Dynamical Brittle Fractures of Nanocrystalline Silicon using Large-Scale Electronic Structure Calculations

Takeo Hoshi and Takeo Fujiwara

Department of Applied Physics, University of Tokyo, Bunkyo-ku, Tokyo 113-8656

(Received March 25, 2003)

A hybrid scheme between large-scale electronic structure calculations is developed and applied to nanocrystalline silicon with more than 10^5 atoms. Dynamical fracture processes are simulated under external loads in the [001] direction. We show how the fracture propagates anisotropically on the (001) plane and reconstructed surfaces appear with asymmetric dimers. Step structures are formed in larger systems, which is explained by the beginning of a crossover between nanoscale and macroscale samples.

Silicon is an ideally brittle material and its fracture behavior is studied intensively, because we can obtain essentially dislocation-free single crystals. A pioneering theory of brittle fractures was given, in the 1920's, by Griffith[1] within a continuum theory, which is the foundation of the present understanding of brittle fractures.[2] The fracture in single crystals should also be investigated in atomistic pictures, on the points of how and why the fracture path is formed and propagates in the crystalline geometry. This point includes surface reconstruction processes. Since fracture is a thermal nonequilibrium process, the atomic structure on a cleavage surface may be different from that on equilibrium clean surfaces. For example, the easiest cleavage plane in macroscale samples of silicon is the (111) plane, in which the surface structure is not the ground-state (7×7) structure but a metastable 2×1 structure.[3,4]

This letter is focused on the study of atomistic fracture behaviors in *nanocrystalline* silicon, particularly, its possible difference from macroscale samples. Such a difference can be expected, as explained below; now a typical atomistic length in silicon is defined as $d_0 \equiv \sqrt[3]{v_0} \approx 3$ Å, where v_0 gives the volume per atom. The essence of the Griffith theory[1] is the energy competition between the energy gain of strain relaxation and the energy loss of surface formation. The former energy is a volume term proportional to $(\text{length})^3$, while the latter energy is a surface term proportional to $(\text{length})^2$. Analogous to the theory of nucleation,[5] dimensional analysis gives the critical crack length for spontaneous fracture propagation. The critical crack length c_G is given as[1,2]

$$c_G \approx \frac{\gamma E}{\sigma^2} \qquad (1)$$

with the stress σ, the Young modulus E ($\approx 10^2$ GPa), and the surface energy per unit area γ. The value of γ was estimated to be on the order of $1 \, \text{J/m}^2$,[6,7] which can be reduced to the bond breaking energy ($\gamma d_0^2 \approx 1$ eV) in the atomistic picture. In a recent experiment with macroscale samples,[8] the stress is $\sigma \approx 10^1$ MPa and eq. (1) gives a macroscale length ($c_G \approx 1$ mm). Since the length c_G is not dependent on the sample size L, the fracture behavior can be expected to be different from the above picture, when the sample size L is smaller than the critical length c_G ($L < c_G$). In this letter, we will discuss such a situation in *nanocrystalline* silicon, in which the numbers of atomic layers for these lengths ($\approx c_G/d_0, L/d_0$) are not macroscale numbers.

For atomistic fracture simulations of silicon crystals, a recent work of classical modelings[9] was carried out with 10^5 atoms. A more recent work,[7] however, pointed out the limited applicability of classical modelings and the importance of electronic structure calculations. On the other hand, there are several *ab initio* calculations with 10^2 atoms.[6,7] Due to the system size of simulations, these investigations are limited in certain situations, such as the preparation of the initial cleavage plane in which the reconstructed surface structure is assumed. Therefore, large-scale electronic structure calculations are essential.

So far, we have developed several order-*N* methods for large-scale electronic structure calculations.[10] The term order-*N* method is the general name of methods in which the computational cost is proportional to the system size (N). We have developed variational and perturbative order-*N* methods based on generalized Wannier states.[10,11] A Wannier state ϕ_i is localized and the index i denotes its localization center. The equation for wave functions is identical in the two methods and is given by the one-body density matrix ($\rho \equiv \sum_i^{\text{occ.}} |\phi_i\rangle\langle\phi_i|$).[10] In the computational algorithm, the perturbative method is simpler than the variational method. Figure 1 demonstrates our large-scale calculation with 10^2–10^6 atoms.

