Parameter Estimation of Mechanistic Differential Equations via Neural Differential Equations

William Bradley, 4th Yr PhD Candidate (Boukouvala Lab)

School of Chemical & Biomolecular Engineering, Georgia Institute of Technology

Background

Motivation

Regressing mechanistic models to experimental data can be intractable due to process nonlinearity and computation issues. Machine learning (ML) tools could reduce the cost of modeling building but offer limited interpretability/robustness. Using ML as a data-driven means to a mechanistic end, **an indirect approach using NODEs is proposed that accelerates parameter estimation of mechanistic models.**

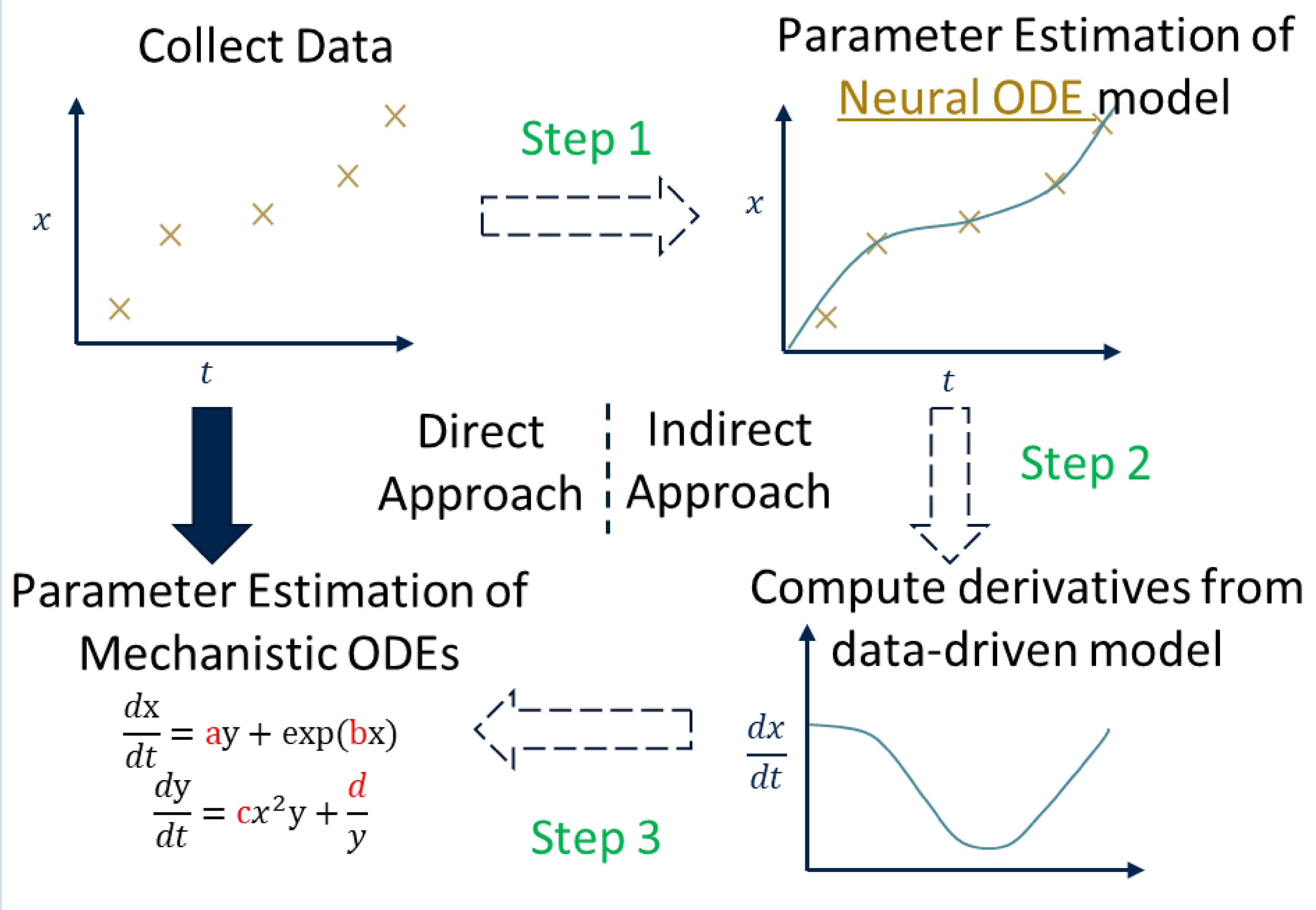
Objectives

- Compare ability of NODEs and NNs to estimate derivatives
- Compare performance of direct approach vs NODE-based indirect approach for fitting parameters

Thesis

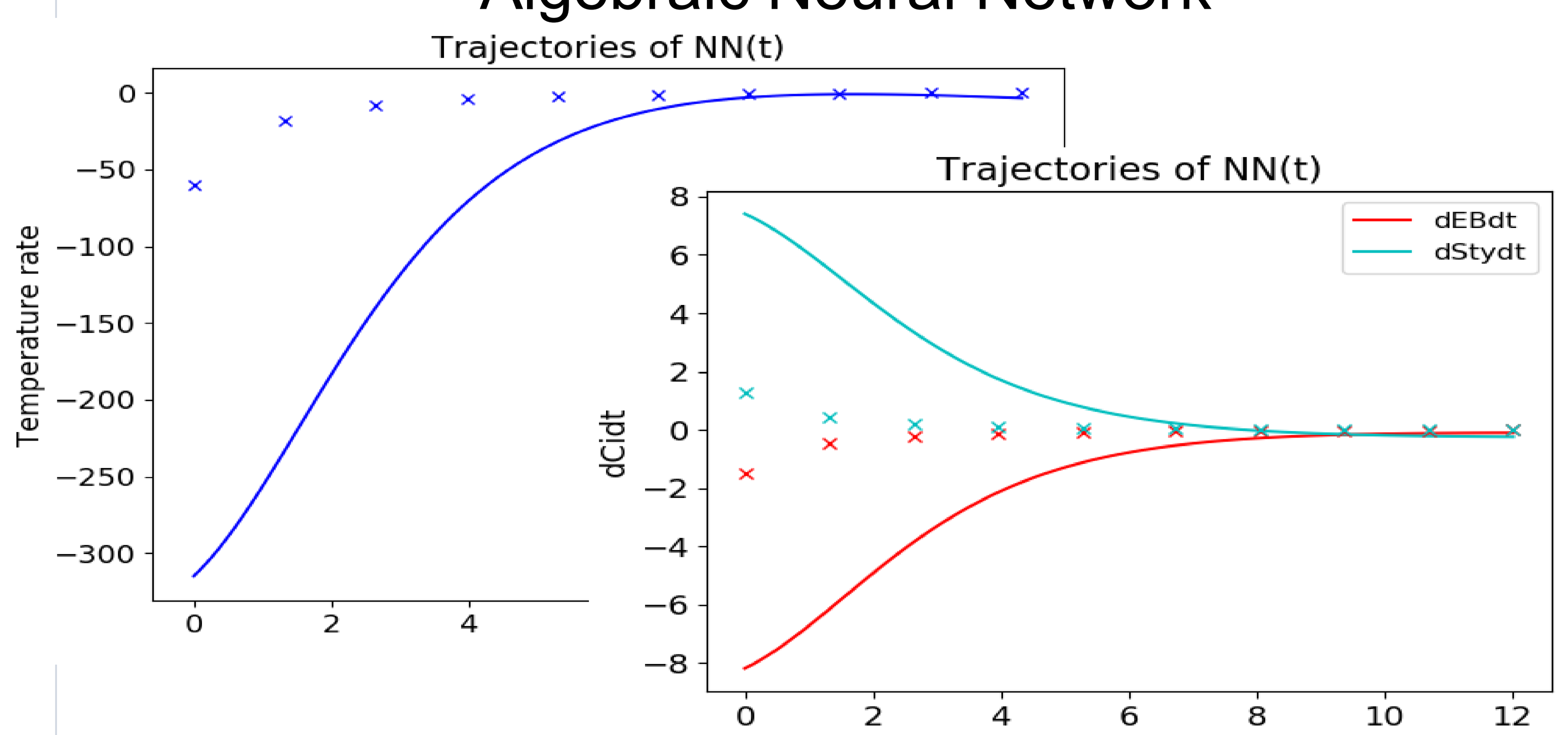
Data-driven models can be used to estimate parameters of mechanistic differential equations, **accelerating creation of interpretable, generalizable** (bio)chemical process models.

