Reconstruction and intermixing in thin Ge layers on Si(001)

L. Nurminen,[1] F. Tavazza,[2] D. P. Landau,[1,2] A. Kuronen,[1] and K. Kaski[1]
[1]*Laboratory of Computational Engineering, Helsinki University of Technology, P.O. Box 9203, 02015 HUT, Finland*
[2]*Center for Simulational Physics, The University of Georgia, Athens, Georgia 30602-2451, USA*
(Received 29 April 2003; published 26 August 2003)

In this work the Monte Carlo method with an empirical potential model for atomic interactions is applied to study reconstruction and intermixing at a Ge-covered Si(001) surface. We investigate the structure and energetics of the $2\times n$ reconstruction which serves as a strain-relief mechanism. The optimal value of n is found to be strongly dependent on the thickness of the Ge overlayer. Si-Ge intermixing is studied using a direct simulation method which includes entropic effects. Ge occupation probabilities in subsurface layers are evaluated as a function of Ge coverage at different temperatures. The results show that strain-relief driven intermixing has a pronounced effect on the surface reconstruction once the Ge coverage reaches a full layer. We also evaluate the effect of temperature on the distribution of Ge in subsurface layers and discuss effects due to kinetic limitations. In agreement with experiments, the study provides a description of the interplay between reconstruction and intermixing at Ge-covered Si(001).

DOI: 10.1103/PhysRevB.68.085326 PACS number(s): 68.35.−p, 05.10.Ln

I. INTRODUCTION

SiGe heterojunctions have generated considerable interest due to their potential applications in novel electronic and optoelectronic devices. Understanding the growth of Ge on Si(001) is crucial for developing new fabrication methods and utilizing the full potential of these materials. From the theoretical point of view, the similar properties of Ge and Si make this system an ideal candidate for investigating the effects of misfit strain in heteroepitaxial systems.

Studying strain-relaxation phenomena, such as reconstruction and intermixing, requires using large-scale simulation methods due to propagation of long-range elastic effects in these systems. Accurate first-principles quantum-mechanical methods are currently limited to systems composed of only a few hundred atoms at best.[1] Tight-binding techniques[2] have extended the feasible system size to a few thousand atoms, but still in many cases classical empirical potentials are the only practical choice for large-scale simulations. In the present case, however, the interactions between atoms are short range, but give rise to long-range elastic effects. When this approach concerning interactions is combined with the Monte Carlo method, it offers a versatile tool for studying various interesting phenomena including effects such as intermixing where considerable configurational rearrangement is required. Furthermore, the empirical-potential based Monte Carlo method can also be used to study systems at nonzero values of temperature where entropic effects may become significant in determining what the state of lowest free energy is.

In this study, the Stillinger-Weber[3] model is used in connection with the Monte Carlo method to investigate the structure and energetics of thin Ge layers deposited on the Si(001) surface. We first consider a system with a fully segregated Ge layer and study the evolution of the $2\times n$ reconstruction which consists of a periodic arrangement of dimer vacancy lines. The optimal periodicity n_{opt} of the $2\times n$ pattern is determined by the balance of the energy gain from strain relief and the energy cost from forming dimer vacancies. We derive n_{opt} as a function of the thickness of the Ge layer and compare the results with experimental measurements and predictions of other theoretical models. Most previous theoretical studies have not accounted for the possibility of Si-Ge intermixing which has complicated direct comparison between theory and experiment. In this work, we include intermixing directly into the simulations and estimate how it affects the surface reconstruction. Knowing that intermixing may be driven by both strain relief and entropic effects, we try to separate these factors by first taking the low-temperature limit where entropic effects do not contribute to the problem and then extending the study to higher temperatures. Kinetic limitations of intermixing are also discussed.

This paper is organized as follows. Computational details of our simulations and surface energy calculations are given in the next section. The results and discussion are presented in Sec. III. Finally, we summarize our conclusions in Sec. IV.

II. COMPUTATIONAL DETAILS

A. Monte Carlo simulations

In this work, constant-pressure Monte Carlo (MC) simulations[4] are used to study temperature-dependent structural properties of the Ge/Si(001) surface. The Si-Si, Ge-Ge, and Si-Ge interactions are modeled using the empirical Stillinger-Weber (SW) potential.[3,5] We have recently tested the suitability of three different empirical potentials for finite-temperature simulations involving the Si(001) surface.[6] The SW potential is used in the current work because it was found give the best overall description of Si(001). For comparison purposes, we also report some results given by the Tersoff potential T3.[7] Details of both potential models (SW and T3) and parametrizations used in this work can be found in Refs. 6, 5, and 7.

In the MC simulations, we use periodic slab geometry in which the unit cell is constructed along the $[110]$, $[1\bar{1}0]$, and $[001]$ directions and periodic boundary conditions are applied in the x and y directions (z normal to the surface).

The simulation slab consists of two surfaces separated by a bulk region. From our earlier study[6] we know that the finite thickness of the slab can significantly affect the results due to effects related to volume changes in constant-pressure simulations. It is important that the dimensions of the system do not change as a result of having a different number of Ge layers on the surface or a different reconstruction (i.e., we want the simulation to correspond to a thin Ge layer on a much larger substrate). In order to minimize the finite size effects, we use thick simulation slabs consisting of 59 atomic layers. Several tests showed that convergence to acceptable accuracy could not be reached using 39 atomic layers, whereas increasing the thickness to 79 layers did not result in significant further changes. The odd number of atomic layers is used in order to have perpendicular dimer directions on the two surfaces so that anisotropic volume changes in the x and y directions can be avoided (see our previous work[6] for details). For the same reason, the top and bottom surfaces of the slab also have an equal amount of Ge and the same reconstruction (in perpendicular directions $2\times n$ and $n\times 2$).

We also performed several runs to check for effects caused by the finite cross section of the slab. For example, two different surface sizes 20×20 and 40×40 were used to calculate the energy per Ge dimer on a 2×10 reconstructed surface with respect to a 2×1 reconstruction (for systems with 1 and 3 ML of Ge). The two sets of simulations produced identical results within the accuracy of the calculation (i.e., 0.0001 eV). Thus the results in Sec. III are obtained using the following surface sizes which are chosen for each simulation according to the periodicity of the $2\times n$ reconstruction 16×16 for ($n=8$ or 16), 18×18 for ($n=6$), 20×20 for ($n=1$, 4, 5 or 10), and 24×24 for ($n=12$). Consequently, the majority of the simulations is performed using systems consisting of 15 000–35 000 atoms, while in the test runs the largest systems reach 100 000 atoms.

