



From Trajectory Reconstruction to Dynamic Gene Regulatory Models in Single Cells



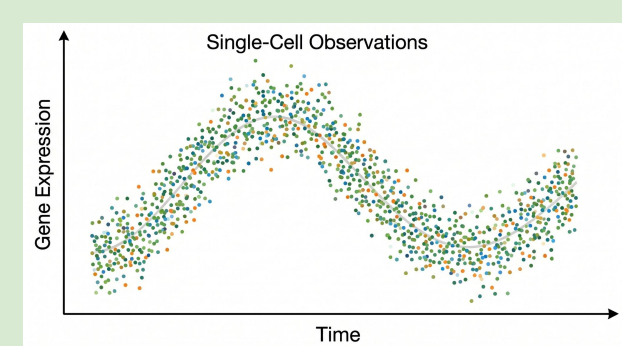
Bhattacharya Lab @ MSU

Samia Islam^{1,3}, Zijun Cui¹, Sudin Bhattacharya^{2,3}

¹Department of Computer Science and Engineering, ²Departments of Biomedical Engineering and Pharmacology & Toxicology,

³Institute for Quantitative Health Science and Engineering,
Michigan State University

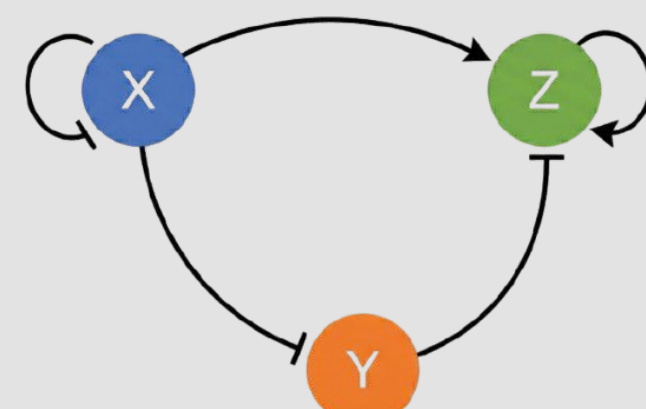
Why ODE discovery is hard in scRNA-seq



Derivatives are hard to estimate from noisy data

scRNA-seq data is noisy, which makes direct dynamic model discovery unstable.

