STOCHASTIC SIMULATION OF BIOCHEMICAL SYSTEMS VIA DISCRETE-TIME CONVERSION

Werner Sandmann

University of Bamberg
Feldkirchenstr. 21, D-96045 Bamberg, Germany

Abstract: A discrete-time conversion is applied to the continuous-time Markov process that describes the dynamics of biochemically reacting systems within the discrete-state stochastic modeling approach (chemical master equation approach). This yields a stochastically identical discrete-time Markov chain and an according formulation of the chemical master equation. Simulating the resulting chain is equivalent to the well-known Gillespie algorithm but requires less effort. Thus, exactness as possessed by the Gillespie algorithm is preserved while the simulation can be performed more efficiently. Numerical examples are presented to compare the Gillespie algorithm and the discrete-time conversion approach.

Keywords: Chemical reactions, Master equation, Uniformization, Discrete-Time Markov chain, Stochastic simulation

INTRODUCTION

Stochastic modeling of biochemical systems is today well-established as it has been often enough pointed out that randomness is present in many systems, see for example (McAdams and Arkin, 1997; McAdams and Arkin, 1999; Turner *et al.*, 2004; Wilkinson, 2006) to mention but a few of the publications dealing specifically with stochastic models. Stochastic simulation is in widespread use to analyze such systems that are governed by stochastic processes, both (probabilistically) exact methods (Gillespie, 1976; Gillespie, 1977; Gibson and Bruck, 2000; Cao *et al.*, 2004) and approximate methods such as the continuously improving tau-leaping simulation (Gillespie, 2001; Cao *et al.*, 2005; Cao and Petzold, 2005; Rathinam *et al.*, 2005; Cao *et al.*, 2006). Accelerating simulations is motivated by the fact that stochastic simulation is inherently costly meaning that the advantage to be able to deal with quite large models has to be paid by a significant amount of computer time. However, accelerated simulation usually comes at the prize of approximation and thus loss of exactness. All the mentioned simulation methods, exact or approximate, deal with generating trajectories of the underlying continuous-time stochastic process that describes the dynamic behavior of chemically reacting systems.

In this paper we apply a discrete-time conversion that we show to yield a discrete-time Markov chain stochastically identical and thus equivalent to the continuous-time process. This discrete-time conversion, also called *uniformization* is particularly suited for stochastic simulation. Hence, the purpose and contribution is twofold. First, the concept of discrete-time conversion is introduced and formally shown to yield an equivalent representation of the original continuous-time stochastic process. Then it is demonstrated how to make use of this for improving the exact stochastic simulation of coupled chemical reactions known as the Gillespie algorithm. The validity and efficiency improvement is empirically demonstrated by numerical examples. We start with briefly explaining

the discrete-state stochastic modeling approach, thereby introducing necessary terminology and notation.

STOCHASTIC MODELING OF BIOCHEMICAL SYSTEMS

Stochastic interpretations of chemically reacting systems can be traced back to (McQuarrie, 1967). A formulation on a physical basis has been provided in (Gillespie, 1976), (Gillespie, 1977) and later on rigorously derived in (Gillespie, 1992). The basic assumptions are that the system is well stirred and thermally equilibrated, meaning that a well stirred mixture of $d \in \mathbb{N}^+$ molecular species $S_1, \ldots, S_d$ inside some fixed volume interact at constant temperature. The system state at any time $t \geq 0$ is a d-dimensional discrete random vector $X(t) = (X_1(t), \ldots, X_d(t))$, where for each species $S_k, k \in \{1, \ldots, d\}$ and $t \geq 0$ a discrete random variable $X_k(t)$ describes the number of molecules of species S_k present at time t. The set $\mathcal{S} \subseteq \mathbb{N}^d$ of all possible system states constitutes the system's state space. The conditional transient (time dependent) probability that the system is in state $x \in \mathcal{S}$ at time t, given that the system starts in an initial state $x_0 \in \mathcal{S}$ at time t_0, is denoted by $p^{(t)}(x) := p^{(t)}(x|x_0, t_0)$, that is

