

Atomic theory of scanning tunneling microscopy

E. Tekman and S. Ciraci

Department of Physics, Bilkent University, Bilkent, 06533 Ankara, Turkey

(Received 11 April 1988; revised manuscript received 3 February 1989)

We present a quantitative analysis of the modifications of the scanning-tunneling-microscopy images due to the local perturbations of the electronic states induced by the tip in close proximity to the sample surface. Using an empirical tight-binding method, we have calculated the electronic states of a prototype tip-sample system consisting of a single-atom tip and the graphite surface, as a function of the tip-sample distance. We find that as the tip approaches the sample, their states start to interact and become laterally confined in the vicinity of the tip at small tip-sample separation. These states influence the tunneling phenomenon by connecting the tip and sample surface electronically. The effect of the tip-induced localized states is discussed, and the expression for the tunneling current is reformulated by incorporating the tip-induced states. Calculations using this expression show that the corrugation amplitude obtained from scanning tunneling microscopy is enhanced and deviates from the proportionality to the local density of states of the bare sample at the Fermi level evaluated at the center of the tip.

I. INTRODUCTION

Real-space imaging capability and atomic-scale resolution are unique features that make scanning tunneling microscopy¹ (STM) a powerful technique in the analysis of surfaces. In the current theory of STM, Tersoff and Hamann² replaced the many-particle wave functions in the Bardeen formalism³ of tunneling by the one-electron states of infinitely separated electrodes. Using this approximation the expression of the tunneling current can be cast into a simple form, and is found² to be proportional to the local density of states of the bare sample at the Fermi level evaluated at the center of the tip, $\rho(\mathbf{r}_0, E_F)$; in the usual STM experiments the tip-to-surface distance, d , is large and thus this theory with the noninteracting electrodes has been used with reasonable success.⁴⁻¹¹ Recent experiments,¹²⁻¹⁴ however, have indicated that the tip-sample interaction is an essential aspect of STM. The STM study of the graphite surface at a very small bias voltage¹² has indicated an elastic deformation of the surface caused by the tip. In this case a tip-to-surface distance as small as ≈ 2 Å is conjectured, whereby a local and strong interaction between the tip and surface sets in, which leads to huge corrugations. Self-consistent force and charge-density calculations by Ciraci and Batra¹⁵ have justified the strong interactions induced by the tip. They showed that at small $d \approx 1.5$ Å the tunneling barrier recedes and a point contact through a chemical bond between the tip and surface atoms is formed. Recently observed STM corrugations on the close-packed (111) surface of the noble¹⁶ and simple¹⁷ metals, which are much larger than the corrugation amplitude obtainable from $\rho(\mathbf{r}_0, E_F)$, provide evidence also for the strong tip-sample interaction, whereby the commonly accepted proportionality to the electronic structure of the bare sample surface is no longer valid. At present the interaction of the tip with the substrate atoms has become the focus of attention on account of new

areas of application brought about by the recent studies^{13,14} of STM combined with atomic force microscopy.¹⁸

As the tip approaches the sample surface, the tip-sample interaction gradually increases, and the potential barrier is lowered. The charge density undergoes a redistribution, and the ions of the tip and sample are displaced from their original positions to attain the lowest total energy. It is therefore expected that the STM images are affected by these modifications over the electronic and atomic structure of bare sample and bare tip. This study presents the first quantitative analysis of the modifications of the STM images induced by the local perturbations of the electronic states. In the first part of the paper, we have investigated the effect of the tip-sample interaction on the electronic states. Based on the empirical tight-binding (ETB) calculations on a prototype graphite surface, we demonstrate that the tip interacting with the sample surface induces localized (or resonance) states. That the tip-induced localized states (TILS) are formed at the vicinity of the tip was pointed out first by ourselves, and reported elsewhere as a preliminary result of the present study.¹⁹ In the second part of this paper we study how the STM images are influenced by the local perturbations in the electronic structure. To this end we reformulate the expression of the tunneling current to include TILS. Calculations using this new expression show that the corrugation amplitude obtained from STM deviates from that related to $\rho(\mathbf{r}_0, E_F)$.

Based on the present analysis we are able to identify three ranges for the tip-sample distance, which lead to three different regimes in the operation of STM for the graphite sample. For large d ($d \gtrsim 4$ Å) the tip-sample interaction is insignificant, and may be represented by the sum of the attractive van der Waals energy and repulsive Pauli exclusion energy yielding a weak attractive interaction. The theory,² which is based on the Bardeen transfer Hamiltonian approach³ and yielding that the tunneling

current is proportional to $\rho(\mathbf{r}_0, E_F)$, can be used safely in this range of d . This mode of operation is called the normal tunneling (or nearly independent electrodes) mode of STM. For relatively smaller d ($2 \lesssim d \lesssim 3.5$ Å) TILS become pronounced, and thus the tunneling current is modified. The mode corresponding to d where TILS are effective is called the TILS mode. Upon further decreasing d , the potential barrier between the tip and sample is perforated by an orifice, whereby the character of TILS changes by the enhanced localization near the dividing plane, and by their energies lying below the Fermi level. The mode of operation corresponding to very small d which is comparable with the interatomic distance ($d \lesssim 2$ Å) is the point-contact mode. In this mode the nature of the conductance is rather different from that occurring at normal tunneling mode, and can be described as a quantum conductance through the channels of the constriction states below E_F . A detailed analysis of this mode will be reported subsequently.

In Sec. II we explain the model for the tip-sample system and the details of the ETB method as used in our study. The results of these calculations are discussed in Sec. III with the emphasis placed on TILS. In Sec. IV we explain the details of the formalism of the tunneling current, in which TILS are incorporated. The enhancement of the corrugation amplitude to be obtained from STM operating in the TILS mode is also exemplified for a typical case in the same section. Our findings are summarized in the concluding section.

