Accurate quadratic-response approximation for the self-consistent pseudopotential of semiconductor nanostructures

Bradley A. Foreman*
Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
(Received 27 March 2007; revised manuscript received 5 June 2007; published 23 July 2007)

Quadratic-response theory is shown to provide a conceptually simple but accurate approximation for the self-consistent one-electron potential of semiconductor nanostructures. Numerical examples are presented for GaAs/AlAs and $In_{0.53}Ga_{0.47}As/InP$ (001) superlattices using the local-density approximation to density-functional theory and norm-conserving pseudopotentials without spin-orbit coupling. When the reference crystal is chosen to be the virtual-crystal average of the two bulk constituents, the absolute error in the quadratic-response potential for Γ_{15} valence electrons is about 2 meV for GaAs/AlAs and 5 meV for $In_{0.53}Ga_{0.47}As/InP$. Low-order multipole expansions of the electron density and potential responses are shown to be accurate throughout a small neighborhood of each reciprocal-lattice vector, thus providing a further simplification that is confirmed to be valid for slowly varying envelope functions. Although the linear response is about an order of magnitude larger than the quadratic response, the quadratic terms are important both quantitatively (if an accuracy of better than a few tens of meV is desired) and qualitatively (due to their different symmetry and long-range dipole effects).

DOI: 10.1103/PhysRevB.76.045326 PACS number(s): 73.21.−b, 73.61.Ey, 71.15.Ap

I. INTRODUCTION

The potential energy of the electron in the single-electron approximation provides the theoretical foundation for most modeling and interpretation of the electronic energy bands of nonmagnetic semiconductors. Implementations of this concept range in sophistication from simple empirical pseudopotentials[1,2] to first-principles quasiparticle self-energies.[3,4] Modern research on semiconductor physics deals mainly with complex systems (such as nanostructures and novel materials), although the spintronic properties of traditional semiconductors are also of strong current interest. It is therefore somewhat surprising that, in this day of supposedly advanced understanding, no clear physical picture of something so basic as the one-electron potential energy at an ideal GaAs/AlAs or GaInAs/InP heterojunction exists.

The key issue is the self-consistent redistribution of charge that must occur at any interface between two bulk semiconductors. First-principles numerical calculations based on density-functional theory and norm-conserving pseudopotentials[5,6] are now routinely performed on small systems (say, a few tens of atoms) using personal computers and free software.[7–9] However, such computations are still not feasible for large quantum dots and quantum wires. Even if they were, there is no means at present for separating this information from a huge numerical calculation. That is, no one has yet demonstrated an accurate method for breaking down the self-consistent potential into simple conceptual units that can be used to *understand* the interface potential and model it using less numerically intensive methods (such as empirical pseudopotentials or envelope-function theory).

The purpose of this paper is to demonstrate that the quadratic-response formalism developed in Refs. 10 and 11 provides a conceptually simple method for constructing the interface potential that is accurate to within a few meV in typical semiconductor heterostructures. In this approach, the heterostructure is treated as a perturbation of a bulk reference crystal, and the self-consistent heterostructure potential is built up as a series of successively smaller corrections to the bulk potential.

The leading term is the linear response, which has the form of a superposition of elementary quasiatomic building blocks. This form is widely used in empirical pseudopotential models (see the literature review in Sec. II), but it is shown here to be accurate to only a few tens of meV. The quadratic response is then added in the form of *diatomic* building blocks, which are qualitatively different because they carry a dipole moment even in cubic semiconductors. Numerical implementations of the theory are presented here for the paradigmatic common-atom and no-common-atom systems GaAs/AlAs and $In_{0.53}Ga_{0.47}As/InP$. The quadratic-response approximation for Γ_{15} valence electrons is shown to be accurate to within about 2 meV for GaAs/AlAs and 5 meV for $In_{0.53}Ga_{0.47}As/InP$.

In typical heterostructures, the low-energy electron and hole states can be represented by slowly varying envelope functions. The properties of such states are determined by the pseudopotential not just at the bulk reciprocal-lattice vectors **G** (as would be the case in a bulk semiconductor) but throughout a small **k**-space neighborhood of each **G**. Within such a neighborhood, an analytic function of **k** can be approximated using a power series (which is equivalent to introducing a multipole expansion of a localized function in coordinate space). This paper also examines the accuracy of such power series and demonstrates that the truncated expansions introduced in Refs. 10 and 11 are sufficiently accurate for slowly varying envelope functions. The application of these multipole expansions in actual envelope-function calculations is presented in another paper.[12]

The paper begins in Sec. II with a review of previous models of the interface potential. Section III describes the choices made in defining the model system used here for numerical calculations. The basic features of the self-consistent density and potential in GaAs/AlAs and

$In_{0.53}Ga_{0.47}As/InP$ superlattices are examined in Sec. IV; here, the valence-band offsets calculated for the model systems are found to be in reasonably good agreement with experiment. Section V outlines the methods used to calculate the linear and quadratic responses to the heterostructure perturbation and discusses the physical significance of the results. Localization of the density and potential responses and the use of multipole expansions at small wave vectors are considered in Sec. VI. Section VII shows that the cumulative errors arising from the quadratic-response and truncated multipole approximations are limited to a few meV for the systems studied. The conclusions of the paper are reviewed in Sec. VIII. Finally, Sec. IX discusses several limitations of the calculations presented here, along with possible extensions of this work that could be used to overcome some of these limitations.

II. PREVIOUS DESCRIPTIONS OF THE INTERFACE POTENTIAL

Two simple non-self-consistent models of the heterostructure pseudopotential are widely used. The simplest is the *planar interface* model, in which the potential is assumed to make an abrupt transition from one periodic bulk potential to another at some predefined interface plane.[13–20] This requires as input only the bulk pseudopotential form factors.[1,2] The assumption of a planar interface provides, in essence, a simple recipe for interpolating the pseudopotential in **k** space between the bulk reciprocal-lattice vectors **G**. A discontinuous potential is, however, clearly far from self-consistent (since by Poisson's equation it would require an infinite electron density), and a planar interface fails to account for the basic physics of the bonds between atoms at a no-common-atom interface.[21]

An arguably more realistic potential (although not all would agree[20]) is provided by the *atomic superposition* model, which is defined as a linear combination of screened atomic pseudopotentials.[22–28] Here, the atomic potentials are assumed to be known as continuous functions of **k**;[29] hence, no interpolation is required. In its simplest form, this model may not accurately reproduce differences between bulk semiconductors,[26] since (for example) the screened As pseudopotential obtained by fitting the properties of bulk GaAs will generally differ from that obtained by fitting AlAs.

However, Mäder and Zunger have overcome this difficulty by using environment-dependent pseudopotentials,[26] in which the replacement of one Ga atom in GaAs by Al is accompanied by small changes in each of the four neighboring As pseudopotentials. The atomic pseudopotential and its associated nearest-neighbor perturbations may be grouped conceptually into a single quasiatomic unit, which has the site symmetry T_d of an Al impurity in GaAs rather than the spherical symmetry of an isolated Al atom. Thus, in this more general *quasiatomic superposition* model,[26] the heterostructure potential is constructed as a linear combination of quasiatomic potentials, each of which has the site symmetry of a substitutional impurity. (The planar interface model may also be interpreted as a superposition of elementary units,[30] but in this case, the units do not have the atomic site symmetry.)

The superposition potential is also manifestly non-self-consistent, if only because the exchange-correlation potential is nonlinear. The assumption of linearity has important physical consequences. As shown in Eq. (10) of Ref. 31, one of these is the absence of direct Γ_1-X_1 intervalley coupling in GaAs/AlAs (001) heterostructures. The same conclusion was reached in Ref. 32 using a superposition of quasiatomic potentials that were constructed by a different method. These conclusions were criticized by Takhtamirov and Volkov[33] on the grounds that the true interface potential does not have the assumed form and that there is consequently no basis for assuming Γ_1-X_1 coupling to be any smaller than Γ_1-X_3 coupling (the latter of which is nonvanishing in the quasiatomic superposition model[31,32]). This criticism is, of course, entirely correct if the self-consistent interface potential is viewed as completely unknown.

Indeed, many derivations of envelope-function models are prefaced by a warning that the potential within a few lattice constants of the interface is too complicated for direct analysis,[34–39] even in the case of an ideal interface (i.e., no interface roughness, defects, or interdiffusion). In early works, it was sometimes concluded that the details of the interface potential have no significant influence on slowly varying envelope functions. Today, however, it is widely accepted that these details are of critical importance in determining the strength of small symmetry-breaking interface terms in the Hamiltonian, such as the above-mentioned intervalley mixing,[31–33] valence-band mixing,[32,40,41] and conduction- and valence-band Rashba couplings[42] (see the following paper[12] for further study of some of these topics). However, the assumption that the self-consistent potential cannot be understood in simple terms remains unchallenged by most envelope-function theorists.

This problem was tackled in a landmark series of papers by Baroni *et al.*, who showed that linear-response theory is very successful in predicting valence-band offsets in semiconductor heterostructures.[43–52] In this approach, the ionic pseudopotential of the heterostructure is treated as a perturbation of a bulk reference crystal. The linear response to this perturbation was shown to give accurate predictions for the valence-band offset (within about 0.01 eV of the value obtained from a direct calculation, and in good agreement with experiment) if the reference crystal was chosen to be the virtual-crystal average of the two bulk materials, since the quadratic response is then the same on both sides of the heterojunction and gives no contribution to the offset. However, if one is interested in the actual position-dependent potential, including the details of its behavior near the interface, the quadratic response is not generally negligible.

