

Fourier and Koopman beat NNs (powergrid data)

Algorithm	Forecast Horizon				Patterns		
	25%	50%	75%	100%	D	W	Y
Koopman Forecast	**0.19**	**0.21**	**0.19**	**0.19**	✓	✓	✓
Fourier Forecast	0.31	0.39	0.33	0.3	✓	✓	✓
LSTM	0.37	0.4	0.42	0.45	✓	×	×
GRU	0.53	0.55	0.52	0.5	✓	×	×
Echo State Network	0.67	0.73	0.76	0.73	✓	×	×
AR(1,12,24,168,4380,8760)	0.75	0.95	1.07	1.13	✓	✓	✓
CW-RNN (data clocks)	1.1	1.14	1.14	1.15	(✓)	×	×
CW-RNN	1.05	1.08	1.08	1.09	(✓)	×	×
AutoARIMA	0.83	1.11	1.18	1.26	×	×	×
Fourier Neural Networks	1.1	1.15	1.21	1.21	✓	×	×

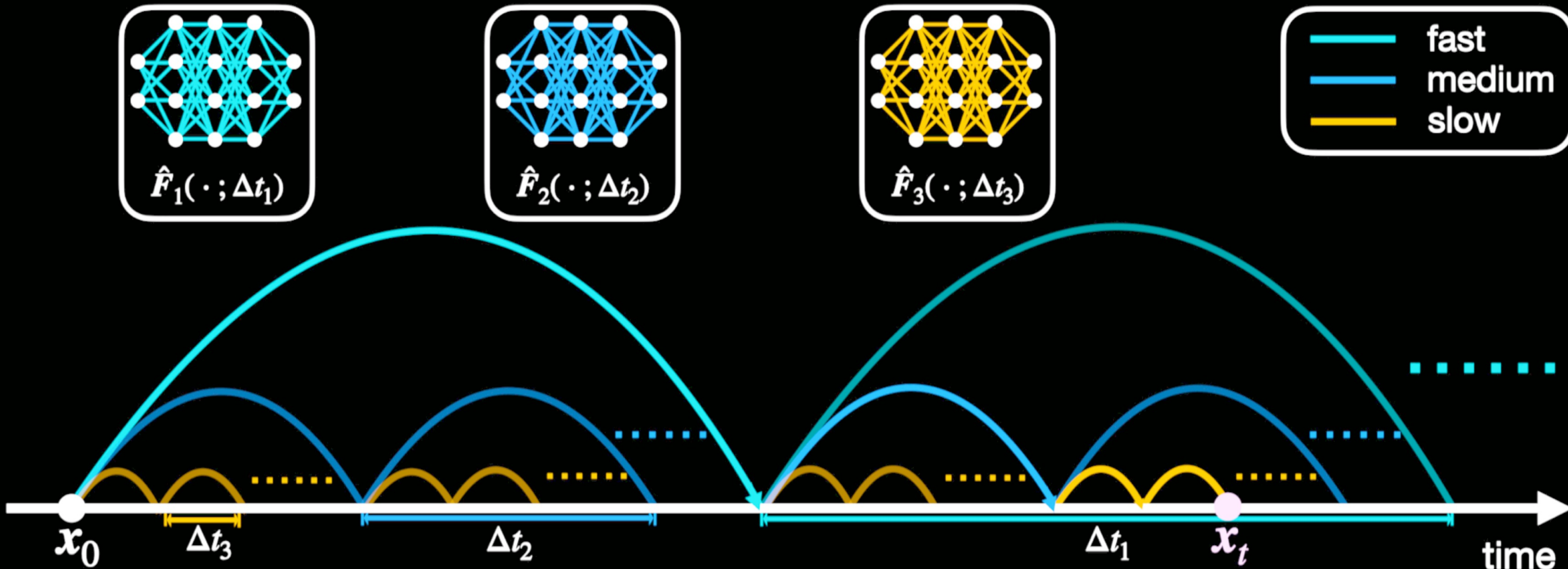
Improved ROMs

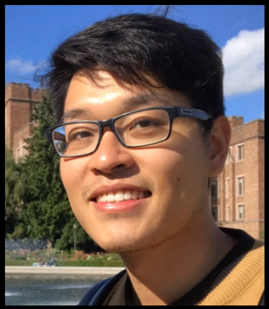
Learn Hierarchical Flow Maps

Xiu et al JCP (2019)

