

Linear-scaling quantum calculations using non-orthogonal localized molecular orbitals

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2008 J. Phys.: Condens. Matter 20 294209

(<http://iopscience.iop.org/0953-8984/20/29/294209>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 29/05/2010 at 13:33

Please note that [terms and conditions apply](#).

Linear-scaling quantum calculations using non-orthogonal localized molecular orbitals

Steven K Burger¹ and Weitao Yang²

¹ Department of Chemistry, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

² Department of Chemistry, Duke University, Box 90346, Durham, NC 27708-0346, USA

E-mail: burgers@univmail.cis.mcmaster.ca and yang@chem.duke.edu

Received 20 February 2008, in final form 22 April 2008

Published 24 June 2008

Online at stacks.iop.org/JPhysCM/20/294209

Abstract

An absolute energy minimum variational principle is used for carrying out linear-scaling calculations with non-orthogonal localized orbitals. Comparing with results based on orthogonal localized molecular orbitals, the method is shown to give significantly more accurate results when the localized molecular orbitals are allowed to be non-orthogonal. This is made possible by introducing a second minimization for approximating the inverse overlap matrix. We also show how an exact line search may be used efficiently with the conjugate gradient method for minimizing the energy functional.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

To address complex problems with electronic structure calculations, there is a desire to study larger systems over longer timescales. However, conventional semi-empirical and *ab initio* methods are constrained by nonlinear-scaling operations associated with constructing the one-electron Hamiltonian. In particular, in order to obtain the eigenstates of a system, the Hamiltonian is diagonalized. This operation scales cubically with the size of the system and, despite a reasonable prefactor, becomes an inefficient operation for large systems.

The first linear-scaling approach was the divide-and-conquer method developed by Yang [1] which is based on exploring the localized nature of the density matrix. This led to a surge of interest among researchers in developing linear scaling electronic structure methods. In general, to achieve linear scaling, solutions for molecular orbitals or density matrix must exist that are localized in real space. The canonical molecular orbitals, which are obtained by diagonalization are known to be highly delocalized, but they can be localized by a unitary transformation. Among the many methods for localizing the density are the minimization of the Boys' orbital spread [2], the Edmiston–Ruedenberg's repulsion energy [3] and von Niessens' charge density overlap [4]. These methods have shown that a good amount of localization can be achieved

with orthogonal orbitals, but that long-range non-localized tails are often observed.

More recently, methods were developed to use non-orthogonal localized orbitals [5–8]. These orbitals, first introduced in electronic structure calculations by Adams and Gilbert [9–11], were shown recently to give much better localization of the electronic density [7, 8]. Also the resulting orbitals are more transferable between systems, since they only include information from the local environment [6].

There have been two main categories of approaches to linear scaling, with density matrix or with localized molecular orbitals. Knowing that localized molecular orbital (LMO) solutions exist has encouraged the development of a number of methods to obtain LMOs [1, 5, 6, 12–30] with an effort that scales as $O(N)$, where N is the number of electrons.

The divide-and-conquer method, which focuses on the density matrix, remains a very effective method [1, 12, 13, 31–34]. The idea is to split up the molecule into a set of smaller subsystems, each of which is solved separately with diagonalization. Stewart [30] applied Jacobi rotations to the Fock matrix to obtain LMOs, while Anikin and co-workers [35] used a penalty function to enforce non-orthogonality while solving for the system variationally. Mauri *et al* [25] used a conjugate gradient search for the orthogonal localized orbitals to achieve linear scaling. The

work of Mauri *et al* was later generalized to overcome the multi-minimum problem [26]. Stechel *et al* used block diagonalization to solve for the eigenstates in a similar way as divide-and-conquer [27]. Additional work was done by Galli and Parrinello [28] in developing a linear-scaling method applicable to plane-wave calculations, which was extended for use in the quantum Monte Carlo method [29]. Additional methods were reviewed in [36, 37].

