DOTTORATO DI RICERCA IN FISICA

IRSLAN ULLAH ASHRAF

MOLECULAR DYNAMICS STUDY OF DEFECT EVOLUTION MECHANISMS IN MATERIAL OF INTERESTS FOR QUANTUM TECHNOLOGIES

TESI DI DOTTORATO

RELATORE:
CHIAR.MO PROF. ELISABETTA PALADINO

CORRELATORE:
CHIAR.MO PROF. ANTONINO LA MAGNA

ANNO ACCADEMICO 2021/2024

Indice

1 Introduction **9**

2 State of the Art **15**

2.1 Experimental Work 15
2.2 Electronic structure of Defect, Spin states 17
2.3 Diffusion Mechanisms and Energetics 19

3 Introduction to the atomistic simulation methods applied in this work **31**

3.1 Ab initio methods 32
3.2 Quantum Many Body problems Born Approximation 33
3.3 Born–Oppenheimer Method as Originally Derived (1927) . . 33
3.3.1 Definition of Coordinates and General Setup 34
3.3.2 Electronic Motion for Stationary Nuclei: Potential Energy Surfaces 35
3.3.3 Expansion of Hamiltonian and Wavefunction 36
3.4 Phase space representation of systems 37
3.4.1 Microcanonical (NVE) Ensemble 38
3.4.2 Canonical (NVT) Ensemble 39
3.4.3 Isothermal Isobaric (NPT) Ensemble 40
3.4.4 Grand Canonical Ensemble 41
3.4.5 Exploration of phase space and Ergodicity 41
3.5 Classical Molecular Dynamics 42
3.5.1 Key Principles 42

3.5.2 Limitations . . . 44
3.5.3 Choice of Potential . . . 45
3.5.4 Periodic boundary conditions . . . 46
3.5.5 Force fields and Force calculation among atoms . . . 47
3.5.6 Tersoff Force Field . . . 47
3.5.7 EDIP force field . . . 49
3.5.8 Integrating Equation of Motion (Dynamics) . . . 50
3.6 Searching Minimum Energy Path (MEP) . . . 52
3.6.1 Introduction . . . 52
3.6.2 Chain of States class of methods and their criticalities 53
3.6.3 Nudging . . . 56
3.7 Stochastic Modeling of Brownian Motion . . . 59
3.7.1 From Discrete Steps to Continuous Diffusion, Einstein's Convolution Derivation . . . 59
3.7.2 Connection to Fick's Law and the continuity equation 62
3.7.3 Mean Square displacement and estimation of Arrhenius type Energy Barrier . . . 63
3.8 Einstein's Derivation and Its Modern Relevance . . . 63
3.9 Statistical Theory of Atomic Transitions for the Study of Defect Evolution . . . 69
3.9.1 Derivation of Fick's Laws with Isotropic Transition Probability . . . 70
3.9.2 Relationship Between Diffusion Coefficient and Mean Square Displacement . . . 73
3.9.3 Frequency rule for Diffusion Coefficient . . . 75

4 Configurations and Energetics of Point Defects in 3C-SiC with Optimization Methods 79
4.1 Computational Methodology and benchmarking . . . 80
4.2 Elastic Properties of Ideal and Defective Configurations . . . 81
4.3 Formation Energies of Point Defects . . . 83
4.4 Study of melting point for 3C-SiC and validation of interatomic potentials . . . 86
4.5 Migration Energy Barriers study using NEB . . . 86
4.6 Summary . . . 94

5 Simulation of Non Equilibrium Processes and Defect Generation 97
5.1 Introduction 97
5.2 Simulation of Non Equilibrium Processes and Defect Generation Using BCA 97
5.3 BCA simulation of ion implantation in Wide Band Gap semiconductors 99
5.3.1 Limitations 101
5.4 Simulation of Non-Equilibrium Processes and Defect Generation using Molecular Dynamics 102
5.4.1 Defect Evolution During the Ballistic Cascade(initial Picosecond scale damage) 102
5.4.2 Post ballistic defect evolution 103
5.5 Molecular Dynamics Simulation Results 106
5.5.1 Simulation Setup 106
5.5.2 Ballistic Regime 108
5.5.3 Energy Dependent Defect Accumulation 109
5.5.4 Temprature dependent Annealing Study 109
5.5.5 Effect of Annealing Duration on Structural Recovery at 1500 K 111
5.6 Conclusions 113

6 Diffusion of point defects in 3C-SiC 117
6.1 Introduction 117
6.2 Diffusion Mechanisms of Point Defects in SiC 118
6.3 Methodology for diffusion studies 124
6.3.1 Exploring Diffusion Coefficient through Mean Square Displacement (MSD) 124
6.3.2 Defect evolution study using frequency rule method (implying COM) 126
6.4 2000 K Case Study: Carbon Vacancy Diffusion 134
6.4.1 Setting up molecular dynamics 134
6.4.2 Processing of mean square displacement (MSD) data 135
6.4.3 Processing Center of Mass (COM) data 135
6.4.4 Fitting of jump time data with Poisson distribution . 138

6.4.5 Fitting of jump time data with waiting time/exponential distribution for V_C . . . 138
6.4.6 Energy Barrier estimation . . . 140
6.5 1200 K Case Study: Carbon Interstitial Diffusion . . . 140
6.5.1 Setting up Molecular dynamics simulation . . . 140
6.5.2 Processing of mean square displacement (MSD) data 141
6.5.3 Processing of center of mass (COM) data . . . 141
6.5.4 Fitting of jump-time data with Poisson distribution . 144
6.5.5 Fitting of jump time data with waiting time/exponential distribution for C_I . . . 144
6.6 MSD Results . . . 146
6.7 Diffusion Energetic . . . 148
6.8 Comparison of Methods, Einstein Diffusion formula (MSD) and frequency rule method (COM) . . . 153
6.8.1 Factors Affecting Diffusivity Estimates . . . 153
6.8.2 Frequency Rule Method Potential Issues . . . 157
6.8.3 MSD Method Potential Issues . . . 157
6.8.4 Evaluation of Diffusivity Estimation Methods . . . 160
6.9 Stabilization of defects systems, annihilation and complex formations . . . 163
6.9.1 Generation of Divacancy . . . 163
6.9.2 Divacancy (VV) Formation . . . 163
6.9.3 Annihilation of Interstitial Carbon and Carbon Vacancy . . . 167

Bibliografia **174**

Appendices **187**

A Parameterization for Force fields **187**

B Tersoff Validity test by Comparison of results with DFT **189**

C Nudged Elastic Band Results **191**
C.1 Spring Constant Optimization for NEB Calculations . . . 191
C.2 Carbon Antisite Vacancy . . . 194
C.3 Silicon Antisite Vacancy . . . 195

D Gaussian Integral Results 197

E StepShiftDetect : Quantifying Jumps in Diffusive Motion via Change Point Analysis 199

F Sampling Error Analysis of MSD Derived Diffusivity 203
F.1 Introduction 203
F.1.1 Simulation Protocol 204
F.1.2 Statistical average 204
F.1.3 Results and Figures 206
F.1.4 Summary Table 212
F.2 Generating 2nn random walk in 3C-SiC and measuring `D_i` over multiple `dt_mean` 213

Ringraziamenti 222

1. NEB study of Carbon Antisite vacancy

2. NEB study of Silicon Antisite Vcancy

Acronym	**Description**
DFT	Density Functional Theory
MD	Molecular Dynamics
FPMD	First Principle Molecular Dynamics
NEB	Nudged Elastic Band Method
MSD	Mean Square Displacement
D	Diffusion Coefficient
E_A	Activation Barrier
PMF	Potential of Mean Force
US	Umbrella Sampling Method
ABF	Adaptive Biasing Force
BCA	Binary Collision Approximation
MC	Monte Carlo
OVPD	Open Volume Point Defects
SIKMC	Super Lattice Kinetic Monte Carlo
ZBL	Ziegler Biersack Littmark
NRT	Norgett Robinson Torrens
COM	Center of Mass
VV	Di Vacancy
V_C ,V_{Si}	Carbon Vacancy, Silicon Vacancy
I_C , I_{Si}	Interstitial Carbon, Interstitial Silicon
KMC	Kinetic Monte Carlo
CAV	Carbon Antisite Vacancy ($V_C C_{Si}$)
TTM	Two Temperature Model
k-ART	Kinetic Atomic Relaxation Technique
PKA	Primary Knock-out Atoms
DFE	Defect Formation Energy

Tabella 1: List of Important Acronyms and their Descriptions

Capitolo 1

Introduction

Understanding defects in wide band gap semiconductors has become increasingly important for applications in quantum technologies and high-power devices. Silicon carbide (SiC) has emerged as a leading third-generation semiconductor and a promising host material for solid-state spin qubits. Its well-studied polytypes, such as 4H, 6H, and 3C, support color centers like silicon vacancies and divacancies that exhibit long coherence times and room-temperature spin control, making them attractive for quantum technologies. However, controlling the formation, stability, and dynamics of such defects remains a major challenge, as their formation mechanisms are still incompletely understood. Investigating how these defects are created, migrate, and interact under different conditions is therefore of both fundamental and applied significance, and has attracted considerable attention from researchers worldwide. Recent experimental and theoretical studies have demonstrated the existence of optically addressable spin defects in SiC, such as the divacancy and the carbon antisite–vacancy (CAV) complex, which share similarities with the well-studied nitrogen-vacancy center in diamond. Considerable progress has been made in characterizing their optical and spin coherence properties, and their formation, stabilization, and reorientation have also been reported [1]. In addition, electronic structure calculations based on density functional theory, as well as enhanced sampling methods coupled with first-principles molecular dynamics, have been employed to investigate these defects [2]. Parallel to this, extensive work has focused on un-

derstanding defect production in SiC under ion implantation and irradiation, given its role in power electronics and radiation hard devices[3, 4, 5]. However, while experiments provide valuable insights, they are often limited by spatial resolution and the difficulty of probing defect dynamics on ultra fast timescales. In this thesis, I employ atomistic simulations, in particular molecular dynamics (MD), to investigate the formation and evolution mechanisms of defects in SiC. The work is focused on three related aspects:

1. Formation key defects in silicon carbide and respective energies,
2. the damage production and relaxation mechanisms during ion implantation, and
3. the diffusivity of key defect species and related energetics, calculated using both the mean square displacement (MSD) method and the Einstein–Brownian relation.

These studies aim to provide a theoretical complement to experimental observations and to clarify atomistic mechanisms that are challenging to access experimentally. Overall our results contribute to understanding both the microscopic origin of color centers and the broader picture of defect dynamics in irradiated SiC.The thesis is organized as follows. Chapter 2 reviews relevant experimental and computational studies on spin defects in SiC. Chapter 3 introduces the theoretical background and simulation methods employed in this work. Chapter 4 presents the molecular dynamics results on defect formation and stability, focusing mainly on the static aspects of the study (excluding diffusion), including the calculation of formation energies and migration barriers. Chapter 5 reports the simulations of ion implantation and examines the mechanisms of damage evolution. Chapter 6 discusses the diffusivity of key defect species using two statistical approaches, evaluates their validity, and proposes the most suitable method for defect studies. An Appendix is provided to support and complement the discussions presented in the thesis.

Although the simulations presented in this thesis are primarily based on classical physics, the behavior of atoms and defects in solids cannot be fully understood without reference to quantum mechanics. In Chapter 3, I introduce the theoretical framework by starting with a brief quantum

mechanical description of lattice vibrations and defect energetics, and then show how classical mechanics can be derived as a limiting case. On this basis, I describe in detail the classical molecular dynamics methods employed to model defect formation, migration, and evolution in SiC.

It is now well established that several point defects in SiC are both spin-active and optically addressable. Examples include silicon vacancies (V_{Si}), nitrogen-vacancy (NV) related centers, antisite-vacancy complexes (e.g. $C_{\mathrm{Si}}V_{\mathrm{C}}$), and divacancies ($V_{\mathrm{C}}V_{\mathrm{Si}}$, often written as VV). Among these, divacancies are particularly promising because they combine optical spin initialization and readout with long coherence times and near infrared spin photon interfaces [6, 7, 8, 9]. These properties have motivated intensive experimental and theoretical work exploring their electronic structure and spin dynamics [10, 11]. However, the atomistic mechanisms that control the selective formation, stabilization, and reorientation of these centers during processing and irradiation remain incompletely understood a gap that this thesis addresses by studying defect formation and diffusion in SiC using atomistic simulations. Simulations provide a complementary route to experiments for probing defect formation and dynamics in SiC, because they can access atomistic processes that are difficult to measure directly. We combine classical molecular dynamics with targeted DFT benchmarking to probe atomistic defect formation and diffusion in SiC, validating MD derived energetics and kinetics against first-principles calculations and experiments where possible.

Despite extensive research on the properties of color centers, much less is understood about their formation and evolution processes. Spin defects, including divacancies, are typically generated through ion implantation, irradiation (e.g. pulsed laser), and subsequent annealing. Annealing plays a crucial role in initializing defect mobility, which is key for VV formation. Several methods, such as electron paramagnetic resonance (EPR), photoluminescence (PL), and deep level transient spectroscopy, have been employed to monitor the formation and spatial localization [12] of these defects [13, 14, 10, 15, 16, 17]. Recent advancements have reported progress in controlling the charge state [18] and production of spin-bearing defect complexes with ion implantation and thermal annealing [19]. However, the conversion-yield of mono vacancies to VV remains low, and spatial localization of VV during annealing remains a challenge.

Moreover, the entire process (placement and formation) is strongly correlated with the initial conditions such as lattice temperature, pre-existing defect concentrations, and the presence of impurity atoms, particularly the spatial distribution of defects. Such challenges necessitates the use of predictive simulation methods, which provide theoretical framework to reproduce experimental environment under controlled conditions. One optimize the computation to access very small time and length scales, thus come up with findings that act as guiding tool priory to experiments.

The evolution of defect complexes presents significant challenges for computational methods. Elevated activation energies, often reaching several electron volts, pose considerable difficulties for *ab initio* methods such as First-Principles Molecular Dynamics (FPMD), thus limiting their applicability. The reason is linked with high activation barrier correspond to rare events in phase space. Simulating such events or better accessing them require intense simulation for extremely long times. Since FPMD is computationally expensive thus limited to the picoseconds timescales and relatively small system sizes it becomes impractical to capture these rare phenomenons directly.

As a result, extensive theoretical investigations have concentrated on the energetics of point defects and the calculation of migration barriers thus skipping the crucial dynamical part. Notable work by Galli[1] has addressed the dynamics of vacancies, offering valuable insights by calculating the kinetics of divacancy formation, divacancy reorientation, and associated migration barriers. Their research was further expanded through the application of ab initio methods to explore the electronic structure of VV and predict the optimal annealing temperature [2]. A lot of work has also been performed in simulating ion implantation using H ion, He ion, Dual ion [20, 21, 22, 4] and subsequent annealing by implementing molecular dynamics [20, 23]. Despite a massive body of research on spin defects their understanding at atomistic level is not completely understood. Especially, dynamics of VV, condition for VV stabilization in the lattice and intermediate process involved.

Our contribution to this ongoing research represents a modest yet meaningful advancement, focusing specifically on the detailed investigation of divacancy (VV) formation and the associated mechanisms. We aim to study the diffusion behavior of point defects using both the mean

square displacement (MSD) method and the frequency rule approach, in order to establish the most suitable methodology for future investigations.

These studies are conducted in a real time atomistic simulation environment, implemented using molecular dynamics (MD) techniques, and further refined through statistical post-processing for desired calculation and trajectory analysis.

Our research begins by replicating and building upon existing studies, then extends these efforts to explore the complex dynamics of critical defect species. This involved scrutinizing force field potentials, reproducing previously established results(Lattice information, Defect formation energies, melting point study). A key objective of this work is to identify and understand the factors that impede divacancy formation as well as the conditions that could promote higher mono vacancy to divacancy conversion yields. Specifically, defect mobility controlled by the diffusivity of vacancies and interstitial is critical factor to induce defect recombination and spin-complex formation. We conduct systematic atomistic simulations to study the diffusivity of these defects and how they participate to enhance or suppress VV formation.

To quantify defect diffusivity, we employed two complementary approaches: the mean square displacement method (D_{MSD}) and the frequency based method ($D_{\text{frequency}}$). While both methods capture the same underlying diffusion processes, discrepancies were observed between their estimates. To investigate this, we simulated a second nearest neighbor random walk in cubic silicon carbide (3C-SiC). Our analysis revealed that ensemble averaged diffusivity values, obtained by averaging over many independent trajectories, consistently converge to $D_{\text{frequency}}$. In contrast, diffusivity extracted from a single trajectory often deviates significantly, reflecting statistical fluctuations and limited sampling.

By accurately calculating the diffusion coefficients of major defects, we have investigated essential processes involved in ion implantation and irradiation that potentially contribute to VV formation. We have verified the diffusion of V_C towards V_{Si} leading to the formation of a stable VV complex, as well as the potential recombination of I_C and V_C, which results in an energy gain of approximately 5 eV. In the latter part of our study, we extended our investigation to the simulation of ion implantation in 3C-SiC, focusing on the real time observation of defect generation

under non equilibrium conditions. This phase of the research provided valuable insight into the dynamic damage mechanisms initiated by ion collisions and subsequent annealing/relaxation processes. The use of direct molecular dynamics simulation allowed us to monitor the evolution of divacancy (VV) formation with enhanced temporal and spatial resolution, enabling precise control over initial implantation conditions and post-implantation thermal treatments.

This methodology provided a more detailed characterization of VV formation mechanisms. In particular, it enabled us to resolve spatial localization during annealing by tracking defect migration on a jump-to-jump basis using the frequency rule method. Furthermore, by combining mean-square displacement (MSD) analysis with potential energy evolution, we could directly correlate decreases in system energy with specific MSD signatures, thereby identifying the onset of VV formation and distinguishing it from other metastable trapping events.By adopting a continuous viscous damping strategy for cooling the system, instead of using phase-separated annealing stages, we obtained a more gradual and consistent removal of excess kinetic energy. Compared to stepwise annealing, this approach reduced artificial thermal fluctuations and avoided abrupt changes in temperature, thereby giving a closer approximation to realistic thermal relaxation. These improvements facilitate a clean strategy and great connection between implantation energy, defect clustering, and annealing efficiency.

Based on the efficiency of statistical sampling, we generally fixed the simulation time to 100 ns for most diffusivity calculations. However, this time scale was extended up to 600 ns when necessary to reach the required number of jump events (at least 70). These measures were adopted to enhance the reliability of the diffusivity estimates and the robustness of defect yield predictions. Our findings provide a predictive framework that may serve experimentalists in optimizing implantation protocols and annealing schedules to control defect formation in SiC and similar materials. The overall objective was to move beyond static defect energetics and gain atomistic insight into intermediate kinetic processes that govern defect mobility, recombination, and divacancy stabilization under realistic non equilibrium conditions.

Capitolo 2

State of the Art

In the introduction chapter, we highlight the importance of spin defects in emerging technologies. The extensive potential applications of these defects have garnered the attention of numerous researchers worldwide, leading to substantial contributions from both theoretical and experimental studies. Here we will review the literature to establish the basis of this work and also provide the theoretical support.

2.1 Experimental Work

Quantum information is physical [24]. Solid-state defect spins are promising candidates for realizing the fundamental unit of quantum information the quantum bit [10]. Among the most studied systems are the nitrogen-vacancy (NV) center in diamond [25],vacancy center in silicon carbide [7] single rare earth ion qubit [26]and the phosphorus (P) donor in silicon [27]. These defects represent opposite behavior in the electronic structure spectrum, while the P-donor in silicon can be modeled as a hydrogenic impurity within the effective mass approximation [28], the NV center in diamond features a strongly localized deep-level state within the bandgap [25]. In addition to these, color centers in silicon [29] and defect complexes in gallium nitride (GaN) [30] are also being explored for their potential in quantum photonics and optoelectronic applications. Several techniques have been developed to engineer spin defects in solid-

state systems, each offering different levels of precision, scalability, and material compatibility. Ion implantation is one of the most widely used methods, allowing controlled introduction of impurity atoms such as nitrogen into diamond [31, 32]. This technique is often followed by thermal annealing to mobilize vacancies and facilitate defect complex formation and widely been used to engineer point defects such as nitrogen-vacancy (NV) centers in diamond [25], color centers in silicon [29], and defect complexes in gallium nitride (GaN) [30], which are critical for quantum technologies and optoelectronic devices. Alternatively, spin defects can be incorporated during crystal growth using methods such as chemical vapor deposition (CVD), which enables the creation of high purity diamond with in-situ nitrogen doping for NV center formation [33, 34]. Focused ion beam (FIB) and laser writing techniques have also emerged as promising methods for spatially precise defect generation [35, 36]. Each fabrication method influences the defect's coherence properties, charge state stability, and optical characteristics parameters that are critical for their use in quantum technologies.

Characterizing spin defects is essential to understanding their electronic structure, coherence properties, and suitability for quantum applications. Optically detected magnetic resonance (ODMR) is one of the most powerful techniques, enabling the detection of spin transitions through changes in fluorescence intensity under microwave excitation [37, 38]. Photoluminescence (PL) spectroscopy is widely used to probe the optical properties of defects, revealing zero-phonon lines and vibronic sidebands that help identify specific defect types [39]. Electron paramagnetic resonance (EPR), including pulsed and continuous-wave modes, provides direct information on the spin Hamiltonian, hyperfine interactions, and local symmetry of the defect [40]. More advanced methods like spin echo and Hahn echo are employed to extract coherence times (T_2) and spin relaxation times (T_1) [41]. Additionally, techniques such as scanning probe microscopy and confocal fluorescence microscopy offer spatial resolution and enable single defect addressability [42]. Together, these tools form a comprehensive toolkit for evaluating spin defects in solid-state platforms.Pavunny employed a mask less focused ion beam technique using Li^+ ions. The fabricated defects were characterized using scanning con-

focal fluorescence microscopy, with a focus on sharp zero-phonon lines (ZPLs) [32]. Several other techniques have also been utilized to investigate defect formation and spatial localization. For instance, Gray Wolfowicz measured photoluminescence under near ultraviolet excitation and developed persistent spatial patterning of defect populations [43]. Similarly, R. Karsthof used proton irradiation for defect generation and performed deep level transient spectroscopy (DLTS) coupled with photoluminescence spectroscopy to characterize the charge states of the defects [15]. David M. Toyli utilized secondary ion mass spectrometry (SIMS) to study the depth profiling of implanted ions, providing detailed three-dimensional spatial profiles of the defects under investigation [12].

2.2 Electronic structure of Defect, Spin states

Color centers are typically intrinsic or extrinsic point defects in wide bandgap semiconductors whose electronic states are optically addressable and spin active [44, 11]. These defects are found in materials such as diamond, silicon carbide (SiC), and gallium nitride (GaN), and can include vacancies, interstitials, and complex configurations like divacancies or carbon antisite vacancy pairs [45, 46]. The Bloch theorem describes how electronic wavefunctions behave in a perfectly periodic crystal lattice, leading to the formation of continuous energy bands [47, 48]. It explains that in such periodic systems, electrons adopt delocalized wavefunctions known as Bloch states, which extend throughout the crystal and obey the translational symmetry of the lattice. However, the introduction of defects breaks this translational symmetry and disrupts the periodic potential. As a result, localized electronic states can form within the band gap states that no longer satisfy the Bloch solutions[49]. When both the ground and excited states of a defect lie inside the band gap, the system behaves like an isolated two-level quantum system, which can be coherently addressed using sub-bandgap optical or microwave frequencies. The emission spectrum may include a zero phonon line (ZPL), characterized by a narrow Lorentzian peak, which occurs without phonon contributions. Additionally, a phonon sideband may arise from transitions to other vibronic levels of the ground state. Color centers are capable of emitting single

photons even at cryogenic temperatures, making them excellent candidates for photonic applications. Understanding spin-dependent energy levels, i.e., how electronic states split due to spin-orbit coupling or magnetic interactions is essential for interpreting the fine structure observed in optical spectra. Equal important is understanding the symmetry properties of these defect states, as the persistence or breaking of specific symmetries (such as inversion or mirror symmetry) governs the selection rules and transition intensities in photoluminescence and absorption measurements. Hybrid Density Functional Theory (DFT) provides an excellent framework for studying the spin states of color centers. For example, carbon-vacancy, carbon antisite, vacancy complexes, such as divacancies (VV), $V_C C_{Si} V_C$, and $[V_C C_{Si} V_C]_N$, have been studied using hybrid DFT [50, 1], and their structures are presented in Figure 2.1. The $[V_C C_{Si} V_C]_N$ complex represents a configuration in which two carbon vacancies are separated by $N = 3$ lattice sites in silicon carbide (SiC), with a silicon atom in between. These complexes introduce localized electronic states within the band gap, primarily originating from the dangling bonds of the atoms adjacent to the vacancies. When two carbon atoms are removed, each contributes four valence electrons, and the surrounding atoms adjust their bonding, leaving behind a total of six unpaired electrons associated with the defect. The interaction between these electrons leads to a spin configuration, and due to exchange interactions and defect symmetry, the ground state preferentially stabilizes as a spin triplet ($S = 1$). This behavior is analogous to the nitrogen-vacancy (NV) center in diamond, where unpaired electrons from the nitrogen and the adjacent vacancy similarly form a stable triplet ground state [25]. Likewise, divacancies in SiC are also known to exhibit triplet ground states due to unpaired electrons from adjacent dangling bonds [1].

From the figure 2.1, one can observe that the ground-state spin density of $V_C C_{Si} V_C$ is more delocalized than that of VV. A spatial delocalization is strongly influenced by the fact that the two vacancies are separated by an antisite. Another similarity between VV and $V_C C_{Si} V_C$ is the presence of unoccupied defect states near the conduction band [51], suggesting that $V_C C_{Si} V_C$ is also a spin defect similar to VV. These defects are generally created during ion implantation, but annealing at optimal temperatures facilitates the conversion of $V_C C_{Si} V_C$ to VV, leading to an eventual

dominant VV concentration [52, 53].

2.3 Diffusion Mechanisms and Energetics

Numerous researchers have explored the implementation of theoretical methods, primarily First Principles calculations within the framework of Density Functional Theory (DFT), to study types of defects in silicon carbide (SiC) poly types and to calculate corresponding formation energies. Additionally, Molecular Dynamics (MD) and Kinetic Monte Carlo (KMC) methods have been employed to investigate diffusion pathways and to calculate migration barriers.

M. Bockstedte et. al. [54] conducted a theoretical investigation of defect energetics in the 4H-SiC polytype using DFT. They examined a range of microscopic defect structures, including V_C, V_{Si}, $V_C C_{Si}$, I_C, and I_{Si} in 4H-SiC. A detailed description of the energetics as a function of the Fermi level is presented in Figure 2.2.They also considered the defect migration mechanisms of interstitials and vacancies, examining how these mechanisms are influenced by doping conditions. For I_{Si}, they reported two key configurations and corresponding migration mechanisms: (1) kick-out migration via split-interstitials and (2) migration via Tetrahedral Si interstitial. These mechanisms are illustrated in Figure 6.2.

M. Bockstedte et. al. [55] also simulated the annealing mechanisms in 4H-SiC and established an associated hierarchy. The principal processes are as follows:

1. Recombination of vacancies and interstitials: (a) Recombination of carbon split interstitial with a carbon vacancy, including first neighbor and second neighbor interactions. (b) Recombination of silicon vacancy with silicon split interstitial or Frenkel pair recombination (limited to n-type conditions).

2. Diffusion to the surface.

3. Diffusion to a sink: Formation of stable complexes, e.g., transformation of V_{Si} into the $V_C C_{Si}$ complex.

A similar ab initio based study was conducted by Xiaopeng Wang et al. to describe the equilibrium concentration and annealing mechanisms based on silicon vacancy diffusion and relevant migration pathways.

They also investigated the effects of charge states, annealing temperature, and crystal orientation [56]. Rurali further explored the migration pathways for V_C and V_{Si}, suggesting that although V_{Si} has a lower migration barrier than V_C, its migration efficiency is reduced by parallel indirect diffusion pathways, making V_C diffusion the most direct event [57]."

R. K. Defo et al. [58] combined Molecular Dynamics (MD), Kinetic Monte Carlo (KMC), and Density Functional Theory (DFT) simulations to investigate the thermodynamics and kinetics of the V_{Si} color center in 4H-SiC. They calculated formation energies for different charge states and subsequently simulated their creation through ion implantation using MD. The annealing process at high temperatures was simulated using KMC to study the diffusion of the produced defects, their conversion into carbon antisite-vacancies (CAV), and recombination. The calculated diffusion barriers are presented in Figures 2.3 and 2.4. Their study showed excellent correspondence between the produced defects and the relevant formation energies, highlighting a monotonic increase in the diffusion barrier with increasing charge state.

Recently, Giulia Galli et al. [1] conducted an outstanding study on spin defects and defect complexes by coupling First-Principles Molecular Dynamics (FPMD) with a neural network-based enhanced sampling method, commonly known as the umbrella sampling method [59, 60]. This approach allows for efficient exploration of the phase space and precise calculation of the free energy landscape, which in turn helps to determine the migration barrier and the stability of defects and defect complexes. The calculated migration barriers are presented in Figure 2.5.

They also performed classical MD simulations to model larger system sizes and longer time scales. This enabled them to calculate defect densities at various temperatures and assess the effect of annealing temperature on defect stabilization, with a particular focus on the formation and stabilization of VV. A summary of defect energetics, including classical MD free energy and enhanced FPMD for different defect structures, is provided in Table 2.1.

Based on previously implemented Enhanced sampling FPMD method and DFT method Galli et.al. (2023)[2] further extended her work to delve in the thermodynamics of defect and defect complex formation processes. They considered number of possible pathways for VV formation that

Process	**Initial state**	**Final state**	**Classical MD free energy barrier, forward (eV)**	**FPMD free energy difference (final – initial state), (eV)**
V_C migration	V_C	$V_C{}'$	2.5	0
V_{Si} migration	V_{Si}	$V_{Si}{}'$	3.5	3.9
Vacancy-antisite conversion	VSi	CSiVC	1.3	-0.6
VV dissociation via V_C	VV	V_{Si} + V_C	2.2	1.3
VV dissociation via V_{Si} migration	VV	V_{Si} + VC	-	3.0
VV reorientation via V_C migration	VV	VV′	1.7	0
VV reorientation via V_{Si} migration	VV	VV′	-	3.1

Tabella 2.1: Comparison of various energy barriers and free energy differences in MD and FPMD simulations. [1]

are presented in figure 2.6. Expanding their study across different temperatures, they computed defect formation energies (E_f), effective energy barriers (E_b, E_b^{EFF}), Gibbs free energy barriers (G_b), entropy differences (∇S), and activation temperatures (T_a). Additionally, they employed the DFT-based nudged elastic band (NEB) method to calculate internal energy changes, utilizing the DDH functional to obtain E_b^{DDH}. Moreover, they calculated the effective barrier E_B^{EFF} for a diffusion event, instead of the simple barrier E_b. Considering shifts in charge or spin states at elevated temperatures, the effective barrier calculation takes into account the charge (q) and spin (s) states of defects during a transformation [61, 62]. For the dissociation of CAV, since it involves multiple steps, E_B^{EFF} is given by the following equation [2]:

$$E_B^{\mathrm{EFF}}(\mathrm{CAV\ migration}) = \Delta E_f + \max\{E_{b,\mathrm{EFF}}(V_{Si}\ \mathrm{migration}), \\ E_{b,\mathrm{EFF}}(V_{Si} \rightarrow \mathrm{CAV})\} \quad (2.1)$$

While $\nabla E_f = E_f(V_{Si}) - E_f(\mathrm{CAV})$. Moreover, if $\nabla E_f < 0$ eV, then V_{Si} is more stable. In that case, equation 2.1 is revised as follows:

$$E_B^{\mathrm{EFF}}(\mathrm{CAV\ migration}) = \max\{E_{b,\mathrm{EFF}}(V_{Si}\ \mathrm{migration}), E_{b,\mathrm{EFF}}(V_{Si})\} \quad (2.2)$$

A significant amount of work has also been conducted to simulate the implantation process and study defect damage. Antoine Jay et al. [63] performed an in depth statistical analysis of the implantation cascade by bombarding a Si atom into bulk silicon. They began by devising a comprehensive approach for studying damage creation and evolution. By comparing their results with neutron damage and repeating the experiment at different energies, they identified employing 1-10 keV **Si** atoms as the optimal implantation process .

They further extended their work [64] and investigated the displacement damage produced by Primary Knock-on Atoms (PKA). This was achieved by coupling Molecular Dynamics (MD) with a two temperature model (TTM), followed by the Kinetic Atomic Relaxation Technique (k-ART). The TTM model accounted for electron stopping power, while k-ART permitted studying relaxation of the damaged configuration

over extended timescales (from nanoseconds in MD to in kinetic monte Carlo), thus simulating real time experimental conditions. As a result, they identified possible stable configurations, including divacancies and tri-interstitials. Identified defect geometries were further studied using first-principles methods [65].

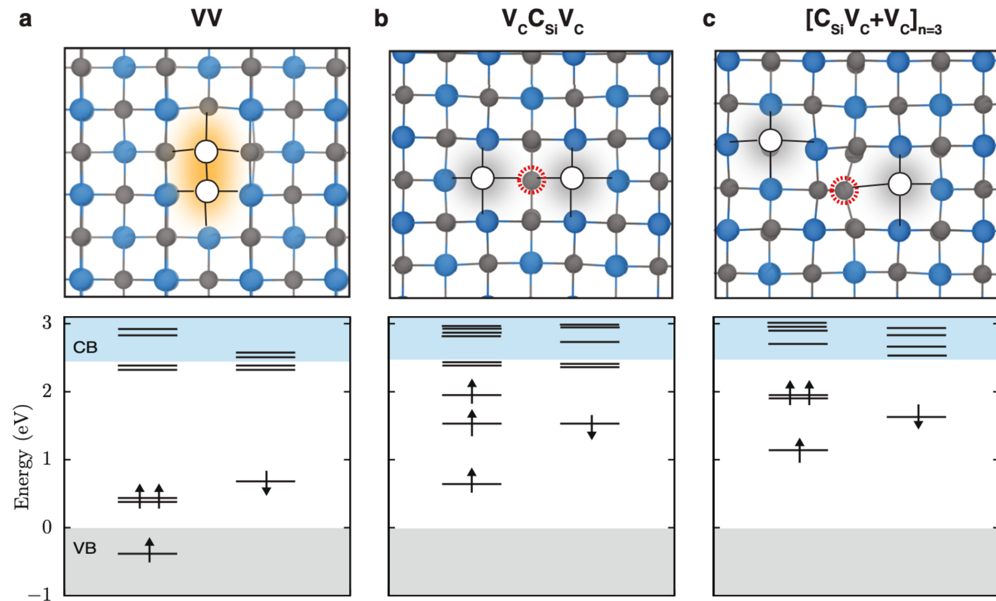


Figura 2.1: Electronic structures of the neutral divacancy (VV^0) and antisite-vacancy complexes, as identified from MD simulations of VV dissociation, are presented. The defect structures (top) and corresponding defect energy-level diagrams (bottom) are shown for (a) VV, (b) $V_C C_{Si} V_C$, and (c) $[C_{Si}V_C + V_C]_{n=3}$. For the $[C_{Si}V_C + V_C]$ defect, the two carbon vacancies are separated by at least $n = 3$ atoms. The lowest electronic configuration for each defect is a triplet state. Shaded gray areas indicate energy levels below the valence band (VB) and above the conduction band (CB). The spin-majority (spin-minority) channel is denoted by upward- (downward-) pointing arrows.(figure obtained from [1])

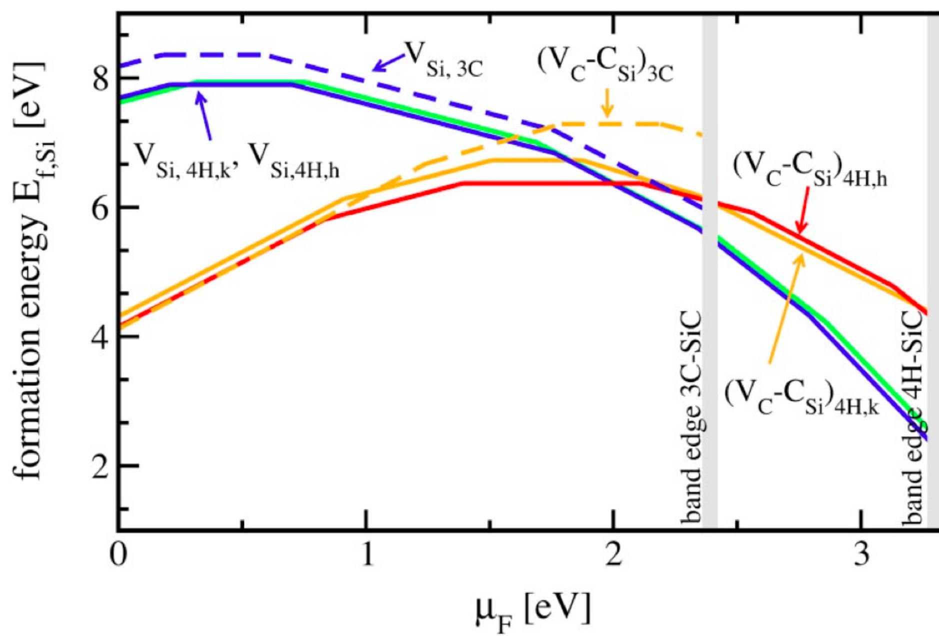


Figura 2.2: Formation energy of the V_{Si} and $V_C C_{Si}$ for 3C-SiC and 4H -SiC. The experimental band edges of 3C- and 4H-SiC are indicated by vertical bars(figure obtained from [54]).

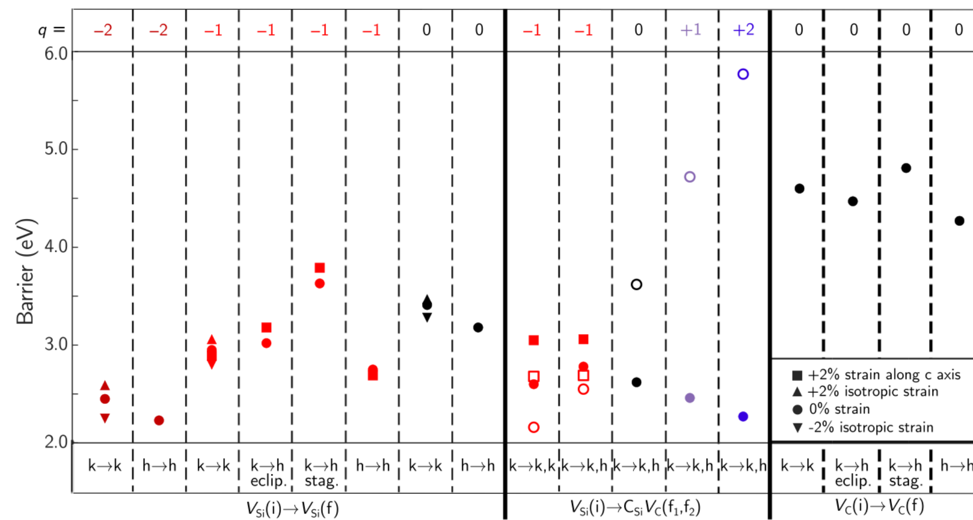


Figura 2.3: The energy barriers for selected transitions between defect states. The term "stag." refers to staggered, and "eclip." refers to eclipsed configurations, as depicted in Fig. 3. The transition is considered in-plane and eclipsed when neither eclipsed nor staggered is specified for a transition between the same defect species. The lattice sites for the initial and final states in the transition are identified as either hexagonal (h) or cubic (k). The forward (reverse) barriers are represented by solid (open) symbols. Reverse barriers are omitted for differences less than 5 percent. Black indicates neutral charge, purple and red signify ±1 charge, and blue and maroon represent ±2 charge.

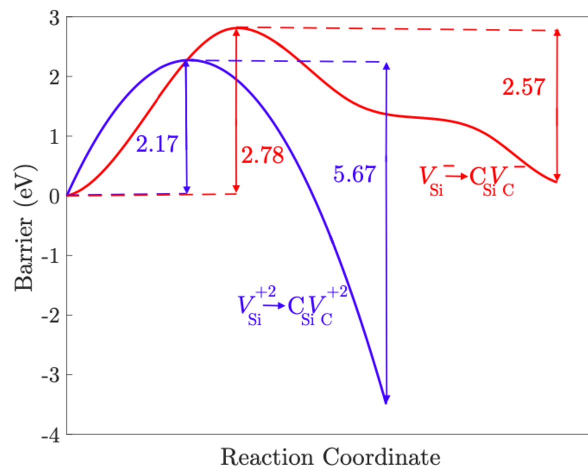


Figura 2.4: The barriers for the conversion $V_{Si} \rightarrow C_{Si}V_C$ for the charge states -1 (red) and $+2$ (blue)

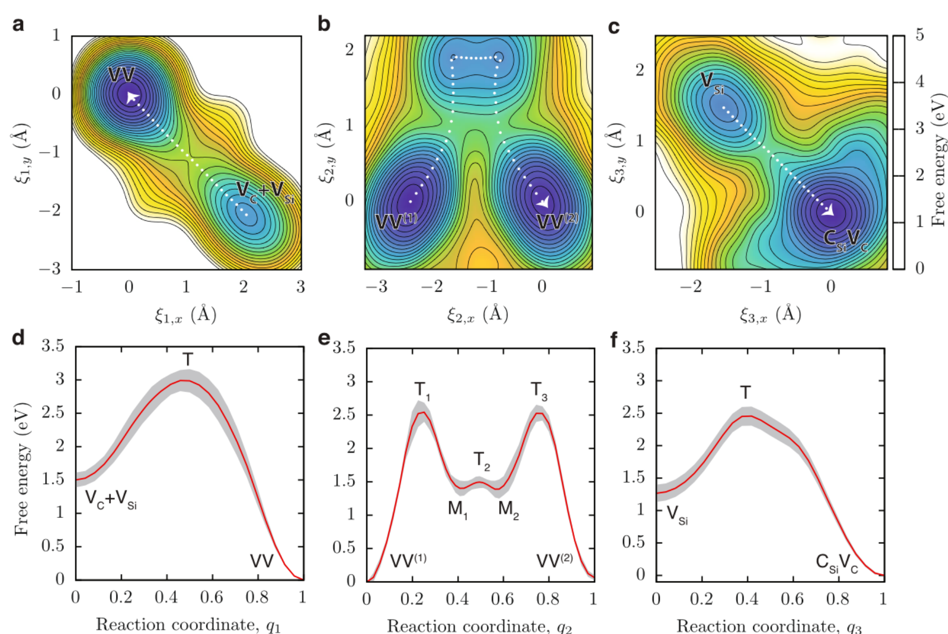


Figura 2.5: Free-energy landscapes of thermally activated (> 1.3 eV)vacancy conversion processes for VV and V_{Si} Potentials of mean force (PMFs) were calculated using enhanced sampling simulations with FP-MD. carbon atom migrating toward a vacancy site (a, d), VV reorientation (b, e), and V_{Si} to $C_{Si}V_C$ conversion processes (c, f). Panels a–c display 2D-PMFs showing free-energy surfaces of vacancy conversion processes at 1500 K, with the minimum free-energy pathways indicated by white dotted lines. Panels d–f show free-energy profiles revealing intermediate states (M_i) and transition states (T_i) along the reaction coordinates. The gray shaded regions denote the error in the PMF determined by block averaging.**Figure and caption obtained from [1]**

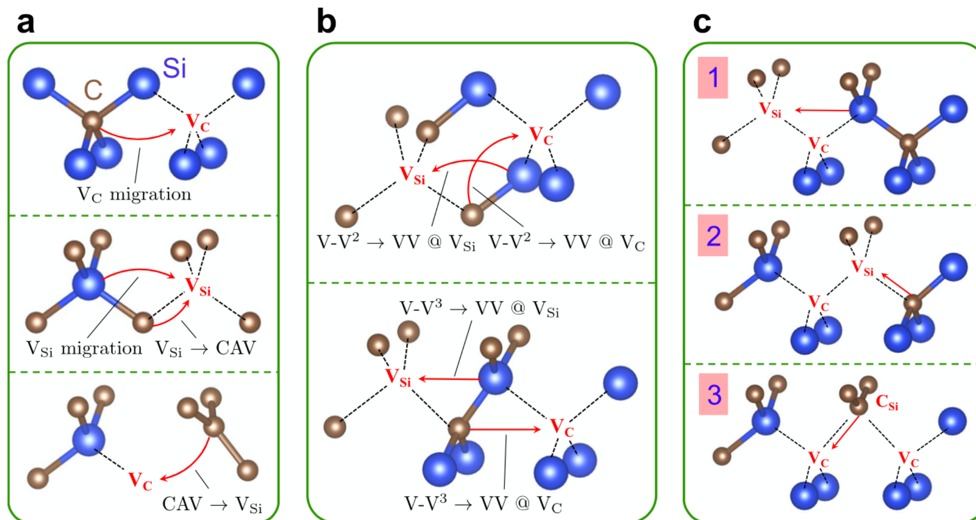


Figura 2.6: Investigated atomic pathways in 3C-SiC. **a** Mono-vacancy dynamics, including carbon (V_C) and silicon (V_{Si}) vacancy migration, as well as the inter-conversion between V_{Si} and carbon antisite vacancy complex (CAV). **b** Pairing of second (V-V_2) and third (V-V_3) neighbors V_C and V_{Si} vacancies to form a divacancy (VV). Only V-V pairs up to third neighbors were considered due to the size limitations of our super-cells. **c** VV migration path with the lowest barrier, where steps 1 to 3 are illustrated. The $V_C C_{Si} V_C$ complex in step 3 is denoted as $V_C V$. **Figure and caption obtained form [2]**

.

Capitolo 3

Introduction to the atomistic simulation methods applied in this work

In this chapter, we present a comprehensive overview of the theoretical and computational methods employed in this study to investigate atomistic processes. We begin with an introduction to *ab initio* methods and the treatment of quantum many body problems, followed by the Born–Oppenheimer approximation in its original formulation and subsequent refinements. The discussion then progresses to phase space representations of physical systems, including the statistical ensembles commonly used in molecular simulations. We further describe the principles of classical molecular dynamics, with emphasis on interatomic potentials, force fields, and integration schemes. Particular attention is given to the Tersoff and EDIP potentials, which play a central role in modeling covalent systems. Methods for exploring potential energy surfaces and searching minimum energy pathways, such as the nudged elastic band approach, are also introduced. In addition, we examine stochastic models of Brownian motion, Einstein's derivations of diffusion, and their modern relevance to transport phenomena. Finally, we address statistical theories of atomic transitions and the simulation of non-equilibrium processes relevant to defect generation and evolution. Together, these sections establish the methodological foundation for the atomistic simulations

performed in this work.

3.1 Ab initio methods

Ab initio methods can be defined as follows: *"Methods based on the solution of the microscopic Hamiltonian, including all interparticle interaction terms [66]"*.Also known as *first-principles* methods, **ab initio** methods aim to describe the properties of materials and molecules by solving the fundamental quantum mechanical equations that govern their constituent electrons and nuclei, without relying on empirical parameters. The starting point is the time-independent, non relativistic Hamiltonian for a many-body system of interacting electrons and nuclei, given (in atomic units) by

$$\begin{aligned}\hat{H} = &-\sum_n \frac{\hbar^2}{2M_n}\nabla^2 - \sum_e \frac{\hbar^2}{2m_e}\nabla^2 \\ &+ \frac{1}{4\pi\varepsilon_0}\sum_{i<j}\frac{e^2}{|r_i - r_j|} - \frac{1}{4\pi\varepsilon_0}\sum_{i,I}\frac{e^2 Z_I}{|r_i - R_I|} + \frac{1}{4\pi\varepsilon_0}\sum_{I<J}\frac{e^2 Z_I Z_J}{|R_I - R_J|}\end{aligned} \tag{3.1}$$

Here, the first and second terms represent the kinetic energy of nuclei (with positions R_I and masses M_n) and electrons (with positions r_i and mass m_e), respectively. The third term corresponds to the Coulomb interaction between electrons, the fourth term accounts for the attractive interaction between electrons and nuclei, and the fifth term represents the repulsion between nuclei. This Hamiltonian does not include relativistic or spin-dependent effects and forms the foundational basis for electronic structure methods such as Hartree–Fock and density functional theory (DFT). These methods do not rely on empirical fitting parameters and are primarily based on physical constants and calculated results.

3.2 Quantum Many Body problems Born Approximation

Studying large systems poses significant computational challenges since one need to address $6N$ degrees of freedom, where N refers to number of particles. Consequently, an exponential divergence in the magnitude of the problem defines an inherent drawback. At larger atomic and molecular scales, the computational cost of capturing correlated electron dynamics increases rapidly, necessitating compromises in both temporal resolution and the treatment of many-body correlation effects. This section seeks to delineate the connection between ab-initio and semi classical approaches .

3.3 Born–Oppenheimer Method as Originally Derived (1927)

In their seminal 1927 paper *On the Quantum Theory of Molecules*, Born and Oppenheimer introduced a systematic perturbative treatment of the molecular Schrödinger equation, exploiting the large mass disparity between nuclei and electrons. In this section, we provide a concise summary of their original reasoning, complete details can be found in the English translation by Blinder (2000) [67]. To enhance accessibility, we re derive the essential steps here, retaining their conceptual framework while adopting modern operator notation for clarity. This approach serves both as historical context and as a self-contained reference for the reader.

The Born–Oppenheimer approximation formalizes an intuitive physical idea: because electrons are much lighter than nuclei, they adjust almost instantaneously to nuclear displacements. Since nuclei are approximately three orders of magnitude heavier than electrons, their motion is significantly slower. This separation of time scales permits the molecular wavefunction to be expressed as a product of an electronic component (depending parametrically on nuclear positions) and a nuclear component. At leading order, one solves the electronic Schrödinger equation with fixed nuclei to obtain a potential energy surface (PES). The nu-

clei then undergo vibrations and rotations on this surface. Higher-order terms in the perturbative expansion provide systematic corrections, capturing couplings between electronic and nuclear motion. The outcome is a hierarchical structure of molecular energies: electronic, vibrational, and rotational.

3.3.1 Definition of Coordinates and General Setup

We denote the electronic mass and coordinates by lowercase symbols

$$m, \quad (x_k, y_k, z_k),$$

and the nuclear masses and coordinates by uppercase symbols

$$M_l, \quad (X_l, Y_l, Z_l).$$

Let M be a representative average of the nuclear masses $\{M_l\}$. The small parameter for perturbation theory is

$$\kappa = \left(\frac{m}{M}\right)^{1/4}, \tag{3.2}$$

which is typically of order 10^{-1}–10^{-2}.

