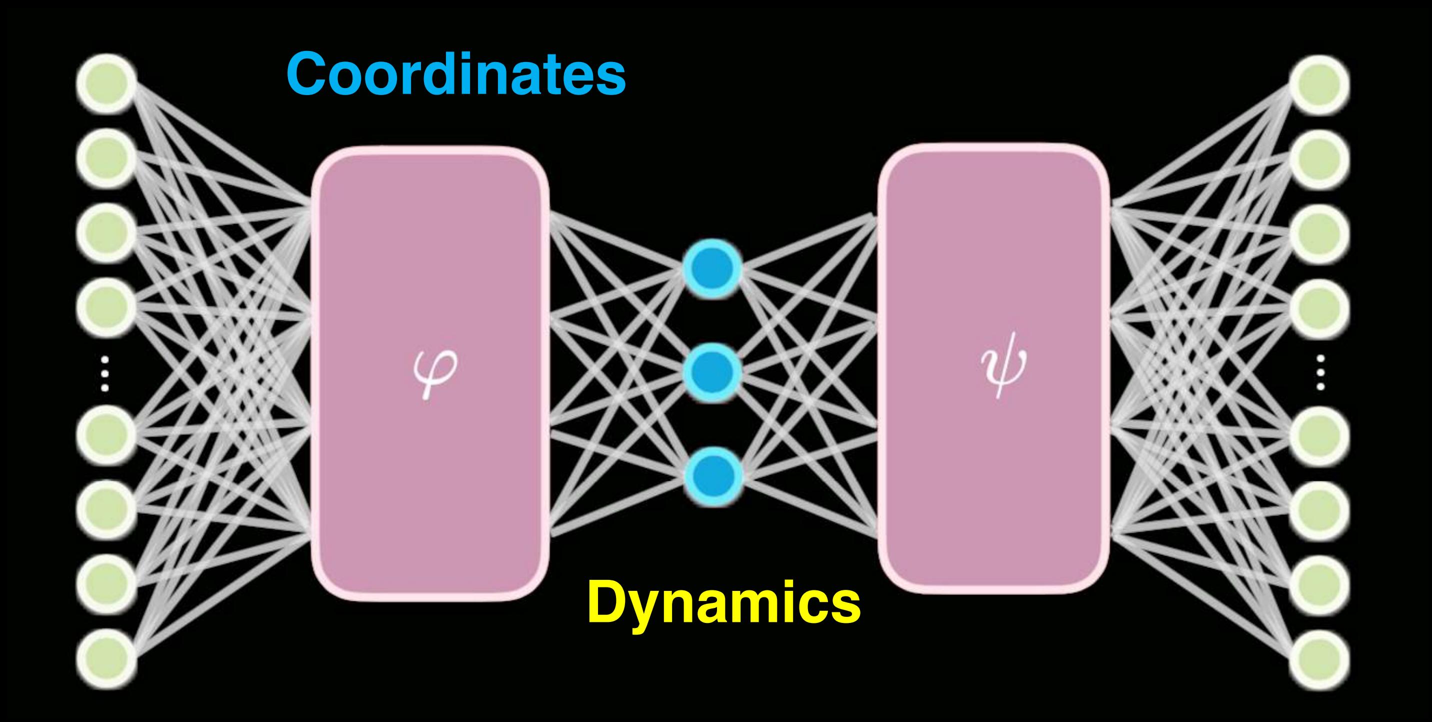
Deep Learning for the Discovery of Parsimonious Physics Models

J. Nathan Kutz

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

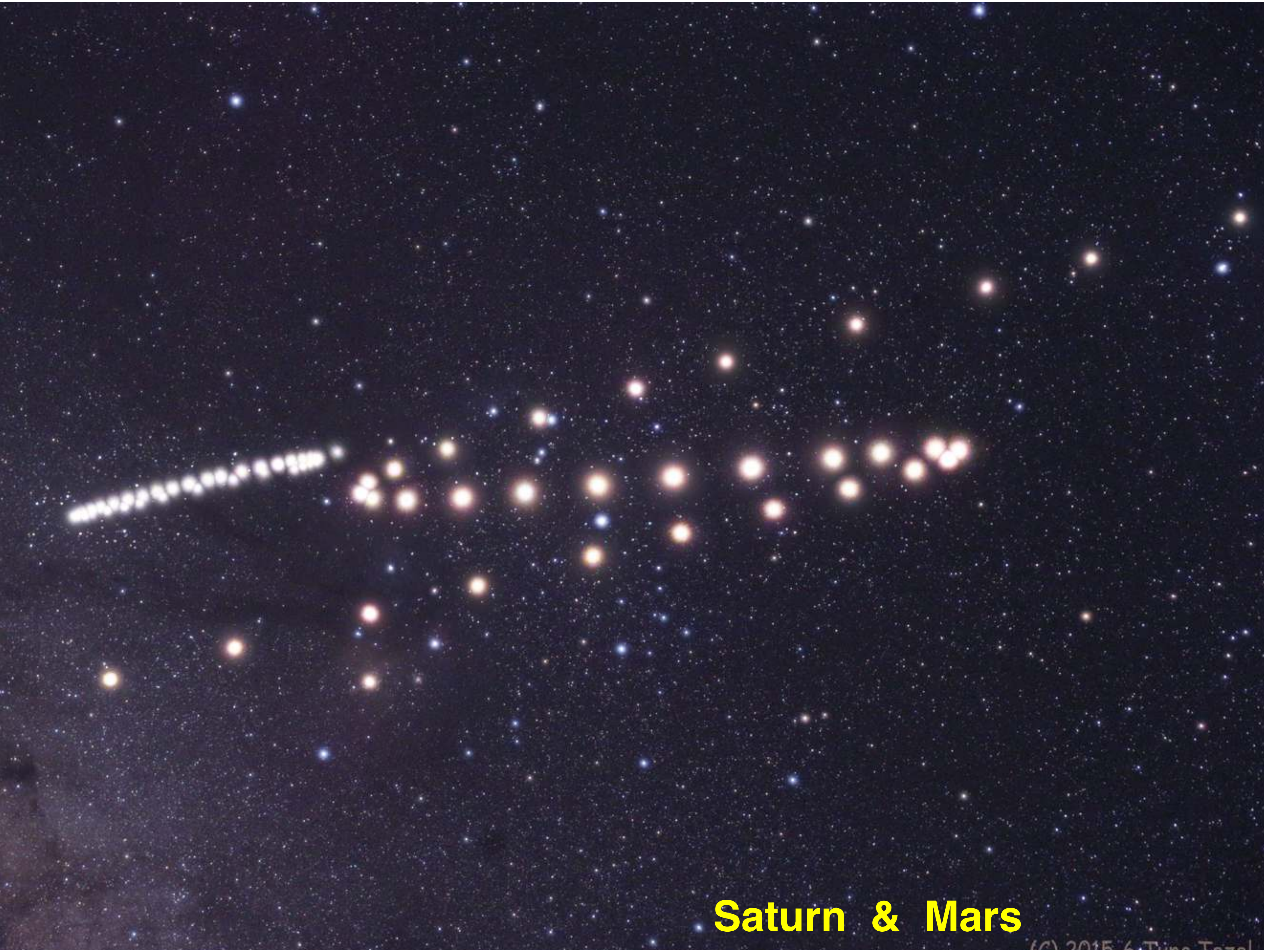
ETH Zurich – November 17, 2021

Coordinates & Dynamics

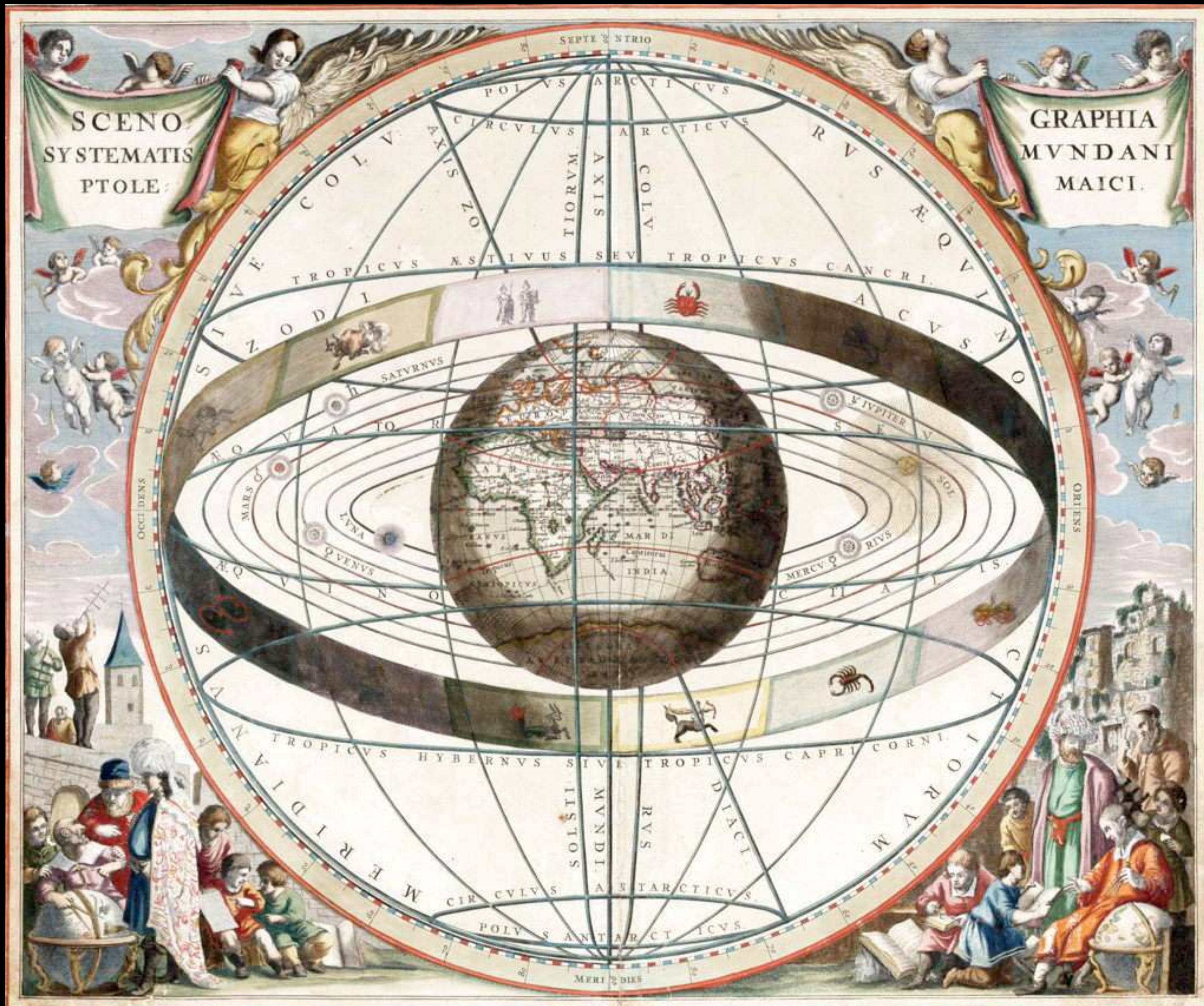


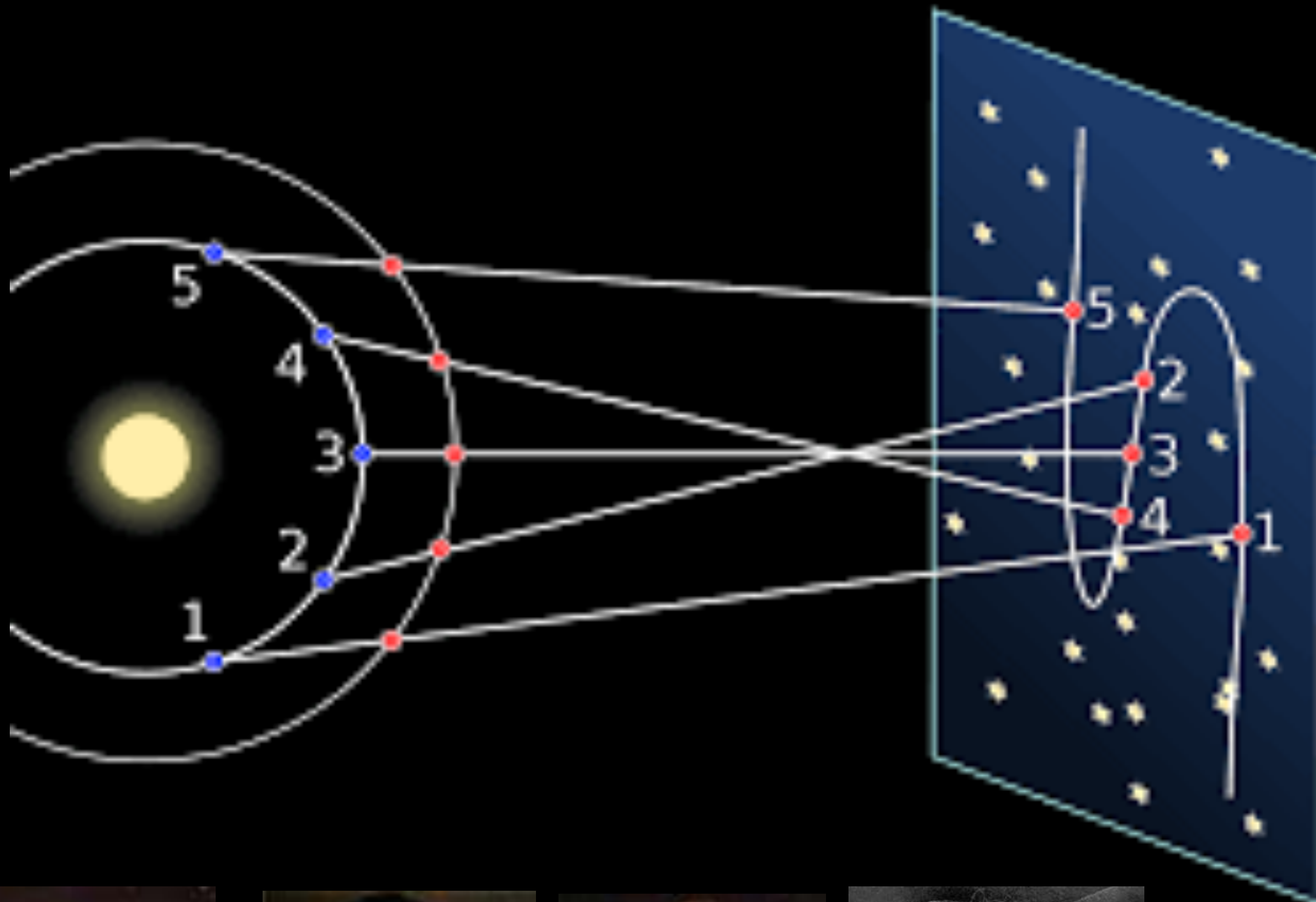
Input Data

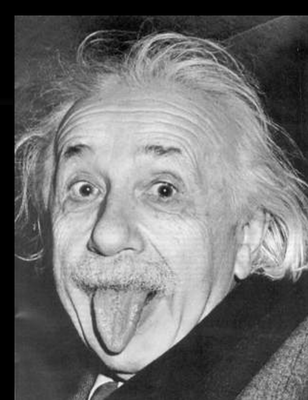
Targeted use of neural networks for discovery coordinate transformations


