Data-driven Discovery of Coordinates and Governing Equations

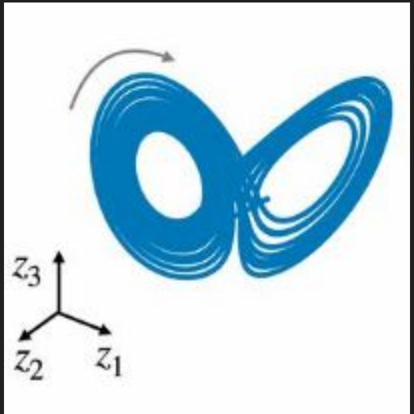
Authors: Kathleen Champion, Bethany Lusch, J. Nathan Kutz, Steven L. Brunton
University of Washington, Seattle, WA

Slides by: Monique Shotande

Motivation

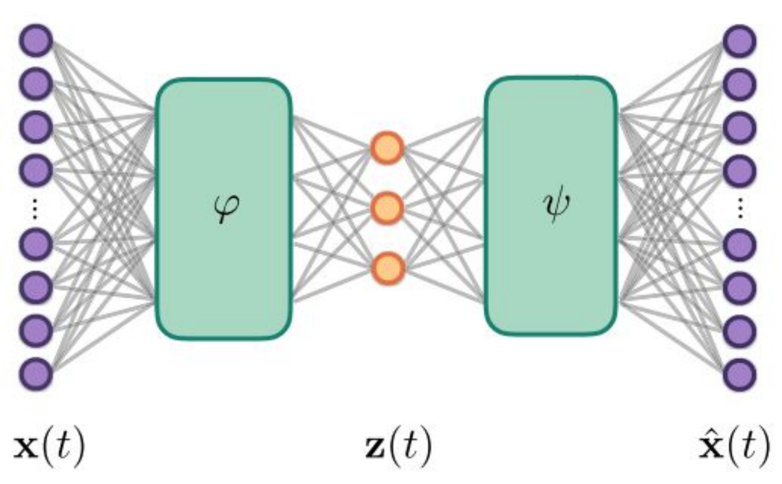
- Deficit of well-characterized quantitative descriptions of most problems
- Balance model complexity with descriptiveness for interpretable and generalizable models

$$\dot{z}_1 = -10z_1 + 10z_2$$
$$\dot{z}_2 = 28z_1 - z_2 - z_1z_3$$
$$\dot{z}_3 = -2.7z_3 + z_1z_2$$



Aims

- Discover sparse and interpretable dynamical models by just observing the dynamical system
- Understand and predict dynamics for complex systems
- Design AEs to automate discovery of coordinate transformation into a reduced space, sparsely representing dynamics



Approach

Incorporate SINDy Algorithm into objective of an autoencoder

SINDy Algorithm

SINDy (Sparse Identification of Nonlinear Dynamics) frames model discovery as a sparse regression problem

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

SINDy Algorithm

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

$$X = \begin{bmatrix} x_1 \\ \dots \\ x_m \end{bmatrix}$$

$$\dot{X} = \begin{bmatrix} \dot{x}_1 \\ \dots \\ \dot{x}_m \end{bmatrix}$$

$$X, \dot{X} \in R^{m \times n}$$

$$x(t) \in R^n$$

$$\dot{x} = \frac{d}{dt}x(t) = f(x(t))$$

SINDy Algorithm

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

$$X = \begin{bmatrix} x_1 \\ \dots \\ x_m \end{bmatrix}$$

$$\dot{X} = \begin{bmatrix} \dot{x}_1 \\ \dots \\ \dot{x}_m \end{bmatrix}$$

$$X, \dot{X} \in R^{m \times n}$$

$$x(t) \in R^n$$

$$\dot{x} = \frac{d}{dt}x(t) = f(x(t))$$

Form of the dynamical system;
f represents dynamical constraints defining the equations

SINDy Algorithm

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

$$\Xi = [\xi_1 \quad \dots \quad \xi_n] \in R^{p \times n}$$

$$\Theta(X) = [\theta_1(X) \quad \dots \quad \theta_p(X)] \in R^{m \times p}$$

$\mathbf{\Xi}$ unknown set of coefficients determining active terms of $\mathbf{\Theta}(\mathbf{X})$

Sparsity promoting regression solves for $\mathbf{\Xi}$ to select a few columns of $\mathbf{\Theta}(\mathbf{X})$

SINDy Algorithm

Library of candidate basis functions

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

Brunton, Steven L et al. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." Proceedings of the National Academy of Sciences of the United States of America vol. 113,15 (2016): 3932-7. doi:10.1073/pnas.1517384113

SINDy Algorithm

$f(x(t))$ can be constructed from a library of p candidate functions

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

$$x(t) \in R^n$$

$$\dot{x} = \frac{d}{dt}x(t) = f(x(t))$$

$$\Theta(X) = [\theta_1(X) \quad \dots \quad \theta_p(X)] \in R^{m \times p}$$

Each θ is a candidate model term

Assume m>>p

Neural Networks

Strengths

- Universal function approximators
- Learn nonlinear transformations

Challenges

- Generalization
- Extrapolation
- Interpretation

Despite challenges, potential to learn general, interpretable dynamic models using *proper constraints*

SINDy AEs

- SINDy to impose sparsity and interpretability
- Discover sparse dynamical models and coordinates enabling simple representations
- NNs for universal function approximation
- SINDy AE performs joint optimization to discover intrinsic coordinates containing associates parsimonious nonlinear dynamical model

