Layer-by-layer analysis of the linear optical response of clean and hydrogenated Si(100) surfaces

Bernardo S. Mendoza,[1] F. Nastos,[2] N. Arzate,[1] and J. E. Sipe[2]
[1]*Department of Photonics, Centro de Investigaciones en Optica, León, Guanajuato, Mexico*
[2]*Department of Physics and Institute for Optical Sciences, University of Toronto, 60 Saint George Street, Toronto, Ontario M5S 1A7, Canada*

(Received 11 May 2006; revised manuscript received 5 July 2006; published 16 August 2006)

We calculate the reflectance anisotropy and the reflectance-difference spectra for a clean Si(100) surface and two hydrogen- (H-) covered Si(100) surfaces. The clean surface is a 2×1 surface reconstruction, characterized by a tilted dimer formed between the two topmost Si atoms. One of the H-covered surfaces is a monohydride surface in which the two dangling bonds of the dimer are H saturated to give a flat dimer, and the other surface is a dihydride surface in which the H saturates each of the two dangling bonds leading to a bulk ideally terminated surface. The optical response is calculated with a pseudopotential framework using the local-density-plus-scissors approximation. A "layer-by-layer" analysis of the response is implemented with the pseudopotential calculation.

DOI: 10.1103/PhysRevB.74.075318 PACS number(s): 73.20.At, 71.20.Mq, 78.68.+m, 71.15.Dx

I. INTRODUCTION

Reflectance anisotropy spectroscopy (RAS) is a versatile surface-optical probe, providing a measure of the surface anisotropy of cubic crystals. The reflectance anisotropy (RA) spectrum is the difference between the linear reflectance spectra of two orthogonal components of nearly normally incident light. Due to the isotropy of the bulk contribution to the reflection, the RA spectrum originates from the lower symmetry surface. Related to RAS is reflectance difference spectroscopy (RDS), where one studies the difference between the near-normal incidence reflectance spectra of two surfaces with identical bulk components. One of the surfaces is used as a reference, and the other is usually the same reference surface but with some modification induced by adsorbed atoms or molecules; thus the reflectance difference (RD) signal captures the difference between these two surfaces.[1]

Theoretical approaches to calculating surface spectra have evolved over the years, involving increasing levels of sophistication. An early phenomenological model is the three-layer model,[2] where the dielectric tensor of the crystal-vacuum interface is modeled by three isotropic layers: a bulk, a surface, and a vacuum. A more popular model is the discrete-dipole model,[3–8] which easily allows local field effects to be taken into account. It represents the extreme tight-binding limit, and its level of success varies greatly from system to system. Moreover, the model depends on parameters that have to be fitted to experiment or found from a more detailed microscopic calculation. Semiempirical tight-binding models are also used with some success.[9–12] Full-band structure calculations based on density functional theory (DFT) and the scissors approximation have the benefit of being nearly *ab initio*, usually requiring the experimental band gap to be known.[13–16] State-of-the-art calculations, based on the *GW* method, include local field and excitonic effects, but are difficult and rare.[17–19]

Recently, Castillo *et al.*[20] introduced a method to decompose the total RA signal into contributions coming from the different microscopic layers of the surface region, and demonstrated the utility of this for the InP(100) surface. The calculation was done within the semiempirical tight-binding approach, and exploited the localization of the wave functions in its numerical implementation. Hogan *et al.*[14] introduced a similar method using a plane-wave basis within a pseudopotential calculation, taking a Si(100) surface as an example to demonstrate the method's usefulness. More recently, in a study of oxygen-covered silicon surfaces, Inczé *et al.*[16] employed this so-called "layer-by-layer" analysis with success. These calculations confirm that most of the RA signal comes from the topmost layers, since they are anisotropic; as one moves into the bulk the contribution to the signal vanishes, since the symmetric environment of the cubic bulk is being sampled.

In this work we study the reflectance spectra of a clean Si(100) surface and two hydrogenated Si(100) surfaces. The clean surface we investigate is the 2×1 reconstruction, characterized by a tilted dimer formed between the two topmost Si atoms. One of the hydrogenated surfaces is the monohydride 2×1 reconstruction, in which one of the two dangling bonds of the Si dimer is H saturated, rendering a flat dimer; the other is a dihydride surface in which hydrogen saturates both dangling bonds, leading to an ideally terminated surface having a 1×1 surface unit cell. To calculate the total RA and RD spectra we use a plane-wave pseudopotential approach based on DFT, within the local-density- (LD-) plus-scissors approximation. We also investigate the layer-by-layer analysis method proposed by Hogan *et al.*,[14] and we provide a microscopic justification for the procedure which complements their phenomenological arguments.

The manuscript is organized as follows. In Sec. II we review the definition of RA and RD signals and present the theoretical derivation of the layer-by-layer analysis of the surface linear-optical response. In Sec. III the details of the calculation are explained, and in Sec. IV we discuss the results of the surface-optical response for the clean and hydrogenated Si(100) surface. Finally, we present our conclusions in Sec. V.

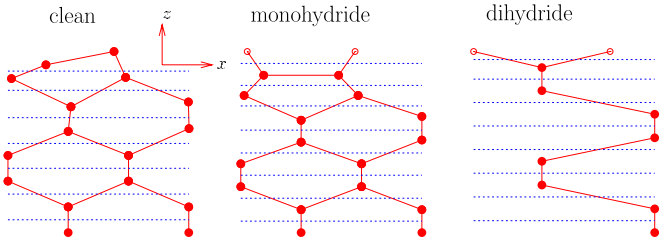

FIG. 1. (Color online) We show the first atomic planes for the unit cells of the slabs used for the three Si(100) surfaces. The coordinates are given in Table I. The solid (empty) circles are the Si (H) atoms, and the solid lines denote the bonds. The dotted lines are the layers, where the upper boundary of the topmost layer (not shown) extends into the vacuum to the midway point between two of the periodically repeated slabs. The dihydride slab has a 1×1 surface reconstruction with only one Si atom in each atomic plane.

II. THEORY

To model the semi-infinite crystal we use a slab consisting of N atomic layers inside a supercell of total height L. This supercell includes the vacuum region required to use a repeated slab scheme. The area of the cell depends on the reconstruction we are considering. The clean and monohydride surfaces are 2×1 reconstructions and the dihydride surface is a 1×1 reconstruction. The surface is parallel to the x-y plane, with the surface normal in the z direction (see Fig. 1 and Table I).