$$p^{(t)}(x) = P\left(X(t) = x \mid X(t_0) = x_0\right).$$

The system state changes due to chemical reactions $R_1, \ldots, R_M$ and the reaction rate of each $R_i, i \in \{1, \ldots, M\}$ is given by the *propensity function* α_i, where $\alpha_i(x)dt$ is the conditional probability that a reaction of type R_i occurs in the infinitesimal time interval $[t, t+dt)$, given that the system is in state x at time t. That is $\alpha_i(x)dt = P\left(R_i \text{ occurs in } [t, t+dt) \mid X(t) = x\right)$. Given that the system starts in an initial state $x_0 \in \mathcal{S}$ at time t_0, the temporal evolution of the system is expressed by the *chemical master equation* (CME)

$$\frac{\partial p^{(t)}(x)}{\partial t} = \sum_{i=1}^{M} \left(\alpha_i(x-v_i)p^{(t)}(x-v_i) - \alpha_i(x)p^{(t)}(x)\right)$$

where $v_i = (v_{i1}, \ldots, v_{id})$ is a *state change vector* and $v_{ik},\ k \in \{1, \ldots, d\}$ denotes the change of molecules of species S_k due to reaction type R_i.

The reaction rates α_i are time-independent since the probability that a reaction occurs within a specific time interval only depends on the length of this interval and not on the interval endpoints. Thus, given a current system state, the next state in the system's time evolution only depends on this current system state and neither on the specific time nor on the history of reactions that led to the current state. Hence, the time evolution of the system is mathematically described by a stochastic process $(X(t))_{t\geq 0}$ with d-dimensional state space $\mathcal{S} \subseteq \mathbb{N}^d$, and due to the just stated independence of time and history this stochastic process is a discrete-state Markov process, also called Markov jump process or continuous-time Markov chain (CTMC). Terminology and notation in the theory of CTMCs, where the state space is typically assumed to be mapped to $\mathbb{N}$ and most expressions are given in vector-matrix notation and can be appropriately handled by linear algebra, is usually rather different from that used to express the CME. However, it can be shown that the CME is equivalent to the Kolmogorov forward differential equations. We will not further expose this, see for example (Bremaud, 1999; van Kampen, 1992) for more information.

Stochastic Simulation

The essential part of any simulation is to imitate the system under consideration. Consequently, stochastic simulation of coupled chemical reactions consists of generating trajectories of the underlying CTMC. With the terminology used in the derivation of the CME this has been introduced to the biochemical literature by Gillespie (Gillespie, 1976; Gillespie, 1977) and is in this context usually referred to as the stochastic simulation algorithm (SSA) or the Gillespie algorithm:

Init $t := t_0$, $x := x_0$ and t_{end}
while $t < t_{end}$

(1) Compute all $\alpha_i(x)$ and $\alpha_{_0}(x) := \sum_{i=1}^{M} \alpha_i(x)$
(2) Generate two random numbers u_1, u_2, uniformly distributed on $(0, 1)$
(3) Generate time τ to next reaction: $\tau = -\ln(u_1)/\alpha_0(x)$
(4) Determine reaction type: $i = \min\{k : \alpha_1(x) + \cdots + \alpha_k(x) > u_2\alpha_{_0}(x)\}$
(5) Set $t := t + \tau$; $x := x + v_i$
(6) Store/Collect/Handle Data

An equivalent version using a different interpretation of the CTMC dynamics is due to (Gibson and Bruck, 2000). However, though this appeared to be more efficient at a first glance, it is now well-known that this not true (Cao *et al.*, 2004).

DISCRETE-TIME CONVERSION

The basic idea of discrete-time conversion is to define an associated discrete-time stochastic process that behaves equivalent to the continuous-time process under consideration. Such an approach already appeared as early as in (Jensen, 1953) and is thus sometimes referred to as Jensen's method. Similar methods are also known as *uniformization* or *randomization* and have been used for computing transient and steady-state solutions

for Markov chains, see for example (Hordijk *et al.*, 1976; Gross and Miller, 1984; Stewart, 1994). Here we apply a discrete-time conversion to the stochastic process formulation describing the temporal evolution of chemically reacting systems. We define a discrete-time process that is stochastically identical to the original process $(X(t))_{t\geq 0}$. More specifically, the original continuous-time process is represented as a discrete-time Markov chain where the times are implicitly driven by a Poisson process. We show that this process indeed behaves equivalent to the original one governed by the CME and we also give an according discrete-time reformulation of the CME.