Nonlinear screened empirical pseudopotentials were introduced by Magri and Zunger,[53] who extended the environment-dependent concept of Mäder and Zunger[26] to the no-common-atom InAs/GaSb system. Nonlinearity arises in this case because the nearest-neighbor perturbations depend on which atoms occupy the neighboring sites. The pseudopotentials can therefore no longer be interpreted as purely quasiatomic units; a complete description requires the inclusion of nearest-neighbor *diatomic* building blocks as well.[54] This represents an important conceptual advance in the modeling of heterostructure pseudopotentials.

The present work develops this concept more thoroughly by extending the first-principles linear-response theory of Baroni *et al.*[45] to include the diatomic quadratic response for arbitrary pairs of atoms. The results show that the environment dependence of the empirical pseudopotentials in Refs. 26 and 53 is qualitatively incomplete, since (for example) the leading contribution to Γ_1-X_1 mixing in GaAs/AlAs (001) superlattices arises from second-nearest-neighbor Ga-Al pairs. (However, this contribution is about 2 orders of magnitude smaller than the linear response, which supports the conclusion that Γ_1-X_1 mixing is much weaker than Γ_1-X_3 mixing.) Also, the diatomic nearest-neighbor response for no-common-atom systems calculated here has a dipole moment that generates long-range electric fields extending over the entire superlattice. Such fields do not occur in the empirical pseudopotential interpolation scheme of Magri and Zunger,[53] which produces only short-range potentials. This suggests that a more complete description of quadratic diatomic response may be of benefit in achieving more realistic empirical pseudopotentials.

III. NUMERICAL CONSIDERATIONS

A. Definition of model system

All of the numerical results in this paper are derived from plane-wave pseudopotential total-energy calculations[5,6] performed using the ABINIT software.[7–9] This package provides a variety of options, but the particular calculations reported here used the local-density approximation (LDA) to density-functional theory and the nonlocal norm-conserving pseudopotentials of Hartwigsen *et al.*[55] with no spin-orbit coupling.

Choosing a particular physical structure and a particular set of technical ingredients (such as plane-wave kinetic-energy cutoffs and **k**-point sampling) defines a model system whose properties can be calculated self-consistently to an accuracy approaching that of machine precision. These "exact" model calculations are used here and in the following paper[12] as a benchmark for comparison with various approximations used in the construction of the first-principles envelope-function theory of Ref. 11. The main objective of this paper is to pin down as closely as possible how much error arises from the quadratic-response approximation and the multipole expansions of the linear and quadratic responses. The approximate potentials derived here are also used directly in a subsequent paper[12] that examines the accuracy of first-principles envelope-function theory in superlattices.

With these goals in mind, relatively low kinetic-energy cutoffs (see below) and numbers of special **k** points were chosen in order to make the exact calculations feasible for the large superlattices in which the envelope-function theory is valid. In particular, two independent **k** points were used for GaAs/AlAs superlattices with D_{2d} symmetry, while four points were used for $In_{0.53}Ga_{0.47}As/InP$ superlattices with C_{2v} symmetry. The total energy is not completely converged at these values, but the valence-band offsets calculated here are nevertheless in good agreement with experiment and with previous first-principles calculations reported in the literature.

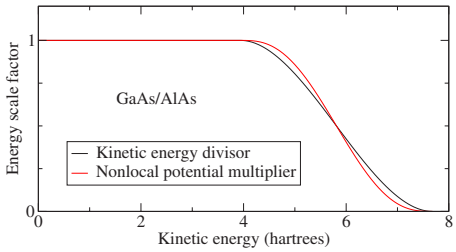

FIG. 1. (Color online) Kinetic-energy divisor (with continuous slope) and nonlocal potential-energy multiplier (with continuous curvature) used to smooth out discontinuities in the momentum-space electron density and pseudopotential. The multiplier used for the local pseudopotential has the same shape, but the energy scale is four times that shown.

In order to calculate multipole moments of the linear and quadratic responses, it is necessary that the electron density and short-range potential responses be well localized within the given large supercells. However, sharp cutoffs in the plane-wave basis produce spurious long-range Gibbs oscillations that are very small but nonetheless large enough to completely swamp the physical multipole values. To suppress these oscillations, the smoothing functions shown in Fig. 1 were applied to the kinetic and potential energies. The chosen kinetic-energy cutoff (i.e., the value at which the divisor in Fig. 1 goes to zero) includes 283 plane waves at the Γ point in bulk material (113 of which are not altered by smoothing) and more than 14 000 plane waves in a 96-atom superlattice.

B. Conventions defining the absolute energy

There are two conventions in the momentum-space total-energy formalism that determine the absolute value of the local potential. First, the mean value of the Hartree potential is set to zero by definition:[56–58]

$$V_{\mathrm{Hartree}}(\mathbf{q}) = (1 - \delta_{\mathbf{q0}})\frac{4\pi}{q^2}n(\mathbf{q}) \tag{1}$$

(where n is the valence electron density), because in a neutral system, the net contribution to the total energy from the Coulomb interaction terms (for the electron-ion system) at wave vector $\mathbf{q}=\mathbf{0}$ is identically zero.

A similar convention[56] is often used for the mean value of the short-range part of the local pseudopotential. This is given by $\bar{V}_{\mathrm{psp}}=\Sigma_j \bar{V}_j$, where[5,6,56,57]

$$\bar{V}_j = \frac{1}{\Omega}\int \left(V_j(r) + \frac{Z_j}{r}\right)d^3r, \tag{2}$$

in which Ω is the unit cell volume and V_j and Z_j are the local pseudopotential and valence charge of ion j. The contribution from $\bar{V}_{\mathrm{psp}}$ to the total energy is just the constant $E_{\mathrm{core}}=Z\Omega\bar{V}_{\mathrm{psp}}$, where $Z=\Sigma_j Z_j$. Hence, it is permissible to set $\bar{V}_{\mathrm{psp}}$ to zero in the Kohn-Sham equations.[56]

However, in order to calculate the nonlinear response correctly, $\bar{V}_{\mathrm{psp}}$ must be included in the local pseudopotential. In the present work, $\bar{V}_{\mathrm{psp}}$ was added to both the local pseudo-

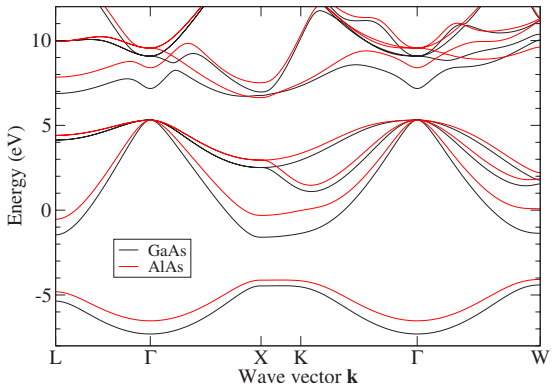

FIG. 2. (Color online) Energy band structure of bulk GaAs and AlAs.

potential and the energy eigenvalues after the conclusion of each self-consistent calculation. Therefore, the absolute energy scale for all figures in this paper (and the following paper[12]) is defined solely by the convention of setting the mean Hartree potential to zero.

IV. VALENCE-BAND OFFSETS IN SUPERLATTICES

Energy band structures for the model system described in Sec. III are presented in Figs. 2 and 3 for the bulk materials GaAs, AlAs, $In_{0.53}Ga_{0.47}As$, and InP. The conduction bands in these figures are clearly not realistic, as the minimum at L lies slightly below the Γ minimum in GaAs and is only slightly above it in $In_{0.53}Ga_{0.47}As$ and InP. The calculated Γ_{1c}-Γ_{15v} energy gaps are 1.850 eV for GaAs and 1.492 eV for $In_{0.53}Ga_{0.47}As$, which are 13% and 61% larger than the experimental values[59] (with spin-orbit splitting removed) of 1.633 and 0.925 eV. Nevertheless, the model does capture many of the gross qualitative trends in experimentally determined band structures,[2] such as the predominance of the X valley in AlAs relative to GaAs.

It is important to note that the difference between the valence-band maxima in Figs. 2 and 3 is not equal to the valence-band offset.[45] To obtain the valence-band offset, one must add to this difference the shift in the macroscopic Hartree potential of the two materials as given by a supercell

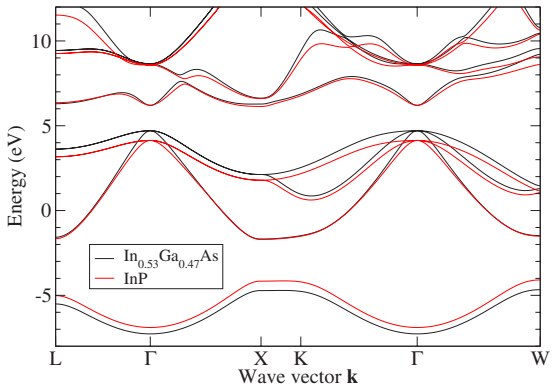

FIG. 3. (Color online) Energy band structure of bulk $In_{0.53}Ga_{0.47}As$ and InP.

