

Kinetic Monte Carlo: Bare Bones and a Little Flesh

UIUC Materials Computation Center
Computational Materials Science Summer School 2001
"Spanning Length and Time Scales"
funded by NSF Current Research–Curriculum Development Grant

*by Duane Johnson, Professor, University of Illinois
Materials Science & Engineering and Physics*

Background Article: "Theoretical foundations of dynamical Monte Carlo simulations",
K.A. Fichtorn and W.H. Weinberg, J. Chem. Phys. **95**, 1090 (1991).
(Used with Permission)

I. Introduction to Kinetic (a.k.a. Dynamical) MC (5 min.)

- A. Time Scales: MD, MC, and KMC
- B. A Hybrid Monte Carlo: MD + KMC

II. Solution of the "Markovian" Master Equation (15 min.)

- A. MC vs. KMC
- B. Detailed Balance Requirements

III. Rudiments of KMC (simulation time vs. real time (45 min.))

- A. Dynamical Hierarchies (time scales via transition rates)
- B. Poisson Processes (a necessary aside)
 - examples: radiation counts, binomial distributions
- C. Simple Example: Adsorption-Desorption in 2-D
(see background article)
- D. Generalization

IV. Algorithms for "Fast" Monte Carlo (30 min.)

- A. Basics of the "n-fold-way" (or "residence-time") algorithm
- B. Improvements using Binary or K-search Methods

V. Laboratory KMC Exercise (20 min.)

- A. Vacancy-assisted Ordering in 2-D Binary Alloy
- B. Special Effects: radiation damage via ballistic jumps

VI. Beware of Solving Master Equation via Approximations (if time permits, 10 min.)

- A. Technical Point: Fokker-Planck Equation is wrong.
- B. Technical Point: First-order Fix-up from Kubo
- C. Example: Non-equilibrium transitions in fcc AB_3 compounds,
see Haider, Bellon and Martin, Phys. Rev. B 42 8274 (1990)

Kinetic Monte Carlo Simulations

(a.k.a. Dynamical Monte Carlo)

Copyright by Duane Johnson (May 2001)

1 Overview: MD vs. MC

Molecular dynamics (MD) simulations explicitly propagate the equations of motion through time (e.g., Verlet algorithm) and, thus, they follow the real dynamics of a many-body system. However, time steps must be chosen small enough to address correctly the physics at hand and large enough to span a large time interval (which is often just *pico-seconds*, far from that to compare to experiments). Such restrictions on computational time limit the scope of dynamical simulations involving interacting particles, molecules, etc. In addition, as system complexity increases, there can be substantial decrease of system sizes and of real time evolution. System complexity can arise from many sources, e.g. gas-surface or condensed phases, which involve direction dependent, many-body interactions. In particular, MD is unable to perform simulations over long times due to the required repeated evaluations of potential and derivatives of potentials (potential-energy surface information). This may be partly overcome by so-called *Hyperdynamics* methods; nonetheless, it still constitutes a significant limitation.

Note on Time Scales in MD: High-frequency motions will determine the integration time step that must be used. For small time steps, low-frequency events cannot be sampled in reasonable computational time. Thus, MD has difficulty accommodating events that occurs with vastly different probabilities or rates.

Direct Monte Carlo (MC) simulation are a quasi-random, i.e. non-deterministic, approach to time evolution. In other words, MC methods evolve according to a *stochastic* algorithm, so as to explore the entire phase space (possible accessible states) of the system. Such stochastic attempts at changing the configurations usually have nothing to do with the underlying energy barriers inhibiting (or favoring) change of state (or motion). The MC methods follow a Markov process to evolve a system towards equilibrium, regardless of pathway. For example, the standard Metropolis MC algorithm makes all transitions of the system to states of lower or equal energy with probability of unity, regardless of the barrier to access this state, the process required or the transition rate. Hence, MC moves based on the Metropolis algorithm cannot be interpreted dynamically as a process that simulates random motion in time! Nonetheless, MC methods do generate states that are correlated – each configuration is correlated because it is produced in a sequential Markov chain process. At present, all practical sampling methods produce correlated configurations (which is also why you use auto-correlation times to remove some clearly wrong correlations during analysis of MC data).

As for the case of diffusion, notice that there is a need for a relation to a real-time

process, which is not necessarily the same as a simulation time, depending on the MC Algorithm used (see later). The real time evolution is required to characterize the system, such as the diffusion constant. Most MC papers report everything in Monte Carlo Step (MCS) being one unit of simulation time. A MCS unit has different values depending on the problem, and it is not necessarily related directly to (real) experimental time.

Note on Time Scales in MC (The BIG Point): If, and only if, the correlations between these quasi-random states in the MC can be interpreted as dynamical correlations in a stochastic sense, can a kinetic interpretation of MC be made; that is, states represent actual motion of the system – not just sampling of the ensemble. If this is the case, then the expectation value can be calculated as time averages! And, importantly, a connection can be made between simulation time, t_{sim} , and real time, t_r , rather than just quoting the simulation time in terms of MC Steps (MCS), as often done. (More on this later.)

Kinetic Monte Carlo (KMC, sometimes called Dynamic MC, and not to be confused with Diffusion MC), in contrast, is meant to address (roughly speaking) materials-related problems whose physical outcomes are very much governed by the energy barriers between various possible states, or local configurations of the system. The rates for these physics-based transitions control the dynamics. KMC is meant to have a direct relationship to real time, rather than MCS. It also follows that KMC does not require potential derivatives, giving it a distinct advantage over MD methods.

Note on Time Scales in KMC: Events with largest rates will dominate the simulation, while low-rate (small probability of occurrence) will be rare. Sometimes such differences in rates can be overcome by ignoring “fast events”; in which case the KMC simulation will have events of similar order.

Obvious examples where KMC is needed: diffusion of an adatom species at the surface of a crystal or a concentration profile inside a bulk material. Each depends on the diffusion barriers associated with the local bonding (chemical environment) and position of other adatoms, and interaction with the substrate or bulk atoms. In these situations, the transition probabilities must occur so that the transition rates are established in terms of models for the rates of microscopic events comprising the overall process. As discussed by Fichthorn and Weinberg [8], when the transition probabilities are formulated as physical-based rates within a Poisson Process, a definite (or unambiguous) relationship can be established between a MC time-step (MCS) and transition rates, if there is a “dynamical hierarchy” established for the possible transitions. Simply stated, a “dynamical hierarchy” requires that there is a spectrum of transition probabilities (when made relative to the maximum rate process) such that the maximum probability is one. Fichthorn and Weinberg [8] concluded that **iff**

1. a so-called “dynamical hierarchy” is established for the transition probabilities which obeys detailed balance criterion;
2. time increments for each successful event are calculated properly (as we will discuss); and
3. independence of each possible events in system can be achieved,

then MC methods simulate a **Poisson Process** and both static and dynamic properties of Hamiltonian may be obtained and interpreted consistently. Not all choices of transition probabilities, however, can satisfy these criteria, as exemplified by Kang and Weingberg [12] and Ray and Baetzold [17](and discussed briefly later).

