A weak second order tau-leaping method for chemical kinetic systems

Yucheng Hu,[a] Tiejun Li,[b] and Bin Min[c]
Laboratory of Mathematics and Applied Mathematics and School of Mathematical Sciences, Peking University, Beijing 100871, People's Republic of China

(Received 20 January 2011; accepted 18 June 2011; published online 14 July 2011)

Recently Anderson and Mattingly [Comm. Math. Sci. **9**, 301 (2011)] proposed a method which can solve chemical Langevin equations with weak second order accuracy. We extend their work to the discrete chemical jump processes. With slight modification, the method can also solve discrete chemical kinetic systems with weak second order accuracy in the large volume scaling. Especially, this method achieves higher order accuracy than both the Euler τ-leaping and mid-point τ-leaping methods in the sense that the local truncation error for the covariance is of order $\tau^3 V^{-1}$ when $\tau = V^{-\beta}$ $(0 < \beta < 1)$ and the system size $V \to \infty$. We present the convergence analysis, numerical stability analysis, and numerical examples. Overall, in the authors' opinion, the new method is easy to be implemented and good in performance, which is a good candidate among the highly accurate τ-leaping type schemes for discrete chemical reaction systems. © *2011 American Institute of Physics*. [doi:10.1063/1.3609119]

I. INTRODUCTION

Researchers are increasingly using discrete stochastic chemical reaction models in cellular biology.[1–4] By treating molecular populations as discrete random variables interacting through stochastic chemical reaction channels, these models often reveal more reasonable and richer behavior than the continuous deterministic approach based on reaction rate equations, which can be viewed as the large-volume limit of the discrete model.[5]

Computationally, simulating the discrete stochastic reaction kinetics is a more challenging task than solving deterministic reaction rate equations.[6] There are methods, such as Gillespie's stochastic simulation algorithm (SSA),[7,8] that can simulate the time evolution of a well-stirred chemically reacting system exactly. But exact methods are too expensive to be applied in practical systems. In 2001, Gillespie proposed an approximate method called τ-leaping that can simulate chemical reaction systems very efficiently.[9] The idea is to lump many reaction events in one update procedure, which can be done by approximating the number of reactions for each channel during a small time interval as a Poisson random variable.

Mathematically, the τ-leaping method can be understood as the Euler-Maruyama scheme for a system of jump stochastic differential equations (SDEs) corresponding to the chemical kinetics.[10,11] These SDEs are different from the conventional SDEs that are driven by Wiener processes in that these are driven by discrete Poisson processes. There are many works trying to improve the accuracy of the τ-leaping, such as the estimated mid-point τ-leaping method proposed by Gillepie in Ref. 9, the Poisson Runge-Kutta method proposed by Burrage and Tian in Ref. 12 and the Gaussian random correction (GRC) method in Ref. 13, etc. The common goal of all these methods is to design higher order methods for solving the SDEs driven by discrete Poisson processes. For convenience, we will call the original τ-leaping scheme as Euler τ-leaping in the later context.

Quite recently, Anderson and Mattingly proposed a method which can solve the chemical Langevin equations with weak second order accuracy.[14] This method does not require the computation of derivatives and is simpler than many other high order schemes solving SDEs. More importantly, they overcome the numerical issue of the square root of the negative propensities introduced by numerical discretization. It has a compact form consisting of two steps: the first is a predicting explicit Euler step and the second is a correction step used to compensate the error made in the first step in order to obtain a higher order, trapezoidal like, approximation to both the drift and diffusion parts.

The main point in this paper is that, with slight modification, the method proposed in Ref. 14 can be successfully applied to discrete chemical reaction systems. The numerical schemes are given in Eqs. (6) and (7) in Sec. III. Our numerical experiences show that the new method is very easy to be used and good in performance. It is more accurate than other methods in our comparative numerical results.

Numerical result is one thing, maybe the most important thing. Theoretical analysis is another thing. As for the theoretical support of the current scheme, we should comment on two convergence concepts under two different scalings for the analysis of τ-leaping methods. One is the convergence under the *traditional scaling* $\tau \to 0$,[10,15] the other is the convergence under the *large volume scaling* when the system size $V \to \infty$.[11] The relation between these two scalings are stated in Ref. 16. Some interesting mathematical results concerning the large volume scaling may be referred to Refs. 11, 17. We mainly take the analysis in the large volume scaling in this paper. But beyond the weak convergence analysis for the compactly supported functions,[11,16] we also investigate the local truncation errors (LTE) for the covariance, which

[a]Electronic mail: huyc@pku.edu.cn.
[b]Electronic mail: tieli@pku.edu.cn.
[c]Electronic mail: minbmath@pku.edu.cn.

is one main motivation of this work. If one only considers the weak convergence for the compactly supported functions in the large volume scaling, the midpoint τ-leaping has already achieved weak second order, which means the current scheme is not an essential improvement. But indeed it improves if one further takes into account the covariance, which is one key point in Ref. 16. We will show that the local truncation error for the covariance of the θ-trapezoidal τ-leaping method is $\mathcal{O}(\tau^3 V^{-1})$ in the large volume scaling, where $\tau = V^{-\beta}$ and $0 < \beta < 1$. While the analysis carried out assumes $\tau = V^{-\beta}$ scaling, numerical evidence suggests that even with constant V the global error, especially for the covariance, is $\mathcal{O}(\tau^2)$ by inspecting the error versus step-size reduction curve in the numerical examples. This result is in contrast with the local truncation error $\mathcal{O}(\tau^2 V^{-1})$ for the covariance of the midpoint τ-leaping method in the large volume scaling, which shows the global first order behavior in the tested numerical examples. The full proof details of the mathematical theorems are left in the Appendix.

The rest of the paper is organized as follows. We will first present some mathematical preliminaries in Sec. II. Then in Sec. III, we formulate the new algorithm for the discrete chemical reaction systems, state the main convergence theorem, and explain why the method works. In Secs. IV and VI, we give the numerical stability analysis and some comparative numerical examples. The effect of the choice of θ on the numerical results are discussed in Sec. VII. Finally we make the conclusion. All of the proof details in Sec. III are left in the Appendix.

II. PRELIMINARY

A. The τ-leaping method

Consider a well-stirred chemical reaction system with N chemical species $\{S_1, \ldots, S_N\}$ interacting through M reaction channels $\{R_1, \ldots, R_M\}$. The system state $\boldsymbol{X}_t = (X_{1t}, \ldots, X_{Nt})^T$ specifies the molecule populations of each reactant specie at time t. For each reaction R_j we know its propensity function $a_j(\boldsymbol{X}_t) \geq 0$ and the state-change vector $\boldsymbol{\nu}_j = (\nu_{1j}, \ldots, \nu_{Nj})^T$. The time evolution of the system state $\boldsymbol{X}_t$ can be described by a discrete state continuous time Markov jump process: in an infinitesimal time dt, the jth reaction will fire with probability $a_j(\boldsymbol{X}_t)dt$, causing a state change of $\boldsymbol{\nu}_j$ to $\boldsymbol{X}_t$. Formally this dynamics can be described by a simple SDEs driven by Poisson processes[5,11]

$$\boldsymbol{X}_t = \boldsymbol{X}_0 + \sum_{j=1}^{M} \boldsymbol{\nu}_j \mathcal{P}_j \left(\int_0^t a_j(\boldsymbol{X}_s) ds \right), \tag{1}$$

where $\mathcal{P}_j$ are independent uni-rate Poisson processes. Another formulation of the SDEs through the Poisson random measure may be referred to Ref. 10, which is a useful set-up for the explanation of the construction of the θ-trapezoidal τ-leaping methods in the continued texts.

Because exact simulation of the above stochastic system is computationally too expensive, in many complex systems one prefers to use an approximate but much more efficient simulation procedure called τ-leaping algorithm, which is proposed by Gillespie in 2001.[9] Mathematically it can be understood as the explicit Euler scheme for the SDEs (1). By freezing the value of $\boldsymbol{X}_t$ during $[t_n, t_n + \tau)$ as $\boldsymbol{X}_{t_n}$, we obtain the increment for each reaction channel

$$\mathcal{P}_j \left(\int_0^{t_n} a_j(\boldsymbol{X}_s) ds + a_j(\boldsymbol{X}_{t_n}) \tau \right) - \mathcal{P}_j \left(\int_0^{t_n} a_j(\boldsymbol{X}_s) ds \right) \sim_d \mathcal{P}_j(a_j(\boldsymbol{X}_{t_n})\tau),$$

where the notation $\sim_d$ means equivalent in distribution sense. Thus we have the following algorithm:

Algorithm 1 (Euler τ-leaping). *Given the state $\boldsymbol{X}_n$ at time t_n and a preselected leap time τ, simulate $\boldsymbol{X}_{n+1}$ by*

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \sum_{j=1}^{M} \boldsymbol{\nu}_j \mathcal{P}_j(a_j(\boldsymbol{X}_n)\tau), \tag{2}$$

where $\mathcal{P}_j(a_j(\boldsymbol{X}_n)\tau)$ are independent Poisson random numbers with parameters $a_j(\boldsymbol{X}_n)\tau$.

In Ref. 9, Gillespie also proposed another scheme, called mid-point τ-leaping algorithm.

Algorithm 2 (mid-point τ-leaping). *Given the state $\boldsymbol{X}_n$ at time t_n and a preselected leap time τ, simulate $\boldsymbol{X}_{n+1}$ by*

$$\begin{cases} \boldsymbol{X}_{n+\frac{1}{2}} = \boldsymbol{X}_n + \sum_{k=1}^{M} \boldsymbol{\nu}_k \frac{1}{2} \tau a_k(\boldsymbol{X}_n), \\ \boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \sum_{k=1}^{M} \boldsymbol{\nu}_k \mathcal{P}_k\left(a_k\left(\boldsymbol{X}_{n+\frac{1}{2}}\right)\tau\right). \end{cases} \tag{3}$$

It was shown that Euler τ-leaping method is of weak first order accuracy for SDEs (1) in both *traditional scaling*[10,15] and *large volume scaling*[11] while the mid-point τ-leaping is of weak second order accuracy in *large volume scaling*.[11] However, further analysis performed by Hu *et al.* in Ref. 16 showed that the local truncation error for the covariance of mid-point τ-leaping is no more accurate than that of Euler τ-leaping. Algorithms to better capture the covariance is still in need.

B. The weak θ-midpoint trapezoidal method

For systems with relatively large molecule populations, the SDEs (1) can be approximated by SDEs driven by Wiener processes, which is called the chemical Langevin equation (CLE),

$$d\boldsymbol{X}_t = \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\boldsymbol{X}_t) dt + \sum_{j=1}^{M} \boldsymbol{\nu}_j \sqrt{a_j(\boldsymbol{X}_t)} dW_t^j, \tag{4}$$

where W_t^j are independent Wiener processes and $\boldsymbol{X}_t \in \mathbb{R}^N$ is now in real space which represents the concentration of the molecules.

Recently, Anderson and Mattingly proposed a weak second order method for the CLE (4). This simple yet novel method is composed of two consequent steps. The first is a predicting explicit Euler step with step size $\theta\tau$, where τ is the objective step size and θ is a preselected parameter whose

value is between 0 and 1. Then in the second step with step size $(1-\theta)\tau$, the information obtained from the first step is used in order to compensate the error already made in the first step. In the end, one obtains a trapezoidal like approximation to both the drift and diffusion parts, thus improve the accuracy to weak second order in solving the CLE (4).

Algorithm 3 (weak θ-midpoint trapezoidal for CLE). *For fixed* $\theta \in (0,1)$, *let*

$$\alpha_1 = \frac{1}{2(1-\theta)\theta}, \quad \text{and} \quad \alpha_2 = \frac{(1-\theta)^2+\theta^2}{2(1-\theta)\theta}.$$

Given $\boldsymbol{X}_n$, *simulate* $\boldsymbol{X}_{n+1}$ *in two continued steps:*

$$\boldsymbol{X}^* = \boldsymbol{X}_n + \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\boldsymbol{X}_n)\theta\tau + \sum_{j=1}^{M} \boldsymbol{\nu}_j \sqrt{a_j(\boldsymbol{X}_n)}\eta_{j,n}^{(1)}\sqrt{\theta\tau},$$

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}^* + \sum_{j=1}^{M} \boldsymbol{\nu}_j (\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))(1-\theta)\tau + \sum_{j=1}^{M} \boldsymbol{\nu}_j \sqrt{(\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))^+}\eta_{j,n}^{(2)}\sqrt{(1-\theta)\tau},$$

where $\eta_{j,n}^{(1)}$, $\eta_{j,n}^{(2)}$ *are two independent standard Gaussian random variables and* $x^+ = \max(0,x)$.

