# *Coordinates, Governing Equations and Limits of Model Discovery*

**_J. Nathan Kutz_**

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

**IPAM at UCLA, October 16, 2019**

# *Coordinates + Dynamics*

Kevrekidis, Coifman et al
Noe et al
Mezic et al
Etc etc etc

# *Featuring*

**Kathleen Champion**

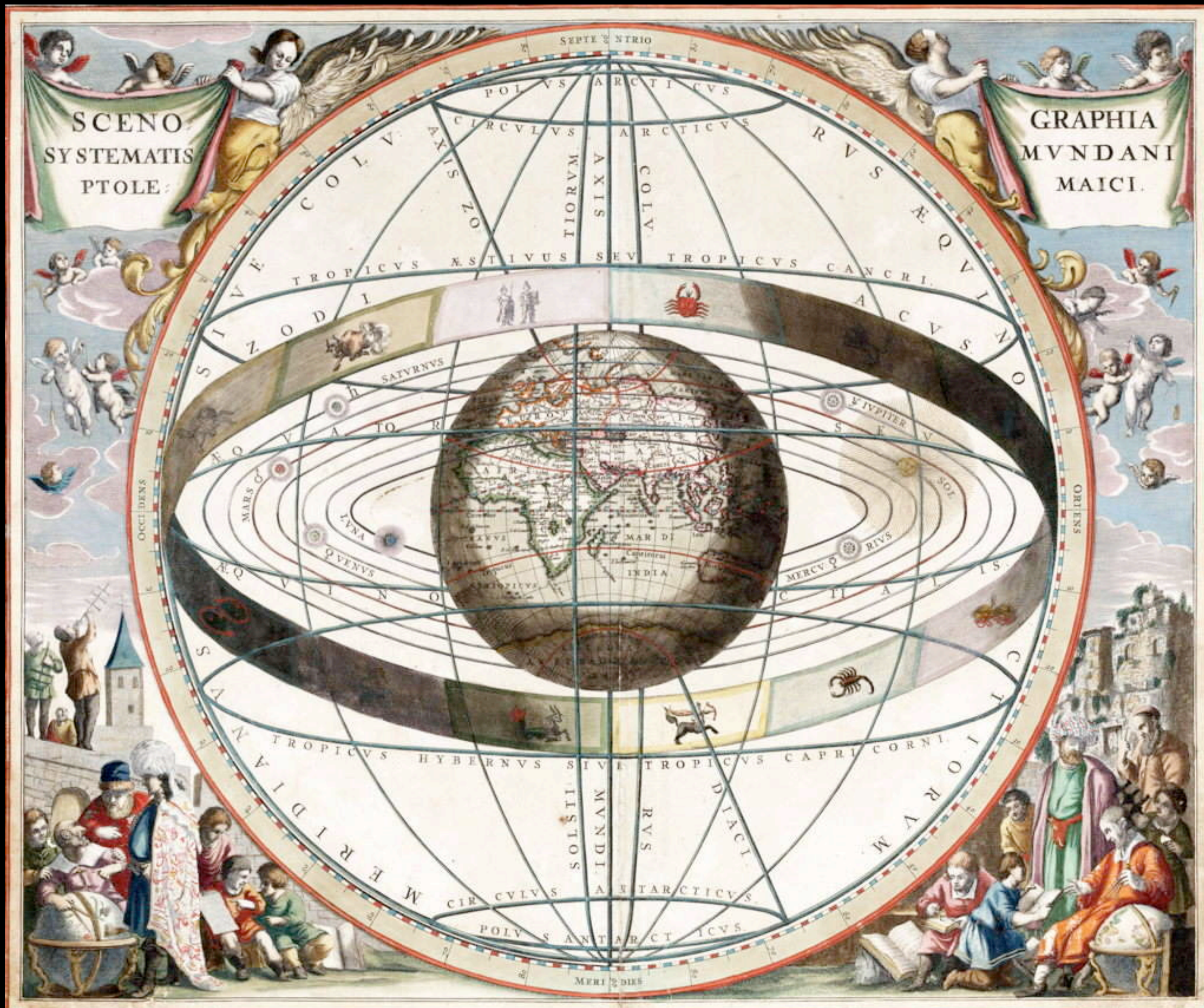
**Brian deSilva**

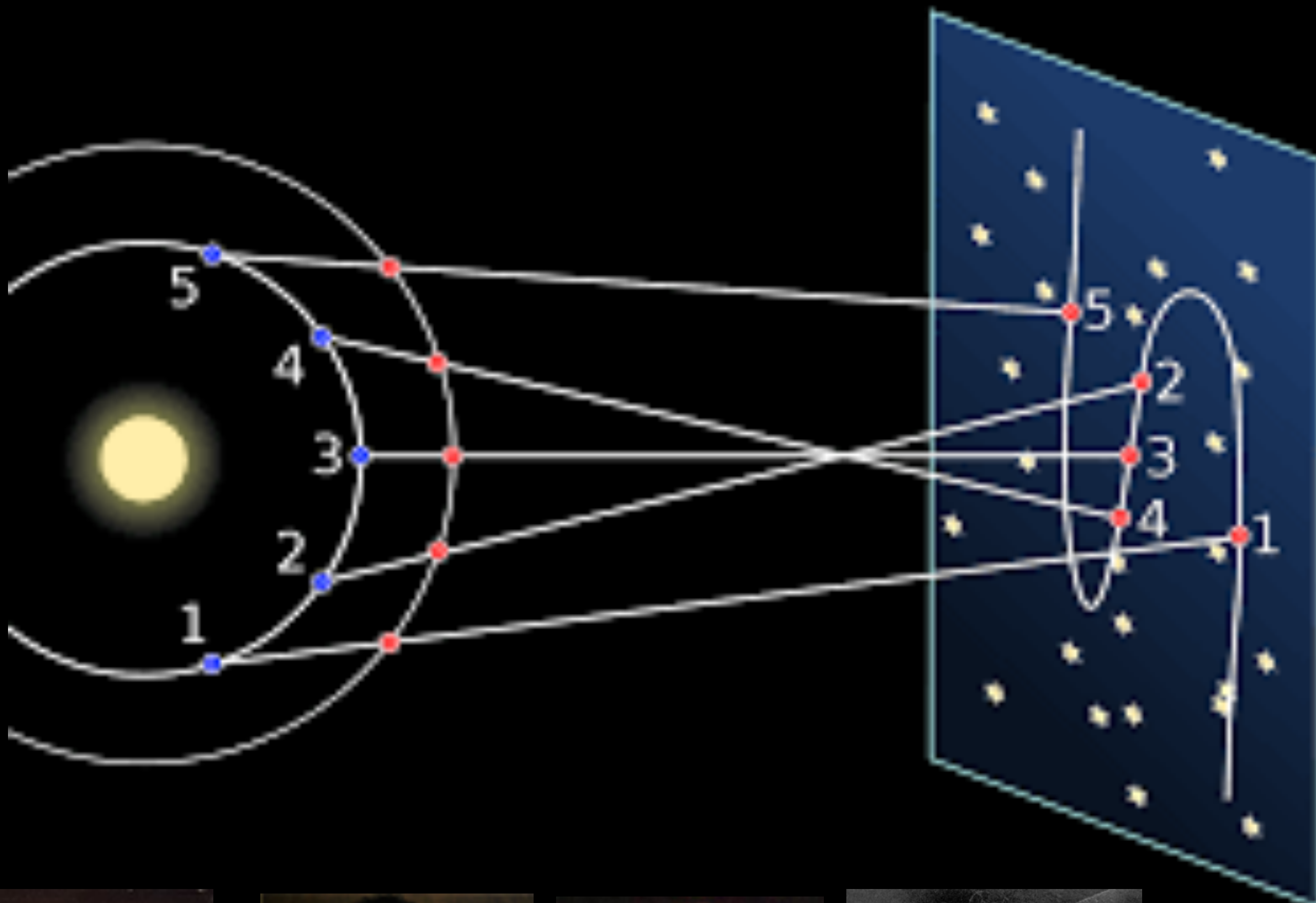
**Bethany Lusch**

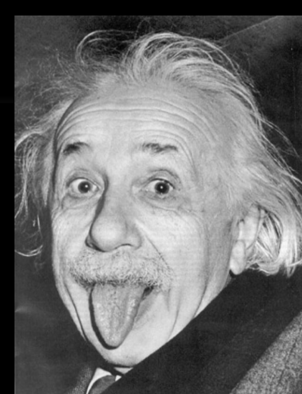
**Steve Brunton**


Saturn & Mars
(C) 2015.6 Tunc Tezel

# Doctrine of the Perfect Circle




5
4
3
2
1
5
2
3
4
1



# Reduced Order Models

$$\frac{d\mathbf{u}(t)}{dt} = L\mathbf{u}(t) + N(\mathbf{u}(t))$$

$$\mathbf{u}(t) \approx \mathbf{\Phi}_r \mathbf{a}(t)$$

$$\frac{d\mathbf{a}(t)}{dt} = \mathbf{\Phi}_r^T L \mathbf{\Phi}_r \mathbf{a}(t) + \mathbf{\Phi}_r^T N(\mathbf{\Phi}_r \mathbf{a}(t))$$