Methods

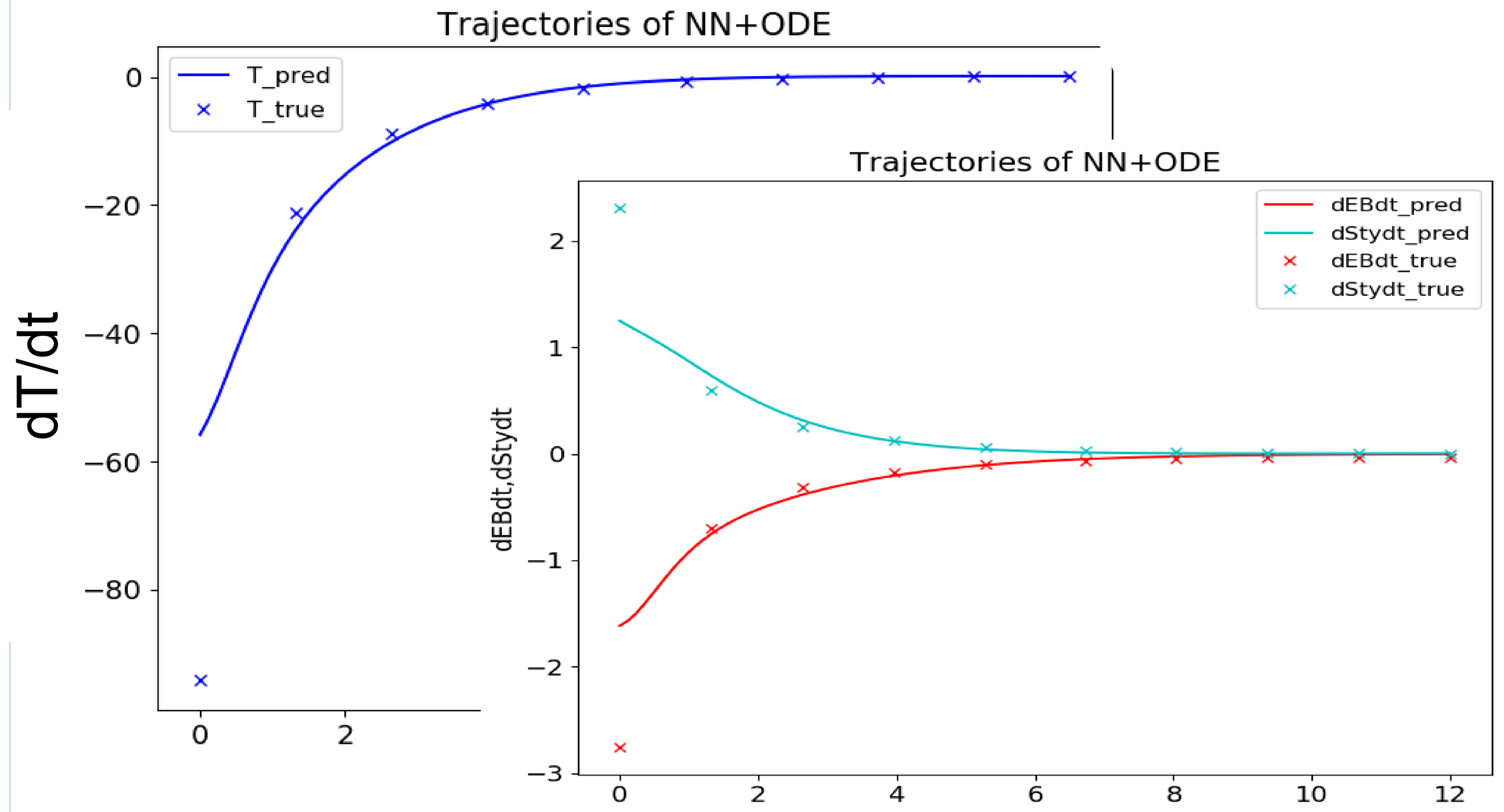


Derivative Estimation: NN vs NODE

Algebraic Neural Network

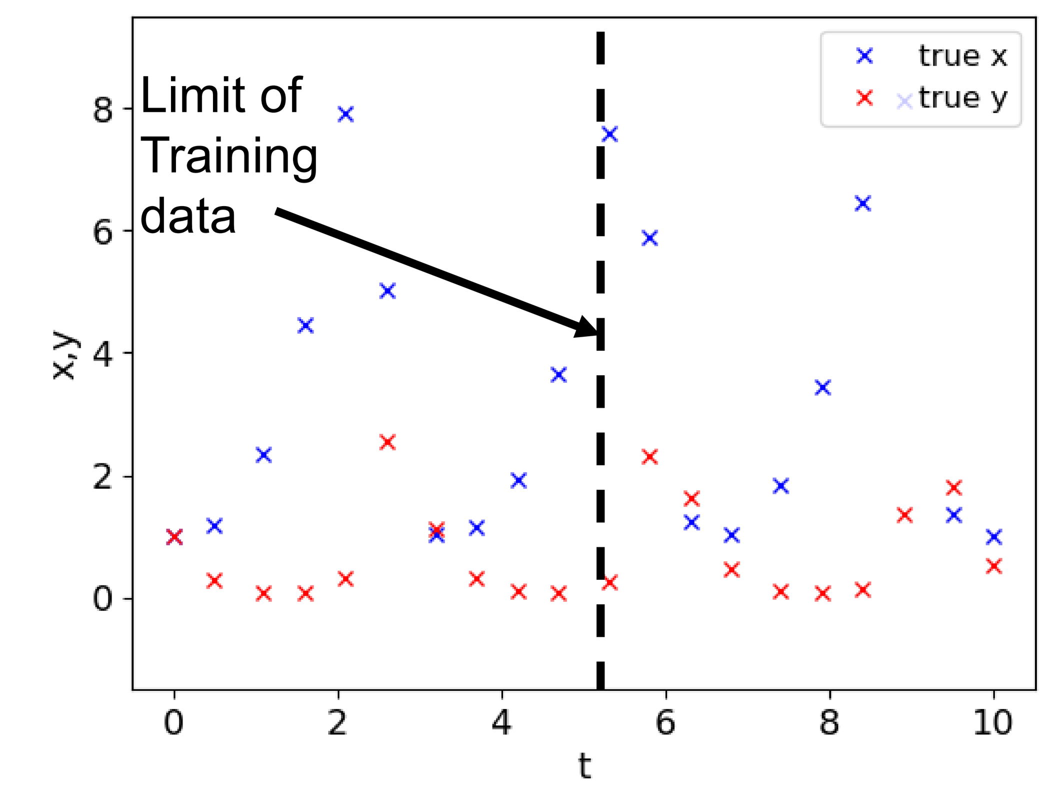


Neural ODE

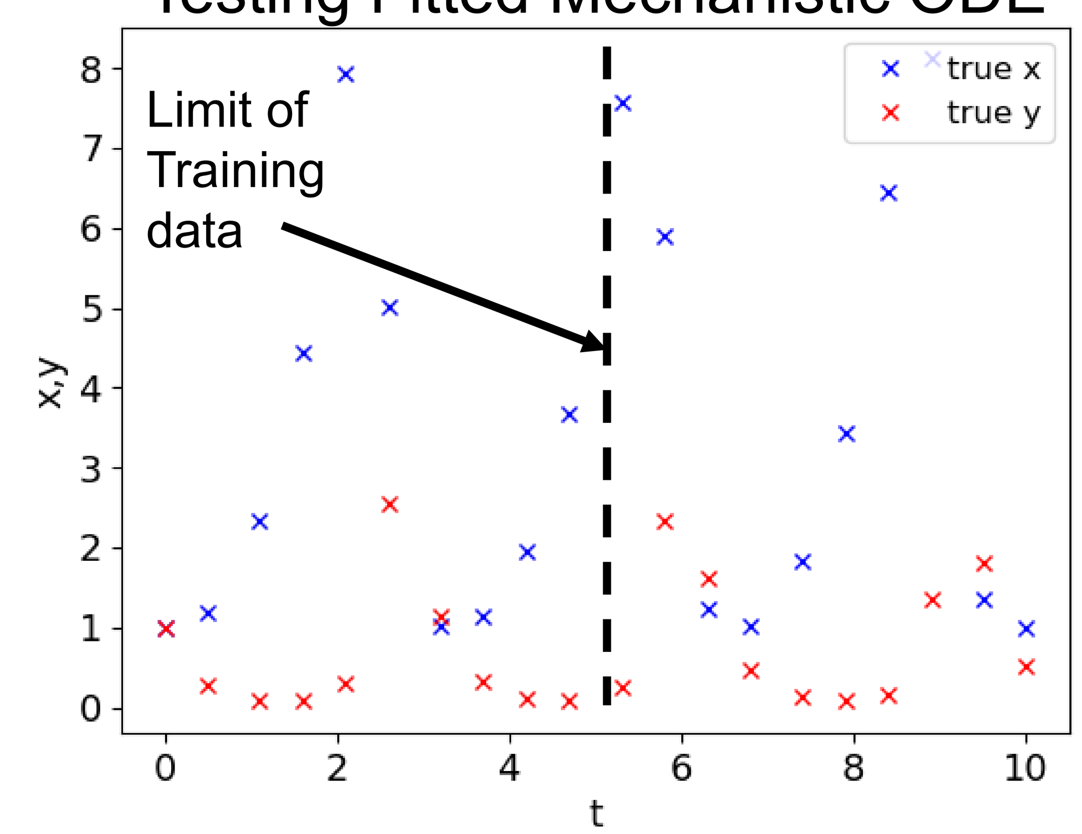


Extrapolation: NODE vs MODE

Testing Data-driven Model



Testing Fitted Mechanistic ODE



Indirect vs Direct Parameter Estimation

	Lotka Volterra (2 states, 3 parameters, 20 dp)	Styrene Reactor (6 states, 3 parameters, 60 dp)	Penicillin fermenter (3 states, 11 parameters, 90 dp)
Direct Approach	Total: 76 s	Total: 352 s	Did not converge
Indirect NODE Approach	Total: 62 s Steps 1&2: 62. s Step 3: 0.009 s	Total: 116 s Steps 1&2: 110 s Step 3: 6.76 s	Total: 183 s Steps 1&2: 181 s Step 3: 1.932 s

dp = data points
NODE = Neural ODE

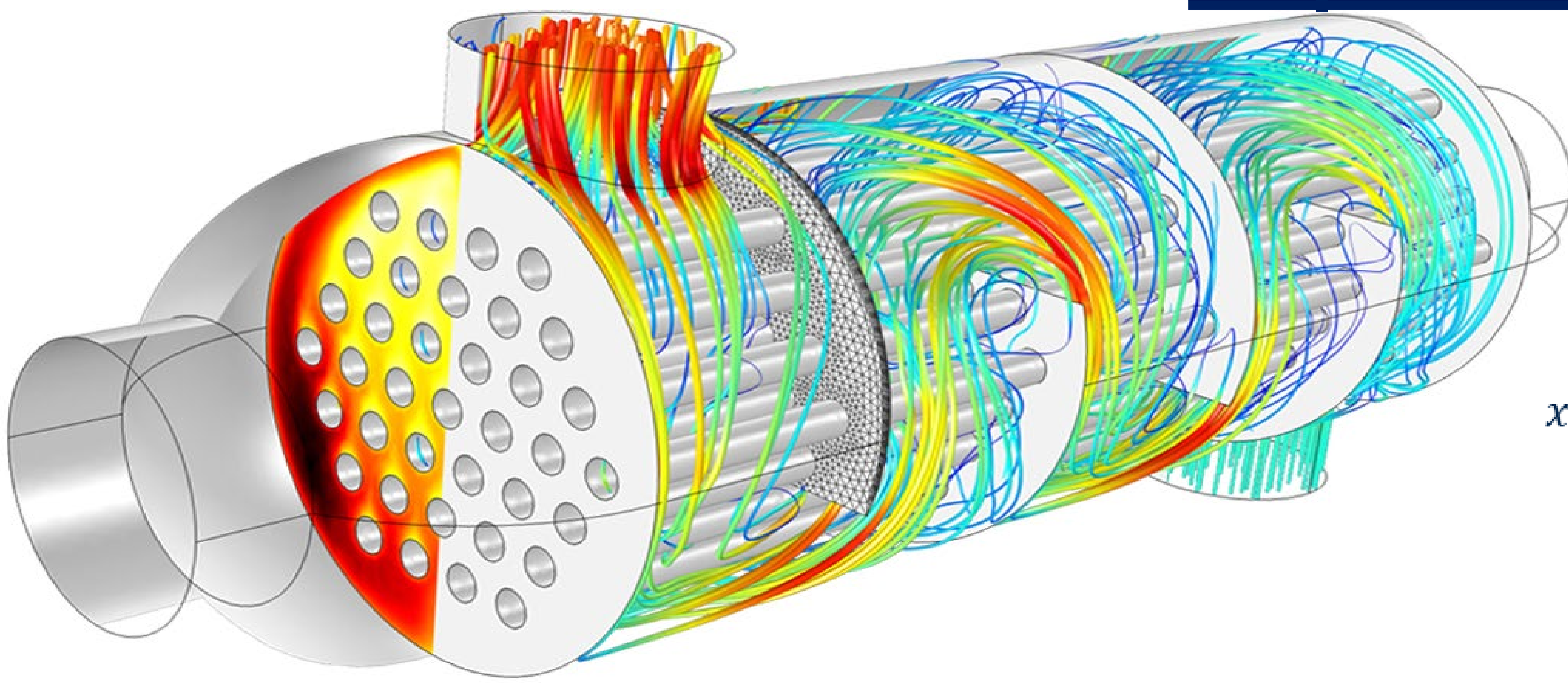
Conclusions

- NODEs estimate state derivatives more accurately than algebraic data-driven models
- Mechanistic ODEs have superior generalizability than Neural ODEs
- NODE Indirect approach can regress parameters of mechanistic models faster than direct approaches

Acknowledgements

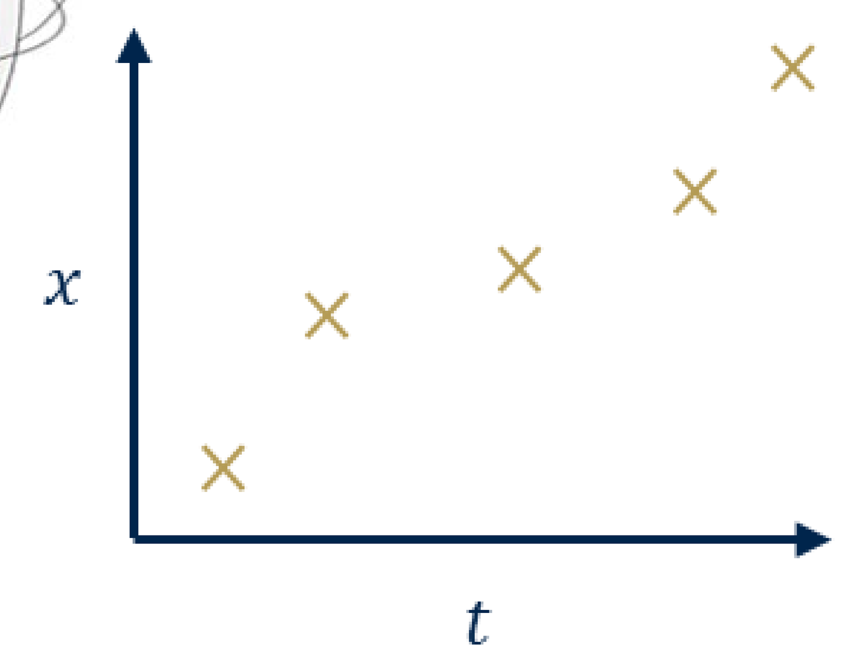
Georgia Tech Startup Grant
RAPID/NNMI Grant #GR10002225

Motivation: Parameter estimation of differential equation models



https://www.machinedesign.com/learning-resources/webinars/webinar/21134601/simulation-of-heat-exchangers

Collect Data



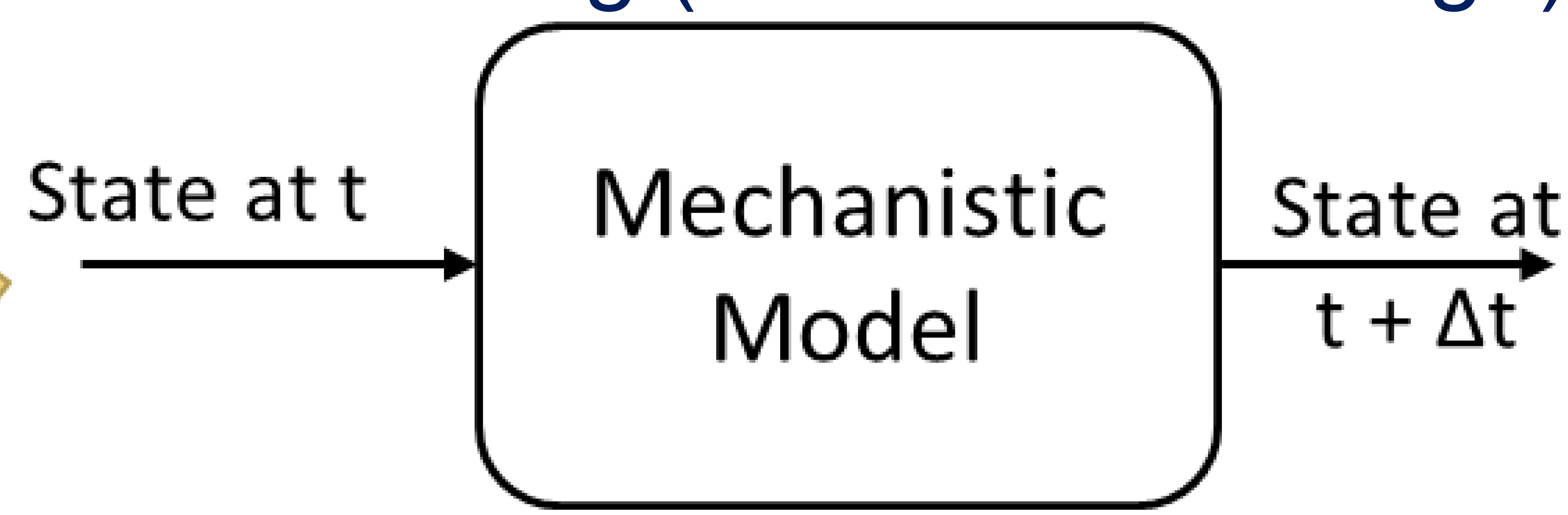
Direct Approach

Parameter Estimation of Mechanistic ODEs

$$\frac{dx}{dt} = ay + \exp(bx)$$

$$\frac{dy}{dt} = cx^2y + \frac{d}{y}$$

Mechanistic Modeling (Domain Knowledge)