Remark 1. *Anderson and Mattingly showed that the above numerical scheme is of second order weak convergence. The proof requires the replacement of the terms of the form* $(\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))^+$ *with* $\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n)$. *They were able to show that it can be done with an error no more than* $\mathcal{O}(\tau^3)$ *for sufficient small step-size* τ, *that is, for any* $p \geq 1$ *and a bounded random variable* Y,

$$\begin{aligned}
&\mathbb{E}[Y(\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))^+] \\
&= \mathbb{E}[Y(\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))] + \mathcal{O}(\tau^p), \\
&\mathbb{E}[Y(\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))^+(\alpha_1 a_k(\boldsymbol{X}^*) - \alpha_2 a_k(\boldsymbol{X}_n))^+] \\
&= \mathbb{E}[Y(\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n)) \\
&\quad (\alpha_1 a_k(\boldsymbol{X}^*) - \alpha_2 a_k(\boldsymbol{X}_n))] + \mathcal{O}(\tau^p).
\end{aligned}$$

III. A SECOND ORDER τ-LEAPING METHOD

A. Algorithm

With slight modification, the weak θ-midpoint trapezoidal method can be extended to solve the discrete chemical kinetics (1). To do this, one only needs to replace the Wiener increments in Algorithm 3 to Poisson increments. Following Anderson and Mattingly, we call the new method θ-trapezoidal τ-leaping.

Algorithm 4 (θ-trapezoidal τ-leaping). *For fixed* $\theta \in (0,1)$, *let*

$$\alpha_1 = \frac{1}{2(1-\theta)\theta}, \quad \text{and} \quad \alpha_2 = \frac{(1-\theta)^2+\theta^2}{2(1-\theta)\theta}. \tag{5}$$

Given $\boldsymbol{X}_n$, *simulate* $\boldsymbol{X}_{n+1}$ *in two continued steps:*

$$\boldsymbol{X}^* = \boldsymbol{X}_n + \sum_{j=1}^{M} \boldsymbol{\nu}_j \mathcal{P}_j(a_j(\boldsymbol{X}_n)\theta\tau), \tag{6}$$

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}^* + \sum_{j=1}^{M} \boldsymbol{\nu}_j \mathcal{P}_j((\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))^+(1-\theta)\tau). \tag{7}$$

Remark 2. *Similar with Anderson and Mattingly's method, if* $\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n) \geq 0$ *we can show that the scheme is of second order weak consistency in the scaling* $\tau \to 0$. *The details are almost the same as those in Appendix A performed in the large volume scaling, which will be omitted in this paper. However, unlike the CLE case, we cannot get rid of the case of* $\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n) < 0$ *by replacing it with 0, because doing so will cost an error of* $\mathcal{O}(\tau)$ *in traditional scaling* $\tau \to 0$ *for the jump processes. As an extreme example, consider the isomerization reaction* $X \to \emptyset$ *with* $a(x) = x$ *and* $\nu = -1$. *Let* $\theta = 1/2$, *which gives* $\alpha_1 = 2$, $\alpha_2 = 1$. *Note that*

$$X^* = X_n - r,$$

where $r = \mathcal{P}(X_n\tau/2)$. *So we have*

$$\begin{aligned}
&\mathbb{E}[(\alpha_1 a(\boldsymbol{X}^*) - \alpha_2 a(\boldsymbol{X}_n))^+] - \mathbb{E}[(\alpha_1 a(\boldsymbol{X}^*) - \alpha_2 a(\boldsymbol{X}_n))] \\
&= \sum_{i > X_n/2} (2i - \boldsymbol{X}_n)\text{Prob}(r = i).
\end{aligned}$$

If $\boldsymbol{X}_n = 1$, *we have*

$$\sum_{i > X_n/2} (2i - \boldsymbol{X}_n)\text{Prob}(r = i) \geq \text{Prob}(r \neq 0) \sim \mathcal{O}(\tau).$$

So strictly speaking, we cannot prove the second order property of the θ*-trapezoidal* τ*-leaping for chemical reaction systems in traditional scaling. Therefore, we will discuss the convergence property of* θ*-trapezoidal* τ*-leaping in large volume scaling in the following context.*

B. Large volume scaling analysis

To perform the weak consistency analysis for τ-leaping methods in the large volume scaling, we borrow some notations from Ref. 11. Let $\mathbb{L}^V = \{\boldsymbol{y} | \boldsymbol{y} = \boldsymbol{X}/V, \boldsymbol{X} \in \mathbb{Z}^N\}$. We also make the same assumptions as introduced in Refs. 11,16.

Assumption 1. *The following basic assumptions are needed in the large volume scaling*

1. *State variables* $\boldsymbol{X}_t = \mathcal{O}(V)$ *for* $i = 1, \ldots, N$ *where* $V \gg 1$.
2. *Propensities* $a_j(\boldsymbol{X}_t)(j = 1, \ldots, M)$ *are* $\mathcal{O}(V)$ *such that the following scaling holds:*

$$a_j(\boldsymbol{X}_t) = V a_j^V(\boldsymbol{X}_t^V), \boldsymbol{X}_t^V = \boldsymbol{X}_t/V \sim \mathcal{O}(1). \tag{8}$$

3. $a_j^V(\boldsymbol{x})(j = 1, \ldots, M)$ *are* C_0^∞ *functions with compact support* Ω *with respect to* $\boldsymbol{x}$.

One key property of the above scaling is if we define

$$\tau = 1/V^{\beta}, \quad 0 < \beta < 1, \tag{9}$$

since $V \gg 1$, we have

$$\tau \sum_k a_k(\boldsymbol{X}) = 1/V^{\beta} V \sum_k a_k^V(\boldsymbol{x}) = \mathcal{O}(V^{1-\beta}) \gg 1,$$

which satisfies the τ-leaping condition automatically. We will consider the relationship of following two rescaled variables:

$$\boldsymbol{X}_t^V := \boldsymbol{X}_t/V, \quad \boldsymbol{X}_n^V := \boldsymbol{X}_n/V, \tag{10}$$

where $\boldsymbol{X}_n$ refers to θ-trapezoidal τ-leaping method. The two theorems below will clarify this relationship. The proofs for them will be postponed to the Appendix.

Theorem 1 (weak second order accuracy). *Consider the chemical reaction system (1) in $[0, T]$ and partition the interval into k equal parts with step $\tau = T/k$ such as $0 < t_1 < \ldots < t_k = T, t_{n+1} - t_n = \tau$. For $f \in C_0^3(\mathbb{R}^N)$ and $\boldsymbol{x} \in \mathbb{L}^V$, there exist constants $C = C(f, T) > 0$ such that*

$$V^{2\beta} \left| \mathbb{E}_{\boldsymbol{x}} f\left(\boldsymbol{X}_{k\tau}^V\right) - \mathbb{E}_{\boldsymbol{x}} f\left(\boldsymbol{X}_k^V\right) \right| \le C \tag{11}$$

for sufficient small $\tau = V^{-\beta}$ as $V \to \infty$.

Remark 3. *The proof is based on the Theorem 5 in Ref. 16, which states that the weak convergence of τ-leaping type methods rests on the property of one-step approximation. And the property of the one-step approximation, also called weak consistency in Ref. 16, can be routinely checked between the infinitesimal generators of continuous and discrete processes. Details are referred to Appendix A.*

Remark 4. *It is instructive to note the difference between the Theorem 1 and the statement in Remark 2. The key obstruction to prove the second order accuracy in traditional scaling is that the condition $\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n) \ge 0$ will not hold in general. While in the large volume scaling, the extreme case considered in Remark 2 is automatically excluded by assuming $X_t \sim \mathcal{O}(V)$. Combing this condition and the weak second order consistency, we obtain the Theorem 1 in large volume scaling.*

Theorem 2 (LTE of covariance). *The LTE for the covariance of θ-trapezoidal τ-leaping method is as follows:*

$$\left| \mathrm{Cov}_x\left(\boldsymbol{X}_\tau^V\right) - \mathrm{Cov}_x\left(\boldsymbol{X}_1^V\right) \right| = \mathcal{O}(V^{-3\beta-1}) = \mathcal{O}(\tau^3 V^{-1}), \tag{12}$$

in the regime $V \to \infty$, where $\tau = V^{-\beta}$.

Remark 5. *This theorem shows that the LTE of θ-trapezoidal τ-leaping method indeed achieves higher order, compared with the $\mathcal{O}(\tau^2 V^{-1})$ truncation error for midpoint τ-leaping.[16] This property is also confirmed by numerical simulations. The proof is based on the similarity between weak Taylor expansion for chemical kinetic processes and Taylor expansion for reaction rate equations. Details are referred to Appendix B. Notice that we do not prove the corresponding global convergence result for covariance, which will be our future work.*

C. Why the method works

Recall the fact that the GRC τ-leaping proposed in Ref. 13 can achieve second order consistency for both the mean and the covariance. It consists of two steps: the first is an Euler step with step-size τ; then conditioned on the random numbers generated in the first step, a correction variable $\tilde{r}_j$ is added to each reaction channel. So the scheme has the form

$$Step1. \quad \boldsymbol{X}^* = \boldsymbol{X}_n + \sum_{j=1}^M \boldsymbol{\nu}_j \mathcal{P}_j(a_j(\boldsymbol{X}_n)\tau),$$

$$Step2. \quad \boldsymbol{X}_{n+1} = \boldsymbol{X}^* + \sum_{j=1}^M \boldsymbol{\nu}_j \tilde{r}_j.$$

If one chooses $\tilde{r}_j$ as Gaussian random variables satisfying

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

where $\eta_{jk} = a_j(\boldsymbol{X}_n + \boldsymbol{\nu}_k) - a_j(\boldsymbol{X}_n)$, and $\sum_{\eta_{jk}<0}$ means the summation with respect to such k that $\eta_{jk} < 0$, and

$$\mathrm{Var}_{\boldsymbol{r}}[\tilde{r}_j] = \frac{\tau^2}{2} \sum_{k=1}^M a_k |\eta_{jk}| \ge 0,$$

then its local consistency can be second order for the mean and variance.

In fact, both the GRC τ-leaping and the θ-trapezoidal τ-leaping can be written as

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \sum_{j=1}^M \boldsymbol{\nu}_j r_j^*, \quad \boldsymbol{r}^* = \boldsymbol{r} + \tilde{\boldsymbol{r}},$$

where $\boldsymbol{r}$ is the predicting random vector, which is usually the τ-leaping part, and $\tilde{\boldsymbol{r}}$ is the correcting random vector which is *dependent* through $\boldsymbol{r}$. To help understand how the method works, in the same manner as in Ref. 14, we can take another SDE formulation of the chemical kinetic process by Poisson random measure.[10] For simplicity, let us consider only one reaction type with propensity $a(\boldsymbol{X}_t) \le A$ and the state change vector $\boldsymbol{\nu}$ as

$$\boldsymbol{X}_{t+\tau} = \boldsymbol{X}_t + \boldsymbol{\nu} \cdot \int_t^{t+\tau} \int_0^A 1_{[0,a(\boldsymbol{X}_s))} \lambda(ds \times da),$$

where $\lambda(ds \times da)$ is the Poisson random measure with the Lebesgue intensity measure. For the details, the readers may be referred to Ref. 10 and the references therein. In order to approximate the stochastic integral term on the right hand side of the above equation, we must approximate the measure $\lambda(A_{[t,t+\tau)}(a(\boldsymbol{X}_s))$, which counts the number of jumps in the area $A_{[t,t+\tau)}(a(\boldsymbol{X}_s))$. Here $A_{[t,t+\tau)}(a(\boldsymbol{X}_s))$ is the region under the pre-fixed curve $a(\boldsymbol{X}_s)$ for $t \le s \le t + \tau$ with an exact sample trajectory of $\boldsymbol{X}$. In Fig. 1 we can see that (a) the Euler τ-leaping scheme tries to approximate the area of $A_{[t,t+\tau)}(a(\boldsymbol{X}_s))$ with the grey rectangle; (b) the GRC τ-leaping first makes a prediction with the rectangular area then it tries to undo the triangle area (stripe) by adding the correction term. This procedure, however, as mentioned in Ref. 14, is analytically difficult to be carried out, and the

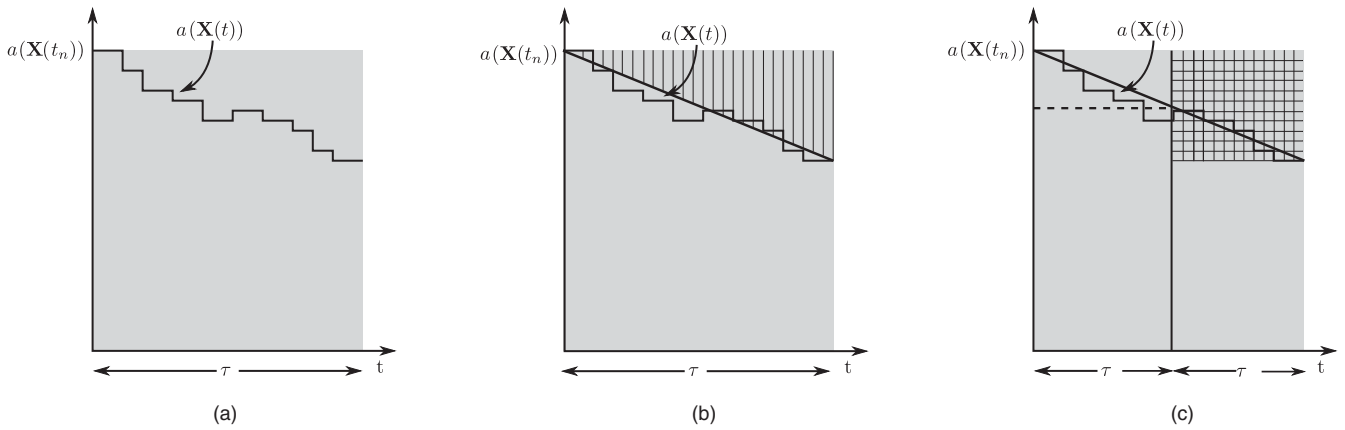

FIG. 1. Graphic representation of different numerical schemes when $\theta = 1/2$. (a) the Euler τ-leaping scheme tries to approximate the measure of $A_{[t,t+\tau)}(a(\boldsymbol{X}_s))$ with the grey rectangles; (b) the GRC τ-leaping first makes a prediction with the grey area then it tries to undo the triangle area (stripe) by adding a correction term, a Gaussian random variable, to match the consistency of mean and variance to second order; (c) the θ-trapezoidal τ-leaping first makes a prediction with step size $\theta\tau$, which is the area of the grey rectangle on the left. Then in the finishing step it removes the grid area and uses the area of the rectangle below. The sum of area of the two rectangles equals to the trapezoidal area under the solid straight line so that it can approximate the area of $A_{[t,t+\tau)}(a(\boldsymbol{X}_s)$ to second order.

