Hybrid Numerical Solution of the Chemical Master Equation

Thomas A. Henzinger[1], Maria Mateescu[2], Linar Mikeev[3], and Verena Wolf[4]

***Short Abstract* — We present a numerical approximation technique for the analysis of continuous-time Markov chains that describe networks of biochemical reactions and play an important role in the stochastic modeling of biological systems.**

Our approach is based on the construction of a stochastic hybrid model in which certain discrete random variables are approximated by continuous deterministic variables.

We compute the solution of the stochastic hybrid model using a numerical algorithm that discretizes time and in each step performs a mutual update of the transient probability distribution of the discrete stochastic variables and the values of the continuous deterministic variables.

***Keywords* — Markov process, biochemical reaction network, chemical master equation, stochastic hybrid model.**

I. INTRODUCTION

THE traditional approach for a dynamical model of cellular reaction networks is based on the assumption that the concentrations of the chemical species change continuously and deterministically in time. During the last decade, however, continuous-time Markov chains (CTMC) that represents the chemical populations as random variables have seen increased interest. The reason is that they take into account the effects of molecular noise in the cell.

The evolution of the CTMC is given by a system of linear ordinary differential equations, known as the chemical master equation (CME). However, if the system under consideration contains large populations, then the numerical algorithms that approximate the solution of the CME perform poorly because the random variables that represent large populations have a large variance. Thus, a large number of states have a significant probability, which renders the numerical approximation of the distribution computationally expensive or infeasible.

II. STOCHASTIC HYBRID MODEL

We propose a stochastic hybrid approach[1] to efficiently approximate the solution of systems containing both small and large populations. More precisely, we maintain the discrete stochastic representation for small populations, but at the same time we exploit the small relative variance of large populations and represent them by continuous deterministic variables. Since population sizes change over time we decide dynamically ("on-the-fly") whether we represent a population by a continuous deterministic variable or keep the discrete stochastic representation. Our criterion for changing from a discrete to a continuous treatment of a variable and vice versa is based on a population threshold.

For the solution of the stochastic hybrid model, we propose a numerical approximation method that discretizes time and performs a mutual update of the distributions of the discrete stochastic variables and the values of the continuous deterministic variables. Hence, we compute the solution of a CME with a reduced dimension as well as the solution of a system of (non-linear) ordinary differential equations. The former describes the distribution of the discrete stochastic variables and the latter the values of the continuous deterministic variables and the two descriptions depend on each other. Assume, for instance, that a system has two chemical species. The two population sizes at time t are represented by the random variables $X(t)$ and $Y(t)$, where $X(t)$ is large and $Y(t)$ is small. Then, we consider for $Y(t)$ all events $Y(t)=y$ that have significant probability, i.e., $Pr(Y(t)=y)$ is greater than a small threshold. For $X(t)$ we consider the conditional expectations $E[X(t)|Y(t)=y]$ and assume that they change continuously and deterministically in time. We iterate over small time steps $h>0$ and, given the distribution for $Y(t)$ and the values $E[X(t)|Y(t)=y]$, we compute the distribution of $Y(t+h)$ and the values $E[X(t+h)|Y(t+h)=y]$. Again, we restrict our computation to those values of y that have significant probability.

To demonstrate the effectiveness of our approach, we have implemented the algorithm and applied it successfully to several examples from systems biology. Our most complex example has 6 different chemical species and 10 reactions. Our results show that significant computational time is saved. We compared our solution to a purely stochastic solution and found that the relative error of the first moment is small (about 1%) for all species. Also higher moments are predicted accurately for those species that are represented by stochastic variables. The method is particularly attractive for stochastic switches but also for oscillatory systems where the population sizes change drastically during one period.

[1]Institute of Science and Technology, Austria. E-mail: tah@ist.ac.at
[2]Ecole Polytechnique Federale de Lausanne, Switzerland. E-mail: maria.mateescu@epfl.ch
[3]Saarland University, Germany. E-mail: mikeev@cs.uni-saarland.de
[4]Saarland University, Germany. E-mail: wolf@cs.uni-saarland.de

REFERENCES

[1] Henzinger T A, Mateescu M, Mikeev L, Wolf V (in press) "Hybrid Numerical Solution of the Chemical Master Equation". In *Proc. Of CMSB*, 2010.