II. ELECTRONIC STRUCTURE OF THE TIP-SAMPLE SYSTEM

It is known that the bulk graphite has a layered structure with a weak interlayer coupling. Since the atomic positions in the consecutive layers are shifted, two inequivalent atomic sites occur in a given layer. These sites are usually denoted as the A and B sites. The A site has carbon atoms directly below and above it in the adjacent layers, whereas the B site does not. In each layer one type of site has three nearest neighbors of the other type site, and thus hexagons are formed from the alternating sites. The centers of these hexagons are denoted as the H site. While an individual graphite layer (unsupported monolayer) has sixfold rotation symmetry with zero charge density at the H site and with the Fermi surface collapsed to a single point at the K point of the Brillouin zone (BZ), the symmetry is lowered to threefold rotation, and the Fermi surface becomes small pockets in the graphite slab. This symmetry lowering due to interlayer interaction has been resolved by STM operating at large d , as such that only three protrusions out of six atomic sites of the surface hexagon (which are formed by the A and B sites) have been observed.²⁰ Based on the Tersoff-Hamann theory² it was argued that¹⁰ the B sites, for which $\rho(\mathbf{r}_0, E_F)$ is larger than that of the A sites, are more likely to be probed by STM. By contrast, both A and B sites can be resolved equivalently in STM (Ref. 21) operating at small d . This implies that the atomic scale interactions occurring between the tip and sample at small d dominate the tunneling current, so that the weak

interlayer coupling distinguishing two different sites becomes only a secondary effect. Since the main features of the electronic structure of graphite are determined by that of a single monolayer, and since the present study deals with a small tip-to-sample distance, the graphite sample is represented by a graphite monolayer in our model. We note that in this simple model the A and B sites become identical and are denoted as the on-top site (or T site) in the rest of the paper. In view of the computational limitations, calculations are performed with periodically repeated supercells. The tip is represented by a single carbon atom, which is periodically repeated in a (3×3) array. Then the whole system is treated in a (3×3) supercell consisting of one carbon atom representing the tip, and nine graphite monolayer cells representing the sample. In this model the repeat period of the tip atoms is 7.25 Å, which is large enough to lead to negligible intertip interaction. The artificial periodic-boundary conditions allow us to use the Bloch sums constructed from the orbitals of carbon atoms. The representation of the tip by a single carbon atom is appropriate and suits the purpose of this study because the sample interaction is determined mainly by the outermost atom of the extended tip at small d .

The electronic band structure of the unsupported graphite monolayer is calculated by using the ETB method. Formally, the wave functions are constructed from the Bloch sums:

$$\chi_{ij}(\mathbf{k}, \mathbf{r}) = \frac{1}{\sqrt{N}} \sum_n e^{i\mathbf{k} \cdot (\mathbf{R}_n + \boldsymbol{\tau}_j)} \phi_i(\mathbf{r} - \mathbf{R}_n - \boldsymbol{\tau}_j), \quad (2.1)$$

where ϕ_i stands for the $2s$ - and $2p_{x,y,z}$ -type Wannier orbitals of the carbon atoms with the position vector $\boldsymbol{\tau}_j$ in the unit cell. \mathbf{R}_n is the Bravais lattice vectors of the hexagonal monolayer lattice. In ETB no specific use of orbitals is made, but the energy parameters for the on-site and nearest-neighbor matrix elements in the secular equation $[\underline{H}_0(\mathbf{k}) - \underline{I}E_0(\mathbf{k})]\underline{a}_0(\mathbf{k}) = 0$ are fitted to the existing band structure. To this end, we used the band energies calculated by Tatar and Rabii²² and determined the energy parameters, which are listed in Table I with the notation given by Slater and Koster.²³ The in-plane orbitals of carbon atoms (ϕ_{2s} , ϕ_{2p_x} , and ϕ_{2p_y}) hybridize into sp^2 -hybrid orbitals and form σ and σ^* bands. The ϕ_{2p_z} orbitals, in turn, form π and π^* bands. In the monolayer geometry these two sets of states are decoupled owing to the reflection symmetry about the layer plane. The Fermi level occurs at the K corner of the graphite BZ, where the π and π^* bands cross and $E_F = -8$ eV.

TABLE I. The empirical tight-binding parameters for the graphite monolayer fitted to the band-structure calculations of Tatar and Rabii (Ref. 22). The notation is taken from Slater and Koster (Ref. 23). Superscript 0 (1) designates the on-site (nearest-neighbor interaction) energies. All parameters are in eV.

E_{ss}^0	$E_{pp\pi}^0$	E_{ss}^1	$E_{sp\sigma}^1$	$E_{pp\sigma}^1$	$E_{pp\pi}^1$
-10.73	-6.13	-5.41	5.59	-2.02	5.84

Having determined the energy parameters of the graphite monolayer, we next consider the tip and sample system in the (3×3) supercell. The tip being represented by a single carbon atom, its orbitals $\phi_{\text{tip}}(\phi_{2s}, \phi_{2p_{x,y,z}})$ interact mainly with the π and π^* states of the monolayer. The interaction energy is given in terms of the matrix element $\langle \phi_{\text{tip}} | \underline{H}_{T+S} | \phi_s \rangle$ of the tip-sample Hamiltonian \underline{H}_{T+S} with respect to the tip (ϕ_{tip}), and the sample (ϕ_s) orbitals. Depending upon the position of the tip atom above the monolayer, these interaction energies may deviate from the nearest-neighbor matrix elements of the bare graphite monolayer given in Table I. Then the interaction energies are determined by scaling the nearest-neighbor matrix elements of the bare graphite with Slater-type orbital functions:

$$\langle \phi_{\text{tip},i} | \underline{H}_{T+S} | \phi_{s,j} \rangle = \langle \phi_{s,i} | \underline{H}_0 | \phi_{s,j} \rangle \exp[-\mu_{ij}(r_{ij} - a_0)], \quad (2.2)$$

where r_{ij} and a_0 are the internuclear distances in the right-hand and left-hand side integrals, respectively. The exponents of the Slater-type orbitals are obtained from the wave functions given by Clementi and Raimondi.²⁴ These are $\mu_{2s,2s} = 0.84 \text{ \AA}^{-1}$, $\mu_{2p,2p} = 0.82 \text{ \AA}^{-1}$, $\mu_{2s,2p} = (\mu_{2s,2s} + \mu_{2p,2p})/2$. In the past, such a scaling of the interaction parameters has been applied to other systems successfully.⁸ As for the on-site matrix elements (or self-energies of the tip orbitals), $E_{ii}^0 = \langle \phi_{\text{tip},j} | \underline{H}_{T+S} | \phi_{\text{tip},i} \rangle$, they are shifted to coincide with the Fermi level of the extended tip, which is biased relative to the monolayer. This way, the features of the macroscopic tip are incorporated into our model, simulating the tip by a single atom.

