

Screened-exchange stress tensor in density functional theory

Michael C. Gibson,* Stuart Brand,[†] and Stewart J. Clark[‡]

Department of Physics, University of Durham, Durham DH1 3LE, United Kingdom

(Received 18 January 2006; published 29 March 2006)

We have derived and implemented the theory for directly computing the contribution to the stress tensor $\sigma_{\alpha\beta}$ from exchange and correlation effects when using the screened exchange method in Kohn-Sham density functional theory (DFT). The method is also applicable when using related functionals such as Hartree-Fock or exact exchange. This paves the way for full geometric relaxation of a crystal structure when using such functionals, without the need for expensive numerical differentiation. The theory is derived in reciprocal space, appropriate for plane-wave based calculations, and has been implemented within an efficient, and fully parallel, DFT code. As a simple demonstration of the validity of this method we calculate the variation in stress as both hydrostatic and shear strains are applied to silicon when using screened exchange.

DOI: [10.1103/PhysRevB.73.125120](https://doi.org/10.1103/PhysRevB.73.125120)

PACS number(s): 71.10.-w, 71.15.Ap, 71.15.Mb

I. INTRODUCTION

Density functional theory (DFT)^{1,2} is, in principle, an exact theory of many-electron systems which allows us to calculate the properties of condensed matter from first principles. In practice, of course, approximations must be made in order that the calculations become computationally feasible. Most commonly used are the local density approximation (LDA),^{2,3} and generalized gradient approximations (GGA),^{4–6} to the exchange-correlation functional, defined within the Kohn-Sham formalism.² The accuracy of these functionals depends very much on the nature of the system under study, and on which particular properties we wish to calculate. More advanced exchange-correlation functionals exist that may provide greater accuracy than the LDA at the expense of increased computational demand. Among these are orbital-based functionals such as screened-exchange (sX-LDA), Hartree-Fock (HF), and exact exchange (EXX)^{7–12} (while HF actually predates the development of DFT, it can still be considered implicitly to be a density functional as there is only ever one solution of the Hartree-Fock equations corresponding to a particular density). The primary application of sX-LDA and EXX, has so far been in the calculation of band structures, where they have generally performed well in overcoming the so called “band gap problem” associated with the LDA and GGA. sX-LDA has also recently been shown to give good results for the energies of defect states.¹³ The primary application of HF has been the study of molecular systems, while in crystalline solids, it is well known to greatly overestimate the electronic band-gap. Calculation of other properties of solids has not yet been extensively explored with these orbital-based functionals.

One very important application of DFT is to calculate the geometry of a crystal structure by finding the structure that minimizes the total energy. To do this efficiently, we need to be able to calculate derivatives of the energy with respect to both the atomic coordinates and the lattice vectors. Derivatives with respect to the coordinates, i.e., the forces on the atoms, can be calculated efficiently using the Hellman-Feynman theorem, and this has no contributions from exchange and correlation. However, derivatives with respect to

the lattice vectors, which relate to the stress tensor,^{14,15} $\sigma_{\alpha\beta}$, do have contributions from exchange and correlation, and this must be taken into account if a new exchange-correlation functional is to be used. While it is always possible to calculate the stress numerically by using finite differences, this will generally require several evaluations of the total energy, which could prove computationally restrictive. Ideally, we would like to evaluate the stress directly from the ground state wave function or density.

Nielsen and Martin¹⁴ described methods to evaluate the contribution to stress from each of the components of the Kohn-Sham energy functional in a plane-wave basis set when using the LDA. In this work, we show how to calculate the contribution from the nonlocal exchange-correlation term when using screened exchange. We note that our method is also applicable when using Hartree-Fock (HF) or exact exchange (EXX), although these functionals require us to deal appropriately with a singularity when integrating over the Brillouin zone, following Gygi and Baldereschi.⁷

In Sec. II, we present a brief review of the theory of screened exchange, as an extension to standard Kohn-Sham DFT. We then use the definition of the stress tensor to obtain our equation for the contribution to stress from the nonlocal exchange-correlation energy. We also discuss how the singularity can be dealt with when using HF and EXX and how this would affect the calculation of stress. In Sec. III we describe our implementation of the screened exchange method, along with the method for calculating stress, within an efficient plane-wave code. Finally, in Sec. IV, we present results of tests we have carried out in which we use our method to calculate the stress tensor in silicon, and we conclude in Sec. V.

