Generalized symmetric Rayleigh–Ritz procedure applied to the closed-shell Hartree–Fock problem

Harold H. Wadleigh III
Department of Mathematics, University of California, Los Angeles, California 90095

Irina V. Ionova and Emily A. Carter[a)]
Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095

(Received 28 August 1998; accepted 1 December 1998)

We present the Generalized Symmetric Rayleigh–Ritz (GSRR) procedure for finding approximate eigenfunctions and corresponding eigenvalues for a linear operator, L, in a finite function space, $\{\phi_i\}_{i=1}^N$. GSRR is derived by minimizing the residual in the norm induced by an inner product, $(\cdot,\cdot)$, under the constraint that the resulting eigenfunctions be mutually orthogonal with respect to another inner product, $(\cdot,\cdot)_a$. When L is the closed-shell Fock operator, f, GSRR is a generalization of the Roothaan equations. We apply this method to f with $(\cdot,\cdot)$ defined by a grid, $\{r_k\}_{k=1}^M$, and $(\cdot,\cdot)_a$ defined by analytic integration, noting that a grid-defined $(\cdot,\cdot)$ lends itself to faster operator evaluation (scaling as MN^2) and effective parallelization. When a grid is used, GSRR scales as pseudospectral methods do; however, it is in the spirit of conventional spectral methods (e.g., GSRR does not use an inverse transform). © *1999 American Institute of Physics.* [S0021-9606(99)30909-0]

I. INTRODUCTION

Variational methods for approximating eigenfunctions, $\{\psi_i\}$, of a linear operator, L, have been very successful for many problems. When L is Hermitian with respect to an inner product, $(\cdot,\cdot)$, distinct eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to $(\cdot,\cdot)$. Moreover, when $(\cdot,\cdot)$ is used in the Rayleigh–Ritz procedure[1] in the case that L is Hermitian, the finite function space (finite basis) approximations have the same property.

For instance, the (linearized, closed-shell) Fock operator, f, is Hermitian with respect to integration over $\mathbb{R}^3$, and so its eigenfunctions are pair-wise orthogonal with respect to integration, and the Rayleigh–Ritz method (applied to the Fock operator, the method gives rise to the Roothaan equations[2]) generates approximate eigenfunctions which are exactly pair-wise orthogonal. Hence, the method is widely used and forms the basis for much of computational quantum chemistry. However, integration can be expensive and it imposes restrictions on the sorts of finite function spaces (basis sets) that might be used for approximation—namely, those amenable to integration.

Thole[3] explored the possibility of performing spectral-type calculations with a grid, beyond simple numerical integration, by ''desymmetrizing'' the problem. ''Dissymmetric'' methods use distinct ''solution'' ($\{\psi\}$) and ''test'' ($\{\chi\}$) spaces. The solution space is that spanned by the basis chosen for the problem. The test space is that with respect to which the solution is a ''weak'' solution. Informally, for an eigenvalue problem, an exact solution will satisfy

$$L\psi=\lambda\psi, \tag{1}$$

but a weak solution need only satisfy

$$(L\psi,\chi)=\lambda(\psi,\chi), \quad \forall\chi. \tag{2}$$

A standard spectral (Rayleigh–Ritz) calculation is simply one which uses the same function space for both solution and test spaces to find a weak solution. (For a Hermitian operator with a complete solution and test space, it can easily be shown that the weak solution is the exact solution.) In Thole's work, he chose $\{\chi\}$ to be made up of weighted Dirac δ-functions, so that it is representable by a grid. Thole chose his solution space, $\{\psi\}$, to be spanned by standard Gaussian basis sets. Thole was not thoroughly encouraged by the results of his tests. This may have been due to his using very coarse grids (at the most a few hundred grid points) combined with the fact that he was comparing results to standard spectral calculations.

Work by Friesner[4–7] on pseudospectral methods took a different approach in a similar spirit, which also resulted in a dissymmetric method that essentially uses a test space of weighted δ-functions to find solutions representable in standard Gaussian basis sets. Pseudospectral methods achieve improvements in computational efficiency by calculating the action of the operator on a spatial (grid) representation of basis functions, then ''back-transforming'' into a ''spectral'' representation (the matrix of coefficients of basis functions).

More generally, work by Sawada *et al.*[8] explored ideas of error minimization in a time-dependent framework.

We introduce the Generalized Symmetric Rayleigh–Ritz (GSRR) procedure which yields approximations to eigenfunctions of L using an arbitrary inner product, $(\cdot,\cdot)$, for projection onto the basis, but maintains the pair-wise orthogonality of the eigenfunctions with respect to another inner product, $(\cdot,\cdot)_a$ (presumably the one with respect to

a)Electronic mail: eac@chem.ucla.edu

which L is Hermitian). Thus, we are free to choose $(\cdot,\cdot)$ to be what is most convenient.

One of the drawbacks of both Thole's method and the pseudospectral method is that preserving the orthogonality of solutions is treated in an *ad hoc* or *a posteriori* manner. In contrast, GSRR preserves symmetry (and therefore orthogonality) in a natural way by keeping the test and solution spaces the same—as in the conventional Rayleigh–Ritz approach—and imposing an orthogonality constraint from the beginning. Doing so simplifies the method of calculation considerably over that of Thole's method. In contrast to pseudospectral methods, GSRR results in a system which is symmetric, so does not need to be symmetrized, as pseudospectral methods do.

Further, GSRR, while sharing similar scaling properties, does not require the use of a generalized inverse collocation transform, as Friesner's approach does. The construction of such a transform is not unique, and, while the nonuniqueness is not an issue when back-transforming the spatial representation of the basis functions, it may have a significant effect when back-transforming the result of the operator's action on the spatial representation.

In order to anticipate the presentation of the implementation of GSRR for f, we will conceptually consider $(\cdot,\cdot)$ to be a ''numerical'' or ''grid-based'' inner product and $(\cdot,\cdot)_a$ to be the ''analytic'' inner product, corresponding to unweighted integration.

II. THE GENERALIZED SYMMETRIC RAYLEIGH–RITZ PROCEDURE

GSRR is derived by minimizing the total residual of approximate eigenfunction/eigenvalue pairs. That is, given $\vec{\psi}$, a vector of approximate eigenfunctions, and $\vec{\lambda}$, the vector of corresponding approximate eigenvalues, we consider the functional

$$I[\vec{\psi},\vec{\lambda}]=\sum_i (L\psi_i-\lambda_i\psi_i, L\psi_i-\lambda_i\psi_i). \tag{3}$$

To this we join the orthogonality constraint:

$$J[\vec{\psi}]=\sum_{i,j} \mu_{ij}[(\psi_i,\psi_j)_a-\delta_{ij}], \tag{4}$$

where the μ_{ij} are Lagrange multipliers.

Functional differentiation (introducing an arbitrary ''direction'' of perturbation, χ_i, for each ψ_i) yields that the stationary points of $I-J$ satisfy (see Appendix A)

$$(L\psi_i-\lambda_i\psi_i, L\chi_i-\lambda_i\chi_i)=\sum_k \mu_{ki}(\psi_k,\chi_i)_a, \tag{5}$$

for all i and χ_i.

A. Finite space approximation

We introduce a finite basis, $\{\phi^i\}_{i=1}^N$, in which we will express the approximate eigenfunctions, $\{\psi_i\}_{i=1}^N$, and which we will use to span the test space. ψ_i has the form

$$\psi_i=\sum_j c_i^j\phi^j, \tag{6}$$

so that the eigenfunctions can be represented by the matrix $\mathbf{C}$, where

$$\mathbf{C}_{ij}=c_i^j. \tag{7}$$

We find that (see Appendix B) the stationary point corresponds to the eigensystem of the matrix $\mathbf{Y}$ where the eigenvectors, $\mathbf{N}$ of $\mathbf{Y}$ are related to $\mathbf{C}$ by

$$\mathbf{C}=\mathbf{G}\mathbf{N}^T, \tag{8}$$

and the eigenvalues of $\mathbf{Y}$, appearing in the diagonal matrix $\mathbf{\Lambda}$, are precisely the approximate eigenvalues of L.

$\mathbf{Y}$ is constructed in the following way: we calculate $\mathbf{P}$ and $\mathbf{S}$, the ''numerical'' and ''analytic'' overlap matrices:

$$\mathbf{P}_{ij}=(\phi^i,\phi^j), \tag{9}$$

$$\mathbf{S}_{ij}=(\phi^i,\phi^j)_a. \tag{10}$$

Then $\mathbf{H}$ (unitary) and $\mathbf{\Xi}$ (diagonal with elements ξ_i) are given by

$$\tilde{\mathbf{P}}\equiv\mathbf{S}^{-1/2}\mathbf{P}\mathbf{S}^{-1/2}=\mathbf{H}^T\mathbf{\Xi}\mathbf{H}. \tag{11}$$

This is possible since $\tilde{\mathbf{P}}$ is symmetric positive definite. We define

$$\mathbf{G}\equiv\mathbf{S}^{-1/2}\mathbf{H}^T, \tag{12}$$

$$\mathbf{Q}\equiv\mathbf{F}^T+\mathbf{F}, \tag{13}$$

where $\mathbf{F}$ has elements:

$$\mathbf{F}_{ij}=(L\phi^i,\phi^j). \tag{14}$$

We note that when L is the Fock operator $\mathbf{Q}$ is a ''symmetrized'' Fock matrix.

