University of Pennsylvania
ScholarlyCommons

Departmental Papers (CBE)

Department of Chemical & Biomolecular Engineering

11-8-2004

Feature Activated Molecular Dynamics: An Efficient Approach for Atomistic Simulation of Solid-State Aggregation Phenomena

Manish Prasad
University of Pennsylvania

Talid R. Sinno
University of Pennsylvania, talid@seas.upenn.edu

Postprint version. Published in *Journal of Chemical Physics*, Volume 121, Issue 18, 8 November 2004, pages 8699-8710.
Publisher URL: http://dx.doi.org/10.1063/1.1804171

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/cbe_papers/6
For more information, please contact repository@pobox.upenn.edu.

Feature Activated Molecular Dynamics: An Efficient Approach for Atomistic Simulation of Solid-State Aggregation Phenomena

Abstract
A new approach is presented for performing efficient molecular dynamics simulations of solute aggregation in crystalline solids. The method dynamically divides the total simulation space into "active" regions centered about each minority species, in which regular molecular dynamics is performed. The number, size and shape of these regions is updated periodically based on the distribution of solute atoms within the overall simulation cell. The remainder of the system is essentially static except for periodic rescaling of the entire simulation cell in order to balance the pressure between the isolated molecular dynamics regions. The method is shown to be accurate and robust for the Environment-Dependant Interatomic Potential (EDIP) for silicon and an Embedded Atom Method (EAM) potential for copper. Several tests are performed beginning with the diffusion of a single vacancy all the way to large-scale simulations of vacancy clustering. In both material systems, the predicted evolutions agree closely with the results of standard molecular dynamics simulations. Computationally, the method is demonstrated to scale almost linearly with the concentration of solute atoms, but is essentially independent of the total system size. This scaling behavior allows for the full dynamical simulation of aggregation under conditions that are more experimentally realizable than would be possible with standard molecular dynamics.

Keywords
molecular dynamics, simulations, crystalline solids, aggregation, silicon, copper, diffusion, vacancies, vacancy clusters

Comments
Postprint version. Published in *Journal of Chemical Physics*, Volume 121, Issue 18, 8 November 2004, pages 8699-8710.
Publisher URL: http://dx.doi.org/10.1063/1.1804171

This journal article is available at ScholarlyCommons: http://repository.upenn.edu/cbe_papers/6

Feature Activated Molecular Dynamics: An Efficient Approach for Atomistic Simulation of Solid-State Aggregation Phenomena

Manish Prasad and Talid Sinno[1]
Department of Chemical and Biomolecular Engineering
University of Pennsylvania
Philadelphia, PA 19104-6393

ABSTRACT

A new approach is presented for performing efficient molecular dynamics simulations of solute aggregation in crystalline solids. The method dynamically divides the total simulation space into "active" regions centered about each minority species, in which regular molecular dynamics is performed. The number, size and shape of these regions is updated periodically based on the distribution of solute atoms within the overall simulation cell. The remainder of the system is essentially static except for periodic rescaling of the entire simulation cell in order to balance the pressure between the isolated molecular dynamics regions. The method is shown to be accurate and robust for the Environment-Dependant Interatomic Potential (EDIP) for silicon and an Embedded Atom Method (EAM) potential for copper. Several tests are performed beginning with the diffusion of a single vacancy all the way to large-scale simulations of vacancy clustering. In both material systems, the predicted evolutions agree closely with the results of standard molecular dynamics simulations. Computationally, the method is demonstrated to scale almost linearly with the concentration of solute atoms, but is essentially independent of the total system size. This scaling behavior allows for the full dynamical simulation of aggregation under conditions that are more experimentally realizable than would be possible with standard molecular dynamics.

[1] Corresponding Author: talid@seas.upenn.edu

1. INTRODUCTION

Solid-state aggregation is a critical component in a tremendous range of technologically important processes. The nucleation and growth behavior of nanoscopic aggregates of point defects and/or impurities in crystalline metals and semiconductors determine the final properties and suitability of the material for any given application. For example, a major factor in the quality of semiconductor silicon substrates is the size and spatial distribution of nanoscopic voids and oxide precipitates, which arise due to the aggregation of vacancies and dissolved oxygen, respectively[1,2]. The presence of extremely low concentrations of impurities such as nitrogen[3,4] or carbon[5,6] can greatly affect these distributions, and depending on how they are introduced, can either greatly improve the material properties or render it useless. Similar phenomena occur in both pure and alloy metallic systems. Microstructural evolution in metals containing intentional dopants often is the most important factor in setting the mechanical and chemical properties of the resulting material, such as in carbon steels and aluminum alloys[7,8]. Finally, embrittlement of nuclear reactor walls due to self-interstitial aggregation has long been a subject of interest for computational study[9,10,11].

As a result, there is an urgent need for process simulators that are able to quantitatively predict microstructural evolution as a function of processing steps. In order for a direct connection between material properties and experimental observables to be feasibly made, such simulators often rely on a continuum, or at least mesoscopic scale models of the relevant physical and chemical processes. Continuum simulators typically are comprised of a system of Master and/or Fokker-Planck equations[12,13,14], while mesoscale tools can be based on a variety of approaches such as Kinetic Monte Carlo (KMC)[15,16,17,18]. While computationally efficient

relative to atomistic simulation methods such as molecular dynamics (MD), the implementation of such models requires the specification of *microscopic*, or atomic scale, physics and parameters in one form or another in order for them to be predictive. Atomistic scale physical elements that must be provided to continuum rate equation models of aggregation, for example, include explicit representations for the coalescence and fragmentation kernels that describe cluster-cluster interactions for all species. These kernels need to account accurately for complex atomistic features of very small atomic clusters, such as irregular geometries, multi-step diffusion mechanisms, and atomic scale concentration heterogeneities.

We have recently developed a two-scale framework[19,20] for systematically increasing the mechanistic accuracy of continuum models of solid-state aggregation by performing parametrically consistent comparisons between atomistic (empirical potential molecular dynamics) and continuum (Master equation) representations[12] of vacancy aggregation in crystalline silicon. The essential feature of the approach was to compute all properties required for the continuum model with the same interatomic potential used to atomistically simulate the aggregation process. By ensuring complete parametric consistency between the two scales, it was possible to develop a sensitive gauge of the mechanistic accuracy of the continuum rate equation model. A major constraint of this approach was the need to simulate atomistically a sufficiently large number of the aggregating species for a long enough time so as to capture enough of the evolution with reasonable statistical quality. Furthermore, an extremely high initial supersaturation of vacancies ($\sim 10^5$) was required to keep the total size of the system reasonable, whereas typical supersaturations encountered in experimental conditions are no more than 10-100. Similar limitations exist in the simulation of multicomponent aggregation. Recently, we

have also considered the co-aggregation of carbon and silicon interstitials[21], in which both species were present at approximately 0.5% concentration. Given the system size constraints, it was not possible to effectively probe the effects of relative concentration of the two species on the resulting size distribution, and therefore make a more direct connection to experimental observations.

In this paper, a new molecular dynamics-based simulation approach is presented that is ideally suited for accelerating the simulation of aggregation in dilute solid solutions. Electrically neutral point defects, impurities, and their clusters in crystalline semiconductors and metals often lead to relatively localized regions of lattice distortion[22,23]. This property, in conjunction with the fact that these species often exist in dilute concentrations in situations of technological interest, implies that most of the lattice atoms in an MD aggregation simulation are unaffected by the presence of defects and impurities. As a result, regular MD simulation of such systems effectively "wastes" much of the computational work on the time integration of "perfect crystal" atomic thermal fluctuations.

1.1 General Features of the FAMD Method

In the present MD approach, henceforth referred to as Feature-Activated Molecular Dynamics, or FAMD, localized regions are created around each defect/impurity atom or cluster. In these regions (termed *active regions*), regular molecular dynamics is performed at the temperature of interest. Active regions are dynamic in nature i.e. they change in shape, diffuse, merge and fragment based on the dynamics of defects embedded in them. Atoms with coordinates within active region boundaries are called *active atoms*. Crystal matrix atoms outside

these regions, however, are not evolved, but instead serve as essentially static boundary conditions to the active regions. Regions of the simulation system that are not active are referred to as *static regions*. As a result, the total computational effort scales with the number of defect/impurity atoms in the system and not with the overall size of the system, allowing for the consideration of arbitrarily large lattices and low concentrations. A schematic showing a collection of active regions is shown in Figure 1.

The proposed approach is conceptually similar to recent multiscale simulation efforts in which multiple representations are woven together in a single simulation[24]. For example, Abraham and coworkers recently developed a multiscale approach for simulating the dynamics of crack propagation in silicon[25]. A tight-binding potential was used to compute the forces for atoms in the vicinity of the crack tip, where all bond breaking takes place, while the remainder of the atoms near the crack were treated with an empirical potential. Further away from the crack, where atomic displacements are relatively small, the atomistic simulation was linked to a continuum finite element representation of the system. Other examples can be found in refs. [26] and [27].

In a FAMD simulation of an ensemble of mobile species that can change morphology, as well as coalesce and fragment, active regions must be allowed to evolve in several ways to correctly capture the physics in the system. The first is that they must be able to move so as to keep the feature centered within the active region. This is necessary to ensure that there are a sufficient number of thermally active lattice atoms between the defects and the static boundary, to screen the features from the static lattice, and thus minimize the effect of the boundary on the

transport and thermodynamic properties of the feature. Further, two active regions that migrate to close proximity of each another must be able to merge into a single active region in order to allow the features within them to interact. Conversely, an active region that contains a cluster that has fragmented into two or more sub-components, must be able to fragment into two or more separate active regions. Finally, the pressure in the overall system must be controllable for NPT ensemble simulations.

Details of how each of these features is implemented are given next in Section 2. In Section 3, the FAMD method is tested and validated using a series of increasingly complex systems all based on the Environment Dependent Interatomic Potential[28,29] (EDIP) potential for silicon. These are: (a) a single vacancy, (b) a vacancy dimer, and (c) a collection of aggregating vacancies in a large host lattice. The latter is compared to previous results obtained for the same system using regular (parallel) molecular dynamics[19]. Section 4 presents a detailed discussion of the performance and scaling behavior of FAMD using simulations of aggregation in silicon and crystalline copper. The latter system is modeled using an Embedded-Atom Method Potential (EAM)[30] as described in Section 5. Finally, conclusions are presented in Section 6.

