Atomistic Characterization of Stress-Driven Configurational Instability and its Activation Mechanisms

by

Ting Zhu

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2004

Author .....
.....
Department of Mechanical Engineering
May 7, 2004

Certified by.....

Sidney Yip
Professor of Nuclear Engineering
and Professor of Materials Science and Engineering
ervisor

Accepted by.....
.....
Ain A. Sonin
Chairman, Department Committee on Graduate Students

Atomistic Characterization of Stress-Driven Configurational Instability and its Activation Mechanisms

by

Ting Zhu

Submitted to the Department of Mechanical Engineering
on May 7, 2004, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

Cleavage decohesion and shear dislocation nucleation are two basic modes of localized deformation in crystal lattices, which normally result from instability of the atomic configuration driven by mechanical forces. The critical state of instability and its thermal activation mechanisms can be quantitatively determined by analyzing the energetics of the lattice system.

In this thesis, the unit processes of configurational instability of crystal lattices under various non-uniform structural and/or chemical environments are characterized by systematically probing the atomistic potential energy landscape of each system using the state of the art configurational space sampling schemes. The problems studied are homogeneous dislocation nucleation in a perfect crystal by nanoindentation, dislocation emission and cleavage decohesion at atomically sharp crack tips, and chemically-enhanced bond breaking in a wet silica nanorod. These processes are studied in a unified manner such that two important types of properties are determined: one is the athermal load at which the instability takes place instantaneously without the aid of thermal fluctuations, and the other is the stress-dependent activation energy used for an estimate of the kinetic rate of transition.

Along the way, important aspects concerning the atomistic characterization of configurational instability are revealed. Of particular note is extending the continuum instability criterion to detect atomic defect nucleation. We demonstrate that a local instability criterion can be applied to identify dislocation nucleation in the case of indentation, considering that the relatively small strain gradient beneath the indenter will lead to a mode of long wavelength phonon instability suitable for a study by the local continuum approach. In addition, the chemical effect on stress-driven lattice instability is revealed via the study on reactivity of a silica nanorod with water. We identify distinct competing mechanisms of hydrolysis which are rate-controlling at different load regimes. The ensuing stress-mediated switch of rate-limiting steps of hydrolysis quantitatively demonstrates the impact of finding the detailed molecular mechanisms on a realistic estimate of the activation rate when configurational instability occurs within a chemically reactive environment. Implications regarding

the analysis of chemically-assisted brittle fracture are also discussed.

Thesis Supervisor: Sidney Yip
Title: Professor of Nuclear Engineering
and Professor of Materials Science and Engineering

Acknowledgments

I would like to express my deepest gratitude to my advisor and mentor, Professor Sidney Yip. I am very grateful to him for his invaluable guidance, inspiration and encouragement. What I have learned from him are not only solutions to specific problems. His insight, attitude, way of conducting research and living will have an everlasting impact on me.

I would like to thank Professors Ali Argon, Nicolas Hadjiconstantinou and Subra Suresh for serving on my thesis committee. My sincere appreciation goes to Professor Argon, who has always been interested in listening to my latest research progress. Discussions with him have been stimulating and educative. I want to thank Professors Hadjiconstantinou and Suresh for their invaluable advice on my thesis study and providing me with various opportunities to present my work.

I owe my special thanks to Professor Ju Li, who has been a constant source of inspiration and support. His contributions to this thesis are tremendous. It has been my great pleasure and privilege to work closely with him. His insight and patience have helped strengthen this thesis work significantly.

I have benefited from collaborations with a number of great researchers whose influences permeate this thesis. I want to thank Professors Lallit Anand, Nora deLeeuw, Shigenobu Ogata, Sam Trickey and Krystyn Van Vliet for their collaborations and support. I owe special thanks to Professor Wei Yang of Tsinghua University, who guided me into the field of mechanics of materials and fostered my interests in atomistic modeling. I also gratefully acknowledge the financial support by NSF-KDI, NSF-ITR(M) and AFOSR.

It has been a great experience to be a part of the MMM group with people who have been good friends and sources of discussion and technical support. I am grateful to Xi Lin, Dongyi Liao, Jinpeng Chang, Xiaofeng Qian, Robert Boyer, Jacob Eapen.

My heartfelt appreciation goes to my parents. Without their consistent emotional and intellectual support, I simply could not have come this far. Finally, I want to thank my wife, Yang, for her love, encouragement and personal sacrifice.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

One of the central properties of a crystalline solid is its strength: the maximum load a crystal can sustain without failing. Failure can occur either by brittle fracture or plastic yielding. Traditionally, people study the onset of material failure as an instability phenomenon within the framework of continuum mechanics (e.g., Rice [75]; Lawn [55]). Though quite successful, this approach has left a number of questions unanswered. For example, a typical problem out of the scope of the continuum analysis is to determine the critical condition governing the transition from brittle to ductile responses of crystalline solids. While this problem is in all aspects a macroscopic phenomenon, the key element that governs it involves the atomic-level processes of dislocation nucleation and migration at crack tips (e.g., Argon [7]). Another example showing limitations of the continuum analysis is the problem of the environmental effect on material strength. It is well known that the strength of silica glass decreases with time when subjected to a static load in an aqueous environment (e.g., Wiederhorn, [112]). This phenomenon, commonly referred to as delayed fracture or static fatigue, results from the environment-assisted quasi-static growth of pre-existing small cracks at the surface of silica glass. Ample experimental evidence proves that the underlying mechanism governing this quasi-static crack growth is the water-enhanced bond breaking at the crack tip. Clearly, in this case, the atomic-level

chemical interactions between silica and water molecules influence the macroscopic strength of silica glass. From these two examples, it is evident that considerations of the atomic-level structural details are not only important but also essential for a fundamental understanding of the strength of solids.

The past few decades have seen a dramatic increase in the use of atomistic simulations in combination with experiments in addressing problems in material science. The problem of material strength is no exception. Among various schemes of atomistic simulations that can be applied to study material strength, the most direct one is the molecular dynamics (MD) method (e.g., Allen and Tildesley [4]; Voter *et al.* [107]), in which an appropriate interatomic potential is chosen to describe the forces between atoms and then the classical equation of motion is integrated with suitable boundary conditions. An appealing feature of MD is that it follows the actual dynamic evolution of the system. However, a significant limitation of MD studies in the accessible simulation time represents a substantial obstacle to making useful predictions with MD. Specifically, resolving individual atomic vibration requires a time step of approximately femtoseconds in the integration of the equation of motion, so that on today's fastest processors, reaching even one microsecond is very difficult for a sufficiently large system.

In the past few years, some new methods, including the quasicontinuum method (Tadmor *et al.* [98]; Shenoy *et al.* [85]); the nudged elastic band method (Jonsson *et al.* [43]); the dimer method (Henkelman and Jonsson [37]), have been developed such that the configurational space of the system can be efficiently sampled, but with the price of losing information about the detailed dynamics. Nevertheless, these methods can be conveniently tailored to study a whole class of material strength problems of interest to us, i.e., the stress-driven, thermally activated, kinetic transitions leading to the localized deformation in crystal lattices such as dislocation nucleation or cleavage decohesion. As will be demonstrated in the following chapters, to quantitatively characterize these processes, the central theme is to determine the critical condition of instability, which requires an efficient sampling of the configurational space with a large number of degrees of freedom. Recent progress in developing the effective

sampling schemes, as listed above, makes possible the direct study of these problems with a reasonable computational cost.

Discussed next is one important idea guiding the present thesis work. That is, under a *low* temperature and *quasi-static* loading condition, the atomic-level study of the critical condition of instability in crystalline solids, e.g., dislocation nucleation or cleavage decohesion, can be carried out by directly probing the stress-mediated potential energy surface and finding the critical saddle point. This principle was developed within the framework of the transition state theory (e.g., Vineyard [103]; Weiner [110]). Specifically, since atoms in crystals are usually tightly packed and the typical temperature of interest is low compared to the melting temperature, the motion of atoms can be assumed to be harmonic. This approximation greatly simplifies the problem such that, instead of probing the free energy landscape, the study of system's instability is reduced to the problem of effectively sampling the configurational space and finding the saddle point on the potential energy surface. In other words, the potential energy plays a critical role in affecting the system's instability and, contributions from the configurational and vibrational entropies are not decisive.

Obviously, the potential energy surface is mediated by the applied stresses. At zero stress, the equilibrated lattice is stable, being in a state of local minimum on the potential energy surface. With an increase in stress, the elastically deformed lattice remains stable. But the energy landscape evolves in such a way that the saddle point gradually approaches the local minimum. As the load reaches the critical value of the athermal load, the saddle point collapses onto the local minimum on the energy surface. It follows that a small perturbation to the system will lead to instantaneous configurational instability. Associated with the instability are finite atomic displacements, leading to breaking of atomic bonds or nucleation of shear dislocations. Below the athermal load, the instability could nevertheless take place if thermal activation supplies enough energy to drive the system from a stable state of local minimum to an unstable state of saddle point. However, in contrast to an instantaneous instability at the athermal load, a waiting time is needed which, within the framework of the transition state theory, exponentially depends on the

energy difference between two states defined as the activation energy barrier. On the edge of the energy well enclosing the local minimum, there may exist several saddle points corresponding to different activation mechanisms. Thus it is important to find the lowest saddle point in order to get a realistic estimate of the waiting time which determines the kinetic rate of the transition. To accurately capture these fine details on the potential energy surfaces, some effective configurational space sampling schemes are desired. Of particular note is that, for the study of mechanical problems in a solid, while structural transitions such as dislocation nucleation or cleavage bond breaking may occur at some local spots, the system size still has to be sufficiently large to allow for the long range elastic relaxation. Hence, to minimize the size effect, sampling should be carried out within the configurational space with a large number of degrees of freedom. The development of effective sampling schemes meets this need.

In summary, the objective of this work is to advance the fundamental understanding of mechanical failures in crystalline solids. The onset of two basic modes of macroscopic failures, i.e., brittle fracture and plastic yielding, will be studied as an instability phenomenon at the atomic scale via systematically exploring the stress-mediated potential energy surface. Recent progress in developing the effective configurational sampling schemes makes possible an accurate probe of the energy surface for a reasonably large-sized solid system.

1.2 Problem Statement

In this thesis, several unit processes of stress-driven configurational instability of crystal lattices are studied under various structural and/or chemical environments. The specific problems studied include: (1) nanoindentation-induced homogeneous dislocation nucleation; (2) dislocation emission and cleavage decohesion at atomically sharp crack tips; (3) chemically-enhanced bond breaking in a wet silica nanorod. As the background and motivation for each problem will be discussed at the beginning of each chapter, it is of interest, here, to reveal some common features and compare

the differences of these problems.

While the problems studied differ in terms of the specific atomic-level details, the underlying physics governing each process, without exception, falls into the framework of the stress-driven configurational instability as elaborated in the last paragraph of section 1.1. It follows that these processes are studied in a unified manner such that two important types of properties are quantitatively determined: one is the athermal load at which the instability takes place instantaneously without the aid of thermal fluctuations, and the other is the stress-dependent activation energy, determined by the barrier height of the saddle point, which can be used to estimate the kinetic rate of transition.

Another common feature of these problems is that each system studied has a certain degree of structural inhomogeneity. As a result, the instability occurs within a non-uniform atomic environment. While this structural inhomogeneity does not affect the atomic-level studies of the instability, which invariably rely on the analysis of the system's energetics, it does influence the applicability of various types of continuum-level instability criterion. This effect can be clearly seen by comparing the dislocation nucleation criterion beneath the indenter with that at the crack tip. For the nanoindentation problem, dislocation nucleation results from the instability occurring within the bulk of a perfect crystal loaded by a foreign object at the surface, i.e, the nanoindenter. Since the radius of the nanoindenter used in simulations, and in real experiments as well, is on the order of 50nm, the strain gradient beneath the indenter is relatively small on the scale of lattice spacing. Thus dislocation nucleation is induced by the local instability within a nearly uniform atomic environment (though with large deformation compared to the natural stress-free state). It follows that the resulting long wavelength phonon instability can be effectively detected via a local continuum-based instability criterion.

In contrast, for dislocation nucleation at an atomically sharp crack tip, rigorously speaking, the continuum stress is not well defined near the crack tip because a highly nonuniform atomic environment exists on the scale of atomic spacing. Clearly, the local instability criterion, which relies on the continuum-based quantities such as

stress and modulus, could not be applied to this situation. However, the nucleation criterion proposed by Rice [80] bypasses the local details at the crack tip and examines the global energy balance between the elastic energy release rate calculated from the J-integral and the internal resistance given by the unstable stacking energy. This global instability criterion can effectively detect dislocation nucleation near the crack tip. Our direct atomistic calculations have confirmed the applicability of the above two continuum-based instability criteria for respective situations. In addition, we note that as the tip radius of the nanoindenter shrinks down to the order of atomic spacing, local nucleation criterion should give way to the global criterion as given in terms of the J-integral. However, in most situations, the value of the J-integral needs to be numerically calculated (e.g., Li [56]). Only for some simple cases such as the crack tip problem, can the value of J be obtained analytically, as it is well known that the stress distribution near the crack tip is characterized by the analytic solution of the K-field which makes the calculation of J straightforward.

Finally, it is of interest to compare two localized processes at crack tips studied in chapter 3 to reveal the effect of bonding characteristics on the crack tip geometry as well as its influence on the fine structures of atomistic energy landscape. The two processes studied are shear dislocation nucleation in an FCC single crystal of Cu and the covalent bond breaking in a diamond-cubic single crystal of Si, respectively. As discussed above, these two crack-tip processes are highly localized compared with that beneath the indenter, which makes the local instability criterion inapplicable at crack tips. However, another level of localized deformation can be clearly seen when comparing atomic displacements between these two processes. Specifically, for dislocation loop emission in Cu, shear displacement distribution across the slip plane at the saddle-point state exhibits a significantly extended feature as shown in Fig. 3-4, while the opening distribution across the cleavage plane in Si at the intermediate state of transition shows a highly localized mode of deformation as given in Fig. 3-8. This difference in the geometry can be correlated to a difference in bonding characteristics. That is, Si atoms have much stronger localized, directional bonds, while the bonding interaction between Cu atoms is more delocalized and shows

little directionality. Besides the geometry, the energetics differs significantly between the two processes. If we consider the crack front bond breakings as the formation of a climb dislocation, the difference manifests as that in the energy barriers to dislocation nucleation. Specifically, shear dislocation in Cu only experiences a single nucleation barrier, while the result for Si shows that the secondary energy barriers exist on the energy surface. These barriers impose substantial resistances to dislocation kink migration, leading to very sluggish motion of the climb dislocation. While these differences have been revealed qualitatively before (e.g., Argon [7]), our calculations demonstrate that the quantitative characterization with full atomistic details is now feasible.

The thesis is organized as follows: In chapter 2, homogeneous nucleation of a dislocation beneath a nanoindenter in single crystal Cu is studied as a strain localization event triggered by elastic instability of the perfect crystal at finite strain. A novel interatomic potential finite element method is employed to simulate both 2D cylindrical and 3D spherical indentations. The critical conditions of dislocation nucleation predicted via the local instability criterion are quantitatively compared with direct MD simulations. In chapter 3, 3D dislocation loop emission from a stressed crack tip in single crystal Cu is first studied. The transition state of dislocation loop emission is found using reaction pathway sampling schemes, the nudged elastic band and dimer methods. The saddle-point geometry of the embryonic dislocation loop and the associated activation energy are quantitatively compared with the continuum estimate. Implications concerning homogeneous dislocation nucleation in the presence of a crack-tip stress field are discussed. Then, the quasi-static crack extension in Si is studied via quantitative characterization of 3D lattice trapping barriers to bond breakings at the crack tip. The physical origin of directional anisotropy in brittle fracture is revealed from the analyses of the energetics of crack font kink-pair formation. In chapter 4, the chemical reaction of hydrolysis in a silica nanorod is studied using quantum mechanical molecular orbital theory calculations. Both the stress-dependent thermodynamics and kinetics for the hydrolysis reaction are quantitatively characterized using reaction pathway sampling schemes. Implications con-

cerning the stress-corrosion crack growth in silica glass are discussed. Finally, chapter 5 summarizes main results of this thesis work.

Chapter 2

Nanoindentation-Induced Homogeneous Dislocation Nucleation

2.1 Introduction

Nanoscale contact of material surfaces provides an opportunity to explore and better understand the elastic limit and incipient plasticity in crystals. A common mode of instability at the elastic limit of a perfect crystal is strain localization, resulting ultimately in the homogeneous nucleation of crystalline defects such as dislocations [38, 75, 115]. In load-controlled nanoindentation experiments, after a preliminary elastic stage, a discontinuity in indenter displacement has been captured for the measured load P versus indentation depth h response for several cubic crystals [72, 69, 31, 21, 95, 35]. The underlying mechanism for the onset of this displacement burst is thought to indicate atomically localized deformation, *i.e.* dislocation nucleation. Recent *in situ* experiments by [35] using the Bragg-Nye bubble raft as an analogue to the {111} plane of face-centered cubic (FCC) crystal clearly demonstrated that nanoindentation of a two-dimensional crystal may lead to homogeneous nucleation of dislocations within the crystal. Atomistic simulations by [97] via the

quasicontinuum method revealed the process of nanoindentation-induced dislocation nucleation near the surface in single crystal aluminum. However, a quantitative understanding of the critical condition for homogeneous dislocation nucleation is still lacking. An accurate characterization of defect nucleation serves two purposes in multiscale materials modeling. First, a reliable defect nucleation criterion can be extracted to seed the initial defect distribution and originate new defects in mesoscopic dislocation dynamics simulations [28, 86]. Second, quantum mechanical ideal strength calculations that reveal bonding and electronic structure characteristics [51, 67] can be connected with the experimental nanoindentation P-h response. Specifically, as *ab initio* calculations are computationally intensive, we may rely on an empirical interatomic potential to identify the nucleation site and the critical stress state, and then analyze the corresponding electronic structure characteristics using *ab initio* calculations. As a continuing effort of predictive modeling of defect nucleation [58, 106], in this chapter [122], we present an atomistically faithful characterization of the critical condition of contact-induced homogeneous dislocation nucleation, based on three essential elements: (1) a computationally efficient continuum method to perform simulations with experimentally relevant length/time scales; (2) an atomistically verified instability criterion to predict dislocation nucleation; and (3) an optimally parameterized interatomic potential to model the constitutive response.

Finite strain elastic instability of perfect crystal during nanoindentation can be studied using either an atomistic or a continuum approach. The most detailed information beneath the indenter is revealed by direct molecular dynamics (MD) simulations [45, 125]. However, due to the computational cost associated with keeping track of large number of atoms, the physical length and time scales for MD simulations are very limited. In contrast, continuum-level finite element method (FEM) calculations with a hyperelastic constitutive relation based on an interatomic potential represent a computationally efficient method to study nonlinear instability at comparatively larger length scales and for realistic boundary conditions, into which crystallographic and nonlinear (de)bonding features of the underlying lattice can be incorporated. The basic premise of this approach is that every point in a continuum corresponds to

a large, uniformly deformed region at the atomic scale. Hence the constitutive relation for each continuum point can be derived within the framework of hyperelasticity with the Cauchy-Born hypothesis [14, 26, 98, 71, 2]. This hypothesis states that the Bravias lattice vectors deform according to the macroscopic deformation gradient. Based on this geometrical connection, one can calculate the local continuum stress by deforming the underlying crystal structure according to the local, atomically uniform deformation gradient, and resorting to the virial sum of interatomic forces [14, 108, 4]. Since the constitutive relation is obtained directly from the atomic lattice sum and interatomic potential, key properties of the crystal such as crystalline anisotropy and nonlinear elastic effects are incorporated automatically. This approach describes the crystal behavior well as long as the spatial variation of deformation is not too large on the atomic scale; it will break down near defects where non-local effects become significant. Interatomic potential-based hyperelasticity is a full-continuum model and represents the local limit of the more general continuum-atomistic framework of the quasicontinuum method [98, 71, 99, 85, 49]. However, for the present investigation of nanoindentation-induced elastic instability, deformation is approximately uniform at the atomic scale - though it varies significantly on the macroscale - and thus it is adequate to describe the crystal behavior accurately [58, 106]. Interatomic potential-based hyperelasticity is significantly more straightforward to implement within any general-purpose FEM package than the more flexible and thus more complex quasicontinuum approach. In this framework, quantitative prediction of defect nucleation may be achieved with the aid of an accurate and reliable defect nucleation criterion.

From the viewpoint of multiscale material modeling, one of the most useful insights that could be gained from atomistic analyses of nanoindentation is the criterion for defect nucleation. A physically rigorous criterion is much needed to replace *ad hoc* criteria that one must otherwise adopt, such as a critical resolved shear stress (CRSS)-based criterion. It is tempting to postulate that the critical stress a perfect crystal can sustain is a robust material constant. We believe this is not the case because the value of critical stress for the criterion depends on other stress components than just the shear stress component acting on the plane. Density functional theory (DFT)

calculations [67] have shown that the ideal shear strength of a Cu perfect crystal depends strongly on other components of the stress. The ideal shear strength of FCC simple metals is defined as the maximum shear stress that the crystal can sustain when sheared on the $\{111\}$ plane in the positive $\langle 11\bar{2} \rangle$ direction. For Cu, the DFT results by Ogata *et al.* [67] gave the maximum shear stress as 2.16 GPa for relaxed shear (all stress components except the principal shear stress are zero) and 3.42 GPa for unrelaxed shear (all strain components except the principal shear strain are zero). This shows that the value of the maximum shear stress depends on specifying the full local stress environment of the nucleation site. Furthermore, the detailed study of dislocation nucleation by Shenoy *et al.* [86] indicated that the stress-based nucleation criterion is able to capture only the qualitative features of the instability point induced via nanoindentation. A significant quantitative difference is found between their stress-based predictions and atomistic simulations. This is another piece of evidence pointing to the limitations of the CRSS-based criterion for predicting homogeneous dislocation nucleation under various loading conditions.

In this study, a localization criterion based on bifurcation analysis with atomistic interactions will be employed to predict nanoindentation-induced dislocation nucleation. This criterion was first proposed by Hill[38, 39]. He showed that the loss of strong ellipticity in the strain energy function is an indication of the instability in a solid. A review of localization analysis and its application for various types of continuum inelastic material models was given by Rice[75]. Several recent studies demonstrated the applicability of this criterion in the atomically-informed continuum model where the link between the movement of atoms and the deformation of the continuum was made via the Cauchy-Born hypothesis. For example, Gao and Klein[29], Klein and Gao [48], and Zhang *et. al.*[118] applied this localization criterion to detect strain localization for virtual internal bond (VIB) material models, by which the cohesive interactions between the material particles are incorporated into the constitutive law with recourse to the Cauchy-Born hypothesis. Van Vliet *et. al.*[106] used this criterion to detect dislocation nucleation in 2D simulations of nanoindentation in a model material characterized by a generic interatomic poten-

tial. As shown in the following section, the application of this localization criterion requires knowledge of the nonlinear stresses and elastic moduli, which can be readily and rigorously calculated within the framework of interatomic potential-based hyperelasticity. Given these atomically informed continuum quantities, the energy-based instability criterion is an effective tool to predict defect nucleation within a crystal, as it can identify the location, character, and stress state required for homogeneous nucleation of a crystalline defect.

The outline of this chapter is as follows. In section 2.2, the formulation of hyperelasticity which incorporates embedded atom model (EAM) interatomic potentials is developed within the framework of continuum mechanics. Next, the elastic instability criterion for defect nucleation is formulated, and the accuracy of EAM potentials is verified by comparison with *ab initio* electronic structure calculations. In section 2.3, the analytic solutions for frictionless indentation on a linear anisotropic elastic, single crystal half space are given for both 2D and 3D loading configurations based on the Stroh formalism, and analytic insights are used to rationalize the finite element simulations. In section 2.4, the 2D finite element simulations of nanoindentation based on hyperelasticity and predictions of dislocation nucleation by the instability criterion are validated via comparison with molecular dynamics simulations. The critical state of nanoindentation-induced dislocation nucleation is quantified with full 3D finite element simulations in section 2.5, and discussions and concluding remarks are presented in section 2.6.

2.2 Interatomic Potential-Based Hyperelasticity of Crystalline Solids

2.2.1 Formulation

An interatomic potential-based constitutive relation can be derived within the framework of hyperelasticity with the Cauchy-Born hypothesis [14, 26, 98, 71, 2]. In continuum mechanics, we may identify a stress-free configuration denoted by $\mathcal{A}$. The

current deformed configuration is denoted by $\mathcal{B}$. A line element $d\mathbf{x}_{\mathcal{A}}$ in configuration $\mathcal{A}$ is deformed into a line element $d\mathbf{x}_{\mathcal{B}}$ in configuration $\mathcal{B}$ by the macroscopic deformation gradient $\mathbf{F} \equiv \partial \mathbf{x}_{\mathcal{B}} / \partial \mathbf{x}_{\mathcal{A}}$. A fundamental postulate of hyperelasticity is the existence of a strain energy function, *e.g.*, Ogden [68]. Then, the symmetric second Piola-Kirchhoff stress is defined by

$$\mathbf{S}^{\mathrm{PK2}} \equiv \frac{\partial \Psi}{\partial \mathbf{E}}, \tag{2.1}$$

with $\mathcal{A}$ as the reference configuration, where Ψ is the strain energy density function (per unit stress-free volume) and the Green strain $\mathbf{E}$ is defined by

$$\mathbf{E} \equiv \frac{1}{2}(\mathbf{F}^{\mathrm{T}}\mathbf{F} - \mathbf{1}). \tag{2.2}$$

The Cauchy stress at the current configuration $\mathcal{B}$ is obtained by

$$\boldsymbol{\sigma} = \frac{1}{\det(\mathbf{F})}\mathbf{F}\mathbf{S}^{\mathrm{PK2}}\mathbf{F}^{\mathrm{T}}. \tag{2.3}$$

The link between the deformation of a continuum and that of the underlying lattice is made through the Cauchy-Born hypothesis [26], which states that Bravais lattice vectors of the crystal, $\mathbf{b}^n$ $(n = 1, 2, 3)$, deform according to the macroscopic deformation gradient $\mathbf{F}$,

$$\mathbf{b}^n_{\mathcal{B}} = \mathbf{F}\mathbf{b}^n_{\mathcal{A}} \,. \tag{2.4}$$

In monatomic crystals like FCC and BCC simple metals, there are no internal degrees of freedom, and the separation $\mathbf{r}$ between any two atoms also transforms as

$$\mathbf{r}_{\mathcal{B}} = \mathbf{F}\mathbf{r}_{\mathcal{A}}. \tag{2.5}$$

When the deformations of the continuum and the lattice are linked via the Cauchy-

Born hypothesis, the continuum strain energy density can be calculated from the interatomic potential energy of underlying lattice. Here, we have assumed temperature $T = 0$, so there is no thermal fluctuation. As all atoms are identical, we may consider the energy of one atom at the origin to be representative. Within the framework of the embedded-atom method [23], the energy per atom $\hat{\Psi}$ is

$$\hat{\Psi} = \frac{1}{2} \sum_i V(|\mathbf{r}^i|) + U(\rho) \tag{2.6}$$

where V is the pair potential, ρ is the ambient electron density for the atom at the origin and U is the energy required to embed this atom into this electron density. In Eq. (2.6), $|\mathbf{r}^i|$ is the interatomic distance and the index i runs over all atoms within a specified cut-off radius R_{cut}. Since each atom occupies a volume of a primitive unit cell, the strain energy density Ψ is related to the energy per atom $\hat{\Psi}$ by $\Psi = \hat{\Psi}/\Omega_{\mathcal{A}}$, where $\Omega_{\mathcal{A}}$ is the volume of the stress-free primitive unit cell. Given Ψ, Cauchy stress $\boldsymbol{\sigma}$ at the current configuration $\mathcal{B}$ can be obtained by substituting Eq. (2.6) into Eq. (2.3),

$$\boldsymbol{\sigma} = \frac{1}{\Omega_{\mathcal{B}}} \sum_i \left[\frac{1}{2} \frac{\partial V}{\partial r} + \frac{\partial U}{\partial \rho} \frac{\partial \rho}{\partial r} \right] \frac{\mathbf{r}^i_{\mathcal{B}} \otimes \mathbf{r}^i_{\mathcal{B}}}{|\mathbf{r}^i_{\mathcal{B}}|}, \tag{2.7}$$

where $\Omega_{\mathcal{B}}$ is the volume of the current primitive unit cell, and all derivatives in Eq. (2.7) are evaluated at $\mathbf{r}^i_{\mathcal{B}}$.

Next, the atomistic expression for the tangent modulus is introduced. Consider a further deformation from the current configuration $\mathcal{B}$ to a new configuration $\mathcal{C}$. A line element $\mathrm{d}\mathbf{x}_{\mathcal{B}}$ in configuration $\mathcal{B}$ is deformed into a line element $\mathrm{d}\mathbf{x}_{\mathcal{C}}$ in configuration $\mathcal{C}$ by the deformation gradient $\mathbf{F}^* \equiv \partial \mathbf{x}_{\mathcal{C}} / \partial \mathbf{x}_{\mathcal{B}}$. The second Piola-Kirchhoff stress with reference to configuration $\mathcal{B}$ is given by $\mathbf{S}^{*\mathrm{PK2}} \equiv \partial \Psi^* / \partial \mathbf{E}^*$, where Ψ^* is the strain energy per unit volume with reference to $\mathcal{B}$ and the Green strain $\mathbf{E}^* \equiv \frac{1}{2}(\mathbf{F}^{*\mathrm{T}} \mathbf{F}^* - 1)$. The energy density Ψ^* is obtained from the energy per atom $\hat{\Psi}$ by $\Psi^* = \hat{\Psi}/\Omega_{\mathcal{B}}$. The rate of $\mathbf{S}^{*\mathrm{PK2}}$ is related to the rate of Green strain $\mathbf{E}^*$ by

$$\dot{\mathbf{S}}^{*\mathrm{PK2}} = \mathbf{C}^* \, \dot{\mathbf{E}}^*, \tag{2.8}$$

where $\mathbf{C}^*$, which is a function of $\mathbf{E}^*$, is defined as

$$\mathbf{C}^* \equiv \frac{\partial^2 \Psi^*}{\partial \mathbf{E}^* \partial \mathbf{E}^*}. \tag{2.9}$$

The tangent modulus $\mathbf{C}$ is $\mathbf{C}^*$ evaluated at $\mathbf{E}^* = \mathbf{0}$,

$$\mathbf{C} \equiv \left. \frac{\partial^2 \Psi^*}{\partial \mathbf{E}^* \partial \mathbf{E}^*} \right|_{\mathbf{E}^*=\mathbf{0}} \tag{2.10}$$

$$= \frac{1}{\Omega_{\mathcal{B}}} \left\{ \sum_i \left[\frac{1}{2} \left(\frac{\partial^2 V}{\partial r^2} - \frac{1}{|\mathbf{r}^i_{\mathcal{B}}|} \frac{\partial V}{\partial r} \right) + \frac{\partial U}{\partial \rho} \left(\frac{\partial^2 \rho}{\partial r^2} - \frac{1}{|\mathbf{r}^i_{\mathcal{B}}|} \frac{\partial \rho}{\partial r} \right) \right] \frac{\mathbf{r}^i_{\mathcal{B}} \otimes \mathbf{r}^i_{\mathcal{B}} \otimes \mathbf{r}^i_{\mathcal{B}} \otimes \mathbf{r}^i_{\mathcal{B}}}{|\mathbf{r}^i_{\mathcal{B}}|^2} \right.$$
$$\left. + \frac{\partial^2 U}{\partial \rho^2} \left(\sum_i \frac{\partial \rho}{\partial r} \frac{\mathbf{r}^i_{\mathcal{B}} \otimes \mathbf{r}^i_{\mathcal{B}}}{|\mathbf{r}^i_{\mathcal{B}}|} \right) \left(\sum_i \frac{\partial \rho}{\partial r} \frac{\mathbf{r}^i_{\mathcal{B}} \otimes \mathbf{r}^i_{\mathcal{B}}}{|\mathbf{r}^i_{\mathcal{B}}|} \right) \right\}. \tag{2.11}$$

In the above expression, all derivatives are evaluated at $\mathbf{r}^i_{\mathcal{B}}$.

The interatomic potential-based constitutive model described above is implemented in the finite element program Abaqus [1] by writing a "user material" subroutine. In the dynamic, explicit computational procedures of this package, the nonlinear response is obtained incrementally, given the internal forces created by the stresses in the elements, as well as the applied external forces at the start of an increment, time t. Finite element procedures solve for the acceleration at the start of the increment by solving the discretized local equations of motion. The velocities at time $t + \Delta t/2$ and the displacements at time $t + \Delta t$ are updated by a central difference time-integration procedure. The deformation gradient for each integration point at time $t + \Delta t$ is then calculated based on the updated displacement field. Given the calculated deformation gradient, a constitutive equation subroutine is required in order to calculate the stress in the element at time $t + \Delta t$. In the implementation of stress calculation according to the Cauchy-Born hypothesis, each material point is represented by an FCC lattice, always larger than a sphere of radius R_{cut}, which deforms according to the local continuum deformation gradient. That is, at the beginning of the calculation ($t = 0$), a set of neighboring atoms is created to represent the atomic environment. The lattice spacing is chosen such that the corresponding stress is zero. For each time increment,

those neighboring atoms update their positions according to the local deformation gradient $\mathbf{F}$, which is generated according to the imposed boundary conditions. Then Cauchy stress $\boldsymbol{\sigma}$ and tangent modulus $\mathbf{C}$ are calculated by substituting the deformed positions of neighboring atoms into Eq. (2.7) and Eq. (2.11), respectively. Thus, material properties depend exclusively on the atomistic description of the system. The efficiency of finite element calculations allows for the simulation of systems which are large compared to those achievable via molecular dynamics calculations, while the computation remains faithful to atomistic interactions at large strains. As such, this method qualifies as a multiscale approach. In the present study, the quasi-static simulations of nanoindentation are made via the dynamic, explicit procedure at low loading rates. The explicit solution method has proven to be effective in solving quasi-static problems, and is particularly efficient for large, three-dimensional simulations dominated by contact.

2.2.2 Instability Criterion for Predicting Defect Nucleation

The homogeneous nucleation of dislocations results from catastrophic elastic instability, as formulated at the continuum level by Hill [38, 39]. Hill showed that the loss of strong ellipticity of the strain energy density function is an indication of elastic instability of a solid, because at this point discontinuous modes of deformation become admissible solutions to the equilibrium equations. Rice [75] derived the same localization condition by admitting an incremental displacement jump across the discontinuous interface. The localized deformation manifests itself along this interface as slip or shear bands in inelastic materials. Here, we summarize the main results based on Rice's derivation and then connect these concepts to the non-linear elastic model that incorporates an appropriate interatomic potential. In the atomistic limit, the shear band represents unit shear, or a single dislocation characterized by its slip plane and slip direction.

The instability criterion given in this section can be applied in both FEM and MD simulations. Since MD simulations do not store information on the starting configuration and deformation history, it will be preferable to derive the localization

condition with reference to the current configuration $\mathcal{B}$. To this end, we first consider a further deformation from the current configuration $\mathcal{B}$ to a new configuration $\mathcal{C}$ and derive the instability criterion at configuration $\mathcal{C}$. Then, we take configuration $\mathcal{C}$ to be coincident momentarily with configuration $\mathcal{B}$ and obtain the instability criterion at the current configuration. The resulting expression will be suitable for both FEM and MD implementations. Rice [75] assumed the following rate relation to describe the material response,

$$\dot{\mathbf{S}}^* = \mathbf{L}\dot{\mathbf{F}}^{*\mathrm{T}} \tag{2.12}$$

where $\mathbf{L}$ is the modulus tensor and the nominal stress tensor $\mathbf{S}^*$ is defined so that $\mathbf{n}\mathbf{S}^*$ is the force acting, per unit area in the configuration $\mathcal{B}$, on a surface element of normal vector $\mathbf{n}$ in $\mathcal{B}$. It is related to the second Piola-Kirchhoff stress $\mathbf{S}^{*\mathrm{PK2}}$, which is defined with reference to $\mathcal{B}$, by $\mathbf{S}^* = \mathbf{S}^{*\mathrm{PK2}}\mathbf{F}^{*\mathrm{T}}$. The localization condition is determined by examining whether a bifurcation within a band of orientation $\mathbf{n}$ is possible such that the continuing kinematical and equilibrium conditions are satisfied. To maintain continuous velocity across the imaginary interface between stable and unstable regions, the change in deformation gradient $\dot{\mathbf{F}}^*$ across the discontinuous surface must be of the form,

$$\Delta\dot{\mathbf{F}}^* = \mathbf{g} \otimes \mathbf{n}, \tag{2.13}$$

where $\mathbf{g}$ is the relative displacement vector across the interface. The continuous equilibrium condition requires that the difference of $\dot{\mathbf{S}}^*$ across the discontinuous surface satisfies,

$$\mathbf{n}\Delta\dot{\mathbf{S}}^* = 0. \tag{2.14}$$

Then, substitution of the constitutive law defined by Eq. (2.12) and the continuing kinematical condition given by Eq. (2.13) into the equilibrium condition Eq. (2.14)

gives,

$$(\mathbf{nLn})\mathbf{g} = 0. \tag{2.15}$$

The onset of localization occurs when a non-trivial solution of $\mathbf{g}$ exists. That is, the acoustical tensor $\mathbf{Q}(\mathbf{n}) \equiv \mathbf{nLn}$ becomes singular [75],

$$\det[\mathbf{Q}(\mathbf{n})] = 0. \tag{2.16}$$

From energy analysis, it can be proven that the acoustical tensor $\mathbf{Q}(\mathbf{n})$ is positive definite when the material is stable, *e.g.*, Rice [75]. Therefore, the determinant function $\det[\mathbf{Q}(\mathbf{n})]$ will approach zero from the positive side. The dislocation will be homogeneously nucleated when the determinant function $\det[\mathbf{Q}(\mathbf{n})]$ attains zero-value for the first time. The associated eigenvector $\mathbf{g}$ calculated from the matrix of $\mathbf{Q}(\mathbf{n})$ will represent the slip vector at the onset of dislocation nucleation.

