



ELSEVIER

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp



Hybrid discrete/continuum algorithms for stochastic reaction networks



Cosmin Safta*, Khachik Sargsyan, Bert Debusschere, Habib N. Najm

Sandia National Laboratories, Livermore, CA 94551-0969, United States

ARTICLE INFO

Article history:

Received 17 December 2013
 Received in revised form 9 July 2014
 Accepted 12 October 2014
 Available online 22 October 2014

Keywords:

Chemical Master Equation
 Fokker–Planck equation
 Finite volume
 Flux splitting
 Hybrid discrete–continuum models

ABSTRACT

Direct solutions of the Chemical Master Equation (CME) governing Stochastic Reaction Networks (SRNs) are generally prohibitively expensive due to excessive numbers of possible discrete states in such systems. To enhance computational efficiency we develop a hybrid approach where the evolution of states with low molecule counts is treated with the discrete CME model while that of states with large molecule counts is modeled by the continuum Fokker–Planck equation. The Fokker–Planck equation is discretized using a 2nd order finite volume approach with appropriate treatment of flux components. The numerical construction at the interface between the discrete and continuum regions implements the transfer of probability reaction by reaction according to the stoichiometry of the system. The performance of this novel hybrid approach is explored for a two-species circadian model with computational efficiency gains of about one order of magnitude.

Published by Elsevier Inc.

1. Introduction

Stochastic noise is prevalent in a wide variety of systems, especially when the system behavior is affected or controlled by phenomena on a nanoscale, molecular level, where thermal noise introduces intrinsic variability in molecular interactions. Common examples of such phenomena are chemical, biochemical, or electrochemical reactions between small numbers of molecules, found in gene regulation, cell signaling, or interfacial electrochemistry, whose models are generally referred to in this paper as Stochastic Reaction Networks (SRNs). Given the relevance of these processes in applications ranging from bioremediation and bioenergy (bacterial behavior), biomedicine (immune system signaling), to electrical storage (electrodes), effective tools are needed for the simulation and analysis of SRNs.

Mathematically, SRNs are continuous time, discrete state Markov processes, governed by the Chemical Master Equation (CME) [1], which describes the time evolution of the probability of the system being in a particular state. Here the system state consists of the number of molecules of each species in the system. While the CME can be solved as a system of linear equations, its solution quickly becomes challenging as the number of possible system states increases exponentially with the number of species. For this reason, many studies rely on the Stochastic Simulation Algorithm (SSA) [2,3], which can efficiently generate sample trajectories of the system state. Based on this fundamental simulation method, many advances have been made over the years [4–12] resulting in a wide range of powerful methods for sampling stochastic systems.

While SSA is a powerful tool for generating sample trajectories of the system state, even in high dimensions, one of its drawbacks is that many samples are needed to construct an accurate probability density function of the system state. As the

* Corresponding author.

E-mail addresses: csafta@sandia.gov (C. Safta), ksargsy@sandia.gov (K. Sargsyan), bjdebus@sandia.gov (B. Debusschere), hnnajm@sandia.gov (H.N. Najm).

CME solution offers a full probability distribution of the system states, rather than just samples, it is able to better capture low probability events, which are often missed by sampling-based methods unless a prohibitively large number of samples is used. In recent years, various approaches have been developed to make the direct solution of the CME computationally more tractable. One approach is *finite state projection* [13–16], which reduces the size of the system of equations to be solved by only keeping track of those states that have a non-negligible probability associated with them. This projected state space is adaptively expanded and shrunk as the state probabilities vary in time. Another approach to reducing the number of equations is to group multiple states together via *aggregation* [17,18]. For a given set of equations, advances in matrix exponentiation [18–21], or taking advantage of time scale separation [22] can speed up the CME solution. Other CME solution approaches rely on a spectral representation of the probabilities of the system states [17,23–27].

The combination of these algorithms has led to dramatic speedups in the solution of the CME, enabling the study of small SRNs by solving their CME directly. For small system (a few species), this approach has been shown to be competitive with and complementary to the SSA sampling based approaches (see e.g. [19]).

