Accuracy of lattice Monte Carlo simulation of biased diffusion models

Gang Guo[1,2,*] and Xiaogang Qiu[1]

[1]*College of Information System and Management, National University of Defense Technology, Changsha 410073, China*

[2]*Institute of Computer Science, University of Rostock, Albert Einstein Str. 22, Rostock 18059, Germany*

(Received 30 December 2013; revised manuscript received 2 May 2014; published 27 October 2014)

The accuracy of lattice Monte Carlo (LMC) simulation of biased diffusion models is of great importance as far as the simulation credibility is concerned. It is known that the fixed time step LMC algorithm can reproduce the mean and the variance of the particle displacement exactly for all discrete time steps. Thereby, we propose to use the skewness and other quantities to measure the accuracy. For the one-dimensional fixed time step LMC simulation, we obtain an explicit expression for the skewness and find that the algorithm always produces a negative skewness that converges to zero in the long-time limit when the velocity is positive. It is proved that the skewness is inversely proportional to the square root of the number of simulation steps and the first step error only depends on the Péclet number. We further discuss several other measures of the accuracy of the approximation based on appropriately defined mean-square errors, leading to interesting, unexpected results. The accuracy measures can exhibit complicated nonmonotonic behavior and the optimal step size may depend on the measure of accuracy used.

DOI: 10.1103/PhysRevE.90.043305 PACS number(s): 02.70.Tt, 05.10.Ln, 02.50.−r

I. INTRODUCTION

Biased diffusion models are frequently encountered in numerous physical, chemical, and biological problems [1]. As an effective method, lattice Monte Carlo (LMC) simulation is usually used to solve such kinds of problems when it is difficult to get analytic solutions [2]. In principle, a LMC simulation will arbitrarily approach the solution of continuous systems if its algorithm is correctly designed and the lattice step is small enough [3].

The key to designing a LMC algorithm is to calculate the first passage time (FPT) distribution [4] and the transition probabilities exactly for biased random walkers on discrete lattices. However, to reproduce the diffusion and drift dynamics for a free particle correctly, it is necessary to use a stochastic time step to recover temporal fluctuations. For numerical techniques that require a fixed time step, it is possible to rescale the mean FPT and the transition probabilities to reproduce an equivalent random-walk process. But this is only the case in one and two dimensions. It is impossible to design such a fixed time step LMC algorithm that would reproduce both drift and diffusion correctly in more than two dimensions if the jumps are only allowed along one axis at each time step [5].

More importantly, the accuracy of such LMC simulations plays an import role in reproducing the exact dynamics of biased diffusion processes that involve simultaneous reactions or interactions between particles [6]. The need for accuracy also motivates the development of stochastic microscopic simulations which track the position and state of each particle individually [7]. For unbiased diffusion, the theory of optimizing the accuracy of the fixed time step LMC simulation is extensively studied in Ref. [3]. It is shown that if the time step or the mesh size remains fixed, the accuracy of the algorithm is optimal for a certain finite value of the other parameter. For biased diffusion, it is proved that the fixed time step LMC simulation will reproduce the mean and the variance of the particle displacement exactly for all discrete time steps [5]. Therefore, we propose to use the skewness and other possible quantities to measure the accuracy.

For the purpose of this paper, we will only consider the accuracy of the one-dimensional fixed time step LMC simulation of biased diffusion models. Our objective is to find some quantities to measure the error convincingly and reveal the main factors that determine the accuracy. The rest of the paper is organized as follows. In Sec. II, we give a brief derivation of the one-dimensional fixed time step LMC algorithm. In Sec. III, we obtain an explicit expression for the skewness by calculating the exact moments of the particle distribution at discrete times directly. We show that the behavior of the skewness is characterized by the inverse power law relation. In addition, we discuss several other measures of the accuracy via the root-mean-square error. In Sec. IV, we further confirm our theoretical analysis by numerical results and argue that different quantities may be useful in different contexts. In this sense, the optimal lattice step may depend on the measure of accuracy. Finally, we summarize our work and draw some conclusions.

II. FIXED TIME STEP LMC ALGORITHM

Consider the one-dimensional Fokker-Planck equation with constant drift q and diffusivity D,

$$\frac{\partial}{\partial t}\rho(x,t) = -q\frac{\partial}{\partial x}\rho(x,t) + D\frac{\partial^2}{\partial x^2}\rho(x,t). \tag{1}$$

With the initial condition $\rho(x,t=0) = \delta(x)$ (where δ is the Dirac delta function), the explicit solution is given by

$$\rho(x,t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left[-\frac{(x-qt)^2}{4Dt}\right]. \tag{2}$$

The mean μ, the variance σ^2, and the skewness η are

$$\mu = qt,\ \sigma^2 = 2Dt,\ \eta = 0. \tag{3}$$

For the biased diffusion equation given by Eq. (1), the FPT probability distribution function (PDF) with interval length ℓ

*hndzgg@aliyun.com

can be written down analytically [8],

$$\varphi(t) = \frac{\pi D(1+e^{\ell q/D})}{2\ell^2}\sum_{n=1}^{\infty} n \sin\left(\frac{n\pi}{2}\right) \times \exp\left(-\frac{n^2\pi^2 Dt}{4\ell^2} - \frac{q^2 t + 2\ell q}{4D}\right). \quad (4)$$

The Laplace transform of the above FPT PDF is [9]

$$\mathscr{L}[\varphi(t),s] = \cosh\left(\frac{\ell q}{2D}\right)\mathrm{sech}\left(\frac{\ell\sqrt{q^2+4Ds}}{2D}\right). \quad (5)$$

Expanding this expression around $s = 0$, we have

$$\mathscr{L}[\varphi(t),s] = 1 - \langle\tau\rangle s + \frac{\langle\tau^2\rangle}{2}s^2 + O(s^3). \quad (6)$$