The RA spectrum $\mathrm{RA}(\omega)$ is defined as

$$\mathrm{RA}(\omega) = \mathcal{R}_x - \mathcal{R}_y, \tag{1}$$

and RD spectrum $\mathrm{RD}(\omega)$ is

$$\mathrm{RD}_{x,y} = \mathcal{R}_{x,y}|_{\mathrm{clean}} - \mathcal{R}_{x,y}|_{\mathrm{covered}} \tag{2}$$

for the incoming light polarized along x or y, and

$$\mathrm{RD}_{upol} = \left.\frac{\mathcal{R}_x + \mathcal{R}_y}{2}\right|_{\mathrm{clean}} - \left.\frac{\mathcal{R}_x + \mathcal{R}_y}{2}\right|_{\mathrm{covered}} \tag{3}$$

for unpolarized incoming light. In these equations, $\mathcal{R}_a$, with $a=x$ or y, is the normal incidence reflectance for light linearly polarized in the a direction. For crystals that exhibit isotropic linear response in the bulk (in the case of RAS), or for surface modifications that do not extend into the bulk (in the case of RDS), $\mathcal{R}_a$ can be replaced by the contribution to that reflectance from the surface region, since bulk contributions cancel in the differences Eqs. (1)–(3). That replacement is related to the calculated optical response of the slab system according to

$$\mathcal{R}_a = 4(\omega/c)\mathrm{Im}\left[\frac{D\chi_{\mathrm{sc}}^{aa}(\omega)}{\chi_B(\omega)}\right], \tag{4}$$

where ω is the angular frequency of the incident light, c is the speed of light, D is half the slab thickness for a bulk material with isotropic linear susceptibility $\chi_B(\omega)$, and $\chi_{\mathrm{sc}}^{aa}(\omega)$ is the linear susceptibility for the fictitious supercell system. The so-called slab polarizabilities $\alpha_{\mathrm{slab}}^{aa}(\omega)$ have dimensions of length and are given by

$$\alpha_{\mathrm{slab}}^{aa}(\omega) = D\chi_{\mathrm{sc}}^{aa}(\omega). \tag{5}$$

We write $\mathcal{R}_a$ in terms of the susceptibility $\chi_{\mathrm{sc}}^{aa}(\omega)$ instead of in terms of the slab polarizabilities as is usually done (see Del Sole[21]), since it reflects our numerical implementation.

TABLE I. Surface-relaxed atomic Cartesian coordinates and the first few bulk atomic coordinates for the clean, monohydride, and dihydride Si(100) surfaces (see Fig. 1) in bohr. We used 5.43 Å for the lattice constant of Si. The slab is obtained by adding more bulk planes and constructing the back surface by using the front surface coordinates and applying centrosymmetry to the full slab.

	Clean 2×1				Monohydride 2×1				Dihydride 1×1		
	x	y	z		x	y	z		x	y	z
				H	0	1.81394	0	H	0	1.81394	0
					6.48035	1.81394	0		4.3884	1.81394	0
Si	0	1.81394	0	Si	0.987039	1.81394	−2.65852	Si	2.1942	1.81394	−1.73952
	4.09973	1.81394	1.32977		5.49331	1.81394	−2.65852		2.1942	−1.81394	−4.22966
	−2.07175	−1.81394	−1.3743		−0.185984	−1.81394	−4.919		5.82208	−1.81394	−6.76977
	4.79944	−1.81394	−1.25138		6.66633	−1.81394	−4.919		5.82208	1.81394	−9.3081
	1.50771	−1.81394	−4.18081		3.24017	−1.81394	−7.69074		2.1942	1.81394	−11.8447
	8.57853	−1.81394	−3.70671		10.4959	−1.81394	−7.2607		2.1942	−1.81394	−14.41
	1.33997	1.81394	−6.64272		3.24017	1.81394	−10.1361		5.82208	−1.81394	−16.9753
	8.61612	1.81394	−6.35769		10.4959	1.81394	−9.89479		5.82208	1.81394	−19.5406
	−2.28341	1.81394	−9.04406		−0.387713	1.81394	−12.5563		2.1942	1.81394	−22.1059
	4.97236	1.81394	−9.04406		6.86806	1.81394	−12.5563		2.1942	−1.81394	−24.6712
	−2.28341	−1.81394	−11.6094		−0.387713	−1.81394	−15.1216		5.82208	−1.81394	−27.2365
	4.97236	−1.81394	−11.6094		6.86806	−1.81394	−15.1216		5.82208	1.81394	−29.8018

The susceptibilities in Eq. (4) can be separated into real and imaginary parts $\chi^{aa}(\omega)=\mathrm{Re}[\chi^{aa}(\omega)]+i\,\mathrm{Im}[\chi^{aa}(\omega)]$. Within the usual independent-particle framework, the imaginary part is given by the integral over the Brillouin zone (BZ),[22]

$$\mathrm{Im}[\chi^{ab}(\omega)] = \gamma(\omega)\sum_{mn}\int \frac{d^3k}{8\pi^3} f_{nm} v^a_{mn}(\mathbf{k}) v^b_{nm}(\mathbf{k})\,\delta[\omega_m(\mathbf{k})-\omega_n(\mathbf{k})-\omega], \tag{6}$$

where $\gamma(\omega)=\pi e^2/\omega^2\hbar$, $\hbar\omega_m(\mathbf{k})$ is the energy eigenvalue of state $|m\mathbf{k}\rangle$, and $f_{nm}=f_n-f_m$ with f_n the Fermi factor equal to 1 if n is a valence band, and zero if n is a conduction band. The superscripts a and b denote Cartesian components. The velocities $\mathbf{v}_{mn}(\mathbf{k})$ in Eq. (6) are found from the matrix elements of the velocity operator $\mathbf{v}=(1/i\hbar)[\mathbf{r},H]$,

$$\mathbf{v}_{mn}(\mathbf{k})\,\delta(\mathbf{k}-\mathbf{k}') \equiv \int d^3r \langle m\mathbf{k}|\mathbf{r}\rangle\langle\mathbf{r}|\mathbf{v}|n\mathbf{k}'\rangle. \tag{7}$$

The Hamiltonian H we use here is the ground-state density-functional-theory Hamiltonian within the LDA, including a scissors correction to correct for the band gap. Neglecting the contribution of the nonlocal pseudopotential in $[\mathbf{r},H]$, one has

$$\langle\mathbf{r}|\mathbf{v}|n\mathbf{k}\rangle \approx \frac{1}{m}\mathbf{P}\psi_{n\mathbf{k}}(\mathbf{r}), \tag{8}$$

where we write $\mathbf{P}$ for $-i\hbar\boldsymbol{\nabla}$, and $\psi_{n\mathbf{k}}(\mathbf{r})=\langle\mathbf{r}|n\mathbf{k}\rangle$.

In a slab calculation one often wants to calculate the response from one of the two slab surfaces. To limit the response to one surface, Reining *et al.*,[23] Mendoza *et al.*,[24] and Mejía *et al.*[25] proposed that one should replace the differential operator $\mathbf{P}$ by $\boldsymbol{\mathcal{P}}$,

$$\mathbf{P}\rightarrow\boldsymbol{\mathcal{P}}\equiv\frac{1}{2}[\mathbf{P}S(z)+S(z)\mathbf{P}], \tag{9}$$

in a slab susceptibility calculation, where $S(z)$ is the so-called "cut function," which is usually taken to be zero over one half the supercell and unity over the other half.[26] The layer-by-layer analysis of Hogan *et al.*[14] and Castillo *et al.*[20] is a generalization of this idea, limiting the calculated current response to a particular layer of the slab. Explicitly, this is done by replacing the cut function $S(z)$ in Eq. (9) by a top-hat cut function $S_\ell(z)$ that selects a given layer,

$$S_\ell(z)=\Theta(z-z_\ell+\Delta^b_\ell)\Theta(z_\ell-z+\Delta^f_\ell), \tag{10}$$

where Θ is the Heaviside function. Here, $\Delta^{f/b}_\ell$ is the distance that the ℓth layer extends towards the front (f) or back (b) from its z_ℓ position. Thus $\Delta^f_\ell+\Delta^b_\ell$ is the thickness of layer ℓ.