# Question #1

## What is the nature of your data?

- quality
- quantity
- observability
- extrapolation vs interpolation

# Model Discovery

## Finding governing equations

# Mathematical Framework

**Dynamics**

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

State-space

Parameters

Dynamics

Stochastic effects

**Measurement**

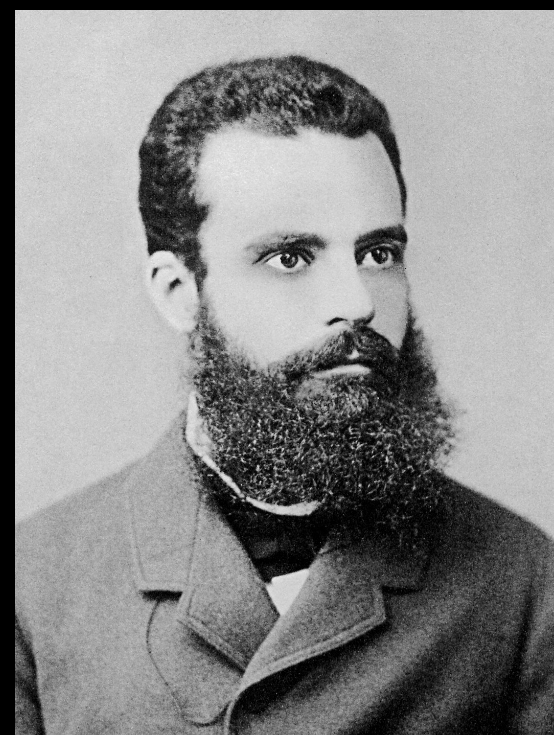
$$\mathbf{y}(t_k) = h(t_k, \mathbf{x}(t_k), \Xi)$$

Measurement model

Measurement noise

# Interpretability

William of Occam



Vilfredo Pareto

# Parsimony

**The Ultimate Physics Regularization**

# of terms
# of dimensions

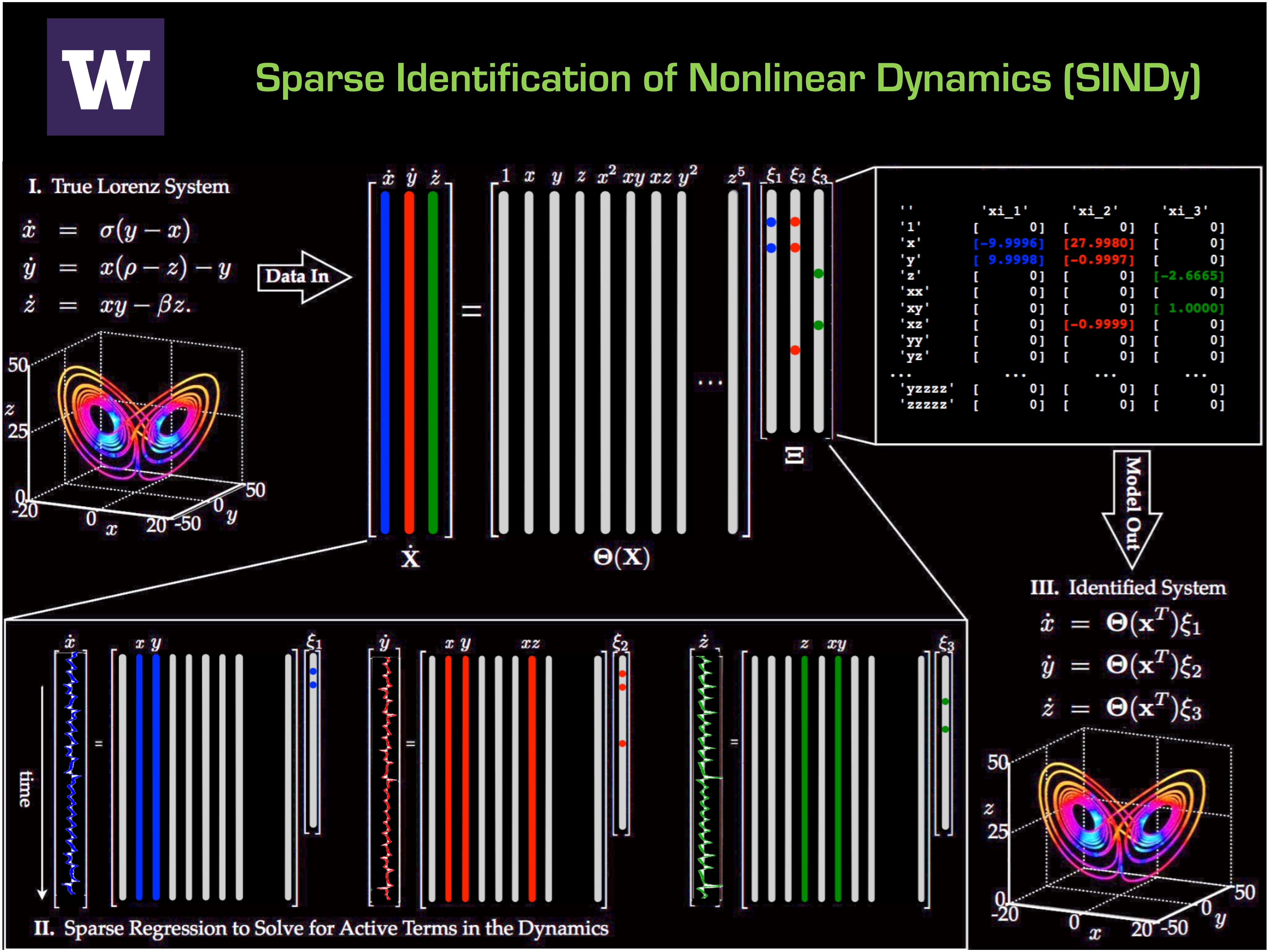
# What Could the Right Side Be?

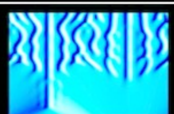
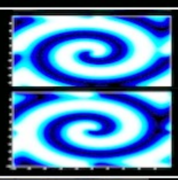
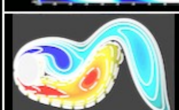
**Limited by your imagination**

$$\Theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & & | & | & | & | & \\ 1 & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \cdots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \sin(2\mathbf{X}) & \cos(2\mathbf{X}) & \cdots \\ | & | & | & | & & | & | & | & | & \end{bmatrix}$$

**2nd degree polynomials**

$$\mathbf{X}^{P_2} = \begin{bmatrix} x_1^2(t_1) & x_1(t_1)x_2(t_1) & \cdots & x_2^2(t_1) & x_2(t_1)x_3(t_1) & \cdots & x_n^2(t_1) \\ x_1^2(t_2) & x_1(t_2)x_2(t_2) & \cdots & x_2^2(t_2) & x_2(t_2)x_3(t_2) & \cdots & x_n^2(t_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_1^2(t_m) & x_1(t_m)x_2(t_m) & \cdots & x_2^2(t_m) & x_2(t_m)x_3(t_m) & \cdots & x_n^2(t_m) \end{bmatrix}$$


Sparse Identification of Nonlinear Dynamics (SINDy)
I. True Lorenz System
$\dot{x} = \sigma(y - x)$
$\dot{y} = x(\rho - z) - y$
$\dot{z} = xy - \beta z.$
Data In
$\dot{x}$ $\dot{y}$ $\dot{z}$
$1$ $x$ $y$ $z$ $x^2$ $xy$ $xz$ $y^2$ ... $z^5$
$\xi_1$ $\xi_2$ $\xi_3$
$\dot{\mathbf{X}}$ $=$ $\Theta(\mathbf{X})$ $\Xi$
'' 'xi_1' 'xi_2' 'xi_3'
'1' [ 0] [ 0] [ 0]
'x' [-9.9996] [27.9980] [ 0]
'y' [ 9.9998] [-0.9997] [ 0]
'z' [ 0] [ 0] [-2.6665]
'xx' [ 0] [ 0] [ 0]
'xy' [ 0] [ 0] [ 1.0000]
'xz' [ 0] [-0.9999] [ 0]
'yy' [ 0] [ 0] [ 0]
'yz' [ 0] [ 0] [ 0]
... ... ... ...
'yzzzz' [ 0] [ 0] [ 0]
'zzzzz' [ 0] [ 0] [ 0]
Model Out
III. Identified System
$\dot{x} = \Theta(\mathbf{x}^T)\xi_1$
$\dot{y} = \Theta(\mathbf{x}^T)\xi_2$
$\dot{z} = \Theta(\mathbf{x}^T)\xi_3$
time
II. Sparse Regression to Solve for Active Terms in the Dynamics