SINDy AEs

Dynamical systems of form: $\dot{x} = \frac{d}{dt}x(t) = f(x(t))$

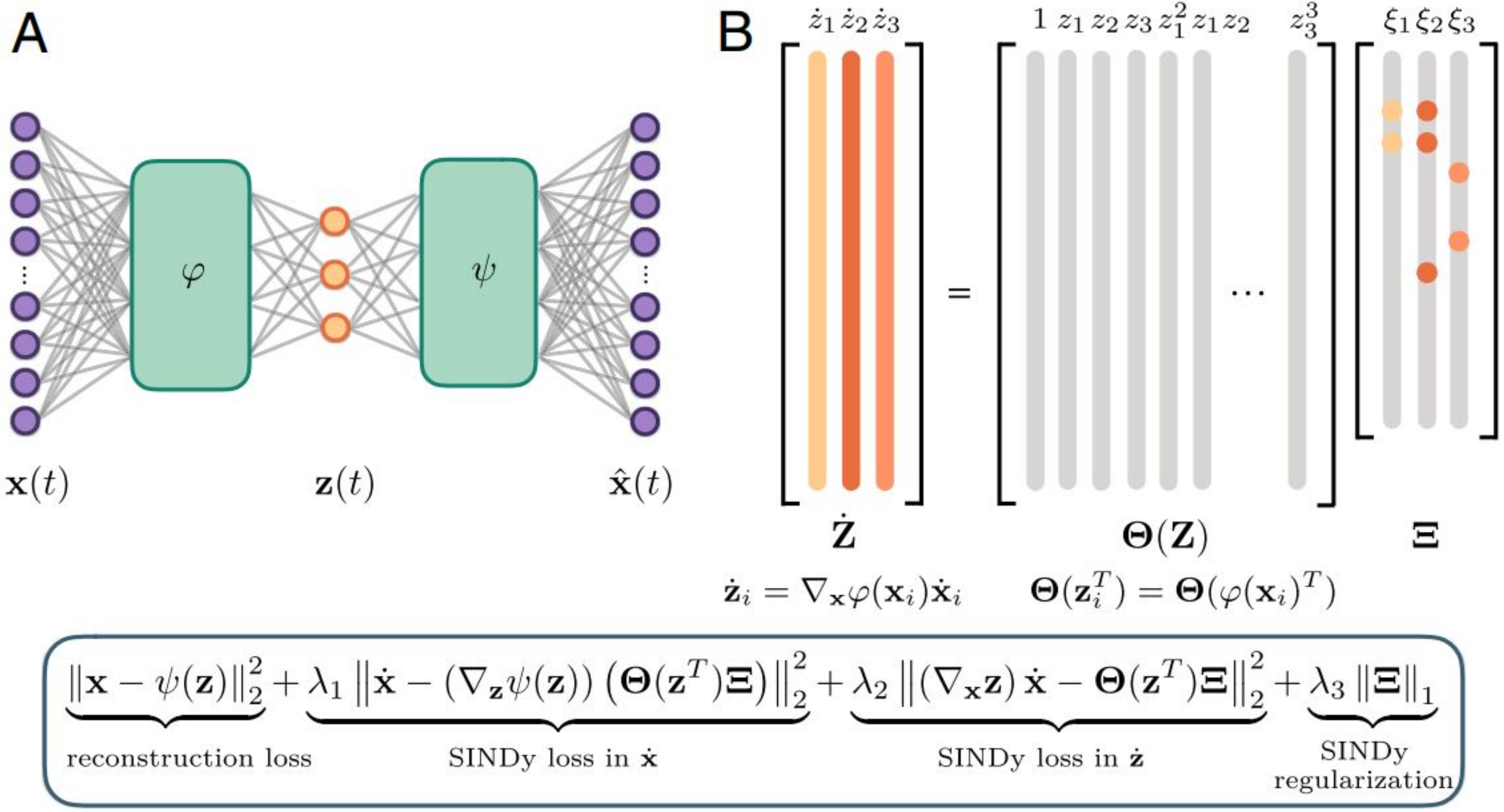
Original measurement coordinates x

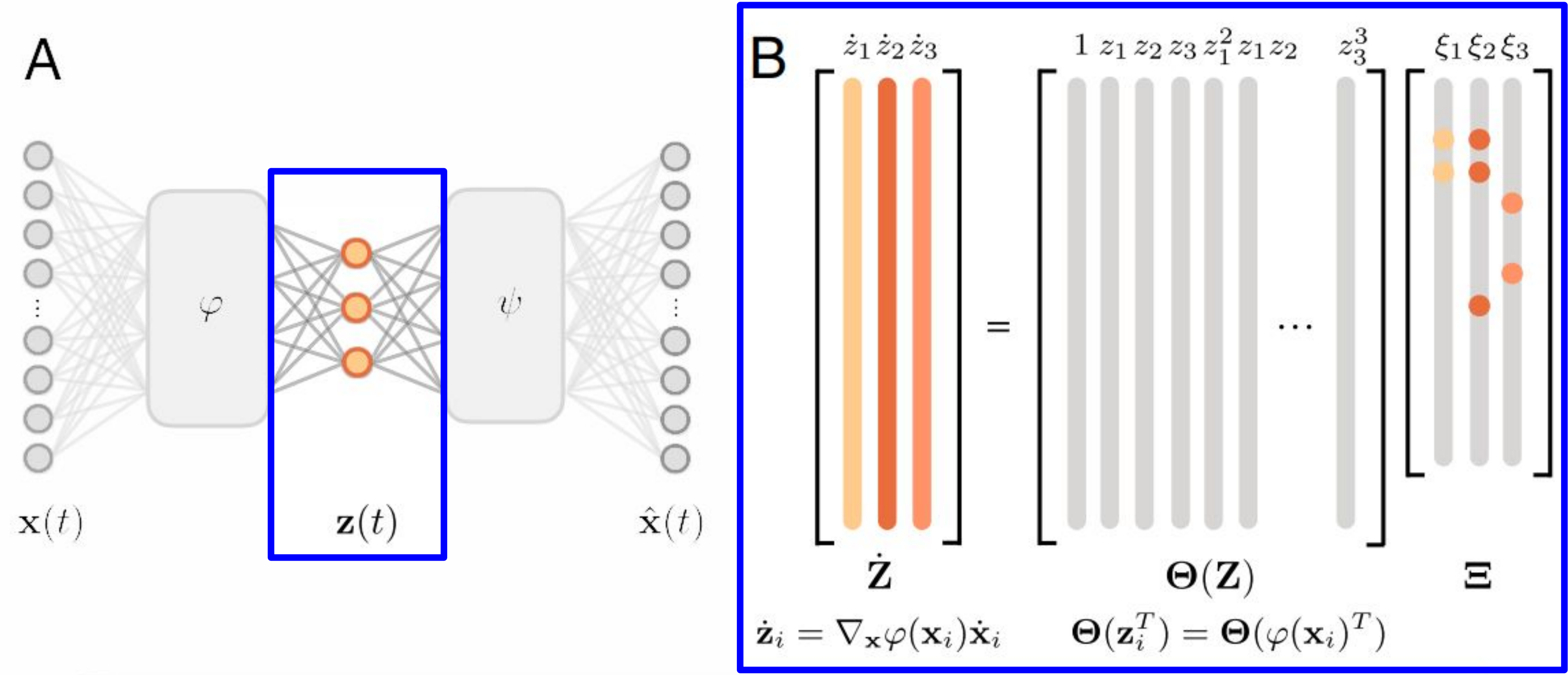
Discover reduced coordinate: $z(t) = \varphi(x(t)) \in R^d \; (d \ll n)$

- Associated dynamical model: $\dot{z} = \frac{d}{dt}z(t) = g(z(t))$
- Yields parsimonious description of dynamics

Coordinate transformations: $\boldsymbol{\varphi}$, $\boldsymbol{\psi}$

- φ: $x \mapsto z$ (encoder)
- ψ: $z \mapsto \sim x$ (decoder)

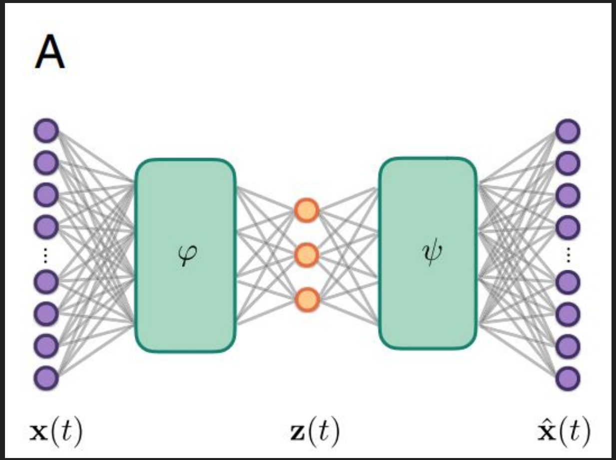

A
φ
ψ
$\mathbf{x}(t)$
$\mathbf{z}(t)$
$\hat{\mathbf{x}}(t)$
B
$\dot{z}_1 \; \dot{z}_2 \; \dot{z}_3$
$1 \; z_1 \; z_2 \; z_3 \; z_1^2 \; z_1 z_2 \; \cdots \; z_3^3$
$\xi_1 \; \xi_2 \; \xi_3$
$=$
$\dot{\mathbf{Z}}$
$\boldsymbol{\Theta}(\mathbf{Z})$
$\boldsymbol{\Xi}$
$\dot{\mathbf{z}}_i = \nabla_{\mathbf{x}} \varphi(\mathbf{x}_i) \dot{\mathbf{x}}_i$
$\boldsymbol{\Theta}(\mathbf{z}_i^T) = \boldsymbol{\Theta}(\varphi(\mathbf{x}_i)^T)$
$\|\mathbf{x} - \psi(\mathbf{z})\|_2^2 + \lambda_1 \left\|\dot{\mathbf{x}} - (\nabla_{\mathbf{z}} \psi(\mathbf{z})) \left(\boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right)\right\|_2^2 + \lambda_2 \left\|(\nabla_{\mathbf{x}} \mathbf{z}) \dot{\mathbf{x}} - \boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right\|_2^2 + \lambda_3 \|\boldsymbol{\Xi}\|_1$
reconstruction loss
SINDy loss in $\dot{\mathbf{x}}$
SINDy loss in $\dot{\mathbf{z}}$
SINDy regularization



