Valence-band mixing in first-principles envelope-function theory

Bradley A. Foreman[*]

Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
(Received 16 January 2007; published 23 July 2007)

This paper presents a numerical implementation of a first-principles envelope-function theory derived recently by the author [B. A. Foreman, Phys. Rev. B **72**, 165345 (2005)]. The examples studied deal with the valence subband structure of GaAs/AlAs, GaAs/Al$_{0.2}$Ga$_{0.8}$As, and In$_{0.53}$Ga$_{0.47}$As/InP (001) superlattices calculated using the local-density approximation to density-functional theory and norm-conserving pseudopotentials without spin-orbit coupling. The heterostructure Hamiltonian is approximated using quadratic-response theory, with the heterostructure treated as a perturbation of a bulk reference crystal. The valence subband structure is reproduced accurately over a wide energy range by a multiband envelope-function Hamiltonian with linear renormalization of the momentum and mass parameters. Good results are also obtained over a more limited energy range from a single-band model with quadratic renormalization. The effective kinetic-energy operator ordering derived here is more complicated than in many previous studies, consisting in general of a linear combination of all possible operator orderings. In some cases, the valence-band Rashba coupling differs significantly from the bulk magnetic Luttinger parameter. The splitting of the quasidegenerate ground state of no-common-atom superlattices has non-negligible contributions from both short-range interface mixing and long-range dipole terms in the quadratic density response.

PACS number(s): 73.21.−b, 73.61.Ey, 71.15.Ap

I. INTRODUCTION

In a recent paper, the author has developed a first-principles multiband envelope-function theory for semiconductor heterostructures.[1] This theory can, in principle, provide accurate numerical predictions of envelope-function Hamiltonian parameters, but only if a reliable quasiparticle self-energy is used as input. On the other hand, if the input potential is taken from a simple density-functional calculation,[2] the numerical values are less accurate, but the theory can still provide deep insight into the basic physics of the interface and clarify various limitations of commonly used envelope-function models. The purpose of the present paper is to explore these topics using numerical examples taken from density-functional calculations of the valence subband structure in semiconductor superlattices.

In order to obtain the most direct comparison with envelope-function theory as it is commonly used, the Hamiltonian was derived in Ref. 1 in the form of a matrix containing differential operators and energy-independent functions of position. Using a differential equation of low order to describe an abrupt heterojunction might seem at first glance to be a gross theoretical blunder, since it is well known that bulk effective-mass theory[3] is valid only for slowly varying perturbing potentials. Yet effective-mass theory is widely accepted as a valid lowest-order approximation for the shallow impurity problem,[4–6] even though the atomic impurity potential is not slowly varying. Its validity in this context is established by the existence of rigorous methods for extending effective-mass theory to include higher-order perturbations,[5–7] including the rapidly varying part of the impurity potential.[8] These produce short-range and long-range corrections to the effective Hamiltonian (each with its own independent parameters), which lead to the well known chemical shift and splitting of effective-mass degeneracies,[5,6,8] thereby bringing the rough estimates of

effective-mass theory into much closer agreement with experiment.

A heterostructure is nothing but an assembly of atoms. It is therefore difficult to imagine a valid argument for accepting the use of generalized effective-mass differential equations in the shallow impurity problem, yet categorically rejecting them for small atomic perturbations in heterostructures. According to linear-response theory, the perturbation generated by a collection of small atomic perturbations is to leading order just the superposition of the individual perturbations. The rapidly varying part of the heterostructure potential[9,10] can thus be handled by the same techniques used in the impurity problem.[8] The validity of low-order differential equations for shallow heterostructures—as for shallow impurities—consequently rests on two fundamental assumptions:[3,8] that the envelope functions satisfying these equations are slowly varying[9] (i.e., have a Fourier transform limited to small wave vectors) within the energy range of experimental interest, and that the atomic perturbations are truly "shallow" in real heterostructures.

The first assumption is confirmed by numerical examples in later sections of this paper, in agreement with prior empirical pseudopotential studies.[11–14] For the second assumption, the definition of shallow is a relative one that depends on the energy separation between the states included explicitly in the envelope-function model and the remote states treated as perturbations [see, e.g., Eq. (II.33) of Ref. 3]. Consider the case of a GaAs/AlAs heterostructure, where the valence-band offset is about 0.5 eV and the energy gap is about 1.5 eV. Treating GaAs/AlAs as a shallow perturbation of the virtual crystal Al$_{0.5}$Ga$_{0.5}$As would be expected to yield marginal results (at best) in a single-band effective-mass model for the degenerate Γ valence states. Such a model would be expected to give good results only for weaker perturbations, such as those in a GaAs/Al$_{0.2}$Ga$_{0.8}$As heterostructure (within the virtual-crystal approximation). On the other hand, treating GaAs/AlAs as a shallow perturbation would

be expected to produce good results in a multiband envelope-function model that includes a few of the nearest conduction-band states explicitly, since the remote bands are then more than 5 eV away from the valence-band maximum. (This method of making moderately "deep" perturbations shallow by including more bands in the Hamiltonian was proposed by Keldysh for the deep impurity problem in bulk semiconductors.[15]) All of the above expectations are confirmed below by numerical examples for GaAs/$Al_{0.2}Ga_{0.8}$As, GaAs/AlAs, and $In_{0.53}Ga_{0.47}$As/InP.

To be more specific, the multiband envelope-function Hamiltonian of Ref. 1 was derived by treating the heterostructure as a perturbation (within the pseudopotential approximation) of a virtual bulk reference crystal. The self-consistent potential energy of the heterostructure was approximated using the linear and quadratic terms of nonlinear-response theory (see the preceding paper[2] for further details). A finite-order envelope-function Hamiltonian was constructed by using Luttinger-Kohn perturbation theory[3,8,16,17] to eliminate the $\mathbf{k} \cdot \mathbf{p}$ and potential-energy coupling to remote bands, working to order k^4 in the bulk reference Hamiltonian,[10] to order k^2 in the linear-response terms,[16,17] and to order k^0 in the quadratic response.[10] This theory is shown here to work well in a 3-state model for the Γ_{15v} valence states (i.e., neglecting spin-orbit coupling) of a GaAs/$Al_{0.2}Ga_{0.8}$As superlattice and in a 7-state $\{\Gamma_{15c}, \Gamma_{1c}, \Gamma_{15v}\}$ model for GaAs/AlAs and $In_{0.53}Ga_{0.47}$As/InP superlattices. Examples illustrating the success of a 4-state $\{\Gamma_{1c}, \Gamma_{15v}\}$ model for $In_{0.53}Ga_{0.47}$As/InP superlattices have been given elsewhere.[18]

To test the limits of the single-band (Γ_{15v}) model, this paper also extends the theory of Ref. 1 to include terms of order $\theta^2 k$ and $\theta^2 k^2$, where θ denotes the heterostructure perturbation. This is shown to yield much better predictions of the position-dependent effective masses of both GaAs/AlAs and $In_{0.53}Ga_{0.47}$As/InP than the $O(\theta^1 k^2)$ theory of Ref. 1. The resulting superlattice band structures are also of reasonably good accuracy, but over a more limited energy range than the multiband models. The principal limitation of the single-band theory seems to be the presence of spurious solutions generated by $O(\theta^0 k^4)$ terms in the bulk reference Hamiltonian.

The extended $O(\theta^2 k^2)$ theory also yields some interesting conclusions regarding operator ordering in the effective kinetic-energy operator T in effective-mass theory. Much previous work has examined various possible ways of choosing the exponents α, β, and γ in the von Roos parametrization[19]

$$T_{vR} = \frac{1}{4}(m^\alpha p m^\beta p m^\gamma + m^\gamma p m^\beta p m^\alpha), \qquad (1)$$

where p is the momentum operator (in one dimension), m is the effective mass, and $\alpha + \beta + \gamma = -1$. Morrow and Brownstein[20,21] have argued that only exponents satisfying $\alpha = \gamma$ are physically permissible in abrupt heterostructures, which would rule out seemingly reasonable possibilities such as[22] $T = \frac{1}{4}(m^{-1}p^2 + p^2 m^{-1})$.

The present theory is not consistent with any single operator ordering of this type. Instead, the terms of order p^2 derived here take the form of a *linear combination* of terms containing position-dependent functions having *all possible* operator orderings with respect to p. The apparent conflict between this result and the theory of Morrow and Brownstein[20,21] is resolved by the fact that these are smooth functions of position with no discontinuity even at an abrupt junction.

The structure of the paper is as follows. The quadratic-response theory used in the preceding article[2] to simplify the self-consistent pseudopotential is briefly reviewed in Sec. II. Section III describes the construction of the envelope-function Hamiltonian, which is also studied from the perspective of the theory of invariants. In Sec. IV, the parameters of the Hamiltonian are calculated and discussed for the material systems GaAs/AlAs, GaAs/$Al_{0.2}Ga_{0.8}$As, and $In_{0.53}Ga_{0.47}$As/InP. The valence subband structures for (001) superlattices of these materials are calculated in Sec. V, which compares the predictions of various approximate envelope-function models with "exact" numerical calculations. Finally, the results of the paper are discussed and summarized in Sec. VI.

II. QUADRATIC-RESPONSE THEORY

The foundations for this work were laid in the preceding article,[2] in which methods are developed for approximating the self-consistent heterostructure pseudopotential using quadratic-response theory. This section presents a brief summary of the main ideas.

All of the numerical results in this paper are taken from self-consistent total-energy calculations[23,24] performed in a plane-wave basis using the ABINIT software[25–27] with nonlocal norm-conserving pseudopotentials and the local-density approximation (LDA) to density-functional theory. Spin-orbit coupling is neglected. This model was chosen in order to permit direct comparisons of envelope-function calculations with exact numerical calculations in the relatively large superlattices where envelope-function theory is valid. Some technical ingredients of the calculations and the justification for choosing this particular model are discussed further in Ref. 2. As shown there, the model system predicts valence-band offsets in reasonably good agreement with experiment, but (as usual for LDA calculations[24]) does not predict accurate conduction bands.

The envelope-function Hamiltonian in this paper is not calculated directly from the exact superlattice pseudopotential (since that would require a separate self-consistent calculation for each new structure), but is instead obtained from the linear and quadratic responses to virtual-crystal perturbations of a bulk reference crystal.[1,2] The reference crystal is defined as the virtual-crystal average of the bulk constituents (e.g., $Al_{0.5}Ga_{0.5}$As for GaAs/AlAs and $In_{0.765}Ga_{0.235}As_{0.5}P_{0.5}$ for $In_{0.53}Ga_{0.47}$As/InP). The perturbation of the heterostructure relative to the reference crystal is defined by the change in pseudopotential

$$\Delta V_{psp}(\mathbf{x}) = \sum_\alpha \sum_\mathbf{R} \theta_\mathbf{R}^\alpha v_{ion}^\alpha(\mathbf{x} - \mathbf{R}_\alpha), \qquad (2)$$

where $v_{ion}^\alpha(\mathbf{x})$ is the ionic pseudopotential of atom α, $\mathbf{R}_\alpha$ is the position of atom α in unit cell $\mathbf{R}$, and $\theta_\mathbf{R}^\alpha$ is the change in

fractional weight of atom α in cell **R** of the heterostructure relative to the reference crystal.

The fundamental assumption of nonlinear-response theory is that the valence electron density n and the self-consistent pseudopotential can be expressed as power series in the perturbation $\theta_\mathbf{R}^\alpha$; e.g.,

$$n(\mathbf{x}) = n^{(0)}(\mathbf{x}) + n^{(1)}(\mathbf{x}) + n^{(2)}(\mathbf{x}) + \cdots, \tag{3a}$$

$$n^{(1)}(\mathbf{x}) = {\sum_{\alpha,\mathbf{R}}}' \theta_\mathbf{R}^\alpha \Delta n_\mathbf{R}^\alpha(\mathbf{x}), \tag{3b}$$

$$n^{(2)}(\mathbf{x}) = {\sum_{\alpha,\mathbf{R}}}' {\sum_{\alpha',\mathbf{R}'}}' \theta_\mathbf{R}^\alpha \theta_{\mathbf{R}'}^{\alpha'} \Delta n_{\mathbf{R}\mathbf{R}'}^{\alpha\alpha'}(\mathbf{x}). \tag{3c}$$

Here, $n^{(0)}(\mathbf{x})$ is the density of the reference crystal, $n^{(1)}(\mathbf{x})$ is the total linear response, and $n^{(2)}(\mathbf{x})$ is the total quadratic response. The coefficients $\Delta n_\mathbf{R}^\alpha(\mathbf{x})$ and $\Delta n_{\mathbf{R}\mathbf{R}'}^{\alpha\alpha'}(\mathbf{x})$ give the linear and quadratic responses to individual monatomic and diatomic perturbations of the reference crystal. The summations in Eqs. (3b) and (3c) are limited to independent values of α.[1,2]

To find all eigenstates of the Hamiltonian, one must know the electron density $n(\mathbf{k}+\mathbf{G})$ for all values of the crystal momentum **k**. However, the properties of slowly varying envelope functions depend on the density only in a small neighborhood of each bulk reciprocal-lattice vector **G**. Within such a neighborhood, it is convenient to approximate $n(\mathbf{k}+\mathbf{G})$ using a power-series expansion. In a superlattice, the allowed values of $\mathbf{k}+\mathbf{G}$ satisfy $\mathbf{k}=k\hat{\mathbf{z}}$ for small $|\mathbf{k}|$, where $\hat{\mathbf{z}}$ is the direction normal to the interface plane. The power series for the monatomic and diatomic response coefficients in Eqs. (3b) and (3c) therefore have the form

$$\Delta n(\mathbf{k}+\mathbf{G}) = \sum_{l=0}^{\infty} (-ik)^l \Delta n_l(\mathbf{G}). \tag{4}$$

Similar expansions are used for the local part of the ionic pseudopotential and the exchange-correlation potential. In the present work, the power series for the potentials are approximated using $0 \leq l \leq 2$ for the linear potential and $l=0$ for the quadratic potential.[1,2]

The nonlocal pseudopotential is also expanded in a similar power series, retaining terms of order k^4 in the reference crystal and order k^2 in the heterostructure perturbation (which is linear by definition). The methods used to obtain this expansion are described in Appendix A.