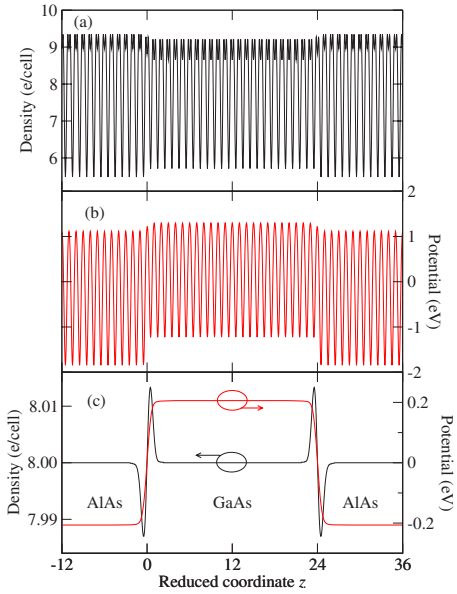

FIG. 4. (Color online) Density and potential of a (001) $(GaAs)_{24}(AlAs)_{24}$ superlattice: (a) Planar average of electron density. (b) Planar average of Hartree potential. (c) Macroscopic average of density and potential.

calculation.[45] The latter shift is shown in Figs. 4 and 5 for a GaAs/AlAs superlattice and in Figs. 6 and 7 for an $In_{0.53}Ga_{0.47}As$/InP superlattice. These figures show planar averages and macroscopic averages[60–62] of the electron density and Hartree potential, with distance given in units of the monolayer spacing $d=\frac{1}{2}a$, where a is the cubic lattice constant.

To calculate the macroscopic average, the microscopic functions were first averaged over the volume of a bulk unit cell, as in Ref. 60. Then, to ensure that the output is smooth and finite even when the input is singular (e.g., in the case of a truncated multipole expansion), an additional averaging was performed with respect to the normalized Gaussian function[61]

$$f(z) = \frac{1}{\sigma} \exp\left[-\pi \left(\frac{z}{\sigma} \right)^2 \right], \tag{3a}$$

which is equivalent to multiplying in momentum space by

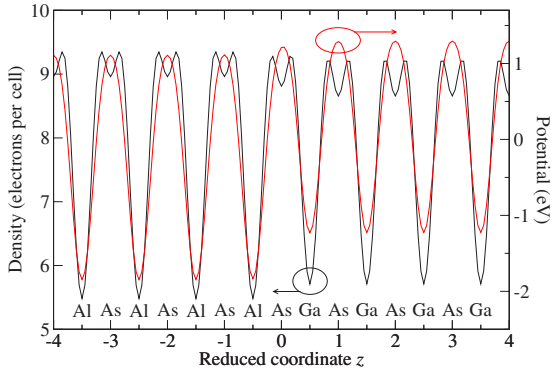

FIG. 5. (Color online) Expanded view of the planar average density and potential from Fig. 4.

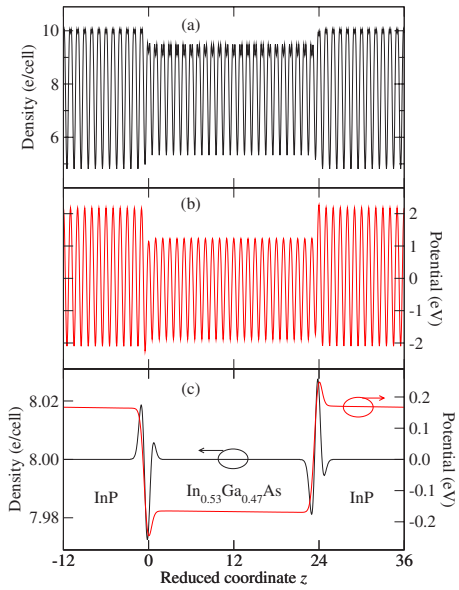

FIG. 6. (Color online) Density and potential of a (001) $(In_{0.53}Ga_{0.47}As)_{24}(InP)_{24}$ superlattice: (a) Planar average of electron density. (b) Planar average of Hartree potential. (c) Macroscopic average of density and potential.

$$g(k) = \exp[-\pi(k\sigma/2\pi)^2]. \tag{3b}$$

In Figs. 4(c) and 6(c), the distance σ was chosen to be $\sigma = d$.

The macroscopic density for GaAs/AlAs in Fig. 4(c) has the familiar interface dipole shape,[45,60] leading to a steplike shift in the Hartree potential of 0.412 eV. The shift predicted by linear-response theory is very close to this value, at 0.416 eV (see below for a more detailed study of the error in this approximation). These results agree well with the shift of 0.41 eV calculated in Ref. 45. Since the difference between the valence-band maxima in Fig. 2 is 18 meV, the valence-band offset for a (001) GaAs/AlAs heterojunction is 0.430 eV for the model system, as compared to the value of 0.45 eV reported in Ref. 45. If this figure is corrected for spin-orbit coupling and quasiparticle effects using the values in Ref. 45, the net offset is roughly 0.56 eV, in good agreement with the experimental value[59] of 0.53 eV.

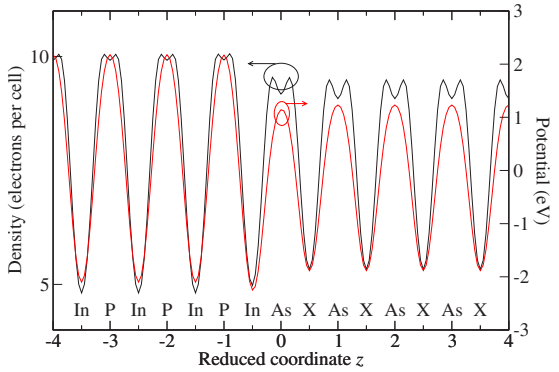

FIG. 7. (Color online) Expanded view of the planar average density and potential from Fig. 6. Here, X represents the virtual atom $In_{0.53}Ga_{0.47}$.

The macroscopic density of the no-common-atom $In_{0.53}Ga_{0.47}As/InP$ superlattice in Fig. 6(c) does not have a simple dipole shape because (to leading order) it is a superposition of two offset dipoles, one for the cation and one for the anion.[45,46] As discussed below, there is also a slight asymmetry of the two interfaces, leading to a macroscopic electric field that is barely visible in Fig. 6(c). The linear Hartree shifts [from Eq. (12) with $l=0$] for cations and anions are +0.366 and −0.697 eV, as compared to the values of +0.34 and −0.58 eV reported in Refs. 45 and 46. The net linear Hartree shift of −0.331 eV agrees well with the mean shift of −0.334 eV in Fig. 6(c). Combining this with the 0.574 eV difference in the valence-band maxima of Fig. 3, the net valence-band offset for the model $In_{0.53}Ga_{0.47}As/InP$ system is 0.240 eV. After adjusting for experimental spin-orbit splitting,[59] the calculated offset of 0.313 eV is close to the experimental value[59] of 0.345 eV.

V. LINEAR AND QUADRATIC RESPONSES

Having established that the model system adopted here provides at least a rough approximation of physical reality, the next step is to calculate the linear and quadratic responses to virtual-crystal perturbations of a bulk reference crystal. The reference crystal is chosen here to be the virtual-crystal average of the bulk constituents (i.e., $Al_{0.5}Ga_{0.5}As$ 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 lattice constant for all calculations was fixed at the value obtained by minimizing the total energy of the reference crystal (i.e., $a=10.5$ bohr for $Al_{0.5}Ga_{0.5}As$ and $a=10.9$ bohr for $In_{0.765}Ga_{0.235}As_{0.5}P_{0.5}$).

The perturbation of the heterostructure relative to the reference crystal is defined by the change in pseudopotential,

$$\Delta V_{\mathrm{psp}}(\mathbf{x}) = \sum_{\alpha} \sum_{\mathbf{R}} \theta_{\mathbf{R}}^{\alpha} v_{\mathrm{ion}}^{\alpha}(\mathbf{x} - \mathbf{R}_{\alpha}), \tag{4}$$

which is written as a local potential for simplicity (see Ref. 11 for the nonlocal case). Here, $v_{\mathrm{ion}}^{\alpha}(\mathbf{x})$ 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 $\mathbf{R}$ of the heterostructure relative to the reference crystal. Since the total change in fractional weight at each site must add to zero, this can be rewritten as[11]

$$\Delta V_{\mathrm{psp}}(\mathbf{x}) = \sum_{\alpha}{}' \sum_{\mathbf{R}} \theta_{\mathbf{R}}^{\alpha} \Delta v_{\mathrm{ion}}^{\alpha}(\mathbf{x} - \mathbf{R}_{\alpha}), \tag{5}$$

where the sum covers only independent values of α (e.g., either Ga or Al—but not both—in GaAs/AlAs) and $\Delta v_{\mathrm{ion}}^{\alpha}$ is the change in $v_{\mathrm{ion}}^{\alpha}$ relative to the reference crystal.

If $n(\mathbf{x})$ is the exact density of the heterostructure, the linear and quadratic responses to virtual perturbations $\theta_{\mathbf{R}}^{\alpha}$ are defined as the derivatives

$$\Delta n_{\mathbf{R}}^{\alpha}(\mathbf{x}) = \frac{\partial n(\mathbf{x})}{\partial \theta_{\mathbf{R}}^{\alpha}}, \tag{6a}$$

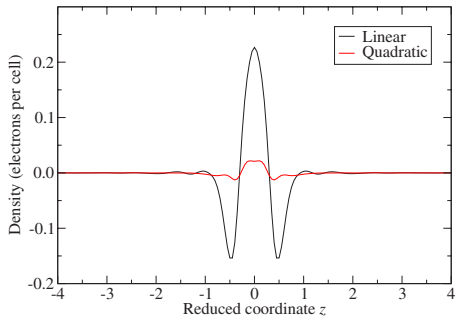

FIG. 8. (Color online) Linear and quadratic density responses to a monatomic Ga perturbation of $Al_{0.5}Ga_{0.5}As$.

$$\Delta n^{\alpha\alpha'}_{\mathbf{R}\mathbf{R}'}(\mathbf{x}) = \frac{1}{2}\frac{\partial^2 n(\mathbf{x})}{\partial\theta^{\alpha}_{\mathbf{R}}\partial\theta^{\alpha'}_{\mathbf{R}'}}. \tag{6b}$$

The total density may then be reconstructed from the power series,

$$n(\mathbf{x}) = n^{(0)}(\mathbf{x}) + n^{(1)}(\mathbf{x}) + n^{(2)}(\mathbf{x}) + \cdots, \tag{7a}$$

$$n^{(1)}(\mathbf{x}) = \sum_{\alpha,\mathbf{R}}{}' \theta^{\alpha}_{\mathbf{R}}\Delta n^{\alpha}_{\mathbf{R}}(\mathbf{x}), \tag{7b}$$

$$n^{(2)}(\mathbf{x}) = \sum_{\alpha,\mathbf{R}}{}' \sum_{\alpha',\mathbf{R}'}{}' \theta^{\alpha}_{\mathbf{R}}\theta^{\alpha'}_{\mathbf{R}'}\Delta n^{\alpha\alpha'}_{\mathbf{R}\mathbf{R}'}(\mathbf{x}), \tag{7c}$$

where $n^{(0)}(\mathbf{x})$ is the density of the reference crystal. In the present work, this power series was truncated at the second order. The derivatives (6) were calculated by the direct supercell method,[43] in which (for example) the $Al_{0.5}Ga_{0.5}As$ reference crystal was perturbed by replacing one or two $Al_{0.5}Ga_{0.5}$ atoms with either $Al_{0.55}Ga_{0.45}$ or $Al_{0.45}Ga_{0.55}$. See the Appendix for further details.