Machine Learning (Data Knowledge)

Mechanistic (M) model

Advantages: High interpretability, good extrapolation

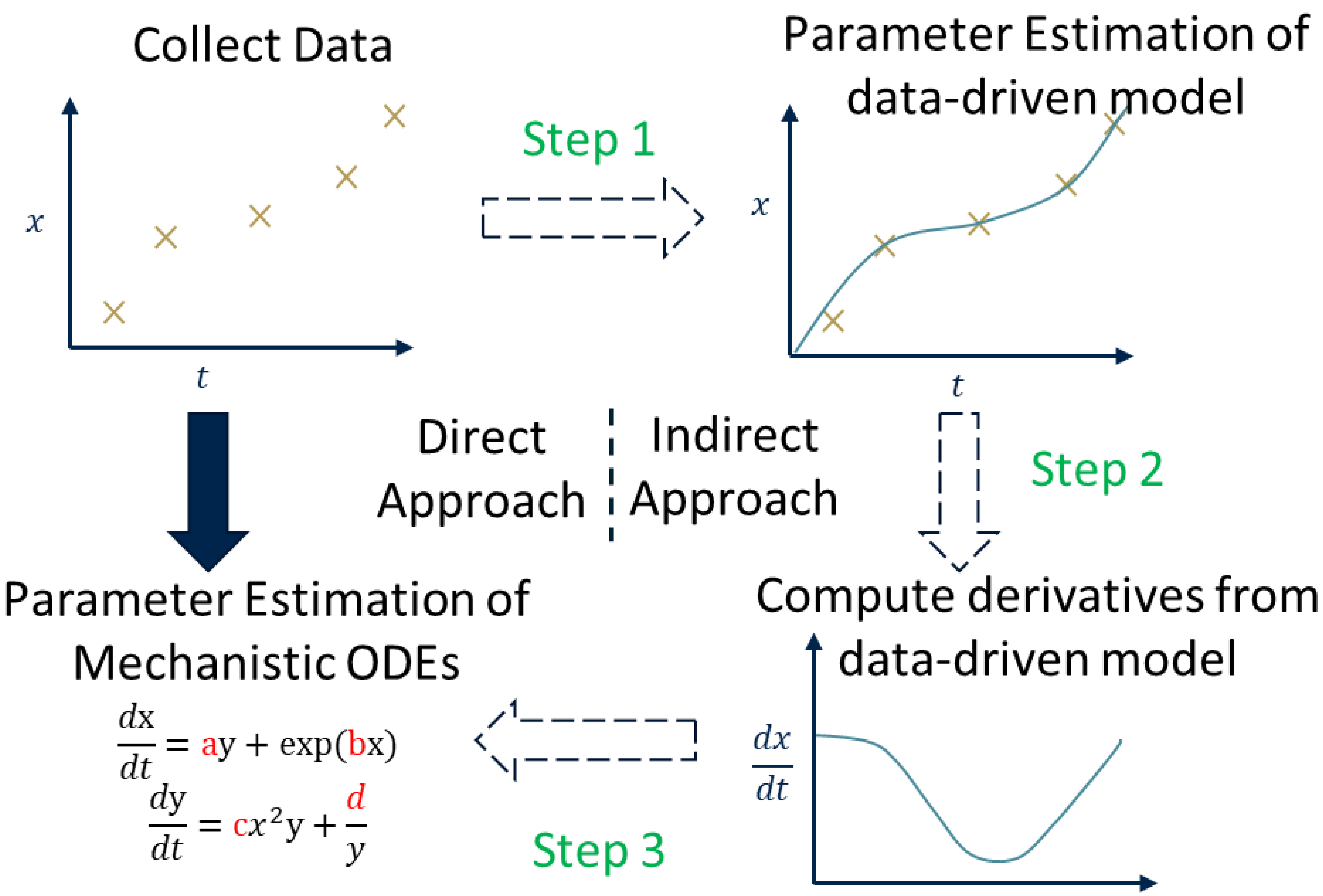
Disadvantages: Computation intensive

Data-driven (DD) model

Advantages: Fast model-building

Disadvantages: Low interpretability, poor extrapolation

Proposed Method: Indirect Approach



Hypotheses:

1. Neural ODEs can predict derivatives better than Algebraic NN models
2. A Neural ODE-based indirect approach can fit mechanistic ODEs faster and more accurately than a direct approach

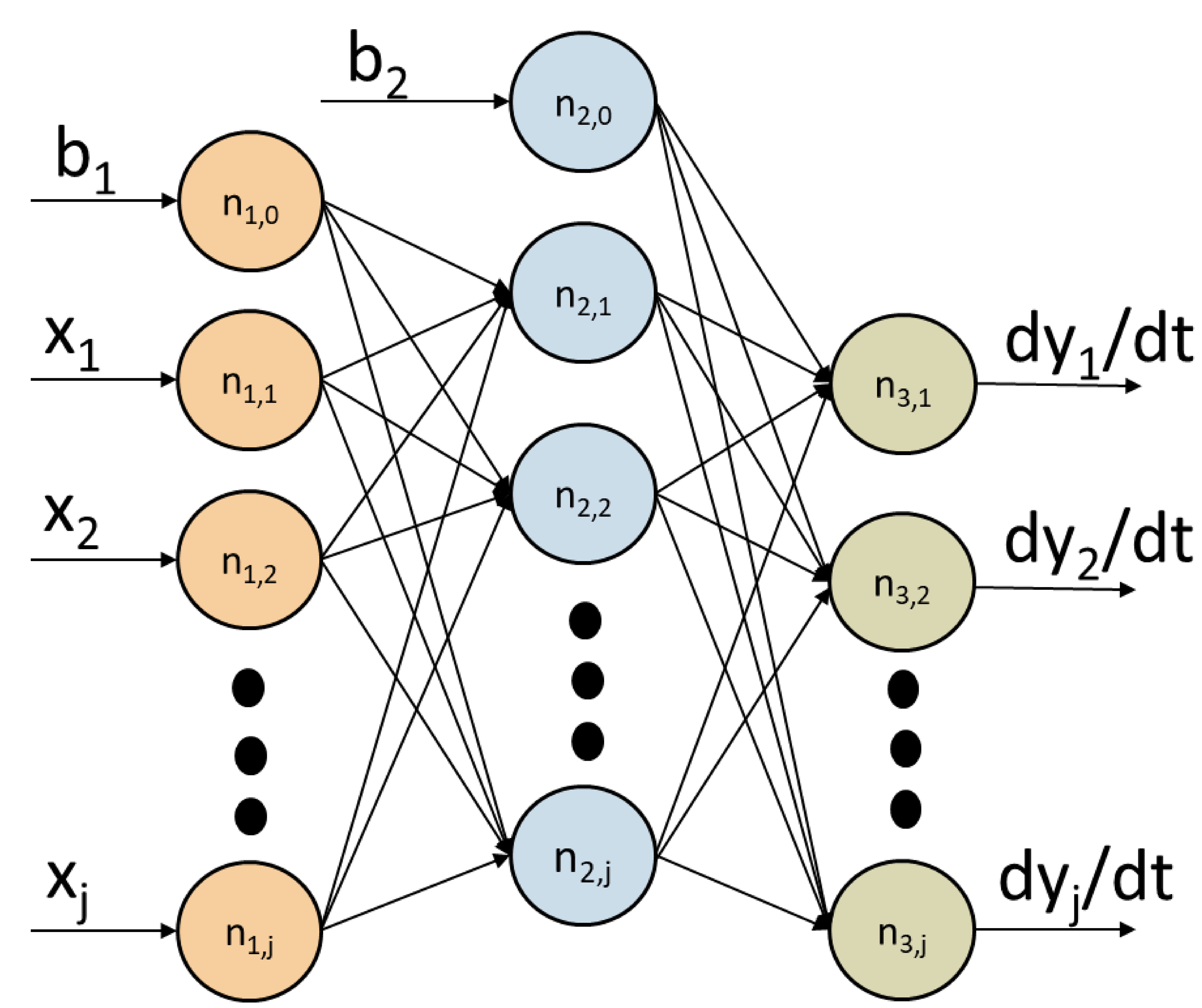
Neural Ordinary Differential Equations (Neural ODEs)

- **Neural Networks** can approximate any nonlinear, continuous algebraic equation

$$y = NN\left(\sum_{k=1}^{s_{i-1}} w_{k,j}^{(i)} x_{(i-1),k} + b_{0,j}^{(i)}\right)$$

- **Neural ODEs** can approximate nonlinear continuous differential equations

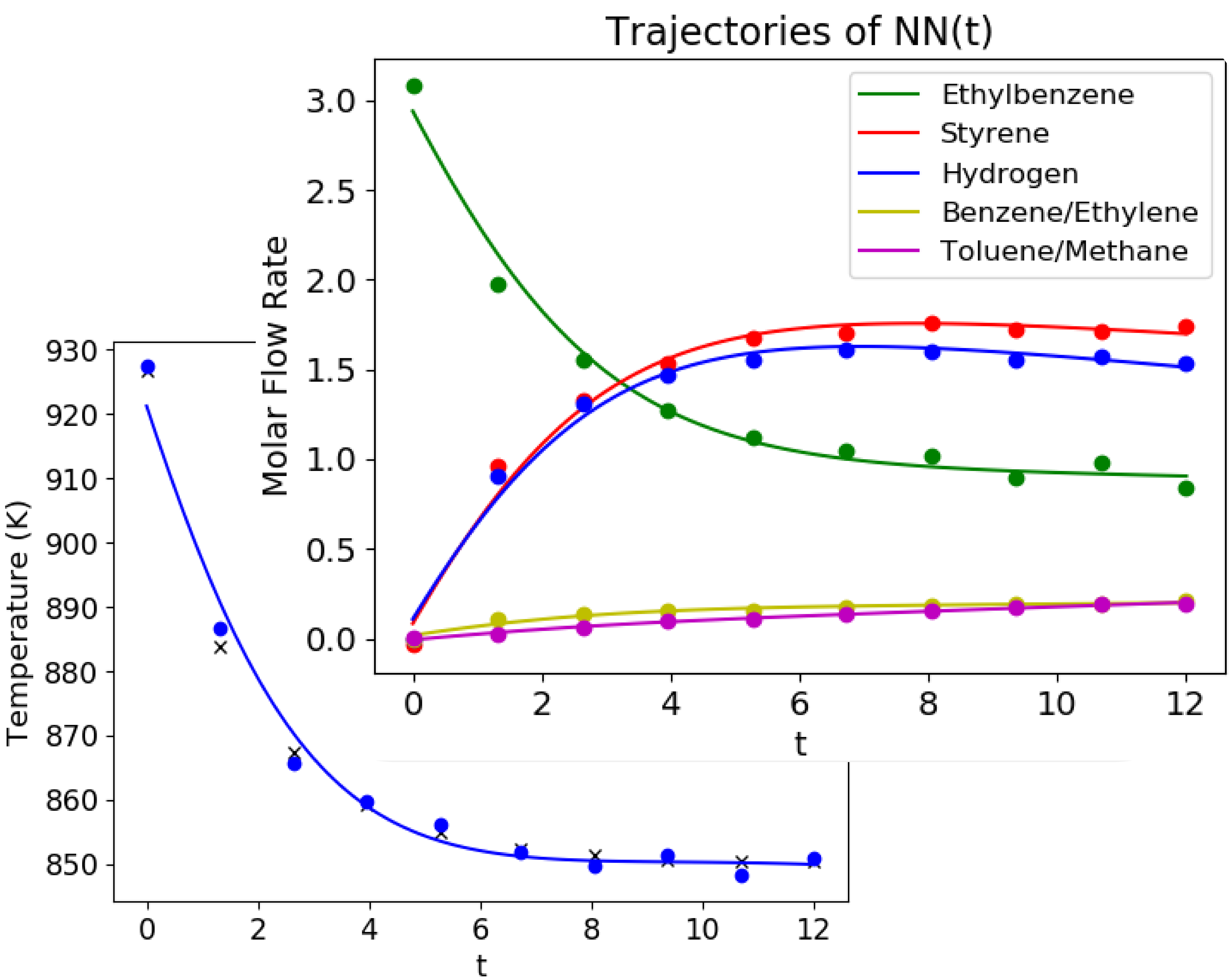
$$\frac{dy}{dt} = NN\left(\sum_{k=1}^{s_{i-1}} w_{k,j}^{(i)} x_{(i-1),k} + b_{0,j}^{(i)}\right)$$