GRC τ-leaping did so by using a Gaussian random variable to match the consistency of mean and variance to second order; (c) the θ-trapezoidal τ-leaping first makes a prediction with step size $\theta\tau$, which is the area of the left rectangle. Then in the second step it removes the grid area and uses the area of the right rectangle below. The sum of the area of the two rectangles needs to be equal to the area of the trapezoidal region under the solid line connecting $a(\boldsymbol{X}_{t_n})$ and $a(\boldsymbol{X}^*)$, which gives a higher order approximation to the area of $A_{[t,t+\tau)}(a(\boldsymbol{X}_s))$. But we note here that this visual graphical interpretation only holds when $\theta = 1/2$. For the general case the weak second order still holds and the readers must be referred to Appendix A for more details.

IV. STABILITY ANALYSIS

Following Rathinam *et al.*,[15] we consider the reversible isomerization reaction system as a test model

$$S_1 \rightleftharpoons [c_2]c_1 S_2. \tag{13}$$

Denote X_t the population of molecule S_1, X^T the total population of S_1 and S_2, and $\lambda = c_1 + c_2$. It can be shown that the mean and variance of X_t satisfy the following linear ODEs:

$$\frac{d\mathbb{E}X_t}{dt} = -\lambda \mathbb{E}X_t + c_2 X^T, \tag{14}$$

$$\frac{d\mathrm{Var}X_t}{dt} = -2\lambda \mathrm{Var}X_t + c_2 X^T + (c_1 - c_2)\mathbb{E}X_t. \tag{15}$$

We have the analytical solution as

$$\mathbb{E}X_t = \frac{c_2 X^T}{\lambda}(1 - e^{-\lambda t}) + e^{-\lambda t}\mathbb{E}X_0,$$

$$\mathrm{Var}X_t = \frac{c_1 c_2 X^T}{\lambda^2}(1 - e^{-2\lambda t}) + e^{-2\lambda t}\mathrm{Var}X_0 + \frac{c_1 - c_2}{\lambda}(e^{-\lambda t} - e^{-2\lambda t})\left(\mathbb{E}X_0 - \frac{c_2 X^T}{\lambda}\right).$$

As t goes to infinity, the asymptotic value of the exact mean $\mathbb{E}X_\infty^{(\mathrm{ex})}$ and the exact variance $\mathrm{Var}X_\infty^{(\mathrm{ex})}$ are

$$\mathbb{E}X_\infty^{(\mathrm{ex})} = \frac{c_2 X^T}{\lambda}, \quad \mathrm{Var}X_\infty^{(\mathrm{ex})} = \frac{c_1 c_2 X^T}{\lambda^2}.$$

Next we compute the asymptotic behavior of mean and variance of the numerically simulated value of X_t at time t_n, denoted as X_n, for a certain numerical scheme applied to the test model.

A. Euler τ-leaping

We have $X_{n+1} = X_n - r_1 + r_2$, where $r_1 \sim \mathcal{P}_1(c_1 X_n \tau)$ and $r_2 \sim \mathcal{P}_2(c_2(X^T - X_n)\tau)$. It is straightforward to obtain that

$$\mathbb{E}X_{n+1} = (1 - z)\mathbb{E}X_n + \frac{c_2 X^T}{\lambda} z, \quad z = \lambda\tau.$$

This imposes the stability condition

$$|1 - z| < 1, \text{ which implies } 0 < z < 2 \tag{16}$$

for the step-size, and by letting $n \to \infty$ we obtain

$$\mathbb{E}X_\infty = \frac{c_2 X^T}{\lambda} = \mathbb{E}X_\infty^{(\mathrm{ex})}.$$

For the variance we have

$$\mathrm{Var}X_{n+1} = (1 - z)^2 \mathrm{Var}X_n + \frac{c_2 X^T}{\lambda} z + \frac{(c_1 - c_2)\mathbb{E}X_n}{\lambda} z.$$

The stability domain as required by the variance is still Eq. (16). The asymptotic value of the variance, however, will be amplified by the Euler τ-leaping method,

$$\mathrm{Var}X_\infty = \frac{1}{1 - z/2}\frac{c_1 c_2 X^T}{\lambda^2} = \frac{1}{1 - z/2}\mathrm{Var}X_\infty^{(\mathrm{ex})}. \tag{17}$$

B. GRC τ-leaping

The GRC τ-leaping has the following form:

$$X_{n+1} = X_n - (r_1 + \tilde{r}_1) + (r_2 + \tilde{r}_2), \tag{18}$$

with

$$
\begin{aligned}
r_1 &\sim \mathcal{P}_1(c_1 X_n \tau),\\
r_2 &\sim \mathcal{P}_2(c_2(X^T - X_n)\tau),\\
\tilde{r}_1 &\sim \mathcal{N}\left(\frac{\tau}{2}(r_2 - r_1)c_1 - \frac{\tau}{2} r_1 c_1 \right.\\
&\quad \left. + \frac{\tau^2}{2} c_1^2 X_n,\ \frac{\tau^2}{2}\left[c_1^2 X_n + c_1 c_2 (X^T - X_n)\right]\right),\\
\tilde{r}_2 &\sim \mathcal{N}\left(\frac{\tau}{2}(r_1 - r_2)c_2 - \frac{\tau}{2} r_2 c_2 \right.\\
&\quad \left. + \frac{\tau^2}{2} c_2^2 (X^T - X_n),\ \frac{\tau^2}{2}\left[c_1 c_2 X_n + c_2^2 (X^T - X_n)\right]\right),
\end{aligned}
$$

where $\mathcal{N}(\mu, \sigma^2)$ is the Gaussian random variable with mean μ and variance σ^2.

We have that the mean of the numerical solution satisfies

$$\mathbb{E}X_{n+1} = \left(1 - z + \frac{z^2}{2}\right)\mathbb{E}X_n + \frac{c_2 X^T}{\lambda}\left(z - \frac{z^2}{2}\right).$$

The asymptotic stability of $\mathbb{E}X_n$ requires

$$\left|1 - z + \frac{z^2}{2}\right| < 1, \text{ which again implies } 0 < z < 2 \qquad (19)$$

and when $n \to \infty$ we have

$$\mathbb{E}X_\infty = c_2 X^T/\lambda = \mathbb{E}X_\infty^{(\mathrm{ex})},$$

which is equal to its exact value. For the variance, after some calculations, we have (for simplicity, we already take $\mathbb{E}X_n = \mathbb{E}X_\infty = c_2 X^T/\lambda$ in the following equation, which does not change the amplification factor and the asymptotic variance)

$$
\begin{aligned}
\mathrm{Var}X_{n+1} &= (1 - z + z^2/2)^2 \mathrm{Var}X_n\\
&+ \frac{c_1 c_2 X^T}{\lambda^2} z\left(2 - 2z + 2c_1 c_2 \tau^2 + \frac{5}{4}\tau^2(c_1^2 + c_2^2)\right).
\end{aligned}
$$

Still, the stability domain for the variance is the same as that for the mean. We note that the stability polynomial $1 - z + z^2/2$ for both the mean and variance coincides with the first three terms of the Taylor expansion of e^{-z}, which reflects the second order property of the methods in some sense. The amplifying factor for the asymptotic value of variance is

$$
\begin{aligned}
\mathrm{Var}X_\infty &= \frac{1 - z + \left(\frac{5}{8} - \frac{c_1 c_2}{4\lambda^2}\right)z^2}{1 - z + z^2/2 - z^3/8}\frac{c_1 c_2 X^T}{\lambda^2}\\
&= \frac{1 - z + \left(\frac{5}{8} - \frac{c_1 c_2}{4\lambda^2}\right)z^2}{1 - z + z^2/2 - z^3/8}\mathrm{Var}X_\infty^{(\mathrm{ex})}. \qquad (20)
\end{aligned}
$$

From the simple inequality $c_1 c_2/\lambda^2 \le \frac{1}{4}$, we have

$$\frac{5}{8} - \frac{c_1 c_2}{4\lambda^2} \ge \frac{1}{2},$$

which means that the asymptotic variance is always amplified. This is similar with the Euler τ-leaping method.

C. θ-trapezoidal τ-leaping

The θ-trapezoidal τ-leaping scheme is

$$X_{n+1} = X_n - (r_1 + \tilde{r}_1) + (r_2 + \tilde{r}_2),$$

with

$$
\begin{aligned}
r_1 &\sim \mathcal{P}_1(c_1 X_n \theta\tau),\\
r_2 &\sim \mathcal{P}_2(c_2(X^T - X_n)\theta\tau),\\
\tilde{r}_1 &\sim \mathcal{P}_3((\alpha_1 c_1(X_n - r_1 + r_2) - \alpha_2 c_1 X_n)(1 - \theta)\tau),\\
\tilde{r}_2 &\sim \mathcal{P}_4\Big((\alpha_1 c_2(X^T - (X_n - r_1 + r_2))\\
&\quad - \alpha_2 c_2(X^T - X_n))(1 - \theta)\tau),
\end{aligned}
$$

where the parameters

$$\alpha_1 = \frac{1}{2(1 - \theta)\theta}, \text{ and } \alpha_2 = \frac{(1 - \theta)^2 + \theta^2}{2(1 - \theta)\theta}.$$

The mean of the numerical solution satisfies

$$\mathbb{E}X_{n+1} = \left(1 - z + \frac{z^2}{2}\right)\mathbb{E}X_n + \frac{c_2 X^T}{\lambda}\left(z - \frac{z^2}{2}\right).$$

Thus the stability polynomial is the same as that of GRC τ-leaping and the same asymptotic limit of mean holds

$$\mathbb{E}X_\infty = \frac{c_2 X^T}{\lambda} = \mathbb{E}X_\infty^{(\mathrm{ex})}.$$

For the variance, we have (the asymptotic limit $\mathbb{E}X_n = c_2 X^T/\lambda$ is also taken below for simplicity)

$$
\begin{aligned}
\mathrm{Var}X_{n+1} &= \left(1 - z + \frac{z^2}{2}\right)^2 \mathrm{Var}X_n\\
&+ \frac{c_1 c_2 X^T}{\lambda^2}\left(1 - z + \frac{z^2}{4\theta}\right)2z.
\end{aligned}
$$

The asymptotic value of variance has the form

$$
\begin{aligned}
\mathrm{Var}X_\infty &= \frac{1 - z + z^2/4\theta}{1 - z + z^2/2 - z^3/8}\frac{c_1 c_2 X^T}{\lambda^2}\\
&= \frac{1 - z + z^2/4\theta}{1 - z + z^2/2 - z^3/8}\mathrm{Var}X_\infty^{(\mathrm{ex})}. \qquad (21)
\end{aligned}
$$

It is interesting to observe the similarity between Eqs. (20) and (21). If we take $\theta = 1/2$, we find the asymptotic variance is actually amplified. And if θ is slightly bigger than $1/2$ (as long as $\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n) \ge 0$, see Remark 1), it is possible that the exact asymptotic variance is preserved. This observation is reminiscent of the numerical results in Anderson and Mattingly's work. It is straightforward to find the best parameter θ^* (in the sense of achieving the exact asymptotic variance) will be

$$\theta^* = \frac{1}{2} + \frac{z}{8 - 2z}, \quad 0 < z < 2. \qquad (22)$$

The above equation tells that we should better always choose θ bigger than 1/2. And the rough guide for setting the parameter θ is that if $z = \lambda\tau$ is very small, θ should stay close to 1/2, otherwise bigger θ is preferable. The realistic choices for $z = 1$ or 1.5 will lead to $\theta^* = 0.67$ or 0.8 according to (22). Though the role of the test equation for the chemical kinetic system is not so decisive as that for the ordinary differential

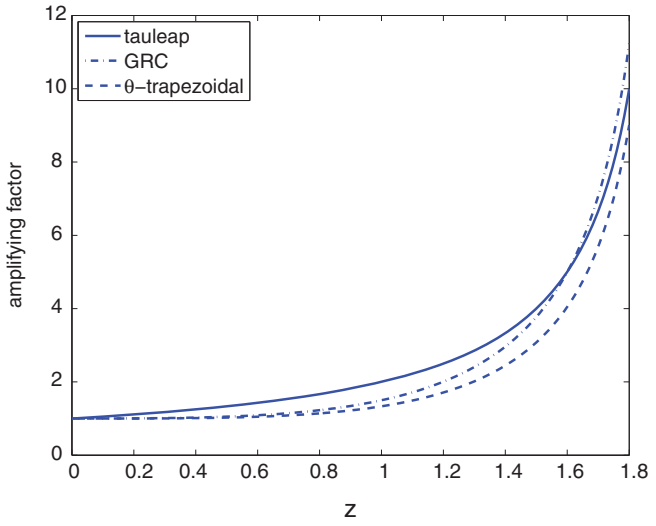

FIG. 2. Amplification factor for the variance of the τ-leaping (17) (solid line), GRC (20) (dashed-dotted line), and θ-trapezoidal τ-leaping (21) (dashed line) when $c_1 = c_2$ and $\theta = 1/2$. The amplification factor for the variance of the θ-trapezoidal method is always under the Euler τ-leaping, indicating the second order property really functions.

equations, we believe the above results provide some insights for the real computations.