We use two kinds of MC algorithms in our simulations. For those simulations where intermixing is not considered, standard constant-pressure MC is used. Each MC step consists of two types of moves: small random displacements of individual particles and random variations of the sides of the simulation cell (in order to keep the pressure constant at $P=0$). The acceptance probabilities are given by the Metropolis form.[8] For studying intermixing, Si-Ge switches are incorporated into the standard algorithm such that at each MC step we randomly select a Si pair for each Ge atom in the system and attempt to exchange the identities of these two atoms (Ge becoming Si and vice versa). The attempted exchange is accepted or rejected using the Metropolis criterion. In this so-called random-switch algorithm, the total number of both Si and Ge atoms is conserved, and thus in a given simulation run we obtain the equilibrium distribution for a fixed amount of Ge in the system.

Typically, a single simulation run consists of 80 000–200 000 MC steps depending on the length of the equilibration (e.g., intermixing requires longer equilibration times). Averages are taken over the last 20 000 MC steps. Unless otherwise indicated, the error bars for the results in Sec. III are smaller than or comparable to the size of the corresponding symbol in the figures.

The random-switch algorithm is a very fast method to achieve large configurational changes in the system which would not be possible using regular single-particle moves or a method such as molecular dynamics (MD) where the particles must follow trajectories dictated by the equations of motion. In our MC simulations, we are not trying to mimic the actual diffusion mechanism in real systems but the purpose of this study is to investigate the thermodynamic driving forces behind experimentally observed changes. Since our simulations are not affected by kinetic limitations, we can use the low temperature limit to determine the relative stability of the different reconstructions and to estimate how it is affected by strain-relief driven intermixing. Experimental observations are of course made at a much higher temperature and the systems are influenced by kinetic factors. For this reason we investigate two limiting cases: first the case where the surface is reconstructed but no intermixing between the Ge layers and the Si substrate takes place and then the case where Si/Ge intermixing is allowed to occur and its effect on the optimal reconstruction of the surface is evaluated. Real systems are likely to lie somewhere in between these two limits, but this approach provides us important information of the interplay between the two main mechanisms of strain relief in the Ge/Si(001) system: reconstruction and intermixing. We are considering ways of extending the approach to higher temperatures, but this would require calculations of free energy differences.

We would like to point out that moving only one particle at a time influences the region of phase space that the system can explore in a reasonable time. The energy landscape of the Ge/Si(001) surface is very complicated due to dimerization and consequent distortions of atomic positions in the underlying layers. The potential barrier for a jump of a single dimer from one binding site to a neighboring site is extremely high, and therefore such events never occur if only small random displacements of individual particles are attempted. This prevents large-scale relaxation of the surface structure (e.g., dimer vacancies are frozen into their initial positions). To overcome these problems we have developed a hybrid MC-MD algorithm in which coupled moves of several particles are incorporated into the standard approach.[9] In the case of the Ge/Si(001) system, the evolution of the surface structure in real systems is governed by the diffusion of dimers. Thus when applied to the study of this surface, the coupled-moves algorithm is based on the introduction of long dimer jumps such that we can overcome the potential barrier between two neighboring binding sites. After the initial displacement, the local neighborhood of the dimer is relaxed using MD. The resulting configuration is then used as a trial configuration in the Metropolis scheme. The algorithm can be applied to study several physically interesting phenomena where dimer diffusion is the main mechanism for structural changes (e.g., formation of dimer-vacancy structures, formation and relaxation of islands or equilibrium shape of step edges). A prerequisite for these investigations is to have a potential model which can accurately describe the essential features of Ge/Si(001). One of the goals of the current study is to provide the necessary information concerning the properties of the SW model.

B. Surface energy calculations

Comparing the energies of surface structures with different stoichiometries has posed problems in earlier studies of Ge/Si(001) surface reconstruction. The total number of Ge atoms in the simulation system changes depending on the periodicity of the surface layer. Consequently, an appropriate chemical potential must be assigned for the Ge atoms to determine the relative stabilities of the different reconstructions. Oviedo *et al.*[10] have recently introduced an intelligent procedure for determining the chemical potential in such a way that close correspondence to experimental conditions is achieved. The procedure is briefly explained here for the cases relevant to this study.

Consider a system with a *fixed* number of Ge atoms. Assuming that the surface is in thermal equilibrium, the probability of observing a particular configuration Γ is proportional to $\exp(-E_\Gamma/k_BT)$ where E_Γ is the energy of the configuration in question. We are interested in the relative stability of different reconstructions, obtained by distributing a fixed number of Ge atoms in different ways. For practical reasons, it is convenient to choose one reference configuration with energy $E_{\rm ref}$ and calculate all the other energies with respect to this reference point. As in Ref. 10, we choose the reference system to be a surface with a perfect 2×1 reconstruction. We now need a computational procedure to determine the energy difference between a given $2\times n$ reconstruction and the reference system. For understanding the correspondence to experimental conditions, consider having a large section of the surface covered with a 2×1 reconstructed layer of Ge. Let E_p denote the energy of a Ge dimer in the 2×1 reconstructed layer (the area of the Ge layer is thought to be so large that edge effects are negligible). Next an appropriate number of Ge atoms is fetched from infinity and placed at the edges of the Ge layer, thus increasing the total energy of the system by E_p per additional Ge dimer. The $2\times n$ reconstruction can now be constructed by removing every nth dimer from each row and taking the removed atoms to infinity (the number of previously added atoms equals the number of removed atoms). The energy cost of forming the missing-dimer vacancies, or the formation energy per vacancy, is denoted E_f. The resulting energy change in going from the initial 2×1 arrangement to the $2\times n$ reconstruction is now given by

$$\xi(n)=\frac{E_f(n)+E_p}{n-1}. \tag{1}$$

In order to determine the energy differences $\xi(n)$, we need to calculate E_p (for each Ge coverage) and E_f (for a range of n values for each Ge coverage). In practice, these values are obtained from simulations in the following way: for a given Ge coverage, (i) E_p is given by the energy difference per Ge dimer between a system with 2×1 reconstruction and a system obtained by removing the top Ge layer and relaxing the system with a 2×1 reconstruction [e.g., if we are calculating the $\xi(n)$ curve for 2 ML Ge coverage, we remove a single layer of Ge from each surface, thus obtaining 1 ML coverage with the total number of atomic layers reduced by 2]. (ii) $E_f(n)$ is given by the energy difference per Ge dimer between a system with a $2\times n$ reconstructed top layer and the reference system with 2×1 reconstruction (note that in both of these calculations the number of Ge atoms in the two systems is different). Having obtained these energy values, we can now determine the relative stability of different $2\times n$ reconstructions using Eq. (1).