The linear and quadratic density responses to a monatomic perturbation in $Al_{0.5}Ga_{0.5}As$ are shown in Fig. 8. The functions plotted here are planar averages of $\Delta n^{\alpha}_{\mathbf{R}}(\mathbf{x})$ and $\Delta n^{\alpha\alpha}_{\mathbf{R}\mathbf{R}}(\mathbf{x})$ for the case α=Ga. Since the supercell used here is the same as that of the (001) superlattice in Figs. 4 and 5, the perturbation actually consists of a *plane* of Ga atoms (or rather an infinitesimal perturbation toward Ga). This perturbation has D_{2d} symmetry, so the planar average shown in Fig. 8 is symmetric. Note that the monatomic quadratic response is about an order of magnitude smaller than the linear response.

The quadratic response to a diatomic perturbation in $Al_{0.5}Ga_{0.5}As$ is shown in Fig. 9. This perturbation also has D_{2d} symmetry, so its qualitative features are similar to those of the monatomic response (except that it is centered on an anion plane rather than a cation plane). The perturbation planes in Fig. 9 are separated by one monolayer (i.e., $\frac{1}{2}a$). The calculations performed in this paper include diatomic perturbations out to a separation of four monolayers. Perturbations beyond this distance were neglected because they yield contributions of less than 0.1 meV.

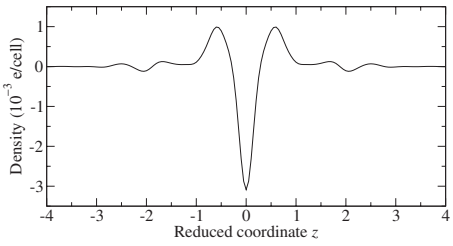

FIG. 9. Quadratic density response to a diatomic Ga-Ga perturbation of $Al_{0.5}Ga_{0.5}As$. The perturbations are applied at $z=-0.5$ and $z=+0.5$.

The second-nearest-neighbor diatomic response shown in Fig. 9 is the leading contribution to Γ_1-X_1 coupling in GaAs/AlAs (001) superlattices. A comparison with Fig. 8 shows that this term is nearly 2 orders of magnitude smaller than the linear response, which is the leading contribution to Γ_1-X_3 coupling. Hence, these results indicate that Γ_1-X_1 coupling is indeed much smaller than Γ_1-X_3 coupling, as suggested in Refs. 31 and 32.

Qualitatively different results are obtained for the perturbation of two different atoms of the $In_{0.765}Ga_{0.235}As_{0.5}P_{0.5}$ reference crystal in Fig. 10. This diatomic perturbation has symmetry C_{2v}, so its planar average is not symmetric. It therefore has a nonvanishing dipole moment, which is readily apparent from the figure. The dipole generates a rapid change in the Hartree potential, as shown in Fig. 10(b). In order to satisfy the periodic boundary conditions, this steplike change in potential must be accompanied by a macroscopic electric field extending over the entire supercell.[63,64] This field polarizes the reference crystal, producing small

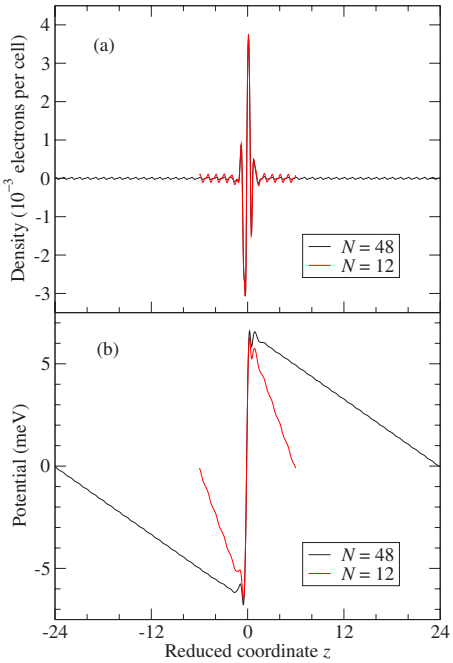

FIG. 10. (Color online) Quadratic response to a diatomic As-Ga perturbation of $In_{0.765}Ga_{0.235}As_{0.5}P_{0.5}$: (a) electron density and (b) Hartree potential. The perturbations are applied at $z=-0.5$ (As) and $z=+0.5$ (Ga). Results are given for supercells with periods N=48 and N=12.

oscillations in the density and potential[65] that are visible away from $z=0$. Figure 10 shows calculations performed on two supercells, one with 48 monolayers (the same as the superlattice in Figs. 6 and 7) and one with 12 monolayers. It is evident that the polarization amplitude is inversely proportional to the period of the supercell.

It is these dipole moments in the diatomic quadratic response that are responsible for the slight interface asymmetry and macroscopic electric field referred to in the discussion of Fig. 6. Such effects have been studied previously in Refs. 66 (numerically) and 10 (analytically).

VI. MULTIPOLE MOMENTS AND LOCALIZATION

A. Definition of multipole moments

As discussed in the Introduction, slowly varying envelope functions probe the density and potential only in a small neighborhood of each reciprocal-lattice vector $\mathbf{G}$. It is therefore convenient to introduce long-wavelength approximations for the density and potential responses in the form of power-series expansions of $n(\mathbf{k}+\mathbf{G})$ and $v(\mathbf{k}+\mathbf{G})$ with respect to $\mathbf{k}$. The expansion coefficients can be found numerically by several methods. One is a simple polynomial fitting of the discrete fast Fourier transform (FFT). Another is to use analytical manipulations of the FFT to obtain expansion coefficients in the form of multipole moments.

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. Hence,

$$n(\mathbf{k}+\mathbf{G}) = \frac{1}{N_F}\sum_{\mathbf{x}} n(\mathbf{x}) e^{-i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}} \tag{8a}$$

$$= \sum_{l=0}^{\infty} (-ik)^l n_l(\mathbf{G}), \tag{8b}$$

where $\mathbf{x}$ is a coordinate on the FFT grid, N_F is the number of such points, and the expansion coefficients are

$$n_l(\mathbf{G}) = \frac{1}{l!N_F}\sum_{\mathbf{x}} z^l n(\mathbf{x}) e^{-i\mathbf{G}\cdot\mathbf{x}} \tag{9a}$$

$$= \frac{1}{l!}\sum_{\mathbf{k}} [\zeta_l(\mathbf{k}-\mathbf{G})]^* n(\mathbf{k}), \tag{9b}$$

where $\zeta_l(\mathbf{k})$ is the FFT of z^l. The coefficient $n_l(\mathbf{G})$ is, to within a factor of $l!$, just the multipole moment[45,61] of order 2^l of the function $n(\mathbf{x})e^{-i\mathbf{G}\cdot\mathbf{x}}$.

B. Order of terms included

For small k, the power series (8b) can be approximated accurately by including only a few terms with small l. This paper adopts the approximation scheme defined in Refs. 10 and 11, in which the power series for the linear and quadratic potential responses $v^{(1)}(\mathbf{k}+\mathbf{G})$ and $v^{(2)}(\mathbf{k}+\mathbf{G})$ are terminated at the upper limits

$$l_v^{(1)} = 2, \quad l_v^{(2)} = 0. \tag{10a}$$

These power series can be used for the pseudopotential and exchange-correlation potential responses, which are analytic functions of $\mathbf{k}$. However, the Hartree potential [Eq. (1)] is nonanalytic, so the expansion (8b) must be applied to the density $n(\mathbf{k}+\mathbf{G})$ instead. To obtain an accuracy equivalent to Eq. (10a), the upper limits for the density power series are extended when $G=0$:

$$l_n^{(1,2)}(\mathbf{G}) = l_v^{(1,2)} + 2\delta_{\mathbf{G0}}, \tag{10b}$$

because of the k^{-2} factor in the Hartree potential [Eq. (1)].

