Physical time scale in kinetic Monte Carlo simulations of continuous-time Markov chains

Santiago A. Serebrinsky*
Center for Industrial Research, TenarisSiderca, Dr. Jorge A. Simini 250, Campana B2804MHA, Buenos Aires, Argentina
(Received 7 December 2010; published 31 March 2011)

We rigorously establish a physical time scale for a general class of kinetic Monte Carlo algorithms for the simulation of continuous-time Markov chains. This class of algorithms encompasses rejection-free (or BKL) and rejection (or "standard") algorithms. For rejection algorithms, it was formerly considered that the availability of a physical time scale (instead of Monte Carlo steps) was empirical, at best. Use of Monte Carlo steps as a time unit now becomes completely unnecessary.

DOI: 10.1103/PhysRevE.83.037701 PACS number(s): 02.70.Tt, 02.50.Ga, 66.30.−h, 68.43.Mn

Markov chains are used to model physical systems in an extremely broad range of scientific disciplines, including cellular biology [1], epidemiology [2], condensed matter [3], and finance [4]. Markov chains, associated with the Monte Carlo method, were originally used for the calculation of thermodynamic properties [5]. Later on, their use was extended to the simulation of dynamics [3]. The essential components of a Markov chain are a state space E, of the available states of the system [6], and a set of transition rate probabilities W, assembled into a master equation, which describes the time evolution of the probability p_i of each state i of E. Tracing the evolution $\{p_i(t)\}$, for a given initial condition, enables a full description of the physical process under consideration, and any observable property of interest can be computed. Nevertheless, when the size of the state space is very large, this becomes an unfeasible task. Then kinetic Monte Carlo (KMC; also called dynamic Monte Carlo) is a method of choice for determination of the evolution of the system [7].

The way to perform KMC simulations is in digital computers, where transitions among different states are performed or attempted at discrete time steps, the Monte Carlo steps (MCSs), regardless of whether the Markov chain is evolving in continuous time or discrete time. The ability of KMC methods to correctly describe Markov chains has often been criticized on several grounds.

In this work we analyze a central aspect behind the application of KMC methods to the simulation of continuous-time Markov chains: the absolute time scale. The importance of the time scale lies in the fact that KMC methods replace the continuous-time Markov chain with a discrete-time Markov chain, and a well-founded linkage between the two corresponding time scales is essential for the correct interpretation of KMC results. For the sake of concreteness, whenever an example is called for we refer to a lattice of spins (Ising system), but this implies no limitation in our derivations at all.

The first objection states that the master equation may not represent the actual dynamics of the physical system of interest [8]. As we do not deal with this problem, our point of departure is a given continuous-time master equation, and we assume that it is faithful to the physical system to be described.

The second objection was raised, arguing that the dynamical behavior of a discrete-time Markov chain can be qualitatively different from the dynamics of the underlying continuous-time Markov chain [9], for a system of spins. Ceccatto [10] clarified this point, showing that the spurious complex behavior [9] was due to a sequential spin sweep updating algorithm. The main point considered here relates to the third criticism, described in the following.

KMC algorithms can be broadly classified into two categories: those with rejection (or sometimes "standard") and those that are rejection-free (or BKL). In rejection methods, in every step a transition among states is selected with some probability and it is accepted or rejected according to a second probability. In rejection-free methods, in every step a transition among states is selected with some probability and it is always executed. Rejection-free methods are computationally more efficient when rejection rates in the first class become very high [11]. It has often been considered that correspondence between the physical (for a continuous-time Markov chain) and the algorithmic (for a discrete-time Markov chain) time scales could be established only for the second class of algorithms. For rejection algorithms, the linkage between time scales is controversial. For instance, Bellon and Martin [12] suggested that the connection between KMC and absolute times in rejection algorithms is empirical. Many other authors have also pointed out, in some way, the lack of a formal connection between time scales [13].

