

# Highly accurate tau-leaping methods with random corrections

Yucheng Hu<sup>a)</sup> and Tiejun Li<sup>b)</sup>*Laboratory of Mathematics and Applied Mathematics and School of Mathematical Sciences, Peking University, Beijing 100871, People's Republic of China*

(Received 20 September 2008; accepted 9 February 2009; published online 24 March 2009)

We aim to construct higher order tau-leaping methods for numerically simulating stochastic chemical kinetic systems in this paper. By adding a random correction to the primitive tau-leaping scheme in each time step, we greatly improve the accuracy of the tau-leaping approximations. This gain in accuracy actually comes from the reduction in the local truncation error of the scheme in the order of  $\tau$ , the marching time step size. While the local truncation error of the primitive tau-leaping method is  $\mathcal{O}(\tau^2)$  for all moments, our Poisson random correction tau-leaping method, in which the correction term is a Poisson random variable, can reduce the local truncation error for the mean to  $\mathcal{O}(\tau^3)$ , and both Gaussian random correction tau-leaping methods, in which the correction term is a Gaussian random variable, can reduce the local truncation error for both the mean and covariance to  $\mathcal{O}(\tau^3)$ . Numerical results demonstrate that these novel methods more accurately capture crucial properties such as the mean and variance than existing methods for simulating chemical reaction systems. This work constitutes a first step to construct high order numerical methods for simulating jump processes. With further refinement and appropriately modified step-size selection procedures, the random correction methods should provide a viable way of simulating chemical reaction systems accurately and efficiently. © 2009 American Institute of Physics. [DOI: 10.1063/1.3091269]

## I. INTRODUCTION

Traditional deterministic modeling of chemical reactions using ordinary differential equations (ODEs) may not be adequate for microscopic chemical kinetic systems such as the reaction networks in a single living cell.<sup>1–4</sup> The deterministic reaction rate equation in the ODE approach is a result of the large volume limit as the number of reacting molecules goes to infinity.<sup>5</sup> When the number of some reactant species is so small that the stochastic effect cannot be neglected, the large volume limit can be misleading and we need to return to chemical kinetics.

Consider a spatially homogeneous well-stirred chemical reaction system. The time evolution of the molecular populations of the reactant species  $X_i$  can be modeled stochastically by a discrete Markov jump process,<sup>6</sup> where the subscript means the dependence of  $X$  on the time  $t$ . The fundamental simulation algorithm for this system is Gillespie's stochastic simulation algorithm (SSA).<sup>7,8</sup> The SSA is essentially exact because it is rigorously based on the same microphysical principles that underlie the chemical master equation (CME).<sup>6</sup> Because it needs to simulate every reaction that fires in the system one at a time, it is extremely inefficient for realistic systems where reactions fire very frequently.

To speed up discrete stochastic simulation, Gillespie<sup>9</sup> proposed the tau-leaping method as an approximate simulation strategy. The basic idea of tau-leaping is as follows. First choose a preselected time step  $\tau$  that encompasses many reaction events, but not so many that the state of the system

will change substantially. Then approximate the number of reaction events during the time interval  $\tau$  corresponding to each reaction channel by a Poisson random number. Such a procedure would allow us to leap along the system's history axis from one subinterval of length  $\tau$  to the next, instead of stepping along from one reaction event to the next. This tau-leaping simulation method can greatly speed up SSA, but it sacrifices some accuracy.

To estimate the numerical error of the tau-leaping approximation, Rathinam *et al.*<sup>10</sup> performed a consistency check for the tau-leaping discretization and showed that its local truncation error (LTE) is  $\mathcal{O}(\tau^2)$  for all moments of  $X_t$ . They also proved that the tau-leaping is of first order weak accuracy for the special case of linear propensity functions. Li<sup>11</sup> extended this result to general propensity functions. Moreover, he showed that the time evolution of  $X_t$  can be described by SDE driven by Poisson type noise and the tau-leaping scheme is just a forward Euler discretization with strong order 1/2 and weak order 1 for this SDE.

Considerable work is being done to improve the accuracy of the tau-leaping method. Gillespie<sup>9</sup> originally proposed the midpoint tau-leaping (MP), which is analogous to the midpoint rule for ODEs. Burrage and Tian<sup>12</sup> proposed the Poisson Runge–Kutta method, which is essentially a reimplementation of the Runge–Kutta methods for ODEs<sup>13</sup> in SDEs driven by Poisson noise. From the numerical results reported, it looks like these methods have significantly improved in accuracy over the primitive tau-leaping. Nevertheless, apparently no rigorous error analysis based on solid mathematical ground has been performed for these methods. The community working on solving SDEs also addresses higher order methods for jump processes. However the results either apply only to the SDEs driven by constant rate

<sup>a)</sup>Electronic mail: huyec@pku.edu.cn.<sup>b)</sup>Electronic mail: tieli@pku.edu.cn.

Poisson processes<sup>14</sup> or remain an abstract formula that is not sufficient for implementation (Ref. 15, p. 986). To the authors' knowledge, all the current simulation methods for chemical reaction systems are of weak order 1.<sup>16</sup>

In this paper we propose some highly accurate methods for solving chemical reaction systems and we call them random correction tau-leaping or RC-tau-leaping methods. Our approach is to add a random correction to the Poisson random number in the primitive tau-leaping method. We will show that with carefully chosen random correction terms, second order accuracy can be achieved at least for the mean and covariance of  $X_t$ , which are the most interesting quantities for its probability distribution function (PDF). The novelty of our method lies in introducing a correlation between the primitive tau-leaping Poisson term and the random correction term, which is crucial for us to design schemes that are of high order accuracy for covariance. By taking advantage of the Taylor expansion of the CME for the probabilities and the consequent results for moments as derived by Rathinam *et al.*<sup>10</sup> and matching the LTE up to the second order in the expansion, we obtain the necessary conditions for generating random correction terms that will lead to high accuracy.

The study of LTE is the first step to analyze the accuracy of a numerical scheme. While the LTE of the primitive tau-leaping is  $\mathcal{O}(\tau^2)$  for all moments,<sup>10</sup> we prove that the Poisson-RC-tau-leaping method (PRC) can reduce LTE for the mean of  $X_t$  to  $\mathcal{O}(\tau^3)$ ; the Gaussian-RC-tau-leaping method version 1 (GRC1) can reduce LTE for both the mean and covariance to  $\mathcal{O}(\tau^3)$ . Milstein and Tretyakov<sup>17</sup> demonstrated that  $\mathcal{O}(\tau^3)$  LTE leads to second order accuracy in SDEs driven by Gaussian white noise. We speculate that this is also true for SDEs driven by Poisson noise, although this property has not yet been strictly proven. Then, from the above results for LTE, the PRC is of second order accuracy for the mean and the GRC1 is of second order accuracy for both mean and covariance. This is consistent with our numerical results. We also proposed a variation in the Gaussian-RC-tau-leaping method (GRC2) that reduces the variance of the random correction and gives more accurate results than GRC1. More research is underway to explain why the GRC2 algorithm is more accurate.

One of our treatments is still open to debate. In GRC1 and GRC2 we choose not to restrict the number of reactions and the population of reactant species to integer values. Otherwise it may happen that the second order conditions in the GRC1 or GRC2 cannot be satisfied. In the authors' opinion, although noninteger valued reaction number and population are physically unrealistic, the simulations remain meaningful in a statistical sense. In other words, the approach is justified by the result that GRC1 and GRC2 do produce a more accurate mean, covariance, and even PDF of the population of reactant species, which is exactly what we want from a pragmatic point of view. In addition, if an integer sample value is needed, one can round  $X_t$  to its nearest integer at the end of each simulation. Similar treatment is also used in some weak methods solving SDEs driven by Gaussian white noise.

Like the tau-leaping algorithm, the new methods also suffer from the negative population problems, in which the unbounded Poisson or Gaussian random variables may lead

to unphysical states with negative species populations. A lot of work addresses this problem for the tau-leaping methods: Chatterjee *et al.*<sup>18</sup> proposed the binomial tau-leaping method; Gillespie and co-workers<sup>19–21</sup> developed more robust and efficient leap-size selection strategies. These ideas can also be applied to our new methods. The new methods will be more useful if they can be used to solve stiff systems which usually involve several widely varying time scales. Stiffness is the norm rather than the exception in chemical reaction systems and a lot of work addresses the issue. Relevant papers include Refs. 22–26. Presumably we could incorporate these strategies into our methods to improve accuracy and handle stiffness simultaneously. Systematic study of these possibilities will be a future research topic.

The rest of this paper is organized as follows. In Sec. II, we will briefly review the SSA, tau-leaping, and MP. In Sec. III, we present some theoretical results and then propose our PRC, GRC1, and GRC2. In Sec. IV, we report the numerical results of different chemical reaction systems. Section V presents our conclusions. Proofs for the most important results in this paper can be found in the Appendixes.

## II. BACKGROUND

Assume that a well-stirred chemical reaction system has  $N$  chemical species  $\{S_1, \dots, S_N\}$  interacting through  $M$  reaction channels  $\{R_1, \dots, R_M\}$ . The state of the system is specified by the vector  $X_t = (X_{1t}, \dots, X_{Nt})^T$ , where  $X_{it}$  denotes the number of molecules of the species  $S_i$  at time  $t$ . Each reaction  $R_j$  is completely characterized by a non-negative propensity function  $a_j(\mathbf{x})$  and a state-change vector  $\mathbf{v}_j = (\nu_{1j}, \dots, \nu_{Nj})^T (j=1, \dots, M)$ . Denote the vector for propensity functions as  $\mathbf{a}(\mathbf{x}) = (a_1(\mathbf{x}), \dots, a_M(\mathbf{x}))^T$ , and the stoichiometry matrix as  $\mathbf{v} = (\mathbf{v}_1, \dots, \mathbf{v}_M)$ . Rules governing the stochastic evolution of such a system are as follows.

- (1) Given the current state  $X_t = \mathbf{x}$ , the reaction  $R_j$  will fire with probability  $a_j(\mathbf{x})dt$  during an infinitesimal time interval  $dt$ , and the reactions are independent of each other.
- (2) If  $R_j$  fires, then the state of the system is updated as  $X_t + \mathbf{v}_j$ .

An exact simulation method for the above chemical reaction system is the SSA, which simulates each chemical reaction event, one at a time.<sup>7,8</sup> It is essentially composed of the following three steps:

### Algorithm 1. Stochastic simulation algorithm (SSA).

- Step 1: Sample the waiting time  $\tau$  as an exponentially distributed random variable with rate  $a_0(X_t) = \sum_{j=1}^M a_j(X_t)$ .
- Step 2: Sample an  $M$  point random variable  $k$  with probability  $a_j(X_t)/a_0(X_t)$  for the  $j$ th reaction.
- Step 3: Update  $X_{t+\tau} = X_t + \mathbf{v}_k$ .

Although SSA is exact and simple to implement, its handling of individual reaction events renders it prohibitively slow for realistic simulations where reactions fire very fre-

quently. An approximate but much faster simulation procedure is the tau-leaping algorithm.<sup>9</sup> By dividing the system's history axis into a set of contiguous subintervals in which the number of firings for each reaction channel is approximated by Poisson random variable, the tau-leaping algorithm can leap from one subinterval to the next rather than requiring an updating of propensity functions at every reaction event.

**Algorithm 2.** Tau-leaping algorithm (tauleap).

- Step 1: Given the state  $\mathbf{X}_n$  at time  $t_n$ , determine a leap time  $\tau$ .
- Step 2: Generate  $\mathbf{r}=(r_1, \dots, r_M)^T$ , where  $r_j = \mathcal{P}_j(a_j(\mathbf{X}_n)\tau)$  are Poisson random variables with rate  $a_j(\mathbf{X}_n)\tau$ .
- Step 3: Update time to  $t_n + \tau$  and  $\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{v} \cdot \mathbf{r}$ .

**Remark 1.** In the tau-leaping algorithm, the adaptive selection of the time step size  $\tau$  is very important for the overall performance of the algorithm. There are different strategies to choose a reasonably large  $\tau$  without introducing large errors. Readers may refer to Refs. 9 and 19–21 for more details.

Essentially, the tau-leaping scheme is the forward Euler discretization for SDEs driven by the state-dependent Poisson processes.<sup>11</sup> In a mathematically nonrigorous form, the SDEs describing the chemical reaction system satisfy

$$d\mathbf{X}_t = \sum_{j=1}^M \mathbf{v}_j \mathcal{P}_j(a_j(\mathbf{X}_{t-})dt) \quad (1)$$

in the infinitesimal sense, where  $\mathcal{P}_j(\lambda_j)$  are independent Poisson random variables with parameter  $\lambda_j$ , and  $\mathbf{X}_{t-} = \lim_{s \rightarrow t, s < t} \mathbf{X}_s$ . A rigorously justified formulation may be found in Refs. 27 and 28. Integration of Eq. (1) from  $t$  to  $t + \tau$  gives

$$\mathbf{X}_{t+\tau} = \mathbf{X}_t + \sum_{j=1}^M \mathbf{v}_j \int_t^{t+\tau} \mathcal{P}_j(a_j(\mathbf{X}_{s-})ds), \quad (2)$$

where  $\tau > 0$  is a small time increment. Generally, the integrals  $\int_t^{t+\tau} \mathcal{P}_j(a_j(\mathbf{X}_{s-})ds)$  cannot be calculated exactly. In the tau-leaping approximation, the propensity functions  $a_j(\mathbf{X}_s)$  are treated as constants in the time interval  $[t, t + \tau)$ . Under this approximation, the integrals in Eq. (2) become

$$\int_t^{t+\tau} \mathcal{P}_j(a_j(\mathbf{X}_{s-})ds) = \mathcal{P}_j(a_j(\mathbf{X}_t)\tau), \quad (3)$$

which explains step 2 in Algorithm 2 and this procedure is analogous to the explicit Euler method for ODEs.

When applying tau-leaping algorithm to simulate Eq. (1), it will introduce some systematic error, i.e., errors that cannot be reduced by increasing the sample size. It has been proven by Rathinam *et al.*<sup>10</sup> and Li<sup>11</sup> that the tau-leaping method is of weak order 1. Considerable work is being done to improve the accuracy of the tau-leaping method. Gillespie<sup>9</sup> originally proposed the MP. The basic implementation of this algorithm is as follows.

