PHYSICAL REVIEW B **74**, 121102(R) (2006)

Comparison of screened hybrid density functional theory to diffusion Monte Carlo in calculations of total energies of silicon phases and defects

Enrique R. Batista,[1,*] Jochen Heyd,[2] Richard G. Hennig,[3] Blas P. Uberuaga,[4] Richard L. Martin,[1] Gustavo E. Scuseria,[2] C. J. Umrigar,[5] and John W. Wilkins[6]
[1]*Theoretical Division, Los Alamos National Laboratory, MS-B268, Los Alamos, New Mexico 87545, USA*
[2]*Department of Chemistry, Rice University, Houston, Texas 77005-1892, USA*
[3]*Department of Physics, Ohio State University, Columbus, Ohio 43210, USA*
[4]*Materials Science and Technology Division, Los Alamos National Laboratory, MS-G755, Los Alamos, New Mexico 87545, USA*
[5]*Cornell Theory Center and Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853, USA*
[6]*Department of Physics, Ohio State University, Columbus, Ohio 43210, USA*

(Received 19 July 2006; published 26 September 2006)

Nearly quantitative agreement between density functional theory (DFT) and diffusion Monte Carlo (DMC) calculations is shown for the prediction of defect properties using the Heyd-Scuseria-Ernzerhof (HSE) screened-exchange hybrid functional. The HSE functional enables accurate computations on complex systems, such as defects, where traditional DFT may be inadequate and DMC calculation computationally unfeasible. The screened-exchange hybrid functional retains the benefits of earlier hybrid functionals in terms of treating strongly correlated insulators but unlike them it can be applied to metallic phases. This study concentrates on the DFT energetic predictions of point defects in silicon and on phase energy differences between the diamond and metallic β-tin phases.

DOI: 10.1103/PhysRevB.74.121102 PACS number(s): 71.15.Mb, 31.15.Ew, 61.72.Bb, 71.10.−w

The prediction of energy landscapes from density functional theory (DFT) has often lacked the predictive power necessary to quantitatively anticipate experimental results such as phase diagrams and properties of point defects. This predictive capability is most needed when experiments are unable to measure desired properties directly, for example indirect probes of the existence and behavior of point defects has led to disagreements regarding the individual contributions of interstitial and vacancy defects to self-diffusion.[1,2] The first two generations of DFT functionals, the local density approximation (LDA) and the generalized gradient approximations (GGA),[3–5] yield structural properties in reasonable agreement with experiments, but the energetics of defects have been known to be poorly reproduced. For example, for both Si and Ge, two group-IV semiconductors, the LDA and GGA predict self-diffusion activation energies that are about 1.5 and 1 eV lower, respectively, than experimental measurements.[6] This is also true for the diffusion of dopants in these materials.[6] For nearly every case considered for these materials, the predicted activation energy is roughly 1 eV lower than that found by experiment. In addition, DFT with the GGA predicts the semiconductor Ge to be metallic. Without quantitative predictions of point defect properties, the modeling of larger-scale phenomena, such as dopant profiles in Si or microstructure in an irradiated material, will be inadequate.

Higher computational accuracy can be achieved with the diffusion Monte Carlo (DMC) approach, although at a greater computational expense. DMC calculations have been performed on key structures, and past work[7] has shown that DMC calculations on interstitial defects in Si find formation energies that are significantly higher than those from DFT with the GGA and thus closer to experimental estimates. A computational method that combines the relative ease of DFT computations with the accuracy of DMC simulation is certainly needed in order to obtain quantitative predictive power of what would be very difficult experiments. In this vein, third- and fourth-generation density functionals have recently been developed but as yet remain untested in this energetic context. In this work two of the latest generation of DFT functionals are compared to highly converged DMC calculations of point defects in solid silicon, and the relative energy between two phases of silicon, the diamond phase and β-tin, which is metallic. The functionals to be examined are the "meta-GGAs," which extend the expansion of the exchange-correlation energy[8] to include the kinetic energy density, and a "hybrid" functional, which contains a portion of the full, nonlocal, Hartree-Fock-type exchange. We find a steady improvement in the quality of the results as we progress along the rungs of Perdew's "Jacob's ladder" to DFT heaven.[9] In fact, the Heyd-Scuseria-Ernzerhof hybrid functional[10] is found to give results in excellent agreement with our DMC benchmark. In addition, the HSE functional has the advantage over previous hybrid functionals of being applicable for systems with a metallic band structure.