In this work we show how Yang's absolute energy variational principle [5] for linear scaling gives significantly more accurate results compared to methods which use orthogonal orbitals. This is due to the increased flexibility in approximating the inverse, which allows the orbitals to be orthogonal and more localized. We also give formulas for exact line searching with conjugate gradient, which significantly improves the convergence of the minimization.

2. Theory

The energy minimization principle for most electronic structure calculations is formulated with the electron density matrix, $\rho(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | \hat{\rho} | \mathbf{r}' \rangle$. Written in terms of the occupied states, the one-electron reduced density operator is given by,

$$\hat{\rho} = 2 \sum_m^{N/2} |\psi_m\rangle \langle \psi_m| \quad (1)$$

where the states $|\psi_m\rangle$ are normalized and orthogonal. These states can be expanded in terms of a set of local basis functions $\{\phi_j\}$,

$$\psi_m(\mathbf{r}) = \sum_j \mathbf{C}_{jm} \phi_j(\mathbf{r}). \quad (2)$$

For Hartree–Fock or density-functional calculations the linear coefficients can be obtained by solving the generalized eigenvalue equation $\mathbf{h}\mathbf{C} = \mathbf{s}\mathbf{C}\epsilon$, where $h_{ij} = \langle \phi_i | \hat{h} | \phi_j \rangle$ and $s_{ij} = \langle \phi_i | \phi_j \rangle$. The canonical molecular orbital solution is highly delocalized, but there are energetically equivalent, orthogonal localized molecular orbitals (OLMO) that can be obtained by a unitary transformation. The existence of such a localized set of the orbitals is necessary for formulating a linear-scaling algorithm, as some form of cutoff must be imposed on the density:

$$\langle \mathbf{r} | \hat{\rho} | \mathbf{r}' \rangle = 0 \quad \text{for } |\mathbf{r} - \mathbf{r}'| > k, \quad (3)$$

where k is a constant. Since we use local atomic orbitals as a basis sets, we impose the cutoff by

$$\mathbf{C}_{jm} = 0 \quad \text{for } |\mathbf{R}_m^{\text{LMO}} - \mathbf{R}_j| > R_c \quad (4)$$

where $\mathbf{R}_m^{\text{LMO}}$ is the center vector of the m th LMO and \mathbf{R}_j is the position vector of the atom that the j th atomic orbital belongs to. The constraint that the orbitals be orthogonal is not necessary, although previously it was thought that releasing this constraint would not be effective for obtaining more localized orbitals [24]. However it was recently shown that using non-orthogonal localized molecular orbitals (NOLMO) gives 10%–30% better localization [7]. Using NOLMOs then, we would expect to obtain a more accurate energy than

with OLMOs when the same cutoff given by equation (4) is imposed.

The density operator in terms of NOLMO can be obtained from a minimum energy principle. This states that the total energy, $E(N)$ can be obtained by minimizing the energy as a function of density, so long as the density satisfies three conditions: (1) Hermiticity, $\hat{\rho} = \hat{\rho}^\dagger$ (2) idempotency, $\hat{\rho}^2 = \hat{\rho}$ and (3) normalization, $2 \text{Tr}[\hat{\rho}] = N$.

To preserve the idempotent condition of the density for non-orthogonal states the density given by equation (1) must include an extra term. The most straightforward approach is to use the inverse of the overlap matrix, which gives,

$$\hat{\rho} = 2 \sum_{ij}^{N/2} |\psi_i\rangle S_{ij}^{-1} \langle \psi_j|. \quad (5)$$

However, solving for \mathbf{S}^{-1} is an $O(N^3)$ operation, so instead approximations to this matrix are used [6, 25]. In [25] a truncated series expansion of $\mathbf{S}^{-1} = \sum_{n=0}^M (\mathbf{I} - \mathbf{S})^n$ is used instead to give an $O(N)$ algorithm, where only odd powers of the series expansion are used. At the lowest order ($M = 1$) the density is

$$\hat{\rho} = 2 \sum_{ij}^{N/2} |\psi_i\rangle (2\mathbf{I} - \mathbf{S})_{ij} \langle \psi_j|. \quad (6)$$