The instability condition based on the above bifurcation analysis will be applied to the current configuration $\mathcal{B}$ to detect dislocation nucleation once Cauchy stress $\boldsymbol{\sigma}$ and tangent modulus $\mathbf{C}$ are calculated through either FEM or MD simulations. To study bifurcation in configuration $\mathcal{B}$, we take configuration $\mathcal{C}$ to be coincident momentarily with configuration $\mathcal{B}$, *i.e.*, $\mathbf{F}^* = \mathbf{1}$. Under this condition, the following relations can be easily derived,

$$\dot{\mathbf{S}}^* = \dot{\mathbf{S}}^{*\mathrm{PK2}} + \boldsymbol{\sigma}\dot{\mathbf{F}}^{*\mathrm{T}}, \tag{2.17}$$

$$\dot{\mathbf{E}}^* = \frac{1}{2}\left(\dot{\mathbf{F}}^* + \dot{\mathbf{F}}^{*\mathrm{T}}\right). \tag{2.18}$$

Substituting Eqs. (2.17) and (2.18) into Eq. (2.12), and using Eq. (2.8), one obtains,

$$L_{ijkl} = C_{ijkl} + \sigma_{il}\delta_{jk}. \tag{2.19}$$

Cauchy stress $\boldsymbol{\sigma}$ in configuration $\mathcal{B}$ arises in the expression of the modulus tensor $\mathbf{L}$ because the current configuration $\mathcal{B}$ is generally a stressed state. Since the modulus tensor $\mathbf{L}$ is calculated according to the atomistically informed Cauchy stress $\boldsymbol{\sigma}$ and tangent modulus $\mathbf{C}$ in configuration $\mathcal{B}$, as defined by Eq. (2.7) and Eq. (2.11), the crystal lattice information is incorporated automatically. This tensor has the symmetry

$$L_{ijkl} = L_{lkji}, \tag{2.20}$$

but,

$$L_{ijkl} \neq L_{jikl}, \quad L_{ijkl} \neq L_{ijlk}, \tag{2.21}$$

as $\mathbf{g}$ and $\mathbf{n}$ are conceptually distinct vectors. Thus, Eq. (2.16) fully defines the site and slip character of dislocations nucleated within simulations of calculable $\boldsymbol{\sigma}$ and $\mathbf{C}$.

Homogeneous nucleation of a defect has been studied in the context of elastic wave instability by Li *et al.* [58]. That is, the defect nucleation is realized by a dynamic process which comprises the progressive steepening of a localized wave front and arrest of the final atomistically sharp wave front in a low-dimensional atomic energy landscape. The free energy-based Λ-criterion for wave instability [58] is equivalent to the defect nucleation criterion given by Eq. (2.16). The wavevector $\mathbf{k}$ and the polarization vector $\mathbf{w}$ for the unstable transverse wave corresponds to the slip plane normal $\mathbf{n}$ and the associated slip vector $\mathbf{g}$, respectively. Though the current study is not concerned with dynamics, the defect nucleation criterion given by Eq. (2.16) is still referred to as the Λ-criterion.

Table 2.1: Elastic constants and intrinsic stacking fault energies for Cu

Method	C_{11} (GPa)	C_{12} (GPa)	C_{44}(GPa)	γ_{sf} (mJ/m^2)
Experiment	168.4	121.4	75.4	45
EAM-Mishin	169.9	122.6	76.2	44.4
EAM-Ackland	168.4	121.3	75.3	46.8

2.2.3 Calibration of EAM Potentials by ab initio Calculation

Recent advances in optimal parametrization of the interatomic potential [64, 65] enable improved accuracy of the constitutive model of interatomic potential-based hyperelasticity, within its intrinsic limitations. The EAM potential for Cu [65] is employed in the present indentation simulations. The elastic constants (at the stress-free state) and stacking fault energies obtained from experiments and calculated via the EAM-Mishin potential are given in Table 2.1. For comparison, the corresponding values calculated from another commonly used EAM potential by Ackland *et al.*[3] are also listed in Table 2.1.

A reliable potential should both fit well the single point properties such as elastic constants and the intrinsic stacking fault energy, and also characterize accurately the deformation path up to the elastic limit. Consider a homogeneous simple shear deformation on the (111) plane along the $[11\bar{2}]$ direction. Figure 2-1 shows the shear stress versus shear strain response predicted by the EAM-Mishin potential. Here, the shear strain is defined by the shear displacement divided by the Shockley partial slip in the $[11\bar{2}]$ direction $a_0/\sqrt{6}$, where a_0 is the lattice constant. In order to test the accuracy of the EAM-Mishin potential, an *ab initio* DFT calculation is performed whereby simple shear deformation is applied to a six-atom supercell of three $\{111\}$ layers. The stress versus strain response in the DFT calculation is obtained by using the generalized gradient approximation (GGA) within an ultrasoft pseudopotential total energy scheme [67]. It can be seen from Fig. 2-1 that the prediction via the EAM-Mishin potential is in good agreement with the *ab initio* calculation. Figure 2-1 also shows the shear stress versus strain curve based on the EAM-Ackland potential.

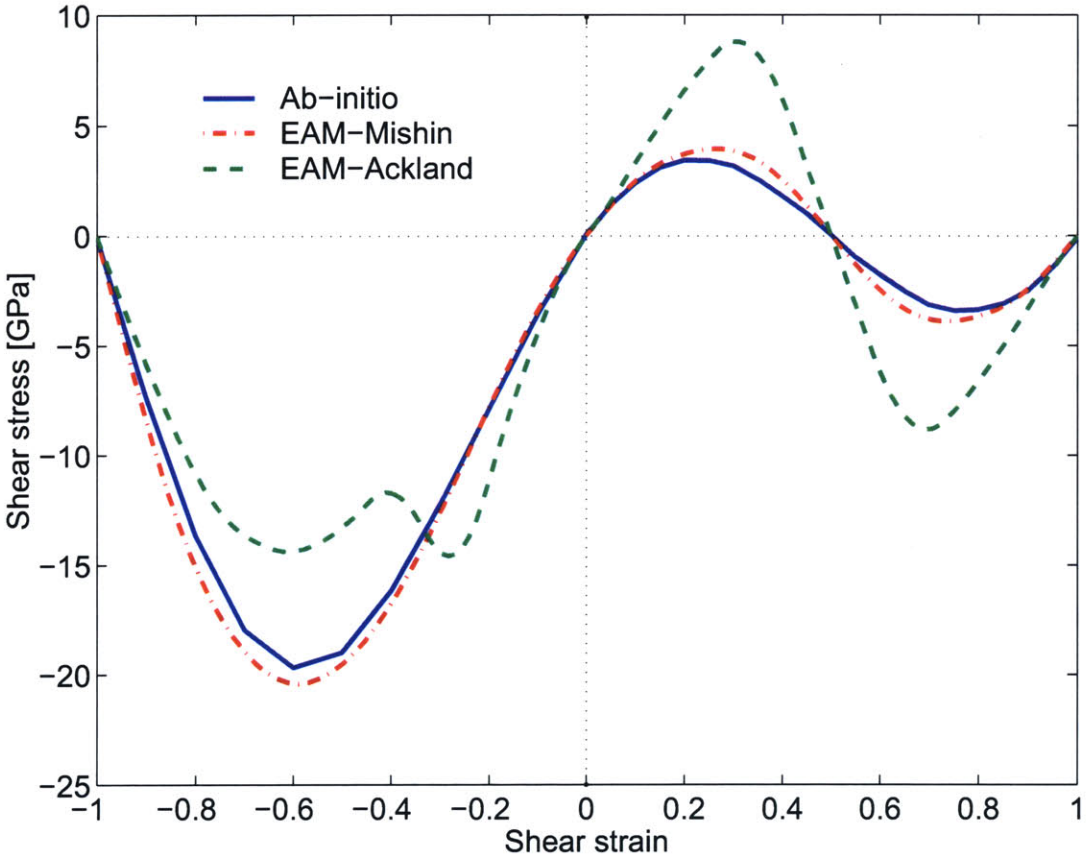


Figure 2-1: Stress versus strain curve for simple shear of Cu.

Compared with the *ab initio* DFT calculation, the EAM-Mishin potential gives a more accurate description of finite shear deformation than does the EAM-Ackland potential. We demonstrate below that the choice of a reliable potential is critical for an accurate prediction of both the load versus displacement response and mode of dislocation nucleation in nanoindentation simulations.

The above simple shear calculations also reveal an important feature at finite deformation, *i.e.*, asymmetry of shear stress response with respect to the shear direction. It can be seen from Fig. 2-1 that the critical shear stress in the $[11\bar{2}]$ direction (3.91 GPa) is much lower than the peak stress when sheared in the opposite direction, *i.e.*, the $[\bar{1}\bar{1}2]$ direction (20.4 GPa). This asymmetry in critical shear stress is due to the effect of finite deformation, and can be understood by examining the relative motion between adjacent close-packed planes in an FCC crystal. In Figure 2-2, the dashed circles represent one atomic layer of the close-packed (111) plane. The solid circles represent atoms above this layer, and are originally located in sites such as the one denoted by a. The shear along the partial slip $[11\bar{2}]$ direction corresponds to motion of the atom from site a to b and the reversed shear along the $[\bar{1}\bar{1}2]$ direction from site a to c. It is immediately evident from Fig. 2-2 that, within the finite shear deformation regime, the reversed shear along the $[\bar{1}\bar{1}2]$ direction requires a larger dilatation normal to the slip plane due to the misfit of adjacent plane atoms, and hence needs to overcome a larger energy barrier. Consequently, the critical shear stress in the $[\bar{1}\bar{1}2]$ direction will be much higher relative to that in the $[11\bar{2}]$ direction. Therefore, not only the magnitude of shear stress but also the direction and the sense of shear deformation is relevant in defining and predicting dislocation nucleation. This dependence of the homogeneous dislocation nucleation barrier on shear direction is in contrast to the conventional crystal plasticity model which assumes the symmetric resistance for the motion of existing dislocations [9]. The assumption of a symmetric slip barrier is based on the statistical nature of the resistances which include forest dislocations threading through the slip plane, solute atoms or second phase particles acting as discrete obstacles in the slip plane.

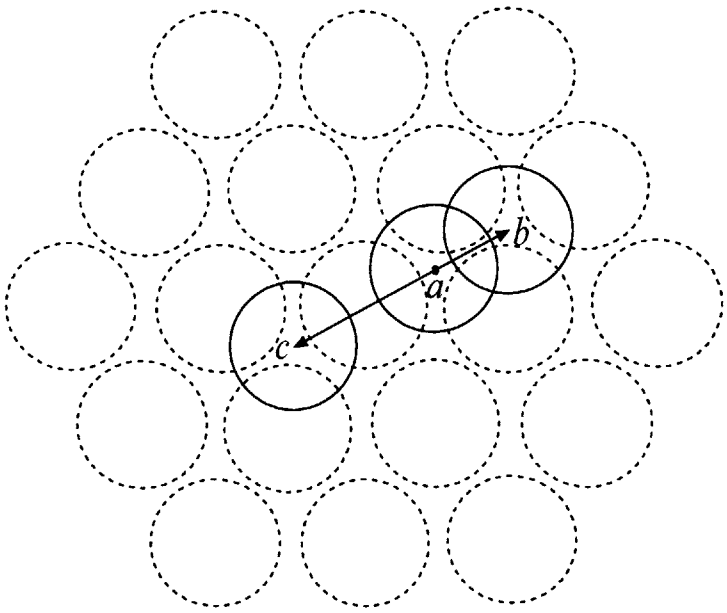


Figure 2-2: Schematics of shear along the Shockley partial slip $[11\bar{2}]$ direction (from site a to b) and reversed shear along the $[\bar{1}\bar{1}2]$ direction (from site a to c) above the close-packed (111) plane.

2.3 Indentation on a Linear Anisotropic Elastic Solid

As a complementary analysis of nanoindentation via interatomic potential-based hyperelasticity, the analytic solutions for frictionless indentation on a linear elastic, single crystal are given in this section for both 2D and 3D configurations. Analytic insights regarding the effects of crystalline anisotropy are discussed, and serve to rationalize the results of interatomic potential-based finite element simulations.

2.3.1 2D Indentation by a Cylindrical Indenter

Consider a frictionless cylindrical indenter of radius R pressed into a linear anisotropic elastic half space, as shown in Fig. 2-3. The coordinate system is oriented such that the x_3 axis is parallel to the axial direction of the cylinder. Within the regime of linear elastic deformation, the resulting displacement field is only a function of in-plane coordinates (x_1, x_2) and has no variation in the out-of-plane x_3 direction. This

is a generalized two-dimensional problem which can be solved by recourse to the Stroh formalism given in Appendix A. Fan and Hwu [27] derived the general solution of punch indentation on a linear anisotropic elastic half space. We apply this general solution to the present geometric configuration. To first order, the cylindrical profile can be approximated by a parabola. By solving a standard Hilbert problem, the derivative of the complex function $\mathbf{f}$, as defined in Appendix A, is given by

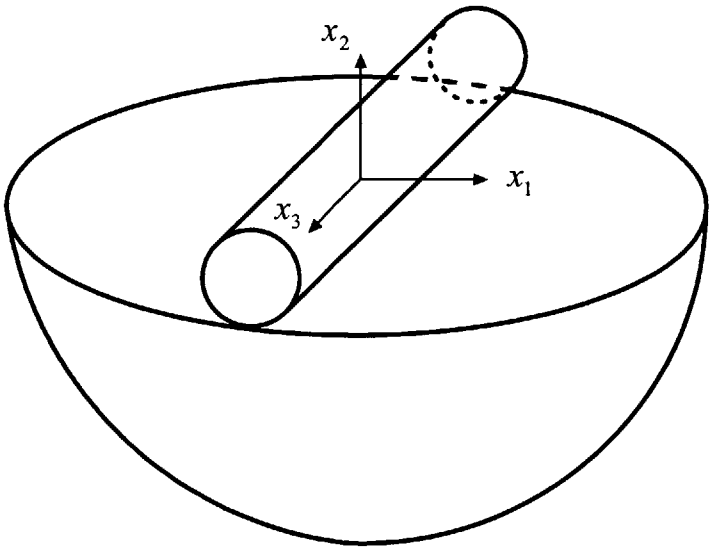


Figure 2-3: Schematics of 2D indentation: a cylindrical indenter on an elastic half space.

$$\mathbf{f}' = \frac{i}{2(\mathbf{M}^{-1})_{22}R} \begin{bmatrix} (\mathbf{B}^{-1})_{12}\left(z_1 - \sqrt{z_1^2 - a^2}\right) \\ (\mathbf{B}^{-1})_{22}\left(z_2 - \sqrt{z_2^2 - a^2}\right) \\ (\mathbf{B}^{-1})_{32}\left(z_3 - \sqrt{z_3^2 - a^2}\right) \end{bmatrix} \tag{2.22}$$

where a denotes contact half-width, and the matrices $\mathbf{B}$ and $\mathbf{M}$, as defined in Appendix A, can be determined uniquely for a given set of the material elastic constants. The subscript in Eq. (2.22) denotes the corresponding component of the matrix. The stress field is obtained by substituting the above expression into Eqs. (A.12) and (A.13). Note that a difficulty arises in deriving the 2D load-penetration relation which is absent in the 3D case: For an elastic half space loaded two-dimensionally,

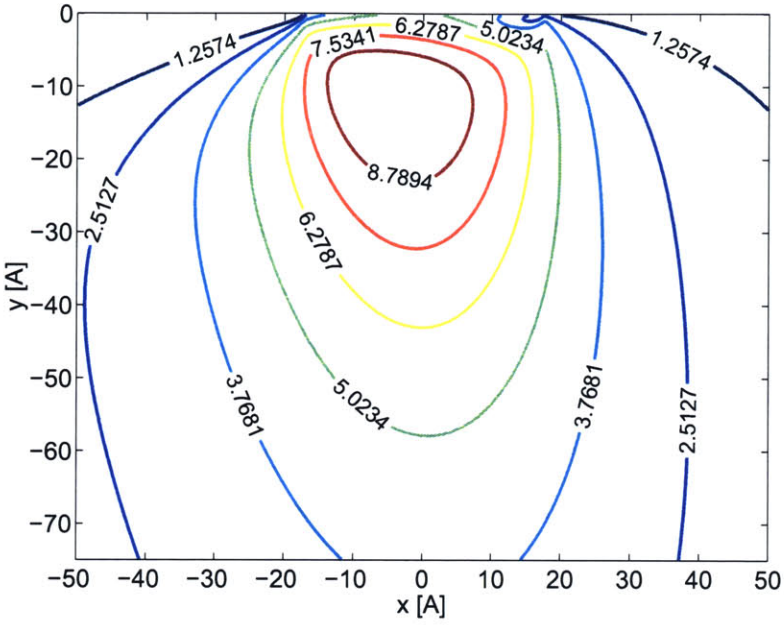
the displacement field decays logarithmically from the contact point [42]. Thus, there exist arbitrary, unresolved constants for the displacement field such that the load-penetration relation depends on the system size and the boundary conditions. However, the analytic expression relating indentation load and contact half-width can be obtained by requiring that the stress singularity at the contact edge vanishes. This explicit relation is given by

$$P = \frac{\pi a^2}{2(\mathbf{M}^{-1})_{22} R}. \tag{2.23}$$

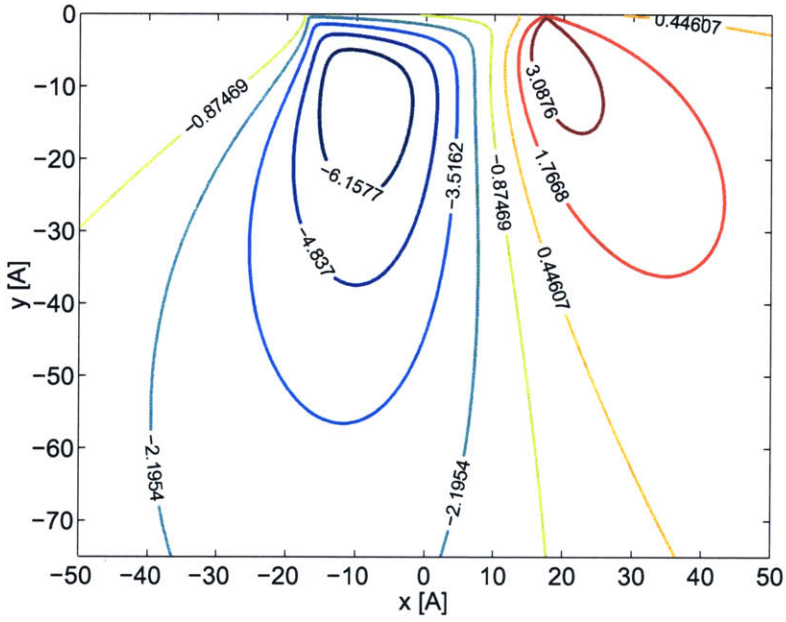
The corresponding solution for isotropic elasticity has the same functional form, except that the value $(\mathbf{M}^{-1})_{22}$ is replaced by $(1-\nu)/\mu$, where μ is shear modulus and ν is Poisson's ratio.

Based on the complex function given by Eq. (2.22), the stress distribution beneath the indenter is calculated for a cylindrical indentation normal to the (111) surface of Cu along the $[1\bar{1}0]$ direction. The coordinate system in Fig. 2-3 is oriented such that the x_1 and x_2 axes correspond with the $[11\bar{2}]$ and [111] directions, respectively. The out-of-plane x_3 axis is along the $[1\bar{1}0]$ direction. The elastic constants for single crystal Cu are taken to be the values fitted by the Mishin potential as listed in Table 2.1. The radius of the indenter R is taken to be 50 Å, and the imposed contact half-width is chosen as $a = 17.4$ Å to facilitate comparison with MD and FEM simulations. The distribution of Mises stress beneath the indenter is shown in Fig. 2-4(a). The effect of material anisotropy is revealed by comparing the Mises stress contour for an isotropic material [42] with that for single crystal Cu. For an isotropic material, this stress distribution is symmetric with respect to x_2. Taking Poisson's ratio to be 0.3, the point of maximum Mises stress is along x_2 axis at a depth of $0.70a$. For single crystal Cu, the anisotropy factor, defined by $2C_{44}/(C_{11} - C_{12})$, is 3.22. As shown in Fig. 2-4(a), the Mises stress contour is not symmetric about x_2 due to this elastic anisotropy. The point of maximum Mises stress is off-center by $0.24a$ and at a depth of $0.66a$. The corresponding maximum Mises stress is 10.16 GPa.

Dislocations tend to nucleate and move on certain close-packed slip planes and



(a)



(b)

Figure 2-4: Stress contour beneath a cylindrical indenter from the 2D analytic solution: (a) Mises stress, (b) Shear stress resolved into $(11\bar{1})[112]$ slip system. The numbers indicated on the contour map are in units of GPa.

directions in crystalline metals. The shear stresses resolved into the FCC primary slip systems of $\{111\}\langle 1\bar{1}0\rangle$ and $\{111\}\langle 11\bar{2}\rangle$ are calculated. For the current orientation of the underlying crystal structure with respect to the indentation loading axis, the maximum magnitude of resolved shear stress is on the Shockley partial slip system $(11\bar{1})[112]$. The slip plane $(11\bar{1})$ is oriented at 19.47° with respect to x_2 axis. Figure 2-4 (b) shows the contour of resolved shear stress on the $(11\bar{1})[112]$ slip system. The point of maximum magnitude of resolved shear stress lies to the left of x_2 axis by $0.54a$ and at a depth of $0.66a$. The corresponding resolved shear stress is -7.55GPa.

Analytic insights gained from the above linear elastic analysis can be used to rationalize numerical simulations. First, for this particular orientation, the out-of-plane displacement is null due to the mirror symmetry of the lattice with respect to the $(1\bar{1}0)$ plane (prior to the possible symmetry breaking caused by the elastic instability). Therefore, this is a true plane strain problem within the elastic deformation regime. No further approximation is made for this generalized plane strain problem when using plane strain elements in 2D finite element simulations. More importantly, the choice of the current indenter geometry will facilitate the study of homogeneous nucleation of dislocation in the bulk by mitigating surface effects. That is, for the current orientation, the resolved shear stress reaches its maximum value at some distance beneath the surface. This is in contrast with the computational nanoindentation study by Tadmor *et. al.* [97], which examined the same crystallographic orientation as considered in the present work, but under a rectangular punch indenter. The sharp corners of such an indenter induce dislocation emission from the contact point at the surface, where the interpretation of atomic stress at free surfaces is unclear [18] and the role of ledge formation is difficult to quantify unambiguously.

2.3.2 3D Indentation by a Spherical Indenter

The analytic load versus displacement relation is given in this subsection for indentation of a spherical indenter on the $\{111\}$ surface of a single crystal half space. Figure 2-5 shows a spherical indenter of radius R in contact with a linear elastic anisotropic half space. The coordinate system is oriented such that x_1, x_2 and x_3

axes correspond with the $[11\bar{2}]$, [111] and [110] directions, respectively. To first order, the spherical indenter can be approximated by a paraboloid. The general elastic solution for the indentation on an anisotropic half space by a paraboloid was first derived by [114] using the Fourier transform technique. Based on the Green's function for an elastic, anisotropic half space [11], the simplified elastic solutions for the indentation load versus displacement response were given by Vlassak and Nix [104, 105], Swadener *et al.* [96] for various indenter profiles. Willis[114] showed that the contact area produced by a paraboloid on an anisotropic, elastic half space is elliptical in shape. For indentation normal to a {111} surface with inherent three-fold rotational symmetry, the contact area is circular. The load P versus penetration h relation is simply given by

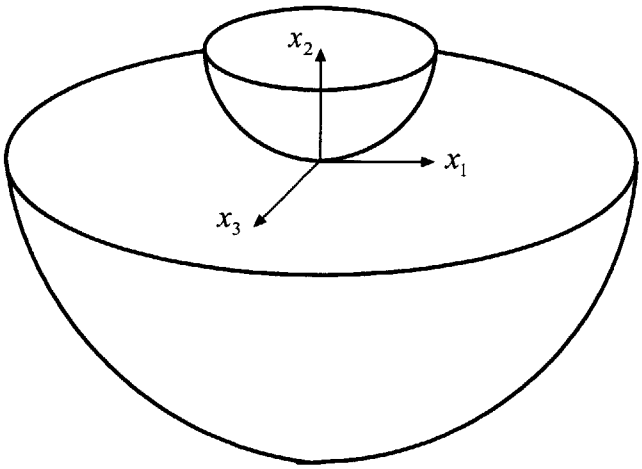


Figure 2-5: Schematics of 3D indentation: a spherical indenter on an elastic half space.

$$P = \frac{4}{3} E^* R^{1/2} h^{3/2} \tag{2.24}$$

where the indentation modulus E^* is

$$E^* = \left(\frac{1}{4\pi} \int_0^{2\pi} \hat{\mathbf{L}}_{22}^{-1}(\gamma)\, d\gamma \right)^{-1} \tag{2.25}$$

In Eq. (2.25), the Barnett-Lothe tensor $\hat{\mathbf{L}}$, defined by Eq. (A.14) in the Appendix A, is calculated in the transformed $x_1' - x_2 - x_3'$ coordinate system, which is obtained by rotating x_1 and x_3 axes in the $x_1 - x_3$ plane an angle γ about x_2 axis. For an isotropic material deformed via a rigid indenter, the indentation modulus E^* reduces to $E/(1-\nu^2)$, where E denotes Young's modulus. The above analytic solution will be used to benchmark 3D FEM simulations. For elastic constants fitted by the EAM-Mishin potential as given in Table 2.1, the value of E^* calculated from Eq. (2.25) is 153 GPa.

2.4 2D Nanoindentation-Induced Dislocation Nucleation

2.4.1 Validation of FEM Simulations

In this subsection, FEM simulations using interatomic potential-based hyperelasticity is validated by comparing 2D cylindrical indentation results with direct MD simulations. The same interatomic potential of Cu fitted by Mishin *et al.* [65] is used for the constitutive input of FEM simulations and MD interatomic force field calculations. The indentation orientation is the same as that for the linear analysis in section 3. That is, a cylindrical indenter along the $[1\bar{1}0]$ direction is imposed normal to the (111) surface.

ABAQUS/Explicit [1] is implemented to simulate 2D indentation. A small system size, 200 × 100 Å, is simulated for direct comparison with MD calculations. Plane strain linear, triangular elements are used, with each element representing a homogeneously deformed crystallite. The boundary conditions are fixed at the bottom, free on the top surface and fixed in the x_1 directions on two sides. The cylindrical indenter is defined by an analytic rigid surface with a radius R of 50 Å. The contact between the indenter and the surface is frictionless. The quasi-static solution is approximated by maintaining the kinetic energy at less than 1% of the internal energy of the system.

MD simulations at the temperature 1K are performed to compare with FEM

calculations. The same in-plane boundary conditions as FEM simulations are used and periodic boundary conditions (PBC) are applied for the out-of-plane x_3 direction. In order to follow the minimum image convention in MD coding, ten layers of atoms are used along the x_3 direction. This ensures that the half-thickness of the PBC box is greater than the Mishin potential cut-off radius. The indenter is regarded as an external repulsive potential interacting with copper atoms at the surface [45],

$$\Psi^{ext}(r) \;=\; AH(R-r)(R-r)^3 \tag{2.26}$$

where A is a force constant and $H(r)$ is the step function. In the present calculations, $A = 10\text{eV}/\text{Å}^3$. The indentation proceeds in displacement-control at a speed of about 1 m/s.

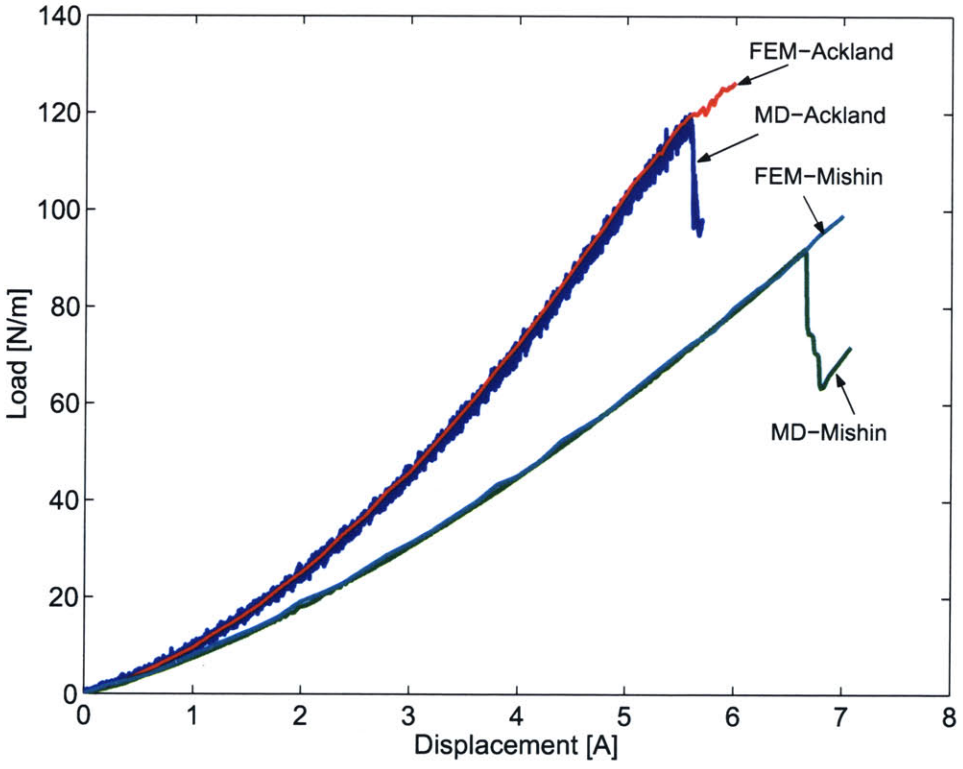


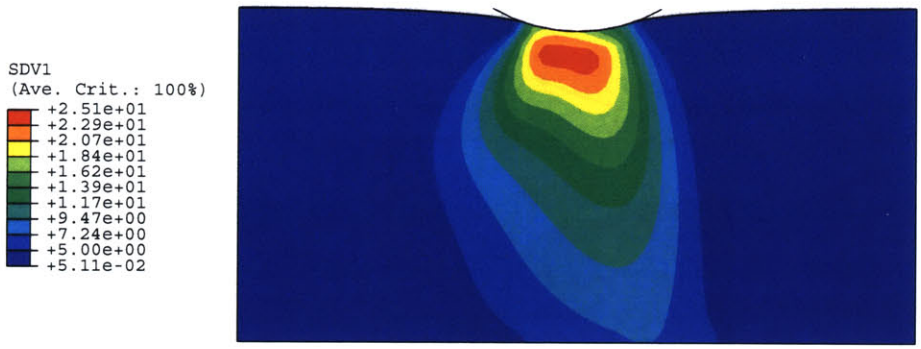
Figure 2-6: Load versus displacement curves of nanoindentation by a cylindrical indenter from FEM and MD simulations.

The calculated nanoindentation responses based on the Mishin potential are given in Fig. 2-6. It can be seen that the P–h responses predicted by FEM and MD are

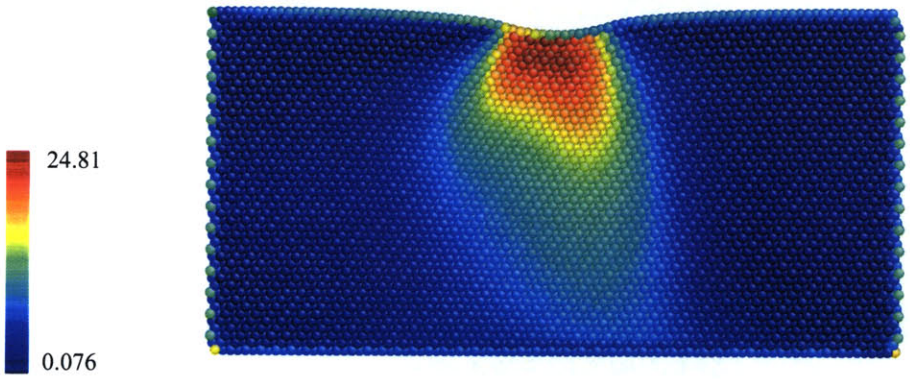
in good agreement. The MD response shows a sharp drop in load at an indentation depth of 6.65 Å, which indicates the onset of homogeneous nucleation of dislocation within the crystal. This load drop was not captured in the corresponding FEM simulation because the 2D setting constrains the out-of-plane displacement mode that the homogeneously nucleated dislocation takes, as shown in the MD simulation. Figure 2-6 also shows the P–h response calculated via the Ackland potential for the same indenter radius and crystallographic orientation. Though agreement between FEM and MD predictions based on the Ackland potential further validates the FEM calculation, the significant differences in the predicted P–h response and critical load drop point given by each of these potentials underscore the importance of optimal parametrization of the interatomic potential.

The calculated stress distribution beneath the indenter is compared between FEM and MD simulations based on the Mishin potential. FEM stress computation at each material point follows the constitutive model of interatomic potential-based hyperelasticity as given by Eq. (2.7), which is dictated by the local deformation gradient. In contrast, for every atom in the MD simulation, atomic Cauchy stresses are calculated by substituting the actual coordinates of neighboring atoms into Eq. (2.7). Thus the positions of neighboring atoms are explicitly tracked for every time step. Though the definition of atomic Cauchy stress is not strictly well-posed in terms of the continuum concept of stress, the value of atomic stress represents an effective measure of the stress state imposed on each atom. Figures 2-7(a) and (b) show the contours of Mises stress calculated from FEM and MD simulations, respectively. The indentation depth is 6.65 Å, corresponding to the indenter displacement immediately preceding dislocation nucleation. It can be seen that the FEM prediction based on interatomic potential-based hyperelasticity is in good agreement with that of the MD simulation. Both simulations give the maximum Mises stress about 25 GPa. However, the Mises stress contours shown in Figs. 2-7 (a) and (b) differ significantly from the linear elastic result for indentation on a 2D half space, as shown in Fig. 2-4 (a) for the same contact half-width as FEM simulations. These differences can be attributed to the elastic nonlinearity of the highly deformed material near the indenter and the

boundary effects due to the finite size of the simulated system.



(a)



(b)

Figure 2-7: Contours of Mises stress (in GPa) beneath a cylindrical indenter: (a) FEM and (b) MD simulations.

2.4.2 MD Simulation of Dislocation Nucleation

Though MD simulations are limited by the achievable length and time scales, this approach has the distinct advantage of allowing defects to nucleate and evolve spontaneously, without prescribing the site and nature of such defects. Thus, the atomic structure of the emerging dislocation is directly revealed by MD simulations. Figure 2-8 (a) shows the embryo of a homogeneously nucleated dislocation via MD simulation using the Mishin potential. Atoms are color-encoded by coordination number N [57].

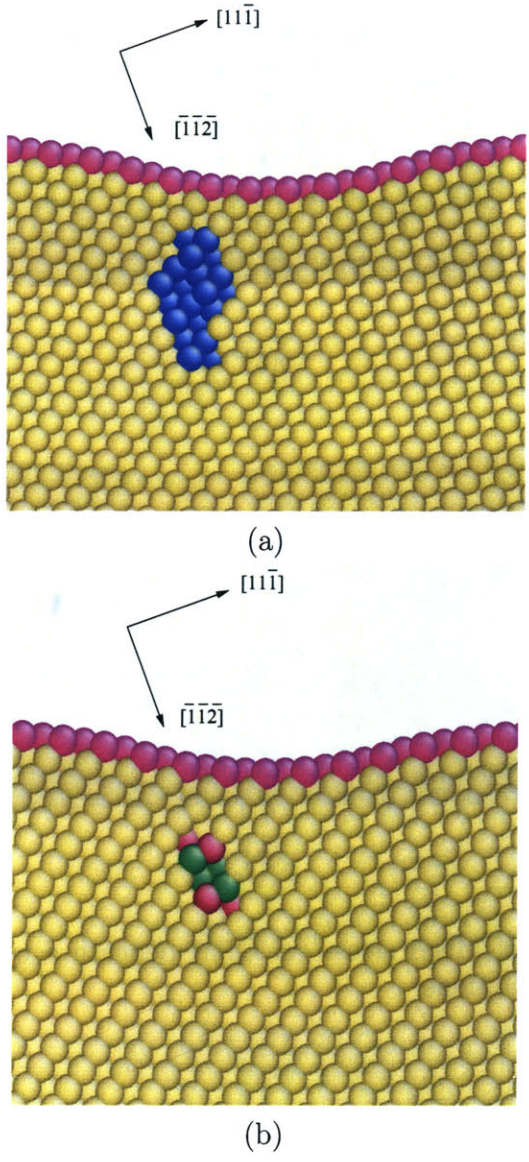


Figure 2-8: MD simulations of nanoindentation by a cylindrical indenter: atomic structure of homogeneously nucleated dislocation on the $(11\bar{1})$ plane. (a) Shockley partial slip along the $[1\bar{2}\bar{1}]$ direction from Mishin potential; (b) Shockley partial slip along the $[\bar{1}\bar{1}\bar{2}]$ direction by Ackland potential.

Perfectly coordinated atoms, $N = 12$, are color-encoded as yellow, and other colored atoms indicate $N \neq 12$. Tracing of atomic trajectories indicates that relative motion between the left and right two $(11\bar{1})$ layer blue atoms represents a Shockley partial slip along the $[1\bar{2}\bar{1}]$ direction within the $(11\bar{1})$ slip plane. This partial slip takes a mixed in-plane and anti-plane shear mode. Due to the 2D nature of MD simulations, the homogeneously nucleated defect is a dislocation line with the $[1\bar{1}0]$ line direction perpendicular to the $x_1 - x_2$ plane. Note that there is another equivalent Shockley partial slip direction $[\bar{2}1\bar{1}]$ due to mirror symmetry of the $(11\bar{1})$ slip plane with respect to the $x_1 - x_2$ plane.