Saturn & Mars

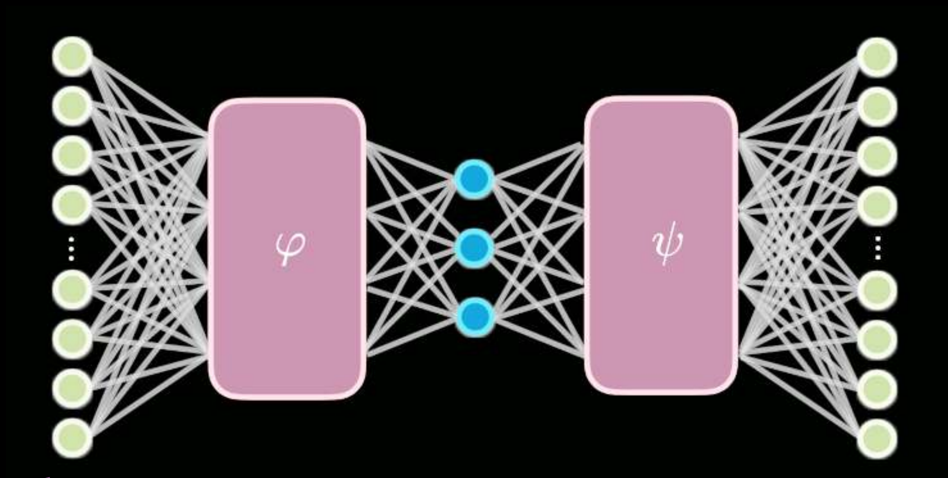
Doctrine of the Perfect Circle

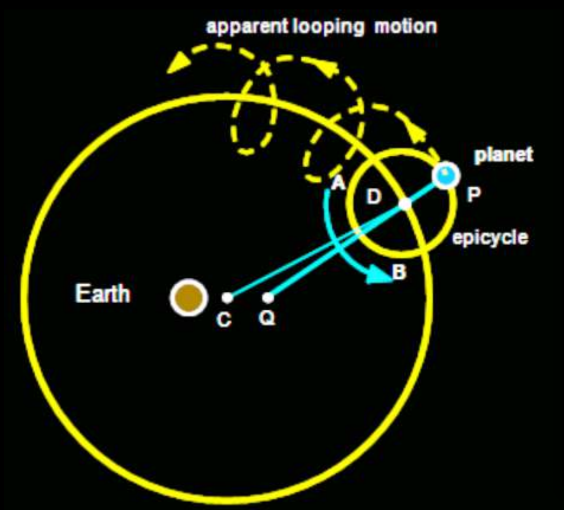



5
4
3
2
1
5
2
3
4
1

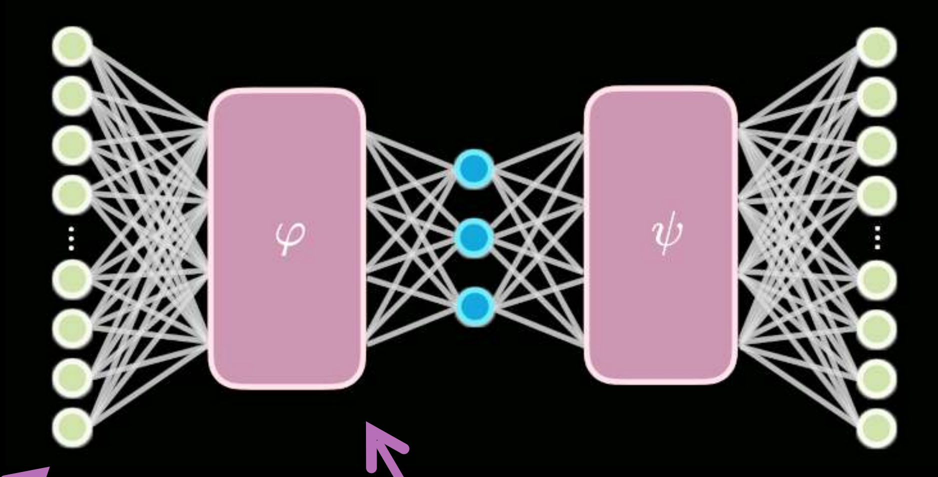


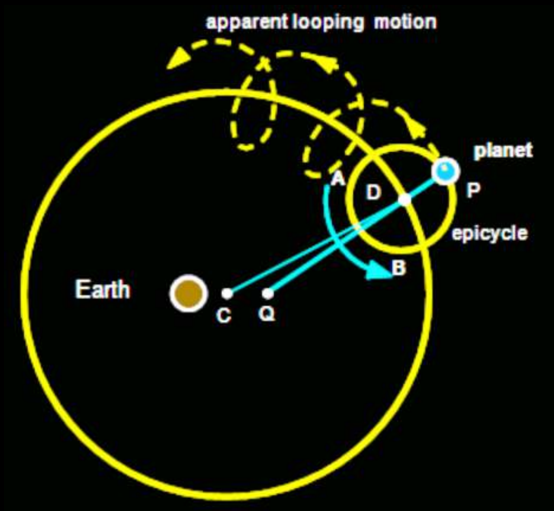
Discovery Paradigm

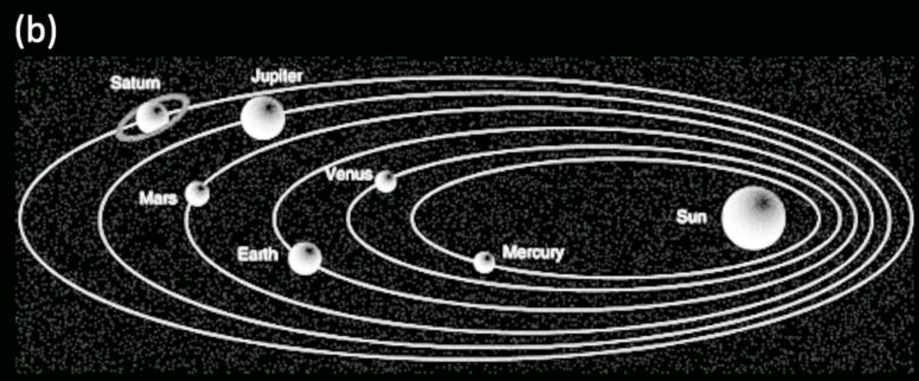




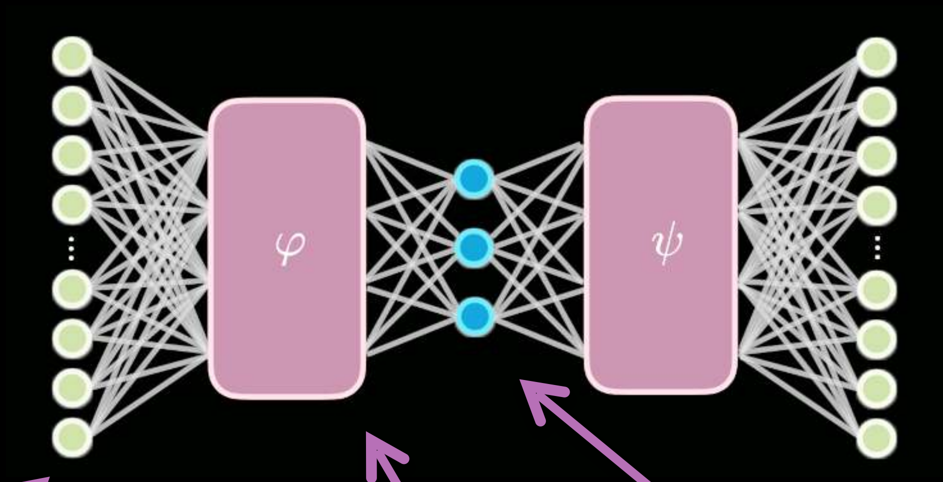
Discovery Paradigm

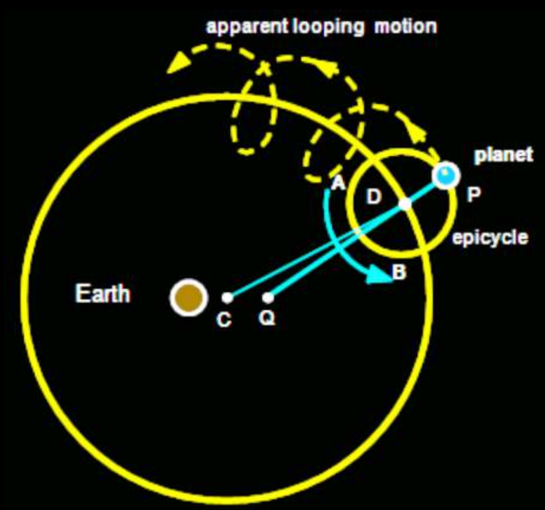


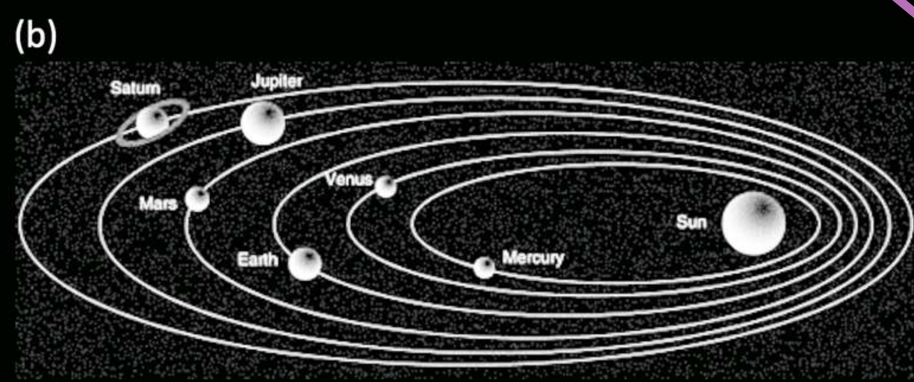




Discovery Paradigm







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

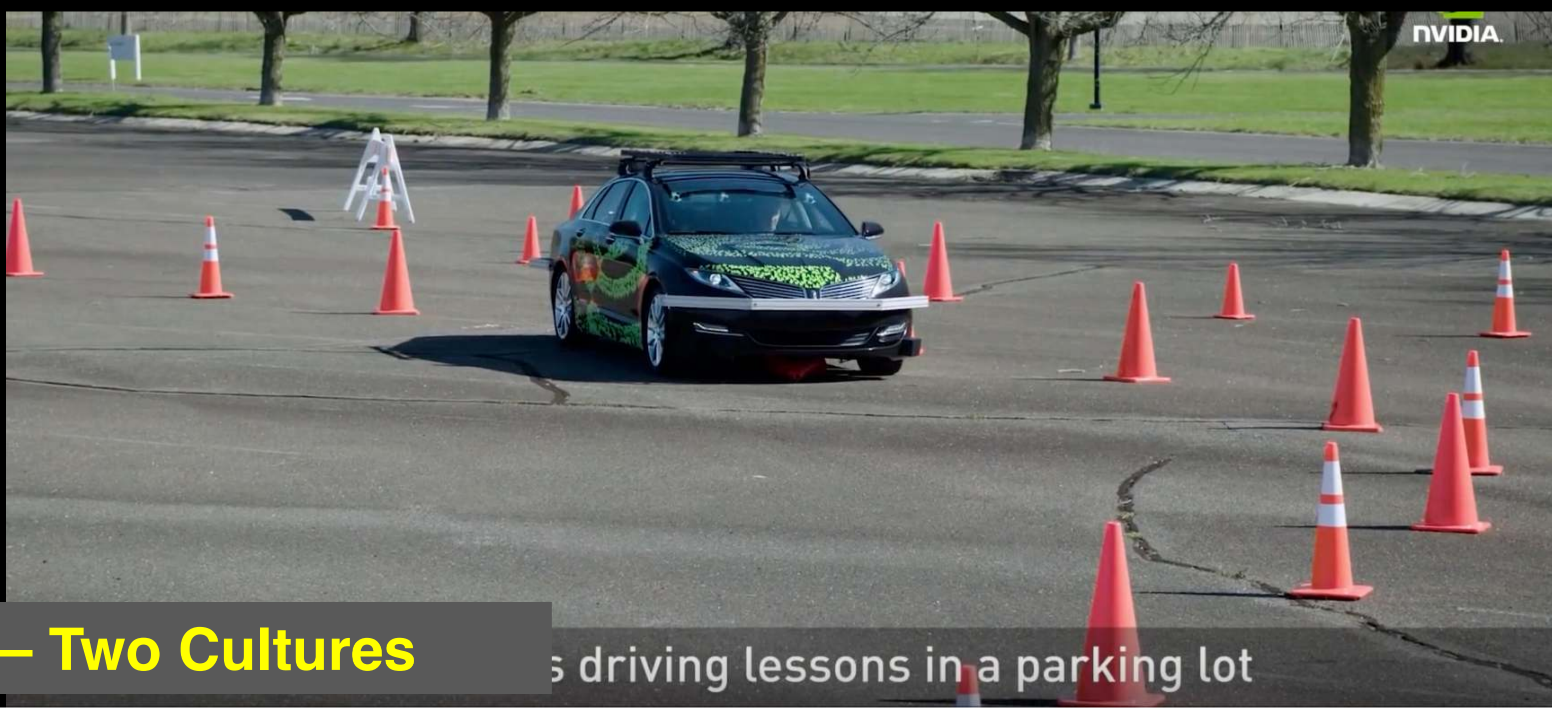
Kepler vs Newton

function approximation (ellipses)

