phys. stat. sol. (b) **224**, No. 3, 727–734 (2001)

Pseudopotential Theory of Semiconductor Quantum Dots

ALEX ZUNGER

National Renewable Energy Laboratory, Golden, CO 80401, USA

(Received July 31, 2000; accepted October 2, 2000)

Subject classification: 71.15.Dx; 73.21.La; S5.11; S5.12; S7.11; S7.12; S8.11; S8.12

This paper reviews our pseudopotential approach to study the electronic structure of semiconductor quantum dots, emphasizing methodology ideas and a survey of recent applications to both "free-standing" and "semiconductor embedded" quantum-dot systems.

1. Introduction: What Are the Bottlenecks Limiting Accurate Theoretical Modeling of Semiconductor Quantum Dots Analyses of the experimental observations on quantum dots reveal that a theory of quantum dots must encompass both single-particle electronic structure and many-body physics.

Regarding the *single-particle electronic* structure, the available methods are either insufficient or impossibly complicated. The prevailing single-particle method for quantum nanostructures – the effective-mass approximation (EMA) and its "$\mathbf{k} \cdot \mathbf{p}$" generalization – were carefully tested by us [1–3], and were found to be insufficiently accurate for our purposes. Specifically, the underlying continuum-like EMA misses the atomistic nature of quantum dots. Indeed, in attempting to expand the dot's wavefunction in terms of only the VBM and CBM of the infinite bulk solid at the Brillouin zone center ($k = 0$) the EMA produces but a rough sketch of the dot's microscopic wavefunctions. The limited accuracy originating from the neglect of multi-band coupling prevents reliable calculations of wavefunction expectation values such as energy levels [1–4], exchange energies [5] and interelectronic Coulomb repulsion [6] in *free-standing* quantum dots. Of course, one can *fit* experimental results within the EMA approach. In other words, this theory can *accommodate* experimental facts, but rarely *predicts* them. For *semiconductor-embedded* dots such as InAs/GaAs, the $\mathbf{k} \cdot \mathbf{p}$ method is much better [26]. The other "single-particle" method – the local density approximation (LDA) – is limited in terms of the number of atoms it can handle. While very small Si particles can be calculated [7], real (e.g., self-assembled dot structures) require $\sim 10^3$–10^5 (including the barrier material) atoms for their description, well outside the range of LDA capabilities. Furthermore, the "LDA errors" in the band gaps and effective masses preclude contact with spectroscopic data on dots.

Regarding the *many-body description* needed, the highly precise Quantum-Monte-Carlo (QMC) approach can unfortunately be used for large dots (10^3–10^5 atoms) only with EMA wavefunctions [8], not with, e.g., pseudopotential wavefunctions. On the other hand, the Configuration Interaction (CI) approach for dots [9] can be used to predict many-body (e.g., multiplet) effects with accurate atomistic wavefunctions.

The "Theoretical Technology" needed to address the electronic structure of quantum dots should thus be capable of incorporating *multi-band coupling* (unlike the EMA), describe *arbitrary shapes and materials*, include *realistic surface* or interface effects, and capture many-body *(multiplet)* effects for objects containing *10^3–10^5 atoms* (including the "barrier" between dots in a supercell). Until recently, such theoretical technology was unavailable.

2. Overcoming Theory Bottlenecks: The "Order-N" Pseudopotential Configuration-Interaction Approach We have developed an accurate, general-purpose electronic structure approach suitable for addressing the leading physical questions pertaining to quantum dots. The method has two parts: A: The single-particle problem and B: the many-body problem. In what follows, we describe the main ideas behind this new method, followed by a discussion of its limitations and our approach to these limitations.

2.1 A: The single-particle problem

1. The shape, size and composition of the dot are accepted as "input", i.e., there is no attempt to predict them from "growth models". This decision stems from the observation that such variables are often not controlled by the thermodynamics of the growth, but that post-growth, it is possible to measure the approximate shape, size and composition, thus use them as "input" to the theory. We can do arbitrary 3D shapes and composition profiles [4].

