

MODELING DEFECT MEDIATED DOPANT DIFFUSION IN SILICON

by
Brian T. Puchala

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Materials Science and Engineering)
in The University of Michigan
2009

Doctoral Committee:

Associate Professor Michael L. Falk, Co-Chair
Associate Professor Krishnakamur R. Garikipati, Co-Chair
Professor Rachel S. Goldman
Assistant Professor Vikram Gavini
Assistant Professor Anton Van der Ven

© Brian T. Puchala
All Rights Reserved
2009

ACKNOWLEDGEMENTS

Throughout my time on this project I have been blessed with the support of many people. My path has been easy and joyful with the loving support of my family and friends. Working with my advisors, Professor Michael Falk and Professor Krishna Garikipati, has been a pleasure for which I am extremely grateful. I also want to thank the members of my Dissertation Committee, Professor Rachel S. Goldman, Professor Vikram Gavini, and Professor Anton Van der Ven for their suggestions and advice. In particular, I am thankful for many useful discussions with Professor Van der Ven.

The work presented in Chapter 3 has been published [1], copyright 2006 by Elsevier Ltd., and [2], copyright 2008 by the American Physical Society. Additionally, much of the work presented in Chapter 7 has been submitted for publication [3], and the American Physical Society retains the copyright upon acceptance.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	vi
LIST OF TABLES	xi
LIST OF ALGORITHMS	xii
LIST OF APPENDICES	xiii
ABSTRACT	xiv
CHAPTER	
1. Introduction	1
1.1 Diffusion in Semiconductors	2
1.2 Multi-Scale Modeling Framework	6
1.2.1 Atomistic parameterization of rare events	6
1.2.2 Modeling collective diffusion behavior	8
1.3 Research Goals	10
1.4 Outline	12
2. Atomistic Parameterization of Rare Events	13
2.1 Diffusion in Solids	13
2.1.1 Review of experimental results	18
2.1.1.1 Diffusion without stress	18
2.1.1.2 Diffusion with stress	23
2.1.2 Review of computational results	25
2.1.2.1 Silicon Vacancy	26
2.1.2.2 Silicon self-interstitial	31
2.1.2.3 Boron	35
2.1.2.4 Arsenic	37
2.1.2.5 Phosphorus and Antimony	40
2.2 Factors Affecting Atomistic Simulations	41
3. Accuracy of Supercell Calculations	44
3.1 Continuum Elasticity Calculations	44
3.1.1 Elastic dipole	45
3.1.2 Formation Energy	46
3.1.3 Formation Volume	48
3.2 Atomistic Calculations	52

3.2.1	Stillinger-Weber Potential	52
3.2.2	Methods	53
3.2.3	Measurements	58
3.2.4	Results	59
3.3	Conclusions	68
4.	Modeling Collective Diffusion Behavior	69
4.1	Calculating Diffusivity	69
4.2	The Master Equation	76
4.3	Kinetic Monte Carlo Simulations	77
4.4	Concentration and Time Dependence	78
4.5	Vacancy-Mediated Arsenic Diffusion in Silicon	82
4.6	Trapping Energy Basins in Kinetic Monte Carlo Simulations	84
5.	Accelerated Kinetic Monte Carlo Theory	89
5.1	Absorbing Markov Chain Analysis	90
5.1.1	First passage time analysis	91
5.1.2	Mean rate method	95
5.1.3	Equilibrating basins	96
5.2	Energy Basin Algorithm	98
5.2.1	Jump-first method	101
5.2.2	Look-first method	105
6.	Accelerated Kinetic Monte Carlo Algorithms	108
6.1	Algorithm Overview	108
6.1.1	Defect grouping	109
6.1.2	Jump-first method	110
6.1.3	Look-first method	114
6.2	Algorithm Details	116
6.2.1	Lattice Representation and Binning	117
6.2.2	Grouping	118
6.2.3	Jump rate parameterization	120
6.2.4	States and superstates	121
6.2.4.1	State saving	121
6.2.4.2	State checking	122
6.2.4.3	Superstates	123
6.2.5	Basins	123
6.2.5.1	Basin merging	124
6.2.5.2	Basin deleting	124
6.2.6	Clustering	124
6.2.6.1	Finding and updating clusters	125
6.2.6.2	Cluster connection saving	127
6.2.6.3	Basin cluster saving	128
6.2.7	Accelerated event determination	129
6.2.7.1	First passage time analysis	130
6.2.7.2	Mean rate method	132
6.2.7.3	Memory only	133
6.2.8	Updating group time	133
6.2.8.1	Memory only	134
6.2.8.2	First passage time analysis	134
6.2.8.3	Mean rate method	135

6.3	Results and Validation: Arsenic-Vacancy Cluster Dissolution	135
7.	Percolation Effects on Vacancy-Mediated Dopant Diffusion in Silicon . . .	144
7.1	Percolating Dopant Interactions	144
7.2	Parameterization Simulations	147
7.3	Standard Diffusion Simulations	149
7.4	Short Diffusion Simulations	155
7.5	Conclusions	156
8.	Conclusions and Future Work	159
	APPENDICES	164
	REFERENCES	180

LIST OF FIGURES

Figure

1.1	Cross section TEM of an Intel 65 nm PMOS transistor, with a 35 nm gate. Dopant implantation occurs in the source and drain, in this case SiGe. Dopant diffusion into the strained silicon channel must be precisely controlled. From Tyagi [6].	3
1.2	Schematic of vacancy and interstitial formation from perfect crystals.	4
1.3	Schematic of transitions between stable states for a vacancy and an interstitial.	5
1.4	The diamond cubic crystal structure of silicon. Figure from [15]	9
1.5	Schematic of a trapping energy basin. The low energy barrier between states 2 and 3 compared to the barriers out of the basin results in many transitions between those states until the system escapes to states 1 or 4.	11
2.1	A one dimensional illustration of a rare event in which the system equilibrates in the first state (1) before crossing a transition state (2) into the final state (3).	16
2.2	Schematics showing the calculation of formation energies and volumes from a supercell. The formation calculation is shown for an (a) isolated Si vacancy and (b) interstitial, and for a dopant-defect complex in (c) and (d). In this case, the reference state is the perfect crystal with dopant A in a substitutional position. $Q = E$ or \bar{V}	26
2.3	The three lowest energy Si self-interstitial defects are the (a) $\langle 110 \rangle$ dumbbell, (b) hexagonal, and (c) tetrahedral interstitials.	31
3.1	Schematic of an elastic dipole. Here the forces, F , are all the same magnitude, separated by the same distance, d , and aligned with the orthonormal basis, e , so it is an isotropic dipole.	45
3.2	Difference between the strain energy of vacancy formation in an infinite body, \mathbb{U}_{∞}^f , and strain energies of vacancy formation in a finite body, $\mathbb{U}_{(\cdot),B_0}^f$, using either periodic (PB) or free (FB) boundary conditions.	49
3.3	A vacancy in silicon viewed along the $\langle 100 \rangle$ direction. The open circles mark positions of atoms in the perfect crystal; i.e., before vacancy formation. Displacements have been scaled 10x for clarity.	54
3.4	Structure of the relaxed Si $\langle 110 \rangle$ dumbbell interstitial with displacements scaled by 3x for clarity. In configuration (A), the dipole tilts and breaks the symmetry about the (110) plane, resulting in a slightly lower energy than in configuration (B). Both (A) and (B) maintain symmetry about the $(1\bar{1}0)$ plane.	55

3.5	The (a) formation energy and (b) trace of the formation volume converge in the large-size limit for all boundary conditions. “PBC CV uncorrected” and “PBC CV” indicate formation energies before and after using Equation 3.11.	60
3.6	The (a) formation energy and (b) trace of the formation volume converge in the large-size limit for all boundary conditions. “PBC CV uncorrected” and “PBC CV” indicate formation energies before and after using Equation 3.11.	61
3.7	The $\langle 110 \rangle$ interstitial (A) formation energy converges in the large-size limit to $E_{SiI, \langle 110 \rangle, (A)}^{f, \infty} = 4.7091$ eV for all boundary conditions. “PBC CV uncorrected” and “PBC CV” indicate formation energies before and after using Equation 3.11.	62
3.8	The convergence of $\text{tr}[V^f]$ for the $\langle 110 \rangle$ interstitial (A) depends on the convergence of the (b) bulk modulus ($K_{inf} = 1.0826 \times 10^{11}$ Pa) and (c) shear modulus ($C_{44, inf} = 6.0256 \times 10^{10}$ Pa).	65
3.9	The formation volume converges to $\text{tr}[V_{SiI, \langle 110 \rangle, (A)}^{f, \infty}] = 13.677 \text{ \AA}^3$ with error decreasing as $1/N$ shown above with \circ . Knowing this convergence rate, we can use Equation 3.34 to estimate a converged value from the two largest systems calculated. This estimate, which is shown with \times , is as converged at $N = 1728$ as the $N = 21,952$ calculation.	66
4.1	Random jumping on a square lattice. Figure as in [18].	70
4.2	Literature results for As diffusivity as a function of dopant concentration. Lines for the Solmi and Nobili data indicate the trend where many data points are reported.	81
4.3	KMC parameterization of the KMC model for vacancy-mediated As diffusion in Si. The binding and migration energies (row \rightarrow column), listed in (a) are from <i>ab initio</i> calculations [85, 45, 65]. The lattice is shown in (b); 3^rNN indicates a position 3NN around a hexagonal ring, and 3^cNN is a position 3NN along a $\{011\}$ chain. The state connectivity is shown in (c), with gray states being saddle states that lead a different V_1As_1 energy basin. The energy landscape is shown for a V moving from a 1NN position with As to the 3^rNN position in (d) (the dotted line indicates the path to a 1NN position by a different direction), and to the 3^cNN position in (e).	83
4.4	Saddle point energy calculation approximation.	84
4.5	Diffusion of V_1As_1 clusters occurs by the ring mechanism. The state and saddle point energies are shown above, and schematics of the cluster configurations below. “E” indicates a pathway for the vacancy to escape the As. Small arrows with “ xn ” indicate that n symmetrically equivalent transitions are possible in the direction of the arrow.	85
4.6	States near the minimum energy V_1As_2 configuration. The state and saddle point energies are shown above, and schematics of the cluster configurations below. “E” indicates a pathway for the vacancy to escape the As, “Rearrange” indicates a transition leading towards a different low energy V_1As_2 configuration. Transitions shown with single arrows indicate that one state is unstable towards the other. Small arrows with “ xn ” indicate that n symmetrically equivalent transitions are possible in the direction of the arrow.	87

4.7	States involved in rotation of vacancy and As around a hexagonal ring in the V_1As_2 cluster. The state and saddle point energies are shown above, and schematics of the cluster configurations below. “E” indicates a pathway for the vacancy to escape the As, “Rearrange” indicates a transition leading towards a different low energy V_1As_2 configuration, “A” and “B” indicates states of the type shown to the left. Transitions shown with single arrows indicate that one state is unstable towards the other. Small arrows with “ xn ” indicate that n symmetrically equivalent transitions are possible in the direction of the arrow.	88
5.1	First passage time analysis for a basin with three states. The figures on top show the state connectivity and energies ($k_B T$). For the case on the right (Case R), the arrows indicate how the exit time is chosen from the occupation probability plot and the exit state is chosen from the exit probability plot.	94
5.2	A “basic energy basin” is a collection of states in which there is a minimum energy saddle point and in moving from the minimum to the other states the saddle point energy does not decrease. Basic energy basins, like Basin 1 and 2, can be merged into combined basins, in which the saddle point energies will no longer be strictly non-decreasing moving away from the minimum.	99
5.3	The jump-first method for a V_1As_1 cluster at 800 °C. At each step, the current state of the system is indicated by a dot. The state connectivity graph is as in Figure 4.3. The gray states indicate different basins. The # ↓ or # ↑ indicate the current state and whether the system is heading down or up the basin. The probability for exiting to different types of states, at the chosen timestep ($\Delta t \nu^{-1}$ s), is listed on the right, and the chosen type of state indicated with a ✓.	102
5.4	The look-first method for a V_1As_1 cluster at 800 °C. The state connectivity graph is as in Figure 4.3. At each step, the current state of the system is indicated by a dot. While exploring a basin, the original state is indicated with a circle. The gray states indicate different basins. The # ↓ or # ↑ indicate the current state and whether the system is heading down or up the basin. A ■ indicates the algorithm has determined that the system left the current basin. The transition energy, E , or probability, P , for exiting to different types of states, at the chosen timestep ($\Delta t \nu^{-1}$ s), is listed on the right, and the chosen type of state indicated with a ✓.	106
6.1	Acceleration achieved during V_1As_1 dissolution. The basin identifying state-saving methods, look-first and jump-first, automatically determine the number of saved states, so they are simply presented as straight lines for comparison with the n -methods. The look-first and jump-first lines are the mean of 10^5 samples. Each data point for the n -methods is the mean of 10^4 samples. Acceleration is relative to the mean of 10^5 unaccelerated samples.	139
6.2	Timing and KMC event histograms for comparison for simulations of V_1As_1 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. Vertical lines in the t^{KMC} plot indicate means, and all fall on top of each other. Magenta symbols in the “# KMC Events” plot are the number of unaccelerated events that would have occurred in the memory only simulations.	140
6.3	Histograms of vacancy and As displacement upon V_1As_1 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first.	141

6.4	Acceleration achieved during V_1As_2 dissolution. The look-first and jump-first lines are the mean of 10^5 samples, and each data point for the n -methods is the mean of 10^4 samples. Acceleration is relative to the mean of 10^5 unaccelerated samples. .	142
6.5	Acceleration achieved during V_1As_3 dissolution. The look-first and jump-first lines are the mean of 10^5 samples, except for memory only which is the mean of 4×10^3 samples. Each data point for the n -methods is the mean of 10^3 samples. The n -method with memory only was too slow to be done. Acceleration is relative to the mean of 5 unaccelerated samples.	143
7.1	The effect of nearby dopants, P, on V formation and migration energies. From [114].	145
7.2	The fraction of time spent in various clusters as a function of simulated time during “Parameterization” simulations with (a) one to (d) four As. The solid lines are simulation data and the dashed green lines is the mean field model fit.	148
7.3	The fraction of time spent by the V in various clusters as a function of simulated time during “Standard” simulations. Thin green lines are mean field model predictions.	151
7.4	As and V diffusivities in simulations with one V ($2 \times 10^{17} \text{ cm}^{-3}$) and varying As concentration. Thin dashed lines are mean field model predictions. Dotted lines indicate the times at which lines A-C are plotted in Figure 7.5.	152
7.5	As diffusivity as a function of As concentration for the three times indicated in Figure 7.4. The As diffusivity is scaled to include the effect of a V concentration increasing with the square of the As concentration.	153
7.6	Histograms showing the time, V r^2 displacement, and maximum saddle point traversed as recorded for 10^5 samples at varying As concentration from when a V is first 1NN with one As until it is 1NN with two or more As. Contour lines are at intervals of 10% of the peak in the top plot. A line of constant diffusivity is shown for reference.	157
C.1	Timing and KMC event histograms for comparison of simulations of V_1As_2 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. Vertical lines in the t^{KMC} plot indicate means, and all fall on top of each other. Magenta symbols in the “# KMC Events” plot are the number of unaccelerated events that would have occurred in the memory only simulations.	176
C.2	Histograms of vacancy and As displacement upon V_1As_2 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. The additional maxima in vacancy displacement are due to vacancies which cross the periodic boundaries before the As are completely dissolved.	177

- C.3 Timing and KMC event histograms for comparison of simulations of V_1As_3 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. Vertical lines in the t^{KMC} plot indicate means, and all fall on top of each other. Magenta symbols in the “# KMC Events” plot are the number of unaccelerated events that would have occurred in the memory only simulations. There is a minimum possible probability for memory only because there are only 4×10^3 samples. The unaccelerated results are estimated from the number of explicit jumps in the memory only simulations. . 178
- C.4 Histograms of vacancy and As displacement upon V_1As_3 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. There is a minimum possible probability for memory only because there are only 4×10^3 samples. The additional maxima in vacancy displacement are due to vacancies which cross the periodic boundaries before the As are completely dissolved. 179

LIST OF TABLES

Table

2.1	<i>Ab initio</i> results for a neutral, isolated Si vacancy. Unless noted, the vacancy has undergone the Jahn-Teller distortion.	28
2.2	<i>Ab initio</i> results for neutral, isolated $\langle 110 \rangle$, hexagonal, and tetrahedral Si interstitials.	32
4.1	Transition probabilities (row \rightarrow column) for a V near As at $T = 800$ °C. The \times indicate the number of different pathways available.	86
7.1	Cluster reaction rate and diffusivity results from the “Parameterization” simulations.	150

LIST OF ALGORITHMS

Algorithm

6.1	A simplified outline of a KMC program using the jump-first method.	111
6.2	The criteria for determining if the system is still in the current basin. Known (K) and unknown (UK) transitions are from the current state only.	112
6.3	The conditional setting of β upon creating a new basin. Known (K) and unknown (UK) transitions are from the current state only.	113
6.4	A simplified outline of a KMC program using the look-first method.	114
6.5	The algorithm for identifying basins in the look-first method.	115
A.1	A detailed outline of a KMC program using the jump-first method.	166
A.2	Save a superstate in the jump-first method.	167
A.3	The conditional setting of β upon creating a new basin. Known (K) and unknown (UK) transitions are from the current state only.	167
A.4	Determining group events for either the jump-first or look-first method.	168
A.5	Perform a V move in the jump-first method.	168
A.6	The criteria for determining if the system is still in the current basin. Known (K) and unknown (UK) transitions are from the current state only.	168
A.7	Updating the basin in the jump-first method.	169
B.1	A detailed outline of a KMC program using the jump-first method.	171
B.2	The algorithm for identifying basins in the look-first method.	172
B.3	Save a superstate in the look-first method.	173
B.4	Determining group events for either the jump-first or look-first method.	173
B.5	Updating the basins in the look-first method.	174
B.6	Perform a V move in the look-first method.	174

LIST OF APPENDICES

Appendix

A.	The Jump-First Method	165
B.	The Look-First Method	170
C.	V_1As_2 and V_1As_3 Dissolution Figures	175

ABSTRACT

MODELING DEFECT MEDIATED DOPANT DIFFUSION IN SILICON

by

Brian T. Puchala

Co-Chairs: Michael L. Falk and Krishnakamur R. Garikipati

The current understanding of dopant diffusion in silicon comes from the synthesis of experimental and computational research. Dopant diffusion is mediated by defects, and the relevant physical phenomena range over many time and length scales, necessitating a multi-scale modeling approach. In this work, we focus on two essential aspects, (1) the accuracy of atomistic methods for calculating defect parameters, and (2) an accelerated kinetic Monte Carlo (KMC) method, which we use to investigate the effects of percolating dopant-defect interactions on diffusion.

We use continuum linear elasticity to quantify the effects of boundary conditions on atomistic calculations of defect energies and volume tensors. It predicts that when using periodic boundary conditions with zero average stress, energies converge with the inverse of system size and relaxation volume tensors are independent of supercell size or symmetry. We verify the linear elastic prediction in the far field of atomistic calculations by calculating the formation energy and volume tensor for vacancy and interstitial defects in silicon using the Stillinger-Weber empirical potential. In practice, both defect energies and relaxation volume tensors converge

with the inverse of system size because changes in the bonding at the defect affect the elastic moduli.

We also introduce an accelerated KMC method which automatically determines which states comprise trapping energy basins, allowing simulations to reach very long times compared to standard KMC simulations. We validate the accelerated method by performing simulations of V-As cluster dissolution and comparing to standard KMC simulations. Then we apply the method to highly time and concentration dependent vacancy-mediated As diffusion in Si. At high As concentrations, percolating dopant interactions lead to limited increased diffusivity, but the effect is limited in magnitude and duration as immobile clusters form quickly. The energy basin algorithms for accelerating KMC simulations may be very useful in a wide variety of applications. By considering issues such as grouping isolated diffusing species and collecting data when the exact location of the system within an energy basin is not resolved, we provide an example that can be followed when applying this method to other systems.

CHAPTER 1

Introduction

Understanding diffusion in crystalline materials is critical in a wide range of technologically important systems. Lithium ion battery improvements, failure prevention of irradiated materials in nuclear reactors, and semiconductor device fabrication all depend on accurate prediction of diffusion. Besides its technological importance, modeling diffusion is an interesting and challenging scientific problem. Accurate diffusion models must account for physical phenomena on time scales ranging from picoseconds for thermal vibrations to years of part operation. Similarly, length scales range from the interatomic spacing for individual diffusive jumps to nanometers and microns for semiconductor devices or centimeters for irradiated structural materials. This necessitates a multi-scale modeling approach in which models of diffusion mechanisms at small time or length scales are used to parameterize diffusion models at larger time and length scales. In this work, we focus on two aspects of multi-scale computational modeling applied to dopant diffusion in silicon.

The current understanding of diffusion in semiconductors has come from a combination of experimental and computational work. Because of the atomic nature of the underlying mechanisms, experimental observations are often difficult or indirect, and valuable contributions have come from direct atomistic computations. This chapter

will provide a broad overview of dopant diffusion and the computational methods used to model it. Later chapters will give a more detailed review of the experimental and computational work that has led to this understanding.

1.1 Diffusion in Semiconductors

Diffusion is of critical importance in the fabrication of semiconductor devices such as the transistors used in computer chips. An example of a 65 nm Intel transistor with a strained silicon channel is shown in Figure 1.1. Device performance depends on precisely controlling the location of impurity atoms, called dopants, implanted in otherwise pure crystalline silicon. Atomic collisions during the implantation process create point defects in the crystal lattice. Silicon or dopant atoms ending up off-lattice between other atoms are called interstitials, and the empty spaces they leave behind are called vacancies, see Figure 1.2. Crystal defects adversely affect electronic conductivity and must be repaired by annealing at temperatures high enough to permit significant atomic diffusion. Annealing allows vacancy and interstitial diffusion so that they combine and annihilate or diffuse to sinks such as surfaces, interfaces, or dislocations, but annealing also causes dopant diffusion. Accurate modeling of dopant diffusion is important to guide device fabrication, especially as transistor dimensions have shrunk to the nanometer scale and high doping concentrations with very abrupt doping profiles are required to keep electrical resistances low [4]. A thorough review of dopant diffusion in silicon has been done by Fahey, Griffin, and Plummer [5].

Atomic motion is driven by the kinetic energy of thermal vibrations and directed by the forces between atoms. Diffusive motion on a crystal lattice involves the transition of the system from one stable atomic configuration across a saddle point in

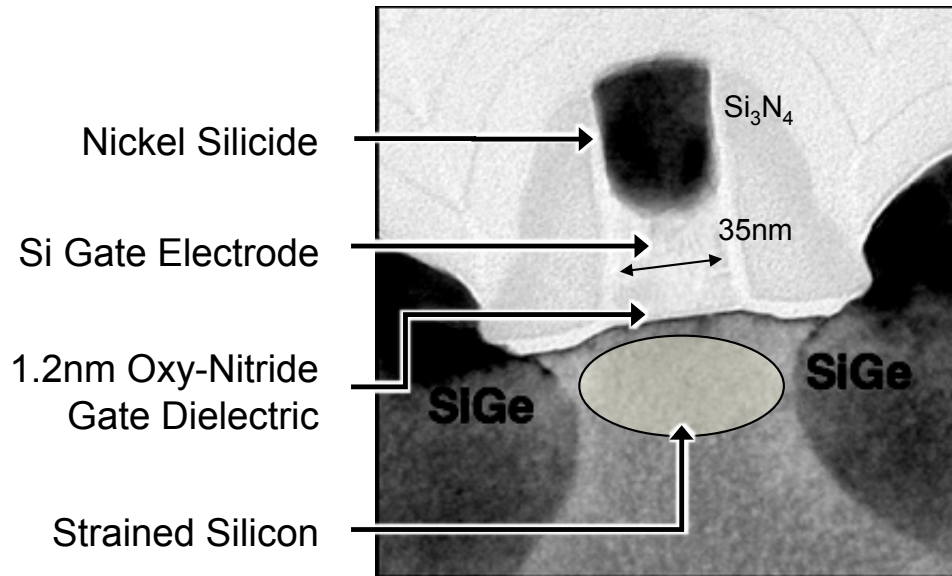


Figure 1.1: Cross section TEM of an Intel 65 nm PMOS transistor, with a 35 nm gate. Dopant implantation occurs in the source and drain, in this case SiGe. Dopant diffusion into the strained silicon channel must be precisely controlled. From Tyagi [6].

the potential energy landscape into a different stable atomic configuration, as shown in Figure 1.3. The individual transition rates are a function of the temperature and the potential energy landscape, which is affected by the atomic configuration, stress, and electronic charge.

Generally in crystal lattices, and particularly in silicon, the energy required for neighboring atoms to swap places is quite large compared to the energy for migration of a pre-existing crystal defect. Therefore, at temperatures well below the melting point most atomic diffusion is defect mediated. In silicon, vacancies and self-interstitials are quite mobile, and tend to cluster with themselves and with dopants due to electronic and stress mediated interactions. Dopants diffuse by jumping into vacancies, or by being knocked off-lattice by an interstitial and then moving between lattice locations until a new lattice position can be occupied, as in Figure 1.3. In this manner, complexes of clustered defects and dopants may diffuse significant distances through the material before breaking up. During complex formation, diffusion, and

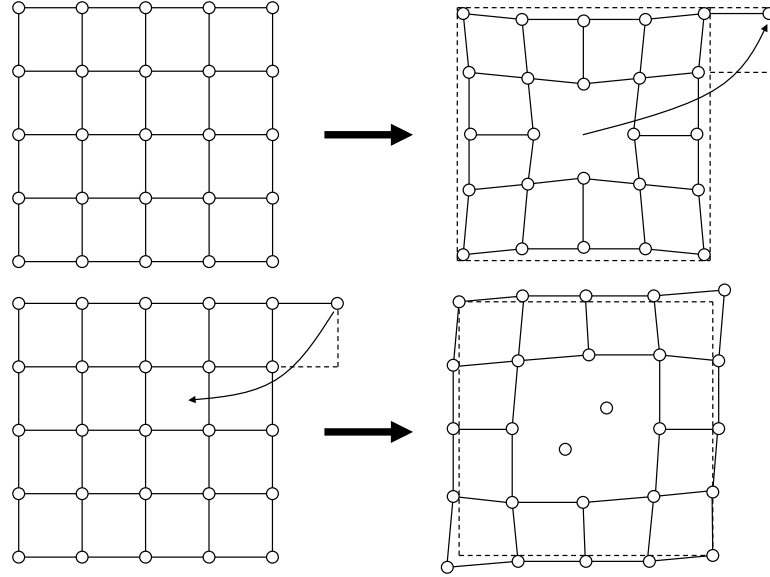


Figure 1.2: Schematic of vacancy and interstitial formation from perfect crystals.

breakup, there are many different possible atomic configurations and transition pathways.

Stress affects diffusion by modifying the free energy for defect formation and migration. When a defect forms or migrates in a crystal, the atoms in the surrounding lattice will generally lower their energy by relaxing inwards or outwards, as in Figures 1.2 and 1.3. The resulting volume change does work in a stress field and must be included in the formation or migration energy of the defect. Stress effects on diffusion become increasingly important in devices with features at very small scales because stress gradients may be higher and dopant concentration profiles must be very precise. Significant and complex stress states can arise from strain engineering, lattice and thermal expansion coefficient mismatch, growth stresses, and defect concentrations [7]. Experiments showing the different effects of hydrostatic and biaxial stress states on dopant diffusivity indicate that a full tensorial formulation for the volume change is necessary [8].

By trapping and releasing electrons or holes, defects and dopants and their com-

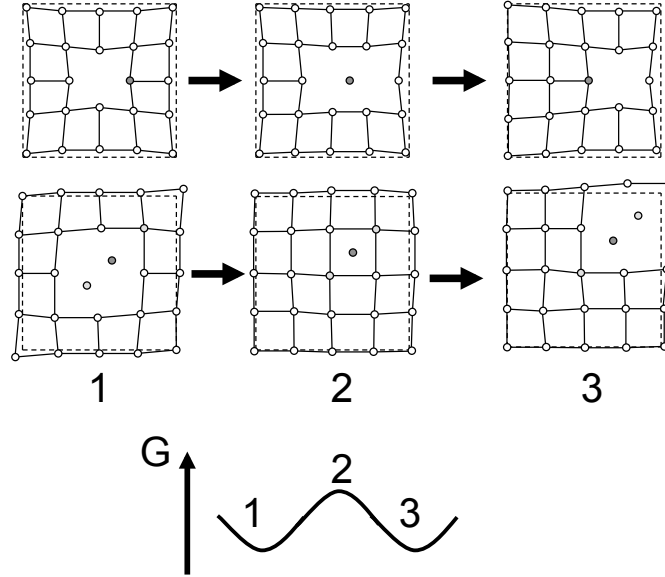


Figure 1.3: Schematic of transitions between stable states for a vacancy and an interstitial.

plexes may take different electronic charge states. This process is generally considered faster than atomic diffusion, so that charge state among the collective defects will equilibrate based on the Fermi level [5]. The electronic charge states can have several consequences. Coulombic interactions may affect defect and dopant clustering. The total carrier concentration, and thus device performance, may be affected. Different atomic structures may be stable at different charge states. The equilibrium quantity of defects and their jump rates may also vary.

All of these effects on individual jump rates modify the complex diffusion behavior that occurs when populations of defects and dopants interact. The tendency of defects and dopants to cluster results in time dependent diffusivity because larger clusters tend to diffuse more slowly than small ones. The cluster diffusion rates are modified by stress, and anisotropy of the individual jump rates caused by anisotropic stresses may result in anisotropic cluster diffusion. Changes in dopant concentration also affect the rate of cluster formation. At very high dopant concentrations, dopant-defect interactions may percolate throughout the material in the sense that a defect

might travel across the full length of the material by a path in which it is always close enough to a dopant to have significant interactions. This makes it impossible to identify distinct clusters and may lead to different diffusion behavior at high dopant concentrations than low concentrations.

1.2 Multi-Scale Modeling Framework

The physical phenomena that must be accounted for in dopant diffusion models range over many time and length scales, but we can roughly identify three scales. At the smallest scales, individual or collective atomic jumps contribute to diffusion. At intermediate scales, species interact with each other forming clusters, diffusing, and breaking up. At the largest scales, average properties can describe the motion and interactions of collections of diffusing species. In a multi-scale modeling approach, the physics at the small scales is parameterized and used in the larger scale models. Accuracy at each level of the model relies on accurate parameterization from the previous level, which can come from experimental observations or computational analysis. Our work has focused on aspects of the computational methodology at the smallest and intermediate scales.

1.2.1 Atomistic parameterization of rare events

Diffusive motion on a crystal lattice generally involves the transition of the system from one stable atomic configuration across a saddle point in the potential energy landscape to a different stable atomic configuration. Often, the energy difference between the stable and saddle states is significantly greater than the average energy available from thermal vibrations. In these cases, the diffusive event can be considered a rare event because thermal fluctuations will provide the requisite energy

infrequently compared to the time scale of thermal vibrations.

Transition state theory, discussed in Section 2.1, allows calculation of jump rates from knowledge of the free energies in the stable and saddle states. The jump rates include the effects of atomic vibrations, so if all of the relevant events can be identified and parameterized the vibrations do not need to be explicitly included in higher level computations. Then the computational work is greatly reduced and much longer times can be reached while reproducing the exact time evolution of the system. The modeling problem is then one of identifying the active (low-energy) transition pathways and calculating the free energies of the stable and saddle states. The energies of the stable states are important for determining the thermodynamic properties of the system, and the differences between the stable and saddle states are important for determining the kinetic properties of the system.

Atomistic calculations begin with approximation of the interaction potential used to calculate potential energy, either using *ab initio* or empirical methods. *Ab initio* methods use quantum mechanics to calculate electron distributions, so they allow calculation of electronic properties and are more accurate, but at much more computational expense. *Ab initio* calculations are usually limited to tens or hundreds of atoms. Empirical methods approximate bonding energies and structures by fitting to experimental observations, and may be less accurate. They are much less computationally expensive, allowing for calculations in systems up to millions of atoms.

The molecular dynamics (MD) method uses the interaction potential to calculate the forces on individual atoms or molecules and uses Newton's equations to directly simulate motion. For the purpose of identifying all active diffusion pathways, MD has the advantage that no restrictions or assumptions are made. It must resolve atomic vibrations, typically at a time scale of 10^{-13} s, so it can only simulate nanoseconds

of physical time even when using empirical potentials. This makes observing and measuring rare events difficult, and often impossible if the accuracy of *ab initio* methods is desired. There exist some MD methods that accelerate the occurrence of rare events in a controlled way so that the normal dynamics can be determined, but they are generally limited to empirical potentials [9].

Most often *ab initio* calculations use the molecular statics (MS) method to calculate the potential energy of atomic configurations at 0 K. Energy minimization techniques are used to adjust the atomic positions until the energies of stable states and saddle states are found. If all transitions are accounted for and the free energy barriers are known, transition state theory can accurately reproduce the system dynamics. The main drawback of MS is that exhaustive searches through configuration space are impossible, so that unexpected important transition pathways may be missed. Another drawback is that entropic effects, included in MD calculations by their nature, must be determined explicitly in MS calculations.

1.2.2 Modeling collective diffusion behavior

Once the individual jump rates have been parameterized, the modeling problem becomes one of calculating the diffusion behavior of collections of interacting species. Dopant diffusion is defect mediated, so it depends on the concentrations defects and dopants, and the rate at which defect-dopant complexes form, diffuse, and dissociate. These things can all be affected by stress in complex ways depending on the anisotropy of the stress field and the lattice on which the species are diffusing. Silicon has a diamond cubic crystal lattice, as depicted in Figure 1.4. Some simple cases, such as vacancy or interstitial diffusion on cubic lattices in the dilute limit, can be modeled as random walk processes [10]. Other approaches are used for diffusion on

anisotropic lattices or diffusion of complexes under the effect of anisotropic stress [11, 12, 13, 14].

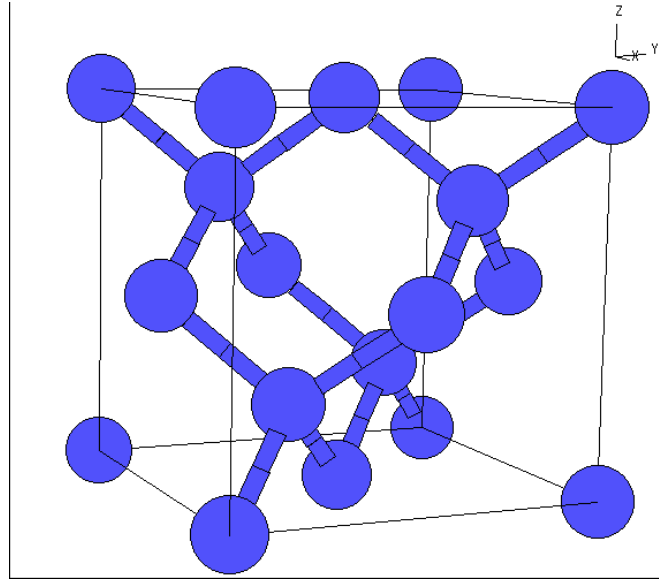


Figure 1.4: The diamond cubic crystal structure of silicon. Figure from [15]

But the situation quickly becomes complicated when anisotropic stress or the presence of a dopant or another defect makes some jumps faster than others. Analytical diffusivity calculations are also difficult for interstitials since they may have several different low energy configurations, with different jump rates, off of the diamond cubic lattice. And in clusters of only a couple defects or dopants there are many configurations possible with complex transitions pathways between them. Especially for highly non-equilibrium processes or at high concentrations when correlations between the diffusing species are important, it may be impossible to exactly calculate the system's evolution from the microscopic jump rates.

In these situations the kinetic Monte Carlo (KMC) method becomes essential. The KMC method simulates the evolution of a system from state to state by numerically sampling the master equation which describes the time evolution of the system through its possible states [16, 17]. For atomic diffusion processes, this usually means

the system evolves one atomic hop at a time and the trajectories of diffusing species can be directly observed. Sometimes, direct simulation is used to measure concentration evolution or dopant clustering. At other times, the KMC results may be averaged and used to parameterize diffusivities or clustering rates for continuum models.

1.3 Research Goals

As part of the effort to develop the multi-scale methodology for modeling dopant diffusion including stress effects, the work in this dissertation focuses on two essential aspects, (1) the accuracy of atomistic methods for calculating defect parameters, and (2) an accelerated KMC method.

The overall accuracy of a multi-scale diffusion model depends on the accurate parameterization of microscopic jump rates. In purely atomistic calculations, a defect is introduced into material with either free or periodic boundary conditions (FBC or PBC). The atomistic system is allowed to relax and after mechanical equilibrium is reached the resulting energy and volume change are measured. These are the defect parameters that determine defect jump rates and equilibrium concentrations including the effects of stress. For the finite sizes currently tractable, calculations using FBC can have significant finite-size and surface effects. With PBC, we consider the computational cell surrounded by an identical cell in the direction of the periodic boundary. Effectively, the system is infinitely repeated and atoms on one side of the cell interact with atoms on the other side. We use the term “supercell” for both the infinite periodic system and the finite computational cell which is repeated. Calculations using PBC in three dimensions do not have finite-size and surface effects, but instead the periodic supercell creates an infinite array of defects which may interact

with each other and affect calculation accuracy. Our work is the first to formally show the elastic effect of boundary conditions on the accuracy of typical atomistic calculations. We model the point defect as a center of expansion or contraction via a force dipole in a finite elastic body and use continuum linear elasticity to quantify the effects of periodic images, supercell size, and symmetry. We verify the linear elastic prediction in the far field of atomistic calculations by calculating the formation energy and volume tensor for vacancy and interstitial defects in silicon using the Stillinger-Weber empirical potential. An empirical potential is useful for studying elastic effects because there are no electrostatic interactions and it allows us to simulate large systems approaching the far field.

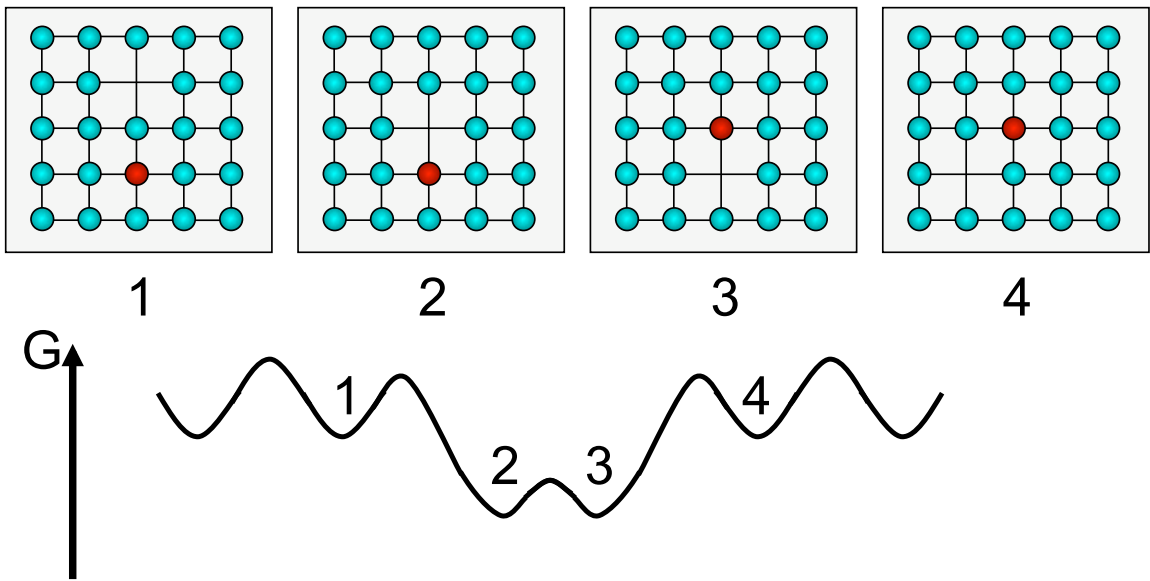


Figure 1.5: Schematic of a trapping energy basin. The low energy barrier between states 2 and 3 compared to the barriers out of the basin results in many transitions between those states until the system escapes to states 1 or 4.

Defect and dopant clustering is a fundamental cause of concentration and time dependent dopant diffusion. The complexity of the atomic picture due to the number of possible configurations during clustering or high concentration diffusion make KMC simulations an essential tool for modeling the collective diffusion behavior.

However, KMC simulations become very inefficient in systems in which there are trapping energy basins, such as the one shown in Figure 1.5. In a trapping energy basin, many fast transitions will occur between the states separated by low energy barriers before the system escapes from the basin over one of the larger barriers. Standard KMC simulations evolve the system state by state so they become trapped inside the basin and may spend many millions of events simply transitioning back and forth inside the trap. This situation is common in clusters of defects and dopants. Significant simulation acceleration is achieved by solving the master equation over the limited number of states comprising the trap and sampling that solution to determine when and where the system escapes. This procedure can be done without any approximation. We have developed an accelerated KMC algorithm which automatically determines which states comprise trapping basins and should be included in the acceleration calculation. We have used this algorithm to model vacancy-mediated arsenic diffusion in silicon, and study the effect of high dopant concentration on diffusivity.

1.4 Outline

In Chapters 2 and 3 we consider the accuracy of atomistic methods for calculating diffusive jump rates. In Chapters 4-7 we present the accelerated technique for simulating defect and dopant diffusion using kinetic Monte Carlo (KMC) and apply it to high concentration vacancy-mediated arsenic diffusion in silicon.

CHAPTER 2

Atomistic Parameterization of Rare Events

Diffusion models depend on the accurate parameterization of individual or collective atomic jumps. In this chapter we review diffusion in solids and transition state theory, which allows calculation of diffusive jump rates from the energetics of the system. Then we will review the experimental and computational work that has been used to determine those parameters. Finally, we will consider some of the factors that affect the accuracy of atomistic calculations of diffusion parameters.

2.1 Diffusion in Solids

As noted in Chapter 1, dopant diffusion is mediated by defects. The equilibrium concentration of defects is

$$(2.1) \quad C^e = C^S \exp(-G^f/k_B T),$$

where C^S is the concentration of sites on which the defect can form, G^f is the Gibbs free energy of formation, k_B is the Boltzmann constant, and T is temperature. The Gibbs free energy of formation is the change in free energy upon the formation of a defect, as in Figure 1.2, and can be written

$$(2.2) \quad G^f = E^f - \bar{\sigma} : \bar{V}^f - TS^f + q\varepsilon_F,$$

where E^f , \bar{V}^f , and S^f denote the energy, volume, and entropy of migration, respectively, q is the electronic charge of the defect and ε_F is the Fermi level. The formation volume is a second order tensor quantity that gives the dependence of the Gibbs free energy on the stress state and is defined by

$$(2.3) \quad V^f \equiv -\frac{\partial G^f}{\partial \bar{\sigma}}.$$

The tensor inner product, “:”, is defined:

$$(2.4) \quad \bar{\sigma} : \bar{V}^f = \sum_{i,j} \sigma_{ij} V_{ij}^f.$$

However, we are often interested in systems and processes in which the concentration of defects is not at equilibrium. In these situations it is more likely that there is a local equilibrium between the concentrations of dopants, defects and dopant-defect complexes. When a complex forms, the Gibbs free energy of binding is

$$(2.5) \quad G_{AX}^b = G_X^f - G_{AX}^f,$$

where A indicates a dopant, X a defect, and AX a dopant-defect complex. Then, in local equilibrium, the concentration of dopant-defect complexes is [5]

$$(2.6) \quad C_{AX} = m_{AX} \frac{C_A C_X}{C_S} \exp\left(\frac{-G_{AX}^b}{k_B T}\right),$$

where C_A is the concentration of isolated dopants, C_X is the concentration of unbound defects, and m_{AX} is the number of orientations of a complex at a single site. For example, $m_{AX} = 4$ for substitutional dopants and vacancies in Si since for each dopant the V could be in one of four nearest neighbor positions.

At a macroscopic level, gradients in chemical potential lead to diffusion. For component i , the chemical potential, μ_i , is the change in Gibbs free energy, G , with the number of moles n_i of component i :

$$(2.7) \quad dG = -SdT + Vdp + \sum_i \mu_i dn_i,$$

where S is entropy, V volume, and p pressure. At equilibrium $dG = 0$, so if there is any spatial variation in the chemical potential the system is not in equilibrium and there will be atomic drift. The drift velocity, v_i , is proportional to the chemical potential gradient,

$$(2.8) \quad \bar{v}_i = \bar{M}_i \bar{\nabla} \mu_i,$$

where \bar{M}_i is the mobility. The atomic flux vector, \bar{J}_i is the product of the concentration, C_i , and the drift velocity,

$$(2.9) \quad \bar{J}_i = C_i \bar{v}_i$$

$$(2.10) \quad = C_i \bar{M}_i \bar{\nabla} \mu_i.$$

Experimentally, chemical potential is not directly observable, but concentrations are and diffusion is described by Fick's Laws. The first,

$$(2.11) \quad \bar{J}_i = -\bar{D}_i^C \bar{\nabla} C_i,$$

relates the atomic flux vector to the concentration gradient through the chemical diffusivity tensor \bar{D}_i^C . The mobility can be related to the diffusivity by [18]

$$(2.12) \quad \bar{D}_i^C = \bar{M}_i kT \left(1 + \frac{d \ln \gamma_i}{d \ln C_i} \right).$$

For ideal or dilute systems the activity coefficient, γ_i , is constant and the bracketed “thermodynamic factor” is unity, giving

$$(2.13) \quad \bar{D}_i^C = \bar{M}_i kT.$$

As we will show in Section 4.1, if diffusion is isotropic and dominated by a single mechanism, jump distances are constant, α , and the system is thermodynamically ideal or concentrations dilute, then the diffusivity can be calculated as

$$(2.14) \quad D_i^C = fg\alpha^2\Gamma_i,$$

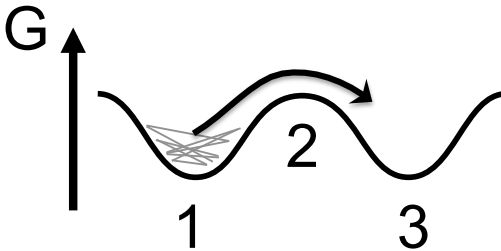


Figure 2.1: A one dimensional illustration of a rare event in which the system equilibrates in the first state (1) before crossing a transition state (2) into the final state (3).

where f is a correlation factor, g is a geometric factor which is $\frac{1}{6}$ for three-dimensional cubic lattices, and Γ_i is the total jump rate for particles of component i . If every particle of component i has z possible jumps with rate R_i , then $\Gamma_i = zR_i$. If a substitutional dopant does not have a defect bound to it, it will not be able to jump, therefore

$$(2.15) \quad \Gamma_A = z \frac{C_{AX}}{C_A^{Tot}} R_{AX},$$

where $C_A^{Tot} = C_{AX} + C_A$ is the total concentration of dopant A.