When $G\neq 0$, the Hartree potential is an analytic function of $\mathbf{k}$ (for small k), and the power series expansion of $|\mathbf{k}+\mathbf{G}|^{-2}$ generates the following Hartree corrections to the analytic potential coefficients:

$$\Delta v_0^{(1,2)}(\mathbf{G}) = \frac{4\pi}{G^2} n_0^{(1,2)}(\mathbf{G}), \tag{11a}$$

$$\Delta v_1^{(1)}(\mathbf{G}) = \frac{4\pi}{G^2}\left[n_1^{(1)}(\mathbf{G}) - \frac{i2G_z}{G^2} n_0^{(1)}(\mathbf{G}) \right], \tag{11b}$$

$$\Delta v_2^{(1)}(\mathbf{G}) = \frac{4\pi}{G^2}\left[n_2^{(1)}(\mathbf{G}) - \frac{i2G_z}{G^2} n_1^{(1)}(\mathbf{G}) + \left(\frac{1}{G^2} - \frac{4G_z^2}{G^4}\right) n_0^{(1)}(\mathbf{G}) \right]. \tag{11c}$$

When $G=0$, analytic potential contributions are also obtained from the terms

$$\Delta v_l(\mathbf{0}) = -4\pi n_{l+2}(\mathbf{0}) \quad (l \geqslant 0), \tag{12}$$

where $0\leqslant l\leqslant 2$ for the linear response and $l=0$ for the quadratic response. In a superlattice, the only nonanalytic terms arise from the density monopole, dipole, and quadrupole coefficients at $G=0$. The quadrupole term is given by $4\pi n_2(\mathbf{0})\delta_{\mathbf{k0}}$, which merely contributes an overall constant shift in potential. The monopole term is zero for the isovalent perturbations considered here, while the linear dipole term vanishes for perturbations with T_d or D_{2d} symmetry. The quadratic dipole term is, however, nonzero for diatomic perturbations with C_{2v} symmetry, as shown above in Fig. 10.

C. Localization

Since the FFT grid is finite, the multipole moment $n_l(\mathbf{G})$ always has a well defined value. Nevertheless, if the response $n(\mathbf{x})$ is not well localized, the value of $n_l(\mathbf{G})$, although finite, does not converge to a meaningful value in the limit of large supercells. It is therefore necessary to check in each case whether the localization is sufficient to use Eq. (9) for the desired values of l.

The localization of the response is not merely a numerical detail; it is fundamental to the physics of the insulating state.[67,68] Phenomenological envelope-function models typically assume that the heterostructure Hamiltonian can be expressed as a power series in $\mathbf{k}$, but the validity of such an assumption crucially depends on the localization of the re-

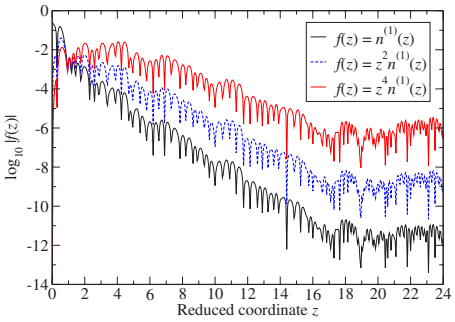

FIG. 11. (Color online) Logarithm-scale plot of linear density response in Fig. 8.

sponse in an insulator.[10,11] If the response is not localized, the entire foundation for the power-series expansion of the Hamiltonian breaks down. Establishing that the response is localized is therefore a key element in the justification of the standard phenomenological approach to envelope-function theory. Numerical error inevitably masks the underlying localization at some point, but it is important to ensure that this error is insignificant for all practical purposes.

This is done in Fig. 11 for the linear density response of $Al_{0.5}Ga_{0.5}As$ in the cases $l=0$, 2, and 4. The response is quasiexponentially localized over nearly 10 orders of magnitude, to a distance of about 17 monolayers from the perturbation. At this point, it flattens out due to the nonanalyticity of the potential-energy smoothing function in Fig. 1. If the potential energy is not smoothed, the quasiexponential localization extends for another order of magnitude, at which point numerical error takes over. However, the localization obtained here is clearly sufficient to calculate accurate values of the linear quadrupole and hexadecapole moments. On the other hand, if the plane-wave cutoff in the kinetic energy is not smoothed, $z^4 n^{(1)}(z)$ is not adequately localized, and the hexadecapole moment ($l=4$) has a nonsensical value for the supercell considered here.

The linear and quadratic terms in those parts of the local potential response that are expected to be localized (i.e., the ionic pseudopotential and the exchange-correlation potential) are shown in Fig. 12. Here, the features seen in the linear response for $|z| \gtrsim 10$ are the remnants of Gibbs oscillations in the local ionic pseudopotential that remain even after smoothing of the plane-wave cutoff. (These Gibbs oscillations are not seen in the quadratic response because the ionic pseudopotential is purely linear.) The localization here is not quite as good as for the density, but it is still sufficient to calculate the desired multipole moments.

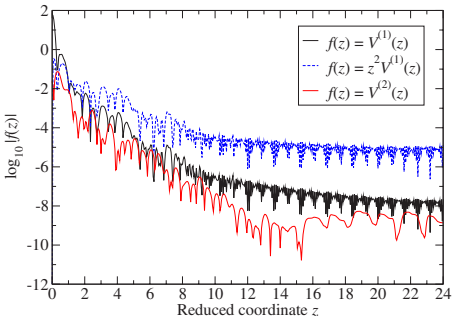

FIG. 12. (Color online) Logarithm-scale plot of linear and quadratic pseudopotential plus exchange-correlation potential for a monatomic Ga perturbation of $Al_{0.5}Ga_{0.5}As$.

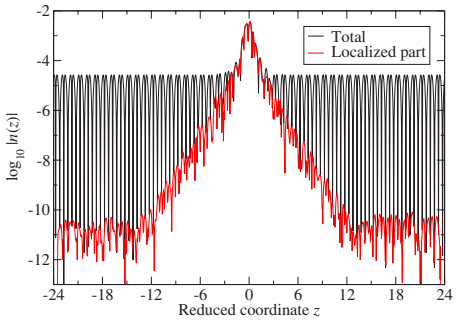

FIG. 13. (Color online) Logarithm-scale plot of the diatomic density response in Fig. 10(a). Both the total density and its localized part (excluding the periodic part) are shown.

The logarithm-scale plot in Fig. 13 of the quadratic density response for C_{2v} diatomic perturbations has additional features of interest. The response is not well localized, but this is a physical effect arising from the periodic oscillations shown in Fig. 10. The periodic part can be separated from the rest by evaluating it in the unit cell farthest from the origin. Subtracting the periodic part from the total leaves a remainder that should be localized. This is confirmed in Fig. 13, which shows both the total quadratic response and the localized remainder.

The localized part of the diatomic response can be used to obtain multipole moments as above. The periodic part is even simpler, because it merely modifies the values of coefficients in the bulk Hamiltonian (as discussed in greater detail below).

D. Calculated multipole moments

The results obtained from the calculated density multipole moments $n_l(\mathbf{G}=\mathbf{0})$ are shown in Figs. 14 and 15 for $Al_{0.5}Ga_{0.5}As$ and $In_{0.765}Ga_{0.235}As_{0.5}P_{0.5}$, respectively. The

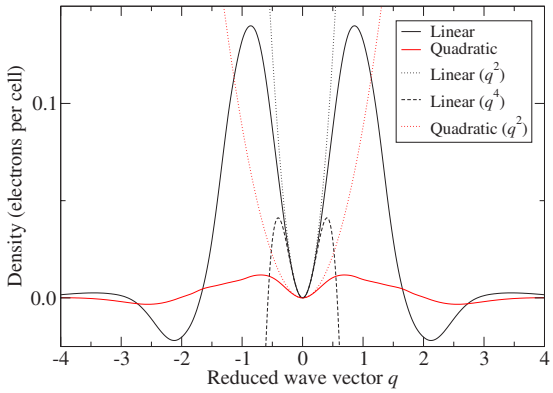

FIG. 14. (Color online) Linear and quadratic density responses to a monatomic Ga perturbation of $Al_{0.5}Ga_{0.5}As$. The solid lines are Fourier transforms of the functions in Fig. 8. The dotted and dashed lines are quadratic and quartic polynomials defined by the quadrupole and hexadecapole moments.

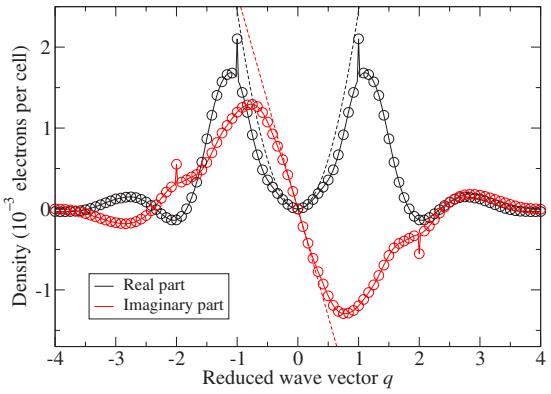

FIG. 15. (Color online) Fourier transform of the diatomic density response in Fig. 10(a). Solid lines are for a supercell with $N=48$ monolayers, while circles are for $N=12$. The FFT is scaled by N so that the results are comparable. Dashed lines are quadratic (real) and linear (imaginary) polynomials corresponding to the calculated quadrupole and dipole moments.

truncated power-series expansions for the **k**-space density response are clearly valid only for small wave vectors. For the C_{2v} perturbation in Fig. 15, the reduced symmetry generates a nonvanishing imaginary part in the FFT. This part has a nonzero slope at the origin, leading to the dipole moment observed previously in Fig. 10.

There are also noticeable spikes in the FFT at nonzero reciprocal-lattice vectors. These correspond to the periodic part of the response discussed above (Sec. VI C), whose physical origin is the polarization of the reference crystal by the macroscopic electric field[63] in Fig. 10. The amplitude of the spikes is obtained most accurately by the method described above, in which the periodic part of the quadratic response is evaluated in coordinate space in the unit cell farthest from the origin and then Fourier transformed. The resulting Fourier series coefficients occur only at the bulk reciprocal-lattice vectors **G**, which means that their contribution to the envelope-function Hamiltonian[12] has the same formal structure as the bulk $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian of the reference crystal. However, these bulk polarization terms have the C_{2v} symmetry of the dipole, not the T_d symmetry of the reference crystal. As shown in Ref. 12, the net effect of the C_{2v} dipoles and their associated bulk polarization is to contribute about one-third of the total splitting of the quasidegenerate X and Y valence states (which would be degenerate under D_{2d} symmetry) in $In_{0.53}Ga_{0.47}As/InP$ superlattices.

