Electronic surface error in the Si interstitial formation energy

Ann E. Mattsson, Ryan R. Wixom and Rickard Armiento

Post Print

N.B.: When citing this work, cite the original article.

Original Publication:
http://dx.doi.org/10.1103/PhysRevB.77.155211
Copyright: American Physical Society
http://www.aps.org/

Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-86296
Electronic surface error in the Si interstitial formation energy

Ann E. Mattsson,1,* Ryan R. Wixom,2,† and Rickard Armiento3,‡

1Multiscale Dynamic Materials Modeling MS 1322, Sandia National Laboratories, Albuquerque, New Mexico 87185-1322, USA
2Energetics Characterization MS 1445, Sandia National Laboratories, Albuquerque, New Mexico 87185-1455, USA
3Physics Institute, University of Bayreuth, D-95440 Bayreuth, Germany

(Received 7 March 2007; revised manuscript received 29 February 2008; published 25 April 2008)

The results for Si interstitial formation energies differ substantially if calculated with quantum Monte Carlo (QMC) or density functional theory (DFT) techniques. In fact, not even DFT results using different exchange-correlation functionals agree well for these energies. We carefully quantify the differences between the DFT results by accurate calculations with large supercells. A similar discrepancy for vacancy formation energies in metals has previously been resolved by introducing the concept of an “electronic surface error,” and this view is adopted and shown relevant also for the present DFT results for interstitials in semiconductors. The origin of the surface error for the Si interstitial is explained by careful examination of the electron density. A postcorrection for the surface error brings all the results obtained with the tested functionals close to the results of the AM05 functional. However, it remains an important puzzle that while the surface error correction aligns the DFT results, they are still in large disagreement with QMC results.

DOI: 10.1103/PhysRevB.77.155211 PACS number(s): 71.15.Mb, 61.72.J–, 73.90.+f

I. INTRODUCTION

The importance of silicon in our technology-based society cannot be overstated; it is at the heart of microelectronics used in everything from entertainment centers to high performance computers and spacecrafts. Knowledge about fundamental microscopic processes in this material is important for an understanding of, e.g., fabrication limitations and radiation damage. The Si interstitial formation energies are examples of important properties needed for this understanding. Experimental results for the Si self-interstitial formation energies (e.g., Refs. 1–3) are obtained through indirect methods which involve processes that are subject to interpretation.3 Experimental results are thus inconclusive, and theoretical values are of great importance.

However, current theoretical predictions of the interstitial formation energies also disagree. Previous works4,5 have pointed out a large discrepancy between quantum Monte Carlo (QMC) and density functional theory (DFT) results. Even between the DFT calculations, different exchange-correlation (XC) functionals give significantly different results. The main focus of the present paper is to better understand the differences between the DFT results.

A discrepancy similar to the one discussed here has previously been observed for metal vacancy formation energies. That issue was resolved by introducing the concept of an “electronic surface error,”7–10 and then correcting this error. The finding unified results for different functionals and brought an increased degree of predictability to defect calculations in metals. The present work brings this concept and a similar correction to the field of semiconductor interstitial formation energies. One of the XC functionals used in this work, AM05,11 is created specifically to minimize the electronic surface error, and hence it constitutes an important tool for understanding the presence of this error in the results obtained with other XC functionals.

The paper is organized as follows. In Sec. II, we give a brief background on DFT and the different XC functionals used in this work. In Sec. III, accurate DFT calculations of the Si interstitial formation energies using large supercells (216/217 atoms) are presented. In Sec. IV, we discuss the electronic surface error and how it enters interstitial formation energies. In Sec. V, we relate the foregoing discussion to our numerical results and devise a correction scheme that aligns the results from all the XC functionals. In Sec. VI, we discuss how the new unified DFT results relate to the quite different results for the formation energies obtained with the QMC method.