With the above two methods, we now construct a novel hybrid scheme, in which the variational method is used only for the wave functions whose centers locate near fracture regions. The regions typically contain 4×10^4 electrons. Some of such wave functions change their character dynamically from the bulk (sp^3 bonding) state to the surface states, as discussed later. The other wave functions, in bulk regions, maintain the character of the bulk bonding state and can be obtained by the perturbative method. The wave functions $\{\phi_i\}$ calculated by the two methods are both used for constructing the one-body density matrix ρ. Any physical quantity is expressed by the density matrix[10] and is well defined in the present hybrid scheme.

The present work is based on a transferable tight-binding Hamiltonian with s and p orbitals.[12] It is used for several crystalline phases and noncrystalline phases, such as liquid[12] and surfaces.[13] Since fracture is the formation of surfaces in a bulk region, the theory should reproduce the atomic
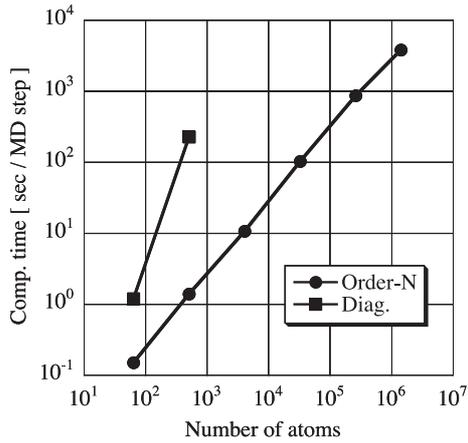
Fig. 1. Computational time for bulk silicon as a function of the number of atoms, up to 1423909 atoms. The CPU time is measured for one time step in the molecular dynamics (MD) simulation. A tight-binding Hamiltonian is used with the perturbative order-N method and the exact diagonalization method. We use a standard workstation with a single CPU and 2 GB of RAM.

structures in both bulk and surface phases, which is satisfied by the present Hamiltonian. The essence of the quantum mechanical freedoms is the fact that sp^3-hybridized bonds are formed in the bulk region, but not on surfaces. To analyze the hybridization freedom, the parameter $f_s^{(j)}$ is defined, for a wave function ϕ_j, as

$$f_s^{(j)} \equiv \sum_I |\langle \phi_j | Is \rangle|^2, \qquad (2)$$

where $|Is\rangle$ is the s orbital of the I-th atom. For example, $f_s = 1/4$ in an ideal sp^3-hybridized state.

In this letter, we focus on the Si(001) surface, a standard template of the modern silicon technology. A characteristic feature of the Si(001) surface is the formation of asymmetric dimers.[14,15] The asymmetric dimer is connected by a 'σ' bonding state. Another state is localized on the 'up' atom, the dimerized atom near the vacuum region. This localized state is called 'π' state, because the direction of its p components is nearly perpendicular to the dimer bond. Here an energy quantity is defined as

$$\Delta \varepsilon_i^{(\mathrm{cov})} \equiv \langle \phi_i | H | \phi_i \rangle - \left[f_s^{(i)} \varepsilon_s + (1 - f_s^{(i)}) \varepsilon_p \right]. \qquad (3)$$

A negative value of $\Delta \varepsilon_i^{(\mathrm{cov})}$ corresponds to the energy gain of covalent bonding. The 'σ' state has the gain of $\Delta \varepsilon_i^{(\mathrm{cov})} \approx -2$ eV, which mainly contributes to the dimerization energy (approximately -2 eV).[15] The 'π' state has much smaller $\Delta \varepsilon_i^{(\mathrm{cov})}$, which is comparable to the energy difference between the asymmetric and symmetric dimers (the order of 0.1 eV).[15]

The simulation details are as follows; the hybrid order-N scheme is used for systems with 10^4 atoms or more. In smaller systems, the variational method is used in the entire region. The samples are isolated tetragonal clusters, whose geometries are labeled with the number of atomic layers, such as $n_{100} \times n_{010} \times n_{001}$ or $n_{110} \times n_{1\bar{1}0} \times n_{001}$. As the boundary condition, the Wannier states on the sample surfaces are terminated by fixed sp^3 bonding states and are not reconstructed. The time step of the molecular dynamics is 3 fs. The total kinetic energy is controlled to be that with 300 K by the Nose thermostat method.[16] Numerical accuracy

is confirmed among bulk, surface and fracture properties, such as the elastic constants, the dimer formation on the clean (001) surface, and the critical stress for fracture. The last quantity is confirmed in smaller samples by the comparison with the result of the exact diagonalization method. The calculated fracture propagation velocity is always in the same order of, but less than, the Rayleigh surface wave velocity (4.5 km/s), as expected from the continuum theory.[2]

