SERIES
SCIENCE

QUANTUM THEORY OF REAL MATERIALS

Edited
by

James R. Chelikowsky
University of Minnesota

and

Steven G. Louie
University of California at Berkeley

KLUWER ACADEMIC PUBLISHERS
Boston / Dordrecht / London

13

FIRST PRINCIPLES AND SECOND PRINCIPLES (SEMIEMPIRICAL) PSEUDOPOTENTIALS

Alex Zunger
National Renewable Energy Laboratory, Golden, CO 80401

Abstract: Screened, transferable atomic pseudopotentials were developed 30 years ago by M. L. Cohen by adjusting the potential to reproduce the observed bulk electronic energies (the "Empirical Pseudopotential Method" or EPM). These potentials are extremely useful in interpreting and guiding the spectroscopy of (a few atoms/cell) bulk materials. The recent emergence of detailed spectroscopy of ($\sim$ 1000-atom) semiconductor nanostructures has created a significant demand for pseudopotentials combining EPM-quality energy levels with LDA-quality wavefunctions. The reason is that for such systems LDA approaches are both too costly and insufficiently accurate, while effective mass-based approaches lack detailed band structure information. We can now improve upon the traditional EPM by a two step process: First, we invert a set of self-consistently determined screened LDA potentials for a range of bulk crystal structures and unit cell volumes, thus determining spherically-symmetric and structurally-averaged atomic potentials (SLDA). These potentials reproduce the LDA band structure to better than 0.1 eV, over a range of crystal structures and cell volumes. Second, we adjust the SLDA to reproduce observed excitation energies. We find that the adjustment represents but a small perturbation over the SLDA potential, so that the ensuing fitted "semiempirical Pseudopotential method" (SEPM) potential still reproduces a $>$ 99.9% overlap with the original LDA wavefunctions despite the excitation energies being distinctly non-LDA. We illustrate the method to Si and CdSe in a range of crystal structures, finding excellent agreement with the experimentally determined band energies, optical spectra, ϵ_2(E), static dielectric constants, deformation potentials, and, at the same time, LDA-quality wavefunctions.

1. INTRODUCTION: THE BASIC IDEOLOGY

The "Empirical Pseudopotential Method" (EPM) developed by Marvin Cohen and collaborators [1, 2, 3] marked an historic milestone in bringing together the spectroscopy of interband transitions in solids with the quantum structure of the solid. The formalism was elegant and simple: The Schrödinger equation for band i at wavevector $\mathbf{k}$ was written in terms of a lattice sum over sites α at $\mathbf{R}_\alpha$, of screened atomic pseudopotentials v^α_{EPM} whose principal Fourier

components were adjusted to fit the positions of the main energy bands:

$$\left\{-\frac{1}{2}\nabla^2 + \sum_{\mathbf{R}_\alpha} v^\alpha_{EPM}(\mathbf{r}-\mathbf{R}_\alpha)\right\}\psi_{i\mathbf{k}}(\mathbf{r}) = \epsilon_{i\mathbf{k}}\ \ \psi_{i\mathbf{k}}(\mathbf{r}). \tag{1}$$

The calculated band energies and the ensuing predicted optical response $\epsilon_2(\omega)$ quickly became the main tool for interpreting the electronic structure of many solids in terms of transferable atomic quantities $v^\alpha_{EPM}(r)$.

In the density functional approach[4, 5], on the other hand, one treats a Schrodinger equation with a *distinct*, non-transferable (self-consistently determined) screening potential V_{HXC}:

$$\left\{-\frac{1}{2}\nabla^2 + \hat{V}_{ext}(\mathbf{r}) + V_{HXC}[\rho(\mathbf{r})]\right\}\psi_{i\mathbf{k}}(\mathbf{r}) = \epsilon_{i\mathbf{k}}\ \ \psi_{i\mathbf{k}}(\mathbf{r}). \tag{2}$$

Here, $\hat{V}_{ext}(\mathbf{r})$ is a fixed external (possibly pseudo) potential determining the chemical and structural identity of the system, and $\rho(\mathbf{r}) = \sum_i^{occ} |\psi_i|^2$ is the charge density of all occupied single particle states ψ_i. The external potential can be written [6, 7] as a sum of a nonlocal part and a local part

$$\hat{V}_{ext}(\mathbf{r}) = \hat{V}_{nonlocal}(\mathbf{r}) + V_{local}(\mathbf{r}) = \hat{V}_{nonlocal}(\mathbf{r}) + \sum_\alpha \sum_{R_\alpha} v^{(\alpha)}_{ps}(|\mathbf{r}-\mathbf{R}_\alpha|) \tag{3}$$

and each of these terms can be represented as a sum over atoms α at lattice sites $\mathbf{R}_\alpha$ of fixed atomic pseudopotentials [1]. (We use V and v to denote, respectively, crystalline and atomic potentials). While $V_{ext}(\mathbf{r})$ can be constructed explicitly for each system once and for all, simply by summing over fixed *atomic* potentials, the (Coulomb + exchange + correlation) screening potential $V_{HXC}[\rho(\mathbf{r})]$ is not a linear sum over atomic quantities, and must be obtained separately for each physical system without any system-to-system transferability. Thus, the EPM approach of Eq. (1) and the LDA approach of Eqs. (2)-(3) represent different views on the same problem: the first approach represents the solid in terms of screened and *transferable* atomic quantities, emphasizing the *spectroscopy* of the solid, while the second approach focuses on self-consistently determined *system-dependent* quantities, emphasizing the *ground state* properties (e.g., charge densities) of the solid.

Having developed in 1975-77 in collaboration with Art Freeman at Northwestern a first-principles total-energy approach for solids [8], I was biased towards the second (density functional) view. Thus, when I came to work with Marvin Cohen at Berkeley in the fall of 1977, I decided to "take a step backwards" relative to the EPM approach of Eq. (1), and determine the bare atomic pseudopotential of Eq. (3) not from fitting experiment but rather from a *microscopic quantum model* of the atom. Continuing the work I started in 1976 at Northwestern [9] with Sid Topiol and Mark Ratner, Marvin and I extracted the bare pseudopotential v^α_{ps} [rather than the screened v^α_{EPM} of Eq. (1)] from an atomic [rather than a solid state] LDA equation:

$$\left\{-\frac{1}{2}\nabla^2 + v^\alpha_{ps}(r) + V_{HXC}[\rho]\right\}\chi_{nl}(r) = \epsilon_{nl}\ \chi_{nl}(r). \tag{4}$$

The ensuing v_{ps}^{α} represented perhaps the earliest LDA-based "first principles pseudopotential" [6, 9]. We immediately applied it to total energy and band structure calculations of bulk Si [6] and bulk Mo and W [10]. These early first principles LDA pseudopotentials, analogous to what Luis Kahn and William Goddard had done earlier in the context of the Hartree-Fock theory [11], were subsequently improved by Kerker [12] and by Haman, Schluter and Chiang [13].

