LA-UR-98-2136

Approved for public release; distribution is unlimited.

Title: ACCELERATING THE DYNAMICS OF INFREQUENT EVENTS: COMBINING HYPERDYNAMICS AND PARALLEL REPLICA DYNAMICS TO TREAT EPITAXIAL LAYER GROWTH

CONF-980405--

Author(s): A. F. Voter
T. C. Germann

Submitted to: Proceedings volume from the 1998 Spring MRS Meeting.

Los Alamos
NATIONAL LABORATORY

DISCLAIMER

ACCELERATING THE DYNAMICS OF INFREQUENT EVENTS: COMBINING HYPERDYNAMICS AND PARALLEL REPLICA DYNAMICS TO TREAT EPITAXIAL LAYER GROWTH

A. F. VOTER and T. C. GERMANN

Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545;
{afv,tcg}@lanl.gov

ABSTRACT

During the growth of a surface, morphology-controlling diffusion events occur over time scales that far exceed those accessible to molecular dynamics (MD) simulation. Kinetic Monte Carlo offers a way to reach much longer times, but suffers from the fact that the dynamics are correct only if all possible diffusion events are specified in advance. This is difficult due to the concerted nature of many of the recently discovered surface diffusion mechanisms and the complex configurations that arise during real growth. Here we describe two new approaches for this type of problem. The first, hyperdynamics, is an accelerated MD method, in which the trajectory is run on a modified potential energy surface and time is accumulated as a statistical property. Relative to regular MD, hyperdynamics can give computational gains of more than 10^2. The second method offers a way to parallelize the dynamics efficiently for systems too small for conventional parallel MD algorithms. Both methods exploit the infrequent-event nature of the diffusion process. After an introductory description of these methods, we present preliminary results from simulations combining the two approaches to reach near-millisecond time scales on systems relevant to epitaxial metal growth.

INTRODUCTION

Atomistic simulations are playing an increasingly important role in materials science. Developments in empirical potentials, dynamical methods, and computer speed have all contributed to the power of simulations in addressing real problems that were insoluble just ten years ago. Molecular dynamics (MD) simulations can access times up to tens of nanoseconds for systems with thousands of atoms. On a large parallel computer, millions of atoms can be simulated. This time scale is long enough to investigate phenomena such as surface sputtering [1], surface and bulk self-diffusion at high temperature [2], surface melting [3], ion-implant cascades [4], and high-strain-rate crack propagation [5].

For activated processes with identifiable transition states, this nanosecond time-scale can be extended considerably using transition state theory (TST) [6] and dynamical corrections [7]. For example, this allows one to study some surface and bulk diffusion mechanisms. For more complex processes, such as diffusion within a submonolayer overlayer on an ideal crystal surface, it is sometimes possible to compute barrier heights for all possible environments that can influence an atomic jump event (e.g., see Figure 1) and use this catalog of rate constants in a stochastic algorithm known as kinetic Monte Carlo [8]. If all diffusion mechanisms are known, this method gives the exact dynamics for the state-to-state evolution of the system.

Despite these advances, however, there remains a class of important problems that cannot be addressed with any of these methods. These problems are characterized by intrinsic time scales that greatly exceed the nanoseconds accessible to direct MD, and a complexity that precludes the use of cataloged, kinetic Monte Carlo methods. In these systems the required catalog is prohibitively large

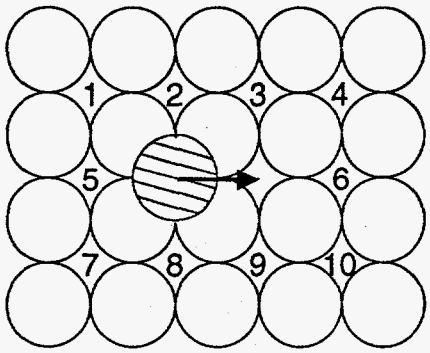

Figure 1. Rate catalog template for adatom hops on metal fcc(100) surface. Each of the ten sites can be empty or occupied, for a total of 2^{10} possible environments affecting the hop rate. The catalog of these 1024 rates, computed using transition state theory, provide input for a kinetic Monte Carlo simulation.

and/or the dynamical evolution involves complicated, many-particle reaction coordinates that defy categorization and cannot be guessed in advance. Processes in this class include vapor-deposited film growth, diffusive motion of a dislocation, and annealing of crystal damage after ion implantation or neutron bombardment.

As an illustration of this complexity problem, consider the fcc(100) surface, perhaps the simplest possible surface of a metal. For years, a hopping mechanism (shown in Figure 1) was assumed to be the primary diffusion mechanism for an adatom on this surface. In 1990, this simple picture was shattered when Feibelman [9] discovered the exchange mechanism (shown in Figure 2), in

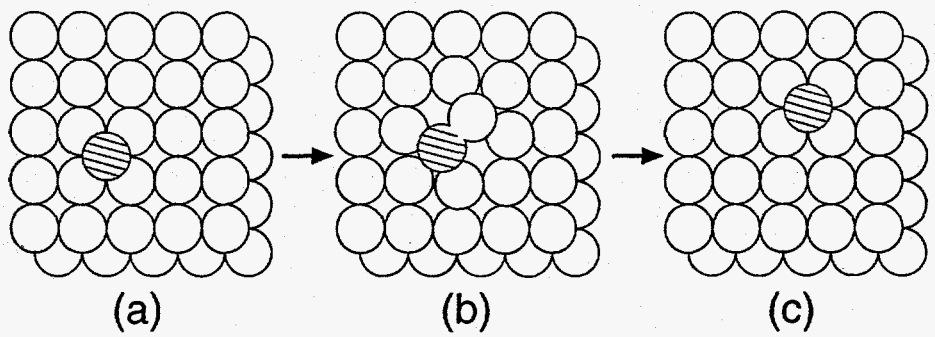

Figure 2. Exchange mechanism on metal fcc(100) surface. Until 1990, adatom diffusion on fcc(100) was thought to occur solely by adatom hops, as indicated in Figure 1. This exchange event is now known to be the dominant diffusion mechanism for some fcc metals, including platinum, iridium and aluminum.

which the adatom plunges into the surface, ejecting a top-layer substrate atom into the second-nearest-neighbor site. (An exchange mechanism was previously known to occur on the channeled fcc(110) face [10–13], but was not expected for the smooth, close-packed fcc(100) face). In field ion microscope experiments, this exchange event was soon shown to be the dominant mechanism

for both Pt(100) [14] and Ir(100) [15]. Complicating matters further, this mechanism is known to be favored on some fcc metals and not others [16], and becomes more or less important depending on the environment. Since then, many complex surface diffusion mechanisms have been discovered (e.g., see [17–24]) shaking our confidence about whether the picture is complete, even now. For example, we now know that whole clusters of atoms (formerly thought to be immobile) can move in a concerted fashion [19,21–23], exchange events can involve three or more atoms [17,25], and the barrier for an adatom to move down a step edge (the Ehrlich-Schwoebel barrier [26]) usually involves exchange [20,27]. Moreover, in real surface growth, complex geometries can form, with interacting steps, intersecting facets, and multiple, partially formed layers. Ensuring that a catalog of rate constants is complete, or even that it contains the essential physics for simulating film growth, is thus a daunting, if not impossible, task.