For comparison, Fig. 2-8 (b) displays the atomic structure of the homogeneously nucleated dislocation via MD simulation using the Ackland potential. It can be seen that the activated slip system of the Shockley partial is on the $(11\bar{1})$ plane and along the $[\bar{1}\bar{1}\bar{2}]$ direction. In contrast to the mixed shear mode of the Shockley partial slip predicted by the Mishin potential, the simulation via the Ackland potential generates a partial slip event with a pure in-plane shear mode. This discrepancy in slip orientation can be correlated with the simple shear calculations given in Fig. 1, which shows that the Ackland potential predicts a fictitiously high critical stress when sheared in the $[11\bar{2}]$ direction and a decrease in critical shear stress when sheared in the reversed direction, with respect to calculations via the Mishin potential. When using the Ackland potential in nanoindentation simulation, due to this artifactual reduction of the shear barrier in the $[\bar{1}\bar{1}\bar{2}]$ direction (equivalent to the $[\bar{1}\bar{1}2]$ direction in Fig. 2-2), the Shockley partial slip is first developed along this direction. In contrast, the high shear barrier along the $[\bar{1}\bar{1}\bar{2}]$ direction for the Mishin potential forces the Shockley partial slip to develop in the $[1\bar{2}\bar{1}]$ direction. This slip direction is equivalent to the $[11\bar{2}]$ direction in Fig. 2-2 and hence has a much lower critical stress. The MD simulations are terminated at this point because further evolution of the dislocation structure, including heterogeneous nucleation of dislocations and interaction among dislocations, depend strongly on the loading rate and the simulated system size. This topic on massive dislocation activity beneath the indenter is addressed by VanVliet *et al.* [106].

2.4.3 Prediction of Dislocation Nucleation by Instability Criterion

Continuum level defect nucleation studies require a reliable instability criterion which should capture the moment, location and nature of homogenously nucleated defects. Below, both the interatomic-potential-based Λ-criterion and the critical resolved shear stress (CRSS) criterion are evaluated by comparing the corresponding defect nucleation predictions with those obtained via direct MD simulations.

Since all the atomic information is channeled through the Cauchy stress $\boldsymbol{\sigma}$ and tangent modulus $\mathbf{C}$, the Λ-criterion can be readily incorporated into finite element analysis. Localization indicator $\det[\mathbf{Q}(\mathbf{n})]$ is calculated at the element level to detect dislocation nucleation. Testing for the formation of a localization surface involves a search over all the possible orientations specified by the vectors $\mathbf{n}$. Given the atomistically informed $\mathbf{L}$ tensor, the general algorithm for searching critical surfaces involves a constrained 6-dimensional minimization by which the localization surface normal $\mathbf{n}$ and the direction of relative displacement vector $\mathbf{g}$, as defined in section 2.2, are determined simultaneously [58, 106]. For the FCC lattice considered here, we simply trace the current surface normals $\mathbf{n}$ of close-packed $\{111\}$ slip planes from their initial surface normals $\mathbf{n}^o$ by the relation $\mathbf{n} = \mathbf{n}^o\mathbf{F}^{-1}$ [9]. When an instability occurs, the acoustical tensor $\mathbf{Q}(\mathbf{n})$ will no longer be positive definite.

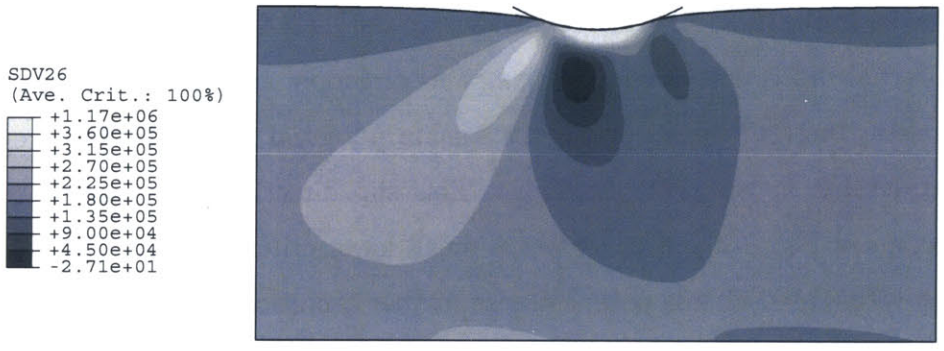
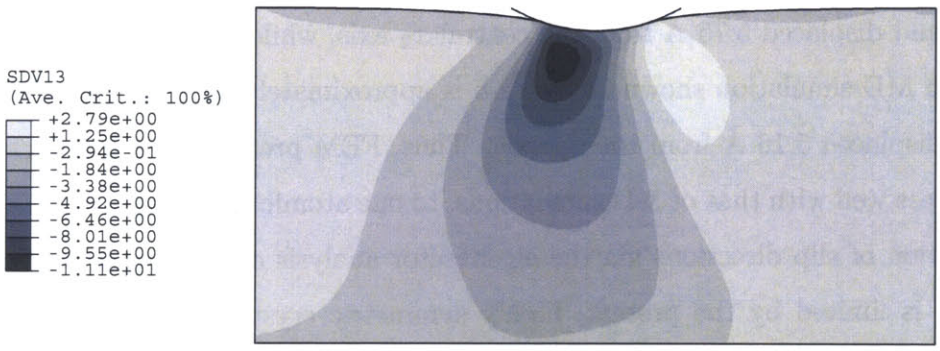


Figure 2-9: Contour of $\det[\mathbf{Q}(\mathbf{n})]$ by 2D FEM simulation.

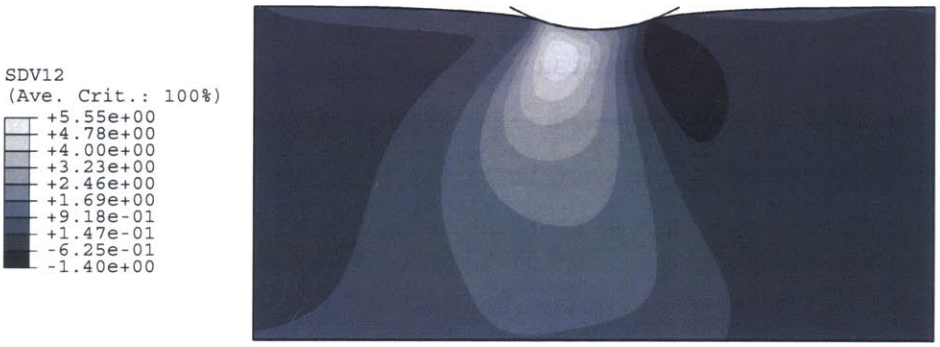
Figure 2-9 shows the contour of $\det[\mathbf{Q}(\mathbf{n})]$ at an indentation depth of 6.68 Å, cor-

responding to the moment when the onset of dislocation nucleation is first detected at one integration point. The critical indentation depth predicted by FEM calculation is in good agreement with that by MD simulation. In calculating $\det[\mathbf{Q}(\mathbf{n})]$, the modulus $\mathbf{L}$ is normalized via the elastic constant $C_{44} = 75.4$ GPa. It can be seen from Fig. 2-9 that the interpolated values of $\det[\mathbf{Q}(\mathbf{n})]$ within the dark grey region beneath the indenter are small negative values, which indicate the site of a homogeneously nucleated dislocation. The position of this dislocation nucleation site is 14.25 Å below the surface, and displaced 5.75 Å from the central x_2 axis, while the center of dislocation core in MD simulation shown in Fig. 2-8 is approximately 16.66 Å below the surface and displaced 5.16 Å from the x_2 axis. Thus, FEM prediction of the nucleation site agrees well with that of MD simulations, to one atomic lattice spacing. The direct prediction of slip directions via the eigenvector analysis of the corresponding matrix $\mathbf{Q}(\mathbf{n})$ is limited by the present, highly symmetric crystallographic orientation. MD simulations indicate that the Shockley partial slip develops along one of two equivalent slip directions, *i.e.*, $[1\bar{2}\bar{1}]$ and $[\bar{2}1\bar{1}]$, within the $(11\bar{1})$ plane. This is a degenerate-eigenvector situation which can be detected numerically by diagonalization of the matrix $\mathbf{Q}(\mathbf{n})$. For the degenerate case, the slip orientations resulting from the instability can be identified from crystallographic symmetry, without resorting to a higher-order (cubic, quartic) analysis of the strain energy.

For comparison, the stress-based dislocation nucleation criterion (CRSS) is also evaluated in FEM calculations using the constitutive model of interatomic potential-based hyperelasticity. The linear elastic analysis indicates that the maximum resolved shear stress occurs on the Shockley partial slip system $(11\bar{1})[112]$, while the MD simulation indicates that the dislocation will nucleate in one of two other equivalent slip systems, $(11\bar{1})[1\bar{2}\bar{1}]$ or $(11\bar{1})[\bar{2}1\bar{1}]$. Since the critical shear stress sensitively depends on other local stress components [67] as discussed in Introduction, the value of CRSS corresponding to the local stress environment associated with nanoindentation is *a priori* unknown. Therefore, we back out the value of CRSS within a nanoindentation-induced stress field from the Λ-criterion. Specifically, the shear stresses resolved into different slip systems are calculated when the indenter is displaced to the critical



(a)



(b)

Figure 2-10: Distribution of resolved shear stress (in GPa) by 2D FEM simulation: (a) $(11\bar{1})[112]$ slip system, (b) $(11\bar{1})[1\bar{2}\bar{1}]$ slip system.

depth for dislocation nucleation as predicted by the Λ-criterion. Figures 2-10 (a) and (b) show contours of resolved shear stress on $(11\bar{1})[112]$ and $(11\bar{1})[1\bar{2}\bar{1}]$ slip systems from FEM calculations. It can be seen from Fig. 2-10(a) that the shear stress along the [112] direction maximizes on the right side of the x_2 axis with a value about 2.8 GPa and the shear stress along the $[\bar{1}\bar{1}\bar{2}]$ direction maximizes on the left side with a value approximately 11.1 GPa. In contrast, Fig. 2-10(b) shows that the shear stress in the $[1\bar{2}\bar{1}]$ direction maximizes on the left side of the x_2 axis with a value about 5.6 GPa and the shear stress along the $[\bar{1}21]$ direction maximizes on the right side with a value about 1.4 GPa. Due to the asymmetric instability threshold for finite shear deformation, the dislocation will nucleate along the $[1\bar{2}\bar{1}]$ direction as shown in MD simulations, though the magnitude of the corresponding resolved shear stress, 5.6 GPa, is smaller than that in the $[\bar{1}\bar{1}\bar{2}]$ direction, 11.1 GPa. Thus the critical shear stress for nanoindentation by a cylindrical indenter is determined to be 5.6 GPa. This value is much higher than that for the unrelaxed simple shear (3.4 GPa) from DFT calculation since the instability occurs within a highly compressive stress environment associated with nanoindentation. This result reinforces the concept that the critical shear stress required to nucleate a dislocation homogeneously is not a unique value, but is in fact quite dependent on the local stress state of the crystal.

The site of maximum resolved shear stress along the $[1\bar{2}\bar{1}]$ direction shown in Figure 2-10(b) is about 9.9 Å below the surface and 9.2 Å on the left of x_2 axis. It differs from the site of minimum $\det[\mathbf{Q}(\mathbf{n})]$ shown in Fig. 2-9. MD simulation of the dislocation nucleation process verifies that the Λ-criterion accurately predicts the location of dislocation nucleation, whereas the CRSS criterion is only approximate. The advantages of the Λ-criterion over the CRSS criterion could be further appreciated as follows: The Λ-criterion is parameter-free for a given lattice structure and interatomic potential. Dislocation nucleation occurs as a natural consequence of losing the positive-definite property of the acoustical tensor $\mathbf{Q}(\mathbf{n})$. Moreover, the influence of the stress state on instability is embedded in the $\mathbf{L}$ tensor. For the CRSS criterion, the threshold needs to be calibrated and a fixed value of CRSS is unable to capture accurately instability for various stress states.

2.5 3D Nanoindentation-Induced Dislocation Nucleation

In this section, we present 3D simulations of nanoindentation by a spherical indenter to quantify the critical state of dislocation nucleation in single crystal Cu. Actual experimental conditions are approached by creating a system size larger than that attainable via atomistic simulations and by maintaining a quasi-static indentation loading rate. Predictions are given as to when and where the dislocation will nucleate within the crystal, and what slip mode the nucleated dislocation will take. The critical stress state at the nucleation site is verified by the DFT calculation. The ideal shear strength probed through nanoindentation, which is defined as the CRSS at the site of the first homogeneous dislocation nucleation event, is evaluated. Finally, 3D MD simulations are performed to verify predictive simulations qualitatively.

2.5.1 Load-Displacement Response

Indentation is simulated for a frictionless spherical indenter pressed into the (111) surface of an anisotropic, elastic half space. The radius of the indenter is 500 Å, the approximate tip radius of a nominally sharp Berkovich indenter used in nanoindentation experiments. Extensive testing is undertaken to assess the effects of geometry of the simulated system, imposed far-field boundary conditions, element type, and node density by comparing with the Hertzian solutions for linear isotropic and anisotropic elastic materials. The size of the simulated system is taken to be $3000 \times 3000 \times 6000$ Å. The choice of the elongated system along the indentation direction is based on the study of FEM calculations of Vickers indentation by Giannakopoulos *et al.*[32], who showed that the boundary condition error is minimized in this manner. The computational cost can be further reduced considering the three-fold rotation symmetry of the (111) surface. In the simulation, the displacement along the bottom of the mesh is constrained to be zero, while the displacements of lateral surfaces are unconstrained. The force on the indenter will increase by about 5% if the traction-free boundary

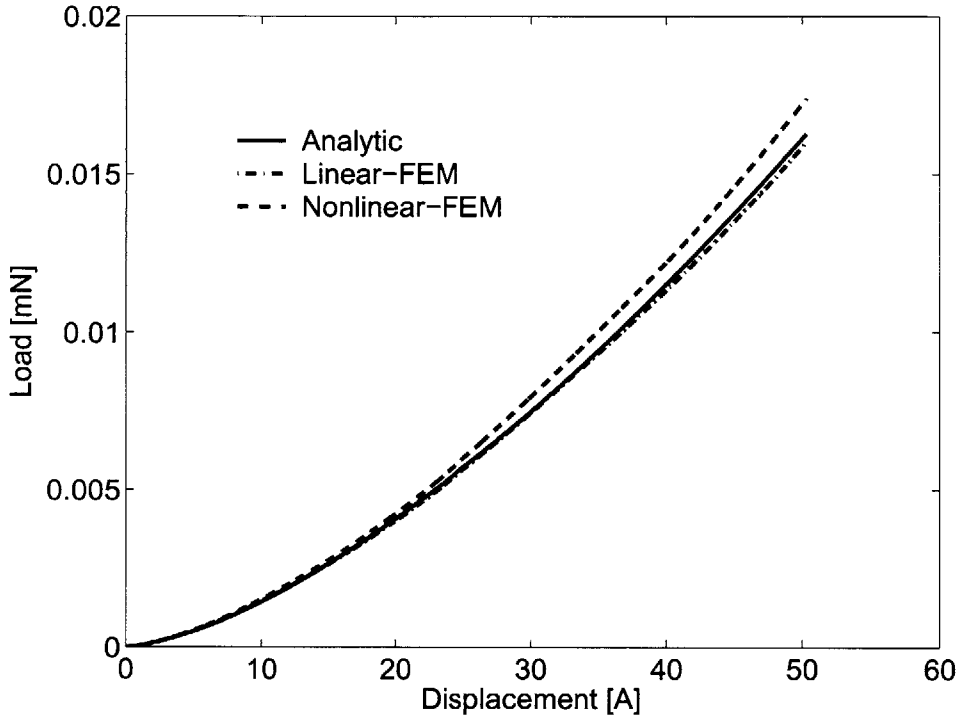


Figure 2-11: Load versus displacement curves of nanoindentation by a spherical indenter.

condition of lateral surfaces is changed to be fixed [53]. The graded mesh comprises 4-node linear brick elements. The typical size of elements near the indenter is about 10 Å, and indentation proceeds in displacement-control. Figure 2-11 compares the indentation responses predicted by the analytic solution given in section 3.2, FEM simulations based on linear anisotropic elasticity and on interatomic potential-based hyperelasticity. The FEM simulation based on linear anisotropic elasticity is in good agreement with the analytic solution, while the force on the indenter from the FEM simulation based on interatomic potential-based hyperelasticity is larger due to the non-linear elastic effect generically termed "pressure-hardening", but which has been discovered to be highly dependent on orientation [67]. The calculations are terminated at an indentation depth of 50.3 Å, when the onset of dislocation nucleation is first detected via the dislocation nucleation criterion. The corresponding indentation load is about 17.4 μN. Load-controlled nanoindentation experiments [95] on polycrystalline Cu films of different film thicknesses showed that the first burst in the P–h

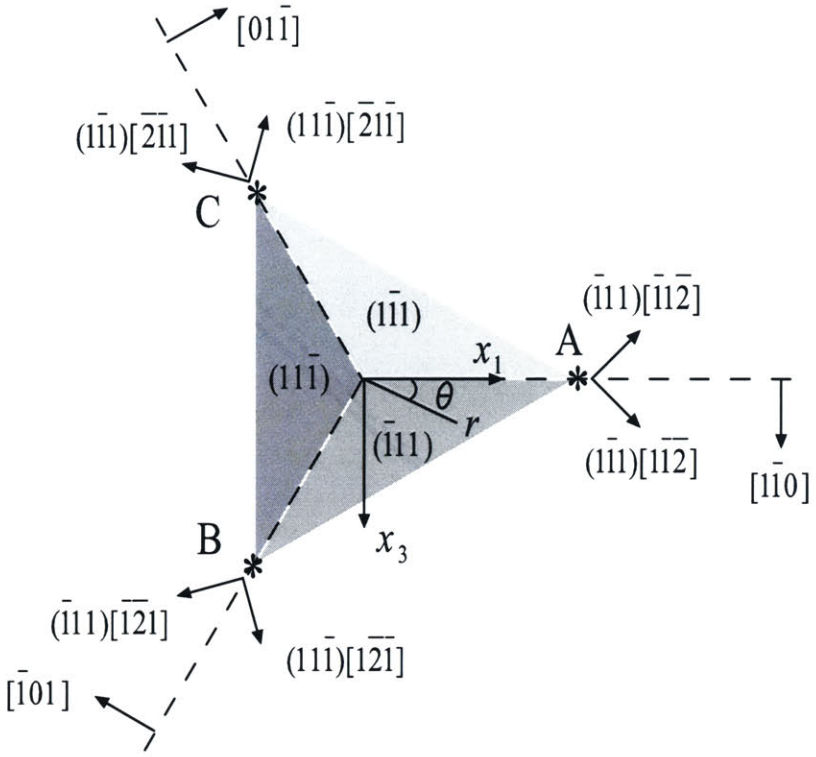


Figure 2-12: Illustration of dislocation nucleation sites and the corresponding slip systems: top view of the indented (111) surface. Some important crystallographic planes and directions are also projected into the (111) plane.

response occurs at a load of 35 ± 10 μN. This discrepancy between the predicted and experimentally determined critical load is mainly due to the idealization of rounded Berkovich indenter tip as a spherical indenter, and/or to various experimental uncertainties such as estimated tip radius, surface oxidation and strong polycrystalline texture in real films.

2.5.2 Prediction of Dislocation Nucleation

The dislocation nucleation sites are identified via the Λ-criterion. Figure 2-12 shows the top view of the indented (111) surface. The global Cartesian coordinate system is the same as that in the linear analysis. For the convenience of positioning nucleation sites, the (r, θ, z) coordinate system is introduced, where (r, θ) are the polar coordinates in the $x_1 - x_3$ plane and the coordinate z is the distance below the (111)

surface. The important crystallographic planes and directions are projected into the (111) plane as shown in Fig. 2-12. The (111) surface has a three-fold rotation symmetry about the [111] direction. The other three close-packed {111}-type planes, represented by the shaded triangles in Fig. 2-12, are positioned symmetrically below the (111) surface, and at an orientation of 120° with respect to the [111] loading axis. For 2D indentation discussed in the previous section, the $(11\bar{1})$ slip plane is favored for dislocation nucleation due to the kinematic constraints imposed by the 2D nature of the simulation. In contrast, three {111}-type slip planes beneath the (111) surface are equivalent in 3D indentation.

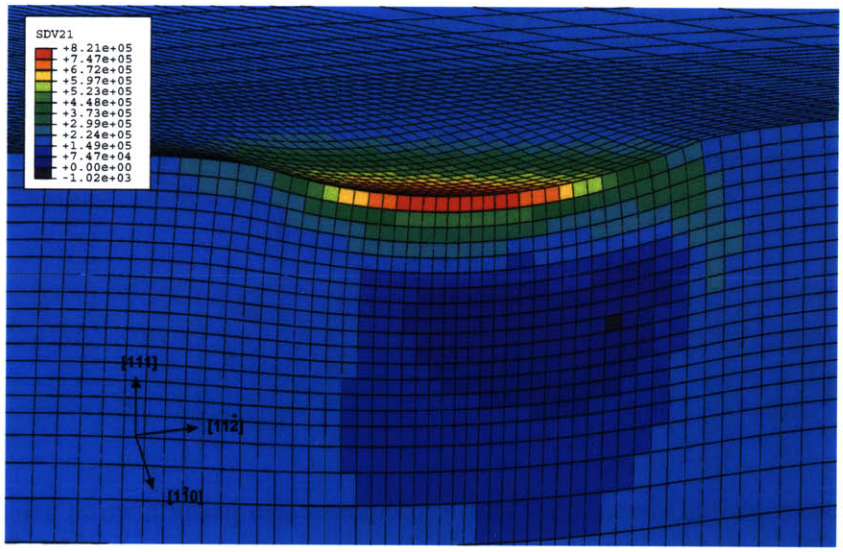


Figure 2-13: Contour of det[$\mathbf{Q}(\mathbf{n})$] by 3D FEM simulation, the cross section plane is $x_1 - x_2$ plane.

Without loss of generality, the critical sites for dislocation nucleation on the $(1\bar{1}1)$ slip plane are first identified. The corresponding spatial distribution of det[$\mathbf{Q}(\mathbf{n})$] is 3D in character, but it is symmetric with respect to the $\theta = 2\pi/3$ plane due to the symmetry of the $(1\bar{1}1)$ slip plane about the $\theta = 2\pi/3$ plane. The distribution of det[$\mathbf{Q}(\mathbf{n})$] within the $\theta = 2\pi/3$ plane presents very similar characteristics to that for the 2D case as shown in Fig. 2-9. As the detected sites deviate from the $\theta = 2\pi/3$ plane, the spatial distribution of det[$\mathbf{Q}(\mathbf{n})$] changes. Two nucleation sites are identified from the critical elements with vanishing det[$\mathbf{Q}(\mathbf{n})$]. The projected positions

of these two critical sites are schematically represented by points A and C in Fig. 2-12. The contour of $\det[\mathbf{Q}(\mathbf{n})]$ for the $(1\bar{1}1)$ slip plane is shown in Fig. 2-13. The black color element with a small negative value of $\det[\mathbf{Q}(\mathbf{n})]$ indicates the site of nucleation. One nucleation site within the $x_1 - x_2$ plane ($\theta = 0$) is visible in Fig. 2-13 and it corresponds to point A in Fig. 2-12. The in-plane coordinate of this nucleation site (in the undeformed coordinate system) is 92.7 Å below the contact surface and off the central axis by 96.5 Å. For the present indentation simulation (indenter radius $R = 500$ Å and critical penetration depth $h = 50.3$ Å), the nominal contact radius $a = \sqrt{Rh} = 158.6$ Å. Thus, the nucleation sites are $0.58a$ below the contact surface and displaced from the central axis by approximately the same distance, $0.61a$.

The slip directions of the homogeneously nucleated dislocation are predicted by the eigenvector analysis of the corresponding matrix $\mathbf{Q}(\mathbf{n})$. The deformation gradient matrix $\mathbf{F}^{\text{Crit}}$ corresponding to the critical element in Fig. 2-13 is

$$\mathbf{F}^{\text{Crit}} = \begin{bmatrix} 0.984 & -0.011 & 0.000 \\ -0.115 & 0.920 & 0.005 \\ 0.000 & -0.001 & 1.064 \end{bmatrix}. \tag{2.27}$$

The Cauchy stress, elastic constant and current slip plane normal are then calculated. The eigenvector analysis predicts that the slip vector $\mathbf{g}$ (in the global coordinate system) is $[0.630\ 0.346\ 0.695]^{\text{T}}$, while the normalized slip vector calculated from the classical relation gives $\mathbf{g} = \mathbf{F}\mathbf{g}^{\circ} = [0.675\ 0.374\ 0.636]^{\text{T}}$, where the initial slip vector $\mathbf{g}^{\circ}$ (in the crystal coordinate system) is $[1\ -1\ -2]^{\text{T}}$. Accordingly, at the nucleation site A, a Shockley partial slip along the $[1\bar{1}\bar{2}]$ direction will develop within the $(11\bar{1})$ slip plane. For the $(11\bar{1})$ slip plane at the nucleation site C, the Shockley partial slip direction is determined to be along the $[\bar{2}\bar{1}1]$ direction by taking into account the symmetry of the $(11\bar{1})$ plane with respect to the $\theta = 2\pi/3$ plane.

Similarly, the nucleation sites and slip modes for the $(\bar{1}11)$ and $(1\bar{1}1)$ slip planes are determined based on three-fold symmetry of the (111) surface. In summary, there are three symmetrically distributed dislocation nucleation sites below the (111)

Table 2.2: Positions and slip systems of homogeneously nucleated dislocations

Nucleation Site	Position (r, θ, z)	Slip systems
A	$0.61a, 0, 0.58a$	$(1\bar{1}1)[1\bar{1}\bar{2}]$ & $(\bar{1}11)[\bar{1}1\bar{2}]$
B	$0.61a, 2\pi/3, 0.58a$	$(\bar{1}11)[\bar{1}\bar{2}1]$ & $(11\bar{1})[1\bar{2}\bar{1}]$
C	$0.61a, 4\pi/3, 0.58a$	$(11\bar{1})[\bar{2}1\bar{1}]$ & $(1\bar{1}1)[\bar{2}\bar{1}1]$

surface, as designated schematically by the points (A, B, C) in Fig. 2-12. At each site, two slip planes are equally likely to be activated. Table 2.2 summarizes the positions and slip systems at each nucleation site. In previous studies [95, 35], the maximum equivalent shear stress criterion is used to identify the dislocation nucleation sites. From the Hertzian contact solution based on linear elasticity [42], the only potential dislocation nucleation site is along the central axis at a depth of $0.48a$.

2.5.3 Ideal Shear Strength of Cu

One of the potential applications of nanoindentation is the experimental characterization of the ideal shear strength of the material. The correlation of the ideal shear strength with the onset of P–h discontinuity has been made by stress analysis from the Hertzian contact solution based on linear isotropic elasticity, see Gerberich *et al.* [31] for Si, Kiely *et al.* [47] for Au, Suresh *et al.* [95] and Gouldstone *et al.* [34] for Al and Cu. A recent study by Krenn *et al.* [50] gave a better connection between the atomistic and experimental estimates of ideal shear strength of W and Mo by taking into account non-linear elastic effects. The detailed *ab initio* DFT calculations for affine shear deformation of Cu and Al by Ogata *et al.* [67] showed that the ideal shear strength strongly depends on the triaxial stress state. In this subsection, the ideal shear strength of Cu is obtained from the critical shear stress resolved in the Shockley partial slip direction at the dislocation nucleation site. Though the value of ideal shear strength is not uniquely defined and depends on the deformation history and state at the point of instability, the ideal strength probed through nanoindentation provides a means to quantify this mechanical parameter experimentally.

Substitution of the deformation gradient matrix Eq. (2.27) into Eq. (2.7) give the Cauchy stress in the global coordinate system

$$\boldsymbol{\sigma}^{\text{Mishin}} = \begin{bmatrix} -0.79 & -5.86 & 0.22 \\ -5.86 & -19.38 & 0.49 \\ 0.22 & 0.49 & -4.61 \end{bmatrix} \text{GPa.} \tag{2.28}$$

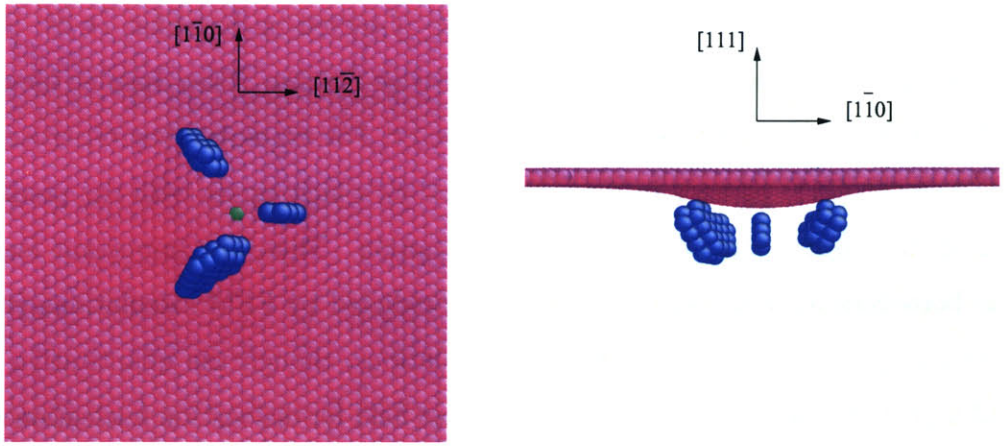
Using the same deformation gradient in Eq. (2.27), the *ab initio* DFT calculation following the scheme by Ogata *et al.* [67] gives the Cauchy stress

$$\boldsymbol{\sigma}^{\text{DFT}} = \begin{bmatrix} -0.59 & -5.90 & 0.21 \\ -5.90 & -19.24 & 0.48 \\ 0.21 & 0.48 & -5.41 \end{bmatrix} \text{GPa.} \tag{2.29}$$

The agreement between the above two calculations demonstrates the accuracy of quantitative characterization of the critical stress state from FEM simulations via the Mishin potential. The ideal strength, *e.g.*, resolved shear stress in the $(1\bar{1}1)[1\bar{1}\bar{2}]$ slip system, is then calculated to be 4.56 GPa, and the corresponding triaxial stress is -8.26 GPa. Compared to the fully relaxed pure shear deformation calculation (the ideal shear strength of 2.16 GPa and triaxial stress of zero) by Ogata *et al.*[67], the ideal shear strength probed by nanoindentation is greater by a factor of two, due to the large triaxial stress at the critical site of dislocation nucleation.

2.5.4 MD Simulation

Direct MD simulations conducted for a smaller system at a comparatively high indentation loading rate (about 6 m/s) qualitatively verify the prediction of dislocation nucleation via the Λ-criterion. A spherical indenter with a radius of 50 Å is pressed into the (111) surface of a Cu cube with an approximate side length of 100 Å. Figures 2-14(a) and (b) show the bottom view (along the [111] direction) and the side



(a) (b)

Figure 2-14: MD simulation of nanoindentation by a spherical indenter: atomic structure of homogeneously nucleated dislocations beneath the (111) surface: (a) bottom view along the [111] direction; (b) side view along the [11$\bar{2}$] direction. The indenter first contacts the green atom.

view (along the [11$\bar{2}$] direction) of three dislocation embryos nucleated on the inclined {111}-type slip planes beneath the surface, respectively. The perfectly coordinated atoms within the bulk (N = 12) have been removed from the images for clarity, such that only surface and imperfectly coordinated atoms are visible. The green atom in Fig. 2-14(a) represents the first contact point between the indenter and the (111) surface. In contrast to the formation of a straight dislocation line in 2D simulation, the dislocation embryo nucleates as a group of atoms, and will expand into a partial dislocation loop. Embryo size is indicative of the sequence of embryo nucleation, with larger embryos nucleating at earlier time increments. In the MD simulation, it is observed that the largest embryo shown in Figs. 2-14(a) and (b) first nucleates along the (1$\bar{1}$1) plane. Then, the medium-sized embryo nucleates along the ($\bar{1}$11) plane almost simultaneously. The site and slip mode of this second embryo is symmetric to that of the first embryo, with respect to $x_1 - x_2$ plane. In the time increment shown in Figs. 2-14(a) and (b), the smallest embryo has just nucleated. It will expand within the same (1$\bar{1}$1) slip plane as the first embryo and coalesce. These three, symmetrically distributed nucleation sites observed in MD simulations agree with predictions from FEM simulations via the Λ-criterion. The same distribution of nucleation sites

and structure of dislocation embryos were also observed by Kelchner *et al.* [45] in MD simulations of indentation on the (111) surface of Au. They further identified the slip vector via the centrosymmetry parameter. For example, they observed that the slip on the $(\bar{1}11)$ plane is in the $[\bar{1}\bar{2}1]$ direction, which agrees with the prediction based on the eigenvector analysis of the Λ-criterion. Thus, the important features of the homogeneously nucleated dislocations predicted by FEM calculations via the Λ-criterion qualitatively agree with MD simulations, though the quantitative state including the critical indentation load and stress field for dislocation nucleation differ due to discrepancies in loading rate and system size.

In MD simulations for Cu given above and for Au by Kelchner *et al.*[45], it is observed that at the very early stage of dislocation nucleation, the glide process always occurs along two of the three inclined $\{111\}$-type slip planes. A three fold symmetric defect structure is never obtained. The loss of the (111) surface three-fold symmetry can be explained by examining the distribution and slip characters of nucleated dislocations shown in Figs. 2-14(a) and (b) with reference to the predictions given in Table 2.2 and Fig. 2-12. Though there exist three equivalent dislocation nucleation sites and two possible slip planes at each site, the first nucleated two dislocations along the $(1\bar{1}1)$ and $(\bar{1}11)$ slip planes occupy each of two possible nucleation sites for the $(11\bar{1})$ slip plane, respectively. Moreover, the two nucleated dislocations with symmetric sites and slip modes could accommodate geometrically the deformation imposed by the indenter penetration. Accordingly, the subsequent activation of the glide process along the $(11\bar{1})$ slip plane is likely to be suppressed, as observed in MD simulations.

2.6 Discussion and Summary

Nanoindentation-induced homogeneous dislocation nucleation in single crystal copper has been analyzed within the framework of hyperelasticity with the Cauchy-Born hypothesis. Homogeneous dislocation nucleation is interpreted as a strain localization event triggered by elastic instability in a perfect crystal at finite strain. Since

the constitutive model of interatomic potential-based hyperelasticity incorporates the key crystal properties such as crystalline anisotropy and nonlinear elastic effects, it is well-suited to the study of defect nucleation at the continuum level while maintaining fidelity with the underlying atomistic response. The efficiency of the finite element formulation permits the simulation of significantly larger systems than would otherwise be possible. However, care must be taken in applying this approach for predictive modeling of defect nucleation because the Cauchy-Born hypothesis is only valid when the spatial variation of the continuum deformation field is gradual on the atomic scale. For problems such as dislocation nucleation at an atomistically sharp crack tip [80], the large strain gradient near the crack tip necessitates the use of the quasi-continuum or direct atomistic methods which can account explicitly for non-local effects. However, in the present study of a relatively slowly varying deformation field imposed by a 50 nm-scale indenter radius, interatomic potential-based hyperelasticity is efficiently and effectively utilized.

As the present simulations idealize indentation normal to an atomically flat surface, it follows that the dislocation nucleates homogeneously within the crystal and moves/grows to approach the free surface. This atomistic process is in contrast with nanoindentation-induced heterogeneous dislocation nucleation from surface steps [47, 125], a process subject to a much lower energy barrier that may significantly decrease the simulated or experimentally measured critical load. Furthermore, we restricted the present dislocation nucleation study to the idealized limit of behavior at 0K, and in the absence of specific effects of thermally activated processes. As a result, the ideal shear strength obtained via 3D indentation simulation represents the mechanical threshold for nanoindentation-induced dislocation nucleation. The effect of thermal motion on homogeneous nucleation of dislocation loops under a simple shear stress state has been studied by Xu and Argon [115] using a variational boundary integral method to determine the saddle point configuration and the corresponding activation energy. They found that for perfect crystals such as Au, Cu, Al and Si, the energy barriers are far too high for thermal motion to play a significant role in dislocation nucleation, even under applied shear stress levels equal to half of the ideal

simple shear strength. The activation energy for homogeneous dislocation nucleation beneath the nanoindenter could be evaluated using the same scheme as Xu and Argon [115]. FEM simulations of nanoindentation using interatomic potential-based hyperelasticity will provide a more realistic stress state beneath the nanoindenter, rather than using a simple shear stress state as an input for activation energy calculation.

We conclude by noting that the study of defect nucleation within the framework of hyperelasticity is not limited to simple FCC lattices and bulk homogeneous material systems. Further extension of interatomic-potential-based hyperelasticity has been made to other type of lattices, *e.g.*, Tadmor *et al.* [99] and Smith *et al.* [89] for nonsymmorphic crystal lattices such as diamond cubic Si, Arroyo and Belytschko [8] for a one-atom thick carbon nanotubes, *etc.* Furthermore, the instability analysis we employed to predict defect nucleation within the bulk can also be generalized to study surface and interface instabilities, as well as homogeneous defect nucleation at coherent grain boundaries with the aid of a generalized instability criterion at interfaces (e.g., Needleman and Ortiz [66]). Such extensions will further demonstrate the applicability of defect nucleation analyses within the framework of hyperelasticity, and will enable the study of a wide range of phenomena at the continuum level, while maintaining explicitly the atomistic interactions which govern the mechanical response.

Chapter 3

Dislocation Emission and Cleavage Decohesion at Crack Tips

3.1 Dislocation Loop Emission from a Crack Tip in Cu

3.1.1 Introduction

A central issue in understanding the ductile versus brittle behavior of solids is the local response of an atomically sharp crack at critical loading [46, 82, 25, 80, 120, 13]. While it is widely recognized that cleavage decohesion and dislocation emission are the two major competing modes of response, atomistically accurate analysis of dislocation loop emission in the presence of a crack-tip stress field has not been carried out. In this section, we perform reaction pathway analysis using a many-body interatomic potential to determine the atomic configurations that make up the activated displacement fields (shear and tensile) surrounding a crack tip under prescribed load, and the associated energetics. In this way we obtain, in full atomistic details, a description of the emission of an embryonic dislocation loop from the crack tip. Comparing our findings with existing continuum-level treatments, we show that inclusion of atomic-level details gives rise to quantitative corrections which are physically understandable.