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

SINDy AEs

Overall loss

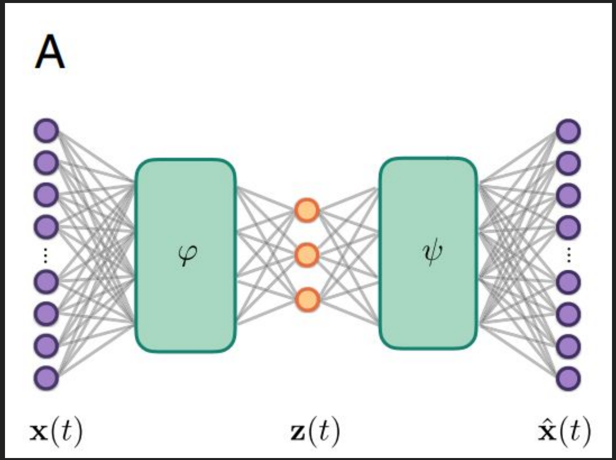


$$\mathcal{L}_{\text{recon}} + \lambda_1 \mathcal{L}_{d\mathbf{x}/dt} + \lambda_2 \mathcal{L}_{d\mathbf{z}/dt} + \lambda_3 \mathcal{L}_{\text{reg}},$$

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

SINDy AEs

Overall loss



$$\mathcal{L}_{\text{recon}} + \lambda_1 \mathcal{L}_{d\mathbf{x}/dt} + \lambda_2 \mathcal{L}_{d\mathbf{z}/dt} + \lambda_3 \mathcal{L}_{\text{reg}},$$

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

SINDy predictions should be able to *reconstruct original time derivatives*

SINDy AEs

Make SINDy prediction useable for reconstruction of original time derivatives

$$\underbrace{\left\| \dot{\mathbf{x}} - \left(\nabla_{\mathbf{z}}\psi(\mathbf{z})\right)\left(\Theta(\mathbf{z}^T)\Xi\right) \right\|_2^2}_{\text{SINDy loss in } \dot{\mathbf{x}}}$$

$$\mathcal{L}_{d\mathbf{x}/dt} = \left\| \dot{\mathbf{x}} - \left(\nabla_{\mathbf{z}}\psi(\varphi(\mathbf{x}))\right)\left(\Theta(\varphi(\mathbf{x})^T)\Xi\right) \right\|_2^2. \qquad \textbf{[4]}$$

SINDy AEs

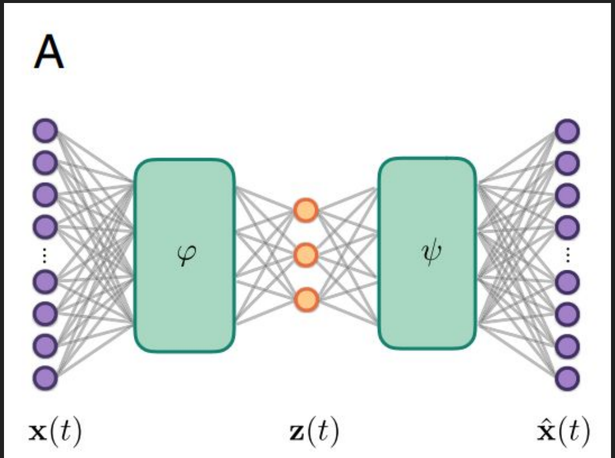
Make SINDy prediction useable for reconstruction of original time derivatives

$$\left\| \dot{\mathbf{x}} - \left(\nabla_{\mathbf{z}} \psi(\mathbf{z})\right) \left(\Theta(\mathbf{z}^T)\Xi\right) \right\|_2^2$$

$$\frac{\mathrm{d}x}{\mathrm{d}t} \qquad \frac{\mathrm{d}x}{\mathrm{d}z} \qquad \frac{\mathrm{d}z}{\mathrm{d}t}$$

SINDy AEs

Overall loss



$$\mathcal{L}_{\text{recon}} + \lambda_1 \mathcal{L}_{d\mathbf{x}/dt} + \lambda_2 \mathcal{L}_{d\mathbf{z}/dt} + \lambda_3 \mathcal{L}_{\text{reg}},$$

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

Use SINDy model with encoder gradient to *encourage prediction of encoder variables' time derivatives*

SINDy AEs

Guarantee learned latent space has associated sparse dynamical model, simultaneously learn SINDy model for the dynamics of the intrinsic coordinates z

$$\underbrace{\left\|(\nabla_{\mathbf{x}}\mathbf{z})\,\dot{\mathbf{x}} - \boldsymbol{\Theta}(\mathbf{z}^T)\boldsymbol{\Xi}\right\|_2^2}_{\text{SINDy loss in } \dot{\mathbf{z}}}$$

$$\mathcal{L}_{d\mathbf{z}/dt} = \left\|\nabla_{\mathbf{x}}\varphi(\mathbf{x})\dot{\mathbf{x}} - \boldsymbol{\Theta}(\varphi(\mathbf{x})^T)\boldsymbol{\Xi}\right\|_2^2. \quad [3]$$

SINDy AEs

Guarantee learned latent space has associated sparse dynamical model, simultaneously learn SINDy model for the dynamics of the intrinsic coordinates z

$$\mathcal{L}_{d\mathbf{z}/dt} = \left\| \nabla_{\mathbf{x}} \varphi(\mathbf{x}) \dot{\mathbf{x}} - \Theta(\varphi(\mathbf{x})^T) \Xi \right\|_2^2 . \quad [3]$$

$$\frac{dz}{dx} \quad \frac{dx}{dt} \qquad \frac{dz}{dt}$$

SINDy AEs

This regularization is achieved by constructing a library $\mathbf{\Theta}(\mathbf{z})$ of candidate basis functions and learning a sparse set of coefficients $\mathbf{\Xi}=[\xi_1 \ldots \xi_d] \in R^{p\times d}$

$$\mathcal{L}_{d\mathbf{z}/dt} = \left\| \nabla_{\mathbf{x}}\varphi(\mathbf{x})\dot{\mathbf{x}} - \mathbf{\Theta}(\varphi(\mathbf{x})^T)\mathbf{\Xi} \right\|_2^2. \qquad [3]$$

$$\dot{z}(t) = \frac{d}{dt}z(t) = g(z(t)) = \Theta(z(t))\Xi$$

SINDy AEs

Regularization achieved by constructing a library $\boldsymbol{\Theta}(\mathbf{z})$ of candidate basis functions and learning a sparse set of coefficients $\boldsymbol{\Xi}$