III. ENVELOPE-FUNCTION HAMILTONIAN

The **k**-space expansion coefficients for the linear and quadratic density and short-range potentials may now be used to construct an envelope-function Hamiltonian.[1] The entire set of expansion coefficients including the local and nonlocal potentials is subjected to the unitary transformation given in Eq. (4.15) of Ref. 1, which converts all matrices from the plane-wave basis to a Luttinger-Kohn basis of zone-center Bloch functions for the reference crystal. Perturbation theory is then used to reduce these 283×283 matrices (where 283 is

the number of Γ plane waves in the reference-crystal basis[2]) to a much smaller dimension by eliminating the coupling between states in the set $\mathcal{A}$ of interest (e.g., $\mathcal{A}=\{\Gamma_{15v}\}$) and all other states. The Luttinger-Kohn unitary transformation method[3,8,16,17,28] is used in order to obtain energy-independent Hamiltonian parameters. The basic perturbation formulas are summarized in Appendix B.

The renormalized Hamiltonian for set $\mathcal{A}$ in the reference crystal is defined by its matrix elements (in the Luttinger-Kohn basis)

$$\langle n\mathbf{k}|\bar{H}^{(0)}|n'\mathbf{k}'\rangle = (E_n\delta_{nn'} + k_i\pi_{nn'}^i + k_ik_jD_{nn'}^{ij} + k_ik_jk_lC_{nn'}^{ijl}$$
$$+ k_ik_jk_lk_mQ_{nn'}^{ijlm})\delta_{\mathbf{k}\mathbf{k}'}, \tag{5}$$

the coefficients of which are defined in Appendix C. Since the set $\mathcal{A}=\{\Gamma_{15v}\}$ will be the focus of subsequent numerical study, symmetry restrictions on the kinetic momentum matrix $\pi_{nn'}^i$ and the inverse effective-mass tensor $D_{nn'}^{ij}$ are of interest. From the symmetry of the zinc-blende crystal, the matrix π^i must have the form

$$\pi^i = -iR|\epsilon_{ijk}|\{I_jI_k\}, \tag{6}$$

where ϵ_{ijk} is the antisymmetric unit tensor, $\{AB\}=\frac{1}{2}(AB+BA)$, and I_j is the j component of the orbital angular momentum in a basis $\mathcal{A}=\{|X\rangle,|Y\rangle,|Z\rangle\}$ of p-like orbitals.[29] The coefficient $R=-i\pi_{XY}^z$ is real by time-reversal symmetry, and the Hermiticity of the Hamiltonian requires $R=-R=0$. Likewise, the D^{ij} matrix is given by[29]

$$D^{ij} = L\delta_{ij}1 + (M-L)\delta_{ij}I_i^2 - N(1-\delta_{ij})\{I_iI_j\} + \frac{i}{2}K\epsilon_{ijk}I_k, \tag{7}$$

where 1 represents the 3×3 unit matrix and the constants $L=D_{XX}^{xx}$, $M=D_{XX}^{yy}$, $N=D_{XY}^{xy}+D_{XY}^{yx}$, and $K=D_{XY}^{xy}-D_{XY}^{yx}$ are real. Here L, M, and N determine the anisotropic Γ_{15v} effective masses, while K is the effective Landé g factor for the valence band.[29]

The renormalized linear-response Hamiltonian for set $\mathcal{A}$ is given by[1]

$$\langle n\mathbf{k}|\bar{H}^{(1)}|n'\mathbf{k}'\rangle = {\sum_\alpha}' \theta_\alpha(\mathbf{k}-\mathbf{k}')H_{nn'}^\alpha(\mathbf{k},\mathbf{k}'), \tag{8}$$

where $\theta_\alpha(\mathbf{k})$ is the Fourier transform[1] of the atomic distribution function $\theta_\mathbf{R}^\alpha$ and

$$H_{nn'}^\alpha(\mathbf{k},\mathbf{k}') = E_{nn'}^\alpha + \Xi^\alpha\delta_{nn'}\delta_{\mathbf{k}\mathbf{k}'} + k_i\pi_{nn'}^{i\alpha} + \pi_{nn'}^{\alpha i}k_i' + k_ik_jD_{nn'}^{ij\alpha}$$
$$+ k_iD_{nn'}^{i\alpha j}k_j' + D_{nn'}^{\alpha ij}k_i'k_j'. \tag{9}$$

Here, $\Xi^\alpha=4\pi n_2^\alpha(\mathbf{0})$ is the nonanalytic contribution[2,30] from the linear quadrupole moment of the electron density, which merely shifts the mean energy by a constant; the remaining terms are defined in Appendix D. The notation has been modified with respect to Ref. 1 in order to draw attention to similarities with the bulk Hamiltonian (5) and avoid undue proliferation of symbols. It should be noted that the superscripts on the coefficients in Eq. (9) indicate how the coordinate and momentum operators are ordered; for example, in

the coordinate representation, the term proportional to $D_{nn'}^{i\alpha j}$ has the form $D_{nn'}^{i\alpha j}p_i\theta_\alpha(\mathbf{x})p_j$, where $\mathbf{p}$ is the momentum operator.[31]

The symmetry of the coefficients in Eq. (9) is the same as the symmetry of site α in the reference crystal. For a zinc-blende reference crystal, the atomic sites have the same point group T_d as the reference crystal, but in nonsymmorphic crystals (e.g., diamond) the site symmetry is lower than the crystal symmetry.[8] Thus, the linear momentum matrix for $\mathcal{A}=\{\Gamma_{15v}\}$ has the same form as in bulk:

$$\pi^{j\alpha\cdot} = -iR^{\cdot\alpha}|\epsilon_{ijk}|\{I_jI_k\}, \tag{10a}$$

$$\pi^{\alpha i\cdot} = -iR^{\alpha\cdot}|\epsilon_{ijk}|\{I_jI_k\}, \tag{10b}$$

where the superscript dots are just placeholders to indicate where $\mathbf{p}$ is positioned with respect to $\theta_\alpha(\mathbf{x})$. As in the bulk case, $R^{\cdot\alpha}=-i\pi_{XY}^{z\alpha}$ and $R^{\alpha\cdot}=-i\pi_{XY}^{\alpha z}$ are real by time-reversal symmetry; however, for the linear response, Hermiticity [i.e., $\pi_{nn'}^{\alpha i}=(\pi_{n'n}^{i\alpha})^*$] requires only that $R^{\cdot\alpha}=-R^{\alpha\cdot}$. Therefore, unlike the bulk case, the linear R coefficients are not required to vanish. As discussed in Ref. 1 (using a different notation), the constant $R^{\cdot\alpha}$ generates a δ-function-like mixing of the $|X\rangle$ and $|Y\rangle$ valence states at a (001) heterojunction.[32] This mixing is considered in greater detail below.

The linear D tensor likewise has the same form as in bulk material:

$$D^{ij\alpha} = L^{\cdot\cdot\alpha}\delta_{ij}1 + (M^{\cdot\cdot\alpha} - L^{\cdot\cdot\alpha})\delta_{ij}I_i^2 - N^{\cdot\cdot\alpha}(1-\delta_{ij})\{I_iI_j\}$$
$$+ \frac{i}{2}K^{\cdot\cdot\alpha}\epsilon_{ijk}I_k, \tag{11}$$

with similar definitions for $D^{i\alpha j}$ and $D^{\alpha ij}$. All of the coefficients $L^{\cdot\cdot\alpha}$, etc., are real by time-reversal symmetry. Hermiticity of the Hamiltonian requires that $D_{nn'}^{\alpha ij}=(D_{n'n}^{jia})^*$ and $D_{nn'}^{i\alpha j}=(D_{n'n}^{j\alpha i})^*$, which yields constraints of the form $K^{\cdot\cdot\alpha}=K^{\alpha\cdot\cdot}$, etc., but does not require any of the constants in Eq. (11) to be zero.

As discussed in the Introduction, for many heterostructures the energy gap is not very large in comparison to the band offsets, which means that the linear approximation for the momentum and mass terms used in Ref. 1 is not very accurate in a single-band model. In an effort to learn more about the limits of single-band models, the perturbative renormalization of the momentum and mass terms was extended to terms quadratic in θ_α, with the results given in Appendix E. The resulting contributions can be written in the form

$$\langle n\mathbf{k}|\Delta\bar{H}^{(2)}|n'\mathbf{k}'\rangle = \sum_{\alpha,\beta}{}'\sum_{\mathbf{k}''}\theta_\alpha(\mathbf{k}-\mathbf{k}'')\theta_\beta(\mathbf{k}''-\mathbf{k}')$$
$$\times H_{nn'}^{\alpha\beta}(\mathbf{k},\mathbf{k}'',\mathbf{k}'), \tag{12}$$

in which

$$H_{nn'}^{\alpha\beta}(\mathbf{k},\mathbf{k}'',\mathbf{k}') = E_{nn'}^{\alpha\beta} + k_i\pi_{nn'}^{i\alpha\beta} + k_i''\pi_{nn'}^{\alpha i\beta} + \pi_{nn'}^{\alpha\beta i}k_i' + k_ik_jD_{nn'}^{ij\alpha\beta}$$
$$+ k_iD_{nn'}^{i\alpha\beta j}k_j' + D_{nn'}^{\alpha\beta ij}k_i'k_j' + k_ik_j''D_{nn'}^{i\alpha j\beta}$$
$$+ k_i''D_{nn'}^{\alpha ij\beta}k_j'' + D_{nn'}^{\alpha i\beta j}k_i''k_j'. \tag{13}$$

Here, again the superscripts indicate the positioning of the associated operators, with (for example) the term proportional to $\pi_{nn'}^{\alpha i\beta}$ having the form $\pi_{nn'}^{\alpha i\beta}\theta_\alpha(\mathbf{x})p_i\theta_\beta(\mathbf{x})$ in coordinate space. Hermiticity of the Hamiltonian gives constraints such as $D_{nn'}^{\alpha i\beta j}=(D_{n'n}^{j\beta i\alpha})^*$. It should be noted that Eqs. (12) and (13) include only those quadratic contributions arising from perturbative renormalization; the other terms arising from direct multipole expansions of the quadratic response have a different form and are considered in Appendix F.

For the Γ_{15v} Hamiltonian, the coefficients in Eq. (13) also have T_d symmetry, so they are given by obvious generalizations of Eqs. (10) and (11). The Hermiticity constraints on the R coefficients are $R^{\alpha\beta\cdot}=-R^{\cdot\beta\alpha}$ and $R^{\alpha\cdot\beta}=-R^{\beta\cdot\alpha}$ (which implies that $R^{\alpha\cdot\alpha}=0$), and one can also choose these coefficients to satisfy $R^{\alpha\beta\cdot}=R^{\beta\alpha\cdot}$ because θ_α and θ_β commute. Likewise, the D coefficients L, M, N, and K all satisfy constraints of the form $K^{\alpha\beta\cdot\cdot}=K^{\cdot\beta\alpha}$, $K^{\alpha\cdot\beta\cdot}=K^{\cdot\beta\cdot\alpha}$, $K^{\alpha\cdot\cdot\beta}=K^{\beta\cdot\cdot\alpha}$, and $K^{\cdot\alpha\beta\cdot}=K^{\cdot\beta\alpha\cdot}$, and one can choose them to satisfy $K^{\alpha\beta\cdot\cdot}=K^{\beta\alpha\cdot\cdot}$ too.

In a (001) heterostructure, where $\theta_\alpha(\mathbf{x})=\theta_\alpha(z)$, the bulk Hamiltonian matrix elements of the form Lp_z^2 and Mp_z^2 are replaced (to order θ^2) by

$$Lp_z^2 \rightarrow L^{(0)}p_z^2 + L^{\cdot\cdot\alpha}(p_z^2\theta_\alpha + \theta_\alpha p_z^2) + L^{\cdot\alpha\cdot}p_z\theta_\alpha p_z$$
$$+ L^{\cdot\cdot\alpha\beta}(p_z^2\theta_\alpha\theta_\beta + \theta_\alpha\theta_\beta p_z^2) + L^{\cdot\alpha\beta\cdot}p_z\theta_\alpha\theta_\beta p_z$$
$$+ L^{\cdot\alpha\cdot\beta}(p_z\theta_\alpha p_z\theta_\beta + \theta_\beta p_z\theta_\alpha p_z) + L^{\alpha\cdot\cdot\beta}\theta_\alpha p_z^2\theta_\beta, \tag{14}$$

where summation on α and β is implicit. This is a linear combination of all but one of the von Roos operators[19] defined in Eq. (1). If the renormalization were extended to cubic order, we would find also a term

$$L^{\alpha\cdot\beta\cdot\gamma}(\theta_\alpha p_z\theta_\beta p_z\theta_\gamma + \theta_\gamma p_z\theta_\beta p_z\theta_\alpha). \tag{15}$$

Hence, the present derivation from perturbation theory supports not a single operator having the form (1) with fixed exponents, but a linear combination of all possible operator orderings. As mentioned in the Introduction, this does not lead to any mathematical or physical difficulties because θ_α is a smooth function of $\mathbf{x}$. The question of whether any particular terms in the linear combination might happen to be negligible is considered in Sec. IV.