The first two moments of the FPT and the total probability of being absorbed at each end are

$$\langle\tau\rangle = \frac{\ell}{q}\tanh\left(\frac{\ell q}{2D}\right), \quad (7)$$

$$\langle\tau^2\rangle = \frac{\ell}{q^2}\left[2\ell\tanh^2\left(\frac{\ell q}{2D}\right) + \frac{2D}{q}\tanh\left(\frac{\ell q}{2D}\right) - \ell\right], \quad (8)$$

$$p_{\pm} = \left[1+\exp\left(\mp\frac{\ell q}{D}\right)\right]^{-1}. \quad (9)$$

For a random walk with a fixed time step T and a probability p_0 to stay put for each jump, the FPT PDF is given by

$$\phi(t) = \sum_{n=1}^{\infty}(1-p_0)p_0^{n-1}\delta(t-nT). \quad (10)$$

Correspondingly, the Laplace transform is

$$\mathscr{L}[\phi(t),s] = \frac{1-p_0}{e^{sT}-p_0}. \quad (11)$$

Similarly, the expansion around $s = 0$ is given by

$$\mathscr{L}[\phi(t),s] = 1 - \frac{T}{1-p_0}s + \frac{(1+p_0)T^2}{2(1-p_0)^2}s^2 + O(s^3). \quad (12)$$

A comparison of the coefficients between Eq. (6) and Eq. (12) reveals the choice

$$T = (1-p_0)\langle\tau\rangle = \frac{\ell q\coth(\ell q/2D) - 2D}{q^2} \quad (13)$$

and

$$p_0 = \frac{\langle\tau^2\rangle}{\langle\tau\rangle^2} - 1 = \frac{2D}{\ell q}\coth\left(\frac{\ell q}{2D}\right) - \mathrm{csch}^2\left(\frac{\ell q}{2D}\right). \quad (14)$$

This is identical to the results given in Ref. [5]. Note that once the parameters ℓ, q, D are given, both the fixed time step and the probability to stay put are determined. Consequently, the transition probabilities must be rescaled,

$$p'_{\pm} = (1-p_0)p_{\pm} = (1-p_0)\left[1+\exp\left(\mp\frac{\ell q}{D}\right)\right]^{-1}. \quad (15)$$

It is advisable to notice that one can get the following parameters in the limit $q \to 0$:

$$T = \frac{\ell^2}{6D},\ p_0 = \frac{2}{3},\ p'_- = \frac{1}{6},\ p'_+ = \frac{1}{6}. \quad (16)$$

These are exactly the same optimal parameters as derived in Ref. [3] for the fixed time step LMC simulation of unbiased diffusion models.

When there is no diffusion, i.e., in the deterministic limit $D \to 0$, we have

$$T = \frac{\ell}{q},\ p_0 = 0,\ p'_- = 0,\ p'_+ = 1. \quad (17)$$

In this case, the fixed time step LMC simulation will also reproduce the pure drift process exactly for all discrete time steps.

III. ACCURACY ANALYSIS

We will first calculate the moments of the particle distribution to analyze the accuracy of the one-dimensional fixed time step LMC simulation. This can be done by calculating the exact moments of the particle distribution at discrete times directly as in Ref. [3].

For a random walk on lattices starting at site 0, the position after N steps is

$$l_N = \ell\sum_{i=1}^{N}\nu_i, \quad (18)$$

where ν_i is +1 (–1) for a move to the right (left) with probability p'_+ (p'_-) and 0 when the particle does not move with probability p_0.

Since different steps are uncorrelated, the first three moments can be determined by

$$\overline{l_N} = \ell\sum_{i=1}^{N}(\overline{\nu_i}) = N\ell(p'_+ - p'_-), \quad (19)$$

$$\overline{l_N^2} = \ell^2\sum_{i,j=1}^{N}(\overline{\nu_i\nu_j}) = \ell^2\left[\sum_{i=1}^{N}\left(\overline{\nu_i^2}\right) + \sum_{i,j=1,i\neq j}^{N}(\overline{\nu_i\nu_j})\right] = N\ell^2[(p'_+ + p'_-) + (N-1)(p'_+ - p'_-)^2], \quad (20)$$

$$\overline{l_N^3} = \ell^3\sum_{i,j,k=1}^{N}(\overline{\nu_i\nu_j\nu_k}) = \ell^3\left[\sum_{i=1}^{N}\left(\overline{\nu_i^3}\right) + 3\sum_{i,k=1,i\neq k}^{N}\left(\overline{\nu_i^2\nu_k}\right) + \sum_{i,j,k=1,i\neq j,j\neq k,i\neq k}^{N}(\overline{\nu_i\nu_j\nu_k})\right] = N\ell^3[(p'_+ - p'_-) + 3(N-1)(p'_+ + p'_-)(p'_+ - p'_-) + (N-1)(N-2)(p'_+ - p'_-)^3]. \quad (21)$$

Therefore, the variance is given by

$$\overline{(l_N-\overline{l_N})^2}=\overline{l_N^2}-\overline{l_N}^2=N\ell^2[(p'_+ + p'_-)-(p'_+ - p'_-)^2]. \tag{22}$$

The third moment of the displacement is

$$\begin{aligned}\overline{(l_N-\overline{l_N})^3}&=\overline{l_N^3}+2\overline{l_N}^3-3\overline{l_N}\,\overline{l_N^2}\\&=N\ell^3(p'_+ - p'_-)[1-3(p'_+ + p'_-)+2(p'_+ - p'_-)^2].\end{aligned} \tag{23}$$

Now it is not difficult to verify that both the drift and the diffusivity are reproduced exactly for all discrete time steps,

$$\frac{\overline{l_N}}{NT}=\frac{\ell(p'_+ - p'_-)}{T}=q, \tag{24}$$

$$\frac{\overline{(l_N-\overline{l_N})^2}}{2NT}=\frac{\ell^2}{2T}[(p'_+ + p'_-)-(p'_+ - p'_-)^2]=D. \tag{25}$$

This means that it is not appropriate to use the mean or the variance to evaluate the accuracy of the fixed time step LMC simulation. Hence we propose to use the skewness to measure the accuracy.