**Algorithm 3.** Midpoint-tau-leaping (MP).

- Step 1: Given the state  $\mathbf{X}_n$  at time  $t_n$ , determine a leap time  $\tau$ .
- Step 2: Compute  $\mathbf{X}^{(1)} = \mathbf{X}_n + \lfloor \frac{1}{2} \mathbf{v} \cdot \mathbf{a}(\mathbf{X}_n) \tau \rfloor$ . Here  $\lfloor \cdot \rfloor$  means round off to the nearest integer.
- Step 3: Generate  $\mathbf{r}=(r_1, \dots, r_M)^T$ , where  $r_j = \mathcal{P}_j(a_j(\mathbf{X}^{(1)})\tau)$ .
- Step 4: Update time to  $t_n + \tau$  and  $\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{v} \cdot \mathbf{r}$ .

The idea behind this algorithm is similar to the midpoint rule for ODEs: in the time interval  $[t_n, t_n + \tau)$ , it takes the estimated value of  $\mathbf{X}$  at  $t_n + \tau/2$  instead of freezing the value of  $\mathbf{X}$  at  $t_n$ . The MP can significantly improve accuracy in many cases. However we will show later that the LTE of MP for the covariance of  $\mathbf{X}$  cannot be  $\mathcal{O}(\tau^2)$ . So at least for covariance MP does not have second order accuracy.

### III. RANDOM CORRECTION TAU-LEAPING METHODS

In this section, some highly accurate methods for simulating chemical reaction systems will be proposed. More precisely, the PRC method is believed to be of second order accuracy only for the mean; the GRC1 and GRC2 methods are believed to be of second order accuracy for both the mean and covariance. First let us state some mathematical definitions and propositions that will help to clarify the rationale behind the new methods.

In designing a numerical scheme for simulating Eq. (1), the goal is to find approximations of the integrals  $r_j^* \approx \int_t^{t+\tau} \mathcal{P}_j(a_j(\mathbf{X}_s)ds)$ ,  $j=1, \dots, M$ . Then Eq. (2) transforms into

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{v} \cdot \mathbf{r}^*, \quad (4)$$

where  $\mathbf{r}^*=(r_1^*, \dots, r_M^*)^T$ . All the numerical schemes in the paper have this form. An analysis of its LTE is the first step for rationalizing the construction of a numerical scheme.

Following Rathinam *et al.*,<sup>10</sup> we define the  $q$ th order weak consistency of the  $p$ th moment of  $\mathbf{X}$  for the numerical scheme (4) as follows.

**Definition 1 (weak consistency).** Let  $\mathbf{X}_{t_n}$  and  $\mathbf{X}_{t_n+\tau}$  be the exact solution of Eq. (1) at time  $t_n$  and  $t_n + \tau$ , respectively;  $\mathbf{X}_n$  and  $\mathbf{X}_{n+1}$  be the simulation of Eq. (1) using numerical scheme (4) at time  $t_n$  and  $t_n + \tau$ , respectively. The  $p$ th moments of the increment  $\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n}$  and  $\mathbf{X}_{n+1} - \mathbf{X}_n$  are  $\mathbb{E}_x[(\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n})^p]$  and  $\mathbb{E}_x[(\mathbf{X}_{n+1} - \mathbf{X}_n)^p]$ , respectively. Here  $\mathbb{E}_x[\cdot]$  denotes the expectation conditioned on  $\mathbf{X}_t = \mathbf{X}_n = \mathbf{x}$  and  $(\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n})^p$  is the shortcut for the  $p$ -fold tensor product. We say that the numerical scheme (4) is weakly consistent for the  $p$ th moment to  $q$ th order, or the LTE of the scheme for the  $p$ th moment is  $\mathcal{O}(\tau^{q+1})$  if there exist  $C > 0$  and  $\delta > 0$  such that  $\forall \tau \in [0, \delta]$ ,

$$\|\mathbb{E}_x[(\mathbf{X}_{n+1} - \mathbf{X}_n)^p] - \mathbb{E}_x[(\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n})^p]\| \leq C\tau^{q+1}. \quad (5)$$

Here the norm  $\|\cdot\|$  can be any suitable norm such as the induced 2-norm for tensors.

**Remark 2.** Straightforwardly, we can define  $q$ th order consistency for the covariance as

$$\|\text{Cov}_x[\mathbf{X}_{n+1} - \mathbf{X}_n] - \text{Cov}_x[\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n}]\| \leq C\tau^{q+1}.$$

Since we have  $\text{Cov}[\mathbf{X}] = \mathbb{E}[\mathbf{X}^2] - (\mathbb{E}[\mathbf{X}])^2$ , if the scheme is  $q$ th order consistent for both the mean and second moment, then it is  $q$ th order consistent for both the mean and covariance. The contrary is also true.

**Remark 3.** The definition of weak consistency and LTE is similar to that in ODEs, where we have the relation: consistency plus stability implies convergence. For SDEs driven by Poisson noise, this relation has not yet been proven with a rigorous mathematical approach, but we speculate that it is true. This would mean if a numerical scheme is stable and  $q$ th order consistent (or with  $\mathcal{O}(\tau^{q+1})$  LTE), then it is of  $q$ th order accuracy. Our numerical results are consistent with this speculation.

Rathinam, *et al.*<sup>10</sup> Taylor expanded the exact  $p$ th moment  $\mathbb{E}_x[(\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n})^p]$  to the second order of  $\tau$  [see Eq. (A1) in Appendix A or Eqs. (3.5) and (3.6) in Ref. 10]. By substituting the tau-leaping approximation  $r_j^* = \mathcal{P}_j(a_j(\mathbf{X}_n)\tau)$  in Eq. (4), they showed that the numerical  $p$ th moment  $\mathbb{E}_x[(\mathbf{X}_{n+1} - \mathbf{X}_n)^p]$  matches the first order terms in the exact expansion for all  $p$ . So the LTE of the tau-leaping method for all moments is  $\mathcal{O}(\tau^2)$ , indicating that the tau-leaping scheme is weakly first order consistent.

To develop weakly higher order accurate methods, naturally one asks its LTE to be  $\mathcal{O}(\tau^3)$  for all moments. Actually we do not know how to accomplish this. So instead, we try to construct numerical schemes that have higher order LTE for the most important parameters of a PDF. More precisely, we present in Sec. III A the Poisson random correction tau-leaping method, whose LTE is  $\mathcal{O}(\tau^3)$  at least for the mean; and in Sec. III B the Gaussian random correction tau-leaping method, whose LTE is  $\mathcal{O}(\tau^3)$  at least for both the mean and covariance. Our approach is to add a random correction  $\tilde{\mathbf{r}}$  to the primitive tau-leaping term  $\mathbf{r}$  so  $\mathbf{r}^* = \mathbf{r} + \tilde{\mathbf{r}}$ . More importantly,  $\tilde{\mathbf{r}}$  is generated conditioning on  $\mathbf{r}$ . We find that the new methods are much more accurate than the primitive tau-leaping algorithm.

## A. Poisson random correction tau-leaping method

Define the  $M \times M$  matrix function  $\boldsymbol{\eta}(\mathbf{x}) = (\eta_{jk}(\mathbf{x}))$  as

$$\eta_{jk}(\mathbf{x}) = a_j(\mathbf{x} + \mathbf{v}_k) - a_j(\mathbf{x}). \quad (6)$$

In what follows, we will denote  $a_j(\mathbf{x})$  as  $a_j$  and  $\boldsymbol{\eta}(\mathbf{x})$  as  $\boldsymbol{\eta}$  when the context is clear. The following proposition provides a guideline to design numerical schemes that are second order consistent for the mean.

**Proposition 1.** Assume that we have a numerical scheme  $\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{v} \cdot \mathbf{r}^*$ , where  $\mathbf{r}^* = \mathbf{r} + \tilde{\mathbf{r}}$ .  $\mathbf{r}$  is a vector with  $M$  mutually independent components  $r_j = \mathcal{P}_j(a_j(\mathbf{X}_n)\tau)$ ,  $j = 1, \dots, M$ . Given  $\mathbf{X}_n = \mathbf{x}$ , if the components of  $\tilde{\mathbf{r}}$  satisfy

$$\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]] = \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{jk} + \mathcal{O}(\tau^3), \quad j = 1, \dots, M, \quad (7)$$

then the scheme is of second order consistency for the mean.

**Remark 4.** Note here and in what follows the conditional expectation of  $\tilde{r}_j$  given the initial  $\mathbf{X}_n = \mathbf{x}$  can be represented

with  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]]$  since  $\tilde{r}_j$  is generated conditioning on  $\mathbf{r}$ , which is conditioning on  $\mathbf{x}$ .

**Remark 5.** In proving all of the propositions in this paper, it will be helpful to keep in mind that the mean and covariance properties of the independent Poisson random variables,

$$\mathbb{E}\mathcal{P}_i(\lambda_i) = \lambda_i, \quad \mathbb{E}(\mathcal{P}_i(\lambda_i)\mathcal{P}_j(\lambda_j)) = \lambda_i\lambda_j + \delta_{ij}\lambda_i.$$

*Proof.* Please see Appendix B.  $\square$

This proposition shows that to achieve second order consistency for the mean, the random correction  $\tilde{\mathbf{r}}$  only needs to satisfy Eq. (7). In fact, there are many such  $\tilde{\mathbf{r}}$  to make Eq. (7) hold. For example, it can be easily verified that  $\tilde{r}_j = \text{sgn}(\lambda_j)\mathcal{P}_j(|\lambda_j|)$  with the following choices of  $\lambda_j$  are acceptable.

- (1)  $\lambda_j = (\tau/2) \sum_{k=1}^M r_k \eta_{jk}$ .
- (2)  $\lambda_j = (1/2) \sum_{k=1}^M (r_k(r_k-1)/a_k) \eta_{jk}$  assuming  $a_k \neq 0$  here.
- (3)  $\lambda_j = (\tau^2/2) \sum_{k=1}^M a_k \eta_{jk}$ .

Note that  $\tilde{\mathbf{r}}$  and  $\mathbf{r}$  are mutually dependent in the first two choices, while they are mutually independent in the last one. In fact, the independent property in the last case allows us to generate  $r_j^* = r_j + \tilde{r}_j$  as a Poisson random variable in one pass, that is,  $r_j^* = \mathcal{P}_j(\mu_j)$ , where

$$\mu_j = a_j\tau + \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{jk}. \quad (8)$$

Thus we have the following numerical scheme:

**Algorithm 4.** Poisson random correction tau-leaping (PRC).

- Step 1: Given the state  $\mathbf{X}_n$  at time  $t_n$ , compute the matrix  $\boldsymbol{\eta}(\mathbf{X}_n)$  and determine a leap time  $\tau$ .
- Step 2: Generate Poisson random variables  $r_j^* = \mathcal{P}_j(\mu_j)$ , where  $\mu_j$  is defined in Eq. (8).
- Step 3: Update time to  $t_n + \tau$  and  $\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{v} \cdot \mathbf{r}^*$ .

**Remark 6.** In step 2 of the above algorithm, we actually assume  $\mu_j \geq 0$  in  $r_j^* = \mathcal{P}_j(\mu_j)$ . Theoretically it is possible that  $\mu_j < 0$  in a system with multiple reactions if  $\lambda_j < 0$  and  $\tau$  is too large. In fact,  $\mu_j < 0$  implies  $|\tilde{r}_j| > r_j$ , i.e., the correction terms dominate the primitive tau-leaping terms, which should never happen because intuitively the correction  $\tilde{r}_j$  should be small compared with  $r_j$ . This problem can be avoided with properly selected  $\tau$ . More precisely, we could impose  $\mu_j > 0$  as a necessary condition to limit the time step size  $\tau$  so that

$$\tau \leq \min_{j, \lambda_j < 0} \left\{ -\frac{2a_j}{\sum_{k=1}^M a_k \eta_{jk}} \right\}.$$

If this admissible  $\tau$  is too small we may switch to SSA. In our numerical tests we use fixed  $\tau$  to test the accuracy of the methods. We find that even for moderate valued  $\tau$  the case  $\mu_j < 0$  seldom occur. If it does happen, for simplicity we just set  $\mu_j = 0$ , whose effect on the sample value can be neglected.

It is easy to check that the LTE of PRC for the mean is reduced to  $\mathcal{O}(\tau^3)$ . For other moments, the LTE is still  $\mathcal{O}(\tau^2)$ . Let us take a closer look at Eqs. (6) and (8). If we take the approximation

$$\eta_{jk}(\mathbf{x}) = a_j(\mathbf{x} + \mathbf{v}_k) - a_j(\mathbf{x}) \approx \sum_{l=1}^N \frac{\partial a_j(\mathbf{x})}{\partial x_l} v_{lk} \quad (9)$$

without considering that the magnitude of  $\mathbf{v}_k$  may not be small, substitute it into Eq. (8), we obtain

$$\mu_j \approx \tau \left( a_j + \sum_{l=1}^N \frac{\partial a_j}{\partial x_l} \sum_{k=1}^M \frac{\tau}{2} a_k v_{lk} \right) \equiv \tau a_j^*.$$

Note that the physical meaning of the term  $\sum_{k=1}^M (\tau/2) a_k v_{lk}$  is the expectation of the amount of change in  $X_l$  in a time length of  $\tau/2$ , assuming the reaction rates are frozen at  $\mathbf{a}(\mathbf{x})$ . So  $a_j^*$  is just the estimated reaction rate at the midpoint of time. In this sense, PRC is another type of MP. While the primitive MP tries to estimate the midpoint value of  $\mathbf{X}$ , the PRC tries to estimate the midpoint value of the reaction rate  $\mathbf{a}^*$ .

The computational effort for these two methods is almost the same. Unlike the MP, where the estimated midpoint value  $\mathbf{X}^{(1)}$  needs to be converted to integers in step 2 of Algorithm 3, we can directly use real-valued estimated reaction rate  $\mathbf{a}^*$ . What is more important, we can prove that the LTE of PRC for the mean is  $\mathcal{O}(\tau^3)$ , while there is no such result for MP.

## B. Gaussian random correction tau-leaping method

Now we construct methods with LTE to be  $\mathcal{O}(\tau^3)$  for both the mean and covariance. This is not so easy because for systems with multiple reactions the reactions are coupled together; hence, intuitively, the number of firings for each reaction channel during the time interval  $[t_n, t_n + \tau)$  should be correlated. In order to achieve higher order accuracy, the components of  $\mathbf{r}^*$  in Eq. (4) should be mutually dependent. Otherwise, as the following proposition shows, in general, a numerical method cannot be second order consistent for the covariance.

**Proposition 2.** For a numerical method solving Eq. (1) with the form (4), if the random variables  $\{r_j^*\}$  are mutually independent, then, in general, it cannot be second order consistent for the covariance of  $\mathbf{X}$ .