VII. ERROR IN THE LINEAR AND QUADRATIC APPROXIMATIONS

A. Electron density and Hartree potential

To test the validity of truncating the power series in the nonlinear response [Eq. (7)], the effects of various truncations were examined by comparing them with the exact superlattice density and potential given previously in Figs. 4–7. The approximations are generally good enough that they are difficult to distinguish visually from the exact results. Therefore, only the error in the approximations is plotted here.

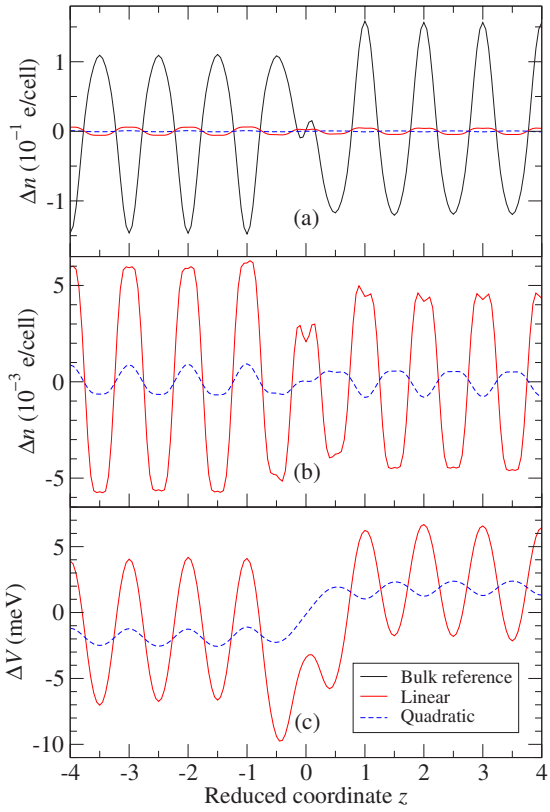

FIG. 16. (Color online) Errors in the electron density and Hartree potential of a GaAs/AlAs superlattice under various approximations: (a) Errors in the reference-crystal density and the densities obtained in the linear and quadratic approximations. (b) Expanded view of the linear and quadratic errors in (a). (c) Errors in the linear and quadratic approximations to the Hartree potential.

The results for the GaAs/AlAs superlattice are shown in Fig. 16. Part (a) shows the error when the superlattice density is approximated by the periodic reference-crystal density, comparing it with the error in the linear and quadratic approximations. The latter are both small on the scale of the reference-crystal error (which is itself small on the scale of the reference density), so the linear and quadratic errors are plotted on an expanded scale in part (b). The corresponding errors in the Hartree potential are shown in part (c). It can be seen that the quadratic approximation is quite accurate, with an error of only about ±2 meV.

The errors in the linear and quadratic approximations for the $In_{0.53}Ga_{0.47}As/InP$ superlattice are shown in Fig. 17. Here, the error in the quadratic Hartree potential has an oscillation with an amplitude of about 3 meV superimposed on a shift of about ±1.5 meV.

B. Truncation of multipole expansions

Using multipole expansions for the density and potential introduces an additional source of error. To illustrate this, Fig. 18 shows the electron density and Hartree potential in coordinate space calculated directly from the FFT of the truncated multipole expansion (with $0 \leqslant l \leqslant 4$ for the linear density and $0 \leqslant l \leqslant 2$ for the quadratic density). Note that the

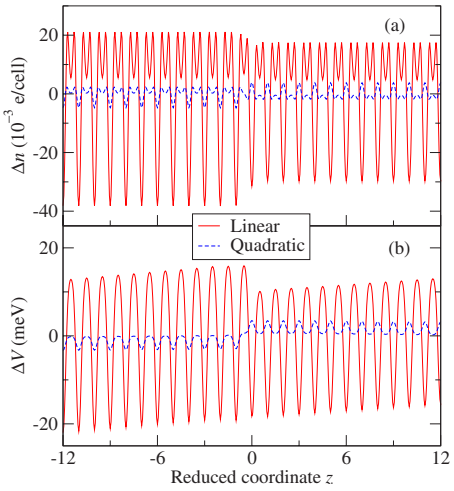

FIG. 17. (Color online) Errors in the electron density and Hartree potential of an $In_{0.53}Ga_{0.47}As/InP$ superlattice under various approximations: (a) Errors in the linear and quadratic densities. (b) Errors in the linear and quadratic Hartree potentials.

scale of this figure is much larger than any of the others. In fact, in a continuous coordinate space, these quantities would be derivatives of the Dirac δ function. Clearly, these multipole expansions yield a very poor approximation of the original density and potential.

However, in an envelope-function model, accuracy is only required at small wave vectors, so it is more appropriate to compare the macroscopic average of the multipole expansions with the macroscopic average of the exact results. This is done in Fig. 19, which shows the errors in the macroscopic density and Hartree potential for a GaAs/AlAs superlattice. Here, the macroscopic averages were calculated using $\sigma = 3d$ in Eqs. (3), which limits momentum transfers to roughly $\frac{1}{3}$ of the bulk Γ-X distance.[69] Such a limit is appropriate for the slowly varying envelope functions considered in the following paper.[12] The macroscopic errors are virtually identical in bulk, but the truncated multipole approximation introduces some additional error near the interfaces. This interface error is negligible for the case $\sigma = 3d$ shown here, although it becomes arbitrarily large in the limit of small σ. This merely reflects the fact that a truncated multipole expansion is valid only for slowly varying envelopes.

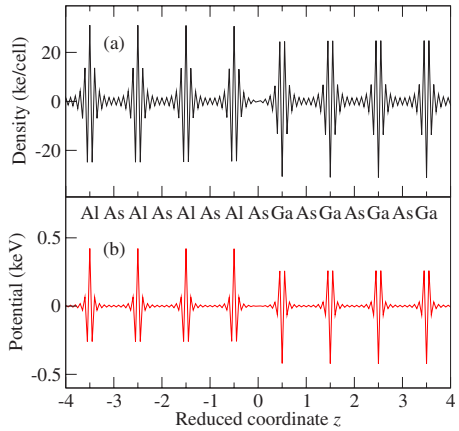

FIG. 18. (Color online) Sum of the linear and quadratic perturbations for (a) the electron density and (b) the Hartree potential constructed from quadrupole and hexadecapole moments in a GaAs/AlAs superlattice. The perturbations are finite only because a discrete coordinate grid is used. The hexadecapole terms are dominant here, although this is no longer true after macroscopic averaging.

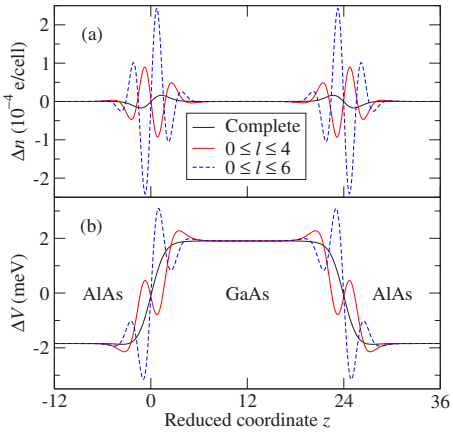

FIG. 19. (Color online) Errors in the macroscopic (a) electron density and (b) Hartree potential of a GaAs/AlAs superlattice under various approximations. The functions shown are the errors in the complete quadratic approximation to the macroscopic density and potential (i.e., the macroscopic averages of the quadratic functions in Fig. 16 with $\sigma = 3d$) and the corresponding errors when the linear and quadratic responses are generated from truncated multipole expansions. The labels $0 \leqslant l \leqslant 4$ and $0 \leqslant l \leqslant 6$ refer to the power series for the linear density; for the quadratic density, the corresponding limits are $0 \leqslant l \leqslant 2$ and $0 \leqslant l \leqslant 4$, respectively.

The corresponding results for $In_{0.53}Ga_{0.47}As/InP$ are shown in Fig. 20. Again, the error generated by the truncation of the multipole series vanishes away from the interfaces. The magnitude of the error is similar to that in GaAs/AlAs. For $In_{0.53}Ga_{0.47}As/InP$, however, there is a slight error in the bulk macroscopic electric fields even before the multipole expansion is truncated.

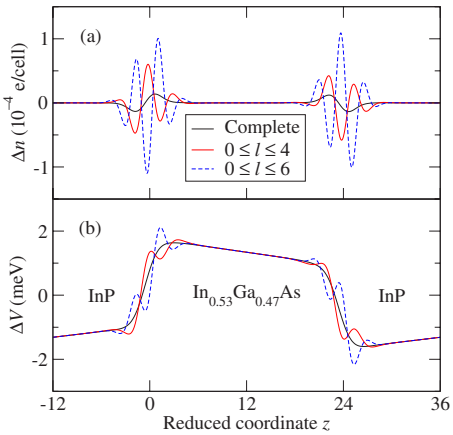

FIG. 20. (Color online) Errors in the macroscopic (a) electron density and (b) Hartree potential of an $InP/In_{0.53}Ga_{0.47}As$ superlattice: comparison of complete quadratic errors with truncated multipole errors, as in Fig. 19.