2. IMPLEMENTATION DETAILS OF THE FEATURE-ACTIVATED MD METHOD

The overall sequence of operations in the execution of a time step within the FAMD method is similar to that in a regular MD simulation. The primary differences are associated with the creation and bookkeeping of information related to the active regions. A flowsheet showing

the major components of the algorithm is presented in Figure 2. In the ensuing discussion, the vacancy-in-silicon system is used to highlight specific technical details of the algorithm.

2.1 Initialization of the FAMD Simulation

An FAMD simulation begins with the specification/identification of all *feature* atom positions. A feature is broadly defined as any species that perturbs (locally) the evolution of the MD system away from perfect crystal behavior; it can be a vacancy, an interstitial, an impurity atom, or even a lattice atom that is displaced from its perfect lattice position by an external driving force. Once the individual feature positions are specified they are grouped according to a Stillinger connectivity criterion[31]. This grouping is performed to partition the simulation system into isolated *MD regions* with each group occupying a unique MD region. An MD region is an algorithmic construct containing all lattice atoms (surrounding a cluster) needed to evaluate the potential energy and forces during molecular dynamics simulation of the active atoms inside it. Note that the grouping cutoff distance is sufficiently large so that each MD region is completely separated from any other and is an independent entity.

The structure of a cubic MD region containing a single feature atom is shown schematically in Figure 3. An MD region contains one active region for every defect feature in it. The various shells in an active region (shown here surrounding a single feature atom for simplicity) represent different temperatures, which are specified using a *temperature envelope function*. The value of the temperature envelope function for each atom in an active region is determined based on that atom's distance from a feature as shown in Figure 4. The thermal activity, T_A, is simply the normalized temperature (with respect to the simulation value), so that

$0 \le T_A \le 1$ for all atoms. The value of the temperature is equal to the desired simulation value up to a distance R_1 and then gradually decreases to zero in the transition interval from R_1 to R_A. Any atoms that lie beyond R_A from any feature atom are assigned a temperature of zero. The radius R_1 is large enough to accommodate any feature-feature interaction distance and includes a buffer for feature diffusion between updates. Atoms that are located within a radius R_1 of any feature atom therefore are *fully active* and are thermostatted to the desired simulation temperature, while atoms that are located in the region $R_1 < r < R_A$ are *transition atoms*. The temperature at the outer edge of the transition layer is truncated at some finite $T_A^{\min}$ (0.50 in this work) in order to reduce the thickness of the transition layer and therefore increase the efficiency of the approach. The effects of this transition layer on the accuracy and efficiency of the FAMD scheme are discussed later.

In the presence of multiple feature atoms, each transition layer atom has a thermal activity that is determined additively based on its distance from all neighboring feature atoms, up to a maximum value of unity; in the latter case the atom simply becomes a fully active atom. Note that all atoms in both the fully active and transition layers are considered as part of an active region and are simulated with normal molecular dynamics. Finally, initial velocities are assigned (based on their temperature) to all atoms in the active region and the simulation is initiated.

2.2 Time Integration in the FAMD Simulation

Time stepping in a FAMD simulation proceeds in much the same way as in a normal MD simulation[32,33], and in this work is based on the 5th-order Gear Predictor-Corrector method.

Bookkeeping lists such as cell-linked lists and Verlet neighbor lists are updated periodically (every 20 time steps in the simulations described in this paper) but are created only for active atoms and their neighbors. Temperature control is imposed using standard velocity rescaling every 10 time steps but now only the velocities of *fully active* atoms are used to compute the average instantaneous kinetic temperature[32]. In the transition region, if present, temperature control is imposed on each atom individually. First, the thermal activity of each atom is computed based on its position relative to the features in order to determine the target temperature and then the velocity is rescaled based on the single-atom kinetic temperature.

In addition to updating the standard MD bookkeeping arrays at 20-time step intervals, the ensemble of MD regions also needs to be reconstructed periodically in order to track the evolution in the simulation. This reconstruction is the primary source of computational overhead in the FAMD approach. MD region reconstruction is performed every t_R time steps, which is implicitly determined by the hopping rate of the feature atoms and also by the thickness of the active region around each feature atom. The MD regions must be reconstructed often enough to ensure that the feature atoms are sufficiently far from any of the active region boundaries. Larger active regions will require less frequent updates of the MD regions, but each time step will be much costlier compared to a smaller active region. Thus, the optimal region size is dependent on the marginal computational costs for simulating extra atoms in a given active region and the costs associated with keeping MD regions information current by identifying lattice hops by defect features. For example, for vacancy diffusion in silicon, the parameters $100 < t_R < 250$ and R_I~12.6 Å provide a good balance between accuracy and overhead.

In general, the optimal approach for locating feature atoms for MD region reconstruction depends on the nature of the system being simulated. The specific case of vacancy aggregation poses unique challenges and the overhead associated with other types of features is expected to be lower. In the present study, vacancies are located in the following manner. First, the energy of each active region is minimized using a conjugate gradient (CG) method[34]. This is equivalent to quenching the atomic coordinates to (near) their zero temperature positions. The layers of static atoms (already in their on-lattice positions) around each active region are not allowed to relax during the energy minimization but instead provide static boundary conditions, which prevent spurious surface reconstructions and bulk translation. The quenched lattice fragment is then compared to an equivalent portion of reference (i.e. perfect) lattice at the same density in order to locate the features. In an idealized system, where the feature atoms and clusters are entirely on-lattice structures, an excess number of atoms corresponding to the number of features would be located in the reference lattice fragment and the positions of the excess atoms in the perfect fragment would correspond closely to the positions of the features in the active region. However, in reality most features, including the present case of vacancies, lead to substantial off-lattice relaxation making the exact specification of vacancy positions impossible.

In order to address this issue, a broadly applicable scheme for identifying feature positions in a perfect crystalline lattice was developed based on the displacement field created by the defect features. For each atom in the perfect crystal fragment a displacement is computed by finding the nearest atom in the actual (quenched fragment). Atoms that are displaced from their perfect lattice positions by more than a specified cutoff, R_c^d (set to below half the bond length, or $0.45r_{bond}$ here), are identified as feature atoms. This approach guarantees that any (non-thermal)

lattice displacement field created by a defect cluster (vacancy or otherwise) is incorporated fully within the active region and simulated with regular MD. A complete specification of the active region extent around a cluster of features therefore requires values for both R_c^d and $R_1 + R_A$. A sensitivity analysis of the FAMD performance and accuracy with respect to these parameters is presented later in Section 4.

Once the new feature positions are identified at each reconstruction interval, the features are re-grouped into clusters, which in turn are assembled into new MD regions. New bookkeeping arrays are then created and thermal activities for all active atoms (i.e. those with $T_A>0$) are updated based on the new feature atom positions and the simulation proceeds with time stepping as discussed earlier. Newly activated atoms are assigned random velocities corresponding to their thermal activity, while atoms to be frozen are restored back to their perfect lattice position with zero velocities. The computational cost of MD region reconstruction is discussed in detail in Section 5.

2.3 Atomic Activation and Deactivation

A collection of vacancies (or any other feature type) can have numerous instantaneous configurations. As a result the distribution of defective atoms found using the above approach can fluctuate significantly between reconstructions even in the absence of center-of-mass diffusion. Such fluctuations can greatly increase the number of atoms that need to be activated (or deactivated) during each reconstruction. This can significantly affect the dynamics because of the operations involved in atom activation/deactivation. As mentioned above, atoms are activated by assigning them random velocities (scaled to the simulation temperature) and therefore are not

in equilibrium with the pre-existing active lattice. Note that activated atoms rapidly become thermally equilibrated with rest of the active lattice during the time integration following activation. Upon deactivation, atoms are moved back to their perfect lattice position, which can also lead to a temporary disturbance of the neighboring active lattice.

It is therefore highly desirable to minimize the number of atoms being activated and deactivated during each reconstruction and the following approaches are used. An atom is activated only if has at least one active neighbor. This reduces the potential impact of sudden activation of large numbers of atoms, which cannot be due to long-range diffusion because of the frequent reconstruction schedule. Because of the possibly significant atomic displacement, the impact of atom deactivation is even larger than that of activation and deactivation is performed only if an atom is consistently found to be outside an active region over the course of several reconstructions. This is achieved by recording a moving average of the atomic thermal activity, T_A^{avg}, defined at the reconstruction intervals. During each reconstruction interval, an atom is tagged for deactivation by reducing its T_A^{avg} by 10%. Note that the actual activity is not affected and T_A^{avg} serves only as a bookkeeping variable. Only if the same atom is tagged for deactivation five consecutive times, leading to the condition that $T_A^{avg} < T_A^{\min} \sim 0.5$, is the atom actually deactivated. If, during any reconstruction, an atom with $T_A^{avg} < 1$ is tagged as active (on the basis of its proximity to a feature atom), then T_A^{avg} is reset to a value of unity. In this way, only persistent feature atom motion (due to diffusion) leads to deactivation, while local configurational fluctuations do not.

2.4 Pressure Balancing Between MD Regions

Impurity and point defect mobilities and thermodynamic properties depend sensitively on pressure[35,36]. For example, the pressure dependence of the single vacancy diffusivity in silicon, as predicted by the EDIP potential, is shown in Figure 5. In the FAMD algorithm outlined above, each active region is mechanically isolated from all others, and the local pressure is determined by the surrounding static lattice density, as well as the concentration of feature atoms within the active region. As active regions merge and exchange defects, the number of defects per active atom can change dramatically, leading to a large change in the local density, and therefore pressure. As shown in Figure 5, such variations can lead to significant changes in the effective diffusivity, which in turn can affect the predicted evolution during a simulation of aggregation. Pressure balancing across the entire system is achieved using the Berendsen barostat[37]

$$\mu_L = [1 - c_r(P_0 - P)]^{1/3} \tag{1}$$

where μ_L is the lattice rescaling factor, c_r is an adjustable constant, and P_0 is the set point pressure. Each t_R timesteps, prior to MD region reconstruction, the pressure moving average (over the last t_R steps) is computed over all existing MD regions and eq. (1) is used to compute the rescaling factor required for the atomic coordinates *in the entire simulation cell*. More frequent updates are unnecessary because the distribution of defects per active atom is unchanged until the active regions themselves are updated. The net effect of the barostat is to provide mechanical communication between all the active regions in the simulation and produce conditions that are a reasonable approximation to constant overall pressure P_0.