The implication of our results is that homogeneous nucleation [117] of a dislocation loop at an atomically smooth crack tip in a ductile crystal like Cu is unlikely to be a dominant process. It also follows that local structural heterogeneities, such as a crack-tip ledge, are likely to govern the brittle-ductile response of a solid, for which an atomic-level description along the lines presented here is now feasible.

From a continuum perspective, crack-tip dislocation emission corresponds to the birth and shedding of a shear-type singularity (the dislocation) from the primary, tensile-type singularity (the crack tip). Atomistically, the embryonic dislocation emerges from the stressed crack tip as a distribution of shear displacements between atoms across the slip plane [5]. As the applied strain energy release rate G increases to the critical value for spontaneous dislocation emission, defined here as the athermal G_{emit}, the incipient shear emanating from the crack tip loses its stability, which leads to a fully formed straight dislocation that moves away from the crack. For loads less than G_{emit}, dislocation emission can occur by overcoming activation energy barriers via thermal fluctuations. This is a very localized process involving the unstable emission of a 3D dislocation loop from a crack tip [82]. Based on the Peierls concept, the saddle-point configurations and the associated activation energies have been calculated by recourse to the interplanar potential (γ-surface) and approximate numerical schemes of finding unstable transition states [84, 81, 116, 117]. Although these continuum models provide great physical insights into the nucleation events, a complete understanding of the thermally activated process of dislocation nucleation necessitates a fully atomistic study of the emission of a 3D dislocation loop from a crack tip.

3.1.2 Geometry and Method

Consider a semi-infinite crack in an otherwise perfect fcc single crystal Cu, schematically shown in Fig. 3-1. The straight crack front, lying on a (111) plane, runs along the $[\bar{1}10]$ direction. The activated slip plane under mode I loading is $(\bar{1}\bar{1}1)$, inclined at $\theta = 70.53°$ with respect to the (111) crack plane. Our simulation cell consists of a cracked cylinder cut from the crack tip, with radius $R = 80$Å. The atoms within

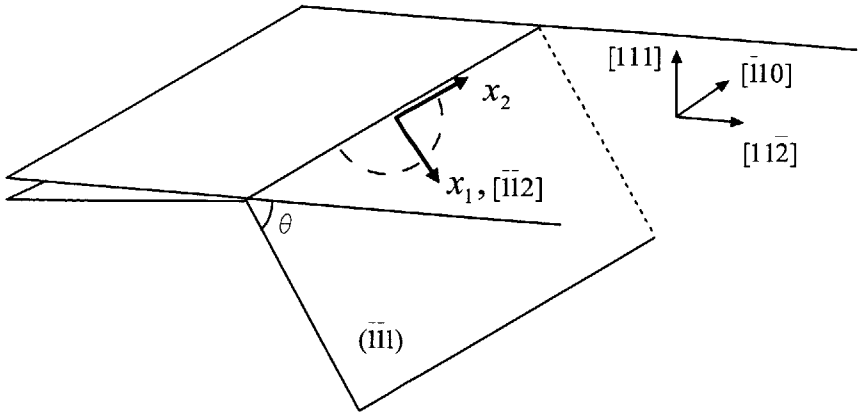


Figure 3-1: Orientations of the crack and the inclined $(\bar{1}\bar{1}1)$ slip plane across which a dislocation loop nucleates.

5Å of the outer surface are fixed according to a prescribed boundary condition while all remaining atoms are free to move. To capture the 3D nature of a dislocation loop emitted from the crack tip, the simulation cell length along the cylinder is taken to be suitably long, 61Å, with periodic boundary condition imposed. The effects of finite size on the simulation results will be discussed later. One layer of atoms is removed to create a crack. The total number of atoms within the system is 103,920. The interatomic interactions are modeled using the embedded atom method (EAM) potential of Cu [65]. The stacking fault energy given by this potential is 44.4mJ/m^2, the experimental value being 45mJ/m^2. The unstable stacking energy γ_{us} given by the potential was fitted to the *ab initio* calculation, $\gamma_{\text{us}} = 158\text{mJ/m}^2$. To demonstrate the accuracy of our atomistic calculations, we compare the stress distribution from atomistic calculation shown in Fig. 3-2(a) with the analytic solution based on the Stroh formalism [93], shown in Fig. 3-2(b) at the same load level $K_{\text{I}} = 0.44\text{MPa}\sqrt{\text{m}}$. Atoms in Fig. 3-2(a) are color-coded according to the atomic stress. It is seen that the atomistic calculation is in agreement with the singular, analytic solution up to a few atomic spacing away from the crack tip. To reveal the overall agreement between two calculations, the upper limit of the contour bar in Fig. 3-2 is set to 5GPa.

Prior to identifying the pathways of thermally activated dislocation nucleation, we first determine the athermal load G_{emit}, the value at which the activation energy

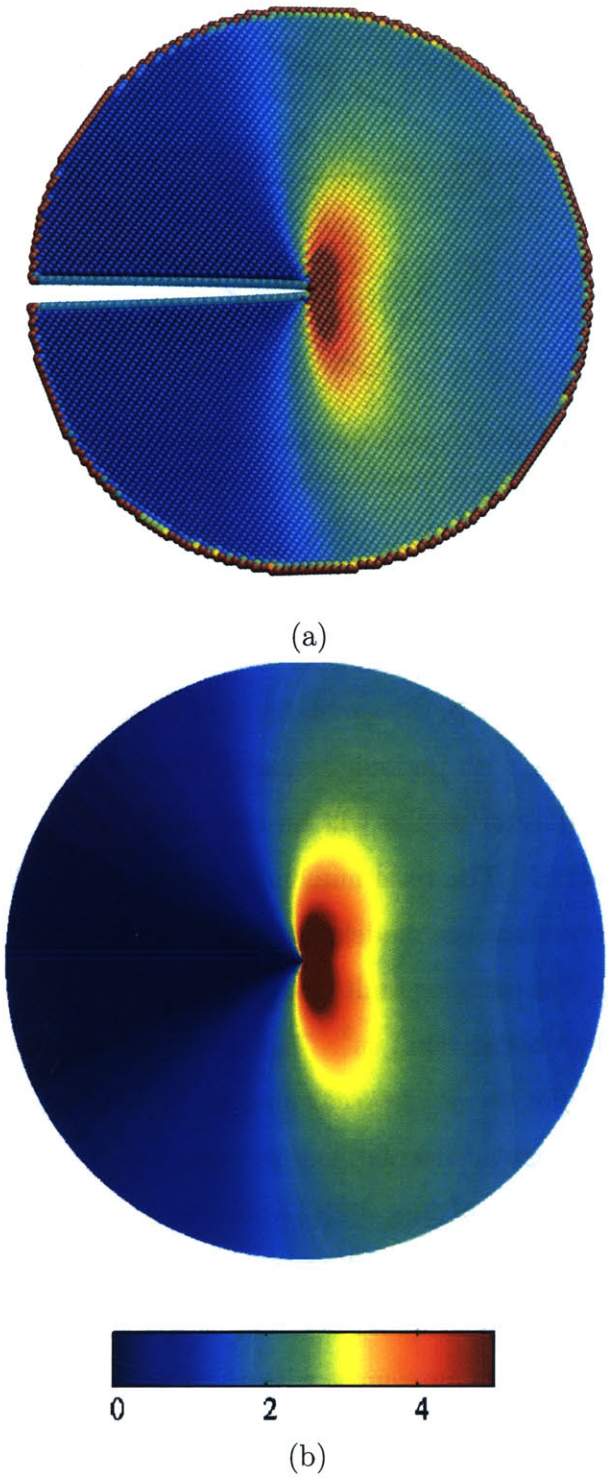


Figure 3-2: Distribution of σ under mode I load $K_{\text{Iemit}} = 0.44\text{MPa}\sqrt{\text{m}}$: (a) atomistic calculation, (b) analytic solution. The stress on the contour bar is in unit of GPa.

barrier for dislocation nucleation vanishes leading to an instantaneous emission of a straight dislocation without thermal fluctuations. We apply incrementally a mode I load G, or the equivalent stress intensity factor K_{I}. At each step, the initial positions of atoms are set according to the anisotropic linear elastic Stroh solution [93]. Then the system is relaxed using conjugate gradient method while the outer boundary is held fixed. As in previous studies [20], at low loads we observed the emergence of an embryonic dislocation in the form of a distribution of shear displacements of atoms across the $(\bar{1}\bar{1}1)$ slip plane in front of the crack tip. This incipient dislocation with a straight core along the crack front is only partially formed in that the maximum shear displacement between two crack-tip atoms across the adjacent $(\bar{1}\bar{1}1)$ slip planes is less than one half the Burgers vector $\mathbf{b} = a_0/6[\bar{1}\bar{1}2]$ of a fully formed Shockley partial dislocation, where $a_0 = 3.615$Å is the lattice constant of Cu. As the applied load increases to $G_{\mathrm{emit}} = 1.629$ J/m^2 (or the equivalent $K_{\mathrm{Iemit}} = 0.508\mathrm{MPa}\sqrt{\mathrm{m}}$ based on the Stroh solution [93]), the metastable state (local energy-minimum) corresponding to the atomic configuration with a partially formed dislocation disappears from the energy landscape (spinodal instability). At this point a straight $a_0/6[\bar{1}\bar{1}2](\bar{1}\bar{1}1)$ Shockley partial dislocation is emitted from the crack tip. The leading edge of the newly formed dislocation moves away from the crack and stops against the fixed outer boundary at about 40Å from the crack tip, laying down a stacking fault in its wake. As both the crack front and outer boundary of the simulation cell are translationally invariant with respect to the x_2 direction, the final energy-minimized state also possesses a straight dislocation parallel to x_2 axis. We note that the athermal G_{emit} from the present atomistic calculation is higher than the value given by analytic criterion $G_{\mathrm{emit}} = 8\gamma_{\mathrm{us}}/[(1+\cos\theta)\sin^2\theta] = 1.067$ J/m^2 [80], as the latter was derived without considering the resistance to crack-front surface production in association with dislocation emission [119, 116, 83].

For a given load below G_{emit}, we use the nudged elastic band (NEB) method [43] to determine the minimum-energy path (MEP) [43] of a partial dislocation loop bowing out from the initially straight crack front. In configurational space, the state with the highest-energy along the MEP is the saddle point on the edge of the energy basin

enclosing the initial state. For each NEB calculation at a given G, the initial state is taken as the relaxed configuration initially set according to the linear Stroh solution as introduced above. It is relative to this state that an embryonic dislocation with a straight core emerges and remains stable near the crack tip. In contrast, the final state contains a fully formed Shockley partial dislocation. This local energy-minimum state is obtained by unloading the simulation cell embedded with a pre-existing Shockley partial dislocation which is generated by loading the system above G_{emit}. Then a discretized path consisting of 15 replicas of the system is constructed to connect the initial and final states. The calculation is considered converged when the potential force on each replica vertical to the path is less than 0.005eV/Å. Based on the result of NEB relaxation, we use the dimer method [36] to refine the calculations of both the saddle-point configuration and the associated activation energy ΔE_{act}. The latter is defined as the energy difference between the saddle point and initial state. The replica with the highest energy from a converged NEB relaxation is taken as the initial input to the dimer method. For the final relaxed dimer, we find the local curvature of energy surface along the dimer direction to be negative and the potential forces on two images in the direction of the dimer are 0.003eV/Å and -0.003eV/Å, respectively, the opposite sign indicating the two images are sitting on two sides of the saddle point. The rotational force, which is the net force acting on one of the images vertical to the direction of the dimer, is less than 0.001eV/Å.

3.1.3 Results and Discussions

The MEP of a dislocation loop bowing out at a typical load $G = 0.75G_{\text{emit}}$ is shown in Fig. 3-3. With the initial-state energy as a reference, the energy variation ΔE along the MEP can be seen. The normalized reaction coordinate is defined as the ratio between l, the hyperspace arc length along the MEP from the initial to the current state, and l_0, the total arc length along the MEP. The continuous energy curve is obtained by cubic-polynomial interpolation of the calculated energies of replicas, indicated by circles, with the aid of the potential force projected in the direction of the path on each replica [37]. The activation energy ΔE_{act} from the refined calculation

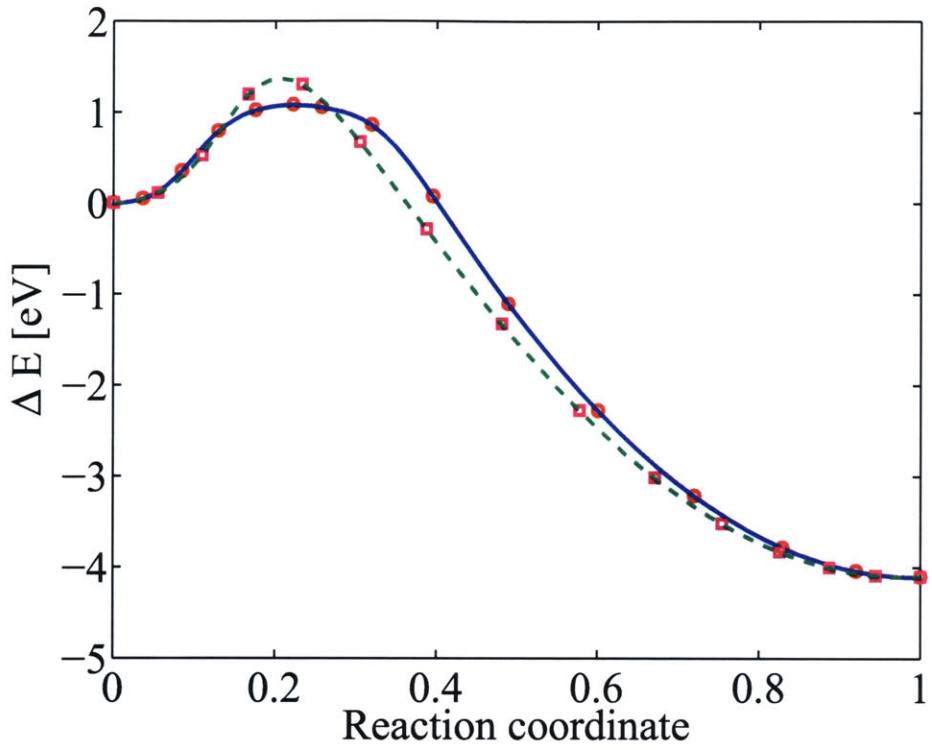
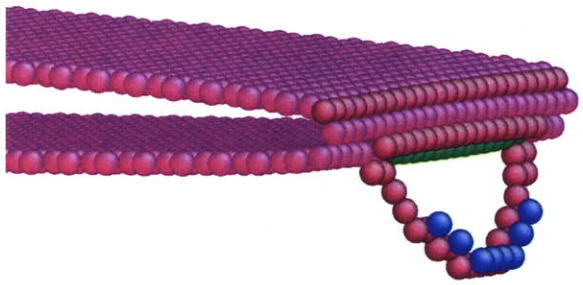
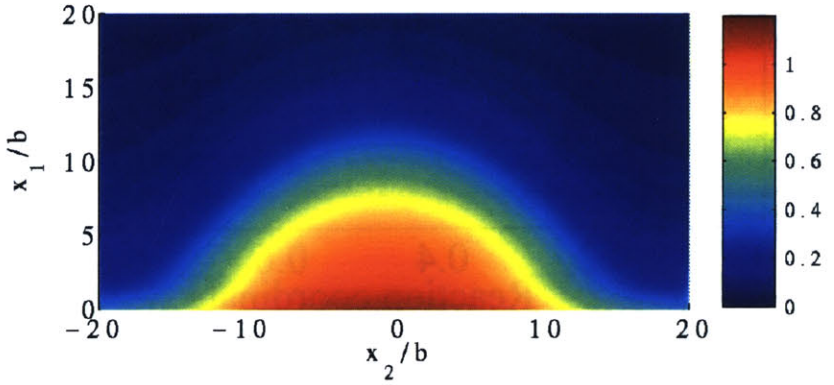


Figure 3-3: The minimum energy paths for nucleating a dislocation loop (solid line) and a straight dislocation (dashed line).

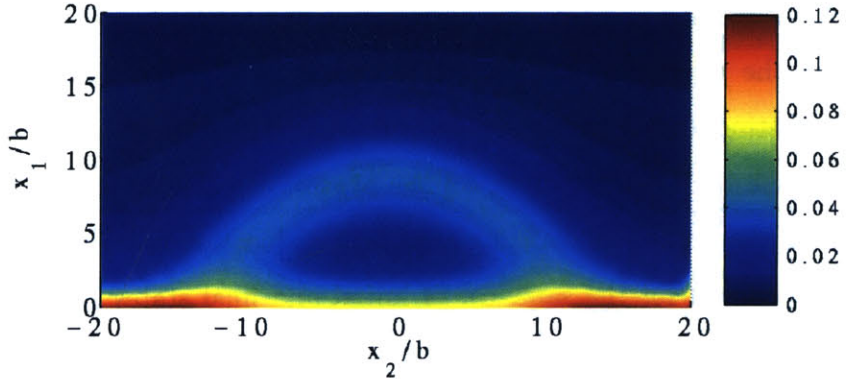
using the dimer method is 1.1eV. For comparison the MEP of nucleating a straight dislocation in the present simulation cell is also shown. Note that the activation energy of nucleating a straight dislocation is known to diverge with increasing length of crack front [84, 81, 116]. However, for the present calculation with a finite size of simulation cell, the process of emitting a straight dislocation represents a competing nucleation mechanism for the same set of initial and final states. As the length of simulation cell along the x_2 direction has already been taken to be sufficiently long, it is seen that activation energy barrier ΔE_{act} for nucleating a 3D dislocation loop is lower than that for a straight dislocation. The difference in ΔE_{act} from the refined calculations using the dimer method is about 0.31eV.



(a)



(b)



(c)

Figure 3-4: The saddle-point configuration at $G = 0.75G_{\text{emit}}$ under mode-I load. (a) Atomic structure of dislocation loop. Atom color indicates coordination number N: light pink, $N = 9$, green, $N = 10$, dark pink, $N = 11$, blue, $N = 13$, atoms with perfect coordination $N = 12$ are made invisible; (b) Contour plot of shear displacement distribution across the slip plane; (c) Contour plot of opening displacement distribution.

Along the MEP of nucleating a 3D dislocation loop, dislocation emission is seen as a localized outward protrusion from the straight crack front. This incipient bulge expands subsequently by spreading out in both the forward and lateral directions. The atomic structure of the saddle-point configuration obtained from the dimer method is shown in Fig. 3-4(a), where all the perfectly coordinated atoms are removed for clarity, and the remaining atoms are those in the core of the protruding dislocation loop, along with those on the crack surfaces. Figure 3-4(b) shows the distribution of shear displacement in the x_1 direction across two adjacent $(\bar{1}\bar{1}1)$ slip planes containing the dislocation loop shown in Fig. 3-4(a). The continuous shear distribution is obtained by cubic-spline interpolation of the calculated shear displacements, normalized by $b = a_0/\sqrt{6} = 1.476$Å, at discrete lattice sites. It is evident from Fig. 3-4(b) that the embryonic dislocation develops in the form a non-uniform distribution of shear across the slip plane. The maximum shear occurs right at the crack front with a value of $1.13b$, while minimum shear along the crack front is $0.41b$. The contour line of $b/2$ shear displacement (in green color) represents approximately the locus of dislocation core as shown in Fig. 3-4(a). Within the region enclosed by the dislocation loop, slips between pairs of atoms across the slip plane are around one b, indicating this part of the crack tip has been swept by a fully formed dislocation. In contrast, for pairs of atoms outside the loop, the slips are less than $b/2$. Additionally, Figure 3-4(a) shows that the coordination numbers of two rows of crack-front atoms within the end points of the loop change from values of 9 and 12 at the initial state to 10 and 11 at the transition state respectively, signifying crack-front surface production [119, 116, 83].

The normal stress across the slip plane is also of interest since it has the effect of reducing the resistance for dislocation nucleation [116]. Figure 3-4(c) shows the corresponding distribution of opening displacement between the same pair of adjacent $(\bar{1}\bar{1}1)$ planes, normalized by the interplanar spacing $h = a_0/\sqrt{3} = 2.087$Å. The maximum opening also occurs at the crack front with a value of $0.12h$. A ring of relatively large opening displacements (in light-blue color) develops ahead of the crack, corresponding to the loop shown in Figs. 3-4(a) and (b). This is the locus of the dislocation core where large shear-induced dilations exist across the slip plane.

Further comparison of the present atomistic results relative to continuum-level findings [81, 116] shows quantitative differences in the contour of the saddle point configuration and the magnitude of the activation energy. One should keep in mind, however, that besides differences in the methods used and the way constitutive behavior are specified, the previous studies were concerned with mode II loading where the activated slip is along the crack plane and the effects of shear-tension coupling and surface-production resistance are known to be weak. Figure 3-4(b) indicates that, at the saddle-point state, the dislocation has extended in the forward direction by about $10b$ and spreaded out along the crack front by about $30b$. Previous results on the saddle-point loop configuration [81, 116, 117] showed forward and lateral extensions of less than $5b$ and $20b$, respectively. For the activation energy ΔE_{act}, the present result of 1.1eV is significantly larger than a first continuum estimate, based on a perturbative approach, of 0.18eV [81] and a second, improved estimate of 0.41eV using a more flexible representation of the incipient dislocation loop [116]. It is physically reasonable that the atomistic treatment should give a higher value than a continuum description since the former is presumably capable of capturing more fully the effects of surface production on the atomic level.

The activation energy allows one to estimate the rate of thermally activated dislocation nucleation. Nucleation from a straight crack front, commonly referred to as homogeneous nucleation [117], should be distinguished from nucleation from discrete heterogeneities along the crack front, such as a ledge, where local crack-front reorientation facilitates nucleation of a screw dislocation, involving no significant surface production and therefore lower thermal activation barrier[121, 117]. For the present problem of homogenous nucleation, the frequency of nucleation events per unit distance along the crack front can be estimated from $v = n(c_{\text{shear}}/b)\exp\left(-\Delta E_{\text{act}}/k_{\text{B}}T\right)$ [81], where $c_{\text{shear}} = 3\text{km/s}$ is the shear wave speed, and $n = 1/30b$ the number of nucleation sites per unit length of crack front. Here, $30b$ is taken in light of the range that the loop spreads out along the crack front laterally. Then one finds $v \approx 1.0/(\text{s}\cdot\text{mm})$ at room temperature. Taking $v \approx 10^6/(\text{s}\cdot\text{mm})$ to be the threshold for thermally activated nucleation in metal in laboratory measurements [81], we conclude that ex-

perimental observation of homogeneous nucleation is unlikely. On the other hand, the presence of any heterogeneity along the crack front can significantly reduce the activation energy for a local process [117, 121]. For any specified heterogeneity, the atomistic method presented in this work should be applicable.

We conclude by commenting on the size effect of our simulation cell. It has been shown that attraction between neighboring loops, due to periodicity along the crack front, leads to an underestimate of ΔE_{act}, and a period of $32b$ is sufficient to obtain an accurate value for the activation energy of an isolated loop [116]. Our choice of simulation cell with a period of about $40b$ is based on this result. Regarding the effect of in-plane radius R of the cell, we have seen in Fig. 3-4 that the dislocation loop at the saddle-point configuration is localized near the crack tip. Hence, the effect of cell boundary on the transition state is expected to be much weaker than that on the final state, where the equilibrium location of a fully formed dislocation sensitively depends on the size of the simulation cell. We have studied finite-size effects in the nucleation of straight dislocation which is computationally less demanding. The 2D nature of the problem requires only a thickness of about 10Å along the crack front following the minimum-image convention. Comparing the values of ΔE_{act} for $R = 80$Å and 120Å, we find the difference is less than 6%. We therefore believe that size effects should not affect the accuracy of the saddle-point configuration and the activation energy reported here.

3.2 3D Lattice Trapping Barriers to Brittle Fracture of Si

3.2.1 Introduction

Ample experimental evidence (e.g., Wiederhorn [112]; Lawn [55]) has indicated that cracks in many brittle solids are able to undergo quasi-static extension under sustained, constant loading, at least for a suitably restricted range of loads. One of the key microscopic steps controlling this time-dependent crack growth is the stress-

mediated, thermally activated bond breakings at the crack tip, either with or without the aid of reactive foreign species from the environment. The microscopic energy barriers that limit the kinetic rate of crack tip bond breaking are the so-called "lattice trapping" barriers (Thomson *et al.* [100]). Specifically, the total energy of a cracked body under stress is not a smooth function of crack length microscopically. Due to the periodicity of the discrete lattice, the energy landscape fluctuates at the atomic scale. As a result, the crack tip may be trapped within a well of local energy minimum on the energy surface, as a dislocation is trapped by the Peierls barrier. The mobility of the crack is then controlled by the stress-mediated activation energies, which are the barrier heights of the saddle points on the separatrices of the present energy well. In addition, it was further recognized that quasi-static crack extension is inherently a 3D process involving a series of localized bond breaking events at the crack front (Lawn [54]; Thomson *et al.* [101]; Marder [60]). Hence, to provide a realistic estimate on the kinetic rate of crack growth, it is essential to characterize the lattice trapping effect quantitatively in a 3D setting.

Despite recent progress in the experimental techniques of probing the atomic structures near crack tips (e.g., Celarie *et al.* [17]; Wiederhorn *et al.* [113]), computer simulation remains the most effective way of studying the atomic process of crack tip bond breaking (e.g., Sinclair and Lawn [87]; Sinclair [88]; Spence *et al.* [91]; Perez and Gumbsch [73, 74]; Bernstein and Hess, [13]). As a continuing effort along this line, we report in this paper the first atomistic calculation of 3D lattice trapping barriers to brittle fracture of Si in vacuum. The study is carried out by systematically probing the potential energy surface of a cracked system calculated using the Stillinger-Weber (SW) interatomic potential (Stillinger and Weber [92]). We quantitatively determine the transition pathways of localized bond breaking at the stressed crack front using an effective configurational space exploration scheme, the nudged elastic band method (Jonsson *et al.* [43]). The calculated energy variations along transition pathways reveal the energetics of crack front kink-pair formation and migration. It follows that the origin of the directional dependence in cleavage crack propagation can be explained in terms of the difference in the energetics of kink-pair formation for two

different crack orientations. Moreover, it is conceivable that, with the geometries and energetics of crack front kink pairs presented in this paper as input, a coarse-grained study of the time evolution of crack front morphology becomes feasible along the lines of Lawn [54] and Cai *et al.* [16, 15].

3.2.2 Model and method

Geometry

The system studied is a (111) plane cleavage crack in Si with a straight crack front along the $[1\bar{1}0]$ and $[11\bar{2}]$ directions, respectively. For the convenience of later comparison, the first crack orientation is denoted as the $(111)[1\bar{1}0]$ crack, while the second is the $(111)[11\bar{2}]$ crack. To contrast the structural differences, figures 3-5(a) and (b) show the relaxed atomic configurations near the crack tip for two crack orientations at respective Griffith critical loads, to be determined in section 3.2.3 and section 3.2.4. For both orientations, the atomistic simulation cell consists of a cylinder cut from the crack tip (Zhu *et al.* [123]), with the cylinder radius $R = 80$Å. Atoms within 5Å of the outer surface are fixed according to a prescribed boundary condition based on the displacement field of anisotropic linear elastic Stroh solution (Stroh [93]; Suo [94]), and all remaining atoms are free to move. To reveal the 3D nature of crack front extension, the simulation cell length along the cylinder is taken to be suitably long, 20 unit cells, with a periodic boundary condition (PBC) imposed. Since the 12-atom unit cell we take for Si is orthorhombic, the length of the simulation cell along the cylinder direction, l, as well as the total number of atoms within the system, N, are different for the two crack orientations. For the $(111)[\bar{1}10]$ crack, $l = 76.8$Å and $N = 77,200$; while for the $(111)[11\bar{2}]$ crack, $l = 133.0$Å and $N = 133,760$. For both orientations, the coordinate system is taken such that the x_3' axis is along the crack front direction and the x_2' axis is $\langle 111 \rangle$. Thus, cracks extend in the direction along the x_1' axis.

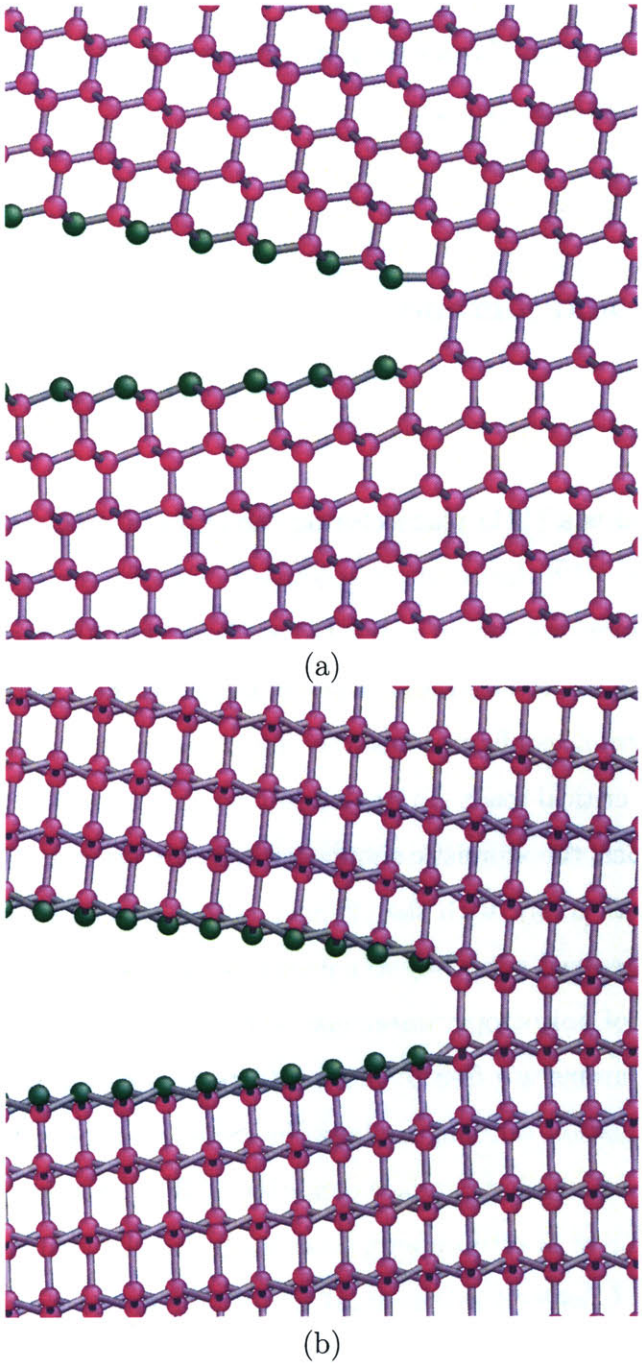


Figure 3-5: Atomic configurations near the crack tip: (a) the $(111)[1\bar{1}0]$ crack; (b) the $(111)[11\bar{2}]$ crack

Table 3.1: Lattice parameter (in Å), elastic constants (in GPa) and relaxed {111} shuffle plane surface energy (in J/m^2) of Si

Potential	a_0	C_{11}	C_{12}	C_{44}	γ_s
Expt.	5.429[a]	167[a]	65[a]	81[a]	1.24[b]
SW	5.431	162	82	60	1.45
EDIP	5.430	175	62	71	1.05

[a]: Balamane *et al.* [10]; [b]: Spence *et al.* [91].

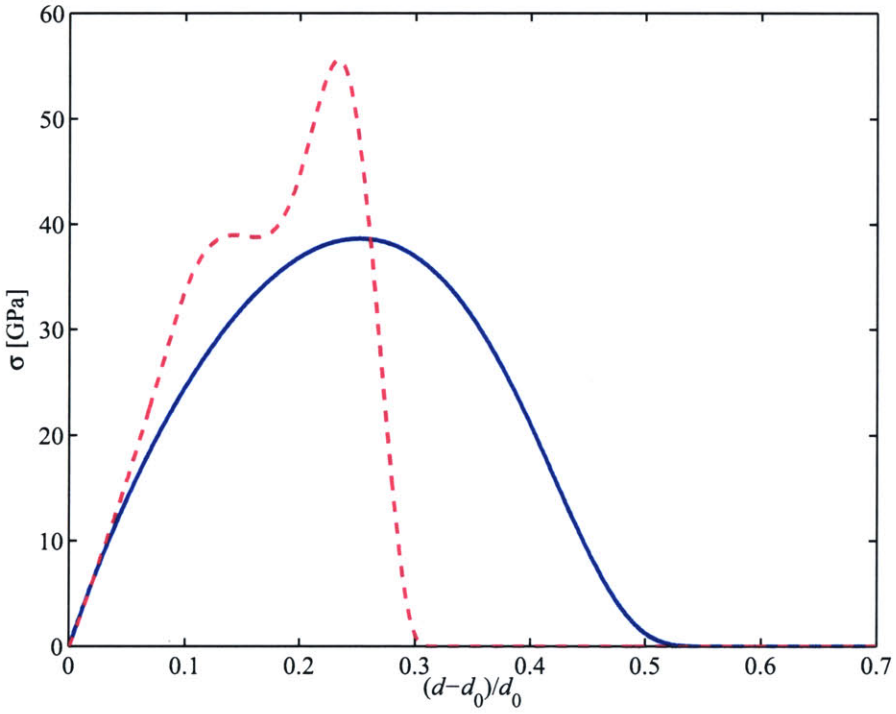


Figure 3-6: Cohesive responses of cleaving a single crystal of Si along the {111} shuffle plane: The solid line is for the SW potential and the dashed line for the EDIP potential.

Interatomic potentials

The main results presented in this paper are calculated using the three-body Stillinger-Weber potential (Stillinger & Weber [92]). The material properties, including lattice parameter, elastic constants, and relaxed {111} shuffle plane[1] surface energy, are listed in table 3.1. For comparison, these properties are also calculated using the environment dependent interatomic potential (EDIP) (Bazant *et al.* [12]; Justo *et al.* [44]) and given in table 3.1. Due to the short-ranged nature of the two potentials (Holland and Marder [41]), the cleaved {111} surfaces are the ideal 1×1 type without reconstruction. Figure 3-6 shows the cohesive responses of uniformly cleaving a perfect crystal of Si along the {111} shuffle plane into two blocks calculated from the two potentials. Here, surface traction is plotted as a function of opening displacement normalized by the equilibrium interplanar spacing $d_0 = \sqrt{3}a_0/4 = 2.35$Å. We first note that, in order to integrate these curves to obtain the correct surface energy, the tractions calculated from the short-ranged SW and EDIP models are generally larger than the real value needed for the same interplanar separation (Holland and Marder [41]). Comparing the two potentials, besides a spurious bend on the rising branch of the EDIP potential, significant differences are seen in the peak stress as well as in the slope on the falling side of the cohesive response. That is, beyond the interplanar separation of respective traction peaks, the traction calculated from the SW potential decreases smoothly to zero, while the response from the EDIP potential exhibits a sharp drop. It has been an intensively investigated topic regarding the influence of the interatomic force law on the lattice trapping effect (Sinclair [88]; Thomson *et al.* [101]; Curtin [22]). Based on the analyses of a series of simplified interatomic force laws, it was found that the slope of the falling side of the force law critically affects the loading range within which the lattice trapping effect exists (Thomson *et al.* [101]). The effect of the SW and EDIP potentials on lattice trapping barriers will be discussed in section 3.2.5.

[1]For Si with a diamond-cubic structure, there are two types of {111} plane: one is the shuffle plane which cuts through single covalent bonds along the direction perpendicular to the {111} plane and the other is the glide plane which cuts through triplets of covalent bonds inclined equally to the {111} plane.

Finding transition pathways of crack extension at different loads

The crack studied is under a pure mode I load as given by the stress intensity factor K_{I}. An atomistically faithful characterization of lattice trapping barriers requires calculations being able to accurately capture the fine structures of the system's atomistic energy landscape. To this end, we employ the Hessian-free Nudged Elastic Band (NEB) method (Jonsson *et al.* [43]) to determine the transition pathways of crack tip bond breaking and calculate the associated activation energy barriers. The NEB calculation requires a knowledge of both initial and final states. In other words, we need to identify two different local energy minima on the same potential energy surface mediated by K_{I}. For the present simulation setup, both the initial and final states are well-defined. That is, they correspond to the equilibrium atomic configurations before and after crack extension in the x_1' direction by one atomic spacing, denoted by Δa. Note that the K-field load is applied via the displacement-controlled method, which requires that the positions of atoms at the outer boundary of the simulation cell are the same for both the initial and final states, as well as the same for those intermediate states along the pathway. After both the initial and final states are identified using the energy minimization scheme, a discrete elastic band consisting of a finite number of replicas of the system is constructed to connect them (Jonsson *et al.* [43]). The final relaxed pathway is the so-called Minimum Energy Path (MEP). The MEP is defined as a continuous path in $3N_f$ dimensional configurational space (N_f is the number of free atoms) with the property that at any point along the path the atomic forces are zero in the $3N_f - 1$ dimensional hyperplane perpendicular to the path (Sorensen *et al.* [90]). The energy maximum along the MEP is the saddle-point energy which gives the activation energy barrier. The calculation is considered converged when the configurational force on each replica vertical to the path is less than 0.002 eV/Å. Of particular note is that, in any NEB calculation, the reaction coordinate is unknown *a priori* and emerges as a natural result from the converged pathway. Specifically, we define the hyperspace arc length along the MEP, i.e., $\int_{\mathrm{MEP}} \sqrt{d\mathbf{x}_{3N_f} \cdot d\mathbf{x}_{3N_f}}$, as the reaction coordinate, where $\mathbf{x}_{3N_f}$ denotes a state in

the hyperspace. Empirically, the bond length of the breaking Si-Si bond has been taken as an approximate reaction coordinate (e.g., Bernstein and Hess [13]). This approach of calculating the energy barrier corresponds essentially to finding the maximum point along a path which is constructed by connecting all the energy minima within a series of hyperplanes vertical to a specific direction, i.e, the direction $\Delta_s \mathbf{x}_{3N_f}$ corresponding to the degree of freedom represented by the Si-Si bond distance in a multidimensional hyperspace. Obviously, it is only in some cases (e.g., $d\mathbf{x}_{3N_f} \cdot \Delta_s \mathbf{x}_{3N_f}$ is always positive) that this path may agree with the MEP defined above and thus give an exact result for the activation energy barrier. But our calculations confirm that, due to localized nature of Si bond breaking, the path obtained from the calculation using the empirical reaction coordinate of the Si-Si bond length is close to the actual MEP in that they give the same transition states and barrier heights. Thus, the relatively abstract, though physically rigorous, reaction coordinate obtained from the NEB calculation can be approximately understood as the bond length of the breaking Si-Si bond. Nevertheless, the NEB method represents a more robust way of calculating the barrier height; it can be directly applied to study more complicated transition pathways when bond breaking involves foreign chemical species (e.g., Zhu *et al.* 2004b, submitted). Also note that for the same initial and final states, there may exist several competing pathways with different activation energy barriers. Along the same pathway, multiple saddle points may exist, and the kinetic barrier is determined by the highest saddle point.

