Theoretical Computer Science 587 (2015) 49–72

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Approximate analysis of biological systems by hybrid switching jump diffusion

Alessio Angius [a], Gianfranco Balbo [a], Marco Beccuti [a], Enrico Bibbona [b], Andras Horvath [a,*], Roberta Sirovich [b]

[a] Università di Torino, Dipartimento di Informatica, Italy
[b] Università di Torino, Dipartimento di Matematica, Italy

ARTICLE INFO

Article history:
Received 16 April 2014
Received in revised form 10 January 2015
Accepted 3 March 2015
Available online 18 March 2015

Keywords:
Diffusion approximation
Jump diffusion
Stochastic differential equations with barriers

ABSTRACT

In this paper we consider large state space continuous time Markov chains arising in the field of systems biology. For a class of such models, namely, for density dependent families of Markov chains that represent the interaction of large groups of identical objects, Kurtz has proposed two kinds of approximations. One is based on ordinary differential equations and provides a deterministic approximation, while the other uses a diffusion process with which the resulting approximation is stochastic. The computational cost of the deterministic approximation is significantly lower, but the diffusion approximation retains stochasticity and is able to reproduce relevant random features like variance, bimodality, and tail behavior that cannot be captured by a single deterministic quantity.

In a recent paper, for particular stochastic Petri net models, we proposed a jump diffusion approximation that aims at being applicable beyond the limits of Kurtz's diffusion approximation in order to cover the case when the process reaches the boundary with non-negligible probability. In this paper we generalize the method so that it can be applied to any density dependent Markov chains. Other limitations of the diffusion approximation in its original form are that it can provide inaccurate results when the number of objects in some groups is often or constantly low and that it can be applied only to pure density dependent Markov chains. In order to overcome these drawbacks, in this paper we propose to apply the jump-diffusion approximation only to those components of the model that are in density dependent form and are associated with high population levels. The remaining components are treated as discrete quantities. The resulting process is a hybrid switching jump diffusion, i.e., a diffusion with hybrid state space and jumps where the discrete state changes can be seen as switches that take the diffusion from one condition to another. We show that the stochastic differential equations that characterize this process can be derived automatically both from the description of the original Markov chains or starting from a higher level description language, like stochastic Petri nets. The proposed approach is illustrated on three models: one modeling the so-called crazy clock reaction, one describing viral infection kinetics and the last considering transcription regulation.

* Corresponding author.
E-mail addresses: angius@di.unito.it (A. Angius), balbo@di.unito.it (G. Balbo), beccuti@di.unito.it (M. Beccuti), enrico.bibbona@unito.it (E. Bibbona), horvath@di.unito.it (A. Horvath), roberta.sirovich@unito.it (R. Sirovich).

http://dx.doi.org/10.1016/j.tcs.2015.03.015
0304-3975/

1. Introduction

Stochastic modeling of the dynamics of biological systems gains in importance as more and more evidence is gathered that randomness plays an important role in many of these phenomena [1,2]. In most cases, as in the pioneering works of Gillespie [3] and Kurtz [4], the stochastic process associated with the evolution of the biological system is a continuous time Markov chain (CTMC). In theory, CTMCs can be analyzed by well-established techniques [5] to characterize both their initial transient period and their long run behavior. In practice however the state space of the CTMC representation of a real phenomenon is often so large that an exact analytical treatment is not feasible.

One approach to the analysis of these models is simulation and, starting from [3], several simulation based techniques have been proposed. The main difficulty lies in the facts that because of the size of the state space many simulation runs are needed to characterize the system, and that often the interactions occur in significantly different time scales. Methods to overcome these difficulties were proposed in [6–8]. Approximate analytical techniques have also been considered. Some examples are the following. In [9] the authors propose a method that dynamically limits the state space to those states that are of non-negligible probability. Since the number of states can be huge even if not all states are considered, in [10,11] approximate randomization methods have been proposed. Another natural approach is aggregation of states which can be done either by aggregating nearby states [12,13] or by exploiting the idea of flow equivalence [14]. Techniques that are based on imposing a special dependency structure on the probabilities of the states were proposed in [15,16].

An important alternative to the above approaches, initiated mainly by Kurtz, is based on constructing a simpler process to approximate the original CTMC when it models the interaction of large groups of identical objects (which can be members of species, or populations, or proteins, or enzymes, etc.). A key concept in these works is the so-called *density dependent* property. For density dependent CTMCs, as it was shown in [4], it is possible to derive a set of ordinary differential equations (ODE) that leads to a good deterministic approximation of the CTMC when the number of interacting objects is large. A stochastic approximation of density dependent processes using diffusion processes, characterized by stochastic differential equations (SDE), was proposed instead in [17]. The ODE based approximation can be strikingly poor when the number of interacting objects is not large enough to rule out variability and when the model involves particular random phenomenon – characterized by bi-modal distributions and/or switching behaviors – that are not possible to capture with a deterministic model. In these cases the diffusion based approximation could give better approximations, although it only works up to the first visit of the boundary of the state space. A recent review on the application of these techniques to model chemical reactions is given in [18].

In [19] we proposed a jump-diffusion approximation that aims at being applicable beyond the limit of Kurtz theory. Namely, since originally the approximation was defined only up to the first time when it reaches a boundary, we added in the approximating model an explicit description of the behavior at the boundaries. When a component attains a boundary, indeed, it stays there for a while and then it jumps back into the interior mimicking the behavior of the original Markov Chain. In this way, the approach is applicable to such systems where boundaries are reached with non-negligible probability as it happens, for example, in ecological models where there are species that can become temporarily extinct.

The scope of this paper is to refine and extend the jump-diffusion approximation further so that it can be applied to get a better approximation of more general Markov Chains. The motivation of this extension is twofold. First, is the case when the size of a subgroup of objects remains constantly low so that a generalized use of the diffusion approximation to the whole system would lead to inaccurate results. Second, is the situation in which the approximation is not applicable because the CTMC underlying the model does not belong to a density dependent family, even if it has a subset of components that interact in manners which enjoy the density dependent property. Our proposal is to approximate such Markov chains with a hybrid process, where our jump-diffusion approximation is applied to those components which correspond to groups with large number of members and which interact according to the property of density dependence. This category of components is referred to as *continuous* or *fluid* components and in the more usual Piecewise Deterministic Markov Processes (PDMP) approach they would be approximated by solving a system of switching ODEs. The remaining components, that we call *discrete*, are treated according to the mechanisms of the original CTMC. The process resulting from our new approach is a hybrid switching jump diffusion (HSJD) (cf. [20]) that uses jumps to handle both the discrete components and the behavior at the boundaries of the state space.

The above described *partial fluidization* is introduced starting from CTMCs. Yet, as models are usually defined in higher level languages, we show that the approximate jump diffusion process can be derived starting from stochastic Petri nets (SPN) as well.

In some simpler cases, the proposed HSJD can be analyzed analytically by solving the Fokker–Planck partial differential equation of the process. Two such simple illustrative examples will be proposed. When the model is more complex, only simulation is feasible. We will describe a simple algorithm for the simulation and will report the results on some models in systems biology.

In the literature several flavors of hybrid models have been proposed and studied in the recent years. The PDMP class, which is strongly related to the class of HSJD processes we propose, has also been generalized in order to include the case in which the fluid component is a diffusion process. A comparison between our approach and those proposed in the literature is deferred to Section 3.2. Petri nets with hybrid state spaces were introduced in [21] and generalized in [22]. The aim of these original proposals was both to handle systems in which the number of objects tend to become exceedingly large and to model intrinsically continuous quantities (like temperatures). Processes with hybrid state space are used also as a mean