

Correcting overbinding in local-density-approximation calculations

A. van de Walle and G. Ceder

Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 27 October 1998; revised manuscript received 5 February 1999)

We introduce a semiempirical method to correct the systematic equilibrium lattice parameters underestimation present in first principles calculations based on the local density approximation. The method consists in performing calculations under a negative pressure such that the calculated equilibrium volume matches the experimentally observed one. We find that elastic properties obtained under these conditions are typically in better agreement with experiment. We also observe that the negative pressure which needs to be applied to crystalline compound can be reliably interpolated by taking the concentration-weighted average of the pressures determined from pure crystals made of each of the elements present in the compound. In a large class of materials, the knowledge of one pressure per element is thus sufficient to correct most of the bias in lattice constants and elastic properties. We finally propose a simple model of the nonlocal contribution to the exchange-correlations energy that is able to explain the observed linear dependence between the required negative pressure and concentration. [S0163-1829(99)01324-7]

I. INTRODUCTION

Over the last three decades, first-principles calculations based on the local density approximation (LDA) of density functional theory (DFT) (Refs. 1 and 2) have been extensively used to successfully predict numerous materials properties.³ While the LDA is known to have many limitations, we will focus here on properties that can, in principle, be exactly determined by DFT. These include any properties that can be derived from the knowledge of the electronic ground state energy for any given position of the nuclei.

One significant limitation of LDA is its “overbinding” of extended solids: Lattice parameters are typically underpredicted, while cohesive energies, phonon frequencies, and elastic moduli are typically overpredicted. In this paper, our main concern will be the biases in the equilibrium lattice constants, the phonon frequencies and the elastic moduli.

The generalized gradient approximation (GGA) (see Ref. 4, and references therein), which augments the LDA exchange-correlation energy with a function of the gradient of the charge density, was introduced to address some of the weaknesses of the LDA. The GGA has been clearly shown to improve agreement with experiments for properties of finite systems such as atoms or molecules.^{4,5} However, the success of the GGA in extended systems has been more controversial:^{6–8} despite many successes,^{4,9} GGA frequently overcorrects LDA’s overbinding, sometimes yielding worse agreement with experiment than LDA. Obtaining accurate cohesive properties of crystalline phases from first-principles calculations thus remains a problematic issue.

This paper describes how LDA’s overbinding can be easily corrected for a large class of materials, by using a small amount of experimental input to correct the results of first-principles calculations. Our approach focuses on obtaining the correct equilibrium volume and relies on the fact that this correction is sufficient to significantly improve the accuracy of calculated elastic properties as well.

We first motivate the use of the experimental equilibrium volume in LDA calculations and describe how LDA calcu-

lations can be combined with experimental input to obtain accurate equilibrium volumes. We then show, through the results of first-principle calculations, that errors in the equilibrium volume follow a trend that can be easily corrected with the use of a minimal amount of experimental input. We finally provide a formal explanation of the success of this simple approach and discuss the implication of our findings for the continuing search for better exchange-correlation functionals.

II. METHOD

A. The importance of volume

It has been previously observed^{10–14} that, in many compounds, phonon frequencies calculated through first-principles are in closer agreement with experiment when the lattice parameters are artificially constrained to the experimental values. To determine whether this is a general trend, we performed a systematic investigation of the calculated bulk modulus in various elements, evaluated at the experimental lattice parameter. Since biases in the bulk modulus are usually correlated with biases in the phonon frequencies, investigating the bulk modulus provides a computationally simple way to investigate the bias in elastic properties in a large number of materials.

Figure 1 shows the result of our calculations, which will be described in detail in a later section. Using the experimental volume appears to essentially correct the bias in the calculated bulk modulus in a large class of materials. For all transition and noble metals tested, a substantial improvement is observed. More importantly, the magnitude of the correction is correlated with the magnitude of the error—indicating that this is clearly not just a random correction in the right direction. Interestingly, performing the same investigation with GGA shows that using the experimental volume for GGA calculations gives worse estimates of the bulk modulus than LDA at the experimental volume (and also worse estimates than GGA at the GGA equilibrium volume). For this

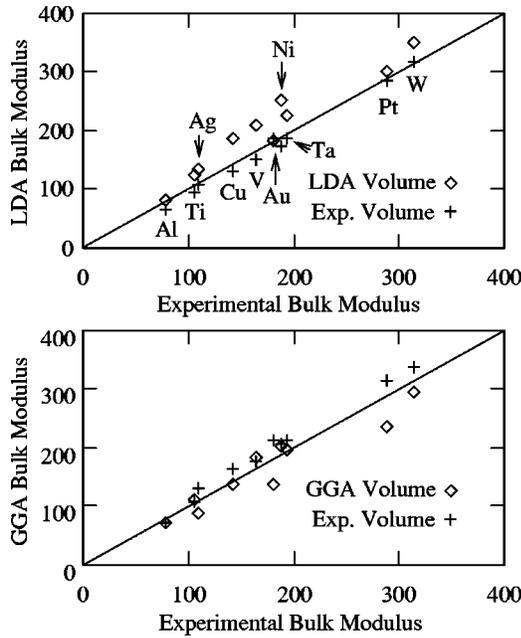


FIG. 1. Calculated versus experimental bulk modulus of selected elements. Evaluating the bulk modulus at the experimental volume considerably improves agreement with experiment. (The data for W and Ta was provided by the authors of Ref. 7.)

reason, we will focus on correcting the LDA equilibrium volume (instead of the GGA equilibrium volume).

Unfortunately, there are exceptions to this trend: for some elements, such as alkali metals, aluminum and silicon, evaluating the bulk modulus through LDA at the experimental lattice constant provides no improvement. The elements that appear to cause problems are those that have a small bulk modulus.

Nevertheless, we can conclude that evaluating elastic properties with LDA at the experimental volume provides significantly better agreement with experiment in a large class of materials including transition metals as well as other compounds studied by previous investigators.¹⁰⁻¹⁴ A method of obtaining the correct lattice parameters from LDA calculations with as little experimental input as possible could therefore be extremely useful. This is the problem which we now turn to.

B. Conceptual framework

In this section, we will first introduce a simple one-parameter correction to the LDA Hamiltonian of the system. This parameter is element dependent and can be tuned to obtain the true equilibrium volume. Such an approach avoids the conceptual problem associated with artificially fixing the volume without modeling the cause of the volume change. We then describe how the knowledge of this parameter for simple systems enables us to extrapolate this parameter for systems where it is unknown. We finally present how we tested the applicability of our method in a large number of systems.

