

LEARNING-BASED RECONSTRUCTION OF CHEMICAL REACTION NETWORKS

Juechun Tang[†] Hongtao Zhong[‡]

[†]Dept of Chemical and Biological Engineering

[‡]Dept of Mechanical and Aerospace Engineering

INTRODUCTION

Dynamics in chemical reaction networks (CRNs) are of central importance for understanding various phenomena including explosion, catalysis, and biological systems.

CRN Formulation

$$\dot{Y}_i(t) = \frac{dY_i}{dt} = \sum_j w_{i,j} = f(Y_i, k_j) \quad i \in S, \quad j \in R$$

Y : species concentrations S : a set of chemical species
 k_j : reaction rate constant R : a set of reactions

CRN Reconstruction:

$k = \mathcal{F}(S, R, Y_i^*(t^*))$
Given experimental species temporal profiles as $Y_i^*(t^*)$, can we determine the rate constants k ?

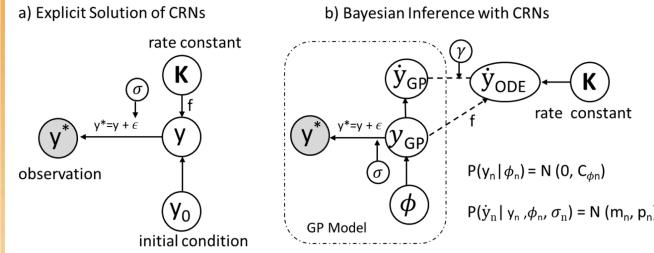
Limitations of Current Methods: manual fitting; complex numerical integration; much prior knowledge.

Why Machine Learning?: More efficient than numerical integration; May provide a distribution of inferred rate constants.

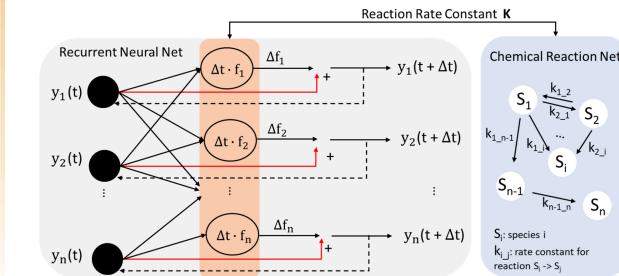
Our Goal: Apply machine learning techniques to reconstruct chemical reaction networks from discrete experimental data, facilitating research in complex chemical kinetics.

METHODS AND APPROACH

Adaptive Gradient Matching with Gaussian Process [2]

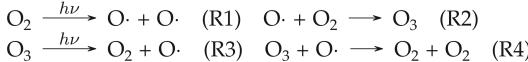


Recurrent Neural Net [3]

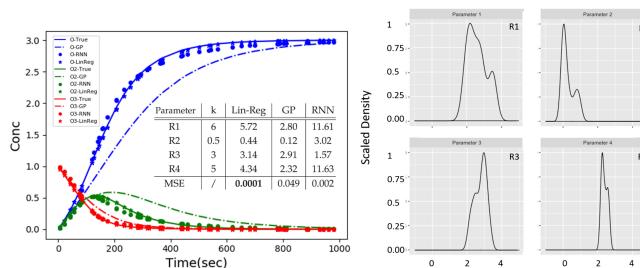


CASE 1: ATMOSPHERIC O₂-O₃-O KINETICS

O₂-O₃-O kinetics is an indispensable component of atmospheric chemistry [1]. The reduced mechanism is written as

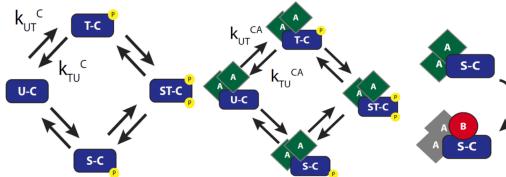


Model Performance and Rate Constant Distribution:

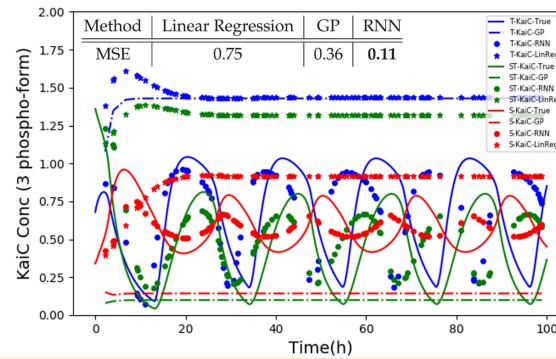


CASE 2: BIOCHEMICAL KINETIC OSCILLATION

We demonstrate the parameter learning process with a real 22h-period circadian oscillator found in cyanobacteria [4].



Model Performance with Inferred Rate Constants :

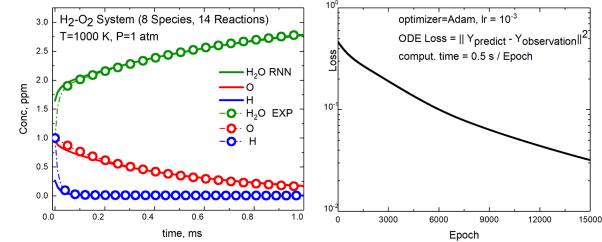


REFERENCES

- [1] L. Barrie, J. Bottemann, R. Schnell, P. Crutzen, and R. Rasmussen. Ozone destruction and photochemical reactions at polar sunrise in the lower arctic atmosphere. *Nature*, 334(6178):138, 1988.
- [2] B. Calderhead, M. Girolami, and N. D. Lawrence. Accelerating bayesian inference over nonlinear differential equations with gaussian processes. In *Advances in neural information processing systems*, pages 217–224, 2009.
- [3] H. Ling, S. Samarasinghe, and D. Kulasiri. Novel recurrent neural network for modelling biological networks: oscillatory p53 interaction dynamics. *Biosystems*, 114(3):191–205, 2013.
- [4] M. Rust, J. Markson, W. Lane, D. Fisher, and E. O’Shea. Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science*, 318(5851):809–812, 2007.

CASE 3: TOWARDS LARGE-SCALE KINETICS

Chemical reaction networks are intrinsically large-scale. Here we present RNN reconstruction of a H₂-O₂ combustion mechanism with 14 reactions.



Model Performance with Inferred Rate Constants:

	Reactions	k	Initial ϵ	RNN ϵ
1	H ₂ + OH → H ₂ O + H	1.3×10^{-12}	0.39	0.02
2	O + H ₂ O → OH + OH	1.0×10^{-13}	0.30	0.36
3	HO ₂ + OH → H ₂ O + O ₂	4.0×10^{-11}	1.90	1.91
...
14	H ₂ O ₂ + H → HO ₂ + H ₂	1.0×10^{-12}	0.32	0.02

k (cm³ / molecule s) retrieved from NIST chemical kinetics database. ϵ is the deviation from the true rate constants defined as $\|(k - k_{\text{true}})/k_{\text{true}}\|$

SUMMARY AND CONCLUSION

Conclusion

- RNN best reserves the system dynamics from small to large-scale systems
- Lin-Reg (with poly-terms) performs best in simple kinetics but fails to capture nonlinear dynamics in real systems

Discussion and Extensions

- This study is limited to a **fully observed system**. CRN reconstruction from a partial observed system would be more applicable and interesting
- Rate constants are assumed to be constant in this work. In reality, there is a strong **nonlinear dependence** between rate constants and temperature, pressure, etc.
- **Large stiffness** is a distinctive feature of CRNs. Rate constants may vary up to several orders of magnitudes. Model reduction is necessary before any learning-based reconstruction.