The full Coulomb interaction $U(x, X)$ defines the molecular potential. The kinetic energy operators are

$$T_E = -\frac{\hbar^2}{2m}\,\nabla_x^2, \tag{3.3}$$

$$T_N = -\frac{\hbar^2}{2M_l}\,\nabla_X^2 \;=\; -\frac{\kappa^4\hbar^2}{2m}\,\nabla_X^2. \tag{3.4}$$

The total Hamiltonian is

$$H = H_0 + \kappa^4 H_1, \tag{3.5}$$

with

$$H_0(x, \partial_x; X) = T_E + U(x, X), \tag{3.6}$$

$$H_1(\partial_X) = -\frac{\hbar^2}{2M_l}\,\nabla_X^2. \tag{3.7}$$

Here H_0 is the clamped-nuclei Hamiltonian, while $\kappa^4 H_1$ accounts for nuclear kinetic energy.

Transforming nuclear coordinates to (ξ_i, θ_i), where ξ_i are internal coordinates and θ_i describe overall translation/rotation, one finds

$$H_1 = H_{\xi\xi} + H_{\xi\theta} + H_{\theta\theta}. \tag{3.8}$$

The stationary Schrödinger equation is

$$\left(H_0 + \kappa^4 H_1 - W\right) \Psi(x, X) = 0. \tag{3.9}$$

3.3.2 Electronic Motion for Stationary Nuclei: Potential Energy Surfaces

Setting $\kappa = 0$ (frozen nuclei), we obtain the electronic eigenvalue problem

$$H_0(x, \partial_x; \xi, \theta)\, \psi_n(x; \xi, \theta) = V_n(\xi)\, \psi_n(x; \xi, \theta), \tag{3.10}$$

where $V_n(\xi)$ defines the *potential energy surface* (PES) on which the nuclei move.

Expansion about equilibrium. For small nuclear displacements

$$\xi_i \to \xi_i + \kappa \zeta_i,$$

the eigenvalue expands as

$$V_n(\xi + \kappa\zeta) = V_n^{(0)} + \kappa V_n^{(1)} + \kappa^2 V_n^{(2)} + \cdots, \tag{3.11}$$

with

$$V_n^{(1)} = \sum_i \zeta_i \frac{\partial V_n}{\partial \xi_i}, \tag{3.12}$$

$$V_n^{(2)} = \frac{1}{2} \sum_{i,j} \zeta_i \zeta_j \frac{\partial^2 V_n}{\partial \xi_i \partial \xi_j}. \tag{3.13}$$

3.3.3 Expansion of Hamiltonian and Wavefunction

The Hamiltonian and wavefunction admit similar expansions:

$$H = H^{(0)} + \kappa H^{(1)} + \kappa^2 H^{(2)} + \cdots, \tag{3.14}$$

$$\Psi_n = \Psi_n^{(0)} + \kappa \Psi_n^{(1)} + \kappa^2 \Psi_n^{(2)} + \cdots. \tag{3.15}$$

Higher-order corrections take the form

$$\Psi_n^{(r)} = \sum_{n'} u_{nn'}^{(r)} \Psi_{n'}^{(0)}, \tag{3.16}$$

where $u_{nn'}^{(r)}$ are polynomials of degree r in the displacements ζ. For instance,

$$u_{nn'}^{(2)} \propto \sum_{i,j} \zeta_i \zeta_j \left\langle \Psi_{n'}^{(0)} \middle| \frac{\partial^2}{\partial \xi_i \partial \xi_j} \Psi_n^{(0)} \right\rangle. \tag{3.17}$$

Substituting the expansions into the electronic eigenvalue equation, multiplying by $\Psi_{n_0}^{(0)*}$ and integrating over x yields

$$u_{nn_0}^{(1)}(V_{n_0}^{(0)} - V_n^{(0)}) + (H^{(1)})_{nn_0} - V_n^{(1)} \delta_{nn_0} = 0, \tag{3.18}$$

$$u_{nn_0}^{(2)}(V_{n_0}^{(0)} - V_n^{(0)}) + (H^{(2)})_{nn_0} + (H^{(1)})_{nn_0} - V_n^{(2)} \delta_{nn_0} = 0, \tag{3.19}$$

and so on. These recursive relations generate successive perturbative corrections.

Expansion of the Nuclear Hamiltonian Using $\partial_\xi = \kappa^{-1} \partial_\zeta$, the nuclear Hamiltonian expands as

$$H_1 = \kappa^2 H_{\zeta\zeta}^{(0)} + \kappa^3 \big(H_{\zeta\theta}^{(0)} + H_{\zeta\zeta}^{(1)}\big) + \kappa^4 \big(H_{\theta\theta}^{(0)} + H_{\zeta\theta}^{(1)} + H_{\zeta\zeta}^{(2)}\big) + \cdots. \tag{3.20}$$

Combining with the electronic expansion gives the full operator

$$\begin{aligned} H = {} & H^{(0)} + \kappa H^{(1)} + \kappa^2 \big(H^{(2)} + H_{\zeta\zeta}^{(0)}\big) \\ & + \kappa^3 \big(H^{(3)} + H_{\zeta\theta}^{(0)} + H_{\zeta\zeta}^{(1)}\big) \\ & + \kappa^4 \big(H^{(4)} + H_{\theta\theta}^{(0)} + H_{\zeta\theta}^{(1)} + H_{\zeta\zeta}^{(2)}\big) + \cdots. \end{aligned} \tag{3.21}$$

Expanding the total energy and wavefunction as

$$W = \sum_{r=0}^{\infty} \kappa^r W^{(r)}, \qquad \Psi = \sum_{r=0}^{\infty} \kappa^r \Psi^{(r)},$$

and matching powers of κ yields

$$(H^{(0)} - W^{(0)})\,\Psi^{(0)} = 0, \tag{3.22}$$

$$(H^{(0)} - W^{(0)})\,\Psi^{(1)} = (W^{(1)} - H^{(1)})\,\Psi^{(0)}, \tag{3.23}$$

$$(H^{(0)} - W^{(0)})\,\Psi^{(2)} = (W^{(2)} - H^{(2)} - H^{(0)}_{\zeta\zeta})\,\Psi^{(0)} + (W^{(1)} - H^{(1)})\,\Psi^{(1)}, \tag{3.24}$$

$$(H^{(0)} - W^{(0)})\,\Psi^{(3)} = (W^{(3)} - H^{(3)} - H^{(0)}_{\zeta\theta} - H^{(1)}_{\zeta\zeta})\,\Psi^{(0)} + \cdots, \tag{3.25}$$

$$(H^{(0)} - W^{(0)})\,\Psi^{(4)} = (W^{(4)} - H^{(4)} - H^{(0)}_{\theta\theta} - H^{(1)}_{\zeta\theta} - H^{(2)}_{\zeta\zeta})\,\Psi^{(0)} + \cdots. \tag{3.26}$$

These recursive relations demonstrate how corrections to the molecular wavefunction and energy can be obtained order by order. In practice, the Born Oppenheimer approximation corresponds to truncating this expansion at leading order, with systematic higher order terms providing a controlled framework for including non adiabatic couplings between nuclear and electronic motion. Thus vibrational contributions appear at $\mathcal{O}(\kappa^2)$, rotational contributions at $\mathcal{O}(\kappa^4)$, and higher-order couplings encode nonadiabatic effects.

3.4 Phase space representation of systems

Part of the results of thesis focuses on classical description of the evolution of classical system composed by many particles. This field is traditionally described by statistical mechanics. Here we review the main concepts relevant for the work addressed in thesis. An important point in this regard is the notion of phase space [68]. Here every degree of freedom of system under study (position and momentum) is described by the axis of multidimensional space. While the system is represented by point in space whose evolution is given by a line in space. So number of dimensions in phase space (Γ space) is linked with number of particles (N).

If we have N number of particles we will expect 6N dimensional Γ space. The instantaneous value of any observable $\mathbf{A}$ will be function of phase space given as $\mathbf{A}(\Gamma)$. This system state evolve over time . To get the time evolution of the observable we need to find the average value $\mathbf{A_{micro}}$ of instantaneous $\mathbf{A}(\Gamma)$ over long time making sure that system have reached to equilibrium. mathematically,

$$\mathbf{A_{micro}} = \lim_{t_{obs}\to\infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} \mathbf{A}(\Gamma(t))\, dt \tag{3.27}$$

One can expect that such long time averaged state represent actual state of the system. This process is called time averaging. This is a general way of approaching system state. Specifically while applying statistical mechanics approach the system is studied in more rigorous way, this demands introducing concept of thermodynamic ensemble. Given the impracticality of accessing all micro states under specific macroscopic conditions, thermodynamic ensembles establish particular thermodynamic constraints to delineate the phase space. This delineation is achieved through the utilization of a partition function, which partitions the phase space with corresponding macroscopic restrictions. Suitable averages are performed on the accessible states to obtain the actual micro-state. The collection of micro states obtained using partition function is called thermodynamic ensemble[69]. We will discuss some commonly used ensembles and they will also be implemented in our work as well.

3.4.1 Microcanonical (NVE) Ensemble

The **microcanonical ensemble**, often denoted as the **NVE ensemble**, refers to a collection of microstates in which the *Number of particles (N)*, the *Volume (V)*, and the total *Energy (E)* of the system remain constant. Hence, the acronym **NVE**. This ensemble describes an *isolated system*, i.e., one that cannot exchange energy or matter with its surroundings. The system is completely determined by the Hamiltonian $H(q, p, t)$, which depends on generalized coordinates q, conjugate momenta p, and possibly on time t. The Hamiltonian equation can be expressed as

$$\frac{dH(q,p,t)}{dt} = \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial t}\dot{t}, \tag{3.28}$$

where $\dot{q}$ and $\dot{p}$ are generalized velocities.

Postulate of Equal A Priori Probability. The fundamental assumption of the microcanonical ensemble is the *postulate of equal a priori probability*, which states that:

> *All accessible microstates of an isolated system with fixed N, V, and E are equally probable.*

This means the system does not prefer one microstate over another, provided that they are consistent with the conservation laws. This postulate forms the statistical foundation for deriving thermodynamic properties.

The microcanonical partition function is defined as

$$Q_{NVE} = \frac{1}{N!}\frac{1}{h^{3N}} \int \delta\big(H(q,p) - E\big)\, dq\, dp, \tag{3.29}$$

where h is Planck's constant and δ is the Dirac delta function, ensuring that only states with energy E contribute. The factor of $1/N!$ accounts for the indistinguishability of identical particles, while the factor of h^{3N} arises from quantum mechanics: the phase space volume element must be normalized in units of h^3 per particle to avoid overcounting microstates.

The corresponding thermodynamic potential of the microcanonical ensemble is the entropy S, given by the Boltzmann formula:

$$S = k_B \ln Q_{NVE}, \tag{3.30}$$

where k_B is the Boltzmann constant. According to Boltzmann [70], an isolated system evolves toward the macrostate with the largest entropy, i.e., the largest number of accessible microstates. This principle underlies the statistical interpretation of the second law of thermodynamics [71].

3.4.2 Canonical (NVT) Ensemble

In the **canonical ensemble**, the system is in thermal contact with a heat bath at a fixed temperature T. Thus, while the number of particles (N) and the volume (V) remain constant, the system energy (E) is allowed to fluctuate due to heat exchange with the reservoir. The probability

distribution of microstates follows the Boltzmann factor. The canonical partition function is given by

$$Q_{NVT} = \frac{1}{N!}\frac{1}{h^{3N}} \int \exp\left(-\beta H(q,p)\right)\, dq\, dp, \tag{3.31}$$

where $H(q,p)$ is the Hamiltonian, h is Planck's constant, and β is defined as

$$\beta = \frac{1}{k_B T}. \tag{3.32}$$

Here k_B is the Boltzmann constant,and $k_B T$ has dimensions of energy, β has dimensions of energy^{-1}. The thermodynamic potential associated with the canonical ensemble is the **Helmholtz free energy** A, related to the partition function as

$$A = -k_B T \ln Q_{NVT}, \tag{3.33}$$

or equivalently,

$$\beta A = -\ln Q_{NVT}. \tag{3.34}$$

The Helmholtz free energy is defined as

$$A = U - TS, \tag{3.35}$$

where U is the internal energy and S is the entropy. At equilibrium, the canonical ensemble minimizes A.

In this work, we employed NVT simulations extensively to study the defect evolution in silicon carbide, since they accurately reproduce the effect of a thermal reservoir on the system.

3.4.3 Isothermal Isobaric (NPT) Ensemble

Isothermal Isobaric ensemble refers to conducting experiments on a system with fixed temperature and pressure. This resembles to NVT ensemble except we also consider fluctuating volume of system as well. The partition function for NPT Ensemble can be given as

$$Q_{NPT} = \frac{1}{N!}\frac{1}{h^{3N}}\frac{1}{V_0} \int \exp -\beta(H(q,p) + PV) dpdqdV \tag{3.36}$$

here β correspond to statistical temperature, while V_0 point out the fixed reference volume. Thermodynamic potential associated with NPT ensemble is Gibbs free energy. Conversely NPT equilibrium refers to minimum value of Gibbs free energy $G = U + PV - TS$.

3.4.4 Grand Canonical Ensemble

Grand canonical Ensemble describes a statistical system in thermal equilibrium that can exchange both energy and number of particles. Average number of particles is determined by the external conditions of the system. Such ensemble is useful for studying system with varying number of particles e.g. liquid and plasma. The γ space for grand canonical ensemble involves N average number of particles with positions q and momenta p. we refereed average number of particles since in practical situation exact number of particles, for a macroscopic system are never precisely known. Density function ($\rho(p, q, N)$) describes the distribution of points in the γ space and is given by following equation

$$Q_{\mu,V,T} = \sum^{N} \frac{1}{N!} \frac{1}{h^{3N}} \exp \beta\mu N \int \exp -\beta(H(p, q))dpdq \tag{3.37}$$

Grand thermodynamic potential is referred as grand potential that does not describe a particular thermodynamic function. It is given by product of pressure and volume

$$-\beta PV = \ln Q_{\mu VT} \tag{3.38}$$

3.4.5 Exploration of phase space and Ergodicity

As already mentioned phase space refers to position momentum (p, q) space where N particle physical system has 6N dimensions $(3p + 3q)$. Time evolution of the physical system can be described as trajectory in such space given that every point refers to one micro state of the system. Usually system passes through each point along trajectory just once, this property is called ergodicity and such systems are called ergodic. whenever you talk of statistics, averaging is one of it's crucial part. It is

Expectation value of quantity to be measured over all possible micro states of the system. Mathematically

$$\langle \mathcal{O}(\Gamma) \rangle = \frac{\int \mathcal{O}(\Gamma) e^{-\beta \mathcal{H}(\Gamma)} \, d\Gamma}{\int e^{-\beta \mathcal{H}(\Gamma)} \, d\Gamma} \tag{3.39}$$

Here Γ represents micro state of the system or it refers to the Γ space. So if one decide to calculate the average of a function with position and momentum of many body system, one can compute using time Averaging (Molecular Dynamics) equation 3.27 or can do ensemble averaging (Monte Carlo) equation 3.39. Moreover it is not required to sample the whole space for the calculation of a well localized quantity in phase space.

3.5 Classical Molecular Dynamics

What molecular Dynamics does, it generate a sample of micro states that represent equilibrium trajectory of system in phase space. Desired properties can be calculated by time average over the generated sample and implying equation 3.27. By **Definition** Molecular dynamics (MD) is a computational technique to model the time evolution of a system of interacting particles by solving Newton's equations of motion. Molecular dynamics update atomic positions and velocities with time thus provides insights into dynamic processes such as melting of a crystal or implantation in the material and relevant formations.

3.5.1 Key Principles

1. **Newton's Second Law**: The motion of each atom i is given by:

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2}, \tag{3.40}$$

   Here $\mathbf{F}_i$ is the force, m_i is the mass, $\mathbf{a}_i$ is the acceleration, and $\mathbf{r}_i$ is the position of atom i.

2. **Force from Potential Energy**: Forces are derived from a potential energy function $U(\mathbf{r})$:

$$\mathbf{F}_i = -\nabla_i U(\mathbf{r}). \tag{3.41}$$

The potential had bonded and non-bonded interaction terms. A generic potential form can be described as:

$$U = \sum_{\text{bonds}} \frac{k_b}{2}(r - r_0)^2 + \sum_{\text{angles}} \frac{k_\theta}{2}(\theta - \theta_0)^2 + \sum_{\text{pairs}} \left[4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) + \frac{q_i q_j}{4\pi\epsilon_0 r}\right] \quad (3.42)$$

representing harmonic bond/angle terms, Lennard-Jones, and Coulomb interactions. This is a general Potential defining all the terms. In our specific case we will use Tersoff and EDIP potential to represent the forces among silicon carbide atom. They have more parametrization to given by the respective authors that fits perfectly to model atomic forces in silicon carbide. The details of these potentials are described in section 3.5.6 and 3.5.7.

3. **Numerical Integration**: Positions and velocities are updated by integrating equation of motion and using algorithms like Velocity Verlet:

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{\mathbf{F}_i(t)}{2m_i}(\Delta t)^2, \quad (3.43)$$

   Here $\Delta t \sim 1\,\text{fs}$ describes the time step or the time after which algorithm should update positions and velocities. Velocity Verlet is briefly discussed in the section 3.5.8.

4. **Thermodynamic Properties**: Trajectories yield macroscopic properties, e.g., temperature:

$$\langle T \rangle = \frac{1}{3Nk_B} \sum_i m_i \langle \mathbf{v}_i^2 \rangle, \quad (3.44)$$

   where N is the number of atoms and k_B is Boltzmann's constant. Similarly desired thermodynamic properties can be calculated. This is definition of ergodicity as discussed in section 3.4.5.

5. **Applications**: MD elucidates systems intractable to experiments, such as drug target binding or nano scale material behavior, by quantifying metrics like root-mean-square deviation and other thermodynamic properties.

This framework enables the study of atomic scale dynamics with high precision, e.g. Defect complex formations , particle diffusion.

3.5.2 Limitations

Classical framework at atomistic scale While implementing Newtonian equations of motion at the atomic level, a natural question arises: given that atomic systems are more accurately described by quantum mechanics rather than classical physics, why are we relying upon simple but rather less accurate classical framework ? why not use the Schrodinger equation instead? The simplest answer is that quantum effects become significant in systems at sufficiently low temperatures. An example of measurable quantum effect in solids is the drop in specific heat capacity of a crystal below the Debye temperature. MD results should be analyzed carefully in these regions. Moreover classical approximations can be valid under certain conditions, which are determined by the de Broglie thermal wavelength (Λ).

De Broglie Thermal Wavelength

The de Broglie thermal wavelength is defined as:

$$\Lambda = \sqrt{\frac{2\pi\hbar^2}{Mk_BT}}$$

where M is the atomic mass while T is temperature, $\hbar$ is the reduced Planck constant, and k_B is the Maxwell Boltzmann constant.

We can postulate about validity of classical mechanics in the limit when Λ is quite smaller than atomic separation. Mathematically $\Lambda \ll a$ refers to a localized particle means, when De Broglie wavelength is smaller, particles are more localized and easily be treated by classical mechanics. An example is for liquids at triple point. $\Lambda = 0.1a$ for light element and it decrease for heavier elements, thus making classical approximation less valid for light element systems e.g H_2, He, Li.

3.5.3 Choice of Potential

In molecular dynamics (MD) simulations, the reliability of predicted material behavior strongly depends on the quality of the interatomic potential employed. Atoms interact and change their relative positions, resulting in varying forces that dictate structural stability, phase behavior, and defect evolution. A simulation is therefore realistic only to the extent that the chosen interatomic forces resemble those experienced by real atoms in the material under investigation.

Several empirical force fields have been developed for SiC, including the Tersoff family of potentials [72, 73], the modified embedded atom method (MEAM) [74], and the environment dependent interatomic potential (EDIP) [75]. The Tersoff formulation has been widely applied to covalent systems and reproduces many bulk and surface properties of SiC, however, it tends to overestimate melting temperatures and often fails to capture defect energetics with sufficient accuracy [76]. MEAM, on the other hand, provides a more transferable description of bonding environments, but its computational cost is higher and reported accuracy for point defect migration barriers in SiC remains limited [77].

The EDIP potential was specifically designed to incorporate environment dependent coordination effects, making it well suited for studying defect formation and diffusion processes in covalent systems. Previous benchmarking studies have demonstrated that EDIP accurately reproduces structural and thermodynamic properties of SiC, including lattice parameters, elastic constants, and defect energetics, while maintaining computational efficiency [75, 1]. In particular, EDIP provides a melting point prediction ($T_m \sim 2620$ K) in close agreement with experimental values ($T_m \sim 2818$ K) [78], whereas Tersoff significantly overestimates this temperature (above 3500 K).

For these reasons, EDIP was selected as the interatomic potential for this work. A more detailed description of its formalism is provided in Section 3.5.5, and its performance is validated in Chapter 4 against structural and thermodynamic benchmarks.

3.5.4 Periodic boundary conditions

Generally, a potential is infinite ranged, but for practical purposes, it is customary to add a cutoff radius R_C and disregard interactions beyond a certain distance. This might cause issues otherwise, imagine a particle pair crossing the cutoff distance. There will be a minute jump in energy, and many such events can spoil energy conservation. These effects of potential truncation are approximated by imposing periodic boundary conditions or treating the system as uniform beyond the cutoff.

Additionally, the simulated system shouldn't simply terminate at the boundary, otherwise, atoms at the boundary will have fewer neighbors than those inside. Moreover, since the number of atoms in a simulation box is much smaller than in real/macroscopic matter (10^{23}), the ratio of surface atoms to the total number of atoms will be much larger, resulting in more dominant surface effects.

Periodic boundary conditions (PBC) solve this problem by replicating the box infinitely through rigid translation in all three directions. For example, having a particle within it actually represents an infinite number of particles located at distances **r** given by:

$$\mathbf{r} + l\mathbf{a} + m\mathbf{b} + n\mathbf{c} \quad (l, m, n = -\infty, \infty) \tag{3.45}$$

Here l, m, n are numbers while a, b, c represent the x, y and z length of cell box. With this assume infite number of particles all moving togeather and there are interbox interactions. Thus, inter particle interactions can go through box boundaries, effectively eliminating surface effects. Meanwhile, the translation of particles relative to particles leaves forces unchanged, this means box boundaries have no effect. Note: one might consider that replicating infinitely might increase the interaction pairs infinitely, but this is not true since the potential usually has a short interaction range. This is governed by what is commonly known as the minimum image criterion. considering a particle i interacting with j, minimum image criterion means means among all the possible images of j, select the nearest and threw away all others.

3.5.5 Force fields and Force calculation among atoms

The time evolution of an ensemble can be described by the Newtonian equations of motion, which are governed by the system's Hamiltonian. Hamiltonian provides a comprehensive framework for describing the dynamics of the system. For instance, the Hamiltonian of a water molecule in the harmonic approximation is expressed as:

$$\hat{H} = \sum_i \frac{p_i^2}{2m_i} + \frac{1}{2}\sum_{\text{bonds}} k_b(b-b_0)^2 + \frac{1}{2}\sum_{\text{angles}} k_\theta(\theta-\theta_0)^2 + \sum_{i>j}\frac{e_i e_j}{4\pi\epsilon_0 r_{ij}} \sum_{i>j} 4\eta_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \quad (3.46)$$

This Hamiltonian includes terms accounting for both bonded interactions (kinetic and potential energy) and non bonded interactions (electrostatic and dispersion forces). By applying the Hamiltonian framework from Equation **??**, the time evolution of the system can be systematically analyzed.

$$\frac{dp_i}{dt} = m\ddot{q} = -\nabla V(q), \quad \frac{dq_i}{dt} = \frac{p_i}{m_i} \quad (3.47)$$

The first equation points out that force is evaluated as the negative gradient of potential (interatomic potential in our case), while the second establishes a connection among position, momentum, and velocities. Thus, force can be evaluated based on atomic distances. For N atoms, there is an evaluation of $\frac{1}{2}[N(N-1)]$ atomic distances for force calculation. This is, by far, the most computationally intensive task in molecular dynamics (MD), as it scales quadratically. Methods have been proposed to address this problem. Thanks to schemes like "using neighbor lists," the problem is mitigated by separating long range and short range interactions, thereby optimizing scaling to N.

3.5.6 Tersoff Force Field

The Tersoff potential [79][72][73] models the total energy E of a system as a sum of pair like interactions, modulated by a bond order term that

depends on the local atomic environment. The total energy is given by:

$$E = \frac{1}{2} \sum_{i} \sum_{j \neq i} V_{ij}, \tag{3.48}$$

where the pair potential V_{ij} between atoms i and j at distance r_{ij} (denoted r in the parameters) is:

$$V_{ij} = f_C(r_{ij}) \left[V_R(r_{ij}) - b_{ij} V_A(r_{ij}) \right]. \tag{3.49}$$

Cutoff Function

The cutoff function $f_C(r_{ij})$ limits the interaction range:

$$f_C(r_{ij}) = \begin{cases} 1, & r_{ij} < R - D, \\ \frac{1}{2} - \frac{1}{2} \sin\left(\frac{\pi (r_{ij} - R)}{2D} \right), & R - D \leq r_{ij} \leq R + D, \\ 0, & r_{ij} > R + D, \end{cases} \tag{3.50}$$

where R (e.g., $\tau_{0,IJ}$) and D define the cutoff range.

Repulsive and Attractive Terms

The repulsive term is:

$$V_R(r_{ij}) = A e^{-\lambda_1 r_{ij}}, \tag{3.51}$$

and the attractive term is:

$$V_A(r_{ij}) = B e^{-\lambda_2 r_{ij}}, \tag{3.52}$$

where A, B, λ_1 (possibly $\lambda_{1,J}$), and λ_2 (possibly $\lambda_{4,J}$) are parameters.

Bond-Order Term

The bond-order term b_{ij} accounts for the local environment:

$$b_{ij} = \left(1 + \beta^n \zeta_{ij}^n \right)^{-1/(2n)}, \tag{3.53}$$

where:

$$\zeta_{ij} = \sum_{k \neq i,j} f_C(r_{ik}) g(\theta_{ijk}) e^{\lambda_3^3 (r_{ij} - r_{ik})^3}, \tag{3.54}$$

and the angular term is:

$$g(\theta_{ijk}) = 1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + (h - \cos\theta_{ijk})^2}, \tag{3.55}$$

with β, n, c, d, h, and λ_3 as parameters, and θ_{ijk} (possibly related to C_{IJK}) as the angle between bonds $i - j$ and $i - k$.

Parameters

The potential includes parameters such as H_J, $Z_{I,J}$, Y_{IJ}, V_{IK}, $\tau_{C,IJ}$, and $\tau_{0,JK}$, which define energy scales, decay rates, and cutoff distances. The details of paramterization is provided in Appendix section A.

3.5.7 EDIP force field

The Environment dependent interatomic potential [80] (EDIP) potential is a many body potential used to model interactions among atoms in materials. It consists of two main parts given as, a two body interaction term and a three body interaction term. The total energy E_i of an atom i is given by:

$$E_i = \sum_{j \neq i} \phi_2(R_{ij}, Z_i) + \sum_{j \neq i} \sum_{k > j} \phi_3(R_{ij}, R_{ik}, Z_i)$$

Two Body Interaction Term

The two body interaction term is given by:

$$\phi_2(r, Z) = A \left(\left(\frac{B}{r} \right)^{\rho} - e^{-\beta Z^2} \right) \exp\left(\frac{\sigma}{r - a} \right)$$

where r is the distance between two atoms, Z is the embedding function, and A, B, ρ, β, σ, and a are material specific constants.

Three Body Interaction Term

The three body interaction term is given by:

$$\phi_3(R_{ij}, R_{ik}, Z_i) = \exp\left(\frac{\gamma}{R_{ij} - a} \right) \exp\left(\frac{\gamma}{R_{ik} - a} \right) h(\cos\theta_{ijk}, Z_i)$$

where R_{ij} and R_{ik} are the distances between atoms i and j, and i and k respectively, θ_{ijk} is the angle between the vectors $\vec{R}_{ij}$ and $\vec{R}_{ik}$, and h is defined as:

$$h(l, Z) = \lambda \left[\left(1 - e^{-Q(Z)(l+\tau(Z))^2}\right) + \eta Q(Z)(l + \tau(Z))^2 \right]$$

with $Q(Z) = Q_0 e^{-\mu Z}$ and $\tau(Z) = u_1 + u_2(u_3 e^{-u_4 Z} - e^{-2u_4 Z})$.

Embedding Function

The embedding function Z_i is given by:

$$Z_i = \sum_{m \neq i} f(R_{im})$$

where f is a cutoff function defined as:

$$f(r) = \begin{cases} 1 & \text{if } r < c \\ \exp\left(\frac{\alpha}{1-(r/c)^3}\right) & \text{if } c < r < a \\ 0 & \text{if } r > a \end{cases}$$

Here, c and a are cutoff radii, and α is a parameter that controls the smoothness of the cutoff. The details of parameterization can be accessed in Appendix section A.

3.5.8 Integrating Equation of Motion (Dynamics)

Once forces among atoms are determined, numerical integration for the newton equation of motion can be performed, that should give the updated positions and momenta. Basically, knowing positions and some of their time derivative at a given time, integrator give us the same quantities at a later time Δt. These iterations can be expanded to long times. Generally integrator are based on finite difference methods where time Δt is discretized on a finite grid. Various integrator are suggested for this purpose generally a good integrator should have following characteristics.

- Robust and memory efficient.

- Provide capability of implementing long integration time step (Δt). Since integration step is our choice, a larger time step allows less evaluation of forces and less integration step for the a given trajectory, making simulation fast. This choice is limited by the variation of DOF, a smart choice of time step shouldn't miss important DOF. This upper bound is imposed by the high temperature and smaller mass.

- Conservation laws of energy and momentum should be followed. It should also respect time reversibility. Energy conservation is strongly linked with another property commonly referred as "symplectic". This tells that integrator should preserve the volume in phase space.

Most common algorithm used for MD is Known as "Velocity Verlet Algorithm" which is updated form of "Verlet" Algorithm. Here is the computational frame work of velocity Verlet. The position update uses the current position, velocity, and acceleration to estimate the new position after a small time step Δt

$$\mathbf{R}(t+\Delta t) = \mathbf{R}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 \tag{3.56}$$

The velocity update (in two steps) first uses the current acceleration to estimate the velocity at an intermediate time $\left(t+\frac{\Delta t}{2}\right)$

$$\mathbf{v}\left(t+\frac{\Delta t}{2}\right) = \mathbf{v}(t) + \frac{1}{2}\mathbf{a}(t)\Delta t \tag{3.57}$$

After updating the position, the force is recalculated to update the acceleration:

$$\mathbf{a}(t+\Delta t) = -\frac{1}{m}\nabla V(\mathbf{R}(t+\Delta t)) \tag{3.58}$$

Finally, the velocity is updated to the next full time step $(t+\Delta t)$

$$\mathbf{v}(t+\Delta t) = \mathbf{v}\left(t+\frac{\Delta t}{2}\right) + \frac{1}{2}\mathbf{a}(t+\Delta t)\Delta t \tag{3.59}$$

so we need 9N memory locations to save this 3N information. Further, Predictor corrector algorithms offers another class of methods to integrate equation of motion.

3.6 Searching Minimum Energy Path (MEP)

3.6.1 Introduction

A significant challenge in theoretical chemistry and condensed matter physics is identifying the lowest energy path for rearranging a group of atoms from one stable configuration to another. This path is often termed the *minimum energy path* (MEP) and is crucial for defining a *reaction coordinate* for various transitions such as chemical reactions, molecular conformation changes, or diffusion processes in solids. The Minimum Energy Path (MEP) between two local minima is crucial for understanding reaction mechanisms. Various methods are used in computational chemistry to find the MEP. The MEP's potential energy maximum, known as the saddle point energy, provides the activation energy barrier, which is essential for estimating transition rates within harmonic transition state theory [81].

Numerous methods have been developed to locate reaction paths and identify saddle points, which are critical for understanding reaction mechanisms. A common class of algorithms initiates at a local minimum on the potential energy surface (PES) corresponding to the reactant state and traces the path of steepest or slowest ascent in a sequential, stepwise manner [82, 83, 84]. However, such approaches do not always converge to a true saddle point. This limitation often arises because the PES is effectively frozen during the path following procedure that is, the algorithm neglects dynamic coupling to orthogonal degrees of freedom and local anharmonicity. As a result, the traced trajectory may veer off the minimum energy reaction path and terminate at a non stationary region or fail to approach a first-order saddle point.

A common approach involves calculating the normal modes of a local harmonic approximation of the potential energy surface and following each mode until a saddle point is found. Each step requires the evaluation and diagonalization of the second derivative matrix, limiting this method to small systems where second derivatives are readily available.

Another family of methods that will be the focus of our interest utilize a two point boundary condition, specifying both initial and final configurations for the transition. These configurations are typically two local

minima on the multidimensional potential energy surface, often obtained through finite temperature molecular dynamics or Monte Carlo simulated annealing methods. The discussed methods require first derivatives of the potential energy, such as the *drag* method or the *reaction coordinate* method. These methods define a progress variable through linear interpolation between the initial and final configurations, incrementing this variable stepwise to minimize the remaining degrees of freedom [85, 86].

While these methods can work well in simple cases, they may fail in more complex scenarios where the path generated can be discontinuous. Additionally, atomic coordinates may *slip* near the saddle point region, potentially missing the saddle point configuration. Even when the initial straight line interpolation is close to the saddle point, relaxation of the remaining degrees of freedom can yield results far from the MEP. For certain values of the drag coordinate, there may be two minima, further complicating the process.

3.6.2 Chain of States class of methods and their criticalities

The earliest method for finding MEPs involves connecting several images to trace the path or what is commonly known as the Plain Elastic Band(PEB) and it belongs to the class of cain of states approaches. In this method, a chain of replicas (states/images) of the system is generated between initial and final configurations. All intermediate images are connected with springs of zero natural length and minimized simultaneously with a suitable minimization procedure. This should result in a discrete distribution of images that form the MEP. Generally the distribution of intermediate image is twice as high in energy than the end point replicas. The path is a minimum energy path if the forces are oriented directly along the path, i.e., $\mathbf{F} = -\nabla V$ and the force in the initial and final configurations is zero. This is commonly known as Plain Elastic Band(PEB) method. Theoretically, this method involves minimizing the energy of objective function between the initial and final images with respect to the intermediate images, while the initial and final images R_0 and R_P will

remain fixed.

$$S^{PEB}(R_1,\dots,R_{P-1}) = \sum_{i=0}^{P} V(R_i) + \sum_{i=1}^{P} \frac{Pk}{2}(R_i - R_{i-1})^2 \tag{3.60}$$

An example is presented in figure 3.2.

We will explain the method by by posing an example problem where an atom B can form bonds with either one of two fixed atoms A or C, and the latter is coupled with a harmonic potential to another atom D. The whole system is represented by what we generally refer as LEPS potential [87]. The force felt by an image i is given by the following equation.

$$\vec{F}_i = -\nabla V(\vec{R}_i) + \vec{F}_i^s \tag{3.61}$$

here

$$\vec{F}_i^s \equiv k_{i+1}\left(\vec{R}_{i+1} - \vec{R}_i\right) - k_i\left(\vec{R}_i - \vec{R}_{i-1}\right). \tag{3.62}$$

Results of 2D LEPS potential, for two spring constants $k = 1.0$ and $k = 0.1$, can be represented as contour plots of the potential energy surface as shown in figure 3.1. The AB distance rAB is presented on x-axis, while the BC distance rBC is represented on the y-axis. When the spring constant is high the elastic band is too stiff, it cuts the corner and MEP miss the saddle point. On the other hand, if the spring constant is very small, now the MEP come close to saddle point but slide down the path to avoid the barrier point again thus skipping the most critical part of MEP, this problem has been discussed previously [89].

To solve this problem we can define a continuum limit where the number of images, P tends to infinity, and the spring constants $k(P)$ can be linked with the number of images P. By increasing the number of images P, $\frac{k(P)}{2P}$ obtains a finite value and we get a stiff elastic band. This becomes a potential reason of corner cutting. It has been observed that even P becomes large, conversely, making $\frac{k(P)}{2P}$ zero, the corner cutting problem still persists.

$$S^{PEB} = P \int d\lambda \left(V(\vec{R}(\lambda)) + \frac{k(P)}{2P} \left| \frac{d\vec{R}}{d\lambda} \right|^2 \right) \tag{3.63}$$

where:

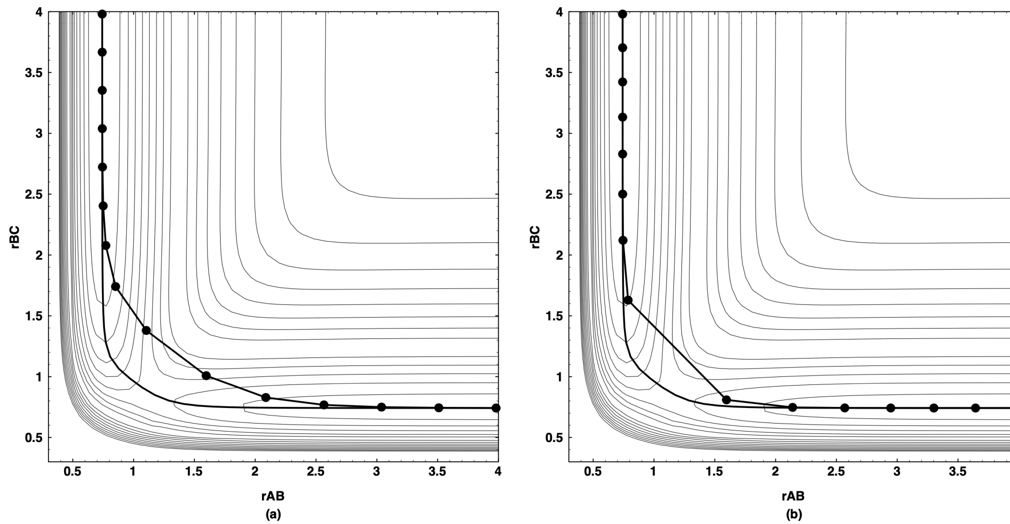


Figura 3.1: Figure represents the contour plot of the potential energy surface for the discussed 3 atoms (A, B, C) problem represented by LEPS potential [87]. rAB and rBC gives the inter atomic distances for given atoms. Solid line represents the MEP obtained by standard NEB in both figures. while the solid line with filled curves gives the orientation of MEP as obtained by plain of states method. Figure a and b overviews the respective problems with plain of states method on extremely small and large value of spring constant k. Figure is used form [88]

- P is the number of images representing the path.
- λ is a parameter varying from 0 to 1 along the path.
- $V(\vec{R}(\lambda))$ is the potential energy along the path.
- $k(P)$ is the spring constant, potentially dependent on P.
- $\left|\frac{d\vec{R}}{d\lambda}\right|^2$ represents the deviation from a straight path.

Here are key additions from this objective function that make it different from the simpler functions.

- Including a term that penalizes deviations from a straight path.
- Handling large P by adjusting the stiffness of the band.

3.6.3 Nudging

The problem with chain of states method is solved in a simple way after diagnosing the reason of it. Off course one has to used sufficient number of images while finding the MEP. Since the images are pulled perpendicularly downward and avoid saddle region in corner cutting, we can avoid this by skipping the perpendicular component of **harmonic force** among images. Similarly problem with skidding down and skipping critical region can be avoided by skipping the parallel component of **True force** $\Delta V(R_I)$ in the direction of MEP. Strategically both partial components balance each other and guide the MEP follow through the critical regions. We will redefine the force as,

$$\vec{F}_i^0 = -\vec{\nabla} V(\vec{R}_i)_\perp + \vec{F}_i^s \cdot \hat{\tau}_\parallel \hat{\tau}_\parallel \tag{3.64}$$

Where $\tau_\parallel$ is the unit vector along the path [90, 91, 92]. This projection of just the perpendicular component of true spring potential and parallel component of spring force accurately decouple the dynamics of path from the distribution of images . Thus images are relaxed staying independent from spring force and helps to follow the correct MEP, This procedure is

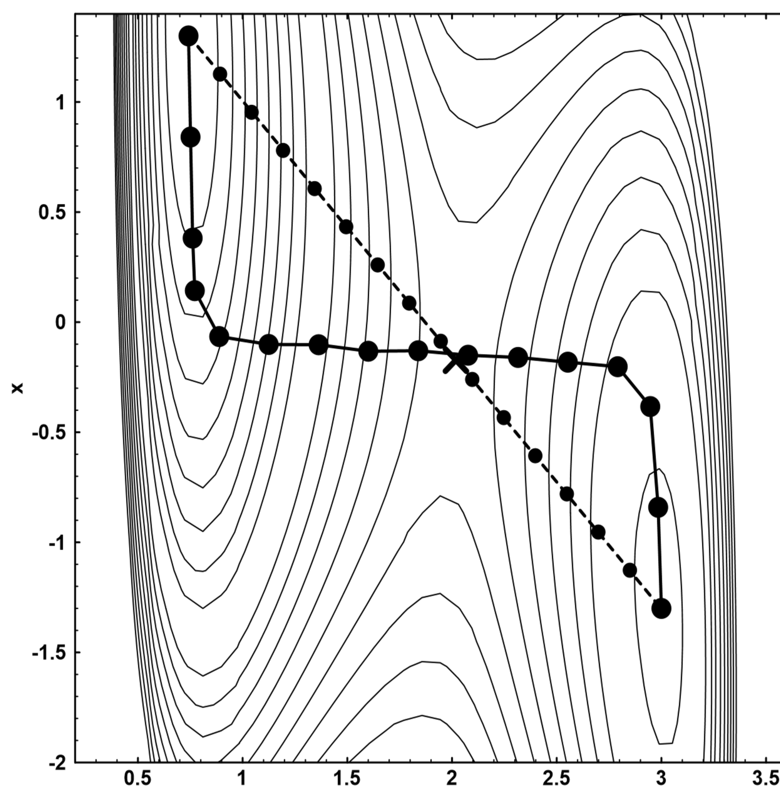


Figura 3.2: The initial and final configuration of NEB with 16 images is presented. Dashed lines with small circles represent the straight line interpolation between initial and final configuration. Solid line with big circles represent the path closer to MEP obtained after minimization of all images. Spring constant of $k = 0.5$ was implemented closer to the end points and and $k = 1.0$ at the middle to magnify the critical central region. Figure is used from reference [88].

commonly referred as **nudging** so the term Nudged Elastic Band method is introduced.

The choice of spring constant is crucial as it controls the distribution of images within the path. If too high, it adds an extra potential, causing the path to miss the real transition state. If too low, the images tend to slide downhill, avoiding the barrier region.

Climbing Nudged Elastic Band method refers to a modification of the NEB method where a particular image with the highest energy is forced to climb to the saddle point while keeping the spring force unaffected [90].

$$\mathbf{F}_{i_{max}} = -\nabla E(R_{i_{max}}) - 2\nabla E(R^{\parallel}_{i_{max}})$$

The objective function in the NEB method is designed to balance the minimization of potential energy with the requirement for a straight path. By carefully tuning the spring constant and using the path deviation term, the NEB method effectively finds the MEP between two states on a potential energy surface. This makes it a powerful tool for studying reaction mechanisms and phase transitions in materials science and chemistry.

Despite optimization of the spring constant and careful preparation of initial and final configurations, the NEB result still exhibits non physical artifacts such as unexpected energy dips and irregular curvature. Below are several possible reasons for the failure of the NEB path in this case:

1. **Inaccurate Tangent Estimation:** NEB relies on the estimation of local tangents to project forces. Poor tangent estimates, especially in regions of high curvature, can cause incorrect projection of spring and true forces, leading to deviation from the MEP.

2. **Corner Cutting and Path Kinking:** If spring forces dominate or are not properly orthogonalized, images may cut corners of the true MEP, producing artificially low energy segments or non physical transitions. This is a known issue in basic implementations of NEB. I couldn't test it in LAMMPS.

3. **Inadequate Initial Path Interpolation:** Linear interpolation between initial and final states may produce geometrically unrealistic intermediate configurations, especially in complex defect systems. This can result in high energy artifacts or false minima.

4. **Inaccurate End State Geometries:** While endpoint energies have been separately verified, the geometries(carbon antisite or Silicon antisite) themselves might correspond to local, not global, minima. This can mislead the path optimization and create nonphysical intermediate states.

5. **Limitations of the Empirical Potential:** The empirical interatomic potential (Tersoff, EDIP) used may not accurately capture defect energetics for transition states. This can lead to qualitatively incorrect energy profiles along the NEB path.

3.7 Stochastic Modeling of Brownian Motion

Einstein's pioneering theoretical study of Brownian particles is the first example of statistical mechanics approaches to establish a connection between the microscopic diffusion coefficient and macroscopic physical quantities, such as mobility. Specifically, by using a probabilistic method, he defined a connection between average value of squared displacement traveled by Brownian particle r(t) in a time interval t and diffusion coefficient D [93, 94].

Einstein's hypothesis was based on the observation that suspended particles in a liquid, which differ from molecules solely in size, must still obey the diffusion equation that governs the chaotic thermal motion of molecules within the liquid.

3.7.1 From Discrete Steps to Continuous Diffusion, Einstein's Convolution Derivation

Einstein's theoretical treatment of Brownian motion is founded on the assumption that each particle undergoes random, uncorrelated displacements over time. These displacements are statistically independent both of the motions of other particles and of previous motions of the same particle, provided that the time interval t is not infinitesimally small. For analytical tractability, the model is reduced to a one dimensional scenario. In the following section we will derive Einstein equation of brownian mo-

tion and prove it exactly taking the form of diffusion equation. Consider a particle undergoing to N discrete displacement steps, each of spatial size ξ, within a time interval τ. We define the probability density function $f_\tau(\xi)$ such that:

$$f_\tau(\xi)\,d\xi = \frac{dN}{N}, \quad \text{or equivalently,} \quad dN = N f_\tau(\xi)\,d\xi \tag{3.65}$$

Note that since $d\xi$ has units of length [L], the probability density

$$f_\tau(\xi)\,d\xi = \frac{dN}{N}$$

must be dimensionless. Hence

$$[f_\tau(\xi)] = [d\xi]^{-1} = \mathrm{L}^{-1},$$

so that $f_\tau(\xi)\,d\xi$ correctly matches the unitless fraction dN/N.

The function $f_\tau(\xi)$ represents the likelihood of a displacement ξ occurring in the interval τ, and due to symmetry in the absence of drift, we require:

$$f_\tau(\xi) = f_\tau(-\xi) \tag{3.66}$$

This condition reflects the fact that forward and backward steps are equally probable.

Let $\rho(x,t)$ be the particles' density at position x and time t. The change in density after an interval τ can be written as:

$$\rho(x, t+\tau) = \int_{-\infty}^{+\infty} \rho(x-\xi, t) f_\tau(\xi)\,d\xi \tag{3.67}$$

Note that the right-hand side is precisely the convolution

$$(\rho * f_\tau)(x,t) = \int_{-\infty}^{\infty} \rho(x-\xi,t)\,f_\tau(\xi)\,d\xi \tag{3.68}$$

Physically, this convolution "collects" at each point x all particles that were at $x-\xi$ at time t, each weighted by the probability $f_\tau(\xi)$ of jumping distance ξ in the interval τ.

This expression accounts for particles arriving at position x from displacements ξ in both directions. Expanding both sides of Equation (3.67) in a Taylor series yields: **Left-hand side (LHS):**

$$\rho(x, t+\tau) = \rho(x,t) + \tau\frac{\partial \rho}{\partial t} + \frac{\tau^2}{2!}\frac{\partial^2 \rho}{\partial t^2} + \cdots \tag{3.69}$$

Right-hand side (RHS):

$$\int_{-\infty}^{+\infty} \rho(x-\xi, t) f_\tau(\xi)\, d\xi = \rho(x,t) \int f_\tau(\xi)\, d\xi - \frac{\partial \rho}{\partial x}\int \xi f_\tau(\xi)\, d\xi + \frac{1}{2!}\frac{\partial^2 \rho}{\partial x^2}\int \xi^2 f_\tau(\xi)\, d\xi + \cdots \tag{3.70}$$

Given the properties:

$$\int_{-\infty}^{\infty} f_\tau(\xi)\, d\xi = 1, \quad \int_{-\infty}^{\infty} \xi f_\tau(\xi)\, d\xi = 0,$$

we simplify Equation (3.70) to:

$$\rho(x,t) + \frac{1}{2}\frac{\partial^2 \rho}{\partial x^2}\int_{-\infty}^{\infty} \xi^2 f_\tau(\xi)\, d\xi + \cdots \tag{3.71}$$

Equating the leading-order terms from Equations (3.69) and (3.71), and neglecting higher-order terms, we arrive at the diffusion equation:

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2} \tag{3.72}$$

where the diffusion coefficient D is related to the second moment of the step size distribution:

$$D = \frac{1}{2\tau}\int_{-\infty}^{\infty} \xi^2 f_\tau(\xi)\, d\xi \tag{3.73}$$

This result is derived under the key assumption that $f_\tau(\xi)$ depends only on the current displacement ξ and is independent of the particle's trajectory history. Such a process is now termed a Markov process in stochastic theory. Einstein's treatment thus forms a cornerstone of the mathematical theory of random processes and stochastic differential equations [93, 95].

3.7.2 Connection to Fick's Law and the continuity equation

The concept of *flux* is central to the description of diffusive transport. In one dimension, the flux $J(x,t)$ is defined as the amount of substance (e.g., particles or mass) crossing a unit area per unit time at position x and time t. It has physical units of $[\text{quantity} \cdot \text{length}^{-2} \cdot \text{time}^{-1}]$.

Building on this, Fick's First Law provides a phenomenological relationship between the particle flux and the concentration gradient. First proposed by Adolf Fick in 1855 [96], the law states that:

$$J(x,t) = -D\frac{\partial \rho(x,t)}{\partial x}, \tag{3.74}$$

where:

- $J(x,t)$ is the particle flux,
- $\rho(x,t)$ is the particle concentration (or number density),
- D is the diffusion coefficient (assumed constant here),
- The negative sign reflects that diffusion occurs from regions of high concentration to low concentration.

To derive the diffusion equation, we apply the principle of local conservation of mass, which is encoded in the one dimensional continuity equation:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial J(x,t)}{\partial x} = 0. \tag{3.75}$$

Substituting Equation (3.74) into the continuity equation (3.75), we obtain:

$$\frac{\partial \rho(x,t)}{\partial t} = D\frac{\partial^2 \rho(x,t)}{\partial x^2}, \tag{3.76}$$

which is the classical diffusion equation governing the time evolution of the concentration field $\rho(x,t)$. This equation forms the theoretical foundation for a wide range of transport phenomena in physics, chemistry, and biology.

