Automatic Discovery of Partial Differential Equation Models from Broadly Spaced Data

UCONN SCHOOL OF ENGINEERING
CHEMICAL & BIOMOLECULAR ENGINEERING

Benjamin Cohen[1], Burcu Beykal[1,2], George Bollas[1]

[1]Department of Chemical and Biomolecular Engineering, University of Connecticut, Storrs, CT; [2]Center for Clean Energy Engineering, University of Connecticut, Storrs, CT;

INTRODUCTION

PARTIAL DIFFERENTIAL EQUATIONS (PDES)

- Predict State Variable in time and space
- Explain changes in time and space
- Discover PDE models from data when space between measurements is small

$$\frac{\partial u}{\partial t}$$ Dynamics

$$\frac{\partial u}{\partial z}$$ Convection

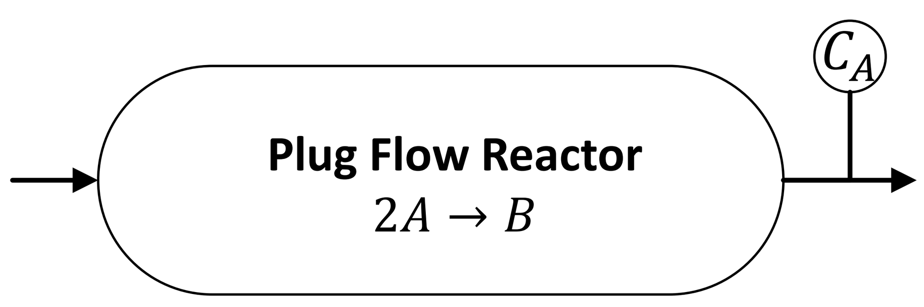
$$\frac{\partial^2 u}{\partial z^2}$$ Diffusion

$$\left(\frac{\partial V}{\partial T}\right)_P$$ Thermal Expansivity

$$\left(\frac{\partial V}{\partial P}\right)_T$$ Isothermal Compressibility

BROADLY SPACED DATA

- Spacing sensors frequently may be difficult or expensive
- Measuring chemical species is often done at the outlet of a reactor



OPPORTUNITY
Create partial differential equation models from broadly spaced measurements

RECOVER UNKNOWN PHYSICS FROM DATA

CHEMICAL REACTOR PHYSICS

$$\frac{\partial C_A}{\partial t} = s\left(C_A, \frac{\partial C_A}{\partial z}, \frac{\partial^2 C_A}{\partial z^2}, v, C_{A,f}\right)$$

PHYSICS

- Dynamics: $\frac{\partial C_A}{\partial t}$
- Reaction: θC_A^{ϕ}
- Convection: $\theta \frac{\partial C_A}{\partial z}$
- Diffusion: $\theta \frac{\partial^2 C_A}{\partial z^2}$