The elements of $\mathbf{Y}$ are given by

$$\mathbf{Y}_{ij}=\frac{\mathbf{G}_j^T\mathbf{Q}\mathbf{G}_i}{\xi_i+\xi_j}. \tag{15}$$

B. Generalization of the Rayleigh–Ritz procedure

We note that, when $(\cdot,\cdot)$ is the same as $(\cdot,\cdot)_a$, and L is Hermitian in $(\cdot,\cdot)$, GSRR reduces to the Rayleigh–Ritz procedure. In the case that the inner products are the same,

$$\mathbf{P}=\mathbf{S}, \tag{16}$$

so that $\tilde{\mathbf{P}}$ is the identity, and thus so is $\mathbf{\Xi}$, and $\mathbf{H}$ can be taken to be the identity as well. In this case,

$$\mathbf{G}=\mathbf{S}^{-1/2}. \tag{17}$$

And thus

$$\mathbf{Y}=\frac{1}{2}\mathbf{S}^{-1/2}\mathbf{Q}\mathbf{S}^{-1/2}. \tag{18}$$

However, since in this case L is symmetric with respect to $(\cdot,\cdot)$, we have that $\mathbf{F}=\mathbf{F}^T$, so that $\mathbf{Q}=2\mathbf{F}$, and thus

$$\mathbf{Y}=\mathbf{S}^{-1/2}\mathbf{F}\mathbf{S}^{-1/2}, \tag{19}$$

the eigensystem of which satisfies

$$\mathbf{F}\mathbf{C}=\mathbf{S}\mathbf{C}\mathbf{\Lambda}, \tag{20}$$

which is the system of equations given by the Rayleigh–Ritz procedure.

C. Grid inner product

Among the more interesting inner products which might be used in GSRR are those defined by a (weighted) grid, $\{r_i\}_{i=1}^{M}$, with corresponding weights $\{w_i\}_{k=i}^{M}$. That is,

$$(g,h)=\sum_i w_i g(r_i)h(r_i). \tag{21}$$

We define $\mathbf{R}$ and $\mathbf{\Omega}$ as

$$\mathbf{R}_{ij}=\phi^j(r_i), \tag{22}$$

$$\mathbf{\Omega}_{ij}=L\phi^j(r_i), \tag{23}$$

so that it is easy to see that in this case

$$\mathbf{P}=\mathbf{R}^T\mathbf{w}\mathbf{R}, \tag{24}$$

$$\mathbf{F}=\mathbf{\Omega}^T\mathbf{w}\mathbf{R}, \tag{25}$$

where $\mathbf{w}$ is diagonal with $\mathbf{w}_{ii}=w_i$.

The bilinear form as defined by (21) satisfies all the properties of an inner product for an arbitrary grid, except possibly positivity. That is,

$$(x,x)=0\rightarrow x=0. \tag{26}$$

And, of course, a grid inner product cannot satisfy (26) on an infinite-dimensional function space or even a finite function space if its dimension exceeds the number of grid points.

When $M\geqslant N$ it is still possible that (26) is not satisfied—precisely when the grid suffers from aliasing problems. That is, the grid cannot ''tell the difference'' between some pair of distinct functions in the space.

In order to avoid the question of aliasing in this present work, we restrict ourselves for the most part to relatively accurate grids which approximate integration—which are less likely to suffer from aliasing problems. The issue of aliasing as it relates to GSRR will be examined in further work.

III. RESULTS

In order to examine GSRR we apply it to the Fock operator, f. When f is linearized (i.e., an initial guess is chosen in order to evaluate the operator), GSRR may be applied within a self-consistent field (SCF) cycle (see Appendix C).

A. Implementation

We implement GSRR to calculate the eigensystem of the linearized f using a grid-based inner product. That is, we form $\mathbf{F}$ as in (25), where (23) becomes

$$\mathbf{\Omega}_{ij}=f\phi^j(r_i)=-\frac{1}{2}\nabla^2\phi^j(r_i)+V_{nuc}\phi^j(r_i)+2J\phi^j(r_i)-K\phi^j(r_i). \tag{27}$$

In our finite basis the Coulomb term becomes

$$J\phi^j(r_i)=\left(\sum_{\alpha=1}^{N_{occ}}\int\frac{|\psi_\alpha(r')|^2}{r_i-r'}dr'\right)\phi^j(r_i) \tag{28}$$

$$=\left(\sum_\alpha\int\frac{\Sigma_{m=1}^N c_\alpha^m\phi^m(r')\Sigma_{n=1}^N c_\alpha^n\phi^n(r')}{r_i-r'}dr'\right)\times\phi^j(r_i) \tag{29}$$

$$=\left(\sum_{\alpha mn}c_\alpha^m c_\alpha^n\int\frac{\phi^m(r')\phi^n(r')}{r_i-r'}dr'\right)\phi^j(r_i) \tag{30}$$

$$=\left(\sum_{mn}\mathbf{D}_{mn}\mathbf{A}_{mn}(r_i)\right)\phi^j(r_i), \tag{31}$$

where

$$\mathbf{D}_{mn}\equiv\sum_{\alpha=1}^{N_{occ}}c_\alpha^m c_\alpha^n, \tag{32}$$

$$\mathbf{A}_{mn}(r_i)\equiv\int\frac{\phi^m(r')\phi^n(r')}{r_i-r'}dr', \tag{33}$$

so that

$$J\phi^j(r_i)=[\mathbf{JR}]_{ij}, \tag{34}$$

with J diagonal and having elements given by

$$\mathbf{J}_{ii}=\mathrm{tr}(\mathbf{DA}(r_i)). \tag{35}$$

The elements of $\mathbf{A}$ are evaluated analytically at each of the grid points, r_i.

For the exchange term we have

$$K\phi^j(r_i)=\sum_{\alpha=1}^{N_{occ}}\psi_\alpha(r_i)\int\frac{\phi^j(r')\psi_\alpha(r')}{r_i-r'}dr' \tag{36}$$

$$=\sum_\alpha\sum_{m=1}^{N}c_\alpha^m\phi^m(r_i)\int\frac{\phi^j(r')\Sigma_{n=1}^N c_\alpha^n\phi^n(r')}{r_i-r'}dr' \tag{37}$$

$$=\sum_{\alpha mn}c_\alpha^m c_\alpha^n\phi^m(r_i)\int\frac{\phi^j(r')\phi^n(r')}{r_i-r'}dr' \tag{38}$$

$$=\sum_{mn}\mathbf{D}_{mn}\mathbf{A}_{jn}(r_i)\mathbf{R}_{im} \tag{39}$$

$$=\sum_n\mathbf{A}_{jn}(r_i)[\mathbf{RD}]_{in} \tag{40}$$

$$=[\mathbf{RDA}(r_i)]_{ij}. \tag{41}$$

An advantage of basing an inner product on a grid is that we may evaluate the elements of $\mathbf{F}$ in a parallel way—we distribute grid points across processors to evaluate the elements of $\mathbf{A}(\vec{r})$ and perform a global sum to evaluate $(\cdot,\cdot)$. Our code is written to take advantage of this sort of parallelization.

B. Accuracy

We examine the accuracy of GSRR when using grids designed to approximate integration. Since in this case the

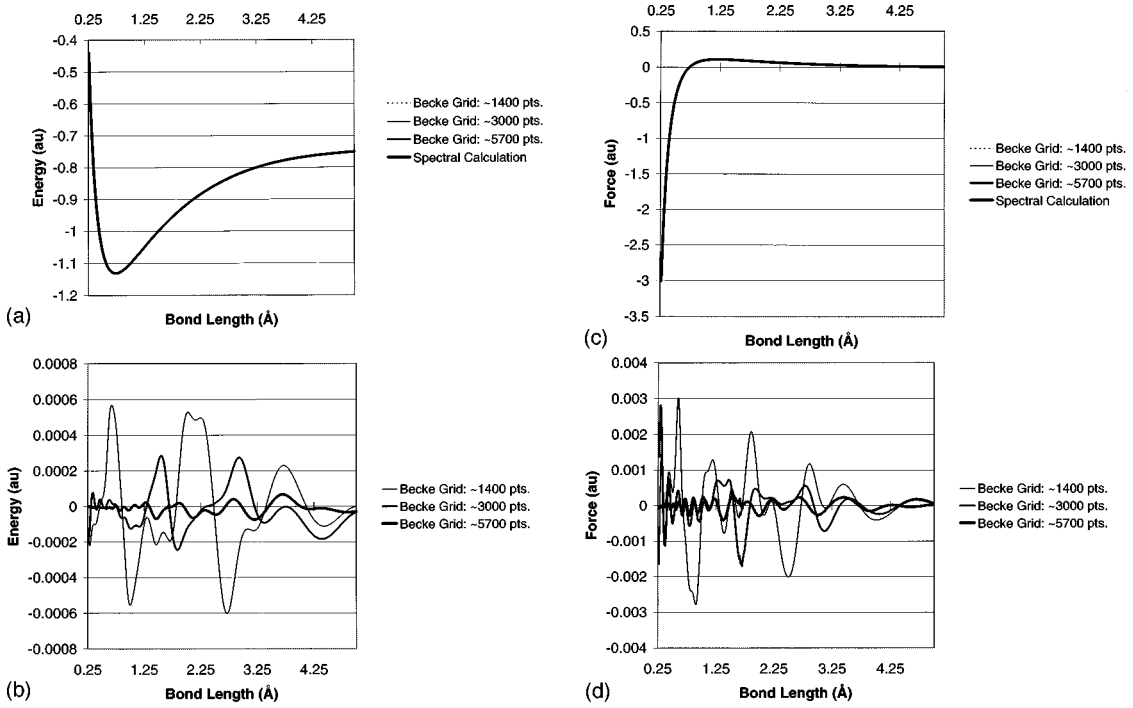

FIG. 1. GSRR calculations for H_2 using Becke-type grids. (a) Total energy; (b) energy deviation; (c) force; and (d) force deviation, in a.u., as a function of the bond length.