The above analysis shows all three methods can obtain right asymptotic numerical mean but the Euler τ-leaping and GRC τ-leaping tend to amplify the variance for the test problem. The θ-trapezoidal τ-leaping has the possibility to give the correct asymptotic variance provided the parameter θ is given suitably. It is also interesting to observe that the amplification factor in (20) explicitly depends on c_1 and c_2 while (21) is not. This indicates that the accuracy of GRC τ-leaping may depend on c_1 and c_2 separately instead of the sum of them. To further illustrate and compare this property for different methods, for the case of $c_1 = c_2$ and $\theta = 1/2$, we plot the theoretical amplification factors against $z = \lambda\tau$ in Fig. 2, which are shown with the solid, dashed-dotted, and dashed lines for the Euler, GRC, and θ-trapezoidal methods, respectively. We can see the amplification factor for the variance of GRC lies in between the Euler τ-leaping and the θ-trapezoidal τ-leaping for small step size, but larger than τ-leaping for large step size. The amplification factor for the variance of the θ-trapezoidal method is always under the Euler τ-leaping, indicating the second order property really functions.

V. COMPARISON WITH IMPLICIT-TRAPEZOIDAL τ-LEAPING SCHEME

The considered θ-trapezoidal τ-leaping scheme is reminiscent of the implicit-trapezoidal τ-leaping method proposed by Cao *et al.*,[18] in which $\boldsymbol{X}_{n+1}$ is obtained by solving the nonlinear equations

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \sum_{j=1}^{M} \nu_j \Big(\mathcal{P}_j(a_j(\boldsymbol{X}_n)\tau) - \frac{1}{2} a_j(\boldsymbol{X}_n)\tau + \frac{1}{2} a_j(\boldsymbol{X}_{n+1})\tau \Big). \tag{23}$$

The idea of the construction of this scheme comes from applying trapezoidal method to the deterministic part of the chemical kinetics.

For the test equation in performing stability analysis, it has been shown that for implicit-trapezoidal τ-leaping method[18]

$$\mathbb{E}X_{n+1} = \frac{2-z}{2+z}\mathbb{E}X_n + \frac{c_2 X^T \tau}{1+z/2}, \tag{24}$$

$$\mathrm{Var}X_{n+1} = \left(\frac{2-z}{2+z}\right)^2 \mathrm{Var}X_n + \frac{1}{(1+z/2)^2} \cdot \left(c_2 X^T \tau + (c_1 - c_2)\tau \mathbb{E}X_n\right). \tag{25}$$

Thus the correct asymptotic mean and variance are kept

$$\mathbb{E}\boldsymbol{X}_\infty = \mathbb{E}\boldsymbol{X}_\infty^{(\mathrm{ex})}, \quad \mathrm{Var}\boldsymbol{X}_\infty = \mathrm{Var}\boldsymbol{X}_\infty^{(\mathrm{ex})}.$$

However, for general nonlinear system this property does not hold.

One interesting thing about the implicit-trapezoidal method for the test equation is that it is second order accurate for the mean and first order accurate for the variance in the traditional scaling. To see this, one can make a direct truncation error analysis based on the recursive relations (24) and (25) and the ODEs (14) and (15). For the mean, the local truncation error is $\mathcal{O}(\tau^3)$. But for the variance, the local truncation error is $\mathcal{O}(\tau^2)$. This result may be responsible for the practical performance of implicit-trapezoidal method as time step-size is relatively small (see the numerical examples in Sec. VI).

To understand the difference between the θ-trapezoidal method and implicit-trapezoidal method more intuitively, we can also explain from the construction of the schemes as in Figure 1. Rewrite (23) as a two step formulation

$$\boldsymbol{X}^* = \boldsymbol{X}_n + \sum_{j=1}^{M} \boldsymbol{\nu}_j \mathcal{P}_j(a_j(\boldsymbol{X}_n)\tau),$$

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}^* + \frac{\tau}{2}\sum_{j=1}^{M} \boldsymbol{\nu}_j (a_j(\boldsymbol{X}_{n+1}) - a_j(\boldsymbol{X}_n)).$$

Though formally it seems that it also admits the trapezoidal interpretation as in Figure 1(c), it is not indeed since the second step is deterministic, which does not help to compensate the variance property of random variables. While the first step is just an Euler τ-leaping step, the overall scheme is only first order for the variance computation.

VI. NUMERICAL EXAMPLES

In this section, we use the Euler τ-leaping method, mid-point τ-leaping method (without truncating the estimated midpoint to integer value[15]), GRC τ-leaping method, θ-trapezoidal τ-leaping method (with $\theta = 1/2$ fixed) and implicit-trapezoidal τ-leaping method to simulate three chemical reaction systems.

We use fixed step-size for each simulation. Although there are many adaptive time step-size control strategies to

gain accuracy and efficiency in solving chemical reaction systems, fixed step-size is more convenient for comparison of the accuracy. In the following three numerical examples, we use step-size $1/2k$ with $k = 1, 2, \cdots, 8$. We compare their statistical properties such as the L^1-norm of histogram distance, the absolute error of central moments. For a general system, usually the large volume scaling is not always maintained in time. In our examples we verify the large volume scaling by inspecting the size of the initial condition and the time history of $a_j(X_t)\delta t$ (see Figure 4). The results show that the new method is of second order accuracy and more accurate and robust than the other methods. Sometimes the implicit-trapezoidal method shows even better performance than the θ-trapezoidal method (see the panels (b) and (d) in Figure 6), but it is not robust and does not show clear order relations. We also notice that the computational effort for implicit-trapezoidal method is large compared with the other explicit methods.

A. System 1: $S \rightarrow 2S$

This system has one reaction with propensity function $a(x) = cx$, where we set the rate constant $c = 1$. The state-change vector is $\nu = 1$. For initial condition $X_0 = 100$, we simulated 10^6 samples until $T = 6$ using the four methods with fixed step-size mentioned above. Note that for this simple system, $a(\boldsymbol{X}_t)\delta t \geq 100/16 \geq 6$, which satisfies the large volume scaling. Figure 3 shows error versus step-size in log-log scale. When an error-curve looks like a straight line, we fit its slope which corresponds to the number in the legend. For this system, it shows that the Euler τ-leaping is even of less than first order accuracy for all of the considered statistics, the mid-point τ-leaping and implicit-trapezoidal method are only second order for the mean, the GRC τ-leaping is second order for both mean and variance and the θ-trapezoidal τ-leaping is clearly more accurate than all the other methods,

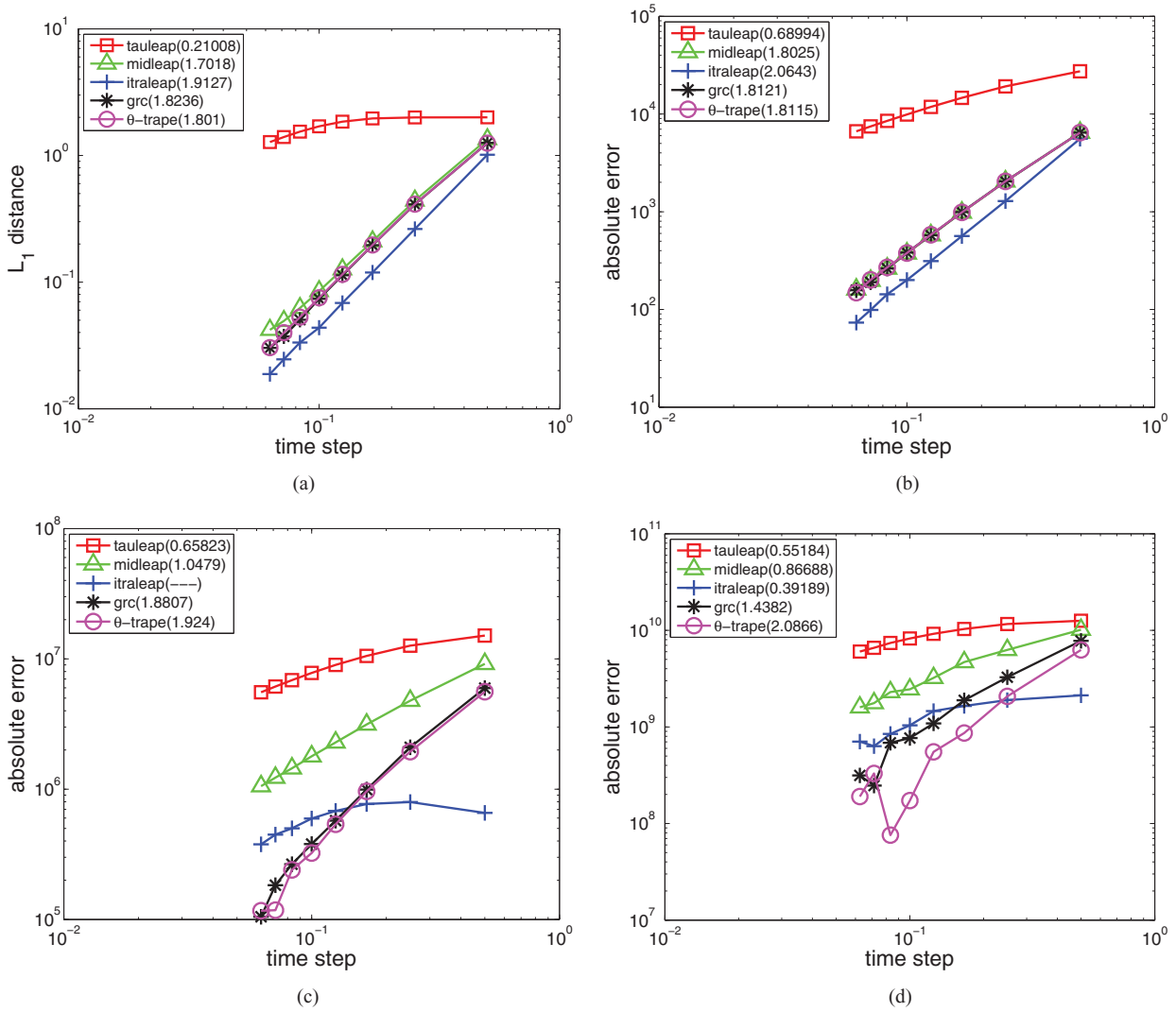

FIG. 3. (System 1) Statistical properties of $X_{T=6}$ drawn from 10^6 independent samples simulated with different methods. The results obtained by using SSA are treated as the exact value. We show the log-log plots of absolute error of (a) L^1-norm of histogram, (b) mean, (c) variance, and (d) third order central moment sampled from results obtained from Euler τ-leaping (red squares), mid-point (green triangles), implicit-trapezoidal (blue plus), GRC (black stars), and θ-trapezoidal (pink circles) using fixed step-sizes as $1/2k$, $k = 1, 2, \cdots, 8$. The numbers in the legend correspond to the fitted slopes if they appear to be straight lines. It shows that the Euler τ-leaping is of first order accuracy, the mid-point is only second order for mean, and the GRC is second order for mean and variance. For the variance computation, the performance of the implicit trapezoidal is similar with those obtained in Ref. 19: it is relatively flat than the other methods though it shows good accuracy with large step-size, but for this system no order relation is shown. From the result for the third-order moment one can tell that θ-trapezoidal is more accurate and robust than the others.

