Adaptive temperature-accelerated dynamics

Yunsic Shim and Jacques G. Amar
Department of Physics & Astronomy
University of Toledo, Toledo, Ohio 43606, USA
(Dated: September 8, 2010)

We present three adaptive methods for optimizing the high-temperature $T_{\text{high}}$ on-the-fly in temperature-accelerated dynamics (TAD) simulations. In all three methods, the high-temperature is adjusted periodically in order to maximize the performance. While in the first two methods the adjustment depends on the number of observed events, the third method depends on the minimum activation barrier observed so far and requires an \textit{a priori} knowledge of the optimal high-temperature $T_{\text{high}}^\text{opt}(E_a)$ as a function of the activation barrier $E_a$ for each accepted event. In order to determine the functional form of $T_{\text{high}}^\text{opt}(E_a)$, we have carried out extensive simulations of submonolayer annealing on the (100) surface for a variety of metals (Ag, Cu, Ni, Pd and Au). While the results for all 5 metals are different, when they are scaled with the melting temperature $T_m$, we find that they all lie on a single scaling curve. Similar results have also been obtained for (111) surfaces although in this case the scaling function is slightly different. In order to test the performance of all 3 methods, we have also carried out adaptive TAD simulations of Ag/Ag(100) annealing and growth at $T = 80$ K and compared with fixed high-temperature TAD simulations for different values of $T_{\text{high}}$. We find that the performance of all three adaptive methods is typically as good as or better than that obtained in fixed high-temperature TAD simulations carried out using the effective optimal fixed high-temperature. In addition, we find that the final high-temperatures obtained in our adaptive TAD simulations are very close to our results for $T_{\text{high}}^\text{opt}(E_a)$. The applicability of the adaptive methods to a variety of TAD simulations is also briefly discussed.

I. INTRODUCTION

Understanding non-equilibrium processes on the atomic scale is challenging due to the complexity involved in various processes. One of the widely used theoretical methods to study such non-equilibrium systems is molecular dynamics (MD). However, one serious limitation of the standard MD method is that the range of time-scales that can be accessed is typically limited due to the requirement of a small time-step for integrating Newton’s equations of motion. As a result, a variety of accelerated dynamics algorithms\textsuperscript{1–12} have been proposed to overcome the time-scale limitation of MD (for recent reviews see Refs. 11,12).

One recently developed accelerated dynamics method is the temperature-accelerated molecular dynamics (TAD) method.\textsuperscript{7,8} The basic idea behind TAD is to carry out a basin-constrained MD simulation at a temperature ($T_{\text{high}}$) which is higher than the desired simulation temperature ($T_{\text{low}}$) in order to accelerate transitions from the initial state. Once sufficient information has been obtained from the high-temperature MD simulation to determine the first transition to occur at the desired low temperature, this transition is accepted, and a high-temperature basin-constrained MD simulation is then carried out with the new state. Thus, the correct dynamics at the desired temperature $T_{\text{low}}$ is recovered while retaining the computational speed-up due to the high-temperature MD simulations.

The TAD method has recently been applied to the study of low-temperature homoepitaxial\textsuperscript{13} and heteroepitaxial\textsuperscript{14} growth at experimental time-scales. It has also been used to study defect diffusion in bulk MgO\textsuperscript{15} and the dissolution of boron interstitial clusters in Si.\textsuperscript{16} More recently, it has also been used to study vacancy formation and strain in low-temperature Cu/Cu(100) growth\textsuperscript{17} and also to investigate defect-grain boundary interaction mechanisms in copper.\textsuperscript{18} In addition, a method of spatially parallelizing the TAD method using synchronous sublattice algorithms\textsuperscript{19} has been proposed in order to extend length scales.\textsuperscript{20} Using the parallel TAD method we have been able to identify many interesting Ag interlayer diffusion processes\textsuperscript{21} which turn out to be crucial for explaining the complex surface-roughening behavior observed experimentally in low-temperature Ag/Ag(100) growth.\textsuperscript{22}

One of the key parameters which controls the speed-up in TAD is the value of the high temperature $T_{\text{high}}$. While one expects that $T_{\text{high}}$ should be significantly higher than $T_{\text{low}}$ for a significant speed-up to be achieved, if the selected value of $T_{\text{high}}$ is too large, then the speed-up will be reduced since many (irrelevant) high-barrier transitions will be observed before the first low-temperature transition is determined. Conversely, if $T_{\text{high}}$ is too low, then there will be few high-temperature transitions but little speed-up will be achieved. Accordingly, one expects that for each configuration there will be an optimal high temperature $T_{\text{high}}^\text{opt}$ which leads to the best performance.

One way to determine an effective optimal value of $T_{\text{high}}$ for a given simulation is to carry out a set of preliminary simulations over a range of different high temperatures, as was recently done in the case of parallel TAD simulations of Cu/Cu(100) growth.\textsuperscript{20} However, obtaining such an optimal fixed temperature can be quite cumbersome and time-consuming. In addition, such a method does not take into account the fact that the optimal tem-
temperature may depend on the system configuration. As a result the effective optimal high-temperature may even vary during the course of a simulation. Therefore, it is desirable to find a general method to adaptively adjust the high-temperature \( T_{\text{high}} \) during the course of a TAD simulation without carrying out additional preliminary simulations to obtain \( T_{\text{high}}^{\text{opt}} \). In this connection, we note that in their original TAD paper, Sørensen and Voter briefly discussed a simple method in which the high temperature was adjusted by comparing the value \( E_{\text{min}} \) of the lowest energy barrier observed so far with \( 4k_B T_{\text{high}} \). However, detailed information about the performance of this method was not presented. In addition, we note that this method may be effective only when the desired high temperature is proportional to the energy barrier for an accepted diffusion process.

Here we first present two general but slightly different methods by which the high temperature \( T_{\text{high}} \) may be adjusted on-the-fly in order to maximize the performance of TAD simulations. A method for “converting” the elapsed high-temperature time in the course of a TAD simulation from one high-temperature \( T_{\text{high},1} \) to a different high-temperature \( T_{\text{high},2} \) is also discussed. We have tested both methods in simulations of submonolayer Ag/Ag(100) annealing and Ag/Ag(100) growth at 80 K. In both cases, we find that both methods automatically lead to a performance which is close to or as good as that obtained using the optimal fixed high-temperature for that case. We also compare our results with those obtained using the adaptive method proposed by Sørensen and Voter.

