Simplified bond-hyperpolarizability model of second harmonic generation

G. D. Powell, J.-F. Wang, and D. E. Aspnes*
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202
(Received 19 November 2001; published 17 May 2002)

We show that the anisotropies of second-harmonic-generation (SHG) intensities of vicinal (111) and (001)Si-SiO_2 interfaces can be described accurately as dipole radiation originating from the anharmonic motion of bond charges strictly along bond directions. This simplified bond-hyperpolarizability model not only substantially simplifies the description of SHG, but also provides a microscopically physical and mathematically more efficient picture of the process than those found in standard phenomenological treatments employing tensor or Fourier coefficients. Using this approach we obtain an analytic solution for the expected response of (111) terraces, and by comparing to data show that the effective angles of incidence and observation for the (111)Si-SiO_2 interface are not those measured in the laboratory but correspond to those refracted at the air-SiO_2 interface. For (111) vicinal interfaces at 765 nm SHG absorption is found to occur mainly for the step bond. The success of this formulation indicates that in many, if not most, cases the description of SHG may be simpler than that of the linear-optical response.

DOI: 10.1103/PhysRevB.65.205320 PACS number(s): 42.65.An, 42.65.Ky, 78.68.+m

I. INTRODUCTION

The development of the femtosecond (fs) laser has generated a significant resurgence of interest in the use of nonlinear-optical (NLO) phenomena as a means of studying bulk materials, thin films, interfaces, and surfaces.[1,2] NLO phenomena possess richer selection rules than their linear-optic equivalent, and are therefore intrinsically more powerful as diagnostic tools.[1–3] In addition, the microscopic symmetry of surfaces and interfaces is lower than that of bulk material, which can provide an additional means of isolating contributions from these regions. For example, with the exception of the bulk quadrupole contribution, which is relatively small for the situations discussed here, second-harmonic generation (SHG) is forbidden in the bulk of materials such as Si that possess inversion symmetry, and in amorphous materials such as SiO_2 where the bonds are oriented in essentially random directions. Thus SHG of oxidized or nitrided Si wafers originates almost entirely from an interface region no more than several atomic layers thick, where there is a regular repeating geometry of the bonds and the bonds are asymmetric. As a result SHG has become a widely used probe for studying the interfaces of this technologically important materials system.[1,2,4–12]

However, NLO spectra remain difficult to acquire relative to their linear-optical counterparts, so with few exceptions[12] SHG data have been limited to relatively narrow spectral ranges. In addition, first-principles theoretical descriptions of NLO phenomena in general and SHG in particular are highly arcane, and being based on one-electron energy-band theory, are not well adapted to providing direct physical insight especially at the level of interface bonding.[13] As a result emphasis has been placed on acquiring dependences of NLO intensities on sample azimuths at a single wavelength or narrow range of wavelengths, and representing these dependences either simply as polar plots[14] or more usually as listings of Fourier coefficients.[5–12] While this approach is at least consistent with general phenomenological treatments,[15–17] it also does not provide much insight, especially at the level of the individual bonds.

Here, we take a more basic approach for describing in particular SHG of materials and interfaces, not necessarily with the objective of outlining an exact, first-principles microscopic formulation but with the goal of providing a framework in which NLO spectra in general and SHG spectra in particular can be represented and analyzed at the bond level in simple terms. This framework also accomplishes the secondary objective of representing these data in terms of parameters that can be calculated from first principles, and thus also provides a convenient interface between experiment and theory. Specifically, we assume that the system consists of charges localized in bonds, then follow a three-step procedure. First, we recognize that the applied field causes a displacement of charge (dipole) at each bond site, which varies anharmonically with time under the action of the applied field and appropriate restoring and dissipation forces. We summarize the nonlinear part of this motion in terms of complex hyperpolarizabilities that can be obtained in principle by solving the standard equation of motion, as discussed for example by Shen.[3] Next, we make the simplifying assumption that for SHG the only relevant anharmonic motion for SHG is that along the bond axis, as discussed in the next paragraphs. Finally, we calculate the far-field intensity as the square of the superposition of fields radiated by these charges in the dipole approximation. For present purposes we also assume that the relevant bond directions are those of the underlying bulk crystalline material. We show that this simplified bond-hyperpolarizability model (SBHM) leads to an excellent description of SHG signals from both (111) and (001)Si-SiO_2 interfaces, and in addition to being mathematically more efficient than previous approaches provides new physical insight at the microscopic level, not only of the interfaces under investigation but also of the physics of SHG itself.

Bond models have long been used to describe optical properties, so we make some comments to place the present work in perspective. Our description of SHG intensity originating from dipole radiation is reminiscent of the extinction

theorem of Ewald and Oseen in linear optics,[18–20] which dates from 1912 and 1915 and will be discussed in more detail below. Early NLO work[21–23] was directed toward describing phenomenologically the dc or low-frequency nonlinear susceptibility of bulk materials. This was given a more detailed theoretical foundation and extended to surfaces of centrosymmetric materials by later workers, most notably Liebsch and Schaich[24] for metals, and Schaich and Mendoza[25] and Wijers *et al.*[26] for covalent materials. In particular, Wijers *et al.* calculate explicitly the far-field intensity as a superposition of radiation from individual (anharmonic) dipoles. Patterson *et al.*[27] also performed cluster calculations to evaluate the nonlinear susceptibility of As-terminated Si and Ge surfaces within a force model. However, the above treatments do not make the distinguishing assumption of the present work and thus are led to expressions that are sufficiently complex to require final descriptions in terms of Fourier coefficients or phenomenological tensor coefficients.

The assumption that the only relevant charge motion is that along the bond axis is equivalent to assuming that the bonds are rotationally symmetric, since symmetric motion, even if anharmonic, can give rise only to odd orders of nonlinearity. The calculations of Patterson *et al.*[27] show that the assumption of cylindrical symmetry for individual bonds is not necessarily a good approximation, but we justify its use because the restricted-motion assumption does provide a highly accurate representation of SHG data with a minimum number of parameters, implying either substantial simplifications or outright cancellations in representations involving effective bonds. A bond-polarizability model that also includes the assumption of cylindrical symmetry for individual bonds was recently presented by Mendoza and co-workers.[28,29] However, Mendoza *et al.* describe SHG as a product of the linear polarizabilities, for which the transverse polarizability must be considered even though the assumption of cylindrical symmetry means that its contribution must vanish in the final expression for SHG. Hence these workers overlooked the essential simplification that results from this assumption.

