Electronic structure of semiconductor quantum films

S. B. Zhang, Chin-Yu Yeh, and Alex Zunger
National Renewable Energy Laboratory, Golden, Colorado 80401
(Received 6 May 1993)

The electronic structure of thin ($\leq$30 Å) free-standing ideal films of Si(001), Si(110), and GaAs(110) is calculated using a plane-wave pseudopotential description. Unlike the expectation based on the simple effective-mass model, we find the following. (i) The band gaps of (001) quantum films exhibit even-odd oscillation as a function of the number N of monolayers. (ii) In addition to sine-type envelope functions which vanish at the film boundaries, some states have cosine envelope functions with extrema at boundaries. (iii) Even-layer Si(001) films exhibit at the valence-band maximum a state whose energy does not vary with the film thickness. Such zero confinement states have constant envelope throughout the film. (iv) Optical transitions in films exhibit boundary-imposed selection rules. Furthermore, oscillator strengths for pseudodirect transitions in the vicinity of *forbidden* direct transitions can be enhanced by several orders of magnitude. These findings, obtained in direct supercell calculations, can be explained in terms of a truncated crystal (TC) analysis. In this approach the film's wave functions are expanded in terms of pairs of bulk wave functions exhibiting a destructive interference at the boundaries. This maps the eigenvalue spectra of a film onto the bulk band structure evaluated at special **k** points which satisfy the boundary conditions. We find that the TC representation reproduces accurately the above-mentioned results of direct diagonalization of the film's Hamiltonian. This provides a simple alternative to the effective-mass model and relates the properties of quantum structures to those of the bulk material.

I. INTRODUCTION

Quantum semiconductor structures[1,2] are generally discussed in terms of three classes — structures periodic in zero dimension (0D quantum boxes), those periodic in one dimension (1D quantum wires), and structures that are periodic in two dimensions (2D quantum wells). Relative to the 3D periodic bulk solids, these structures exhibit localization effects in three, two, and one dimension, respectively. There is, however, another type of 2D periodic structure which is discussed less frequently, namely a free-standing quantum film. It can be characterized by its thickness L and layer orientation $\hat{\epsilon}$. Unlike the conventional 2D periodic quantum structure, the quantum well, a film can exhibit a far stronger confining potential since all conduction and valence states that lie below the vacuum level experience a common potential well whose depth (>1 rydberg) equals the valence-band width plus the work function [Fig. 1(a)]. In contrast, states in a quantum well are confined just by the potential discontinuity for a particular band state which is generally of the order of 0.1–1 eV [Fig. 1(b)]. Quantum film can hence be thought of as the limiting case of a semiconductor/insulator quantum well with a giant band offset. This paper deals with the electronic structure of semiconductor quantum films. To articulate the unexpected features of such systems relative to the standard model of quantum confinement, i.e., effective-mass particle-in-a-box,[3,4] we first summarize the basic assumptions and predictions of the standard model, as applied to films.

Consider a (001) = z oriented quantum film experiencing an infinite potential well outside $z = 0$ and $z = L$ and the quasiperiodic potential $V_{\text{periodic}}(\mathbf{r})$ (characterizing the film material) inside the interval $(0, L)$ [Fig. 2(a)]. Neglecting surface effects, the effective-mass particle-in-a-box description of such a system (= particle-in-an-empty-film) proceeds with the following steps.

(i) Renormalize away the quasiperiodic microscopic potential $V_{\text{periodic}}(\mathbf{r})$, replacing it by a constant potential in the $(0, L)$ interval [Fig. 2(b)]. In replacing the "particle-in-a-film" problem [Fig. 2(a)] by the "particle-in-an-empty-film" problem [Fig. 2(b)], the true bulk band

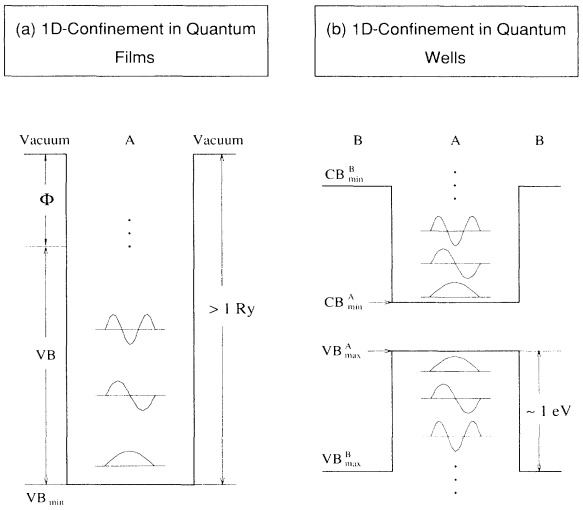

FIG. 1. Schematic depiction of quantum confinement in (a) films and (b) quantum wells. The subscripts A and B denote the well and barrier materials, respectively. Φ is the work function.

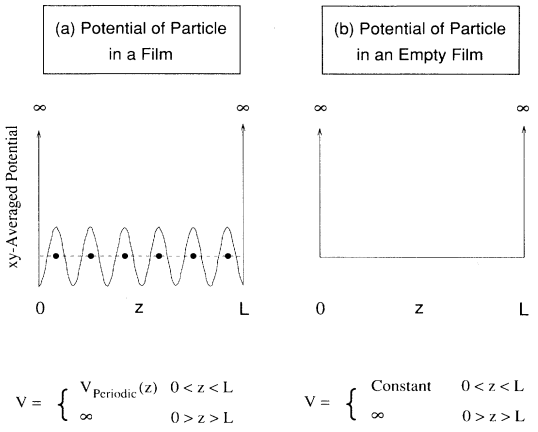

FIG. 2. Schematic depiction of the xy-planar averaged potentials for (a) a quantum film and (b) a particle in an empty film.

structure $\epsilon^{\text{bulk}}_{n,\mathbf{k}}$ of the film material

$$\epsilon^{\text{bulk}}_{n,\mathbf{k}} = \langle n\mathbf{k}| -\tfrac{1}{2}\nabla^2 + V_{\text{periodic}}(\mathbf{r})|n\mathbf{k}\rangle \tag{1a}$$

is replaced by the pure kinetic energy form

$$\epsilon^{\text{EMA}}_{n,\mathbf{k}} = \epsilon_{n,\mathbf{k_0}} + \frac{\hbar^2(\mathbf{k}-\mathbf{k_0})^2}{2m^*_{n,\mathbf{k_0}}} \quad , \tag{1b}$$

where $\epsilon_{n,\mathbf{k_0}}$ and $m^*_{n,\mathbf{k_0}}$ are the band edge energy and the effective mass of the nth band at $\mathbf{k_0}$, respectively. This approximation removes (a) the coupling between the band structure effects of the quasiperiodic potential and the confinement effects of the external potential (b) the microscopic (layer dependent) symmetry of the films.

(ii) The true bulk crystal eigenstates $\psi^{\text{bulk}}_{n,\mathbf{k}}(\mathbf{r})$ which satisfy

$$(-\tfrac{1}{2}\nabla^2 + V_{\text{periodic}})\psi^{\text{bulk}}_{n,\mathbf{k}}(\mathbf{r}) = \epsilon^{\text{bulk}}_{n,\mathbf{k}}\,\psi^{\text{bulk}}_{n,\mathbf{k}}(\mathbf{r}) \tag{2a}$$

are represented by a product of an envelope function $f_{\mathbf{k}-\mathbf{k_0}}(\mathbf{r})$ and the periodic piece $u_{n,\mathbf{k_0}}(\mathbf{r})$ of $\psi^{\text{bulk}}_{n,\mathbf{k_0}}(\mathbf{r})$,

$$\psi^{\text{EMA}}_{n,\mathbf{k}}(\mathbf{r}) = f_{\mathbf{k}-\mathbf{k_0}}(\mathbf{r})u_{n,\mathbf{k_0}}(\mathbf{r}) \quad . \tag{2b}$$

A generalization of Eq. (2b) is to use a few bands at $\mathbf{k_0}$ (Luttinger's model[5])

$$\psi^{\text{EMA}}_{m,\mathbf{k}}(\mathbf{r}) = \sum_n f_{n,m,\mathbf{k}-\mathbf{k_0}}(\mathbf{r})\; u_{n,\mathbf{k_0}}(\mathbf{r}) \quad . \tag{2c}$$

Consistent with the decoupling in (i), the external potential is then permitted to modify the envelope function $f_{\mathbf{k}-\mathbf{k_0}}(\mathbf{r})$ and its energy $\hbar^2(\mathbf{k}-\mathbf{k_0})^2/2m^*_{n,\mathbf{k_0}}$, but not $u_{n,\mathbf{k_0}}(\mathbf{r})$ and $\epsilon_{n,\mathbf{k_0}}$.

(iii) The film boundary problem then becomes a particle-in-an-empty-film problem. In the single band approximation the film's wave function is represented by a standing wave created by a destructive interference between two running waves of opposite directions, $(\mathbf{k}-\mathbf{k_0})$ and $[k_x - k_{x0}, k_y - k_{y0}, -(k_z - k_{z0})]$,

$$\begin{aligned}\psi^{\text{EMA}}_{n,\mathbf{k}}(\mathbf{r}) &= u_{n,\mathbf{k_0}}(\mathbf{r})[f_{\mathbf{k}-\mathbf{k_0}}(\mathbf{r}) - f_{[k_x-k_{x0},k_y-k_{y0},-(k_z-k_{z0})]}(\mathbf{r})]/2i \\ &= u_{n,\mathbf{k_0}}(\mathbf{r})e^{i[(k_x-k_{x0})x+(k_y-k_{y0})y]}\sin(k_z-k_{z0})z \quad .\end{aligned} \tag{3}$$

Intervalley (e.g., Γ and X) coupling is neglected. The boundary conditions on the film eigenstate $\psi^{\text{EMA}}(z=0) = \psi^{\text{EMA}}(z=L) = 0$ then lead to the quantization condition

$$k_z - k_{z0} = k^*_z = \frac{\pi}{L}j; \quad j = 1,2,3,\ldots \quad . \tag{4}$$

Note that the lowest index j is one, not zero. The more general quantization conditions for the particle-in-an-empty-film problem at $\mathbf{k_0} = 0$ are

$$\mathbf{k} = (k^*_x, k^*_y, k^*_z) + \mathbf{G} \tag{5a}$$

with

$$k^*_z = \frac{\pi}{L}j, \tag{5b}$$

where

$$j = \begin{cases} 1,2,3,\ldots,j_L & \text{if } \mathbf{G} = \mathbf{0} \\ 0,1,2,\ldots,j_L & \text{if } \mathbf{G} \neq \mathbf{0} \end{cases} \quad . \tag{5c}$$

Here j_L is the maximum value of j (determined by the thickness L), k^*_x and k^*_y are restricted to the first Brillouin zone of the bulk material, and $\mathbf{G}$ are the reciprocal lattice vectors (in what follows, we will consider the case $k^*_x = k^*_y = 0$, i.e., states at $\bar{\Gamma}$ of the 2D Brillouin zone). Note that in Eq. (5c) the solution $j = 0$ is allowed for $\mathbf{G} \neq \mathbf{0}$, unlike Eq. (4). The more general quantization conditions (5) are, however, deemed inappropriate in the EMA since $(\mathbf{k}-\mathbf{k_0})$ needs to be small for Eqs. (2b) and (2c) to be valid, and $\mathbf{G} \neq \mathbf{0}$ violates this restriction.

(iv) Substituting the quantization condition [Eq. (4)] into the bulk eigenvalue spectrum [Eq. (1b)] yields $\epsilon^{\text{EMA}} \propto 1/L^2$ and

$$\epsilon^{\text{EMA}}_g = E_g + \frac{1}{L^2}\left(\frac{1}{m^*_e} + \frac{1}{m^*_h}\right)\frac{\hbar^2\pi^2}{2} \quad , \tag{6}$$

where E_g is the bulk band gap, and e and h stand for electron and hole.

The EMA particle-in-a-box description of a quantum film makes a number of clear predictions.