An important subtlety arises in the measurement of the pressure within each MD region. In a standard periodic MD cell, the system pressure can be evaluated as a single value across the entire MD cell using[32]

$$P = \sum_i \frac{1}{3}\left(\sum_{j \neq i} r_{ij} \cdot f_{ij} + m_i v_i^2 \right). \tag{2}$$

where r_{ij} is the distance between particles i and j and f_{ij} is the force on particle i due to particle j. However, in an FAMD simulation, active atoms near the periphery of an active region contain inactive atoms in their neighbor lists and therefore the summation necessarily includes active-inactive atom pairs. While the first term in eq. (2) can be evaluated for such pairs, the second cannot because velocity cannot be defined (i.e. it is formally set to zero) for a static atom. This inconsistency leads to an ill-defined pressure contribution from active atoms near the boundary of an active region. In order to circumvent this problem, the atomic scale stresses $\sigma_i^{\alpha\beta}$ (α,β=x,y,z) on each atom[38] are first evaluated. Then pressure in the active region is computed by averaging over a subset of the atoms, i.e. excluding the ones near the active region periphery, as explained below:

$$\sigma_i^{\alpha\beta} = -\frac{1}{\Omega_i}\left(\frac{1}{2}\sum_j f_{ij}^{\alpha} r_{ij}^{\beta} + m_i v_i^{\alpha} v_i^{\beta} \right), \tag{3}$$

$$\left\langle \sigma^{\alpha\beta} \right\rangle = \frac{1}{V}\sum_i \Omega_i \sigma_i^{\alpha\beta} , \tag{4}$$

$$P = -\frac{1}{3}\sum_{\alpha}\left\langle \sigma^{\alpha\alpha} \right\rangle . \tag{5}$$

In eqs. (3)-(5), $\sigma_i^{\alpha\beta}$ is the ($\alpha\beta$) component of the stress tensor acting on atom *i*. The radial distribution of the virial component of the pressure (first term in eq. (2)) around a single vacancy is shown in Figure 6 in both a standard periodic cell and within an FAMD active region. Note that in the FAMD case, the virial terms in the vicinity of the boundary are adversely affected by the presence of static atoms – atoms in this region are not included in the calculation of the active region pressure. In the remainder of the active region, both MD cells exhibit identical virial components, showing that the boundary effect is localized, and provides evidence for the validity of the overall FAMD approach. Also note that the vacancy itself leads to a very large local stress field, but the number of atoms that experience this variation are small compared to the total number of atoms in the region and therefore are included in the overall region pressure estimate. The higher effective concentration of vacancies in the pressure evaluation is a potential source of error in FAMD but one that can be reduced systematically by increasing the active region radius around each vacancy.

Finally, the effect of including a transition region on the pressure distribution was considered. Interestingly, the presence of a transition region (i.e. $R_A>R_1$) was found to result in additional pressure (virial term) fluctuations at the active-transition region interface and did not seem to offer any advantages vis-à-vis a system without a transition region – at least for the transition region temperature functions considered here. In the remainder of this paper (as well as in Figure 6) the transition region size is set to zero i.e. $R_1=R_A$.

3. TESTING AND VALIDATION OF THE FAMD METHOD

The FAMD approach was tested extensively with respect to accuracy and efficiency as applied to vacancy diffusion and aggregation in silicon using the EDIP potential. In particular, various parameters, such as the active region radius, R_1, and MD region update frequency t_R, were tuned to optimize the simulation. It is expected that only slight modifications would be needed to consider other species and host lattices, and the application of this method to another system will be considered later in this paper.

3.1 Single Vacancy Diffusion

The random walk of a single vacancy was used to investigate the effect of active region size. A sequence of (NVT ensemble) diffusion measurement runs was performed using varying active region radii, R_1. In each simulation, the zero pressure density was first computed using the Berendsen barostat as discussed in Section 2.3. The single vacancy diffusivity, D_{1V}, was estimated using the Einstein relation as in a regular MD simulation, where

$$D_{1V} = \frac{1}{6}\frac{[r_{1V}^{CM}(t) - r_{1V}^{CM}(0)]^2}{t}. \tag{6}$$

In eq. (6), $r_{1V}^{CM}(t)$, is the center-of-mass position of the vacancy (also applicable to vacancy clusters) at time t, which for a single vacancy, is taken to be equilibrium lattice position of the corresponding missing atom. The calculated diffusivity as a function of active region radius is shown in Figure 7. Also shown is the value obtained from standard MD simulations.

Clearly, an insufficient number of active atoms in the active region leads to a coupling between the vacancy dynamics and the static atoms at the edge of the active region, and therefore a lowered diffusion coefficient. About 400 active atoms are required to remove the coupling and an active region radius of 12.6 Å is used in all the following calculations based on the EDIP potential. Larger active regions continue to make very small differences; smaller than the uncertainty associated with the MD measurement of the diffusion coefficient, which is about 10% in this case. Note that in a simulation cell with periodic boundary conditions, only about 200 atoms are needed to isolate the vacancy from finite cell size effects, but periodic boundary conditions are not appropriate here because they completely decouple an active region from the surrounding lattice.

The validity of this active region radius choice was further confirmed by performing a constrained path energy minimization along the diffusion path of the vacancy, as shown in Figure 8. During the minimization only the atoms within the active regions R_1 were minimized. The relaxed energy profiles are indistinguishable throughout the hop further confirming that the vacancy is unaffected by the static boundary and the defect strain fields are confined well within the active region. Clearly, the minimum size of the active region required for a single feature atom depends on several factors including the type of feature atom as well as the nature of the interatomic potential. Specifically, the range of the potential, and whether it includes three-body interactions, can play significant roles in determining the number of active atoms that must be considered. These issues will be discussed in Section 6, where FAMD calculations using another interatomic potential are presented.

3.2 Vacancy Dimer Dynamics

The vacancy dimer is the simplest structure that can be used to test the remaining components of the FAMD algorithm, i.e. active region coalescence and fragmentation. The evolution of the number of active atoms, N_A, the vacancy-vacancy separation, $R_{V\text{-}V}$, and total mean-square displacement, MSD, are shown in Figure 9 for a two-vacancy system at a temperature of 1600 K. During the simulation, two distinct regimes can be observed that clearly delineate the periods in which the dimer is bound and those in which the two vacancies are isolated. The number of active atoms fluctuates between 800 (2x400 per active region) when the active regions are fully separated, and about 470 when the two active regions merge into a single region and the two vacancies are in the nearest neighbor configuration.

Note that the vacancy dimer remains bound as a single cluster entity for vacancy separations up to 7.8 Å[22,23]. The extended vacancy-vacancy interaction distance allows the vacancy dimer (and larger clusters) to diffuse quite rapidly and to assume a variety of configurations during diffusion[19]. The various configurations can be identified in Figure 9 in both the evolution of the vacancy-vacancy separation distance, $R_{V\text{-}V}$, and the number of active atoms. In other words, as the dimer size increases temporarily during diffusion, the number of active atoms is also seen to increase even though a single active region is present. Occasionally, the two vacancies diffuse away from each other and the dimer fragments. If their separation becomes large enough, the single active region also fragments into two separate regions. The two distinct slopes of the MSD curve shown in Figure 9 correspond well with values for single and dimer vacancy diffusivities measured with standard MD simulations[19].

3.3 FAMD Simulation of Vacancy Aggregation in Silicon

Once the various components of the FAMD simulation were tested for single vacancy and dimer diffusion in EDIP silicon, it was used to simulate vacancy aggregation in a large simulation box. Two systems, each containing 216,000 silicon lattice sites were investigated. In the first, 1,000 vacancies were placed at uniformly separated sites, while in the second 343 vacancies were included, also in uniformly separated sites. The 1,000-vacancy system was previously simulated using a parallel implementation of standard MD (PMD) based on the EDIP potential[19] and therefore provides a rigorous test of the overall accuracy of the FAMD approach. The simulation temperature in both cases was fixed at 1600 K and the pressure set to zero. This high temperature was chosen to accelerate the aggregation process and thereby maximize the extent of the comparison. The time step in both simulations was 0.77 *f*s. These conditions are identical to those used in the PMD simulation described in ref. [19]. In both of the FAMD simulations the active region radius was determined by activating all atoms within 12.6 *Å* of a vacancy. This is based on the fact that for silicon vacancy aggregates as described by EDIP potential, the interaction distance is not a function of the aggregate size, i.e. it is roughly the same as that for a single vacancy[20].

The evolution of the vacancy size distribution was monitored according to the method outlined above and in ref. [19] in which periodic quenches of atomic coordinates with a conjugate-gradient algorithm[34], followed by a comparison of the atomic coordinates to a reference perfect crystal were used to locate each vacancy. Once the vacancies were located they were assigned to physical clusters based on the Stillinger criterion, in which a vacancy-vacancy interaction distance of 7.8 A was assumed, which corresponds to the 4th-nearest neighbor shell

along the (110) direction (4NN-110) as determined in earlier work[19]. Note that cluster tracking (every time the active region topology is updated) is part of the FAMD algorithm so that no extra work is needed to compute the vacancy size distribution in this case.