Our results [6, 7, 10] of 1978-79 created a dilemma: On one hand, our pseudopotentials produced (in conjunction with the just developed pseudopotential total energy method [7]) remarkable agreement with the measured *ground state* properties of Si, Mo, and W, yet the excited state properties (e.g., the Si band gap) were far inferior to what had been achieved earlier with the EPM. In the intervening 15 years we have "learned to live" with this dichotomy—we use first-principles LDA pseudopotentials whenever we need accurate ground state properties, and we use the EPM, empirical tight-binding or *ab initio* quasiparticle calculations whenever we need accurate spectroscopy.

It is now becoming clear that this 'solution' has its own problems. High resolution spectroscopy, especially of quantum (dot/wire/film) nanostructures shows that *transition matrix elements* (thus, wavefunctions) need to be accurately determined to explain the rich spectrocopy of such structures [14]. Unfortunately, the EPM is not trained to do so. There is thus a contemporary need to produce EPM-like, transferable pseudopotentials that represent well the excited state *energies*, having, at the same time, LDA-quality wavefunctions. Such pseudopotentials can then be used to understand the fascinating properties of the $\approx$ 1000-atom nanostructures for which the LDA is both too costly and insufficiently accurate, while the effective-mass band approach is inappropriate when these structures are too small. Figure 1 illustrates the calculated [15] optical absorption spectra of spherical Si particles containing up to 1500 atoms.

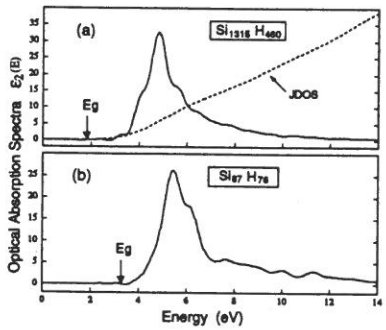


Figure 1: Optical absorption $\epsilon_2(E)$ of spherical Si quantum dots saturated by hydrogens, as obtained from new semiempirical pseudopotentials. Vertical arrows denote the minimum gap. From Ref. [15].

Such calculations are now easily performed with the new generation of (semi)empirical pseudopotentials to be described here. What made such $\mathcal{O}(10^3)$ atom nanostructure calculations feasible was the realization [15] that if one needs to know only the near band-gap states, it is not necessary to solve Eq. (1)

(which, because of the orthogonality condition forces one to solve for *all* eigenstates below the physically interesting near-gap states). Instead, one can 'fold' the spectrum of Eq. (1) about a reference energy ϵ_{ref} so that the lowest energy solutions of

$$\left(-\frac{1}{2}\nabla^2 + V_{EPM} - \epsilon_{ref}\right)^2 \psi_{i\mathbf{k}}(\mathbf{r}) = \left(\epsilon_{i\mathbf{k}} - \epsilon_{ref}\right)^2 \psi_{i\mathbf{k}}(\mathbf{r}) \tag{5}$$

are the states closest to ϵ_{ref}. By placing this 'pointer' ϵ_{ref} inside the gap, one is guaranteed to find the VBM or CBM without spending effort looking for the few thousand lower energy levels. This "folded spectrum method" (and the generalized moment method [15] used to calculate Fig.1) enables $\mathcal{O}(10^3)$ atom calculations with an effort that is linear in the system's size, *provided one has a good* V_{EPM}.

L. W. Wang and I have recently worked on the problem of finding a good V_{EPM} using a two-step process [16]. First, we have explored the possibility of constructing from Eqs. (2)-(3) *transferable* and *fixed* (i.e., screened) *spherical* and *local* atomic potentials $v^\alpha(r)$ such that the solutions of

$$\{-\frac{1}{2}\nabla^2 + \hat{V}_{nonlocal}(\mathbf{r}) + \sum_\alpha \sum_{R_\alpha} v^{(\alpha)}(|\mathbf{r} - \mathbf{R}_\alpha|)\}\tilde{\psi}_i(\mathbf{r}) = \tilde{\epsilon}_i \tilde{\psi}_i(\mathbf{r}) \tag{6}$$

reproduce to within a good approximation the solutions of the LDA Eqs. (2)-(3), i.e, $\tilde{\psi}_i$ will have large overlaps with ψ_i and $\tilde{\epsilon}_i$ will be close to ϵ_i. For simplicity, the nonlocal LDA potential $\hat{V}_{nonlocal}(r)$ is kept the same in Eqs. (2)-(3) and in Eq. (6). In the second step, we introduce changes into v^α such that the energy levels $\tilde{\epsilon}_i$ will match *experiment* (rather than LDA) while the wavefunctions $\tilde{\psi}$ will still maintain a high ($> 99\%$) overlap with the LDA wavefunctions.

Our first step involves two approximations: First, while $V_{HXC}(\mathbf{r})$ of Eq. (2) can be written rigorously as a sum of *non-spherical*, atom-centered potentials, we use instead in Eq. (6) a *spherical* representation. Second, while $V_{HXC}[\rho(\mathbf{r})]$ is system dependent [through the system's $\rho(\mathbf{r})$], we will assume that the $v^\alpha(r)$ of Eq. (6) are fixed atomic potentials and hence transferable from one system to the other. The combined nonspherical and nontransferability error will be examined by comparing the solutions $\{\epsilon_i, \psi_i\}$ of Eqs. (2)-(3) with the solutions $\{\tilde{\epsilon}_i, \tilde{\psi}_i\}$ of Eq. (6) for a few systems covering a range of coordination numbers and volumes. At this step we have a *spherical* (S) and approximately *transferable* LDA potential $v^{(\alpha)}_{SLDA}(r)$ which we will term "SLDA". We will show that this potential provides good approximations to the LDA results for bulk systems in that the band structures are reproduced to within better than 0.1 eV over a considerable energy range. Thus, the SLDA approach, like the LDA itself, suffers from poor reproduction of the *observed* excitation energies. Therefore, in the second step we will empirically adjust $v^\alpha_{SLDA}(r)$ to reproduce the *experimentally* observed excitation energies. The amount of adjustment needed (and its momentum dependence) will be explicit indicators of "LDA errors" in the Kohn-Sham effective single-particle potential. It is interesting to note that *relatively minor changes are required in* $v^{(\alpha)}_{SLDA}(r)$ *to reproduce the observed excitation energies.* Thus the resulting wavefunctions retain a large overlap with

LDA wavefunctions. This approach yields what we term the "second principles", or "semiempirical pseudopotential method" (SEPM).