What is dynamical hierarchy? For example, in a simplistic adsorption-desorption of a “atom” at a surface, which we return to again, a transition rates for adsorption at any site i , W_{Ai} , and for desorption, W_{Di} , of one type of molecule at a surface will be proportional to the overall rates for these events, r_A and r_D and the site-dependent coverage, if for nothing else than intuitively. The probabilities for these events would be $P_A = r_A/R$, where $R=r_A + r_D$, and similarly for desorption. An order for these probabilities, or dynamical heirarchy, can be created most easily by defining $W_i = r_i/r_{max}$. This means, if $r_A > r_D$, that $W_A=1$ and $W_D = r_D/r_A$. Thus, $W_A > W_D$, a heirarchy exist [8]. This generalizes for more processes. You can be begin to intuit that time will be reflected in these rates, for the more probable the less time passes before the event occurs.

Recall that a number of definitions of transition probabilities will give detailed balance. As such, an additional criteria is required so that the transition probabilities reflect unique transition rates, and, hence, simulate the dynamical phenomena within the model. It is the formulation of a dyamical heirarchy for the transition rates that is the additional criteria, see Ref. [8] for more discussion.)

1.1 Synopsis of KMC

Thus, main difference in KMC is to evolve a system using a stochastic (or random-walk) algorithm that directly takes into account the physical energy barriers that govern the evolution of a system and that translate to a real time scale (t_r , rather than just reporting some simulation time (t_{sim})). Importantly, the KMC time for a given step will be scaled by the average time required to observe the particular stochastic event chosen to occur that time step. Thus, each KMC time step will have widely different magnitudes (sometime by orders of magnitude) depending on event, temperatures, barrier changes, and so on. Notably, KMC simulations are usually on a lattice (or ordered mesh) as many problems, such as surface diffusion, can always be mapped to an effective lattice.

Technically, independent events, such as in adhesion within a low-flux atmosphere, can be considered Poisson processes which have certain transition probabilities associated with them. Each of the possible transition probabilities depend upon configuration and environment, and evolve as a function of time. It is the up-date per time step of these transition probabilities that become the time consuming portion of any Kinetic Monte Carlo. Basically, three criteria are needed to simulate a Poisson Process:

1. transitions must satisfy *detailed balance*,
2. there is a proper correspondence between t_{sim} and t_r , and
3. the events comprising the sequence of processes are independent.

It can be shown that any set of Poisson Processes, (or, based on a dynamical interpretation, any Monte Carlo calculations), are governed by a so-called Master Equation which encompasses all possible transitions allowed from a given set of configurations. As such, KMC allows simulation of non-equilibrium processes, not just simulation of equilibrium averages as typical from Monte Carlo.

1.2 KMC Problems and Algorithms

Generally speaking, every simulation method has two key components: the physical problem and the algorithm (and data structure) to solve the model or equations representing that problem. It is important to keep in mind that for any simulation the efficiency of the algorithm may be more important than the physical model. For KMC, this is also true and there have been, and continue to be, research into developing efficient algorithms for updates of the stochastic rates each time step. These methods, e.g., "n-fold-way", binary-tree, and K-level search algorithms, and so on, combined with clever organization of data can speed up simulations by orders of magnitude, allowing significant improvement in statistics. We discuss briefly the important algorithm and data structures for KMC in a later section. Here we just cursorily state a few example areas that KMC have addressed. (Note that the references are listed in alphabetical order, hence, the references are not labelled sequentially.)

The paper of Fichthorn and Weinberg [8] is a very good introduction to the fundamental underpinnings to KMC. For the effects within the simulation regarding physical energy barriers and stochastic-based algorithms, as well as the connection of t_r and t_{sim} , see Kang and Weinberg [12]. Clark et al. [5] have discussed a hybrid method that combines methodologies (MD and KMC) to address processes with substantially varying rates, in their case Chemical Vapor Deposition (CVD).

Probably, although I have performed no exhaustive search, one of the first materials simulations using KMC for studies of surface diffusivity is that by Reed and Ehrlich at the University of Illinois [18]. This paper is probably not as well known because Reed and Ehrlich did not give the method a fancy name, like *Kinetic Monte*

Carlo. Since that time there have been numerous publications in physics, chemistry, biology, surface science, materials science, etc., using the KMC approach. These include topics and physical processes such as:

1. surface diffusion (e.g. [18, 17]),
2. MBE (molecular beam epitaxy) growth (e.g. [15]),
3. CVD (chemical vapor deposition) growth (e.g. [5]),
4. vacancy diffusion in alloys (e.g. [1, 2]),
5. compositional patterning in alloys driven by irradiation (e.g. [6]),
6. polymers: topological constraints vs. entanglements (e.g. [16]),
7. coarsening of domain evolution (e.g.[11]), and
8. dislocation motion (e.g. [14]).

(N.B. *Although I am familiar with some of these these items, they are not listed as endorsement or importance of work. They are provided as examples from different areas.*)

Note that, for instance, in thin-film growth there are multi-events occurring with varying time scales, e.g., deposition ~ 1 monolayer/sec or hopping ~ 1 atom/site/sec. Each problem may have multiple processes with different rates, and, you need to know what are the (relative) rates of each. There are difficulties inherent. For reconstruction, one needs to keep track of top atom and its bonding partners, dimer bond energies, dimer-dimer energy (e.g., does it prefer rows, not random?). The basic idea – describe system with a set of equations in terms of rates of reaction; iteratively solve equation in time; and, use data from a few atomistic calculations (or experiment or phenomenology). We will see this in our example problem of adsorption-desorption. Such Kinetic Rate Equation methods provide a simple means to describe a wide variety of systems, but one gives up spatial resolution of lattice MC, yet gains ability to model larger systems.

KMC may be used to study non-equilibrium situations, which may perhaps be best developed within a so-called Master-Equation Approach. For example, for studying an order-disorder transition in an Ising-type alloy on a lattice, KMC via a Master Equation approach has been used by Haider, Bellon, and Martin [10]. In this study, a non-equilibrium phase transition were studied where there were two competing processes: thermally-activated jumps and ballistic jumps, like those that occur under irradiation (which we shall consider for homework). A mean-field approximation was used for the kinetic equations for the evolution of the *order parameter* of interest, in their case of binary alloy ordering it was long-range order. The validity of KMC for non-equilibrium cases is guaranteed by establishing a “dynamical hierarchy” and transition based on Poisson processes [8].

2 The Markovian Master Equation

As stated earlier, MC methods are correlated in that they are generate sequentially in a Markov process. The time evolution of the probability density is given by a Markovian Master Equation. This Master Equation for evolution of dynamical processes is quite intuitive and is given by

$$\frac{\partial P(\sigma_f, t)}{\partial t} = \sum_{\sigma_i} W(\sigma_i \rightarrow \sigma) P(\sigma_i, t) - \sum_{\sigma_i} W(\sigma_f \rightarrow \sigma_i) P(\sigma_f, t) . \quad (1)$$

Here σ_i and σ_f are the successive (e.g., intial and final) states of the system, usually a set of 3-D vector for all M particles. $P(\sigma, t)$ is the probability that the system is in state σ at time t , and $W(\sigma_i \rightarrow \sigma_f)$ is the transition probability per unit time that the system will undergo a transition from state σ_i to σ_f . (Sometime W_{ij} is explicitly given as $\omega\tau^{-1}$, where τ is a characteristic MCS and ω is a probability. Note that in Markov processes typically $\sum_j W_{ij} = 1$ so that the total probability of doing an event (or not) is 100%). Generally, KMC can be regarded as algorithms that numerically solve the Eq. (1) by choosing randomly among various possible transitions and accepting (or rejecting) particular transitions with appropriate probabilities. Upon each successful transition (or, in some cases, each attempted transition), time is incremented in varying values τ (in standard MC these typically are integral values of MCS). The time scale is incorporated via the transition rates (discussed in the next Section). But, importantly, these rates must be determined by some method, such as phenomenological arguments, experiment, or MD simluations, for they are not intrinsic to the KMC method.