Various components of the vacancy size distribution as predicted by the PMD and FAMD simulations are shown in Figures 10(a) and 10(b) for the 1,000-vacancy case. Specifically, the number of monomers (X_1) and dimers (X_2) are shown in Figure 10(a) for both the PMD and FAMD simulations. The 1,000 vacancy PMD and FAMD simulations are clearly in excellent agreement. The deviation in the dimer concentration at very early times is due to statistical fluctuations between individual runs, which can be large when very few clusters are present, and the dimer profiles are observed to converge rapidly once the number of dimers increases. The agreement between the two 1,000 vacancy simulations is equally good in Figure 10(b) which shows the evolution of the mean cluster size (M_2/M_1) and total number of clusters (M_0), where $M_n = \sum_s s^n X_s$ are the nth-order moments of the size distribution, and s is an index that runs over all cluster sizes observed in the simulation.

The evolution of the 343-vacancy system (not shown) is qualitatively similar to that of the 1,000-vacancy case, but exhibits a longer lag at the start of the evolution due to the increased diffusion distance between vacancies. Note that even these conditions represent a very large vacancy concentration (0.2%) relative to typical experimental conditions found during silicon crystal growth and wafer thermal annealing[39]. More dilute concentrations will be considered in the future with a parallel version of the FAMD algorithm. Snapshots of sections of the 343-vacancy system after 2 *n*s of evolution are shown in Figure 11. Active atoms, inactive atoms, and

feature atoms are delineated with different shades. Clearly, most of the atoms in the system are inactive implying a large savings in computational effort due to the fact that very few single vacancies remain at this stage of the aggregation process. The active regions are mainly spherical implying that cluster morphology is generally compact and three-dimensional. Several instances of recently merged active regions can be observed in which two compact clusters have diffused close enough to activate atoms between them, allowing for a possible coalescence event to take place.

4. PERFORMANCE AND EFFICIENCY OF THE FAMD METHOD

The evolution of the total number of active atoms in the 343-vacancy system is shown in Figure 12. At the beginning of the simulation, only about 65 % of the total atoms in the cell are active, corresponding to 343 isolated active regions – one for each vacancy in the system. By comparison, in the 1,000-vacancy system, the vacancy concentration is sufficiently high to lead to a 100% activation fraction, although this number decreases as the aggregation process evolves. As clusters are formed, defect features have an increasing number of overlapping active atoms, therefore the numbers of active regions and total number of active atoms decrease rapidly, and by the time the 343-vacancy simulation is terminated at about 3 million time steps, the total number of active atoms is less than 35,000, or about 16% of the total system.

The computational time needed for performing a single time step is shown in Figure 13 as a function of the number of active atoms for both EDIP (343 vacancies in 216,000 sites) and EAM (64 vacancies in 23,328 sites) simulations. The latter are discussed further below in

Section 5. The overall computational effort scales as $N_A^{0.75}$ for EDIP and $N_A^{0.85}$ for the EAM simulation, both of which are somewhat lower than the expected linear scaling. The primary reason for this deviation is that additional (inactive) atoms must be considered when the forces on the active atoms are being evaluated.

The less favorable scaling of the EDIP calculations is due to the specific structure of the EDIP potential, which requires the consideration of a substantial number of inactive atoms in a shell around each active region. This requirement arises from the fact that the energy E_i (and therefore forces $\boldsymbol{f}_i$) are functions of both inter-particle separation and an *effective coordination number* Z_i which appears in both the two and three-body components of the potential as

$$\begin{aligned} E_i &= \sum_j V_2(r_{ij}, Z_i) + \sum_{j \neq i} \sum_{k \neq i, k > j} V_3(\vec{r}_{ij}, \vec{r}_{ik}, Z_i) \\ Z_i &= \sum_{i \neq j} f(r) \end{aligned} \tag{7}$$

and

$$f(r) = \begin{cases} 1 & r < c \\ \alpha/(1 - x^{-3}) & c < r < b \\ 0 & r > b \end{cases}, \tag{8}$$

where $x = (r - c)/(b - c)$. The cutoff distances b and c lie in between the first and second nearest neighbor positions. Therefore, evaluation of the total force $\boldsymbol{f}_i$ for a particle i at the border of the active region (for example) requires computation of Z_i and also Z_j for all the neighboring j's i.e. information is needed for atoms up to the twice the potential cutoff distance from the active

region boundary. In other words, even for an inactive neighbor j of an active particle i, it is necessary to loop over all j's neighbors, k, in order to evaluate Z_j and the force component $\boldsymbol{f}_{ji}$.

The performance of the FAMD algorithm, using the EDIP potential, is shown in Figure 14 as a function of the vacancy concentration. At each of the three vacancy fractions considered, 1.59 x10^{-3}, 3.64 x10^{-4}, and 1.25 x10^{-4}, the speedup relative to identical standard MD simulations is shown to increase once the vacancy cluster size distribution coarsens. As the systems become more dilute the initial speedup relative to MD increases, but the onset of aggregation and concomitant increase in efficiency is delayed towards later times, due to greater initial vacancy-vacancy separations. At a vacancy fraction of 1.25x10^{-4} vacancies per site, the FAMD method is approximately 12 times faster per time step relative to standard MD at the beginning of the simulation and the speedup approaches a factor of 20 by about 5 million time steps. Note that these simulation timescales would not be readily accessible using standard MD on a single processor for a system of this size and demonstrates how the FAMD method can be used to extend the scope of aggregation simulations.

As mentioned earlier, the most significant contribution to overhead in FAMD relative to regular MD is the identification of feature positions in order to keep current the MD region information. The conjugate gradient (CG) minimization scheme takes approximately 5-10 steps to quench the atomic coordinates to a level sufficient to identify the vacancy positions. This is roughly equivalent to 25 force evaluations. For the conservative choice of t_R=100, this would imply that the total simulation cost per FAMD step is about 1.25 times the regular MD step. In Figure 14, this overhead is factored out so that the relative computational effort for a single time-

step can be compared. Note that this overhead source is required only for the case of vacancies, which are notoriously difficult to isolate. Other feature atoms such as impurities, or even interstitials are likely to be identifiable "on-the-fly" and in this case, the defect-location overhead would be negligible. More computationally expedient approaches for tracking vacancies will be discussed in a future publication.

The cost of performing bookkeeping work (in order to construct the hybrid lattice), such as identification of feature clusters, creation/destruction of MD regions, and thermalization of atoms within the MD regions represents only a small additional source of overhead. Figure 15 shows the distribution of these costs, along with cost of CG minimization, for a system with single vacancy in various lattices containing 1728-4.096x10^6 host sites. The time required to perform a single MD step (T_{MD}), as well as time to quench the active lattice for location of defect positions (T_{CG}) stays nearly constant with increasing lattice sizes. The slight increase that is apparent is likely due to increasing array lengths but is much weaker than linear. The remaining bookkeeping overhead cost (T_{HY}), although linearly dependent on the total lattice size, is small, and even for a realistic concentration of 10^{-6} it represents less than 5% of the total computational cost. Other costs associated with creating cell linked lists and Verlet neighbor lists also are small because they are constructed only for atoms in the active regions and are not dependent on the actual lattice size.

4.1 FAMD Stability and Sensitivity Analysis

Smaller simulation systems containing 125 vacancies in host lattices of 27,000 silicon atoms were simulated (also at 1600 K and zero pressure) using both FAMD and standard PMD

in order to study the effect of various parameters e.g. t_R, R_c^d, R_1, R_A on the efficiency and accuracy of the FAMD simulation. The predicted average cluster sizes up to 7x10^6 time-steps are shown in Figure 16 for both the base case FAMD parameter set (with t_R=100, $R_1 = R_A = 12.6$ Å and $R_c^d = 0.45 r_{bond}$) and standard MD. The agreement is good throughout the simulations, however, discrete effects become important near the end where few clusters remain and the average cluster size begins to evolve in a stepwise fashion due to coalescence. In fact, at the end of the FAMD simulation, less than 8000 (< 30%) atoms are active and there are fewer than 20 clusters. Note also that the overall power-law evolution is identical to that for the larger 1000 vacancy/216,000-atom system previously shown in Figure 10(b), which is expected because the vacancy fraction is the same in both simulations.

The selective activation and deactivation scheme proposed in Section 2.3 plays an important role in the stability of the FAMD method. Initial simulations in which atoms were activated and deactivated purely based on their instantaneous proximity to feature atoms resulted in a gradual deterioration of the FAMD aggregation dynamics, exhibiting (reproducible) rapid cluster coalescence and a divergence in the average cluster size, as shown by the triangles in Figure 16 (denoted by arrow). This divergence was observed even for simulations in which the active region size, R_1, was increased and the threshold for identifying defective atoms, R_c^d, was decreased. More frequent reconstructions with $t_R = 20$ actually resulted in an earlier appearance of the divergent behavior. In order to examine the hypothesis that the observed deviation was due to inadequately represented long-range displacement fields, a modified simulation was developed in which the entire static lattice was relaxed during MD region reconstruction, but no effect on the anomalous dynamics was found. These tests conclusively indicated that the problem

was not arising because of insufficient relaxation inside active regions, but rather was due to the effects of frequent atomic activation and deactivation. The scheme presented in Section 2.3 addresses the inherent uncertainty associated with vacancy position identification and the resulting active region "wobble" by effectively assigning *inertia* to the activity of atoms.

Finally, the robustness of the approach with respect to the parameters R_1 , R_c^d was tested by performing aggregation runs with three different pairs of R_1 and R_c^d that gave approximately the same number of active atoms. The parameter values used in the three additional runs were (all in units of Å) {9.40, 0.20}, {10.5, 0.42} and {11.5, 0.84} as compared to the base case values {12.6,1.05}. It was found that each of the three test runs led to essentially identical evolution profiles, confirming that the FAMD approach is robust with respect to choices of these parameters if a sufficient number of active atoms is used.