CONTROLS

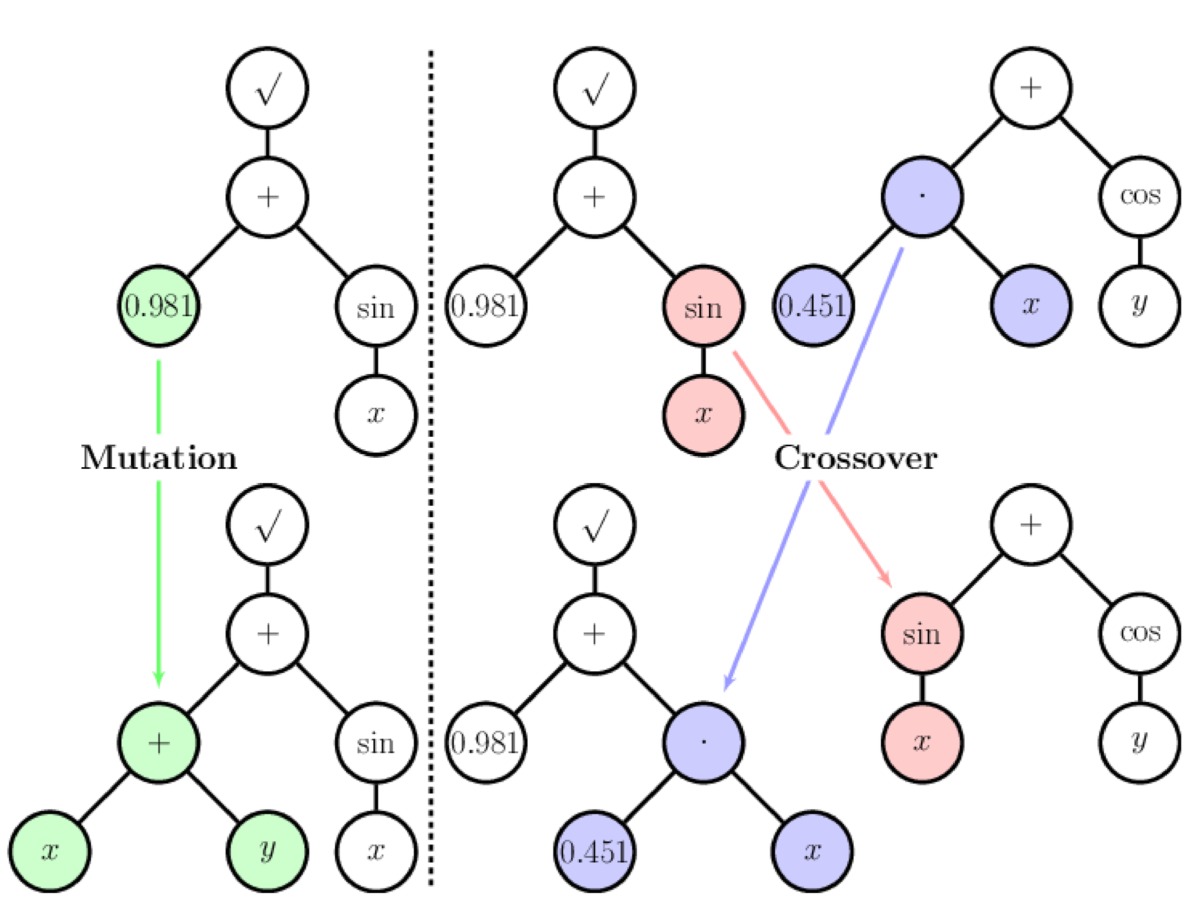
- Flow Velocity: θv
- Feed Concentration: $\theta C_{A,f}$

- Initial Condition: $C_A = 0$ at $t = 0$
- Inlet Condition: $C_A = C_{A,f}$ at $z = 0$
- Outlet Condition: $\frac{\partial C_A}{\partial z} = 0$ at $z = l$