It is noteworthy that each possible transition rate for an event depends upon *configuration and environment* and *evolve as a function of time*. (For example, the simple adsorption-desorption model discussed later, the fractional surface coverage evolves in time and affects the rate of adsorption and desorption at each time-step. Hence, the UPDATE per TIME-STEP of the transition rates is the time-consuming portion of KMC. These issues also highlight the importance of data structure, search and update algorithms in the KMC methods, which are discussed later.

Fichthorn and Weinberg [8] have discussed that KMC formulated as a set of (independent) Poisson processes is valid for non-equilibrium Monte Carlo simulations if the transition rates satisfy the “dynamical hierarchy” criterion (or scaled rates, such that maximum probabilities are 1) mentioned earlier and these transition probabilities (based on the rates) obey detailed balance (or microscopic reversibility).

2.1 Detailed Balance

Note that at steady-state (not necessarily equilibrium) the L.H.S. of Eq. (1) is zero and the sum of all transitions into a particular state σ equals the sum of all transitions out of the state. You must be careful to ensure this if using phenomenological arguments (which can be done by requiring detailed balance). Although not actually guaranteed to give solution of the Master Equation at any given time, detailed-balance

criterion, i.e.,

$$W(\sigma_i \rightarrow \sigma_f)P(\sigma_i, t = \infty) = W(\sigma_f \rightarrow \sigma_i)P(\sigma_f, t = \infty) \quad , \quad (2)$$

must be imposed so that the MC transition probabilities are consistent with the Boltzmann distribution,

$$P(\sigma, t = \infty) = \frac{e^{-\beta H(\sigma)}}{Z} \quad , \quad (3)$$

expected at thermal equilibrium (steady state). The detailed-balance criterion is clearly much more strict here than is required. However, as in the Metropolis algorithm, it was used in order to require no knowledge of the transition matrix in the Random Walk, and to guarantee Boltzmann-like behavior. There are some cases of driven systems, such as under irradiation, where the detailed-balance criterion cannot be obeyed while solving the Master Equation, see, e.g., the Refs. [1, 2, 10].

So, generally speaking, the choice of transition probabilities in MC simulations can affect the ratio of time scales [12], and gives a temperature dependence for the time scales. Most notably, under non-equilibrium situations the time scales for different transition probabilities (i.e., different types of MC simulations) results in **non-linear time scales** between these different MC methods. In addition, other important quantities are affected, such as the growth exponents and the diffusion coefficients, especially at low temperatures. For example, using the Kawasaki dynamics within Monte Carlo (where two "atoms" are swapped simultaneously so as to maintain stoichiometry, or number of spins), energy barriers are taken into account BUT there can NOT be a correspondence between MC time and real time for most physical processes - not unless there is a single rate operable in the physics under consideration [12, 17]. In other words, there can sometimes be a dynamical hierarchy for simple cases; but, the transition probabilities for Kawasaki dynamics cannot be generalized for complex potentials to produce a "dynamical hierarchy", so time (diffusion constants, etc.) must be carefully considered due to the non-linear ratio of time scales.

2.2 Possible Choices for $W(\sigma_i \rightarrow \sigma_f)$

Let $\Delta E = H(\sigma_f) - H(\sigma_i)$. We have mentioned 3 choices, which are representative. (Here I follow ideas discussed by Kang and Weinberg [12].)

2.2.1 Metropolis Walk

$$W_M(\sigma_i \rightarrow \sigma_f) = \frac{\omega_M(\sigma_i \rightarrow \sigma_f)}{\tau} = \begin{cases} e^{-\beta \Delta E} & \text{for } \Delta E > 0 \\ 1 & \text{for } \Delta E \leq 0 \end{cases} \quad , \quad (4)$$

For Metropolis, τ is the usual MCS. Note, again, that the acceptance of all states that lower energy is why Metropolis algorithm does not express the real dynamics

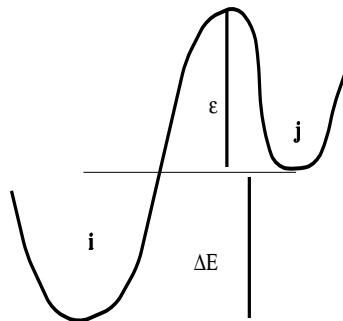


Figure 1: Example for barrier between two sites i and j with energy ΔE between minima and additional ϵ barrier. Note to hop from i to j would cost more energy than j to i , which is an example of microscopic reversibility without equal probability.

of system. Another way to view this is that, unlike MD which diffuses through phase space based on real potential barriers and transitions, Metropolis based MC samples phase randomly – jumping here and there regardless of barriers required to reach a particular lower-energy minimum. In addition, this algorithm for reaching equilibrium clearly does not allow one to establish a *dynamical hierarchy* for the transition rates.

2.2.2 Kawasaki Dynamics

In Kawasaki dynamics two particles, or two spins, etc., are exchanged simultaneously so as to conserve particle (spin) number, which yields a factor of 2.

$$W_K(\sigma_i \rightarrow \sigma_f) = \frac{\omega_K(\sigma_i \rightarrow \sigma_f)}{\tau} = \frac{e^{-\beta\Delta E/2}}{e^{-\beta\Delta E/2} + e^{\beta\Delta E/2}} \quad , \quad (5)$$

For a surface diffusion, for example, the attempt frequency $\tau = e^{\beta Q(T)}/\nu$, where Q is the average energy barrier a particle must jump and it is T dependent because σ_i is. The frequency, ν , of frustrated translation motion (parallel) on surface. Typically, in Kawasaki dynamics, τ is taken as unit of time and the Q dependence is lost.

2.2.3 Thermally-Excited Process (barrier controlled)

$$W_E(\sigma_i \rightarrow \sigma_f) = \frac{\omega_E(\sigma_i \rightarrow \sigma_f)}{\tau} = e^{-\beta E_b(\sigma_i, \sigma_f)} = e^{-\beta(\epsilon + \Delta E)} \quad , \quad (6)$$

where the barrier depends on *initial* and *final* positions. That is, it is not symmetric (e.g., $W_E(\sigma_f \rightarrow \sigma_i) = e^{-\beta\epsilon}$) but does obey microscopic reversibility or detailed balance. Figure 1 shows the example barrier. Here, τ is simply $1/\nu$ and ω_E is a physical quantity: the probability of success for thermally excited barrier crossing for average time τ .