5. APPLICATION OF FAMD TO EAM POTENTIAL SYSTEMS

The FAMD simulation approach was applied to the simulation of vacancy diffusion and clustering in FCC copper using an Embedded-Atom Method (EAM) potential in order to confirm the transferability of the approach. This system was chosen because the EAM potential differs in several fundamental ways to the EDIP potential and has wide applicability in a variety of metallic and non-metallic systems, therefore providing a good test of the general usefulness of the FAMD approach. The EAM potential is a linear scaling, multi-atom, semiempirical approach to modeling of interatomic forces in pure and multi-component metallic systems. In the EAM, as

originally proposed by Daw and Baskes[40], the total energy E is a sum over energies E_i of N atoms in the system, where E_i is given by,

$$E_i = \frac{1}{2}\sum_j V(r_{ij}) + F(\overline{\rho_i})$$
$$\overline{\rho_i} = \sum_j \rho(r_{ij}) \tag{9}$$

where, $V(r_{ij})$ is the pair potential, $F(\overline{\rho_i})$ is the embedding potential and $\rho(r_{ij})$ is the electronic density function. The EAM potential used here was developed by Mishin and coworkers[30] for simulation of copper and was shown to reproduce well the properties of the bulk crystal as well as non-equilibrium structures well as predicted by *ab initio* calculations. The potential and forces were efficiently evaluated using tables constructed from analytical functional in eq. (9) and its derivatives. A total of 8,192 elements each were used over the entire range of r and ρ values. The resulting potential evaluations ran 3 times faster when compared to the analytical function evaluations, with relative error of less than 10^{-5} in energy evaluations, and in the remainder of this section, the table interpolation evaluation method is used for all EAM simulations.

The implementation of an EAM potential differs from the EDIP version only in the actual incorporation of the potential into the code (as would be the case in a regular MD code). In addition, potential-specific parameters, namely the size of the active region around a single feature atom (R_A), must be optimized. EAM also has computational issues similar to those mentioned for EDIP in Section 5, i.e. the complete specification of forces on atom i requires evaluation of the derivatives of $F(\overline{\rho_i})$ and $F(\overline{\rho_j})$ for all its neighbor j's. But, this does not lead

to as large a computational penalty because both the terms in EAM potential are essentially pair-wise interactions and angular forces are not present. This enables a single *a priori* computation of the density, ρ_{ij}, and its derivatives due to inactive-inactive atom pairs for the entire time span of t_R time steps, which is then used to compute $\overline{\rho}$, $F(\overline{\rho})$, and their derivatives for inactive particles seen by active particles. Therefore, the neighbor list contains only the active-active and active-inactive pairs for which the potential and force evaluations are performed.

5.1 Parameterization of Cu-EAM FAMD

The EAM-FAMD simulation was tuned using single vacancy diffusion in the temperature range of 1000-1200K employing active region sizes between 10 and 18 Å. The MD region reconstruction interval, t_R, was set to 100 time steps as in the EDIP case and once again, no transition region was imposed (i.e. $R_1 = R_A$). Active region sizes smaller than 10 Å are not physically possible because the potential cutoff distance of 5.5 Å and a nearest-neighbor bond length of 2.56 Å imply that a single feature (diffusive) hop will bring it into direct interaction with the static lattice at the MD region edge. Indeed, an active region with radius of 9 Å was observed to lead to a large, uncontrollable pressure in the MD region, whereas a radius of 10 Å lead to a well-behaved pressure distribution in the MD region, as shown in Figure 6 for the EDIP case. A slightly larger active region was needed in order to reproduce the single vacancy diffusion coefficient robustly as shown in Figure 17 for several different temperatures, and an active region radius of 12.5 Å was used in all subsequent results. The cutoff for identifying excessively displaced (feature) atoms was once again set to ~$0.45r_{bond}$ or 1.1 Å. The consistency between the FAMD and MD diffusivities also show that the FAMD approach is valid across a range of temperatures.

5.2 Application of EAM-FAMD to Vacancy Aggregation in Copper

Large-scale vacancy aggregation runs were performed using the EAM system. The goal was to evaluate the applicability of FAMD method for EAM potential systems and its computational advantages over regular MD. One additional challenge in these simulations was the overall computational expense of the EAM potential relative to EDIP, which arises because of the relatively long range of the potential. As a result somewhat smaller systems were used in order to make feasible the runs on a single computational node. A system of 62,500 FCC sites containing 216 vacancies at a temperature of 1200 K and zero pressure was evolved using both FAMD and standard parallel MD. The evolution of the average cluster size, M_2/M_1 and total cluster number, M_0 as defined previously in Figure 10 are shown in Figure 18 and are in very good agreement throughout the simulation, once again validating the FAMD approach, making it applicable to a broad range of EAM potential systems[30,41,42,43,44].

The evolution of the number of active atoms and the computational cost per time step for an EAM simulation are shown in Figure 13 (for a system containing 64 vacancies in a 23,328-site lattice), along with the results from the EDIP simulations. Once again, the number of active atoms decays in time according to a power-law (exponent=0.85), following the coarsening dynamics. The simulation is observed to accelerate by over an order-of-magnitude during the course of the run. The improved scaling relative to the EDIP case is due to the smaller number of static atoms that must be considered during the evaluation of forces.

6. CONCLUSIONS

A new approach for performing MD simulations of solid-state aggregating systems has been presented, which takes advantage of the localized fields exhibited by electrically neutral point defects, impurities, and their respective clusters. It is shown that for appropriately selected active region sizes and update frequency, the predictions of the FAMD method are essentially indistinguishable from those of standard MD predictions. The method is particularly suited towards the simulation of dilute systems in which many of the atoms are not influenced by the presence of defects and/or impurity atoms.

The FAMD method scales almost linearly with the number of feature atoms and is essentially independent of the total system size allowing for the simulation of arbitrarily large (and correspondingly dilute) atomic systems. Even for relatively concentrated systems of aggregating vacancies in silicon and copper, as modeled by the EDIP and EAM potentials respectively, the FAMD approach is up to 20 times faster than standard MD, with no loss in accuracy. The method is expected to be broadly applicable to a wide range of microstructural evolution phenomena in crystalline materials and should allow for a substantial increase in the length scale accessible to MD simulations of solid-state aggregation phenomena.

ACKNOWLEDGEMENTS

We gratefully acknowledge support from the National Science Foundation (CTS01-34418).

REFERENCES:

1 K. Nakamura, T Saishuoji, T. Kubota, and T. Tanimura, Mat. Sci. Engg. B, **36**, 22 (1996).

2 V. V. Voronkov and R. Falster, J. Cryst. Growth, **204**, 462 (1999).

3 F. Shimura and R. S. Hockett, App. Phys. Lett., **48**, 224 (1986).

4 A. Karoui, F. Sahtout Karoui, G. A. Rozgonyi, M. Hourai, and K. Sueoka, J. Electrochem. Soc. **150**, G771 (2003).

5 P. A. Stolk, D. J. Eaglesham, H.-J. Gossmann, and J. M. Poate, Appl. Phys. Lett., **66**, 1370 (1995).

6 R. F. Scholz, P. Werner, U.Gösele and T. Y. Tan, Appl. Phys. Lett., **74**, 392 (1999).

7 Y. N. Osetsky, D. J. Bacon, A. Serra, B. N. Singh, S. I. Golubov, J. Nucl. Mat. **276**, 65 (2000).

8 R. Phillips, *Crystals, Defects and Microstructures: Modeling Across Scales*, (Cambridge University Press, Cambridge, 2001).

9 W. J. Phythian and C. A. English, J. Nucl. Mater. **205**, 162 (1993).

10 Yu N. Osetsky and A. Serra, Philos, Mag. A, **73**, 249 (1996) and **75**, 1097 (1997).

11 J. J. Blackstock and G. J. Ackland, Philos. Mag. A, **81**, 2127 (2001).

12 M. von Smoluchowski, Z. Phys. Chem, **92**, 129 (1917); Phys. Z. **17**, 583 (1916).

13 H. Risken, *The Fokker-Planck Equation: Methods of Solutions and Applications* (Springer-Verlag, Berlin, New York, 1989).

14 T. Sinno, R. A. Brown, W. von Ammon, and E. Dornberger, J. Electrochem. Soc. **145**, 302 (1998).

15 K. A. Fichthorn and W. H. Weinberg, J. Chem. Phys., **95**, 1090 (1991).

16 A. La Magna, and S. Coffa, Comput. Mat. Sci. **17**, 21 (2000).

17 M. Jaraiz, L. Pelaz, E. Rubio, J. Barbolla, G. H. Gilmer, D. J. Eaglesham, H.-J. Gossmann, J. M. Poate, Mater. Res. Soc. Proc. **568,** 43 (1998).

18 C. C. Battaile, D. J. Srolovitz and J. E. Butler, J. Appl. Phys., **82**, 6293 (1997).

19 M. Prasad and T. Sinno, Phys. Rev. B, **68**, 045206 (2003).

20 M. Prasad and T. Sinno, Phys. Rev. B, **68**, 045207 (2003).

21 S. S. Kapur, M. Prasad, and T. Sinno, Phys. Rev. B, **69**, 155214 (2004).

22 A. Bongiorno and L. Colombo, Phys. Rev. B, **57**, 8767 (1998).

23 T. Sinno and M. Prasad, Mol. Phys., **102**, 395 (2004).

24 R. E. Rudd and J. Q. Broughton, Phys. Stat. Sol. B, **215**, 251 (2000).

25 F. F. Abraham, J. Q. Broughton, N. Bernstein and E. Kaxiras, Comput. Phys. **12**, 538 (1998); Europhys. Lett., **44**, 783 (1998).

26 E. B. Tadmor, M. Ortiz and R Phillips, Phil. Mag. A, **73**, 1529 (1996).

27 M. Ortiz and R. Phillips, Adv. Appl. Mech., **36**, 1 (1999).

28 M. Z. Bazant, E. Kaxiras, J. F. Justo, Phys. Rev. B **56** 8542 (1997).

29 J. F. Justo, M. Z. Bazant, E. Kaxiras, V. V. Bulatov, and S. Yip, Phys. Rev. B **58**, 2359 (1998).

30 Y. Mishin, M. J. Mehl, D. A. Papaconstantopoulos, A. F. Voter and J. D. Kress, Phys. Rev. B, **63**, 224106 (2001).

31 F. H. Stillinger, J. Chem. Phys., **38**, 1486 (1963).

32 M. P Allen, and D. J. Tildesley, *Computer Simulations of Liquids* (Clarendon Press, Oxford, 1987).

33 D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, San Diego, 1996).

34 J. C. Gilbert, and J. Nocedal, SIAM J. on Optim., **2**, 21 (1992).

35 A. Antonelli, E. Kaxiras and D. J. Chadi, Phys. Rev. Lett., **81**, 2088 (1998).

36 T. Halicioglu and D. M. Barnett, Surf. Sci, **441**, 265 (1999).

37 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys., **81**, 3684 (1984).