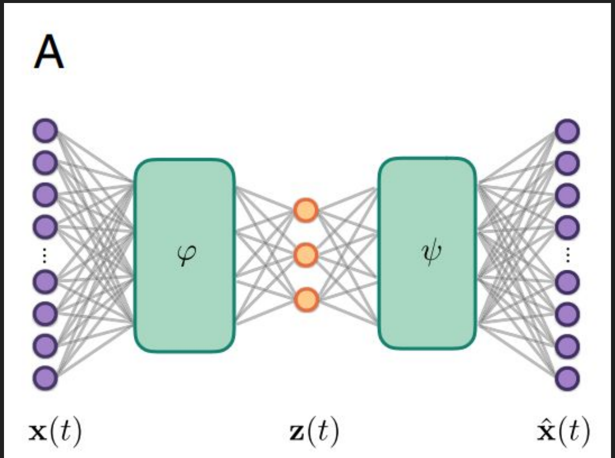
$$\mathcal{L}_{d\mathbf{z}/dt} = \left\| \nabla_{\mathbf{x}} \varphi(\mathbf{x}) \dot{\mathbf{x}} - \boldsymbol{\Theta}(\varphi(\mathbf{x})^{T}) \boldsymbol{\Xi} \right\|_2^2. \quad [3]$$

$$\dot{z}(t) = \frac{d}{dt} z(t) = g(z(t)) = \Theta(z(t))\Xi$$

$$\Xi = [\xi_1 \quad \dots \quad \xi_d] \in R^{p \times d}$$

SINDy AEs

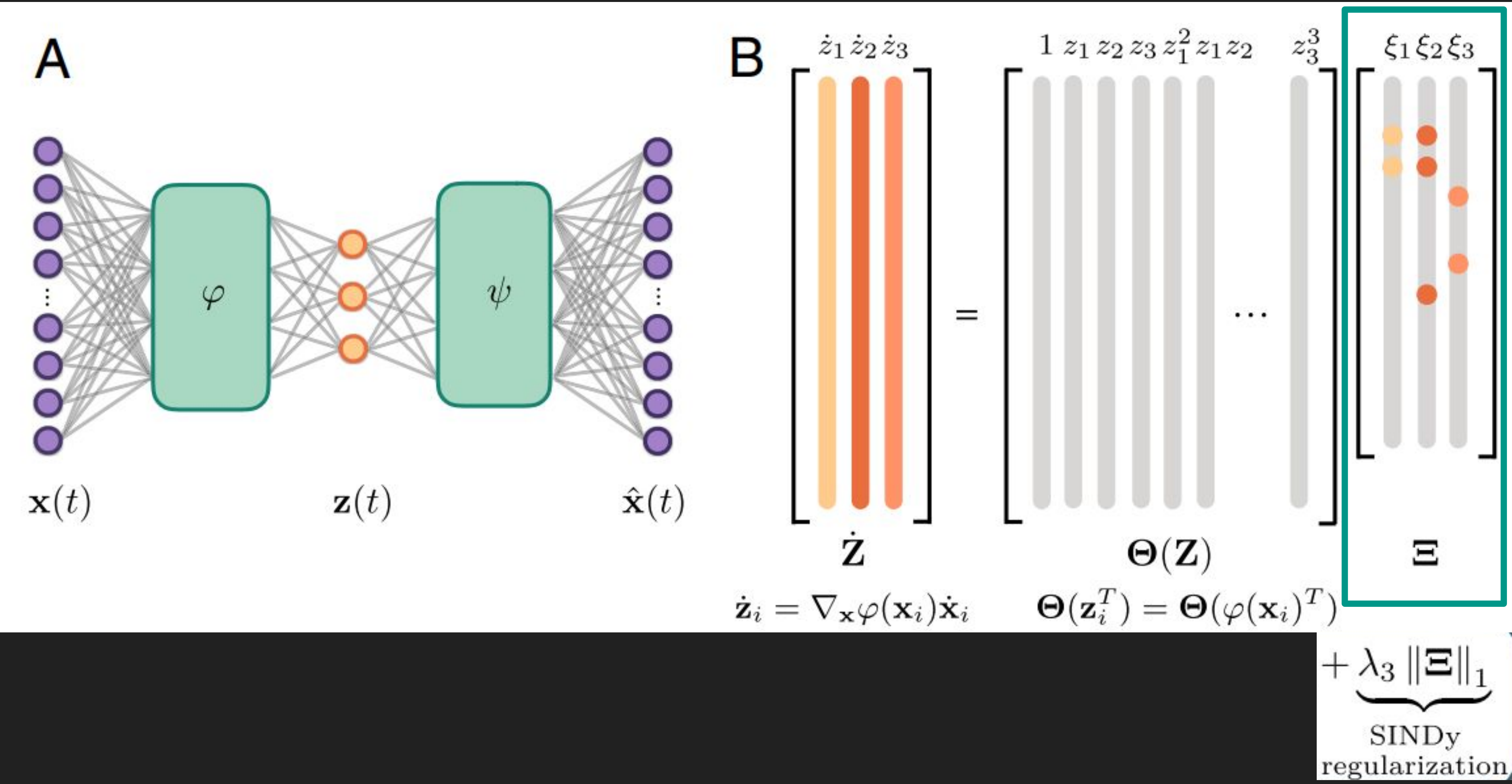
Overall loss



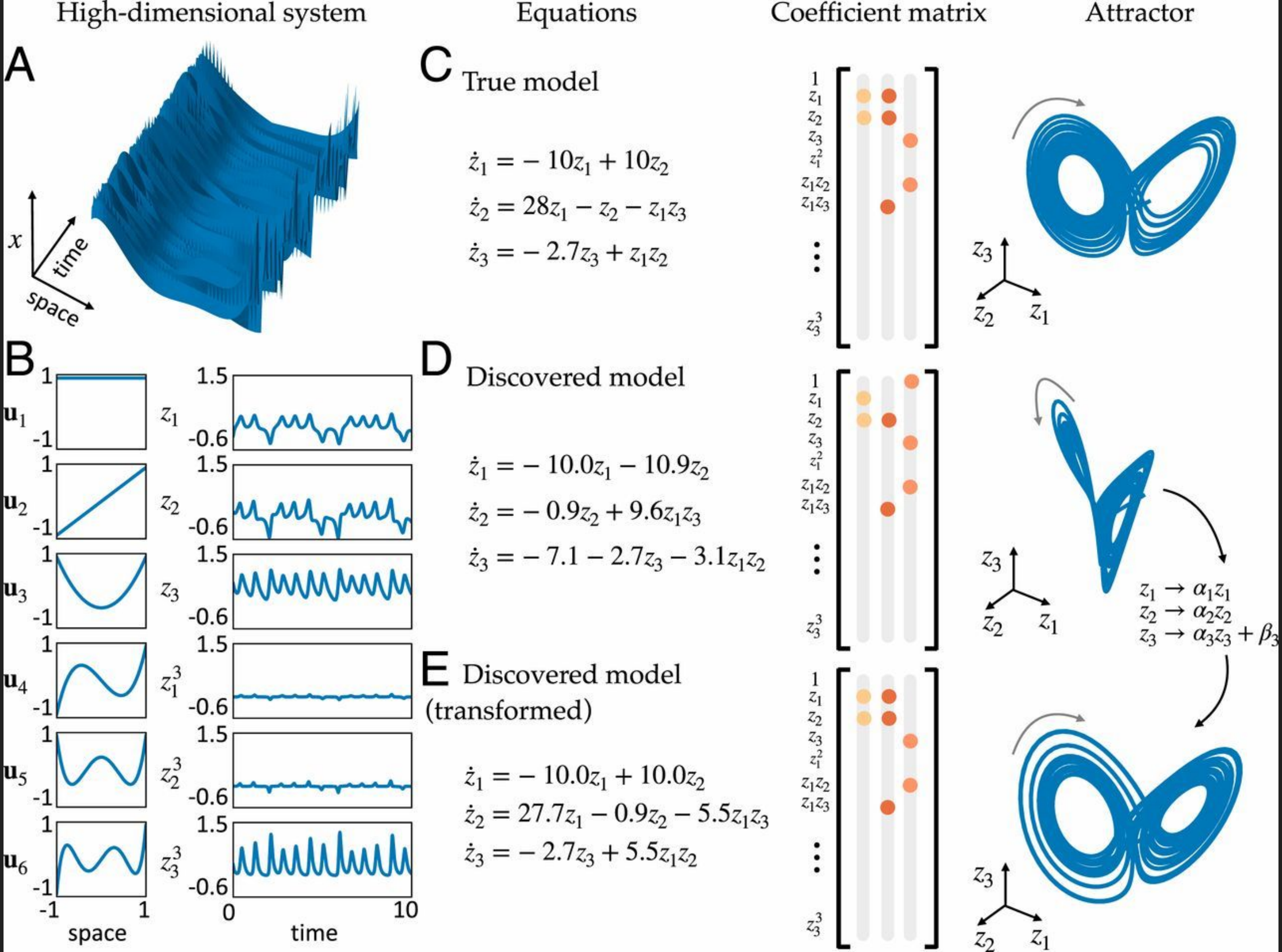
$$\mathcal{L}_{\text{recon}} + \lambda_1 \mathcal{L}_{d\mathbf{x}/dt} + \lambda_2 \mathcal{L}_{d\mathbf{z}/dt} + \lambda_3 \mathcal{L}_{\text{reg}},$$