This procedure permits simple, non-iterative electronic structure calculations [via Eq. (6)] for $\approx$ 1000-atom nanostructures [15]. This approach is analogous to the EPM [1, 2]. Unlike the EPM, in which $v^{(\alpha)}(r)$ was adjusted to fit the single-particle excitation spectra, our approach produces also a large (in practice $\sim$ 99.9%) overlap $< \tilde{\psi}_i | \psi_i >$ with the *LDA* wavefunctions, while reproducing *experimental* excitation energies. Furthermore, unlike the EPM, which produces only discrete form factors and is hence suitable only for a particular crystal structure and lattice constant, we will develop a *continuous* $v^{(\alpha)}_{SEPM}(r)$ which can be used for different structures and volumes with good transferability. In essence, instead of using the LDA equation of an isolated *atom* [Eq. (4)] to generate atomic pseudopotentials [6,9,11-13] we will use the LDA equations of *periodic solids* [Eqs. (2)-(3)] to extract effective, screened solid state potentials $v^{(\alpha)}(r)$ [Eq. (6)]. This approach affords a systematic procedure for generating transferable effective potentials from a set of bulk LDA calculations, retaining LDA-like wavefunctions but adjusting the potential to produce accurate excitation energies. I will illustrate the method for a covalent solid (Si) and for a more ionic case (CdSe), considering in each case a number of crystal structures and a range of unit cell volumes, thus providing information on the transferability of these potentials.

2. CONSTRUCTION OF SEMIEMPIRICAL PSEUDOPOTENIALS FROM BULK LDA CALCULATIONS

2.1 The spherical LDA potential (SLDA)

Any periodic total potential such as $\hat{V}_{ext} + V_{HXC}$ of Eq. (2) can be rigorously expressed as a sum over atom-centered *non-spherical* functions. For example, the *local* part of the total LDA potential for crystal structure σ can be written as

$$V^{(\sigma)}_{LDA}(\mathbf{r}) \equiv V^{(\sigma)}_{local}(\mathbf{r}) + V^{(\sigma)}_{HXC}(\mathbf{r}) = \sum_{\alpha} \sum_{\mathbf{R}_{\alpha,\sigma}} v^{(\alpha,\sigma)}_{LDA}(\mathbf{r} - \mathbf{R}_{\alpha,\sigma}), \qquad (7)$$

where $v^{(\alpha,\sigma)}_{LDA}(\mathbf{r})$ is a non-spherical potential that can be expanded as

$$v^{(\alpha,\sigma)}_{LDA}(\mathbf{r}) = \sum_{l} v^{(\alpha,\sigma,l)}_{LDA}(|\mathbf{r}|)\ K_l(\hat{\mathbf{r}}), \qquad (8)$$

where $K_l(\hat{\mathbf{r}})$ are the Kubic harmonics of angular momentum l for crystal structure σ. For the diamond structure, for example, the symmetry-allowed l values are 0,3,4 ... In the following, we will introduce two simplifying approximations intended to remove the dependence of $v^{(\alpha,\sigma,l)}_{LDA}(r)$ on angular momentum l and on the crystal structure σ:

(i) *The spherical approximation:* We retain only the spherically-symmetric ($l = 0$) part of Eq. (8). This means that we will write Eq. (7) without the

vector notation, i.e,

$$V_{SLDA}^{(\sigma)}(\mathbf{r}) = \sum_{\alpha} \sum_{\mathbf{R}_{\alpha,\sigma}} v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{r} - \mathbf{R}_{\alpha,\sigma}|), \tag{9}$$

where the subscript "SLDA" denotes "spherically approximated LDA". We define as $\mathbf{G}_\sigma$ the reciprocal lattice vectors of structure σ with unit cell volume Ω and as $V(\mathbf{G})$ the Fourier transform of $V(\mathbf{r})$. The α-th sublattice structure factors of structure σ is

$$S^{(\alpha,\sigma)}(\mathbf{G}_\sigma) = \frac{1}{\Omega} \sum_{\mathbf{R}_{\alpha,\sigma}} e^{i\mathbf{G}_\sigma \cdot \mathbf{R}_{\alpha,\sigma}}. \tag{10}$$

To implement the spherical approximation, we will first solve Eqs. (2)-(3) within the LDA (using *ab initio nonlocal* pseudopotentials) for a set of structures and unit cell volumes denoted collectively by $\{\sigma\}$, finding the self-consistent Fourier coefficients $V_{LDA}^{(\sigma)}(\mathbf{G}_\sigma)$ of Eq. (7). We then [following Eq. (9)] attempt the equality

$$V_{LDA}^{(\sigma)}(\mathbf{G}_\sigma) = \sum_{\alpha} S^{(\alpha,\sigma)}(\mathbf{G}_\sigma) \;\; v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|) \tag{11}$$

for each structure σ and reciprocal lattice vector $\mathbf{G}_\sigma$ and find the real-valued $v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|)$. The solution is not always exact (or even possible) because the matrix of the linear equations Eq. (11) could be singular. A simple example is when $S^{(\alpha,\sigma)}(\mathbf{G}_\alpha) = 0$ but $V_{LDA}^{(\sigma)}(\mathbf{G}_\sigma) \neq 0$. These are the "forbidden reflections". For these $\mathbf{G}_\sigma$ values, $v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|)$ is either approximated [when $S^{(\alpha,\sigma)}(\mathbf{G}_\sigma) \neq 0$, but the determinent is singular] or not calculated at all [when $S^{(\alpha,\sigma)}(\mathbf{G}_\sigma) = 0$]. We fit the points $\{v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|), |\mathbf{G}_\sigma|\}$ *for a given structure* σ to a parametrized form

$$u_{SLDA}^{(\alpha,\sigma)}(q) = \sum_{n} C_{SLDA}^{(\alpha,\sigma)}(n) \;\; e^{-(q-a_n)^2/b_n^2}. \tag{12}$$