The only assumption is that all propensity functions are finite [1], that is $\forall x \in \mathcal{S} : \alpha_i(x) < \infty$. Thus, also for all x the sum of all propensity functions is finite, $\alpha_0(x) := \alpha_1(x) + \cdots + \alpha_M(x) < \infty$. Obviously, this assumption is trivially given for finite state spaces and therefore not a serious restriction since there should be a finite number of molecules in practical situations. Notice that in the original process $(X(t))_{t\geq 0}$, governed by the CME and usually simulated with the Gillespie algorithm, for all states $x \in \mathcal{S}$ the conditional probability of reaction type R_i given that any reaction occurs is $\frac{\alpha_i(x)}{\alpha_0(x)}$. In more general Markov process terminology these are the jump probabilities given that a transition out of state x occurs.

Define a *uniformization constant* λ such that

$$\sup_x\{\alpha_0(x)\} \leq \lambda < \infty.$$

Then

$$\forall x \in \mathcal{S} :\ 0 \leq \frac{\alpha_i(x)}{\lambda} \leq 1, \quad i \in \{0, \ldots, M\}.$$

Hence, in that way division by λ uniformizes the reaction rates which explains the term of uniformization. Now, define a discrete-time Markov chain $(Y_n)_{n\in\mathbb{N}}$ with the same initial distribution as $(X(t))_{t\geq 0}$ (that is for example the distribution assigning an initial probability of one to state x_0 at time t_0 and initial probabilities of zero to all other states) and with transition probabilities $\frac{\alpha_i(x)}{\lambda}$ for (unconditional) transitions out of state x. Since a transition out of state x means that a reaction occurs these transition probabilities shall be taken as reaction probabilities. Obviously,

$$0 \leq \sum_{i=1}^{M} \frac{\alpha_i(x)}{\lambda} = \frac{\alpha_0(x)}{\lambda} \leq 1$$

which means that there may be a probability of less than one for leaving state x, that is that any reaction occurs, and there is a remaining probability of $1 - \frac{\alpha_0(x)}{\lambda}$ for staying in state x meaning that no reaction occurs. Notice that this construction induces that the times between transitions in $(Y_n)_{n\in\mathbb{N}}$ are all exponentially distributed with the same mean $1/\lambda$. In that sense these times are randomized by a Poisson process with rate λ which explains the alternative term of randomization. To see that $(Y_n)_{n\in\mathbb{N}}$ is stochastically identical to $(X(t))_{t\geq 0}$ consider for $(Y_n)_{n\in\mathbb{N}}$ the conditional reaction probabilities given that a reaction occurs. Using as a short-hand notation R_i for "reaction of type R_i occurs" and R for "any reaction occurs" and noticing that $R_i \Rightarrow R$ these are given by

$$P(R_i|R) = \frac{P(R_i \wedge R)}{P(R)} = \frac{P(R_i)}{P(R)} = \frac{\frac{\alpha_i(x)}{\lambda}}{\frac{\alpha_0(x)}{\lambda}} = \frac{\alpha_i(x)}{\alpha_0(x)}$$

and indeed the conditional probabilities for the processes $(X(t))_{t\geq 0}$ and $(Y_n)_{n\in\mathbb{N}}$ are the same. This implies that the original process and the discrete-time Markov chain with randomized times between transitions behave equivalently and are stochastically identical. Formally, we get

$$\begin{aligned} P(X(t) = x) &= \sum_{k=0}^{\infty} P(Y_k = x \mid k \text{ reactions in [0,t]}) \\ &= \sum_{k=0}^{\infty} P(Y_k = x)\ e^{-\lambda t}\ \frac{(\lambda t)^k}{k!} \end{aligned}$$

where the latter equality is due to the fact that for the randomized chain the number of transitions in any time interval of length t has a Poisson distribution with rate λ. In principle, the above derivation also offers an alternative method for non-simulative computation of the transient probabilities $p^{(t)}(x) = P(X(t) = x)$. Instead of numerically solving the CME, one can try to approximate the above sum. However, this infinite sum must be definitely truncated and moreover, numerical difficulties and large models may cause serious problems requiring a lot of specific extra investigations that are not the topic of the present paper. Nevertheless, before we demonstrate how the above can be used for stochastic simulation, we give some more theory and evidence by an according reformulation of the CME.