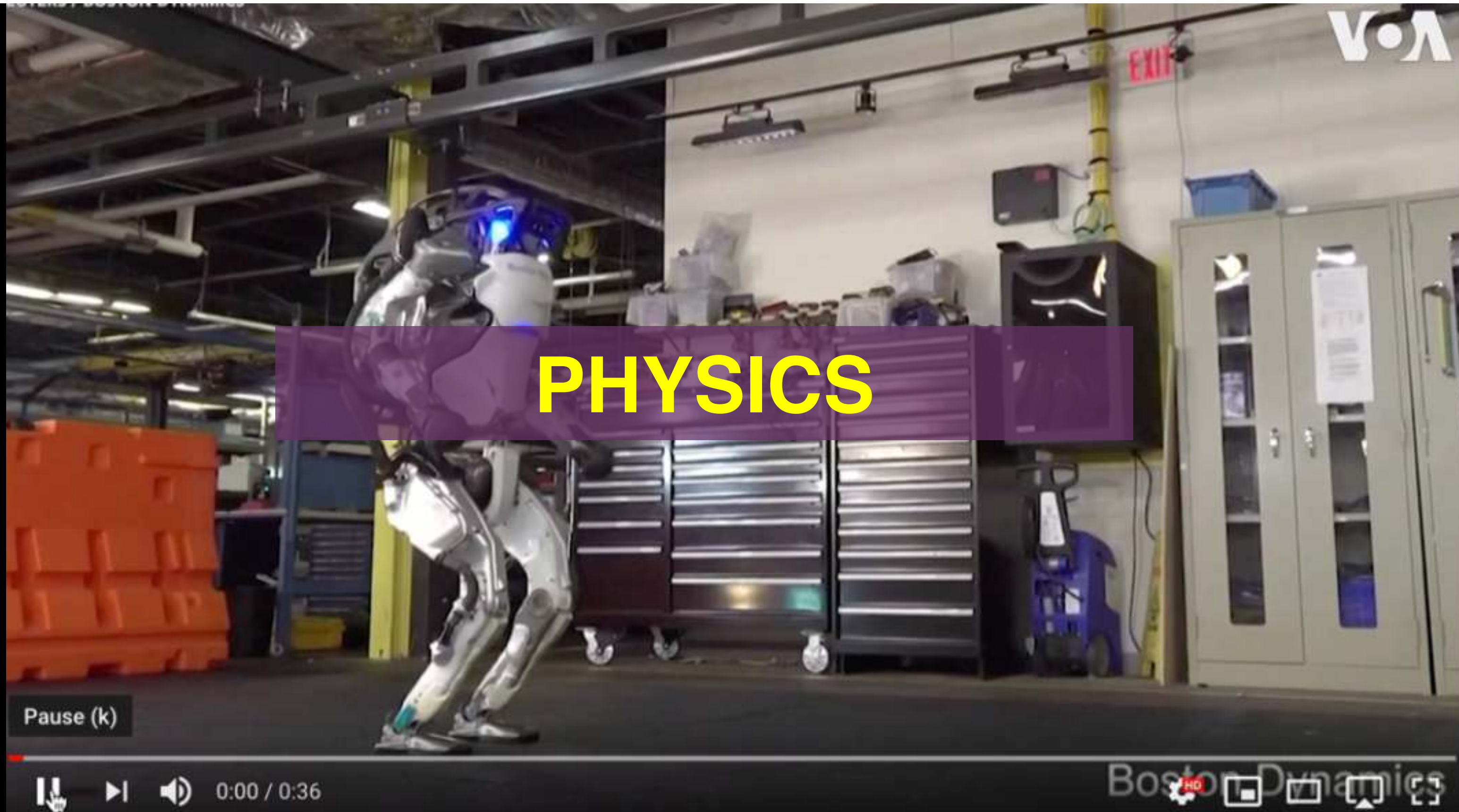
F=ma (ellipses)

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

Question #1

What is the nature of your data?

- quality
- quantity
- observability
- extrapolation vs interpolation

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

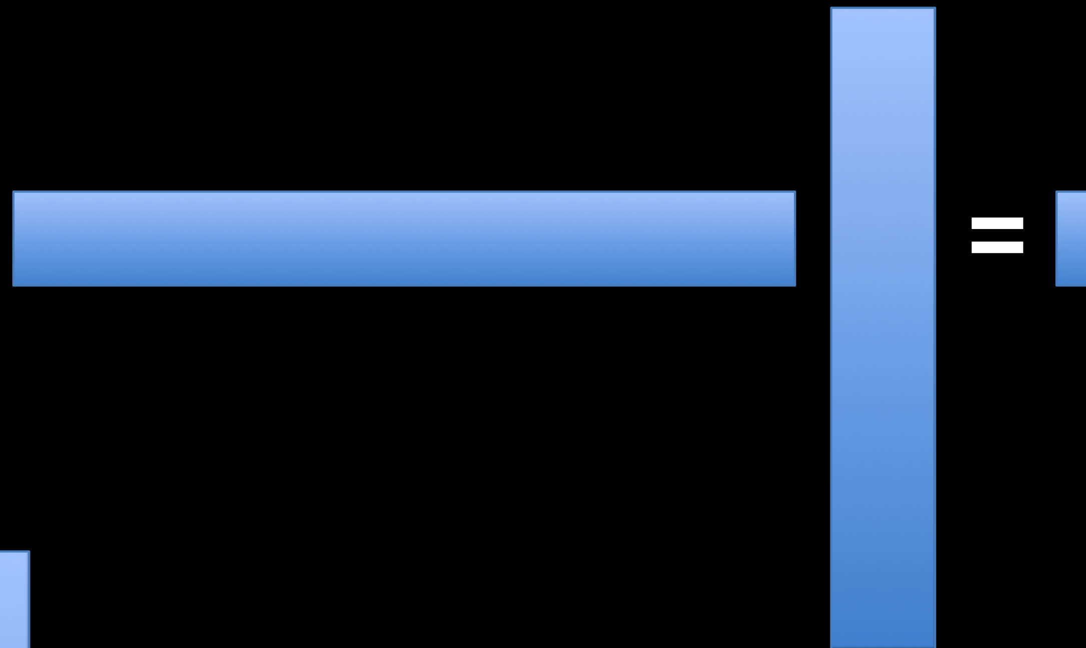
Model Discovery

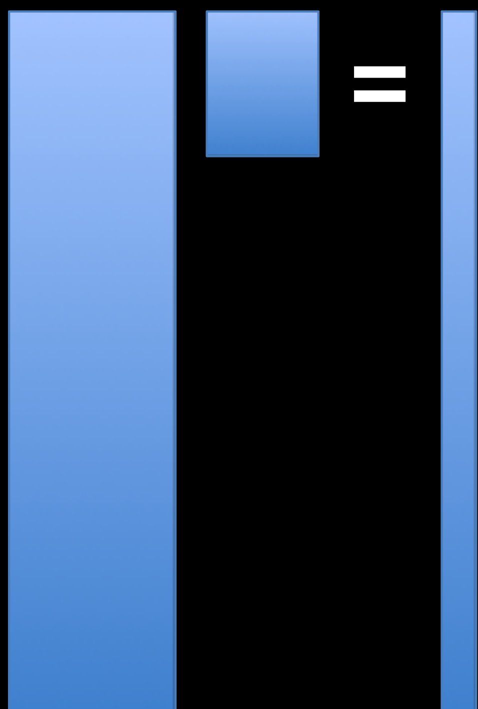
Finding governing equations

$$Ax=b$$

Data Science Today

Under





Over

- \
- pinv
- Lasso
- Ridge
- Elastic net
- Robust fit

$$Ax=b$$

subject to

$$\min g(x)$$

$$f(A,x)=b$$

subject to

$$\min g(x)$$

Governing Dynamical Systems

Generic nonlinear , time-dependent, parametric system

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

Measurements (assimilation)

$$G(\mathbf{x}, t_k) = 0$$

What Could the Right Side Be?

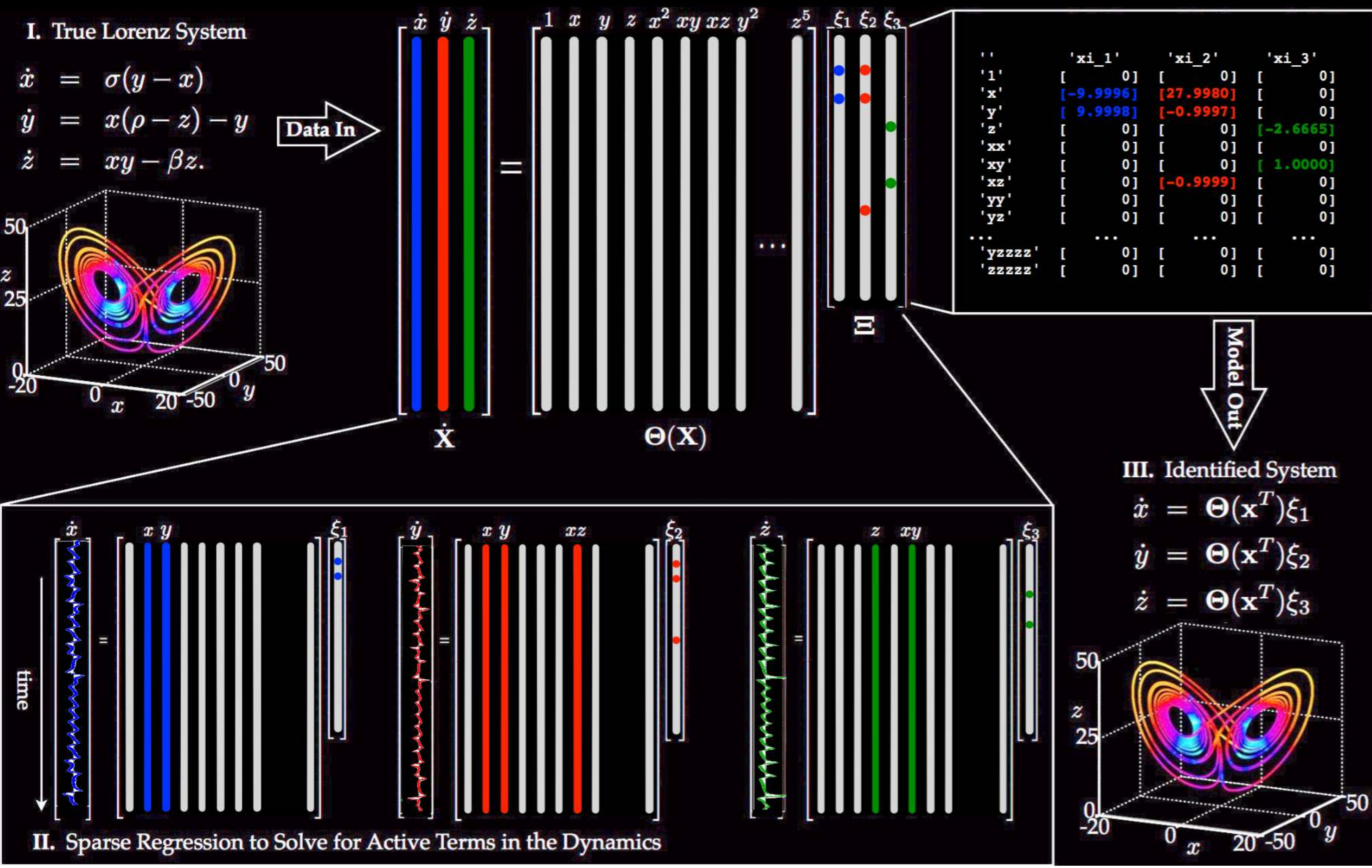
Limited by your imagination

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

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)



Nonlinear Systems ID

1. Collect Data

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}^T(t_1) \\ \mathbf{x}^T(t_2) \\ \vdots \\ \mathbf{x}^T(t_m) \end{bmatrix} = \overset{\text{state}\ \longrightarrow}{\begin{bmatrix} x_1(t_1) & x_2(t_1) & \cdots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \cdots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_m) & x_2(t_m) & \cdots & x_n(t_m) \end{bmatrix}} \Big\downarrow \text{time}$$

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{\mathbf{x}}^T(t_1) \\ \dot{\mathbf{x}}^T(t_2) \\ \vdots \\ \dot{\mathbf{x}}^T(t_m) \end{bmatrix} = \begin{bmatrix} \dot{x}_1(t_1) & \dot{x}_2(t_1) & \cdots & \dot{x}_n(t_1) \\ \dot{x}_1(t_2) & \dot{x}_2(t_2) & \cdots & \dot{x}_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{x}_1(t_m) & \dot{x}_2(t_m) & \cdots & \dot{x}_n(t_m) \end{bmatrix}.$$

