Electronic structure analysis of self-consistent embedding theory for quantum/molecular mechanics simulations

Xu Zhang,[1,2] Chong-Yu Wang,[1,3] and Gang Lu[2,*]
[1]*Department of Physics, Tsinghua University, Beijing 100084, China*
[2]*Department of Physics and Astronomy, California State University—Northridge, Northridge, California 91330-8268, USA*
[3]*The International Centre for Materials Physics, Chinese Academy of Sciences, Shenyang 110016, China*
(Received 7 September 2008; revised manuscript received 15 November 2008; published 16 December 2008)

The self-consistent embedding theory provides a rigorous framework for quantum/molecular mechanics simulations of materials. By using crystalline aluminum as an example, we present a critical analysis on the accuracy of the embedding theory, focusing on the electronic structure of the primary quantum mechanics region. We examine the influence of embedding potential on the structural energy, local density of states, total electronic density, and electronic states at the Fermi energy of the primary quantum system. The analysis illustrates the success and limitations of the embedding scheme in describing the electronic structure of the primary quantum region.

DOI: 10.1103/PhysRevB.78.235119 PACS number(s): 71.15.Dx, 71.15.Mb, 73.20.At

Despite the ever increasing computational capacity, modeling and simulation of complex material systems at the atomic level still remain a challenge.[1] For example, quantum mechanics (QM) is necessary for a proper treatment of bond breaking, charge transfer, electron excitation, magnetism, etc., in materials; however, owing to its demanding computational cost, the application of QM has been limited to relatively small systems consisting of up to a few hundreds of atoms. On the other hand, atomistic simulations based on empirical interatomic potentials are often capable of describing small-amplitude vibrations and torsions, elastic deformation and electrostatic interactions, etc., in many materials and biological systems. Termed as molecular mechanical (MM) methods, these empirical atomistic approaches can treat millions of atoms or more. Algorithms that combine quantum mechanics and molecular mechanics (QM/MM) offer a promising solution to the computational challenge in materials science and they have attracted great interest in the past decade.[1,2] In general, QM/MM coupling schemes can be divided into two categories, mechanical coupling and quantum coupling, depending on how the interaction energy between QM and MM is formulated.[2] With the mechanical coupling, the interaction energy is calculated at the MM level, whereas the interaction energy is calculated at the QM level in the quantum coupling.

In a typical QM/MM simulation, QM is often used to treat a small region which is embedded into a much larger region dealt by MM simulations. Therefore, how the boundary conditions are imposed at the interface between the two regions is of crucial of importance for QM/MM algorithms. The self-consistent embedding theory[2–8] is a quantum coupling framework which provides rigorous boundary conditions for the primary QM region embedded into a larger environment. Although in general the environment could be treated at either QM or MM level, we focus on an MM description of the environment in this paper. Central to the embedding theory is the embedding potential, which is defined as a functional derivative of the interaction energy (between the primary region and environment) with respect to the electron density in the primary region. The single-particle embedding potential represents the quantum-mechanical effects of the environment, including both nuclei and electrons, on the behavior of the electrons in the primary region. The embedding potential is included in the Hamiltonian of the electrons in the primary region, for which the quantum-mechanical problem is solved self-consistently. Wesolowski and Warshel[3] were among the first to use this scheme for describing chemical processes in solutions. Govind *et al.*[4,5] utilized the method to explore the electronic structure and excited states of molecules adsorbed on metal surfaces. Recently, the embedding method has been implemented in the context of QM/MM simulations of metallic materials.[6,7] In the self-consistent embedding QM/MM theory, the Kohn-Sham (KS) density-functional theory (DFT) (KS-DFT) (Ref. 9) is employed to perform QM calculations. The electrons and nuclei in the MM region provide the embedding potential which needs to be determined self-consistently based on DFT. The ultimate goal of the QM/MM embedding theory is to ensure that the resultant physical quantities (such as electronic structure, energetics, dynamics, etc.) are the same as those with the QM calculations applied to the entire system. To what extent can such QM/MM embedding scheme reproduce the full QM results is the main objective of this paper. In particular, we will focus on electronic structure in the primary QM region and examine how well it is described by the QM/MM scheme. Since there has been no systematic study on the subject matter in the literature, this paper provides a much-needed critical assessment of the theory, which could motivate more research in this direction, both in terms of methodology development and applications.