According to Eqs. (22) and (23), the skewness of the fixed time step LMC simulation at discrete times can be calculated by

$$\begin{aligned}\eta_N&=\overline{(l_N-\overline{l_N})^3}\,[\overline{(l_N-\overline{l_N})^2}]^{\frac{3}{2}}\\&=\frac{(p'_+ - p'_-)+2(p'_+ - p'_-)^3-3(p'_+ - p'_-)(p'_+ + p'_-)}{N^{\frac{1}{2}}[(p'_+ + p'_-)-(p'_+ - p'_-)^2]^{\frac{3}{2}}}.\end{aligned} \tag{26}$$

This proves that the skewness converges to zero in the limit $N\to+\infty$ and the rate of convergence is inversely proportional to $\sqrt{N}$.

To proceed further, let us suppose $q>0$ for simplicity. With the dimensionless Péclet number

$$\xi=\frac{\ell q}{2D}, \tag{27}$$

Eq. (26) reduces to

$$\begin{aligned}\eta_N&=N^{-\frac{1}{2}}\xi^{\frac{3}{2}}\left[\coth(\xi)-\frac{1}{\xi}\right]^{-\frac{1}{2}}\\&\quad\times\left[1-\frac{\coth(\xi)}{\xi}+\frac{2}{\xi^2}-\coth^2(\xi)\right]\\&=N^{-\frac{1}{2}}\xi^{\frac{3}{2}}\left[\coth(\xi)-\frac{1}{\xi}\right]^{-\frac{1}{2}}\\&\quad\times\left[\frac{4\sinh^2(\xi)-\xi\sinh(2\xi)-2\xi^2}{2\xi^2\sinh^2(\xi)}\right].\end{aligned} \tag{28}$$

Now let us only consider the numerator in the last parentheses. One can show that

$$\begin{aligned}\Delta_\xi&=4\sinh^2(\xi)-\xi\sinh(2\xi)-2\xi^2,\quad \Delta_{\xi=0}=0,\\ \dot{\Delta}_\xi&=3\sinh(2\xi)-2\xi\cosh(2\xi)-4\xi,\quad \dot{\Delta}_{\xi=0}=0,\\ \ddot{\Delta}_\xi&=4[\cosh(2\xi)-\xi\sinh(2\xi)-1],\quad \ddot{\Delta}_{\xi=0}=0,\\ \dddot{\Delta}_\xi&=4[\sinh(2\xi)-2\xi\cosh(2\xi)],\quad \dddot{\Delta}_{\xi=0}=0,\\ \ddddot{\Delta}_\xi&=-16\xi\sinh(2\xi)\leqslant 0.\end{aligned} \tag{29}$$

Consequently, it is proved that

$$\dddot{\Delta}_\xi\leqslant 0,\quad \ddot{\Delta}_\xi\leqslant 0,\quad \dot{\Delta}_\xi\leqslant 0,\quad \Delta_\xi\leqslant 0. \tag{30}$$

Finally, it leads to the fact that $\eta_N\leqslant 0$ from which we conclude that the fixed time step LMC simulation always produces a negative skewness for biased diffusion when the velocity $q>0$. Since the skewness of the Gaussian solution of the diffusion problem is identically zero, we can use the nonzero skewness of the fixed time step LMC simulation as a measure of its accuracy.

It also shows that the maximum error introduced in the skewness of the fixed time step LMC simulation is only determined by the Péclet number. The larger the Péclet number is, the larger the maximum error is. This systematic error does not vanish nor decrease as the number of simulation runs increases. However, it converges to zero in the long-time limit. Therefore, to achieve a given tolerance $\varepsilon\ll 1$ in the magnitude of the skewness, we suggest choosing the lattice step ℓ so that $-\varepsilon<\eta_{N=1}<0$ according to Eq. (28). Theoretically, it will be sufficient to guarantee that the error introduced in the skewness is always within the tolerance if the numerical round-off errors are not considered.

As far as the accuracy of the one-dimensional fixed time step LMC simulation is concerned, another straightforward way to obtain a quantitative estimate of the accuracy is defined by

$$\Delta_I(t)=\left(\sum_{i=-\infty}^{+\infty}\left\{\int_{i\ell-\frac{\ell}{2}}^{i\ell+\frac{\ell}{2}}\left[\rho(x,t)-\frac{P(i,t)}{\ell}\right]^2dx\right\}\right)^{\frac{1}{2}}. \tag{31}$$

In this definition, the difference between the continuous distribution [cf. Eq. (2)] and the piecewise-constant discontinuous distribution produced by the fixed time step LMC simulation is measured by the summation of integrals on lattices. Clearly, it does not depend on the length of the interval on which the simulation is run once this length exceeds the width of the distribution significantly. However, a potential problem with this quantity is that it is always nonzero for $\ell>0$, no matter how accurate the algorithm is. A potentially better possibility is using

$$\Delta_{\rm II}(t)=\left\{\sum_{i=-\infty}^{+\infty}\left[\int_{i\ell-\frac{\ell}{2}}^{i\ell+\frac{\ell}{2}}\rho(x,t)dx-P(i,t)\right]^2\right\}^{\frac{1}{2}}. \tag{32}$$

With this quantity, the error can actually be zero even for finite $\ell > 0$, if the number of particles in each bin is correct. One may also argue that each term in the above equation should be further divided by the lattice step and hence propose another quantity,

$$\Delta_{\rm III}(t) = \left\{ \frac{1}{\ell} \sum_{i=-\infty}^{+\infty} \left[\int_{i\ell-\frac{\ell}{2}}^{i\ell+\frac{\ell}{2}} \rho(x,t)dx - P(i,t) \right]^2 \right\}^{\frac{1}{2}} = \frac{\Delta_{\rm II}(t)}{\sqrt{\ell}}. \tag{33}$$