A Toy System



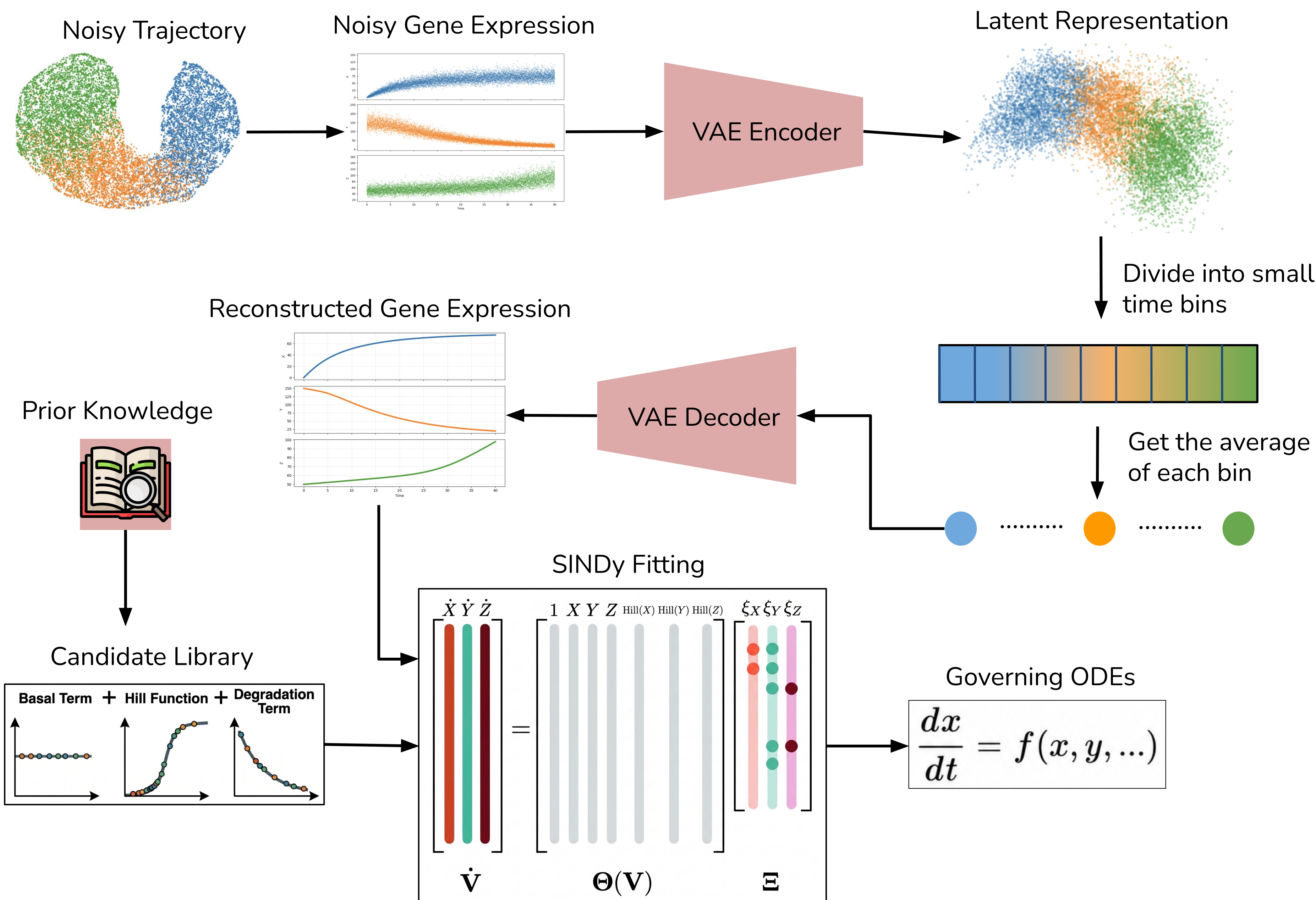
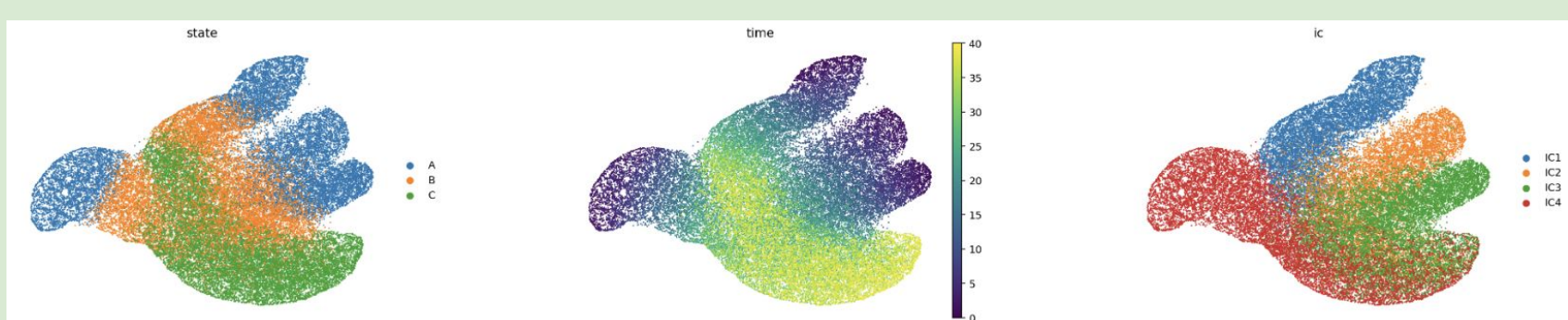