Although the dynamical evolution of a system can be accelerated by raising the temperature of the simulation, in general this causes the system to proceed along different pathways. For example, the growth mode of a metal crystal during vapor deposition can change from layer-by-layer to three-dimensional growth and then back to layer-by-layer as the temperature is raised [28], due to the delicate interplay of the different diffusion barriers.

This paper describes two new approaches for investigating systems of this type. The first, hyperdynamics,is a direct dynamical approach with a built-in accelerator, eliminating the problem of explicitly knowing all the possible escape paths from a each state. This method has the character of a regular molecular dynamics trajectory, in that it "finds" its way out of a potential energy basin without knowing anything in advance about the possible escape paths. In the second approach, the parallel replica method, the infrequent-event nature of the system is exploited in a new way to design a parallel processing implementation that is efficient even for relatively small systems. After an introductory description of the two methods, we present preliminary results from simulations combining the two approaches to acheive nearly a millisecond of simulation time on a system relevant to epitaxial growth on Ag(100). This extension of the time scale by a few orders of magnitude is still not sufficient to simulate realistic deposition rates, but it is a big step closer.

The treatment here is intentionally pedagogical, omitting many details and some rigor. The reader can consult the original references [29–31] for a full presentation.

HYPERDYNAMICS

As discussed above, the goal is to develop a method for extending the time available to a molecular dynamics (MD) simulation. We focus on the category of "infrequent event" systems, in which the trajectory spends the vast majority of its time simply vibrating in a basin of the potential energy surface. Occasionally, it finds an escape path, making a transition to a new potential basin by passing over the ridge-top separating the two states (this ridge-top boundary is known as the dividing surface). For a short time following this initial escape, the system still has additional energy localized in the reaction coordinate, and may make additional crossings of this same dividing surface, or may pass through another dividing surface to yet another state.. These are referred to as correlated dynamical events. For example, an atom diffusing on a surface will sometimes make a double jump, rather than simply hopping to the nearest-neighbor site. Nonetheless, after a brief period of time (a couple of picoseconds for a metal), the excess energy dissipates into other vibrational modes in the system, and the trajectory settles into some potential basin, losing all memory of how it arrived in that state. We will refer to this time until memory is lost as the correlation time. The large disparity between the correlation time and the average time between transitions is what defines an infrequent event system. This separation of time scales also gives rise to some useful statistical properties, which we exploit to derive the two methods.

Assuming that one knows the position of this ridge-top dividing surface surrounding a state, a good approximation to the escape rate can be calculated without ever running a trajectory, by appealing to transition state theory (TST) [32,33]. In the TST approximation, the rate constant is taken as the outgoing flux through the dividing surface. This is an equilibrium property of the system, depending on the shape and energy of the many-dimensional basin, and on the shape and energy of the ridge top. If the system never executes correlated dynamical events (i.e., if every time a trajectory crosses a dividing surface, it settles down in the new state without any recrossings), the TST approximation is exact (e.g., see [34]). For diffusive processes in materials, TST is an excellent approximation, with the correlated dynamical events causing only a few-percent effect on the rate constant. With the TST approach, the average time until an escape event occurs can be computed directly, even if this escape time is much longer than the feasible simulation time for a trajectory.

In situations where the *exact* rate constant is desired, it can be obtained by computing a dynamical correction factor that multiplies the TST rate constant. This correction factor is straightforward to calculate from the results of trajectories that are initiated at the dividing surface (with properly Boltzmann-sampled starting points) and evolved for a time a bit longer than the correlation time [7,35,36]. An appealing aspect of computing the rate constant in this way is that the product of the TST rate with the dynamical correction factor gives the exact rate constant regardless of the location of the dividing surface. Thus, for example, if the dividing surface does not lie along the peak of the ridge, the TST rate constant will increase, but the dynamical correction factor will decrease to compensate. I.e., some trajectories that pass through this "bad" dividing surface (which contribute to the flux, and hence the TST rate) will not actually make it over the ridge top. These trajectories that fall back into the initial state contribute negatively to the dynamical correction factor. Thus, transition state theory and dynamical corrections theory make a robust framework from which to develop alternative approaches for simulating the dynamical evolution in a many-state system. Perhaps surprisingly, this even allows construction of methods that accelerate the dynamical evolution of systems in which the locations of the dividing surfaces are unknown.

In the hyperdynamics method, we assume that there are no correlated recrossing events – i.e., that TST is exact. As mentioned above, this is an excellent approximation for the types of materials processes we wish to study. The basic concept is to use the properties of transition state theory to define a modified dynamical system that gives accelerated transitions from state to state. We derive it qualitatively here, while the full derivation can be found in Refs. [29,30].

Consider a trajectory in a basin (state A) of a potential energy surface. It wanders around in this basin for many vibrational periods (perhaps thousands or millions) until it happens to find the dividing surface that leads to an adjacent state. From transition state theory, we know that the average time until it finds this surface is simply a property of the shape and energy of the dividing surface relative to the shape and energy of the basin of state A. More precisely, it is proportional to the ratio of the partition function of the 3N-1 dimensional dividing surface to the partition function of the 3N dimensional basin. If we modify the basin of state A to have less configuration space, thus lowering its partition function, we raise the escape rate in a predictable way.

A schematic example of this type of modification is shown in Figure 3. The energy of the state-A basin has been raised by adding a bias potential, $\Delta V_b(\mathrm{R})$. In addition, the basins of states B and C have been raised, as have the other states of the system. Note that we have taken care not to change the energy along the dividing surfaces. This is a requirement of the hyperdynamics method, and a key point is that we can accomplish this even when we do not know in advance where the dividing surfaces may be. This is demonstrated and discussed below. The other requirement on the bias potential is that it preserve the TST-obeying nature of the dynamics; i.e., whether a trajectory

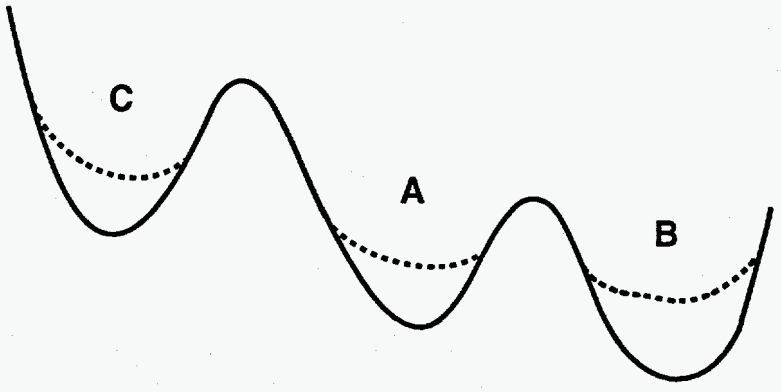

Figure 3. Illustration of the hyperdynamics method. A trajectory is evolved on the biased potential (dashed line) rather than the original potential (solid line). Relative escape probabilities are preserved because the bias potential is zero at the dividing surfaces. The accelerated time scale is calculated as the simulation proceeds.

is run on the original potential surface or on the biased potential surface, correlated crossing events should not occur. (In practice, correlated events may occur occasionally, but the formal derivation assumes they never do.)