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

# Discrepancy Models

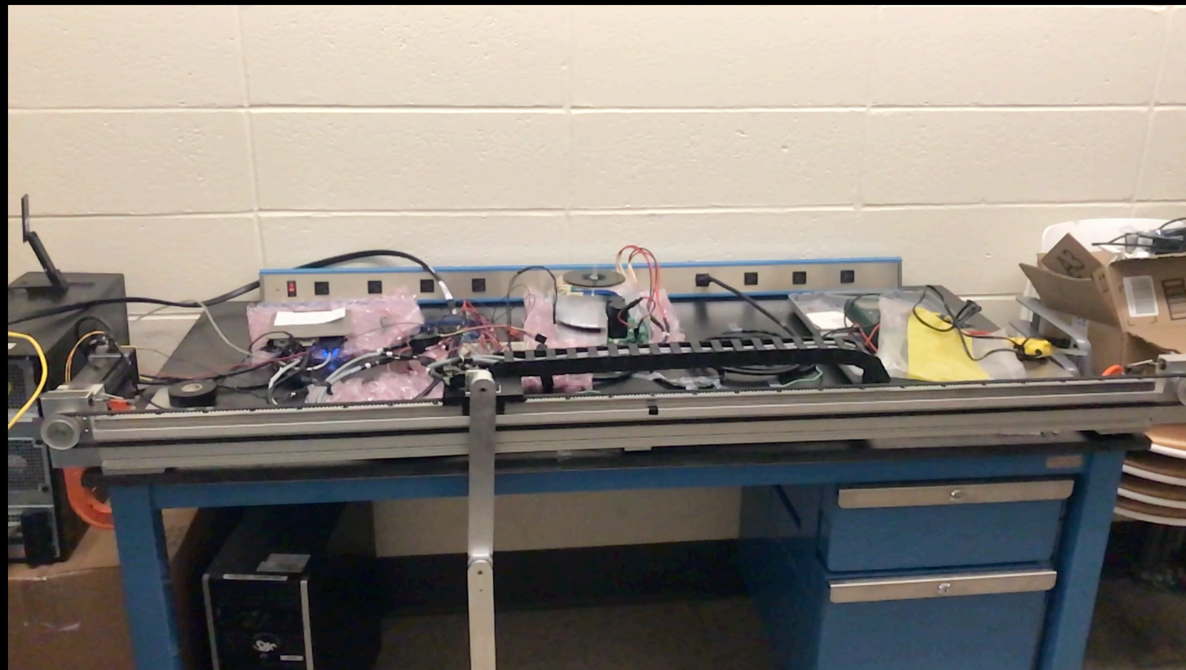
I'm not dumb

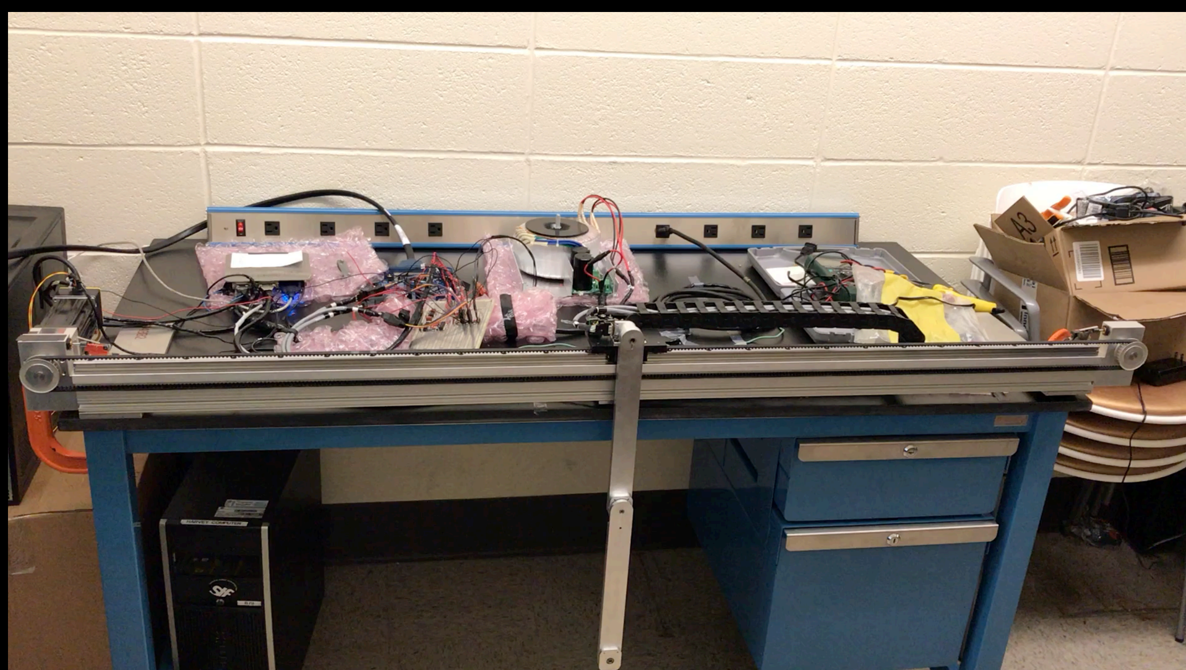
Instead of model discovery from scratch…

…we often start with partial knowledge of the physics

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

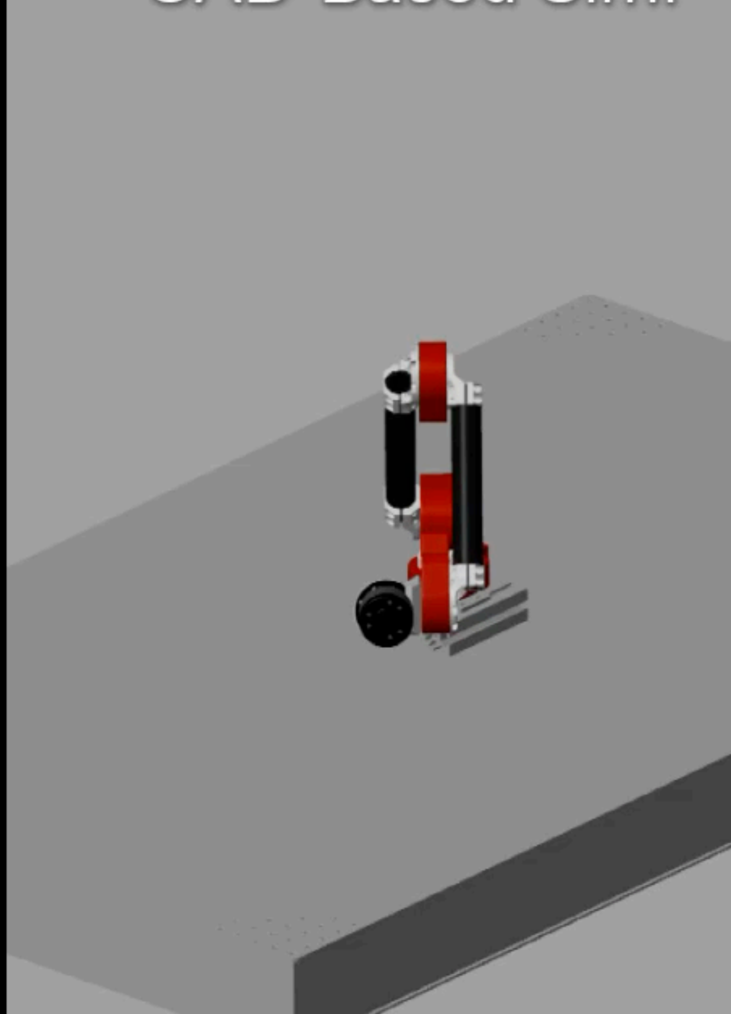
$$\underbrace{\boxed{\frac{d}{dt}\mathbf{x} = \mathbf{f}(\mathbf{x})}}_{\text{Imperfect model}} + \underbrace{\boxed{\boldsymbol{\delta}\mathbf{g}(\mathbf{x})}}_{\text{Discrepancy}}$$



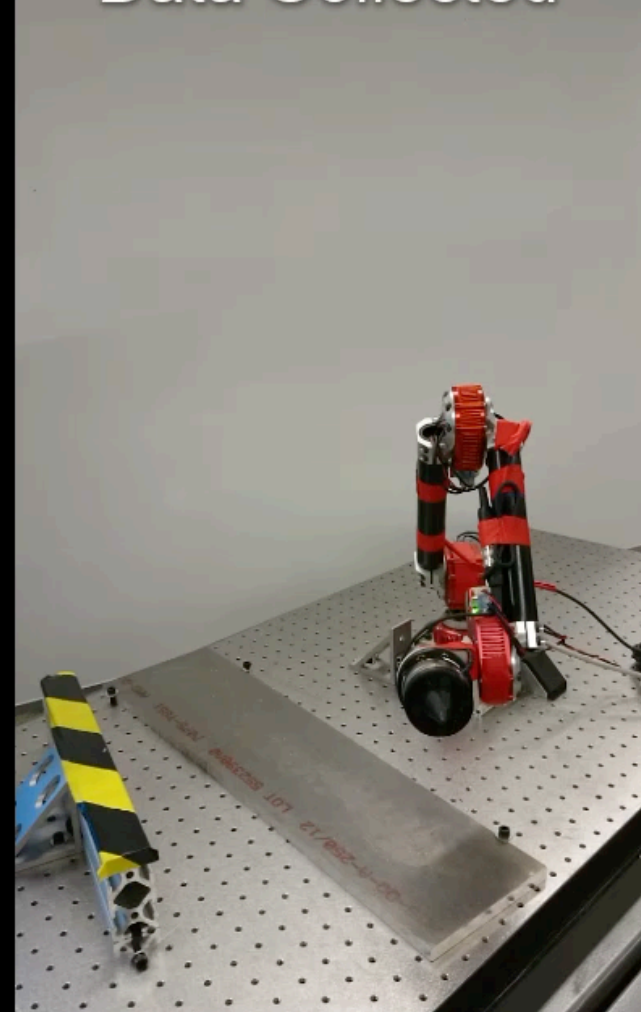


# Digital Twins

CAD-Based Sim.