Reformulation of the chemical master equation

In addition to the already introduced notation let $p_{s,s'},\ s, s' \in \mathcal{S}$ denote the probability of a transition from state s to state s' in the uniformized chain. More specifically, in order to appropriately reformulate the CME the probabilities $p_{x-v_i,x}$ for a reaction of type R_i in state $x - v_i$ (which leads to state x) and $p_{x,x}$ for staying in state x are needed. That is

$$p_{x-v_i,x} = \frac{\alpha_i(x - v_i)}{\lambda}, \quad p_{x,x} = 1 - \frac{\alpha_0(x)}{\lambda}.$$

[1] This may be even dropped in planned further investigations.

Now, the CME can be rewritten as

$$
\begin{aligned}
&\sum_{i=1}^{M}\Big(\alpha_i(x-v_i)p^{(t)}(x-v_i)-\alpha_i(x)p^{(t)}(x)\Big)\\
&=\sum_{i=1}^{M}\alpha_i(x-v_i)p^{(t)}(x-v_i)-\sum_{i=1}^{M}\alpha_i(x)p^{(t)}(x)\\
&=\sum_{i=1}^{M}\alpha_i(x-v_i)p^{(t)}(x-v_i)-p^{(t)}(x)\sum_{i=1}^{M}\alpha_i(x)\\
&=\sum_{i=1}^{M}\alpha_i(x-v_i)p^{(t)}(x-v_i)-p^{(t)}(x)\alpha_0(x)\\
&=\sum_{i=1}^{M}\alpha_i(x-v_i)p^{(t)}(x-v_i)\\
&\quad-p^{(t)}(x)(\alpha_0(x)+\lambda-\lambda)\\
&=\sum_{i=1}^{M}\alpha_i(x-v_i)p^{(t)}(x-v_i)\\
&\quad-\lambda p^{(t)}(x)+p^{(t)}(x)(\lambda-\alpha_0(x))\\
&=\sum_{i=1}^{M}\alpha_i(x-v_i)p^{(t)}(x-v_i)\\
&\quad-\lambda p^{(t)}(x)+p^{(t)}(x)\lambda(1-\frac{\alpha_0(x)}{\lambda})\\
&=\sum_{i=1}^{M}\alpha_i(x-v_i)p^{(t)}(x-v_i)\\
&\quad-\lambda p^{(t)}(x)+p^{(t)}(x)\lambda p_{x,x}\\
&=\lambda\sum_{i=1}^{M}\frac{\alpha_i(x-v_i)}{\lambda}p^{(t)}(x-v_i)\\
&\quad-\lambda p^{(t)}(x)+p^{(t)}(x)\lambda p_{x,x}\\
&=\lambda\sum_{i=1}^{M}p_{x-v_i,x}p^{(t)}(x-v_i)\\
&\quad-\lambda p^{(t)}(x)+p^{(t)}(x)\lambda p_{x,x}.
\end{aligned}
$$

Now, setting $v_0 = \underline{0}$ which expresses that if no reaction occurs the state change vector is simply the null vector, we have

$$p^{(t)}(x)\lambda p_{x,x} = p^{(t)}(x-v_0)\lambda p_{x-v_0,x}$$

and altogether the CME becomes

$$\frac{\partial p^{(t)}(x)}{\partial t} = -\lambda p^{(t)}(x)+\lambda\sum_{i=0}^{M}p_{x-v_i,x}p^{(t)}(x-v_i).$$

NEW SIMULATION ALGORITHMS

The probabilistic interpretation of the dynamics of the uniformized chain directly leads us to a very simple method for stochastic simulation of coupled chemical reactions. In the following an according algorithm is formulated. Notice that due to the derivations in the previous section this algorithm is exact in the same sense as the Gillespie algorithm but does not require generation of any exponentially distributed time.

