Chain model of Si(111)2×1 surface: Optical properties and surface-state excitons

R. Del Sole
Dipartimento di Fisica, II Università di Roma, Rome, Italy

Annabella Selloni
Dipartimento di Fisica, Università La Sapienza, Rome, Italy
(Received 23 September 1983)

The optical properties of the chain model of Si(111)2×1 are computed both within a single-particle picture and with the inclusion of excitonic and local-field effects. The presence of bound exciton states depends on the magnitude of the screening of electrons in surface states. For weak screening the singlet exciton is strongly bound (~0.3 eV) and gives rise to a narrow Lorentzian peak which dominates the absorption spectrum. For strong screening, bound excitons are not present. The resulting absorption line shape is asymmetrical, in qualitative agreement with experiment, but almost twice as broad. Polarization-dependent reflectance measurements are suggested as an important test for the various models.

I. INTRODUCTION

The atomic and electronic structure of Si(111)2×1 is still an open question though it has been extensively studied for many years. The buckling model originally proposed by Hanemann,[1] and for a long time believed to be correct, has been strongly questioned after angle-resolved photoemission measurements[2,3] on single-domain samples showed large discrepancies with respect to the calculated[4,5] surface-state dispersion. On the other hand, the π-bonded chain model proposed by Pandey[6] yields electron states in agreement with photoemission data and is energetically favored,[6,7] but it does not seem to agree with the dynamical low-energy electron diffraction (LEED) measurements.[8] Other models have been proposed, such as the antiferromagnetic[9] and the molecular[10] models, but at present none of them seem capable to account for all the experimental results.

The optical properties of the Si(111)2×1 surface measured with unpolarized light[11,12] show that transition between surface states are strongly peaked at about 0.45 eV, with an asymmetric line shape broader on the high-energy side. The random-phase-approximation— (RPA-) calculated reflectivity for the buckling model[9(b)] is in qualitative agreement with experiment. One of the purposes of this paper is to extend these calculations to the chain model, in order to check its ability to describe the observed optical properties in the frequency range of the forbidden energy gap of bulk Si. We also compare the polarization dependence of the reflectance for the various models—chain, buckling, and molecular—and find clear-cut differences between the first one and the others. We suggest that reflectivity experiments performed with polarized radiation may bring strong evidence in favor of some of these models.

The microscopic theory of optical properties of crystal surfaces is still developing.[13–15] Recently, a method has been proposed to compute the reflectivity of semiconductor and insulator surfaces, including local-field and exci-

tonic effects within the framework of the many-body perturbation technique.[16] We apply it here for the first time to a realistic problem, namely the chain model of Si(111)2×1. The method is formally similar to the one developed by Hanke and Sham for the study of the optical properties of bulk materials.[17] These authors also demonstrated the importance of including local-field and excitonic effects in a consistent manner in order to adequately describe optical spectra. Although we shall attempt a detailed comparison between theory and experiment, we shall be mainly concerned with the study of the influence of these effects on surface optical properties.

An essential ingredient of our calculation is the screening of the electron-hole interaction at the surface, of which, however, very little is known to date. For this reason various screening models are considered, which may or may not yield bound-exciton states. Very different absorption line shapes and strengths are obtained for the two cases, while the polarization dependence is to a large extent independent of many-body effects. If the screening of electrons in surface states is neglected, a strongly bound exciton is found (~0.3 eV). A significant contribution to the binding energy is due in this case to correlation effects, which are not included in the conventional exciton theory.[18] The absorption basically consists of a narrow Lorentzian peak, in disagreement with experiment. If, on the contrary, a strong screening contribution from surface states is assumed, no bound state exists; the calculated absorption line shape is qualitatively similar to the experiment, but almost twice as broad.

This paper is organized as follows. In Sec. II we summarize the theoretical framework underlying the calculation of the surface optical properties. In Sec. III we describe our model band structure and compute the RPA reflectivity. (We shall conventionally refer to the RPA as the approximation in which local-field as well as excitonic effects are neglected.) The calculation of optical properties including local-field and excitonic effects is described in Sec. IV. A simple molecular model is considered in

Sec. V, to study in detail some features of the electron-hole interaction. Results are presented and compared to reflectivity experiments with unpolarized light in Sec. VI. Finally, in Sec. VII the polarization dependence of the reflectivity for various surface reconstruction models is discussed.

II. THEORETICAL FRAMEWORK

We consider light normally incident on a semi-infinite crystal occupying the half-space $z > 0$ and denote the direction of light polarization by α. The surface contribution to the external reflectivity at frequency ω is given by[15]

$$\frac{\Delta R_\alpha(\omega)}{R(\omega)} = \frac{4\omega}{c} \, \mathrm{Im} \, \frac{\Delta\widetilde{\epsilon}_{\alpha\alpha}(\omega)}{\epsilon_b(\omega) - 1} , \tag{1}$$

where $\epsilon_b(\omega)$ is the bulk dielectric function and $\Delta\widetilde{\epsilon}_{\alpha\alpha}$ is defined as

$$\Delta\widetilde{\epsilon}_{\alpha\alpha}(\omega) = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} dz' \left[\Delta\epsilon_{\alpha\alpha}(z,z';\omega) - \int_{-\infty}^{\infty} dz'' \int_{-\infty}^{+\infty} dz''' \Delta\epsilon_{\alpha z}(z,z'';\omega)\epsilon_{zz}^{-1}(z'',z''';\omega)\Delta\epsilon_{z\alpha}(z''',z';\omega) \right] , \tag{2}$$

with

$$\Delta\epsilon_{\alpha\beta}(z,z';\omega) = \epsilon_{\alpha\beta}(z,z';\omega) - \delta_{\alpha\beta}\delta(z-z')\epsilon_0(z;\omega) . \tag{3}$$

Here, $\epsilon_{\alpha\beta}$ is a component of the macroscopic dielectric tensor[16] of the semi-infinite crystal—ϵ_{zz}^{-1} being the inverse of its (zz) component—and $\epsilon_0(z;\omega)$ is the dielectric function of the sharp (ideal) vacuum-crystal interface,

$$\epsilon_0(z;\omega) = \epsilon_b(\omega)\Theta(z) + \Theta(-z) , \tag{4}$$

which gives rise to the classical Fresnel formulas of reflectivity. Therefore, $\Delta\epsilon_{\alpha\beta}(z,z';\omega)$ represents the surface contribution to the macroscopic dielectric tensor.

We obtain $\epsilon_{\alpha\beta}(z,z';\omega)$ through Fourier transformation of $\epsilon_{\alpha\beta}(\vec{q}_{||},k_z,k_z';\omega)$ in the limit $\vec{q}_{||} \rightarrow 0$, $\vec{q}_{||}$ denoting the surface component of the light wave vector. Using a localized-orbital representation for the electronic states,[17] we have[16]

$$\epsilon_{\alpha\beta}(\vec{q}_{||},k_z,k_z';\omega) = \delta_{\alpha\beta}\left[\delta(k_z - k_z') - \frac{2e^2}{m_0\omega^2}\rho_0(k_z - k_z') \right] - \frac{2e^2}{A_0} \frac{1}{\hbar^2 m_0^2\omega^2} \sum_{s_1,s_2} p_{s_1}^\alpha e^{ik_z R_{1z}} S_{s_1 s_2}^a (\vec{q}_{||},\omega) p_{s_2}^{*\beta} e^{-ik_z' R_{2z}} , \tag{5}$$

where

$$\rho_0(k_z - k_z') = A_S^{-1} \int d^3r \, \rho_0(\vec{r})e^{i(k_z - k_z')z} . \tag{6}$$

$\rho_0(\vec{r})$ is the electron density, $A_S = N_S A_0$ the surface area, and A_0 the surface unit-cell area. The label s_1—and similarly s_2—in Eq. (5) denotes an ordered pair of one conduction, a_{v_1}, and one valence, $a_{v_1'}$, Wannier functions, located in the cells $(\vec{R}_{1||},R_{1z})$ and $(\vec{R}'_{1||},R'_{1z})$, respectively. The momentum matrix element between these Wannier functions is denoted by $\vec{p}_{s_1}$. The matrix $S_{s_1 s_2}^a$ is a modified two-particle Green's function which is obtained by solving the Bethe-Salpeter equation[17]

$$S^a = N^0[1 - (V^a - \tfrac{1}{2}V^s)N^0]^{-1} , \tag{7}$$

where all quantities are matrices in the Wannier function representation. The matrix N^0 is the RPA polarizability, V^a is the Coulomb electron-electron interaction (electron-hole exchange) without long-range part, and V^s is the electron-electron screened exchange (electron-hole Coulomb interaction). The matrix S^a describes both single-particle and collective excitations of the semi-infinite crystal; in particular, $\det(S^a)^{-1} = 0$ determines the (singlet) bound-exciton states.[17] Note that triplet exciton energies can be obtained in a similar way, by simply setting $V^a = 0$ in (7).