First, we give a brief introduction to the self-consistent QM/MM embedding theory. The entire system is partitioned into two regions in real space: a QM region (region I) and an MM region (region II) with the former embedded into the latter. Region I is normally treated by KS-DFT calculations while region II by classical atomistic simulations. We associate each MM atom in region II with a valence electron density (ρ^{at}) and a pseudopotential; both of them are constructed *a priori* and remain fixed during a QM/MM simulation.[7] The charge density of region I, ρ^{I}, is the degree

of freedom and is determined self-consistently by minimizing the energy functional

$$E[\mathbf{R}^{\mathrm{tot}}] = \min_{\rho^{\mathrm{I}}}\{E_{\mathrm{DFT}}[\rho^{\mathrm{I}};\mathbf{R}^{\mathrm{I}}] + E_{\mathrm{OF}}^{\mathrm{int}}[\rho^{\mathrm{I}},\rho^{\mathrm{II}};\mathbf{R}^{\mathrm{I}},\mathbf{R}^{\mathrm{II}}]\}. \quad (1)$$

Here $\mathbf{R}^{\mathrm{tot}} \equiv \mathbf{R}^{\mathrm{I}} \cup \mathbf{R}^{\mathrm{II}}$ and $\mathbf{R}^{\mathrm{I}}$ and $\mathbf{R}^{\mathrm{II}}$ denote atomic coordinates in regions I and II, respectively. The charge density of region II, ρ^{II}, is defined as the superposition of atomic-centered charge densities ρ^{at} via $\rho^{\mathrm{II}}(\mathbf{r}) = \Sigma_{i\in\mathrm{II}}\rho^{\mathrm{at}}(\mathbf{r}-\mathbf{R}_i)$, which only changes upon the relaxation of region II ions. The interaction energy between regions I and II, $E_{\mathrm{OF}}^{\mathrm{int}}$, formulated by orbital-free density-functional theory (OFDFT) (Refs. 10–12) is defined as follows:

$$E_{\mathrm{OF}}^{\mathrm{int}}[\rho^{\mathrm{I}},\rho^{\mathrm{II}};\mathbf{R}^{\mathrm{I}},\mathbf{R}^{\mathrm{II}}] = E_{\mathrm{OF}}[\rho^{\mathrm{tot}};\mathbf{R}^{\mathrm{tot}}] - E_{\mathrm{OF}}[\rho^{\mathrm{I}};\mathbf{R}^{\mathrm{I}}] - E_{\mathrm{OF}}[\rho^{\mathrm{II}};\mathbf{R}^{\mathrm{II}}], \quad (2)$$

where $\rho^{\mathrm{tot}} = \rho^{\mathrm{I}} + \rho^{\mathrm{II}}$. The corresponding embedding potential $\mu_{\mathrm{emb}}(\mathbf{r})$ is defined as

$$\mu_{\mathrm{emb}}(\mathbf{r}) \equiv \frac{\delta E_{\mathrm{OF}}^{\mathrm{int}}[\rho^{\mathrm{I}},\rho^{\mathrm{II}};\mathbf{R}^{\mathrm{I}},\mathbf{R}^{\mathrm{II}}]}{\delta\rho^{\mathrm{I}}}, \quad (3)$$

which represents the effective single-particle potential that region I electrons feel due to the presence of region II (Refs. 3–8); it is through $\mu_{\mathrm{emb}}(\mathbf{r})$ that the QM/MM coupling is achieved quantum mechanically at the level of OFDFT. The embedding potential depends on ρ^{I} and is updated self-consistently. For QM modeling of region I, we employ KS-DFT based plane-wave pseudopotential method implemented in Vienna *ab initio* simulation package (VASP).[13–15] In this manner, the modified KS Hamiltonian is given by $H = T + V_{\mathrm{loc}} + V_{\mathrm{NL}}$, where T is the KS kinetic-energy operator. The local potential is given by