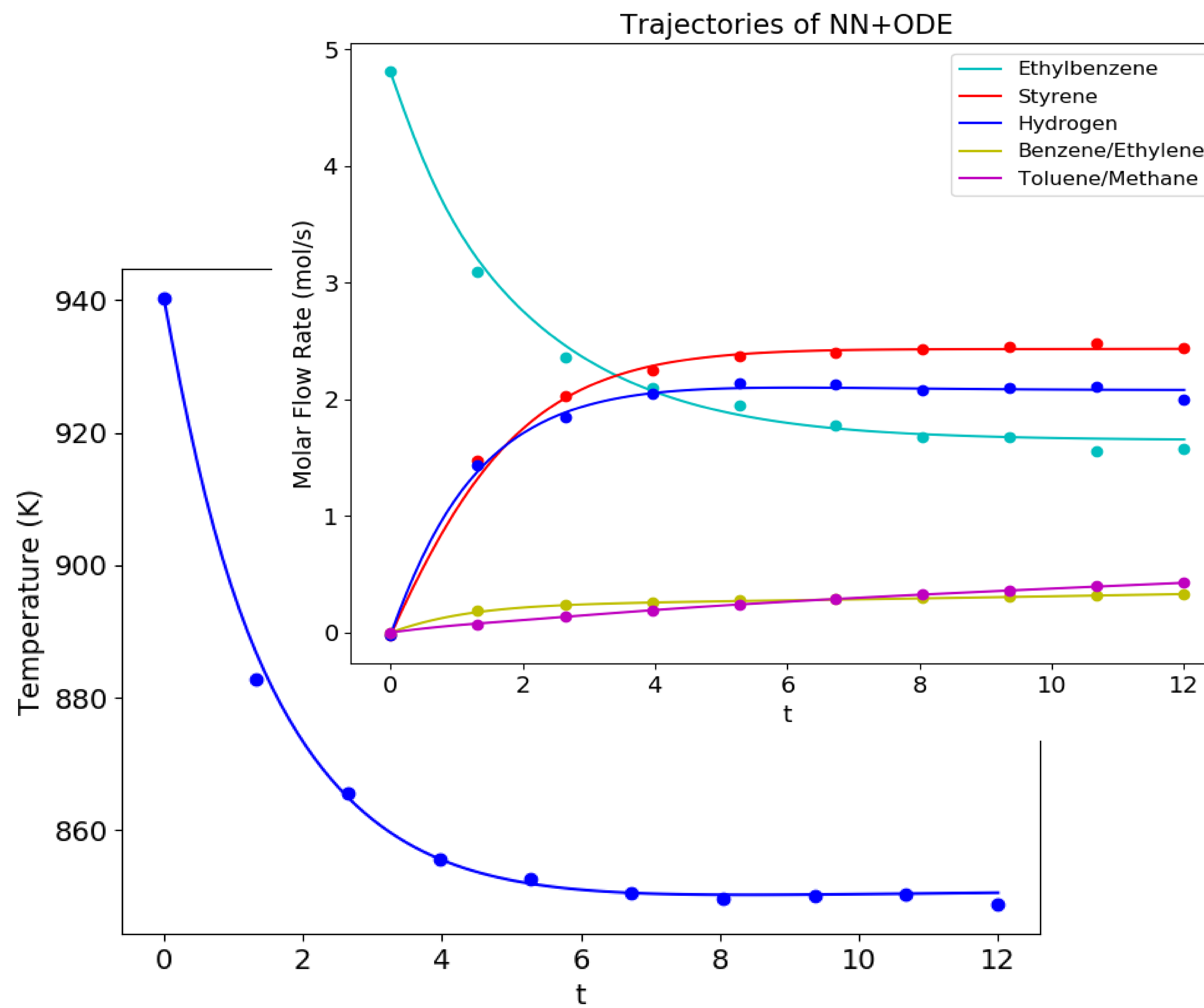


Results: NN vs Neural ODE state estimates of styrene reactor

Neural Network[1]: $y = NN(t, w)$

Neural ODE[2]: $\frac{dy}{dt} = NN(x(t), w)$



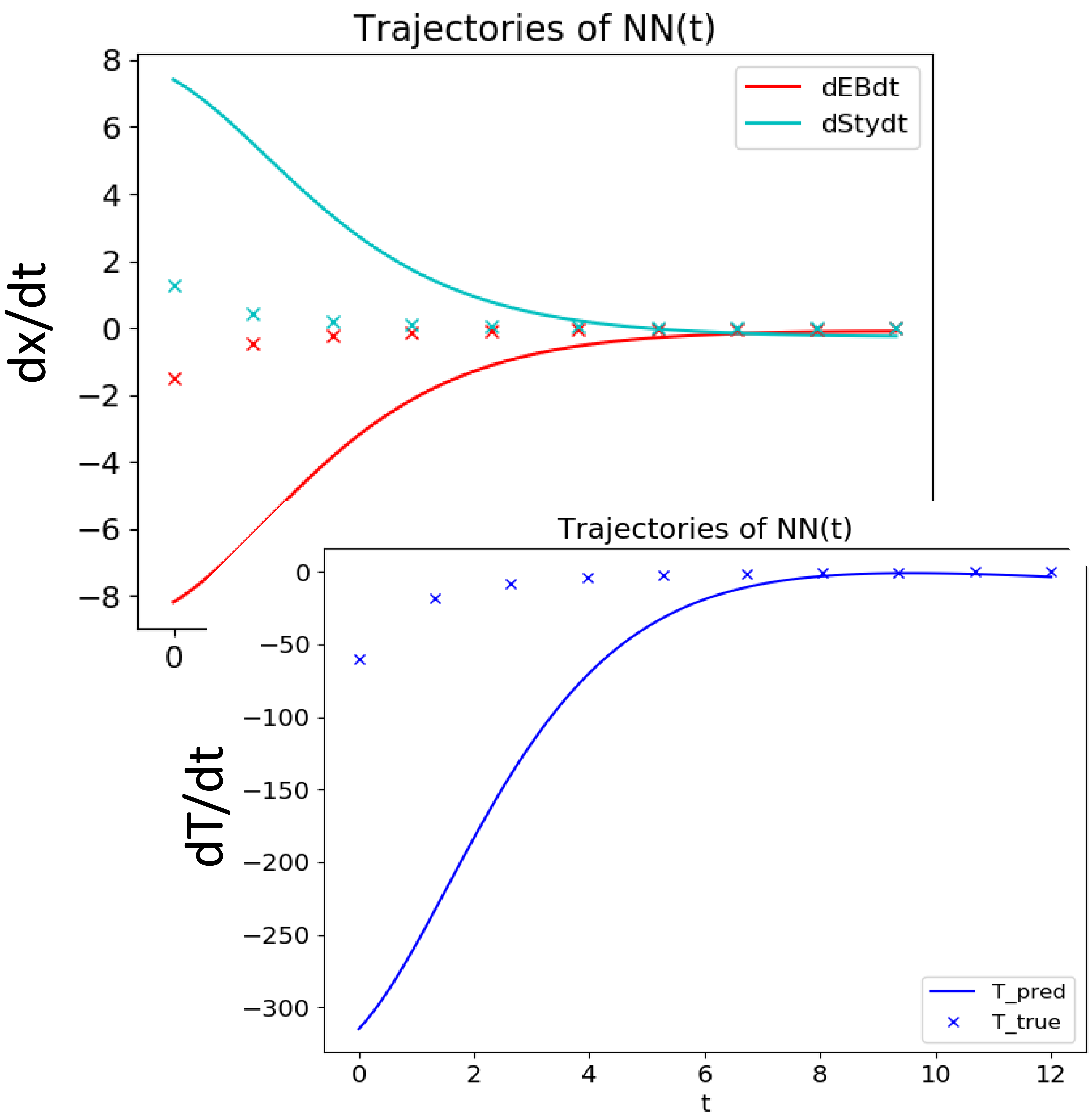


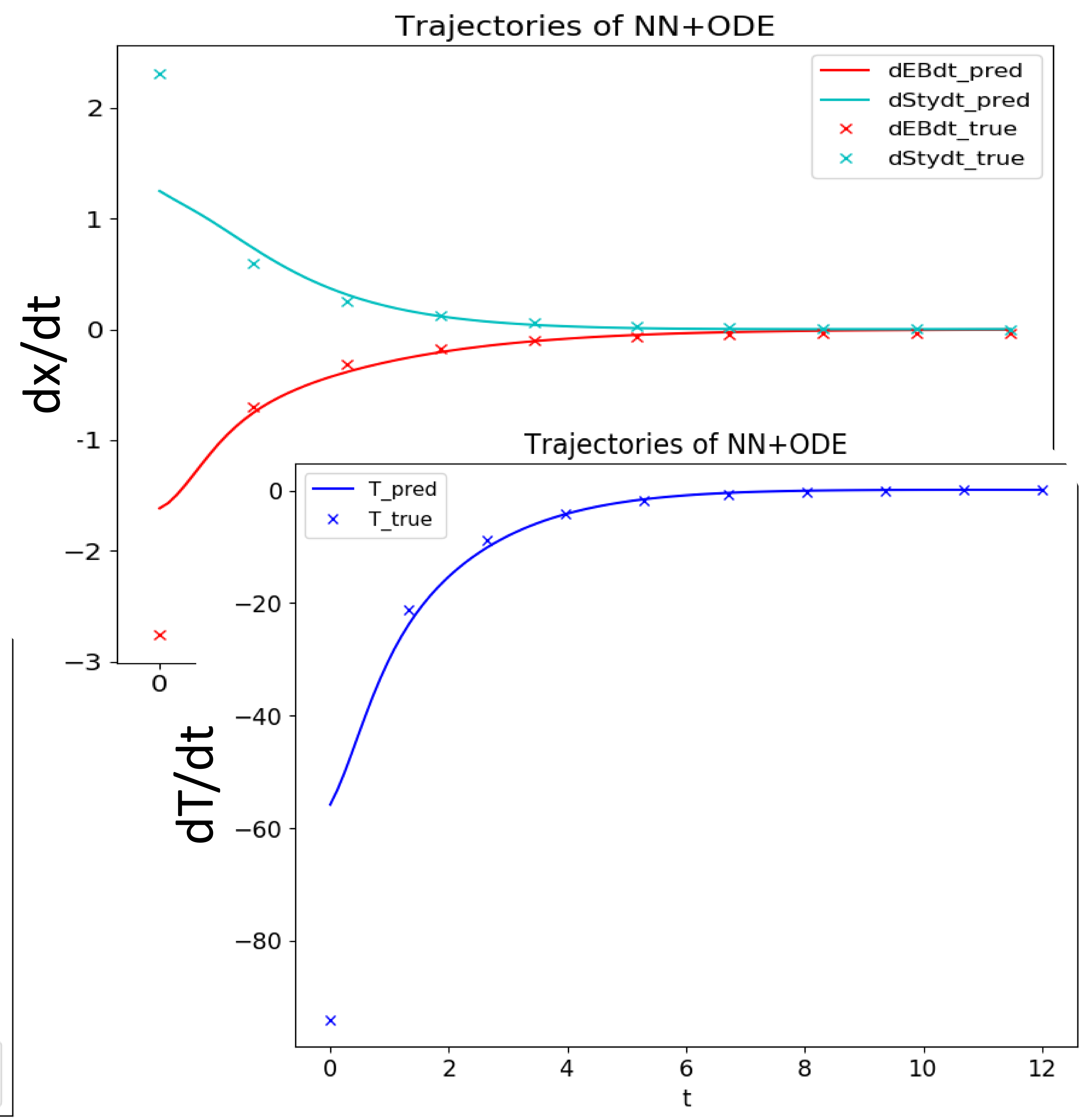
1) Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems, 2*(4), 303-314. doi:10.1007/bf02551274
2) Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. (2018). Neural Ordinary Differential Equations. *arXiv e-prints*. Retrieved from https://ui.adsabs.harvard.edu/abs/2018arXiv180607366C

