System Identification of Large-Scale Coupled Dynamics Using Differentiable Programming

Joshua Everts
Department of Bioengineering
UC Berkeley – UCSF
josh_everts@berkeley.edu

Abhay Shinde
Deep Forest Sciences, Inc.
abhay@deepforestsci.com

Bharath Ramsundar
Deep Forest Sciences, Inc.
bharath@deepforestsci.com

Sandya Subramanian
Dept. of Computational Precision Health
University of California, Berkeley
sandyas@berkeley.edu

Abstract

Synchronized nonlinear dynamical systems underpin a wide range of phenomena, from lasers to brainwaves to traffic. Nonlinear ODEs describe these dynamics well in simplified regimes, but they remain challenging to scale to real-world, high-dimensional data. The Kuramoto model for coupled oscillatory systems is a good example, as it has been widely applied, yet performing system identification remains challenging [1,6]. Maximum likelihood estimation (MLE) has shown promise for Kuramoto parameter estimation [5] but requires specialized implementation and remains challenging to apply to real data. We address this gap by using differentiable programming [4] as a new paradigm to scale system identification of the Kuramoto model to high dimensions. We show in synthetic data that our algorithm compares favorably to maximum-likelihood (MLE) estimation, with significantly improved noise and time-scale invariance. This suggests that the differentiable programming approach can be expanded to the study of other nonlinear dynamical systems applied to real-world data.

1 Introduction

Many phenomena can be described as dynamical systems with synchrony, where the phases of oscillatory signals entrain over time. Examples include photon interactions that create lasers, neural signals that enable thought, and even the formation of traffic jams [1, 6]. One canonical ODE for these effects is the Kuramoto model, which uses a system of oscillators parameterized by intrinsic frequencies ω_i, a phase lag α, and a cross-coupling matrix K_{ij} that describes dynamic relationships within complex systems. [1, 5, 7].

$$\dot{\theta}_i = \omega_i + \sum_{j \neq i} K_{ij} \sin(\theta_j - \theta_i - \alpha_i), \qquad i = 1, \ldots, N \tag{1}$$

Maximum likelihood estimation of K_{ij} provides a rigorous framework for parameter inference but requires explicit likelihood formulations, sampling assumptions, and does not estimate all parameters[4]. We instead chose to adapt an approach implemented via differentiable programming, which enables gradient-based optimization of all parameters by automatic differentiation [2]. Differentiable programming based parameter estimation works for a general function using timestep-wise backpropagation of the loss between forward-integrated trajectories and the current estimate [3]. The loss function can then be minimized using standard optimization frameworks combined with automatic differentiation and backpropagation [2]. We apply this estimation technique to the Kuramoto-Sakaguchi ODE and show that it offers distinct benefits compared to classical MLE.

Preprint. Under review.

2 Methods

We randomly initialized K_{ij}, w_i, and α_i from a Gaussian distribution and simulated the system using a fourth-order Runge-Kutta integrator with $dt = .05s$. For the differential programming method we used the Adam optimizer for gradient descent and allowed both α and w_i to be learned. For standard estimation we used 1000 epochs and a linearly scheduled learning rate with range $[1e-3, 1e-4]$. We simulated trajectories for the MLE with fixed w_i and no α because the method we followed[4] cannot recover w_i and uses a different form of α. This allowed us to directly compare K_{ij} recovery.

Algorithm 1 N-Oscillator Algorithm for differentiable parameter estimation

Require: Initial estimates $K^{(0)}, \alpha^{(0)}, w_i^{(0)}$, observations y, learning rate η

Initialize K, α, w_i
for $e = 1, \ldots, E$ **do**
  $\hat{\theta} \leftarrow \mathrm{RK4}(K, \alpha, \theta_0, w_i)$
  $L \leftarrow \|\hat{\theta} - y\|_2^2$
  Compute $\nabla_K L, \nabla_\alpha L$ by automatic differentiation
  Update (K, α, w_i) using Adam
end for
return K, α