In a similar fashion, bulk matrix elements of the form Np_xp_z are replaced in a (001) heterostructure by

$$Np_xp_z \rightarrow N^{(0)}p_xp_z + (N^{\alpha\cdot} + 2N^{\cdot\alpha})p_x\{p_z,\theta_\alpha\} + (N^{\alpha\beta\cdot} + N^{\{\alpha\cdot\beta\cdot\}}$$
$$+ 2N^{\cdot\alpha\beta})p_x\{p_z,\theta_\alpha\theta_\beta\} + (N^{\alpha\cdot\cdot\beta} + N^{\{\alpha\cdot\beta\cdot\}})p_x\theta_\alpha p_z\theta_\beta, \tag{16}$$

where $N^{\{\alpha\cdot\beta\cdot\}}=\frac{1}{2}(N^{\alpha\cdot\beta\cdot}+N^{\beta\cdot\alpha\cdot})$. Note that the final term proportional to $\theta_\alpha p_z\theta_\beta$ is not found in the usual symmetrization recipe for envelope-function Hamiltonians.[33,34]

In a (001) heterostructure, the contribution of the R terms to the Hamiltonian is

$$H_R = -i2\{I_x I_y\}(R^{\cdot\alpha}[p_z,\theta_\alpha] + R^{\cdot\alpha\beta}[p_z,\theta_\alpha\theta_\beta] + R^{\alpha\cdot\beta}\theta_\alpha[p_z,\theta_\beta]), \quad (17)$$

which mixes the $|X\rangle$ and $|Y\rangle$ valence states because

$$2\{I_x I_y\} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (18)$$

The function $\theta_\alpha(z)$ is a smooth steplike function, hence $i[p_z,\theta_\alpha] = d\theta_\alpha/dz$ is localized at interfaces and has the form of a macroscopic average of the Dirac δ function. The function $\theta_\alpha[p_z,\theta_\beta]$ is also localized, but the associated term in Eq. (17) cannot be written as a simple derivative because $R^{\alpha\cdot\beta} = -R^{\beta\cdot\alpha}$. Mixing of the type (17) has been studied by Ivchenko et al.[32]

The contribution of the K terms in a (001) heterostructure is very similar:

$$H_K = \frac{i}{2}(p_x I_y - p_y I_x)\Big(K^{\cdot\alpha\cdot}[p_z,\theta_\alpha] + (K^{\alpha\beta\cdot} + K^{\{\alpha\cdot\beta\}})[p_z,\theta_\alpha\theta_\beta]$$

$$+ (K^{\alpha\cdot\beta} - K^{\beta\cdot\alpha})\theta_\alpha[p_z,\theta_\beta]\Big). \quad (19)$$

Here, the operator $(p_x I_y - p_y I_x)$ is analogous to the Rashba coupling[17,35,36] $(p_x\sigma_y - p_y\sigma_x)$ in the Γ_6 conduction band, so contributions of the form (19) have been referred to as the valence-band Rashba coupling.[28,37,38] As discussed in Ref. 1, this type of coupling was introduced in Ref. 39 under an approximation that is equivalent to assuming that the Rashba coefficient is the same as the effective Landé factor K in bulk material. This has the advantage of reducing the number of unknown parameters, since K is known from magnetoabsorption experiments (see, e.g., Ref. 40). However, the bulk Landé factor to order θ^2 is given by

$$K = K^{(0)} + \theta_\alpha K^\alpha + \theta_\alpha\theta_\beta K^{\alpha\beta}, \quad (20a)$$

where

$$K^\alpha = K^{\cdot\alpha\cdot} + 2K^{\cdot\cdot\alpha}, \quad (20b)$$

$$K^{\alpha\beta} = K^{\alpha\beta\cdot} + 2K^{\cdot\cdot\alpha\beta} + K^{\cdot\alpha\beta} + K^{\alpha\cdot\beta} + K^{\alpha\cdot\cdot\beta}, \quad (20c)$$

which shows that the Rashba coupling (19) is generally independent of the bulk Landé factor. To linear order, replacing the Rashba coefficient with K would be a good approximation if $2|K^{\cdot\cdot\alpha}| \ll |K^{\cdot\alpha\cdot}|$. As shown in Sec. IV, this is true in some materials but not in others; hence, it cannot be presumed to hold true in general.

IV. EFFECTIVE-MASS PARAMETERS

In this section, the numerical values of the envelope-function parameters calculated for the model system are examined to see whether any general conclusions can be drawn regarding the structure of the Hamiltonian in Sec. III. Values of the Luttinger parameters[29] L, M, N, and K calculated for various bulk materials are listed in Table I. This table shows

TABLE I. Luttinger parameters for several bulk compounds and their virtual-crystal averages.

	GaAs	Al$_{0.5}$Ga$_{0.5}$As	AlAs
L	−5.028	−3.648	−2.863
M	−1.511	−1.326	−1.126
N	−6.146	−4.676	−3.792
K	−2.086	−0.993	−0.514
	In$_{0.53}$Ga$_{0.47}$As	In$_{0.765}$Ga$_{0.235}$As$_{0.5}$P$_{0.5}$	InP
L	−5.596	−4.187	−3.317
M	−1.385	−1.253	−1.119
N	−6.590	−5.071	−4.087
K	−2.797	−1.609	−0.974

that the parameters are of order 1 (in atomic units), and that their dependence on material composition is not linear. For example, the change (relative to the reference crystal) in K is about -1 for GaAs but only $\frac{1}{2}$ for AlAs. It should be noted that the bulk K values reported here include only the contributions from $\mathbf{k}\cdot\mathbf{p}$ renormalization, since the asymmetric terms in the nonlocal pseudopotential[41] cannot be determined by polynomial fitting.[42]

To obtain a measure of the accuracy of the linear and quadratic approximations, the change in effective-mass parameters for various bulk compounds relative to the reference crystal was calculated in these approximations using expressions of the form (20) and compared with the exact changes. The results of these calculations are given in Table II. The top part of the table considers the changes in GaAs and Al$_{0.2}$Ga$_{0.8}$As relative to the reference crystal Al$_{0.1}$Ga$_{0.9}$As. Here, the linear approximation is accurate to better than 10%, while the quadratic approximation is accurate to better than 1%. Since the linear changes are already a small perturbation in this case, the linear approximation for GaAs/Al$_{0.2}$Ga$_{0.8}$As should be adequate for most purposes.

Note that M, which determines the mass of heavy holes along the $\langle 100\rangle$ directions, is much more accurate than L, N, and K. This happens because the coupling of Γ_{15v} to Γ_{1c} does not contribute to M,[43] so the remote bands in M are more remote and the heterostructure perturbation for M is "shallower" than for the other parameters. As discussed in the Introduction, one can achieve a similar effect for L, N, and K by including Γ_{1c} in the set $\mathcal{A}$.[18]

For the case of GaAs/AlAs (with Al$_{0.5}$Ga$_{0.5}$As as the reference crystal), the linear approximation is off by nearly 50% in some cases, while the quadratic approximation is accurate to within 13% for K and to within 8% for the other parameters. The quadratic approximation is more accurate for In$_{0.53}$Ga$_{0.47}$As/InP, with a maximum error of less than 4%. Although these results are not perfect, an accuracy of around 10% in a small perturbation is good enough in many cases, and as shown in Sec. V, the quadratic approximation for $\mathcal{A} = \{\Gamma_{15v}\}$ yields reasonably good band structures over a limited range of energies (although not as good as for a multiband Hamiltonian).

Calculated values of various parameters in the Γ_{15v} linear-response Hamiltonian (9) are listed in Table III. The linear

TABLE II. Difference between Luttinger parameters of bulk crystals and average reference crystal: Comparison of linear and quadratic approximations with exact differences. Error columns give percent relative error.

	Linear		Quadratic		Exact Value	Linear		Quadratic		Exact Value
	Value	Error (%)	Value	Error (%)		Value	Error (%)	Value	Error (%)	
	GaAs					$Al_{0.2}Ga_{0.8}As$				
ΔL	−0.329710	−7.22	−0.353495	−0.526	−0.355363	0.329710	7.21	0.305924	−0.526	0.307542
ΔM	−0.035986	0.99	−0.035632	−0.003	−0.035634	0.035986	−0.98	0.036339	−0.004	0.036341
ΔN	−0.347108	−6.83	−0.370680	−0.501	−0.372548	0.347108	6.75	0.323536	−0.498	0.325154
ΔK	−0.273357	−8.70	−0.297525	−0.623	−0.299389	0.273357	8.99	0.249189	−0.644	0.250803
	GaAs					AlAs				
ΔL	−0.9968	−27.8	−1.2706	−7.93	−1.3800	0.9968	27.0	0.7230	−7.85	0.7846
ΔM	−0.1933	4.5	−0.1855	0.24	−0.1850	0.1933	−3.6	0.2011	0.27	0.2005
ΔN	−1.0919	−25.7	−1.3609	−7.43	−1.4701	1.0919	23.5	0.8229	−6.93	0.8841
ΔK	−0.7004	−35.9	−0.9828	−10.05	−1.0926	0.7004	45.9	0.4179	−12.92	0.4799
	$In_{0.53}Ga_{0.47}As$					InP				
ΔL	−1.1067	−21.5	−1.3650	−3.12	−1.4091	1.1067	27.2	0.8483	−2.53	0.8703
ΔM	−0.1338	2.1	−0.1319	0.63	−0.1311	0.1338	−0.6	0.1358	0.77	0.1347
ΔN	−1.2192	−19.7	−1.4758	−2.84	−1.5189	1.2192	24.0	0.9626	−2.13	0.9835
ΔK	−0.8778	−26.1	−1.1431	−3.77	−1.1879	0.8778	38.1	0.6125	−3.63	0.6355

changes in the bulk Luttinger parameters are determined by constants L^α, M^α, N^α, and K^α of the form (20b). As noted below Eq. (19), the linear contribution to the valence-band Rashba coupling is just $K^{\cdot\alpha\cdot}$.

Many envelope-function calculations in the literature use the BenDaniel-Duke operator ordering[9,33,34,44,45] in which it is assumed that $|L^{\cdot\alpha\cdot}| \gg 2|L^{\cdot\cdot\alpha}|$, $|M^{\cdot\alpha\cdot}| \gg 2|M^{\cdot\cdot\alpha}|$, and $|N^{\cdot\alpha\cdot}|$

TABLE III. Linear parameters in the Γ_{15v} Hamiltonian. Here, RC stands for reference crystal, and the labels light and heavy holes refer to the bulk properties in the $\langle 100 \rangle$ directions.

| RC | | $Al_{0.5}Ga_{0.5}As$ | $In_{0.765}Ga_{0.235}As_{0.5}P_{0.5}$ | |
α		Ga	As	Ga
Light hole	L^α	−1.984	−1.806	−0.847
	$L^{\cdot\alpha\cdot}$	−1.341	−0.586	−1.002
	$L^{\cdot\cdot\alpha}$	−0.321	−0.610	+0.077
Heavy hole	M^α	−0.387	−0.329	+0.130
	$M^{\cdot\alpha\cdot}$	−0.039	−0.109	+0.093
	$M^{\cdot\cdot\alpha}$	−0.174	−0.110	+0.018
k^2 mixing	N^α	−2.181	−2.074	−0.771
	$N^{\cdot\alpha\cdot}$	−1.542	−0.550	−1.085
	$N^{\cdot\cdot\alpha}$	−0.320	−0.762	+0.157
Landé	K^α	−1.399	−1.287	−0.993
Rashba	$K^{\cdot\alpha\cdot}$	−1.372	−0.464	−1.089
	$K^{\cdot\alpha\cdot\cdot}$	−0.013	−0.411	+0.048
δ mixing	$R^{\cdot\alpha}$	−0.028	−0.017	−0.038

$\gg 2|N^{\cdot\cdot\alpha}|$. Inspection of Table III shows that this is perhaps a tolerable approximation in some cases (e.g., light holes in GaAs/AlAs), but it is a bad approximation in others (e.g., heavy holes in GaAs/AlAs). It should be noted that Bastard[45] and Burt[9] both derive the BenDaniel-Duke ordering using variations of Löwdin perturbation theory,[46] which yields energy-dependent mass parameters. This type of perturbation theory cannot be used to draw conclusions about operator ordering in a Hamiltonian with energy-independent parameters, since Luttinger-Kohn perturbation theory is qualitatively different. A detailed comparison of the two theories is outside the scope of this paper, but it will be noted that a direct application of Löwdin perturbation theory (using a power-series expansion to treat the energy dependence of the denominators) to the present first-principles calculations yields values of $L^{\cdot\cdot\alpha}$, $M^{\cdot\cdot\alpha}$, and $N^{\cdot\cdot\alpha}$ that are smaller than those in Table III, but still not generally negligible.

As mentioned in Sec. III, it is also common practice to estimate the Rashba coupling $K^{\cdot\alpha\cdot}$ by the approximation[39] $K^{\cdot\alpha\cdot} \simeq K^\alpha$, which amounts to an extension of the BenDaniel-Duke hypothesis to the antisymmetric terms in the Luttinger Hamiltonian.[29] Table III shows that this is a good approximation for cationic perturbations in GaAs/AlAs and $In_{0.53}Ga_{0.47}As$/InP, but not for anionic perturbations in $In_{0.53}Ga_{0.47}As$/InP. Hence, as noted in Ref. 1, this approximation can only be relied on in general to produce a rough order-of-magnitude estimate of the Rashba parameter $K^{\cdot\alpha\cdot}$.