2. Atoms are relaxed to their strain-minimizing positions using an atomistic "force field" fit to LDA [10]. We currently use a three-parameter generalized "Keating model" that fits C_{11}, C_{12}, C_{44} and reproduces LDA relaxation of energies of related ordered compounds. We avoid the less-accurate continuum elasticity approach for strain minimization, in which the atomistic point-group symmetry of the dot is often overlooked [4].

3. The single-particle screened pseudopotential is fit to bulk solids: The total screened pseudopotential is written as a superposition of atomically screened potentials at sites R_n for atom type α,

$$V_{\text{ext}} + V_{\text{scr}} = \sum_{n,\alpha} V_\alpha(r - R_{n,\alpha})\,. \tag{1}$$

The screened potentials $\{V_\alpha\}$ are fit to the measured band structure and anisotropic effective masses of the underlying bulk materials, as well as to the LDA-calculated bulk wavefunctions, deformation potentials, and band offsets [11, 12]. We thus avoid "band gap LDA errors". We have developed a few schemes permitting us to achieve "LDA quality wave functions" [11, 12] with "experimental quality band gaps and masses". Spin–orbit splitting is included. For certain elements α, a non-local potential is used. A separate potential is fit for the surface atoms [13], assuring reproduction of the electronic structure obtained from model "slab calculations".

4. The wavefunctions are expanded in plane-waves, thus affording a microscopic description (not just a macroscopic, "envelope-function" description), as well as representing multi-band coupling (i.e., many bands at Γ) and inter-valley (Γ–X–L) couplings [10].

5. The pseudopotential-plane wave Hamiltonian is diagonalized incredibly rapidly via the order-N "Folded Spectrum Method" (FSM) [14, 15]. Our philosophy is that much of the physical interest in quantum dots centers around energy levels in the vicinity of the band gap, but that current computational approaches do not take advantage of this fact. For example, a 10000 atom Si dot has 20000 occupied levels, but only the few highest are "interesting". Still, quantum mechanics forces us to calculate all of the 19999 levels below the VBM (= level number 20000) or the CBM (= level number 20001). This is wasteful, not to mention time consuming. The Folded Spectrum Method [14] thus "folds" the lowest ≈19999 levels to very high energies, leaving the VBM or CBM as the lowest solutions of the modified Hamiltonian $(H - A)^2$. Finding the eigensolutions to the squared Hamiltonian is both sufficient (as the VBM and CBM are captured) and very fast. These solutions are identical to those that are obtained by diagonalizing brute-force the full Hamiltonian. In practice, we search for ≈10 eigensolutions in the neighborhood of the VBM and CBM. Another modification of this method, the Linear Expansion in Bloch Bands (LCBB) [16], affords an even faster diagonalization, applicable literally to million-atom quantum dots. These methods run on a parallel computer platform for extra speed, as described in Ref. [15].

2.2. B: The many-body problem

1. Inter-electronic integrals are computed numerically: Given the single-particle wave functions $\{\psi_i\}$ from the pseudopotential calculation, we calculate the screened inter-electronic Coulomb J_{ij} and exchange K_{ij} [9, 17], integrals numerically using a multi-grid approach. The screening includes both ionic and electronic parts. The latter is described phenomenologically via literature models, and depends on the size R of the quantum dot, as well as on the inter-particle separation $|r_e - r_h|$. More details are given in Refs. [5, 6, 9].

2. Many-body effects are included via configuration interaction: The wave function of an exciton (or multi-exciton) is described as a superposition of a limited number of Slater determinants corresponding to many-particle excitations. This “configuration-interaction” many-body Hamiltonian is diagonalized to yield the many-body energy and the multiplet structure of the dot. Convergence is examined by increasing the number of determinants (currently up to 40000). More details are given in Ref. [9].

2.3 What have we neglected?

1. The exact atomic positions in dots are often unknown. The atomic positions used are only as good as the experimental input. However, it now appears that inside large (>1000-atom) dots, the measured interatomic distances are very close to the bulk values. Hence, uncertainties in exact positions are confined to the surface area. We use realistic models for the surface atomic structure based on LDA surface (slab) calculations. Nevertheless, experimental uncertainties in shape and composition profile is a real limitation for the theory.