3.2.3 Formation and migration of kink pair for a $(111)[1\bar{1}0]$ crack

In this subsection, we present results on the study of a (111) crack with a straight crack front along the $[1\bar{1}0]$ direction using the SW potential. This crack orientation has been used in most 2D simulations of Si cleavage fracture in the literature.

Griffith critical load K_{IG}

We first determine the Griffith load K_{IG}, at which the change in the system's total energy ΔE is zero when the crack extends in the $[11\bar{2}]$ direction by one atomic spacing Δa, with borders fixed. Corresponding to the present setup with 20 unit cells along the $[\bar{1}10]$ direction (PBC), 20 bonds need to be broken along the crack front to accomplish this process. The value of K_{IG} can be estimated from the relaxed surface energy γ_s using the Griffith relation in linear elastic fracture mechanics (LEFM) (e.g., Rice [76]). That is,

$$\Delta E = (\mathcal{G}_c - 2\gamma_s)\Delta A = 0, \tag{3.1}$$

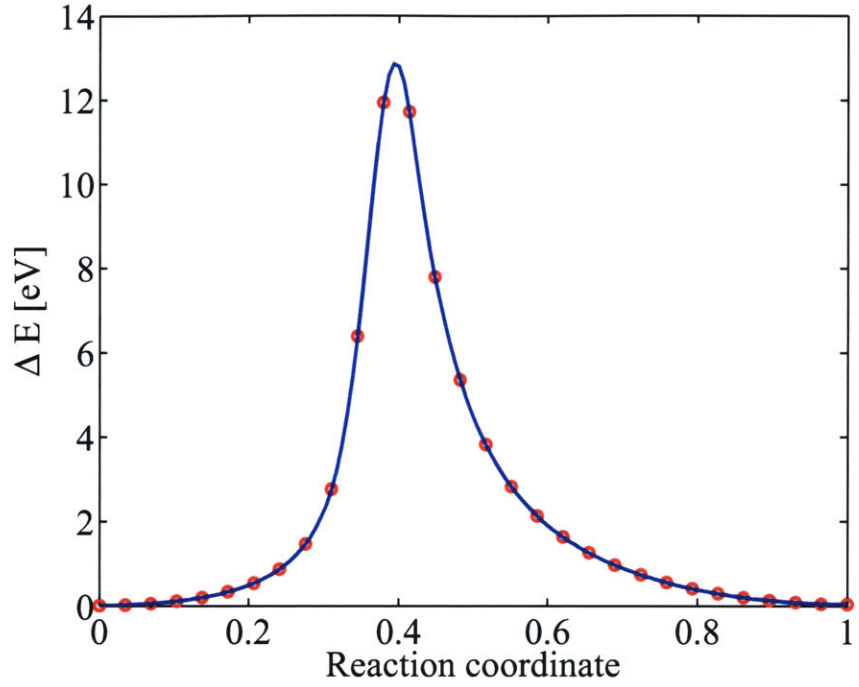
where $\mathcal{G}_c$ denotes the critical strain energy release rate and $\Delta A = l\Delta a$ is the area of newly created surface. For a linear elastic, anisotropic solid, $\mathcal{G}_c$ is related to K_{IG} by

$$\mathcal{G}_c = \frac{1}{4}\mathrm{H}_{2'2'}\, K_{\mathrm{IG}}^2, \tag{3.2}$$

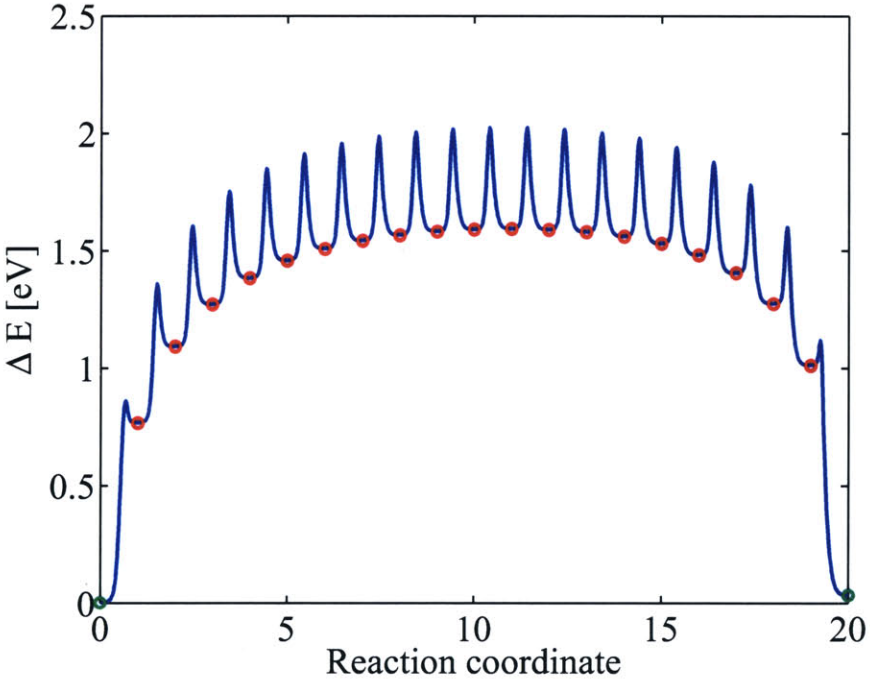
where $\mathrm{H}_{2'2'}$ is the effective compliance of the cracked system within the $x_1' - x_2'$ plane and it is the 22 component of the matrix $\mathbf{H}$ defined by equation (2.13) in the paper of Suo [94]. Note that $\mathbf{H}$ was defined in terms of elastic constants $C_{i'j'}$ in the global coordinate system (x_1', x_2', x_3') where the crack front is along the x_3' axis. For the $(111)[1\bar{1}0]$ crack, the computed $\mathrm{H}_{2'2'}$ is $2.692 \times 10^{-11}\mathrm{Pa}^{-1}$ using the elastic constants given by the SW potential as listed in table 3.1. Then substitution of $\mathrm{H}_{2'2'}$ and surface energy γ_s into equations (3.1) and (3.2) gives an estimate of the Griffith load $K_{\mathrm{IG}} = 0.656\ \mathrm{MPa}\sqrt{\mathrm{m}}$. On the other hand, direct atomistic calculation to be detailed in the next subsection gives a numerical result of $K_{\mathrm{IG}} = 0.646\ \mathrm{MPa}\sqrt{\mathrm{m}}$. The good agreement in K_{IG} attests to the accuracy of atomistic calculation.

2D vs. 3D crack extension at K_{IG}

Two competing pathways in our atomistic simulation are studied that advance the crack along the $[11\bar{2}]$ direction by one atomic spacing. In this subsection, we focus on



(a)



(b)

Figure 3-7: Energy variation along the MEPs of breaking 20 crack front bonds at the Griffith load $K_{\mathrm{IG}} = 0.646$ MPa$\sqrt{\mathrm{m}}$: (a) simultaneous breaking; (b) sequential breaking.

comparing the overall features of the two pathways and leave the detailed discussions on the 3D pathway into the next subsection.

The first pathway is characterized by breaking every bond along the crack front *simultaneously*. This is essentially a 2D fracture mode since what happens within a unit cell is repeated along the crack front direction. The corresponding MEP can be simply calculated using a 2D crack configuration. That is, only one unit cell is needed along the crack front with PBC imposed. Obviously, the activation energy barrier will grow linearly with the thickness of the simulation cell in the crack-front direction. For later comparisons, figure 3-7(a) shows energy variation along the MEP of simultaneously breaking 20 bonds along the crack front. Here, the circles represent the calculated energies of the replica configurations and the continuous curve is constructed by cubic-polynomial interpolation. The reaction coordinate is the normalized hyperspace arc length along the MEP. It is first seen that, in advancing the crack by one atomic spacing along the $[11\bar{2}]$ direction, the net change in total energy between the initial and final states is zero, at $K_{\mathrm{IG}} = 0.646\mathrm{MPa}\sqrt{\mathrm{m}}$. This value is obtained by trial-and-error searches in balancing the initial with final state total energies. Evidently, at the Griffith load K_{IG}, energy change is zero due to the cancellation between the elastic energy decrease and the surface energy increase. Moreover, the MEP shown in figure 3-7(a) has only a single peak with the activation energy barrier of about 12.85 eV, corresponding to an average of 0.64 eV per crack-front bond. We note that the present estimate of the barrier height per bond for the 2D fracture mode, obtained by using a reasonably good empirical potential and a highly accurate scheme of finding the transition state, is considerably higher than those predicted by Sinclair [88] using highly simplified interatomic force laws, which are in the range of 0.01 to 0.03 eV at the Griffith load corresponding to the interatomic potential he used.

In contrast to the first pathway which is 2D in nature, the second pathway involves breaking 20 bonds *sequentially* along the crack front and hence the process is inherently 3D. Energy variation along the 3D MEP is shown in figure 3-7(b). Here, the normalized reaction coordinate s is defined such that breaking one bond will in-

crease s by one. Thus each integer number s labels a locally equilibrated state with s broken bonds on the crack front and the intermediate coordinate between s and $s+1$ is the normalized hyperspace arc length along the MEP of further breaking the $(s+1)$th bond. Correspondingly, each red circle in figure 3-7(b) represents the energy, in reference to that of the equilibrium state with a straight crack front, of a metastable state of local equilibrium with s broken bonds on the crack front. The curve with a single peak connecting two adjacent circles gives the energy variation along the MEP of further breaking one more bond. Note that, for clarity, we only plot the interpolated curve indicating the continuous energy variation along the MEP and leave out the discrete data points corresponding to the calculated energies of 15 relaxed replica configurations between two adjacent circles. The detailed analysis on energy variation along this 3D transition pathway will be given in the next subsection. However, the overall kinetic rate for the second mechanism can be determined by the maximum barrier height which corresponds to the saddle point on the MEP of breaking the 11th bond. Hence, the kinetic barrier is about 2.02 eV for the second pathway.

Comparing energy variations along the two transition pathways at K_{IG}, which have the same initial and final states, we quantitatively confirm that even for the present simulation cell with nanoscale thickness (76.8 Å) along the crack front, the 3D extension mode is far more favorable kinetically than the 2D mode for cleavage crack propagation. Therefore, in the rest of the paper, we will focus on the energetics of 3D extension under different load levels and for different crack orientations.

Geometry and energetics of the crack-front kink pair at K_{IG}

The detailed discussion on the 3D cleavage mode is given in this subsection. Figure 3-8 shows the geometry of the crack front for a representative state of local energy minimum with 10 broken bonds. A continuous field of crack opening displacement across the two adjacent (111) planes are rendered by cubic-spline interpolation of the opening displacements at discrete lattice sites, normalized by the interplanar spacing d_0. Readers should be aware that this is just an effective way of visualizing completely

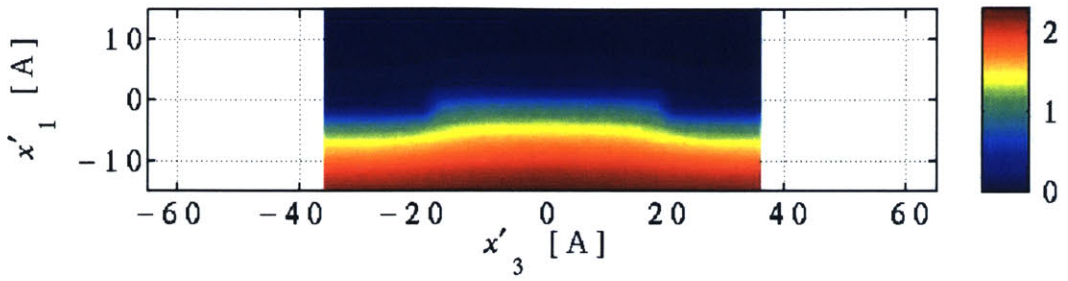
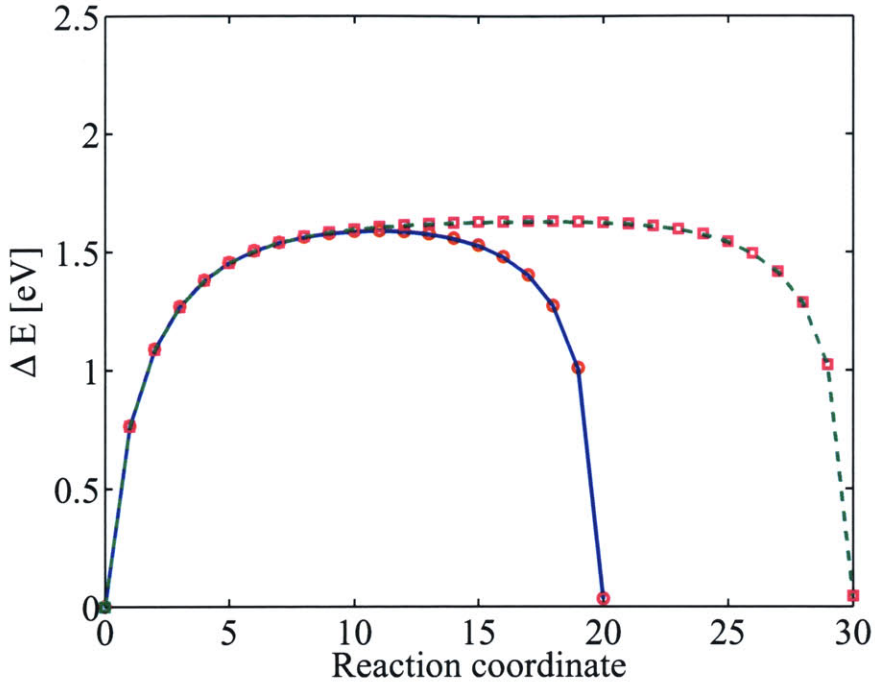


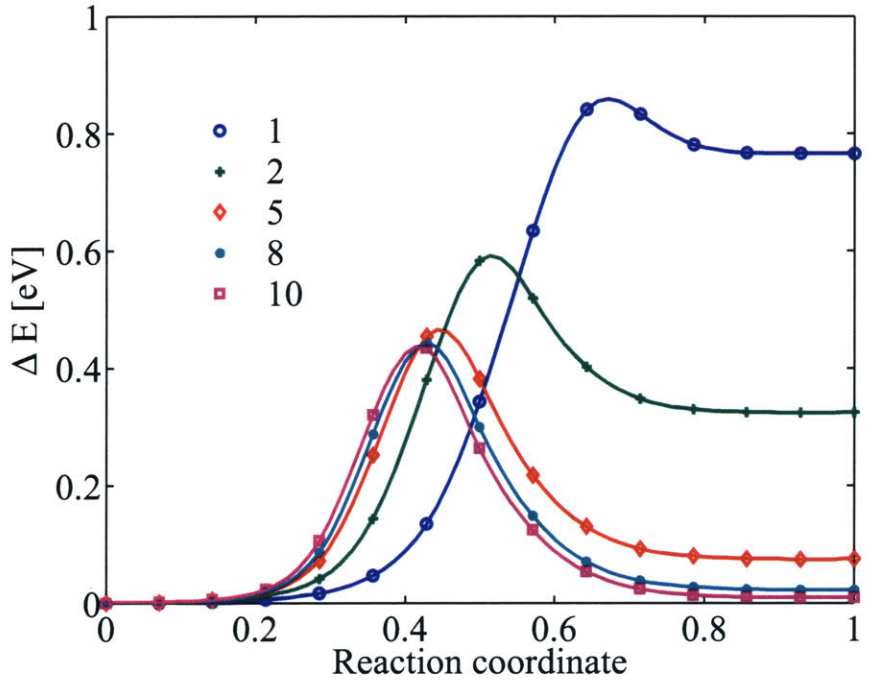
Figure 3-8: Distribution of opening displacement across the (111) cleavage plane for the (111)[1$\bar{1}$0] crack. The kink pair consists of 10 broken bonds on the crack front.

discrete atomistic calculation results. It is seen that two kinks of opposite signs with sharp features are developed locally on each side of the advanced crack front. This highly localized mode of crack opening distribution is in distinct contrast with our previous result of crack-tip nucleation of embryonic dislocation loop in an FCC Cu crystal which exhibits a significant spreading of shear fault distribution across the glide plane (Zhu *et al.* [123]). More importantly, this demonstration of localized bond-breaking in a covalent solid lends support to our unit-process studies of bond-strength reduction by water molecules to address the problem of environmental effects on crack-tip bond breaking (Zhu *et al.* 2004, in preparation).

Since the crack front with s broken bonds is characterized by a pair of kinks with opposite signs, it is convenient to describe the evolution of crack-front morphology in terms of kink-pair formation and kink migration. Consider first the kink-pair energies as given by the circles in figure 3-7(b). Now we extract those data points and replot them in figure 3-9(a). In other words, the bottom envelope of figure 3-7(b) is shown in figure 3-9(a). Also, we show in figure 3-9(a) the kink-pair energies for a larger calculation consisting of 30 unit cells along the crack front ($l = 115.2$Å), but with otherwise identical setup. It is seen that, in both cases, to break the first bond beyond the straight crack front requires an input of extra energy from the thermal reservoir. This trend of energy increase is sustained in the subsequent bond breaking. However, the energy increment associated with further bond breaking will gradually decrease as the separation of kink pair, denoted by d_{Kink}, increases. Consequently, the kink-pair energy approaches an asymptotic constant as d_{Kink} becomes large enough. To model



(a)



(b)

Figure 3-9: Lattic trapping barriers at K_{IG}: (a) Kink pair formation energies for the simulation cell with 20 (solid line) and 30 (dashed line) bonds along the front, respectively; (b) Selective MEPs of breaking 1st, 2nd, 5th, 8th, 10th bond on the crack front.

the effect of kink-pair separation, the system's total energy (in reference to a straight crack front) can be partitioned into d_{Kink}-independent and d_{Kink}-dependent parts. The latter is the interaction energy $E_{\text{Kink-Int}}$ between two kinks of opposite signs separated by d_{Kink}, and the former is twice the self energy $E_{\text{Kink-Self}}$ of an individual kink. That is to say, we write

$$\Delta E = 2E_{\text{Kink-Self}} + E_{\text{Kink-Int}}(d_{\text{Kink}}) \tag{3.3}$$

for the ideal situation of a kink pair embedded in an infinitely long, and otherwise perfectly straight, crack front. Following the theory of dislocation kinks, we conjecture that $E_{\text{Kink-Int}}(d_{\text{Kink}})$ is of algebraic form $-Cd_{\text{Kink}}^{-\alpha}$ with $\alpha > 1$, and the constant C should depend only on the elastic constants of the material and geometric characteristics of the kinks. In other words, $E_{\text{Kink-Int}}(d_{\text{Kink}})$ is asymptotically a smooth function describing the elastic interaction between two opposite kinks. $E_{\text{Kink-Self}}$ then contains all the remaining atomistic energetic information, and can be interpreted as the *formation energy* of an isolated kink on crack front.

From figure 3-9(a), the self energy $E_{\text{Kink-Self}}$ estimated from the plateau of the solid line, which comes from the simulation cell of $l = 76.8$ Å, is about 0.80 eV. This value is very close to an estimate of 0.81 eV from the thicker, $l = 115.2$ Å simulation cell. Furthermore, figure 3-9(a) shows that a further increase of d_{Kink} beyond $l/2$ will decrease the kink-pair energy. This arises due to image interactions, as we are not actually in the ideal situation of equation (3.3) of an infinitely long crack front. Specifically, when $d_{\text{Kink}} > l/2$, the attractive interaction from the image of the other kink in the neighboring simulation cell will begin to dominate, leading to a decrease of the total energy. Because by construction, the initial and final states in advancing the crack by one atomic spacing in the x_1' direction, i.e., the crack front at $x_1' = 0 \times \Delta a$ and $1 \times \Delta a$, have the same energy at K_{IG}, it is straightforward to show that after the image interactions have been taken into account, the kink pair energy in the PBC cell should be roughly symmetric with respect to $d_{\text{Kink}} = l/2$. Fortunately, it appears from numerical results that $E_{\text{Kink-Int}}(d_{\text{Kink}})$ is a rapidly decreasing function with d_{Kink}, so

even with image interactions present, we can still unambiguously determine $E_{\text{Kink}-\text{Self}}$ to be around 0.8 eV.

Consider next the kinetic barriers for kink migration. Recall that each curved segment connecting two adjacent circles in figure 3-7(b) represents the MEP of migrating a kink laterally by one atomic spacing. To facilitate comparison of kink migration barriers at different d_{Kink} values, we show in figure 3-9(b) the MEP of breaking the 1st, 2nd, 5th, 8th, 10th bond, respectively. These curves in figure 3-9(b) are essentially the folded and magnified presentations of the "cusps" in figure 3-7(b), with the energy reference for each curve being that of the initial state of breaking an individual bond, so all curves start from zero. Also shown in figure 3-9(b) are the data points of the computed energies of replica configurations along each pathway. It is seen first that breaking each individual bond from the 1st to the 10th is thermodynamically unfavorable in that the final state has a higher energy than the initial state. Moreover, the kinetic barrier for a forward transition will decrease, from 0.86 eV of breaking the first bond to 0.44 eV of breaking the 10th bond. As the number of broken bonds, i.e., the separation between two kinks d_{Kink}, increases, the migration barrier will approach an asymptotic value corresponding to the kinetic barrier of migrating an isolated kink, denoted as $E_{\text{Kink}-\text{Migr}}$. This value is given approximately by the barrier height of breaking the 10th bond determined to be 0.44 eV. In addition, we compare the present d_{Kink}-dependent kinetic barrier of breaking a bond with the barrier height per bond for the 2D pathway in section 3.2.3 in which all bonds at the crack front break simultaneously. It is of interest to observe that the kinetic barrier of 0.64 eV per bond from the 2D calculation is close to the average of the upper and lower limits from the 3D calculation.

Loading effect and lattice trapping range

In the literature, the Griffith critical load K_{IG} is also termed the neutral load (e.g., Sinclair [88]), because the initial and final states in advancing the crack by one atomic spacing have the same energy. As the applied load increases beyond K_{IG}, the potential energy landscape of the system will be tilted such that a forward transition becomes

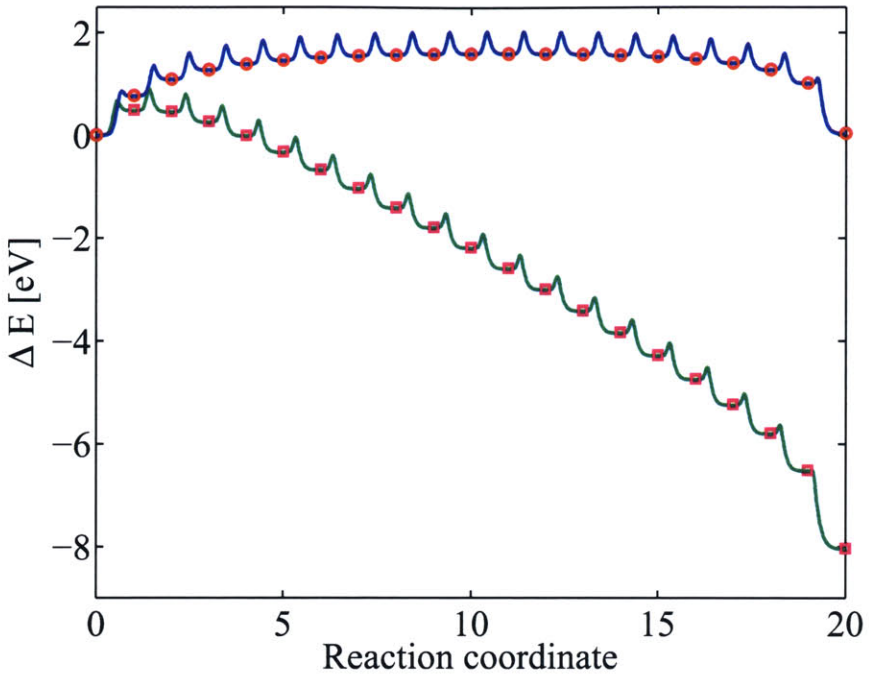
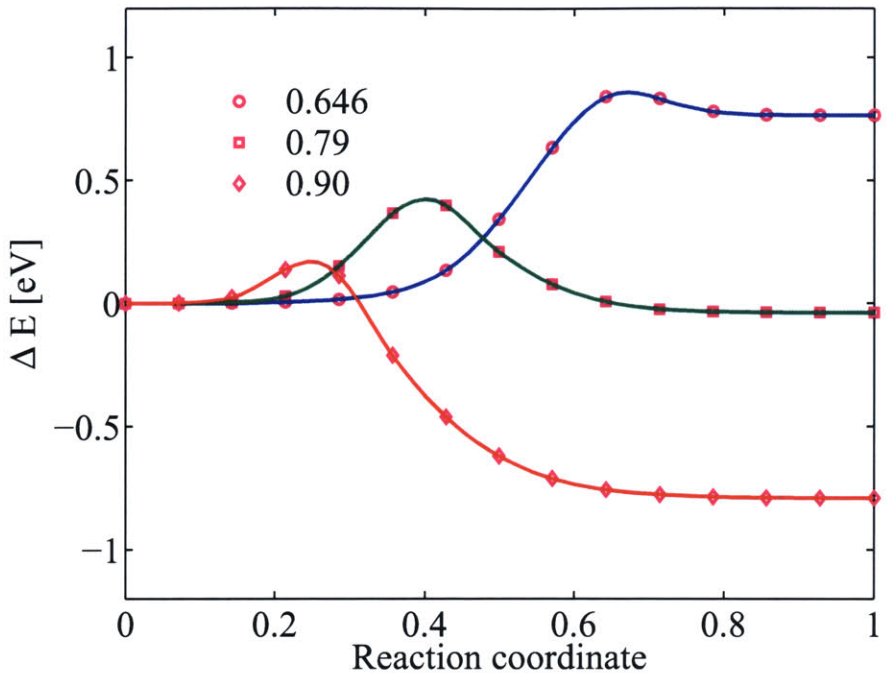


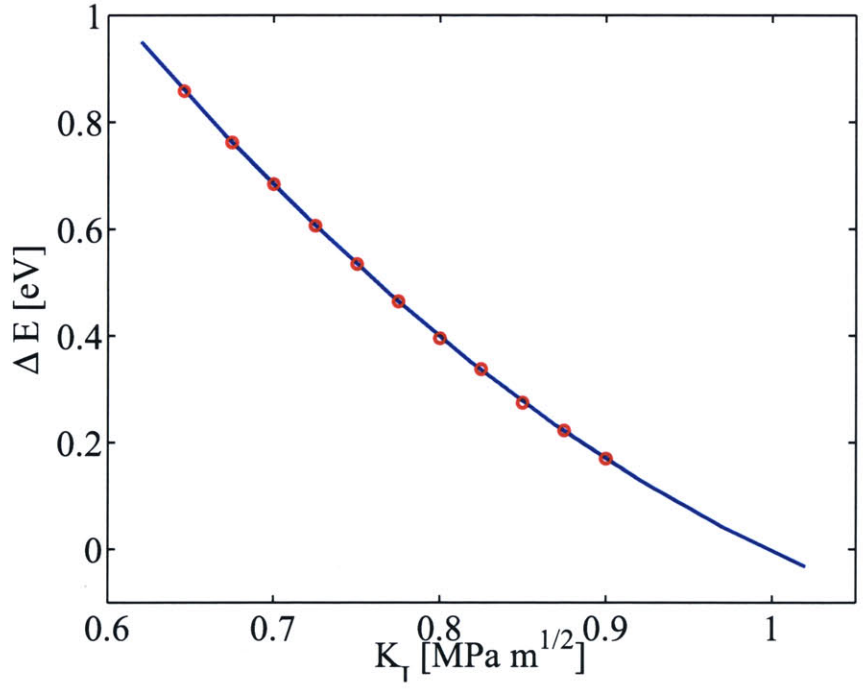
Figure 3-10: The MEPs at two load levels: the blue line is for $K_{\mathrm{I}} = 0.646$ MPa$\sqrt{\mathrm{m}}$ and the green line is for $K_{\mathrm{I}} = 0.70$ MPa$\sqrt{\mathrm{m}}$.

favorable both thermodynamically and kinetically. This loading effect can be clearly seen from the MEP of sequentially breaking 20 bonds at a higher load of $K_{\mathrm{I}} = 0.70$ MPa$\sqrt{\mathrm{m}}$ as shown in figure 3-10. For comparison, the same MEP shown in figure 3-7(b) at the Griffith load $K_{\mathrm{IG}} = 0.646$ MPa$\sqrt{\mathrm{m}}$ is replotted in figure 3-10, but adjusted to the new energy scale. It is seen from the MEP at $K_{\mathrm{I}} = 0.70$ MPa$\sqrt{\mathrm{m}}$ that there exists a critical distance between a pair of kinks, $d^{\mathrm{Crit}}_{\mathrm{Kink}}(K_{\mathrm{I}})$, above which further separation of the kink pair will decrease the system's energy and thus becomes thermodynamically favorable. Clearly, $d^{\mathrm{Crit}}_{\mathrm{Kink}}(K_{\mathrm{I}})$ is finite if and only if $K_{\mathrm{I}} > K_{\mathrm{IG}}$. At $K_{\mathrm{I}} = 0.70$ MPa$\sqrt{\mathrm{m}}$, this state corresponds to that with 1 broken bond on the front as shown in figure 3-10. In addition, the kinetic barrier of kink migration (moving in the direction of separating the kink pair) decreases compared to the corresponding barrier height with the same d_{Kink}, but at a lower load of K_{IG}.

To characterize K_{I}-dependent energetics of kink-pair formation and migration, two more critical loads, beyond the Griffith load K_{IG}, can be defined: The first one, denoted by $K^{\mathrm{f}}_{\mathrm{I}}$, is related to the thermodynamic energy balance of kink formation.



(a)



(b)

Figure 3-11: Energetics of breaking the first bond out of a straight crack front: (a) The MEPs at three typical loads of K_{I}; (b) The activation energy barriers as a function of load K_{I}.

In other words, at K_I^f, the energy for nucleating the smallest possible kink pair, i.e., breaking the first bond beyond the straight front, vanishes. For the present simulation setup, $K_I^f \approx 0.79$ MPa$\sqrt{\text{m}}$. The MEP of breaking the first bond at K_I^f is shown in figure 3-11(a). Evidently, the energy change is zero between the initial and final states. But there is still a kinetic barrier of about 0.43 eV at this load level.

The second critical load, denoted by K_I^m, is the athermal load which is related to the kinetics of kink-pair formation. That is, at K_I^m, the activation energy barrier for breaking the first bond at the straight crack front vanishes. More importantly, our calculations of MEPs at different load levels have indicated that, compared to the barrier heights of later breaking other bonds, the kinetic barrier of breaking the first bond is the slowest decreasing one as K_I increases. Hence, above K_I^m, the first kink pair nucleation as well as subsequent kink migration can take place spontaneously without the aid of thermal fluctuations, leading to instantaneous cleavage fracture.

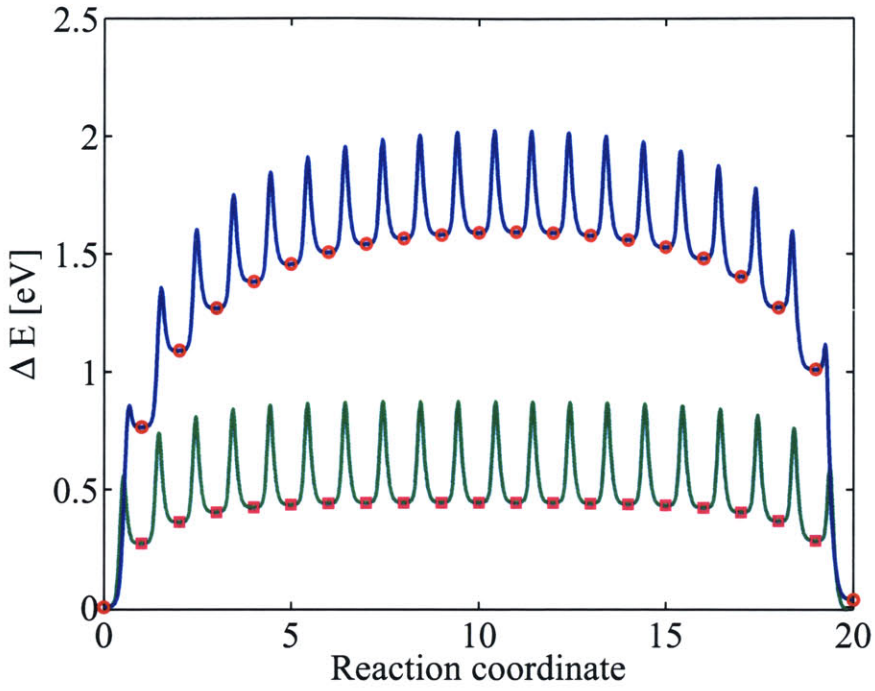
In numerically determining K_I^m, we apply the load incrementally to the system. At each load level, the corresponding MEP of breaking the first bond is obtained via the NEB calculation. Figure 3-11(a) shows energy variations along MEPs at three typical loads. Here, markers give the calculated energies of replica configurations along the pathway and the continuous curves are constructed via interpolations. We find that, as the applied load increases above $K_I \approx 0.90$ MPa$\sqrt{\text{m}}$, we are unable to obtain the relaxed initial state that is geometrically similar to those at lower loads. In other words, the system tends to relax to another type of local energy-minimum state which corresponds to a different deformation mode, i.e, amorphization by forming new crack-tip ring structures. Since the focus of this work is to study a particular kind of transition pathway, i.e., cleavage fracture along the (111) plane via crack-front bond breaking, we leave this interesting topic of competitions among different deformation modes to future investigation by noting that the study of these two competing modes requires a force field capable of accurately describing both types of deformation. Nevertheless, we can estimate K_I^m by extrapolating the barrier heights calculated at lower load levels as shown in figure 3-11(b), where circles are the calculated data points and the solid line is the polynomial fitting curve. It is seen

that the activation energy barrier for breaking the first bond shows a slight nonlinear dependence on K_{I}. The value of $K_{\mathrm{I}}^{\mathrm{m}}$, obtained by extrapolating the curve down to zero activation energy, is about 1.0 MPa$\sqrt{\mathrm{m}}$. This athermal load is the upper limit of lattice trapping. Hence, the lattice trapping range given by the SW potential, defined here in terms of the ratio $K_{\mathrm{I}}^{\mathrm{m}}/K_{\mathrm{IG}}$, is about 1.55. In particular, we note that since the 3D pathway is more favorable kinetically than the corresponding 2D pathway at the same load level, the present estimate of lattice trapping range should be a lower bound of that estimated from a 2D calculation. Furthermore, we point out that the lattice trapping range predicted by the SW potential may still differ significantly from the true response of Si considering the approximate nature of empirical potential. However, the self-consistent calculations as well as the detailed analyses we have presented above represent a significant improvement to an atomistically accurate characterization of the 3D, load-mediated lattice trapping barriers to brittle fracture.

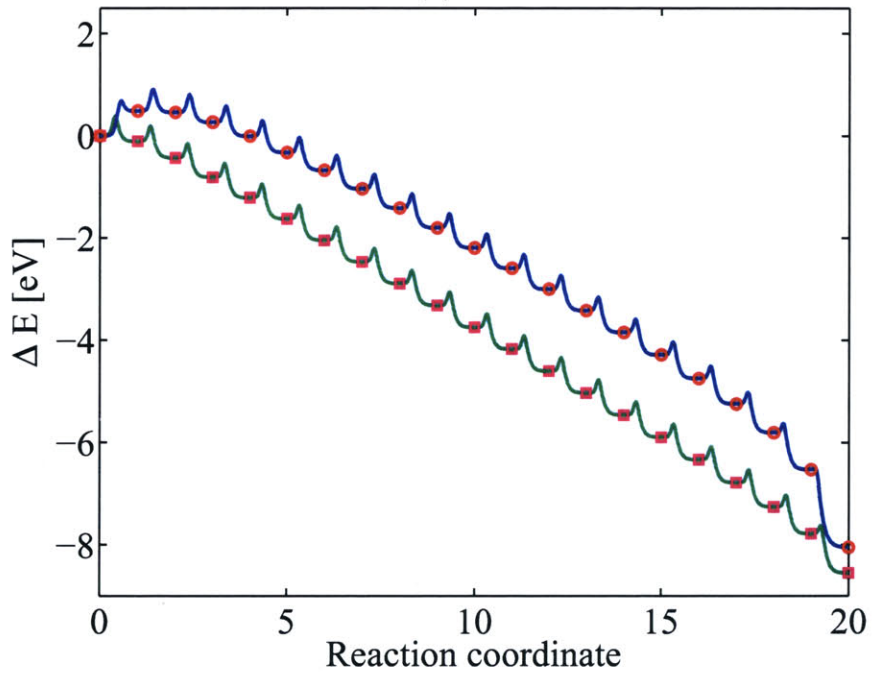
3.2.4 Propagation anisotropy

Although the $\{111\}\langle 1\bar{1}0\rangle$ crack (propagating in the $\langle 11\bar{2}\rangle$ direction) is the most frequently studied crack orientation, experimental fractography observations of the cleavage surface indicate that, instead, a $\{111\}$ cleavage crack prefers to propagate in the $\langle 110\rangle$ direction (e.g., George and Michot [30]). This preference of a crack propagation direction *within* a certain cleavage plane is called propagation anisotropy. Perez and Gumbsch [73, 74] have studied propagation anisotropy in Si by comparing orientation-dependent lattice trapping ranges of 2D crack configurations. Here, we provide a more detailed explanation of propagation anisotropy in terms of the orientation-dependent energetics of crack front kink-pair formation and migration.

