



PERGAMON

Solid State Communications xx (0000) xxx-xxx

**solid
state
communications**

www.elsevier.com/locate/ssc

A semi-empirical approach to electron correlations based on a non-local formalism

M.D. Watson*, C.Y. Fong

Department of Physics, University of California, Davis, CA 95616-8677, USA

Received 6 August 2002; received in revised form 17 August 2002; accepted 26 August 2002 by M. Cardona

Abstract

We present a new correlation function based on variational Monte Carlo results. The function is, for ease of implementation, semi-local, spin-independent, and semi-empirical. This approach is applied to semiconductor systems and the quantity of interest is the fundamental gap energy. We found that except for InN, these gap energies improved remarkably when compared to local density approximation results. © 2002 Published by Elsevier Science Ltd.

PACS: 71.20.Nr; 71.15.Mb

Keywords: A. Semiconductors; D. Electron–electron interactions

1. Introduction

In applying density functional theory (DFT) [1] to physical problems, the knowledge of the form of the relevant functionals is necessary in order to get proper results. Among them, the form of the exchange–correlation (XC) functional is still being constantly improved [2–6]. While the two parts of this functional are easily understood in terms of physical pictures, this still leaves the task of actually forming an expression that describes the correlation effects. Since the correlation plays a significant role in semiconducting systems, a more accurate description of its effects would lead to better theoretical predictive power [7]. The objective of this paper is to report the introduction of a new correlation functional within DFT which, when implemented appropriately, significantly improves upon the fundamental gap energy of some semiconducting materials predicted by the local density approximation (LDA) [2]. When compared to experimental values, our results were more accurate than the LDA, and attained at the same level of computational efficiency.

2. Introduction to exchange–correlation effects

The XC effect can be understood by visualizing the interactions between electrons with parallel or antiparallel spins. Consider a single test electron in a solid with spin up, being approached by another spin-up electron. The Pauli principle prevents these two electrons from getting too close spatially; hence, there is a region around the test electron from which a same-spin electron is excluded. Therefore, in terms of the correlated density, this region is same-spin depleted; this is known as the Fermi hole.¹ What happens if the electron that approaches has opposite spin? In that case, the Pauli principle is not a consideration. In principle, this electron is allowed to get as close to the test electron as it wishes. However, the Coulomb interaction now separates electrons, and this brings about another density ‘hole’, known as the Coulomb hole.² Both types of holes are

¹ In truth, the name ‘Fermi hole’ arises from quantum chemistry’s description of the exchange hole as determined by HF. Here, the exchange is calculated using DFT and the KS scheme, so technically speaking there is a difference. However, for this work, it is hereby redefined to mean the latter.

² Strictly speaking, the mapping of the spin decomposition to the exchange and correlation effects is not one-to-one. In general, the correlation effects also play a part when the spins are parallel, but for this description we can consider them to be washed out by the comparatively stronger exchange effects.

* Corresponding author.

E-mail address: mdwatson@lifshitz.ucdavis.edu (M.D. Watson).

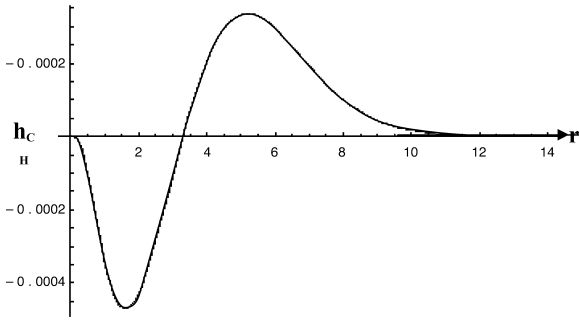


Fig. 1. A plot of the correlation function $h_{CH}(r)$ for a typical value of $r_s = 2$ for semiconductors.

characterized by an effective radius, which is the distance from the test electron where the excluded charge begins to pile up. These polarizations contribute to the electronic energy of the system. The Fermi hole is easily treatable with Hartree–Fock theory, while the correlation energy (defined as the difference between the exact energy and the HF energy, including the Coulomb hole polarization energy) has proven more difficult to handle. The well-known LDA was the first to offer a solution to the difficulty by assuming the XC functional at a point \mathbf{r} depends solely on the charge density at that point.