2. Build Library of Candidate Nonlinearities

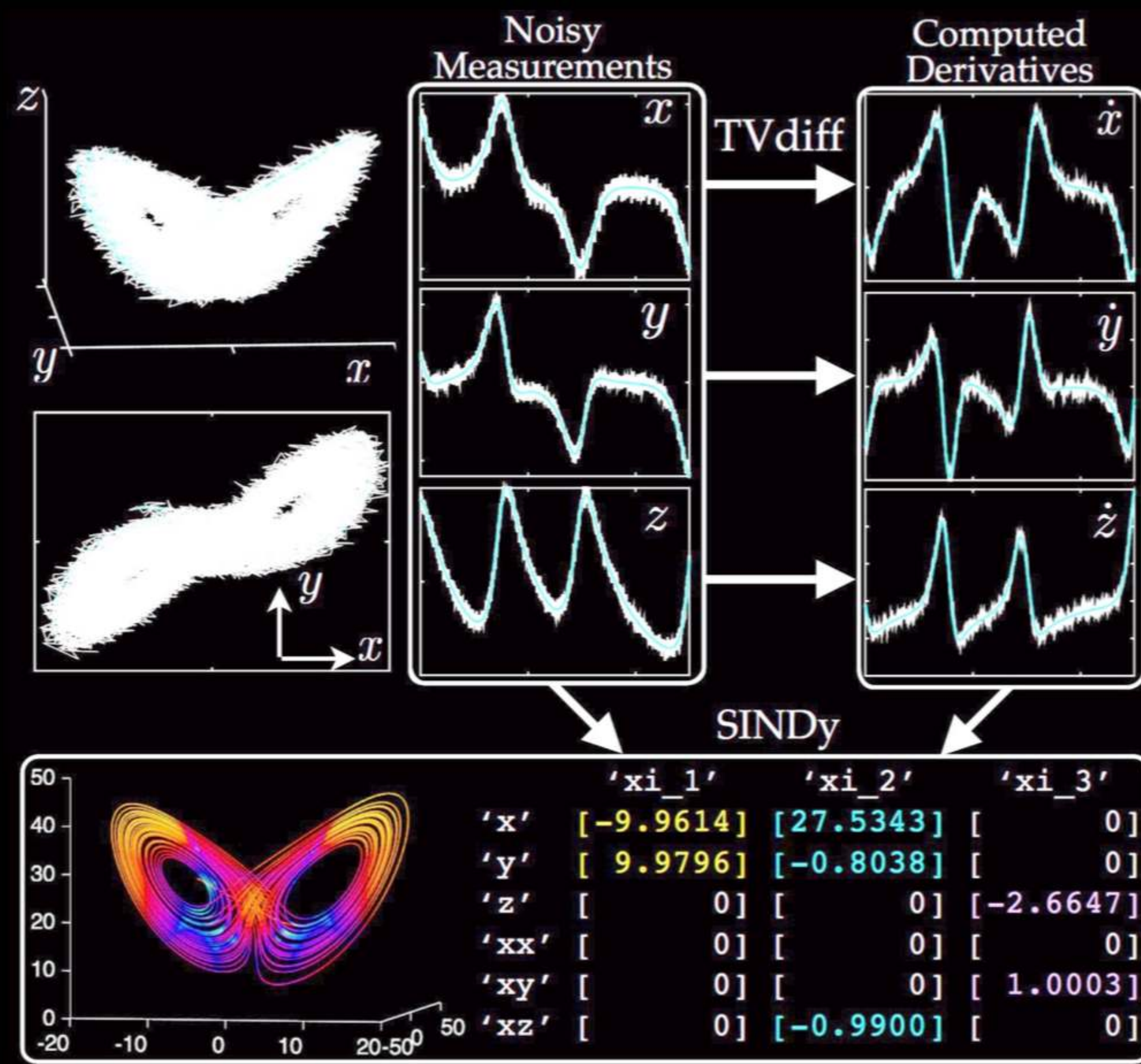
$$\Theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & & | & | & \\ \mathbf{1} & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \cdots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \cdots \\ | & | & | & | & & | & | & \end{bmatrix}.$$

3. Sparse Regression to Find Active Terms

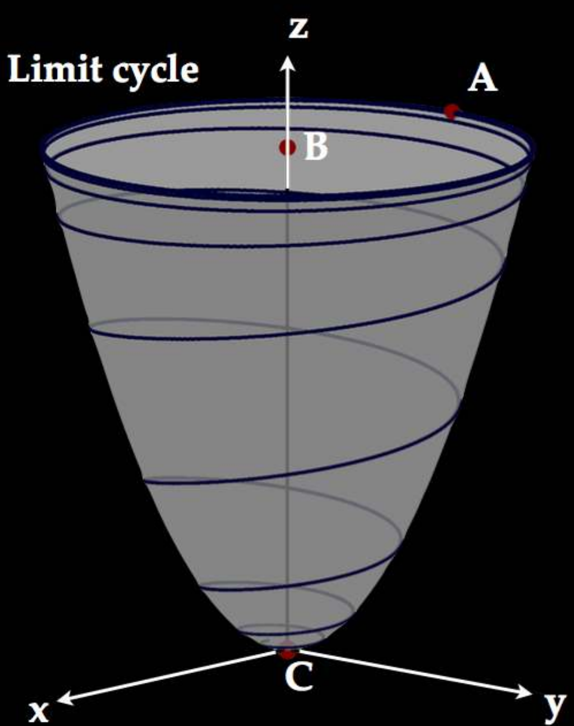
$$\dot{\mathbf{X}} = \Theta(\mathbf{X})\Xi.$$

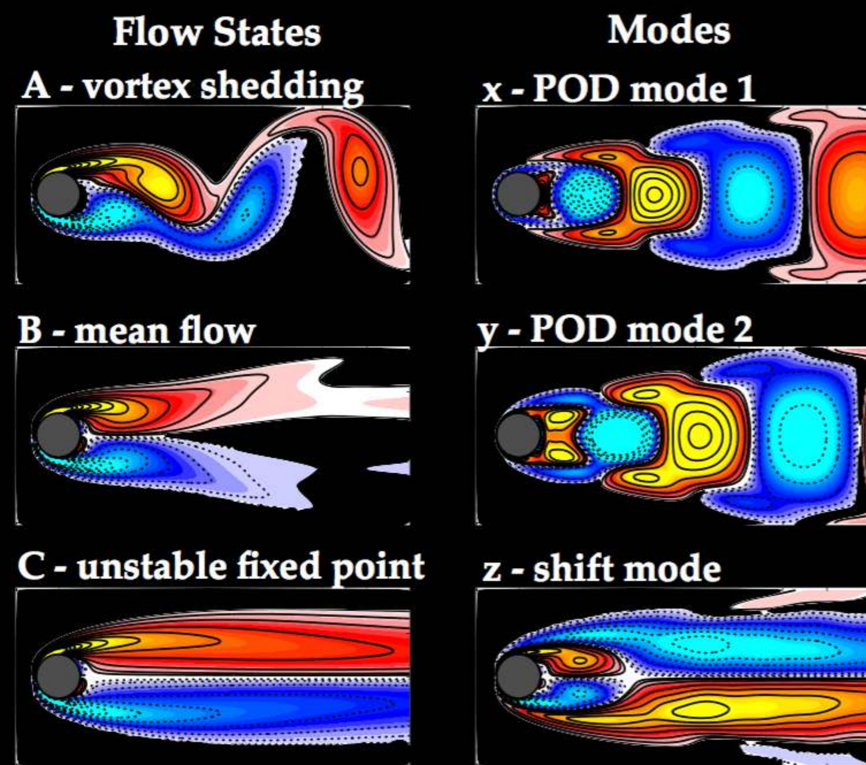
4. Nonlinear Model

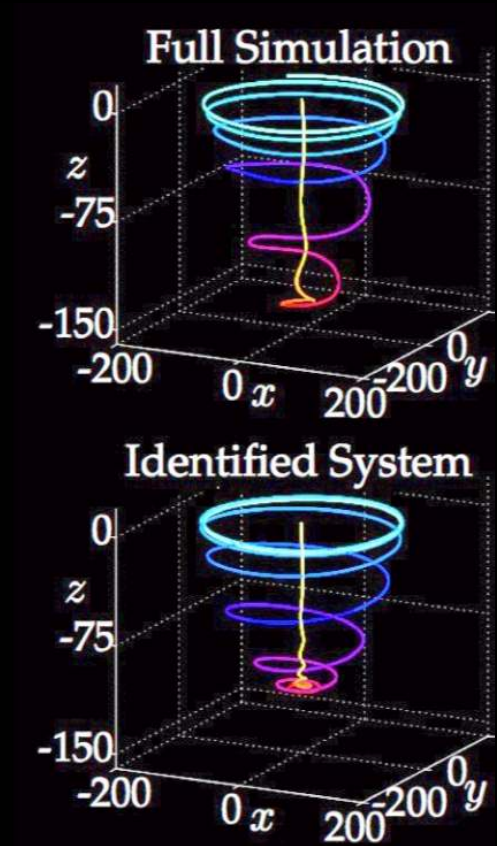
$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) = \Xi^T(\Theta(\mathbf{x}^T))^T$$



Identifying Slow Manifolds







30 years of progress

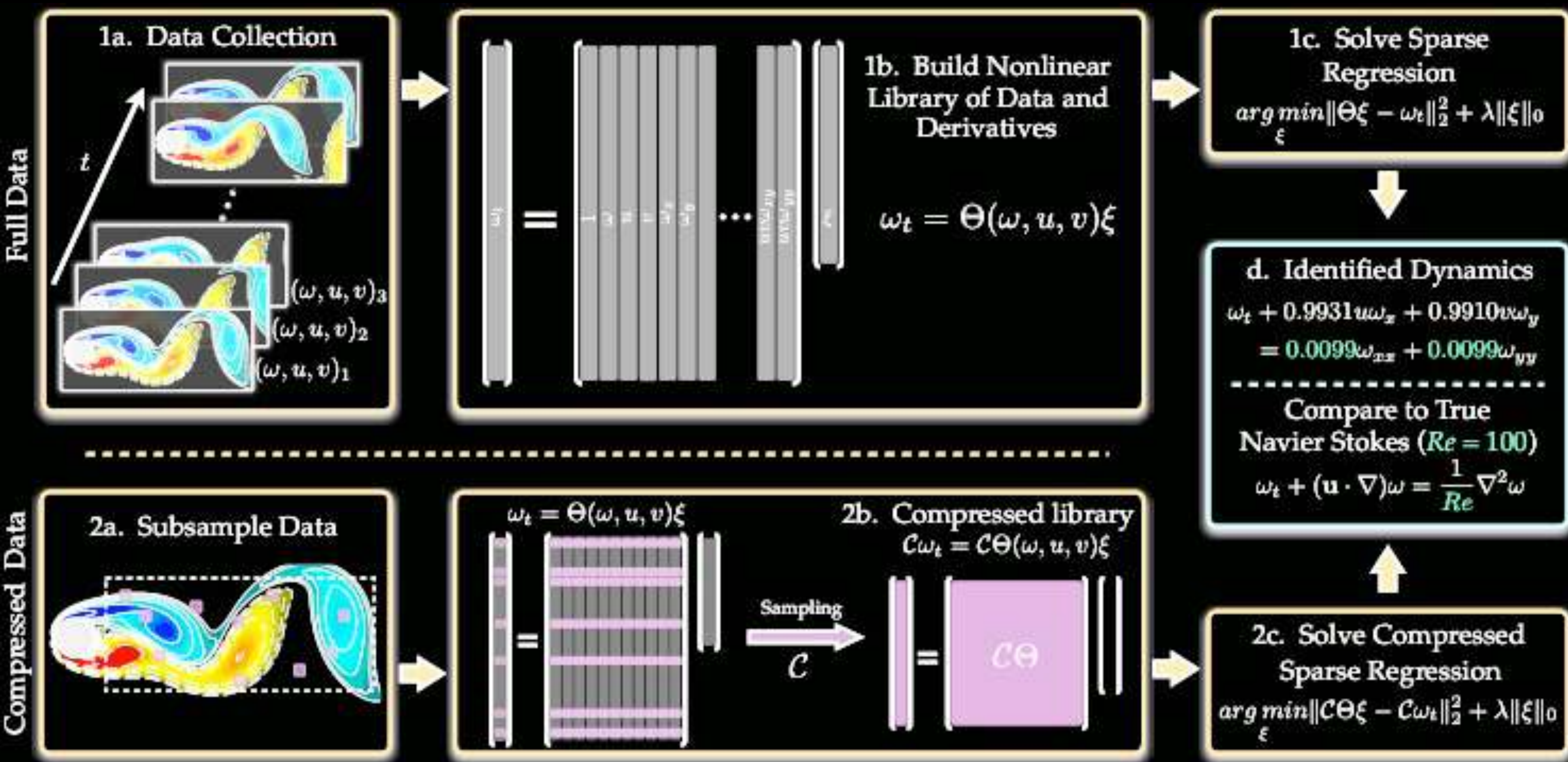
$$
\begin{aligned}
\dot{x} &= \mu x - \omega y + Axz \\
\dot{y} &= \omega x + \mu y + Ayz \\
\dot{z} &= -\lambda(z - x^2 - y^2).
\end{aligned}
$$

1. Hopf bifurcations as path to turbulence
Ruelle & Takens, *Communications in Mathematical Physics*, 1971

2. Vortex shedding and Hopf bifurcation
Jackson, *Journal of Fluid Mechanics, 1987.*

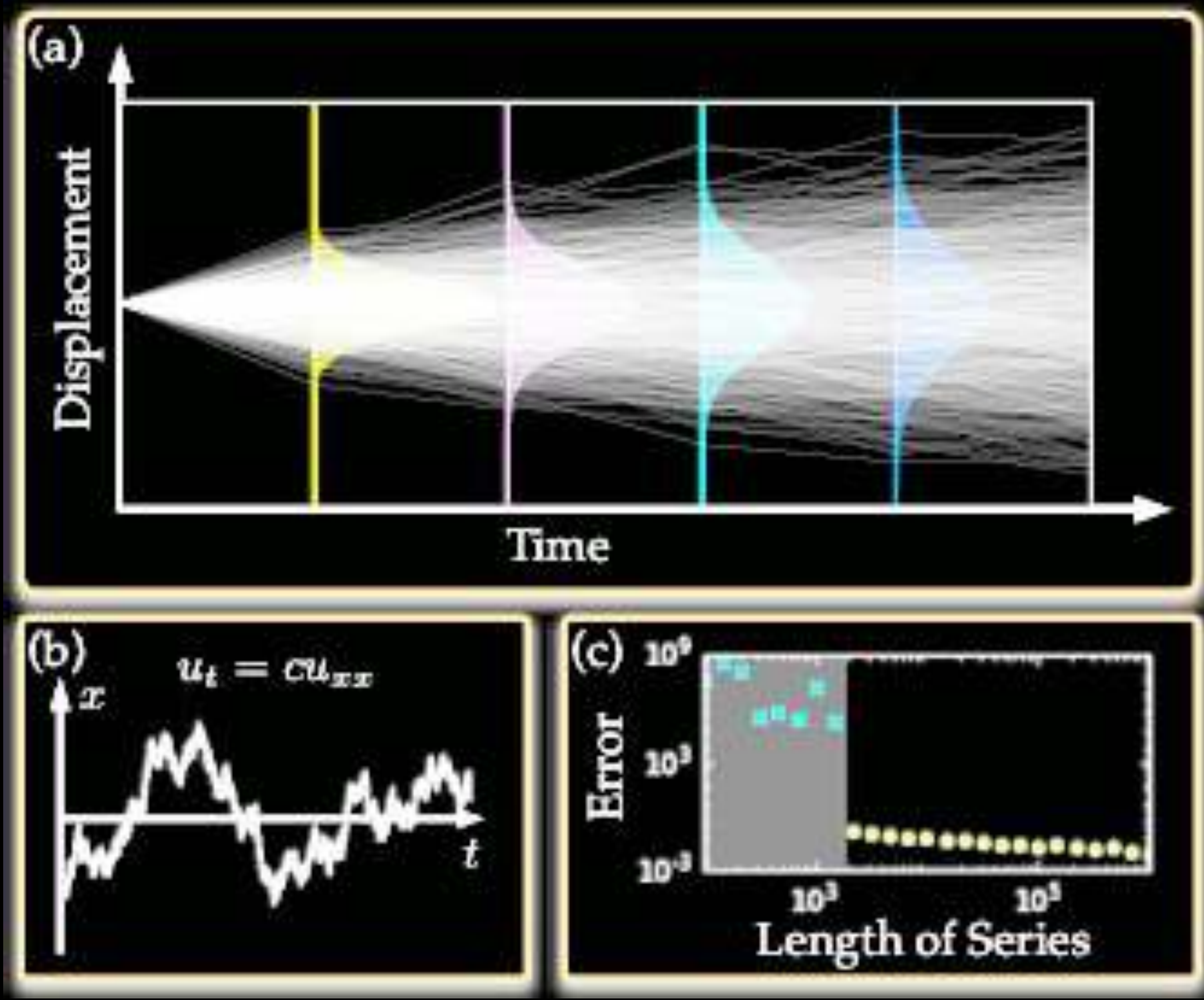
3. Mean-field model with slow manifold
Noack, Afanasiev, Morzynski, Tadmor, & Thiele,
Journal of Fluid Mechanics, 2003.