We prove here that rejection-free algorithms are simply a limiting case of a generic class of KMC algorithms, which include the rejection algorithms as well. Simultaneously with the proof, we naturally derive a physical time scale for this generic class of algorithms, which is thus valid for both rejection and rejection-free KMC. Therefore, rejection algorithms give a correct description of the dynamics of physical systems, whenever due care is taken for the selection of exit rates (see below). It is worth noting that most of the articles using rejection algorithms actually use the correct time scale, and we provide here a solid ground for it.

The starting point is a (continuous) transition function $P_{ij}(t)$ that gives the probability that the random variable X describing the state of a continuous-time Markov chain at time t is j, given that the state at time $t = 0$ is i, $\Pr\{X(t) = j | X(0) = i\}$. P satisfies the forward and backward Kolmogorov equations [14]. The q matrix of P is defined as $Q = (q_{ij}) = P'(0)$, and it satisfies

$$q_{ij} \geqslant 0 \quad \text{for} \quad i \neq j, \tag{1}$$

$$-q_{ii} \geqslant 0. \tag{2}$$

*sserebrinsky@gmail.com

Conservation of probability implies

$$\sum_{j\neq i} q_{ij} = r_i \quad \text{for all} \quad i\ (r_i = -q_{ii}). \tag{3}$$

A square matrix Q is called a q matrix if it satisfies (1) and (3) [thus, (2) is implied]. For a given q matrix Q, there exists a unique transition function $P(t)$ such that Q is the q matrix of P. Therefore, specifying Q is equivalent to specifying P.

The (continuous-time) master equation associated with P and Q is

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \sum_{j\neq i}[q_{ji}\ p_j(t) - q_{ij}\ p_i(t)]. \tag{4}$$

Two essential properties of a continuous-time Markov chain are as follows.

(1) The *holding time* T_i is the random variable that represents the time that the process remains in state i after an observation of the process in state i. It has an exponential distribution with exit rate r_i and mean residence time $\tau_i = 1/r_i$, that is, $P_{T_i}(t) = \Pr\{T_i > t | X(0) = i\} = \mathrm{e}^{-r_i t}$.

(2) When the process exits state i, it enters another state with a probability distribution given by $\Pr\{X(T_i) = j | X(0) = i\} = q_{ij}/r_i\ (j \neq i)$.

The Markov chain S *embedded* in P is given by

$$S_{ij} = \begin{cases} \delta_{ij} & \text{if} \quad r_i = 0 \quad \text{(absorbing state)}, \\ 0 & \text{if} \quad r_i > 0 \quad \text{and} \quad j = i, \\ q_{ij}/r_i & \text{if} \quad r_i > 0 \quad \text{and} \quad j \neq i. \end{cases} \tag{5}$$

S gives the probability distribution when a jump is effected, but it does not give the holding time. We can say that S forgets the rates of the process $\boldsymbol{r} = (r_i)$ when normalizing q_{ij} with r_i, that is, when transforming the matrix Q into the Markov chain S, which has one less independent variable per row than Q (giving precisely the rates r_i). If there are no absorbing states, as in a lattice of spins, the case $r_i = 0$ does not appear, and

$$S_{ij} = \begin{cases} 0 & \text{if} \quad j = i, \\ q_{ij}/r_i & \text{if} \quad j \neq i. \end{cases} \tag{6}$$

The two properties above were given as an alternative definition of a continuous-time Markov chain [15]; that is, if a process satisfies them, then it is a continuous-time Markov chain, and its Q (and therefore P) is uniquely determined by $(S,\boldsymbol{r})$. Therefore, the following descriptions correspond to the same continuous-time Markov chain—(i) P, (ii) Q, (iii) $(S,\boldsymbol{r})$—and there is a one-to-one correspondence among the three entities. These properties are the essence of rejection-free KMC: if the system is in state i at a given time step n, (i) the next state $j \neq i$ is generated with probability S_{ij}, and (ii) the time is increased by a random value Δt_n with an exponential distribution with parameter r_i. Selection of state j requires knowledge of all S_{ij}, as $\sum_{j\neq i} S_{ij} = 1$ (for nonabsorbing states).