(a) The film band gap decreases *monotonically* with the film thickness, as $1/L^2$ [Eq. (6)]. Figure 3(a) gives a schematic graphical interpretation of the way the film band gap can be read off the parabolic EMA dispersion relation of Eq. (1b): as the film becomes thicker, the band edge states approach the bulk Γ point and $\epsilon^{\text{EMA}}_g \to E_g$. We see, therefore, that this monotonic behavior is not a

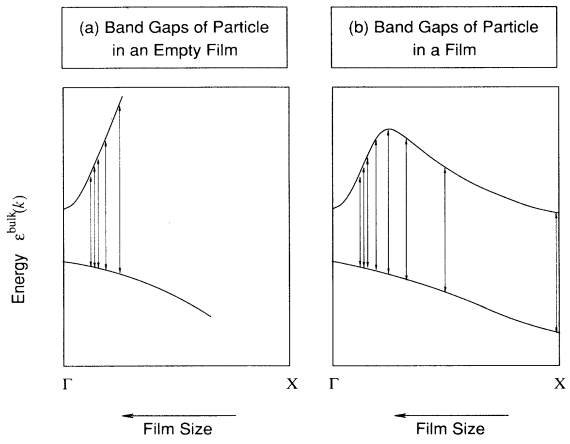

FIG. 3. Schematic drawings showing the relation between the bulk energy dispersion $\epsilon^{\text{bulk}}(\mathbf{k})$ and the direct band gap ϵ^{film} of the corresponding films of different sizes L for (a) a particle in an empty film with parabolic bands and (b) a real semiconductor film where the energy dispersion for large k deviates significantly from the parabolic band. The arrows indicate film band gaps for different film sizes. They are not evenly spaced because $k \propto 1/L$. Note that in part (a) ϵ^{film} decreases monotonically with size while in part (b) the change is nonmonotonic.

consequence of quantum confinement per se, but rather a consequence of using the parabolic dispersion relation [Eq. (1b)]. We will see below that other dispersion relations [e.g., Fig. 3(b)] can lead to a nonmonotonic variation of the band gap with film sizes.

(b) The envelope function $\sin(k_z - k_{z0})z$ in Eq. (3) has a minimum amplitude (= 0) at the boundaries (0, L). There are no solutions in which the envelope function has maxima at the boundaries.

(c) All energy eigenvalues of Eq. (1) depend on the film size. This is a consequence of the fact that only $j \neq 0$ solutions are allowed in Eq. (4). Thus, $k_z - k_{z0} \propto 1/L$ and $\epsilon \propto 1/L^2$ for all states.

The coupling between different bands [Luttinger's model: see Eq. (2c)] does not alter the results (a)–(c) in any qualitative way.

There are two purposes to this paper. First, we wish to find to what extent the EMA predictions (a)–(c) are real properties of semiconductor films, as opposed to being consequences of the EMA particle-in-a-box approximations (i)–(iv). To address this question we will establish the "exact solution" by directly diagonalizing the film Hamiltonian with the complete potential shown schematically in Fig. 2(a) using pseudopotential band structure techniques. In doing so we avoid the decoupling of the external confining potential from the microscopic periodic potential $V_{\text{periodic}}(\mathbf{r})$, circumvent the effective-mass approximation to the energy [Eq. (1b)] and wave function [Eq. (2b)], and use a multiband description of the film states [rather than Eq. (3)]. *Application to Si and GaAs quantum films of different thicknesses and layer orientations reveals that the basic EMA predictions (a)–(c) are often not borne out.* In particular, (a) deviations from $\epsilon \propto 1/L^2$ and even nonmonotonic oscillations of ϵ with L are possible, (b) cosine-type envelope functions with maxima at the boundaries can occur, and (c) the energies of some levels do not change at all with the film size ("zero-confinement states"). Given these qualitative failures of the EMA, our second objective in this paper is to develop an approach which captures the main features of the direct calculation yet is essentially as simple as the effective-mass approach. This "truncated crystal approach" builds on an earlier idea[6] whereby periodic band structure solutions are mapped onto those of finite-size quantum structures. *The nonsurface eigenvalue spectrum of the finite quantum structure is then found to correspond to a subset of the eigenvalue spectrum of the infinite periodic solid, evaluated at special k points.* Our method extends such ideas from simple one orbital per atom representations[6] and single band models[7–10] to a general-purpose description which reliably reproduces the main features of the direct calculations on quantum films.

II. THE TRUNCATED CRYSTAL REPRESENTATION

We describe the truncated crystal (TC) approach in a parallel manner used in the preceding section to describe the EMA. The basic point is that the periodic part $u_{n,\mathbf{k_0}}$ of the Bloch wave function in Eq. (2b) or (2c) need not be limited to $\mathbf{k} = \mathbf{k_0}$. With this generalization, the film wave functions can be expanded in terms of the exact bulk states, $\{\psi^{\text{bulk}}_{n,\mathbf{k}}(\mathbf{r})\}$. The method is constructed as follows.

(i) We use the actual bulk dispersion relation $\epsilon^{\text{bulk}}_{n,\mathbf{k}}$ of Eq. (1a) and wave functions $\psi^{\text{bulk}}_{n,\mathbf{k}}(\mathbf{r})$ of Eq. (2a), rather than their EMA counterparts in Eqs. (1b) and (2b). Hence, no assumption that $\mathbf{k} - \mathbf{k_0}$ needs to be small is made here.

(ii) As a consequence, we can use the complete quantization conditions of Eq. (5), rather than the form of Eq. (4) restricted to $\mathbf{G} = \mathbf{0}$. For $\mathbf{G} \neq \mathbf{0}$, we can have a $j = 0$ solution, which is forbidden in the EMA.

(iii) We select a set of basis functions in which the quantized k_z^* is retained, but the $\mathbf{G}$ values of the pure sine basis [Eq. (5)] is replaced by the band index n. In analogy with Eq. (3), we now construct a truncated crystal basis set representing a destructive interference, namely,

$$\chi^{\text{TC}}_{n,k_z^*}(\mathbf{r}) = \begin{cases} \psi^{\text{bulk}}_{n,k_z^*}(\mathbf{r}) - \psi^{\text{bulk}}_{n,-k_z^*}(\mathbf{r}) & \text{if } 0 < z \leq L \\ 0 & \text{if } 0 > z \geq L \end{cases} , \tag{7}$$

Since

$$\psi^{\text{bulk}}_{n,-k_z^*}(\mathbf{r}) = [\psi^{\text{bulk}}_{n,k_z^*}(\mathbf{r})]^* , \tag{8}$$

Eq. (7) can be rewritten for $0 < z < L$ as

$$\chi^{\text{TC}}_{n,k_z^*}(\mathbf{r}) = \sqrt{2}[u^R_{n,k_z^*}(\mathbf{r}) \sin(k_z^* z) + u^I_{n,k_z^*}(\mathbf{r}) \cos(k_z^* z)] , \tag{9}$$

where $u^R_{n,k_z^*}(\mathbf{r})$ and $u^I_{n,k_z^*}(\mathbf{r})$ are the real and imaginary parts of the bulk Bloch wave $u_{n,k_z^*}(\mathbf{r})$ in $\psi^{\text{bulk}}_{n,\mathbf{k}}(\mathbf{r}) =$

$u_{n,\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$, respectively. Equation (9) is clearly different from the pure sine form of Eq. (3). Note that the boundary conditions at $z=0$ and $z=L$ can be satisfied either by the trigonometric functions appearing in Eq. (9) or by the Bloch-periodic parts, if the latter have nodal planes at the boundaries. The periodic function $u_{n,k_z^*}(\mathbf{r})$ can have such nodes if its orbital character is p, d, etc., but not s. A constant phase factor is implicit in $\psi^{\text{bulk}}_{n,k_z^*}(\mathbf{r})$ in Eq. (7), as well as in $u_{n,k_z^*}(\mathbf{r})$ in Eq. (9). It has to be determined by the boundary conditions.

(iv) Using this complete TC basis set, a film eigenstate (f) at $\bar{\Gamma}$ can be expanded as

$$\psi^{\text{TC}}_{f,\bar{\Gamma}}(\mathbf{r}) = \sum_n \sum_{k_z^*} a_{n,f}(k_z^*)\chi^{\text{TC}}_{n,k_z^*}(\mathbf{r}) \quad . \tag{10}$$

Since $k_z^* \propto j$ and $j = 0,1,2,\ldots,j_L$ we have a total of j_L+1 states, hence one of these states is unphysical. We return to this point in Sec. III E.

One can now proceed and directly diagonalize the film Hamiltonian [with the potential in Fig. 2(a)] using the representation of Eq. (10). This will produce the exact film eigenvalues

$$\epsilon^{\text{TC}}_{f,\bar{\Gamma}} = \sum_n \sum_{k_z^*} |a_{n,f}(k_z^*)|^2 \epsilon^{\text{bulk}}_{n,k_z^*} \quad . \tag{11}$$

Our central observation here is that such a matrix representation of Eqs. (10) and (11) is essentially diagonal in the band index n and wave vector k_z^*. This is evidenced by the fact (see below) that a single TC basis function can reproduce well the results of the direct diagonalization. The truncated crystal *approximation* to Eqs. (10) and (11) thus consists of retaining in these expansions just the dominant terms. In its simplest form, we retain in these equations just a single basis function with $f=(n,j)$, so

$$\psi^{\text{TC}}_{f,\bar{\Gamma}}(\mathbf{r}) \approx \chi^{\text{TC}}_{n,k_z^*}(\mathbf{r}) = \sqrt{2}[\, u^R_{n,k_z^*}(\mathbf{r})\sin(k_z^* z) + u^I_{n,k_z^*}(\mathbf{r})\cos(k_z^* z)] \tag{12}$$

and

$$\epsilon^{\text{TC}}_{f,\bar{\Gamma}} \approx \epsilon^{\text{bulk}}_{n,k_z^*} \quad , \tag{13}$$

where

$$k_z^* = \frac{\pi}{L} j \quad . \tag{14a}$$

Note that Eq. (14a) is derived for a special case where the film orientation $\hat{\boldsymbol{\epsilon}}$ is in the z direction. More generally, k_z^* in Eqs. (12) and (13) should be replaced by

$$\mathbf{k}^*_{\hat{\boldsymbol{\epsilon}}} = \frac{\pi}{L} j\hat{\boldsymbol{\epsilon}} \quad . \tag{14b}$$

In either case, j is given by

$$j = \begin{cases} 1,2,3,\ldots,j_L & \text{if } n=1 \quad , \\ 0,1,2,\ldots,j_L & \text{if } n\neq 1 \quad . \end{cases} \tag{14c}$$

For $n \neq 1$, it turns out that the lowest energy state (at $\mathbf{k}=\mathbf{k_0}$) should be deleted from this set (see Sec. III E below). In general, however, this state may not be at the Γ point. In such a case ($\mathbf{k_0} \neq \mathbf{0}$), the counting of j in Eq. (14) should start at $\mathbf{k}^* = \mathbf{k_0}$, not at Γ. To avoid confusion, we reserve the symbol j for $\mathbf{k_0}=0$ (Γ) and use j' for $\mathbf{k_0} \neq \mathbf{0}$. Equations (12)–(14) define the simplest TC approximation. These expressions should be contrasted with the EMA Eqs. (3), (6), and (4), respectively.

To the extent that the TC approximation is sufficiently accurate (see below), it provides an exceedingly simple and useful result: *it predicts a one-to-one mapping between the film energy eigenvalues and those of the periodic bulk $\epsilon^{\text{bulk}}_{n,k_z^*}$ at some special k points k_z^*.* This is precisely the procedure followed empirically in the early TC approaches,[6–9] where it was demonstrated (via direct tight binding Hamiltonian diagonalizations) that the eigenvalues of finite clusters form a subset at special k points of the eigenvalues of the periodic crystal. This was used in the past to predict the bulk band structure from finite cluster calculations.[6–9] Here, we will do the reverse: since we are able to reliably calculate (or fit) bulk band structures, we will use the latter, via Eqs. (12)–(14), to predict the properties of finite quantum structures. A somewhat similar approach[11] has been adopted earlier in an *ad hoc* fashion. It was proposed that the EMA approach can be extended to a nonparabolic energy dispersion by replacing the EMA bands with the actual bulk energy dispersion of Eq. (1a). The rationale behind our approach is quite different: (i) in Ref. 11 the wave function was not constructed from Eq. (12) and (ii) the EMA quantization condition (4) was used instead of the TC condition Eq. (14). This approach thus missed most of the interesting features of the quantum films, i.e., items (a)–(c) listed in the end of Sec. I.

The accuracy of the simple TC approximation of Eqs. (12)–(14) can be examined directly by comparing these solutions to an "exact" numerical diagonalization of the film Hamiltonian [using a potential such as that shown schematically in Fig. 2(a)]. The results of such direct diagonalizations will be described in the next section, where we find that the valence states of (001)-oriented Si films are accurately described by a single TC basis function $\chi^{\text{TC}}_{n,k_z^*}(\mathbf{r})$ while the conduction bands, which have a minimum at $\Delta_{1c} = 0.85X_{1c}$, require $\sim$2 basis functions. Similar results hold for Si(110) and GaAs(110) films. However, a larger number of TC basis functions might be needed when (i) the film involves lower symmetry orientations, (ii) when a film state with a large off-$\bar{\Gamma}$ $\bar{\mathbf{k}}$-vector is considered, or (iii) when one is interested in lower-dimension quantum systems (such as wires or dots).