In short, the SBHM (1) incorporates at the most basic level the intrinsic symmetry of a given system, (2) allows NLO phenomena to be described with only a very small number of bonds (four for the tetrahedrally bonded semiconductors and their interfaces of interest here), (3) yields physical insight by providing a direct connection between the observed response and the microscopic physical properties of the individual bonds, (4) is sufficiently simple to allow analytic expressions to be derived, which we show leads to further physical insights, and (5) provides a convenient and simple interface between theory and experiment. For NLO phenomena that originate from a region no more than a few atomic layers thick a formulation on the basis of bond properties is clearly more natural, and desirable, than a band-structure approach where the wave functions are extended over the entire crystal and the necessary localization can be achieved only by the coherent superposition of many waves.[13] In the configurations investigated here four bonds are used but two of the four are effectively equivalent, reducing the number of complex hyperpolarizabilities required to three. We treat these phenomenologically as fitting parameters, although they could be calculated from first principles.

II. THEORY

A. Basic formulation

Consistent with our assumption that the only relevant charge motion is that along the bond axis, we consider a simple one-dimensional force model as discussed by Shen.[3] We let the direction of the jth bond in a unit cell or other appropriate set of bonds be defined by the unit vector $\hat{\mathbf{b}}_j$. We suppose an applied field $\vec{\mathbf{E}}e^{-i\omega t}$, noting that the field $\vec{\mathbf{E}}_j$ at the jth bond site may be different from $\vec{\mathbf{E}}$ owing to contributions from other induced dipoles in the area (local-field effect). However, for simplicity we assume here that $\mathbf{E}_j = \mathbf{E}$. Then the motion of the charge q_j along $\hat{\mathbf{b}}_j$ can be described by the equation of motion

$$F = q_j\vec{\mathbf{E}}\cdot\hat{\mathbf{b}}_j e^{-i\omega t} - \kappa_1(x-x_0) - \kappa_2(x-x_0)^2 - b\,dx/dt = m\,d^2x/dt^2, \tag{1}$$

where x is the position of the charge, x_0 is its equilibrium position, κ_1 and κ_2 are the harmonic (Hooke's Law) and anharmonic spring constants, respectively, and b is a frictional coefficient representing losses. Assuming that x can be written

$$x = x_0 + \Delta x_1 e^{-i\omega t} + \Delta x_2 e^{-i2\omega t}, \tag{2}$$

it follows that to lowest order the linear $\mathbf{p}_{1j}$ and first-order nonlinear $\mathbf{p}_{2j}$ parts of the induced dipole $\vec{\mathbf{p}}_j = q_j\Delta\vec{\mathbf{x}}_j$ corresponding to the jth charge are given in this model by

$$p_{1j} = q_j\Delta x_1 = \frac{q_j^2\vec{\mathbf{E}}\cdot\hat{\mathbf{b}}_j}{\kappa_1 - m\omega^2 - ib\omega} = \alpha_{1j}(\hat{\mathbf{b}}_j\cdot\vec{\mathbf{E}}), \tag{3a}$$

$$p_{2j} = q_j\Delta x_2 = \frac{q_j\kappa_2\Delta x_1^2}{\kappa_1 - 4m\omega^2 - ib2\omega} = \alpha_{2j}(\hat{\mathbf{b}}_j\cdot\vec{\mathbf{E}})^2, \tag{3b}$$

where α_{1j} and α_{2j} are the microscopic (first-order) polarizability and (second-order) hyperpolarizability, respectively, of the jth bond. Since the displacement Δx_2 is assumed to be along the bond axis, we can write the associated polarization per unit volume $\vec{\mathbf{P}}$ as the sum of the individual polarizations, which with the above assumptions becomes to within a scaling factor

$$\vec{\mathbf{P}} = \frac{1}{V}\sum_j \vec{\mathbf{p}}_j = \frac{1}{V}\sum_j (\alpha_{1j}\hat{\mathbf{b}}_j\hat{\mathbf{b}}_j)\cdot\vec{\mathbf{E}} + \frac{1}{V}\sum_j (\alpha_{2j}\hat{\mathbf{b}}_j\hat{\mathbf{b}}_j\hat{\mathbf{b}}_j)\cdot\cdot\,\vec{\mathbf{E}}\vec{\mathbf{E}} \tag{4a}$$

$$= \chi_1\cdot\vec{\mathbf{E}} + \chi_2\cdot\cdot\,\vec{\mathbf{E}}\vec{\mathbf{E}}, \tag{4b}$$

where V is the volume and χ_1 and χ_2 are the linear and second-order-nonlinear susceptibility tensors, respectively, of the system. Thus in this model the different orders of susceptibilities can be defined in intrinsic terms as sums of dyadic, triadic, etc. products of the bonds, with the influence of the

applied fields and the description of the observed radiation both represented as external operators. Extensions to higher orders are obvious.

For SHG we are not interested in the susceptibility tensors χ_1 and χ_2 directly, but instead in the far-field radiated intensity. Using the assumption that the bond charges radiate as dipoles, we can express the radiation observed in the far-field regime by the field $\vec{\mathbf{E}}_{ff}$, where

$$\vec{\mathbf{E}}_{ff}k^2=\frac{e^{ikr}}{r}\left[\sum_j \mathbf{p}_j-\hat{\mathbf{k}}\left(\hat{\mathbf{k}}\cdot\sum_j \mathbf{p}_j\right)\right] \tag{5a}$$

$$=k^2\frac{e^{ikr}}{r}[\mathbf{I}-\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}]\sum_j \mathbf{p}_j, \tag{5b}$$

where $\mathbf{I}$ is the unit tensor and $\mathbf{k}=k\hat{\mathbf{k}}$ is the unit vector in the direction of the observer. By Eqs. (4) and (5b), χ_2 and the observed SHG intensity are connected by the sum over bonds, which is common to both expressions. For this reason SHG can be described in terms of a bulk susceptibility even though in our case the sum over bonds is used for a different purpose.