A dynamical simulation on a surface modified in this way has two desirable properties. First, the escape rate from each state is enhanced. This is obvious, as the wells are not as deep as for the original potential. Second, the relative rates of escape from state A to the various adjacent states (B and C in this case) are not affected. This is not so obvious, but results from the fact that each rate is a ratio of the dividing surface partition function to the basin partition function. Taking the ratio of two rates thus eliminates any dependence on the basin itself, leaving only the ratio of dividing surface partition functions. The consequence of these two properties is that a trajectory run on this surface will make transitions at an accelerated rate, and will execute a sequence of state-to-state transitions that is indistinguishable from that of a long trajectory on the original potential surface. The details of the dynamical behavior within any of the basins are irretrievably changed by the bias potential, but this is a sacrifice we happily make to achieve the accelerated transition rate.

The simulation time scale, which is accelerated in a nonlinear fashion (the bias potential will, in general, increase the escape rates more for some states than others) can be recovered from an extremely simple procedure. At each integration step during the trajectory, the boosted time clock is advanced by the normal MD time step (Δt_{MD}) multiplied by the inverse Boltzmann factor for the value of the bias potential at that point in configuration-space. For a trajectory that has evolved for n steps, the time on the MD clock is

$$t_{MD} = \sum_{i}^{n} \Delta t_{MD}, \tag{1}$$

while the total boosted time (or hypertime) is

$$t_{hyper} = \sum_{i}^{n} \Delta t_{MD} \exp(\Delta V_b(\mathbf{R}_i))/k_B T), \tag{2}$$

where $\mathbf{R}_i$ is the position of the trajectory at step i, k_B is the Boltzmann constant, and T is the temperature. In the course of a simulation, this instantaneous boost factor ($\exp(\Delta V_b(\mathbf{R})/k_B T)$) fluctuates (possibly wildly), but over a long run, gives a well defined average boost factor. This definition for the boosted time converges on the exact time in the long-time limit, even though the

system may escape from each state long before adequate sampling of the bias potential is obtained, and even if every state of the system is different [29].

To summarize the hyperdynamics formalism: assuming a bias potential can be constructed that raises the potential basins without affecting the dividing surfaces, a trajectory run on this biased potential surface evolves from state to state in a proper sequence. Time is no longer an independent variable, but is instead estimated statistically as the simulation runs, in a way that gives zero relative error at long simulation times. In contrast to a kinetic Monte Carlo approach, hyperdynamics has the character of a regular molecular dynamics simulation. It simply wanders around in each potential basin until it accidentally passes through a dividing surface into another state. Despite not knowing anything about the other possible escape routes, it picks one with exactly the correct Boltzmann probability. On the other hand, hyperdynamics is still considerably slower than kinetic Monte Carlo.

This hyperdynamics formalism is only useful if it is possible to design a computationally efficient bias potential that is zero at the dividing surfaces without advanced knowledge of their location. In principle, the exact position of a dividing surface can be determined by following steepest-descent paths downhill from a number of different geometries. The dividing surface separates those paths that fall into one basin from those that fall into another basin. While this procedure is mathematically well defined, we want a definition that can tell us how far the trajectory is from a dividing surface, so that the bias potential can be turned off smoothly as the dividing surface is approached. This type of definition can be created using the local slope and curvature properties of the potential surface. Specifically, the gradient vector $\mathbf{g}$ gives the slope of the surface, and the the matrix of second derivatives (the Hessian, $\mathbf{H}$) gives the curvature information.

As an example, consider the two-dimensional model potential shown in Figure 4a. (The form of this potential, and the details of this calculation, are given in Ref. [29]) In this case, the dividing surface is the vertical line defined by x=0. At any point on the potential surface, there are two principle curvatures, which can be found by diagonalizing the Hessian matrix. Because by definition the lowest principle curvature (the lowest eigenvalue) is negative at any saddle point, a relatively simple definition for the bias potential is one that smoothly turns on as the lowest eigenvalue of the Hessian goes positive, and is zero whenever it is negative. The result of this simple approach is shown in Figure 4b. While the potential near the minimum is altered significantly, the dividing surface is hardly affected. Figure 5 shows the evolution of the boosted time on this biased potential surface at a reduced temperature of 0.1 for a short section of hyperdynamics trajectory. Notice that it evolves very nonlinearly, indicating that the trajectory moves in and out of the regions with high bias potential. At this temperature, the average boost factor is 3435. Figure 6 shows an Arrhenius plot of the diffusion constants computed at a number of temperatures. The average boost values (shown next to each point) are seen to increase as the temperature is lowered. This is a general characteristic of this approach, resulting from the inverse Boltzmann form of the instantaneous boost factor. The hyperdynamics results are seen to agree perfectly with those from direct molecular dynamics. At the lower temperatures, direct MD is unfeasible, but the exact dynamical result can be computed using dynamically corrected TST [37], as discussed above.

For higher-dimensional systems, such a simple form for the bias potential is inadequate, because the Hessian has all positive eigenvalues in a vanishingly small fraction of the total space. Thus, for realistic systems, we use a more complicated function of the gradient and Hessian information, by exploiting an approximate, local definition for the dividing surface position proposed by Sevick, Bell, and Theodorou [38]. In their approximation, the system is at a dividing surface if the lowest Hessian eigenvalue is negative and the gradient projected onto the lowest Hessian eigenvector is zero. This definition is exact at the saddle points of the system, and is a good approximation nearby.