We first note that it is fundamentally difficult to rationalize propagation anisotropy by only thermodynamic arguments (Perez and Gumbsch, 2000*a*). That is, the Griffith criterion may suggest the preferential cleavage plane to be the one with the lowest surface energy γ_s. But, within the same cleavage plane, two cracks of different orientations would experience the same resistance γ_s to crack propagation and thus have the same critical energy release rate $\mathcal{G}_c$ as can be clearly seen from equation (3.1).



(a)



(b)

Figure 3-12: Comparison of lattice trapping barriers for two crack orientations: (a) the MEP at respective Griffith load; (b) the MEPs at a higher load level $K_{\mathrm{I}} = 0.70$ MPa$\sqrt{\mathrm{m}}$. In both figures, the blue lines are MEPs for the $(111)[1\bar{1}0]$ crack and the green lines are MEPs for the $(111)[11\bar{2}]$ crack.

Indeed, there may exist a difference in the critical stress intensity factor K_{IG} which arises from a variation of the effective compliance within the x_1'-x_2' plane due to elastic anisotropy for different crack orientations. However, as will be numerically shown next, this difference is so small that K_{IG} could not serve as a robust measure of propagation anisotropy.

On the other hand, consider a realistic situation in which the material may come with an assortment of flaws with distributions in sizes and aspect ratios. It seems unlikely that a special pre-crack orientation is selected *every time* to first trigger unstable propagation, just due to the condition of the initial flaws. A much more plausible scenario is that, with the increase in macroscopic load, an ensemble of crack orientations exceed the Griffith limit at approximately the same time; they all start to grow, aided by thermal activations, since the respective load K_{I} (here, for simplicity, we only discuss mode I load) is still less than the corresponding $K_{\mathrm{I}}^{\mathrm{m}}$. However, there is a special orientation whose front grows the fastest due to smaller kink-pair formation and migration energies. If the system is under load control, K_{I} will increases with a in a positive feedback, until this particular crack front reaches the athermal threshold, $K_{\mathrm{I}}^{\mathrm{m}}$, first. At that point, runaway dynamical fracture occurs that proceeds at near sound speed. Also, this orientational competition may occur not only at physically separate crack fronts, but also at different places in one contiguous crack front, leading to morphological evolution that favors faster growth of one orientation until $K_{\mathrm{I}}^{\mathrm{m}}$ is triggered first in that orientation. Therefore, we see that although the lattice trapping ranges could be quite narrow in brittle materials, it is necessary to do a careful study because a physically important question of which orientation triggers dynamical fracture first could depend on the exact numbers.

Consider a (111) crack with a straight crack front along the $[11\bar{2}]$ direction. For this new orientation, the crack will extend in the close-packed $[\bar{1}10]$ direction. The crack-tip atomic structure is shown in figure 3-5(b). Similar to section 3.2.3, we can estimate the Griffith load using the analytic LEFM Stroh solution and γ_s. The effective compliance calculated for the new crack orientation is $\mathrm{H}_{2'2'} = 2.71 \times 10^{-11}$ Pa^{-1}. Substitution of $\mathrm{H}_{2'2'}$ and (111) surface energy γ_s into equations (3.1) and (3.2)

gives an estimate of the Griffith load $K_{IG} = 0.654$ MPa$\sqrt{m}$. On the other hand, the value of K_{IG} determined from direct atomistic calculations is 0.643 MPa$\sqrt{m}$. Evidently, the value of K_{IG} for the $(111)[11\bar{2}]$ crack is very close to that of 0.646 MPa$\sqrt{m}$ for the $(111)[1\bar{1}0]$ crack, which numerically proves that the Griffith criterion cannot be applied to explain propagation anisotropy.

Though the difference in K_{IG} for different crack orientations is small, the difference in K_I^m may be large. We begin by studying the energetics of crack front kink-pair formation and migration for the $(111)[11\bar{2}]$ crack at its Griffith load $K_{IG} = 0.643$ MPa$\sqrt{m}$. Figure 3-12 shows the MEP of sequentially breaking 20 bonds along the crack front. For comparison, the corresponding MEP for the $(111)[1\bar{1}0]$ crack at its Griffith load $K_{IG} = 0.646$ MPa$\sqrt{m}$ is also shown in figure 3-12. Here, the circles/squares are the locally stable kink-pair energies which are the sums of kink-pair interaction energy and twice of the kink self-energy as given by equation (3.3), while the saddle points between circles/squares give the kink-pair migration barriers. Comparing the two MEPs, a significant difference is seen in the kink-pair formation energies between two crack orientations, while the kink migration barriers are similar. Specifically, for the $(111)[11\bar{2}]$ crack, the kink self-energy $E_{Kink-Self}$, which can be estimated from the asymptotic value of the lower envelope curve connecting the squares, has a much lower value about 0.22 eV. Moreover, the interaction between kink pairs is relatively weak. In contrast, for the $(111)[1\bar{1}0]$ crack, the kink self-energy is much higher with a value of about 0.8 eV.

As the applied loads increase by the same ratio, our calculations indicate that the kinetic barriers of extending the $(111)[11\bar{2}]$ crack are always lower than those of the $(111)[1\bar{1}0]$ crack, as demonstrated selectively in figure 3-12 using two MEPs at $K_I = 0.70$ MPa$\sqrt{m}$. Finally, in order to estimate the lattice trapping range for the $(111)[11\bar{2}]$ crack, the athermal load K_I^m needs to be identified. Similar to the situation for the $(111)[1\bar{1}0]$ crack, the system tends to relax to a state of crack-tip amorphization at higher loads. So we use the same method to estimate K_I^m as for the $(111)[1\bar{1}0]$ crack. Namely, we extrapolate the kinetic barriers of breaking the first bond at different K_I's down to zero barrier and find $K_I^m = 0.88$ MPa$\sqrt{m}$.

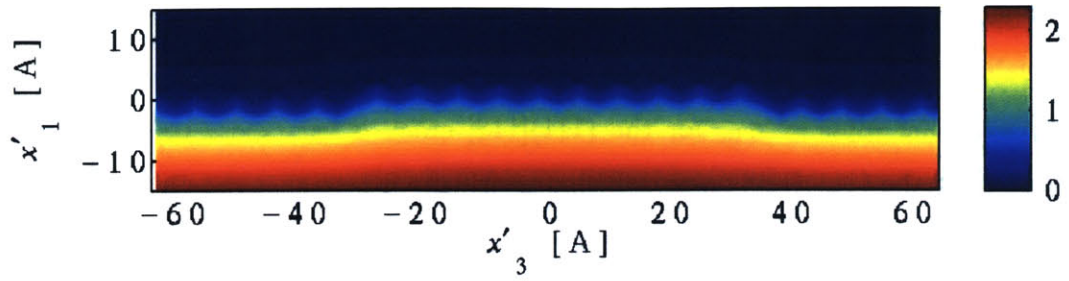


Figure 3-13: Distribution of opening displacement across the (111) cleavage plane for the (111)[11$\bar{2}$] crack. The kink pair consists of 10 broken bonds on the crack front.

Correspondingly, the ratio of $K_{\mathrm{I}}^{\mathrm{m}}/K_{\mathrm{IG}}$ is 1.37, which indicates a narrower lattice trapping range for the (111)[11$\bar{2}$] crack compared with that of the (111)[1$\bar{1}$0] crack. We note that the 2D simulations by Perez and Gumbsch (2000a, b) on {110} cleavage cracks also demonstrate orientation-dependent lattice trapping ranges. However, the present 3D calculations represent a more realistic probe of K_{I}-mediated, orientation-dependent lattice trapping barriers.

The orientation-dependent kink-pair formation energy is closely related to the bond densities along different directions within the (111) plane. The geometric profile of the kinked crack front for the (111)[11$\bar{2}$] crack at $K_{\mathrm{IG}} = 0.643$ MPa$\sqrt{\mathrm{m}}$ is shown in figure 3-13, where the distribution of opening displacement across the cleavage plane is plotted for the state with 10 broken bonds on the crack front. It is seen, first, that both the leading and trailing edges of the crack front have a zigzag profile, in contrast to the atomically smooth crack front for the (111)[1$\bar{1}$0] crack as shown in figure 3-8. The zigzag interpolated field indicates that the second array of bonds immediately adjacent to the crack front also have relatively large opening displacements. This can be explained as follows. Assuming the LEFM solution is approximately correct near the crack front, it follows that the bond opening displacement (in the x_2' direction) sensitively depends on its distance (in the x_1' direction) from the crack front. The close packing of atoms along the x_1' direction for the (111)[11$\bar{2}$] crack leads to a smaller distance of the second array of bonds from the crack front. As a result, larger opening displacements are necessary for the second array of bonds, leading to a zigzag crack front profile. In addition, as the second array of bonds becomes closer to the first

array of bonds by acquiring larger equilibrium opening displacements, we also expect that the atomistic energy barrier separating two arrays of crack front bonds, which is in analogous to the barrier of the Peierls ridges in the dislocation theory, will become smaller. It follows that the energy penalty for the crack front kinks, which is crossing the analogous Peierls ridges, is smaller, leading to a lower kink formation energy and the narrower lattice trapping range as well.

3.2.5 Discussions

Effect of interatomic potential

We selectively repeat some calculations using the EDIP potential to study the influence of the interatomic potential on results. The Griffith loads are first determined for the two crack orientations. Atomistic calculations give $K_{\mathrm{IG}} = 0.58$ MPa$\sqrt{\mathrm{m}}$ for the $(111)[1\bar{1}0]$ crack and $K_{\mathrm{IG}} = 0.59$ MPa$\sqrt{\mathrm{m}}$ for the $(111)[11\bar{2}]$ crack. Similar to the SW potential, the close values of K_{IG} between the two crack orientations arises due to a slight variation of the effective compliance within the $x_1' - x_2'$ plane. However, the SW and EDIP potentials give different predictions for the Griffith load due to different values of elastic constants and surface energy as listed in table 3.1.

Figure 3-14 shows the energy variations along the MEPs at the respective Griffith load for the two crack orientations calculated using the EDIP potential. Comparing to results from the SW potential as shown in figure 3-12, we find that the $(111)[11\bar{2}]$ crack orientation is also favored kinetically. However, despite the general agreement in trend, there are some quantitative differences in both the kink pair formation and kink migration energies. Specifically, the EDIP potential gives a lower kink formation energy, but a higher migration energy. Furthermore, our calculations of MEPs at different load levels using the EDIP potential indicate that the lattice trapping range is larger for both crack orientations compared to the corresponding result obtained from the SW potential. This difference in lattice trapping range between the two potentials can be correlated with the different cohesive responses of uniformly cleaving a perfect crystal as shown in figure 3-6, while keeping in mind that the actual crack tip bond

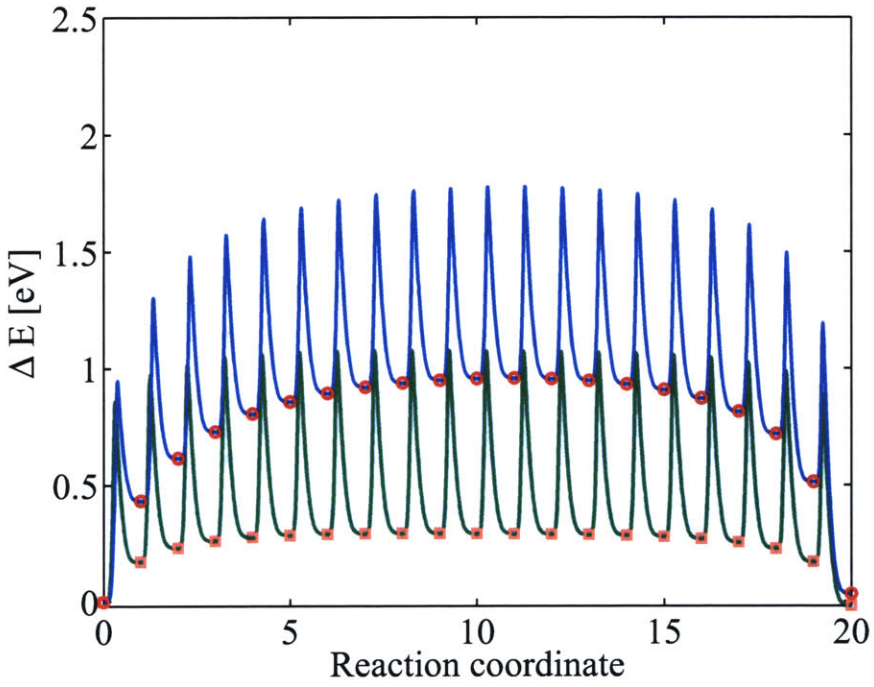


Figure 3-14: The MEPs calculated using the EDIP potential for two crack orientations at respective Griffith load: the blue line is for the $(111)[1\bar{1}0]$ crack and the green line is for the $(111)[11\bar{2}]$ crack.

breaking occurs under a non-uniform stress environment with a large stress gradient. Recall that figure 3-6 shows a significant difference in the slope on the backside of the cohesive response for the two potentials. Previous analyses from a series of simplified interatomic force laws indicated that a steep slope on the backside of the force law will lead to a large lattice trapping range (Thomson *et al.* [101]). Our present calculations show the same trend quantitatively. Finally, we note that a recent 2D study by Bernstein and Hess [13] on Si shows that the SW potential overestimates lattice trapping barriers compared to the results obtained from a multiscale simulation, i.e., the tight-binding description of bonding near the crack tip embedded in an empirical potential (EDIP) region. Nevertheless, the present study represents the first effort of mapping out 3D lattice trapping energy barriers using the atomistic approach. It is conceivable that a calculation using a more accurate force field along the lines presented here will give a more realistic estimate on the 3D lattice trapping barriers.

Other competing transition pathways

In the present work, we focus on studying the lattice trapping barriers along a specific 3D transition pathway, i.e., breaking crack front bonds sequentially, to extend the crack in the x_1' direction. Obviously, for a given K-field load, there exist other competing pathways to extend a crack. However, among a few possible scenarios we have studied, the present pathway is the most favorable kinetically. For example, given s broken bonds on the front, we have studied the following two schemes of further breaking bonds: The first one is to break the next two bonds at the front simultaneously instead of breaking them one by one. This scheme essentially corresponds to the situation in which the system evolves along different kinetic pathways, but will reach the same final state. The calculated MEP for the new scheme gives a much larger activation barrier. The second scheme is to break a bond on the advanced front instead of on the trailing front. For this case, both the saddle point and final state of the new scheme have higher energies than the corresponding values of the original scheme. Hence, this scheme is thermodynamically unfavorable; needless to say, the larger kinetic barrier is. Moreover, the study of morphological stability of an

initially straight crack front in a linear elastic solid by Rice [77] indicates that, for a crack under a displacement-controlled boundary condition, small perturbations from straightness should decay with time. Assuming this long wavelength analysis works for atomic-scale kinks, an initially straight crack front should maintain straightness during quasi-static extension, which agrees with our atomistic calculations.

Chapter 4

Stress-Dependent Reactivity of a Silica Nanorod with Water

4.1 Introduction

The rates of chemical reactions involving solids are dependent on the mechanical stress levels [33, 40, 55]. This effect on reaction kinetics can be attributed to a change of activation energy barrier in the presence of stress. A typical example of stress-dependent reaction kinetics is stress corrosion of silica glass by water. Specifically, the strength of silica glass will decrease with time when subjected to a static load in an aqueous environment [112]. This phenomenon, commonly referred to as delayed failure or static fatigue, results from slow growth of pre-existing surface flaws due to stress corrosion by water from environment. It has long been recognized that the coupling between mechanics and chemistry at the crack tip plays a central role in stress corrosion of silica glass. From a microscopic viewpoint [62, 63], the intrusive water molecules chemically attack the strained siloxane (Si-O-Si) bonds at the crack tip. Due to the elevated stress state near the crack tip, the hydrolysis reaction is significantly enhanced, leading to the breaking of siloxane bonds and subsequent formation of terminal silanol (Si-OH) groups. These newly formed terminal groups separate from each other due to mutual repulsion, thereby completing the bond rupture process. This molecular level mechanism of stress-enhanced chemical reactivity

will ultimately affect the macroscopic kinetics of quasi-static crack motion, leading to final catastrophic failure.

While recognizing the importance of molecular mechanism of crack tip chemistry, most early studies of the silica glass-water fracture system relied on phenomenological models to address the effect of stress corrosion by water. One of the earliest theories of chemistry in brittle fracture was put forward by Orowan [70], who extended Griffith's idea of equilibrium crack to study water-assisted crack growth by considering environmental effect on surface energy reduction. Within a more general framework of irreversible thermodynamics, Rice [76] derived an inequality governing the relation between the crack extension rate and the associated thermodynamic driving force in the presence of a chemically reactive environment. This relation is essentially a global restriction on the detailed molecular kinetics of crack growth. Beyond a purely thermodynamic perspective, Hillig and Charles [40] considered crack motion as a stress-enhanced thermal activation process and formulated an Arrhenius-type relation to describe stress corrosion rate based on reaction rate theory. This model was applied successfully by Wiederhorn [112] to correlate the experimental measurement of crack velocity as a function of stress intensity factor. In addition, Charles and Hillig's idea of stress-dependent thermal activation was also extended by Chuang and Fuller [19] to study morphological evolution of notch tip and its implication for stress corrosion cracking.

Despite the usefulness of the phenomenological models, a complete understanding of stress corrosion in silica glass requires a detailed study of molecular mechanism of crack tip chemistry and a quantitative characterization of stress-enhanced reaction kinetics. From the standpoint of molecular orbital theory, Michalske and Freiman [62, 63] first proposed a molecular interpretation of stress corrosion in vitreous silica. The chemical interaction between the intrusive water molecule and strained siloxane bonds was envisaged to proceed in a three-step sequence: (1) adsorption; (2) reaction and (3) separation. By considering the electron and proton-donating capacity of the intrusive chemical species and the polarity of the bridging bond, the Michalske and Freiman (MF) model explained why water and other chemical species such as

ammonia could be effective in promoting crack growth. However, the MF model did not study the effect of stress on chemical reactivity of solids. This deficiency was addressed by Michalske and Bunker [61]. Based on the molecular orbital calculation which shows that bond angle deformations are most effective in increasing the chemical reactivity, they suggested that the increasing chemical reactivity associated with deformation could be attributed to the stress-induced pinching of O-Si-O bond angle to form chemically active kink site. On the other hand, in contrast to Michalske and Bunker's approach which studied stress-dependent chemistry through directly imposing deformation on a certain type of siloxane ring, West and Hench [111] compared chemical reactivities of different ring structures, *i.e.*, various membered cyclosiloxane rings. From semi-empirical molecular orbital calculation (AM-1 method), they obtained quantitative reaction pathways for three- and four-fold silica rings and a five-fold ring-chain structure. Their results showed that three-membered ring is more chemically reactive due to its lower energy barrier associated with larger ring strain. In addition, the study of chemical reactivity of different ring structures is also useful in understanding the kinetics of hydrolyzing silica surface where various types of ring structures exist due to surface reconstruction, e.g., see Walsh *et al.* [109].

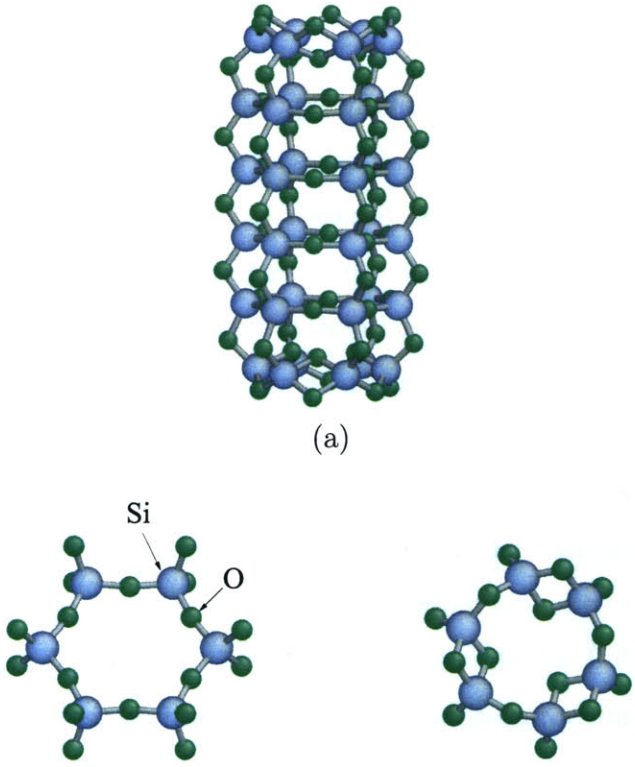
Although the above mentioned works provide great insights into the coupling between mechanics and chemistry at the molecular level, to our knowledge, no systematic studies have been made on the explicit calculation of reaction barriers as a function of mechanical stress to directly reveal stress-dependent chemical reaction kinetics. In this chapter, we present a detailed investigation of stress-enhanced hydrolysis reaction via a model system, *i.e.*, a single water molecule interacting with a silica nanorod structure. We construct the nanorod as a representative structure unit cut from a crack tip. Considering the discrete, localized nature of bond-breaking in covalent solids [55], we study the unit process of hydrolyzing a stretched siloxane bond at the surface of a nanorod as our first effort to reveal some fundamental aspects of hydrolysis reaction taking place at a stressed crack tip. In addition, we rely on this unit process study to address the problem of stress corrosion at a crack tip based on the assumption that intrusive water molecules have direct access to individual

crack-tip siloxane bonds. In other words, crack-tip hydrolysis reaction is the only rate limiting step for stress corrosion and other transport kinetic processes such as bulk diffusional flow and water diffusion at fractured surface are relatively fast. Furthermore, the silica nanorod is constructed as a mechanically meaningful structure, which is in contrast to the commonly used model clusters, e.g., $H_6Si_{18}O_7$ by Lindsay *et al.* [59]. The major structural characteristics of the nanorod is its geometric uniformity. This property allows for a well-defined stress and strain to be associated with the nanorod. Thus, we are able to explicitly map out families of reaction pathways, parametrized by the continuous nominal stress. It follows that our further analysis of stress-mediated reaction pathways can clearly reveal the stress-dependence of thermodynamic driving force and kinetic barriers of hydrolysis reaction. Through this analysis, we take a step closer to a truly fundamental description of thermodynamics and kinetics of quasi-static crack motion in silica glass.

4.2 Model and Method

4.2.1 Nanorod Structure

A silica nanorod composed of 108 atoms (36 SiO_2 units) is shown in Fig. 4-1(a) [124]. Our construction of a nanorod first involves the formation of planar six-membered rings, Si_6O_{18}. As shown in Fig. 4-1(b), each Si_6O_{18} ring contains six corner-linked SiO_4 tetrahedra, where each tetrahedron has two bridging oxygen atoms that are shared with two neighboring tetrahedra. The nanorod is then assembled by stacking such rings (layers) one over another to form a rod-like network of SiO_4 tetrahedra. For the structure to be free from dangling bonds, the nanorod is capped at the ends by 'terminating' rings in which the six Si atoms are connected alternately by a bridging oxygen or two oxygens, in the manner shown in Fig. 4-1(c). The resulting structure is a stoichiometric silica nanorod with the bulk built up of corner-sharing SiO_4 tetrahedra, where all Si and O atoms are four- and two-coordinated respectively, but capped by three pairs of edge-sharing SiO_4 tetrahedra at each end. The size and



(a)

(b)

(c)

Figure 4-1: Structure of (a) a relaxed nanorod, (b) an interior ring, (c) an end ring.

geometry of the nanorod are similar to the channels in α-quartz and the outer wall of the rod resembles the α-quartz (0001) surface, where all Si and O atoms are fully coordinated without the presence of dangling bonds. Moreover, the hexagonal planar six-membered rings of the nanorod exist in natural cyclosilicate minerals such as beryl which contains those rings linked by Be and Al ions.

4.2.2 Force Field

The study of stress-enhanced hydrolysis kinetics entails the proper description of interatomic interactions. Empirical potentials have been proven to have difficulty in treating correctly the effects of charge transfer when bonds are broken and formed. On the other hand, it is very computationally expensive to apply first principle quantum mechanical calculation to explore various reaction pathways under different stress states. As a reasonable compromise between accuracy and efficiency, we choose to use semi-empirical molecular orbital calculation (PM3 method). The energies for different nanorod configurations given below are calculated using MOPAC 2000. The comparison study for hydrolysis reaction of various silica clusters between density functional theory (DFT) calculation and semi-empirical molecular orbital method (AM1 and PM3) has been given by Laurence and Hillier [52]. They found that both optimized structures and transition states (for interaction with single water molecule) from semi-empirical molecular orbital methods are qualitatively similar to those obtained via DFT calculations, though there exist some quantitative differences. Considering that semi-empirical method is significantly faster than higher level first principle calculations, we apply semi-empirical calculations as our first step to explore stress-mediated reaction pathways.

4.2.3 Mechanical Deformation

Mechanical deformation of the nanorod is imposed via the displacement controlled method. We use 15 atoms in each end-ring as grips, with each atom given a prescribed displacement in the axial direction of the rod, and all other atoms in the rod are

allowed to relax using conjugate gradient method. To characterize the stress state of the nanorod we take the sum of all forces acting on the cross section divided by the nominal cross section area $A_0 = 4\pi d_{OO}^2 = 88.9\text{Å}^2$ where d_{OO} denotes the side length of a SiO_4 tetrahedron and it is also the edge length of the hexagon composed of six bridging Os within the planar six-membered ring. The nominal cross section of the nanorod is taken as the circle enveloping a hexagon with a radius $R = 2d_{OO}$. To measure the tensile strain applied to the nanorod, we take the elongation of the rod divided by the initial gage length (distance between two grips) $l_0 = 4d_{OO} = 13.3\text{Å}$.

4.2.4 Finding Reaction Pathway and Transition State

The study of stress-enhanced hydrolysis reaction kinetics is carried out within the framework of Transition State Theory (TST) (e.g., Vineyard [103]). The problem then becomes that of identifying the reaction mechanism and finding the free energy barrier. Within the harmonic approximation to TST, the problem is further reduced to finding minimum energy path (MEP) on the potential energy surface. The MEP is defined as a continuous path in $3N$ dimensional configuration space (where N being the number of free atoms) with the property that at any point along the path the atomic forces are zero in the $3N - 1$ dimensional hyperplane perpendicular to the path [90]. The energy maximum along the MEP is the saddle-point energy which gives the activation energy barrier. Since the potential energy surface will evolve with stress, the stress-enhanced kinetics manifests as a stress-dependent activation energy barrier.

The minimum energy paths of hydrolysis reactions are calculated using the nudged elastic band (NEB) method [43, 37]. Prior to NEB calculations, we first identify a physisorbed state which corresponds to a local energy minimum along the reaction pathway. Here, physisorption refers to the process of hydrogen bond formation between H_{wat} of water molecule and the bridging O_{br} of nanorod. During physisorption, there is no real chemical bond breaking and/or formation. In contrast, hydrolysis reaction is a chemisorbed process since it involves dissociation of water molecule as well as formation of new chemical bonds, *i.e.*, silanol surface groups.

The NEB calculations proceed in two steps: we first calculate the physisorption pathway by choosing as the initial state where the water molecule is separated from a stressed nanorod by about 15Å. The final state is one of the physisorbed states, while keeping extension of grip ends the same as that of the initial state. Then the chemisorption pathway is located by choosing as two endpoints the initial physisorbed state and the final chemisorbed state, respectively. The elastic band consists of seven equally-spaced images of the system forming a discretization of the path between two fixed endpoints. The calculation is considered converged when the force on each image is less than 0.05eV/Å. A continuous MEP is then obtained by cubic polynomial interpolation of the calculated energies with the aid of the potential force projected in the direction of the path on each image [37]. Since the spacing between adjacent images is fixed during NEB relaxation, the transition state may not be coincident with any of the relaxed images. Hence we show the transition state in section 4.3 using the relaxed image closest to the saddle point. Additionally, for the same set of initial and final states, there may exist several competing reaction pathways corresponding to different reaction mechanisms. As the potential energy landscape evolves with stress, the same type of initial guess, e.g., straight line interpolation of intermediate images in configuration space between two endpoints, may relax to distinct pathways at different stress levels. Therefore, in order to follow the evolution of the same transition mechanism with an increase of stress, we first identify one pathway of interest at zero stress. Then, in all subsequent NEB searches at other stress levels, the final relaxed path from a previous search will be taken as an initial path for relaxation at a new stress level, but the system will be uniformly scaled according to the current extension of the nanorod. Since external loading is applied via the displacement-controlled method, during each NEB relaxation, we keep the positions of atoms at grip ends for all intermediate images fixed at that of the initial state. The tensile stress within the nanorod at the initial state is then taken as the nominal stress for the calculated activation energy barrier.

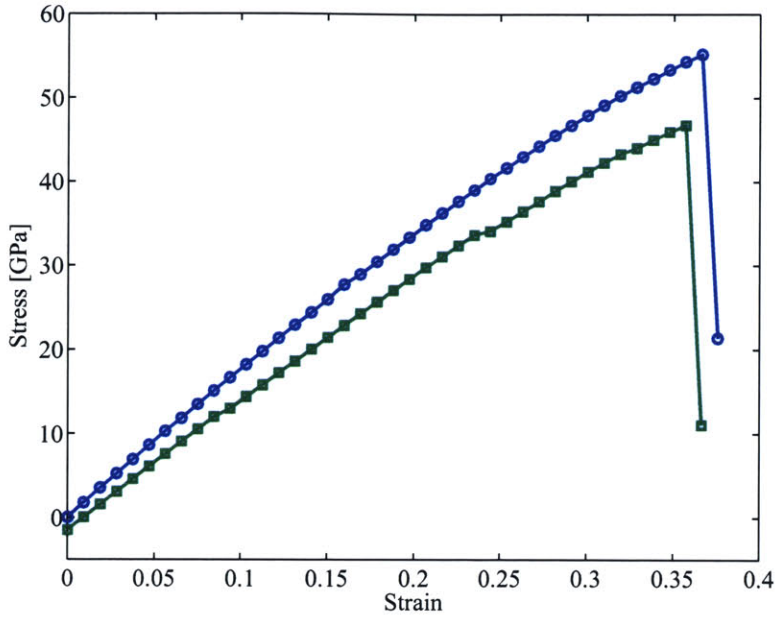


Figure 4-2: Stress vs. strain curves for uniaxial tension of the dry nanorod (the blue/black line) and hydrolyzed nanorod (the green/gray line).

4.3 Results

4.3.1 Uniaxial Tension

The stress-strain response for uniaxial tension of a nanorod without water is first calculated as shown in Fig. 4-2 by the blue line. It can be seen that the nanorod deforms linearly in the early stages. The corresponding tension modulus is about 183GPa. Then the stress-strain variation shows a slight nonlinearity up to the point of failure where the stress drops precipitously. The breaking stress at the onset of failure σ_{cr} is about 55.2GPa and the corresponding critical strain ϵ_{cr} is 0.36. The value of σ_{cr} represents nanorod's athermal mechanical strength, *i.e.*, the maximum resistance to fracture in the absence of thermal fluctuations. This athermal strength σ_{cr} will be used later as a strength-normalizing parameter for the calculated activation energy barriers at different stress levels. As the inter-ring siloxane bonds are the primary load carriers, we focus on geometric changes for a representative inter-ring siloxane bond located in the middle section of the nanorod. The calculated Si-O bond length immediately preceding bond-breaking is 1.827Å compared to 1.686Å at

the fully relaxed state. Correspondingly, the Si-O-Si bond angle changes from the initial value of 114.3° to 146.0°.

We also show in Fig. 4-2 the uniaxial stress-strain response of a hydrolyzed nanorod. The atomic structure of the hydrolyzed nanorod with silanol surface groups is shown in Fig. 4-4(d). Similarly, stress increases almost linearly with strain, but with a lower tension modulus of 154GPa. More importantly, there exists a compressive stress of about 1.44GPa within the hydrolyzed nanorod at zero strain. The origin of the residual stress at zero strain lies in the repulsive interaction between two fully relaxed silanol groups. As will be demonstrated later, the existence of this residual stress will lead to a term, which is linearly dependent on the applied stress, in the thermodynamic driving force of hydrolysis reaction.

4.3.2 Physisorption and Chemisorption

Physisorption involves formation of two hydrogen bonds with two bridging O_{br}s of the nanorod. We identify two states of physisorption. One corresponds to an asymmetric binding form with different hydrogen bond lengths as shown in Fig. 4-4(b). At zero stress, the calculated bond lengths are 1.75Å and 2.50Å, respectively. Another physisorbed state takes a symmetric binding form as shown in Fig. 4-6(b). Two hydrogen bond lengths at zero stress are both 1.79Å. The energy variation associated with physisorption is defined as hydration energy, E_{hydr}, of the nanorod [24]. That is,

$$E_{\text{hydr}} = E_{\text{phys-rod-wat}} - (E_{\text{rod}} + E_{\text{wat}}), \tag{4.1}$$

where $E_{\text{phys-rod-wat}}$ is the total energy of the nanorod with a physisorbed water molecule, E_{rod} is the energy of the nanorod without water, but at the same stress level of the hydrated nanorod, and E_{wat} is the self-energy of a water molecule. At zero stress, the calculated E_{hydr} for the asymmetric and symmetric binding forms are -0.57eV and -0.63eV, respectively. Comparing the two physisorbed states, it is evident that a lower hydration energy leads to stronger hydrogen bonding as in-

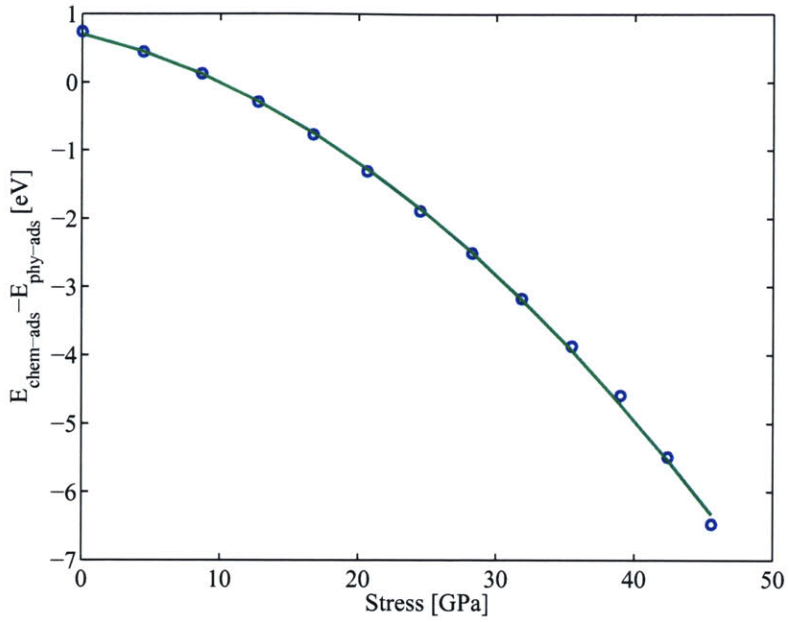


Figure 4-3: Stress-dependent chemisorption energy. The circles represent the calculated data points and the solid line corresponds to the polynomial fit given by Eq. (4.3).

dicated by a shorter hydrogen bond length. We also calculate the hydration energy as a function of stress and find that stress only affects hydration energies slightly for both physisorbed states. Take the asymmetrically adsorbed state as an example: when stress is increased to about 35GPa, the hydration energy only reduces by about 0.03eV down to -0.60eV. In addition, we also calculate the MEPs of physisorption and find that no saddle point exists along the MEPs irrespective of stress levels. Therefore, physisorption of a single water molecule to the silica nanorod is a thermodynamically favorable process without kinetic barrier.

In contrast, chemisorption involves formation of two silanol surface groups. At a given stress level, the relaxed structure of the hydrolyzed nanorod may take various forms with different arrangements of relative position of the two silanol groups. These geometrically similar states correspond to local minimums with very close energy levels on the potential energy surface. To examine the effect of stress, we focus on a representative type of relaxed structure as shown in Figs. 4-4(d)/4-6(f)/4-8(d). This representative structure of chemisorption will later be taken as the final state

to identify various pathways of hydrolysis. With an elongation of the hydrolyzed nanorod, the total energy of the system increases since more elastic energies are stored within the system. More importantly, as shown in Fig. 4-3, the chemisorption energy E_{chem} will change from positive values at low stress levels to negative ones at high stress levels. Here, E_{chem} is defined as the difference in total energy between the chemisorbed state $E_{\text{chem-rod-wat}}$ and the water-nanorod non-interacting state at the same stress level. That is,

$$E_{\text{chem}} = E_{\text{chem-rod-wat}} - (E_{\text{rod}} + E_{\text{wat}}). \tag{4.2}$$

This trend of increasing chemisorption energy with stress indicates that the chemisorption will evolve from a thermodynamically unfavorable process to a favorable one with an increase of mechanical tensile stress. Hence the simulation result about stress effect on chemisorption is consistent with the concept of stress-enhanced hydrolysis reaction from the thermodynamic perspective. Moreover, the functional dependence of chemisorption energy E_{chem} on stress can be revealed by a polynomial fit up to the second power of the nominal stress, σ. The solid line in Fig. 4-3 corresponds to the fitting formula given by

$$E_{\text{chem}} = 0.7004 - 0.0472\sigma - 0.0024\sigma^2 \tag{4.3}$$

where E_{chem} and σ are in units of eV and GPa, respectively. The physical meaning of the coefficients in Eq. (4.3) as well as implications for understanding the effect of stress on hydrolysis reaction will be discussed in section 4.4.