3.7.3 Mean Square displacement and estimation of Arrhenius type Energy Barrier

We can easily verify that the solution of the diffusion equation (3.72) with the initial condition $\rho(x,0) = \delta(x)$ is the Gaussian distribution:

$$\rho(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \tag{3.77}$$

This solution maintains normalization and describes the spreading of particles' probability over time. The profile is initially a Dirac delta function at the origin and gradually broadens with time, while conserving the total probability:

$$\int_{-\infty}^{\infty} \rho(x,t)\,dx = 1$$

Furthermore, the mean square displacement can be computed as:

$$\langle x^2(t) \rangle = \int_{-\infty}^{\infty} x^2 \rho(x,t)\,dx = 2Dt \tag{3.78}$$

or

$$D = \frac{\langle x^2(t) \rangle}{t} \tag{3.79}$$

which quantifies the average spread of the particle cloud and provides experimental access to the diffusion coefficient D.

Einstein's elegant formulation not only describes the microscopic basis for macroscopic diffusion but also introduces the idea of random processes and sets the stage for modern probability theory and stochastic calculus. Using the familiar result of gaussian integrals one can check that solution is properly normalized. Gaussian integrals are presented in Appendix section D.

3.8 Einstein's Derivation and Its Modern Relevance

Einstein's 1905 theory of Brownian motion provides a probabilistic model for the motion of suspended particles, introducing a fundamental link

between microscopic particle dynamics and macroscopic diffusion. The derivation starts from the convolution of probability distributions for discrete random steps, leading to a Gaussian distribution in the limit of continuous motion This yields a relationship between the **mean square displacement (MSD)** and the **diffusion coefficient**:

$$\langle r^2(t) \rangle = 2nDt \tag{3.80}$$

where n is the spatial dimension. The diffusion coefficient D is related to the particle's mobility μ by the Einstein Smoluchowski relation:

$$D = \mu k_B T = \frac{k_B T}{\gamma} \tag{3.81}$$

Diffusion of Charged Particles Einstein Smoluchowski Relation

In systems involving charged particles, such as ions in a solution or electrons in a semiconductor, the diffusion process is often influenced by external electric fields. In such cases, the particles not only undergo random thermal motion but also drift in response to the field.

The drift velocity v_d of a charged particle under an electric field $\vec{E}$ is given by:

$$v_d = \mu \vec{E} \tag{3.82}$$

where μ is the particle's mobility, defined as the ratio of drift velocity to the applied field.

Albert Einstein originally derived a relationship between this mobility and the diffusion coefficient D, which was later extended to include charged species. The resulting expression is known as the **Einstein–Smoluchowski relation**:

$$D = \mu k_B T \tag{3.83}$$

where D is the diffusion coefficient, μ is the electrical mobility of the particle, k_B is the Boltzmann constant, and T is the absolute temperature. This relation links thermally driven diffusion to field driven transport and is particularly important in describing ionic conduction in electrolytes, electrophoretic mobility, and charge transport in semiconductor materials.

Diffusion in Viscous Media Stokes–Einstein–Sutherland Equation

When particles diffuse in a viscous medium (e.g., colloids in water, proteins in cytoplasm, or nanoparticles in polymers), hydrodynamic interactions and viscous drag must be accounted for. This leads to a modification of the diffusion coefficient that includes the effects of the medium's viscosity.

Assuming a spherical particle moving in a continuum fluid at low Reynolds number, the hydrodynamic drag force is described by Stokes' law:

$$F_{\text{drag}} = 6\pi\eta R v \tag{3.84}$$

where η is the dynamic viscosity of the medium, R is the hydrodynamic radius of the spherical particle and v is the velocity of the particle. Combining this with Einstein's fluctuation dissipation concept yields the **Stokes–Einstein–Sutherland equation**:

$$D = \frac{k_B T}{6\pi\eta R} \tag{3.85}$$

This equation highlights how diffusion is inversely proportional to the viscosity of the medium and the size of the diffusing particle. It plays a fundamental role in characterizing diffusion in soft matter systems, biological fluids, and nanomaterials.

It is important to note that this relation is derived under the assumptions of spherical symmetry. Deviations from these assumptions may require modified or generalized models.

In this study, we use LAMMPS to perform classical molecular dynamics (MD) simulations and track the motion of particles over time.

First, denote by $\mathbf{r}_i(t)$ the position vector of particle i at time t. For any time lag Δt, its displacement is

$$\Delta\mathbf{r}_i(t;\Delta t) = \mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t). \tag{3.86}$$

We choose many starting times t along a total simulation of length t_{tot}, so that the number of intervals is $N_{\Delta t} = t_{\text{tot}} - \Delta t + 1$. Summing the squared displacements over all N particles and averaging over those intervals

yields the total mean-squared displacement (TMSD):

$$\mathrm{TMSD}(\Delta t) = \sum_{i=1}^{N} \frac{1}{N_{\Delta t}} \sum_{t=0}^{t_{\mathrm{tot}}-\Delta t} \left\| \Delta \mathbf{r}_i(t;\Delta t) \right\|^2. \tag{3.87}$$

Dividing by N gives the per-particle MSD,

$$\mathrm{MSD}(\Delta t) = \frac{\mathrm{TMSD}(\Delta t)}{N}. \tag{3.88}$$

In the long-time (diffusive) regime, MSD grows linearly with Δt, so

$$\mathrm{MSD}(\Delta t) \approx 2dD\,\Delta t, \tag{3.89}$$

and we extract the diffusion coefficient D from the slope of MSD vs. Δt. Here d is the dimensionality (e.g. $d = 3$). which quantifies the collective mobility over the interval Δt. During an MD run of total duration t_{tot}, multiple ($N_{\Delta t}$) such intervals of length Δt are available, each starting at a different time t (with $\Delta t < t_{\mathrm{tot}}$). Since particle displacement over Δt reflects mobility, the total mean squared displacement (TMSD) across all mobile particles is given by:

$$\mathrm{TMSD}(\Delta t) = \sum_{i=1}^{N} \left\langle \left[\mathbf{r}_i(\Delta t) - \mathbf{r}_i(0)\right]^2 \right\rangle = \sum_{i=1}^{N} \frac{1}{N_{\Delta t}} \sum_{t=0}^{t_{\mathrm{tot}}-\Delta t} \left[\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)\right]^2, \tag{3.90}$$

where the angle brackets denote time averaging over all $N_{\Delta t}$ segments. This temporal ensemble averaging improves statistical reliability and accuracy in estimating diffusional properties.

The mean squared displacement (MSD) per ion, used to estimate diffusivity, is obtained by normalizing TMSD by the number of mobile particles:

$$\mathrm{MSD}_{per-ion}(\Delta t) = \frac{1}{N}\mathrm{TMSD}(\Delta t), \tag{3.91}$$

where N denotes the number of diffusing particles.

Generally, the MSD exhibits a linear dependence on Δt when sufficient diffusive motion is captured. The atomic diffusivity D is then extracted from the slope of the MSD curve using the Einstein relation:

$$D = \frac{\mathrm{MSD}(\Delta t)}{2d\Delta t} + D_{\mathrm{offset}}, \tag{3.92}$$

where d is the dimensionality of the diffusion process, and D_{offset} accounts for potential systematic deviations or finite size corrections. For completeness, the Nernst Einstein relation links D to ionic conductivity σ by

$$\sigma = \frac{nq^2 D}{k_{\mathrm{B}} T}, \tag{3.93}$$

but we do not compute σ directly in our MD runs. Instead, we report D values and discuss conductivity qualitatively.

Arrhenius Behavior and Transition State Theory in Diffusion

In solid-state diffusion (e.g., vacancy migration in SiC), transition state theory (TST) provides a microscopic foundation for the Arrhenius rate law. Under conditions of thermal equilibrium and classical over barrier hopping, the defect jump rate follows

$$k = \nu \exp\left(-\frac{E_a}{k_B T}\right),$$

where E_a is the activation energy (the energy barrier or migration enthalpy required to form the activated complex at the saddle point) and ν is the Arrhenius prefactor (attempt frequency) [97]. TST assumes a quasi-equilibrium between the initial state and the activated complex, and that, once the system has crossed the saddle point, it proceeds to the new site without recrossing . The prefactor ν reflects the vibrational attempt frequency of the diffusing atom and can be calculated via the harmonic TST (Vineyard) formula as the ratio of the product of normal mode frequencies in the initial state to that in the transition state (excluding the imaginary frequency) [97]. This classical TST picture is widely applicable and enables the prediction of diffusion parameters in solids, for example, first-principles studies of defects in SiC using harmonic TST yield both activation energies and diffusion pre factor in agreement with experimentally measured diffusivities [98].

The temperature dependence of diffusion processes in solids is commonly described by an Arrhenius type expression, which originates from the framework of Transition State Theory (TST). The diffusion coefficient D is then expressed as:

$$D = D_0 \exp\left(-\frac{E_A}{k_B T}\right) \tag{3.94}$$

Here,D is the diffusion coefficient at temperature T,D_0 is the pre exponential (or attempt frequency) factor,E_A is the activation energy barrier for diffusion and k_B is the Boltzmann constant. The logarithmic form of Eq. (3.94) is often used for data fitting:

$$\ln D = \ln D_0 - \frac{E_A}{k_B T} \tag{3.95}$$

In the present study, the diffusion coefficient D is determined from the time evolution of the mean square displacement (MSD), defined in d-dimensional space as:

$$D = \frac{\langle r^2(t) \rangle}{2d\,\Delta t} \tag{3.96}$$

By computing D at multiple temperatures, one can fit the data to the Arrhenius relation (Eq. 3.95) using a least squares regression. This procedure enables the estimation of both the activation energy E_A and the pre exponential factor D_0, providing insight into the thermally activated nature of the diffusion mechanism.

This method has been widely employed in molecular dynamics simulations and experimental studies of diffusion processes [99].

Correction of D_{MSD} for Single-Defect Diffusion

In the standard Einstein relation for diffusion, the mean squared displacement (MSD) is averaged over all diffusing particles in the system. This is appropriate when many particles are mobile and contribute equally to the overall displacement. However, in our case, we consider a single mobile defect embedded in a crystalline lattice of $N-1$ atoms that primarily vibrate around their equilibrium positions.

When using LAMMPS to compute the MSD, the default calculation averages the squared displacements over *all* atoms in the simulation supercell. For our $6 \times 6 \times 6$ cubic supercell ($N = 1728$ atoms in the pristine

case), this means that the contribution of the single mobile defect is divided by N. As a result, the reported MSD is underestimated by a factor of N compared to the true MSD of the defect.

Physically, this underestimation arises because only one atom (the defect species) exhibits long range diffusive motion, while the remaining atoms perform localized thermal vibrations that average to zero over time. The MSD from vibrations adds a small background but does not grow linearly with time, so the defect's contribution is effectively diluted.

Let the MSD reported by LAMMPS be:

$$\mathrm{MSD}_{\mathrm{LAMMPS}} = \frac{1}{N}\sum_{i=1}^{N}\langle \Delta r_i^2 \rangle. \tag{3.97}$$

If only one atom (the defect) contributes significantly to diffusion, then:

$$\left\langle \Delta r^2_{\mathrm{defect}} \right\rangle \approx N \times \mathrm{MSD}_{\mathrm{LAMMPS}}. \tag{3.98}$$

The Einstein relation for the defect then reads:

$$D_{\mathrm{defect}} = \frac{1}{2d}\frac{d}{dt}\langle \Delta r^2_{\mathrm{defect}} \rangle = \frac{1}{2d}\frac{d}{dt}\left[N \times \mathrm{MSD}_{\mathrm{LAMMPS}}\right], \tag{3.99}$$

where d is the dimensionality of the system (e.g., $d = 3$ in our case).

In our simulations, the total number of atoms is $N = 1727$ for the vacancy case (V_C) and $N = 1729$ for the interstitial case (I_C). Without the N-factor correction, the diffusion coefficient obtained from LAMMPS output would be underestimated by this same factor.

It is critical to note that only the *linear* region of the MSD time plot corresponds to the diffusive regime and should be used for fitting. We extract D_{defect} by applying a least squares fit (LSF) to this linear segment.

3.9 Statistical Theory of Atomic Transitions for the Study of Defect Evolution

While the previous derivation of the diffusion coefficient from the mean square displacement (MSD) using the Brownian motion formalism provides a direct link to Fick's second law, it assumes an isotropic and homogeneous medium. In real crystalline solids particularly those with

complex crystal structures, anisotropic environments, and discrete defect sites atomic migration is better described within a statistical kinetic framework.

At this stage, our aim is to derive Fick's laws of diffusion starting from the transition probability formalism, which offers a more general and microscopic perspective. Re deriving Fick's law in this context is important because it explicitly connects the continuum description of diffusion to the underlying stochastic jump processes, ensuring that the macroscopic transport equation naturally emerges from the atomistic picture. This is particularly valuable for systems where anisotropy, complex lattice geometries, or correlated jumps can modify the effective transport behavior.

Within the transition probability method, the diffusion coefficient D is shown to be proportional to the mean square atomic displacement per unit time. For a defect undergoing random jumps between equivalent lattice sites, this displacement can be directly expressed in terms of the jump frequency. This leads naturally to the so called *frequency rule*, in which the diffusion coefficient is determined directly from the atomic jump rate.

The treatment presented here is restricted to relatively simple model systems, but the methodology is sufficiently general to illustrate the statistical mechanical principles underlying defect diffusion in both isotropic and anisotropic crystalline environments.

3.9.1 Derivation of Fick's Laws with Isotropic Transition Probability

Let us consider a volume element d^3r (units: m^3) whose position vector is $\mathbf{r}$ (units: m) at the center of any medium. We want to estimate diffusion from this volume element, represented by the number of particles passing through d^3r per unit time. We define the *conditional transition probability density per unit time* $\Lambda(\mathbf{r} \mid \mathbf{R})$. By definition,

$$\Lambda(\mathbf{r} \mid \mathbf{R})\, d^3R\, dt$$

is the **probability** (dimensionless) that a particle located at $\mathbf{r}$ will travel by a displacement vector $\mathbf{R}$ (units: m) into a displacement space volume

element d^3R (units: m^3) in a time interval dt (units: s). Because this probability must be dimensionless, the units of Λ are:

$$[\Lambda] = \mathrm{m}^{-3}\,\mathrm{s}^{-1}.$$

Let $C(\mathbf{r}, t)$ denote the particle concentration (units: m^{-3}). The number of particles in d^3r is

$$C(\mathbf{r}, t)\, d^3r \quad \text{(units: particles).}$$

The number of particles **leaving** d^3r in time dt is obtained by multiplying the number of particles in d^3r by the probability of moving in any displacement direction. Mathematically,

$$\int_{\mathbb{R}^3} C(\mathbf{r}, t)\, d^3r\, \Lambda(\mathbf{r} \mid \mathbf{R})\, d^3R\, dt \tag{3.100}$$

Here:

$$[C\, d^3r] = \text{particles}, \quad [\Lambda\, d^3R\, dt] = \text{dimensionless probability.}$$

Now, to estimate the number of particles **entering** d^3r, consider particles located at $\mathbf{r} - \mathbf{R}$ that jump by $\mathbf{R}$ into $\mathbf{r}$. The number of such particles coming from a displacement volume d^3R in time dt is

$$C(\mathbf{r} - \mathbf{R}, t)\, d^3r\, \Lambda(\mathbf{r} - \mathbf{R} \mid \mathbf{R})\, d^3R\, dt \tag{3.101}$$

Integrating equation (3.101) over all $\mathbf{R}$ gives the total number of particles entering d^3r per time interval dt.

$$\int_{\mathbb{R}} C(r - R, t)\Lambda(r - R|R) dr dR dt \tag{3.102}$$

The difference between 3.102 and 3.100 should give the added number of particles in dr in time dt, since the difference is nothing else the difference of particles moving out of dr and number of particles moving into dr. Dividing the difference by dr and dt gives the rate of increase in particle concentration at position r and time t. Mathematically this is given as,

$$\frac{\partial C}{\partial t} = \int_{\mathbb{R}} [C(r - R, t)\Lambda(r - R|R) - C(r, t)\Lambda(r|R)] dR \tag{3.103}$$

Considering that transition probability is quite smaller for big R in comparison to the diffusion distance. This would be the case where diffusoin jumps are of the order of atomic dimensions. Similarly its true for the liquids and gases where R is equivalent to mean free path. We can drive Fick's law by implying 3.103 where we apply Taylor series to expand first two terms and retain it up to the second order. We will get following,

$$C(r-R,t)\Lambda(r-R) = C(r,t)\Lambda(r|R) - \Lambda(r|R)\mathbf{R}\cdot\nabla C + \frac{1}{2}\Lambda(r|R)(\mathbf{R}\cdot\nabla)^2 C \tag{3.104}$$

Let's consider cases where transition probability is independent of position vector r we can substitue 3.104 into 3.103 and obtain

$$\frac{\partial C}{\partial t} = -\int_{\mathbb{R}} \Lambda(r|R)\mathbf{R}\cdot\nabla C dR + \frac{1}{2}\int_{\mathbb{R}} \Lambda(r|R)(\mathbf{R}\cdot\nabla)^2 C dR \tag{3.105}$$

We can ignore the first integral in 3.105 because, concentration are generally measured at r and they are independent of R. Moreover, we have already assumed transitional probabilities as independent of position and we can take them same for R and $-R$ thus killing the first integral.

$$\frac{\partial C}{\partial t} = \frac{1}{2}\int_{\mathbb{R}} \Lambda(r|R)(\mathbf{R}\cdot\nabla)^2 C dR \tag{3.106}$$

Let's assume the components of $\mathbf{R}$ and r in a Cartesian coordinate system is given by (X_1, X_2, X_3) and (x_1, x_2, x_3), respectively. Then

$$(\mathbf{R}\cdot\nabla)^2 C = \frac{1}{2}\sum_{i,j=1}^{3} X_i X_j \frac{\partial^2 C}{\partial x_i \partial x_j} \tag{3.107}$$

We can rewrite equation (3.106) as

$$\frac{\partial C}{\partial t} = \frac{1}{2}\sum_{i,j=1}^{3} \overline{X_i X_j}\,\frac{\partial^2 C}{\partial x_i\,\partial x_j} \tag{3.108}$$

where

$$\overline{X_i X_j} = \int_{\mathbb{R}^3} X_i X_j\,\Lambda(\mathbf{r}\mid\mathbf{R})\,d^3R \tag{3.109}$$

Units of each term:

- X_i, X_j are components of the displacement vector $\mathbf{R}$, each with units m.
- $\Lambda(\mathbf{r} \mid \mathbf{R})$ has units $\mathrm{m}^{-3}\,\mathrm{s}^{-1}$.
- d^3R has units m^3.

Therefore, in equation (3.109):

$$[X_i X_j] \cdot [\Lambda] \cdot [d^3R] = \mathrm{m}^2 \cdot \mathrm{m}^{-3}\,\mathrm{s}^{-1} \cdot \mathrm{m}^3 = \mathrm{m}^2\,\mathrm{s}^{-1}.$$

Thus $\overline{X_i X_j}$ represents the **mean quadratic displacement per unit time**, with the correct physical units of m^2/s.

Equation (3.108) is therefore dimensionally consistent:

$$\left[\frac{\partial C}{\partial t}\right] = \mathrm{m}^{-3}\,\mathrm{s}^{-1}, \quad [\overline{X_i X_j}] \cdot \left[\frac{\partial^2 C}{\partial x_i \partial x_j}\right] = \mathrm{m}^2/\mathrm{s} \cdot \mathrm{m}^{-5} = \mathrm{m}^{-3}\,\mathrm{s}^{-1}.$$

By comparing with equation (3.109), we identify the components of the diffusion tensor:

$$D_{rs} = \frac{1}{2}\,\overline{X_r X_s} \tag{3.110}$$

Since $\overline{X_r X_s}$ has units m^2/s, it follows immediately that $[D_{rs}] = \mathrm{m}^2/\mathrm{s}$, the standard units of a diffusion coefficient.

Remark: The factor 1/2 in equation (3.110) is dimensionless. The numerical prefactor does not alter the units, it only rescales the physical magnitude. Thus, D_{rs} is *not* a dimensionless number it is a transport coefficient with physical units m^2/s, directly proportional to the mean square displacement of particles per unit time.

3.9.2 Relationship Between Diffusion Coefficient and Mean Square Displacement

If one only choose the principal axes of diffusion as coordinate axes, then only the diagonal terms in 3.110 are nonzero and equation 3.108 takes the following form.

$$\frac{\partial C}{\partial t} = \frac{1}{2}\sum_{i=1}^{3}\overline{X_i^2}\frac{\partial^2 C}{\partial x_i^2} \tag{3.111}$$

and the principle diffusion coefficients will be

$$D_i = \frac{1}{2}\overline{X_i^2} \tag{3.112}$$

Thus, 3.111 establishes a fundamental connection between the diffusion coefficient and the mean square displacement per unit time. From this study we can also develop the understanding of microscopic diffusion coefficient in terms of atomic mechanisms.

If all D_i are equal (e.g. in isotropic crystals), 3.111 becomes

$$\frac{\partial C}{\partial t} = \overline{X^2}\nabla^2 C \tag{3.113}$$

Here, $\overline{X^2} = \overline{X_1^2} = \overline{X_2^2} = \overline{X_3^2}$. And the diffusion coefficient can be expressed in terms of $\overline{R^2}$ instead of $\overline{X^2}$. This is because,

$$\overline{R^2} = \overline{X_1^2} + \overline{X_2^2} + \overline{X_3^2} \tag{3.114}$$

For cubic medium, since they are isotropic we have,

$$\overline{R^2} = 3\overline{X^2} \tag{3.115}$$

and equation 3.113 can be rewritten as,

$$\frac{\partial C}{\partial t} = D\nabla^2 C \tag{3.116}$$

where

$$D = \frac{1}{6}\overline{R^2} \tag{3.117}$$

Comparison of 3.116 to the equation of continuity gives the The statistical meanings of Fick's first law. So for cubic case we immediately obtain,

$$J = -D\nabla C \tag{3.118}$$

with D given by 3.117.

An important point to recall is the very **strong assumption** that is made when deriving these results. It tells that conditions probability is anyways isotropic and it is independent of particle positions. Without this assumption, we may not be able to derive Fick's laws. So the relevant physical implications must be considered. An isotropic transition probability forces the elimination of any external fields and temperature gradients. Moreover, the transition probability is also unaffected by concentration changes.

3.9.3 Frequency rule for Diffusion Coefficient

The diffusion coefficient D relates the mean square particle displacement per unit time to the diffusion process. In isotropic or cubic media, the diffusion coefficient can be expressed in terms of the mean square displacement:

$$D = \frac{1}{6}\overline{R^2}$$

Atomic diffusion in crystalline systems takes place as discrete atomic hope from site to site unlike fluid flow in liquid phase. An example could be a vacancy or interstitial defect in silicon carbide crystal. Generally in such system the diffusing species spend most of the time executing small vibrations across the mean position until momentum added by local thermal fluctuations surpasses the respective migration barrier. This result in diffusion jump to the first nearest neighbor thus defect adopts a tortuous path inside crystal consisting of a large number of random jumps. The studied vacancy diffusion in our case, had close relevance with the explained phenomenon. For carbon vacancy diffusion carbon atom diffuse to the nearest carbon cite or second nearest neighbor traveling a diffusion jump of 3.08 angstrom. For the interstitial case the Diffusion distance is even shorter. For all atomic diffusion the length of total displacement R is the sum of all elementary jump distances r that are of the same lengths. Thus **R** would be,

$$\mathbf{R} = \sum_{i=1}^{\Gamma} \mathbf{r}_i \tag{3.119}$$

Here $\mathbf{r}_i$ is the elementary jump distance for the jump i while Γ is the jump frequency.

Since diffusion coefficient D is proportional to r^2 and We have to square equation 3.119 to satisfy this condition.

$$\mathbf{R}^2 = \sum_{i,j=1}^{\Gamma} \mathbf{r}_i \cdot \mathbf{r}_j \tag{3.120}$$

Further simplification is performed by pairing out the similar terms or the terms with $i = j$:

$$\mathbf{R}^2 = \sum_{i=1}^{\Gamma} \mathbf{r}_i^2 + \sum_{i=j} \mathbf{r}_i \cdot \mathbf{r}_j \tag{3.121}$$

We will rewrite the second term on the right as partial sums.

$$\sum_{i \neq j} \mathbf{r}_i \cdot \mathbf{r}_j = 2 \sum_{i=1}^{\Gamma-1} \mathbf{r}_i \cdot \mathbf{r}_{i+1} + 2 \sum_{i=1}^{\Gamma-2} \mathbf{r}_i \cdot \mathbf{r}_{i+2} + \dots \tag{3.122}$$

The factor of 2 came since the sum is taken on both i and j seperately. Substituting 3.122 in 3.121 we get .

$$\mathbf{R}^2 = \sum_{i=1}^{\Gamma} \mathbf{r}_i^2 + 2 \sum_{i=1}^{\Gamma-1} \sum_{j=i+1}^{\Gamma} \mathbf{r}_i \cdot \mathbf{r}_j \tag{3.123}$$

The expression is further simplified by restricting our approach to the cubic monoatomic crystals. For such crystals all the jump vectors have same length thus converting first term into r^2 times Γ. The second term can makes the following form.

$$\mathbf{r}_i \cdot \mathbf{r}_{i+j} = r^2 \cos \theta_{i,i+j} \tag{3.124}$$

Here $\theta_{i,i+j}$ is the angle between mentioned ith and the $(i+j)$th jump. Equation 3.123 will take the following form.

$$\mathbf{R}^2 = \Gamma r^2 + 2r^2 \sum_{j=1}^{\Gamma-1} \sum_{i=1}^{\Gamma-j} \cos \theta_{i,i+j} \tag{3.125}$$

Here R defines the magnitude of jump displacement for a single atom in a unit time. To get the value of mean square displacement for transition probability theory, it's required to make an average over many atoms that are doing Γ jumps in unit time. Thus

$$\overline{\mathbf{R}^2} = \Gamma r^2 f \tag{3.126}$$

where

$$f = 1 + \frac{2}{\Gamma} \sum_{j=1}^{\Gamma-1} \sum_{i=1}^{\Gamma-j} \cos\theta_{i,i+j} \tag{3.127}$$

Here f is the correlation factor. It is worth mentioning that this kind of treatment is solely for our special case of solid state diffusion. Moreover f can be given as:

$$f = \frac{1+q}{1-q} = \frac{1+\overline{\cos\theta}}{1-\overline{\cos\theta}} \tag{3.128}$$

As can be seen correlation factor is linked with crystal structure thus strongly correlated with diffusion mechanism. Estimation of f is difficult in general but for simple diffusion mechanisms as our case it is constantly unity. For example in the case of interstitial diffusion in cubic crystal environment, clearly $\overline{\cos\theta} = 0$. The reason lie in the fact that Every possible jump in a given direction will be countered by another jump in the opposite direction thus $f = 1$. This kind of diffusion is defined as uncorrelated random flight of impurity that is interstitial in present case. One key understanding from this aspect is that sequential jumps by the impurity atoms have no physical correlation.

$$\mathbf{D} = \mathbf{f_d} \frac{\mathbf{\Gamma \Delta^2}}{6} \tag{3.129}$$

This is a simplified approximation for calculation of D in the particular case of point defects. As mentioned before $\mathbf{\Gamma}$ is the jump frequency while $\mathbf{\Delta}$ signifies the diffusion jump distance. Our approach yields distinct advantages by focusing exclusively on the study of diffusing species at varying temperatures. Thus, enabling the precise calculation of the diffusion coefficient(D) .

Capitolo 4

Configurations and Energetics of Point Defects in 3C-SiC with Optimization Methods

This chapter presents a systematic study aimed at validating the computational methodology and establishing initial benchmarks for 3C-SiC, primarily within static and quasi-static simulation frameworks. The overarching goal is to define a robust and reliable computational protocol that will serve as the foundation for subsequent dynamic simulations and defect evolution studies.

We begin by reproducing selected reference results from the literature, including lattice constants, elastic constants, melting points, and defect formation energies, using molecular dynamics simulations with Tersoff and Environment Dependent Interatomic Potentials (EDIP) implemented in LAMMPS. The comparison of these results allows for the critical assessment of interatomic potentials, ultimately leading to the selection of EDIP for subsequent defect energetics investigations.

Following the potential selection, we investigate migration energy barriers for key single-vacancy-type defects carbon vacancy (V_{C}), silicon vacancy (V_{Si}), carbon antisite–vacancy complex ($C_{\mathrm{Si}}V_{\mathrm{C}}$), and silicon antisite vacancy complex ($Si_{\mathrm{C}}V_{\mathrm{Si}}$) using the Nudged Elastic Band (NEB) method. In particular, we employ the NEB implementation available in the

LAMMPS computational package [100, 101, 102, 103]. For implementation details, see the LAMMPS NEB documentation at `https://docs.lammps.org/neb.html`.

4.1 Computational Methodology and benchmarking

The selection of an appropriate interatomic potential is a critical first step in molecular dynamics studies, as it directly affects the accuracy of simulated material properties. Each potential embodies a specific functional form and set of empirical parameters, which must be evaluated for the material under investigation. For silicon carbide (SiC), several interatomic potential models are available , including Tersoff [72],[73],[79], Vashishta [104], EDIP [80],[75] and various others. To assess the validity and reliability of these potential models, we embarked on a bench marking process, involving the reproduction of results from prior studies as provided by the respective authors. Furthermore, a validation was undertaken by comparing these potentials results with results obtained from *ab initio* calculations and experimental investigations. The initial stage of this convergence test involved the determination of structural parameters and the calculation of defect formation energies in three key poly types of silicon carbide, namely 3C-SiC, 4H-SiC, and 6H-SiC. The study was then extended to reproduce migration barriers for key defect complexes and melting point for cubic silicon carbide.

Based on the prescribed theoretical framework in chapter 3, formation energies were calculated following full geometry optimization as implemented in LAMMPS. A conjugate gradient minimization algorithm was employed, with box relaxation performed prior to atomic position minimization. The convergence criterion for energy minimization was set to 10^{-7} eV.

Calculations were carried out for a series of point defects, including silicon vacancies (V_{Si}), carbon vacancies (V_{C}), and carbon antisites (C_{Si}). Each analysis was repeated using three interatomic potentials: the Environment-Dependent Interatomic Potential (EDIP) [80, 75], the Tersoff potential

[79, 72], and the Tersoff potential with the Ziegler–Biersack–Littmark (ZBL) correction [105, 106].

Once defect structures were validated, migration pathways were investigated using the Nudged Elastic Band (NEB) method. The initial NEB calculations employed ten intermediate replicas generated from DFT optimized initial and final geometries, without considering the displacement of neighboring atoms. The procedure was then refined in three correction stages (cor1–cor3) by: (1) increasing the number of replicas, (2) tightening the minimization convergence criteria on the extreme geometries, and (3) considering the neighbor atom displacement while generating replicas to avoid nonphysical configuration.

For the final NEB results, we employed 20 intermediate images to model migration pathways. The first nearest neighbor jump (2.0 Å) was considered for $C_{\mathrm{Si}}V_{\mathrm{C}}$ and $Si_{\mathrm{C}}V_{\mathrm{Si}}$ defects, while the second nearest neighbor jump (3.4 Å) was studied for V_{C} and V_{Si} defects. The initial and final states were energy minimized to a tolerance of 10^{-5} eV prior to interpolation.

To account for the influence of neighboring atoms during pathway generation, interpolation was applied not only to the diffusing atom but also to its nearest neighbors. NEB replicas are connected via a spring potential that constraints them aginest deviation from the transition path. With the fact that spring constant is an external imposed potential one needs to optimize it prior to implementation to ensure numerical stability and physical accuracy. A systematic study of the harmonic spring constant revealed that a value of 3 eV provided stable convergence with negligible effect on the calculated migration barriers (see Figure C.1 and C.2).

4.2 Elastic Properties of Ideal and Defective Configurations

We first analyzed the structural characteristics of ideal SiC polytypes by calculating bond lengths, bond angles, and lattice parameters using the EDIP and Tersoff interatomic potentials. In all cases, the Si–C bond length remains approximately 1.9 Å, while the tetrahedral bond angle is close to

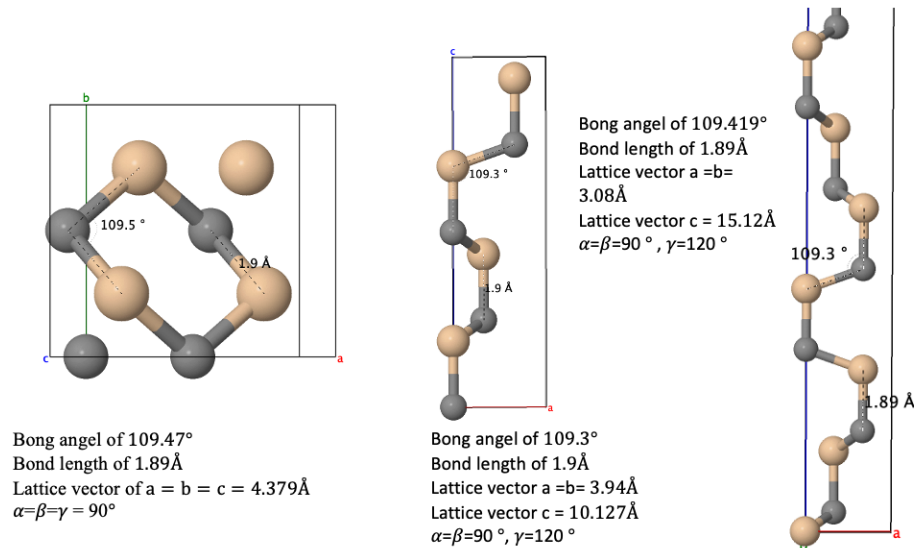


Figura 4.1: Optimized geometric parameters of SiC polytypes. In all cases, the Si–C bond length is 1.89 Å and the tetrahedral bond angle is 109.4°. For **3C-SiC** (cubic), the lattice constant is $a = 4.38$ Å. For **4H-SiC** (hexagonal), the lattice parameters are $a = b = 3.94$ Å, $c = 10.127$ Å, with $\alpha = \beta = 90°, \gamma = 120°$. For **6H-SiC** (hexagonal), the lattice parameters are $a = b = 3.08$ Å, $c = 15.12$ Å, with $\alpha = \beta = 90°, \gamma = 120°$.

109.5°. However, the lattice constants differ between polytypes due to their distinct crystal systems:

- **3C-SiC (cubic, zincblende)**: Lattice constant $a \approx 4.379$ Å.
- **4H-SiC (hexagonal)**: two lattice constants, $a \approx 3.07$ Å and $c \approx 10.127$ Å.
- **6H-SiC (hexagonal)**: two lattice constants, $a \approx 3.08$ Å and $c \approx 15.12$ Å.

The differences in c reflect the stacking sequence of Si–C bilayers along the [0001] axis. A schematic comparison of the 3C, 4H, and 6H unit cells is shown in Fig. 4.1.

4.3 Formation Energies of Point Defects

The defect formation energy (DFE) can be precisely defined as the aggregate energy requirement for introducing a specific type of defect into a given crystal structure. In the context of classical molecular dynamics, DFEs for vacancies and interstitials can be calculated using the following expressions:

$$E^{\text{form}}_{\text{vacancy}} = E_{\text{host + vacancy}} - \left(E_{\text{host}} - E_{\text{atom removed}}\right), \tag{4.1}$$

$$E^{\text{form}}_{\text{interstitial}} = E_{\text{host + interstitial}} - \left(E_{\text{host}} + E_{\text{atom added}}\right). \tag{4.2}$$

Validation of Methodology Before investigating complex defects, our methodology was validated by reproducing key results from the original parameterizations of the interatomic potentials. Table 4.1 compares stoichiometric defect complex energies, lattice constants, and elastic constants obtained with the Tersoff potential against the values reported in the original study. The close agreement underscores the consistency of our implementation.

Stoichiometric Defect Complexes in 3C-SiC (eV)		
	Tersoff [72]	Our results
$V_{Si} + V_C$	7.4	7.4
$C_{Si} + Si_C$	7.2	6.76
Lattice constant (Å) and Elastic constants (Mbar)		
	Tersoff [72]	Our results
a	4.32	4.32
c_{11}	4.2	4.26
c_{12}	1.2	1.10
c_{44}	2.6	2.53

Tabella 4.1: Validation of the Tersoff potential in 3C-SiC by reproducing defect complex formation energies, lattice constant, and elastic constants reported in the original publication.

To strengthen this validation, Table 4.2 compares elastic constants and moduli across multiple empirical potentials (Tersoff, EDIP, Vashishta, and

Tersoff/ZBL) and ab initio data. The agreement with literature values confirms that both Tersoff and EDIP are suitable candidates for defect studies in SiC.

	Tersoff	EDIP	Vashishta	Tersoff/ZBL	Ab-initio [107]
Bulk Modulus (GPa)	214.98	225.87	225.10	204.48	228.48
Shear Modulus 1 (GPa)	252.54	170.47	136.93	207.81	189.00
Shear Modulus 2 (GPa)	157.98	128.25	123.75	149.92	187.00
Poisson's ratio	0.20	0.26	0.26	0.21	0.13
c_{11}	425.62	396.89	390.11	404.39	420.00
c_{12}	109.67	140.36	142.61	104.54	132.00
c_{44}	252.54	170.47	136.94	207.81	267.00

Tabella 4.2: Comparison of elastic constants and moduli of 3C-SiC obtained using empirical potentials and ab-initio data. The close agreement validates the quality of the potentials.

Benchmarking Defect Energetics in 3C-SiC Having validated the methodology, we next computed the formation energies of key point defects in 3C-SiC using Tersoff and EDIP potentials. Table 4.3 summarizes these results, covering single vacancies, antisite-vacancy complexes, divacancies, and diinterstitials.

Tabella 4.3: Formation energies (E_{form}) and corresponding optimized supercell energies ($E_{\text{optimized SC}}$) for point defects in 3C-SiC, calculated with EDIP and Tersoff potentials. Bold values highlight the highest formation energy within each set. All energies are given in eV.

Defect Type	EDIP Potential		Tersoff Potential	
	E_{form} (eV)	$E_{\text{optimized SC}}$	E_{form} (eV)	$E_{\text{optimized SC}}$
V_C	9.27	-10942.12	11.44	-10637.76
V_{Si}	5.33	-10942.11	11.99	-10641.25
$C_{Si}V_C$	3.95	-10943.50	13.86	-10639.39
$Si_C V_{Si}$	—	—	—	—
VV ($V_C + V_{Si}$)	**20.98**	-10934.31	**21.68**	-10631.58
II ($I_C + I_C$)	**2.92**	-10954.31	**2.92**	-10954.31

Extension to Polytypes To assess the transferability of the potentials, the study was extended to other SiC polytypes. Table 4.4 reports the formation energies of carbon and silicon vacancies in 3C-, 4H-, and 6H-SiC using Tersoff, Tersoff/ZBL, and EDIP potentials. The results show consistent trends across polytypes, while highlighting systematic shifts depending on the choice of potential.

Polytype	Defect Type	Tersoff (eV)	Tersoff/ZBL (eV)	EDIP (eV)
3C-SiC	V_C	3.296	8.239	4.961
	V_{Si}	4.051	4.523	1.869
4H-SiC	V_C	3.307	8.239	4.603
	V_{Si}	4.056	4.531	1.873
6H-SiC	V_C	3.316	8.236	4.606
	V_{Si}	4.059	4.536	1.879

Tabella 4.4: Defect formation energies (eV) for V_C and V_{Si} in the three main polytypes of SiC, calculated using Tersoff, Tersoff/ZBL, and EDIP potentials. Results are grouped by polytype for clarity.

4.4 Study of melting point for 3C-SiC and validation of interatomic potentials

The provided analysis holds true under the condition that the system remains in an equilibrium state. This consideration demands a study of melting point behavior to confine the prescribed analysis to a limit far below the melting point and avoid the phase change complexities. To perform this test, MD calculations were conducted using the NPT ensemble and considering the Tersoff along with EDIP potential. Calculated results for EDIP can be examined in Figure 4.2.This figure presents an overview of the solid to liquid phase transition (melting) for SiC, characterized by a sharp expansion in volume. The transition is captured by 20 ns MD simulations under an NPT ensemble at 1 atm, employing the EDIP empirical potential. One can observe a sudden expansion in volume at 2620K (after waiting 20ns), indicating a phase transition from solid to liquid, thus setting a limit for stopping analysis before 2600K. Moreover, the calculated melting point aligns well with experimental results [78], unlike the Tersoff value, which estimated a melting point above 3500K. This also establishes the validity of EDIP over Tersoff and supports its selection particularly for diffusion studies. Our results are consistent with a recent study of Galli[1] et. al who performed a complete benchmark of potentials' predictive features by studying the density of SiC with a variety of potentials and a temperature range from 0 to 3500K. By means of comparisons with experimental data they pointed out the EDIP is the optimal choice for defects evolution study in SiC, especially when high temperatures are considered. An example of their results are presented in figure 4.3. Results of the Galli's team were reproduced successfully by us for further confirmation of the methodology reliability. The agreement can be demonstrated comparing figure 4.4, obtained by our simulations, and figure 4.3 of ref. [1].

4.5 Migration Energy Barriers study using NEB

The Nudged Elastic Band (NEB) approach, formally introduced in Sec. 3.6, is regarded as a state of the art method for estimating the energetics asso-

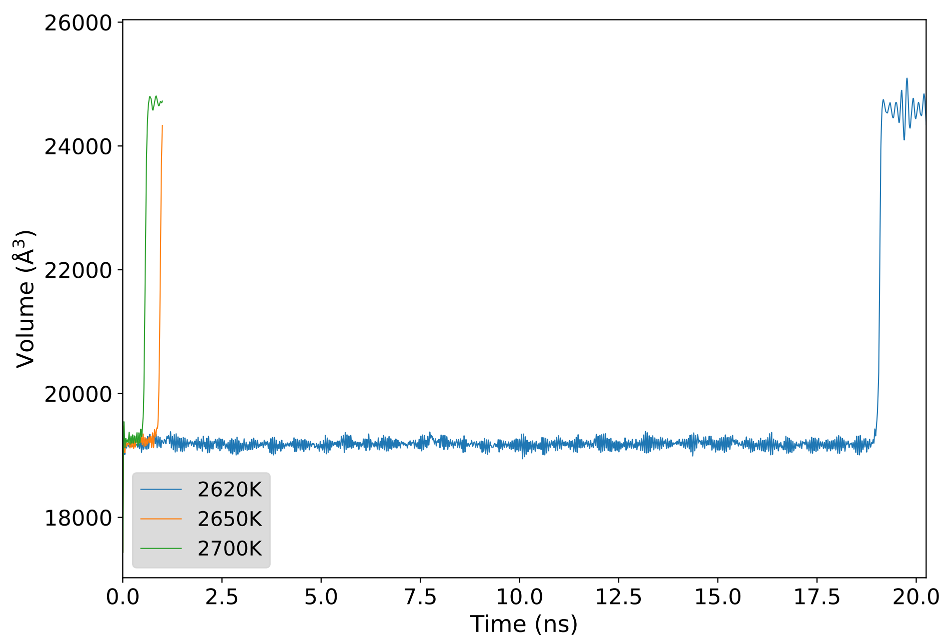


Figura 4.2: Volume of 3C-SiC as a function of temperature during 20 ns molecular dynamics simulations in the NPT ensemble at 1 atm, using the EDIP empirical potential. The data show a sharp increase in volume at ∼2650 K, indicating the solid liquid phase transition. Two regimes are visible: rapid melting within < 2 ns above the transition temperature, and slower melting over ∼20 ns below it. For comparison, melting point estimates from density functional theory (2620 K) and experiments (2818 K) are indicated.

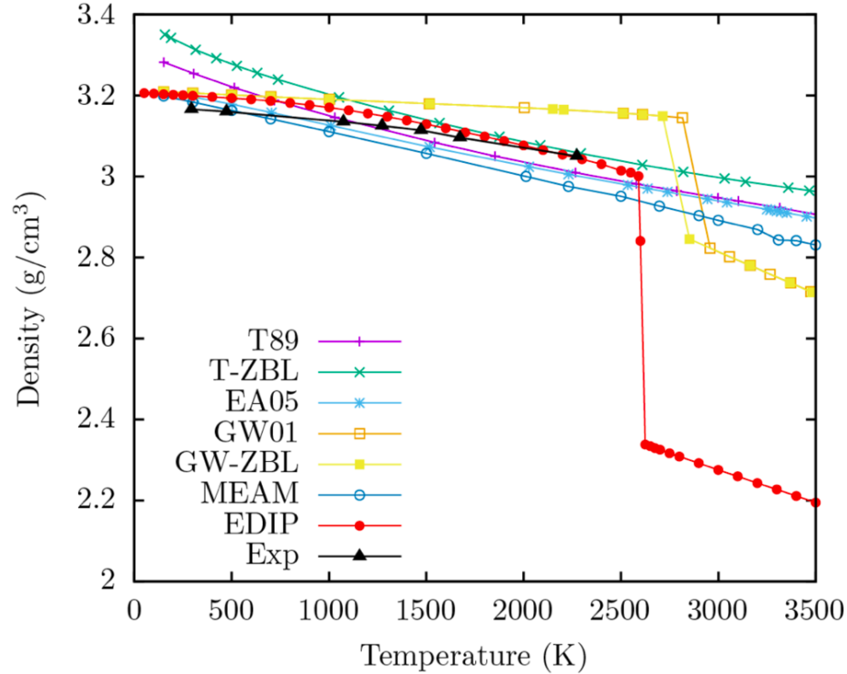


Figura 4.3: Empirical inter atomic potential benchmark study figure reused from nature communication [1]: Density study for 3C-SiC at a given temperature was performed from a 1 ns long classical MD simulation in the NPT ensemble at 1 atm using 4096 atom cells. A bench of empirical force fields were employed : "T89" (in purple) for the Tersoff potential [72, 73], "T-ZBL" (in green) for the Tersoff potential with the Ziegler-Biersack-Littmark (ZBL) screened nuclear repulsion [105], "EA05" (in light blue) for the Erhart-Albe potential [77], "GA01" (in orange) for the Gao-Weber potential [76], "GW-ZBL" (in yellow) for the Gao-Weber potential with the ZBL screened nuclear repulsion [108], "MEAM" (in dark blue) for the modified embedded atom method [74], and "EDIP" (in red) for the environment-dependent interatomic potential [75], while black line correspond to experimental work from [109]. EDIP was picked for the study form all the potentials considering its reasonable agreement with experimental data for temperature-dependent density up to 2200 K and the Melting temperature, e.g., T_m 2620 K using EDIP and 2818 K from experiments [78].

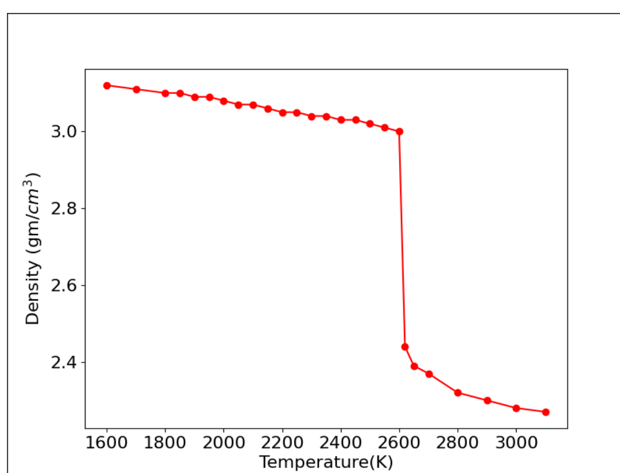


Figura 4.4: This figure illustrates the study of decomposition temperature as determined by classical molecular dynamics simulations using the EDIP [75] parameterization and an NPT ensemble with a 1728-atom supercell. The density decreases with increasing temperature up to 2620 K, where a sudden drop in density, associated with decomposition, is observed.

ciated with metastable states and saddle point configurations in physical and chemical processes. It enables tracing a reaction pathway and calculating the minimum energy configurational profile along that path. In this work, we applied the NEB method to determine the energy barriers required for the formation of defect structures. Several representative cases are discussed individually. For benchmarking purposes, the barriers were calculated using two empirical interatomic potentials, namely Tersoff and EDIP, as previously defined. The analysis is presented for both potentials.
In the following subsections, we present the results of Nudged Elastic Band (NEB) calculations for selected defect configurations in 3C-SiC, focusing on the carbon antisite vacancy ($C_{Si}V_C$), silicon antisite vacancy ($Si_C V_{Si}$), silicon vacancy (V_{Si}), and carbon vacancy (V_C). These defects were chosen due to their relevance for the formation of spin-bearing defect complexes in implantation processes. For each defect type, both EDIP and Tersoff interatomic potentials were employed to evaluate migration barriers, to enable comparison with values reported in the literature. The results highlight key challenges in accurately capturing defect energetics, including discrepancies between potentials, unexpected barrier heights, and their sensitivity to local relaxation effects.

Carbon Antisite Vacancy ($C_{Si}V_C$)

Carbon Anti-site Vacancy (CAV) can be described as a composed defect made by a carbon atom in a silicon-like lattice location C_{Si} (anti-site defect) and a carbon vacancy V_C as the nearest neighbor. A silicon vacancy can trigger the formation of this defect trough the displacement of one of four C atoms present in its nearest neighbor shell, $V_{Si} \rightarrow C_{Si}V_C$. The calculated barrier from both potentials is shown in figure C.3. Carbon antisite is a stable configuration it is even more stable than silicon vacancy [1].The NEB results calculated from LAMMPS pertaining to the carbon antisite exhibit certain ambiguities. It is difficult to accept that two distinct configurations (initial and final)would yield precisely the same energy. Furthermore, the observation of an exceptionally high energy barrier of 9.6 eV, in contrast to the established literature value of approximately 1.4 eV [1], is perplexing. Rodrick Defo calculated a forward energy

barrier of 2.5 eV and reverse energy barrier of 3.5 eV [58].To address this issue, various corrections were systematically explored to identify and rectify potential calculation errors originating from the user's side. These corrections, denoted as corr1, corr2, and corr3 in figure C.3 as presented in Appendix , were applied and examined. None of these corrections produced a significant alteration in the obtained results.