II. THEORY

In standard Kohn-Sham DFT, the basic expression for the total energy of a condensed matter system, with ground-state electron density $\rho(\mathbf{r})$ is given by

$$E[\rho] = T_S + V_H + V_{\text{ext}} + V_{L-I} + E_{\text{XC}}, \quad (1)$$

where T_S is the kinetic energy of a system of noninteracting electrons of density $\rho(\mathbf{r})$, V_H is the Hartree energy, i.e., the

Coulombic self-energy of a nonquantized distribution of charge with density $\rho(\mathbf{r})$, V_{ext} is the potential energy of the electrons due to the external potential created by the positively charged ions or nuclei, and $V_{L,I}$ is the potential energy due to the mutual repulsion of the ions or nuclei. The remaining term E_{XC} is the exchange-correlation energy that contains all the many-body effects that are not included in the other terms. This term must always be approximated to some degree in practical calculations.

In the LDA, for example, the exchange-correlation energy is approximated such that the energy per electron at a given point \mathbf{r} is simply a function of the density at \mathbf{r}

$$\varepsilon_{\text{XC}}(\mathbf{r}) = \varepsilon_{\text{XC}}^{\text{HEG}}[\rho(\mathbf{r})], \quad (2)$$

where $\varepsilon_{\text{XC}}^{\text{HEG}}(\rho)$ is the exchange-correlation energy per electron in a homogeneous electron gas of uniform density ρ .

In screened exchange, the exchange-correlation energy is split into two parts

$$E_{\text{XC}} = E_{\text{XC}}^{\text{NL}} + E_{\text{XC}}^{\text{LOC}} \quad (3)$$

where $E_{\text{XC}}^{\text{NL}}$ is the nonlocal exchange-correlation energy which, in real space, is given by

$$E_{\text{XC}}^{\text{NL}} = -\frac{1}{2} \sum_{\mathbf{ikjq}} \int \int d\mathbf{r} d\mathbf{r}' \frac{\phi_{\mathbf{ik}}^*(\mathbf{r}) \phi_{\mathbf{ik}}(\mathbf{r}) \phi_{\mathbf{jq}}(\mathbf{r}) \phi_{\mathbf{jq}}^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} e^{-k_s |\mathbf{r} - \mathbf{r}'|}, \quad (4)$$

where the $\phi_{\mathbf{ik}}(\mathbf{r})$ are the Kohn-Sham orbitals, and k_s is the reciprocal screening length. In reciprocal space, this contribution is given by¹¹

$$E_{\text{XC}}^{\text{NL}} = -\frac{2\pi}{\Omega} \sum_{\mathbf{ikjq}} \sum_{\mathbf{G}} \frac{|C_{\mathbf{jqik}}(\mathbf{G})|^2}{|\mathbf{q} - \mathbf{k} + \mathbf{G}|^2 + k_s^2}, \quad (5)$$

where Ω is the volume of the unit cell, and the $C_{\mathbf{jqik}}(\mathbf{G})$ are defined as

$$C_{\mathbf{jqik}}(\mathbf{G}) \equiv \sum_{\mathbf{G}'} c_{\mathbf{ik}}^*(\mathbf{G}') c_{\mathbf{jq}}(\mathbf{G} + \mathbf{G}'), \quad (6)$$

where the $c_{\mathbf{ik}}(\mathbf{G})$ are the reciprocal space coefficients of the Kohn-Sham orbitals.

The screening length k_s is usually given a value of ~ 0.764 bohr⁻¹, which is the value of the Thomas-Fermi screening length in an homogeneous electron gas of “natural” density, i.e., the density for which the total energy per electron is minimized. In this case k_s is a universal constant that does not depend on the density of the particular system under study.

The local part of the exchange-correlation energy $E_{\text{XC}}^{\text{LOC}}$ is defined in a very similar way to the LDA, where the local exchange-correlation energy per electron is given by

$$\varepsilon_{\text{XC}}^{\text{loc}}(\mathbf{r}) = \varepsilon_{\text{XC}}^{\text{HEG}}(\rho(\mathbf{r})) - \varepsilon_{\text{XC}}^{\text{NLHEG}}[\rho(\mathbf{r})] \quad (7)$$

where $\varepsilon_{\text{XC}}^{\text{NLHEG}}$ is the nonlocal exchange-correlation energy per electron in an homogeneous electron gas. Contributions to the stress tensor from this local energy can be treated in exactly the same way as in the LDA.