For fracture propagations, external loads in the [001] direction are imposed. During the simulations, the external loads are dynamically controlled by the atoms on the sample surfaces in the z direction. These atoms are fixed or under artificial constant-velocity motions in the z axis. The velocity, typically 10^{-2} km/s, is much smaller than the observed fracture propagation velocities (km/s). As a seed of fractures, a short range repulsive potential is imposed on one particular pair of atoms, as a defect bond. For smaller samples, the simulations begin without initial deformations. The fracture always occurs with the external loads on the order of $\sigma \approx 1$ GPa, which corresponds to the strain energy of $\sigma d_0^3 \approx 0.1$ eV per atom. For larger samples, the simulations begin with initial static deformations in the above magnitude of external loads. The length c_G in eq. (1) is calculated as $c_G \approx 100$ nm, which is longer than the present sample sizes ($L \leq 20$ nm).

In results, a two-stage reconstruction process is commonly observed as the elementary process during successive bond breakings; in Fig. 2(a), we monitor the one-electron energy $\varepsilon_i \equiv \langle \phi_i | H | \phi_i \rangle$ and the hybridization freedom $f_s^{(i)}$ of a Wannier state $|\phi_i\rangle$. Before bond breaking ($t < 0$ ps), the wave function $|\phi_i\rangle$ is a bonding state in the bulk region, deformed due to the external load. At $t \approx 0$ ps, a bond breaking occurs and the wave function $|\phi_i\rangle$ loses the bonding character with a rapid increase in bond length. Then (0 ps $< t < 0.2$ ps), a twofold coordinated surface atom is formed, since another bond is broken almost simultaneously. The wave function $|\phi_i\rangle$ forms a lone pair state that is stabilized by an increase in $f_s^{(i)}$ ($0.6 \rightarrow 0.8$). The corresponding energy gain is estimated to be $-0.2 \times (\varepsilon_p - \varepsilon_s) \approx -1.3$ eV, which explains the energy gain in the figure ($\varepsilon_i = -2.7$ eV $\rightarrow -3.8$ eV). In other words, the bond breaking process is caused by the local electronic instability, that is, the energy competition between the *loss* of the bonding (transfer) energy and the *gain* due to the increase in the weight of the s orbitals (f_s). Finally, after the thermal motions with a finite time
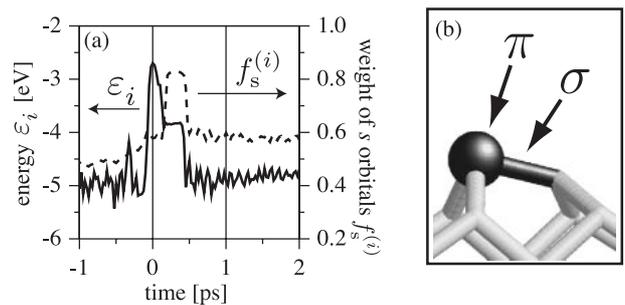