The justification of quantities $\Delta_{\rm II}$ and $\Delta_{\rm III}$ can be briefly given as follows. Suppose for simplicity that there are only n nonzero terms in the sum in Eq. (32) and each of them is equal to a^2. Then $\Delta_{\rm II} = n^{1/2}a$. However, if instead we take the absolute value and remove the square root, then we get $\Delta_{\rm II} = na$, which is quite different. If we are comparing two distributions, then, if we decrease ℓ by a factor of 2, a will decrease by a factor of 2 and n will increase by a factor of 2. So na will not change, and $\Delta_{\rm II}$ does not depend on ℓ, but $\Delta_{\rm III}$ does. However, if $\Delta_{\rm II} = n^{1/2}a$ is considered instead, then this quantity will decrease by a factor of $\sqrt{2}$ and therefore the opposite conclusion is reached: $\Delta_{\rm III}$ does not depend on ℓ while $\Delta_{\rm II}$ does.

The justification of $\Delta_{\rm II}$ is simply that it is a convenient quantity, since it only depends on ξ for the LMC algorithm considered in this paper. First of all, let us consider the quantity $\Delta_{\rm II}$ defined by Eq. (32). According to the analytic solution given by Eq. (2) and the fixed time step given by Eq. (13), we have

$$\begin{aligned}
\Delta_{\rm II}(t = NT+) &= \left\{ \sum_{i=-\infty}^{+\infty} \left[\int_{i\ell-\frac{\ell}{2}}^{i\ell+\frac{\ell}{2}} \rho(x,NT)dx - P(i,NT) \right]^2 \right\}^{\frac{1}{2}} \\
&= \left(\sum_{i=-\infty}^{+\infty} \left\{ \int_{i\ell-\frac{\ell}{2}}^{i\ell+\frac{\ell}{2}} \frac{1}{\sqrt{4\pi NDT}} \exp\left[-\frac{(x-NqT)^2}{4NDT} \right] dx - P(i,NT) \right\}^2 \right)^{\frac{1}{2}} \\
&= \left\{ \sum_{i=-\infty}^{+\infty} \left[\Phi\left(\frac{i\ell + \frac{\ell}{2} - NqT}{\sqrt{2NDT}} \right) - \Phi\left(\frac{i\ell - \frac{\ell}{2} - NqT}{\sqrt{2NDT}} \right) - P(i,NT) \right]^2 \right\}^{\frac{1}{2}} \\
&= \left[\sum_{i=-\infty}^{+\infty} \left(\Phi\left\{ \frac{i + \frac{1}{2} - N\left[\coth(\xi) - \frac{1}{\xi}\right]}{\sqrt{\frac{N}{\xi}\left[\coth(\xi) - \frac{1}{\xi}\right]}} \right\} - \Phi\left\{ \frac{i - \frac{1}{2} - N\left[\coth(\xi) - \frac{1}{\xi}\right]}{\sqrt{\frac{N}{\xi}\left[\coth(\xi) - \frac{1}{\xi}\right]}} \right\} - P(i,NT) \right)^2 \right]^{\frac{1}{2}}, \\
N &= 1,2,3,\ldots,
\end{aligned} \tag{34}$$

where Φ is the cumulative distribution function of the standard normal distribution and the "+" sign in $t = NT+$ indicates that all the quantities are measured after the jump, rather than before.

Upon using the combined generating function and Fourier transform [4], we can obtain

$$P(i,NT) = \sum_{m=0}^{N-|i|} \frac{N!(p_0)^m (p'_+)^{\frac{N-m+i}{2}} (p'_-)^{\frac{N-m-i}{2}}}{m!\left(\frac{N-m+i}{2}\right)!\left(\frac{N-m-i}{2}\right)!}, \quad N = 1,2,3,\ldots. \tag{35}$$

Note that the sum over m in Eq. (35) should only include the values of m of the same parity as $N + i$. According to Eqs. (14) and (15), it is clear that the probabilities to move to the nearest neighbor sites as well as to stay put at each step only depend on the Péclet number. Consequently, it is proved that $\Delta_{\rm II}$ only depends on the Péclet number and the number of simulation steps.

However, the error calculated by $\Delta_{\rm II}$ exhibits rather complicated nonmonotonic behavior as a function of the Péclet number and the number of simulation steps. In the limit $\xi \to 0$, the first step error of $\Delta_{\rm II}$ approaches a constant. Therefore, the behavior of the first step error of $\Delta_{\rm III}$ for small lattice steps is characterized by

$$\Delta_{\rm III}(t = T+) \sim \ell^{-\frac{1}{2}}, \quad \ell \to 0. \tag{36}$$

Note that the first step error is not necessarily the maximum error during simulation. One can further show that the maximum of $\Delta_{\rm II}$ also approaches a constant in the limit $\ell \to +\infty$ or $\xi \to +\infty$. For $\xi \to +\infty$, it follows from Eqs. (14) and (15) that $p_0 \approx 1/\xi$, $p'_+ \approx 1 - 1/\xi$, and p'_- is negligible (exponentially small). Then only the $m = N - i$ term survives in Eq. (35) and

$$P(i,NT) \approx \frac{N!(1/\xi)^{N-i}(1-1/\xi)^i}{(N-i)!i!}. \tag{37}$$

Suppose $N = \alpha\xi$, where $\alpha \sim 1$, and $i = N - j$, where j is a small integer. According to Eq. (13), this corresponds to the time

$$t = NT \approx \alpha\xi\frac{\ell}{q}\left(1 - \frac{1}{\xi}\right) = \alpha\left(\frac{\ell^2}{2D} - \frac{\ell}{q}\right). \tag{38}$$