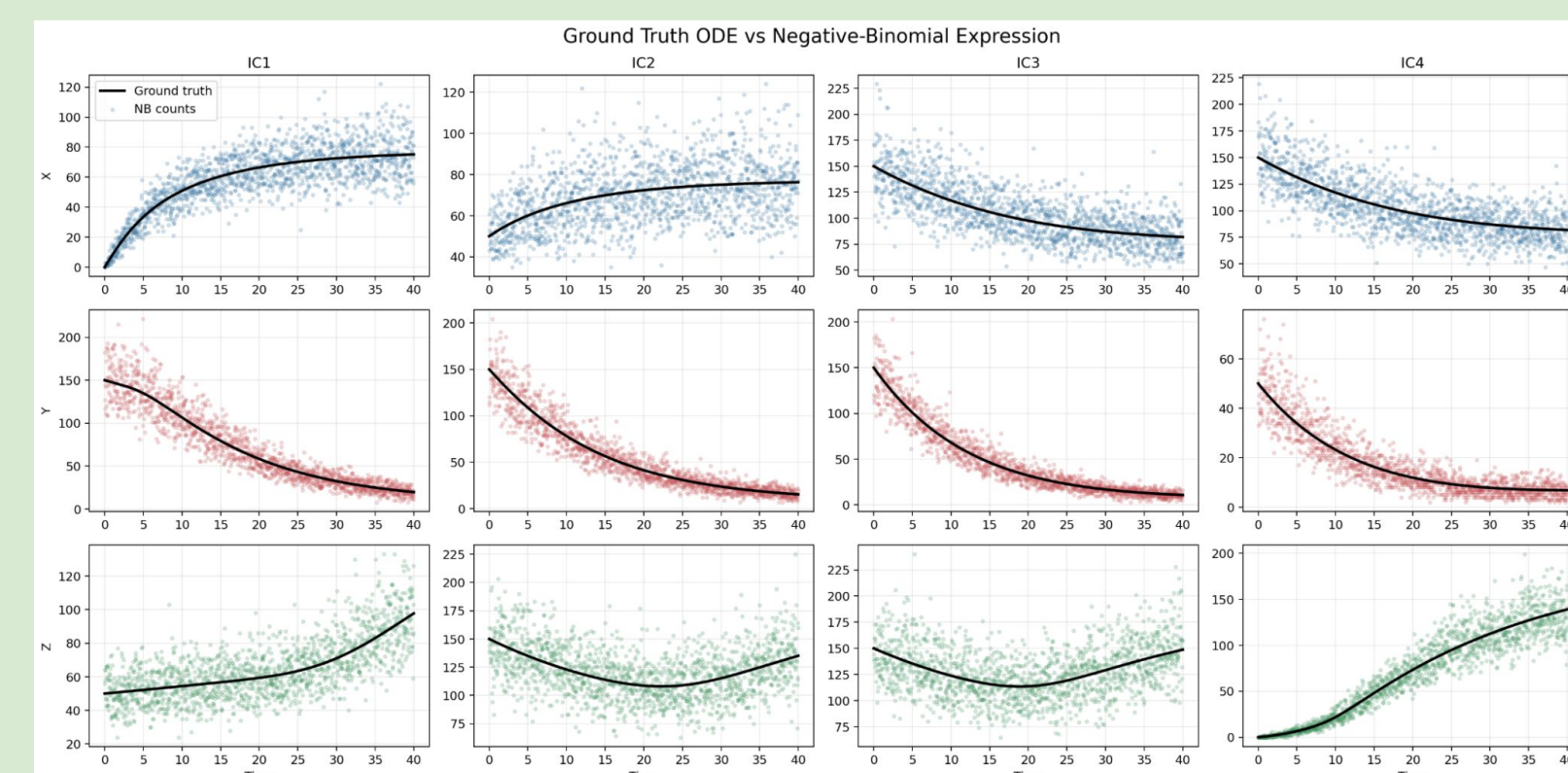
$$\frac{dX}{dt} = 0.3 + 8.14 \cdot \text{Rep}(X; 50, 2) - 0.035X$$