III. TIP-INDUCED LOCALIZED STATES

The electronic energy structure is calculated for the system consisting of the graphite monolayer and the tip atom located above various places of the surface (T site, H site, and bridge site) with varying d . The bias voltage between the tip and the monolayer is taken to be infinitesimal. Depending upon the value of d , we can identify three different regimes (i.e., the point contact, TILS, and nearly independent electrode or normal tunneling regimes) which intermix in a wide range of the border value of d .

The interaction energy $\langle \phi_{\text{tip}} | \underline{H}_{T+S} | \phi_s \rangle$ decays exponentially with the increasing tip-sample distance, and is negligible for the tip heights $d \gtrsim 3.5 \text{ \AA}$. As a result, the band structure of the tip-sample system is just the projection of the tip bands on the band structure of the graphite monolayer. The tip and sample can be considered independent, and thus the formalism developed by Tersoff and Hamann² for the tunneling current works well. This regime is called the nearly independent electrode regime. On the other hand, for $d \lesssim 2.0 \text{ \AA}$, a strong chemical bond forms between the tip and sample atoms. In compliance with the results of the self-consistent field calculations,¹⁵ this small d region is identified as the chemical bond (or point-contact) regime, in which the tip and sample are connected by an orifice in the potential barrier. The

states localized at this orifice can be considered as the constriction states leading to the quantum conductance.²⁵ Therefore the character of the conductance in this regime undergoes a dramatic change. The point-contact regime itself has many interesting features^{13,14} such as the one-dimensional conductivity, but it is beyond the scope of the present study.

For $2.0 \lesssim d \lesssim 3.5 \text{ \AA}$ the interaction energy is not strong enough to form a chemical bond, but is significant to lead to the intermixing of the tip atom and monolayer states. Owing to the supercell used in the calculations, the K point of the graphite BZ is folded into the Γ point of the supercell BZ. The band structures of the tip-graphite monolayer system are shown in Fig. 1 for d ranging from 2 to 2.75 \AA . It is clear that the tip and sample states hybridize if $|E_s - E_{\text{tip}}|$ is small and the matrix element $\langle \psi_{\text{tip}} | \underline{H}_{T+S} | \psi_s \rangle$ is large. For the model described in Sec. II these two conditions are satisfied around the center of the (3×3) BZ (Γ point). At this point, we observe that the states TS_1 and TS_2 approach the Fermi level as d increases. The analysis of the wave functions of these states reveals that they have more weight in the vicinity of the tip, and have weak bonding and weak antibonding characters, respectively. Therefore, the states (TS_1 and TS_2 near Γ point) are identified as TILS, and the range of d , for which TILS are active, is called the TILS regime of STM. In Fig. 2, the energy of the band, TS_2 at the Γ point is shown as a function of d . It is clear that the

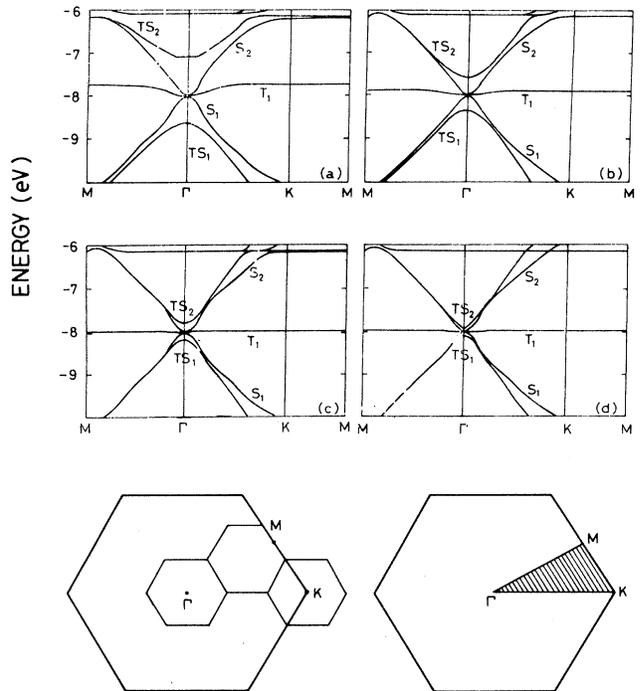


FIG. 1. Band structure for the tip located at the on-top position: For tip heights (a) 2.0 \AA , (b) 2.25 \AA , (c) 2.50 \AA , and (d) 2.75 \AA . The labeling is explained in the text. The graphite BZ and the 3×3 folding is shown in the lower left corner and the irreducible part of the supercell BZ is shown in the lower right corner.

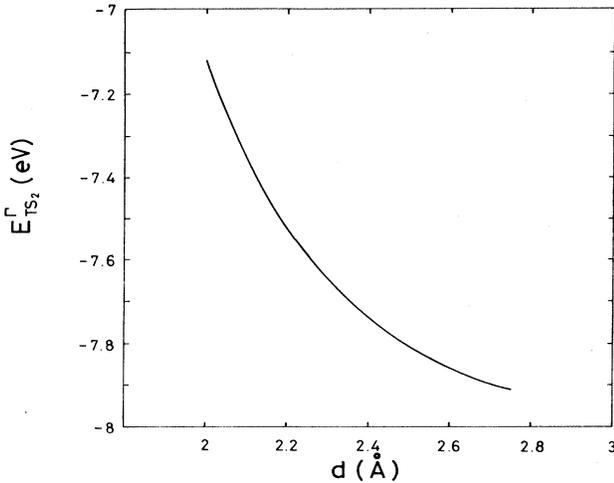


FIG. 2. Energy of TS_2 at the center of BZ (Γ) as a function of the tip-surface distance d . The tip is at the on-top position.

difference between the TILS energy and the Fermi level of the monolayer is an exponentially decaying function of d . The decay exponent is found to be 0.83 \AA^{-1} which is very close to the value of μ_{pp} used for scaling of the ETB parameters, confirming that the interaction is strongly localized in the vicinity of the tip atom.

The formation of TILS is more than a weak perturbation, and involves the mixing of the sample and tip states. Because of the potential barrier (which is lower than the vacuum level, but is still above the Fermi level) TILS have usually more weight either on the tip atom, or on the sample side, when their energies are close to the Fermi level. However, the TILS located far above the Fermi level may have significant weight across the barrier. If the tip-sample interaction gets stronger and the potential barrier is collapsed by a local orifice, certain TILS dip in the Fermi level and become strongly localized. This corresponds to the transformation of TILS into the constriction states, and thus to the transition from the TILS regime to the point-contact regime. Another important point we would like to emphasize is that TILS become resonance states when their energies lie in the band continua of the sample or the tip. This is the situation one encounters at the graphite sample. However, TILS are expected to be more localized, and to lead to more dramatic effects for a semiconductor sample.