Then

$$P(i,NT) = P(N-j,\alpha\xi T) \approx \frac{(\alpha\xi)!(1-1/\xi)^{\alpha\xi-j}}{j!(\alpha\xi-j)!\xi^j}$$
$$\approx \frac{(\alpha\xi)^j(1-1/\xi)^{\alpha\xi}}{j!\xi^j} \approx \frac{(\alpha\xi)^j e^{-\alpha}}{j!\xi^j} = \frac{\alpha^j e^{-\alpha}}{j!}, \quad (39)$$

i.e., the Poisson distribution. Note that ξ does not enter the final expression. According to Eq. (2), the continuous distribution $\rho(x,t)$ at this time is Gaussian with the mean $qt \approx \ell\alpha[\ell q/(2D) - 1] = \ell\alpha(\xi - 1)$ and the width $\sqrt{2Dt} \approx \sqrt{\alpha(\ell^2 - 2D\ell/q)} \approx \ell\sqrt{\alpha}$. Thus,

$$\rho(x,\alpha\xi T) \approx \frac{1}{\ell\sqrt{2\pi\alpha}} \exp\left(-\frac{(x-\ell\alpha(\xi-1))^2}{2\alpha\ell^2}\right). \quad (40)$$

Note that

$$\pi_j(\alpha) \equiv \int_{(N-j-\frac{1}{2})\ell}^{(N-j+\frac{1}{2})\ell} \rho(x,\alpha\xi T)dx \quad (41)$$

only depends on j and α, but not on ℓ, ξ, D, or q. Therefore,

$$\Delta_{\rm II}(\alpha\xi T) = \left\{\sum_{j=-\infty}^{+\infty} [\pi_j(\alpha) - P(N-j,\alpha\xi T)]^2\right\}^{\frac{1}{2}} \quad (42)$$

is a function of $\alpha = t/(\xi T)$ only. Consequently, the maximum of $\Delta_{\rm II}$ is a universal constant that does not depend on ξ. This also proves that $\Delta_{\rm II}$ as a function of $\alpha = t/(\xi T) \approx 2Dt/\ell^2$ becomes universal when $\xi \to +\infty$. It follows that the behavior of the maximum of $\Delta_{\rm III}$ for large lattice steps is characterized by

$$\max\{\Delta_{\rm III}(t)\} \sim \ell^{-\frac{1}{2}}, \quad \ell \to +\infty. \quad (43)$$

Since it is not easy to get more information by further mathematical analysis, we will show more details about the quantities by using numerical methods in the next section. According to the definitions, both $\Delta_{\rm II}$ and $\Delta_{\rm III}$ can be numerically calculated before or during simulation.

IV. NUMERICAL RESULTS

According to Eqs. (13)–(15), we implement the fixed time step LMC simulation and set up the one-dimensional biased diffusion problem with constant diffusivity and drift for the purpose of testing.

The total simulation time is set to 100 s and the total number of initial particles is set to 10^6. Since the error introduced in the skewness is only related with the Péclet number, we use a fixed diffusivity $D = 1\ \mu\text{m}^2/\text{s}$ and a set of other parameters, i.e., $\ell = \{1,2,4\}\ \mu$m, $q = \{0.8,1.6,3.2\}\ \mu$m/s, for a comparative study. For each configuration, we run the simulation 100 times and analyze the average result.

We use the Mersenne Twister random number generator to generate random numbers subject to uniform distribution during simulation [10]. We report that both the mean and the mean-square displacement are rather accurate as one can expect. Thus we only plot the skewness in Fig. 1. Note that we use the log-log scale to plot the curve of $-\eta_t$ versus simulation time t for clarity. It turns out that the results agree with our theoretical predictions very well.

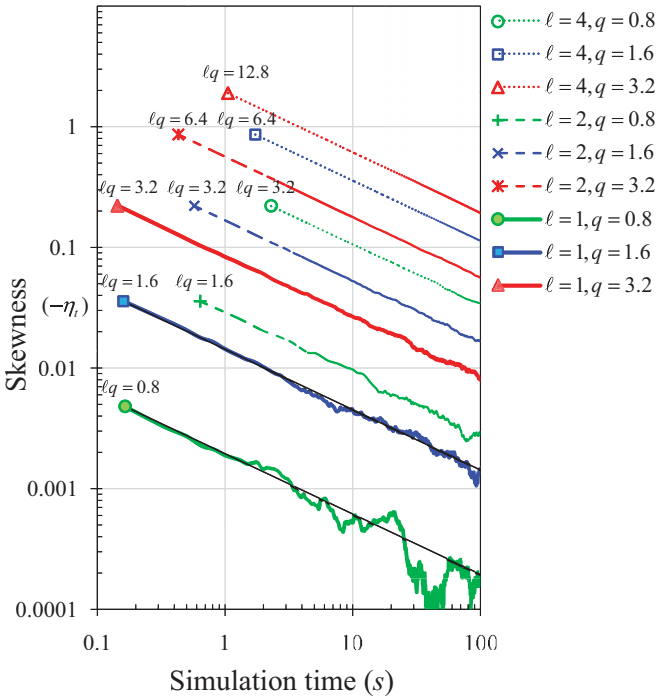

FIG. 1. (Color online) Errors measured by skewness $(-\eta_t)$ as a function of simulation time (s). Values of ℓq (proportional to the Péclet number) are shown on each curve; individual values of ℓ and q are also given.

Clearly, it shows that the fixed time step LMC simulation yields a negative skewness although it can reproduce the mean and the variance exactly. It is also obvious that the first step error only depends on the Péclet number and the error decreases significantly as the Péclet number becomes small. In this sense, the choice of lattice step can be loosed for weakly biased diffusion without significant degradation of accuracy, and vice versa.

It is confirmed that the skewness is inversely proportional to the square root of the number of simulation steps. Note the time step is also related with the lattice step and the drift coefficient according to Eq. (13).

It should be mentioned that the skewness measured during the simulation may vary slightly, especially for small Péclet number and large simulation time. In effect, we have two variants of the LMC simulation: the particle-based approach and the numerically exact master equation approach. The latter can be used instead of the former to avoid the use of random numbers by solving the master equation

$$n_i(t+T) = p_0 n_i(t) + p'_+ n_{i-1}(t) + p'_- n_{i+1}(t), \quad (44)$$

where n_i is the mean particle number at site i.