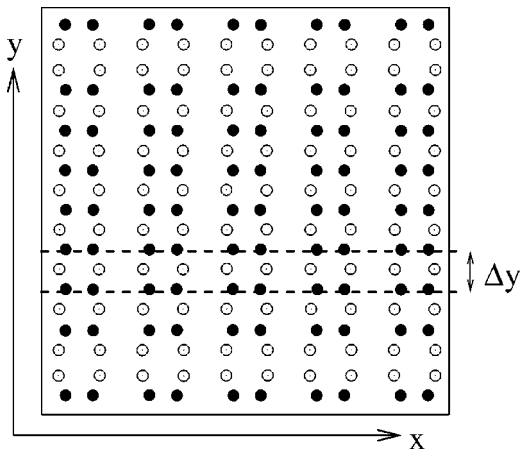

FIG. 1. $2\times n$ reconstruction viewed from above ($n=8$). Filled circles are for atoms in the topmost layer (dimers) and open circles for atoms in the layer beneath. Δy shows the direction in which separation between rows of dimers is measured in Fig. 3.

For full-monolayer coverages, the reference system consists of the preassigned number of Ge layers (1, 2, or 3 ML) with a 2×1 reconstructed surface. For submonolayer coverages, however, the surface layer consists of a random mixture of Ge-Ge, Si-Ge, and Si-Si dimers with the percentage of Ge atoms equal to $\theta_{\rm Ge}$. We checked that optimizing the Si-Ge mixture in the surface layer results in a negligible decrease in energy (the intermixing algorithm was used for atoms in the surface layer). Similarly, the reference system for 1.5 ML Ge coverage consists of a full top layer of Ge and a random mixture in the layer beneath. In all cases, the $2\times n$ reconstruction is obtained from the reference system by removing every nth dimer from each row (vacancy lines are preformed in the initial configurations but rebonding of the second-layer atoms occurs during the simulation).

III. RESULTS AND DISCUSSION

A. $2\times n$ reconstruction

The $2\times n$ reconstruction is formed during the growth of a wetting layer when Ge is deposited on Si(001).[11] In this structural rearrangement, vacancies are formed on the surface and their mutual interactions lead to the formation of a periodic line structure where every nth dimer in each row is replaced by a vacancy (see Fig. 1). The missing-dimer trenches relieve part of the misfit strain in the system by offering space for the outward relaxation of the Ge overlayer. Due to rebonding of the second-layer atoms, the formation of vacancies does not cost a large amount of energy. The balance between the energy gain from the strain relaxation and the energy cost from forming the vacancies determines the optimal periodicity of the $2\times n$ reconstruction.

The optimal value of n can also be affected by strain relief due to possible Si-Ge intermixing. Several recent experiments have suggested that significant Ge migration into subsurface regions may occur at typical growth temperatures (500 °C–700 °C).[12–16] In a scanning tunneling microscopy study,[17] empty-state imaging was successfully applied to identify intermixing sites for submonolayer Ge coverages deposited on Si(001). The study confirms that a measurable degree of Si-Ge place exchange occurs in the first layer at temperatures as low as 57 °C (330 K). The stoichiometry of the entire wetting layer is, however, very difficult to determine precisely from experimental measurements.[11,18] Consequently, several questions related to intermixing have not been fully answered.

Our aim, here, is to investigate the evolution of the $2\times n$ reconstruction by first considering the case where all Ge remains in the topmost layers. In this section, we calculate the optimal value of n as a function of Ge coverage without intermixing and compare the results to experimental measurements and other theoretical calculations. Then Si-Ge intermixing is discussed in the next section where we calculate the compositional profiles for Ge occupation in subsurface layers and determine the effect of intermixing on the optimal surface reconstruction. The simulations in this section and in Secs. III B 1 and III B 2 are performed at a very low temperature $k_BT=0.001$ eV (11 K), because determining free energy differences at higher temperatures is not a straightforward task, especially when comparing systems with different stoichiometries. Finally, the temperature dependence of intermixing is discussed in Sec. III B 3.

The following reference energies $E_p(\theta_{\mathrm{Ge}})$ are obtained for comparing the relative stability of different $2\times n$ reconstructions -8.668, -8.316, -8.079, -7.607, and -7.612 eV (per dimer) for Ge coverages of 0.5, 0.8, 1.0, 2.0 and 3.0 ML respectively. The small but surprising decrease in E_p when the Ge coverage increases from 2 to 3 ML is explained by subtle effects due to the relaxation of the underlying Ge layers and the Si substrate. Interestingly a similar change was observed in tight binding calculations by Li *et al.*[19] For each coverage, the formation energy $E_f(n)$ is calculated for different values of n ranging from 4 to 16, and using the obtained values for E_p and E_f, we then use Eq. (1) to determine the $\xi(n)$ values for comparing the relative stability of the different reconstructions.

In Fig. 2, we show the energy difference per Ge dimer between a $2\times n$ reconstruction and the reference surface with 2×1 ordering. We notice that for the 0.5 ML coverage, vacancy lines are energetically favorable only for very large values of n and $\xi(n)$ curve has no minimum for $n\leqslant16$. This indicates that while a low concentration of dimer vacancies is favorable at small submonolayer coverages of Ge, the vacancy-vacancy interactions are likely to be too weak for an organized line pattern to appear on the surface. When the Ge coverage is increased to 0.8 ML, a weak minimum appears in the $\xi(n)$ curve between $n=12$–14. Further increase in the thickness of the Ge overlayer leads to a shift in the optimal value of n such that at 1.0 ML coverage the minimum is at $n=10$ and at about 2 ML coverage a saturation point is reached after which the optimal value remains at $n=8$. The

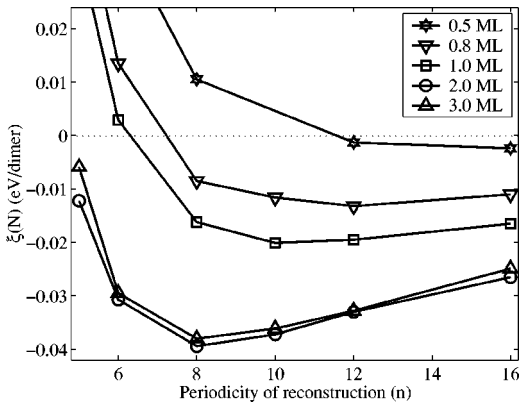

FIG. 2. Energy change per Ge dimer to form a $2\times n$ reconstruction from a 2×1 reconstructed surface obtained from MC simulations at $k_BT=0.001$ eV (11 K).

shift toward smaller n values is accompanied by an increasing gain in energy until 2 ML coverage is reached. Further reduction of the periodicity beyond 2 ML coverage is energetically unfavorable due to a repulsive interaction between dimer vacancy lines. The repulsion is mediated by subsurface atoms under the vacancy line which are strained due to a distortion from the usual diamond lattice positions. The interaction between the two local strain fields results in an effective repulsion between the vacancy lines.