Numerical values for the quadratic renormalization terms in Eq. (13) are listed in Table IV. The bulk values in this table are defined by expressions of the form (20c). It should be noted that the present calculations on (001) supercells do not provide separate values for the constants $N^{\{\alpha\cdot\beta\cdot\}}$, $N^{\cdot\alpha\beta\cdot}$, and $N^{\alpha\cdot\cdot\beta}$, since these terms always appear together in the

TABLE IV. Quadratic parameters in the Γ_{15v} Hamiltonian. Here, $\bar{N}^{\cdot\alpha\beta\cdot}=N^{\cdot\alpha\beta\cdot}+N^{\{\alpha\cdot\beta\cdot\}}$, $\bar{N}^{\cdot\alpha\cdot\cdot\beta}=N^{\cdot\alpha\cdot\cdot\beta}+N^{\{\alpha\cdot\beta\cdot\}}$, $N^{\{\alpha\cdot\beta\cdot\}}=\frac{1}{2}(N^{\alpha\cdot\beta\cdot}+N^{\beta\cdot\alpha\cdot})$, and $N^{[\alpha\cdot\beta\cdot]}=N^{\alpha\cdot\beta\cdot}-N^{\beta\cdot\alpha\cdot}$, with similar definitions for K.

| RC | $Al_{0.5}Ga_{0.5}As$ | $In_{0.765}Ga_{0.235}As_{0.5}P_{0.5}$ | | |
(α,β)	(Ga,Ga)	(As,As)	(Ga,Ga)	(As,Ga)
$L^{\alpha\beta}$	−1.087	−0.633	−0.668	−0.262
$L^{\cdot\alpha\beta\cdot}$	−0.718	−0.640	−1.118	−0.111
$L^{\cdot\cdot\alpha\beta}$	+0.051	+0.164	+0.190	+0.026
$L^{\alpha\cdot\cdot\beta}$	−0.007	−0.081	−0.002	+0.010
$L^{\alpha\cdot\beta\cdot}$	−0.232	−0.120	+0.036	−0.235
$L^{\cdot\alpha\cdot\beta}$	−0.232	−0.120	+0.036	+0.020
$M^{\alpha\beta}$	+0.0310	+0.0014	−0.0265	+0.0129
$M^{\cdot\alpha\beta\cdot}$	−0.0040	−0.0119	−0.1111	+0.0009
$M^{\cdot\cdot\alpha\beta}$	+0.0301	+0.0209	+0.0463	+0.0065
$M^{\alpha\cdot\cdot\beta}$	−0.0004	−0.0077	+0.0018	+0.0050
$M^{\alpha\cdot\beta\cdot}$	−0.0124	−0.0104	−0.0050	−0.0066
$M^{\cdot\alpha\cdot\beta}$	−0.0124	−0.0104	−0.0050	+0.0007
$N^{\alpha\beta}$	−1.074	−0.638	−0.679	−0.252
$\bar{N}^{\cdot\alpha\beta\cdot}$	−0.965	−0.787	−1.195	−0.225
$N^{\cdot\cdot\alpha\beta}$	+0.069	+0.183	+0.243	+0.035
$\bar{N}^{\alpha\cdot\cdot\beta}$	−0.248	−0.216	+0.031	−0.097
$N^{[\alpha\cdot\beta\cdot]}$	0.0	0.0	0.0	−0.271
$K^{\alpha\beta}$	−1.128	−0.660	−0.629	−0.278
$\bar{K}^{\cdot\alpha\beta\cdot}$	−0.933	−0.746	−0.966	−0.217
$K^{\cdot\cdot\alpha\beta}$	+0.009	+0.133	+0.151	+0.019
$\bar{K}^{\alpha\cdot\cdot\beta}$	−0.213	−0.180	+0.035	−0.099
$K^{[\alpha\cdot\beta\cdot]}$	0.0	0.0	0.0	−0.248
$R^{\cdot\alpha\beta\cdot}$	−0.0076	+0.0034	−0.0185	−0.0034
$R^{\alpha\cdot\beta}$	0.0	0.0	0.0	−0.0048

sums $N^{\alpha\beta}+N^{\{\alpha\cdot\beta\cdot\}}$ and $N^{\alpha\cdot\cdot\beta}+N^{\{\alpha\cdot\beta\cdot\}}$ (the same is true for K). Table IV is not discussed here beyond a brief comment that although BenDaniel-Duke operator ordering is not a very good approximation in any case, it is typically better for light holes than for heavy holes, and better for cation perturbations than for anion perturbations. (Of course, since the position-dependent corrections to the bulk value of M are rather small, heavy-hole calculations are also less sensitive to the choice of operator ordering.)

Finally, it should be emphasized that the numbers reported here are not expected to be accurate. They are merely intended to provide a crude picture of some of the qualitative features that would be found in a more accurate quasiparticle calculation.

V. SUPERLATTICE VALENCE SUBBAND STRUCTURE

In this section, exact model calculations of the valence subband structure of (001) superlattices are used to evaluate the accuracy of various approximate envelope-function models. As a starting point, the bulk band structure of $Al_{0.5}Ga_{0.5}As$ (used as a reference crystal for GaAs/AlAs) is
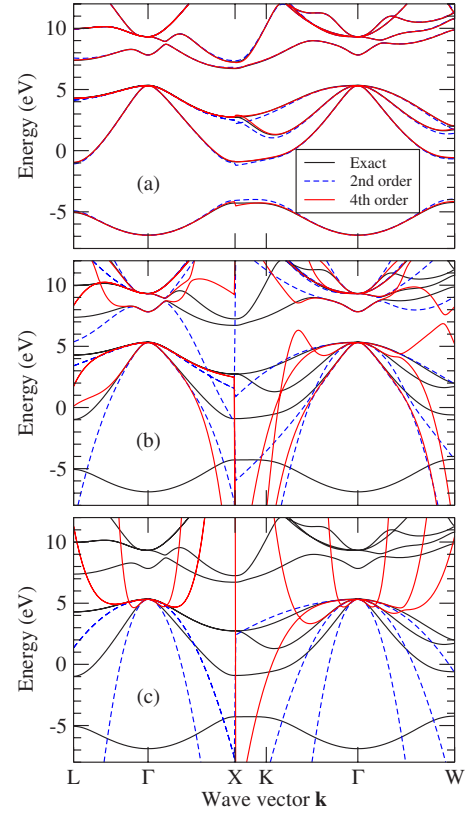

FIG. 1. (Color online) Energy band structure of bulk $Al_{0.5}Ga_{0.5}As$: comparison of exact calculation with $O(k^2)$ and $O(k^4)$ $\mathbf{k}\cdot\mathbf{p}$ models. (a) 283-state $\mathbf{k}\cdot\mathbf{p}$ model, (b) 7-state $\mathbf{k}\cdot\mathbf{p}$ model, and (c) 3-state $\mathbf{k}\cdot\mathbf{p}$ model.

considered in Fig. 1. Part (a) shows the results when all 283 states of the plane-wave basis are retained in the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. Here, it makes little difference whether the polynomials in the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian are terminated at order k^2 or k^4; both cases provide a good description throughout the Brillouin zone. In part (b), the set $\mathcal{A}=\{\Gamma_{15c},\Gamma_{1c},\Gamma_{15v}\}$ contains 7 states (or 14 with spin[47,48]). The description of the band structure is still accurate near Γ, although spurious solutions within the energy gap do occur for both the $O(k^2)$ and $O(k^4)$ $\mathbf{k}\cdot\mathbf{p}$ models.

Part (c) gives the results for $\mathcal{A}=\{\Gamma_{15v}\}$. Here, there are no spurious solutions in the $O(k^2)$ effective-mass model, but the spurious solutions for the $O(k^4)$ model occur at rather small wave vectors. The critical point at which the light-hole band has zero slope is about $\frac{1}{6}$ of the distance to the X point. To prevent problems with spurious solutions, the envelope functions in a (001) superlattice should contain no Fourier components beyond this point.[49] Therefore, a superlattice with a period of 48 monolayers[2] was chosen as the standard test case, since this permits the inclusion of 9 envelope-function plane waves within the region $|k_z|\leqslant\frac{4}{48}(4\pi/a)=\frac{1}{6}(2\pi/a)$. Previous calculations on empirical pseudopotential models show that this is sufficient to achieve reasonably accurate results.[11–14]

The features shown in Fig. 1(c) have a direct analog in the Dirac equation for relativistic electrons, for which the dispersion relation is
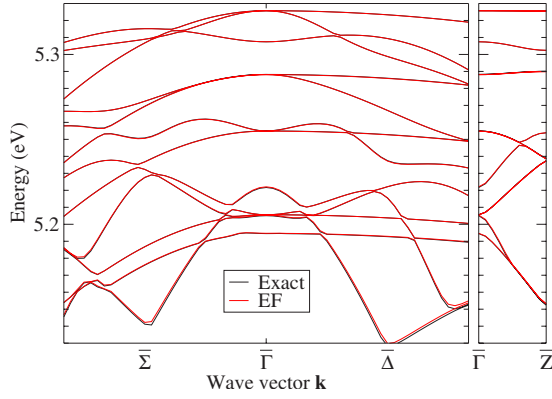
FIG. 2. (Color online) Valence-band structure of a (001) $(GaAs)_{24}(Al_{0.2}Ga_{0.8}As)_{24}$ superlattice: comparison of exact and 3-state envelope-function (EF) calculations. The EF model is the same as that of Ref. 1, and the calculation includes nine plane waves (PWs). Symmetry labels are defined in Refs. 51 and 52; the $\bar{\Sigma}$ axis corresponds to the bulk Δ axis, and the border of the figure in the $\bar{\Sigma}$ direction is 6% of the distance to the bulk X point (the border in the $\bar{\Delta}$ or bulk Σ direction is the same physical distance).

$$E = \sqrt{p^2 c^2 + m^2 c^4} \tag{21a}$$

$$= mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3 c^2} + \cdots. \tag{21b}$$

Here, the power series converges only when $|p| < mc$. If the series is terminated at order p^4, the slope of $E(p)$ changes sign at $p = \sqrt{2}mc$, which lies outside the region of validity of the power series. Although the convergence radius of Luttinger-Kohn perturbation theory for the $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian is not known, one would expect it also to be of the same order of magnitude as the critical point where $E(k)$ changes sign. Therefore, it is reasonable to choose this as the cutoff for plane-wave expansions,[49] even though it may lie slightly outside the convergence radius of the perturbation power series.

The GaAs/$Al_{0.2}Ga_{0.8}As$ material system historically provided one of the first direct comparisons between experiment and effective-mass theory in heterostructures.[50] Figure 2 shows the top 12 valence subbands in a (001) $(GaAs)_{24}(Al_{0.2}Ga_{0.8}As)_{24}$ superlattice. The exact model calculations are compared here with a 3-state Γ_{15v} envelope-function model based on the theory of Ref. 1, in which terms of order k^3 and k^4 are included only for the bulk reference crystal, the mass and momentum terms are linear in θ, and the potential is quadratic in θ. The envelope-function results are in excellent agreement with the exact calculations; the mean error in each of the first ten subbands does not exceed 0.1 meV. Note that the valence-band offset in this system is only 104 meV, which means that the heterostructure perturbation is indeed shallow[4] even in a single-band Γ_{15v} model. These results confirm that the theory of Ref. 1 works very well in the weak-perturbation limit under which it was derived.
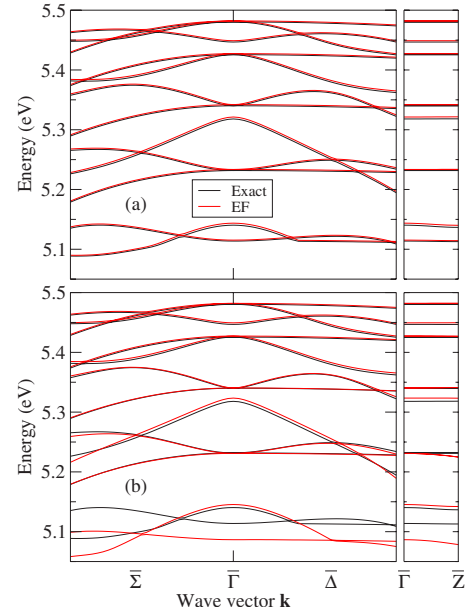

FIG. 3. (Color online) Valence-band structure of a (001) $(GaAs)_{24}(AlAs)_{24}$ superlattice: comparison of exact and 283-state EF calculations. (a) 25 EF plane waves and (b) 9 EF plane waves. Both EF calculations use $O(k^4)$ bulk dispersion.

The following examples provide a more detailed study of the effects of various approximations in systems containing stronger perturbations. Figure 3 shows the top 12 valence subbands in a (001) $(GaAs)_{24}(AlAs)_{24}$ superlattice. The exact model calculations are compared here with a 283-state envelope-function model that includes all zone-center Bloch functions explicitly, corresponding to Fig. 1(a) in the bulk case. Figure 3(a) includes 25 plane waves in the envelope-function model. The results here are very accurate, with an error corresponding to about a 2 meV shift that is nearly the same for all subbands. (To be more precise, the error in the ground state is +1.6 meV, which is almost the same as the +1.5 meV error in the bulk valence-band edge of GaAs calculated in Sec. VII C of the preceding paper.[2]) Figure 3(b) shows the effect of reducing the number of plane waves from 25 to 9. There is little change for energies close to the valence-band maximum, but the error in subband 12 is significant. Note that the energy range here is much wider than that in Fig. 2, although it still corresponds approximately to the valence-band offset.[2]

The effect of reducing set $\mathcal{A}$ to the seven states in Γ_{15c}, Γ_{1c}, and Γ_{15v} is shown in Fig. 4. Part (a) is just the multiband theory of Ref. 1, the single-band version of which was used previously in Fig. 2. The results here are even slightly better than those in Fig. 3(a), which can only be attributed to a fortuitous cancellation of errors. In part (b), the $O(k^4\theta^0)$ terms are dropped. The results are still fairly accurate near the valence-band maximum; however, the peculiar behavior of (what should be) the ground state near $\bar{Z}$ shows that the plane-wave cutoff in this case is not quite sufficient to eliminate all effects of the $O(k^2)$ spurious solutions in Fig. 1(b). In Fig. 4(c), the $O(k^2\theta^1)$ terms are dropped, so that the mass and momentum matrices are approximated by those of the refer-
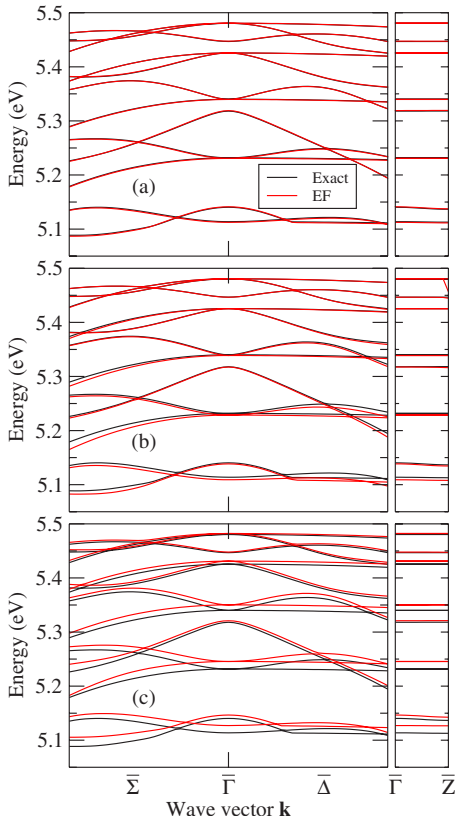
FIG. 4. (Color online) Valence-band structure of a (001) $(GaAs)_{24}(AlAs)_{24}$ superlattice: comparison of exact and 7-state EF calculations. The EF models are (a) $O(k^4\theta^0+k^2\theta^1+k^0\theta^2)$, (b) $O(k^2\theta^1+k^0\theta^2)$, and (c) $O(k^2\theta^0+k^0\theta^2)$. All EF calculations use 25 plane waves and $O(\theta^2)$ potentials.

ence crystal. Here, the error becomes significant even at fairly small energies, which shows the importance of including linear terms for multiband models.