In order to gain a better understanding of the dependence of the optimal high-temperature on configuration, we also have determined the optimal high-temperature at \( T_{\text{low}} = 80 \) K for ten different configurations of submonolayer islands on Ag/Ag(100) for which the activation barriers (for the accepted event) span the range from 0.1 eV to 0.36 eV. For comparison, the corresponding results have also been obtained for a variety of other metal(100) systems including Cu/Cu(100), Ni/Ni(100), Pd/Pd(100), and Au/Au(100). While the dependence of the optimal temperature \( T_{\text{high}}^{\text{opt}} \) on activation energy \( E_a \) is different for all five metals, we find that the scaled optimal high-temperature \( T_{\text{high}}^{\text{opt}} / T_m \) (where \( T_m \) is the melting temperature for the corresponding metal) collapses onto a single “universal” curve when plotted as a function of the scaled activation energy \( E_a / k_B T_m \). The resulting scaling function can also be accurately represented using a simple analytical expression. By comparing with additional simulations at \( T_{\text{low}} = 40 \) K and 160 K, we have also found that the dependence of the optimal high-temperature on \( T_{\text{low}} \) is relatively weak.

We have also carried out similar calculations for typical submonolayer metal(111) configurations including Ag/Ag(111), Cu/Cu(111), Ni/Ni(111), and Pd/Pd(111). In this case, we again find reasonably good scaling for the dependence of \( T_{\text{high}}^{\text{opt}} / T_m \) as a function of \( E_a / k_B T_m \), although the corresponding scaling function is slightly different from that for the (100) surface. Based on these results, we present a third but less general method for optimizing the high-temperature in adaptive TAD simulations. As a test, we have also applied this method to simulations of submonolayer Ag/Ag(100) annealing and growth at 80 K. We find that the performance is generally as good as or slightly better than the more general methods.

The organization of this paper is as follows. In Sec. II, we first briefly review the original TAD method and then present the three adaptive methods discussed here. In Sec. III, we then describe our TAD simulations in more detail and also present results for the scaling of the optimal temperature as a function of the energy barriers for diffusion events for five different fcc metals - Ag, Cu, Ni, Pd, and Au. We then present our results for the performance of all three adaptive methods in TAD simulations of Ag/Ag(100) annealing and growth. Finally, in Sec IV we summarize our results and also discuss some additional possible methods for carrying out adaptive TAD.

II. METHODS

A. Review of TAD method

The TAD method is based on the harmonic approximation to transition state theory, which states that the rate \( k_i \) of an activated event \( i \) with energy barrier \( E_i \) may be described using the Arrhenius law,

\[
k_i = \nu_i \exp\left[ -\frac{E_i}{k_B T} \right],
\]

where \( \nu_i \) is the prefactor, \( k_B \) is Boltzmann’s constant and \( T \) is the temperature. A TAD simulation then involves performing a basin-constrained MD simulation at a high-temperature \( T_{\text{high}} \) which is higher than the temperature \( T_{\text{low}} \) at which one wishes to study the time evolution of the system. At some point in time \( t_i \), during the high-temperature MD simulation, the system makes a transition (with activation energy \( E_i \)) from the initial basin state to another basin state. The corresponding activation energy and high-temperature time can then be used to determine the time \( t_i \) that this transition would have occurred at low temperature using the Arrhenius expression,

\[
t_i = t_i \exp\left[ \frac{E_i}{k_B (\beta_{\text{low}} - \beta_{\text{high}})} \right],
\]

where \( \beta = 1/k_B T \). Once such a transition has been observed, the system is then returned to the original basin and the high-temperature MD simulation is continued. Figure 1 shows an example in which two attempted events with energy barriers \( E_1 \) and \( E_2 \) occurred at high temperatures \( t_{1,\text{high}} \) and \( t_{2,\text{high}} \), corresponding to low temperature times \( t_{1,\text{low}} \) and \( t_{2,\text{low}} \), respectively.

This process continues until one is confident, with confidence level \( 1 - \delta \) (where \( \delta \ll 1 \)), that no further attempted transitions will be found with extrapolated low-
temperature event times shorter than the earliest extrapolated low-temperature event-time \( t_{\text{low,short}} \) found so far. In particular, if one assumes that there is a lower bound \( \langle \nu_{\text{min}} \rangle \) on the frequency prefactors \( \{\nu_i\} \) for all possible transitions, then the high-temperature MD run can be terminated once the high-temperature simulation time \( t_{\text{high}} \) reaches the value,\(^7\)\(^8\)

\[
t_{\text{high, stop}} = (1/\nu_{\text{min}}^*) (\nu_{\text{min}}^{*} t_{\text{low,short}})^{1/\nu_{\text{min}}^*} , \tag{3}
\]

where \( \nu_{\text{min}}^* = \nu_{\text{min}}^*/\ln(1/\delta) \). In the case shown in Fig. 1, \( t_{\text{low,short}} = t_2 \).\(^2\)\(^5\) Once this occurs, the transition corresponding to time \( t_{\text{low,short}} \) is accepted and the whole process is repeated starting in the new basin state. We note that Eq. 3 is equivalent to the graphical construction shown in Fig. 1, corresponding to the intersection of the straight-line joining the points \( (0, \ln(1/\nu_{\text{min}}^*)) \) and \( (1/T_{\text{low}}, \ln(t_{\text{low,short}})) \) and the vertical axis at \( x = 1/T_{\text{high}} \).

We note that in typical TAD simulations, the high-temperature \( T_{\text{high}} \) is fixed during the simulation. However, in adaptive TAD simulations, it is desirable to change the high-temperature \( T_{\text{high}} \) (e.g. from \( T_{\text{high,1}} \) to \( T_{\text{high,2}} \)) before the first low-temperature event has been determined, in order to improve the performance. In this case the elapsed high-temperature time at the new high-temperature \( T_{\text{high,2}} \) satisfies the expression,

\[
t_{\text{high,2}} = (1/\nu_{\text{min}}) (\nu_{\text{min}}^{*} t_{\text{high,1}})^{1/\nu_{\text{min}}^*} , \tag{4}
\]

which corresponds to the graphical construction shown in Fig. 2.