B. Bulk nonlinear susceptibility of tetrahedral semiconductors

As an example we consider the bulk polarization response of the cubic group IV and III-V semiconductors having four tetrahedral bonds per unit cell. In the bulk these bonds are all equivalent except for direction, whence $\alpha_j=\alpha$ for all j. However, for the second-order calculation (but not the first) it is important that the directions be all defined consistently, for example, pointing from an atom of type *A* toward an atom of type *B*, since it follows from Eq. (4a) that for even susceptibilities a reversal of the direction of the jth bond is equivalent to a change of sign of α_j. Accordingly, placing an atom of type *A* at the origin of a standard face-centered-cubic coordinate system we have the four unit vectors

$$\hat{\mathbf{b}}_1=(\hat{\mathbf{x}}+\hat{\mathbf{y}}+\hat{\mathbf{z}})/\sqrt{3}, \tag{6a}$$

$$\hat{\mathbf{b}}_2=(\hat{\mathbf{x}}-\hat{\mathbf{y}}-\hat{\mathbf{z}})/\sqrt{3}, \tag{6b}$$

$$\hat{\mathbf{b}}_3=(-\hat{\mathbf{x}}+\hat{\mathbf{y}}-\hat{\mathbf{z}})/\sqrt{3}, \tag{6c}$$

$$\hat{\mathbf{b}}_4=(-\hat{\mathbf{x}}-\hat{\mathbf{y}}+\hat{\mathbf{z}})/\sqrt{3}. \tag{6d}$$

It follows immediately that $\alpha_1\Sigma_j\hat{\mathbf{b}}_j\hat{\mathbf{b}}_j=4\alpha_1\mathbf{I}/3$, and $\alpha_2\Sigma_j\hat{\mathbf{b}}_j\hat{\mathbf{b}}_j\hat{\mathbf{b}}_j=(8\sqrt{3}\alpha_2/3)\hat{\mathbf{x}}\hat{\mathbf{y}}\hat{\mathbf{z}}$, where α_1 and α_2 are the linear polarizability and first-order nonlinear hyperpolarizability, respectively, parallel to the bonds. The former result is perhaps the simplest proof that the tetragonally bonded cubic semiconductors are optically isotropic even when the microscopic polarizable species themselves are completely anisotropic. The latter result exhibits the tensor character expected for χ_2 from the bulk of these materials. For crystals with inversion symmetry $\alpha_2=0$, so SHG does not occur.

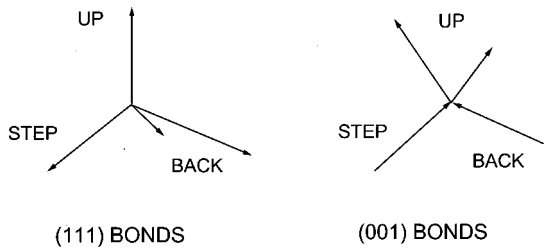

FIG. 1. Schematic of the bonds used in the calculations.

C. Interface response: Analytic solutions for SHG intensities for singular (111) and (001) Si interfaces

We consider next SHG signals from simple interfaces, specifically those SHG intensities that are expected to arise from interfaces formed on singular (111) and (001) surfaces. In the former case there are two classes of interface bonds: a single bond perpendicular to the interface plane, and three equivalent bonds between the interface-plane atoms and those in the plane below as shown on the left side of Fig. 1. Calculations are done most efficiently in a coordinate system where one of the bonds and the *z* axis are both normal to the interface and where one of the three equivalent bonds lies in the *xz* plane. In this case we assign to the vertical (up) bond a complex hyperpolarizability α_u and to the back (down) bonds a hyperpolarizability α_d. In fact for vicinal interfaces, one of the three down bonds will become inequivalent to the other two, including both step and terrace contributions. We recognize this inequivalency by defining in Fig. 1 and below a step bond to perform this function, although in deriving the analytic result we take it to be equivalent to the other two. While some differences in orientation can be expected with respect to the bulk-bond directions, for simplicity and on the basis of a likely result of statistical averaging we assume that the bond vectors at the interface are the same as those in the bulk. In this case the four interface bonds are described as the following.

Up bond:

$$\hat{\mathbf{b}}_1=\hat{\mathbf{z}}, \tag{7a}$$

step bond:

$$\hat{\mathbf{b}}_2=\frac{\sqrt{8}}{3}\hat{\mathbf{x}}-\frac{1}{3}\hat{\mathbf{z}}, \tag{7b}$$

back bonds:

$$\hat{\mathbf{b}}_3=-\frac{\sqrt{2}}{3}\hat{\mathbf{x}}+\frac{\sqrt{6}}{3}\hat{\mathbf{y}}-\frac{1}{3}\hat{\mathbf{z}}, \tag{7c}$$

$$\hat{\mathbf{b}}_4=-\frac{\sqrt{2}}{3}\hat{\mathbf{x}}-\frac{\sqrt{6}}{3}\hat{\mathbf{y}}-\frac{1}{3}\hat{\mathbf{z}}. \tag{7d}$$

Note that, as in the bulk of III-V materials, we are using the convention that all bonds, including the lower three, point away from the Si atom in the outermost plane.

The two normal modes, *s* and *p* polarization, of the incoming and outgoing beams lead to four observational com-

binations conventionally labeled *p-p*, *p-s*, *s-p*, and *s-s*, where the first and second letters refer to the polarizations of the incident and emerging beams, respectively. Assuming an incoming angle of incidence θ_i, the incoming *s*- and *p*-polarized beams are given by $\vec{\mathbf{E}}_s = E_s\hat{\mathbf{y}}$ and $\vec{\mathbf{E}}_p = E_p(-\hat{\mathbf{x}}\cos\theta_i + \hat{\mathbf{z}}\sin\theta_i)$, respectively. Assuming an observation angle θ_o, the *k* vector of the outgoing wave is given by $\hat{\mathbf{k}} = -\hat{\mathbf{x}}\sin\theta_o + \hat{\mathbf{z}}\cos\theta_o$. Using these values with the above expressions for the bond directions and Eqs. (4a) and (5b), we obtain the following relations for the radiated field in the far-field region in terms of unit input field.