The dimer vacancy lines act as a strain relief mechanism by providing space for the expansion of the Ge overlayer. Figure 3 shows the average separation between rows of dimers (in the direction of the dimer bond) on a 2×10 reconstructed surface. For comparison, we also show the average row separation on a 2×1 reconstructed surface. The row separation is measured in the direction perpendicular to the dimer bonds (and perpendicular to the vacancy lines) (see Fig. 1 for illustration). Two vacancy lines are located at the edges of the surface area shown in the figure. We notice that the separation between the Ge atoms increases significantly

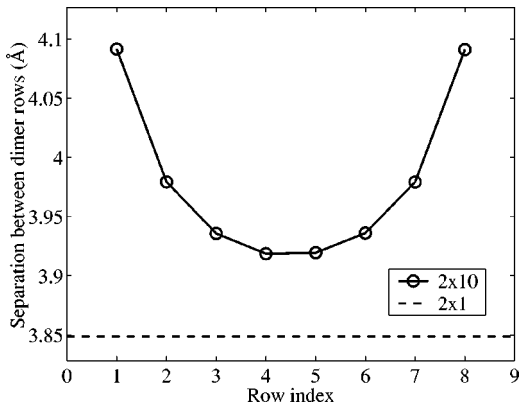

FIG. 3. Average separation between rows of dimers (in the direction of the dimer bond) on 2×10 and 2×1 reconstructed surfaces with 2.0 ML Ge coverage. The row separation is measured in the direction perpendicular to the dimer bonds (marked as Δy in Fig. 1). The vacancy lines in the 2×10 reconstruction are situated at positions with indeces 0 and 9. The simulation temperature is $k_BT=0.001$ eV (11 K).

in the whole region between the two vacancy lines when compared to the 2×1 surface. The relaxation is largest near the vacancy lines. Decreasing the periodicity to $n=8$ (optimal for 2 ML coverage) provides further room for relaxation.

Very similar results have been reported in experimental studies of Ge on Si(001).[18,20,21] All studies have indicated that the $2\times n$ reconstruction appears when approximately 0.8 ML of Ge has been deposited on the surface. In addition, continued deposition leads to a decrease in the average value of n until saturation is reached around 2 ML Ge coverage. Different groups have reported slightly different values of n depending on the growth conditions. For example, at 0.8 ML coverage, the measured values range from $n=17$ (Ref. 20) to $n=11$.[18] Around 1.5 ML coverage, the values have decreased to $n=9$–10 (Refs. 18,20) and convergence to $n\approx8$ around 2 ML coverage has been reported by several groups (Refs. 20,21 and references therein). Because further reduction of the periodicity is energetically unfavorable, the $2\times n$ reconstruction becomes less effective in reducing the strain in the system as the Ge film thickens. Experiments show that other structures begin to appear on the surface beyond 2 ML coverage (e.g., a so-called $m\times n$ pattern emerges where additional trenches form perpendicular to the dimer vacancy lines in the $2\times n$ reconstruction.[11]). In addition, experimental evidence shows that the distribution of n is broad for low Ge coverages and becomes narrower with increasing coverage.[18] The same trend can be seen in Fig. 2 where the energies for $n=12$–16 at 0.8 ML coverage are almost degenerate, but as the Ge coverage increases, the $\xi(n)$ curve becomes steeper around the minimum. The deeper and narrower potential well indicates that the distribution of n will be narrower on surfaces with larger Ge coverages.

The distribution of n is related to the number of kinks which appear in the individual vacancy lines. Experiments show that when the vacancy lines first appear on the surface, they have a highly disordered structure.[18] The ordering of missing dimers into straight lines becomes energetically more favorable as the thickness of the Ge layer increases. This straightening of the vacancy lines is associated with the narrowing of the distribution of the average line separation. In this work, we have only considered the idealized case of straight lines with a fixed line spacing. However, we are currently investigating possibilities of simulating a changing vacancy-line configuration with kinks and no restrictions on the line spacing as an application of the hybrid MC-MD algorithm mentioned in Sec. II.

The $2\times n$ reconstruction has been previously investigated using, e.g., the empirical Tersoff potential,[21] the tight binding (TB) method,[19] and even first-principles calculations.[10] The results obtained in these studies are not directly comparable because varying approaches have been used to compare the relative stability of $2\times n$ reconstructions. The method used in the present work and, in our view, the one corresponding most closely to experimental conditions (see Sec. II for details) was introduced recently by Oviedo *et al.*[10] and further modified by Li *et al.*[19] In the latter tight-binding study, the values $n=8$, 6, and 4 were obtained for Ge coverages of 1, 2, and 3 ML, respectively. Surprisingly, our empirical-potential based approach produces results in much closer agreement with experimental observations than do the TB calculations. The authors of Ref. 19 explain that the discrepancy between their TB calculations and experiments may be due to the fact that the experimentally observed structures may not correspond to thermal equilibrium due to kinetic reasons and experiments may also be influenced by intermixing of Ge into subsurface layers. In our MC simulations, however, the surface is also in thermal equilibrium as in the TB calculations and yet our results correspond closely to experimentally observed changes in the $2\times n$ reconstruction. For comparison purposes, we also performed the same calculations using the Tersoff potential T3 which gives the following optimal periodicities for the $2\times n$ reconstruction: $n=10$, 6, and 6 for Ge coverages of 1, 2, and 3 ML, respectively. These values are closer to those predicted by the SW model ($n=10$, 8, and 8), although the T3 model gives a deeper and narrower potential minimum for 2 and 3 ML coverages than does SW. We note that since the same simulation method was used for the SW and T3 calculations, the differences result only from the use of different potential models, whereas the TB results were obtained from static calculations of much smaller systems at constant volume, which may lead to other effects in addition to the properties of the interaction model itself. To further investigate this issue, questions related to effects of intermixing are studied in the next section.

B. Intermixing

Studying the effects of intermixing is complicated due to an interplay of thermodynamic driving forces, effects of entropy, and kinetic limitations. Our approach is to first consider the system at the low-temperature limit, where strain relaxation is the main driving force for intermixing. Equilibrium Ge distributions are obtained by direct simulation where, in addition to the usual MC moves, randomly selected pairs of Si and Ge atoms can exchange positions. All Ge atoms are initially placed in the top layers. Some additional runs were performed to check that the same average distribution is obtained regardless of the initial state (e.g., we can start with a random mixture of Si and Ge atoms or cool the system down from a higher temperature). We can now estimate how strain-relief driven intermixing affects the relative stability of the different $2\times n$ reconstructions. This is done by calculating the energy difference between the intermixed state and the system with all Ge placed in the top layers. Secondly, we consider how temperature affects the intermixed distributions. The surface intermixing distributions are obtained at various temperatures and the results are compared with experimental measurements and with previous theoretical calculations.