The macroscopic character of the dielectric tensor (5) is

related to the use of the truncated Coulomb interaction V^a and to the restriction

$$|k_z|, |k_z'| < k_c \tag{8a}$$

for the allowed range of values of k_z and k_z'.[16] Here, k_c is a cutoff wave vector—the same used to define V^a in Sec. IV—satisfying

$$\omega/c \ll k_c \ll d^{-1} , \tag{8b}$$

where d is the interlayer spacing in the z direction.

In the limit $\vec{q}_{||} \rightarrow 0$, the diagonal components $\epsilon_{\alpha\alpha}$ satisfy[16]

$$\lim_{\vec{q}_{||} \rightarrow 0} \lim_{k_z,k_z' \rightarrow 0} \epsilon_{\alpha\alpha}(\vec{q}_{||},k_z,k_z';\omega)$$

$$= \lim_{q_\alpha \rightarrow 0} \lim_{k_z,k_z' \rightarrow 0} \epsilon^{LL}(q_\alpha,k_z,k_z';\omega) , \tag{8c}$$

where ϵ^{LL} denotes the longitudinal component of the macroscopic dielectric tensor. Because of the nonanalyticity of ϵ^{LL}, it is important that the limit $q_\alpha \rightarrow 0$ follows the one for k_z and k_z' in the right-hand side of (8c). When the off-diagonal elements $\Delta\epsilon_{z\alpha}$ appearing in the second term on the right-hand side of (2) vanish, Eq. (8c) enables us to treat optical properties simply in terms of the longitudinal response, as is usual in the case of cubic crystals.[17] In the formalism in Eq. (5), ϵ^{LL} is given by

$$\epsilon^{LL}(\vec{q}_{||},k_z,k_z';\omega) = \delta(k_z - k_z') - \frac{2e^2}{A_0} \sum_{s_1 s_2} \mu_{s_1}^L(\vec{q}_{||} + \vec{k}_z)e^{ik_z R_{1z}} S_{s_1 s_2}^a (\vec{q}_{||},\omega) \mu_{s_2}^L(\vec{q}_{||} + \vec{k}'_z)e^{-ik_z' R_{2z}} , \tag{9}$$

where $\mu^L_{s_1}(\vec{q}_{||}+\vec{k}_z)$ is the component along $\vec{q}_{||}+\vec{k}_z$ of the dipole matrix element $\vec{\mu}_{s_1}$ between the Wannier functions of the pair s_1.

A complication arising in the calculation of Eqs. (1)—(5) is that a large number of electron states is, in general, required to represent the matrix S^a. However, in our case the frequency interval of interest for the evaluation of $\Delta R\,(\omega)/R\,(\omega)$ is small $(0.2 \lesssim \hbar\omega \lesssim 1$ eV), so that a simple approximation—effectively involving a small number of states—is likely to work rather well. We assume that $\epsilon_{\alpha\beta}$ is the sum of the unperturbed bulk dielectric function plus the contribution of the dangling-bond (DB) states. This assumption amounts to treating the back bonds as unperturbed bulk bonds and to neglecting transitions between back-bond and DB states. Although the first approximation is usually not well justified, transitions involving back-bond-like surface states correspond to energies of a few electron volts,[4,6] and therefore their contri-

bution to the imaginary part of $\epsilon_{\alpha\beta}$ is vanishing for ω values in the forbidden energy gap of bulk Si (unless unlikely strong excitonic effects occur). With the above assumption the off-diagonal elements $\epsilon_{\alpha z}$ of the dielectric tensor vanish for the reconstructed surface models considered in this paper (see Sec. III), and $Im\Delta\tilde{\epsilon}_{\alpha\alpha}(\omega)$ in Eq. (1) reduces to

$$Im\Delta\tilde{\epsilon}_{\alpha\alpha}(\omega) = -\frac{4\pi e^2}{A_0}Im\left[\sum_{s_1 s_2}\mu^\alpha_{s_1}S^a_{s_1 s_2}(0;\omega)\mu^\alpha_{s_2}\right], \quad (10)$$

where the sum over s_1 and s_2 is now restricted to filled and empty DB states only.

Calculations can be further simplified when working within RPA (and neglecting local fields). In this case the two-particle propagator matrix $S^a_{s_1 s_2}(\vec{q}_{||},\omega)$ reduces to $N^0_{s_1 s_2}(\vec{q}_{||},\omega)$, as shown by Eq. (7). This is given by[17]

$$N^0_{s_1 s_2}(\vec{q}_{||},\omega) = N^{-1}_S \sum_{nn'}\sum_{\vec{k}_{||}} c^*_{n\nu_1}(\vec{k}_{||})c_{n'\nu'_1}(\vec{k}_{||}+\vec{q}_{||})c^*_{n'\nu'_2}(\vec{k}_{||}+\vec{q}_{||})c_{n\nu_2}(\vec{k}_{||})e^{i(\vec{k}_{||}+\vec{q}_{||})\cdot(\vec{R}'_{1||}+\vec{R}'_{2||})}$$

$$\times \frac{f_{n'}(\vec{k}_{||}+\vec{q}_{||})-f_n(\vec{k}_{||})}{E_{n'}(\vec{k}_{||}+\vec{q}_{||})-E_n(\vec{k}_{||})-\hbar\omega-i\eta}, \quad (11)$$

where n,n' are band indices and the $c_{n\nu}(\vec{k}_{||})$'s are coefficients of the expansion of the Bloch function $\psi_n(\vec{k}_{||},\vec{r})$ into Wannier functions. Since we are restricted to dangling-bond states, there is only one filled (v) and one empty (c) band which enter the sum in (11). From (2) and (5) we obtain

$$Im\Delta\tilde{\epsilon}_{\alpha\alpha}(\omega) = \frac{8\pi^2 e^2}{\hbar^2 m^2_0\omega^2}\int_{SBZ}\frac{d^2k}{(2\pi)^2}|p^\alpha(\vec{k})|^2$$

$$\times\delta(\hbar\omega - E_{cv}(\vec{k})), \quad (12)$$

where the integral is over the surface Brillouin zone (SBZ), $E_{cv}(\vec{k})=E_c(\vec{k})-E_v(\vec{k})$ is the direct gap between DB bands at $\vec{k}$, and

$$p^\alpha(\vec{k}) = \langle\psi_c(\vec{k})|p^\alpha|\psi_v(\vec{k})\rangle$$

is the α component of the momentum matrix elements between the Bloch states $\psi_c(\vec{k})$ and $\psi_v(\vec{k})$. (From now on we shall assume $\vec{k}\equiv\vec{k}_{||}$). Within the RPA, Eq. (12) is simpler since it does not require calculation of the Wannier functions.

III. RPA ABSORPTION SPECTRUM

The geometry of the chain model for the (2×1)-reconstructed (111) surface of silicon is characterized by the surface atoms being close as bulk nearest neighbors $(d\sim2.35$ Å) and forming "chains" along the $[1\bar{1}0]$ direction, while interchain distances are rather large (see Fig. 1). The DB bands resulting from slab calculations[6] are characterized by a large dispersion along the $\bar{\Gamma}\bar{J}$ direction of the SBZ, parallel to the chain direction, whereas the

bands are almost flat along the perpendicular direction, $\bar{\Gamma}\bar{J}'$ and $\bar{J}\bar{K}$. These features are related to strong DB interactions within a chain, interchain interactions being, on the other hand, negligible. A small gap is obtained along
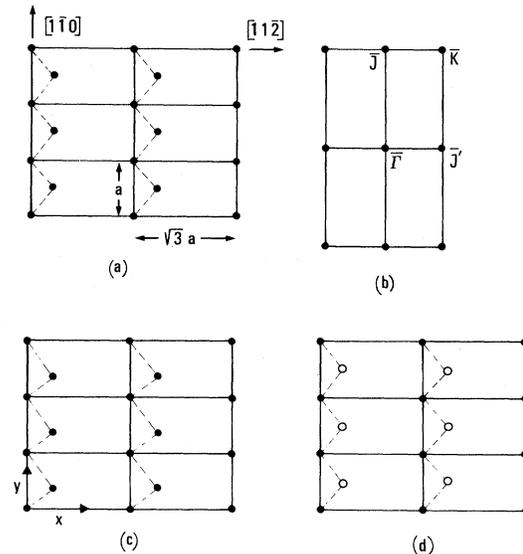

FIG. 1. Atomic arrangement in the surface plane for the π-bonded chain model of Si(111)2×1. Solid circles denote atomic positions, and dashed lines represent bonds in the surface plane. (a) The symmetric chain model. (b) The surface Brillouin Zone. (c) The dimerized-chain model (for better visualization, the dimerization has been strongly magnified). (d) The buckled-chain model, where open circles indicate downward relaxed atoms. The lattice constant along the y (chain) direction is $a = 3.85$ Å. The bond length of the symmetric chain model is taken $d = 2.35$ Å, with bond angles $\theta = 109.5°$.