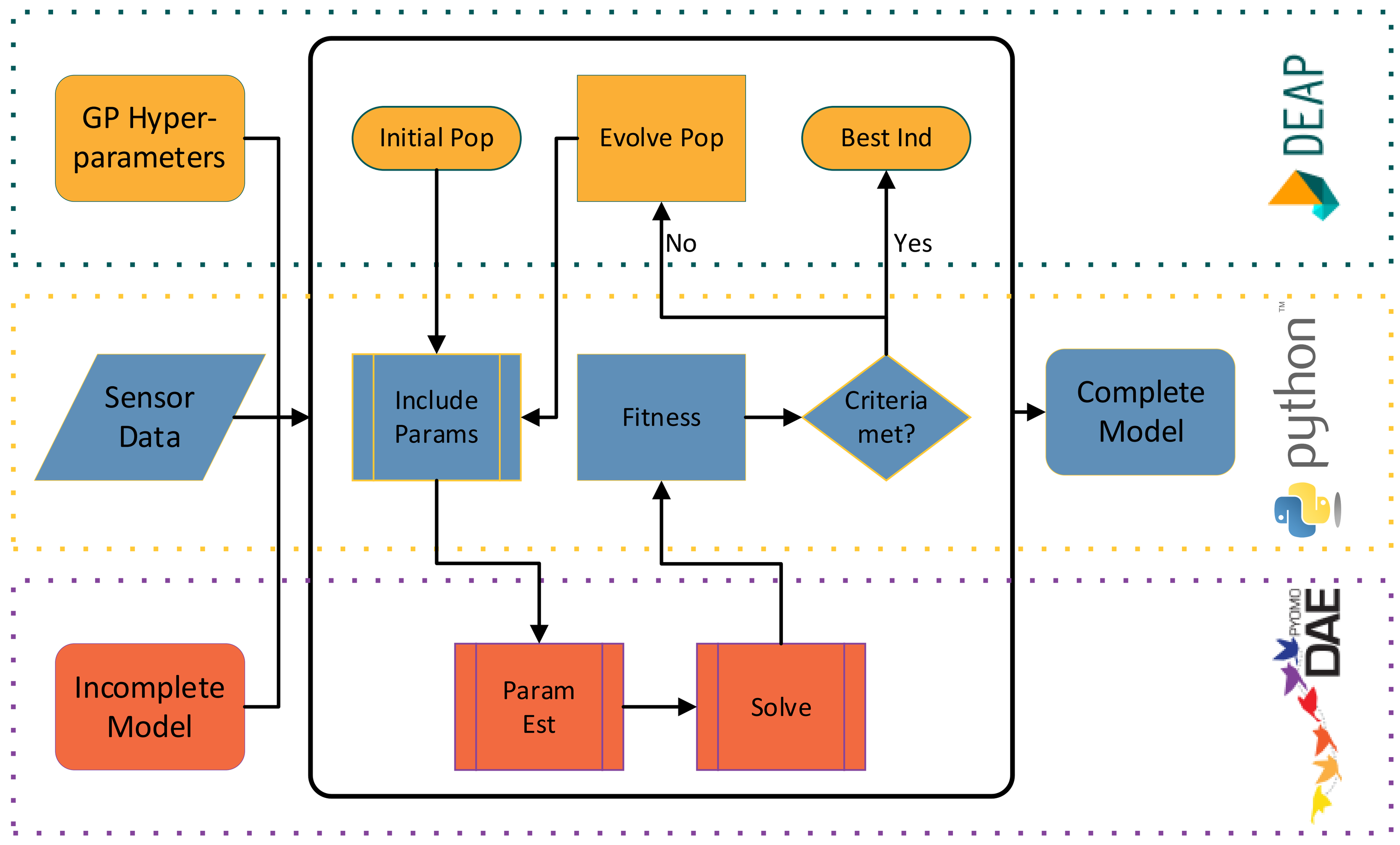
SYMBOLIC REGRESSION VIA GENETIC PROGRAMMING

Symbolic Regression
- Discover model shape
- Physics-informed
- Data-driven
- Identify parameters