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

L1 regularization on SINDy coefficients to *promote sparsity*


A
φ
ψ
x(t)
z(t)
x̂(t)
B
ż1 ż2 ż3
1 z1 z2 z3 z1^2 z1z2 ... z3^3
ξ1 ξ2 ξ3
=
Ż
Θ(Z)
Ξ
ż_i = ∇_x φ(x_i) ẋ_i
Θ(z_i^T) = Θ(φ(x_i)^T)
+ λ3 ||Ξ||_1
SINDy
regularization

Results



High-dimensional system
Equations
Coefficient matrix
Attractor
A
x
time
space
B
$\mathbf{u}_1$
$\mathbf{u}_2$
$\mathbf{u}_3$
$\mathbf{u}_4$
$\mathbf{u}_5$
$\mathbf{u}_6$
z_1
z_2
z_3
z_1^3
z_2^3
z_3^3
1
-1
1.5
-0.6
space
time
0
10
C True model
$\dot{z}_1 = -10z_1 + 10z_2$
$\dot{z}_2 = 28z_1 - z_2 - z_1z_3$
$\dot{z}_3 = -2.7z_3 + z_1z_2$
D Discovered model
$\dot{z}_1 = -10.0z_1 - 10.9z_2$
$\dot{z}_2 = -0.9z_2 + 9.6z_1z_3$
$\dot{z}_3 = -7.1 - 2.7z_3 - 3.1z_1z_2$
E Discovered model (transformed)
$\dot{z}_1 = -10.0z_1 + 10.0z_2$
$\dot{z}_2 = 27.7z_1 - 0.9z_2 - 5.5z_1z_3$
$\dot{z}_3 = -2.7z_3 + 5.5z_1z_2$
1
z_1
z_2
z_3
z_1^2
z_1z_2
z_1z_3
z_3^3
$z_1 \to \alpha_1 z_1$
$z_2 \to \alpha_2 z_2$
$z_3 \to \alpha_3 z_3 + \beta_3$

Creation of Training Data

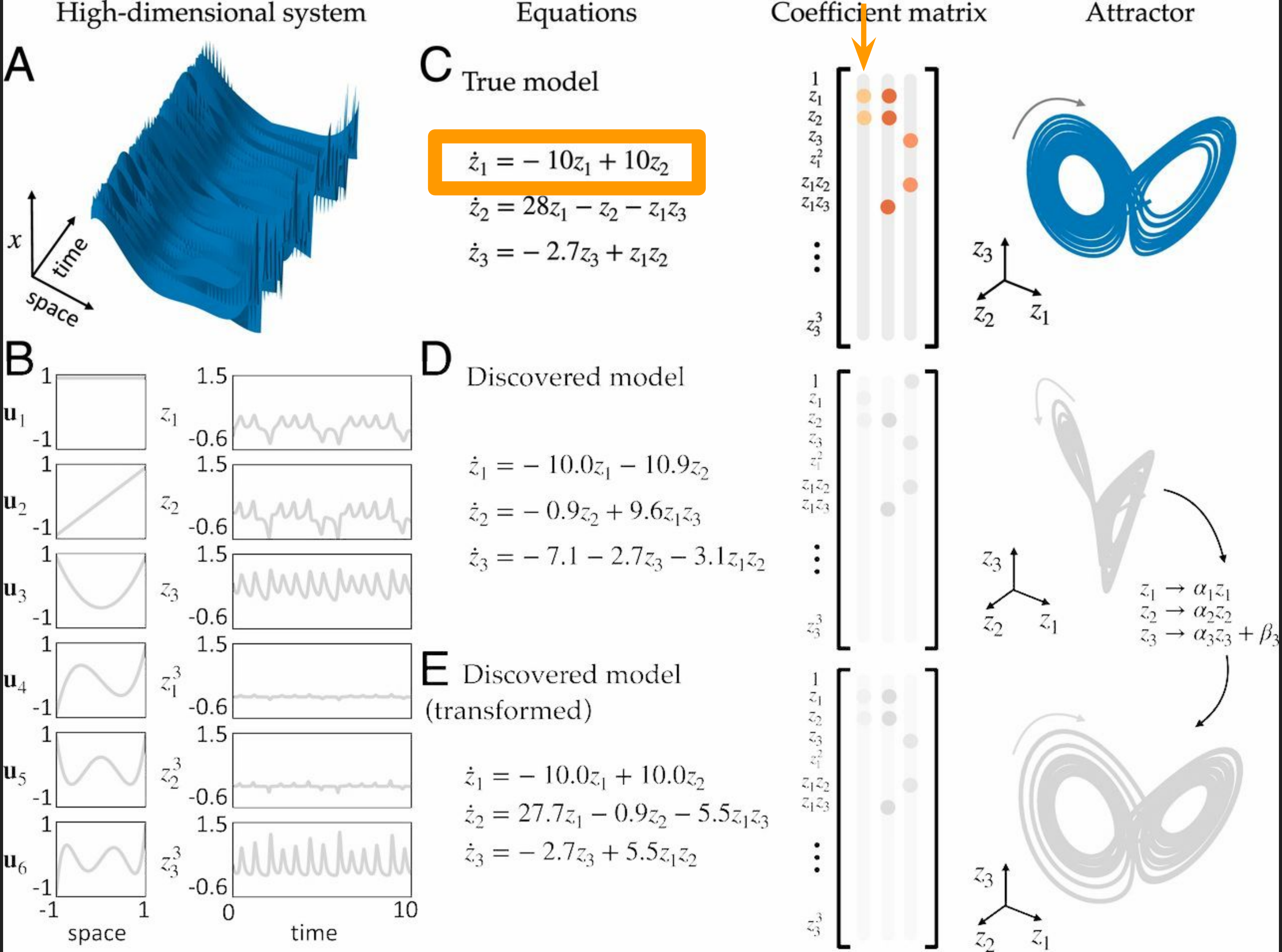
- True dynamics —------------------->

$$\dot{z}_1 = -10z_1 + 10z_2$$
$$\dot{z}_2 = 28z_1 - z_2 - z_1z_3$$
$$\dot{z}_3 = -2.7z_3 + z_1z_2$$