As a result, it will remove the "noise" from Fig. 1. Note that only two solid straight lines are plotted in the figure to give an impression of the result obtained by solving the above master equation. Clearly, it matches the theoretical result exactly.

Now let us consider the error $\Delta_{\rm II}$ defined by Eq. (32). According to Eqs. (34) and (35), we plot the error $\Delta_{\rm II}$ as a function of the number of simulation steps with different Péclet numbers in Fig. 2.

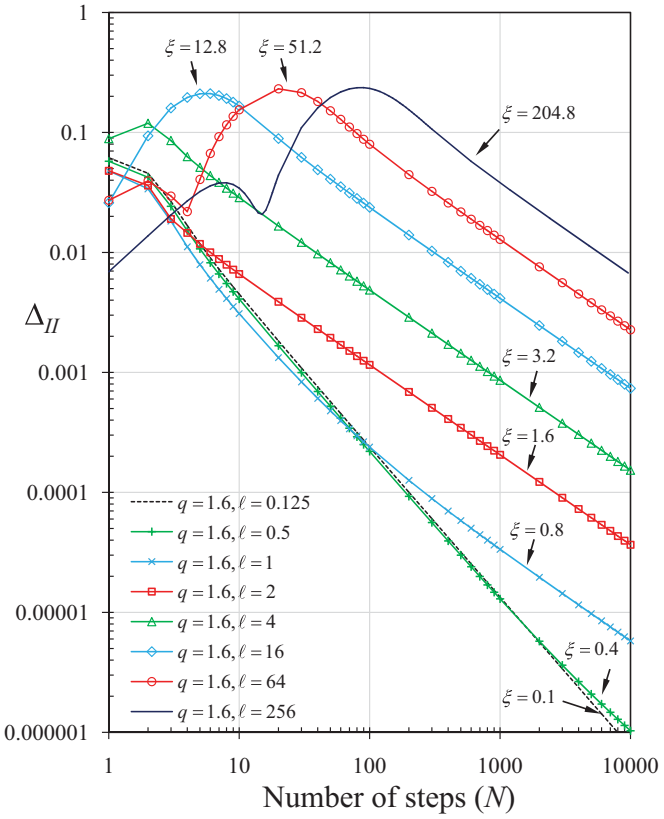

FIG. 2. (Color online) Errors calculated by Δ_{II} versus number of simulation steps with different Péclet numbers. Note that the log-log scale is used in the plot.

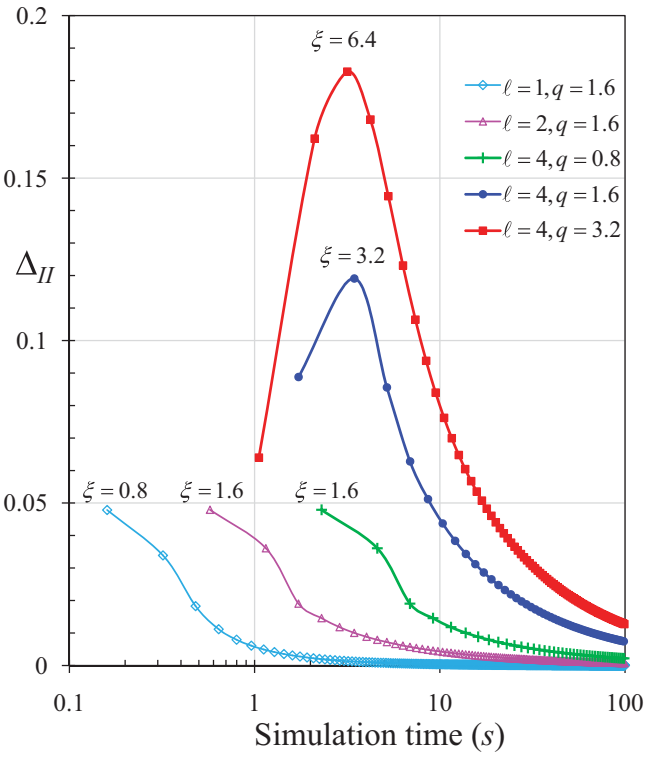

FIG. 3. (Color online) Errors calculated by Δ_{II} versus simulation time. Note fitting curves connecting the data points are used in the plot.

According to the numerical results, one can roughly infer that Δ_{II} decreases monotonically when the Péclet number is small or the number of simulation steps is large. The behavior of Δ_{II} as a nonmonotonic function of the number of simulation steps is rather complicated, especially when the Péclet number is large. It is seen that at $\xi = 3.2$ or $\xi = 12.8$, Δ_{II} first goes up before going down monotonically. It is also obvious that at even higher $\xi = 51.2$ or $\xi = 204.8$ an additional maximum appears.

It shows that the dependence of the error Δ_{II} on the number of simulation steps is power law in the limit $N \to +\infty$. Numerical calculations suggest that there are two power laws: $-5/4$ for intermediate times and $-3/4$ for large times. The crossover between them shifts downwards with increasing ξ. For small ξ, the crossover is at very large times, so that only the $-5/4$ law is visible, and for large ξ, the crossover is at very small times, so only the $-3/4$ law is visible.

According to Fig. 2, it is further confirmed that the maximum of Δ_{II} approaches a constant in the limit $\xi \to +\infty$ which is already stated in the previous section. In fact, it turns out that both in the limit $\xi \to 0$ and in the limit $\xi \to +\infty$ the curves merge when plotted against Dt/ℓ^2.

It is worth mentioning that the time step depends on the lattice step ℓ, the diffusivity D, and the drift q instead of the single Péclet number ξ. Therefore, the figure will be different if Δ_{II} is plotted against the simulation time instead of the number of simulation steps.