2. The single-particle wave functions are not self-consistent. This is a controlled error, since self-consistency can be implemented. Our tests, documented in Ref. [6] demonstrate only small (<5 meV) errors. Nevertheless, some critical parts of the dot (e.g., next to impurity atom) do need self-consistency.

3. The CI expansion is truncated. Although we perform what is normally termed “Full CI” (all Slater determinants that evolve from the bound states in the dot are accounted for), this neglects the contribution of continuum states to the correlation energy. The Quantum-Monte Carlo (QMC) method does include continuum states, but can be currently applied only to single-band EMA models. Thus, to get an estimate of our errors, we applied both the QMC and the CI approaches to a single-band model of a dot with finite potential barriers [18]. A typical exciton total energy (for a dot with 40 Å diameter, $m_e = 0.1$, $m_h = 0.5$) is 140 meV. The correlation energy is 2 meV. QMC gets the full 2 meV, while our CI gets only 1 meV of the correlation energy. Note that we can apply CI to a multi-band (e.g., pseudopotential) Hamiltonian, while QMC is currently limited to EMA.

4. The screening function is phenomenological. We have tested all screening functions available in the literature and find some differences between them. We are continuing to search for the “best” screening function.

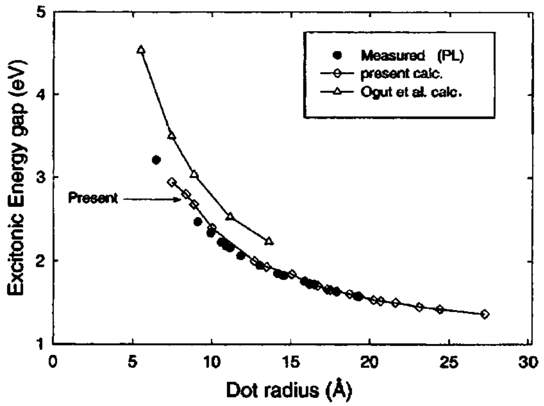

Fig. 1. Calculated (◇) and measured (●) excitonic gaps in Si dots (Ref. [20] and references therein). The triangles show LDA results of other authors

3. Recent Pseudopotential Results on the Theory of Quantum Dots

The electronic structure of quantum dots involves:

(i) "High-energy problems" (i.e., the 1–3 eV range excitonic energies);

(ii) "intermediate-energy problems" (i.e., the ~0.1 eV range charging energies); and

(iii) "low-energy problems" (i.e., the 0.01–0.001 eV splittings due to electron–hole exchange and different multi-exciton multiplets.

3.1 The main accomplishments to date for the "high-energy problems"

1. *Reproduced accurately the excitonic gap for "free-standing" Si dots* (Ref. [19, 20]), *InP dots* (Ref. [12, 13]) *and CdSe dots* (Ref. [21, 22]). *Our results are closer to experiment than other models* even though other models (e.g., EMA and **k** · **p**) often fit the experimental quantum dot data directly. For CdSe, we calculated not only the lowest exciton, but also the seven excited excitons above it [22]. The agreement with experiment was remarkable, as illustrated in Fig. 1 (Si), Fig. 2 (InP), and Fig. 3 (CdSe). No fit is involved beyond the bulk bands.

2. *Established new quantum-size scaling laws.* Textbook models predict that band gaps Eg scale with quantum-size as $E_g \sim R^{-2}$, whereas electron–hole Coulomb energies J_{eh} scale as $J_{e,h} \sim R^{-1}$. These results hold in the single-band effective-mass approach which retains kinetic energies of confined particles but neglect potential-energy-induced multi-band coupling. Our multi-band results reveal new, material-dependent scaling laws, e.g., for Si [19, 20] $E_g \sim R^{-1.4}$ while for InP [12, 13] $E_g \sim R^{-1.3}$, whereas Coulomb energies scale as [6] $J_{eh} \sim R^{-0.8}$. These new scaling laws are important as they reveal, for the first time, the interplay between kinetic-energy (EMA-like) effects and multi-band coupling effects.