Flow maps & numerical schemes

$$x(t+\Delta t) = \boldsymbol{F}(x(t), \Delta t)$$



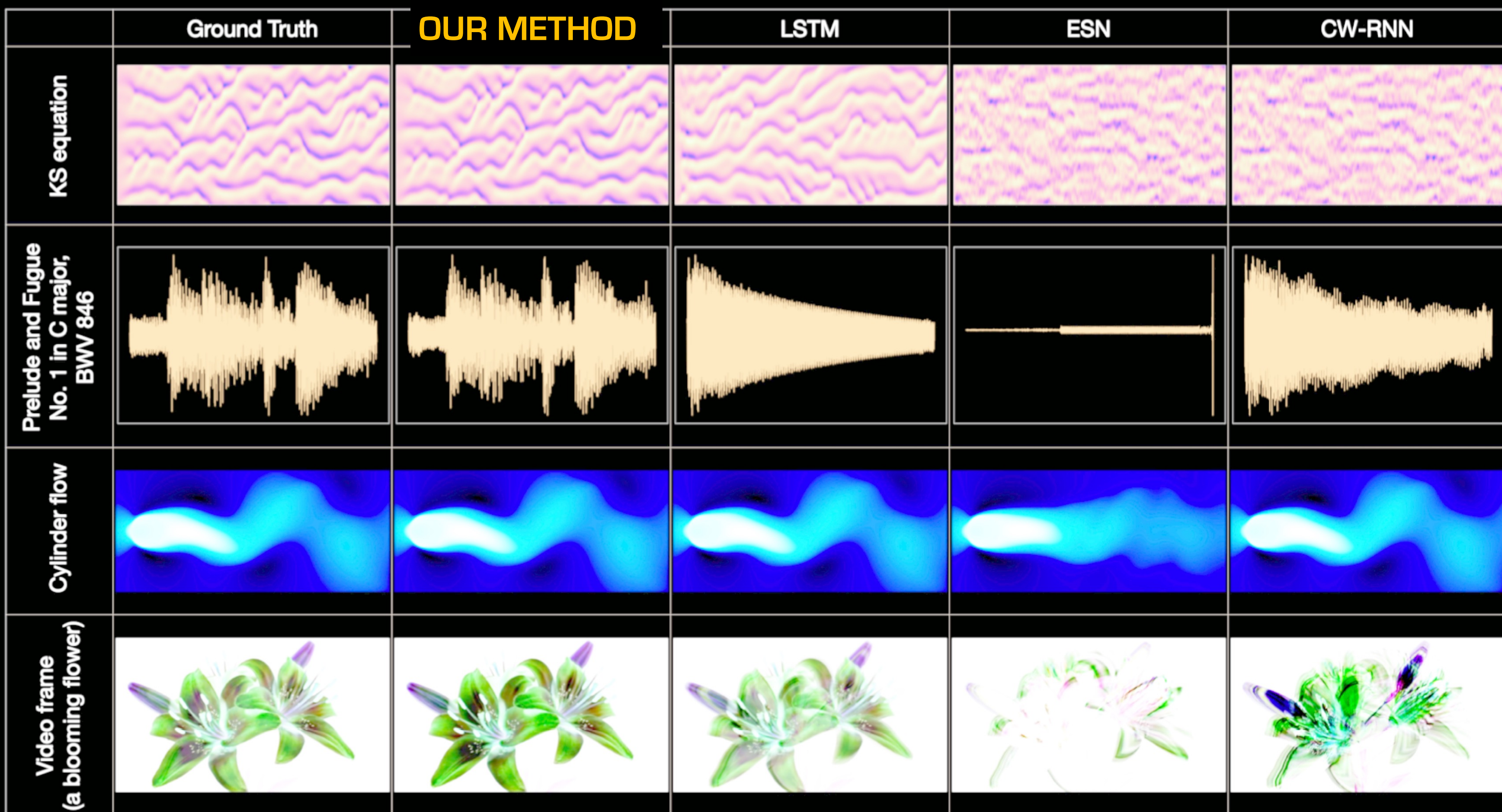


SINDy on Poincare/Flow Maps: Bramburger & Kutz PRE (2020)