In addition to the TILS, states (labeled as S_1 , T_1 , and S_2) which are dominated by one type of the electrode (tip or sample) are also visible in the energy bands shown in Fig. 1. At the center of the supercell BZ, three of the states at the Fermi level originate from the sample states only. In the absence of the tip, at the Fermi level, π and π^* states are degenerate and four linearly independent solutions can be written in the form

$$\psi_K(\mathbf{r}) = \left(\frac{2}{N} \right)^{1/2} \sum_n e^{i\mathbf{k}_K \cdot (\mathbf{R}_n + \boldsymbol{\tau}_j)} \phi_{p_z}(\mathbf{r} - \mathbf{R}_n - \boldsymbol{\tau}_j), \quad (3.1)$$

where the orbitals of the alternating sites of the hexagons in the graphite slab (either only A or only B sites) are in-

cluded in the summation and \mathbf{k}_K is the wave vector corresponding to one of the K points of graphite BZ. (It can be shown that only four of these six functions are linearly independent.²⁶) Upon the inclusion of the tip at the B site, three linear combinations of these states (two A -site states and the antisymmetric combination of the B -site states) are found such that they do not interact with the tip. The fourth linear combination (the symmetric combination of the B -site states) interacts with the tip orbitals and forms the TILS. This combination has the maximum amplitude on the atom just below the tip (i.e., for the on-top position of the tip), and the amplitude decreases away from this site. This shows that the tip-induced states are localized.

For the tip atom located at the H site the degeneracy of the π and π^* states do not split. This is a direct consequence of the symmetry of the model used to simulate the tip-sample system. The effect of the tip at the H site seems to be small for the carbon atom representing the tip. The tip atom located at the bridge site yields the effect similar to that at the B site. However, the strength of the interaction energy is smaller owing to the larger internuclear separation. In the present study, the effect of the tip on the electronic structure of the sample has been discussed only for the states near the Γ point of the supercell BZ, which are more likely to interact. However, important interaction may occur in the other regions of the BZ (for example, near the M point) depending upon the character of the tip atom²⁷ and the value of d as well.

IV. THE TUNNELING CURRENT IN THE PRESENCE OF TILS

In the theory of STM developed by Tersoff and Hamann,² the tip-sample interactions were disregarded, and thus the bare (unperturbed) sample and bare tip states were used. Consequently, the application of this theory is restricted to large d , which we identify as the normal tunneling regime. As revealed by the analysis presented in the preceding section, however, the atomic scale interactions become dominant leading to significant perturbations of some sample and tip states. We show that TILS form as a result of the tip-sample interaction. Therefore one expects that the tunneling current deviates from that predicted by the Tersoff-Hamann theory. In the past the formulation of the tunneling phenomenon between the interacting electrodes has been the major effort. The application of the existing, more rigorous theories, which go beyond the transfer Hamiltonian approach (THA) to treat the interacting electrodes, are hindered by the boundary conditions.²⁸ Here we develop a formalism for the tunneling between the interacting tip and sample by extending the THA to incorporate the effects of TILS. We start with the THA, because we note that in a real STM the tip creates a local perturbation on a large sample surface. Accordingly, TILS, which appear as localized in the vicinity of the tip, contribute to the tunneling current by modifying the local density of states at the Fermi level.

In view of the above arguments, the time-dependent wave function describing the tunneling event can be writ-

ten as

$$\Psi(\mathbf{r}, t) = a_{\text{tip}}(t) e^{-iE_{\text{tip}}t/\hbar} \psi_{\text{tip}} + \sum_s b_s(t) e^{-iE_s t/\hbar} \psi_s + \sum_{\text{TILS}} b_{\text{TILS}}(t) e^{-iE_{\text{TILS}}t/\hbar} \psi_{\text{TILS}}. \quad (4.1)$$

Here, the wave functions ψ_{tip} and ψ_s are the eigenstates of the bare tip and bare sample Hamiltonians, respectively. The states strongly localized around the tip are treated as TILS. However, the dividing line is not sharp for weakly perturbed states (e.g., bands originated from the tip $2p_x$ and $2p_y$ orbitals in Sec. II) and these are grouped under $\{\psi_{\text{tip}}\}$ and $\{\psi_s\}$. Nevertheless, including these into $\{\psi_{\text{TILS}}\}$ will not effect the essential results. The extension to the Bardeen formalism is the inclusion of TILS, which satisfies

$$\underline{H} \psi_{\text{TILS}}(\mathbf{r}) = E_{\text{TILS}} \psi_{\text{TILS}}(\mathbf{r}), \quad \mathbf{r} \in R_{\text{TILS}} \quad (4.2)$$

R_{TILS} denoting the local environment of the tip atom, used in Sec. II. It is clear that ψ_{TILS} can be written as a linear combination of $\{\psi_{\text{tip}}\}$ and $\{\psi_s\}$, which already form an overcomplete basis set. However, to investigate the effects of the atomic scale interactions on tunneling, we write them explicitly as the third term in Eq. (4.1). This form is also consistent with the original THA, where different tip or sample states, being orthogonal, are treated independently in the current calculation, but overlapping tip and sample states are coupled to each other by a transfer Hamiltonian. In our study the states denoted as TILS are *good solutions*³ of the Schrödinger equation [Eq. (4.2)] in a region different from the tip and sample regions (R_{tip} and R_s , respectively) so they are not orthogonal to the existent basis $\{\psi_{\text{tip}}\} \cup \{\psi_s\}$.