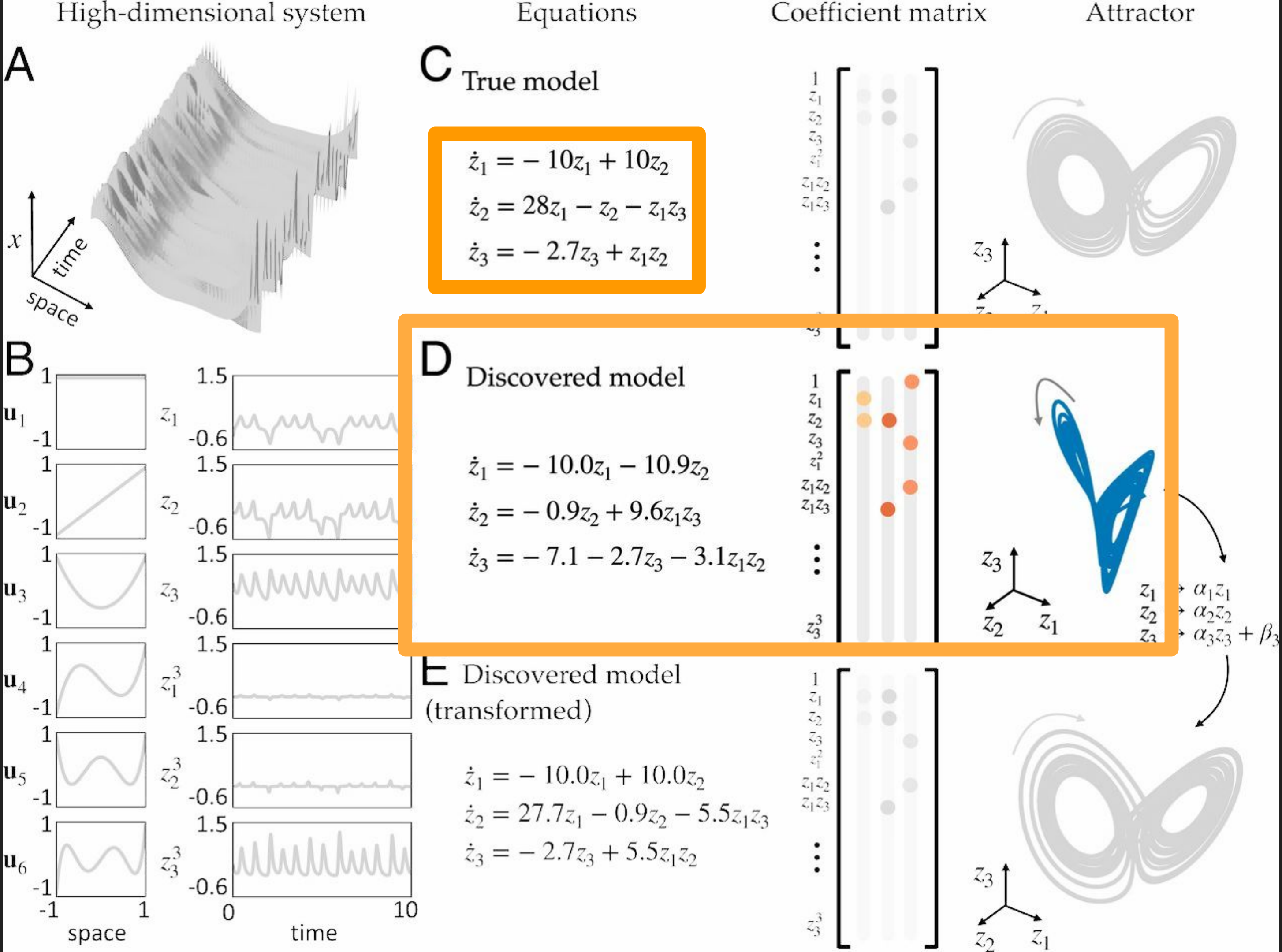
- Create data set:
  - Integrate system in time
  - Transform state (z1, z2, z3) using an expansive, non-linear transformation. Each dimension of this expanded representation is of the form:

$$\mathbf{x}(t) = \mathbf{u}_1 z_1(t) + \mathbf{u}_2 z_2(t) + \mathbf{u}_3 z_3(t) + \mathbf{u}_4 z_1(t)^3 + \mathbf{u}_5 z_2(t)^3 + \mathbf{u}_6 z_3(t)^3.$$

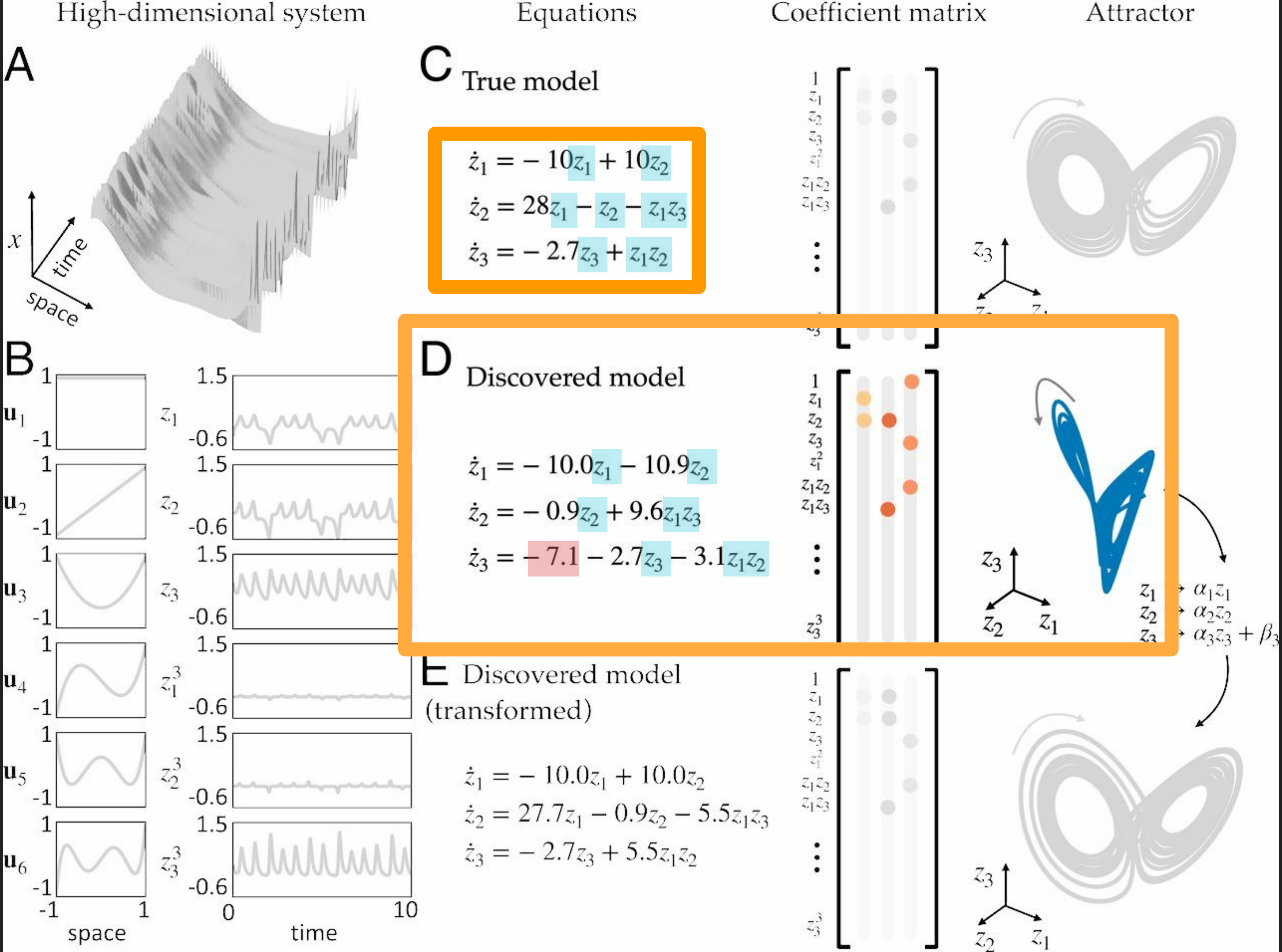
  - The model never sees z1, z2, z3!