$\overline{J}\overline{K}$, resulting from the nonequivalence of the two top-layer atoms within a surface unit cell with respect to the fourth-layer atoms.

The relevant features of these surface bands can be reasonably well reproduced by a nearest-neighbor tight-binding model including DB states only. Within this approach, however, degenerate bands along $\overline{J}\overline{K}$ are obtained for the symmetric chain model which was originally proposed for Si(111)2×1,[6] unless the nonequivalence of the two DB states within the surface unit cell is assumed a priori. In our approach such nonequivalence can result, for instance, from some buckling, leading to charge transfers and, hence, to a partially ionic surface. We shall refer to this model as to "the buckled chain." An alternative choice, also leading to a gap along $\overline{J}\overline{K}$, is "the dimerized-chain model," which consists of chains with alternating short (contracted) and long (stretched) bonds.[19] For this second model, which is strictly covalent, the opening of the gap along $\overline{J}\overline{K}$ can be related to the breaking of the reflection symmetry through the $(1\overline{1}0)$ (which we shall also refer to as the xz) plane, as shown in Fig. 1(c). We shall discuss the RPA optical spectrum for both of these models. The tight-binding parameters will be determined by the following requirements: (i) the valence bandwidth is ~ 0.6 eV, as observed in angle-resolved photoemission,[2,3] and (ii) the peak of the absorption spectrum for unpolarized light occurs at 0.45 eV, as observed experimentally.[11,12] This peak is associated with the single-particle gap along $\overline{J}\overline{K}$, which therefore should be taken to be 0.45 eV if excitonic effects are excluded.

A. Buckled chain

We denote the on-site energies of the two nonequivalent DB's within the surface unit cell by ϵ_1 and ϵ_2 and the nearest-neighbor hopping integral by t. The single-particle dispersion relations are

$$E_{c,v}(\vec{k}) = \frac{\epsilon_1 + \epsilon_2}{2} \pm \left[\left[\frac{\epsilon_1 - \epsilon_2}{2} \right]^2 + 4t^2 \cos^2 \frac{k_y a}{2} \right]^{1/2}, \quad (13)$$

where y is the chain direction and $a = 3.85$ Å is the lattice constant along y. Using requirements (i) and (ii) above, we take $E_g = \epsilon_1 - \epsilon_2 = 0.45$ eV and $t = -0.4$ eV, where E_g is the constant gap along $\overline{J}\overline{K}$. The gap along $\overline{\Gamma}\overline{J}'$ is also constant and is given by $E_g' = (E_g^2 + 16t^2)^{1/2} = 1.66$ eV. We calculate the matrix element of the momentum operator between Bloch functions using the commutation property

$$\vec{p} = \frac{im_0}{\hbar}[H, \vec{r}], \quad (14)$$

with H and $\vec{r}$ being the effective DB Hamiltonian and position operator, respectively. We assume that

$$\langle \varphi_i(\vec{r} - \vec{R}_m) | \vec{r} | \varphi_j(\vec{r} - \vec{R}_{m'}) \rangle = \delta_{ij} \delta_{mm'} \vec{R}_{im}, \quad (15)$$

where $\vec{R}_{im}$ denotes the position of the ith atom $(i = 1,2)$ in the mth surface unit cell $(\vec{R}_m)$. This approximation turns out to be correct up to second order in the nearest-neighbor DB overlap, which we estimate to be of the order 0.1.[5] The momentum matrix element is

$$\langle \psi_v(\vec{k},\vec{r}) | \vec{p} | \psi_c(\vec{k},\vec{r}) \rangle = i\frac{m_0}{\hbar} t[c_{1v}^2(k_y) + c_{2v}^2(k_y)]$$

$$\times [(\vec{R}_{20} - \vec{R}_{10}) - (\vec{R}_{11} - \vec{R}_{20})$$

$$\times 2\cos(k_y a)], \quad (16)$$

where $c_{in}(k_y)$ $(n = v,c; i = 1,2)$ are expansion coefficients of the Bloch eigenstates into Bloch sums of DB orbitals. Equation (16) shows that $p^z(k)$ is not vanishing if some buckling is present. The y-z coupling—in the sense of Eq. (2)—is, however, forbidden since p^z and p^y belong to different irreducible representations of the surface-lattice point group. The x-y coupling is, in principle, present; since this is quadratic in the buckling, it is, however, about 2 orders of magnitude smaller than the first term in (2). This justifies the use of Eq. (12) to calculate the absorption spectrum of both x- and y-polarized light. We obtain

$$\mathrm{Im}\Delta\tilde{\epsilon}_{xx}(\omega) = \frac{4\pi e^2}{\sqrt{3}a^2} X_0^2 \frac{1}{\hbar\omega} \left[\frac{(\hbar\omega)^2 - E_g^2}{E_g'^2 - (\hbar\omega)^2} \right]^{1/2}, \quad (17)$$

$$\mathrm{Im}\Delta\tilde{\epsilon}_{yy}(\omega) = \frac{4\pi e^2}{\sqrt{3}a^2} Y_0^2 \frac{E_g^2}{(\hbar\omega)^3} \left[\frac{E_g'^2 - (\hbar\omega)^2}{(\hbar\omega)^2 - E_g^2} \right]^{1/2}, \quad (18)$$

where X_0 and Y_0 denote the x and y components of $\vec{R}_{20} - \vec{R}_{10}$, respectively. From Fig. 1(c), we see that $Y_0 = a/2$ and $X_0 = a/2\sqrt{2}$ (taking $|\vec{R}_{20} - \vec{R}_{10}|$ to be approximately equal to the first-neighbor distance in bulk Si). Equations (17) and (18) are plotted in Fig. 2 together with the analogous results for the dimerized-chain model. For polarized light parallel to the chain direction (y) a sharp peak in the absorption spectrum is found at $\hbar\omega \sim E_g$; the absorption decreases quickly with increasing frequency and is strictly zero at the upper edge $\hbar\omega = E_g'$. Just the opposite trend is predicted for x-polarized light: Here, $\Delta\tilde{\epsilon}_{xx}$ is zero at E_g, increases with increasing x and has a singularity at $\hbar\omega \sim E_g'$. The typical one-dimensional
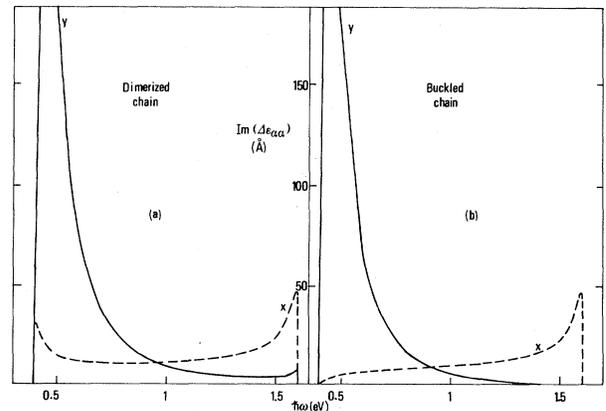

FIG. 2. Imaginary part of the RPA surface dielectric function for light polarized along (y) or perpendicular (x) to the chain. (a) Dimerized-chain model with $t_1 = -0.5$ eV and $t_2 = -0.3$ eV. (b) Buckled-chain model with $\epsilon_1 - \epsilon_2 = 0.4$ eV and $t = -0.4$ eV.

character, $(\hbar\omega - E_g)^{-1/2}$, of the singularities is the result of the bonding topology of the chain model and is not significantly altered by interaction with back-bond states.[20] Since x and y are principal axes for the present model, the absorption spectrum for any polarization direction α in the surface plane is simply given by

$$\text{Im}\Delta\tilde{\epsilon}_{\alpha\alpha}(\omega) = \text{Im}\Delta\tilde{\epsilon}_{xx}(\omega)\cos^2\theta + \text{Im}\Delta\tilde{\epsilon}_{yy}(\omega)\sin^2\theta , \quad (19)$$

where θ is the angle between the α and the x directions. In particular, the spectrum for unpolarized light, which is obtained by angular averaging of (19),

$$\text{Im}\Delta\tilde{\epsilon}(\omega) = \tfrac{1}{2}[\text{Im}\Delta\tilde{\epsilon}_{xx}(\omega) + \text{Im}\Delta\tilde{\epsilon}_{yy}(\omega)] , \quad (20)$$

can be used to evaluate the absolute value of the peak in the surface reflectivity occurring at $\hbar\omega \sim E_g$,

$$(\Delta R/R)_{\text{peak}} \sim \frac{2\pi e^2}{\sqrt{3}(\epsilon_b - 1)c}\left[\frac{E_g'^2 - E_g^2}{2E_g}\right]^{1/2}(\hbar\omega - E_g)^{-1/2} . \quad (21)$$