One of the primary factors determining the efficiency of the TAD method is the high-temperature \( T_{\text{high}} \). For example, if \( T_{\text{high}} \) is too low, the efficiency will be low since a large amount of MD time will be required before any activated events are observed. However, if \( T_{\text{high}} \) is too high, then too many high-barrier attempted events will occur before an event is accepted, thus also decreasing the efficiency. As a result, we expect that for a given configuration and energy landscape there exists an optimal high-temperature \( T_{\text{high}}^{\text{opt}} \) which results in the best performance.

Due to the stochastic nature of attempted events, as well as the unknown energy landscape (which may contain transitions with a wide range of energy barriers and prefactors) developing an efficient methodology for adjusting the high-temperature \( T_{\text{high}} \) on-the-fly during a high temperature MD simulation is a non-trivial optimization problem. Here, we have considered three different adaptive TAD methods. In all three methods a high-temperature TAD simulation is carried out with a fixed \( T_{\text{high}} \) for a fixed period or cycle-time \( \tau = 30 \) ps, before considering whether or not to increase or decrease \( T_{\text{high}} \) at the beginning of the next cycle.\(^2\)\(^4\) The decision to change \( T_{\text{high}} \) in the next cycle is then based on 3 quantities: (a) the number \( n \) of different attempted transitions or pathways observed in the current cycle, (b) the number \( n' \) of different attempted transitions or pathways observed in the current cycle which have not been previously observed in previous cycles, and (c) the total number \( N \) of different attempted transitions or pathways observed since the last accepted transition. Once a stop-time \( t_{\text{high,stop}} \) has been reached and a transition has been accepted, then the high-temperature is reset to an initial value \( T_{\text{high}}^1 \) starting at the beginning of the next cycle.

**B. Adaptive TAD method**

![FIG. 1: Schematic diagram of TAD method. Time moves up the vertical lines shown. The corresponding low-temperature event time for each attempted event observed at a high temperature \( T_{\text{high}} \) is determined by Eq. 2 and drawn with a dotted line. When the high-temperature MD time reaches \( t_{\text{high,stop}} \), then the earliest low-temperature event observed so far (in this example, the 2nd event) is accepted and the simulation clock is advanced by \( t_{\text{low,short}} \).](image1)

![FIG. 2: Diagram showing calculation of elapsed high-temperature time \( t_{T_{\text{high,2}}} \) at temperature \( T_{\text{high,2}} \) (point labeled \( t_2 \)) after a time \( t_{T_{\text{high,1}}} \) has elapsed at temperature \( T_{\text{high,1}} \) (point labeled \( t_1 \)).](image2)
1. Method I

Method I is based on the observation that statistically, a low-barrier event tends to occur first. Thus, by starting a high-temperature MD simulation with $T_{\text{high}}$ equal to $T_{\text{low}}$, one may be able to scan and accept the lowest-barrier event quickly for a given configuration without seeing a large number of attempted higher barrier events. If at the end of a cycle-time $\tau$, at least one attempted event has been observed during that cycle ($n \geq 1$), then the high-temperature $T_{\text{high}}$ is not changed. However, if no attempted events have been observed during the current cycle ($n = 0$) then the high-temperature is increased according to the following rules:

(a) If the total number $N$ of different transitions observed so far during all cycles is equal to zero ($N = 0$), then $T_{\text{high}}^{\text{new}} = \alpha T_{\text{high}}^{\text{old}}$ with $\alpha = 1.5$.

(b) If $N \neq 0$ then $T_{\text{high}}^{\text{new}} = T_{\text{high}}^{\text{old}}(1 + \beta/N)$ with $\beta = 0.33$.

2. Method II

In Method II one starts with an intermediate high-temperature, such as $T_{\text{high}} = 500$ K. The main idea behind this method is to quickly observe attempted transitions and then either increase or decrease $T_{\text{high}}$ so that $T_{\text{high}}$ converges to a desirable high-temperature. If no attempted events were observed during the current cycle ($n = 0$), then the temperature is increased following the rules given for Method I. However, if attempted events were observed but the number of new attempted events $n'$ is equal to 0, or the total number $N$ of different events observed so far over all cycles is less than 3, then the temperature is not changed. Finally, if $n' \neq 0$ and $N \geq 3$, then the temperature is decreased following the expression,

$$T_{\text{high}}^{\text{new}} = T_{\text{high}}^{\text{old}}(1 - \gamma n/N) \quad (5)$$

where $\gamma = 0.15$. Otherwise the high-temperature is not changed. Note that in Eq. 5, $n$ ($N$) correspond to the number of different transitions observed in the previous cycle (over all previous cycles).

3. Method III

While Methods I and II do not require any prior information about the system, method III is based on the idea that for each system configuration there exists an optimal high-temperature $T^\text{opt}_{\text{high}}(E_a)$ which is a function of the energy barrier $E_a$ of the corresponding accepted transition. The main idea of this method is to observe an attempted event quickly and then adjust the high-temperature according to $T_{\text{high}} = T_{\text{high}}^\text{opt}(E_a^{\text{min}})$, where $E_a^{\text{min}}$ is the lowest energy barrier observed so far and $T_{\text{high}}^\text{opt}(E_a^{\text{min}})$ is the optimal temperature corresponding to $E_a^{\text{min}}$. The detailed functional form of $T_{\text{high}}^\text{opt}$ for homoeptaxial cases will be presented in Sec. III A. We note that while determining the form of $T_{\text{high}}^\text{opt}$ is quite cumbersome, once it is known this method is the most efficient algorithm for adjusting $T_{\text{high}}$. In Method III we again start with an intermediate temperature, such as $T_{\text{high}} = 500$ K as in Method II, and then change $T_{\text{high}}$ according to the following relatively simple rules:

1. If at the end of the current cycle no attempted events have yet been observed ($N = 0$), then $T_{\text{high}}^{\text{new}} = \alpha T_{\text{high}}^{\text{old}}$ with $\alpha = 1.5$.

