Statistical Inference of Distributed Time Delay in Stochastic Chemical Reaction Networks: Application in Gene Regulation Process and Infectious Disease Transmission

Boseung Choi[1,2,3]

[1]Division of Big Data Science,
Korea University, Republic of Korea
cbskust@korea.ac.kr

[2]Biomedical Mathematics Group,
Institute for Basic Science, Republic of Korea
[3]College of Public Health,
The Ohio State University, USA

Advances in experimental or imaging techniques and sampling processes have provided unprecedented insights into stochastic reaction networks. However, many reactions yield outcomes that remain concealed or are only indirectly observable. Unobserved reactions can be substituted with time delays to mitigate this, reducing model dimensionality and simplifying inference. However, resulting non-Markovian models necessitate novel inference techniques. Current estimation methods for inferring reaction rates rely on marginalization over unobserved processes and states. This approach, while powerful, poses computational challenges, yielding large uncertainties and lacking robustness in parameter estimates.

In response, we present a Bayesian inference framework based on unobserved reactions with time delays. This novel approach extends Markov Chain Monte Carlo (MCMC) methods to estimate reaction constants efficiently and delay distribution parameters using synthetic and experimental data, demonstrating robust parameter inference even with limited measurements.

Our Bayesian inference method extends in two significant ways. First, simultaneous recordings of network dynamics across large populations reveal substantial species variability even within clonal lines. Quantifying parameter variability is crucial for understanding behavior, leading us to propose a non-Markovian, hierarchical Bayesian inference framework. This framework quantifies the variability of reaction network processes within and across species in a population. Second, despite experimental advances, we can only measure a fraction of the process. Therefore, we develop a simulation-based Bayesian MCMC method using an approximate likelihood for efficient and accurate inference of reaction network parameters when only some products are observed. We applied these methods to transcriptional and translational regulation, as well as to the COVID-19 pandemic in Seoul, Korea, and demonstrated our approach's capability to accurately estimate delay and kinetic parameters, even when only a subset of the data is observable.

In conclusion, our proposed methods offer a robust and efficient means of inferring parameters in stochastic reaction networks, addressing challenges posed by unobserved reactions and parameter variability within populations. The application of these methods to diverse scenarios, from biological processes to epidemiological studies, underscores their versatility and effectiveness.

Keywords: Bayesian inference, Stochastic reaction network, Time delay, GRN, COVID-19