Init $n := 0,\ x := x_0,\ t_{end},\ \lambda,\ N := \lambda t_{end}$
while $n < N$

(1) Generate random number u, uniformly distributed on $(0, 1)$
(2) Determine reaction type: $i = \min\{k : \alpha_1(x) + \cdots + \alpha_k(x) > \lambda u\}$
(3) Set $n := n + 1;\ x := x + v_i$
(4) Store/Collect/Handle Data

In the following we present three examples to demonstrate the validity of our algorithm and the efficiency improvement compared to the Gillespie algorithm. We focus on relatively small examples that are sufficient for this purpose but at the same time easy enough to verify for the reader. All algorithms have been implemented in C++.

Example 1: We start with an example that has been a reference model for testing tau-leaping methods in (Cao and Petzold, 2005), the simple model

$$S_1 \overset{c_1}{\to} S_2 \overset{c_2}{\to} S_3$$

with the initial numbers of molecules of the different species $X_1(0) = 9,\ X_2(0) = 2\cdot 10^4,\ X_3(0) = 0$ and rate constants $c_1 = 10, c_2 = 0.1$ yielding the propensity functions $\alpha_1(x) = c_1x_1 = 10x_1$ and $\alpha_2(x) = c_2x_2 = 0.1x_2$. It is easy to see that $\mathcal{S} = \{0,\ldots,9\}\times\{2\cdot 10^4+9\}$ and for all $x\in\mathcal{S}:\ \alpha_1(x)+\alpha_2(x) < 2091$. Thus, $\lambda = 2091$ is an appropriate choice for the uniformization constant. The model was simulated from time $t_0 = 0$ to time $t_{end} = 0.1$ performing 10^6 runs of both algorithms. Table 1 shows that the mean values $\bar{x}_1(0.1), \bar{x}_2(0.1), \bar{x}_3(0.1)$ for the number of molecules of each species obtained via the Gillespie algorithm and via the discrete-time conversion almost perfectly coincide and that the latter algorithm reduces the runtime.

Method	$\bar{x}_1(0.1)$	$\bar{x}_2(0.1)$	$\bar{x}_3(0.1)$	Runtime
Gillespie	3.314	19763.1	199.02	199s
Discrete	3.311	19763.1	199.04	108s

Table 1. Mean values for the number of molecules of each species at time 0.1 and runtime required for 10^6 simulation runs for the Gillespie algorithm and the discrete-time conversion method.

Example 2: Now, we extend the example to

$$S_1 \overset{c_1}{\to} S_2 \overset{c_2}{\to} S_3 \overset{c_3}{\to} S_4 \overset{c_4}{\to} S_5$$

with the same initial molecule numbers as in the previous example and two different sets of rate constants $c_1 = 10, c_2 = \cdots = c_5 = 0.1$ and $c_1 = \cdots = c_5 = 0.1$, respectively. Again, 10^6 simulation runs of both algorithms were performed, each up to time 0.1. The uniformization constant remained

unchanged. The computed mean values and the runtimes are shown in Tables 2 and 3, respectively (for obvious reasons with a rotated orientation compared to Table 1). Again, the mean values almost perfectly coincide and the discrete-time conversion is faster.

	Gillespie	Discrete-Time
$\bar{x}_1(0.1)$	3.314	3.311
$\bar{x}_2(0.1)$	19763.0	19763.1
$\bar{x}_3(0.1)$	198.04	198.05
$\bar{x}_4(0.1)$	0.992	0.989
$\bar{x}_5(0.1)$	0.00330	0.00337
Runtime	215s	113s

Table 2. Mean values and runtimes for Example 2 with choice of rate constants $c_1 = 10, c_2 = \cdots = c_5 = 0.1$.

	Gillespie	Discrete-Time
$\bar{x}_1(0.1)$	8.911	8.911
$\bar{x}_2(0.1)$	19762.3	19762.3
$\bar{x}_3(0.1)$	198.00	198.02
$\bar{x}_4(0.1)$	0.990	0.991
$\bar{x}_5(0.1)$	0.00323	0.00331
Runtime	215s	112s

Table 3. Mean values and runtimes for Example 2 with choice of rate constants $c_1 = c_2 = \cdots = c_5 = 0.1$.