In Figs. 5 and 6, the set $\mathcal{A}$ is reduced even further to only the three Γ_{15v} states. Figure 5 shows the bands on the same energy scale as before, while Fig. 6 shows an expanded view of the region near the band edge. Part (a) uses the linear mass model of Ref. 1 (the same as in Fig. 2). A close inspection of the top three subbands in Fig. 6(a) shows evidence of the errors in the linear mass approximation displayed in Table II. These errors are corrected in part (b), which includes quadratic corrections to the mass as well as (see Appendix E) terms of order θ^4 in the potential. This yields a noticeable improvement, although the quantitative failure for higher excitations (due in large part to the use of only nine plane waves) is still present. In part (c), the $O(k^4\theta^0)$ terms are dropped; since spurious solutions are no longer a problem, the number of plane waves is increased to 25. It can be seen that the top three subbands are still quite accurate under this approximation.

Figure 7 shows the valence subband structure of a (001) $(In_{0.53}Ga_{0.47}As)_{24}(InP)_{24}$ superlattice. Part (a) gives the results obtained from the original 283-state basis. The error here is larger than in the analogous calculation for GaAs/AlAs in Fig. 3(a); the ground-state error is +5.1 meV, which is close to the error of +4.5 meV in the bulk valence-
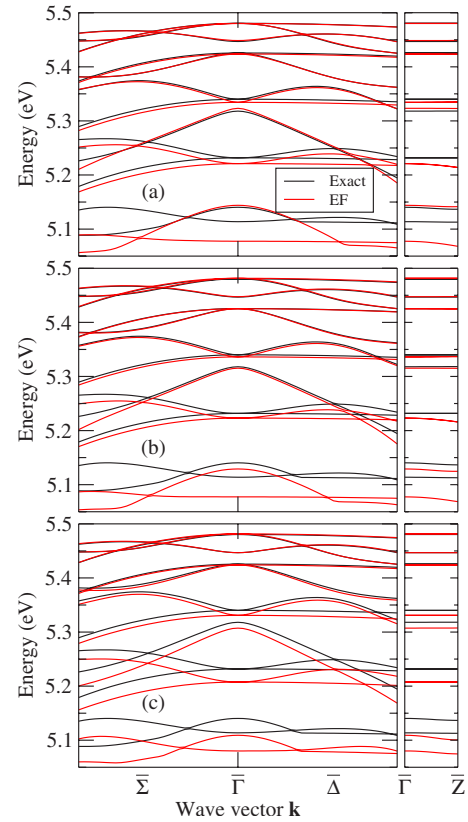

FIG. 5. (Color online) Valence-band structure of a (001) $(GaAs)_{24}(AlAs)_{24}$ superlattice: comparison of exact and 3-state EF calculations. The EF models are (a) $O(k^4\theta^0+k^2\theta^1+k^0\theta^2)$, 9 PWs, (b) $O(k^4\theta^0+k^2\theta^2+k^0\theta^4)$, 9 PWs, and (c) $O(k^2\theta^2+k^0\theta^4)$, 25 PWs.

band edge of $In_{0.53}Ga_{0.47}As$ calculated in Sec. VII C of Ref. 2. The deviation from a constant shift of about 5 meV is negligible for the first ten subbands, which covers the full range of the valence-band offset.[2]

In Fig. 7(b), the basis is reduced to seven states using the linear mass renormalization of Ref. 1. The results are quite close to the exact calculation, although again the improvement over part (a) is fortuitous. This is demonstrated in part (c), which includes additional terms of order $\theta^2 k^2$ and $\theta^4 k^0$. Most of the bands are shifted slightly upward, returning nearly to the result from part (a). The shift is mainly due to $O(\theta^3 k^0)$ corrections in the renormalized potential [see Eq. (E11)]. Hence, neglecting $O(\theta^3 k^0)$ terms in perturbative renormalization [Fig. 7(b)] approximately compensates for the neglect of cubic response terms in the original Hamiltonian.

Figure 2 of Ref. 18 shows the results of calculations that are the same as Fig. 7(b), but for the four-dimensional set $\mathcal{A}=\{\Gamma_{1c},\Gamma_{15v}\}$. The results with $O(\theta^0 k^4)$ terms are almost identical to Fig. 7(b). The effect of dropping the $O(\theta^0 k^4)$ terms is somewhat more significant than in Fig. 4(b), however.

Finally, Fig. 8 shows the predictions of the single-band Γ_{15v} model for $In_{0.53}Ga_{0.47}As/InP$. When $O(k^4)$ terms are included in the bulk Hamiltonian, the critical point of zero slope in the light-hole $\langle 100 \rangle$ dispersion is closer to Γ than it was in GaAs/AlAs, so that now only seven plane waves can
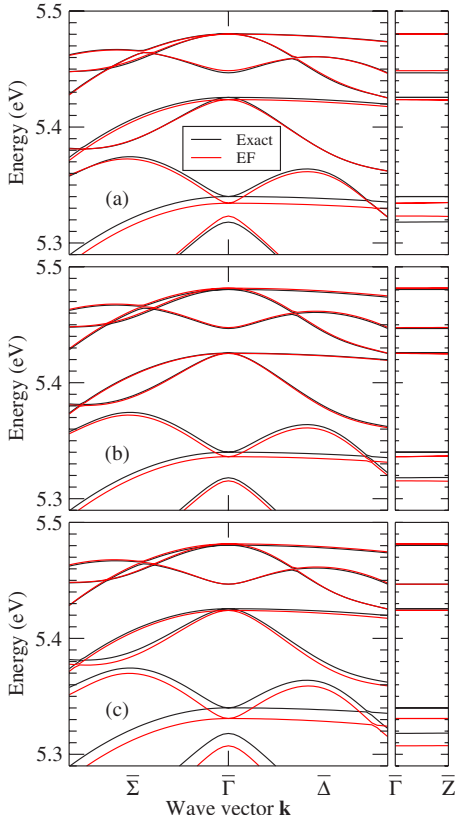
FIG. 6. (Color online) Expanded view of the top eight valence states from Fig. 5.

FIG. 7. (Color online) Valence-band structure of a (001) $(In_{0.53}Ga_{0.47}As)_{24}(InP)_{24}$ superlattice: comparison of exact and EF calculations. (a) 283-state EF model, (b) 7-state EF model, $O(k^4\theta^0+k^2\theta^1+k^0\theta^2)$, and (c) 7-state EF model, $O(k^4\theta^0+k^2\theta^2+k^0\theta^4)$. All EF calculations use 25 plane waves and $O(k^4)$ bulk dispersion.

be included in the envelope functions if one wishes to avoid spurious solutions. Apart from this restriction, the envelope-function models used in parts (a), (b), and (c) of Fig. 8 are the same as those used in the corresponding parts of Figs. 5 and 6. It can be seen that in this case the limitations of using only seven plane waves are sufficiently severe that, for the first seven subbands, one is better off omitting the $O(k^4)$ terms and including more plane waves. It should be noted that the predictions of the single-band Hamiltonian for real $In_{0.53}Ga_{0.47}As/InP$ superlattices would likely be substantially worse than what is shown here (perhaps even qualitatively incorrect), since the energy gap of $In_{0.53}Ga_{0.47}As$ in the model system is 61% larger than the experimental value (see Sec. IV of the preceding paper[2]).

Although it is not visible on the scale of these figures, the double degeneracy of the $\bar{\Gamma}$ ground state in GaAs/AlAs is removed in $In_{0.53}Ga_{0.47}As/InP$ due to the reduction in symmetry from D_{2d} to C_{2v}. The primary cause of the splitting is mixing of the $|X\rangle$ and $|Y\rangle$ valence states due to the short-range interface mixing in Eq. (17) and the long-range interface dipole potential in Fig. 10 of the preceding paper.[2] The splitting of the quasidegenerate ground state calculated exactly and in various envelope-function models is presented in Table V. It can be seen that all of the envelope-function models give a reasonably good estimate of the splitting (which means that they provide a satisfactory description of the microscopic wave function in the interface region). However, when the diatomic dipole terms in Fig. 10 of Ref. 2 (and their associated polarization of the bulk reference crys-
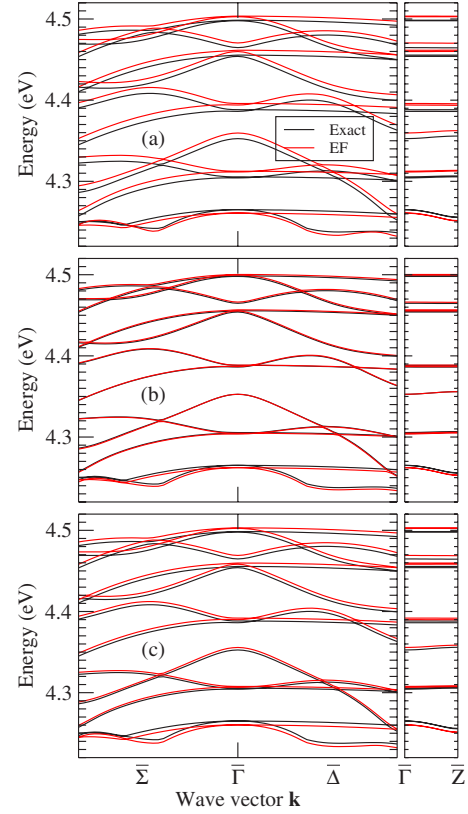
tal) are omitted, the splitting of the ground state in the seven-dimensional envelope-function model is reduced by about one-third. This shows that the practice of fitting experimental splitting data to short-range interface terms only[32,53,54] may give an incorrect description of the basic physics and overestimate the magnitude of the short-range terms.

VI. CONCLUSIONS

This paper has presented a numerical implementation of the first-principles envelope-function theory of Ref. 1 in a model system based on superlattice LDA calculations with norm-conserving pseudopotentials. The electron density and potential energy of the superlattice were approximated by retaining only the linear and quadratic responses to the heterostructure perturbation. This approximation worked well, with a net error of about 2 meV in GaAs/AlAs and 5 meV in $In_{0.53}Ga_{0.47}As/InP$.[2] The principal effect of this error was simply a constant shift of the superlattice energy eigenvalues.

The density and short-range potentials were then approximated further using truncated multipole expansions (i.e., power series in k), retaining terms of order k^2 in the linear potential and k^0 in the quadratic potential. This had no effect on the macroscopic density and potential in bulk, but it gen-
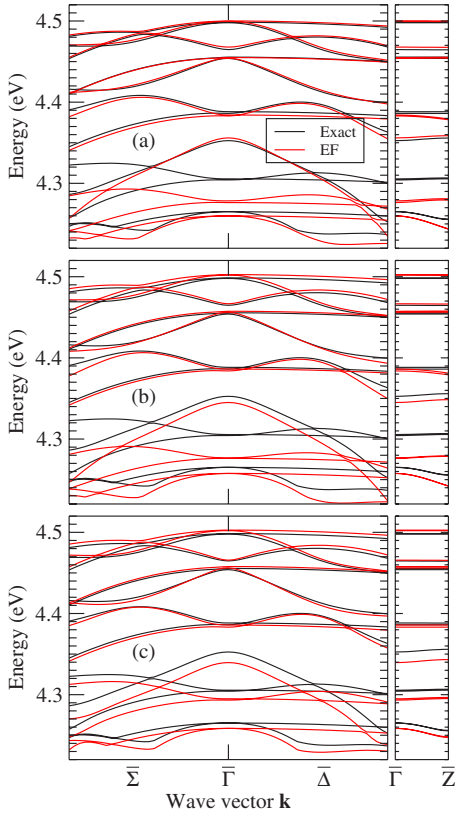
FIG. 8. (Color online) Valence band structure of a (001) $(In_{0.53}Ga_{0.47}As)_{24}(InP)_{24}$ superlattice: comparison of exact and 3-state EF calculations. The EF models are (a) $O(k^4\theta^0+k^2\theta^1+k^0\theta^2)$, 7 PWs, (b) $O(k^4\theta^0+k^2\theta^2+k^0\theta^4)$, 7 PWs, and (c) $O(k^2\theta^2+k^0\theta^4)$, 25 PWs.

erated some additional error (due primarily to the truncation of the linear density response) in a narrow region near the interfaces.[2] This error was confirmed to be negligible for slowly varying envelope functions.