GSRR results ought to follow those of a conventional spectral calculation, we compare the results of GSRR with the spectral solution.

1. Hydrogen molecule potential

Our first test was to calculate total energies for H_2 for bond lengths between 0.25 Å and 5.0 Å. We used the 6-31G** basis (10 basis functions) and calculated the energies using both GSRR and the spectral method. For GSRR, we used three Becke-type[9] grids of varying coarseness which are constructed to approximate integration.

The results shown in Figs. 1(a) and 1(b) suggest that GSRR does, in fact, generalize the spectral method. Figure 1(a) shows no noticeable deviation for each of the grids, while Fig. 1(b) shows that none of the results for any of the GSRR calculations deviates from the spectral calculation by any more than 0.0006 a.u.

The deviations shown in Fig. 1(b) have a periodic character which can be explained by the method of constructing the Becke grids. A Becke grid is constructed by locating radial grids at atom sites. These radial grids are joined together in the final grid by truncating them at the surfaces of the Voronoi polyhedra defined by the atom centers and modifying the grid weights to compensate. The radial grids are more dense near the atom centers. As the H–H bond is stretched from shorter bond lengths, radial grid points will appear more rapidly in the final Becke grid than when the bond is longer and the faces of the Voronoi polyhedra are farther from the atom centers where the radial grid is less dense. This gives rise to the oscillatory behavior of increasing period in the results.

For another measure of accuracy, we calculated (using central differences on the energy data) the force as a function of bond length. There are two reasons why comparing forces is a useful measure of accuracy: forces are insensitive to global energy shifts, which are generally unimportant, and the measure of accuracy is local.

Figure 1(c) again indicates that the GSRR results are close to those of the spectral calculation, and Fig. 1(d) shows no deviation from the spectral calculation of more than 0.003 a.u.

2. GSRR calculations with pseudospectral grids

We ran the same series of calculations using grids constructed for use in a pseudospectral calculation. Figures 2(a) through 2(d) show that the grids are prone to give unpredictable results, although they all seem to give stable results in the neighborhood of the equilibrium bond length. We note, however, that in this case the spectral calculation is not a good point of comparison, since the inner product given by the grid is not meant to approximate integration. The only true measure of accuracy for results using nonintegrating grids (and, in fact, any grid) is an exact (complete basis) spectral calculation.

Pseudospectral grids are often constructed with a set of dealiasing functions to improve results, which allows pseudospectral methods to use coarser grids. The grids used were very coarse, so the results indicate that if a grid of similar coarseness is to be used to obtain reasonable results with GSRR, that grid will have to be constructed carefully.

Grid issues in general, and aliasing issues in particular, will be considered in future work.

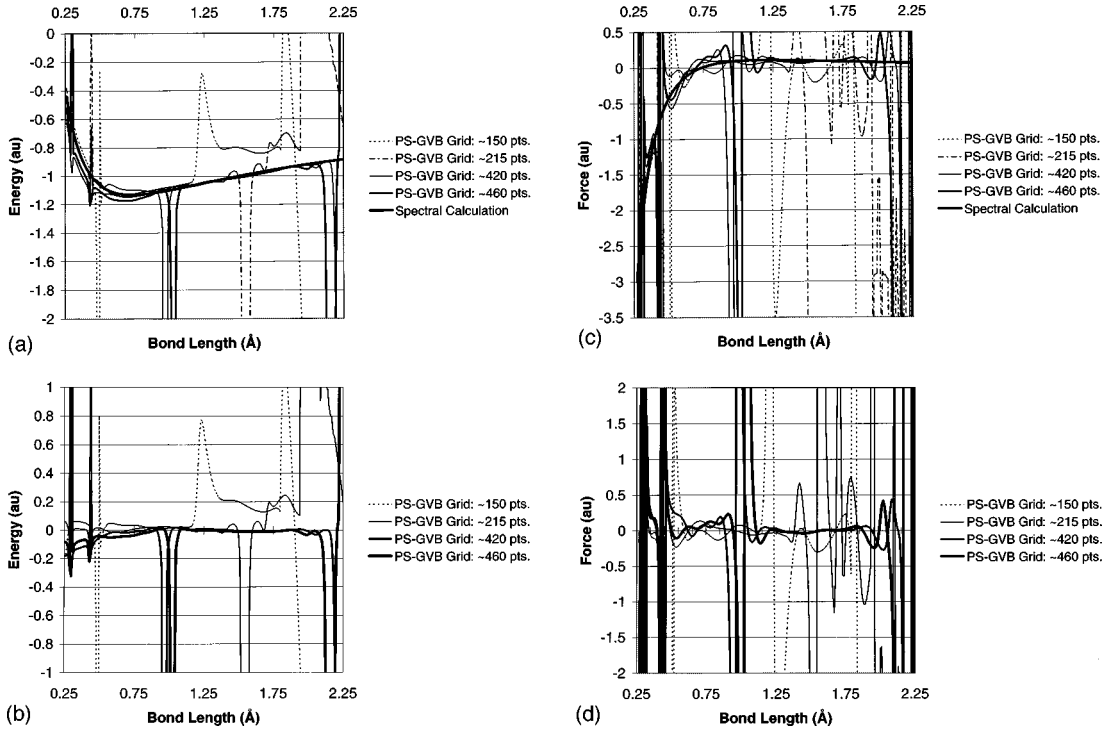

FIG. 2. GSRR calculations for H_2 using pseudospectral grids. (a) Total energy; (b) energy deviation; (c) force; and (d) force deviation, in a.u., as a function of the bond length.

3. Potential energy surfaces for larger molecules

In order to examine how GSRR performs on larger molecules, we performed calculations for ethylene and 1,1,2,2-tetrafluoroethane. For both molecules, 7 geometries were selected by first finding the spectral Hartree–Fock equilibrium geometry, which was selected as one of the geometries. The others were chosen by stretching the C–C bond length while leaving the relative positions of the remaining atoms fixed. For each molecule, we used the 6-31G** basis (50 and 100 basis functions, respectively) and three Becke grids.

Figure 3(a) shows that the energy curves for ethylene are close to one another. Figure 3(b) shows that, while the coarse

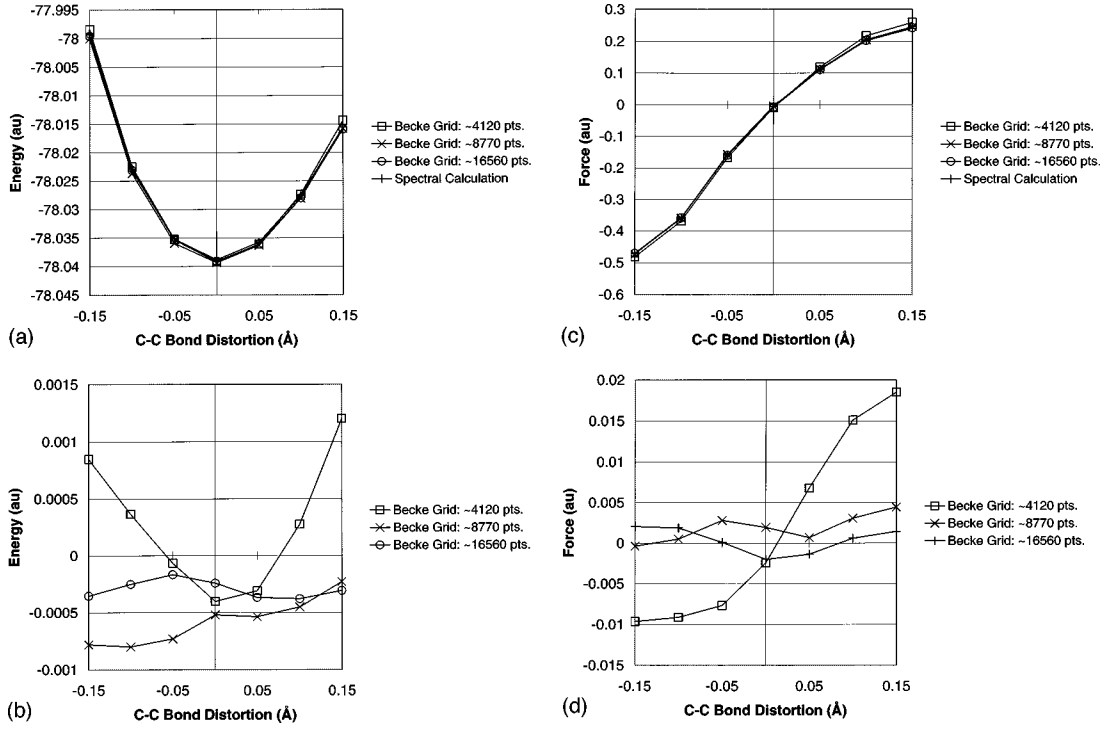

FIG. 3. GSRR calculations for ethylene. (a) Total energy; (b) energy deviation; (c) force; and (d) force deviation, in a.u., as a function of C–C bond distortion.