Results: NN vs Neural ODE derivative estimates of styrene reactor

Neural Network: $y = NN(t, w)$

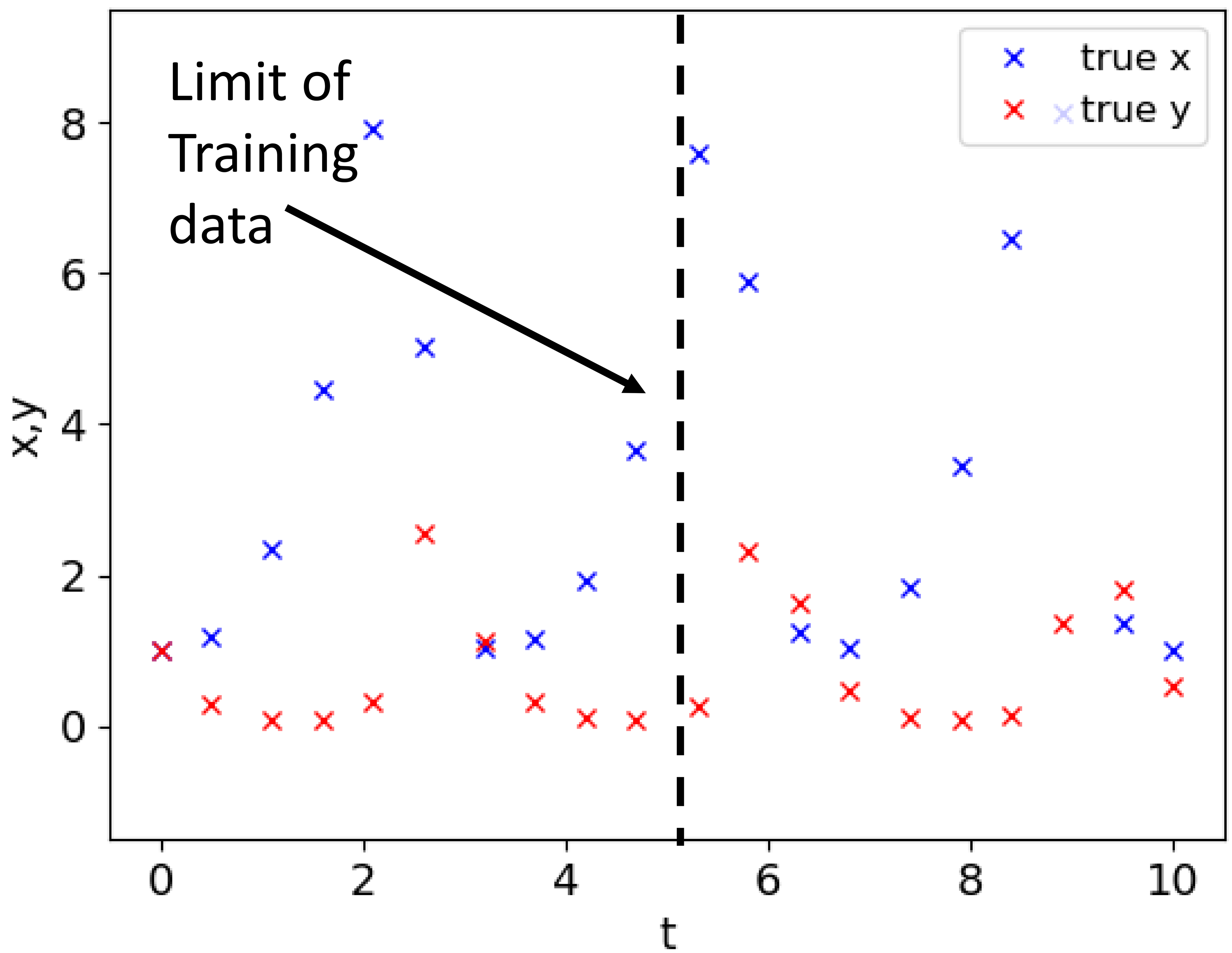
Neural ODE: $\frac{dy}{dt} = NN(x(t), w)$



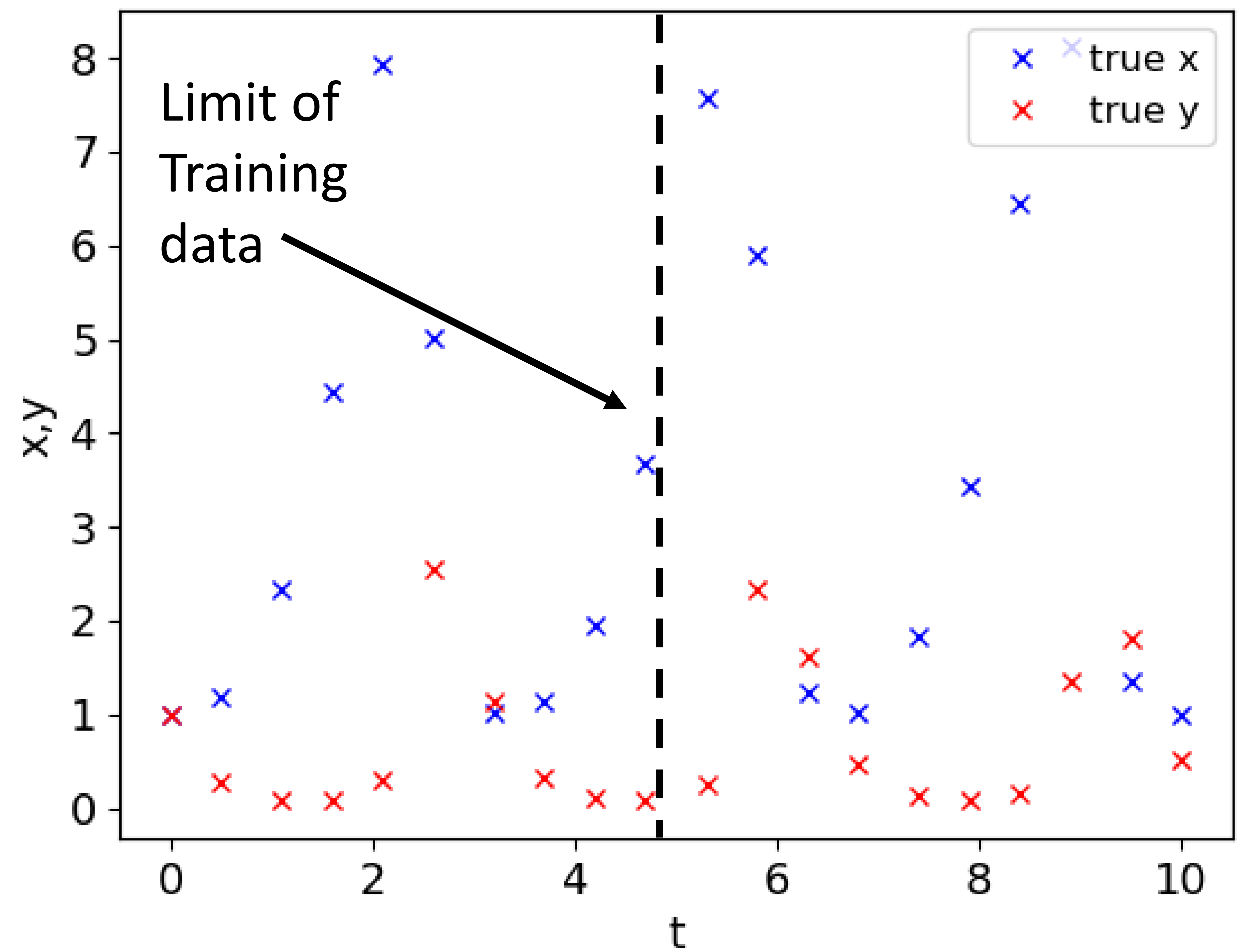


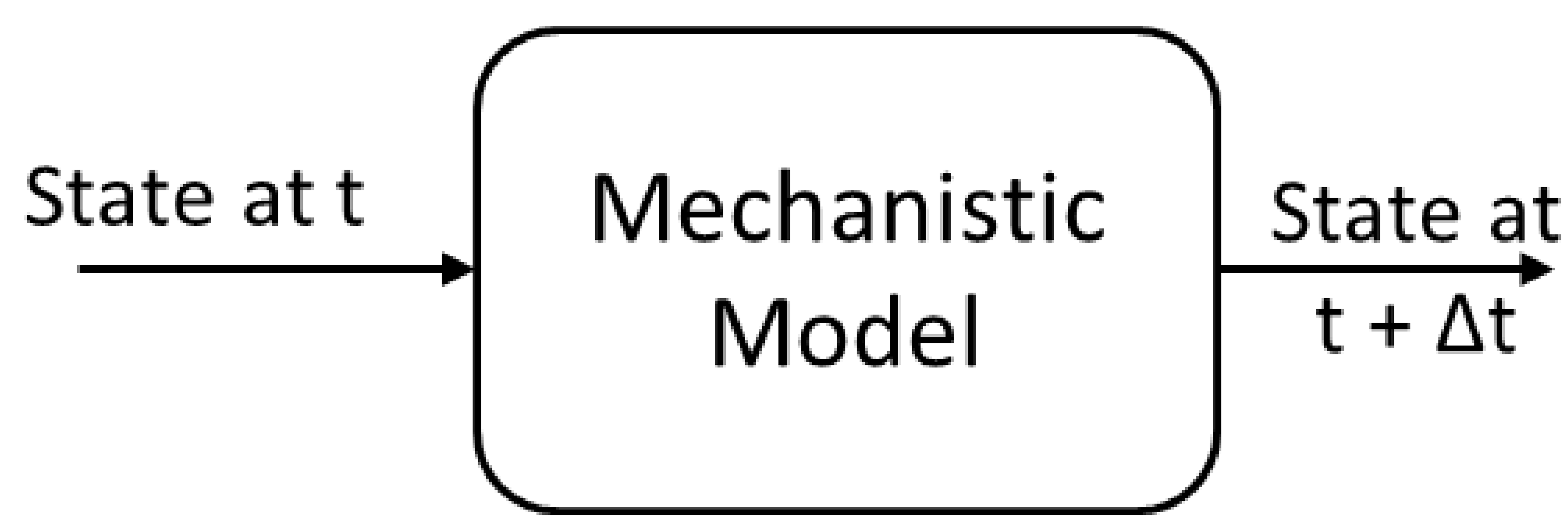
Results: Extrapolation of NODE vs Mechanistic ODE

Testing NODE Model



Testing Fitted Mechanistic Model (parameters fitted w/NODE Approach)





Conclusions

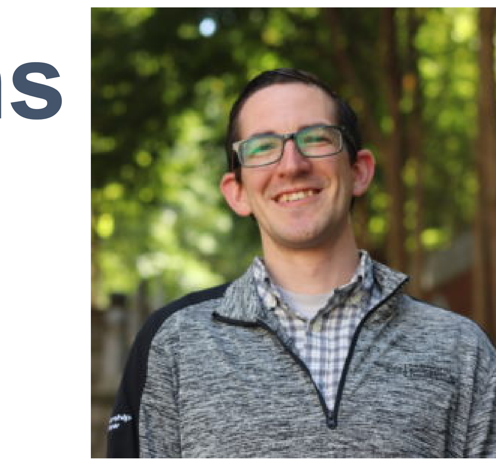
	Predator-Prey System (2 states, 3 parameters, 20 dp)	Styrene Reactor (6 states, 3 parameters, 60 dp)	Penicillin Fermenter (3 states, 11 parameters, 90 dp)
Direct Approach	Total: 76 s	Total: 352 s	Did not converge
Indirect NODE Approach	Total: 62 s Steps 1&2: 62. s Step 3: 0.009 s	Total: 116 s Steps 1&2: 110 s Step 3: 6.76 s	Total: 183 s Steps 1&2: 181 s Step 3: 1.932 s

- Neural ODEs estimate state derivatives more accurately than algebraic data-driven models (e.g. Neural Networks)
- Mechanistic ODEs have superior generalizability than Neural ODEs
- NODE-based indirect approach can estimate parameters of mechanistic models faster than direct approaches

Email: wbradley@gatech.edu
Website: boukouvala.chbe.gatech.edu

Parameter Estimation of Mechanistic Differential Equations via Neural Differential Equations

William Bradley, 4th Yr PhD Candidate (Boukouvala Lab)
School of Chemical & Biomolecular Engineering, Georgia Institute of Technology



Background

Motivation

Regressing mechanistic models to experimental data can be intractable due to process nonlinearity and computation issues. Machine learning (ML) tools could reduce the cost of modeling building but offer limited interpretability/robustness. Using ML as a data-driven means to a mechanistic end, **an indirect approach using NODEs is proposed that accelerates parameter estimation of mechanistic models.**