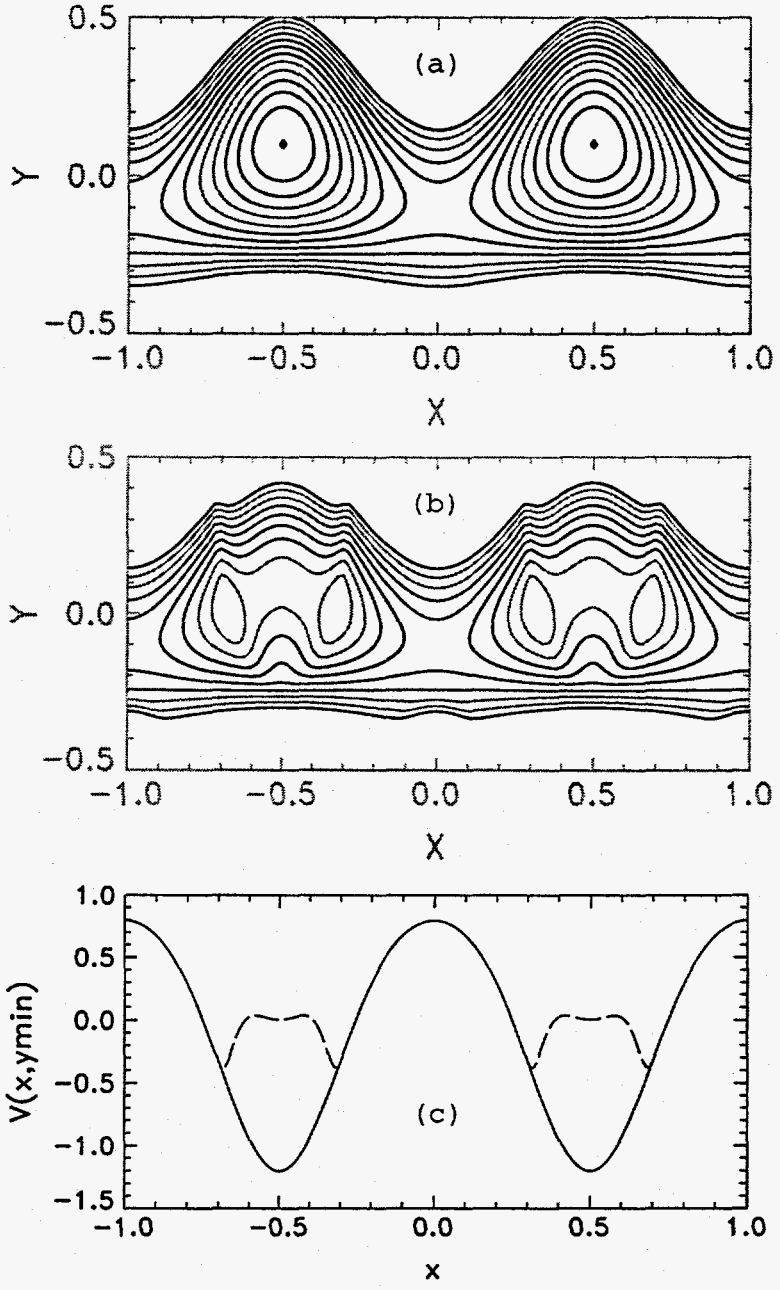

Figure 4. Two-dimensional model potential. The potential is periodic in the x direction and parabolic in the y direction. Two periods are shown. (a) contour plot of original potential. (b) contour plot of the potential modified by a bias potential based on the lowest eigenvalue of the Hessian matrix. Note that the dividing surface (the x=0 line) is slightly corrupted. (c) projection into one dimension: minimum energy along y for each value of x, for both original potential (solid line) and biased potential (dashed line). (Adapted from Ref. [29].)

This expression, and its incorporation into a form that defines a smooth bias potential, is described in Ref. [30].

In this approach, one must choose a value for the strength (h) of the bias potential that is lower than the lowest barrier in the system. In real applications, where the mechanisms and barriers are not known, the lowest barrier can be estimated from the length of time the simulation has run

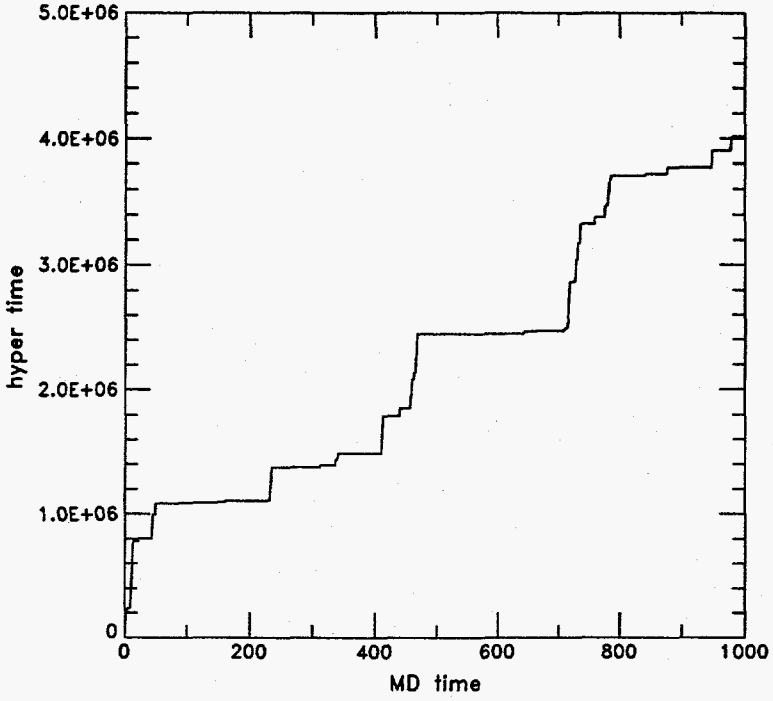

Figure 5. Evolution of the boosted time for a small section of the trajectory at kT=0.1. Both time and temperature are expressed in reduced units. The average boost factor is 3435. (Adapted from Ref. [29].)

without any observed event. For example, assume the rate constant for escape over the lowest barrier has the standard Arrhenius form (as would result from a Vineyard calculation [39]),

$$k = n_{path}\nu exp(-E_a/kT), \tag{3}$$

where E_a is the barrier height, ν is the preexponential factor, and n_{path} is a degeneracy factor accounting for the number of equivalent pathways. The average time between transitions of this type is given by the inverse of k. If a simulation has run for a time t without any transition occurring, an approximate lower bound on the lowest barrier can be calculated from Eq. (3) as

$$E_a > k_B T ln(n_{path}\nu t). \tag{4}$$

If we assume that ν has a typical value of $10^{12}-10^{13}$ sec^{-1}, and take n_{path}=1 (the most conservative choice), the bound on the barrier is completely defined. Setting h smaller than this lower bound on the barrier gives a bias potential that is unlikely to block any transition regions. This procedure can be applied iteratively. At first the simulation is performed with direct MD, until a useful lower bound is obtained. Hyperdynamics simulations then produce successively higher bounds on the lowest barrier, each of which allows the bias potential to be made stronger, causing the hypertime to evolve more rapidly. This "safe" procedure for turning up the bias potential is demonstrated below. In the next two examples, h is simply fixed according to known barriers in the system.

Although the bias potential depends formally on the second derivatives of the potential (and the forces neeeded for MD would thus depend on the third derivatives), an iterative algorithm [30] allows propagation of the hyperdynamics trajectory using only first derivatives, just as for normal molecular dynamics. The necessary quantities are accumulated by making a few dozen evaluations of the interatomic forces at positions very close to the instantaneous position of the trajectory. In

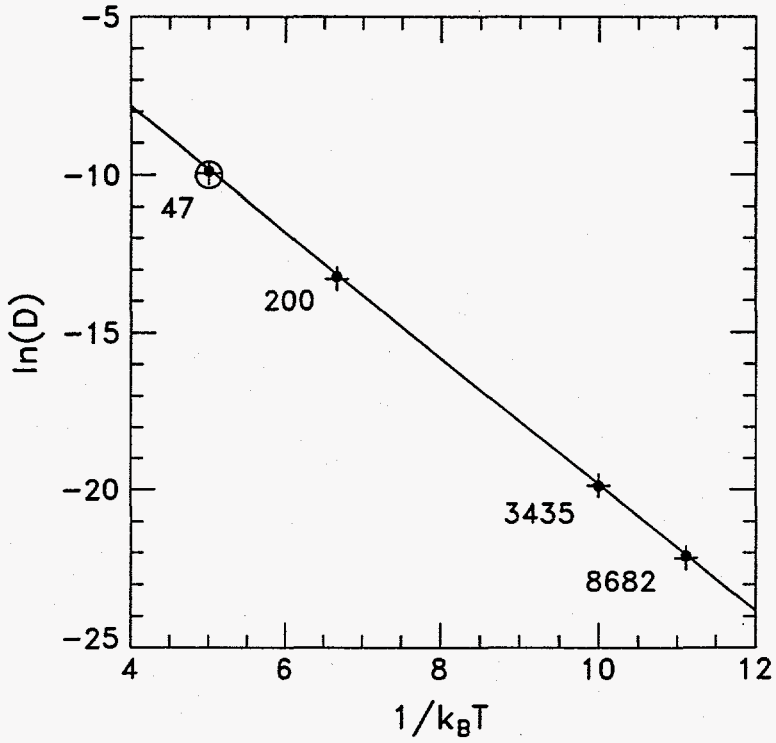

Figure 6. Arrhenius plot of diffusion constants computed for the 2-D potential shown in Figure 4, using direct molecular dynamics (open circle), hyperdynamics (filled circles), transition state theory with dynamical corrections (plus symbols), and harmonic transition state theory (solid line). (Adapted from Ref. [29].)