2.2.4 A Comparison of Probabilities

For a single barrier jump, one can show (see Figure 1 for definition of energies) that $\omega_K(i \rightarrow j) = A\omega_E(i \rightarrow j)$, where

$$A = \frac{e^{\beta\Delta E/2 + \beta|\epsilon|}}{e^{\beta\Delta E/2} + e^{-\beta\Delta E/2}} \quad . \quad (7)$$

In an interacting lattice-gas model, the difference in energy for each pair of sites will depend on the local configuration around the pair of sites. So, each type of neighborhood will have a specific value of A. When you have no additional interaction in the lattice-gas, i.e. no dependence on other neighbors and so forth, then there is only one value of A. As a result, one can use Kawasaki dynamics and still obtain real time. For Kawasaki dynamics, the real time is $t_r = A\tau S/N$; whereas, for kinetic version with excitation barriers, real time is $t_r = \tau S/N$, where S is the number of MCS.

The only reason this may be important is if we consider normalizing the Thermal Excited transition probabilities with respect to hopping forward and backwards. We find that

$$\frac{\omega_E(i \rightarrow j)}{\omega_E(i \rightarrow j) + \omega_E(j \rightarrow i)} = \omega_K(i \rightarrow j) \quad . \quad (8)$$

Thus Kawasaki dynamics renormalizes the probabilities of barrier crossings so that $\omega_K(i \rightarrow j) + \omega_K(j \rightarrow i) = 1$. It follows immediately that for barriers where $\omega_E(i \rightarrow j) + \omega_E(j \rightarrow i) < 1$ the hopping rates are increased and, where the sum is > 1 , the rates are decreased. Because of this renormalization, Kawasaki dynamics requires shorter computational times for processes where ω_E are very small. When ω_E are small, a hop has to be attempted many times before it is successful and computer time can be prohibitive. But, again, this depends also on algorithm used in simulation. For thermally excited KMC case using n-fold-way algorithm, this advantage of Kawasaki dynamics is no longer an issue. One may, however, wish to utilize the advantage of Kawasaki dynamics for maintaining stoichiometry because it swaps two things at a time. This may be incorporated into thermal KMC approach too.

3 Poisson Processes and Real Time

3.1 Poisson Statistics

If one adopts a dynamical interpretation of Monte Carlo with a perspective that we will resolve time such that events never occur simultaneously, then our MC algorithm must establish a chronological sequence of distinct events that are separated by various intervening times. Because the microscopic dynamics giving the exact times are not modelled directly, the events and associated times must be constructed from appropriately weighted probability distributions of all possible outcomes. This course-graining of time and events, with its inherent expectations, assumes that each event, i , can be characterized by an average transition rate, r_i . Given this, it is assumed that any particular transition which is possible at t can again be possible at some later time $t + \Delta t$ with a uniform probability based on its rate and is independent of previous events. This is, by definition, a Poisson process and is given by the Poisson distribution, i.e.

$$P(n_i) = \frac{(r_i t)^{n_i}}{n_i!} e^{-r_i t} \quad , \quad (9)$$

where n is a random variable counting the number of events i which have occurred within time t . Notice the expected number of events occurring within a time t is $\langle n_i \rangle = r_i t$, which recovers the average occurrence rate of the process. It can be shown that the Poisson process is consistent with the Master Equation (N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1990).

A useful feature of the Poisson process is that an ensemble of independent Poisson processes will behave as one, large Poisson process. Thus, the statistical properties of the ensemble may be formulated in terms of the dynamics of the individual processes. For example, if there are two types of events with rates r_1 and r_2 , then the total number of events of both types ($n = n_1 + n_2$) that occur in time t is a Poisson distribution with a total rate of $R = r_1 + r_2$. (For N types of events, $R = \sum_{i=1}^N r_i$ and $n = \sum_{i=1}^N n_i$) And, it follows that

$$P(n) = \frac{(Rt)^n}{n!} e^{-Rt} \quad . \quad (10)$$

These ideas and developments of Poisson statistics may be found in W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1975); E. Cinlar, *An Introduction to Stochastic Processes* (Prentice-Hall, Englewood Cliffs, 1975); and, H. Stark and J.W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers* (Prentice-Hall, Englewood Cliffs, 1986).

3.2 Mean Time between Events and Real Time

Mean Time: An important additional attribute of the Poisson process is that the probability density of time t between successive events is

$$P(t) = R e^{-Rt} \quad . \quad (11)$$

From this probability density, the mean time between successive events is $\langle t \rangle = 1/R$. In a non-stationary situation (evolving towards equilibrium) all still holds but the **rates are time dependent**, see below.

Real Time via MC: If a system with processes under study can be cast as a set of independent events with known rates, then the time between events has an exponential distribution (Eq. 11) that can be sampled to determine the time that it takes for an event to occur. To see this from Eq. (11), note that the time to an event is

$$T(\tau) = \int_0^\tau dt' R e^{-Rt'} = 1 - e^{-R\tau} \quad , \quad (12)$$

which lies between [0,1]. So, a random variable $U = e^{-R\tau}$ is uniformly distributed between [0,1], and this relation allows us to obtain the real time, τ (in units of MCS), between successive events:

$$\tau = \frac{-\ln U}{R} \quad \text{Real Time for Event Occurance} \quad . \quad (13)$$

This random sampling of the Poisson time distribution for each chosen event ensures that a direct and unambiguous relationship between MCS and real time step is established.

Time in the General Case: Note that the basic ideas leading to the Poisson distribution of times and events are equally applicable to systems which are non-stationary and evolving towards equilibrium. The non-stationary Poisson process has overall rates that are a function of time. As a result, if the MC algorithm is made to simulate the Poisson process, then the relationship between MC time and real time can be given a firm basis in both static and dynamic situations. For the general case of N species (e.g. [8]), k events are characterized by the rates $\{r\} = \{r_1, r_2, \dots, r_k\}$, which are consistent with average dynamic barriers arising from the potential energy surface. The N species can partition among various possible transition events as $\{n\} = \{n_1, n_2, \dots, n_k\}$, where n_i is the number of species capable of undergoing a transition with a rate r_i and $N = \sum_{i=1}^k n_i$. If a sufficiently large system is used to achieve independence of the various events, then the MC algorithm effectively simulates the Poisson process and the passage of real time can be maintained in terms of $\{r\}$ and $\{n\}$. For each trial i in which an event is realized, the time should be updated with an increment τ_i selected from an exponential distribution with $R_i = \sum_{j=1}^k n_j r_j$.

Appendix: More Formal Derivation of Unambiguous Time

Note that $X(t) = 1 - T(t)$ is the time for an event **not to occur**. The Functional Inverse of $X(t)$, i.e. $X^{-1}(t)$, gives a relation of time to sampling distribution. And, clearly, $X^{-1}(\tau) = \ln X(\tau)$, so $\ln X(\tau) = -R\tau$. As $X(\tau)$ is a random number between [0,1], let it be U , a pseudo-random number from a uniform distribution. This gives the results for real time, τ , (in units of MCS) given in Eq. 13 and guarantees the Poisson distribution. See Appendix of M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Oxford Press, 1987).

Appendix: Derivation of Poisson distribution

Let us follow radioactivity example taken from M.L. Boas, *Mathematical Methods in the Physical Sciences* 2nd Edition (Wiley, 1983). Consider an event such as radioactive particle emission. We count the particles emitted during our observation. Assume that the average count rate does not decrease during our observation (i.e., it is less than 1/2-life of substance).