II. BACKGROUND: DENSITY FUNCTIONAL THEORY AND EXCHANGE-CORRELATION FUNCTIONALS

The Kohn–Sham (KS) DFT scheme, in principle, accounts for all many-body effects of the Schrödinger equation. In practice, errors are introduced by, on the one hand, implementation-related approximations and, on the other hand, the use of an approximate XC functional. However, the implementation-related approximations (e.g., basis sets, pseudopotentials, approximate matrix diagonalization methods, etc.) can all be successively improved by increasing the computational expense.12 In contrast, there is no known way, not even in principle, to systematically improve the XC functional. Hence, it is important to continuously revisit different XC approximations and carefully compare their performance. This is especially true when different theoretical methods do not agree, as in the present case for the Si interstitial formation energies.

In the present work, we calculate three Si interstitial formation energies with four different XC functionals: the local density approximation (LDA),13 which is the most straightforward but still effective XC functional; two of the most popular generalized gradient approximations (GGAs), Perdew–Burke–Ernzerhof (PBE)13 and PW91,14 and a functional recently developed by two of the present authors, AM05.11 The GGAs are constructed from the principle of fulfilling constraints on the separate exchange and correlation parts,
but where the form of the exchange part is guided by a real-space cutoff procedure. LDA and AM05 are both constructed from the principle that the local physics in a specific part of a system is described by a similar model system. LDA is based on the uniform electron gas. AM05 uses the subsystem functional scheme to further include the Airy part of a system is described by a similar model system constructed from the principle that the local physics in a specific electronic surface error to discuss the difference between the results of the different XC functionals. As the AM05 functional is created specifically to minimize this error, it is highly relevant in this context. Second, AM05 has recently been shown to give substantial improvements of the description of lattice constants and bulk moduli for solids in general and semiconductors in particular, and thus is an excellent candidate for calculating semiconductor interstitial formation energies.

III. NUMERICAL CALCULATIONS

For the calculations, we used the SOCORRO plane-wave code with the same functional dependent Si pseudopotentials (PPs) as used and specified in Ref. 11. The interstitial formation energy is calculated from \( E_{\text{form}} = E_{\text{int}} - E_{\text{bulk}}(N+1)/N \), where \( E_{\text{int}} \) and \( E_{\text{bulk}} \) are the total energies for interstitial and bulk cells, respectively, and \( N \) is the number of atoms in the bulk supercell. The calculations are performed with supercells of 216/217 atoms for the bulk/interstitial configurations. The atoms in the fixed volume interstitial supercell are geometry relaxed to their minimal energy configuration with quenched minimization using a root-mean-square force cutoff of \( 5 \times 10^{-5} \) Ry/bohr. The volume is fixed to the value derived from the functional dependent optimal lattice constant obtained in the corresponding bulk calculation.

The optimal lattice constant obtained with AM05, LDA, PBE, and PW91, are 10.264 (5.431), 10.174 (5.384), 10.333 (5.468), and 10.318 (5.460) bohr (Å), respectively. The experimental lattice constant is 5.43 Å (Ref. 22) and our numbers agree well with previously published results. Notice the excellent agreement between the AM05 result and the experimental value. The optimal lattice constant was obtained by fitting seven energy vs volume points to the Murnaghan equation of state. The wave-function cutoff was 20.0 Ry in all calculations and a Monkhorst-Pack \( 24 \times 2 \times 2 \) k-point sampling with \( 2 \times 2 \times 2 \) points was used. We used a Fermi smearing of \( 7.3 \times 10^{-3} \) Ry.