applications so far using embedded atom method interatomic potentials (a good description for fcc transition metals [40]), this procedure has been found to require about 20 to 50 times as much work per hyperdynamics step as a regular MD step would.

As an example, in a hyperdynamics simulation of Ag adatom diffusion on Ag(100) at T=400K, the bias potential strength h was set to 0.3 eV based on the known hopping and exchange barrier heights, which are slightly above 0.5 eV. The resulting average boost factor was 1356 and the extra work per integration step step (compared to normal MD) was about 30 times. Thus, the same simulation (9.89 microseconds total boosted time) performed with direct MD would have required 1356/30 = 45 times more computational work. During this 9.89 microseconds, 23 hops and 16 exchanges were observed, corresponding to rate constants that agree [30] with the values estimated from full harmonic transition state theory [39] (direct molecular dynamics was unfeasible).

Figure 7 shows an application of hyperdynamics to a more complex system, the diffusive motion of a 10-atom Ag cluster on the Ag(111) surface [30]. Using a bias potential with parameters similar to those employed in the Ag/Ag(100) case, and allowing two substrate layers to move (for a total of 70 moving atoms), a boost factor of 8310 was obtained at a cost of about 20 times the work per integration step. In a few weeks of processor time on a scientific workstation, the system evolved for 221.2 microseconds. The nonlinear evolution of the boosted time is shown in Figure 8. The cluster does not move far in this time, but we can observe a number of complicated, concerted mechanisms that might be omitted from a kinetic Monte Carlo simulation. For example, At 9.5 microseconds, all 10 atoms jump in the same direction to move the cluster from hcp to fcc registry (Figures 8a and b), a mechanism discovered in 1992 [19]. At 94.5 microseconds, the cluster breaks out of its compact configuration, via a mechanism analogous to the dimer-shearing

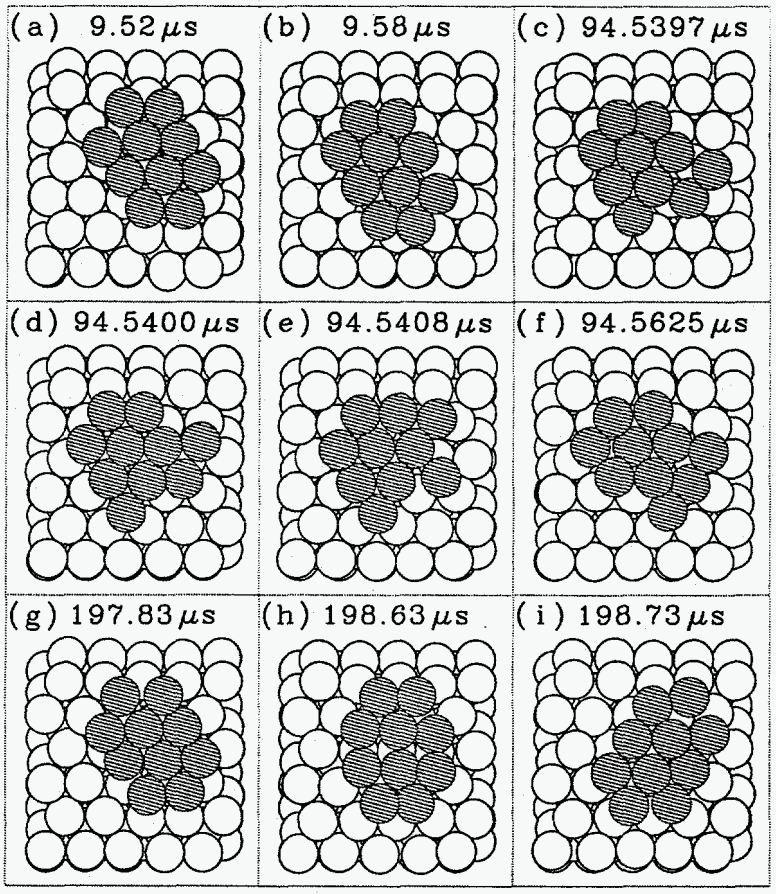

Figure 7. Selected snapshots from hyperdynamics simulation of 10-atom Ag cluster on Ag(111) surface. Three diffusive mechanisms can be observed, as discussed in text. Boosted time values are shown. (Reproduced with permission from Ref. [30].)

on fcc(100) [24], allowing adatom motion at the cluster periphery (Figures 10c-10f). At 197.8 microseconds (Figures 10g-10i), there is a dislocation event [23], resulting in a rotation of the cluster orientation.

This example serves to illustrate the generality of the hyperdynamics approach and the time scales accessible with it. While for this system, all the observed events (and perhaps their relative frequencies) could have been predicted based on extrapolation downward from direct MD simulations at higher temperatures, such an extrapolation becomes impossible for a system as complex as a growing surface. Hyperdynamics should be helpful for understanding these more complicated systems, on time scales approaching those that occur between deposition events. With further development and faster computers, direct simulation of realistic deposition rates may be possible in certain cases, especially when used in conjunction with the new parallel method described below.

Because the average boost factor will be large for systems in which the lowest barrier is high relative to the temperature, hyperdynamics is best employed on systems where a regular molecular dynamics simulation can reach many nanoseconds. Another way to state this is that activated processes in most physical systems have a standard preexponential factor of about one inverse picosecond. If the temperature is raised to infinity, the events will not occur any faster than about once per picosecond. Similarly, if the bias potential is so effective that the well depth is reduced to zero, the rate will be increased to once per picosecond. The greater the difference between the unboosted rate constant and this limiting rate, the greater the possible boost factor.