Discovering PDEs

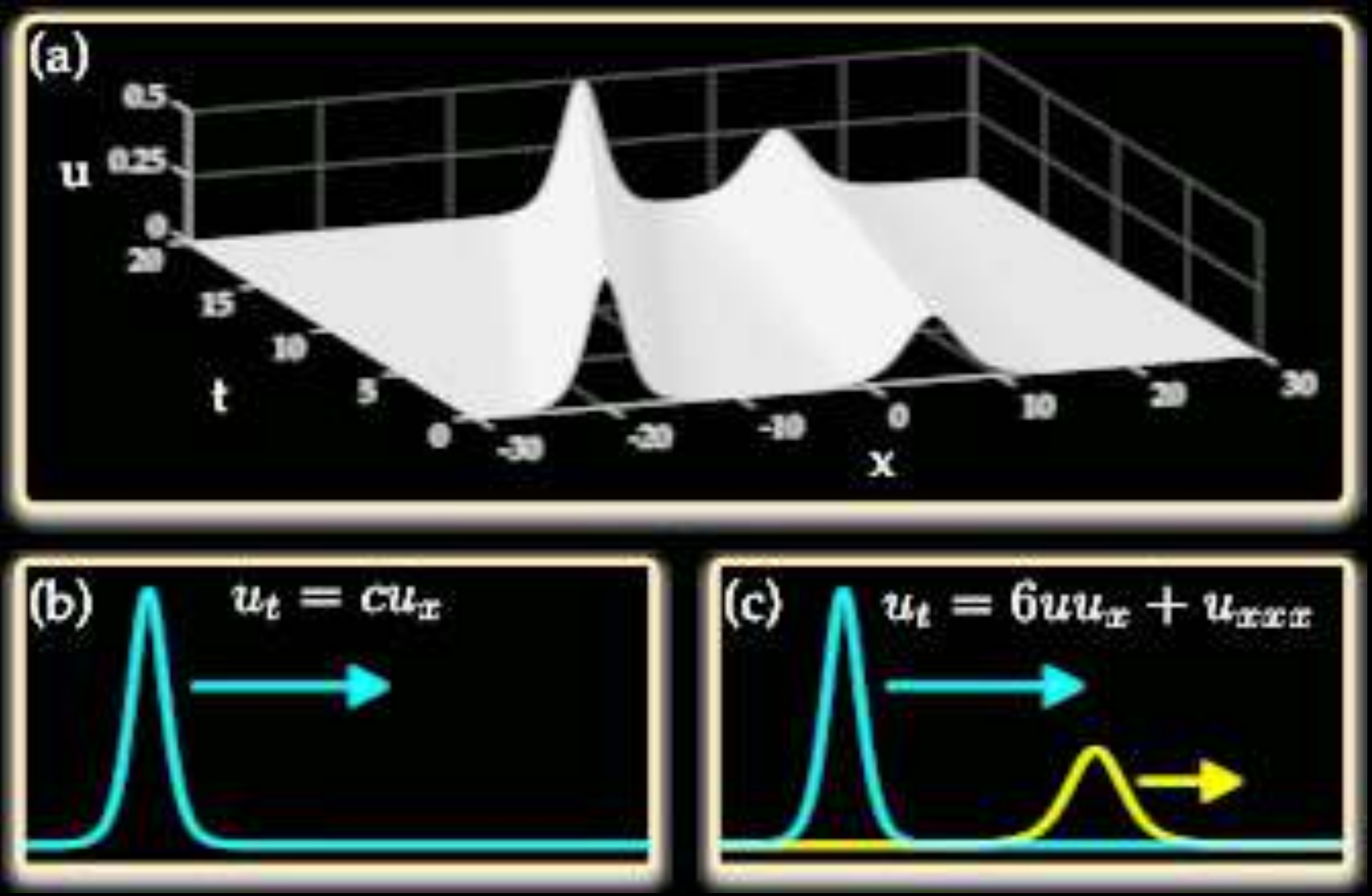


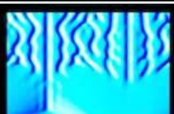
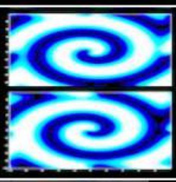
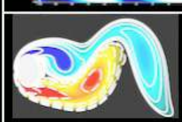
Sam Rudy

Lagrangian Measurements

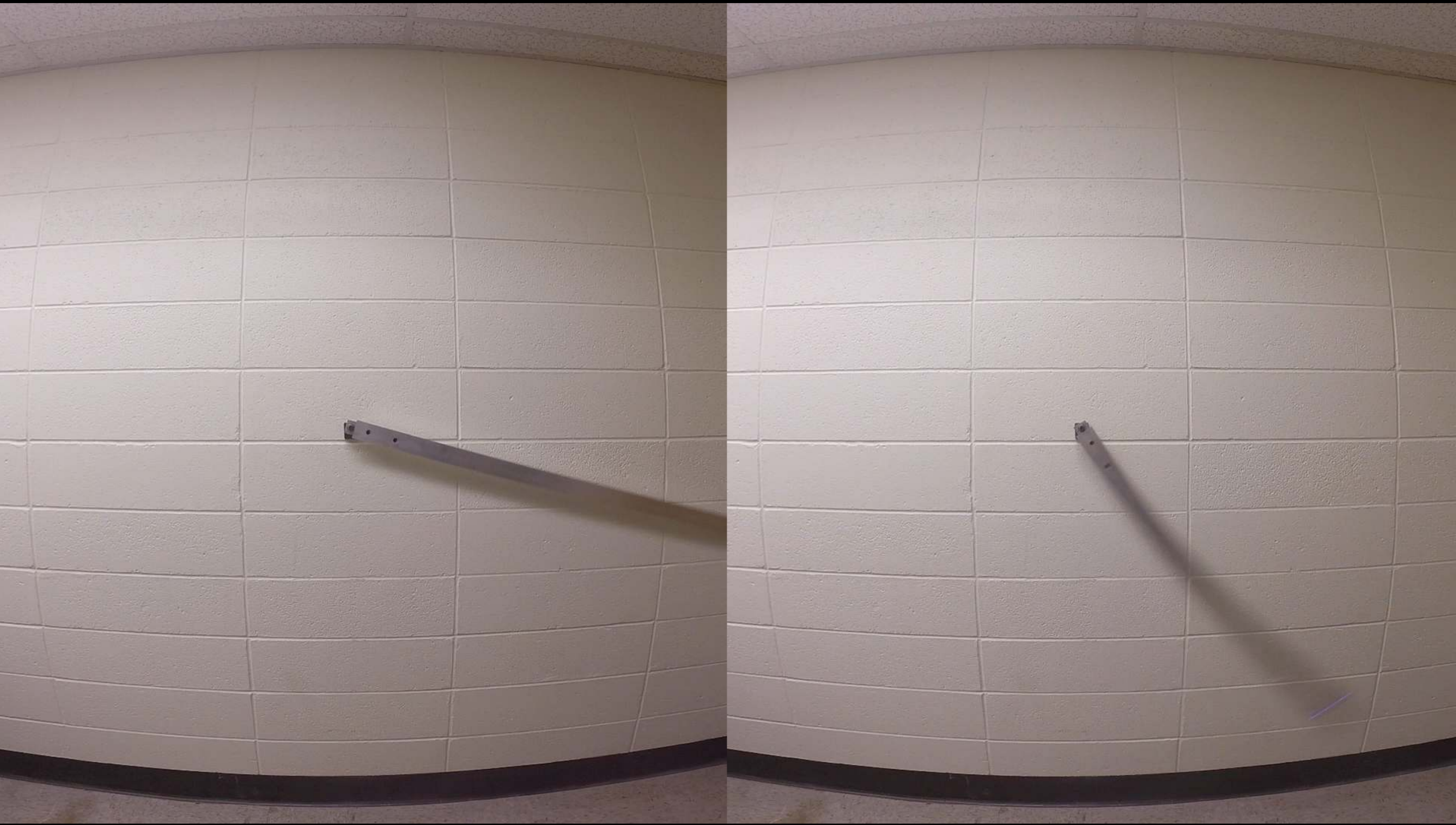


Disambiguation



PDE	Form	Error (no noise, noise)	Discretization
 KdV	$u_t + 6uu_x + u_{xxx} = 0$	$1\%\pm0.2\%, 7\%\pm5\%$	$x\in[-30, 30], n{=}512,\ t\in[0, 20], m{=}201$
 Burgers	$u_t + uu_x - \epsilon u_{xx} = 0$	$0.15\%\pm0.06\%, 0.8\%\pm0.6\%$	$x\in[-8, 8], n{=}256,\ t\in[0, 10], m{=}101$
 Schrodinger	$iu_t + \frac{1}{2}u_{xx} - \frac{x^2}{2}u = 0$	$0.25\%\pm0.01\%, 10\%\pm7\%$	$x\in[-7.5, 7.5], n{=}512,\ t\in[0, 10], m{=}401$
 NLS	$iu_t + \frac{1}{2}u_{xx} + \|u\|^2u = 0$	$0.05\%\pm0.01\%, 3\%\pm1\%$	$x\in[-5, 5], n{=}512,\ t\in[0, \pi], m{=}501$
 KS	$u_t + uu_x + u_{xx} + u_{xxxx} = 0$	$1.3\%\pm1.3\%, 70\%\pm27\%$	$x\in[0, 100], n{=}1024,\ t\in[0, 100], m{=}251$
 R-D	$u_t = 0.1\nabla^2u + \lambda(A)u - \omega(A)v$ $v_t = 0.1\nabla^2v + \omega(A)u + \lambda(A)v$ $A{=}u^2{+}v^2, \omega{=}-\beta A^2, \lambda{=}1-A^2$	$0.02\% \pm 0.01\%,\ 3.8\% \pm 2.4\%$	$x, y\in[-10, 10], n{=}256,\ t\in[0, 10], m{=}201$ subsample $3 \cdot 10^5$
 Navier Stokes	$\omega_t + (\mathbf{u} \cdot \nabla)\omega = \frac{1}{Re}\nabla^2\omega$	$1\% \pm 0.2\%$, $7\% \pm 6\%$	$x\in[0, 9], n_x{=}449,\ y\in[0, 4], n_y{=}199,$ $t\in[0, 30], m{=}151$, subsample $3 \cdot 10^5$

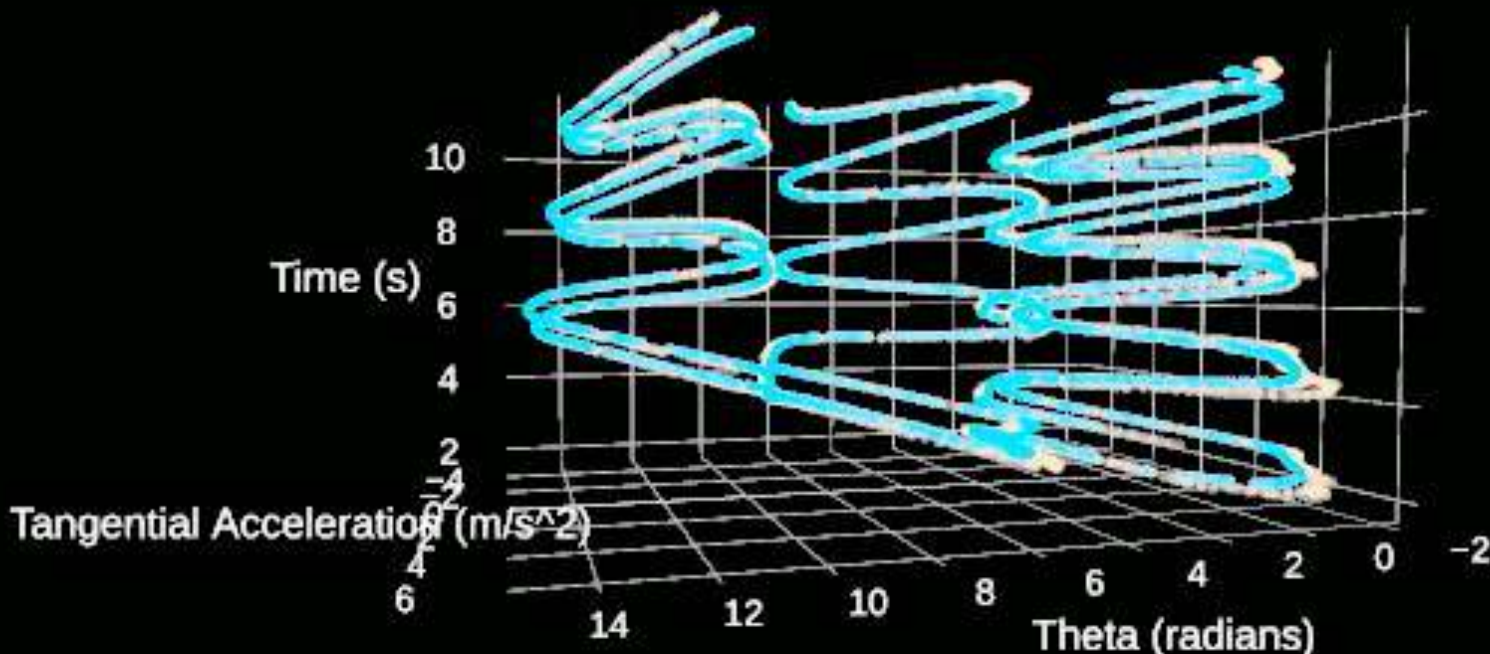
Experiments



Arduino Magic

Taren Gorman

Data vs. SINDy Plot



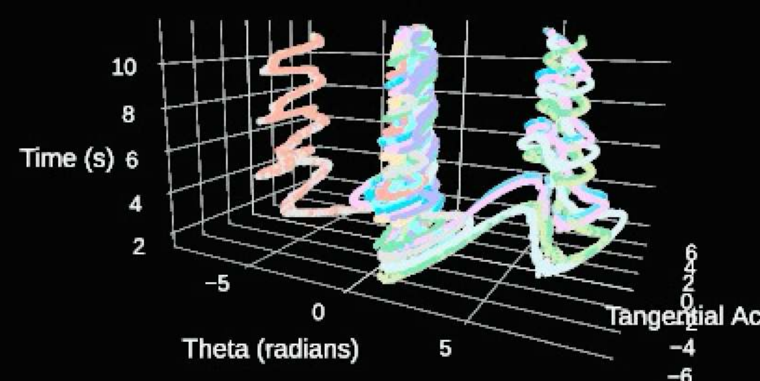

```
/home/taren/ana
ning:

divide by zero

/home/taren/ana
ning:

divide by zero

(77854, 2) (778
With  -1 jobs, fit and predict STRidge took 5.747981 seconds.
dx_0 / dt =  1.0*x_1
dx_1 / dt =  -0.1460697460858498*x_1+-3.9120253716489075*sin(x_0)
```