Interstitial formation energies for Si have been calculated for three different interstitial types, as presented in Table I. These results cannot be directly compared with the DFT results given in Ref. 5 due to technical differences (size of supercells, level of relaxation, and choice of basis sets). Thus, we also cannot compare to another, more expensive and generally not available in solid state codes, “pure” functional tested in that work: the metaGGA functional by Tao, Perdew, Staroverov, and Scuseria (TPSS). However, Fig. 2 of Ref. 5 shows the results of TPSS and PBE as being comparable. Large supercell PW91 results for the interstitials are given in Ref. 26. These results are obtained by also optimiz-

<table>
<thead>
<tr>
<th>System</th>
<th>AM05</th>
<th>LDA</th>
<th>PBE</th>
<th>PW91</th>
<th>QMCa</th>
<th>QMCb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetrahedral</td>
<td>3.399</td>
<td>3.562</td>
<td>3.908</td>
<td>4.091</td>
<td>5.40</td>
<td>5.05</td>
</tr>
<tr>
<td>Hexagonal</td>
<td>3.253</td>
<td>3.424</td>
<td>3.617</td>
<td>3.768</td>
<td>4.82</td>
<td>5.13</td>
</tr>
</tbody>
</table>

aReference 4.
bReference 5.

IV. ELECTRONIC SURFACES IN INTERSTITIAL SYSTEMS

In Ref. 17, an electronic surface is defined as the surface traced out by the crossing points of the chemical potential and the effective potential. The physics at such a surface is distinctively different from the physics of a uniform electron gas. While the low density regions in a crystal are not typically exhibiting explicit electronic surfaces, the effective potential is still higher than the KS orbital energy of a significant number of occupied orbitals, and the electron density strongly decays. Commonly used XC functionals make a larger error in these implicit surface regions than in the rest of the system, since they do not properly handle the different electron physics present in these regions.

The electronic surface error is important for formation energy calculations when two structures with different amounts of electronic surface are compared because in that case the error made in the removed or added surface regions will not be canceled out. The findings of Refs. 8–10 for metal vacancies will be summarized in the following. In a metal, the electron density is fairly homogeneous. When an atom is removed to create a vacancy, a hole is created in the electron density. By estimating the surface area of this hole and matching the density profile with theoretical surfaces on which the errors of the different functionals are easy to calculate, the error can be estimated and removed by a postcorrection. This discovery led to a reevaluation of calculated vacancy formation energies in metals, and we will discuss the need to reevaluate interstitial formation energies the same way in the following. The rest of Sec. IV will discuss why and how we expect the electronic surface error to be present for these systems, and then in Sec. V, we will motivate this interpretation from our numerical results.

The basic idea of how the electron surface error enters Si interstitial calculations is that a normal Si crystal contains large “holes” in its charge density, as shown in Fig. 1. This is a common feature of semiconductors. If an interstitial Si atom is inserted into one of these low density sites, the added electron density will fill the hole. This means that the density around the new atom becomes more homogeneous, and, as a result, some surface is removed. The calculation of the for-
V. CORRECTING THE NUMERICAL RESULTS
FOR THE ELECTRONIC SURFACE ERROR

In Sec. IV, we explained how the electronic surface error is expected to enter interstitial formation energies. In Sec. V, we will connect this interpretation to the calculated numerical results reported in Sec. III. A first numerical indication of the presence of electronic surface errors is found in the PBE and PW91 results in Table I. While PBE and PW91 often perform similarly, their respective electronic surface errors are different. Hence, the fact that these functionals give almost identical lattice constants for the Si crystals (10.33\, bohr, respectively), but give noticeable different formation energies, strongly suggests that the results are influenced by the electronic surface error.

When the results of all XC functionals in Table I are considered, there is a clear trend in the formation energies throughout all the interstitials. The AM05 functional gives the lowest energy, while LDA, PBE, and PW91 give successively larger energies. Reference 10 presents the electronic surface error for LDA, PBE, and PW91 derived from the surface energy of jellium;\(^{18,19}\) LDA has a relatively small surface error, while PBE’s error is larger, and PW91’s is the largest. Furthermore, the reported differences are of similar relative proportions as found in the numerical results of this work. In addition to this, the AM05 functional is specifically designed to have a minimal electronic surface error.\(^{11}\) Hence, the numerical trend in Table I perfectly matches the theoretical trend of a system where the dominant error is due to the electronic surface error. In the following, we will formalize this reasoning into an explicit correction scheme.