The approximate Hamiltonian was transformed from the original plane-wave basis to a Luttinger-Kohn basis using zone-center Bloch functions of the reference crystal. A Luttinger-Kohn unitary transformation was then used to eliminate the $\mathbf{k} \cdot \mathbf{p}$ and potential-energy coupling between the $\mathcal{A}$ states of interest and the remote $\mathcal{B}$ states. The resulting basis is material dependent (due to the potential-energy terms) and approximates the position dependence of the quasi-Bloch functions in the heterostructure. The perturba-

TABLE V. Splitting of the $\bar{\Gamma}$ ground-state degeneracy in a (001) $(In_{0.53}Ga_{0.47}As)_{24}(InP)_{24}$ superlattice.

	Diatomic dipole included?	
Model	Yes	No
Exact	0.639 meV	
EF (283 states)	0.722 meV	
EF (7 states)	0.626 meV	0.426 meV
EF (3 states)	0.585 meV	

tion theory of Ref. [1] was extended to account for quadratic renormalization of the mass and momentum parameters.

A 7-state $\{\Gamma_{15v}, \Gamma_{1c}, \Gamma_{15c}\}$ envelope-function model with linear momentum and mass renormalization was shown to give a very good description of the Γ_{15} valence subband structure of GaAs/AlAs and $In_{0.53}Ga_{0.47}As/InP$ (001) superlattices, although the good results were partly due to a fortuitous cancellation of errors. Calculations reported elsewhere[18] show that similar results for Γ_{15v} can be obtained from a simpler 4-state $\{\Gamma_{15v}, \Gamma_{1c}\}$ model. A 3-state Γ_{15v} model gave fairly good results over a more limited energy range (although it probably would not work as well in real $In_{0.53}Ga_{0.47}As/InP$ superlattices, since the energy gap of $In_{0.53}Ga_{0.47}As$ in the model system was significantly greater than the experimental value). The primary limitation of this single-band model is a conflict between the need for $O(k^4)$ bulk terms in order to achieve better accuracy in the excited states and the sometimes rather severe plane-wave cutoff needed to avoid spurious solutions generated by the $O(k^4)$ terms. The 3-state model did, however, give excellent results for $GaAs/Al_{0.2}Ga_{0.8}As$, where the band offset is small enough to satisfy Kohn's definition ($\lesssim 0.1$ eV) (Ref. [4]) of a shallow perturbation.

Dipole terms in the quadratic response were found to produce interface asymmetry and macroscopic electric fields in the no-common-atom $In_{0.53}Ga_{0.47}As/InP$ system.[2] These terms, which have C_{2v} symmetry, produce a significant fraction of the splitting of the quasidegenerate ground state in such systems. Fitting this splitting to only short-range interface XY mixing terms may therefore overestimate the short-range terms and omit important physics.

Numerical results for $In_{0.53}Ga_{0.47}As/InP$ and GaAs/AlAs indicate that the linear valence-band Rashba coupling parameter is well approximated by the bulk effective Landé factor K for cationic perturbations, but that there is a wide disparity for anionic perturbations. Therefore, using bulk magnetoabsorption measurements to evaluate interface parameters such as the Rashba coupling[55,56] cannot generally be relied upon to provide anything better than a rough order-of-magnitude estimate. Of course, the particular numbers generated by the present model would likely change significantly in a more realistic quasiparticle calculation, but the discrepancy between the Rashba and Landé coefficients is unlikely to vanish.

The operator ordering of the effective-mass terms at a heterojunction was found to be more complicated than in many previous models. Instead of having a single von Roos kinetic-energy operator of the form shown in Eq. (1), perturbation theory yields a linear combination of terms with all possible operator orderings. Certain terms are larger than others, however. As shown by Leibler,[16,17] to linear order only the BenDaniel-Duke operator[44] $T_{BD} = \frac{1}{2}pm^{-1}p$ and the Gora-Williams operator[22] $T_{GW} = \frac{1}{4}(m^{-1}p^2+p^2m^{-1})$ arise. In a simple model where the matrix $E_{nn'}^{\alpha}$ in Eq. (11) is assumed diagonal, the former arises in third-order perturbation theory from the position-dependent energies of remote bands in set $\mathcal{B}$, whereas the latter comes from the position dependence of the bands in set $\mathcal{A}$ (see Appendix D). The Zhu-Kroemer operator[57] $T_{ZK} = \frac{1}{2}m^{-1/2}p^2m^{-1/2}$ appears as one of several

terms in quadratic renormalization, and the most general von Roos operator does not occur until cubic order.

Actually, with repeated use of the commutator $[p, f(z)] = -i(df/dz)$, one can move the momentum operators into any desired position. For example, Morrow and Brownstein have shown that upon replacing $\alpha \to \alpha + \epsilon$ and $\gamma \to \alpha - \epsilon$ in Eq. (1), the von Roos operator can be rewritten as[21]

$$T_{\mathrm{vR}} = T_{\mathrm{H}} - \frac{1}{2m}\left(\frac{d[\ln(m^{\epsilon})]}{dz}\right)^2, \qquad (22)$$

where $T_{\mathrm{H}} = \frac{1}{2}m^{\alpha} p m^{\beta} p m^{\alpha}$ is the Harrison operator[58] and the second term has the form of a potential energy. However, one can continue this process indefinitely, for example, by writing

$$T_{\mathrm{H}} = T_{\mathrm{BD}} - \frac{d}{dz}\left(\frac{1}{2m}\frac{d[\ln(m^{\alpha})]}{dz}\right) + \frac{1}{2m}\left(\frac{d[\ln(m^{\alpha})]}{dz}\right)^2. \qquad (23)$$

Hence, the operator ordering in the effective kinetic energy is nothing but an arbitrary convention, as long as one takes care to retain all of the effective potential-energy terms generated by changing conventions.[59]

The effective kinetic-energy operator given by the perturbation theory in this paper does have the advantage that the position-dependent functions $\theta_{\alpha}(\mathbf{x})$ appearing in it are smooth steplike functions (although it should be noted that the position of the step is different for cations and anions). One could reduce it to the conventional BenDaniel-Duke form, or any other desired form, if one were willing to deal with functions having a more complicated position dependence near the interface.

This suggests that it may be possible—at least at the phenomenological level—to extend the perturbative scheme described here to arbitrarily high order in θ.[60] From this perspective, as long as (i) the bulk materials of the heterostructure are accurately described by an effective-mass equation and (ii) the heterostructure perturbation series eventually converges at *some* finite order, one can rearrange the operators into some standard order, yielding a standardized Hamiltonian with parameters that have a complicated but, in principle, calculable position dependence.

Of course, this is unlikely to provide a useful first-principles calculation method unless the series converges at a fairly low order. Fortunately, the examples given here demonstrate that linear (in the multiband case) or quadratic (in the single-band case) renormalization of the momentum and mass parameters is sufficient to achieve good results in several typical III-V heterostructures. Given that linear-response theory[61] has produced accurate predictions of valence-band offsets in many other lattice-matched and lattice-mismatched systems (see the bibliography of Ref. 1), it is likely that the present envelope-function theory can be applied successfully in many systems too.

APPENDIX A: NONLOCAL PSEUDOPOTENTIAL

The nonlocal part of the pseudopotential was handled by polynomial fitting in $\mathbf{k}$ space. For the bulk reference crystal, the entire plane-wave Hamiltonian matrix $H(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') \equiv H_{\mathbf{GG}'}(\mathbf{k})$ was evaluated at 57 points near $\mathbf{k}=0$, including $\mathbf{k}=0$ itself and points of the form $\langle 100 \rangle$, $\langle 110 \rangle$, $\langle 111 \rangle$, $\langle 200 \rangle$, and $\langle 210 \rangle \times \Delta$, where Δ is some specified interval (usually set equal to half the magnitude of the smallest superlattice reciprocal-lattice vector). It is important to choose a set of fitting points with O_h symmetry so that the fitted coefficients maintain the T_d symmetry and time-reversal symmetry of the Hamiltonian. The Hamiltonian was fitted to a general polynomial of order k^4 with 35 independent coefficients. (For general $\mathbf{G}$ and $\mathbf{G}'$, there is no special symmetry that can be used to reduce the number of coefficients.)

For the linear response, the nonlocal pseudopotential $V^{\mathrm{nl}}(\mathbf{k}+\mathbf{g}, \mathbf{k}+\mathbf{g}')$ may be written as $V^{\mathrm{nl}}_{\mathbf{GG}'}(k_x, k_y, k_z + \frac{1}{2}(\Delta g_z + \Delta g_z'), \Delta g_z - \Delta g_z')$, in which $\mathbf{g}$ is a reciprocal-lattice vector of the superlattice and $\Delta \mathbf{g} \equiv \mathbf{g} - \mathbf{G}$ is an integer multiple of $(2\pi/Nd)\hat{\mathbf{z}}$, where N is the number of monolayers in the superlattice. The function $V^{\mathrm{nl}}_{\mathbf{GG}'}(k_1, k_2, k_3, k_4)$ was fitted to a general polynomial of order k^2 with 15 independent coefficients using a set of 33 points arranged on a cubic grid. A larger grid was tested using fitting polynomials of order k^4, but the difference was negligible so only the simpler method was used. The nonlocal pseudopotential is purely linear, so no fitting of the quadratic response was necessary.

APPENDIX B: PERTURBATION THEORY

This appendix defines a set of functions that offer a convenient way to describe operator ordering in fourth-order Luttinger-Kohn perturbation theory. These functions are merely an alternative way of writing the expressions given on p. 205 of Winkler's monograph.[28]

In Luttinger-Kohn perturbation theory,[3,8] the total Hamiltonian of the system is written as $H = H_0 + H'$, where H_0 has matrix elements $(H_0)_{mm'} = E_m \delta_{mm'}$. The states of the unperturbed Hamiltonian H_0 are divided into a set $\mathcal{A}$ containing the states of interest and a set $\mathcal{B}$ containing all other states. It is assumed that the energies of $\mathcal{A}$ and $\mathcal{B}$ do not overlap. A unitary transformation $\bar{H} = e^{-S} H e^{S}$ is used to eliminate the coupling between $\mathcal{A}$ and $\mathcal{B}$ to any desired order in the perturbation H'.

The notation used here is defined as follows. $M_{\mathcal{AB}}$ is the block of the matrix M that has rows in set $\mathcal{A}$ and columns in set $\mathcal{B}$. The matrix G is defined by

$$(G_{\mathcal{AB}})_{nn'} = (E_n - E_{n'})^{-1}, \quad G_{\mathcal{BA}} = (G_{\mathcal{AB}})^{\mathrm{T}}, \qquad (B1)$$

where T denotes the transpose. A dot indicates element-by-element multiplication of congruent matrices:

$$(A \cdot B)_{nn'} = A_{nn'} B_{nn'}, \qquad (B2)$$

whereas juxtaposition denotes ordinary matrix multiplication.

The renormalized Hamiltonian $\bar{H}$ for states in $\mathcal{A}$ is given to fourth order in H' by[28]

$$\bar{H}_{\mathcal{A}\mathcal{A}} = H_{\mathcal{A}\mathcal{A}} + P_2(H',H') + P_3(H',H',H')$$
$$+ P_4(H',H',H',H'), \qquad (B3)$$

where the functions P_2, P_3, and P_4 are defined by

$$P_2(H^1,H^2) = \tfrac{1}{2}[(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^2_{\mathcal{B}\mathcal{A}} + H^1_{\mathcal{A}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^2_{\mathcal{B}\mathcal{A}})],$$
$$(B4)$$

$$P_3(H^1,H^2,H^3) = \tfrac{1}{2}\{[(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^2_{\mathcal{B}\mathcal{B}}] \cdot G_{\mathcal{A}\mathcal{B}}\}H^3_{\mathcal{B}\mathcal{A}}$$
$$+ \tfrac{1}{2}H^1_{\mathcal{A}\mathcal{B}}\{G_{\mathcal{B}\mathcal{A}} \cdot [H^2_{\mathcal{B}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^3_{\mathcal{B}\mathcal{A}})]\}$$
$$- \tfrac{1}{2}\{[H^1_{\mathcal{A}\mathcal{A}}(G_{\mathcal{A}\mathcal{B}} \cdot H^2_{\mathcal{A}\mathcal{B}})] \cdot G_{\mathcal{A}\mathcal{B}}\}H^3_{\mathcal{B}\mathcal{A}}$$
$$- \tfrac{1}{2}H^1_{\mathcal{A}\mathcal{B}}\{G_{\mathcal{B}\mathcal{A}} \cdot [(H^2_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})H^3_{\mathcal{A}\mathcal{A}}]\},$$
$$(B5)$$