In Eq. (4.1), the initial conditions such as $a_{\text{tip}}(-\infty) = 1$, $b_s(-\infty) = 0$ are imposed in order that the wave function $\Psi(\mathbf{r}, t)$ to represent a transition from the state ψ_{tip} to the sample state ψ_s . Hence, the tunneling probability into the state ψ_s is $p_{\text{tip} \rightarrow s} = |b_s(\infty)|^2$. To understand the roles of strongly perturbed states both occupied and unoccupied TILS (e.g., TS_1 and TS_2 discussed earlier, respectively) are considered. An important point to emphasize at this stage is that the periodically repeating tip-sample system, and the tip represented by a single atom are the only approximations involved in determining TILS. For a very large supercell with a single tip atom, TILS appear as resonances around $E_{TS_1}^\Gamma$ and $E_{TS_2}^\Gamma$ in the density of states. It should be noted that such bands may occur not only along ΓK , but also in the other regions of the BZ. In fact, self-consistent field pseudopotential calculations on the Al-tip and graphite sample have revealed that at certain tip-sample distance TILS near the M point of the graphite BZ can have even more dramatic effects.²⁷ We next include the effects of the transfer Hamiltonian by using the time-dependent perturbation theory. The Schrödinger equation describing this tunneling event is

$$\underline{H} \Psi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t). \quad (4.3)$$

In this equation and in Eq. (4.2), \underline{H} is the total Hamiltonian consisting of interacting electrode contributions and a transfer Hamiltonian responsible for tunneling. Thus, it differs from the local interaction Hamiltonian \underline{H}_{T+S} used in Sec. II. Inserting the wave function given by Eq. (4.1) in Eq. (4.3) one gets

$$a_{\text{tip}}(t) e^{-iE_{\text{tip}}t/\hbar} (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}} + \sum_s b_s(t) e^{-iE_s t/\hbar} (\underline{H} - E_s) \psi_s + \sum_{\text{TILS}} b_{\text{TILS}}(t) e^{-iE_{\text{TILS}}t/\hbar} (\underline{H} - E_{\text{TILS}}) \psi_{\text{TILS}} = i\hbar \dot{a}_{\text{tip}} e^{-iE_{\text{tip}}t/\hbar} \psi_{\text{tip}} + \sum_s i\hbar \dot{b}_s e^{-iE_s t/\hbar} \psi_s + \sum_{\text{TILS}} i\hbar \dot{b}_{\text{TILS}} e^{-iE_{\text{TILS}}t/\hbar} \psi_{\text{TILS}}. \quad (4.4)$$

Using a first-order time-dependent perturbation approach, one sets $a_{\text{tip}}(t) = 1$, $\dot{a}_{\text{tip}}(t) = 0$, and $b_s(t) = 0$. Also for unoccupied (occupied) TILS $b_{\text{TILS}}(t) = 0$ [$b_{\text{TILS}}(t) = 1$ and $\dot{b}_{\text{TILS}}(t) = 0$]. Therefrom one obtains

$$e^{-iE_{\text{tip}}t/\hbar} (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}} + \sum_{\text{TILS}'} e^{-iE_{\text{TILS}'}t/\hbar} (\underline{H} - E_{\text{TILS}'}) \psi_{\text{TILS}'} = \sum_s i\hbar \dot{b}_s e^{-iE_s t/\hbar} \psi_s + \sum_{\text{TILS}} i\hbar \dot{b}_{\text{TILS}} e^{-iE_{\text{TILS}}t/\hbar} \psi_{\text{TILS}}. \quad (4.5)$$

It is apparent that the occupied TILS are behaving like tip states (from which tunneling occurs), and the unoccupied ones are coupled to the sample states (into which tunneling occurs). We will find out below that this resemblance is manifested in the current expression as well. Multiplying Eq. (4.5) from left by $e^{iE_s t/\hbar} \psi_s$ (or by $e^{iE_{\text{TILS}'}t/\hbar} \psi_{\text{TILS}'}$ in the case of unoccupied TILS) and integrating over the whole space, two sets of equations are found:

$$e^{-i(E_{\text{tip}} - E_s)t/\hbar} \int d^3r \psi_s^* (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}} + \sum_{\text{TILS}'} e^{-i(E_{\text{TILS}'} - E_s)t/\hbar} \int d^3r \psi_s^* (\underline{H} - E_{\text{TILS}'}) \psi_{\text{TILS}'} = i\hbar \dot{b}_s + i\hbar \sum_{\text{TILS}}^{\text{unocc}} \dot{b}_{\text{TILS}} e^{-i(E_{\text{TILS}} - E_s)t/\hbar} \int d^3r \psi_s^* \psi_{\text{TILS}}, \quad (4.6a)$$

$$e^{-i(E_{\text{tip}} - E_{\text{TILS}'})t/\hbar} \int d^3r \psi_{\text{TILS}'}^* (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}} + \sum_{\text{TILS}'} e^{-i(E_{\text{TILS}'} - E_{\text{TILS}'})t/\hbar} \int d^3r \psi_{\text{TILS}'}^* (\underline{H} - E_{\text{TILS}'}) \psi_{\text{TILS}'} = i\hbar \dot{b}_{\text{TILS}'} + i\hbar \sum_s \dot{b}_s e^{-i(E_s - E_{\text{TILS}'})t/\hbar} \int d^3r \psi_{\text{TILS}'}^* \psi_s. \quad (4.6b)$$

Finally Eq. (4.6) can be cast into a matrix form:

$$\begin{pmatrix} \underline{I} & \underline{S}^\dagger \\ \underline{S} & \underline{I} \end{pmatrix} \begin{pmatrix} \underline{\dot{b}}_s \\ \underline{\dot{b}}_{\text{TILS}} \end{pmatrix} = \frac{1}{i\hbar} \left[\begin{pmatrix} \underline{M}_{s\text{-tip}} \\ \underline{M}_{\text{TILS-tip}} \end{pmatrix} + \sum_{\text{TILS}'}^{\text{occ}} \begin{pmatrix} \underline{M}_{s\text{-TILS}'} \\ \underline{M}_{\text{TILS-TILS}'} \end{pmatrix} \right], \quad (4.7)$$

where \underline{S} is $(N_{\text{TILS}} \times N_s)$; $\underline{\dot{b}}_s$, $\underline{M}_{s\text{-tip}}$ and $\underline{M}_{s\text{-TILS}'}$ are $(1 \times N_s)$, and $\underline{\dot{b}}_{\text{TILS}}$, $\underline{M}_{\text{TILS-tip}}$, and $\underline{M}_{\text{TILS-TILS}'}$ are $(1 \times N_{\text{TILS}})$ matrices. Their elements are defined as

$$S = e^{-i(E_s - E_{\text{TILS}})t/\hbar} \int d^3r \psi_{\text{TILS}}^* \psi_s, \quad (4.8a)$$