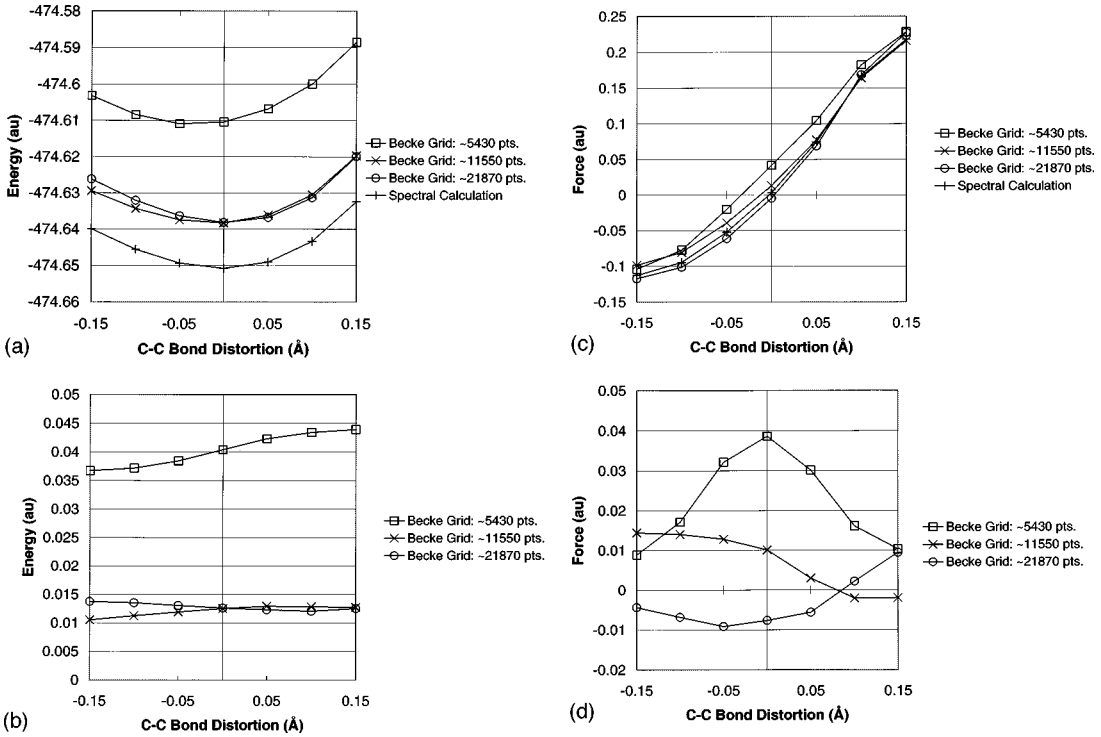

FIG. 4. GSRR calculations for 1,1,2,2-tetrafluoroethane. (a) Total energy; (b) energy deviation; (c) force; and (d) force deviation, in a.u., as a function of C–C bond distortion.

grid varies more dramatically relative to the other two calculations, the calculations for the medium and the fine grids are comparable, with the fine grid being generally more accurate. Figures 3(c) and 3(d) show that this also holds true for the estimation of error based on calculated forces.

Figure 4(a) shows the GSRR calculations for 1,1,2,2,-tetrafluoroethane varying significantly from the spectral calculation in terms of total energies. However, since the curves in Fig. 4(b) are very nearly flat, the results are still comparable to those of the spectral calculation modulo an energy shift. The force calculations, as shown in Figs. 4(c) and 4(d), show a much closer agreement. For instance, the force calculation for the finest grid deviates from that of the spectral calculation by no more than 0.01 a.u.

4. Further tests of GSRR

To test how GSRR would respond to a more widely varying set of geometries, we performed calculations on water and again on ethylene. In both cases, we selected 49 geometries with all atoms fixed except for one hydrogen atom, which in each of the geometries occupied different positions in the plane of the molecule. Again, one of the geometries corresponded to the spectral Hartree–Fock equilibrium geometry. In the case of water, the atom centers were located in the *y*-*z* plane, with the plane of symmetry passing through the *z*-axis and the hydrogen atoms lying in the upper half-plane. The hydrogen atom lying in the left half-plane was chosen to be perturbed. For ethylene, the *z*-axis passed through the C–C bond with the molecule lying in the *y*-*z* plane. The hydrogen chosen for perturbation was that lying in the first quadrant of the *y*-*z* plane.

Figure 5(a) shows the potential surface for water given by the spectral method, while Figs. 5(b), 5(c), and 5(d) show how the energies (and forces) calculated with GSRR deviate from the spectral results. The variations of the GSRR calculations get smaller as the grid becomes more refined, although there seems to be a fixed deviation of about -0.007 a.u. (see Table I). We thus note that the GSRR results will not always be upper bounds to the spectral results (which does not mean that they are not upper bounds to the exact energies). The force magnitude error results [Figs. 5(e) through 5(g), Table I] again show a decrease in total magnitude as the grid is refined. Comparable behavior is seen in the results for ethylene [see Figs. 6(a) through 6(g), Table I].

We also note, as we can observe through all of the calculations, that the results improve noticeably between the coarse and medium grids, but not so much between the medium and fine grids, suggesting that the calculations with the medium grids will give accurate results while achieving significant savings.

C. Performance

In order to have a practical gauge of how our implementation scales with the number of basis functions, grid points, and processors, we used timings obtained by systematically varying parameters.

1. Serial performance

We first consider how the method scales with respect to basis functions, grid points, and iterations (since the solution of the Roothaan equations requires an iteration process).

We chose a serial timing model with the following form:

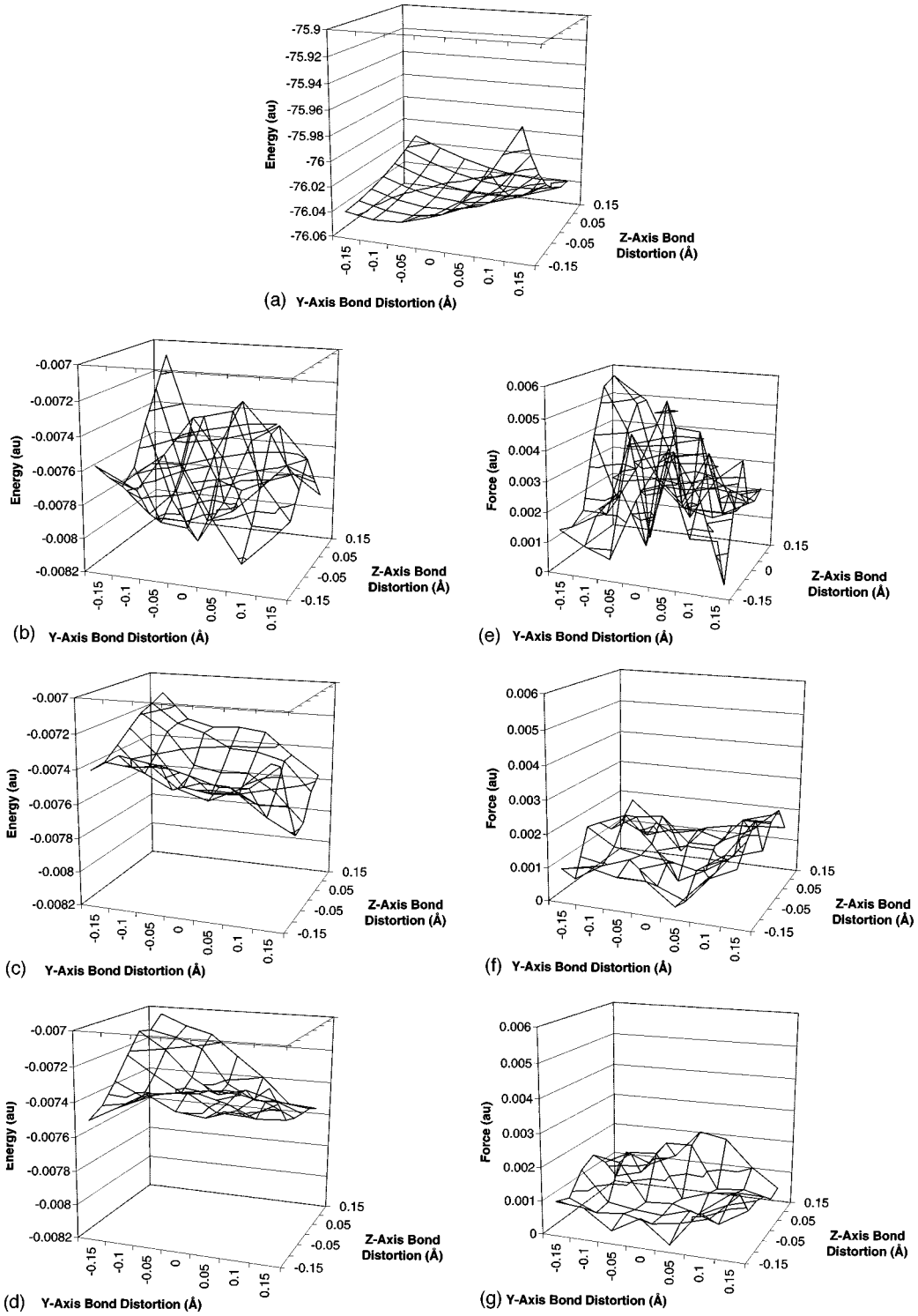

FIG. 5. GSRR calculations for water. (a) Total energy (spectral calculation); (b) energy deviation (~2090 point Becke grid); (c) energy deviation (~4440 point Becke grid); (d) energy deviation (~8400 point Becke grid); (e) force magnitude deviation (~2090 point Becke grid); (f) force magnitude deviation (~4400 point Becke grid); and (g) force magnitude deviation (~8400 point Becke grid), in a.u., as a function of O–H bond distortion.

$$T = C(\alpha + (1-\alpha)I)N^{\beta}M^{\delta}, \tag{42}$$

where T is the time of the calculation, C is a prefactor, I is the number of iterations taken. α is the fraction of the non-SCF code (that which is executed outside the SCF cycle), and β and δ are the powers of scaling for the number of basis functions, N, and grid points, M, respectively.