Silicon Antisite Vacancy($Si_C V_{Si}$)

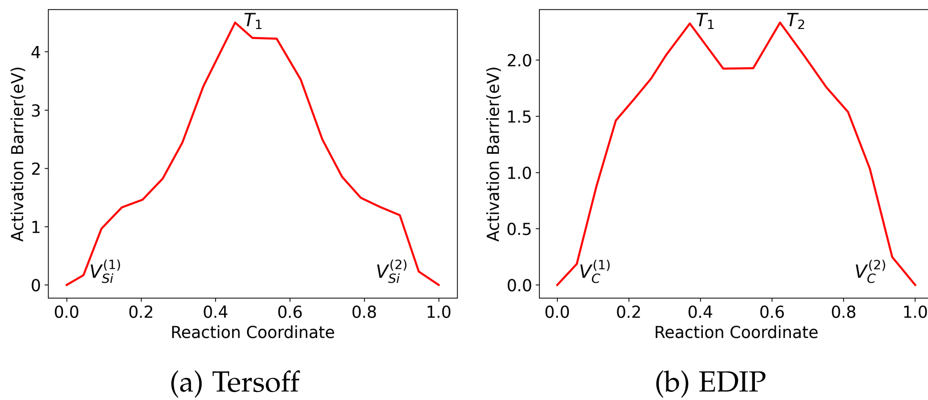
In line with our investigation of carbon antisite vacancy, a similar effort was made to analyze silicon antisite. Silicon antisite is the formation of silicon antisite with silicon vacancy as nearest neighbor (SAV defect). Just as in the case of carbon antisite, we exercised the same precautions due to the inherent calculation challenges. However, the results obtained proved to be notably ambiguous, rendering it challenging to draw meaningful conclusions. One can't expect lower energy configurations in the NEB path as produce by Tersoff. EDIP exhibited very unstable final configuration. Furthermore, there is a dearth of support within the existing literature regarding the difficulties in the formation of silicon antisite with a silicon vacancy as suggested by Bokstedte et.al [54]. The origin of this difficulty lies in the size effect. Both silicon and silicon antisite undergo massive outward relaxation. In the case of the antisite, this relaxation is followed by the optimization of the Si-Si bond to the neighboring atoms, striving to achieve the ideal Si-Si value. This results in strained bonds. Moreover, the vacancy provides additional space, allowing the antisite to relax more towards it. It is possible that the energy of the dangling bond towards the carbon vacancy is reduced by its interaction with other carbon atoms, which might result in a diminishing barrier towards vacancy [110] as depicted in Figure C.4b. These circumstances have limited our ability to make definitive assertions regarding the evolution of silicon antisite. The NEB results for silicon antisite using Tersoff and EDIP are given in figure C.4.

Carbon vacancy V_C

The migration of a carbon vacancy involves interactions with next nearest neighbors, i.e. the second neighbors shell where two equivalent V_C configurations can be present in the 3C-SiC lattice, while we already presented in 4.5 the NEB analysis of first nearest neighbor displacements, which leads to the formation of SAV defect. We determined a migration (diffusion) barrier of 4.5 eV using the Tersoff potential and a barrier of 2.5 eV using the EDIP potential. Giulia Galli et. al [1] arrived at the same value of 2.5 eV using the classical MD and the same EDIP potential and further conducted first principles molecular dynamics simulations to refine the calculation, obtaining a value of 3.9 eV [1]. The barrier calculated using the Tersoff potential is presented in figure 4.5 and aligns well with the findings in literature[17]. Unlike carbon antisite vacancy and silicon antisite vacancy, carbon vacancy produced reliable results showing reliable literature correspondence. The reason why vacancy diffusion barriers in our runs are roughly symmetric while antisite (Si↔C) barriers are not is not yet fully clear. Several plausible explanations can be given: antisite motion changes the local chemistry and bonding environment (an atom of a different species sits on a lattice site), which breaks the simple inversion symmetry of a vacancy hop and can produce intrinsically asymmetric energy landscapes, empirical force fields may treat cross species bonds and charge transfer poorly, giving asymmetric energetics that are an artifact of the potential, and practical setup issues (bad image interpolation) can further distort the computed profile. Because these possibilities have very different implications, the problem should be treated as unresolved until checked.g., by manually interpolating replicas, including more local atoms in the NEB group, testing different potentials, or validating select points with DFT.

Silicon vacancy V_{Si}

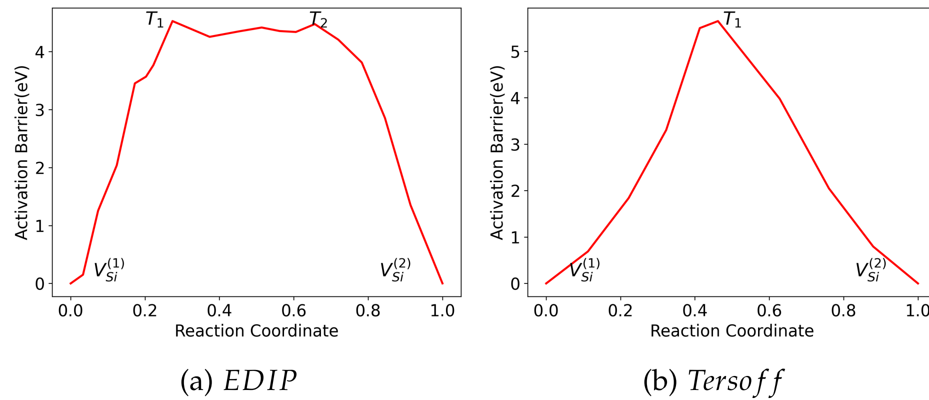
Similarly to the V_C silicon vacancy next nearest neighbor displacements or $V_{Si} \rightarrow V_{Si}$ could allow effective migration, while next neighbor displacements is constrained by the antisite transformation given as $V_{Si} \rightarrow C_{Si}V_C$. It is important to note that the silicon vacancy represents a meta stable configuration in the latter transformation process, ultimately leading



(a) Tersoff

(b) EDIP

Figura 4.5: Energy profile for carbon vacancy (V_C) diffusion in 3C-SiC obtained from Nudged Elastic Band (NEB) calculations using Tersoff (left) and EDIP (right) interatomic potentials. The horizontal axis represents the reaction coordinate along the migration path, while the vertical axis shows the relative energy (eV). The initial configuration corresponds to a carbon vacancy located at its original lattice site, and the final configuration corresponds to the vacancy occupying the nearest neighbor carbon site after migration. The calculated migration barriers are 4.7 eV for Tersoff and 2.7 eV for EDIP.



(a) $EDIP$ (b) $Tersoff$

Figura 4.6: Energy profile for silicon vacancy (V_{Si}) diffusion in 3C-SiC from Nudged Elastic Band (NEB) calculations using the Tersoff (left) and EDIP (right) interatomic potentials. The reaction coordinate on the horizontal axis traces the migration pathway, while the vertical axis shows the relative energy in eV. The initial state corresponds to a silicon vacancy at its original lattice position, and the final state corresponds to the vacancy relocated to the nearest neighbor silicon site. Migration barriers are 4.3 eV for Tersoff and 5.2 eV for EDIP.

to the creation of a CAV defects. Calculated barrier for V_{Si} transition is illustrated in 4.6. It is notably higher than the value reported in existing literature, which is 3.5 eV as determined through ab initio calculations [54, 58] . A similar energy barrier of 3.0 eV is obtained using FPMD [1].

A comparative analysis of calculated results with corresponding literature [1] can be obtained from table 4.5. It is noteworthy, that energy barrier associate with silicon vacancy is symmetric unlike carbon and silicon antisite with the reason stated in subsection carbon vacancy.

4.6 Summary

In summary, the quasi static analysis reproduced key static properties of 3C-SiC, including lattice constants and defect formation energies, in good agreement with literature values. Among the tested potentials, EDIP exhibited the best overall correspondence with both classical molecular dy-

Defect Type	Forward $E_{Barrier}$ (eV)		$\Delta E_{Final-Initial}(eV)$		Literature
	Tersoff	EDIP	Tersoff	EDIP	
V_C	4.8	2.6	-	-	3.5 [54],2.5 [1],4.10 [111]
V_{Si}	4.5	6.4	-	-	3.0-3.5 [1],2.35 [111]
$V_C C_{Si}$	9.6	9.3	1.8	-1.6	1.9-2.4 [54],1.3 [1]
$V_{Si} Si_C$	1.7	1.8	1.3	1.4	-

Tabella 4.5: Comparative assessment of diffusion barriers determined via the Nudged Elastic Band (NEB) method and compared against data extracted from the literature.

namics and *ab initio* results, and uniquely reproduced the experimental melting point of silicon carbide a critical validation point. This finding is further reinforced by recent work by Giulia Galli and co workers, published in *Nature Communications*, which also adopted the EDIP potential for modeling SiC [1].

A noted limitation of EDIP is its relatively poor estimation of migration barriers when using the Nudged Elastic Band (NEB) method. However, this shortcoming is not unique to EDIP, the Tersoff potential also failed to provide complete agreement with reference barrier values. Given that activation energies for diffusion will ultimately be determined via the Arrhenius equation in the dynamic simulations, this limitation is acceptable for the scope of this work.

Overall, the quasi static framework proved effective for validating static properties but less suited for accurately capturing energetics of saddle point configurations and migration pathways. Building on these findings, the next chapter will focus on the diffusion of defects in 3C-SiC using the EDIP potential, enabling a comprehensive assessment of defect mobility under realistic thermal conditions.

Capitolo 5

Simulation of Non Equilibrium Processes and Defect Generation

5.1 Introduction

Ion implantation is a key non equilibrium process for introducing dopants and defects in SiC. High energy ions penetrating the lattice create cascades of atomic displacements (Frenkel pairs: vacancies and interstitials) and complex defects. Such irradiation can degrade device performance in high radiation environments, so understanding defect generation and evolution is critical [112, 113, 4].

5.2 Simulation of Non Equilibrium Processes and Defect Generation Using BCA

The binary collision approximation (BCA) is the traditional Monte Carlo method for simulating ion implantation. In BCA, each dopant ion is tracked as it undergoes successive two body collisions with target atoms in a (typically static) crystal lattice [112].

Popular process simulators implement BCA to predict dopant range, damage, and activation. For example, Synopsys Sentaurus Process is an industry standard TCAD tool that supports all processing steps on sili-

con carbide substrates: it recognizes multiple SiC polytypes (2H, 4H, 6H, 3C) and can perform Monte Carlo implantation (using its Taurus MC engine or the integrated Crystal TRIM). Crystal TRIM (sometimes referred to as TRICAD TRIM) is an extension of the well known SRIM/TRIM code that incorporates crystalline lattice geometry and channeling effects, in contrast to the amorphous only approximation of the original TRIM. Sentaurus also supports subsequent steps such as dopant diffusion and oxidation [114].

Similarly, Silvaco Victory Process is a commercial TCAD suite that models ion implantation and annealing, including advanced dopant activation for SiC (Victory provides specialized models for high concentration acceptor activation in SiC [115]). General purpose multiphysics solvers such as COMSOL can, in principle, be used to model diffusion or stress fields after implantation, but they lack built in BCA physics for ion penetration and channeling. In such cases, COMSOL must be manually scripted or coupled to external BCA data.

Academic and open source BCA codes also exist. Crystal TRIM [116] is a Monte Carlo engine that explicitly includes the crystal lattice for channeling simulations (Sentaurus Process can use Crystal TRIM under the hood [114]). scat-GUI (from Hosei University) is a graphical BCA code tailored for ion channeling and scattering, it implements advanced stopping models (e.g., Firsov and Oen–Robinson) and even includes a many body electronic stopping option. For example, recent releases of scat-GUI include SiC specific parameters (the electronic stopping β was set to 0.45 for SiC) [117] to improve accuracy for wide bandgap targets. The open source SIIMPL code[118] is a MATLAB/C implementation of BCA for crystalline targets, developed specifically for SiC. MulSKIPS, by contrast, is not a BCA code for implantation but a kinetic Monte Carlo super lattice simulator for epitaxial growth and defect kinetics in SiC and related materials [119]. We include MulSKIPS here because it is part of the recent MUNDFAB consortium tools for SiC device fabrication, it handles doping and defect evolution during growth rather than ballistic implantation.

Each of these BCA tools can generate dopant depth profiles and estimate damage. For example, Sentaurus Process provides parameter sets and an "Advanced Calibration 4H-SiC" option for specific dopants (B, Al, N) with [0001] orientation [114]. In Victory Process, new impurity

activation models for SiC also account for doping saturation [115]. scat-GUI and SIIMPL explicitly model crystallographic effects (channeling) by positioning lattice atoms and allowing off axis collisions.

BCA simulators efficiently compute dopant ranges and preliminary damage and can easily handle large numbers of ions at keV–MeV energies. However, they assume a static or thermally broadened lattice, so channeling behavior is approximated. Damage and defect generation are estimated via empirical formulas (e.g., Kinchin–Pease) and thus do not capture intra cascade recombination or dynamic annealing. These limitations are particularly critical in wide bandgap materials like SiC, where calibration is required [114, 115].

5.3 BCA simulation of ion implantation in Wide Band Gap semiconductors

Wide bandgap semiconductors (e.g. 4H-SiC, GaN, AlN, Ga_2O_3, diamond) are typically doped by ion implantation because diffusion is limited. Binary collision approximation (BCA) Monte Carlo simulators (both commercial and open source) are widely used to predict implanted dopant and damage profiles. In these codes, an energetic ion is tracked through successive two body collisions against lattice atoms, nuclear collisions above a displacement threshold create vacancies/interstitials (damage), while energy loss via electronic stopping is modeled by empirical formulas. Modern BCA tools remain indispensable for implantation studies. They can quickly explore process windows (ion species, energies, angles, dose, penetration depth) and give first order estimates of dopant distributions and damage. Here I will review some of the work where BCA based implantation is used to fabricate defects in the crystals.

Silicon carbide (SiC) is a widely studied wide bandgap (WBG) material with numerous applications in electronics. Linnarsson et al. [120] implanted $^{51}V^{+}$ ions at 50-300 keV into 4H-SiC(0001) at room temperature (and up to 500 °C) along various tilt angles (0–17° off the [000$\bar{1}$] axis) with fluences in the range of $\sim 10^{12}$–10^{14} cm^{-2}. Using the SIIMPL Monte Carlo BCA code, they found that channeling dominates. Even small tilts could not eliminate deep penetration. In three-dimensional si-

mulations, SIIMPL predicted that “ions will be steered into the $[000\bar{1}]$ and $\langle 11\bar{2}3 \rangle$ directions and therefore penetrate deep into the sample.” Elevated temperature (up to 500 °C) increased lattice vibrations and dechanneling, slightly reducing penetration depth. Measured SIMS dopant profiles confirmed these simulations, showing that implanted V formed long channeling “tails” deep in the wafer.Similarly, Linnarsson et al. [120] examined the channeling effect in 4H-SiC by implanting 100 keV Al^+ and 50 keV B ions along the [0001] direction at 600 K. They observed that channeling implants produced much deeper penetration profiles than random implants. Moreover, the penetration depth for Al decreased with increasing temperature. The channeling effect was successfully captured and explained by SIIMPL simulations. In another study, Muting et al. [121] (2020) used the Synopsys Sentaurus process to perform BCA-based implantation simulations of 180 keV Al^+ ions in 4H-SiC at 500 °C with doses ranging from 5×10^{13} to 5×10^{15} cm^{-2}. These simulations successfully tracked point defects such as vacancies and interstitials, and showed that defect concentration increases with dose and saturates at the maximum dose. In summary, BCA tools vary in fidelity. Commercial tools and SRIM are convenient but often assume amorphous substrates by default. Open tools like SIIMPL and scatGUI explicitly include the crystal geometry, enabling channeling studies. All these BCA codes generate vacancy/interstitial counts (often via a Kinchin Pease or NRT model) to produce defect depth profiles, but none inherently simulate post implantation annealing or defect clustering. Past studies have highlighted that not only ion irradiation but also epitaxial growth processes can introduce significant numbers of intrinsic defects in wide bandgap semiconductors. Awschalom et al. [8] reported the incorporation of open volume point defects (OVPDs) during the growth of 4H-SiC, with densities on the order of $10^{12}\,\mathrm{cm}^{-3}$. Building on this evidence, La Magna *et al.* (2025) [122] extended the investigation to both planar and three dimensional 3C-SiC substrates. Using a superlattice kinetic Monte Carlo (SlKMC) framework implemented within the MulSKIPS code.They modeled epitaxial growth under non equilibrium conditions. These simulations account for a diverse set of point defects, including carbon and silicon vacancies (V_C, V_{Si}), antisites (Si_C, C_{Si}), vacancy antisite complexes (SAVs, CAVs), divacancies, and stacking faults arising from polytype instability. The SlKMC-based modeling predicts

that OVPDs are not only generated but also redistributed near the surface region of the 3C-SiC substrate. Two notable implications emerge from these findings: first, epitaxial growth alone can generate high densities of point defects particularly V_C and V_{Si}without any subsequent processing (e.g., implantation or irradiation). Second, by tuning fixed process parameters during growth, it may be possible to manipulate the abundance or depletion of near surface point defects, opening pathways for defect engineering.

5.3.1 Limitations

Despite their utility, BCA methods suffer from several limitations. It is understood that BCA approach treats the lattice as static during implantation simulations, thereby neglecting generic in situ defect evolution processes such as recombination and migration. As a result, all created defect cascades (vacancies and interstitials) remain in the lattice and do not recombine. In reality, however, most defects recombine almost instantly, especially at elevated temperatures, leading to an overestimation of damage creation. In such cases, molecular dynamics (MD) approaches are more appropriate.

Similarly, crystal based BCA tools (e.g., SIIMPL, scatGUI, Sentaurus with crystal options) can model channeling, but their accuracy depends strongly on the choice of scattering potentials and the treatment of thermal vibrations. Simplifications, such as neglecting multi atom collision cascades, may lead to mispredictions of channeling peaks. Linnarsson et al. [123] reported that fast BCA simulations predicted channeling peaks closer to the surface than observed experimentally, indicating an underestimation of deep channeling. They achieved improved agreement by enabling the full cascade option (tracking all recoils), though at a significantly higher computational cost. By contrast, simpler amorphous codes such as SRIM completely miss channeling effects. Nishimura et al. [124] demonstrated that SRIM profiles differed considerably from experimental data, whereas the crystal aware scatGUI reproduced the measurements more accurately.

Another limitation is the reliance on the Kinchin Pease or NRT models with fixed displacement thresholds in BCA codes. Recent studies have

shown that near threshold collisions may be incorrectly counted as vacancies by SRIM, whereas the actual defect production could be smaller. Thus, none of the BCA codes can naturally predict defect clustering, highlighting the need for atomistic simulations or experimental validation [125].

5.4 Simulation of Non-Equilibrium Processes and Defect Generation using Molecular Dynamics

Classical molecular dynamics (MD) has become an important tool for this, allowing atomistic modeling of collision cascades in 3C-, 4H-, and 6H-SiC. These simulations typically use bond order potentials (e.g., Tersoff) with short-range Ziegler–Biersack–Littmark (ZBL) corrections to model high-energy collisions [126]. Systems are equilibrated (often by NPT) and then cascade events are run under NVE, some studies also include electronic stopping (e.g., via a friction term) to account for energy loss to electrons [127]. Simulations are performed at various temperatures (often 100–450 K) and implantation geometries (normal vs. glancing incidence) to capture thermal effects on defect creation [5, 126]. In this section, we investigate the ion implantation as non-equilibrium mechanism responsible for the generation, migration, and evolution of defects in silicon carbide (SiC). These processes, which deviate significantly from thermodynamic equilibrium, often lead to the formation of metastable defect configurations, complex defect clusters, and kinetic pathways that are not accessible under equilibrium conditions. Such phenomena play a critical role in determining the material's micro structural integrity and functional properties, particularly in high radiation and high temperature environments.

5.4.1 Defect Evolution During the Ballistic Cascade(initial Picosecond scale damage)

During the ballistic cascade stage (microcanonical, NVE), the implanted ion transfers its kinetic energy through a series of two body collisions,

generating Frenkel pairs (vacancy–interstitial) and small defect clusters along its path. In MD simulations, the defect count rises sharply as the primary knock on atom (PKA) slows down, peaking once its kinetic energy is fully dissipated. This trend is also evident in our results (Fig. 5.2). This behavior is also strongly supported in the literature. For instance, Chen et al. (2023) [3] reported that the Frenkel pair population peaks within 20 ps for keV scale Si PKAs in 6H–SiC. Likewise, Bonny et al. (2020)[128] reported that a 1–100 keV Si PKA in 3C–SiC generates roughly 10–1000 Frenkel pairs, with the final defect numbers agreeing within about a factor of two with predictions from the binary collision approximation (BCA) models, such as TRIM (Transport of Ions in Matter) and the Norgett–Robinson–Torrens (NRT) model. Once the PKA stops, the Frenkel count stabilizes and any remaining recombination begins (the "thermal spike" cools). In Chen's simulations the defect count drops slightly after 21 ps as the cascade thermalizes, then levels off to the stable damage state [128, 3].MD cascades create all primary SiC defect species: Si and C vacancies (V_{Si}, V_C), interstitials (I_{Si}, I_C) and antisites (Si_C, C_{Si}). Because carbon atoms are lighter and have a lower displacement threshold, more C vacancies/interstitials form than Si ones. Bonny et al. observed in 3C–SiC that the C:Si ratio of both vacancies and interstitials is significantly >1 (i.e. cascades eject more C than Si atoms)[128]. The cascade briefly heats the local lattice (up to 10³ K), enabling partial annealing. However, much of the ballistic damage persists. Notably, higher-energy cascades exhibit lower Frenkel-pair recombination fractions: Chen et al. show the recombination rate drops as PKA energy rises[3] . Thus very energetic ions leave more lasting damage in the first few ps. (Subsequent annealing or diffusion at longer times can heal many of the smaller defects, but the primary damage topology is set by this microsecond to nanosecond cascade.)

5.4.2 Post ballistic defect evolution

MD simulations of SiC cascades consistently report formation of all point defect types. The primary Frenkel pairs are carbon and silicon vacancies (V_C, V_{Si}) paired with carbon and silicon interstitials (I_C, I_{Si}). Due to the binary lattice, irradiation also produces antisite defects (C_{Si} and

Si_C). Several studies highlight that vacancies and antisites dominate cluster formation, while interstitials tend to remain isolated. For example, Fung et al. (2018) simulated high energy irradiation of 3C-SiC and observed that many interstitials scatter broadly through the lattice, whereas the less mobile vacancies congregate into clusters [127]. Antisite defects (Si_C and C_{Si}) appear in many cascades: in Chen's 6H-SiC study these were counted alongside vacancies [113]. In Si, high energy implantation creates tri-interstitial clusters (W-centers) used as quantum emitters. Gennetidis et al. showed Ga implantation in Si generates increasing Si tri-interstitial (W-center) clusters with energy, and that annealing preserves these clusters while removing other defects [5]. As implantation dose increases, defects accumulate and clusters grow. Fan et al. (2023) performed MD of Xe implantation in 4H-SiC (coupled with Raman experiments) and found a roughly linear increase in V and I counts with dose up to 2×10^{14}ions$/cm^2$, beyond that, the growth rate slows and defects begin to saturate [4].This indicates that repeated cascades eventually create overlap of cascades and amorphous zones. Fan et al. also noted that heavy Xe ions penetrate deeply, causing severe internal damage: their MD showed that Si vacancies could not form as perfect isolated centers (we can say the lattice is too much disrupted) [4].By contrast, lighter ions (H, He) tend to create shallower damage and allow point defects to survive. Similarly Chen et.al also examined how temperature and geometry affect the defect creation in 6H-SiC lattice. They found linear relation between Frankel pair counts and temperature. We can argue that atoms with warm lattice are easy to displace [113]. They also found that implant incident at 45 degree yields the fewest of total defects, this yield increases by lowering incidence.

Annealing and Defects diffusion After implantation, subsequent diffusion and annealing can change defect populations. MD (often combined with thermal spike models) has been used to study dynamic recovery. Notably, Backman et al. (2013) used a thermal spike MD model for swift heavy ion (electronic energy) irradiation of 3C-SiC. They found that the localized heating can anneal pre existing damage. Point defects were significantly healed and amorphous regions partly recrystallized at interfaces [129]. We can correspond strong self electronic excitation with

annealing as well. On the annealing side, Wu et al. (2021) simulated 3C-SiC after Al implantation with subsequent thermal annealing. They observed recrystallization proceeding from the crystalline interface into the damaged region. Compressive stress built up at the amorphous/-crystal interface and new dislocations formed there [126]. Importantly, Wu et al. reported that annealing (especially after high temperature implants) tends to leave behind carbon vacancies: their MD predicted that C vacancies are especially likely to remain after annealing of high temperature implanted samples [126]. In 6H-SiC, Chen et al. noted that raising temperature increased Frenkel pair recombination up to a point, but very high local heating could also create larger cascades (over 1000 K peaks) that again promote defect recombination [113].This complex behavior shows that MD can capture non linear recovery after damage. modest heating helps recombine defects (increase in recombination "centers"), while too much heating generates new local cascades that eventually anneal again [113]. As summarized in Table 5.1, the Binary Collision Ap-

Aspect	BCA (Monte Carlo)	MD (Atomistic)
Accuracy	Approximate, empirical, misses cascade dynamics [126, 58]	High fidelity, captures cascades, annealing, mobility [58, 126]
Materials	Broad, calibrated for Si, extended for SiC (Sentaurus, Victory) [114, 115]	Limited by potentials, mainly Si, SiC, some WBGs
Openness	Mostly commercial, some open codes exist (SIIMPL, scat-GUI) [112, 117]	Mostly open-source (LAMMPS, custom MD)
Use Cases	Process-level TCAD, dopant profiles, dose calibration [114]	Fundamental damage, defect dynamics, annealing [4, 58]

Tabella 5.1: Comparison of BCA and MD methods for ion implantation in Si/SiC.

proximation (BCA) remains useful for rapid, processflevel implantation studies, particularly in Technology ComputerfAided Design (TCAD) frameworks where dopant depth profiles and dose calibration are required

[114, 115]. However, its approximate, empirical nature inherently neglects cascade dynamics, defect recombination, and annealing processes that are critical in widefbandgap semiconductors. In contrast, Molecular Dynamics (MD) simulations, despite being computationally more demanding and limited by the accuracy of interatomic potentials, provide a highffidelity atomistic description of ion–solid interactions. MD captures collision cascades, defect mobility, and dynamic annealing events in real time, enabling direct insight into defect generation and evolution at the atomic scale [58, 126, 4]. These advantages make MD a superior choice when the research objective is to understand fundamental damage mechanisms and defect formation in Si, SiC, and other wide band gap materials, positioning it as a necessary complement to or refinement of BCA based approaches.

5.5 Molecular Dynamics Simulation Results

Having established the theoretical background and methodological framework for ion implantation and non equilibrium defect dynamics in SiC in previous sections, we now present the molecular dynamics results obtained for a 13,825-atom supercell subjected to implantation energies ranging from 0.1 to 100 keV. These simulations track both the ballistic generation of defects and their subsequent evolution during annealing, allowing us to resolve the dependence of defect accumulation and recovery on implantation energy and thermal treatment.

5.5.1 Simulation Setup

In practice, MD simulations of ion implantation in SiC use large supercells (periodic or with free surfaces) and impart a chosen kinetic energy to a primary knock-on atom (PKA) to initiate a cascade.

To establish the methodology for simulating ion implantation processes, we performed a series of preliminary molecular dynamics (MD) calculations. A 13 824 ($12 \times 12 \times 12$) atom supercell of silicon carbide was constructed as the simulation domain. An external silicon (Si) atom was selected as the implanted ion (totaling 13 825 atoms) , introduced with

a relatively low implantation energy range (100 eV,130 ev,140 ev,155 ev) to coap the smaller box size. Impinging silicon atom was initialized at a zero incidence angel to the (001) plane normal. The integration time step for the simulations was set to 0.1 fs to ensure accurate resolution of high-energy collision dynamics. Thus direction of implantation was aligned perpendicular to the *xy*-plane to mimic vertical ion bombardment. To model energy dissipation and ensure physical boundary conditions, the supercell was partitioned into three distinct regions:

1. **Newtonian region:** Located centrally, this region was evolved under Newtonian dynamics (microcanonical ensemble, NVE) to conserve the momentum of the implanted ion and accurately capture its interactions with the host lattice. Implant atom was also simulated with microcanonical ensemble.

2. **Thermostat region:** Surrounding the Newtonian zone, this region was coupled to a Nose Hoover thermostat to facilitate heat dissipation and stabilize local temperature variations caused by ion impacts.

3. **Boundary region:** The outermost layer of the supercell was fixed or weakly coupled to maintain structural integrity and suppress spurious reflections of phonons.

The implanted silicon ion was initially coupled to a canonical ensemble (NVT) to control its thermal energy during the early stages of penetration. Following ion insertion, a 0.1 ps simulation was conducted under the NVE ensemble to capture non equilibrium momentum transfer and defect formation that lies within the ballistic collision regime of 0.1-0.4 ps [130]. System was subsequently followed by a 10 ps annealing phase using the NVT ensemble to allow the system to relax and facilitate defect recombination, migration, or clustering. Defect populations were analyzed using the Wigner–Seitz cell method, as implemented in OVITO, by comparing atomic positions against a reference perfect lattice. This enabled the reliable identification of point defects, including vacancies, interstitials, antisites, and divacancies.

These preliminary simulations serve as the foundation for systematic investigations of defect formation under controlled non equilibrium conditions.

5.5.2 Ballistic Regime

In order to resolve the short time dynamics of defect generation at lower recoil energy, we simulated a 155 eV implantation event in the $12 \times 12 \times 12$ SiC supercell for 5 ps. The results, shown in Figure 5.1, exhibit the characteristic initial spike in defect production immediately following the ion impact. A transient burst of vacancies and interstitials is observed within the first picoseconds, accompanied by the formation of bi vacancies. As the cascade dilutes, the overall defect count decreases, reflecting partial recombination and defect annihilation processes. These results highlight the fast, highly non equilibrium nature of defect formation and early stage recovery in the ballistic regime. Such trends are consistent with previous reports by Bonny *et al.* (2020) [128] and Chen *et al.* (2023) [3], as discussed in Section 5.4.1.

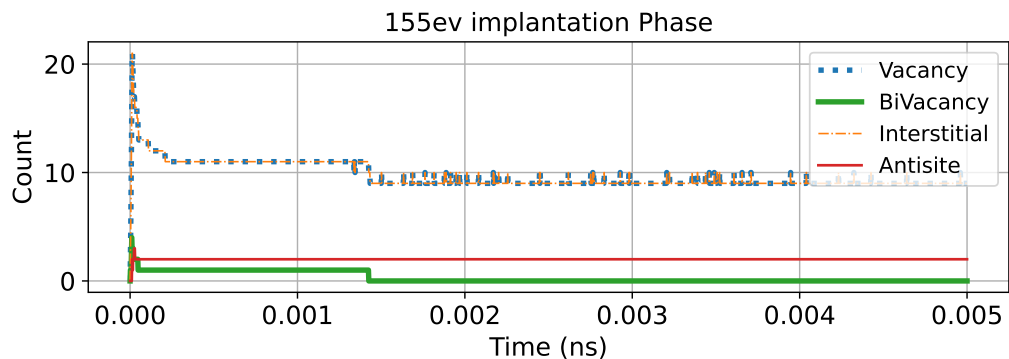


Figura 5.1: Defect evolution during the ballistic regime of a 155 eV implantation in a 13 825 atoms SiC supercell over 5 ps. An initial spike in vacancies and interstitials is observed, along with the formation of bi vacancies. Subsequent dilution of the cascade leads to a gradual reduction in the total defect population due to recombination and defect annihilation.

The results also got support from literature. Gennetidis et al. (2024) simulated Ga ion cascades in Si and found that Si tri interstitial clusters (W-center defects) form more readily at higher ion energies [5].They showed a linear increase in number and extention depth with energy.

5.5.3 Energy Dependent Defect Accumulation

To investigate the defect generation and evolution in silicon carbide (SiC) under non equilibrium conditions, we employed molecular dynamics simulations using a $12 \times 12 \times 12$ cubic supercell consisting of 13,825 atoms. A single silicon ion implantation event was simulated, with ion kinetic energies varied between 100 eV and 155 eV in order to assess the dependence of defect production on implantation energy. The simulation protocol was structured into distinct phases to capture dynamic thermal effects, and computational details are reported in Section 5.5.1. The results show a clear correlation between implantation energy and defect production. At 100 eV, the defect count rapidly increases during the first few picoseconds before stabilizing to approximately four vacancies and interstitials, accompanied by one divacancy and one antisite. At 130 eV, the system initially produces around Ten defects, which stabilize to six to seven vacancies and interstitials, along with two divacancies and two antisites. At 150 eV, the steady state defect population consists of eight to nine vacancies and interstitials, with two divacancies and two antisites. At the highest energy studied (155 eV), the system exhibits a sharp initial rise to nearly 20 vacancies and interstitials, which gradually decrease and stabilize to about 11, together with two divacancies and more than two antisites. These observations demonstrate a clear trend of increasing defect generation with implantation energy, while also highlighting the transient surge in defect populations during the very early picoseconds of the simulation (see Fig. 5.2).

5.5.4 Temprature dependent Annealing Study

Following the ballistic implantation stage, the system was equilibrated using the NVT ensemble at the target temperatures of 1500 K, 1900 K, 2100 K, and 2300 K to investigate the effect of annealing temperature on

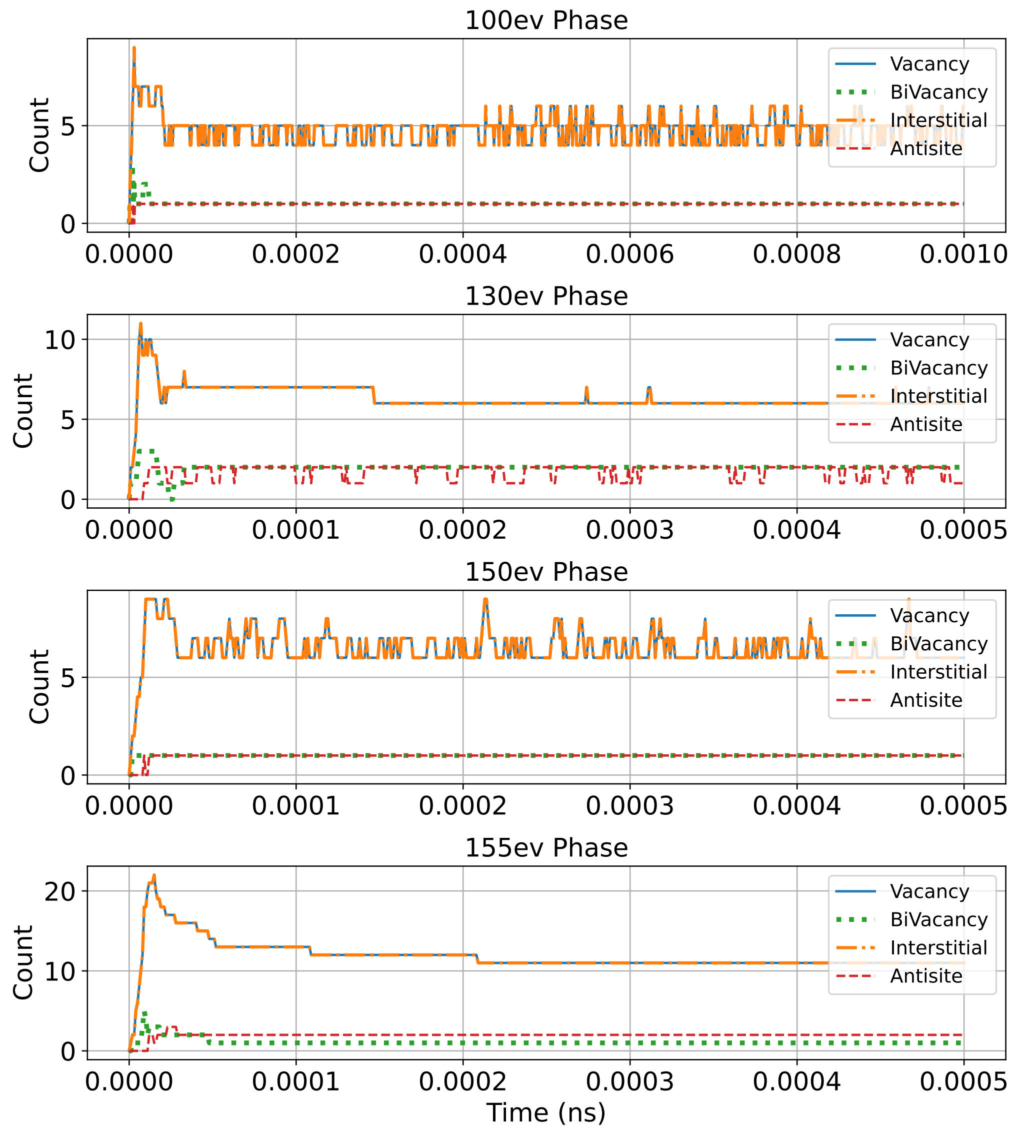


Figura 5.2: Defect evolution in SiC under ion implantation at different energies (100–155 eV). The plots show the time dependent populations of vacancies, interstitials, divacancies, and antisites obtained using the Wigner–Seitz cell method. A pronounced transient increase in defect count is observed during the initial picoseconds, followed by stabilization at energy dependent steady state values. Higher implantation energies yield larger defect populations, with the most significant damage observed at 155 eV.

defect evolution. The defect populations were analyzed up to 10 ps using the Wigner–Seitz method, allowing for the identification of vacancies, interstitials, antisites, and divacancies.

Across all temperatures, the simulations reveal a pronounced transient increase in defect populations during the very early picoseconds, followed by a progressive decrease as recombination and annihilation processes set in. Interstitials exhibit the highest mobility and rapidly recombine or escape from cascades, while vacancies and antisites display greater persistence, often forming small clusters. Consequently, both vacancy and interstitial counts decrease with time, with the reduction being more pronounced at elevated temperatures.

Temperature elevation generally enhances recombination efficiency (dynamic annealing), accelerating the reduction of Frenkel pairs. However, higher temperatures also promote more initial defect formation due to increased atomic mobility and localized thermal spikes, a feature particularly evident in the 2100 K and 2300 K cases. This dual role of temperature initially amplifying damage production but subsequently enhancing defect recovery highlights the complex balance between defect generation and healing under thermal treatment.

These observations are consistent with molecular dynamics trends reported in the literature, where heavy ions have been shown to produce transient thermal spikes that can heal part of the damage, while simultaneously driving deeper amorphization when implantation energies are sufficiently high. Compared with the ballistic regime, the annealing simulations clearly demonstrate the critical role of temperature in moderating the long term survival and stability of point defects in SiC.

5.5.5 Effect of Annealing Duration on Structural Recovery at 1500 K

To further explore the temporal evolution of defects during thermal annealing, the 155 eV implantation case was extended to 0.5, 1.0, and 1.5 ns at 1500 K. The defect populations, presented in Figure 5.4, show clear signatures of dynamic recombination and defect restructuring beyond the ballistic regime.

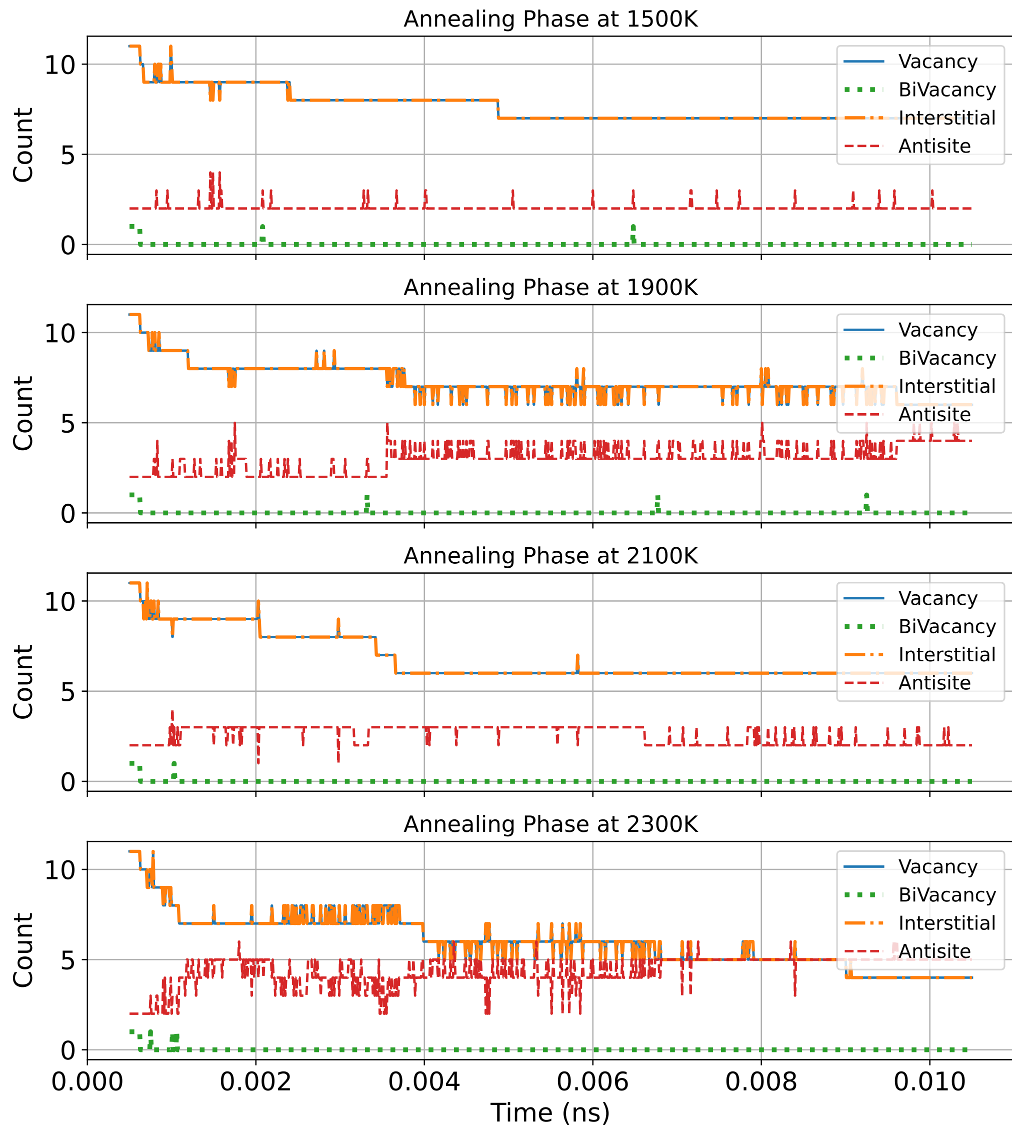


Figura 5.3: Defect evolution during post implantation annealing in SiC at 1500–2300 K following a 155 eV recoil. The initial ballistic phase produces a transient surge in vacancies, interstitials, divacancies, and antisites, which rapidly evolve within the first few 5.1. Interstitials recombine or escape due to their high mobility, while vacancies and antisites persist and partially cluster. Elevated annealing temperatures enhance recombination efficiency (dynamic annealing), leading to a progressive reduction in surviving Frenkel pairs, although higher temperatures also generate larger initial defect populations.

At 0.5 ns, both vacancies and interstitials remain present at moderate levels, with occasional divacancy formation events and a small population of antisites ($\sim$2). By 1.0 ns, the system undergoes a marked reduction in Frenkel pairs, with vacancy–interstitial annihilation occurring around 0.9 ns, leading to a noticeable decrease in surviving defects. During this period, divacancies exhibit transient stability, forming and dissolving intermittently, while the antisite population gradually increases. By 1.5 ns, the recombination of vacancies and interstitials continues, leaving fewer Frenkel pairs, whereas antisites further increase from two to approximately four, indicating progressive lattice reconfiguration. Meanwhile, divacancies persist as dynamic entities, fluctuating between presence and annihilation.

Overall, these observations suggest that at 1500 K the annealing process is dominated by vacancy interstitial recombination, while antisites emerge as more stable long lived defects. Divacancy dynamics highlight the transient clustering tendency of point defects, although their stability is limited within the 1.5 ns timeframe. These results emphasize the interplay between annihilation and defect transformation, underscoring the importance of annealing timescales in determining the surviving defect populations.

5.6 Conclusions

In this work, molecular dynamics simulations were carried out to investigate ion implantation and defect evolution in 3C-SiC using a $12 \times 12 \times 12$ supercell comprising 13,825 atoms. Implantation energies ranging from 100 ev to 155 eV were explored, providing a comprehensive picture of ballistic damage creation and subsequent defect evolution during annealing.

The ballistic phase of implantation was found to generate a large number of vacancy–interstitial pairs, with the peak defect population occurring within the first picoseconds as the primary knock on atom dissipated its kinetic energy. At low recoil energies (50 eV), the resulting cascades created only a handful of Frenkel pairs and small complexes, which quickly recombined during annealing. At higher recoil energies (150–200 eV),

cascades produced substantially larger damage zones, resulting in persistent populations of vacancies, interstitials, and antisites even after long annealing runs. The nonlinear increase in defect retention with implantation energy highlights the sensitivity of residual disorder to ion energy, consistent with previous computational and experimental studies.

Thermal annealing strongly influenced the post implantation defect landscape. Both phase separated annealing (heating, equilibration, cooling) and viscous damping approaches demonstrated significant dynamic recovery, with defect counts decreasing over nanosecond timescales. However, complete recombination was not achieved, and stable antisites and a reduced number of Frenkel pairs remained in the lattice. This persistence of antisites supports earlier reports that antisite formation provides a low energy pathway for interstitials, particularly in SiC where such configurations are relatively stable.

Overall, these simulations provide atomistic insight into the interplay of implantation energy, cascade dynamics, and annealing protocols in governing the stability of irradiation induced defects in SiC. The findings corroborate literature trends regarding energy dependent defect yields, the role of antisites as stable end products of interstitial activity, and the critical influence of temperature on recombination kinetics. Future work should extend these studies to larger supercells, longer annealing times, and inclusion of electronic stopping effects to bridge closer to experimental implantation conditions. Such simulations will be crucial for predicting long term radiation resilience of SiC in advanced nuclear and electronic applications.

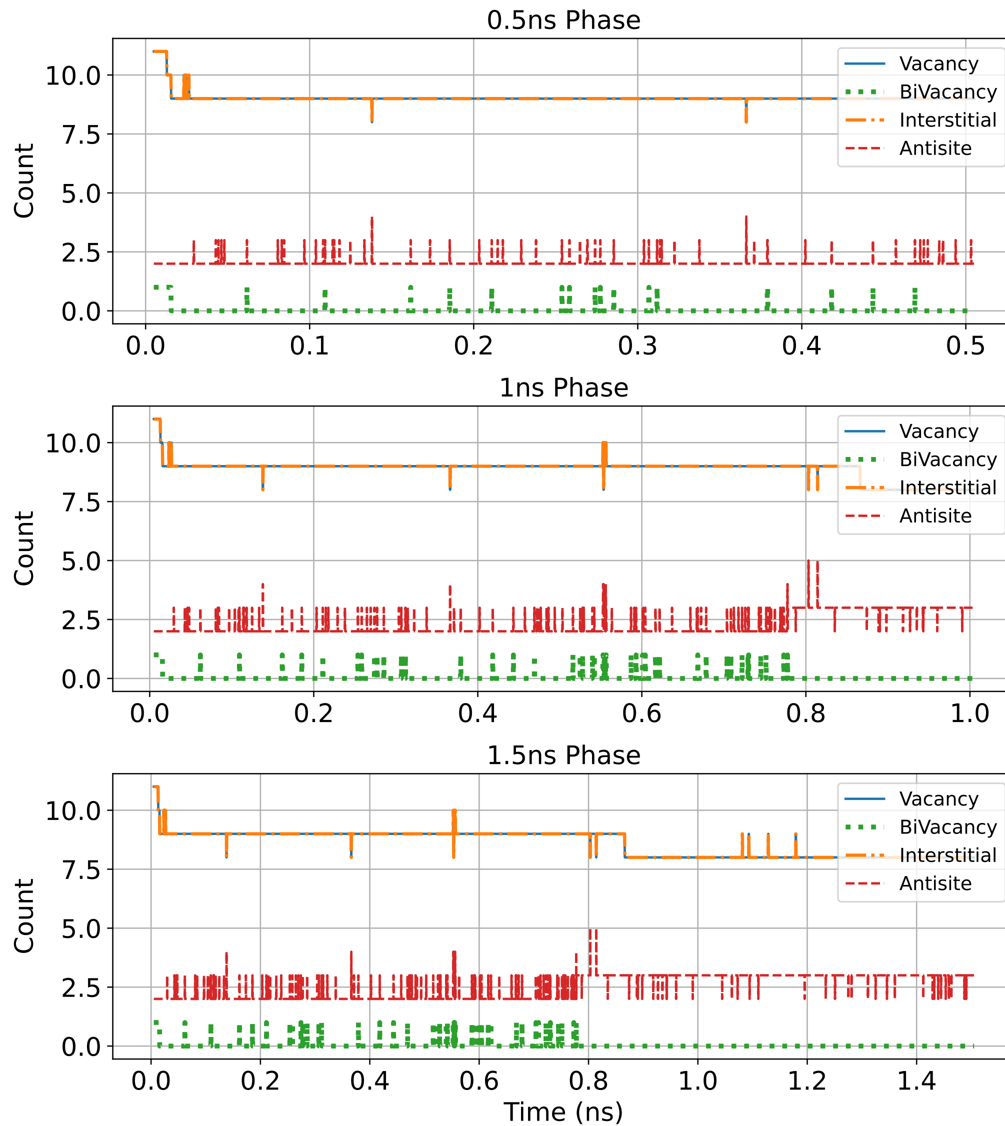


Figura 5.4: Defect evolution during annealing at 1500 K following a 155 eV implantation in a $12 \times 12 \times 12$ SiC supercell. Results are shown for 0.5, 1.0, and 1.5 ns. Vacancy–interstitial recombination becomes evident around 0.9 ns, leading to a reduction in Frenkel pairs. Antisites increase progressively from two to four over the simulation time, while divacancies exhibit dynamic formation and dissolution.

Capitolo 6

Diffusion of point defects in 3C-SiC

6.1 Introduction

In the previous chapter, research efforts were undertaken to test and establish a computational framework upon which the core analysis of atomic scale defect diffusion, associated pathways, formation mechanisms, and corresponding energetics can be built. In this chapter, we aim to ensure the robustness of established framework by extending the defect energetics analysis to include the evaluation of diffusion coefficients through the study of mean square displacement and center of mass motion(frequency rule method). The estimated diffusivity D is subsequently used to determine the migration energy barrier via Arrhenius plots. Comparison of results from both methods suggested that frequency rule method provides a more stable estimate of diffusivity. The formation of divacancy (VV) complexes and the annihilation processes involving V_C and I_C are presented as key diffusion outcomes.