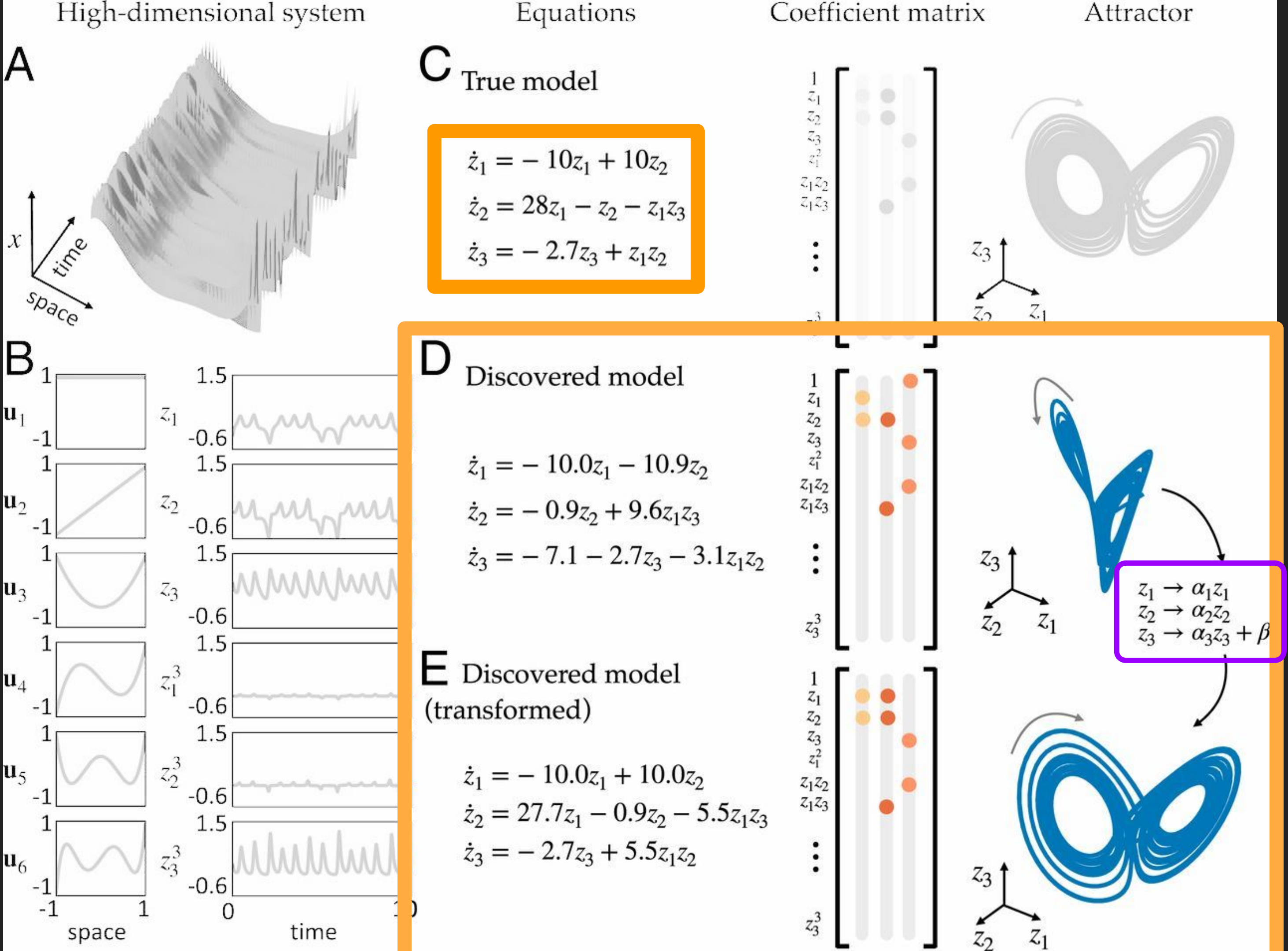
High-dimensional system
Equations
Coefficient matrix
Attractor
A
x
time
space
B
C
True model
$\dot{z}_1 = -10z_1 + 10z_2$
$\dot{z}_2 = 28z_1 - z_2 - z_1z_3$
$\dot{z}_3 = -2.7z_3 + z_1z_2$
D
Discovered model
$\dot{z}_1 = -10.0z_1 - 10.9z_2$
$\dot{z}_2 = -0.9z_2 + 9.6z_1z_3$
$\dot{z}_3 = -7.1 - 2.7z_3 - 3.1z_1z_2$
E
Discovered model (transformed)
$\dot{z}_1 = -10.0z_1 + 10.0z_2$
$\dot{z}_2 = 27.7z_1 - 0.9z_2 - 5.5z_1z_3$
$\dot{z}_3 = -2.7z_3 + 5.5z_1z_2$
$z_1 \to \alpha_1 z_1$
$z_2 \to \alpha_2 z_2$
$z_3 \to \alpha_3 z_3 + \beta_3$



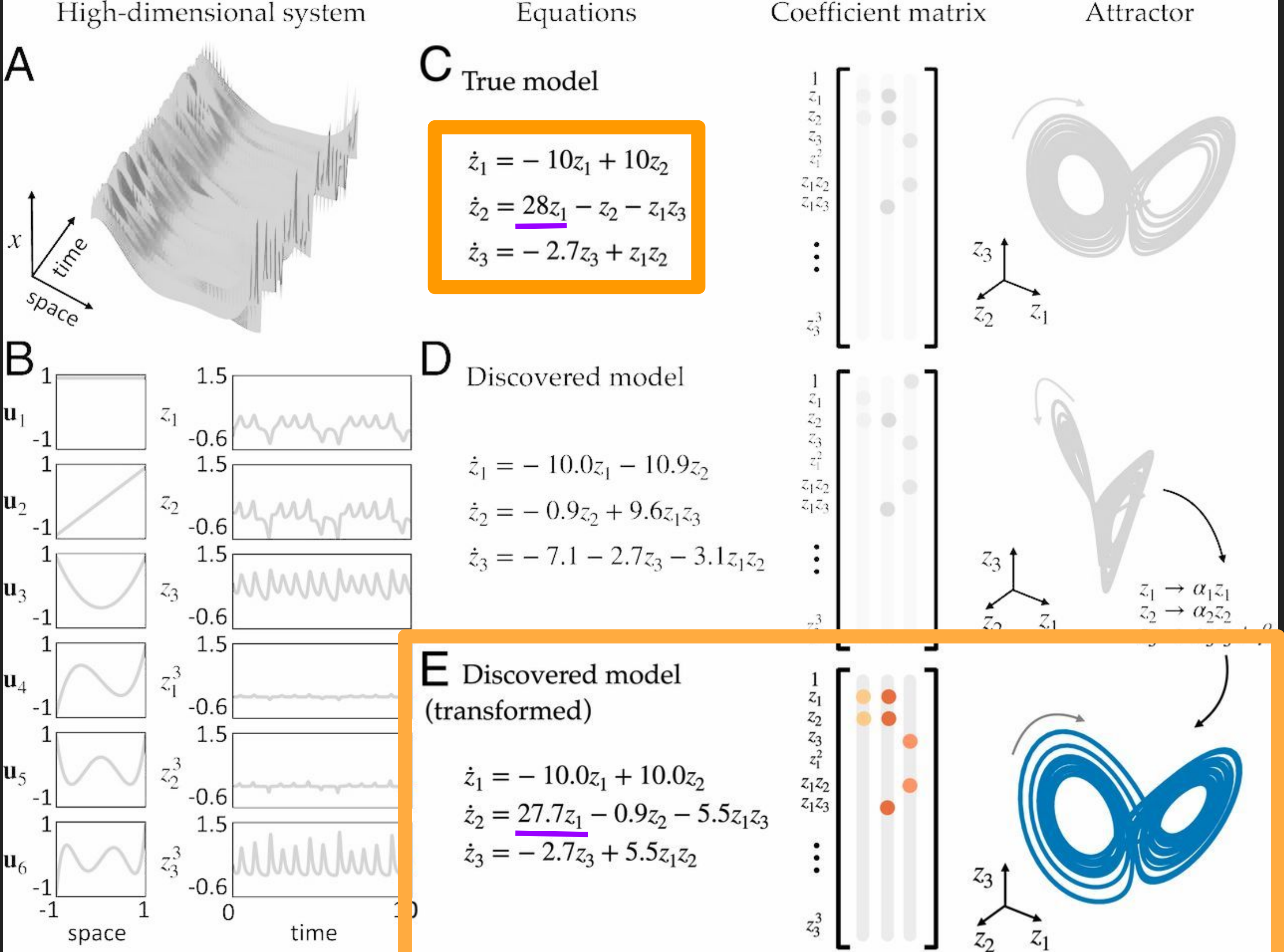
High-dimensional system
Equations
Coefficient matrix
Attractor
A
x
time
space
B
$\mathbf{u}_1$
$\mathbf{u}_2$
$\mathbf{u}_3$
$\mathbf{u}_4$
$\mathbf{u}_5$
$\mathbf{u}_6$
z_1
z_2
z_3
z_1^3
z_2^3
z_3^3
1.5
-0.6
-1
1
space
0
10
time
C
True model
$\dot{z}_1 = -10z_1 + 10z_2$
$\dot{z}_2 = 28z_1 - z_2 - z_1z_3$
$\dot{z}_3 = -2.7z_3 + z_1z_2$
D
Discovered model
$\dot{z}_1 = -10.0z_1 - 10.9z_2$
$\dot{z}_2 = -0.9z_2 + 9.6z_1z_3$
$\dot{z}_3 = -7.1 - 2.7z_3 - 3.1z_1z_2$
E
Discovered model
(transformed)
$\dot{z}_1 = -10.0z_1 + 10.0z_2$
$\dot{z}_2 = 27.7z_1 - 0.9z_2 - 5.5z_1z_3$
$\dot{z}_3 = -2.7z_3 + 5.5z_1z_2$
1
z_1
z_2
z_3
z_1^2
z_1z_2
z_1z_3
z_3^3
$z_1 \to \alpha_1 z_1$
$z_2 \to \alpha_2 z_2$
$z_3 \to \alpha_3 z_3 + \beta_3$