Assuming a Lorentzian broadening of half-width $\gamma = 20$ meV we find $(\Delta R/R)_{\text{peak}} \sim 2\%$, to be compared with the experimental value $(\Delta R/R)_{\text{peak}} \sim 3\%$ of Ref. 12. A further and more useful comparison with experiment—not dependent of γ—is obtained by evaluating the effective number $N_{\text{eff}}(\omega_M)$ of electrons per atom contributing to optical transitions up to a given energy $\hbar\omega_M$. The experimental $N_{\text{eff}}(\omega_M)$, which is proportional to the integrated reflectivity, is given in Ref. 21, where the definition

$$\int_0^{\omega_M} d\omega\,\omega\,\text{Im}\Delta\tilde{\epsilon}(\omega) = \frac{\pi}{2}\frac{4\pi e^2}{m_0 A_\Omega}N_{\text{eff}}(\omega_M) \quad (22)$$

is used, A_Ω being the area per surface atom. Using (20) and $\omega_M = 1$ eV, we find $N_{\text{eff}}(\omega_M) \sim 0.14$: this is approximately the effective number of electrons per atom participating to the transitions peaked at 0.45 eV, to be compared with the experimental value $N_{\text{eff}} \sim 0.2$.[21]

B. Dimerized chain

We denote ϵ the DB on-site energy, t_1 the hopping integral between DB's along the contracted bond, and t_2 the hopping along the stretched bond. Our model band structure is now

$$E_{cv}(\vec{k}) = \epsilon \pm [t_1^2 + t_2^2 + 2t_1 t_2\cos(k_y a)]^{1/2} , \quad (23)$$

resulting in $E_g' = 2|t_1 + t_2|$ (gap along $\overline{\Gamma J'}$) and $E_g = 2|t_1 - t_2|$. We choose $t_1 = -0.5$ eV and $t_2 = -0.3$ eV, which approximately satisfy both photoemission and optical-absorption data. The momentum matrix element is

$$\langle\psi_v(\vec{k},\vec{r})|\vec{p}|\psi_c(\vec{k},\vec{r})\rangle = i(m_0 t_1/\hbar)(\vec{R}_{10} - \vec{R}_{20})\cos\theta_k$$

$$+ i(m_0 t_2/\hbar)(\vec{R}_{11} - \vec{R}_{20})$$

$$\times\cos(\theta_k + k_y a) , \quad (24)$$

with

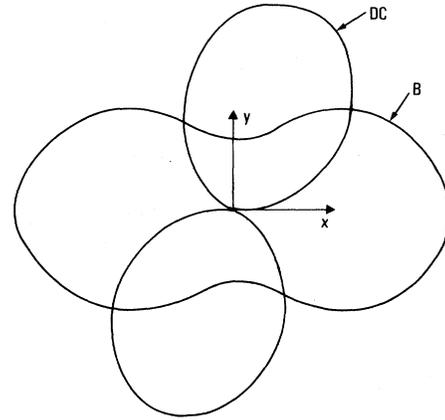$$\theta_k = \tan^{-1}\{t_2\sin(k_y a)/[t_1 + t_2\cos(k_y a)]\} .$$

FIG. 3. Angular plot of the DB optical absorption at $\hbar\omega = 0.45$ eV. The directions x and y are as in Fig. 1(c). The curve labeled DC is for the dimerized-chain model, with the same parameters as in Fig. 2(a). The curve B is for the buckling model, computed according to Ref. 9.

Equation (24) has a zero z component since we are assuming that no buckling is present. Hence, $\Delta\epsilon_{xz}$ and $\Delta\epsilon_{yz}$ vanish and the use of (12) to calculate optical properties is justified. For x- and y-polarized light we find

$$\Delta\tilde{\epsilon}_{xx}(\omega) = \frac{4\pi e^2}{\sqrt{3}a^2}X_0^2\hbar\omega[(E_g'^2 - \hbar^2\omega^2)(\hbar^2\omega^2 - E_g^2)]^{-1/2} , \quad (25)$$

$$\Delta\tilde{\epsilon}_{yy}(\omega) = \frac{4\pi e^2}{\sqrt{3}a^2}Y_0^2\frac{E_g^2 E_g'^2}{(\hbar\omega)^3}[(E_g'^2 - \hbar^2\omega^2)(\hbar^2\omega^2 - E_g^2)]^{-1/2} . \quad (26)$$

The spectra above are very similar to those of the buckled chain, as shown in Fig. 2. The only difference is that now $\Delta\tilde{\epsilon}_{xx}$ and $\Delta\tilde{\epsilon}_{yy}$ have singularities at both E_g and E_g'. The y-x intensity ratio is, however, $2(E_g'/E_g)^2$ at $\omega = E_g$ and $2(E_g/E_g')^2$ at $\omega = E_g'$, with $(E_g'/E_g)^2 \sim 16$.

For the dimerized-chain model, x and y are not principal axes, and thus the simple formula (19) does not hold. However, Eq. (20) for unpolarized light is still valid. In Fig. 3 we show a plot of the intensity of the peak at 0.45 eV as a function of the azimuthal angle of the light-polarized vector (angle formed with x axis). It is interesting to note that the maximum intensity occurs somewhat midway between the y direction and the direction of the contracted bond. In the extreme dimerization limit ($t_2 \to 0$, i.e., $E_g' \to E_g$), the spectra (25) and (26) reduce to a δ function peak at E_g and the maximum intensity occurs along the "molecule" direction.

The absolute peak intensity and the integrated reflectivity have values almost equal to those of the buckled-chain model, as it is evident from the similarity of Figs. 2(a) and 2(b).

IV. EXCITONIC AND LOCAL-FIELD EFFECTS: CALCULATIONS

The possibility of interpreting the observed optical spectrum of Si(111)2×1 as due to a surface-state exciton

was first suggested by Del Sole and Tosatti.[22] Their estimated binding energy—within effective-mass theory—was ~ 0.32 eV. In this and the next two sections we investigate excitonic effects in the absorption spectrum of the dimerized-chain model of Si(111)2×1. From the general discussion of Ref. 6, it seems indeed reasonable to infer that the dimerized model, which is purely covalent, should be energetically more favorable than the buckled-chain model, in which charge transfers between surface atoms occur.

Our study is based on the many-body perturbation technique (MBPT), proposed by Hanke and Sham,[17] leading to Eqs. (5) and (10). By this approach both many-body final-state interactions and local-field effects[23]—intended as those related to charge distribution inhomogeneities—are accounted for. The relevance of these effects in the absorption spectrum of bulk Si was demonstrated in Ref. 17.

In this section we describe the calculations needed for the determination of the optical response function in Eq. (10). In Sec. IV A we calculate the Wannier functions which are needed as the basis set for the construction of the matrices N^0, V^a, and V^s from which S^a is obtained through Eq. (7). The matrix N^0, as given by (11), is also calculated in Sec. IV A, while the Coulomb- and exchange-interaction matrices V^s and V^a are evaluated in Sec. IV B.