2. Otherwise the high-temperature is set to $T_{\text{high}}^{\text{new}} = T_{\text{high}}^\text{opt}(E_a^{\text{min}})$ where $E_a^{\text{min}}$ is the lowest barrier so far observed.

We note that in methods II and III, a small change in the high-temperature has a relatively weak effect on the performance. Thus, in order to avoid the computational overhead associated with frequent small changes of temperature we do not change the high temperature if $T_{\text{high}}^\text{new} < T_{\text{high}}^\text{old}$ and $T_{\text{high}}^\text{old} - T_{\text{high}}^\text{new} < 10$ K. We also note that although the optimal values of the parameters $\alpha, \beta, \gamma$ may depend on system configuration, we find that the values given above, e.g. $\alpha = 1.5, \beta = 1.33, \gamma = 0.15$ yield good performance in realistic cases in which the range of energy barriers for surface diffusion is very wide.

III. Results

In order to study the performance of adaptive TAD, we have carried out simulations of both Ag/Ag(100) sub-monolayer relaxation as well as homoepitaxial growth of 1 ML of Ag/Ag(100). In our simulations, a Voter-Chen (VC) embedded-atom-method (EAM) potential was used, while the minimum prefactor was set to $\nu_{\text{min}} = 5 \times 10^{11}$ s$^{-1}$ and the uncertainty $\delta = 0.05$ was used. Our substrate was composed of three moving layers and three fixed (bottom) layers, with lateral size equal to 5$a$ where $a$ is the lattice constant. Periodic boundary conditions were applied in the lateral (x and y) directions. Our high-temperature MD simulations were carried out using a Langevin thermostat with a friction coefficient of $1.0 \times 10^{12}$ s$^{-1}$ and a time step of 4 fs. The energy barriers for attempted events were calculated using the climbing image nudged elastic band (NEB) method and the Vineyard prefactors were also measured. In most of our simulations a low-temperature $T_{\text{low}} = 80$ K was assumed, although we have also carried out some simulations with $T_{\text{low}} = 40$ K and $T_{\text{low}} = 160$ K.

In our simulations of Ag/Ag(100) homoepitaxial growth, atoms were deposited randomly on the surface with a deposition rate of $F = 1$ ML/s while an initial kinetic energy of 0.2 eV (corresponding to $2k_BT_m$ where $T_m = 1235$ K is the melting temperature of Ag) was used for the depositing atoms. The depositing atoms were initially placed at a height above the substrate equal to the
A. Dependence of the optimal high-temperature on activation barrier

Before presenting our simulation results for all three adaptive TAD methods, we first present results for the optimal fixed high-temperature $T_{\text{high}}$ as a function of the activation energy $E_a$ of the accepted event. These results were obtained by carrying out extensive TAD simulations with $T_{\text{low}} = 80$ K for each of 10 different representative submonolayer configurations (see Fig. 4). (We note that these configurations correspond to a sequence of states which occur at low-temperature.) For each configuration TAD simulations with fixed $T_{\text{high}}$ were carried out until the first event was accepted. The resulting simulation times were then averaged over 10 runs to obtain the total simulation time required until a transition is accepted. By carrying out simulations for different values of $T_{\text{high}}$ the optimal value of $T_{\text{high}}$ for each configuration was determined.

Figure 3 shows typical results (for the configurations (1) and (7) shown in Fig. 4) for the simulation time as a function of $T_{\text{high}}$ along with the estimated optimal value of $T_{\text{high}}$ (see arrow) in each case. We note that these two examples span a relatively wide range of energy-barriers for the accepted transition, e.g. edge-zipping with $E_a = 0.097$ eV and dimer edge diffusion with $E_a = 0.359$ eV. As can be seen in Fig. 3, the optimal temperature for the high-barrier transition shown in Fig. 4.7 (715 K) is significantly higher than for the low-barrier transition shown in Fig. 4.1 (313 K). Fig. 3 also indicates that the increase in the performance at $T_{\text{high}}$ compared to that for the ‘worst’ value of $T_{\text{high}}$ studied is significant for the ranges of $T_{\text{high}}$ tested (28% and 63% for (a) and (b), respectively). We have also determined the optimal temperatures for the same edge-zipping transition shown in Fig. 3(a) for both lower and higher values of $T_{\text{low}}$. Our results indicate that the dependence of the optimal temperature on $T_{\text{low}}$ is relatively weak, e.g. we find $T_{\text{high}}^{\text{opt}} = 325$ K (300 K) for $T_{\text{low}} = 40$ K (160 K).

Fig. 5 (a) shows a summary of our results for $T_{\text{high}}^{\text{opt}}$ as a function of $E_a$ based on the 10 Ag/Ag(100) configurations shown in Fig. 4. Also shown are results obtained using the same set of configurations for Cu/Cu(100), Ni/Ni(100), Pd/Pd(100), and Au/Au(100). As can be seen, in all cases the optimal high-temperature increases with increasing activation energy $E_a$. We note that for small $E_a$ the behavior is approximately linear, e.g. $T_{\text{high}}^{\text{opt}} \sim E_a$. This may be explained both qualitatively and semi-quantitatively by the Arrhenius behavior shown in Eq. 1. In particular, if one assumes that for low $E_a$ only one attempted event (corresponding to the accepted event) is observed at the optimal temperature during a cycle-time $\tau$, then one may write $\tau = \frac{1}{\nu} e^{E_a/k_B T_{\text{high}}^{\text{opt}}}$. Assuming a typical prefactor $\nu = 10^{12} \text{sec}^{-1}$ and a cycle-time $\tau = 30$ ps, as is used in our simulations, one obtains $T_{\text{high}}^{\text{opt}} \approx (3400K/eV)E_a$ which is in relatively good agreement with the results for Ag/Ag(100) shown in Fig. 5.

highest height on the surface plus the EAM cutoff distance $d_{\text{cut}}$ (where $d_{\text{cut}} = 5.542$ Å for Ag). Most of our annealing simulations were carried out using a 2 GHz Intel dual-core Mac mini. However, our growth simulations were carried out using the 2.6 GHz IBM 1350 Glenn cluster at the Ohio Supercomputing Center (OSC). As already noted, the adaptive parameters used were $\alpha = 1.5$, $\beta = 0.33$ and $\gamma = 0.15$ along with a cycle-time $\tau = 30$ ps. In all three adaptive methods, $T_{\text{high}}$ was not allowed to be lower than $T_{\text{low}}$ and was also not allowed to be higher than the maximum high-temperature $T_{\text{high}}^{\text{max}} = 750$ K.
We note however, that as \( E_a \) increases the increasing
cost of observing additional higher-barrier attempted
events which will not be accepted causes \( T_{\text{high}}^{\text{opt}} \) to deviate
from this linear behavior and to saturate for high \( E_a \).

