Discovering interpretable and generalizable dynamical systems from data

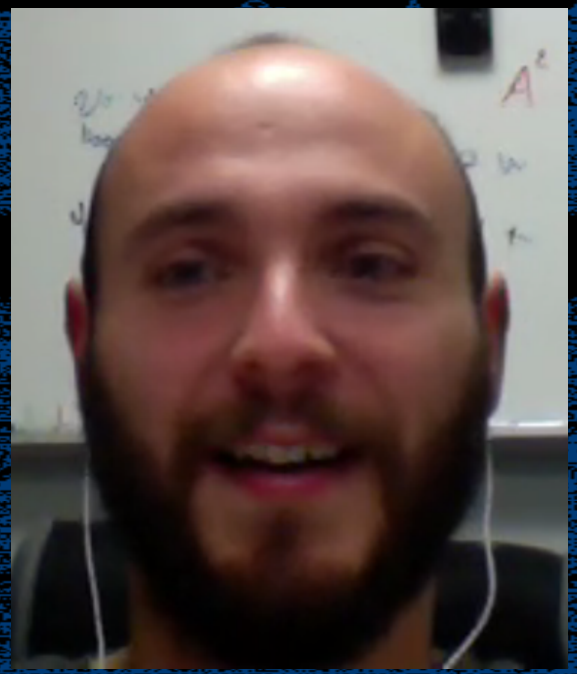
Nathan Kutz

Josh Proctor

Bing Brunton

Bernd Noack

J-Ch. Loiseau

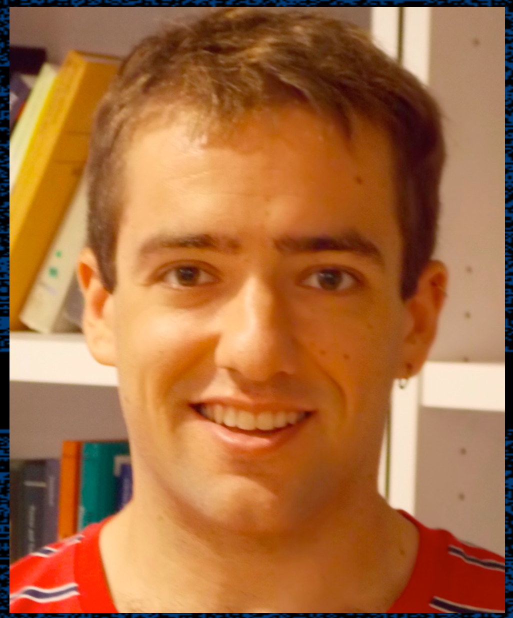


Eurika Kaiser

Bethany Lusch

Cardy Kadierdan

Sam Rudy



Jared Callaham

Ben Strom

Kathleen Champion

Often Equations are Unknown or Too complex to work with:

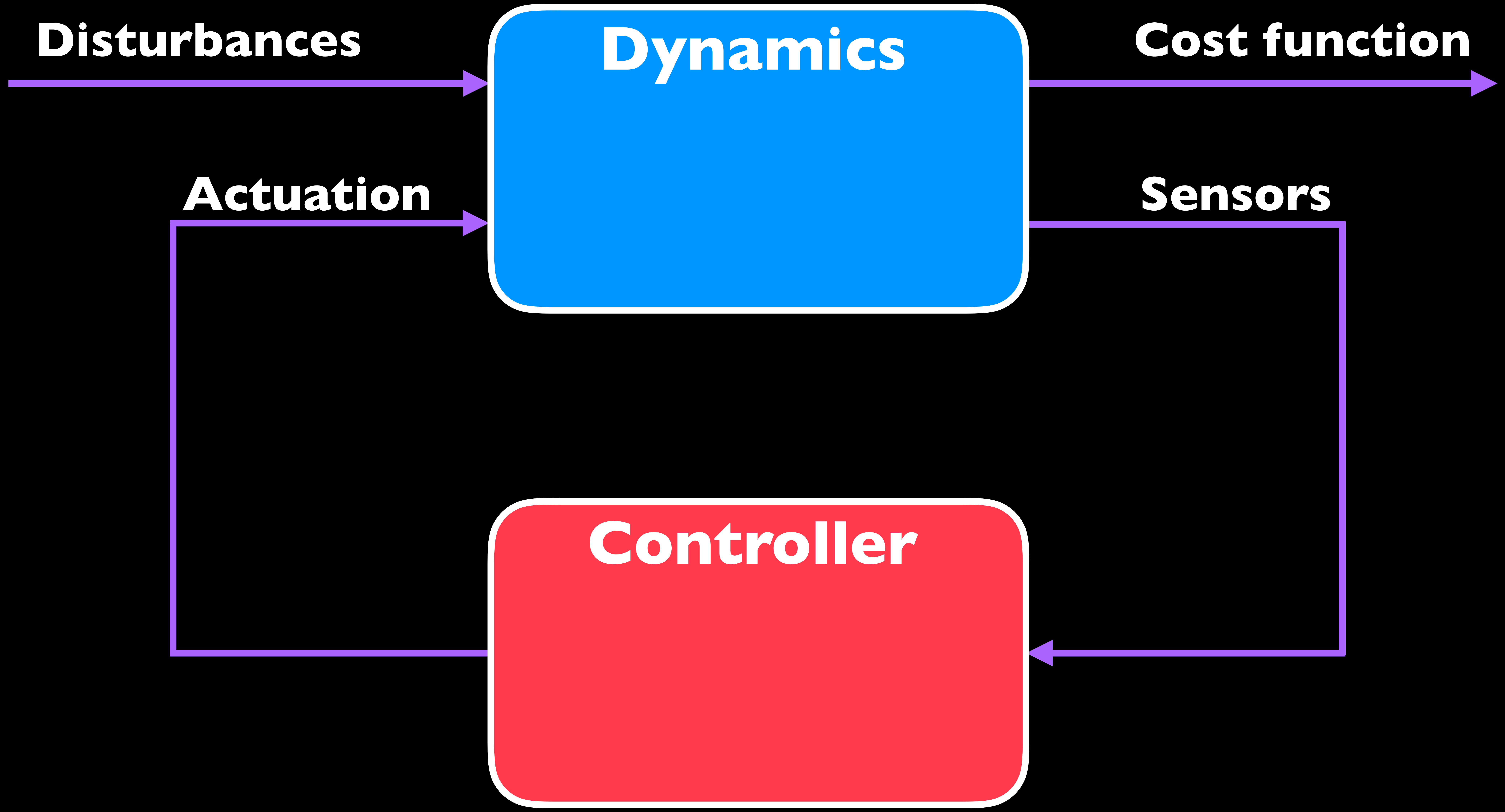
- Model discovery with machine learning
- Discover Reduced Order Models with machine learning

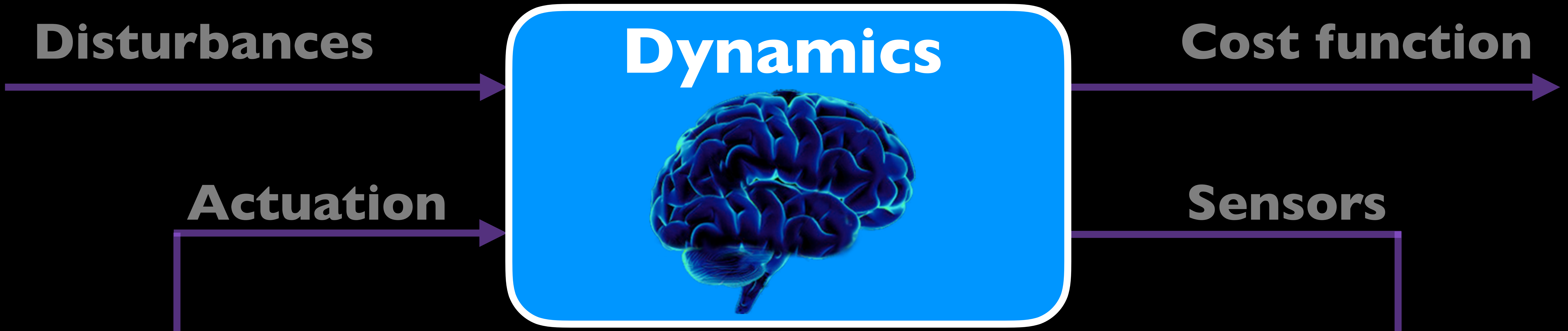
Dynamics are Nonlinear and High-Dimensional:

- Coordinate transformations to linearize dynamics
- Patterns facilitate sparse measurements

Proposed approach:

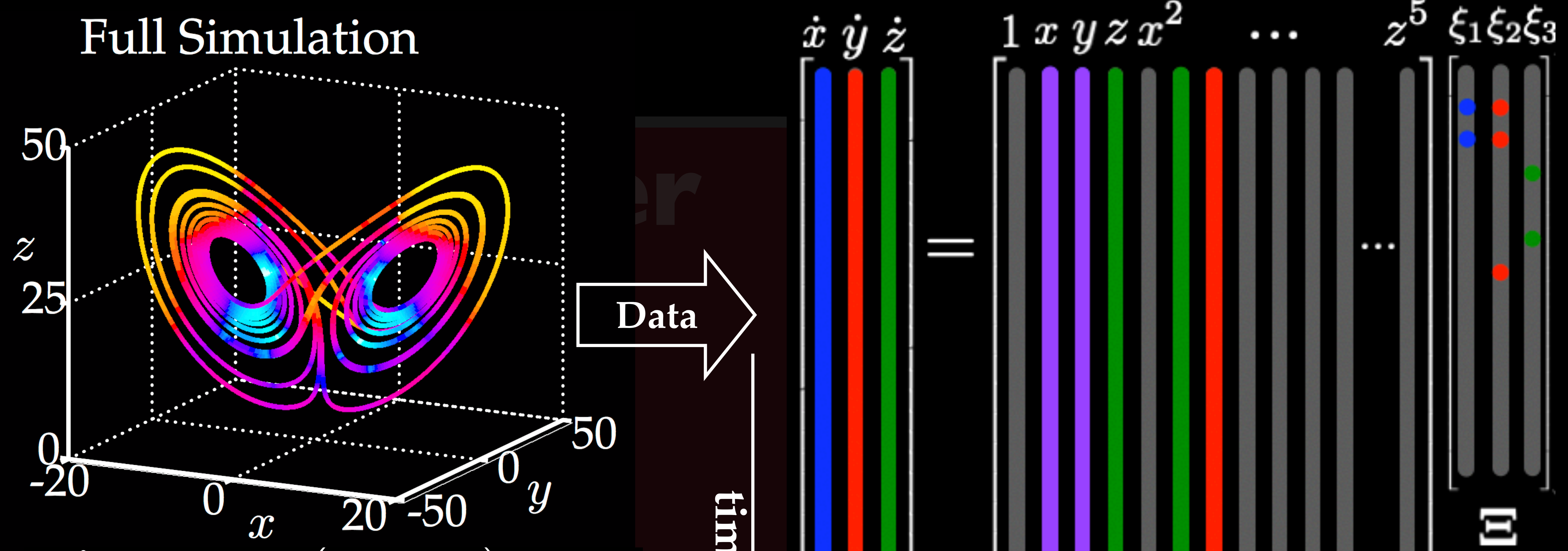
- Learn physics from data: interpretable & generalizable
- Respect known, or partially known, physics


Disturbances
Dynamics
Cost function
Actuation
Sensors
Controller



Learning Physics from Data:

- Interpretable
- Generalizable

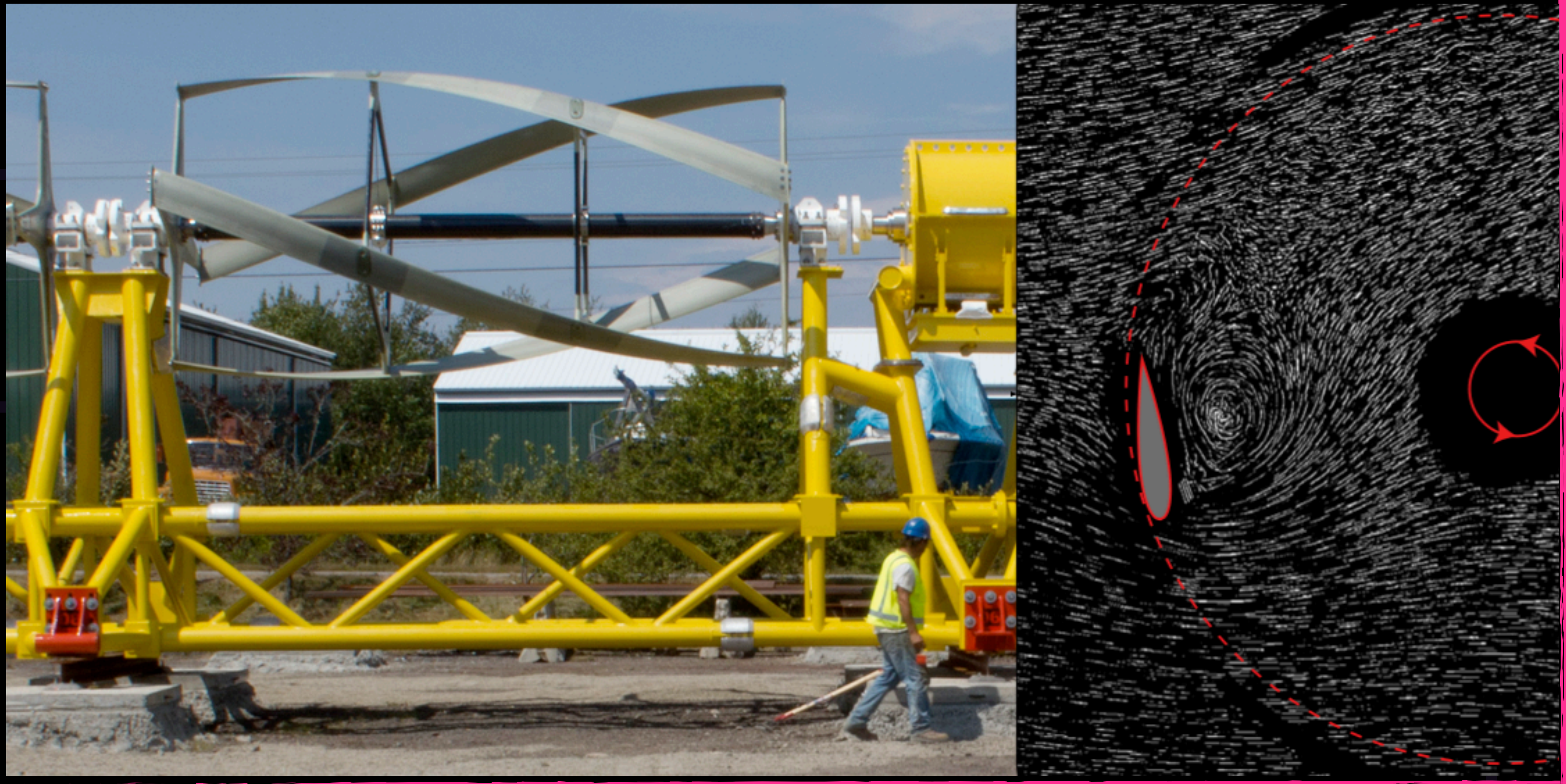


$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = x(\rho - z) - y$$
$$\dot{z} = xy - \beta z.$$

SLB, Proctor, Kutz, *PNAS* 2016.

Control and Optimization:

- Nonlinear
- Use ML

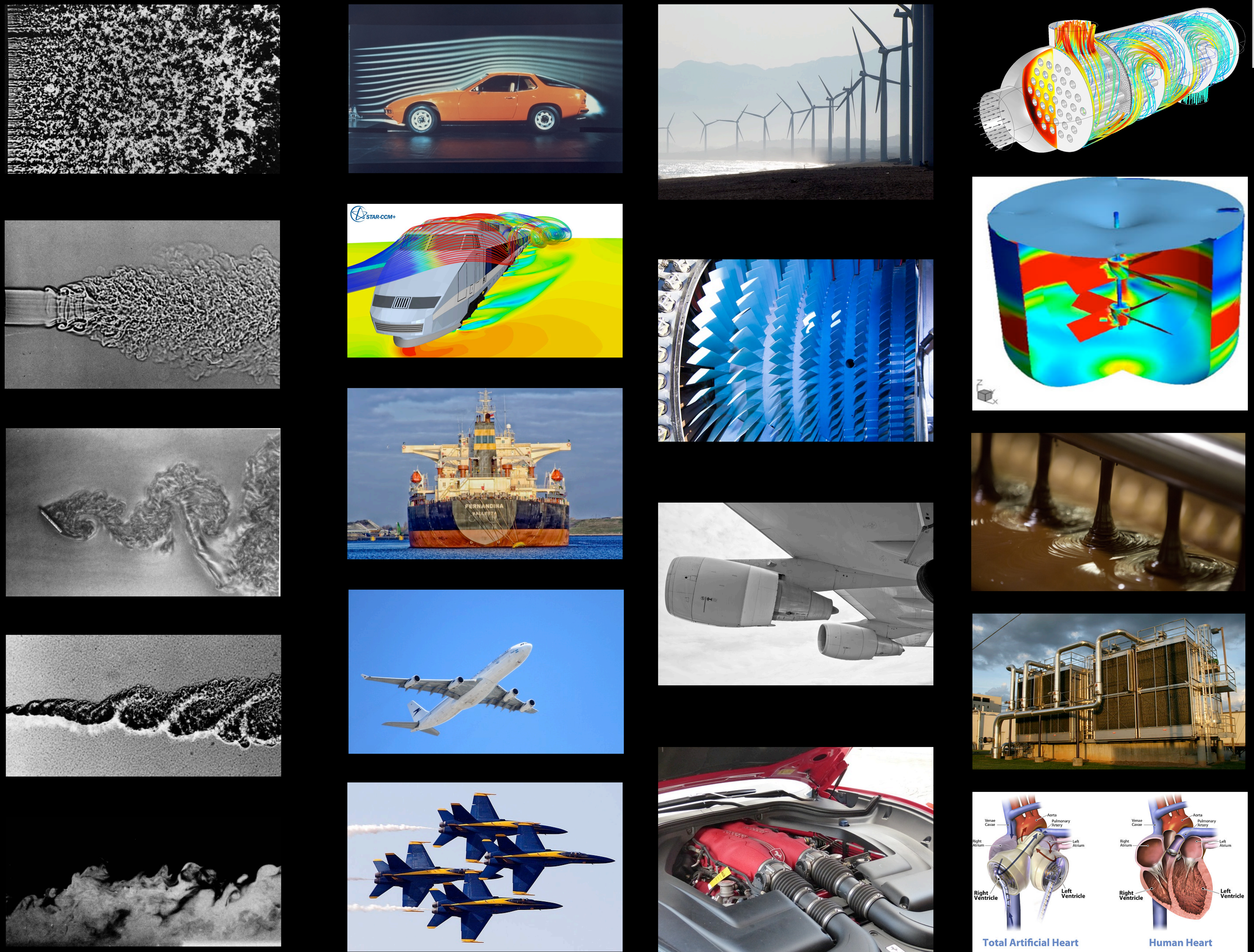


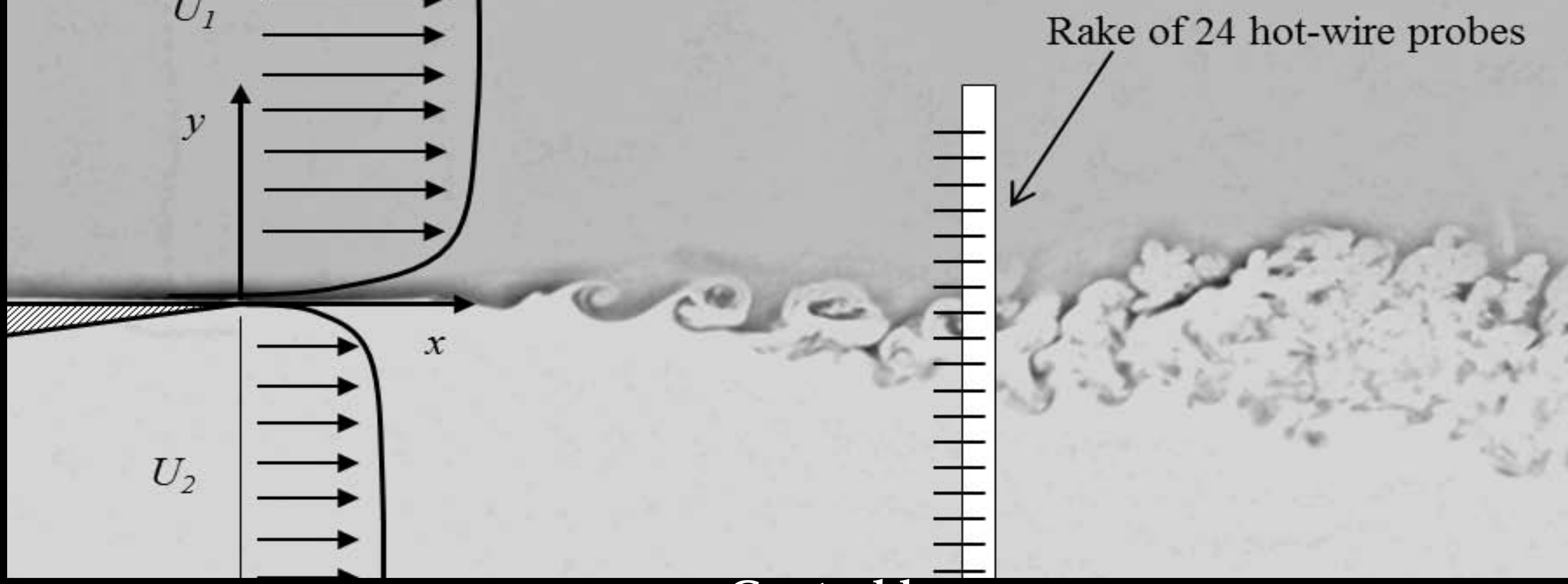
Strom, SLB, Polagye, *Nature Energy* 2017.