6.2 Diffusion Mechanisms of Point Defects in SiC

We begin with a brief discussion of the various types of intrinsic defects reported in silicon carbide. These include the carbon vacancy (V_C), silicon vacancy (V_{Si}), carbon interstitial (I_C), silicon interstitial (I_{Si}), antisites, and their complexes with respective vacancies ($C_{Si}V_C$, Si_CV_{Si}). First principles studies have mapped out the dominant migration pathways and energy barriers for these defects. In general, vacancies migrate by atom vacancy exchange within their own sublattices, whereas interstitials move via split interstitial and kick out mechanisms [54]. We next discuss the respective migration mechanisms.

Carbon vacancy migration. Carbon vacancy diffusion occurs within the carbon sub lattice through a second nearest neighbor hop. The migration barrier is reported to be approximately 4.0 eV in the literature and is confirmed by our NEB calculations, as presented in Section 4.5. This indicates that the carbon vacancy remains relatively immobile at low temperatures until a substantial increase in temperature activates the hopping mechanism. Bathen et al. (2019) explicitly report an activation energy of ≈ 4.0 eV for V_C migration in 3C-SiC [17], consistent with earlier DFT studies. In MD simulations and doped material calculations, even when Si vacancies are created, a neighboring carbon atom often hops into the site, effectively converting the defect into carbon vacancy and carbon antisite complex. This observation supports the conclusion that V_C rather than V_{Si} is the dominant migrating vacancy species under typical annealing conditions [54]. The detailed description of mono and divacancy migration pathways can be obtained from figure 2.6

Silicon vacancy migration Silicon vacancies (V_{Si}) diffuse within the silicon sublattice following a similar pattern to carbon vacancies, but with a somewhat lower migration barrier. Despite the lower barrier of approximately (3) eV compared to V_C, V_{Si} is less likely to participate in diffusion because it is metastable [1]. Bockstedte et al. explained that the stability of V_{Si} is hindered by its rapid conversion into a carbon antisite vacancy

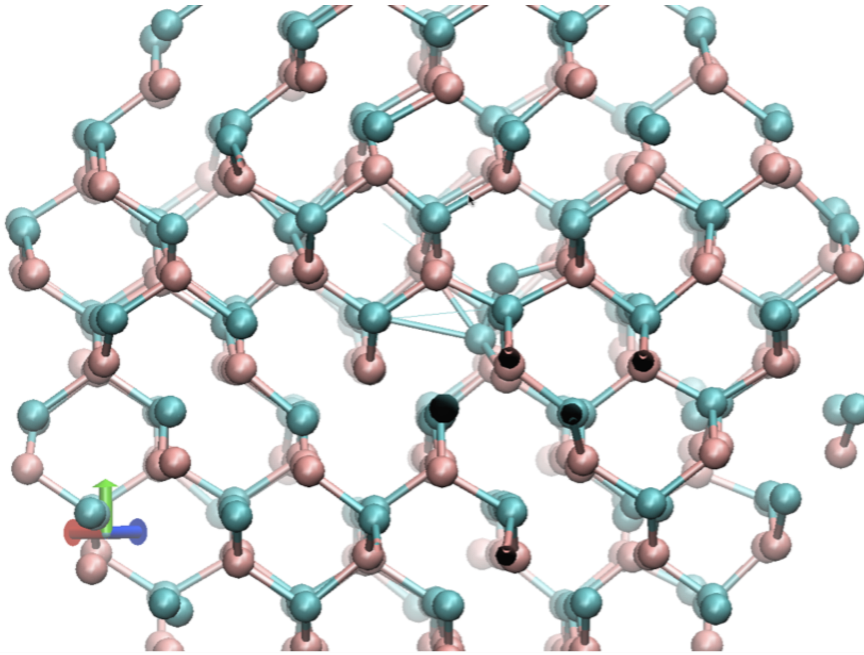
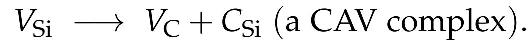


Figura 6.1: Formation of a dumbbell shaped diinterstitial complex: The introduction of an interstitial carbon atom induces the displacement of a lattice carbon atom, generating a carbon vacancy. This facilitates the formation of a stable dumbbell shaped diinterstitial complex, which exhibits mobility and diffuses through the crystal lattice.

(CAV) complex, thereby immobilizing the vacancy [54].



$$V_{\mathrm{Si}} \longrightarrow V_{\mathrm{C}} + \mathrm{C}_{\mathrm{Si}} \text{ (a CAV complex).}$$

Even under extreme n-type conditions, the formation of this complex suppresses silicon vacancy migration. The conversion is also confirmed my our Molecular Dynamics simulation results.

Carbon interstitials Carbon interstitials (I_{C}) exist as split interstitial pairs, forming a dumbbell shaped rotating as we expressed in figure 6.1. Calculations suggest a migration barrier of approximately 0.7 eV for the long-range diffusion of I_{C} [131]. Such a low barrier implies rapid diffusion during annealing, as will be discussed in Section 4.5. This observation is

consistent with the hypothesis proposed by Bockstedte et al., who argued that split interstitials diffuse significantly faster than [54].

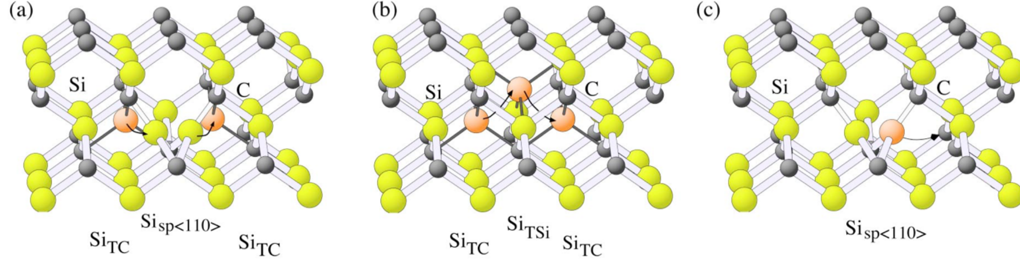
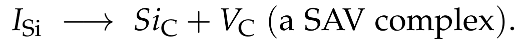


Figura 6.2: Migration of silicon interstitials: (a) kick out migration of Si_{TC} via the split interstitial $Si_{sp<110>}$, (b) migration via the tetrahedral interstitial sites Si_{TC} and Si_{TSi} , and (c) jumps between adjacent $Si_{sp<110>}$ configurations(figure obtained from [54]).

Silicon Interstitials Silicon interstitials (I_{Si}) also migrate relatively easily, although their diffusion involves more complex pathways (e.g., transitions between tetrahedral and split interstitial configurations). Zhao et al. report a comparable migration barrier of approximately 0.73 eV for I_{Si} during long range diffusion [131]. Similar to I_{C}, these values are much smaller than vacancy migration barriers, indicating that I_{Si} can diffuse readily when present. However, silicon interstitials generally possess higher formation energies in equilibrium, making them less prevalent than I_{C} under intrinsic or lightly doped conditions.

Interstitial defects in SiC are inherently complex because a single defect can potentially occupy several nonequivalent interstitial sites within the lattice. However, our analysis indicates that among these possibilities, only one configuration is energetically stable. In this configuration, the carbon interstitial pairs with a displaced lattice carbon atom as its nearest neighbor, forming a characteristic dumbbell shaped structure. The two carbon atoms rotate around their common axis, a behavior that can be clearly visualized in Figure 6.3.



$$I_{\mathrm{Si}} \longrightarrow Si_{\mathrm{C}} + V_{\mathrm{C}} \text{ (a SAV complex).}$$

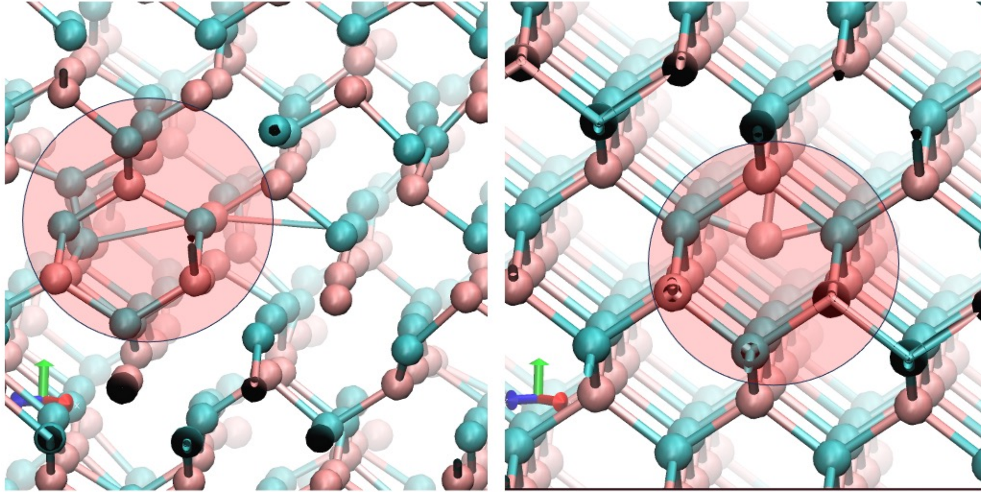


Figura 6.3: Initial and final configurations extracted from the molecular dynamics trajectory, showing the transformation of a silicon interstitial (I_{Si}) into a silicon antisite (Si_C) accompanied by the creation of a dumbbell shaped carbon interstitial (I_C). The specific region of the supercell where the transformation occurs is highlighted with a circle for clarity.

We also find indirect computational evidence in the literature that supports this argument. Lento *et al.* (2004) [132] reported that the silicon interstitial (I_{Si}) has a relatively high formation energy (∼6 eV) compared to antisite defects, implying that antisite formation is thermodynamically more favorable once appropriate pathways are accessible. In addition, Zheng *et al.* (2013) [133], through *ab initio* calculations of various interstitial antisite recombination processes in 3C-SiC, explicitly demonstrated that I_{Si} can transform into antisites and associated interstitials. Taken together, these findings suggest that the conversion of I_{Si} into antisites represents a likely fate of silicon interstitials under dynamic conditions. This so called kick-out mechanism is presented in figure 6.3 that we extracted from our simulation video. The evolution of defects during the simulation at 1600 K was also analyzed by tracking the defect counts as a function of time. Initially, the introduction of a silicon interstitial results

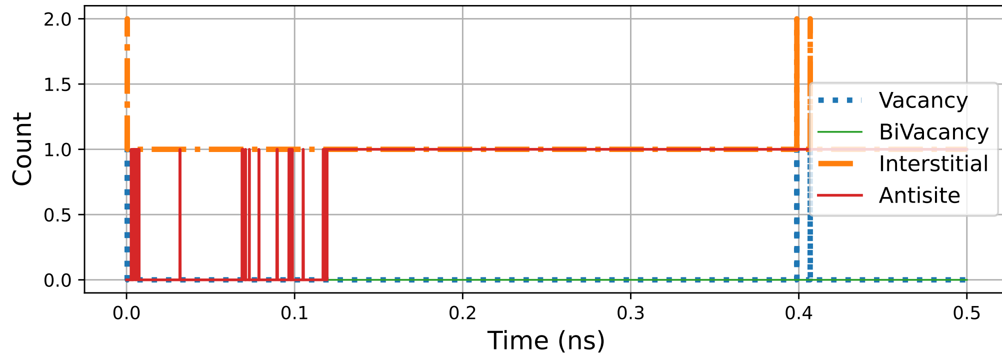


Figura 6.4: Time evolution of defect counts at 1600 K illustrating the transformation of a silicon interstitial (I_{Si}) into a silicon antisite (Si_C) accompanied by a carbon interstitial (I_C). In the early stage of the simulation, two interstitials and one vacancy coexist with a fluctuating antisite configuration. After approximately 0.12 ns, the system relaxes to a stable state characterized by one antisite and a single interstitial. The appearance of a vacancy whenever two interstitials are present highlights the dumbell formation.

in the transient formation of two interstitials, one vacancy, and a fluctuating antisite. After approximately 0.12 ns, the system stabilizes into a configuration consisting of a single antisite and one interstitial. The correlation between the presence of two interstitials and the appearance of a vacancy indicates that the interstitialcy mechanism is functional, where the I_{Si} displaces a lattice carbon atom, producing a Si_C antisite and a mobile carbon interstitial. This dynamic is clearly visible in the defect count evolution as shown in figure 6.4. This dumbbell configuration demonstrates well defined diffusion behavior, consistent with the expected mobility of carbon interstitials. In other words, the system naturally evolves so that the displaced carbon continues to diffuse in a dumbbell configuration, effectively mirroring the intrinsic carbon interstitial mechanism. Such a dumbbell type configuration has also been proposed by Gao et al. [111] based on atomistic simulations of interstitial defects in SiC, which provides additional support for the validity of our observations. The stability and reproducibility of this structure, along with its recurring role in

both intrinsic and interstitialcy mediated processes, underscore its central importance to diffusion in SiC.

Antisites and antisite-vacancy complexes Carbon antisite (C_{Si}) and silicon antisite (Si_C) defects are substitutional in nature and, by themselves, do not migrate. More relevant are their antisite-vacancy complexes. The CAV complex (a carbon vacancy adjacent to a carbon-on-silicon antisite) corresponds to the $V_C - C_{Si}$ configuration that forms when a V_{Si} transforms [54]. First principles free energy calculations reveal that the barrier for the conversion of V_{Si} into a $C_{Si}V_C$ complex ($\sim$ 1.3 eV) is comparable to that for divacancy (VV) formation ($\sim$ 1.5 eV)[1]. Moreover, the activation energy for the formation of $C_{Si}V_C$ from V_{Si} is lower than that for the migration of V_{Si} (barrier $\sim$ 3.0 eV). Thus, the formation of the $C_{Si}V_C$ complex is more likely than long range silicon vacancy migration. Notably, the free energy of $C_{Si}V_C$ is lower than that of V_{Si} by approximately 1.3 eV at 1500 K [1]. Rodrick Defo further calculated a forward energy barrier of 2.5 eV and a reverse energy barrier of 3.5 eV for this transformation [58]. Our NEB estimation of the CAV complex did not yield results sufficient to characterize antisite diffusion. Thus, we suggest that such complexes remain relatively immobile unless they undergo dissociation. Similarly, a hypothetical silicon antisite vacancy pair (often referred to as SAV, i.e., $V_{Si} + Si_C$) would represent the analogous defect on the silicon sub lattice however, V_{Si} typically does not persist long enough. Thus the barrier associated with SAV exhibit strange barrier shape as shown in section 4.5 In summary, antisite vacancy pairs are energetically unfavorable to migrate and thus contribute negligibly to diffusion during annealing process.

Taken together, both the literature and our simulation results indicate that interstitial mediated diffusion dominates self diffusion in 3C-SiC. Vacancies migrate only slowly, and V_{Si} is often deactivated by transforming into a CAV [54]. In contrast, carbon interstitials (I_C) with migration barriers of approximately 0.7 eV [131] and carbon vacancies (V_C) with barriers near 4.0 eV [54] are the primary defects driving diffusion under typical annealing conditions. At sufficiently high temperatures, even the comparatively slow V_C defects become mobile, whereas silicon-related defects remain largely immobile or convert into C-containing complexes. Our molecular dynamics simulations consistently show that any V_{Si} in-

troduced will rapidly capture a neighboring carbon atom, leading to the formation of a carbon vacancy and thereby establishing V_C diffusion as the dominant vacancy mechanism.

Therefore, the overall conclusion is that only carbon interstitials (I_C) and carbon vacancies (V_C) govern the evolution of the defect population during post implantation or annealing in 3C-SiC. All other intrinsic defects either migrate too slowly or transform into these carbon based defects, in agreement with DFT-NEB results reported in the literature.

6.3 Methodology for diffusion studies

The Einstein hypothesis of Brownian motion of particles establishes a formalism to calculate the macroscopic quantity of the diffusion coefficient from the microscopic quantity of mean square displacement. Complete theory is discussed in Section 3.7.

Current understanding of point defect diffusion in solids relies on Einstein's theory, which is a classical diffusion model. This model portrays defect transport as the hopping of individual point defects from one lattice site to another through interconnected diffusion pathways in the crystal structure. The crystal lattice determines the energy landscape of point defect migration. During defect diffusion, a mobile point defect traverses this energy landscape, with the highest energy point along the diffusion path defining the activation energy E_a of point defect diffusion. Achieving high defect diffusion σ necessitates a low activation energy E_a and a high concentration n_c of mobile defect carriers (such as vacancies or interstitials), which is proportional to $n_c \cdot \exp\left(-\frac{E_a}{k_B T}\right)$ at temperature T.

6.3.1 Exploring Diffusion Coefficient through Mean Square Displacement (MSD)

The diffusion coefficient can generally be measured by appropriately fitting the mean square displacement against time, as given by the Einstein relation 3.92:

diffusion coefficient for a given diffusion event at various temperature one can generate a plot correlating D with temperature (T). Subsequently implementation of a suitable fitting procedure enables to determine the unknown coefficients in Arrhenius equation 3.94 . This approach has been previously employed for the diffusion study[99] . We employed two statistical approaches to calculate the diffusion coefficient, taking into account the respective pros and cons. First one relays on collecting MSDs while second one is based on following COM movement and it is discussed in section 6.3.2. The mean square displacement (MSD) provides a measure of the average atomic displacement over time and is commonly used to estimate diffusion coefficients. A detailed derivation of the Einstein relation and its connection to diffusion is given in Section 3.8. We will add a correction by multiplying MSD with factor N to get the diffusion solely from the effective mobile species, details are described in section 3.8.

To collect the Mean Square Displacement (MSD) data, 2 ns NPT calculation was carried out followed by long NVT simulation. A time step of 0.1 fs was considered for all calculations. NPT calculation is important to perform the volume correction that will ensure correct MSD calculation. NVT calculations were conducted over a time span ranging from 50 ns to 100 ns, covering temperatures from 1300 K to 2400 K. The choice of temperature and duration is linked to the specific defect type, with a correlation based on the defect activation energy (V_C exhibiting a higher activation energy, approximately 4 eV, necessitating simulations at elevated temperatures to ensure statistical significance). Specifically, a temperature range of 1200 K to 1450 K was selected for $C_{Interstitial}$, while a broader temperature range (1600 K to 2350 K) was utilized for V_C. The simulation duration was adjusted to capture diffusion jump statistics within the range of 70 to more than 100 jumps. Complete details of NVT simulation and corresponding Diffusion results for V_C and $C_{Interstitial}$ at respective temperature can be found in table 6.1 One significant advantage of employing Mean Square Displacement (MSD) methods is their inherent simplicity in analysis. With MSD, individual diffusion events can be readily captured without the need to delve into the specifics of each jump event, as required in the Center of Mass (COM) technique, where counting individual jumps is necessary. The MSD method involves plotting the Mean Square Displacement and applying least square fitting procedure, which

effectively characterizes the diffusion process. This approach streamlines the analysis process and provides valuable insights into the dynamics of particle movement within the material system.

6.3.2 Defect evolution study using frequency rule method (implying COM)

An accurate and comprehensive description of atomic diffusion can be obtained by tracking the movement of specific defects over time and space. We conducted a similar statistical analysis based on MSD in the previous section 6.3. The MSD method offers a straightforward and rapid analysis while also capturing individual diffusion events, however, MSD calculations are inherently collective. A more precise and detailed description can be achieved by monitoring the center of mass (COM) motion, which accurately correlates with the evolution of individual defects within 3C-SiC. The complete theoretical framework for this description is detailed in section 3.9 . One can utilize equation 3.129 to calculate the diffusion coefficient [134] as explained in section 6.3.2. Thus it is straightforward to measure the number of jumps per unit time or jump frequency $\mathbf{\Gamma}$ for a given fixed distance $\Delta = 3.06 \times 10^{-8}$ cm to estimate $\mathbf{D}$ eventually.

The details for running simulation are explained in section 6.3.1. To determine the jump frequency, the COM coordinates were extracted throughout the simulation by implementing the COM computation.

Void−Volume analysis for vacancy defects

To identify and monitor the positions of vacancy defects, Voronoi cells around each atom in a bulk crystal was constructed using VORO++ [135, 136]. The position of the vacancy defect was defined by the centeroid of an empty Voronoi cell after aligning the atoms in the bulk crystal with those in the MD snapshot. 3D void volume at the vacancy defect site was determined by computing the distribution of voids on a discretized 3D grid with a resolution of 0.15 Å × 0.15 Å × 0.15 Å. The excluded volume for each atom was approximated as a sphere with a radius R (1.6 Å for carbon and 2.0 Å for silicon), where the ratio of the C and Si radii was consistent with the ratio of their respective Vander Waals radii. Voronoi

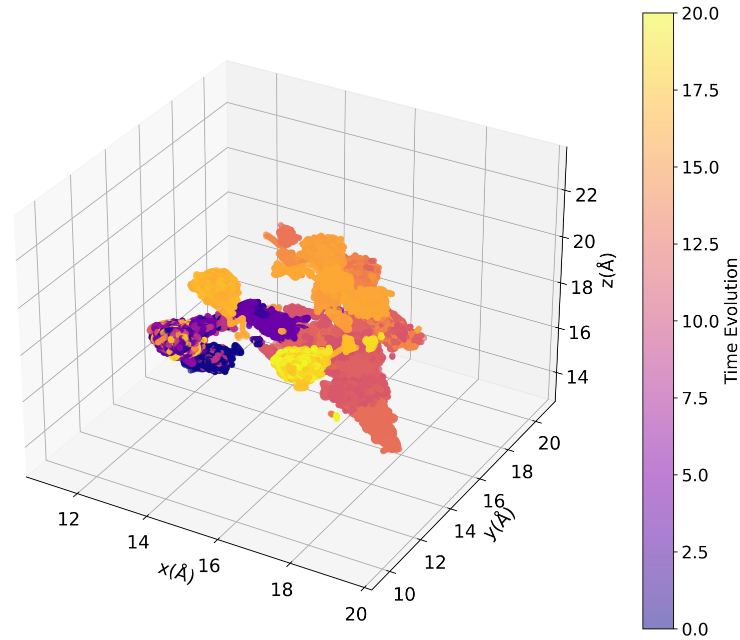


Figura 6.5: Zoomed 3D view of vacancy evolution in bulk SiC at 1750 K. The blue dots represent the crystal lattice structure, while the color bar indicates the temporal progression of carbon vacancies through inter lattice jumps. The representation of defect evolution shows a strong correlation with the center of mass (COM) dynamics as depicted in figure 6.6, and the mean squared displacement (MSD) evolution in figure 6.7. Each edge of the SiC simulation box measures 25 Å, with a small portion magnified for enhanced visibility.

compute per atom volume data that can be utilized for the calculation of desired defects and this is already implemented in LAMMPS. This data can also be utilized to provide a fantastic visualization of vacancy evolution in time that can be inspected by a 2D plot or a 3D animation. A snip of voronoi result in presented in figure 6.5 for just one case.

Jump Count analysis

Post-processing of COM data can be performed to count jumps and determine the jump frequency γ. The center of mass (COM) evolution in the total space is directly linked with defect movement in space. Due to the planar geometry of defect evolution in an FCC lattice, one can observe changes in just two axes (e.g., x and y, y and z, or z and x) at a time to define a jump. Visual inspection can be used to count the jumps within a given time frame.

The Rupture package [137] can be employed to perform this task effectively. It facilitates the offline detection of multiple change points in multivariate time series data. The detection algorithm relies on three key components: a cost function, a search method, and an external complexity penalty/constraint denoted as $\text{pen}(\cdot)$ (P2). We used the PELT method to perform segmentation analysis, introducing a penalty to balance the goodness-of-fit. The selection of the complexity penalty is crucial for determining the amplitude of changes to be detected. If the penalty is too low compared to the goodness-of-fit, numerous change points, including those resulting from noise, are detected. Conversely, if the penalty is too high, only the most significant changes, or potentially none, are identified. We optimized the penalty parameter and chose penalty $= 10$ for our study.

Rupture provides excellent functionality for counting jumps and saving respective times data. The time data was used to count inter jump intervals and study the distribution trend of jump data by fitting statistical distributions (Poisson, exponential). This will be discussed further in coming section. Rupture also offers excellent visual inspection tools to check for errors. One rupture result for V_C evolution during a 50ns NVT simulation at a temperature of 1800 K is shown in Figure 6.6. Each jump event is associated with a change in color, which aids in analysis

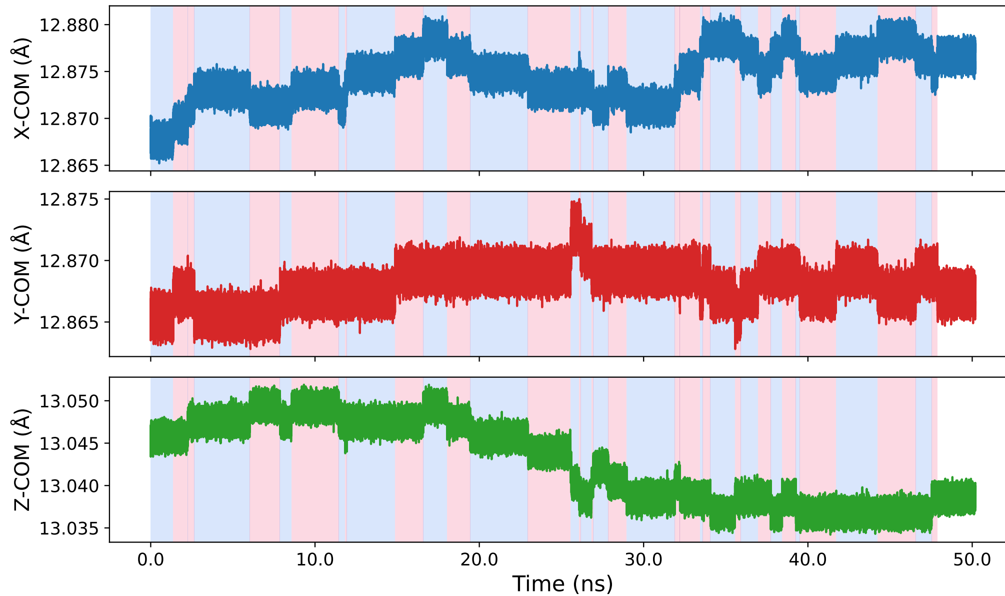


Figura 6.6: The three figures, arranged from top to bottom, illustrate the center of mass movement along the x, y, and z axes, respectively associated with V_C diffusion at 1800K during 50ns NVT simulation. The blue lines denote the relative displacement of the center of mass associated with the diffusion of vacancies (V_C) within the plane.It is notable that diffusion events involve simultaneous movement along two axes while the third axis remains unchanged, reflecting anisotropic diffusion characteristics.

and reduces effort significantly. At low temperature simulation simple visual inspection as shown in figure 6.8 is more optimal rather than rupture study. Graphical representation of the center of mass (COM) over time provides a clear depiction of the jump frequency. This can also be corroborated by observing the evolution of vacancy coordinates by voronoi analysis or video recording of the simulation for further validation and studying the kinetics of defects. Furthermore, center of mass coordinates can be re scaled to illustrate real time jumps in the crystal environment. Thus statistical theory of atomic transition allows specific tracking of each vacancy jump, and provide a simple way to study the nano (ti-

me and length) scale processes. One can see how interesting defects like divacancy can form as we will ply this for further analysis.

Calculation of statistical error in D

Most physical quantities typically cannot be estimated through a sole direct measurement but instead require a two step procedure. Initially, we measure one or more directly observable quantities from which the desired quantity can be calculated. Subsequently, we use the measured values of these quantities to compute the desired quantity. For instance, to determine the area of a rectangle, one measures its length (l) and height (h), then calculates the area (A) as $A = l \times h$. A little reflection reveals that almost all meaningful measurements involve these two distinct steps: direct measurement followed by calculation. In such cases of two step measurement, the estimation of uncertainties also involves a two-step process. First, we must estimate the uncertainties in the directly measured quantities. Then, we determine how these uncertainties propagate through the calculations to produce an uncertainty in the final result.

Depending on how the final result is obtained from the primary measurements (sum, difference, product, quotient), we use different formulas to determine the propagation of uncertainties. For the sum and difference of any number of quantities, the uncertainties in those quantities simply add.

Conversely, when we multiply or divide two measured quantities, x and y, the fractional uncertainty in the result can be calculated by summing the fractional uncertainties in x and y. In other words, when two quantities are multiplied or divided, their fractional uncertainties are added. For example, if $q = \frac{x}{u}$, then the uncertainty will be

$$\frac{\Delta q}{q} = \frac{\Delta x}{x} + \frac{\Delta u}{u} \tag{6.1}$$

Equation 6.1 applies to our case of the Einstein formula 3.79 for the calculation of the diffusion coefficient, allowing us to propagate the uncertainty in time to calculate the uncertainty in the diffusion coefficient $D = f_d \frac{\Gamma \Delta^2}{6}$ here Gamma is jump frequency is $\Gamma = \frac{No.of jumps}{totaltime}$ in the

following way

$$\frac{\Delta D}{D} = \frac{\Delta t}{t} \tag{6.2}$$

This requires us to find the uncertainty in jump times given by Δt

Estimation of Uncertainty in Hopping Times Δ t

Estimating the uncertainty in the diffusion coefficient D necessitates understanding the uncertainty in jump time statistics. The diffusion of defects or hops is a stochastic process, and uncertainties in such processes are termed random errors since all sources of uncertainty are inherently random. Uncertainty in random processes can be acquired by repeating measurements n times, a process commonly referred to as determining the mean or average (sd), which provides the most probable value. Consider measuring a quantity x, N times using the same equipment and procedures, resulting in N values $x_1, x_2, x_3, \ldots, x_N$. The best estimate of the measurement is given by the mean:

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \tag{6.3}$$

Given that we have calculated the best measure value, the next question arise about how a particular measure value i is deviated from the best measured value $\bar{x}$. This is given by difference or deviation $d_i = x_i - \bar{x}$. To measure the reliability of N measurement $x_1, x_2, x_3, ..., x_N$ one use the concept of deviation for calculation of what commonly known as Standard Deviation(SD) denoted by σ_x. SD can be described as

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{6.4}$$

To determine the SD in jump times, we need to extract the inter jump times between all consecutive hops. We can calculate the mean of the inter jump times data and use equation 6.4 to measure the uncertainty in the jump times calculations.

$$\Delta D = D\frac{\Delta t}{t} \tag{6.5}$$

Prescribed analysis is performed and the results are presented in table 6.2 as well as in figure 6.14.

Displaying Data as Histogram/Distribution

For measurements repeated multiple times, one of the key challenges is displaying or handling the obtained results appropriately. Utilizing a distribution or histogram is one of the convenient solutions. Since the mean/average can be given by Equation 6.3, it can be rewritten as:

$$\bar{x} = \frac{\sum_k x_k n_k}{N} \tag{6.6}$$

Moreover,

$$\sum_k n_k = N \tag{6.7}$$

Similar to summing over all values in Equation 6.3, here we again sum over all obtained values, multiplying each value by the number of times it occurred, or its respective frequency. This type of summation is referred to as a weighted sum, as each value x_k is weighted by the number n_k of times it occurred.

This can be further simplified by introducing the notion of fraction. We can define the fraction as:

$$F_k = \frac{n_k}{N} \tag{6.8}$$

This describes the fraction of N measurements that result in x_k. Equation 6.6 can be rewritten in terms of the fractional sum as:

$$\bar{x} = \sum_k F_k x_k \tag{6.9}$$

Each value x_k in the mean is weighted by the fraction of time F_k it occurred in the total results. Using Equation 6.7 we can write:

$$\sum_k F_k = 1 \tag{6.10}$$

This indicates that the sum of all fractions for all possible results x_k should be equal to 1. Equation 6.10 is also known as the normalization condition. An example of a bin histogram from the results is shown in Figure 6.9.

This is referred to as a bin histogram since the distribution of calculated data is represented as vertical bins above the measured values x_k. The fraction of measurements in each bin is equal to the area of the rectangle drawn above the bin. When the number of measurements increases, the histogram begins to take on a definite simple shape that can be represented as a continuous curve, commonly known as the limiting distribution. This curve is a theoretical construct that cannot be measured exactly, only by increasing the number of measurements infinitely can one make the histogram approximate the limiting distribution. This limiting distribution defines a function $f(x)$ whose value at any interval $(x, x + dx)$ can be obtained by integrating $f(x)dx$ between the respective limits.

Another more appropriate way to define this is following, $f(x)dx$ is the probability that a single measurement will yield a result between x and $(x + dx)$. Here is the conclusion of the discussion and the ultimate goal of this estimation: **If one knows the limiting distribution $f(x)$ of a measured quantity x, then the probability of obtaining x within any interval $a \leq x \leq b$ can be determined.**. Total probability of obtaining an answer between infinite bounds must be equal to unity. This argument impose an important condition to the limiting distribution given as

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{6.11}$$

In the present study, the limiting distribution of jump time data was obtained from molecular dynamics simulations and represented using histograms. These histograms were then fitted to theoretical models to assess the underlying statistical process. Specifically, the Poisson distribution was employed to describe the probability of observing discrete jump counts within a given interval, while the corresponding waiting time statistics were analyzed using the exponential distribution. The close agreement between the simulation data and these theoretical fits validates the assumption that diffusion jumps follow a memoryless Poisson process, with exponentially distributed waiting times.

6.4 2000 K Case Study: Carbon Vacancy Diffusion

In this section, we outline the methodology used to obtain dynamical statistics at a reference temperature of 2000 K, which serves as the model case. The same workflow is later extended to other temperatures and defect types. The analysis begins with the evaluation of the mean squared displacement (MSD), which is employed to estimate the diffusion coefficient of the system. In parallel, the center-of-mass (COM) trajectories of the defect are tracked to extract inter jump times, from which the corresponding jump frequency can be determined. These quantities provide a microscopic understanding of defect migration. Furthermore, the distribution of jump times is statistically analyzed by fitting to both exponential and Poisson models, allowing us to characterize the stochastic nature of the diffusion process and validate its agreement with random walk behavior. Estimated diffusion coefficient will further be utilized in Arrhenius equation to measure energy barrier and methods evaluation .

6.4.1 Setting up molecular dynamics

All the calculations were conducted through LAMMPS[138] and implying EDIP Force field. We started simulation by choosing a cubic super cell housing 1728 atoms to model bulk 3C-SiC lattice with one carbon vacancy (V_C). 3D periodic boundary conditions were employed as well. At first energy minimization was performed with a minimization criteria of e^{-6} eV. This provided a stable geometry to start. Equilibration of bulk 3C-SiC structure was achieved at target temperature(T) of 2000 K and 1 atm pressure in the NPT ensemble over 2 ns and utilizing time step of 1.0 fs. This provided volume correction to super cell. soon after further 2 ns of relaxation was performed with NVT ensemble. Nose Hoover thermostat and barostate was utilized to regulate the temperature and pressure respectively. After this 100 ns long production run was performed to collect the MSD data, COM data and voronoi [135, 136] data for prescribed super cell with (1) mono vacancy, (2) mono interstitial. The next step involved post processing on gathered data for further analysis.

6.4.2 Processing of mean square displacement (MSD) data

The mean squared displacement (MSD) analysis provides a fundamental approach to quantify defect mobility in crystalline systems. For the carbon vacancy (V_C) case, the MSD trajectory was computed over the entire simulation window and plotted as a function of time. A linear square fit (LSF) was applied to the MSD curve using the Brownian diffusion model, enabling the extraction of the diffusion coefficient via the Einstein relation. The fitting methodology is discussed in detail in Section 6.3, and a correction for the systematic underestimation of MSD-based diffusion was introduced following the approach outlined in Section 3.8. The final diffusion coefficient for the vacancy system is presented in Figure 6.7, providing a reference point for comparison with interstitial diffusion.

6.4.3 Processing Center of Mass (COM) data

The evolution of the defect's center of mass (COM) provides a robust way to identify individual migration events and to quantify the associated jump statistics. For the carbon vacancy (V_C) system, the COM trajectory was monitored throughout the production run. Each discrete shift in the COM position corresponds to a vacancy migration event, and the number of such jumps can be extracted directly from the trajectory. This information is then used to estimate the jump frequency, which, in turn, allows the diffusion coefficient to be calculated through Equation **??**.

To automate the detection of jumps, the rupture package [137] was employed to segment the COM data and identify discontinuities corresponding to migration events. In parallel, a simpler visual inspection of the COM plot was also performed for validation. The processed jump statistics provide a complementary measure of vacancy diffusion, and the results obtained using rupture are reported in Figure 6.8. we processed the data in chunks while analyzing with rupture, this is because rupture requirement for random access memory(RAM) increases linearly with data magnitude. Rupture results can be plotted for further validation of jumps frequency as presented in figure 6.6. Using $\Delta = 3.06e^{-8}cm$,we estimated a D value of $2.67e^{-7}cm^2/sec$ with COM evaluation. COM along with MSD for this simulation are presented in figure. 6.16

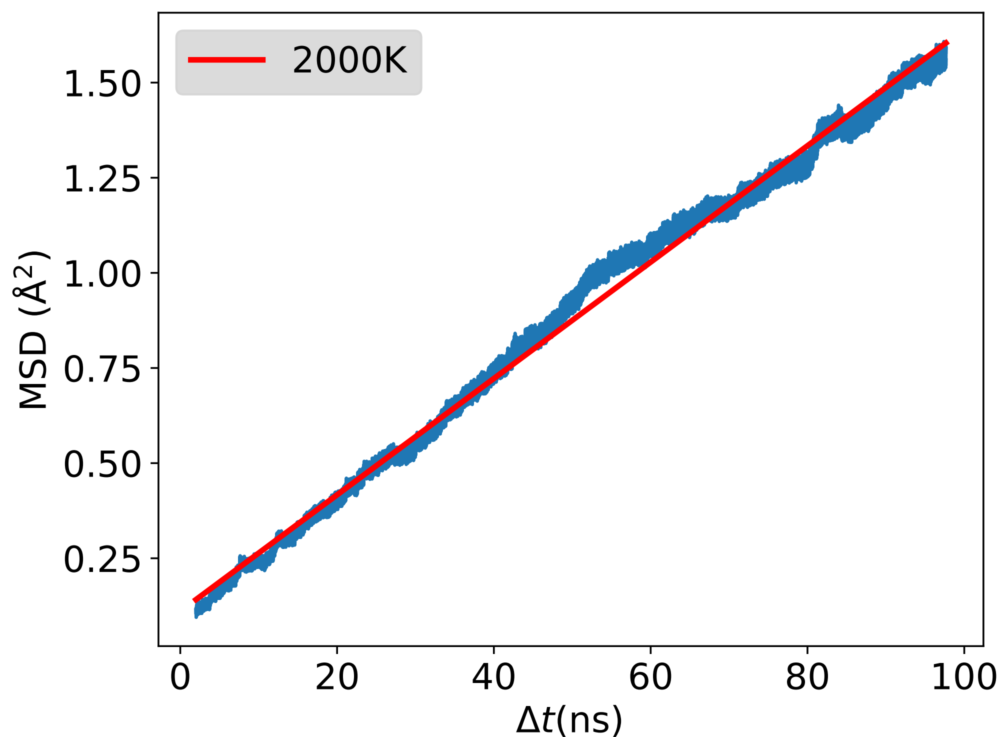


Figura 6.7: Mean squared displacement (MSD) as a function of time for carbon interstitial diffusion at 1200 K, obtained from 100 ns NVT molecular dynamics simulations. The linear region of the MSD indicates diffusive behavior, and the diffusion coefficient is determined from the slope using the Einstein relation, yielding $D \approx 1.91 \times 10^{-8}$ cm^2/s.

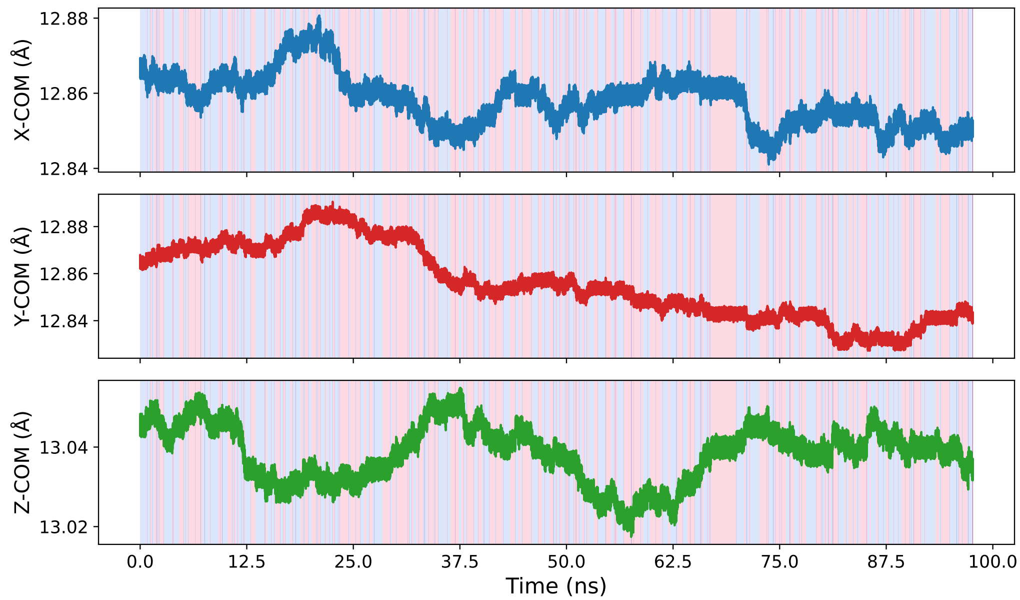


Figura 6.8: The three figures, arranged from top to bottom, illustrate the center of mass movement along the x, y, and z axes, respectively associated with V_C diffusion at 2000K during 100ns NVT simulation. The blue, red and green plot denote the relative displacement of the center of mass associated with the diffusion of vacancy (V_C) within the plane.It is notable that diffusion events involve simultaneous movement along two axes while the third axis remains unchanged, reflecting anisotropic diffusion characteristics.

6.4.4 Fitting of jump time data with Poisson distribution

To statistically characterize the migration behavior of carbon vacancies (V_C), the time intervals between successive jumps were analyzed. A Poisson process provides a natural framework for describing random, independent events occurring at a constant average rate. The probability of observing N jumps in a total time T can therefore be modeled by the Poisson distribution, provided that the jump rate per unit time, α, remains constant.

To implement this, the total simulation time T was divided into a large number of intervals, $N = \frac{T}{\Delta t} \gg 1$, where in each interval Δt, a jump event occurs with probability $p = \alpha \Delta t$. The probability of no event is then $1 - \alpha \Delta t$, while the likelihood of more than one event in Δt remains negligible.

Since in practice the distribution of vacancy jump times is continuous rather than discrete, the analysis was extended by considering the waiting-time distribution, which corresponds to an exponential probability distribution. This formulation provides a more accurate description of the temporal statistics of vacancy migration events and enables direct comparison of simulation results with theoretical random walk models.

6.4.5 Fitting of jump time data with waiting time/exponential distribution for V_C

The exponential distribution is widely employed to describe the time interval until a random event occurs and is therefore often referred to as the waiting time distribution. A key assumption of this model is the memoryless property, namely that the expected future waiting time is independent of the past. The probability density function is expressed as:

$$f(t) = \frac{1}{\tau} \exp\left(-\frac{t}{\tau}\right), \tag{6.12}$$

where τ represents the expected waiting time and t is the observed inter jump interval. In this model, τ is the only fitting parameter, obtained as the average of all measured waiting times. For the present vacancy study, we determined τ to be 3.28×10^{-10} s.

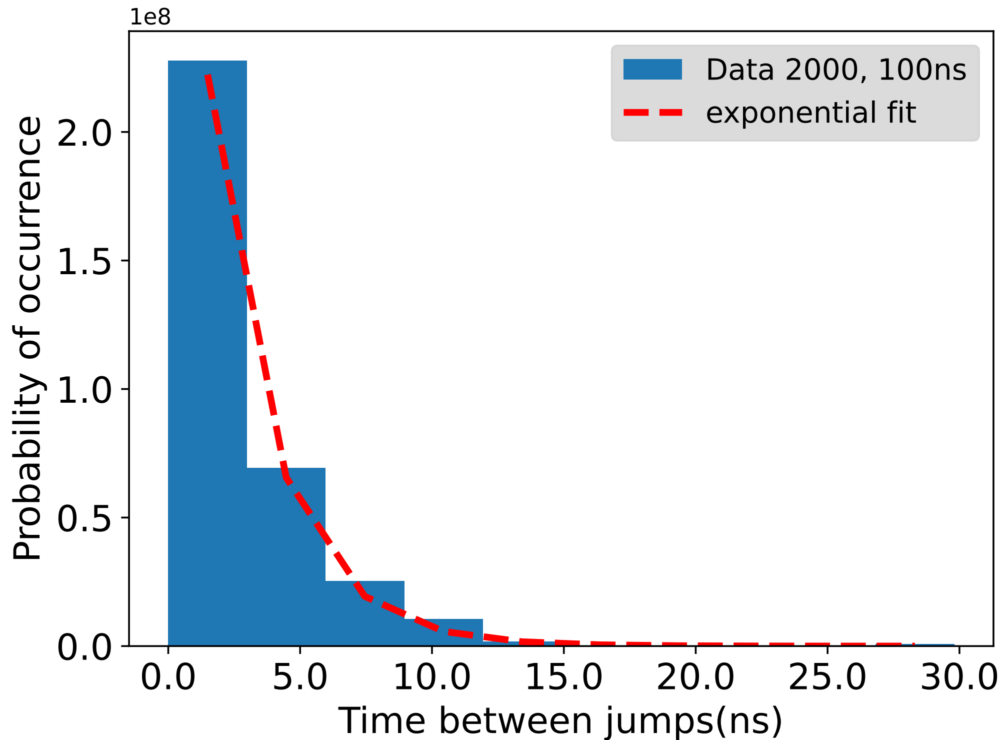


Figura 6.9: Analysis of time intervals between diffusion jumps at 2000K, based on statistical data collected over a 100 ns simulation. The inter jump times are fitted to an appropriate distribution, with results indicating that the data closely follows an exponential distribution. Additionally, the estimated standard deviation is provided for further statistical insight.

To quantify the distribution of inter jump times, the COM trajectories were analyzed to identify 170 vacancy migration events within a 100 ns simulation at 2000 K. Subtracting consecutive jump times yielded a total of 168 waiting time intervals. These intervals were assembled into a histogram and subsequently fitted with the exponential function, as shown in Figure 6.9. The results demonstrate close agreement between the simulation data and the exponential model, with the variance providing an estimate of statistical uncertainty. This approach not only characterizes the stochastic temporal behavior of vacancy jumps but also enables direct comparison with the Poisson process framework discussed earlier.

The analysis confirms that vacancy migration in 3C-SiC at 2000 K can be reliably modeled as a memoryless random process.

6.4.6 Energy Barrier estimation

The diffusion properties of point defects are strongly correlated with their respective activation energies. We will utilize the Arrhenius equation, $D = D_0 \exp\left(\frac{-E_A}{k_B T}\right)$, to determine the activation barrier. The theory underlying this approach is explained with details in Section 3.7. Since the activation barrier is estimated by considering diffusion parameters over a range of temperatures, we will elaborate on this estimation process in the subsequent sections.

6.5 1200 K Case Study: Carbon Interstitial Diffusion

In this part of the study, we focus on the behavior of carbon interstitials (C_I) in the 3C-SiC lattice at 1200 K. The introduction of an additional carbon atom as a point defect significantly alters the local atomic environment and diffusion characteristics, making it an essential counterpart to the vacancy case. To capture the underlying transport mechanisms, the molecular dynamics trajectories were analyzed by extracting mean squared displacement (MSD) and center of mass (COM) data, which were further employed to evaluate defect jump times. The resulting statistics were subsequently modeled using Poisson and exponential distributions to provide a quantitative framework for describing the stochastic nature of interstitial migration. The diffusion coefficient was further utilized in Arrhenius equation for barrier estimation.

6.5.1 Setting up Molecular dynamics simulation

All simulations were carried out using LAMMPS[138] with the EDIP potential. The system was initialized with a cubic supercell containing 1727 atoms to represent bulk 3C-SiC, into which a single carbon interstitial

(C_I) was introduced as an additional impurity atom. Periodic boundary conditions were imposed along all three dimensions. An initial energy minimization was performed with a convergence criterion of 10^{-6} eV, ensuring a well-relaxed starting configuration.

Subsequently, equilibration of the bulk 3C-SiC structure was conducted at a target temperature of 2000 K and ambient pressure (1 atm) within the NPT ensemble for 2 ns, using a time step of 1.0 fs. This stage allowed the supercell volume to adjust to equilibrium. Following this, an additional 2 ns relaxation run was performed under the NVT ensemble. Temperature and pressure control were maintained using Nose–Hoover thermostat and barostat algorithms, respectively.

A production run of 100 ns was then executed to accumulate mean squared displacement (MSD), center-of-mass (COM), and Voronoi data [135, 136] for the defected supercell containing (1) a monovacancy and (2) a mono-interstitial defect.

The collected data were subsequently processed for detailed analysis.

6.5.2 Processing of mean square displacement (MSD) data

The mean squared displacement (MSD) provides a direct measure of atomic mobility and is widely used to evaluate diffusion properties. In our analysis, the MSD data were plotted as a function of time, and a linear square fit (LSF) was applied to the data points using the Brownian diffusion model. The slope of this fit was then used to extract the diffusion coefficient according to the Einstein relation. Details of the fitting procedure are given in Section 6.3, while a correction for systematic underestimation was applied following the method described in Section 3.8. The resulting diffusion coefficient for the present case study is reported in Figure .