The basis for parameterizing event rates, R , is transition state theory [9]. Transition state theory treats situations in which the temperature and energy landscape are such that the system will spend a long enough time in one state that it loses any memory of where it came from before transitioning to the next state. Then the system will essentially equilibrate within a state, and one event is not correlated with any other event. In this case the system will cross transition states with equilibrium probability. Figure 2.1 is a one dimensional illustration, but in reality the system is $3N$ dimensional, N being the number of atoms. The transition rate, can be calculated from the probability of being in the transition state with momenta taking the

system across the $3N - 1$ dimensional dividing surface between the states:

$$(2.16) \quad R = \int_{\bar{p}^*} \frac{\bar{p}}{m} P(\bar{r}^*, \bar{p}) d\bar{p} = \frac{\int_{\bar{p}^*} \frac{\bar{p}}{m} \exp(-H(\bar{r}^*, \bar{p})/k_B T) d\bar{p}}{\int \int \exp(-H(\bar{r}, \bar{p})/k_B T) d\bar{r} d\bar{p}},$$

where \bar{r} and \bar{p} are the system position and momentum coordinates, an $*$ indicates a position in the transition state and momenta taking the system out of the current state, H is the Hamiltonian giving the system energy, k_B is the Boltzmann constant, and T is temperature. A common approximation is that the potential energy landscape is harmonic in the stable state and at the transition state in the direction of a transition. In this case, Equation 2.16 becomes

$$(2.17) \quad R = \nu \exp(-G^m/k_B T),$$

where the Gibbs free energy for migration, G^m , is the energy barrier height going from the minimum of the stable state to the transition state, and the attempt frequency, ν , is

$$(2.18) \quad \nu = \frac{\prod_i^{3N} \nu_i^{min}}{\prod_i^{3N-1} \nu_i^*}.$$

Here ν_i^{min} are the normal mode frequencies at the minimum of the stable state, and ν_i^* are the normal mode frequencies at the transition state. The Gibbs free energy for migration can be written as

$$(2.19) \quad G^m = E^m - \bar{\sigma} : \bar{V}^m - TS^m,$$

where $\bar{\sigma}$ is the stress tensor, and E^m , \bar{V}^m , and S^m denote the energy, volume, and entropy of migration, respectively. The migration volume is defined similarly as the formation volume,

$$(2.20) \quad V^m \equiv -\frac{\partial G^m}{\partial \bar{\sigma}}.$$

2.1.1 Review of experimental results

Examining the results of the previous section, we see that the parameters that must be determined are the prefactors and G^f and G^m . Experimentally, diffusion profiles are measured using a technique such as secondary ion mass spectroscopy (SIMS). Then the chemical diffusivity is determined by fitting the profiles to the appropriate solution of Fick's laws. The prefactors and Arrhenius exponents can be determined from diffusion profiles at different times and temperatures, but it is often difficult to identify different mechanisms or the migration and formation energies separately. Even if more than one diffusion mechanism is operating, experiments in equilibrium may be well-described with a single activation energy. To obtain an equilibrium dopant diffusivity we combine Equation 2.1, setting $C_S = m_{AX}C_A^{Tot}$, and Equations 2.14, 2.15, and 2.17 to get

$$(2.21) \quad D_A^e = fg\alpha^2 z m_{AX} \nu \exp(-G^*/k_B T),$$

with activation energy $G_{AX}^* = G_{AX}^f + G_{AX}^m$. Then, analogous to the migration and formation free energies,

$$(2.22) \quad G^* = E^* - \bar{\sigma} : \bar{V}^* - TS^*.$$

In experiments at zero stress, E^* is determined from the Arrhenius exponent, and the change with varying experimental pressure determines \bar{V}^* . Entropic effects modify the prefactors and are often considered small and neglected.

2.1.1.1 Diffusion without stress

The equilibrium temperature dependence of the self-diffusion coefficient of silicon has been determined by Bracht *et al.* from isotope heterostructure experiments using

SIMS to observe diffusion profiles [19]. That work gives

$$D_{Si}^e = (530_{-170}^{+250}) \exp\left(-\frac{4.75 \pm 0.04 \text{ eV}}{k_B T}\right) \text{ cm}^2 \text{ s}^{-1}$$

which they show is in good agreement with previous experiments over more limited temperature ranges, and it has also been confirmed by Ural *et al.* [20]. Vacancy-mediated diffusion, interstitial-mediated diffusion, and direct exchange of silicon atoms could all contribute to silicon self-diffusion and dopant diffusion. The relative contributions of these mechanisms are still somewhat uncertain and must be determined in order to accurately predict non-equilibrium behavior.

A common approach for determining the contribution of each mechanism involves injecting excess point defects by oxidation or nitridation and determining the change in diffusivity for self-diffusion and dopant diffusion [5, 21, 22, 23]. With E standing for the direct exchange mechanism, the change in diffusivity due to non-equilibrium point defect concentrations is

$$\begin{aligned} D_A &= D_{AV} \frac{C_{AV}}{C_A^{Tot}} + D_{AI} \frac{C_{AI}}{C_A^{Tot}} + D_{AE} \\ &= D_{AV} \frac{C_{AV}^e}{C_{AV}^e} \frac{C_{AV}}{C_A^{Tot}} + D_{AI} \frac{C_{AI}^e}{C_{AI}^e} \frac{C_{AI}}{C_A^{Tot}} + D_{AE} \\ \frac{D_A}{D_A^e} &= \frac{D_{AV}}{D_A^e} \frac{C_{AV}^e}{C_A^{Tot}} \frac{C_{AV}}{C_{AV}^e} + \frac{D_{AI}}{D_A^e} \frac{C_{AI}^e}{C_A^{Tot}} \frac{C_{AI}}{C_{AI}^e} + \frac{D_{AE}}{D_A^e}. \end{aligned}$$

With the condition of local equilibrium and that C_{AX} is much less than C_A it holds that [5],

$$\frac{C_{AX}}{C_{AX}^e} = \frac{C_X}{C_X^e}.$$

Then, defining the component fractions of diffusion by

$$\sum_X f_{AX}^e = 1, \quad f_{AX}^e \equiv \frac{D_{AX}}{D_A^e} \frac{C_{AX}^e}{C_A^{Tot}},$$

results in

$$(2.23) \quad \frac{D_A}{D_A^e} = f_{AV}^e \frac{C_{AV}}{C_{AV}^e} + f_{AI}^e \frac{C_{AI}}{C_{AI}^e} + f_{AE}^e.$$

Using SIMS to measure profile evolution, the change in diffusivity of the species can be measured, and under identical oxidation and nitridation treatments the excess point defect concentrations are expected to be identical for silicon and for different dopants. Then, with only conservative assumptions about the excess point defect concentrations, an underdetermined system of equations can be solved for bounds on the fractions of diffusion. Using this method, Ural *et al.* have shown that Si self-diffusion and As dopant diffusion have contributions from both vacancy and interstitial mechanisms and possibly, to a lesser extent, direct exchange. This method has also shown that Sb diffusion is dominated by the vacancy mechanism, while B and P diffusion are dominated by the interstitial mechanism [5, 20, 21, 22, 23]. In those calculations the contribution of the direct exchange mechanism was assumed to be zero because an *ab initio* calculation showed that it has a higher activation energy than interstitial or vacancy mechanisms in B and P [24]. The fact that injection of one of the defects retards diffusion for B, P and Sb shows experimentally that direct exchange is not the dominant diffusion mechanism for these dopants [5]. Since one mechanism dominates B, P, and Sb diffusion, the measured activation energy can be considered a good estimation of the activation energy for the dominant mechanism. As compiled by Fahey for diffusion under intrinsic conditions, the measured activation energy for B ranges from $E_B^* = 3.25$ to 3.87 eV, for P from $E_P^* = 3.51$ to 3.67 eV, for Sb from $E_{Sb}^* = 3.89$ to 4.05 eV, and for As from $E_{As}^* = 4.05$ to 4.34 eV. For silicon self-diffusion and As diffusion, the activation energies of multiple mechanisms need to be parameterized.

The activation energies of the vacancy and interstitial mechanisms for silicon self-diffusion are still somewhat uncertain. From the silicon self-diffusion coefficient and Au or Zn diffusion results [25, 26, 27], Bracht calculates $G_{SiI}^* = 4.84$ to 4.95 eV and

$G_{SiV}^* = 4.1$ to 4.7 eV [28]. On the basis of the point defect injection experiments, Ural *et al.* calculate $G_{SiI}^* = 4.68$ eV and $G_{SiV}^* = 4.86$ eV [20]. From those results, and also measurements of defects in silicon crystal growth and vacancy profiles in rapid thermal annealing (RTA) wafers, Voronkov and Falster calculate $G_{SiI}^* = 5.04$ eV and $G_{SiV}^* = 4.5$ eV [29].

The final step is to determine formation and migration energies, but observation of silicon point defects and dopant-defect complexes is quite difficult so few experimental results are available. At low temperatures, electron paramagnetic resonance (EPR) can be used to determine migration energies. For dopant diffusion, studies with EPR determined the B-interstitial migration energy to be $E_{BI}^m \approx 0.6$ eV, the P-vacancy migration energy to be $E_{PV}^m = 0.94$ eV, the As-vacancy migration energy to be $E_{AsV}^m = 1.07$ eV, and the Sb-vacancy migration energy to be $E_{SbV}^m = 1.28$ eV [30, 31, 32]. For Si self-diffusion, EPR studies found silicon vacancy and interstitial migration energies to be $E_{SiV,SiI}^m \leq 0.5$ eV [33]. Voronkov and Falster [29] use this data along with measurements of defect profiles resulting from high temperature crystal growth and RTA to determine, with an estimated 30% uncertainty,

$$D_{SiV} = (0.002) \exp\left(-\frac{0.38\text{eV}}{k_B T}\right) \text{cm}^2\text{s}^{-1},$$

$$C_{SiV}^e = (2.4 \times 10^3) \exp\left(-\frac{4.12\text{eV}}{k_B T}\right) \text{cm}^{-3}.$$

They deduce that the interstitial migration energy must be comparably small and estimate,

$$D_{SiI} = (0.004) \exp\left(-\frac{0.3\text{eV}}{k_B T}\right) \text{cm}^2\text{s}^{-1},$$

$$C_{SiI}^e = (1.3 \times 10^6) \exp\left(-\frac{4.74\text{eV}}{k_B T}\right) \text{cm}^{-3}.$$

Contradictory to these figures, Bracht *et al.* report proton radiation enhanced self-diffusion (RESA) results at 780 to 872 °C that give vacancy formation and migration

enthalpies of $E_{SiV}^f = 2.1 \pm 0.5$ eV and $E_{SiV}^m = 1.8 \pm 0.5$ eV [27]. From Zn diffusion, Bracht determined the high temperature interstitial migration enthalpy $E_{SiI}^m = 1.77 \pm 0.12$ eV, which gives $E_{SiI}^f = 3.18 \pm 0.15$ eV [27]. They accept the low temperature EPR studies as reliable and conclude that the migration enthalpies increase with temperature as might be predicted by an extended defect [34]. Dannefaer *et al.* reported positron annihilation experiments that gave $E_{SiV}^f = 3.6 \pm 0.2$ eV, but this could not be replicated in later studies because the vacancy concentration is lower than the detection limit [35, 36, 37, 38].

To our knowledge, experimentally determined activation energies for interstitial and vacancy mechanisms for As diffusion have not been reported separately, though they might be deduced from the experimental results for As-vacancy complex (V_mAs_n) formation and migration energies. Arsenic doping has received significant attention because it becomes deactivated at high concentrations so that the free-carrier concentration saturates at around 3×10^{20} cm⁻³ [39, 40, 41]. Fair and Weber examined how the inactive As affected the solubility and diffusivity of As donors and proposed the formation of V_1As_2 complexes. Arsenic deactivation is proposed to occur when up to four arsenic atoms cluster around a vacancy so that each As atom is only covalently bonded to three neighbors and it becomes more energetically favorable for each As to keep its remaining two electrons [40, 42]. Arsenic-vacancy complexes are therefore important in determining both diffusion and deactivation. For As diffusion, some information about formation and migration energies has been reported by Ranki *et al.* from positron annihilation experiments [35, 41]. Positrons become trapped at neutral and negatively charged vacancies due to the missing positive ion core causing an increase in the average lifetime that gives a concentration measurement. Electron momentum distribution shape depends on

the valence and core electrons surrounding the vacancy and can be used to determine chemical composition. For 30 min anneals at temperatures from 300 to 900 K, they find that V_1As_1 migrates and runs into other As to form V_1As_2 at 450 K or higher and that V_1As_2 migrate to form V_1As_3 at 700 K or higher. From the temperatures at which the clusters form they estimate the migration energies of V_1As_1 and V_1As_2 to be $E_{V_1As_1}^m = 1.3 \pm 0.2$ eV and $E_{V_1As_2}^m = 2.0 \pm 0.2$ eV [43]. The change in complex concentration with temperature allows for a measurement of the formation energy, and concluding that complexes being measured are V_1As_1 they find that $E_{V_1As_1}^f = 1.1 \pm 0.2$ eV. Using an V_1As_1 binding energy found from simulation they found that $E_{S_iV}^f = 2.8 \pm 0.3$ eV [35, 44, 45]. Similar measurements showed that P and Sb form di- and tri-dopant-vacancy complexes similarly, though formation and migration energies were not reported [41].

2.1.1.2 Diffusion with stress

The last experimental data we review are measurements of stress dependent dopant diffusion. There are an increasing number of studies of diffusion in strained Si and $Si_{1-x}Ge_x$, but still few quantitative reports of activation volumes. The most complete study to date is that of Aziz *et al.* on the stress dependence of B and Sb diffusion in silicon [46, 47, 48]. They used high temperature, high pressure fluid argon to create hydrostatic pressure in their B and Sb doped Si superlattice samples. They also compare their results to measurements of B and Sb diffusion in biaxially strained Si and $Si_{1-x}Ge_x$ done by other groups [49, 50, 51, 52, 53]. In order to separate the effects of composition and strain, biaxial strain is controlled independently of Ge concentration by depositing the diffusion heterostructure on top of a buffer layer. Lower Ge concentration in the buffer will create compressive strained layers

and higher Ge concentration in the buffer will create layers with tensile strain [51].

Aziz *et al.* report an apparent activation volume which includes the possible effect of changes in the frequency factor and attempt frequency with pressure. In the case that these are pressure independent, the apparent activation volume is identical to the “true” activation volume. The scalar apparent activation volume is the trace of the tensorial apparent activation volume. They found that hydrostatic compression increases B diffusion and decreases Sb diffusion, with apparent scalar activation volumes $V_B^* = -3.2 \pm 1 \text{ \AA}^3$ and $V_{Sb}^* = 1.2 \pm 0.8 \text{ \AA}^3$ (we calculated these from the reported values and the atomic volume of silicon, $\Omega \approx 20 \text{ \AA}^3$). The biaxial strain studies report that compressive strains decrease B diffusion and increase Sb diffusion perpendicular to the biaxial strain plane. Taking into account the form of the full volume tensor, Aziz *et al.* find some agreement between the measured hydrostatic and biaxial stress dependencies for Sb, but not for B. Nygren *et al.* found that hydrostatic pressure increased the diffusivity of As in Si, characterized by an activation volume ranging from -7.8 to -12.3 \AA^3 [54]. In other qualitative studies on P and As diffusion in strained layers, their diffusivities decreased with compressive biaxial strain after accounting for Ge concentration [53, 55]. Additionally, Dilliway *et al.* found increased As diffusion in tensile strained Si, though in their analysis they did not recognize that strain can affect both formation and migration energies [56].

This review of the experimental data available for defect and dopant diffusion is meant to highlight a few main points. First, the most accurate and available experimental data is of the equilibrium activation energies. Second, the type of intrinsic point defects that mediate dopant and self-diffusion are known, but especially in the case of As and Si self-diffusion, the relative contributions of each mechanism

are somewhat uncertain. Third, even when it is known if diffusion is mediated by vacancies or interstitials, there can be multiple diffusing species in multiple charge states each with different formation and migration energies that depend on local composition. Fourth, separately determining the formation and migration energies of the diffusing species is difficult experimentally and is thus rather uncertain and unavailable. Fifth, experimentally determining activation, formation, and migration volumes is difficult due to the combination of uncertainties in stress measurements, diffusivities, and separating formation and migration volume effects and thus the data is mostly unavailable.

2.1.2 Review of computational results

Because of the experimental difficulties, simulation has found a significant role in determining diffusion mechanisms, energies and volumes. Simulation helps explain experimental results and fills gaps in the experimental data to allow for parameterization of process models. In atomistic simulations, E and \bar{V} are measured from the change in energy and volume of a supercell, as shown in Figure 2.2. The formation quantities are calculated with conservation of mass, which gives rise to the addition or subtraction of the energy or volume of a single silicon atom. The without conservation of mass, the change alone is called the relaxation energy, E^r , or volume, \bar{V}^r . For zero-pressure simulations in the large supercell limit, the energy and volume of the supercell is expected to approach the value for an isolated defect in an infinite medium at zero-pressure. There have been many atomistics simulation studies done to determine the structure and energetics of silicon vacancies and interstitials and dopant-defect complexes using both *ab initio* methods and empirical potentials. In the following sections we review the experimental and computational work done to

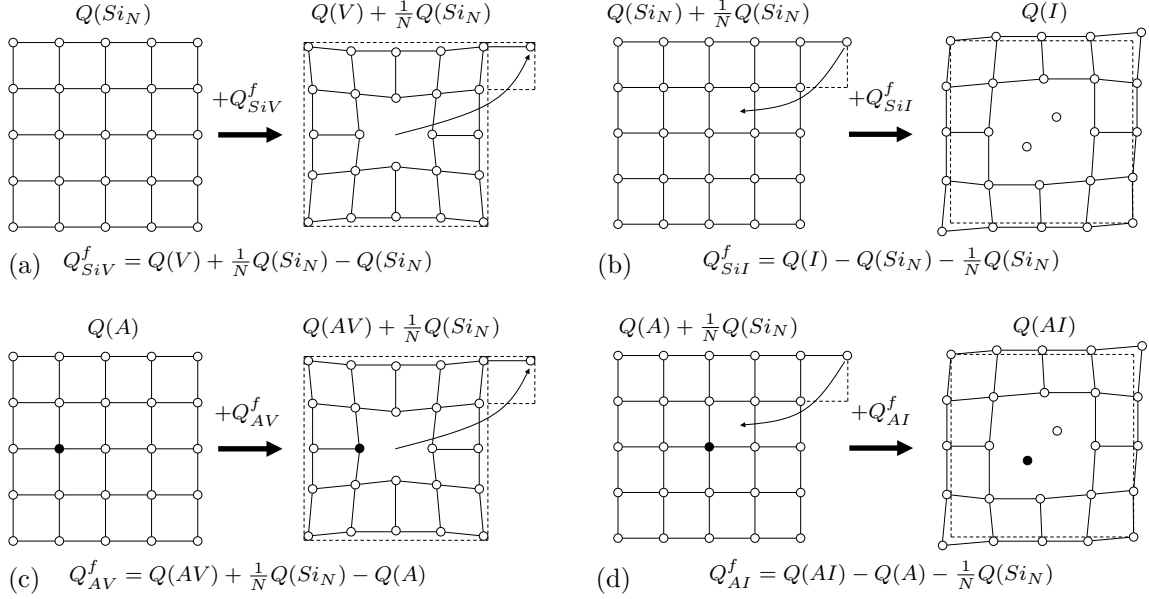


Figure 2.2: Schematics showing the calculation of formation energies and volumes from a supercell. The formation calculation is shown for an (a) isolated Si vacancy and (b) interstitial, and for a dopant-defect complex in (c) and (d). In this case, the reference state is the perfect crystal with dopant A in a substitutional position. $Q = E$ or \bar{V} .

determine these parameters.

2.1.2.1 Silicon Vacancy

There have been many DFT calculations of the energetics of silicon vacancies and interstitials. For instance, formation energies calculated by energy minimization at 0 K for the neutral vacancy range from about $E_{SiV}^f = 3.1$ to 4.1 eV, with most falling between $E_{SiV}^f = 3.2$ and 3.7 eV. In agreement with experiment [33], *ab initio* simulations that are large enough show that the relaxed vacancy undergoes a Jahn-Teller distortion [57] in which neighboring atoms pair and move apart from each other, reducing the symmetry of the vacancy. We have compiled *ab initio* data for the neutral silicon vacancy in Table 2.1, including the largest supercell result from each study. Unless mentioned in the notes accompanying Table 2.1, the simulated vacancy has undergone the Jahn-Teller distortion. The migration energies, ranging

from $E_{SiV}^m = 0.28$ to 0.59 eV, are in good agreement with the low temperature EPR studies of Watkins. Calculating an activation energy for Si diffusion by a vacancy mechanism from the simulation results gives a range of $E_{SiV}^* = 3.5$ to 4.3 eV. This could be in agreement with the range of $E_{SiV}^* = 4.1$ to 4.7 determined by Bracht *et al.*, except that from RESD they determined a migration energy $E_{SiV}^m = 1.8 \pm 0.5$ eV and a formation energy of $E_{SiV}^f = 2.1 \pm 0.7$ eV, quite different from the simulation results. They proposed a migration enthalpy increase with temperature due to an extended defect, but this does not explain the discrepancy in the formation energies.

There are relatively few *ab initio* formation volume calculations for the isolated silicon vacancy, and only one vacancy migration volume calculation. An early calculation by Antonelli and Bernholc found the formation volume for a neutral vacancy to be $V_{SiV}^f = 15 \text{ \AA}^3$ and the neutral tetrahedral interstitial to be $V_{SiI_T}^f = -5.6 \text{ \AA}^3$ [58]. They also calculated the pressure dependence of the formation energy of those defects in the doubly charged state and for bond-centered interstitials and the concerted exchange mechanism. Since then, the volume change of the defect has been quantified by calculating the change in the volume of the tetrahedron of neighboring atoms [24, 59, 60, 61, 62, 63], but this gives no information about the long-range elastic relaxation. Antonelli *et al.* revisited the vacancy about ten years later and found two distortions with similar energy at zero pressure, but different formation volumes $V_{SiV}^f = -1.7 \text{ \AA}^3$ for the first and $V_{SiV}^f = 4.0 \text{ \AA}^3$ for the second [64]. These formation volumes were calculated for hydrostatic pressure from supercell cells fixed to remain cubic, and give no information on the anisotropy of the Jahn-Teller distortion.

To our knowledge there are two *ab initio* studies that have report the full volume tensor for vacancies in Si, and only Centoni *et al.* report a full volume tensor for Si interstitials [65]. Windl *et al.* performed constant-pressure minimization of supercells

Table 2.1: *Ab initio* results for a neutral, isolated Si vacancy. Unless noted, the vacancy has undergone the Jahn-Teller distortion.

Year	Authors	Method	PW (Ry)	N	k-points	BC	E_f (eV)	E^m (eV)	$\text{tr}[V^f]$ (\AA^3)	$\text{tr}[V^m]$ (\AA^3)	Ω (\AA^3)
2005	¹ Centoni <i>et al.</i>	GGA	11	256	see ref	CP	3.69	0.28	-0.4	-5.9	20.34
2003	Mueller <i>et al.</i>	GGA	12	216	2 ³		3.66				20.35
2003	Liu <i>et al.</i>	GGA		216	2 ³		3.7				20.37
2002	² Goedecker <i>et al.</i>	GGA	35	216	Γ		3.17				
2001	² Kumeda <i>et al.</i>	GGA	12	216	Γ	CV	3.336	0.586			20.18
2004	² Diebel and Dunham	GGA	29	64	2 ³	CV	3.38				20.01
1989	Antonelli and Bernholc	GGA	6-16	32		m-CV	4.4		15		
2004	El-Mellouhi <i>et al.</i>	LDA		512	2 ³	CV	3.29	0.4			19.57
2003	³ Lento and Nieminen	LDA	15	256	(1/4,1/4,1/4)		3.6				19.57
2003	⁴ Lento and Nieminen	sx-LDA		256	(1/4,1/4,1/4)		3.515				20.12
2002	² Goedecker <i>et al.</i>	LDA	25	216	Γ		3.56				
2001	² Kumeda <i>et al.</i>	LDA	15	216	Γ	CV	3.32				19.57
2001	⁵ Windl <i>et al.</i>	LDA, GGA	11-13	216	2 ³	CP	3.6		-2.0		
1998	⁶ Puska <i>et al.</i>	LDA	15	216	Γ (2 ³)	CV	3.27 (3.31)				19.57
1998	⁷ Antonelli <i>et al.</i>	LDA	16	216	(1/4,1/4,1/4)	m-CV	3.49		-1.7 (+4.0)		19.7
1998	Zywietz <i>et al.</i>	LDA	13-2	216	4 ³		3.503				19.68
1998	Mercer <i>et al.</i>	LDA	12-20	128	2 ³		3.65				19.49
2003	² Ramanarayanan <i>et al.</i>	LDA	20	64	2 ³	m-CV	3.31		-1.15		19.57
2003	Al-Mushadani	LDA	24	64	3 ³		3.53				19.54
1999	² Xie and Chen, J Phys. D	LDA	12	64	2 ³		3.57	0.31			19.68
1996	⁶ Seong and Lewis	LDA	10	64	Γ		3.29				19.63
1992	² Sugino and Oshiyama	LDA		64		m-CV	4.0	0.2	-6.7	-2.7	
1996	^{2,8} Ramamoorthy and Pantelides	LDA	15	32	see ref		3.78	0.5			19.9
1989	^{2,9} Nichols <i>et al.</i>	LDA	see ref	32	see ref		3.5	0.3			19.98
1997	^{2,10} Pankratov <i>et al.</i>	see ref	12	64	see ref	see note	3.6				
1993	Blöchl <i>et al.</i>	see ref	8-14	64	3 ³	CV	4.1				
1997	² Tang <i>et al.</i>	TBMD		216		m-CV	3.97	0.1	0.6		20.26

Method: GGA = Generalized gradient approximation; LDA = Local density approximation; sx-LDA = Screened-exchange LDA; TBMD = Tight-binding molecular dynamics

PW: Plane-wave basis set energy cutoff

BC: Boundary condition: CV = Constant volume; m-CV = Multiple CV calculations to find minimum energy lattice parameter; CP = Constant pressure; if not mentioned, likely CV

Notes: ¹Calculate full volume tensors for neutral vacancy and interstitials, calculate energies and $\text{tr}[V^f]$ for other charge states, k-points "equivalent to 43 in an 8-atom cell". ²Vacancy symmetry not mentioned. ³Also calculates other charge states. Calculates vacancy volume from surrounding tetrahedron.

⁴Relaxes outward. ⁵ $\text{tr}[V^f]$ obtained from an average over 216, 512, and 1000 atom supercells. ⁶Calculates vacancy volume from surrounding tetrahedron. ⁷Found two structures with different volumes but nearly equal formation energy. ⁸Not clear if E^m is in bulk or near As dopant. ⁹No Jahn-Teller distortion. ¹⁰Positions of atoms at the edge of the supercell were fixed.

containing up to 1000 atoms and calculated the vacancy formation volume tensor to be [66],

$$\bar{V}_{SiV}^f = \begin{bmatrix} -9.3 & 0 & 0 \\ 0 & -9.3 & 0 \\ 0 & 0 & 16.5 \end{bmatrix} \text{Å}^3.$$

This shows that the Jahn-Teller distortion of the vacancy induces a strongly anisotropic volumetric change in the surrounding material. They also examine the change of bulk modulus with supercell size and find a $1/N$ form, N being the number of atoms, about ten times stronger for the vacancy than the $\langle 110 \rangle$ interstitial for which they do not report a formation volume. In small supercells the bulk modulus change is due to the high concentration of defects and could significantly affect the formation volume. Centoni *et al.* calculated formation energies and scalar volumes for defects in various charge states using constant volume calculations fit to an equation of state, but then they also calculated full volume tensors for neutral defects by allowing all of the vectors defining the supercell to relax. Using the first method they find $V_{SiV}^f = -1.4 \text{Å}^3$, which is in excellent agreement with the calculations of Antonelli and Bernholc. After allowing full relaxation of the supercell vectors, they find that $\text{tr}[V_{SiV}^f] = -0.4 \text{Å}^3$, only a small change. The full relaxation allows them to pick up the effect of the Jahn-Teller distortion and (transformed to Cartesian coordinates) results in

$$\bar{V}_{SiV}^f = \begin{bmatrix} -10.0 & -0.1 & 0 \\ -0.1 & -10.0 & 0 \\ 0 & 0 & 19.6 \end{bmatrix} \text{Å}^3.$$

This agrees quite well with the result of Windl *et al.*. The migration of a vacancy occurs when a neighboring atom jumps into the vacancy, and it is intuitive that the

saddle point of this process will occur when the atom is halfway between its original position and its final position. This is called the “split vacancy” and can actually be the stable position for doubly-negative charged vacancies, but for the neutral vacancy it is the saddle point for migration. For a split vacancy along a [111] bond, Centoni *et al.* report its relaxation tensor in terms of its principle directions, from which we can calculate its formation tensor (transformed to Cartesian coordinates),

$$\bar{V}_{SiV,split}^f = \begin{bmatrix} -2.1 & -5.2 & -5.2 \\ -5.2 & -2.1 & -5.2 \\ -5.2 & -5.2 & -2.1 \end{bmatrix} \text{Å}^3.$$

The difference between the formation volume of the split vacancy and the lattice vacancy allows us to calculate the strongly anisotropic migration volume for a [111] vacancy jump,

$$\bar{V}_{SiV}^m = \begin{bmatrix} -7.9 & -5.1 & -5.2 \\ -5.1 & -7.9 & -5.2 \\ -5.2 & -5.2 & -21.7 \end{bmatrix} \text{Å}^3.$$

The strong anisotropy of the migration volume suggests strong anisotropy of the diffusivity. However, much of the anisotropy comes from the orientation of the Jahn-Teller distortion and El-Mellouhi *et al.* [67] calculate that the reorientation energy is small, $E_{SiV}^{JT,rot} = 0.2$ eV. This confirms an experimental result of $E_{SiV}^{JT,rot} = 0.23$ eV. In the experiment [68], vacancies are originally oriented by applying a stress at high temperature. Then, the temperature is reduced to 20 K, the stress relieved, and the reorientation of vacancies observed upon illumination. Therefore, at typical processing temperatures the low barrier to reorientation is likely to decrease the anisotropic effect of the Jahn-Teller distortion on the diffusivity, though the anisotropic effect of the transition state formation volume would remain.

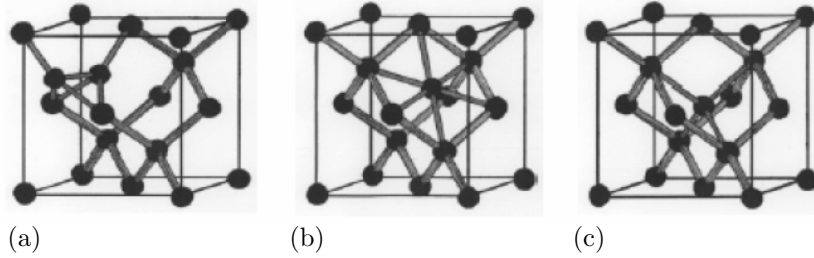


Figure 2.3: The three lowest energy Si self-interstitial defects are the (a) $\langle 110 \rangle$ dumbbell, (b) hexagonal, and (c) tetrahedral interstitials.

2.1.2.2 Silicon self-interstitial

Several stable self-interstitial structures have been found in empirical and *ab initio* simulations. We will consider the three structures, shown in Figure 2.3, that *ab initio* simulations have found to be lowest in energy. The $\langle 110 \rangle$ dumbbell interstitial is formed by placing two Si atoms aligned in a $\langle 110 \rangle$ direction astride a single lattice point. The hexagonal interstitial is formed by placing an interstitial atom at the center of one of the six-member rings in the Si lattice. In the cubic unit cell with normalized lattice parameter and basis atoms at $(0,0,0)$ and $(1/4, 1/4, 1/4)$, one such location is at $(3/8, 5/8, 3/8)$. The tetrahedral interstitial is located at $(1/2, 1/2, 1/2)$, where it is equally distanced from the four 'internal' atoms in a diamond cubic lattice. We have compiled *ab initio* data for the neutral silicon interstitial in Table 2.2, including the largest supercell result from each study. The DFT studies using the local density approximation (LDA) to the exchange and correlation function result in lower energies than the generalized gradient approximation (GGA). Both methods are in agreement that the $\langle 110 \rangle$ and hexagonal defects are similar in energy to each other and lower in energy than the tetrahedral interstitial. One tight-binding (TB) result is included in Table 2.2 that indicates that TB finds a formation energy for hexagonal interstitials that is much larger than the formation energy for a $\langle 110 \rangle$ interstitial.

Table 2.2: *Ab initio* results for neutral, isolated $\langle 110 \rangle$, hexagonal, and tetrahedral Si interstitials.

Year	Authors	Method	PW (Ry)	N	k-points	BC	Ω (\AA^3)	E^f (eV)	$\langle 110 \rangle$		Hexagonal		Tetrahedral	
									$\text{tr}[V^f]$ (\AA^3)	E^f	$\text{tr}[V^f]$	E^f	$\text{tr}[V^f]$	
1999	¹ Leung <i>et al</i>	DMC		54		CV	20.00	4.96		4.82		5.40		
2005	² Centoni <i>et al</i>	GGA	11	256		CP	20.34	3.73		3.79	-7.3	4.06	-12.8	
2002	Goedecker <i>et al</i>	GGA	35	216	Γ			3.31	-6.5	3.31				
2001	³ Kumeda <i>et al</i>	GGA	12	216	Γ	CV	20.18	4.406		4.456				
1999	⁴ RJ Needs	GGA	12	128	2^3	CV	20.00	3.84		3.8		4.01		
2005	⁵ Lin <i>et al</i>	GGA	15	64	see ref	CV	20.31					4.0		
2004	Diebel and Dunham	GGA	29	64	2^3		20.01	3.78						
1989	Antonelli and Bernholc	GGA	6-16	32		m-CV						3.6	-5.6	
2002	Goedecker <i>et al</i>	LDA	25	216	Γ			2.88		2.87				
1999	⁶ RJ Needs	LDA	12	128	2^3	CV	20.00	3.31		3.31			3.43	
2003	Al-Mushadani	LDA	24	64	3^3		19.54	3.40		3.45				
1996	Zhu <i>et al</i>	LDA		64	see ref	CV		3.2						
1997	Clark and Ackland	LDA, MD	11	64	4^3	CV		2.16		2.45				
1993	Blöchl <i>et al</i>	see ref	8-14	64	3^3	CV		3.3						
1997	Tang <i>et al</i>	TBMD		216		m-CV	20.25	3.8		4.93			4.39	

Method:

GGA = Generalized gradient approximation; LDA = Local density approximation; sx-LDA = Screened-exchange LDA; TBMD = Tight-binding molecular

dynamics

PW: Plane-wave basis set energy cutoff

BC: Boundary condition: CV = Constant volume; m-CV = Multiple CV calculations to find minimum energy lattice parameter; CP = Constant pressure; if

not mentioned, likely CV

Notes:

¹Used DMC to calculate energy of structures obtained from LDA simulations reported by Needs. ²Calculate full volume tensors for neutral vacancy and

 interstitials, calculate energies and $\text{tr}[V^f]$ for other charge states, k-points "equivalent to 43 in an 8-atom cell." ³Transition state $\langle 110 \rangle \rightarrow \text{hex } E^f = 4.629$,

 $\langle 110 \rangle$ reorientation=0.749. ⁴Energies include a correction for the PW cutoff, $\langle 110 \rangle \rightarrow \text{hex } E^m = 0.20$, hex \rightarrow hex $E^m = 0.18$. ⁵Includes E^f for +1 and

 +2 charged tetrahedral. ⁶Energies include a correction for the PW cutoff, $\langle 110 \rangle \rightarrow \text{hex } E^m = 0.15$, hex-hex $m = 0.03$. ⁷Effective interstitial $E^m = 1.37$ eV,

 $V^f = -2.0 \text{ \AA}^3$.

Leung *et al.* reported a calculation using diffusion kinetic Monte Carlo (DMC) [69, 70]. DMC uses stochastic methods to solve the Schrödinger equation, making it much more computationally intensive than DFT calculations, but it is expected to be more accurate than using LDA or GGA. Currently DMC has not been used to relax atomic positions, so Leung *et al.* found relaxed structures using LDA and then calculated the energy of those structures with DMC. The resulting energies, $E_{SiI,(110)}^f = 4.96$ eV, $E_{SiI,H}^f = 4.82$ eV, and $E_{SiI,T}^f = 5.40$ eV, are nearly 50% greater than the LDA results. They also calculated a migration energy for transitions between the $\langle 110 \rangle$ and hexagonal sites that is quite low, $E_{SiI}^m \leq 0.2$ eV. Combining these results gives an activation energy for interstitial diffusion that agrees well with the range of activation energies determined experimentally by Bracht *et al.* or Voronkov, $E_{SiI}^* = 4.84$ to 5.04 eV. Windl found that energy corrections for band gap and finite-size effects proposed elsewhere might explain the difference between LDA and DMC results [71] He finds that these corrections might not be necessary for the vacancy, but that there is too much uncertainty in the defect bands of Jahn-Teller distorted vacancies to be sure. If corrections are not necessary, the higher vacancy formation energy values reported ($E_{SiV}^f = 3.2$ to 3.7 eV) combined with the calculated migration energy values ($E_{SiV}^m = 0.28$ to 0.59 eV) would give an activation energy in good agreement with the lower end of the range determined by Bracht *et al.* ($E_{SiV}^* = 4.1$ to 4.7 eV). Additionally, Leung *et al.* calculated the energy for concerted exchange to be $E_{SiE}^* = 5.78$ eV. This is in good agreement with the experimental result that concerted exchange is a small contribution to Si self-diffusion.

The suggestion by Bracht *et al.* that migration enthalpies increase with temperature can not be determined from the 0 K *ab initio* calculations reviewed up to this point. Short *ab initio* MD simulations which might address this question were per-

formed by Blochl *et al.* and Clark and Ackland, but the computational requirements limited the studies to short times and high temperatures near the melting point [72, 73]. Blochl *et al.* found that interstitial diffusion dominated, but could not directly calculate the migration energy. Clark and Ackland found the $\langle 110 \rangle$ interstitial to be lowest in energy at 0 K and that the quenched structure of the transition state was a hexagonal interstitial, which supports the $\langle 110 \rangle$ to hexagonal to $\langle 110 \rangle$ diffusion path for Si self-interstitial diffusion. They could not calculate a high temperature migration energy either.

Besides an older calculation by Antonelli and Bernholc [58] (the study with the vacancy result revised 10 years later), the only formation volumes calculated for isolated Si self-interstitials come from Centoni *et al.* Transforming their relaxation volumes into formation volumes in Cartesian coordinates gives

$$\begin{aligned} \bar{V}_{SiI,\langle 110 \rangle}^f &= \begin{bmatrix} -0.8 & 5.6 & 0 \\ 5.6 & -0.8 & 0 \\ 0 & 0 & -4.9 \end{bmatrix} \text{ \AA}^3, \\ \bar{V}_{SiI,H}^f &= \begin{bmatrix} -2.4 & -0.3 & -0.3 \\ -0.3 & -2.4 & -0.3 \\ -0.3 & -0.3 & -2.4 \end{bmatrix} \text{ \AA}^3, \\ \bar{V}_{SiI,T}^f &= \begin{bmatrix} -4.3 & 0 & 0 \\ 0 & -4.3 & 0 \\ 0 & 0 & -4.3 \end{bmatrix} \text{ \AA}^3. \end{aligned}$$

Since the migration path seems to consist of jumps between $\langle 110 \rangle$ and hexagonal sites, the differences in their formation volume might be used to estimate a migration volume. However, Leung *et al.* report a transition energy for going between $\langle 110 \rangle$ and hexagonal sites so that transition state might also determine the migration volume.

2.1.2.3 Boron

Simulation work done to determine dopant diffusion mechanisms and energetics is largely focused on B and As, with the goal of understanding transient enhanced diffusion (TED) and electrical deactivation. TED is a phenomenon in which point defect supersaturation due to ion implantation results in significantly enhanced dopant diffusion until the point defect concentration re-equilibrates [74]. The implanted ions create Frenkel pairs of vacancies and interstitial, but because the ions tend to take substitutional positions interstitials dominate vacancies by the amount of the ion dose. Dopants that diffuse by an interstitial mechanism (B, P, and to a lesser extent As) can form many more mobile complexes than in equilibrium and their diffusivity is greatly increased. It is desirable to know the formation, binding, and migration energies of the mobile complexes in order to model TED.

The initial ab initio simulation results for B diffusion suggested a “kick-out” mechanism in which a Si interstitial knocks a substitutional B atom, B_S , into an interstitial position, B_I , where it diffuses rapidly until it returns to a substitutional position by creating a new Si interstitial [24, 75]. These simulations were limited to neutral defects, and found to accurately model B diffusion over a small range of temperatures and annealing times [76]. The calculation was revisited by Sadigh *et al.* and Windl *et al.* to include charged defects [76, 77]. These studies both found B to diffuse by the “interstitialcy” mechanism, meaning as a B-interstitial complex. Windl *et al.* found the diffusion path for neutral defects to be $B_S + Si_{I,T} \rightarrow B_{I,H} \rightarrow B_S + Si_{I,T}$, with a migration energy of $E_{BI}^m = 0.2$ (0.4) eV, for the GGA (LDA) method. For positively charge systems they found two competing diffusion paths, both $B_S + Si_{I,T}^+ \rightarrow B_S + Si_{I,T}^+$, but with different saddle points and migration energies of $E_{BI}^m = 0.8$ (1.2) eV or $E_{BI}^m = 1.0$ (1.3) eV, respectively. For

negatively charge systems the diffusion path was $B_{I,\langle 110 \rangle}^{-1} \rightarrow B_{I,\langle 100 \rangle}^{-1} \rightarrow B_{I,\langle 110 \rangle}^{-1}$ with migration energy $E_{BI}^m = 0.6$ (0.5) eV, where the $\langle 100 \rangle$ interstitial is a metastable dumbbell position aligned in the $\langle 100 \rangle$ direction. The formation energy for the neutral $B_S + Si_{I,T}$ complex was $E_{BI}^f = 2.8$ (2.5) eV. The formation energies for the charged complexes are Fermi-level dependent. Using the values from near midgap and combining all the mechanisms, they report an activation energy of $E_{BI}^* = 3.2$ to 3.6 eV, in excellent agreement with the experimental values of $E_{BI}^* = 3.25$ to 3.87 eV, and a migration energy of $E_{BI}^m = 0.4$ to 0.7 eV, also in excellent agreement with the experimental value of $E_{BI}^m = 0.6$ eV.

The simulation results of Windl *et al.* were qualitatively confirmed by *ab initio* MD simulations at 1200 and 1500 K by Sadigh *et al.* Sadigh *et al.* also calculated scalar activation volumes for the both the interstitialcy and kick-out mechanisms, reporting values of $V_{BI,i}^* = -3.0$ (-2.2) \AA^3 and $V_{BI,k}^* = -0.46$ (0.12) \AA^3 , respectively. Compared with the experimental value found by Aziz *et al.* of $V_{BI,k}^* = -3.2 \pm 1$ \AA^3 , this lends further support for the interstitialcy mechanism. An activation volume tensor for B-interstitial diffusion was reported by Laudon *et al.* for the saddle point which is B in a hexagonal position, referenced to isolated substitutional B [78],

$$\bar{V}_{BH}^* = \begin{bmatrix} -1.77 & 6.5 & 4.4 \\ 6.5 & -1.77 & 0.3 \\ 4.4 & 0.3 & -1.77 \end{bmatrix} \text{\AA}^3.$$

The trace, $\text{tr}[\bar{V}_{BH}^*] = -5.3$ \AA^3 , compares well to the hydrostatic experimental results and the fixed-cubic results of Sadigh *et al.*, but if all jumps have similar anisotropy it does not agree with the results of (100)-oriented biaxial strain tests that tend to show diffusivity increasing with biaxial strain.

There have also been two recent studies of B diffusion under strain using the same

migration paths as Windl *et al.* and Sandigh *et al.* The calculations of Dunham *et al.* [79] are reported in terms of a concentration dependent induced strain instead of volumes which makes comparison more difficult, but appear to show that the trace of the activation volume tensor is positive and not aligned in the same direction as the results of Laudon *et al.* The results of Lin *et al.* [80] are limited to biaxial tension and show a reduction in migration energy within the plane. This is a positive migration volume, opposite in sign from the negative activation volume. Therefore, the effect of stress on diffusion is sensitive to whether or not defects can equilibrate with a free surface. If they do equilibrate, then negative the activation volume will decrease diffusion in tensile stress. If they do not equilibrate, the positive migration volume will increase diffusion in tensile stress.

2.1.2.4 Arsenic

Understanding of As deactivation and diffusion improved thanks to several studies that calculated formation and migration energies for As-vacancy clusters. From experiments, discussed above, and simulation it is clear that there is a strong binding energy between As and vacancies. Simulations show that each additional As around a vacancy provides an energy reduction of about $E_{AsV}^b = 1.2$ to 1.5 eV [44, 45, 81, 82]. Three studies have reported migration energies for an As-vacancy complex. For the As to diffuse, the vacancy starting at a 1st nearest neighbor (1NN) position has to migrate around a six-member ring and approach the As from a new direction. The saddle point for the ring mechanism is between the 2NN and 3NN. Pankratov *et al.* calculated the formation energy to be $E_{AsV}^f = 2.4$ eV and the migration energy to be $E_{AsV}^m = 0.81$ eV [83]. Xie and Chen found the formation energy to be $E_{AsV}^f = 2.37$ eV and the migration energy to be $E_{AsV}^m = 1.19$ eV [45]. Nichols *et*

al. found the formation energy to be $E_{AsV}^f = 2.3$ eV and the migration energy to be $E_{AsV}^m = 1.07$ eV [24]. Xie and Chen calculated migration energies for the case of a vacancy diffusing between As atoms located at 5NN or 6NN positions from each other and found that the migration barriers were significantly lowered to $E_{AsV}^m = 0.45$ or 0.84 eV. This explained an experimental result that found significantly increased As diffusion at high concentrations during short anneals [84]. Then, the As will quickly cluster and the diffusivity decreases because the migration barrier for V_1As_2 diffusion is calculated to be 2.0 eV [85]. At high enough temperatures, V_1As_2 will diffuse and form V_1As_3 clusters, which judging from the positron annihilation experiments are even more immobile since there is little evidence of V_1As_4 clusters.

While the defect injection experiments find that As-interstitial complexes contribute significantly to As diffusion, few simulations have investigated the structure or energetics. Harrison *et al.* reported that the lowest activation energy, 3.26 eV, is for the neutral As-interstitial complex in a $\langle 110 \rangle$ dumbbell structure, with a formation energy of 3.11 eV and a migration energy of 0.15 eV [86]. This is in good agreement with the work of Nichols *et al.* that found a formation energy of 3.2 eV and migration energy of 0.4 eV.

Dunham *et al.* calculated the energies of As-interstitial complexes and induced strains for substitutional As and the As-interstitial and As-vacancy transition states. They reported that the lowest energy interstitial is the $\langle 110 \rangle$ at $E_{AsI}^f = 3.1$ eV, and found a migration energy of $E_{AsI}^m = 0.5$ eV. They report an isotropic induced strain for the As-interstitial transition state with activation volume $\text{tr}[\bar{V}_{AsI}^*] = -1.4$ Å³. They also wrote that they found "... that the AsV transition state is equivalent to (the) vacancy transition state in self-diffusion, and can be treated as isotropic." However, we know from the results of Centoni *et al.*, and intuition, that the vacancy

transition state is quite anisotropic, aligned in the $\langle 110 \rangle$ jump direction. They report an isotropic volume tensor because they do not allow the vectors defining their supercell to deviate from orthonormality.

Sugino and Oshiyama calculated activation energies and pressure dependence for P, As, Sb, and Si diffusion by vacancy mechanisms, and formation energies and pressure dependence for P, As, and Sb dopant-interstitial complexes [59]. They determined the activation energy for As diffusion by the vacancy mechanism to be $E_{AsV}^* = 3.9$ eV, and $E_{AsI}^* = 4.3$ eV for the interstitial mechanism. These are high compared to the other calculations, but in agreement with the range of reported experimental values reviewed by Fahey *et al.* [5]. They also determined the change in vacancy mechanism activation energy when a pressure of approximately 6 GPa is applied by decreasing the lattice constant of the supercell by 2.0%. Following the example of Aziz *et al.* [46], we can calculate the activation volume for the vacancy mechanism from their activation energy change by including a $+p\Omega$ work term for vacancy formation. This shows that their value corresponds to $V_{AsV}^* = 4 \text{ \AA}^3$. They also calculated the formation energy change with pressure for an As-interstitial pair and found $V_{AsI}^* = 1 \text{ \AA}^3$. Qualitatively similar results were obtained for the other dopants and silicon. The As-vacancy result is qualitatively opposite the experimental results of Nygren *et al.* However, if we examine the pressure dependence of the Si-vacancy activation enthalpy we find it has an activation volume of $V_{SiV}^* = -9.4 \text{ \AA}^3$. This is somewhat larger than the $V_{SiV}^* = -6.3 \text{ \AA}^3$ calculated by Centoni *et al.*, and Table 2.1 shows that the migration and formation volumes disagree, suggesting that Sugino and Oshiyama's calculations might be best viewed qualitatively.