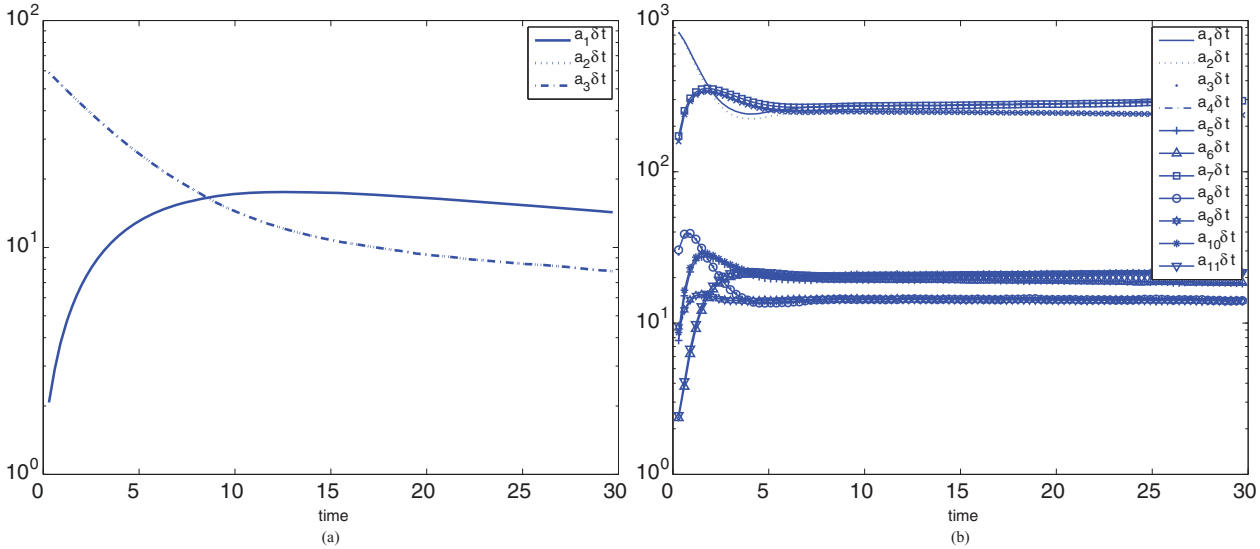

FIG. 4. The left and right panels show the averaged value of $a\delta t$ for each reaction in systems 2 and 3, respectively. Here δt corresponds to the smallest step-size that will be used, which is 1/16. It shows most of the time the value of $a\delta t$ is around 10, which satisfies the large volume scaling.

especially for third-order moments. The performance of the implicit-trapezoidal method is similar with those obtained in Ref. 19: the error versus step-size curve is relatively flat compared with the other methods though it shows good accuracy with large step-size. When time step-size is relatively small, implicit-trapezoidal τ-leaping exhibits first order property for the variance and higher moments. These numerical results are consistent with the analysis in Ref. 16 and the current paper.

B. System 2: 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

$$R_1 : S_1 + S_2 \to S_3$$
$$R_2 : S_3 \to S_1 + S_2$$
$$R_3 : S_3 \to S_2 + S_4.$$

Detailed simulation of this system can be found in Ref. 20. In our implementation, the rate constant $\boldsymbol{c} = (1 \times 10^{-4}, 0.5, 0.5)^T$, and the propensity function is $\boldsymbol{a}(\boldsymbol{x}) = (c_1 x_1 x_2, c_2 x_3, c_3 x_3)^T$. For initial condition $\boldsymbol{X}_0 = (1000, 200, 2000, 0)^T$, we simulated 10^6 samples of $X_{T=6}$ using the four methods with fixed step-size mentioned above. For this example we compared the result of X_1 for different methods. First of all, Fig. 4 shows even for the smallest $\delta t = 1/2k, k = 8$ we used, the large volume limit is still satisfied since $a\delta t$ are around 10 for each reactions. From Fig. 5 We can see that the θ-trapezoidal method gives the most accurate histogram. For the mean it seems the mid-point can give better result than the GRC but for the variance it is even worse than the Euler τ-leaping.

C. System 3: A more complicated system

This system involves 8 species and 12 reactions.[21] The chemical reactions, propensity functions, and initial values used in our implementation are listed in Tables I and II. We simulate 10^6 samples of the system to $T = 30$ for each fixed time step $\tau = 1/2k, k = 1, 2, \ldots, 8$. Figure 4 shows even for the smallest $\delta t = 1/2k, k = 8$ we used, the large volume limit is still satisfied since $a\delta t$ are above 10 for all reactions after a initial transition period. Fig. 6 shows the L_1-norm distance of histograms and error of variance of X_1 and X_5. The mid-point τ-leaping seems to give very accurate result for X_1 but for X_5 the error is big. The implicit-trapezoidal method shows better performance for the computation of X_5 than X_1, but it does not show clear order relations for X_5. For GRC and θ-trapezoidal there is no such unpredictable behavior and it also shows the θ-trapezoidal is very robust and accurate compared with the other methods. The result is similar for the mean, variance, and higher moment, but the statistical fluctuation is more violent for higher moments for this system, so we do not put these results here.

VII. DISCUSSION ON THE CHOICE OF θ

Now let us discuss the effect of choosing different θ to the final results. As observed in Sec. IV about the best θ^* to minimize the deformation of the variance, we will make a numerical check to test our results based on simulating System 2 considered in Sec. IV. But except the accuracy, another important issue is about the degeneracy, which means we should take $(\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n))^+ = 0$ when the left hand side is negative. The more we take this truncation, the bigger the bias will be. Different choices of θ will affect the degeneracy in different degrees.

We take System 2 as the test problem. To check the effect of θ on degeneracy, we fix the step-size as 0.25 and let θ change from 0.1 to 0.9. 10^6 samples are simulated for each θ. In each time step and for each reaction $R_j (j = 1, 2, 3)$, if $\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n) < 0$, a degenerate event will be counted. The result is plotted in the panel (a) of Figure 7. The total number of degenerate events is big when

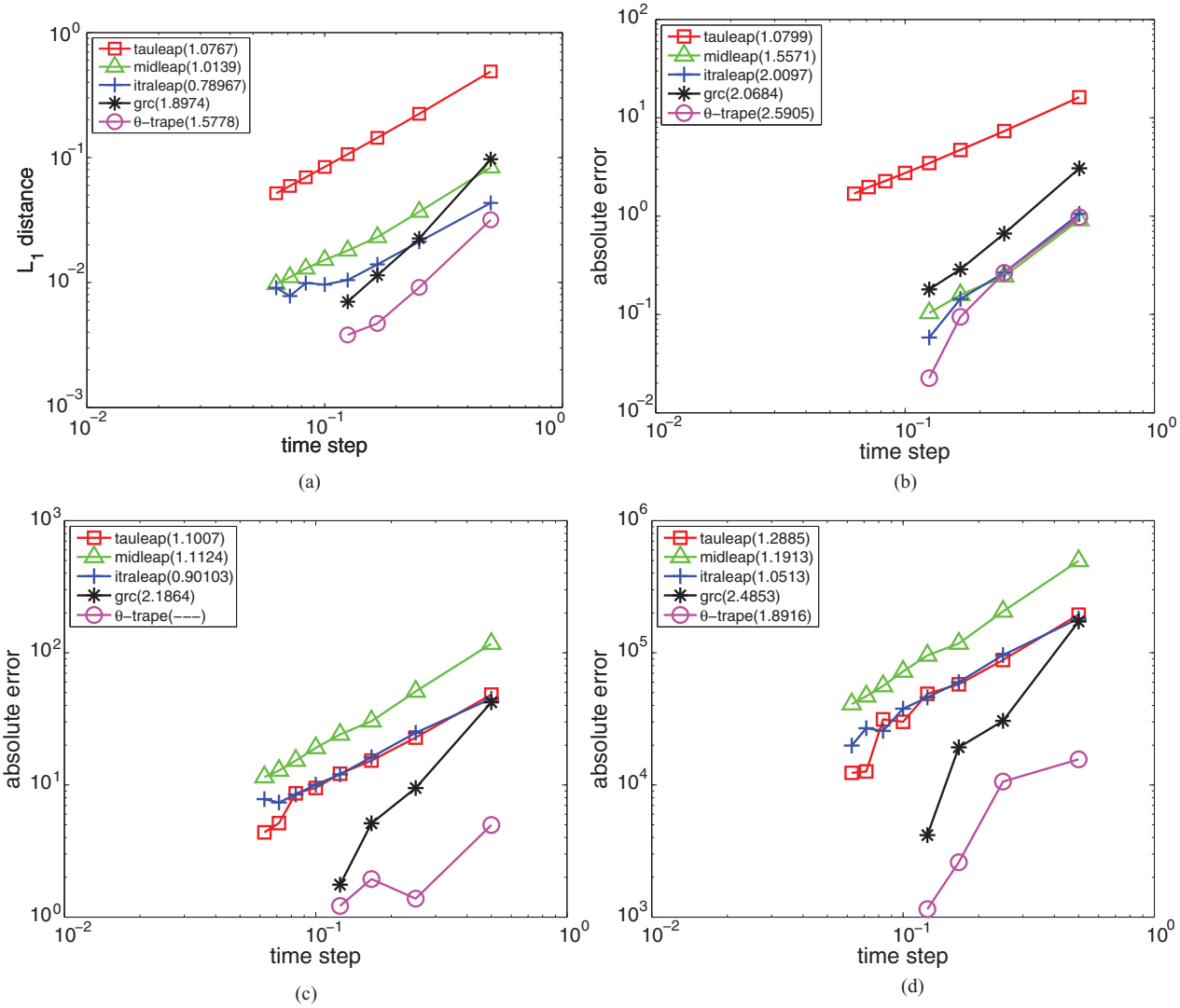

FIG. 5. (System 2) Statistical properties of X_1 at $T = 6$ drawn from 10^6 independent samples simulated with different methods. The results obtained by using SSA are treated as the exact value. We show the log-log plots of absolute error of (a) L^1-norm of histogram, (b) mean, (c) variance, and (d) 4th order central moment sampled from results obtained from Euler τ-leaping (red squares), mid-point (green triangles), implicit-trapezoidal (blue plus) GRC (black stars), and θ-trapezoidal (pink circles) using fixed step-sizes as $1/2k$, $k = 1, 2, \cdots, 8$. For $k \geq 5$ result of GRC and θ-trapezoidal are not shown here because for these small step-size, the error of the scheme is already below the statistical fluctuation of this system. Correspondingly the slopes are fitted using partial data. One can see that the θ-trapezoidal method gives the most accurate histogram. For the mean it seems the mid-point can give better result than the GRC but for the variance it is even worse than the Euler τ-leaping. The reason is not very clear but the numerical convergence order validates our analysis, which is the same as that of Euler τ-leaping.

TABLE I. List of reactions and propensity functions for system 3.

	Reaction	Propensity	Rate constant
1.	$E_A \to E_A + A$	$a_1 = c_1[E_A]$	$c_1 = 15$
2.	$E_B \to E_B + B$	$a_2 = c_3[E_B]$	$c_2 = 15$
3.	$E_A + B \to E_AB$	$a_3 = c_3[E_A][B]$	$c_3 = 0.0001$
4.	$E_AB \to E_A + B$	$a_4 = c_4[E_AB]$	$c_4 = 0.6$
5.	$E_AB + B \to E_AB_2$	$a_5 = c_5[E_AB][B]$	$c_5 = 0.0001$
6.	$E_AB_2 \to E_AB + B$	$a_6 = c_6[E_AB_2]$	$c_6 = 0.6$
7.	$A \to \emptyset$	$a_7 = c_7[A]$	$c_7 = 0.5$
8.	$E_B + A \to E_BA$	$a_8 = c_8[E_B][A]$	$c_8 = 0.0001$
9.	$E_BA \to E_B + A$	$a_9 = c_9[E_BA]$	$c_9 = 0.6$
10.	$E_BA + A \to E_BA_2$	$a_{10} = c_{10}[E_BA][A]$	$c_{10} = 0.0001$
11.	$E_BA_2 \to E_BA + A$	$a_{11} = c_{11}[E_BA_2]$	$c_{11} = 0.6$
12.	$B \to \emptyset$	$a_{12} = c_{12}[B]$	$c_{12} = 0.5$

TABLE II. List of species and their initial value (in number of molecules) for system 3.

	Species	Initial value
1.	A	2000
2.	B	1500
3.	E_A	950
4.	E_B	950
5.	E_AB	200
6.	E_AB_2	50
7.	E_BA	200
8.	E_BA_2	50