It was shown in [5] that the generalized inverse of the overlap matrix, \mathbf{S}^- , defined by $\mathbf{S}\mathbf{S}^- \mathbf{S} = \mathbf{S}$ can also be used. A truncated approximation to the matrix \mathbf{S}^- can be obtained with $O(N)$ operations by taking advantage of the relation,

$$\text{Tr}[\mathbf{B}\mathbf{S}^-] = \min_{\mathbf{X}=\mathbf{X}^\dagger} \text{Tr}[\mathbf{B}(2\mathbf{X} - \mathbf{X}\mathbf{S}\mathbf{X})] \quad (7)$$

where \mathbf{B} is any positive definite matrix and \mathbf{X} is constrained to be Hermitian. This gives the density as,

$$\hat{\rho} = 2 \sum_{ij}^{N/2} |\psi_i\rangle (2\mathbf{X} - \mathbf{X}\mathbf{S}\mathbf{X})_{ij} \langle \psi_j|. \quad (8)$$

The same density expression was also given in [6], although it was derived differently with the constraint that \mathbf{X} is positive definite. To ensure that the method remains linear for matrix multiplications, in the same way we defined R_c we set a sparsity pattern for \mathbf{X}

$$\mathbf{X}_{ml} = 0 \quad \text{for } |\mathbf{R}_m^{\text{LMO}} - \mathbf{R}_l^{\text{LMO}}| > R_X. \quad (9)$$

Using any of these expressions for the density, an energy functional of $N/2$ orbitals can be defined as,

$$W[\{\psi\}, \mathbf{X}] = \langle E \rangle - 2\eta \langle N \rangle + \eta N. \quad (10)$$

For the non-interacting case $\langle E \rangle = 2 \text{Tr} \hat{h} \hat{\rho}$, $\langle N \rangle = 2 \text{Tr} \hat{\rho}$ and η is a parameter which should be larger than the largest eigenvalue of the canonical solution. Using equation (10) it is shown in [5] that,

$$E(N) = \min_{\mathbf{C}} \min_{\mathbf{X}=\mathbf{X}^\dagger} 2 \text{Tr} [(\mathbf{H} - \eta\mathbf{S})(2\mathbf{X} - \mathbf{X}\mathbf{S}\mathbf{X})] + \eta N \quad (11)$$

where $\mathbf{H} = \mathbf{C}^\dagger \mathbf{h} \mathbf{C}$ and $\mathbf{S} = \mathbf{C}^\dagger \mathbf{s} \mathbf{C}$. When sparse matrix operations are used for the matrix multiplications

then equation (11) can be used as a linear-scaling method. Equation (11) appears in [5] and is the first of two variational principles. In the case given here the rank of the matrix \mathbf{C} is restricted to be equal to the number of occupied orbitals, $N/2$, so that unoccupied orbitals are not used in the optimization. The second variational principle developed by Yang [5] allows the use of unoccupied orbitals and thus permits long-range density rearrangements during the optimization. This may make the optimization less prone to being trapped in local minima, although this problem was not observed for the system tested in this paper. Thus the present paper only focuses on using equation (11).

3. Implementation

There are a number of ways to do the minimizations in equation (11). In [5] it is shown that the variational principle is valid for any fixed \mathbf{X} . This implies that we can perform the minimization in either order or together as one vector. Ideally both variables would be minimized concurrently, but we found that this proves to be inefficient because of the large number of terms that arise in determining the length of the line search. Instead we minimized the functional with respect to \mathbf{C} , minimizing the function with respect to \mathbf{X} at each step.