2.1.2.5 Phosphorus and Antimony

There are relatively few simulation studies of P and Sb formation and migration energies [24, 59, 87, 88]. For P diffusion, Nichols *et al.* report activation energies of $E_{PI}^* = 3.8$ eV and $E_{PV}^* = 3.4$ eV for the interstitial and vacancy mechanisms, respectively, while Sugino and Oshiyama reported $E_{PV}^* = 4.2$ eV for the vacancy mechanism. More recently, Nelson *et al.* reported a neutral P-vacancy activation energy of $E_{PV}^* = 3.69$ eV, which is equivalent to their vacancy formation energy because they found that the P-V binding energy and exchange energy are equal at $E_{PV}^b = E_{PV}^x = 1.05$ eV. This seems to assume that the ring migration energy is less than 1.05 eV, which is not reported but could not be more than 1.23 eV, the sum of the binding energy and vacancy migration energy, $E_V^m = 0.18$ eV. It is confirmed by Liu *et al.* who calculated the ring migration energy to be 0.8 eV, and confirmed the activation energy for the neutral case. However, they found a lower activation energy for the negatively charged complex, $E_{PV,(-1)}^* = 3.4$ eV. They suggest that inaccuracies in the band structure due to GGA or LDA give a result that is too low. Point defect injection experiments show that P diffusion is dominated by the interstitial mechanism and Liu *et al.* found an activation energy for the P-interstitial mechanism in the range $E_{PI}^* = 3.1$ to 3.5 eV, depending on the charge state.

For Sb diffusion, Nichols *et al.* report activation energies of $E_{SbI}^* = 4.9$ eV and $E_{SbV}^* = 3.6$ eV for the interstitial and vacancy mechanisms, respectively, while Sugino and Oshiyama report $E_{SbV}^* = 3.9$ eV for the vacancy mechanism. Nelson *et al.* only report the vacancy binding energy, $E_{SbV}^b = 1.45$ eV, which gives a formation energy of $E_{SbV}^f = 2.24$ eV. They find that since the Sb dopant is larger than the host Si atoms the lowest energy position for the Sb-V complex is with the Sb half-way between lattice points.

Sugino and Oshiyama calculated activation volumes for vacancy diffusion of P and Sb, finding $V_{P^*V}^* = -4.0 \text{ \AA}^3$ and $V_{Sb^*V}^* = 1.3 \text{ \AA}^3$. The activation energy for Sb diffusion is in very good agreement with the experimental value, $V_{Sb^*V}^* = 1.2 \pm 0.8 \text{ \AA}^3$, of Aziz *et al.*

2.2 Factors Affecting Atomistic Simulations

As discussed by Probert and Payne [63] or Puska *et al.* [61], the observed variations in calculation results can come from differences in the exchange-correlation functional, the pseudopotential used, the shape and size of the supercell, symmetry restrictions, and the Brillouin zone sampling used. Probert and Payne elaborate a methodology for obtaining well converged *ab initio* calculations which starts by converging the basis-set size, followed by the Brillouin-zone integration, the supercell shape and size, and the calculated Hellmann-Feynman forces. When the Hellmann-Feynman forces are converged the atomic positions are relaxed and the defect structure converged.

Due to interactions between the defect and its periodic images the finite supercell size affects the calculated energy and volume, and it can also affect the symmetry of the defect structure. For example, Zywietz *et al.* [89] and Puska *et al.* show that the Jahn-Teller distortion of the vacancy becomes stable after increasing the supercell size. It is most common to treat the size convergence by increasing the supercell size until the energy, defect symmetry and atomic positions are converged to some limit. However, this does not account for any systematic effects of periodic boundary conditions and generally has not included a separate volume convergence, which may be slower than energy convergence. Probert and Payne considered the atomic displacements and observed that near the defect the difference in relaxation

of successive atomic shells oscillates near the defect. They then treated the volume relaxation by showing that with a large enough supercell there is some convergence of the relaxation of successive shells of atoms before the shell farthest from the defect. While the calculation is certainly not converged if this was not the case, it does not guarantee that there are not systematic elastic effects causing convergence to the wrong solution. By the nature of periodic boundary conditions the difference in relaxation of successive shells diminishes to zero at the periodic boundary.

Probert and Payne also noted that supercell shape affects the convergence rate of the formation energy of an unrelaxed silicon vacancy. The face-centered cubic (FCC) supercells converge the most slowly, and the body-centered cubic (BCC) supercells converge the most quickly. They wrote that “...if we simply order the different possible supercells in terms of the total number of atoms (or equivalently, the defect-image distance), we will be misled as the defect density will be changing in a nonmonotonic manner...” and provided plots of unrelaxed formation energy as a function of number of atoms and defect density. However, this is backwards, it is the defect-image distance and the number of images at that distance that changes with supercell shape, not the defect density which is just $(\Omega N)^{-1}$ for an unrelaxed defect. They demonstrated there is a defect-defect interaction which perturbs the electronic charge density. If this perturbation aligns with the bonding, as it does in the BCC supercell, there is less of a localized charge-density difference and hence less of an effect on the unrelaxed formation energy. This explains the observed difference in convergence rates for unrelaxed formation energies, but it does not give us any indication of how supercell shape might affect energies or volumes through the elastic fields of relaxed defects.

The convergence of basis-set size, Brillouin-zone integration and calculated Hellmann-

Feynmann forces are particular to DFT calculations. The atomic relaxation and convergence of the defect structure with supercell shape and size are sensitive to quantum mechanical effects, but are also sensitive to longer-range elastic effects. For example, the results of Centoni *et al.* showed that there is a change in the trace of the volume tensor when the supercell is allowed to fully relax, and the vacancy migration volume tensor results of Dunham *et al.* showed that boundary conditions can drastically change the volume tensor symmetry. The supercell approximation creates a periodic array of defects, each defect surrounded by a strain field that it induces in the bulk material. Near the defect the atomic relaxation is highly non-linear, but farther away it is well-described by linear elasticity. For small supercells, the non-linear relaxations near the defect interact with those of its periodic images and affect the calculation of energies and volumes. As the supercell size increases those short range interactions vanish, but there will still be a long-range elastic field that could interact with the long-range elastic fields of its periodic images. It is not immediately apparent how these elastic fields couple with the shape and size of a supercell to affect the calculated energy and volume, especially when considering anisotropic defects. These questions are addressed in the work presented in Chapter 3.

CHAPTER 3

Accuracy of Supercell Calculations

In this chapter we use continuum linear elasticity to quantify the effect of boundary conditions on atomistic calculations to test the validity of formation energies and volumes obtained using the supercell approximation. Then we compare linear elasticity's predictions to atomistic calculations using an empirical potential to reach the far-field.

3.1 Continuum Elasticity Calculations

When a defect is created in a computational cell it induces relaxations in the surrounding material. It is natural to treat the interactions between periodic images of the defect in the supercell using continuum elasticity. Close to the defect atomic relaxations depend significantly on the bonding changes and may be very non-linear, but in the far-field the strain is well described by linear elasticity. Therefore, we use linear elasticity to quantify the effects of the boundary conditions as follows. Within linear elasticity point defects are modeled as centers of contraction or expansion by an elastic dipole. The solution for stress and strain due to an elastic dipole in an infinite linear elastic medium can be obtained using Green's functions. Then, through cutting and welding operations in the manner of Eshelby [90, 91], a solution can be found for a defect in a body with arbitrary boundary conditions. This allows

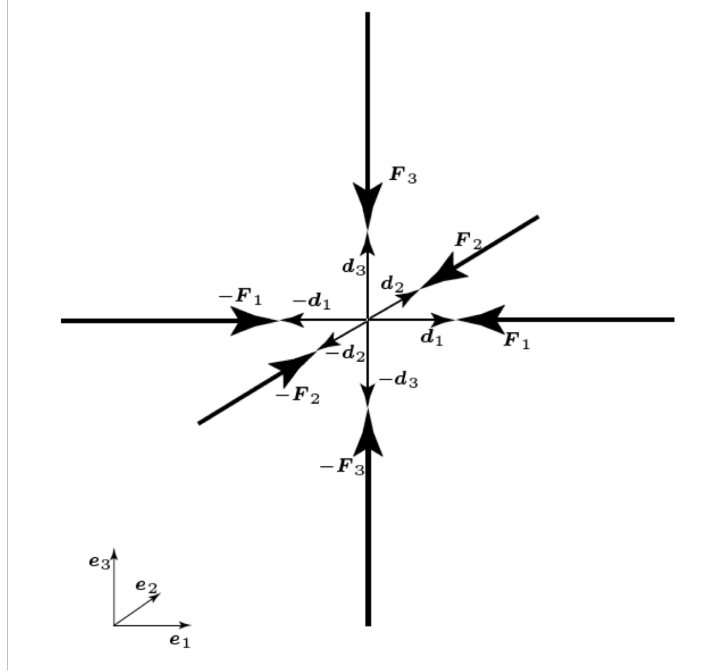


Figure 3.1: Schematic of an elastic dipole. Here the forces, F , are all the same magnitude, separated by the same distance, d , and aligned with the orthonormal basis, e , so it is an isotropic dipole.

us to determine the effect of PBC conditions on atomistically determined formation or migration energies and volumes.

3.1.1 Elastic dipole

An elastic dipole is composed of balanced point forces separated by a small distance, as shown in Figure 3.1.1. The associated dipole tensor is

$$(3.1) \quad \bar{\bar{D}} = \sum_{i=1}^3 \bar{F}_i \otimes 2\bar{d}_i,$$

From the dipole, a Green's function approach can be used to determine the resulting displacement field. The Green's function for an anisotropic, linear infinite body satisfies

$$(3.2) \quad \mathbb{C}_{ijkl} \frac{\partial^2 G_{km}}{\partial x_j \partial x_l} + \delta_{im} \delta(\bar{x} - \bar{x}') = 0,$$

where \mathbb{C}_{ijkl} is the elastic stiffness tensor, G_{km} is the k -component of the displacement at the position \bar{x} due to a unit point force at \bar{x}' in the m direction, and δ_{im} is the Kronecker delta. Then, the displacement field in the infinite body at \bar{x} given a point force at \bar{x}'' is determined from the Green's function by

$$(3.3) \quad \bar{u}^\infty(\bar{x}; \bar{x}'') = \int_{\mathbb{R}^3} \bar{\bar{D}}\bar{G}(\bar{x} - \bar{x}') \bar{\nabla}_{\bar{x}'} \delta(\bar{x}' - \bar{x}'') dV',$$

where \bar{x}' is the variable of integration. From this, the strain field can be determined as

$$(3.4) \quad \varepsilon_{il}^\infty = \frac{1}{2} \left(\frac{\partial^2 G_{ij}}{\partial x_k \partial x_l} + \frac{\partial^2 G_{lj}}{\partial x_k \partial x_i} \right) D_{jk},$$

and the stress field as $\sigma_{ij}^\infty = \mathbb{C}_{ijkl} \varepsilon_{kl}^\infty$. Evaluation of the infinite body displacement, stress, and strain can be obtained using analytical expressions for $\bar{\bar{G}}$ and its derivatives obtained by Barnett [92].

3.1.2 Formation Energy

From the solutions for stress and strain in an infinite body, the resulting strain energy can be calculated from

$$(3.5) \quad \mathbb{U}_\infty = \int_{\mathbb{R}^3} \frac{1}{2} \sigma_{ij}^\infty \varepsilon_{ij}^\infty dV.$$

Let us split the infinite body into two regions, B_0 with surface S_0 enclosing the defect, and B_1 including the rest of the body. The two regions have strain energy \mathbb{U}_{∞, B_0} and \mathbb{U}_{∞, B_1} , respectively. To compare the formation energies of atomistic calculations with various boundary conditions we use \mathbb{U}_{∞, B_0} as a reference state. Using the Divergence Theorem since B_1 is divergence-free, the volume integral above can be related to a surface integral,

$$(3.6) \quad \mathbb{U}_\infty = \mathbb{U}_{\infty, B_0} - \int_{S_0} \frac{1}{2} \sigma_{ij}^\infty n_j u_i^\infty dA,$$

where \bar{n} is the unit outward normal to S_0 .

Free boundary conditions on B_0 are obtained by holding the traction fixed at $\sigma_{ij}^\infty n_j$ on S_0 , cutting along the surface, and then applying the traction $\sigma_{ij}^F n_j = -\sigma_{ij}^\infty n_j$ to S_0 . Then, \mathbb{U}_{F,B_0} , the strain energy in B_0 with FBC is

$$(3.7) \quad \mathbb{U}_{F,B_0} = \mathbb{U}_{\infty,B_0} - \int_{B_0} \frac{1}{2} \sigma_{ij}^F \varepsilon_{ij}^F dV.$$

Since σ_{ij}^F is divergence-free in B_0 , this can also be related to a surface integral,

$$(3.8) \quad \mathbb{U}_{F,B_0} = \mathbb{U}_{\infty,B_0} - \int_{S_0} \frac{1}{2} \sigma_{ij}^F n_j u_i^F dA,$$

where u_i^F is the displacement arising from the traction $\sigma_{ij}^F n_j$.

An atomistic calculation with PBC and constant volume (CV) would correspond to taking the FBC result and displacing the boundary by $u_i^P = -(u_i^\infty + u_i^F)$, with associated stress and strain, σ_{ij}^P and ε_{ij}^P . Therefore, the strain energy of B_0 in the PBC case is

$$(3.9) \quad \mathbb{U}_{P,B_0} = \mathbb{U}_{F,B_0} + \int_{B_0} \frac{1}{2} \sigma_{ij}^P \varepsilon_{ij}^P dV.$$

For similar reasons as above, this can also be written

$$(3.10) \quad \mathbb{U}_{P,B_0} = \mathbb{U}_{F,B_0} + \int_{S_0} \frac{1}{2} \sigma_{ij}^P n_j u_i^P dA.$$

To obtain an average zero pressure PBC, an additional traction is applied equal to the opposite of the average traction on each surface in the constant volume PBC condition. In an atomistic calculation, the constant volume system would feel the opposite uniform stress/strain field, $\sigma_{ij}^{CV} = \mathbb{S}_{ijkl} \varepsilon_{kl}^{CV}$, which performs the return transformation. Since the average stress in the relaxed supercell is zero, the average stress measured in the constant volume supercell is $\sigma_{ij}^{av} = \sigma_{ij}^{CV}$. Therefore, the

zero-stress PBC formation energy, E^f , can be related to the constant volume PBC formation energy by

$$(3.11) \quad \begin{aligned} E^f &= E^{f,CV} - \int_B \frac{1}{2} \sigma_{ij}^{CV} \varepsilon_{ij}^{CV} dV, \\ E^f &= E^{f,CV} - V^{ref} \frac{1}{2} \sigma_{ij}^{av} \mathbb{S}_{ijkl} \sigma_{kl}^{av}, \end{aligned}$$

where $V^{ref} = B_0$, the volume of the finite body before the defect has been introduced or any transformations performed.

Our collaborators have numerically evaluated the quantities \mathbb{U}_{F,B_0} and \mathbb{U}_{P,B_0} at zero average pressure for cubic B_0 using the finite element method. The value for $\bar{\bar{D}}$ was obtained from atomistic calculations as described in the following section. The results, shown in Figure 3.2 along with atomistic calculations described in Section 3.2, demonstrate that both \mathbb{U}_{F,B_0} and \mathbb{U}_{P,B_0} converge to \mathbb{U}_∞ as the volume of the B_0 , $V_{B_0} \rightarrow \infty$. This indicates that formation energies in atomistic calculations will converge with system size.

3.1.3 Formation Volume

A somewhat similar process is used to determine the predicted continuum linear elastic effect of boundary conditions on atomistically calculated relaxation volumes. For any boundary conditions the observed tensorial volume change in a finite body, B , from the original defect-free state, ΔV_{kl} , is

$$(3.12) \quad \Delta V_{kl} = \int_B \varepsilon_{kl} dV.$$

From the stress-strain relation this is

$$(3.13) \quad \Delta V_{kl} = \mathbb{S}_{kl ij} \int_B \sigma_{ij} dV.$$

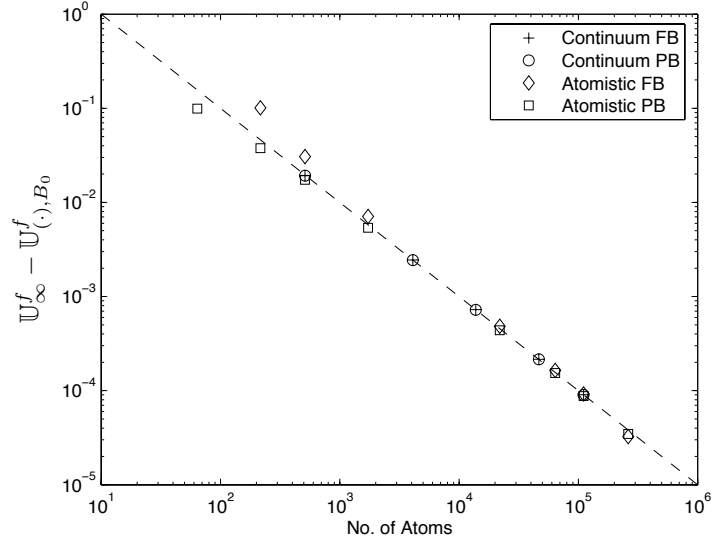


Figure 3.2: Difference between the strain energy of vacancy formation in an infinite body, U_∞^f , and strain energies of vacancy formation in a finite body, $U_{(\cdot), B_0}^f$, using either periodic (PB) or free (FB) boundary conditions.

where $S_{kl ij}$ is the constant compliance tensor. Note that Equation 3.13 is not directly comparable to atomistic calculations because σ_{ij} is the continuum stress field and contains a singularity, which does not exist in atomistic calculations. To evaluate the effect of boundary conditions we wish to use the divergence theorem to relate this to a surface integral. Therefore we write Equation 3.13 as

$$(3.14) \quad \Delta V_{kl} = S_{kl ij} \int_B \left[\frac{\partial(x_i \sigma_{mj})}{\partial x_m} - x_i \frac{\partial \sigma_{mj}}{\partial x_m} \right] dV.$$

It will become useful to divide the finite body, B , into a region enclosing the center of expansion or contraction, B_0 and a region surrounding it, B_1 . In this case, the body is finite so B_1 is a finite region. Then we write

$$(3.15) \quad \Delta V_{kl} = S_{kl ij} \left\{ \int_{B_0} \left[\frac{\partial(x_i \sigma_{mj})}{\partial x_m} - x_i \frac{\partial \sigma_{mj}}{\partial x_m} \right] dV + \int_{B_1} \left[\frac{\partial(x_i \sigma_{mj})}{\partial x_m} - x_i \frac{\partial \sigma_{mj}}{\partial x_m} \right] dV \right\}.$$

The equilibrium equation for the elastic dipole is

$$(3.16) \quad \frac{\partial \sigma_{jm}}{\partial x_m} + D_{jm} \frac{\partial \delta(\bar{x} - \bar{x}')}{\partial x_m} = 0.$$

Substituting this in Equation 3.15 gives

$$(3.17) \quad \Delta V_{kl} = \mathbb{S}_{klij} \left\{ - \int_{B_0} x_i D_{jm} \frac{\partial \delta(\bar{x} - \bar{x}')}{\partial x_m} dV + \int_B \frac{\partial (x_i \sigma_{mj})}{\partial x_m} dV - \int_{B_1} x_i \frac{\partial \sigma_{mj}}{\partial x_m} dV \right\}.$$

We evaluate the first term using the standard result for the spatial derivative of the Dirac-delta function, and the second term using the symmetry of the stress tensor and the Divergence Theorem. The third term drops out due to the fact that σ_{mj} is divergence-free over B_1 . Using these three results in Equation 3.17 we obtain

$$(3.18) \quad \Delta V_{kl} = \mathbb{S}_{klij} D_{ij} + \mathbb{S}_{klij} \int_S x_i \sigma_{jm} n_m dA.$$

With FBC, the surface integral vanishes and the volume change defines the relaxation volume tensor,

$$(3.19) \quad \Delta V_{kl} \equiv \mathbb{S}_{klij} D_{ij}.$$

The difference in ΔV_{kl} between FBC and arbitrary boundary conditions can be determined by evaluating the integral in Equation 3.18. However, in atomistic calculations with PBC it is more convenient to compare with a volume integral. Therefore we rewrite the second term of Equation 3.18 using a stress field $\langle \sigma_{jm} \rangle$, such that $\langle \sigma_{jm} \rangle = \sigma_{jm}$ in B_1 , and $\langle \sigma_{jm} \rangle$ is non-singular and divergence-free in B_0 . We then use the fact that $\langle \sigma_{jm} \rangle$ so-defined is divergence-free over all of B to obtain the expression

$$(3.20) \quad \Delta V_{kl} = V_{kl}^r + \mathbb{S}_{klij} \int_B \langle \sigma_{ij} \rangle dV,$$

where V_{kl}^r is the relaxation volume tensor. The field $\langle \sigma_{ij} \rangle$ is a good model for an atomistic stress field since it is non-singular and divergence-free over B and is equal to the elastic field in B_1 . In what follows, we will use the virial formulation for $\langle \sigma_{ij} \rangle$.

To compare Equation 3.20 to atomistic calculations with PBC we note that supercell relaxation occurs by changing the magnitude and direction of the vectors that

define the supercell. The change in these vectors defines an average strain relative to the perfect structure, so we find

$$(3.21) \quad \Delta V_{kl} = V^{ref} \varepsilon_{kl}^{av.} = V_{kl}^r + \mathbb{S}_{kl ij} \sigma_{ij}^{av.} V^{ref}.$$

The stress $\sigma_{ij}^{av.}$ is the average over B , as would be measured using the virial formulation, and since Equation 3.21 is derived using the assumptions of linear elasticity theory the integral is carried out over the entire undeformed volume, V^{ref} . This result shows that there are no elastic image effects, and it holds for defects of any anisotropy, and supercells of any shape or size. The observed volume change in atomistic calculations is the relaxation volume plus a correction term that arises if the system is not allowed to fully relax. For PBC with zero average stress, linear elasticity predicts that the observed volume change is exactly the relaxation volume. For self-equilibrated stress states in parallelepiped-shaped supercells the average surface stress must equal the resolved virial stress, so Equation 3.21 is equally valid if $\sigma_{ij}^{av.}$ is measured from the forces crossing the supercell boundaries. Using this fact, Equation 3.21 can also be derived for PBC by evaluating the surface integral in Equation 3.18.

This results of this section are a validation that the supercell approach does not introduce any systematic errors due to elasticity. In the following section we present atomistic calculations to show that in practice system size does effect the observed volume change. FBC introduce surfaces and PBC an infinite array of defects that change the elastic moduli and result in deviations from elasticity's prediction at small cell sizes.

3.2 Atomistic Calculations

We calculated the formation energy and volume tensor of a vacancy and a $\langle 110 \rangle$ dumbbell interstitial in Stillinger-Weber silicon by energy minimization using the conjugate gradient method. The Stillinger-Weber potential is a commonly used empirical potential for silicon [93, 94] and as such it is not as accurate as *ab initio* calculations near the defect, but is useful for our purposes since the decreased computational costs allow us to use the large system sizes necessary to check the predictions of continuum linear elasticity in the far-field.

3.2.1 Stillinger-Weber Potential

The Stillinger-Weber potential consists of two- and three-body terms:

$$(3.22) \quad \Phi = \sum_{i < j}^N v_2(r_{ij}) + \sum_{i < j < k}^N v_3(\bar{r}_i, \bar{r}_j, \bar{r}_k),$$

$$(3.23) \quad v_2(r_{ij}) = e f_2 \left(\frac{r_{ij}}{\xi} \right),$$

$$(3.24) \quad v_3(\bar{r}_i, \bar{r}_j, \bar{r}_k) = e f_3 \left(\frac{\bar{r}_i}{\xi}, \frac{\bar{r}_j}{\xi}, \frac{\bar{r}_k}{\xi} \right),$$

where the energy and length units e and ξ are chosen to give f_2 a minimum value of -1 if its argument is $2^{1/6}$. The two-body function, f_2 , depends only on the distance $r_{ij} = |\bar{r}_i - \bar{r}_j|$ between a pair of atoms with position vectors \bar{r}_i and \bar{r}_j and has a cut-off at $r = a$ without discontinuities in any derivatives with respect to r :

$$(3.25) \quad f_2(r) = \begin{cases} A(Br^{-p} - r^{-q}) \exp\left(\frac{1}{r-a}\right) & r < a \\ 0 & r \geq a \end{cases}$$

The three-body function, f_3 , depends on the scalar distances between the three atoms and also on the angles subtended at the vertices of the triangle formed by the three atoms. In the following relations Θ_{jik} is the angle subtended at vertex i between

atoms with position vectors \bar{r}_j and \bar{r}_k :

$$(3.26) \quad f_3 \left(\frac{\bar{r}_i}{\xi}, \frac{\bar{r}_j}{\xi}, \frac{\bar{r}_k}{\xi} \right) = h \left(\frac{r_{ij}}{\xi}, \frac{r_{ik}}{\xi}, \Theta_{jik} \right) + h \left(\frac{r_{ji}}{\xi}, \frac{r_{jk}}{\xi}, \Theta_{ijk} \right) + h \left(\frac{r_{ki}}{\xi}, \frac{r_{kj}}{\xi}, \Theta_{ikj} \right),$$

(3.27)

$$h(r_1, r_2, \Theta) = \begin{cases} \lambda \exp \left(\frac{\gamma}{r_1 - a} + \frac{\gamma}{r_2 - a} \right) \left(\cos \Theta + \frac{1}{3} \right)^2 & r_1 < a \text{ and } r_2 < a \\ 0 & \text{otherwise} \end{cases}$$

The equilibrium bond angle in the tetrahedral structure of silicon satisfies $\cos \Theta = -1/3$. The energetic contribution of any bond angle distortions are thus represented by the trigonometric term in h . The lattice spacing and bond energy of silicon at 0 K is obtained with $\xi = 2.0951 \text{ \AA}$ and $e = 2.3146 \text{ eV}$. Note that this follows other work in multiplying the value of e in the original Stillinger-Weber paper by 1.068 to match the cohesive energy of the experimental value [94, 95].

The optimized set of parameters for this potential is

$$\begin{aligned} A &= 7.049556277 & B &= 0.6022245584 \\ p &= 4, q = 0 & a &= 1.80 \\ \lambda &= 21.0 & \gamma &= 1.20 \end{aligned}$$

3.2.2 Methods

We calculated formation energies and volumes for cubic systems ranging in size from 64 to 262,144 atoms for the vacancy and 64 to 110,592 atoms for the interstitial. Note that creating the defect changes the number of atoms by one, but we will refer to system sizes by the number of atoms in a perfect crystal reference state. Separate calculations were run with periodic and free boundary conditions. The vacancy was constructed by removing an atom near the center of the simulation cell. The relaxation around a vacancy is inwards, see Figure 3.3, so we relaxed the four nearest neighbors to the vacancy towards the vacancy by 24% of the equilibrium atomic

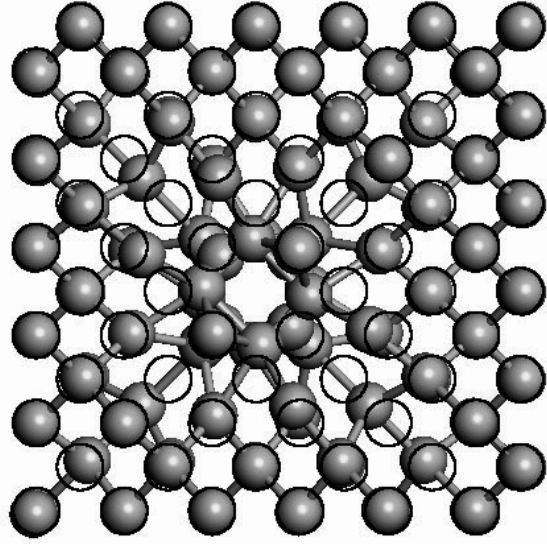


Figure 3.3: A vacancy in silicon viewed along the $\langle 100 \rangle$ direction. The open circles mark positions of atoms in the perfect crystal; i.e., before vacancy formation. Displacements have been scaled 10x for clarity.

distance to accelerate convergence and avoid the metastable configuration where atoms remain in the perfect crystal configuration. The $\langle 110 \rangle$ dumbbell interstitial was constructed by displacing an atom near the center of the simulation cell by $(-0.162, -0.162, +0.1325)$ unit cells and adding an interstitial that is displaced $(+0.162, +0.162, +0.1325)$ from the first atom's original position. Upon relaxation the atoms composing the dumbbell relax in the z direction away from the neighboring atoms in the $\langle 110 \rangle$ chain, so to speed convergence the dumbbell atoms were given the initial z displacement indicated above. The sign of the z displacement depends on which atomic site the dumbbell is located.

The $\langle 110 \rangle$ dumbbell interstitial in Stillinger-Weber silicon was found to have two different configurations with nearly equal formation energy, as shown in Figure 3.4. The major difference between the two is that the lower energy configuration, which we call (A), had non-zero V_{xz}^f and V_{yz}^f , while for the higher energy configuration,

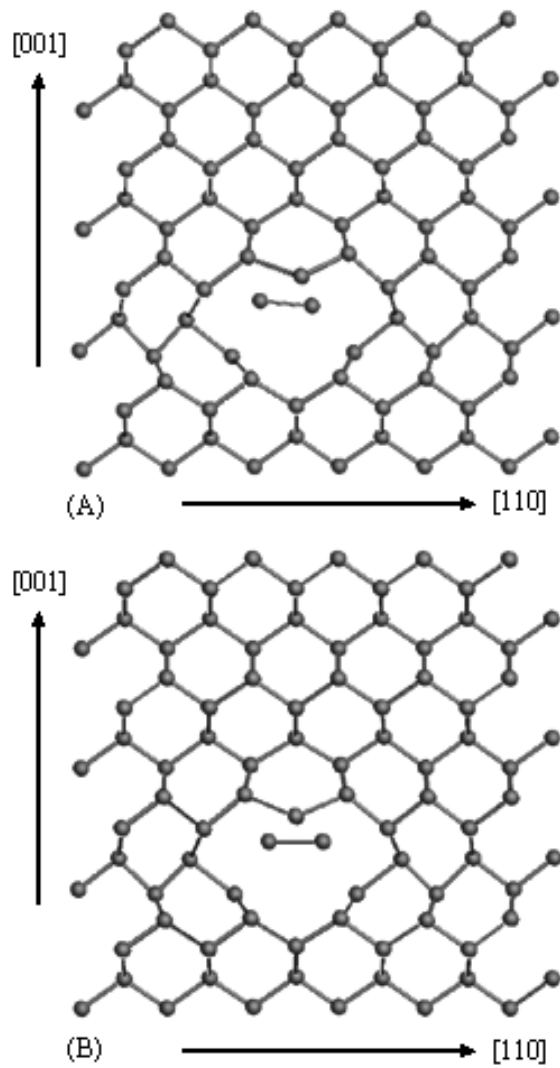


Figure 3.4: Structure of the relaxed Si $\langle 110 \rangle$ dumbbell interstitial with displacements scaled by 3x for clarity. In configuration (A), the dipole tilts and breaks the symmetry about the (110) plane, resulting in a slightly lower energy than in configuration (B). Both (A) and (B) maintain symmetry about the $(\bar{1}10)$ plane.

which we call (B), V_{xz}^f and V_{yz}^f are zero. As can be seen in Figure 3.4, the non-zero V_{xz}^f and V_{yz}^f is manifested locally by the dumbbell tilting and breaking the symmetry about the (100) plane. Both structures maintain symmetry about $(1\bar{1}0)$. For both FBC and PBC an initially perfect crystal with a defect, small random displacements, and less than 512 atoms relaxed to (A), with 512 atoms the crystal relaxed to (A) or (B), and with greater than 512 atoms the crystal became stuck in configuration (B). We attempted several schemes of increasing complexity to ensure minimization to the lower energy configuration (A) at large system sizes, and the successful method involved taking a relaxed (A) configuration at one system size and adding atoms at the surface to construct the next largest system size. The new atoms were positioned according to the final average strain state of the previous system.

Three boundary conditions were used: FBC, PBC at constant pressure (PBC CP), and PBC at constant volume (PBC CV) equal to the volume of the relaxed defect-free system. We check that there are no elastic image effects by comparing the FBC and PBC calculations, and check the second term in Equation 3.21 and Equation 3.11 by comparing PBC CP and PBC CV calculations. We also imposed small random initial displacements of approximately 1% of the atomic spacing on all the atoms and created ten samples for each system size, except for the largest vacancy systems for which there were five samples. The energy of the system was then minimized from this starting configuration using the conjugate gradient method. The minimization was considered complete when fourteen sequential iterations each resulted in less than a 1 neV reduction in energy, with the last seven sequential iterations also producing less than a 10^{-4} \AA^3 change in volume.

As noted in the previous section, at mechanical equilibrium with the parallelepiped-shaped computational cells used, a relaxation to zero surface traction is identical to

relaxation to zero average volumetric stress as measured by the virial formulation. However, in practice we found it preferable to use the virial stress for two reasons. First, in diamond cubic silicon before mechanical equilibrium is reached, xy -shearing results in a state of alternating positive and negative stress between (001) planes as the interpenetrating FCC lattices attempt to relax internally relative to each other. If the stress is only calculated at a single boundary, this gives an inaccurate measurement of the overall stress state and impedes convergence to zero average stress. Second, given the same nominal stress convergence criteria, the virial stress is stricter because it averages over the entire cell while the boundary stress only averages over the boundary. At cell sizes from 64 up to 21,952 atoms the energies and volumes measured using the zero virial stress condition matched the energies and volumes resulting from using the zero average surface traction condition in the periodic case or relaxation in the absence of constraints for free surface boundary conditions. However, the spread in values was reduced when the zero virial stress condition was employed. For these reasons we used the virial formulation to calculate an average stress tensor in the computational cell and zero average stress ($\pm 10^{-2}$ Pa) was maintained by scaling atomic positions and, if present, periodic boundaries. The elastic moduli of Stillinger-Weber silicon were used to adjust the strain on the system in order to maintain zero average stress during the relaxation process. In the FBC case, after the energy was minimized in this way, rescaling toward the zero average stress condition was discontinued and the energy was again minimized to reach a zero surface traction condition. We found that this method reduced scatter in the formation volume of FBC samples.

3.2.3 Measurements

As discussed previously, formation volume measurements are straightforward for periodic boundary conditions. Strain is defined by the position of the periodic boundaries, and each component of the relaxation volume is determined by multiplying the corresponding strain component by the perfect reference volume

$$(3.28) \quad V_{ij}^r = \varepsilon_{ij}^{av} \cdot V^{ref}.$$

Then, the formation volume is

$$(3.29) \quad V_{ij}^f = \varepsilon_{ij}^{av} \cdot V^{ref} \pm \frac{1}{3} \Omega \delta_{ij},$$

where (+/−) is for a (vacancy/interstitial) defect. For FBC, the volume change of an elastic body must be determined from the displacement of the surface of the sample according by

$$(3.30) \quad V_{ij}^f = \int_S \frac{1}{2} (u_i n_j + u_j n_i) dA \pm \frac{1}{3} \Omega \delta_{ij},$$

where u is the displacement, and n is the surface normal. In an atomistic simulation, this is a finite sum of individual atomic displacements and the differential area is the average surface area per surface atom. This method is not appropriate for PBC because it does not take into account strain between atoms on either side of the periodic boundary, which is small, but significant when multiplied over the area of the boundary.

We calculated elastic moduli for defect-free Stillinger-Weber silicon by using a single perfect unit cell with periodic boundaries. The bulk modulus, K , was calculated by measuring the volume change under hydrostatic pressure of $\sigma_h = -100$ MPa and $\sigma_h = 100$ MPa. Then,

$$(3.31) \quad K = \frac{\sigma_h}{\Delta V} V(\sigma = 0).$$

Additionally, σ_{xy} was measured for $\varepsilon_{xy} = \varepsilon_{yz} = 0.001$ so that C_{44} could be calculated,

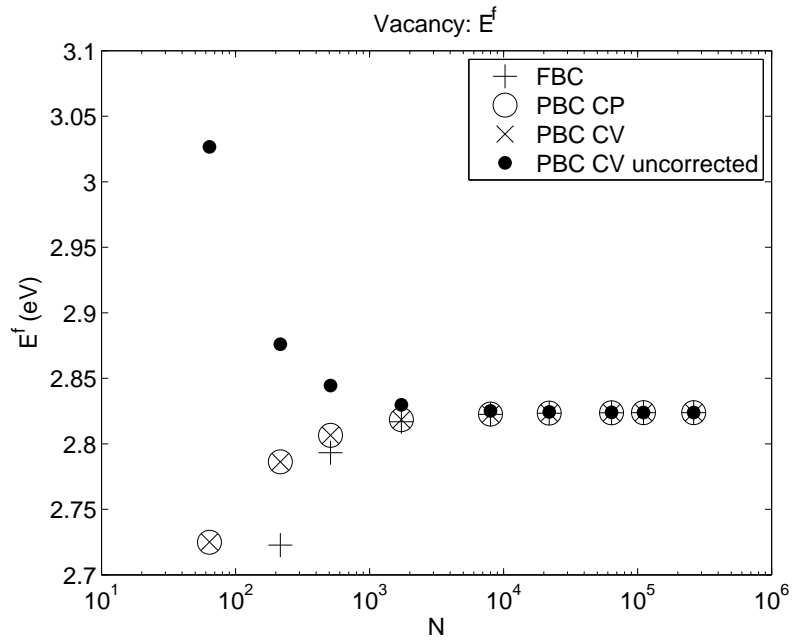
$$(3.32) \quad C_{44} = \frac{\sigma_{xy}}{2\varepsilon_{xy}}.$$

The diamond cubic structure can be thought of as two interpenetrating FCC lattices, which can relax internally and produce an internal strain. We measured C_{44} with internal strain because it does occur after a defect is introduced. We also calculated the bulk and shear moduli as a function of system size for both FBC and PBC systems with an interstitial. These systems were created for $N = 216$ to 110,592 atoms without any random initial displacements and tested as above, except stress control had to be used for the FBC case. In this case the strains were calculated using a relaxed system with an interstitial and the same number of atoms as a reference state.

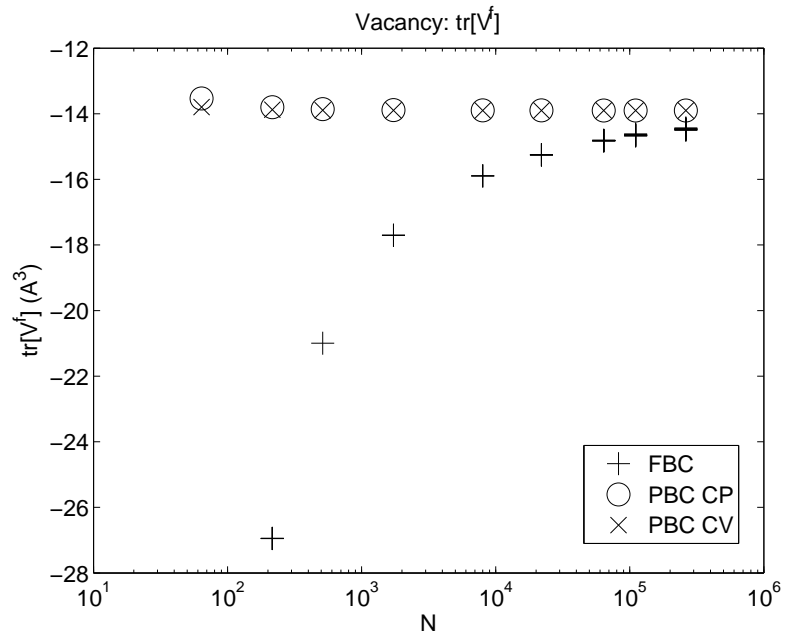
3.2.4 Results

As predicted by continuum linear elasticity, the formation energies and volume tensors calculated with FBC and PBC, at both constant pressure and volume, converge to the same value in the large-size limit, as shown in Figure 3.5 for the vacancy and Figure 3.6 for the $\langle 110 \rangle$ interstitial (A). The trace of the formation volume is shown, but each component of the volume tensor converged similarly. The 64 atom FBC samples underwent surface reconstructions and are not included in the plots or analysis. No other samples underwent surface reconstructions, and in the Stillinger-Weber potential the equilibrium distance and angle is not coordination dependent, so surface stress is not a factor in these results.

The formation energy converges more rapidly with PBC than FBC, as seen in Figures 3.5(a) and 3.6(a). As predicted by continuum linear elasticity, and shown in

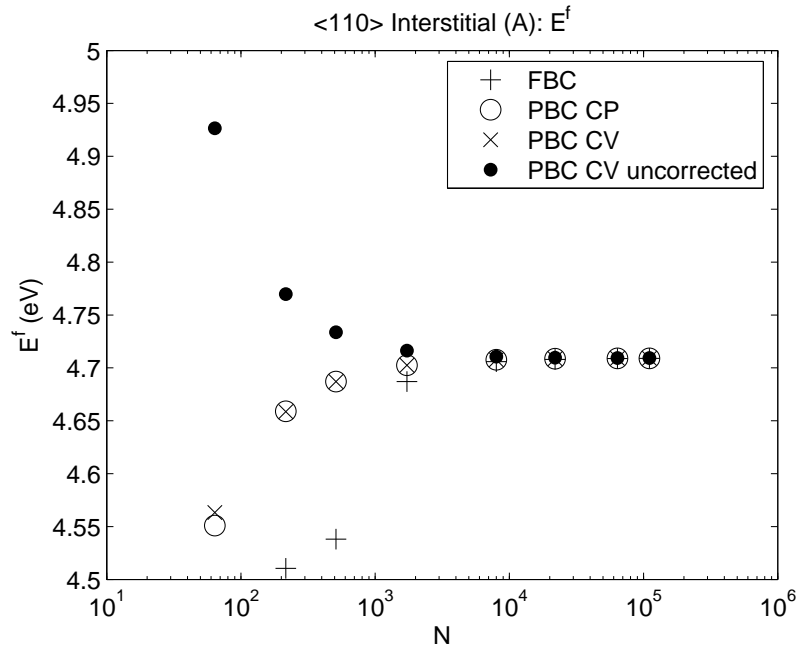


(a)

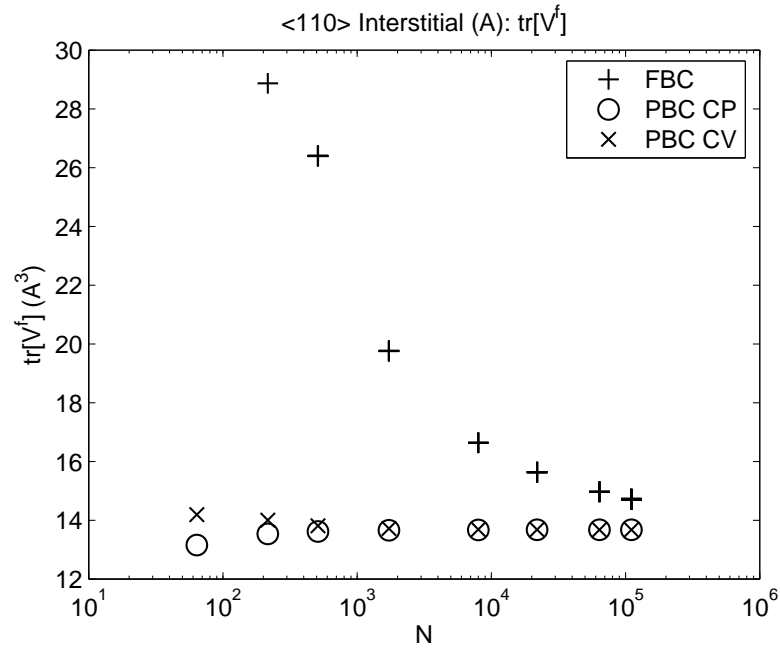


(b)

Figure 3.5: The (a) formation energy and (b) trace of the formation volume converge in the large-size limit for all boundary conditions. “PBC CV uncorrected” and “PBC CV” indicate formation energies before and after using Equation 3.11.



(a)



(b)

Figure 3.6: The (a) formation energy and (b) trace of the formation volume converge in the large-size limit for all boundary conditions. “PBC CV uncorrected” and “PBC CV” indicate formation energies before and after using Equation 3.11.

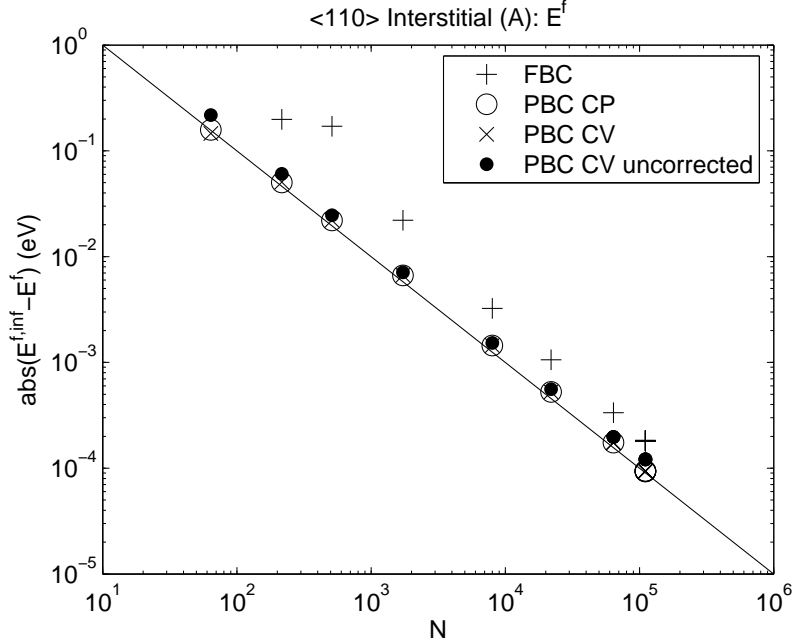


Figure 3.7: The $\langle 110 \rangle$ interstitial (A) formation energy converges in the large-size limit to $E_{SiI, \langle 110 \rangle, (A)}^{f, \infty} = 4.7091$ eV for all boundary conditions. “PBC CV uncorrected” and “PBC CV” indicate formation energies before and after using Equation 3.11.

Figure 3.7, formation energy converges in the large-size limit, with the error decreasing as $1/N$,

$$(3.33) \quad \log[E^f(N) - E^{f, \infty}] = \text{const} - \log(N).$$

The analogous result for the vacancy is shown in Figure 3.2. The converged formation energies, $E^{f, \infty}$, were determined from the formation energies at a given system size, $E^f(N)$, by fitting the data to Equation 3.34. The converged energy for the vacancy are $E_{SiV}^{f, \infty} = 2.8239$ eV. The converged energies for the interstitial are $E_{SiI, \langle 110 \rangle, (A)}^{f, \infty} = 4.7091$ eV, and $E_{SiI, \langle 110 \rangle, (B)}^{f, \infty} = 4.7122$ eV. These quantities are with uncertainty no greater than 10^{-4} eV.

The formation energies are in agreement with the values calculated elsewhere for the Stillinger-Weber potential [94, 95, 96, 97]. Some of the literature seems to confuse the $\langle 110 \rangle$ dumbbell and what is generally called the “extended” interstitial. The

extended interstitial is lower energy than the $\langle 110 \rangle$ dumbbell in empirical calculations [96, 98], but was found to be metastable in an *ab initio* calculation [75]. No other Stillinger-Weber results are known for the full $\langle 110 \rangle$ formation volume tensor.