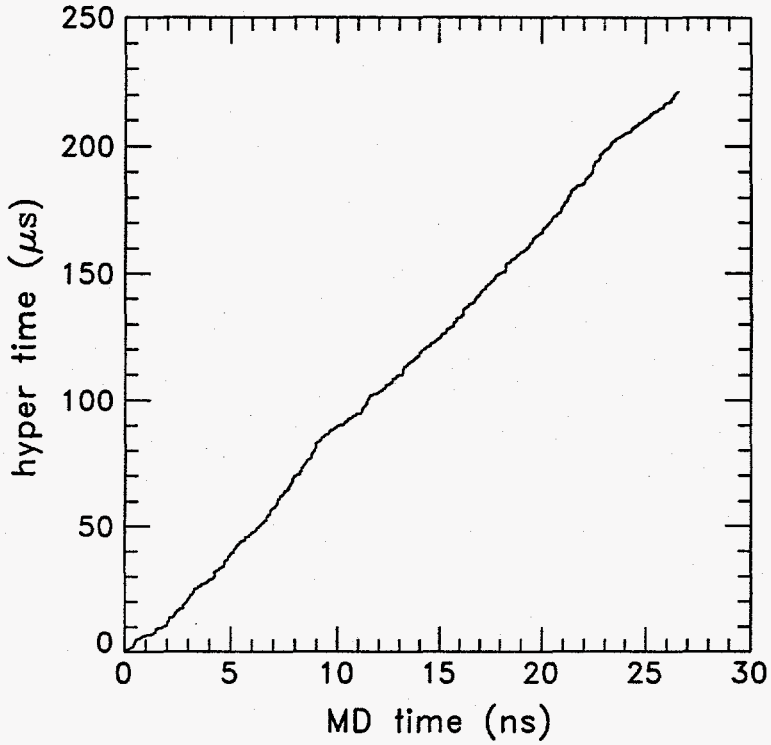

Figure 8. Evolution of the cumulative boosted time during the hyperdynamics simulation of a 10-atom Ag cluster on Ag(111). The average boost factor is 8310.

A consequence of this argument is that hyperdynamics is most effective when applied to small systems with simple potentials on fast computers. Harnessing parallel processing to make the simulation run even faster would be ideal. However, existing parallel molecular dynamics algorithms work by subdividing physical space, assigning the atoms within each region to to a single processor. For systems smaller than a few thousand atoms, the surface-to-volume ratio of each subspace is so high that communication time dominates and the efficiency is severely degraded. For infrequent event systems, there is another way to organize the work, circumventing this problem. We describe this next.

PARALLEL REPLICA DYNAMICS

In the parallel replica method [31], we exploit the fact that escape from a potential basin is a first-order process. Assuming that any correlated events during the transition to the present basin have already finished, the probability distribution for the time (t) until the next escape is given by

$$p(t) = k_{tot} \exp(-k_{tot} t) \,, \tag{5}$$

where k_{tot} is the sum of all rate constants for escape from this basin. We now consider M replicas of the entire system running independently on M equivalent processors. The key point is that these M replicas behave like a new physical system with M times as many escape paths and a total escape rate that is Mk_{tot}. If the M-replica simulation is followed until the first transition occurs on *any* of the processors, a relatively simple manipulation [31] shows that the correct escape-time distribution (Eq. (5)) is recovered, if the time is defined as that accumulated on all the processors up to the instant of the transition. Because this replication does nothing to disrupt the relative probabilities of the available escape paths, a simulation in which this approach is repeated for each new state of the system gives a sequence of states and transition times that are indistinguishable

from those of a long trajectory on a single processor.

The computational procedure can be summarized as follows. The entire system is replicated onto each of the M processors. The replicas are decorrelated by performing a short series of independent momentum randomizations. On each processor, a constant-temperature trajectory for the entire system is propagated, while periodically checking for a transition (discussed below). When a transition to a new basin is detected (on processor i, say), the other processors are stopped, and the master simulation clock is advanced by the sum of the trajectory times accumulated on all the processors up to the time of the transition. The trajectory on processor i is then continued for a short time (a picosecond or two) to allow completion of correlated dynamical events (e.g., an adatom hopping on a surface may make another jump before its excess energy is dissipated). This additional correlation time is added to the simulation clock, and the procedure is repeated by replicating the new geometry of processor i onto all M processors.

With this approach, infrequent-event dynamics proceed nearly M times faster than a single-processor trajectory, provided the average time between transitions is much longer than the time required to evolve the correlated events on one processor. Communication time is typically insignificant, although it could become important in implementations on hundreds of processors.

A requirement of this approach is that transitions can be detected. This is accomplished by periodically performing a steepest-descent (or conjugate gradient) minimization, and comparing the resulting configuration to a reference geometry corresponding to the minimum of the basin. If, after some prechosen number (n_{SD}) of steepest-descent steps (e.g., n_{SD}=100), the position of any atom still differs from the reference geometry by more than a threshold amount (e.g., 1 Å), a transition is declared. This procedure can be performed without slowing the trajectory too much, because usually only a few steps are required to recognize that no transition has occurred.

Although it is critical that escape events be detected when they occur, an attractive feature of this parallel method is that a false positive in the transition detection stage does not invalidate the dynamics. A false positive might occur occasionally if n_{SD} is not large enough to allow the system to relax to a position near the minimum of the basin. In this case, all the processors will be stopped, the trajectory will continue to wander in the same basin during the correlated-event integration stage (since it never left that basin), and the new geometry will then be transmitted to all the processors. The net result will be indistinguishable from a true escape event that recrossed back into the same basin, and the dynamical evolution is still correct.

COMBINED HYPERDYNAMICS AND PARALLEL REPLICA DYNAMICS

As implied above, the trajectory on each processor can be run using hyperdynamics instead of regular MD, resulting in a multiplicative gain in the simulation time per wall clock time. We are currently testing the use of the combined hyperdynamics and parallel replica dynamics methods with the goal of developing a general tool for materials simulations. Test runs look promising. Figure 9 shows the simulation time plotted versus wall clock time for the initial stage of a simulation of bulk vacancy diffusion in silver. This simulation employed 16 Pentium Pro 200 MHz processors. For comparison, the time scale resulting from regular molecular dynamics (on one processor), and the parallel replica method (16 times faster) are also shown. The bias strength in the parallel+hyperdynamics simulation is ramped up slowly (beginning from zero – i.e., regular MD), to prevent accidentally blocking an escape path, using the procedure described above. After each transition (in this case, a vacancy jump) the bias potential was reset to zero. (The first transition occurred long after the period of time shown in Figure 9.) After an induction period of a little less than one hour, the bias potential strength is sufficiently strong that the hyper+parallel time climbs exponentially away from the other curves. In 65 hours, the simulation clock accumulated 1.9 mil-