Data Collected



# KEY CHALLENGES

- Limited measurements & data
- Noise
- Multi-scale physics
- Latent variables
- Parametric dependencies
- Stochastic systems

# Manifolds and Embeddings

## Observables & Coordinates

# Bernard Koopman 1931

**Definition: Koopman Operator** (Koopman 1931): *For a dynamical system*

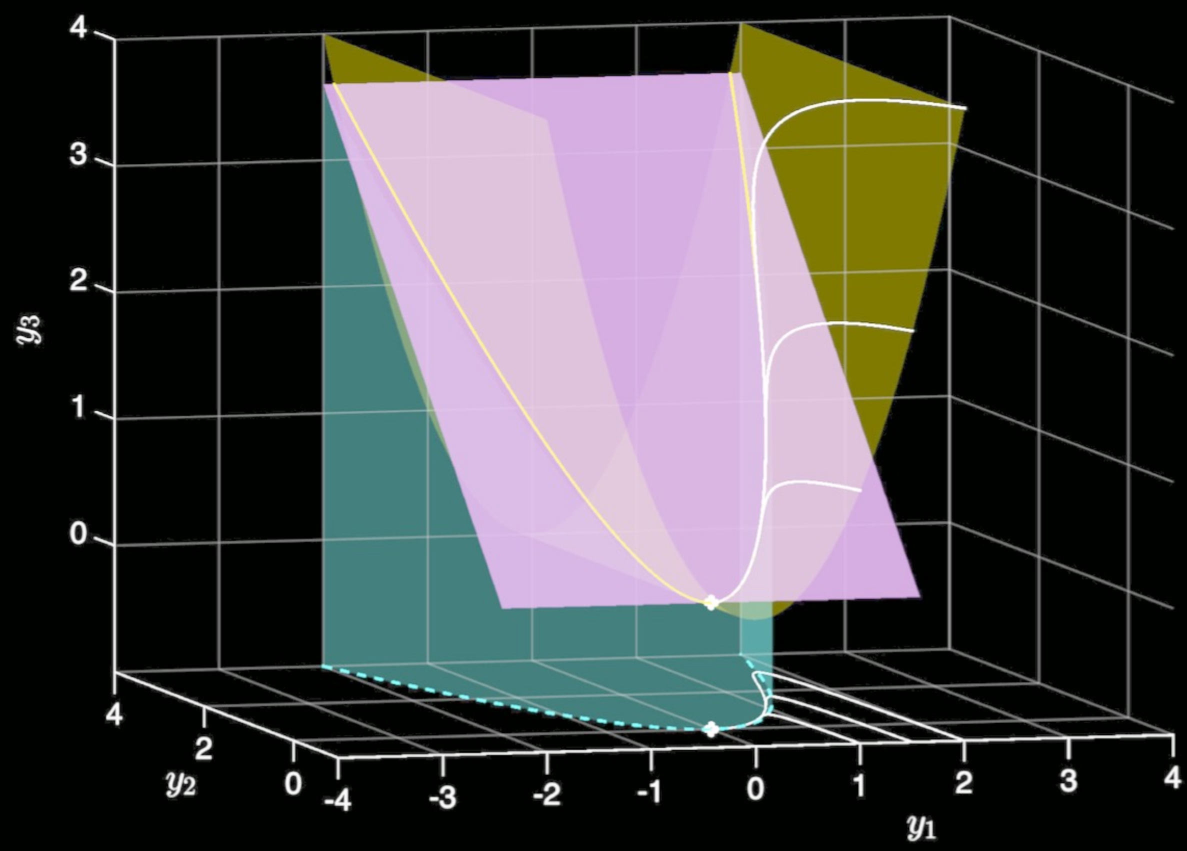
$$\frac{d\mathbf{x}}{dt} = \mathbf{N}(\mathbf{x}),$$

*where* $\mathbf{x} \in \mathbb{R}^n$ *is in a state space* $\mathbf{x} \in \mathcal{M}$. *The Koopman operator* $\mathcal{K}$ *acts on a set of scalar observable variables* $g_j$ *which comprise the vector* $\mathbf{g} : \mathcal{M} \to \mathbb{C}$ *so that*

$$\mathcal{K}\mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{N}(\mathbf{x})) .$$

# Koopman Invariant Subspaces

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



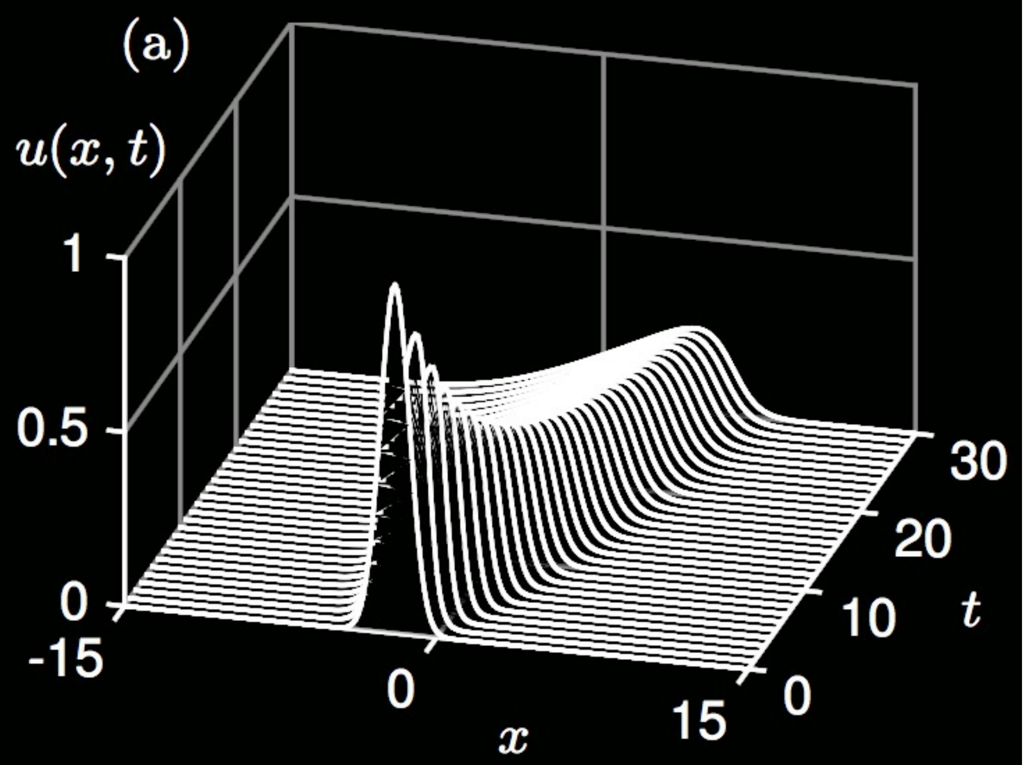
# Burgers' Equation

$$u_t + uu_x - \epsilon u_{xx} = 0 \qquad \epsilon > 0,\ x \in [-\infty, \infty]$$