4.3.3 Reaction Pathways of Hydrolysis

In transition from the initial to final state of chemical dissociation, three distinct reaction mechanisms are identified with the characteristic initial processes as wa-

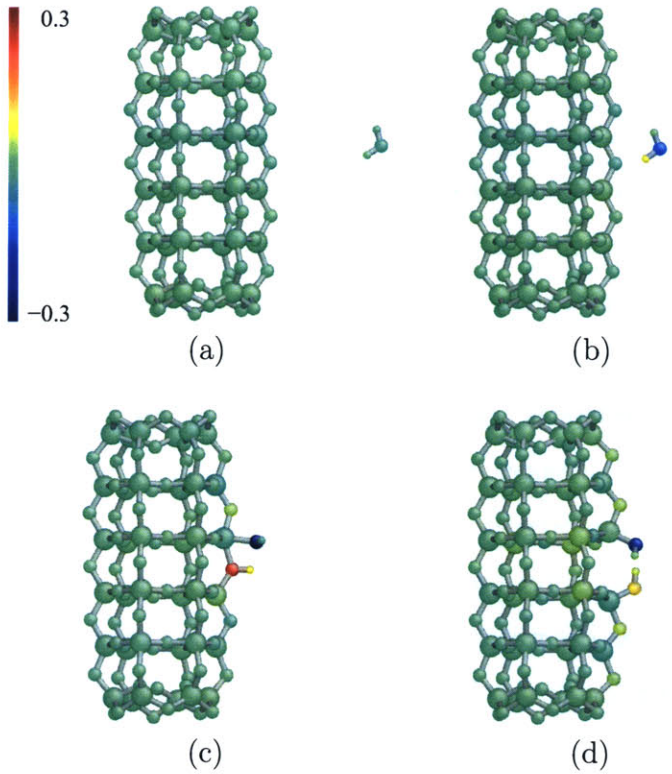


Figure 4-4: Atomic configurations along the transition pathway for mechanism I: (a) initial state, (b) physisorbed state, (c) saddle-point configuration, (d) final chemisorbed state.

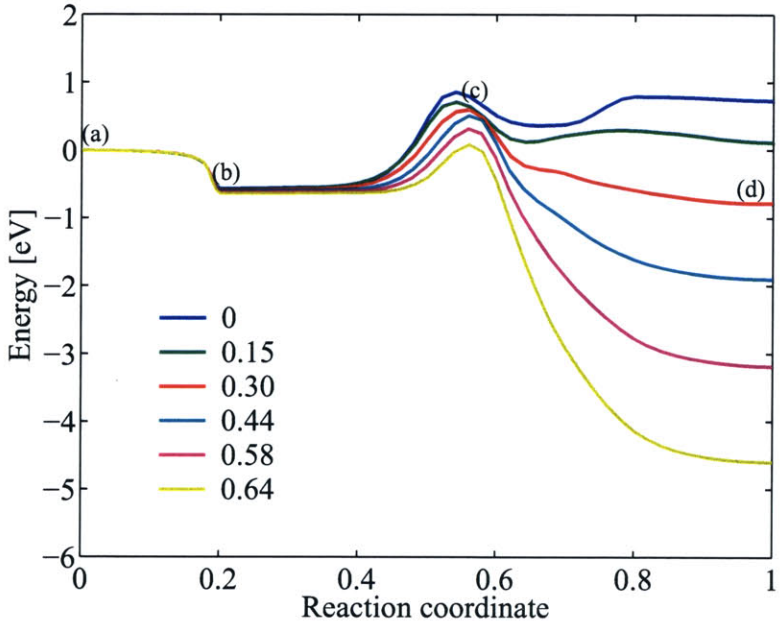


Figure 4-5: Stress-dependent minimum energy paths for mechanism I.

ter dissociation, water molecular chemisorption, and direct siloxane bond-breaking, respectively.

The first mechanism is characterized with the following sequence of reaction steps: (1) physisorption; (2) subsequent water dissociation which involves proton transfer and bond formation between Si-O_{wat}; and (3) final bond breaking between Si and bridging O_{br} of the siloxane bond. We show in Fig. 4-4 a series of snapshots along the reaction pathway at a stress level of 16.7GPa ($0.3\sigma_{cr}$). The corresponding location on the MEP for each frame is marked in Fig. 4-5. Atoms in Fig. 4-4 are color-coded according to the differences in atomic charge (Coulson charge calculated from MOPAC2000) between the present value and the one at the initial state. Thus, the charge redistribution during processes of bond breaking and formation can be revealed clearly. It is seen that, as the water molecule approaches the nanorod, physisorption first occurs with formation of two asymmetric hydrogen bonds as shown in Fig. 4-4(b). A small color variation indicates that there is only a slight charge transfer from in association with hydrogen bond formation. Then a concerted reaction occurs in which proton transfer to the bridging O_{br} is followed immediately with electron transfer from O_{wat} to Si atom. During this process, a significant charge variation is observed as shown in the saddle-point configuration of Fig. 4-4(c). Specifically, as the bridging O_{br} forms a new bond with the incoming proton, it loses electron and becomes positively charged. On the other hand, as Si atom forms a new bond with O_{wat}, it gains electrons and thereby becomes negatively charged. The reaction is finally accomplished by forming two new silanol groups as shown in Fig. 4-4(d).

The stress-dependent MEPs corresponding to the first reaction mechanism are shown in Fig. 4-5. As the nanorod is loaded with an incremental stretch of 0.5Å, the MEPs are parametrized by the corresponding nominal stresses, normalized by σ_{cr}. To facilitate comparison at different stress levels, we take the energy of the initial state as zero and plot energy variation relative to the initial state as a function of the reaction coordinate s. Here, the reaction coordinate is defined as follows: for any given extension of the nanorod, we take as $s = 0$ the initial state at which the nanorod and water are non-interacting and as $s = 1$ the final chemisorbed state.

Within a $3N$ configurational space, where N is the total number of free atoms within the system, the hyperspace arc length along the MEP between the initial state $\mathbf{R}_i$ and an intermediate state $\mathbf{R}$ is given by

$$l \equiv \int_{\mathbf{R}_i}^{\mathbf{R}} \sqrt{d\mathbf{R} \cdot d\mathbf{R}}. \tag{4.4}$$

Then, the normalized reaction coordinate s is defined as

$$s \equiv \frac{l}{l_f} \tag{4.5}$$

where l_f denotes the hyperspace arc length between the initial and final states. It is seen from Fig. 4-5 that, in reference to the initial state, the energies at saddle point as well as that at the final equilibrium state will decrease as the stress increases. The barrier height as a function of stress is plotted in Fig. 4-10 and will be compared with other mechanisms at the end of this section.

The second reaction mechanism is different from the first one in that a metastable, molecularly adsorbed state exists along the reaction pathway going from the physisorbed to the final hydrolyzed state. We show in Fig. 4-6 snapshots along the reaction pathway at stress of 16.7GPa ($0.3\sigma_{\text{cr}}$) and plot in Fig. 4-7 the calculated MEPs at different stress levels. Similarly, the corresponding location on the MEP for each frame in Fig. 4-6 is also marked in Fig. 4-7. The atomic configuration of molecularly adsorbed state which corresponds to a local minimum on the MEP is shown in Fig. 4-6(d). Note that this metastable state is absent on the MEP at zero stress and it starts to appear as the applied stress increases to about 5GPa. At this molecularly adsorbed state, Si atom adopts five-fold coordination and forms a new bond with O_{wat}. For example, the bond distance between Si and O_{wat} is 1.916Å at a stress of 16.7GPa. This value falls into the range of an empirical chemical bond length. Due to the existence of such a metastable state on the MEP, there are two energy barriers that need to be overcome in going from the initial state to the final state

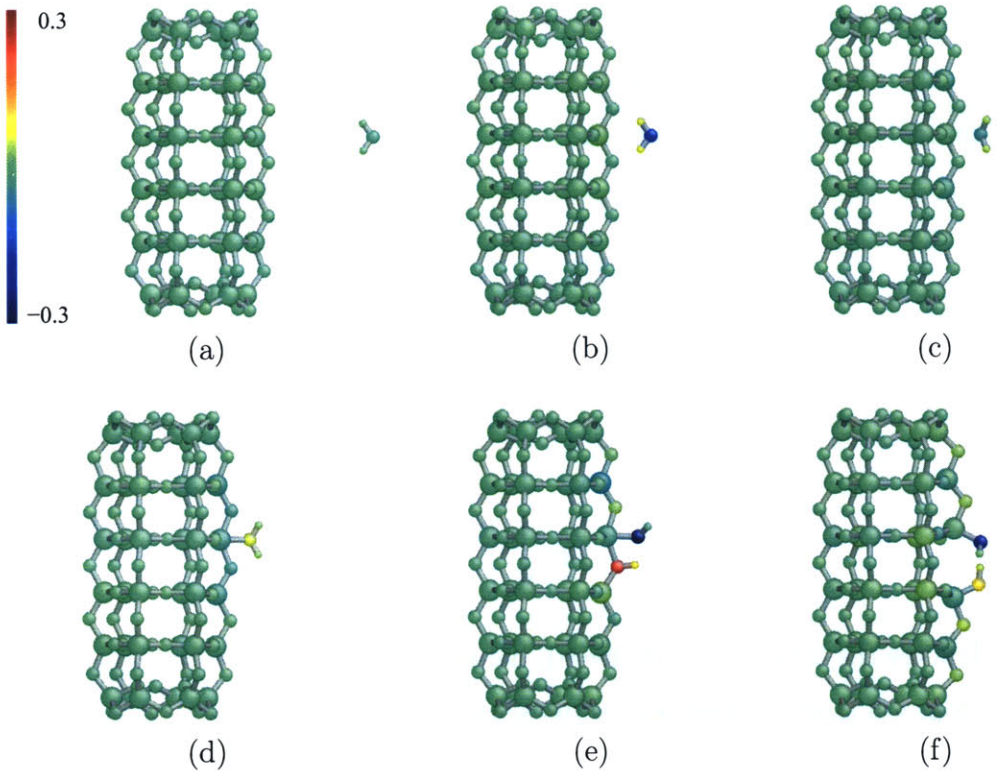


Figure 4-6: Atomic configurations along the transition pathway for mechanism II: (a) initial state, (b) physisorbed state, (c) the first saddle-point configuration, (d) metastable molecularly adsorbed state, (e) the second saddle-point configuration, (f) final chemisorbed state.

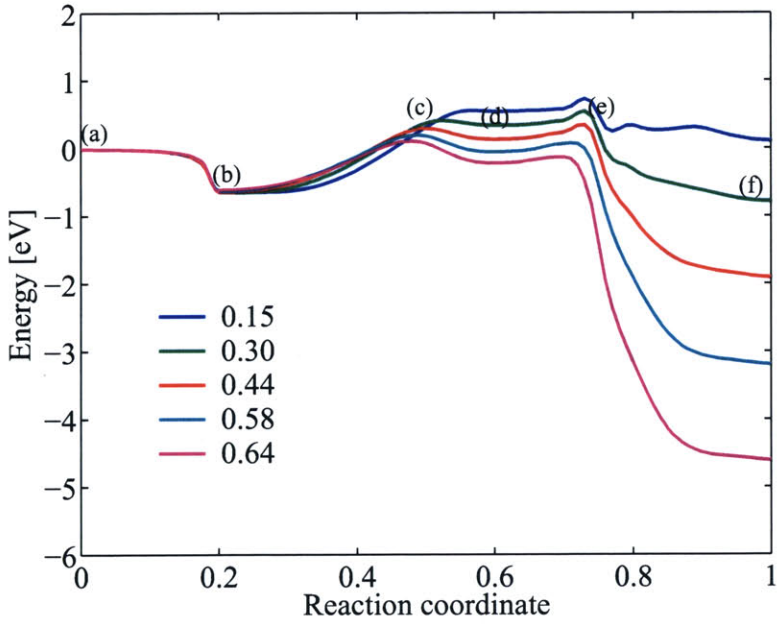


Figure 4-7: Stress-dependent minimum energy paths for mechanisms II.

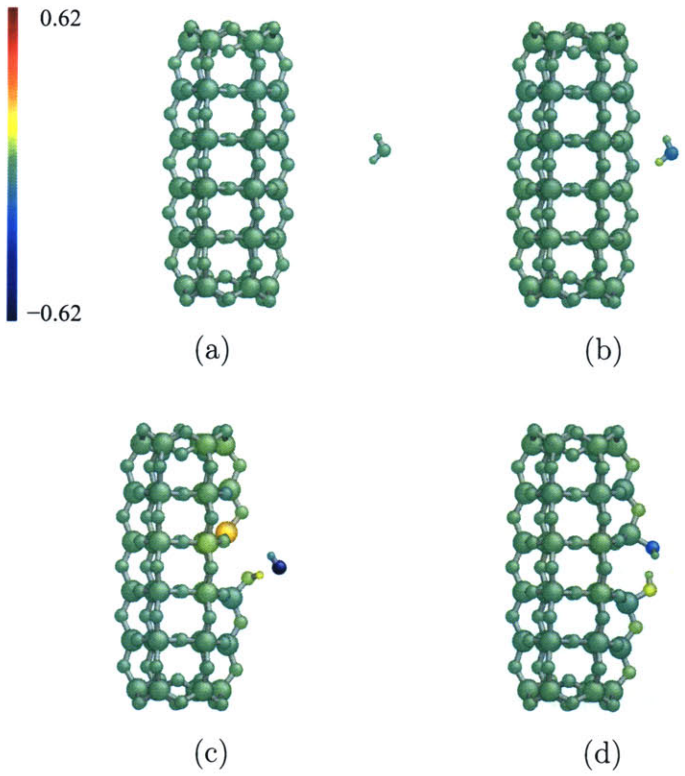


Figure 4-8: Atomic configurations along the transition pathway for mechanism III: (a) initial state, (b) physisorbed state, (c) saddle-point configuration, (d) final chemisorbed state.

of chemical dissociation. Two saddle-point configurations along the reaction pathway are shown in Fig. 4-6(c) and (e), respectively. The barrier height corresponding to each saddle point is also plotted in Fig. 4-10 as a function of stress. It is of interest to note that the two curves cross at stress about 27GPa. This crossing indicates that the rate limiting step for the second reaction mechanism will switch from the process of water molecule dissociation followed by Si-O_{wat} bond-breaking at low stresses to the process of water molecularly binding with penta-coordinated Si at high stresses. We will compare stress-dependent barrier heights with other mechanisms at the end of this section.

Lastly, a third reaction mechanism of hydrolyzing a strained nanorod is identified. This mechanism distinguishes itself from the other two in that the reaction first involves breaking a Si-O_{br} bond within the nanorod, and then terminating the dangling bond via the dissociated water molecule. Likewise, we show in Fig. 4-8 snapshots

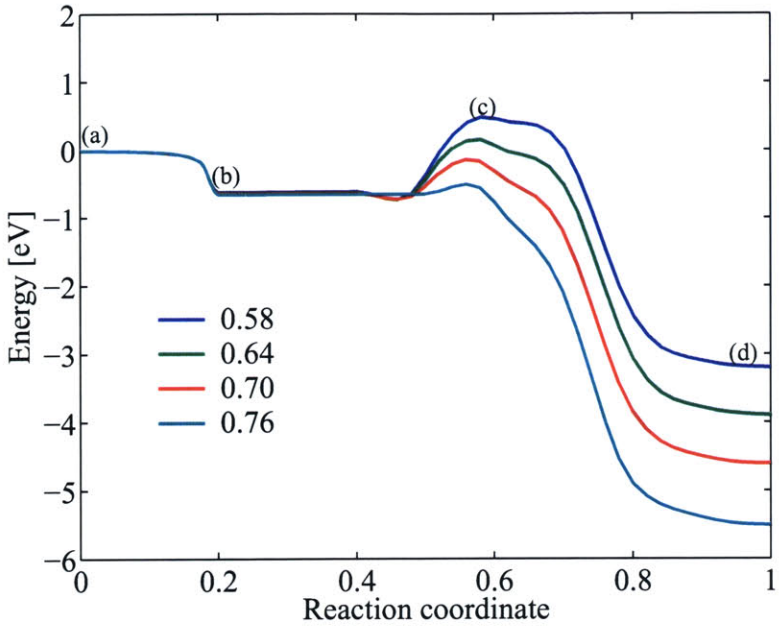


Figure 4-9: Stress-dependent minimum energy paths for mechanism III.

along the reaction pathway at stress of 31.8GPa ($0.58\sigma_{cr}$) and plot in Fig. 4-9 the calculated MEPs at different stress levels. The configuration close to the saddle point is shown in Fig. 4-8 (c). It can be seen that the dangling O_{br} bond is being terminated by a proton transferring from a dissociated water molecule. For the third mechanism, the energy variation has a stronger dependence on the deformation state of the nanorod. Hence, we plot different MEPs in Fig. 4-9 at smaller stress intervals.

Figure 4-10 compares the barrier height as a function of stress for the three different mechanisms of hydrolysis reaction. Evidently, tensile stress will reduce the activation energy barrier for any specific reaction mechanism. More importantly, as the relative barrier height of different mechanisms changes with an increase of stress, the switching of rate-limiting steps will occur either within one type of reaction pathway, as discussed earlier for the second reaction mechanism, or among different reaction mechanisms. Thus, combining the results on stress-mediated barrier height with those of the stress-dependent chemisorption energy, a complete picture of stress-dependent hydrolysis reaction kinetics can be revealed. At the stress level below about 20% of athermal strength σ_{cr}, the negative value of chemisorption en-

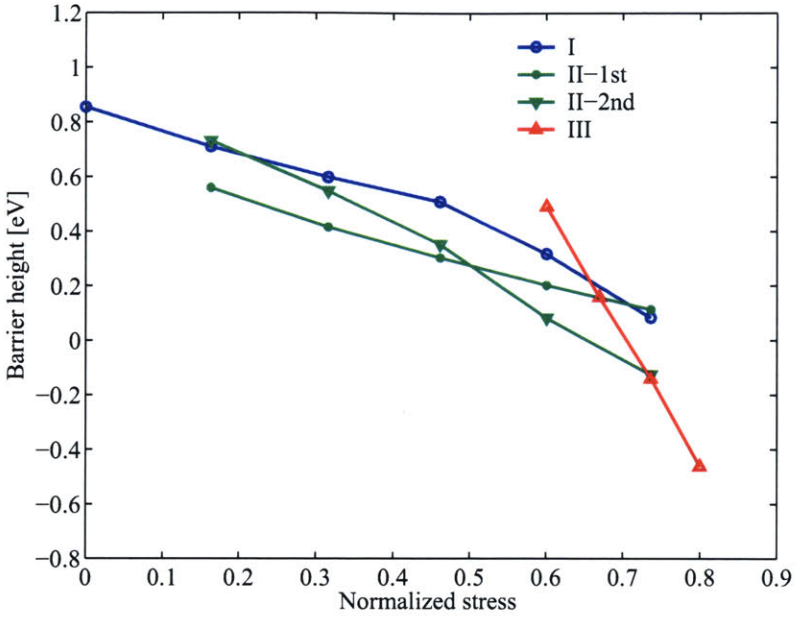


Figure 4-10: Stress-dependent activation energy barriers.

ergy indicates that the hydrolysis reaction is thermodynamically unfavorable. For the applied stress within $20 - 75\%$ of athermal strength, the second reaction mechanism will be the rate-limiting process. Note that there will be a switching of rate limiting step within the pathway at about 50% of athermal strength. Finally, as the applied stress increases to above 75% of athermal strength, the third reaction mechanism will have lowest activation energy barriers. However, the negative values of barrier height indicate that the reaction has become kinetically barrierless.

4.3.4 Bond Breaking in a Dry Nanorod

While the present study focuses on stress-dependent hydrolysis reaction, it is also of interest to compare the differences in breaking a siloxane bond with and without water. We show in Fig. 4-11 the MEPs of breaking a siloxane bond in a dry nanorod at two representative stress levels of 31.8GPa and 35.5GPa, respectively. It can be seen that, when stress is applied up to 31.8GPa, siloxane bond-breaking without water is still thermodynamically unfavorable, while hydrolysis reactions under the same stress level are all thermodynamically favorable irrespective of detailed

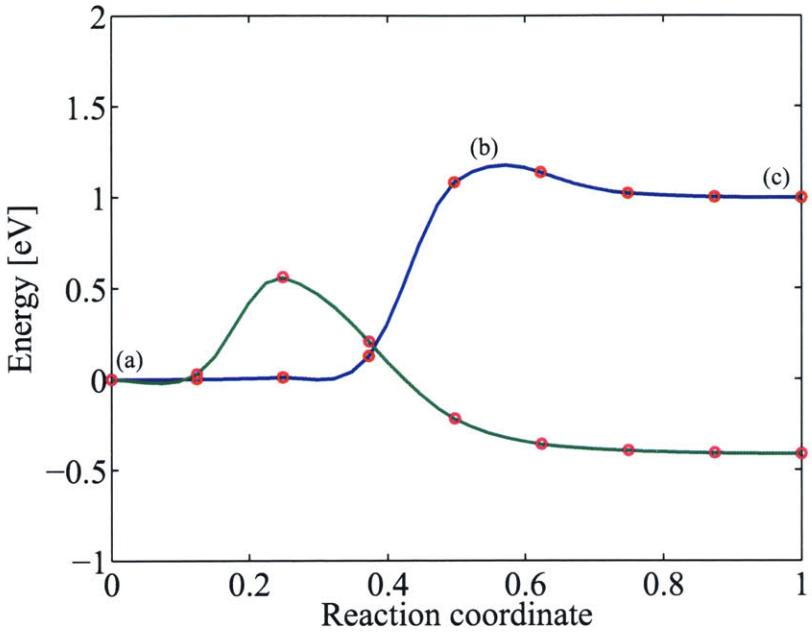


Figure 4-11: Minimum energy paths for breaking a siloxane bond in a dry nanorod. The blue and green curves correspond to the load levels at 31.8GPa and 35.5GPa, respectively.

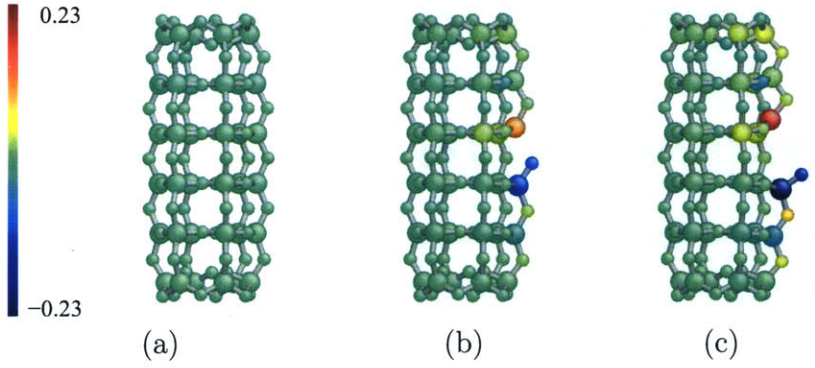


Figure 4-12: Snapshots along the transition pathway of siloxane bond breaking in the absence of water, the nominal stress is 35.5GPa: (a) initial state, (b) saddle point configuration, (c) final state.

transition mechanisms. Evidently, the presence of the water molecule facilitates siloxane bond-breaking. As the applied stress increases, the MEP of breaking a siloxane bond without water changes dramatically. When stress is increased to 35.5GPa, the corresponding MEP shown in Fig. 4-11 by the dashed line indicates that siloxane bond-breaking becomes thermodynamically favorable with an activation energy barrier of about 0.6eV. Figure 4-12 shows the images of atomic configurations along the MEP at stress 31.8GPa. It is seen that, at the final state, there is a significant charge variation in association with the broken siloxane bond. In contrast, the charge variation of the final hydrolyzed states shown in Figs. 4-4(d)/4-6(f)/4-8(d) are relatively small because the dangling siloxane bond is terminated by the dissociated water molecule.

4.4 Interpretations

As we focus on the molecular interpretations of energy variation along the MEPs in section 4.3, we consider next an elementary analysis of stress-dependence of the thermodynamic driving force and the kinetic barrier of chemical reaction using the nanorod as an example. To our knowledge, this is the first time that an explicit analysis of the stress/strain-dependent MEP is carried out and applied to analyze the atomistic simulation results, although a similar approach has been taken to study continuum inelastic deformation (*e.g.*, see Rice [78] and Argon [6]). More importantly, the present analysis helps to elucidate the crack extension driving force of brittle fracture for the situation in which the effect of surface relaxation is significant, such as stress corrosion by water. Specifically, we find that an important term, which is linearly dependent on the external loading, is commonly left out in formulating the crack extension driving force. Our simulation results indicate that this term may play an important role in determining the Griffith's critical condition of brittle fracture initiation. In the following discussions, we will mainly focus on analyzing the strain-dependent MEP, considering that the nanorod is loaded via the displacement-controlled method in our simulations. A brief discussion on the stress-dependent

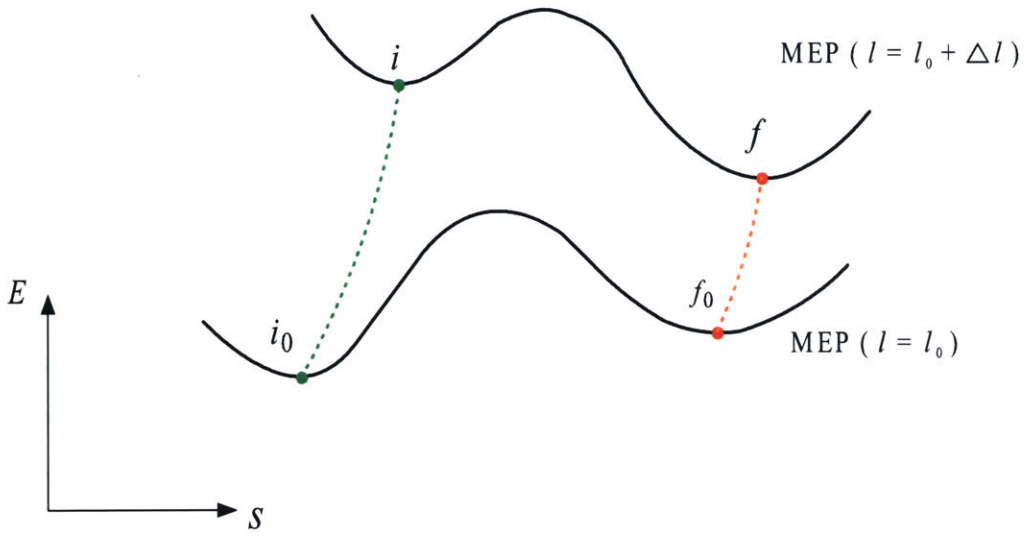


Figure 4-13: Schematics of the minimum energy paths at different load levels.

MEP will be given in section 4.4.4 to contrast the difference between two loading methods.

4.4.1 Strain-Dependent MEPs

Figure 4-13 schematically plots two MEPs corresponding to situations in which the nanorod is extended by zero and ΔL, respectively. When the nanorod is deformed, the total energy of the nanorod system will increase due to the accumulating elastic energy associated with mechanical deformation. Thus, the upper curve in Fig. 4-13 should represent the MEP of hydrolyzing a deformed nanorod, while the lower curve the MEP of hydrolyzing a undeformed nanorod. The local minima i on the MEPs correspond to the initial states in which the nanorod and single water molecule are non-interacting; and the local minima f correspond to the final chemisorbed states. The x-axis in Fig. 4-13 represents a generically defined reaction coordinate, i.e., the locations of the initial/saddle-point/final states may drift along the MEP at different load levels.

Prior to analyzing energy variation along the MEPs, we first specify the scheme of labeling states on the MEPs at different load levels. Here, we follow the most

intuitively direct labeling scheme as given in section 4.3, while an alternative scheme will be discussed in section 4.4.4. Recall that, in section 4.3, a normalized reaction coordinate s is defined, as given by Eq. (4.5), such that $s \in [0,1]$ labels all states connecting state i to state f on the MEP for any given nanorod extension. According to this definition, the initial and final states are fixed at $s = 0$ and $s = 1$, respectively. But the reaction coordinate of the saddle point may drift along the MEP at different load levels.

For any frozen state s, the energy of the nanorod system at strain ϵ can be related to that at $\epsilon = 0$ by

$$E(\epsilon, s) = E(0, s) + \sigma_0(s)\epsilon V + \frac{1}{2}k_0(s)\epsilon^2 V + \mathcal{O}(\epsilon^2) \tag{4.6}$$

where $V = A_0 L_0$ is the nominal volume of the nanorod, $\sigma_0(s)$ is the residual stress within the nanorod at $(\epsilon = 0, s)$, and $k_0(s)$ is the nanorod stiffness at $(\epsilon = 0, s)$. In Eq. (4.6), we keep terms to second-order accuracy in ϵ based on the fact that, as shown in Fig. 4-2, both the initial ($s = 0$) and final states ($s = 1$) exhibit nearly linear stress-strain responses up to the athermal strength limit. Hence, Eq. (4.6) represents a reasonable approximation to the energy variation for both the initial and final states. However, caution has to be taken in using Eq. (4.6) for any intermediate state s, because the effective range, with which Eq. (4.6) is a good representation of energy variation, may reduce. In addition, we note that, since the initial state at $\epsilon = 0$ is stress free, the relation $\sigma_0(0) = 0$ holds.

Denote $\Delta E(\epsilon, s)$ as the energy change in transition from the initial state i to an intermediate state s, while keeping extension of the nanorod fixed. From Eq. (4.6), one obtains

$$\begin{aligned} \Delta E(\epsilon, s) &\equiv E(\epsilon, s) - E(\epsilon, 0) \\ &= \Delta E(0, s) + \sigma_0(s)\epsilon V + \frac{1}{2}\Delta k_0(s)\epsilon^2 V + \mathcal{O}(\epsilon^2). \end{aligned} \tag{4.7}$$

where $\Delta k_0(s)$ denotes the stiffness change from state i to s, i.e., $\Delta k_0(s) \equiv k_0(s) - k_0(0)$. It is seen from Eq. (4.7) that the energy variation going from state i to s at a fixed ϵ can be partitioned into strain-independent and strain-dependent contributions. The former, as represented by the term $\Delta E(0, s)$, is the energy difference between state s and i at zero strain. For the strain-dependent part, the linear term in ϵ arises from the residual stress $\sigma_0(s)$ in association with a partially hydrolyzed nanorod at zero strain, while the quadratic term in ϵ results from the change of nanorod stiffness when a siloxane bond is partially broken.

4.4.2 Thermodynamic Driving Force

To correlate the above general analysis of energy variation along the MEP with simulation results, we consider the energy difference between two thermodynamically equilibrated states i and f. As $\Delta E(\epsilon, 1)$ represents the energy increase in transition from the initial to final state, its negative value is commonly defined as the thermodynamic driving force of transition, denoted by $D(\epsilon)$. That is,

$$\begin{aligned} D(\epsilon) &\equiv -\Delta E(\epsilon, 1) \\ &= -\Delta E(0,1) - \sigma_0(1)\epsilon V - \frac{1}{2}\Delta k_0(1)\epsilon^2 V \\ &= -\Delta E(0,1) - \frac{\sigma_0(1)}{k_0(0)}\sigma V - \frac{1}{2}\frac{\Delta k_0(1)}{[k_0(0)]^2}\sigma^2 V \end{aligned} \tag{4.8}$$

where $\Delta E(0, 1)$ is the energy variation associated with reaction at zero strain. This strain-independent energy term is greater than zero for hydrolysis reaction. Hence, it can be regarded as the internal resistance to reaction. As two newly formed silanol surface groups repel against each other, the residual stress $\sigma_0(1)$ is compressive, i.e., $\sigma_0(1) < 0$, when we hold the two ends of the nanorod fixed. In addition, the nanorod stiffness will decrease, i.e., $\Delta k_0(1) < 0$, when a strongly connected siloxane bond is substituted with a pair of weakly interacted silanol groups. Hence, in the second step of Eq. (4.8), both the linear and quadratic terms in ϵ will contribute to increase

the driving force $D(\epsilon)$. The third step in Eq. (4.8) is given to facilitate correlation with the fitting formula of Eq. (4.3), which is expressed in terms of the nominal stress $\sigma = k_0(0)\epsilon$. Comparing with Eq. (4.8), the physical meaning of each term in Eq. (4.3) becomes transparent. In addition, combining Eq. (4.8) and Eq. (4.3), we can estimate values of both the residual stress $\sigma_0(1)$ and the nanorod stiffness change $\Delta k_0(1)$, once the initial stiffness $k_0(0)$ is known. For the present nanorod system, the initial value of $k_0(0) = 183$GPa leads to estimates of $\sigma_0(1) = -1.43$GPa and $\Delta k_0(1) = -27$GPa, comparing with direct calculation results of $\sigma_0(1) = -1.5$GPa and $\Delta k_0(1) = -29$GPa.

To prepare for the discussion on stress-dependence of the crack extension driving force in section 4.4.3, we also calculate the critical condition when the thermodynamic driving force of hydrolysis reaction vanishes. In other words, we need to determine the critical load above which the forward transition of bonding-breaking starts to be thermodynamically favorable. From Eq. (4.3), the critical load σ is solved to be 10.2GPa. To demonstrate the importance of linear stress dependence, we intentionally drop the linear term in Eq. (4.3) and then recalculate the critical load. This corresponds to the situation in that the accurate stiffness change has been estimated, but the effect of linear stress term is not taken into account. Then the critical value of σ becomes 17.1GPa. Hence, there is an approximately 68% difference in the estimated critical load if the linear stress-dependent term is ignored.

4.4.3 Implications for Brittle Fracture Analysis

The analysis of energy change associated with nanorod bond-breaking can help elucidate the thermodynamic driving force of brittle fracture. We first give a brief review of the concept of lattice trapping in the atomic theory of brittle fracture [100, 55]. Consider a simplified model system with an atomically sharp crack embedded in an otherwise perfect lattice. The quasi-static crack growth corresponds to a sequence of localized bond-breaking process at the instantaneous crack tip. Under a given external loading, the total energy of the system can be written as a function of crack length. Because of the lattice discreteness, the energy landscape fluctuates at the atomic scale along the crack length direction. As a result, crack extension may cease

temporarily when the system is trapped in a energy well. Thus, an activation energy is required to overcome the local energy barrier. Each time the crack extends by one lattice spacing, the system will move from one state of local energy minimum to an adjacent one. The energy difference between the two local energy-minimum states determines the instantaneous driving force of crack extension and the local maximum in between gives the activation energy barrier, which determines the instantaneous kinetic rate of crack growth.

A connection between the concept of lattice trapping at a crack tip and that of bond-breaking in a nanorod can be established if we consider breaking a siloxane bond in a perfect nanorod as effectively creating a small notch crack. Thus, the targeted siloxane bond can be also thought as being lattice-trapped. Correspondingly, the bond-breaking driving force given by Eq. (4.8) will represent the thermodynamic driving force of creating a notch crack. Although the detailed atomic structure and stress state are different for two situations, we believe that the general trend revealed via the study on nanorod bond-breaking is instructive in understanding the hydrolysis reaction at the crack tip, considering the discrete, highly localized nature of covalent bond breaking involved.

Now suppose a single siloxane bond occupies an effective area Δa, we may write the crack extension driving force as

$$D(\epsilon) = [\mathcal{G}(\epsilon) - \gamma]\Delta a, \tag{4.9}$$

where the strain-dependent terms of energy change are lumped together to give the elastic energy release rate $\mathcal{G}(\epsilon)$

$$\mathcal{G}(\epsilon) = -\frac{\sigma_0(1)\epsilon V + \frac{1}{2}\Delta k(1)\epsilon^2 V}{\Delta a}, \tag{4.10}$$

and the strain-independent energy variation corresponds to the surface energy increase rate γ, i.e.,

$$\gamma = \frac{\Delta E(0,1)}{\Delta a}. \tag{4.11}$$

It is well known that, for a Griffith crack within a linear elastic continuum, the energy release rate $\mathcal{G}(\epsilon)$ is only a function of ϵ^2. But the energy release rate $\mathcal{G}(\epsilon)$ given in Eq. (4.10) contains a linear term in ϵ. We have demonstrated in the case of nanorod bond-breaking that this linear term is important in determining the critical condition at which the forward transition of bonding-breaking starts to become thermodynamically favorable. Recall that, under the fixed-displacement boundary condition, this linear term arises from the residual stress, which is generated in association with relaxation of two newly-formed silanol surface groups at zero strain. As will be demonstrated in section 4.4.5, under the fixed-force boundary condition, there is also a corresponding linear term in the stress that should be taken into account. However, this linear term is absent in the conventionally defined driving force of crack extension. The reason is that it is a common practice to assign an energy penalty, *i.e.*, surface energy, to associate with creation of a fresh surface. However, the corresponding relaxation at the newly created surface are commonly considered to be negligible. Consequently, their effects in coupling with the applied external load to change the system's energy, *i.e.*, the thermodynamic driving force of quasi-static crack extension, are ignored. Therefore, in applying the Griffith condition to determine whether a crack will grow or heal in a surface reactive environment, caution has to be taken in formulating the correct form of the thermodynamic driving force. This is particularly important when surface relaxation involves spatial accommodation of foreign species.

4.4.4 Extremum Point Drift

In this subsection, we provide an alternative scheme of labeling various states on the MEPs at different load levels. For this new scheme, the reaction coordinates of both initial and final states will not be fixed. The only requirement is that, for

states with the same reaction coordinate s on the MEPs at different load levels, the instantaneous tensile modulus must be the same. Accordingly, an extremum point may drift along the MEP at different load levels, and there will be a one-to-one correspondence between the reaction coordinate of the initial/saddle-point/final state and applied external loading. We note that introduction of such a scheme is motivated by a well-known result in fracture mechanics, which can be readily derived using the J-integral (*e.g.*, see Rice [79]). That is, there exists a one-to-one correspondence between the crack-tip bond opening displacement, δ, and the applied far-field load, which can be expressed in terms of elastic energy release rate $\mathcal{G}$. If we take the crack-tip opening displacement as the reaction coordinate[1], the relation between δ and $\mathcal{G}$ clearly indicates that the initial state will not stay at a fixed reaction coordinate.