The HF method is very similar to sX-LDA, except that k_s is set to zero, and there is no treatment of correlation. Both

methods can be considered to be density functionals within generalized Kohn-Sham (GKS) schemes. In such schemes the orbitals are no longer solutions of a purely local potential, but rather a potential in which the exchange-correlation part has a nonlocal matrixlike contribution. In practice, this potential does not have to be explicitly calculated, as the solution can be obtained by minimizing the total energy with respect to the orbitals. These methods can still be described as density functionals because there is only ever one solution corresponding to a particular density. The EXX method is similar to HF in the sense that k_s is set to zero. However, in EXX correlation also is included, usually in a LDA-like fashion, as is done in sX-LDA. The difference between EXX and the other two schemes is that it is a pure Kohn-Sham method in which the orbitals are solutions in a purely local potential.

In order to obtain the nonlocal contribution to the stress, we start with the equation for the total stress tensor in reciprocal space.¹⁴ This is defined in terms of a scaling of space $\epsilon_{\alpha\beta}$

$$\mathbf{K}_\alpha \rightarrow \sum_{\beta} (\delta_{\alpha\beta} - \epsilon_{\alpha\beta}) \mathbf{K}_\beta, \quad (8)$$

where \mathbf{K} is a general vector in reciprocal space. The total stress tensor is given by

$$\sigma_{\alpha\beta} = \frac{1}{\Omega} \frac{\partial E}{\partial \epsilon_{\alpha\beta}}. \quad (9)$$

This means that to obtain the contribution from nonlocal exchange and correlation, we essentially have to differentiate Eq. (5) with respect to the reciprocal lattice vectors. The expression for the nonlocal exchange-correlation energy includes both lattice vectors and orbital coefficients. However, we do not have to differentiate the coefficients because the total ground state energy is stationary with respect to these. We thus have

$$\begin{aligned} \frac{1}{\Omega} \frac{\partial E_{\text{XC}}^{\text{NL}}}{\partial \epsilon_{\alpha\beta}} = & -\frac{2\pi}{\Omega^2} \sum_{\mathbf{ikjq}} \sum_{\mathbf{G}} \frac{|C_{\mathbf{jqik}}(\mathbf{G})|^2}{|\mathbf{q} - \mathbf{k} + \mathbf{G}|^2 + k_s^2} \\ & \times \left(\frac{2(\mathbf{q} - \mathbf{k} + \mathbf{G})_\alpha (\mathbf{q} - \mathbf{k} + \mathbf{G})_\beta}{|\mathbf{q} - \mathbf{k} + \mathbf{G}|^2 + k_s^2} - \delta_{\alpha\beta} \right), \end{aligned} \quad (10)$$

which is the contribution to the stress tensor from nonlocal exchange and correlation when using screened exchange.

This is also applicable to HF and EXX calculations, in which the screening constant is set to zero. However, with zero screening constant, a singularity appears both in the expression for the exchange energy and the corresponding contribution to the stress. If we do not take this into account then a rather coarse sampling of the Brillouin zone may not be adequate. Following Gygi and Baldereschi⁷ the singularity can be dealt with by adding the following term to the expression for the exchange energy:

$$\frac{2\pi}{\Omega} \sum_{ikj} |C_{j\mathbf{k}\mathbf{i}\mathbf{k}}(\mathbf{0})|^2 \left(\sum_{\mathbf{q}} F(\mathbf{q} - \mathbf{k}) - \int_{\text{B.Z.}} d\mathbf{q} F(\mathbf{q} - \mathbf{k}) \right), \quad (11)$$

where $F(\mathbf{p})$ is a function sharing periodicity with the reciprocal lattice, and which has the same singularity as the exchange energy, i.e., $F(\mathbf{p}) \rightarrow 1/p^2$ as $p \rightarrow 0$. The choice of F is essentially arbitrary, except that it should be smoothly varying away from the singularity to minimize the size of the Monkhorst-Pack grid required for summation. So long as the grid is converged, then the form chosen for F should not affect the results. The form that we use for F is as follows:

$$F(\mathbf{p}) = \sum_{\mathbf{G}} \frac{S(|\mathbf{p} - \mathbf{G}|)}{|\mathbf{p} - \mathbf{G}|^2}, \quad (12)$$

where S is a sinusoidal envelope function of the following form:

$$S(x)_{|x| \leq w} = \frac{1}{2} \left[\cos\left(\frac{\pi x}{w}\right) + 1 \right], \quad (13)$$

$$S(x)_{|x| > w} = 0.$$

The larger the width of this envelope w the smoother F is away from the singularity; w should therefore be set substantially larger than the typical separation of the \mathbf{k} points. The addition of this term to the exchange energy must also be carried through to the calculation of the stress, which essentially means differentiating F with respect to \mathbf{p} . It is therefore useful to choose a form for F , such as that given in Eq. (12), which is easily differentiable.