III. ACCURACY OF THE TRUNCATED CRYSTAL APPROXIMATION

A. Direct calculations for quantum films

An exact diagonalization of the film problem involves solving

$$[-\tfrac{1}{2}\nabla^2 + V^{\text{film}}(\mathbf{r})]\psi^{\text{film}}_f(\mathbf{r}) = \epsilon^{\text{film}}_f \psi^{\text{film}}_f(\mathbf{r}) \quad , \tag{15}$$

where $V^{\text{film}}(\mathbf{r})$ is the potential of the film [shown schematically in Fig. 2(a)]. It includes a quasiperiodic

piece $V_{\text{periodic}}(\mathbf{r})$ inside the film and potential walls outside it. In practice, we construct $V^{\text{film}}(\mathbf{r})$ from a superposition of atomic pseudopotentials. Equation (15) is solved by imposing periodic boundary conditions on the N-layer film straddled by N_{vacuum} layers of vacuum. This transforms the film problem into a Bloch-periodic band structure problem, solved here by expanding ψ_f^{film} in plane waves. We checked that the results are independent of the number of vacuum layers to within 0.02 eV for $N_{\text{vacuum}} \sim 8d$, where d is the interlayer spacing. The solution of Eq. (15) is repeated for films with different thickness L and different layer orientations, $\hat{\epsilon}_{(001)}$ and $\hat{\epsilon}_{(110)}$.

Figure 4(a) depicts the xy-planar averaged potential $\bar{V}^{\text{film}}(z)$ for a 12 atomic layer Si(001) film embedded in 8 layers of vacuum. This potential was constructed from a superposition of empirical local Si atomic pseudopotentials fitted to the bulk band structure and the film's work function [see the Appendix (Refs. 12–19)]. In the interior of the film the potential is quasiperiodic, having an average value V_{av} [thick horizontal line in Fig. 4(a)]. Outside the film the potential approaches the vacuum level. The position of the valence-band maximum in bulk Si is indicated in this figure by the horizontal dashed line. It

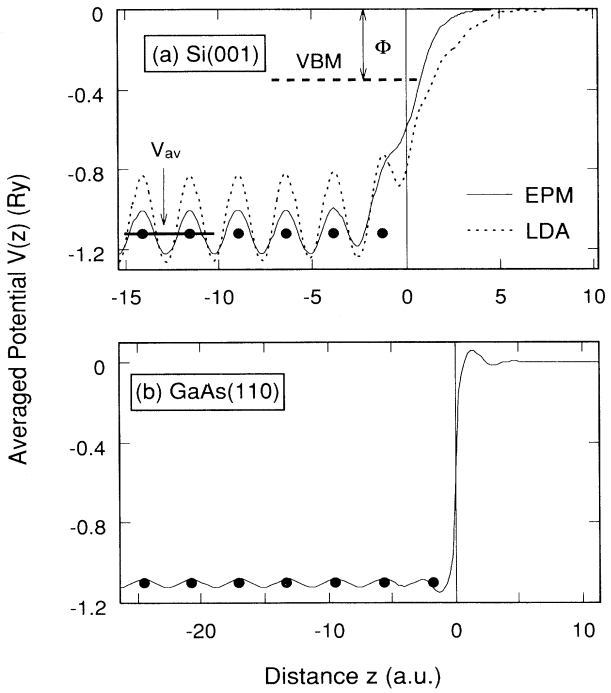

FIG. 4. (a) The xy-planar averaged potential for a 12-layer Si(001) film. The solid line gives the EPM result whereas the dotted line is the local part of the self-consistent LDA potential. The origin of the potential is at the vacuum level and the origin of the z axis is at a position half interlayer spacing outside the surface atom (the "truncated crystal boundary"). V_{av} is the z averaged potential at the interior of the film. This is used to determine the position of the valence-band maximum (VBM) and the work function Φ (Ref. 15). (b) The xy-planar averaged EPM potential for a 14-layer GaAs(110) film. Solid dots indicate the positions of atomic planes.

is determined by aligning V_{av} in the interior of the film with the same quantity calculated in bulk periodic Si (a procedure analogous to that used in determining heterojunction band offsets[20,21]). The work function Φ (~4.9 eV) is then given by the distance of the VBM from the vacuum level. Since the Γ_{1v}-$\Gamma_{25'v}$ valence-band width of Si is ~12.6 eV, the total confining potential [Fig. 1(a)] is 12.6 + 4.9 = 17.5 eV. A similar plot of the averaged potential for a 14 layer GaAs(110) film is shown in Fig. 4(b).

These "exact" film problems differ from ideal textbook depictions [e.g., Fig. 2(a)] of a film in two ways. First, the potential barriers are finite. We will therefore consider only film states that are well confined within the potential barrier, i.e., all valence bands and conduction bands with energies at least about 2 eV below the vacuum level. Second, the potential barriers are nonabrupt. Both effects contribute to the coupling to surface states. This is undesirable in the present study which aims at testing simple, surfaceless models such as the EMA and TC. We have minimized the coupling effects by deliberately using a non-self-consistent pseudopotential description, rather than a self-consistent one; the latter would propagate the unwanted surface effects [notably, the surface dipole potential in Fig. 4(a)] deeper into the film. In any event, surface states are identified in the present study by examining their wave functions ψ_f^{film}; they are then omitted from the comparison with EMA or TC states.

The same atomic pseudopotential is used to solve the bulk band structure, thus providing $\psi_{n,\mathbf{k}}^{\text{bulk}}(\mathbf{r})$ and $\epsilon_{n,\mathbf{k}}^{\text{bulk}}$ [Eqs. (2a) and (1a), respectively] needed to construct the truncated crystal solutions. Like the film eigenstates, those of the bulk are expanded in plane waves, i.e.,

$$\psi_{n,\mathbf{k}}^{\text{bulk}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} B_n(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} \quad , \tag{16}$$

where $\mathbf{G}$ are reciprocal lattice vectors and $B_n(\mathbf{G})$ are the variational expansion coefficients.

The direct solution of the film problem [Eq. (15)] avoids the approximations underlying the effective-mass particle-in-a-box model in that (i) the solutions are valid at any $\mathbf{k}$ value (not just near a band edge), (ii) the boundary conditions imposed by the external potential can couple to the quasiperiodic potential, and (iii) intervalley coupling is allowed. The direct solution also avoids the truncated crystal approximation, i.e., use of a single TC basis function. In what follows, we will compare the predictions of the simple EMA and TC approaches with the "exact" film solutions, thus assessing the utility of the former methods.

B. Structure and symmetry of (001) and (110) films

Equation (14) shows that the wave vectors $\mathbf{k}^*$ entering the TC representation depend on the film's geometry through its layer orientation $\hat{\epsilon}$ and thickness L.

Figure 5 illustrates the relation between the film's geometries and the bulk crystal structure shown in a cubic

cell of lattice constant a. The bulk translation vector $\mathbf{T}_{\hat{\epsilon}}$ and the shortest nonzero projection $\mathbf{a}_{\hat{\epsilon}}$ of the primitive lattice vectors in the film direction $\hat{\epsilon}$ are indicated. Intuitively, one might think that the translation vector $\mathbf{T}_{\hat{\epsilon}}$ determines how the film evolves when additional layers of Si are added. However, it turns out that this is determined solely by the projection $\mathbf{a}_{\hat{\epsilon}}$, not by $\mathbf{T}_{\hat{\epsilon}}$. For example, for (001) films [Fig. 5(a)], $a_{(001)} = 2d_{(001)}$ where $d_{(001)}$ is the interlayer spacing, thus even- and odd-layered films have different symmetry properties: the point group of (001) films oscillates between D_{2h} for even-layer films and D_{4h} for odd layer films. These symmetries do not follow $T_{(001)}$ which equals $4d_{(001)}$. For $\hat{\epsilon} = (110)$, since $a_{(110)} = d_{(110)}$ [Fig. 5(b)] the point group symmetry is D_{2h} (Si) and C_{2v} (GaAs) independent of the number of layers N, despite the fact that $T_{(110)} = 2d_{(110)}$. The relations between $\mathbf{T}_{\hat{\epsilon}}$, $\mathbf{a}_{\hat{\epsilon}}$, L, and $d_{\hat{\epsilon}}$ are summarized in Table I, which also gives the TC $\mathbf{k}^*$ vectors derived from these structural data. We see that for (001) films the special $\mathbf{k}^*$ vectors $\frac{2\pi}{a}(0, 0, \frac{2j}{N})$ lie along the Γ-X line in the bulk Brillouin zone (i.e., direction Δ), whereas for (110) films the special $\mathbf{k}^*$ vectors $\frac{2\pi}{a}(\frac{j}{N}, \frac{j}{N}, 0)$ lie along the Γ-K line (direction Σ).

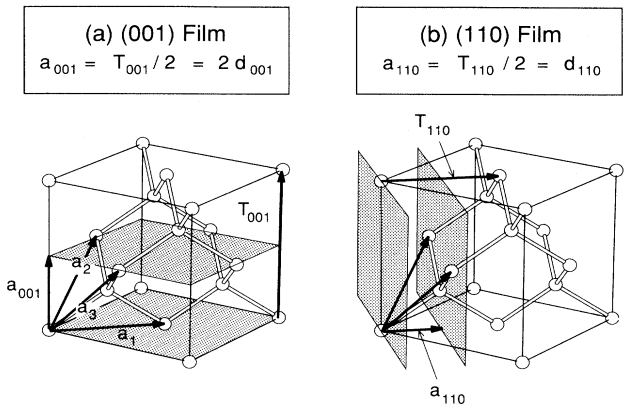

FIG. 5. Relations between bulk translation vector $\mathbf{T}_{\hat{\epsilon}}$ and the projection vector $\mathbf{a}_{\hat{\epsilon}}$ of the primitive lattice vectors $\{\mathbf{a}_i\}$ in the film direction $\hat{\epsilon}$ for (a) (001) and (b) (110) films. See Table I.

C. Mapping bulk states onto states of quantum films: Deconfinement effects

The truncated crystal approximation of Eqs. (12)–(14) provides a simple way for guessing the energies of quantum structures from the bulk dispersion relation. Table I gives the TC wave vectors $\mathbf{k}^*$ of Eq. (14) in terms of N and a for (001) and (110) films. Denoting by "HVB" the highest valence band along the corresponding line (Table I) in the bulk Brillouin zone, and by "LCB" the lowest conduction band along this line, the minimum band gap of an N-layer (001)-oriented film is given by

$$\epsilon_g^{\mathrm{TC}}[(001), N] = \epsilon_{\mathrm{LCB}}^{\mathrm{bulk}}\left[\mathbf{k_0} \pm \frac{2\pi}{a}\left(0, 0, \frac{2}{N}\right)\right] - \epsilon_{\mathrm{HVB}}^{\mathrm{bulk}}\left[\frac{2\pi}{a}\left(0, 0, \frac{2}{N}\right)\right], \tag{17a}$$

whereas that of a (110) film is

$$\epsilon_g^{\mathrm{TC}}[(110), N] = \epsilon_{\mathrm{LCB}}^{\mathrm{bulk}}\left[\mathbf{k_0} \pm \frac{2\pi}{a}\left(\frac{1}{N}, \frac{1}{N}, 0\right)\right] - \epsilon_{\mathrm{HVB}}^{\mathrm{bulk}}\left[\frac{2\pi}{a}\left(\frac{1}{N}, \frac{1}{N}, 0\right)\right], \tag{17b}$$

where $\mathbf{k_0}$ is the conduction-band minimum and sign (±) is chosen to obtain the lowest energy gap. For the direct gap, $\mathbf{k_0} = 0$. As the number of layers N increases, the wave vector k_z^* approaches zero, so the band gaps approach the bulk values [Fig. 3(b)]. Equations (17) establish a simple relation between the geometry of the film (i.e., the number N and orientation $\hat{\epsilon}$ of the layers) and its eigenvalue spectrum. These relationships replace Eq. (6) in the effective-mass particle-in-a-box approximation. The two approaches become identical only when the true bulk dispersion relation is parabolic along all of these directions. These equations can hence be used to determine the minimum size L_c for which this condition

TABLE I. Structural information on (001) and (110) films. Here, $\mathbf{a_1}$, $\mathbf{a_2}$, and $\mathbf{a_3}$ are the primitive lattice vectors, and a is the cubic lattice parameter. The bulk translation vector in the direction $\hat{\epsilon}$ is given by $\mathbf{T}_{\hat{\epsilon}}$, whereas $\mathbf{a}_{\hat{\epsilon}}$ is the shortest nonzero projection of the primitive lattice vectors along the same direction. See Fig. 5 for a graphical depiction of $\mathbf{T}_{\hat{\epsilon}}$ and $\mathbf{a}_{\hat{\epsilon}}$. The film length is $L = Nd_{\hat{\epsilon}}$, where N is the number of layers and $d_{\hat{\epsilon}}$ is the interlayer spacing. In the last two lines, the wave vector $\mathbf{k}^*$ of the film, according to the quantization condition of Eq. (14), and the maximum j (starting at Γ) value for even-layer (001) and for all (110) films are given.