Let us define a structure to be a specific stable or metastable periodic arrangement of atoms of given types. Such a structure has well defined energy, equilibrium volume, elastic moduli and phonon frequencies. The fact that, in general,

LDA correctly models elastic properties at the experimental volume suggests the following approximation for the total energy of a structure in the neighborhood of its equilibrium geometry:

$$E = E^{\text{LDA}}(\{\mathbf{c}_i\}, \mathbf{a}_j) + \Delta E_{\text{xc}}(\Omega),$$

where Ω is the average atomic volume, E^{LDA} is the energy obtained using LDA, for given atomic positions \mathbf{c}_i and lattice vectors \mathbf{a}_j , while ΔE_{xc} is the correction introduced to model the effect of the nonlocal contribution to the exchange-correlation energy (which is not accounted for within LDA). The approximation introduced is to allow ΔE_{xc} to depend on the average atomic volume only. This approximation rests on the assumptions that LDA usually predicts reasonably accurate unit cell *shapes* (except for an isotropic scaling factor) and phonon frequencies, when evaluated at the experimental volume. If ΔE_{xc} were strongly dependent on the cell shape, the LDA cell shape would be significantly biased and if ΔE_{xc} were strongly dependent on atomic positions, the LDA would give incorrect estimates of the phonon frequencies.

We do not know ΔE_{xc} and seek to approximate it with a truncated Taylor series. Since we are interested in the equation of state in the neighborhood of the true equilibrium volume, we expand ΔE_{xc} in the neighborhood of the experimental volume Ω_{exp} :

$$E = E^{\text{LDA}}(\{\mathbf{c}_i\}, \mathbf{a}_j) + \Delta E_{\text{xc}}(\Omega_{\text{exp}}) + \frac{\partial \Delta E_{\text{xc}}(\Omega_{\text{exp}})}{\partial \Omega} (\Omega - \Omega_{\text{exp}}). \quad (1)$$

The truncation of the Taylor series after the first order can be justified from the fact that LDA seems to give accurate values of the bulk modulus when it is evaluated at the experimental volume, as mentioned before. In structures having a small bulk modulus, this assumption is more questionable, as the second order term is more likely to be non-negligible compared to the second derivative of E^{LDA} . This is probably the source of the inaccuracies in the bulk modulus we observed in soft elements.

While the term $\Delta E_{\text{xc}}(\Omega_{\text{exp}})$ may have an influence on cohesive energies, it leaves the equilibrium volume as well as elastic properties unaffected. We will thus focus on the quantity $\partial \Delta E_{\text{xc}}(\Omega_{\text{exp}})/\partial \Omega$ which has a direct impact on the equilibrium volume. It can be interpreted as the pressure that needs to be applied to the LDA solid so that its volume equals the experimental volume. This pressure, hereafter called the nonlocal exchange-correlation pressure (P_{xc}), is expected to be structure-dependent and we now turn to the problem of determining this pressure for a given structure using the minimum amount of experimental input.

The value of P_{xc} for each element, in its crystalline form, significantly changes from one element to the next. Elements that have large charge density inhomogeneities, such as 3d metals, require a large corrective pressure while free-electron-like metals, such as alkali metals or aluminum, require a small pressure. It is often stated that LDA performs poorly in the case of alkali metals, since it significantly underpredicts their equilibrium volume. But the magnitude of the error introduced by the LDA, in terms of energy per unit volume (as measured by P_{xc}), is actually small. This small

error in the energy translates into a large volume change only because alkali metals have such a low bulk modulus.

Obviously, using one P_{xc} per element, one can obtain the correct equilibrium volume for all elements. It is now interesting to check whether the value of P_{xc} in elements can help predict the value of P_{xc} in compounds. Remarkably, we have observed that taking the concentration-weighted average of P_{xc} of each element in the compound provides a simple and effective method to predict P_{xc} in a compound.

Since we are linearly interpolating the nonlocal exchange-correlation pressure, this naturally raises the following question: Why not linearly interpolate either the volume or the lattice parameter instead? The answer is that compounds that have the same composition can have slightly different volumes. For example, alloys typically expand by a fraction of a percent upon disordering. These volume relaxations are energetically significant and have an important impact on phonon frequencies¹⁵⁻¹⁷ and mixing energies. Usually, LDA correctly predicts these relative volume changes and preventing them (by fixing the volume) would neglect an effect that LDA is actually able to model. Moreover, applying a pressure instead of fixing the volume has a firmer conceptual basis: Applying a pressure corresponds to introducing a perturbation to the Hamiltonian which causes a volume change. Fixing the volume imposes a volume change without modeling its cause.

C. Calculations

In order to benchmark this approach of linearly interpolating P_{xc} , we compute P_{xc} for a large number of compounds and compare it to the pressure interpolated from the P_{xc} of elements. The value P_{xc} can be determined by artificially constraining the volume to the experimental value and by calculating the pressure acting on the unit cell. For crystals of noncubic symmetry, the cell shape is allowed to relax (at constant volume) until the pressure is isotropic. In all cases, we allow the internal degrees of freedom (i.e., the atom positions) to relax. This relaxation step is introduced so that the pressure is calculated in the same conditions as those in which it would be used, that is, without prior knowledge of cell shape and atomic positions.

The experimental volume used must be the one at absolute zero temperature. This is achieved by taking the experimental volume at room temperature (or higher) listed in Refs. 18,19 and extrapolating it down to absolute zero using the thermal expansion data found in Ref. 20. For intermetallics for which the thermal expansion down to absolute zero has not been determined experimentally, we use a concentration weighted average of the values of the pure elements. For compounds that exhibit an allotropic or a magnetic transformation between the temperature at which the volume is measured and absolute zero, we approximate the thermal expansion of the high-temperature phase at low temperature by the thermal expansion of the low-temperature phase and disregard the volume change taking place at the phase transition. Note that, in most compounds, these corrections are quite small (typically 0.2% change in the lattice constant) compared to the LDA error (1% to 2%). Hence, any inaccuracies in these extrapolations have a small impact on our results. In fact, calculations at the room temperature volume exhibit essentially the same behavior.