There are generally two different ways to improve the LDA results without resorting to a more computationally intensive Green function approach [8]. There is the method of add-on corrections, which essentially include higher-order density terms, such as gradient [3] and Laplacian [4]. This has been explored extensively. The other way is to add a non-local functional to the local functional $F(n)$ of Kohn and Sham [2]. These non-local functionals are generally based on the pair correlation function [5,6].

The pair correlation function can be formulated in terms of the Coulomb hole function $h(\mathbf{r}_1, \mathbf{r}_2)$, which reduces to its asymptotic limits at high and low density (based on a many-body approach):

$$\varepsilon_c = \frac{1}{2} \int d\mathbf{r}_2 \frac{h(\mathbf{r}_1, \mathbf{r}_2)}{r_{12}}. \quad (1)$$

In general, $h(\mathbf{r}_1, \mathbf{r}_2)$ is constructed to be the difference between the pair correlation function based on the exact wavefunction and the pair correlation function based on a reference wavefunction (e.g. HF, KS, etc.). The idea of the pair correlation function is that it describes the probability of finding two particles separated by a certain distance, r_{12} . This physical description necessarily puts limits on the behavior of the function. For example, the total effect of the Coulomb hole, as described above, is to exclude probability density characterizing the presence of a nearby electron from coming within a certain radius. If the density is being depleted from the vicinity near the test electron, then the charge must be going somewhere due to charge conservation. Thus, one condition on the Coulomb hole function is

that it must integrate to zero:

$$\int_0^\infty dr_{12} r_{12}^2 h(\mathbf{r}, r_{12}) = 0. \quad (2)$$

An example of a non-local pair-correlation function treating the uniform electron gas (therefore not applicable to semiconductor systems) is given below, taken from the first paper in Refs. [5,6]:

$$h^{\sigma\sigma'}(\mathbf{r}, r_{12}) = -n^{\sigma'}(\mathbf{r}) \frac{(1 - ar^2) \exp(-r_{12}^2/(r_c^{\sigma\sigma'})^2)}{1 + 3r_{12}^2/(2c_0(0)(r_s^{\sigma\sigma'})^3)}. \quad (3)$$

In Eq. (3), σ is the spin index, a is a normalization constant, r_{12} is the interelectronic spacing, r_c is the effective correlation radius, $c_0(0)$ is the leading coefficient in the high-density expansion of the electron gas, and r_s is taken from the definition of the electron number density. These parameters were determined from the uniform electron gas. We did not directly implement this formula into our program because this function is non-local (due to r_{12}), as well as being spin-dependent. Furthermore, the theory is based entirely on a short-ranged approximation to the Coulomb hole. This makes the long-ranged components unimportant, even in systems with large long-range contributions, such as the uniform electron gas. In addition, the behavior of the function near the test electron is much too strong, meaning that the hole depth is too large.

3. Implementation and calculation methods

Recently, a study has been done to compare the angular averaged correlation holes of second row atoms, except Na, as described by DFT and Variational Monte Carlo (VMC) [9]. This study particularly highlighted the XC hole and the pair correlation function of the valence shell in the silicon atom. It seems to correctly mimic the important behavior of the correlation function, both short- and long-range. The picture of the correct behavior was supplied (Fig. 1), but the function itself was not provided. Hence, we modified Eq. (3) so that it fits the picture provided by Cancio et al. [10]. This was accomplished through matching the picture shown in Fig. 1 to a mathematical expression. Fig. 1 shows the typical correlation function found in the VMC study done in Ref. [10], and all elements studied could be scaled to fit the same shape. The individual correlation functions are quite different due to the size of the atoms, number of valence electrons, etc. However, by scaling each function to an effective interaction radius instead of simply using the usual radial distance, it was found that each element's correlation function was identical, reflecting the pairwise nature of the correlation. Hence, the features of the correlation effect are universal. By using Eq. (3) as a starting point, we were able to adjust the behavior to match the graph of the correlation

Table 1
Tabulation of data using LDA, experiment, and the correlation function presented in this work. All data are in electron volts (eV)

	Si	GaAs	SiC	NaCl	InN
LDA	0.46	0.59	1.93	4.69	−0.1
CH	1.11	1.26	2.36	5.55	−0.2
Experiment	1.12 ^a	1.46 ^a	2.39 ^b	8.97 ^a	0.8 ^c

^a Ref. [16].