Yuying Liu

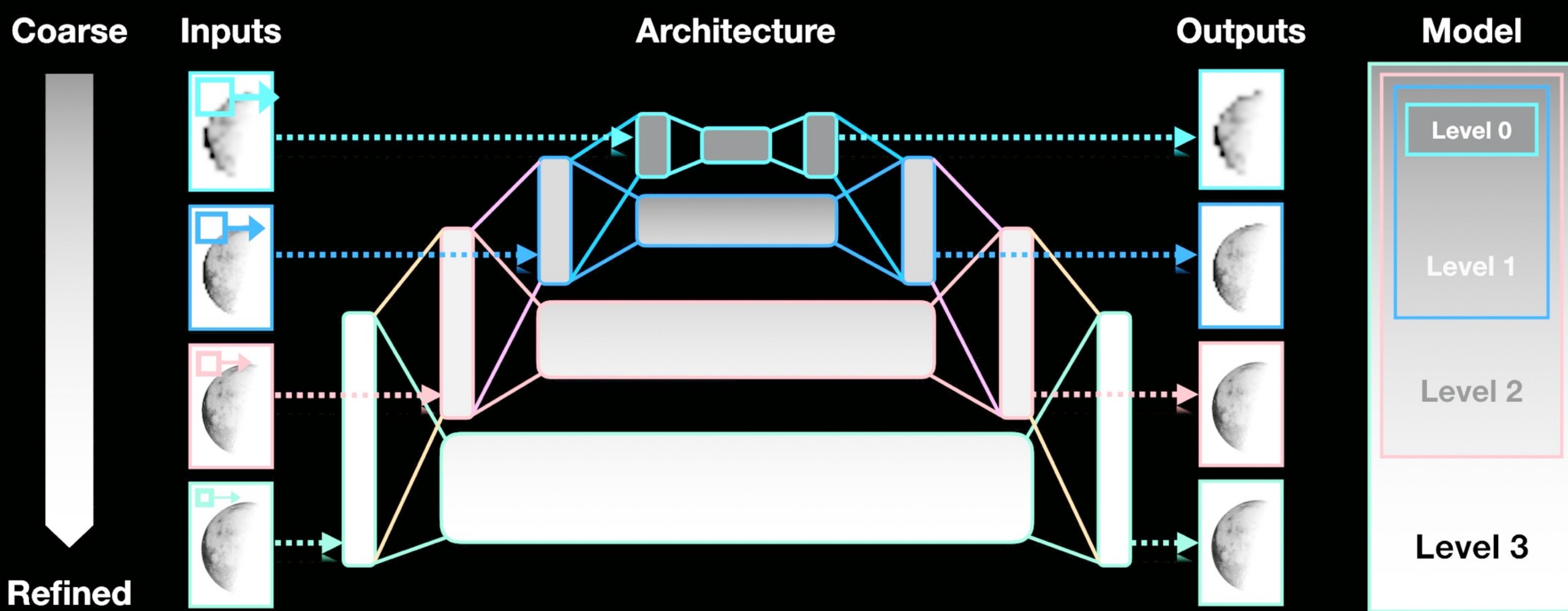
Liu et al, arxiv (2020)