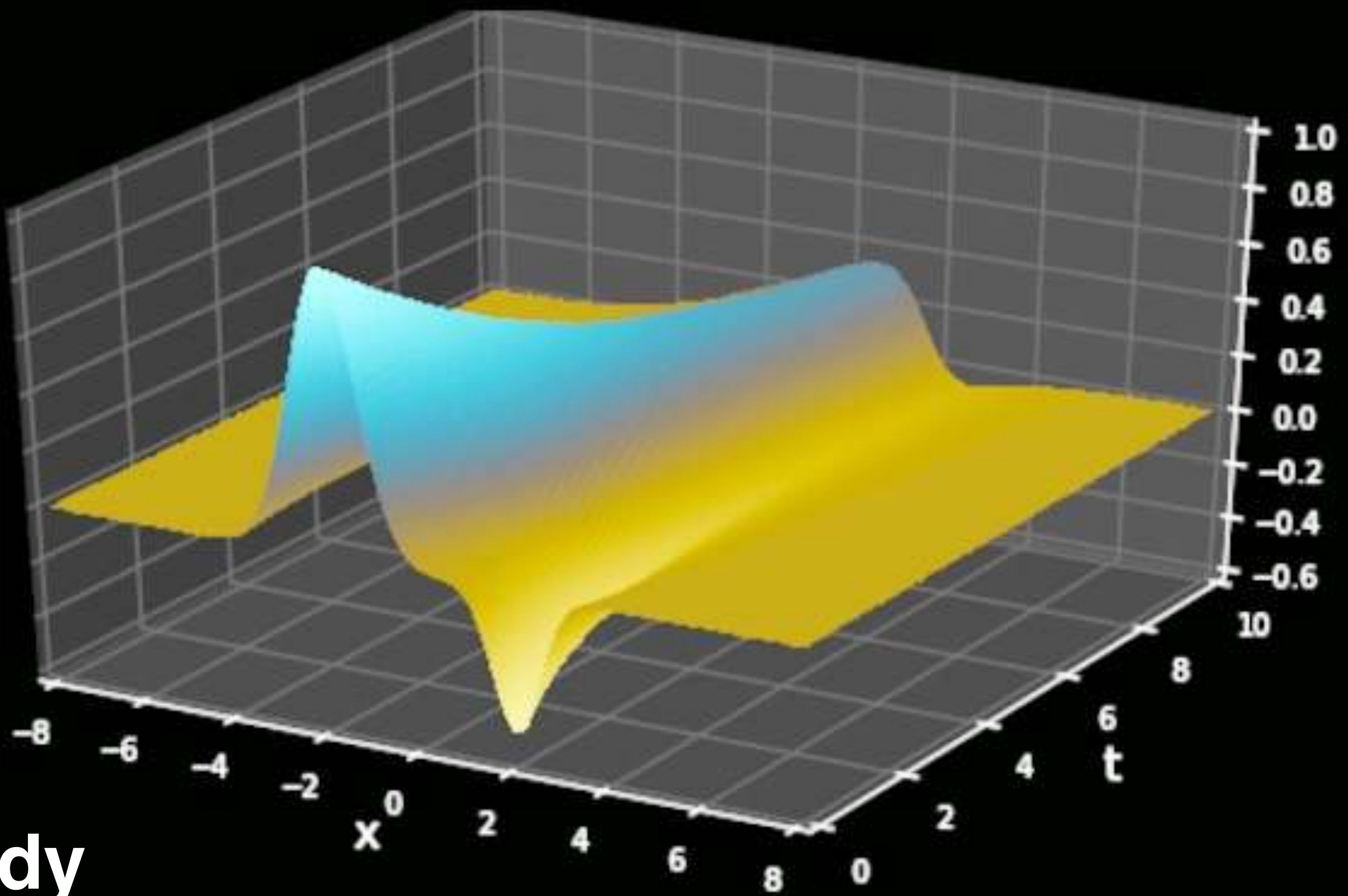
KEY CHALLENGES

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

Parametric Systems

Parametric Burgers

$$u_t + \left(1 + \frac{1}{4}\sin(t)\right) u u_x - D u_{xx} = 0$$

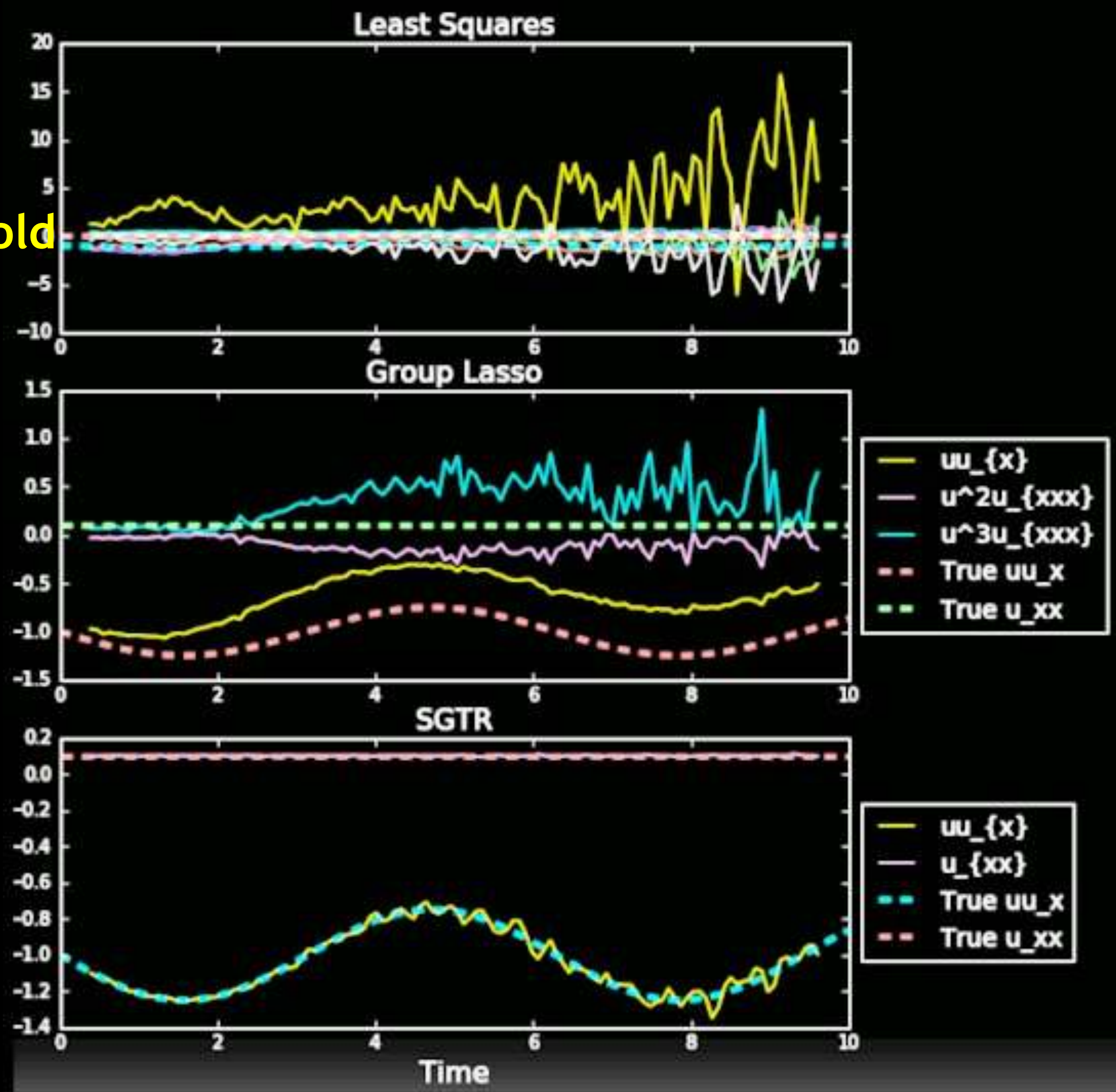


Sam Rudy

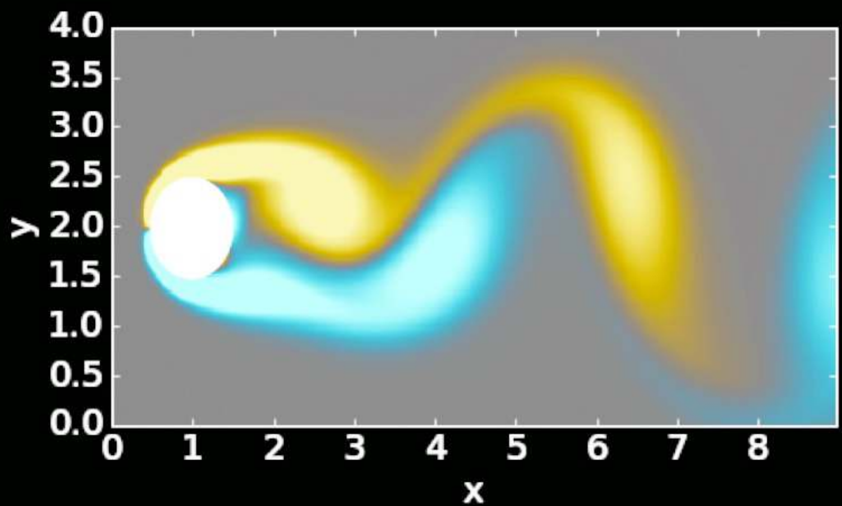
Parametric Discovery

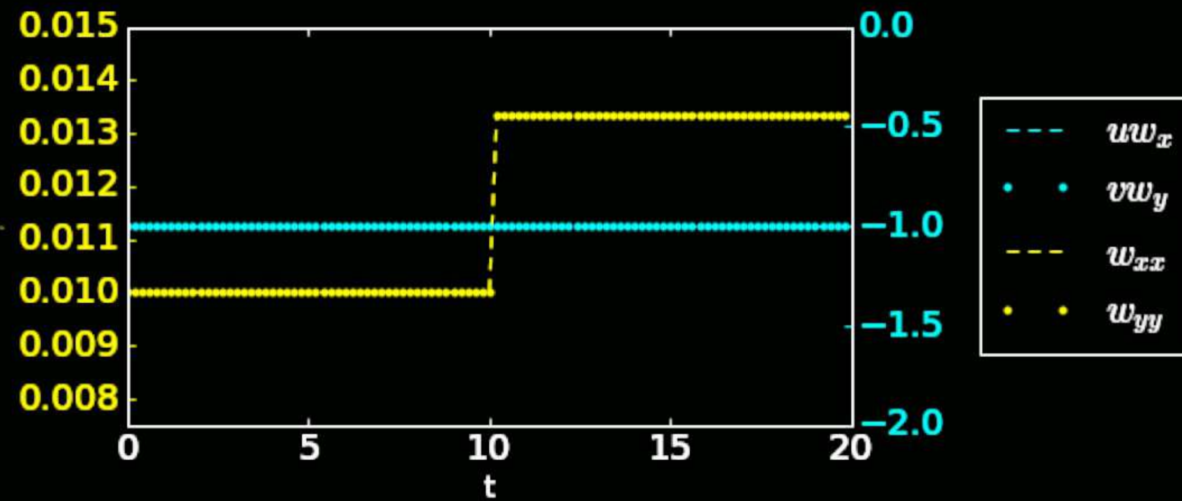
Group LASSO vs
Sequential Group Threshold
Regression (SGTR)

Our innovation: SGTR
(works amazingly well!)