Controller

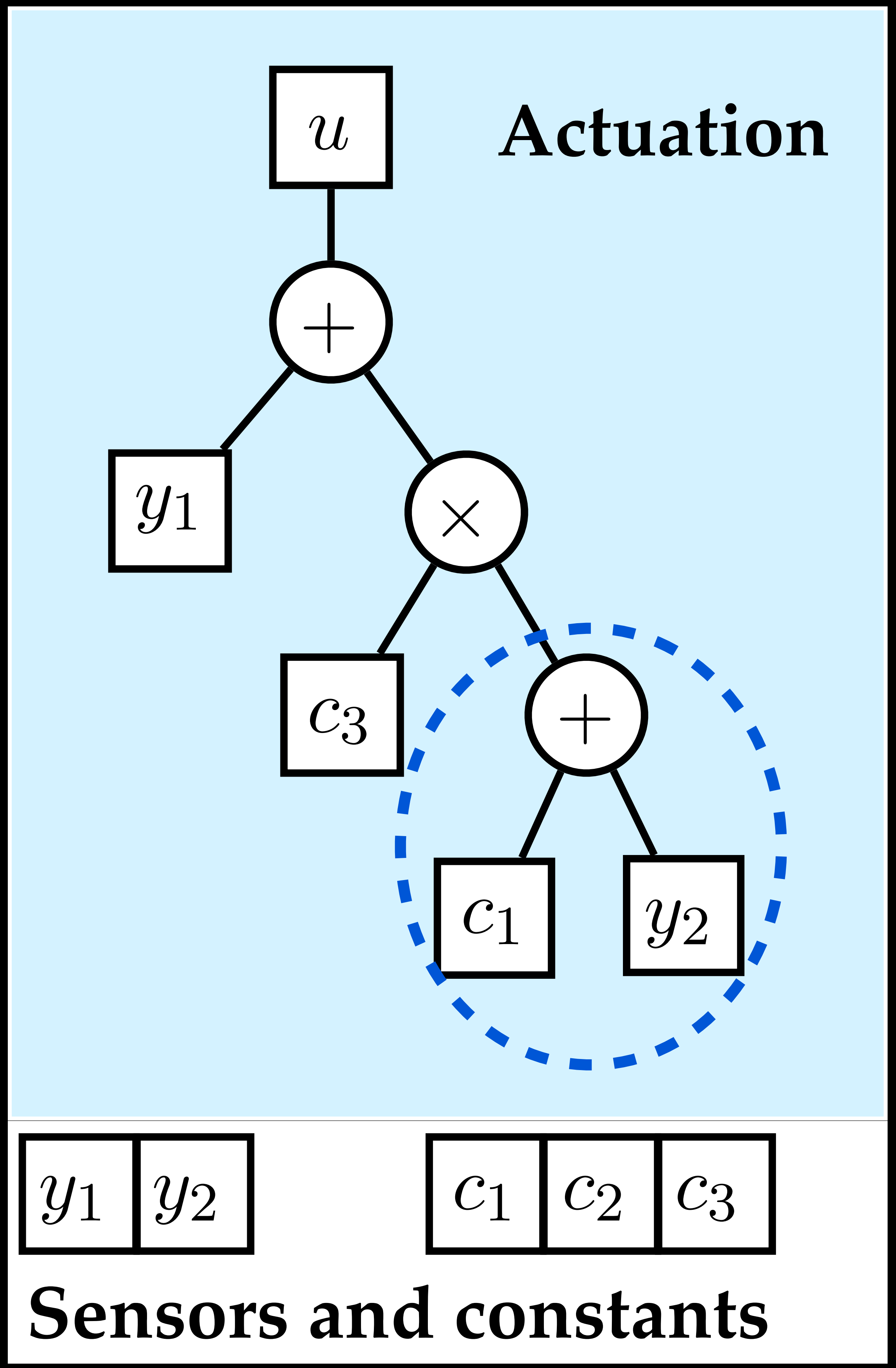
FLOW CONTROL

SLB, Noack, *AMR*, 2015
Duriez, SLB, Noack, Springer *2016*



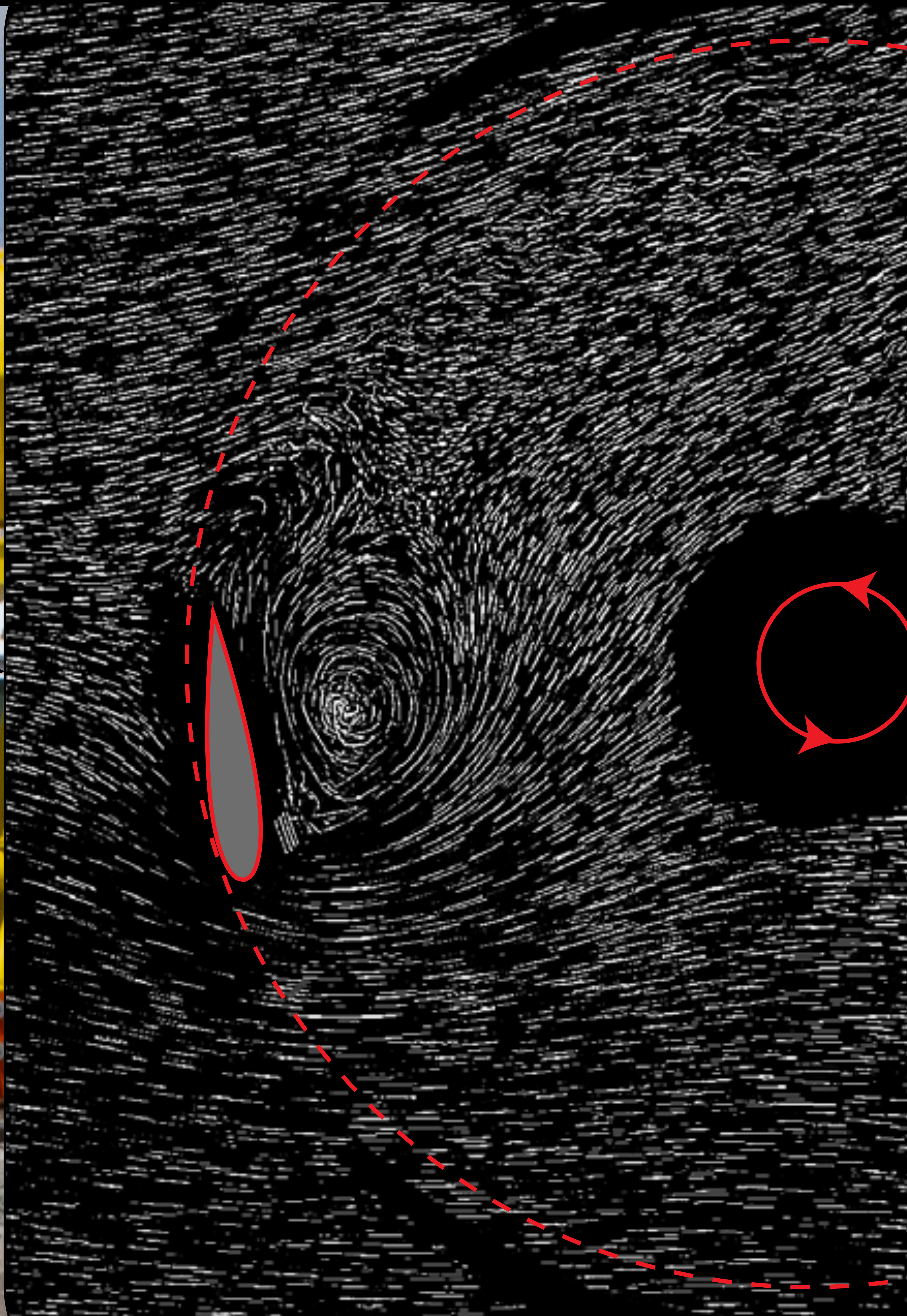


Control law



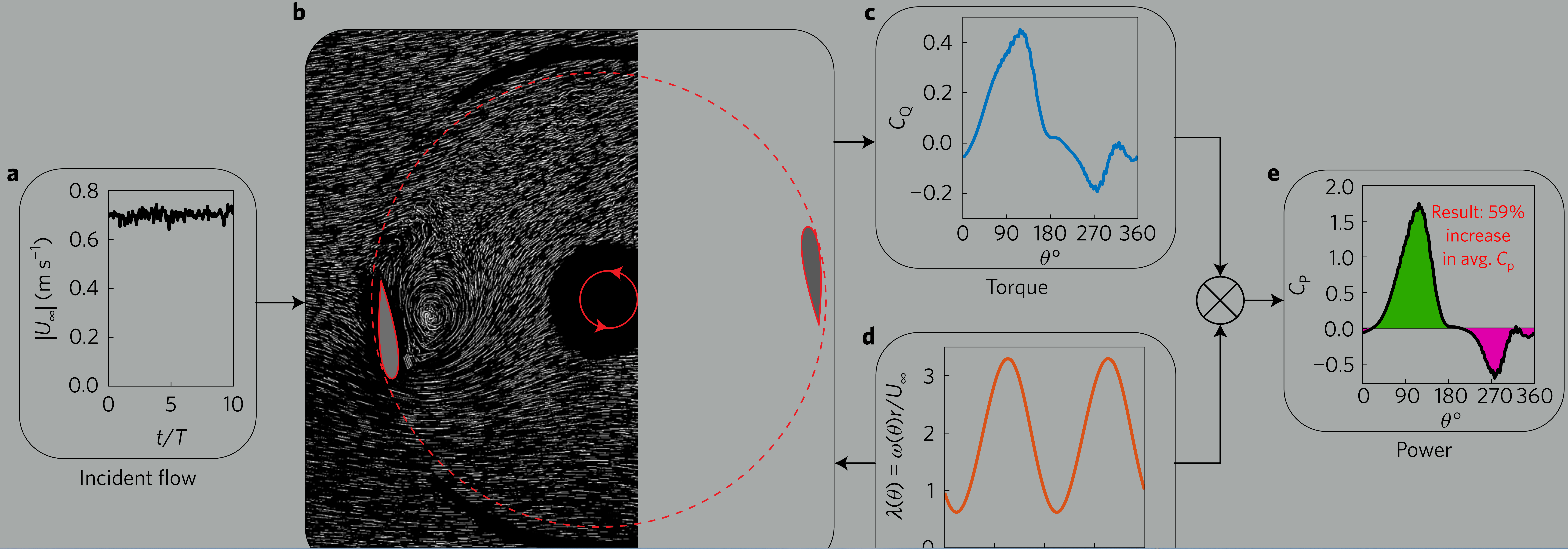
59% Power increase in lab-scale cross-flow turbine experiment using gradient simplex optimization

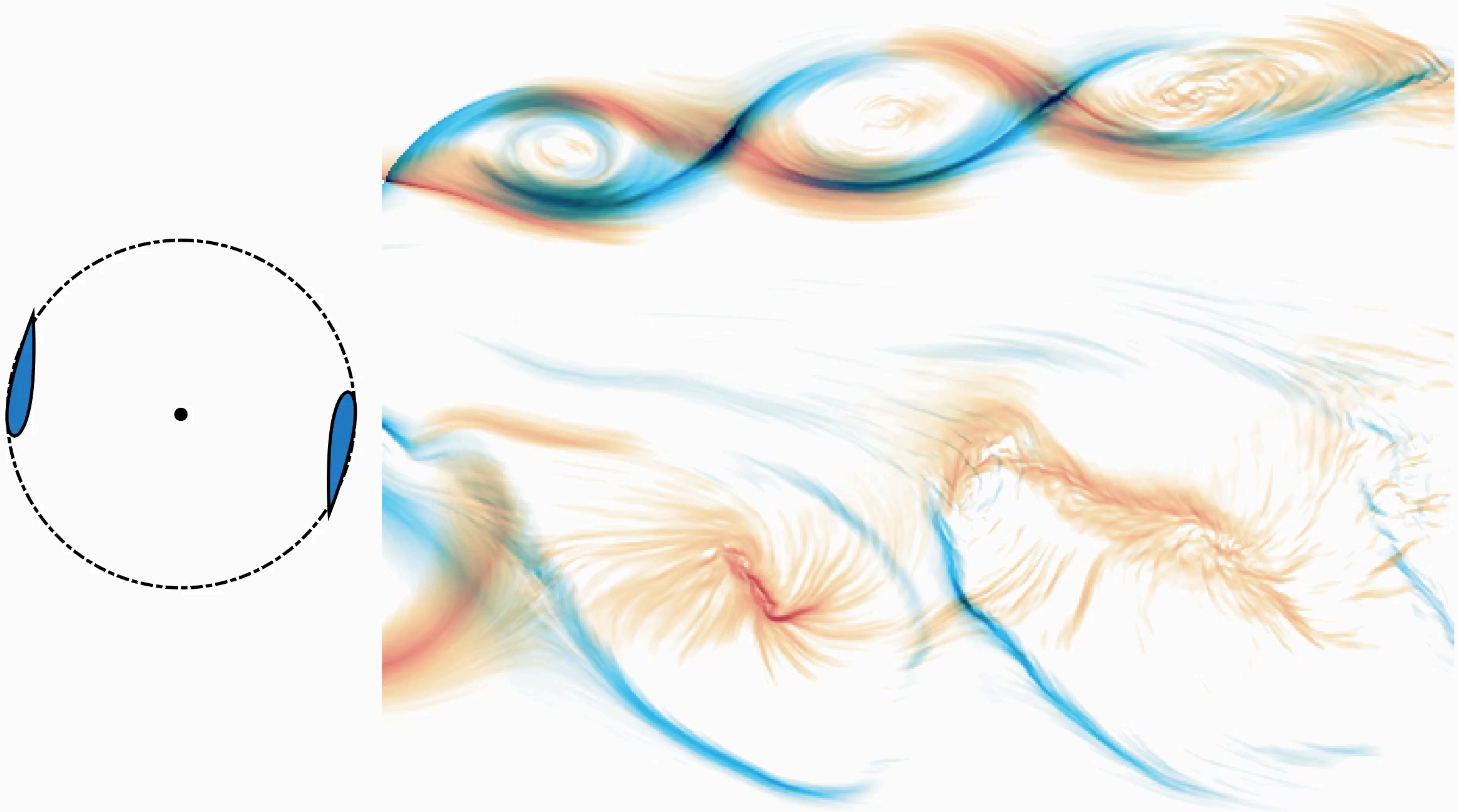
Nature Energy, 2017
Strom, SLB, Polagye



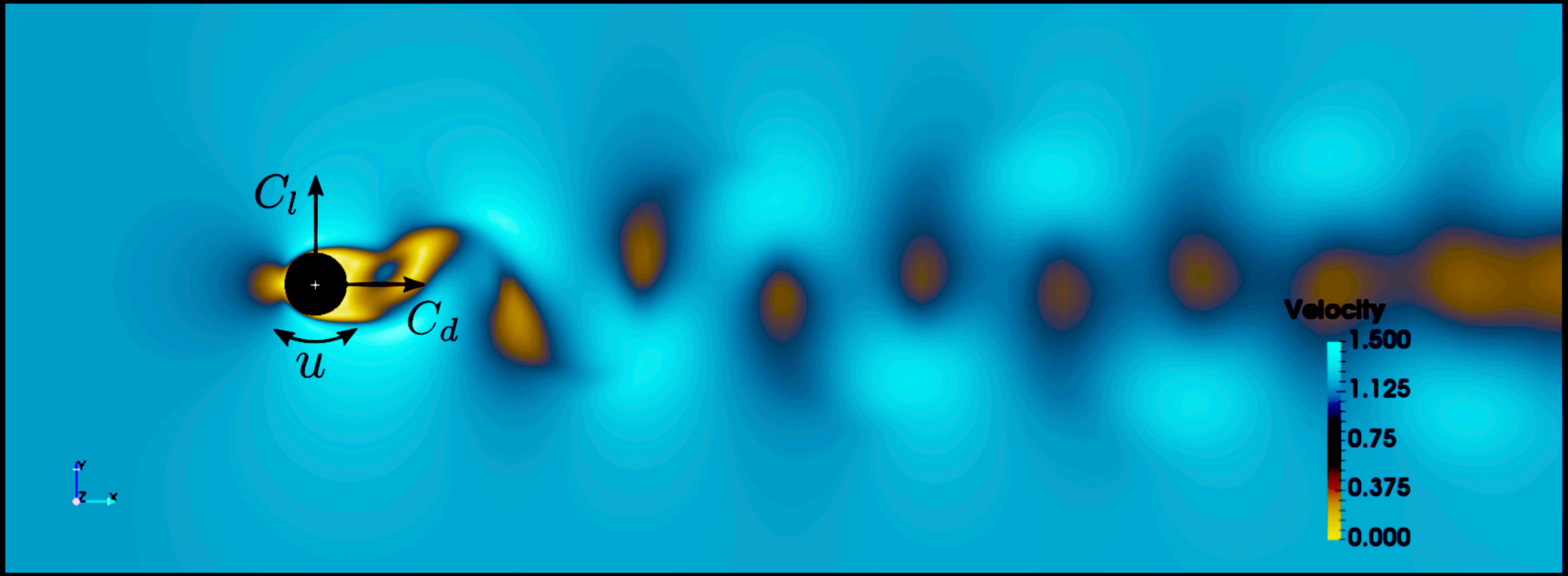
59% Power increase in lab-scale cross-flow turbine experiment using gradient simplex optimization

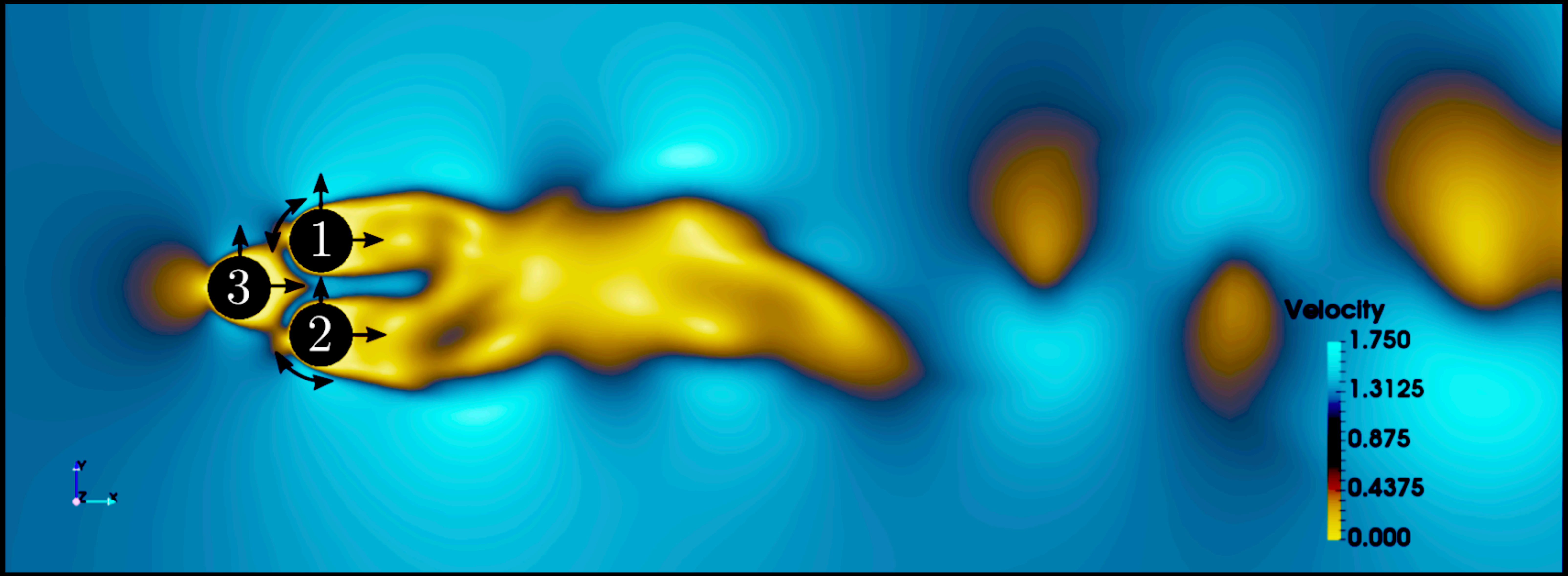
Nature Energy, 2017
Strom, SLB, Polagye





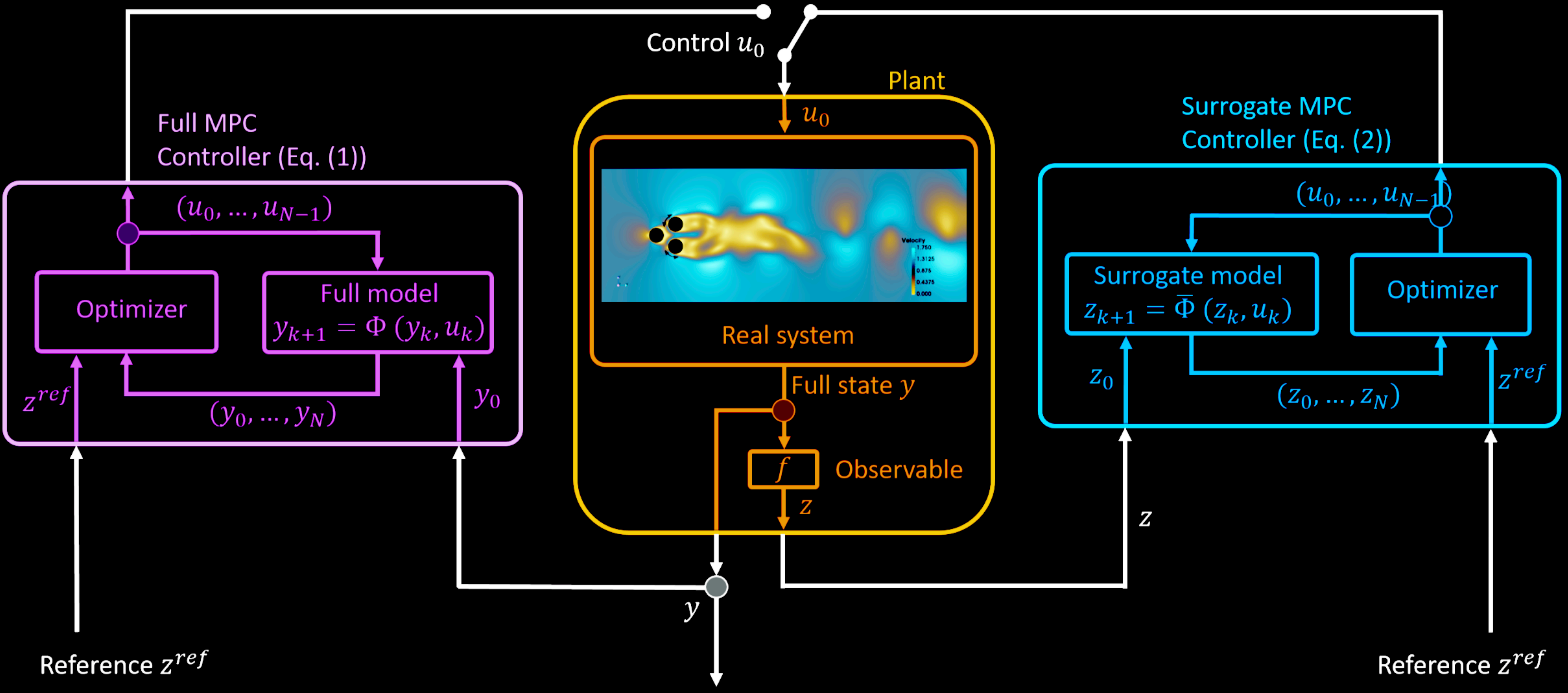
Deep MPC for Fluid Flow Control





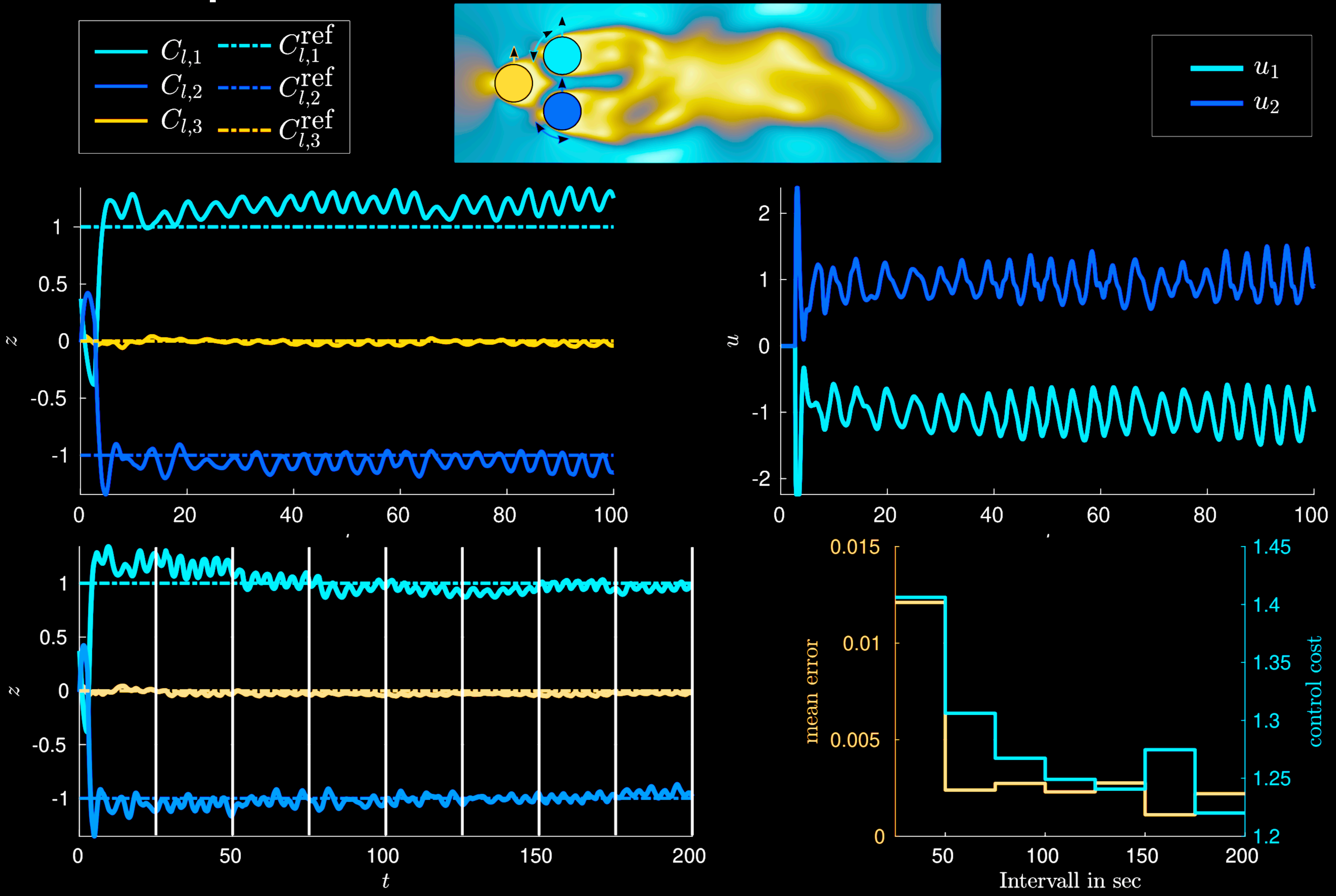
Bieker, Peitz, Brunton, Kutz, Dellnitz, *arXiv*, 2019.