Contrary to the prediction of continuum linear elasticity, the formation volume tensor was also found to converge with system size, the error decreasing as $1/N$, as shown in Figure 3.8(a) for the $\langle 110 \rangle$ interstitial (A). The trace of the formation volume tensor converges much more rapidly with PBC than FBC. The convergence of each component of the volume tensor is not shown but similar. The converged values of the formation volume tensor were determined similarly to the formation energies, and are

$$\begin{aligned} \bar{V}_{SiV}^{f,\infty} &= \begin{bmatrix} -4.633 & 0 & 0 \\ 0 & -4.633 & 0 \\ 0 & 0 & -4.633 \end{bmatrix} \text{Å}^3, \\ \bar{V}_{SiI,\langle 110 \rangle,(A)}^{f,\infty} &= \begin{bmatrix} 9.585 & 7.972 & \pm 2.176 \\ 7.972 & 9.585 & \pm 2.176 \\ \pm 2.176 & \pm 2.176 & -5.493 \end{bmatrix} \text{Å}^3, \\ \bar{V}_{SiI,\langle 110 \rangle,(B)}^{f,\infty} &= \begin{bmatrix} 8.871 & 7.296 & 0 \\ 7.296 & 8.871 & 0 \\ 0 & 0 & -4.849 \end{bmatrix} \text{Å}^3. \end{aligned}$$

with uncertainty in the components of $\bar{V}^{f,\infty}$ no greater than 10^{-3}Å^3 . The (\pm) for (A) indicates that it is physically equivalent for the tilt to be in either direction since the sign of shear strains is arbitrary.

The principle axes of the $\langle 110 \rangle$ (A) formation volume tensor are tilted 7.5° off the (001) plane. To our knowledge, this is the first report of the tilted $\langle 110 \rangle$ interstitial. This may be due to a focus on the energy of the defect rather than on the structure

of the relaxation in previous Stillinger-Weber studies. In a recent *ab initio* study [65] there was not any tilting in the $\langle 110 \rangle$ interstitial, despite allowing the full relaxation of the supercell, indicating that the tilt is likely to be an artifact of the Stillinger-Weber potential.

Continuum linear elasticity's inability to predict the formation volume tensor convergence with system size is due to its assumption that the elastic moduli are constant. The slow convergence of the formation volume tensor with FBC is caused by the slow convergence of elastic moduli, as shown in Figure 3.8. Decreased coordination of surface atoms results in decreased stiffness. The similarity to the formation volume convergence is apparent. Note that Figure 3.8 plots only the absolute values of the convergence, therefore the direction of the convergence can not be determined from the figure. We observed that in systems with FBC the elastic moduli increase with system size, matching the observed decrease with system size in the magnitude of the volume relaxation. In other words, as the moduli increase with system size the outward relaxation around the interstitial decreases in order to reduce the strain energy in the surrounding system. In systems with PBC, due to bonding changes at the defect there is a slight decrease in the bulk modulus and a slight increase in the shear modulus with increasing system size. The trace of the formation volume tensor shows an increase in the magnitude of the relaxation around the interstitial with increasing system size, corresponding to the decrease in the bulk modulus. Thus, the convergence of the elastic moduli in the system with a defect to the elastic moduli of the defect-free system is an indication of the convergence of the formation volume tensor and a form for the error can be expected. The elastic moduli convergence is in agreement with the results of Windl *et al.* [66] and suggests that either the precision of those calculations or electrostatic effects are hiding the associated formation

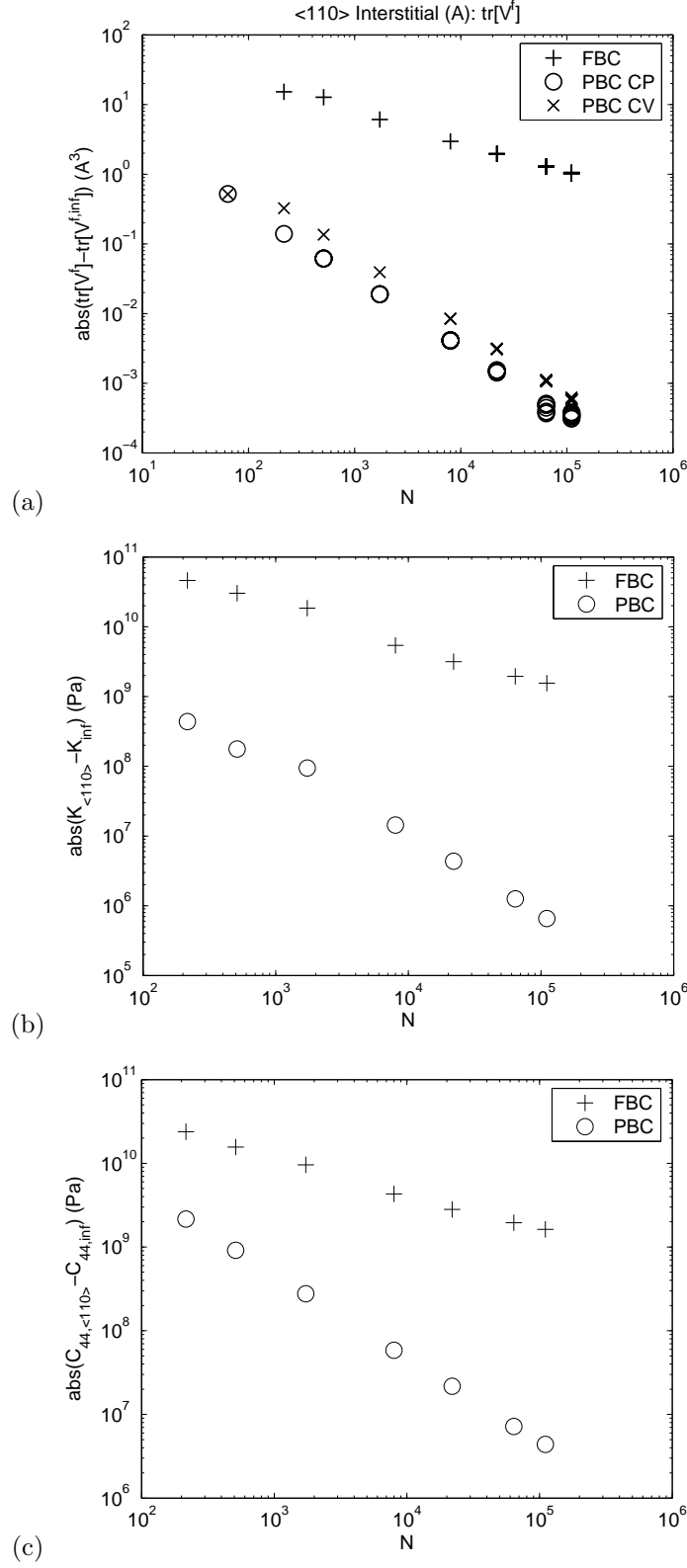


Figure 3.8: The convergence of $\text{tr}[V^f]$ for the $\langle 110 \rangle$ interstitial (A) depends on the convergence of the (b) bulk modulus ($K_{inf} = 1.0826 \times 10^{11}$ Pa) and (c) shear modulus ($C_{44,inf} = 6.0256 \times 10^{10}$ Pa).

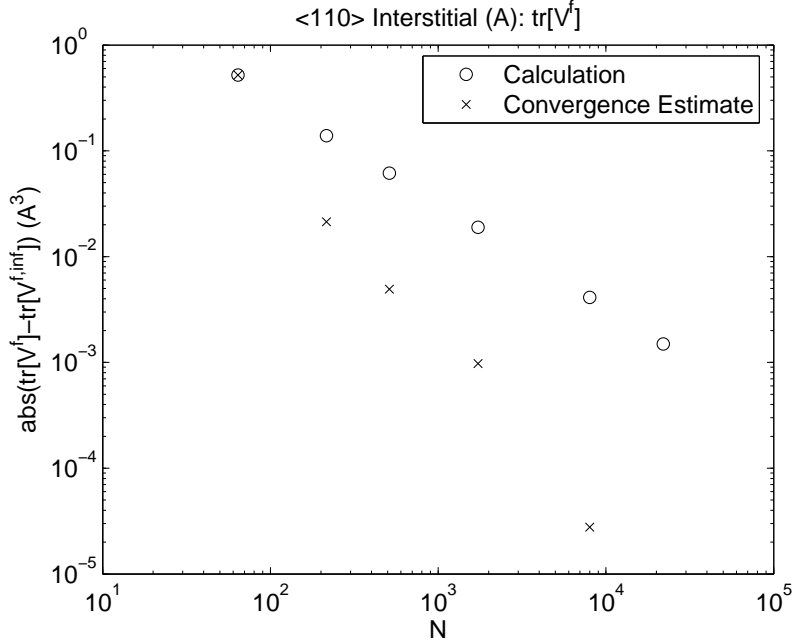


Figure 3.9: The formation volume converges to $\text{tr}[V_{SiI,(110),(A)}^{f,\infty}] = 13.677 \text{ \AA}^3$ with error decreasing as $1/N$ shown above with \circ . Knowing this convergence rate, we can use Equation 3.34 to estimate a converged value from the two largest systems calculated. This estimate, which is shown with \times , is as converged at $N = 1728$ as the $N = 21,952$ calculation.

volume convergence. Similar to the approach of Castleton and Mirbt [99], we can use the convergence rate to estimate the final converged formation volume tensor. At this convergence rate, with formation volume calculations for systems with N_1 and N_2 atoms, the estimated formation volume in the large-size limit is

$$(3.34) \quad V^{f,\infty}(N_2) = \frac{N_1 V^f(N_1) - N_2 V^f(N_2)}{N_1 - N_2}.$$

Figure 3.9 shows how this estimate converges with system size for the $\langle 110 \rangle$ interstitial (A), allowing us to estimate the converged formation volume within the accuracy of the measurements, 10^{-3} \AA^3 , by extrapolating from the $N_1 = 512$ and $N_2 = 1728$ systems.

Finally, the agreement between calculations with PBC CP and PBC CV in Figures 3.5 and 3.6 demonstrates that correction terms calculated using linear elasticity do hold, at least to a good approximation. When the system is not allowed to fully

relax we can adjust the formation volume tensor by using Equation 3.21, and we can adjust the formation energy to account for elastic strain energy by using Equation 3.11. At small system sizes, Figures 3.5 and 3.6 show that there are small differences between PBC CP and PBC CV which can be attributed to using the elastic moduli of a perfect system in Equations 3.21 and 3.11, rather than the actual moduli of the system with a defect, to strain dependence of the moduli, or strain dependence of the defect's strength. As the system size increases these effects decrease and the calculations converge in the large-size limit.

A nonlinear elasticity theory might be parameterized to model the convergence of the volume tensors with cell size. Finite strains occur in the near field of the defect, are rapidly varying, and are strongly dependent on the type of defect. Therefore, a nonlinear elasticity model, most of which have only a small number of free parameters, would need to be fit afresh for each type of defect. The predictive capability of commonly used nonlinear elasticity models therefore is somewhat limited in the near field of defects. Rather than try to predict the convergence rate and extrapolate from fully atomistic calculations, it might be simpler to use a “hybrid” method that links a fully atomistic calculation around the defect to a continuum mechanics description of the surrounding material [100]. Currently, hybrid methods using accurate density functional theory (DFT) descriptions of the atomistic region are limited to metals [101], and semiempirical tight binding is used for non-metals [102]. The size-dependence of the elastic constants is also a reminder that defect density affects elastic constants and that changes in elastic constants affect formation volume tensors. At high defect concentrations, continuum simulations of defect formation and diffusion may need to take this into account using data from calculations at high defect densities.

3.3 Conclusions

The work presented in this chapter validates the supercell approach for calculating formation energies and volume tensors. We have formally shown that, according to linear elasticity, the formation energy converges with the system size and that the calculated relaxation volume tensor of any anisotropic defect in any shape or size supercell is not affected by PBC at zero average stress. This rigorously demonstrates why the supercell approach can provide accurate calculations of defect parameters. When the supercell is not allowed to fully relax, the average stress can be used to calculate the relaxation volume tensor and formation energy. Atomistic calculations verify linear elasticity's predictions in the far-field for both an isotropic vacancy and an anisotropic $\langle 110 \rangle$ interstitial in Stillinger-Weber silicon, and show that, in practice, bonding changes at the defect result in elastic moduli changes. The observed $1/N$ decrease of the error of the relaxation volume tensor is due to the convergence of the elastic moduli. Knowledge of this convergence rate allows for accurate estimation of the relaxation volume tensor with relatively modest simulation sizes.

CHAPTER 4

Modeling Collective Diffusion Behavior

As we noted in the introduction, KMC simulations are an essential tool for modeling the complex behavior exhibited by collections of diffusing species. Dopant diffusion is significantly affected by concentration, stress, time and the underlying crystal lattice. In some cases, KMC simulations are used to calculate diffusivities that can be used in continuum models. In other cases, KMC simulations are used to directly determine dopant distributions. This may be especially relevant under non-equilibrium conditions and at small scales where deviations from average behavior are more important. In this chapter we will first discuss how continuum diffusivities are calculated from microscopic hops rates and numerical simulations of atomic motion. Then we will review the experimental and computational literature for dopant diffusion as a function of concentration, time, and stress. Finally, we will parameterize a model for vacancy-mediated As diffusion from *ab initio* literature and consider the difficulties that occur when using the standard KMC method to calculate diffusivity.

4.1 Calculating Diffusivity

Although atomic motion is driven by chemical potential gradients rather than concentration gradients, it is concentration that is typically directly observable. Fick's

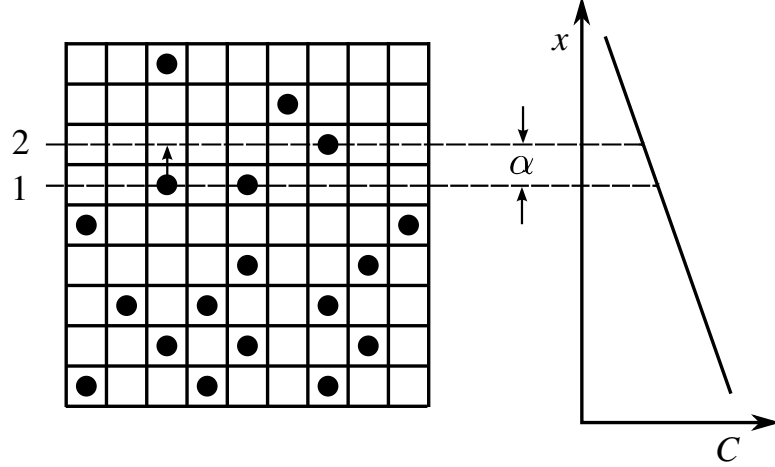


Figure 4.1: Random jumping on a square lattice. Figure as in [18].

first law,

$$(4.1) \quad \bar{J} = -\bar{D}^C \bar{\nabla} C,$$

relates the atomic flux vector \bar{J} to the concentration gradient $\bar{\nabla} C$ through the chemical diffusivity tensor \bar{D}^C .

Fick's first law essentially describes how random particle motion causes a net flux down a concentration gradient. Equation 4.1, an empirical observation that holds in many systems, can be derived by considering random atomic hopping on a cubic lattice [10, 18]. Let there be two neighboring planes, as in Figure 4.1, one on the left with a concentration n_1 atoms per m^2 and the other on the right with n_2 atoms per m^2 , separated by a distance α m. The atoms jump from plane to neighboring plane with a jump rate Γ s^{-1} , and $\Gamma = 6R$, where R is the transition rate of an individual jump. Then, the number of atoms per m^2 per s jumping from left to right is $J_{1 \rightarrow 2} = \frac{1}{6} n_1 \Gamma$, where the factor $\frac{1}{6}$ comes from an assumption that in three dimensions it is equally likely for an atom to jump in any of the six directions. Similarly, the number of atoms per m^2 per s jumping from right to left is $J_{2 \rightarrow 1} = \frac{1}{6} n_2 \Gamma$. The net flux of atoms per m^2 per s from left to right is $J = J_{1 \rightarrow 2} - J_{2 \rightarrow 1} = \frac{1}{6} (n_1 - n_2) \Gamma$.

The concentration of atoms per m^3 , C , is related to the number of atoms per m^2 on a plane by $n = \alpha C$. Since the concentration gradient is $\partial C / \partial x = (C_2 - C_1) / \alpha$, it follows that

$$(4.2) \quad J = - \left(\frac{1}{6} \Gamma \alpha^2 \right) \frac{\partial C}{\partial x}.$$

This is Fick's first law, with isotropic diffusivity

$$(4.3) \quad D^C = \frac{1}{6} \Gamma \alpha^2,$$

and holds for any cubic lattice if α is the individual jump distance, as we'll show below following the example of Einstein [18, 10, 103]. In order to derive Equation 4.3 off a simple cubic lattice, we will relate the mean squared displacement of a discrete random walker to that of the continuum solution for diffusion away from a point. This exercise will also help illuminate the general approach for determining continuum diffusivity from microscopic hops.

First, we calculate the squared displacement of a random walker that has made n jumps, by

$$(4.4) \quad R_n^2 = \left(\sum_i^n \bar{r}_i \right)^2,$$

where \bar{r}_i is the vector of magnitude r_i for jump i . Multiplying out the sums and regrouping the terms we can write

$$(4.5) \quad R_n^2 = \sum_i^n r_i^2 + 2 \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \bar{r}_i \cdot \bar{r}_{i+j} = \sum_i^n r_i^2 + 2 \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} r_i r_{i+j} \cos \theta_{i,i+j}.$$

If each jump is of equal magnitude α , and we average over the ensemble of diffusing particles, then the mean squared displacement is

$$(4.6) \quad \langle R_n^2 \rangle = n \alpha^2 \left(1 + \frac{2}{n} \left\langle \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \cos \theta_{i,i+j} \right\rangle \right).$$

The term in parentheses is called the correlation factor, f^* , and $\langle \rangle$ indicates an ensemble average. For random walkers on a cubic lattice each jump direction is uncorrelated with the others and the average of the double sum will be zero, so that $f^* = 1$ and

$$(4.7) \quad \langle R_n^2 \rangle = n\alpha^2.$$

Next, we determine the continuum solution for diffusion from a point. Mass conservation is given by the condition

$$(4.8) \quad \frac{\partial C}{\partial t} = -\bar{\nabla} \cdot \bar{J},$$

which with Equation 4.1 gives Fick's second law,

$$(4.9) \quad \frac{\partial C}{\partial t} = \bar{\nabla} \cdot (\bar{D}^C \bar{\nabla} C).$$

When there are multiple diffusing species Equation 4.9 is generalized

$$(4.10) \quad \begin{aligned} \frac{\partial C_A}{\partial t} &= \bar{\nabla} \cdot (\bar{D}_{AA}^C \bar{\nabla} C_A) + \bar{\nabla} \cdot (\bar{D}_{AB}^C \bar{\nabla} C_B) \\ \frac{\partial C_B}{\partial t} &= \bar{\nabla} \cdot (\bar{D}_{BA}^C \bar{\nabla} C_A) + \bar{\nabla} \cdot (\bar{D}_{BB}^C \bar{\nabla} C_B), \end{aligned}$$

where A and B indicate different species. When the chemical diffusivity is constant with concentration, $\bar{D}^C = D^C$, Fick's second law for a single species can be written

$$(4.11) \quad \frac{\partial C}{\partial t} = D^C \nabla^2 C,$$

where ∇^2 is the Laplace operator. Consider a fixed concentration of particles located at the origin $r = 0$ at time $t = 0$ which diffuse outward in three dimensions. It can be verified that a solution to Equation 4.11 in this case is

$$(4.12) \quad C(r, t) = (A/t^{3/2}) \exp(-r^2/4D^C t),$$

where A is a constant. Then the probability of finding a particle at time t a distance between r and $r + dr$ from the origin is

$$(4.13) \quad \rho(r, t)dr = \frac{C(r, t)4\pi r^2 dr}{\int_V C(r, t)} = \frac{r^2}{2\sqrt{\pi}(D^C t)^{3/2}} \exp(-r^2/4D^C t)dr.$$

The mean square displacement can be calculated as

$$(4.14) \quad \langle r^2 \rangle = \int_0^\infty r^2 \rho(r, t)dr = 6D^C t.$$

Equating Equations 4.7 and 4.14 and substituting $\Gamma = n/t$ returns the original result, Equation 4.3. In this simple example of a random walker the continuum chemical diffusivity is calculated from the microscopic jumps through the mean squared displacements.

Of course, this derivation of Equation 4.3 does not hold when jumps are correlated. Correlation effects may arise from anisotropy or non-random jumping. Even if each individual hop rate is the same in a non-cubic lattice, the diffusion rates will be different in different directions due to anisotropy [11, 13]. Similarly, correlation effects arise when jump rates are anisotropic due to the lattice or anisotropic stress. Correlation also occurs at high concentrations, when atoms can no longer be viewed as isolated walkers on a lattice because some neighboring spots are occupied [12]. The high concentration limiting case is when diffusion occurs by a vacancy mechanism. In this case, jumping is not random because after an atom jumps into the vacancy it is more likely for it to jump back into the vacancy again than to jump in a different direction. Additionally, when the system is not thermodynamically ideal, gradients in chemical potential will result in diffusion that is not random.

The standard approach [104, 105, 106] for generally determining the diffusion coefficient in terms of microscopic hops is similar in concept, if not technique, to the approach presented above. The continuum diffusivity is related to the dissipa-

tion of a concentration fluctuation at equilibrium, because the same phenomenon, thermally activated motion, is responsible for macroscopic transport and for the concentration fluctuations that occur in equilibrium. This allows for use of the fluctuation-dissipation theorem [107] to relate the generalized forces in the diffusion equation, $\nabla^2 C$, to time correlation functions of the concentration fluctuations. The concentration fluctuation correlation functions are in turn related to individual particle trajectory correlation functions. A thermodynamic factor is included to account for differences between the concentration gradient, which is more readily observed experimentally, and the chemical potential gradient, which actually drives diffusion. We simply present the results here and do not repeat the derivation.

The chemical diffusion coefficient is a function of the jump diffusion coefficient, \bar{D}^J , and the thermodynamic factor, Θ , which was defined in Section 2.1, by

$$(4.15) \quad \bar{D}^C = \bar{D}^J \Theta.$$

The diffusion coefficients and thermodynamic factor may be concentration dependent. When the system is thermodynamically ideal or in the dilute limit $\Theta = 1$. The jump diffusion coefficient tensor is

$$(4.16) \quad D_{jk}^J = \frac{1}{2t} \left\langle \frac{1}{N} \left(\sum_i^N x_j(i) \right) \left(\sum_i^N x_k(i) \right) \right\rangle,$$

where $x_j(i)$ is the j -component of the displacement vector $\bar{x}(i)$ of the i particle in time t , and N is the number of particles. If the diffusion is isotropic, then

$$(4.17) \quad D^J = \frac{1}{2dt} \left\langle \frac{1}{N} \left(\sum_i^N \bar{x}(i) \right)^2 \right\rangle,$$

where d is the dimension of the space. Another commonly used quantity is the tracer diffusion coefficient, in tensor

$$(4.18) \quad D_{jk}^* = \frac{1}{2t} \frac{1}{N} \sum_i^N \langle x_j(i) x_k(i) \rangle,$$

or scalar form

$$(4.19) \quad D^* = \frac{1}{2dt} \frac{1}{N} \sum_i^N \langle x(i)^2 \rangle,$$

where $x(i)$ is the magnitude of $\bar{x}(i)$. The values of D^J and D^* are only equivalent when the cross terms in Equation 4.17 average to zero, meaning there are no correlations between different particles. The tracer diffusion coefficient can be related to the chemical and jump diffusion coefficient through a correlation factor, f , by

$$(4.20) \quad D^C = D^J \Theta = f D^* \Theta.$$

In this case the jumps are not restricted to all being the same magnitude, so f is not exactly equivalent to f^* as defined in Equation 4.6.

For the interdiffusion of multiple species, Equation 4.15 can be generalized to

$$(4.21) \quad D_{ABkl}^C = D_{ABkl}^J \Theta_{AB},$$

where A and B run over different diffusing species, and k and l run over the coordinates. In this case, the thermodynamic factor is

$$(4.22) \quad \Theta_{AB} = \frac{1}{k_B T} \frac{\partial \mu_A}{\partial C_B},$$

and the jump diffusion coefficient becomes

$$(4.23) \quad D_{ABkl}^J = \frac{1}{2tV} \left\langle \left(\sum_i^{N_A} x_k^A(i) \right) \left(\sum_j^{N_B} x_l^B(j) \right) \right\rangle,$$

where V is the volume.

With the results for the jump diffusion coefficient, the calculation of continuum diffusivities depends on determining individual particle trajectories. The use of the fluctuation-dissipation theorem in the derivation of these expressions relies on time correlation functions in equilibrium at long time and extended length scales limit.

Especially at the extremely small length scales and in the highly non-equilibrium systems present in dopant diffusion, it is not clear how well these results for continuum diffusivities hold up. In many cases, direct simulation of dopant distributions may be desired. In either case, if continuum diffusivities or direct simulations are desired, the task of computational methods becomes one of calculating the individual particle trajectories.

4.2 The Master Equation

Since dopant diffusion is the result of many rare event transitions, it is natural to model it as a discrete state system. The time evolution of a system through its possible states is determined by the master equation

$$(4.24) \quad \frac{\partial}{\partial t} P_j(t) = \sum_i (R_{i \rightarrow j} P_i(t) - R_{j \rightarrow i} P_j(t)),$$

where i runs over all states except j ; P_j is the probability of being in state j , and is a function of time, t ; and $R_{i \rightarrow j}$ is the jumping rate from state i to state j . The solution of the master equation, $P_j(t)$ for all j , gives the probability of being in any state of the system as a function of time. At long times, the probability of being in any state will equal the equilibrium probability, P_j^{eq} . If we desire a direct simulation of dopant distribution as a function of time, we could sample the states of the system proportionally to $P_j(t)$. If we desire continuum diffusivities, the solution allows calculation of the ensemble averages needed for the jump diffusion coefficient.

For example, Equation 4.17 becomes

$$(4.25) \quad D^J = \frac{1}{2dt} \sum_j^{N_S} \left(P_j^{eq} \sum_k^{N_S} \left(\frac{P_k(t|j, 0)}{N} \left(\sum_i^N \bar{x}_{j \rightarrow k}(i) \right)^2 \right) \right),$$

where N_S is the number of states in the system, P_j^{eq} is the probability of being in state j at equilibrium, $P_k(t|j)$ is the probability of being in state k at time t given

that the system was in state j at time $t = 0$, and $\bar{x}_{j \rightarrow k}(i)$ is the displacement vector of particle i from state j to k .

Of course, in most interesting systems the number of possible states is too large to allow for a full solution of the $P_j(t)$. It can be solved for a limited number of states to approximate correlation factors [11, 12]. Ishioka and Koiwa [108] presented a method for determining diffusion coefficients from transition rates. Daw *et al.* [14] generalized the work of Dederichs and Schroeder [13] to calculate stress-affected anisotropic diffusivities for defects or defect-dopant clusters by expanding the master equation in plane waves and looking at their evolution in the long time and extended length scale limit. This is an alternative method for calculating continuum diffusivities from the microscopic hops. It is limited to dilute systems and does not account for cluster formation or dissolution.

4.3 Kinetic Monte Carlo Simulations

For direct simulations and situations too complex for analytical solutions, KMC simulations are a useful way of numerically sampling the solution to the master equation. With the standard KMC method [16, 17, 109, 110], each simulation is one random sample of the possible trajectories through states of the system over time. The trajectory is determined by advancing the system in time state by state according to appropriate probabilities. The probability of jumping from state i to state j is

$$(4.26) \quad P_{i \rightarrow j} = \frac{R_{i \rightarrow j}}{\sum_k R_{i \rightarrow k}}.$$

At each step, all of the possible events taking the system away from the current state are enumerated and their rates are summed, $Q = \sum_i R_i$. By choosing a random number, $rand1 \in (0, Q]$, the next event can be determined probabilistically by picking

the event j such that

$$(4.27) \quad \sum_{i=1}^{j-1} R_i < rand1 \leq \sum_{i=1}^j R_i.$$

Then, assuming the events are Poisson processes, the timestep Δt is determined by

$$(4.28) \quad \Delta t = -\frac{\ln(rand2)}{Q},$$

using a second random number, $rand2 \in (0, 1]$. By averaging over many particles, or many simulations, diffusivities and other quantities can be calculated.

4.4 Concentration and Time Dependence

The interactions of many diffusing particles often leads to concentration and time dependent diffusion behavior. In particular, concentration and time effects are important in modeling transient enhanced diffusion (TED) and understanding the concentration dependence of dopant diffusion at very high doping levels.

We briefly introduced the idea transient enhanced diffusion (TED) in Section 2.1.2.3 in the context of boron diffusion parameterization. Cowern and Rafferty [74] review many of the causes and effects of TED. During ion implantation a supersaturation of Si interstitials results in the formation of Si interstitial clusters, or extended rodlike $\{311\}$ defects in which Si interstitials cluster in $\langle 110 \rangle$ directions on $\{311\}$ planes. During annealing the clusters and $\{311\}$ defects emit Si interstitials and eventually dissolve. Until the dissolution is complete, the Si interstitial concentration remains above the equilibrium level and there is a corresponding increase in B diffusion. Additionally, Si interstitials can bind to multiple B atoms and form immobile clusters. The binding with B is tighter than to other Si interstitials, so Si interstitials in those clusters are emitted last. Small predominantly B clusters remain until they capture Si interstitials and emit mobile B-interstitial pairs. In order to predict B diffusion

the formation, and eventual dissolution of all of these clusters must be understood and modeled as a function of implantation energy and dose, annealing time and temperature, and device geometry. KMC simulations have been used for this purpose.

Pelaz *et al.* [111] and Caturla *et al.* [112] have performed KMC simulations of the interaction and diffusion of B and interstitials after implantation. These models are similar in that they model the diffusion and clustering of B and Si interstitials and vacancies. The inputs include migration energies or rates for the different species or complexes and reaction energies which are used to determine clustering and dissociation rates. Pelaz *et al.* used a combination of *ab initio* calculations and empirical fits to experimental observations to determine the energetics. Caturla *et al.* used atomistic calculations to determine energetics and experimental data for diffusion prefactors. Pelaz *et al.* and Caturla *et al.* used a implantation simulation to determine the initial defect and dopant positions. Then the KMC simulations determine the evolution of the species and complex concentration profiles. Pelaz *et al.* found that BI_2 clusters are important precursors to the formation of relatively stable and immobile large B clusters. A BI_2 cluster are not stable for long times, but long enough that it captures other mobile B interstitials to create larger more stable clusters. Caturla *et al.* showed that initially after implant about 56% of B is in BI and BI_2 clusters, and vacancies are in small clusters. Then, the vacancy clusters dissolve and vacancies combine with the BI and BI_2 clusters resulting in about 98% of B taking substitutional positions. After the vacancies have all recombined with interstitials or the surface, interstitial and B-interstitial clusters form and B diffusion greatly increases. This indicates that a significant amount of TED diffusion may occur during the initial supersaturation of interstitials before extended interstitial clusters form and then dissolve. Both groups found that they were able

to reproduce experimental data from their calculations. These results demonstrate that KMC simulations are useful for determining the active mechanisms in highly non-equilibrium processes at scales that are difficult to observe experimentally. The resulting B diffusion depends greatly on the concentration of B and intrinsic defects and their evolution with time.

Concentration and time dependence is also important in dopant diffusion at very high doping levels. Several experiments with different results have been made for As diffusion as a function of dopant concentration. Fair and Weber [39] observed increasing As diffusivity with As concentration up to about $3 \times 10^{20} \text{ cm}^{-3}$ when it began to decrease. Their diffusion anneals were at 1000 °C for hours. Larsen *et al.* [84] found that As diffusivity rapidly increased at P doping levels above $2 \times 10^{20} \text{ cm}^{-3}$. They used rapid thermal annealing at 1050 °C for 10 s. Solmi and Nobili [113] reported As diffusivity remained constant or decreased for concentrations above about $3 \times 10^{20} \text{ cm}^{-3}$ for anneals at 1050 °C and 900 °C, respectively, for 15 min. to 1 hr. Data from these studies is plotted in Figure 4.2.

Mathiot and Pfister [114] considered that at high doping levels V may feel interactions with multiple As, lowering their formation energy and migration barriers. The combination of higher V concentrations and mobilities would lead to an increase in diffusivity at the concentration when the dopant-V interactions begin to percolate throughout the material. They showed if the dopant-V interaction ranges 3 nearest neighbor positions this could occur at a dopant concentration of 3^{20} cm^{-3} in diamond cubic silicon which has a site concentration of $5 \times 10^{22} \text{ cm}^{-3}$. Xie and Chen [85] reviewed these results and performed *ab initio* calculations of vacancy formation and migration energies near As. Their results confirmed that placing a second As nearby lowered the migration barriers for a vacancy to travel from one to the other.

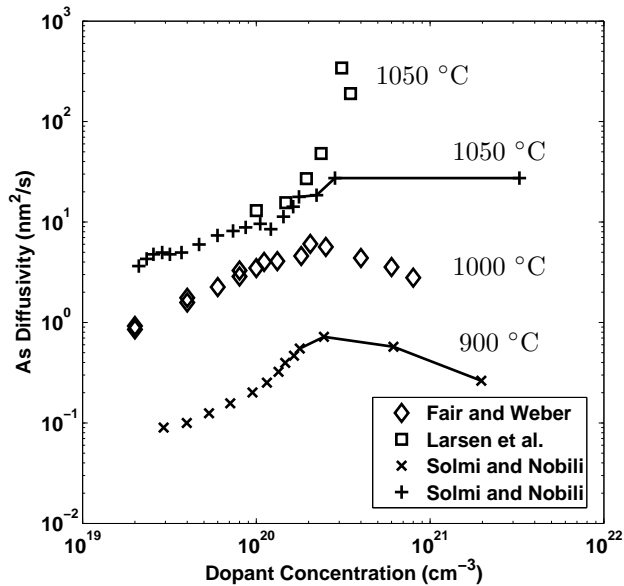


Figure 4.2: Literature results for As diffusivity as a function of dopant concentration. Lines for the Solmi and Nobili data indicate the trend where many data points are reported.

They also found that the migration barrier for a V_1As_2 cluster is 2.0 eV, compared to 1.19 eV for a V_1As_1 cluster. Therefore, they proposed that it is the combination of the concentration and clustering over time that produces the observed differences in diffusion behavior. Larsen *et al.* observed enhanced diffusivity because at very short times low vacancy migration barriers due to percolating dopant-vacancy interactions would enhance diffusivity. At longer times, such as in the experiments of Fair and Weber clusters would form and the high migration barriers for clusters would decrease diffusivity. The low temperature experiments of Solmi and Nobili would behave similarly, but in the higher temperature experiments some of the clusters could become mobile and keep diffusivity constant. Time dependent clustering and decreasing diffusivity for larger clusters is consistent with the positron annihilation experiments of Ranki *et al.* discussed in Section 2.1.1.1.

KMC simulations by Dunham and Wu [115] showed enhanced dopant diffusivity at high dopant concentrations. Their simulations used random dopant placements

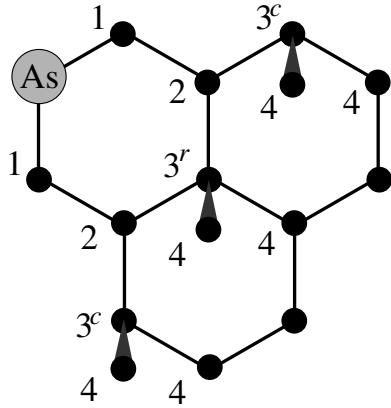
and, based on those positions and the As-vacancy interaction energy, equilibrium vacancy concentrations. They do not report simulation times, only noting that diffusivity began to drop at longer times so the reported values are for short times. List and Ryssel [116] performed longer time simulations (reported up to 4×10^6 $1/f_0$, f_0 being the vacancy jump frequency in pure silicon, so likely on the order of 1 μ s), and found dopant diffusivity dropped at high concentrations. Their simulations attempted to use equilibrium dopant distributions and vacancy concentrations. It is not necessarily clear that defect concentrations in the experiments are the equilibrium values. More recently, Haley and Gronbech-Jensen [117] performed KMC simulations for times up to milliseconds which varied the temperature, As concentration, vacancy concentration, As-V concentration ratio, and the range of the As-V interactions. Their results clearly show that increasing numbers of clustering over time have significant effects on dopant diffusivity and that the type and number of clusters strongly depends on the overall As and vacancy concentrations.

4.5 Vacancy-Mediated Arsenic Diffusion in Silicon

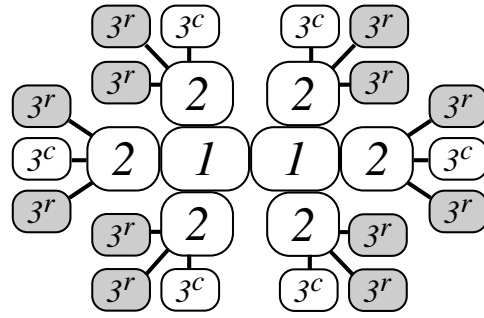
In order to study concentration and time dependent diffusivity, we have parameterized a KMC model for vacancy-mediated As diffusion in Si using *ab initio* energies available in the literature. Vacancy jump rates are based on calculations by Xie and Chen [45, 85] near As, and the isolated V migration energy, $E^{m,iso} = 0.28$ eV, comes from Centoni *et al.* [65]. We use those results, shown in Figure 4.3, for a V hopping within 3 nearest neighbor (3NN) positions of a single dopant. When there are multiple dopants within 3NN of a V, before or after it jumps, we calculate the formation energy in the vacancy's initial and final positions, E_{final}^f and E_{init}^f by summing the binding energy of the V with all i As within 3NN, $E^f = \sum_i E_i^b$. Then the migration

E^b (eV)	1NN	2NN	3^r NN	3^c NN	≥ 4 NN
	-1.21	-0.37	-0.25	-0.18	0.00
E^m (eV)	1NN	2NN	3^r NN	3^c NN	≥ 4 NN
	0.55	0.92	-	-	-
	0.08	-	0.35	0.46	-
	-	0.23	-	-	0.405
	-	0.27	-	-	0.37
	-	-	0.09	0.19	0.28

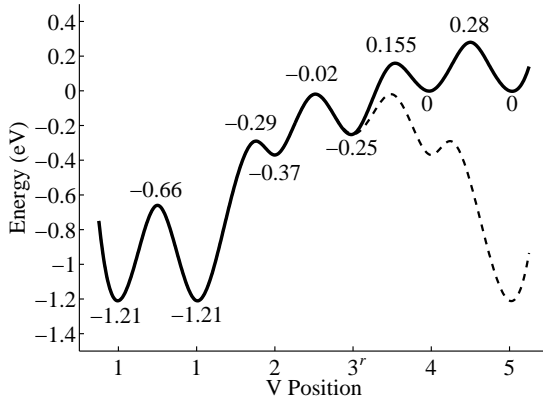
(a)



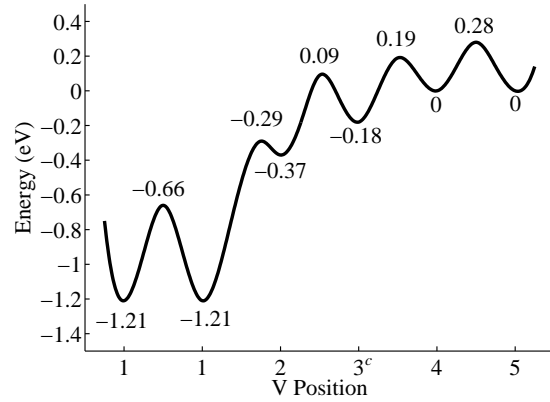
(b)



(c)



(d)



(e)

Figure 4.3: KMC parameterization of the KMC model for vacancy-mediated As diffusion in Si. The binding and migration energies (row→column), listed in (a) are from *ab initio* calculations [85, 45, 65]. The lattice is shown in (b); 3^r NN indicates a position 3NN around a hexagonal ring, and 3^c NN is a position 3NN along a $\{011\}$ chain. The state connectivity is shown in (c), with gray states being saddle states that lead a different V_1As_1 energy basin. The energy landscape is shown for a V moving from a 1NN position with As to the 3^r NN position in (d) (the dotted line indicates the path to a 1NN position by a different direction), and to the 3^c NN position in (e).

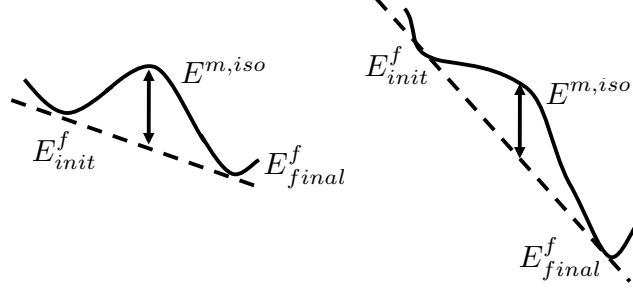


Figure 4.4: Saddle point energy calculation approximation.

energy is calculated from the difference, $\Delta E^f = E_{final}^f - E_{init}^f$, using the common [115, 116, 117] approximation

$$(4.29) \quad E^m = \max \left[\max [0, \Delta E^f], \frac{\Delta E^f}{2} + E^{m,iso} \right].$$

As shown in Figure 4.4, this approximation is the result of assuming the barrier can be calculated from a linear gradient between the states plus $E^{m,iso}$. If that sum is less than the higher energy state, then there is no barrier and the states are combined into a single state. Arsenic only moves by exchanging with V, and we assume a constant vacancy jump attempt frequency of $\nu = 10^{13} \text{ s}^{-1}$. The vacancy jump rates are calculated by

$$(4.30) \quad R = \nu \exp(-E^m/k_B T).$$

4.6 Trapping Energy Basins in Kinetic Monte Carlo Simulations

Preliminary KMC simulations using the model presented in Section 4.5 immediately demonstrate that standard KMC simulations are inadequate for modeling anything but the initial stages of diffusion. The simulations become very inefficient as soon as V-As clustering begins. To understand why, consider the diffusion of a V_1As_1 pair. Pair diffusion takes place by a mechanism, shown in Fig. 4.5, in which the V must traverse a hexagonal ring in order to approach the As from a

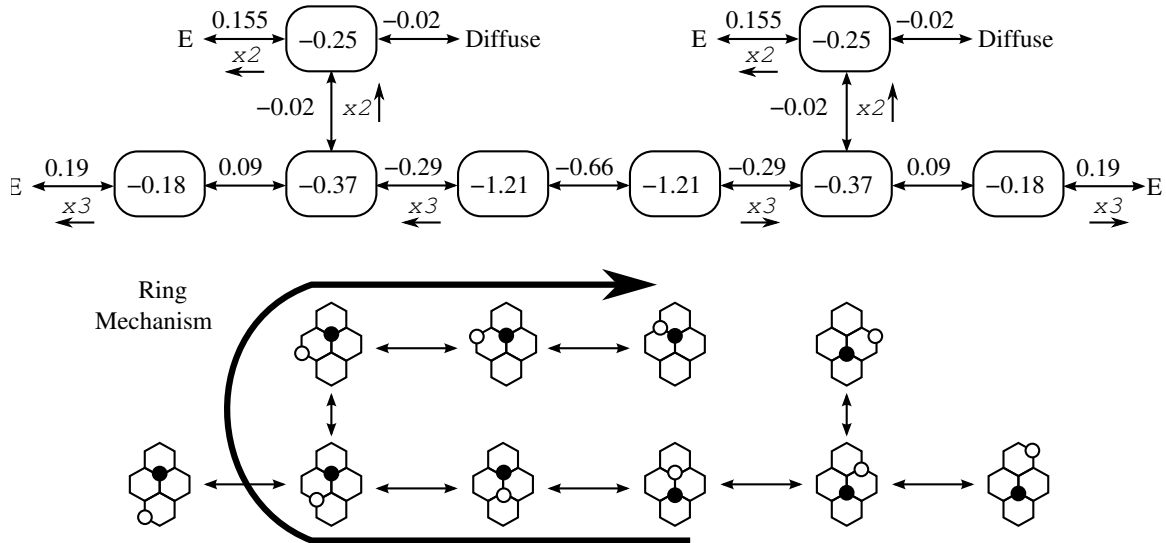


Figure 4.5: Diffusion of V_1As_1 clusters occurs by the ring mechanism. The state and saddle point energies are shown above, and schematics of the cluster configurations below. “E” indicates a pathway for the vacancy to escape the As. Small arrows with “ xn ” indicate that n symmetrically equivalent transitions are possible in the direction of the arrow.

different direction. The energy required to traverse the ring is much larger than the energy required for the V-As exchange, see also Figure 4.3(c). Using Equation 4.26, the transition probabilities have been calculated at 800 °C and listed in Table 4.1. The transition probability for V-As exchange, $P_{1 \rightarrow 1} = 0.948$, is much higher than the probability for a 1NN vacancy to traverse a ring in four direct jumps, $3P_{1 \rightarrow 2} \times 2P_{2 \rightarrow 3r} \times P_{3r \rightarrow 2} \times P_{2 \rightarrow 1} = 0.002$. This is not an exact calculation of the relative probabilities for exchange versus ring traversal; for that we would have to include all possible transition pathways. But, it does indicate that V-As exchange will occur many times before the V traverses the ring and a diffusive motion occurs. At lower temperatures, ring traversal would be even more unlikely compared to V-As exchange. We refer to any situation where the energy landscape, as in Figure 1.5, determines that many fast transitions over low energy barriers will occur before important slow transitions over high energy barriers a “trapping energy basin”.

The situation is much more complicated for As_2V , as shown in Figs. 4.6 and 4.7

P	1NN	2NN	3 ^r NN	3 ^c NN	≥ 4 NN
1NN	0.948×1	0.017×3	-	-	-
2NN	0.889×1	-	0.048×2	0.015×1	-
3 ^r NN	-	0.435×2	-	-	0.065×2
3 ^c NN	-	0.496×1	-	-	0.168×3

Table 4.1: Transition probabilities (row \rightarrow column) for a V near As at $T = 800$ °C. The \times indicate the number of different pathways available.

for two of the six lowest energy configurations. For V_1As_2 there are more possible diffusive pathways, but there are also more low energy configurations that act as trapping energy basins. For example, the lowest energy configuration, shown in Fig. 4.6, is when the vacancy is in between both As. There is also a configuration in which the As are on opposite sides of a hexagonal ring, with a vacancy in between, shown in Fig. 4.7. Relative to the minimum energy states, rotation around the ring requires at least 0.84 eV, while rearrangement to a different configuration off the ring requires at least 1.24 eV. This causes standard KMC simulations to spend many events simply rotating around the ring. Because the barriers inside the trapping energy basin are low relative to the escape barriers, the timesteps associated with the jumps are much smaller. This prevents the KMC simulation from reaching long times. Larger clusters have more and deeper trapping energy basins, creating worse efficiency problems. Therefore, to simulate As-V cluster diffusion in a reasonably efficient manner it is necessary to use an accelerated KMC method to escape the trapping basins.