There has been some question as to which, if not both, of the velocity matrix elements of Eq. (6) should be modified by the cut functions. Hogan *et al.*[14] have shown from a phenomenological description of the microscopic nonlocal dielectric susceptibility that only one of the velocity matrix elements should be modified, and Castillo *et al.*[20] use a heuristic argument, which distinguishes the perturbation from the response, to assign which of the velocities should be replaced. Here we derive this result from a microscopic approach. From perturbation theory,[22] the density matrix $c_{mn}(\mathbf{k},t)$ of the system with the applied optical field is

$$c_{mn}(\mathbf{k},t)=\frac{-ie}{\hbar}\int d\omega \frac{f_{nm}v^b_{mn}(\mathbf{k})}{\omega_{mn}(\mathbf{k})[\omega_{mn}(\mathbf{k})-\omega]}E^b(\omega)e^{-i\omega t}, \tag{11}$$

to first order in the field strength, where $\mathbf{E}(\omega)=\int dt\mathbf{E}(t)\exp(i\omega t)$ is the Fourier transform of the applied optical field, and the repeated index b is summed over. The microscopic current density that is induced throughout the slab is then given by

$$\mathbf{j}(\mathbf{r},t)=\Omega\int\frac{d^3k}{8\pi^3}\sum_{mn}c_{mn}(\mathbf{k},t)\mathbf{j}_{nm}(\mathbf{k};\mathbf{r}), \tag{12}$$

where Ω is the unit cell volume and $\mathbf{j}_{nm}(\mathbf{k};\mathbf{r})\delta(\mathbf{k}-\mathbf{k}')$ is the matrix element of the microscopic current operator $\mathbf{j}$,[27] given by

$$\mathbf{j}=\frac{e}{2}[\mathbf{v}|\mathbf{r}\rangle\langle\mathbf{r}|+|\mathbf{r}\rangle\langle\mathbf{r}|\mathbf{v}], \tag{13}$$

in a single-particle formulation, neglecting the nonlocal terms[28] that arise in a pseudopotential calculation. Integrating the microscopic current $\mathbf{j}(\mathbf{r},t)$ over the entire cell and dividing by the cell volume gives the total macroscopic current density, whose Fourier components $\mathbf{J}(\omega)$ are related to the susceptibility by $J^a(\omega)=-i\omega\chi^{ab}(-\omega,\omega)E^b(\omega)$. However, if we want the contribution from only one region of the unit cell towards the total current, we can integrate $\mathbf{j}(\mathbf{r},t)$ over the desired region. The contribution to the current density from the ℓth layer of the slab is given by

$$\frac{1}{\Omega}\int d^3r S_\ell(z)\mathbf{j}(\mathbf{r},t) \equiv \mathbf{J}^{(\ell)}(t)=\int\frac{d^3k}{8\pi^3}\sum_{mn}c_{mn}(\mathbf{k},t)\int d^3r S_\ell(z)\mathbf{j}_{nm}(\mathbf{k};\mathbf{r}), \tag{14}$$

where $\mathbf{J}^{(\ell)}(t)$ is the microscopic current in the ℓth slab. From Eq. (13) we obtain

$$\int d^3r S_\ell(z)\mathbf{j}_{nm}(\mathbf{k};\mathbf{r}) = \int d^3r S_\ell(z)\frac{e}{2}[\langle n\mathbf{k}|\mathbf{v}|\mathbf{r}\rangle\langle\mathbf{r}|m\mathbf{k}\rangle+\langle n\mathbf{k}|\mathbf{r}\rangle\langle\mathbf{r}|\mathbf{v}|m\mathbf{k}\rangle] \equiv e\mathcal{V}^{(\ell)}_{nm}(\mathbf{k}), \tag{15}$$

which defines the layer matrix elements of the velocity operator $\mathcal{V}^{(\ell)}_{nm}(\mathbf{k})$. One immediately sees that the matrix element $\mathbf{v}_{mn}(\mathbf{k})$ appearing in Eq. (11) should not be replaced by the layer matrix element $\mathcal{V}^{(\ell)}_{mn}(\mathbf{k})$. Only the velocity matrix element $\mathbf{v}_{nm}(\mathbf{k})$ directly associated with the generated current needs to be replaced by $\mathcal{V}^{(\ell)}_{nm}(\mathbf{k})$. Using Eq. (8) in Eq. (15), we

can write, for any function $S(z)$ used to identify the response from a region of the slab, that

$$\mathcal{V}_{nm}(\mathbf{k}) = \frac{1}{m}\int d^3r\{S(z)[i\hbar\,\boldsymbol{\nabla}\,\psi^*_{n\mathbf{k}}(\mathbf{r})]\psi_{m\mathbf{k}}(\mathbf{r}) + S(z)\psi^*_{n\mathbf{k}}(\mathbf{r}) \times(-i\hbar\,\boldsymbol{\nabla}\,)\psi_{m\mathbf{k}}(\mathbf{r})\} \tag{16a}$$

$$= \frac{1}{m}\int d^3r\psi^*_{n\mathbf{k}}(\mathbf{r})\left[\frac{S(z)\mathbf{P}+\mathbf{P}S(z)}{2}\right]\psi_{m\mathbf{k}}(\mathbf{r}) \tag{16b}$$

$$= \frac{1}{m}\int d^3r\psi^*_{n\mathbf{k}}(\mathbf{r})\mathcal{P}\psi_{m\mathbf{k}}(\mathbf{r}) \equiv \frac{1}{m}\mathcal{P}_{nm}(\mathbf{k}). \tag{16c}$$

Here an integration by parts is performed on the first term of the right-hand side of Eq. (16a); since the $e^{-i\mathbf{k}\cdot\mathbf{r}}\psi_{n\mathbf{k}}(\mathbf{r})$ are periodic over the unit cell, the surface term vanishes. From Eqs. (16) we see that the replacement introduced in Eq. (9), and its generalization using Eq. (10), is a natural result of the microscopic treatment of the induced current given in Eq. (14).

The Fourier components of the microscopic current $\mathbf{J}^{(\ell)}(\omega)$ can be related to a position-dependent susceptibility as

$$J^{(\ell)a}(\omega) = -i\omega\chi^{(\ell)ab}(-\omega,\omega)E^b(\omega), \tag{17}$$

which from Eq. (14) we obtain for the imaginary part of the ℓth layer susceptibility

$$\mathrm{Im}[\chi^{(\ell)ab}(\omega)] = \gamma(\omega)\sum_{mn}\int\frac{d^3k}{8\pi^3}f_{nm}\mathcal{V}^{(\ell)a}_{mn}(\mathbf{k})v^b_{nm}(\mathbf{k}) \times\delta[\omega_m(\mathbf{k})-\omega_n(\mathbf{k})-\omega], \tag{18}$$

where $\mathcal{V}^{(\ell)}_{mn}(\mathbf{k})$ is given by Eq. (15). Using $\chi^{(\ell)ab}(\omega)$ for the surface susceptibility $\chi^{ab}_{\mathrm{sc}}(\omega)$ in the definition [Eq. (4)] of the reflectivity allows one to calculate the contribution to a linear optical response, such as the RA or RD spectrum, for a specified layer.