Probability 1 particle is emitted during a small time interval Δt is $\mu\Delta t = \text{constant}$, if Δt is short enough so that the probability of two particles being emitted is negligible. Let $P_n(t)$ be the probability of n counts in time interval $t + \Delta t$. For $n > 0$, this is sum of probabilities of two mutually exclusive events:

1. “ $n + 0$ particles in t , 0 in Δt ”
2. “ $n - 1$ particles in t , 1 in Δt ”. Therefore,

$$P_n(t + \Delta t) = P_n(t)P_0(\Delta t) + P_{n-1}(t)P_1(\Delta t) \quad (14)$$

$$P_1(\Delta t) = \mu\Delta t \quad (1 \text{ particle in this time})$$

$$P_0(\Delta t) = 1 - \mu\Delta t \quad (0 \text{ particle in this time}) \quad . \quad \text{Thus,}$$

$$P_n(t + \Delta t) = P_n(t)[1 - \mu\Delta t] + P_{n-1}(t)[\mu\Delta t]. \quad (15)$$

For $\Delta t \rightarrow 0$, Eq. 15 re-arranges to

$$\frac{dP_n(t)}{dt} = \mu(P_{n-1}(t) - P_n(t)). \quad (16)$$

For $n = 0$, Eq. 16 simplifies since the only possible event is “no particle in t , no particles in Δt ”; so, $dP_0(t)/dt = -\mu P_0(t)$ and $P_0(t) = e^{-\mu t}$, because the probability that no particle is emitted during a zero time interval is $P_0(0) = 1$. Substituting this result into Eq. 16 gives $P_1(t) = \mu t e^{-\mu t}$. Continued substitution gives $P_2(t)$, $P_3(t)$, ..., $P_n(t)$, where $P_n(t)$ is found to be Eq. 10. With $t = 1$, n is the counts per unit time, $\mu = \langle n \rangle$.

Example 1: Consider 1800 radiation counts over 10 hours.

Question: During how many 1-minute intervals should we expect to observe no particles, 1 particle, etc.? We have $\langle n \rangle = \mu = 1800 \text{ counts}/(10 \text{ hrs} * 60 \text{ min}) = 3 \text{ counts/min}$. From the Poisson distribution,

$$P_n = \frac{3^n}{n!} e^{-3} \quad \text{and, hence,} \quad P_0 = e^{-3} = 0.05. \quad (17)$$

In other words, no particles in about 5% of the 600 1-minute intervals, or 30 1-minute intervals.

Example 2: Approximate binomial distribution

Question: If $n = 1000$ people select a random number between 1 and 500, what is probability that 3 people select the number 29?

Normally we would use the binomial distribution to find this with probability $p = 1/500$ and $n=1000$. That is,

$$\frac{1000!}{3!997!} \left(\frac{1}{500}\right)^3 \left(\frac{499}{500}\right)^{997} = 0.1806 \ .$$

But we can make an excellent approximate to the binomial distribution by using the Poisson distribution, which is much easier to evaluate. This is only good for a large n and small $np = \langle n \rangle$. Let $p \ll 1$ and $np \ll n$ such that $n \rightarrow \infty$, $p \rightarrow 0$ and $np = \langle n \rangle = \text{constant}$. From values given above, $\langle n \rangle = np = 1000(1/500) = 2$, we find that

$$P(n = 3; \langle n \rangle = 2) = \frac{(\langle n \rangle)^n}{n!} e^{-\langle n \rangle} = \frac{(2)^3}{3!} e^{-2} = 0.1804 \ .$$

The Poisson distribution is indeed a very good approximation.

3.3 Adsorption-Desorption in Lattice Gas Model

A simple example following Fichtorn and Weinberg [8].

Consider adsorption of a gas-phase species A with a solid, single-crystalline surface. Equilibrium occurs when chemical potentials of gas-phase and chemisorbed A are equal. Equilibrium is established from a kinetic point of view when the net rate of chemisorption of gas-phase A is equal to net rate of desorption of A to gas-phase.

Analytic Solution - Using the simple lattice-gas model, we can assume adsorbed molecules do not interact and each molecule requires one site for adsorption. In addition, we assume the molecules arrive at surface at random, uncorrelated times and can be characterized by an average rate r_A . A similar situation applies to atoms chemisorbed on the surface so that all microscopic influences induce a desorption which occurs with an average rate r_D . Thus, this kinetically maintained balance follows

$$\frac{d\theta}{dt} = r_A(1 - \theta) - r_D\theta \ , \quad (18)$$

where θ is the fractional surface coverage of A (here the initial condition is no coverage or $\theta(t=0)=0$). So the obvious solution of the above equation is

$$\theta(t) = \frac{r_A}{r_A + r_D} (1 - e^{-(r_A+r_D)t}) \underset{t \rightarrow \infty}{=} \frac{r_A}{r_A + r_D} \ , \quad (19)$$

where the last equality shows the equilibrium coverage in terms of rate arising from microscopic phenomena, as well as detailed balance. It should be obvious that the coverage θ plays the same role as probability.

KMC Algorithm - In the flowchart (Fig 2) of the adsorption-desorption the transition probabilities W_A and W_D are chosen to obey detailed balance because they are chosen at random and independently such that a successful adsorption is $W_A(1 - \theta_i)$ and desorption is $W_D\theta_i$. Thus, the total probability of success per trial i is $W_A(1 - \theta_i) + W_D\theta_i \leq 1$, where $\theta_i = M_i/N$ for M_i occupied sites out of N . Detailed balance is mathematically guaranteed because for T trials of adsorption events the probability is

$$P(n_A) = \left(\frac{T}{n_A}\right) [W_A(1 - \theta_i)]^{n_A} [1 - W_A(1 - \theta_i)]^{T-n_A} ,$$

so that average expected adsorption events in T trials is

$$\langle n_A \rangle = W_A(1 - \theta_i)T , \quad (20)$$

with similar result for $\langle n_D \rangle$. Therefore, when steady-state is attained, i.e. when $\langle n_A \rangle = \langle n_D \rangle$, we have $W_A(1 - \theta_{ss}) = W_D\theta_{ss}$, which is *detailed balance*! The actual rate to be used at each time-step is $r_{Ai} = r_A(1 - \theta_i)$ for adsorption and $r_{Di} = r_D\theta_i$ for desorption. Clearly, at each accepted event the rates are updated (but in this case it is a trivial update). Notice also in the flowchart that at each *trial* when some event happens, either desorption or adsorption, the *increment of time* is different for each MCS and determined from sampling *randomly* the Poisson time-distribution, where $R_i = r_{Ai} + r_{Di}$. This may be viewed in the appendix where a code is reproduced which gives the results of simple model and it follows the flowchart in Fig 2. In this code, the acceptance and rejection method is clear as is the sampling of the Poisson distribution in time. Notably, if τ^{-1} is the attempt frequency, then the *real time* for 1 MCS = τ/N , where N is the total number of particles and assuming that each particle is excited only once.

Following this KMC algorithm, we find the results in Fig. 3. Notice that the analytic and KMC results agree very well and reach steady-state (although with fluctuations) in the same real time. The code to derive these results is given in Appendix (or download). You may see the flowchart is followed in the code and that Time is incremented according to the Poisson distribution. Take the code and alter the time increment (e.g., make it the same quasi-random number that decides on the event, adsorption or desorption), you will see that the theory curve and KMC simulation will not agree! Note that the simple code and problem does not require the “n-fold-way” and fancy data structures because it is so fast, even for modest size lattices (400x400).