As shown in Fig. 5 (b), we also find that if both the op-
timal high-temperature \( T_{\text{high}}^{\text{opt}} \) and the activation energy
\( E_a / k_B \) are scaled by the experimental melting temperature \( T_m \) then the results for the (100) surfaces of all
five metals Ag, Cu, Pd, Ni, and Au collapse into a single
curve. Here we have used the experimental melting
temperatures \( T_m = 1235 \) K, 1356 K, 1728 K, 1828 K,
and 1337 K for Ag, Cu, Ni, Pd, and Au, respectively. A
simple fit to this curve leads to the form,

\[
T_{\text{high}}^{\text{opt}}(E_a)/T_m = a_1 [1 - \exp(-a_2 E_a/k_B T_m)]
\] (6)

where \( a_1 = a_2 = 0.64 \) for (100) surfaces. We note that
we have also tried scaling with other quantities, such as
the cohesive energy \( E_c \) and/or theoretical melting tem-
peratures, but the scaling results are not quite as good
as when scaling with \( T_m \).

We have also carried out similar simulations for
the case of submonolayer annealing on Ag/Ag(111),
Cu/Cu(111), Ni/Ni(111), and Pd/Pd (111). In this case
we considered 11 different configurations with energy-
barriers (for the first low-temperature accepted transition)
ranging from 0.001 eV to 0.27 eV for Ag/Ag(111).

We note that these configurations actually correspond
to a sequence of low-temperature states which starts
with an initial configuration consisting of two monomers
along with a nearby five-atom and eight-atom cluster.

As the system evolves this initial configuration goes
through a series of transitions with increasing energy-
barriers including monomer diffusion, edge-zipping and
atom-attraction\(^2\) as well as monomer and dimer edge-
diffusion. In this case, we again find reasonably good
scaling (see Fig. 5 (b)) by scaling both \( T_{\text{high}}^{\text{opt}} \) and \( E_a / k_B \)
by the experimental melting temperature \( T_m \), although
the scaling function deviates somewhat from that for
(100) surfaces. We note that these results can also be fit
reasonably well using Eq. 6 (with \( a_1 = 1 \) and \( a_2 = 0.285 \))
as shown by the dashed line in Fig. 5 (b).

We note that finding the optimal temperatures on
the (111) surface is computationally quite demanding
due to the relatively large fluctuations in the measured
timings for this case, as well as the existence of com-
plex saddle-points, sometimes involving transitions to
hcp sites, which make the NEB calculations more com-
plex and time-consuming. In addition, due to the ex-
istence of low-barriers, multiple transitions may occur
before any transition can be detected. This leads to the
problem of “intermediate” states which need to be re-
solved once a transition has been detected. For config-
urations with small \( E_a \), the probability of having such
states is relatively small because \( T_{\text{high}}^{\text{opt}} \) tends to be low
and the required high-temperature MD simulation time 
\( t_{\text{high}, \text{stop}} \) is small. However, as \( E_a \) increases, the effect
of such intermediate states becomes more important as
can be seen in Fig. 5 (b). In particular, \( T_{\text{high}}^{\text{opt}} \) tends
to be lower to minimize the probability of having such
states and to reduce their effect on timing. As a result,
for intermediate values of the energy-barrier the optimal
scaled temperature for (111) surfaces tends to be some-
what lower than that for (100) surfaces. On the other
hand, for transitions with sufficiently high barriers, the
relative computational cost associated with intermediate
states decreases with increasing \( E_a \) due to the relatively
large high-temperature MD simulation time required to
observed any attempted transitions.

B. Performance of adaptive TAD

We now consider the performance of all three adap-
tive TAD methods in the case of Ag/Ag(100) annealing
and growth. First we define the relative performance
\( P(T_{\text{high}}) \) corresponding to the ratio of the “simulation
speed” using adaptive TAD to that using a fixed high-
temperature $T_{\text{high}}$ as,

$$P(T_{\text{high}}) = \frac{t(T_{\text{high}})}{t_{\text{adaptive}}}$$  \hspace{1cm} (7)

where $t(T_{\text{high}})$ is the execution time at a fixed $T_{\text{high}}$ and $t_{\text{adaptive}}$ is the execution time obtained using adaptive TAD. We note that in our simulations using method III, the expression given in Eq. 6 was used to calculate the optimal temperature $T_{\text{high}}^{\text{opt}}(E_a)$. For comparison, we have also tested the performance of the adaptive method proposed by Sørensen and Voter\(^7\) in which the high temperature is adjusted by comparing the value of the lowest energy barrier ($E_a$) observed for a given configuration with $4k_B T_{\text{high}}$, such that if $E_a < 4k_B T_{\text{high}}$, then $T_{\text{high}} = E_a/(4k_B)$. As in Ref. 7, we used the initial high temperature $T_{\text{high}}^{\text{opt}} = 1100$ K when testing their method.

In our adaptive TAD simulations of Ag/Ag(100) annealing we have started with the submonolayer configuration shown in Fig. 4 (a), which contains a range of energy barriers for surface diffusion. Starting with this configuration, we then carried out both fixed and adaptive TAD simulations at $T_{\text{low}} = 80$ K until the system progressed through the sequence of configurations shown in Fig. 4. We note that the accepted events in these simulations include edge-zipping (transitions 1 and 2), atom attraction (3 and 6), edge diffusion (transitions 4, 5, 8, 9, and 10) and dimer edge diffusion (7). While most of these events are low-barrier events (except for dimer edge diffusion with $E_a = 0.359$ eV) the distribution of energy barriers corresponding to attempted events is typically very broad (0.097 – 1.03 eV at $T_{\text{high}} = 600$ K) and there is also a wide range of prefactors ($2.4 \times 10^{12} \text{ s}^{-1} - 2.5 \times 10^{13} \text{ s}^{-1}$).