For the *p-p* case:

$$\begin{aligned}\vec{\mathbf{E}}_{ff} = [\hat{\mathbf{x}}\cos\theta_o + \hat{\mathbf{z}}\sin\theta_o][\,&\alpha_u \sin^2\theta_i \sin\theta_o \\ &+ \alpha_d(\cos^3\beta \sin^2\theta_i \sin\theta_o \\ &+ \tfrac{3}{4}\sin\beta\sin 2\beta(\cos^2\theta_i\sin\theta_o - \sin 2\theta_i\cos\theta_o) \\ &+ \tfrac{3}{4}\sin^3\beta\cos^2\theta_i\cos\theta_o\cos 3\phi)],\end{aligned} \tag{8a}$$

for the *p-s* case:

$$\vec{\mathbf{E}}_{ff} = \hat{\mathbf{y}}\,\tfrac{3}{4}\,\alpha_d \sin^3\beta\cos^2\theta_i\sin 3\phi, \tag{8b}$$

for the *s-p* case:

$$\begin{aligned}\vec{\mathbf{E}}_{ff} = [\hat{\mathbf{x}}\cos\theta_o + \hat{\mathbf{z}}\sin\theta_o][\,&\tfrac{3}{4}\,\alpha_d(-\sin^3\beta\cos\theta_o\cos 3\phi \\ &+ \sin\beta\sin 2\beta\sin\theta_o)],\end{aligned} \tag{8c}$$

for the *s-s* case:

$$\vec{\mathbf{E}}_{ff} = -\hat{\mathbf{y}}\,\tfrac{3}{4}\,\alpha_d \sin^3\beta\sin 3\phi, \tag{8d}$$

where $\beta = 109.47°$ is the bond angle, and ϕ is measured from the *xz* plane. To calculate the intensity we take the absolute square of these fields.

For the singular, that is, single-domain (001) interface the equivalent bonds occur in pairs, as shown on the right side of Fig. 1, and thus the (001) result cannot be obtained simply as a rotation of the (111) result above. In principle this could be done if we assume that a second step bond is incorporated into the formalism, but we proceed by a different route. Considering first a single-domain interface we use a coordinate system where the upper and lower bonds lie in the *xz* and *yz* planes, respectively, and have complex hyperpolarizabilities α_u and α_l, respectively. Again, taking the bonds to be the same as those in the bulk the upper $\hat{\mathbf{b}}_{1,2}$ and lower $\hat{\mathbf{b}}_{3,4}$ bond vectors can be written

$$\hat{\mathbf{b}}_{1,2} = \pm\sqrt{\frac{2}{3}}\hat{\mathbf{x}} + \sqrt{\frac{1}{3}}\hat{\mathbf{z}}, \tag{9a,b}$$

$$\hat{\mathbf{b}}_{3,4} = \pm\sqrt{\frac{2}{3}}\hat{\mathbf{y}} + \sqrt{\frac{1}{3}}\hat{\mathbf{z}}. \tag{9c,d}$$

Here, we choose directions that make the *z*-axis projections of all bonds positive (all bonds pointing toward the interface) because we will also use the resulting equations to describe macroscopically double-domain (001) interfaces. Performing the same calculation as before yields, in simplest mathematical form, the following expressions for the far fields.

For the *p-p* case:

$$\begin{aligned}\vec{\mathbf{E}}_{ff} = [\hat{\mathbf{x}}\cos\theta_o + \hat{\mathbf{z}}\sin\theta_o][\,&(\alpha_u + \alpha_l)(2\cos^3\beta\sin^2\theta_i\sin\theta_o \\ &- \tfrac{1}{2}\sin\beta\sin 2\beta(\sin 2\theta_i\cos\theta_o - \cos^2\theta_i\sin\theta_o)) \\ &- \tfrac{1}{2}(\alpha_u - \alpha_l)\sin\beta\sin 2\beta(\sin 2\theta_i\cos\theta_o \\ &- \cos^2\theta_i\sin\theta_o)\cos 2\phi],\end{aligned} \tag{10a}$$

for the *p-s* case:

$$\vec{\mathbf{E}}_{ff} = -\hat{\mathbf{y}}\frac{1}{2}(\alpha_u - \alpha_l)\sin\beta\sin 2\beta\sin 2\theta_i\sin 2\phi, \tag{10b}$$

for the *s-p* case:

$$\begin{aligned}\vec{\mathbf{E}}_{ff} = [\hat{\mathbf{x}}\cos\theta_o + \hat{\mathbf{z}}\sin\theta_o]\,&\tfrac{1}{2}\sin\beta\sin 2\beta\sin\theta_o[(\alpha_u + \alpha_l) \\ &- (\alpha_u - \alpha_l)\cos 2\phi],\end{aligned} \tag{10c}$$

and for the *s-s* case:

$$\vec{\mathbf{E}}_{ff} = 0, \tag{10d}$$

where α_u and α_l are the hyperpolarizabilities of the upper and lower bonds, respectively. The intensity is calculated by taking the absolute square.

Real on-axis (001) Si interfaces are not singular but consist of statistically equal areas of two inequivalent domains that are rotated by 90° with respect to each other. A macroscopic beam averages over many such domains, thereby sampling statistically equal areas of both. The effect of such a macroscopic average in Eqs. (10) is to replace both α_u and α_l with their average value $(\alpha_u + \alpha_l)/2$. The result is that the azimuthal dependences vanish, consistent with group-theoretical arguments. The effective equivalency of all four bonds in this case is the reason why we chose this direction convention for the (001) interface. For vicinal interfaces the situation is more complicated, since the lower bonds associated with the step atoms become inequivalent, as also indicated in Fig. 1.

In least-squares comparisons of the predictions of the SBHM to data on vicinal samples, we evaluate the SBHM numerically, tilting the bonds appropriately and assuming for both nominal (111) and (001) orientations the most general case of independent complex hyperpolarizabilities for all four bonds. This provides a cross check on the data and calculations, since the back and up bonds are expected to be equivalent for vicinal samples in the two cases.