$$P_4(H^1,H^2,H^3,H^4) = \tfrac{1}{2}[(H^1_{\mathcal{A}\mathcal{A}}\{G_{\mathcal{A}\mathcal{B}} \cdot [H^2_{\mathcal{A}\mathcal{A}}(G_{\mathcal{A}\mathcal{B}} \cdot H^3_{\mathcal{A}\mathcal{B}})]\}) \cdot G_{\mathcal{A}\mathcal{B}}]H^4_{\mathcal{B}\mathcal{A}} + \tfrac{1}{2}H^1_{\mathcal{A}\mathcal{B}}[G_{\mathcal{B}\mathcal{A}} \cdot (\{[(H^2_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})H^3_{\mathcal{A}\mathcal{A}}] \cdot G_{\mathcal{B}\mathcal{A}}\}H^4_{\mathcal{A}\mathcal{A}})]$$
$$- \tfrac{1}{2}H^1_{\mathcal{A}\mathcal{B}}[G_{\mathcal{B}\mathcal{A}} \cdot (\{[H^2_{\mathcal{B}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^3_{\mathcal{B}\mathcal{A}})] \cdot G_{\mathcal{B}\mathcal{A}}\}H^4_{\mathcal{A}\mathcal{A}})] - \tfrac{1}{2}H^1_{\mathcal{A}\mathcal{B}}[G_{\mathcal{B}\mathcal{A}} \cdot (H^2_{\mathcal{B}\mathcal{B}}\{G_{\mathcal{B}\mathcal{A}} \cdot [(H^3_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})H^4_{\mathcal{A}\mathcal{A}}]\})]$$
$$- \tfrac{1}{2}(H^1_{\mathcal{A}\mathcal{A}}\{G_{\mathcal{A}\mathcal{B}} \cdot [(H^2_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^3_{\mathcal{B}\mathcal{B}}]\}) \cdot G_{\mathcal{A}\mathcal{B}}]H^4_{\mathcal{B}\mathcal{A}} - \tfrac{1}{2}[(\{[H^1_{\mathcal{A}\mathcal{A}}(G_{\mathcal{A}\mathcal{B}} \cdot H^2_{\mathcal{A}\mathcal{B}})] \cdot G_{\mathcal{A}\mathcal{B}}\}H^3_{\mathcal{B}\mathcal{B}}) \cdot G_{\mathcal{A}\mathcal{B}}]H^4_{\mathcal{B}\mathcal{A}}$$
$$- \tfrac{1}{3}(\{[(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^2_{\mathcal{B}\mathcal{A}}](G_{\mathcal{A}\mathcal{B}} \cdot H^3_{\mathcal{A}\mathcal{B}})\} \cdot G_{\mathcal{A}\mathcal{B}})H^4_{\mathcal{B}\mathcal{A}} - \tfrac{1}{3}H^1_{\mathcal{A}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot \{(H^2_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})[H^3_{\mathcal{A}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^4_{\mathcal{B}\mathcal{A}})]\})$$
$$- \tfrac{1}{6}\{[(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})(H^2_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})H^3_{\mathcal{A}\mathcal{B}}] \cdot G_{\mathcal{A}\mathcal{B}}\}H^4_{\mathcal{B}\mathcal{A}} - \tfrac{1}{6}\{[H^1_{\mathcal{A}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^2_{\mathcal{B}\mathcal{A}})(G_{\mathcal{A}\mathcal{B}} \cdot H^3_{\mathcal{A}\mathcal{B}})] \cdot G_{\mathcal{A}\mathcal{B}}\}H^4_{\mathcal{B}\mathcal{A}}$$
$$- \tfrac{1}{6}H^1_{\mathcal{A}\mathcal{B}}\{G_{\mathcal{B}\mathcal{A}} \cdot [H^2_{\mathcal{B}\mathcal{A}}(G_{\mathcal{A}\mathcal{B}} \cdot H^3_{\mathcal{A}\mathcal{B}})(G_{\mathcal{B}\mathcal{A}} \cdot H^4_{\mathcal{B}\mathcal{A}})]\} - \tfrac{1}{6}H^1_{\mathcal{A}\mathcal{B}}\{G_{\mathcal{B}\mathcal{A}} \cdot [(H^2_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})(H^3_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^4_{\mathcal{B}\mathcal{A}}]\}$$
$$+ \tfrac{1}{24}(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})(H^2_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})(H^3_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^4_{\mathcal{B}\mathcal{A}} + \tfrac{1}{24}H^1_{\mathcal{A}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^2_{\mathcal{B}\mathcal{A}})(G_{\mathcal{A}\mathcal{B}} \cdot H^3_{\mathcal{A}\mathcal{B}})(G_{\mathcal{B}\mathcal{A}} \cdot H^4_{\mathcal{B}\mathcal{A}})$$
$$+ \tfrac{1}{8}(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})(H^2_{\mathcal{B}\mathcal{A}} \cdot G_{\mathcal{B}\mathcal{A}})H^3_{\mathcal{A}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^4_{\mathcal{B}\mathcal{A}}) + \tfrac{1}{8}(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^2_{\mathcal{B}\mathcal{A}}(G_{\mathcal{A}\mathcal{B}} \cdot H^3_{\mathcal{A}\mathcal{B}})(G_{\mathcal{B}\mathcal{A}} \cdot H^4_{\mathcal{B}\mathcal{A}})$$
$$+ \tfrac{1}{2}[(\{[(H^1_{\mathcal{A}\mathcal{B}} \cdot G_{\mathcal{A}\mathcal{B}})H^2_{\mathcal{B}\mathcal{B}}] \cdot G_{\mathcal{A}\mathcal{B}}\}H^3_{\mathcal{B}\mathcal{B}}) \cdot G_{\mathcal{A}\mathcal{B}}]H^4_{\mathcal{B}\mathcal{A}} + \tfrac{1}{2}H^1_{\mathcal{A}\mathcal{B}}[G_{\mathcal{B}\mathcal{A}} \cdot (H^2_{\mathcal{B}\mathcal{B}}\{G_{\mathcal{B}\mathcal{A}} \cdot [H^3_{\mathcal{B}\mathcal{B}}(G_{\mathcal{B}\mathcal{A}} \cdot H^4_{\mathcal{B}\mathcal{A}})]\})].$$
$$(B6)$$

This way of expressing $\bar{H}_{\mathcal{A}\mathcal{A}}$ is particularly useful when H' is a sum of operators that do not commute, and one wishes to keep track of the order of the various terms. In the present case, H' is a sum of $\mathbf{k} \cdot \mathbf{p}$ terms and potential-energy matrix elements. Examples are given in the appendixes below.

APPENDIX C: BULK RENORMALIZATION

In terms of the functions defined in Appendix B, the renormalized coefficients of order k^2, k^3, and k^4 in the bulk reference Hamiltonian for set $\mathcal{A}$ are given by[1]

$$D^{ij}_{\mathcal{A}\mathcal{A}} = \widetilde{D}^{ij}_{\mathcal{A}\mathcal{A}} + P_2(\pi^i,\pi^j), \qquad (C1)$$

$$C^{ijk}_{\mathcal{A}\mathcal{A}} = \widetilde{C}^{ijk}_{\mathcal{A}\mathcal{A}} + P_2(\pi^i,\widetilde{D}^{jk}) + P_2(\widetilde{D}^{ij},\pi^k) + P_3(\pi^i,\pi^j,\pi^k),$$
$$(C2)$$

$$Q^{ijkl}_{\mathcal{A}\mathcal{A}} = \widetilde{Q}^{ijkl}_{\mathcal{A}\mathcal{A}} + P_4(\pi^i,\pi^j,\pi^k,\pi^l) + P_3(\widetilde{D}^{ij},\pi^k,\pi^l)$$
$$+ P_3(\pi^i,\widetilde{D}^{jk},\pi^l) + P_3(\pi^i,\pi^j,\widetilde{D}^{kl}) + P_2(\widetilde{D}^{ij},\widetilde{D}^{kl})$$
$$+ P_2(\pi^i,\widetilde{C}^{jkl}) + P_2(\widetilde{C}^{ijk},\pi^l). \qquad (C3)$$

Here, a tilde denotes a quantity before renormalization, which is obtained by fitting the reference Hamiltonian to a polynomial of order k^4 (as described in Appendix A). The tilde is omitted on π^i because it does not change under renormalization.

APPENDIX D: LINEAR RENORMALIZATION

The terms in the renormalized $\mathcal{A}$ Hamiltonian that are linear in θ_α are given by[1]

$$\pi^{i\alpha}_{\mathcal{A}\mathcal{A}} = \widetilde{\pi}^{i\alpha}_{\mathcal{A}\mathcal{A}} + P_2(\pi^i,E^\alpha), \qquad (D1)$$

$$\pi^{\alpha i}_{\mathcal{A}\mathcal{A}} = \widetilde{\pi}^{\alpha i}_{\mathcal{A}\mathcal{A}} + P_2(E^\alpha,\pi^i), \qquad (D2)$$

$$D^{\alpha ij}_{\mathcal{A}\mathcal{A}} = \widetilde{D}^{\alpha ij}_{\mathcal{A}\mathcal{A}} + P_2(\widetilde{\pi}^{\alpha i},\pi^j) + P_2(E^\alpha,\widetilde{D}^{ij}) + P_3(E^\alpha,\pi^i,\pi^j),$$
$$(D3)$$

$$D^{i\alpha j}_{\mathcal{A}\mathcal{A}} = \widetilde{D}^{i\alpha j}_{\mathcal{A}\mathcal{A}} + P_2(\widetilde{\pi}^{i\alpha},\pi^j) + P_2(\pi^i,\widetilde{\pi}^{\alpha j}) + P_3(\pi^i,E^\alpha,\pi^j),$$
$$(D4)$$

$$D^{ij\alpha}_{\mathcal{A}\mathcal{A}} = \widetilde{D}^{ij\alpha}_{\mathcal{A}\mathcal{A}} + P_2(\pi^i,\widetilde{\pi}^{j\alpha}) + P_2(\widetilde{D}^{ij},E^\alpha) + P_3(\pi^i,\pi^j,E^\alpha).$$
$$(D5)$$

These are the same as the expressions given in Appendix C of Ref. 1, although written in a different notation. Once again, a tilde denotes a quantity before renormalization, which is obtained from a multipole expansion of the linear density and short-range potentials (Sec. II and Ref. 2) and from fitting the linear nonlocal potential to a polynomial of order k^2 (Appendix A). The tilde is omitted on E^α because it does not change under renormalization.

APPENDIX E: QUADRATIC RENORMALIZATION

Perturbative renormalization also generates terms that are quadratic in θ_α. The only term included in the Hamiltonian of Ref. 1 was the renormalized potential

$$E_{\mathcal{A}\mathcal{A}}^{\alpha\beta} = P_2(E^\alpha, E^\beta). \tag{E1}$$

Some of the present calculations also include quadratic renormalization of the momentum matrix,

$$\pi_{\mathcal{A}\mathcal{A}}^{\alpha\beta i} = P_2(E^\alpha, \widetilde{\pi}^{\beta i}) + P_2(\widetilde{E}^{\alpha\beta}, \pi^i) + P_3(E^\alpha, E^\beta, \pi^i), \tag{E2}$$

$$\pi_{\mathcal{A}\mathcal{A}}^{\alpha i\beta} = P_2(E^\alpha, \widetilde{\pi}^{i\beta}) + P_2(\widetilde{\pi}^{\alpha i}, E^\beta) + P_3(E^\alpha, \pi^i, E^\beta), \tag{E3}$$

$$\pi_{\mathcal{A}\mathcal{A}}^{i\alpha\beta} = P_2(\widetilde{\pi}^{i\alpha}, E^\beta) + P_2(\pi^i, \widetilde{E}^{\alpha\beta}) + P_3(\pi^i, E^\alpha, E^\beta), \tag{E4}$$

and of the effective masses,

$$D_{\mathcal{A}\mathcal{A}}^{\alpha\beta ij} = P_2(E^\alpha, \widetilde{D}^{\beta ij}) + P_2(\widetilde{E}^{\alpha\beta}, \widetilde{D}^{ij}) + P_3(E^\alpha, E^\beta, \widetilde{D}^{ij})$$
$$+ P_3(E^\alpha, \widetilde{\pi}^{\beta i}, \pi^j) + P_3(\widetilde{E}^{\alpha\beta}, \pi^i, \pi^j) + P_4(E^\alpha, E^\beta, \pi^i, \pi^j), \tag{E5}$$

$$D_{\mathcal{A}\mathcal{A}}^{i\alpha\beta j} = P_2(\widetilde{\pi}^{i\alpha}, \widetilde{\pi}^{\beta j}) + P_3(\widetilde{\pi}^{i\alpha}, E^\beta, \pi^j) + P_3(\pi^i, E^\alpha, \widetilde{\pi}^{\beta j})$$
$$+ P_3(\pi^i, \widetilde{E}^{\alpha\beta}, \pi^j) + P_4(\pi^i, E^\alpha, E^\beta, \pi^j), \tag{E6}$$

$$D_{\mathcal{A}\mathcal{A}}^{ij\alpha\beta} = P_2(\widetilde{D}^{ij\alpha}, E^\beta) + P_2(\widetilde{D}^{ij}, \widetilde{E}^{\alpha\beta}) + P_3(\widetilde{D}^{ij}, E^\alpha, E^\beta)$$
$$+ P_3(\pi^i, \widetilde{\pi}^{j\alpha}, E^\beta) + P_3(\pi^i, \pi^j, \widetilde{E}^{\alpha\beta}) + P_4(\pi^i, \pi^j, E^\alpha, E^\beta), \tag{E7}$$

$$D_{\mathcal{A}\mathcal{A}}^{\alpha i\beta j} = P_2(\widetilde{\pi}^{\alpha i}, \widetilde{\pi}^{\beta j}) + P_2(E^\alpha, \widetilde{D}^{i\beta j}) + P_3(E^\alpha, \pi^i, \widetilde{\pi}^{\beta j})$$
$$+ P_3(E^\alpha, \widetilde{\pi}^{i\beta}, \pi^j) + P_3(\widetilde{\pi}^{\alpha i}, E^\beta, \pi^j) + P_4(E^\alpha, \pi^i, E^\beta, \pi^j), \tag{E8}$$

$$D_{\mathcal{A}\mathcal{A}}^{i\alpha j\beta} = P_2(\widetilde{\pi}^{i\alpha}, \widetilde{\pi}^{j\beta}) + P_2(\widetilde{D}^{i\alpha j}, E^\beta) + P_3(\widetilde{\pi}^{i\alpha}, \pi^j, E^\beta)$$
$$+ P_3(\pi^i, \widetilde{\pi}^{\alpha j}, E^\beta) + P_3(\pi^i, E^\alpha, \widetilde{\pi}^{j\beta}) + P_4(\pi^i, E^\alpha, \pi^j, E^\beta), \tag{E9}$$

$$D_{\mathcal{A}\mathcal{A}}^{\alpha ij\beta} = P_2(\widetilde{\pi}^{\alpha i}, \widetilde{\pi}^{j\beta}) + P_2(\widetilde{D}^{\alpha ij}, E^\beta) + P_2(E^\alpha, \widetilde{D}^{ij\beta})$$
$$+ P_3(\widetilde{\pi}^{\alpha i}, \pi^j, E^\beta) + P_3(E^\alpha, \widetilde{D}^{ij}, E^\beta) + P_3(E^\alpha, \pi^i, \widetilde{\pi}^{j\beta})$$
$$+ P_4(E^\alpha, \pi^i, \pi^j, E^\beta). \tag{E10}$$

In these expressions, several approximations are used. No fitting of the momentum and mass terms in the original quadratic Hamiltonian was performed here; consequently, the unrenormalized parts are set to zero. Also, the term $\widetilde{E}^{\alpha\beta}$ is the mean unrenormalized diatomic short-range potential summed over all diatomic perturbations with the given values of α and β.