Fig. 2. (a) Elementary reconstruction process with the one-electron energy ε_i and the weight on s orbitals $f_s^{(i)}$. (b) An asymmetric dimer on a resultant crack. The black rod and black ball correspond to the 'σ' and 'π' states, respectively.
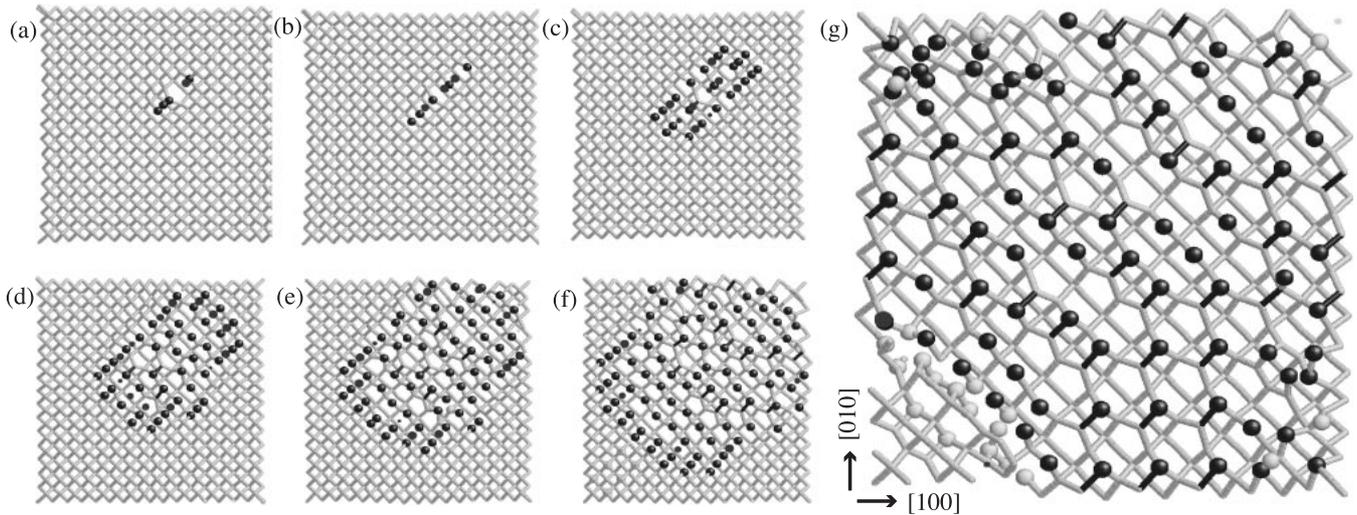
Fig. 3.   Snapshots of a fracture process in the (001) plane. The sample size is $n_{100} \times n_{010} \times n_{001} = 33 \times 33 \times 33$ (4501 atoms). The time interval between two successive snapshots is 0.3 ps, except that between (f) and (g) (approximately 1.3 ps). A set of connected black rod and black ball corresponds to an asymmetric dimer, as in Fig. 2(b). The lower left area has not yet been fractured.

($t \approx 0.4$ ps), a pair of twofold coordinated atoms forms an asymmetric dimer with a σ bonding state $|\phi_i\rangle$. The corresponding covalent-bonding energy, defined in eq. (3), is $\Delta\varepsilon_i^{(cov)} \approx -1.9$ eV. This energy explains the gain in the figure ($\varepsilon_i = -3.8$ eV $\rightarrow -4.8$ eV) and the energy *loss* (approximately 1.3 eV) due to a decrease in $f_s^{(i)}$ ($0.8 \rightarrow 0.6$). This asymmetric dimer is preserved until the end of the simulation, during a couple of pico seconds. Figure 2(b) shows an example of the observed asymmetric dimers.

Figure 3 shows the fracture process of a cubic sample with 4501 atoms. Each Wannier state is classified from its weight distribution into a bonding or atomic orbital, which is shown as a rod or a ball in the figures, respectively. A *black* rod or ball corresponds to a bonding or atomic orbital in the layer that contains the defect bond. One almost flat (001) surface is being created with many asymmetric dimers. The surface contains, however, many twofold coordinated atoms that have two back bonds (white rods) and a lone pair state (black ball). This is because a lone pair state are metastable, as discussed above. In Fig. 3, an anisotropic bond-breaking propagation is seen in the [110] and [1$\bar{1}$0] directions, particularly in the early snapshots. In the [110] direction, successive bond breakings propagate along *nearest neighbor* bond sites, which forms a zigzag path, shown as black rods in Fig. 4(a). A bond breaking process significantly weakens *nearest neighbor* bonds, due to the local electronic instability, as observed in Fig. 2(a). Therefore, the successive bond breakings propagate easily in the [110] direction. In the [1$\bar{1}$0] direction, on the other hand, bond-breaking paths are not connected, shown as red rods in Fig. 4(a). In this direction, the bond breakings propagate by the local strain relaxation, not by the local electronic instability. As a result, the bond breaking propagation along the *nearest neighbor* bond sites (in the [110] direction of the present surface) is faster than that in the perpendicular direction (in the [1$\bar{1}$0] direction), due to the difference in successive bond-breaking mechanisms. Note that a flat (001) surface is also obtained by a similar simulation *without the initial defect bond*, in which the fracture begins at the sample edges.