$$V_{\mathrm{loc}} = V_{\mathrm{loc}}^{\mathrm{ion}} + V_H[\rho^{\mathrm{I}}] + V_{\mathrm{xc}}[\rho^{\mathrm{I}}] + \mu_{\mathrm{emb}}[\rho^{\mathrm{I}},\rho^{\mathrm{II}}], \quad (4)$$

where $V_{\mathrm{loc}}^{\mathrm{ion}}$, $V_H[\rho]$, and $V_{\mathrm{xc}}[\rho]$ are local part of the ionic pseudopotential, the Hartree potential, and the exchange-correlation potential, respectively. The nonlocal potential V_{NL} depends on the local potential.[14,15] For a given charge density ρ^{I}, one can determine the band-structure energy ϵ_n and the KS orbital ϕ_n from the modified KS eigenvalue equations. Then the energy of Eq. (1) is calculated by

$$E[\mathbf{R}^{\mathrm{tot}}] = \sum_n f_n\epsilon_n + E_{\mathrm{dc}} + \gamma_{\mathrm{Ewald}}[\mathbf{R}^{\mathrm{I}}], \quad (5)$$

where f_n is the occupation number in band n, γ_{Ewald} is the Madelung energy, and the double-counting energy term is

$$E_{\mathrm{dc}} = -E_H[\rho^{\mathrm{I}}] + E_{\mathrm{xc}}[\rho^{\mathrm{I}}] - \int d\mathbf{r}\, V_{\mathrm{xc}}(\mathbf{r})\rho^{\mathrm{I}}(\mathbf{r}) + E_{\mathrm{OF}}^{\mathrm{int}}[\rho^{\mathrm{I}},\rho^{\mathrm{II}};\mathbf{R}^{\mathrm{I}},\mathbf{R}^{\mathrm{II}}] - \int d\mathbf{r}\, \mu_{\mathrm{emb}}(\mathbf{r})\rho^{\mathrm{I}}(\mathbf{r}). \quad (6)$$

Following the same numerical schemes as implemented in VASP, ρ^{I} can be self-consistently determined for a given ionic configuration. The Wang-Govind-Carter kinetic-energy functional[11] with a density-dependent kernel is used for the OFDFT. The Perdew-Zunger local-density approximation (LDA) (Ref. 16) is used to evaluate the exchange-correlation energy. More details of the self-consistent QM/MM embedding method can be found in Refs. 6 and 7.

In this work, we use aluminum as an example in the analysis of the coupling method because of the existence of the excellent kinetic-energy functional of Al. First, we examine the energetic convergence with respect to the volume of region I. Specifically, we determine the critical dimension of region I, so that the local energy of the central atom in region I is converged; i.e., its energy is identical to the bulk value. To this end, we vary the size of region I as $na_0 \times na_0 \times a_0$, in which n changes from 1 to 6. a_0 is the equilibrium lattice constant of Al. For each n value, we perform three different calculations: (i) *bulk*—where region I is treated by a stand-alone KS-DFT calculation with periodic boundary conditions in x, y, and z directions; (ii) *vacuum*—where region I is treated by a stand-alone KS-DFT calculation as an isolated system, and a periodic boundary condition is applied in z direction and vacuum is placed in x and y directions; and (iii) *embedding*—the present embedding method which incorporates the influence of region II onto region I. A periodic boundary condition is applied in z direction for region I, so that the QM/MM coupling is carried out only in x and y directions. If the embedding scheme worked perfectly, the embedding calculation would reproduce the same results of the bulk calculation. On the other hand, the discrepancies between them can be attributed to the coupling errors, which are always present for realistic calculations. In addition, there are certain physical quantities that cannot be reproduced by the present method even if the method was made exact. This is a subtle point that deserves a close scrutiny in this paper. A plane-wave cutoff of 300 eV is used in the calculations and the k points are sampled according to the Monkhorst-Pack method[17] with a $1\times1\times9$ mesh in the vacuum and the embedding calculations; in the bulk calculations, k mesh of $9\times9\times9$ is used for $n=1$, $5\times5\times9$ for $n=2$, $3\times3\times9$ for $n=3$, and $2\times2\times9$ for remaining n values. The computational cost is similar between the vacuum and the embedding calculations for the same size of region I.