Film orientation	(001)	(110)
Translation vector	$\mathbf{T_{(001)}} = -\mathbf{a_1} + \mathbf{a_2} + \mathbf{a_3}$	$\mathbf{T_{(110)}} = \mathbf{a_1}$
Length of $\mathbf{T}_{\hat{\epsilon}}$	$T_{(001)} = 4d_{(001)}$	$T_{(110)} = 2d_{(110)}$
Shortest projection	$\mathbf{a_{(001)}} = \mathbf{T_{(001)}}/2$	$\mathbf{a_{(110)}} = \mathbf{T_{(110)}}/2$
length of $\mathbf{a}_{\hat{\epsilon}}$	$a_{(001)} = 2d_{(001)}$	$a_{(110)} = d_{(110)}$
Film length	$L = Nd_{(001)} = N\frac{1}{4}a$	$L = Nd_{(110)} = N\frac{\sqrt{2}}{4}a$
Film's $\mathbf{k}^*$-vector	$\frac{\pi}{L}j\hat{\epsilon}_{(001)} = \frac{2\pi}{a}(0, 0, \frac{2j}{N})$	$\frac{\pi}{L}j\hat{\epsilon}_{(110)} = \frac{2\pi}{a}(\frac{j}{N}, \frac{j}{N}, 0)$
Maximum j	$j_L = L/a_{(001)} = N/2$	$j_L = L/a_{(110)} = N$

holds, thus finding the limits of the applicability of the EMA.

Equations (17a) and (17b) illustrate an important point on the nature of "quantum confinement": In the EMA one finds that the band gaps of 2D periodic systems scale as $\epsilon_g \propto 1/L^2$ [Eq. (6)]. This, however, reflects the choice of a simplified, parabolic dispersion relation in Eq. (1b). If such a parabolic dispersion is inserted in Eqs. (17a) and (17b) instead of $\epsilon^{\rm bulk}_{n,\mathbf{k}}$, we will find too that $\epsilon \propto 1/L^2$ [Fig. 3(a)]. However, the more general Eqs. (17a) and (17b) show that such a scaling is *not* a general property of 2D periodic systems. If the dispersion relation of the HVB and the LCB is like that shown schematically in Fig. 3(b), the band gap will first increase with L, then, at smaller L values, it will turn around and *decrease* with L.[11] We term such a behavior "deconfinement." Equations (17a) and (17b) thus provide guidelines on how to select a bulk material such that a film made from it will exhibit deconfinement.

D. Comparison of truncated crystal and direct calculations

We next examine the accuracy of the truncated crystal approach by comparing with the results of direct calculations.

1. Energy levels

The left hand side of Fig. 6(a) shows as horizontal lines the directly calculated energy eigenvalues $\epsilon^{\rm film}_{f,\bar\Gamma}$ for an $N = 12$ layer Si(001) film. Surface states, appearing at 1.2–1.6 eV above the VBM, are discarded from the following discussion. The bulk bands are indexed in Fig. 6 by the band index n in increasing energy order: $n = 1$, 2, 3, and 4 for the four valence bands and $n = 5$ and 6 for the two lowest conduction bands. The TC approximation to the directly calculated levels is given [Table I and Eq. (13)] by $\epsilon^{\rm bulk}_n[\frac{2\pi}{a}(0, 0, \frac{2j}{N})]$. These TC results are indicated in Fig. 6(a) by the intersections of the dotted vertical lines taken at $k^*_z = \frac{2\pi}{a}(0, 0, \frac{2j}{12})$ with the bulk bands along Γ-X. The error in the TC approximation is hence given by the amount by which the crossings of the vertical dashed lines and the bulk bands miss the directly calculated values $\epsilon^{\rm film}_{f,\bar\Gamma}$ (solid dots in Fig. 6). We see that the TC method provides an accurate and natural classification of film states in terms of the bulk band structure.

The quantitative mapping between the TC wave function $\psi^{\rm TC}_{f,\bar\Gamma}(\mathbf{r})$ [Eq. (12)] and the exact film state $\psi^{\rm film}_{f,\bar\Gamma}(\mathbf{r})$ [Eq. (15)] can be obtained by calculating the projection $|a_{n,f}(k^*_z)|^2$ of Eq. (10), i.e.,

$$a_{n,f}(k^*_z) = \langle \psi^{\rm film}_f(\mathbf{r}) | \chi^{\rm TC}_{n,k^*_z}(\mathbf{r}) \rangle \tag{18}$$

depicted in Tables II and III for valence and conduction bands of (001)-oriented Si film, respectively. The $N = 12$ monolayer film contains 12 atoms, hence 24 doubly occupied valence bands. Of these, two are surfacelike bands, leaving 22 intrinsic film valence bands whose energies at the film's $\bar\Gamma$ point are shown in the tables. We see that

TABLE II. Projections (in %) of the valence-band film (f) states $\psi^{\rm film}_{f,\bar\Gamma}(\mathbf{r})$ to the TC basis functions $\chi^{\rm TC}_{n,j}(\mathbf{r})$ [Eq. (18)]. The last column gives the energy difference (in eV) of Eq. (19), for a 12-layer Si(001) film. The sign (-) indicates that the projection $|a_{n,f}(j)|^2 < 0.1\%$ for that particular $(f; n, j)$.

	$n=1$: sine-type						$n=2$: sine-type						$n=3$: cosine-type						$n=4$: sine-type						
f	$j=0$	1	2	3	4	5	6	$j=0$	1	2	3	4	5	6	$j=0$	1	2	3	4	5	6	$j=0$	1	2	3	4	5	6	$\delta\epsilon$ (eV)
1	-	99.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.02
2	-	-	99.8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.03
3	-	-	-	99.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.04
4	-	-	-	-	99.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02
5	-	-	-	-	-	98.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.04
6	-	-	-	-	-	-	97.8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.06
7	-	-	-	-	-	0.1	-	-	-	-	0.2	-	96.4	-	-	-	-	-	-	99.9	-	-	-	-	-	-	-	-	-0.12
8	-	-	-	-	-	-	0.4	-	-	0.3	-	94.5	-	-	-	-	-	-	99.8	-	-	-	-	-	-	-	-	-	-0.20
9	-	-	-	-	-	0.1	-	-	0.3	-	92.1	-	0.8	-	-	-	-	99.4	-	-	-	-	-	-	-	-	-	-	-0.28
14	-	-	-	-	-	-	0.3	-	-	89.5	-	1.5	-	-	-	-	98.9	-	-	-	-	-	-	-	-	-	-	-	-0.32
19	-	-	-	-	-	0.1	-	-	88.8	-	1.8	-	0.4	-	-	98.0	-	-	-	-	-	-	-	-	-	-	-	-	-0.26
10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	98.5	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.03
17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.03
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.01
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.9	-	0.0
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.9	-	-	0.02
16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.8	-	-	-	0.03
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.7	-	-	-	-	0.04
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.8	-	-	-	-	-	0.01

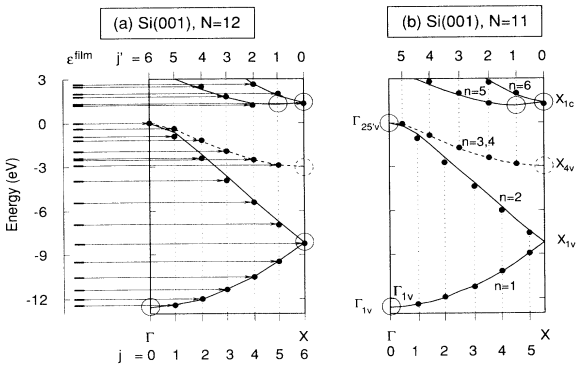

FIG. 6. Mapping of the directly calculated film energy levels (solid dots) at the 2D Brillouin zone center $\bar{\Gamma}$ onto the truncated crystal energy levels for (a) a 12-layer Si(001) film and (b) an 11-layer Si(001) film. The vertical dotted lines indicate the quantized k_z^* values. Thus, the intersections between the bulk dispersion (solid line for nondegenerate states and dashed line for double degenerate states) and the vertical dotted lines give the truncated crystal energy levels. The difference between these and the solid dots represent the errors in the TC approach. A solid (dashed) large open circle indicates one (two) spurious TC state(s) to be removed from the TC spectrum. The quantized k_z^*'s are indexed both by j (starting from Γ) and by j' (starting from X). Note that for the 11-layer film, the verticle dotted lines representing j and j' do not coincide with each other. In this case k_z^*'s are band index dependent.

for film valence-band states the projection of Eq. (18) selects but a single, dominant TC basis function $\chi^{\rm TC}_{n,k_z^*}$ with a typical coefficient $|a_{n,f}(k_z^*)|^2 = 98$–$99\,\%$, with the exception of film states derived from the bulk $n = 2$ band, for which wave function analysis (not shown) showed that the coupling to surface states is non-negligible. The TC error in the energy eigenvalue

$$\delta\epsilon(f;n,j) = \epsilon^{\rm film}_{f,\bar{\Gamma}} - \epsilon^{\rm bulk}_n\left(\frac{\pi}{L}j\right) \tag{19}$$

is given in the last column of Table II. For $n \neq 2$, it is generally as small as 0.05 eV.

For the film conduction-band states, the projection of Eq. (18) selects *two* TC basis functions (see Table III). The two, as can be seen in Fig. 6(a), are nearly degenerate. For example, in bulk Si

$$\epsilon^{\rm bulk}_{n=5,j=2} \approx \epsilon^{\rm bulk}_{n=6,j=4} \tag{20a}$$

and

$$\epsilon^{\rm bulk}_{n=5,j=3} \approx \epsilon^{\rm bulk}_{n=6,j=5} \quad . \tag{20b}$$

In addition, the dispersion of the lowest ($n = 5$) conduction band in bulk Si is such that

$$\epsilon^{\rm bulk}_{n=5,j=4} \approx \epsilon^{\rm bulk}_{n=5,j=6} \quad . \tag{21}$$

In these cases, the degenerate pairs (n, k^*_{z1}) and (m, k^*_{z2}) couple, leading to

$$\psi^{\rm TC}_{f,\bar{\Gamma}}(\mathbf{r}) \approx a_{n,f}(k^*_{z1})\chi^{\rm TC}_{n,k^*_{z1}}(\mathbf{r}) + a_{m,f}(k^*_{z2})\chi^{\rm TC}_{m,k^*_{z2}}(\mathbf{r}) \tag{22}$$

and the energy levels are given by

$$\epsilon^{\rm TC}_{f,\bar{\Gamma}}(L) \approx |a_{n,f}(k^*_{z1})|^2\epsilon^{\rm bulk}(n,k^*_{z1}) + |a_{m,f}(k^*_{z2})|^2\epsilon^{\rm bulk}(m,k^*_{z2}) \quad . \tag{23}$$

Table III shows that the sum of the projections of the low lying conduction states $|a_{n_1,f}(k^*_{z_1})|^2 + |a_{n_2,f}(k^*_{z_2})|^2 \approx 99\%$. For higher energy states, a smaller projection can be expected since TC basis functions near vacuum level were not included in the projection.

A similar mapping between directly calculated and TC eigenvalues is shown in Fig. 7 for (110) oriented Si and GaAs films. As indicated in Table I and Eq. (13), here the TC predictions are $\epsilon^{\rm bulk}_n[\frac{2\pi}{a}(\frac{j}{N}, \frac{j}{N}, 0)]$ with $j = 0, 1, 2, \ldots, N$ and $N = 6$. These TC states are shown in Fig. 7 as the intersections between dotted vertical lines at $k^* = \frac{2\pi}{a}(\frac{j}{6}, \frac{j}{6}, 0)$ with the bulk bands. The states obtained by direct calculations are shown as solid dots. Again, good agreement is obtained between the direct calculation and the TC approach.