$$M_{s\text{-tip}} = e^{-i(E_{\text{tip}} - E_s)t/\hbar} \int d^3r \psi_s^* (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}}, \quad (4.8b)$$

$$M_{\text{TILS-tip}} = e^{-i(E_{\text{tip}} - E_{\text{TILS}})t/\hbar} \times \int d^3r \psi_{\text{TILS}}^* (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}}, \quad (4.8c)$$

$$M_{s\text{-TILS}'} = e^{-i(E_{\text{TILS}'} - E_s)t/\hbar} \times \int d^3r \psi_s^* (\underline{H} - E_{\text{TILS}'}) \psi_{\text{TILS}'}, \quad (4.8d)$$

$$M_{\text{TILS-TILS}'} = e^{-i(E_{\text{TILS}'} - E_{\text{TILS}})t/\hbar} \times \int d^3r \psi_{\text{TILS}'}^* (\underline{H} - E_{\text{TILS}'}) \psi_{\text{TILS}'}. \quad (4.8e)$$

In Eq. (4.8) the subscripts corresponding to the specific elements of the matrices in Eq. (4.7) are suppressed.

The square matrix on the left-hand side of Eq. (4.7) can be inverted to find $\underline{\dot{b}}_s$ and $\underline{\dot{b}}_{\text{TILS}}$. Then the tunneling current can be found by adding all the average transition rates for each tip state, and integrating this quantity over the Fermi surface of the tip. The resulting expression is complicated and requires a detailed account of the electronic structure. Further in this section we will present a simplified form of it to show the qualitative trends. Before doing this we would like to comment on some important points.

(i) For a finite bias, it is not possible to have TILS between the Fermi energies of the two electrodes since otherwise there would be a continuous flow of electrons between the electrodes. However, such a situation may arise in the point-contact regime, in which the current between the two electrodes has a nontunneling character, and requires a treatment beyond THA. In the rest of this section we disregard this situation by neglecting the second term on the right-hand side of Eq. (4.7).

(ii) Equation (4.7) manifests two different effects of empty TILS on tunneling. First, the nonorthogonality of sample and TILS states changes the current via the inversions of the left-hand side coefficient matrix in Eq. (4.7). In addition to this, the interaction of TILS and tip states is also present in the final expression, denoted by $\underline{M}_{\text{TILS-tip}}$. These contributions disappear when $\underline{S} \rightarrow 0$, which indicates the independent electrode regime. For this case the unoccupied TILS are not localized states in the vicinity of the tip, but are unperturbed tip states. Another possibility is to have some elements of \underline{S} equal to

unity, which leads to a zero determinant for the coefficient matrix. For this case TILS are unperturbed sample states. Therefore, the caution has to be exercised in counting the sample states. Upon inversion of the left-hand side matrix and integration over time, the time-dependent exponents in Eq. (4.8) will combine to give Dirac δ functions of energy differences.

(iii) Inverting Eq. (4.7) one gets $\underline{\dot{b}}_{\text{TILS}} \propto e^{-i(E_{\text{tip}} - E_{\text{TILS}})t}$ and the time integration leads to $P_{\text{tip} \rightarrow \text{TILS}} = |b_{\text{TILS}}(\infty)|^2 \propto \delta(E_{\text{tip}} - E_{\text{TILS}}) = 0$ since $E_{\text{tip}} = E_F \neq E_{\text{TILS}}$. Thus the current sinks into the sample states, but not the TILS. This shows that the energy is conserved. Nevertheless, being nonorthogonal TILS and the tip (sample) states interact. The energy conservation is relaxed for time intervals smaller than $(|E_{\text{TILS}} - E_F|/\hbar)^{-1}$ and unoccupied TILS act as virtual states through which electrons can transfer from the tip into the sample.

To demonstrate the effect of TILS on the observed corrugation we will use the results of Sec. II in a simplified form to calculate the tunneling current for a graphite sample. It is clear from Fig. 1 that the states related to the tunneling event are the sample states at the Fermi level and unoccupied TILS around Γ denoted as TS_2 . To obtain the exact result for the tunneling current, one has to include all these localized states in the matrices in Eq. (4.7). Here we will rather use an approximate method to simplify calculations, and thus to reveal the qualitative trends. Owing to Bloch normalization of the sample states [Eq. (2.1)] and the localized nature of TILS and the tip states, $|S|$ and $|M_{s\text{-tip}}|$ are on the order of $1/\sqrt{N}$, that is the effect of the individual TILS is negligibly small in Eq. (4.7). Nevertheless, the tunneling current is enhanced significantly when all TILS in the whole BZ are concerned. To this end, we will approximate the band labeled TS_2 by a flat band at energy $E_{TS_2}^\Gamma$ with the corresponding TILS wave function $\psi_{\text{TILS}}(\mathbf{r})$. This approximation will lead to an overestimation of the effects of TILS, since as seen in Fig. 1, the TS_2 band approaches to the unperturbed sample π^* band away from Γ point of the BZ.

Using the flat-band approximation and a single sample state, the square matrix on the left-hand side of Eq. (4.7) will have a single sample entry and N TILS entries. Since $|S|$ is the same for all points in the BZ, the determinant of this $[(N+1) \times (N+1)]$ matrix is $(1 - N|S|^2)$. Apparently $N|S|^2$ is on the order of 1 and thus, the TILS enhancement is considerable. Carrying out the inversion and integration over the BZ for the TILS band, one finds $\underline{\dot{b}}_s$ as

$$\underline{\dot{b}}_s = \frac{1}{i\hbar} \frac{1}{(1 - N|S|^2)^2} (M_{s\text{-tip}} - S^* M_{\text{TILS-tip}}), \quad (4.9)$$

where S is the overlap matrix element for the TILS at Γ point, and clearly $M_{\text{TILS-tip}}$ is the same for all points in the BZ. Defining s , $m_{s\text{-tip}}$, and $m_{\text{TILS-tip}}$ as the time-independent factors of the quantities defined in Eq. (4.8), the tunneling probability from the tip state ψ_{tip} into the sample state ψ_s is found by the integration with respect to time

$$P_{\text{tip} \rightarrow s} = \frac{2\pi}{\hbar(1-N|s|^2)^2} |m_{s\text{-tip}} - s^* m_{\text{TILS-tip}}|^2 \delta(E_{\text{tip}} - E_s). \quad (4.10)$$