A. Wannier functions

For nondegenerate bands we can apply the straightforward definition

$$a_n(\vec{r}-\vec{R}_m)=N_S^{-1/2}\sum_{\vec{k}}e^{-i\vec{k}\cdot\vec{R}_m}\psi_n(\vec{k},\vec{r}) \quad (n=v,c) , \quad (27)$$

where the Bloch functions are

$$\psi_n(\vec{k},\vec{r})=N_S^{-1/2}\sum_m e^{i\vec{k}\cdot\vec{R}_m}\sum_{i=1,2}c_{in}(\vec{k})\varphi_i(\vec{r}-\vec{R}_m) , \quad (28)$$

with coefficients $c_{in}(k)=\pm\exp(\pm i\theta_k/2)$. This choice makes $\psi_n(\vec{k},\vec{r})$ real at the center of the contracted bond of our origin unit cell. The calculation of (27) reduces to evaluating the integrals

$$\mathscr{I}_m=N_S^{-1}\sum_{\vec{k}}e^{\pm i\theta_k/2}e^{-i\vec{k}\cdot\vec{R}_m} , \quad (29)$$

where m is an integer labeling the unit cell along a given chain. We approximate $\exp(i\theta_k/2)$ by its first-order series expansion in the ratio t_2/t_1, leading to

$$a_{v0}\equiv a_v(\vec{r}-\vec{R}_0)$$

$$=\frac{1}{\sqrt{2}}[\alpha(-\varphi_{1\bar{1}}+\varphi_{2\bar{1}})+(\varphi_{10}+\varphi_{20})-\alpha(-\varphi_{11}+\varphi_{21})] , \quad (30)$$

$$a_{c0}\equiv a_c(\vec{r}-\vec{R}_0)$$

$$=\frac{1}{\sqrt{2}}[\alpha(\varphi_{1\bar{1}}+\varphi_{2\bar{1}})+(-\varphi_{10}+\varphi_{20})-\alpha(\varphi_{11}+\varphi_{21})] ,$$

where $\alpha=t_2/4t_1$, and for simplicity we have introduced

the notation $\varphi_{im}\equiv\varphi_i(\vec{r}-\vec{R}_m)$. Equation (30) shows that the lower (upper) Wannier function a_{v0} (a_{c0}) in the reference cell "0" is made up of the bonding (antibonding) combination of DB orbitals in the same cell in addition to an antibonding (bonding) contribution from the adjacent cells along the chain (cells labelled $+1$ and -1). With the parameters of Sec. III B we find $\alpha=0.15$, which justifies our linearization procedure (other choices of the parameters accounting also for the exciton binding can only reduce this value). For example, from successive terms in the expansion, we have found the contribution of second-neighbor cells to have an amplitude of $5t_2^2/32t_1^2$.

The approximate Wannier functions (30) satisfy the condition

$$\langle a_{nm} \mid a_{n'm'}\rangle=\delta_{nn'}\delta_{mm'} \quad (31)$$

to the second order in α. The Wannier state energies are

$$\epsilon_v\equiv\langle a_{vm} \mid H \mid a_{vm}\rangle=\epsilon+t_1 , \quad (32a)$$

$$\epsilon_c\equiv\langle a_{cm} \mid H \mid a_{cm}\rangle=\epsilon-t_1 , \quad (32b)$$

while

$$\langle a_{vm} \mid H \mid a_{cm'}\rangle\equiv 0 , \quad (32c)$$

and

$$\langle a_{vm} \mid H \mid a_{vm\pm 1}\rangle=\tfrac{1}{2}t_2 , \quad (33a)$$

$$\langle a_{cm} \mid H \mid a_{cm\pm 1}\rangle=-\tfrac{1}{2}t_2 . \quad (33b)$$

With the above relations, we find the approximate band structure

$$E_v(k)=\epsilon+t_1+t_2\cos(k_y a) , \quad (34a)$$

$$E_c(k)=\epsilon-t_1-t_2\cos(k_y a) , \quad (34b)$$

which can also be obtained from (23) by a first-order expansion in t_2/t_1. Equations (34) and (23) have the same value along $\bar{J}\bar{K}$ and $\bar{\Gamma}\bar{J}'$.

To calculate $\mathrm{Im}\Delta\bar{\epsilon}_{\alpha\alpha}(\omega)$ [Eq. (10)] we require the dipole matrix elements between lower and upper Wannier functions. To first order in t_2/t_1 the only nonvanishing elements are

$$\langle a_{vm} \mid \vec{r} \mid a_{cm}\rangle=\tfrac{1}{2}(\vec{R}_{2m}-\vec{R}_{1m}) \quad (35a)$$

and

$$\langle a_{vm} \mid \vec{r} \mid a_{cm\pm 1}\rangle=-\alpha a\hat{y} . \quad (35b)$$

As the basis set for the calculation of the matrix N^0, and similarly for V^s, V^a, and S^a, we choose the following (ordered) pairs of Wannier functions:

$$(a_{v0},a_{c-1}), \quad (a_{v0},a_{c0}), \quad (a_{v0},a_{c1}) , \quad (36a)$$

$$(a_{c0},a_{v-1}), \quad (a_{c0},a_{v0}), \quad (a_{c0},a_{v1}) . \quad (36b)$$

In this way we neglect the long-range part of the electron-hole interaction V^s, retaining only central-cell and nearest-neighbor Wannier-function interactions. This approximation works well for strongly bound excitons—as expected at surfaces,[22,24] and for excitonic effects in the continuum spectrum,[17] but it fails to describe large-radius exciton states.

Using expression (11), it appears that the RPA polarizability N^0 has vanishing matrix elements between the two groups of pairs in (36a) and (36b). For the 3×3 matrix corresponding to set (36a) we obtain

$$N^0_1(\omega)=\frac{2\text{sgn}(2\,|\,t_1\,|\,-\omega)}{[(2\,|\,t_1\,|\,-\widetilde{\omega})^2-4t^2_2]^{1/2}}\begin{bmatrix}1 & W & W^2 \\ W & 1 & W \\ W^2 & W & 1\end{bmatrix},\qquad(37)$$

where $\widetilde{\omega}=\omega+i\gamma$, γ being the broadening, and

$$W=(2\,|\,t_2\,|\,)^{-1}$$
$$\times\{(\widetilde{\omega}-2\,|\,t_1\,|\,)+\text{sgn}(2\,|\,t_1\,|\,-\omega)$$
$$\times[(2\,|\,t_1\,|\,-\widetilde{\omega})^2-4t^2_2]^{1/2}\}\;.\qquad(38)$$

The factor 2 in front of (37) is due to spin. Projection

onto set (36b) gives rise to the antiresonant part, N^0_2, of N^0,

$$N^0_2(\omega)=[N^0_1(-\omega)]^*\;.$$

Only minor differences are present between the linearized RPA spectrum (37) and the "exact result" of Sec. III B. Coupling between the two sets (36a) and (36b) is provided by the Coulomb and exchange interactions within the MBPT approach, while it vanishes in the conventional excitonic model.[18]

B. Coulomb- and exchange-interaction matrices

The matrix elements of the Coulomb- and exchange-interaction operators between pairs of Wannier functions are defined by [17]

$$V^a_{s_1s_2}(\vec{q}_{||})=N^{-1}_S\sum_m e^{-i\vec{q}_{||}\cdot\vec{R}_m}\int d^3r\,d^3r'a^*_{v'_1}(\vec{r}-\vec{R}'_1-\vec{R}_m)a_{v_1}(\vec{r}-\vec{R}_m)v^a(\vec{r}-\vec{r}')a^*_{v_2}(\vec{r}')a_{v'_2}(\vec{r}'-\vec{R}'_2)\qquad(39)$$

and

$$V^s_{s_1s_2}(\vec{q}_{||})=N^{-1}_S\sum_m e^{-i\vec{q}_{||}\cdot\vec{R}_m}\int d^3r\,d^3r'a^*_{v'_1}(\vec{r}-\vec{R}_1-\vec{R}_m)a_{v_1}(\vec{r}'-\vec{R}_m)v^s(\vec{r},\vec{r}')a^*_{v_2}(\vec{r}')a_{v'_2}(\vec{r}-\vec{R}'_2)\;,\qquad(40)$$

respectively. Here $\vec{q}_{||}$ is a vector in the SBZ—the limit $\vec{q}_{||}\rightarrow0$ is taken in our calculation—the R's are surface translation vectors, while $v^a(\vec{r}-\vec{r}')$ and $v^s(\vec{r},\vec{r}')$ are Coulomb, without long-range tails, and exchange potentials. We take

$$v^s(\vec{r},\vec{r}')=e^2/\epsilon_s\,|\,\vec{r}-\vec{r}'\,|\;,\qquad(41)$$

where ϵ_s is the surface screening function, including the contribution of DB's.[25] The cutoff Coulomb potential is best defined in terms of its Fourier transform,[16]

$$v^a(\vec{q}_{||},k_z)=\begin{cases}0\text{ if }\vec{q}_{||}\text{ is in the first SBZ and }|\,k_z\,|<k_c, \\ 4\pi e^2/\epsilon_{\text{bg}}(q^2_{||}+k^2_z)/\text{ otherwise,}\end{cases}\qquad(42)$$

where ϵ_{bg} is the background dielectric constant, $\epsilon_{\text{bg}}=(\epsilon_b+1)/2$, describing (bulk) polarization effects not directly included in our treatment.