Ideal for Numerical Stiffness

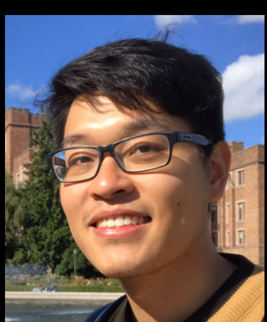


Multiscale Physics

Multiscale Physics



Exploits transfer learning & multi-grid methods

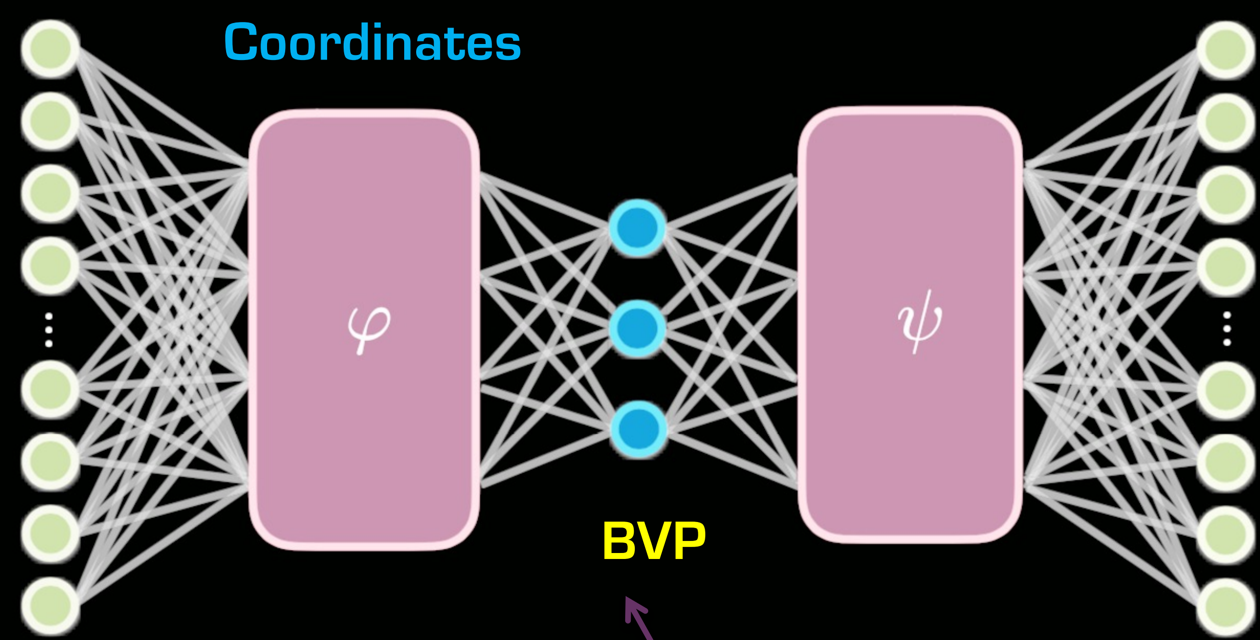


Yuying Liu

Liu et al, arxiv (2020)

Boundary Value Problems

Coordinates & BVPs



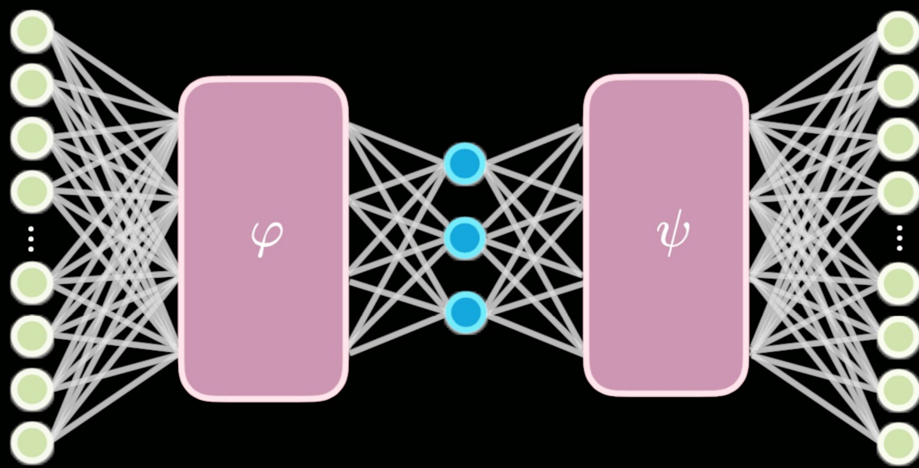
Input data
(measurements)

- SINDy: Discover ODEs or PDEs
- Green's Functions

Dan Shea

Shea et al. arxiv (2020)

Conclusion: Parsimony is the Physics Regularizer



Dynamics/BVP

State-space

Parameters

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, t, \Theta, \Omega)$$

Dynamics

Stochastic effects

Measurement

$$\mathbf{y}(t_k) = h(t_k, \mathbf{x}(t_k), \Xi)$$

Measurement model

Measurement noise