D. Comparison with experiment: Vicinal (111) Si

We assess the validity of the model by applying it to the data reported in Ref. 8 by Lüpke, Bottomley, and van Driel on an oxidized vicinal (111) Si wafer. These data are appropriate because all four combinations of polarizations are reported. Also, the data are given as a table of Fourier coefficients of the field (Table 6 of Ref. 8), thereby allowing these data to be reconstructed numerically as either fields or as the as-measured quantities, the intensities. These data are normalized to the amplitude of the 3ϕ Fourier coefficients of the fields, so the intrinsic relative-intensity information among the different polarization configurations is lost and compari-

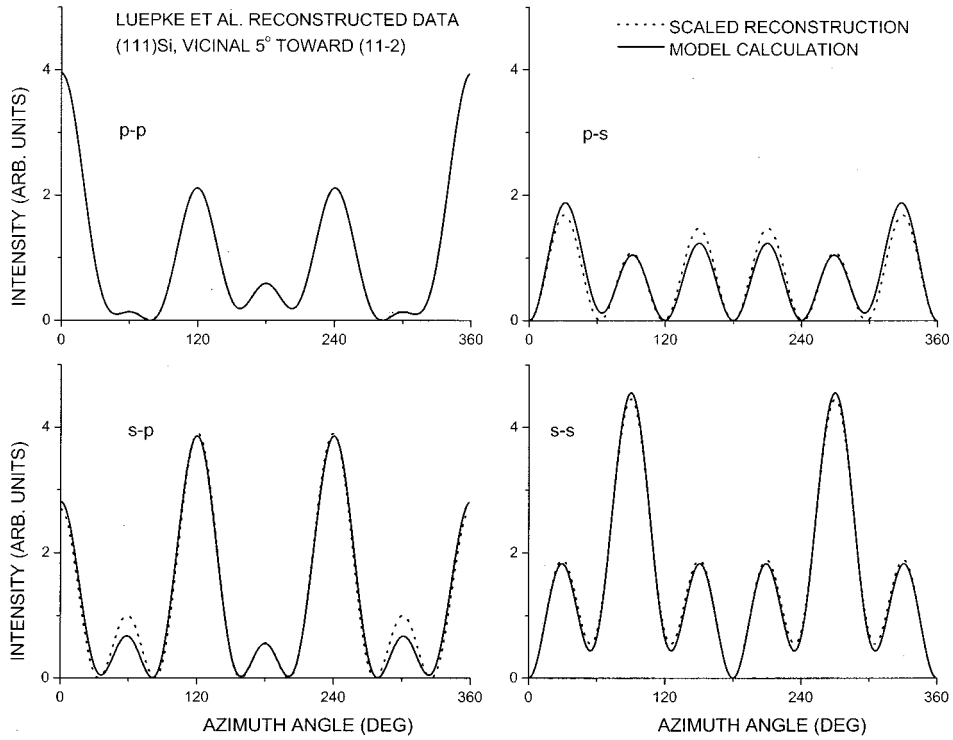

FIG. 2. Comparisons between the SHG data of Lüpke *et al.* (points, Ref. 8) and azimuthal dependences calculated with the SBHM (lines) for a 5° vicinal (111) Si-SiO_2 interface. Upper left: Result of least-squares fitting the *p-p* data. Remaining panels: Dependences predicted for *p-s*, *s-p*, and *s-s* using parameters determined for *p-p*.

sons must be done on the basis of lineshapes alone. Experimental conditions were as follows: wavelength $\lambda = 765$ nm, nominal miscut angle$=5°$ toward the $[11\bar{2}]$ direction, nominal angle of incidence$=45°$, and nominal angle of observation$=45°$. Using their Fourier coefficients we have reconstructed their data as intensities, normalizing them equivalently in our case to the 6ϕ coefficients.

We first determine the (complex) bond hyperpolarizabilities by least-squares fitting the azimuthal variation of the SHG intensity predicted by the model to the *p-p* data. This will also determine whether the SBHM adequately spans the necessary mathematical space, i.e., whether it implicitly contains the Fourier coefficients needed to represent the data as given in Ref. 8. This will provide the first critical test of the model. The second critical test follows by examining the resulting parameters to see whether they are physically reasonable. The third critical test will be to examine whether the parameters determined for the *p-p* configuration can predict the azimuthal dependences observed for the SHG intensity for the *p-s*, *s-p*, and *s-s* configurations.

However, it is first necessary to decide which angles of incidence and observation should be used in Eqs. (4a) and (5b). One possibility is the 45° angle used in the laboratory. The others are the values determined by refraction at the outer surface of the SiO_2 overlayer as calculated by Snell's law, noting that the refractive indices of SiO_2 at $\lambda = 765$ and 382.5 nm are slightly different. Specifically, these indices are $n_{0x} = 1.454$ and 1.472, respectively. In the second case, given the 45° laboratory angles, we would expect the interface angles to be 29.1° and 28.7°, respectively. In practice this difference is too small to be relevant, and if this is the only factor we can assume the angles of incidence and observation to be equal. We make this assumption, but check it by using the value of the incidence and observation angle as a least-squares fitting parameter along with the vicinal angle and the complex polarizabilities of the four bonds.

The results of the fitting procedure are shown in the upper left panel of Fig. 2. The fit is essentially exact, with a residual of $\delta = 0.0039$. Thus the SBHM not only has the necessary mathematical breadth to represent the data, passing the first test, but also does it in a mathematically more efficient manner: excluding normalization, the SBHM representation requires only three hyperpolarizability parameters, as opposed to the four Fourier coefficients listed in Ref. 8. The parameters obtained are 29.5° for the angles of incidence and observation, 1.12° for the vicinal angle, and $1.77 + i0.17$, $2.76 + i1.43$, $2.19 - i0.00$, and 2.19 for the hyperpolarizabilities of the up, step, and two nominally equivalent back bonds, respectively. The imaginary part of the hyperpolarizability of one of the back bonds is arbitrarily set equal to zero, since the absolute phase cannot be determined from an intensity. The low value of the residual essentially makes the uncertainties in the determined parameters meaningless, but we list them anyway: 0.04° for the vicinal angle, 0.01° for the angles of incidence and observation, 0.02 for the hyperpolarizability of the up bond, and about 0.005 for the hyperpolarizabilities of the remaining bonds. We note that these hyperpolarizabilities scale with the data, hence the use of arbitrary units for the SHG intensity means that the obtained hyperpolarizabilities have relative value only.