We now consider the exchange matrix V^s. We keep only terms up to second order in α so that the sum over R_m must be restricted to $m\le2$. By using the expansion (30) of the Wannier functions in terms of DB orbitals, and restricting ourselves to one- and two-center integrals, all matrix elements (40) can be expressed in terms of the following interaction integrals:

$$U\equiv\int d^3r\,d^3r'\,|\,\varphi_{im}(\vec{r})\,|^2\,|\,\varphi_{im}(\vec{r}')\,|\,v^s(\vec{r},\vec{r}')\;,$$

$$V_{ij}(R)\equiv\int d^3r\,d^3r'\,|\,\varphi_i(\vec{r})\,|^2\,|\,\varphi_j(\vec{r}'-\vec{R})\,|\,^2v^s(\vec{r},\vec{r}')\;,$$
$$\qquad(43)$$
$$V_{\text{dip}}\equiv\int d^3r\,d^3r'\varphi_{1m}(\vec{r})\varphi_{2m}(\vec{r})v^s(\vec{r},\vec{r}')\varphi_{1m}(\vec{r}')\varphi_{2m}(\vec{r}')\;.$$

Here U is the intrasite Coulomb repulsion, $V_{ij}(R)$ is the interaction between DB charge distributions at different sites, and V_{dip} is a dipole-type self-interaction between

charge distribution on nearest-neighbor sites (in principle, we should distinguish sites connected by a short or long bond, but here we shall ignore this small difference). In view of the well-known difficulties of a direct calculation of U,[26] we shall treat this term as an adjustable parameter.[27] The Coulomb integrals $V_{ij}(R)$ are calculated using a screened point-charge approximation, which is an exact result for large enough R's. However, this can also be justified for nearest neighbors because of the lateral localization of DB states.[7] Finally, V_{dip} is estimated to be given by the self-interaction of the overlap charge between nearest-neighbor DB's: taking $S\sim0.1$ we find $V_{\text{dip}}\sim0.2$ eV.

We next turn to the matrix V^a. In this case there is no restriction on the values of R_m in the sum (39). For the term $V^a_{(v_0,c_0)(v_0,c_0)}$ we have to evaluate a lattice sum of dipole-dipole—type interactions, similar to the one giving rise to the classical Lorentz field. In three dimensions and with the full Coulomb potential v, instead of v^a, this term is nonanalytic for $\vec{q}\rightarrow0$. This leads to the longitudinal-transverse splitting of singlet excitons.[18] In contrast, when v^a is used, the result is independent of the orientation of $\vec{q}$ and equal to the (full v) transverse sum.[28] In two dimensions, on the other hand, the longitudinal-transverse splitting vanishes as $\vec{q}_{||}\rightarrow0$,[29] so that the distinction between v^a and v is irrelevant in our calculation. With expression (30) for the Wannier functions and the two-center and point-charge approximations for the integrals between DB orbitals, our result is

$$V^a_{(v_0,c_0)(v_0,c_0)}=\tfrac{1}{2}U-0.54(e^2/\epsilon_{\text{bg}}a)\;,\qquad(44)$$

where the value (-0.54) results from the sum of the interchain contribution (-0.99) and the intrachain contribution $(+0.45)$.

V. EXCITON BINDING ENERGY IN THE EXTREME DIMERIZATION LIMIT

In this section we treat the extreme dimerization limit, namely $t_2/t_1 \to 0$. Dangling bonds are coupled in pairs (molecules) and optical excitations are of the Frenkel-exciton type, with only the dipole-dipole interaction surviving between different pairs. Although this model does not reproduce the Si(111)2×1 DB bandwidth, it is sufficiently simple to allow a better understanding of the electron-hole interaction given by the many-body perturbation technique. We can also derive from it the polarization dependence of the optical absorption of other molecular models, such as the one proposed by Chadi.[10]

In the molecular limit, the only nonvanishing dipole moment is between lower and upper Wannier functions in the same cell, as given by (35a), and it is directed along the molecular axis. Therefore, a general feature of molecular DB models is the $\cos^2\theta$ dependence of light absorption, where θ is the angle between the electric field direction and the molecular axis.

The matrix S^a, to be inverted, factorizes into three 2×2 matrices, with only one being related to optically allowed transitions,

$$(S^a)^{-1} = \begin{vmatrix} \omega - 2|t_1| + V_{00} & \widetilde{V}_{00} \\ \widetilde{V}_{00} & -\omega - 2|t_1| + V_{00} \end{vmatrix} . \qquad (45)$$

Here the first row (or column) refers to the pair of Wannier functions (a_{v0}, a_{c0}), while the second refers to (a_{c0}, a_{v0}). A resonant RPA polarizability $(\omega - 2|t_1|)^{-1}$ is associated to the former, and a nonresonant polarizability $(-\omega - 2|t_1|)^{-1}$ to the latter, $2|t_1|$ being the single-particle gap. The coupling between them,

$$\widetilde{V}_{00} = -\tfrac{1}{2}U + 2E_M - \tfrac{1}{2}V_c(\tau) , \qquad (46)$$

is introduced within the framework of the MBPT,[17,30] and is due to correlation, as discussed below. Here E_M is a Madelung-like term accounting for dipole-dipole interactions, and $V_c(\tau) = e^2/\epsilon_s\tau$, τ being the molecular distance.

The energy of the Frenkel exciton is given by the zero of the determinant of $(S^a)^{-1}$,

$$\omega_{\text{ex}}^2 = (2|t_1| - V_{00})^2 - \widetilde{V}_{00}^2 . \qquad (47)$$

If the usual excitonic model is used, $\widetilde{V}_{00} = 0$ and the exciton binding energy is

$$V_{00} = -\tfrac{1}{2}U + 2E_M + \tfrac{1}{2}V_c(\tau) - V_{\text{dip}} . \qquad (48)$$

This is the electron-hole central-cell interaction, that is, the e-e exchange $[\tfrac{1}{2}U + \tfrac{1}{2}V_c(\tau)]$ subtracted, of twice the e-e Coulomb interaction $(\tfrac{1}{2}U + \tfrac{1}{2}V_{\text{dip}} - E_M)$. The inclusion of the coupling $\widetilde{V}_{00}$ between resonant and nonresonant terms decreases the excitation energy, so that the exciton binding energy increases. This effect has been discussed by Anderson[31] for the case of interatomic correlation, assuming *ab initio* correlated atomic levels. In (46), in addition to the term $2E_M$ accounting for intercell dipole-dipole correlation similar to Anderson's, we have also an intramolecular contribution,

$$\widetilde{V}_{\text{intra}} = -[\tfrac{1}{2}U + \tfrac{1}{2}V_c(\tau)] .$$

The origin of this term can be studied by considering, for simplicity, a single molecule with one bonding and one antibonding oribital. Let E_N ($N = 2$) be the energy of the Hartree-Fock ground state, which is decreased by correlation to the value (we use asterisks to indicate quantities including correlation) $E_N^* = E_N - E_{\text{corr}}$.

The correlation energy can be easily computed in this simple model, and is found to be due to the interaction with the double excited state. It turns out that such interaction is just $\widetilde{V}_{\text{intra}}$. The exciton state, of energy E_{ex}, has no correlation energy, since it couples neither to the ground state (Brillouin's theorem[32]) nor to the double excited state. Though this is a characteristic of the two-level model, it is also qualitatively correct, in general, since electron correlation is less effective in an excited state, where the excited and valence electrons are in different orbitals.

The optical excitation energy

$$\omega_{\text{ex}}^* = E_{\text{ex}} - E_N^* = E_{\text{ex}} - E_N + E_{\text{corr}} \qquad (49)$$

is increased by intramolecular correlation, in apparent contrast with our finding based on Eq. (47). However, Eq. (47) is referred to the gap in the single-particle spectrum $E_{\text{gap}}^* = 2|t_1|$, which by definition is already affected by correlation. The single-particle gap is the difference between the conduction-electron level and the hole level,

$$E_{\text{gap}}^* = E_c^* - E_v^* = (E_{N+1}^* - E_N^*) - (E_N^* - E_{N-1}^*) ,$$

where $E_{N\pm1}^*$ denotes the ground-state energy of the system with $N\pm1$ electrons. In our case,

$$E_{\text{gap}}^* = E_{N+1} + E_{N-1} - 2E_N + 2E_{\text{corr}} ,$$

since E_{N+1} and E_{N-1} are not affected by correlation. The result is that the gap E_{gap}^* is increased by twice the ground-state correlation energy, while the excitation energy ω_{ex}^* is increased only by E_{corr}. The increase of exciton-binding energy previously derived from the MBPT is then recovered.

We note that the e-h interaction (V^s and V^a) alters the energy of the transition and thus its oscillator strength. This is not in contrast with the f-sum rule, since we consider only transitions between two levels, which do not exhaust the entire spectrum.