38 S.-P. Chen, T. Egami and V. Vitek, Phys. Rev. B, **37**, 2440, (1998)

39 R. Falster, V. V. Voronkov and F. Quast, Phys. Stat. Sol., **222**, 219 (2000).

40 M. S. Daw and M. I. Baskes, Phys. Rev. B, **29**, 6443 (1984).

41 Y. Mishin, D. Farkas, M. J. Mehl, and D. A. Papaconstantopoulos, Phys. Rev. B, **59**, 3393 (1999).

42 Y. Mishin, M. J. Mehl and D. A. Papaconstantopoulos, Phys. Rev. B, **65**, 224114 (2002).

43 S. M. Foiles, M. I . Baskes and M. S. Daw, Phys. Rev. B, **33**, 7983 (1986).

44 Y. Li, D. J. Siegel, J. B. Adams and X.-Y. Liu, Phys. Rev. B, **67**, 125101 (2003).

Figure 1: Schematic representation of active regions surrounding a distribution of aggregating particles.

Figure 2: Flowsheet of the basic operations in a FAMD simulation.

Figure 3: Various components of an *MD region* construct. (1) Fully active atom region – radius = R_1, (2) Transition region – radius = R_A, (3) Inactive neighbor region – radius = R_N. Atoms that are further than R_N from the nearest feature are included in the MD region but do not contribute to the overall computational effort.

Figure 4: Thermal activity in an MD region as a function of radial distance from a single feature atom.

Figure 5: Vacancy diffusivity in bulk crystalline silicon as a function of dimensionless pressure as predicted by the EDIP potential.

Figure 6: Radial evolution of the virial component of the pressure around a vacancy placed in a (a) standard periodic MD cell, and (b) FAMD active region surrounded by a static lattice. Atoms in region (1) are used to compute the overall pressure of the active region, while those in region (2) are discounted due to their interaction with the active region boundary.

Figure 7: Single vacancy diffusivity as a function of active region radius, R_1. The number of active atoms for each of the cases shown are 274, 328, 416, and 1014, respectively.

Figure 8: Constrained path minimized energy along a vacancy hop in a standard MD and FAMD cell (R_1=12.6 Å).

Figure 9: Evolution of (a) Number of active atoms, N_1, (b) Vacancy-vacancy separation distance, $R_{V\text{-}V}$, and (c) mean-square displacement in a two-vacancy system simulated using FAMD.

Figure 10: (a) Evolution of the number of single vacancies (X_1) and vacancy dimers (X_2) as a function of time for (i) PMD (circles), and (ii) FAMD (squares) simulations; (b) Evolution of the (i) mean cluster size (M_2/M_1) and (ii) total cluster number (M_0) of vacancy clusters as a function of time.

Figure 11: Snapshot of a slab taken from the 343-vacancy FAMD cell (t = 2 *n*s) showing active regions (dark gray atoms are active) centered about feature atoms (black spheres are vacancies). Small light gray atoms are static.

Figure 12: Evolution of the fraction of active atoms in an FAMD simulation of 343 vacancies in a 216,000-site silicon crystal.

Figure 13: Scaling behavior of the computational effort per time step as a function of the number of active atoms in the simulation cell. EDIP potential simulations (343 vacancies in 216,000 sites) – circles; EAM simulations (64 vacancies in 23,328 sites) - Squares. Power law fit exponents are 0.72 and 0.85, respectively.

Figure 14: Percent speedup of the FAMD algorithm as compared to standard MD for several vacancy concentrations.

Figure 15: Overhead costs associated with the FAMD algorithm for single vacancy in lattices ranging from 1728-4.096x10^6 concentrations. (a) Squares – time to perform regular MD (T_{MD}) on active atoms, (b) Triangles – time to perform lattice minimization (T_{CG}) and (c) Circles – time to construct the hybrid lattice (T_{HY}). Basis: 100 time steps.

Figure 16: Evolution of the mean cluster size (M_2/M_1) as a function of time for a system of 125 vacancies in a 27,000-site silicon lattice. (a) Squares – standard PMD, (b) Circles – FAMD with activation/deactivation inertia, and (c) Triangles – FAMD without reconstruction inertia (highlighted by arrow).

Figure 17: Single vacancy diffusivity in FCC copper as a function of temperature for standard MD (squares) and FAMD (diamonds). Symbol size corresponds to one standard deviation from the mean values.

Figure 18: Evolution of the mean cluster size (M_2/M_1) and total cluster number (M_0) as a function of time for a system of 216 vacancies in a 62,500-site FCC copper lattice. (a) Squares – standard PMD, (b) Circles – FAMD.

