Observers for Stochastic Chemical Kinetics

David Thorsley, Robert Egbert, Eric Klavins

***Short Abstract* — Time-lapse movies are a powerful method for studying intracellular stochastic phenomena, but they are limited by the fact that most chemical species are unobservable. We apply the control-theoretic notion of an *observer* and stochastic chemical kinetic models to propose a method for estimating the dynamically changing molecular populations of unobservable species from time-lapse fluorescence data and demonstrate how the observer can be applied to the problems of state estimation, diagnosis, and hypothesis testing.**

I. Motivation

RANDOMNESS pervades the biochemical processes that make up life at the cellular level. A standard formalism for describing the variable behavior of stochastic processes inside the cell is stochastic chemical kinetics [1], [2]. The *state* of a stochastic chemical kinetic model is a vector where each element is the molecular population of one of the species in the biochemical process. Using such models, we can make predictions regarding the distributions of dynamic behaviors of cellular processes. These predictions can then be tested using single-cell experiments.

However, it is not experimentally possible to completely observe the evolution of the state vector as it varies with time. Time-lapse movies collected using single-cell fluorescence microscopy allow the experimenter to make estimates of the dynamic populations of a few fluorescent species. Observations of dynamic behavior are not made continuously but instead at intermittent time points. A well-designed experiment ensures that these limited observations provide enough useful information to draw conclusions about the behavior of unobservable species.

In control theory, the standard approach to dealing with the problem of estimating a dynamically changing state from limited observations is to construct an *observer*. An observer is a system that receives measurements from a system being monitored and computes an estimate of the monitored system's state. In this work, we develop an observer structure specialized for systems described using the formalism of stochastic chemical kinetics and specialized for quantitative analysis of time-lapse movies made using fluorescence microscopy. We show how observers can be applied to the following problems:

- *State estimation*: Given a sequence of observations y^n, what is the *a posteriori* probability distribution vector $\mathbf{p}(t \mid y^n)$ of the molecular populations?
- *Diagnosis*: Given a sequence of observations, what is the *a posteriori* probability that an event of interest occurred?
- *Hypothesis Testing*: Given a sequence of observations and a set of hypotheses, what is the *a posteriori* probability of each hypothesis being the correct one?

The authors are with the Department of Electrical Engineering, University of Washington, Seattle, WA, 98195 E-mail: {thorsley,robere,klavins}@ee.washington.edu

II. Results

We consider two possible observation paradigms: an ideal case in which observations of the system are noise-free and made continuously, and a practical case in which observations are noisy and made intermittently. For both cases, we derive an observer with discrete and continuous updates that calculates the probability distribution vector $\mathbf{p}(t \mid y^n)$ that is the solution to the state estimation problem. The expected value of this probability distribution is the optimal mean-square error estimator of the system state. We then show how the diagnosis and hypothesis testing problems can be treated as straightforward extensions of the state estimation problem.

We illustrate the performance of the observer on two systems: one in simulation and one in experiment. In simulation, we consider a model of the stochastic Pap epigenetic switch in *E. coli*, which controls the expression of pili on the bacterial surface. We show that by observing the population change of PapI protein, we can use the observer and diagnoser to make probabilistic inferences as to the configuration of the underlying switch.

In experiment, we consider an IPTG-inducible GFP production circuit in *E. coli*. We evaluate the performance of the observer as hypothesis tester by growing the bacteria in different concentrations of IPTG and measuring the dynamic GFP expression with time-lapse microscopy. Using the expression data, we compare the observer's calculation of the most likely concentration of IPTG with the true level. This experiment demonstrates that it is possible to indirectly estimate the population of an unobservable species inside a cell by measuring the dynamic behavior of downstream species, using only a simple model of the stochastic chemical process. Full results will soon be available in [3].

References

[1] D. McQuarrie, "Stochastic approach to chemical kinetics," *Journal of Applied Probability,* vol. 4, pp. 413–478, 1967.

[2] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Third Edition (North-Holland Personal Library). North Holland, 2007.

[3] D. Thorsley, R. Egbert, and E. Klavins, "State estimation in stochastic gene regulatory networks," in preparation.