To obtain reasonable statistics all of our timing results were obtained from averages over at least 10 runs. Due to the broad range of energy barriers, in a few runs (especially for large $T_{\text{high}}$) there was a slight change in the sequence of transitions. Accordingly, in our timing measurements for annealing we only considered cases in which the diffusion sequence followed exactly the sequence provided in Fig. 4 in order to make a precise comparison. We note that in our simulations the total low-temperature time $t_{\text{low}}$ at $T_{\text{low}} = 80$ K was about $10^{10}$ sec due to the relatively high barrier for dimer edge-diffusion. In contrast, the total high-temperature time $t_{\text{high}}$ was about four nsec, thus resulting in a huge boost of about $10^{10}$.

The performance results for all three adaptive TAD methods are shown in Fig. 6 (a) along with the corresponding results obtained using the method of Sørensen and Voter.\(^7\) As can be seen, the optimal fixed high-temperature ($T_{\text{high}}^{\text{opt}} = 650 – 700$ K) is relatively high, due to the relatively large barrier for dimer edge-diffusion (which took about one third of the total execution time). We note that the performance of methods II and III is actually somewhat better than that obtained using the optimal fixed high-temperature, while for method I it is
approximately the same \((P(T_{\text{opt}}^{\text{high}}) \simeq 0.98)\). On the other hand, Sørensen and Voter’s method yields a significantly lower performance. In addition, we note that the relative performance of all three adaptive methods increases significantly as the high-temperature deviates from the optimal fixed high-temperature. For example, in the case of method III the adaptive TAD simulations are approximately 1.6 (1.35) times faster than the corresponding fixed high-temperature simulations at \(T_{\text{high}} = 550\) K (800 K).

For comparison with our results for \(T_{\text{opt}}^{\text{high}}(E_a)\) in Fig. 4 we have also recorded the final high-temperature \(T_{\text{opt}}^{\text{high}}(E_a)\) reached in our adaptive TAD simulations just before each accepted event. As shown in Fig. 6 (b), the resulting values for \(T_{\text{opt}}^{\text{high}}(E_a)\) obtained using method II are in good agreement with \(T_{\text{opt}}^{\text{high}}(E_a)\), while the values obtained using method I tend to be somewhat higher. This is consistent with the fact that the performance for method II is close to that obtained using method III while the performance using method I is not quite as good. Fig. 6 (b) also indicates that the simple adaptive method proposed by Sørensen and Voter\(^7\) leads to a final high-temperature \(T_{\text{high}}^{\text{opt}}(E_a)\) which is significantly lower than \(T_{\text{opt}}^{\text{high}}(E_a)\). Similar results (not shown) have also been obtained for Cu/Cu(100).

In summary, while method III shows the best performance, in general we find that all 3 adaptive methods yield a performance which is as good (or better) than that obtained using the optimal fixed temperature \(T_{\text{opt}}^{\text{high}}\). It should also be noted that while the results discussed above using methods II and III were all obtained using a fixed starting value of the high-temperature \((T_{\text{high}} = 500 K)\), as a test we have carried out additional simulations with both a lower initial value \((T_{\text{high}} = 360 K)\) as well as a somewhat higher value \((T_{\text{high}} = 600 K)\). In both cases we find that the performance depends only weakly on the initial value of \(T_{\text{high}}\).

We now consider the performance of adaptive TAD in the case of homoepitaxial thin-film growth. We note that in order to optimize the performance in growth simulations, it is useful not only to consider the time \(t_{\text{low,short}}\) until the first low-temperature activated event, but also the time \(t_{\text{dep}}\) until the next deposition event. If \(t_{\text{dep}}\) is very large (as in annealing) then the optimal high-temperature is determined solely by the activation barriers for activated events in the system, as discussed in Sec. III.A. On the other hand, if a deposition event is to occur relatively soon, then increasing the high-temperature (to increase the number of attempted events observed) may slightly decrease the performance, since none of these events are likely to be accepted before the deposition event occurs.

In order to take this into account, we treat deposition in the same manner as an activated event. In particular, assuming an average prefactor \(\bar{\nu}\) and an average time-interval between depositions \(\bar{t}_{\text{dep}}\) (determined by the deposition rate and the system size) we may associate an activation energy \(E_a^{\text{dep}} = k_B T_{\text{low}} \ln(\bar{\nu} \bar{t}_{\text{dep}})\). Using the results for annealing shown in Fig. 5, the corresponding optimal temperature \(T_{\text{opt}}^{\text{dep}}\) may be obtained. We then modify our adaptive TAD simulations in the following way. As long as the earliest low-temperature attempted event time \(t_{\text{low,short}}\) is less than the deposition time \(t_{\text{dep}}\), or if no attempted events have yet been observed, then the high-temperature \(T_{\text{high}}\) is adjusted as already described in Sec. III.B.1 - III.B.3. However, if \(t_{\text{dep}}\) is less than \(t_{\text{low,short}}\) then the high-temperature is set to the value \(T_{\text{opt}}^{\text{dep}}\).

In order to test the performance we have carried out adaptive TAD simulations of Ag/Ag(100) growth. Our growth simulations were carried out at \(T_{\text{low}} = 80 K\) with flux \(F = 1 \text{ML/s}\) and corresponded to 1 ML of deposition. In this case, assuming \(T_{\text{low}} = 80 K\) and \(\bar{\nu} \approx 5 \times 10^{12} \text{sec}^{-1}\) we obtain an associated activation energy \(E_a^{\text{dep}} \approx 0.17 \text{eV}\) which corresponds, using the results in Fig. 5(a), to an optimal temperature \(T_{\text{opt}}^{\text{dep}} \approx 500 K\). As in our previous annealing simulations, our growth simulations were carried out on a square substrate with 50 fcc sites in the top layer and periodic boundary conditions in the lateral direction, while the substrate consisted of 3 moving and 3 fixed layers. We note that in these simulations the typical boost factor was approximately \(10^{16}\), e.g. the total high-temperature MD time to simulate 1 sec of growth at \(T = 80 K\), was approximately 1 ns. In addition, the distribution of energy barriers for attempted events was quite broad \((0.078 - 0.66 \text{eV})\) along with a relatively wide range of prefactors \((1.4 \times 10^{12} \text{ s}^{-1} - 1.1 \times 10^{14} \text{ s}^{-1})\).