III. COMPUTATIONAL DETAILS

Our layer-by-layer analysis is done using the ABINIT plane-wave code.[29] A self-consistent calculation is first made to determine the Kohn-Sham potential for the relaxed surface structures, and then the matrix elements are determined for the **k** points used in the integration of Eq. (18). We use a different grid of **k** points in the irreducible BZ for each of the surface calculations (see below), and a grid of 3654 points for the bulk calculations. For the bulk calculation a scissor shift of 0.98 eV is used to adjust the theoretical bulk gap to the experimental value of 3.5 eV at the Γ point of the BZ. For the surface calculation a scissor shift of 0.68 eV is used in order to have the first negative peak of the RA to coincide with the experimental results of Shioda *et al.*[30] for the clean surface (see the negative peak at 1.68 eV in Figs. 5 and 8 below). We take an energy cutoff of 15 Ha and adopt LDA Troullier-Martins pseudopotentials for Si and H. We find converged results for all the quantities of interest in this work.

The code provides the plane-wave coefficients $C_{n\mathbf{k}}(\mathbf{G})$ of the expansion

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{n\mathbf{k}}(\mathbf{G})\exp[i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}]. \tag{19}$$

The velocities $\mathcal{V}^{(\ell)}_{mn}(\mathbf{k})$ are calculated through Eq. (15) using the approximation of Eq. (8), so that

$$\mathcal{V}^{(\ell)}_{mn}(\mathbf{k}) = \frac{\hbar}{2}\sum_{\mathbf{G},\mathbf{G}'}C^*_{m\mathbf{G}'}(\mathbf{k})C_{n\mathbf{G}}(\mathbf{k})[2\mathbf{k}+\mathbf{G}+\mathbf{G}'] \times\delta_{\mathbf{G}_\parallel,\mathbf{G}'_\parallel}f_\ell(G_\perp - G'_\perp), \tag{20}$$

where the reciprocal lattice vectors **G** are decomposed into components parallel to the surface $\mathbf{G}_\parallel$, and perpendicular to the surface $G_\perp\hat{\mathbf{z}}$, so that $\mathbf{G}=\mathbf{G}_\parallel+G_\perp\hat{\mathbf{z}}$ and

$$f_\ell(g) = \frac{1}{L}\int_{z_\ell-\Delta^b_\ell}^{z_\ell+\Delta^f_\ell}e^{igz}dz. \tag{21}$$

The double summation over the **G** vectors can be efficiently done by creating a pointer array to identify all the plane-wave coefficients associated with the same $\mathbf{G}_\parallel$. We take z_ℓ at the center of an atom that belongs to layer ℓ, and thus Eq. (18) gives the ℓth atomic-layer contribution to the imaginary part of the polarizability of the slab. The procedure taken in Eq. (20) avoids the calculation of the matrix elements of the cut functions, as done by Castillo *et al.*,[20] Mendoza *et al.*,[24] and Mejía *et al.*,[25] and permits a faster and more direct calculation of the matrix elements of $\mathcal{V}^{(\ell)}$ required in Eq. (18). Note that if we take the cut function $S(z)$ to be unity through the whole slab, then $f_\ell(g)=\delta_{g,0}$ and from Eq. (20) one recovers the usual result for the matrix elements of v^a_{mn}.

Although the way in which the slab is partitioned into layers is arbitrary, for both the monohydride and dihydride surfaces we have taken the topmost layer to include the hydrogen atom, and for the clean surface we have taken the topmost layer to include the dimer alone. This topmost layer extends into the vacuum to the midway point between two of the periodically repeated slabs (see Fig. 1). Since the atomic planes are parallel to each other, we have chosen each layer below the surface to include one atomic plane where all the Si atom coordinates have the same z value. The qualitative features that will follow depend on this choice.

The nonlocality of the pseudopotentials we use introduces a contribution to the velocity matrix elements, which has been discussed before, even in the context of slab calculations.[15] These contributions are demanding to compute. To estimate the error in neglecting the nonlocal contribution when using Eq. (8), we compare the velocity matrix elements computed from our pseudopotential calculation, using Eq. (20) with $S(z)=1$, to the matrix elements from an all-electron augmented plane-wave plus local-orbital (APW+lo) method, which uses the full local-crystal potential.[31] The relevant parameter for the accuracy of the calculation is $RK_{\max}=4.5$, which is a product of the "muffin-tin" radius R and the maximum value for the plane-waves vectors $K_{\max}$.

The diagonal components of the velocity matrix elements satisfy

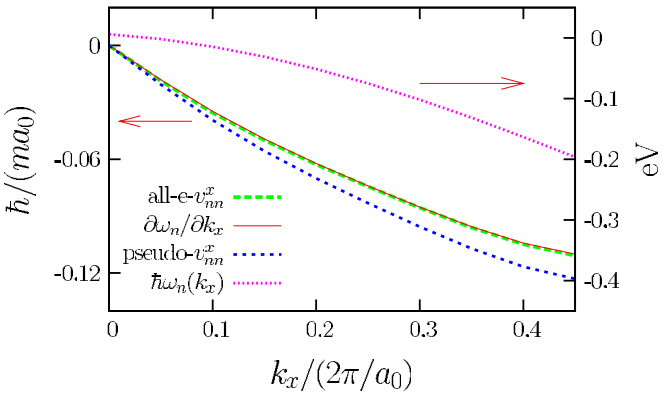

FIG. 2. (Color online) We show $\partial\omega_n(\mathbf{k})/\partial k_x$ and $v^x_{nn}(\mathbf{k})$ vs k_x along the Γ-J line of the two-dimensional (2D) BZ of the clean Si(100)2×1 surface. The thin line is the topmost valence energy band with the scale in the right vertical axis (see text for details).

$$\frac{\partial\omega_n(\mathbf{k})}{\partial k_i} = v^i_{nn}(\mathbf{k}). \tag{22}$$

In Fig. 2 we show $\partial\omega_n(\mathbf{k})/\partial k_x$ and $v^x_{nn}(\mathbf{k})$ vs k_x, where n corresponds to the topmost valence band of the electronic structure of the clean Si(100)2×1 surface. The pseudopotential and all-electron band structures are essentially the same, so there is only one plot of $\hbar\omega(\mathbf{k})$ and $\partial\omega_n(\mathbf{k})/\partial k_x$. We see that the all-electron energies and velocities satisfy Eq. (22), whereas the pseudopotential ones deviate from it. Yet the error is rather small. We have checked that the same holds true for other components of the velocities $\mathbf{v}_{nn}$ and other directions through the BZ.