Our calculations are performed with the VASP (Refs. 21,22) package which uses ultrasoft²³ pseudopotentials²⁴ and a plane-wave basis. The equilibrium volume can be sensitive to the quality of the pseudopotentials used. We thus verified that the equilibrium lattice constants we obtain do not differ significantly²⁵ from previous full-electron calculations in elements.^{7,9,4,8} The energy cutoff used ranges from 200 to 360 eV depending on the elements present in the structure. For ultrasoft pseudopotentials, these cutoffs are largely sufficient to obtain pressures with a precision of about 1% for transition metals and simple metals. To minimize the effect of Pulay stress, calculations are performed at a sequence of various volumes ranging from -18 to $+12\%$ around the experimental volume, at a constant energy cutoff. At each volume, all degrees of freedom are fully relaxed and the resulting energies are fitted to a polynomial of degree 4. The resulting equation of state is then used to obtain the quantities of interest. The Brillouin zone is sampled using Monkhorst-Pack special points and integrated with the help of the tetrahedron method with Bloech corrections.²⁶ For all structures, the total number of k points in the Brillouin zone is chosen to be around $(15)^3$ divided by the number of atoms in the unit cell. LDA calculations relied on the Ceperley-Alder exchange-correlation functional,²⁷ as parametrized by Perdew and Zunger,²⁸ while GGA calculations relied on the Perdew-Wang functional.⁴ In all calculations, the same exchange-correlation functional is used for both the pseudopotential generation step and the actual calculations. Scalar-relativistic effects are accounted for in the pseudopotential. Nonspin polarized calculations are used for all compounds which are paramagnetic at room temperature. Spin-polarized calculations are used for metallic nickel, the only magnetic compound considered here.

III. RESULTS

We choose binary systems where the number of well characterized phases is sufficiently large to provide a rigorous test of the linear dependence of P_{xc} on concentration. Systems exhibiting intermetallic phases are ideal tests for our method, as the known structures are well characterized experimentally and have small unit cells, which makes calculations easily tractable.

The nonlocal exchange-correlation pressures for all the intermetallic phases tested are shown in Fig. 2. In this figure, the accuracy of the volume determined with the help of an exchange-correlation pressure relative to the LDA volume can be noted by observing that the data points are much closer to the interpolating line than to the zero axis. It is remarkable that the linear relationship holds even when the phases do not share a common parent lattice.²⁹ In particular, the different crystalline structures of titanium exhibit almost the same P_{xc} . This is also the case with lithium, whose hcp and bcc phases have a nonlocal exchange-correlation pressure of -1.48 and -1.39 GPa, respectively.

The same linear relationship appears to hold for ionic and covalent solids as well, as shown in Figs. 3 and 4. While the agreement is less satisfactory in the case of alkali halides, our method still reduces LDA's error on the equilibrium volume by 50%. The case of the ionic compounds is a particularly stringent test of our method because the pure elements

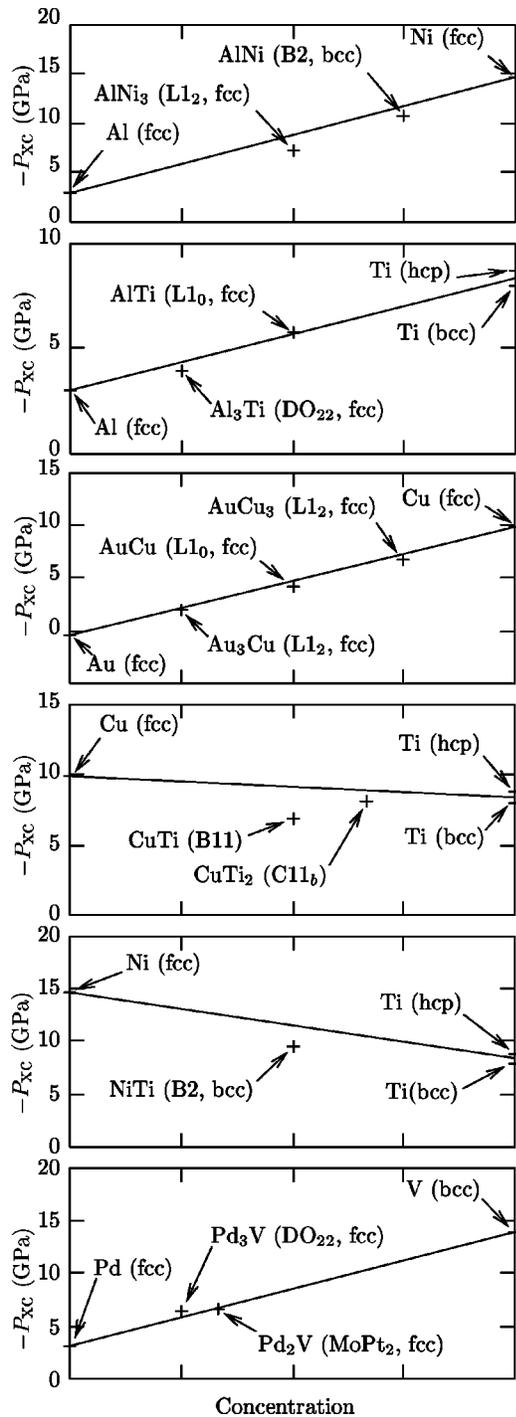


FIG. 2. Nonlocal exchange-correlation pressure as a function of concentration in metallic systems.

used to predict the pressures for the compounds are chemically very different from these compounds: on one side, they are metals (Ca, Mg, Na, K, and Cs), while on the other side, they are covalently bound molecules held together by van der Waals forces (O_2 and Cl_2).

Our method may appear less successful in the case of the covalent crystal SiC. However, our method predicts the lattice constant of SiC within 0.46%, which is more accurate than both the LDA lattice constant (-0.85% error) and the lattice constant obtained by taking either the mean of the lattice parameter of C and Si (3.2% error) or the mean of

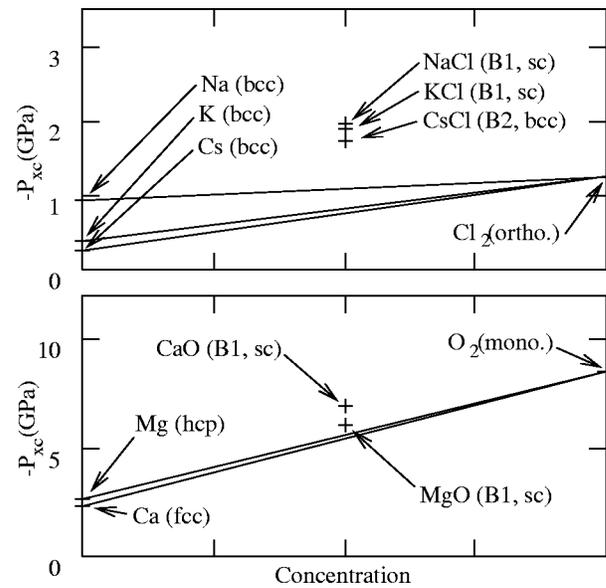


FIG. 3. Nonlocal exchange-correlation pressure as a function of concentration in ionic systems. The pressures for oxygen and chlorine are obtained from their solid crystalline structure at low temperature (23 K for oxygen and 100 K for chlorine).

their molar volume (7.4% error).