Deep MPC for Fluid Flow Control



Bieker, Peitz, Brunton, Kutz, Dellnitz, *arXiv*, 2019.

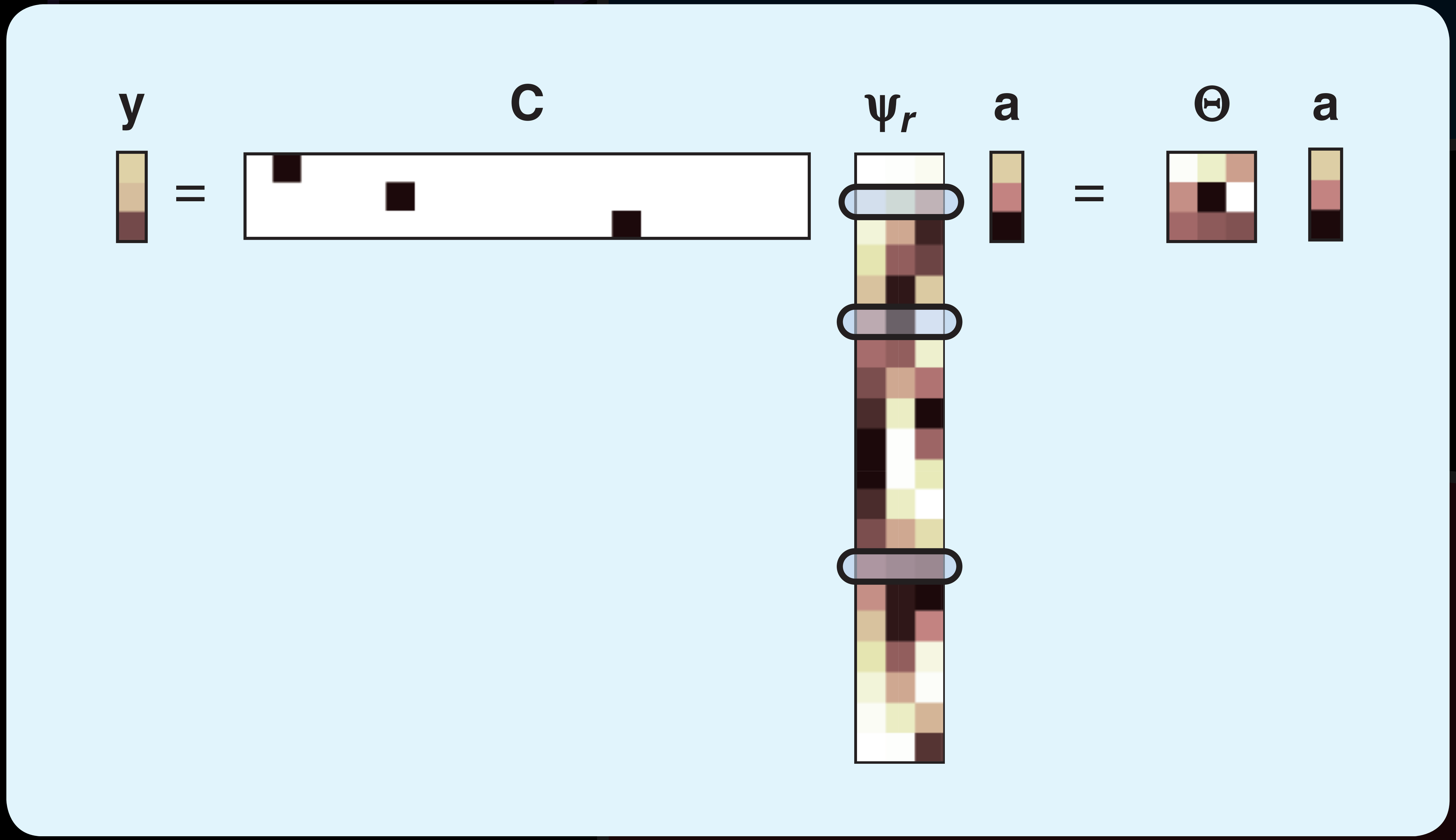
Deep MPC for Fluid Flow Control



Bieker, Peitz, Brunton, Kutz, Dellnitz, *arXiv*, 2019.

Sparse Sensor Optimization:

- **Patterns facilitate sparse sampling**



B. Brunton, SLB, Proctor, Kutz, *SIAM SIAP* 2016.
Manohar, B Brunton, Kutz, SLB, *IEEE CSM* 2017.

Cost function

Sensors

BIO-INSPIRED

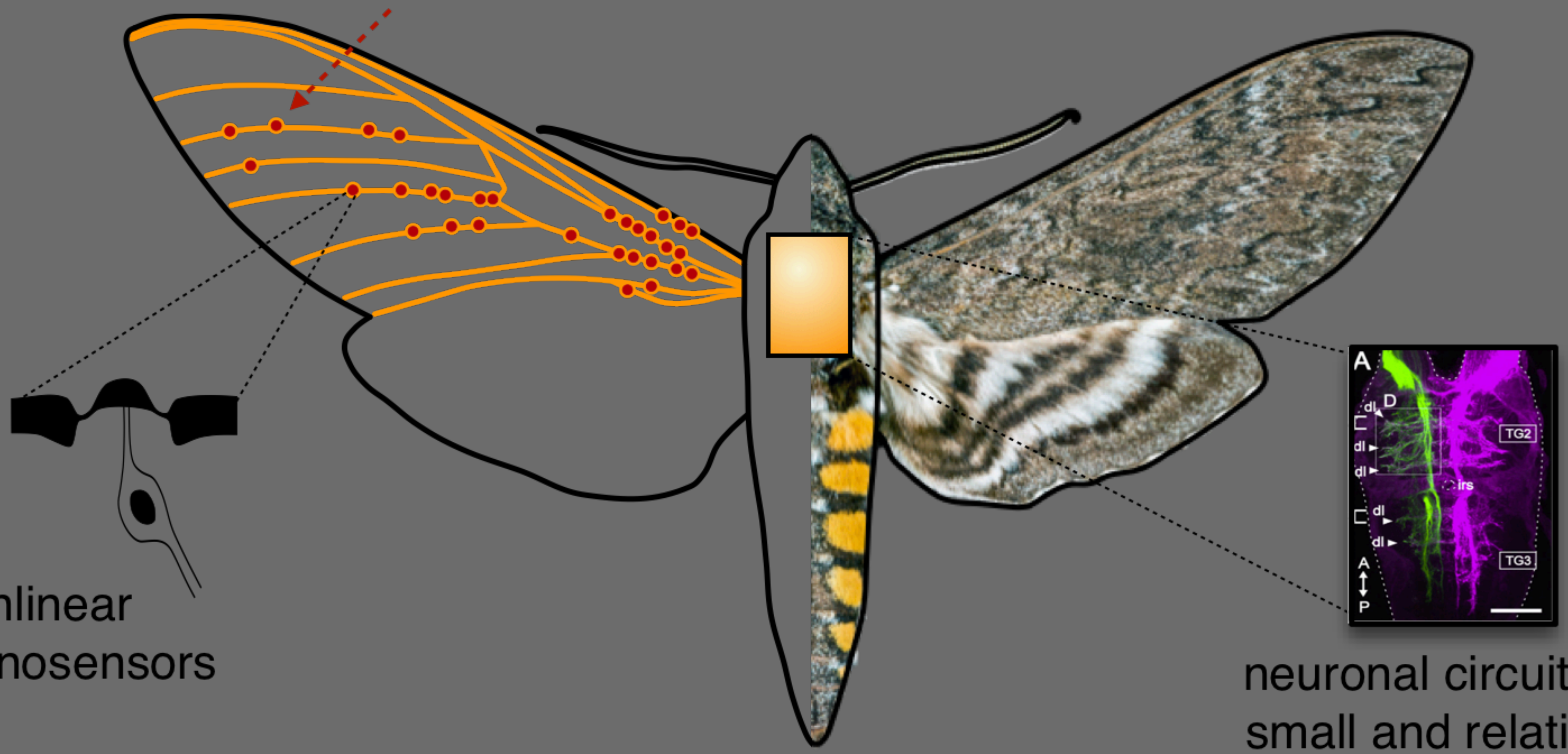
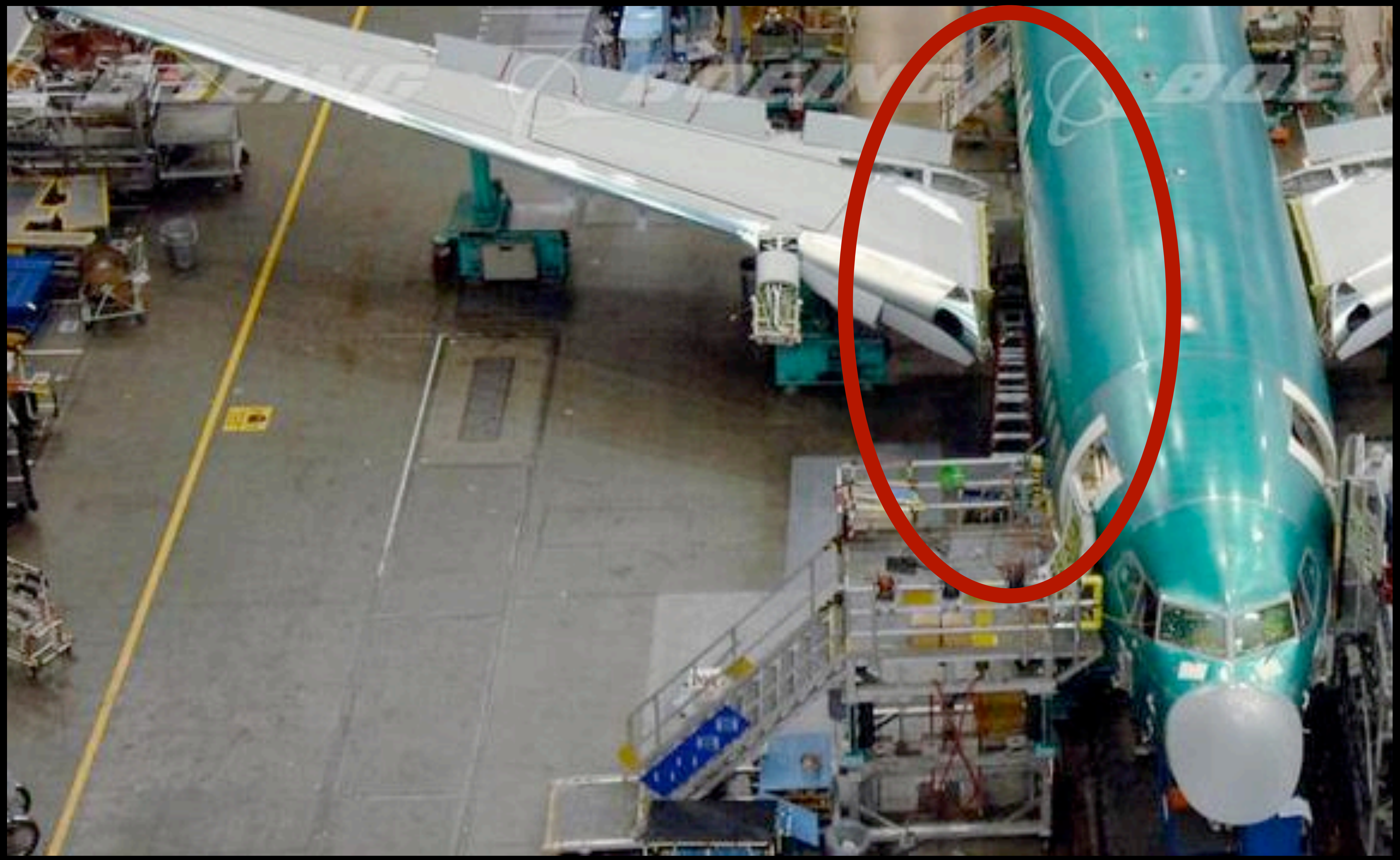


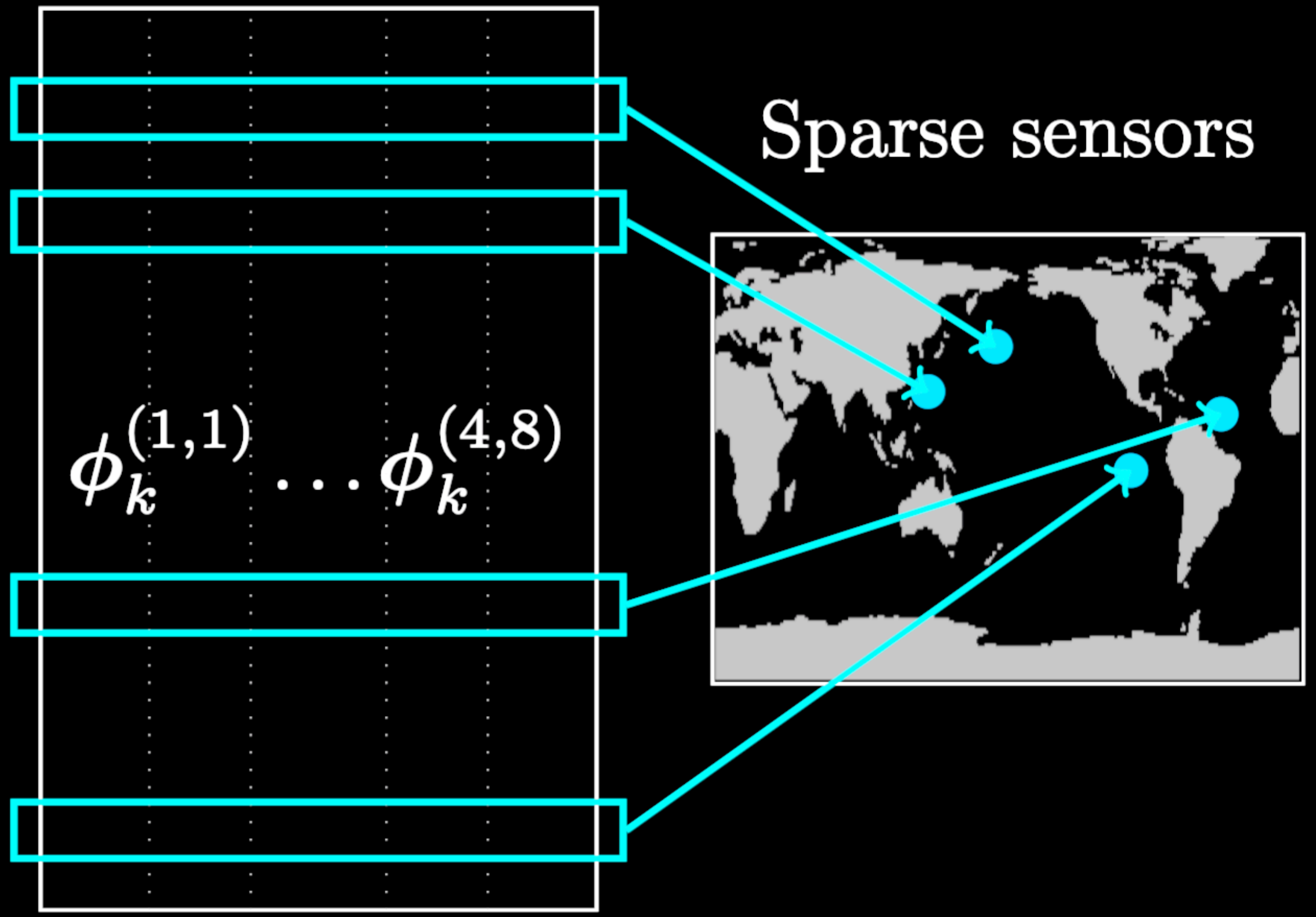
diagram adapted from Ali Weber

anatomy adapted from Ando et al. 2011.

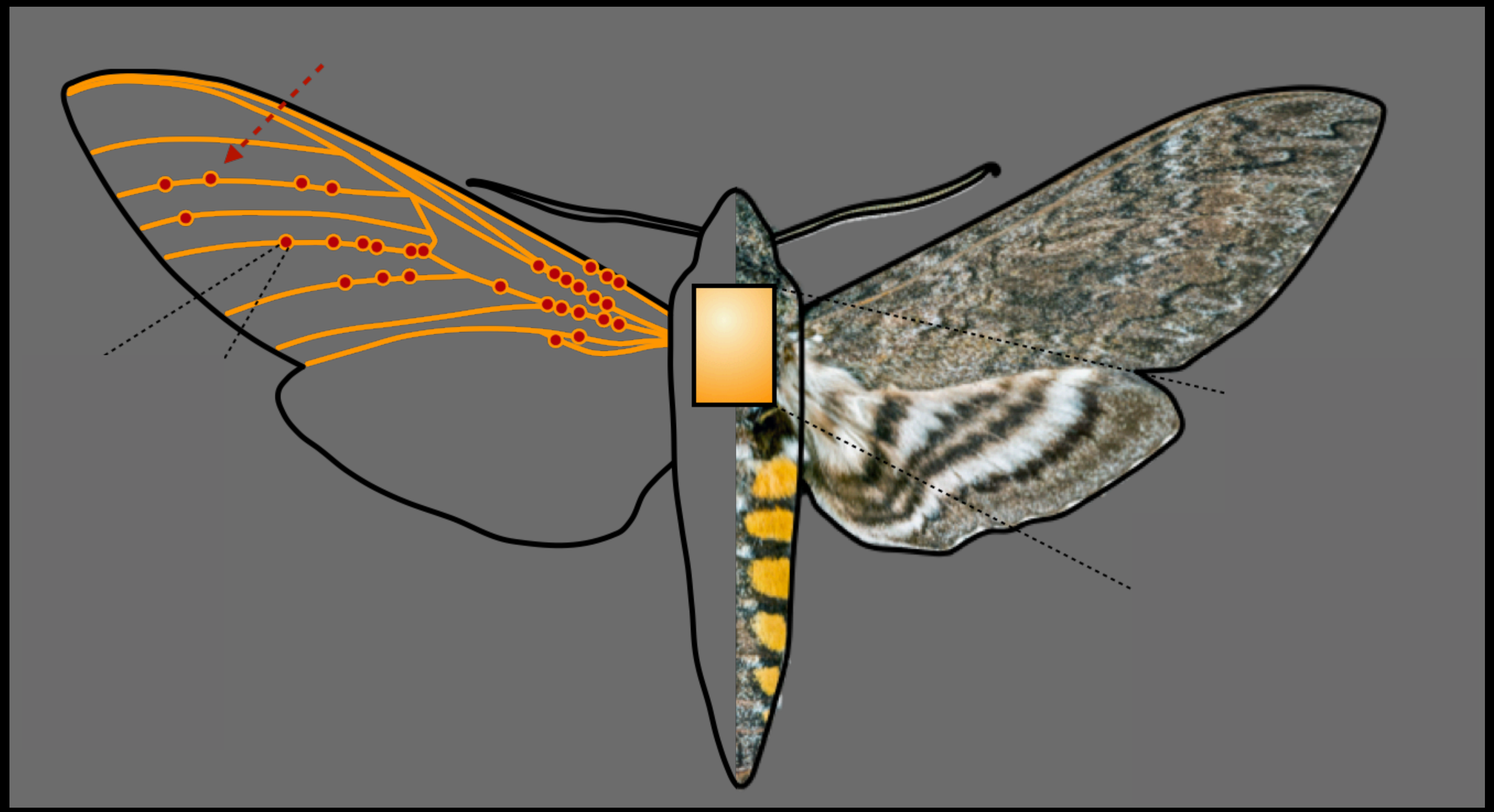
Applications of Sparse Sensing



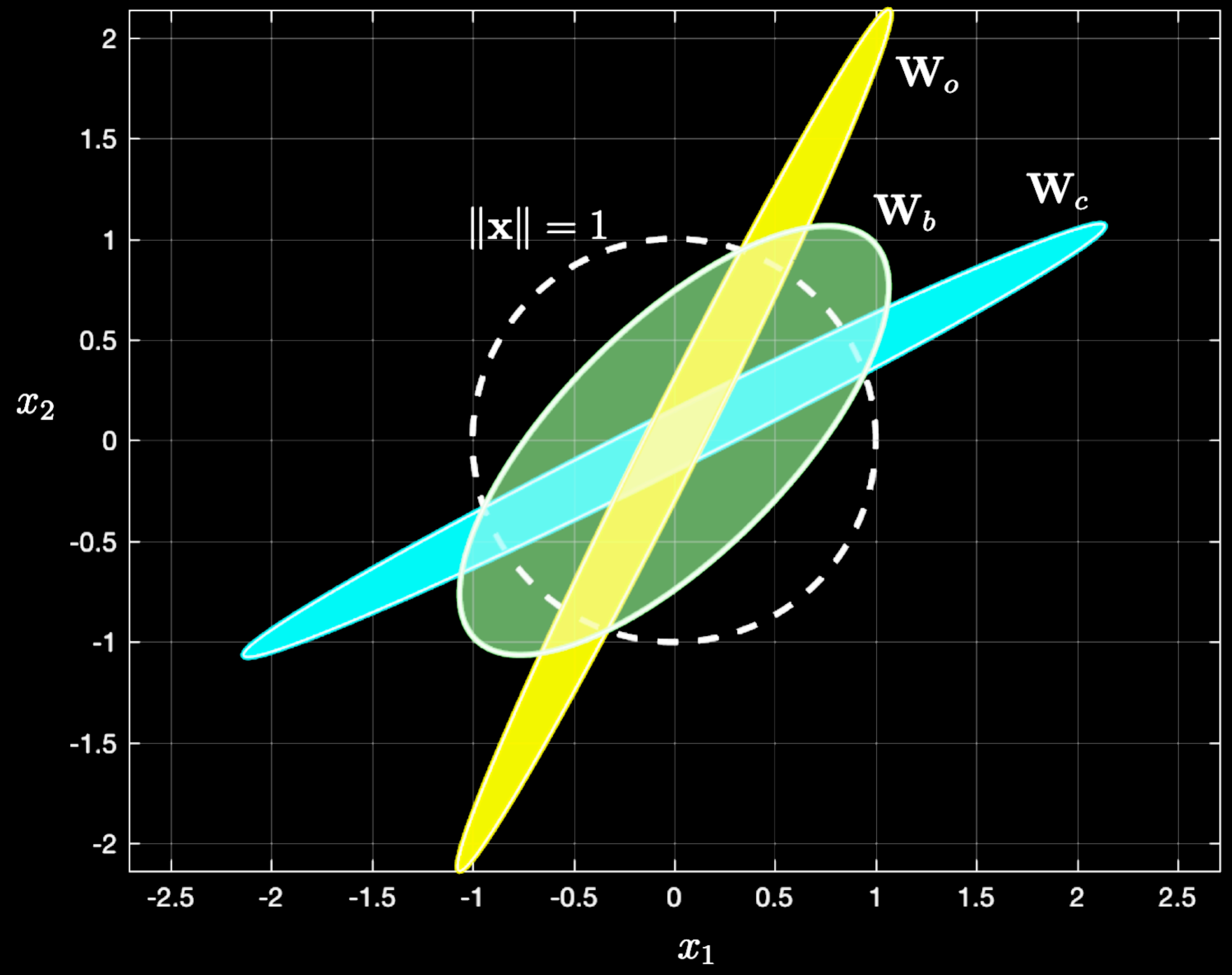
Manohar, et al, *J. Manufacturing Systems*, 2018.