To judge the overall consequences of calculating $\mathbf{v}_{nm}$ without the nonlocal contribution, we show in Fig. 3 the imaginary part of the bulk dielectric function $\epsilon_B(\omega)$ calculated for both schemes, using a broadening of 30 meV. The calculation of $\epsilon_B(\omega)$ is done using the usual extension of Eq. (6) to the infinite crystal, with $\epsilon(\omega)=1+4\pi\chi(\omega)$. As we can see, the spectral features are the same for both calculations; however, our pseudopotential calculation gives a larger value for the bulk dielectric function. Since the pseudopotential band structure is nearly identical to the all-electron band

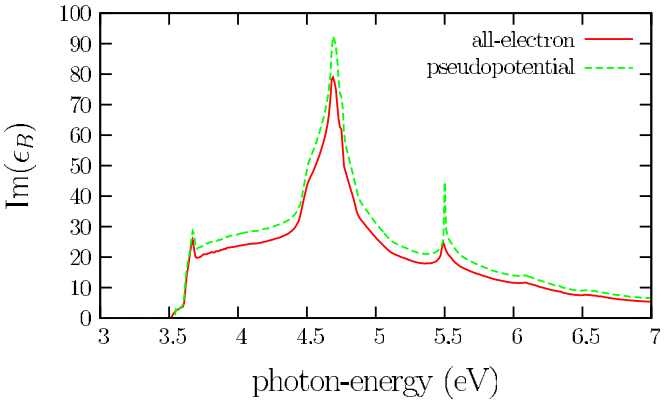

FIG. 3. (Color online) Plot of the imaginary part of the bulk dielectric function for both codes.

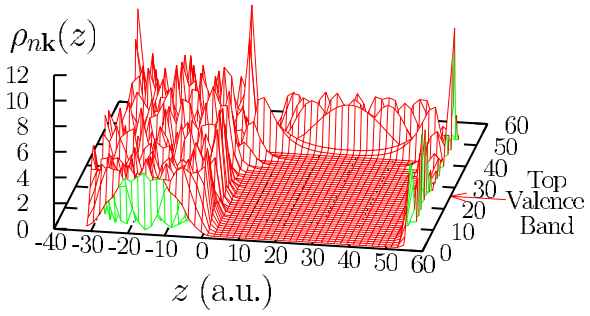

FIG. 4. (Color online) Plot of $\rho_{n\mathbf{k}}(z)$ vs z and band index for $\mathbf{k}=0$, i.e., the Γ point in the 2D BZ of a Si(100) dihydride surface with 12 Si planes.

structure, with typical differences less than 0.1 meV, the disagreement is due to the omission of the nonlocal part of the velocity matrix elements. We find similar behavior for the calculated slab susceptibilities for our Si(100) surfaces. Since it is prohibitively complicated to implement a full layer-by-layer analysis using the APW+lo basis, from now on we use only the pseudopotential scheme, as it readily permits the evaluation of $\boldsymbol{\mathcal{V}}^{(\ell)}_{mn}(\mathbf{k})$ through Eq. (20).

In repeated-slab calculations it is important to have a sufficiently large vacuum region between the slabs. A large vacuum region insures that the slabs are independent. It is also important to have enough atomic planes, so that the two surfaces of one slab do not interact. At high enough energies, the calculated conduction-band states can always be expected to extend between slabs, and one expects the appearance of ionization states as is seen in Fig. 4. For the dihydride slab, we plot the planar-integrated charge density $\rho_{n\mathbf{k}}(z)=\int dxdy|\psi_{n\mathbf{k}}(\mathbf{r})|^2$ for the Γ-point states as a function of z and band index n. For this figure we have set the vacuum region to $55a_0$ and used 12 atomic planes of Si with a_0 Bohr's radius. One can see that by $n=42$ the conduction-band state extends between the surfaces of neighboring slabs. Increasing the number of atomic planes used in the calculation increases the energy at which tunneling begins. This is because adding more atomic planes, thus increasing the bulk region, deepens the electrostatic potential associated with the slab. Also, increasing the vacuum region for a fixed number of atomic planes affects the decay into the vacuum region of the higher-conduction states.

The details of the ionization states that arise in a calculation such as ours are obviously artifacts of using a slab to model the behavior of a surface, and any contribution those states make to a calculated optical response is necessarily suspect. To investigate the effect of these ionization states on the RA spectrum of clean Si(100), we make two calculations, each using 16 layers of Si atoms and a vacuum equivalent to the distance of 20 Si layers. In the first we include the first 14 conduction bands; none of these include ionization states. In the second we use 66 conduction bands; some of the higher-energy states here are ionization states. The results are shown in Fig. 5; in this figure, and from now on, we use a smearing of 300 meV to simulate the resolution typical in RAS and RDS. The results of the two calculations agree up to about 3.5 eV; at higher energies the difference signals the onset of

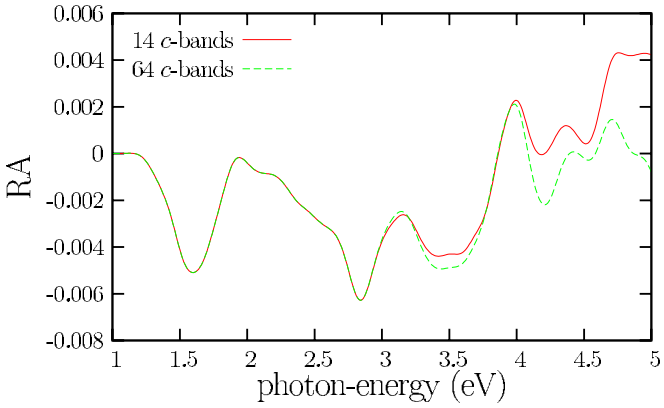

FIG. 5. (Color online) Plot of RA using two values of the total number of conduction bands used in the sum of Eq. (18). The slab is for a clean Si(100) surface (see the text for details).

contributions from the higher 52 bands. Nonetheless, the differences are generally small and only quantitative up to 5 eV; the situation is similar for the dihydride and monohydride Si(100) surfaces. Thus up to 5 eV the states of the upper 52 bands, including the ionization states, make only a small contribution to the RA spectrum. Of course, we cannot assert with certainty that the ionization states in a true surface calculation would make only a small contribution to the RA spectrum in this energy range, since they would differ from the ionization states in our slab calculation. Still, we take the results shown in Fig. 5 as an indication that our calculations of the RA and RD spectra for the surfaces studied are likely physically significant up to 5 eV, and present them up to this photon energy.

IV. RESULTS

We calculate the optical response of a clean Si(100) surface and two H-covered Si(100) surfaces. The clean surface we consider is a 2×1 surface reconstruction characterized by a tilted dimer formed between the two topmost Si atoms. These asymmetric buckled surface dimers are along the $[01\bar{1}]$ crystallographic direction that we take as x, and thus the dimer rows along the surface are along the [011] or y direction. One of the H-covered surfaces is a monohydride surface in which the two dangling bonds of the dimer are H saturated (rendering a flat dimer), and the other is a dihydride surface in which the H saturates each of the two dangling bonds, leading to a bulk ideally terminated surface which exhibits a 1×1 surface-unit cell. This dihydride surface is thought to be created if enough H is added to the surface.[32] The structures that we obtain are calculated using the standard DFT-LDA molecular dynamics technique, and agree very well with previous studies.[13]

In Fig. 6 we show the imaginary part of $\chi^{(\ell)aa}(\omega)$, with $a=x$ and y, for a dihydride Si(100) slab with 32 Si atomic planes, 2 H atomic planes, and a vacuum equivalent to 10 Si layers. The number of **k** points in the irreducible BZ is 196. Since there are 34 atomic layers, we use 17 unique layers in the upper half of the slab for the analysis. By inversion symmetry the 17 layers in the lower half of the slab give the same response. We group the 17 layers into three sets. The set containing $\ell=1$ to $\ell=6$, which we refer to as the surface region, is seen to be highly anisotropic, with $\chi^{(\ell)xx}(\omega)$ very different from $\chi^{(\ell)yy}(\omega)$. We note, too, that the anisotropy of the optical response is the largest for $\ell=2$, corresponding to the topmost Si atomic plane; the H layer ($\ell=1$) also has a rather large anisotropy. This is not unreasonable, since the H atoms form chains oriented along the x direction. The next set, containing $\ell=7$ to $\ell=12$, is anisotropic but less so, and this region can be identified with the transition from the surface to the bulk, at least as far as the linear optical response is concerned. Finally, the last set containing $\ell=13$ to $\ell=17$ is almost isotropic, as expected, since these layers are deep within the bulk of the system. This behavior of $\boldsymbol{\chi}^{(\ell)}(\omega)$ is physically expected, and shows how the optical response of a semi-infinite crystal changes as we move from the surface into the bulk of the system. Similar results for the decomposition of $\boldsymbol{\chi}^{(\ell)}(\omega)$ are obtained for the clean and the monohydride surface, but instead of presenting them we focus on the RA and RD signals, which could be compared with experiments.