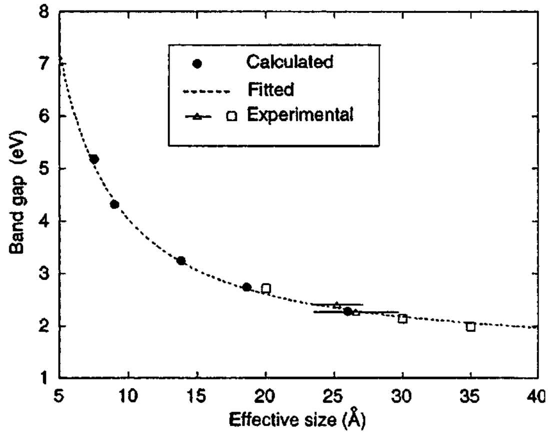

Fig. 2. Calculated (●) and measured (□ and line) excitonic gaps in InP dots (Ref. [5] and references therein)

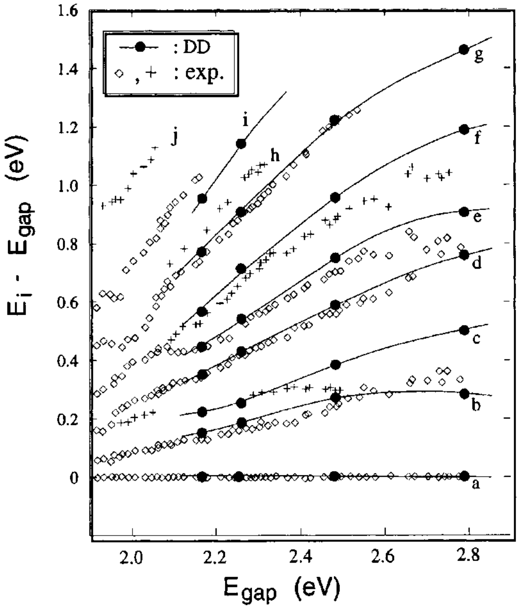

Fig. 3. Calculated (●) and measured (other symbols) excitonic transitions in CdSe dots (Ref. [22] and references therein)

3. Established an energy-level model for the "semiconductor embedded" self-assembled InAs/GaAs dots [23–29]. Perhaps the most studied quantum dots are the "self-assembled", strained InAs/GaAs dots. We provided the first comprehensive theory for interpreting a broad range of experimentally determined energy levels in such structures resulting from PL and PLE measurements, as well as inter-subband spectroscopy and Stark effect. The results include the determination of excitonic energies; spacings between electron levels, hole levels; electron and hole binding energies; wetting-layer energies, the magnitude of the excitonic dipols, and their pressure dependence. We believe that such results could become a "standard energy level model" for these types of dots. We show in Fig. 4 the realistic wavefunctions that result from such atomistic calculations. Note that they do not look at all as EMA (s, p, d) states.

3.2 The main accomplishments to-date for the "intermediate-energy problems"

1. Predicted the electron-addition energies in freestanding InAs [30, 31] *and CdSe* [32] *dots*, in excellent agreement with recent STM-carrier injection experiments. Figure 5 illustrates the predictability of the model for both addition energies and quasi-particle gaps of InAs. More importantly, by calculating the total energies for different electronic configurations of the charged dot, we have.

2. Established deviations from Hund's rule as well as "spin-blockade" [32]. Such deviations represent the competition between gain of exchange energy (favoring Hund's rule) and Coulomb repulsion. Our recently published predictions [32] await experimental testing.

3.3 The main accomplishments to-date, for the "low-energy problems"

Established theory of electron–hole exchange in dots [5, 9, 17]. We tackled a long-standing fundamental problem in quantum dot physics: Does the electron–hole exchange have just a short-range part (resulting from wavefunction overlap) or could it have a long-range component? In bulk solids, the long-range component results from dipole interactions which average to zero in a spherical dot. It was thus believed that in dots, there is but a short-range exchange. Indeed the measured e–h exchange was always fit to a short-range formula with the scaling $\Delta_x \sim R^{-3}$.