Fig. 7 shows the performance results in each case. As can be seen, the performance of all 3 adaptive TAD methods is at least 4–8% better than the best result obtained with the optimal fixed high-temperature. Interestingly, in this case we find that there is very little difference between the performance of the 3 methods, although method III still appears to be somewhat more efficient than the other two methods. We note that in this case the optimal fixed high-temperature \((T_{\text{opt}}^{\text{high}} = 550 K)\) is somewhat lower than for Ag/Ag(100) annealing, while the performance for the case of fixed high-temperature depends more strongly on the value of \(T_{\text{high}}\). For example, our adaptive TAD methods lead to a performance which is approximately 50% better than that obtained using a fixed high-temperature \(T_{\text{high}} = 600 K\) and approximately twice as fast as that obtained with \(T_{\text{high}} = 700 K\). For comparison, we have also carried out adaptive TAD simulations using methods I, II, and III without taking into account the time \(t_{\text{dep}}\) until the next deposition event. In this case, the performance was approximately 10% worse than the results shown in Fig. 7. On the other hand, the method of Sørensen and Voter yields a significantly lower performance.
We have presented three adaptive TAD methods in which the high-temperature $T_{\text{high}}$ is adjusted on-the-fly in order to maximize the performance of TAD simulations. In the first two adaptive methods (I and II) the high-temperature is adjusted based on the number of attempted events observed. In particular, Method I starts at the desired low-temperature $T_{\text{low}}$ and then increases the temperature if an insufficient number of transitions have been observed within a specified cycle-time $\tau$. The motivation behind this method is to start at a low value of $T_{\text{high}}$ to avoid the possibility of intermediate states or a large number of high-barrier attempted transitions and then to increase $T_{\text{high}}$ if necessary. Method II is similar, however it starts at a significantly higher temperature and then either increases or decreases the temperature based on the number of observed attempted events. The third method (method III) is based on the assumption that there exists an optimal temperature which depends primarily on the activation energy of the lowest-barrier transition for a given configuration, and not on other factors such as transition prefactors or the activation barriers for other transitions. However, this last method relies on an a priori knowledge of the dependence of the optimal high-temperature $T_{\text{high}}^{\text{opt}}$ on activation energy $E_a$.

In order to test the performance of these three methods, we have carried out simulations of submonolayer Ag/Ag(100) annealing as well as Ag/Ag(100) growth at $T_{\text{low}} = 80$ K. Despite the additional overhead associated with adaptive TAD, we have found that the performance of all three adaptive methods is typically at least as good as the best timing result obtained using the optimal high-temperature ($T_{\text{high}}$) in the case of fixed high-temperature TAD simulations. In addition, the improvement of adaptive TAD compared to fixed high-temperature TAD increases significantly as $T_{\text{high}}$ deviates from the optimal fixed high temperature. Since in general the optimal high-temperature is not known (and may also change in the course of a long simulation) this suggests that the adaptive TAD methods discussed here should be useful in improving the performance of future simulations.

We have also tested the simple adaptive method proposed by Sørensen and Voter. However, due to the fact that the resulting high-temperature $T_{\text{high}}^{\text{opt}} = E_a^{\text{min}}/6k_B$ is significantly lower than that corresponding to the optimal high-temperature $T_{\text{high}}^{\text{opt}}(E_a)$ this method leads to relatively poor performance. In contrast, we find that both methods I and II automatically converge to a final high-temperature which is close to the optimal high-temperature $T_{\text{high}}^{\text{opt}}(E_a)$. However, since method II allows the high-temperature to either increase or decrease it converges to a temperature which is closer to $T_{\text{high}}^{\text{opt}}(E_a)$ and also leads to a performance which is somewhat better than for method I. We note that, in general, we have found that method III leads to a performance which is slightly better than obtained using the other methods. However, as already noted this method requires an a priori knowledge of the dependence of the optimal high-temperature on the minimum activation-barrier $E_a$.

In order to determine the dependence of the optimal high-temperature $T_{\text{high}}^{\text{opt}}(E_a)$ on activation-barrier $E_a$ we have also carried out extensive simulations of submonolayer annealing at $T_{\text{low}} = 80$ K for a variety of configurations for Ag/Ag(100), Cu/Cu(100), Ni/Ni(100), Pd/Pd(100), and Au/Au(100). We note that if the optimal high-temperature depends only on the activation energy $E_a$ of the accepted transition, and not on the prefactor or the activation energy of other higher-barrier transitions, then one would expect the values of $T_{\text{high}}^{\text{opt}}(E_a)$ for each metal to lie on the same curve. However, our results show that this is clearly not the case. Nevertheless, we find that when both the optimal temperature and activation energy are scaled with the melting temperature $T_m$ for each metal, then the results converge to a single curve which may be fit using the expression given in Eq. 6. We note that reasonably good scaling is also obtained for Ag/Ag(100), Cu/Cu(100), Ni/Ni(100), and Pd/Pd(100) if one replaces $T_m$ by $E_1/k_B$ (where $E_1$ is the barrier for monomer hopping on a (100) terrace). However, such a choice does not lead to good scaling for Au/Au(100). While we do not completely understand the good scaling obtained with $T_m$, these results suggest that it may be possible in the future to find a more general method to determine $T_{\text{high}}^{\text{opt}}(E_a)$.

We have also obtained similar results for the case of submonolayer annealing on Ag/Ag(111), Cu/Cu(111), Ni/Ni(111), and Pd/Pd(111) although for the (111) surface the scaling curve is slightly different. In particular, it is slightly lower than the scaling curve for (100) surfaces in the range $0.5 < E_a/k_BT < 2$. One possible explanation for this is the existence of a variety of low-barrier transitions on (111) surfaces which leads to additional overhead due to intermediate states and thus reduces the optimal temperature. However, further work will be needed to understand this in more detail.