The hyperpolarizabilities appear to be reasonable. In particular the hyperpolarizabilities determined for the two back bonds are essentially identical, as expected. It is interesting to note that the step bond exhibits the greatest phase difference, which is perhaps not surprising since the (effective) step bond is expected to exhibit the greatest difference with respect to bulk properties. However, absorption is expected to be a strong function of energy, and the absorption that occurs at 765 nm may not be representative of that which occurs at other energies. When theoretical calculations become available the prediction of a large absorption for the step bond, small absorption for the up bond, and essentially no absorption for the back bonds will provide another critical test of the model. The angles of incidence and observation are consistent with refraction at the Si-SiO_2 interface, but this agreement should not be taken too seriously at this point since we find that the *p-p* fitting results are rather insensitive to this parameter. The only significant discrepancy between expected and obtained parameters occurs with the vicinal angle, for which we have no immediate explanation.

We now assess whether the hyperpolarizabilities determined for the *p-p* configuration predict the results observed for the remaining three configurations. These are also shown in Fig. 2. As can be seen, the SBHM provides a very good representation of the reconstructed intensities if they are scaled by the least-squares-determined values 1.40, 1.93, and 2.26, respectively. The need for, and physical meaning of, these scaling factors can be understood with reference to Eqs. (8). Although Eqs. (8) were derived on the basis of a singular sample and the data are for a nominally 5° vicinal sample, we expect that the main effect of a miscut will be to generate additional harmonics, and that the primary (3ϕ and 6ϕ) components of the singular interface will be affected only to second order.

Equations (8) show that the amplitudes of the 6ϕ coefficients of all four configurations share the common factor $[3\alpha_d \sin^3\beta/4]$ and differ only in that the *p-p*, *p-s*, *s-p*, and *s-s* coefficients are multiplied in addition by $\cos^2\theta_i \cos\theta_0$, $\cos^2\theta_i$, $\cos\theta_0$, and 1, respectively. Thus if we again make the reasonable assumption that the angle of incidence equals the angle of observation the scaling coefficients given in the previous paragraph should equal 1, $1/\cos^2\theta$, $1/\cos^4\theta$, and $1/\cos^6\theta$, respectively, relative to their values for the *p-p* data.

The best-fit scaling parameters yield θ values of 32.5°, 32.0°, and 29.2° for the *p-s*, *s-p*, and *s-s* configurations, respectively. These values are not only mutually consistent but are also in excellent agreement with the values of 29.1° and 28.7° calculated for refraction at the air-SiO_2 interface using the refractive indices of 1.454 and 1.472 for SiO_2 at 765 and 382.5 nm, respectively. From this we can also conclude that for this oxidized (111)Si interface the SHG signal originates mainly from the SiO_2 side. Since we can account for the scaling parameters, the SBHM reproduces the *p-s*, *s-p*, and *s-s* anisotropies from the same four parameters determined for the *p-p* configuration, in contrast to the seven additional Fourier coefficients that are needed to describe the associated fields in Ref. 8.

The absolute square of Eq. (8c) shows that a fifth test is also possible from the relative amplitudes of the 3ϕ and 6ϕ coefficients of the azimuthal dependence for the *s-p* configuration. Assuming that the coefficients *b* and *c* of the 3ϕ and 6ϕ terms have the same phase, it is possible to show that $(\frac{1}{2})\tan\beta\cot\theta_o = c/b = 1.00/0.354$. Using the bond angle $\beta = 109.47°$ we find $\theta_o = 26.6°$, again in good agreement with the above. Thus taking advantage of the capability of the SBHM to lead to an analytic solution, we obtain direct information not only about the bond hyperpolarizabilities themselves but also about the basic physical origin of interface SHG.

E. Comparison with experiment: (001) Si

A thorough analysis of SHG signals obtained in the *s-p* configuration for singular and vicinal (001) Si samples, including those for a clean single-domain surface, has been given by Lüpke *et al.*[7] These data clearly show the fourfold rotationally symmetric bulk quadrupole contribution for the singular surface and various lower-order contributions for vicinal surfaces. As mentioned above, owing to the areal fractions of the two domains of a singular (001) Si sample being statistically equal, the lowest-order NLO contribution on a macroscopic scale is of the form $\cos(4\phi)$, which requires a fourth-rank tensor for its description and can therefore only be observed for nonlinearities at and above THG. Thus no interface contribution is expected in SHG for singular Si-SiO_2 samples. However, SHG signals can be obtained on vicinal (001) *c*-Si samples miscut by at least several degrees toward one of the nearest (111) axes.[7] Since Lüpke *et al.* report SHG anisotropy data for only the *s-p* configuration and since we believe that a critical test of the SBHM requires such data for more than one configuration, we acquired SHG anisotropy data on vicinal samples oxidized as described above for both *p-p* and *p-s* configurations. The data and fitting results are illustrated in Fig. 3. The angles of incidence and observation were 45° and the wavelength was 830 nm. The *p-s* data were multiplied by a factor of 4 to compensate for the different pulse durations and intensities that resulted as a consequence from the need to use different experimental conditions for the *p-p* and *p-s* data.

It is clear that, as with the (111) interface, the SBHM gives a good representation of the data. Here, the values of the least-squares fitting parameters are 12° for the angles of incidence and observation, 7.28° for the vicinal angle, and $105+i30$, $197+i55$, $133+i0$, and 131 for the hyperpolarizabilities of the four bonds, again in arbitrary units, where the phase of the last bond was arbitrarily set equal to zero. The values here are good to within ± 4. The step bond consistently shows the highest polarizability and absorption. The *p-s* anisotropy is found to be particularly sensitive to the angles of incidence and observation, and we estimate that the 12° value is determined here to within 3°.