Unfortunately, our linear interpolation scheme does fail in some systems. During our benchmarking, we found three systems where this approach was unsuccessful: In the Ga-As and Li-Co-O systems, the pressure was clearly not linear in concentration, while in the case of pure carbon, the pressure associated with diamond and graphite were radically different. In all cases, the failure is associated with the presence of a markedly anisotropic bonding in some of the compounds, which makes the use of an isotropic pressure inappropriate. The value of the stress in graphite, when calculated with the LDA at the experimental lattice parameters, best illustrates the nature of the problem. While the stress perpendicular to the graphitic layers (0.09 GPa) differs from the isotropic component of the stress in the diamond structure (14.4 GPa), the stress along the graphitic layers (16.0 GPa) agrees very well with the value obtained in diamond.

We have up to now only focused on predicting the lattice parameters, but in principle, our correction to the LDA Hamiltonian can just as well be used to improve the accuracy of calculated formation energies. As an additional test, we now show that in two cases where LDA is known to predict the wrong ground state, our correction to the LDA Hamiltonian is sufficient to obtain the correct ground state.

If our correction to the LDA is indeed a function of volume and concentration only, the correction to the energy of

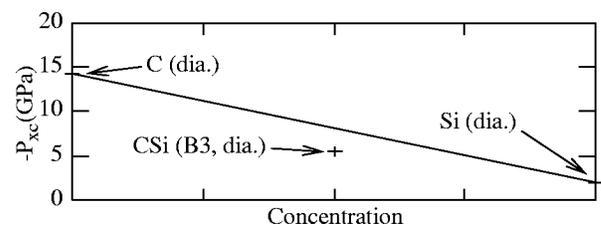


FIG. 4. Nonlocal exchange-correlation pressure as a function of concentration in SiC, a covalent system.

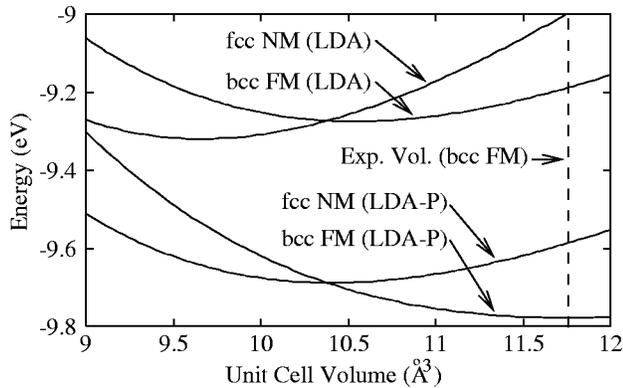


FIG. 5. Equation of state of iron calculated with the LDA only and with the LDA using an exchange-correlation pressure (LDA-P).

two structures which share the same composition should be of the same form. It follows that the volume-independent terms in Eq. (1) do not affect the relative stability of the two structures. The volume-dependent part, $P_{xc}\Omega$, which does affect the relative stability, can be determined as before, from the pressure calculated with the LDA at the known experimental volume of one of the structures. The same quantity $P_{xc}\Omega$ is then added to the equation of state of both structures.

One of LDA's most notorious failures is its prediction that the ground state of iron is a nonmagnetic fcc structure instead of the observed bcc ferromagnetic structure. As shown in Fig. 5, the problem is readily corrected by adding a term of the form $P_{xc}\Omega$ determined from the knowledge of the experimental lattice parameter of bcc iron.

Another example is the LiMnO_2 compound used in lithium ion batteries³⁰ whose ground state is known to be an orthorhombic antiferromagnetic structure. LDA incorrectly predicts a so-called layered structure to be the ground state. Once again, our simple correction, based on the knowledge of the true cell volume of the orthorhombic structure, restores the correct ground state (see Fig. 6).

IV. DISCUSSION

Our observations give us more than a practical way to correct LDA's overbinding. They also provide useful infor-

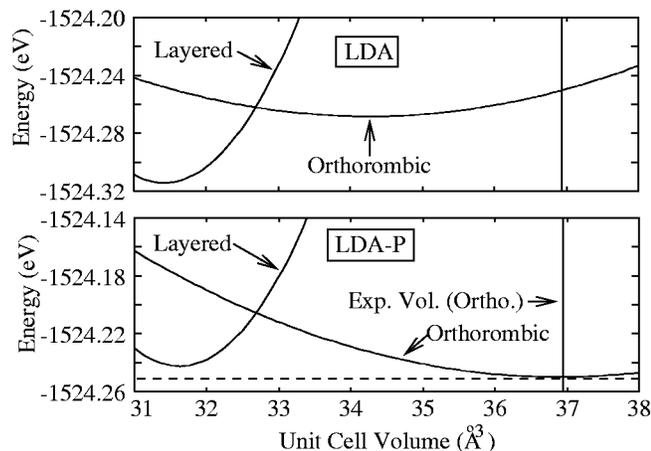


FIG. 6. Equation of state of LiMnO_2 calculated with the LDA only and with the LDA using an exchange-correlation pressure (LDA-P).

mation regarding the structure of the nonlocal contribution to the exchange-correlation energy. Our results indicate that a large part of the error on the total energy obtained through LDA takes the form of a linear function of the volume, where the constant of proportionality is itself linear in concentration. Only very specific corrections to the LDA exchange-correlation energy can give rise to such a contribution. This section is thus devoted to determining the most general form of exchange-correlation functional that is compatible with our observations. We will motivate such a functional form based on the results of previous calculations as well as on reasonable physical assumptions.

Admittedly, our description of the nonlocal contribution to the exchange correlation energy is highly simplified: It is not intended to model strong electron correlations effects. We rather seek to remove LDA's overbinding problem, which is present in all materials, even when electron correlations effects are relatively weak. Nevertheless, we believe that this analysis provides important insight into the continuing search for more accurate exchange-correlation functionals.

This discussion is organized as follows. We first determine which form the correction to the LDA exchange-correlation functional must take in order to give rise to an energy contribution that is a linear function of volume only. We will then examine the implications of constraining the exchange-correlation pressure to be linear in concentration. In particular, we will argue that this behavior is possible only if the correction to the LDA mainly arises from the large nonuniformity in the charge density near the atom nuclei. This picture naturally provides a physical interpretation to the exchange-correlation pressure in terms of the zero-order coefficient of the functional expansion of the exchange-correlation energy, when the point of expansion is taken to be the atomic charge density.