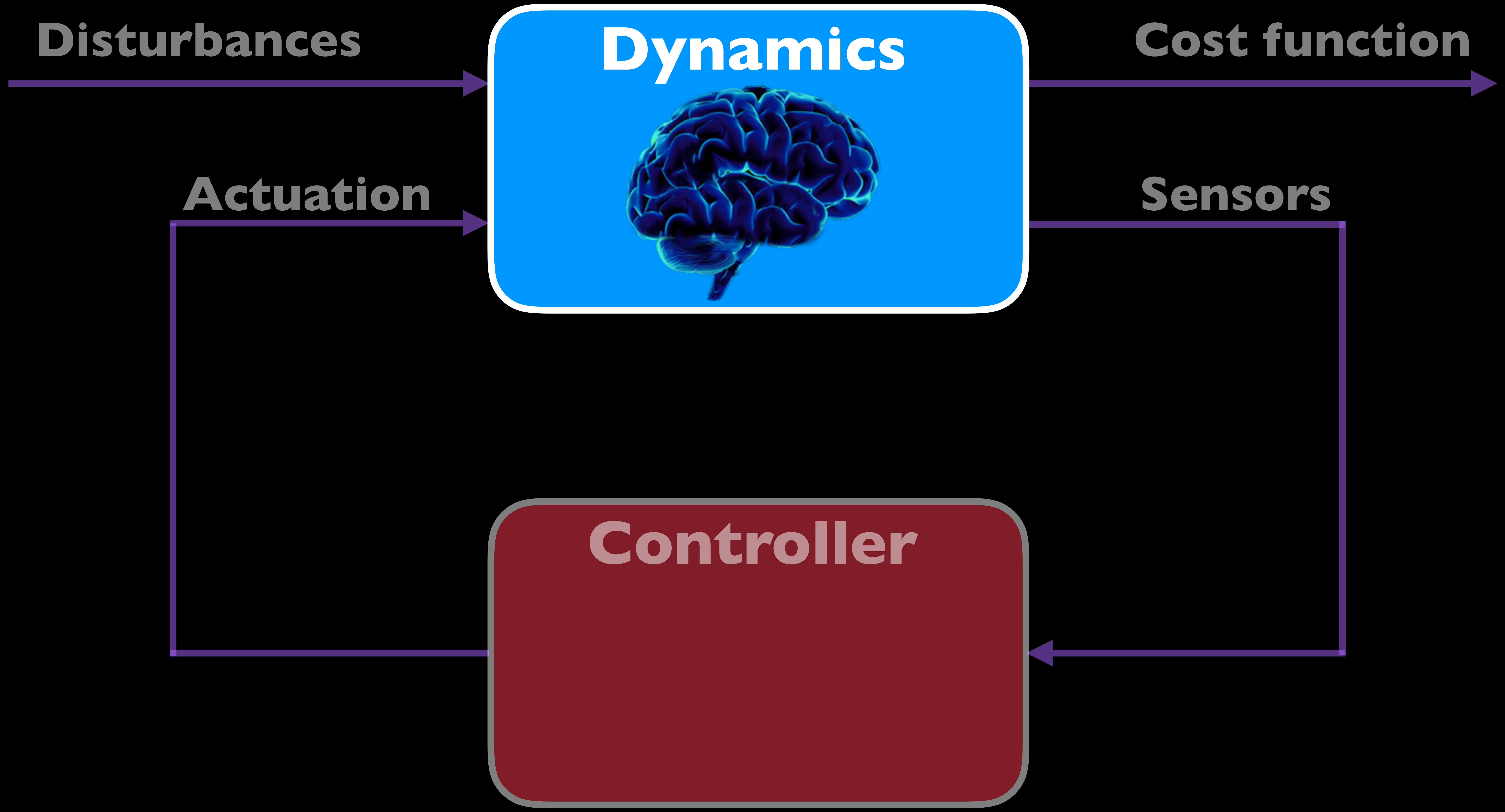
Manohar, et al, *SIAM MMS*, 2019.



Mohren, et al, *PNAS*, 2018.



Manohar, et al, *arXiv*, 2018.


Disturbances
Dynamics
Cost function
Actuation
Sensors
Controller

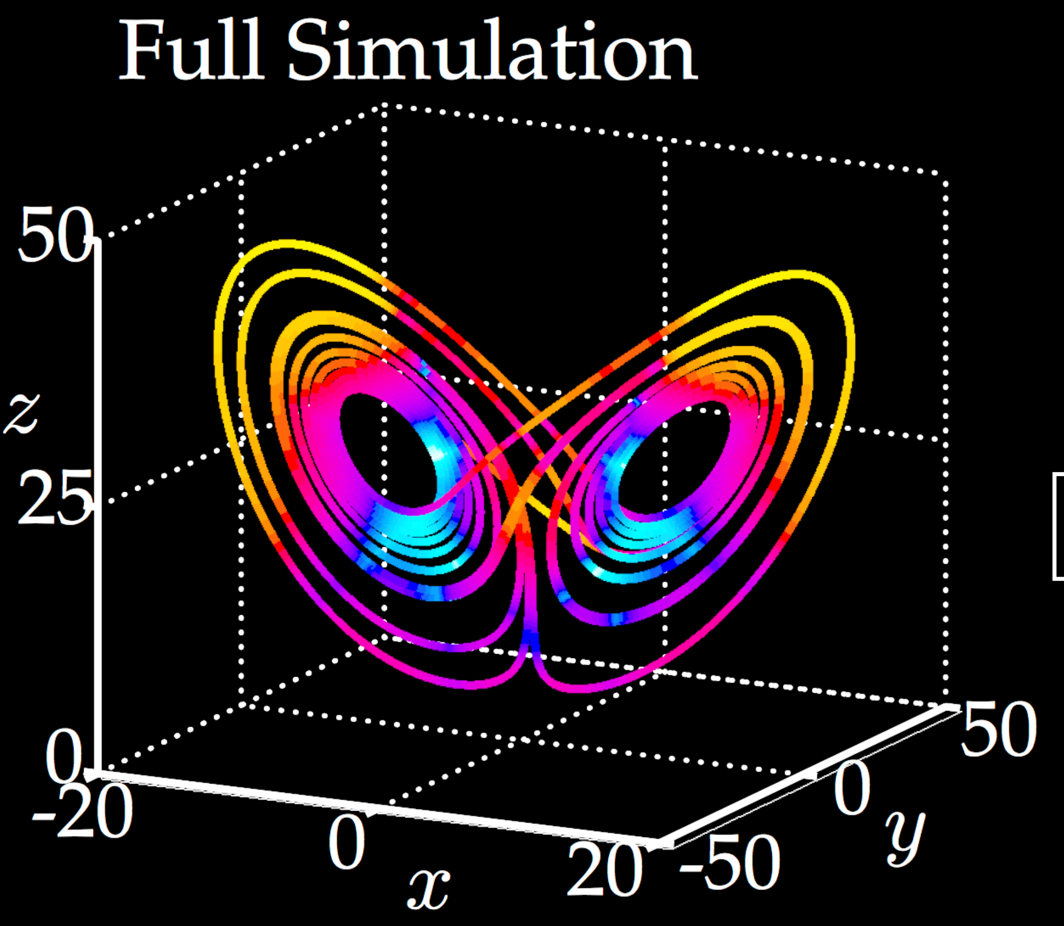
MODEL DISCOVERY

Lots of great work:

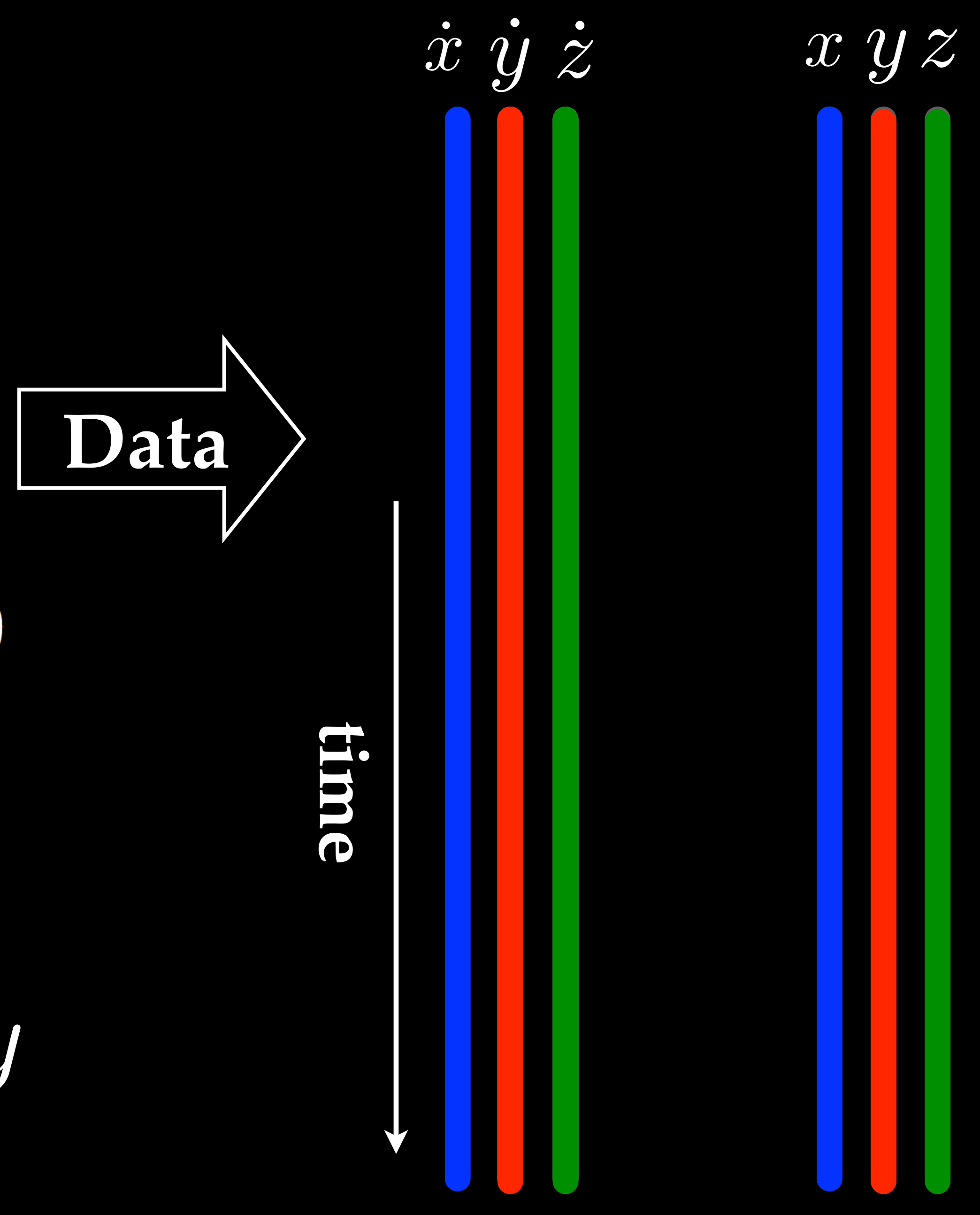
Gonzalez-Garcia, Rico-Martinez, Kevrekidis, *Comp. Chem. Eng.* 1998
Yao and Bollt, *Physica D*, 2007
Bongard and Lipson, *PNAS* 2007
Schmidt and Lipson, *Science* 2009
Wang, Yang, Lai, Kovanis, Grebogi, *PRL* 2011
Bright, Lin, Kutz, *Phys. Fluids*, 2013
Schaeffer, Caflisch, Hauck, Osher, *PNAS* 2013
Noe and Nuske, *MMS* 2013
Nuske, Keller, Perez-Hernandez, Mey, Noe, *JCTC* 2014
Noe, et al., Molecular dynamics, 2013-2016
Schaeffer, *PRSA*, 2017
Schaeffer, Tran, Ward, *SIAP*, 2018
Boninsegna, Nuske, Clementi, *JChemPhys* 2018
Hoffmann, Frohner, Noe, *JChemPhys* 2019
Raissi, Perdikaris, Karniadakis, *JCompPhys* 2019
… and many more!!!

Sparsity/parsimony in dynamics

Sparse Identification of Nonlinear Dynamics (SINDy)

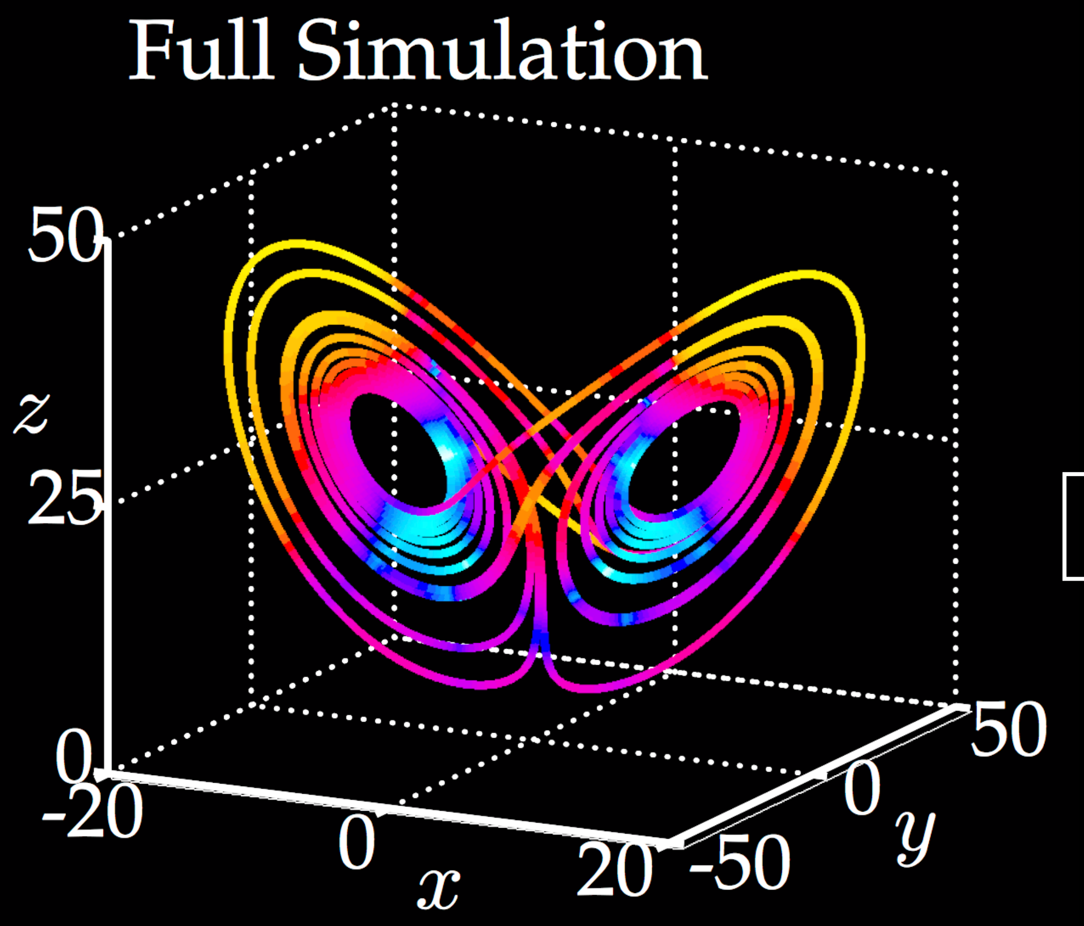


$$\begin{aligned}\dot{x} &= \sigma(y - x)\\ \dot{y} &= x(\rho - z) - y\\ \dot{z} &= xy - \beta z.\end{aligned}$$

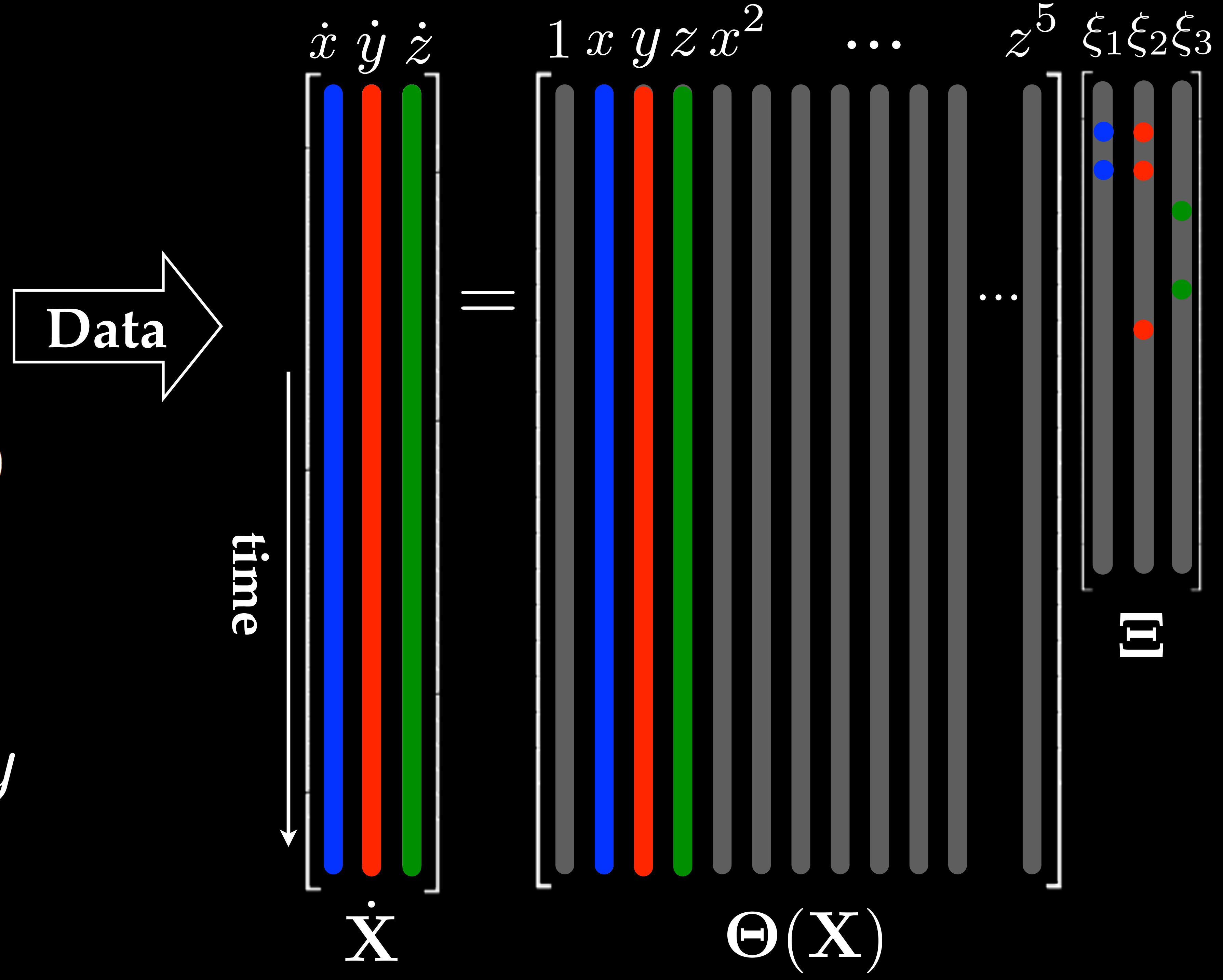


SLB, Proctor, Kutz, *PNAS* 2016.

Sparse Identification of Nonlinear Dynamics (SINDy)



$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = x(\rho - z) - y$$
$$\dot{z} = xy - \beta z.$$



SLB, Proctor, Kutz, *PNAS* 2016.