**Cole-Hopf**

$$u = -2\epsilon v_x / v$$

$$v_t = \epsilon v_{xx}$$



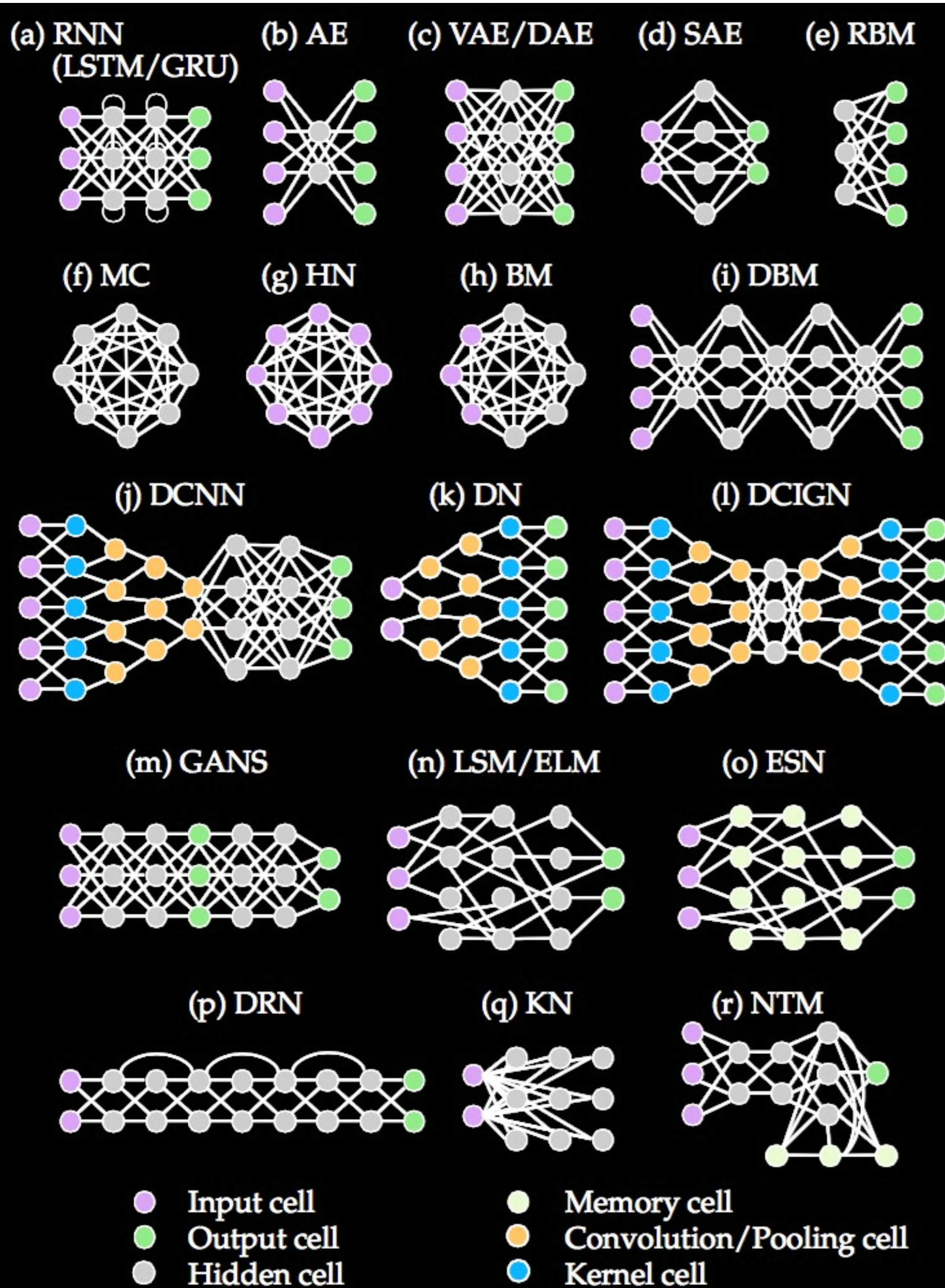

Kutz, Proctor & Brunton, Complexity (2018)

# Neural Nets

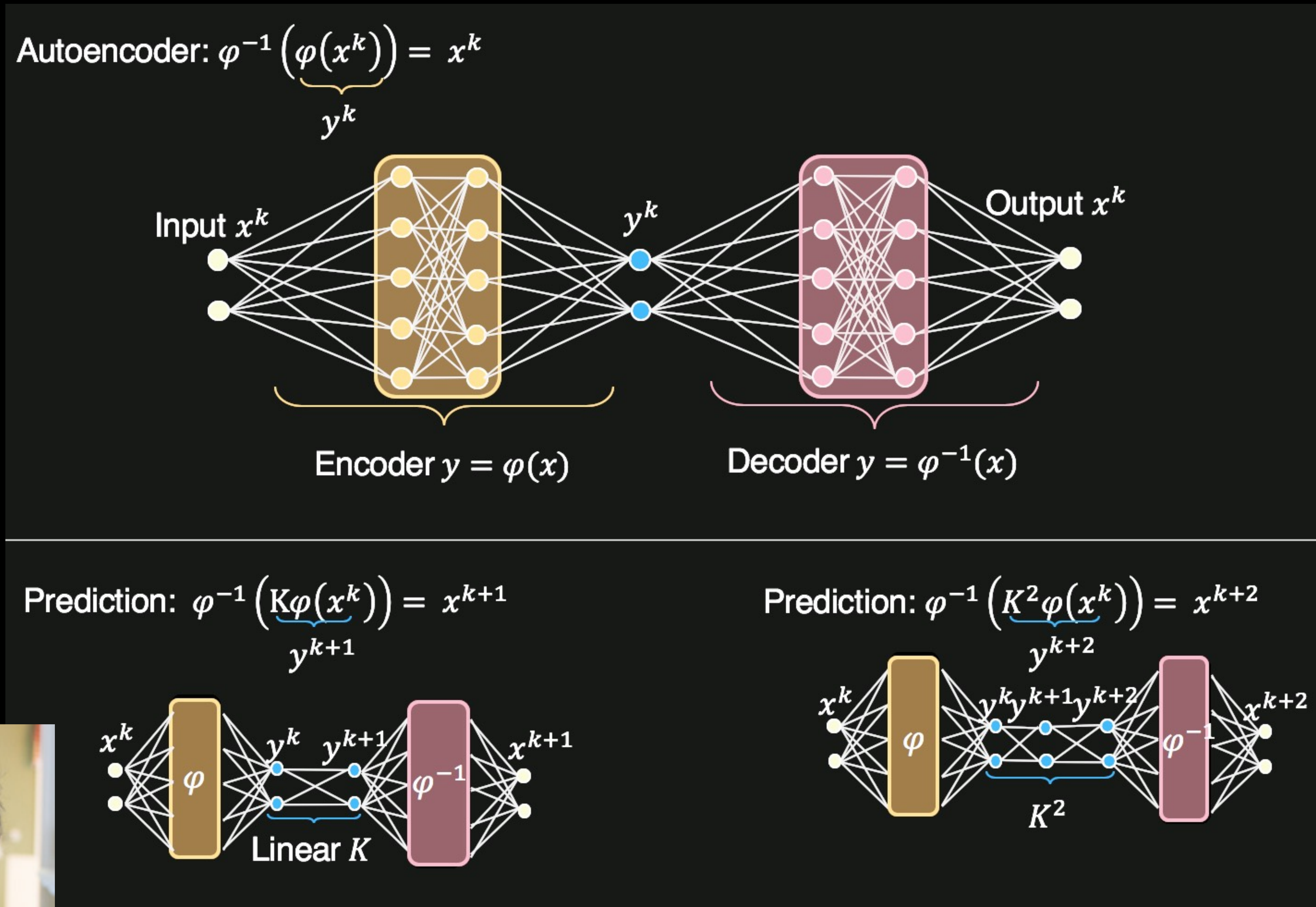
"Supervised learning is a high-dimensional interpolation problem."

S. Mallat, PRSA (2016)

# NN Zoo


(a) RNN (LSTM/GRU)
(b) AE
(c) VAE/DAE
(d) SAE
(e) RBM
(f) MC
(g) HN
(h) BM
(i) DBM
(j) DCNN
(k) DN
(l) DCIGN
(m) GANS
(n) LSM/ELM
(o) ESN
(p) DRN
(q) KN
(r) NTM
Input cell
Output cell
Hidden cell
Memory cell
Convolution/Pooling cell
Kernel cell

# NNs for Koopman Embedding




**Bethany Lusch**

Lusch et al. Nat. Comm (2018)