To test the convergence of the truncated multipole expansions used here, the power series (8b) was extended to include terms up to $l=6$ in the linear density and up to $l=4$ in the quadratic density. The resulting macroscopic errors (shown in Figs. 19 and 20) actually increase slightly, but the change in the Hartree potential remains negligible. For the case $\sigma=6d$ (not shown here), which includes the most important part of the potential for slowly varying envelopes, the macroscopic Hartree error of both truncated expansions is visually almost indistinguishable from that of the complete quadratic response.

As a direct practical test of the convergence of the multipole expansion, the extended limits (including also terms of the same order in the ionic pseudopotential and exchange-correlation potential) were used to recalculate the superlattice subband structures shown in Sec. V of the following paper.[12] No difference was observable in any case. Therefore, the truncated multipole expansions defined in Eqs. (10) are very well converged for slowly varying envelopes.

C. Bulk energy eigenvalues at Γ

The error analysis to this point has focused on the Hartree potential, with the multipole truncation analysis limited to the macroscopic ($G=0$) case. However, the Hamiltonian for slowly varying envelope functions is also sensitive to the $G\neq 0$ terms, the quadratic-response error in the exchange-correlation potential, and the multipole truncation error in this potential and the local and nonlocal ionic pseudopotentials. (There is no quadratic error in the latter because the ionic pseudopotential is purely linear.)

These sources of error are not analyzed separately here, but to leading order, their net effect is to generate an error in the bulk energy bands of Figs. 2 and 3. This error can be calculated easily at the Γ point by adding linear and quadratic bulk perturbations to the Γ Hamiltonian of the reference crystal, diagonalizing the resulting matrices, and comparing the approximate Γ energies with the exact values for the given materials. For example, the quadratic error in the Γ_{15v} energy is −0.4 meV for GaAs and +0.5 meV for AlAs. Combining these values with the corresponding bulk macroscopic errors of +1.9 and −1.8 meV for the Hartree potential in Fig. 19(b), there is a net error of +1.5 meV in the bulk valence-band edge of GaAs and a net error of 2.7 meV (or 0.6%) in the GaAs/AlAs valence-band offset.

Likewise, the quadratic error in the Γ_{15v} energy is +3.15 meV for $In_{0.53}Ga_{0.47}As$ and −4.45 meV for InP. The nominal positions of the interfaces in Fig. 20 are $z=-0.25$ and $z=23.75$. The errors in the macroscopic Hartree potential at the positions halfway between the interfaces are +1.35 meV for $In_{0.53}Ga_{0.47}As$ and −1.32 meV for InP. Therefore, there is a net quadratic error of +4.5 meV in the bulk valence-band edge of $In_{0.53}Ga_{0.47}As$ and a net error of 10.3 meV (or 4.4%) in the $In_{0.53}Ga_{0.47}As/InP$ valence-band offset. Note that the quadratic error in each component of the $In_{0.53}Ga_{0.47}As/InP$ valence-band offset is only about 1%, but the total quadratic error is larger because the Hartree shift and bulk energy shift have opposite signs, while the corresponding errors have the same signs.

VIII. CONCLUSIONS

This paper has investigated numerically a quadratic-response approximation[10,11] for the self-consistent one-electron potential within 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 absolute error for the Γ_{15} valence states of about 2 meV in GaAs/AlAs and 5 meV in $In_{0.53}Ga_{0.47}As/InP$. As shown in the following paper,[12] the principal effect of this error for slowly varying envelope functions is 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 generated some additional error (due primarily to the truncation of the linear density response) in a narrow region near the interfaces. This error was confirmed to be negligible for slowly varying envelope functions.

Although the quadratic response is about an order of magnitude smaller than the linear response, it must be included if the self-consistent heterostructure potential is to be predicted to an accuracy of better than a few tens of meV. The quadratic response is also qualitatively important because of its different symmetry. 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. As shown in the following paper,[12] these terms, which have C_{2v} symmetry, produce a significant fraction of the splitting of the quasidegenerate ground state in such systems. Therefore, to provide a realistic description of such interface-induced symmetry-breaking effects, existing empirical pseudopotentials would need to be modified to include these contributions.

IX. LIMITATIONS AND POSSIBLE EXTENSIONS OF THE THEORY

A. Quasiparticle effects

The LDA model system used in the present study is, of course, not completely realistic because it fails to provide accurate predictions of the conduction-band states. This model could be improved by including quasiparticle self-energies[3,4] and by replacing norm-conserving pseudopotentials with projector augmented waves[70–72] (which would provide a better description of core electrons while remaining within a pseudopotential-like formalism needed for the applicability of the present perturbation scheme). Despite the well-known inaccuracies of LDA conduction bands,[6] LDA wave functions do have a very high overlap with *GW* quasiparticle wave functions.[73] This similarity of the wave functions was used by Wang and Zunger[27] to construct empirical pseudopotentials with accurate energy gaps and effective masses using only a small (first-order) perturbation of the

LDA Hamiltonian. This strongly suggests that, even though a numerical implementation of quasiparticle self-energies would be substantially more complicated than the present LDA model, the basic qualitative conclusion of the present study—namely, that the quadratic response provides an accurate approximation of the one-electron heterostructure potential—would remain valid in this theory also.

B. Atomic relaxation

Another useful extension of the present study would be to allow the superlattice ions to relax to their equilibrium positions. For example, it is well known that, although $In_{0.53}Ga_{0.47}As$ and InP have the same bulk lattice constant, the bond lengths for the InAs or $In_{0.53}Ga_{0.47}P$ bonds that form at a heterojunction are significantly different.[46,66,74] The resulting displacement of the ionic planes from their present "clamped" reference-crystal positions would generate additional dipole, quadrupole, and higher-order moments that have not been included here. The lowest-order contribution is a dipole term arising from the displacement of a single plane of ions, which generates a potential shift[46,75] $\Delta V = 4\pi Z^* u/A\epsilon_\infty$, where Z^* is the transverse effective ionic charge, u is the displacement, A is the area per ion in the plane, and ϵ_∞ is the static electronic dielectric constant. The qualitative effects of such strain-induced interface dipoles are the same as those of the purely "chemical" dipoles in the clamped-ion system studied here, but accurate predictions of the properties of physical heterostructures could only be achieved by including both contributions.

Atomic relaxation is readily calculated using any of several structural optimization algorithms already included in the ABINIT software package used here.[7–9] The main difficulty in handling strained superlattices is the same as that encountered in strained bulk crystals—namely, that the electronic boundary conditions for the strained system are not the same as those for the unstrained system, so that the strain cannot be treated by the ordinary methods of perturbation theory.[76,77] This difficulty can be surmounted by an extension[78,79] of the usual bulk coordinate-transformation technique[76] in which the coordinates of one (or possibly more than one) atom per unit cell in the strained system are mapped onto the coordinates of the same atom in the reference crystal. The coordinate transformation is straightforward in principle but carries with it an additional heavy layer of mathematical formalism. Therefore, in the interests of simplicity, this complication was excluded by fiat in the present work, although it may be taken up in future studies.

C. Quantum wires or dots

The final extension discussed here is the application of the present methods to quantum wires or quantum dots. This paper was limited to superlattices primarily for practical reasons, since, as mentioned in Sec. III A, these calculations are intended to be used in the following paper[12] to compare envelope-function predictions of valence subband structure directly with numerical LDA results. Such a direct comparison is possible for superlattices containing on the order of 100 atoms, which are large enough for slowly varying envelopes to exist and yet small enough for LDA plane-wave calculations to be feasible. However, this clearly would not be practicable (at present) for a large quantum dot containing $\sim 10^6$ atoms.

Nevertheless, even if direct comparisons are currently out of reach for large quantum dots and wires, the quadratic-response approximation can still be used to construct envelope-function Hamiltonians for such systems. Inspection of Fig. 15 shows that the response to monatomic and diatomic perturbations is essentially independent of the size of the supercell as long as the supercell is large enough for interactions between perturbations in adjacent cells to be negligible. Therefore, one can construct the Hamiltonian from calculations on relatively small supercells containing one or two atomic perturbations (as opposed to the perturbations comprising one or two planes of atoms in the present study). Indeed, such calculations have already been performed for monatomic perturbations in 16-atom fcc supercells in Refs. 45 and 46, although somewhat larger supercells would probably be necessary for an accurate treatment of diatomic perturbations.

It should be noted, however, that the one-dimensional multipole expansions developed in Sec. VI for superlattices are not directly applicable in quantum wires or dots. For such cases, one would need to use the more general three-dimensional multipole expansions developed in Refs. 10 and 11. These would lead to additional terms in the envelope-function Hamiltonian, as discussed in Ref. 11 and in footnote 30 of the following paper.[12]

ACKNOWLEDGMENT

This work was supported by Hong Kong RGC Grant No. 600905.

APPENDIX

If we consider for simplicity a function of only two θ variables, the density and potential are written as quadratic functions

$$f(\theta_1,\theta_2) = c + c_1\theta_1 + c_2\theta_2 + c_{11}\theta_1^2 + c_{22}\theta_2^2 + 2c_{12}\theta_1\theta_2. \tag{A1}$$

Here, the coefficients c, c_i, and c_{ij} were calculated by the direct supercell method.[43] First, the self-consistent density and potential were found for a supercell consisting of the reference crystal with a monatomic perturbation $\theta_i = \pm\Delta_i$ (e.g., in $Al_{0.5}Ga_{0.5}As$, one $Al_{0.5}Ga_{0.5}$ atom was replaced by either $Al_{0.55}Ga_{0.45}$ or $Al_{0.45}Ga_{0.55}$). This provides the values of the coefficients

$$c = f_{00}, \tag{A2a}$$

$$c_1 = \frac{f_{10} - f_{\bar{1}0}}{2\Delta_1}, \tag{A2b}$$

$$c_{11} = \frac{f_{10} + f_{\bar{1}0} - 2f_{00}}{2\Delta_1^2}, \tag{A2c}$$

where $f_{\bar{1}0}=f(-\Delta_1,0)$, etc. Next, diatomic perturbations were used to find the value of c_{12}. Two possible formulas were considered:

$$c_{12}=\frac{f_{11}-f_{1\bar{1}}-f_{\bar{1}1}+f_{\bar{1}\bar{1}}}{8\Delta_1\Delta_2}, \tag{A3a}$$

$$c_{12}=\frac{f_{11}-f_{10}-f_{01}+f_{00}}{2\Delta_1\Delta_2}. \tag{A3b}$$

The symmetric formula (A3a) is presumably more accurate, but very little difference was found in comparison with the asymmetric formula (A3b). Since the latter requires only one-fourth the computation time and storage of the former, Eq. (A3b) was used for all calculations reported here.