If we consider again the dimerized chain with $t_2 \neq 0$ the $\widetilde{V}_{00}$ term survives as central-cell correlation tending to increase the exciton binding energy. It may play an important role in cases with a small gap, as it is evident from Eq. (45). In our calculations the effect of $\widetilde{V}_{00}$ is found to be quantitatively relevant ($\sim 20\%$) for the weak screening model reported below.

VI. EXCITONIC AND LOCAL-FIELD EFFECTS: RESULTS AND DISCUSSION

In order to calculate the reflectivity of the dimerized-chain model, including excitonic and local-field effects, we have to specify the screening of the e-e exchange interaction [Eq. (41)]. For this we shall use two rather extreme models.

The first model, hereafter referred to as the "weak-screening model," completely neglects the screening of DB electrons and approximates the semi-infinite semiconductor as a classical dielectric, using the image-charge method. The resulting surface-screening constant is $\epsilon_s = \epsilon_{bg} = 6.5$. Strong excitonic effects result from this model. To fit the experimental peak position we take $t_2 = -0.3$ eV and $t_1 = -0.7$ eV, yielding $E_g = 0.8$ eV and $U = 1.5$ eV. The exciton binding energy turns out to be 0.32 eV, in substantial agreement with the effective-mass calculation of Del Sole and Tosatti.[22] The computed absorption is shown in Fig. 4 for polarizations parallel (y) and perpendicular (x) to the chains. The high-energy singularity—at $E'_g = 2 \mid t_1 + t_2 \mid = 2$ eV within the RPA—is shifted to lower energies by about 0.6 eV. The effect of $\tilde{V}_{00}$ is mainly on this peak, which originates from e-h pairs more localized in the central cell. The low-energy peak (the bound exciton) is weakly dependent on the value of U, because the electron and hole are mainly localized on first-neighbor cells. However, the triplet exciton becomes unstable as U increases beyond $U \sim 1.5$ eV. The line shape does not show the experimentally observed asymmetry and the inclusion of higher exciton states should not modify it. Higher exciton states (independently calculated including ~ 200 Wannier states) indeed appear close to the unperturbed band edge, i.e., about 0.3 eV above the main peak, and have a small oscillator strength. The effective number of electrons participating to the transitions up to $\omega_M = 1$ eV [see Eq. (22)] turns out to be $N_{eff} = 0.042$, about a factor of 5 smaller than the experimental value.[21] Exciton-phonon coupling can alter the symmetric line shape and make it more similar to experiment,[33] but it does not change the oscillator strength. The weak-screening model therefore seems inappropriate to describe the optical properties of Si(111)2×1 within the chain model.

In the second model, namely the "strong-screening" model, we account for the screening generated by DB electrons. The electron-hole interaction at large distances is not affected by DB screening, because the dielectric function of a two-dimensional (2D) semiconductor tends to unity at small wave vectors.[34] Similarly, electrons and holes interact via the bare Coulomb potential as their distance tends to zero. However, it is reasonable to assume that at some intermediate distance, of the order of those involved in virtual transitions between DB states, large screening is produced because of the small gap. We try to describe this situation by neglecting the e-h interaction $V_c(R)$ at the distances of first, second, and third neighbors, but retaining the intra-DB U term and the Madelung-type term E_M. The same approximation has been employed in describing excitons within the antiferromagnetic buckling model of Si(111)2×1, resulting in very small excitonic and local-field effects on optical properties.[35] Here the total e-h interaction becomes repulsive when this picture is used. To fit the absorption peak position we use the parameters $t_1 = -0.5$ eV and $t_2 = -0.3$ eV corresponding to $E_g = 0.4$ eV — and $U = 1$ eV. The results are shown in Fig. 5. As in the previous (weak-screening) case the dependence on polarization is not substantially changed with respect to the RPA. The effective number of electrons up to 1 eV is 0.13, quite smaller than the experimental value of $N_{eff} = 0.2$.[21] In the spectrum for x-polarized light, an exciton is found above the upper band edge E'_g, due to the repulsive e-h interaction. This peak is not expected to be significant for the purpose of comparison with experimental data since many other transitions can occur at this energy. However, it might partially account for the high-energy structure observed in external reflectivity.[12] The line shape of the y polarization is now more similar to the experimental one. However, the width is larger (about twice), which could be due to the extreme character of the "strong-screening" model.

In order to check this possibility, we further used the "intermediate-screening" function given by Keldysh[36] for the macroscopic three-layer model. Here the three layers are bulk silicon, the DB layer of thickness d_s, and the
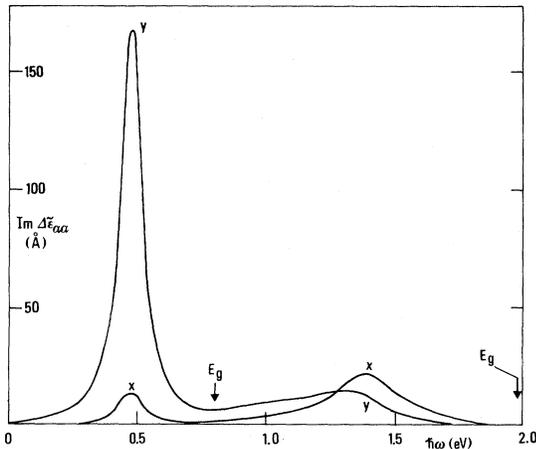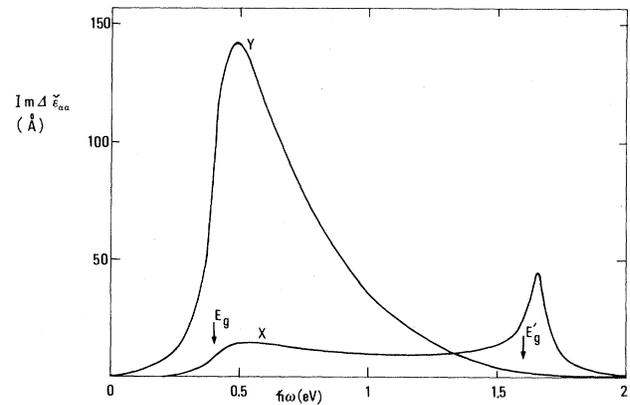

FIG. 4. Imaginary part of the surface dielectric function for the dimerized-chain model including excitonic and local-field effects, computed using, $t_1 = -0.7$ eV, $t_2 = -0.3$ eV, $U = 1.5$ eV, and the "weak-screening" assumption. The broadening is 50 meV.

FIG. 5. Imaginary part of the surface dielectric function for the dimerized-chain model including excitonic and local-field effects, computed using $t_1 = -0.5$ eV, $t_2 = -0.3$ eV, $U = 1$ eV, and the "strong-screening" assumption. The broadening is 50 meV.

vacuum. Large screening occurs for electron-hole distance ρ in the range $d_s \ll \rho \ll \bar{\epsilon} d_s /(\epsilon_b + 1)$, where the interaction depends logarithmically on ρ. $\bar{\epsilon}$ is the static dielectric constant of the surface layer. Taking $d_s = 1.5$ Å, we estimate $\bar{\epsilon} = 94$ from the experimental optical spectrum.[21] This leads to $\epsilon_s \cong 15$, 12, and 9 for first, second, and third neighbor e-e exchange screening, respectively. This screening model yields a weakly-bound-exciton state ($E_{bind} \sim 0.1$ eV), but the absorption line shape and strength (we find $N_{eff} = 0.073$) are similar to those of the strongly bound exciton in Fig. 4, apart from a slightly larger high-energy tail due to band-to-band transitions.[37]

Summarizing, our results show that the line shape and strength of the surface optical absorption are very sensitive to the details of the screening of electrons in surface states. Surface-state screening is still very poorly known,[36,38] and for this reason, rather crude models have been used. Our "best fit" to experimental data,[11,12] is that of Fig. 5, based on the strong-screening approximation. For a more direct comparison with experiment, in Fig. 6 we show the calculated surface reflectivity $\Delta R /R \sim \omega \epsilon(\omega)$ for unpolarized radiation. In order to provide a quantitative estimate of many-body effects in the strong-screening model, the RPA reflectivity is also shown. Incidentally, we note that the latter is in better agreement with experiment than the one including many-body effects. For this the agreement with experiment is only qualitative since, as mentioned earlier, the line shape is 2–3 times too broad and the oscillator strength 40% too small. Also, the peak height is about 3 times too small. We stress, however, that these results should not be taken as a "proof" that the chain model is inconsistent with optical experiments on Si(111)2×1. The above discrepancies—although quan-

titatively large—may probably be overcome by an improved treatment of both band-structure and e-h interaction effects. In addition, electron-phonon coupling also influences surface optical properties. In particular, as shown in Ref. 33, the observed absorption spectrum can be described fairly well assuming a weakly bound exciton coupled to the lattice. A more stringent test for the chain model, as well as for other models of surface reconstruction, is discussed in Sec. VII.