A. Linearity of ΔE_{xc} in volume

We can divide the volume of the solid in two parts: "core" regions where the charge density is essentially identical to the charge density of a free atom and "interstitial" regions where this approximation ceases to apply. The core regions are simply considered to be spheres centered on each nuclei. The interstitial regions include all points outside of these spheres. While the charge density in the core regions is clearly nonuniform, it is essentially frozen. Hence, even if the LDA is a poor approximation in the core regions, the error it introduces is essentially constant and has little influence on cohesive properties and in particular, on the equilibrium volume. We therefore seek the source of LDA's volume underestimation in the inaccurate description of the exchange-correlation energy in the interstitial region. Our approach is to correct the exchange-correlation energy in the interstitial regions only, leaving the correction to be made in the core regions unspecified. We simply need to assume that the exchange-correlation potential correction in the core regions is independent of the environment of the core and is a relatively smooth continuation of the correction in the interstitial regions, in order not to dramatically affect the wave functions in the interstitial region.

In the remainder of this section, we will describe how our results lead to the conclusion that most of LDA's bias can be

corrected by adding to the LDA exchange-correlation energy per unit volume $e_{xc}^{LDA}(\mathbf{r})$ in the interstitial region, a term of the form

$$\Delta e_{xc}(\mathbf{r}) = \Delta e_{xc} + \Delta \mu_{xc} \rho(\mathbf{r}), \quad (2)$$

where $\rho(\mathbf{r})$ is the charge density at point \mathbf{r} and Δe_{xc} and $\Delta \mu_{xc}$ are corrections to the exchange-correlation energy and potential. We take Δe_{xc} and $\Delta \mu_{xc}$ to be uniform over the interstitial region for the following reasons.

The observation that our non-self-consistent correction successfully predicts cohesive properties indicates that the correction to the exchange-correlation potential is relatively uniform over the interstitial region, or at least that the deviations away from a uniform correction are sufficiently small to allow a first-order perturbative treatment (i.e., a non-self-consistent correction). Any large nonuniformities in the correction to the exchange-correlation potential would inevitably lead to large corrections to the LDA wave functions, leaving the success of our non-self-consistent approach unexplained.

Prior investigations also support the validity of this non-self-consistent approach. The error in the LDA exchange-correlation potential was shown to be surprisingly uniform in systems where the exact exchange-correlation potential can be obtained at a moderate computational cost: in light atoms³¹ and in light atoms dimers.³² In these systems, one can see that a simple shift is sufficient to dramatically improve the accuracy of the calculated exchange-correlation potential in the region that would correspond to the interstitial region in extended systems. A non-self-consistent approach can also be justified from the fact that GGA, when implemented in a non-self-consistent way by simply using the GGA exchange-correlation functional with the LDA charge density as an input, yields results that are very similar to a fully self-consistent procedure.^{33,34} The magnitude of GGA's corrections over LDA results are comparable to the magnitude of the corrections devised in the context of the present work and we thus expect the perturbative approach to be appropriate.

A uniform shift in the exchange-correlation potential is modeled by taking $\Delta \mu_{xc}$ to be uniform over the interstitial region and independent of the interstitial charge density, which prohibits the inclusion in Eq. (2) of higher-order terms in ρ . Note that the value of $\Delta \mu_{xc}$ has little effect on the cohesive properties: to the extent that the charge density in the core region is truly frozen, the total charge in the interstitial region is also constant and the integral of $\Delta \mu_{xc} \rho(\mathbf{r})$ over the interstitial regions is volume independent.

On the contrary, the shift in the exchange-correlation energy per unit volume Δe_{xc} directly gives rise to the exchange correlation pressure P_{xc} and, in fact, is numerically equal to it. The fact that the exchange-correlation pressure for structures which have the same composition but do not share a common parent lattice is nearly identical constrains the correction to the exchange-correlation energy to be uniform in the interstitial region. Since the shape of the interstitial space strongly depends on the type of lattice, any nonuniformity would make the P_{xc} lattice dependent. A convenient consequence of the uniformity of Δe_{xc} is that the exchange-

correlation pressure is independent of the precise radius chosen to delimit the core regions.

B. Importance of the core charge density

Using the observation that most of LDA's error on the total energy takes the form of a correction that is linear in volume enabled us to considerably restrict the type of correction to the LDA exchange-correlation functional we are focusing on. Using the fact that the exchange-correlation pressure is linear in atomic concentration lets us be more specific about what determines the unknown constants Δe_{xc} and $\Delta \mu_{xc}$. We have already indicated that these constants should be essentially independent of the interstitial charge density. Analyzing the linearity of P_{xc} with concentration will make it clear that these constants are essentially determined by the charge density in the *core* regions.

Limiting ourselves to the conservative assumption that $\Delta e_{xc}(\mathbf{r})$ at a point \mathbf{r} in the interstitial region only depends on the charge density in the interstitial region (as done in the GGA, for example), makes it difficult to explain our simple linear relationship between P_{xc} and concentration. First, the charge density in the interstitial region is highly dependent on the exact location of neighboring atoms—information the concentration by itself seems unable to convey. Second, even if we assume that atoms sit exactly on the sites of a parent lattice common to all phases in the system of interest *and* approximate the charge density in the interstitial regions by overlapping atomic charge densities, we still face a problem. The exchange-correlation functional is a nonlinear function of the charge density and summing atomic charge densities does not translate into summing the exchange-correlation contributions coming from each neighboring atom. The only possibility is that Δe_{xc} can be approximated by a linear functional. But then, we obtain a ΔE_{xc} that is volume independent since, as we have indicated before, the total interstitial charge density is essentially volume independent.