Figures 4(b) and 4(c) show larger samples with step formations.[17] In the two cases, all the conditions are the same, except for the sample size. To observe the step structures clearly, the broken bond sites are shown as rods in the *ideal* crystalline geometry. The defect bond is located in the center of the drawn area. The anisotropic fracture propagation in one (001) plane increases the anisotropic strain energy.[18] The anisotropy originates from the inequivalence between the [110] and [1$\bar{1}$0] directions within *one* (001) layer. Since the above inequivalence does not appear within *two* successive layers, a step formation between them will release the anisotropic strain energy. In Fig. 4(b), a step is formed between the layer of black rods and that of red rods. In the [110] direction, the bond-breaking propagation reaches the sample surfaces without step formations. In the [1$\bar{1}$0] directions, the bond breakings propagate slower and a step is formed in the central area at an early period of the crack propagation. After that, the fracture propagates among the two atomic layers. Since the two layers are symmetrically equivalent, the resultant step formation path is almost a line in the [100] or [010] direction as the boundary of the fractured areas between the two layers.

In Fig. 4(c), the largest sample used in the present letter shows that the above line structure does not reach the sample surfaces but is canceled with additional step formations in complicated paths. The sample size dependence of the step structures is explained by the beginning of the crossover between nanoscale and macroscale samples; if the sample contains so many atoms, the geometry of the resultant crack will be almost circular, as in Fig. 4(c), so as to minimize the anisotropic strain energy.[18] If not, the strain energy is accumulated only within the confined bulk region due to the finite sample size. The resultant fracture behavior is directly related to the anisotropic atomic structure of the cleaved surface, as in Fig. 4(b).

Since the above mechanism of step formations is two-dimensional, the present samples may be nanoscale 'thin' samples. In larger or thicker samples, an expected fracture behavior is the bending of the fracture plane into the (111) plane, the easiest cleavage plane in macroscale samples, which is the crossover in the present context. In a *sufficiently*
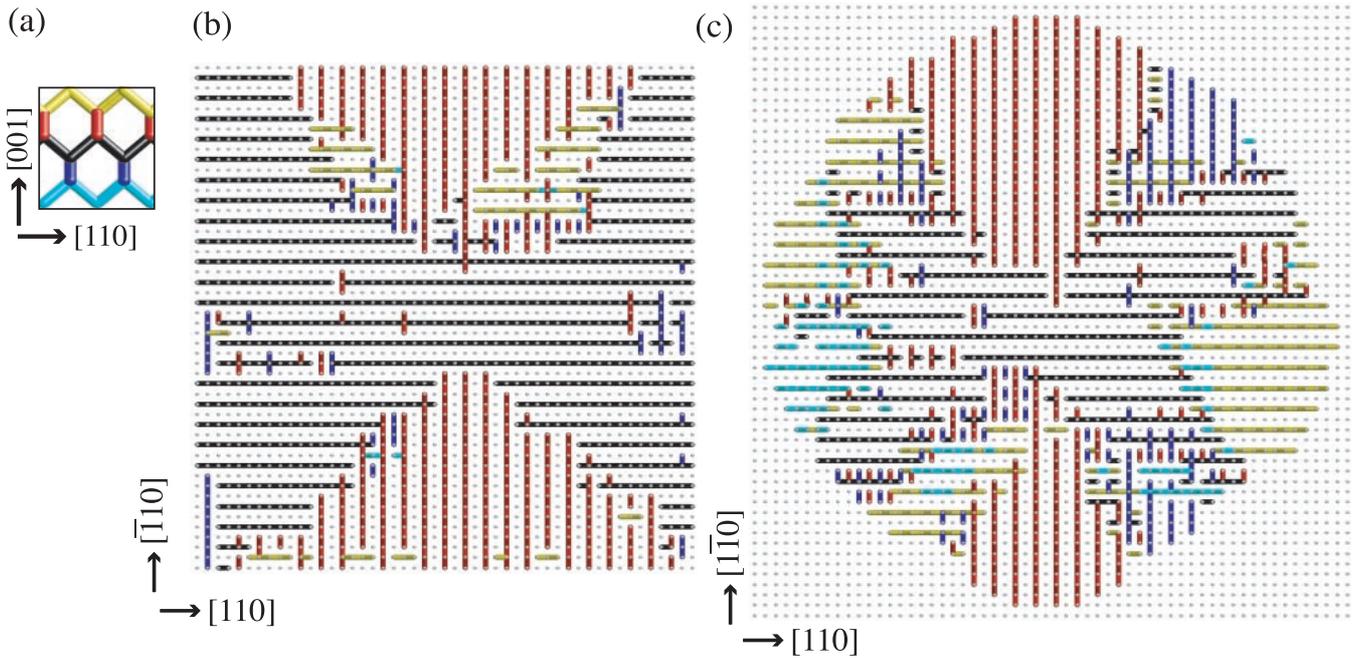
Fig. 4.   (a) Ideal diamond structure with colored bond sites. (b)(c) Geometry of resultant cracks in the (001) plane. The broken bond sites are plotted as colored rods in the ideal (crystalline) geometry. Rods within one layer are painted within the same color, as in (a). The layer of black rods contains the defect bond at its central area. Atoms are plotted as dots. The sample sizes in (b) and (c) are $n_{110} \times n_{1\bar{1}0} \times n_{001} = 49 \times 50 \times 49$ (30025 atoms) and $97 \times 100 \times 49$ (118850 atoms), respectively. In (c), only the central area ($n_{110} \times n_{1\bar{1}0} = 58 \times 60$) of the sample is shown. Note that the length of $n_{110} = 50$ atomic layers is approximately 10 nm.