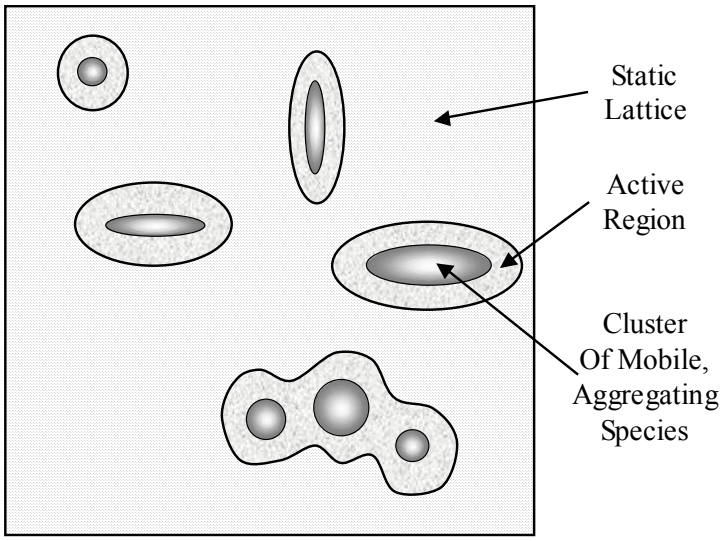

Figure 1: Prasad and Sinno, Journal of Chemical Physics

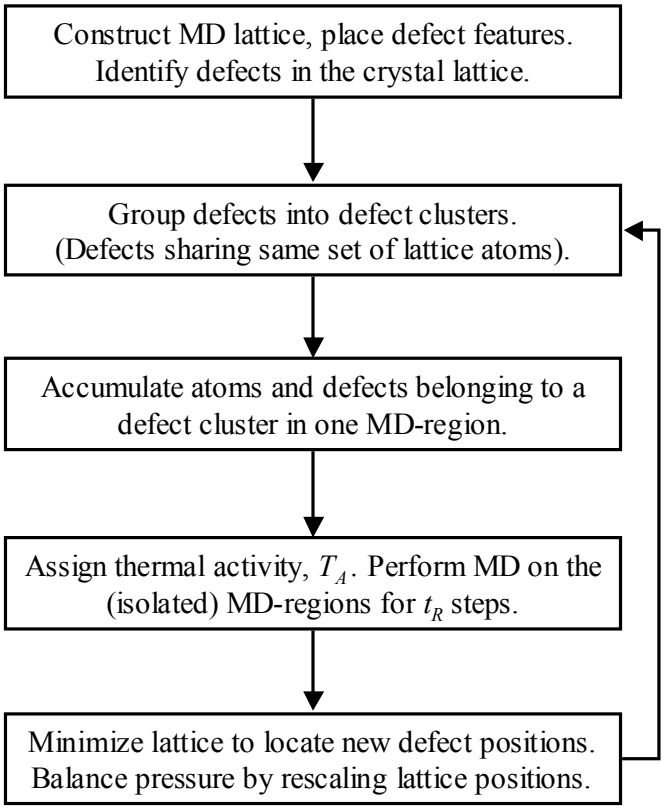

Figure 2: Prasad and Sinno, Journal of Chemical Physics

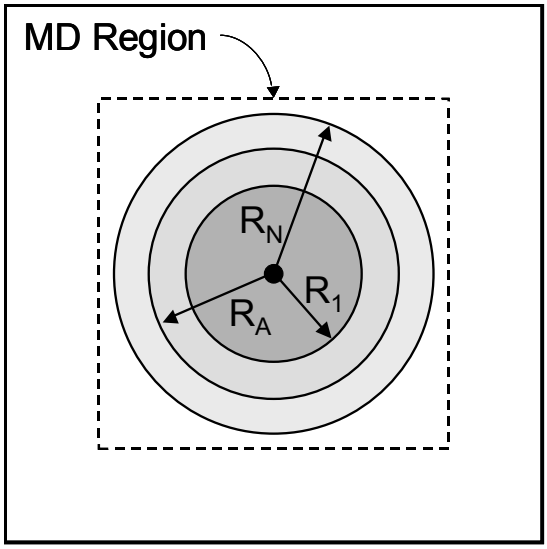

Figure 3: Prasad and Sinno, Journal of Chemical Physics

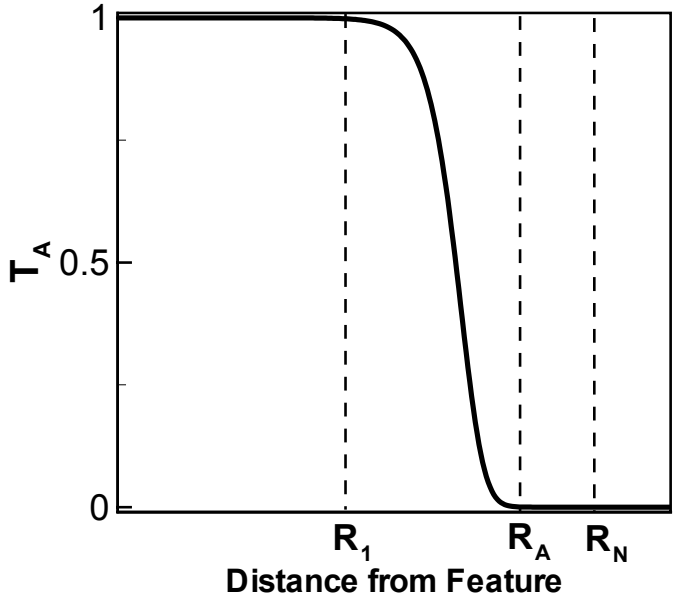

Figure 4: Prasad and Sinno, Journal of Chemical Physics

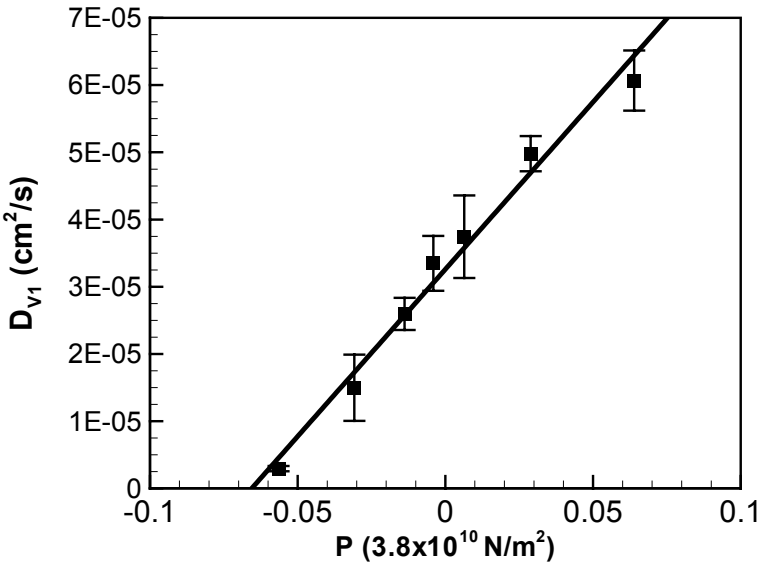

Figure 5: Prasad and Sinno, Journal of Chemical Physics

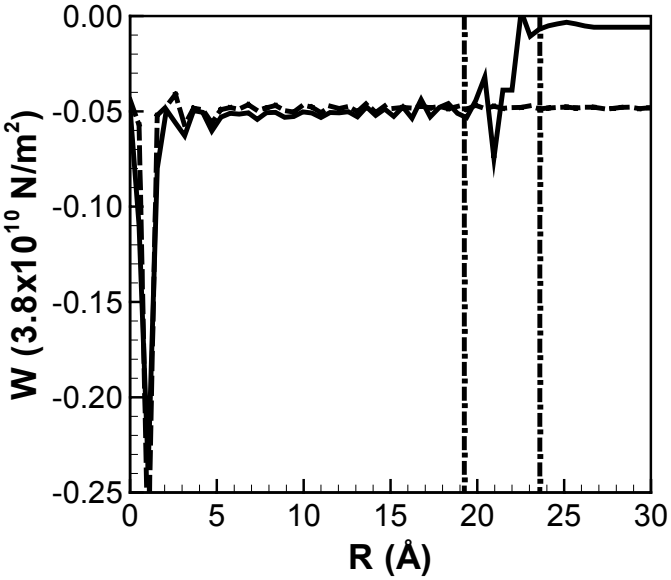

Figure 6: Prasad and Sinno, Journal of Chemical Physics

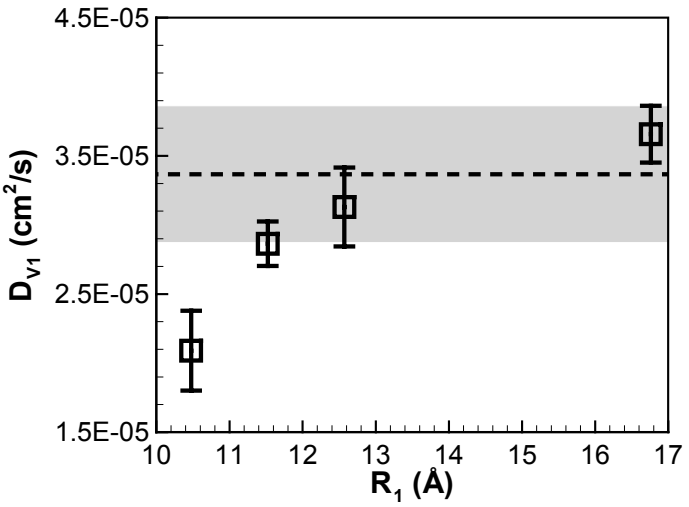

Figure 7: Prasad and Sinno, Journal of Chemical Physics

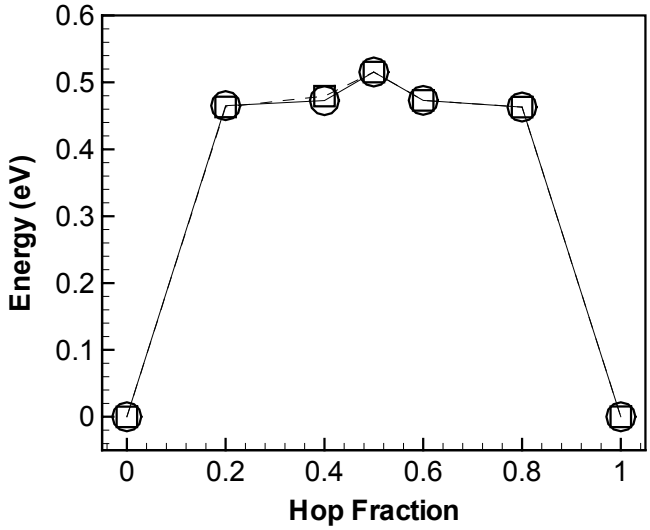

Figure 8: Prasad and Sinno, Journal of Chemical Physics

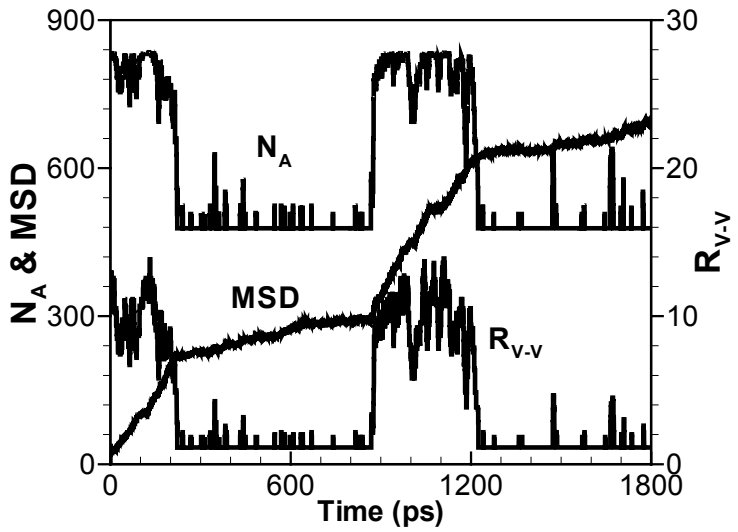

Figure 9: Prasad and Sinno, Journal of Chemical Physics

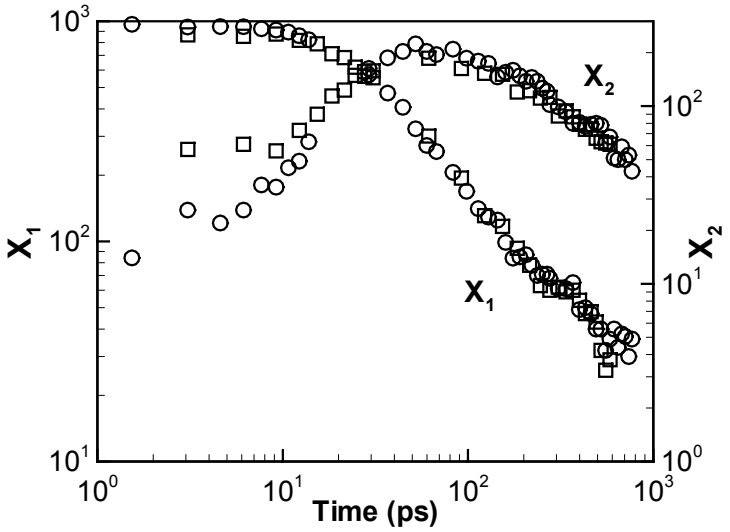