For the subsequent formulas we shift the fock matrix to make it negative definite, by setting $\mathbf{h} = \mathbf{h} - \eta\mathbf{s}$. The gradient required to minimize equation (11) with respect to \mathbf{C} is,

$$\nabla_{\mathbf{C}}E(\mathbf{C}, \mathbf{X}) = 8\mathbf{h}\mathbf{C}\mathbf{X} - 4(\mathbf{h}\mathbf{C}\mathbf{X}\mathbf{S}\mathbf{X} + \mathbf{s}\mathbf{C}\mathbf{X}\mathbf{H}\mathbf{X}), \quad (12)$$

and the gradient with respect to \mathbf{X} is,

$$\nabla_{\mathbf{X}}E(\mathbf{C}, \mathbf{X}) = 4\mathbf{H} - 2(\mathbf{H}\mathbf{X}\mathbf{S} + \mathbf{S}\mathbf{X}\mathbf{H}). \quad (13)$$

The gradient terms can be used with the limited memory BFGS method or conjugate gradient minimization. These methods generate a search direction, Δ , for which a length, λ needs to be determined. Traditional bracketed line search methods such as those found in [38] can be used. However since we have the exact analytic form, we can work out λ exactly. For \mathbf{X} we can expand equation (11), assuming \mathbf{C} is constant,

$$E(\mathbf{C}, \mathbf{X}) = 2\text{Tr}[\mathbf{H}(2(\mathbf{X} + \Delta_{\mathbf{X}}\lambda_{\mathbf{X}}) - (\mathbf{X} + \Delta_{\mathbf{X}}\lambda_{\mathbf{X}})\mathbf{S}(\mathbf{X} + \Delta_{\mathbf{X}}\lambda_{\mathbf{X}}))] \quad (14)$$

setting $\frac{dE(\mathbf{C}, \mathbf{X})}{d\lambda_{\mathbf{X}}} = 0$ we can solve for $\lambda_{\mathbf{X}}$ giving,

$$\lambda_{\mathbf{X}} = \frac{\text{Tr}[\mathbf{H}\Delta_{\mathbf{X}}(\mathbf{I} - \mathbf{S}\mathbf{X})]}{\text{Tr}[\mathbf{H}\Delta_{\mathbf{X}}\mathbf{S}\Delta_{\mathbf{X}}]}. \quad (15)$$

The convergence behavior for \mathbf{X} should be good, since equation (11) is a quadratic function of \mathbf{X} . For \mathbf{C} though, equation (11) is a quartic. This means that we must solve a cubic equation to determine $\lambda_{\mathbf{C}}$ and the convergence behavior is not expected to be as good as \mathbf{X} . Expanding $E(\mathbf{C}, \mathbf{X})$ in terms of $\Delta_{\mathbf{C}}$ and setting $\frac{dE(\mathbf{C}, \mathbf{X})}{d\lambda_{\mathbf{C}}} = 0$ we get,

$$\begin{aligned} &\text{Tr}[2\mathbf{H}\mathbf{X}\mathbf{S}_{\mathbf{C}}\mathbf{X} - 4\mathbf{H}_{\mathbf{C}}\mathbf{X} + 2\mathbf{H}_{\mathbf{C}}\mathbf{X}\mathbf{S}\mathbf{X}] + 2\text{Tr}[\mathbf{H}\mathbf{X}\mathbf{S}_{\mathbf{C}}\mathbf{X} \\ &+ 4\mathbf{H}_{\mathbf{C}}\mathbf{X}\mathbf{S}_{\mathbf{C}}\mathbf{X} - 2\mathbf{H}_{\mathbf{C}}\mathbf{X} + \mathbf{H}_{\mathbf{C}}\mathbf{X}\mathbf{S}\mathbf{X}]\lambda_{\mathbf{C}} \\ &+ 3\text{Tr}[2\mathbf{H}_{\mathbf{C}}\mathbf{X}\mathbf{S}_{\mathbf{C}}\mathbf{X} + 2\mathbf{H}_{\mathbf{C}}\mathbf{X}\mathbf{S}_{\mathbf{C}}\mathbf{X}]\lambda_{\mathbf{C}}^2 \\ &+ 4\text{Tr}[\mathbf{H}_{\mathbf{C}}\mathbf{X}\mathbf{S}_{\mathbf{C}}\mathbf{X}]\lambda_{\mathbf{C}}^3 = 0 \end{aligned} \quad (16)$$