6.5.3 Processing of center of mass (COM) data

In the case of the carbon interstitial (C_I), the center of mass (COM) trajectory was analyzed to capture defect migration dynamics. Here, a sudden displacement in the COM trace corresponds to an interstitial jump between neighboring sites. By counting these events, the jump frequency

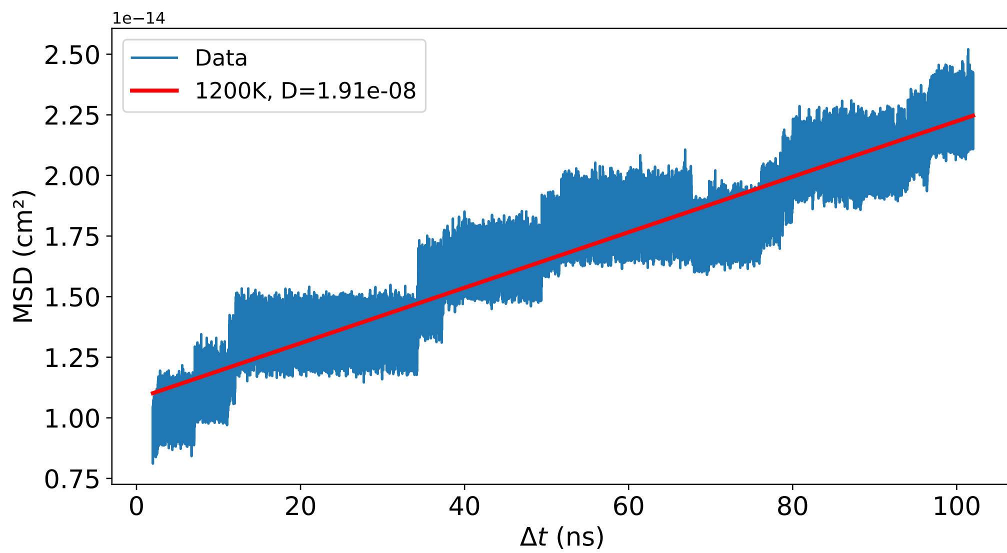


Figura 6.10: Mean squared displacement (MSD) as a function of time for carbon interstitial diffusion at 1200 K, obtained from 100 ns NVT molecular dynamics simulations. The linear region of the MSD indicates diffusive behavior, and the diffusion coefficient is determined from the slope using the Einstein relation, yielding $D \approx 1.91 \times 10^{-8}$ cm^2/s.

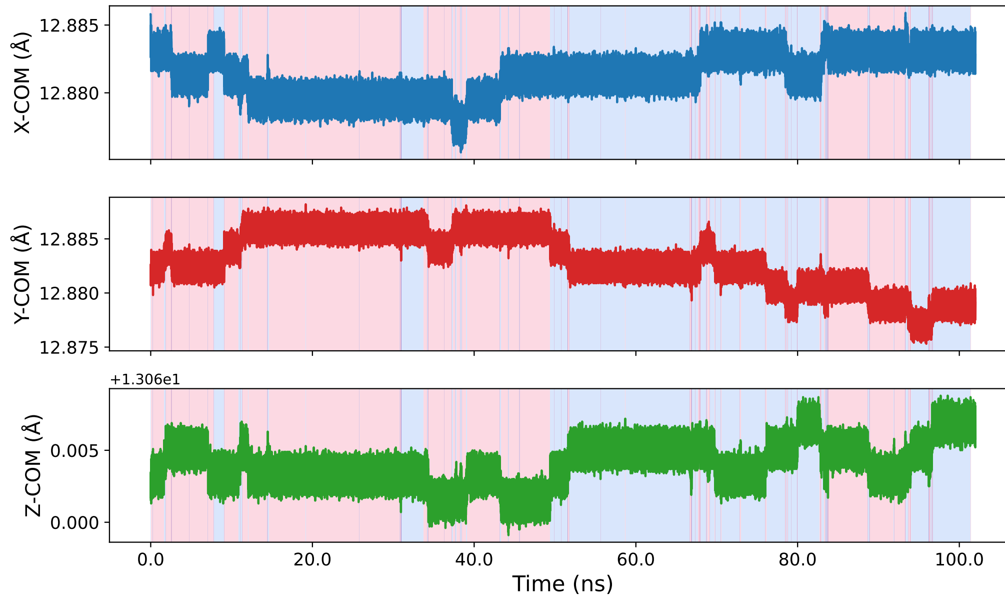


Figura 6.11: Center of mass trajectories along the x, y, and z directions for carbon interstitial diffusion at 1200 K, obtained from 100 ns NVT molecular dynamics simulations. The blue, red, and green curves represent displacement along the respective axes. The plots reveal that diffusion events typically involve correlated motion along two axes while the third axis remains nearly unchanged, indicating anisotropic diffusion behavior.

was determined and subsequently used to compute the diffusion coefficient through Equation **??**. As with the vacancy case, the rupture package [137] was utilized to detect discontinuities in the COM time series, enabling a reliable estimate of the number of interstitial jumps. Visual inspection of the COM trajectory was also performed to cross check the automated results. The extracted jump statistics provide insight into the stochastic migration of interstitials, and the outcomes of this analysis are summarized in Figure 6.11.

6.5.4 Fitting of jump-time data with Poisson distribution

To statistically characterize the migration behavior of carbon interstitials (I_C), the time intervals between successive jumps were analyzed. A Poisson process provides a natural framework for representing random, independent events occurring at a constant average rate. The probability of observing N jumps within a total time T can be described by the Poisson distribution, provided the jump rate per unit time, α, remains constant.

For implementation, the simulation time T was divided into a large number of intervals, $N = \frac{T}{\Delta t} \gg 1$, where in each interval Δt, a jump event occurs with probability $p = \alpha \Delta t$. The probability of no event is $1 - \alpha \Delta t$, while the likelihood of multiple events within Δt is negligible.

Because the actual distribution of interstitial jump times is continuous, the analysis was extended to the waiting time distribution, corresponding to an exponential probability distribution. This approach provides an accurate description of the temporal statistics of interstitial migration and enables direct comparison with theoretical random walk models.

6.5.5 Fitting of jump time data with waiting time/exponential distribution for C_I

For the carbon interstitial (C_I) system, the time intervals between successive migration events were also analyzed using the exponential (waiting time) distribution. This distribution is well suited for describing random processes where the probability of an event occurring in the next time interval is independent of its past history. The probability density function is defined as:

$$f(t) = \frac{1}{\tau} \exp\left(-\frac{t}{\tau}\right), \tag{6.13}$$

with τ denoting the mean waiting time and t the observed inter-jump interval. The parameter τ was determined directly as the average of the measured jump intervals. As before, the total time T was partitioned into $N = T/\Delta t \gg 1$ intervals, with the probability of a single jump in Δt given by $p = \alpha \Delta t$. The probability of observing no event is $1 - \alpha \Delta t$, while the chance of multiple jumps in the same interval is negligible.

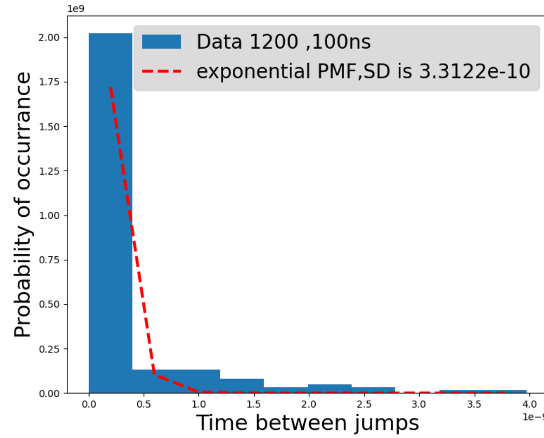


Figura 6.12: Distribution of time intervals between successive diffusion jumps for carbon interstitials at 1200 K, obtained from a 100 ns NVT molecular dynamics simulation. The data are fitted to an exponential distribution, and the corresponding standard deviation is reported to quantify the statistical spread.

However, since interstitial jump times are not strictly discrete and instead exhibit continuous variability, the waiting time formulation of the Poisson process was employed. This version is called*exponential distribution*, which accurately captures the temporal statistics of interstitial migration and provides a consistent framework for extracting jump frequency parameters from the molecular dynamics trajectories. From the COM trajectory analysis, a total of N interstitial jumps were detected within the 100 ns simulation at 1200 K, yielding $(N-2)$ independent waiting-time intervals. These intervals were binned into a histogram and subsequently fitted to the exponential distribution. An example of this fit is presented in figure 6.12, where the simulation data exhibit excellent agreement with the theoretical exponential model.

The results confirm that interstitial migration in 3C-SiC follows a stochastic process consistent with memoryless exponential behavior, providing a solid statistical foundation for estimating jump frequencies and diffusion coefficients.

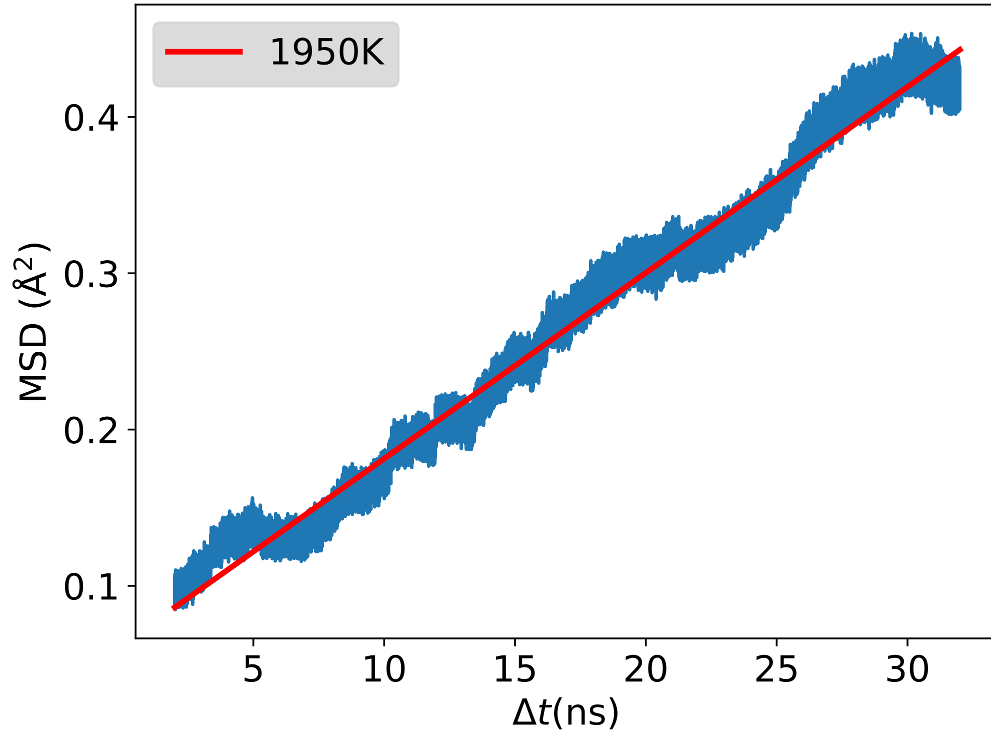


Figura 6.13: Mean squared displacement (MSD) versus time for carbon vacancy diffusion at 1950 K, obtained from 30 ns NVT molecular dynamics simulations. The linear region of the MSD is used to extract the diffusion coefficient according to $D = \frac{1}{2d}\frac{d\langle r^2\rangle}{dt}$ with $d = 3$ for bulk diffusion.

6.6 MSD Results

A detailed analysis of diffusion events for point defects has been conducted using the discussed method for defects such as V_C, V_{Si}, and I_C. Our observations indicate that these defects diffuse as independent single species, with respective diffusion coefficients provided in Table 6.1. Below, I present a sample result for V_C diffusion at 1950 K (Figure 6.13).

The MSD evolution is depicted smoothly, with each diffusion event indicated by a step on the graph. The inter jump time can be inferred

Temperature K	$D_{\text{Carbon Vacancy}}$ cm^2/s	Temperature K	$D_{\text{Carbon Interstitial}}$ cm^2/s
1600	1.85×10^{-8}	1150	1.3178×10^{-8}
1700	5.88×10^{-8}	1175	3.4811×10^{-8}
1800	1.28×10^{-7}	1200	1.9070×10^{-8}
1850	1.47×10^{-7}	1230	4.3312×10^{-8}
1900	2.78×10^{-7}	1250	9.1275×10^{-8}
1950	4.07×10^{-7}	1275	1.0720×10^{-7}
2000	2.65×10^{-7}	1300	1.0463×10^{-7}
2050	7.48×10^{-7}	1350	1.8350×10^{-7}
2100	2.73×10^{-7}	1375	1.9958×10^{-7}
2150	1.12×10^{-6}	1400	2.2938×10^{-7}
2200	1.64×10^{-6}	1430	3.0062×10^{-7}

Tabella 6.1: Comparison of diffusion coefficients (D) for carbon vacancies and interstitials over their respective optimal temperature ranges. The temperature ranges are chosen based on different diffusion barriers, for example, the range of 1600–2200 K is selected for carbon interstitials (C_I) due to the lower diffusion barrier and enhanced diffusion rates.

from the smooth line segments between consecutive steps. This analysis was repeated over a temperature range of 1600 to 2350 K to obtain comprehensive statistics.

Similar calculations of Mean Square Displacement (MSD) were performed for $C_{interstitial}$, and the diffusion coefficient was analyzed using the Einstein formula 3.79.

The scope of the analysis was extended to a temperature range from 1150 to 1400 K, in intervals of 50 K. In the MSD plot, diffusion events are clearly indicated by steps in MSD. Consecutive jumps appear as significant steps. The complete range of analysis is presented in table 6.1.

At higher temperature e.g. 2500K for V_C, system might loose equilibrium, we can expect a massive contribution from lattice vibrations and other thermodynamics factors. That argues stopping analysis below elevated temperature and also necessitated the melting point study that has been displayed in bench marking potential study in section 4.1 and figure 4.2. Here are suggestions to be considered while doing this analysis.

- Simulation should be large enough to capture a large number of diffusion events. One can relate simulated time with diffusion coefficient to fulfill this requirement.

- Least square Fitting to the Einstein relation should be limited to the linear region of MSD vs Δt plot

6.7 Diffusion Energetic

In this work, we systematically test and establish the procedures to obtain the diffusion properties and their statistical variances from Classical MD simulations. Statistical variance of diffusivity is directly linked with total number of ion hops during simulation.On this basis, MD simulations should be long enough to collect sufficient number of defect hopes. A minimum diffusivity of approximately $10^{-7}\,\text{cm}^2/\text{s}$ is necessary for a material to be accessible by MD simulations, and a diffusivity of around $10^{-5}\,\text{cm}^2/\text{s}$ may allow for a more precise calculation of D). Additionally this is also crucial to include the statistical error bound for each diffusivity point while fitting to Arrhenius relation/equation as explained in section 3.7. The error bound to D should be incorporated as weight to linear fitting of $\log(D)$ vs. $\frac{1}{t}$ as can be seen in figure 6.14. In the following sections we will present in detail estimation of diffusion energetics with respective barriers.

Energy Barrier Estimation for V_C

Certainly, it was ascertained that the migration of carbon vacancies occurs through second nearest neighbor jumps. systematic tests were performed to establish the procedure for obtaining the diffusion properties and respective statistical variance from said methodology and MD simulations. Employing the designated methodology, estimation of the diffusion coefficient (D) was undertaken for carbon vacancies across an extensive range of temperatures. Detailed insights of this study are presented in table 6.2.The resulting dataset was subsequently utilized within the Arrhenius equation to derive the energy barrier of diffusion and the pre exponential factor.

Tabella 6.2: Computational results for V_C diffusion in SiC across a temperature range. The table summarizes simulation times, number of jumps, average jump times, diffusion coefficients, standard deviation of jump times, and associated errors in diffusion coefficients.

Temperature (K)	Simulation Time (s)	No. of Jumps	Average Jump Time (s)	Diffusion Coefficient (cm^2/s)	SD_time (s)	Error_D
1600	600.0×10^{-9}	71	6.77×10^{-9}	1.08×10^{-7}	7.05×10^{-9}	1.26×10^{-9}
1700	600.0×10^{-9}	226	2.78×10^{-9}	3.43×10^{-7}	2.59×10^{-9}	1.48×10^{-9}
1800	100.0×10^{-9}	81	5.47×10^{-10}	7.37×10^{-7}	5.94×10^{-10}	4.38×10^{-9}
1850	100.0×10^{-9}	106	5.05×10^{-10}	9.64×10^{-7}	5.16×10^{-10}	4.98×10^{-9}
1900	100.0×10^{-9}	170	2.20×10^{-10}	1.55×10^{-6}	2.63×10^{-10}	4.06×10^{-9}
1950	100.0×10^{-9}	257	1.94×10^{-10}	2.34×10^{-6}	1.79×10^{-10}	4.18×10^{-9}
2000	100.0×10^{-9}	375	2.53×10^{-10}	3.41×10^{-6}	3.00×10^{-10}	1.02×10^{-8}
2050	100.0×10^{-9}	484	2.08×10^{-10}	4.40×10^{-6}	2.18×10^{-10}	9.60×10^{-9}
2100	100.0×10^{-9}	600	1.71×10^{-10}	5.46×10^{-6}	1.77×10^{-10}	9.66×10^{-9}
2150	100.0×10^{-9}	716	6.94×10^{-11}	6.51×10^{-6}	6.85×10^{-11}	4.46×10^{-9}
2200	100.0×10^{-9}	1,054	4.84×10^{-11}	9.59×10^{-6}	5.02×10^{-11}	4.81×10^{-9}
2250	100.0×10^{-9}	1,216	4.19×10^{-11}	1.11×10^{-5}	4.83×10^{-11}	5.34×10^{-9}
2300	100.0×10^{-9}	1,440	3.47×10^{-11}	1.31×10^{-5}	3.42×10^{-11}	4.48×10^{-9}
2350	50.0×10^{-9}	901	2.77×10^{-11}	1.64×10^{-5}	2.73×10^{-11}	8.95×10^{-9}

The graphical representation of the Arrhenius plot for carbon vacancies can be observed in 6.14. Statistical error can be studied in the form of error bar for each D value. Extending the scope of this analysis to encompass a broad temperature range and timescales spanning up to 100 ns facilitates the capture of significant diffusion events (a minimum of 70) and substantiates the accuracy of our barrier measurements. For two particular cases of 1600 K and 1700 K, simulation time was enhanced up to 600 ns. The detailed description of result statistics is presented in table 6.2. Similar analysis is performed by collecting the diffusion coefficient data from MSD at respective temperatures. The details of data can be found in table 6.1, and visual presentation is given in figure 6.14. Polynomial fitting was performed using SciPy, resulting in an activation energy value of 2.06 eV for the energy barrier associated with carbon vacancies in 3C-SiC. The outcomes of this calculation are visually represented in 6.14.

Energy barrier estimation for interstitial defects

Interstitial defects are inherently intricate due to the potential existence of multiple interstitial sites for a given defect. After analyzing the results, it was discerned that there is existence of only one stable configuration. In this configuration a carbon interstitial is bonded to another displaced carbon atom as first neighbor and form a rotating dumbbell shaped structure, that can be observed from figure 6.1.

This dumbbell configuration exhibits clear diffusion behavior. Remarkably, introduction of silicon interstitial immediately kicks out a carbon giving rise to SiC antisite, while the carbon continues to rotate in dumbbell shape arrangement. Thus, it also follows the same carbon interstitial behaviour. This kind of Dumbbell shape is also suggested by Gao [111] for interstitial, this further reinforce the validity of our findings. The configuration is also displayed in figure 6.1 and discussed ins section 6.2. Energy barrier determination was conducted following the methodology outlined in Section 6.7. This process involved performing Molecular Dynamics (MD) simulations using the NVT ensemble. The simulation covered a range of 11 distinct temperatures, spanning from 1150 to 1430 Kelvin. This temperature range differs from that of V_C due to the comparatively smaller diffusion barrier and larger diffusion coefficient. The

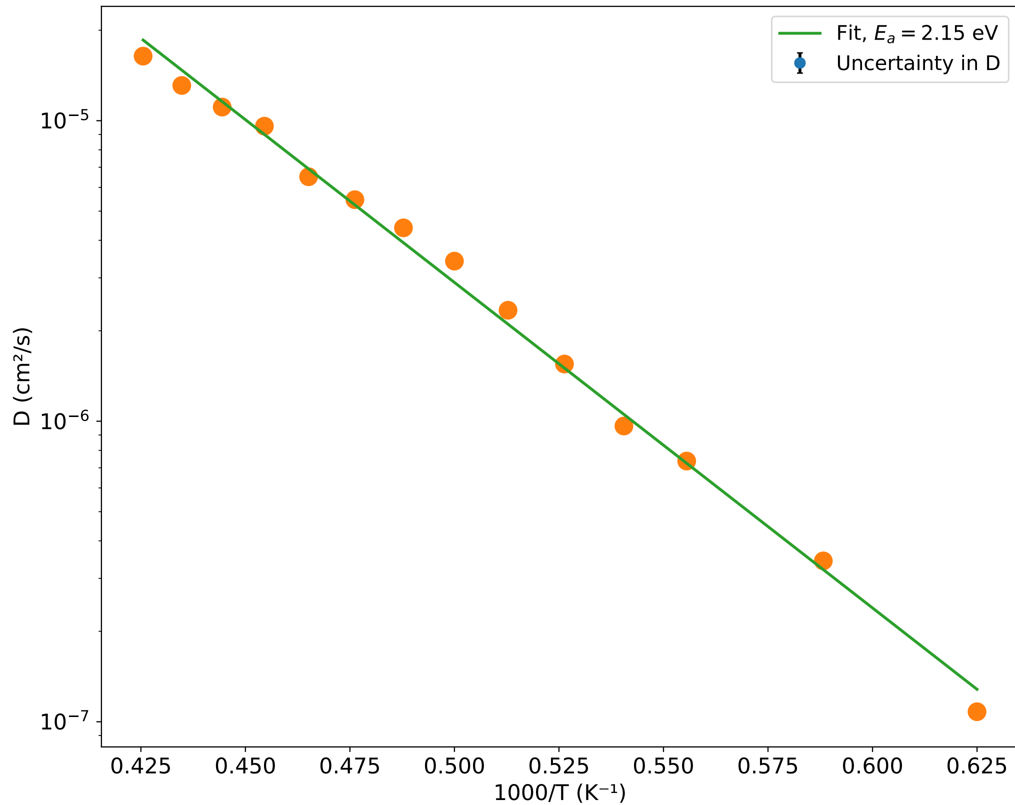


Figura 6.14: Plot of V_C diffusion coefficient versus temperature, used for energy barrier estimation based on the Arrhenius behavior of the **D vs. Temperature** relationship. The diffusion coefficient (D) is derived from the analysis of defect jump events over a 100 ns simulation period. Error bars represent the calculated statistical uncertainties in the diffusion coefficient at each corresponding temperature point.

details of diffusion, along with error estimation, are presented in Table 6.3.

Temp. (K)	Simulation Time (s)	No. of Jumps	Average Jump Time (s)	Diffusion Coefficient (cm^2/sec)	Error in Time (SD) (s)	Error in D (relative error)
1150	100.0×10^{-9}	107	4.27×10^{-9}	2.68×10^{-7}	8.00×10^{-9}	2.14×10^{-8}
1175	100.0×10^{-9}	122	3.32×10^{-9}	3.05×10^{-7}	6.66×10^{-9}	2.03×10^{-8}
1200	100.0×10^{-9}	136	3.31×10^{-9}	3.40×10^{-7}	6.65×10^{-9}	2.26×10^{-8}
1230	100.0×10^{-9}	165	2.51×10^{-9}	4.13×10^{-7}	4.74×10^{-9}	1.96×10^{-8}
1250	100.0×10^{-9}	198	1.86×10^{-9}	4.95×10^{-7}	3.73×10^{-9}	1.85×10^{-8}
1275	100.0×10^{-9}	232	1.77×10^{-9}	5.80×10^{-7}	3.54×10^{-9}	2.05×10^{-8}
1300	100.0×10^{-9}	278	1.49×10^{-9}	6.95×10^{-7}	2.74×10^{-9}	1.90×10^{-8}
1350	100.0×10^{-9}	301	1.04×10^{-10}	7.53×10^{-7}	1.98×10^{-10}	1.49×10^{-9}
1375	100.0×10^{-9}	368	1.01×10^{-9}	9.20×10^{-7}	1.78×10^{-9}	1.64×10^{-8}
1400	100.0×10^{-9}	403	9.29×10^{-10}	1.01×10^{-6}	1.67×10^{-9}	1.68×10^{-8}
1430	100.0×10^{-9}	440	7.21×10^{-10}	1.10×10^{-6}	1.26×10^{-9}	1.38×10^{-8}

Tabella 6.3: Computational details for interstitial defect diffusion across different temperatures, showing simulation time, number of jumps, average jump time, diffusion coefficient, and errors in time (SD) and diffusion coefficient (relative error).

Calculations were accomplished using a time step of 0.0001 ps and cumulative time for MD simulation amounted to 25-30 ns. Analysis at

elevated temperatures specifically in the range above 1430 K was omitted. Given that in the vicinity of melting point, it is less likelihood for system to display uniform behavior. Consequently, the computed energy barrier was given as 1.12 eV. The graphical representation of diffusion statistics along with uncertainty can be perceived in figure 6.15 .

6.8 Comparison of Methods, Einstein Diffusion formula (MSD) and frequency rule method (COM)

A comparative analysis of two statistical methods is presented to elucidate their respective diffusion validation approaches. The Mean Square Displacement (MSD) method, detailed in Section 6.3, involves estimating MSD over time to derive the diffusion coefficient. In contrast, the frequency rule method, discussed in Section 6.3.2, is grounded in statistical kinetic theory and offers more precise control over atomic jump events. It relates the diffusion coefficient (D) to the mean square of atomic jump displacements per unit time. Although both methods are designed to yield the same diffusivity D, based on the alignment between center of mass (COM) jumps (which reflect diffusion events) and mean square displacement (MSD) steps, discrepancies emerge in practice. A pictorial comparison of defect dynamics using the frequency-rule method versus the MSD based approach is shown in Figure 6.16 for I_C diffusion. A natural correspondence between the rise in MSD and COM jumps can be observed, as theoretically expected.

However, a clear mismatch for diffusion coefficient arises for both cases, as illustrated in Figure 6.17 likely due to unidentified factors.

6.8.1 Factors Affecting Diffusivity Estimates

This section outlines key factors influencing diffusivity estimates within their respective theoretical frameworks.

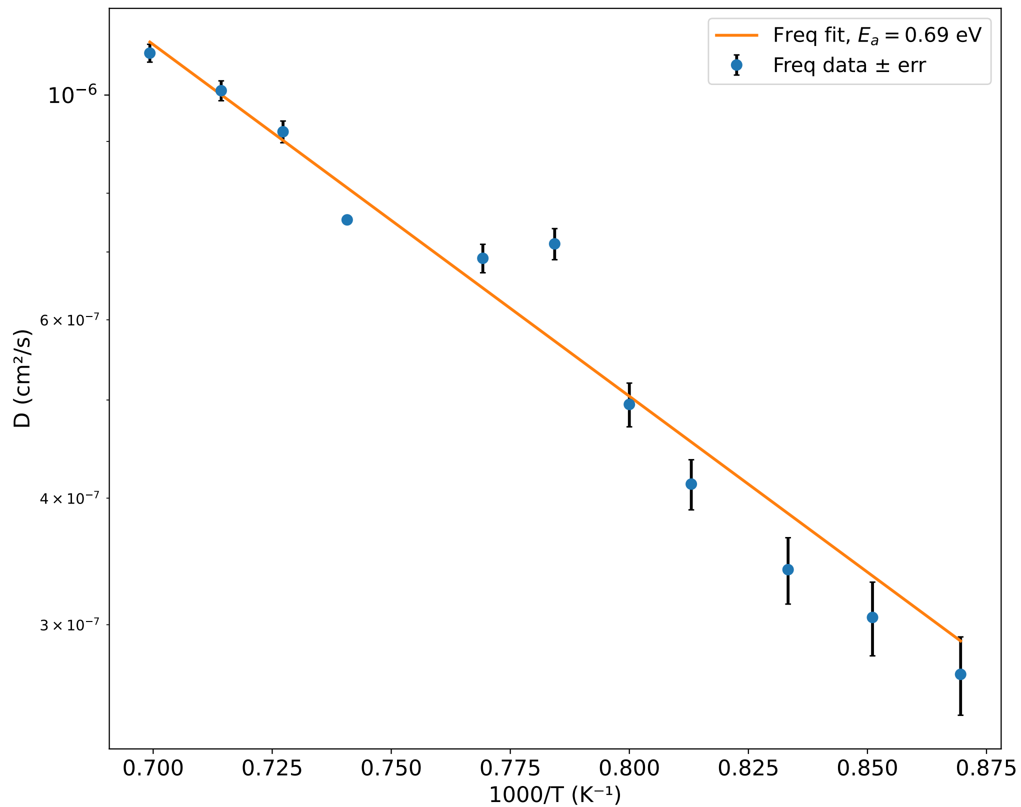


Figura 6.15: Diffusion coefficient for I_C as a function of temperature, illustrating the energy barrier estimation for C-Interstitial diffusion following the Arrhenius trend of **D vs. Temperature**. The diffusion coefficient (D) is computed by analyzing atomic displacement events over a 100 ns simulation. The associated statistical uncertainties in D are shown as error bars at each respective temperature measurement.

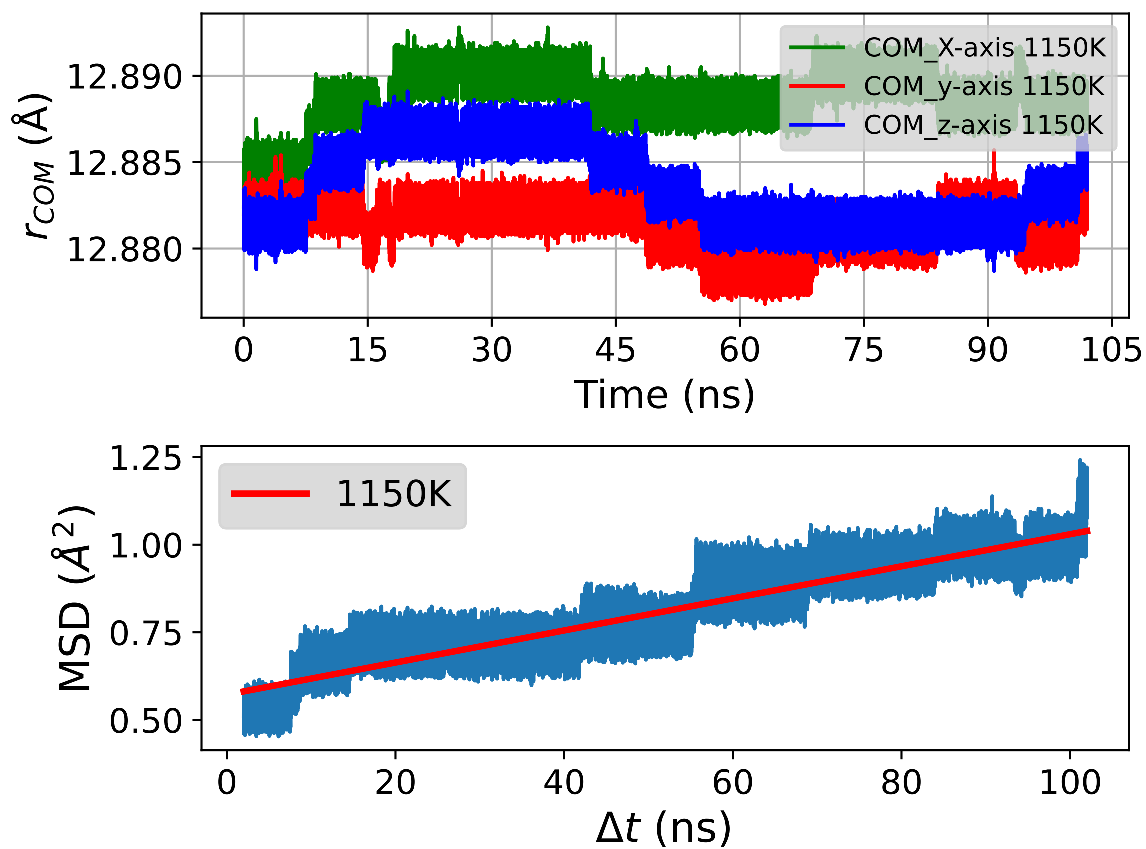


Figura 6.16: Figure make a comparison of MSD vs COM evolution for 100 ns simulation. I_C Diffusion jumps can be correlated in both figures.

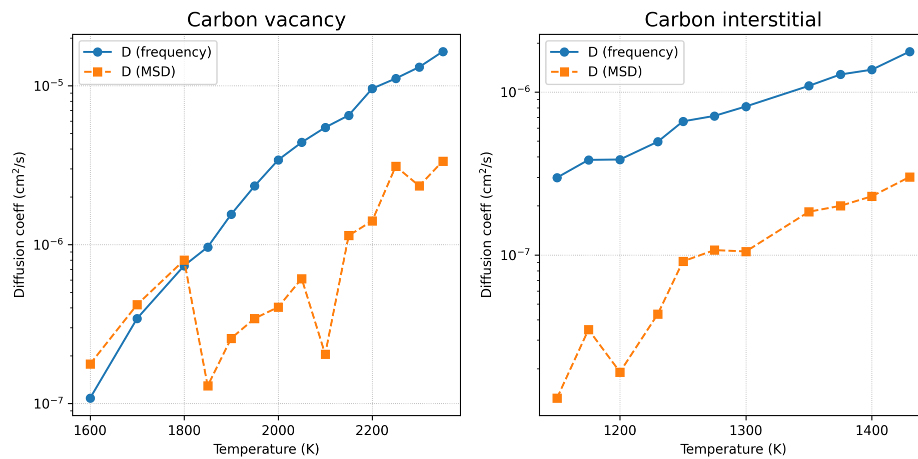


Figura 6.17: Temperature dependence of diffusion coefficients (D) estimated using two approaches: the frequency jump count model ("Frequency Rule") and the Einstein relation derived from mean squared displacement (MSD) for carbon vacancy and carbon interstitial diffusion. Both methods show increasing D with temperature , consistent with thermally activated diffusion. However, a marked divergence appears at higher temperatures for carbon vacancy diffusion (> 2100 K), where the frequency-based estimates significantly exceed those from the MSD approach, and overall for carbon interstitial suggesting methodological sensitivity to jump statistics or non-equilibrium effects.

6.8.2 Frequency Rule Method Potential Issues

Jump Counting Errors : Inaccurate identification of the number of jumps (Γ) due to algorithmic thresholding or sampling frequency can directly affect $D_{\text{frequency}}$. To mitigate this, all estimated jumps were manually verified via visual inspection. Although the error carries low statistical weight, it has been rigorously handled using the `ruptures` package.

Inconsistent Jump Lengths: Variability in jump lengths (ℓ) for example, due to non uniform spacing between interstitial or vacancy sites can introduce inaccuracies. However, this effect is minimal in the context of vacancy diffusion within a periodic lattice.

6.8.3 MSD Method Potential Issues

Non Random Jumps/back-jumps: Correlated or oscillatory motions reduce net displacement, leading to slower MSD growth. Consequently, D_{MSD} underestimates diffusion even when jump activity is high. In contrast, $D_{\text{frequency}}$ still counts these events as valid contributions. This is a hypothesis but a detailed voronoi analysis suggested no evidence of back/correlated jumps thus giving no physical grounds to this hypothesis.

Ensemble vs. Global Averaging: MSD is typically computed via ensemble or time averaging, whereas the frequency rule method employs a global average over all detected jumps. This methodological difference can result in varying levels of statistical noise.

Diffusion estimation using the MSD method is elaborated in Section 6.3, while the frequency rule approach is described in Section 6.3.2.

To assess overall drift or collective motion, we compute:

$$\text{MSD}(t) = \left\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \right\rangle, \quad \mathbf{R}_{\text{COM}}(t) = \frac{1}{N} \sum_{i=1}^{N} \left(\mathbf{r}_i(t) - \mathbf{r}_i(0) \right)$$

and visualize both:

- $\text{MSD}(t)$

- $|\mathbf{R}_{\mathrm{COM}}(t)|$

to compare net diffusion and drift (Figure 6.18). Distinct intervals of non-increasing MSD alongside COM fluctuations suggest local rather than net diffusive motion.

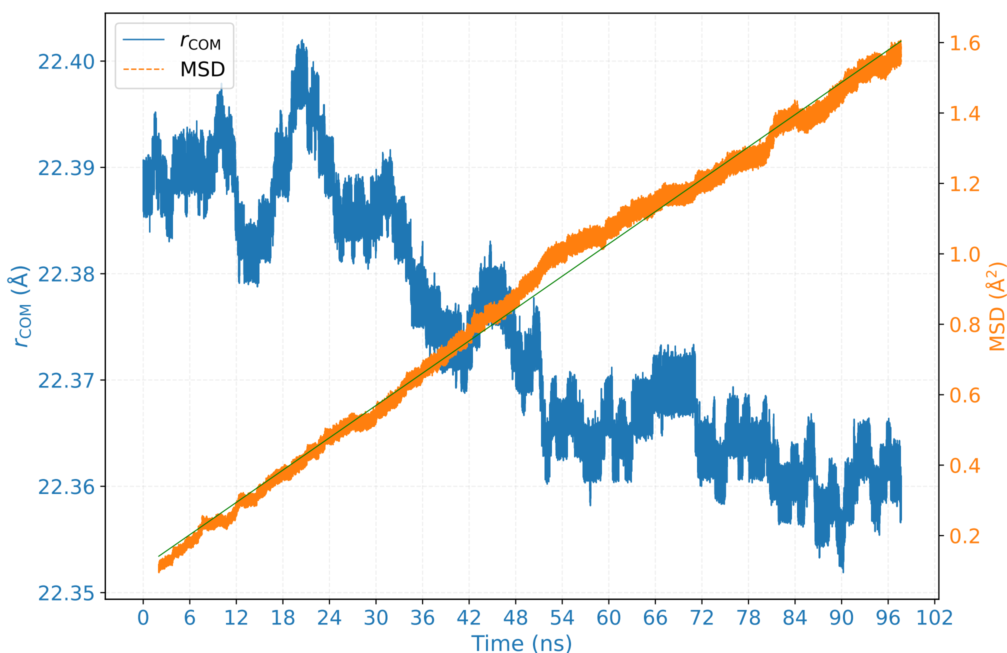


Figura 6.18: Comparison of mean squared displacement (MSD) and center-of-mass (COM) motion over time. MSD plateause.g., at 10 ns, 25 ns, 68 ns, and 85 ns indicate phases of negligible net displacement. However, concurrent COM displacements imply localized jumps, not captured by MSD, emphasizing the need for complementary diffusivity measures.

To remind MSD analysis includes a correction to account for collective averaging. Since MSD is calculated across all atoms, contributions from non diffusing atoms dilute the signal. Consequently, D_{MSD} underestimates the true diffusivity by a factor of N. This is corrected by multiplying with correction factor N, as outlined in Section 3.8. A comparison of corrected and uncorrected values is shown in Table 6.4 and Table 6.5.

Temperature (K)	$D_{\text{frequency}}$ (cm^2/s)	D_{MSD} (cm^2/s)
1600	1.08×10^{-7}	1.77×10^{-7}
1700	3.43×10^{-7}	4.19×10^{-7}
1800	7.37×10^{-7}	7.98×10^{-7}
1850	9.64×10^{-7}	1.29×10^{-7}
1900	1.55×10^{-6}	2.57×10^{-7}
1950	2.34×10^{-6}	3.43×10^{-7}
2000	3.41×10^{-6}	4.04×10^{-7}
2050	4.40×10^{-6}	6.11×10^{-7}
2100	5.46×10^{-6}	2.04×10^{-7}
2150	6.51×10^{-6}	1.14×10^{-6}
2200	9.59×10^{-6}	1.41×10^{-6}
2250	1.11×10^{-5}	3.11×10^{-6}
2300	1.31×10^{-5}	2.34×10^{-6}
2350	1.64×10^{-5}	3.35×10^{-6}

Tabella 6.4: Comparison of diffusion coefficients (D) for carbon vacancy (V_C) diffusion in 3C-SiC, estimated using two approaches: the frequency rule method (D_{FR}) based on jump statistics , and the Einstein relation derived from mean squared displacement (D_{MSD}). The values highlight methodological agreement at lower temperatures and reveal increasing divergence at higher temperatures, indicating sensitivity of the MSD method to dynamic averaging, noise, or non equilibrium effects. This comparison underscores the importance of method selection when characterizing diffusion processes across a broad temperature range.

Despite corrections, discrepancies in D for V_C diffusion necessitate deeper methodological justification, as discussed in Section 6.8.4.

Temperature (K)	$D_{\text{frequency}}$ (cm^2/s)	D_{MSD} (cm^2/s)
1150	2.68×10^{-7}	1.32×10^{-8}
1175	3.05×10^{-7}	3.48×10^{-8}
1200	3.40×10^{-7}	1.91×10^{-8}
1230	4.13×10^{-7}	4.33×10^{-8}
1250	4.95×10^{-7}	9.13×10^{-8}
1275	5.80×10^{-7}	1.07×10^{-7}
1300	6.95×10^{-7}	1.05×10^{-7}
1350	7.53×10^{-7}	1.84×10^{-7}
1375	9.20×10^{-7}	2.00×10^{-7}
1400	1.01×10^{-6}	2.29×10^{-7}
1430	1.10×10^{-6}	3.01×10^{-7}

Tabella 6.5: Diffusion coefficients of carbon interstitial in 3C-SiC at 1150–1430 K, obtained from the frequency rule ($D_{\text{frequency}}$) and mean squared displacement method (D_{MSD}).

6.8.4 Evaluation of Diffusivity Estimation Methods

To evaluate the variability in diffusivity estimates, we performed a large scale ensemble simulation of second-neighbor random walks in a 3C-SiC lattice using a Python based model. The simulation incorporated 10,000 independent trajectories, each with 2,000 jumps, and used a fixed mean waiting time ($\Delta t_{\text{mean}} = dt$ ns) extracted from previous MD simulations. Diffusivity (D_i) was computed for each trajectory under four stochastic models: (i) exponential waiting times without directional memory, (ii) Gaussian waiting times without memory, (iii) exponential times with directional back-jump bias, and (iv) Gaussian times with memory. This setup allowed a statistical comparison of MSD derived diffusivities against the theoretical prediction from the frequency rule.

Each simulation used the MD-extracted mean jump interval Δt_{mean}, and ensemble trajectories (10,000) of 2,000 jumps were generated to esti-

mate D_i from MSD. (iii) and (iv) test our proposed hypothesis of possible back jumps making the walk non random. This Analysis enabled statistical quantification of variability in D_{MSD} and comparison against the theoretical D_{FR}.

We demonstrate how the calculated diffusion coefficient D, derived from mean-square displacement (MSD), varies with changes in the average time step dt_mean or, equivalently, temperature T. Although the ensemble-averaged diffusivity $\langle D \rangle$ converges to the expected theoretical value given a large number of trajectories, individual molecular dynamics (MD) simulations often yield significant deviations due to finite sampling noise and the inherent randomness of atomic scale diffusion. Even under fixed simulation settings, single trajectory estimates can diverge notably from the mean, especially for short simulations or limited jump statistics. This highlights the necessity of averaging over many independent runs to obtain reliable diffusivity estimates. Our findings also clarify discrepancies between diffusivities derived via the frequency rule and those from the Einstein relation, offering a framework to quantify variability and assess uncertainty in computational diffusion analyses.

Our findings suggest that frequency based estimation is robust and accurate, especially under moderate temperatures. At elevated temperatures, increased jump frequency and data noise make jump counting more difficult, though still feasible. In contrast, MSD based estimation provides microscopic insight but suffers from variability across trajectories. While ensemble averaging can recover expected values, individual runs often deviate widely especially when temperature dependent dt_mean introduces uncertainty.

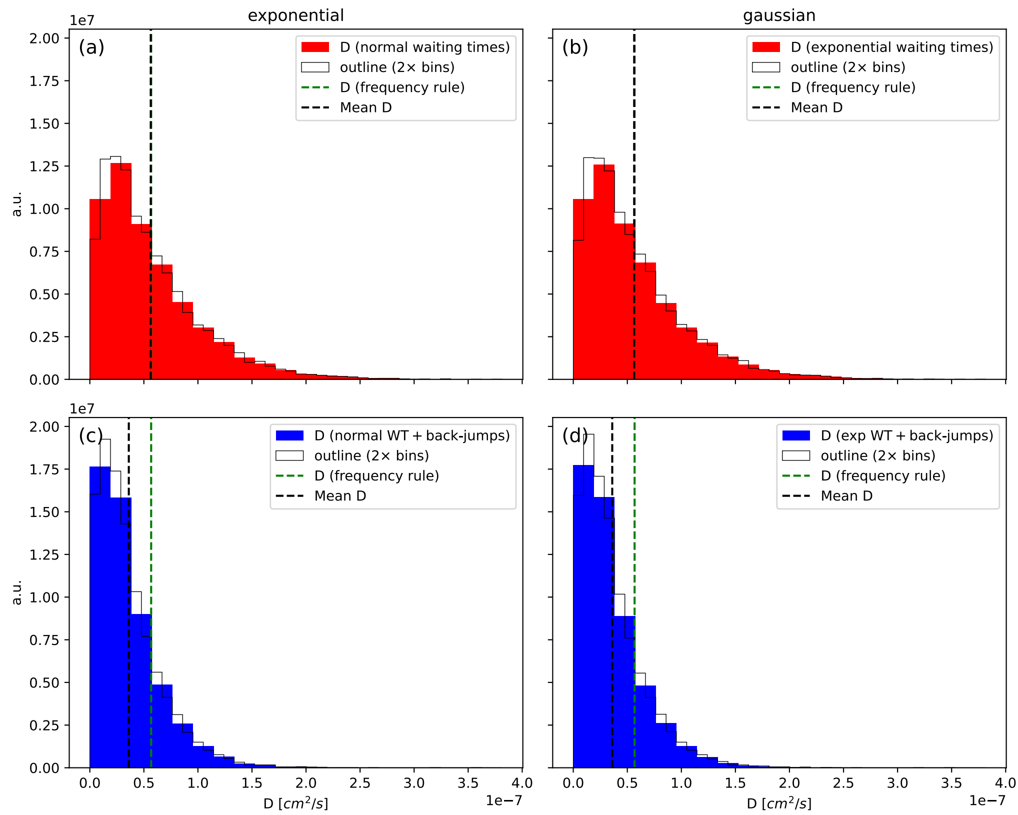


Figura 6.19: Distribution of diffusivity values (D) from ensemble random walks in 3C-SiC ($T = 1700$ K, $\Delta t_{\text{mean}} = 2.78$ ns). Each panel shows 10,000 trajectories under different stochastic models: (a) Exponential waiting times (unbiased), (b) Gaussian waiting times (unbiased), (c) Exponential with directional memory, (d) Gaussian with directional memory. Green dashed lines show theoretical D_{FR}, black dashed lines show $\langle D \rangle$. Directional memory alters the effective diffusivity.

Summary of Key Observations

- Jump counting statistics generally offer more accurate and robust diffusivity estimates, particularly at moderate temperatures.
- MSD based estimation is sensitive to statistical fluctuations and may show large deviations in single run simulations.

- Appendix results confirm that variability in MSD based diffusivity is intrinsic even when dt_mean is accurately known.

Based on these conclusions, we prioritize the frequency rule approach for primary diffusivity estimates, supported by MSD data as a secondary validation strategy.

6.9 Stabilization of defects systems, annihilation and complex formations

6.9.1 Generation of Divacancy

Among the many point defect complexes in SiC with potential qubit activity, two are of particular interest, the divacancy ($V_C V_{Si}$, often abbreviated as VV) and the carbon antisite vacancy pair ($C_{Si} V_C$). Several recent studies both experimental and theoretical have hinted that VV centers may emerge almost naturally during the dynamical evolution of irradiated or annealed SiC lattices [2, 1].

Here, the focus is not on forced defect creation but rather on whether such complexes can form incidentally under realistic thermodynamic conditions conditions similar to those experienced during post irradiation annealing. Two representative processes are examined: (i) the spontaneous formation of divacancies and (ii) the annihilation of interstitial vacancy pairs.

6.9.2 Divacancy (VV) Formation

To get a clearer sense of how VV complexes actually form, we set up molecular dynamics simulations in which two independent vacancies a carbon vacancy (V_C) and a silicon vacancy were introduced into the same supercell, about 6.3 Å apart. The temperature was fixed at 1600 K, not arbitrarily, but because it sits in an interesting window: high enough to allow noticeable defect motion, yet not so high that the lattice loses structural integrity. Below roughly 1400 K, most defects barely move, above

1800 K, things become too chaotic, with clustering and even partial sublimation of carbon atoms. At 1600 K, diffusion is active but still disciplined an ideal regime for observing defect interactions in something resembling thermal equilibrium.

At this temperature, the defects do not all move equally. Silicon vacancies (V_{Si}) remain stubbornly immobile due to their large migration barriers, often beyond 3 eV. Carbon vacancies (V_C), by contrast, show moderate mobility, typically requiring 2.0–2.2 eV to hop between sites. Carbon interstitials (I_C) are in a league of their own fast and restless, with barriers well below 1 eV, they can creep through the lattice even around 1200 K. Silicon interstitials fall somewhere in between. This mobility hierarchy $I_C > I_{Si} > V_C > V_{Si}$ naturally makes V_C the main actor in the eventual VV formation, drifting toward the relatively anchored V_{Si}.

Figure 6.21 shows how this process unfolds. For quite some time, the system's total potential energy barely changes. The vacancies move, but not because they are being "pulled" by any strong force instead, the motion appears to be a random, thermally driven wandering through a nearly flat energy landscape. In this regime, migration seems mostly entropic, more about temperature induced agitation than about an energetic incentive. Only when the two vacancies get within a few lattice spacings does the system suddenly take into account their mutual presence. The energy then drops, modestly but clearly, marking the formation of a more stable VV complex. The mean square displacement (MSD) rises sharply around the same moment, signaling the successful encounter. The final energy gain, about 1.2 eV, indicates that once the two vacancies are close enough, combining into a divacancy is indeed energetically favorable.

The same trend is captured in Figure 6.20, which tracks defect counts over time. The system starts with two well separated vacancies, confirmed by 3D visualization. After roughly 8 ns, a single divacancy signal emerges. Interestingly, this formation time is longer than what a simple diffusion estimate (based on the Einstein relation) would suggest roughly 0.5 ns for a direct 6.3 Å hop. The reason, as the trajectory plots later reveal, is that diffusion in a real lattice rarely follows a straight line. The carbon vacancy performs several detours, short jumps, and even partial reversals before finally settling next to V_{Si}, lengthening the overall timescale of coalescence.