TABLE III. Projections (in %) of the conduction-band film (f) states $\psi^{\rm film}_{f,\bar{\Gamma}}(\mathbf{r})$ to the TC basis functions $\chi^{\rm TC}_{n,j}(\mathbf{r})$ [Eq. (18)], for a 12-layer Si(001) film. The last column gives the energy difference (in eV) of Eq. (19) with the TC energies weighted by the projection $|a_{n,f}(j)|^2$. The sign (-) indicates that $|a_{n,f}(j)|^2 < 0.1\%$ for that particular $(f;n,j)$. An asterisk denotes mixed states between a surface and a bulk state. Each of them contains a fraction $\rho_f = \sum_{n,j}|a_{n,f}(j)|^2$ of the bulk state and, ideally, $\rho_{f1} + \rho_{f2} = 1$.

f	$n=5$					$n=6$			$\sum_{n,j}\|a_{n,f}(j)\|^2$	$\delta\epsilon$
	$j=2$	3	4	5	6	$j=4$	5	6		(eV)
					cosine-type					
24	-	-	57.3	-	42.5	-	-	-	99.8	0.0
27	-	54.1	-	-	-	-	44.1	-	98.2	-0.04
29	35.2	-	-	-	-	59.2	-	-	94.4	-0.23
					sine-type					
23(26)*	-	-	31.0(25.1)	-	-	-	-	30.3(13.6)	61.3 + (38.7) = 100.0	-
28(25)*	-	37.9(15.0)	-	-	-	-	24.9(20.4)	-	62.8 + (35.4) = 98.2	-
30	32.1	-	-	-	-	48.4	-	-	80.5	0.3

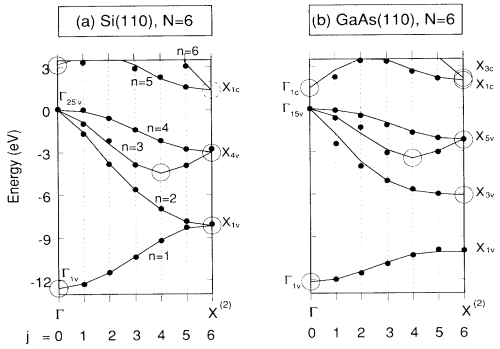

FIG. 7. Mapping of the directly calculated film energy levels (solid dots) at the 2D Brillouin zone center $\bar{\Gamma}$ onto the truncated crystal energy levels for (a) a 6-layer Si(110) film and (b) a 6-layer GaAs(110) film. The legends are the same as in Fig. 6. $X^{(2)}$ indicates the X point in the *second* bulk Brillouin zone.

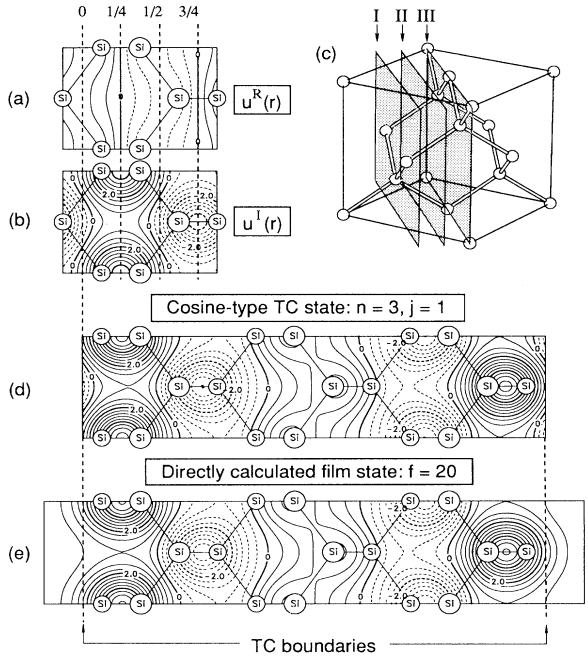

FIG. 8. Wave function contour plots for the cosine-type TC state: $n = 3, j = 1$. Solid (dashed) lines are positive (negative) contours. Parts (a) and (b) show the real and imaginary parts of the bulk Bloch state $u_{n=3,j=1}(\mathbf{r})$, scaled by $\sqrt{2}$. The contours in these panels are shown in a (110) plane [denoted as plane II in part (c); This plane contains no atoms. The atoms shown in parts (a) and (b) are projections of those in planes I and III]. Note that $u^I_{n=3,j=1}(\mathbf{r})$ has bonding character, with wave function maxima and minima at the bond center sites. Its nodal planes intersect the (001) direction at $z = 0$ and $1/2$ in units of lattice parameter a. The function $u^R_{n=3,j=1}(\mathbf{r})$ has antibonding character with zero amplitude at the sites where $u^I_{n=3,j=1}(\mathbf{r})$ has extremal values. The nodal planes for $u^R_{n=3,j=1}(\mathbf{r})$ intersect the (001) direction at $z = 1/4$ and $3/4$. Part (d) shows the TC wave function constructed according to Eq. (12) using the u's shown in parts (a) and (b). Part (e) shows the directly calculated wave function for this state. Contour step is $0.4\ /\sqrt{V_{\text{cell}}}$ where $V_{\text{cell}} = 399.86$ Å^3 is the volume of the supercell containing 12 layers of Si embedded in 8 layers of vacuum.

2. *Wave functions*

We next compare the shapes of valence wave functions of the $N = 12$ layer Si(001) film as obtained by the single basis function TC approximation [Eq. (12)] and the direct calculation [Eq. (15)]. We will illustrate the comparison separately for cosine-type (Fig. 8) and sine-type (Fig. 9) wave functions. The TC wave function $\psi^{\text{TC}}_{n,\mathbf{k}^*_{\bar{\Gamma}}}(\mathbf{r})$ is calculated from Eq. (12), given the periodic parts $u^R(\mathbf{r})$ and $u^I(\mathbf{r})$ of the bulk wave functions. The latter are the real and imaginary parts of the sum in Eq. (16). Normally, $u^R(\mathbf{r})$ and $u^I(\mathbf{r})$ are not uniquely defined since $u(\mathbf{r}) = u^R(\mathbf{r}) + iu^I(\mathbf{r})$ is defined only up to a constant complex phase factor. We fixed the phase factor by setting the origin at a bond center site so that real $B_n(\mathbf{G})$'s can be used in Eq. (16). Figures 8(a) and 8(b) show $u^R(\mathbf{r})$ and $u^I(\mathbf{r})$ for the bulk state $n = 3$, $j = 1$, i.e., $\mathbf{k}^* = \frac{2\pi}{a}(0, 0, \frac{1}{6})$. These are plotted on the plane denoted as "plane II" in Fig. 8(c) (planes I and III, which contain Si atoms, are nodal planes for this wave function). Two features of u^R and u^I are noteworthy: (i) $u^I_{n=3,j=1}(\mathbf{r})$ has curved nodal planes intersecting the (001) $= z$ direction at $z = 1/4a$ and $3/4a$. As we will see below, these will satisfy approximately the boundary conditions of the film's wave function at $z = 0$ and $z = L$. (ii) The

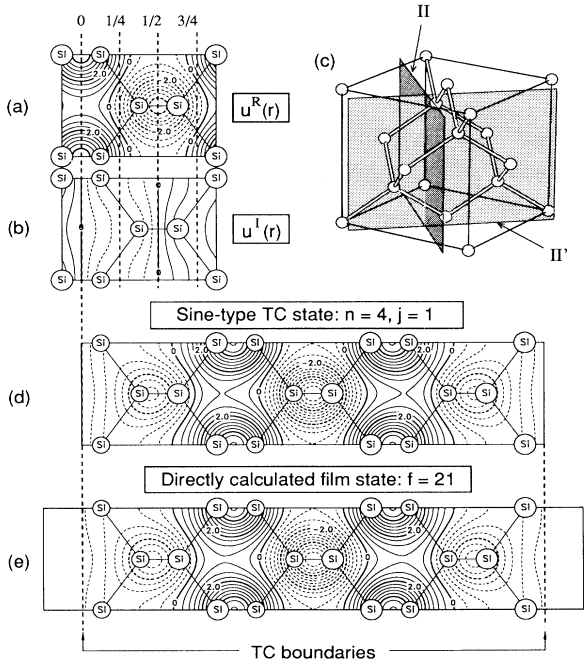

FIG. 9. Wave function contour plots for the sine-type TC state: $n = 4, j = 1$. The legends and contour step are the same as in Fig. 8. The plane on which the contours are shown is, however, not the (110) plane of Fig. 8(c) but the $(\bar{1}10)$ plane [plane II′ in (c)]. For this state, all the $(\bar{1}10)$ atomic planes are the nodal planes. In bulk the $n = 3, j = 1$ and $n = 4, j = 1$ states are doubly degenerate. This degeneracy was lifted in the construction of the TC states by introducing a very small lattice distortion (~0.01% in units of a) in the [110] direction.

function $u^R_{n=3,j=1}(\mathbf{r})$ has a low amplitude over the entire plane; its maximum amplitude is much smaller than that of $u^I_{n=3,j=1}(\mathbf{r})$. Hence, for $n = 3$, $j = 1$ the TC wave function of Eq. (12) is essentially cosinelike,

$$\psi^{\rm TC}_{n=3,j=1}(\mathbf{r}) \sim \sqrt{2} u^I_{n=3,j=1}(\mathbf{r}) \cos(k^*_z z) \quad . \tag{24}$$

This type of envelope function (having maxima at the boundaries) is absent in an EMA description [Eq. (3)].

The two functions $u^R_{n=3,j=1}(\mathbf{r})$ and $u^I_{n=3,j=1}(\mathbf{r})$ are combined according to Eq. (12) to give the TC wave function $\psi^{\rm TC}_{n=3,j=1}(\mathbf{r})$ shown in Fig. 8(d). It is compared with the directly calculated film wave function $\psi^{\rm film}_{f,\Gamma}(\mathbf{r})$ shown in Fig. 8(e). We see that the single TC basis function approximation provides an adequate representation of the film wave function. In particular, the boundary conditions at $z = 0$ and $z = L$ are approximately satisfied in the TC representation by the nodal surface of $u^I_{n=3,j=1}(\mathbf{r})$, not by the envelope function.

A similar plot is shown in Fig. 9 for the sinelike $n = 4$, $j = 1$ film state. Notice that $u^R_{n=4,j=1}(\mathbf{r})$, not $u^I_{n=4,j=1}(\mathbf{r})$, is the dominant term here, so

$$\psi^{\rm TC}_{n=4,j=1}(\mathbf{r}) \sim \sqrt{2} u^R_{n=4,j=1}(\mathbf{r}) \sin(k^*_z z) \quad . \tag{25}$$

The u's in Figs. 9(a) and 9(b) are combined according to Eq. (12) to give the TC wave function $\psi^{\rm TC}_{n=4,j=1}(\mathbf{r})$ shown in Fig. 9(d). In this case, the boundary conditions at $z = 0$ and $z = L$ are satisfied by the sine envelope function, $\sin(k^*_z z)$, not by the nodal surfaces of $u^R_{n=4,j=1}(\mathbf{r})$. Comparison between Fig. 9(d) and Fig. 9(e), which gives the directly calculated film state, shows again how accurate a single TC basis function representation is.

3. *Thickness dependence*

The calculations above pertain to a film with a fixed number of layers, N. We have repeated them as a function of N as well as for layer orientation $\hat{\boldsymbol{\epsilon}} = (001)$ and (110) for Si and (110) GaAs. The dependence of the highest valence-band energy level ($j = 1$) and the lowest conduction-band energy level on the layer number N and orientation $\hat{\epsilon}$ is shown in Fig. 10, where the TC results are compared with those of the direct calculation.[22] To facilitate (later) comparison with the effective-mass results, in Fig. 10 we have omitted the $j = 0$ ZCS states which have no correspondence in the EMA approach. We find excellent agreement between the two calculations for Si(001) and (110), while for GaAs(110) there is a good agreement for the valence-band states but a nearly fixed TC error of 0.4 eV in the conduction band. This might reflect interaction with surface states present only in the direct calculation.

We thus conclude that the simple TC approach of Eqs. (12)–(14) provides a reasonable approximation to the far more complex direct film calculation. Moreover, these direct numerical calculations show that the sum in Eq. (10) is rapidly converged and only one to two terms are sufficient for an accuracy of 0.02 eV. Hence, the use of the TC basis set affords high numerical accuracies without the need for large basis sets. This should be contrasted with a direct calculation, e.g., with a plane wave basis[23,24] where a much larger number of basis functions is required to achieve the same accuracy.

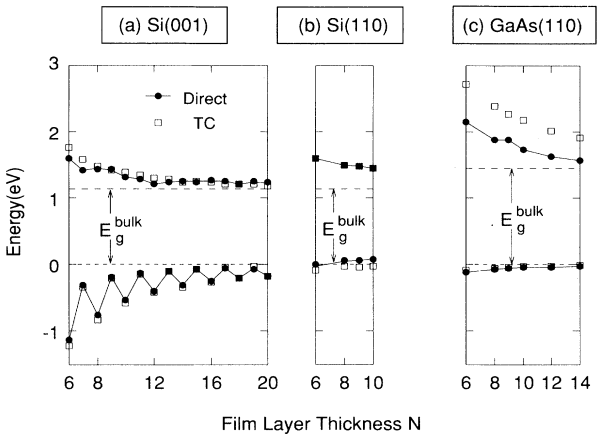

FIG. 10. The size dependence of the highest (non-ZCS) valence-band state and the lowest conduction-band state for (a) Si(001), (b) Si(110), and (c) GaAs(110). The solid dots connected by solid lines are the results of the direct calculations and the open squares are the TC predictions. The zero of energy is at the bulk VBM.