In Fig. 7 we show the RA signal, decomposed into the same layers as those shown in Fig. 6, for the dihydride Si(100)1×1 surface; we also show the total RA signal,

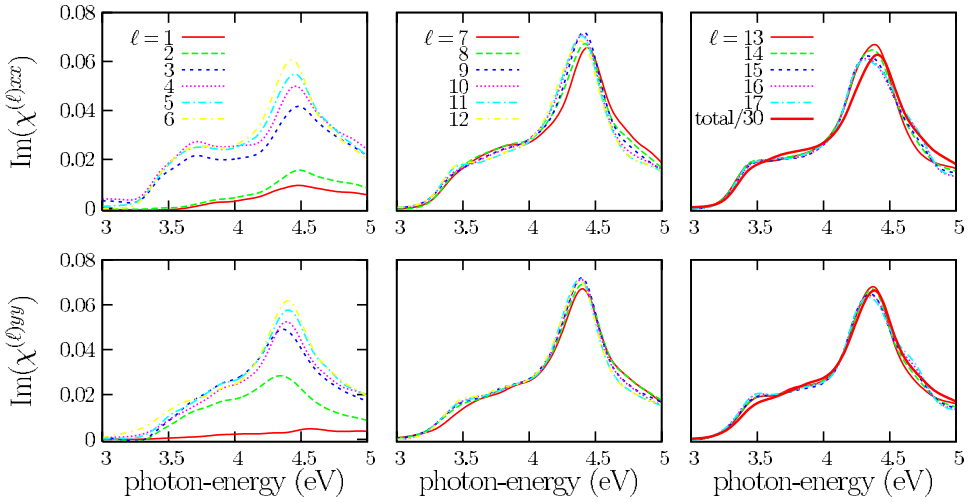

FIG. 6. (Color online) Plot of $\mathrm{Im}[\chi^{(\ell)xx}(\omega)]$ and $\mathrm{Im}[\chi^{(\ell)yy}(\omega)]$ vs ℓ for a dihydride Si(100) slab with 32 Si atomic planes.

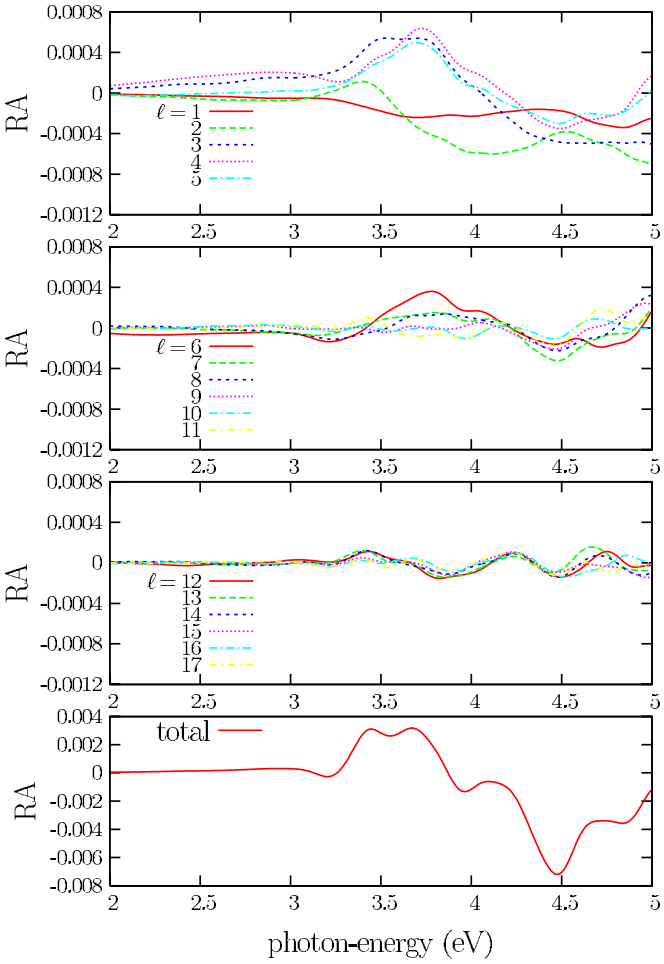

FIG. 7. (Color online) RA for the dihydride Si(100)1×1 surface as a function of the layer ℓ, where $\ell=1$ corresponds to the H. The bottom panel shows the total RA signal with a different vertical scale.

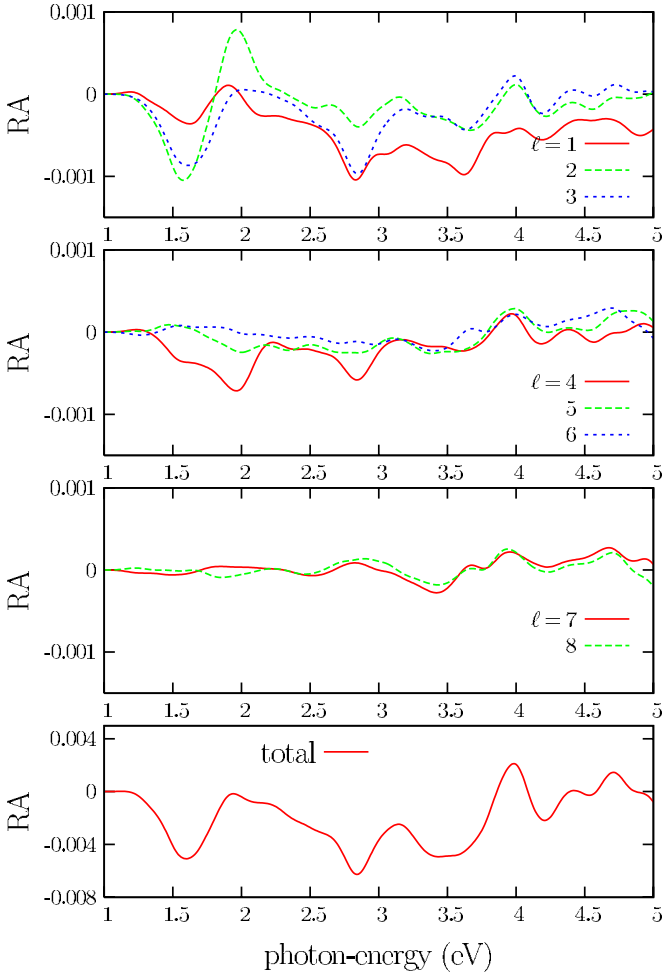

FIG. 8. (Color online) RA for the clean Si(100)2×1 surface as a function of the layer ℓ, where $\ell=1$ corresponds to the Si dimer layer. The bottom panel shows the total RA signal with a different vertical scale.