**Failure!**

*(obviously)*

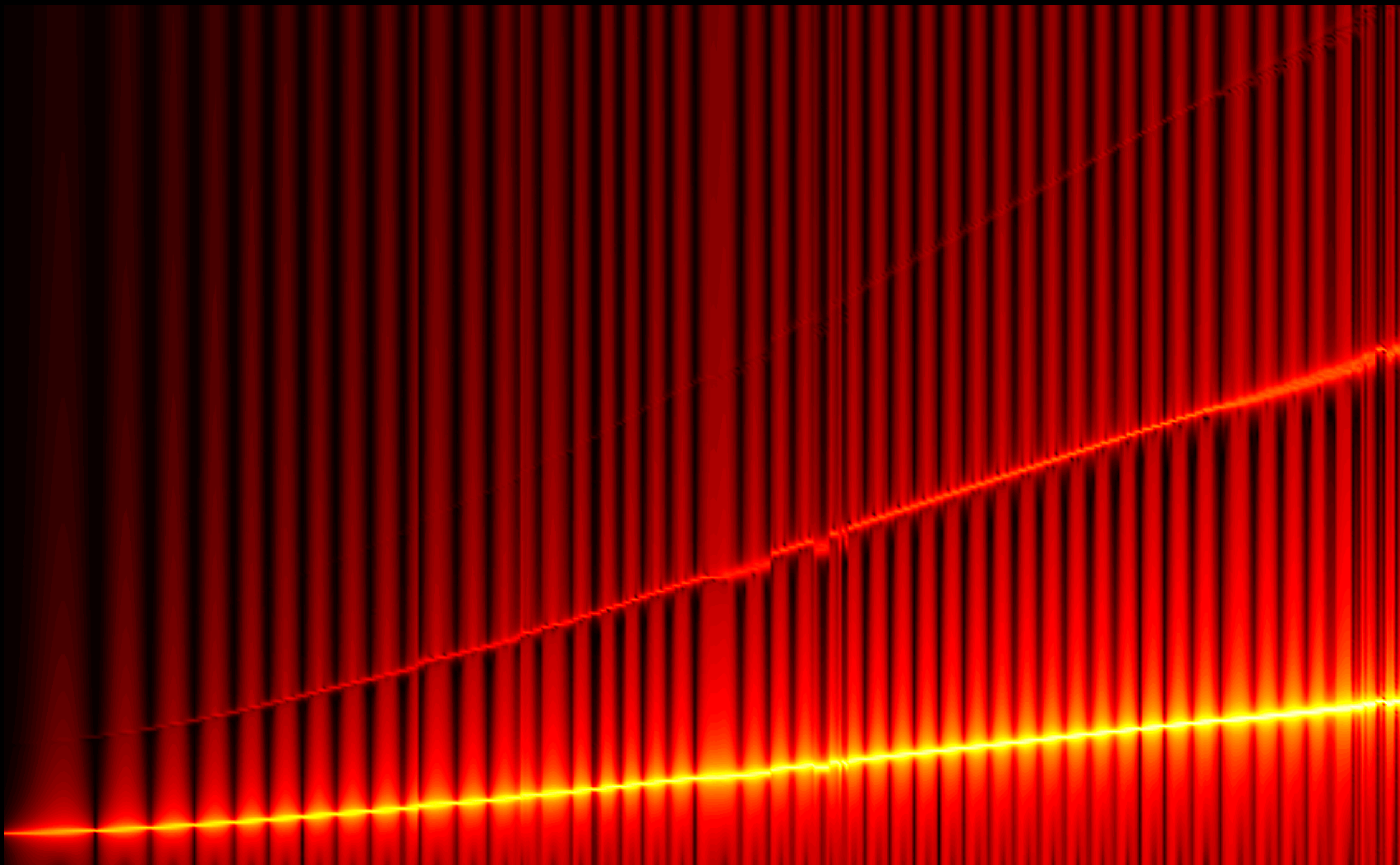
# Duffing Oscillator

$$\frac{\partial^2 u}{\partial t^2} + u + \epsilon u^3 = 0$$

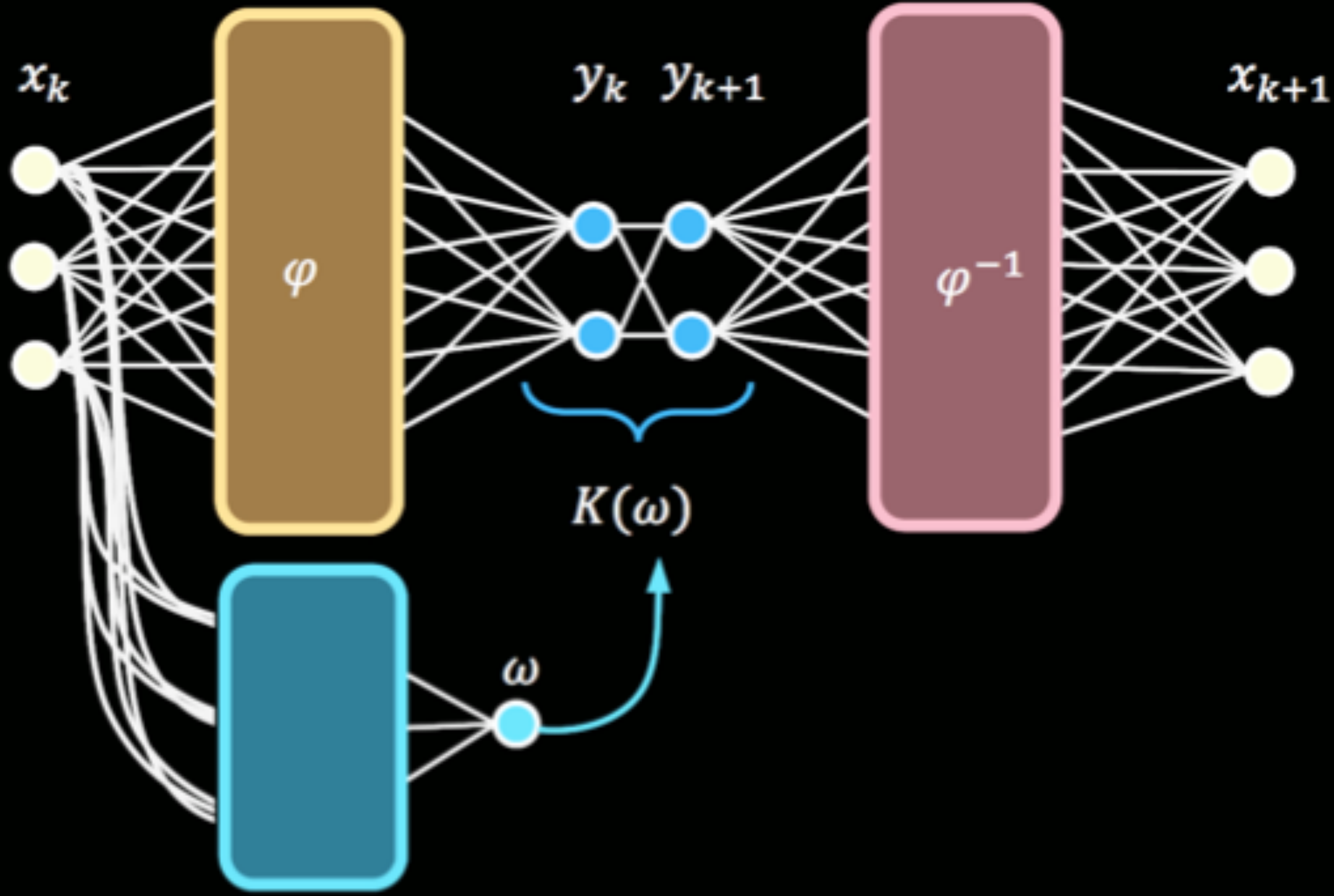
**Nonlinearity: Shifts Frequencies + Generates Harmonics**