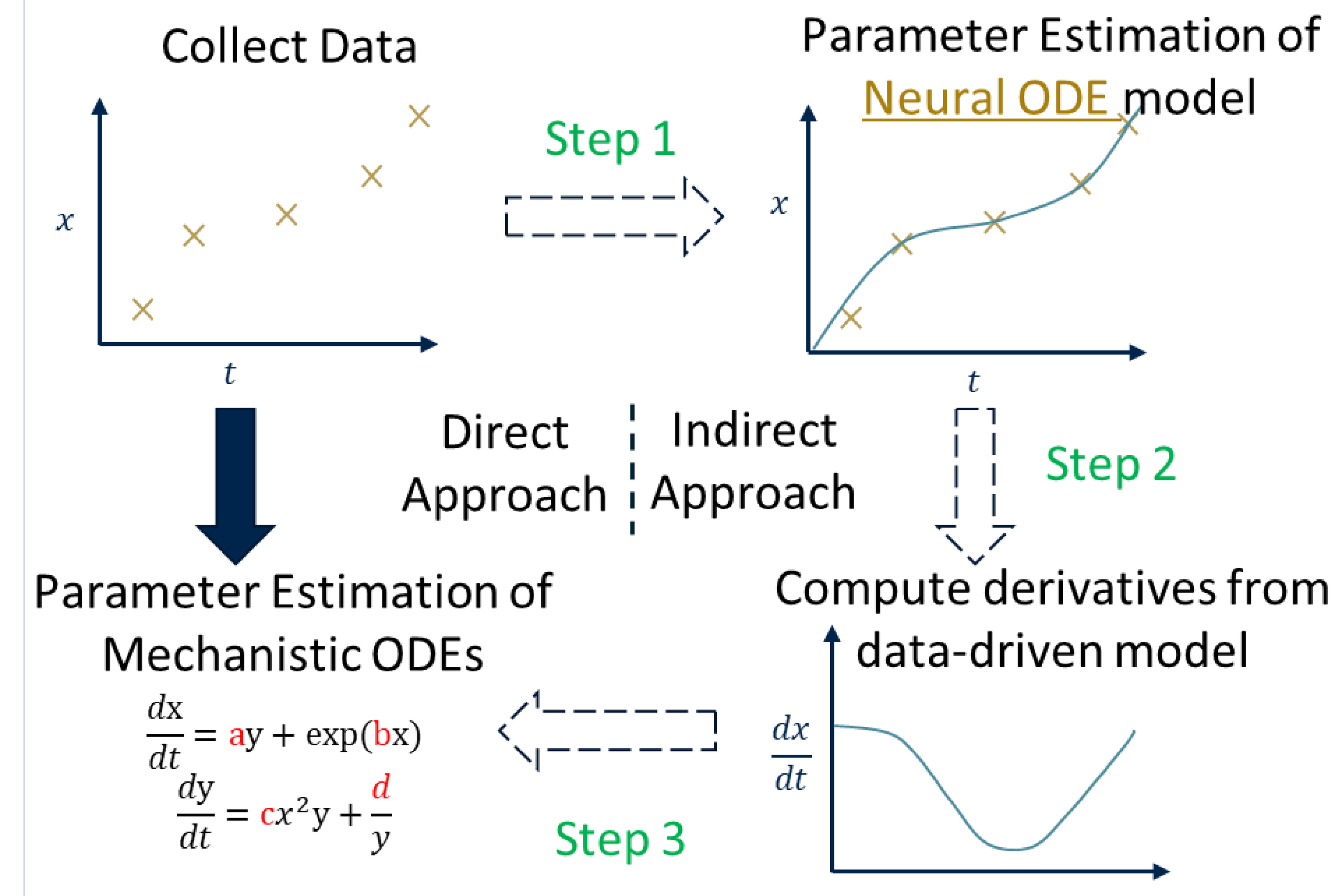
Objectives

- Compare ability of NODEs and a-NNs to estimate derivatives
- Compare performance of direct approach vs NODE-based indirect approach for fitting parameters

Thesis

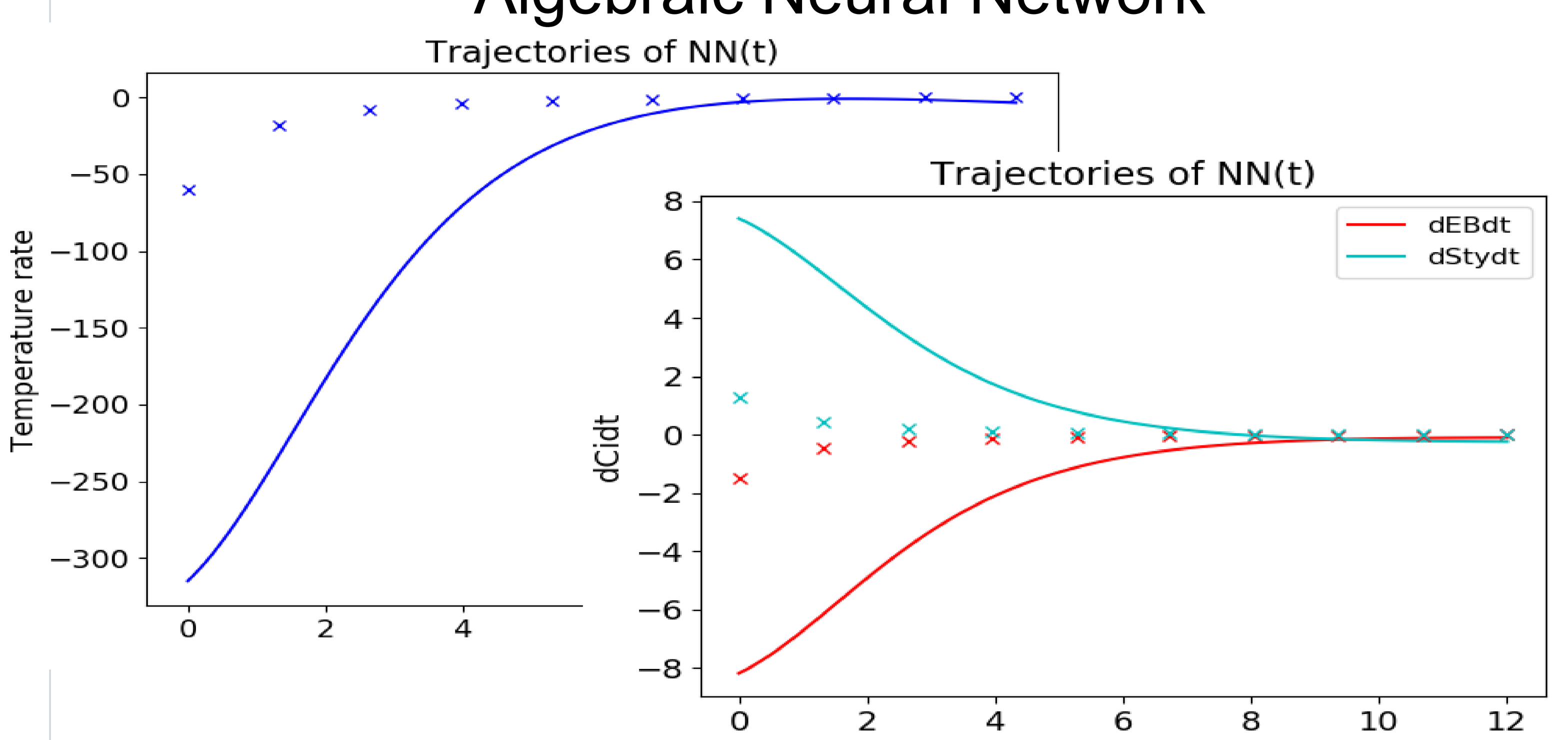
Data-driven models can be used to estimate parameters of mechanistic differential equations, **accelerating creation of interpretable, generalizable** (bio)chemical process models.