Following the lead of the vacancy corrections of Refs. 8–10, we expect the basic principle of an interstitial energy correction to be

\[
E_{\text{form}}^{\text{Corrected}} = E_{\text{form}}^{\text{DFA}} - A \Delta\sigma_{\text{XC}}^{\text{DFA}} - \Delta\sigma_{\text{XC}}^{\text{DFA}}.
\]

where DFA denotes “density functional approximation” and is either LDA, PBE, or PW91, \(\Delta\sigma_{\text{XC}}^{\text{DFA}}\) is the electronic surface error of the XC functional per area, and \(A\) is the electronic surface area that was removed by the interstitial (since the different functionals give almost identical densities, we can safely assume that \(A\) is the same for all functionals). The correction term has an opposite sign compared to what has been used for vacancies, since, here, electronic surface area is removed rather than added. The problem at this point, compared to previous postcorrection schemes for the surface error,\(^{8–10}\) is that we have no estimate of the surface area \(A\),...
FIG. 3. (Color) Investigating the bonds of the (110)-split interstitial. (a) The density on a sphere with a radius of half the perfect crystal interatomic distance around an (110)-split interstitial atom (upper blue in Fig. 2). This interstitial atom is four coordinated but only has one strong, high density, bond of the type seen for the four-coordinated atoms in a bulk lattice shown in (b). (c) The density around a five-coordinated atom in the (110)-split interstitial (yellow in Fig. 2) shows three bulk type bonds and two weak, smeared out, bonds to the two interstitial atoms (blue). \( \theta \) is the angle from the positive \( z \) axes, while \( \phi \) is the angle in the \( x-y \) plane measured counterclockwise from the \( x \) axes. The density shown is calculated with AM05, but does not differ substantially from densities calculated with other functionals. The colored dots show the direction to the equally colored atoms in Fig. 2.

nor do we know if the dependency of the surface correction \( \Delta \sigma_{\text{XC}}^{\text{DFA}} \) on the jellium-surface-related “bulk density” parameter is similar to what was found for vacancies.9,10

However, there is an interesting general relation between the surface corrections of different XC functionals that can be observed in Fig. 4 of Ref. 10. While the surface correction per area depends on a bulk density parameter, the relative proportions between the different corrections remain remarkably consistent as 0.29:0.76:1 for LDA:PBE:PW91 over the range of relevant densities. Hence, the individual corrections for different XC functionals can all be expressed in the correction of only one of the functionals,

\[
\begin{align*}
\Delta \sigma_{\text{XC}}^{\text{LDA}} &= 0.29 \Delta \sigma_{\text{XC}}^{\text{PW91}}, \\
\Delta \sigma_{\text{XC}}^{\text{PBE}} &= 0.76 \Delta \sigma_{\text{XC}}^{\text{PW91}}.
\end{align*}
\]

The discussion in the beginning of Sec. V suggests that the electronic surface error is the dominant error. We thus assume that the corrected interstitial formation energies \( E_{\text{form}}^{\text{corrected}} \) from different functionals are approximately equal. Below, we will be able to \textit{ex facto} reexamine this assumption.

For each interstitial, there are now three equations:

\[
\begin{align*}
E_{\text{form}}^{\text{corrected}} &= E_{\text{form}}^{\text{LDA}} - 0.29 \Delta \sigma_{\text{XC}}^{\text{PW91}}, \\
E_{\text{form}}^{\text{corrected}} &= E_{\text{form}}^{\text{PBE}} - 0.76 \Delta \sigma_{\text{XC}}^{\text{PW91}}, \\
E_{\text{form}}^{\text{corrected}} &= E_{\text{form}}^{\text{PW91}} - \Lambda \Delta \sigma_{\text{XC}}^{\text{PW91}}.
\end{align*}
\]

This is an overdetermined set of equations for the two unknowns, \( E_{\text{form}}^{\text{corrected}} \) and \( \Lambda \Delta \sigma_{\text{XC}}^{\text{PW91}} \) (i.e., the area multiplied with the error correction for PW91 is regarded as one combined unknown). The set of equations can be solved by finding the least-squares solution.