Some of the calculations also include renormalized short-range potential terms of order θ^3 and θ^4, which were approximated using the expressions

$$E_{\mathcal{A}\mathcal{A}}^{\alpha\beta\gamma} = P_2(\widetilde{E}^{\alpha\beta}, E^\gamma) + P_2(E^\alpha, \widetilde{E}^{\beta\gamma}) + P_3(E^\alpha, E^\beta, E^\gamma), \tag{E11}$$

$$E_{\mathcal{A}\mathcal{A}}^{\alpha\beta\gamma\delta} = P_2(\widetilde{E}^{\alpha\beta}, \widetilde{E}^{\gamma\delta}) + P_3(E^\alpha, E^\beta, \widetilde{E}^{\gamma\delta}) + P_3(E^\alpha, \widetilde{E}^{\beta\gamma}, E^\delta)$$
$$+ P_3(\widetilde{E}^{\alpha\beta}, E^\gamma, E^\delta) + P_4(E^\alpha, E^\beta, E^\gamma, E^\delta). \tag{E12}$$

Corrections of the same order arising from the long-range terms in the Hartree potential were neglected. For all of the numerical examples treated here, the $O(\theta^4)$ terms in Eq. (E12) were found to be negligible.

It should be emphasized that the results presented here do not provide a fully consistent perturbation scheme according to the criteria given by Takhtamirov and Volkov,[10] in which the mean kinetic energy of the states of interest is assumed to be comparable to the heterostructure potential-energy perturbation. According to this scheme, if one includes the $O(k^2\theta^2)$ and $O(k^0\theta^3)$ terms shown here, one should also include terms of order $k^4\theta^1$ and $k^6\theta^0$.

However, since these require the use of sixth-order perturbation theory, such terms were judged to be not worth the effort in a preliminary investigation of this nature. Therefore, the results obtained by adding only the $O(k^2\theta^2)$ and $O(k^0\theta^3)$ terms are not expected to be valid for kinetic energies covering the full range of the band offset, but only for kinetic energies small in comparison to the band offset. This is indeed what was found in the numerical calculations of Sec. V. Likewise, the $O(k^4\theta^0)$ terms were found to be less important for states of small kinetic energy. These results are merely a reflection of the fact that a wide quantum well, unlike a hydrogenic impurity,[8] does have states in which the mean kinetic energy is small compared to the mean perturbing potential. Thus, in this sense the theory of low-energy excitations of wide quantum wells is actually simpler than the corresponding theory of shallow impurities, because terms of high order in k are of lesser importance.

APPENDIX F: QUADRATIC RESPONSE

The method used here to handle the quadratic response differs slightly from that of Ref. 1. The quadratic potential response is given by Eq. (3.14) of Ref. 1:

$$V^{(2)}(\mathbf{x}, \mathbf{x}') = \sum_{\alpha, \mathbf{R}}' \sum_{\alpha', \mathbf{R}'}' \theta_{\mathbf{R}}^\alpha \theta_{\mathbf{R}'}^{\alpha'} \Delta v_{\mathbf{R}\mathbf{R}'}^{\alpha\alpha'}(\mathbf{x}, \mathbf{x}'), \tag{F1}$$

which has the same form as the quadratic density response in Eq. (3c). The translation symmetry of the reference crystal allows $\Delta v_{\mathbf{R}\mathbf{R}'}^{\alpha\alpha'}(\mathbf{x}, \mathbf{x}')$ to be written as

$$\Delta v_{\mathbf{R}\mathbf{R}'}^{\alpha\alpha'}(\mathbf{x}, \mathbf{x}') \equiv \Delta v^{\alpha\alpha', \mathbf{R}'-\mathbf{R}}(\mathbf{x} - \overline{\mathbf{R}}_{\alpha\alpha'}, \mathbf{x}' - \overline{\mathbf{R}}_{\alpha\alpha'}), \tag{F2}$$

in which $\overline{\mathbf{R}}_{\alpha\alpha'}$ is the midpoint of the two atoms:

$$\overline{\mathbf{R}}_{\alpha\alpha'} = \frac{1}{2}(\mathbf{R} + \mathbf{R}' + \boldsymbol{\tau}_\alpha + \boldsymbol{\tau}_{\alpha'}). \tag{F3}$$

In Eq. (F2), the coordinate reference is taken to be $\overline{\mathbf{R}}_{\alpha\alpha'}$, whereas in Eq. (3.17) of Ref. 1, it was chosen to be $\mathbf{R}_\alpha$

$\equiv \mathbf{R} + \boldsymbol{\tau}_\alpha$. The Fourier transform of Eq. (F1) is[1]

$$V^{(2)}(\mathbf{k}, \mathbf{k}') = N \sum_{\alpha, \alpha', \mathbf{R}'}{}' \theta^{\alpha \alpha' \mathbf{R}'}(\mathbf{k} - \mathbf{k}') \Delta v^{\alpha \alpha' \mathbf{R}'}(\mathbf{k}, \mathbf{k}'), \quad \text{(F4)}$$

where

$$\theta^{\alpha \alpha' \mathbf{R}'}(\mathbf{k}) = \frac{1}{N} \sum_{\mathbf{R}} \theta^{\alpha}_{\mathbf{R}} \theta^{\alpha'}_{\mathbf{R}+\mathbf{R}'} e^{-i\mathbf{k} \cdot [\mathbf{R}_\alpha + (\mathbf{R}' + \tau_{\alpha'} - \tau_\alpha)/2]}. \quad \text{(F5)}$$

The coordinate reference (F3) is arbitrary, but it is sometimes more convenient for analyzing symmetry properties than the choice used in Ref. 1.

In the LDA model used here, the quadratic potential is local [i.e., $\Delta v^{\alpha \alpha' \mathbf{R}'}(\mathbf{k}, \mathbf{k}') = \Delta v^{\alpha \alpha' \mathbf{R}'}(\mathbf{k} - \mathbf{k}')$] since the ionic pseudopotential is purely linear. The diatomic potential $\Delta v^{\alpha \alpha' \mathbf{R}'}(\mathbf{q})$ was approximated for small $\mathbf{q}$ by keeping only the dipole and quadrupole terms in the electron density and the $l=0$ term in the power-series expansion (4) of the short-range potential (which in the quadratic case consists only of the exchange-correlation potential). For $\mathbf{q}$ near $\mathbf{G} \neq \mathbf{0}$, only the $l=0$ terms were retained in both the density and short-range potential; see Sec. VI B of the preceding paper[2] for further details.

*phbaf@ust.hk

[1] B. A. Foreman, Phys. Rev. B **72**, 165345 (2005).

[2] B. A. Foreman, preceding paper, Phys. Rev. B **76**, 045326 (2007).

[3] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).

[4] W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1957), Vol. 5, pp. 257–320.

[5] F. Bassani, G. Iadonisi, and B. Preziosi, Rep. Prog. Phys. **37**, 1099 (1974).

[6] S. T. Pantelides, Rev. Mod. Phys. **50**, 797 (1978).

[7] L. J. Sham, Phys. Rev. **150**, 720 (1966).

[8] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974); see Sec. 27 for corrections to elementary effective-mass theory.

[9] M. G. Burt, J. Phys.: Condens. Matter **4**, 6651 (1992).

[10] É. E. Takhtamirov and V. A. Volkov, JETP **89**, 1000 (1999).

[11] J.-B. Xia, Phys. Rev. B **39**, 3310 (1989).

[12] L.-W. Wang, A. Franceschetti, and A. Zunger, Phys. Rev. Lett. **78**, 2819 (1997).

[13] L.-W. Wang and A. Zunger, Phys. Rev. B **56**, 12395 (1997).

[14] L.-W. Wang and A. Zunger, Phys. Rev. B **59**, 15806 (1999).

[15] L. V. Keldysh, Sov. Phys. JETP **18**, 253 (1964).

[16] L. Leibler, Phys. Rev. B **12**, 4443 (1975).

[17] L. Leibler, Phys. Rev. B **16**, 863 (1977).

[18] B. A. Foreman, Phys. Rev. B **75**, 235331 (2007).

[19] O. von Roos, Phys. Rev. B **27**, 7547 (1983).

[20] R. A. Morrow and K. R. Brownstein, Phys. Rev. B **30**, 678 (1984).

[21] R. A. Morrow and K. R. Brownstein, Phys. Rev. B **31**, 1135 (1985).

[22] T. Gora and F. Williams, Phys. Rev. **177**, 1179 (1969).

[23] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[24] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).

[25] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, and D. C. Allan, Comput. Mater. Sci. **25**, 478 (2002).

[26] X. Gonze, G.-M. Rignanese, M. Verstraete, J.-M. Beuken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, Ph. Ghosez, M. Veithen, J.-Y. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D. R. Hamann, and

D. C. Allan, Z. Kristallogr. **220**, 558 (2005).

[27] The ABINIT code is a common project of the Université Catholique de Louvain, Corning Incorporated, and other contributors (URL: http://www.abinit.org).

[28] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*, Springer Tracts in Modern Physics Vol. 191 (Springer, Berlin, 2003).

[29] J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).

[30] If one considers single atomic perturbations rather than planar perturbations, there are additional nonanalytic contributions from the linear octopole and hexadecapole moments of the electron density (Ref. 1). For the present case of planar perturbations, these are simply absorbed into the definition of the short-range potential using Eq. (12) of Ref. 2.

[31] For examples illustrating the transformation of operators from momentum to coordinate space, see Eqs. (22.10b) and (22.14c) of Bir and Pikus (Ref. 8) and Eqs. (4.7) and (4.8) of Winkler (Ref. 28).

[32] E. L. Ivchenko, A. Yu. Kaminski, and U. Rössler, Phys. Rev. B **54**, 5852 (1996).

[33] Y. R. Lin-Liu and L. J. Sham, Phys. Rev. B **32**, 5561 (1985).

[34] R. Eppenga, M. F. H. Schuurmans, and S. Colak, Phys. Rev. B **36**, 1554 (1987).

[35] É. I. Rashba, Sov. Phys. Solid State **2**, 1109 (1960).

[36] Yu. A. Bychkov and E. I. Rashba, J. Phys. C **17**, 6039 (1984).

[37] R. Winkler, Phys. Rev. B **62**, 4245 (2000).

[38] R. Winkler, H. Noh, E. Tutuc, and M. Shayegan, Phys. Rev. B **65**, 155303 (2002).

[39] B. A. Foreman, Phys. Rev. B **48**, R4964 (1993).

[40] C. R. Pidgeon and R. N. Brown, Phys. Rev. **146**, 575 (1966).

[41] B. A. Foreman, J. Phys.: Condens. Matter **18**, 1335 (2006).

[42] They could be obtained from analytical derivatives of the nonlocal pseudopotential in $\mathbf{k}$ space, but that was not done here since the primary focus of this paper is the physics of abrupt heterojunctions. The value of K also depends on the choice of coupling between the nonlocal potential and the magnetic field (Ref. 41).

[43] E. O. Kane, in *Narrow Gap Semiconductors: Physics and Applications*, edited by W. Zawadzki, Lecture Notes in Physics, Vol. 133 (Springer, Berlin, 1980), pp. 13–31.

[44] D. J. BenDaniel and C. B. Duke, Phys. Rev. **152**, 683 (1966).

[45] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Wiley, New York, 1988).

[46] P.-O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).

[47] M. E. Flatté, P. M. Young, L.-H. Peng, and H. Ehrenreich, Phys. Rev. B **53**, 1963 (1996).

[48] P. Pfeffer and W. Zawadzki, Phys. Rev. B **53**, 12813 (1996).

[49] W. Yang and K. Chang, Phys. Rev. B **72**, 233309 (2005).

[50] R. Dingle, W. Wiegmann, and C. H. Henry, Phys. Rev. Lett. **33**, 827 (1974).

[51] L. J. Sham and Y.-T. Lu, J. Lumin. **44**, 207 (1989).

[52] G. F. Koster, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1957), Vol. 5, pp. 173–256.

[53] W. H. Lau and M. E. Flatté, Appl. Phys. Lett. **80**, 1683 (2002).

[54] F. Szmulowicz, H. Haugan, and G. J. Brown, Phys. Rev. B **69**, 155321 (2004).

[55] Rodina and Alekseev (Ref. 56) have suggested dividing the heterostructure Hamiltonian into a bulk part and a surface part, in which the valence-band Rashba parameters for the bulk part are defined to be the same as the magnetic Luttinger parameters. However, since by symmetry the surface part contains terms with the same mathematical structure as the bulk part, this division does not carry any meaning unless one can make some definite statement about the magnitude of the surface part. As shown by the model calculations presented here, the surface part in this scheme is not generally negligible.

[56] A. V. Rodina and A. Yu. Alekseev, Phys. Rev. B **73**, 115312 (2006).

[57] Q.-G. Zhu and H. Kroemer, Phys. Rev. B **27**, 3519 (1983).

[58] W. A. Harrison, Phys. Rev. **123**, 85 (1961).

[59] K. Young, Phys. Rev. B **39**, 13434 (1989).

[60] This scheme would be limited to the low-energy excitations of a wide quantum well or superlattice, as discussed in the remarks at the end of Appendix E.

[61] S. Baroni, R. Resta, A. Baldereschi, and M. Peressi, in *Spectroscopy of Semiconductor Microstructures*, edited by G. Fasol, A. Fasolino, and P. Lugli, NATO Advanced Studies Institute, Series B: Physics (Plenum, New York, 1989), pp. 251–272.