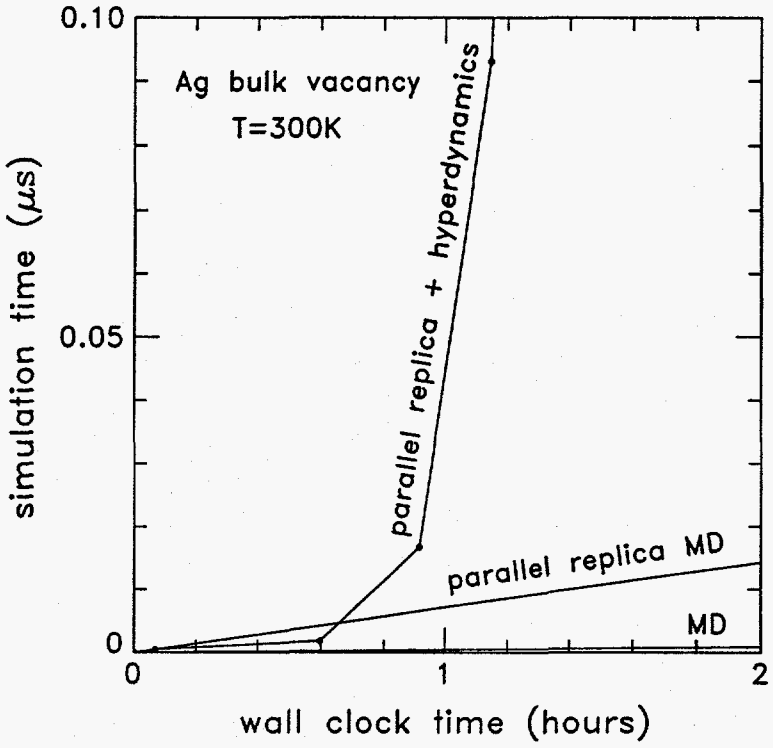

Figure 9. Time evolution during simulation of a bulk silver vacancy, combining the hyperdynamics method with parallel replica method. Only a fraction of the time leading up to the first vacancy hop is shown, as discussed in the text. The simulation achieved 1.9 milliseconds in 65 hours of wall clock time on 16 Pentium Pro processors.

liseconds, and the vacancy made 10 diffusive jumps, corresponding to the correct rate constant, which can be computed directly for this simple case.

As a test of the applicability of this combined method for epitaxial growth processes, we have simulated the rearrangement at T=300K of two monolayers of silver randomly deposited on a small patch of Ag(100) surface. In addition to the two deposited layers (64 atoms), the top three substrate layers were allowed to move (for a total of 160 moving atoms), while two deeper substrate layers were held fixed.

For this simulation, we could not predict how the system would evolve, so the safe biasing procedure was critical. After some initial tests, we decided on a modified version of the procedure in which h was reset to 0.2 eV after each transition, rather than zero.

We ran this simulation as a sequence of 6 runs using various numbers of processors (from 40 to 126) of a 128-node Cray/SGI Origin 2000. The total simulation used 5242 processor-hours of CPU time, and resulted in 796 events over a total boosted simulation time of 182.5 μs.

A sequence of snapshots from this simulation are shown in Figure 10. We note that due to the small (periodic) size of this system, the initial configuration with 2 ML deposited has actually filled in one complete monolayer, so that the system may be simply described as a six-adatom and six-vacancy surface. During the 182.5 μs simulation, three of these adatoms filled in vacancies, in three separate transitions (shown in Figures 11 to 13) occuring at 2.5, 25.5, and 28.1 μs. Following the last of these, the majority of the simulation consists of simple rotations of the 3-adatom surface cluster, with one of the end atoms moving into the free site opposite the apex atom, while achieving no actual migration towards the 3-atom vacancy cluster. This transition has a relatively low barrier (xxxx eV), limiting the maximum boost that is possible in the hyperdynamics. Further work is

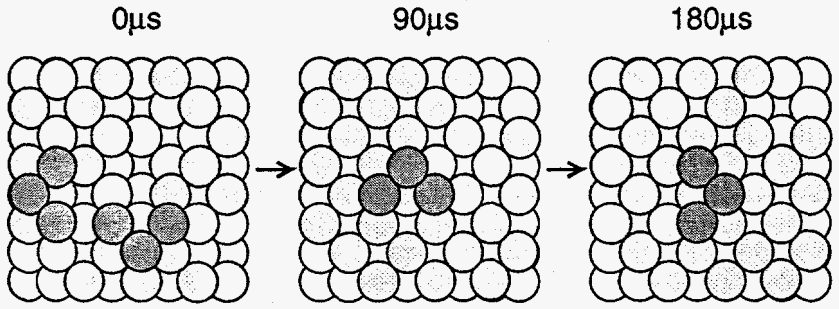

Figure 10. Overview of 2 ML Ag/Ag(100) diffusion simulation carried out using the combined hyperdynamics and parallel replica dynamics methods. Snapshots, from left to right, are taken at 0, 90, and 180 μs, respectively.

needed to find efficient methods for avoiding such cases, where a low-barrier process dominates, effectively hiding more interesting events which may have only slightly higher barriers. Note that while the hyperdynamics boost is reduced by this effect, the full parallel gain is still observed.

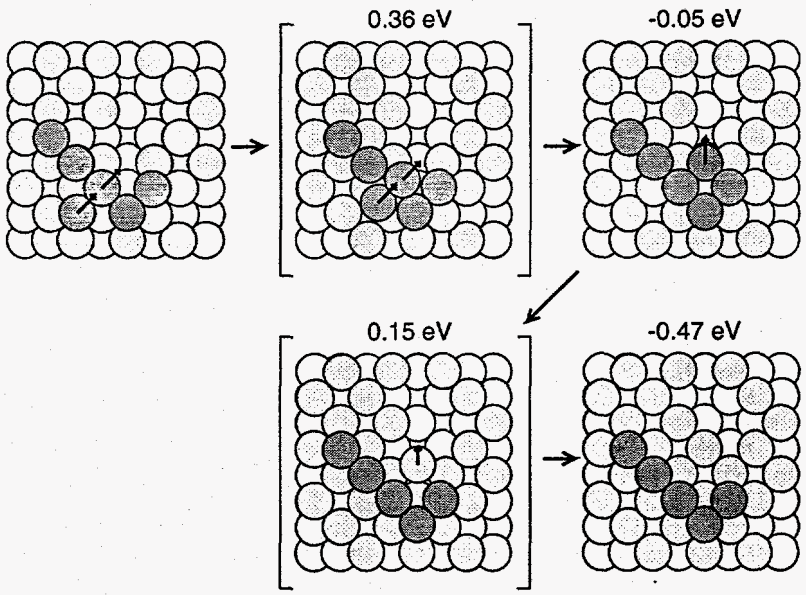

Figure 11. First event pair which results in an adatom filling a surface vacancy, at $t = 2.53\mu$s and $t = 2.60\mu$s. Energies relative to the initial configuration are shown for the subsequent two minima and two transition states.

Still, this preliminary trial is already quite encouraging. The boosted simulation time of 182.5 μs would have taken over 152 thousand hours of CPU time (nearly 18 years) running ordinary MD on a single Origin 2000 processor node, while our simulation was performed in a wall-clock time of 73 hours. This overall computational speedup of 2080 may be approximately partitioned into a factor of 70 from the parallel replicas (the average number of processors being used), and a factor of 30 from the hyperdynamics.

It is interesting to note that all three of the interlayer smoothing events (Figures 11-13) involve concerted motion that might be hard to include in a kinetic Monte Carlo simulation. xxx more

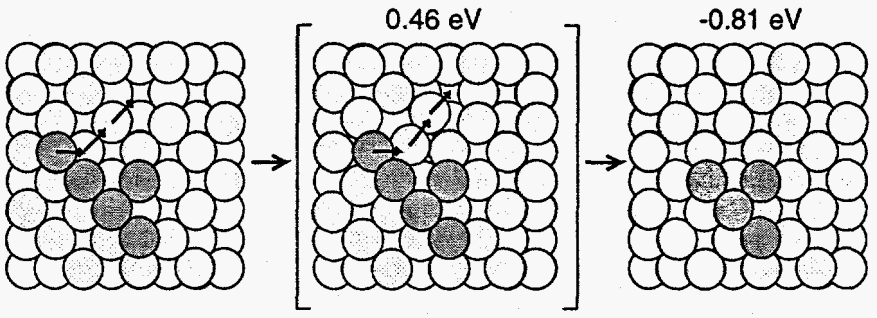

Figure 12. Second event in which an adatom fills a vacancy, involving the concerted motion of three atoms, at $t = 25.46\mu$s. Energies are relative to the initial configuration on the left.