We define the structural energy of atom l in a solid as[18]

$$E_l = \int_{-\infty}^{E_F} E\sum_{\alpha} n_{\alpha l}(E)dE, \quad (7)$$

where $n_{\alpha l}(E)$ is the local density of states (DOS) (LDOS) of atomic orbital α of atom l and E_F is the Fermi energy. The structural energy E_l represents the energy contribution of the atom l to the total band-structure energy of the system and is dominated by the bonding energy contribution from its neighboring atoms. Therefore if region I is large enough for the central atom, the surfaces would not affect E_l of the central atom, and thus as far as the central atom is concerned, the vacuum calculation is equivalent to the embedding calculation. However, as shown later, this conclusion is not true for the global properties of the material.

In Fig. 1(a), we show the structural energy of the central atom as a function of n for both the vacuum and the embedding calculations. The horizontal line represents the structural energy from the bulk calculations. It is found that for $n \geq 4$, the vacuum and the embedding calculations yield the

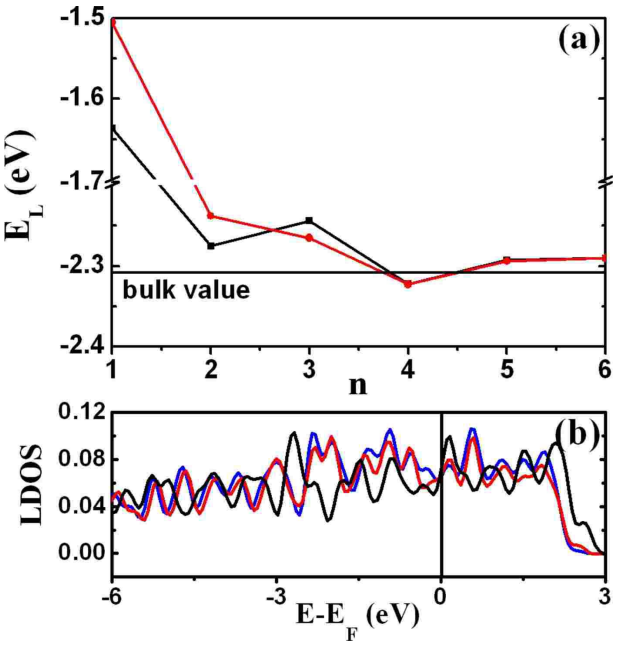

FIG. 1. (Color online) (a) The structural energy E_l of the atom l which is located at the center of the primary region with n varying from 1 to 6. The black and red solid lines represent the structural energy from the vacuum and the embedding calculations, respectively. (b) The LDOS of the atom at the interface for $n=4$. The blue, red, and black solid lines denote the results from the vacuum, the embedding, and the bulk calculations, respectively. The Fermi level is shifted to zero.