III. IMPLEMENTATION

The screened exchange and Hartree-Fock methods, along with the method for calculating stress introduced here in Eq. (10) have been implemented within the CASTEP code.^{16,17} Now with the ability to calculate stress when using screened exchange or Hartree-Fock it will be possible to perform geometry optimization and variable-cell molecular dynamics, in addition to the usual band structure calculations that have so far been the main application of these functionals.

To perform screened exchange calculations efficiently we make extensive use of Fourier transforms, largely following Chawla and Voth,¹¹ however we extend this work to include multiple \mathbf{k} points and screening. The functions, $C_{j\mathbf{q}\mathbf{i}\mathbf{k}}(\mathbf{G})$, defined in Eq. (6) can be calculated from the Kohn-Sham orbitals in real space

$$C_{j\mathbf{q}\mathbf{i}\mathbf{k}}(\mathbf{G}) = \text{FT}[\phi_{j\mathbf{q}}(\mathbf{r})\phi_{i\mathbf{k}}^*(\mathbf{r})]. \quad (14)$$

Therefore calculation of the full set of functions, and hence the screened exchange energy, scales roughly as $N_b^2 N_k^2 N_p \log N_p$, where N_b is the number of occupied bands, N_k is the number of \mathbf{k} points, and N_p is the number of plane waves in the basis set. This is much more efficient than the $N_b^2 N_k^2 N_p^2$ implied by Eq. (6). Similarly, the gradient of the screened exchange energy with respect to the orbitals, required for self-consistent solution of the Kohn-Sham equations, can be obtained via

$$\frac{\delta E_{\text{XC}}^{\text{NL}}}{\delta c_{i\mathbf{k}}^*(\mathbf{G})} = \frac{8\pi}{\Omega} \text{FT} \left\{ \sum_{j\mathbf{q}} \phi_{j\mathbf{q}}(\mathbf{r}) \text{FT} \left[\frac{C_{j\mathbf{q}\mathbf{i}\mathbf{k}}^*(\mathbf{G})}{|\mathbf{q} - \mathbf{k} + \mathbf{G}|^2 + k_s^2} \right] \right\}, \quad (15)$$

which again scales as $N_b^2 N_k^2 N_p \log N_p$. Examining Eq. (10) we see that calculation of the stress also scales as $N_b^2 N_k^2 N_p \log N_p$, and does not add significantly to the cost of the calculation.

In terms of memory, the largest objects we need to store scale in size as $N_b^2 N_k^2 N_p$. Calculation of stress does not involve the storage of extra arrays of this size and so the extra memory requirement is negligible.

Due to the large computational cost of the screened exchange functional, parallelization across large numbers of processors is particularly desirable. Our implementation of screened exchange, including the ability to calculate stress, allows simultaneous distribution of data storage and processing across both \mathbf{k} points and \mathbf{G} vectors. Distribution across \mathbf{k} points is the more efficient because the calculations for each \mathbf{k} point are almost independent; scaling is approximately linear with respect to the number of processors. Distribution across \mathbf{G} vectors is less efficient because communication between processors is required whenever a Fourier transform is applied. Nevertheless, significant increase in speed can be achieved from \mathbf{G} vector distribution, as well as a reduction in memory requirements per processor.

Further gains in efficiency can be achieved by making use of crystal symmetries that may be present in the structure under study. In a LDA or GGA calculation, if two or more \mathbf{k} points are related by symmetry, then we need only explicitly treat one of these points, reducing processor time and memory accordingly. When using screened exchange and related functionals, involving Eq. (5), the symmetry is partially broken by terms involving two different \mathbf{k} points. In our implementation we define a full Monkhorst-Pack grid of “ \mathbf{q} points” along with the corresponding symmetry-reduced set of \mathbf{k} points. For reasons of efficiency, we generate the full set of orbitals for the \mathbf{q} points and so symmetry does not reduce memory requirements. However we do still get a reduction in processor time because we only need to include the symmetry-reduced \mathbf{k} points in the summation.

IV. TESTS ON SILICON

We have verified the validity of our method for calculating stress and its computational implementation with a simple test on silicon. We use an eight-atom cube as the unit cell, and use a plane wave cut-off energy of 350 eV, sampling the Brillouin zone with a $4 \times 4 \times 4$ Monkhorst-Pack grid of \mathbf{k} points, which is adequately converged. Standard nonlocal, norm-conserving, pseudopotentials of the Kleinman-Bylander form were used,^{18,19} and it should be noted that these were generated using the LDA.