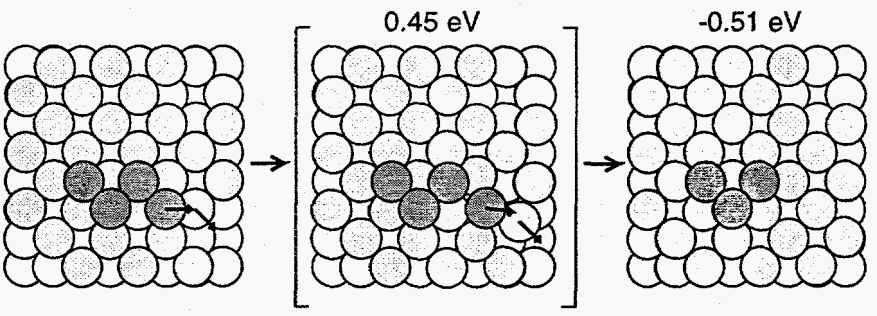

Figure 13. Third event in which an adatom fills a vacancy, involving the concerted motion of two atoms, at $t = 28.13\mu$s. Energies are relative to the initial configuration on the left.

here?

CONCLUSIONS

The hyperdynamics method and the parallel replica method each extend substantially the simulation times accessible to molecular dynamics. Combining these approaches multiplies the boost factors, allowing simulations on the millisecond time scale for small systems. This approach looks promising as a tool for studying the details of epitaxial layer growth. Although a few milliseconds is still far short of the simulation time needed to study surface growth at realistic deposition rates, with further development of the methods and anticipated advances in computer technology, such simulations may be possible in another decade. Already this combined approach should be useful as a companion to kinetic Monte Carlo, assisting in finding missing reaction pathways. Overcoming the "low-barrier" problem, in which the lowest barrier in the system limits the hyperdynamics boost, remains as a challenge. The parallel replica method does not suffer from this effect, unless the lowest barrier is extremely low.

ACKNOWLEDGEMENTS

This work was supported by the United States Department of Energy, Office of Basic Energy Sciences. Computing time on the LOKI parallel computer (16 Pentium Pro processors) in the Theoretical Division at Los Alamos, and assistance in this regard from Mike Warren, are gratefully

acknowledged.

... xxxx acknowledge ASCI machine...

REFERENCES

1. F. Karetta and H.M. Urbassek, J. Appl. Phys. **71**, 5410 (1992).
2. R.M. Lyndenbell, Surf. Sci. **259**,129 (1991).
3. E.T. Chen, R.N. Barnett, and U. Landman, Phys. Rev. B **41**, 439 (1990).
4. T.D. De La Rubia, A. Caro, and M. Spaczer, Phys. Rev. B **47**, 11483 (1993).
5. W. Zhong, Y. Cai, and D. Tomanek, Nature **362**, 435 (1993)
6. A.F. Voter and J.D. Doll, J. Chem. Phys. **80**, 5832 (1984).
7. A.F. Voter and J.D. Doll, J. Chem. Phys. **82**, 80 (1985).
8. A.F. Voter, Phys. Rev. *B* 34, 6819 (1986).
9. P.J. Feibelman, Phys. Rev. Lett **65**, 729 (1990).
10. D.W. Bassett and P.R. Webber, Surf. Sci. **70**, 520 (1978).
11. T. Halichioglu and G.M. Pound, Thin Solid Films **57**, 241 (1979).
12. J.D. Wrigley and G. Ehrlich, Phys. Rev. Lett. **44**, 661 (1980).
13. R.T. Tung and W.R. Graham, Surf. Sci. **97**, 73 (1980).
14. G.L. Kellogg and P.J. Feibelman, Phys. Rev. Lett. **64**, 3143 (1990).
15. C. Chen and T.T. Tsong, Phys. Rev. Lett. **64**, 3147 (1990).
16. C.L. Liu, J.M. Cohen, J.B. Adams, and A.F. Voter, Surf. Sci., **253**, 334 (1991).
17. L. Hansen, P. Stoltze, K.W. Jacobsen, and J.K. Norskov, Phys. Rev. B **44**, 6523 (1991).
18. G.L. Kellogg and A.F. Voter, Phys. Rev. Lett. **67**, 622 (1991).
19. C.L. Liu and J.B. Adams, Surf. Sci. **268**, 73 (1992).
20. R. Wang and K.A. Fichthorn, Molec. Sim. **11**, 105 (1993).
21. M. Villarba and H. Jónsson, Phys. Rev. B **49**, 2208 (1994).
22. S. Liu, Z. Zhang, J. Norskov, and H. Metiu, Surf. Sci. **321**, 161 (1994).
23. J.C. Hamilton, M.S. Daw, and S.M. Foiles, Phys. Rev. Lett. **74**, 2760 (1995).
24. Z-P. Shi, Z. Zhang, A.K. Swan, and J.F. Wendelken, Phys. Rev. Lett. **76**, 4927 (1996).
25. J.M. Cohen, Surf. Sci. Lett. **306**, L545 (1994).
26. G. Ehrlich and F.G. Hudda, J. Chem. Phys. **44**, 1039 (1966); R.L. Schwoebel J. Appl. Phys. **40**, 614 (1969).
27. P. Stoltze and J.K. Norskov, Phys. Rev. B **48**, 5607 (1993).
28. W.F. Egelhoff, Jr. and I. Jacob, Phys. Rev. Lett. **62**, 921 (1989).
29. A.F. Voter, J. Chem. Phys. **106**, 4665 (1997).
30. A.F. Voter, Phys. Rev. Lett. **78**, 3908 (1997).
31. A.F. Voter, Phys. Rev. B, in press, 1998.
32. R. Marcelin, Ann. Physique **3**, 120 (1915).
33. H. Eyring, J. Chem. Phys. **3**, 107 (1935).
34. J.B. Anderson, Adv. Chem. Phys. **91**, 381 (1995).
35. C.H. Bennett, in *Algorithms for Chemical Computation*, edited by R.E. Christofferson (American Chemical Society, Washington, DC, 1977), p. 63.
36. D. Chandler, J. Chem. Phys. **68**, 2959 (1978); J.A. Montgomery, Jr., D. Chandler, and B.J. Berne, J. Chem. Phys. **70**, 4056 (1979).
37. A.F. Voter, J.D. Doll, and J.M. Cohen, J. Chem. Phys. **90**, 2045 (1989).
38. E.M. Sevick, A.T. Bell, and D.N. Theodorou, J. Chem. Phys. **98**, 3196 (1993).
39. G.H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).

40. M.S. Daw, S.M. Foiles, and M.I. Baskes, Mater. Sci. Reports **9**, 251 (1993).