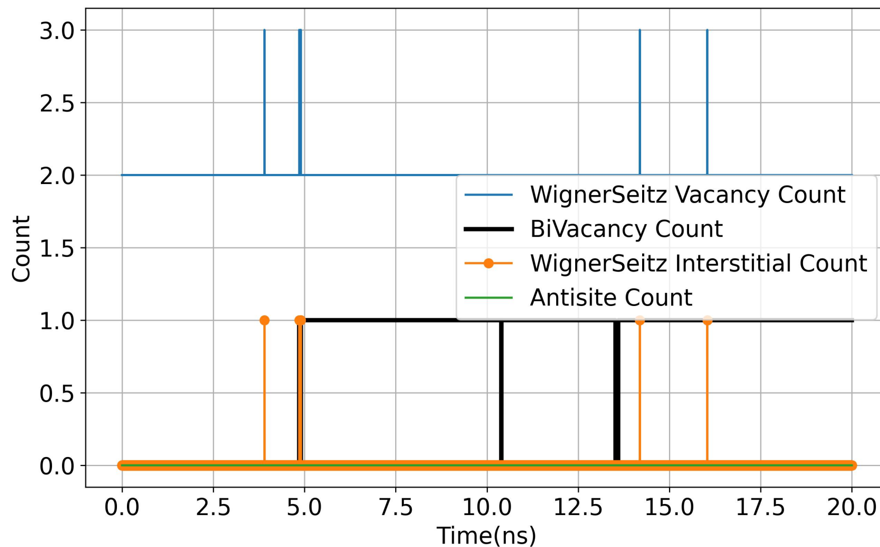


Figura 6.20: Time evolution of defect counts at 1600 K during divacancy (VV) formation. Initially, two independent vacancies (V_C and V_{Si}) are present in the supercell. Around 8 ns, the appearance of a VV complex is detected (black line), while the total number of vacancies remains constant since aggregation conserves the overall count.

The 3D trajectory mapping in Figure 6.22 tells the story even more clearly. The carbon vacancy meanders across the lattice, occasionally pausing or looping back, while the silicon vacancy stays rooted in place. There's something almost hesitant in the way V_C approaches V_{Si} as if the system explores several microstates before settling into the energetically preferred configuration. Notably, the energy landscape remains largely flat until the very end, the lattice doesn't "push" the vacancies together. Instead, thermal agitation does the work until proximity triggers a local stabilization event. This behavior supports earlier DFT studies that proposed VV centers as energetically stable configurations [2, 1]. What our molecular dynamics approach adds is the kinetic narrative: how long it takes, how indirect the path can be, and how much temperature governs the balance between mobility and recombination. At lower temperatu-

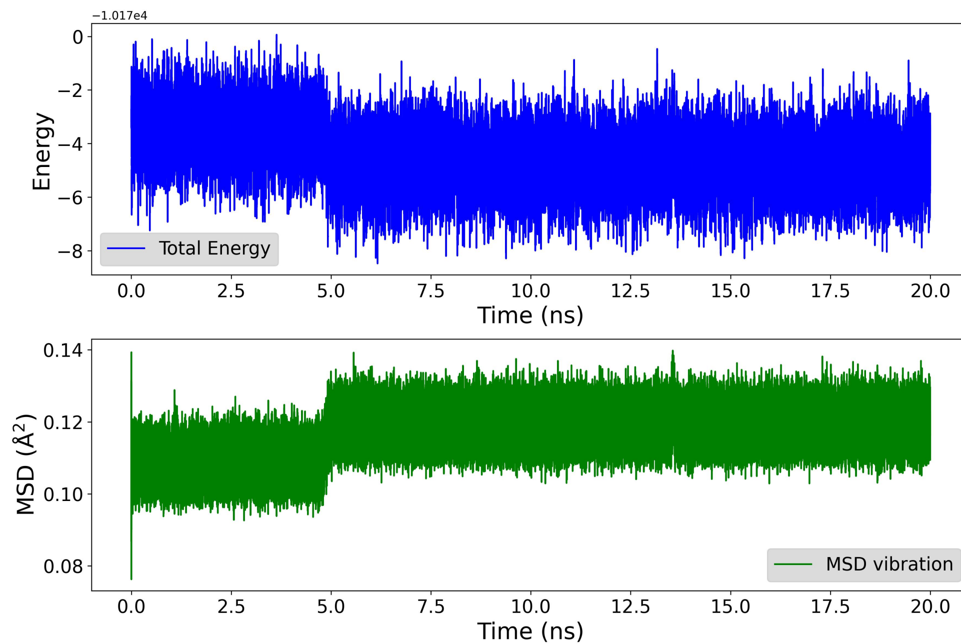


Figura 6.21: Energy and mean square displacement (MSD) profiles during VV formation at 1600 K. The MSD jump corresponds to V_C migration toward V_{Si}. The subsequent energy decrease reflects the thermodynamically favorable bonding of the two vacancies. The absence of early energy loss confirms that diffusion proceeds without a net driving force until recombination occurs.

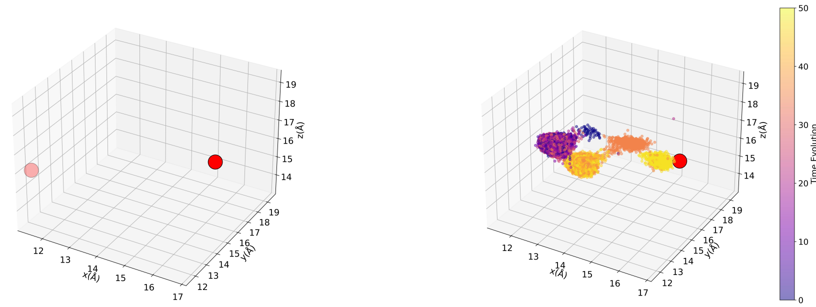


Figura 6.22: Three dimensional visualization of the VV formation process at 1600 K. (a) Initial configuration showing two vacancies separated by 6.3 Å. (b) Final configuration after coalescence. The color scale tracks simulation time, outlining the migration path of the carbon vacancy toward the stationary silicon vacancy.

res, these same defects would likely remain isolated indefinitely, at much higher ones, they might over diffuse and annihilate with interstitials instead. The chosen 1600 K, then, is not just a convenient midpoint but a kind of “sweet spot” where migration is visible yet controlled.

In the next subsection, we turn to the complementary mechanism interstitial–vacancy annihilation to see how these two processes, formation and recombination, jointly dictate the stabilization of the SiC lattice under thermal treatment.

6.9.3 Annihilation of Interstitial Carbon and Carbon Vacancy

The previous sections have shown that defect mobility in 3C-SiC depends strongly on both temperature and defect type. Among these, carbon interstitials (I_C) and carbon vacancies (V_C) stand out for their contrasting behaviors: one restless, the other almost inert. The interstitials, with comparatively low migration barriers, roam easily through the lattice, while vacancies move sluggishly, often remaining pinned unless temperature provides sufficient activation. This contrast becomes particularly significant at elevated temperatures, where diffusion does not merely drive aggregation or divacancy (VV) formation but can also lead to spontaneous

annihilation a process by which complementary defects meet and restore local crystalline order. The present study focuses on such an event, the annihilation of a carbon interstitial with a carbon vacancy, and how thermal energy facilitates this seemingly simple yet physically rich process.

The choice of 1400 K was deliberate. At lower temperatures (below $\sim$1200 K), I_C diffusion is still possible but slow, and V_C mobility is almost entirely frozen, making defect encounters improbable on nanosecond timescales. Raising the temperature much beyond 1600 K, however, introduces competing mechanisms: defect clustering, antisite formation, and even partial graphitization of carbon layers in some cases. 1400 K therefore represents a balanced regime high enough for interstitial motion to become statistically meaningful, yet moderate enough to prevent large scale lattice disorder. At this temperature, diffusion events can be observed clearly without the system drifting into uncontrolled defect dynamics or melting like behavior.

It is useful to recall the hierarchy of migration barriers that define defect mobility in SiC. Carbon interstitials (I_C) typically migrate with barriers around 0.8 eV, while silicon interstitials (I_{Si}) require nearly twice that energy ($\sim$1.6 eV). Vacancies, in contrast, are far more sluggish, V_C has a migration barrier of roughly 2.0-2.2 eV, and V_{Si} often exceeds 3 eV [55, 58]. This order of mobility $I_C > I_{Si} > V_C > V_{Si}$ implies that at moderate to high temperatures, diffusion in SiC is almost entirely governed by interstitial dynamics. Vacancies mostly serve as static traps, waiting for mobile counterparts to find them.

To explore this, the simulation cell was initialized with a single I_C–V_C pair separated by 7.85 Å. During the initial few nanoseconds, the total potential energy of the system remained nearly constant (Figure 6.23), even as the MSD indicated steady motion of the interstitial. This steady energy level is telling: it means that diffusion itself is not driven by an energetic gradient. The system neither gains nor loses potential energy as the interstitial wanders, motion occurs purely through thermal agitation. In other words, there is no long range "force" pulling the interstitial toward the vacancy only stochastic atomic vibrations enabling local hops through a mostly flat energy landscape.

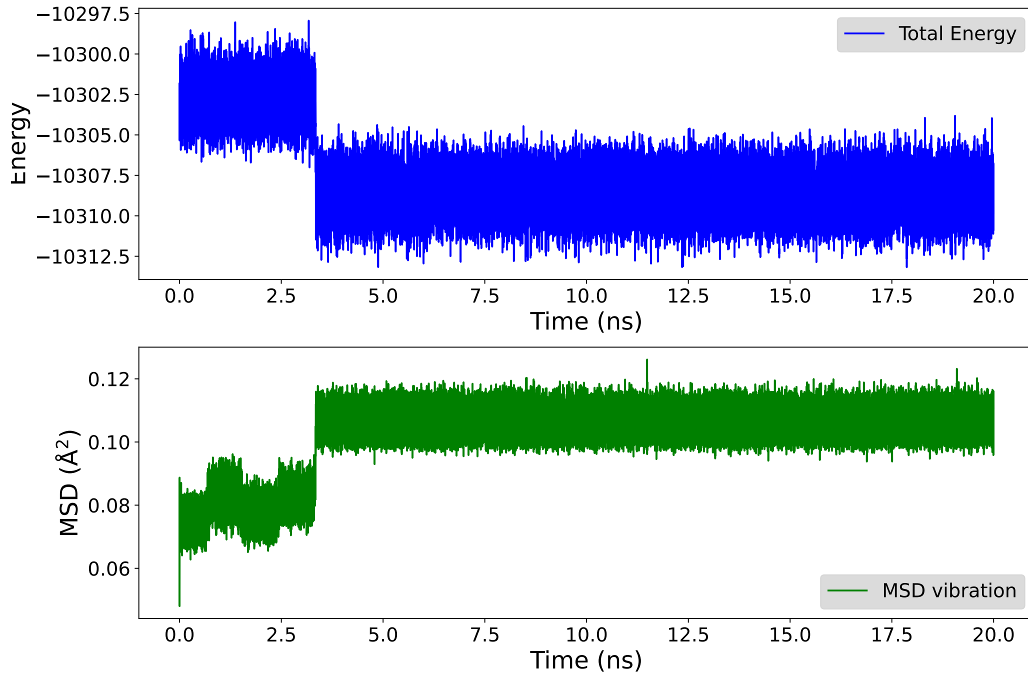


Figura 6.23: Correlation between potential energy and mean square displacement (MSD) during the recombination of I_C and V_C at 1400 K. The interstitial's migration produces a gradual MSD increase until ∼3 ns, followed by a sharp jump and simultaneous energy drop that mark the annihilation event.

Once the I_C happens to enter the capture radius of the V_C, however, the situation changes abruptly. Around 3 ns, both the MSD and the energy curves show synchronized transitions: the MSD spikes sharply as the interstitial jumps into the vacancy site, and the potential energy drops by about 5 eV. That discontinuity represents the system's shift from a metastable defect configuration to a relaxed, nearly perfect lattice. The recombination is both kinetically accessible and energetically favorable.

Figure 6.24 confirms this interpretation, one interstitial and one vacancy exist independently at $t = 0$, then both vanish from the defect inventory near 3 ns. The lattice, in effect, repairs itself. The energy released during this single annihilation event (roughly 5 eV) is substantially larger than that observed for divacancy formation (∼1.2 eV), emphasizing how

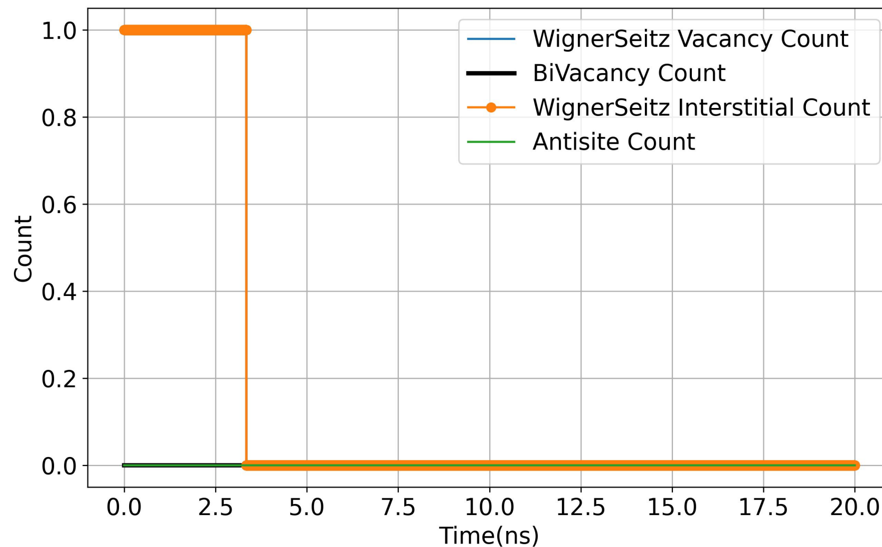


Figura 6.24: Temporal evolution of defect counts during I_C–V_C recombination at 1400 K. Initially, one interstitial and one vacancy exist independently. Around 3 ns, the two defects annihilate, releasing ∼5 eV and restoring lattice integrity.

recombination acts as a much stronger stabilization mechanism. This also explains why, at temperatures where both V_C and I_C are mobile, annihilation dominates over complex formation the kinetics simply favor the faster moving species finding and eliminating their slower counterparts.

Three dimensional trajectory mapping of this process (Figure 6.26) provides a vivid view of the mechanism. The carbon interstitial executes several short range hops, briefly stabilizing at metastable sites before making a final jump into the vacancy. The path is not direct rather, it appears diffusive and somewhat hesitant, suggesting a rugged potential landscape where the atom must navigate multiple local minima. The vacancy, on the other hand, remains nearly static throughout, as expected given its high migration barrier. The moment of recombination corresponds to a distinct lattice relaxation, visible as a local contraction and a smooth energy descent.

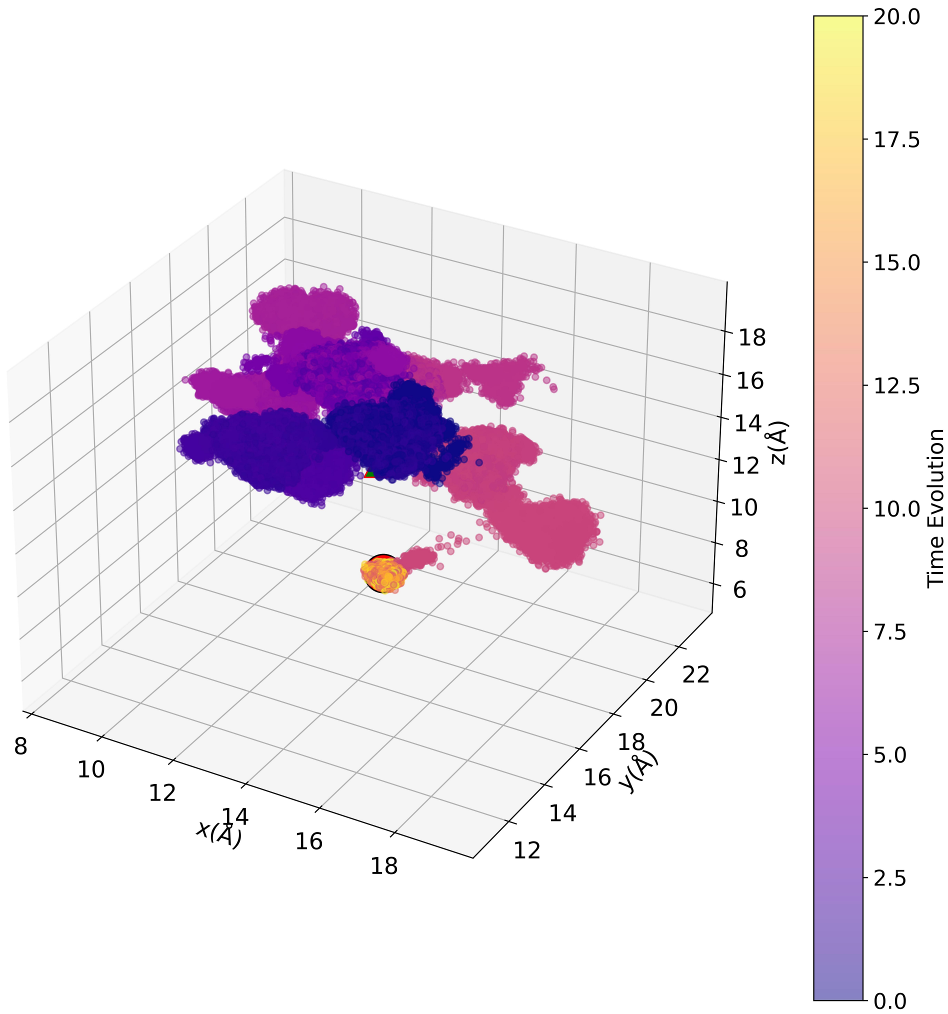


Figura 6.25: Three dimensional trajectory of I_C–V_C recombination at 1400 K. The color scale denotes simulation time. The interstitial undergoes multiple local hops before finally merging with the vacancy site, after which the lattice regains near perfect order.

From a broader perspective, these observations reinforce that defect mobility and annihilation in SiC are governed less by directional driving forces and more by probabilistic encounters. Diffusion supplies the opportunity, energy minimization follows only once two complementary

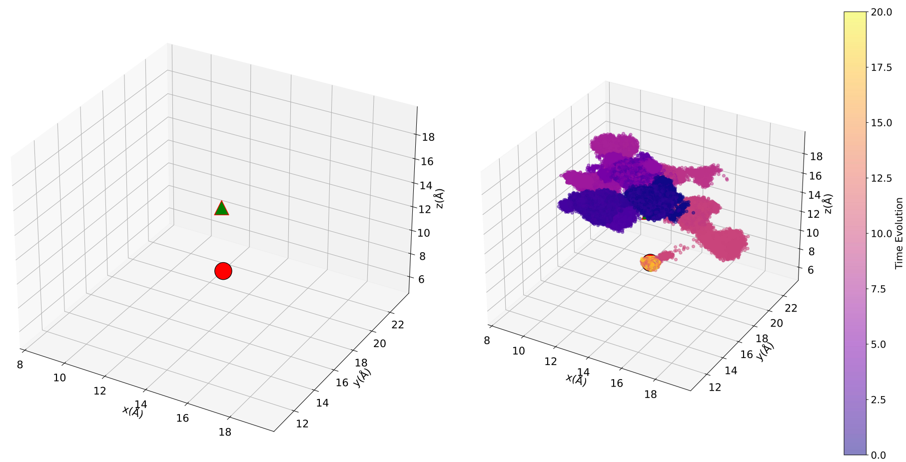


Figura 6.26: Temporal and spatial evolution of I_C to aggregate with V_C resulting in an energy gain of 5 eV. The temporal progress is represented by a color bar. Thus complete picture of annihilation process can be analyzed.

defects meet. The process, while stochastic in nature, plays a vital role in the material's long term stability, high temperature diffusion is not purely destructive but also the mechanism through which SiC self heals.

For simulations of radiation damage or post irradiation annealing, this balance between mobility and stability is crucial. The present results highlight that while vacancy clusters and divacancies may form under certain conditions (as discussed in Section 6.9.2), annihilation events such as I_C–V_C recombination often dominate the late stage recovery process. Together, these insights outline how defects in wide bandgap semiconductors evolve, compete, and eventually restore structural integrity under thermal treatment.

Conclusion

The formation and evolution of defects in SiC were investigated using molecular dynamics simulations with an environment dependent interatomic potential. Our findings highlight the probability and pathways for the aggregation of point defects, driven by their respective diffusion coef-

ficients, which can lead to the formation of spin bearing defect complexes relevant for quantum and electronic applications.

The discrepancy between D_{MSD} and $D_{\text{frequency}}$ estimates motivated further investigation using random walk simulations in the 3C-SiC lattice. These studies suggest that intrinsic noise, stochastic fluctuations contribute to the observed mismatch. While individual trajectory based diffusion estimates can diverge significantly, averaging over many trajectories yields convergence towards $D_{\text{frequency}}$. Non Markovian back jumps could also lead to different diffusivity estimates. However, this more complex mechanism has not clearly been evidenced in our analyses, and its eventual occurrence needs additional studies aimed at analyzing possible kinetic asymmetries of the free energy basin during the hopping.

The extracted diffusion coefficients were incorporated into the Arrhenius relation (Eq. 3.94) to estimate migration barriers. Results show that the carbon interstitial (I_C) exhibits a lower migration barrier compared to the carbon vacancy (V_C), consistent with its higher diffusivity. Furthermore, key defect processes were identified, including I_C–V_C annihilation and divacancy (V_C–V_{Si}) formation, which dominate the defect evolution under non equilibrium conditions.

Outlook and Future Directions

This study provides fundamental insights into defect dynamics in SiC under irradiation and implantation conditions. The combined use of atomistic simulations, statistical analysis, and trajectory based diffusion measurements establishes a framework for connecting microscopic defect motion with macroscopic transport properties. Nevertheless, several opportunities remain for improvement. Larger supercell sizes and longer simulation times would reduce finite size effects and better capture rare events such as extended defect formation. The role of charge states, electronic excitations also warrant future exploration.

From a broader perspective, understanding defect migration and recombination pathways in SiC has direct implications for improving the performance of power electronics and quantum devices, where radiation tolerance and defect engineering are critical. Future work integrating molecular dynamics with machine learning potentials could revolutioni-

ze this field by offering predictive capability for technologically relevant conditions. Specifically, this study can be extended by investigating thermodynamic parameters (e.g., optimal temperature favoring spin defect formation) using a combination of MD and DFT, thereby refining predictions for experimental validation by engineers. These predictions can be tested through in-situ implantation simulations. Other spin defects, such as the negatively charged boron vacancy (V_B^-) in hBN, are also actively studied, and the present methodology can be replicated for these candidate materials. Moreover, the diffusion study can be extended to applications in energy storage devices, such as electrode materials.

Bibliografia

[1] Elizabeth MY Lee, Alvin Yu, Juan J de Pablo, and Giulia Galli, Nature communications **12**, 6325 (2021).

[2] Cunzhi Zhang, Francois Gygi, and Giulia Galli, Nature Communications **14**, 5985 (2023).

[3] Yaolin Chen, Hongxia Liu, Tianzhi Gao, and Hao Wei, Micromachines **14**, 455 (2023).

[4] Yexin Fan, Zongwei Xu, Chengjuan Yang, Zhen Yang, Kun Zhang, and Shuxian Zheng, Ceramics International **49**, 26654 (2023).

[5] Christos Gennetidis, Patrice Chantrenne, and Thomas Wood, Physical Review B **109**, 075428 (2024).

[6] William F Koehl, Bob B Buckley, F Joseph Heremans, Greg Calusine, and David D Awschalom, Nature **479**, 84 (2011).

[7] David J Christle, Paul V Klimov, Charles F de las Casas, Krisztián Szász, Viktor Ivády, Valdas Jokubavicius, Jawad Ul Hassan, Mikael Syväjärvi, William F Koehl, Takeshi Ohshima, et al., Physical Review X **7**, 021046 (2017).

[8] Hosung Seo, Abram L Falk, Paul V Klimov, Kevin C Miao, Giulia Galli, and David D Awschalom, Nature communications **7**, 12935 (2016).

[9] Christopher P Anderson, Elena O Glen, Cyrus Zeledon, Alexandre Bourassa, Yu Jin, Yizhi Zhu, Christian Vorwerk, Alexander L Crook, Hiroshi Abe, Jawad Ul-Hassan, et al., Science advances **8**, eabm5912 (2022).

[10] Gary Wolfowicz, F Joseph Heremans, Christopher P Anderson, Shun Kanai, Hosung Seo, Adam Gali, Giulia Galli, and David D Awschalom, Nature Reviews Materials **6**, 906 (2021).

[11] David D Awschalom, Ronald Hanson, Jörg Wrachtrup, and Brian B Zhou, Nature Photonics **12**, 516 (2018).

[12] David M Toyli, Christoph D Weis, Gregory D Fuchs, Thomas Schenkel, and David D Awschalom, Nano letters **10**, 3168 (2010).

[13] Tobias Lühmann, Nicole Raatz, Roger John, Margarita Lesik, Jasper Rödiger, Marc Portail, Dominik Wildanger, Felix Kleissler, Kai Nordlund, Alexander Zaitsev, et al., Journal of Physics D: Applied Physics **51**, 483002 (2018).

[14] Tobias Lühmann, Roger John, Ralf Wunderlich, Jan Meijer, and Sébastien Pezzagna, Nature communications **10**, 4956 (2019).

[15] Robert Karsthof, Marianne Etzelmüller Bathen, Augustinas Galeckas, and Lasse Vines, Physical Review B **102**, 184111 (2020).

[16] Patrick Carlsson, NT Son, A Gali, J Isoya, N Morishita, T Ohshima, B Magnusson, and Erik Janzén, Physical Review B—Condensed Matter and Materials Physics **82**, 235203 (2010).

[17] Marianne Etzelmüller Bathen, José Coutinho, HM Ayedh, J Ul Hassan, Ildiko Farkas, Sven Öberg, Ymir Kalmann Frodason, Bengt Gunnar Svensson, and Lasse Vines, Physical Review B **100**, 014103 (2019).

[18] Christopher P Anderson, Alexandre Bourassa, Kevin C Miao, Gary Wolfowicz, Peter J Mintun, Alexander L Crook, Hiroshi Abe, Jawad Ul Hassan, Nguyen T Son, Takeshi Ohshima, et al., Science **366**, 1225 (2019).

[19] Felipe Fávaro de Oliveira, Denis Antonov, Ya Wang, Philipp Neumann, Seyed Ali Momenzadeh, Timo Häußermann, Alberto Pasquarelli, Andrej Denisenko, and Jörg Wrachtrup, Nature communications **8**, 15409 (2017).

[20] Tianze Sun, Zongwei Xu, Jintong Wu, Yexin Fan, Fei Ren, Ying Song, Long Yang, and Pingheng Tan, Ceramics International **49**, 7452 (2023).

[21] Yexin Fan, Ying Song, Zongwei Xu, Jintong Wu, Rui Zhu, Qiang Li, and Fengzhou Fang, Nanotechnology **33**, 125701 (2021).

[22] Yexin Fan, Zongwei Xu, Ying Song, and Tianze Sun, Diamond and Related Materials **119**, 108595 (2021).

[23] Wei Zhao, Zongwei Xu, Fei Ren, Bing Dong, Junlei Zhao, and Pengfei Wang, Diamond and Related Materials **132**, 109683 (2023).

[24] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, ADDRESS, 2002).

[25] Marcus W Doherty, Neil B Manson, Paul Delaney, Fedor Jelezko, Jörg Wrachtrup, and Lloyd CL Hollenberg, Physics Reports **528**, 1 (2013).

[26] Mouktik Raha, Songtao Chen, Christopher M Phenicie, Salim Ourari, Alan M Dibos, and Jeff D Thompson, Nature communications **11**, 1605 (2020).

[27] Yu He, SK Gorman, Daniel Keith, Ludwik Kranz, JG Keizer, and MY Simmons, Nature **571**, 371 (2019).

[28] Wayne M Witzel, Malcolm S Carroll, Andrea Morello, Łukasz Cywiński, and S Das Sarma, Physical Review Letters **105**, 187602 (2010).

[29] John W Steeds et al., Materials Science in Semiconductor Processing **9**, 622 (2006).

[30] Michael A Reshchikov and Hadis Morkoç, Journal of applied physics **97**, (2005).

[31] J Meijer, B Burchard, M Domhan, C Wittmann, Torsten Gaebel, I Popa, F Jelezko, and J Wrachtrup, Applied Physics Letters **87**, (2005).

[32] Shojan P Pavunny, Andrew L Yeats, Hunter B Banks, Edward Bielejec, Rachael L Myers-Ward, Matthew T DeJarld, Allan S Bracker, D Kurt Gaskill, and Samuel G Carter, Scientific reports **11**, 3561 (2021).

[33] Norikazu Mizuochi, T Makino, H Kato, D Takeuchi, M Ogura, H Okushi, M Nothaft, P Neumann, A Gali, Fedor Jelezko, et al., Nature photonics **6**, 299 (2012).

[34] Kevin Ohno, F J Heremans, Lee C Bassett, Bryan A Myers, David M Toyli, Ania C Bleszynski Jayich, Chris J Palmstrøm, and David D Awschalom, Applied Physics Letters **101**, 082413 (2012).

[35] Tim Schröder, Sara L Mouradian, Jing Zheng, Matthew E Trusheim, Michael Walsh, Edward H Chen, Linbo Li, Igal Bayn, and Dirk Englund, Nature Communications **8**, 15376 (2017).

[36] Yu-Chen Chen, Patrick S Salter, Sebastian Knauer, Laiyi Weng, Angelo C Frangeskou, Colin J Stephen, Shazeaa N Ishmael, Philip R Dolan, Sam Johnson, Ben L Green, et al., Nature Photonics **11**, 77 (2017).

[37] A Gruber, A Drabenstedt, C Tietz, L Fleury, Joerg Wrachtrup, and C von Borczyskowski, Science **276**, 2012 (1997).

[38] Fedor Jelezko, T Gaebel, I Popa, A Gruber, and Jorg Wrachtrup, Physical review letters **92**, 076401 (2004).

[39] Elke Neu, David Steinmetz, Janine Riedrich-Möller, Stefan Gsell, Martin Fischer, Matthias Schreck, and Christoph Becher, New Journal of Physics **13**, 025012 (2011).

[40] NB Manson, JP Harrison, and MJ Sellars, Physical Review B—Condensed Matter and Materials Physics **74**, 104303 (2006).

[41] Gopalakrishnan Balasubramanian, Philipp Neumann, Daniel Twitchen, Matthew Markham, Roman Kolesov, Norikazu Mizuochi, Junichi Isoya, Jocelyn Achard, Johannes Beck, Julia Tissler, et al., Nature materials **8**, 383 (2009).

[42] JR Rabeau, Patrick Reichart, Grigori Tamanyan, DN Jamieson, Steven Prawer, Fedor Jelezko, Torsten Gaebel, Iulian Popa, Michael Domhan, and J Wrachtrup, Applied Physics Letters **88**, (2006).

[43] Gary Wolfowicz, Christopher P Anderson, Andrew L Yeats, Samuel J Whiteley, Jens Niklas, Oleg G Poluektov, F Joseph Heremans, and David D Awschalom, Nature communications **8**, 1876 (2017).

[44] Alexander M Zaitsev, *Optical properties of diamond: a data handbook* (Springer Science & Business Media, ADDRESS, 2013).

[45] JR Weber, WF Koehl, JB Varley, Anderson Janotti, BB Buckley, CG Van de Walle, and David D Awschalom, Proceedings of the National Academy of Sciences **107**, 8513 (2010).

[46] Jörg Wrachtrup and Fedor Jelezko, Journal of Physics: Condensed Matter **18**, S807 (2006).

[47] Steven H Simon, *The Oxford solid state basics* (OUP Oxford, ADDRESS, 2013).

[48] Charles Kittel and Paul McEuen, *Introduction to solid state physics* (John Wiley & Sons, ADDRESS, 2018).

[49] Gang Zhang, Yuan Cheng, Jyh-Pin Chou, and Adam Gali, Applied Physics Reviews **7**, (2020).

[50] Jonathan H Skone, Marco Govoni, and Giulia Galli, Physical Review B **89**, 195112 (2014).

[51] L Gordon, Anderson Janotti, and Chris G Van de Walle, Physical Review B **92**, 045208 (2015).

[52] Pavel G Baranov, Ivan V Ilyin, Alexandra A Soltamova, and Eugene N Mokhov, Physical Review B—Condensed Matter and Materials Physics **77**, 085120 (2008).

[53] TL Petrenko and VP Bryksa, Physica B: Condensed Matter **517**, 47 (2017).

[54] Michel Bockstedte, Alexander Mattausch, and Oleg Pankratov, Physical Review B **68**, 205201 (2003).

[55] Michel Bockstedte, Alexander Mattausch, and Oleg Pankratov, Physical Review B **69**, 235202 (2004).

[56] Xiaopeng Wang, Mingwen Zhao, Hongxia Bu, Hongyu Zhang, Xiujie He, and Aizhu Wang, Journal of Applied Physics **114**, 123 (2013).

[57] R Rurali, E Hernández, P Godignon, J Rebollo, and P Ordejón, Computational materials science **27**, 36 (2003).

[58] Rodrick Kuate Defo, Xingyu Zhang, David Bracher, Gunn Kim, Evelyn Hu, and Efthimios Kaxiras, Physical Review B **98**, 104103 (2018).

[59] Emre Sevgen, Ashley Z Guo, Hythem Sidky, Jonathan K Whitmer, and Juan J De Pablo, Journal of chemical theory and computation **16**, 1448 (2020).

[60] Elizabeth MY Lee, Thomas Ludwig, Boyuan Yu, Aayush R Singh, François Gygi, Jens K Nørskov, and Juan J de Pablo, The Journal of Physical Chemistry Letters **12**, 2954 (2021).

[61] U Gerstmann, E Rauls, and H Overhof, Physical Review B—Condensed Matter and Materials Physics **70**, 201204 (2004).

[62] Fabien Bruneval and Guido Roma, Physical Review B—Condensed Matter and Materials Physics **83**, 144116 (2011).

[63] Mélanie Raine, Antoine Jay, Nicolas Richard, Vincent Goiffon, Sylvain Girard, Marc Gaillardin, and Philippe Paillet, IEEE Transactions on Nuclear Science **64**, 133 (2016).

[64] Antoine Jay, Mélanie Raine, Nicolas Richard, Normand Mousseau, Vincent Goiffon, Anne Hémeryck, and Pierre Magnan, IEEE Transactions on Nuclear Science **64**, 141 (2016).

[65] Antoine Jay, Anne Hémeryck, Nicolas Richard, Layla Martin-Samos, Mélanie Raine, Alexandre Le Roch, Normand Mousseau, Vincent Goiffon, Philippe Paillet, Marc Gaillardin, et al., IEEE Transactions on Nuclear Science **65**, 724 (2018).

[66] Andrew R Leach, *Molecular modelling: principles and applications* (Pearson education, Pearson,United Kingdom, 2001).

[67] Max Born and Robert Oppenheimer, *Quantum Chemistry: Classic Scientific Papers* (World Scientific, ADDRESS, 2000), pp. 1–24.

[68] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz, *Fluid Mechanics: Landau and Lifshitz: Course of Theoretical Physics, Volume 6* (Elsevier, baku Azerbaijan, 2013), Vol. 6.

[69] J Willard Gibbs, Yale University Press **0**, 1 (1902).

[70] Ludwig Boltzmann, *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie:(vorgelegt in der Sitzung am 8. Februar 1866)* (Staatsdruckerei, Germany, 1866).

[71] Peter Atkins, *Shriver and Atkins' inorganic chemistry* (Oxford University Press, USA, USA, 2010).

[72] JJPRB Tersoff, Physical review B **39**, 5566 (1989).

[73] J Tersoff, Physical Review B **41**, 3248 (1990).

[74] Hanchen Huang, Nasr M Ghoniem, Jimmy K Wong, and M Baskes, Modelling and Simulation in Materials Science and Engineering **3**, 615 (1995).

[75] G Lucas, M Bertolus, and L Pizzagalli, Journal of Physics: Condensed Matter **22**, 035802 (2009).

[76] Fei Gao, Eric J Bylaska, William J Weber, and L René Corrales, Physical Review B **64**, 245208 (2001).

[77] Paul Erhart and Karsten Albe, Physical Review B—Condensed Matter and Materials Physics **71**, 035211 (2005).

[78] Hugh O Pierson, Handbook of refractory carbides and nitrides **chapter 8**, 137 (1996).

[79] Jerry Tersoff, Physical Review B **38**, 9902 (1988).

[80] Joao F Justo, Martin Z Bazant, Efthimios Kaxiras, Vasily V Bulatov, and Sidney Yip, Physical review B **58**, 2539 (1998).

[81] George H Vineyard, Journal of Physics and Chemistry of Solids **3**, 121 (1957).

[82] Charles J Cerjan and William H Miller, The Journal of chemical physics **75**, 2800 (1981).

[83] Dzung T Nguyen and David A Case, The Journal of Physical Chemistry **89**, 4020 (1985).

[84] Wolfgang Quapp, Chemical physics letters **253**, 286 (1996).

[85] Thomas A Halgren and William N Lipscomb, Chemical Physics Letters **49**, 225 (1977).

[86] Michael J Rothman and Lawrence L Lohr Jr, Chemical Physics Letters **70**, 405 (1980).

[87] JC Polanyi and WH Wong, The Journal of Chemical Physics **51**, 1439 (1969).

[88] HANNEs JóNssoN, GREG MILL, and KARsTEN W JAcosEN, book:Classical and Quantum Dynamics in Condensed Phase Simulations **0**, 385 (1998).

[89] Richard E Gillilan and Kent R Wilson, The Journal of chemical physics **97**, 1757 (1992).

[90] Daniel Glen Sheppard, Rye Terrell, and Graeme A Henkelman, The Journal of chemical physics **128 13**, 134106 (2008).

[91] Hannes Jonsson, G. Mills, and {Karsten Wedel} Jacobsen, in *Nudged elastic band method for finding minimum energy paths of transition* (Enrico Fermi Summer School proceeding, ADDRESS, 1997).

[92] Vilhjálmur Ásgeirsson and Hannes Jónsson, Handbook of Materials Modeling (2020).

[93] Albert Einstein, Annalen der Physik **18**, 67 (1905).

[94] Albert Einstein et al., Annalen der physik **17**, 208 (1905).

[95] Igor M Sokolov and Joseph Klafter, Chaos: An Interdisciplinary Journal of Nonlinear Science **15**, (2005).

[96] Adolf Fick, Annalen der physik **170**, 59 (1855).

[97] Joseph Lefèvre López, Normand Mousseau, Gilles Adjanor, and Christophe Domain, Physical Review Materials **8**, 013609 (2024).

[98] Olga Natalia Bedoya-Martínez and Guido Roma, Physical Review B—Condensed Matter and Materials Physics **82**, 134115 (2010).

[99] XW Zhou, Reese E Jones, and Jacob Gruber, Computational materials science **128**, 331 (2017).

[100] Graeme Henkelman, Blas P Uberuaga, and Hannes Jónsson, The Journal of chemical physics **113**, 9901 (2000).

[101] Graeme Henkelman and Hannes Jónsson, The Journal of chemical physics **113**, 9978 (2000).

[102] Aiichiro Nakano, Computer Physics Communications **178**, 280 (2008).

[103] Emile Maras, Oleg Trushin, Alexander Stukowski, Tapio Ala-Nissila, and Hannes Jonsson, Computer Physics Communications **205**, 13 (2016).

[104] Priya Vashishta, Rajiv K Kalia, Aiichiro Nakano, and José Pedro Rino, Journal of applied physics **101**, 0 (2007).

[105] Ramaswami Devanathan, T Diaz De La Rubia, and William J Weber, Journal of nuclear materials **253**, 47 (1998).

[106] James F Ziegler, Matthias D Ziegler, and Jochen P Biersack, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms **268**, 1818 (2010).

[107] M Iuga, G Steinle-Neumann, and J Meinhardt, The European Physical Journal B **58**, 127 (2007).

[108] German D Samolyuk, YN Osetsky, and Roger E Stoller, Journal of Nuclear Materials **465**, 83 (2015).

[109] EL Kern, DW Hamill, HW Deem, and HD Sheets, *Silicon Carbide–1968* (Elsevier, University Park, Pennsylvania, 1969), pp. S25–S32.

[110] L Torpo, Risto M Nieminen, KE Laasonen, and S Pöykkö, Applied physics letters **74**, 221 (1999).

[111] Fei Gao, William J Weber, Matthias Posselt, and V Belko, Physical Review B **69**, 245205 (2004).

[112] Mark T Robinson, Radiation Effects and Defects in Solids **130**, 3 (1994).

[113] Yaolin Chen, Hongxia Liu, Cong Yan, and Hao Wei, Micromachines **14**, 2126 (2023).

[114] Synopsys Inc., Sentaurus Process: TCAD Process Simulation, 2025, accessed: 2025-08-29.

[115] Silvaco Inc., Victory Process 3D TCAD Simulator, 2025, accessed: 2025-08-29.

[116] Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Crystal-TRIM, 2025, accessed: 2025-08-29.

[117] Hosei University, scat-GUI: Ion Beam Analysis Simulation Software, 2025, accessed: 2025-08-29.

[118] M. S. Janson, SIIMPL: Simulation of Ion Implantation in Materials, 2025, accessed: 2025-08-29.

[119] MulSKIPS Collaboration, MulSKIPS: Multi Structures Kinetic Process Simulator, 2025, accessed: 2025-08-29.

[120] Margareta K Linnarsson, L Vines, and Anders Hallén, Journal of Applied Physics **130**, (2021).

[121] Johanna Müting, Viktor Bobal, Marc Georg Willinger, Ali Baghi Zadeh, Steffen Reidt, Lasse Vines, and Ulrike Grossner, IEEE Transactions on Electron Devices **67**, 4360 (2020).

[122] Gaetano Calogero, Ioannis Deretzis, Giuseppe Fisicaro, Damiano Ricciarelli, Rosario Gaetano Viglione, and Antonino La Magna, Advanced Quantum Technologies e2500160 (2025).

[123] MK Linnarsson, Anders Hallén, and Lasse Vines, Semiconductor Science and Technology **34**, 115006 (2019).

[124] Tomoaki Nishimura and Tetsu Kachi, Applied Physics Express **14**, 116502 (2021).

[125] S Agarwal, Y Lin, C Li, RE Stoller, and SJ Zinkle, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms **503**, 11 (2021).

[126] Jintong Wu, Zongwei Xu, Lei Liu, Alexander Hartmaier, Mathias Rommel, Kai Nordlund, Tao Wang, Rebecca Janisch, and Junlei Zhao, Journal of Materials Chemistry C **9**, 2258 (2021).

[127] Ka Yu Fung, Yan Ru Lin, Pei Jun Yu, Ji Jung Kai, and Alice Hu, Journal of Nuclear Materials **508**, 292 (2018).

[128] G Bonny, L Buongiorno, A Bakaev, and N Castin, Scientific Reports **10**, 10483 (2020).

[129] Marie Backman, Marcel Toulemonde, Olli H Pakarinen, Niklas Juslin, Flyura Djurabekova, Kai Nordlund, Aurelien Debelle, and William J Weber, Computational materials science **67**, 261 (2013).

[130] Kai Nordlund, Steven J Zinkle, Andrea E Sand, Fredric Granberg, Robert S Averback, Roger E Stoller, Tomoaki Suzudo, Lorenzo Malerba, Florian Banhart, William J Weber, et al., Journal of Nuclear Materials **512**, 450 (2018).

[131] Shangquan Zhao, Guang Ran, Fei Gao, and Sihan Ma, Available at SSRN 3770989 (2021).

[132] JM Lento, Leena Torpo, Torsten EM Staab, and RM Nieminen, Journal of Physics: Condensed Matter **16**, 1053 (2004).

[133] Ming-Jie Zheng, Narasimhan Swaminathan, Dane Morgan, and Izabela Szlufarska, Physical Review B—Condensed Matter and Materials Physics **88**, 054105 (2013).

[134] Louis A Girifalco, *Statistical mechanics of solids* (OUP USA, Oxford New York, 2003), Vol. 58.

[135] Chris H Rycroft, Gary S Grest, James W Landry, and Martin Z Bazant, Physical review E **74**, 021306 (2006).

[136] Chris Rycroft, chaos **chaos 19**, 041111 (2009).

[137] Charles Truong, Laurent Oudre, and Nicolas Vayatis, Signal Processing **167**, 107299 (2020).

[138] Steve Plimpton, Journal of computational physics **117**, 1 (1995).

[139] K Kamitani, M Grimsditch, JC Nipko, C-K Loong, M Okada, and I Kimura, Journal of applied physics **82**, 3152 (1997).

Appendice A

Parameterization for Force fields

Tabella A.1: Comparison of Tersoff potential parameters for C and Si as given in LAMMPS and in Tersoff's formulations.

Parameter	LAMMPS(C)	LAMMPS (Si)	Tersoff[73] (CCC)	Tersoff[73] (SiSiSi)
m	0	0	3	3
γ	–	–	1	1
λ_3	0	0	0	0
c	3.8049×10^4	1.0039×10^5	38049	100390
d	4.3841	1.6217×10^1	4.3484	16.217
$\cos\theta$	–	–	-0.57058	-0.59825
n	7.2751×10^{-1}	7.8734×10^{-1}	0.72751	0.78734
β	1.5724×10^{-7}	1.1000×10^{-6}	1.5724×10^{-7}	1.1×10^{-6}
λ_2 (1/Å)	–	–	2.2119	1.73222
B (eV)	3.467×10^2	4.7118×10^2	346.7	471.18
R (Å)	1.8	2.7	1.95	2.85
D (Å)	–	–	0.15	0.15
λ_1 (1/Å)	3.4879	2.4799	3.4879	2.4799
A (eV)	1.3936×10^3	1.8308×10^3	1393.6	1830.8

Tabella A.2: Comparison of EDIP potential parameters for C and Si systems as implemented in LAMMPS.

Parameter	**C(LAMMPS)**	**CCC[75]**	**SiSiSi[75]**	**SiCC[75]**
A (eV)	7.982173	10.222599	5.488043	7.535967
B (Å)	1.5075463	0.959814	1.446435	1.177019
Cutoff a	3.121382	2.212263	2.941586	2.534972
Cutoff c	2.5609104	1.741598	2.540193	1.973974
α	3.1083847	1.962090	3.066580	2.507738
β	0.0070975	0.025661	0.008593	0.015347
η	0.2523244	0.275605	0.589390	0.432497
γ	1.1247945	1.084183	1.135256	1.191567
λ	1.4533108	3.633621	2.417497	3.025559
μ	0.6966326	0.594236	0.629131	0.611684
τ	–	2.827634	1.343679	2.061835
σ	0.5774108	0.536561	0.298443	0.423863
Q_0	312.1341346	289.305617	208.924548	249.115082
u_1	-0.165799	-0.165799	-0.165799	-0.165799
u_2	32.557	32.557	32.557	32.557
u_3	0.286198	0.286198	0.286198	0.286198
u_4	0.66	0.66	0.66	0.66

Appendice B

Tersoff Validity test by Comparison of results with DFT

Elastic constants comparison			
	Experimental[139]	DFT[107]	Tersoff
c11	501	534	505.78
c12	111	96	86.113
c13	52	50	44.485
c33	553	574	546.796
c44	163	171	188.94
c66	195	219	209.83

Tabella B.1: Quantitative comparison of elastic parameters calculated using the Tersoff potential, DFT, and experimental data. This table demonstrates the agreement between theoretical and experimental results.

Appendice C

Nudged Elastic Band Results

C.1 Spring Constant Optimization for NEB Calculations

The spring constant in the NEB method controls the artificial harmonic force between neighboring replicas, ensuring a smooth distribution along the reaction pathway. To assess the sensitivity of migration barrier calculations to this parameter, we performed systematic tests using both the EDIP and Tersoff potentials for the $C_{Si}V_C$ defect in 3C-SiC. In each case, 20 intermediate images were used, and spring constants ranging from 1 eV/Å^2 to 7 eV/Å^2 were applied. Figure C.1 presents the energy profiles for $V_{Si} \rightarrow V_C C_{Si}$ migration in cubic SiC, obtained from NEB calculations using the EDIP potential for different spring constants (1, 3, 5, 7, and 9 eV/Å^2). The initial state (V_{Si}) and final state ($V_C C_{Si}$) are labeled along the reaction coordinate. While very low spring constants (e.g., 1 eV/Å^2) produce a smoother path but under sample the barrier region, excessively high values (e.g., 9 eV/Å^2) can result in excessive image crowding and nonphysical oscillations in the energy profile due to numerical stiffness. Based on this analysis, intermediate spring constants (3–5 eV/Å^2) provide a good compromise between stability and resolution of the energy barrier. For the EDIP potential, the energy profiles.

In the case of the Tersoff potential, the spring constant variation (Figure C.2) results in negligible differences in the computed barrier height.

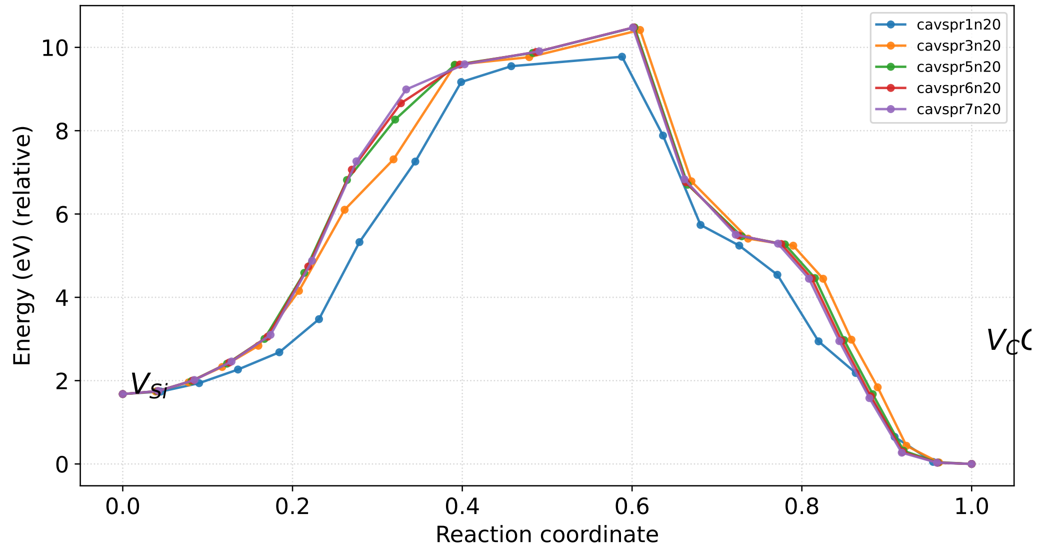


Figura C.1: Effect of varying spring constant on the NEB energy profile for the $C_{Si}V_C$ defect in 3C-SiC using the EDIP potential. Spring constants from 1 eV/Å^2 to 7 eV/Å^2 were tested with 20 intermediate images. The migration barrier shows negligible dependence on the spring constant.