Our accurate wavefunctions permitted an assessment of long versus short-range components. We found in Si, InP, and CdSe an unequivocal long-range component which

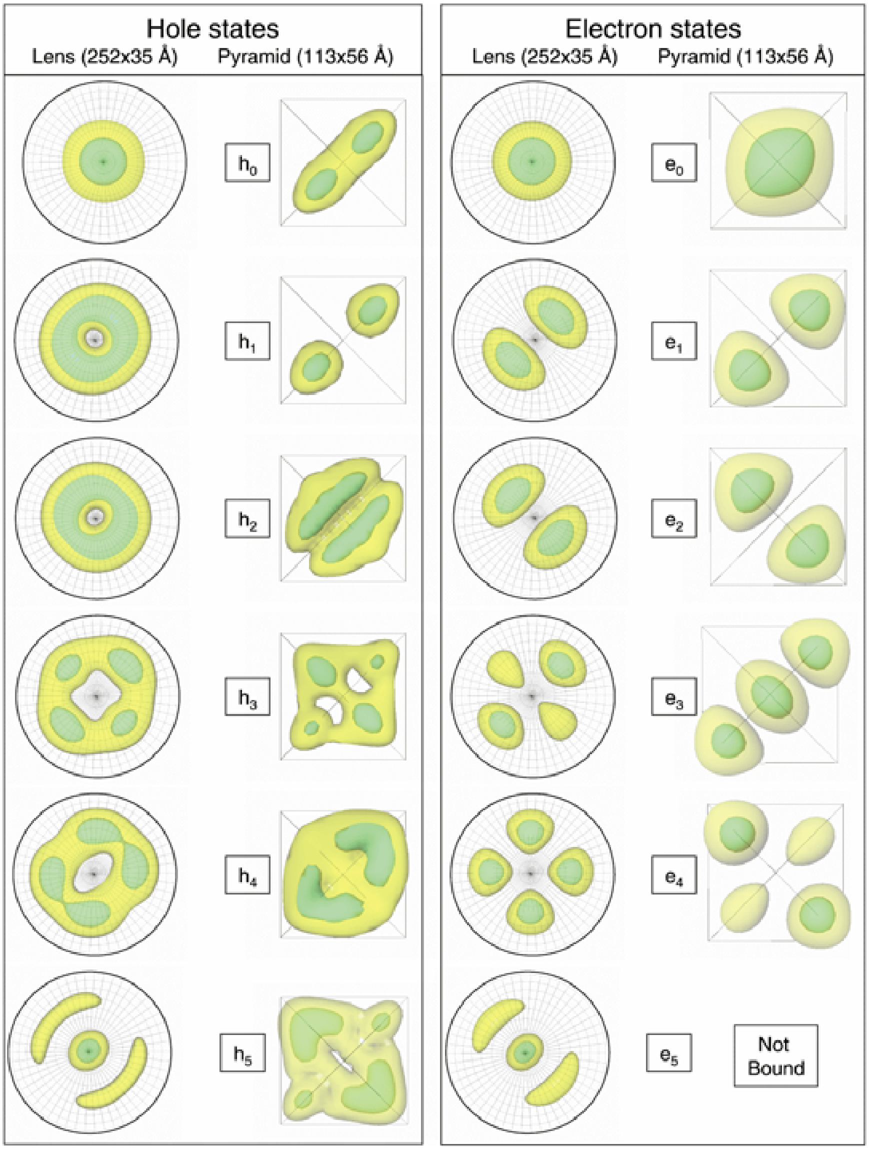

Fig. 4. Color top view of calculated electron and hole wavefunctions squared for lens-shaped and pyramidal InAs dot embedded in GaAs; Ref. [23]

dominates over the short-range component [5, 9, 17]. Our subsequent analytical study [17] revealed that the long-range exchange originates from the previously neglected monopol-monopol interactions. Our calculated exchange splitting versus size agreed very well with the measured data for InP [9] and CdSe [9] (without any fit). This work established a new view on the microscopic nature of electron–hole exchange interactions in nanostructures. A similar recent study on multi-excitons in InAs/GaAs dots [33] establishes the nature of many-body multiplet effects in the spectra.