1. Strain relaxation

Let us now consider the low-temperature limit where Si-Ge intermixing is driven by strain relaxation. The purpose is to investigate the energetical reasons for intermixing and its effect on the optimal periodicity of the $2\times n$ reconstruction. Our MC simulations do not have kinetic limitations, but

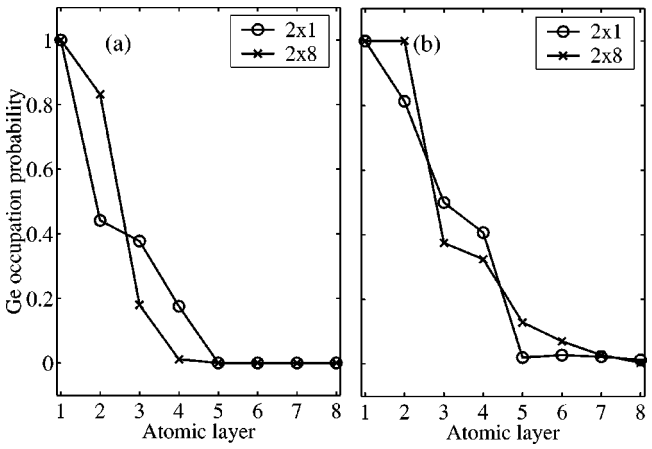

FIG. 4. Strain-relief driven intermixing: Ge occupation probability as a function of distance from the surface (in atomic layers) for (a) 2 ML Ge coverage and (b) 3 ML Ge coverage. Data for 2×1 and 2×8 reconstructed surfaces are shown. The distributions were obtained at $T=11$ K.

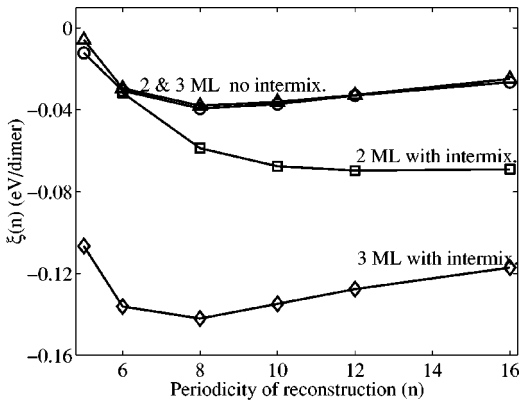

FIG. 5. Effect of strain-relief driven intermixing on the relative stability of different $2\times n$ reconstructions. Results for 2 and 3 ML coverages are shown. The data for the case without intermixing are the same as shown in Fig. 2. For coverages up to 1 ML, the surface structure does not change because all Ge remains in the top layer at $T=11$ K.

equilibration times become longer at low temperatures because once a near-equilibrium configuration is formed, it becomes increasingly unlikely that a favorable Si-Ge pair is randomly selected for the attempted exchange. Consequently, reaching full equilibrium requires simulation runs up to 200 000 MC steps (averages are calculated over the last 20 000 steps).

In Figure 4 we show the average distribution of Ge atoms as a function of distance from the surface for 2 and 3 ML Ge coverages and for 2×1 and 2×8 reconstructed surfaces. The temperature is $k_BT=0.001$ eV (11 K), and thus the distributions correspond to configurations where Ge atoms occupy lattice sites with lowest energy. Coverages of 1 ML and below are not shown for clarity because there all Ge remains in the top layer due to its lower surface energy. For the larger coverages, we notice that the Ge distribution is dependent on both the thickness of the Ge layer and the reconstruction of the surface. In addition, Ge atoms show clear preference to certain lattice sites due to local variations in subsurface strain which are caused by the surface reconstruction (see Sec. III B 2 for details). The stepped shape of the distribution curves follows from this strong site preference and from the fact that these distributions are obtained at a very low temperature where entropic effects do not smooth out the differences in energy.

For 2 ML coverage, less than 20% of the second-layer Ge atoms have moved to the third layer when the surface is 2×8 reconstructed, compared to over a half in the 2×1 case. This indicates that the $2\times n$ reconstruction is effective in relieving large part of the strain in system, and thus the need for intermixing decreases when compared to the 2×1 surface. On the other hand, it can be energetically more favorable to increase the periodicity of the reconstruction while having more intermixing in the system. Figure 5 shows that this is indeed the case: for 2 ML coverage the optimal value of n shifts from $n=8$ to larger values when intermixing occurs (a very weak minimum is observed at $n=12$, but taking into account the accuracy of the calculations, the same energy values are obtained for $n=12$ and 16). STM measurements[18] show similar behavior: for 1.5 ML of Ge grown at a lower temperature, post-deposition annealing at 760 °C changes the observed value of n from 9 to 14. The annealing enables the system to overcome at least part of the kinetic barriers affecting subsurface diffusion of Ge, and consequently the periodicity of the vacancy-line structure changes as a response to the strain relaxation provided by intermixing.

For 3 ML coverage, Fig. 4 shows that the second layer is fully occupied by Ge provided that the surface is 2×8 reconstructed, whereas large part of third-layer atoms have moved to the four substrate layers beneath the Ge overlayer. The Ge migration into layers 5–7 is enabled by the vacancy lines on the surface because such accumulation does not occur in the 2×1 reconstructed system. Figure 5 shows that for the 3 ML coverage, intermixing leads to a large gain in energy but does not affect the optimal periodicity of the $2\times n$ reconstruction. This is due to the fact that intermixing provides additional strain relief which cannot be accomplished by further reducing the periodicity of the reconstruction. In experimental systems, the degree of intermixing increases rapidly as a function of temperature. Thus based on the need for additional strain relief, it is possible that intermixing could change the surface morphology in the 2–3 ML regime by postponing the onset of 2D islanding and the formation of the $m\times n$ pattern.

2. Site selectivity

Now we will discuss the site selectivity of Si-Ge intermixing which is induced by the two main factors: dimer rows and vacancy lines. As an example, in Fig. 6 we show a snapshot of a typical configuration obtained from a simulation of intermixing in a system composed of 3 ML of Ge on Si(001). In this we observe that the region under the vacancy lines is clearly unfavorable to Ge occupation. This is due to the fact that the vacancy lines induce atomic displacements in their vicinity, which in turn leads to a local compressive

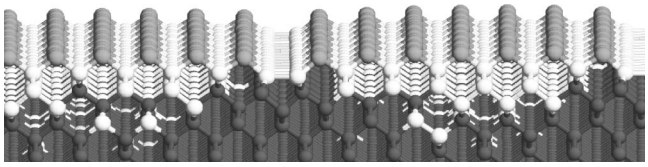

FIG. 6. Snapshot of a typical configuration obtained as a result of intermixing in a system with 3 ML of Ge on Si(001). Si atoms are shown in dark gray and Ge atoms in lighter gray (dimers) or white (underlying layers). The view direction is along the *x* axis shown in Fig. 1 (dimer bonds are directed toward the viewer).

strain in this region. In addition, the dimer rows produce similar local fluctuations in the strain field which results in alternating Si and Ge atoms in the third and fourth layers. Consequently, the optimal surface structure is determined by a complex interplay of vacancy lines and subsurface distribution of Ge.