High-dimensional system
Equations
Coefficient matrix
Attractor
A
x
time
space
B
u1
u2
u3
u4
u5
u6
z1
z2
z3
space
time
C
True model
ż1 = − 10z1 + 10z2
ż2 = 28z1 − z2 − z1z3
ż3 = − 2.7z3 + z1z2
D
Discovered model
ż1 = − 10.0z1 − 10.9z2
ż2 = − 0.9z2 + 9.6z1z3
ż3 = − 7.1 − 2.7z3 − 3.1z1z2
E
Discovered model
(transformed)
ż1 = − 10.0z1 + 10.0z2
ż2 = 27.7z1 − 0.9z2 − 5.5z1z3
ż3 = − 2.7z3 + 5.5z1z2
z1 → α1z1
z2 → α2z2
z3 → α3z3 + β3



High-dimensional system
Equations
Coefficient matrix
Attractor
A
x
time
space
B
space
time
C True model
$\dot{z}_1 = -10z_1 + 10z_2$
$\dot{z}_2 = 28z_1 - z_2 - z_1z_3$
$\dot{z}_3 = -2.7z_3 + z_1z_2$
D Discovered model
$\dot{z}_1 = -10.0z_1 - 10.9z_2$
$\dot{z}_2 = -0.9z_2 + 9.6z_1z_3$
$\dot{z}_3 = -7.1 - 2.7z_3 - 3.1z_1z_2$
$z_1 \to \alpha_1 z_1$
$z_2 \to \alpha_2 z_2$
$z_3 \to \alpha_3 z_3 + \beta$
E Discovered model (transformed)
$\dot{z}_1 = -10.0z_1 + 10.0z_2$
$\dot{z}_2 = 27.7z_1 - 0.9z_2 - 5.5z_1z_3$
$\dot{z}_3 = -2.7z_3 + 5.5z_1z_2$



High-dimensional system
Equations
Coefficient matrix
Attractor
A
x
time
space
B
u1
u2
u3
u4
u5
u6
z1
z2
z3
z1^3
z2^3
z3^3
1.5
-0.6
1
-1
space
time
0
C True model
ż1 = − 10z1 + 10z2
ż2 = 28z1 − z2 − z1z3
ż3 = − 2.7z3 + z1z2
D Discovered model
ż1 = − 10.0z1 − 10.9z2
ż2 = − 0.9z2 + 9.6z1z3
ż3 = − 7.1 − 2.7z3 − 3.1z1z2
z1 → α1z1
z2 → α2z2
E Discovered model (transformed)
ż1 = − 10.0z1 + 10.0z2
ż2 = 27.7z1 − 0.9z2 − 5.5z1z3
ż3 = − 2.7z3 + 5.5z1z2
1
z1
z2
z3
z1^2
z1z2
z1z3
z3^3

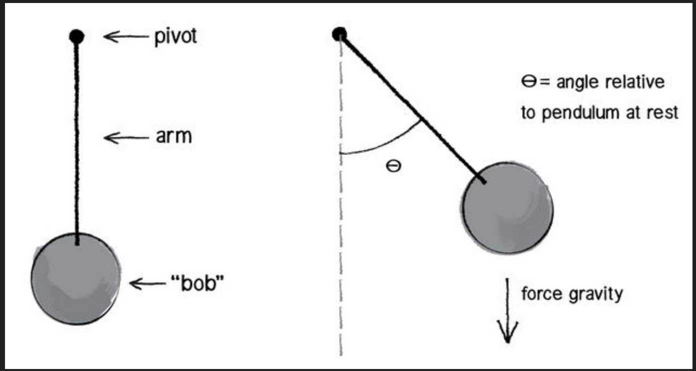
Lorenz Take-Aways

- Able to infer a low-dimensional representation of the state dynamics
- The coordinate frames between the original model and the inferred model do not have to be the same!
  - z1 does not have to match the model's z1
  - In practice (for this case), the relationship between the two spaces can be described as a rotation

Pendulum Example

- Minimal true state representation: angle + angular velocity
- Input to model: video showing the end of a pendulum in space
- Model must transform image to a low-dimensional state description

https://www.youtube.com/watch?v=WHhDgxkiR9c



www.khanacademy.org

Results

Test trajectories from 100 initial conditions sampled from the training distribution

Relative L_2 prediction errors

	x	**x_dot**	**z**	**z_dot**
Lorenz System	3x10^-5	2x10^-4		7x10^-4
Reaction-diffusion	.016	.016	1x10^-4	.002
Nonlinear pendulum	8x10^-4	x_ddot:3x10^-4		z_ddot:.02

Discussion

- AEs allow discovery of nonlinear, compressive coordinate transformations
- Quantitative and objective choice of coordinate measurements
- Approach is generalizable
- Approach yields classically interpretable models
- In regards to new scientific breakthroughs, must careful to ensure valid conclusions are drawn from results.
    - One approach: combine ML approaches with well-established domain knowledge.
    - Methods providing interpretable models potentially enable new discoveries in data-rich fields

Limitations

- Requires clean data
- SINDy requires reasonable derivative estimates

QUESTIONS?

SINDy AEs

Derivatives $\dot{x}(t)$ either available or computable

- calculate derivative of encoder variables
  - $\dot{z}(t) = \nabla_x \varphi(x(t))\dot{x}(t)$
- Enforce accurate prediction of dynamics by incorporating the following loss
  - $\mathcal{L}_{\frac{\mathrm{d}x}{\mathrm{d}t}} = \left|\left|\nabla_x \varphi(x)\dot{x} - \Theta\left(\varphi(x)^T\right)\Xi\right|\right|_2^2$
- Ensure SINDy predictions can be used to reconstruct time derivatives of x
  - $\mathcal{L}_{\frac{\mathrm{d}x}{\mathrm{d}t}} = \left|\left|\dot{x} - (\nabla_z \psi(\varphi(x)))\Theta\left(\varphi(x)^T\right)\Xi\right|\right|_2^2$