The prefactor $(1-N|s|^2)^2$ is the nonorthogonality term described above. In the $s \rightarrow 0$ limit, the expression reduces to that of Tersoff-Hamann,² which is the independent electrode approximation. In this form it is possible to make further simplifications. Following Bardeen's derivation³ the first matrix element can be found as

$$m_{s\text{-tip}} = \frac{-\hbar^2}{2m} \int_{S_b} d\mathbf{S} \cdot (\psi_s^* \nabla \psi_{\text{tip}} - \psi_{\text{tip}} \nabla \psi_s^*), \quad (4.11)$$

where the integration surface S_b is lying entirely in the barrier region. For the second one, one has to assume that the TILS is localized in R_{TILS} such that $\psi_{\text{TILS}} \rightarrow 0$ outside R_{TILS} very quickly. Then one has

$$m_{\text{TILS-tip}} = \int d^3r \psi_{\text{TILS}}^* (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}} \quad (4.12a)$$

$$\simeq \int_{R_{\text{TILS}}} d^3r \psi_{\text{TILS}}^* (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}} \quad (4.12b)$$

$$\simeq \int_{R_{\text{TILS}}} d^3r \psi_{\text{TILS}}^* (\underline{H} - E_{\text{tip}}) \psi_{\text{tip}} - \int_{R_{\text{TILS}}} d^3r \psi_{\text{tip}} (\underline{H} - E_{\text{TILS}}) \psi_{\text{TILS}}^* \quad (4.12c)$$

$$\simeq (E_{\text{TILS}} - E_{\text{tip}}) \int_{R_{\text{TILS}}} d^3r \psi_{\text{TILS}}^* \psi_{\text{tip}} - \hbar^2/2m \int_{S_t} d\mathbf{S} \cdot (\psi_{\text{TILS}}^* \nabla \psi_{\text{tip}} - \psi_{\text{tip}} \nabla \psi_{\text{TILS}}^*), \quad (4.12d)$$

where the integration surface S_t for the second integral lies in the intersection of R_{tip} and R_{TILS} . Note that Eq. (4.12c) is obtained by adding to Eq. (4.12b) a term equal to zero in order to symmetrize the integrand.³

If one adds all the possible tunneling events by using the appropriate statistical distribution function for infinitesimal bias and takes the zero temperature limit $T \rightarrow 0$, the final expression for current becomes

$$I = \frac{2\pi e^2 V}{\hbar(1-N|s|^2)^2} \sum_{s,\text{tip}}^{E_{\text{tip}}=E_F} |m_{s\text{-tip}} - s^* m_{\text{TILS-tip}}|^2 \times \delta(E_{\text{tip}} - E_s). \quad (4.13)$$

Introducing some further modifications we can put this equation in a quantitatively tractable form. According to Eq. (4.11), $m_{s\text{-tip}}$ is just the matrix element used in Tersoff-Hamann theory.² Therefore, it is related to the current due to the local density of states, $\rho(\mathbf{r}_0, E_F)$. In addition to this, s can be expressed in terms of the orbital contribution coefficients calculated by using ETB method in Sec. II. An appropriate expression for $m_{\text{TILS-tip}}$ can be found by analyzing Eq. (4.12d). Since ψ_{TILS} can be written as a linear combination of the tip and sample states, the defining integral has tip-tip and tip-sample contributions. The tip-sample part is exponentially small, whereas the tip-tip contribution is of major interest. Thus, $m_{\text{TILS-tip}}$ is approximately proportional to c_t , the tip orbital contribution in the TILS. Finally the complex

phases of the $m_{s\text{-tip}}$ and $s^* m_{\text{TILS-tip}}$ terms can be determined by using the relation between the transition-matrix elements and quantum mechanical current operator.³ Since the sample-tip and TILS-tip interactions are of tunneling and nontunneling character, respectively, the transition-matrix elements will be real and pure imaginary, respectively. Thus, the relative phase between these terms is $e^{i\pi/2}$. Then the resulting expression for total current is

$$I \simeq (1-N|s|^2)^{-2} (I_{\text{LDOS}} + N|s|^2 g |c_t|^2), \quad (4.14)$$

where $G = Ng$ is a scaling function for the TILS-tip interaction. Since we have used several approximations in the derivation of Eq. (4.14), it is not worthwhile to find an exact expression for G . Instead of that, we will use Eq. (4.12d) to find a scaling argument. It is clear that the first term contributing to G is proportional to $(E_{\text{TILS}} - E_F)^2$. Upon summation over the Fermi surface of the sample and tip one gets a factor $D_t(E_F)D_s(E_F)$, $D(E_F)$ being the density of states at the Fermi level for the corresponding electrode. Using the dependence of $(E_{\text{TILS}} - E_F)$ on d (from Fig. 2) one finds $G \propto D_t(E_F)\rho(\mathbf{r}_0, E_F)$ and $g \propto I_{\text{LDOS}}$ for the on-top site position for which the proportionality constant can be calculated. In turn, this constant is used for all lateral positions of the tip. This completes the calculation of the tunneling current using the results of ETB method.

Following the above calculations, the enhancement effect of the TILS on the tunneling current is shown in Fig. 3, by varying the lateral position of the tip at $d=2$ Å. The peak near the on-top position and the shoulder near the hollow site position are due to the approximations involved in the calculation of G , and have no physical significance. It is clear that near the on-top site the enhancement is as large as a factor of 5 and there is no enhancement around the hollow site. The enhancement of the tunneling current is certainly overestimated owing to the nature of the simplifications in Eq. (4.14). Moreover, since the bare sample states forming TILS are counted as ψ_{TILS} , the value of I_{LDOS} becomes smaller

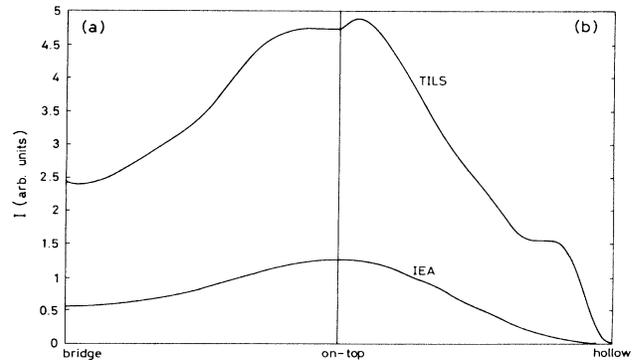


FIG. 3. Tunneling current I along (a) the line connecting bridge site to the on-top site and (b) along the line connecting the on-top site to the hollow site. The tip height is taken to be 2.0 Å. Labels: TILS, this work; IEA, independent electrode approximation.

than that calculated in the independent electrode approximation. In spite of all these the present result provides evidence that the tunneling current is enhanced due to the tip-sample interaction. For example, as the tip is scanned from the H site to the T site the tunneling current increases not only due to the increasing $\rho(\mathbf{r}_0, E_F)$ but also due to the presence of TILS. Neglecting the pathological singularities of the graphite monolayer leading to an infinite corrugation,²⁹ this implies an additional corrugation (≈ 3 Å) over that of $\rho(\mathbf{r}_0, E_F)$. Clearly, at small d (small V) the corrugation is amplified by TILS, but it is still smaller than the observed huge corrugation.