To obtain a better understanding of the dependence of $T_{\text{high}}^{\text{opt}}$ on $E_a$ shown in Fig. 5, we note that the typical transition time $\tau_i$ for an accepted transition with activation barrier $E_i$ and prefactor $\nu_i$ at temperature $T_{\text{high}}^{\text{opt}}$ is given by, $1/\tau_i = \nu_i e^{-E_i/k_BT_{\text{high}}^{\text{opt}}}$. This implies that $T_{\text{high}}^{\text{opt}}$ satisfies

$$T_{\text{high}}^{\text{opt}} = E_i/k_BT_{\text{high}}^{\text{opt}} \ln(\nu_i \tau_i)$$

As already noted, if we assume that the optimal temperature is determined by requiring that the average transition-time $\tau_i$ be a constant which is of the order of the cycle-time $\tau$ and is independent of the activation barrier $E_i$, then the assumption of a constant prefactor $\nu_i = \nu$ leads to a linear dependence of $T_{\text{high}}^{\text{opt}}$ on $E_i$, e.g. $k_BT_{\text{high}}^{\text{opt}} = E_i/\ln(\nu/\tau)$, in disagreement with our simulation results. On the other hand if the prefactor $\nu_i$ increases significantly with $E_i$, then this could lead to the sort of saturation behavior for $T_{\text{high}}^{\text{opt}}(E_a)$ shown in Fig. 5.
typically of the order of the cycle-time $\tau$. It appears that the observed dependence of the optimal high-temperature on the activation barrier $E_i$.

Furthermore, a plot of the average transition time at the optimal high-temperature as a function of activation barrier $E_i$.

In order to reduce the effects of fluctuations, variable-temperature MD simulation may be repeated several times and the corresponding average temperature used as the optimal fixed high-temperature for a TAD simulation. We note that in our tests, we carried out the variable-temperature MD simulation only twice to minimize computational cost. As a result, occasionally the final high-temperature is close to the desired low-temperature $T_{low}$, resulting in poor performance. To avoid this problem, we used a minimum-temperature cutoff of 280 K. In tests of Ag/Ag(100) annealing at $T = 80$ K for the sequence shown in Fig. 4, the resulting performance (0.93) was close to that found using adaptive TAD.

Finally, we briefly discuss the applicability of the adaptive methods discussed here to a variety of TAD simulations. Although method III yields the best results in our test simulations of both annealing and growth, it is limited to those cases in which the dependence of the optimal temperature on the activation barrier of the first accepted event is known. Accordingly, while we would expect it to be useful in simulations of metal homoepitaxial growth and/or annealing, in more complex cases such as heteroepitaxial growth or in other cases where $T_{high}(E_a)$ is not known, the other adaptive methods (methods I and II) are likely to be more suitable.

FIG. 8: (a) Prefactors for accepted events shown in Fig. 4 observed in Ag/Ag(100) annealing simulations at $T_{low} = 80$ K. (b) Corresponding transition times $\tau_i(T_{opt}(E_i))$ at the optimal high-temperature as a function of activation barrier $E_i$.

In our test simulations of both annealing and growth, the cycle-time $\tau$ tends to increase with increasing $E_i$, although the prefactors do tend to increase significantly with increasing $E_i$ over the range $0.2$ eV $< E_i < 0.36$ eV where $T_{high}$ tends to saturate. Furthermore, a plot of the average transition time at the optimal high-temperature $\tau_i = e^{E_i/k_B T_{opt}}/\nu_i$ as a function of the barrier for accepted events (see Fig. 8(b)) indicates that $\tau_i$ increases by more than an order of magnitude as $E_i$ increases from 0.1 eV to 0.36 eV. We note, however, that except for the two smallest barriers $\tau_i$ is typically of the order of the cycle-time $\tau$. Thus, it appears that the observed dependence of $T_{opt}^{high}$ on activation energy cannot be completely explained by the assumption that the expected transition time should be a constant of the order of the cycle-time $\tau$. Instead, a more plausible explanation is that while the optimal temperature depends primarily on the rate of the first accepted event, it also depends on the rates of attempted events. In particular, we expect that with increasing $E_i$ (and $T_{opt}^{high}$) the number of attempted events corresponding to "nearby"-barrier transition pathways will also increase. Thus, while $T_{opt}^{high}(E_i)$ tends to increase with $E_i$, the increase will be somewhat reduced in order to limit the additional overhead due to additional attempted events. As a result, the typical transition-time $\tau_i$ for the first-accepted event at the optimal high-temperature tends to increase with increasing $E_i$.

The increase of $T_{opt}^{high}(E_i)$ with $E_i$ also suggests a possible simple alternative method to optimize the high-temperature in the case of annealing. In this method, rather than continuously adapting the high-temperature, a preliminary simulation is carried out to estimate the optimal fixed high-temperature for each new configuration. This may be done by carrying out a “variable-temperature” MD simulation in which the temperature is periodically increased starting with a temperature equal to $T_{low}$. The MD simulation is stopped when an attempted event is observed, and the final temperature is then chosen as the optimal fixed high-temperature for a regular TAD simulation. Several possible forms for the ramping temperature were tested, and we found that the nonlinear form $T(i) = 80 + 100(i - 1)^{1/2}$ (where each value of $i$ corresponds to a 1 ps block and a check for transitions is performed at the end of each block) yields a reasonably good performance.

Acknowledgments

We would like to acknowledge useful discussions with Arthur F. Voter. This work was supported by NSF grant DMR-0907399 as well as by a grant of computer time from the Ohio Supercomputer Center.
23 We note that the slope of the line joining $\left(0, \ln(1/\nu^\omega_{\min})\right)$ and $\left(1/T_{\text{high}}, \ln(t_{\text{high}})\right)$ gives a lower bound, with confidence $1 - \delta$, on the activation energy for any transitions not observed so far.
24 In our TAD simulations the time-interval or “block-time” between tests for a transition to a different energy basin was 2 ps. Our cycle-time therefore corresponds to 15 block-times.