Genetic Programming
- Inspired by evolution
- Mutation
- Cross-over



FRAMEWORK



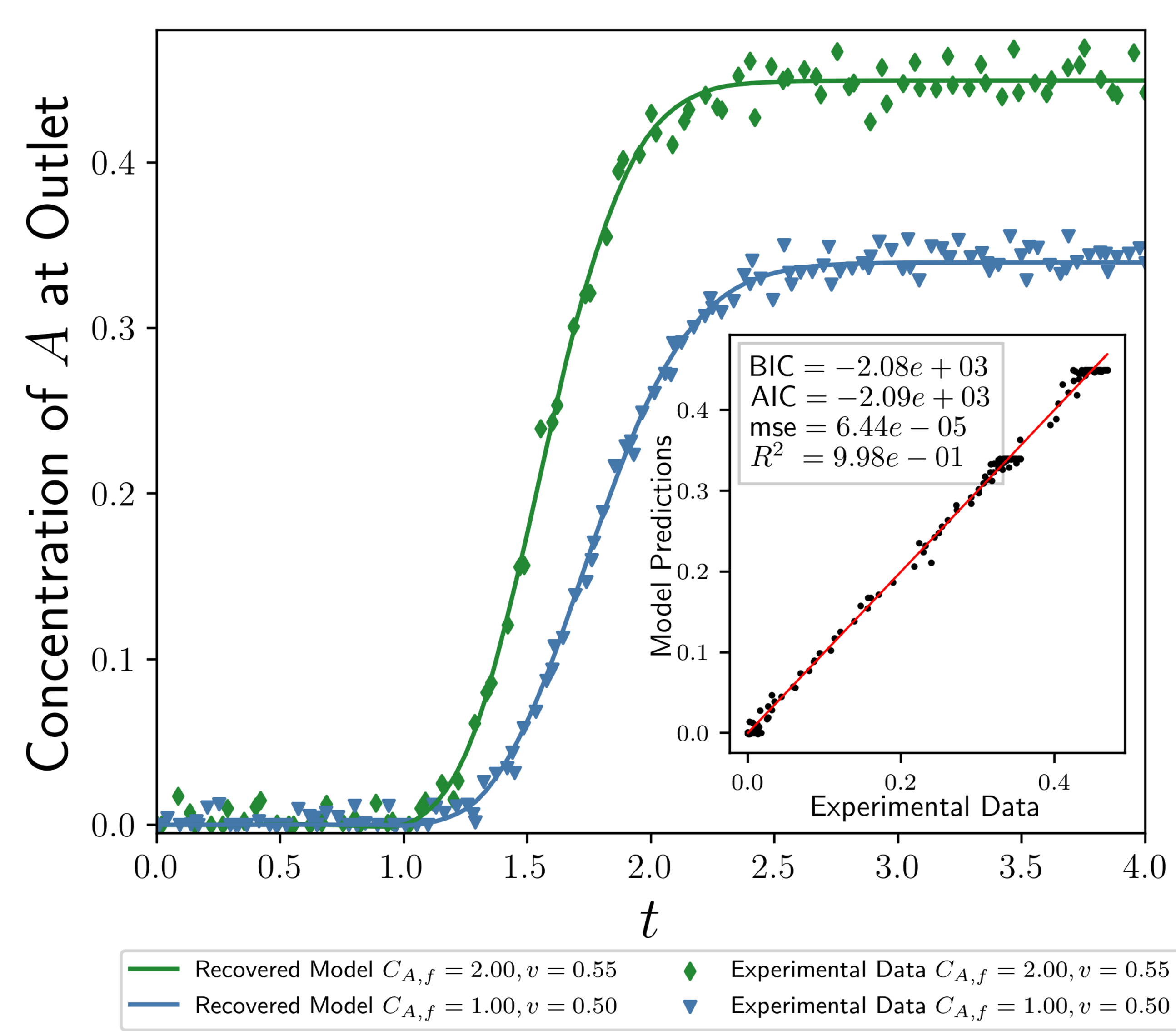
$$s^* \in \arg\min_{s \in S} I(s)$$
$$s.t.$$
$$\theta^* \in \arg\min_{\theta \in \mathbb{R}} \frac{1}{t_f}\int_0^{t_f} (\hat{C}_A - C_{A,meas})dt$$
$$s.t.$$
$$\frac{\partial C_A}{\partial t} - s = 0$$
$$C_A = 0 \text{ at } t = 0$$
$$C_A = C_{A,f} \text{ at } z = 0$$
$$\frac{\partial C_A}{\partial z} = 0 \text{ at } z = l$$
$$t \in [0, t_f], z \in [0, l]$$

- Minimize Information Theoretic Criterion, $I(s)$
- Find least complex model that agrees with data
- Gradient-based parameter estimation scheme

RESULTS

PLUG FLOW REACTOR (PFR)

$$\frac{\partial C_A}{\partial t} = 0.01\frac{\partial^2 C_A}{\partial z^2} - 1.00v\frac{\partial C_A}{\partial z} - 1.00C_A^{2.00}$$
$$\frac{\partial C_A}{\partial t} = 0.01\frac{\partial^2 C_A}{\partial z^2} - 1.00v\frac{\partial C_A}{\partial z} - 1.02C_A^{1.95}$$



- Successfully recovered model structure and parameters
- No predefined combinations of argument set terms

CONCLUSIONS AND FUTURE WORK

CONCLUSIONS

- Framework can recover simple PDE models
- Framework can discard terms from argument set
- Framework does not need predefined combinations of terms

FUTURE WORK

- Increase system model complexity (non-isothermal reactors)
- Leverage information theory for optimal design of experiments

REFERENCES

- Fortin, F.-A., et al. (2012). DEAP: Evolutionary Algorithms Made Easy. Journal of Machine Learning Research, https://github.com/DEAP/deap
- Quade, M., et al. (2016). Prediction of dynamical systems by symbolic regression. *Physical Review E*, https://doi.org/10.1103/PhysRevE.94.012214
- Messenger, D.A., and Bortz, D.M. (2020). Weak SINDy for Partial Differential Equations. *Journal of Computational Physics*, https://doi.org/10.1016/j.cp.2021.110525