As the focus of this paper is an evaluation of functionals, as opposed to direct comparison with experiment, diffusion Monte Carlo calculations provide our standard of accuracy. The DMC simulation is a many-body method based on correlated wave functions and stochastic computational techniques that solves the full many-body Schrödinger equation to high accuracy. It involves only two essential approximations, namely, the fixed-node approximation and, for heavy atoms, the pseudopotential approximation. It typically recovers over 95% of the correlation energy, provided that well-optimized wave functions are employed. Calculations for periodic systems also include a finite-size extrapolation.

The new functionals examined are the meta-GGA of Tao, Perdew, Scuseria, and Staroverov[9] (TPSS) and the screened-exchange hybrid functional of Heyd, Scuseria, and

©2006 The American Physical Society

Ernzerhof.[10] For completeness the results for the LDA and GGA [Perdew and Wang 1991 (Refs. 3 and 4) (PW91) and Perdew-Burke-Ernzerhof (Ref. 5) (PBE) functionals] are also reported. The TPSS functional augments the GGA with kinetic-energy-density-dependent terms. The HSE functional mixes in a fraction (25%) of exact screened Hartree-Fock exchange as done in the PBE hybrid[11] (PBEh) functional.[23] The key difference in the HSE, relative to the PBEh functional, is that the exact exchange term is screened at large distances, effectively giving a hybrid functional only at short range and GGA behavior at long range. The optimal screening length of 3.5 Å was chosen to best reproduce molecular formation enthalpies on a training set of molecules.[10] As a result the HSE functional delivers the same accuracy in molecular energies already reported for hybrid functionals. Somewhat surprisingly, we have recently demonstrated that it provides band gaps in good agreement with experiment for a benchmark set of 40 semiconductors.[12] In this regard the HSE functional is similar in spirit to the screened exchange methodology of Bylander and Kleinman.[13] Most importantly, since the radius of convergence of the exact exchange interaction scales inversely proportionally to the band gap of the solid, previous hybrid functionals are more computationally demanding than the HSE and are even unusable for metallic systems. This problem is eliminated with the screening introduced in the HSE functional, allowing us to simulate not only the semiconductor phases of silicon but also the β-tin phase, which is metallic, and point defects, which introduce energy levels in the gap region.

Identical atomic configurations are used for all the methodologies explored. For the volume of the diamond phase we use 20 Å^3 per atom and for the β-tin phase we chose a value of 15 Å^3 per atom and a c/a ratio of 0.55 close to the experimental values. The structure of the point defects is calculated using the PW91 functional,[3,4] which in periodic systems usually results in geometries close to those found in experiment. These calculations are performed with VASP,[14] using an ultrasoft Vanderbilt-type pseudopotential[15,16] with a cutoff energy of 250 eV and an $8\times8\times8$ k-point mesh. The HSE calculations are carried out using the development version of the GAUSSIAN suite of programs[17] with a modified 6-311G(2d) basis set with a sampling of 800 k points. To avoid redundancies in the description of the electron density of the solid state, the smallest Gaussian exponent is set to 0.12 a.u.

The DMC calculations are performed using the CHAMP code developed by Umrigar and Filippi. The $1s$, $2s$, and $2p$ electrons of Si are eliminated using a Hartree-Fock pseudopotential.[18] A Slater-Jastrow-type wave function is used as the trial wave function. The orbitals of the Slater determinant are taken from a DFT calculation with the CPW2000 code of Martins using the PBE functional. The parameters of the Jastrow function describing electron-electron and electron-nuclear correlations are optimized using a recently developed energy minimization method.[19]

The relative energy of the Si(I) (diamond) and Si(II) (β-tin) phases assesses the accuracy of our functional for predicting bulk energies. The diamond phase is the ground state of silicon; it is a semiconductor with a band gap of 1.12 eV. At a pressure of about 10 GPa the diamond phase transforms into the metallic β-tin phase. The atomic arrangements of these two phases are shown in Fig. 1.