*Proof.* Please see Appendix C.  $\square$

It is easy to check that the  $\{r_j^*\}$  in MP are mutually independent, so as we mentioned earlier that MP cannot be second order consistent for the covariance. To introduce correlations to the components of  $\mathbf{r}^*$ , our strategy is to generate  $\tilde{\mathbf{r}}$  conditioning on  $\mathbf{r}$ , while still keeping the components in  $\tilde{\mathbf{r}}$  to be mutually independent. The next proposition gives a sufficient condition of generating such  $\tilde{\mathbf{r}}$  so that Eq. (4) satisfies second order consistency for both the mean and covariance.

**Proposition 3.** Assume that we have a numerical scheme  $\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{v} \cdot \mathbf{r}^*$ , where  $\mathbf{r}^* = \mathbf{r} + \tilde{\mathbf{r}}$ .  $\mathbf{r}$  is a vector with  $M$  mutually independent components  $r_j = \mathcal{P}_j(a_j(\mathbf{X}_n)\tau)$ ,  $j = 1, \dots, M$ . Given  $\mathbf{X}_n = \mathbf{x}$ , if the components of  $\tilde{\mathbf{r}}$  satisfy

- (1)  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]] = (\tau^2/2) \sum_{k=1}^M a_k \eta_{jk} + \mathcal{O}(\tau^3)$ ;
- (2) for  $j \neq k$ ,  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j \tilde{r}_k]] = \mathcal{O}(\tau^3)$ ;
- (3) for  $j \neq k$ ,  $\mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_k]] = (\tau^2/2) a_j \eta_{kj} + \mathcal{O}(\tau^3)$ ; and
- (4)  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j^2]] + 2\mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_j]] = (\tau^2/2) \sum_{k=1}^M a_k \eta_{jk} + \tau^2 a_j \eta_{jj} + \mathcal{O}(\tau^3)$ ,

then the scheme is second order consistent for both the mean and covariance.

*Proof.* Please see Appendix D.  $\square$

Based on this proposition, we now propose a scheme for choosing  $\tilde{\mathbf{r}}$  so that it satisfies all four conditions above. Given  $\mathbf{r}$  as stated in the proposition, let  $\tilde{\mathbf{r}}$  be a vector with  $M$  mutually independent components  $\tilde{r}_j$  with the mean and variance conditioning on  $\mathbf{r}$  to be, respectively,

$$\mathbb{E}_r[\tilde{r}_j] = \frac{\tau}{2} \sum_{k=1}^M r_k \eta_{jk} + \frac{\tau}{2} \sum_{\eta_{jk} < 0} \left( \frac{a_k}{a_j} r_j - \tau a_k \right) \eta_{jk}, \quad (10)$$

where  $\sum_{\eta_{jk} < 0}$  means the summation with respect to such  $k$  that  $\eta_{jk} < 0$ , and

$$\text{Var}_r[\tilde{r}_j] = \frac{\tau^2}{2} \sum_{k=1}^M a_k |\eta_{jk}| \geq 0. \quad (11)$$

It is demonstrated in Appendix E that such an  $\tilde{\mathbf{r}}$  satisfies all four conditions in Proposition 3.

**Remark 7.** For consistency it is required that if  $a_j(\mathbf{x}) = 0$  then  $\eta_{jk} \geq 0$  in Eq. (10). This is true since when  $a_j(\mathbf{x}) = 0$ , we have  $\eta_{jk} = a_j(\mathbf{x} + \mathbf{v}_k) - a_j(\mathbf{x}) = a_j(\mathbf{x} + \mathbf{v}_k) \geq 0$  from Eq. (6).

**Remark 8.** Now given the mean and variance as in Eqs. (10) and (11), we want to generate an appropriate random variable  $\tilde{r}_j$ . The physical meaning of  $r_j^*$  would demand an integer-valued  $\tilde{r}_j$ . However, by our approach,  $\tilde{r}_j$  is generated conditioning on  $r_j$  as a compensator to the primitive tau-leaping terms. It may happen that the mean and variance required for  $\tilde{r}_j$  are both very small such that no integer-valued random variable could satisfy these conditions. For example, if a random variable  $X$  has mean  $1/2$  and second moment  $1/3 > 1/4$ , then it cannot be integer valued since it must satisfy  $\mathbb{E}X^2 \geq \mathbb{E}X$ . So in our implementation,  $\tilde{r}_j$  is simply generated as a real-valued Gaussian random variable, which also leads  $\mathbf{r}^*$  and  $\mathbf{X}$  to be real valued. This treatment is still open to debate. We think it is acceptable in simulations because stochastic trajectories should be understood in a statistical sense. One trajectory cannot really describe the evolution of a system. The new methods can give a more accurate mean, covariance, and even PDF, and that is exactly what we want. Finally, in any case, we can enforce the positive integer condition for  $\mathbf{X}$  just by rounding them after the simulations.

Now we have the following GRC1 method, which is weakly second order consistent for both the mean and covariance.

**Algorithm 5.** Gaussian random correction tau-leaping version 1 (GRC1).

- Step 1: Given the state  $\mathbf{X}_n$  at time  $t_n$ , compute the matrix  $\boldsymbol{\eta}(\mathbf{X}_n)$ , and determine a leap time  $\tau$ .

- Step 2: Generate the random vector  $\mathbf{r}$  whose components are mutually independent Poisson random variables  $r_j = \mathcal{P}_j(a_j(\mathbf{X}_n)\tau)$ .
- Step 3: Generate random vector  $\tilde{\mathbf{r}}$  conditioning on  $\mathbf{r}$ , whose components are mutually independent Gaussian random variables with mean  $\mathbb{E}_r[\tilde{r}_j]$  and variance  $\text{Var}_r[\tilde{r}_j]$  as in Eqs. (10) and (11), respectively.
- Step 4: Update time to  $t_n + \tau$  and  $\mathbf{X}_{n+1} = \mathbf{X}_n + \boldsymbol{\nu} \cdot (\mathbf{r} + \tilde{\mathbf{r}})$ .

During the implementation of this algorithm, we identified a modification that makes it much more accurate for chemical reaction systems with multiple reactions. Note that the second term on the right hand side of Eq. (10) makes no contribution to  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]]$ ,  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j\tilde{r}_k]]$ , and  $\mathbb{E}_x[r_j\mathbb{E}_r[\tilde{r}_j]]$  up to  $\mathcal{O}(\tau^2)$ . The function of this term is to ensure that  $\text{Var}_r[\tilde{r}_j] \geq 0$ . Without this term  $\text{Var}_r[\tilde{r}_j]$  becomes  $\tau^2 \sum_{k=1}^M a_k \eta_{jk} / 2$ , which may be negative! Actually, it is easy to check that each term like  $\tau(a_k r_j / a_j - \tau a_k) \eta_{jk} / 2$  contributes to  $\text{Var}_r[\tilde{r}_j]$  a positive part  $-\tau^2 a_k \eta_{jk}$  if  $\eta_{jk} < 0$ . So in Eq. (11) we have  $\text{Var}_r[\tilde{r}_j] = (\tau^2 / 2) \sum_{k=1}^M a_k |\eta_{jk}|$ , which is definitely non-negative.

Since we only need  $\text{Var}_r[\tilde{r}_j] \geq 0$ , not as much as the  $(\tau^2 / 2) \sum_{k=1}^M a_k |\eta_{jk}|$ , a modified algorithm is to first compute

$$\mathbb{E}_r[\tilde{r}_j] = \frac{\tau}{2} \sum_{k=1}^M r_k \eta_{jk}, \quad (12)$$

$$\text{Var}_r[\tilde{r}_j] = \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{jk}. \quad (13)$$

If  $\text{Var}_r[\tilde{r}_j] \geq 0$ , we directly generate  $\tilde{r}_j$  with this mean and variance. Otherwise, we select  $k$  such that  $\eta_{jk} < 0$ , add the term  $\tau(a_k r_j / a_j - \tau a_k) \eta_{jk} / 2$  to the right hand side of Eq. (12), add  $-\tau^2 a_k \eta_{jk}$  to the right hand side of Eq. (13), and then check if  $\text{Var}_r[\tilde{r}_j] \geq 0$ . We iterate in  $k$  until a non-negative  $\text{Var}_r[\tilde{r}_j]$  is obtained. Such an iteration will definitely end because in the worst case we would get  $\text{Var}_r[\tilde{r}_j] = (\tau^2 / 2) \sum_{k=1}^M a_k |\eta_{jk}| \geq 0$ , which is the same as the GRC1. The  $\tilde{r}_j$  thus obtained in the modified approach may have a smaller variance yet all four conditions in Proposition 3 still hold. We name the modified algorithm Gaussian random correction tau-leaping version 2 (GRC2).

**Algorithm 6.** Gaussian random correction tau-leaping version 2 (GRC2).

- Step 1: Given the state  $\mathbf{X}_n$  at time  $t_n$ , compute the matrix  $\boldsymbol{\eta}(\mathbf{X}_n)$ , and determine a leap time  $\tau$ .
- Step 2: Generate the random vector  $\mathbf{r}$  whose components are mutually independent Poisson random variables  $r_j = \mathcal{P}_j(a_j(\mathbf{X}_n)\tau)$ .
- Step 3: Calculate

$$\mathbb{E}_r[\tilde{r}_j] = \frac{\tau}{2} \sum_{k=1}^M r_k \eta_{jk},$$

$$\text{Var}_r[\tilde{r}_j] = \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{jk}.$$

TABLE I. Comparison of numerical schemes simulating the chemical reaction system (1). The second (third) column compares the order of consistency for the mean (covariance). The last column compares the number of random variables needed in each time step, indicating the computational effort needed by each method. Here  $M$  is the number of reactions in the system.

Scheme	Mean	Covariance	No. of random variables
Tauleap	First	First	$M$
MP	Not clear	Not clear	$M$
PRC	Second	First	$M$
GRC1	Second	Second	$2M$
GRC2	Second	Second	$2M$

- Step 4: For  $k=1$  to  $M$  such that  $\eta_{jk} < 0$ ,
  - (1) if  $\text{Var}_r[\tilde{r}_j] \geq 0$ , go to step 5;
  - (2) otherwise

$$\mathbb{E}_r[\tilde{r}_j] := \mathbb{E}_r[\tilde{r}_j] + \frac{\tau}{2} \left( \frac{a_k}{a_j} r_j - \tau a_k \right),$$

$$\text{Var}_r[\tilde{r}_j] := \text{Var}_r[\tilde{r}_j] - \tau^2 a_k \eta_{jk}.$$

- Step 5: Generate the random vector  $\tilde{\mathbf{r}}$ , whose components are mutually independent Gaussian random variables with the mean  $\mathbb{E}_r[\tilde{r}_j]$  and variance  $\text{Var}_r[\tilde{r}_j]$  given above.
- Step 6: Update time to  $t_n + \tau$  and  $\mathbf{X}_{n+1} = \mathbf{X}_n + \boldsymbol{\nu} \cdot (\mathbf{r} + \tilde{\mathbf{r}})$ .

For the convenience of the reader, we list some important features of the tau-leaping, MP, PRC, GRC1, and GRC2 in Table I.

## C. Implementation issues

### 1. Leap-size selection

So far we have not discussed how to determine the time step size  $\tau$ . For realistic systems, using an adaptive leap-size selection procedure can greatly improve the efficiency. There are several leap-size selection procedures for the tau-leaping method. Here we briefly review two of them that will be tested with our new methods.

Define  $a_0 = \sum_{j=1}^M a_j$ , the  $M \times M$  matrix  $\mathbf{f} = (f_{jk})$ , where

$$f_{jk} \equiv \sum_{l=1}^N \frac{\partial a_j}{\partial x_l} \nu_{lk}, \quad (14)$$

and the vector  $\boldsymbol{\mu}^{(1)} = (\mu_1^{(1)}, \dots, \mu_M^{(1)})^T$  and  $\boldsymbol{\mu}^{(2)} = (\mu_1^{(2)}, \dots, \mu_M^{(2)})^T$ , where

$$\mu_j^{(1)} \equiv \sum_{k=1}^M f_{jk} a_k, \quad (15)$$

$$\mu_j^{(2)} \equiv \sum_{k=1}^M (f_{jk})^2 a_k. \quad (16)$$

Gillespie<sup>9</sup> originally proposed a leap-size selection procedure which is to take

$$\tau = \min \left\{ \frac{\epsilon a_0}{\mu_1^{(1)}}, \dots, \frac{\epsilon a_0}{\mu_M^{(1)}} \right\} \quad (17)$$

for a given parameter  $\epsilon \ll 1$ . Gillespie and Petzold<sup>19</sup> proposed an improved version in which

$$\tau = \min \left\{ \frac{\epsilon a_0}{\mu_1^{(1)}}, \frac{\epsilon^2 a_0^2}{\mu_1^{(2)}}, \dots, \frac{\epsilon a_0}{\mu_M^{(1)}}, \frac{\epsilon^2 a_0^2}{\mu_M^{(2)}} \right\}. \quad (18)$$

The difference between the two procedures is that only the change in the mean of  $\mathbf{a}$  is considered in the first, while the changes in both the mean and variance of  $\mathbf{a}$  are considered in the second. Cao *et al.*<sup>21</sup> proposed a more robust version, but since in this work our major interest is the scheme itself, not the step size selection, we will not incorporate this in our code.

From Eqs. (9) and (14) we can see that the matrix  $\boldsymbol{\eta}$  is a discrete approximation of the matrix  $\mathbf{f}$ . To save computational cost, we use  $\boldsymbol{\eta}$  to replace  $\mathbf{f}$  in Eqs. (15) and (16) and then use Eqs. (17) and (18) to select the leap-size  $\tau$ . The same  $\boldsymbol{\eta}$  is used to determine the random corrections in the new methods.

## 2. Computational efficiency

The computational costs needed by the tau-leaping, the MP, and the PRC are relatively the same. In PRC, we need the  $M \times M$  matrix  $\boldsymbol{\eta}$  defined in Eq. (6) to estimate the mean of  $\mathbf{r}^*$ . However the same  $\boldsymbol{\eta}$  can be used in the leap-size selection procedure to determine  $\tau$ . There will be no noticeable computational overhead compared with the more time-consuming random number generating processes.

Approximately the computational costs are doubled for the GRC1 and GRC2 compared with the tau-leaping because the random variables needed by them are doubled in each time step. However, the GRC1 and GRC2 can be more efficient than other methods because they have much better accuracy. This will be shown in Sec. IV.

## IV. NUMERICAL RESULTS

We apply the tau-leaping, MP, PRC, GRC1, and GRC2 to four chemical reaction systems. In the first two systems, exact solutions for the mean and variance can be obtained. In the other two systems, the mean and variance sampled from the SSA are considered to be exact values. Note that since there is only one reaction in the first two systems, the GRC1 and GRC2 are exactly the same, and they are simply called GRC. For the other two systems, the GRC1 and GRC2 are effectively different.