Figure 7 shows typical lattice sites which are favorable to Ge occupation. The degree of Ge occupation in the various sites depends on the amount of Ge on the surface (e.g., only first layer sites are occupied at submonolayer coverages). In Fig. 7(a) the surface is viewed from the side along the direction of the dimer rows. The dimerization of the atoms in the surface layer results in significant distortions in the positions of the atoms in subsurface layers. Consequently, the third and fourth layer sites under the troughs (labeled sites 3b and 4b) are clearly more favorable to Ge occupation than are the sites under the dimer rows (labeled 3a and 4a). This preference for sites 3b and 4b was also predicted by the empirical Tersoff potential;[22] but in contrast to those results which show a strong reduction in the second-layer Ge concentration, we find that, for the SW potential, the Ge occupation in layer 2 is clearly above that in layers 3 and 4 (counting both a and b sites in layers 3 and 4). Our results are in agreement with recent density functional theory (DFT) calculations[16,23] which also predict Ge occupation in the second layer and strong preference to sites under the troughs in layers 3 and 4.

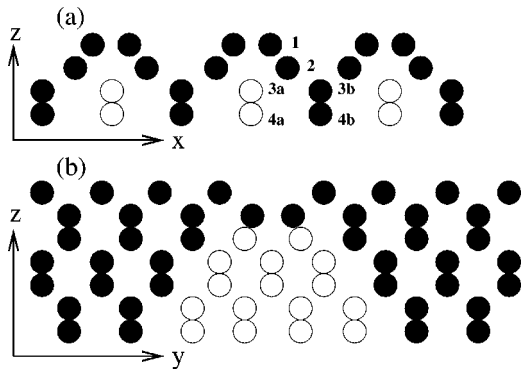

FIG. 7. Site selectivity in intermixing (schematic). Lattice sites which are under tensile strain and thus favorable for Ge atoms are marked with filled circles and sites under compression and thus favoring Si occupation are marked with open circles. (a) Side view of the surface along the *y* axis (in the direction of the dimer rows). (b) Side view of the surface along the *x* axis (in the direction of the dimer vacancy lines). See Fig. 1 for top view of the surface showing the coordinates.

TABLE I. Energy cost of substituting a Ge atom for Si in different sites near the Si(001) surface relative to the first layer (in eV). For the DFT results the difference between sites 1a (buckling-up) and 1b (buckling-down) is also given. The DFT calculations are from Ref. 16. Results using both the local density approximation (LDA) and the PW91 exchange-correlation function are quoted. The surface is 2×1 reconstructed in all cases.

Layer	SW	Tersoff	LDA	PW91
1a	0	0	0	0
1b			0.134	0.149
2	0.244	0.230	0.314	0.363
3a	0.289	0.133	0.361	0.415
3b	0.183	0.121	0.292	0.344
4a	0.274	0.173	0.344	0.419
4b	0.199	0.103	0.291	0.348

This picture is further supported by several experiments[12–16] which indicate Ge diffusion down to at least the fourth layer with no reduction in the second-layer occupation.

In Fig. 7(b) the surface is viewed from the other side along the direction of the dimer bonds. A dimer vacancy line in the $2\times n$ reconstruction is visible on the surface. The formation of missing dimers causes significant movement of a large number of atoms in the vicinity of the vacancy line as the exposed second-layer atoms rebond in order to reduce the number of dangling bonds. The inward relaxation leads to a compressively strained region near the vacancy line which becomes unfavorable to Ge occupation. The effect of the vacancy line is reflected down to at least the seventh layer beneath the surface. Note that energetically less favorable sites in deeper layers are only occupied after the total amount of Ge becomes sufficiently high. [See Fig. 6 as an example showing 3 ML of Ge on Si(001)]. On a 2×1 surface, there is no preference to sites in layers 5–7 due to the absence of dimer vacancy lines as shown by the occupation probabilities in Fig. 4.

In order to be able to directly compare our results to earlier studies, we used a static energy minimization method to calculate the energy cost of substituting a Ge atom in different sites near the Si(001) surface. For comparison purposes, we calculated the binding energies using both the SW and the Tersoff[7] potentials. Table I summarizes the results obtained using the empirical potentials (SW and Tersoff) and earlier DFT results. All calculations show that the first layer is most favorable to Ge atoms. The compressively strained sites under the dimer rows (3a and 4a) are clearly higher in energy than are the tensile sites under the troughs (3b and 4b). The difference comes when the binding energy of the second-layer sites is compared to the energies of the compressive sites 3a and 4a. The Tersoff potential predicts that the second-layer sites are higher in energy than both the tensile and compressive sites in layers 3 and 4. In contrast, the SW potential gives qualitatively similar results to the DFT calculations which give an energy value in between the a and b sites in layers 3 and 4. As mentioned earlier, experiments support the latter result.

We note that the binding energies shown in Table I do not include correlation effects between multiple Ge atoms or effects resulting from the growing strain when several Ge layers are placed on the surface. In addition, significant effects due to the $2\times n$ reconstruction are disregarded in these calculations. In contrast, the strength of the MC approach is that all these factors are directly included in the simulation, and as an outcome we obtain the equilibrium distribution for a given Ge coverage. The occupation probabilities in Fig. 4 give the low-temperature limit of the distribution. We now extend the study to higher temperatures in order to include entropic effects.

3. Effect of temperature

In our MC simulations, the movement of the Ge atoms is not limited by kinetic factors as in real systems since we do not attempt to mimic the actual diffusion mechanisms. In the full thermodynamic limit, where intermixing is allowed to occur in the whole system, Si/Ge alloy is formed in the bulk region with constant average ratio of Si and Ge. The fraction of Ge in the bulk alloy increases linearly as a function of the total amount of Ge in the system after initial accumulation to surface areas (most Ge will stay near the surface when the total amount of Ge is very low). In addition, the distribution of Ge is also affected by temperature: the lower the temperature, the higher is the Ge occupation in surface layers and consequently the bulk fraction decreases; e.g., for 20% total Ge content, the bulk fraction of Ge is $\sim$0.16 at 600 °C, $\sim$0.15 at 300 °C and only $\sim$0.02 close to the static limit ($kT=0.001$ eV).

In real systems, bulk diffusion of Ge is practically negligible at usual growth conditions, but as mentioned, several experiments have indicated that near the surface Ge penetrates down to at least the fourth layer at temperatures as low as 400°C. The physical origin of this much higher diffusion rate in the surface region is explained by the different nature of diffusion mechanisms near the surface. Ge diffuses in bulk Si by two main mechanisms: substitutional-interstitial exchange and vacancy diffusion. The exchange process, in which Ge diffusion is assisted by a Si interstitial, has been shown to be clearly faster in the surface region because the formation energy of Si interstitials is significantly lower near the surface than in the bulk.[16] Thus, in the MC simulations, we must limit the intermixing to a physically suitable region near the surface, which will allow us to study the quasiequilibrium situation where the surface layers are in thermodynamic equilibrium but no Si-Ge exchange to deeper layers occurs. Following both experimental evidence in Refs. 12–16 and the DFT calculations in Ref. 16, we chose the intermixing region to extend four or five layers below the surface depending on the Ge coverage.