Serial timing data was collected by running calculations on water, ethylene, and 1,1,2,2-tetrafluoroethane with the same basis sets as in the previous calculations. For each of these, we used six Becke grids, and ran each of these calculations for 1, 2, 4, 8, and 16 iterations. The results for the least-squares fit to this model are shown in Table II. We note that β and δ are very nearly 2 and 1, respectively, as in the

TABLE I. Summary of water and ethylene calculations.

		Energy deviations (a.u.)			Force deviations (a.u.)	
	Grid	Maximum	Average	Total	Maximum	Average
Water	Coarse	8.195×10^{-3}	-7.701×10^{-3}	1.101×10^{-3}	5.764×10^{-3}	2.656×10^{-3}
	Medium	7.816×10^{-3}	-7.442×10^{-3}	6.937×10^{-4}	2.573×10^{-3}	1.236×10^{-3}
	Fine	7.568×10^{-3}	-7.368×10^{-3}	5.242×10^{-4}	2.015×10^{-3}	8.894×10^{-4}
Ethylene	Coarse	2.902×10^{-3}	-1.323×10^{-3}	3.011×10^{-3}	9.801×10^{-3}	5.882×10^{-3}
	Medium	9.016×10^{-4}	-3.980×10^{-4}	9.893×10^{-4}	4.151×10^{-3}	1.778×10^{-3}
	Fine	3.303×10^{-4}	-1.641×10^{-4}	3.438×10^{-4}	2.157×10^{-3}	9.815×10^{-4}

pseudospectral methods. When these are constrained to be exactly 2 and 1, only the prefactor changes significantly. We also note that most of the calculation ($>95\%$ in the test cases) takes place before the iteration cycle. This is consistent with the expectation that most of the effort occurs in calculating the one-electron integrals at each of the grid points [the elements of $\mathbf{A}(\vec{r})$], which happens before the SCF cycle.

2. Parallel performance

In order to measure the parallelization of our implementation of GSRR, we ran our code on an IBM SP2, using 1, 8, 16, and 24 processors. The calculation used was that of 1,1,2,2-tetrafluoroethane with the finest Becke grid, iterating to convergence.

Two measures of time were used. The first was the maximum CPU time (CT) for each of the processes. The second was a more accurate measure, which consisted of measuring CPU time for most of the calculation, but measuring real (wall) time during sections of the code where communication occurs. (We will refer to this measure of time as PT.)

The first model we considered has the form

$$T = C\left((1-\alpha) + \frac{\alpha}{P} \right), \tag{43}$$

where C is the prefactor, P is the number of processors, and α is the fraction of code which is parallelized. The fits using both CT and PT measures of time are shown in Table III for this first model. We note that the results indicate that more than 98% of the code is parallelized. This means, in this model, that if infinitely many processors were used, the run time would be less than 2% of the one-processor run time.

We also considered a second model:

$$T = CP^{\alpha}, \tag{44}$$

the results for which are shown Table III, as well. We see that the estimated scaling factor is approximately -0.94, compared to -1 for perfect parallelization.

Figure 7 shows the parallelization efficiency plotted against the number of processors. Parallelization efficiency is defined as follows:

$$E \equiv \frac{T_1}{nT_n}, \tag{45}$$

where E is the efficiency and T_n is the time taken for a calculation with n processors. Figure 7 shows that efficiency does not deteriorate significantly as the number of processors increases.

The code is not completely parallelized, since the matrix operations (e.g., the eigenproblem that is solved in the SCF cycle) are not parallelized. In fact, the parallelization is exploited only in evaluating $(\cdot,\cdot)$, but since the inner product evaluations are what give rise to the leading term in the scaling for the method, efficient parallelization of evaluations of $(\cdot,\cdot)$ is, for practical purposes, all that is necessary. That our implementation was measured to be 98% parallelized supports this claim, and we do not expect that parallelization of the matrix operations will result in a significant improvement in parallelization.

IV. CONCLUSIONS AND FUTURE DIRECTIONS

The GSRR procedure is a symmetric generalization of the Rayleigh–Ritz procedure (i.e., by minimizing the residual of approximate eigenfunction/eigenvalue pairs), allowing one to choose $(\cdot,\cdot)$ freely while maintaining the orthogonality of the eigenfunctions. Thus GSRR provides an alternative to solution of the Roothaan equations using numerical integration or pseudospectral or other grid-based methods.

In our implementation, we noticed that Becke-type grids of medium coarseness produced results comparable to those of the Roothaan equations. We also note that in our implementation the method scales like pseudospectral methods, i.e., as $O(MN^2)$, and is well parallelized.

GSRR applied to f can potentially be made into an easily parallelized $O(N)$ method.[10–33] One condition for $O(N)$ scaling to be achieved is that GSRR use a basis of compactly supported functions (the support having a measure bounded by a constant). This is approximately true for standard atom-centered basis sets, and can be made exactly true with the appropriate use of cutoffs. If the density of grid points and the number of basis functions per atom is bounded, then $O(M)=O(N)$ and $\mathbf{R}$ and $\mathbf{A}(\vec{r})$ have special properties. $\mathbf{R}$ has $O(N)$ elements with the critical property that the number of elements in any column or row is bounded by a constant independent of N, which means that multiplications involving $\mathbf{R}$ involve the same order of operations as those involving diagonal matrices, i.e., $O(N)$ operations. $\mathbf{A}(\vec{r})$ has an even better property: for each grid point i, the number of nonzero elements of $\mathbf{A}(r_i)$ is bounded by a constant indepen-