Since the initial, saddle-point and final states all belong to the category of the extremum points on the MEPs, we next derive a general relation of the extremum-point drift, under the only assumption that the instantaneous tensile modulus of the nanorod at a frozen state of s is constant. Denote s_0^* as the reaction coordinate of one extremum point on the MEP at zero strain. For state s near s_0^* on the same MEP, the energy is given by

$$E(0,s) = E(0,s_0^*) + \frac{1}{2}\kappa_0(s_0^*)(s-s_0^*)\epsilon^2 V + \mathcal{O}(\epsilon^2). \tag{4.12}$$

where $\kappa_0(s)$ denotes the local curvature at s and it is given by

$$\kappa_0(s) = \frac{\partial^2 E(0,s)}{\partial s^2}. \tag{4.13}$$

Consider state s on the MEP at strain ϵ. Note that Eq. (4.6) is still valid to

[1]The definition of reaction coordinate using crack-tip opening displacement is only approximately valid because the reaction of bond-breaking may proceed along a rather complicated pathway which can not be simply characterized by crack-tip bond distance. However, for the simple situation involving highly localized bond breaking in vacuum, the distance of a pair of crack-tip bond may be sufficient to be taken as a reaction coordinate.

connect energies between states at different MEPs with the same reaction coordinate s. Substitution of Eq. (4.12) into Eq. (4.6), one relates the energy at the point (ϵ, s) to that at $(0, s_0^*)$ by

$$E(\epsilon, s) = E(0, s_0^*) + \frac{1}{2}\kappa_0(s_0^*)(s - s_0^*)\epsilon^2 V + \sigma_0(s)\epsilon V + \frac{1}{2}k_0(s)\epsilon^2 V. \tag{4.14}$$

Denote s^* as the saddle-point state on the MEP at strain ϵ. That is,

$$\frac{\partial E(\epsilon, s)}{\partial s}|_{s=s^*} = 0. \tag{4.15}$$

Substitution of Eq. (4.14) into Eq. (4.15), one obtains

$$\kappa_0(s_0^*)(s^* - s_0^*) + \sigma_0'(s^*)\epsilon + \frac{1}{2}k_0'(s^*)\epsilon^2 = 0. \tag{4.16}$$

where prime denotes the derivative with respect to s. From Eq. (4.16), one obtains

$$s^* = s_0^* - \frac{\sigma_0'(s_0^*)}{\kappa_0(s_0^*)}\epsilon + \mathcal{O}(\epsilon). \tag{4.17}$$

Eq. (4.17) gives the first order estimate on the extremum-point drift based on information at s_0^* on the MEP of zero strain. Substitution of Eq. (4.17) into Eq. (4.14), one obtains the relation between the extremum-point energy at strain ϵ and that at $\epsilon = 0$,

$$E(\epsilon, s^*) \;=\; E(0, s_0^*) + \sigma_0(s_0^*)\epsilon V + \frac{1}{2}\left[k_0(s_0^*) - \frac{\sigma_0'^2(s_0^*)}{\kappa_0(s_0^*)}\right]\epsilon^2 V + \mathcal{O}(\epsilon^2). \tag{4.18}$$

Evidently, the extremum-point drift along the MEP contributes to energy variation on the second order in ϵ. Denote s_0^{ini} and s_0^{sad} as the reaction coordinate of the initial

and saddle-point states on the MEP of zero strain, respectively. From Eq. (4.18), the activation energy $\Delta E(\epsilon)$ at strain ϵ is related to the activation energy ΔE_0 at zero strain by

$$\Delta E(\epsilon) = \Delta E_0 + \Delta\sigma_0 \epsilon V + \frac{1}{2}\left[\Delta k_0 + \frac{\sigma'^2_0(s_0^{\mathrm{sad}})}{\kappa_0(s_0^{\mathrm{sad}})} - \frac{\sigma'^2_0(s_0^{\mathrm{ini}})}{\kappa_0(s_0^{\mathrm{ini}})}\right]\epsilon^2 V + \mathcal{O}(\epsilon^2). \qquad (4.19)$$

where $\Delta\sigma_0$ denotes the difference in stress between the saddle-point and initial states on the MEP of zero strain, Δk_0 the corresponding difference in the system's stiffness. From Eq. (4.19), the activation energy $\Delta E(\epsilon)$ at strain ϵ can be estimated using information obtained from the MEP of zero strain. Practically speaking, we have to make a few more calculations beyond the MEP of zero strain in order to estimate the higher order quantities such as the system's stiffness. Nevertheless, this perturbative analysis reveals the physical origin of external load-dependent terms in modifying kinetic barriers of reaction.

4.4.5 Stress-Dependent MEP

Although the above analysis is given using the nanorod's strain as a control variable, the procedure of derivation and results about the strain-dependent thermodynamic driving force and activation barrier can be easily adapted to the stress-controlled situation. When the nanorod is deformed under stress control, the total energy of the system, $G(\sigma, s)$, involves contributions from both atomic configuration energy and potential energy of the external stress, i.e., $G = E - \sigma\epsilon V$. Following the same scheme of labeling various states on the MEPs as that discussed in section 4.4.1, in parallel to Eq. (4.6), the energy at stress σ is related to that at zero stress by

$$\Delta G(\sigma, s) = \Delta G(0, s) + \epsilon_0(s)\sigma V - \frac{1}{2}\Delta c_0(s)\sigma^2 V.$$

where $\epsilon_0(s)$ is the strain within the nanorod due to the hydrolysis reaction at zero stress, $\Delta c_0(s)$ denotes the compliance change from state i to s. That is, $\Delta c_0(s) \equiv c_0(s) - c_0(0)$, where $c_0(s)$ is the compliance of the nanorod at an intermediate state s and it is the inverse of the stiffness $k_0(s)$.

The connection between two loading conditions can be seen when Eq. (4.7) is rewritten in terms of the nominal stress σ, where σ is related to the nominal strain by $\sigma = k_0(0)\epsilon$. That is,

$$\Delta E(\epsilon, s) = \Delta E(0, s) + \frac{\sigma_0(s)}{k_0(s)} \frac{k_0(s)}{k_0(0)} \sigma V - \frac{1}{2} \frac{k_0(s)}{k_0(0)} \Delta c_0(s) \sigma^2 V. \tag{4.20}$$

Comparing Eq. (4.20) with Eq. (4.7), it is seen that if the change in stiffness at state s is small with respect to the initial stiffness $k_0(0)$, i.e., $k_0(s)/k_0(0) \sim 1$, Eq. (4.20) and Eq. (4.7) will be equivalent, because the relations $\epsilon_0(s) = \sigma_0(s)/k_0(s)$ and $\Delta E(0, s) = \Delta G(0, s)$ always hold. Therefore, as a first approximation, the energy variation associated with bonding breaking is independent of the type of boundary conditions imposed. This is reminiscent of the well known fact in fracture mechanics: The expression of energy release rate for crack extension is independent of loading methods. However, for a finite size system such as a nanorod, there are only six siloxane bonds intersecting each cross section plane of the nanorod. Breaking a single bond will approximately reduce the effective tension stiffness of nanorod by one sixth. Because of this significant stiffness change, the effect of different loading methods is expected to be appreciable. In the future, it would be of interest to quantitatively evaluate the size effect on the stress-dependence of thermodynamics and kinetics of chemical reaction revealed from the present study.

Chapter 5

Summary

In this thesis, the unit processes of configurational instability of crystal lattices under various non-uniform structural and/or chemical environments are studied by systematically probing the atomistic potential energy landscape of each system using the state of the art configurational space sampling schemes. The main results are summarized as follows.

In chapter 2, homogeneous nucleation of a dislocation beneath a nanoindenter is studied as a strain localization event triggered by elastic instability of the perfect crystal at finite strain. The finite element calculation, with a hyperelastic constitutive relation based on an interatomic potential, is employed as an efficient method to simulate the process of nanoindentation. This implementation facilitates the study of dislocation nucleation at length scales that are large compared to atomic dimensions, while remaining faithful to the nonlinear interatomic interactions. An local instability criterion based on bifurcation analysis is incorporated into the finite element calculation to predict homogeneous dislocation nucleation. This criterion is superior to that based on the critical resolved shear stress in terms of its accuracy of prediction for both the nucleation site and the slip character of the defect. Finite element calculations of nanoindentation of single crystal copper by a cylindrical indenter and predictions of dislocation nucleation are validated by comparing with direct molecular dynamics simulations governed by the same interatomic potential. In addition, analytic 2D and 3D linear elasticity solutions based on the Stroh formalism are used to

benchmark the finite element results. The critical configuration of homogeneous dislocation nucleation under a spherical indenter is quantified with full 3D finite element calculations. The prediction of the nucleation site and slip character is verified by direct molecular dynamics simulations. The critical stress state at the nucleation site obtained from the interatomic potential is in quantitative agreement with *ab initio* density functional theory calculation.

In chapter 3, we first report on the atomistic calculation of the saddle-point configuration and activation energy for the nucleation of a 3D dislocation loop from a stressed crack tip in single crystal Cu. The transition state is found using reaction pathway sampling schemes, the nudged elastic band and dimer methods. For the $\{111\}\langle\bar{1}10\rangle$ crack, loaded typically at 75% of the athermal critical strain energy release rate for spontaneous dislocation nucleation, the calculated activation energy is 1.1eV, significantly higher than the continuum estimate. The implication for such a high nucleation barrier is that experimental observation of homogeneous nucleation is unlikely on the laboratory time scale. On the other hand, we expect that the presence of any heterogeneity along the crack front can significantly reduce the activation energy for dislocation nucleation. For any specified heterogeneity, the atomistic method presented in this work should be applicable. In the second part of this chapter, we take up the atomistic calculation of 3D lattice trapping barriers to brittle fracture of Si. The system studied is a (111) cleavage crack with a straight crack front along the $[1\bar{1}0]$ and $[11\bar{2}]$ directions, respectively. The activation processes of crack front bond breaking are quantitatively characterized using a reaction pathway sampling scheme, the nudged elastic band method. The calculated energy variations along transition pathways reveal the energetics of 3D crack front kink-pair formation and migration. We find that the physical origin of directional anisotropy in cleavage crack propagation can be attributed to a difference in the kink-pair formation energy for different crack orientations. In addition, the effect of interatomic potentials on results is studied by comparing the Stillinger-Weber (SW) potential and the environment dependent interatomic potential (EDIP).

In chapter 4, stress-corrosion processes of silica with water are studied by system-

atically exploring the stress-dependent potential energy surfaces computed quantum mechanically at the molecular orbital level. A bond-saturated silica nanorod with clearly defined nominal tensile stress is constructed to model a structural unit of the stressed crack tip. Three competing pathways of hydrolysis reaction are identified with the characteristic initial processes as water dissociation, pentavalent chemisorption, and direct siloxane bond-breaking, which are the rate-limiting steps at low, intermediate, and high stress levels, respectively. From our calculation results, important aspects concerning the stress-dependent chemical reactivity are revealed. Of particular note is the linear stress dependence of the thermodynamic driving force, which was not commonly considered in the Griffith's criterion of brittle fracture initiation. Its physical origin lies in the coupling between the applied stress and volumetric expansion of the atomic structure due to surface relaxation associated with bond breaking. Our nanorod example quantitatively demonstrates that this linear stress-dependent driving force is particularly important when the bond-breaking process involves space accommodation of foreign species. Furthermore, we carry out a perturbative analysis on the stress-mediated minimum energy paths, from which the physical origin of the stress dependence of both thermodynamic driving force and kinetic energy barrier is revealed.

Appendix A

Stroh Formalism of Anisotropic Elasticity

Basic equations for Stroh formalism, which are used for deriving 2D and 3D analytic indentation solutions, are given in this appendix [102]. A general solution for displacement $\mathbf{u}$ and stress function Φ is given by

$$\mathbf{u} = \mathbf{A}\mathbf{f}(z) + \overline{\mathbf{A}\mathbf{f}(z)} \tag{A.1}$$

$$\Phi = \mathbf{B}\mathbf{f}(z) + \overline{\mathbf{B}\mathbf{f}(z)} \tag{A.2}$$

where

$$\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3] \tag{A.3}$$

$$\mathbf{B} = [\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3] \tag{A.4}$$

$$\mathbf{f} = [f_1(z_1)\ f_2(z_2)\ f_3(z_3)]^{\mathrm{T}} \tag{A.5}$$

$$z_\alpha = x + p_\alpha y, \quad \alpha = 1, 2, 3 \tag{A.6}$$

$\mathbf{f}$ is a complex function vector to be determined by satisfying the boundary conditions of the problem considered. Given the elastic constants tensor C_{ijkl} which is expressed in the global coordinates, p_α and $\mathbf{a}_\alpha$ can determined by solving the following eigenvalue problem

$$\left[\mathbf{Q} + (\mathbf{R} + \mathbf{R}^{\mathrm{T}})p + \mathbf{T}p^2\right] \mathbf{a} = 0 \tag{A.7}$$

where

$$Q_{ik} = C_{i1k1}, \; R_{ik} = C_{i1k2} \text{ and } T_{ik} = C_{i2k2} \tag{A.8}$$

The non-zero solution of $\mathbf{a}$ requires that

$$\det\left[\mathbf{Q} + (\mathbf{R} + \mathbf{R}^{\mathrm{T}})p + \mathbf{T}p^2\right] = 0 \tag{A.9}$$

Equation (A.9) has six roots which form three conjugate pairs. The root p_α has a positive imaginary part, and thus $\mathbf{b}_\alpha$ is obtained by

$$\mathbf{b}_\alpha = \left(\mathbf{R}^{\mathrm{T}} + p_\alpha \mathbf{T}\right) \mathbf{a}_\alpha \tag{A.10}$$

The stresses are given by

$$\{\sigma_{2j}\} = 2\mathrm{Re}\sum_{\alpha=1}^{3} \mathbf{b}_\alpha f'_\alpha(z_\alpha) \tag{A.11}$$

$$\{\sigma_{1j}\} = -2\mathrm{Re}\sum_{\alpha=1}^{3} \mathbf{b}_\alpha p_\alpha f'_\alpha(z_\alpha) \tag{A.12}$$

The impedance matrix $\mathbf{M}$ appears in the 2D indentation solution. It is defined by

$$\mathbf{M} = -i\mathbf{B}\mathbf{A}^{-1} \tag{A.13}$$

$\mathbf{M}$ is a Hermitian matrix and independent of how $\mathbf{a}_\alpha$ and $\mathbf{b}_\alpha$ are normalized. The inverse of impedance matrix $\mathbf{M}$ is also commonly used [94] and given by

$$\mathbf{Y} = i\mathbf{A}\mathbf{B}^{-1} = \mathbf{M}^{-1} \tag{A.14}$$

For isotropic material, $\mathbf{Y}$ is

$$\mathbf{Y} = \begin{bmatrix} \frac{1-\nu}{\mu} & i\frac{1-2\nu}{2\mu} & 0 \\ -i\frac{1-2\nu}{2\mu} & \frac{1-\nu}{\mu} & 0 \\ 0 & 0 & \frac{1}{\mu} \end{bmatrix}. \tag{A.15}$$

where μ denotes shear modulus and ν Possion's ratio. The Barnett-Lothe tensor $\hat{\mathbf{L}}$ is employed in deriving the 3D indentation solution. It is defined by

$$\hat{\mathbf{L}} = -2i\mathbf{B}\mathbf{B}^{\mathrm{T}} \tag{A.16}$$

where the vectors $\mathbf{a}_\alpha$ and $\mathbf{b}_\alpha$ are normalized by

$$2\mathbf{a}_\alpha \cdot \mathbf{b}_\alpha = 1 \tag{A.17}$$

Bibliography

[1] *ABAQUS/Explicit Reference Manuals.* Hibbitt, Karlsson and Sorensen Inc., Pawtucket, R.I., 2001.

[2] R. Abeyaratne, K. Bhattacharya, and J. K. Knowles. Strain-energy functions with multiple local minima: Modeling phase transformations using finite thermo-elasticity. In Y. Fu and R. W. Ogden, editors, *Nonlinear Elasticity: Theory and Applications*, pages 433 480. Cambridge University Press, 2001.

[3] G. J. Ackland, D. J. Bacon, A. F. Calder, and T. Harry. Computer simulation of point defect properties in dilute fe-cu alloy using a many-body interatomic potential. *Phil. Mag. A*, 57:713–732, 1997.

[4] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Clarendon Press, Oxford, 1987.

[5] A. S. Argon. Brittle to ductile transition in cleavage fracture. *Acta Metall.*, 35:185–196, 1987.

[6] A. S. Argon. Rate processes in plastic deformation of crystalline and noncrystalline solids. In M. A. Meyers, R. W. Armstrong, and H. Kirchner, editors, *Mechanics and Materials: Fundamentals and Linkages*, pages 175 230. Jonh Wiley & Sons, Inc, 1999.

[7] A. S. Argon. Mechanics and physics of brittle to ductile transitions in fracture. *J. Eng. Mater. Tech.*, 123:1–11, 2001.

[8] M. Arroyo and T. Belytschko. An atomistic-based finite deformation membrane for single layer crystalline films. *J. Mech. Phys. Solids*, 50:1941–1977, 2002.

[9] R. J. Asaro. Crystal plasticity. *J. Appl. Mech.*, 50:921–934, 1983.

[10] H. Balamane, T. Halicioglu, and W. A. Tiller. Comparative study of silicon empirical interatomic potentials. *Phys. Rev. B*, 46:2250–2279, 1992.

[11] D. M. Barnett and J. Lothe. Line force loadings on anisotropic half-spaces and wedges. *Phys. Norvegica*, 8:13–22, 1975.

[12] M. Z. Bazant, E. Kaxiras, and J. F. Justo. Environment-dependent interatomic potential for bulk silicon. *Phys. Rev. B*, 56:8542–8552, 1997.

[13] N. Bernstein and D. W. Hess. Lattice trapping barriers to brittle fracture. *Phys. Rev. Lett.*, 91:025501, 2003.

[14] M. Born and K. Huang. *Dynamical Theory of Crystal Lattices.* Clarendon, Oxford, 1956.

[15] W. Cai. *Atomistic and Mesoscale Modeling of Dislocation Mobility.* PhD thesis, Massachusetts Institute of Technology, 2001.

[16] W. Cai, V. V. Bulatov, J. F. Justo, A. S. Argon, and S. Yip. Intrinsic mobility of a dissociated dislocation in silicon. *Phys. Rev. Lett.*, 84:3346–3349, 2000.

[17] F. Celarie, S. Prades, D. Bonamy, L. Ferrero, E. Bouchaud, C. Guillot, and C. Marliere. Glass breaks like metal, but at the nanometer scale. *Phys. Rev. Lett.*, 90:075504, 2003.

[18] K. S. Cheung and S. Yip. Atomic-level stress in an inhomogeneous system. *J. Appl. Phys.*, 70:5688–5690, 1994.

[19] T. J. Chuang and E. R. Fuller. Extended charles-hillig theory for stress corrosion cracking of glass. *J. Amer. Ceram. Soc.*, 75:540–545, 1992.

[20] F. Cleri, S. Yip, D. Wolf, and S. R. Phillpot. Atomic-scale mechanism of crack-tip plasticity: dislocation nucleation and crack-tip shielding. *Phys. Rev. Lett.*, 79:1309–1312, 1997.

[21] S. G. Corcoran, R. J. Colton, E. T. Lilleodden, and W. W. Gerberich. Anomalous plastic deformation at surfaces: Nanoindentation of gold single crystals. *Phys. Rev. B*, 55:R16057–R16060, 1997.

[22] W. A. Curtin. On lattice trapping of cracks. *J. Mat. Res.*, 5:1549–1560, 1990.

[23] M. S. Daw and M. I. Baskes. Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals. *Phys. Rev. Lett.*, 50:1285, 1983.

[24] N. H. de Leeuw, Z. Du, J. Li, S. Yip, and T. Zhu. A computer modelling study of the effect of hydration on the stability of a silica nano-tube. *Nato. Lett.*, 3:1347 1352, 2003.

[25] B. deCelis, A. S. Argon, and S. Yip. Molecular dynamics simulation of crack tip processes in alpha-iron and copper. *J. Appl. Phys.*, 54:4864–4878, 1983.

[26] J. L. Ericksen. The cauchy and born hypotheses for crystals. In M. E. Gurtin, editor, *Phase Transformations and Material Instabilities in Solids*, pages 61–78. Academic Press, 1984.

[27] C. W. Fan and C. B. Hwu. Energetics of homogeneous nucleation of dislocation loops under a simple shear stress in perfect crystals. *J. Appl. Mech.*, 63:69–76, 1996.

[28] M. C. Fivel, C. F. Robertson, G. R. Canova, and L. Boulanger. Three-dimensional modeling of indent-induced plastic zone at a mesoscale. *Acta Mater.*, 46:6183–6194, 1998.

[29] H. Gao and P. Klein. Numerical simulation of crack growth in an isotropic solid with randomized internal cohesive bonds. *J. Mech. Phys. Solids*, 46:187–218, 1998.

[30] A. George and G. Michot. Dislocation loops at crack tips: nucleation and growth an experimental study in silicon. *Mater. Sci. Eng. A*, 164:118 134, 1993.

[31] W. W. Gerberich, J. C. Nelson, E. T. Lilleodden, P. Anderson, and J. T. Wyrobek. Indentation induced dislocation nucleation: the initial yield point. *Acta Mater.*, 44:3585–3598, 1996.

[32] A. E. Giannakopoulos, P. L. Larsson, and R. Vestergaard. Analysis of vickers indentation. *Int. J. Solid Struct.*, 31:2679–2708, 1994.

[33] J. J. Gilman. *Electronic Basis of the Strength of Materials.* Cambridge University Press, 2003.

[34] A. Gouldstone, H. J. Koh, K. T. Zeng, A. E. Giannakopoulos, and S. Suresh. Discrete and continuous deformation during nanoindentation of thin films. *Acta Mater.*, 48:2277 2295, 2000.

[35] A. Gouldstone, K. J. Van Vliet, and S. Suresh. Nanoindentation: Simulation of defect nucleation in a crystal. *Nature*, 411:656, 2001.

[36] G. Henkelman and H. Jónsson. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.*, 111:7010 7022, 1999.

[37] G. Henkelman and H. Jónsson. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.*, 113:9978 9985, 2000.

[38] R. Hill. Acceleration waves in solids. *J. Mech. Phys. Solids*, 10:1 16, 1962.

[39] R. Hill. On the elasticity and stability of perfect crystals at finite strain. *Math. Proc. Camb. Phil. Soc.*, 77:225 240, 1975.

[40] W. B. Hillig and R. J. Charles. Surfaces, stress-dependent surface reactions, and strength. In V. F. Zackay, editor, *High-Strength Materials*, pages 682–705. Wiley, New York, 1964.

[41] D. Holland and M. Marder. Cracks and atoms. *Adv. Mater.*, 11:793–806, 1999.

[42] K. L. Johnson. *Contact Mechanics.* Cambridge University Press, 1985.

[43] H. Jonsson, G. Mills, and K. W. Jacobsen. Nudged elastic band method for finding minimum energy paths of transitions. In B. J. Berne, G. Ciccotti, and D. F. Coker, editors, *Classical and Quantum Dynamics in Condensed Phase Simulations*, pages 385–404. World Scientific, 1998.

[44] J. F. Justo, M. Z. Bazant, E. Kaxiras, V. V. Bulatov, and S. Yip. Interatomic potential for silicon defects and disordered phases. *Phys. Rev. B*, 57:2539–2550, 1998.

[45] C. L. Kelchner, S. J. Plimpton, and J. C. Hamilton. Dislocation nucleation and defect structure during surface indentation. *Phys. Rev. B*, 58:11085–11088, 1998.

[46] A. Kelly, W. R. Tyson, and A. H. Cottrell. Ductile and brittle crystals. *Phil. Mag.*, 15:567–586, 1967.

[47] J. D. Kiely and J. E. Houston. Nanomechanical properties of au (111), (001), and (110) surfaces. *Phys. Rev. B*, 57:12588–12594, 1998.

[48] P. Klein and H. Gao. Crack nucleation and growth as strain localization in a virtual-bond continuum. *Eng. Fract. Mech.*, 61:21–48, 1998.

[49] J. Knap and M. Ortiz. An analysis of the quasicontinuum method. *J. Mech. Phys. Solids*, 49:1899–1923, 2001.

[50] C. R. Krenn, D. Roundy, M. L. Cohen, D. C. Chrzan, and J. W. Morris. Connecting atomistic and experimental estimates of ideal strength. *Phys. Rev. B*, 65:134111, 2002.

[51] C. R. Krenn, D. Roundy, J. W. Morris, and M. L. Cohen. The non-linear elastic behavior and ideal shear strength of al and cu. *Mater. Sci. Eng. A*, 317:44–48, 2001.

[52] P. R. Laurence and I. H. Hillier. Towards modelling bioactive glasses: Quantum chemistry studies of the hydrolysis of some silicate structures. *Comp. Mater. Sci.*, 28:63–75, 2003.

[53] T. A. Laursen and J. C. Simo. A study of the mechanics of microindentation using finite elements. *J. Mat. Res.*, 7:618–626, 1992.

[54] B. R. Lawn. Atomistic aspects of brittle fracture. *J. Mater. Sci.*, 10:469–480, 1975.

[55] B. R. Lawn. *Fracture of Brittle Solids.* Cambridge University Press, 1993.

[56] J. Li. *Modeling Microstructural Effects on Deformation Resistance and Thermal Conductivity.* PhD thesis, Massachusetts Institute of Technology, 2000.

[57] J. Li. Atomeye: an efficient atomistic configuration viewer. *Modelling Simul. Mater. Sci. Eng.*, 11:173, 2003.

[58] J. Li, K. J. Van Vliet, T. Zhu, S. Yip, and S. Suresh. Atomistic mechanisms governing elastic limit and incipient plasticity in crystals. *Nature*, 418:307–310, 2002.

[59] C. G. Lindsay, G. S. White, S. W. Freiman, and W. Wong-Ng. Molecular orbital study of an environmentally enhanced crack growth process in silica. *J. Amer. Ceram. Soc.*, 77:2179–2187, 1994.

[60] M. Marder. Energies of a kinked crack line. *J. Stat. Phys.*, 93:511–525, 1998.

[61] T. A. Michalske and B. C. Bunker. Slow fracture model based on strained silicate structures. *J. Appl. Phys.*, 56:2686–2693, 1984.

[62] T. A. Michalske and S. W. Freiman. A molecular interpretation of stress corrosion in silica. *Nature*, 295:511, 1982.

[63] T. A. Michalske and S. W. Freiman. A molecular mechanism for stress corrosion in vitreous silica. *J. Amer. Ceram. Soc.*, 66:284–288, 1983.

[64] Y. Mishin, D. Farkas, M. J. Mehl, and D. A. Papaconstantopoulos. Interatomic potentials for monoatomic metals from experimental data and ab initio calculations. *Phys. Rev. B*, 59:3393–3407, 1999.

[65] Y. Mishin, M. J. Mehl, D. A. Papaconstantopoulos, A. F. Voter, and J. D. Kress. Structural stability and lattice defects in copper: Ab initio, tight-binding, and embedded-atom calculations. *Phys. Rev. B*, 63:224106, 2001.

[66] A. Needleman and M. Ortiz. Effect of boundaries and interfaces on shear-band localization. *Int. J. Solid Struct.*, 28:859–877, 1991.

[67] S. Ogata, J. Li, and S. Yip. Ideal pure shear strength of aluminum and copper. *Science*, 298:807–811, 2002.

[68] R. W. Ogden. *Non-linear Elastic Deformations.* John Wiley and Sons, 1984.

[69] W. C. Oliver and G. M. Pharr. An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *J. Mat. Res.*, 7:1564–1583, 1992.

[70] E. Orowan. The fatigue of glass under stress. *Nature*, 154:341–343, 1944.

[71] M. Ortiz and R. Phillips. Nanomechanics of defects in solids. In T. Y. Wu and E. van der Giessen, editors, *Advances in Applied Mechanics*, volume 36, pages 1–79. Academic Press, 1999.

[72] T. F. Page, W. C. Oliver, and C. J. McHargue. The deformation-behavior of ceramic crystals subjected to very low load (nano)indentations. *J. Mat. Res.*, 7:450–473, 1992.

[73] R. Perez and P. Gumbsch. Directional anisotropy in the cleavage fracture of silicon. *Phys. Rev. Lett.*, 84:5347–5350, 2000.

[74] R. Perez and P. Gumbsch. An *ab initio* study of the cleavage anisotropy in silicon. *Acta Mater.*, 48:4517–4530, 2000.

[75] J. R. Rice. The localization of plastic deformation. In W. T. Koiter, editor, *Theoretical and Applied Mechanics*, pages 207–220. North Holland Pub. Co., 1976.

[76] J. R. Rice. Thermodynamics of the quasi-static growth of griffith cracks. *J. Mech. Phys. Solids*, 26:61–78, 1978.

[77] J. R. Rice. First-order variation in elastic fields due to variation in location of a planar crack front. *J. Appl. Mech.*, 52:571–579, 1985.

[78] J. R. Rice. Continuum mechanics and thermodynamics of plasticity in relation to microscale deformation mechanisms. In A. S. Argon, editor, *Constitutive Equations in Plasticity*, pages 23–79. M.I.T. Press, Boston, 1987.

[79] J. R. Rice. Mechanics of brittle cracking of crystal lattices and interfaces. In R. M. Latanision and R. H. Jones, editors, *Chem. Phys. Fract., Proc. of NATO Advanced Research Workshop*, pages 22–43. Martinus Nijhoff Publishers, Dordrecht, 1987.

[80] J. R. Rice. Dislocation nucleation from a crack tip: an analysis based on the peierls concept. *J. Mech. Phys. Solids*, 40:239–271, 1992.

[81] J. R. Rice and G. E. Beltz. The activation energy for dislocation nucleation at a crack. *J. Mech. Phys. Solids*, 42:333–360, 1994.

[82] J. R. Rice and R. Thmoson. Ductile versus brittle behavior of crystals. *Phil. Mag.*, 29:73–97, 1974.

[83] J. Schiotz and A. E. Carlsson. The influence of surface stress on dislocation emission from sharp and blunt cracks in fcc metals. *Phil. Mag. A*, 80:69–82, 2000.

[84] G. Schoeck and W. Püschl. The formation of dislocation loops at crack tips in three dimensions. *Phil. Mag. A*, 64:931 949, 1991.

[85] V. B. Shenoy, R. Miller, E. B. Tadmor, D. Rodney, R. Phillips, and M. Oriz. An adaptive finite element approach to atomic-scale mechanics - the quasicontinuum method. *J. Mech. Phys. Solids*, 47:611–642, 1999.

[86] V. B. Shenoy, R. Phillips, and E. B. Tadmor. Nucleation of dislocations beneath a plane strain indenter. *J. Mech. Phys. Solids*, 48:649–673, 2000.

[87] J. E. Sinclair. An atomistic study of cracks in diamond-structure crystals. *Proc. Roy. Soc. Lond. A*, 59:83–103, 1972.

[88] J. E. Sinclair. Influence of interatomic force law and of kinks on propagation of brittle cracks. *Phil. Mag.*, 31:647–671, 1975.

[89] G. S. Smith, E. B. Tadmor, N. Bernstein, and E. Kaxiras. Multiscale simulations of silicon nanoindentation. *Acta Mater.*, 49:4089–4101, 2001.

[90] M. R. Sorensen, Y. Mishin, and A. F. Voter. Diffusion mechanisms in cu grain boundaries. *Phys. Rev. B*, 62:3658 3673, 2000.

[91] J. C. H. Spence, Y. M. Huang, and O. Sankey. Lattice trapping and surface reconstruction for silicon cleavage on (111) - *ab initio* quantum molecular-dynamics calculations. *Acta Mater.*, 41:2815–2824, 1993.

[92] F. H. Stillinger and T. A. Weber. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B*, 31:5262–5271, 1985.

[93] A. N. Stroh. Dislocations and cracks in anisotropic elasticity. *Phil. Mag.*, 7:625–646, 1958.

[94] Z. Suo. Singularities, interfaces and cracks in dissimilar anisotropic media. *Proc. Roy. Soc. Lond. A*, 427:331–358, 1990.

[95] S. Suresh, T. G. Nieh, and B. W. Choi. Nano-indentation of copper thin films on silicon substrates. *Scripta Mater.*, 41:951–957, 1999.

[96] J. G. Swadener and G. M. Pharr. Indentation of elastically anisotropic half-spaces by cones and parabolas of revolution. *Phil. Mag. A*, 81:447–466, 2001.

[97] E. B. Tadmor, R. Miller, R. Phillips, and M. Ortiz. Nanoindentation and incipient plasticity. *J. Mat. Res.*, 14:2233–2250, 1999.

[98] E. B. Tadmor, M. Ortiz, and R. Phillips. Quasicontinuum analysis of defects in solids. *Phil. Mag. A*, 73:1529–1563, 1996.

[99] E. B. Tadmor, G. S. Smith, N. Bernstein, and E. Kaxiras. Mixed finite element and atomistic formulation for complex crystals. *Phys. Rev. B*, 59:235–245, 1999.

[100] R. Thomson, C. Hsieh, and V. Rana. Lattice trapping of fracture cracks. *J. Appl. Phys.*, 42:3145–3160, 1971.

[101] R. Thomson, V. K. Tewary, and K. Masuda-Jindo. Theory of chemically induced kink formation on cracks in silica. i. 3-d crack green's functions. *J. Mat. Res.*, 5:619–630, 1987.

[102] T. C. T. Ting. *Anisotropic Elasticity: Theory and Applications.* Oxford Univ. Press, New York, 1996.

[103] G. H. Vineyard. Frequency factors and isotope effects in solid state rate processes. *J. Phys. Chem. Solids*, 3:121–127, 1957.

[104] J. J. Vlassak and W. D. Nix. Indentation modulus of elastically anisotropic half-spaces. *Phil. Mag. A*, 67:1045–1056, 1993.

[105] J. J. Vlassak and W. D. Nix. Measuring the elastic properties of anisotropic materials by means of indentation experiments. *J. Mech. Phys. Solids*, 42:1223 1245, 1994.

[106] K. J. Van Vliet, J. Li, T. Zhu, S. Yip, and S. Suresh. Quantifying the early stages of plasticity through nanoscale experiments and simulations. *Phys. Rev. B*, 67:104105, 2003.

[107] A. F. Voter, F. Montalenti, and T. C. Germann. Extending the time scale in atomistic simulation of materials. *Annu. Rev. Mater. Res.*, 32:321–346, 2002.

[108] D. C. Wallace. *Thermodynamics of Crystals.* Wiley, New York, 1972.

[109] T. R. Walsh, M. Wilson, and A. P. Sutton. Hydrolysis of the amorphous silica surface. ii. calculation of activation barriers and mechanisms. *J. Chem. Phys.*, 113:9191 9201, 2000.

[110] J. H. Weiner. *Statistical Mechanics of Elasticity.* Wiley, New York, 1983.

[111] J. K. West and L. L. Hench. Silica fracture, part ii. a ring-opening model via hydrolysis. *J. Mater. Sci.*, 29:5808–5816, 1994.

[112] S. M. Wiederhorn. Influence of water vapor on crack propagation in soda-lime glass. *J. Amer. Ceram. Soc.*, 50:407–414, 1967.

[113] S. M. Wiederhorn, A. Dretzke, and J. Rodel. Near the static fatigue limit in glass. *Int. J. Fract. Mech.*, 121:1–7, 2003.

[114] J. R. Willis. Hertzian contact of anisotropic bodies. *J. Mech. Phys. Solids*, 14:163–176, 1966.

[115] G. Xu and A. S. Argon. Energetics of homogeneous nucleation of dislocation loops under a simple shear stress in perfect crystals. *Mater. Sci. Eng. A*, 319:144–147, 2001.

[116] G. Xu, A. S. Argon, and M. Ortiz. Nucleation of dislocations from crack tips under mixed modes of loading: implications for brittle against ductile behavior of crystals. *Phil. Mag. A*, 72:415–451, 1995.

[117] G. Xu, A. S. Argon, and M. Ortiz. Critical configurations for dislocation nucleation from crack tips. *Phil. Mag. A*, 75:341 367, 1997.

[118] P. Zhang, P. Klein, Y. Huang, H. Gao, and P. D. Wu. Numerical simulation of cohesive fracture by the virtual-internal-bond model. *Comp. Model. Eng. Sci.*, 3:263–277, 2002.

[119] S. J. Zhou, A. E. Carlsson, and R. Thomson. Crack blunting effects on dislocation emission from cracks. *Phys. Rev. Lett.*, 72:852–855, 1994.

[120] S. J. Zhou, P. S. Lomdahl, R. Thomson, and B. L. Holian. Dynamic crack processes via molecular dynamics. *Phys. Rev. Lett.*, 76:2318–2321, 1996.

[121] S. J. Zhou and R. Thomson. Dislocation emission at ledges on cracks. *J. Mat. Res.*, 6:639–653, 1991.

[122] T. Zhu, J. Li, K. J. Van Vliet, S. Yip, and S. Suresh. Predictive modeling of nanoindentation-induced homogeneous dislocation nucleation in copper. *J. Mech. Phys. Solids*, 52:691–724, 2004.

[123] T. Zhu, J. Li, and S. Yip. Atomistic study of dislocation loop emission from a crack tip. *Phys. Rev. Lett.*, 92, 2004.

[124] T. Zhu, J. Li, S. Yip, R. J. Bartlett, S. B. Trickey, and N. H. deLeeuw. Deformation and fracture of a sio_2 nanorod. *Molec. Simul.*, 29:671–676, 2003.

[125] J. A. Zimmerman, C. L. Kelchner, P. A. Klein, J. C. Hamilton, and S. M. Foiles. Surface step effects on nanoindentation. *Phys. Rev. Lett.*, 87:165507, 2001.