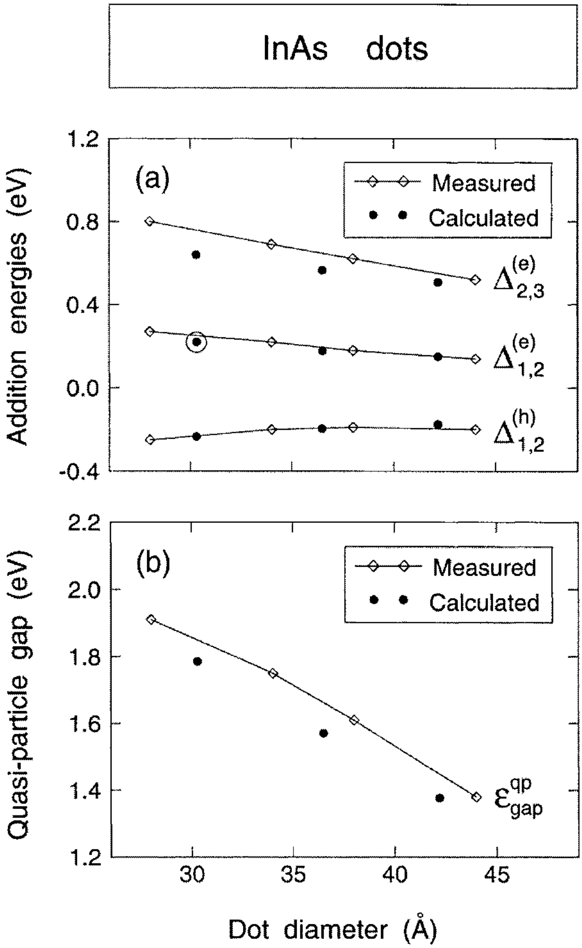

Fig. 5. a) Calculated (●) and measured electron (e) and hole (h) energies of InAs dots. b) Calculated and STM measured quasiparticle gaps; from Ref. [30] and references therein

4. Summary The pseudopotential approach is a practical method for describing atomistic and microscopic aspects of the electronic structure of nanostructures. When such aspects are not important (e.g., large systems), an atomistic approach is not needed, and conventional continuum approaches are fine. Furthermore, continuum approaches can be conveniently used to *fit* known data even for small nanostructures, although the *predictive* ability of these methods is questionable.

References

[1] D. M. Wood, A. Zunger, and D. Gershoni, Europhys. Lett. **33**, 383 (1996).
D. M. Wood and A. Zunger, Phys. Rev. B **53**, 7949 (1996).
[2] L. W. Wang and A. Zunger, Phys. Rev. B **54**, 11417 (1996).
L. W. Wang, A. J. Williamson, A. Zunger, H. Jiang, and J. Singh, App. Phys. Lett. **76**, 339 (2000).
[3] H. Fu, L. W. Wang, and A. Zunger, Appl. Phys. Lett. **71**, 3433 (1997).
H. Fu, L. W. Wang, and A. Zunger, Phys. Rev. B **57**, 9971 (1998).
[4] C. Pryor, J. Kim, L. W. Wang, A. Williamson, and A. Zunger, J. Appl. Phys. **83**, 2548 (1998).