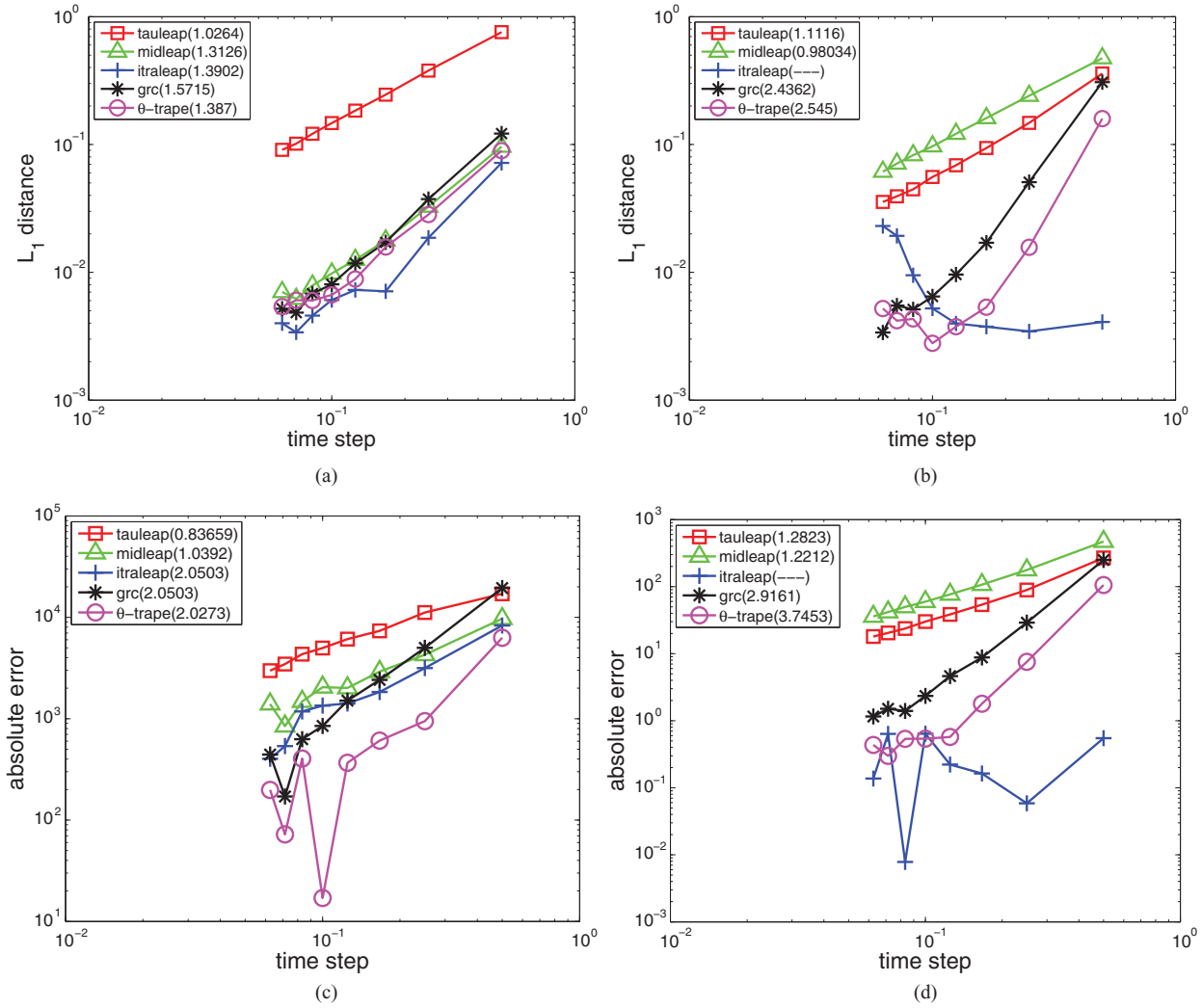

FIG. 6. (System 3) Log-log plots of L_1-norm distance between histogram of X_1 and X_5 sampled from 10^6 independent simulations using different methods with fixed step-sizes $1/2k$, $k = 1, 2, \cdots, 8$. The sample histogram using SSA are treated as exact. Euler τ-leaping (red squares) still shows first order accuracy. Mid-point (green triangles) seems to have very good accuracy for X_1, but at the same time it has the worst accuracy for X_5. The implicit-trapezoidal (blue plus) shows better performance for the computation of X_5 than the others, but it does not show clear order relations for X_5. GRC (black stars) and θ-trapezoidal (pink circles) show higher order convergence for histogram distance, while the latter seems to be slightly more accurate.

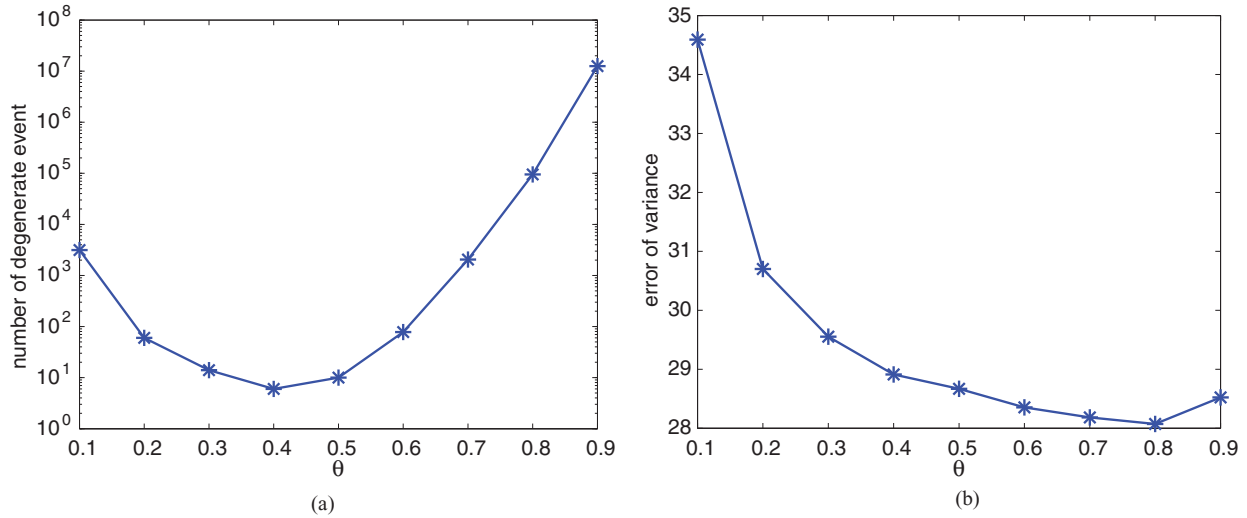

FIG. 7. Effect of θ on degeneracy and accuracy with System 2 as the numerical example. 10^6 samples are simulated for each θ. Panel (a): In each time step and for each reaction $j = 1, 2, 3$, if $\alpha_1 a_j(\boldsymbol{X}^*) - \alpha_2 a_j(\boldsymbol{X}_n) < 0$, a degenerate event will be counted. The total number of degenerate events is big when θ is small or big, indicating the second order property of the θ-trapezoidal scheme is compromised for those choice of θ. Panel (b): The error of variance of X_2 is decreasing with the increasing of θ in general, except for very large θ when the second order property is severely compromised.

θ is small or big, indicating the second order property of the θ-trapezoidal scheme is compromised for those choices of θ.

To check the effect of θ on the accuracy, we take the same setup and look at the tendency for the error of the variance of X_2. The result is shown in the panel (b) of Figure 7. We observe that the error of variance of X_2 is decreasing with the increasing of θ in general, except for very large θ when the second order property is severely compromised. This is consistent with our theoretical estimation in Eq. (21).

Overall, there will be trade off between the accuracy and degeneracy. The theoretical and numerical results suggest us to take θ slightly larger than 0.5. But further numerical studies are needed for the concrete choice strategy for any typical system and parameter setup.

VIII. CONCLUSION

We extend the weak θ-midpoint trapezoidal method proposed by Anderson and Mattingly for solving the chemical Langevin equations to the τ-leaping type methods for the chemical kinetic systems. Different from the midpoint τ-leaping method, which has been proved to be weakly second order in the large volume scaling, we further find that the local truncation error for the covariance is $\mathcal{O}(\tau^3 V^{-1})$ for the θ-trapezoidal τ-leaping method based on the theory in Ref. 16, where $\tau = V^{-\beta}$. This is in contrast with the $\mathcal{O}(\tau^2 V^{-1})$ result for the midpoint method. We proved the second order weak convergence for the compactly supported smooth functions of the random variable $\boldsymbol{X}$ for the new method in the large volume scaling. The numerical examples show that the new method is quite promising. How to extend the current algorithm to higher order weak methods will be an interesting issue for the future study.

ACKNOWLEDGMENTS

The authors acknowledge the anonymous referees for the careful reading of the paper. This work is supported by the Natural Science Foundation of China under Grant No. 10871010.

APPENDIX A: PROOF OF THEOREM 1

To prove Theorem 1, we only need to check the second order weak consistency (Definition 9 in Ref. 16) of θ-trapezoidal method according to Theorem 5 in Ref. 16. We adopt the notations in Ref. 16 here. The second order weak consistency can be described as follows.

The second order weak consistency: *For any* $\boldsymbol{x}_0 \in \mathbb{L}^V$ *and* $g(\boldsymbol{x}) \in D_0^3(\mathbb{R}^N)$ *(Definition* 8 *in Ref. 16), there exist constants* $K > 0$, *independent of* V *and* $\boldsymbol{x}_0$, *such that*

$$\left|\mathbb{E}_{\boldsymbol{x}_0} g\left(\boldsymbol{X}_\tau^V\right) - \mathbb{E}_{\boldsymbol{x}_0} g\left(\boldsymbol{X}_1^V\right)\right| \le K\tau^3, \tag{A1}$$

for sufficiently small $\tau = V^{-\beta}$ *as* $V \to \infty$.

To prove above weak consistency, we first notice that

$$\begin{cases} (\boldsymbol{X}^*)^V = \boldsymbol{x}_0 + \sum_{j=1}^M \dfrac{\boldsymbol{v}_j}{V} \mathcal{P}_j(V a_j^V(\boldsymbol{X}_n^V)\theta\tau), \\ \boldsymbol{X}_{n+1}^V = (\boldsymbol{X}^*)^V + \sum_{j=1}^M \dfrac{\boldsymbol{v}_j}{V} \mathcal{P}_j\left(\left(\alpha_1 V a_j^V((\boldsymbol{X}^*)^V) - \alpha_2 V a_j^V(\boldsymbol{x}_0)\right)^+ (1-\theta)\tau\right). \end{cases} \tag{A2}$$

For simplicity, we denote $\boldsymbol{X}_0^V, (\boldsymbol{X}^*)^V, \boldsymbol{X}_1^V$ in (A2) as $\boldsymbol{x}_0$, $\boldsymbol{y}^*$, $\boldsymbol{y}(\tau)$, respectively, which mimics the notations in Ref. 14. So we have the following schematics:

- the exact solution $\boldsymbol{x}_0 \to [A]\tau \boldsymbol{x}(\tau)$,
- the numerical solution$\boldsymbol{x}_0 \to [B_1]\theta\tau\, \boldsymbol{y}^* \to [B_2](1-\theta)\tau\, \boldsymbol{y}(\tau)$,

where A, B_1, B_2, defined as follows, are infinitesimal generators corresponding to the rescaled form of (1), the first equation in (A2) and the second equation in (A2), respectively. We have

$$(Ag)(\boldsymbol{x}) = \sum_k a_k^V(\boldsymbol{x}) g^{V,k}(\boldsymbol{x}),$$

$$(B_1 g)(\boldsymbol{x}) = \sum_k a_k^V(\boldsymbol{x}_0) g^{V,k}(\boldsymbol{x}),$$

$$(B_2 g)(\boldsymbol{x}) = \sum_k \left[\alpha_1 a_k^V(\boldsymbol{y}^*) - \alpha_2 a_k^V(\boldsymbol{x}_0)\right]^+ g^{V,k}(\boldsymbol{x}),$$

where we take the difference notation

$$g^{V,k}(\boldsymbol{x}) = V[g(\boldsymbol{x} + \boldsymbol{v}_k/V) - g(\boldsymbol{x})]$$

as in Ref. 16. Note that $(Ag)(\boldsymbol{x}_0) = (B_1 g)(\boldsymbol{x}_0)$. For any integer $k \ge 2$ we define $(A^k g)(\boldsymbol{x}) = (A(A^{k-1}g))(\boldsymbol{x})$, and the same for B_1 and B_2. We can define pathwise continuous time extension $\bar{\boldsymbol{X}}_t^V$ ($t \in [0, \theta\tau]$) associated with the infinitesimal generator B_1 as

$$\bar{\boldsymbol{X}}_t^V = \boldsymbol{x}_0 + \sum_{k=1}^M \frac{\boldsymbol{v}_k}{V} \mathcal{P}_k\left(\int_0^t V a_k^V(\boldsymbol{x}_0) dt\right). \tag{A3}$$

We have $\boldsymbol{y}^* = \bar{\boldsymbol{X}}_{\theta\tau}^V$. Similarly,we can define pathwise continuous time extension $\boldsymbol{Y}_t^V$, $t \in [0, (1-\theta)\tau]$ associated with the

infinitesimal generator B_2 as

$$\boldsymbol{Y}_t^V = \boldsymbol{y}^* + \sum_{k=1}^{M} \frac{\boldsymbol{v}_k}{V} \mathcal{P}_k\left(\int_0^t \left[\alpha_1 V a_k^V(\boldsymbol{y}^*) - \alpha_2 V a_k^V(\boldsymbol{x}_0)\right]^+ dt\right).$$

We have $\boldsymbol{y}(\tau) = \boldsymbol{Y}^V_{(1-\theta)\tau}$.

By Ito-Dynkin formula and the condition $g \in D_0^3(\mathbb{R}^N)$, we obtain

$$\mathbb{E}_{\boldsymbol{x}_0} g(\boldsymbol{x}(\tau)) = g(\boldsymbol{x}_0) + (Ag)(\boldsymbol{x}_0)\tau + (A^2 g)(\boldsymbol{x}_0)\frac{\tau^2}{2} + \mathcal{O}(\tau^3), \tag{A4}$$

$$\mathbb{E}_{\boldsymbol{x}_0}[\mathbb{E}_{\theta\tau} g(\boldsymbol{y}(\tau))] = \mathbb{E}_{\boldsymbol{x}_0} g(\boldsymbol{y}^*) + \mathbb{E}_{\boldsymbol{x}_0}(B_2 g)(\boldsymbol{y}^*)(1-\theta)\tau + \mathbb{E}_{\boldsymbol{x}_0}(B_2^2 g)(\boldsymbol{y}^*)\frac{(1-\theta)^2\tau^2}{2} + \mathcal{O}(\tau^3), \tag{A5}$$

where $\mathbb{E}_{\theta\tau}$ means the conditional expectation with respect to fixed $\boldsymbol{y}^*$.