The calculation of diagonal stress components was tested by applying hydrostatic strain to the silicon crystal, i.e., by varying the lattice parameter, a ($=b=c$), which defines the eight-atom cube of the unit cell. Three quantities were calculated as a function of a : the total energy per cell, the total

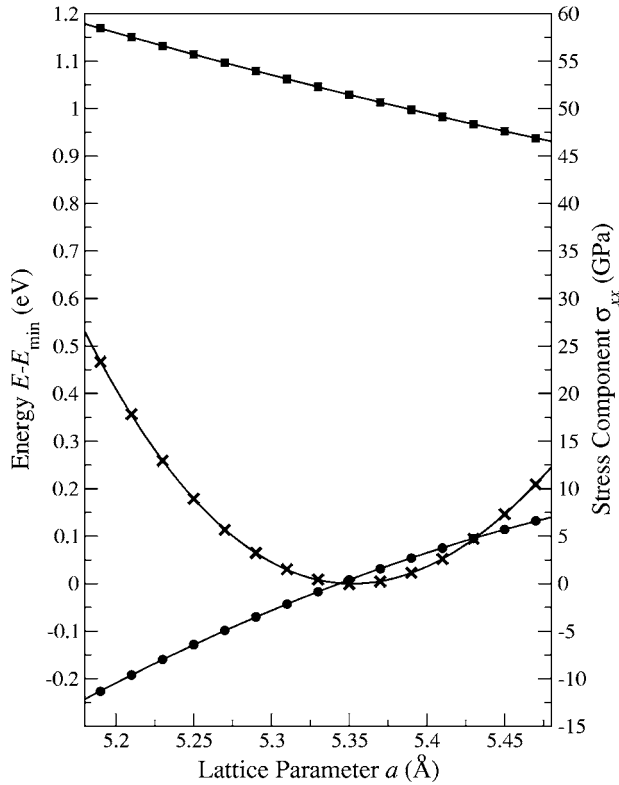


FIG. 1. Total energy per eight-atom unit cell, calculated using screened exchange, relative to the energy at equilibrium lattice constant (crosses), along with the diagonal component of the total stress tensor σ_{xx} (dots), and the contribution to this stress component from the nonlocal screened exchange term (squares), as a function of the lattice parameter a ($=b=c$) under hydrostatic strain.

stress tensor, and the nonlocal screened exchange contribution to the stress tensor. Due to the symmetry of the crystal, all three diagonal components of the stress are equal, and the off-diagonal terms are all zero. Figure 1 shows these three quantities plotted together on the same graph. The total energy is plotted as a difference from its equilibrium value. This allows us to easily compare the point of zero stress with the point of minimum energy, which should coincide exactly in a perfectly converged calculation. We find that these points differ by ~ 0.01 Å, which is of similar order to what we find when using the LDA at this cut-off energy. We also observe that the nonlocal screened exchange contribution to the stress, and its gradient, both contribute significantly to the total.

The calculation of the off-diagonal terms was tested by applying shear strain to the crystal, i.e., by applying a displacement in the y direction proportional to the x coordinate. Again, we calculated the total energy, the total stress tensor, and the nonlocal screened exchange contribution to the tensor. As shear strain is applied, nonzero off-diagonal terms appear in the stress tensor, and these are plotted in Fig. 2. In this case the coincidence of the zero points is trivial due to symmetry, but we see that the nonlocal screened exchange contribution to the off-diagonal component is very small, making up only $\sim 4\%$ of the total. We