Sparse Identification of Nonlinear Dynamics (SINDy)

$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = x(\rho - z) - y$$
$$\dot{z} = xy - \beta z.$$

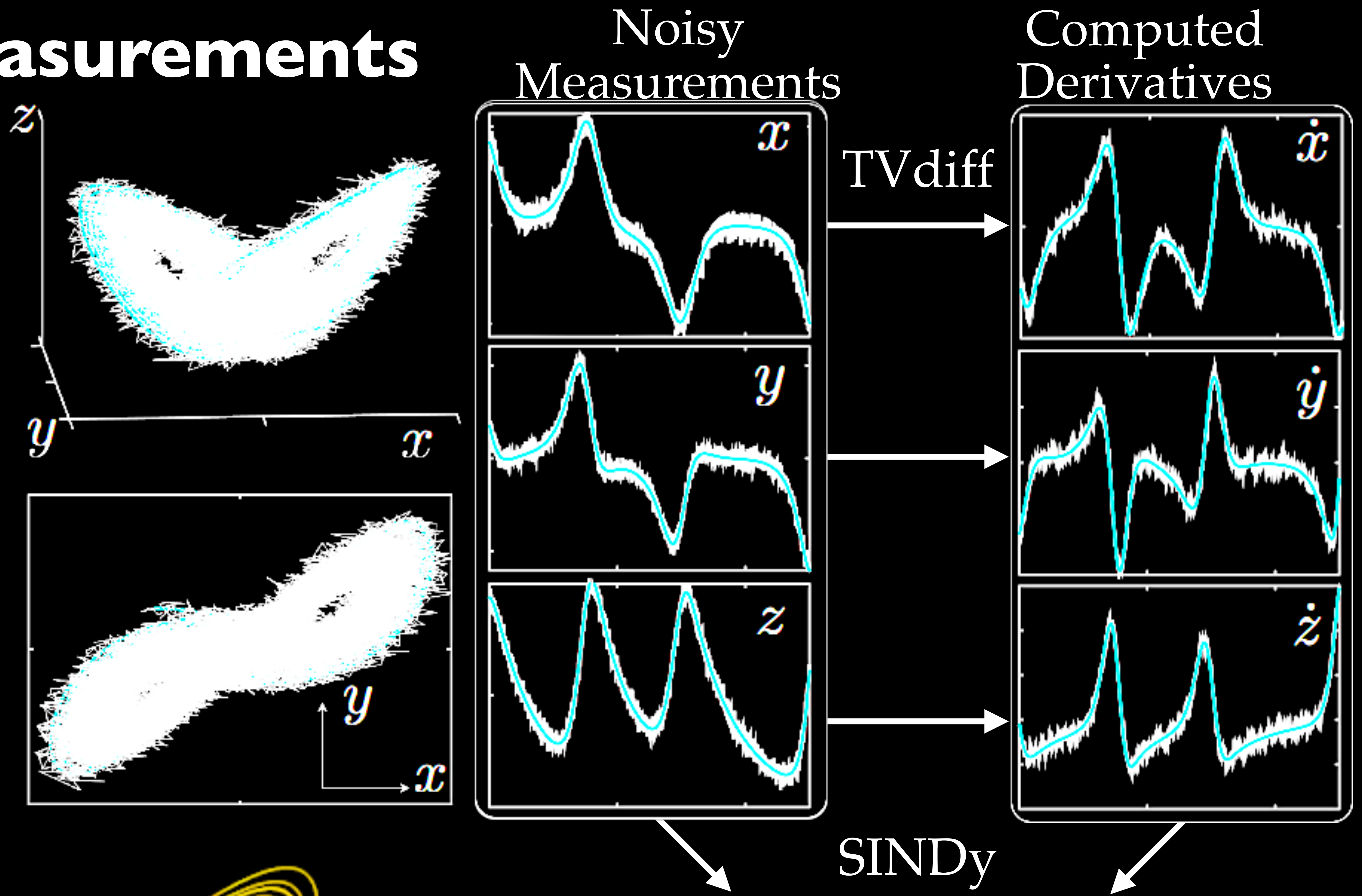
$$\dot{\mathbf{X}} = \Theta(\mathbf{X})\,\Xi$$

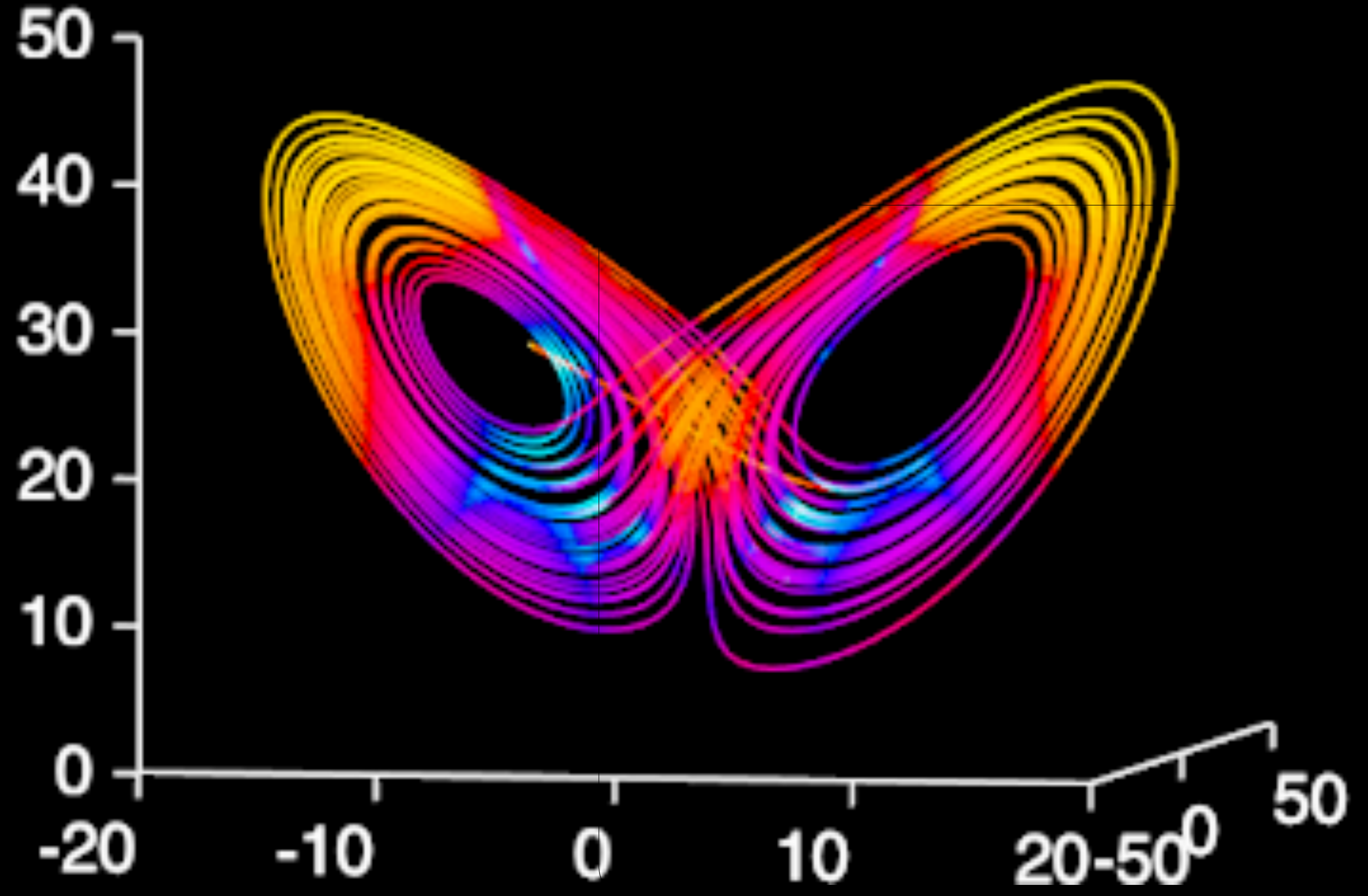
''	'xi_1'	'xi_2'	'xi_3'
'1'	[0]	[0]	[0]
'x'	[-9.9996]	[27.9980]	[0]
'y'	[9.9998]	[-0.9997]	[0]
'z'	[0]	[0]	[-2.6665]
'xx'	[0]	[0]	[0]
'xy'	[0]	[0]	[1.0000]
'xz'	[0]	[-0.9999]	[0]
'yy'	[0]	[0]	[0]
'yz'	[0]	[0]	[0]
'zz'	[0]	[0]	[0]
...	...	...	...
'yzzzz'	[0]	[0]	[0]
'zzzzz'	[0]	[0]	[0]

SLB, Proctor, Kutz, *PNAS* 2016.

SINDy: Noisy State Measurements

Rudin, Osher, Fatemi, *Physica D*, 1992.
SLB, Proctor, Kutz, *PNAS* 2016.



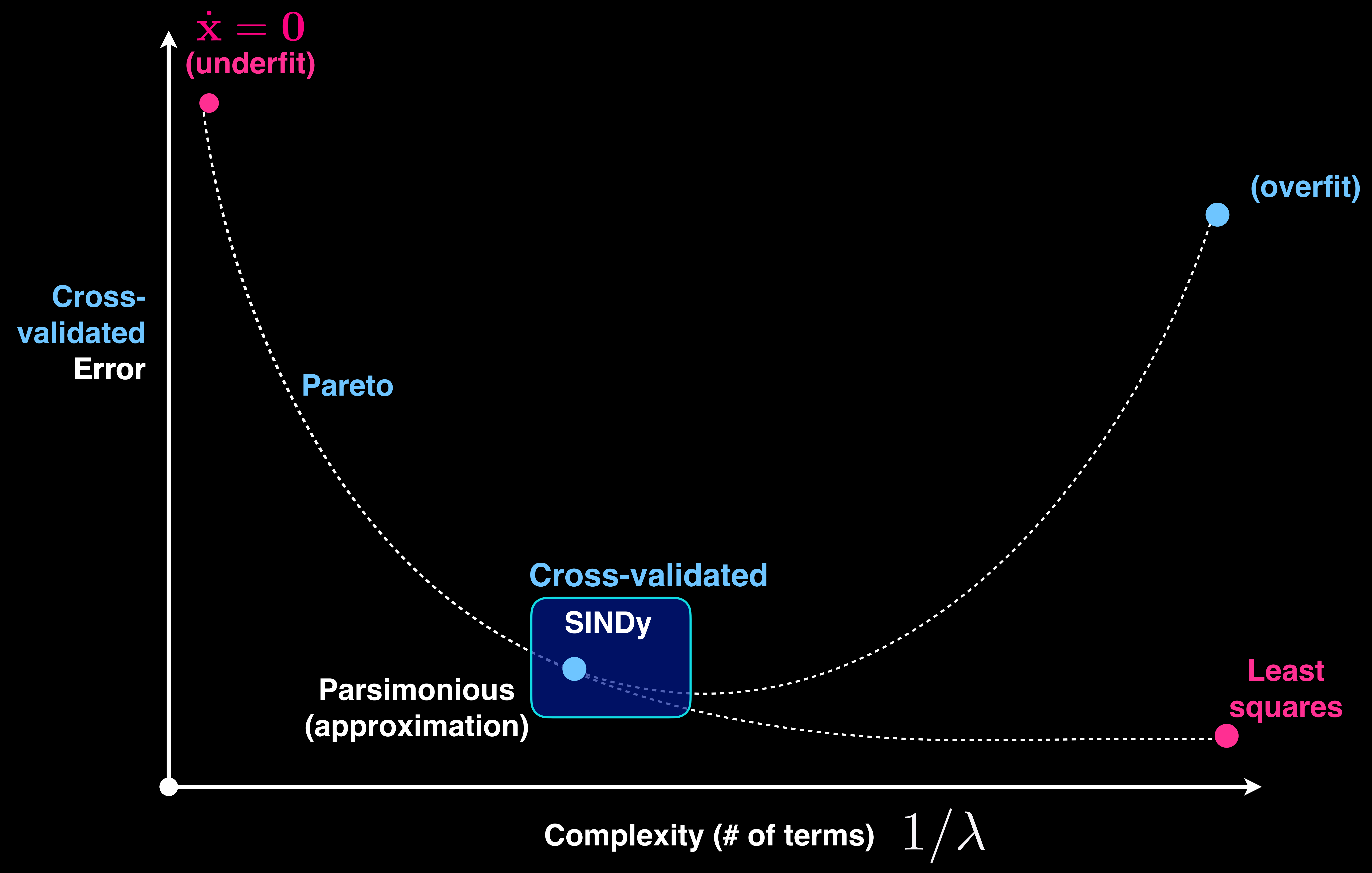


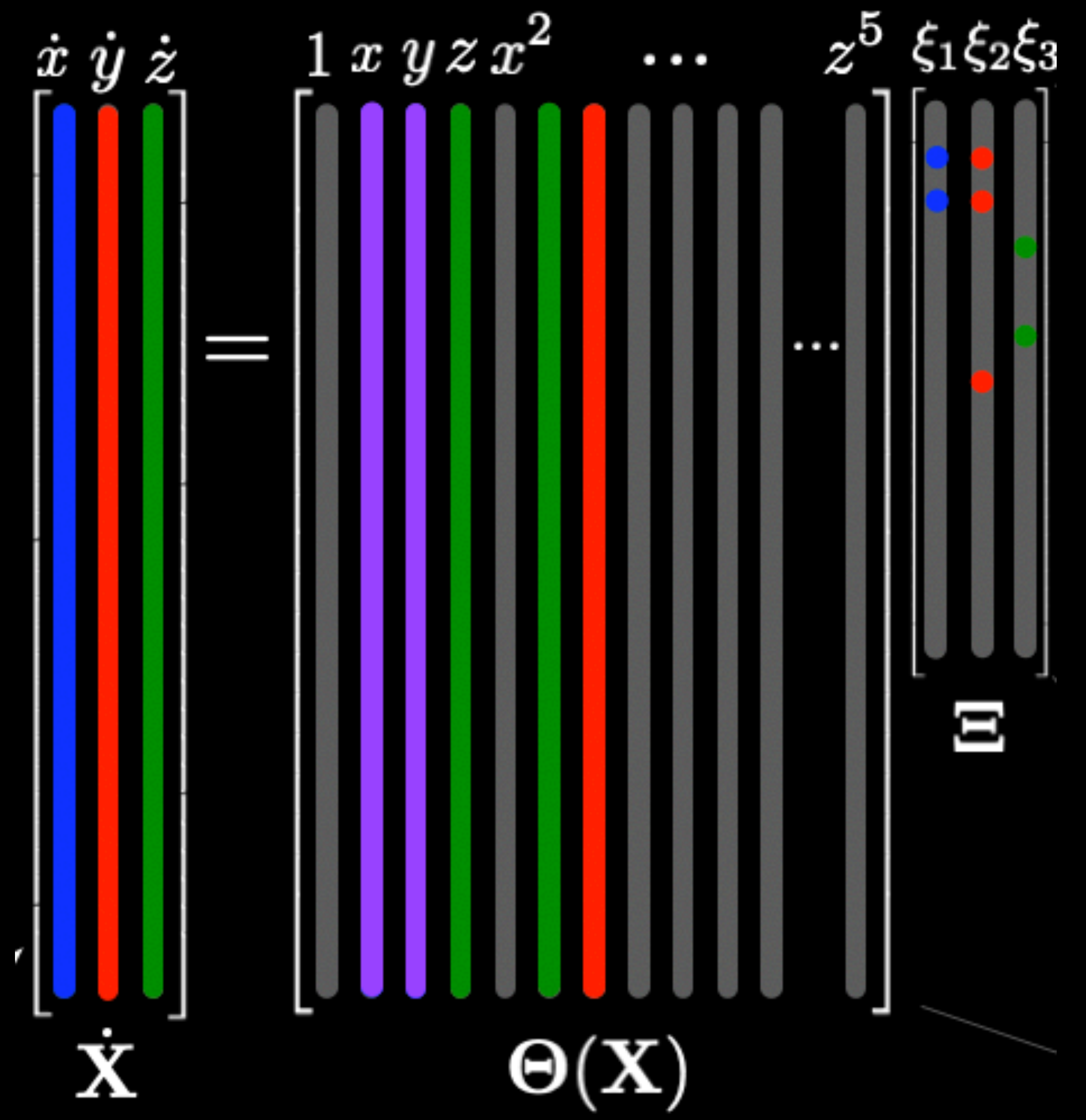
	'xi_1'	'xi_2'	'xi_3'
'x'	[-9.9614]	[27.5343]	[0]
'y'	[9.9796]	[-0.8038]	[0]
'z'	[0]	[0]	[-2.6647]
'xx'	[0]	[0]	[0]
'xy'	[0]	[0]	[1.0003]
'xz'	[0]	[-0.9900]	[0]


PDEs
Rudy, SLB, Proctor, Kutz
Science Advances, 2017
Full Data
1a. Data Collection
t
$(\omega, u, v)_3$
$(\omega, u, v)_2$
$(\omega, u, v)_1$
ω_t
=
1 ω u v ω_x ω_y … $uv\omega_{xy}$ $uv\omega_{yy}$
ξ
1b. Build Nonlinear Library of Data and Derivatives
$\omega_t = \Theta(\omega, u, v)\xi$
1c. Solve Sparse Regression
$\arg\min_{\xi} \|\Theta\xi - \omega_t\|_2^2 + \lambda\|\xi\|_0$
d. Identified Dynamics
$\omega_t + 0.9931u\omega_x + 0.9910v\omega_y = 0.0099\omega_{xx} + 0.0099\omega_{yy}$
Compare to True Navier Stokes ($Re = 100$)
$\omega_t + (\mathbf{u}\cdot\nabla)\omega = \frac{1}{Re}\nabla^2\omega$
Compressed Data
2a. Subsample Data
$\omega_t = \Theta(\omega, u, v)\xi$
=
Sampling
$\mathcal{C}$
2b. Compressed library
$\mathcal{C}\omega_t = \mathcal{C}\Theta(\omega, u, v)\xi$
=
$\mathcal{C}\Theta$
2c. Solve Compressed Sparse Regression
$\arg\min_{\xi} \|\mathcal{C}\Theta\xi - \mathcal{C}\omega_t\|_2^2 + \lambda\|\xi\|_0$

Parsimonious modeling

$$\|\dot{\mathbf{X}} - \Theta(\mathbf{X})\Xi\| + \lambda\|\Xi\|_0$$

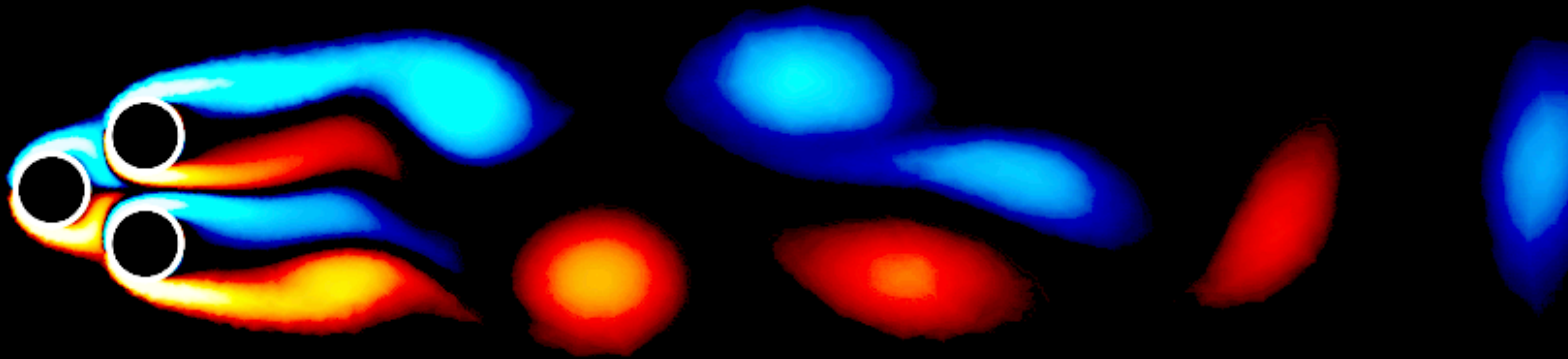




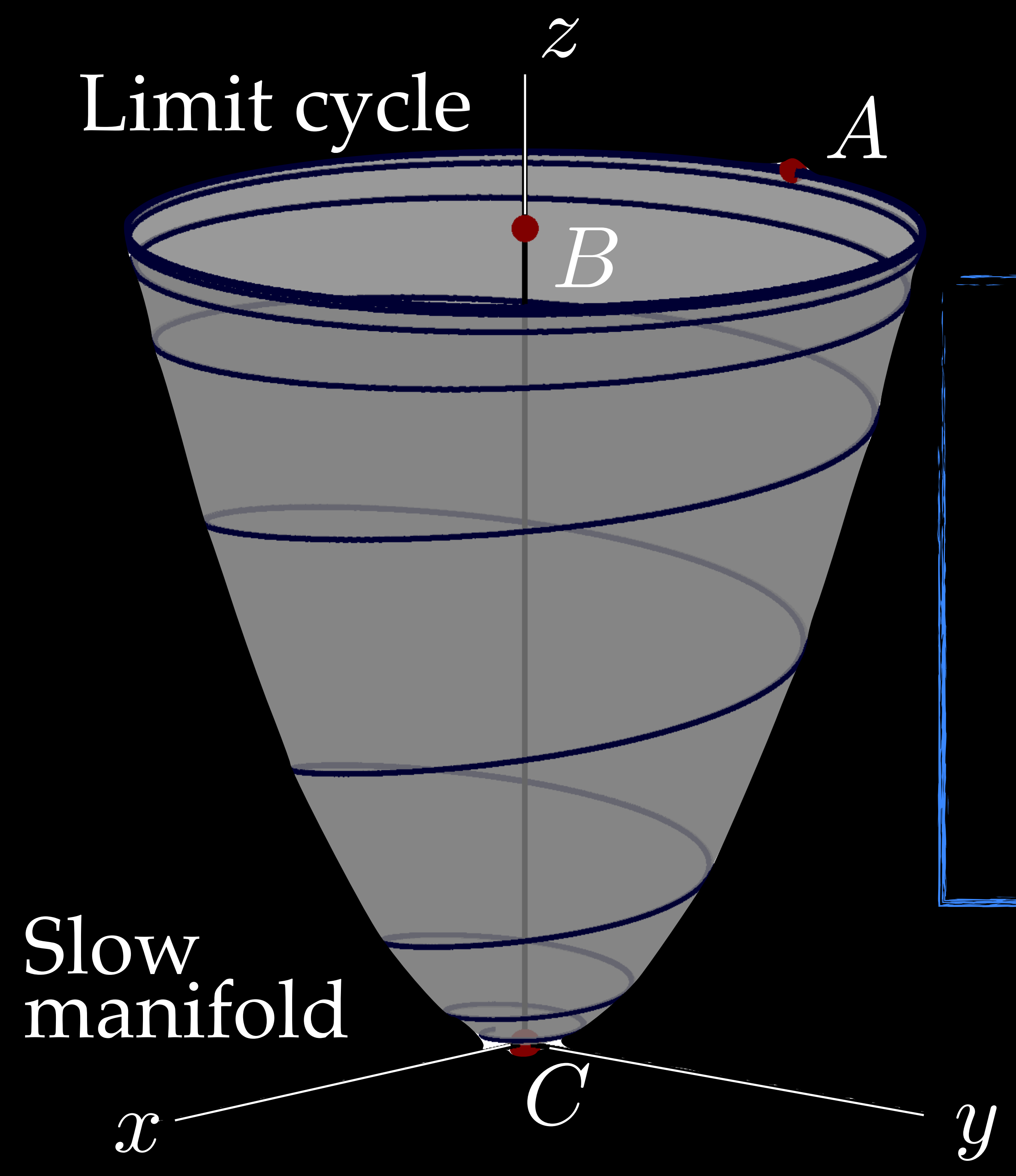
Optimization (SR3): Zheng, Askham, SLB, Kutz, Aravkin IEEE Access, 2019
Model selection: Mangan, SLB, Kutz, Proctor, *PRSA 2017*