Flowhart for Adsorption-Desorption KMC Model

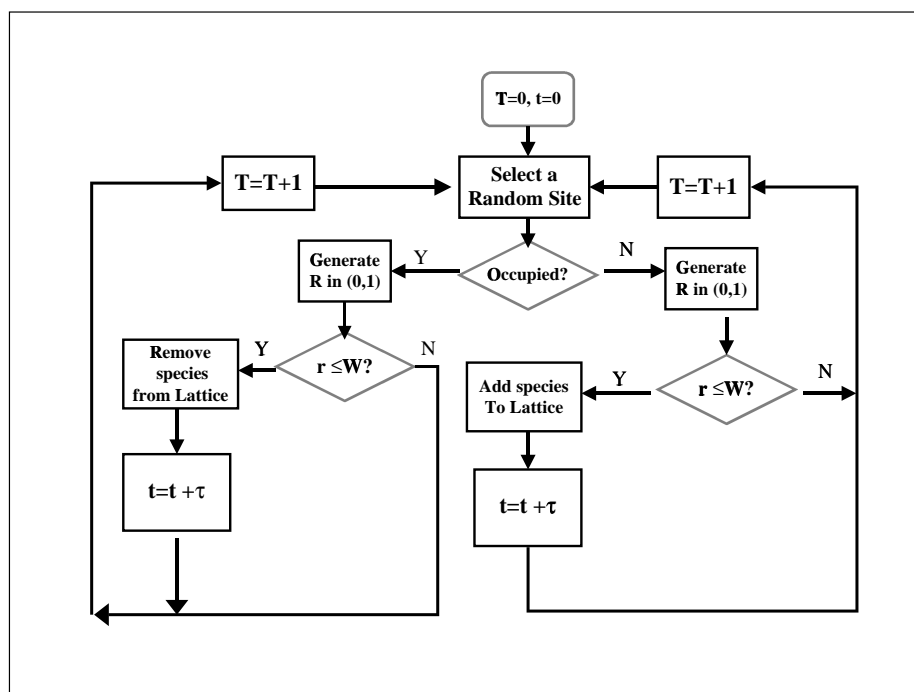


Figure 2: Flowchart from simple lattice-gas from Ref. [8] showing how real time is incremented (with varying amounts of time depending on the transition rates in a Poisson process) only when transition occurs. For complex cases, the efficiency of this approach can be highly improved by using a residence-time algorithm.

Results for Adsorption-Desorption Model

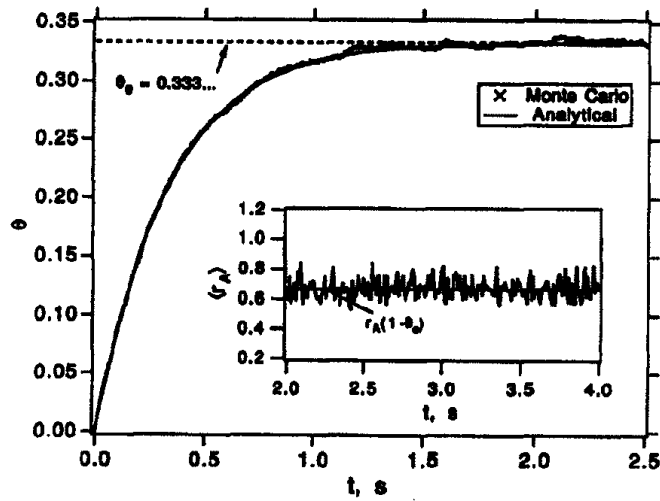


FIG. 2. The transient solution of Eq. (12) for the initial condition of an empty surface [$\theta(t=0) = 0$] provided by both the analytical form of Eq. (13) and the Monte Carlo algorithm of Fig. 1 with $r_A = 1.0$ (site s) $^{-1}$ and $r_D = 2.0$ (site s) $^{-1}$ ($W_A = 1/2$ and $W_D = 1$). The inset depicts the rate of adsorption measured from the Monte Carlo simulation at steady-state and the continuum steady-state rate.

Figure 3: KMC results for simple adsorption-desorption lattice-gas model (analytic vs. KMC) [8]. A code for re-producing their results can be downloaded.

3.4 The KMC Bottom Line

KMC simulations exhibit rate changes with each sampling because the configuration has changed (e.g., coverage or barriers); the possible events change each MCS; and, the elapsed time since the last event is adjusted depending on the probability of the transition. For KMC, the time increment must be scaled by the average time between successful events in the conventional MC scheme, which, as we have seen, is given by a Poisson distribution from the time may be sampled each time step. In general, you can make set up the KMC such that every trial some event happens (there are no wasted MCS), if improved algorithms and data structures are used (see below). Conventional MC simulation time is incremented by constant amount after each attempted event, whether or not event was successful.

4 Algorithms for “Faster” MC

In terms of algorithms, Kinetic or equilibrium Monte Carlo simulations of a system of size M is based on the so-called “ n -fold-way” by Bortz, Kalos, and Liebowitz [4] (or, in more modern references, the “residence-time algorithm”). Such an algorithm scales as $O(M^{1/2})$ (see below). More recently, these have been supplanted by improved algorithms that scale as $O(\log_2 M)$ or $O(M^{1/K})$ that use clever data structures and K -level searches, e.g., [15, 7, 19, 3]. Keep in mind that efficient algorithms and data structures can be just as important as the physical problem, speeding up simulations by orders of magnitude. Hence, in the following, I cover a few of the important algorithmic ideas for performing KMC type calculations.

4.1 MC for Barriers: n -fold-way

In the case of surface diffusion, for example, using the W_E or ω_E distribution, one requires many attempts before a surface atom hop is successful to a neighboring site. In such a case, which occur at low T , MC should not be done by picking particle at random. Rather, a better approach is as follows (see, e.g., Ref. [3]). Let M be the number of possible MC events (e.g. adsorption, hops, or desorption), and let r_i be the rate for the i^{th} event. The total rate (as before) is $R = \sum_{i=1}^M r_i$. For a proper simulation, the i^{th} event should occur with probability, $P_i = r_i/R$, as should be obvious. (Keep in mind that in standard MC, the $\{r_i\}$ are transition probabilities, W_i , not rates.) We then have the following procedure:

1. GROUP all possible events according $\{r_i\}$ (or W_E),
2. GET a uniform quasi-random number $U \in [0, 1]$ so that $UR \in [0, R)$.
3. FIND the particular event according to relative rates (or relative probability in MC):

$$\sum_{i=1}^{\nu-1} r_i \leq UR < \sum_{i=1}^{\nu} r_i \ .$$

(Note that the, by dividing by R , you are using Probabilities in uniform quasi-random number between $[0,1)$.)

4. CARRY out the event ν , which was selected at random.
5. UPDATE those r_i that have changed as a result of event ν ; update R and any data structures being used.
6. INCREMENT TIME by τ from the Poisson distribution.

Steps 2,4, and 6 are independent of M . Steps 3 and 5 are time consuming. Depending on the data structure needed for the search algorithm, step 5 may or may not grow with M . For example, if a simple linear search method is used then steps 3

scales as $O(M)$ and step 5 scales at most as $O(M)$. That is, this is a linear-scaling method. This algorithm for MC simulation determines the probability of success of an event before the event is picked (rather than after). It was first developed by Bortz, Kalos and Lebowitz [4] and is commonly referred to as the *n-fold-way*, or *residence-time algorithm*.