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

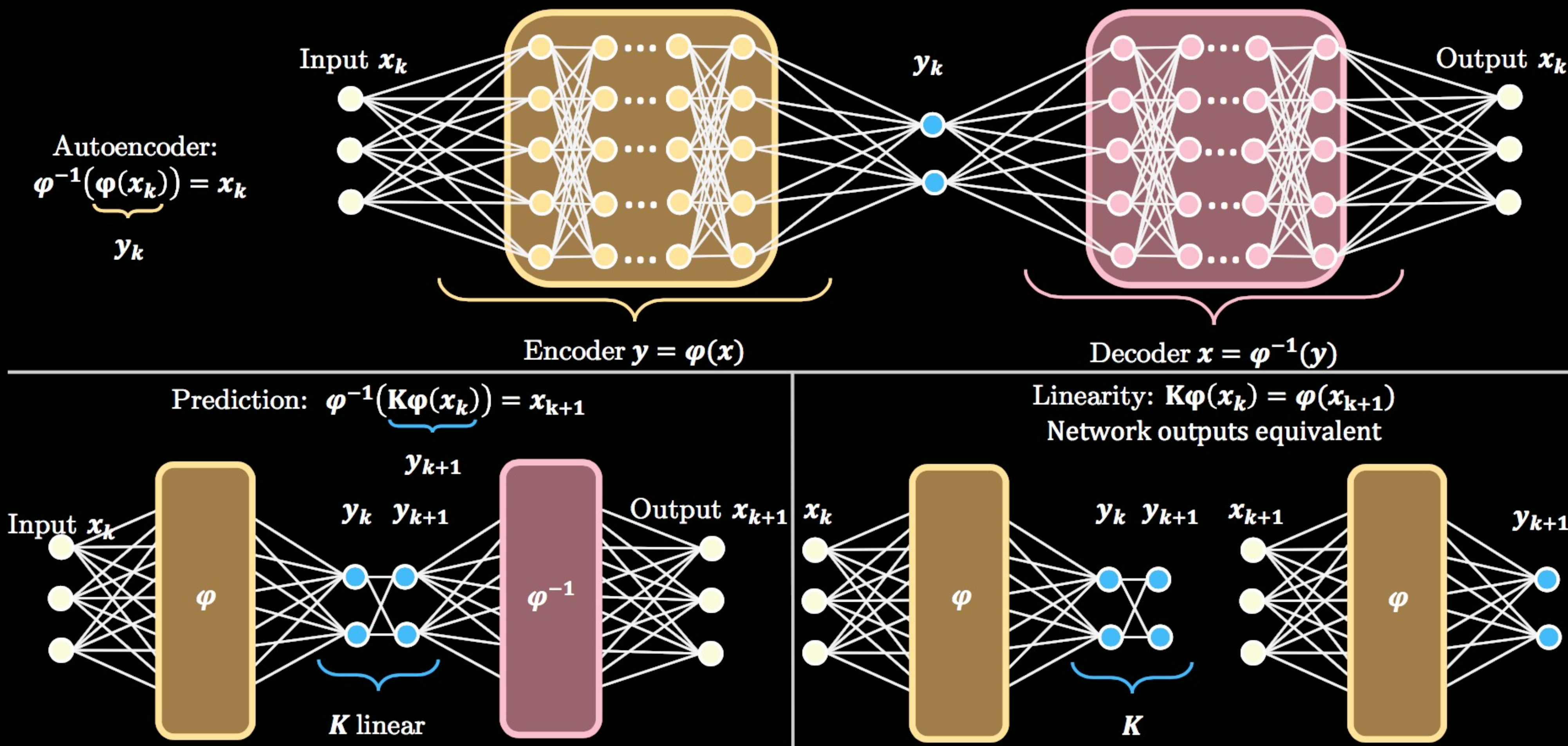
# Spectrogram



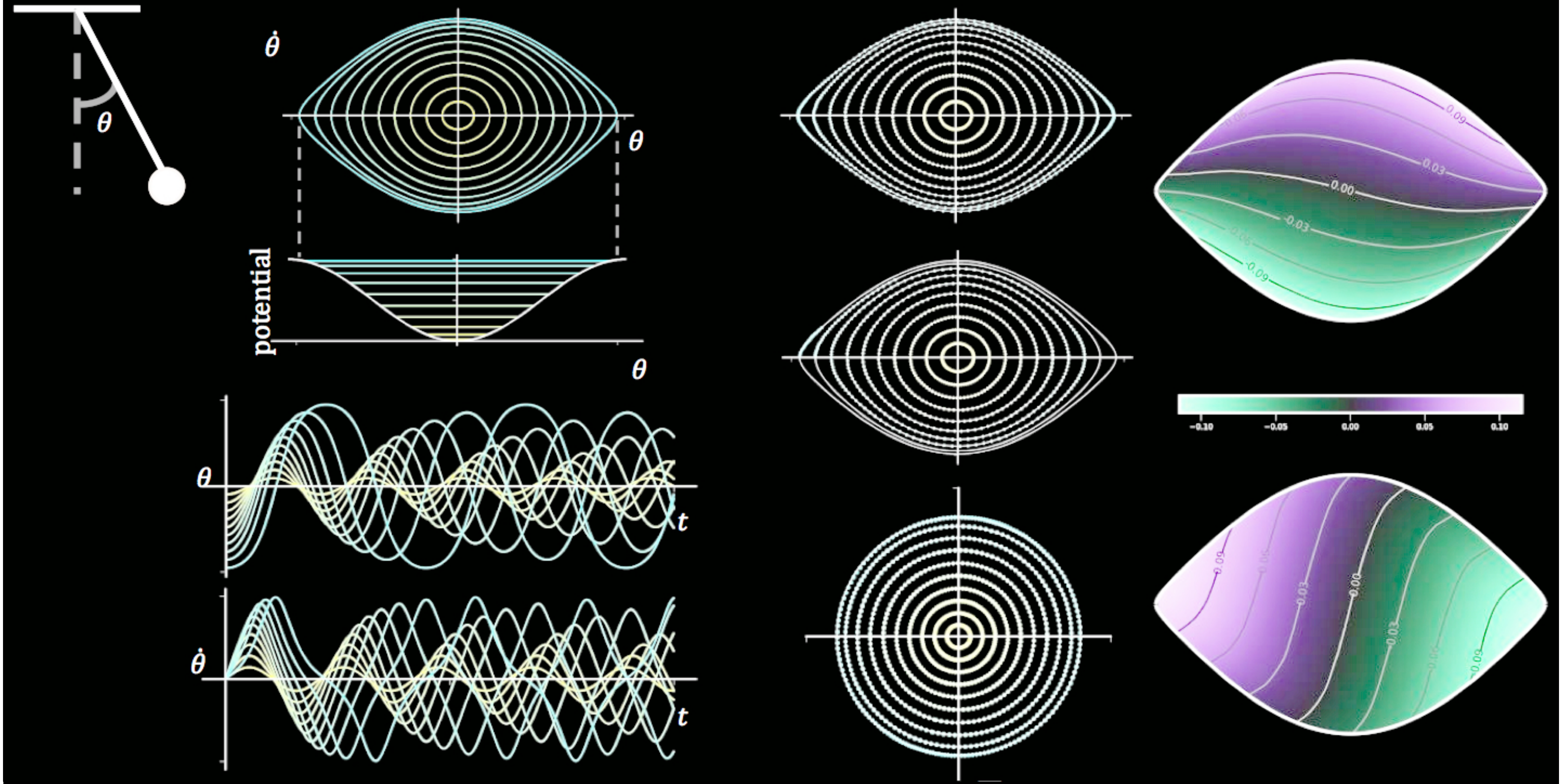
# Handling the Continuous Spectra



# Training Loss Function

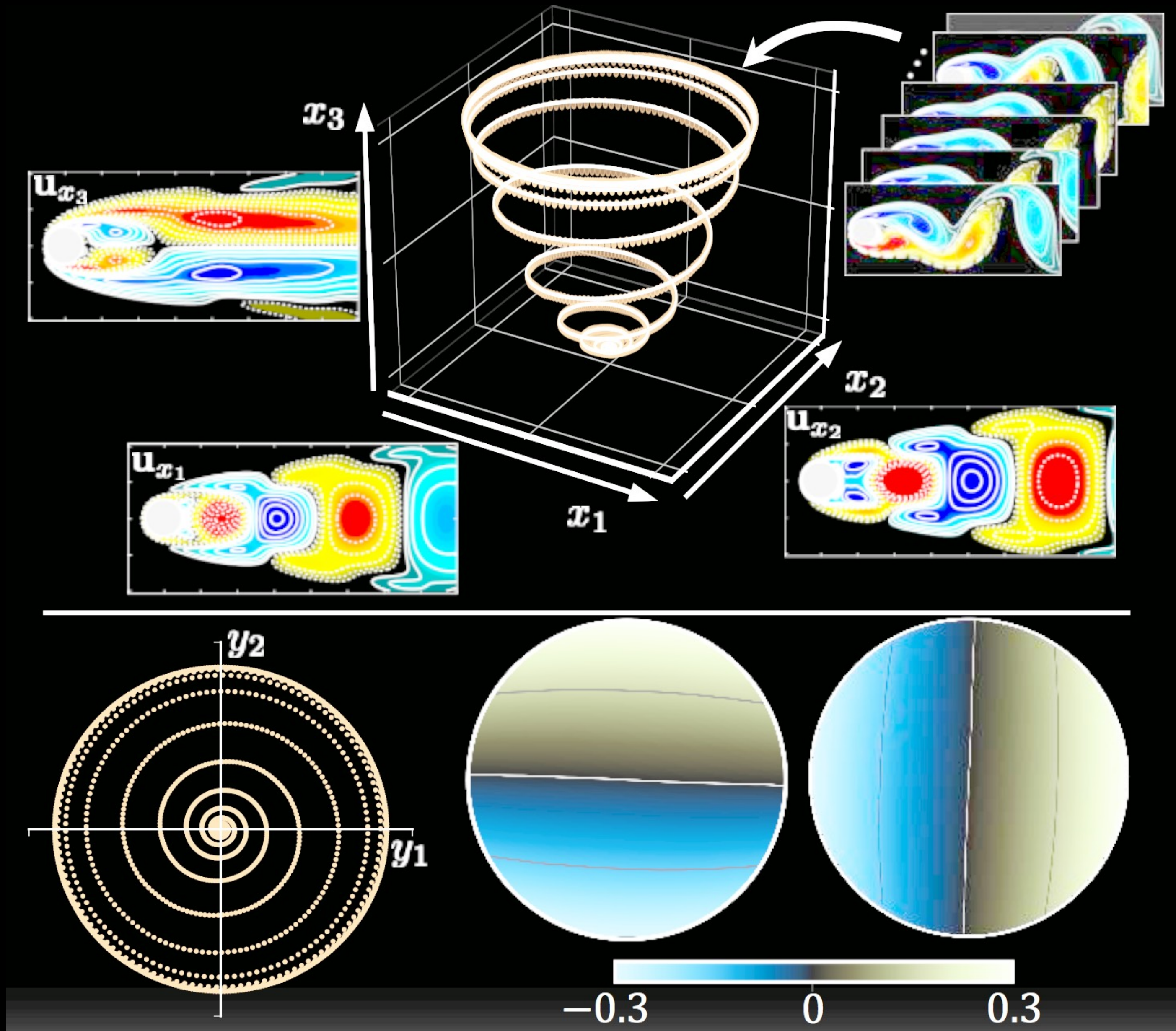


# The Pendulum



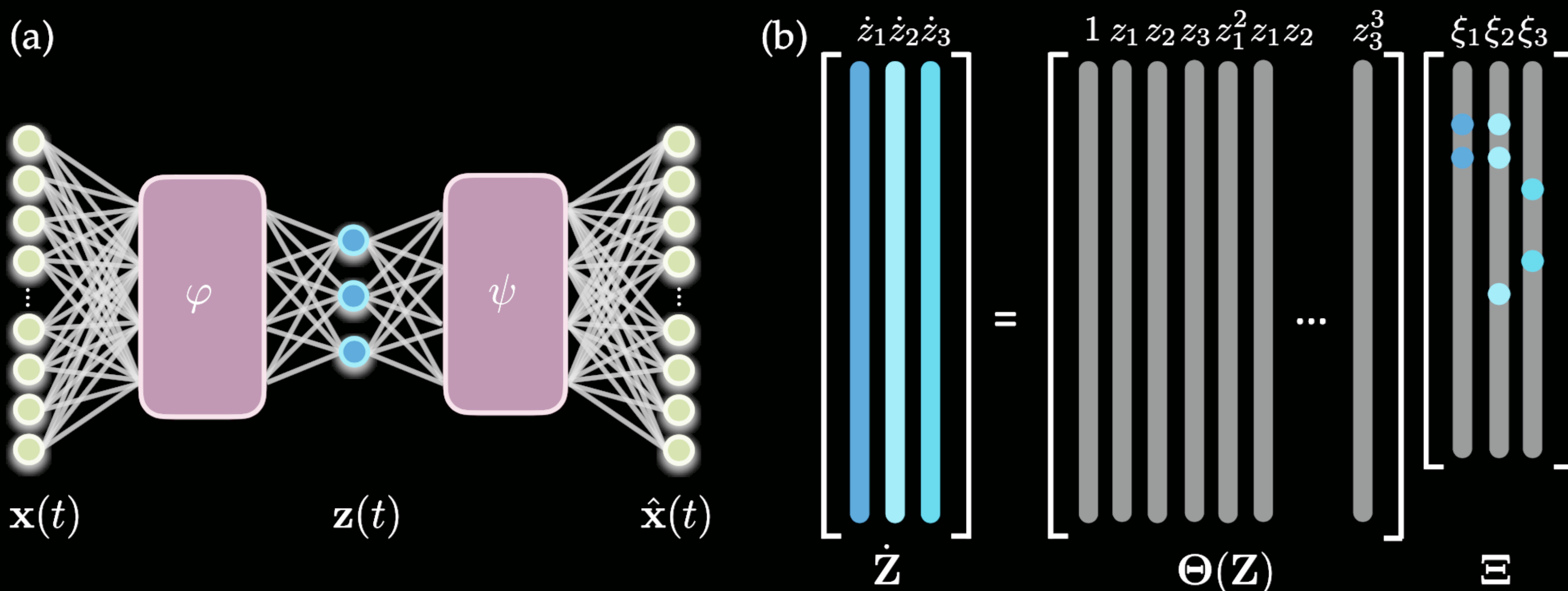

Lusch et al. Nat. Comm (2018)

# Flow Around a Cylinder



# Relax Koopman

# Coordinates + Dynamics

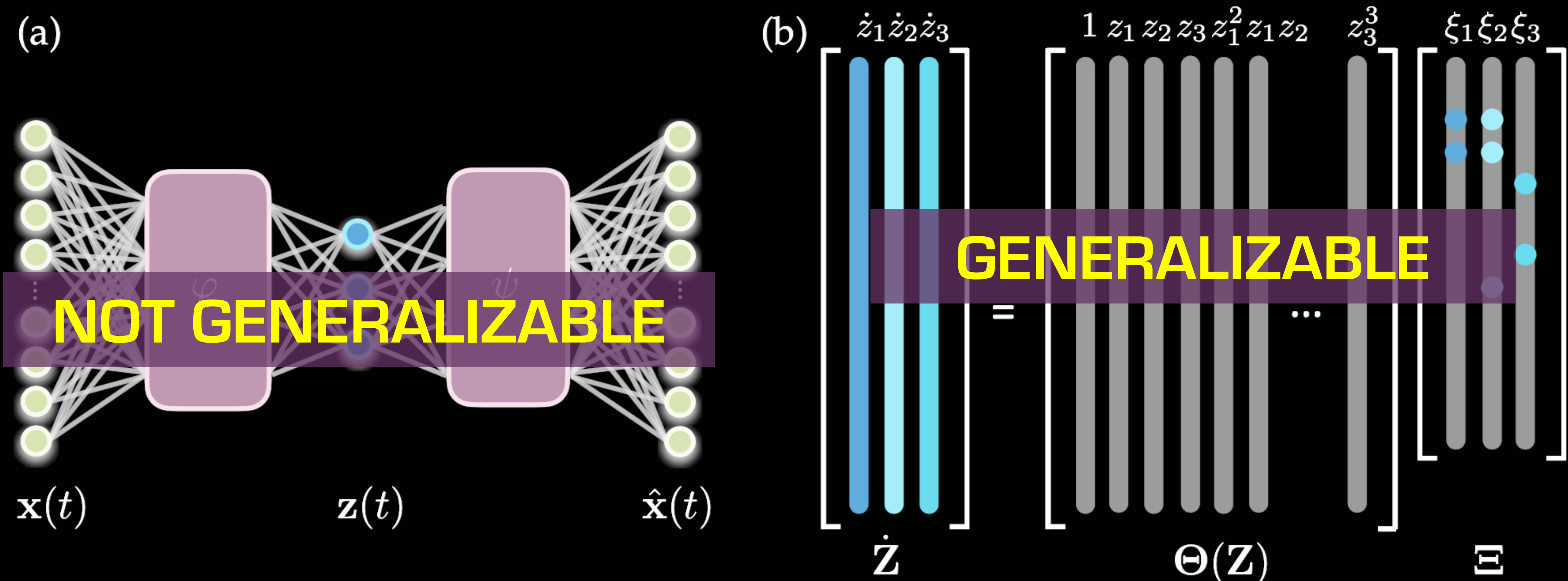



$$\dot{\mathbf{z}}_i = \nabla_{\mathbf{x}}\varphi(\mathbf{x}_i)\dot{\mathbf{x}}_i \qquad \boldsymbol{\Theta}(\mathbf{z}_i^T) = \boldsymbol{\Theta}(\varphi(\mathbf{x}_i)^T)$$