To summarize, for each interstitial, the above procedure removes as much energy from each calculated result, \( E_{\text{form}}^{\text{DFA}} \),
as can be motivated from rescaling the energy differences obtained with different functionals for a pure surface model (jellium surfaces\textsuperscript{18,19}). The outcome is a “fitted energy,” $E_{\text{form}}^{\text{corrected}}$, which represents the common formation energy value to which the correction scheme aligns the different XC functional results. However, individual corrected values for each functional are also available by using the fitted value for $\Delta E_{\text{xc}}^{\text{PW91}}$ in Eqs. (3)–(5). How well these individual values agree is an indication of how well the differences between LDA, PBE, and PW91 fall into the predicted 0.29:0.76:1 proportions, and thus, how good the assumption is that the surface error is the dominating difference between the results obtained with different functionals for the same system.

In Table II, we present the results of applying this correction scheme to the interstitial formation energies calculated with the LDA, PBE, and PW91 functionals. The corrected results of the different XC functionals agree well, which validate our assumption that the surface error is the dominant error. In addition, the results obtained with AM05 functional, which has not been used in this scheme, show very good agreement with the aligned results of the other functionals. AM05 is constructed to minimize its surface error in such a way that its correction would be zero.

The correction scheme described above is general and can be used for any kind of energy calculation where the electronic surface error is the dominant error. To obtain surface error corrected results with this scheme, only the results obtained with LDA, PBE, and PW91 for the same system are needed. (One could, in fact, calculate a correction using only the results of two XC functionals, but that would sacrifice the consistency check from the overdetermined system of equations.) The general applicability of this scheme will be discussed in future publications.

### VI. CONCLUSIONS

We have shown that the differences in Si interstitial formation energies calculated with different DFT XC functionals can be explained by using the concept of an electronic surface error. By correcting the results for this error, they align to give a unified DFT picture of the formation energies for tetrahedral (3.5 eV), hexagonal (3.28 eV), and (110)-split (3.23 eV) interstitials, with a spread between the results of the different XC functionals of the order of 0.05 eV. However, the corrected and unified DFT results disagree substantially with published QMC results.

The disagreement between DFT and QMC for the Si interstitial formation energy has been thoroughly discussed in previous works.\textsuperscript{43} Since QMC accounts for the full XC energy, the QMC results are generally regarded as more accurate than DFT results. Thus, the accepted view is that the approximate DFT XC functionals make an unusually large error for the Si interstitial formation energies. Further motivation for this picture has been drawn from agreement with indirect experimental data,\textsuperscript{1,2} even though these data are not fully conclusive.\textsuperscript{3} If one adopts this picture, our results point to some properties in the Si interstitial system that all the tested semilocal XC functionals are equally unsuccessful in treating. This source of error, however, must be fundamentally different from the source of the electronic surface error discussed in this paper. This interpretation challenges DFT XC functional developers to find, explain, and handle this unknown source of error.

However, the shared XC functional error in the above interpretation has unusual properties. Our corrected XC functional results suggest that the error has a magnitude of about 1.5 eV for the Si interstitial formation energies, yet the error is shared equally between all the tested functionals within just $\pm 0.05$ eV. Furthermore, it is much larger for the Si interstitial formation energy than, e.g., for other defect energies in bulk Si (Ref. 28) and at Si surfaces.\textsuperscript{29} For these calculations, highly precise experiments are available, the results of which agree to within a few tenths of an eV. Of course, these facts are not conclusive evidence that the DFT Si formation energies should be equally accurate, but motivate a discussion that does not uncritically assume the QMC results to be the “exact” benchmarks to which we should strive to align results for DFT XC functionals.