E. Eliminating the spurious TC states

As noted in Sec. II, the quantization condition $k^*_z = \frac{\pi}{L} j$, $j = 0, 1, 2, \ldots, j_L$, includes $j_L + 1$ states per band n, so one of these must be spurious. The projection technique of Eq. (18) permits a straightforward identification of such "extra TC states" that need to be deleted from the spectrum. These states, which have a null projection, are indicated in Figs. 6 and 7 by large open circles. *We see that the spurious TC states are at the energy extrema ($\mathbf{k} = \mathbf{k_0}$) of a given band.*

The elimination of one of the two degenerate zone boundary $j = j_L$ states, X_{1v} and X_{1c}, from the TC spectrum of the 12-layer Si(001) film in Fig. 6(a) is understandable: the destructive interference between two running waves can generate only *one* state satisfying the boundary condition at $z = 0$ and $z = L$. There are, however, TC states of true energy minima also being eliminated in Fig. 6(a). These are the $(n, j) = (1, 0)[\Gamma_{1v}]$; $(3, 6)[X_{4v}]$; $(4, 6)[X_{4v}]$; and $(5, 5)[\Delta_{1c}]$ states. They are eliminated from the TC basis set (similar to the elimination of the $n = 1$, $j = 0$ state in the EMA approach) because their wave functions constructed according to Eq. (12) are nodeless in the confined direction (see Fig. 11, for example, for the two degenerate X_{4v} TC states), and hence these states cannot satisfy the film's boundary conditions.

Figure 6(b) shows that the null projection rules also hold for the odd-layer ($N = 11$) Si(001) film. The $n = 2$ band in this case is an exception where, due to the fact that the film thickness $L = Nd$ is not a multiple

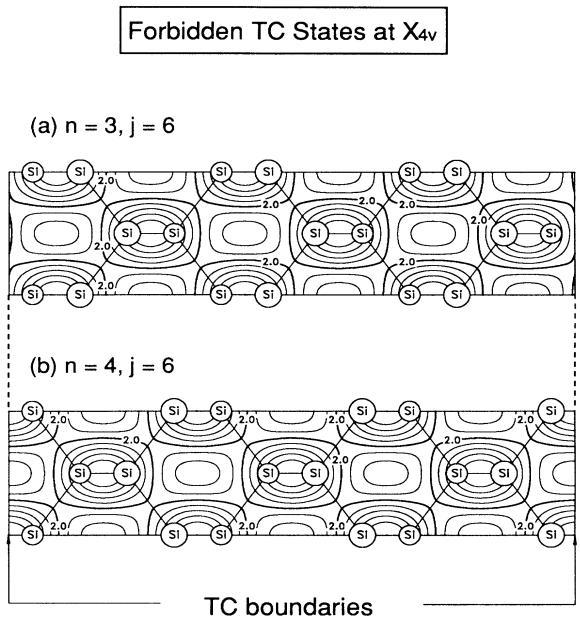

FIG. 11. Wave functions for the two degenerate forbidden TC states at X_{4v} in Fig. 6(a). (a) The $n = 3, j = 6$ state "belongs" to the same cosine band as the $n = 3, j = 1$ state in Fig. 8 and is hence plotted in the (110) plane, whereas (b) the $n = 4, j = 6$ state "belongs" to the same sine band as the $n = 4, j = 1$ state in Fig. 9 and is plotted in the $(\bar{1}10)$ plane. The contour step is $0.4/\sqrt{V_{\text{cell}}}$ where $V_{\text{cell}} = 399.86$ Å^3 is the volume of the supercell containing 12 layers of Si embedded in 8 layers of vacuum. These are spurious TC states since they do not vanish near the TC boundaries.

of $a_{(001)} = 2d$, there is no TC basis function at X_{1v} for the $n = 2$ band. The null projection in this case is governed, instead, by minimization of the band energy, i.e., eliminating the highest energy state at $\Gamma_{25'v}$.

The same rules for elimination of the extra TC states apply to the (110)-oriented films: they are the lowest energy states in each band (large open circles in Fig. 7).

IV. QUANTUM FILM PROPERTIES ABSENT IN AN EMA DESCRIPTION

Having established the close similarity between the TC and directly calculated states of quantum films, we now contrast the TC results with those of the "standard model," i.e., the effective-mass particle-in-a-box predictions (a)–(c) discussed in Sec. I.

A. Cosinelike envelope functions with maxima at the boundaries

Equation (3) shows that the EMA envelope function is sinelike, hence it reaches its minimum amplitude (zero) at the film boundaries. In contrast, Eq. (12) shows that the TC function can have both sinelike and cosinelike envelope functions. Table II classifies the directly calculated film states according to the dominant envelope function character. We see that for a (001) oriented Si film all film states originating from valence bands $n = 1$, 2, and 4 are sinelike, whereas all film states associated with the third valence band ($n = 3$) are cosinelike, hence unrepresentable by the EMA. An example of a cosinelike film state is the $(n, j) = (3, 1)$ level of Eq. (24) and Fig. 8. As discussed above, a cosinelike state can exist when $u^R_{n,\mathbf{k}^*}(\mathbf{r})$ is small for all $\mathbf{r}$ and the boundary conditions at $z = 0, L$ are satisfied by the nodal planes of $u^I_{n,\mathbf{k}^*}(\mathbf{r})$ rather than by its envelope. In contrast, in the effective-mass particle-in-a-box approach, the nodal structure of the periodic piece $u_{n,\mathbf{k_0}}(\mathbf{r})$ of the wave function of Eq. (3) is not taken into consideration. In fact, it was argued[4] that if a nodal structure exists, it will not affect the sine-envelope for a smooth varying function $\psi^{\text{EMA}}(\mathbf{r})$ near the boundaries. This is not generally true, as exampled by the existence of cosine-type states in the direct calculations [Fig. 8(e)].

B. Size-independent "zero-confinement states"

A special case of the cosine-type envelope function is the "zero-confinement state" (ZCS). This is the $k^*_z = 0$ state (i.e., $j = 0$) of Eq. (24). The discussion of Sec. III E showed that the $j = 0$ solution of a quantum film is permitted if the bulk energy at this point is not the minimum energy of that band. Hence, the $j = 0$ valence-band maximum in even-N Si(001) [Fig. 6(a)], Si(110)[Fig 7(a)], and GaAs(110) [Fig 7(b)] are all allowed solutions in a quantum film, in contrast with the EMA result of Eq. (3). With $j = 0$, Eq. (24) reads

$$\psi^{\text{TC}}_{n=3,j=0}(\mathbf{r}) \sim \sqrt{2} u^I_{n=3,j=0}(\mathbf{r}) \quad . \tag{26a}$$

Likewise, Eq. (13) gives

$$\epsilon^{\text{TC}}_{n=3,j=0} = \epsilon^{\text{bulk}}_{\text{VBM}} \quad . \tag{26b}$$

This ZCS is hence predicted to have a constant envelope throughout the film and its energy does not depend on the size N. Figure 12 shows the ZCS wave function and energies found in direct calculation for $N = 12$ Si(001). It is clear from Figs. 12(a) and 12(b) that this state is not a surface state, as it has a constant amplitude throughout the film. Figure 12(c) shows the energies of the ZCS as obtained in the direct calculation for a few even-N Si(001) films [note: odd-N Si(001) films do not have ZCS]. This figure shows that the energy of this state is indeed size-independent [solid dots in Fig. 12(c)]. Of course, the effective-mass model [represented by the dotted line in Fig. 12(c)] always predicts size dependent energies. Direct calculations for (110)-oriented films (solid dots in Fig. 13) predict a small size dependence (~25 meV for Si and ~10 meV for GaAs). This reflects the fact that for (110) films the imaginary part of the Bloch function $u^I_{n=3,k^*_x=k^*_y=0}(\mathbf{r})$ at the film boundaries has a low amplitude but not an exact nodal plane. Consequently, for (110) films inclusion of a single bulk state [Eq. (24)] does not suffice to satisfy the film boundary conditions and a small admixture of other states occurs. This leads

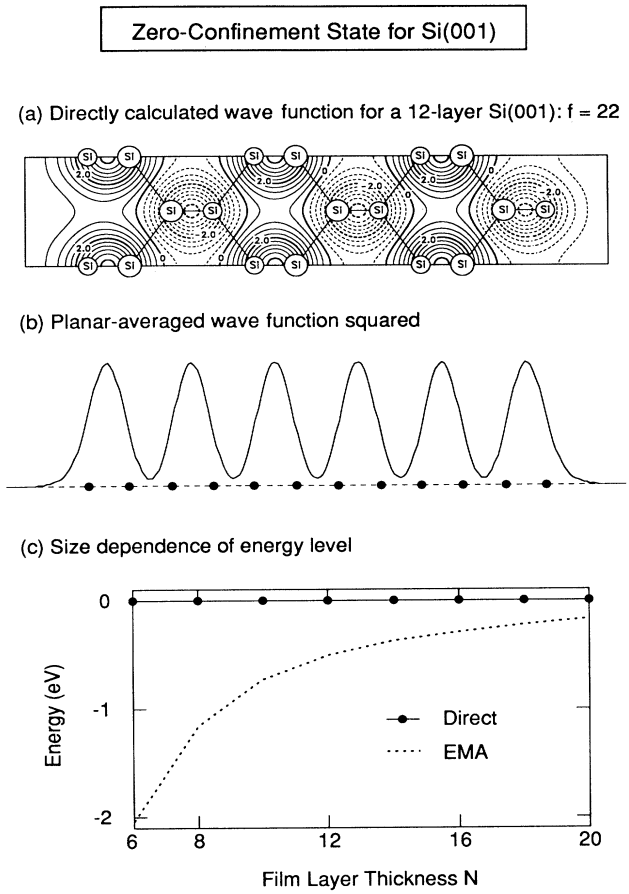

FIG. 12. The zero-confinement state for Si(001). (a) The directly calculated ZCS wave function for a 12-layer Si(001) is shown in a (110) plane [plane II in Fig. 8(c)]. The legends and contour step are the same as in Fig. 8. (b) The same wave function is squared and averaged over the xy plane perpendicular to the [001] direction. The solid dots indicate the positions of the atomic planes. Note the constant amplitude throughout the film. (c) The size dependence of the energy of the ZCS is shown as solid dots and contrasted with the EMA predictions (the dotted line).

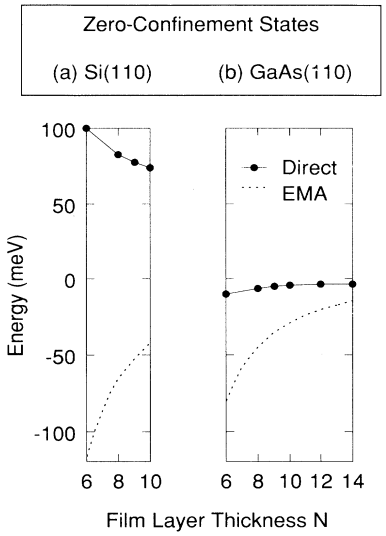

FIG. 13. The size dependence of the energy of the directly calculated zero-confinement states (solid dots) for (a) Si(110) and (b) GaAs(110). The EMA predictions are shown as dotted lines.

to a weak N dependence of the ZCS. This N-dependence, however, is far weaker than that expected from conventional quantum confinement (dotted lines in Fig. 13).

C. Even-odd energy oscillations in (001) films

We have seen in Fig. 10(a) that the Si(001) film exhibits marked even-odd energy oscillations for the non-ZCS highest valence state ($n = 4, j = 1$). It shows that the TC approach captures closely the magnitude of these oscillations, as revealed by the direct calculations. These oscillations, as expected, are absent in the effective-mass description, shown in Fig. 14. Energy level oscillations were previously noted in (001) superlattices, e.g., the L-derived and X-derived *conduction*-band states in $(AlAs)_n/(GaAs)_n$ (Refs. 20, 25, and 26) and the L-derived conduction-band states in Si_n/Ge_m.[27] The oscillations shown in Fig. 10(a) and Fig. 14 are, however, in the *valence* band. Their amplitudes are about fivefold larger than those for superlattices. The origin of such oscillations lies in the change of the point group symmetry as the number of layers N changes from even (D_{2h} symmetry) to odd (D_{4h} symmetry). In the EMA, however, the quasiperiodic potential $V_{\text{periodic}}(\mathbf{r})$ is removed (see Sec. I) so the symmetry information responsible for these oscillations is lost.