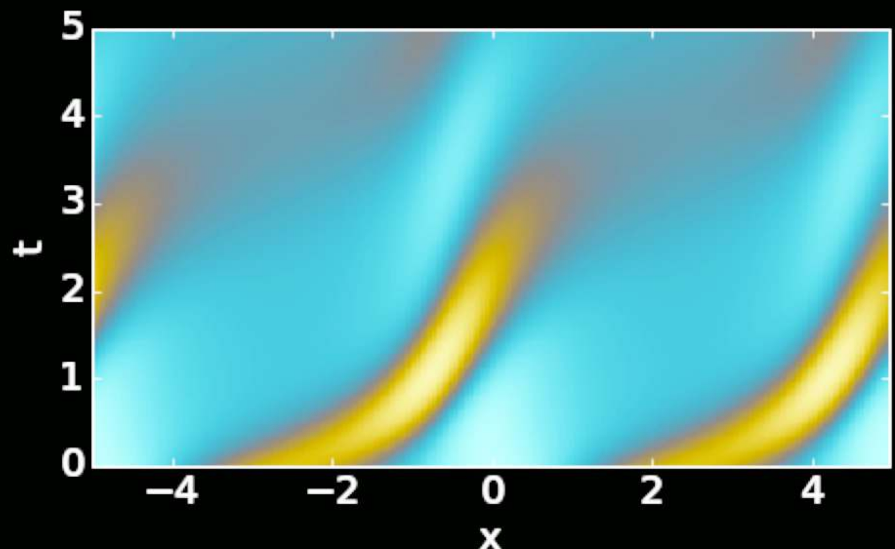


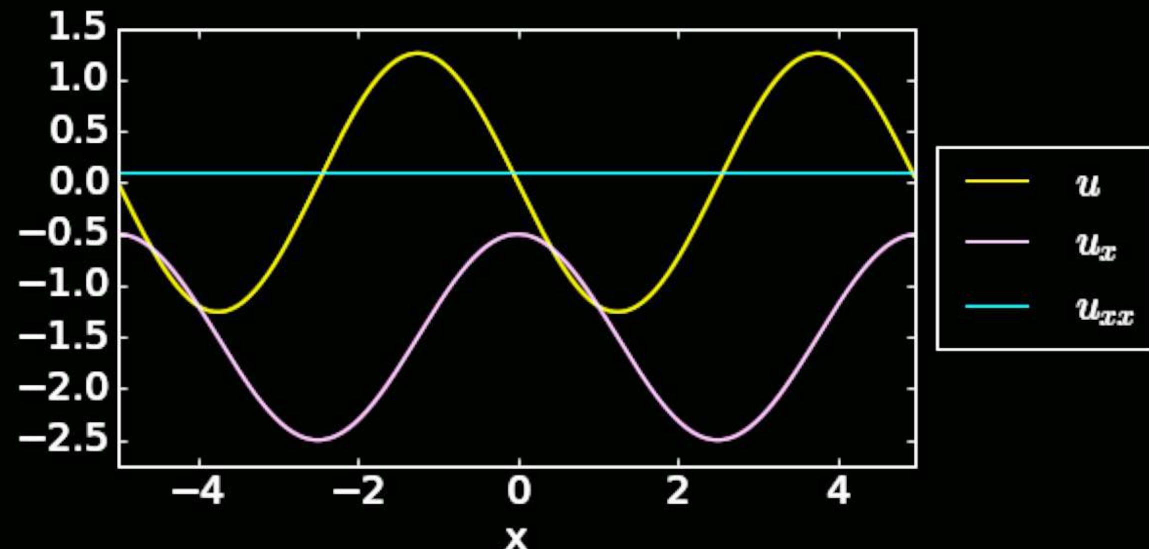
Parametric Dependence





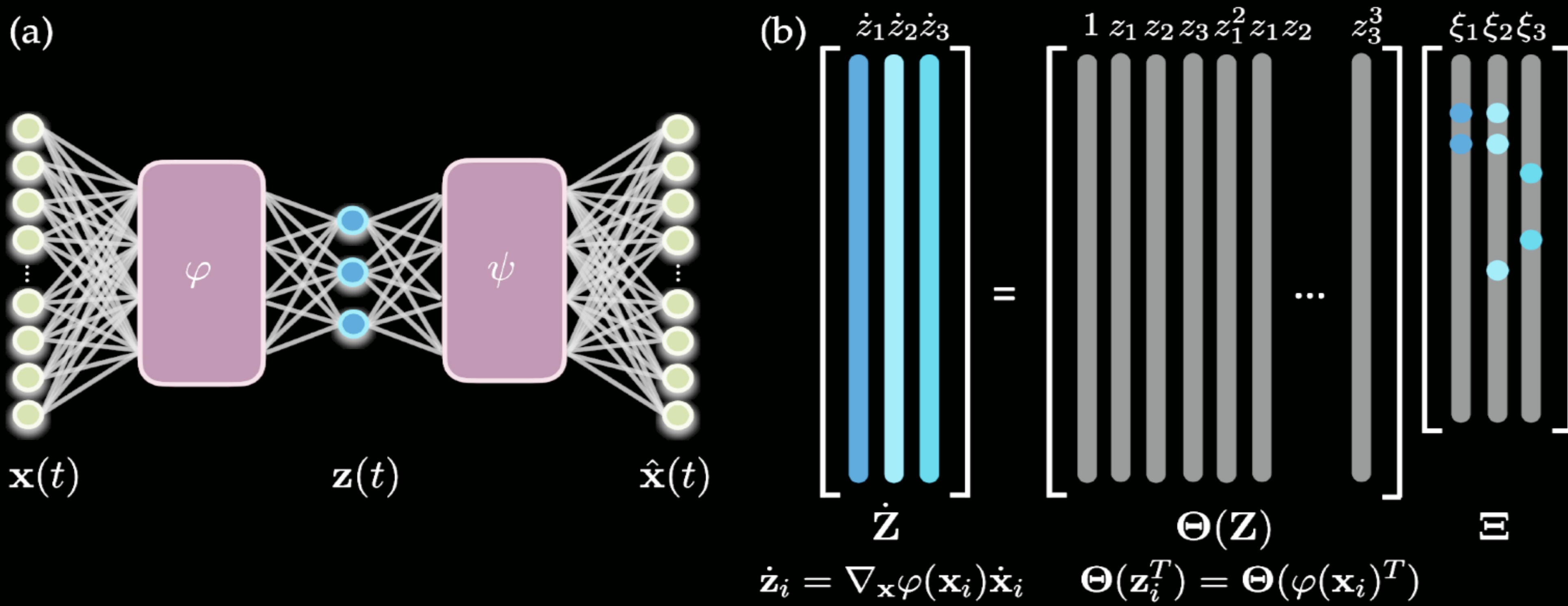
$$u_t = (c(x)u)_x + \epsilon u_{xx}$$





Coordinates
&
Dynamics

Coordinates + Dynamics

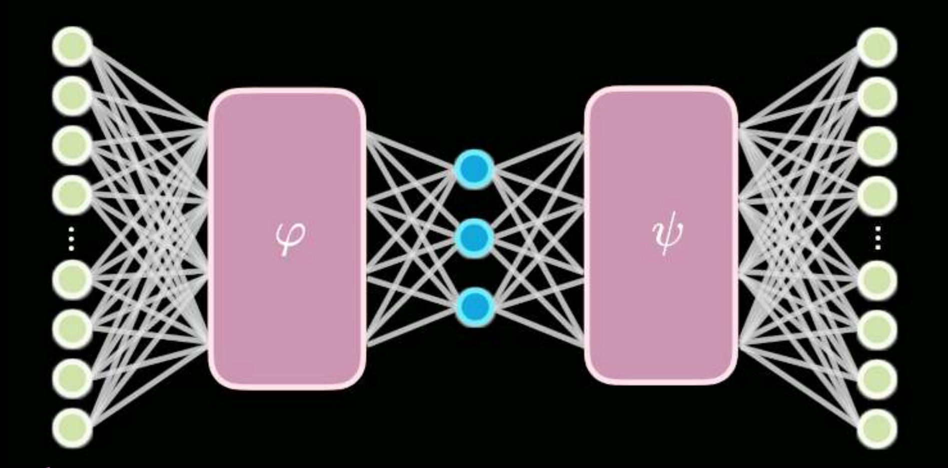


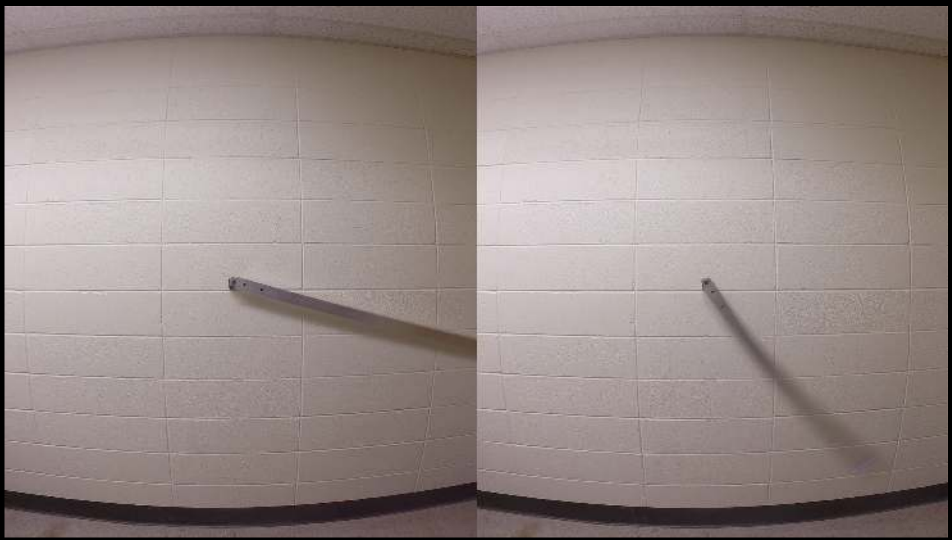
$$\underbrace{\|\mathbf{x} - \psi(\mathbf{z})\|_2^2}_{\text{reconstruction loss}} + \underbrace{\lambda_1 \left\|\dot{\mathbf{x}} - (\nabla_{\mathbf{z}}\psi(\mathbf{z}))\left(\boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right)\right\|_2^2}_{\text{SINDy loss in } \dot{\mathbf{x}}} + \underbrace{\lambda_2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

Kathleen Champion

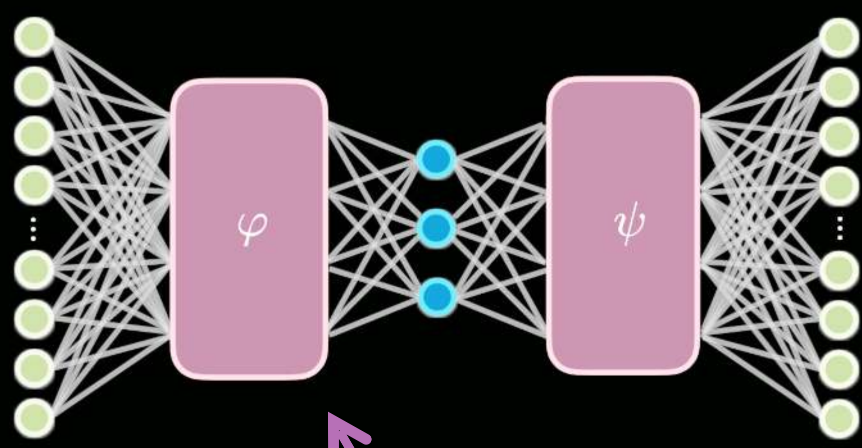
Champion, Lusch, Kutz, Brunton, PNAS (2019)
Zheng et al, SR3 – IEEE Access (2019)