Despite these advances in solving the CME, a direct solution is still intractable for systems with many species, especially if some of those species are present in large numbers of molecules, thereby significantly increasing the size of the state space. For species that are present in large molecule counts, it is often appropriate to use a continuum representation for the system state [28], and rely on the Fokker–Planck equation (FPE) [29,30], which is the continuum equivalent to the discrete CME, or use deterministic rate equations. In this context, Sjöberg et al. [31] developed a hybrid approach that uses the FPE for species that are present in large numbers of molecules, while keeping the discrete CME for species that only have a few molecules.

Salis and Kaznessis [32], Haseltine and Rawlings [7], and references therein proposed hybrid approaches that partition the stochastic reaction set into slow and fast reactions. The approach introduced in [32] approximates the fast reaction subset as continuous Markov processes using the chemical Langevin equation, while the slow processes are still modeled via the SSA approach. Haseltine and Rawlings [7,33] proposed a hybrid approach that splits the reaction set into fast and slow subsets with the help of a reaction extent coordinate. The conversion makes use of conditional probability densities to derive approximate master equations from the original CME. Other hybrid approaches use the discrete SSA approach for some species, and continuum, deterministic rate equations for others [34].

While the hybrid approaches offer large computational savings for systems where a clean separation is possible between the species that should be handled with a discrete formulation, and those that can be handled with a continuum approach, such a separation is not always feasible, and may vary as species molecule counts vary in time. Rather than relying on a static separation, where species are handled with either a discrete or continuum formulation, the current work varies the formulation over different regions in the species phase space. In the region where all species are present with few molecules, the CME is used for all species, since discreteness is important in that region. In regions corresponding to a large number of molecules for all species the FPE formulation is used. In regions where some species are present in small numbers, but others in large numbers, the formulation by Sjöberg et al. [31] is used. We propose flux formulations and numerical discretizations that conserve the transfer of probability between discrete (CME), semi-discrete (Sjöberg), and continuous (FPE) formulations while preserving the order of accuracy of the numerical scheme across interfaces.

By using the discrete CME formulation only where it is needed, the algorithm presented in this paper can allocate computational resources (*i.e.* grid density) more effectively. Also, as will be shown, the resulting set of equations for the hybrid approach is very similar to the original set of linear equations corresponding to the CME. This means that the hybrid algorithm can be used in synergy with other approaches such as finite state projection and improved matrix exponentiation for further computational gains. We note that our approach could, in theory, be coupled with the algorithms proposed by Salis and Kaznessis [32] and Haseltine and Rawlings [7]. In regions where the discrete CME formulation is needed, the system could be split between slow and fast reactions to allow the use of efficient modeling approaches for time advancement of the solution in these parameter space regions.

For proof of concept, the approach proposed in this paper was implemented for systems with two molecular species, and applied to the simulation of SRNs describing a simple metabolite model, and a circadian rhythm model. Canonical tests show second order convergence and good accuracy compared to full CME simulation, with an order of magnitude gain in computational efficiency.

This paper is organized as follows. Section 2 describes the CME approach to study SRNs and compares its direct solution with results based on the SSA. Section 3 introduces the Fokker–Planck approximation to the CME, and results based on a finite volume discretization of the FPE are compared to full CME results. The hybrid CME–FPE formulation is formulated in Section 4, numerical experiments are presented in Section 5, and conclusions are presented in Section 6.

2. Chemical Master Equation

In this section we describe the Chemical Master Equation (CME) approach to study the dynamics of Stochastic Reaction Networks. The CME models the evolution of the probability of the state of the system.

Consider a stochastic reaction network (SRN) for a system consisting of two species, X and Y . The 2-dimensional state space: $\mathbf{n} = (n_x, n_y) \in \mathbb{N}_0^2$, corresponds to molecule counts for species X and Y where \mathbb{N}_0 is the set of all non-negative integers. Denote the (jump, propensity) pairs by $\{(\mathbf{v}_r, w_r(\mathbf{n}))\}_{r=1}^R$, where R is the number of reactions. More specifically, $\mathbf{n} \rightarrow \mathbf{n} + \mathbf{v}_r$ with probability $w_r(\mathbf{n})$ per unit time.

Download English Version:

<https://daneshyari.com/en/article/6931981>

Download Persian Version:

<https://daneshyari.com/article/6931981>

[Daneshyari.com](https://daneshyari.com)