Schmidt and Lipson, *Science* 2009
Nuske, Noe, et al., 2013-16

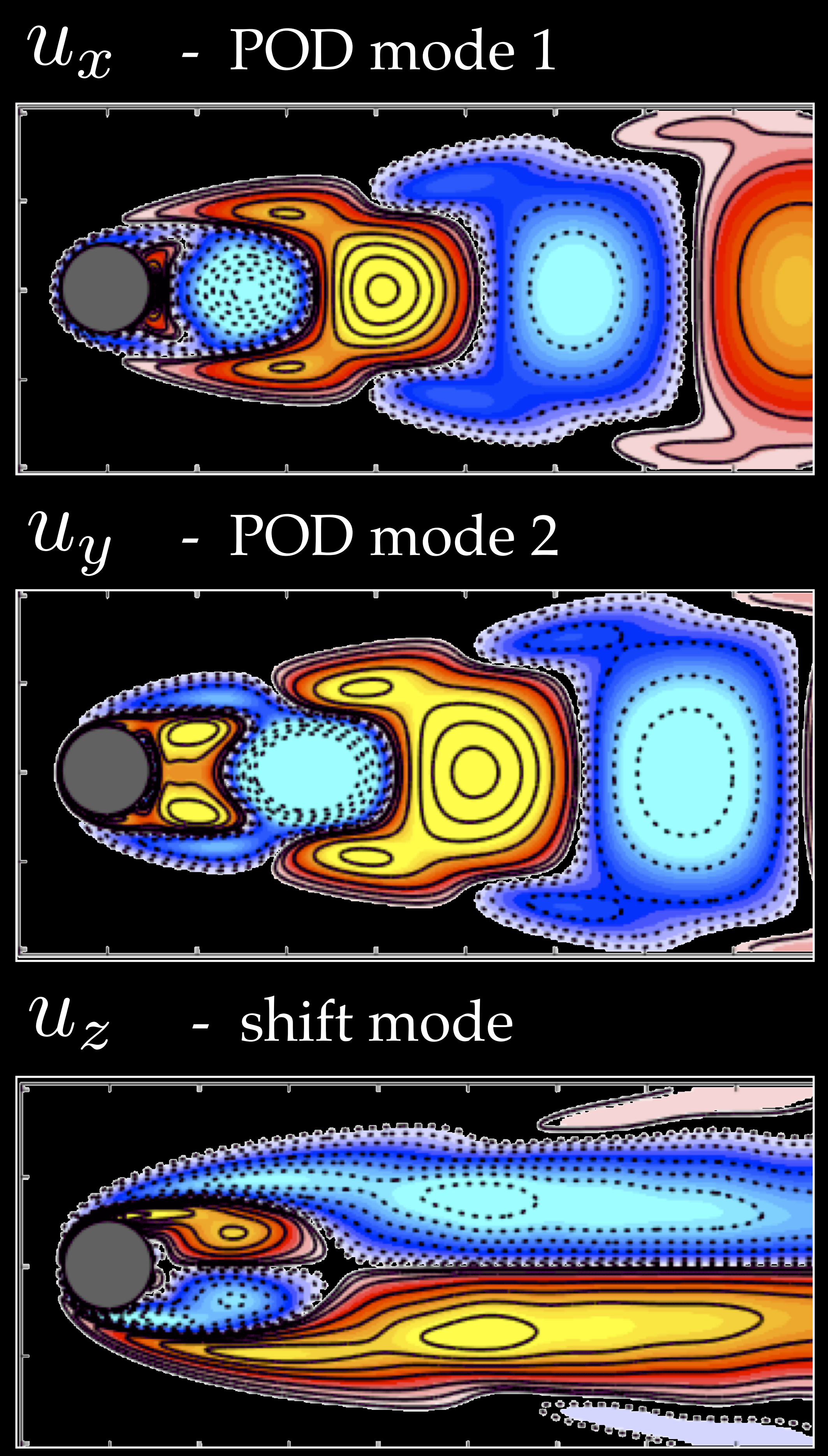
FLUIDS

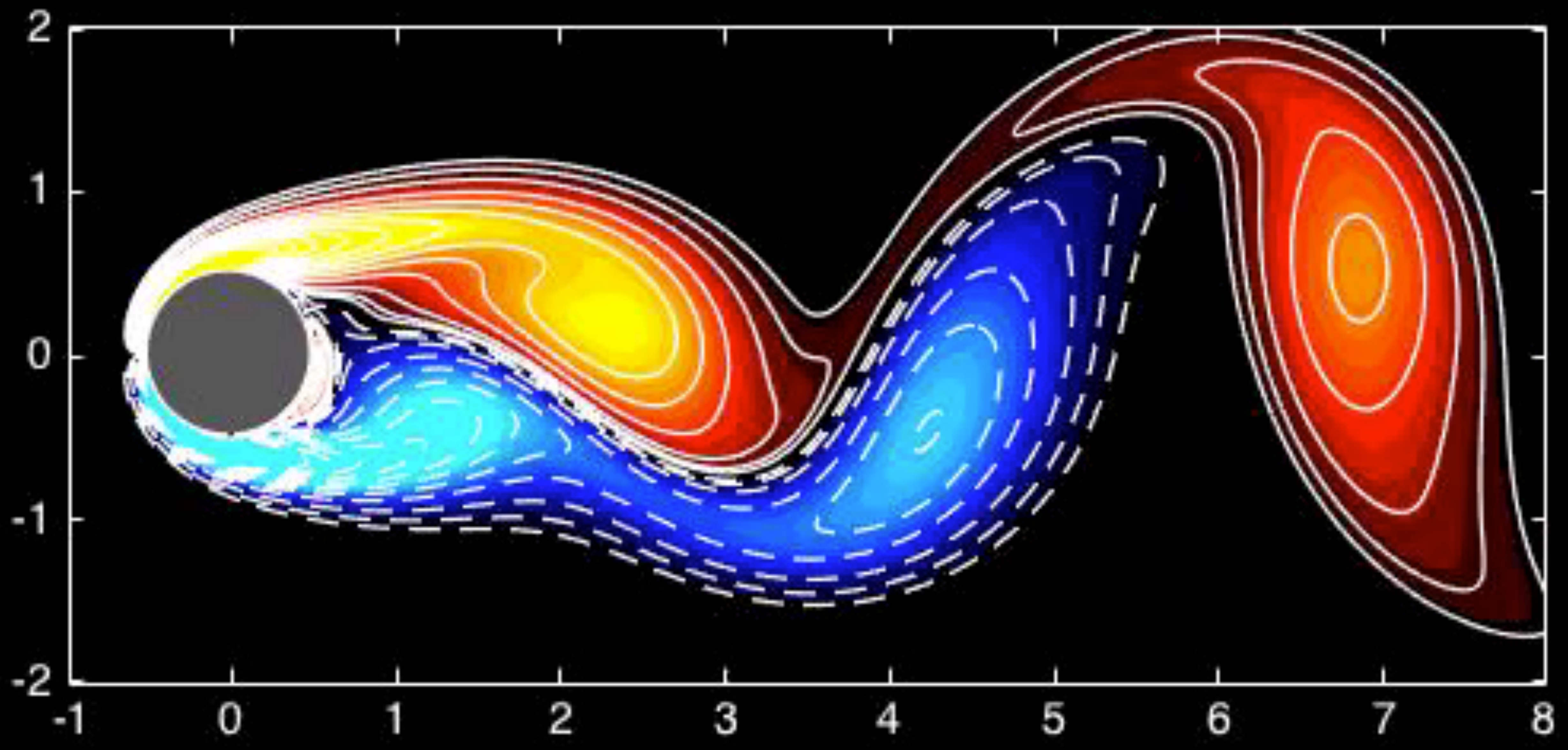


SINDy: Vortex Shedding Past a Cylinder



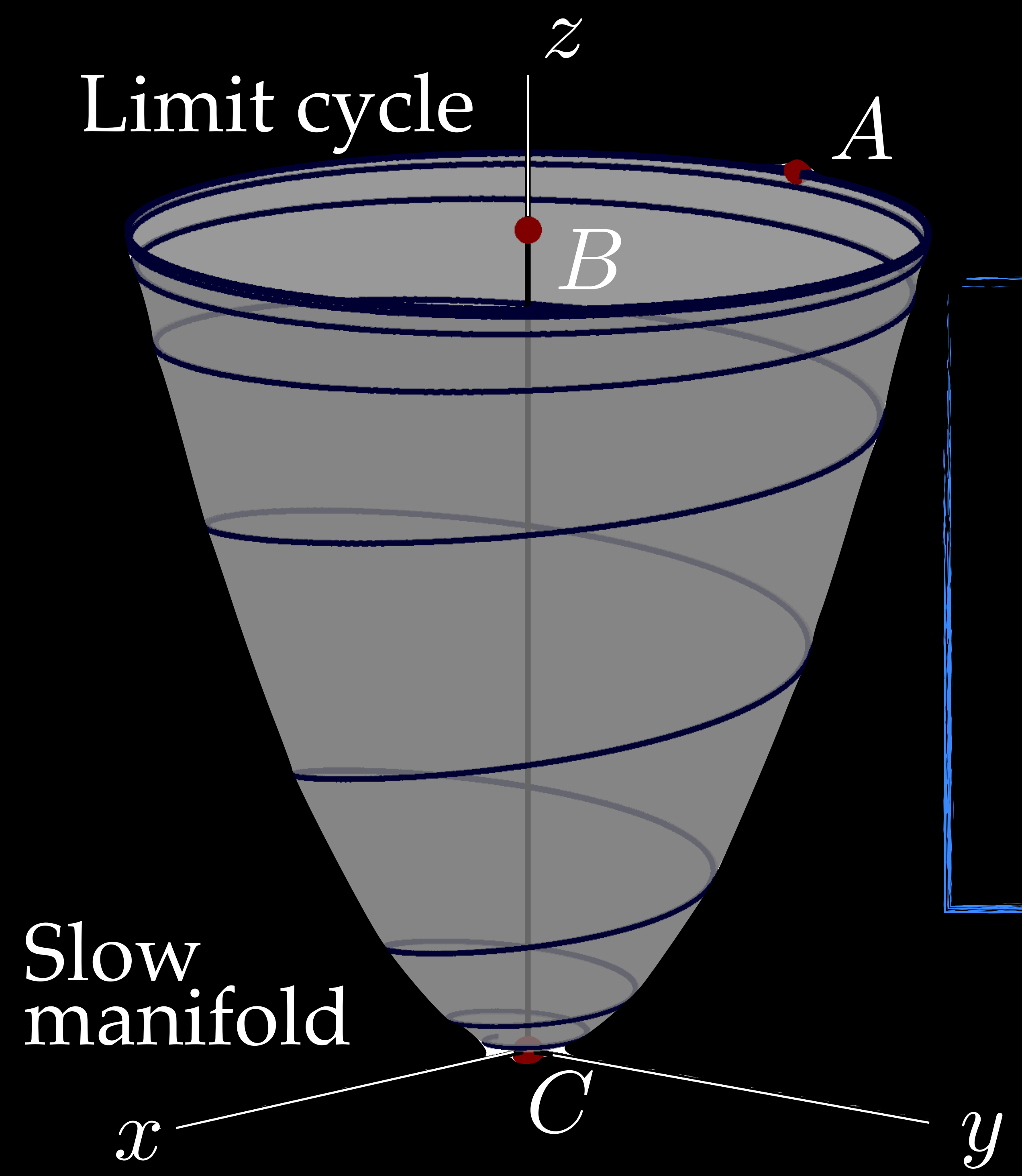
$$
\begin{aligned}
\dot{x} &= \mu x - \omega y + Axz \\
\dot{y} &= \omega x + \mu y + Ayz \\
\dot{z} &= -\lambda(z - x^2 - y^2).
\end{aligned}
$$



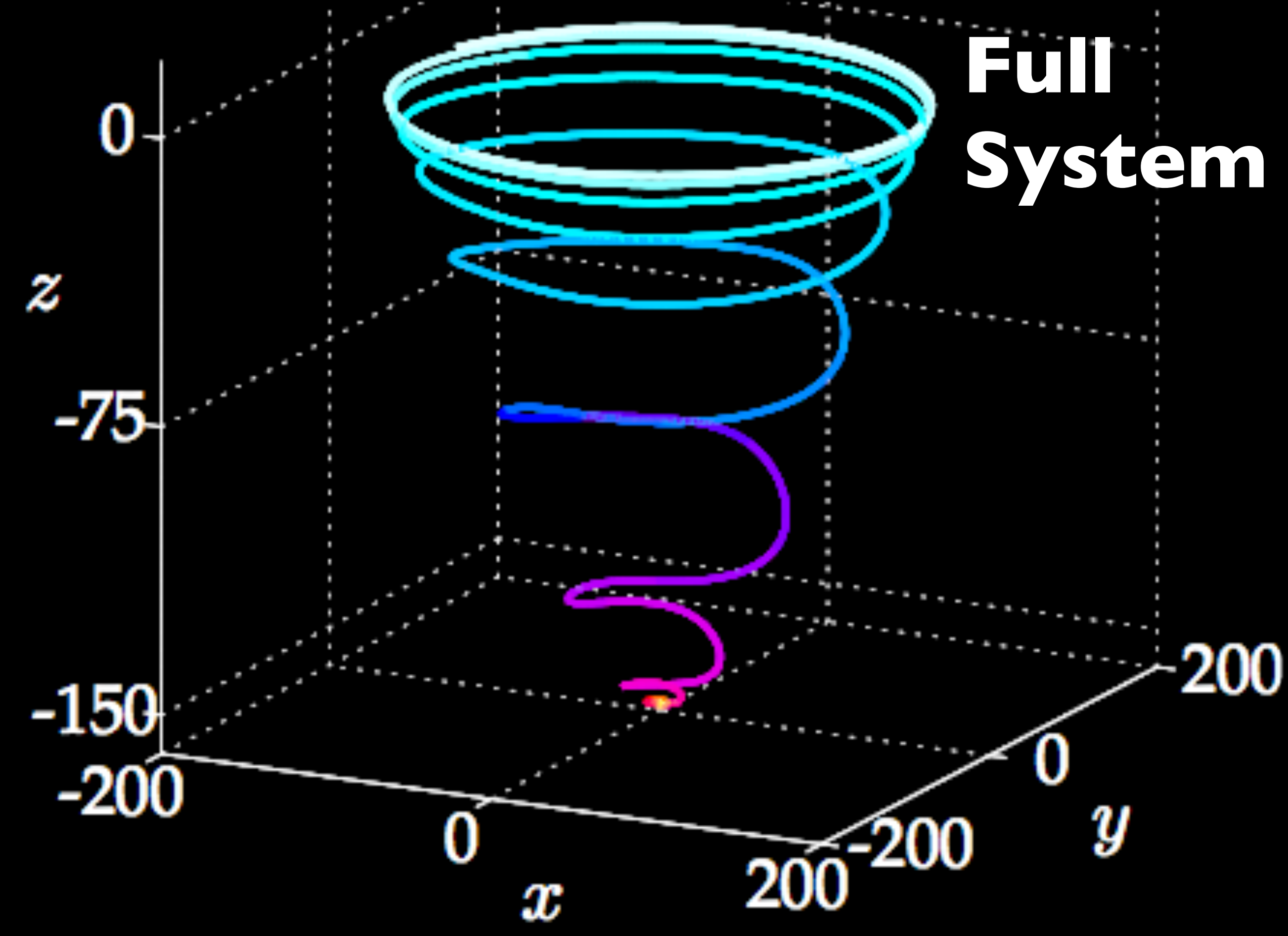


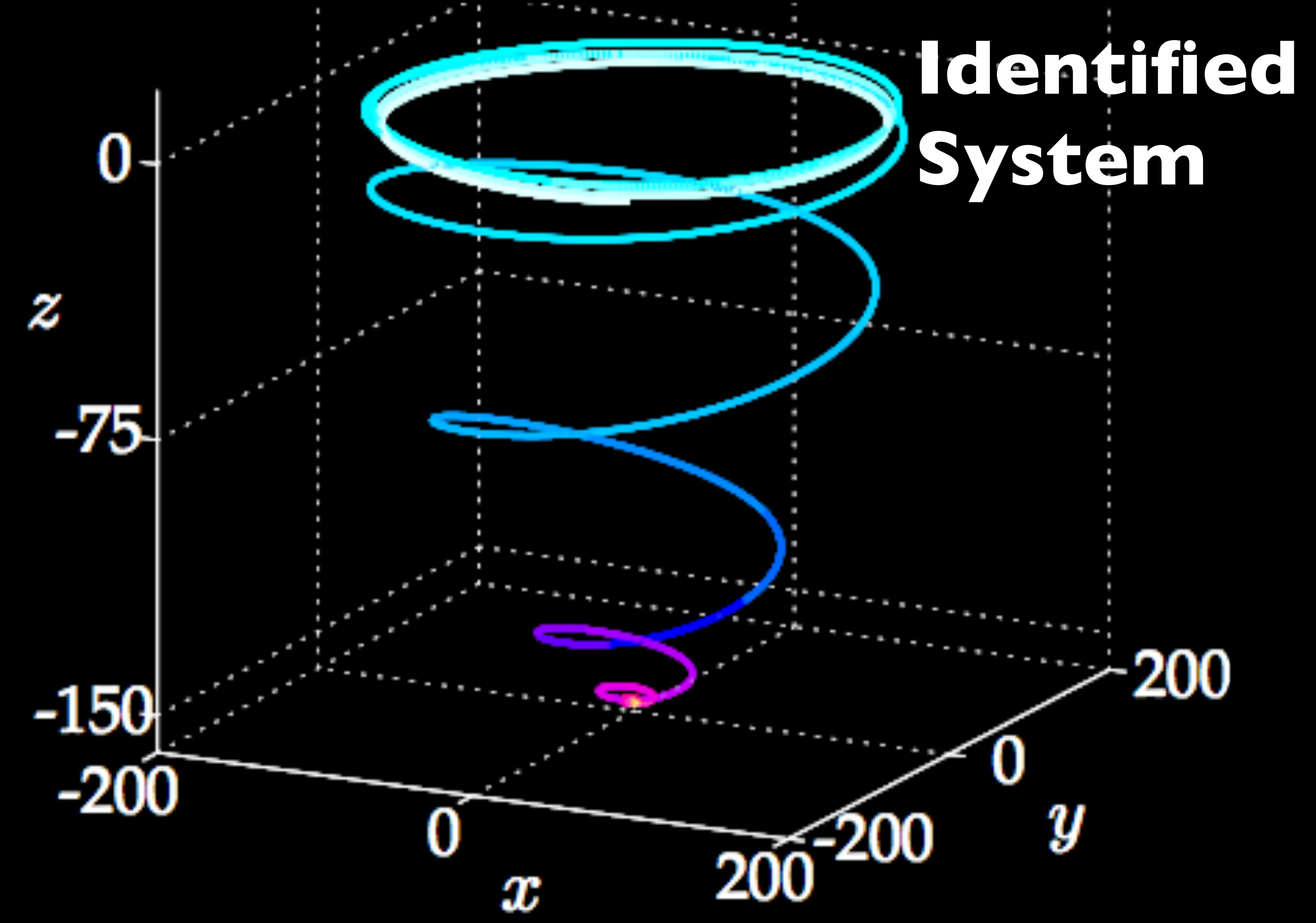
Ruelle and Takens, 1971
Zebib, 1987 and Jackson, 1987
Noack *et al.*, JFM 2003.

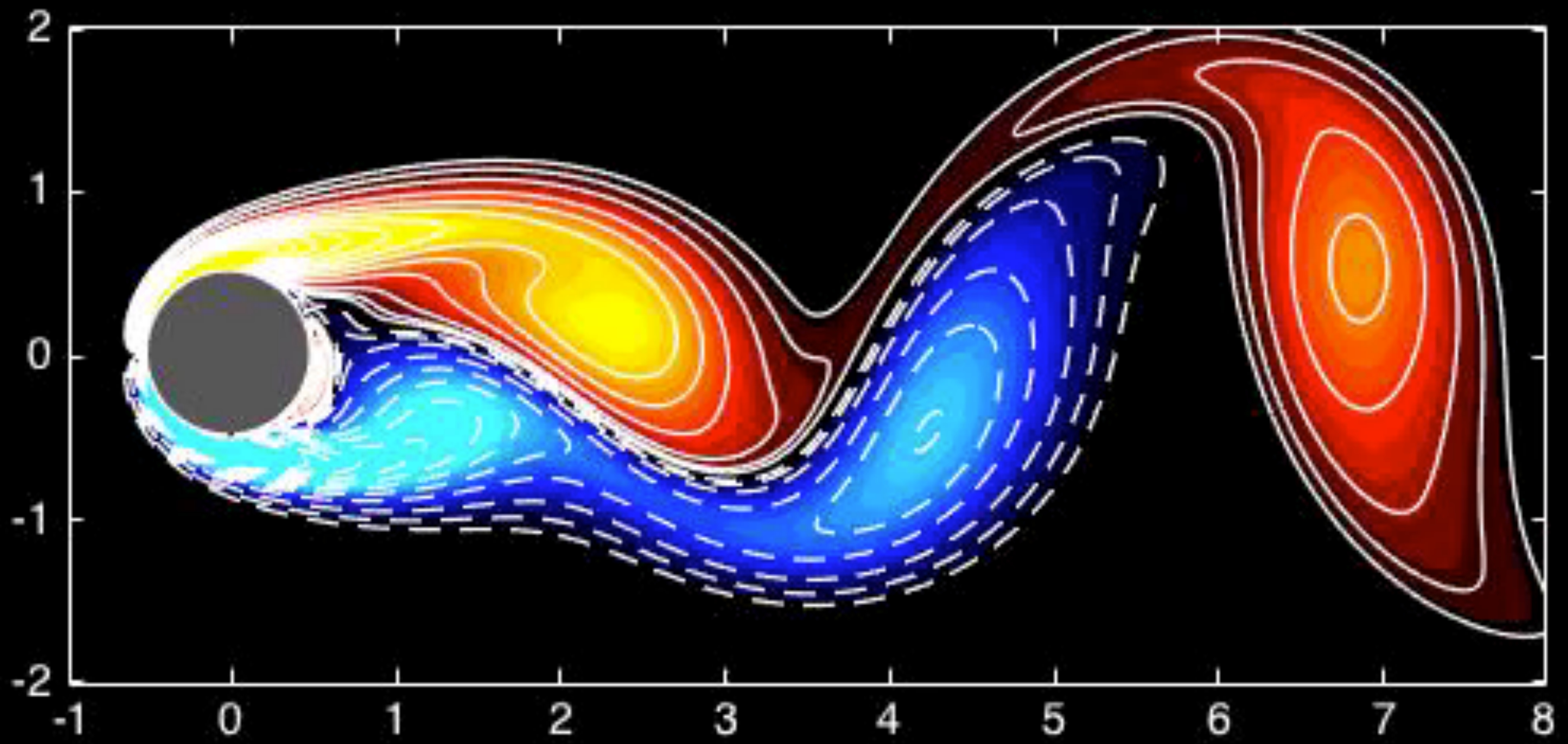
SINDy: Vortex Shedding Past a Cylinder



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

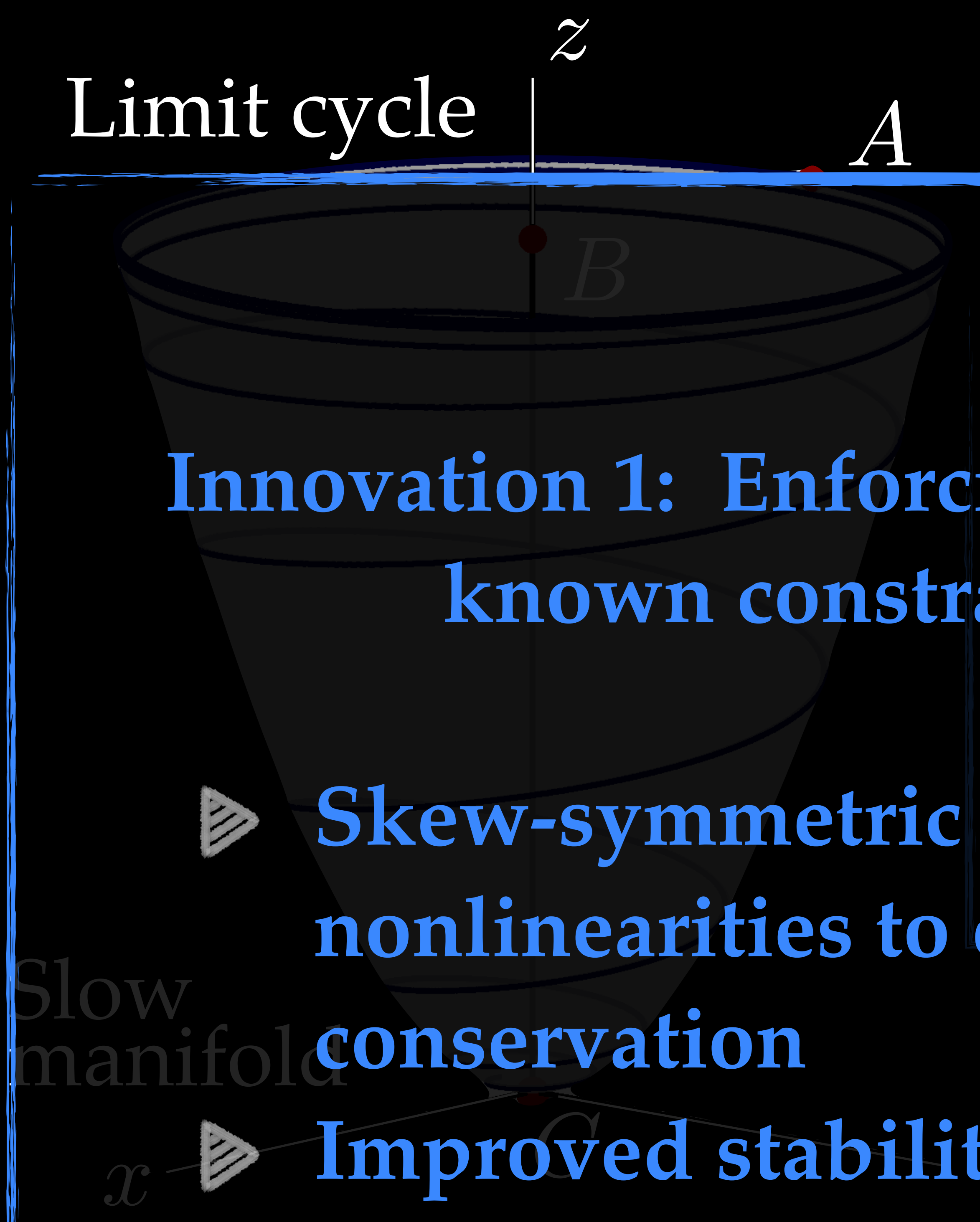






Ruelle and Takens, 1971
Zebib, 1987 and Jackson, 1987
Noack *et al.*, JFM 2003.
SLB, Proctor, Kutz, PNAS 2016.