In Fig. 8 we show the occupation probabilities of Ge in layers 1–4 as a function of Ge coverage at (a) 76 °C and (b) 600 °C. A 39-layer slab was used for these simulations because it was found to be sufficient for distribution averages. The data show that the total amount of intermixing increases significantly as the temperature is raised. At the lower temperature, the third and fourth layers become populated by Ge atoms only after 1 ML coverage is reached. At the higher

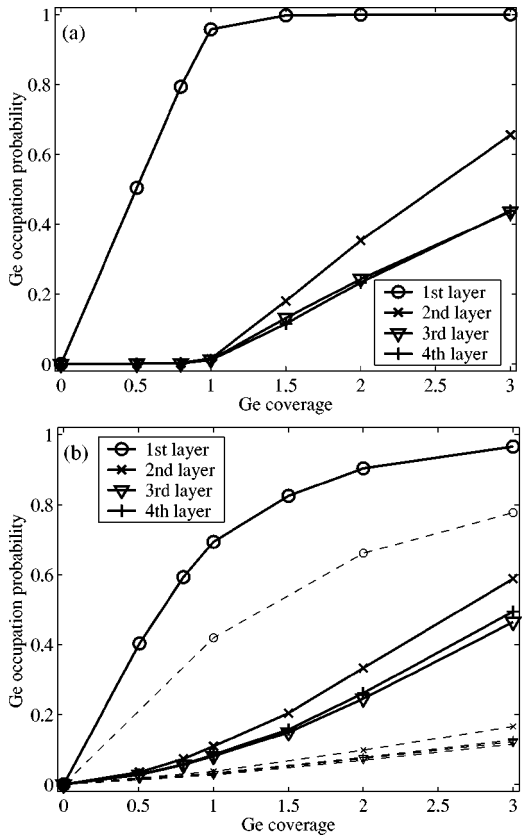

FIG. 8. Ge occupation probabilities in each atomic layer as a function of Ge coverage obtained from MC simulations at (a) $k_BT=0.03$ eV (76 °C) and (b) $k_BT=0.075$ eV (600 °C). The surface is 2×10 reconstructed. The intermixing region covers layers 1–4 (up to 1 ML Ge coverage) or layers 1–5 (for larger coverages). The dashed lines in (b) show the effect of allowing intermixing to occur in the whole bulk region.

temperature, a notable amount of Ge is present throughout the surface region with third and fourth layers becoming populated even at low coverages. This agrees well with experiments[12–14,16] which show Ge occupation down to at least the fourth layer with second layer occupation comparable to the occupation of the tensile sites in layers 3 and 4. In addition, the dashed lines in Fig. 8(b) show the effect of allowing intermixing to occur in the whole system. (Note that here the results depend on the thickness of the simulation slab which determines the total amount of Ge in the system. Here the total amount of Ge is $\sim$15% for the 3 ML coverage.) Figure 8(b) can also be compared directly to a corresponding figure in Ref. 15 which shows the results for experimental measurements at 500 °C for 0.2–1.5 ML Ge coverages. The shape and order of the curves in the two figures is very similar, showing clearly largest Ge concentrations in the first and second layers with the growth rate decreasing in the first layer and increasing in the second layer. Within the accuracy of both of these studies, the calculated and measured concentrations are well of the same order.

Experiments have also indicated that Ge occupation is larger in the tensile sites in layers 3 and 4 (3b, 4b).[16] We have already seen that the energy of substituting a Ge atom in the different sites in layers 3 and 4 is clearly lower for the

TABLE II. Fraction of Ge atoms occupying tensile b sites in layers 3 and 4. The surface is 2×10 reconstructed. The statistical errors are of the order 0.001.

	76 °C			600 °C		
site	0.8 ML	1.5 ML	2.0 ML	0.8 ML	1.5 ML	2.0 ML
3b	1.000	0.951	0.922	0.820	0.785	0.747
4b	1.000	0.861	0.815	0.696	0.666	0.644

tensile b sites than for the compressively strained a sites. At very low temperatures (e.g., for the data in Fig. 4 obtained at 11 K), only tensile sites in layers 3 and 4 are occupied by Ge. Table II shows the fraction of Ge atoms in layers 3 and 4 which occupy tensile b sites. Data for a range of Ge coverages at two different temperatures is given. Note that the total amount of Ge in the 3rd and 4th layers depends strongly on the total Ge coverage. At low coverages there are very few Ge atoms in the 3rd and 4th layers and at low temperatures they occupy mainly the energetically more favorable b sites. At higher coverages, the amount of Ge in the 3rd and 4th layers increases and consequently some Ge atoms are driven to occupy the compressive a sites even at low temperatures. At higher temperatures the amount of Ge atoms in the a sites increases for all coverages, but in agreement with the experimental observations, a majority of the Ge atoms remain in the tensile b sites even at 600 °C. This leads to an ordered structure in subsurface layers as has been observed in growth experiments for Si/Ge alloys on Si(001).[24]

We would like to point out that detailed quantitative comparison of calculations and experimental data is not viable at the moment due to rather large variations in different sets of experimental measurements. These variations are explained in part by the difficulties in obtaining experimental data on subsurface intermixing and in part by different experimental conditions. In addition, kinetic factors are not included in the theoretical calculations, but they may have a significant influence on the experimental measurements.

In several DFT studies, the Ge occupation probabilities have been obtained by first calculating the energy of substituting a single Ge atom at the various sites in subsurface layers and then using either the Fermi-Dirac distribution (number of Ge atoms in the unit cell not fixed)[16,25] or the Boltzmann distribution (fixed number of Ge atoms in the unit cell)[23] to calculate the occupation probabilities at different temperatures. Two main simplifications are made in these studies: (i) The coverage dependence of the Ge binding energy is neglected (the occupation probability of a given site is assumed not to depend on the configuration of other sites) and (ii) $2\times n$ reconstruction is not included. Table III shows a comparison of our results and results obtained in the three DFT studies mentioned above. Two different cases are considered: 1.0 ML of Ge at 600 °C and 1.6 ML at 700 °C. The surface is 2×1 reconstructed in all cases. The agreement is found to be fairly good, especially when taking into account the differences between simulation methods and the techniques of obtaining the occupation probabilities. Furthermore, in Ref. 16 the DFT results for the 1.6 ML case were compared with Auger electron diffraction (AED) measurements of a corresponding experimental system and the results were found to be in good qualitative agreement. Since the differences between the DFT and our data are very small, our results would give similar agreement with the experiment. The main difference between our results and the different DFT calculations is that we obtain slightly lower Ge occupation in the second layer and slightly higher occupation in the third and fourth layers, which could be due to Ge-Ge correlation effects.