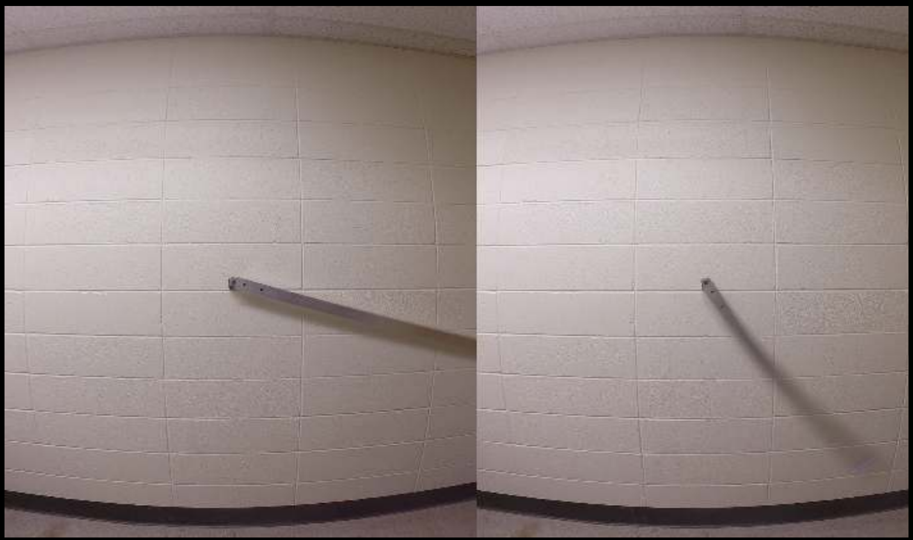
Discovery Paradigm

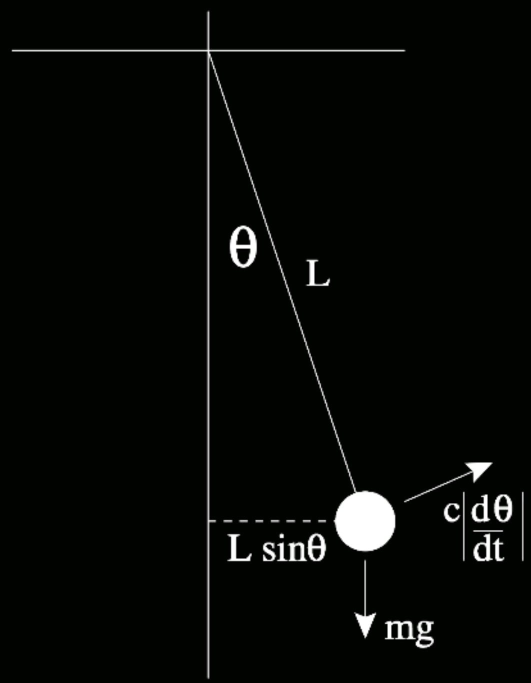




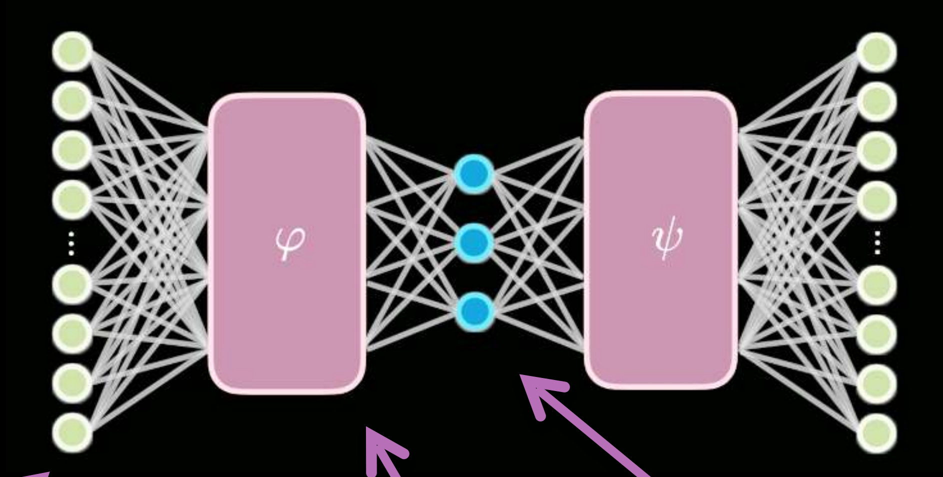
Discovery Paradigm

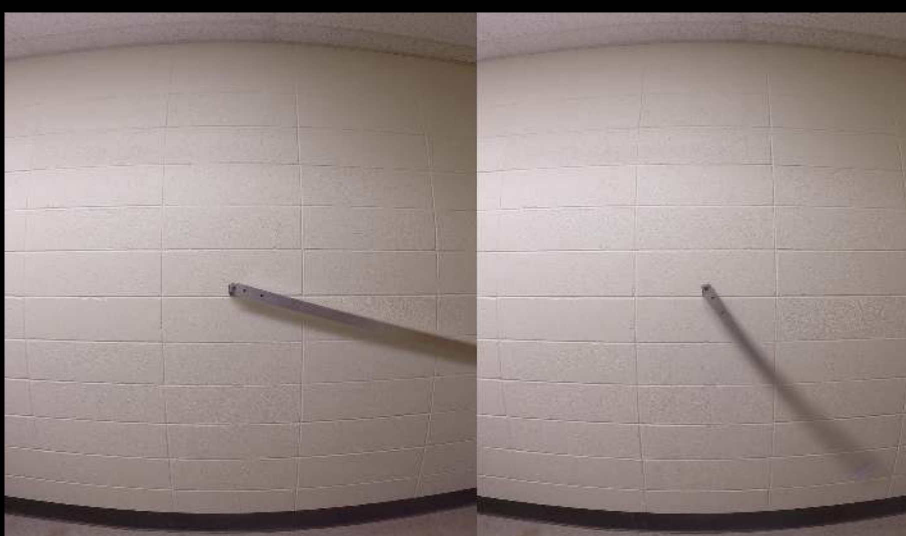


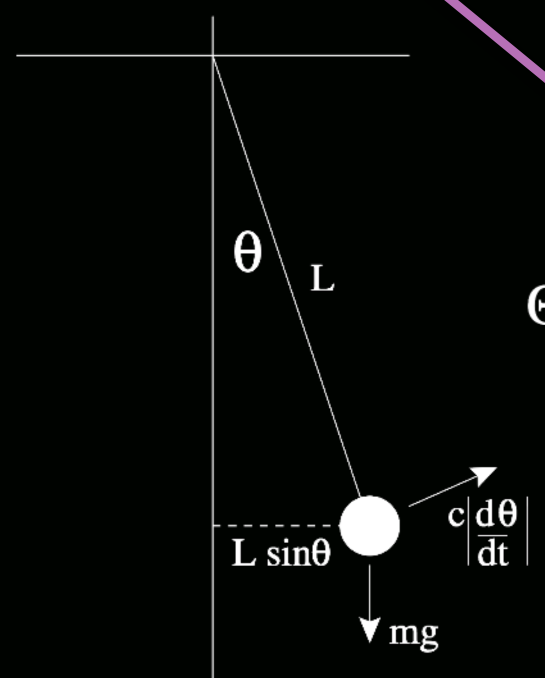




Discovery Paradigm







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

Discrepancy Modeling

Instead of model discovery from scratch…

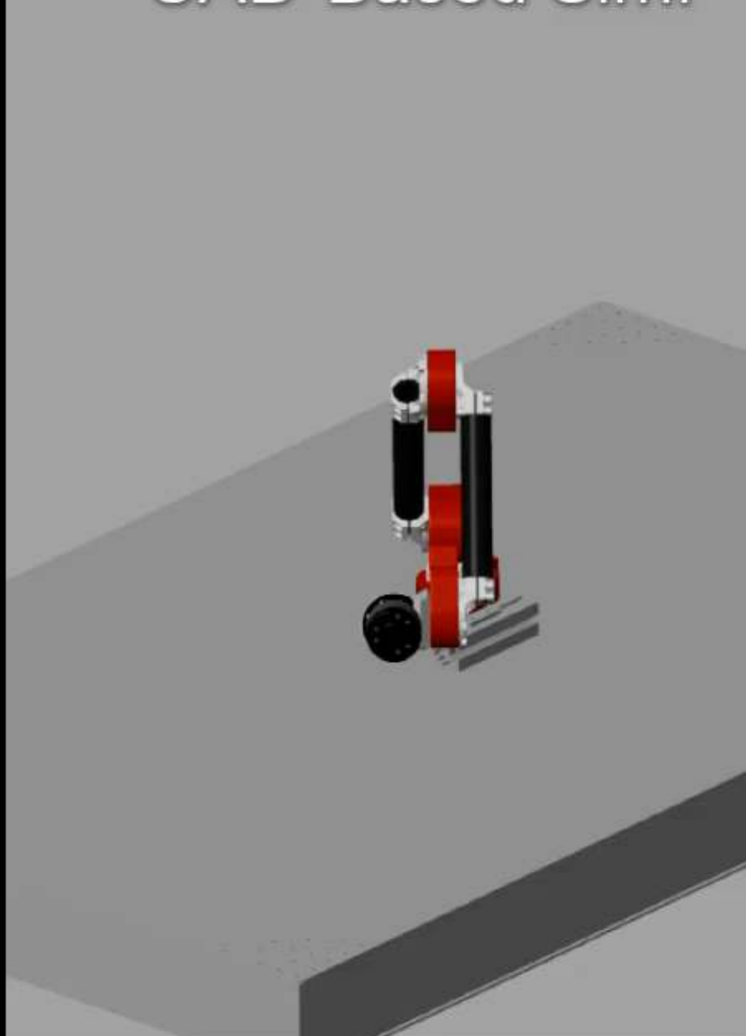
…we often start with partial knowledge of the physics

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

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

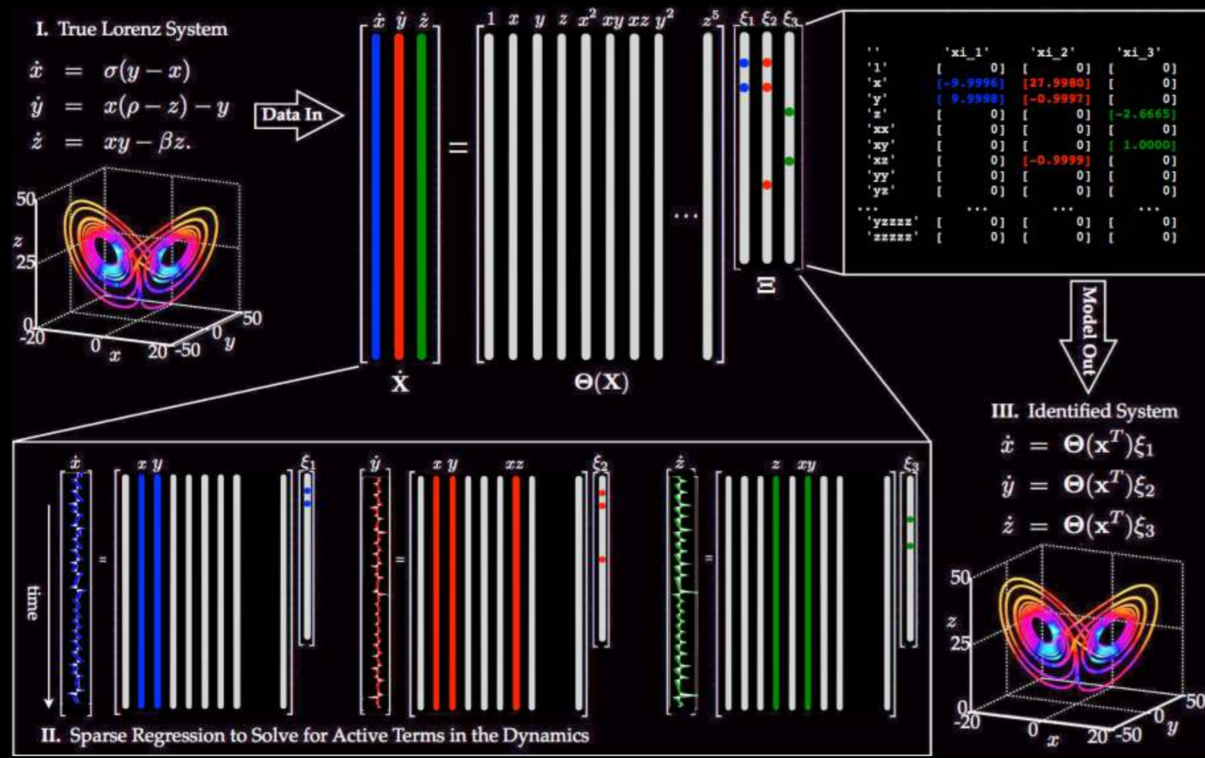
Digital Twins

CAD-Based Sim.



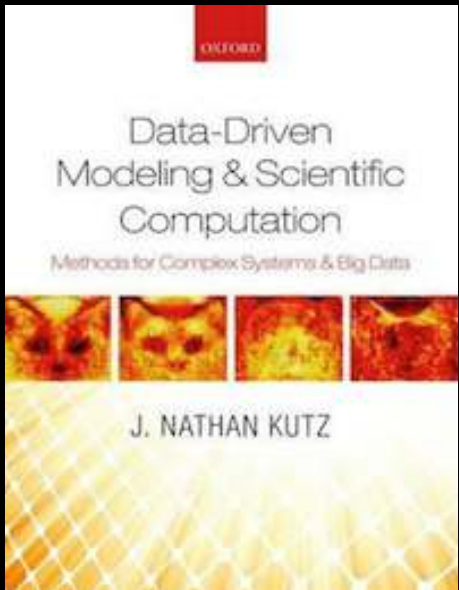
Data Collected

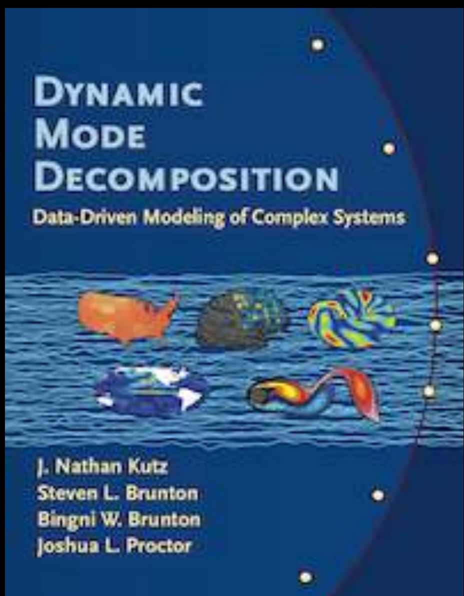
Sparse Identification of Nonlinear Dynamics (SINDy)

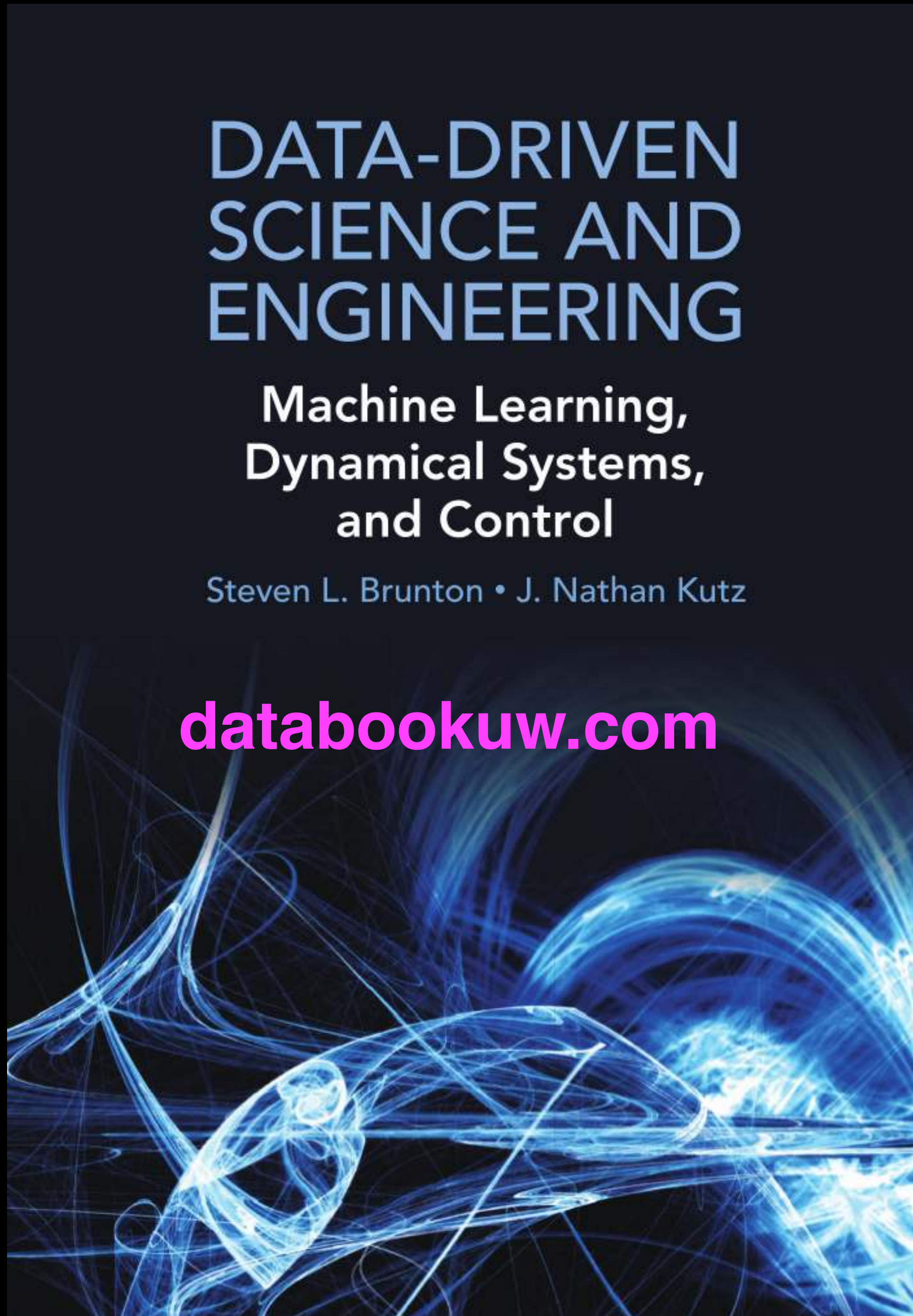


Modular, flexible and adaptive

- 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)







databookuw.com

YouTube Resources & Open Source Code