(ii) *The structural average approximation:* We will average over a number of different structures $\{\sigma\}$. Thus, we perform the replacement

$$v_{SLDA}^{(\alpha,\sigma)}(r) \rightarrow \; < v_{SLDA}^{(\alpha)}(r) >, \tag{13}$$

where the angular brackets denote "structural average". To implement the structural average of step (ii), we include in the fit of Eq. (12) *all structures* $\{\sigma\}$. The resulting curve (now dropping the superscript σ) can be written as:

$$u_{SLDA}^{(\alpha)}(q) = \sum_{n} C_{SLDA}^{(\alpha)}(n) \;\; e^{-(q-a_n)^2/b_n^2}. \tag{14}$$

At this stage we have a spherically-symmetric and structural-averaged, screened atomic LDA potential $u_{SLDA}^{(\alpha)}$; this will be called SLDA in the following discussions. We will examine the effect of the total error by solving the band structure of Eq. (6) [replacing $v^{(\alpha)}(|\mathbf{r} - \mathbf{R}_\alpha|)$ by $u_{SLDA}^{(\alpha)}(|\mathbf{r} - \mathbf{R}_\alpha|)$] and comparing the

resulting eigenvalues $\tilde{\epsilon}_i$, wavefunctions $\tilde{\psi}_i$ and related properties with those obtained by solving self-consistently the original LDA Eqs. (2)-(3) for the same structures.

2.2 Using the SLDA to construct the "semiempirical pseudopotential method" (SEPM)

Our next task is to adjust the SLDA potential of Eq. (14) so that the ensuing wavefunctions of Eq. (6) will still retain a high overlap with the LDA wavefunctions of Eqs. (2)-(3), yet the eigenvalues will fit the *experimental* (or quasiparticle calculated) excitations. In this process, the *ab initio* nonlocal potential $\hat{V}_{nonlocal}(\mathbf{r})$ of Eqs. (3) and (6) is kept unchanged as in the reference LDA calculations. Furthermore, we use the same analytical form of Eq. (14) for both $u^{(\alpha)}_{SLDA}(q)$ and $u^{(\alpha)}_{SEPM}(q)$: the coefficients will change from $C^{(\alpha)}_{SLDA}(n)$ to $C^{(\alpha)}_{SEPM}(n)$ while b_n and c_n will be kept fixed.

The fit to the observed excitations is done as follows. If P_i denotes the physical property that we wish to reproduce and $M^i_{n,\alpha} = \partial P_i / \partial C^{(\alpha)}_n$ is its derivative with respect to the fitting coefficients $C^{(\alpha)}_n$ of Eq. (14), we will minimize the "cost function"

$$F = \sum_i w_i \left(P_i^{exp} - P_i^{SLDA} - \sum_{n,\alpha} M^i_{n,\alpha} \, \Delta C^{(\alpha)}_n \right)^2 + \sum_{n,\alpha} |\Delta C^{(\alpha)}_n|^2 \, \omega_{n,\alpha} \quad , \quad (15)$$

where $\Delta C^{(\alpha)}_n \equiv C^{(\alpha)}_{SEPM}(n) - C^{(\alpha)}_{SLDA}(n)$ are solutions of the linear equations Eq. (15) and w_i and $\omega_{n,\alpha}$ are predetermined weight functions. As will be shown later, the changes from $C^{(\alpha)}_{SLDA}(n)$ to $C^{(\alpha)}_{SEPM}(n)$ are rather small, thus, the use of a *linear* representation Eq. (15) for the fitting process is adequate. This closeness of $u^{(\alpha)}_{SEPM}(q)$ to $u^{(\alpha)}_{SLDA}(q)$ implies that many properties of SEPM follow that of SLDA. These include wavefunctions, deformation potentials, the transferability between different structures, etc.

3. APPLICATIONS TO BULK CdSe AND Si

In this section we will apply the approach outlined in Sec. II to covalent Si and to partially ionic CdSe, thus covering a range of semiconductor systems.

3.1 Calculating the spherical LDA potential (SLDA)

Using *ab initio* pseudopotentials [17] and the Ceperley-Alder exchange correlation function as parameterized by Perdew and Zunger [18], we have performed self-consistent LDA calculations for five CdSe and five Si systems. The five CdSe systems are: (1) the wurtzite structure with atomic volumes Ω_0^{CdSe} and $0.944\Omega_0^{CdSe}$; (2) the zincblende structure with atomic volume Ω_0^{CdSe}; (3) the rocksalt structure with atomic volumes $0.8\Omega_0^{CdSe}$ and Ω_0^{CdSe}. Here Ω_0^{CdSe} is the experimental equilibrium volume per atom of wurtzite CdSe at ambient pressure (lattice constants a=4.3Åand $c/a = \frac{2}{3}\sqrt{6}$). Note that both the wurtzite and the zincblende structures have coordination number of four, while

rocksalt is sixfold coordinated. The five Si systems are: (1) the diamond structure with atomic volumes Ω_0^{Si} and $0.92\Omega_0^{Si}$; (2) the simple cubic structure with atomic volume $0.82\Omega_0^{Si}$; (3) the β-tin structure with atomic volume $0.72\Omega_0^{Si}$ and c/a ratio 0.552; and (4) the simple hexagonal structure with atomic volume $0.69\Omega_0^{Si}$ and c/a ratio 0.94. Here, Ω_0^{Si} is the measured cell volume per atom of diamondlike Si at ambient pressure (lattice constants $a = 5.43\mathring{A}$). The diamond structure of Si is a fourfold coordinated semiconductor, while Si in the simple hexagonal structure is an eightfold coordinated metal.

The self-consistently screened LDA potentials $V_{LDA}^{(\alpha)}(\mathbf{G}_\sigma)$ are used to solve for $v_{SLDA}^{(\alpha,\sigma)}$ in Eq. (11). The solutions $v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|)$ *vs* $|\mathbf{G}_\sigma|$ for CdSe and Si are shown as the diamond symbols in Fig. 2.