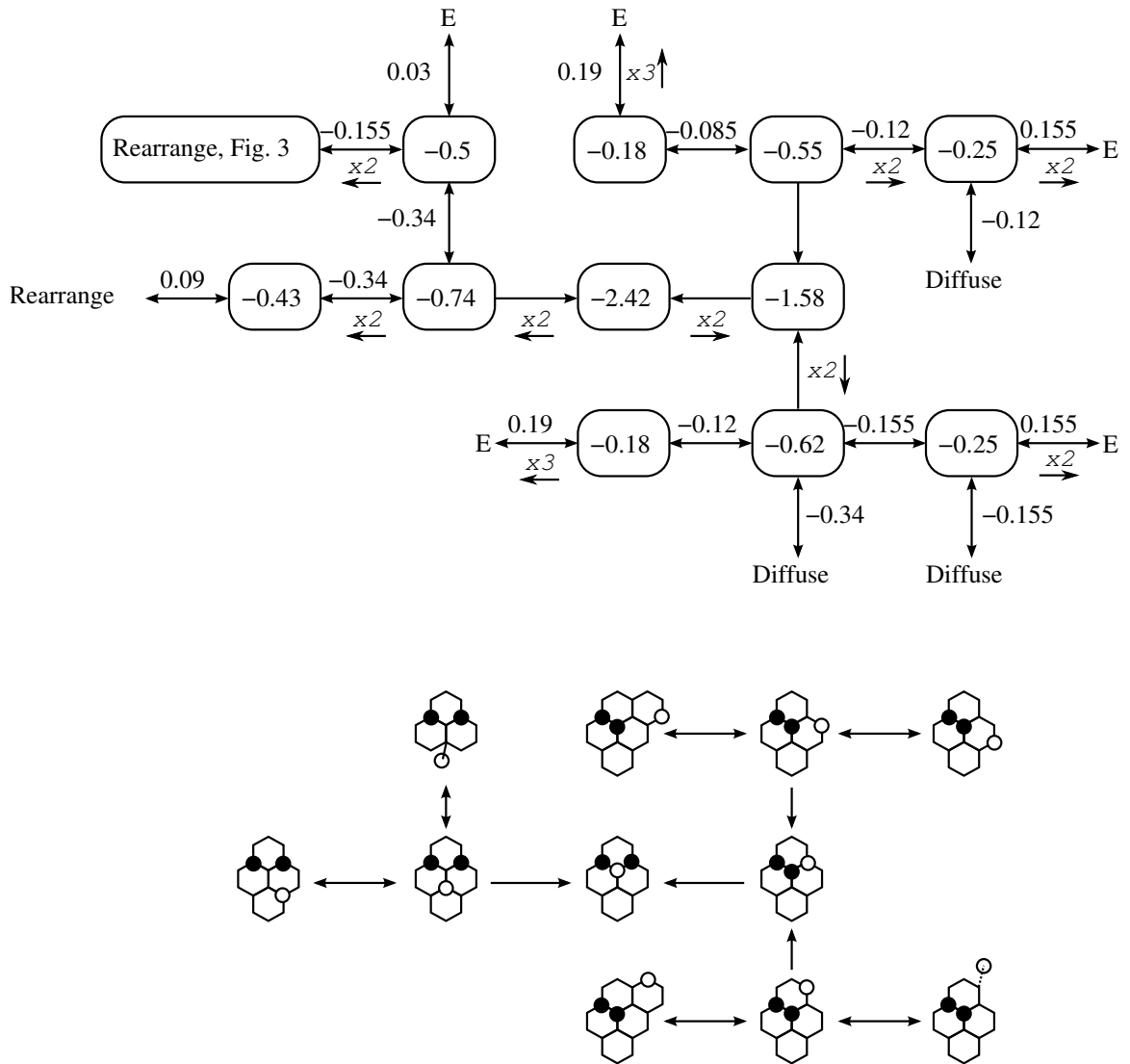


Figure 4.6: States near the minimum energy V_1As_2 configuration. The state and saddle point energies are shown above, and schematics of the cluster configurations below. “E” indicates a pathway for the vacancy to escape the As, “Rearrange” indicates a transition leading towards a different low energy V_1As_2 configuration. Transitions shown with single arrows indicate that one state is unstable towards the other. Small arrows with “ xn ” indicate that n symmetrically equivalent transitions are possible in the direction of the arrow.

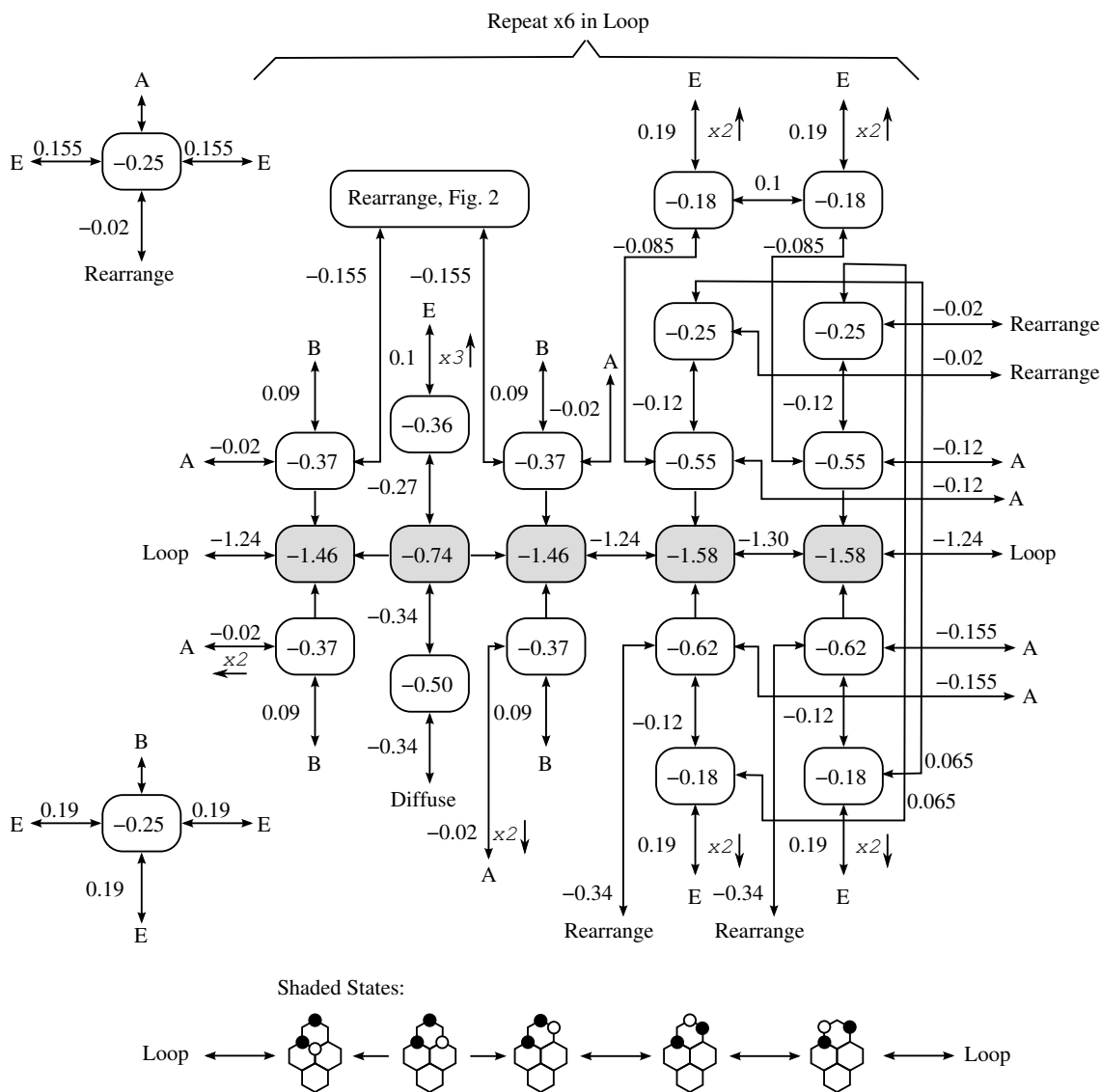


Figure 4.7: States involved in rotation of vacancy and As around a hexagonal ring in the V_1As_2 cluster. The state and saddle point energies are shown above, and schematics of the cluster configurations below. “E” indicates a pathway for the vacancy to escape the As, “Rearrange” indicates a transition leading towards a different low energy V_1As_2 configuration, “A” and “B” indicates states of the type shown to the left. Transitions shown with single arrows indicate that one state is unstable towards the other. Small arrows with “ xn ” indicate that n symmetrically equivalent transitions are possible in the direction of the arrow.

CHAPTER 5

Accelerated Kinetic Monte Carlo Theory

In Section 4.6, we observed that KMC simulations have difficulty reaching long times in systems where the energy landscape determines that many fast transitions between certain repeatedly visited states occur before important slower events. To treat this situation, several methods [118, 119, 120, 121, 122, 123, 124, 125, 126], ranging from exact to approximate, have been developed to accelerate the calculation of escapes from trapping energy basins. There are also some related methods useful for accelerating diffusion of rarely reacting species on a flat energy landscape [127, 128, 129], and others more generally used for transition path sampling [130, 131, 132]. The key idea behind the acceleration methods is that it is possible to fully solve the master equation over the limited number of repeatedly visited states using Markov chain analysis. By sampling the solution, a KMC event can be chosen which moves the system out of the repeatedly visited states in one jump. Acceleration is achieved if the calculation can be performed faster than explicitly moving between states until the system escapes the basin. Lacking from the literature has been a systematic way to identify, as the simulation runs and for any energy landscape, which states are in the energy basin and should be included in the calculation for efficient acceleration. Here we briefly review the acceleration methods, and introduce an algorithm that

uses the energy landscape to create Markov chains. In Chapter 6 we will detail the algorithms as applied to the model of vacancy-mediated arsenic diffusion in silicon presented in Section 4.5.

5.1 Absorbing Markov Chain Analysis

A Markov process is one in which previous events have no bearing on the probability of subsequent events. The current state of the system alone determines transition probabilities. The master equation as presented in Equation 4.24 describes a Markov process since the transition rates, $R_{i \rightarrow j}$, are constant. Atomic diffusion in crystalline materials is well modeled as a discrete state Markov process since the system tends to equilibrate within a state and lose memory of where it came from before transitioning to the next state. But even when some correlated jumps exist, the system can be modeled as a Markov process if the correlations can be properly accounted for in the transition rates. A Markov chain is the set of states that make up a Markov process.

The standard KMC method becomes trapped in energy basins because it is limited to choosing events that advance the system one state at a time. It then repeatedly visits states inside the basin many times before reaching the states outside the basin. It is much more efficient if the KMC events consist of moving the system from a state within the basin to an external state outside of the basin. To correctly reproduce the system dynamics, we need these accelerated events to properly account for the probability and time when the system *first* exits the basin. By solving the master equation for the states in the basin, with the condition that the system stops as soon as it reaches an external state, we can determine the exact probabilities and first passage time distributions. In the language of Markov processes, the repeatedly

visited internal states are referred to as “transient states” and the external states are called “absorbing states”; together they comprise an absorbing Markov chain.

5.1.1 First passage time analysis

We will refer to exact methods for solving the master equation of the absorbing Markov chain as first passage time analysis (FPTA). This was initially used with KMC by Novotny to study Ising model nucleation events at very low temperatures[118], and has since been described for dislocation kink nucleation [120], and with a limited number of states for vacancy diffusion in alloys [121].

The basic idea is that the system starts in the initial state, s^{init} , and then we calculate the probability as a function of time that it is in each of the transient or absorbing states. By sampling the probability distribution we can determine which absorbing state the system enters at what time and use that as the next KMC event. Unlike standard KMC with one state, here the accelerated exit time and event are correlated. This is clear when you consider that at the very shortest times it must be more likely to exit to an absorbing state from s^{init} than from other transient states, because of the time it takes to travel between transient states. Since FPTA is exact, for the sake of accuracy it does not matter which states are included in the calculation. In Section 5.2 we will discuss how it does matter for KMC algorithm efficiency.

Generally, the master equation can be written in terms of a state probability vector, \bar{P} , and transition rate matrix, \bar{M} , as

$$(5.1) \quad \frac{\partial \bar{P}}{\partial t} = -\bar{M}\bar{P},$$

$$(5.2) \quad M_{ij} = \begin{cases} -R_{j \rightarrow i} & \text{if } i \neq j \\ \sum_k R_{i \rightarrow k} & \text{if } i = j \end{cases} .$$

For each transient state i , $R_{i \rightarrow j}$ is determined from the model, for example Equation 4.30. For each absorbing state i , $R_{i \rightarrow j} = 0$. The solution of the system of differential equations is

$$(5.3) \quad \bar{P}(t) = e^{-\bar{M}t} \bar{P}(0).$$

Moler and Van Loan [133] have written about the advantages and disadvantages various ways to compute the exponential of a matrix, and it is the eigenvector approach that has been used previously [118, 120] for accelerating KMC. This approach gives

$$(5.4) \quad \bar{P}(t) = \bar{V} \bar{\Lambda} \bar{V}^{-1} \bar{P}(0).$$

where $\bar{\Lambda}_{ij} = e^{-\lambda_i t} \delta_{ij}$, λ_i being the eigenvalues of \bar{M} , and \bar{V} is the matrix whose columns are the eigenvectors of \bar{M} with the i th column relating to λ_i .

The system starts in the initial transient state with probability one, and eventually as $t \rightarrow \infty$ it will end up in an absorbing state. Therefore, the probability of being in an absorbing state, P^a , is a non-decreasing function with time and $P^a(t = 0) = 0$ and $P^a(t = \infty) = 1$. We can sample the exit time, t^{exit} , by choosing a random number, $rand \in (0, 1]$ and solving for $P^a(t^{exit}) = rand$. The transition rate matrix \bar{M} only needs to include the transient states because by summing over the i transient states we can calculate $P^a(t) = 1 - \sum_i P_i(t)$. In practice, we find it convenient to combine all the absorbing states into one for the purposes of FPTA and including it in \bar{M} so that $P^a(t)$ can be calculated from that one state.

After choosing the exit time, we must choose which transition the system took out of the transient states. The rate $R_{i \rightarrow j}$ at which the system exits from transient state i by transition j must be weighted by the probability that the system is in the transient state at the exit time. So we can calculate appropriately weighted rates

by $P_i(t^{exit})R_{i \rightarrow j}$. By setting $Q = \sum_{i,j} P_i(t^{exit})R_{i \rightarrow j}$, Equation 4.27 can be used to choose the next state. The chosen state and time t^{exit} comprise the next KMC event.

Figure 5.1 graphically demonstrates two cases of FPTA for a simple trapping basin. In the case on the left (Case L, $6k_B T$) the energy barriers between transient states are lower than the case on the right (Case R, $9k_B T$). The state and saddle point energies are shown on the top in units of $k_B T$. In the middle are the occupation probabilities for being in each state as a function of time, Equation 5.4. On the bottom are the probabilities of exiting to a particular absorbing state as a function of time, $P_j^{exit} = \sum_i P_i(t)R_{i \rightarrow j}/Q$. The exit time can be chosen by picking a random number, $rand \in (0, 1]$, and following it across the occupation probability plot until it intercepts the heavy line indicating the total probability of being in an absorbing state. The exit state can be chosen by picking a random number, $rand \in (0, 1]$, and using the lower graph, finding which state's area it falls in at the exit time. Arrows indicating this process are shown with Case R.

In both cases, the system is initially located in state 1 with probability 1 and ends up in the absorbing states with probability 1, as it must. In Case L, the probability of the system moving between all three transient states is greater. Because of this, there is a nearly equal probability of ending up in absorbing 4 or 10, while in Case R it is more likely to end up in 4 than 10. This also causes Case L to be more likely to escape to states 9 or 10 than it is in Case R. This occurs despite state 1 being symmetrically equivalent to state 3, because the system has to travel through states to get from one side of the basin to the other. This demonstrates that the probability of exiting to a particular state does depend on the initial state. Additionally, the exit probability plots clearly demonstrate how the exit state and time are correlated. At short times there is a significantly greater probability of exiting to states 4 or 5

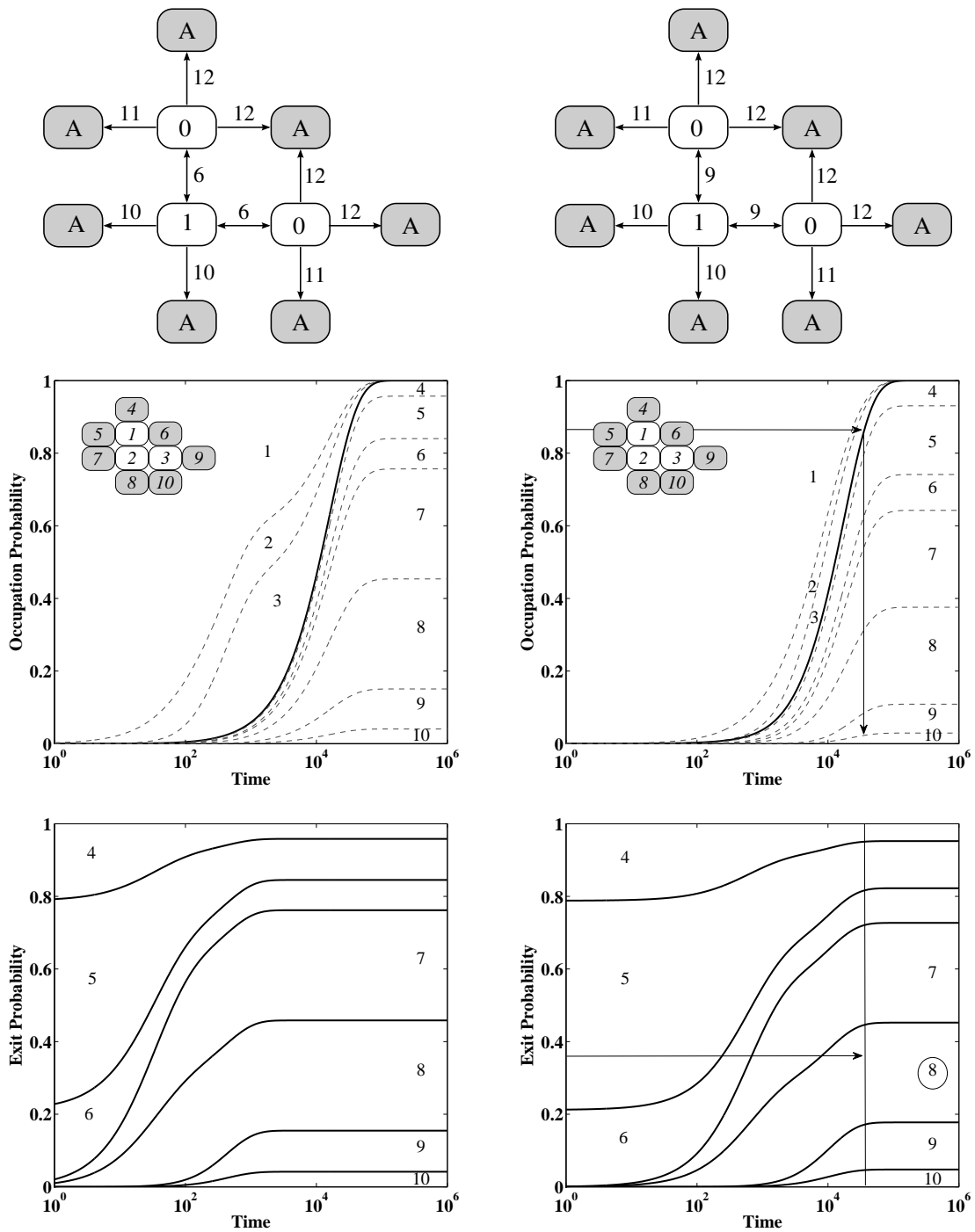


Figure 5.1: First passage time analysis for a basin with three states. The figures on top show the state connectivity and energies ($k_B T$). For the case on the right (Case R), the arrows indicate how the exit time is chosen from the occupation probability plot and the exit state is chosen from the exit probability plot.

than to states 9 or 10.

5.1.2 Mean rate method

For a given absorbing Markov chain, it is possible to calculate the mean rate of escape from the transient states to any absorbing state, and the exact probability of exiting to each absorbing state without also calculating the full time distribution. We will refer to any method that does this in full or in part as a mean rate method. Various descriptions of this have been given in specific [122, 121] cases and in general [119, 134], and the graph transformation method [123] also gives the same result. The mean rate method is faster than FPTA because it replaces matrix exponential calculation with one matrix inversion and simply uses the mean exit time instead of iteratively solving for t^{exit} . After many events either just using the mean time or always sampling from the time distribution, the resulting time will be the same because the differences are a random variable with mean zero. However, it remains unclear that using only mean rates instead of the full time distribution will never affect the simulation outcome. When measuring average quantities such as diffusivities after many events our experience is that this always gives the same answer as FPTA, but it is possible that under some situations it neglects important effects from time correlated events.

Generally, the system will jump back and forth through the transient states before escaping into an absorbing state. For a given absorbing Markov chain, the mean rate method calculates τ_i , the mean time per escape spent in each transient state i . To do this, we create the transition probability matrix, \bar{T} , with entries

$$(5.5) \quad T_{ji} = \frac{R_{i \rightarrow j}}{\sum_k R_{i \rightarrow k}} = \tau_i^{-1} R_{i \rightarrow j},$$

being the probabilities for exiting from transient state i to transient state j , where

k runs over all transient and absorbing states and $\tau_i^1 = (\sum_k R_{i \rightarrow k})^{-1}$ is the mean residence time in state i each time it is occupied. Then, the occupation probability vector, $\bar{\Theta}(m)$, which gives the probability of being in any transient state after jump m and before jump $m + 1$, is

$$(5.6) \quad \bar{\Theta}(m) = \bar{T}^m \bar{\Theta}(0),$$

$$(5.7) \quad \Theta_i(0) = \begin{cases} 1 & \text{if } i = s^{init} \\ 0 & \text{otherwise} \end{cases}.$$

Summing the occupation probabilities over all possible numbers of jumps gives

$$(5.8) \quad \bar{\Theta}^{Tot} = \sum_{m=0}^{\infty} \bar{T}^m \bar{\Theta}(0) = (I - \bar{T})^{-1} \bar{\Theta}(0).$$

With the summed occupation probabilities known, multiplying by the mean residence time gives the desired result,

$$(5.9) \quad \tau_i = \tau_i^1 \bar{\Theta}_i^{Tot}.$$

Once the τ_i are obtained, the mean rate of exiting to absorbing state j from transient state i is

$$(5.10) \quad \langle R_{i \rightarrow j} \rangle = \frac{\tau_i}{\sum_k \tau_k} R_{i \rightarrow j},$$

with k summing over all transient states. After summing over all possible exits, setting $Q = \sum_{i \rightarrow j} \langle R_{i \rightarrow j} \rangle$, the absorbing state occupied in the next KMC event is determined in a way analogous to Equation 4.27, using the accelerated rates $\langle R_{i \rightarrow j} \rangle$. Finally, the mean escape time, $\sum_i \tau_i$, is used for the KMC timestep.

5.1.3 Equilibrating basins

Unlike in FPTA or the mean rate method, the choice of transient states does matter for the accuracy of the equilibrating basin approach. It is an $O(n)$ calculation,

n being the number of transient states, so it is much faster than the other methods. This method [124, 125] assumes that it is so much more likely to transition between all of the transient states than it is to exit to any absorbing state that the system will equilibrate within the transient states before escaping. If this is the case, then the probability of being in a particular state is determined by the Boltzmann distribution and

$$(5.11) \quad P_i(t) = P_i = \frac{e^{-E_i/kT}}{\sum_j (e^{-E_j/k_B T})},$$

where E is the energy of transient state i , k_B is the Boltzmann constant, T is temperature, and the summation runs over all j transient states.

As with FPTA, the escape rate from transient state i into the absorbing state j is given by the product $P_i R_{i \rightarrow j}$. With $Q = \sum_{i,j} P_i R_{i \rightarrow j}$, the absorbing state occupied in the next KMC event is determined using Eq. 4.27 with the escape rates $P_i R_{i \rightarrow j}$. Because the system is assumed to equilibrate within the transient states, the exit time can be chosen using the standard approach of Eq. 4.28.

Because this method assumes the system equilibrates within the transient states it tends to give increased diffusivities in the case that the system does not fully equilibrate before exiting. The deeper the energy trap, the better the approximation. The occupation and exit probability plots in Figure 5.1 show that the equilibrating basin approximation would be a good approximation in Case L and a bad approximation in Case R. In Case L, the steady state in the exit probability plot is reached when the occupation probability plot indicates it is still about 90% likely for the system to be in the transient states. In Case R, the steady state exit probabilities are not reached until it is about 90% likely for the system to be in an absorbing state. We are not aware of work describing how to quickly and accurately determine on-the-fly when it is “safe” to use Equilibrating Basins. It would be useful to have a simple

test that uses the energy landscape, connectivity geometry, and error tolerance in order to decide when to use equilibrating basins and when to use FPTA or the mean rate method.

5.2 Energy Basin Algorithm

In order to make efficient use of the methods described in the previous section we need an algorithm to determine which states should be included in the absorbing Markov chain. If too few states are included, the system will not escape the trapping energy basin, because the most probable event will be to one inside the basin. If states outside the basin are included, the calculation is larger than it needs to be and, as shown in Section 6.3, the simulation efficiency decreases. We create absorbing Markov chains by systematically identifying trapping energy basins. A “basic energy basin” is defined as a group of states in which there is (i) a minimum energy saddle point, and (ii) in moving away from the minimum the saddle point energies do not decrease between successive moves. See Figure 5.2 for an example of energy basins. Additionally, the system could have a “basin of basins” out of which it needs to be accelerated. In that case we can merge two or more basins together to create a larger basin. In the merged basins there will not necessarily be one minimum saddle point moving away from which the saddle point energies are strictly non-decreasing. In what follows, we use the term “basin” simply to indicate any set of states which have been grouped together by one of the following algorithms. The states need not form a true “basic energy basin” if we use FPTA or the mean rate method since they calculate event probabilities and either the time distribution or mean timestep exactly, but for efficiency that is the goal. In the rest of this section FPTA is assumed, but the approach for the mean rate method is identical unless noted otherwise. We

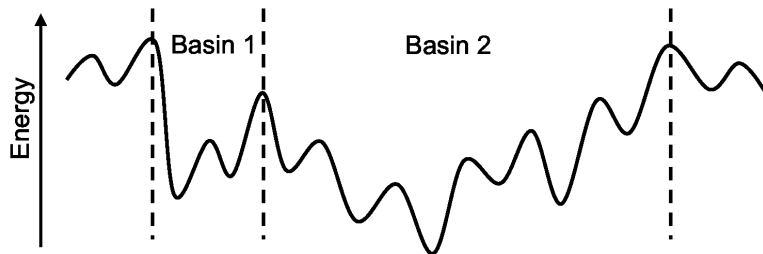


Figure 5.2: A “basic energy basin” is a collection of states in which there is a minimum energy saddle point and in moving from the minimum to the other states the saddle point energy does not decrease. Basic energy basins, like Basin 1 and 2, can be merged into combined basins, in which the saddle point energies will no longer be strictly non-decreasing moving away from the minimum.

have introduced equilibrating basins because they may be very useful in the right circumstances, but because of the unknown approximation error we only use them in one specific situation noted in Section 6.2.7.1 when the error can be demonstrated to be very small.

Once basins are identified, FPTA is used on the states in the basin. As with any KMC simulation, it is necessary to know all transition rates out of states the system enters. However, we assume that it takes non-negligible time for the KMC program to identify states and calculate transition rates as the computation proceeds. We consider states to be “known” (K) if their energy and exit rates have been calculated and saved in memory, all other states are considered “unknown” (UK). We assume that it is not known which states a transition leads to unless the transition has actually been made during the computation. Transitions that have not been taken are considered UK, and when the system takes a transition it becomes K. For FPTA on a basin, the states within a basin are the transient states, and transitions either to unknown states (whether they are actually to states in or out of the current basin) or to states in other basins are considered transitions to absorbing states. Within each basin the acceleration calculation is performed once, and if the system returns

to a previous basin we can reload the solution (eigenvalues and eigenvectors of the rate matrix) from memory. Then the next escape time and event may be chosen probabilistically without having to perform the full calculation again.

The three main ideas for basin identification are (1) that the saddle points within energy basins are non-decreasing moving away from the minimum, (2) that it is likely the system will transition to the bottom of a basin first, and (3) that the system is not likely to visit states in the basin above the minimum energy escape transition. To help identify the basin we set a variable, β , that gives the current direction of the system through the states of the basin, either down (\downarrow) or up (\uparrow). When the system first enters a new basin, β is usually set to \downarrow because we assume it will head down to the bottom. Each time we enter a new state we check if there are any UK transitions that are lower energy than the energy of the transition the system entered by. If there are, β remains \downarrow . If the minimum UK transition energy is greater than or equal to the incoming transition energy, then the system is beginning to head up and we set $\beta = \uparrow$. Once $\beta = \uparrow$, when the system enters a new state with an UK transition energy less than or equal to the transition energy, we have left the previous basin.

There are two cases when β is set to \uparrow to begin a basin. The first case is when there are more than one equally low minimum energy UK transitions from the first state. Those transitions could lead to different basins with the current state being a saddle state. By setting $\beta = \uparrow$ we begin a new basin if the system heads down into one of those basins. The second case is when the minimum UK transition from the first state is greater than the energy of the incoming transition. This usually occurs when all of the states in a basic energy basin were not included initially, either because a lower energy escape exists but the system has re-entered the basin,

or because the system did not reach the bottom before beginning to head up the basin.

We use two different methods to identify the basins. The first, which we call the “jump-first” method, advances the system one state at a time, at each step using FPTA, adding a single state to the current basin until we detect we have left it and entered a new basin. The other approach, which we call the “look-first” method, leaves the system in the initial state while it explores the surrounding states. At each step the look-first method follows the lowest energy UK transition out of all the states identified up to that point. Once it has identified the states in the basin, FPTA is used once to escape. The jump-first method has the advantage that it does not spend time in states that the system does not visit, but it performs more FPTA calculations. The look-first method minimizes the number of FPTA calculations, at the expense of possibly identifying more states than necessary and having a larger FPTA calculation. In the following sections we will demonstrate the jump-first and look-first methods on a simple system. In Sections 6.1.2 and 6.1.3 we will fully detail the algorithms.

5.2.1 Jump-first method

We demonstrate the jump-first method for the case of a V_1As_1 cluster at 800 °C in Figure 5.3. The system trajectory presented in Figure 5.3 was obtained directly from one instance of the jump-first method. At each step, Figure 5.3 shows the starting state of the system, β , the timestep until the next event, and the probability of exiting to different types of states at that exit time. The probability of exiting to different states of the same type is not equal and this is treated appropriately, but we do not resolve the differences in the figure due to lack of space. The type of state the






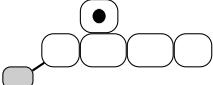
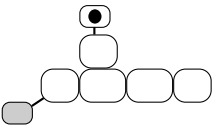
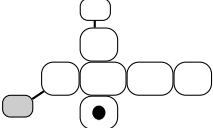
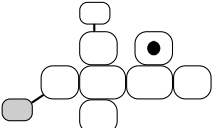
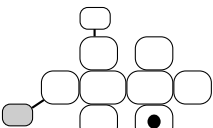
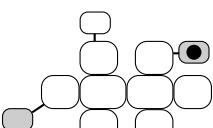
FPTA	1)		$3^r \uparrow$	$\Delta t = 1.18$	$P \rightarrow 2 = 8.69\text{E-}01$	$P \rightarrow 4 = 1.31\text{E-}01$	✓		
	2)		$2 \downarrow$	$\Delta t = 2.81$	$P \rightarrow 1 = 8.89\text{E-}01$	$P \rightarrow 3^r = 9.60\text{E-}02$	$P \rightarrow 3^c = 1.46\text{E-}02$	✓	
FPTA	3)		$1 \downarrow$	$\Delta t = 342.99$	$P \rightarrow 1 = 9.63\text{E-}01$	$P \rightarrow 2 = 3.52\text{E-}02$	$P \rightarrow 3^r = 1.70\text{E-}03$	$P \rightarrow 3^c = 2.59\text{E-}04$	✓
	4)		$1 \uparrow$	$\Delta t = 7532.68$	$P \rightarrow 2 = 9.78\text{E-}01$	$P \rightarrow 3^r = 1.89\text{E-}02$	$P \rightarrow 3^c = 2.87\text{E-}03$	✓	
FPTA	5)		$2 \uparrow$	$\Delta t = 406.61$	$P \rightarrow 2 = 9.48\text{E-}01$	$P \rightarrow 3^r = 4.55\text{E-}02$	$P \rightarrow 3^c = 6.92\text{E-}03$	✓	
	6)		$2 \uparrow$	$\Delta t = 0.17$	$P \rightarrow 2 = 7.01\text{E-}05$	$P \rightarrow 3^r = 8.68\text{E-}01$	$P \rightarrow 3^c = 1.32\text{E-}01$	✓	
FPTA	7)		$3^c \uparrow$	$\Delta t = 26525.38$	$P \rightarrow 2 = 9.02\text{E-}01$	$P \rightarrow 3^r = 8.70\text{E-}02$	$P \rightarrow 3^c = 8.79\text{E-}03$	$P \rightarrow 4 = 2.24\text{E-}03$	✓
	8)		$2 \uparrow$	$\Delta t = 18271.25$	$P \rightarrow 2 = 8.19\text{E-}01$	$P \rightarrow 3^r = 1.59\text{E-}01$	$P \rightarrow 3^c = 1.81\text{E-}02$	$P \rightarrow 4 = 3.09\text{E-}03$	✓
FPTA	9)		$2 \uparrow$	$\Delta t = 444.49$	$P \rightarrow 2 = 6.71\text{E-}01$	$P \rightarrow 3^r = 2.89\text{E-}01$	$P \rightarrow 3^c = 3.58\text{E-}02$	$P \rightarrow 4 = 4.08\text{E-}03$	✓
	10)		$2 \uparrow$	$\Delta t = 78478.23$	$P \rightarrow 3^r = 8.78\text{E-}01$	$P \rightarrow 3^c = 1.11\text{E-}01$	$P \rightarrow 4 = 1.13\text{E-}02$	✓	
FPTA	11)		$3^c \uparrow$	

Figure 5.3: The jump-first method for a V_1As_1 cluster at 800 °C. At each step, the current state of the system is indicated by a dot. The state connectivity graph is as in Figure 4.3. The gray states indicate different basins. The $\# \downarrow$ or $\# \uparrow$ indicate the current state and whether the system is heading down or up the basin. The probability for exiting to different types of states, at the chosen timestep ($\Delta t \nu^{-1}$ s), is listed on the right, and the chosen type of state indicated with a ✓.

system actually transitions to is indicated with a \checkmark . It is useful to refer to Figure 4.3 in order to view the energy landscape while following the steps in Figure 5.3. The state connectivity graphs in Figure 5.3 show the states that the system has visited and are saved in memory. They are part of the larger connectivity graph, shown in Figure 4.3(c). States are colored gray or white to distinguish separate basins.

The system begins in a 3^r NN state, with a vacancy 3NN around a hexagonal ring from an As, as shown in Figure 4.3(b). This is a saddle state between two different energy basins, so initially $\beta = \uparrow$. For the first event, the FPTA calculation is performed, but there is only one state so it is identical to standard KMC. The system quickly jumps to a 2NN state by a -0.02 eV transition. Since there is a lower (-0.29 eV) UK transition out of the 2NN state, the system is in a new basin and $\beta = \downarrow$. The FPTA is performed, but again there is only one state in the new basin, so it is identical to standard KMC. The system quickly jumps down to the 1NN state which is added to the basin. β remains \downarrow because the minimum UK transition is lower energy (-0.66 eV) than the incoming transition (-0.29 eV). Now there are two states in the FPTA calculation, and the system makes the transition to the other 1NN state by V-As exchange. At this point, step (4), the minimum UK transition (-0.29 eV) is higher energy than the incoming transition (-0.66 eV) so $\beta = \uparrow$. The next timestep obtained from FPTA, $\Delta t = 7532.68 \nu^{-1}$ s, is significantly greater. This event probably includes time spent by the system repeatedly visiting the 1NN states, though the actual path is not resolved by FPTA. The following event, (6 \rightarrow 7), is a relatively improbable jump to a 3^c NN state. Because the chosen timestep is improbably small, it makes it very likely that the system never left the 2NN state and becomes much more probable that the system exits to a 3NN state, as reflected in the exit probabilities. Next, the system returns to the basin and enters

another 2NN state, step (8). This state is lower energy than the previous 3^cNN state, but the algorithm is looking at the incoming transition pathway (-0.29 eV), not the last state energy, so β remains \uparrow and the algorithm keeps identifying the current basin. This continues for a few more steps until all of the 2NN states are included in the basin and the system enters a 3^rNN state in step (11). Here the minimum UK transition (-0.02 eV) is equal to the incoming transition, so the algorithm determines that the system has left the basin and a new basin is started with $\beta = \uparrow$.

The simulation proceeds in a similar manner. Whenever the system re-enters a basin that has already been saved, a basin occupation counter is incremented. If two neighboring basins are occupied more than some given number of times, the basins are combined so that the system can escape the combined basin. In this manner the system can accelerate out of complicated “basins of basins” automatically.

A few other observations can be made about the jump-first method in Figure 5.3. First, the system did reach the bottom of the basin before escaping. Second, the system did only reach one of the six higher energy 3^cNN states before escaping by a lower energy transition. While this is only one sample, it is representative of the quality of those assumptions and is supported by the exit probabilities. Third, by step (8) it is about 15% probable that the system escapes to a 3^rNN state before identifying the last two 2NN states. By step (9) it is about 30% probable that the system escapes without finding the last 2NN state. As noted previously, the FPTA calculation is done repeatedly for 1 to 9 states. If the 9 states were identified first and one calculation done, it might be more efficient. However, if the system escaped before entering all of those states, which we see is not too unlikely, some of the time spent identifying all 9 states would be wasteful and the calculation would be larger than necessary.

5.2.2 Look-first method

We demonstrate the look-first method for a V_1As_1 cluster at 800 °C in Figure 5.4. The system trajectory presented in Figure 5.4 was obtained directly from one instance of the look-first method. The information presented is the same as for the jump-first method, except while exploring the energy landscape inside a basin, where the UK transition energies are presented rather than exit probabilities.

In the look-first method we keep the system in the initial state and explore the energy landscape to identify the rest of the energy basin. At each step the algorithm follows the lowest energy UK transition out of all the states identified up to that point. Once it has identified the states in the basin, FPTA is used once to escape. For the first step there are two equally low minimum UK transitions. In this case, the algorithm does not try to guess which way the system would jump and considers the single state the complete basin. FPTA is performed, time is incremented, and the systems jumps into a 2NN state. This is a new basin and there is a lower energy UK transition, so $\beta = \downarrow$. Now the system explores the basin by following the lowest energy pathways without incrementing time. The original state is remembered and indicated in Figure 5.4 with a circle on the state connectivity graph. When the system reaches the second 1NN state in step (2c) there are no more lower UK transitions, so $\beta = \uparrow$ and the system begins exploring up the basin. It visits and saves all of the 2NN states until in step (2i) it jumps into a new 3^rNN state and has left the basin, indicated with a ■. It is returned to the original state, the state outside of the basin is forgotten and FPTA is performed. The timestep, $\Delta t = 2.69 \times 10^5 \nu^{-1}$ s, is large because it accounts for all the time the system spent in the basin. The exit state begins a new basin and the simulation continues in a similar manner. The procedure for counting the number of times a basin has been occupied and merging frequently

FPTA	1)		3^r	$\Delta t = 8.39$	$P \rightarrow 2 = 0.869$	\checkmark
					$P \rightarrow 4 = 0.131$	
min[UK]	2a)		$2 \downarrow$	$\Delta t = 0$	$E \rightarrow 1 = -0.29$	\checkmark
					$E \rightarrow 3^r = -0.02$	
min[UK]	2b)		$1 \downarrow$	$\Delta t = 0$	$E \rightarrow 1 = -0.66$	\checkmark
					$E \rightarrow 2 = -0.29$	
min[UK]	2c)		$1 \uparrow$	$\Delta t = 0$	$E \rightarrow 3^r = -0.02$	\checkmark
					$E \rightarrow 3^c = 0.09$	
min[UK]	2d)		$2 \uparrow$	$\Delta t = 0$	$E \rightarrow 2 = -0.29$	\checkmark
					$E \rightarrow 3^r = -0.02$	
min[UK]	2e - 2g)	...	$2 \uparrow$	$\Delta t = 0$	$E \rightarrow 2 = -0.29$	\checkmark
					$E \rightarrow 3^r = -0.02$	
min[UK]	2h)		$2 \uparrow$	$\Delta t = 0$	$E \rightarrow 3^r = -0.02$	\checkmark
					$E \rightarrow 3^c = 0.09$	
Load State 1	2i)		$3^r \blacksquare$	$\Delta t = 0$	$E \rightarrow 2 = -0.02$	\blacksquare
					$E \rightarrow 4 = 0.155$	
FPTA	2j)		2	$\Delta t = 268921.20$	$P \rightarrow 3^r = 0.868$	\checkmark
					$P \rightarrow 3^c = 0.132$	
FPTA	3a)		$3^r \uparrow$

Figure 5.4: The look-first method for a V_1As_1 cluster at 800 °C. The state connectivity graph is as in Figure 4.3. At each step, the current state of the system is indicated by a dot. While exploring a basin, the original state is indicated with a circle. The gray states indicate different basins. The $\# \downarrow$ or $\# \uparrow$ indicate the current state and whether the system is heading down or up the basin. A \blacksquare indicates the algorithm has determined that the system left the current basin. The transition energy, E , or probability, P , for exiting to different types of states, at the chosen timestep ($\Delta t \nu^{-1}$ s), is listed on the right, and the chosen type of state indicated with a \checkmark .

visited basins is identical for both jump-first and look-first methods.

It is clear that the look-first method has the advantage that it performs fewer FPTA calculations than the jump-first method. Another advantage, discussed in Section 6.2.7.1, is that there seem to be fewer very badly conditioned FPTA calculations when the look-first method is used. This may be a consequence of identifying all the states in the basin and therefore not leaving leaks to absorbing states that are orders of magnitude faster than other exits. However, the look-first method also visits an extra state each basin in order to determine when it has left a basin. This, along with larger than necessary FPTA calculations when the system would have left the basin before visiting all of the states, are disadvantages to the look-first method.

CHAPTER 6

Accelerated Kinetic Monte Carlo Algorithms

Implementation of an accelerated KMC method with the energy basin state saving algorithm is significantly more involved than implementation of the standard KMC method. Much of the complexity comes from bookkeeping involved with state and basin saving and tracking dopant and defect clustering. In this chapter we first give an overview of the algorithm for both the jump-first and look-first methods. Then we will detail the data structures and functions used in our practical implementation. Finally, the we validate the algorithm through simulations of V-As cluster dissolution.

6.1 Algorithm Overview

As with the standard KMC algorithm, the accelerated algorithm is a loop of determining all of the possible events, calculating their rates, and then choosing which event occurs at which time. The primary difference between the outline of accelerated and standard algorithms is that after an event occurs there is a state saving step in the accelerated algorithms. We found it more efficient to allow random walks on a flat energy landscape than to use FPTA, so whenever V are isolated we simply skip the state saving steps and perform standard KMC. Once a V interacts with other defects we begin using the accelerated KMC algorithms. Before continuing with the algorithms, it is necessary to be more specific with the definition of “states”

in KMC.

6.1.1 Defect grouping

Usually for KMC simulations of particle systems, the state of the system is the configuration of all particles. More generally, the states must include everything that would affect jump rates out of the state, and in order to use the acceleration methods presented in Chapter 5 the jump rates must not vary in time. For instance, this means that external stress or electric fields that affect jump rates must be constant over the time the system is in the transient states. If there are long range interactions between particles such that moving one particle affects the energetics of every other particle, then a state must be the configuration of all particles in the system. However, in many systems interactions are short range and it may be possible to create separate groups of interacting particles that are independent from every other group. This is the case in the V-As model when there is a low enough V concentration that we can identify groups of interacting V which are not affected by V outside of the group. Then, a state is the configuration of V in a group and the As with which they are interacting.

A KMC event is chosen for each group and the events are performed in order of occurrence. The system time is set to the time at which the latest event occurred, and each group time is set to when the last group event occurred. Until the group's event occurs, the group may transition between transient states and the accelerated KMC method does not resolve which state it is actually in. Therefore, when two groups approach each other it must be determined which transient state they are in. The lagging group is evolved through its transient states until the group time matches the leading group's time. Once they are synced, they are grouped so that

they will stay at the same time and can interact appropriately. As the total system evolves the groups must be continually updated. Despite the additional bookkeeping involved, it may be useful to break a system into smaller independent groups when possible because it reduces the number of states that must be saved and gives smaller, therefore faster, acceleration calculations.

Finally, it was noted in Section 4.5 that in cases when the migration energy approximation, Equation 4.29, finds no barrier between states, they are combined into a single state. In general, we refer to the stable state and all unstable states that fall down to the stable state as a “superstate”. The system always jumps out of a superstate in one event. It is possible for a state to be an unstable saddle state between multiple superstates. Whenever the system jumps to or falls into an unstable saddle state, we randomly choose which superstate it falls to. After every event in which the system jumps into a state with unstable transitions up or down, we must fully identify the superstate before continuing. In the following sections, we will generally just use “state” and save the use of “superstate” for situations in which having unstable states makes an important difference.

6.1.2 Jump-first method

A simplified outline of the jump-first method is presented in Algorithm 6.1. Specific details for each step will be given in Section 6.2. As discussed in Section 5.2 and demonstrated in Section 5.2.1, the jump-first method advances the system one state at a time. At each step it uses FPTA and adds a single state to the current basin until the algorithm detects the system has entered a new basin. Then it saves the current state in a new basin and continues. The algorithm assumes that the system will head down to the bottom of the basin and then will head up and out.

```

1 Calculate event rates;
2 Save first state, begin first basin, set  $\beta$  with Algorithm 6.3;
3 Determine first event;
4 repeat
5   Perform next event, advance time;
6   Update event rates;
7   if In a previously saved state then
8     Update state connectivity;
9     if In a different basin then
10      Update basin connectivity;
11      Check if basins should be merged;
12    end
13  else
14    Save new state;
15    Update state connectivity;
16    Decide if we are in the same basin, Algorithm 6.2;
17    if Still in current basin then
18      Add state to current basin;
19    else
20      Add state to new basin, set  $\beta$  with Algorithm 6.3;
21      Update basin connectivity;
22    end
23  end
24  Determine next event;
25 until Simulation finishes ;

```

Algorithm 6.1: A simplified outline of a KMC program using the jump-first method.

The current direction of the system through the states of the basin, either down (\downarrow) or up (\uparrow), is stored in the variable β . When the system enters a new state, β and the energies of transitions out of the state are used in Algorithm 6.2 to determine if the system is in the current basin or a new one. The criteria are based on the non-decreasing nature of saddle point energies moving away from the minimum of the basin, as shown in Figure 5.2. If the system is heading down, but the minimum UK transition energy out of the new state is greater than or equal to the K transition energy it entered by, then the system has reached the bottom of the basin and is now heading up. If the system is heading up, but the minimum UK transition out of the new state is less than or equal to the K transition it entered by, the system has escaped the basin and entered a new one. The conditional criteria are set to “ \leq ” and “ \geq ” so that basins are not saved on a flat energy landscape.

```

1 if  $\beta = \uparrow$  and  $\min[UK] \leq K$  then
2   | In a different basin;
3 else
4   | Still in current basin;
5   if  $\beta = \downarrow$  and  $\min[UK] \geq K$  then
6     |  $\beta = \uparrow$ ;
7   end
8 end

```

Algorithm 6.2: The criteria for determining if the system is still in the current basin. Known (K) and unknown (UK) transitions are from the current state only.

Upon creating a new basin, β is set using the criteria of Algorithm 6.3. Usually, $\beta = \downarrow$. There are two cases when β is set to \uparrow to begin a basin. The first case is when there are more than one equally low minimum energy UK transitions from the first state ($\#\min[UK] > 1$). Those transitions could lead to different basins with the current state being a saddle state. By setting $\beta = \uparrow$ we begin a new basin if the system heads down into one of those basins. The second case is when the minimum UK transition from the first state is greater than or equal to the energy of

the incoming transition. This usually occurs when all of the states in a basic energy basin were not included initially, either because a lower energy escape exists but the system has re-entered the basin, or because the system did not reach the bottom before beginning to head up the basin.