In order to demonstrate the order of accuracy, we follow a procedure that is widely used in the numerical study of ODEs. We simulate  $X_t$  from time  $t=0$  to  $t=T$ , advancing by a fixed time step size  $\tau$ . If the sample size is large enough, the statistical fluctuations in  $E_j^\tau$ , the sample mean of  $X_j$  with step size  $\tau$ , can be neglected. Then we doubled the step size to  $2\tau$  and obtain the sample mean  $E_j^{2\tau}$ . If the simulation method has  $p$ th order accuracy for the mean, there exists a constant  $C$  such that

$$|E_j^\tau - \mathbb{E}(X_j)| \approx C\tau^p, \quad j = 1, \dots, M.$$

It follows

$$\frac{|E_j^{2\tau} - \mathbb{E}(X_j)|}{|E_j^\tau - \mathbb{E}(X_j)|} \approx \frac{C(2\tau)^p}{C\tau^p} \approx 2^p, \quad j = 1, \dots, M.$$

For example, if  $p=2$ , then  $|E_j^{2\tau} - \mathbb{E}(X_j)| \approx 4|E_j^\tau - \mathbb{E}(X_j)|$ , i.e., the absolute error of the mean will be approximately four times larger when the step size is doubled. In addition, the log-log plot of the absolute error of the mean over the step size will be a straight line with slope 2. The same is true for the variance. These substantiate the order of accuracy.

### A. System 1: $S \rightarrow \emptyset$

For this system the propensity function is  $a(x)=cx$ , where the rate constant  $c=0.1$ . The state-change vector is  $\nu=-1$ . The initial condition is set to  $X_0=10\,000$ . The mean and variance can be solved explicitly,

$$\mathbb{E}[X_T] = X_0 e^{-cT},$$

$$\text{Var}[X_T] = X_0(e^{-cT} - e^{-2cT}).$$

For this system the GRC1 and GRC2 are identical, in which  $\tilde{r}$  satisfies

$$E_r[\tilde{r}] = \tau\boldsymbol{\eta}\tilde{r} - \frac{\tau^2}{2}a\boldsymbol{\eta} \quad (19)$$

since  $\boldsymbol{\eta}=-0.1 < 0$ . We simulate the reaction from time 0 to  $T=10.4$  using different step sizes.

We plot the absolute errors of mean and variance in Fig. 1. The sample size is as large as  $10^8$  so that the magnitude of statistical fluctuation is small. It shows that for this system, tauleap has first order accuracy for mean and variance; PRC has second order accuracy for mean and first order accuracy for variance; GRC has second order for mean, for variance it is not a straight line (the error of variance for GRC is so small for this system that statistical fluctuation become significant); MP has first order accuracy for variance, while for the mean it shows second order accuracy only when  $\tau$  is large. This phenomenon that the second order property is lost as  $\tau$  gets small is perhaps due to the rounding of  $X^{(1)}$ . An interesting discussion of this irregular behavior of rounding may be found in Ref. 10.

In Fig. 2 we plot the histograms of  $X_T$  for different methods. The histogram of tauleap has large deviations from SSA; GRC shows the best performance in capturing the PDF.

### B. System 2: $S \rightarrow 2S$

This system has one reaction with propensity function  $a(x)=cx$ , where the rate constant  $c=0.1$ . The state-change vector is  $\nu=1$ . The initial condition is  $X_0=400$ . The exact mean and variance of  $X_T$  can be solved explicitly,

$$\mathbb{E}[X_T] = X_0 e^{cT},$$

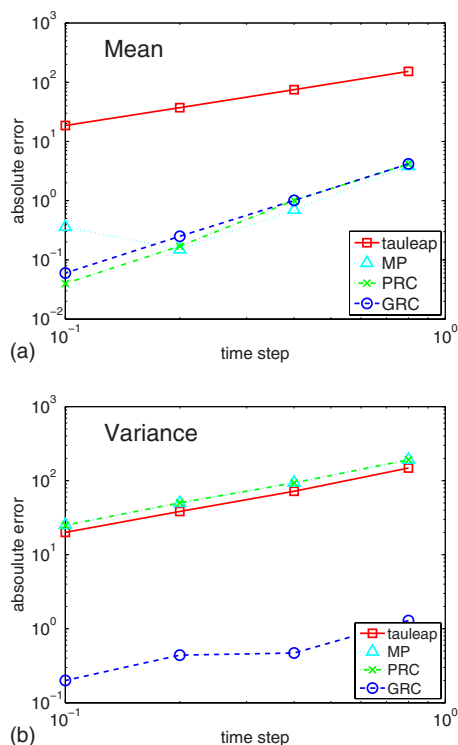


FIG. 1. (Color online) (System 1) Comparison of the absolute errors of sample mean and variance of  $X$  at  $\tau=10.4$ , with a sample size of  $10^8$ . For the mean, tauleap has first order accuracy; PRC and GRC both have second order accuracy; the behavior of MP is less regular than the others. For the variance, tauleap, MP, and PRC all have first order accuracy; GRC has extremely small error and statistical fluctuation become evident, especially when step size is small. Error values of mean and variance are presented in Tables IV and V.

$$\text{Var}[X_T] = X_0(e^{2cT} - e^{cT}).$$

The major difference from system 1 is that we have  $\eta=0.1 > 0$ , which causes the second term in Eq. (19) to vanish so that

$$E_r[\bar{r}] = \frac{\tau}{2} \eta r. \quad (20)$$

We plot the absolute errors of mean and variance in Fig. 3. The sample size is  $10^8$  for each sample mean and variance. For the mean, PRC and GRC are almost identical, showing convergence of the order of 2; MP seems to have no order relation for this system. For the variance, all methods fit in straight lines (for GRC there is a little statistical fluctuation when the error is extremely small). It clearly shows that GRC improve the order of accuracy of the variance compared with other methods. We can see that the errors of the sample mean and variance for GRC when  $\tau=1.6$  are  $-4.171$  and  $-47.846$ , respectively. To achieve the same accuracy using the tauleap method, one needs  $\tau=0.1$  when the errors are  $-5.830$  and  $-47.564$ , respectively, which means for this particular case, GRC is eight times more efficient than tauleap.

In Fig. 4 we plot the histograms of  $X_T$  for different methods. The histogram of tauleap has large deviations from SSA; GRC shows the best performance in capturing the PDF.

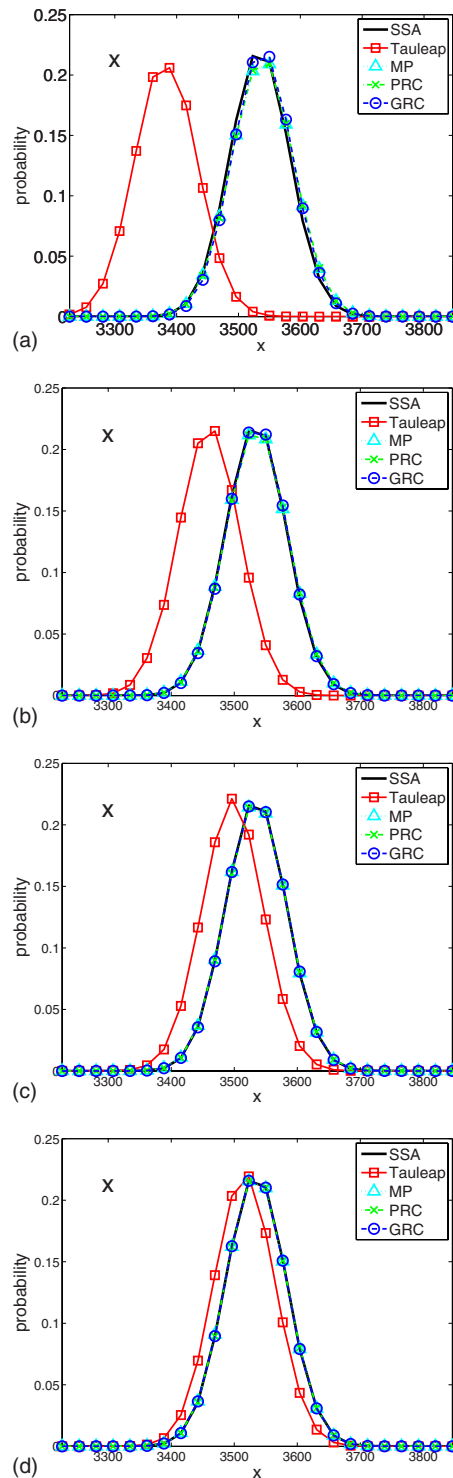
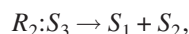
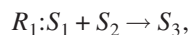


FIG. 2. (Color online) (System 1) Comparison of the histograms obtained from  $10^6$  samples. We can see that tauleap shows significant deviations from SSA. The other three methods can capture the PDF quite well. The histogram distance appears in Table VI.

### C. System 3: The Michaelis–Menten system

The Michaelis–Menten system describes the kinetics of some enzymes. It involves four species participating in three chemical reactions. The chemical reactions are





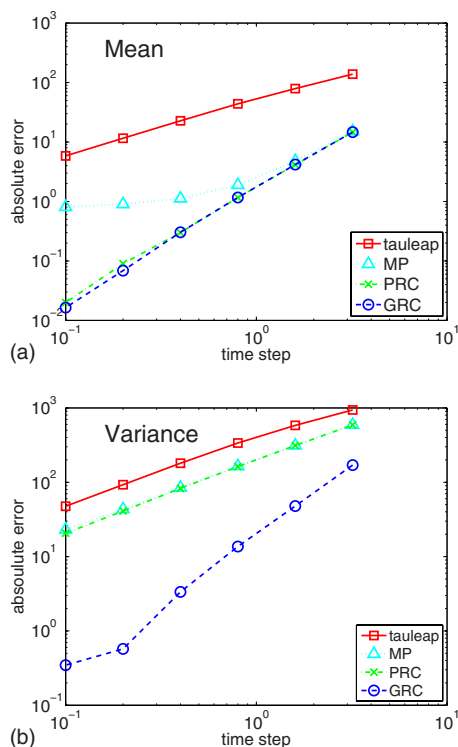


FIG. 3. (Color online) (System 2) Comparison of the absolute errors of  $X$  at  $t=10.4$ , with a sample size of  $10^8$ . For the mean, tauleap shows first order accuracy; PRC and GRC both show second order accuracy; there is still no clear order of accuracy for MP. For the variance, tauleap, MP, and PRC all show first order accuracy; GRC clearly shows second order accuracy. Error values are presented in Tables VII and VIII.

$$R_3:S_3 \rightarrow S_2 + S_4.$$

Detailed simulation of this system can be found in Ref. 29. In our implementation, the rate constant  $\mathbf{c}=(1 \times 10^{-4}, 0.5, 0.5)^T$ , the propensity function is  $\mathbf{a}(\mathbf{x})=(c_1 x_1 x_2, c_2 x_3, c_3 x_3)^T$ , and the initial value of  $\mathbf{X}$  is  $\mathbf{X}_0=(1000, 200, 2000, 0)^T$ . We simulate the system in the time interval  $[0, 6]$  with sample size equal to  $10^6$ .

For this system, GRC1 and GRC2 behave differently. As Figs. 5 and 6 show, the absolute errors of mean of both methods are quite similar, but for variance GRC2 is more accurate than GRC1. GRC1 shows second order accuracy for the variance of both  $X_1$  and  $X_2$ . Due to the statistical fluctuation, the error of variance for GRC2 does not form a straight line. Strangely, when  $\tau$  goes from 0.4 to 0.8, the absolute error of the variance of  $X_2$  for GRC2 grows 12 times. Above all, for both GRC1 and GRC2, the overall picture of second order accuracy for both mean and variance is still valid.

Histograms obtained from  $10^6$  samples of  $X_1$  and  $X_2$  using different step sizes are plotted in Figs. 7 and 8. For relatively large  $\tau$ , it is evident from the figures that GRC2 achieves outstanding performance in capturing the PDF.

#### D. System 4: A more complicated system

This system involves 8 species and 12 reactions.<sup>30</sup> The chemical reactions, propensity functions, and initial values used in our implementation are listed in Tables II and III. We simulate the system in the time interval  $[0, 3]$  with sample size equal to  $10^6$ .

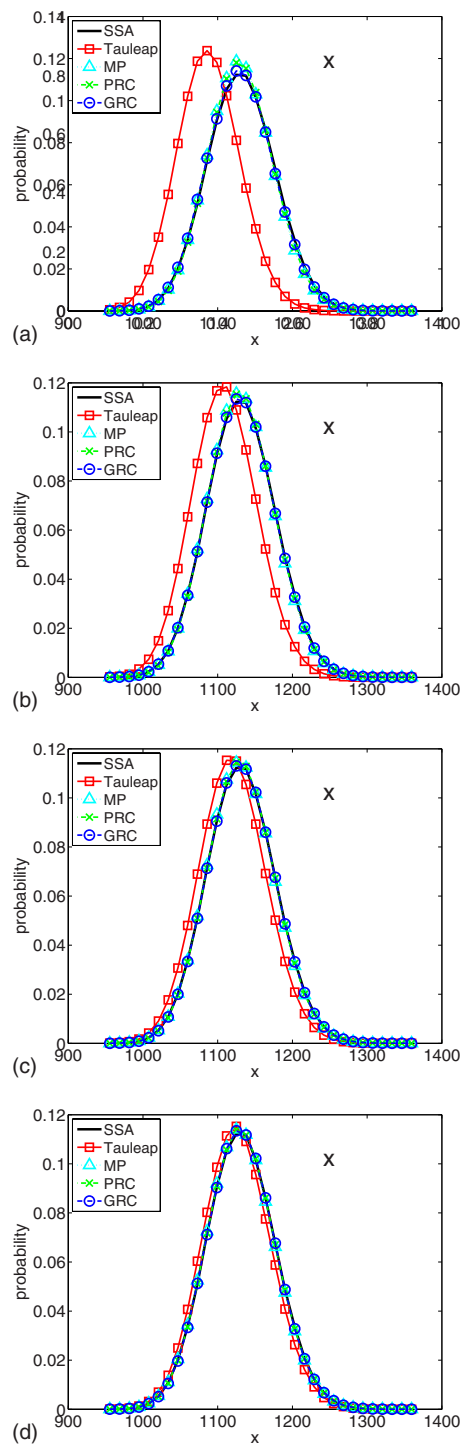


FIG. 4. (Color online) (System 2) Comparison of histograms obtained from  $10^6$  samples. Compared with SSA histogram, that of GRC is much closer than tauleaps and slightly closer than MPs and PRCs. The histogram distance appears in Table IX.