4.2 Binning Methods

The Binning method proposed by Maksym [15] allows a search to be made over partial sums of rates (or probabilities). Choose sums containing g rates, such that there are M/g partial sums given by

$$S_j^{(2)} = \sum_{i=1+(j-1)g}^{jg} r_i \quad j=1,\dots,M/g \quad . \quad (21)$$

Now we have a 2-Level (binary chopping) search in Step 3, as suggest by the superscript on the partial sum, S, where $K=2$. (i) Search for the correct bin (an $O(M/g)$ time cost) to find the type of event, and, once the type of event is found, (ii) Search within a bin or group (an $O(g)$ time cost) to find the site for event. Updating in Step 5 is independent of M . Minimizing with respect to g (bin sizes) leads to $O(M^{1/2})$: The time for update of the group is $\propto M/g$, while the site search is $\propto g$. The total time is then $T = aM/g + bg$, where a and b are constants. Clearly, by minimizing w.r.t. g , we get optimal group size $g = \sqrt{aM/b}$ and optimal time $T = 2\sqrt{abM}$. Hence the $O(M^{1/2})$ scaling for $K=2$ Binning Searches.

This method was generalized to a K -Level search by Blue, Beichl, and Sullivan [3]. For $K=2$, it is the method of Maksym [15]. Higher-levels of binning are defined recursively, namely,

$$S_j^{(k)} = \sum_{i=1+(j-1)g}^{jg} S_i^{(k-1)} \quad , \quad j=1,\dots,M/g^{K-1}, \quad k=3,\dots,K \quad . \quad (22)$$

Minimizing time for searches yields g is $O(M^{1/K})$ and the total time is $O(KM^{1/K})$ for step 3, whereas step 5 is $O(K)$ because K partial sums need to be updated. Of course, there is more storage for data, being $O(M^{1-1/K} + M^{1-2/K} + \dots + M^{1/K})$.

For a fixed M , the best time can be achieved by largest K that gives only $g = 2$ terms in each partial sum, i.e. $2 = M^{1/K}$ or $K = \log_2 M$. This is a binary tree search and both step 3 and 5 are $O(\log_2 M)$. Storage space is then $O(M)$. Please see Ref.[3] for algorithm for recursive tree searches.

The results of such a K -level simulation for $N \times N$ epitaxial growth was given by Blue et al. [3]. It is reproduced here for convenience., see Fig. 4 The substantial reduction in MC time versus system size is apparent. The binary-tree search method was originally proposed by Wong and Easton [19] and is equivalent to $K=5$, although for larger system sizes K may have to be increased to continue this trend.

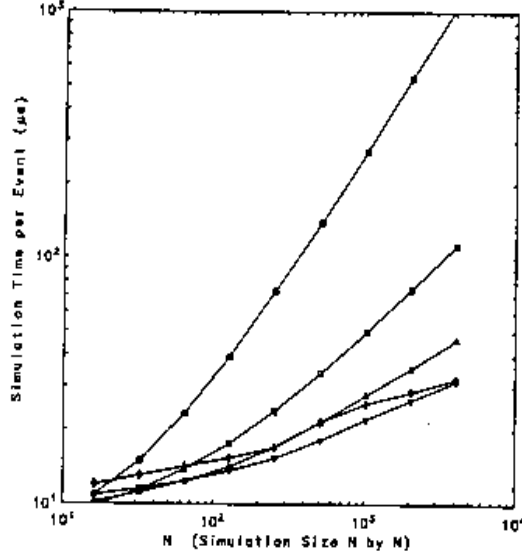


FIG. 1. Log-log plot of time per event for different search schemes. \circ , $K=2$; \square , $K=3$; \triangle , $K=4$; ∇ , $K=5$; \diamond , binary tree

Figure 4: From Ref. [3], note the log-log plot shows the dramatic reduction in CPU time. The binary-tree search algorithm (curve approaching the bottom $K=5$ curve) is an efficient way to organize and search data structure within KMC algorithm [15, 7, 4], e.g., “n-fold-way”.

4.3 Example for $K=2$: MBE growth

Let us make a connection to previous model and scaling discussed above using Maksym’s discussion of MBE growth. In this MBE growth model there are adsorption (a), hopping (h) and desorption (d) events possible, similar to the simple model discussed earlier. Recall that we should select time, type of event, and site of event beforehand. For M surface sites, we have, as before, the total rate

$$R = \sum_{i=1}^M (a_i + h_i + d_i) . \quad (23)$$

Again time increment will be from this rate, $\tau = -\ln U/R$. The type and site of event are selected from the conditional probabilities found from the total rates of adsorption, hopping, or desorption, i.e.

$$A = \sum_{i=1}^M a_i \quad H = \sum_{i=1}^M h_i \quad D = \sum_{i=1}^M d_i . \quad (24)$$

Thus their probabilities are A/R , H/R , and D/R . Similarly, the conditional probabilities for event to occur at a specific site i are a_i/R , h_i/R , and d_i/R .

The procedure outlined in the last section can be numerically done as follows. Pick a second random number $U_2 \in [0, 1]$. The type of event is chosen as:

1. ARRIVAL if $RU_2 < A$.
2. HOPPING if $A \leq RU_2 < H$.
3. DESORB if $RU_2 \geq H$.

After type of event is determined, the site i of event is chosen as, for example:

1. If ARRIVAL, $RU_2 \geq \sum_{j=1}^{i-1} a_j$.
2. If HOPPING, $RU_2 - A \geq \sum_{j=1}^{i-1} h_j$.
3. IF DESORB, $RU_2 - (A + H) \geq \sum_{j=1}^{i-1} d_j$.

This method was first discussed actually by Ref. [7].

The data search is be done as K=2 Level search (binary chopping) to find the type of event. Once the type of event is found, a search to find the site can be performed within that event group. As given above, the time for update of the group is $\propto M/g$, while the site search is $\propto g$. The total time is then $T = aM/g + bg$, so that the optimal group size and time are $g = \sqrt{aM/b}$ and time $T = 2\sqrt{abM}$, respectively. Hence the $O(M^{1/2})$ scaling mentioned in the beginning.

Maksym [15] exemplified this on a 100×100 lattice by plotting *Time per Event* in milli-seconds versus Logarithm Base 10 of Group Size. He found a factor of 100 reduction even for this small system size. The K-level version is even more efficient, as seen in Figure above.

In closing this section, it is worth mentioning that Hassold and Holmes [11] has developed the n-fold-way algorithm for domain growth in materials based upon a Pott's Model (an Ising Model with N different spin orientations). In such a model, each spin flip alters the domain structure. In section C of their paper they discuss how to efficiently select flips based on their relative probabilities. This requires setting up the algorithm outlined above for the n-fold-way, but it is helpful to see it applied to different models. They provide section of coding to see the method and update at work. You may also compare this to 2-D vacancy diffusion in a binary alloy with and without ballistic jumps, as given in Homework.