The important result demonstrated in Fig. 2 is that $v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|)$ *for different structures and unit cell volumes all fall on a nearly universal curve.* The structurally averaged and least square fitted curves $u_{SLDA}^{(\alpha)}(q)$ of Eq. (14) are shown in Fig. 2(a)-(c) as solid lines. The fit is excellent.

To examine the overall accuracy of the SLDA, we compare the band structures obtained using in Eq. (6) $u_{SLDA}^{(\alpha)}(q)$ with the band structure obtained in direct LDA calculations. Table I shows the RMS and maximum band energy differences between SLDA and LDA calculations of Si. Here we include

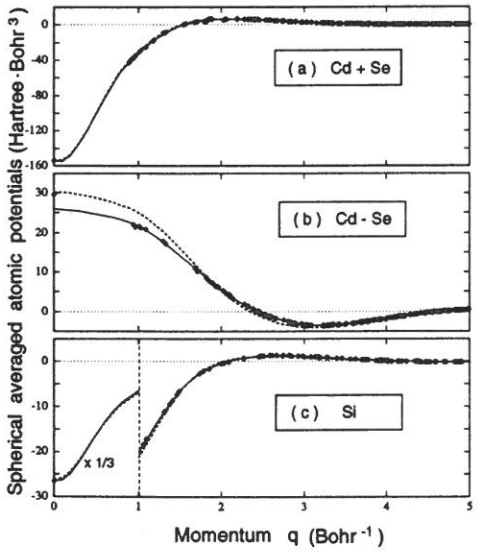
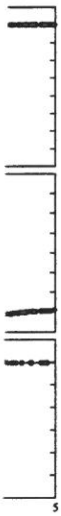


Figure 2: The spherical LDA (SLDA) potential $v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|)$ as obtained by inverting the self-consistent bulk LDA calculations on five crystal structures and cell volumes. Diamond symbols show the numerical results obtained by solving Eq. (11) for all five structures. Solid lines represent least square fits of all the diamond symbols using the analytic expression of Eq. (14). Dashed lines represent the empirically adjusted potential (SEPM) fit to the experimental excitations. For CdSe, we give the symmetric $v^{Cd} + v^{Se}$ potential as well as the antisymmetric $v^{Cd} - v^{Se}$ part. From Ref. [16].

(1) the diamond struc-
le cubic structure with
tomic volume $0.72\Omega_0^{Si}$
cture with atomic vol-
asured cell volume per
stants $a = 5.43\mathring{A}$). The
nductor, while Si in the
netal.
$(\mathbf{G}_\sigma)$ are used to solve
$\mathbf{G}_\sigma|$ for CdSe and Si are

t $v_{SLDA}^{(\alpha,\sigma)}(|\mathbf{G}_\sigma|)$ *for dif-*
ly universal curve. The
$v_{LDA}^{(\alpha)}(q)$ of Eq. (14) are

ompare the band struc-
d structure obtained in
maximum band energy
Here we include



$_A(|\mathbf{G}_\sigma|)$ as obtained by
ve crystal structures and
ults obtained by solving
least square fits of all
Eq. (14). Dashed lines
fit to the experimental
potential as well as the

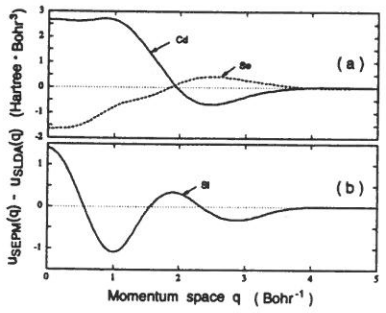


Figure 3: Potential difference $u_{SEPM}^{(\alpha)} - u_{SLDA}^{(\alpha)}$ in momentum space for (a) Cd, Se and (b) Si. These are the $l = 1$ pseudopotentials used here as the local part. From Ref. [16].

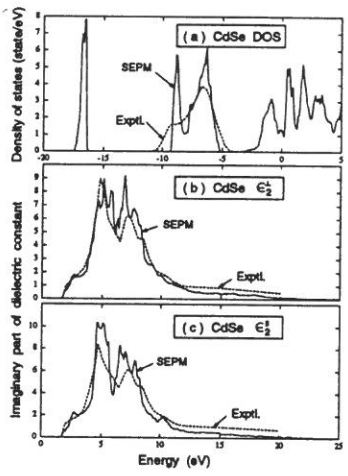


Figure 4: Density of states (DOS) and imaginary part of dielectric constant (ϵ_2) of CdSe in the wurtzite structure. The solid lines are SEPM results. The spin-orbit interaction is included. $\perp$ and $\parallel$ denote perpendicular and parallel to the c axis of the wurtzite structure, respectively. The experimental curve for the DOS is from Ref.[22]. The experimental curves for ϵ_2 are from Ref.[24]. From Ref. [16].

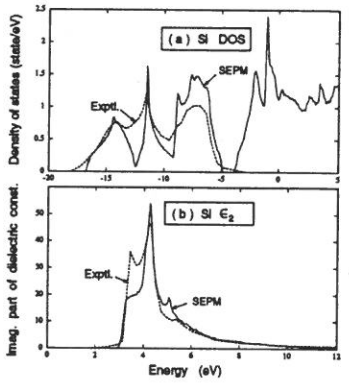


Figure 5: Density of states (DOS) and imaginary part of dielectric constant (ϵ_2) of Si in the diamond structure. Solid lines are SEPM results. The spin-orbit interaction is not included. The experimental curves for the DOS and ϵ_2 are from Ref.[23] and Ref.[26], respectively. From Ref. [16].

Table 1: RMS and maximal band structure error $\epsilon_i^{SLDA} - \epsilon_i^{LDA}$ (in meV) for Si systems. We include in the statistics most high symmetry states, (e.g, L, Γ, X and K for the diamond structure) and average over band energies up to ∼4 eV above the CBM or Fermi level. The SLDA *vs* LDA error in the hydrostatic deformation potential in column four is about 18% and the relative error of the phonon deformation potentials in column five is about 14%. The BC8 and fcc structures were not included in the fit of the SLDA potential, so the results are pure predictions.