So far, so good. Although the runtime is not as much reduced as reported for some cases of tau-leaping methods, as we already emphasized, the discrete-time conversion is exact. Moreover, research on the specific application to chemically reacting systems is still at quite an early stage and further improvements shall further speed up the algorithm. Before discussing that, we address an important problem that we neglected till now and give a suitable modification of the discrete-time conversion algorithm. The reader may have noticed that in case of propensity functions that differ significantly (which is in particular the case for stiff reaction sets) the uniformization constant is much larger than many of the propensity functions. Then in the uniformized chain a lot of transitions occur from a state to itself meaning that no reaction occurs and thus computer time is wasted by simulating something that is in reality simply not present. Fortunately, this can be circumvented, again by a probabilistic argument that preserves exactness. Since our uniformized chain is a discrete-time Markov chain, the difference between staying in a state (no reaction) and leaving a state (reaction) can be expressed as a Bernoulli trial where leaving a state is interpreted as a success. Hence, because the number of unsuccessful Bernoulli trials until the first success is geometrically distributed with mean (expected value) $(1-p)/p$ where p is the success probability, in our case (where success means leaving a state) the number of transitions that do not change the current state x is geometrically distributed with $p = \alpha_0(x)/\lambda$ and mean $\lambda/\alpha_0(x) - 1$. Hence, instead of simulating all the "non-reactions" we can compute the expected value and just update the number of steps in our algorithm accordingly. Then a "real reaction", that is a state transition to another state occurs with the jump probabilities $\alpha_i(x)/\alpha_0(x)$ just as in the Gillespie algorithm but again we avoid the time-consuming generation of exponentially distributed times, now via computing the number of steps in the uniformized chain that do not change the state. Hence, the modified algorithm to deal with the problems due to significantly different propensity function is given as follows

Init $n := 0,\ x := x_0,\ t_{end},\ \lambda,\ N := \lambda t_{end}$
while $n < N$

(1) Generate random number u, uniformly distributed on $(0, 1)$
(2) Compute number of "no reactions": $n_0 = \lambda/\alpha_0(x) - 1$
(3) Set $n := n + n_0$; **if** $n \geq N$ exit
(4) Determine reaction type: $i = \min\{k : \alpha_1(x) + \cdots + \alpha_k(x) > \lambda u\}$
(5) Set $n := n + 1;\ x := x + v_i$
(6) Store/Collect/Handle Data

Now we are ready to present an example of a stiff reaction set.

Example 3: As the third example we consider the decaying dimerization as a representative for very stiff systems that was also chosen for instance in (Gillespie, 2001; Rathinam *et al.*, 2005; Cao *et al.*, 2005) to study the performance of different tau-leaping methods. The reaction set is given by

$$S_1 \xrightarrow{c_1} \emptyset, \quad S_1 + S_1 \underset{c_3}{\overset{c_2}{\rightleftharpoons}} S_2, \quad S_2 \xrightarrow{c_4} S_3$$

and the parameter choices in order to achieve a high degree of stiffness according to the aforementioned literature is $c_1 = 1, c_2 = 10, c_3 = 1000, c_4 = 0.1$ for the rate constants and thus the propensity functions are $\alpha_1(x) = x_1,\ \alpha_2(x) = 5x_1(x_1 - 1),\ \alpha_3(x) = 1000x_2,\ \alpha_4(x) = 0.1x_2$. The model is simulated from time $t_0 = 0$ up to time $t_{end} = 0.2$ with the initial numbers of molecules $X_1(0) = 400,\ X_2(0) = 798,\ X_3(0) = 0$. Table 4 shows the mean values and runtimes for 10^4 simulation runs with the Gillespie algorithm and the (modified) discrete-time conversion method. Once more, the results coincide and the discrete-time conversion is faster.

Method	$\bar{x}_1(0.1)$	$\bar{x}_2(0.1)$	$\bar{x}_3(0.1)$	Runtime
Gillespie	386.96	749.55	15.55	3554s
Discrete	386.35	747.89	16.09	2527s

Table 4. Mean values and runtimes of 10^4 runs for Example 4.