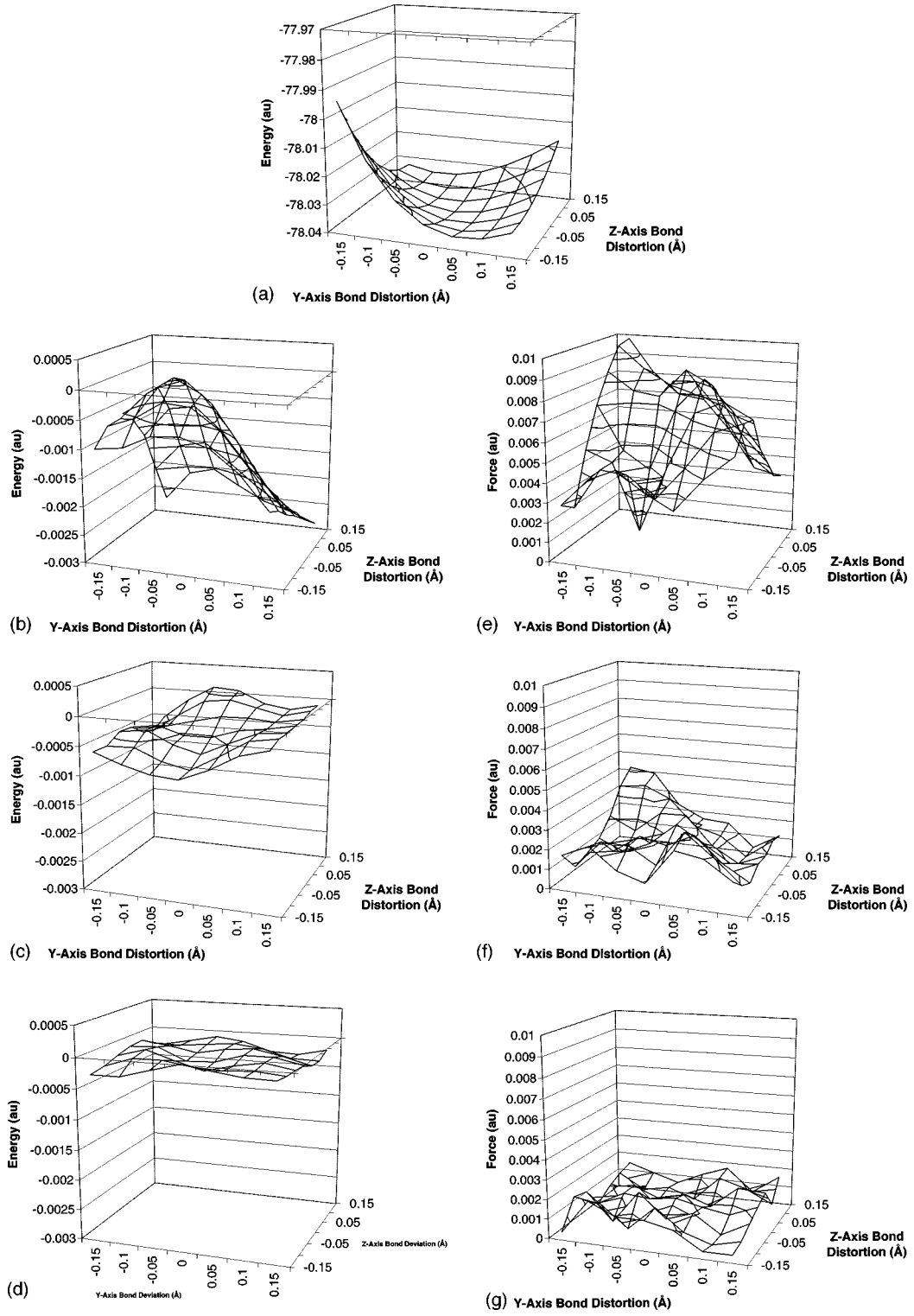

FIG. 6. GSRR calculations for ethylene. (a) Total energy (spectral calculation); (b) energy deviation (~4110 point Becke grid); (c) energy deviation (~8760 point Becke grid); (d) energy deviation (~16560 point Becke grid); (e) force magnitude deviation (~4110 point Becke grid); (f) force magnitude deviation (~8760 point Becke grid); and (g) force magnitude deviation (~16560 point Becke grid), in a.u., as a function of C–H bond distortion.

dent of N. These properties, along with a $O(N)$-sparse density matrix, **D**, allow for the operator evaluations in (35) and (41) to occur in $O(N)$ time. **D** is known to be sparse for systems with a band gap, making such problems a potential candidate for this performance-enhanced version of GSRR. We also note that GSRR is $O(MN^2)$ because of the con-

TABLE II. Serial model parameter fits.

	Unconstrained	Constrained
C (prefactor)	9.72×10^{-6}	7.53×10^{-6}
β (basis function scaling)	1.951	2.000
δ (grid point scaling)	0.998	1.000
α (non-SCF Fraction)	0.956	0.955

TABLE III. Parallelization model parameter fits.

		CT	PT
Model 1	C (prefactor)	1435.093	1441.691
	α (fraction parallelized)	0.985	0.985
Model 2	C (prefactor)	1435.255	1441.870
	α (processor scaling)	−0.943	−0.944

struction of $\mathbf{A}(\vec{r})$, which, as we have seen, takes up approximately 95% of the calculation time. Given that the per-atom numbers of basis functions and grid points are bounded, $\mathbf{A}(\vec{r})$ can be constructed in $O(N)$ time without requiring the sparsity of $\mathbf{D}$, making for drastic savings in the full range of problems—effectively becoming a $O(N)$ method for a much wider range and size of systems.

The extension to density functional theory (DFT) is trivial, since it can be accomplished simply by replacing the exchange operator with an exchange-correlation potential. This will be pursued in further work.

No work has yet been done on calculating efficient grids for this method in either a static or adaptive matter, and such work is a natural next step.

ACKNOWLEDGMENTS

We thank Professor Dennis Salahub for providing the DeMon DFT code used to generate the Becke-type grids used in this work. Partial support from the Office of Naval Research and the National Science Foundation is gratefully acknowledged.

APPENDIX A: STATIONARY POINT CONDITION

We seek the minimum stationary point of

$$I[\vec{\psi},\vec{\lambda}]=\sum_i I_i - J, \tag{A1}$$

where

$$I_i=(L\psi_i-\lambda_i\psi_i, L\psi_i-\lambda_i\psi_i), \tag{A2}$$

$$J=\sum_{i,j}\mu_{ij}[(\psi_i,\psi_j)_a-\delta_{ij}]. \tag{A3}$$

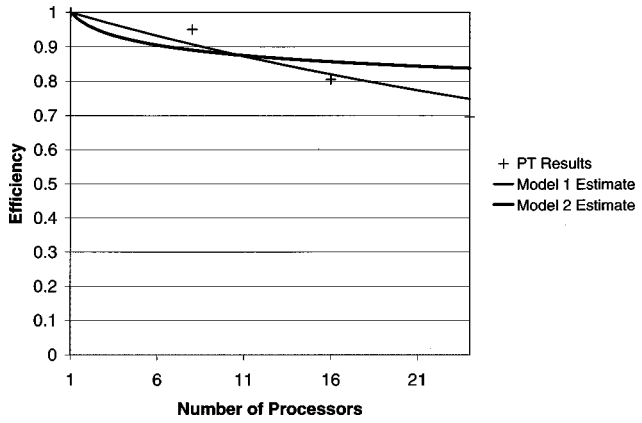

FIG. 7. Parallelization efficiency versus processors.

We note that the symmetry of δ_{ij} and $(\cdot,\cdot)_a$ allows us to choose $\mu_{ij}=\mu_{ji}$.

For the moment we neglect J and consider $I^*\equiv\Sigma_i I_i$. Fix a $\vec{\psi}$. We note that the stationary points of I^* with respect to $\vec{\lambda}$ satisfy

$$\frac{\partial I^*}{\partial\lambda_i}=\frac{\partial I_i}{\partial\lambda_i}=\frac{\partial}{\partial\lambda_i}[(L\psi_i,L\psi_i)-2\lambda_i(L\psi_i,\psi_i)+\lambda_i^2(\psi_i,\psi_i)]$$
$$=-2(L\psi_i,\psi_i)+2\lambda_i(\psi_i,\psi_i) \tag{A4}$$

and

$$\frac{\partial^2 I^*}{\partial\lambda_i^2}=2(\psi_i,\psi_i)>0, \tag{A5}$$

using the positivity of the inner product. This implies that the stationary point, $\vec{\lambda}$, is unique and a minimum for fixed $\vec{\psi}$, and the stationary point is given by

$$\frac{\partial I^*}{\partial\lambda_i}=0\Rightarrow\lambda_i=\frac{(L\psi_i,\psi_i)}{(\psi_i,\psi_i)}, \tag{A6}$$

which is the Rayleigh quotient.

We now substitute this expression for λ_i into I_i, so that

$$I_i=(L\psi_i,L\psi_i)-\frac{(L\psi_i,\psi_i)^2}{(\psi_i,\psi_i)}. \tag{A7}$$

We drop the subscript and consider

$$I'\equiv I_1-\frac{I_2^2}{I_3}, \tag{A8}$$

where

$$I_1\equiv(L\psi,L\psi), \tag{A9}$$

$$I_2\equiv(L\psi,\psi), \tag{A10}$$

$$I_3\equiv(\psi,\psi). \tag{A11}$$

We define the functional derivative:

$$F_\epsilon[f]\equiv\frac{\partial F}{\partial\epsilon}[f+\epsilon g]\bigg|_{\epsilon=0}, \tag{A12}$$

for an arbitrary g. It is easy to see that the functional derivative satisfies the usual rules of differentiation. Thus, we easily compute

$$I'_\epsilon=I_{1_\epsilon}-2\frac{I_2 I_{2_\epsilon}}{I_3}+\frac{I_2^2 I_{3_\epsilon}}{I_3^2}, \tag{A13}$$

$$I'_{\epsilon\epsilon}=I_{1_{\epsilon\epsilon}}-2\frac{I_2 I_{2_{\epsilon\epsilon}}}{I_3}+\frac{I_2^2 I_{3_{\epsilon\epsilon}}}{I_3^2}-\frac{2}{I_3}\left[I_{2_\epsilon}-\frac{I_2 I_{3_\epsilon}}{I_3}\right]^2. \tag{A14}$$

Noting that

$$\lambda=\frac{I_2}{I_3}, \tag{A15}$$

we have, after substituting,

$$I'_\epsilon=I_{1_\epsilon}-2\lambda I_{2_\epsilon}+\lambda^2 I_{3_\epsilon}, \tag{A16}$$

$$I'_{\epsilon\epsilon}=I_{1_{\epsilon\epsilon}}-2\lambda I_{2_{\epsilon\epsilon}}+\lambda^2 I_{3_{\epsilon\epsilon}}-\frac{2}{I_3}[I_{2_\epsilon}-\lambda I_{3_\epsilon}]^2. \tag{A17}$$

We easily compute the functional derivatives of I_1, I_2, and I_3:

$$I_{1_\epsilon}=2(L\psi,L\chi), \tag{A18}$$

$$I_{1_{\epsilon\epsilon}}=2(L\chi,L\chi), \tag{A19}$$

$$I_{2_\epsilon}=(L\psi,\chi)+(L\chi,\psi), \tag{A20}$$

$$I_{2_{\epsilon\epsilon}}=2(L\chi,\chi), \tag{A21}$$

$$I_{3_\epsilon}=2(\psi,\chi), \tag{A22}$$

$$I_{3_{\epsilon\epsilon}}=2(\chi,\chi), \tag{A23}$$

for arbitrary χ. Upon substituting and simplifying, we have

$$I'_\epsilon=2(L\psi-\lambda\psi,L\chi-\lambda\chi), \tag{A24}$$

$$I'_{\epsilon\epsilon}=2(L\chi-\lambda\chi,L\chi-\lambda\chi)-\frac{2}{(\psi,\psi)}[(L\psi-\lambda\psi,\chi)+(L\chi-\lambda\chi,\psi)]^2. \tag{A25}$$