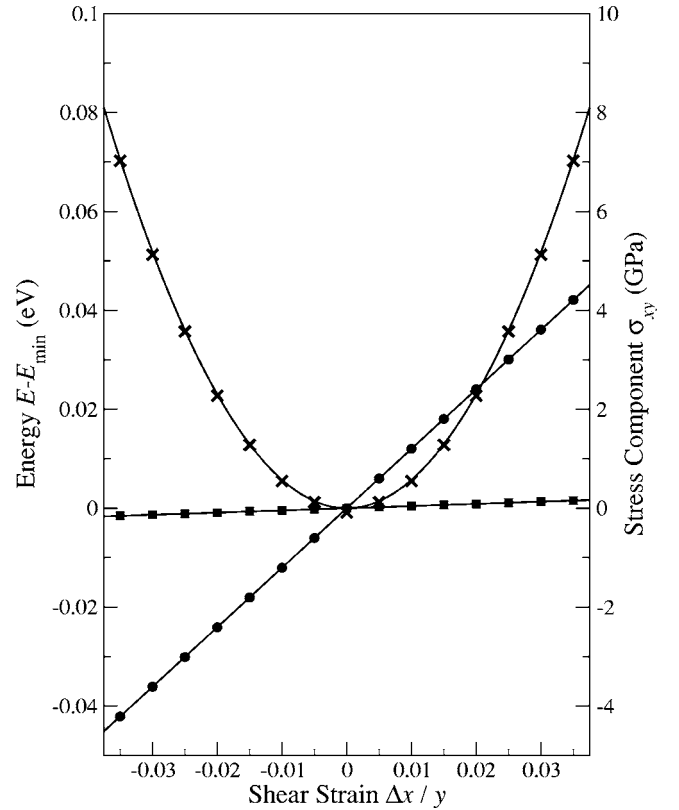


FIG. 2. Total energy per eight-atom unit cell, calculated using screened exchange, relative to the energy at equilibrium lattice constant (crosses), off-diagonal component of the total stress tensor σ_{xy} (dots), and the contribution to this stress component from the nonlocal screened exchange term (squares), as a function of shear strain $\Delta x/y$.

note that, when using the LDA, the off diagonal contribution to the stress tensor from exchange-correlation effects is always zero for any system.

V. CONCLUSIONS

We have shown how to calculate the contribution to the stress tensor from the nonlocal exchange-correlation energy when using screened exchange. This paves the way for efficient geometry optimization and variable-cell molecular dynamics when using this functional, without the need to take finite differences. The method has been efficiently implemented within the CASTEP code and demonstrated with a simple test on silicon. Discrepancies between the stress inferred numerically, and that calculated directly using our method, can be ascribed to finite basis effects since they are of a similar order to those found when using the LDA at the same cut-off energy. The equilibrium lattice parameters obtained from screened exchange are somewhat lower than both the LDA results and experimental results.²⁰ This may be a property of the functional itself, after all it has not been extensively tested as to its accuracy when calculating geometries. It may also be due to the fact that we are using LDA pseudopotentials which may not be appropriate for screened exchange calculations. However the question of how best to

define a pseudopotential appropriate for screened exchange remains open.

As we have mentioned, this method for calculating the contribution to the stress from screened exchange is also applicable to Hartree-Fock and exact exchange. It is likely that most, if not all, major developments in exchange-correlation functionals in the future will use EXX to treat exchange. Any such calculations involving efficient geom-

etry optimization, or variable-cell molecular dynamics, will require calculation of the exchange contribution to the stress tensor. We therefore expect the future scope of this work to be extensive.

ACKNOWLEDGMENT

M.C.G. acknowledges the EPSRC for financial support.

*Electronic address: m.c.gibson@durham.ac.uk

†Electronic address: stuart.brand@durham.ac.uk

‡Electronic address: s.j.clark@durham.ac.uk; URL: <http://cmt.dur.ac.uk/sjc>

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

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

³J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

⁴J. P. Perdew and Y. Wang, Phys. Rev. B **33**, R8800 (1986).

⁵J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 16533 (1996).

⁶J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

⁷F. Gygi and A. Baldereschi, Phys. Rev. B **34**, 4405 (1986).

⁸A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).

⁹A. Görling, Phys. Rev. B **53**, 7024 (1996).

¹⁰M. Städele, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. Lett. **79**, 2089 (1997).

¹¹S. Chawla and A. G. Voth, J. Chem. Phys. **108**, 4697 (1998).

¹²M. Städele, M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. B **59**, 10031 (1999).

¹³K. Xiong, J. Robertson, M. C. Gibson, and S. J. Clark, Appl. Phys. Lett. **87**, 183505 (2005).

¹⁴O. H. Nielsen and R. M. Martin, Phys. Rev. B **32**, 3792 (1985).

¹⁵R. M. Martin, *Electronic Structure: Basic Theory and Methods* (Cambridge University Press, Cambridge, United Kingdom, 2004).

¹⁶M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Condens. Matter **14**, 2717 (2002).

¹⁷S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne, Zeit. Kryst. **220**, 567 (2005).

¹⁸L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

¹⁹M. H. Lee, Ph.D. thesis, University of Cambridge, 1994.

²⁰*Semiconductors—Basic Data*, edited by O. Madelung (Springer, New York, 1996).