Struct.:	Diam.	Diam.	Diam.	Def. diam.	sc
Volume:	Ω_0	$0.92\Omega_0$	$\{\Omega_0\} - \{0.92\Omega_0\}$	$\{R+\delta\} - \{R\}$	$0.82\Omega_0$
RMS error:	46	69	59	66	10
max error:	97	136	201	176	22
Struct.:	sh	β-tin	BC8	fcc	
Volume:	$0.69\Omega_0$	$0.72\Omega_0$	$0.91\Omega_0$	$0.72\Omega_0$	
RMS error:	16	15	75	75	
max error:	43	41	270	122	

the high symmetry points for each structure (e.g, L, Γ, X and K for the diamond structure) and all eigenstates up to ∼4 eV above the conduction band minimum for semiconductors and ∼4 eV above the Fermi energy for metals. The average SLDA *vs* LDA energy difference is about 60 meV. Notably, the energy errors for simple cubic, simple hexagonal and β-tin structures are much smaller (∼ 15 meV) than those for the diamond structure. The hydrostatic deformation potential difference between SLDA and LDA is about 18% for the diamond structure. Table I also gives the energies of the diamond structure with randomly displaced Si atoms (by 10% of the bond length). The error in the phonon deformation potential is about 14%. Table I also shows predictions for additional crystal structures, *not used in our fits*. These include the fourfold coordinated BC8 structure [19], which has eight Si atoms per unit cell and an atomic volume of $0.91\Omega_0^{Si}$. The average SLDA *vs* LDA band energy error is 75 meV, only slightly larger than the errors for the structures included in the fit. We also tested the simple face centered cubic structure, which is a 12-fold coordinated system with an atomic volume of $0.72\Omega_0^{Si}$. The average SLDA *vs* LDA energy error is also 75 meV.

To summarize the *energy errors*, we can say that the spherical and structurally averaged potential $u_{SLDA}^{(\alpha)}(q)$ reproduces the original self-consistent LDA band energies to within ∼ 0.1 eV or better for a range of different crystal structures, including those not used in its construction.

We next discuss the accuracy of $u_{SLDA}^{(\alpha)}(q)$ in reproducing LDA *wavefunctions* and related properties. To this end, we have calculated the overlap between the SLDA wavefunctions $\tilde{\psi}_i$ [Eq. (6)] and the LDA wavefunctions ψ_i [Eqs. (2)-(3)]. The results for CdSe in the wurtzite structure and for Si in the diamond structure show that for both systems the average overlap $< \tilde{\psi}_i | \psi_i >$ (over states of a few special k points and within 4eV above the CBM) is above 99.99%, while the minimum overlap is about 99.97%. This agreement is excellent.

A more stringent test of wavefunction quality is to compare the momentum matrix elements $M_{if} = < \psi_i|\hat{p}|\psi_f >$, since a small change in the wavefunction can lead to a large change in this quantity. We find that M_{if} calculated by the SLDA is very close to the LDA values with typical errors less than 1%.

In summary, the wavefunctions, transition matrix elements (and actually also the dielectric constants) obtained with the spherical and structural averaged atomic potential $u^{(\alpha)}_{SLDA}(q)$ reproduce the original self-consistent LDA calculations extremely well.

3.2 Using the SLDA potential to construct the SEPM

We next will fit $u^{(\alpha)}_{SLDA}(q)$ to get the semiempirical potential $u^{(\alpha)}_{SEPM}(q)$ using Eqs. (14) and (15). The resulting $u^{(\alpha)}_{SEPM}(q)$ is shown as dashed lines in Fig. 2(a) and (b) for CdSe and in Fig. 2(c) for Si. The change from $u^{(\alpha)}_{SLDA}(q)$ to $u^{(\alpha)}_{SEPM}(q)$ [Fig. 1(a)-(c)] is very small. Figure 3 shows the difference $u^{(\alpha)}_{SEPM} - u^{(\alpha)}_{SLDA}$ in momentum space on an enlarged scale. Note that the change is confined to $q < 3\ au^{-1}$. *the main LDA error in the potential is an underestimation (overestimation) of the low-momentum components in the cation (anion).*

3.3 SEPM results and comparison with the traditional EPM

In this section, we will show our SEPM results and compare them with the results of traditional Bergstresser and Cohen [20] local empirical pseudopotential for CdSe and with the Chelikowsky and Cohen [21] nonlocal empirical pseudopotential for Si. The comparison is done for (a) the band structure, (b) the wavefunctions (c) the optical properties.

(a) *SEPM band energies*: The fitted experimental quantities P_i and their SEPM results are listed in Table II for CdSe and in Table III for Si. To see the modification from SLDA to SEPM, we also listed the physical quantities calculated from SLDA. For Si, our SEPM band structure has a similar quality as the (already accurate) Chelikowsky-Cohen nonlocal empirical pseudopotential (note below, however, the different quality of wavefunctions). For CdSe, our SEPM result represent an improvement especially for the upper valence band width, compared to the Bergstresser-Cohen local empirical pseudopotential. Note that both the hydrostatic and phonon deformation potential of the SEPM closely follow those of the LDA and the SLDA. This is one of the advantage of using SLDA as our starting potential in the fitting procedure. The SEPM inherits many of the correct properties of the SLDA (hence LDA) without explicit fitting.

(b) *SEPM wavefunctions*: We examine the average overlap for all states up to CBM + 4 eV. For CdSe, the overlap of the Bergstresser-Cohen wavefunction with the LDA wavefunction is on average only 97%, the minimum overlap being 92%. The overlap of the SEPM wavefunction with LDA (and SLDA) wavefunction is 99.8% on average, the minimum being 99.5%. For Si, the Chelikowsky-Cohen nonlocal EPM wavefunction overlap with LDA result

Table 2: Comparison for CdSe of band energies (in eV), band gap E_g (in eV) and effective masses (in units of electron mass) as obtained in the (i) LDA, (ii) experiment (references will be given in publication [16]), (iii) LDA with spherically symmetric local potential (SLDA), (iv) the semiempirical SEPM and (v) the traditional empirical pseudopotential method (EPM) [20]. Unless otherwise stated, the assumed structure is wurtzite and the energies are measured from top of the valence band $\Gamma_{6v}(A)$. ZB and RS denote the zincblende and rocksalt structures, respectively. Quantities denoted by asterisk show particularly noted improvement in SEPM relative to EPM.