We now consider J. We introduce the perturbation

$$\psi_i\rightarrow\psi_i+\epsilon_i\chi_i, \tag{A26}$$

where χ_i ranges over a space of test functions. Thus

$$J\rightarrow\sum_{i,j}\mu_{ij}[(\psi_i,\psi_j)_a+\epsilon_i(\chi_i,\psi_j)_a+\epsilon_j(\chi_j,\psi_i)_a+\epsilon_i\epsilon_j(\chi_i,\chi_j)_a-\delta_{ij}], \tag{A27}$$

so

$$\frac{\partial J}{\partial\epsilon_k}=\frac{\partial}{\partial\epsilon_k}\left[\begin{array}{l}\displaystyle\sum_j\mu_{kj}\left[\begin{array}{l}(\psi_k,\psi_j)_a+\epsilon_k(\chi_k,\psi_j)_a+\epsilon_j(\chi_j,\psi_k)_a\\ \quad+\epsilon_j\epsilon_k(\chi_k,\chi_j)_a-\delta_{jk}\end{array}\right]\\ \displaystyle+\sum_i\mu_{ik}\left[\begin{array}{l}(\psi_i,\psi_k)_a+\epsilon_i(\chi_i,\psi_k)_a+\epsilon_k(\chi_k,\psi_i)_a\\ \quad+\epsilon_i\epsilon_k(\chi_i,\chi_k)_a-\delta_{ik}\end{array}\right]\end{array}\right]$$

$$=2\sum_i\mu_{ik}\frac{\partial}{\partial\epsilon_k}\left[\begin{array}{l}(\psi_i,\psi_k)_a+\epsilon_i(\chi_i,\psi_k)_a+\epsilon_k(\chi_k,\psi_i)_a\\ \quad+\epsilon_i\epsilon_k(\chi_i,\chi_k)_a-\delta_{ik}\end{array}\right]$$

$$=2\sum_i\mu_{ik}[(\chi_k,\psi_i)_a+\epsilon_i(\chi_i,\chi_k)_a], \tag{A28}$$

and so

$$\left.\frac{\partial J}{\partial\epsilon_i}\right|_{\epsilon_i=0}=\frac{\partial J}{\partial\psi_i}=2\sum_k\mu_{ik}(\chi_i,\psi_k)_a. \tag{A29}$$

We also note that

$$\frac{\partial^2 J}{\partial\epsilon_i^2}=0. \tag{A30}$$

The stationary point satisfies, by (A24) and (A29),

$$(L\psi_i-\lambda_i\psi_i,L\chi_i-\lambda_i\chi_i)=\sum_k\mu_{ki}(\psi_k,\chi_i)_a, \tag{A31}$$

for all i and χ_i. (A25) and (A30) show that the stationary point is a minimum when (A25) is positive for all χ.

APPENDIX B: FINITE SPACE SOLUTIONS

We introduce a finite function space, $\{\phi^j\}$, which we use for both test and solution spaces. The elements of our problem become

$$\begin{aligned}&(L\psi_i,L\chi)\rightarrow\sum_j c_i^j(L\phi^j,L\phi^k),\\ &(L\psi_i,\chi)\rightarrow\sum_j c_i^j(L\phi^j,\phi^k),\\ &(L\chi,\psi_i)\rightarrow\sum_j c_i^j(L\phi^k,\phi^j),\\ &(\psi_i,\chi)\rightarrow\sum_j c_i^j(\phi^j,\phi^k),\\ &(\psi_i,\chi)_a\rightarrow\sum_j c_i^j(\phi^j,\phi^k)_a,\end{aligned} \tag{B1}$$

$\forall k$, so that, for each i,

$$\mathbf{AC}_i-\lambda_i(\mathbf{F}+\mathbf{F}^T)\mathbf{C}_i+\lambda_i^2\mathbf{PC}_i=\mathbf{SCM}_i \tag{B2}$$

which we note is just the matrix version of (A31), with

$$\begin{aligned}&\mathbf{A}_{jk}=(L\phi^j,L\phi^k),\\ &\mathbf{F}_{jk}=(L\phi^j,\phi^k),\\ &\mathbf{P}_{jk}=(\phi^j,\phi^k),\\ &\mathbf{M}_{jk}=\mu_{jk},\\ &\mathbf{S}_{jk}=(\phi^j,\phi^k)_a,\\ &\mathbf{C}_{jk}=c_k^j.\end{aligned} \tag{B3}$$

We also note that the orthonormality condition is equivalent to

$$\mathbf{C}^T\mathbf{SC}=\mathbf{I}. \tag{B4}$$

If $\mathbf{C}$ is square and nonsingular (i.e., we calculate the entire eigensystem), then (B2) is equivalent to

$$\mathbf{C}_j^T\mathbf{AC}_i-\lambda_i\mathbf{C}_j^T(\mathbf{F}+\mathbf{F}^T)\mathbf{C}_i+\lambda_i^2\mathbf{C}_j^T\mathbf{PC}_i=\mathbf{C}_j^T\mathbf{SCM}_i=\mu_{ij}, \tag{B5}$$

$\forall i,j$, where, in the second equality, we use (B4). Switching the roles of i and j and using the symmetry of $\mathbf{M}$, we have

$$\mathbf{C}_j^T\mathbf{A}\mathbf{C}_i-\lambda_i\mathbf{C}_j^T(\mathbf{F}+\mathbf{F}^T)\mathbf{C}_i+\lambda_i^2\mathbf{C}_j^T\mathbf{P}\mathbf{C}_i$$
$$=\mathbf{C}_j^T\mathbf{A}\mathbf{C}_i-\lambda_j\mathbf{C}_j^T(\mathbf{F}+\mathbf{F}^T)\mathbf{C}_i+\lambda_j^2\mathbf{C}_j^T\mathbf{P}\mathbf{C}_i \tag{B6}$$

$$\Updownarrow$$

$$(\lambda_i-\lambda_j)\mathbf{C}_j^T(\mathbf{F}+\mathbf{F}^T)\mathbf{C}_i=(\lambda_i^2-\lambda_j^2)\mathbf{C}_j^T\mathbf{P}\mathbf{C}_i, \tag{B7}$$

$$\Updownarrow$$

$$\mathbf{C}_j^T(\mathbf{F}+\mathbf{F}^T)\mathbf{C}_i=(\lambda_i+\lambda_j)\mathbf{C}_j^T\mathbf{P}\mathbf{C}_i\,. \tag{B8}$$

We note that (B8) is not exactly equivalent to (B7), but is a slightly stronger statement, with equivalence if and only if $\lambda_i\neq\lambda_j$. The effect of degeneracies will need to be examined in further work. As a matrix equation, (B8) is

$$\mathbf{C}^T\mathbf{Q}\mathbf{C}=\mathbf{\Lambda}\mathbf{C}^T\mathbf{P}\mathbf{C}+\mathbf{C}^T\mathbf{P}\mathbf{C}\mathbf{\Lambda}, \tag{B9}$$

where $\mathbf{\Lambda}$ is diagonal with elements:

$$\mathbf{\Lambda}_{ii}=\lambda_i \tag{B10}$$

and

$$\mathbf{Q}\equiv\mathbf{F}+\mathbf{F}^T. \tag{B11}$$

We define

$$\tilde{\mathbf{C}}\equiv\mathbf{S}^{1/2}\mathbf{C}, \tag{B12}$$

so that

$$\tilde{\mathbf{C}}^T\tilde{\mathbf{C}}=\mathbf{I}. \tag{B13}$$

With (B12), (B9) becomes

$$\tilde{\mathbf{C}}^T\tilde{\mathbf{Q}}\tilde{\mathbf{C}}=\mathbf{\Lambda}\tilde{\mathbf{C}}^T\tilde{\mathbf{P}}\tilde{\mathbf{C}}+\tilde{\mathbf{C}}^T\tilde{\mathbf{P}}\tilde{\mathbf{C}}\mathbf{\Lambda}, \tag{B14}$$

where

$$\tilde{\mathbf{Q}}\equiv\mathbf{S}^{-1/2}\mathbf{Q}\mathbf{S}^{-1/2}, \tag{B15}$$

$$\tilde{\mathbf{P}}\equiv\mathbf{S}^{-1/2}\mathbf{P}\mathbf{S}^{-1/2}. \tag{B16}$$

We used the symmetry of $\mathbf{S}$ in deriving (B14). Pre- and post-multiplying (B14) by $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{C}}^T$, respectively, and using (B13), we have

$$\tilde{\mathbf{Q}}=\mathbf{X}\tilde{\mathbf{P}}+\tilde{\mathbf{P}}\mathbf{X}, \tag{B17}$$

where

$$\mathbf{X}\equiv\tilde{\mathbf{C}}\mathbf{\Lambda}\tilde{\mathbf{C}}^T. \tag{B18}$$