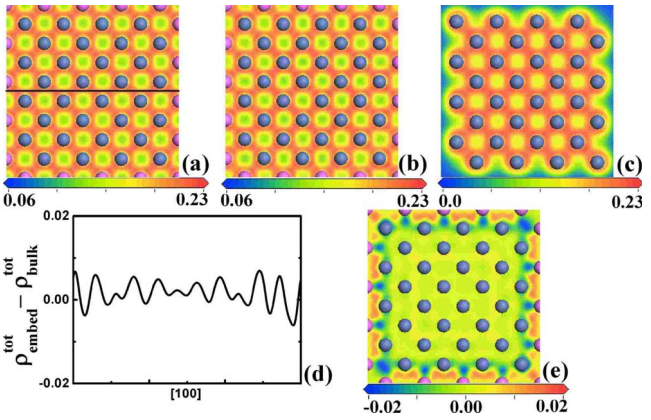

FIG. 2. (Color online) The total charge density (Å^{-3}) in the primary region for $n=4$ from (a) the bulk calculation, (b) the embedding calculation, and (c) the vacuum calculation. (d) The difference of the total charge density between the embedding and the bulk calculations along a horizontal line shown in (a). (e) The difference of the charge density ρ^{I} between the embedding and the vacuum calculations. The blue and magenta (partially visible) spheres represent regions I and II atoms, respectively.

identical E_l for the central atom, indicating that the surfaces do not affect the central atom for $n \geq 4$ in terms of the structural energy. In this analysis, we take the central atom as the sole atom of interest whose properties are to be determined accurately by the QM/MM approach. In general, a region of "central atoms" ("central region") may be required; as a result, the critical size determined here needs to be increased by an amount which corresponds to the volume of the central region. Nevertheless, the results are useful for providing the lower bound of the vacuum dimension for similar material systems.

A critical observation from Fig. 1(a) is that both the converged vacuum and the embedding results ($n \geq 4$) deviate from the bulk value by ~0.01 eV. The discrepancy can be understood by realizing that the LDOS in Eq. (7) depends on the energy spectrum of the total Hamiltonian of the system. In other words, it does not only depend on the electronic states of the primary region which are determined accurately but also on the eigenstates of the environment, which are not considered explicitly in the present theory—only the electron density of region II is included in the formulation. Therefore, the present theory cannot give correct LDOS and the structural energy. The reasons that the present theory focuses on the coupling of the electron densities as opposed to the electronic wave functions are the following: (i) the electron density is the essential physical quantity—a correct determination of the electron density ensures the correct description of other physical properties of the system, particularly, the total energy and atomic structure—and (ii) it is more convenient to deal with the electron density than wave functions because the latter often depend on the band index and k point, in addition to the spatial dependence. In particular, a wave-function-based embedding scheme would require the coupling of every electronic bands unless further approximations are involved. The downside of the electron-density-based coupling method is that the information of the electronic states in region II is missing, and as result certain physical quantities such as LDOS cannot be correctly accounted for. The embedding theory becomes exact only in the limit when the size of the primary region approaches infinity. This is because although the QM/MM coupling is supposed to eliminate fictitious surface effects, it does not render the finite primary region infinite. Therefore certain electronic fine structure of an extended system, such as Van Hove singularities of DOS (Ref. 19) cannot be reproduced by a QM/MM calculation of a finite system. For example, as shown in Fig. 1(b), the LDOS for an atom at the surface of the vacuum calculation traces closely to that of the corresponding atom at the QM/MM interface of the embedding calculation for $n=4$; but both deviate significantly from the LDOS of the bulk system. This result confirms that the absence of the electronic states in region II leads to an incorrect LDOS in spite of the presence of the embedding potential.

Next, we examine the results of electron density from the bulk, embedding, and vacuum calculations. As shown in Figs. 2(a), 2(b), and 2(d), there is a rather good agreement of the total charge density between the bulk and the embedding calculations. The minor density errors at the interface are due to the fact that the superposition of the atomic-centered charge densities does not reproduce the bulk charge density exactly. Our recent work has shown that a better approximation of the atomic-centered density could reduce the density errors significantly. On the other hand, as expected, the total charge density of the vacuum calculation shown in Fig. 2(c) differs significantly from the bulk result, with the greatest