CdSe	LDA	Exptl.	SEPM	SLDA	EPM
$\Gamma_{1,6v}$[work func.]	-6.29	−5.35	-5.24	-4.61	–
Cry. Field Split.	0.031	0.025	0.027*	0.052	-0.076
Spin-Orb. Split.	0.46	0.43	0.40	0.43	–
E_g	0.73	1.74	1.72	0.73	1.78
Γ_{3v}	-4.27	−5.20	−4.06*	-4.59	-2.53
Γ_{5v}	-0.83	−1.20	-0.84	-0.91	-0.58
$\Gamma_{3c} - \Gamma_{5v}$	3.11	4.30	3.96	3.00	4.24
$\Gamma_{6c} - \Gamma_{5v}$	6.45	8.5	7.12	6.26	7.75
Γ'_{1v}	-12.19	−11.4	−12.08*	-12.27	-14.73
$M_{3c} - M_{4v}$	4.04	5.2	4.88	4.07	5.10
M_{1c}	3.39	4.5	4.17	3.27	4.90
M'_{3c}	4.83	6.25	5.23	4.71	6.41
M_{4c}	6.35	7.50	6.95	6.24	7.73
M_{4v}	-0.74	−1.20	-0.81	-0.90	-0.55
M_{3v}	-1.39	−1.70	−1.35*	-1.59	-0.70
M_{2v}	-1.74	−2.45	−1.79*	-1.98	-1.17
M_{1v}	-2.55	−3.20	−2.51*	-2.87	-1.51
M'_{3v}	-3.65	−3.60	−3.37*	-3.94	-2.36
M'_{1v}	-4.19	−4.90	−3.88*	-4.50	-2.70
$m_c(\|\|)$	0.07	0.13	0.13	0.07	0.18
$m_c(\perp)$	0.08	0.13	0.16	0.08	0.17
$m_{vA}(\|\|)$	1.64	1.30	1.83	1.26	–
$m_{vA}(\perp)$	0.18	0.36	0.29	0.15	–
$m_{vB}(\perp)$	0.20	0.9(?)	0.44	0.21	–
E_g (ZB)	0.70	1.68	1.59	0.59	–
$X_{1c} - \Sigma_{4v}$ (RS)	-0.62	0.70	0.69	-0.80	–
$X_{1c} - L_{3v}$ (RS)	-0.60	0.70	0.69	-0.82	–

band gap E_g (in eV)
d in the (i) LDA, (ii)
iii) LDA with spher-
irical SEPM and (v)
20]. Unless otherwise
s are measured from
ncblende and rocksalt
ow particularly noted

EPM
–
-0.076
–
1.78
-2.53
-0.58
4.24
7.75
-14.73
5.10
4.90
6.41
7.73
-0.55
-0.70
-1.17
-1.51
-2.36
-2.70
0.18
0.17
–
–
–
–
–
–

Table 3: Comparison for Si of band energies (in eV), band gap E_g (in eV) and effective masses (in units of electron mass) as obtained in the (i) LDA, (ii) experiment, (iii) LDA with spherically symmetric local potential (SLDA), (iv) the semiempirical SEPM and (v) the traditional empirical pseudopotential method (EPM) [21]. The energies are measured from top of the valence band $\Gamma_{25'v}$. Spin-orbit interaction is omitted. $m^{(i)}_{\Gamma-X}(h)$ and $m^{(i)}_{\Gamma-L}(h)$ stand for the non-spin-coupled effective hole mass [defined as $(\hbar k)^2/2\Delta E$] in the $\Gamma - X$ and $\Gamma - L$ directions, where i denotes the band degeneracy.

Si	LDA	Exptl.	SEPM	SLDA	EPM
$\Gamma_{25'v}$[work func.]	-5.2	−4.9	-4.90	-4.74	–
Γ_{1v}	-11.92	−12.5	-11.79	-11.98	-12.36
Γ_{15c}	2.55	3.35	3.01	2.70	3.42
Γ_{2c}	3.14	4.15	4.12	3.75	4.16
X_{4v}	-2.83	−2.9	-2.78	-2.93	-2.88
X_{1c}	0.62	1.13	1.26	0.76	1.14
L_{2v}	-9.58	−9.3	-9.52	-9.64	-9.90
L_{1v}	-6.96	−6.8	-6.83	-7.06	-7.10
L_{3v}	-1.17	−1.2	-1.18	-1.25	-1.23
L_{1c}	1.47	2.04	1.96	1.58	2.34
L_{3c}	3.30	3.9	3.88	3.52	4.34
E_{gap}	0.51	1.124	1.114	0.630	1.009
Σ_{min}	-4.42	−4.48	-4.26	-4.55	-4.47
$m_L(e)$	0.97	0.92	0.93	0.92	0.88
$m_T(e)$	0.19	0.19	0.21	0.19	0.19
$m^{(2)}_{\Gamma-X}(h)$	0.26	0.34	0.31	0.27	0.31
$m^{(1)}_{\Gamma-X}(h)$	0.17	0.15	0.19	0.18	0.20
$m^{(2)}_{\Gamma-L}(h)$	0.54	0.69	0.64	0.59	0.74
$m^{(1)}_{\Gamma-L}(h)$	0.11	0.11	0.15	0.13	0.12

for 99.4% on average, the minimum being 98.7%. On the other hand, the overlap of SEPM wavefunction with LDA (and SLDA) wavefuntion is 99.97% on average, the minimum being 99.90%. Thus, SEPM wavefunctions are 10 times closer to their LDA counterparts than the traditional EPM wavefunctions.

Furthermore, the matrix elements $|<\psi_i|\hat{p}|\psi_f>|^2$ of SEPM follow closely the LDA and SLDA results. On the other hand, we find that the some matrix elements of the traditional EPM's differ from the LDA results by a factor of two.