TABLE II. Uncorrected and corrected Si interstitial formation energies in eV obtained with the LDA (Refs. 6 and 25), PBE (Ref. 13), and PW91 (Ref. 14) functionals as compared to $E_{\text{form}}^{\text{corrected}}$ and the AM05 results.

<table>
<thead>
<tr>
<th></th>
<th>Tetrahedral</th>
<th>Hexagonal</th>
<th>(110) split</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calculated</td>
<td>Corrected</td>
<td>Calculated</td>
</tr>
<tr>
<td>LDA</td>
<td>3.56</td>
<td>3.35</td>
<td>3.42</td>
</tr>
<tr>
<td>PBE</td>
<td>3.91</td>
<td>3.34</td>
<td>3.62</td>
</tr>
<tr>
<td>PW91</td>
<td>4.09</td>
<td>3.35</td>
<td>3.77</td>
</tr>
<tr>
<td>$E_{\text{form}}^{\text{corrected}}$</td>
<td>3.35</td>
<td>3.28</td>
<td></td>
</tr>
<tr>
<td>AM05 (uncorrected)</td>
<td>3.40</td>
<td>3.25</td>
<td>3.16</td>
</tr>
</tbody>
</table>
Can the discrepancy be explained by known technical differences? The QMC calculations are performed on LDA or PW91 structures with LDA or Hartree–Fock PPs; the atoms in the QMC calculations are not relaxed to their minimum energy configuration and the resulting energies should be expected to be larger than the energies obtained in our relaxed calculations. In order to estimate what effect the relaxation and the choice of PPs have on the energies, we have performed an unrelaxed PBE calculation on the (110)-split LDA structure and using the LDA PPs. The resulting formation energy is 3.774 eV, which is 0.228 eV above the relaxed result obtained with the appropriate PPs presented in Table I. Hence, this still does not explain the full difference of more than 1.5 eV between DFT and QMC results.

In the present paper, we have discussed how different errors for bulk and electronic surface regions obtained when using traditional DFT XC functionals influence the accuracy of calculated Si interstitial formation energies. Could there possibly be some surface related problem also hiding in the QMC interstitial formation energy results? Recent articles explain the inaccuracies of some QMC calculations made for jellium surfaces\(^\text{30,31}\) by suggesting that the fixed node approximation makes a larger error for jellium slab systems than for bulk systems. This explanation suggests a similarity with the situation of DFT for Si interstitials discussed here, since our electronic surface error correction is based on relating the electron physics in relevant regions of a Si crystal to the electron physics at jellium surfaces. While this suggested source of error in the QMC calculations for the Si interstitial formation energies is highly speculative, we hope that future work in the QMC field will address this issue and investigate if it is relevant.

It is not the goal of the present paper to conclusively explain the difference between DFT and QMC results, or evaluate if DFT or QMC is giving the best results for the formation energies of Si interstitials. Our primary contribution is instead a very careful quantification of the differences between the DFT results obtained with different XC functionals for these formation energies, and a discussion of how the concept of the electronic surface error explains these differences. However, our electronic surface error corrected results and this concluding discussion suggest that the theoretical value of the Si interstitial formation energy should not be regarded as a completely settled matter yet. Further studies that use independent, improved, or experimental methods will be highly relevant.

\section*{Acknowledgments}

We thank T. R. Mattsson for fruitful discussions, P. A. Schultz for comments on the paper, and N. A. Modine for initial structures. R.A. acknowledges support from the Alexander von Humboldt Foundation. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy’s National Nuclear Security Administration under Contract DE-AC04-94AL85000.

\begin{thebibliography}{26}
\bibitem{21} SOCORRO is developed at Sandia National Laboratories and available from (http://dft.sandia.gov/Socorro/).
\bibitem{26} S. A. Centoni, B. Sadigh, G. H. Gilmer, T. J. Lenosky, T. Díaz de
\end{thebibliography}