We next show that there exist pairs $(V,\boldsymbol{w})$ of Markov chain V and sets of exit rates $\boldsymbol{w} = (w_i)$, alternative to $(S,\boldsymbol{r})$, that describe the same continuous-time Markov chain as P. These $(V,\boldsymbol{w})$ form a continuous set, which, in a limit, tends to $(S,\boldsymbol{r})$.

We take fictitious exit rates $w_i \geqslant r_i$, rejection probabilities $\gamma_i = 1 - r_i/w_i$, and the Markov chain

$$V_{ij}[\boldsymbol{\gamma}(\boldsymbol{w})] = \begin{cases} \gamma_i & \text{if} \quad j = i, \\ (1-\gamma_i)S_{ij} & \text{if} \quad j \neq i, \end{cases} \tag{7}$$

where $\boldsymbol{\gamma} = (\gamma_i)$. Clearly,

$$(S,\boldsymbol{r}) = \lim_{w_i \to r_i^+} (V_{ij}[\boldsymbol{\gamma}(\boldsymbol{w})], \boldsymbol{w}).$$

This process, when observed in state i, remains in state i without transitioning for a time, with exponential distribution, exit rate w_i, and mean residence time $s_i = 1/w_i$. When it jumps from state i, it enters state j with probability V_{ij}. The difference between S and V is that V allows for jumps from state i onto itself; that is, $V_{ii} \neq 0$ in general. Rejection methods are based on this description: (i) the next state $j \neq i$ is generated and accepted with probability $V_{ij} \leqslant S_{ij}$ [see Eq. (7)], and (ii) the time is increased by a random value Δt_n with exponential distribution with parameter w_i. Selection of state j does not require knowledge of all V_{ij}, as $\sum_{j\neq i} V_{ij} \leqslant 1$.

To prove that $(V,\boldsymbol{w}) \equiv (S,\boldsymbol{r})$, it is enough to prove that (1) the *net exit rate* from state i (i.e., the rate at which the process actually jumps from state i onto a state $j \neq i$) accords with property 1 above, and (2) the probability distribution among states $j \neq i$ upon exiting state i accords with property 2 above. Part (2) is trivial, since $V_{ij}/\sum_{j\neq i} V_{ij} = V_{ij}/(1-\gamma_i) = S_{ij}$.

We denote the event "the jump number m is from state i to state j" with (i,j,m). Then, the holding time has probability distribution

$$\begin{aligned} P_{T_i}(t) = \Pr\{T_i > t | X(0) = i\} &= \Pr\{N(t) = 0\} + \Pr\{N(t) \\ &= 1,(i,i,1)\} + \Pr\{N(t) = 2,(i,i,1),(i,i,2)\} + \dots, \end{aligned} \tag{8}$$

where $N(t)$ is the number of jumps by time t. Due to the independence of successive jumps

$$\Pr\{N(t) = n,(i,i,1),\dots\} = \Pr\{M(t) = n\}\ {V_{ii}}^n,$$

where $\{M(t)\}$ is a Poisson process with rate w_i. Therefore,

$$\begin{aligned} P_{T_i}(t) &= \sum_{n=0}^{\infty} \Pr\{M(t) = n\}\ {V_{ii}}^n \\ &= \sum_{n=0}^{\infty} \frac{(w_i t)^n}{n!} \mathrm{e}^{-w_i t} {V_{ii}}^n \\ &= \mathrm{e}^{-w_i t} \mathrm{e}^{\gamma_i w_i t} = \mathrm{e}^{-(1-\gamma_i)w_i t} = \mathrm{e}^{-r_i t}, \end{aligned} \tag{9}$$

which proves part (1).

Our analysis is a more general case of the *uniformization* technique [14], whereby a given continuous-time Markov chain [specified by $(S,\boldsymbol{r})$, P, or Q] is associated with a $(V,\boldsymbol{w})$ with uniform $w_i = w$. Under this condition, the transition function of $(V,\boldsymbol{w})$ is

$$\tilde{P}_{ij}(t) = \sum_{n=0}^{\infty} (V^n)_{ij}\, \mathrm{e}^{-wt} \frac{(wt)^n}{n!}. \tag{10}$$

It is easily shown that the q matrix of $\tilde{P}$ is $\tilde{Q} = \tilde{P}'(0) = Q$, that is,

$$\tilde{P'}_{ij}(0) = \begin{cases} -r_i & \text{if} \quad j = i, \\ r_i S_{ij} = q_{ij} & \text{if} \quad j \neq i, \end{cases} \tag{11}$$

and therefore $\tilde{P} = P$.