where $\mathbf{S}_{\mathbf{C}} = (\mathbf{C}^T\mathbf{s}\Delta_{\mathbf{C}} + \Delta_{\mathbf{C}}^T\mathbf{s}\mathbf{C})/2$, $\mathbf{S}_{\mathbf{C}\mathbf{C}} = (\Delta_{\mathbf{C}}^T\mathbf{s}\Delta_{\mathbf{C}})$, $\mathbf{H}_{\mathbf{C}} = (\mathbf{C}^T\mathbf{h}\Delta_{\mathbf{C}} + \Delta_{\mathbf{C}}^T\mathbf{h}\mathbf{C})/2$, and $\mathbf{H}_{\mathbf{C}\mathbf{C}} = (\Delta_{\mathbf{C}}^T\mathbf{h}\Delta_{\mathbf{C}})$.

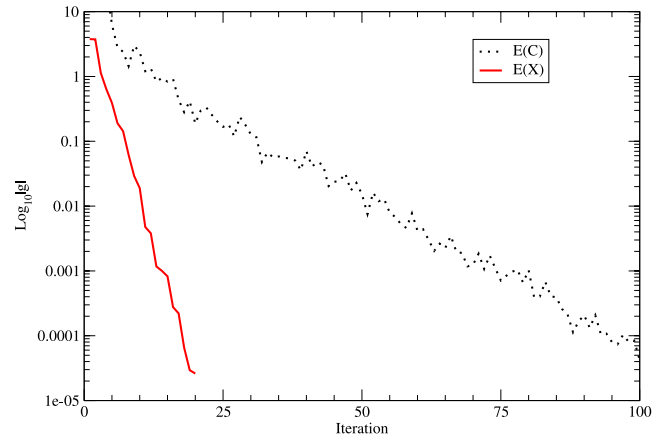


Figure 1. Convergence log plot for $g = \nabla_{\mathbf{C}}E(\mathbf{C}, \mathbf{X})$ and $g = \nabla_{\mathbf{X}}E(\mathbf{C}, \mathbf{X})$ for the molecule $\text{C}_{18}\text{H}_{38}$ with $R_c = 1 \text{ \AA}$ and $R_X = 2 \text{ \AA}$. The minimization is done with conjugate gradient and exact line searching.

4. Numerical results

In order to test our implementation we focused on octadecane ($\text{C}_{18}\text{H}_{38}$). We used Gaussian03 [39] to generate the Fock and overlap matrices from a HF/STO-3G calculation. Since the eigenvalues of the generalized eigenvalues problem are bounded above by 0.96 a.u. we used $\eta = 1$ a.u. The fock and overlap matrix elements were truncated at 4 Å, but the intermediate matrix products were not truncated. SparseKit [40] was used for all matrix operations.

There were 120 atomic orbitals used for this system. We started out with random elements for the matrix \mathbf{C} and then scaled the rows of \mathbf{C} , such that $2\text{Tr}[\mathbf{C}^T\mathbf{s}\mathbf{C}] = N$. The LMOs were centered at bonds and lone pairs according to the Lewis structure. Initially for \mathbf{X} , we set $\mathbf{X} = \mathbf{I}$. The overall minimization was done for \mathbf{C} with \mathbf{X} minimized at each step. Using this scheme with a well defined chemical structure we did not observe any local minima other than the global minimum which approximates the ground state energy.