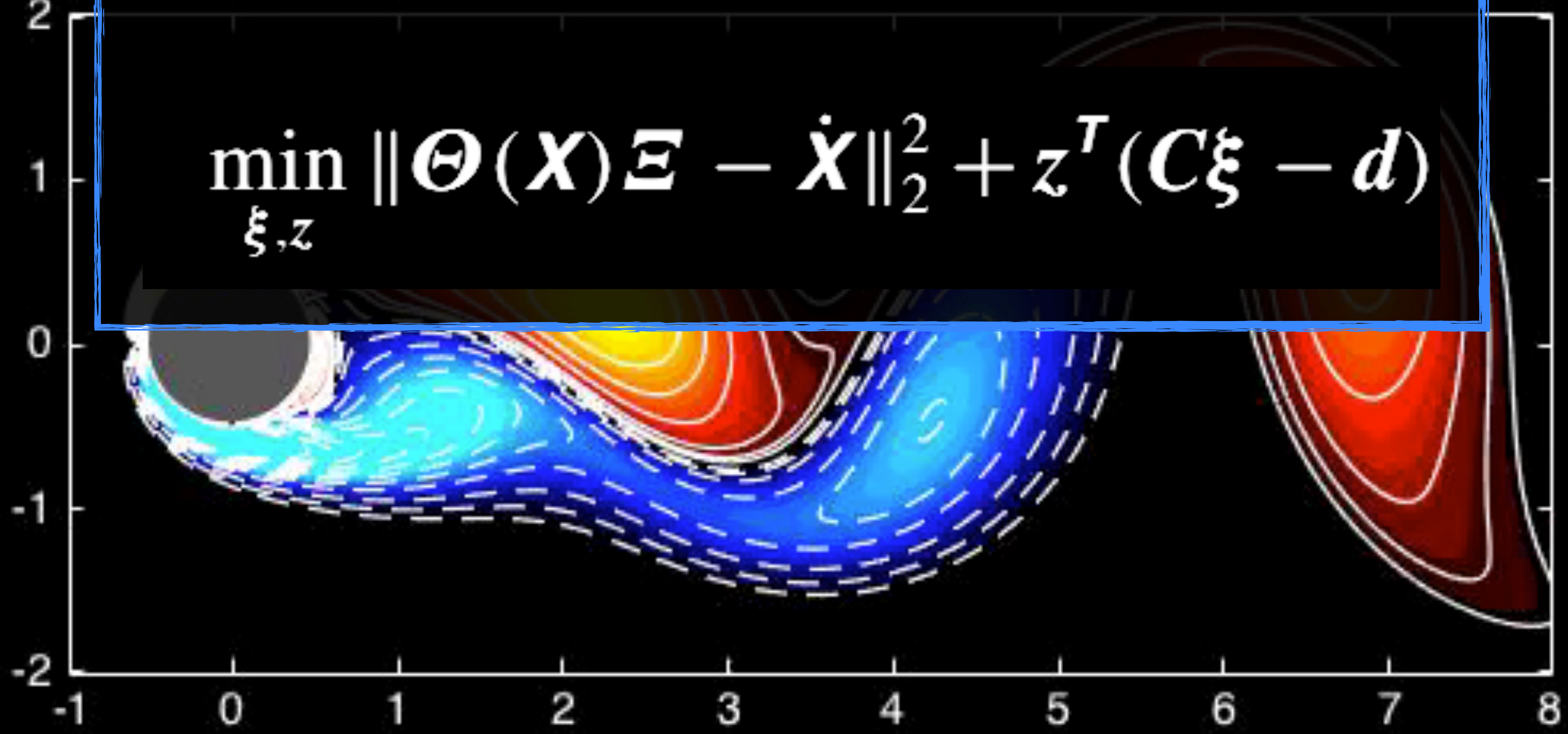
SINDy: Vortex Shedding Past a Cylinder

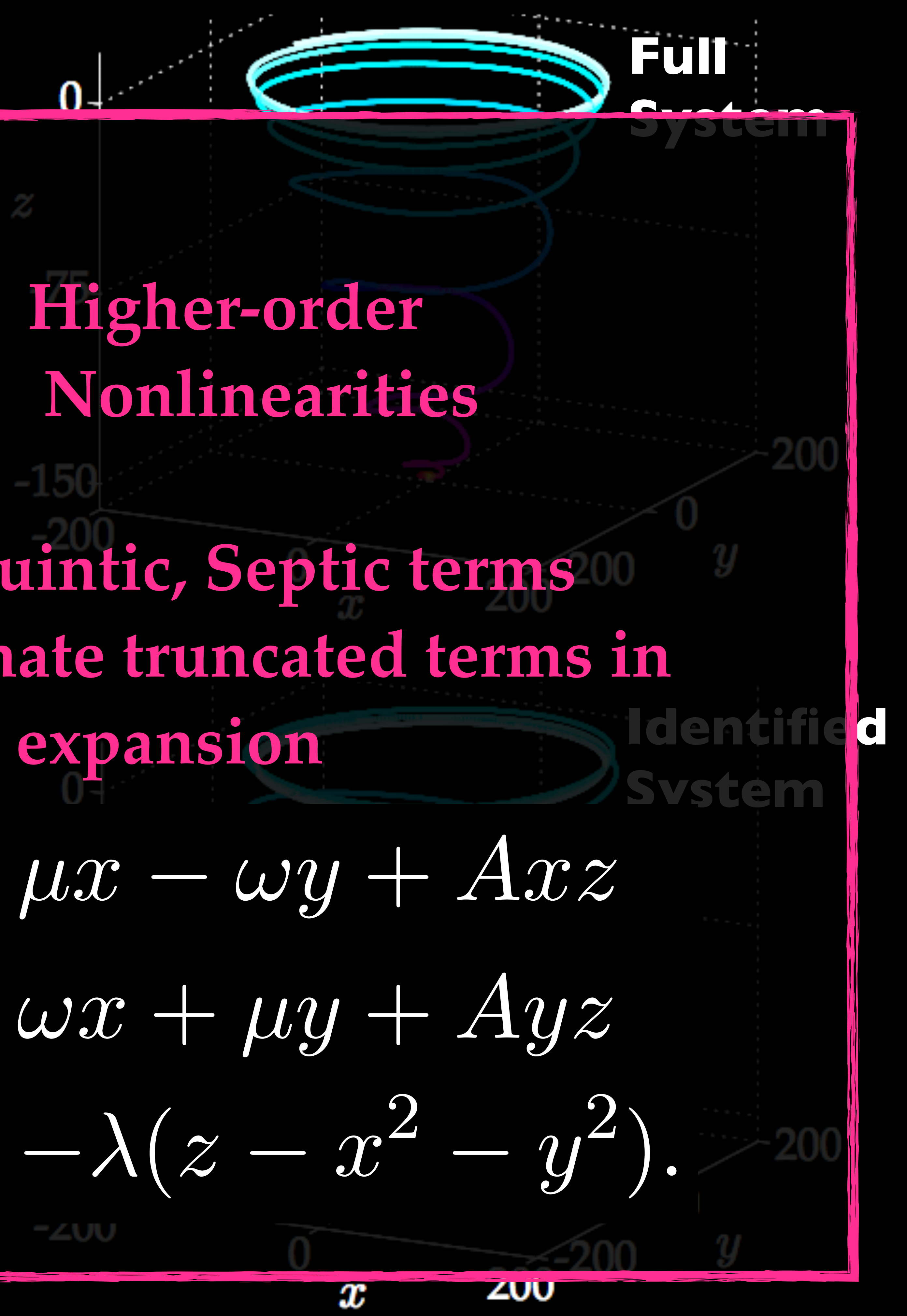


Innovation 1: Enforcing known constraints

- Skew-symmetric quadratic nonlinearities to enforce energy conservation
- Improved stability

$$\min_{\xi, z} \|\Theta(X)\Xi - \dot{X}\|_2^2 + z^T(C\xi - d)$$





Innovation 2: Higher-order Nonlinearities

- Cubic, Quintic, Septic terms approximate truncated terms in Galerkin expansion

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

Ruelle and Takens, 1971
Zebib, 1987 and Jackson, 1987
Noack *et al.*, JFM 2003.
SLB, Proctor, Kutz, PNAS 2016.
Loiseau and SLB, JFM 2017.

Constrained Sparse Galerkin Regression

Innovation 1: Enforcing known constraints

$$\int_\Omega \boldsymbol{u} \cdot (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} \, \mathrm{d}\Omega = 0 \quad \Rightarrow \quad \boldsymbol{a} \cdot \mathcal{N}(\boldsymbol{a}) = 0$$

- Skew-symmetric quadratic nonlinearities to enforce energy conservation
- Improved stability

$$0 = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} \xi_4^{(a_1)} a_1 & \xi_5^{(a_1)} a_1 + \xi_7^{(a_1)} a_2 & \xi_6^{(a_1)} a_1 + \xi_9^{(a_1)} a_3 \\ \xi_4^{(a_2)} a_1 + \xi_5^{(a_2)} a_2 & \xi_7^{(a_2)} a_2 & \xi_8^{(a_2)} a_2 + \xi_9^{(a_2)} a_3 \\ \xi_4^{(a_3)} a_1 + \xi_6^{(a_3)} a_3 & \xi_7^{(a_3)} a_2 + \xi_8^{(a_3)} a_3 & \xi_9^{(a_3)} a_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

$$+ \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} \xi_8^{(a_1)} a_2 a_3 \\ \xi_6^{(a_2)} a_1 a_3 \\ \xi_5^{(a_3)} a_1 a_2 \end{bmatrix} .$$

Loiseau & SLB, *JFM* 838, 2018

Constrained Sparse Galerkin Regression

Innovation 1: Enforcing known constraints

$$\int_{\Omega} \boldsymbol{u} \cdot (\boldsymbol{u} \cdot \nabla) \boldsymbol{u} \, \mathrm{d}\Omega = 0 \quad \Rightarrow \quad \boldsymbol{a} \cdot \mathcal{N}(\boldsymbol{a}) = 0$$

- Skew-symmetric quadratic nonlinearities to enforce energy conservation
- Improved stability

$$\left.\begin{aligned} \xi_8^{(a_1)} + \xi_6^{(a_2)} + \xi_5^{(a_3)} &= 0, \\ \xi_4^{(a_1)} = \xi_7^{(a_2)} = \xi_9^{(a_3)} &= 0, \\ \xi_5^{(a_1)} &= -\xi_4^{(a_2)}, \\ \xi_7^{(a_1)} &= -\xi_5^{(a_2)}, \\ \xi_6^{(a_1)} &= -\xi_4^{(a_3)}, \\ \xi_9^{(a_1)} &= -\xi_6^{(a_3)}, \\ \xi_8^{(a_2)} &= -\xi_7^{(a_3)}, \\ \xi_9^{(a_2)} &= -\xi_8^{(a_3)}, \end{aligned}\right\}$$

Loiseau & SLB, *JFM* 838, 2018

Constrained Sparse Galerkin Regression

Innovation 1: Enforcing known constraints

$$\int_\Omega \boldsymbol{u}\cdot(\boldsymbol{u}\cdot\nabla)\boldsymbol{u}\,\mathrm{d}\Omega = 0 \;\Rightarrow\; \boldsymbol{a}\cdot\mathcal{N}(\boldsymbol{a}) = 0$$

- Skew-symmetric quadratic nonlinearities to enforce energy conservation
- Improved stability

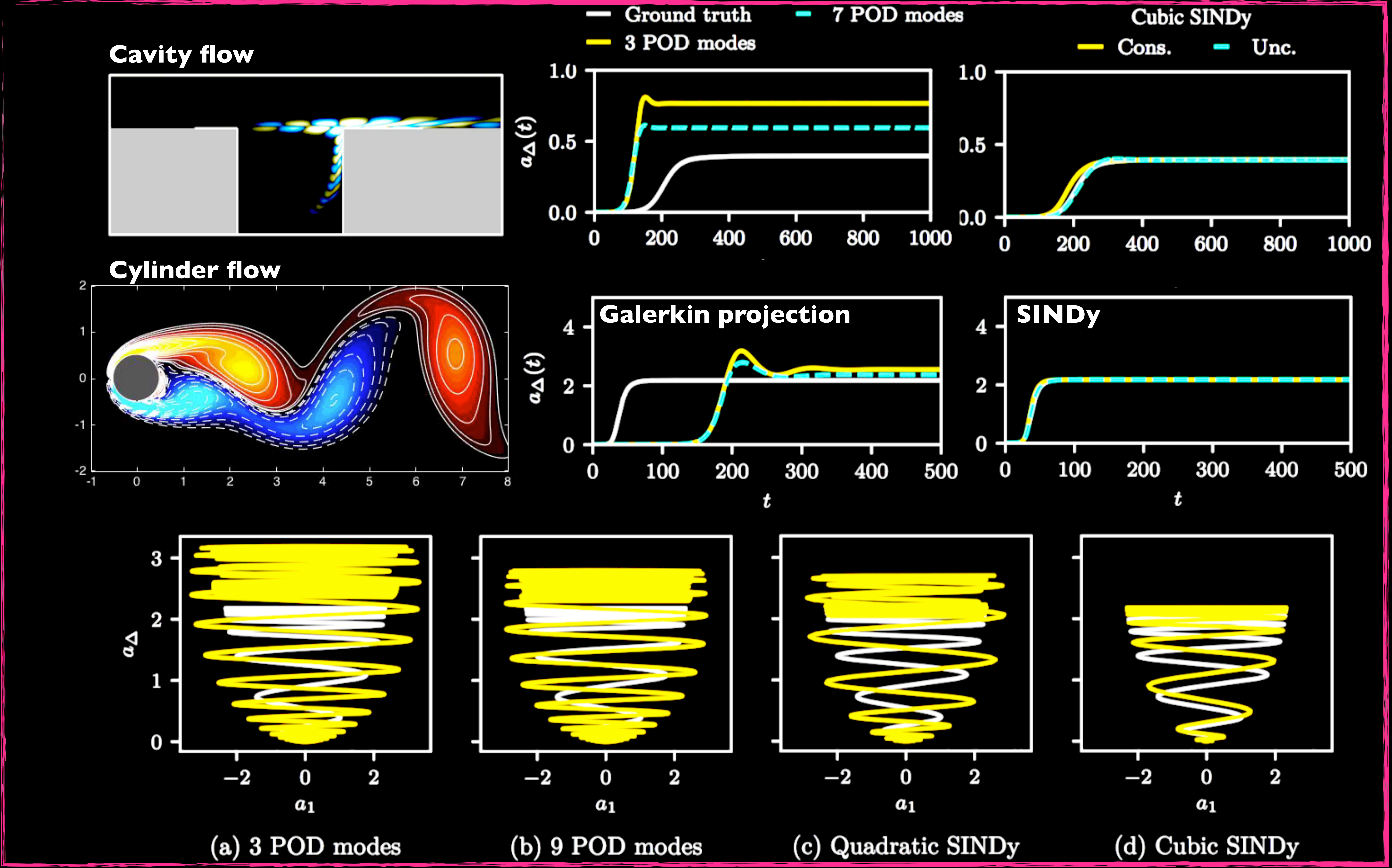
$$\min_{\xi,z} \|\boldsymbol{\Theta}(\boldsymbol{X})\boldsymbol{\Xi} - \dot{\boldsymbol{X}}\|_2^2 + z^T(C\xi - \boldsymbol{d})$$

$$\begin{bmatrix} 2\hat{\boldsymbol{\Theta}}(\boldsymbol{X})^T\hat{\boldsymbol{\Theta}}(\boldsymbol{X}) & \boldsymbol{C}^T \\ \boldsymbol{C} & 0 \end{bmatrix} \begin{bmatrix} \xi \\ z \end{bmatrix} = \begin{bmatrix} 2\hat{\boldsymbol{\Theta}}(\boldsymbol{X})^T\dot{\boldsymbol{X}}(:) \\ \boldsymbol{d} \end{bmatrix}$$

$$\left.\begin{aligned}
\xi_8^{(a_1)} + \xi_6^{(a_2)} + \xi_5^{(a_3)} &= 0,\\
\xi_4^{(a_1)} = \xi_7^{(a_2)} = \xi_9^{(a_3)} &= 0,\\
\xi_5^{(a_1)} &= -\xi_4^{(a_2)},\\
\xi_7^{(a_1)} &= -\xi_5^{(a_2)},\\
\xi_6^{(a_1)} &= -\xi_4^{(a_3)},\\
\xi_9^{(a_1)} &= -\xi_6^{(a_3)},\\
\xi_8^{(a_2)} &= -\xi_7^{(a_3)},\\
\xi_9^{(a_2)} &= -\xi_8^{(a_3)},
\end{aligned}\right\}$$

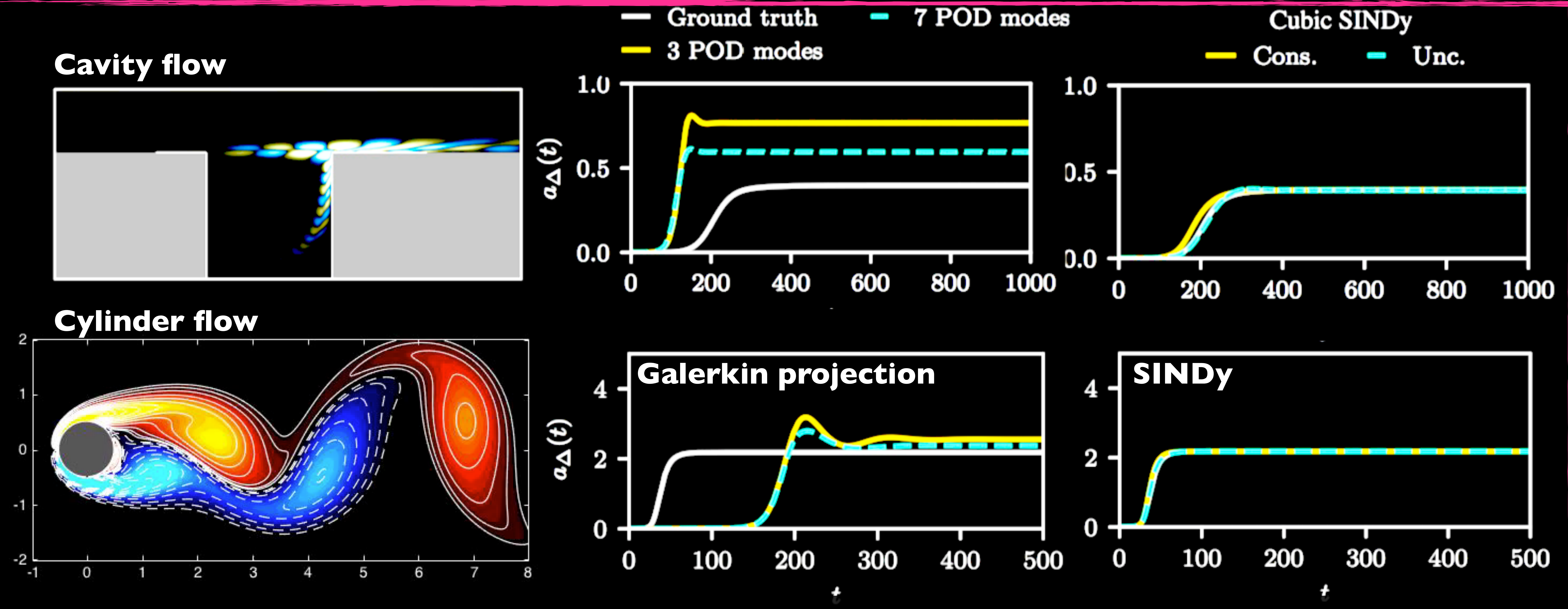
Loiseau & SLB, *JFM* 838, 2018

Constrained Sparse Galerkin Regression



Loiseau & SLB,
JFM 838, 2018

Constrained Sparse Galerkin Regression

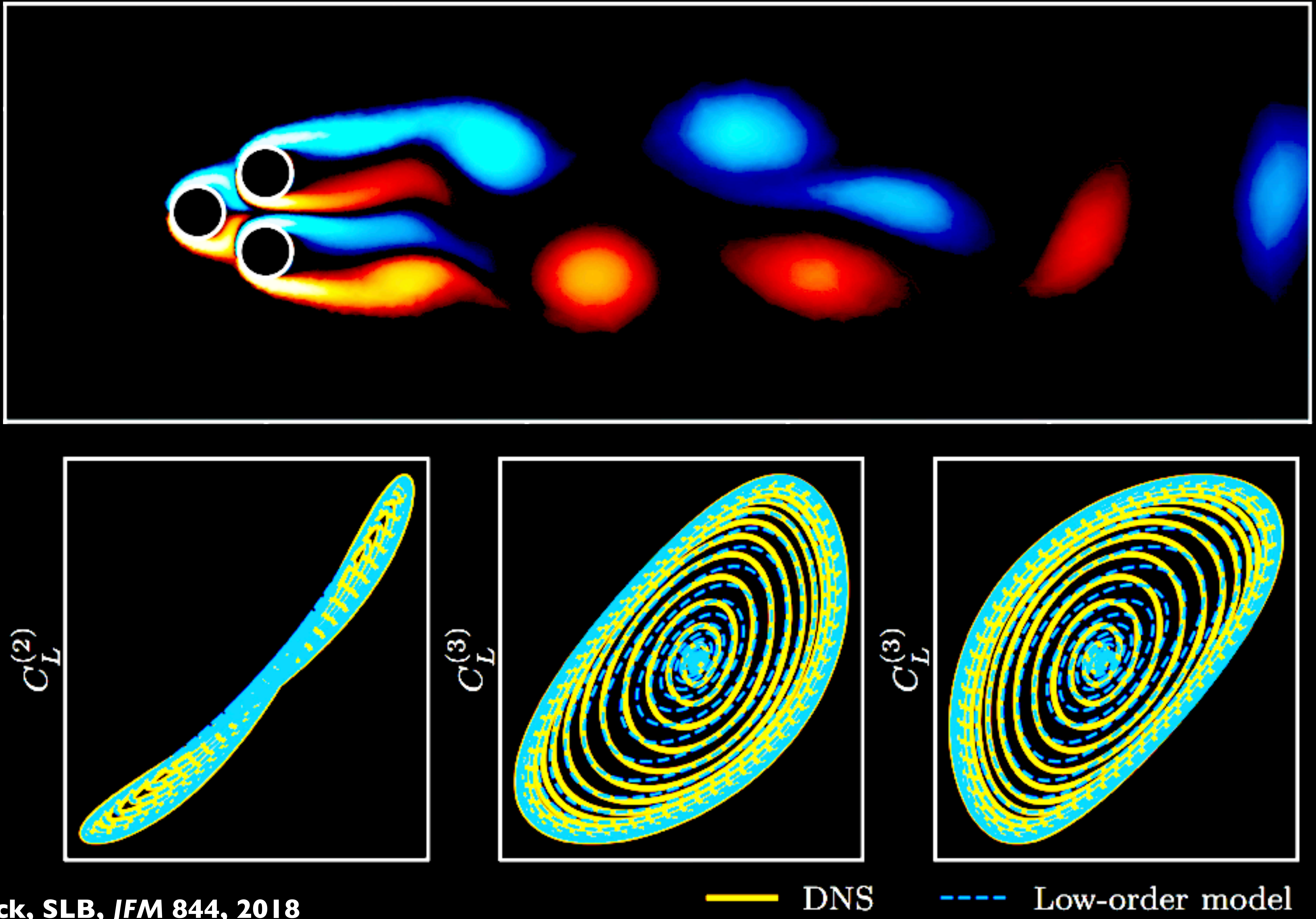


$$\ddot{x} - \underbrace{\left(0.2 - 0.24x^2 - 0.15\dot{x}^2\right)}_{k(x,\dot{x})}\dot{x} + 1.26x = 0$$

Spring-Mass Damper with Nonlinear Damping!