Now we define the modified operator

$$(\tilde{B}_2 g)(\boldsymbol{x}) = \sum_k \left[\alpha_1 a_k^V(\boldsymbol{y}^*) - \alpha_2 a_k^V(\boldsymbol{x}_0)\right] g^{V,k}(\boldsymbol{x}) \tag{A6}$$

which corresponds to remove the positivity truncation in the operator B_2. From Proposition 2 in Appendix C, the condition $g \in D_0^3(\mathbb{R}^N)$ and Assumption 1, we can replace B_2 with $\tilde{B}_2$ in (A5) by introducing the error

$$|\mathbb{E}_{\boldsymbol{x}_0}(B_2 g)(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}(\tilde{B}_2 g)(\boldsymbol{y}^*)| \le C\tau^2,$$
$$|\mathbb{E}_{\boldsymbol{x}_0}(B_2^2 g)(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}(\tilde{B}_2^2 g)(\boldsymbol{y}^*)| \le C\tau.$$

And we have the following results by direct calculations

$$\mathbb{E}_{\boldsymbol{x}_0} g(\boldsymbol{y}^*) = g(\boldsymbol{x}_0) + (B_1 g)(\boldsymbol{x}_0)\theta\tau + \left(B_1^2 g\right)(\boldsymbol{x}_0)\frac{\theta^2\tau^2}{2} + \mathcal{O}(\tau^3). \tag{A7}$$

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}_0}(\tilde{B}_2 g)(\boldsymbol{y}^*) &= \alpha_1[(Ag)(\boldsymbol{x}_0) + \theta\tau(B_1(Ag))(\boldsymbol{x}_0)] \\ &\quad -\alpha_2[(B_1 g)(\boldsymbol{x}_0) + \theta\tau(B_1(B_1 g))(\boldsymbol{x}_0)] \\ &\quad + \mathcal{O}(\tau^2) \\ &= (\alpha_1 - \alpha_2)(Ag)(\boldsymbol{x}_0) + \theta\tau\alpha_1(A^2 g)(\boldsymbol{x}_0) \\ &\quad - \theta\tau\alpha_2(B_1^2 g)(\boldsymbol{x}_0) + \mathcal{O}(\tau^2).\end{aligned} \tag{A8}$$

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}_0}\left(\tilde{B}_2^2 g\right)(\boldsymbol{y}^*) &= \left(\alpha_1^2 - 2\alpha_1\alpha_2 + \alpha_2^2\right) \\ &\quad \times \sum_{j,k} a_j^V(\boldsymbol{x}_0) a_k^V(\boldsymbol{x}_0) g^{V,jk}(\boldsymbol{x}_0) + \mathcal{O}(\tau) \\ &= (\alpha_1 - \alpha_2)^2\left(B_1^2 g\right)(\boldsymbol{x}_0) + \mathcal{O}(\tau).\end{aligned} \tag{A9}$$

Taking advantage of A7, A8, and A9, and $(Ag)(\boldsymbol{x}_0) = (B_1 g)(\boldsymbol{x}_0)$, we get

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}_0} g(\boldsymbol{y}(\tau)) &= g(\boldsymbol{x}_0) + (\theta + (1-\theta)(\alpha_1 - \alpha_2))(Ag)(\boldsymbol{x}_0)\tau \\ &\quad + \alpha_1\theta(1-\theta)(A^2 g)(\boldsymbol{x}_0)\tau^2 \\ &\quad + (\theta^2/2 - \alpha_2\theta(1-\theta) \\ &\quad + (\alpha_1 - \alpha_2)^2(1-\theta)^2/2)(B_1^2 g)(\boldsymbol{x}_0)\tau^2 + \mathcal{O}(\tau^3).\end{aligned}$$

Using the expressions of α_1, α_2 as in (5), we have

$$\theta + (1-\theta)(\alpha_1 - \alpha_2) = 1,$$
$$\alpha_1\theta(1-\theta) = \tfrac{1}{2},$$
$$\frac{\theta^2}{2} - \alpha_2\theta(1-\theta) + (\alpha_1 - \alpha_2)^2\frac{(1-\theta)^2}{2} = 0.$$

Thus finally we obtain

$$\mathbb{E}_{\boldsymbol{x}_0} g(\boldsymbol{y}(\tau)) = g(\boldsymbol{x}_0) + (Ag)(\boldsymbol{x}_0)\tau + (A^2 g)(\boldsymbol{x}_0)\frac{\tau^2}{2} + \mathcal{O}(\tau^3). \tag{A10}$$

This finishes the matching process between (A4) and (A5) until the second order. The proof is completed.

APPENDIX B: PROOF OF THEOREM 2

Define the infinitesimal generator in the large volume scaling as

$$(\mathcal{L}g)(\boldsymbol{x}) = \sum_k a_k^V(\boldsymbol{x}) g^{V,k}(\boldsymbol{x}). \tag{B1}$$

We begin with a lemma concerning with covariance expansion.

Lemma 1. *Assume $\boldsymbol{X}_t^V$ is a Markov process with infinitesimal generator (B1). Then for any $n \in \mathbb{N}$, $f, g \in C^n(\mathbb{R}^N)$, the coefficient of the term τ^n in the weak Taylor expansion of*

$$\mathbb{E}_{\boldsymbol{x}_0}(fg)\left(\boldsymbol{X}_\tau^V\right) - \mathbb{E}_{\boldsymbol{x}_0} f\left(\boldsymbol{X}_\tau^V\right)\mathbb{E}_{\boldsymbol{x}_0} g\left(\boldsymbol{X}_\tau^V\right) \tag{B2}$$

is $\mathcal{O}(V^{-1})$ for sufficient large V. The coefficient may depend on $\boldsymbol{x}_0$ in this setting.

Proof. First we have $\mathcal{L}(fg) = \mathcal{L}f \cdot g + \mathcal{L}g \cdot f + \mathcal{O}(V^{-1})$. Using the above rule recursively, we get for any $n \in \mathbb{N}$,

$$\mathcal{L}^n(fg)(\boldsymbol{x}_0) = \sum_{k=0}^{n} C_n^k \mathcal{L}^{(n-k)} f(\boldsymbol{x}_0) \cdot \mathcal{L}^k g(\boldsymbol{x}_0) + \mathcal{O}(V^{-1}) \tag{B3}$$

which is anolog to the Newton-Leibniz formula in calculus. Then, the coefficient of the term τ^n in the weak Taylor expansion of (B2) takes the form

$$\frac{1}{n!}\mathcal{L}^n(fg) - \sum_{k=0}^{n} \frac{\mathcal{L}^{(n-k)} f}{(n-k)!}\frac{\mathcal{L}^k g}{k!} = \mathcal{O}(V^{-1})$$

by (B3). The proof is completed. □

Lemma 2. *Given $0 < \beta < 1$, $\mathrm{Cov}_{\boldsymbol{x}_0}(\boldsymbol{X}_1^V)$ has the following expansion:*

$$Cov_{\boldsymbol{x}_0}\left(\boldsymbol{X}_1^V\right) = \tau V^{-1}\sum_j a_j^V \boldsymbol{v}_j^{(2)} + \frac{\tau^2}{2} V^{-1}\sum_{j,k} a_j^V a_k^{V,j} \boldsymbol{v}_j^{(2)} + \tau^2 V^{-1}\sum_{j,k} a_k^V a_j^{V,k} \boldsymbol{v}_j \boldsymbol{v}_k + \mathcal{O}(V^{-3\beta-1}),$$

where $\tau = V^{-\beta}$ and the coefficient of the $\mathcal{O}(V^{-3\beta-1})$ term may depend on $\boldsymbol{x}_0$.

Proof. We will take the same notational convention as in Appendix A to denote $\boldsymbol{X}_0^V$, $(\boldsymbol{X}^*)^V$, $\boldsymbol{X}_1^V$ as $\boldsymbol{x}_0$, $\boldsymbol{y}^*$, $\boldsymbol{y}(\tau)$, respectively. According to Proposition 2 in Appendix C, for $g(\boldsymbol{x}) = \boldsymbol{x}$, $\boldsymbol{x}^{(2)}$ the following estimates hold:

$$|\mathbb{E}_{\boldsymbol{x}_0}(B_2 g)(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}(\tilde{B}_2 g)(\boldsymbol{y}^*)| \le C(\boldsymbol{x}_0)\tau^p,$$
$$|\mathbb{E}_{\boldsymbol{x}_0}(B_2^2 g)(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}(\tilde{B}_2^2 g)(\boldsymbol{y}^*)| \le C(\boldsymbol{x}_0)\tau^p,$$

where p can be arbitrarily large. Then by definition and weak Taylor expansion

$$\begin{aligned}\mathrm{Cov}_{\boldsymbol{x}_0}(\boldsymbol{y}(\tau)) &= \mathbb{E}_{\boldsymbol{x}_0}\Big[(\boldsymbol{y}^*)^{(2)} + (1-\theta)\tau B_2(\boldsymbol{x}^{(2)})(\boldsymbol{y}^*) \\ &\quad + \frac{(1-\theta)^2\tau^2}{2} B_2^2(\boldsymbol{x}^{(2)})(\boldsymbol{y}^*)\Big] \\ &\quad - \Big[\mathbb{E}_{\boldsymbol{x}_0}\boldsymbol{y}^* + (1-\theta)\tau\mathbb{E}_{\boldsymbol{x}_0}B_2(\boldsymbol{x})(\boldsymbol{y}^*)\Big]^{(2)} \\ &= \mathbb{E}_{\boldsymbol{x}_0}\Big[(\boldsymbol{y}^*)^{(2)} + (1-\theta)\tau \tilde{B}_2(\boldsymbol{x}^{(2)})(\boldsymbol{y}^*) \\ &\quad + \frac{(1-\theta)^2\tau^2}{2} \tilde{B}_2^2(\boldsymbol{x}^{(2)})(\boldsymbol{y}^*)\Big] \\ &\quad - \Big[\mathbb{E}_{\boldsymbol{x}_0}\boldsymbol{y}^* + (1-\theta)\tau\mathbb{E}_{\boldsymbol{x}_0}\tilde{B}_2(\boldsymbol{x})(\boldsymbol{y}^*)\Big]^{(2)} \\ &\quad + \mathcal{O}(\tau^p), \end{aligned} \tag{B4}$$

where $\boldsymbol{x}^{(2)}$ means the 2-fold tensor product of $\boldsymbol{x}$, and $B_2(\boldsymbol{x}^{(2)})(\boldsymbol{y}^*)$ means the action of the operator B_2 to each component of $\boldsymbol{x}^{(2)}$ then taking value at $\boldsymbol{y}^*$. To prove the expansion of the covariance, we take two steps.

Step 1. Obtain the $\mathcal{O}(\tau)$ and $\mathcal{O}(\tau^2)$ terms explicitly.

Similar with the procedure for proving weak consistency result (A10), we can easily show that the first part in the parenthesis on the right hand side of (B4) equals

$$\boldsymbol{x}_0^{(2)} + \tau(A(\boldsymbol{x}^{(2)}))(\boldsymbol{x}_0) + \frac{\tau^2}{2}(A^2(\boldsymbol{x}^{(2)}))(\boldsymbol{x}_0) + \mathcal{O}(\tau^3)$$

and the second part equals

$$\left[\boldsymbol{x}_0 + \tau(A(\boldsymbol{x}))(\boldsymbol{x}_0) + \frac{\tau^2}{2}(A^2(\boldsymbol{x}))(\boldsymbol{x}_0) + \mathcal{O}(\tau^3)\right]^{(2)},$$

where the coefficients before τ^3 may depend on $\boldsymbol{x}_0$. After some simple calculations, we get

$$\begin{aligned}\mathrm{Cov}_{\boldsymbol{x}_0}(\boldsymbol{y}(\tau)) &= \tau V^{-1}\sum_j a_j^V \boldsymbol{v}_j^{(2)} + \frac{\tau^2}{2}V^{-1}\sum_{j,k} a_j^V a_k^{V,j}\boldsymbol{v}_j^{(2)} \\ &\quad + \tau^2 V^{-1}\sum_{j,k} a_k^V a_j^{V,k}\boldsymbol{v}_j\boldsymbol{v}_k + \mathcal{O}(\tau^3) + \mathcal{O}(\tau^p).\end{aligned}$$

Step 2. Prove the coefficient of the $\mathcal{O}(\tau^3)$ term above is $\mathcal{O}(V^{-1})$.