$$\frac{dY}{dt} = 0.02 + 9.26 \cdot \text{Rep}(X; 40, 4) - 0.08Y$$

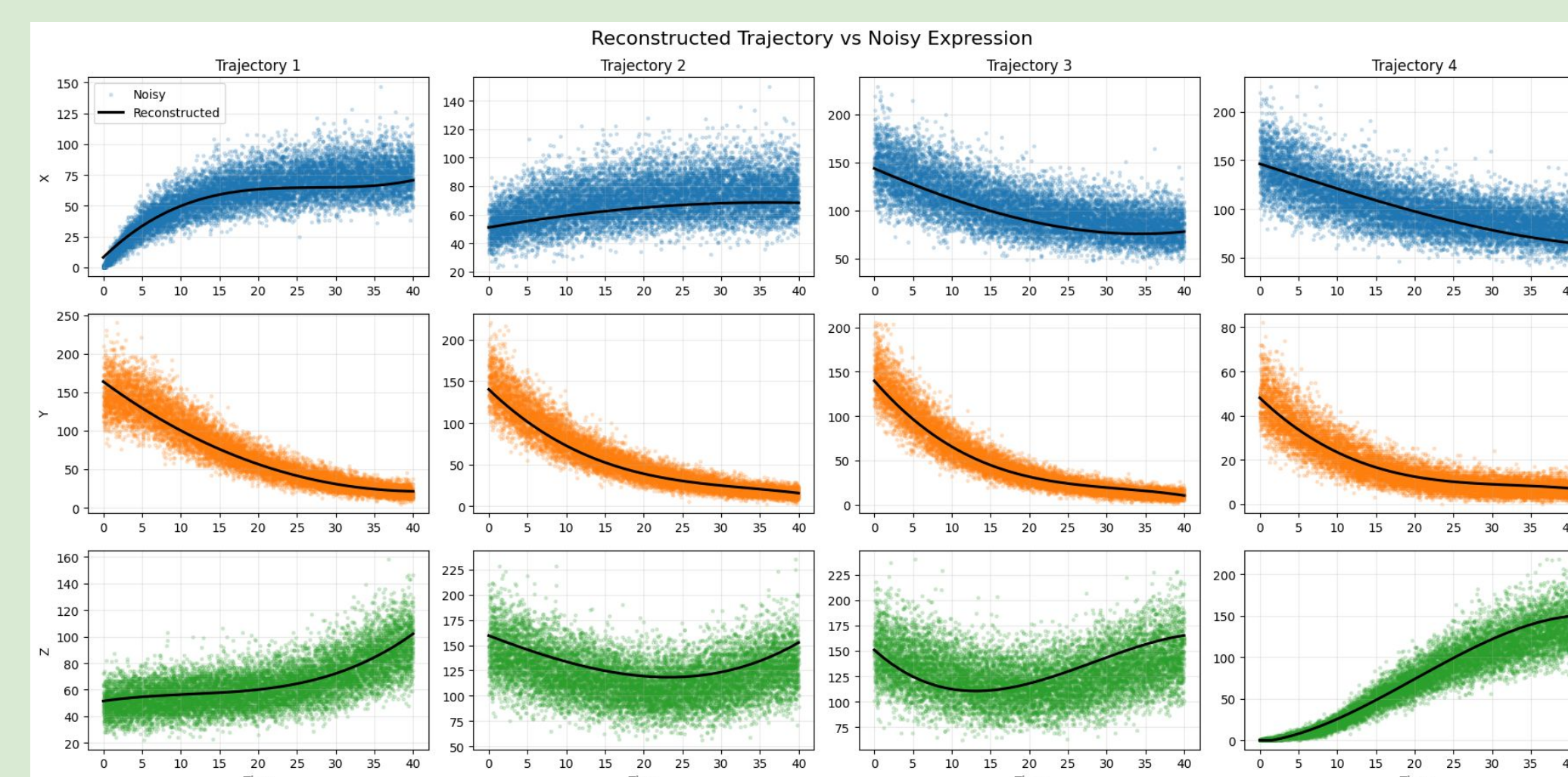
$$\frac{dZ}{dt} = 0.0028 + 0.28 \cdot \text{Act}(X; 50, 3) + 5 \cdot \text{Rep}(Y; 30, 4) + 2.5 \cdot \text{Act}(Z; 20, 4) - 0.04Z$$

Data Generation from the Toy System

Simulate four trajectories from the toy system
→ sample negative-binomial counts



Trajectory Reconstruction



Prior-informed SINDy

- ODE form:

$$\frac{dX_i}{dt} = \text{basal} + \sum_j \alpha_{ij} H_{ij}(X_j; K, n) - \beta_i X_i$$