Loiseau & SLB,
JFM 838, 2018

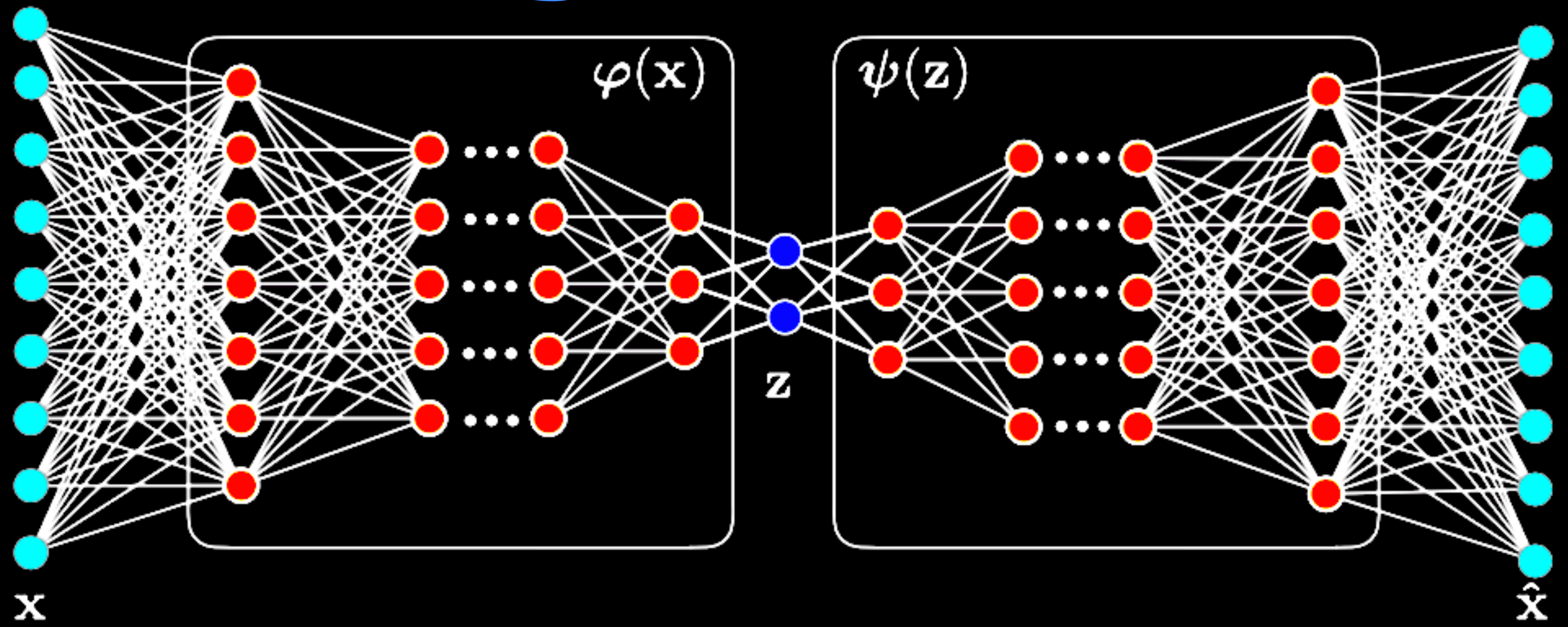
More Complex Flow: Fluidic Pinball



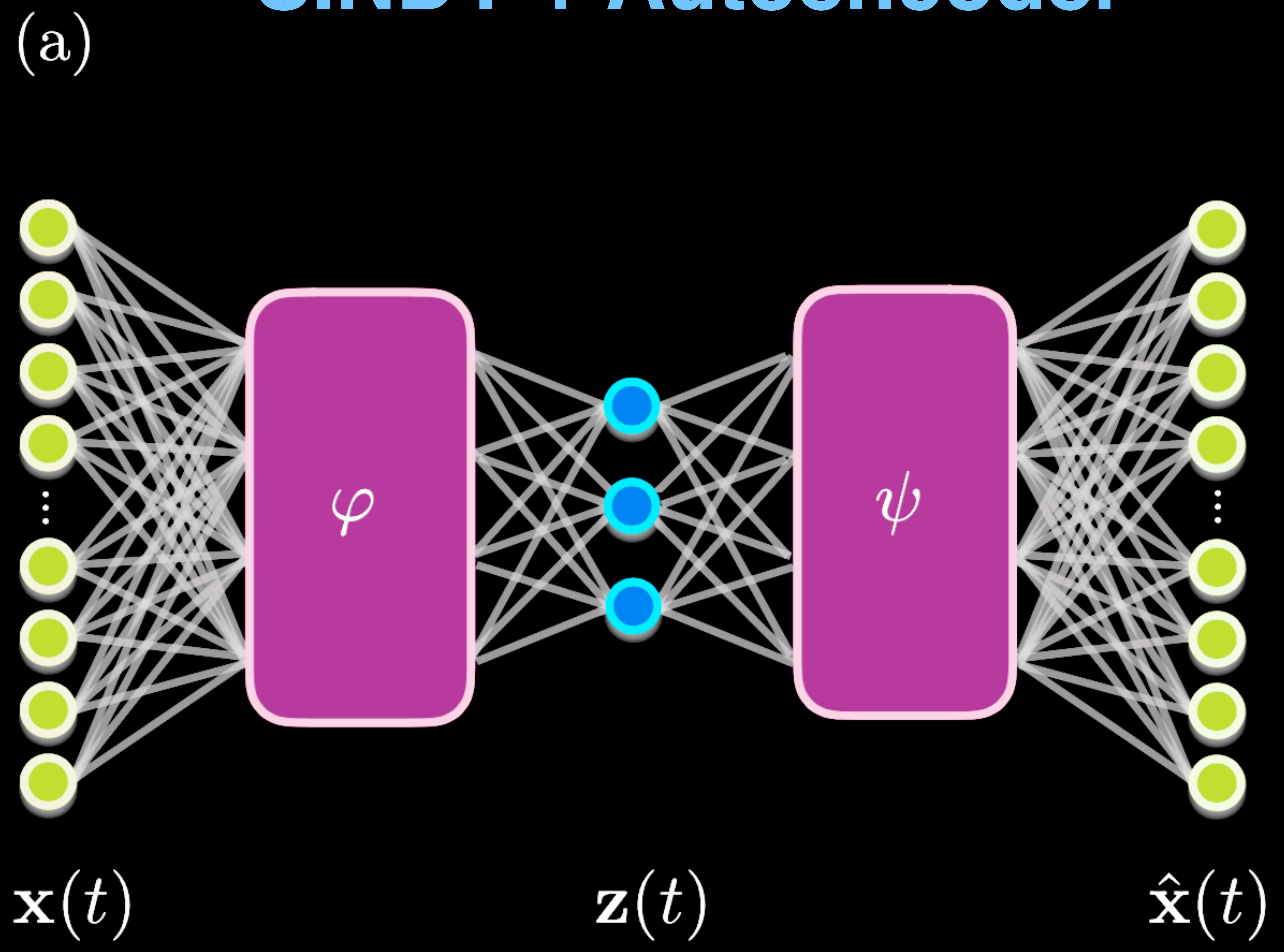
Loiseau, Noack, SLB, *JFM* 844, 2018

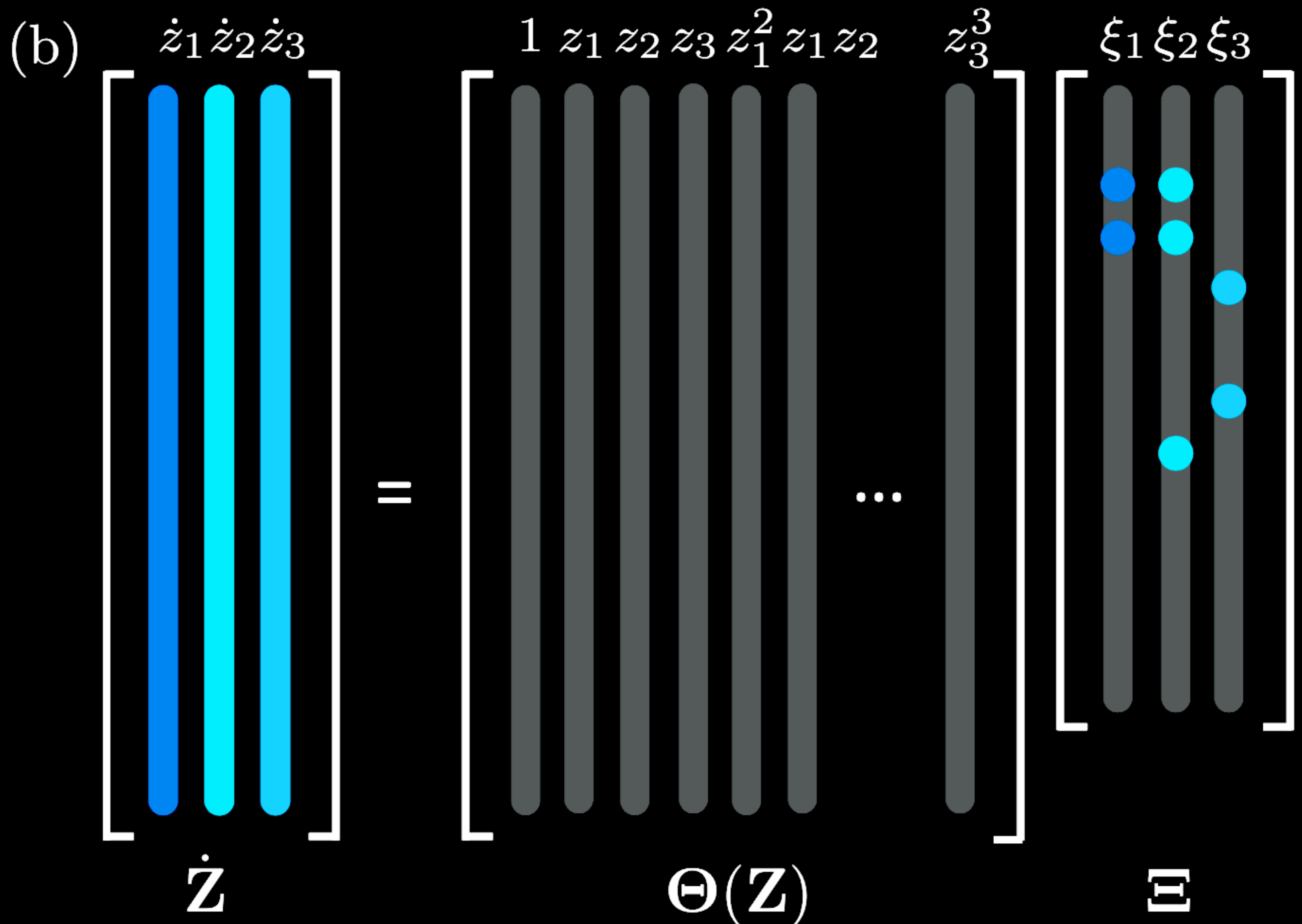
LATENT VARIABLES

DEEP LEARNING



SINDY + Autoencoder



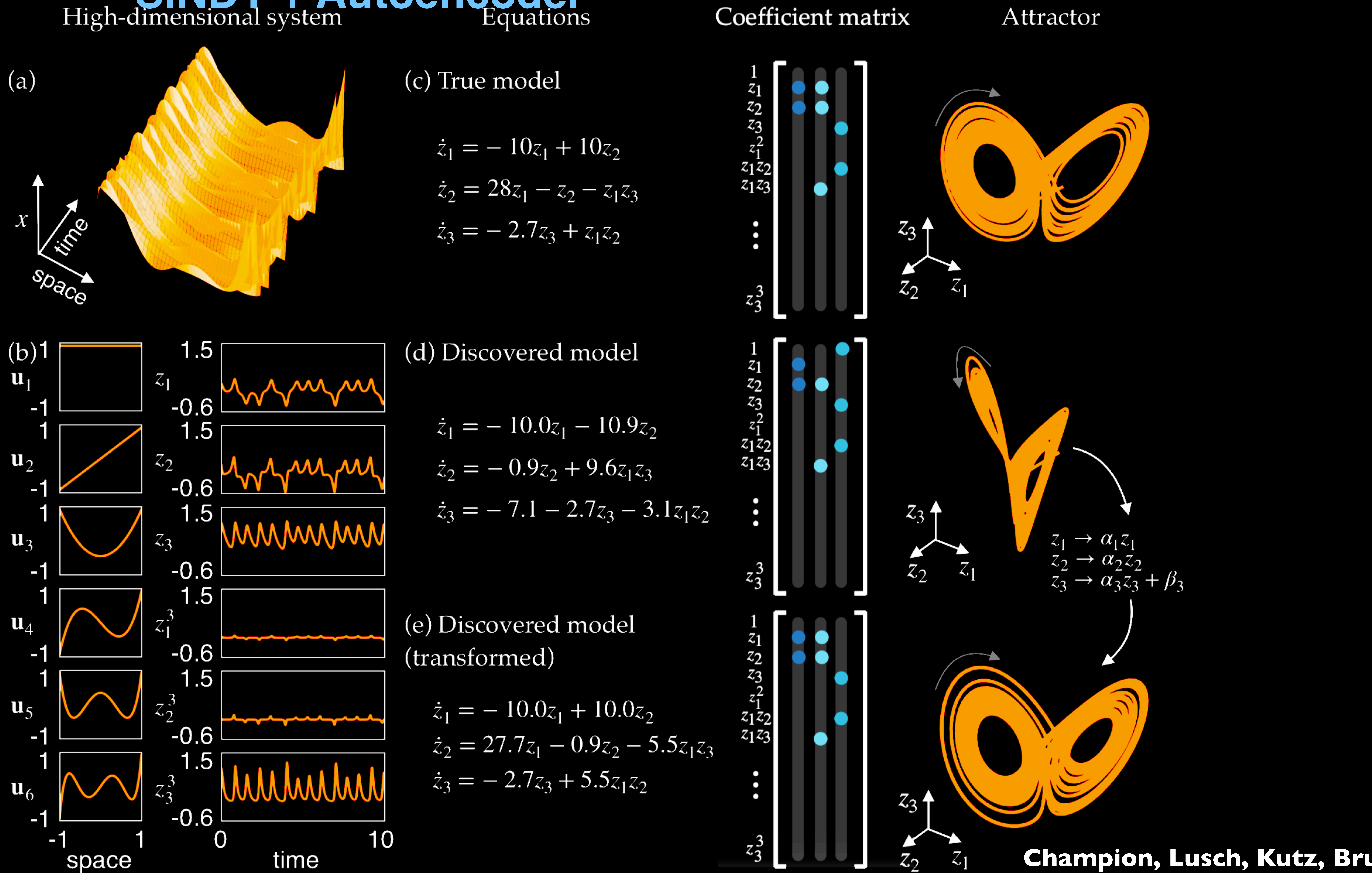


$\dot{\mathbf{z}}_i = \nabla_{\mathbf{x}}\varphi(\mathbf{x}_i)\dot{\mathbf{x}}_i \qquad \boldsymbol{\Theta}(\mathbf{z}_i^T) = \boldsymbol{\Theta}(\varphi(\mathbf{x}_i)^T)$

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

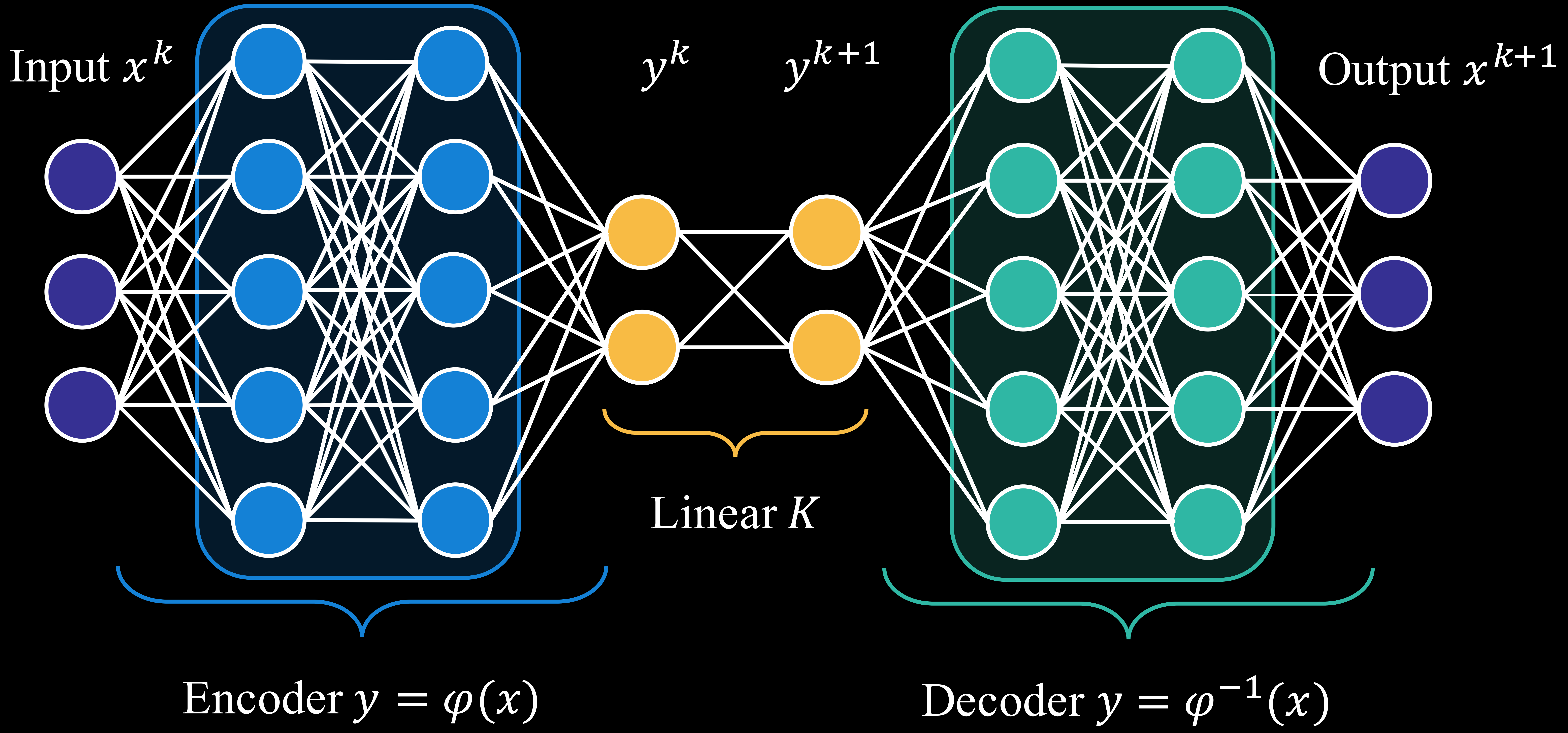
Champion, Lusch, Kutz, Brunton. *PNAS*, 2019

SINDY + Autoencoder



Champion, Lusch, Kutz, Brunton. *PNAS*, 2019

Neural Network to Find Koopman Eigenfunctions



Related work on Deep Learning Koopman:

- **Mardt, Pasquali, Wu, and Noé, *Nat. Comm.* 2018**
  arXiv:1710.06012, 2017
- **Wehmeyer and Noé, *J. Chem. Phys.* 2018**
  arXiv:1710.11239, 2017
- **Yeung, Kundu, Hodas (PNNL),**
  arXiv:1708.06850, 2017
- **Takeishi, Kawahara, and Yairi (RIKEN), *NIPS* 2017**
  arXiv:1710.04340
- **Otto and Rowley, *SIADS 2019***
  arXiv:1712.01378

Lusch, Kutz, Brunton, *Nat. Comm.* 2018

THEMES

Often Equations are Unknown or Too complex to work with:

- **Model discovery with machine learning**
- **Discover Reduced Order Models with machine learning**

Dynamics are Nonlinear and High-Dimensional:

- **Coordinate transformations to linearize dynamics**
- **Patterns facilitate sparse measurements**

Proposed approach:

- **Learn physics from data: interpretable & generalizable**
- **Respect known, or partially known, physics**