which equals the sum of RA(ℓ) from $\ell=1$ to $\ell=17$. In agreement with the analysis of Fig. 6, we see that RA(ℓ) gets smaller as we move into the bulk, and that the surface-related RA gives the largest contribution to the total signal. The different features seen in the total RA can be identified with the particular layers. For instance, the feature between 2.6 and 4 eV has a strong contribution from the first three Si layers; the H layer has only a small contribution, and the bulk layers also contribute much less than the surface layers or the transition layers. In principle, the RA of the bulk layers should be zero; deviations from this are likely due to the still limited number of planes in the slab. We have checked that the calculation with 28 Si layers gives similar results to the one with 32 Si layers. However, for this particular dihydride surface, it seems that 32 Si layers is the minimum size of the slab that captures the expected behavior of the RA signal in the bulk.

In Fig. 8 we show the RA signal for the clean Si(100)2×1 surface, which is characterized by buckled surface dimers. Here 16 atomic layers were used with a vacuum equivalent to 10 Si layers, and 66 **k** points in the irreducible BZ, which capture the essential physics for this surface. This was verified by repeating the calculations for 12 layers and the same vacuum, and obtaining similar results. Indeed, we see that all the features in the total signal are well correlated with the first four layers of Si, and that the dimer layer ($\ell=1$) has a sizable contribution. Before layer-by-layer analyses were carried out in RA studies, it was customary to assume that features at low energies are related to surface states, since they usually lie in the bulk gap; only a decomposition of the transitions from surface or bulk valence states to surface or bulk conduction states was then carried out.[11] But, from Fig. 8 we can see how each layer contributes, and it is interesting to note that the fourth layer has a rather large contribution at low energies, meaning that the anisotropy of this surface extends to deep layers. We also see that the contribution of the layers deeper than the $\ell=4$ layer is only marginal. The total RA signal is in agreement with the calculation of Palummo *et al.*[13] and the experimental results of Shioda *et al.*[30] The same qualitative features are also seen for the monohydride Si(100)2×1 surface as shown in Fig. 9. Here we see that the second Si layer mainly dominates the

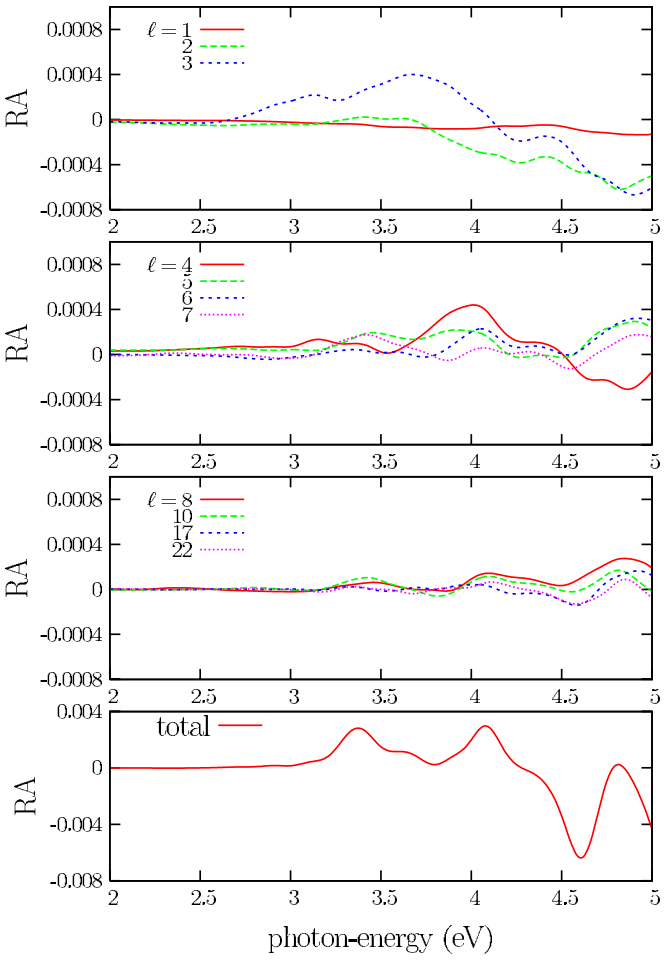

FIG. 9. (Color online) RA for the monohydride Si(100)2×1 surface as a function of the layer ℓ, where $\ell=1$ is the H layer and $\ell=2$ corresponds to the dimer. The bottom panel shows the total RA signal with a different vertical scale (see the text for details).

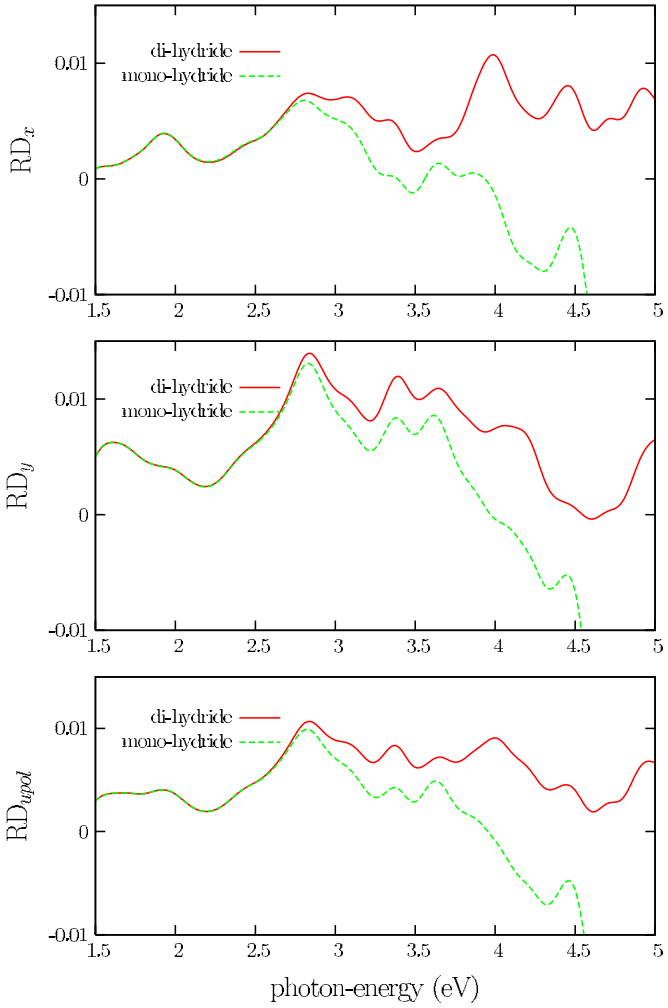

FIG. 10. (Color online) Polarized and unpolarized RD, taken the clean Si(100)2×1 surface as the clean reference, and the monohydride and dihydride surfaces as the H-covered surfaces.