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

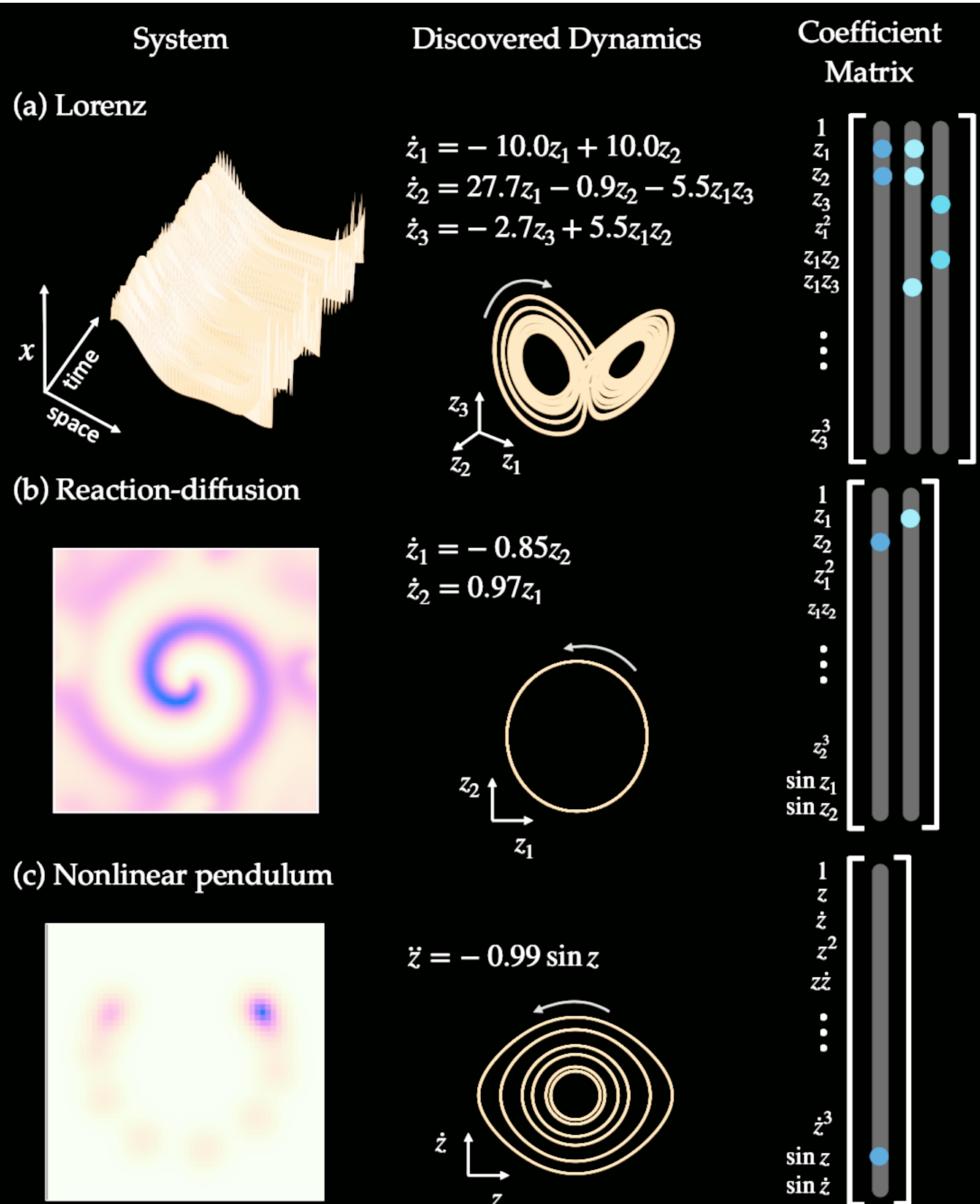
Champion et al (arxiv 2019)

# Coordinates + Dynamics




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

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


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
$z$
$\dot{z}$
$z^2$
$z\dot{z}$
$\dot{z}^3$
$\sin z$
$\sin \dot{z}$

# Generalization
# &
# Limits

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

S. Mallat, PRSA (2016)

# Parsimony

The Ultimate Physics Regularization

# of terms
# of dimensions