The convergence results for conjugate gradient minimization are shown in figure 1. The results of minimizing $E(\mathbf{C}, \mathbf{X})$ with respect to \mathbf{X} with \mathbf{C} held fixed are clearly very efficient, with the minimization converging in only about 20 steps. The minimization with respect to \mathbf{C} with \mathbf{X} fixed takes about five times as long. The combined minimization is done by performing a minimization of \mathbf{X} for each step of the \mathbf{C} minimization. The convergence is slower than the individual parts but fortunately the \mathbf{X} minimization typically only requires 3 or 4 steps since the last values of \mathbf{X} can be used as an initial guess for the next step.

In figure 2 we see the convergence results for different values of R_c and R_X . The case of $\mathbf{X} = \mathbf{I}$ corresponds to assuming the orbitals are orthogonal by using the first order approximation of \mathbf{S}^{-1} . Any finite cutoff of R_X allows the orbitals to be non-orthogonal. The most dramatic difference in energy occurs between $\mathbf{X} = \mathbf{I}$ and $R_X = 2 \text{ \AA}$. For $R_c = 1 \text{ \AA}$ using $\mathbf{X} = \mathbf{I}$, the converged result is off by $6.906E_h$, while for $R_X = 3 \text{ \AA}$ the energy is only off by $0.278E_h$. This

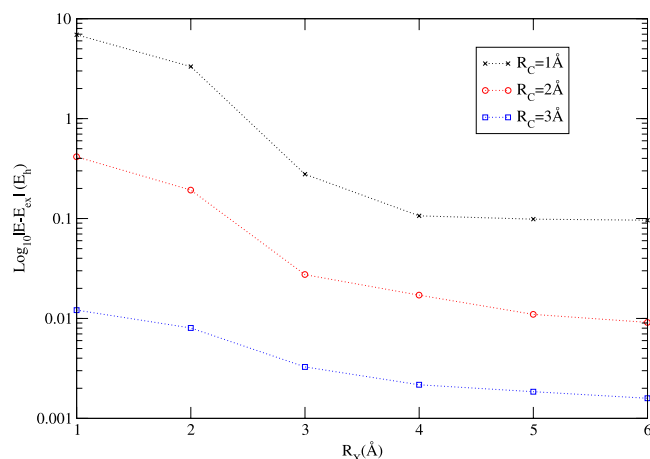


Figure 2. Converged difference to the exact energy for the system $C_{18}H_{38}$ with different values of R_c and R_X . $E_{xc} = -417.223 E_h$ is the exact energy. The point $R_X = 1 \text{ \AA}$ corresponds to the results for $X = I$.

order of magnitude improvement seems consistent for each set of R_c curves. For these calculations the exact energy result is obtained by solving the generalized eigenvalue problem without using the 4 \AA truncation for the elements of the Fock and overlap matrices.

5. Summary

Our results show that using NOLMO in the energy functional for linear-scaling calculations give much more accurate values for the energy, compared to using orthogonal orbitals. This is due to the increased ability of the orbitals to localize in a region of space, without the constraint of orthogonalization, made possible by the flexibility in the approximation of the inverse of the overlap matrix. Conjugate gradient minimization with an exact line search designed here gives rapid convergence of the orbital optimization.

Acknowledgments

Support from the National Institute of Health, the National Science Foundation (WY) and the Hamilton Community Foundation (SKB) is gratefully appreciated. Thanks to Felipe Bulat for help obtaining the Fock and overlap matrices from Gaussian03.