However, if we consider the possibility that $e_{xc}(\mathbf{r})$ at a point in the interstitial region depends on the charge density in a neighborhood large enough to include the core regions of the neighboring atoms, we will see that a simple explanation for our findings emerges. While the dependence of Δe_{xc} on the interstitial charge density needs to be assumed linear, no such restriction needs to be imposed in the case of the dependence on the core charge density. Assuming that our volume-dependent correction to the LDA exchange-correlation energy in the interstitial region does not arise from the small nonuniformities of the interstitial charge density, but rather from the large nonuniformities of the core charge densities, avoids most of the difficulties mentioned previously. First, the charge density in the core regions is independent of the exact location of neighboring atoms. Second, the problem of the nonlinearity of the exchange-correlation energy is irrelevant since, the core charge density is not obtained by a sum of the contributions of neighboring atoms but by the contribution of one atom only. The only assumption needed is that the corrections arising from each neighboring atom are additive, which we will motivate shortly. Previous calculations in systems simple enough to allow an accurate evaluation of the exchange-correlation en-

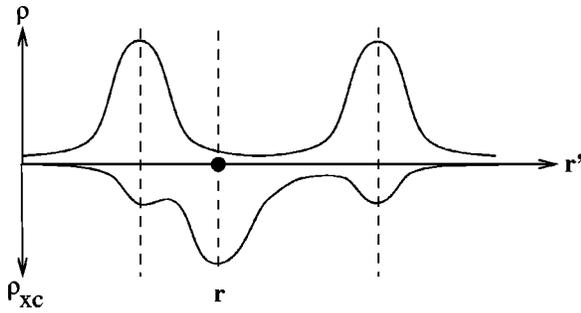


FIG. 7. Schematic shape of the exchange-correlation hole (ρ_{xc}) for a reference electron located at point \mathbf{r} in the interstitial region. The exchange-correlation hole exhibits extra peaks where the hole overlaps the core charge density (peaks in the graph of ρ) of neighboring atoms.

ergy provide substantial evidence that the core charge density has an important effect on the exchange-correlation energy in the interstitial region.

The spatial extent of the exchange-correlation hole is a measure of the size of the neighborhood which is expected to influence the exchange-correlation energy at a point. Accurate calculations of the exchange-correlation hole in real systems show that it extends up to distances that are comparable to the interatomic distances. This has been directly observed in calculations in silicon,³⁵ while calculations of the electron pair-correlation function in diamond,³⁶ and metallic lithium³⁷ indicate that a similar feature is present in those systems.³⁸ Additionally, the radius of the exchange-correlation hole obtained through LDA, $\langle 1/r_{xc} \rangle = \epsilon_{xc}^{LDA}$, provides an estimate of the spatial extent of the true exchange-correlation hole. For all the metals investigated in this work, this approximate exchange-correlation hole radius is comparable to the distance separating neighboring atoms.

Another indication of the influence of the core charge density on the interstitial exchange-correlation energy is that the exchange-correlation hole exhibits a surprising feature when a reference electron is located in the neighborhood of an isolated atom or a molecule. The exchange-correlation hole possesses not only a peak centered on the reference electron but also exhibits additional peaks where the exchange-correlation hole overlaps the core of the neighboring atoms, as illustrated in Fig. 7. Whenever the reference electron is located in regions of relatively low electronic density, the exchange-correlation hole tends to “leak” to the nearest region where the charge density is large. This effect is especially large for isolated atoms,³⁹ molecules,⁴⁰ and surfaces.⁴¹ This is responsible for the incorrect asymptotic behavior of the LDA exchange correlation as the reference electron is moved to infinity.⁴⁰ LDA predicts an exchange-correlation hole which follows the reference electron as it is moved away, whereas the true exchange-correlation hole actually largely remains in the atom, the molecule, or near the surface. This is clearly an indication that regions of large charge density can in principle have an impact on the exchange correlation energy up to arbitrarily large distance. Accurate quantum Monte Carlo calculations have shown a similar effect in extended solids. Calculations in silicon³⁵ show the presence of these multiple peaks while calculations on metallic lithium³⁷ and in diamond³⁶ show multiple peaks

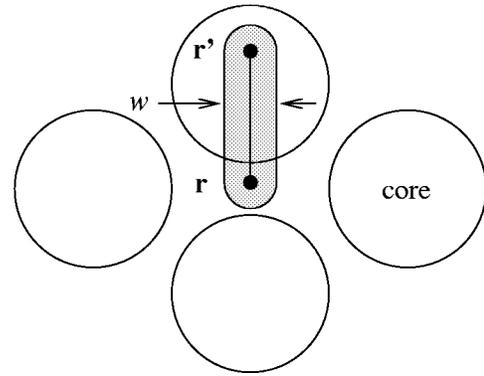


FIG. 8. Idealized shape of the region whose charge density is expected to influence the electron pair correlation function between points \mathbf{r} and \mathbf{r}' .

in the electron pair correlation function from which multiple peaks in the exchange-correlation hole can be inferred.

C. Linearity of P_{xc} in concentration

While we have motivated that Δe_{xc} in the interstitial region is mainly determined by the charge density in the cores we still have to describe how a linear dependence on concentration arises. Let us first define a few quantities. The exchange-correlation energy per unit volume $e_{xc}(\mathbf{r})$ at a point \mathbf{r} can be expressed in terms of the electron pair-correlation function

$$e_{xc}(\mathbf{r}) = \frac{1}{2} \int g(\mathbf{r}, \mathbf{r}') d^3 \mathbf{r}', \quad (3)$$

where $g(\mathbf{r}, \mathbf{r}')$ is a functional of the charge density that embodies all the information regarding the electron pair-correlation function. More specifically,

$$g(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} \rho(\mathbf{r}) (\bar{h}[\mathbf{r}, \mathbf{r}'] - 1) \rho(\mathbf{r}'),$$

where $\bar{h}(\mathbf{r}, \mathbf{r}')$ is the coupling-constant-averaged electron pair correlation of the system, as defined, for example, in Ref. 42. [The explicit functional dependence of $e_{xc}(\mathbf{r})$ and $g(\mathbf{r}, \mathbf{r}')$ on the charge density ρ has been omitted to simplify the notation].

The difference between the true and the LDA exchange-correlation follows a similar structure:

$$\Delta e_{xc}(\mathbf{r}) = \frac{1}{2} \int \Delta g(\mathbf{r}, \mathbf{r}') d^3 \mathbf{r}', \quad (4)$$

where $\Delta g(\mathbf{r}, \mathbf{r}')$ is defined as the difference between the true value of $g(\mathbf{r}, \mathbf{r}')$ and its approximation obtained with the LDA. With these definitions in hand, we now introduce a plausible restriction to the functional dependence of $\Delta g(\mathbf{r}, \mathbf{r}')$ on the charge density which guarantees that the shift in the exchange-correlation energy density in the interstitial region is linear in concentration. We consider $\Delta g(\mathbf{r}, \mathbf{r}')$ to be a functional of the charge density in a relatively narrow region surrounding the segment joining points \mathbf{r} and \mathbf{r}' . The width w of this region is assumed to be small relative to interatomic distances, as illustrated in Fig. 8.