Figure 10(a): Prasad and Sinno, Journal of Chemical Physics

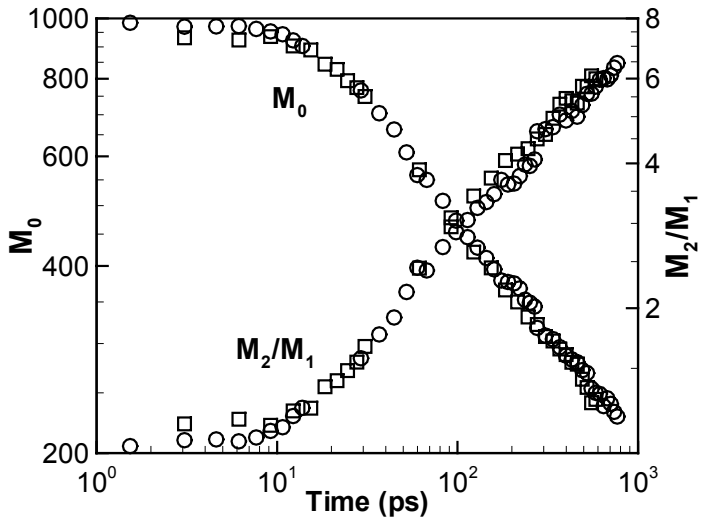

Figure 10(b): Prasad and Sinno, Journal of Chemical Physics

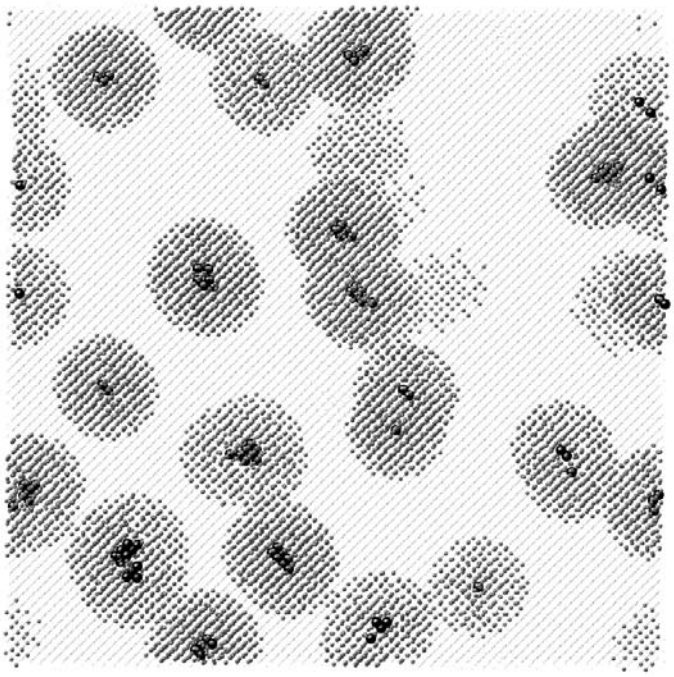

Figure 11: Prasad and Sinno, Journal of Chemical Physics

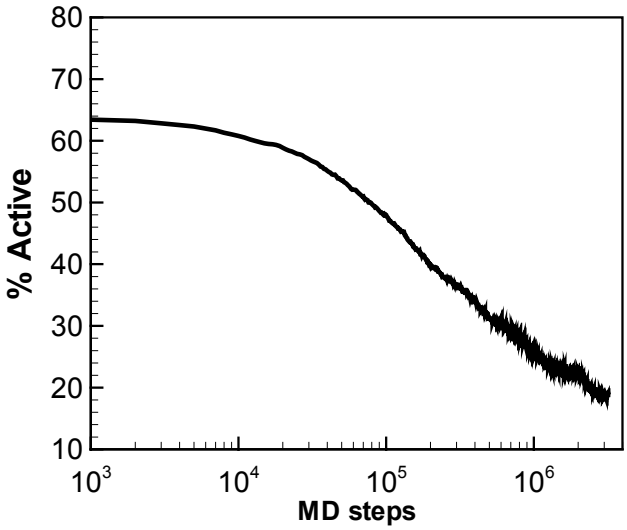

Figure 12: Prasad and Sinno, Journal of Chemical Physics

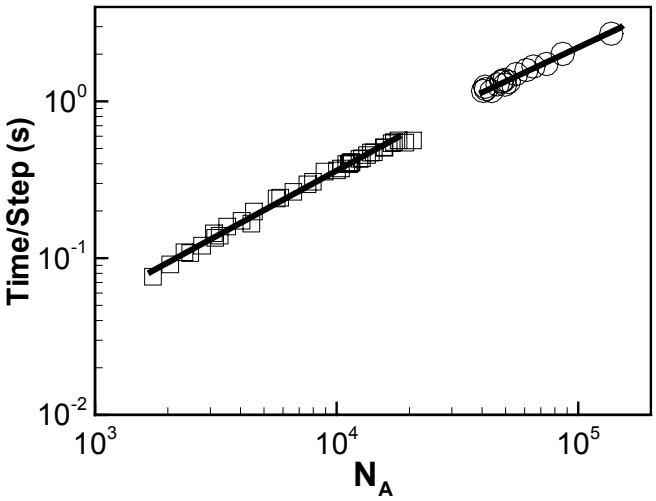

Figure 13: Prasad and Sinno, Journal of Chemical Physics

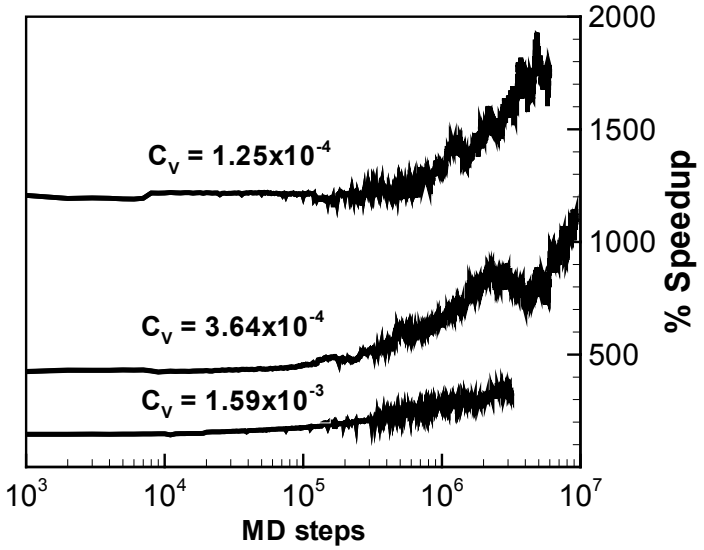

Figure 14: Prasad and Sinno, Journal of Chemical Physics

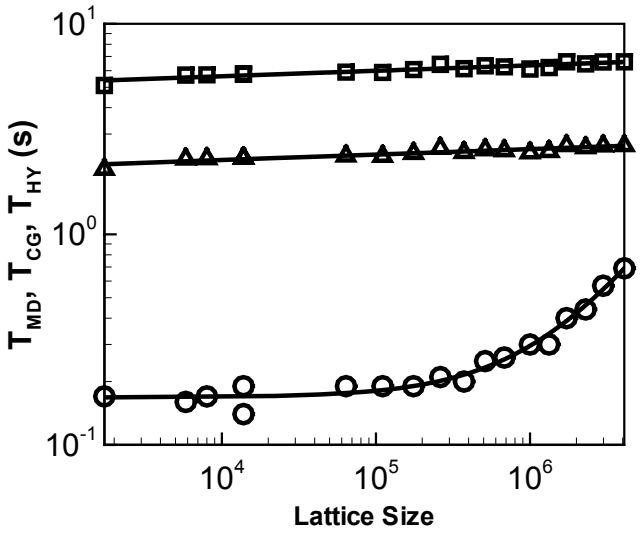

Figure 15: Prasad and Sinno, Journal of Chemical Physics

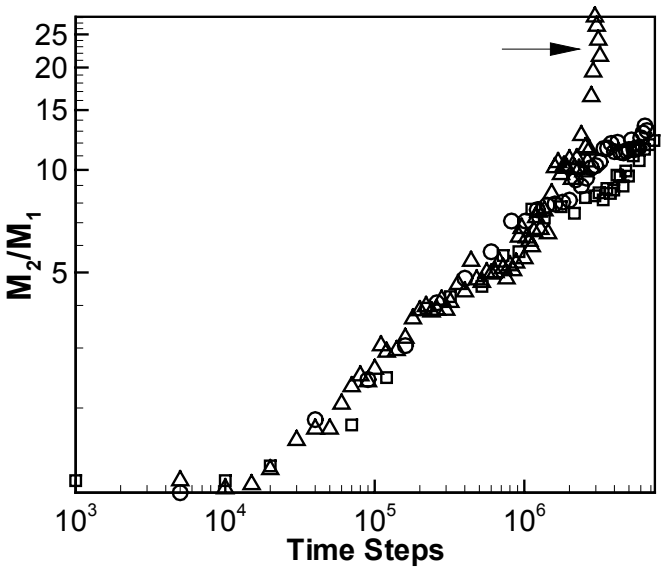

Figure 16: Prasad and Sinno, Journal of Chemical Physics

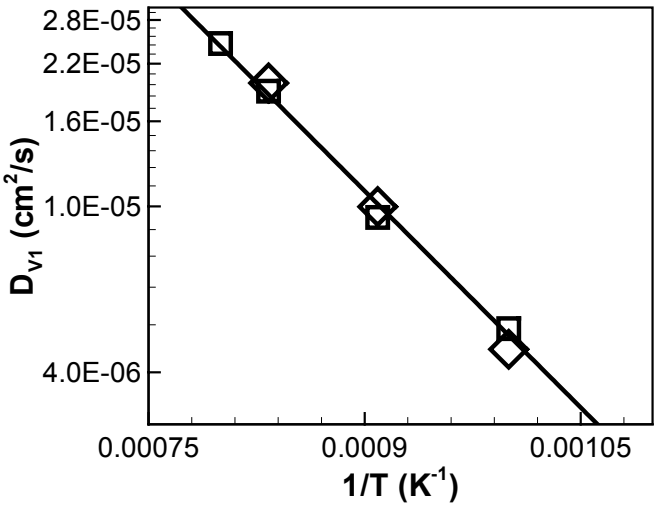

Figure 17: Prasad and Sinno, Journal of Chemical Physics

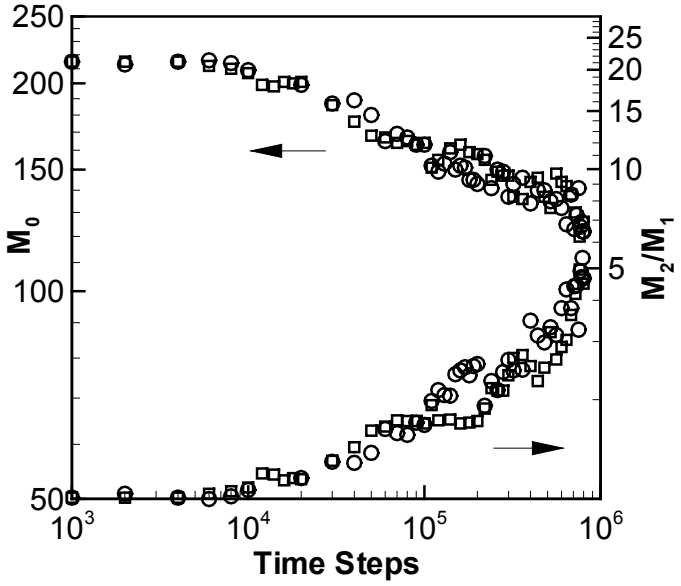

Figure 18: Prasad and Sinno, Journal of Chemical Physics