The jump-first algorithm includes checks after each KMC event to determine whether the system is in a new state or a previously saved state. We need to track the state connectivity in order to properly create the \bar{M} or \bar{T} matrices for the acceleration calculation. If the system enters a saved state in the current basin from a new transition pathway, then the state connectivity is updated for the next acceleration calculation. If we did not check if the current state was already saved, the acceleration calculation would still be correct, but we could not escape from “looped” basins, such as in Figure 4.7. If the system enters a saved state that is in another basin, the state connectivity is updated and the basin connectivity is updated. We need to track basin connectivity to allow neighboring basins to merge. If the system is trapped between basins and will repeatedly transition between them, they are merged into a single basin. This process of creating basins, and then when necessary merging them into larger basins allows the system to escape from any trap.

```

1 if #min[UK] > 1 or min[UK] ≥ K then
2 |   β = ↑;
3 else
4 |   β = ↓;
5 end

```

Algorithm 6.3: The conditional setting of β upon creating a new basin. Known (K) and unknown (UK) transitions are from the current state only.

```

1 Calculate event rates;
2 Save first basin, use Algorithm 6.5;
3 Determine first event;
4 repeat
5   | Perform next event, advance time;
6   | Update event rates;
7   | if In a previously saved state then
8     | | Update state connectivity;
9     | | if In a different basin then
10    | | | Update basin connectivity;
11    | | | Check if basins should be merged;
12    | | end
13    | else
14    | | Save new basin, use Algorithm 6.5;
15    | end
16    | Determine next event;
17 until Simulation finishes ;

```

Algorithm 6.4: A simplified outline of a KMC program using the look-first method.

6.1.3 Look-first method

A simplified outline of the look-first method is presented in Algorithm 6.4. The outline is identical to the jump-first method, except that where the jump-first method saves a single state, the look-first method saves an entire basin using Algorithm 6.5. As discussed in Section 5.2 and demonstrated in Section 5.2.2, the algorithm follows the lowest energy UK transition out of all the states identified up to that point. Once it has identified the states in the basin, FPTA is used once to escape.

Generally $\beta = \downarrow$ to start. However, if the first state has multiple equally low minimum energy UK transitions and they are less than the incoming K transition energy, we do not want to guess which one the system would take, so we stop. If they are greater than the incoming K transition energy, that means the basin was not fully identified previously, perhaps due to another escape path, but now the system has returned so it behooves us to finish going up the basin and β is set to \uparrow .

Like the jump-first method, the look-first method is an approximate method for identifying basic energy basins because it stops as soon as the first exit to a different

```

1 Save first state in new basin;
2 stop = false;
3 if #min[UK] > 1 and min[UK] ≤ K then
4 | stop = true;
5 else if min[UK] ≥ K then
6 | β = ↑;
7 else
8 | β = ↓;
9 end
10 while stop == false do
11 | Considering all states in current basin, take minimum energy UK transition;
12 | if In a previously saved state then
13 | | Update state connectivity;
14 | | if In a different basin then
15 | | | Update basin connectivity;
16 | | | stop = true;
17 | | end
18 | | else
19 | | | Update event rates;
20 | | | if β = ↑ and min[UK] ≤ K then
21 | | | | stop = true;
22 | | | else
23 | | | | if β = ↓ and min[UK] ≥ K then
24 | | | | | β = ↑;
25 | | | | end
26 | | | | Save state in current basin;
27 | | | | Update state connectivity;
28 | | | end
29 | | end
30 end
31 Place system in first state in current basin;

```

Algorithm 6.5: The algorithm for identifying basins in the look-first method.

basin is found. For a KMC simulation this is sensible because it is likely the system will choose that exit. Time spent identifying the higher energy states in a basic energy basin is often wasted. If it is desirable to include states above the minimum energy exit, Algorithm 6.5 could be modified to only stop when the entire basin is found. Or it could wait to stop when the minimum energy UK transition out of the basin is a certain magnitude above the known minimum energy basin exit.

6.2 Algorithm Details

Algorithms 6.1 and 6.4 give an important outline of the accelerated KMC algorithm, but there are many details lacking and complications left untreated. For instance, we noted that groups of isolated vacancies must be regrouped continually as the simulation runs. When regrouping is necessary, the groups must be synced in time, then their saved states are deleted, then the groups are combined, and finally the initial state (if jump-first) or basin (if look-first) is saved. Another example is collecting data as the simulations run. In order to study time and concentration dependent diffusion we desire data on the formation and diffusion of clusters of As and V. Simply looking at the defect and dopant positions at some sampled times is not sufficient because sampling errors occur when we are not in equilibrium and some clusters exist for much shorter times than others. It is not trivial to determine the exact amount of time spent in each type of cluster in an accelerated KMC simulation because FPTA does not resolve the amount of time spent in each transient state. If the clustering changes within an energy basin, the amount of time spent in different clusters can not be allocated appropriately. Also, determining the clusters at every step can be very time consuming and ruin simulation efficiency if not done intelligently. Finally, when there are many saved states, checking if the system is in

a previously saved state can be very time consuming if the search is done by simply checking all defect and dopant positions in all the saved states. We limit the maximum number of saved states and also perform some sorting to reduce the number of comparisons necessary. In this section, we will detail the data structures we used to save simulation data and show how these operations can be done efficiently. For some of these operations there may be more highly optimized computer science solutions available in the literature, but the solutions presented here were readily implemented and “good enough”.

The jump-first and look first method are detailed in algorithmic form in Appendices A and B. The algorithms contain references to where the steps are discussed in this chapter. The algorithmic details are somewhat tailored to the specific needs of our model for vacancy-mediated As diffusion in Si, but are readily generalized to most crystalline systems. These methods are especially suitable for systems with a low concentration of actively jumping defects but large variations in the energy landscape. Our V-As model does not include V-V interactions, but the KMC algorithms are designed to include those interactions once the parameterization is included.

6.2.1 Lattice Representation and Binning

The V-As diffusion model is limited to Si, As, and V all on diamond cubic lattice spots. All of our simulations have periodic boundary conditions, i.e. it is an endlessly repeating computational cell in which jumping off one side brings a V back on the other side. Only V move by themselves, and the energetics of their jumps are only affected by As within 3NN jumps. Therefore, we represent the system by keeping a list of all of the V and As positions and assuming Si atoms on every other lattice spot. In the following we will use “defect” to refer to a species which could be either

V or As, and if necessary specify that this refers to V or As only. Defect positions are specified by seven values: $(x_1, x_2, x_3, s, j_1, j_2, j_3)$, where x_1, x_2, x_3 are the x, y, z coordinates of the unit cell containing the defect, s is the index of the occupied lattice spot within the unit cell, and j_1, j_2, j_3 are the number of jumps the defect has made across the periodic boundaries. Jumps in the positive direction are a +1 and jumps in the negative direction are a -1. This allows for absolute displacements to be measured across the periodic boundaries. Another list contains the possible moves between lattice spots by specifying $(\Delta x, \Delta y, \Delta z, s)$, the change in unit cell positions and the new lattice spot. There are four possible neighbor jumps for any point on the diamond cubic lattice, and eight possible lattice spots, so there are thirty-two different moves specified. If off-lattice spots and jumps are desired, they can simply be added to the list of positions and allowable moves.

This representation scheme makes it necessary to search through the list of defects to find those that are close enough to interact and affect jump rates. To expedite the search we bin defects by region of space. This entails dividing the computational cell into regions of identical cubic size and creating a list containing all of the defects in each region. When the bin size is greater than the interaction range, we can find all possibly interacting defects by searching the $3 \times 3 \times 3$ bin region surrounding and including the defect under consideration. For our model with interactions ranging 3NN, bins can be as small as a single unit cell. Binning is done once initially before the main loop of the algorithm and updated as defects move.

6.2.2 Grouping

If there are multiple V, prior to starting the main loop of the KMC algorithm we find all the possibly mutually interacting V and place them in the same group.

Each group consists of a data structure specifying which V are in the group, the saved states belonging to the group, the basins belonging to the group, and some of the information about the clustering of defects in the saved states. We check for V - V interactions by using the bins. Each V has a $3 \times 3 \times 3$ bin region of influence. Any V with overlapping regions of influence are in the same group. A list of groups is maintained and groups can be identified by their index in the list, $g_{\#}$. In order to determine when two different groups are close enough that they might begin to interact and must be merged, we track which group is influencing each bin and keep a count of the number of influences for each bin. There is a $+1$ influence that comes from each V in the current state of the system, but we also need to include a $+1$ in the region of influence for each V in each saved transient state. Each time a state is saved or deleted or a V switches bins in the current state of the system, the bin influence count is modified. If a V begins to influence a bin already influenced by another group those two groups must be synced in time (see Section 6.2.8), all their saved states and basins deleted, and then merged. Provided that the bin length is greater than the defect interaction range, the $3 \times 3 \times 3$ bin region of influence provides a buffer so that two defects belonging to different groups will not interact before the groups are merged.

In the look-first method, two groups cannot be allowed to impinge upon each other during basin identification. If an attempted move would cause impingement, then the move is not made and basin identification is stopped. The system is replaced in the original state, FPTA done, and the chosen move made. At that point, group merging takes place as usual.

Groups are separated if the V belonging to the group move apart. For each V in the group, we set $grp = -1$. Then grp is set to $g_{\#}$ for the first V and it is checked

to see with which other V it shares an overlapping region of influence. Then grp is set to $g_{\#}$ for each of those, they are checked, and so on recursively. At the end, any V with $grp = -1$ is in a different group and must be regrouped. The ungrouping check does not need to be made every KMC event, that would just lower efficiency, so it is done either upon group merging or after every 10^3 events.

6.2.3 Jump rate parameterization

Jump rates are parameterized using the model described in Section 4.5. Rates can be determined efficiently by using a lookup table of migration and binding energies created before the main loop of the KMC algorithm. When the jumping rates for a specific V need to be determined, the $3 \times 3 \times 3$ bin region surrounding it are searched for other defects. The x,y,z displacement to the other defect i is used as indices into the lookup table and a set of energy parameters saved into a list. The energy parameters obtained from the lookup table include the appropriate E^m from Figure 4.3, the appropriate E^b from Figure 4.3 in the current state and after each possible V jump, and the number NN jumps from the V to defect i . If the defect is in a 1NN position we make a note of which defect and in which jump direction. If the defect is another V the move is not allowed, and if it is an As then that As will be swapped into the current V position if this jump is chosen.

After all the nearby defects have been found, the jump rates for the V can be calculated as follows. A loop is performed over each possible jump, j . If there is more than one defect within 3NN before or after jump j , then the formation energy in the current state and each possible final state are calculated and Equation 4.29 is used to calculate E^m . Otherwise, E^m is used directly from Figure 4.3. Using the value of E^m , R_j can be calculated from Equation 4.30. The saddle point energies are

the sum of the formation and migration energies. All of these quantities are saved to the current V.

The jump rate parameterization must be done for every V before the main loop of the KMC algorithm. During the simulation, the jump rates are only updated for the V that jumped during the last move and any interacting V within the jumping V's $3 \times 3 \times 3$ bin region.

6.2.4 States and superstates

6.2.4.1 State saving

States are saved within the group object as a separate object for each V in the group. The saved state contains the V's position, and the state connectivity: whether the possible V jumps lead to UK or K states, if K then to which state, and if any defects are in 1NN positions from the V. The saved state also includes the energy in the state and the energy of the saddle point for each transition. In order to account for the energy of all the vacancies in the group, the energy in the state is determined from the migration energies and the zero point reference is the energy of the first saved state.

The positions of As also need to be saved. To avoid saving the positions of every As, we only save those As that have exchanged with V from this group since the first state in the group was saved. When an As jumps for the first time we also save the original position before it jumped. When we load a state from memory, we load in the all of the saved As positions, and then place any As that jumped since the state was saved in their original positions. The saved and original As positions are stored in the group object. A limit is placed on the number of allowed jumped As, and when the limit is exceeded, all the states are cleared from memory. Currently, we

set the limit at $\max[20, 4N_V(g_{\#})]$, where $N_V(g_{\#})$ is the number of V in the group.

We also limit the total number of saved states, in part to avoid using too much memory, but mostly for efficiency when checking if the system is in a new state or a previously saved state. If saving a new state would exceed the limit, we first delete a basin as described in Section 6.2.5.2 and then save the state. Currently, we set the limit at 50,000 states, but in practice this limit is never reached because the limit on the number of basins is reached first.

6.2.4.2 State checking

When the system jumps into a new state by an UK transition, we need to check if it is really in a new state, or has entered a previously saved state. If we did not check if the current state was already saved, the acceleration calculation would still be correct, but the system could not escape from “looped” basins, such as in Figure 4.7. To help the efficiency of checking saved states, we bin them by V position. A good binning spreads the states evenly among the bins. We assume that the V position on the lattice will change the most often since it changes with every V jump, so we use that to bin. If there are 3 or more V in the group, the bin indices are $(s(V_1), s(V_2), s(V_3))$. There are 512 bins total. If there are fewer than 3 V in the group we use a combination of the V lattice spots and the V x_1 and x_2 positions modulo 8. After narrowing the search down to the bin corresponding to the current system, we simply compare all of the V and A positions in the saved state to the current state.

6.2.4.3 Superstates

After every event in which the system jumps into a state with unstable transitions up or down, we must fully identify the superstate before continuing. A list of superstates is kept in the group object. We identify the superstate using an algorithm similar to the “look-first” method: moving down to the bottom and then moving back up to the top. The difference is that for superstates we continue to the top of the superstate so that all states and all exits are included and don’t stop when we find the lowest exit. After all the states are identified we create a data structure listing the states and which one is the stable state. Additionally, we check each state to see if it is unstable to more than one superstate. If it is, we save to the state object the index of each superstate it falls to, the probability that it will fall to each superstate, and the probability it will fall to an UK exit. These probabilities are necessary for properly determining the rates between superstates in the acceleration calculations, as discussed in Section 6.2.7.

6.2.5 Basins

The basin data structure contains a list of the superstates in the basin, β , the basin occupation count, and the basin connectivity: which basins are connected and the saddle point energy of the connecting transition. It is possible for two basins to be connected at more than one transition, but we only store the lowest saddle point energy. The basin also saves FPTA or mean rate calculation results, as discussed in Section 6.2.7, and a list of clustering as discussed in Section 6.2.6.3.

6.2.5.1 Basin merging

Each time a basin is re-entered its occupation count is incremented. If the occupation count is greater than a certain number the occupation count is also checked for the neighboring basin connected by the lowest energy transition. If the neighboring basin's occupation count is also greater than the limiting value, then the basins are merged. Currently the merging limit is set at 100. When the basins are combined, the occupation count of the new basin is set at the sum of the original basins. This makes it quicker to fill in basins of basins, such as for the configuration in Figure 4.7.

6.2.5.2 Basin deleting

We limit the number of basins for similar reasons as the limiting the number of saved states: space in memory and state checking efficiency. Currently the maximum number of basins is set to 20. If saving a new basin would make the number of basins exceed the limit, then one of the current basins is deleted. The best basin to delete is the one that the system is least likely to revisit and does not break the overall basin connectivity if it is deleted. To find this basin, we traverse the basin connectivity graph beginning at the current basin and at each step following the lowest energy basin connection. The last basin reached is deleted.

6.2.6 Clustering

In order to understand the time dependence of vacancy-mediated As diffusion, we collected data on the clustering of V and As. A cluster is defined as all V and As that are within $n_c = 1$ NN jumps of each other, but the definition can be varied up to 4 NN jumps. The clustering is determined before the main loop of the KMC algorithm begins, and then updated as defects move. Efficient cluster updating is facilitated

by maintaining lists of cluster connections. Any two defects that are within n_c NN jumps of each other are said to have a cluster connection. Each defect has a list of all its cluster connections. Each cluster has a list of each defect in the cluster. The cluster can be considered a connected graph in which each defect in the cluster is a node and each cluster connection is an edge. At the moment the cluster is formed, the system time is saved to cluster data structure. When the cluster breaks up, the total time the cluster has been active is saved according to type, where the type is the number of V and As in the cluster.

It is important to note that clustering is independent of grouping. Two different groups may be in the same cluster if there is a background of As connecting isolated groups of V. Also, different clusters may be in the same group. All the vacancies in the same group need not be within n_c NN jumps of each other.

6.2.6.1 Finding and updating clusters

We found that for large, dense systems a recursive approach to identifying clusters would result in errors. Instead, a stack approach was implemented as follows. A loop iterates over every defect i . If defect i has not yet been added to a cluster, the cluster index $c_{\#}$ is incremented, and defect i is added to a stack of defects to search for cluster $c_{\#}$. While there are defects in the stack, a defect j is added to cluster $c_{\#}$ and searched for connections. The search is done over the $3 \times 3 \times 3$ bin region around the defect, and a cluster connection connection is added between defect j and any defect k found within n_c NN jumps if the connection does not already exist. If that defect k has yet not been added to a cluster, then it is also added to the stack. Once the search of defect j 's $3 \times 3 \times 3$ bin region is complete, the defect j is removed from the stack.

This process will find all of the clusters and create all the cluster connections, but it is too computationally intensive to be done every step. Instead, after a defect moves we check its cluster connections to see if any of them have broken and check the surrounding $3 \times 3 \times 3$ bin region to see if any new cluster connections should be formed. If cluster connections form, they may be within the same cluster or to a different cluster. If cluster connections break, the cluster may separate into different clusters or the overall cluster connectivity may remain intact. So if there are any cluster connectivity changes we need to check if the cluster has changed. This is easy to check for new connections, because each defect saves which cluster it is in. If two defects in different clusters become connected, then we merge those clusters. In practice we check for connections, then perform any mergers, then check for disconnections, then check for broken clusters, and finally separate any broken clusters.

It is not trivial to check if a cluster has broken. After disconnections, we check for cluster connectivity as follows. Call the moving defect d_1 . Then make a starting list that includes every defect, d_s , that d_1 disconnected from. If we can traverse the cluster connectivity graph from each d_s back to d_1 , the cluster is still connected. If we cannot traverse the graph back to d_1 , then in the process of trying we will identify all of the defects in a new cluster.

A rather efficient way to traverse the graph is to perform a directed depth-first search. At each step, the chosen connection is the one that maximizes the dot product $\hat{r}(d_c \rightarrow d_1) \cdot \hat{r}(d_c \rightarrow d_k)$, where d_c is the current defect, d_k is each of the possible connecting defects, and $\hat{r}(d_i \rightarrow d_j)$ is the unit vector connecting defect i and j . While traversing the graph we set $c_{\#}(d_c) = -(s + 1)$, where s is the index of d_s in the starting list, save d_c in a list of traversed nodes called *visit_list*, and remember

the previous node d_{prev} . Indices all start with 0, so using $-(s + 1)$ assures that while checking for connectivity we are using a negative number distinguishable from the actual cluster numbers. If at any point we come to a node with a negative cluster number that is not the current $-(s + 1)$, we know that we can stop looking because some other d_s has found d_1 . If a connection is to a defect with $c_{\#} = -(s + 1)$, the search has already been there and the move is not allowed. If no connections from a node are allowed we return to $d_{prev}(d_c)$ and continue the search.

At the end, if each d_s found d_1 , the cluster is still connected and we reset $c_{\#}$ for every defect in *visit_list*. Otherwise, the cluster has broken up. Then we can run through the list of defects in the original cluster, and the defect's current $c_{\#}$ can be used to create the new clusters.

6.2.6.2 Cluster connection saving

The different states within a superstate might have different clustering when we use the n_c NN jumps criteria. Therefore, we define that for a superstate there is a cluster connection between two defects if they are within n_c NN jumps of each other in *any* state in the superstate. The clustering must be checked for changes after the superstate is fully identified. If defects disconnect, we need to check if the cluster broke by traversing the cluster graph from each d_s to *every* defect that moved, d_{move} . So in the algorithm above, the goal of finding d_1 is replaced with a goal of finding every d_{move} .

Whenever a superstate is completed and saved, the cluster connections are saved to the superstate object. The number of cluster connections and which defect the connection leads to is saved for each V in the group, and each As that has jumped since the first state was saved. Then whenever the superstate is loaded, all the saved

connections are added if they do not currently exist, and any current connections that do not exist in the saved state are deleted.

6.2.6.3 Basin cluster saving

As noted at the beginning of Section 6.2, the acceleration calculations do not resolve the amount of time spent in each transient state. If the clustering changes within an energy basin, the amount of time spent in different clusters can not be allocated appropriately. Therefore, every superstate must be checked to see if the clustering changes. If it does, the basin is stopped prematurely and it can not be merged with basins that have different clustering. This is an additional criteria added to Line 1 or Algorithm 6.2 in the jump-first method, and Line 20 in Algorithm 6.5.

Prematurely stopping the basin identification and preventing basin merging can cripple the efficiency of the accelerated algorithm. Therefore, we continue to prevent basins from forming with different clustering initially, but save the clustering in the basin so that it can be quickly loaded upon re-entering a basin. Additionally, we allow basin merging as usual even if the clustering is different. In this case, the clustering within the basin is defined similarly as within the superstates: if a cluster connection exists between defects in any superstate in the basin, then the defects are considered connected in the basin. The updated cluster connections are still saved within each superstate object.

The method for saving and loading the clustering is fairly straightforward, and similar to the method for saving As positions discussed in Section 6.2.4.1. When a cluster breaks up or combines to form a new cluster, instead of deleting the old data structure we simply make it inactive, store the time it was active for, and add it to a list of changed clusters. We also record whether the cluster was active or inactive

when the list started in the list *active_init*. Then when a basin is saved, we also save to the basin object the number of changed clusters, N_{cc} and the current state (active/inactive) of the clusters in the changed clusters list. For each active cluster, a counter in the cluster object, *num_saved*, is also incremented +1. When the system re-enters a basin, along with loading the cluster connections for the current superstate, we set the state of each changed cluster. For each cluster i in the global changed cluster list, if $i < N_{cc}$ the state of cluster i is set by the basin's saved cluster activity list. For each cluster $i \geq N_{cc}$ the state of the cluster is set by *active_init*. The start time is set for each of the newly active clusters and for each defect in the cluster we set $c_{\#}$. The total active time for each of the deactivated clusters is stored. Whenever a basin is deleted *num_saved* is decremented for each of the clusters that were active when the basin was saved. If *num_saved* reaches zero and the cluster is not currently active it can be deleted. The size of the list of changed clusters is limited and when that limit is reached, all the saved states, basins, and changed clusters are cleared from memory. Currently the limit is set at 100. The changed clusters list is also cleared any other time all the saved states and basins are cleared.

6.2.7 Accelerated event determination

When there are several independent groups of V a KMC event is chosen for each group. Then in every loop of the KMC algorithm the next occurring event is performed. After the event is performed that group's next event is chosen. If all the V in the group are isolated from other defects, then the standard KMC method in Section 4.3 is used to determine the event. When the V interact with A s we begin saving states and using the accelerated methods.

6.2.7.1 First passage time analysis

The FPTA method is used as introduced in Section 5.1.1. Each time FPTA is calculated for a basin we create the \bar{M} matrix from the basin's list of superstates and the superstates' list of states and state connectivity. All UK transitions from superstates in the basin lead to a single absorbing state. All of the connections between superstates are determined by checking the state connectivity to determine which states they are connected to, and from that which superstates. Because a state may be unstable to multiple superstates it is important to know which superstates it is unstable to and with what probability it will fall to each superstate. These are determined whenever the superstates are created. For example, assume state 1 in superstate A is connected to state 2 which unstably falls 50% to K superstate B, 25% to K superstate C, and 25% to an UK exit. Then in the \bar{M} matrix, the rate $R_{A \rightarrow B} = 0.5R_{A \rightarrow 2}$, $R_{A \rightarrow C} = 0.25R_{A \rightarrow 2}$, and $R_{A \rightarrow \alpha} = 0.25R_{A \rightarrow 2}$, where α is the absorbing state. In this example, state 2 must not be in superstate A.

We use the LAPACK library [135] in order to calculate eigenvalues, eigenvectors, and the inverse eigenvector matrix. When all of the saddle point energies are the same for transitions in either direction detailed balance should hold within the transient states, and we expect that all the eigenvalues should be real and positive. In practice, precision errors can result in eigenvalues with very small imaginary components, usually less than 10^{-14} . Currently we carry the imaginary component through the calculation as long as possible until a real valued time or probability is used, but we have not observed a difference without it. After the eigenvalues, eigenvectors, and inverse eigenvector matrix are calculated they are saved to the basin. Any time the system returns to the basin, the FPTA solution can be used to choose a new KMC event. However, if the system jumps out of the basin by an UK transition that leads

back to the same basin, then the state connectivity must be updated and the FPTA solution recalculated so that this jump does not occur again.

Sometimes, the differences between rates connecting different states range over many orders of magnitude and create conditioning problems that can affect the state probabilities. We check for these problems in a couple of different ways. At time zero the system is in the initial state, *init*, with probability 1 and every other state with probability 0, so we can write Equation 5.4 as

$$(6.1) \quad P_\alpha(t) = \sum_i a_i e^{-\lambda_i t},$$

$$a_i = \sum_j V_{\alpha,j} V_{j,init}^{-1}.$$

From this, and $\lambda_\alpha = 0$, we can calculate the mean exit time,

$$(6.2) \quad \langle t^{exit} \rangle = \frac{\sum_{i \neq \alpha} a_i / \lambda_i}{\sum_{i \neq \alpha} a_i}.$$

We know that $P_\alpha(\infty) = a_\alpha = 1$, $P_\alpha(0) = \sum_{i \neq \alpha} a_i = 0$, and $\langle t^{exit} \rangle > 0$, which gives us three different requirements to check. These requirements are often not all met when the problem is ill-conditioned.

In this situation there are two options. The first option is to rid the basin of the fastest transitions by finding the states that are creating the problem and combining them into a single state using the equilibrating basin approximation. The equilibrating basin approximation is only used (1) after there is an ill-conditioning error, and (2) when we find a set of superstates between which transition rates are 10^7 times faster than any transition rate to a superstate outside of the set. If the second condition holds, the equilibrating basin approximation ought to be very good. Then the set of superstates is replaced in the \bar{M} matrix with a single state, and the transition rates to other states are adjusted by the equilibrium probability of Equation 5.11.

When the equilibrating basin approximation can not be used, the second option is to use the memory only method described in Section 6.2.7.3 to explicitly jump through the basin until the system escapes. Although these problems happen infrequently in comparison to the total number of events, being forced to use memory only in the worst trapping basins can significantly decrease simulation efficiency.

Usually the exit time is determined by choosing a random number and solving the equation

$$(6.3) \quad rand \in (0, 1] = \sum_i a_i e^{-\lambda_i t^{exit}}$$

for t^{exit} . This is solved numerically by making a starting approximation $t^{exit} = \langle t^{exit} \rangle$ and iteratively improving the approximation until the bounds $t^{min} < t^{exit} < t^{max}$ satisfy

$$(6.4) \quad 10^{-5} > \frac{t^{max} - t^{min}}{(t^{max} + t^{min})/2}.$$

Convergence is typically attained in 20-30 steps using a bracketing and bisection method [136]. We choose the exit event based on the exit time as described in Section 5.1.1, but in two steps. The first step chooses which superstate the system exits from. The second step chooses which state within the superstate the system exits by, and which V jumps in what direction.

6.2.7.2 Mean rate method

The mean rate method is calculated for a basin as introduced in Section 5.1.2. As in FPTA, we create the \bar{T} matrix from the basin's list of superstates and superstates' list of states and state connectivity. There are two options for determining the solution. We can either (1) solve for only the solution given the particular initial state the system is in, or (2) solve for the system starting in any state. Using

Equation 5.8, it is quicker to just get the solution for our particular initial state by solving for $\bar{\Theta}^{Tot}$ from

$$(6.5) \quad (I - \bar{T})\bar{\Theta}^{Tot} = \bar{\Theta}(0).$$

But if we fully solve for $(I - \bar{T})^{-1}$ we get the solution for the system starting in any state. We can use that solution any subsequent time the system re-enters the basin. As in FPTA, if the system jumps out of the basin by an UK transition that leads back to the same basin, then the state connectivity must be updated and the mean rate solution recalculated so that this jump does not occur again. The optimal method, (1) or (2), depends on how often the system re-enters basins and how often the solution must be recalculated due to connectivity changes. Currently we use option (1). We solve for $\bar{\Theta}^{Tot}$ using Gauss-Jordan elimination and conveniently $(I - \bar{T})$ never has zeroes on the diagonal.

6.2.7.3 Memory only

In order to compare the efficiency gained from acceleration calculations to the efficiency gained from reducing the number of jump rate calculations, we have introduced the memory only method. This method explicitly evolves the system through the states of the basin using the standard KMC method for choosing timesteps and events. At each step we sum the timesteps in order to calculate the final escape time. Once the system leaves the basin, the escape time and transition are chosen for the next KMC event.

6.2.8 Updating group time

Occasionally, we have already chosen a group's KMC event, but before the event occurs something happens and we need to determine in which transient state the

system is located at a given time. In our simulations this occurs when two groups impinge upon each other or when we decide to output simulation data at a given time.

6.2.8.1 Memory only

In the memory only method, the system is updated by evolving it through the transient states as normal, but transitions into an absorbing state are not allowed. If a timestep for a jump puts the group time past the desired stop time, then that jump is not allowed and the system stays in the current transient state. Afterwards, it is no longer valid to use the original KMC event because that was chosen without taking into account where the system was located at the update time. The additional constraint that the system was located in a given transient state at a certain time changes the exit time and event probabilities. However, because the events are Poisson processes we can calculate the next KMC event exactly as we normally would using the updated state and time as the starting positions.

6.2.8.2 First passage time analysis

The updating procedure for FPTA is analogous to the updating procedure for memory only. We calculate the state occupation probabilities as a function of time, but do not allow any transitions to the absorbing state. Then we can use the state occupation probabilities at the desired update time to randomly choose which transient state the system is in. Afterwards, a new KMC event is chosen by the normal method using the updated state and time as the starting positions.

6.2.8.3 Mean rate method

The mean rate method can not be used to determine updated system positions. The mean occupation times it calculates for each transient state are not up to a given time, but until the system escapes the basin. Consider the case in which data is output at regular time intervals less than the mean escape time from the current basin. Even if memory only or FPTA was used to update the group to the output time, if the next event was determined using the mean rate method it would never escape the basin. One solution is to use FPTA or memory only to update the group time, allowing it to escape the current basin, and then use FPTA or memory only to also calculate the next event after the output time. Instead, we make the approximation that each group is still in the current basin's initial state and do not recalculate the group's next KMC event. This should not affect average quantities like diffusivities, because the affect of shortening one data output time interval just increases the next and will average out over many samples and data points.

6.3 Results and Validation: Arsenic-Vacancy Cluster Dissolution

We have performed a series of accelerated KMC simulations of As-V cluster dissolution in Si to compare the results of the several acceleration and state-saving methods. We performed simulations without acceleration and with FPTA, the mean rate method, or the memory only method. For state-saving we used both the jump-first and look-first methods as shown in Appendices A and B and described in this chapter. We also used “ n -jump-first” and “ n -look-first” methods in which, rather than stopping state-saving based on the energy landscape we always save n states, with n varying from 2 to 50. Practically, in the unaccelerated simulations we still need to save states to create superstates, so we use the memory only method with

$n = 1$ -jump-first state saving. For the n -jump-first method, once the state limit is reached we delete a state using the basin deleting algorithm in Section 6.2.5.2 before saving a new state. For the n -look-first method, we save a new basin with n states each step.

Cubic systems 8 unit cells on a side with periodic boundary conditions were initialized with 1 V surrounded by 1 to 3 As dopants in 1NN positions. The simulations were run at 800 °C until the cluster completely dissolved as defined by every V or As being more than 4NN jumps from every other. Once the cluster dissolved, the final displacements, simulated KMC time (t^{KMC}), elapsed CPU time (t^{CPU}), and number of KMC events were recorded. Simulation efficiency is defined by the ratio $t^{KMC}:t^{CPU}$. All simulations of a given type were performed on the same computer platform. Acceleration is calculated from the ratio of simulation efficiency obtained using acceleration methods to the simulation efficiency obtained using KMC simulations with no acceleration or state-saving. The absolute acceleration number depends very significantly on the simulation temperature; the lower the temperature the greater the observed acceleration.

Acceleration results for V_1As_1 cluster dissolution are given in Figure 6.1. The mean rate method is the fastest, followed by memory only and then FPTA. The fact that memory only provides nearly as much acceleration as the mean rate method indicates that efficiency loss in standard KMC simulations of this cluster is mostly due to the repeated calculations of jumping rates in repeatedly visited states. The automatic basin finding algorithms are generally faster than n -methods for any n . The exception is memory only with jump-first state-saving, in which the basin finding and n -methods are nearly equal. The n -methods generally peak at $n = 8 - 10$ states. This corresponds to the 8 states in the energy basin of clustered V_1As_1 : two 1NN

states, and six 2NN states. The exception is jump-first FPTA, which peaks at $n = 3$. Because FPTA is a more involved calculation and the V_1As_1 basin is not too deep at 800 °C, it is more disadvantageous to redo the FPTA each time a state is added than it is to not include every state in the basin. In every case, the total acceleration falls off as n passes the 8-10 states needed to fill the first basin. The look-first method tends to fall off more steeply than the jump-first method because at large n it is more likely to include many states that the system never visits. The memory only jump-first method falls off very slowly at large n since the memory only calculation does not increase with n if there are unvisited states.

Further analysis of the V_1As_1 dissolution results in Figure 6.2 confirms that the acceleration methods are working properly. The plot of elapsed CPU time, t^{CPU} , confirms the acceleration results of Figure 6.1. The plot of the number of KMC events shows that using either the jump-first or look-first method with any acceleration method produces the same distribution of number of KMC events. Additionally, by counting the number of explicit jumps made through the transient states in the memory only simulations, we see that it exactly reproduces the distribution of number of KMC events in unaccelerated simulations. Finally, the plot of simulated KMC time, t^{KMC} , shows that the distribution of cluster dissolution times is exactly reproduced by the FPTA and memory only simulations. The distribution is not reproduced by the mean rate method simulations, which are clustered near the mean. However, the mean dissolution time is the same for all methods, indicated by the vertical lines that all fall on top of each other. The accuracy of the acceleration methods in determining the vacancy and As displacements upon dissolution is confirmed in Figure 6.3. The x -displacement upon dissolution is identical for all simulation methods, and by crystal symmetry is the same for all three directions. The high probability for no As

displacement shows that at 800 °C the V_1As_1 cluster dissolves more readily than it diffuses as a complex.

The power of the KMC acceleration methods and the automatic basin finding algorithms is more evident in the simulation of deeper trapping basins. Results for the dissolution of V_1As_2 and V_1As_3 clusters are shown in Figures 6.4 and 6.5. These results show that the accelerated KMC methods are orders of magnitude faster than unaccelerated KMC, the value of the acceleration depending on the temperature. In the deeper traps of V_1As_2 and V_1As_3 , the advantage of FPTA and the mean rate method over memory only becomes very significant, as most of the KMC inefficiency comes from bouncing between trapping states rather than repeating rate calculations. These results also show the significant advantage of the automatic basin finding state-saving methods over the n -methods with fixed numbers of saved states. Particularly, note the increase in efficiency that comes in the V_1As_2 dissolution simulations when $n = 24$. This is due to the configuration of two As opposite each other around a hexagonal ring, with a vacancy in between, as shown in Figure 4.7. As the vacancy causes the As to rotate around the ring, there are 24 states in an energy basin comprising the entire rotational motion. Including all of these states in an acceleration calculation provides a significant advantage. However, most other basins are smaller so using 24 states all the time is undesirable. As the number of defects and dopants in a cluster increases, the number of unique basins increases and the flexibility to use just the right number of saved states becomes increasingly advantageous. For completeness, figures of timing and displacement for V_1As_2 and V_1As_3 dissolution analogous to Figures 6.2 and 6.3 have been included in Appendix C.

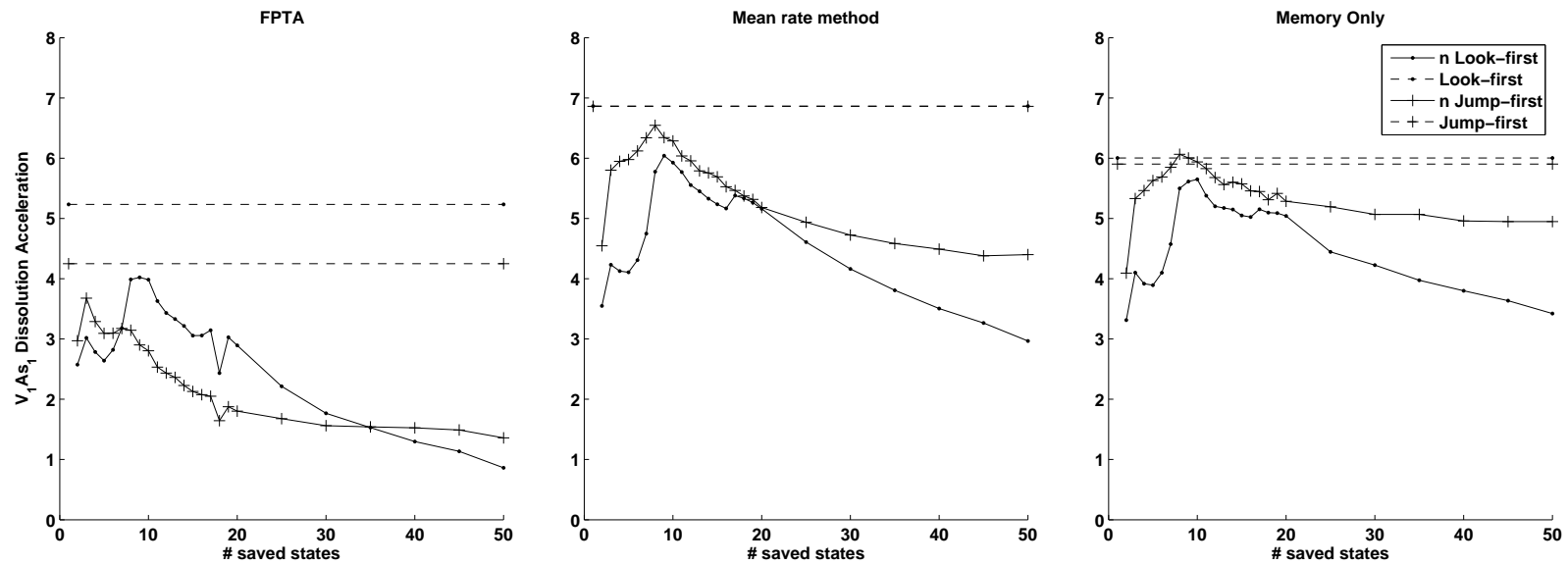


Figure 6.1: Acceleration achieved during V_1As_1 dissolution. The basin identifying state-saving methods, look-first and jump-first, automatically determine the number of saved states, so they are simply presented as straight lines for comparison with the n -methods. The look-first and jump-first lines are the mean of 10^5 samples. Each data point for the n -methods is the mean of 10^4 samples. Acceleration is relative to the mean of 10^5 unaccelerated samples.

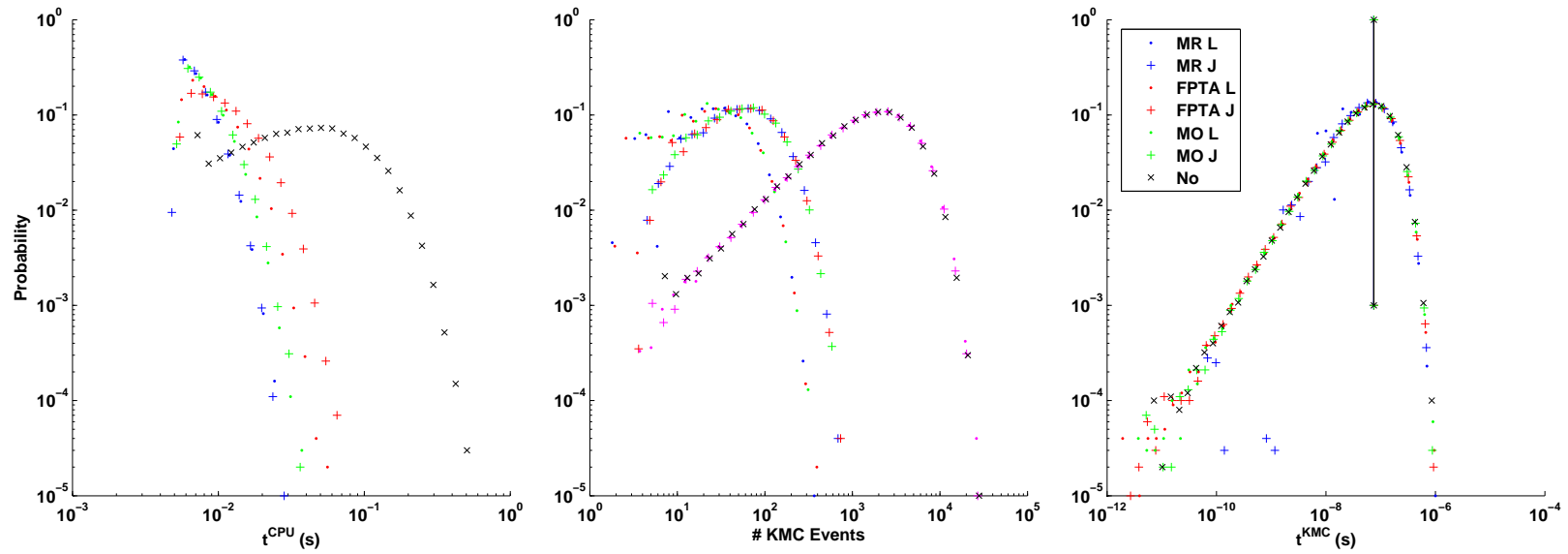


Figure 6.2: Timing and KMC event histograms for comparison for simulations of V_1As_1 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. Vertical lines in the t^{KMC} plot indicate means, and all fall on top of each other. Magenta symbols in the “# KMC Events” plot are the number of unaccelerated events that would have occurred in the memory only simulations.

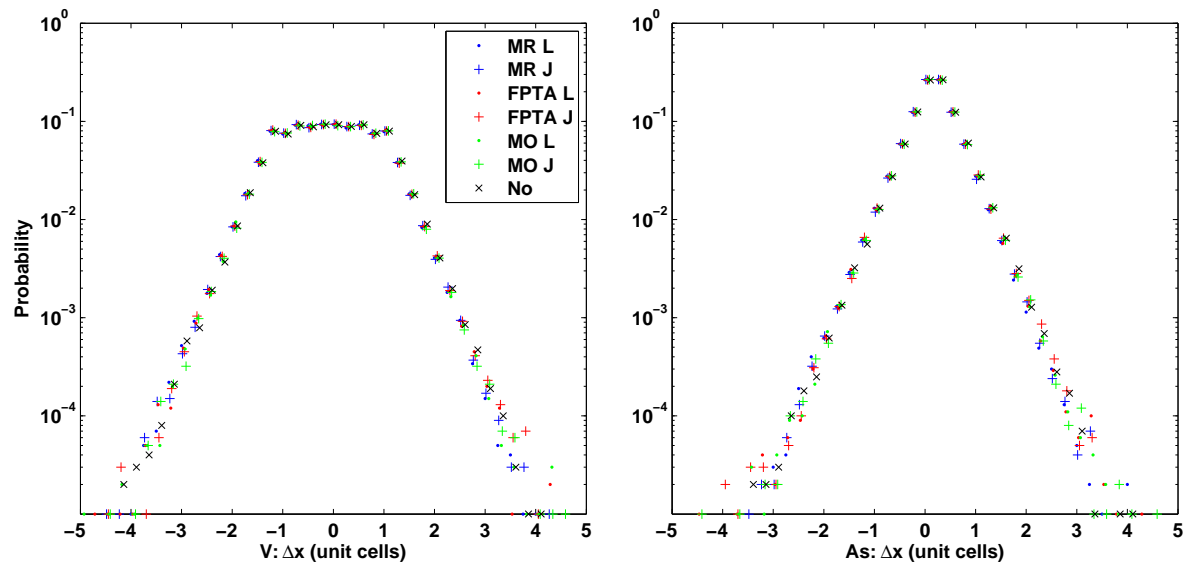


Figure 6.3: Histograms of vacancy and As displacement upon V_1As_1 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first.

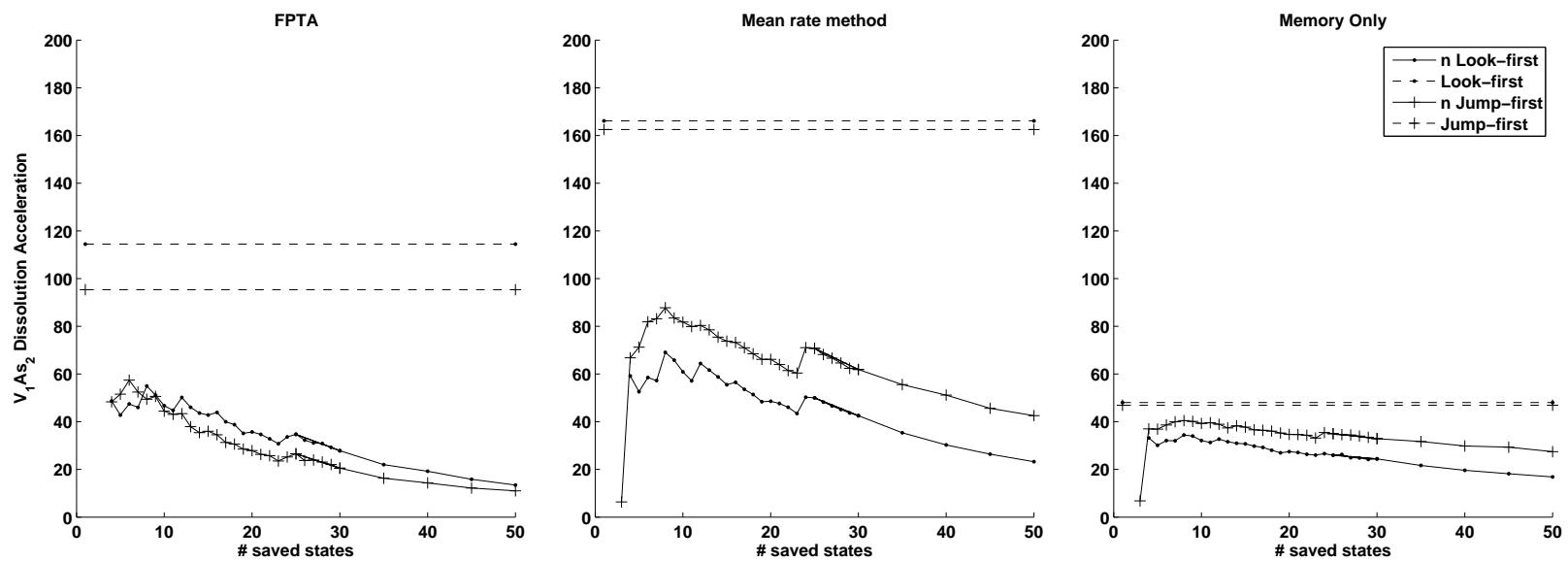


Figure 6.4: Acceleration achieved during V_1As_2 dissolution. The look-first and jump-first lines are the mean of 10^5 samples, and each data point for the n -methods is the mean of 10^4 samples. Acceleration is relative to the mean of 10^5 unaccelerated samples.