VII. POLARIZATION DEPENDENCE OF REFLECTANCE

An important test for surface-reconstruction models may be provided by polarization-dependent reflectance measurements. Clear-cut differences between the various models, not depending on the details of the exciton states, are in fact present.

In order to establish a unique notation for the various models, we call x the direction where the periodicity is doubled, and y the perpendicular direction, in agreement with the notations used in the preceding sections.

For the chain model the optical absorption at the 0.45-eV peak is substantially due to y-polarized light, the ratio between y and x polarization being 1:0 for the buckled chain and of the order 10:1 for the dimerized chain.

The optical properties of the buckled model of Si(111)2×1 have been computed by Chadi and Del Sole.[9] The calculation was for the antiferromagnetic buckling model, but it holds also for the usual nonmagnetic configuration. Excitonic and local-field effects on optical properties are found to be small in the "strong-screening" approximation.[35] The y:x polarization ratio turns out to be 1:3 (see Fig. 3), reflecting the geometry of this model.

For Chadi's molecular model[10] the absorption is expected to be maximum for light polarized along the molecular axis at a small angle with the x direction. Interactions between different molecules, of the same order of those considered in the buckling model, i.e., 0.1 eV, which are neglected here, should not alter this result considerably.

In conclusion, reflectance measurements using polarized light should provide a large amount of information about the Si(111)2×1 (and also Ge) structure. Previous measurements,[39] performed not at the main peak energy but at a metastable lower-energy peak, seem to favor the buckling model, but are certainly incomplete. Measurements at the peak frequency—or better, the polarization dependence of the whole spectrum—can, on the other hand discriminate well between the chain and the buckling or molecular models. Smaller differences between the two latter models, as well as between the buckled and dimerized chains, may probably be detected as well.

Note added. After submission of this paper, polarization-dependent reflectivity measurements were performed on Si(111)2×1 by P. Chiaradia et al.[40] The observed y:x polarization ratio at the 0.45-eV peak is 1:0 within experimental errors. Similar results have been obtained by Olmstead and Amer[41] by photothermal displacement spectroscopy. Therefore, these results—besides
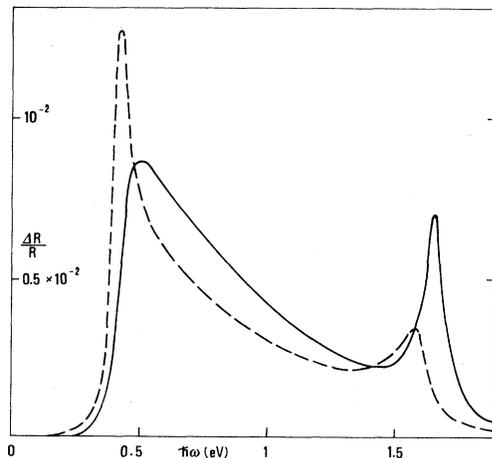

FIG. 6. Surface reflectivity of the dimerized-chain model including excitonic and local-field effects (solid line) and within the RPA (dashed line). The parameters of the calculation are $t_1 = -0.5$ eV, $t_2 = -0.3$ eV, and $U = 1$ eV. The "strong-screening model" is used to describe the screening of electrons in surface states. To improve the fitting to experiment an energy-dependent broadening has been used ($\gamma = 20$ meV below 1 eV and $\gamma = 50$ meV above it).

ruling out the buckling and molecular models—favor the symmetric (slightly buckled) chain model over the dimerized one. Discrimination between the two versions of the chain model is, however, almost at the limit of experimental error. A detailed interpretation of the polarization-dependent reflectivity data will be published elsewhere.

[1]D. Haneman, Phys. Rev. 121, 1093 (1961).

[2]F. J. Himpsel, P. Heimann, and D. E. Eastman, Phys. Rev. B 24, 2003 (1981).

[3]R. I. G. Uhrberg, G. V. Hansson, J. M. Nicholls, and S. A. Flodstrom, Phys. Rev. Lett. 48, 1032 (1982).

[4]K. C. Pandey and J. C. Phillips, Phys. Rev. Lett. 32, 1433 (1974).

[5]F. Casula and A. Selloni, Solid State Commun. 37, 495 (1981).

[6]K. C. Pandey, Phys. Rev. Lett. 47, 1913 (1981); 49, 223 (1982).

[7]J. Northrup and M. L. Cohen, Phys. Rev. Lett. 49, 1349 (1982).

[8]R. Feder, Solid State Commun. 45, 51 (1983).

[9](a) R. Del Sole and D. J. Chadi, Phys. Rev. B 24, 1120 (1981); (b) D. J. Chadi and R. Del Sole, J. Vac. Sci. Technol. 21, 319 (1982).

[10]D. J. Chadi, Phys. Rev. B 26, 4762 (1982).

[11]G. Chiarotti, S. Nannarone, R. Pastore, and P. Chiaradia, Phys. Rev. B 4, 3398 (1971).

[12]P. Chiaradia, G. Chiarotti, S. Nannarone, and P. Sassaroli, Solid State Commun. 26, 813 (1978).

[13]P. J. Feibelman, Prog. Surf. Sci. 12, 287 (1982).

[14]A. Bagchi and A. K. Rajagopal, Solid State Commun. 31, 127 (1979); A. Bagchi, R. C. Barrera, and A. K. Rajagopal, Phys. Rev. B 20, 4824 (1978).

[15]R. Del Sole, Solid State Commun. 37, 537 (1981).

[16]R. Del Sole and E. Fiorino, Phys. Rev. B 29, 4631 (1984).

[17]W. Hanke and L. J. Sham, Phys. Rev. B 21, 4656 (1980).

[18]F. Bassani and G. Pastori-Parravicini, in Electronic States and Optical Transitions in Solids, edited by R. A. Ballinger (Pergamon, New York, 1975).

[19]K. C. Pandey, Phys. Rev. B 25, 4338 (1982).

[20]A. Selloni and C. M. Bertoni, Solid State Commun. 45, 475 (1983).

[21]S. Nannarone and S. Selci, Phys. Rev. B 28, 5930 (1983).

[22]R. Del Sole and E. Tosatti, Solid State Commun. 22, 307 (1977).

[23]S. Adler, Phys. Rev. 126, 413 (1982); N. Wiser, 129, 62 (1983).

[24]M. Altarelli, G. B. Bachelet, V. Bouché, and R. Del Sole, Surf. Sci. 129, 447 (1983).

[25]W. Kohn, in Many Body Physics, edited by C. De Witt and R. Balian (Gordon and Breach, New York, 1968).

[26]G. Allan and M. Lannoo, Surf. Sci. 63, 11 (1977).

[27]Note, however, that the triplet-exciton binding energy depends critically on the value of U. The condition of stability of the ground state against triplet-exciton formation actually determines the value of U quite unambiguously.

[28]P. V. Giaquinta, M. Parrinello, E. Tosatti, and M. P. Tosi, J. Phys. C 9, 2031 (1976).

[29]M. H. Lee and A. Bagchi, Phys. Rev. B 22, 2645 (1980).

[30]E. Fiorino and R. Del Sole, Phys. Status Solidi B 119, 315 (1983).

[31]P. W. Anderson, Concepts in Solids (Benjamin, New York, 1963), p. 132.

[32]J. C. Slater, Quantum Theory of Matter (McGraw-Hill, New York, 1968).

[33]C. D. Chen, A. Selloni, and E. Tosatti (unpublished).

[34]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. 54, 437 (1982).

[35]R. Del Sole (unpublished).

[36]L. V. Keldysh, Pis'ma Zh. Eksp. Fiz. 29, 716 [JETP Lett. 29, 653 (1979)].

[37]For this case our results are only indicative, because our basis set is too small to adequately represent shallow excitons. However, calculations with a large basis set (200 Wannier states) show that the exciton levels are similar to those of the weak-screening model, with the lowest exciton bound by 0.2 eV, while excited levels occur close to the gap.

[38]C. H. Wu and W. Hanke, Solid State Commun. 23, 829 (1977).

[39]J. Assmann and W. Monch, Surf. Sci. 99, 34 (1980).

[40]P. Chiaradia, A. Cricenti, S. Selci, and G. Chiarotti, Phys. Rev. Lett. 52, 1145 (1984).

[41]M. A. Olmstead and N. M. Amer, Phys. Rev. Lett. 52, 1148 (1984).