To reveal more details about Δ_{II}, we also plot Δ_{II} as a function of simulation time for some typical parameters in Fig. 3. According to the definition of Δ_{II}, it is clear that Δ_{II} is not piecewise constant because $P(i,t)$ remains constant between the jumps while $\rho(x,t)$ does not. Strictly speaking, Δ_{II} is smooth between the jumps. Therefore, fitting curves connecting the data points are used in the plot.

According to Fig. 3, it shows that the error calculated by Δ_{II} evolves exactly the same as a function of the number of simulation steps if the Péclet number is the same. However, the trajectories of Δ_{II} plotted against the simulation time may not overlap because of the different time steps.

As far as the accuracy of the fixed time step LMC simulation is concerned, it would be more interesting to investigate into the maximum of Δ_{II}. For comparison, both the maximum error and the first step error calculated by Δ_{II} are plotted against the Péclet number in Fig. 4.

It is observed that both the maximum error and the first step error calculated by Δ_{II} approach the same constant (approximately 0.062) in the limit $\xi \to 0$. This is consistent with the observation that Δ_{II} decreases monotonically when the Péclet number is small. The maximum of Δ_{II} approaches approximately 0.238 in the limit $\xi \to +\infty$. In addition, it is found that the behavior of the first step error calculated by Δ_{II} is power law in the limit $\xi \to +\infty$ and the exponent is -1.

It is worth noting that the value at which the maximum of Δ_{II} reaches its minimum corresponds to a fixed Péclet number which can be found by numerical method that gives $\xi \approx 1.21$. Consequently, the optimal lattice step can be determined for the fixed time step LMC simulation if it is required to minimize the

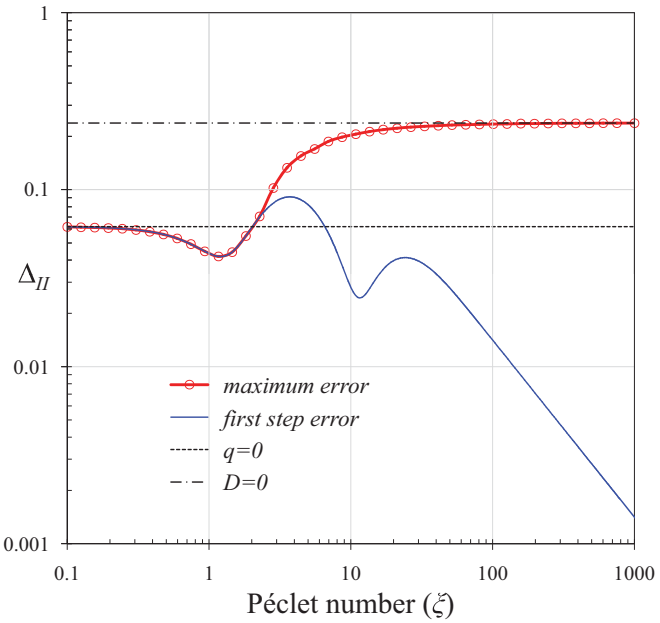

FIG. 4. (Color online) Maximum error and first step error calculated by Δ_{II} versus Péclet number. Note fitting curves connecting the data points are used in the plot.

maximum of Δ_{II}. One can also find that the two curves separate at $\xi \approx 2.25$. This means that Δ_{II} only decreases monotonically when $\xi < 2.25$. Otherwise, Δ_{II} is nonmonotonic.

Then let us consider the maximum error and the first step error calculated by Δ_{III} which are plotted in Fig. 5. Note that Δ_{III} is not only determined by the Péclet number. Therefore, we plot Δ_{III} against the lattice step for some typical parameters.

Clearly, the maximum of Δ_{III} is no longer a constant in the limit $\ell \to 0$ or $\ell \to +\infty$. Instead, it shows that the dependence

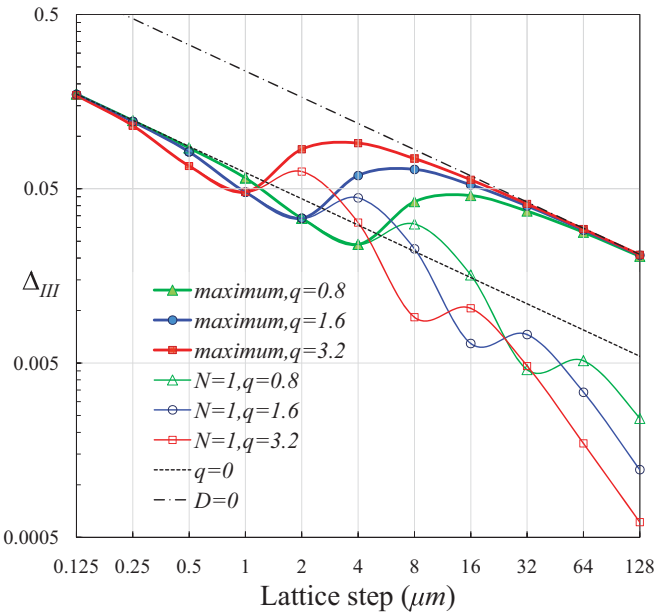

FIG. 5. (Color online) Maximum error and first step error calculated by Δ_{III} versus lattice step. Note the curves connecting the data points are fitting curves.

of the maximum of Δ_{III} on the lattice step is the inverse power law relation in both limits. The reason for this is clear since the maximum of Δ_{II} approaches a constant in the limit $\ell \to 0$ or $\ell \to +\infty$. Consequently, the behavior of the maximum of Δ_{III} is dominated by the coefficient $\ell^{-1/2}$ in both limits. In the limit $\ell \to 0$, the first step error of Δ_{III} is dominated by the coefficient $\ell^{-1/2}$. In the limit $\ell \to +\infty$, the behavior of the first step error of Δ_{III} is also power law but the exponent is $-3/2$.