CONCLUSION AND DISCUSSION

The presented discrete-time conversion yields an equivalent discrete-time Markov chain representing the original continuous-time process. Simulating the resulting uniformized chain is still exact as the Gillespie algorithm but requires less effort if appropriately applied. However, research on uniformization for simulating biochemical systems is at an early stage. For stiff systems one improvement has been already done and promising approaches are currently under investigation. The choice of the uniformization constant λ is one major factor influencing the the efficiency of uniformization. At any time, λ must be an upper bound on the current $\alpha_0(x)$ but determining a fixed λ in advance may be both difficult and inefficient for large and/or stiff models. It is thus highly desirable to determine λ automatically which can be done, for instance, by a few pre-simulations. Moreover, dynamically updating λ during each simulation run appears to be promising. In some sense, similar challenges as for tau-leaping are posed. Nevertheless, other than tau-leaping the discrete-time conversion remains exact, and moreover these challenges may be easier tackled in discrete-time than for tau-leaping. Apart from specific challenges for uniformization we emphasize the necessity to judge a method's accuracy without comparing to the Gillespie algorithm. Given that improved methods are motivated by the large amount of computer time required by the Gillespie algorithm in many models of practical interest the Gillespie algorithm will not provide reliable results in reasonable time. Thus we consider statistical output analysis such as forming confidence intervals. Since the width of confidence intervals relies on the variance of the statistical estimator (defined by the simulation method) variance reduction is highly desirable. Importance sampling as a technique with the great potential of yielding zero-variance in its optimal case has been recently shown in (Sandmann, 2007) to be applicable to the Gillespie algorithm, and applicability to the discrete-time conversion is immediate.

REFERENCES

Bremaud, P. (1999). *Markov Chains.*. Springer.

Cao, Y. and L.R. Petzold (2005). Trapezoidal tau-leaping formula for the stochastic simulation of biochemical systems. In: *Proc. FOSBE 2005*. pp. 149–152.

Cao, Y., D.T. Gillespie and L.R. Petzold (2005). Avoiding negative populations in explicit Poisson tau-leaping. *J. Chem. Phys.* **123**, 054104.

Cao, Y., D.T. Gillespie and L.R. Petzold (2006). Efficient stepsize selection for the tau-leaping simulation. *J. Chem. Phys.* **124**, 044109.

Cao, Y., H. Li and L.R. Petzold (2004). Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J. Chem. Phys.* **121**(9), 4059–4067.

Gibson, M.A. and J. Bruck (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem.* **A 104**, 1876–1889.

Gillespie, D.T. (1976). A general method for numerically simulating the time evolution of coupled chemical reactions. *J. Comp. Phys.* **22**, 403–434.

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **71**(25), 2340–2361.

Gillespie, D.T. (1992). A rigorous derivation of the chemical master equation. *Physica A* **188**, 404–425.

Gillespie, D.T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**, 1716–1732.

Gross, D. and D. Miller (1984). The randomization technique as a modeling tool and solution procedure for transient Markov processes. *Oper. Res.* **32**(2), 926–944.

Hordijk, A., D.L. Iglehart and R. Schassberger (1976). Discrete time methods for simulating continuous time Markov chains. *Adv. Appl. Prob.* **8**, 772–788.

Jensen, A. (1953). Markoff chains as an aid in the study of Markoff processes. *Skandinavisk Aktuarietidskrift* **36**, 87–91.

McAdams, H.H. and A. Arkin (1997). Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA* **94**, 814–819.

McAdams, H.H. and A. Arkin (1999). It's a noisy business!. *Trends in Genetics* **15**(2), 65–69.

McQuarrie, D.A. (1967). Stochastic approach to chemical kinetics. *J. Appl. Prob.* **4**, 413–478.

Rathinam, M., L.R. Petzold, Y. Cao and D.T. Gillespie (2005). Consistency and stability of tau leaping schemes for chemical reaction systems. *SIAM Multiscale Modeling and Simulation* **4**, 867–895.

Sandmann, W. (2007). Applicability of importance sampling to coupled molecular reactions. In: *Proc. 12th Int. Conf. Applied Stoch. Models and Data Analysis*. 8 pages, CD ed.

Stewart, W.J. (1994). *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press.

Turner, T.E., S. Schnell and K. Burrage (2004). Stochastic approaches for modelling in vivo reactions. *Comp. Bio. Chem.* **28**, 165–178.

van Kampen, N. (1992). *Stochastic Processes in Physics and Chemistry*. Elsevier.

Wilkinson, D.J. (2006). *Stochastic Modelling for Systems Biology*. Chapman & Hall.