The log-log plots and histograms of  $X_1$  and  $X_7$  are given in Figs. 9–12. It shows that the four improved methods—MP, PRC, GRC1, and GRC2—all have comparable accuracy for the mean of both  $X_1$  and  $X_7$ , and also for the variance of  $X_1$ . However, for the variance of  $X_7$ , the accuracy of MP and PRC are even worse than that of tauleap, while GRC1 and GRC2 still show very good accuracy. It appears that MP and PRC can significantly improve the accuracy of the mean, but the errors of variance for these methods are somehow unre-

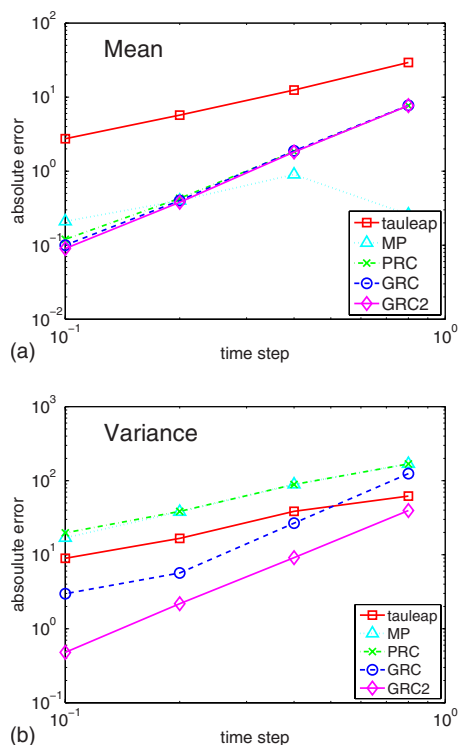


FIG. 5. (Color online) (System 3,  $X_1$ ) Absolute errors of sample mean and variance of  $X_1$  at  $t=6$  with a sample size of  $10^6$ . The sample mean and variance of  $10^6$  simulations using the SSA is considered the exact value. For this system, GRC2 differs from GRC1. For the mean, tauleap shows first order accuracy, while the PRC, GRC1, and GRC2 show second order accuracy and MP shows no clear order relation. For variance, tauleap, MP, and PRC show first order accuracy, while GRC1 and GRC2 are of second order accuracy. Note that even though MP and PRC give a more accurate mean than tauleap, they are less accurate than tauleap for variance. GRC1 and, especially GRC2, have good accuracy for this system. Error values are presented in Tables X and XI.

dictable. On the other hand, GRC1 and GRC2 have approximately second order accuracy for both the mean and variance. So if we demand small error in both the mean and variance, using GRC1 and, especially, GRC2 will be more effective.

### E. System 4 with adaptive leap-size selection

In all the above examples, the step sizes in each sampling are fixed. Here we will test the performance of the new schemes with adaptive leap-size selection. Two leap-size selection procedures are reviewed in Sec. III C 1.

For system 4, both procedures (17) and (18) work for the tauleap method if  $\epsilon$  is not too big. However for the GRC1 and GRC2, the difference between these two strategies is drastic: if we use Eq. (17), negative population is likely to happen during the simulation even when the parameter  $\epsilon$  is very small; if we use Eq. (18), then there will be no such problem. This example reminds us that when applying the new methods, we should be careful in selecting the leap size. It should be possible to design a more robust leap-size selection strategy for GRC1 and GRC2. We speculate that it must be related to their numerical stability. For now, we just recommend the leap-size selection procedure (18) and plan to study this in the future.

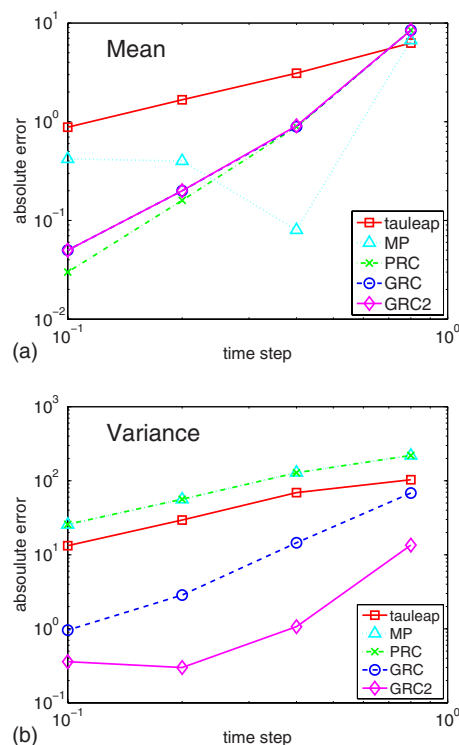


FIG. 6. (Color online) (System 3,  $X_2$ ) Absolute errors of sample mean and variance of  $X_2$  at  $t=6$  with a sample size of  $10^6$ . The behavior of these methods is essentially the same as the case for  $X_1$ . Note that the accuracy of GRC2 for variance is much better than the others. Error values are presented in Tables XII and XIII.

We test all the five methods for system 4 using adaptive leap-size selection procedure (18) with  $\epsilon=0.01$  and  $\epsilon=0.005$ , respectively. For a sample size of  $10^6$ , the errors of mean and variance for each species are presented in Tables XXII–XXV. For the mean, it seems that, the four improved methods have comparable accuracy (statistical fluctuations are relatively large for this system) and their improvements over tauleap are evident. For the variance, the errors of variance of  $X_5$  and  $X_7$  are very large for MP and PRC, but they are small for GRC1 and GRC2, which shows their advantage. For  $\epsilon=0.01$ , GRC1 seems less accurate than MP and PRC. However for  $\epsilon=0.005$ , GRC1 is more accurate than both. GRC2 is more accurate than MP and PRC even for  $\epsilon=0.01$  and it is the most accurate scheme of all.

### V. CONCLUSION

Some highly accurate methods for simulating spatially homogeneous well-stirred 3 chemical reaction systems are proposed. The idea of the new methods is to add random correction to the primitive tau-leaping term. More precisely, in the new methods, the vector  $\mathbf{r}^*$  approximating the number of firings of each reaction channel in a time interval  $[t_0, t_0 + \tau)$  is generated in two steps. The first is to generate Poisson random vector  $\mathbf{r}$ , which is the same as the primitive tau-leaping. Next is to generate a random correction vector  $\tilde{\mathbf{r}}$  according to certain conditions. The final approximation is the sum of the two parts,  $\mathbf{r}^* = \mathbf{r} + \tilde{\mathbf{r}}$ . The novelty of our methods is that the

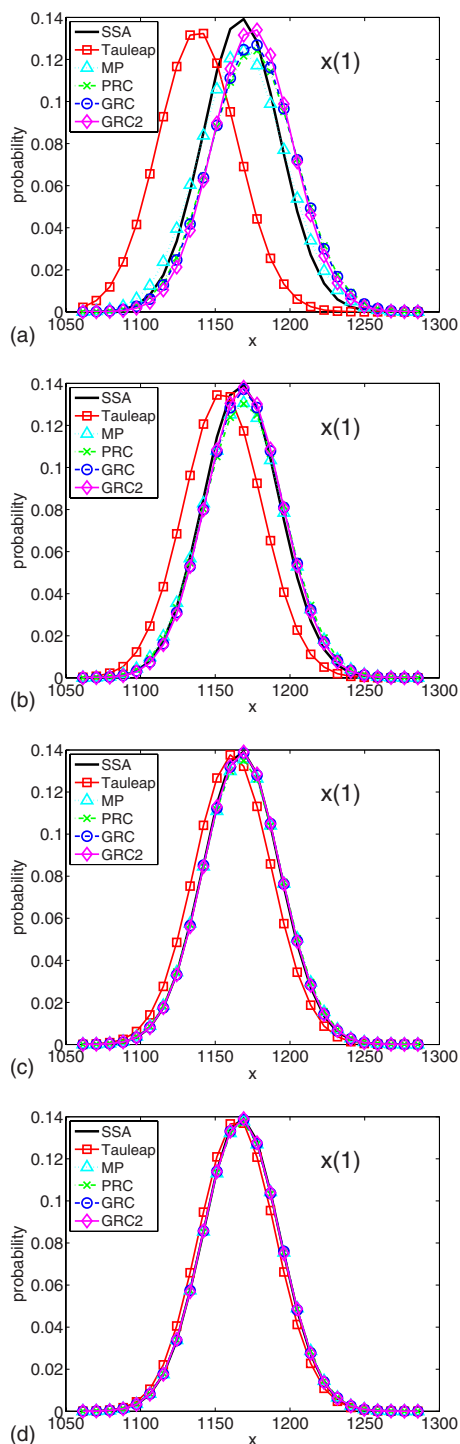


FIG. 7. (Color online) (System 3,  $X_1$ ) Histograms of  $X_1$  at  $t=6$  obtained from  $10^6$  samples. The histograms of GRC2 appear to be the closest one to SSAs for all step sizes. The histogram distance appears in Table XIV.

random correction  $\tilde{r}$  is generated conditioning on  $r$ . This allows us to design schemes that are of higher order accuracy at least for the mean and covariance of  $X_r$ , which is one first step towards high order accuracy method for simulating jump processes.

In this paper three examples of RC-tau-leaping schemes are proposed, namely, the PRC, GRC1, and GRC2. Their improvement in accuracy comes from the fact that these schemes have higher order LTE for moments of  $X_r$ . It has been shown that the LTE of the primitive tau-leaping is

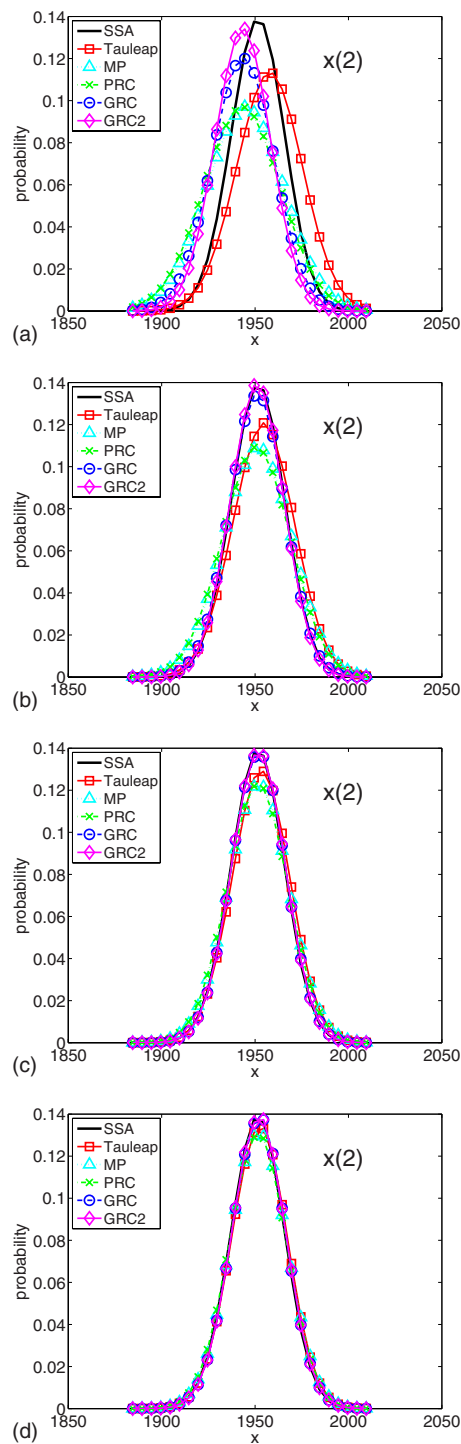


FIG. 8. (Color online) (System 3,  $X_2$ ) Histograms of  $X_2$  at  $t=6$  obtained from  $10^6$  samples. It clearly shows that the GRC1 and, especially GRC2, capture the PDF of  $X_2$  better than the other methods. For  $\tau=0.2$ , GRC1 and GRC2 are almost identical with SSA, while for tauleap, MP, and PRC the error is quite obvious. The histogram distance appears in Table XV.

$\mathcal{O}(\tau^2)$  for all moments.<sup>10</sup> Our first objective is to reduce the LTE of a scheme to  $\mathcal{O}(\tau^3)$  just for the mean. This is a relatively simple task and there are many alternatives for choosing such a random correction. In the PRC scheme, we adapted a simple form in which  $r^*$  can be generated in one pass. At the end of Sec. III A we demonstrated that, from another perspective, the PRC can be considered a variation in the MP. The major difference between the two is that

TABLE II. List of reactions and propensity functions for system 4.

	Reaction	Propensity	Rate constant
1	$E_A \rightarrow E_A + A$	$a_1 = c_1[E_A]$	$c_1 = 150$
2	$E_B \rightarrow E_B + B$	$a_2 = c_3[E_B]$	$c_2 = 150$
3	$E_A + B \rightarrow E_A B$	$a_3 = c_3[E_A][B]$	$c_3 = 0.001$
4	$E_A B \rightarrow E_A + B$	$a_4 = c_4[E_A B]$	$c_4 = 6$
5	$E_A B + B \rightarrow E_A B_2$	$a_5 = c_5[E_A B][B]$	$c_5 = 0.001$
6	$E_A B_2 \rightarrow E_A B + B$	$a_6 = c_6[E_A B_2]$	$c_6 = 6$
7	$A \rightarrow \emptyset$	$a_7 = c_7[A]$	$c_7 = 5$
8	$E_B + A \rightarrow E_B A$	$a_8 = c_8[E_B][A]$	$c_8 = 0.001$
9	$E_B A \rightarrow E_B + A$	$a_9 = c_9[E_B A]$	$c_9 = 6$
10	$E_B A + A \rightarrow E_B A_2$	$a_{10} = c_{10}[E_B A][A]$	$c_{10} = 0.001$
11	$E_B A_2 \rightarrow E_B A + A$	$a_{11} = c_{11}[E_B A_2]$	$c_{11} = 6$
12	$B \rightarrow \emptyset$	$a_{12} = c_{12}[B]$	$c_{12} = 5$

while the MP tries to estimate the midpoint of the population of reactant species  $X_i$ , the PRC tries to estimate the midpoint of the propensity functions  $\mathbf{a}$ . Numerical results show that the PRC is more accurate than the MP.

The next goal, which is more interesting and challenging, is to reduce the LTE of a scheme to  $\mathcal{O}(\tau^3)$  for both the mean and covariance of  $\mathbf{X}_i$ . We proved in Proposition 2 that this goal cannot be achieved unless the components in  $\mathbf{r}^*$  are mutually dependent. We introduced a correlation among components in  $\mathbf{r}^*$  by generating  $\tilde{\mathbf{r}}$  conditioned on the previous approximation  $\mathbf{r}$ . This is the central idea behind the RC tau-leaping methods. By matching the coefficients in the Taylor expansion of the exact mean and covariance of  $\mathbf{X}_i$  up to  $\mathcal{O}(\tau^2)$ , we presented in Proposition 3 a sufficient condition for  $\tilde{\mathbf{r}}$  that leads to second order consistency for both the mean and covariance.