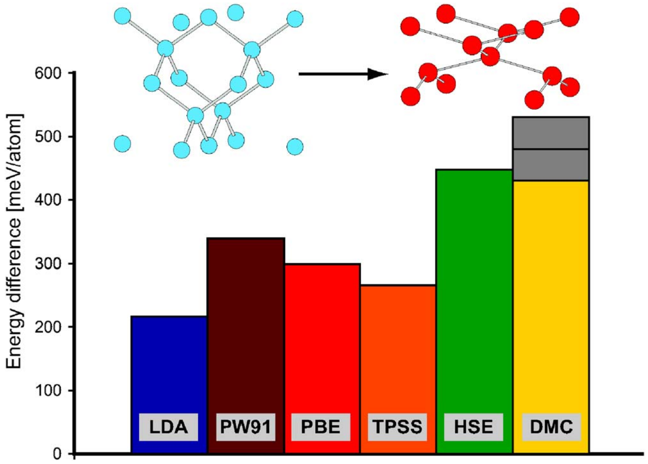

FIG. 1. (Color online) Difference in energy per atom in the diamond phase and in the β-tin phase of Si. The DMC energy of 480 ± 50 meV/atom benchmarks the accuracy of the different DFT functionals. The gray (top) region of the DMC bar represents the uncertainty of the DMC energy. The HSE functional agrees with the DMC results while the other functionals underestimate the energy difference by 140–250 meV/atom.

We calculate the energy difference between the diamond and the β-tin phases of Si using the DMC method. For these crystal phases the finite-size error is significant. Extrapolating the energy from calculations for simulation cells of 16–250 atoms results in an energy difference of 480 ± 50 meV. The total error of the energy difference between the diamond and β-tin phases includes a finite-size error of about 30 meV/atom and a combined statistical, time-step and population control error of about 10 meV/atom. We estimate the fixed-node and pseudopotential error to be of the order of 10 meV/atom, resulting in the total error of about 50 meV/atom. Earlier DMC calculations[18] predicted an energy difference between the diamond and β-tin phase of 505 ± 10 meV/atom. Our error estimate of 50 meV/atom is larger than their estimate because we also include estimates of systematic errors in addition to the statistical error. Our slightly lower energy for the β-tin phase might be related to the use of the improved energy optimization method for the trial wave function.[19]

The energies found with the LDA, GGA, and HSE functionals of density functional theory are compared relative to the DMC value of 480 ± 50 meV/atom. The pure density functionals underestimate the energy difference by 55% for the LDA (216 meV/atom), 29% for the GGA PW91 (339 meV/atom), 38% for the GGA PBE (299 meV/atom), and 45% for the meta-GGA TPSS (266 meV/atom). We choose to report the calculations with both these GGA parametrizations because they do not predict the same energy difference between the two phases; in fact they differ from each other by 12%. The hybrid HSE functional is within the error bar of the DMC calculation with a value of 447 meV/atom. While the parameters of the hybrid functional that control the fraction of exact exchange interaction and the screening function are adjusted to reproduce molecu-

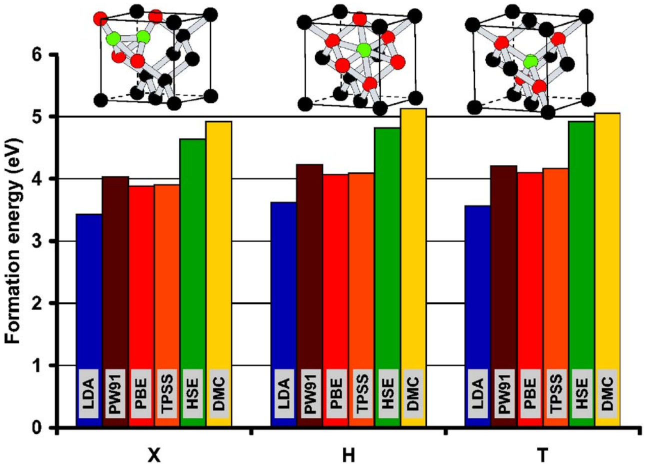

FIG. 2. (Color online) Formation energy of the three lowest-energy single interstitials in silicon (X, H, and T). Diffusion Monte Carlo calculations provide benchmarks for the accuracy of current density functionals based on the local density approximation, the generalized gradient (PW91), the kinetic energy density (TPSS), and hybrid (HSE) techniques. The DFT formation energies vary by as much as 2 eV between different approximations of the exchange-correlation functional. Comparison to the DMC results demonstrate a steady improvement of the accuracy of the functionals as the order of density expansion increases, with quantitative agreement for the hybrid HSE functional.

lar bond energies, this result shows the excellent applicability of this functional to the solid phase.