These results are interesting for several reasons. First, as with the (111) interface, the SBHM is found to be capable of providing a good mathematical representation of the data. Second, the vicinal angle is recovered to within 2°. Third, as with the (111) interface two bonds are expected and found within experimental uncertainty to be equivalent. Fourth, the angles of incidence and observation are consistent with

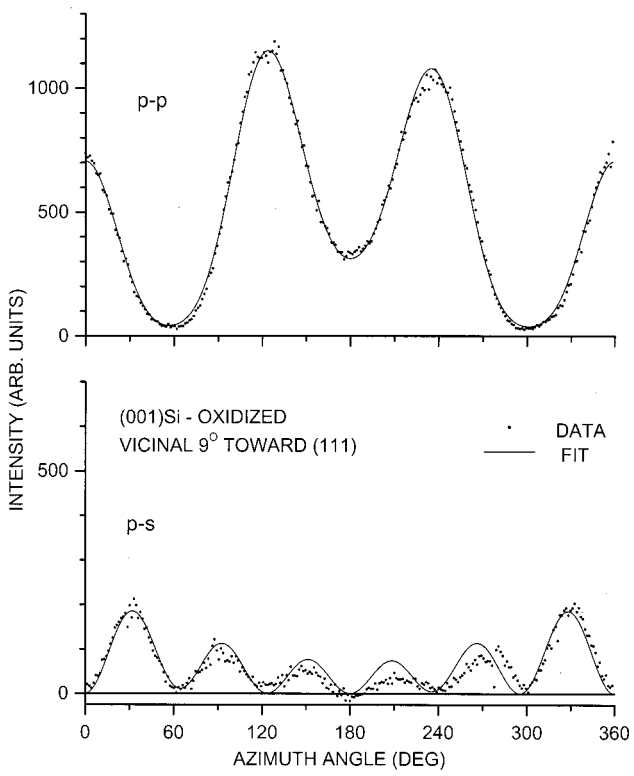

FIG. 3. Result of least-squares fitting for a 9° vicinal Si-SiO_2 interface.

Snell's law with a refractive index of 3.4, which indicates that the SHG signal here is originating predominantly from the Si side of the interface. This difference with respect to the (111) data is probably not unreasonable since the outermost Si atoms on the (001) terrace are bonded to at least two O atoms, as opposed to the (111) terrace atoms where only one O atom is involved.

This argument clearly needs refinement, since the assignment of the observed anisotropy to terraces is obviously an oversimplification. In fact on the basis of macroscopic symmetry it may be more reasonable to assign the signal entirely to steps. Nevertheless our linear-anisotropy (reflectance-difference/anisotropy) data on a series of annealed vicinal (001)Si-SiO_2 interfaces show a step phase transition that is similar to that observed for clean surfaces in ultrahigh vacuum.[30] This is probably driven by a similar mechanism although, given the strength of the Si-O bond, it must be one that maximizes the density of Si-O bonds rather than minimizes the density of dangling bonds. Although various authors now appear to agree on the (terrace) reconstruction that takes place at the singular (001) Si-SiO_2 interface,[31,32] to our knowledge no equivalent theoretical calculations are available for steps. If macroscopic terrace symmetry were overridden one would expect contributions from both terraces and steps. We are attempting to address the relative step/terrace contribution as well as other issues experimentally, and will discuss this elsewhere.

III. DISCUSSION

Although with the exception of Ref. 26, radiating-dipole models appear to not have been used previously in the context of NLO, the concept of observed intensities originating as radiation from dipoles driven by incident waves goes back in linear optics at least as far as the work of Ewald and Oseen, in 1912 and 1915, respectively.[18,19] The (largely forgotten) Ewald-Oseen extinction theorem describes specular reflection as the result of the coherent superposition of radiation from dipoles in the medium driven by the incident wave, which is assumed to penetrate the substrate uniformly (no absorption). The dipoles in the bulk generate a radiated wave there that completely cancels the incident wave (thus the term ''extinction''), and a polarization wave that obeys Snell's law. The reflected beam itself therefore originates entirely from the outermost layer of dipoles. An immediate consequence is a direct physical interpretation of the vanishing of the reflectance for *p*-polarized light incident at Brewster's angle: the observer is simply viewing the radiating interface dipoles head-on.

The only recent treatments that are related are those of Wijers *et al.*[26] and that of Sipe.[33] The latter presented a formalism for representing SHG intensities in terms of a polarization layer together with the use of standard boundary conditions for relating the electric fields of the incident and transmitted waves. While this approach leads to a mathematical connection between the polarization of the layer and the emitted wave, it does not connect the observed SHG intensity to the microscopic properties of the layer.

Using a more general representation that included additional bonds, allowed different hyperpolarizabilities for each bond, and allowed hyperpolarizability contributions both along and orthogonal to a bond, Patterson *et al.* performed a cluster calculation to estimate both linear and nonlinear restoring forces and therefore the second-order nonlinear susceptibility of As- and Ga-terminated Si surfaces.[27] The results indicated that the orthogonal contributions at the level of the individual bonds could be important in SHG, although in our case such contributions would be included in the effective hyperpolarizabilities of the four tetrahedral bonds. The results obtained by Patterson *et al.* were in reasonable agreement with experiment, supporting the general treatment of SHG in terms of anharmonically polarizable bonds.

We comment finally on the evidence for SHG absorption in anisotropy data. It is easy to show in the SBHM by arbitrarily setting all imaginary parts of the hyperpolarizabilities to zero that SHG absorption can be recognized by the failure of certain features in the azimuthal dependences of the *s-p*, *p-s*, and *s-s* configurations to reach zero or nearly zero for intermediate azimuth angles. Examples include the local minima at 50°, 155°, 205°, and 310° in the *p-p* data of Fig. 2, and those at 60°, 120°, 240°, and 300° in the *s-p* data of Fig. 3. These features should be valuable for determining SHG absorption processes at Si-dielectric interfaces. Also, the relative imaginary values of the hyperpolarizabilities of the vicinal (001)Si-SiO_2 interface at 830 nm are found to be substantially smaller than those for the (111)Si-SiO_2 interface at 765 nm. Whether this is due to different reconstructions or the fact that these data were obtained at different wavelengths highlights the diagnostic possibilities that will exist when it becomes possible to obtain SHG data over extended spectral ranges.

IV. CONCLUSIONS

We have shown that a simplified bond-hyperpolarizability model (SBHM) provides a relatively straightforward and apparently fairly complete way of representing SHG anisotropy data in terms of parameters that have direct physical significance on the microscopic scale, in contrast to the tensor or Fourier coefficients of previous phenomenological approaches. Even in the absence of first-principles theoretical estimates of the relevant bond hyperpolarizabilities, the SBHM allows us to extract some of the previously inaccessible physics of SHG at interfaces, specifically providing direct evidence of SHG absorption and allowing us to determine the effective angles of incidence and observation. These in turn indicate the part of the interface from which the SHG signals originate.