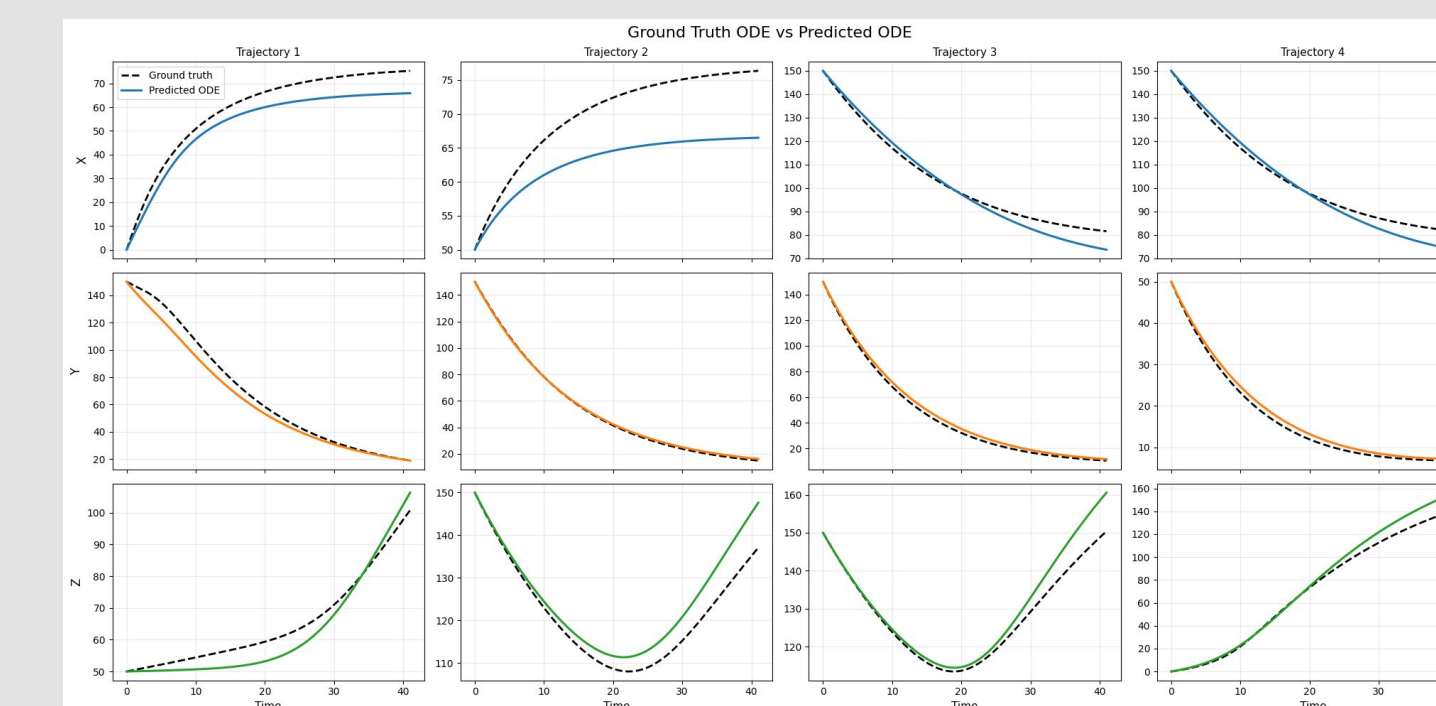
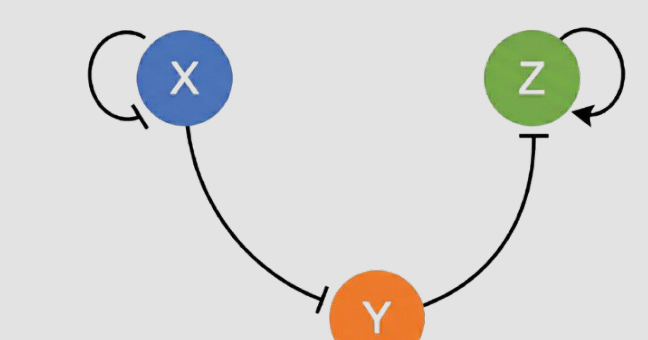
- Prior knowledge:** allowed Hill terms are defined by known activation/repression links.
- Grid search:** Hill parameters K, n are selected by minimizing error.

Predicted Dynamics

$$\frac{dX}{dt} = 0.404125 - 0.0267433X + 5.79453 \text{Rep}(X; 50, 4)$$

$$\frac{dY}{dt} = 0.124928 - 0.0754461Y + 4.87418 \text{Rep}(X; 40, 4)$$

$$\frac{dZ}{dt} = 0.346258 - 0.0501905Z + 4.41363 \text{Act}(Z; 50, 2) + 5.96386 \text{Rep}(Y; 30, 4)$$



Quantitative Results