^b Ref. [17].

^c Ref. [15].

function, resulting in Eq. (4):

$$h_{\text{CH}}(r) = Ar \frac{(1 - ar^2)e^{-(r/r_c)^2} (e^{(r/2r_c)^2} - 1)}{2c_0 r_s^3 + 3r^{5/2}}. \quad (4)$$

In Eq. (4), A is an adjustable constant, and all other variables follow those given in Eq. (3). For ease of implementation, the function $h_{\text{CH}}(r)$ was assumed to be semi-local, in the sense that only nearby electrons (within r_c) were considered, and spin-independent (as opposed to Eq. (3)). With Eq. (4), the implementation is straightforward, and computation time is not significantly increased above the time used by the LDA treatment. Since the study leading to this function was based on silicon, this particular function is designed to empirically reproduce (through the adjustable constant, A) the fundamental gap of bulk silicon. From this expression, other materials can be explored.

Our first principles LDA code was based on the Ceperley–Alder [11] XC form, as parameterized by Perdew and Zunger [12]. In running these calculations of bulk materials, a two-atom basis arranged in the usual zinc-blende structure with the appropriate unit cell, or a four-atom basis for the Wurtzite structure was used. Each lattice constant was determined through energy minimization, and thus turned out to be generally between 2 and 3% smaller than the experimental values. The kinetic energy cutoff for the plane wave basis was set between 25 and 35 Ry, after determining that more plane waves did not change the energy of the structures.³ Norm-conserving pseudopotentials were used to treat the electron–ion interaction [13]. The code was run at the Lawrence Livermore National Laboratory computing center, on DEC alpha machines with 600MB of memory. The calculations were run until the electronic self-consistent energy convergence was at least 10^{-6} Ry. The modifications required to the main program consisted of nothing more difficult than a simple integration routine, without any matrix diagonalization required. The total addition was in the order of 200 lines (some of which were the pre-packaged integrating routine). The relative simplicity of the calculations in this modification means that no

³ Our required convergence precision here was 10^{-5} Ry.

significant extra time was needed to run the program. On the average, the time difference between using LDA [12] and this modified Coulomb hole routine for bulk energy gap calculations was usually small (in the order of minutes). For the possible gains that can be attained, this seems to be quite reasonable.

4. Results

The new correlation function presented here was designed to improve the fundamental gap of materials, primarily semiconductors. The constant A of Eq. (4) is the only adjustable parameter in the function, and its sole purpose is to scale the function by a multiplicative factor while not changing the overall shape. This is important because the shape is the fundamental idea behind the design of this function. It is the shape that determines how well the interelectronic correlations are modelled, not the overall scale. This new function was tested on materials whose energy gaps have been accurately measured. Five such materials were considered, including silicon, silicon carbide, sodium chloride (an insulator), gallium arsenide, and indium nitride. Table 1 shows a summary of the results including the present work, the LDA values, and the experimental values for comparisons. In most cases, the present work shows a significant improvement over the LDA.

The silicon gap is highly improved, but that is mainly because this particular gap was used as the benchmark, to help ‘fine tune’ the constant (a plot of the Si charge density is included as Fig. 2). The maximum charge density in the bond itself is slightly increased above the LDA value by 6%. In this large charge density region, screening shields the effect of the Coulomb hole, while in the interstitial region (where the charge is less dense) the charge density is more easily polarized. Thus, charge from the interstitial region is pushed into the bond. This scenario is consistent with the calculations done by QMC in Ref. [10] to describe the effect of the Coulomb hole on atoms on the left and right sides of the periodic table. This also explains how the gap correction is more effective for covalent than for ionic materials.