RA below 3 eV, and that the dimer has a moderate contribution only above 3 eV. The H layer has a negligible contribution towards the RA; however, it is responsible for suppressing the surface states seen in the clean surface below 2 eV. For this surface we had to use a rather large number of Si atomic planes in order to find converged results; indeed, Fig. 9 was obtained with 44 Si planes, a vacuum equivalent to 20 Si layers, and 66 **k** points in the irreducible BZ. Notice that for the layers deep within the bulk of this slab, we have only plotted a selected set of layers, with $\ell=22$ the deepest bulk-like layer. So again, the layer-by-layer analysis shows that the asymmetric optical response is not concentrated exclusively on the topmost layer, as might have been thought, but it extends to subsurface layers.

In Fig. 10 we show the RD spectrum of the monohydride and dihydride surfaces, taking the clean Si(100)2×1 surface as the reference. We notice that below 2.75 eV both RD signals are identical since the H surfaces are not optically active, as can be seen from the RA displayed in Figs. 7 and 9. Above 2.75 eV, the dihydride signal is larger than the monohydride signal, and the two signals are different; there are features seen in the dihydride surface that are absent in the monohydride surface and vice versa. This result is in qualitative agreement with recent measurements of Borensztein *et al.*,[32,33] where the RD spectrum of H-covered surfaces referenced to a clean Si(100) surface was investigated.

V. CONCLUSIONS

We have presented a method for calculating the layer-by-layer contribution to the linear optical-surface response of a semiconductor surface following the ideas presented by Hogan *et al.*[14] and Castillo *et al.*[20] We have directly derived the way in which the velocity matrix elements should be modified in such an approach. We have shown that the nonlocal contribution to the momentum-matrix elements, neglected in the formalism used, is very small. We have also investigated the important issue of how many atomic planes and how large a vacuum region should be used for calculating the layer-by-layer response.

Experimental studies and theoretical analyses of RA and RD signals provide insight into the anisotropy of surfaces.

The layer-by-layer analysis of such a signal allows us to determine the size of the surface region that contributes to that anisotropy, at least as revealed by these optical spectroscopies. For the clean and hydrogenated Si(100) surfaces studied here, five atomic layers are mainly responsible for the RA and RD signal. Our results agree qualitatively with available experimental results, and in particular they support the idea that RDS could be used to optically study the possible breaking of the surface dimer due to adsorbed molecules. It would be desirable to incorporate a detailed analysis of how the electronic transitions contribute to the optical response of the layers, and to extend the present formalism to include many-body, excitonic, and local field effects.

ACKNOWLEDGMENTS

B.S.M. acknowledges the National Science and Engineering Research Council of Canada and Photonics Research Ontario and the Department of Physics of the University of Toronto, for the support during a sabbatical year. F.N. acknowledges support by the Ontario Graduate Scholarship program. N.A. acknowledges support by CONCYTEG 05-04-k117-026-A03 and CONACYT SEP-2003-C02-42576. J.E.S. acknowledges the National Science and Engineering Research Council of Canada and Photonics Research Ontario.

[1] *Proceedings of the 5th International Conference on Optics of Surfaces and Interfaces*, Berlin, Germany, edited by B. S. Mendoza (Wiley-VCH, Berlin, 2003) [Phys. Status Solidi C **0**, 2915 (2003)].
[2] J. D. E. McIntyre and D. E. Aspnes, Surf. Sci. **24**, 417 (1971).
[3] N. Arzate, B. S. Mendoza, and R. Vázques, J. Phys.: Condens. Matter **16**, s4259 (2004).
[4] N. Arzate and B. S. Mendoza, Phys. Rev. B **63**, 113303 (2001).
[5] C. D. Hogan and C. H. Patterson, Phys. Rev. B **57**, 14843 (1998).
[6] B. S. Mendoza and W. L. Mochán, Phys. Rev. B **55**, 2489 (1997).
[7] C. M. J. Wijers, P. L. de Boeij, C. W. van Hasselt, and T. Rasing, Solid State Commun. **93**, 17 (1995).
[8] W. L. Mochán and R. G. Barrera, Phys. Rev. Lett. **55**, 1192 (1985).
[9] X. López-Lozano, O. Pulci, C. Noguez, K. Fleischer, R. Del Sole, and W. Richter, Phys. Rev. B **71**, 125337 (2005).
[10] R. A. Vázquez-Nava, B. S. Mendoza, and C. Castillo, Phys. Rev. B **70**, 165306 (2004).
[11] B. S. Mendoza, N. Esser, and W. Richter, Phys. Rev. B **67**, 165319 (2003).
[12] B. S. Mendoza, R. Del Sole, and A. I. Shkrebtii, Phys. Rev. B **57**, R12709 (1998).
[13] M. Palummo, G. Onida, R. Del Sole, and B. S. Mendoza, Phys. Rev. B **60**, 2522 (1999).
[14] C. Hogan, R. Del Sole, and G. Onida, Phys. Rev. B **68**, 035405 (2003).
[15] O. Pulci, G. Onida, R. Del Sole, and A. J. Shkrebtii, Phys. Rev. B **58**, 1922 (1998).
[16] A. Incze, R. Del Sole, and G. Onida, Phys. Rev. B **71**, 035350 (2005).
[17] P. H. Hahn, W. G. Schmidt, and F. Bechstedt, Phys. Rev. Lett. **88**, 016402 (2002).
[18] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **83**, 856 (1999).
[19] O. Pulci, G. Onida, R. Del Sole, and L. Reining, Phys. Rev. Lett. **81**, 5374 (1998).
[20] C. Castillo, B. S. Mendoza, W. G. Schmidt, P. H. Hahn, and F. Bechstedt, Phys. Rev. B **68**, 041310(R) (2003).
[21] R. Del Sole, *Photonic Probes of Surfaces* (Elsevier Science, New York, 1995).
[22] J. E. Sipe and A. I. Shkrebtii, Phys. Rev. B **61**, 5337 (2000).
[23] L. Reining, R. Del Sole, M. Cini, and J. G. Ping, Phys. Rev. B **50**, 8411 (1994).
[24] B. S. Mendoza, M. Palummo, G. Onida, and R. Del Sole, Phys. Rev. B **63**, 205406 (2001).
[25] J. Mejía, C. Salazar, and B. S. Mendoza, Rev. Mex. Fis. **50**, 134 (2004).
[26] The $S(z)$ function was introduced (Refs. 23–25) in the context of second harmonic generation; however, it can be used in linear response as well.
[27] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics* (Wiley, New York, 1977), Vol. 1, p. 239.
[28] H. Zhong, Z. H. Levine, D. C. Allan, and J. W. Wilkins, Phys. Rev. B **48**, 1384 (1993).
[29] First-principles computation of material properties: the ABINIT software project, X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet *et al.*, Comput. Mater. Sci. **25**, 478 (2002); www.abinit.org
[30] R. Shioda and J. van der Weide, Phys. Rev. B **57**, R6823 (1998).
[31] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, and J. Luitz, *An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties* (Technische Universität Wien, Austria, 2001).
[32] Y. Borensztein, O. Pluchery, and N. Witkowski, Phys. Rev. Lett. **95**, 117402 (2005).
[33] Y. Borensztein and N. Witkowski, J. Phys.: Condens. Matter **16**, S4301 (2004).