These oscillations can be analyzed in terms of the TC approach. In Figs. 8 and 9, we showed two typical film's wave functions for a 12-layer (001) film: the bulk periodic function $u_{n,\mathbf{k}}(\mathbf{r})$ in one case ($n = 3$) possesses

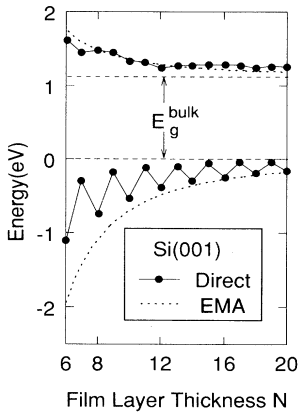

FIG. 14. Comparison between the results of the direct calculation (solid dots connected by solid lines) and the EMA predictions (dotted lines) for the size dependence of the energy of the highest (non-ZCS) valence-band state and the lowest conduction-band state for Si(001).

nodal (N) planes at the boundaries, (N,N), whereas in the other case ($n=4$), it reaches extrema [E = either a peak or a minimum] at the boundaries, (E,E). The (E,E) and (N,N) combinations correspond to sine- and cosine-type envelopes, respectively. For odd-layer films, both (E,E) and (N,N) are, however, symmetry forbidden. Instead, we have an (E,N) and an (N,E) combination. These imply that, for even-layer films, the quantized k_z^* is given by $(k_z^*)^{\text{even}} = \frac{\pi}{L}j$, whereas for odd-layer films, $(k_z^*)^{\text{odd}} = (k_z^*)^{\text{even}} - \frac{2\pi}{L}$. Hence, the energy pattern $\epsilon_{n=3,4}^{\text{bulk}}(k_z^*)$ changes from even to odd.

This pattern change can be phrased in a slightly different way: since the spurious states to be removed (Sec. III E) are always at the energy minima, one can postulate a band dependent $k_z'^*$, opposite to what has been done so far, with its minimum always tied to the energy minimum, same as in the EMA. Using this notion, the quantized states in Figs. 8 and 9 for a 12-layer (001) film correspond to $j' = 5$, not 1, with $k_z^* = k^{(X)} - k_z'^* = \frac{2\pi}{a}(\frac{2}{12})$. A prime for j and k_z^* is used here to indicate that they do not start from Γ. For the $N = 11$ odd-layer film shown in Fig. 6(b), the smallest k_z^* for the $n = 3$ and 4 bands is $\frac{2\pi}{a}(\frac{1}{11})$ given by $j' = 5$, instead of the familar form $\frac{2\pi}{a}(\frac{2}{11})$. This leads to the same energy pattern change discussed above. Of course, since this pattern change (= energy oscillation) depends on the existence of a unique (even or odd) N, they are expected to exist only in ideally abrupt films.

D. What can one expect from approaches other than the EPM?

Our forgoing discussion illustrated the results using the EPM. We now discuss how the results could change if other band structure representations are used.

(i) *Tight binding scheme.* The above discussion shows that in order to obtain the correct solutions for boundary problems, it is essential to allow the wave functions to adjust variationally to the imposed boundary conditions. In view of this, the tight binding scheme, in which the degree of freedom of the wave functions is limited to that of the orbital coefficients, may lead to results that differ from those obtained when $\psi(\mathbf{r})$ is variational at each $\mathbf{r}_i$.

(ii) *LDA calculation.* (a) With an ideal film surface, an LDA calculation produces results that closely resemble those of the EPM. Figure 15 shows the xy-planar averaged wave functions squared for the $f = 20$ (cosine-type), $f = 21$ (sine-type), and $f = 22$ (cosine-type ZCS) states for the 12-layer Si(001) film as obtained by the EPM and self-consistent LDA calculations. While the wave functions are very much the same, the ZCS state no longer has the constant envelope function in the LDA calculation and its energy hence shows some size-dependence. The reason for this is that the self-consistent procedure in the LDA approach modifies the surface potentials such that in the surface regions they deviate from a superposition of atomic potentials [see dashed line in Fig. 4(a)]. In essence, the energy level of the ZCS state can go either up or down depending on the details of the surface potentials. However, these variations are weaker than what

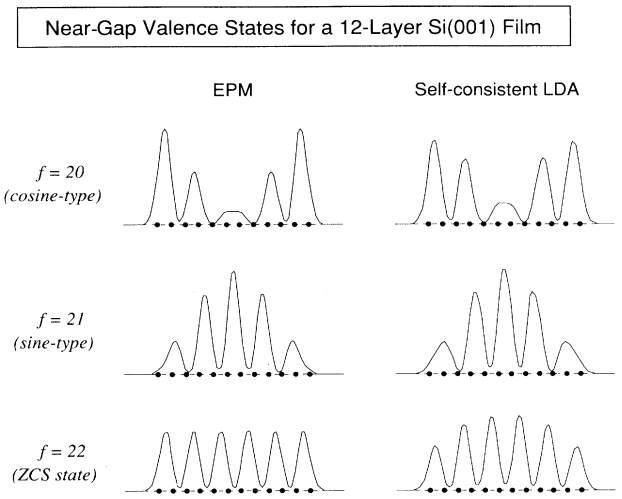

FIG. 15. xy-planar averaged wave functions squared given by the EPM and self-consistent LDA calculations. The solid dots are the atomic planes of Si.

is expected from quantum confinement ($1/L^2$). (b) With surface reconstructions or chemisorptions. In such cases, we expect that the single basis function TC approach may not suffice to describe all the film states. Mixing between different TC basis functions that are close in energy needs to be considered. Even in such cases, the single basis function TC approach provides valuable insights to the problem as it classifies the unperturbed bulklike states into two subgroups: (1) sinelike states with low amplitude at the surface which are, therefore, surface insensitive, and (2) cosinelike states with large amplitudes on the surface which are more sensitive to surface changes. This division justifies the use of the EMA approach for quantum films, when the sinelike states happen to be dominant near the band gap.

V. MATRIX ELEMENTS AND OPTICAL PROPERTIES

In this section, we discuss optical properties of Si(001) films obtained from direct calculations. We show that the TC approach provides a way of systematizing the trends in the optical data which otherwise seem random. The transitions at the $\bar{\Gamma}$ point of the 2D Brillouin zone can be classified as either direct or pseudodirect. Generally speaking, these transitions do not obey the bulk $\mathbf{k}$ conservation rule and thus pseudodirect transitions may have nonzero matrix elements. By mapping the directly calculated film states onto the the TC states, two features are observed: (i) A dipole matrix element between an initial state $(n,j_1) = (n,i)$ and a final state $(m,j_2) = (m,f)$ (where n and m index the bands and j_i indexes k_z^*) is nonzero only for a subset of (i,f). Explicitly, it is zero either for all even $(i-f)$ or for all odd $(i-f)$ for a given n,m. (ii) Direct transition in $\mathbf{k}$-space may have vanishingly small matrix elements. However, pseudodirect transitions that occur near the energy of a forbidden

direct transition can be as strong as allowed direct transitions. These points are illustrated in what follows.

A. Intraband matrix elements

The dipole matrix element between the states (n,i) and (m,f) is given by

$$M[(n,i);(m,f)] = \langle \psi_{n,i}|\mathbf{p}|\psi_{m,f}\rangle = -i\int \psi^*_{n,i}(\mathbf{r})\nabla\psi_{m,f}(\mathbf{r})d^3r \quad . \tag{27}$$

Here, $\mathbf{p} = -i\nabla$ is a momentum operator, which is odd with respect to an inversion ($\mathbf{r} \rightarrow -\mathbf{r}$) operation; its components (p_x, p_y, p_z) are odd with respect to mirror reflections ($x \rightarrow -x$; $y \rightarrow -y$; $z \rightarrow -z$).

It is instructive to first examine Eq. (27) using a single Fourier component approximation to the film's wave functions. This is equivalent to examining matrix elements of an empty film. Let us use m_G to index the reciprocal lattice vector $\mathbf{G}$. Thus, the empty film's wave function for $n = 1$ (or $m_G = 0$) is

$$\psi_{0,j}(z) = \sin\left(\frac{j\pi z}{L}\right) \quad , \tag{28}$$

where the quantum number j [= either i or f in Eq. (27)] indexes $k_z^* = \frac{\pi}{L}j$. Inserting Eq. (28) into Eq. (27) leads to the intraband matrix element for the empty film model ($n = m = 1$)

$$M[(1,i);(1,f)] = \frac{1-(-1)^{i-f}}{L}\frac{2if}{i^2-f^2} \quad . \tag{29}$$

This equation reveals that $M[(1,i);(1,f)]$ vanishes when $(i-f)$ is even. The reason for this is that each $\psi_{0,j}(z)$ has a definite parity with respect to a mirror reflection ($z \rightarrow -z$) at the center of the film: even parity for j = odd and odd parity for j = even. The nonzero matrix elements, according to Eq. (29), are functions of both i and f. For a fixed $f = 1$, the matrix element squared $M^*M = |M|^2$ is proportional to (i) $1/L^2$ if $i \sim f = 1$; but (ii) to $1/L^4$ if $i \sim j_L >> f = 1$ because j_L is also proportional to L.

Figure 16 shows the intraband matrix elements squared $|M|^2$ for the $n = 1$ and 2 valence bands of the 12-layer Si(001) film in Fig. 6(a). The indexes i and f here for the initial and final states correspond to the quantum number j in Fig. 6(a). We show results of the empty film model [Fig. 16(a)] and of direct calculations [Fig. 16(b)]. (Of course, since matrix elements between occupied valence-band states are considered here, these do not correspond to actual optical transitions. We use Fig. 16 only to discuss the nature of the dipolar $\mathbf{p}$ coupling.) This figure is presented using the extended zone scheme in which the $n = 2$ band is part of the $n = 1$ band with $j = 7, 8, 9, \ldots, 12$. We see from Fig. 16 that $|M|^2$ decreases as i increases or as f decreases. The agreement between the direct calculation [Fig. 16(b)] and the empty film model [Fig. 16(a)] is reasonable. In general, however, the intraband matrix elements squared depend not only on i and f but also on band indexes n and m.

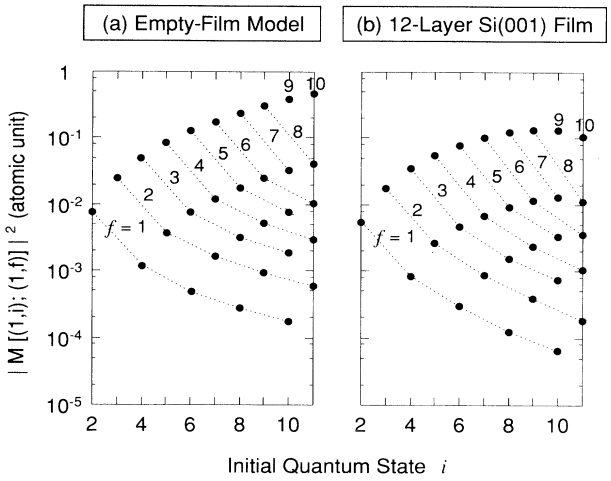

FIG. 16. The nonzero intraband matrix elements squared between state i and f calculated from (a) an empty film model and (b) a 12-layer Si(001) film with $n = 1$ and 2.

B. Interband transitions

Turning next to the actual optical interband transitions, Fig. 17 shows the directly calculated $|M[(n,i);(m,f)]|^2$ for transitions from an initial state i in the $n = 5$ and 6 conduction bands to the final $f = 1$ states in the $m = 3$ and 4 valence bands for a 12-layer Si(001) film. For transitions between $(n,i)\rightarrow(m,f)$ = (5,i)→(3,1) and (6,i)→(4,1), $|M|^2$ vanishes at even $(i-f)$, the same as the empty film model, Eq. (29). However, since there are two types of film states (sine- and cosine-type in Tables II and III), $|M|^2$ may also vanish at odd $(i-f)$. This is seen in Fig. 17 for transitions between bands of different types, i.e., (6,i)→(3,1) (sine- to cosine-type) and (5,i)→(4,1) (cosine- to sine-type). All transitions in Fig. 17 are polarized in the xy-plane.

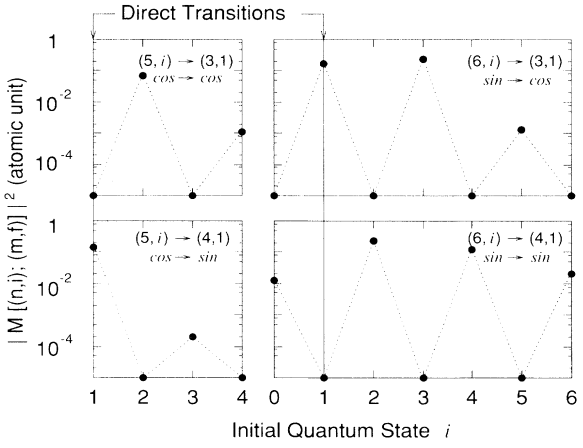

FIG. 17. The directly calculated matrix elements squared for interband transitions in a 12-layer Si(001) film. These are transitions from the two lowest conduction bands $n = 5$ and 6 with initial state index i to the highest (non-ZCS) $n = 3$ and 4 valence-band states with final state index $f = 1$ (see Fig. 6). In the plot, $i = 1$ corresponds to direct bulk transitions.