The behavior of the pure density functionals (LDA, PW91, and TPSS) can be understood by the slowly varying density expansion.[8] The metallic β-tin phase exhibits a more homogeneous electron density than the covalently bound semiconducting diamond phase. The LDA does not account for the effect of density gradients on the exchange and correlation energy and has a tendency to favor more homogeneous systems. Hence it describes the β-tin better, being a fairly homogeneous charge distribution, but diamond not so well, since it is not so homogeneous. Since β-tin lies above diamond in energy, the difference is smaller than it should be. The GGA PW91 functional[3,4] has the correct slowly varying density expansion to second order in the density gradients and the meta-GGA to fourth order.[9] In contrast to the LDA, GGA PBE, and meta-GGA TPSS, the GGA PW91 functional does not correctly reproduce the linear response of the uniform electron gas. Both the GGA and meta-GGA improve the description of the relative energy of the Si phases over the LDA with the meta-GGA TPSS falling between the LDA and GGA results. In contrast, the HSE screened hybrid functional agrees very closely with the DMC results. The origin of this is likely the improved description of covalent bonds by the HSE functional which results in a more accurate energy for the covalently bound diamond phase.

After establishing the accuracy of the HSE functional for crystalline solids, we examine the different density functionals on interstitial defect configurations in the diamond phase. Figure 2 shows the structure of the three lowest-energy interstitial defects in Si: X (split), H (hexagonal site), and T (tetrahedral site). The interstitial defect structures are obtained by relaxing the atomic configurations using the GGA (PW91) functional at fixed volume with a lattice constant of 5.432 Å. The defects are placed in a 16-atom simulation cell, which results in energies converged to about 0.1 eV.[7] Calculations for larger cells with up to 64 atoms show that the DMC finite-size error is small for these defects.

Figure 2 compares the DMC and DFT energies of the X, H, and T interstitial defects.[24] The discrepancies between DFT calculations at the LDA and GGA levels and experimental data and DMC results are already well documented in the literature. The 16-atom cell formation energies for X, H, and T defects predicted by the DMC method are 4.94, 5.13, and 5.05 eV, respectively, with a statistical error of 0.05 eV. Relative to the DMC result, the LDA underestimates the formation energy of the defects by about 1.5 eV, the GGAs by 0.9 eV (PW91) and 1.0 eV (PBE), and the meta-GGA (TPSS) by roughly 0.95 eV. The HSE hybrid functional improves the agreement dramatically, with an average discrepancy of 0.25 eV.[25] In spite of the differences in the predicted formation energies, all the DFT functionals and the DMC method predict that the X interstitial defect has the lowest formation energy. The energy difference between the X and H defects is insensitive to the choice of the functional and is about 0.2 eV. The H and T defects are nearly degenerate for all the functionals and their energy difference is within the error of the calculations. The formation energy of the T defect is more sensitive than that of the H defect to the lattice constant and the supercell size.

The HSE functional yields interstitial formation energies that agree significantly better with both DMC and experimental results than do the energies obtained from the LDA, GGA, and meta-GGA. The experimental overall diffusion activation energy (defect formation + migration barrier) for defects in silicon is 4.75 eV.[1,2] This includes contributions from interstitials and vacancies. Upon deconvolution, different experimental groups estimate diffusion activation energy for interstitials of $4.68^{+0.12}_{-0.15}$ (Ref. 2) and 4.95 eV.[1] Assuming an energy barrier of 0.1–0.3 eV,[20] the results of the HSE functional fall within the range of experimental values and the DMC method slightly overestimates the formation energy. Part of this overestimation could be attributed to using the GGA geometry. In spite of this, both methods provide a much better agreement with experiment than the predictions from the pure density functionals.