As an example, consider an Ising system with applied magnetic field, undergoing single spin flips with Metropolis transition rate probabilities

$$q_{ij} = \nu\, \mathrm{e}^{-\beta \Delta E} \min(\mathrm{e}^{-\beta(H_j - H_i)}, 1), \tag{12}$$

where ν is an attempt frequency, $\beta = 1/kT$ is the reciprocal temperature, ΔE is an activation energy, and H_i is the energy in state i. The algorithm loops through three steps for evolving from state i at time step n: (i) select a state j among the b_i candidate states communicating with i, with probability $1/b_i$; (ii) accept the transition to j with probability $\min(\mathrm{e}^{-\beta(H_j - H_i)}, 1)$ [5]; and (iii) add $\Delta t = -\ln \chi / w_i$, where χ is a uniformly distributed random variable in $(0,1]$, $w_i = r_i / \sum_{j \neq i} V_{ij} = \nu\, \mathrm{e}^{-\beta \Delta E} b_i$, $r_i = \sum_{j \neq i} q_{ij} = \nu\, \mathrm{e}^{-\beta \Delta E_0^{\mathrm{m}}} \sum_{j \leftrightarrow i} \min(\mathrm{e}^{-\beta(H_j - H_i)}, 1) \leqslant w_i$, and $V_{ij} = \min(\mathrm{e}^{-\beta(H_j - H_i)}, 1)/b_i \leqslant 1/b_i$. In each iteration, only one V_{ij} and w_i are computed.

The comparison of the results of simulations using rejection-free KMC and an appropriately derived rejection KMC for the same problem tests the linkage between time scales. Such a favorable comparison is implicit in developments of rejection-free algorithms [11,15].

In conclusion, we have provided a formal proof of the existence of a physical time scale for rejection KMC algorithms. Any simulation with a rejection KMC algorithm that is related to the intended master equation as indicated here is guaranteed to provide the correct results.

[1] K.-H. Lee, A. R. Dinner, C. Tu, G. Campi, S. Raychaudhuri, R. Varma, T. N. Sims, H. Burack, W. Richard Wu, J. Wang, O. Kanagawa, M. Markiewicz, P. M. Allen, M. L. Dustin, A. K. Chakraborty, and A. S. Shaw, Science **302**, 1218 (2003).

[2] C. J. Rhodes and R. M. Anderson, Nature **381**, 600 (1996).

[3] W. M. Young and E. W. Elcock, Proc. Phys. Soc. London **89**, 735 (1966).

[4] B. Eraker, J. Bus. Econ. Stat. **19**, 177 (2001).

[5] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[6] We defer the consideration of subtleties about the size of the state space, focusing on finite-state spaces.

[7] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, 2nd ed. (Cambridge University Press, Cambridge,2000).

[8] K. Binder and D. W. Heermann, *Monte Carlo Algorithms in Statistical Physics—An Introduction*, 3rd ed., Springer Series in Solid-State Sciences, Vol. 80 (Springer, Berlin, 1997).

[9] M. Y. Choi and B. A. Huberman, Phys. Rev. B **28**, 2547 (1983); **29**, 2796 (1984).

[10] H. A. Ceccatto, Phys. Rev. B **34**, 4889 (1986).

[11] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).

[12] P. Bellon and G. Martin, Mater. Sci. Forum **155-156**, 209 (1994).

[13] M. A. Gallivan, D. G. Goodwin, and R. M. Murray, Phys. Rev. B **70**, 045409 (2004).

[14] S. Ross, *Introduction to Probability Models*, 7th ed. (Academic Press, San Diego, CA, 2000).

[15] A. Sadiq, J. Comput. Phys. **55**, 387 (1984).