5 Beware of Approximations to Master Equations

Haider, Bellon and Martin [10] discuss several aspect of KMC for driven alloy systems. In particular, they discuss the effects of ballistics atomic jump (as perhaps obtained under irradiation, in addition to the usual thermal jumps already discussed. Within a very simple alloy model (using mean-field free energies, etc.) they discuss how a kinetic model and a thermodynamic model can be made to obtain the roughly same results. In other words, in the infinite time limit, kinetic equilibrium yields the the thermodynamic expectation, even in the case of ballistic jumps.

This result is obtained, however, only when the probability distributions are sharply peaked and the deviations from the average state are very small. Hence, it is clear this result does not have to be generally true. In fact, it is known not to be true in general. For example, in 1-dimension, the exact solution of the Master Equation is known. By Taylor's series expansion of the Master Eqs. (both $P(x + \epsilon, t)$ and $W(x + \epsilon \rightarrow x, t)$) about such a sharply peaked distribution centered at x and truncating to first-order in ϵ , one obtains the well-known Fokker-Planck equation. The Fokker-Planck equation does not converge to exact solution for large but finite system, as might be expected. In multidimensions, the Fokker-Planck equation does not even fulfill detailed balance, or also the condition of zero-probability flux. (N.B. For a 1-D random walker, or single particle experiencing white (or Gaussian distributed) noise, The Fokker-Planck equation does provide an exact solution to the Master Equation, or Langevin equation. But, with more than 1 particle, or colored noise, or multi-dimensions, this is not the case. So that is why the caution for the user to beware.)

This occurs because $P(x) \propto e^{-N\phi(x)}$, where ϕ is an intensive quantity for a system of N particles. Thus, a Taylor's series expansion of the Master Eq. yields

$$P(x + \epsilon) = \sum_{n=0}^{\infty} \frac{1}{n!} (\epsilon \cdot \nabla)^n P(x) . \quad (25)$$

Surprisingly, if you consider each term, they have contributions of $O(N^0)$, so the series does not converge. Hence, by truncating to any order in ϵ , you have thrown out equally important terms.

Kubo et al. [13] that one must handle ill-convergent functionals with care. Consider expanding intensive variable about the extremal state x , where $x + \epsilon$ is the final state not far away from x ,

$$\phi(x + \epsilon) = \sum_{n=0}^{\infty} \frac{1}{n!} (\epsilon \cdot \nabla)^n \phi(x) . \quad (26)$$

Putting this into Master Eq. gives

$$\frac{\partial \phi(x, t)}{\partial t} = \sum_{\epsilon} W(x \rightarrow x + \epsilon) [e^{\epsilon \cdot \nabla \phi(x)} - 1] . \quad (27)$$

By performing expansion in this manner, one can show that the Master Eq. (using this Kubo method) agrees with the thermodynamic approach except for term of $O(N^{-1})$. You can show that if you just expand $P(x)$ in Master Eq. then the approaches agree only in the vicinity of the extrema of $P(x)$.

In actual (binary A-B) alloy problems, the probability could be directly related to variable of interest. For example, as related to a Homework problem, P could be site concentration of A atoms, i.e., how much A atoms are on given sublattice. This depends on time, thermal history, local environment (the local barriers to diffusion change), and so on. Thus, the Master Equation provides the evolution of the site concentration given the transition probabilities. Additional events can be incorporated which affect this site concentration. Bellon et al. include ballistic jumps as cause by irradiation, for example. (See Homework KMC code.)

Copyright by Duane Johnson May 2001

Note to the consumer:

These notes were created as an overview and introduction to concepts involved in Kinetic Monte Carlo. They are, in some sense, a distillation of many papers, and a means to point the interested to references that may be useful, or to more in-depth information. Invariably, while preparing these notes, errors or unclear statements may have been unintentionally created. As always, use with care (don't believe everything you read, unless you sorted it out for yourself!); nonetheless, they should be helpful in getting started.

Duane Johnson

References

- [1] M. Athenes, P. Bellon, G. Martin, F. Haider, "A Monte Carlo Study of B2 Ordering and Precipitation via Vacancy Mechanism in BCC Lattices," *Acta mater.* **44**, 4739 (1996).
- [2] M. Athenes, P. Bellon, G. Martin, "Identification of novel diffusion cycles in B2 ordered phase by Monte Carlo simulations," *Phil. Mag. A* **76**, 565 (1997).
- [3] J.L. Blue, I. Beichl and F. Sullivan, "Faster MC Simulations," *Phys. Rev. E* **51** R867 (1995). This uses a generalization of Maksym's method, which, in the K th-level limit, behaves like a *Binary Tree Search* Algorithm of Wong and Easton.
- [4] A.B. Bortz, M.H. Kalos, J.L. Lebowitz, *J. Comput. Phys.* **17**, 10 (1975). The "residence-time algorithm".
- [5] M.M. Clark, L.M. Raff, and H.L. Scott, "Hybrid Monte Carlo method for off-lattice simulation of processes involving steps with widely varying rates," *Computers in Physics* **10**, 584 (1996).
- [6] R.A. Enrique, P. Bellon, "Compositional patterning in alloys under irradiation," *Phys. Rev. B* **63**, 13 411 (2001).
- [7] W. Fawcett, A.D. Boardman, S. Swain, *J. Phys. Chem. Solids* **31**, 1963 (1970), originally suggested by in electron transport simulations.
- [8] K.A. Fichthorn and W.H. Weinberg, "Theoretical Foundation of Dynamic Monte Carlo Simulations," *J. Chem. Phys.* **95**, 1090 (1991).
- [9] D.T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).
- [10] F. Haider, P. Bellon and G. Martin, "Non-equilibrium transitions in driven AB_3 compounds on the fcc lattice: A multivariant master-equation approach," *Phys. Rev. B* **42**, 8274 (1990).
- [11] G.N. Hassold and E.A. Holmes, "A fast serial algorithm for the finite temperature quenched Potts model," *Comp. in Physics* **7**, 97 (1993)
- [12] Kang and Weinberg, "Dynamic Monte Carlo with Proper Energy Barriers: surface diffusion...", *J. Chem. Phys.* **90**, 2824 (1989).
- [13] Kubo et al., *J. Stat. Phys.* **9**, 51 (1973)
- [14] K. Lin, D.C. Chrzan, "Kinetic Monte Carlo simulation of dislocation dynamics," *Phys. Rev. B* **60**, 3799 (1999).
- [15] P.A. Maksym, "Fast MC Simulation of MBE Growth," *Semicond. Sci. & Technol.* **3**, 594 (1988). This method uses a two-level search algorithm, like a tree with only two branches, as originally suggested by Fawcett, Boardman and Swain (see ref.) in electron transport.
- [16] M. Muller, J.P Wittmer, J-L Barrat, "On two intrinsic length scales in polymer physics. Topological constraints vs. entanglement length," *Europhysics Letters* **52**, 406 (2000).

- [17] L.A. Ray and R.C. Baetzold, “A Monte Carlo estimation of surface diffusion by simulating laser-induced thermal desorption,” *J. Chem. Phys.* **93**, 2871 (1990).
- [18] Reed and Ehrlich, “Surface Diffusivity and Time Correlation of Concentration Fluctuations,” *Surface Sci.* **105**, 603 (1981).
- [19] Wong and Easton, *SIAM J. Comput.* **9** 111 (1980). The original *Binary Tree Search* Algorithm, similar to K-level search.