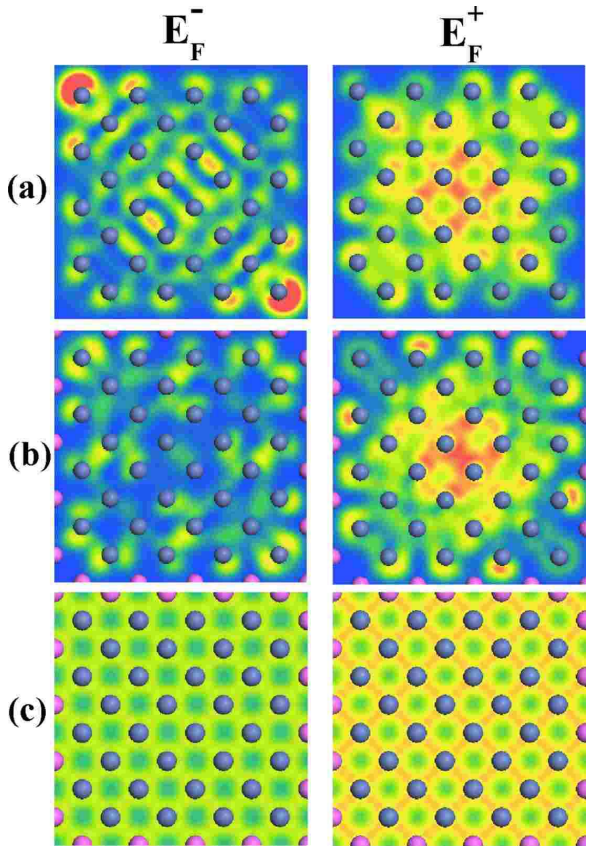

FIG. 3. (Color online) The band-decomposed charge density at E_F^- and E_F^+ for $n=4$ from (a) the vacuum calculation, (b) the embedding calculation, and (c) the bulk calculation. Contour scale (in $Å^{-3}$) ranges from 0.0 (blue) to 0.007 (red). The contour scale at E_F^- in (c) ranges from 0.0 (blue) to 0.002 (red). The blue and magenta (partially visible) spheres represent regions I and II atoms, respectively.

errors occurring at the surfaces. Specifically, the surface charge density does not resemble the directional bonding that exists in bulk Al.[7,20] In Fig. 2(e), we plot the differential charge density between the embedding and the vacuum calculations for region I, $\rho^{\mathrm{I}}_{\mathrm{embed}}(\mathbf{r})-\rho^{\mathrm{I}}_{\mathrm{vacuum}}(\mathbf{r})$. It is evident that the charge-density difference is primarily localized at the QM/MM interface, and it is due to the nearest-neighbor directional bonding. The result shows the benefit of coupling charge densities as opposed to wave functions because "charge densities are nearsighted and wave functions are not."[21] Therefore, the embedding potential can reproduce the directional bonding that is characteristic of bulk Al, and its influence is short ranged.

Finally, we analyze the band-decomposed charge density to examine the effect of the embedding potential on the electronic states near the Fermi energy. The band-decomposed charge density at E_F^- and E_F^+ for the vacuum and the embedding calculations are presented in Figs. 3(a) and 3(b), respectively. Here, the band-decomposed charge density at E_F^- (E_F^+) is evaluated in the energy range between $E_F-0.1$ eV and E_F (E_F and $E_F+0.1$ eV), including the contributions of all k points. We focus on these states because they constitute possible "frontier orbitals" and could play critical roles in chemical reactions for certain material systems. It turns out that the majority contributions to the band-decomposed charge density at E_F^- come from the surface atoms in the vacuum calculation; the electronic states at E_F^- are mainly localized surface states. On the other hand, the electronic states at E_F^+ are much more extended and spread all over the primary region. Although the surface states are clearly visible in Fig. 3(a) of the vacuum calculation, they are significantly suppressed by the embedding potential as shown in Fig. 3(b). Further analysis reveals that the electronic states associated with the boundary atoms are shifted into lower energy levels in the embedding calculation. As expected, the short-ranged embedding potential has little effect on the extended states at E_F^+; there is minor difference between Figs. 3(a) and 3(b) in terms of E_F^+ states. Therefore the embedding potential could suppress or eliminate the localized surface states arising from the vacuum calculation, but it has little effect on the extended states that are inherent to the bulk system. More importantly, like the vacuum calculation, the band-decomposed charge density of the embedding calculation does not reproduce the periodic charge density of the bulk calculation shown in Fig. 3(c). This is again due to the finite size of the primary region. However, it should be realized that the present embedding method is *not* developed to treat perfect lattice but rather to treat systems where the translational symmetry is broken, such as lattice defects. For these systems, the embedding method is quite useful because no matter how large a bare cluster is chosen for a QM calculation, the surface states are always present and show up close to the Fermi energy as frontier orbitals. These fictitious states can be eliminated by a moderate size embedding calculation. On the other hand, the fictitious surface states cannot be suppressed or removed by mechanical-based QM/MM schemes[2,22,23] even though these schemes are capable of predicting reliable atomic structures. Another advantage of the present method over the mechanical-based methods is that for certain magnetic systems, the spin-polarized DFT calculation of a bare cluster (which is done in mechanical-based methods) can often have difficulties to converge while the present method does not suffer from this problem.