References

- [1] Yang W 1991 *Phys. Rev. Lett.* **66** 1438
- [2] Boys S F 1960 *Rev. Mod. Phys.* **32** 296
- [3] Edmiston C and Ruedenberg K 1963 *Rev. Mod. Phys.* **35** 457

- [4] Niessen W V 1972 *J. Chem. Phys.* **56** 4290
- [5] Yang W 1997 *Phys. Rev. B* **56** 9294
- [6] Hierse W and Stechel E B 1994 *Phys. Rev. B* **50** 17811
- [7] Liu S, Perez-Jorda J M and Yang W 2000 *J. Chem. Phys.* **112** 1634
- [8] Feng H, Bian J, Li L and Yang W 2004 *J. Chem. Phys.* **120** 9458
- [9] Adams W 1961 *J. Chem. Phys.* **34** 89
- [10] Adams W 1962 *J. Chem. Phys.* **37** 2009
- [11] Gilbert T and Lykos P 1961 *J. Chem. Phys.* **34** 2199
- [12] Lee T, York D and Yang W 1996 *J. Chem. Phys.* **105** 2744
- [13] Yang W and Lee T 1995 *J. Chem. Phys.* **105** 5674
- [14] Gillan M, Bowler D, Torralba A and Miyazaki T 2007 *Comput. Phys. Commun.* **177** 14
- [15] Mostofi A A S C and Haynes P D 2007 *Mol. Simul.* **33** 551
- [16] Hernández E, Gillan M J and Goringe C M 1996 *Phys. Rev. B* **53** 7147
- [17] Kohn W 1996 *Phys. Rev. Lett.* **76** 3168
- [18] Ordejón P, Drabold D, Grumbach M and Martin R M 1993 *Phys. Rev. B* **48** 14646
- [19] Ordejón P, Drabold D, Martin R M and Grumbach M P 1995 *Phys. Rev. B* **51** 1456
- [20] Li X-P, Nunes R W and Vanderbilt D 1993 *Phys. Rev. B* **47** 10891
- [21] Voter A F, Kress J D and Silver R N 1996 *Phys. Rev. B* **53** 12733
- [22] Goedecker S and Colombo L 1994 *Phys. Rev. Lett.* **73** 122
- [23] Horsfield A P and Bratkovsky A M 1996 *Phys. Rev. B* **53** 15381
- [24] Sundberg K R, Bicerano J and Lipscomb W N 1979 *J. Chem. Phys.* **71** 1515
- [25] Mauri F, Galli G and Car R 1993 *Phys. Rev. B* **47** 9973
- [26] Kim J, Mauri F and Galli G 1995 *Phys. Rev. B* **52** 1640
- [27] Stechel E B, Williams A P and Feibelman P J 1993 *Phys. Rev. B* **49** 3898
- [28] Galli G and Parrinello M 1992 *Phys. Rev. Lett.* **69** 3547
- [29] Reboredo F A and Williamson A 2005 *Phys. Rev. B* **71** 121105
- [30] Stewart J 1964 *Int. J. Quantum Chem.* **58** 133
- [31] Akama T, Fujii A, Kobayashi M and Nakai H 2007 *Mol. Phys.* **103** 2799
- [32] Kobayashi M, Imamura Y and Nakai H 2007 *J. Chem. Phys.* **127** 074103
- [33] Hu H, Lu Z, Elstner M, Hermans J and Yang W 2007 *J. Phys. Chem. A* **111** 5685
- [34] Nakano A, Kalia R, Nomura K, Sharma A, Vashishta P, Shimojo F, van Duin A, Goddard W, Biswas R and Srivastava D 2007 *Comput. Mater. Sci.* **38** 642
- [35] Anikin N, Anisimov V, Bugaenko V, Bobrikov V V and Andreyev A M 2004 *J. Chem. Phys.* **121** 1266
- [36] Goedecker S 1999 *Rev. Mod. Phys.* **71** 1085
- [37] Yang W and Pérez-Jordá J M 1998 *Encyclopedia of Computational Chemistry* ed P Schleyer (New York: Wiley) pp 1496–513
- [38] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in C* (Cambridge: Cambridge University Press)
- [39] Frisch M J *et al* 2003 *Gaussian 03* (Pittsburgh, PA: Gaussian)
- [40] Saad Y 1994 *Sparsekit: A Basic Tool Kit for Sparse Matrix Computations* Computer Science Department, University of Minnesota