[5] H. Fu and A. Zunger, Phys. Rev. B **57**, R15064 (1998).
O. I. Micic, H. M. Cheong, H. Fu, A. Zunger, J. R. Sprague, A. Mascarenhas, and A. J. Nozik, J. Phys. Chem. B **101**, 4904 (1997).
F. A. Reboredo, A. Franceschetti, and A. Zunger, Phys. Rev. B **61**, 13073 (2000).
[6] A. Franceschetti and A. Zunger, Phys. Rev. Lett. **78**, 915 (1997).
[7] B. Delley and E. F. Steigmeier, Phys. Rev. B **47**, 1397 (l993).
[8] J. Shumway et al., Physica E **8**, 260 (2000).
[9] A. Franceschetti, H. Fu, L. W. Wang, and A. Zunger, Phys. Rev. B **60**, 1819 (1999).
A. Zunger, MRS Bull. **23**, 35 (1998) (special issue).
[10] A. Zunger, 23rd Internat. Conf. Physics of Semiconductors, Vol. 2, World Scientific Publ. Co., Singapore, 1996 (pp. 1341–1348).
L. W. Wang, A. Franceschetti, and A. Zunger, Phys. Rev. Lett. **78**, 2819 (1997).
[11] K. Mader, L. W. Wang, and A. Zunger, Phys. Rev. Lett. **74**, 2555 (1995).
L. W. Wang and A. Zunger, Phys. Rev. B **51**, 17398 (1995).
[12] H. Fu and A. Zunger, Phys. Rev. B **55**, 1642 (1997).
A. Zunger, in: Real Materials, Eds. J. R. Chelikowsky and S. G. Louie, Kluwer Academic Publ., Boston 1996 (pp. 173–187).
[13] H. Fu and A. Zunger, Phys. Rev. B **56**, 1496 (1997).
[14] L. W. Wang and A. Zunger, J. Chem. Phys. **100**, L2394 (1994).
L. W. Wang and A. Zunger, J. Phys. Chem. **98**, 2158 (1994).
[15] A. Canning, L. W. Wang, A. Williamson, and A. Zunger, J. Comput. Phys. **160**, 29 (2000).
[16] L. W. Wang and A. Zunger, Phys. Rev. B **59**, 15806 (1999).
[17] A. Franceschetti, L. W. Wang, H. Fu, and A. Zunger, Phys. Rev. B **58**, R13367 (1998).
[18] J. Shumway, A. Franceschetti, and A. Zunger, Phys. Rev. B, submitted.
[19] A. Franceschetti, L. W. Wang, and A. Zunger, Phys. Rev. Lett. **83**, 1269 (1999).
L. W. Wang and A. Zunger, Phys. Rev. Lett. **73**, 1039 (1994).
[20] F. Reboredo, A. Franceschetti, and A. Zunger, Appl. Phys. Lett. **75**, 2972 (1999).
L. W. Wang and A. Zunger, in: Studies in Surface Science and Catalysis, Eds. P. V. Kamat and D. Meisel, Elsevier Publ. Co., Amsterdam 1996 (pp. 161–207).
A. Zunger and L. W. Wang, Appl. Surf. Sci. **102**, 350 (1996).
[21] L. W. Wang and A. Zunger, Phys. Rev. B **53**, 9579 (1996).
[22] L. W. Wang and A. Zunger, J. Phys. Chem. **102**, L6449 (1998).
[23] A. Williamson, L. W. Wang, and A. Zunger, Phys. Rev. B **62**, 12963 (2000).
[24] J. Kim, L.W. Wang and A. Zunger, Phys. Rev. B **56**, R15541 (1997).
J. Kim, L. W. Wang, and A. Zunger, Phys. Rev. B **57**, R9408 (1998).
A. J. Williamson, A. Zunger, and A. Canning, Phys. Rev. B **57**, R4253 (1998).
[25] H. Fu and A. Zunger, Phys. Rev. Lett. **80**, 5397 (1998).
[26] A. J. Williamson and A. Zunger, Phys. Rev. B **58**, 6724 (1998).
[27] L. W. Wang, J. Kim, and A. Zunger, Phys. Rev. B **59**, 5678 (1999).
[28] A. J. Williamson and A. Zunger, Phys. Rev. B **59**, 15819 (1999).
[29] A. J. Williamson, A. Franceschetti, H. Fu, L. W. Wang, and A. Zunger, J. Electron. Mater. **28**, 414 (1999).
A. Franceschetti, A. Williamson, and A. Zunger, J. Phys. Chem. B **104**, 3398 (2000).
[30] A. Franceschetti and A. Zunger, Phys. Rev. B **62**, 2614 (2000).
[31] A. Franceschetti and A. Zunger, Appl. Phys. Lett. **76**, 1731 (2000).
[32] A. Franceschetti and A. Zunger, Europhys. Lett. **50**, 243 (2000); Phys. Rev. B **62**, R16287 (2000).
[33] A. J. Williamson, A. Franceschetti, and A. Zunger, Europhys. Lett. **53**, 59 (2001).