In summary, we have performed critical analysis on the self-consistent embedding QM/MM theory for material simulations, using fcc aluminum as an example. We compare the energetics and the electronic structure between the vacuum, embedding, and bulk calculations in aluminum. We find that the embedding scheme reproduces accurately the total electron density of the bulk system. The embedding method could eliminate the localized surface states at the Fermi energy which originate from the dangling bonds at the surface. Since the embedding method does not involve the electronic states from the environment, it cannot provide the correct LDOS of the system. Although we believe the qualitative conclusions drawn from the present study are applicable to other materials as well, details remain to be assessed in future studies. In any case, the assessment of the embedding QM/MM method provides guidance for further development and application of the method in material problems.

The work at Tsinghua University was supported by "973" Project from the Ministry of Science and Technology of China (Grant No. 2006CB605102). The work at California State University—Northridge was supported by NSF under Grant No. DMR-0611562 and DOE under Grant No. DE-FC02-06ER25791.

*ganglu@csun.edu

[1] G. Lu and E. Kaxiras, in *Handbook of Theoretical and Computational Nanotechnology*, edited by M. Rieth and W. Schommers (American Scientific, Stevenson Ranch, CA, 2004), Chap. 22.
[2] H. Lin and D. G. Truhlar, Theor. Chem. Acc. **117**, 185 (2007).
[3] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).
[4] N. Govind, Y. A. Wang, A. J. R. da Silva, and E. A. Carter, Chem. Phys. Lett. **295**, 129 (1998).
[5] T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter, Phys. Rev. Lett. **86**, 5954 (2001).
[6] N. Choly, G. Lu, W. E, and E. Kaxiras, Phys. Rev. B **71**, 094101 (2005).
[7] X. Zhang and G. Lu, Phys. Rev. B **76**, 245111 (2007).
[8] P. Cortona, Phys. Rev. B **44**, 8454 (1991).
[9] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[10] L. W. Wang and M. P. Teter, Phys. Rev. B **45**, 13196 (1992).
[11] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **60**, 16350 (1999).
[12] P. Garcia-Gonzalez, J. E. Alvarellos, and E. Chacon, Phys. Rev. B **53**, 9509 (1996).
[13] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993); **49**, 14251 (1994).
[14] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[15] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[16] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[17] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[18] Wang Chong-yu, Liu Sen-ying, and Han Lin-guang, Phys. Rev. B **41**, 1359 (1990).
[19] N. W. Ashcroft, Phys. Rev. B **19**, 4906 (1979).
[20] K. M. Carling, G. Wahnström, T. R. Mattsson, N. Sandberg, and G. Grimvall, Phys. Rev. B **67**, 054101 (2003).
[21] W. Kohn, Int. J. Quantum Chem. **56**, 229 (1995).
[22] D. Bakowies and W. Thiel, J. Phys. Chem. **100**, 10580 (1996).
[23] Y. Liu, G. Lu, Z. Chen, and N. Kioussis, Modell. Simul. Mater. Sci. Eng. **15**, 275 (2007).