V. CONCLUSIONS

In summary, we showed that the tip-sample interaction in the small tip-sample distance is significant. Owing to this interaction the electron states of the bare sample and bare tip may be disturbed strongly leading to the local-

ized states in the vicinity of the tip. The expression of the tunneling current is reformulated to include the tip induced localized states. A qualitative analysis of this expression indicates that the tunneling current is affected in the presence of these tip-induced states. The STM images, which under conventional circumstances are related to the local density of states at the Fermi level of the clean surface are distorted. Therefore, the experimental results of STM cannot directly reproduce neither the topographical nor the electronic structure of the bare surface. The atomic-scale structure of the tip and surface become important in order to make an analysis of the problem.

ACKNOWLEDGMENTS

We wish to acknowledge valuable discussions with Professor A. Baratoff, Professor R. Ellialtıoglu, and Professor C. Yalabik.

-
- ¹G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* **49**, 57 (1982); G. Binnig and H. Rohrer, *Helv. Phys. Acta.* **55**, 726 (1982); *IBM J. Res. Develop.* **30**, 355 (1986).
- ²J. Tersoff and P. R. Hamann, *Phys. Rev. Lett.* **50**, 1998 (1983); *Phys. Rev. B* **31**, 805 (1985).
- ³J. Bardeen, *Phys. Rev. B* **6**, 57 (1961).
- ⁴A. Baratoff, *Physica B+C* **127B**, 143 (1984).
- ⁵A. Selloni, P. Carnevalli, E. Tosatti, and C. D. Chen, *Phys. Rev.* **31**, 2602 (1984).
- ⁶N. Lang, *Phys. Rev. Lett.* **56**, 1164 (1986); *Phys. Rev. B* **34**, 5947 (1986); *Phys. Rev. Lett.* **58**, 45 (1987); *IBM J. Res. Develop.* **30**, 374 (1986).
- ⁷J. Schneir, R. Sonnefeld, P. K. Hansma, and J. Tersoff, *Phys. Rev. B* **34**, 4979 (1986).
- ⁸A. Bryant, D. P. E. Smith, G. Binnig, W. A. Harrison, and C. F. Quate, *Appl. Phys. Lett.* **49**, 936 (1986).
- ⁹R. M. Feenstra and P. Martenson, *Phys. Rev. Lett.* **61**, 447 (1988).
- ¹⁰I. P. Batra, N. Garcıa, H. Rohrer, H. Salemink, E. Stoll, and S. Ciraci, *Surf. Sci.* **181**, 126 (1987); D. Tománek, S. G. Louie, H. J. Mamin, D. W. Abraham, R. E. Thomson, E. Ganz, and J. Clarke, *Phys. Rev. B* **35**, 7790 (1987).
- ¹¹S.-I. Park and C. F. Quate, *Appl. Phys. Lett.* **48**, 112 (1986).
- ¹²J. M. Soler, A. M. Baro, N. Garcıa, and H. Rohrer, *Phys. Rev. Lett.* **57**, 444 (1986).
- ¹³U. Dürig, J. K. Gimzewski and D. Pohl, *Phys. Rev. Lett.* **57**, 2403 (1986).
- ¹⁴J. K. Gimzewski and R. Möller, *Phys. Rev. B* **36**, 1284 (1987); N. D. Lang, *ibid.* **8173** (1987).
- ¹⁵S. Ciraci and I. P. Batra, *Phys. Rev. B* **36**, 6194 (1987); I. P. Batra and S. Ciraci, *J. Vac. Sci. Technol. A* **6**, 313 (1988).
- ¹⁶V. M. Hallmark, S. Chiang, J. F. Rabot, J. D. Swalen, and R. J. Wilson, *Phys. Rev. Lett.* **59**, 2879 (1987).
- ¹⁷J. Winterlin, J. Wiechers, H. Brune, T. Gritsch, H. Höfer, and R. J. Behm, *Phys. Rev. Lett.* **62**, 59 (1989).
- ¹⁸G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).
- ¹⁹E. Tekman and S. Ciraci, *Proceedings of the Seventh General Conference of the Condensed Matter Division of the European Physical Society, Pisa, Italy, 1987*, edited by F. Bassani *et al.* [*Phys. Scr.*, Topical Issue **19A&B** (1987)]; *Phys. Scr.* **38**, 486 (1988); E. Tekman, M. Sc. thesis, Bilkent University (1988).
- ²⁰G. Binnig, H. Fuchs, Ch. Gerber, H. Rohrer, E. Stoll, and E. Tossatti, *Europhys. Lett.* **1**, 31 (1985).
- ²¹G. Binnig (private communication).
- ²²R. C. Tatar and S. Rabii, *Phys. Rev. B* **25**, 4126 (1982).
- ²³J. C. Slater and G. F. Koster, *Phys. Rev.* **94**, 1498 (1954).
- ²⁴E. Clementi and D. L. Raimondi, *J. Chem. Phys.* **48**, 2686 (1963).
- ²⁵E. Tekman and S. Ciraci, *Phys. Rev. B* **39**, 8772 (1989).
- ²⁶H. A. Mizes, S.-I. Park, and W. A. Harrison, *Phys. Rev. B* **36**, 4491 (1987).
- ²⁷S. Ciraci, A. Baratoff, and I. P. Batra (unpublished).
- ²⁸M. S. Chung, T. E. Feuchtwang, and P. H. Cutler, *Surf. Sci.* **187**, 559 (1987).
- ²⁹J. Tersoff, *Phys. Rev. Lett.* **57**, 440 (1986).