In our present treatment we have analyzed data that have been obtained at only two wavelengths. Extension to SHG spectra would be extremely interesting, not only as a means of obtaining greater insight concerning the energy dependence and therefore the origin of the observed SHG absorption, but also for additionally critically assessing the SBHM model. In particular, an assessment of the effective angles of incidence and observation as a function of wavelength would provide additional insight into the specific interface region responsible for the SHG signal.

Finally, the assumption that the only relevant charge motion occurs along the bond axis for covalent materials cannot be made for any odd-order susceptibility, including linear optics and third-order susceptibility as well as higher-order effects such as bulk quadrupole generation in centrosymmetric materials. However, we showed in Sec. II that maximally anisotropic bonds still lead within this formulation to expected macroscopic symmetries for the bulk responses of tetrahedrally bonded semiconductors, isotropic for linear optics and of the form χ_{ijk} for SHG. Our focus in the present work is on SHG. Whether this consistency is retained for higher orders of nonlinearity remains to be established, and will be the topic of future investigations.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the Office of Naval Research for continuing support of this work and for providing the funding by which our SHG equipment was purchased. We would also like to acknowledge further support by the Army Research Office and the Max-Planck-Gesellschaft for an Award for International Cooperation. Discussions with G. Lucovsky are also greatly appreciated.

*Corresponding author. Email: aspnes@unity.ncsu.edu; Fax: US 919-515-1333.

[1] J. F. McGilp, Prog. Surf. Sci. **49**, 1 (1995).
[2] G. Lüpke, Surf. Sci. Rep. **35**, 75 (1999).
[3] Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984).
[4] W. Daum, H.-G. Krause, U. Reichel, and H. Ibach, Phys. Rev. Lett. **71**, 1234 (1993).
[5] C. H. Bjorkman, C. E. Shearon, Jr., Y. Ma, T. Yasuda, G. Lucovsky, U. Emmerichs, C. Meyer, K. Leo, and H. Kurz, J. Vac. Sci. Technol. A **11**, 964 (1993).
[6] C. H. Bjorkman, T. Yasuda, C. E. Shearon, Jr., Y. Ma, G. Lucovsky, U. Emmerichs, C. Meyer, K. Leo, and H. Kurz, J. Vac. Sci. Technol. B **11**, 1521 (1993).
[7] G. Lüpke, D. J. Bottomley, and H. M. van Driel, Phys. Rev. B **47**, 10 389 (1993).
[8] G. Lüpke, D. J. Bottomley, and H. M. van Driel, J. Opt. Soc. Am. B **11**, 33 (1994).
[9] J. I. Dadap, X. F. Hu, M. H. Anderson, M. C. Downer, J. K. Lowell, and O. A. Aksipetrov, Phys. Rev. B **53**, R7607 (1996).
[10] C. Meyer, G. Lüpke, Z. G. Lũ, A. Götz, H. Kurz, and G. Lucovsky, J. Vac. Sci. Technol. B **14**, 3107 (1996).
[11] S. T. Cundiff, W. H. Knox, F. H. Baumann, K. W. Evans-Lutterodt, and M. L. Green, J. Opt. Soc. Am. A **16**, 1730 (1998).
[12] G. Erley and W. Daum, Phys. Rev. B **58**, R1734 (1998).
[13] See, e.g., M. Cini, Phys. Rev. B **43**, 4792 (1991).
[14] T. F. Heinz, M. M. T. Loy, and W. A. Thompson, J. Vac. Sci. Technol. B **3**, 1467 (1985).
[15] J. E. Sipe, D. J. Moss, and H. M. van Driel, Phys. Rev. B **35**, 1129 (1987).
[16] V. Mizrahi and J. E. Sipe, J. Opt. Soc. Am. B **5**, 660 (1988).
[17] J. E. Sipe and A. I. Shkrebtii, Phys. Rev. B **61**, 5337 (2000).
[18] P. P. Ewald, Ph.D. thesis, Munich, 1912; Ann. Phys. (Leipzig) **49**, 1 (1916).
[19] C. W. Oseen, Ann. Phys. (Leipzig) **48**, 1 (1915).
[20] The most convenient discussion of the Ewald-Oseen extinction theorem is probably found in M. Born and E. Wolf, *Principles of Optics*, 5th ed. (Pergamon, Oxford, 1975), pp. 100 ff.
[21] N. Bloembergen and P. S. Pershan, Phys. Rev. **128**, 606 (1962).
[22] N. Bloembergen, R. K. Chang, S. S. Jha, and C. H. Lee, Phys. Rev. **174**, 813 (1968).
[23] B. F. Levine, Phys. Rev. B **7**, 2600 (1973).
[24] A. Liebsch and W. L. Schaich, Phys. Rev. B **40**, 5401 (1989).
[25] W. L. Schaich and B. S. Mendoza, Phys. Rev. B **45**, 14 279 (1992).
[26] C. M. J. Wijers, P. L. de Boeij, C. W. van Hasselt, and Th. Rasing, Solid State Commun. **93**, 17 (1995).
[27] C. H. Patterson, D. Weaire, and J. F. McGilp, J. Phys.: Condens. Matter **4**, 4017 (1992).
[28] B. S. Mendoza and W. L. Mochán, Phys. Rev. B **55**, 2489 (1997).
[29] N. Arzate and B. S. Mendoza, Phys. Rev. B **63**, 113303 (2001).
[30] See, e.g., O. L. Alerhand, A. N. Berker, J. D. Joannopoulos, D. Vanderbilt, R. J. Hamers, and J. E. Demuth, Phys. Rev. Lett. **64**, 2406 (1990).
[31] R. Buczko, S. J. Pennycook, and S. T. Pantelides, Phys. Rev. Lett. **84**, 943 (2000).
[32] J. B. Neaton, D. A. Muller, and N. W. Ashcroft, Phys. Rev. Lett. **85**, 1298 (2000).
[33] J. E. Sipe, J. Opt. Soc. Am. **4**, 481 (1987).