The patterns in Fig. 17 are the consequence of parities of the film wave functions under inversion. The final valence-band states $m = 4$, $f = 1$ and $m = 3$, $f = 1$ have an even and an odd parity, respectively, whereas the parities of the initial conduction-band states are determined by its band index n and the momentum quantum number i. That is, $(-1)^i$ for state i in the $n = 5$ band and $(-1)^{i+1}$ if the state is in the $n = 6$ band (the minus sign here indicates an odd parity). When the overall parity, tabulated in Table IV, of the integrand in Eq. (27) is odd, $M[(n,i);(m,1)]$ is zero.

One notices from Fig. 17 that certain pseudodirect transitions may have $|M(n,i;m,1)|^2$ comparable to the direct transitions ($i = f = 1$). Two factors are responsible for this: (i) The transition may actually be direct, since a directly calculated conduction-band film state (although shown in Fig. 6 as a point) generally involves two TC basis functions (see Table III). One example of this is the (6,3) state which has a large projection onto the (5,1) TC basis function. Transition (5,1)→(3,1) in Fig. 17 is not pseudodirect but direct, as is the (6,3)→(3,1) transition. (ii) The spectral weight of the direct transition ($i = f = 1$) is shifted to the neighboring pseudodirect transitions ($i = f \pm 1 = 0,2$) when the direct transition is dipole forbidden. Examples of the forbidden direct transitions in Fig. 17 are the (5,1)→(3,1) (cosine- to cosine-type) and (6,1)→(4,1) (sine- to sine-type) transitions. *This implies that for Si films there will be a strong pseudodirect transition at an energy below the direct transition of ~3.4 eV.*

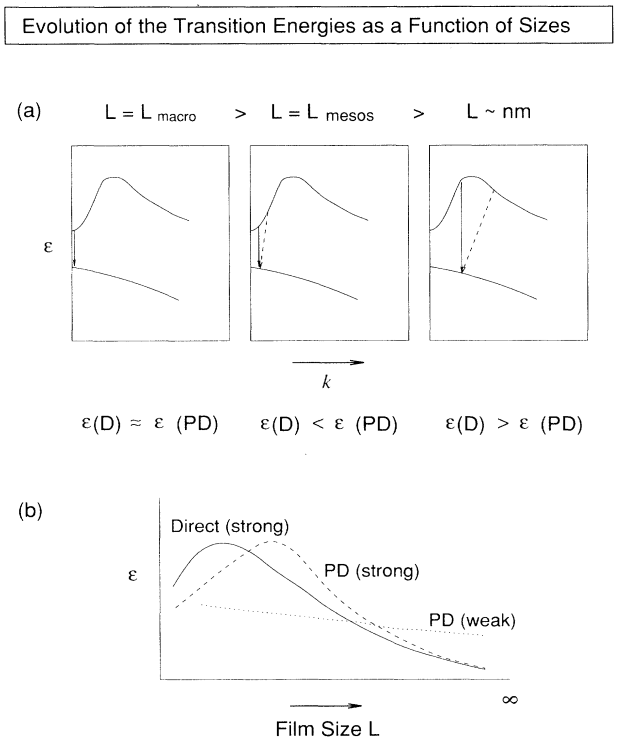

FIG. 18. (a) Schematic drawings showing that a degenerate bulk direct transition splits into a direct and a strong pseudodirect transition in quantum films. (b) Band gaps as a function of size for the band structure shown in part (a).

TABLE IV. Parities for $\psi^*_{n,i}(\mathbf{r})\nabla\psi_{m,f}(\mathbf{r})$ for interband transitions. The parities of the initial and final states are given in square brackets. The plus (+) and minus (−) stand for even and odd parity, respectively.

	(5,i) [$(-1)^i$]	(6,i) [$(-1)^{i+1}$]
(3,1) [−]	$(-1)^i$	$(-1)^{i+1}$
(4,1) [+]	$(-1)^{i+1}$	$(-1)^i$

Now, with the knowledge about matrix elements, we have a more complete picture about the deconfinement effect, i.e., the increase and then decrease of the direct band gap, resulting from film size reduction. In addition to the discussion in Sec. III C, the finite size splits the direct bulk transitions (across the gap) into a direct and a pseudodirect branch, as shown in Fig. 18(a). For large enough sizes, the pseudodirect transition energy is always larger than that of the direct transition. But when the size is reduced to nanoscale, the pseudodirect transition will have a lower transition energy. Figure 18(b) shows schematically the evolution of (i) the direct, (ii) the strong pseudodirect (originated from bulk direct transitions), and (iii) low energy but weak pseudodirect transition (originated from the bulk indirect transition).

VI. SUMMARY

Our pseudopotential band structure calculations reveal several features of the semiconductor quantum films that are unexpected on the basis of the effective-mass particle-in-a-box model. These include (i) nonmonotonic size dependence of the band energies with even-odd energy oscillations for the top of the valence-band states, (ii) the presence of states with cosine-type envelope functions, nonvanishing at the boundaries, and (iii) the zero-confinement states pinned at constant energies. The presence of surface boundaries selectively enhances certain pseudodirect optical transitions by several orders of magnitude — a useful feature for optical device design. We explain all these results in terms of a simple truncated crystal approach, which provides a one-to-one mapping between the eigenstates, eigenvalue spectra of a film, and those of the corresponding bulk material at a special subset of the bulk **k** points.

ACKNOWLEDGMENTS

We would like to thank S. Froyen and L. W. Wang for many helpful discussions on the subject. This work was supported by the Office of Energy Research (OER) [Division of Materials Science of the Office of Basic Science (BES)], U.S. Department of Energy, under Contract No. DE-AC02-83-CH10093.

APPENDIX: FILM PSEUDOPOTENTIALS

The local empirical pseudopotential used previously for Si (Ref. 12) and GaAs (Ref. 13) has the reciprocal-space form

$$V(q) = a_1(q^2 - a_2)/(e^{a_3(q^2-a_4)} + 1) \quad , \tag{A1}$$

where $\{a_i\}$ are four parameters derived from fitting the pseudopotential form factors to experiment. This potential, originally designed for bulk materials, does not produce the correct work function (which is a surface property). Based on Eq. (A1), we derive a different analytic potential for Si (cubic lattice parameter: $a = 5.43$ Å),

$$V(q) = b_1(q^2 - b_2)/(b_3 e^{b_4 q^2} - 1) \tag{A2}$$

with $b_1 = 0.53706$, $b_2 = 2.19104$, $b_3 = 2.05716$, and $b_4 = 0.48716$ in rydberg atomic unit. The four parameters $\{b_i\}$ in Eq. (A2) are determined from both the three empirical pseudopotential form factors[14] [used to derive $\{a_i\}$ in Eq. (A1)] and the experimental work function of Si as an additional constraint. The work function (Φ) depends on surface orientation. However, the dependence of the confined states on Φ is rather weak (a change of a few tenths of an eV in the work function affects the energy levels 2 eV below the vacuum level by only ~0.01 eV). Here we used the experimentally determined value $\Phi = 4.9$ eV (Ref. 15) for the Si(110) surface. Table V compares for Si the bulk energy levels, calculated using this potential and a kinetic energy cutoff of 4.5 Ry, with experiment.[16] The results agree to within a few tenths of an eV.

Equation (A2) does not produce a good fit for GaAs. We use instead Eq. (A1) with $a_1 = 1.22$ (0.35), $a_2 = 2.45$ (2.62), $a_3 = 0.54$ (0.93), and $a_4 = -2.71$ (1.57) for Ga (As), respectively, taken from Ref. 13. To get also the correct work function, we have added an external potential which has the real space form

$$V_{\text{ext}}(z) = \delta V \{\text{erf}[\alpha(z_{bl} - z)] + \text{erf}[\alpha(z - z_{bu})]\} \quad , \tag{A3}$$

TABLE V. Comparison between the EPM and experimental (Ref. 16) energy levels for Si and GaAs (in eV).

	Si			GaAs	
	EPM	Expt.		EPM	Expt.
$\Gamma_{25'v}$	0.00	0.0	Γ_{15v}	0.00	0.0
Γ_{15c}	3.24	3.35	Γ_{1c}	1.45	1.63
$\Gamma_{2'c}$	4.11	4.15	Γ_{15c}	4.26	4.72
X_{4v}	-2.98	-2.9	X_{5v}	-2.21	-2.80
X_{1c}	1.28	1.13 (Δ)	X_{1c}	1.95	2.18
			X_{3c}	2.11	2.58
$L_{3'v}$	-1.25	-1.2	L_{3v}	-0.91	-1.3
L_{1c}	2.18	2.04	L_{1c}	1.67	1.85
L_{3c}	4.02	3.9			

where the error function erf(z) with a damping factor $\alpha = 5$ is used to reduce numerical noise, and z_{bl} and z_{bu} are the boundary positions of the GaAs film. Using $\delta V = 0.69$ Ry in (A3) gives a work function $\Phi = 4.9$ eV. Comparison between the results of the TC and direct calculation for Si films suggests that the boundaries (z_{bl} and z_{bu}) for GaAs should be placed half interlayer spacing outside the surface atomic planes. A 4.5 Ry kinetic energy cutoff is used also for GaAs. The calculated energy levels are compared with photoemission and photoelectron spectroscopy data[17,18] in Table V. They are compiled in Ref. 19. The agreement is not as good as for Si. The important energies in these calculations, however, are not the absolute energies but the energy differences between the direct film calculation and the TC approach. The latter uses also the directly calculated wave functions and band structure (for the bulk) as input. The relevant energy difference is thus the difference between results of two direct calculations and is accurate to within a few hundredths of an eV.

[1] E. Corcoran, Sci. Am. **263** (5), 122 (1990).
[2] M. L. Steigerwald and L. Brus, Acc. Chem. Res. **23**, 183 (1990).
[3] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Editions de Physique, Les Ulis, 1988), p. 63.
[4] M. G. Burt, J. Phys. Condens. Matter **4**, 6651 (1992).
[5] J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).
[6] A. Zunger, J. Phys. C **7**, 76 (1974).
[7] R. P. Messmer, Phys. Rev. B **15**, 1811 (1977).
[8] O. Bilek and L. Skala, Czech. J. Phys. B **28**, 1003 (1978).
[9] P. Kadura and L. Kunne, Phys. Status Solidi B **88**, 537 (1978).
[10] R. A. Smith, *Wave Mechanics of Crystalline Solids* (Wiley, New York, 1961), p. 62.
[11] M. V. Rama Krishna and R. A. Friesner, Phys. Rev. Lett. **67**, 629 (1991); J. Chem. Phys. **96**, 873 (1992).
[12] M. Schluter, J. R. Chelikowsky, S. G. Louie, and M. L. Cohen, Phys. Rev. B **12**, 4200 (1975).
[13] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **20**, 4150 (1979).
[14] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976).
[15] F. G. Allen, J. Phys. Chem. Solids **8**, 119 (1959).
[16] D. R. Masovic, F. R. Vukajlovic, and S. Zekovic, J. Phys. C **16**, 6731 (1983).
[17] T. C. Chiang, J. A. Knapp, M. Aano, and D. E. Eastman, Phys. Rev. B **21**, 3513 (1980).
[18] D. Straub, M. Skibowski, and F. J. Himpsel, Phys. Rev. B **32**, 5237 (1985).
[19] *Semiconductors: Group IV Elements and III-V Compounds*, edited by O. Madelung (Springer-Verlag, Berlin, 1991).
[20] S. B. Zhang, M. L. Cohen, S. G. Louie, D. Tomanek, and M. S. Hybertsen, Phys. Rev. B **41**, 10 058 (1990).
[21] R. G. Dandrea, C. Duke, and A. Zunger, J. Vac. Sci. Technol. B **10**, 1744 (1992).
[22] The reason that the valence-band energies for (110) oriented films have a weaker overall N dependence relative to (001) oriented films is their heavier hole effective mass: neglecting spin orbit, the effective masses obtained in our empirical pseudopotential calculations are $m_h^*(110) = 2.42m_0$ and $3.26m_0$ for Si and GaAs, respectively, whereas $m_h^*(001) = 0.28m_0$ for Si.
[23] E. Kaxiras, Phys. Rev. Lett. **64**, 551 (1990).
[24] L. W. Wang and A. Zunger (unpublished).
[25] S. H. Wei and A. Zunger, J. Appl. Phys. **63**, 5794 (1988).
[26] L. J. Sham and Y. T. Lu, J. Lumin. **44**, 207 (1989).
[27] S. Froyen, D. M. Wood, and A. Zunger, Phys. Rev. B **36**, 4547 (1987).