3 Results

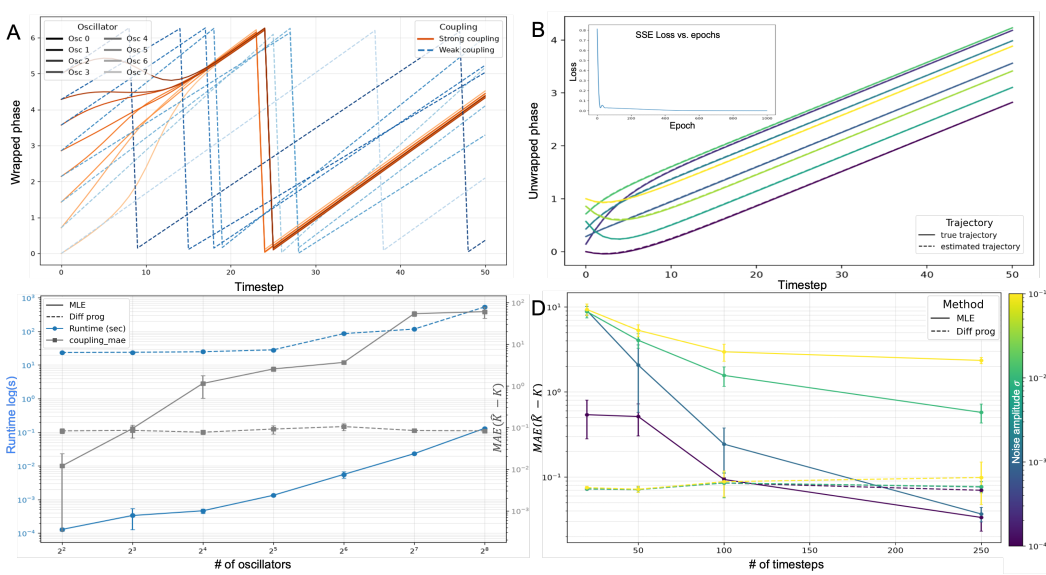

Figure 1: A. Example dynamics of large and small coupling K_{ij} terms. B. Coupled dynamics fitted with differential programming. C. Runtime and coupling error scaling with oscillator dimension between MLE and diff prog. D. Coupling recovery error between MAE and differential programming with noise and sample size.

Fig. 1A shows that Kuramoto systems can exhibit a wide range of dynamics, with coupling demonstrated at higher K_{ij} values. The differentiable programming approach faithfully reconstructs the dynamics of a coupled system (Fig. 1B), and compares favorably in parameter recovery accuracy (Fig. 1C), even when noise is added and number of samples is reduced (Fig. 1D). It is also much more accurate in high dimensions compared to the MLE estimate (Fig. 1C). However, the MLE estimate is much faster to run (Fig. 1C).

4 Discussion

The accurately reconstructed dynamics for systems of many oscillators demonstrate that differentiable programming can scale Kuramoto parameter estimation to higher dimensions. Furthermore, the estimates remain robust under increasing noise levels and perform well with short trajectories, often many orders of magnitude better than the classical MLE approach. However, this approach also comes with substantially higher computational expense, as runtimes were approximately 1000× slower than MLE on CPU. Overall, we demonstrate a proof-of-concept in synthetic data that differentiable programming approaches make high-dimensional nonlinear ODE parameter fitting significantly more feasible. This may allow the study of dynamic coupled systems using noisy data with short time scales, enabling real-time parameter estimation.

References

[1] J. A. Acebrón, L. L. Bonilla, C. J. P. Vicente, F. Ritort, and R. Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1):137–185, 2005.

[2] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind. Automatic differentiation in machine learning: A survey. *Journal of Machine Learning Research*, 18(153):1–43, 2018.

[3] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems 31*, pages 6571–6583, 2018.

[4] T. Hwang et al. A kuramoto-von mises time series model for probabilistic modeling of coupled oscillators. *arXiv preprint arXiv:2606.15012*, 2026.

[5] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer-Verlag, Berlin, 1984.

[6] F. A. Rodrigues, T. K. D. Peron, P. Ji, and J. Kurths. The Kuramoto model in complex networks. *Physics Reports*, 610:1–98, 2016. ISSN 0370-1573. doi: 10.1016/j.physrep.2015.10.008. URL `https://www.sciencedirect.com/science/article/pii/S0370157315004408`.

[7] H. Sakaguchi and Y. Kuramoto. A solvable active rotator model showing phase transitions via mutual entrainment. *Progress of Theoretical Physics*, 76(3):576–581, 1986.