InN is a technologically important semiconductor that, until recently, has not been described well by ab initio theoretical methods. A new DFT–LDA calculation has been able to reproduce an experimental value for the energy gap found in an InN thin film [14]. The calculation used pseudopotentials and included an add-on quasiparticle correction. Another study [15] has also confirmed that the experimental value of the energy gap in InN is smaller than previously found. Using the present correlation method, we found that the conduction band has dropped below the top valence band, producing a negative gap. There could be

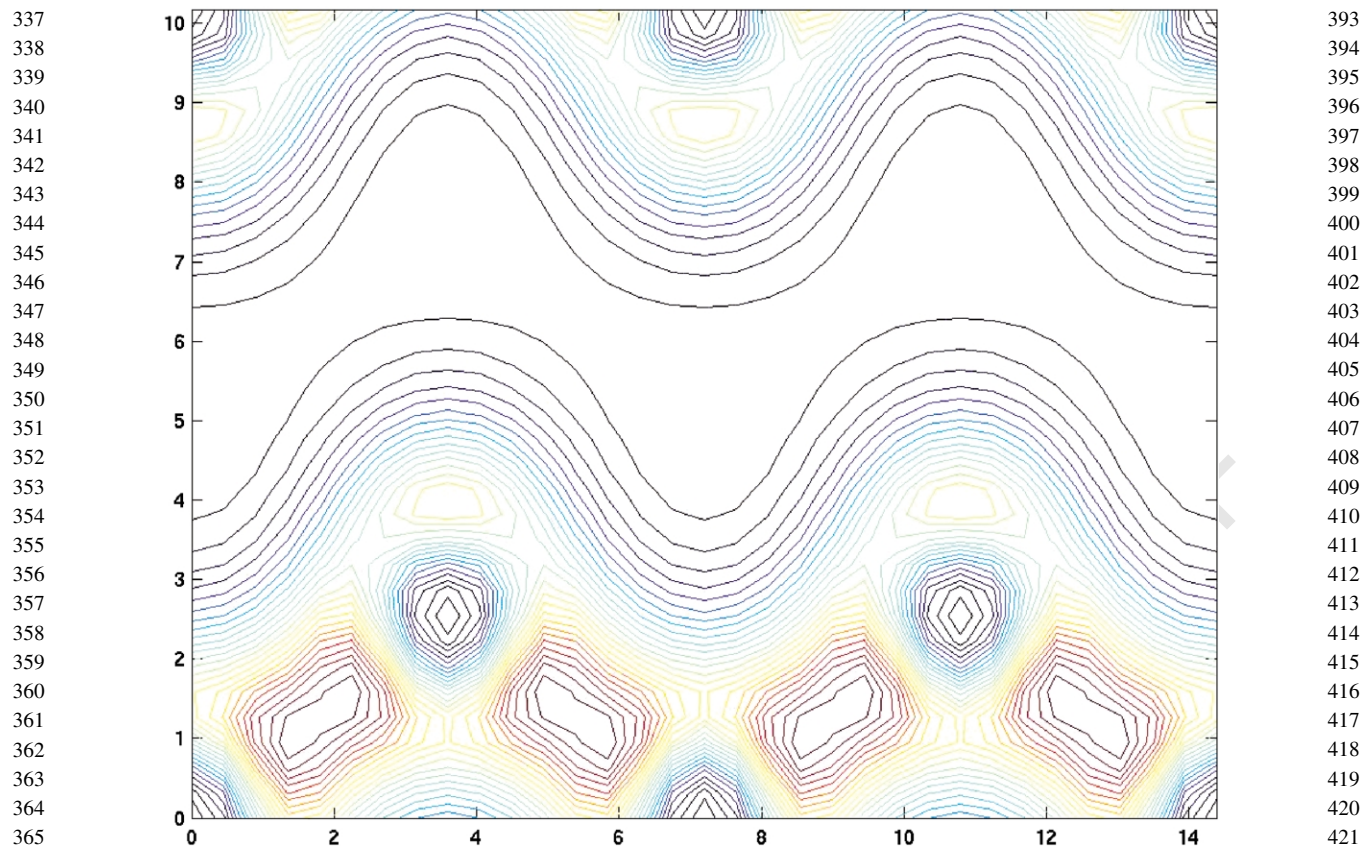


Fig. 2. A plot of Si charge density as calculated with the correlation function presented in the text. The bond charge (red) is clearly represented between the atoms (blue). The maximum charge density in the bond is higher than in LDA, which contributes to more charge localization. This effect would lead to a larger energy gap, which is what occurs.

several reasons as to why this particular compound was not well described by this correlation interaction. For example, the original Coulomb hole picture derived in Ref. [10] was done for second row atoms (i.e. Mg, Al, etc.). Since indium is not close to this set of elements, it is easy to see that it may not be encompassed by this particular interaction. In fact, it is also possible that this particular function works better for second row elements/compounds.