However, at very low spring constants ($\leq$ 1 eV/Å^2), a slight unevenness in image distribution can be observed, potentially leading to less accurate interpolation of the transition state geometry. Increasing the spring constant to around 3 eV/Å^2 or higher produces consistently smooth energy profiles without over constraining the path, confirming this as a robust choice for subsequent NEB calculations.

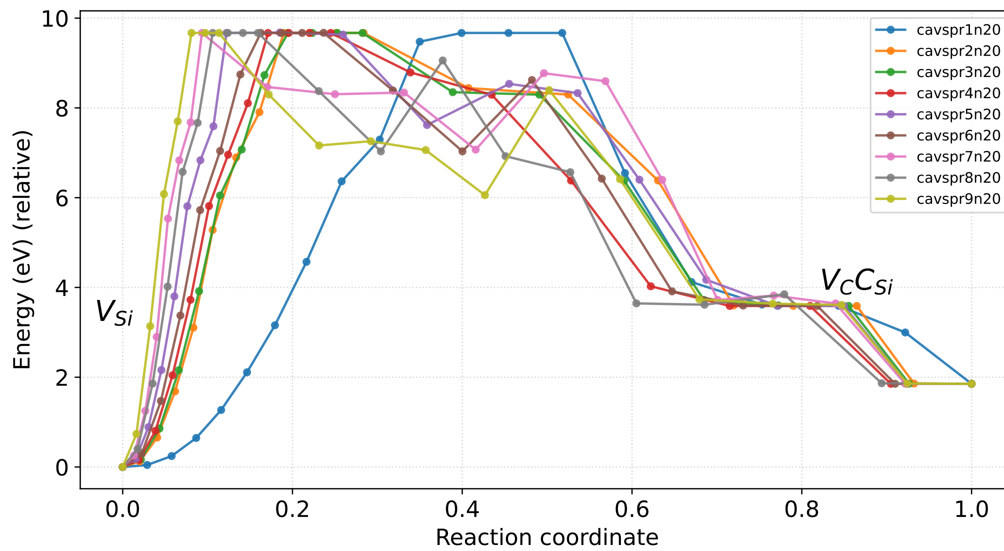
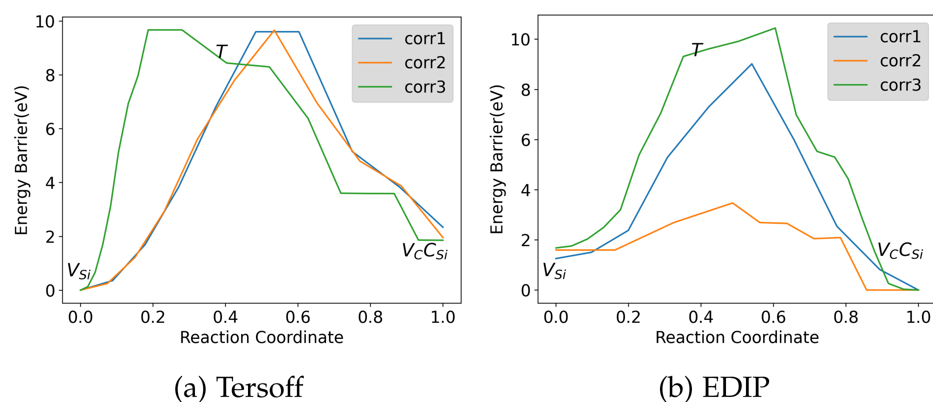


Figura C.2: Effect of varying spring constant on the NEB energy profile for the $C_{\mathrm{Si}}V_{\mathrm{C}}$ defect in 3C-SiC using the Tersoff potential. While the barrier height remains stable across the tested range, a spring constant of around 3 eV/Å^2 yields the smoothest and most evenly distributed images along the path.

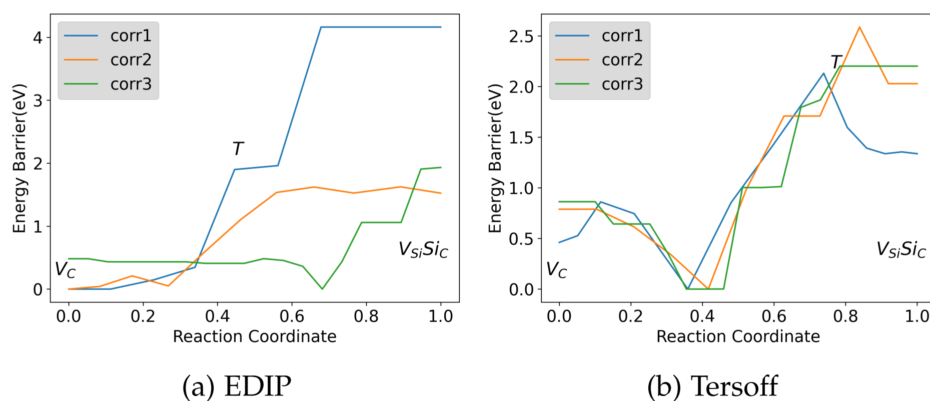
C.2 Carbon Antisite Vacancy



(a) Tersoff (b) EDIP

Figura C.3: NEB energy profile for the $C_{Si}V_c$ transformation, contrasting the results computed with the Tersoff (left) and EDIP (right) potentials. The corrections labeled as *corr1*, *corr2*, and *corr3* pertain to the methodology adjustments, including the use of forcefield optimized initial and final states, accounting for the displacement of neighboring atoms during replica generation, and the adoption of an optimized spring constant.

C.3 Silicon Antisite Vacancy



(a) EDIP (b) Tersoff

Figura C.4: NEB energy landscape for the $Si_C V_{Si}$ conversion, comparing results obtained using the Tersoff (left) and EDIP (right) potentials. The terms *corr1*, *corr2*, and *corr3* refer to the corrections applied to the methodology, as discussed in the text.

Appendice D

Gaussian Integral Results

For completeness, we summarize the standard Gaussian integrals used throughout the analysis:

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}, \tag{D.1}$$

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2}\,dx = \frac{\sqrt{\pi}}{2\,a^{3/2}}, \tag{D.2}$$

$$\int_{-\infty}^{\infty} e^{-a(x+b)^2}\,dx = \sqrt{\frac{\pi}{a}}. \tag{D.3}$$

Appendice E

StepShiftDetect : Quantifying Jumps in Diffusive Motion via Change Point Analysis

```
import pandas as pd
import numpy as np
import ruptures as rpt
import os
import time

# Function to get user input for parameters
def get_user_input():
    Temp = input("Enter the temperature (e.g., 2100): ")
    tme = input("Enter the time (e.g., 100ns): ")
    cut = int(input("Enter the cut value (e.g., 10): "))
    chunk_size = int(input("Enter the chunk size (e.g., 500): "))
    model = input("Enter the model (e.g., rbf): ")
    cs = input("Enter the cs value (e.g., cs5h): ")
    return Temp, tme, cut, chunk_size, model, cs

# Get user input for parameters
Temp, tme, cut, chunk_size, model, cs = get_user_input()

# Start time
```

```
21 s_time = time.time()
22
23 # Load your center of mass data
24 try:
25     data = pd.read_csv(f'cent_mass_{Temp}_c.dat.{tme}', skiprows=30, skip
26     print(data)
27 except FileNotFoundError:
28     print(f"File␣'cent_mass_{Temp}_c.dat.{tme}'␣not␣found.␣Please␣check␣t
29     exit()
30
31 # Perform change point detection using Pelt algorithm
32 all_results = np.empty((0, 3))  # Create an empty array to append the cal
33 com = data[['x', 'y', 'z']].to_numpy()
34
35 # Specify the file name
36 file_path = f'result.rpt.{Temp}.{tme}.chunks.cut{cut}.{cs}'
37 print("File␣Path:", file_path)
38
39 # Check if the file exists before removing it
40 if os.path.exists(file_path):
41     os.remove(file_path)
42     print(f"Old␣file␣'{file_path}'␣removed.")
43
44 for start in range(0, len(com), chunk_size):
45     end = min(start + chunk_size, len(com))
46     chunk_data = com[start:end]  # Extracting the current chunk of data
47     algo = rpt.Pelt(model=model).fit(chunk_data)  # Fit the model on the
48     result = algo.predict(pen=cut)
49     if len(chunk_data) in result:  # To miss the start and end point nois
50         result[-1] = np.nan
51     else:
52         continue
53     result = [value for value in result if not np.isnan(value)]
54     # Apply chunk_index * chunk_size to each element in result
55     chunk_index = start // chunk_size
56     result = [value + chunk_index * chunk_size for value in result]
57     df = pd.DataFrame(result)
58     try:
```

```
59          df.to_csv(file_path, mode='a', header=False, index=True)
   # Write to the file
60      except Exception as e:
61          print("Error occurred while writing to file:", e)
62
63  # End time
64  e_time = time.time()
65  t_time = (e_time - s_time) / 60
66  print(f'Total computation time is {t_time} minutes')
```

Appendice F

Sampling Error Analysis of MSD Derived Diffusivity

F.1 Introduction

In this appendix, we illustrate how estimates of the diffusion coefficient D, calculated from the mean square displacement (MSD), fluctuate as the mean time step dt_mean (or, equivalently, the temperature T) is varied. While the ensemble averaged diffusivity $\langle D \rangle$ converges reliably to the theoretical value in the limit of many trajectories, individual simulation runs can exhibit substantial deviations. This variability reflects both the intrinsic statistical noise inherent to finite sampling and the stochastic character of diffusion processes at the atomic scale.

Notably, even when all simulation parameters are held constant, estimates of D obtained from single molecular dynamics (MD) trajectories often differ appreciably from the ensemble mean. Such fluctuations become more pronounced at shorter simulation times or when fewer jump events are sampled. This analysis underscores the importance of averaging over a sufficiently large number of independent realizations to achieve accurate and reproducible estimates of diffusivity.

The results presented here help to clarify the origin of the discrepancies observed between diffusivity values derived from the frequency rule method and those obtained using the Einstein relation. By systematically

quantifying the degree of variation across multiple runs, we provide an independent framework for interpreting these differences and for assessing the uncertainty associated with diffusion estimates in computational studies.

F.1.1 Simulation Protocol

- **Model system:** 3C-SiC, second-nearest-neighbor random walk.
- **Algorithm:** At each step, choose one of the 12 second-nearest sites uniformly at random, integrate positions and compute MSD.
- **Parameters:**
  - dt_mean $\in$ $\{6.77, 2.78, 54.7, 50.5, 22.0, 19.4\}$ ns. dt_mean is mean waiting time, estimated from real time MD simulation and it relates the respective temperature value.
  - Number of trajectories per dt_mean: $traj = 10000$
  - Number of steps per trajectory: 2000
- **Data collection:** For each trajectory i, estimated

$$D_i = \frac{1}{2d} \frac{d}{dt} \langle r^2(t) \rangle$$

  using $r^2(t)$. and fixed value of D$-$frequency

F.1.2 Statistical average

For each dt_mean, we collect the ensemble $\{D_i\}_{i=1}^{M}$ and compute:

$$\langle D \rangle = \frac{1}{M} \sum_{i=1}^{M} D_i, \quad \sigma_D = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M} (D_i - \langle D \rangle)^2}.$$

F.1.3 Results and Figures

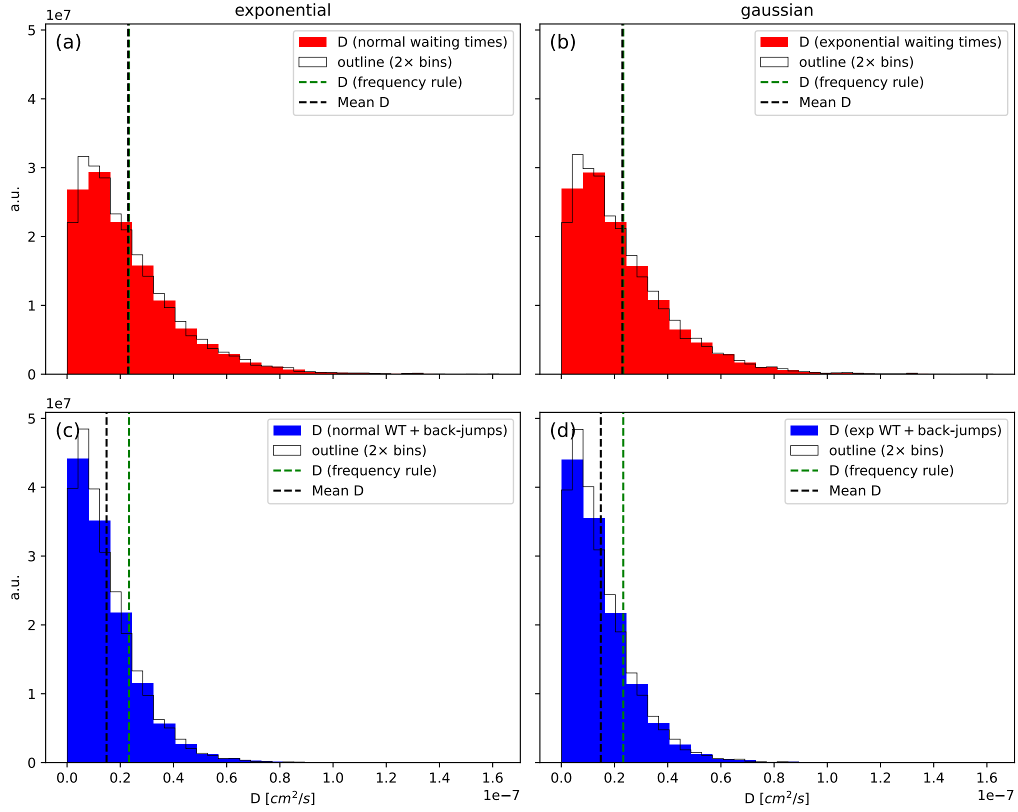


Figura F.1: Distribution of diffusivity values (D) obtained from ensemble simulations of second-neighbor random walks on a 3C-SiC lattice for a mean jump interval $\Delta t_{\text{mean}} = 6.77$ ns at temperature $T = 1600$ K . Each subplot shows histograms of D calculated from 10,000 independent trajectories of 2,000 jumps, using different stochastic time models: (a) exponential waiting times without directional bias, (b) Gaussian-distributed waiting times without directional bias, (c) exponential waiting times with directional memory (biased inverse jump), and (d) Gaussian waiting times with directional memory. Vertical dashed green lines indicate the theoretical diffusivity $D_{\text{freq}} = \lambda^2 / \left(6\langle \Delta t \rangle\right)$, while black dashed lines show the ensemble-averaged diffusivity $\langle D \rangle$ for each case. The directional memory bias introduces asymmetry in the random walk, altering the effective diffusivity. .

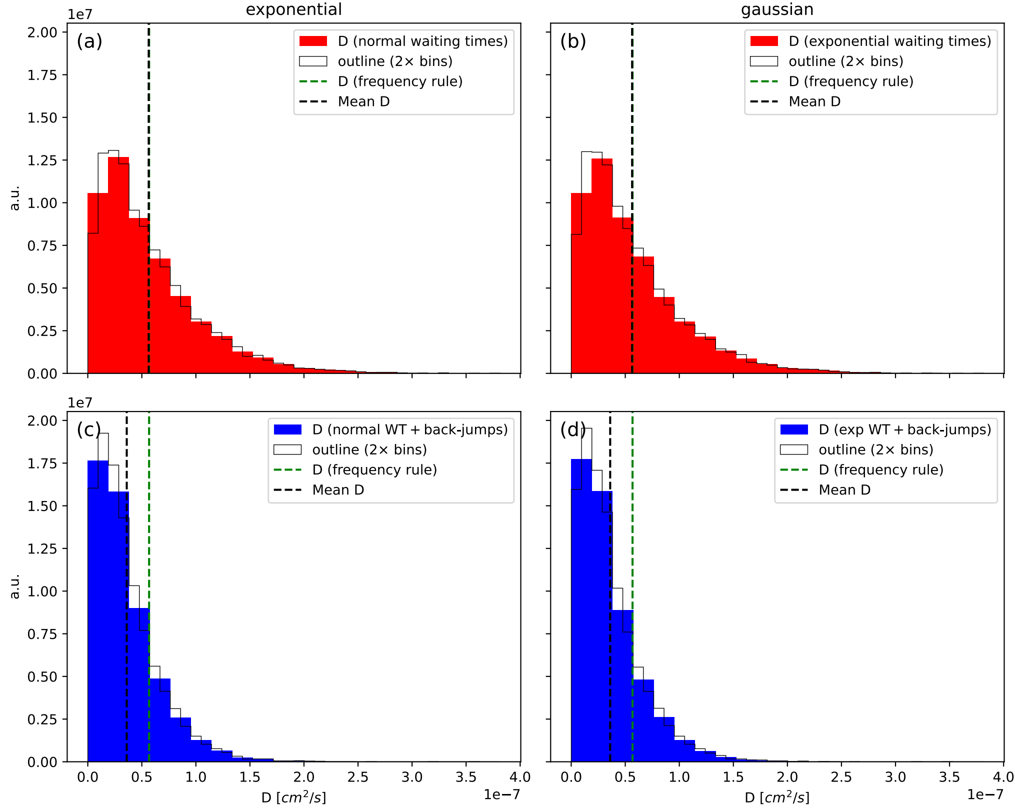


Figura F.2: Distribution of diffusivity values (D) obtained from ensemble simulations of second-neighbor random walks on a 3C-SiC lattice with a mean jump interval $\Delta t_{\text{mean}} = 2.78$ ns at temperature $T = 1700$ K. Each subplot presents histograms of D computed from 10,000 independent trajectories of 2,000 jumps, employing different stochastic time models: (a) exponential waiting times without directional bias, (b) Gaussian-distributed waiting times without directional bias, (c) exponential waiting times with directional memory (biased inverse jump), and (d) Gaussian waiting times with directional memory. Vertical dashed green lines denote the theoretical diffusivity $D_{\text{freq}} = \lambda^2 / \left(6\langle\Delta t\rangle\right)$, while black dashed lines indicate the ensemble-averaged diffusivity $\langle D \rangle$ for each case. The directional memory bias introduces asymmetry into the random walk, thereby altering the effective diffusivity.

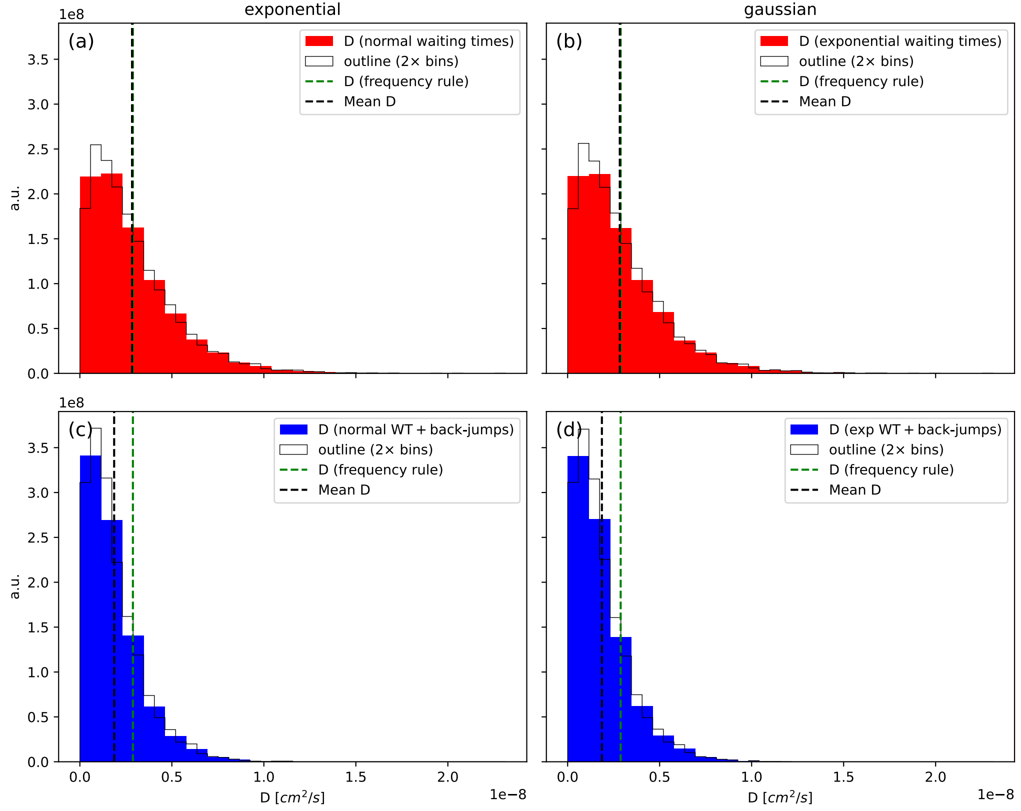


Figura F.3: Distribution of diffusivity values (D) obtained from ensemble simulations of second-neighbor random walks on a 3C-SiC lattice, with a mean jump interval $\Delta t_{\mathrm{mean}} = 54.7$ ns at temperature $T = 1800$ K. Each subplot displays histograms of D computed from 10,000 independent trajectories comprising 2,000 jumps, using four different stochastic time models: (a) exponential waiting times without directional bias, (b) Gaussian-distributed waiting times without directional bias, (c) exponential waiting times with directional memory (biased inverse jumps), and (d) Gaussian waiting times with directional memory. Vertical dashed green lines mark the theoretical diffusivity $D_{\mathrm{freq}} = \lambda^2/(6\langle\Delta t\rangle)$, while black dashed lines indicate the ensemble-averaged diffusivity $\langle D\rangle$. Incorporating directional memory introduces asymmetry in the random walk and modifies the effective diffusivity.

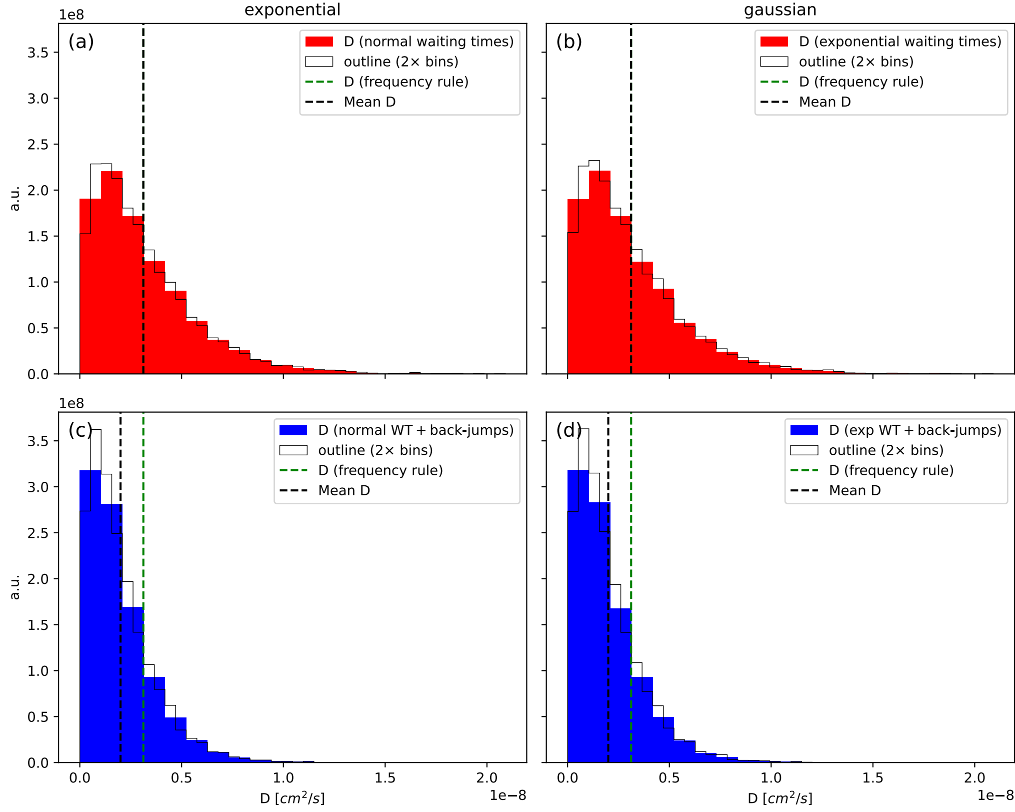


Figura F.4: Histogram distributions of diffusivity values (D) obtained from ensemble simulations of second-neighbor random walks on a 3C-SiC lattice. Simulations were conducted using a mean jump interval of $\Delta t_{\text{mean}} = 50.5$ ns at a temperature of $T = 1850$ K. Each subplot presents results from 10,000 independent trajectories, each comprising 2,000 jumps, evaluated under four distinct stochastic timing models: (a) exponential waiting times without directional bias, (b) Gaussian waiting times without directional bias, (c) exponential waiting times incorporating directional memory (biased inverse jumps), and (d) Gaussian waiting times with directional memory. Vertical dashed green lines denote the theoretical diffusivity $D_{\text{freq}} = \lambda^2/\left(6\langle\Delta t\rangle\right)$, whereas black dashed lines indicate the ensemble-averaged diffusivity $\langle D\rangle$ for each case. Directional memory introduces asymmetry in the trajectories, thereby modifying the effective diffusivity.

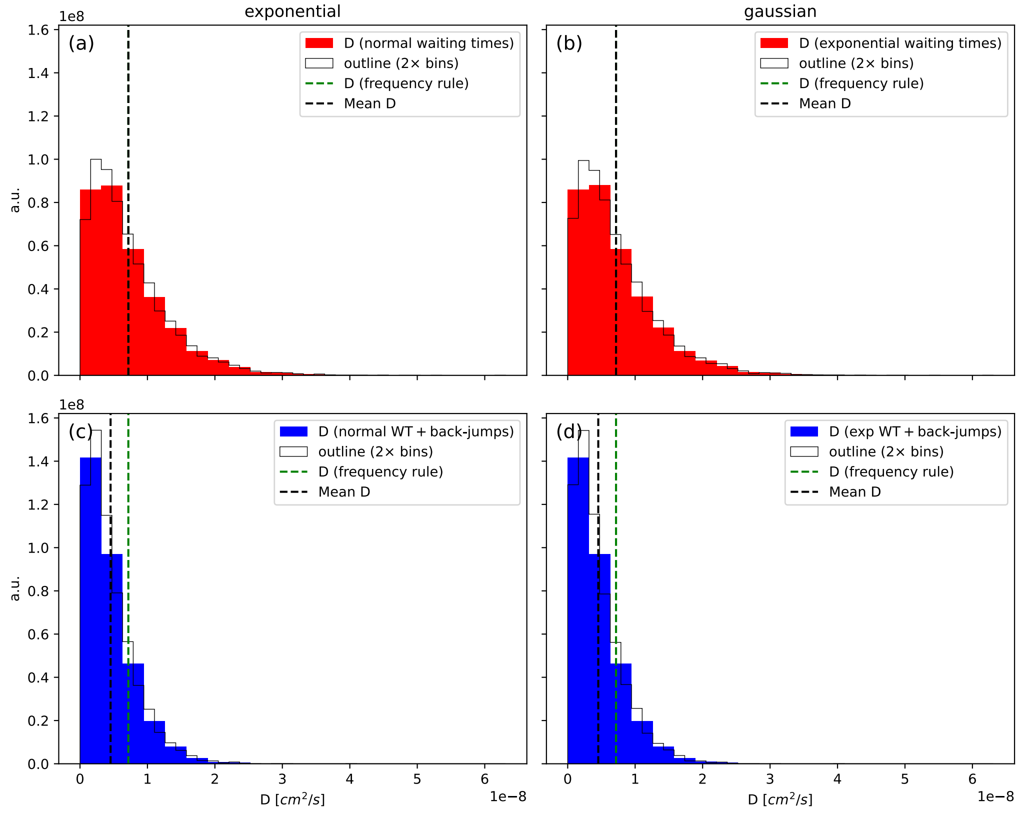


Figura F.5: Histogram distributions of diffusivity values (D) derived from ensemble simulations of second-neighbor random walks on a 3C-SiC lattice. The simulations employed a mean jump interval of $\Delta t_{\text{mean}} = 50.5$ ns at a temperature of $T = 1850$ K. Each panel illustrates results from 10,000 independent trajectories, each consisting of 2,000 jumps, analyzed under four different stochastic time models: (a) exponential waiting times without directional bias, (b) Gaussian waiting times without directional bias, (c) exponential waiting times incorporating directional memory (biased toward inverse jumps), and (d) Gaussian waiting times with directional memory. The vertical dashed green lines mark the theoretical diffusivity $D_{\text{freq}} = \lambda^2 / \left(6\langle\Delta t\rangle\right)$, while the black dashed lines represent the mean diffusivity $\langle D \rangle$ computed across the ensemble. The presence of directional memory alters trajectory symmetry, thereby influencing the observed effective diffusivity.

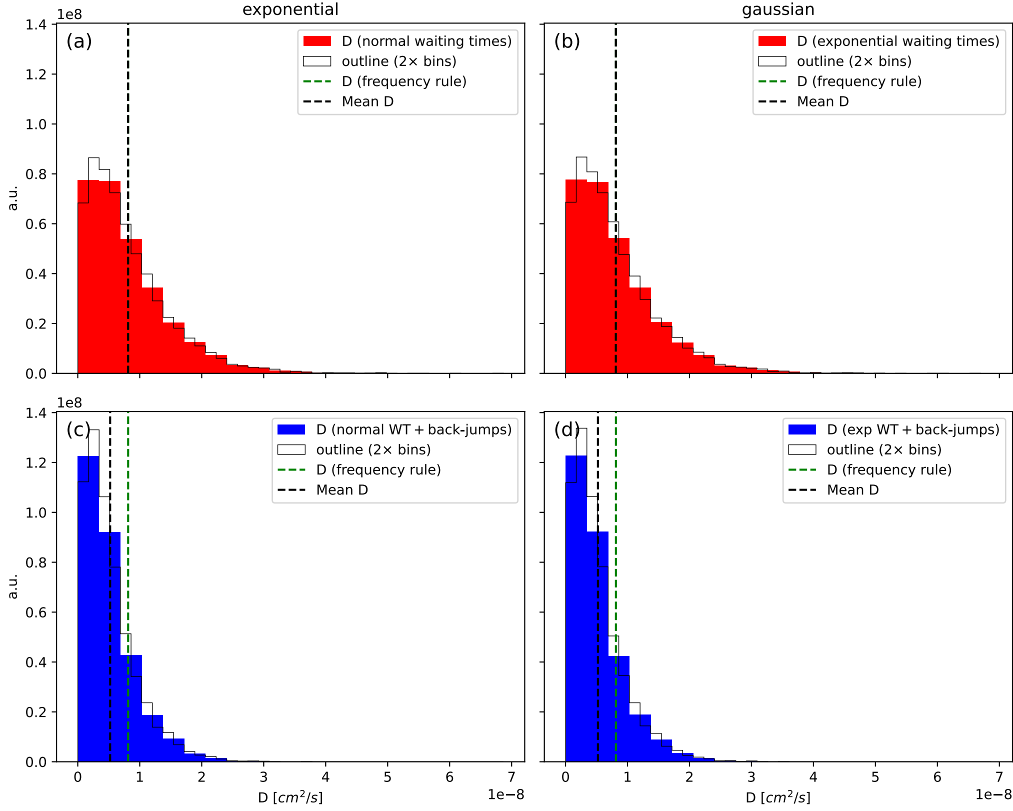


Figura F.6: Histogram distributions of diffusivity values (D) obtained from ensemble simulations of second-neighbor random walks on a 3C-SiC lattice, performed with a mean jump interval of $\Delta t_{\text{mean}} = 19.4$ ns at a temperature of $T = 1950$ K. Each panel presents results from 10,000 trajectories comprising 2,000 jumps, evaluated across four stochastic time models: (a) exponential waiting times without directional bias, (b) Gaussian waiting times without bias, (c) exponential waiting times incorporating directional memory (favoring inverse jumps), and (d) Gaussian waiting times with directional memory. Vertical dashed green lines represent the theoretical diffusivity $D_{\text{freq}} = \lambda^2/(6\langle\Delta t\rangle)$, while black dashed lines indicate the ensemble-averaged diffusivity $\langle D\rangle$. The introduction of directional memory induces asymmetry in the walks, thereby influencing the effective diffusivity.

F.1.4 Summary Table

Tabella F.1: Computed diffusivity values corresponding to each average jump interval and temperature. The columns, in order from left to right, represent: (1) the mean waiting time obtained from molecular dynamics simulations, $\mathrm{dt_mean}$ (in nanoseconds), (2) the temperature T (in Kelvin), (3) the analytical diffusivity estimated using the frequency-based formula, $D_{\mathrm{freq}} = 10^{-5}, \lambda^2 / \left(6, \langle \Delta t \rangle\right)$ (in cm^2/s), (4) the ensemble-averaged diffusivity derived from mean squared displacement (MSD) using exponential sampling, $\langle D_{\mathrm{MSD}}^{\mathrm{expon}} \rangle$ (in cm^2/s), and (5) the ensemble-averaged MSD diffusivity using normal sampling, $\langle D_{\mathrm{MSD}}^{\mathrm{Norm}} \rangle$ (in cm^2/s).

dt_mean (ns)	$Temp$ (K)	$\langle D_{freq} \rangle$ (cm^2/s)	$\langle D_{MSD}^{expon} \rangle$ (cm^2/s)	$\langle D_{MSD}^{Norm} \rangle$ (cm^2/s)
6.77	1600	$2.329 \times e^{-08}$	$2.329 \times e^{-08}$	$2.329 \times e^{-08}$
2.78	1700	$5.672 \times e^{-08}$	$5.739 \times e^{-08}$	$5.739 \times e^{-08}$
54.7	1800	$2.888 \times e^{-09}$	$2.862 \times e^{-09}$	$2.863 \times e^{-09}$
50.5	1850	$3.122 \times e^{-09}$	$3.240 \times e^{-09}$	$3.240 \times e^{-09}$
22.0	1900	$7.167 \times e^{-09}$	$7.171 \times e^{-09}$	$7.168 \times e^{-09}$
19.4	1950	$8.128 \times e^{-09}$	$8.130 \times e^{-09}$	$8.127 \times e^{-09}$

Tabella F.2: Calculated diffusivities for each mean jump interval and temperature when one-step directional memory is included. Columns, from left to right, are the MD-derived mean waiting time dt_mean (ns), temperature T (K), analytical frequency-based diffusivity $D_{\text{freq}} = 10^{-5}\,\lambda^2/(6\langle\Delta t\rangle)$ (cm^2/s), ensemble-averaged MSD diffusivity with exponential sampling $\langle D_{\text{MSD}}^{\text{expon}}\rangle$ (cm^2/s), and ensemble-averaged MSD diffusivity with normal sampling $\langle D_{\text{MSD}}^{\text{Norm}}\rangle$ (cm^2/s) under a back-jump probability of $p_{\text{back}} = 1/12 + d_{\text{prp}}$.

dt_mean (ns)	$Temp$ (K)	$\langle D_{freq}\rangle$ (cm^2/s)	$\langle D_{MSD}^{expon}\rangle$ (cm^2/s)	$\langle D_{MSD}^{Norm}\rangle$ (cm^2/s)
6.77	1600	$2.329 \times e^{-08}$	$1.499 \times e^{-08}$	$1.499 \times e^{-08}$
2.78	1700	$5.672 \times e^{-08}$	$3.662 \times e^{-08}$	$3.660 \times e^{-08}$
54.7	1800	$2.888 \times e^{-09}$	$1.871 \times e^{-09}$	$1.870 \times e^{-09}$
50.5	1850	$3.122 \times e^{-09}$	$1.990 \times e^{-09}$	$1.989 \times e^{-09}$
22.0	1900	$7.167 \times e^{-09}$	$4.505 \times e^{-09}$	$4.503 \times e^{-09}$
19.4	1950	$8.128 \times e^{-09}$	$5.279 \times e^{-09}$	$5.278 \times e^{-09}$

F.2 Generating 2nn random walk in 3C-SiC and measuring `D_i` over multiple `dt_mean`

```
START TIMER t0

// 1. SET UP PARAMETERS
set lattice_constant a <- 0.435  // nm
compute jump_length lambda <- a / sqrt(2)
set attempt_frequency Gamma <- 1e12   // unused here
set n_traj <- 10000               // number of independent walks
set n_jumps <- 2000               // max jumps per walk
set mean_wait dt_mean <- 19.4     // ns

// Precompute jump vectors for 12 second-neighbor directions
build array jump_vectors[12] from lattice geometry
build inverse_index array inv_jump[12]

// Compute maximum simulation time (unused later)
max_time <- n_jumps * dt_mean * 1e-9   // s

```

```
18 // Frequency-rule diffusivity in cm^2/s
19 D_freq <- 1e-5 * lambda^2 / (6 * dt_mean)
20
21 INITIALIZE arrays D_msd_arr, D_msd_arr1 of length n_traj
22
23 // 2. ENSEMBLE OF TRAJECTORIES, NO MEMORY EFFECT
24 FOR each trajectory i = 1 to n_traj:
25     t_acc_exp <- 0
26     t_acc_gauss <- 0
27     pos_vector <- zero vector
28
29     REPEAT n_jumps times:
30         // draw waiting times
31         dt_exp <- random exponential with mean dt_mean
32         dt_gauss <- random normal with mean dt_mean
33
34         // accumulate time
35         t_acc_exp <- t_acc_exp + dt_exp
36         t_acc_gauss <- t_acc_gauss + dt_gauss
37
38         // perform a random jump
39         choose random index k in [0..11]
40         pos_vector <- pos_vector + jump_vectors[k]
41
42     // compute squared displacement
43     msd <- dot(pos_vector, pos_vector)
44
45     // compute MSD-based D in cm^2/s
46     D_msd_arr[i]  <- 1e-5 * msd / (6 * t_acc_exp)
47     D_msd_arr1[i] <- 1e-5 * msd / (6 * t_acc_gauss)
48 END FOR
49
50 // Compute summary statistics of these distributions
51 compute D_msd_av  <- mean(D_msd_arr)
52 compute D_msd_av1 <- mean(D_msd_arr1)
53 compute Dmin, Dmax from both arrays
54
55 // 3. ADD MEMORY ("back-jump") EFFECT
56 set back_jump_probability increment dprp <- 0.2
57 compute phigh <- 1/12 + dprp
58 compute plow  <- (1 - phigh) / 11
59
60 INITIALIZE arrays D_msd_arr2, D_msd_arr3 of length n_traj
```

```
61  pick initial direction ndir <- random 0..11
62  set inv_dir <- inv_jump[ndir]
63  initialize random number generator rng
64
65  FOR each trajectory i = 1 to n_traj:
66      t_acc_exp <- 0
67      t_acc_gauss <- 0
68      pos_vector <- zero vector
69
70      REPEAT n_jumps times:
71          // draw waiting times
72          dt_exp  <- random exponential with mean dt_mean
73          dt_gauss<- random normal with mean dt_mean
74
75          // accumulate time
76          t_acc_exp   <- t_acc_exp + dt_exp
77          t_acc_gauss <- t_acc_gauss + dt_gauss
78
79          // build jump-probability array
80          set all 12 entries of pjump to plow
81          set pjump[inv_dir] to phigh
82
83          // select new direction according to pjump
84          generate uniform prand in [0,1)
85          ndir <- smallest index where cumulative sum(pjump) >=
                prand
86          inv_dir <- inv_jump[ndir]
87
88          // perform jump
89          pos_vector <- pos_vector + jump_vectors[ndir]
90
91      // compute squared displacement
92      msd <- dot(pos_vector, pos_vector)
93
94      // compute D with back-jumps
95      D_msd_arr2[i] <- 1e-5 * msd / (6 * t_acc_exp)
96      D_msd_arr3[i] <- 1e-5 * msd / (6 * t_acc_gauss)
97  END FOR
98
99  // Update global min/max if needed
100 update Dmin, Dmax from D_msd_arr2 and D_msd_arr3
101 compute D_msd_av2 <- mean(D_msd_arr2)
102 compute D_msd_av3 <- mean(D_msd_arr3)
```

```
103
104 // 4. PLOT HISTOGRAMS
105 create 2x2 subplot grid
106 for each of the four D arrays (no-memory vs memory x exp vs
        gauss):
107      plot histogram over [Dmin, Dmax]
108      overlay outline histogram (double bins)
109      draw vertical line at D_freq (frequency rule)
110      draw vertical line at corresponding mean D
111
112 label subplots, axes, legends, annotations (a), (b), ...
113
114 adjust layout, save figures as PNG and PDF
115
116 // 5. FINISH
117 stop timer tfin <- time.time()
118 print CPU time tfin - t0
```

Listing F.1: Pseudocode for ensemble diffusion simulation

```
1  import numpy as np
2  import os
3  from tqdm import tqdm
4  import matplotlib.pyplot as plt
5  import time
6  t0 = time.time()
7  # Parameters
8  lambda_jump = a/np.sqrt(2) # Second-neighbor distance
9  Gamma = 1e12 # it has no meaning for our scope
10 n_traj = 10000 # we will simulate ensemble or many independent trajectori
11 n_jumps = 2000 # Max jumps per traj, we will use many independent walks a
12 #if we want uniform jumps we can use fix time for each jump, no need rate
13 # Build the 12 s e c o n d neighbor jump vectors thanks to np.roll functi
14 jump_vectors = np.array([np.roll([s1, s2, 0], i) for i in range(3) for s1
   in [1,-1] for s2 in [1,-1]]) * (a/2) ## raw verctors give magnitude of sq
15 inv_jump = np.array([3, 2, 1, 0, 7, 6, 5, 4, 11, 10, 9, 8], dtype=int)
16 #print(jump_vectors[inv_jump [0]])
17 #print(inv_jump)
18 dt_mean=1.# ns this should be the value estimated by MD
19 max_time = n_jumps * dt_mean *1.e-9 # we multiply total_steps and avg-tim
20 D_freq = 1.e-5 *lambda_jump 2 / (dt_mean * 6) # Diffusivity 1e-5 factor t
```

```
21 fig, axs = plt.subplots(2, 2,sharex=True, sharey=True)
22
23 D_msd_arr=np.zeros(n_traj)
24 D_msd_arr1=np.zeros(n_traj)
25 for itraj in range(n_traj) :
26     t_acc = 0.
27     t_acc1= 0.
28     pos = 0.
29     for ijump in range(n_jumps):
30         dt = np.random.exponential(dt_mean)
31         dt1 = np.random.normal(dt_mean)
32         t_acc= t_acc + dt
33         t_acc1= t_acc1 + dt1
34         pos += jump_vectors[np.random.randint(12)]
35     msd=np.dot(pos, pos)
36     D_msd=1.e-5* msd/(6.*t_acc)
37     D_msd1=1.e-5* msd/(6.*t_acc1)
38     D_msd_arr[itraj]=D_msd
39     D_msd_arr1[itraj]=D_msd1
40 Dmax=np.max(D_msd_arr)
41 Dmin=np.min(D_msd_arr)
42 Dmax_n=np.max(D_msd_arr1)
43 if(Dmax_n>Dmax):Dmax=Dmax_n #updating D_max from new calculated D
44 Dmin_n=np.min(D_msd_arr1)
45 if(Dmin_n<Dmax):Dmin=Dmin_n
46 D_msd_av=np.mean(D_msd_arr)
47 #D_msd_av=np.mean(D_msd_arr)
48 D_msd_av1=np.mean(D_msd_arr1)
49 #D_msd_av1=np.mean(D_msd_arr1)
50 print('standard','D_freq=',D_freq,'D_msd_exp=',D_msd_av,'D_msd_gau=',D_msd_av1)
51
52 dprp  = 0.2
53 if(1./12.+dprp > 1):
54     print('too high jump probability')
55     quit()
56 phigh = 1./12. + dprp
57 plow  = (1. - phigh)/11.
58 ndir=np.random.randint(12)  # initialization
```

```
59 inv_dir=inv_jump[ndir]
60 rng=np.random.default_rng()
61 
62 D_msd_arr2=np.zeros(n_traj)
63 D_msd_arr3=np.zeros(n_traj)
64 for itraj in range(n_traj) :
65     pos = 0.
66     t_acc=0.
67     t_acc1=0.
68     for ijump in range(n_jumps):
69         dt = np.random.exponential(dt_mean)
70         dt1 = np.random.normal(dt_mean)
71         t_acc= t_acc + dt
72         t_acc1= t_acc1 + dt1
73         pjump=np.full((12),plow) # uniform low probability
74         pjump[inv_dir]=phigh     # high probability in inverse direction
75         ndir=0  # select jump direction according pjump
76         prand=rng.random()
77         pcurr=pjump[0]
78         while (pcurr < prand):
79             ndir+=1
80             pcurr=pcurr+pjump[ndir]
81     #    print(ndir)  # check if works for high dprp position oscillation
82         pos += jump_vectors[ndir]
83         inv_dir=inv_jump[ndir]
84     msd=np.dot(pos, pos)
85     D_msd2=1e-5*msd/(6.*t_acc)
86     D_msd3=1e-5*msd/(6.*t_acc1)
87     D_msd_arr2[itraj]=D_msd2 #collecting many D_exponential
88     D_msd_arr3[itraj]=D_msd3 #collection many D_normal
89 #    ndir=0
90 Dmax_n=np.max(D_msd_arr2)
91 if(Dmax_n>Dmax):Dmax=Dmax_n
92 Dmin_n=np.min(D_msd_arr2)
93 if(Dmin_n<Dmax):Dmin=Dmin_n
94 Dmax_n=np.max(D_msd_arr3)
95 if(Dmax_n>Dmax):Dmax=Dmax_n
96 Dmin_n=np.min(D_msd_arr3)
```

```
97  if(Dmin_n<Dmax):Dmin=Dmin_n
98  D_msd_av2=np.mean(D_msd_arr2)
99  D_msd_av3=np.mean(D_msd_arr3)
100
101 nbin = 20
102 #Dmin=0.
103
104 axs[0, 0].hist(D_msd_arr, nbin,range=(Dmin,Dmax), density=True, facecolor='r', a
105 axs[0, 0].hist(D_msd_arr, nbin*2,range=(Dmin,Dmax), density=True, histtype='step
106 axs[0, 1].hist(D_msd_arr1, nbin,range=(Dmin,Dmax), density=True, facecolor='r',
107 axs[0, 1].hist(D_msd_arr1, nbin*2,range=(Dmin,Dmax), density=True, histtype='ste
108 axs[0, 0].axvline (x=D_freq, color='green', linestyle='--')
109 axs[0, 1].axvline (x=D_freq, color='green', linestyle='--')
110 axs[0, 0].axvline (x=D_msd_av, color='black', linestyle='--')
111 axs[0, 1].axvline (x=D_msd_av1, color='black', linestyle='--')
112 axs[1, 0].hist(D_msd_arr2, nbin,range=(Dmin,Dmax), density=True, facecolor='b',
113 axs[1, 0].hist(D_msd_arr2, nbin*2,range=(Dmin,Dmax), density=True, histtype='ste
114 axs[1, 1].hist(D_msd_arr3, nbin,range=(Dmin,Dmax), density=True, facecolor='b',
115 axs[1, 1].hist(D_msd_arr3, nbin*2,range=(Dmin,Dmax), density=True, histtype='ste
116 axs[1, 0].axvline (x=D_freq, color='green', linestyle='--')
117 axs[1, 1].axvline (x=D_freq, color='green', linestyle='--')
118 axs[1, 0].axvline (x=D_msd_av2, color='black', linestyle='--')
119 axs[1, 1].axvline (x=D_msd_av3, color='black', linestyle='--')
120 axs[0, 0].set_title('exponential')
121 axs[0, 1].set_title('gaussian')
122 print('memory','D_freq=',D_freq,'D_msd_exp=',D_msd_av2,'D_msd_gau=',D_msd_av3)
123 for ax in axs.flat:
124     ax.set(xlabel=r'D␣[$cm^2␣/␣s$]', ylabel='a.u.')
125
126 # Hide x labels and tick labels for top plots and y ticks for right plots.
127 for ax in axs.flat:
128     ax.label_outer()
129 fig.subplots_adjust(wspace=0.05)
130 fig.savefig('Dnt'+str(n_traj)+'nj'+str(n_jumps)+'.png')
131 fig.savefig('Dnt'+str(n_traj)+'nj'+str(n_jumps)+'.pdf')
132 fig.show()
133 tfin = time.time()
134 print("tCPU␣=",tfin-t0)
```

```
135 #pjmup=np.full
136
137
138 # (Save ‘results‘ to file or use it to make the above plots.)
```

Acknowledgments

It is difficult to find words that truly capture the depth of my gratitude to the *University of Catania* and especially to the *Department of Physics and Astronomy "Ettore Majorana" (DFA)* for granting me the opportunity to pursue my doctoral studies. The department has offered not just resources, but a vibrant and stimulating environment that encouraged curiosity and critical thinking at every step.

I owe my deepest thanks to my supervisor, **Prof. Elisabetta Paladino**, whose intellectual clarity and patient guidance have shaped this work in more ways than I can count. Her mentorship not only refined my scientific discipline but also helped me to truly grasp and appreciate the fundamental ideas that underpin physics itself. Many of the concepts that once felt abstract or elusive became clear under her thoughtful explanations. I have often found myself inspired by her precision, her calm persistence, and her ability to see elegance in complexity.

My sincere appreciation also goes to the *Institute for Microelectronics and Microsystems of the National Research Council (CNR-IMM), Italy*, for welcoming me into its computational framework and allowing me to conduct part of my research there. The time I spent within that environment was immensely formative both technically and personally.

I am especially grateful to **Prof. Antonino La Magna**, whose insight and constructive feedback often arrived at just the right moment. His openness to discussion, his generosity with knowledge, and his patience when I inevitably stumbled have all left a deep impression on me. I have learned not only from his scientific expertise but also from his steady and

generous way of engaging with others.

To my parents, **Farhat Jabeen** and **Ashraf Asim**, My sisters **Sobia, Sundas and Madiha** I owe everything. Their love, quiet strength, and countless sacrifices made it possible for me to follow this path. It may sound simple, but their faith in me often stronger than my own has carried me through moments of uncertainty.

To my wife, **Laiba Ejaz**, my daughter **Mirha Irslan** I extend my heartfelt gratitude. Laiba patience, warmth, and understanding have sustained me during the long hours and unpredictable days of this journey. She has been both my anchor and my encouragement, reminding me of balance and perspective when I was too deep in thought or too buried in code.

Lastly, I would like to thank all my colleagues and the staff at both DFA and CNR-IMM. Many of them may not realize how their small acts of help, conversations over coffee, or words of support made a difference but they truly did. And to my entire family, whose support has always felt quietly present even from afar, I am deeply thankful.