According to the maximum of Δ_{III} as well as the plot in Fig. 5, one would most likely conclude that the lattice step should be chosen as large as possible to minimize the error introduced in the fixed time step LMC simulation. This is obviously counterintuitive at first, as one would expect the algorithm to become less accurate as the lattice is coarser. However, this is mainly due to the fact that Δ_{III} calculated by Eq. (33) is dominated by the coefficient $\ell^{-1/2}$ in the limit $\ell \to 0$ or $\ell \to +\infty$ as we have stated above.

Suppose there are two normalized distributions and we want to estimate how much they differ from each other using Δ_{II} or Δ_{III}. If the squaring of each term is replaced by the absolute value and the square root is removed, then Δ_{II} is exactly the nonoverlapping area covered by the two distributions while Δ_{III} equals Δ_{II} divided by ℓ. Clearly, Δ_{II} does not depend on ℓ while Δ_{III} does. Most importantly, Δ_{II} has a finite upper bound. In this case, it is not strange that Δ_{III} will be infinitely small in the limit $\ell \to +\infty$. The situation is similar when we come back to the original definitions. Loosely speaking, a larger lattice step will in general produce an even wider and flatter distribution. For simplicity, let us suppose ℓ is *large enough*. If ℓ is increased, but q and D are kept the same, then T will increase according to Eq. (13). As a result, the width of the Gaussian distribution $\sqrt{2NDT}$ will increase. Consequently, the difference measured by Δ_{III} will be smaller.

According to Fig. 3, one can also infer the behavior of Δ_{III} as a function of the simulation time by taking the coefficient $\ell^{-1/2}$ into account. The only difference is that the relative vertical relation between the trajectories with different lattice steps changes. Thus the plot is not presented here.

To sum up, the accuracy of the fixed time step LMC simulation can be measured by different quantities. Roughly speaking, the skewness only characterizes the crude overall shape of the distribution and it may be zero even when the simulated distribution is not identical to the correct distribution. This is the case, in particular, for $q = 0$ and $\eta_N = 0$, but in fact the distribution obtained in LMC simulation is obviously not identical to the Gaussian distribution, and indeed, Δ_{II} and Δ_{III} are nonzero. Therefore, looking at the skewness alone may be misleading, since other characteristics of the error do not necessarily vanish in the limit $\xi \to 0$.

On the other hand, Δ_{II} and Δ_{III} look at the errors in every bin and sum them up. For a large ξ and a small N, Δ_{II} and especially Δ_{III} may be quite small, since most particles are within one bin in both the simulated distribution and the theoretical one, but the overall shape of the distribution is very different from Gaussian and the distribution is very asymmetric, which is reflected in a large η_N. Therefore, it is the meaning of the quantities that makes the difference.

Despite the fact that different quantities may exhibit rather complicated nonmonotonic behavior and lead to various optimal lattice steps, it is believed that these quantities will be appropriate for different purposes or make sense in different contexts.

V. CONCLUSION

The accuracy of the fixed time step LMC simulation is always a matter of concern, especially for biased diffusion models. However, it is never a trivial matter to measure the accuracy of such LMC simulations quantitatively and convincingly. We propose to use the skewness and other quantities to measure the accuracy. We have obtained an explicit expression for the skewness of the one-dimensional fixed time step LMC simulation. It reveals that the accuracy as measured by the skewness only depends on the Péclet number and converges to zero in the long-time limit. We also confirm our theoretical analysis by numerical results.

Moreover, we have discussed several other quantities defined by the root-mean-square error that can also be used to measure the accuracy of the fixed time step LMC simulation. We have revealed some complicated nonmonotonic behavior of these quantities by using numerical calculations. It is found that different quantities may lead to different optimal lattice steps although sometimes it seems to be counterintuitive. Therefore, other possible quantities that can be used to measure the accuracy of such stochastic simulations should be further investigated.

As the importance of introducing a wait time and a probability to stay put is already well known in the LMC simulation [11], we would like to emphasize the importance of choosing an optimal lattice step and better quantities to measure the accuracy of such stochastic LMC simulations whenever it is possible. More generally, the accuracy of such LMC simulations, either uncorrelated or correlated, should be extensively studied because many real applications that involve reactions or interactions between particles depend greatly on the accuracy of simulation [12].

The accuracy not only determines the simulation credibility but also affects the user's confidence. While this work by itself only reveals a little about this topic, we hope it would help to give more insight into the accuracy of such stochastic simulation.

ACKNOWLEDGMENTS

This research was supported by the Natural Science Foundation of China (Grant No. 91024030) and Application of High Performance Computing program from National University of Defense Technology. We gratefully acknowledge the anonymous referees for their careful review and valuable comments, especially for pointing out the errors in the manuscript.

[1] R. M. Mazo, *Brownian Motion: Fluctuations, Dynamics and Applications* (Clarendon, Oxford, 2002).
[2] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon, Oxford, 1999).
[3] M. V. Chubynsky and G. W. Slater, Phys. Rev. E **85**, 016709 (2012).
[4] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, 2001).
[5] M. G. Gauthier and G. W. Slater, Phys. Rev. E **70**, 015103(R) (2004).
[6] A. Donev, V. V. Bulatov, T. Oppelstrup, G. H. Gilmer, B. Sadigh, and M. H. Kalos, J. Comput. Phys. **229**, 3214 (2010).
[7] J. S. van Zon and P. R. ten Wolde, J. Chem. Phys. **123**, 234910 (2005).
[8] G. W. Slater, Electrophoresis **14**, 1 (1993).
[9] M. Vigelius and B. Meyer, PLoS One **7**, e33384 (2012).
[10] M. Matsumoto and T. Nishimura, ACM Trans. Model. Comput. Simul. **8**, 1 (1998).
[11] H. W. de Haan, M. G. Gauthier, M. V. Chubynsky, and G. W. Slater, Comput. Phys. Commun. **182**, 29 (2011).
[12] S. Lampoudi, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **130**, 094104 (2009).