We note that $\tilde{\mathbf{P}}$ is positive definite, so

$$\tilde{\mathbf{P}}=\mathbf{H}^T\mathbf{\Xi}\mathbf{H}, \tag{B19}$$

where $\mathbf{H}$ is unitary and $\mathbf{\Xi}$ is diagonal, having elements:

$$\mathbf{\Xi}_{ii}=\xi_i\,. \tag{B20}$$

We define

$$\tilde{\tilde{\mathbf{Q}}}\equiv\mathbf{H}\tilde{\mathbf{Q}}\mathbf{H}^T, \tag{B21}$$

so that, from (B17),

$$\tilde{\tilde{\mathbf{Q}}}=\mathbf{Y}\mathbf{\Xi}+\mathbf{\Xi}\mathbf{Y}, \tag{B22}$$

with

$$\mathbf{Y}\equiv\mathbf{H}\mathbf{X}\mathbf{H}^T, \tag{B23}$$

but then

$$\mathbf{Y}=\mathbf{H}\mathbf{X}\mathbf{H}^T \tag{B24}$$

$$=(\mathbf{H}\tilde{\mathbf{C}})\mathbf{\Lambda}(\tilde{\mathbf{C}}^T\mathbf{H}^T) \tag{B25}$$

$$=(\mathbf{H}\mathbf{S}^{1/2}\mathbf{C})\mathbf{\Lambda}(\mathbf{C}^T\mathbf{S}^{1/2}\mathbf{H}^T), \tag{B26}$$

so we see that the eigenvalues of $\mathbf{Y}$ are the elements of $\mathbf{\Lambda}$, and the matrix of corresponding eigenvectors, $\mathbf{N}$, is related to $\mathbf{C}$ by

$$\mathbf{C}=\mathbf{S}^{-1/2}\mathbf{H}^T\mathbf{N}^T. \tag{B27}$$

We define

$$\mathbf{G}\equiv\mathbf{S}^{-1/2}\mathbf{H}^T, \tag{B28}$$

and note that $\mathbf{G}$ and $\mathbf{\Xi}$ satisfy the generalized eigenvalue problem:

$$\mathbf{P}\mathbf{G}=\mathbf{S}\mathbf{G}\mathbf{\Xi}. \tag{B29}$$

As (B15), (B22), and (B28) show, $\mathbf{Y}$ can be constructed by

$$\mathbf{Y}_{ij}=\frac{\mathbf{G}_j^T\mathbf{Q}\mathbf{G}_i}{\xi_i+\xi_j}. \tag{B30}$$

APPENDIX C: THE SELF-CONSISTENT FIELD CYCLE

Applying GSRR to finding the eigenfunctions of f requires the linearization of f, and therefore the necessary introduction of an iterative cycle.

Once a finite function subspace (a basis) is chosen, $\mathbf{S}$ can be calculated, along with the kinetic energy and nuclear attraction terms.

Once the inner product (defined by a grid in our implementation) is chosen, $\mathbf{R}$ can be evaluated, along with $\mathbf{P}$. $\mathbf{G}$ can then be calculated from $\mathbf{P}$ and $\mathbf{S}$. The elements of $\mathbf{A}(\vec{r})$ can also be calculated before the self-consistent field (SCF) cycle.

The simple SCF cycle is as follows:

(1) Given $\mathbf{C}$ and $\mathbf{A}(\vec{r})$, calculate $\mathbf{Q}$.
(2) Form $\mathbf{Y}$, and calculate its eigensystem: $\mathbf{N}$ and $\mathbf{\Lambda}$.
(3) Extract $\mathbf{C}$ from $\mathbf{N}$.
(4) Repeat to convergence.

In practice, however, convergence will generally only be achieved by modifying $\mathbf{C}$ in step (3) via damping and extrapolation. In our implementation we use a DIIS method.[34]

[1] R. Courant and D. Hilbert, *Methods of Mathematical Physics (Wiley, New York, 1989), Vol. 1, Chap. 4, pp. 174–176.*
[2] C. C. J. Roothaan, ''New developments in molecular orbital theory,'' Rev. Mod. Phys. **23**, 69 (1951).
[3] B. T. Thole, ''Least-squares numerical Rayleigh-Ritz and minimum-variance methods for molecular calculations,'' Int. J. Quantum Chem. **XXVIII**, 535 (1985).
[4] R. A. Friesner, ''Solution of self-consistent field electronic structure equations by a pseudospectral method,'' Chem. Phys. Lett. **116**, 39 (1985).
[5] R. A. Friesner, ''Solution of the Hartree-Fock equations by a pseudospectral method: Application to diatomic molecules,'' J. Chem. Phys. **85**, 1462 (1986).
[6] R. A. Friesner, ''Solution of the Hartree-Fock equations for polyatomic molecules by a pseudospectral method,'' J. Chem. Phys. **86**, 3522 (1987).
[7] R. A. Friesner, ''An automatic grid generation scheme for pseudospectral

self-consistent field calculations on polyatomic molecules,'' J. Phys. Chem. **88**, 3091 (1988).
[8] S-I. Sawada, R. Heather, B. Jackson, and H. Metiu, ''A strategy for time dependent quantum mechanical calculations using a Gaussian wave packet representation of the wave function,'' J. Chem. Phys. **83**, 3009 (1985).
[9] A. D. Becke, ''Multicenter numerical integration scheme for polyatomic molecules,'' J. Chem. Phys. **88**, 2547 (1988).
[10] W. T. Yang, ''Direct calculation of electron density in density-functional theory,'' Phys. Rev. Lett. **66**, 1438 (1991).
[11] W. T. Yang, ''Analytical energy gradients and geometry optimization in the divide-and-conquer method for large molecules,'' J. Chem. Phys. **102**, 9598 (1995).
[12] W. T. Yang, ''A local projection method for the linear combination of atomic orbital implementation of density-functional theory,'' J. Chem. Phys. **94**, 1208 (1991).
[13] M. Parrinello, ''Large scale electronic structure calculations,'' Phys. Rev. Lett. **69**, 3547 (1992).
[14] G. Galli, ''Electronic-structure calculations and molecular-dynamics simulations with linear system-size scaling,'' Phys. Rev. B **50**, 4316 (1994).
[15] G. Galli and R. Car, ''Orbital formulation for electronic-structure calculations with linear system-size scaling,'' Phys. Rev. B **47**, 9973 (1993).
[16] D. A. Drabold, R. M. Martin, and M. P. Grumbac, ''Linear system-size scaling methods for electronic-structure calculations,'' Phys. Rev. B **51**, 1456 (1995).
[17] D. A. Drabold, M. P. Grumbach, and R. M. Martin, ''Unconstrained minimization approach for electronic computations that scales linearly with system size,'' Phys. Rev. B **48**, 14646 (1993).
[18] F. Mauri and G. Galli, ''Total-energy global optimizations using nonorthogonal localized orbitals,'' Phys. Rev. B **52**, 1640 (1995).
[19] E. B. Stechel, ''Order-*N* methods in self-consistent density-functional calculations,'' Phys. Rev. B **50**, 17811 (1994).
[20] A. R. Williams and P. J. Feibelman, ''*N*-scaling algorithm for density-functional calculations of metals and insulators,'' Phys. Rev. B **49**, 10088 (1994).
[21] M. J. Gillan, ''Self-consistent first-principles technique with linear scaling,'' Phys. Rev. B **51**, 10157 (1995).
[22] W. T. Yang, ''Absolute-energy-minimum principles for linear-scaling electronic-structure calculations,'' Phys. Rev. B **56**, 9294 (1997).
[23] L. Colombo, ''Efficient linear scaling algorithm for tight-binding molecular dynamics,'' Phys. Rev. Lett. **73**, 122 (1994).
[24] M. Teter, ''Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals,'' Phys. Rev. B **51**, 9455 (1995).
[25] M. Head-Gordon, ''Chebyshev expansion methods for electronic structure calculations on large molecular systems,'' J. Chem. Phys. **107**, 10003 (1997).
[26] D. A. Drabold, ''Order-*N* projection method for first-principles computations of electronic quantities and Wannier functions,'' Phys. Rev. B **57**, 6391 (1998).
[27] J. D. Kress and R. N. Silver, ''Linear-scaling tight binding from a truncated-moment approach,'' Phys. Rev. B **53**, 12 733 (1996).
[28] H. Roeder, A. F. Voter, and J. D. Kress, ''Kernel polynomial approximations for densities of states and spectral functions,'' J. Comput. Phys. **124**, 115 (1996).
[29] W. Nunes and D. Vanderbilt, ''Density-matrix electronic-structure method with linear system-size scaling,'' Phys. Rev. B **47**, 10891 (1993).
[30] M. S. Daw, ''Model for energetics of solids based on the density matrix,'' Phys. Rev. B **47**, 10895 (1993).
[31] D. Vanderbilt, ''Generalization of the density-matrix method to a nonorthogonal basis,'' Phys. Rev. B **50**, 17611 (1994).
[32] C. Z. Wang, K. M. Ho, and C. T. Chan, ''Tight-binding molecular dynamics with linear system-size scaling,'' J. Phys.: Condens. Matter **6**, 9153 (1994).
[33] W. Kohn, ''Density functional and density matrix method scaling linearly with the number of atoms,'' Phys. Rev. Lett. **76**, 3168 (1996).
[34] I. V. Ionova and E. A. Carter, ''Orbital-based direct inversion in the iterative subspace for the generalized valence bond method,'' J. Chem. Phys. **102**, 1251 (1995).