Both the GRC1 and GRC2 satisfy the second order condition in Proposition 3, and numerical results indeed show second order accuracy for them. Their difference lies in the procedure for generating the random correction  $\tilde{\mathbf{r}}$ . In the GRC1,  $\tilde{\mathbf{r}}$  is generated according to Eqs. (10) and (11), which is in a concrete form and easier to understand. The modification in the GRC2 aims to reduce the variance of the components in  $\tilde{\mathbf{r}}$ . The approach is not very systematic, but the numerical result is much better than those of GRC1. We speculate that keeping the variance of the components in  $\tilde{\mathbf{r}}$  as small as possible is beneficial for the method's accuracy and stability. However more study is needed to understand this mechanism and to develop methods that are even faster and more accurate.

TABLE III. List of species and their initial value (in number of molecules) for system 4.

	Species	Initial value
1	A	2000
2	B	1500
3	$E_A$	950
4	$E_B$	950
5	$E_A B$	200
6	$E_A B_2$	50
7	$E_B A$	200
8	$E_B A_2$	50

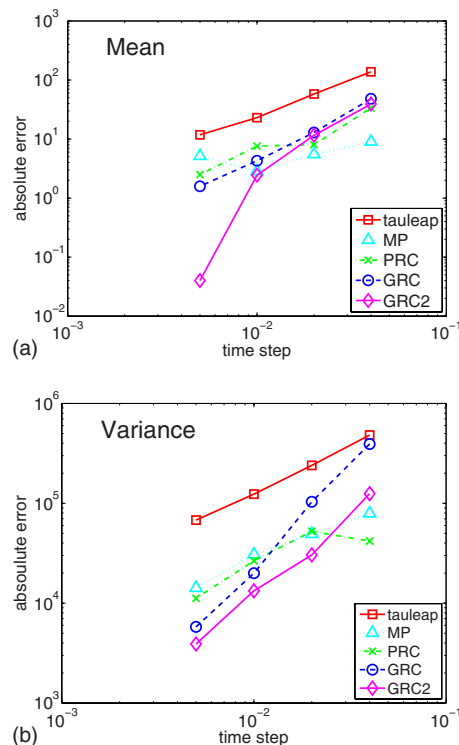


FIG. 9. (Color online) (System 4,  $X_1$ ) Absolute errors of sample mean and variance of  $X_1$  at  $t=3$  with a sample size of  $10^6$ . The mean and variance sampled from the SSA are considered exact value. The improvement in the accuracy of variance by using GRC1 and GRC2 is quite obvious. Error values are presented in Tables XVI and XVII.

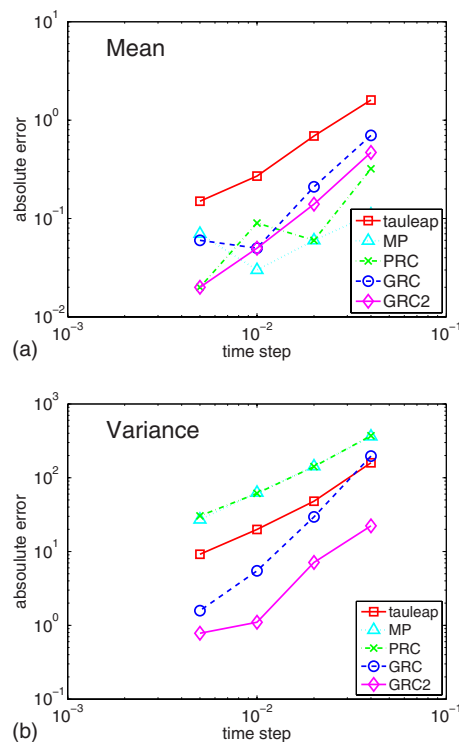


FIG. 10. (Color online) (System 4,  $X_7$ ) Absolute errors of sample mean and variance of  $X_7$  at  $t=3$  with a sample size of  $10^6$ . Here we can see that for the variance of  $X_7$ , the performances of MP and PRC are worse than those of tauleap. The GRC1 and GRC2, however, still preserve second order accuracy for both the mean and variance. Error values are presented in Tables XVIII and XIX.

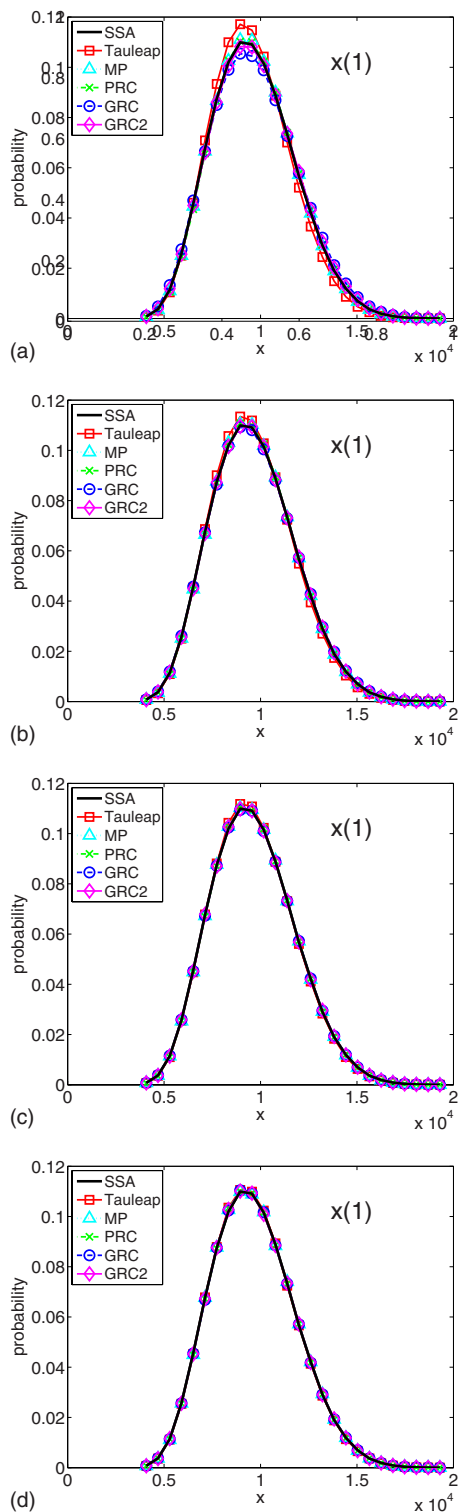


FIG. 11. (Color online) (System 4,  $X_1$ ) Histograms of  $X_1$  at  $t=3$  obtained from  $10^6$  samples. The histogram distance appears in Table XX.

The new methods presented in this paper can be directly applied to chemical reaction systems for which the tau-leaping method is suitable. The PRC, which improves accuracy while requiring almost the same computational effort, can be used as a replacement for MP. The GRC2, which is probably better than the GRC1, seems to be more promising for simulating chemical reaction systems. It doubles the computational cost of the tau-leaping method, but greatly

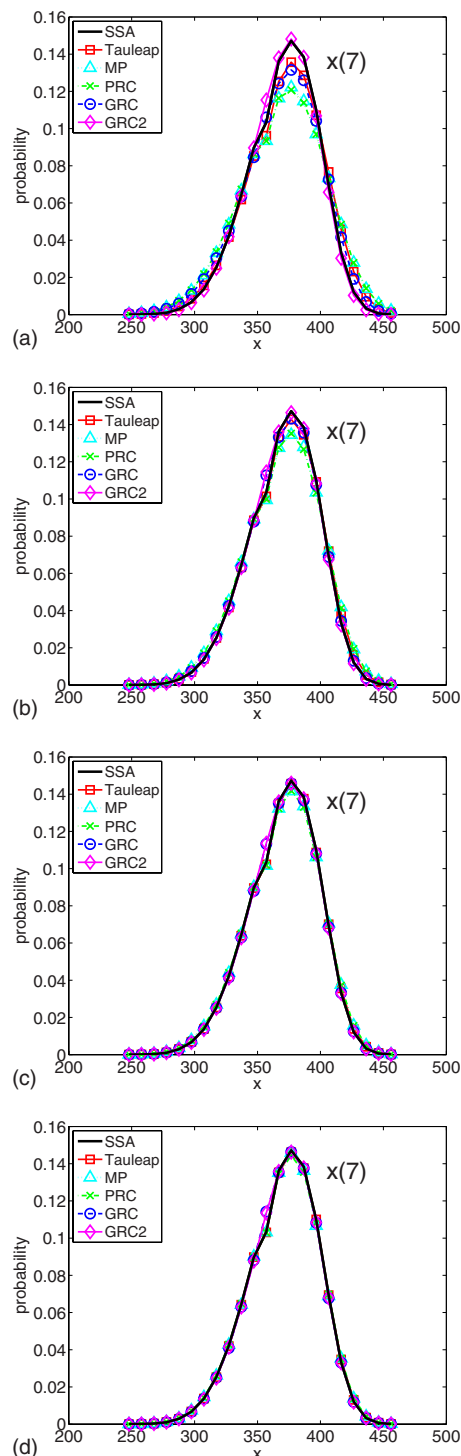


FIG. 12. (Color online) (System 4,  $X_7$ ) Histograms of  $X_7$  at  $t=3$  obtained from  $10^6$  samples. The histogram distance appears in Table XXI.

improves accuracy. Numerical results show that it can be more efficient overall than other schemes. Finally we mention that, in simulating complex chemical reaction systems, one still needs an adaptive step-size selection procedure. There exist several step-size selection strategies, but they have only been tested with the tau-leaping method. As seen in Sec. IV E, for system 4, the leap-size selection procedure (17) works fine in the tau-leaping method but suffers from the negative population issue if applied to GRC1 and GRC2. We believe that analyzing the stability property will provide

TABLE IV. (System 1) Error of the mean, as shown by Fig. 1(a). Note that here the sign of error is kept while in the figures it is their absolute value. Numbers in the brackets are the increasing ratio of the error when the step size is doubled.

$\tau$	Tauleap	MP	PRC	GRC
0.8	-152.03(NaN)	3.84 (NaN)	4.15 (NaN)	4.17(NaN)
0.4	-74.72(2.03)	0.70(5.49)	1.01(4.11)	1.01(4.13)
0.2	-37.14(2.01)	-0.15(-4.67)	0.17(5.94)	0.25(4.04)
0.1	-18.56(2.00)	-0.36(0.42)	-0.04(-4.25)	0.06(4.17)

TABLE V. (System 1) Error of the variance, as shown by Fig. 1(b).

$\tau$	Tauleap	MP	PRC	GRC
0.8	148.30(NaN)	191.70(NaN)	191.60(NaN)	1.29 (NaN)
0.4	71.89(2.06)	94.03(2.04)	93.28(2.05)	0.47(2.74)
0.2	38.36(1.87)	49.88(1.89)	50.20(1.86)	-0.44(-1.07)
0.1	20.03(1.92)	25.21(1.98)	24.97(2.01)	-0.20(2.20)

TABLE VI. (System 1) The  $L^1$ -distance for the histograms compared with the SSA. For tauleap, the distance is doubled when step size is doubled. This is reasonable because it has first order accuracy for all moments. For other methods there is no obvious order relation.

$\tau$	Tauleap	MP	PRC	GRC
0.8	1.750	0.069	0.073	0.063
0.4	1.112	0.022	0.023	0.026
0.2	0.600	0.010	0.011	0.017
0.1	0.306	0.007	0.005	0.004

TABLE VII. (System 2) Error of the mean, as shown by Fig. 3(a).

$\tau$	Tauleap	MP	PRC	GRC
3.2	-138.099 (NaN)	-15.074 (NaN)	-14.640 (NaN)	-14.634 (NaN)
1.6	-79.125(1.745)	-4.780(3.153)	-4.150(3.528)	-4.171(3.509)
0.8	-43.816(1.806)	-1.892(2.526)	-1.160(3.578)	-1.165(3.579)
0.4	-22.699(1.930)	-1.120(1.690)	-0.300(3.867)	-0.303(3.851)
0.2	-11.562(1.963)	-0.901(1.243)	-0.090(3.333)	-0.069(4.411)
0.1	-5.830(1.983)	-0.807(1.117)	-0.020(4.500)	-0.016(4.183)

TABLE VIII. (System 2) Error of the variance, as shown by Fig. 3(b).

$\tau$	Tauleap	MP	PRC	GRC
3.2	-939.632 (NaN)	-592.069 (NaN)	-591.310 (NaN)	-170.086 (NaN)
1.6	-582.306(1.614)	-312.914(1.892)	-311.940(1.896)	-47.846 (3.555)
0.8	-336.894(1.728)	-164.700(1.900)	-162.080(1.925)	-13.645 (3.506)
0.4	-180.311(1.868)	-84.503(1.949)	-82.850(1.956)	-3.348 (4.075)
0.2	-92.724(1.945)	-43.492(1.943)	-41.000(2.021)	-0.574 (5.837)
0.1	-47.564(1.949)	-23.118(1.881)	-20.640(1.986)	-0.346 (1.657)

TABLE IX. (System 2) The  $L^1$ -distance for the histograms compared with the SSA. The improvement in the accuracy of MP, PRC, and especially GRC over the tauleap is significant for this system.

$\tau$	Tauleap	MP	PRC	GRC
0.8	0.769	0.050	0.044	0.028
0.4	0.387	0.020	0.023	0.014
0.2	0.203	0.020	0.015	0.005
0.1	0.107	0.021	0.014	0.005

TABLE X. (System 3) Error of the mean of  $X_1$ , as shown by Fig. 5(a).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.8	-29.35(NaN)	0.26(NaN)	7.66(NaN)	7.73(NaN)	7.67(NaN)
0.4	-12.45(2.36)	0.90(0.29)	1.83(4.19)	1.88(4.11)	1.81(4.24)
0.2	-5.71(2.18)	0.40(2.25)	0.43(4.26)	0.40(4.70)	0.38(4.76)
0.1	-2.74(2.08)	0.21(1.90)	0.12(3.58)	0.10(4.00)	0.09(4.22)

TABLE XI. (System 3) Error of the variance  $X_1$ , as shown by Fig. 5(b).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.8	61.99(NaN)	168.66(NaN)	168.44(NaN)	124.62(NaN)	39.47(NaN)
0.4	38.53(1.61)	89.16(1.89)	88.88(1.90)	26.71(4.67)	9.12(4.33)
0.2	16.63(2.32)	37.98(2.35)	38.51(2.31)	5.67(4.71)	2.18(4.18)
0.1	8.95(1.86)	16.91(2.25)	19.74(1.95)	2.97(1.91)	0.48(4.54)

TABLE XII. (System 3) Error of the mean of  $X_2$ , as shown by Fig. 6(a).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.8	6.27(NaN)	-6.70 (NaN)	-8.47(NaN)	-8.46(NaN)	-8.44(NaN)
0.4	3.10(2.02)	-0.08(83.75)	-0.89(9.52)	-0.89(9.51)	-0.91(9.27)
0.2	1.67(1.86)	0.40(-0.20)	-0.16(5.56)	-0.20(4.45)	-0.20(4.55)
0.1	0.88(1.90)	0.42(0.95)	-0.03(5.33)	-0.05(4.00)	-0.05(4.00)

insight in designing more robust leap-size selection procedure for the new methods, rendering them even more efficient.