The results presented above show that, for a set of representative problems in silicon, the screened-exchange hybrid DFT HSE functional significantly improves the agreement between DFT and DMC calculations over the highly used LDA and GGA functionals. Compared to other hybrid functionals, such as the B3LYP or PBEh functional, the HSE functional has the advantage of being applicable to metallic phases, such as silicon β-tin, and the point defects. While these conclusions must be verified more thoroughly in subsequent research, it appears that this opens up the possibility of pursuing more complex calculations involving larger simulation cells. Problems within the range of semiquantitative understanding may include activation barriers of diffusion, interactions among point defects, and the properties of point defect clusters.[21] It also opens the possibility of studies of other systems where pure DFT methods have failed to reproduce the band structure, such as Ge and UO_2, which the LDA and GGA predict to be metallic, but where the HSE functional accurately describes the gap.[22] Our results on defects are especially noteworthy as radiation damage and the resultant point defects play an important role in the properties of these technologically important materials.

We thank J.-L. Martins for the use of his CPW2000 plane-wave code. The work at Rice University was supported by DOE Grant No. DE-FG02-01ER15232 and The Welch Foundation. This work was also supported by DOE (Grant No. DE-FG02-99ER45795) and NSF (Grants No. EAR-0530301, and No. DMR-0205328) and Sandia National Laboratory. Computational resources were provided by OSC, NERSC, and NCSA. The work at LANL was supported by the LDRD program. B.P.U. acknowledges funding from the U.S. Department of Energy, Office of Basic Energy Science.

---

*Electronic address: erb@lanl.gov

[1] H. Bracht, E. E. Haller, and R. Clark-Phelps, Phys. Rev. Lett. **81**, 393 (1998).

[2] A. Ural, P. B. Griffin, and J. D. Plummer, Phys. Rev. Lett. **83**, 3454 (1999).

[3] J. P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie-Verlag, Berlin, 1991), p. 11.

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

[5] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); J. P. Perdew, K. Burke, and M. Ernzerhof, *ibid.* **78**, 1396(E) (1997).

[6] B. P. Uberuaga, Ph.D. thesis, University of Washington, Seattle, WA, USA, 2000.

[7] W. K. Leung, R. J. Needs, G. Rajagopal, S. Itoh, and S. Ihara, Phys. Rev. Lett. **83**, 2351 (1999).

[8] P. S. Svendsen and U. von Barth, Phys. Rev. B **54**, 17402 (1996).

[9] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

[10] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003); J. Heyd and G. E. Scuseria, *ibid.* **120**, 7274 (2004); J. Heyd and G. E. Scuseria, *ibid.* **121**, 1187 (2004).

[11] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **110**, 5029 (1999).

[12] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, J. Chem. Phys. **123**, 174101 (2005).

[13] D. M. Bylander and L. Kleinman, Phys. Rev. B **41**, 7868 (1990).

[14] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993); **49**, 14251 (1994); G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 16 (1996); Phys. Rev. B **54**, 11169 (1996).

[15] D. Vanderbilt, Phys. Rev. B **41**, R7892 (1990).

[16] G. Kresse and J. Hafner, J. Phys.: Condens. Matter **6**, 8245 (1994).

[17] M. J. Frisch *et al.*, GAUSSIAN, Development Version, Revision C.01 (Gaussian, Inc., Wallingford, CT, 2004).

[18] D. Alfe, M. J. Gillan, M. D. Towler, and R. J. Needs, Phys. Rev. B **70**, 214102 (2004).

[19] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).

[20] R. J. Needs, J. Phys.: Condens. Matter **11**, 10437 (1999).

[21] D. A. Richie, J. Kim, S. A. Barr, K. R. A. Hazzard, R. Hennig, and J. W. Wilkins, Phys. Rev. Lett. **92**, 045501 (2004).

[22] K. N. Kudin, G. E. Scuseria, and R. L. Martin, Phys. Rev. Lett. **89**, 266402 (2002).

[23] Also known as PBE0 and PBE1PBE in the literature.

[24] The defect formation energy was defined as the energy cost for a bulk atom to turn into an intestitial. This energy corresponds to the relative energy between the configuration with the defect and a perfect bulk with an extra atom at a kink site.

[25] The computed formation energies for the *X*, *H*, and *T* interstitial defects are, respectively, for the LDA 3.43, 3.62, and 3.56 eV, for the GGA PW91 4.03, 4.23, and 4.21 eV, for the GGA PBE 3.88, 4.06, and 4.09 eV, for meta-GGA (TPSS) 3.90, 4.09, and 4.16 eV, and for the screened hybrid HSE functional 4.64, 4.82, and 4.92 eV.