TABLE III. Comparison of Ge occupation probabilities in subsurface layers obtained using the SW potential based MC method (current study) and DFT calculations from Refs. 25 (DFT1), 23 (DFT2) and 16 (DFT3). The first set of results (DFT1, DFT2, SW1) is for 1.0 ML Ge coverage at 600 °C and the second set (DFT3, SW2) is for 1.6 ML coverage at 700 °C. The surface is 2×1 reconstructed in all cases. The statistical errors for the SW data are of the order 0.001.

	1.0 ML at 600 °C			1.6 ML at 700 °C	
Layer	DFT1	DFT2	SW1	DFT3	SW2
1	0.705	0.648	0.694	0.893	0.836
2	0.111	0.131	0.095	0.257	0.251
3a	0.067	0.028	0.041	0.156	0.137
3b	0.133	0.082	0.168	0.303	0.367
4a	0.048	0.032	0.063	0.150	0.183
4b	0.121	0.078	0.149	0.293	0.338

IV. CONCLUSION

In this paper, we have investigated the two principal strain-relief mechanisms, namely, reconstruction and Si-Ge intermixing, in a heteroepitaxial system of Ge on the Si(001) surface. Strain-relief related phenomena, such as the rearrangement of the reconstruction and Si-Ge intermixing, occur over large distances, and can be studied with an empirical-potential based MC method for much larger systems than those feasible to *ab initio* or TB calculations. As a result, we have been able to investigate the combined effects of intermixing and reconstruction at a Ge-covered Si(001) surface.

In close agreement with experimental observations, our results show that the surface undergoes a structural evolution as the thickness of the Ge layer increases. In the absence of Si-Ge intermixing (e.g., at low growth temperatures), the optimal periodicity of the $2\times n$ reconstruction first decreases as Ge is deposited on the surface and then saturates to $n=8$ at about 2 ML coverage. The effect of intermixing on the optimal reconstruction is found to be strongly dependent on the thickness of the Ge overlayer. At about 2 ML coverage, most of the strain can be relieved either by formation of dimer vacancy lines ($2\times n$ reconstruction) or by Si-Ge intermixing. Thus the optimal periodicity of the $2\times n$ reconstruction increases as a result of intermixing, as also observed experimentally. For coverages above 2 ML, the need for additional strain relaxation grows further. Consequently the optimal periodicity of the reconstruction remains unchanged but intermixing leads to a large energy gain. In addition, the surface

reconstruction induces atomic-scale variations in the strain field near the surface, which in turn leads to a strongly site selective distribution of Ge in the subsurface layers.

This study has focused on investigating the thermodynamic driving forces behind structural changes during the growth of Ge on Si(001). We have shown that the evolution of the surface is characterized by an interplay of different strain-relief mechanisms. In addition, kinetic effects play a significant role during growth which is inherently a nonequilibrium process. In this work, we have considered two limiting cases concerning intermixing: full segregation of Ge and full thermal equilibrium in surface layers. Most experiments are likely to lie somewhere in between these two limits. Nevertheless, being in good agreement with the trends observed in experiments, our study indicates that the MC approach with a good interatomic potential could provide a plausible framework for describing the incipient growth of Ge on Si(001).

ACKNOWLEDGMENTS

We would like to thank B. P. Uberuaga for helpful comments. This work was supported in part by the Academy of Finland, project on Computational Research of Semiconductor Materials, Project No. 1169043 (Finnish Center of Excellence Program 2000-2005), and in part by NSF, Grant No. DMR-0094422. L.N. would also like to acknowledge support of the Finnish Cultural Foundation and the Finnish Academy of Science and Letters.

[1] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[2] C.M. Goringe, D.R. Bowler, and E.H. Hernández, Rep. Prog. Phys. **60**, 1447 (1997).
[3] F.H. Stillinger and T.A. Weber, Phys. Rev. B **31**, 5262 (1985).
[4] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000).
[5] M. Laradji, D.P. Landau, and B. Dünweg, Phys. Rev. B **51**, 4894 (1995).
[6] L. Nurminen, F. Tavazza, D.P. Landau, A. Kuronen, and K. Kaski, Phys. Rev. B **67**, 035405 (2003).
[7] J. Tersoff, Phys. Rev. B **39**, 5566 (1988).
[8] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[9] F. Tavazza, L. Nurminen, D.P. Landau, A. Kuronen, and K. Kaski (unpublished).
[10] J. Oviedo, D.R. Bowler, and M.J. Gillan, Surf. Sci. **515**, 483 (2002).
[11] B. Voigtländer, Surf. Sci. Rep. **43**, 127 (2001).
[12] M. Sasaki, T. Abukawa, H.W. Yeom, M. Yamada, S. Suzuki, S. Sato, and S. Kono, Appl. Surf. Sci. **82-83**, 387 (1994).
[13] A. Ikeda, K. Sumitomo, T. Nishioka, T. Yasue, T. Koshikawa, and Y. Kido, Surf. Sci. **385**, 200 (1997).
[14] H.W. Yeom, M. Sasaki, S. Suzuki, S. Sato, S. Hosoi, M. Iwabuchi, K. Higashiyama, H. Fukutani, M. Nakamura, T. Abukawa, and S. Kono, Surf. Sci. **381**, L533 (1997).
[15] K. Nakajima, A. Konishi, and K. Kimura, Nucl. Instrum. Methods Phys. Res. B **161-163**, 452 (2000).
[16] B.P. Uberuaga, M. Leskovar, A.P. Smith, H. Jonsson, and M. Olmstead, Phys. Rev. Lett. **84**, 2441 (2000).
[17] X.R. Qin, B.S. Swartzentruber, and M.G. Lagally, Phys. Rev. Lett. **84**, 4645 (2000).
[18] F. Liu, F. Wu, and M.G. Lagally, Chem. Rev. (Washington, D.C.) **97**, 1045 (1997).
[19] K. Li, D.R. Bowler, and M.J. Gillan, Surf. Sci. **526**, 356 (2003).
[20] B. Voigtländer and M. Kästner, Phys. Rev. B **60**, R5121 (1999).
[21] F. Liu and M.G. Lagally, Phys. Rev. Lett. **76**, 3156 (1996).
[22] P.C. Kelires and J. Tersoff, Phys. Rev. Lett. **63**, 1164 (1989).
[23] J.-H. Cho and M.-H. Kang, Phys. Rev. B **61**, 1688 (2000).
[24] F.K. LeGoues, V.P. Kesan, S.S. Iyer, J. Tersoff, and R. Tromp, Phys. Rev. Lett. **64**, 2038 (1990).
[25] Y. Yoshimoto and M. Tsukada, Surf. Sci. **423**, 32 (1999).