Methods

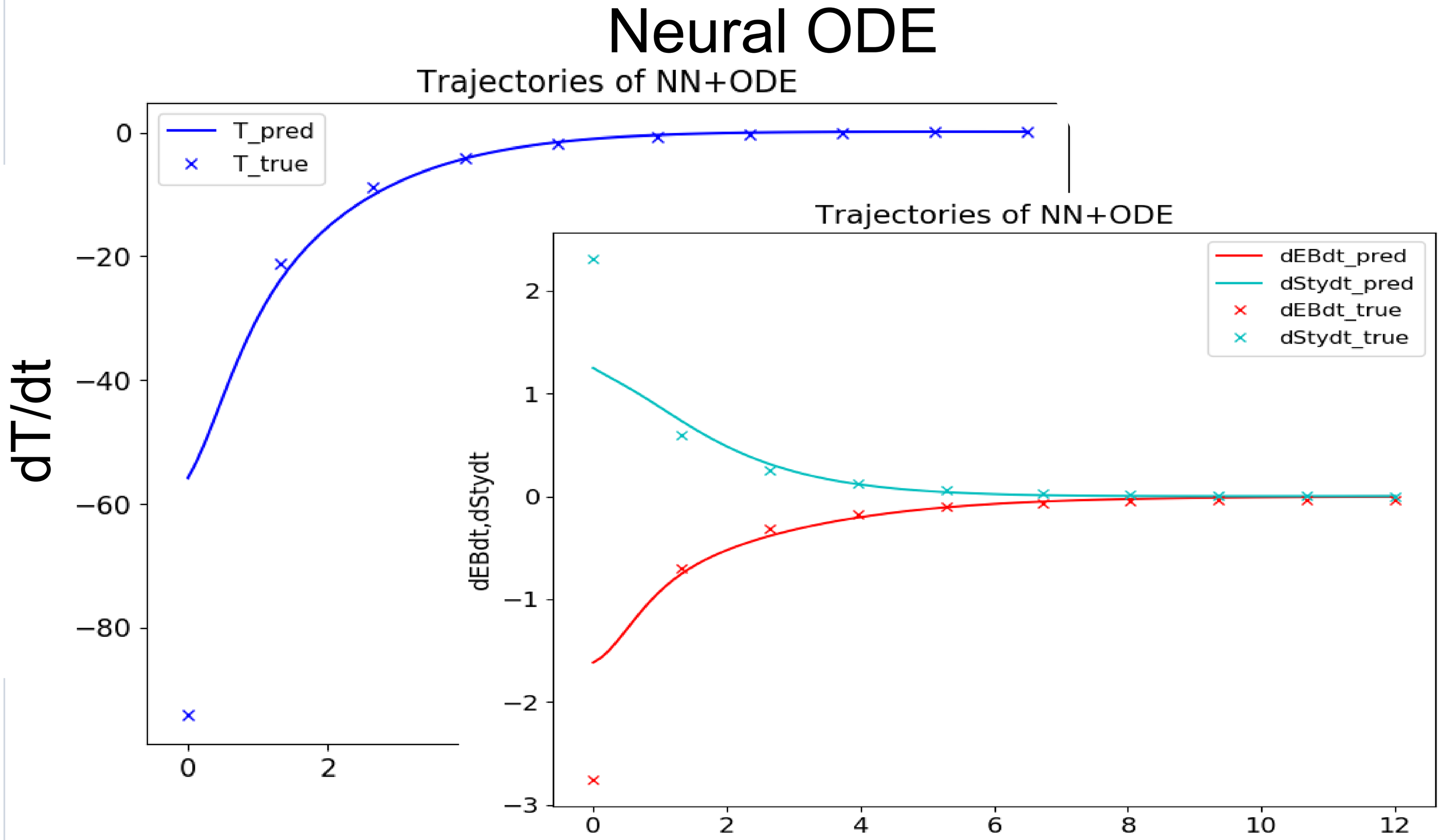


Derivative Estimation: a-NN vs NODE

Algebraic Neural Network

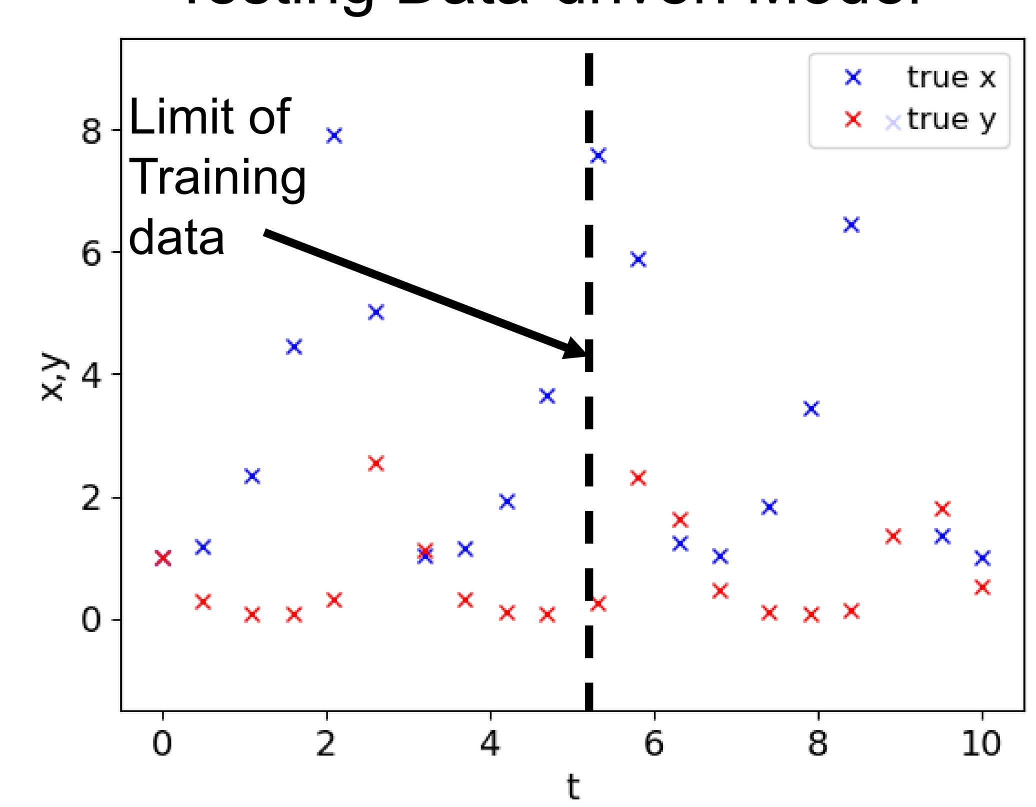


Neural ODE

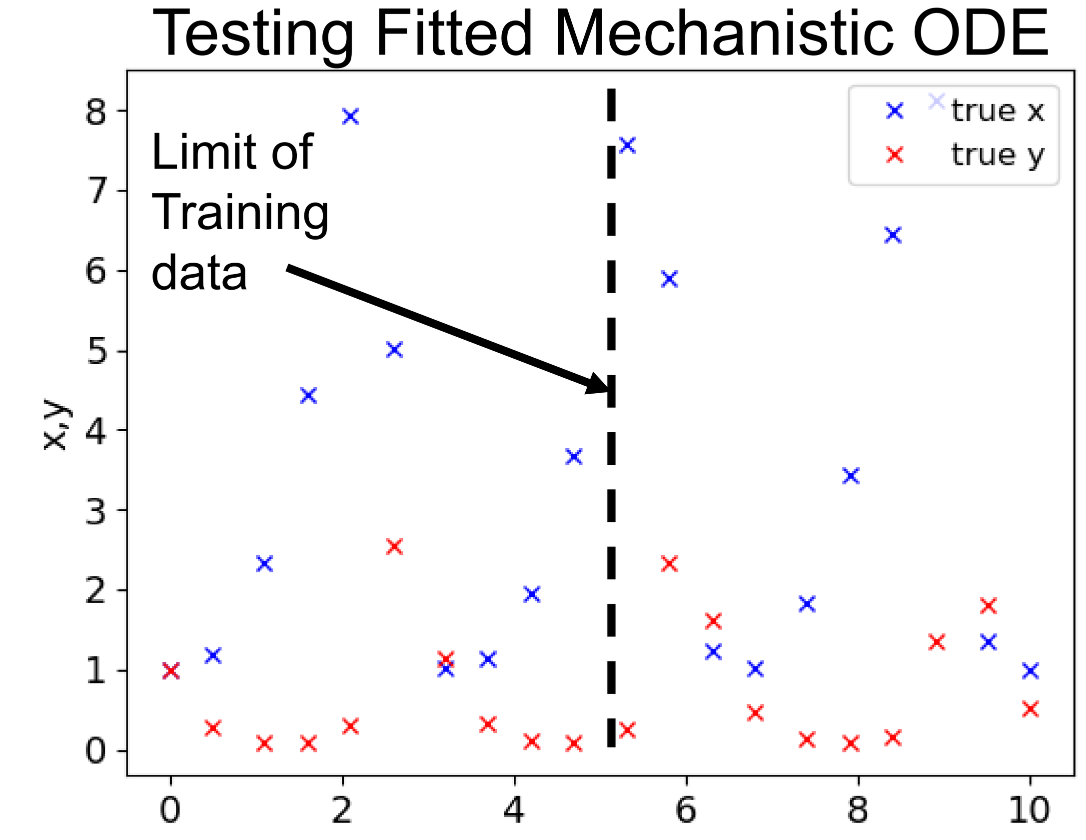


Extrapolation: NODE vs MODE

Testing Data-driven Model



Testing Fitted Mechanistic ODE



Indirect vs Direct Parameter Estimation

	Lotka Volterra (2 states, 3 parameters, 20 dp)	Ethylbenzene (6 states, 3 parameters, 60 dp)	Penicillin (3 states, 11 parameters, 90 dp)
Direct Approach	Total: 76 s	Total: 352 s	Did not converge
Indirect NODE Approach	Total: 62 s Steps 1&2: 62. s Step 3: 0.009 s	Total: 116 s Steps 1&2: 110 s Step 3: 6.76 s	Total: 183 s Steps 1&2: 181 s Step 3: 1.932 s

dp = data points
NODE = Neural ODE

Conclusions

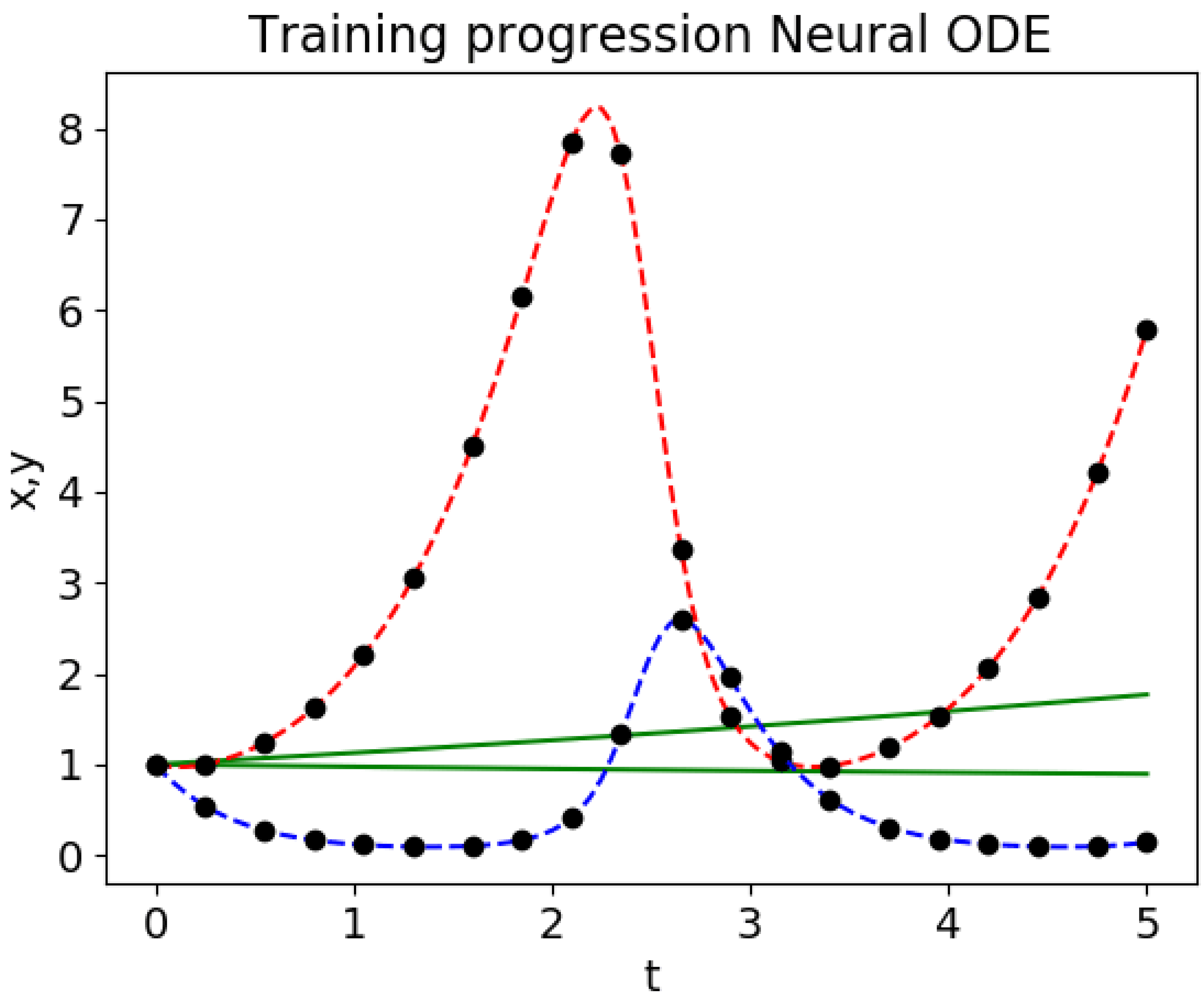
- NODEs estimate state derivatives more accurately than algebraic data-driven models
- Mechanistic ODEs have superior generalizability than Neural ODEs
- NODE Indirect approach can regress parameters of mechanistic models faster than direct approaches

Acknowledgements

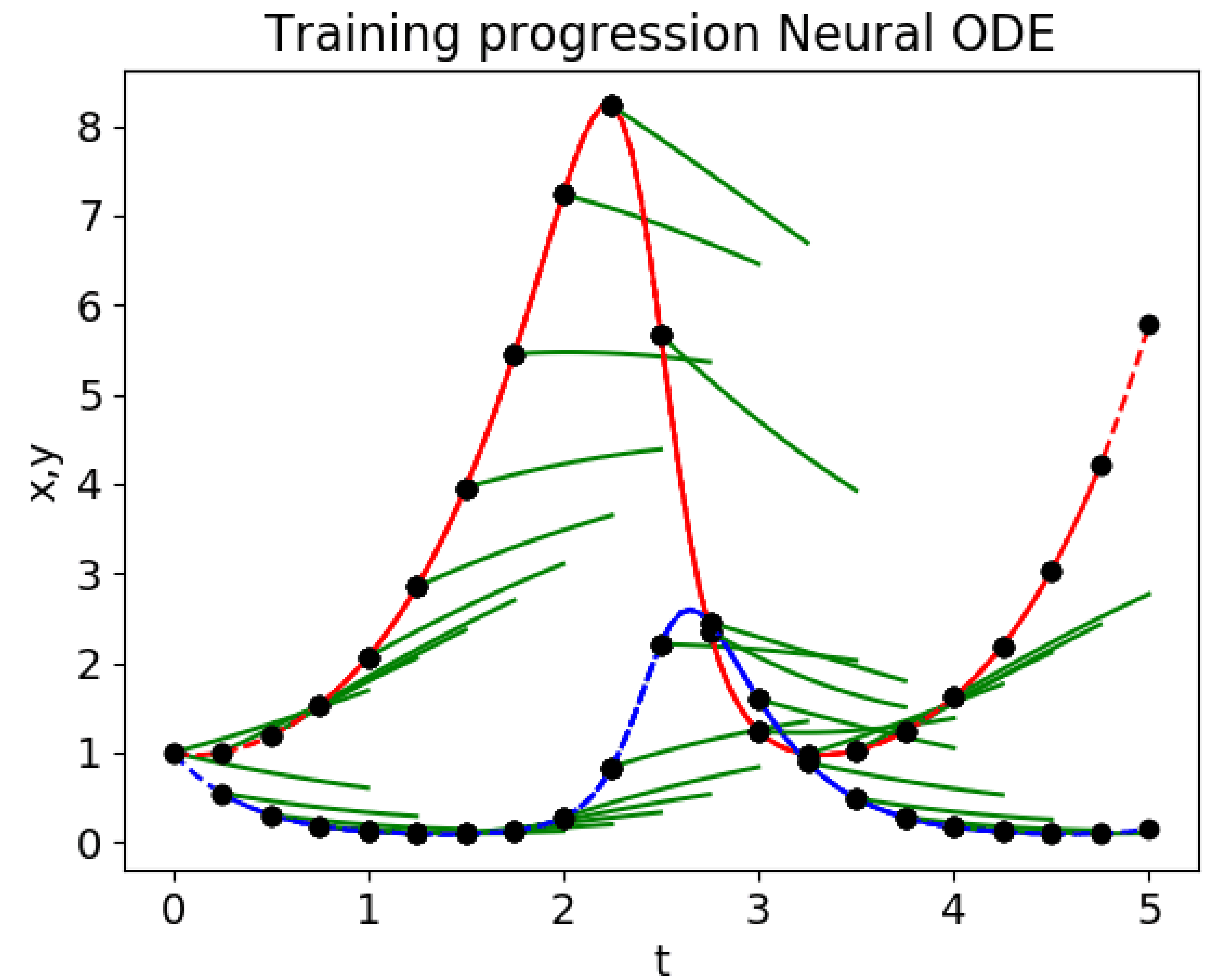
Georgia Tech Startup Grant
RAPID/NNMI Grant #GR10002225