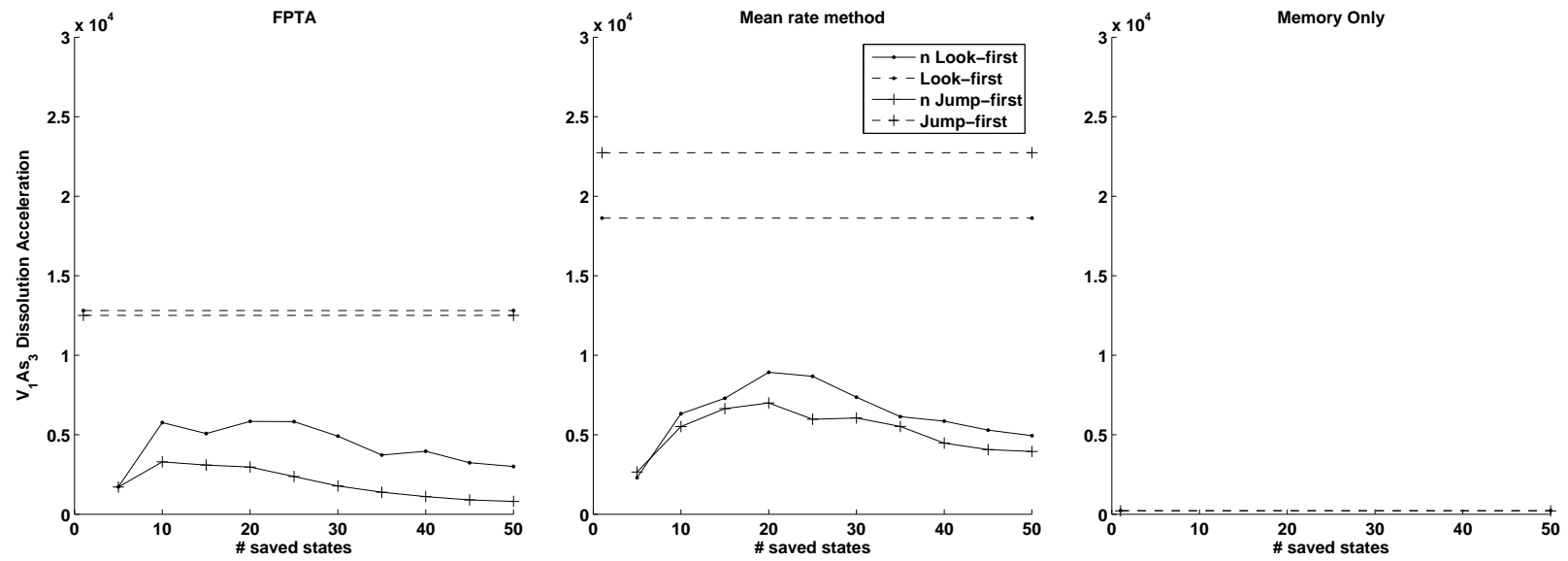


Figure 6.5: Acceleration achieved during V_1As_3 dissolution. The look-first and jump-first lines are the mean of 10^5 samples, except for memory only which is the mean of 4×10^3 samples. Each data point for the n -methods is the mean of 10^3 samples. The n -method with memory only was too slow to be done. Acceleration is relative to the mean of 5 unaccelerated samples.

CHAPTER 7

Percolation Effects on Vacancy-Mediated Dopant Diffusion in Silicon

At large length scales and moderate to low concentrations, diffusion is often well-described by continuum laws. But, at high concentrations and atomic length scales continuum theories break down. In this chapter we use the accelerated KMC algorithms presented in Chapter 6 to study one instance of this breakdown at high concentration using the model for V mediated As diffusion in silicon presented in Section 4.5.

7.1 Percolating Dopant Interactions

From the work reviewed in Chapter 2 and Section 4.4 we know that the presence of dopants in silicon can lower the formation energy of nearby-mediating defects. In the case of As and V this interaction ranges at least three lattice spaces [45, 83, 85, 117, 137], creates the tendency to cluster [42, 43, 138, 139, 140], and leads to time and concentration dependent diffusivity [39, 84, 85, 113]. Mathiot and Pfister [114] recognized that at high concentrations these interactions will percolate throughout the material, and pair or cluster based diffusion models must fail because there are no longer isolated pairs or clusters of V and As. They proposed that the additional interaction with nearby As, as in Figure 7.1, will lower V formation ener-

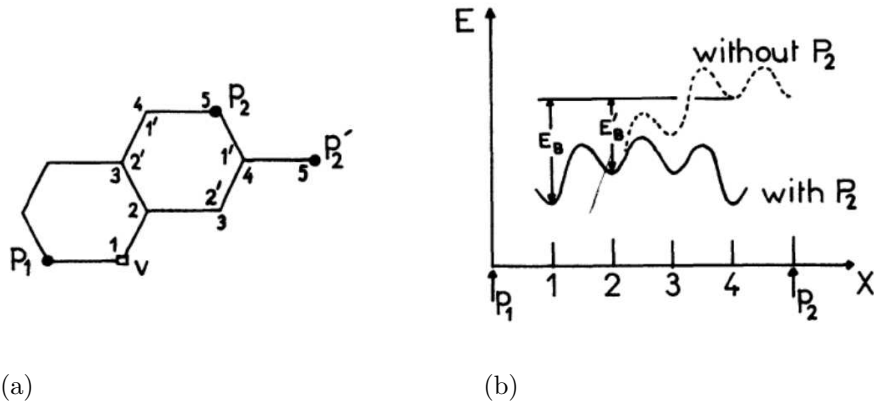


Figure 7.1: The effect of nearby dopants, P, on V formation and migration energies. From [114].

gies, increasing the equilibrium number of V, and lower migration energies, allowing a V to “freely diffuse” throughout the “infinite cluster”. With *ab initio* calculations, Xie and Chen [85] confirmed that placing a second As 5 or 6 nearest neighbor (NN) jumps away from the first lowers the migration barrier for a V to diffuse from one to the other. They noted that this is consistent with experimental results showing enhanced diffusion for short times [84]. However, they also considered that more slowly diffusing As_2V clusters should quickly form as As find their neighbors, slowing diffusion. This is consistent with longer time experiments which show As diffusivity leveling off [113] or decreasing significantly [39] at high doping levels. Kinetic Monte Carlo (KMC) simulations provide a direct way to test these ideas and determine how much percolation of dopant interactions may enhance V diffusivity, how much As diffusivity is increased in turn, and how long and over what distances such effects persist.

Previous KMC work [115, 116, 117] has focused on setting equilibrium V concentrations based on the As concentration, but V may not be in equilibrium in many systems of interest, and this makes it difficult to separate the effects of V concentration and V-V interactions from the effects of the percolating As interactions. In this

work, we focus on the effects of percolating As interactions on V and As diffusivity using a fixed V concentration and non-interacting V by performing simulations with a single V. We determine diffusivities by averaging over hundreds or thousands of identical simulations. KMC simulations in the literature have been limited to relatively short times and may not account for all of the time dependent clustering effects. By using the accelerated KMC algorithm of Chapter 6 we can perform these simulations on timescales up to minutes of simulated time.

We use small-scale simulations of cluster diffusion to determine reaction rates for cluster formation and breakup in a mean field model that describes well the low As concentration time dependent diffusion in larger-scale simulations. Around the critical As concentration identified in experiments, $2-3 \times 10^{20} \text{ cm}^{-3}$, the mean field model becomes inadequate to describe the diffusivity. We follow the trajectories of individual V and find that at high As concentrations individual V do not diffuse freely through the material, but rather become quickly bound in clusters.

All of our simulations have been run in cubic systems with periodic boundaries at 800 °C. During the course of the simulations we keep track of dopant clustering, using the methods described in Section 6.2.6. We define a cluster as all As and V that are 1NN with each other. Varying this clustering definition up to 4NN can affect the cluster formation and breakup reaction rates used in the mean field model, but cause no qualitative differences, so for simplicity we use the 1NN definition and only report those results here. Our simulations track the total time spent in each type of cluster over the entire data output interval, which is important for avoiding sampling errors. We report tracer diffusivities calculated as in Equation 4.19. Data is collected at logarithmic intervals at the same time in every sample so that we can calculate averages. All of the simulations used the look-first state saving method.

7.2 Parameterization Simulations

In “Parameterization” simulations, we randomly placed one V and one to four As into systems 8 unit cells on a side. The simulations with one to three As were run long enough for equilibrium to be reached. The four As cluster simulations did not equilibrate until past 10^5 s which made running a statistically significant number of samples difficult even with the KMC acceleration. Statistics were collected from runs of 10 s, at which point V_1As_3 clusters have formed and a quasi-steady state occurs.

The mean field model for clustering includes binary reactions of the form $V_1As_m + As_n \longleftrightarrow V_1As_{m+n}$, with forward and backward cluster reaction rates k_m^f and $k_{m,n}^b$, to model the cluster population evolutions as

$$(7.1) \quad \frac{d[V_1As_m]}{dt} = \sum_{0 < i \leq m} (k_{m-i}^f [V_1As_{m-i}] [As_i] - k_{m-i,i}^b [V_1As_m]) + \sum_{0 < n} (-k_m^f [V_1As_m] [As_n] + k_{m,n}^b [V_1As_{m+n}]).$$

We included up to V_1As_4 clusters in the mean field model and solved the resulting system of equations using ode15s in MATLAB [141].

However, to determine the cluster reaction rates from the results of the small scale “Parameterization” simulations it is not possible to use Equation 7.1 directly. The mean field model assumes that there is a continuum of cluster concentrations, but the simulations are actually discrete. For example, in the one As simulations, the defects can either be clustered or unclustered. In the two As simulations there are four states: V_1As_2 , or V_1 & As_2 , or V_1As_1 & As_1 , or V_1 & $2As_1$. Since it is a discrete state system, the probability evolution of the system through the states can be calculated using the same methods used for FPTA, but in this case there is no absorbing state. The transition rates in the \bar{M} matrix that fit the simulation data are the cluster reaction rates, k_m^f and $k_{m,n}^b$. To fit the reaction rates, we started with

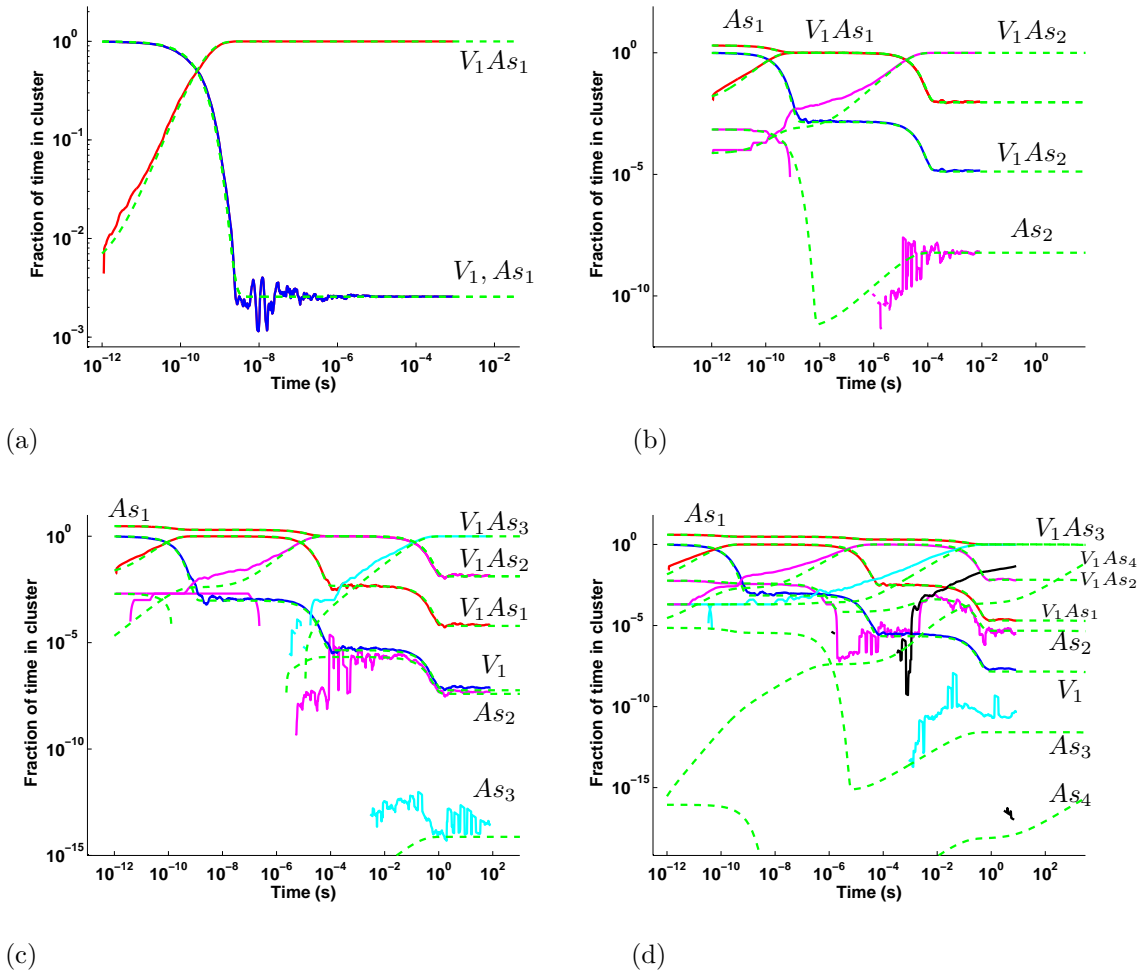


Figure 7.2: The fraction of time spent in various clusters as a function of simulated time during “Parameterization” simulations with (a) one to (d) four As. The solid lines are simulation data and the dashed green lines is the mean field model fit.

the one As simulations. The steady state provides a ratio of forward to backwards reaction rates and then the one remaining free parameter is fit to the transient curve, as shown in Figure 7.2(a).

In the two As simulations the $V_1As_1 + As_1 \rightarrow V_1 + 2As_1$ rate should be the same as the $V_1As_1 \rightarrow V_1 + As_1$ rate in the one As simulations because it is simply the V_1As_1 dissolution rate. We found the best fit for this dissolution rate varies slightly, less than 10% between the one As and four As simulations. Similarly, the $V_1 + As_2 \rightarrow V_1As_2$ formation rate can be assumed identical to the $V_1 + As_1 \rightarrow V_1As_1$

formation rate because it depends on the V_1 diffusivity. It is reasonable to consider that the rates might differ if As_2 presents a larger cross section to the V_1 , but we found that the approximation that they are identical is good. The $V_1 + 2As_1 \rightarrow V_1As_1 + As_1$ formation rate is twice the $V_1 + As_1 \rightarrow V_1As_1$ formation rate because the concentration of As_1 is doubled. This leaves three rates left to be parameterized: the $V_1As_2 \rightarrow V_1 + As_2$ breakup rate, the $V_1As_2 \rightarrow V_1As_1 + As_1$ breakup rate, and the $V_1As_1 + As_1 \rightarrow V_1As_2$ formation rate. They can be fit using the process of determining forward to backward ratios from the steady state and then fitting the transient curves. The resulting fit is shown in Figure 7.2(b). The deviations between the data and model for the V_1As_2 curves arise from the binary reaction assumption. The simulation data includes some reactions that are essentially $V_1 + As_1 + As_1 \rightarrow V_1As_2$, but the model does not. The process was continued for the three and four As simulations, shown in Figure 7.2(c,d).

Finally, the steady-state data was also used to calculate the cluster diffusivities from

$$(7.2) \quad D^{As,V} = \sum_m f_{V_1As_m}^{As,V} D_{V_1As_m},$$

where $D^{As,V}$ is the overall (As,V) diffusivity, while $D_{V_1As_m}$ is the diffusivity of the V_1As_m cluster and $f_{V_1As_m}^{As,V}$ is the fraction of (As,V) time spent in the V_1As_m cluster. The resulting values for k_m^f , $k_{m,n}^b$ and $D_{V_1As_m}$, are listed in Table 7.1. Diffusivities for V_1As_3 and V_1As_4 clusters were too small to be measured.

7.3 Standard Diffusion Simulations

In ‘‘Standard’’ simulations, we randomly placed one V and a varying number of As, corresponding to As concentrations from 10^{19} to 10^{21} cm^{-3} , into large systems 32 unit cells on a side. Diffusivity and clustering were measured in 10^2 – 10^3 samples

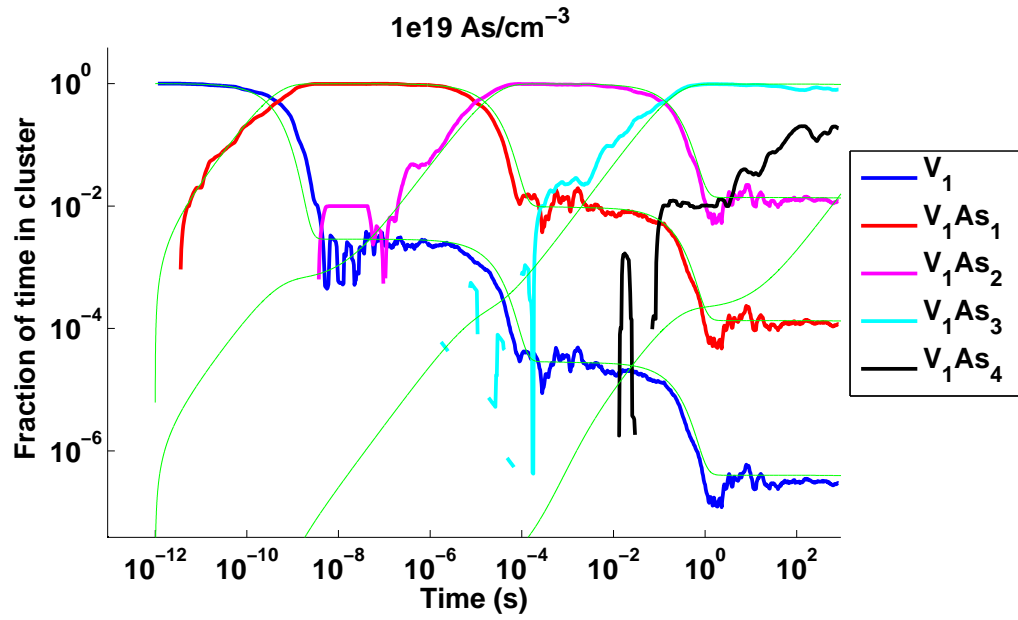
Table 7.1: Cluster reaction rate and diffusivity results from the “Parameterization” simulations.

m	k_m^f (cm^3s^{-1})	$k_{m,1}^b$ (s^{-1})	$k_{m,2}^b$ (s^{-1})	$k_{m,3}^b$ (s^{-1})	$k_{m,4}^b$ (s^{-1})	$D_{V_1As_m}$ (nm^2s^{-1})
0	1.354×10^{12}	7.5×10^6	16	2×10^{-5}	9×10^{-6}	1.785×10^{10}
1	2.2×10^7	4.0×10^2	4.3×10^{-2}	5×10^{-6}	-	4.69×10^5
2	2.6×10^3	2.5×10^{-2}	~ 0	-	-	90.3

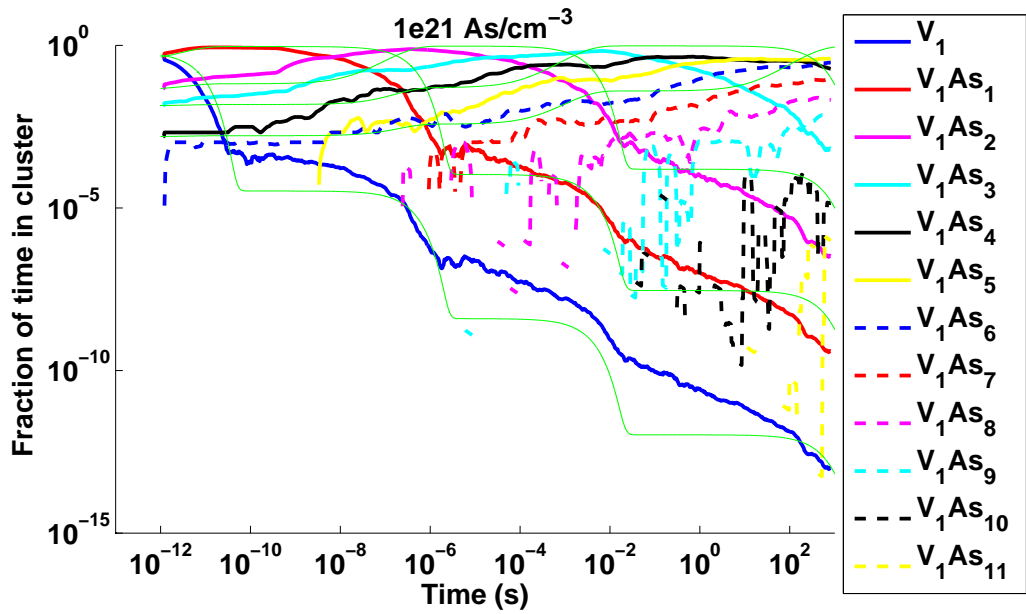
each run for 10^3 s of simulated time.

Figure 7.3 shows the results of the “Standard” KMC simulations for clustering compared to the mean field model predictions. The mean field model describes the clustering well at low concentrations, and the results at $1 \times 10^{19} \text{ cm}^{-3}$ in Figure 7.3(a) are representative. Clustering occurs with time in steps at well-defined average times when the next largest cluster forms. At high concentrations, as in Figure 7.3(b), larger clusters form more continuously in time and the mean field model predictions do not fit as well.

Figure 7.4 shows the the results of the “Standard” KMC simulations for diffusivity compared to the mean field model predictions. At low concentrations, diffusivity drops with time in steps at well-defined average times corresponding to the formation of the next largest cluster, and the mean field model predictions fit the KMC data well. At high concentrations, diffusivity drops continuously with time as clusters form continuously, and is higher than that predicted by the mean field model, except at the longest times. Generally, diffusivity of both V and As drops with increasing As concentration. For V diffusivity this results from binding more often, and for As diffusivity this occurs because there is only one V to mediate diffusion and more As to average over. Variations of V and As diffusivity above the mean field model predictions increase at high As concentrations, and at some times ($\sim 10^{-9}$ and 10^{-5} s) V diffusivity increases slightly with As concentration.

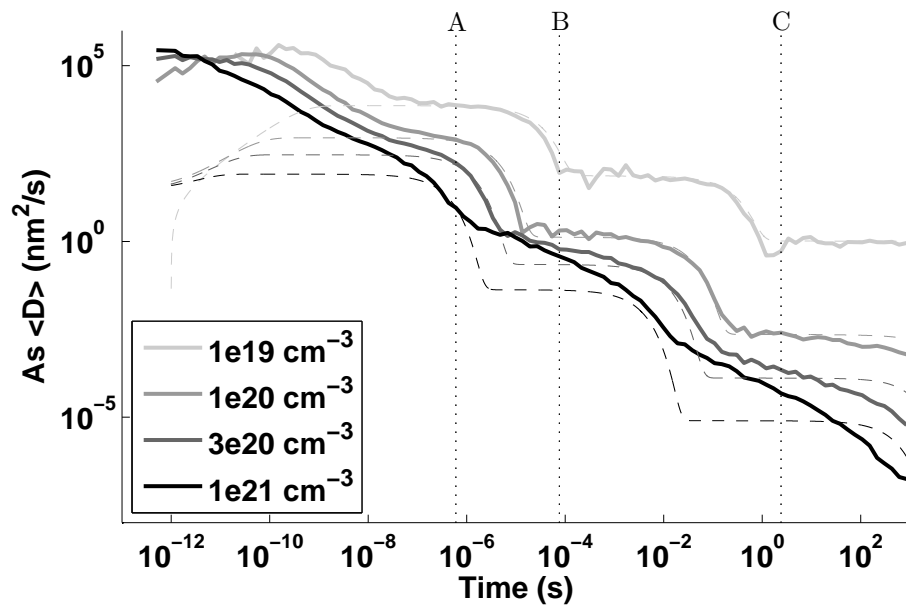


(a)

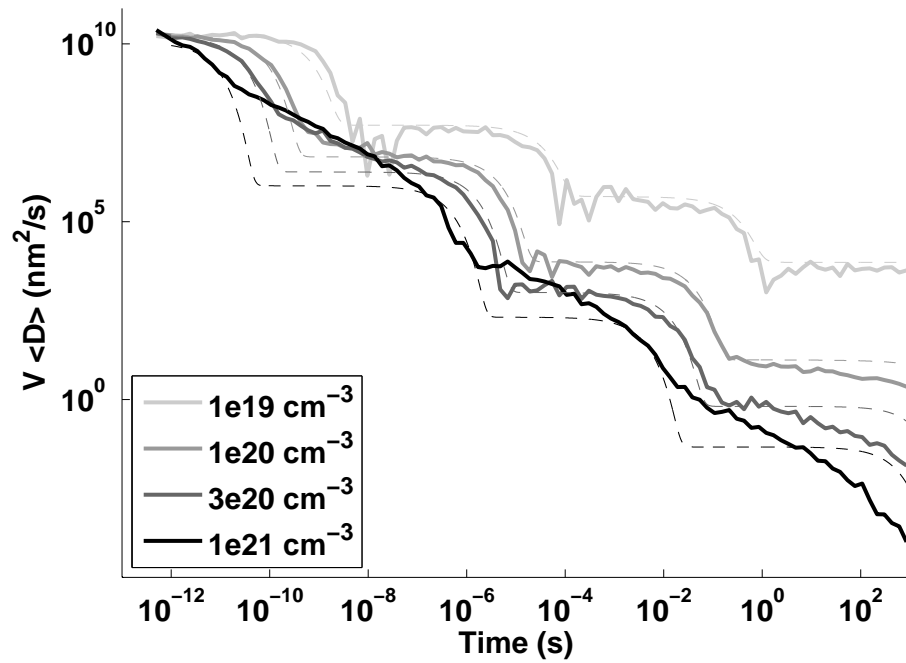


(b)

Figure 7.3: The fraction of time spent by the V in various clusters as a function of simulated time during “Standard” simulations. Thin green lines are mean field model predictions.



(a)



(b)

Figure 7.4: As and V diffusivities in simulations with one V ($2 \times 10^{17} \text{ cm}^{-3}$) and varying As concentration. Thin dashed lines are mean field model predictions. Dotted lines indicate the times at which lines A-C are plotted in Figure 7.5.

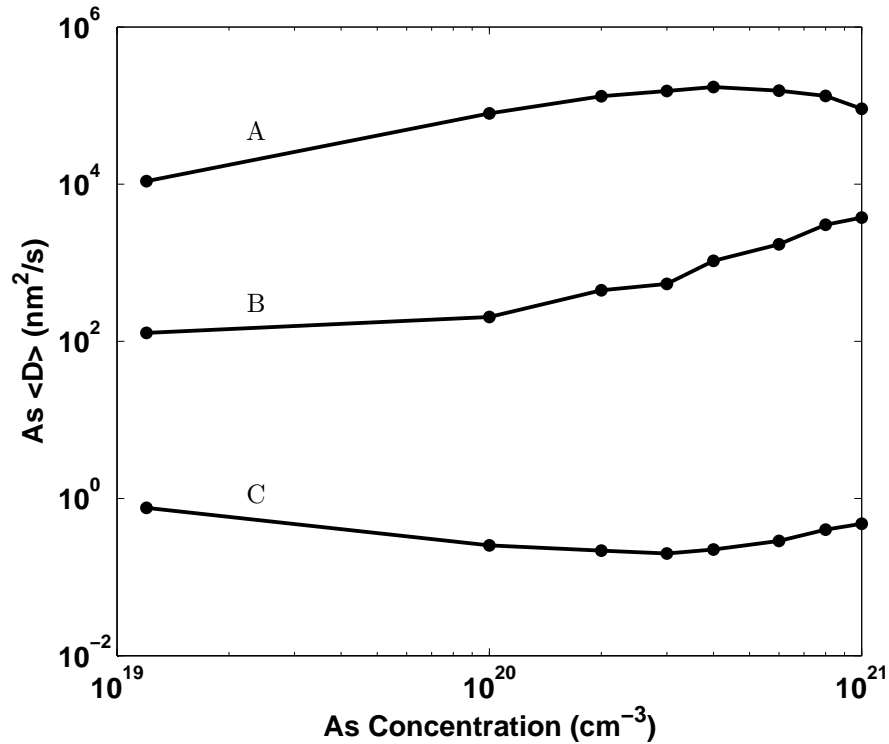


Figure 7.5: As diffusivity as a function of As concentration for the three times indicated in Figure 7.4. The As diffusivity is scaled to include the effect of a V concentration increasing with the square of the As concentration.

Differences in Figure 7.4 between the mean field predictions and KMC results for As diffusivity at the shortest times arise from V-As exchanges that do not contribute to diffusion. In order for an AsV to diffuse, at least 1.19 eV is needed to move a V around a hexagonal ring from a 1NN to at least a 3NN position and then back to a different 1NN position. Simple V-As exchanges only require 0.66 eV but by themselves cannot contribute to diffusion. Similar effects also happen in larger clusters. The KMC simulations measure the r^2 displacements of the non-diffusive exchanges which appear as high diffusivities because they happen at very short times. However, the mean field model only includes diffusive motions, creating the discrepancy.

With a fixed concentration of non-interacting V, the “Standard” simulations are designed to investigate the effect of percolating defect-dopant interactions on diffusiv-

ity at a constant V concentration. In order to fully replicate experimental conditions and results for As diffusivity as a function of dopant concentration, both V concentration effects and V-V interactions must be included. The effect of V-V interactions is unclear: they might decrease diffusivity due to formation of V_m clusters, but it might also increase diffusivity if V_mAs_n clusters are more mobile than V_1As_n clusters. Ignoring V-V interactions, we can make a rough comparison to experimental results by scaling the simulations results by the expected experimental V concentration. The simulations have a V concentration of $2 \times 10^{17} \text{ cm}^{-3}$, which is high compared to expected experimental values, especially at low As concentrations. The plots would all shift up or down based on the actual V concentration. Additionally, the V concentration can be expected to vary based on the As concentration. Increasing donor concentration lowers the Fermi level and results in increasing equilibrium negatively charged free V concentrations [5]. The equilibrium total V concentration would also increase with As concentration since the V formation energy is lowered near As. Therefore, it is reasonable to expect the equilibrium V concentration might increase proportionally to the square of the dopant concentration. If we make this assumption, and scale the As diffusivity results of Figure 7.4 (a) accordingly, we get the results in Figure 7.5. In this figure the As diffusivity results are plotted as a function of As concentration for the three times indicated in Figure 7.4 (a). The experimentally observed trend that As diffusivity increases proportionally with As concentration at concentrations below $2 \times 10^{20} \text{ cm}^{-3}$, see Figure 4.2, is only reproduced at the shortest times, as in line A, when diffusion is dominated by V_1As_1 clusters. In the experimental results of Fair *et al.* this trend persisted to times as long as hours. In the simulations, once the lone V becomes trapped in a V_1As_2 or larger cluster, on the order of a microsecond, this trend stops. However, in experi-

ments, free vacancies could diffuse in to a region in which V_1As_2 or larger clusters have formed so that the number of V_1 and V_1As_1 is not depleted as in the simulations. This could lead to a persistence of the increasing diffusivity trend to longer times as observed experimentally. When V_1As_2 or V_1As_3 clusters form, this trend is no longer observed, but as in line B and C, diffusivity enhancement above As concentrations of $2 \times 10^{20} \text{ cm}^{-3}$ is observed for a short period of time immediately after the average time those clusters form. At longer times, As diffusivity drops with As concentration.

7.4 Short Diffusion Simulations

In order to understand what is happening at the highest concentrations we turn to the results of the “Short” simulations. In “Short” simulations, we used the same setup as for “Standard”, but only ran the simulation until the V found itself 1NN to two or more As. We performed 10^5 such trials at each concentration, measuring the r^2 displacement, time elapsed, and maximum saddle point energy traversed starting from when the V was first 1NN with one As to when it was 1NN with two or more As.

When As are isolated it takes 1.21 eV for a V to unbind. Percolation enhanced diffusion is proposed to occur when this energy is lowered due to a second As nearby. Because the binding energies are additive in our simulations, if the V is 1NN to two As it will take 1.21 eV for it to migrate to a position 1NN with a single As. This is as great as the energy to fully unbind a V from an isolated As, so the potential diffusion enhancement from percolating dopant interactions will be gone. Certainly, additional As in 2NN and 3NN positions will also modify the migration energies, but not to the extent of the 1NN As. For this reason, we consider the V becoming 1NN to two or more As a stopping point for any percolation enhanced diffusion.

Histograms in Figure 7.6 show the r^2 displacement and time elapsed starting from when the V was first 1NN with one As to when it was 1NN with two or more As. We see that increasing the As concentration shortens the r^2 displacements, since the average distance between dopants is lower. But the finite nature of the lattice means the displacements can not decrease forever, and near the critical concentrations $2-3 \times 10^{20} \text{ cm}^{-3}$ the range of observed displacements is approaching the size of the unit cell. Above this concentration, the histograms shift left as the displacement can not be reduced much more, but the clustering time decreases, effectively a small diffusivity increase. The tail of data to the left at very short times comes from V that find As that are already clustered. Additionally, we see that increasingly the maximum saddle point traversed is lower than the unbound maximum of 0.28 eV, a hallmark of the percolation model. However, the small r^2 values demonstrate that at high concentrations the V do not diffuse freely through the material from As to As, but instead either quickly find an existing As_2 pair or rearrange existing single As into a cluster. Qualitatively similar results, shifted to higher diffusivity, have been obtained at 1000 °C and 1200 °C.

7.5 Conclusions

In conclusion, the work presented in this chapter shows that at low As concentrations a mean field model predicts well the diffusivity of V and As observed in KMC simulations and can be used as a baseline for determining diffusivity changes at high As concentrations. On short time scales, percolation of dopant interactions does lead to limited increased diffusivity for individual V, and in turn this causes increased As diffusivities. The effect is limited in magnitude and we do not observe V “freely diffusing” through the material as a result of percolating dopant interactions,

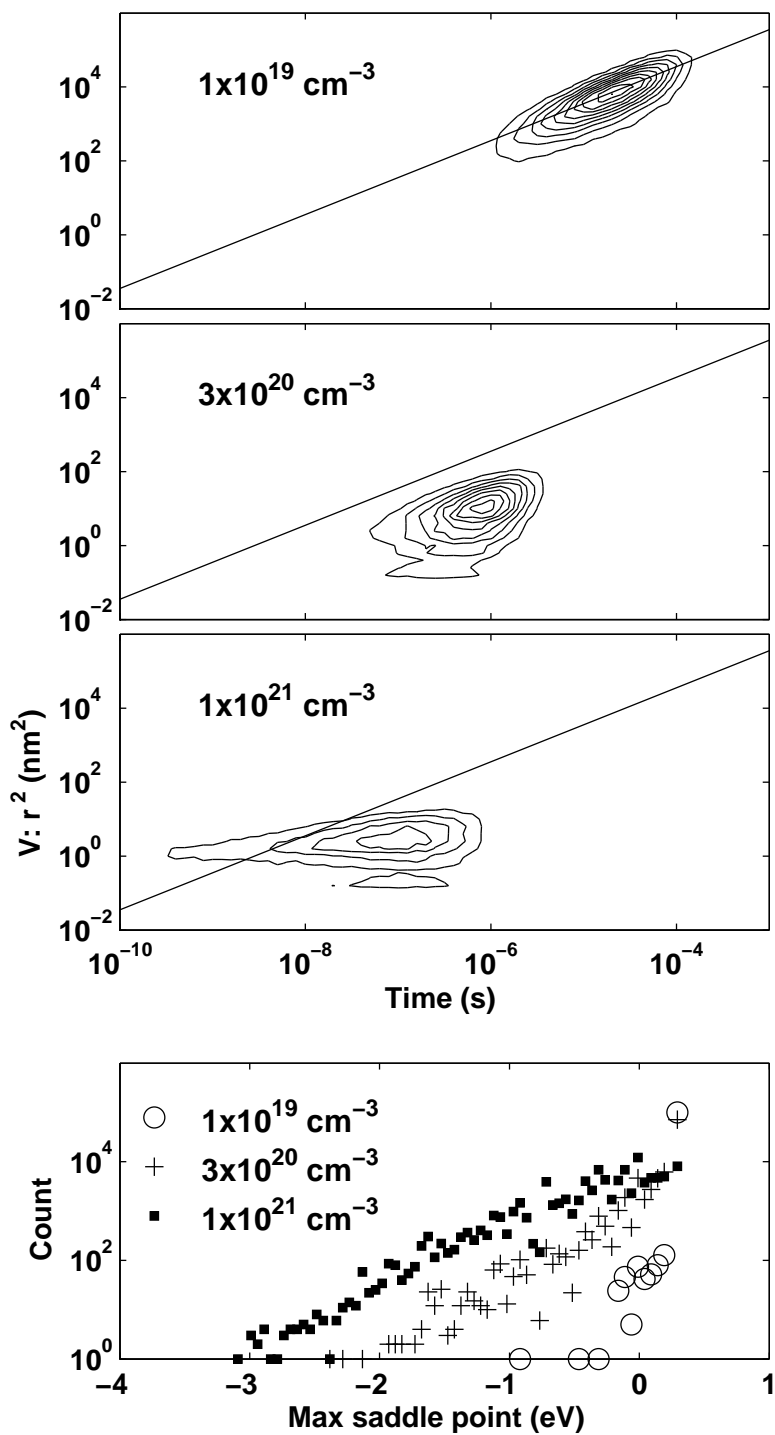


Figure 7.6: Histograms showing the time, $V r^2$ displacement, and maximum saddle point traversed as recorded for 10^5 samples at varying As concentration from when a V is first 1NN with one As until it is 1NN with two or more As. Contour lines are at intervals of 10% of the peak in the top plot. A line of constant diffusivity is shown for reference.

but rather observe quick local clustering that decreases diffusivity on timescales of seconds to minutes.

CHAPTER 8

Conclusions and Future Work

From the work we have reviewed, it is clear that the current understanding of dopant diffusion in silicon comes from the synthesis of experimental and computational research. Atomistic calculations provide valuable understanding and parameterization of the mechanisms that contribute to diffusion. The relevant physical phenomena range over many time and length scales, which necessitates a multi-scale modeling approach. Our work has contributed to two different aspects of the effort to model dopant diffusion, (1) the accuracy of atomistic methods for calculating defect parameters, and (2) an accelerated KMC method, which we used to investigate the effects of percolating dopant-defect interactions on diffusion.

By using continuum linear elasticity to formally show the effect of boundary conditions on supercell calculations we have validated the supercell approach for atomistic calculations. Our comparisons with empirical atomistic calculations verify the linear elastic prediction in the far-field, and provide useful insights into the convergence of atomistic calculations. For instance, we found that although linear elasticity predicts periodic boundary conditions at zero average stress have no effect on relaxation volume tensor calculations, in practice bonding changes result in elastic moduli changes that cause the relaxation volume tensor to converge as $1/N$.

For future atomistic calculations, knowledge of this convergence rate allows for accurate estimation of the relaxation volume tensor from modest simulation sizes. The elastic moduli might also be measured and compared to the value expected for a perfect crystal in order to gauge convergence. We have shown that the thermodynamically relevant volume change is that of the entire body, so relaxation volumes should be reported accordingly, rather than simply describing local relaxations. Additionally, even if the computational cell is fixed at a constant volume the relaxation volume tensor can be simply calculated from the average stress in the cell.

We have introduced an accelerated KMC method that automatically determines which states comprise trapping energy basins. The accelerated KMC method has allowed us to study the highly time dependent diffusion of arsenic in silicon. We demonstrated the use of the accelerated method to determine cluster diffusivities and reaction rates which were used in a mean field model. We observed that percolating dopant interactions do lead to limited increased diffusivity, but that the effect is limited in magnitude and duration as immobile clusters form quickly.

Future work on the accelerated KMC method could focus on several different aspects. Perhaps the first thing to consider is optimizing the efficiency using the current method. Several variables have been introduced without optimization, such as the number of basins or saved states allowed, or the number of times basins are occupied before merging. Since the optimal values may vary not only between systems and simulations, but even for different situations within the same simulation, it might worthwhile to vary them based on the current situation. It should be feasible to vary the limits during the course of a simulation and automatically optimize them using a Monte Carlo method. Some efficiency gains could also be realized by improving on the bracketing and bisection method used to determine exit times in

FPTA. There may also be improvements possible with the implementation of the bracketing and bisection method. For instance, during the selection of the exit time with FPTA, the probability of being in the absorbing state is calculated at several different times. Saving those results in memory so that they can be used when the system re-enters the basin could also increase efficiency.

Larger efficiency gains might be possible with an efficient method to characterize how “trapping” a basin is, i.e. how good the equilibrating basin approximation is. For instance, if simple information about a basin could be used to place bounds on the errors introduced from using the equilibrating basin acceleration method, then it might be used more often. We expect that it is a function of some generalized basin width and depth. The fewer transitions it takes for the system to jump between all the states in the basin, the “narrower” it is and the better the equilibrating basin approximation should be. The higher the internal transition energies are compared to the transitions to the absorbing states are, the “shallower” the basin is and the worse the equilibrating basin approximation should be. The initial state of the system within the basin is also important. The closer the system starts to the transient state that it spends the most time in on average, the better the equilibrating basin approximation should be. Similar information may provide an efficient way to determine ahead of time which basins should be merged without waiting for the system to repeatedly visit them. It might also help avoid ill-conditioning problems in FPTA calculations by allowing the certain states to be merged into equilibrating basins so that very large differences between rates in the absorbing Markov chain can be avoided.

Future work may also attempt to use the accelerated method for systems with high concentrations of diffusing defects. At high concentrations, many defects may

interact with each other and result in a multitude of states that must be saved. The acceleration methods can theoretically handle any number of states, but it remains to be seen how many states are practically treatable before efficiency drops. The accelerated algorithm's success may depend on the system of interest and whether the energy landscape consists of "wide", "shallow" basins or "narrow", "deep" basins. At high concentrations and very large basins, there might also be more complications collecting data. The larger the basin is, the more likely that relevant quantities vary within the basin. It may be advantageous to segment the large basin into smaller regions and perform FPTA on each of them separately. Data could be collected for each of the regions and then the escape calculated for the overall basin. At some point when the basin becomes very wide and shallow, this would begin to resemble the first-passage time methods of DeVita *et al.* [127] and Opplestrup *et al.* [128].

In fact, a combination of those methods and our method may be very effective. Our energy basin algorithm can be used to calculate cluster diffusivity and reaction rates from simulations that include all of the atomic jumps. Then those rates could be used with the DeVita or Opplestrup methods to simulate the motion of reacting clusters. This kind of approach would likely be best at lower concentrations when isolated diffusing clusters can be identified.

One final consideration is that the properties of energy basins may themselves be valuable for describing a system. The eigenvalues calculated during the course of FPTA provide information on the timescale of relevant mechanisms. This may provide as much information on the onset of clustering or other phenomena as more visually understandable, but practically difficult, criteria such as cluster connection distances. By basing data collection on the underlying energy landscape the most relevant mechanisms may be revealed and the basin finding algorithms will not be

slowed down.

The energy basin algorithms presented for accelerating KMC simulations may be very useful in a wide variety of applications. Particularly in systems with low concentrations of diffusing species but deep trapping energy basins, our accelerated algorithm will allow simulations to reach very long times compared to standard KMC simulations. This may be especially useful for determining the diffusivity or dissolution rates of clusters of crystalline defects. By considering issues such as grouping isolated diffusing species and collecting data when the exact location of the system within an energy basin is not resolved, we have provided an example that can be followed when applying this method to other systems.

APPENDICES

APPENDIX A

The Jump-First Method

```

1 Input defect positions and bin defects, see Section 6.2.1;
2 Set initial system time;
3 Determine the next output time;
4 Determine the clustering, see Section 6.2.6.1;
5 Determine the grouping, set initial group time, see Section 6.2.2;
6 Create parameterization lookup tables and calculate jump rates, see Section 6.2.3;
7 for Each group do
8   | Save the first state, see Algorithm A.2;
9   | Determine the group events, see Algorithm A.4;
10 end
11 repeat
12   | Choose the group with next occurring event;
13   if The next event occurs past the next output time then
14     | if Using FPTA or Using memory only then
15       | Update all groups to the next output time, see Section 6.2.8;
16       | Re-determine the group events, see Algorithm A.4;
17     end
18     | Output data and get the next output time;
19   else
20     | Update the system time and chosen group time to the event time;
21     if Move by a K transition then
22       | Load the state the system is jumping to, see Section 6.2.4.1;
23     else
24       | Load the state the system is jumping from, see Section 6.2.4.1;
25       | Perform the move, Algorithm A.5;
26     end
27     if Impinging upon other groups then
28       | Update those groups to the system time, see Section 6.2.8;
29       | Update jump rates, see Section 6.2.3;
30       | Clear saved states and basins, and regroup, see Section 6.2.2;
31       | Save the state, see Algorithm A.2;
32     else
33       | Update jump rates, see Section 6.2.3;
34       | Update the basin, see Algorithm A.7;
35     end
36     | Determine group events for chosen group and any new groups, see Algorithm A.4;
37   end
38 until Simulation finishes ;

```

Algorithm A.1: A detailed outline of a KMC program using the jump-first method.

```

1 while Identifying states in the superstate, see Section 6.2.4.3 do
2   if #States > limit then
3     | Delete a basin, see Section 6.2.5.2;
4   end
5   Save the state, see Section 6.2.4.1;
6   if Grouping changes while identifying the superstate then
7     | if Heading up the superstate then
8       | Place the system in the stable state;
9     | end
10    | Clear saved states and basins, and regroup, see Section 6.2.2;
11    | Restart save state, this Algorithm;
12  | end
13  | if Too many As jump while identifying the superstate, see Section 6.2.4.1 then
14    | if Heading up the superstate then
15      | Place the system in the stable state;
16    | end
17    | Clear saved states and basins;
18    | Restart save state, this Algorithm;
19  | end
20 end
21 Place the system in the stable state;
22 Create a superstate, save the list of states to the superstate, see Section 6.2.4.3;
23 Update clusters, see Section 6.2.6.1;
24 Save the cluster connections in the superstate, see Section 6.2.6.2;
25 if First superstate then
26   | Create a new basin, see Section 6.2.5;
27   | Update basin connectivity, see Section 6.2.5;
28   | Set  $\beta$ , see Algorithm A.3;
29   | Save the current clusters to the basin, see Section 6.2.6.3;
30 end

```

Algorithm A.2: Save a superstate in the jump-first method.

```

1 if  $\#min[UK] > 1$  or  $min[UK] \geq K$  then
2   |  $\beta = \uparrow$ ;
3 else
4   |  $\beta = \downarrow$ ;
5 end

```

Algorithm A.3: The conditional setting of β upon creating a new basin. Known (K) and unknown (UK) transitions are from the current state only.

```

1 if Falling to UK exit, see Algorithm A.7 then
2 |   The UK exit is the next move and no time is incremented;
3 else if No saved states, see Section 4.3 then
4 |   Use standard KMC to pick the V, jump direction, and jump time;
5 else if Using FPTA, see Section 6.2.7.1 then
6 |   if FPTA solution does not exist for the current basin then
7 |     Calculate the FPTA solution and save to the basin;
8 |   end
9 |   Use the FPTA solution to pick the event time;
10 |  Use the event time to pick the escape move;
11 else if Using the mean rate method, see Section 6.2.7.2 then
12 |  Use the mean rate method to pick the escape move;
13 |  Use the mean time for the event time;
14 else if Using memory only, see Section 6.2.7.3 then
15 |  Use standard KMC to move through the basin until the system escapes;
16 |  Use the escape time, and move for the event;
17 end

```

Algorithm A.4: Determining group events for either the jump-first or look-first method.

```

1 Move V position and check V bin, see Section 6.2.1;
2 Note any groups impinged upon, see Section 6.2.2;
3 if Exchange with As then
4 |   Move As position and check As bin, see Section 6.2.1;
5 |   if This As is not already in the jumped As list, see Section 6.2.4.1 then
6 |     if #Jumped As > limit then
7 |       Add this As to the list, and record its original position;
8 |     else
9 |       Clear saved states and basins;
10 |    end
11 |  end
12 end

```

Algorithm A.5: Perform a V move in the jump-first method.

```

1 if  $(\beta = \uparrow \text{ and } \min[UK] \leq K)$  or Clustering changed then
2 |   In a different basin;
3 else
4 |   Still in current basin;
5 |   if  $\beta = \downarrow \text{ and } \min[UK] \geq K$  then
6 |      $\beta = \uparrow$ ;
7 |   end
8 end

```

Algorithm A.6: The criteria for determining if the system is still in the current basin. Known (K) and unknown (UK) transitions are from the current state only.


```

1 if Move was by UK transition then
2   if In a previously saved state, see Section 6.2.4.2 then
3     if Still in the same basin and Using FPTA then
4       Clear FPTA solution, see Section 6.2.7.1;
5     end
6   end
7 end
8 if If K transition or In a previously saved state then
9   Update state connectivity, see Section 6.2.4.1;
10  Fall to either K superstate stable state or to UK exit, see Section 6.2.4.3;
11  if Fall to UK exit then
12    Load superstate cluster connections, see Section 6.2.6.2;
13    Load basin clusters, see Section 6.2.6.3;
14  else if Fall to K Superstate stable state then
15    Load superstate stable state, see Section 6.2.4.3;
16    Load superstate cluster connections, see Section 6.2.6.2;
17    Load basin clusters, see Section 6.2.6.3;
18    Update basin connectivity, see Section 6.2.5;
19    Increment basin occupation count;
20    Check if basins should merge, see Section 6.2.5.1;
21  end
22 else
23   Occasionally check if still grouped, see Section 6.2.2;
24   if Still grouped then
25     if #Changed clusters > limit, see Section 6.2.6.3 then
26       Clear saved states and basins;
27       Save first state, see Algorithm A.2;
28     else
29       Save new state, see Algorithm A.2;
30       Update state connectivity, see Section 6.2.4.1;
31       Decide if we are in the same basin, Algorithm A.6;
32       if Still in current basin then
33         Add state to current basin;
34       else
35         if #Basins > limit then
36           Delete a basin, see Section 6.2.5.2;
37         end
38         Add state to a new basin, see Section 6.2.5;
39         Update basin connectivity, see Section 6.2.5;
40         Set  $\beta$ , see Algorithm A.3;
41         Save the current clusters to the basin, see Section 6.2.6.3;
42         Update basin connectivity, see Section 6.2.5;
43       end
44     end
45   end
46 end

```

Algorithm A.7: Updating the basin in the jump-first method.