## ACKNOWLEDGMENTS

The authors are supported by the National Science Foundation of China under Grant No. 10871010 and the National Basic Research Program under Grant No. 2005CB321704. We would like to thank the reviewers for their constructive comments and Professor Sharon Murrell for helping us with the English.

## APPENDIX A: TAYLOR EXPANSION OF $\mathbb{E}_{\mathbf{x}}[(X_{t_n+\tau} - X_{t_n})^p]$

Here we state some results that have been reported in Ref. 10. By first Taylor expanding the CME of the chemical reaction system, and then taking the  $p$ th moment of the increment  $X_{t_n+\tau} - X_{t_n}$ , one obtains that for  $p \geq 1$ ,

$$\begin{aligned} \mathbb{E}_{\mathbf{x}}[(X_{t_n+\tau} - X_{t_n})^p] &= \tau \sum_{j=1}^M \mathbf{v}_j^p a_j - \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j^p a_j a_k - \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j^p a_j a_k (\mathbf{x} + \mathbf{v}_k) \\ &+ \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M (\mathbf{v}_j + \mathbf{v}_k)^p a_j a_k + \mathcal{O}(\tau^3). \end{aligned} \quad (\text{A1})$$

Note that here  $\mathbf{v}_j \mathbf{v}_k$ , which should be interpreted as the matrix  $\mathbf{v}_j \mathbf{v}_k^T$ , is not equal to  $\mathbf{v}_k \mathbf{v}_j$ , which should be interpreted as the matrix  $\mathbf{v}_k \mathbf{v}_j^T$ .

Letting  $p = 1, 2$ , substituting Eq. (6) and simplifying, we obtain

TABLE XIII. [System (3)] Error of the variance  $X_2$ , as shown by Fig. 6(b).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.8	103.31 (NaN)	218.76 (NaN)	219.98 (NaN)	68.20 (NaN)	13.52 (NaN)
0.4	69.14 (1.49)	128.85 (1.70)	128.13 (1.72)	14.53 (4.69)	-1.07 (-12.64)
0.2	29.45 (2.35)	56.06 (2.30)	56.30 (2.28)	2.86 (5.08)	-0.30 (3.57)
0.1	13.27 (2.22)	25.51 (2.20)	25.79 (2.18)	0.96 (2.98)	-0.36 (0.83)

TABLE XIV. (System 3,  $X_1$ ) The  $L^1$ -distance for the histograms compared with the SSA. The MP, PRC, GRC1, and GRC2 give similar results on histogram distance. Their improvement over tauleap is evident.

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.8	0.845	0.111	0.241	0.237	0.235
0.4	0.376	0.065	0.080	0.058	0.054
0.2	0.177	0.029	0.030	0.014	0.012
0.1	0.086	0.013	0.015	0.008	0.007

TABLE XV. (System 3,  $X_2$ ) The  $L^1$ -distance for the histograms compared with the SSA. When  $\tau=0.8$ , the distances for all the methods are close, but as  $\tau$  gets smaller, GRC1 and GRC2 give better results.

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.8	0.354	0.434	0.481	0.418	0.434
0.4	0.203	0.234	0.236	0.039	0.024
0.2	0.106	0.117	0.119	0.017	0.016
0.1	0.057	0.058	0.060	0.026	0.025

TABLE XVI. (System 4) Error of the mean of  $X_1$ , as shown by Fig. 9(a).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.04	-137.42(NaN)	9.01(NaN)	33.55(NaN)	48.02(NaN)	39.14 (NaN)
0.02	-57.88(2.37)	5.52(1.63)	8.02(4.18)	12.91(3.72)	11.57(3.38)
0.01	-22.99(2.52)	2.85(1.94)	7.58(1.06)	4.29(3.01)	2.44(4.74)
0.005	-11.79(1.95)	5.16(0.55)	2.50(3.03)	1.58(2.72)	0.04(61.00)

TABLE XVII. (System 4) Error of the variance  $X_1$ , as shown by Fig. 9(b).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.04	-481 902.97(NaN)	-79 168.90(NaN)	-41 920.25(NaN)	392 130.51 (NaN)	124 655.17 (NaN)
0.02	-239 708.57(2.01)	-49 304.18(1.61)	-52 396.39(0.80)	103 463.86(3.79)	30 397.63(4.10)
0.01	-123 456.45(1.94)	-30 703.45(1.61)	-26 553.29(1.97)	19 967.65(5.18)	13 283.21(2.29)
0.005	-67 855.77(1.82)	-14 223.73(2.16)	-11 196.76(2.37)	-5785.42(-3.45)	-3896.82(-3.41)

TABLE XVIII. (System 4) Error of the mean of  $X_7$ , as shown by Fig. 10(a).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.04	1.60(NaN)	-0.11(NaN)	-0.32(NaN)	-0.70(NaN)	-0.47(NaN)
0.02	0.69(2.32)	-0.06(1.83)	-0.06(5.33)	-0.21(3.33)	-0.14(3.36)
0.01	0.27(2.56)	-0.03(2.00)	-0.09(0.67)	-0.05(4.20)	-0.05(2.80)
0.005	0.15(1.80)	-0.07(0.43)	-0.02(4.50)	-0.06(0.83)	-0.02(2.50)

TABLE XIX. (System 4) Error of the variance  $X_7$ , as shown by Fig. 10(b).

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.04	159.39(NaN)	362.48(NaN)	370.83(NaN)	196.34 (NaN)	-22.30(NaN)
0.02	48.09(3.31)	142.70(2.54)	141.37(2.62)	29.50(6.66)	-7.12(3.13)
0.01	19.91(2.42)	62.80(2.27)	61.38(2.30)	5.47(5.39)	-1.10(6.47)
0.005	9.18(2.17)	27.08(2.32)	30.37(2.02)	-1.58(-3.46)	-0.78(1.41)

TABLE XX. (System 4,  $X_1$ ) The  $L^1$ -distance for the histograms compared with the SSA.

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.04	0.066	0.012	0.015	0.041	0.017
0.02	0.031	0.008	0.008	0.012	0.006
0.01	0.015	0.005	0.006	0.004	0.004
0.005	0.009	0.005	0.005	0.004	0.004

TABLE XXI. (System 4,  $X_7$ ) The  $L^1$ -distance for the histograms compared with the SSA.

$\tau$	Tauleap	MP	PRC	GRC1	GRC2
0.04	0.099	0.190	0.194	0.104	0.029
0.02	0.033	0.086	0.085	0.031	0.022
0.01	0.014	0.040	0.040	0.022	0.021
0.005	0.007	0.018	0.019	0.022	0.021

TABLE XXII. The error of the mean for  $X=(X_1, X_2, \dots, X_8)$  at  $t=3$  with  $\epsilon=0.01$ , where  $\epsilon$  is the parameter in the leap-size selection procedure. The sample size is  $10^6$ . Still, the sample mean and variance obtained from the SSA is considered as the exact value. "Steps" is the averaged time steps needed in one simulation.

Method	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	Steps
Tauleap	-29.94	14.23	-1.20	0.50	0.17	0.93	0.40	-0.90	124.41
MP	-0.96	-4.35	0.02	-0.12	-0.05	-0.08	0.03	0.09	126.01
PRC	0.87	-5.38	0.13	-0.17	-0.07	-0.16	0.02	0.15	125.99
GRC1	6.02	4.26	0.28	0.21	-0.30	-0.08	-0.16	-0.04	122.78
GRC2	-3.30	2.63	-0.10	0.13	-0.12	0.12	-0.01	-0.13	121.55

TABLE XXIII. The error of the variance with  $\epsilon=0.01$ . The result of MP and PRC is very good, except for  $X_5$  and  $X_7$  where the errors are still large for both methods. The error of the variance of GRC1 is still very large here. The result of GRC2 is more accurate than others.

Method	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$
Tauleap	-168 088.49	-103 188.48	-209.30	-105.55	104.74	-216.88	110.03	-174.23
MP	-57 954.72	-47 756.14	7.06	25.45	189.52	-7.91	210.15	24.66
PRC	-51 985.41	-41 925.67	21.98	38.21	190.80	11.16	210.75	46.35
GRC1	149 537.88	109 554.75	235.22	186.80	53.02	283.72	77.46	283.88
GRC2	13 307.03	8857.88	22.18	13.74	-12.96	23.66	-22.98	14.57

TABLE XXIV. The error of the mean with  $\epsilon=0.005$ . Here all the methods can capture the mean very well, except for tauleap where there is still some noticeable error in  $X_1$ .

Method	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	Steps
Tauleap	-6.97	1.99	-0.29	0.10	0.03	0.17	0.10	-0.20	398.01
MP	5.15	-5.55	0.20	-0.15	-0.02	-0.28	-0.09	0.24	397.86
PRC	1.65	-2.30	0.08	-0.03	-0.03	-0.15	-0.03	0.06	397.90
GRC1	0.79	-1.16	0.05	0.00	-0.08	-0.07	-0.06	0.06	397.56
GRC2	2.75	-3.13	0.12	-0.05	-0.04	-0.18	-0.04	0.09	397.51



TABLE XXV. The error of the variance with  $\epsilon=0.005$ . Here GRC1 and GRC2 achieve relatively the same accuracy, which is much better than the other three methods.

Method	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$
Tauleap	-60 025.15	-39 320.80	-76.77	-42.25	21.32	-91.97	21.24	-72.42
MP	-13 060.29	-16 769.64	-2.31	-4.59	45.73	-10.69	50.52	-2.51
PRC	-16 376.93	-14 449.99	-4.37	3.29	45.81	-8.75	52.15	8.44
GRC1	2340.39	-3872.82	3.16	-2.34	1.89	1.00	2.62	-0.94
GRC2	-2252.49	-4748.28	-4.86	-2.71	-0.93	-7.51	-1.98	-5.27

$$\mathbb{E}_x[\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n}] = \tau \sum_{j=1}^M \mathbf{v}_j a_j + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j a_k \eta_{jk} + \mathcal{O}(\tau^3) \quad (\text{A2})$$

for the mean and

$$\begin{aligned} \mathbb{E}_x[(\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n})^2] &= \tau \sum_{j=1}^M \mathbf{v}_j^2 a_j + \left( \tau \sum_{j=1}^M \mathbf{v}_j a_j \right)^2 + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j^2 a_k \eta_{jk} \\ &+ \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k a_j \eta_{kj} + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k a_k \eta_{jk} + \mathcal{O}(\tau^3) \end{aligned} \quad (\text{A3})$$

for the second moment.

## APPENDIX B: PROOF OF PROPOSITION 1

To prove the second order consistency for the mean, we just need to verify that Eq. (5) holds for  $p=1$ ,  $q=2$ . From Eq. (4), we have

$$\begin{aligned} \mathbb{E}_x[\mathbf{X}_{n+1} - \mathbf{X}_n] &= \mathbb{E}_x[\mathbb{E}_r[\mathbf{v} \cdot \mathbf{r} + \mathbf{v} \cdot \tilde{\mathbf{r}}]] \\ &= \mathbb{E}_x[\mathbf{v} \cdot \mathbf{r}] + \mathbb{E}_x[\mathbb{E}_r[\mathbf{v} \cdot \tilde{\mathbf{r}}]] \\ &= \sum_{j=1}^M \mathbf{v}_j \mathbb{E}_x[r_j] + \sum_{j=1}^M \mathbf{v}_j \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]] \\ &= \tau \sum_{j=1}^M \mathbf{v}_j a_j + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j a_k \eta_{jk} + \mathcal{O}(\tau^3). \end{aligned}$$

Comparing it with Eq. (A2) we see that the difference is of  $\mathcal{O}(\tau^3)$ . Thus Eq. (5) has been verified for  $p=1$ ,  $q=2$ , which completes the proof.

## APPENDIX C: PROOF OF PROPOSITION 2

By Definition 1, if the numerical scheme for solving chemical reaction system (1)

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{v} \cdot \mathbf{r}^*$$

is second order consistent for the covariance of  $\mathbf{X}$ , then

$$\text{Cov}_x[\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n}] = \text{Cov}_x[\mathbf{X}_{n+1} - \mathbf{X}_n] + \mathcal{O}(\tau^3).$$

On the one hand, from Eqs. (A2) and (A3), we have

$$\begin{aligned} \text{Cov}_x[\mathbf{X}_{t_n+\tau} - \mathbf{X}_{t_n}] &= \tau \sum_{j=1}^M \mathbf{v}_j^2 a_j + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j^2 a_k \eta_{jk} + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k a_j \eta_{kj} \\ &+ \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k a_k \eta_{jk} + \mathcal{O}(\tau^3) \equiv \Lambda[P + Q]\Lambda^T + \mathcal{O}(\tau^3), \end{aligned}$$

where  $\Lambda = (\mathbf{v}_1, \dots, \mathbf{v}_M)$  is a  $N \times M$  matrix,

$$P = \text{diag} \left( \tau a_1 + \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{1k}, \dots, \tau a_M + \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{Mk} \right)$$

is a  $M \times M$  diagonal matrix and

$$Q = \begin{pmatrix} \tau^2 a_1 \eta_{11} & \frac{\tau^2}{2} (a_2 \eta_{12} + a_1 \eta_{21}) & \dots & \frac{\tau^2}{2} (a_M \eta_{1M} + a_1 \eta_{M1}) \\ \frac{\tau^2}{2} (a_2 \eta_{12} + a_1 \eta_{21}) & \ddots & \dots & \vdots \\ \vdots & \dots & \ddots & \vdots \\ \frac{\tau^2}{2} (a_M \eta_{1M} + a_1 \eta_{M1}) & \dots & \dots & \tau^2 a_M \eta_{MM} \end{pmatrix}$$

is a  $M \times M$  symmetric matrix. On the other hand, by using the condition that  $r_j^*$  and  $r_k^*$  are mutually independent for  $j \neq k$ , we obtain

$$\begin{aligned} \text{Cov}_x[\mathbf{X}_{n+1} - \mathbf{X}_n] &= \text{Cov}_x \left[ \sum_{j=1}^M \mathbf{v}_j r_j^* \right] \\ &= \sum_{j=1}^M \mathbf{v}_j^2 \text{Var}_x[r_j^*] = \Lambda D \Lambda^T, \end{aligned}$$

where

$$D = \text{diag}(\text{Var}_x[r_1^*], \dots, \text{Var}_x[r_M^*])$$

is an  $M \times M$  diagonal matrix. So to achieve second order consistency, we need,

$$\Lambda[P + Q - D]\Lambda^T = \mathcal{O}(\tau^3).$$

This is a system of  $N \times N$  equations with  $(\text{Var}_x[r_1^*], \dots, \text{Var}_x[r_M^*])$  being the  $M$  unknowns. Since  $N \times N$  may be greater than  $M$ , in general, there could be no solution which means second order consistency cannot be achieved. Another derivation is that if the diagonal matrix  $D$  has expansion  $D = \tau D_1 + \tau^2 D_2 + \mathcal{O}(\tau^3)$ , then for general choices of  $\Lambda$  there will be no solution for the matrix equations obtained by comparing each order of  $\tau$ . This completes the proof.