large sample, the fracture mode with the easiest cleavage plane will grow regardless of sample shape and details of conditions. Note that the dynamical simulation with 10^5 atoms is the practical limitation within a single CPU workstation. A program code with parallel computations is now being developed for simulations with large samples.

This letter shows a possible difference in fracture behaviors among nanoscale and macroscale silicon crystals. Its origin is the size dependence of the energy competition between bulk and surface regions. The electronic structures between the two regions are essentially different and can be described by the present method with the well-defined total energy. This energy competition is also inherent in other phenomena, such as crystal growth and self-organizations, which may be candidates for applications of the present method.

1) A. A. Griffith: Philos. Trans. R. Soc. London, Ser. A **221** (1920) 163.
2) See textbooks of fracture, such as B. Lawn: *Fracture of Brittle Solids* (Cambridge University Press, 1993) 2nd ed.
3) K. C. Pandey: Phys. Rev. Lett. **47** (1981) 1913.
4) F. Ancilotto *et al.*: Phys. Rev. Lett. **65** (1990) 3148.
5) For example, L. D. Landau and E. M. Lifshitz: *Statistical Physics* (Pergamon Press, Oxford, 1980) 3rd ed., Part I.
6) J. C. H. Spence, Y. M. Huang and O. Sankey: Acta Metall. Mater. **41** (1993) 2815.
7) R. Pérez and P. Gumbsch: Phys. Rev. Lett. **84** (2000) 5347.
8) T. Cramer, A. Wanner and P. Gumbsch: Phys. Rev. Lett. **85** (2000) 788.
9) D. Holland and M. Marder: Phys. Rev. Lett. **80** (1998) 746.
10) T. Hoshi and T. Fujiwara: J. Phys. Soc. Jpn. **69** (2000) 3773.
11) T. Hoshi and T. Fujiwara: Surf. Sci. **493** (2001) 659.
12) I. Kwon *et al.*: Phys. Rev. B **49** (1994) 7242.
13) For example, C.-C. Fu, M. Weissmann and A. Saúl: Surf. Sci. **494** (2001) 119.
14) D. J. Chadi: Phys. Rev. Lett. **43** (1979) 43.
15) A Ramstad, G. Brocks and P. J. Kelly: Phys. Rev. B **51** (1995) 14504.
16) S. Nose: Mol. Phys. **52** (1984) 255.
17) The steps on the (001) surface are classified into four types [D. J. Chadi: Phys. Rev. Lett. **59** (1987) 1691]. The present surfaces, however, contain unreconstructed domains and are different from the above ones.
18) The elastic property of silicon crystal shows only a small anisotropy within the (001) plane; the [110] and [$\bar{1}$10] directions are equivalent and the values of the Young modulus are different by only approximately 30% in the [100] and [110] directions.