Extra Slides

Results: Avoiding convergence to local minimum during NODE training



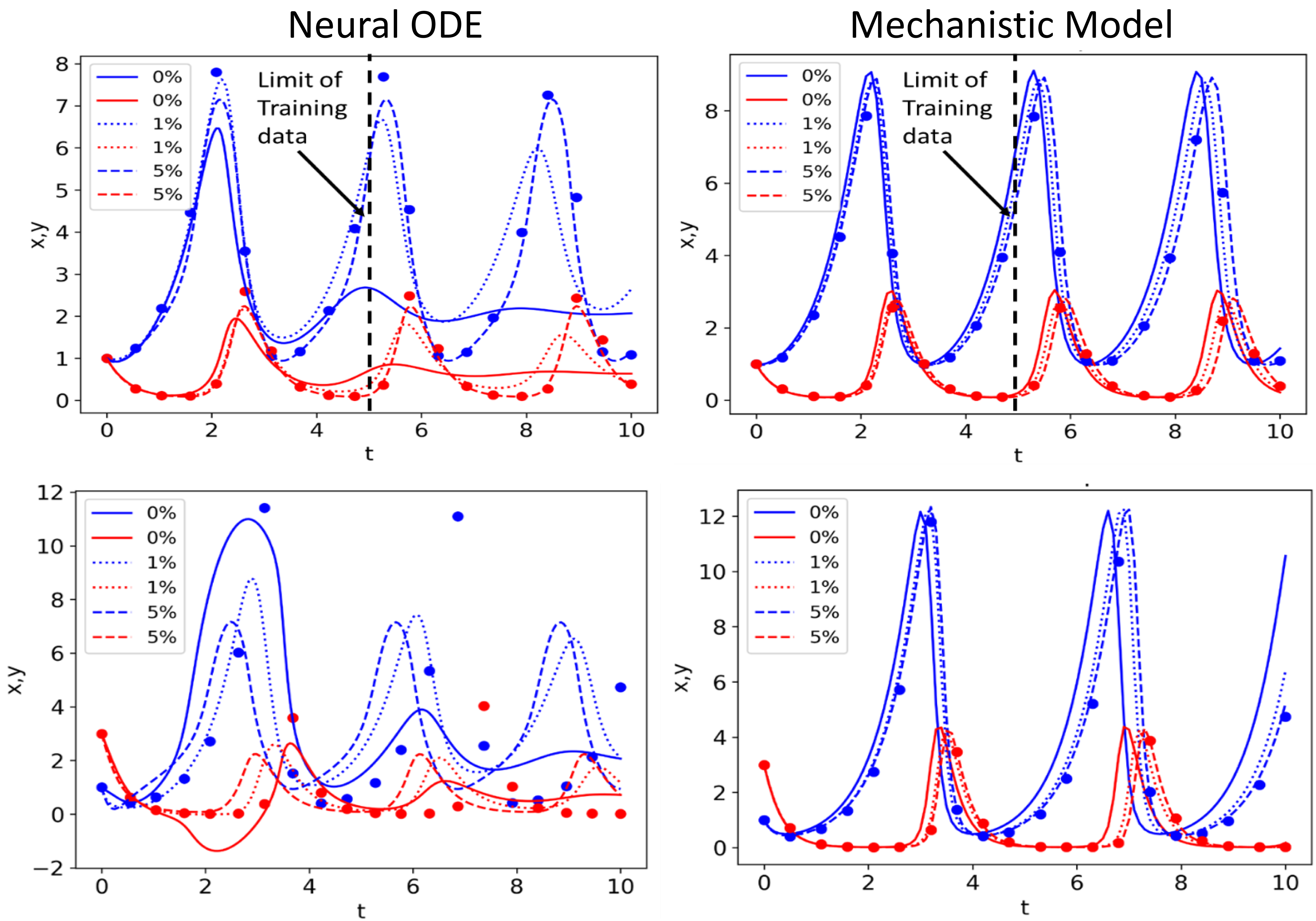
Integration from t = 0 to 5



Integration from $t = t_o$ to $t_o + 4\Delta t$

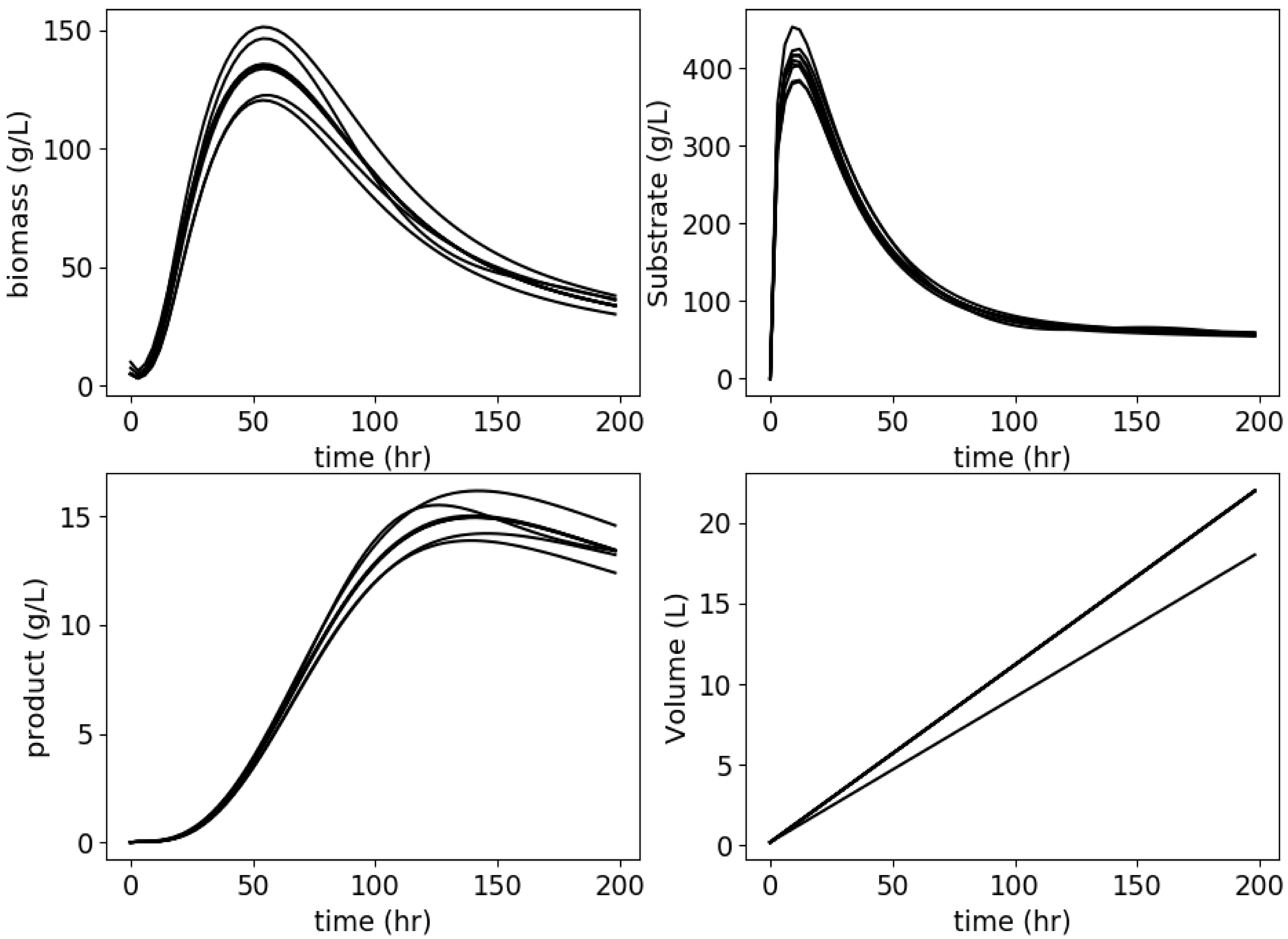
$t_o = \{0, 1\Delta t, 2\Delta t, 3\Delta t, \ldots, t_f\}$

Results: Extrapolation of NODE vs Mechanistic ODE

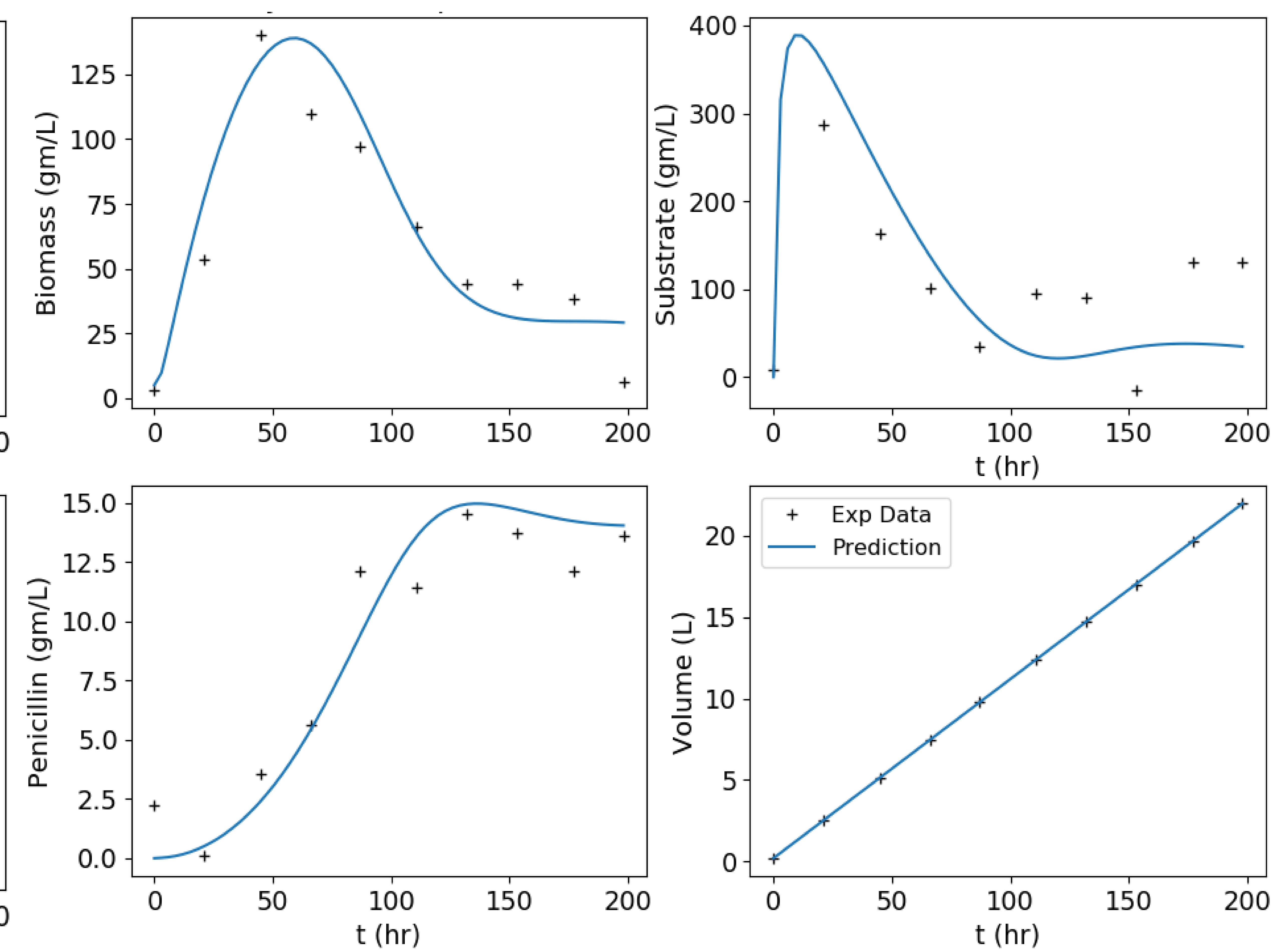


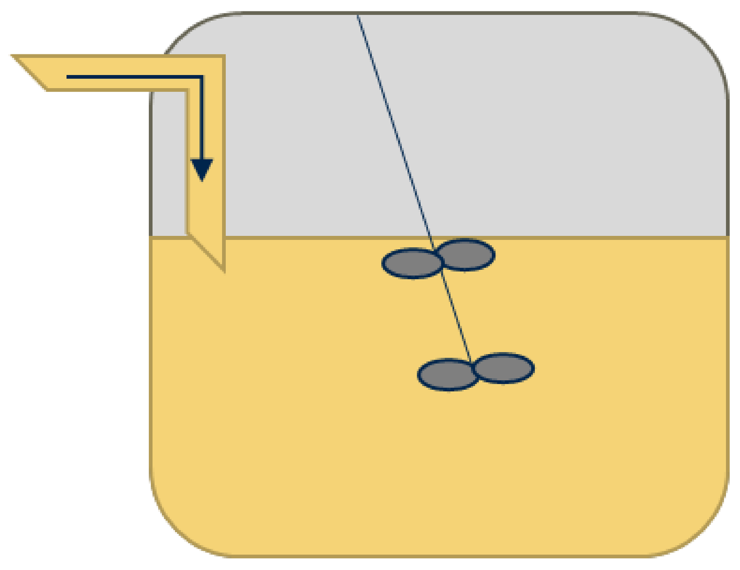
Results: indirect approach can fit ODE models with many parameters

Simulation of mechanistic ODEs with true parameters

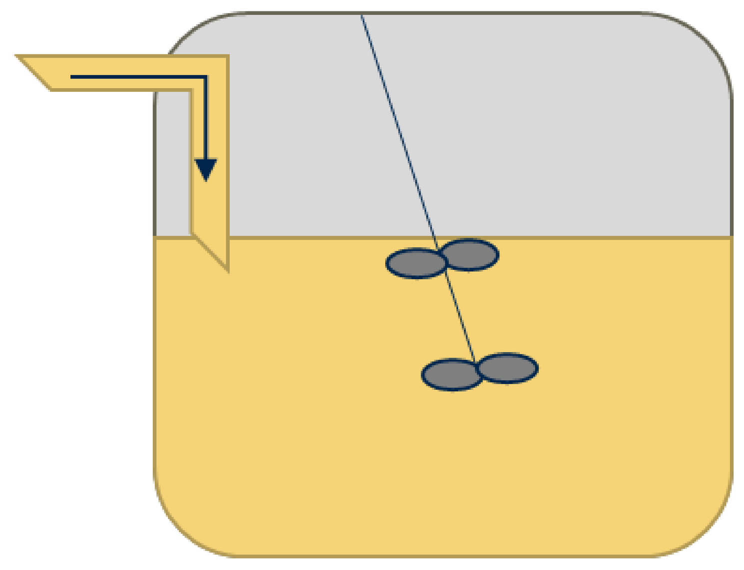


Simulation of mechanistic ODE; parameters fitted with noisy data via indirect approach





$$\frac{dX}{dt} = X\mu - XD$$



$$\frac{dX}{dt} = X\color{red}{\mu}\color{black} - XD$$