APPENDIX B

The Look-First Method

```

1 Input defect positions and bin defects, see Section 6.2.1;
2 Set initial system time;
3 Determine the next output time;
4 Determine the clustering, see Section 6.2.6.1;
5 Determine the grouping, set initial group time, see Section 6.2.2;
6 Create parameterization lookup tables and calculate jump rates, see Section 6.2.3;
7 for Each group do
8   | Save the first basin, see Algorithm B.2;
9   | Determine the group events, see Algorithm B.4;
10 end
11 repeat
12   | Choose the group with next occurring event;
13   if The next event occurs past the next output time then
14     | if Using FPTA or Using memory only then
15       | Update all groups to the next output time, see Section 6.2.8;
16       | Re-determine the group events, see Algorithm B.4;
17     end
18     | Output data and get the next output time;
19   else
20     | Update the system time and chosen group time to the event time;
21     if Move by a K transition then
22       | Load the state the system is jumping to, see Section 6.2.4.1;
23     else
24       | Load the state the system is jumping from, see Section 6.2.4.1;
25       | Perform the move, Algorithm B.6;
26     end
27     if Impinging upon other groups then
28       | Update those groups to the system time, see Section 6.2.8;
29       | Update jump rates, see Section 6.2.3;
30       | Clear saved states and basins, and regroup, see Section 6.2.2;
31       | Save the basin, see Algorithm B.2;
32     else
33       | Update jump rates, see Section 6.2.3;
34       | Update the basins, see Algorithm B.5;
35     end
36     | Determine group events for chosen group and any new groups, see Algorithm B.4;
37   end
38 until Simulation finishes ;

```

Algorithm B.1: A detailed outline of a KMC program using the jump-first method.

```

1 Save first state, see Algorithm B.3;
2 Update clusters, see Section 6.2.6.1;
3 if #Basins > limit then
4 | Delete a basin, see Section 6.2.5.2;
5 end
6 Add state to new basin;
7 Update basin connectivity, see Section 6.2.5;
8 stop = false;
9 if #min[UK] > 1 and min[UK] ≤ K then
10 | stop = true;
11 else if min[UK] ≥ K then
12 | β = ↑;
13 else
14 | β = ↓;
15 end
16 while stop == false do
17 | Considering all states in current basin, choose minimum energy UK transition;
18 | Perform a move, see Algorithm B.6;
19 | if In a previously saved state, see Section 6.2.4.2 then
20 | | Update state connectivity, see Section 6.2.4.1;
21 | | if In a different basin then
22 | | | Update basin connectivity, see Section 6.2.5;
23 | | | stop = true;
24 | | end
25 | else
26 | | Update jump rates, see Section 6.2.3;
27 | | Save a state, see Algorithm B.3;
28 | | Check if clustering changed, see Section 6.2.6.1;
29 | | if (β = ↑ and min[UK] ≤ K) or Clustering changed or Groups impinged or Too
30 | | many As jumped then
31 | | | Delete the state;
32 | | | stop = true;
33 | | else
34 | | | if β = ↓ and min[UK] ≥ K then
35 | | | | β = ↑;
36 | | | end
37 | | | Update state connectivity, see Section 6.2.4.1;
38 | | | Save the state in the current basin;
39 | | | Save the cluster connections in the superstate, see Section 6.2.6.2;
40 | | end
41 | end
42 Place system in first state in current basin;
43 Save the current clusters to the basin, see Section 6.2.6.3;

```

Algorithm B.2: The algorithm for identifying basins in the look-first method.

```

1 while Identifying states in the superstate, see Section 6.2.4.3 do
2   if #States > limit then
3     | Delete a basin, see Section 6.2.5.2;
4   end
5   Save the state, see Section 6.2.4.1;
6   if Impinging upon other groups while identifying the superstate then
7     | if First superstate in the basin then
8       | Update impinged groups to the system time, see Section 6.2.8;
9       | Clear saved states and basins, and regroup, see Section 6.2.2;
10      | Re-start save the basin, see Algorithm B.2;
11     else
12       | Delete the states in this superstate;
13       | Groups impinged in Algorithm B.2;
14       | Return;
15     end
16   end
17   if Too many As jump while identifying the superstate, see Section 6.2.4.1 then
18     | if First superstate in the basin then
19       | Increase As jump limit;
20     else
21       | Delete the states in this superstate;
22       | Too many As jumped in Algorithm B.2;
23       | Return;
24     end
25   end
26 end
27 Place the system in the stable state;
28 Create a superstate, save the list of states to the superstate, see Section 6.2.4.3;

```

Algorithm B.3: Save a superstate in the look-first method.

```

1 if Falling to UK exit, see Algorithm A.7 then
2   | The UK exit is the next move and no time is incremented;
3 else if No saved states, see Section 4.3 then
4   | Use standard KMC to pick the V, jump direction, and jump time;
5 else if Using FPTA, see Section 6.2.7.1 then
6   | if FPTA solution does not exist for the current basin then
7     | Calculate the FPTA solution and save to the basin;
8   end
9   | Use the FPTA solution to pick the event time;
10  | Use the event time to pick the escape move;
11 else if Using the mean rate method, see Section 6.2.7.2 then
12  | Use the mean rate method to pick the escape move;
13  | Use the mean time for the event time;
14 else if Using memory only, see Section 6.2.7.3 then
15  | Use standard KMC to move through the basin until the system escapes;
16  | Use the escape time, and move for the event;
17 end

```

Algorithm B.4: Determining group events for either the jump-first or look-first method.

```

1 if Move was by UK transition then
2   if In a previously saved state, see Section 6.2.4.2 then
3     if Still in the same basin and Using FPTA then
4       | Clear FPTA solution, see Section 6.2.7.1;
5     end
6   end
7 end
8 if If K transition or In a previously saved state then
9   | Update state connectivity, see Section 6.2.4.1;
10  | Fall to either K superstate stable state or to UK exit, see Section 6.2.4.3;
11  if Fall to UK exit then
12    | Load superstate cluster connections, see Section 6.2.6.2;
13    | Load basin clusters, see Section 6.2.6.3;
14  else if Fall to K Superstate stable state then
15    | Load superstate stable state, see Section 6.2.4.3;
16    | Load superstate cluster connections, see Section 6.2.6.2;
17    | Load basin clusters, see Section 6.2.6.3;
18    | Update basin connectivity, see Section 6.2.5;
19    | Increment basin occupation count;
20    | Check if basins should merge, see Section 6.2.5.1;
21  end
22 else
23   | Occasionally check if still grouped, see Section 6.2.2;
24   if Still grouped then
25     | if #Changed clusters > limit, see Section 6.2.6.3 then
26       | Clear saved states and basins;
27     end
28     | Save a new basin, see Algorithm B.2;
29   end
30 end

```

Algorithm B.5: Updating the basins in the look-first method.

```

1 Move V position and check V bin, see Section 6.2.1;
2 Note any groups impinged upon, see Section 6.2.2;
3 if Exchange with As then
4   | Move As position and check As bin, see Section 6.2.1;
5   if This As is not already in the jumped As list, see Section 6.2.4.1 then
6     | if #Jumped As > limit then
7       | Clear saved states and basins;
8     else
9       | Add this As to the list, and record its original position;
10    end
11  end
12 end

```

Algorithm B.6: Perform a V move in the look-first method.

APPENDIX C

V_1As_2 and V_1As_3 Dissolution Figures

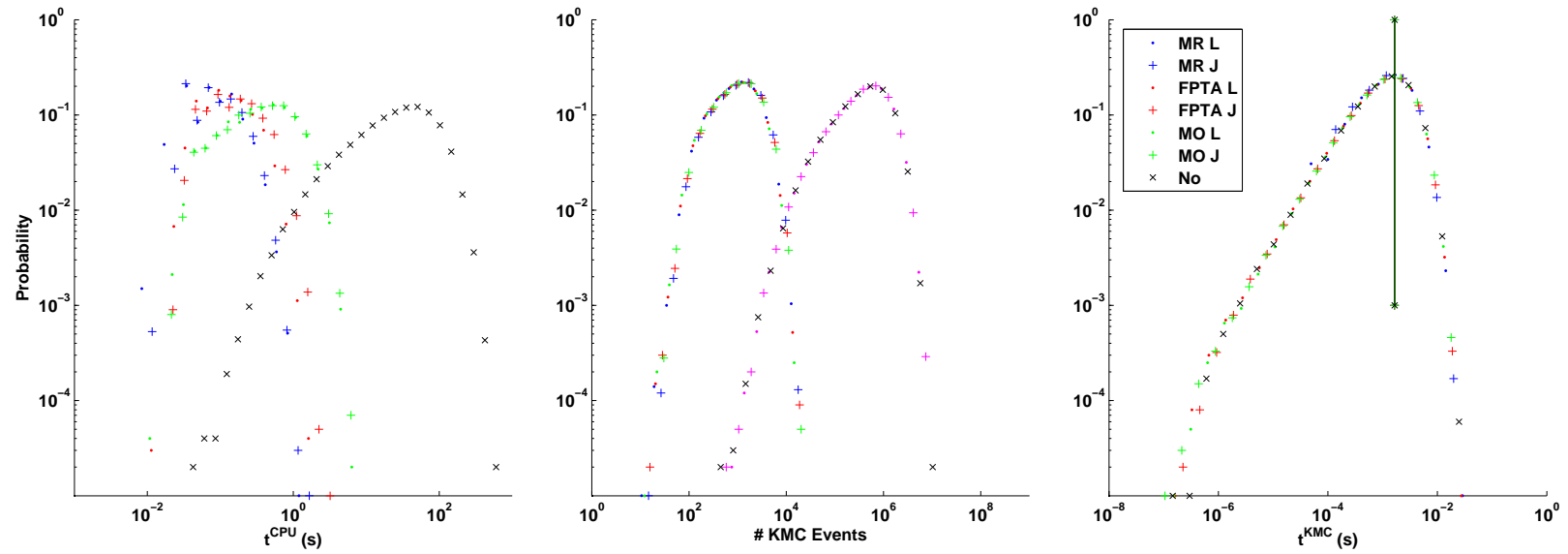


Figure C.1: Timing and KMC event histograms for comparison of simulations of V_1As_2 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. Vertical lines in the t^{KMC} plot indicate means, and all fall on top of each other. Magenta symbols in the “# KMC Events” plot are the number of unaccelerated events that would have occurred in the memory only simulations.

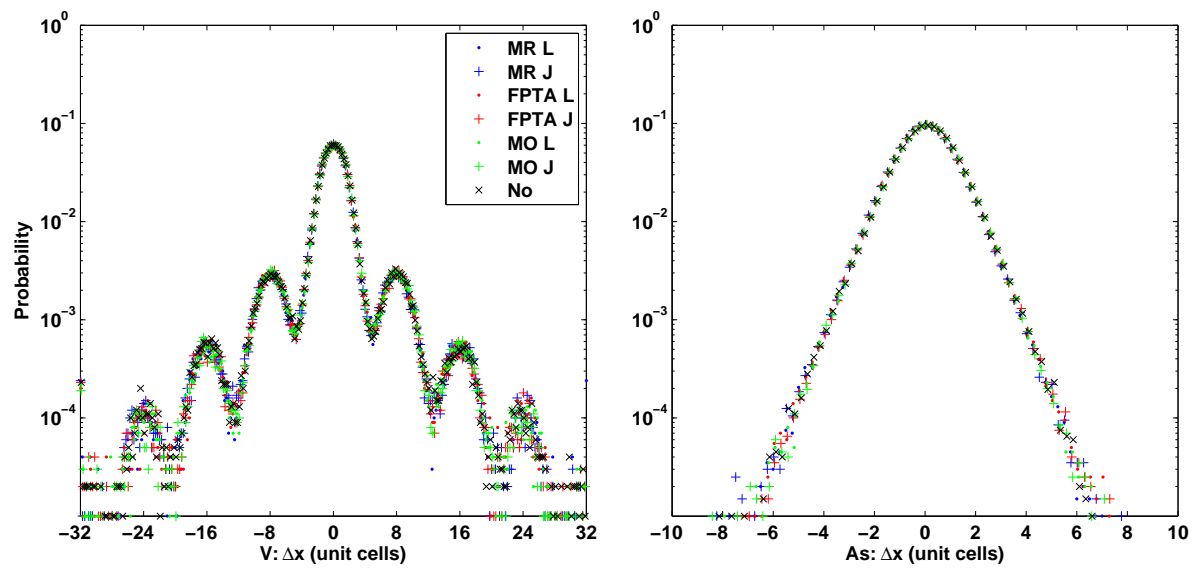


Figure C.2: Histograms of vacancy and As displacement upon V_1As_2 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. The additional maxima in vacancy displacement are due to vacancies which cross the periodic boundaries before the As are completely dissolved.

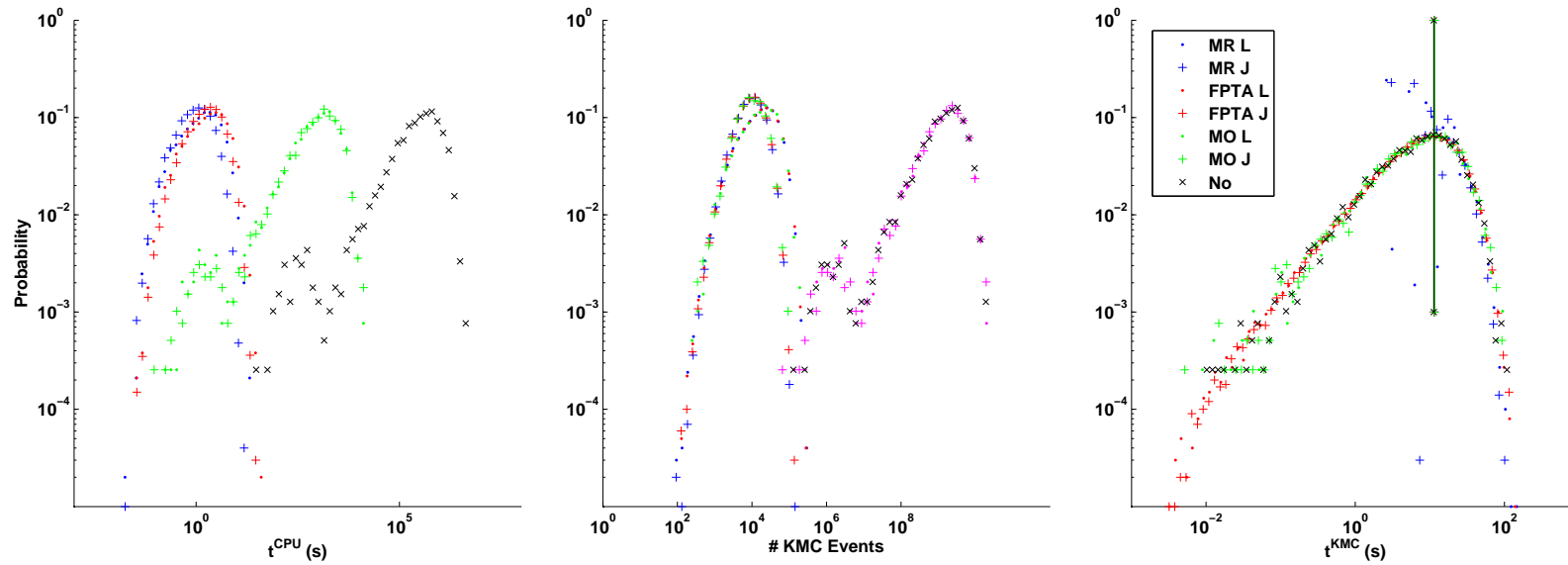


Figure C.3: Timing and KMC event histograms for comparison of simulations of V_1As_3 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. Vertical lines in the t^{KMC} plot indicate means, and all fall on top of each other. Magenta symbols in the “# KMC Events” plot are the number of unaccelerated events that would have occurred in the memory only simulations. There is a minimum possible probability for memory only because there are only 4×10^3 samples. The unaccelerated results are estimated from the number of explicit jumps in the memory only simulations.

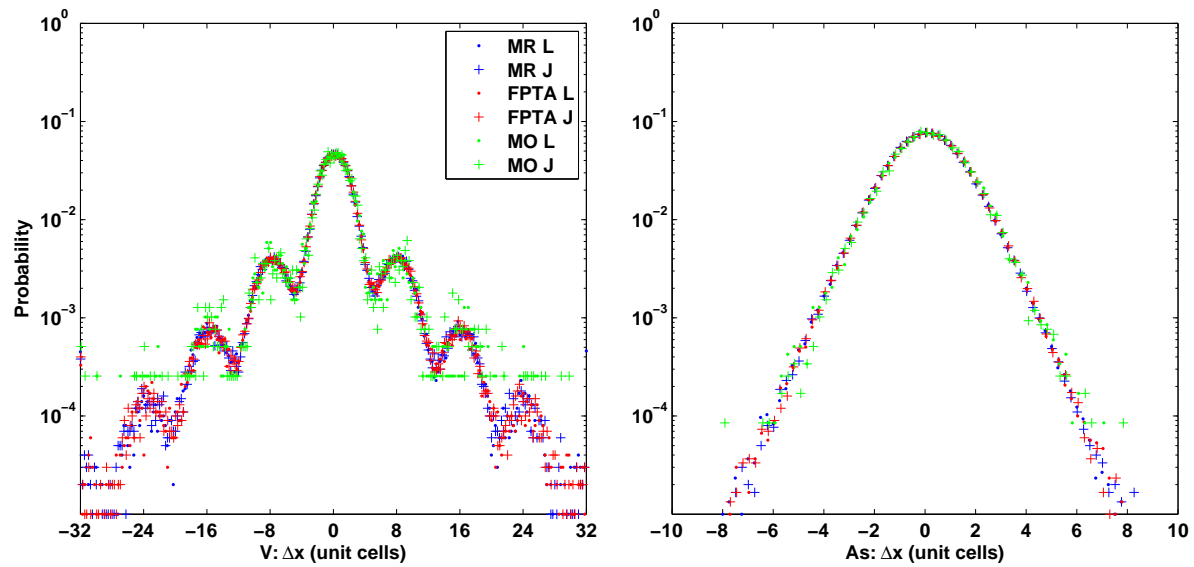


Figure C.4: Histograms of vacancy and As displacement upon V_1As_3 dissolution. The labels “MR” = mean rate method, “MO” = memory only, “No” = no acceleration, “L” = look-first, and “J” = jump-first. There is a minimum possible probability for memory only because there are only 4×10^3 samples. The additional maxima in vacancy displacement are due to vacancies which cross the periodic boundaries before the As are completely dissolved.

REFERENCES

REFERENCES

- [1] K. Garikipati, M. Falk, M. Bouville, B. Puchala, and H. Narayanan. The continuum elastic and atomistic viewpoints on the formation volume and strain energy of a point defect. *J. Mech. Phys. Sol.*, 54:1929, 2006.
- [2] B. Puchala, M. L. Falk, and K. Garikipati. Elastic effects on relaxation volume tensor calculations. *Phys. Rev. B*, page 174116, 2008.
- [3] B. Puchala, M. L. Falk, and K. Garikipati. Accelerated kinetic Monte Carlo simulations of percolation effects on vacancy-mediated dopant diffusion in Si. submitted to *Phys. Rev. B*, 2009.
- [4] P. A. Packan. Scaling transistors into the deep-submicron regime. *MRS Bull.*, 25:18, 2000.
- [5] P. M. Fahey, P. B. Griffin, and J. D. Plummer. Point defects and dopant diffusion in silicon. *Rev. Mod. Phys.*, 61:289, 1989.
- [6] S. Tyagi, C. Auth, P. Bai, G. Curello, H. Deshpande, S. Gannavaram, O. Golonzka, R. Heussner, R. James, C. Kenyon, S-H Lee, N. Lindert, M. Liu, R. Nagisetty, S. Natarajan, C. Parker, J. Sebastian, B. Sell, S. Sivakumar, A. St. Amour, and K. Tone. An advanced low power, high performance, strained channel 65nm technology. In *2005 IEEE International Electron Devices Meeting (IEDM) December 5-7*, page 245, 2005.
- [7] H. Park, K. S. Jones, J. A. Slinkman, and M. E. Law. Effects of hydrostatic pressure on dopant diffusion in silicon. *J. Appl. Phys.*, 78:3664, 1995.
- [8] M. J. Aziz. Thermodynamics of diffusion under pressure and stress: Relation to point defect mechanisms. *Appl. Phys. Lett.*, 70:2810, 1997.
- [9] A. F. Voter, F. Montalenti, and T. C. Germann. Extending the time scale in atomistic simulation of materials. *Annu. Rev. Mater. Res.*, 32:321, 2002.
- [10] P. Shewmon. *Diffusion in Solids*. The Minerals, Metals & Materials Society, 1989.
- [11] J. G. Mullen. Effect of Bardeen-Herring correlation on vacancy diffusion in anisotropic crystals. *Phys. Rev.*, 6:1723, 1961.
- [12] K. Compaan and Y. Haven. Correlation factors for diffusion in solids. *Trans. Faraday Soc.*, 52:786, 1956.
- [13] P. H. Dederichs and K. Schroeder. Anisotropic diffusion in stress fields. *Phys. Rev. B*, 17:2524, 1978.
- [14] M. S. Daw, W. Windl, N. N. Carlson, M. Laudon, and M. P. Masquelier. Effect of stress on dopant and defect diffusion in Si: A general treatment. *Phys. Rev. B*, 64:045205, 2001.
- [15] Figure from <http://cst-www.nrl.navy.mil/lattice/struk/a4.html>.

- [16] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. A new algorithm for Monte Carlo simulation of Ising spin systems. *J. Comput. Phys.*, 17:10, 1975.
- [17] K. A. Fichtorn and W. H. Weinberg. Theoretical foundations of dynamical Monte Carlo simulations. *J. Chem. Phys.*, 95:1090, 1991.
- [18] D. A. Porter and K. E. Easterling. *Phase Transformation in Metals and Alloys, 2nd edition*. Nelson Thornes Ltd, 2001.
- [19] H. Bracht, E. E. Haller, and R. Clark-Phelps. Silicon self-diffusion in isotope heterostructures. *Phys. Rev. Lett.*, 81:245502, 1998.
- [20] A. Ural, P. B. Griffin, and J. D. Plummer. Self-diffusion in silicon: similarity between the properties of native point defects. *Phys. Rev. Lett.*, 83:3454, 1998.
- [21] A. Ural, P. B. Griffin, and J. D. Plummer. Experimental evidence for a dual vacancy-interstitial mechanism of self-diffusion in silicon. *Appl. Phys. Lett.*, 73:1706, 1998.
- [22] A. Ural, P. B. Griffin, and J. D. Plummer. Fractional contributions of microscopic diffusion mechanisms for common dopants and self-diffusion in silicon. *J. Appl. Phys.*, 85:6440, 1998.
- [23] H. J. Gossmann, T. E. Haynes, P. A. Stolk, D. C. Jacobson, G. H. Gilmer, J.M. Poate, H.S. Luftman, T.K. Mogi, and M.O. Thompson. The interstitial fraction of diffusivity of common dopants in Si. *Appl. Phys. Lett.*, 71:3862, 1997.
- [24] C. S. Nichols, C. G. Van de Walle, and S. T. Pantelides. Mechanisms of dopant impurity diffusion in silicon. *Phys. Rev. B*, 40:5484, 1989.
- [25] N. A. Stolwijk, B. Schuster, J. Holzl, H. Mehrer, and W. Frank. Diffusion and solubility of gold in silicon. *Physica B*, 116:335, 1983.
- [26] F. Morehead, N. A. Stolwijk, W. Meyberg, and U. Gosele. Self-interstitial and vacancy contributions to silicon self-diffusion determined from the diffusion of gold in silicon. *Appl. Phys. Lett.*, 42:690, 1983.
- [27] H. Bracht, N. A. Stolwijk, and H. Mehrer. Properties of intrinsic point defects in silicon determined by zinc diffusion experiments under nonequilibrium conditions. *Phys. Rev. B*, 52:16542, 1995.
- [28] H. Bracht. Diffusion mediated by doping and radiation-induced point defects. *Physica B*, 376-377:11, 2006.
- [29] V. V. Voronkov and R. Falster. Intrinsic point defects in silicon: a unified view from crystal growth, wafer processing and metal diffusion. *Sol. St. Phen.*, 108-109:1, 2005.
- [30] G. D. Watkins. Defects in irradiated silicon: EPR and electron-nuclear double resonance of interstitial boron. *Phys. Rev. B*, 12:5824, 1975.
- [31] M. Hirata, M. Hirata, and H. Saito. The interactions of point defects with impurities in silicon. *J. Phys. Soc. Japan*, 27:405, 1969.
- [32] E. L. Elkin and G. D. Watkins. Defects in irradiated silicon: electron paramagnetic resonance and electron-nuclear double resonance of the arsenic- and antimony-vacancy pairs. *Phys. Rev.*, 174:881, 1968.
- [33] G. D. Watkins and J. W. Corbett. Defects in irradiated silicon: electron paramagnetic resonance and electron-nuclear double resonance of the Si-e center. *Phys. Rev.*, 134:A1359, 1964.
- [34] A. Seeger and K. P. Chik. Diffusion mechanisms and point defects in silicon and germanium. *Phys. Stat. Sol.*, 29:455, 1968.

- [35] V. Ranki and K. Saarinen. Formation of thermal vacancies in highly As and P doped Si. *Phys. Rev. Lett.*, 93:255502, 2004.
- [36] S. Dannefaer, P. Mascher, and D. Kerr. Monovacancy formation enthalpy in silicon. *Phys. Rev. Lett.*, 56:2195, 1986.
- [37] J. Throwe, T. C. Leung, B. Nielsen, H. Huomo, and K. G. Lynn. Search for thermally generated monovacancies in silicon using monoenergetic positrons. *Phys. Rev. B*, 40:12037, 1989.
- [38] R. Wurschum, W. Bauer, K. Maier, A. Seeger, and H.-E. Schaefer. Defects in semiconductors after electron irradiation or in high-temperature thermal equilibrium, as studied by positron annihilation. *J. Phys.: Condens. Matter*, 1:SA33, 1989.
- [39] R. B. Fair and G. R. Weber. Effect of complex formation on diffusion of arsenic in silicon. *J. Appl. Phys.*, 44:273, 1973.
- [40] P. M. Rousseau, P. B. Griffin, and J. D. Plummer. Electrical deactivation of arsenic as a source of point defects. *Appl. Rev. Lett.*, 65:578, 1994.
- [41] V. Ranki, A. Pelli, and K. Saarinen. Formation of vacancy-impurity complexes by annealing elementary vacancies introduced by electron irradiation of As-, P-, and Sb-doped Si. *Phys. Rev. B*, 69:115205, 2004.
- [42] K. C. Pandey, A. Erbil, G. S. Cargill III, R. F. Boehme, and D. Vanderbilt. Annealing of heavily arsenic-doped silicon: Electrical deactivation and a new defect complex. *Phys. Rev. Lett.*, 61:1282, 1988.
- [43] V. Ranki, J. Nissilä, and K. Saarinen. Formation of vacancy-impurity complexes by kinetic processes in highly As-doped Si. *Phys. Rev. Lett.*, 88:105506, 2002.
- [44] C. D. Mueller, E. Alonso, and W. Fichtner. Arsenic deactivation in Si: electronic structure and charge states of vacancy-impurity clusters. *Phys. Rev. B*, 68:045208, 2003.
- [45] J. Xie and S. P. Chen. *Ab initio* calculations of the structure and energetics of As-vacancy complexes in silicon. *J. Phys. D: Appl. Phys.*, 32:1252, 1999.
- [46] M. J. Aziz, Y. Zhao, H. J. Grossman, S. Mitha, S. P. Smith, and D. Schiferl. Pressure and stress effects on the diffusion of B and Sb in Si and Si-Ge alloys. *Phys. Rev. B*, 73:054101, 2006.
- [47] Y. Zhao, M. J. Aziz, H.-J. Gossmann, S. Mitha, and D. Schiferl. Activation volume for boron diffusion in silicon and implications for strained films. *Appl. Phys. Lett.*, 74:31, 1999.
- [48] Y. Zhao, M. J. Aziz, H.-J. Gossmann, S. Mitha, and D. Schiferl. Activation volume for antimony diffusion in silicon and implications for strained films. *Appl. Phys. Lett.*, 74:31, 1999.
- [49] P. Kringhoj, A. Nylandsed Larsen, and S. Yu. Shirayev. Diffusion of Sb in strained and relaxed Si and SiGe. *Phys. Rev. Lett.*, 76:3372, 1996.
- [50] A. Yu. Kuznetsov, J. Cardenas, D. C. Schmidt, and B. G. Svensson. Sb-enhanced diffusion in strained $\text{Si}_{1-x}\text{Ge}_x$: dependence on biaxial compression. *Phys. Rev. B*, 59:7274, 1999.
- [51] A. Portavoce, P. Gas, I. Berbezier, A. Ronda, J.S. Christensen, A. Yu. Kuznetsov, and B. G. Svensson. Sb lattice diffusion in $\text{Si}_x\text{Ge}_{1-x}/\text{Si}(001)$ heterostructures: chemical and stress effects. *Phys. Rev. B*, 69:155415, 2004.
- [52] P. Kuo, J. L. Hoyt, J. F. Gibbons, J. E. Turner, and D. Lefforge. Effects of strain on boron diffusion in Si and $\text{Si}_x\text{Ge}_{1-x}$. *Appl. Phys. Lett.*, 66:580, 1995.

- [53] N. R. Zangenberg, J. Fage-Pedersen, J. Lundsgaard Hansen, and A. Nylandsted Larsen. Boron and phosphorus diffusion in strained and relaxed Si and SiGe. *J. Appl. Phys.*, 94:3883, 2003.
- [54] E. Nygren, M. J. Aziz, D. Turnbull, J. M. Poate, D. C. Jacobson, and R. Hull. Pressure dependence of arsenic diffusivity in silicon. *Appl. Phys. Lett.*, 47:105, 1985.
- [55] S. Uppal, J. M. Bonar, J. Zhang, and A. F. W. Willoughby. Arsenic diffusion in Si and strained $\text{Si}_x\text{Ge}_{1-x}$ alloys at 1000 °C. *Mat. Sci. Eng. B*, 114-115:346, 2004.
- [56] G. D. M. Dilliway, A. J. Smith, J. J. Hamilton, J. Benson, Lu Xu, P. J. McNally, G. Cooke, H. Kheyrandish, and N.E.B. Cowern. Transient enhanced diffusion and deactivation of ion-implanted As in strained Si. *Nucl. Instrum. Meth. B*, 237:131, 2005.
- [57] H. A. Jahn and E. Teller. Stability of polyatomic molecules in degenerate electronic states. I. Orbital degeneracy. *Proc. Roy. Soc. Lond. Ser. A*, 161:220, 1937.
- [58] A. Antonelli and J. Bernholc. Pressure effects on self-diffusion in silicon. *Phys. Rev. B*, 40:10643, 1989.
- [59] O. Sugino and A. Oshiyama. Microscopic mechanism of atomic diffusion in Si under pressure. *Phys. Rev. B*, 46:012335, 1992.
- [60] H. Seong and L. J. Lewis. First-principles study of the structure and energetics of neutral divancies in silicon. *Phys. Rev. B*, 53:9791, 1996.
- [61] M. J. Puska, S. Poykko, M. Pesola, and R. M. Nieminen. Convergence of supercell calculations for point defects in semiconductors: vacancy in silicon. *Phys. Rev. B*, 58:1318, 1998.
- [62] J. Lento and R. M. Nieminen. Non-local screened-exchange calculations for defects in semiconductors: vacancy in silicon. *J. Phys.: Condens. Matter*, 15:4387, 2003.
- [63] M. I. J. Probert and M. C. Payne. Improving the convergence of defect calculations in supercells: an ab initio study of the neutral vacancy in silicon. *Phys. Rev. B*, 67:075204, 2003.
- [64] A. Antonelli, E. Kaxiras, and D. J. Chadi. Vacancy in silicon revisited: structure and pressure effects. *Phys. Rev. Lett.*, 81:2088, 1998.
- [65] S. A. Centoni, B. Sadigh, G. H. Gilmer, T. J. Lenosky, T. Diaz de la Rubia, and C. G. Musgrave. First-principles calculation of intrinsic defect formation volumes in silicon. *Phys. Rev. B*, 72:195206, 2005.
- [66] W. Windl, M. S. Daw, N. N. Carlson, and M. Laudon. Multiscale modeling of stress-mediated diffusion in silicon - volume tensors. In *Mat. Res. Soc. Symp. Proc. Vol. 677*, page AA9.4.1, 2001.
- [67] F. El-Mellouhi, N. Mousseau, and P. Ordejon. Sampling diffusion paths of a neutral vacancy in silicon with quantum mechanical calculations. *Phys. Rev. Lett.*, 70:205202, 2004.
- [68] G. D. Watkins. In S. T. Pantelides, editor, *Deep Centers in Semiconductors*, chapter 3. Gordon and Breach, 1992.
- [69] W.-K. Leung, R. J. Needs, and G. Rajagopal. Calculations of silicon self-interstitial defects. *Phys. Rev. Lett.*, 83:2351, 1999.
- [70] D. M. Ceperley. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566, 1977.
- [71] W. Windl. Diffusion in silicon and the predictive power of *ab initio* calculations. *Phys. Stat. Sol. (B)*, 241:2313, 2004.

- [72] P. E. Blochl, E. Smargiassi, Car R., D. B. Laks, W. Andreoni, and S. T. Pantelides. First-principles calculations of self-diffusion constants in silicon. *Phys. Rev. Lett.*, 70:2435, 1993.
- [73] S. J. Clark and G. J. Ackland. *Ab initio* calculations of the self-interstitial in silicon. *Phys. Rev. B*, 56:47, 1997.
- [74] N. Cowern and C. Rafferty. Enhanced diffusion in silicon processing. *MRS Bull.*, June:39, 2000.
- [75] J. Zhu, T. Diaz de la Rubia, L. H. Yang, C. Mailhot, and G. H. Gilmer. *Ab initio* pseudopotential calculations of B diffusion and pairing in Si. *Phys. Rev. B*, 54:4741, 1996.
- [76] W. Windl, M. M. Bunea, R. Stumpf, S. T. Dunham, and M. P. Masquelier. First-principles study of boron diffusion in silicon. *Phys. Rev. Lett.*, 83:4345, 1999.
- [77] B. Sadigh, T. J. Lenosky, S. K. Theiss, M.-J. Caturla, T. Diaz de la Rubia, and M. A. Foad. Mechanism of boron diffusion in silicon: an *ab initio* and kinetic Monte Carlo study. *Phys. Rev. Lett.*, 83:4341, 1999.
- [78] M. Laudon, N. N. Carlson, M. P. Masquelier, M. S. Daw, and W. Windl. Sb-enhanced diffusion in strained $\text{Si}_{1-x}\text{Ge}_x$: dependence on biaxial compression. *Appl. Phys. Lett.*, 78:201, 2001.
- [79] S. T. Dunham, M. Diebel, C. Ahn, and C. L. Shih. Calculations of effect of anisotropic stress/strain on dopant diffusion in silicon under equilibrium and nonequilibrium conditions. *J. Vac. Sci. Technol. B*, 24(1):456, 2006.
- [80] L. Lin, T. Kirichenko, B. R. Sahu, G. S. Hwang, and S. K. Banarjee. Theoretical study of B diffusion with charged defects in strained Si. *Phys. Rev. B*, 72:205206, 2005.
- [81] M. Ramamoorthy and S. T. Pantelides. Complex dynamical phenomena in heavily arsenic doped silicon. *Phys. Rev. Lett.*, 76:4753, 1996.
- [82] A. Satta, E. Albertazzi, G. Lulli, and L. Colombo. *Ab initio* structures of As_mV complexes and the simulation of Rutherford backscattering channeling spectra in heavily As-doped crystalline silicon. *Phys. Rev. B*, 72:235206, 2005.
- [83] O. Pankratov, H. Huang, T. Diaz de la Rubia, and C. Mailhot. As-vacancy interaction and ring mechanism of diffusion in Si. *Phys. Rev. B*, 56:13172, 1997.
- [84] A. Nylandsted Larsen, K. Kylesbach Larsen, P. E. Andersen, and B. G. Svenson. Heavy doping effects in the diffusion of group IV and V impurities in silicon. *J. Appl. Phys.*, 73:691, 1993.
- [85] J. Xie and S. P. Chen. Diffusion and clustering in heavily arsenic-doped silicon: Discrepancies and explanation. *Phys. Rev. Lett.*, 83:1795, 1999.
- [86] S. A. Harrison, T. F. Edgar, and G. S. Hwang. Structure, stability, and diffusion of arsenic-silicon interstitial pairs. *Appl. Phys. Lett.*, 87:231905, 2005.
- [87] J. S. Nelson, P. A. Schultz, and A. F. Wright. Valence and atomic size dependent exchange barriers in vacancy-mediated dopant diffusion. *Appl. Phys. Lett.*, 73:247, 1998.
- [88] X.-Y. Liu, W. Windl, K. M. Beardmore, and M. P. Masquelier. First-principles study of phosphorus diffusion in silicon: interstitial- and vacancy-mediated diffusion mechanisms. *Appl. Phys. Lett.*, 82:1839, 2003.
- [89] A. Zywiets, J. Furthmuller, and F. Bechstedt. Neutral vacancies in group-IV semiconductors. *Phys. Stat. Sol. (B)*, 210:13, 1996.

- [90] J. D. Eshelby. The determination of the elastic field of an ellipsoidal inclusion and related problems. *Proc. Roy. Soc. Lond. Ser. A*, 241:376, 1957.
- [91] J. D. Eshelby. Elastic inclusions and inhomogeneities. In R. Hill and J. Sneddon, editors, *Progress in Solid Mechanics*, page 89. North Holland Publishing Co., 1961.
- [92] D. M. Barnett. The precise evaluation of derivatives of anisotropic elastic Green's functions. *Phys. Stat. Sol.*, 59:741, 1972.
- [93] F. H. Stillinger and T. A. Weber. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B*, 31:5262, 1985.
- [94] H. Balamane, T. Halicioglu, and W. A. Tiller. Comparative study of silicon empirical interatomic potentials. *Phys. Rev. B*, 46:2250, 1992.
- [95] M. Nastar, V. V. Bulatov, and S. Yip. Saddle-point configurations for self-interstitial migration in silicon. *Phys. Rev. B*, 53:13521, 1996.
- [96] H. R. Schober. Extended interstitials in silicon and germanium. *Phys. Rev. B*, 39:13013, 1989.
- [97] D. Maroudas and R. A. Brown. Calculation of thermodynamic and transport properties of intrinsic point defects in silicon. *Phys. Rev. B*, 47:15562, 1993.
- [98] G. H. Gilmer, T. Diaz de la Rubia, D. M. Stock, and M. Jaraiz. Diffusion and interactions of point defects in silicon: molecular dynamics simulations. *Nucl. Instrum. Meth. B*, 102:247, 1995.
- [99] C. W. M. Castleton and S. Mirbt. Finite-size scaling as a cure for supercell approximation errors in calculations of neutral native defects in inp. *Phys. Rev. B*, 70:195202, 2004.
- [100] R. E. Miller and E. B. Tadmor. Hybrid continuum mechanics and atomistic methods for simulating materials deformation and failure. *MRS Bull.*, 32:920, 2007.
- [101] V. Gavini, Bhattacharya K., and M. Ortiz. Quasi-continuum orbital-free density-functional theory: A route to multi-million atom non-periodic DFT calculation. *J. Mech. Phys. Sol.*, 55:627, 2007.
- [102] J. Q. Broughton, F. F. Abraham, N. Bernstein, and E. Kaxiras. Concurrent coupling of length scales: Methodology and application. *Phys. Rev. B*, 60:2391, 1999.
- [103] A. Einstein. *Investigations into the Theory of the Brownian Movement*. Dover Publications, Inc., 1956.
- [104] R. Gomer. Diffusion of adsorbates on metal surfaces. *Rep. Prog. Phys.*, 53:917, 1990.
- [105] A. Van der Ven, G. Ceder, M. Asta, and P. D. Tepeesch. First-principles theory of ionic diffusion with nondilute carriers. *Phys. Rev. B*, 64:184307, 2001.
- [106] Y. Zhou and G. H. Miller. Green-Kubo formulas for mutual diffusion coefficients in multi-component systems. *J. Phys. Chem.*, 100:5516, 1996.
- [107] R. Kubo. The fluctuation-dissipation theorem. *Rep. Prog. Phys.*, 29:255, 1966.
- [108] S. Ishioka and M. Koiwa. Diffusion coefficient in crystals with multiple jump frequencies. *Philos. Mag. A*, 52:267, 1985.
- [109] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81:2340, 1977.

- [110] A. F. Voter. Classically exact overlayer dynamics: Diffusion of rhodium clusters on Rh(100). *Phys. Rev. B*, 34:6819, 1986.
- [111] L. Pelaz, M. Jaraiz, G. H. Gilmer, H.-J. Gossmann, C. S. Rafferty, D. J. Eaglesham, and J. M. Poate. B diffusion and clustering in ion implanted Si: The role of B cluster precursors. *Appl. Phys. Lett.*, 70:2285, 1997.
- [112] M. J. Caturla, M. D. Johnson, and T. Diaz de la Rubia. The fraction of substitutional boron in silicon during ion implantation and thermal annealing. *Appl. Phys. Lett.*, 72:2736, 2004.
- [113] S. Solmi and D. Nobili. High concentration diffusivity and clustering of arsenic and phosphorus in silicon. *J. Appl. Phys.*, 83:2484, 1998.
- [114] D. Mathiot and J. C. Pfister. High concentration diffusion of P in Si : a percolation problem? *J. Phys. (Paris) Lett.*, 43:L-453, 1982.
- [115] S. T. Dunham and C. D. Wu. Atomistic models of vacancy-mediated diffusion in silicon. *J. Appl. Phys.*, 78:2362, 1995.
- [116] S. List and H. Ryssel. Atomistic modeling of high-concentration effects of impurity diffusion in silicon. *J. Appl. Phys.*, 83:7595, 1998.
- [117] B. P. Haley and N. Gronbeck-Jensen. Vacancy-assisted arsenic diffusion and time-dependent clustering effects in silicon. *Phys. Rev. B*, 71:195203, 2005.
- [118] M. A. Novotny. Monte Carlo algorithms with absorbing Markov chains: Fast local algorithms for slow dynamics. *Phys. Rev. Lett.*, 74:1, 1995.
- [119] M. A. Novotny. A tutorial on advanced dynamic Monte Carlo methods for systems with discrete state spaces. *Annual Reviews of Computational Physics IX*, page 153, 2001.
- [120] C. S. Deo and D. J. Srolovitz. First passage time Markov chain analysis of rare events for kinetic Monte Carlo: double kink nucleation during dislocation glide. *Modelling Simul. Mater. Sci. Eng.*, 10:581, 2002.
- [121] D. R. Mason, R. E. Rudd, and A. P. Sutton. Stochastic kinetic Monte Carlo algorithms for long-range hamiltonians. *Comput. Phys. Commun.*, 160:140, 2004.
- [122] M. Athènes, P. Bellon, and G. Martin. Identification of novel diffusion cycles in B2 ordered phases by Monte Carlo simulation. *Phil. Mag. A*, 76:565, 1997.
- [123] S. A. Trygubenko and D. J. Wales. Graph transformation method for calculating waiting times in markov chains. *J. Chem. Phys.*, 124:234110, 2006.
- [124] A. La Magna and S. Coffa. Accelerated Monte Carlo algorithms for defect diffusion and clustering. *Comput. Mater. Sci.*, 17:21, 2000.
- [125] C. DeW. Van Sichen. Stochastic method for accommodation of equilibrating basins in kinetic Monte Carlo simulations. *J. Phys.: Condens. Matter*, 19:072201, 2007.
- [126] S. X. Sun. Path summation formulation of the master equation. *Phys. Rev. Lett.*, 96:210602, 2006.
- [127] J. P. DeVita, L. M. Sander, and P. Smereka. Multiscale kinetic Monte Carlo algorithm for simulating epitaxial growth. *Phys. Rev. B*, 72:205421, 2005.
- [128] T. Opplestrup, V. V. Bulatov, G. H. Gilmer, M. H. Kalos, and B. Sadigh. First-passage Monte Carlo algorithm: Diffusion without all the hops. *Phys. Rev. Lett.*, 97:230602, 2006.
- [129] C.-C. Chou and M. L. Falk. Multiscale diffusion Monte Carlo simulation of epitaxial growth. *J. Comput. Phys.*, 217:519, 2006.

- [130] W. Cai, M. H. Kalos, M. de Koning, and V. V. Bulatov. Importance sampling of rare transition events in Markov processes. *Phys. Rev. E*, 66:046703, 2002.
- [131] M. de Koning, W. Cai, B. Sadigh, T. Opplestrup, M. H. Kalos, and V. V. Bulatov. Adaptive importance sampling Monte Carlo simulation of rare transition events. *J. Chem. Phys.*, 122:074103, 2005.
- [132] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu. Rev. Phys. Chem.*, 53:291, 2002.
- [133] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM REVIEW*, 45:3, 2003.
- [134] J. G. Mullen. Isotope effect in intermetallic diffusion. *Phys. Rev.*, 6:1649, 1961.
- [135] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999.
- [136] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, third edition, 2007.
- [137] O. Sugino and A. Oshiyama. Microscopic mechanism of atomic diffusion in Si under pressure. *Phys. Rev. B*, 46:12335, 1992.
- [138] M. Ramamoorthy and S. T. Pantelides. Complex dynamical phenomena in heavily arsenic doped silicon. *Phys. Rev. Lett.*, 76:4753, 1996.
- [139] V. Ranki, K. Saarinen, J. Fage-Pedersen, J. L. Hansen, and A. N. Larsen. Electrical deactivation by vacancy-impurity complexes in highly As-doped Si. *Phys. Rev. B*, 67:041201(R), 2003.
- [140] D. C. Mueller, E. Alonso, and W. Fichtner. Arsenic deactivation in Si: Electronic structure and charge states of vacancy-impurity clusters. *Phys. Rev. B*, 68:045208, 2003.
- [141] MATLAB R2007a, Natick, Massachusetts: The MathWorks Inc., 2007.