#### APPENDIX D: PROOF OF PROPOSITION 3

From Remark 2, we need to check Eq. (5) for  $p=1$  and 2 with  $q=2$ . The case of  $p=1$  has already been proved in Proposition 1 if condition 1 holds. So the proof would be finished if we can check

$$\mathbb{E}_x[(\mathbf{X}_{n+1} - \mathbf{X}_n)^2] = \mathbb{E}_x[(\mathbf{X}_{n+\tau} - \mathbf{X}_n)^2] + \mathcal{O}(\tau^3). \quad (\text{D1})$$

On the left hand side,

$$\begin{aligned} \mathbb{E}_x[(\mathbf{X}_{n+1} - \mathbf{X}_n)^2] &= \mathbb{E}_x \left[ \left( \sum_{j=1}^M \mathbf{v}_j r_j + \sum_{j=1}^M \mathbf{v}_j \tilde{r}_j \right)^2 \right] \\ &= \sum_{j=1}^M \mathbf{v}_j^2 \mathbb{E}_x[r_j^2] + \sum_{j=1}^M \sum_{k=1, k \neq j}^M \mathbf{v}_j \mathbf{v}_k \mathbb{E}_x[r_j r_k] \\ &\quad + \sum_{j=1}^M \mathbf{v}_j^2 \mathbb{E}_x[\tilde{r}_j^2] + \sum_{j=1}^M \sum_{k=1, k \neq j}^M \mathbf{v}_j \mathbf{v}_k \mathbb{E}_x[\tilde{r}_j \tilde{r}_k] \\ &\quad + \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k \mathbb{E}_x[r_j \tilde{r}_k] + \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k \mathbb{E}_x[r_k \tilde{r}_j]. \end{aligned}$$

(D2) which is just the first condition.

Substituting  $r_j = \mathcal{P}_j(a_j \tau)$  and applying  $\mathbb{E}_x[r_j r_k] = \mathbb{E}_x[r_j] \mathbb{E}_x[r_k]$  for  $j \neq k$ , the second line of Eq. (D2) becomes

$$\tau \sum_{j=1}^M \mathbf{v}_j^2 a_j + \left( \tau \sum_{j=1}^M \mathbf{v}_j a_j \right)^2.$$

By condition 2, the second term in the third line is of the order of  $\mathcal{O}(\tau^3)$  and can be dropped off. Substituting conditions 3 and 4 into the above equation, we obtain

$$\begin{aligned} \mathbb{E}_x[(\mathbf{X}_{n+1} - \mathbf{X}_n)^2] &= \tau \sum_{j=1}^M \mathbf{v}_j^2 a_j + \left( \tau \sum_{j=1}^M \mathbf{v}_j a_j \right)^2 + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j^2 a_k \eta_{jk} \\ &\quad + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k a_j \eta_{kj} + \frac{\tau^2}{2} \sum_{j=1}^M \sum_{k=1}^M \mathbf{v}_j \mathbf{v}_k a_k \eta_{jk} + \mathcal{O}(\tau^3). \end{aligned}$$

Comparing it with Eq. (A3), we see that all the  $\mathcal{O}(\tau)$  and  $\mathcal{O}(\tau^2)$  terms are identical, which completes the proof.

#### APPENDIX E: VERIFY CONDITIONS IN PROPOSITION 3 FOR GRC1

Here we will verify that under conditions (10) and (11),  $\tilde{r}$  satisfy all the conditions in Proposition 3.

*Condition 1.*  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]] = (\tau^2/2) \sum_{k=1}^M a_k \eta_{jk}$ .

Taking expectation  $\mathbb{E}_x$  on both sides of Eq. (10), we have

$$\begin{aligned} \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]] &= \mathbb{E}_x \left[ \frac{\tau}{2} \sum_{k=1}^M r_k \eta_{jk} \right] + \mathbb{E}_x \left[ \frac{\tau}{2} \sum_{\eta_{jk} < 0} \frac{a_k}{a_j} r_j \eta_{jk} \right] \\ &\quad - \frac{\tau}{2} \sum_{\eta_{jk} < 0} \tau a_k \eta_{jk}. \end{aligned}$$

Since  $a_j$  and  $\eta_{jk}$  depend only on  $\mathbf{X}$ , when taking expectation  $\mathbb{E}_x$  they can be treated as constants, which gives

$$\begin{aligned} \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]] &= \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{jk} + \frac{\tau^2}{2} \sum_{\eta_{jk} < 0} a_k \eta_{jk} - \frac{\tau^2}{2} \sum_{\eta_{jk} < 0} a_k \eta_{jk} \\ &= \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{jk}, \end{aligned}$$

Condition 2. For  $j \neq k$ ,  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j \tilde{r}_k]] = \mathcal{O}(\tau^3)$ .

$$\begin{aligned} \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j \tilde{r}_k]] &= \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j] \mathbb{E}_r[\tilde{r}_k]] = \frac{\tau^2}{4} \left( \sum_{l=1}^M \sum_{m=1}^M \eta_{jl} \eta_{km} \mathbb{E}_x[r_l r_m] + \sum_{l=1}^M \sum_{\eta_{km} < 0} \frac{a_m}{a_k} \eta_{jl} \eta_{km} \mathbb{E}_x[r_l r_k] - \sum_{l=1}^M \sum_{\eta_{km} < 0} \tau a_m \eta_{jl} \eta_{km} \mathbb{E}_x[r_l] \right. \\ &+ \sum_{\eta_{jl} < 0} \sum_{m=1}^M \frac{a_l}{a_j} \eta_{jl} \eta_{km} \mathbb{E}_x[r_j r_m] + \sum_{\eta_{jl} < 0} \sum_{\eta_{km} < 0} \frac{a_l a_m}{a_j a_k} \eta_{jl} \eta_{km} \mathbb{E}_x[r_j r_k] - \sum_{\eta_{jl} < 0} \sum_{\eta_{km} < 0} \frac{a_l a_m}{a_j} \tau \eta_{jl} \eta_{km} \mathbb{E}_x[r_j] \\ &\left. - \sum_{\eta_{jl} < 0} \sum_{m=1}^M \tau a_l \eta_{jl} \eta_{km} \mathbb{E}_x[r_m] - \sum_{\eta_{jl} < 0} \sum_{\eta_{km} < 0} \frac{a_l a_m}{a_k} \tau \eta_{jl} \eta_{km} \mathbb{E}_x[r_k] + \sum_{\eta_{jl} < 0} \sum_{\eta_{km} < 0} \tau^2 a_l a_m \eta_{jl} \eta_{km} \right). \end{aligned}$$

It is easy to check that each term on the right hand side in the above equation is at least of the order of  $\mathcal{O}(\tau)$ , thus the second condition is satisfied.

Condition 3. For  $j \neq k$ ,  $\mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_k]] = (\tau^2/2) a_j \eta_{kj} + \mathcal{O}(\tau^3)$ . Substituting Eq. (10), we have

$$\begin{aligned} \mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_k]] &= \frac{\tau}{2} \sum_{l=1}^M \mathbb{E}_x[r_j r_l] \eta_{kl} + \frac{\tau}{2} \sum_{\eta_{kl} < 0} \frac{a_l}{a_k} \mathbb{E}_x[r_j r_k] \eta_{kl} \\ &- \frac{\tau}{2} \sum_{\eta_{kl} < 0} \mathbb{E}_x[r_j] \tau a_l \eta_{kl}. \end{aligned}$$

All the terms are of  $\mathcal{O}(\tau^3)$  except for the first term with  $l=j$ , so it gives

$$\mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_k]] = \frac{\tau^2}{2} a_j \eta_{kj} + \mathcal{O}(\tau^3),$$

which is just the third condition.

Condition 4.  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j^2]] + 2\mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_j]] = (\tau^2/2) \sum_{k=1}^M a_k \eta_{jk} + \tau^2 a_j \eta_{jj} + \mathcal{O}(\tau^3)$ .

It can be checked in the same way as checking  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j \tilde{r}_k]] = \mathcal{O}(\tau^3)$  that  $\mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j^2]] = \mathcal{O}(\tau^3)$ . From Eq. (11) it follows that

$$\begin{aligned} \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j^2]] &= \mathbb{E}_x[\text{Var}_r[\tilde{r}_j]] + \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j]^2] \\ &= \frac{\tau^2}{2} \sum_{k=1}^M a_k |\eta_{jk}| + \mathcal{O}(\tau^3). \end{aligned} \quad (\text{E1})$$

Substituting Eq. (10), we have

$$\begin{aligned} \mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_j]] &= \frac{\tau}{2} \sum_{k=1}^M \mathbb{E}_x[r_j r_k] \eta_{jk} + \frac{\tau}{2} \sum_{\eta_{jk} < 0} \frac{a_k}{a_j} \mathbb{E}_x[r_j^2] \eta_{jk} \\ &- \frac{\tau}{2} \sum_{\eta_{jk} < 0} \mathbb{E}_x[r_j] \tau a_k \eta_{jk} \\ &= \frac{\tau^2}{2} a_j \eta_{jj} + \frac{\tau^2}{2} \sum_{\eta_{jk} < 0} a_k \eta_{jk} + \mathcal{O}(\tau^3). \end{aligned} \quad (\text{E2})$$

With Eqs. (E1) and (E2), we obtain

$$\begin{aligned} \mathbb{E}_x[\mathbb{E}_r[\tilde{r}_j^2]] + 2\mathbb{E}_x[r_j \mathbb{E}_r[\tilde{r}_j]] &= \frac{\tau^2}{2} \sum_{k=1}^M a_k |\eta_{jk}| + \tau^2 a_j \eta_{jj} + \frac{\tau^2}{2} \sum_{k=1}^M a_k (\eta_{jk} - |\eta_{jk}|) + \mathcal{O}(\tau^3) \\ &= \frac{\tau^2}{2} \sum_{k=1}^M a_k \eta_{jk} + \tau^2 a_j \eta_{jj} + \mathcal{O}(\tau^3), \end{aligned}$$

which is exactly the last condition.

- <sup>1</sup>A. Arkin, J. Ross, and H. McAdams, *Genetics* **149**, 1633 (1998).
- <sup>2</sup>H. McAdams and A. Arkin, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 814 (1997).
- <sup>3</sup>D. Endy and R. Brent, *Nature (London)* **409**, 391 (2001).
- <sup>4</sup>M. Elowitz, A. Levine, E. Siggia, and P. Swain, *Science* **297**, 1183 (2002).
- <sup>5</sup>N. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, The Netherlands, 1992).
- <sup>6</sup>D. Gillespie, *Markov Processes: An Introduction for Physical Scientists* (Academic, Boston, 1992).
- <sup>7</sup>D. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).
- <sup>8</sup>D. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- <sup>9</sup>D. Gillespie, *J. Chem. Phys.* **115**, 1716 (2001).
- <sup>10</sup>M. Rathinam, L. Petzold, Y. Cao, and D. Gillespie, *Multiscale Model. Simul.* **4**, 867 (2005).
- <sup>11</sup>T. Li, *Multiscale Model. Simul.* **6**, 417 (2007).
- <sup>12</sup>K. Burrage and T. Tian, Proceedings of the Third International Workshop on Scientific Computing and Applications, City University of Hong Kong, 2003 (unpublished).
- <sup>13</sup>E. Hairer, S. Norsett, and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems* (Springer, Berlin 1993).
- <sup>14</sup>N. Bruti-Liberati and E. Platen, QFRC Research Paper 164, University of Technology, Sydney, Australia, 2005.
- <sup>15</sup>N. Bruti-Liberati and E. Platen, *J. Comput. Appl. Math.* **205**, 982 (2007).
- <sup>16</sup>Y. Cao, personal communication (June, 2006).
- <sup>17</sup>G. Milstein and M. Tretyakov, *Stochastic Numerics for Mathematical Physics* (Springer, Berlin, 2004).
- <sup>18</sup>A. Chatterjee, D. Vlachos, and M. Katsoulakis, *J. Chem. Phys.* **122**, 024112 (2005).
- <sup>19</sup>D. Gillespie and L. Petzold, *J. Chem. Phys.* **119**, 8229 (2003).
- <sup>20</sup>Y. Cao, D. Gillespie, and L. Petzold, *J. Chem. Phys.* **123**, 054104 (2005a).
- <sup>21</sup>Y. Cao, D. Gillespie, and L. Petzold, *J. Chem. Phys.* **124**, 044109 (2006).
- <sup>22</sup>Y. Cao, L. Petzold, and D. Gillespie, *J. Chem. Phys.* **122**, 014116 (2005b).
- <sup>23</sup>W. E, D. Liu, and E. Vanden-Eijnden, *J. Chem. Phys.* **123**, 194107 (2005).
- <sup>24</sup>M. Rathinam, L. Petzold, Y. Cao, and D. Gillespie, *J. Chem. Phys.* **119**, 12784 (2003).
- <sup>25</sup>Y. Cao and L. Petzold, in Proceedings of the AIChE Conference on

Foundations of Systems Biology in Engineering (FOSBE), 2005 (unpublished), pp. 149–152.

<sup>26</sup>M. Rathinam and H. El-Samad, *J. Comput. Phys.* **224**, 897 (2007).

<sup>27</sup>S. Ethier and T. Kurtz, *Markov Processes: Characterization and Conver-*

*gence* (Wiley, New York, 1986).

<sup>28</sup>D. Anderson, *J. Chem. Phys.* **127**, 214107 (2007).

<sup>29</sup>D. Higham, *SIAM Rev.* **50**, 347 (2008).

<sup>30</sup>T. Marquez-Lago and K. Burrage, *J. Chem. Phys.* **127**, 104101 (2007).