*phbaf@ust.hk

[1] M. L. Cohen and T. K. Bergstresser, Phys. Rev. **141**, 789 (1966).
[2] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976); **30**, 4828 (1984).
[3] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, in *Solid State Physics*, edited by H. Ehrenreich and F. Spaepen (Academic, San Diego, 2000), Vol. 54, pp. 1–218.
[4] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[5] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[6] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).
[7] 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).
[8] 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).
[9] The ABINIT code is a common project of the Université Catholique de Louvain, Corning Incorporated, and other contributors, http://www.abinit.org
[10] B. A. Foreman, Phys. Rev. B **72**, 165344 (2005).
[11] B. A. Foreman, Phys. Rev. B **72**, 165345 (2005).
[12] B. A. Foreman, following paper, Phys. Rev. B **76**, 045327 (2007).
[13] A. C. Marsh and J. C. Inkson, J. Phys. C **17**, 6561 (1984).
[14] C. Mailhiot and D. L. Smith, Phys. Rev. B **33**, 8360 (1986).
[15] S. Brand and D. T. Hughes, Semicond. Sci. Technol. **2**, 607 (1987).
[16] J.-B. Xia, Phys. Rev. B **39**, 3310 (1989).
[17] D. L. Smith and C. Mailhiot, Rev. Mod. Phys. **62**, 173 (1990).
[18] J. P. Cuypers and W. van Haeringen, J. Phys.: Condens. Matter **4**, 2587 (1992).
[19] G. C. Dente and M. L. Tilton, J. Appl. Phys. **86**, 1420 (1999).
[20] G. C. Dente and M. L. Tilton, Phys. Rev. B **66**, 165307 (2002).
[21] R. Magri and A. Zunger, Phys. Rev. B **68**, 155329 (2003).
[22] W. Andreoni, A. Baldereschi, and R. Car, Solid State Commun. **27**, 821 (1978).
[23] W. Andreoni and R. Car, Phys. Rev. B **21**, 3334 (1980).
[24] J.-B. Xia, Phys. Rev. B **38**, 8358 (1988).
[25] P. Friedel, M. S. Hybertsen, and M. Schlüter, Phys. Rev. B **39**, 7974 (1989).
[26] K. A. Mäder and A. Zunger, Phys. Rev. B **50**, 17393 (1994).
[27] L.-W. Wang and A. Zunger, Phys. Rev. B **51**, 17398 (1995).
[28] H. Fu and A. Zunger, Phys. Rev. B **55**, 1642 (1997).
[29] A. O. E. Animalu and V. Heine, Philos. Mag. **12**, 1249 (1965).
[30] L. Kleinman, Phys. Rev. B **24**, 7412 (1981).
[31] L.-W. Wang and A. Zunger, Phys. Rev. B **56**, 12395 (1997).
[32] B. A. Foreman, Phys. Rev. Lett. **81**, 425 (1998).
[33] É. E. Takhtamirov and V. A. Volkov, JETP **90**, 1063 (2000).
[34] G. Bastard, J. A. Brum, and R. Ferreira, in *Solid State Physics*, edited by H. Ehrenreich and D. Turnbull (Academic, Boston, 1991), Vol. 44, pp. 229–415.
[35] B. Laikhtman, Phys. Rev. B **46**, 4769 (1992).
[36] M. V. Kisin, B. L. Gelmont, and S. Luryi, Phys. Rev. B **58**, 4605 (1998).
[37] F. T. Vasko and A. V. Kuznetsov, *Electronic States and Optical Transitions in Semiconductor Heterostructures* (Springer, New York, 1999).
[38] A. V. Rodina, A. Yu. Alekseev, Al. L. Efros, M. Rosen, and B. K. Meyer, Phys. Rev. B **65**, 125302 (2002).
[39] A. V. Rodina and A. Yu. Alekseev, Phys. Rev. B **73**, 115312 (2006).
[40] E. L. Ivchenko, A. Yu. Kaminski, and U. Rössler, Phys. Rev. B **54**, 5852 (1996).
[41] É. E. Takhtamirov and V. A. Volkov, JETP **89**, 1000 (1999).
[42] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*, Springer Tracts in Modern Physics Vol. 191 (Springer, Berlin, 2003).
[43] R. Resta, S. Baroni, and A. Baldereschi, Superlattices Microstruct. **6**, 31 (1989).
[44] S. Baroni, R. Resta, and A. Baldereschi, in *Proceedings of the 19th International Conference on the Physics of Semiconductors, Warsaw, 1988*, edited by W. Zawadzki (Institute of Physics, Polish Academy of Sciences, Warsaw, 1988), pp. 525–528.
[45] 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), Vol. 206, pp. 251–272.
[46] M. Peressi, S. Baroni, A. Baldereschi, and R. Resta, Phys. Rev. B **41**, 12106 (1990).
[47] M. Peressi, S. Baroni, R. Resta, and A. Baldereschi, Phys. Rev. B **43**, 7347 (1991).
[48] L. Colombo, R. Resta, and S. Baroni, Phys. Rev. B **44**, 5572 (1991).
[49] S. Baroni, M. Peressi, R. Resta, and A. Baldereschi, in *Proceedings of the 21st International Conference on the Physics of Semiconductors, Beijing, 1992*, edited by P. Jiang and H.-Z. Zheng (World Scientific, Singapore, 1992), pp. 689–696.
[50] M. Peressi and S. Baroni, Phys. Rev. B **49**, 7490 (1994).
[51] M. Peressi, B. Montanari, S. Baroni, and E. Molinari, in *Proceed-*

ings of the 23rd International Conference on the Physics of Semiconductors, Berlin, 1996, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), pp. 943–946.
[52] B. Montanari, M. Peressi, S. Baroni, and E. Molinari, Appl. Phys. Lett. **69**, 3218 (1996).
[53] R. Magri and A. Zunger, Phys. Rev. B **65**, 165302 (2002).
[54] Although it was not originally expressed in this way, the pseudopotential interpolation scheme defined in Eq. (A4) of Magri and Zunger (Ref. 53) is equivalent to the inclusion of diatomic building blocks.
[55] C. Hartwigsen, S. Goedecker, and J. Hutter, Phys. Rev. B **58**, 3641 (1998).
[56] J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979); **13**, 3095 (1980).
[57] M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 3259 (1982).
[58] M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 5668 (1982).
[59] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001).
[60] A. Baldereschi, S. Baroni, and R. Resta, Phys. Rev. Lett. **61**, 734 (1988).
[61] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).
[62] F. N. H. Robinson, *Macroscopic Electromagnetism* (Pergamon, Oxford, 1973).
[63] If the supercell has a nonvanishing dipole moment, a crystal consisting of a finite number of supercells in vacuum does not satisfy periodic boundary conditions because of the macroscopic electric field (depolarization field) generated by macroscopic polarization charges on the surface of the crystal. This depolarization field can be canceled by embedding the crystal in a perfect metal, thus establishing a physical context for the use of periodic boundary conditions (see Refs. 60 and 64). The macroscopic electric field obtained by averaging over the volume of a supercell is therefore identically zero under periodic boundary conditions. However, the macroscopic electric field obtained by averaging over the volume of a primitive zinc-blende unit cell is not everywhere zero. This is apparent in Fig. 10(b), where the macroscopic part of the electric field in the regions away from the dipole at the origin is just the field due to the induced surface charge on the metal (plus the linear response to this field).
[64] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson, Phys. Rev. B **53**, 1814 (1996).
[65] K. Kunc and R. Resta, Phys. Rev. Lett. **51**, 686 (1983).
[66] R. G. Dandrea, C. B. Duke, and A. Zunger, J. Vac. Sci. Technol. B **10**, 1744 (1992).
[67] W. Kohn, Phys. Rev. **133**, A171 (1964).
[68] D. Vanderbilt and R. D. King-Smith, Phys. Rev. B **48**, 4442 (1993).
[69] At this value of k, the Gaussian function in Eq. (3b) has a value of $\exp(-\pi/4) \simeq 0.456$.
[70] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[71] N. A. W. Holzwarth, G. E. Matthews, R. B. Dunning, A. R. Tackett, and Y. Zeng, Phys. Rev. B **55**, 2005 (1997).
[72] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[73] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[74] M. S. Hybertsen, J. Vac. Sci. Technol. B **8**, 773 (1990).
[75] R. M. Martin and K. Kunc, Phys. Rev. B **24**, 2081 (1981).
[76] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974).
[77] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).
[78] M. G. Burt, in *Band Structure Engineering in Semiconductor Microstructures*, edited by R. A. Abram and M. Jaros, NATO Advanced Studies Institute, Series B: Physics (Plenum, New York, 1989), Vol. 189, pp. 99–109.
[79] L.-W. Wang and A. Zunger, Phys. Rev. B **59**, 15806 (1999).