(c) *SEPM optical properties*: The SEPM density of states and the dielectric constants $\epsilon_2(E)$ are shown in Fig. 4 for CdSe and Fig. 5 for Si. The density of states compares very well with the experimental data, especially the peak positions. For CdSe, however, there are a number of discrepancies caused by the neglect of the Cd 4d orbitals in our pseudopotential treatment: (i) The width of the CdSe upper valence band in Fig. 4(a) is smaller than the experimental value. This is partly caused by our neglect of the Cd 4d states. Because of p-d coupling, the 4d state will push up the top of the valence band resulting in

an increased upper valence band width. (ii) The optical absorption spectrum $\epsilon_2(E)$ of CdSe agrees very well with the experimental data [24] for the peak positions. However, in the $E \sim 14$ eV region, the experimental $\epsilon_2(E)$ has larger values than the SEPM results. This too is because we have neglected the Cd 4d states in our CdSe SEPM calculations. The 4d electrons has a contribution to $\epsilon_2(E)$ right in the 14 eV region[25].

For Si, the SEPM $\epsilon_2(E)$ misses the first peak apparent in the experimental result [26]. This peak comes from a large excitonic effect in the Si system, which cannot be described by our single electron representation. (The same is true for all other single electron methods including LDA and traditional EPM).

4. CONCLUSIONS

Our principal results are:
(i) First-principles structure factors for a range of crystal structures, coordination numbers, and unit cell volumes of a given compound all fall on a simple, nearly universal curve, with little scatter (diamond symbols in Fig. 2). This constitutes an *ex post facto* modern explanation of Marvin Cohen's 30 year old EPM idea.
(ii) Spherically-symmetrized and structurally-averaged LDA structure factors (SLDA given by the solid lines in Fig. 2) reproduce the full LDA band structure (Tables I and II) to within $\simeq$0.1 eV or better over a range of crystal structures.
(iii) The corrections in the SLDA potential needed to reproduce *observed* excitation energies are both systematic and small (Fig. 3). Thus the resulting wavefunctions retain a large overlap with the original LDA wavefunctions.

Our procedure of empirically fitting the SLDA to obtain the SEPM can be contrasted with the *ab initio* GW correction to LDA results [27]: both approaches change the LDA potential, such that the ensuing band energies agree with the experimental results. Our SEPM wavefunctions have larger than 99.9% overlap with the LDA wavefunctions. This provides some insight into a long-standing puzzle: as reported by Hybertson *et al.* [27], the GW single particle wavefunction also has larger than 99.9% overlaps with the LDA wavefuntion (this has, apparently, never been understood). Since in both modifications of LDA (GW and SEPM), the ensuing wavefunctions have $\geq$ 99.9% overlaps with the original LDA wavefunction, the reason for this must be rather general, and not a unique property of the GW correction. The reason might be that the original LDA band energies are close enough to the experimental results, so that small Hamiltonian modifications ΔH are enough to correct them by first order perturbation $\Delta\epsilon_i = < \psi_i|\Delta H|\psi_i >$ without changing the wavefunctions.

ACKNOWLEDGMENTS

This work was supported by OER-BES-DMS under contract No. DE-AC02-83CH10093.

References

[1] M.L. Cohen and T.K. Bergstresser, Phys. Rev. **141**, 789 (1966).

[2] D. Brust, J.C. Phillips and F. Bassani, Phys. Rev. Lett. **9**, 94 (1962); D. Brust, M.L. Cohen and J.C. Phillips, Phys. Rev. Lett. **9**, 389 (1962).

[3] M. L. Cohen and V. Heine, Solid State Physics, **24**, 38 (1970).

[4] P. Hohenberg and W. Kohn, Phys. Rev. **136 B**, 864 (1964).

[5] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).

[6] A. Zunger and M. L. Cohen, Phys. Rev. Lett. **41**, 53 (1978); Phys. Rev. B **18**, 5449 (1978); Ibid. **20**, 4082 (1979).

[7] J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).

[8] A. Zunger and A. J. Freeman, Int. J. Quantum Chem. S**10**, 383 (1976); Phys. Rev. B **15** 4716 (1977); Ibid. **15**, 5049 (1977).

[9] S. Topiol, A. Zunger, and M. Ratner, Chem. Phys. Lett. **49**, 367 (1977).

[10] A. Zunger and M. L. Cohen, Phys. Rev. B **19**, 568 (1979).

[11] L. Kahn and W. A. Goddard, Chem. Phys. Lett. **7**, 667 (1968); L. R. Kahn, P. Baybutt, and D. G. Tikhlar, J. Chem. Phys. **65**, 3826 (1976).

[12] G. Kerker, J. Phys. C **13**, L189 (1980).

[13] D.R. Hamann, M. Schluter and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[14] See *Physics of low-dimensional semiconductor structures*, Edts. P. Butcher, N. H. March, and M. P. Tosi, Plenum, NY (1993), and *Intersubband Transitions in Quantum Wells*, NATO ASI **228** (1992).

[15] L.W. Wang and A. Zunger, J. Chem. Phys. **100**, 2394 (1994); *ibid*, J. Phys. Chem. **98**, 2158 (1994); *ibid* Phys. Rev. Lett. **73**, 1039 (1994).

[16] L. W. Wang and A. Zunger (unpublished results).

[17] N. Trouillier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[18] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980); J. Perdew and A. Zunger, Phys. Rev. B, **23**, 5075 (1981).

[19] J.S. Kasper and S.M. Richards, Acta Crystallogr. **17**, 752 (1964). U.O. Karlsson, D. Straub, S.A. Flodstrom and F.J. Himpsel, Phys. Rev. B **36**, 6566 (1987).

[20] T. K. Bergstresser and M. L. Cohen, Phys. Rev. **164**, 1069 (1967).

[21] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B, **10**, 5095 (1974).

[22] L. Ley, R.A. Pollak, F.R. McFeely, S.P. Kowalczyk and D.A. Shirley, Phys. Rev. B, **9**, 600 (1974).

[23] L. Ley, S.P. Kowalczyk, R.A. Pollak, D.A. Shirley, Phys. Rev. Lett. **29**, 1088 (1972).

[24] J.L. Freeouf, Phys. Rev. B, **7**, 3810 (1973).

[25] R.L. Hengehold and F.L. Redrotti, Phys. Rev. B **6**, 2262 (1972).

[26] D.E. Aspnes and A.A. Studna, Phys. Rev. B **27**, 985 (1983); G.E. Jellison, Jr and F.A. Modine, *ibid*, **27**, 7466 (1983).

[27] M.S. Hybertson and S. G. Louie, Phys. Rev. B, **34**, 5390 (1986); see also: Z.H. Levine and D.C. Allan, Phys. Rev. Lett. **63**, 1719 (1989).