5. Conclusions

In summary, we present a mathematical description of a correlation function based on the physical pictures in Ref. [10]. It is a semi-local, spin-independent function implemented within the LDA of DFT. With a single adjustable parameter referenced to silicon, we found that for several semiconducting materials this approach gave more accurate results of the fundamental gap energy than the LDA by an average of 73.5% in the cases where improvement was shown, while not adding significantly to the computation time.

Acknowledgements

This work was supported in part by the Campus Laboratory Collaboration grant from the UC president's office and NSF grant No. INT-9875-23. C.Y.F. was also supported by SDSC and NERSC at LBNL.

References

- [1] P. Hohenberg, W. Kohn, Phys. Rev. B 136 (1964) 864.
- [2] W. Kohn, L.J. Sham, Phys. Rev. A 140 (1965) 1133.
- [3] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1966) 3865.
- [4] J.P. Perdew, S. Kurth, A. Zupan, et al., Phys. Rev. Lett. 82 (1999) 2544.
- [5] M. Filatov, W. Thiel, Int. J. Quant. Chem. 62 (1997) 603.
- [6] O.V. Gritsenko, A.A. Bagaturjants, V.B. Kazansky, Int. J. Quant. Chem. 29 (1986) 1799.
- [7] A.J. Williamson, R.Q. Hood, R.J. Needs, G. Rajagopal, Phys. Rev. B 57 (1998) 12140.
- [8] L. Hedin, Phys. Rev. (1965) 139.
- [9] A.C. Cancio, C.Y. Fong, J.S. Nelson, Phys. Rev. A 62 (2000) 062507.

ARTICLE IN PRESS

M.D. Watson, C.Y. Fong / Solid State Communications xx (0000) xxx–xxx

5

449	[10] A.C. Cancio, C.Y. Fong, J.S. Nelson, in: A. Gonis, N. Kioussis, M. Ciftain (Eds.), <i>Electron Correlations and Materials Properties</i> , Plenum Press, New York, 1999, p. 509.	
450		
451	[11] D.M. Ceperley, B. Alder, <i>Phys. Rev. Lett.</i> 45 (1980) 566.	
452	[12] J.P. Perdew, A. Zunger, <i>Phys. Rev. B</i> 23 (1981) 5048.	
453	[13] G.B. Bachelet, D.R. Hamann, M. Schlüter, <i>Phys. Rev. B</i> 26 (1982) 4199.	
454		
455		
456		
457		
458		
459		
460		
461		
462		
463		
464		
465		
466		
467		
468		
469		
470		
471		
472		
473		
474		
475		
476		
477		
478		
479		
480		
481		
482		
483		
484		
485		
486		
487		
488		
489		
490		
491		
492		
493		
494		
495		
496		
497		
498		
499		
500		
501		
502		
503		
504		
	[14] V.Yu. Davydov, et al., <i>Phys. Status Solidi B</i> 229 (2002) R1–R3.	505
		506
	[15] J. Wu, et al., <i>Appl. Phys. Lett.</i> 80 (2002) 3967.	507
	[16] M.L. Cohen, V. Heine, <i>Solid State Phys.</i> 24 (1970) 37.	508
	[17] W.J. Choyke, D.R. Hamilton, L. Patrick, <i>Phys. Rev. A</i> 133 (1964) 1163.	509
		510
		511
		512
		513
		514
		515
		516
		517
		518
		519
		520
		521
		522
		523
		524
		525
		526
		527
		528
		529
		530
		531
		532
		533
		534
		535
		536
		537
		538
		539
		540
		541
		542
		543
		544
		545
		546
		547
		548
		549
		550
		551
		552
		553
		554
		555
		556
		557
		558
		559
		560