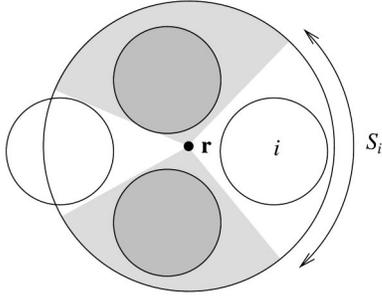


FIG. 9. The exchange-correlation energy at a point \mathbf{r} in the interstitial region can be written as a sum of integrals over disjoint sectors S_i , each of which represents the contribution coming from one particular neighboring atom i .

Under this assumption, the integration of $\Delta g(\mathbf{r}, \mathbf{r}')$ over \mathbf{r}' in Eq. (3) is a linear operation: $\Delta e_{xc}(\mathbf{r})$ reduces to a sum of integrals over disjoint sectors S_i (see Fig. 9), each of which representing the contribution coming from one particular neighboring atom i :

$$\Delta e_{xc}(\mathbf{r}) = \sum_i \int_{S_i} g(\mathbf{r}, \mathbf{r}') d^3 \mathbf{r}'.$$

If w is not too large compared to interatomic distances, there will be only a small coupling between the charge density of neighboring sectors and a term corresponding to sector S_i will mainly depend on the charge density in sector S_i .

These approximations allow the difference between the LDA exchange-correlation energy per electron and the true one $\Delta \epsilon_{xc}$ in the interstitial region to be approximated by a sum of contributions coming from nearby atoms. The linearity of $\Delta \epsilon_{xc}(\mathbf{r})$ in the atomic concentration is thus a consequence of the well-known fact that only the spherical average of the exchange-correlation hole plays a role in determining the exchange-correlation energy.

D. Atom-specific correction

We can now give a simple interpretation to the constants Δe_{xc} and $\Delta \mu_{xc}$ introduced earlier. Since the correction to the LDA exchange correlation in the interstitial region can be approximated by a sum of *independent* contributions from neighboring atoms, we can approximate the correction to the exchange-correlation energy at a point in the interstitial region by replacing the exact contributions coming from each neighboring atom by the contributions of isolated atoms centered at the same locations, as illustrated in Fig. 10. Even though this procedure may seem approximate, including nonlocal effects, even in this simplified way, is an improvement over not including them at all.

Consider an isolated neutral atom i centered at the origin with charge density $\rho_i(\mathbf{r})$. This atom has a nonzero non-local exchange-correlation energy density $\Delta e_{xc}^i(\mathbf{r})$ at every point in space, which can be integrated over space to give its total nonlocal exchange-correlation energy

$$\Delta E_{xc}^i = \int \Delta e_{xc}^i(\mathbf{r}) d^3 \mathbf{r}.$$

The quantity ΔE_{xc}^i will of course change if the charge density in the neighborhood of this atom is perturbed by an

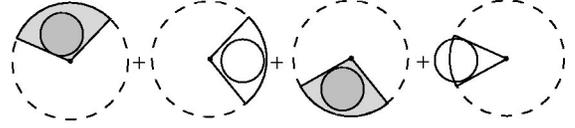


FIG. 10. Approximating the nonlocal exchange-correlation energy as a sum of independent contributions.

amount $\delta \rho(\mathbf{r})$. Provided that $\delta \rho(\mathbf{r})$ is not too large, we can approximate the true dependence of ΔE_{xc}^i on $\delta \rho(\mathbf{r})$ by an expansion of the true exchange-correlation function in terms of a series of homogeneous functional of $\delta \rho(\mathbf{r})$ of degree $0, 1, 2, \dots$

The zero order term of this expansion is simply the value of ΔE_{xc}^i at the ‘‘point’’ of expansion, that is, evaluated at the atomic charge density $\rho_i(\mathbf{r})$. Thus the zero-order correction to LDA consists in adding the energy $\Delta e_{xc}^i(\mathbf{r})$ at every point in the neighborhood of atom i . In accordance with our earlier assumptions we consider the dependence on \mathbf{r} to be weak outside of the core. The shift in exchange-correlation energy in the interstitial region introduced earlier is thus simply the sum of the atom-specific shift of the neighboring atoms

$$\Delta e_{xc} = \sum_i \Delta e_{xc}^i.$$

The first order term of this functional expansion can be expressed in terms of the functional derivative of the nonlocal exchange-correlation energy evaluated at the atomic charge density

$$\begin{aligned} \delta \Delta E_{xc}^i &= \int \left. \frac{\delta \Delta E_{xc}^i}{\delta \rho} \right|_{\rho=\rho_i} \delta \rho(\mathbf{r}) d\mathbf{r} \\ &\equiv \int \Delta \mu_{xc}^i(\mathbf{r}) \delta \rho(\mathbf{r}) d\mathbf{r}, \end{aligned}$$

which defines an atom-specific correction to the exchange-correlation potential $\Delta \mu_{xc}^i(\mathbf{r})$. As before, assuming a weak dependence on \mathbf{r} and summing the contributions of neighboring atoms provides the shift in the exchange-correlation potential

$$\Delta \mu_{xc} = \sum_i \Delta \mu_{xc}^i.$$

In the same spirit as in the beginning of this section, higher order terms of the functional expansion are neglected.

The nonlocal exchange-correlation pressure (P_{xc}) is thus more than a fitting parameter in an empirical equation of state. Under plausible assumptions, this pressure can be interpreted as the sum of the zero-order corrections to the LDA exchange-correlation energy of each atom present in a given compound.

V. CONCLUSION

This paper’s first objective is to emphasize that LDA’s overestimation of phonon frequencies and elastic moduli is essentially a consequence of its underestimation of the equilibrium volume. In a large class of materials, calculated elastic properties are in good agreement with experiment when

the LDA calculations are performed at the experimental volume.

Given the importance of obtaining the correct equilibrium volume, we propose a simple and yet effective way to correct LDA's bias: apply a negative pressure such that the equilibrium volume agrees with experiment. We have argued that this so-called nonlocal exchange-correlation pressure originates from the nonlocal contribution to the exchange-correlation energy per unit volume in the interstitial region of the solid.

This method would be of limited usefulness if one needed a different nonlocal exchange-correlation pressure for each compound. Fortunately, we have observed that the nonlocal exchange-correlation pressure of a compound can be accurately determined by taking the concentration-weighted average of the nonlocal exchange-correlation pressure of elements. This linear relationship holds for nearly all the metallic, ionic, and covalent systems we have investigated and we are confident of its wide applicability.

We then propose a simple explanation for this surprising linear behavior which relies on the following main assumptions.

(i) The nonlocal contribution to the exchange-correlation energy in the interstitial region mainly originates from the large nonuniformity of the core charge density and not from the weak nonuniformities of the interstitial charge density.