According to (B4), we have

$$\begin{aligned}\mathrm{Cov}_{\boldsymbol{x}_0}(\boldsymbol{y}(\tau)) &= \big[\mathbb{E}_{\boldsymbol{x}_0}(\boldsymbol{y}^*)^{(2)} - (\mathbb{E}_{\boldsymbol{x}_0}\boldsymbol{y}^*)^{(2)}\big] \\ &\quad + (1-\theta)\tau\big[\mathbb{E}_{\boldsymbol{x}_0}\tilde{B}_2(\boldsymbol{x}^{(2)})(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}\boldsymbol{y}^* \\ &\quad \cdot \mathbb{E}_{\boldsymbol{x}_0}\tilde{B}_2(\boldsymbol{x})(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}\tilde{B}_2(\boldsymbol{x})(\boldsymbol{y}^*)\cdot\mathbb{E}_{\boldsymbol{x}_0}\boldsymbol{y}^*\big] \\ &\quad + \frac{(1-\theta)^2\tau^2}{2}\big[\mathbb{E}_{\boldsymbol{x}_0}\tilde{B}_2^2(\boldsymbol{x}^{(2)})(\boldsymbol{y}^*) \\ &\quad - 2(\mathbb{E}_{\boldsymbol{x}_0}\tilde{B}_2(\boldsymbol{x})(\boldsymbol{y}^*))^{(2)}\big] + \mathcal{O}(\tau^p). \end{aligned} \tag{B5}$$

Now we use Lemma 1 to investigate each term in above expansion. Note that the first part in the parenthesis on the right hand side in (B5) can be expressed as the form shown in the formula (B2) with $f(\boldsymbol{x}) = g(\boldsymbol{x}) = \boldsymbol{x}$, thus Lemma 1 applies. Define the vector functions $f^j(\boldsymbol{x}) = (\alpha_1 a_j^V(\boldsymbol{x}) - \alpha_2 a_j^V(\boldsymbol{x}_0))\boldsymbol{v}_j$, $g(\boldsymbol{x}) = \boldsymbol{x}$. From the fact

$$[(\boldsymbol{x})^{(2)}]^{V,j} = \boldsymbol{v}_j\boldsymbol{x} + \boldsymbol{x}\boldsymbol{v}_j + \boldsymbol{v}_j^{(2)}/V,$$

we have the following representation for the second part:

$$\begin{aligned}&\sum_j\big[\big(\mathbb{E}_{\boldsymbol{x}_0}f^j(\boldsymbol{y}^*)g(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}f^j(\boldsymbol{y}^*)\mathbb{E}_{\boldsymbol{x}_0}g(\boldsymbol{y}^*)\big) \\ &\quad + \big(\mathbb{E}_{\boldsymbol{x}_0}g(\boldsymbol{y}^*)f^j(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}g(\boldsymbol{y}^*)\mathbb{E}_{\boldsymbol{x}_0}f^j(\boldsymbol{y}^*)\big)\big] + \mathcal{O}(V^{-1}).\end{aligned}$$

We can observe that Lemma 1 still applies to the two parts in the parenthesis above, thus the coefficients of τ^n term in the second part of (B5) are also $\mathcal{O}(V^{-1})$.

From the fact

$$[\boldsymbol{x}^{(2)}]^{V,jk} = \boldsymbol{v}_j\boldsymbol{v}_k + \boldsymbol{v}_k\boldsymbol{v}_j,$$

we have the following representation for the third part:

$$2\sum_{jk}\big[\mathbb{E}_{\boldsymbol{x}_0}f^j(\boldsymbol{y}^*)f^k(\boldsymbol{y}^*) - \mathbb{E}_{\boldsymbol{x}_0}f^j(\boldsymbol{y}^*)\mathbb{E}_{\boldsymbol{x}_0}f^k(\boldsymbol{y}^*)\big].$$

Then Lemma 1 applies and the coefficients of τ^n term in the third part of (B5) are also $\mathcal{O}(V^{-1})$.

For $\tau = V^{-\beta}$, where $0 < \beta < 1$, we expand the Taylor series of the covariance until the power $q > 1/\beta + 3$ and take the exponent $p > 1/\beta + 3$ in Eq. (B4) to ensure the remainder term is $\mathcal{O}(\tau^3 V^{-1})$. By combing all of the obtained results above, we get the desired estimate (B4), which completes the proof of the Lemma. □

Thanks to the result for the expansion of $\mathrm{Cov}_{\boldsymbol{x}_0}(\boldsymbol{X}_\tau^V)$ in Ref. 16, we have

$$\begin{aligned}\mathrm{Cov}_{\boldsymbol{x}_0}(\boldsymbol{x}(\tau)) &= \tau V^{-1}\sum_j a_j^V\boldsymbol{v}_j^{(2)} + \frac{\tau^2}{2}V^{-1}\sum_{j,k}a_j^V a_k^{V,j}\boldsymbol{v}_j^{(2)} \\ &\quad + \tau^2 V^{-1}\sum_{j,k}a_k^V a_j^{V,k}\boldsymbol{v}_j\boldsymbol{v}_k + \mathcal{O}(V^{-3\beta-1}).\end{aligned}$$

Together with Lemma 2 we finish the proof of Theorem 2.

APPENDIX C: EFFECT OF REPLACING $[\alpha_1 a_j^V(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0)]^+$ WITH $\alpha_1 a_j^V(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0)$

The proof of Theorem 1 and Theorem 2 requires the replacement of the terms of the form $[\alpha_1 a_j^V(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0)]^+$ with $\alpha_1 a_j^V(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0)$. The following two propositions show that this can be done at the cost of an error $\mathcal{O}(\tau^p)$ where p is an arbitrary large integer. It may be understood as the generalization of the corresponding results in Ref. 14 for chemical kinetic systems in the large volume scaling. At first let us quote the following result, which is exactly the Proposition A.1 in Ref. 14.

Proposition 1. *Let X and Y be two real valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $|XY|_{L^p(\Omega)} < \infty$ for some $p \in (1, \infty]$. Then $|\mathbb{E}Y[X]^+ - \mathbb{E}YX| \le |YX|_{L_p(\Omega)}(\mathbb{P}\{X < 0\})^{1/q}$ where p and q satisfy $1/p + 1/q = 1$. Similarly if X, Y, Z are real valued random variables with $|XYZ|_{L^p(\Omega)} < \infty$ and $A = \{X < 0\} \cup \{Z < 0\}$ then $|\mathbb{E}Y[X]^+[Z]^+ - \mathbb{E}YXZ| \le 2|YXZ|_{L^p(\Omega)}(\mathbb{P}\{A\})^{1/q}$.*

Proposition 2. *Assume Y_0 is a random variable with $|Y_0| < C$ a.s. for some C or $Y_1 = \boldsymbol{y}^*$. Then under the Assumption 1 and $\tau = V^{-\beta}$, for any $p \ge 1$ there exists an τ_0 so that for $i = 0, 1$*

$$\mathbb{E}_{\boldsymbol{x}_0}\left[Y_i(\alpha_1 a_j^V(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0))^+\right]$$

$$= \mathbb{E}_{\boldsymbol{x}_0}\left[Y_i(\alpha_1 a_j(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0))\right] + \mathcal{O}(\tau^p),$$

$$\mathbb{E}_{\boldsymbol{x}_0}\left[Y_i(\alpha_1 a_j^V(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0))^+(\alpha_1 a_k^V(\boldsymbol{y}^*) - \alpha_2 a_k^V(\boldsymbol{x}_0))^+\right]$$

$$= \mathbb{E}_{\boldsymbol{x}_0}\left[Y_i(\alpha_1 a_j^V(\boldsymbol{y}^*) - \alpha_2 a_j^V(\boldsymbol{x}_0))(\alpha_1 a_k^V(\boldsymbol{y}^*) - \alpha_2 a_k^V(\boldsymbol{x}_0))\right] + \mathcal{O}(\tau^p)$$

for all $\tau \in (0, \tau_0]$ and $j, k \in \{1, \ldots, M\}$, where the constant before τ^p term is independent of $\boldsymbol{x}_0$ for Y_0 but dependent of $\boldsymbol{x}_0$ for Y_1.

Proof. Define the event $A_k = \{a_k^V(\boldsymbol{y}^*) < \frac{\alpha_2}{\alpha_1} a_k^V(\boldsymbol{x}_0)\}$ and $\hat{a}_j^V(\boldsymbol{x}) = \alpha_1 a_j^V(\boldsymbol{x}) - \alpha_2 a_j^V(\boldsymbol{x}_0)$. To use Proposition 1, we have the following L^2 bounds for Y_0:

$$|Y_0 \hat{a}_j^V(\boldsymbol{y}^*)|_{L_2(\Omega)} < \infty, \quad |Y_0 \hat{a}_j^V(\boldsymbol{y}^*)\hat{a}_k^V(\boldsymbol{y}^*)|_{L_2(\Omega)} < \infty$$

by Assumption 1. For Y_1, we have

$$\left(\mathbb{E}_{\boldsymbol{x}_0}\left|Y_1 \hat{a}_j^V(\boldsymbol{y}^*)\right|^2\right)^{1/2} \le C(\mathbb{E}_{\boldsymbol{x}_0}|Y_1|)^{1/2},$$

$$\left(\mathbb{E}_{\boldsymbol{x}_0}\left|Y_1 \hat{a}_j^V(\boldsymbol{y}^*)\hat{a}_k^V(\boldsymbol{y}^*)\right|^2\right)^{1/2} \le C(\mathbb{E}_{\boldsymbol{x}_0}|Y_1|)^{1/2},$$

where C are generic constants. By definition of $\boldsymbol{y}^*$ (A3) in Appendix A, we have

$$\mathbb{E}_{\boldsymbol{x}_0}|Y_1|^2 \le C\left(|\boldsymbol{x}_0|^2 + \frac{\theta\tau}{V}\sum_j \left|a_j^V(\boldsymbol{x}_0)\boldsymbol{v}_j\right|^2 + (\theta\tau)^2 \sum_{j,k} a_j^V(\boldsymbol{x}_0)a_k^V(\boldsymbol{x}_0)|\boldsymbol{v}_j||\boldsymbol{v}_k|\right) < \infty.$$

With the above bounds, now it is sufficient to show that for any $p > 1$ there exists a C_p such that $\mathbb{P}_{\boldsymbol{x}_0}(A_k) \le C_p \tau^p$. With Assumption 1 there exists a positive constant C such that

$$a_k^V(\boldsymbol{y}^*) - \frac{\alpha_2}{\alpha_1} a_k^V(\boldsymbol{x}_0) > \left(1 - \frac{\alpha_2}{\alpha_1}\right) a_k^V(\boldsymbol{x}_0) - C|z|,$$

where

$$z = \boldsymbol{y}^* - \boldsymbol{x}_0 = \sum_{j=1}^M \frac{\boldsymbol{v}_j}{V} \mathcal{P}_j(V a_j^V(\boldsymbol{x}_0)\theta\tau).$$

Let $D = (1 - \alpha_2/\alpha_1)a_k^V(\boldsymbol{x}_0)/C$. We know that $D \sim \mathcal{O}(1)$ and

$$\mathbb{P}_{\boldsymbol{x}_0}\{|z| > D\} \le \sum_{j=1}^M \mathbb{P}_{\boldsymbol{x}_0}\left\{\mathcal{P}_j(V a_j^V(\boldsymbol{x}_0)\theta\tau) \ge V\tilde{C}_j\right\},$$

where $\tilde{C}_j \sim \mathcal{O}(1)$, $j = 1, \ldots, M$. Using Chebyshev inequality, we get

$$\mathbb{P}_{\boldsymbol{x}_0}\{|z| > D\} \le \sum_{j=1}^M \frac{\mathbb{E}_{\boldsymbol{x}_0}\left(\mathcal{P}_j(V a_j^V(\boldsymbol{x}_0)\theta\tau)\right)^p}{V^p \tilde{C}_j^p} \sim \mathcal{O}(\tau^p)$$

as $\tau = V^{-\beta}(0 < \beta < 1)$ and $V \to \infty$. Now we complete the proof. □

[1] A. Arkin, J. Ross, and H. McAdams, Genetics **149**, 1633 (1998).
[2] H. McAdams and A. Arkin, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997).
[3] D. Endy and R. Brent, Nature (London) **409**, 391 (2001).
[4] M. Elowitz, A. Levine, E. Siggia, and P. Swain, Science **297**, 1183 (2002).
[5] S. Ethier and T. Kurtz, *Markov Processes: Characterization and Convergence* (Wiley, New York, 1986).
[6] D. Gillespie, *Markov Processes: An Introduction for Physical Scientists* (Academic Press, Boston, USA, 1992).
[7] D. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[8] D. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[9] D. Gillespie, J. Chem. Phys. **115**, 1716 (2001).
[10] T. Li, Multiscale Model. Simul. **6**, 417 (2007).
[11] D. Anderson, A. Ganguly, and T. Kurtz, Ann. Appl. Probab. arXiv: 0909.4790 (2009).
[12] K. Burrage and T. Tian, *Third International Workshop on Scientific Computing and Applications* (City University of Hong Kong, Hong Kong, 2003).
[13] Y. Hu and T. Li, J. Chem. Phys. **130**, 124109 (2009).
[14] D. Anderson and J. Mattingly, Commun. Math. Sci. **9**, 301 (2011).
[15] M. Rathinam, L. Petzold, Y. Cao, and D. Gillespie, Multiscale Model. Simul. **4**, 867 (2005).
[16] Y. Hu, T. Li, and B. Min, "The weak convergence analysis of tau-leaping methods: Revisited," Commun. Math. Sci. (in press).
[17] K. Ball, T. Kurtz, L. Popovic, and G. Rempala, Ann. Appl. Probab. **16**, 1925 (2006).
[18] Y. Cao, L. Petzold, M. Rathinam, and D. Gillespie, J. Chem. Phys. **121**, 12169 (2004).
[19] Y. Cao and L. Petzold, *Proceedings of Foundations of Systems Biology in Engineering (FOSBE 2005)*, 149 (2005).
[20] D. Higham, SIAM Rev. **50**, 347 (2008).
[21] T. Marquez-Lago and K. Burrage, J. Chem. Phys. **127**, 104101 (2007).