(ii) This nonlocal correction can be approximately expressed as a sum of independent contributions coming from nearby atoms.

(iii) These contributions take the form of a nearly uniform, atom-specific, and charge-density-independent correction to the exchange-correlation energy density in the interstitial region.

These assumptions, which are crucial in obtaining a linear dependence between the nonlocal exchange-correlation pressure and concentration, are guided by the results of previous investigations which have determined the exact exchange-correlation energy in simple systems.

Our results have implications at two different levels. (i) A simple linear interpolation scheme provides a simple and accurate way to correct LDA's volume underestimation. Furthermore, this correction is often sufficient to dramatically improve the accuracy of calculated elastic properties as well. (ii) The fact that such a simple scheme performs so well provides strict constraints on the behavior of the nonlocal corrections to the LDA and provides helpful clues in the continuing search for better exchange-correlation functionals. Atom-specific corrections determined from isolated-atom calculations would appear a promising approach.

ACKNOWLEDGMENTS

This work was sponsored in part by the National Science Foundation under Contract No. DMR9501856. Axel van de Walle acknowledges support from the Natural Sciences and Engineering Research Council of Canada.

¹P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).

²W. Kohn and L.J. Sham, *Phys. Rev.* **140**, A1133 (1965).

³R.O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).

⁴J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, and C. Fiolhais, *Phys. Rev. B* **46**, 6671 (1992).

⁵G. Ortiz and P. Ballone, *Phys. Rev. B* **43**, 6376 (1991).

⁶I.-H. Lee and R.M. Martin, *Phys. Rev. B* **56**, 7197 (1997).

⁷A. Khein, D.J. Singh, and C.J. Umrigar, *Phys. Rev. B* **51**, 4105 (1995).

⁸C. Filippi, D.J. Singh, and C.J. Umrigar, *Phys. Rev. B* **50**, 14 947 (1994).

⁹V. Ozolins and M. Körling, *Phys. Rev. B* **48**, 18 304 (1993).

¹⁰R. Stadler, W. Wolf, R. Podloucky, G. Kresse, J. Furthmüller, and J. Hafner, *Phys. Rev. B* **54**, 1729 (1996).

¹¹R.E. Cohen, *Nature (London)* **358**, 136 (1992).

¹²W. Zhong, D. Vanderbilt, and K.M. Rabe, *Phys. Rev. Lett.* **73**, 1861 (1994).

¹³W. Zhong and D. Vanderbilt, *Phys. Rev. Lett.* **74**, 2587 (1995).

¹⁴A. Dal Corso, A. Pasquarello, and A. Baldereschi, *Phys. Rev. B* **56**, R11 369 (1997).

¹⁵G.J. Ackland, in *Alloy Modelling and Design*, edited by G. Stocks and P. Turchi (The Minerals, Metals and Materials Society, Pittsburgh, PA, 1994), pp. 149–153.

¹⁶A. van de Walle, G. Ceder, and U.V. Waghmare, *Phys. Rev. Lett.* **80**, 4911 (1998).

¹⁷J.D. Althoff, D. Morgan, D. de Fontaine, M. Asta, S.M. Foiles, and D.D. Johnson, *Phys. Rev. B* **56**, R5705 (1997).

¹⁸W.B. Pearson, P. Villars, and L. Calvert, *Pearson's Handbook of Crystallographic Data for Intermetallic Phases* (American Society for Metals, Metals Park, OH, 1985).

¹⁹*Powder Diffraction File: Inorganic Volume* (JCPDS International Centre for Diffraction Data, Swarthmore, PA, 1970).

²⁰Y.S. Touloukian, *Thermophysical Properties of Matter, TPRC Data Series* (IFI/Plenum, New York, 1970–1979).

²¹G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11 169 (1996).

²²G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).

²³D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).

²⁴J.C. Phillips and L. Kleinman, *Phys. Rev.* **116**, 287 (1959).

²⁵Lattice constants of Ag, Al (LDA), Au, Cu, Na, Ni, Pd, Pt, Si, and V differed by less than 0.6%. Lattice constant of Ti (LDA and GGA) and Al (GGA) differed by about 1%.

²⁶For relaxation steps the Methfessel-Paxton smearing method of order 1 with a width of 0.1 eV is used instead, to avoid the problem that accurate forces cannot be obtained with the tetrahedron method.

²⁷D.M. Ceperley and B.J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).

²⁸J.P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).

²⁹The parent lattice is defined as the set of all atomic positions in the crystal, disregarding the type of atom occupying each site. As is commonly done, we also disregard small relaxation away from the ideal lattice positions when labeling parent lattices. For example, the $L1_0$ structure is considered to have an fcc parent lattice despite its tetragonal distortion.

³⁰S.K. Mishra and G. Ceder, *Phys. Rev. B* **59**, 6120 (1999).

³¹C.-O. Almbladh and A.C. Pedroza, *Phys. Rev. A* **29**, 2322 (1984).

- ³²O.V. Gritsenko, R. van Leeuwen, and E.J. Baerends, Phys. Rev. A **52**, 1870 (1995).
- ³³M. Fuchs, M. Bockstedte, E. Pehlke, and M. Scheffler, Phys. Rev. B **57**, 2134 (1998).
- ³⁴X.J. Kong, C.T. Chan, K.M. Ho, and Y.Y. Ye, Phys. Rev. B **42**, 9357 (1990).
- ³⁵R.Q. Hood, M.Y. Chou, A.J. Williamson, G. Rajagopal, R.J. Needs, and W.M.C. Foulkes, Phys. Rev. Lett. **78**, 3350 (1997).
- ³⁶S. Fahy, X.W. Wang, and S.G. Louie, Phys. Rev. Lett. **65**, 1478 (1990).
- ³⁷G. Sugiyama, G. Zerah, and B.J. Alder, Physica A **156**, 144 (1989).
- ³⁸The electron pair-correlation function is directly related to the exchange-correlation hole (Ref. 3).
- ³⁹M. Ernzerhof, J.P. Perdew, and K. Burke, Top. Curr. Chem. **180**, 1 (1996).
- ⁴⁰R. van Leeuwen, O.V. Gritsenko, and E.J. Baerends, Top. Curr. Chem. **180**, 107 (1996).
- ⁴¹N.D. Lang, in *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N.H. March (Plenum Press, New York, 1983), Chap. 5, pp. 309–376.
- ⁴²W. Kohn and P. Vashishta, in *Theory of the Inhomogeneous Electron Gas* (Ref. 41), Chap. 2, pp. 79–147.