Efficient multiscale algorithms for solution of self-consistent eigenvalue problems in real space

Nimal R. Wijesekera,[1,2] Guogang Feng,[1] and Thomas L. Beck[1,2,*]
[1]*Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221-0172, USA*
[2]*Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221-0011, USA*
(Received 17 July 2006; revised manuscript received 12 December 2006; published 2 March 2007)

Real-space multiscale methods provide efficient algorithms for large-scale electronic structure calculations. In this paper, we present multigrid strategies for solving self-consistent problems in density functional theory. The full approximation scheme (FAS) formulation of the multigrid method allows for transfer of the expensive orthogonalization and Ritz projection operations to coarse levels. In addition, the effective potential may be updated on coarse levels during multiscale processing of the eigenfunctions. We investigate modifications of a previously proposed algorithm which are necessary to yield robust convergence rates. With these modifications, rapid convergence is observed without orthonormalization or Ritz projection for the full occupied subspace on the fine level. Calculations comparing the various algorithms are performed on three many-electron examples: benzene, benzenedithiol, and the amino acid glycine. The modified algorithm is also illustrated on several larger test cases. Recently developed relativistic separable dual-space Gaussian pseudopotentials are utilized to remove the core electrons.

DOI: 10.1103/PhysRevB.75.115101 PACS number(s): 71.15.−m

I. INTRODUCTION

The solution of large-scale electronic structure problems is a computationally demanding task. In recent years, a great deal of effort has been devoted to the development of efficient numerical methods for tackling these problems. The intense interest is stimulated by challenging applications arising in the realistic modeling of systems encountered in physics, chemistry, biophysics, and the emerging nanoscale sciences. The main existing approaches for large-scale *ab initio* calculations can be loosely categorized as plane-wave basis set,[1] localized (Wannier) function[2–5] or Gaussian basis set,[6] and real-space methods.[7–14] One appeal of the Wannier-function and/or real-space methods is the localized nature of the representation. This property has led to the development of methods whose complexity scales linearly with the number of electrons.[4,5,15–17]

In real-space calculations, the discretized partial differential equations are represented on grids using either finite-element[8,18–20] or finite-difference[7,9,21] formulations. The solutions can be obtained using iterative techniques due to the highly banded nature of the Hamiltonian. On single grids, however, stalling is a major problem for the iterative methods. The stalling results from the long-wavelength modes of the errors, which are not efficiently removed on the fine scale.

Multigrid methods overcome the stalling problem by decimating errors on multiple length scales. The short-wavelength errors are smoothed on fine levels, while the long-wavelength components are removed on coarse levels during the multigrid cycles. In the past, a number of real-space multigrid solvers have been developed to solve the Kohn-Sham equations of density functional theory.[3,21–26]

Previously we implemented a nonlinear full approximation scheme [27–29] (FAS) multigrid method[30] in solving the self-consistent Kohn-Sham equations.[23] Reference 23 showed that utilization of the FAS approach leads to an improvement in convergence efficiency compared with linearized multigrid methods.[22,25] In the original eigenvalue method of Ref. 30, the expensive orthonormalization and Ritz projection operations are performed on the fine level.

Costiner and Ta'asan[31] advanced the FAS method by also moving the fine-scale separation techniques (orthonormalization and Ritz projection) to coarse levels. The FAS representation of the Ritz projection leads to a generalized eigenvalue problem on the coarse levels. They also introduced a back-rotation step designed to prevent rotations in degenerate subspaces, sign changes, rescalings, and permutations of the eigenvectors. The overall algorithm is termed generalized Ritz and back rotation (GRBR). The back-rotation step is necessary since the fine-scale functions are corrected following the GRBR process on the coarse grid, and therefore the coarse-level eigenfunctions must correspond to their fine-level counterparts.

In a second contribution,[32] Costiner and Ta'asan adapted their eigenvalue method to solve large-scale self-consistent problems. They considered two approaches for handling the self-consistency. First, the Poisson and eigenvalue problems were solved in a sequential fashion: once the effective potential was obtained on the finest level via solution of the Poisson equation (given a fixed updated charge density), the restricted potential was used during the multigrid solution for the eigenfunctions. Thus, the method cycles back and forth between the Poisson and eigenvalue problems much as in a traditional self-consistency code. Second, they updated the potential on coarse levels simultaneously with updates of the eigenfunctions. In this way, the eigenfunctions and potential evolve toward the self-consistent solution together on the coarse levels. An alternative method for updating the potential along with the eigenfunctions on coarse levels has appeared recently.[33]

Here we present work aimed at applying the methods of Refs. 31 and 32 to nonperiodic systems involving tens of wave functions. Relativistic separable dual-space Gaussian pseudopotentials[34] are utilized to remove the core electrons, leaving 15–21 eigenfunctions for our comparative test cases [benzene(15), glycine(15), and benzenedithiol(21)]. Further,

we illustrate our modified algorithms on larger systems: the glycine dimer(26), glutamine(29), phenylalanine(32), tyrosine(35), and the C_{20} molecules(40).

The Gram-Schmidt orthonormalization and Ritz projection steps on the fine level scale as $q^2N_g^l$, where q is the number of eigenfunctions and N_g^l is the total number of grid points. The finest level is labeled by l, while coarser levels are labeled by k (the smallest value of k is for the coarsest level). This scaling assumes the eigenfunctions span the whole physical domain. Moving these steps to coarser levels via the GRBR algorithm reduces the cost by a factor of 8 for each level. Even though the coarse-level generalized Ritz process scales as $q^2N_g^k$, Costiner and Ta'asan observed an effective qN_g^l algorithmic scaling, corresponding to the cost of updating the eigenfunctions on the fine level.

In our earlier paper,[24] we applied the GRBR algorithm to fixed-potential eigenvalue problems and self-consistent electronic structure problems. The fixed-potential and small test self-consistent problems possessed clearly defined eigenvalue cluster structures; the eigenvalue clusters consist of degenerate or near-degenerate subspaces of the full occupied subspace. Later in this paper we use the term cluster to denote any ordered and partitioned collection of eigenvalues. We showed that the GRBR method converged for the fixed-potential and small self-consistent problems such as the CO molecule and the Ne atom, all of which possess clearly defined eigenvalue clusters. The convergence rates slowed, however, after a few self-consistency iterations. The convergence rates could be restored by performing Gram-Schmidt orthonormalization occasionally on the fine level (once in every three to five self-consistency cycles). For larger systems such as benzenedithiol, which does not possess such clearly defined eigenvalue clusters, the GRBR algorithm in its original form stalled prior to full convergence. The method does converge (slowly) with periodic fine-scale orthonormalization for the benzene molecule with its clear eigenvalue cluster structure (unpublished results).

These results led us to conclude that, for large molecules having tens of states and ambiguous eigenvalue cluster structures, the GRBR procedure alone cannot bring sufficient separation of the wave functions on the fine level, and this results in stalling. The reasons for this are likely due to the relatively large number of wave functions, vague eigenvalue near-degeneracies (clusters), and the complexities of the potentials relative to those considered in Refs. 31 and 32.

In this paper, we describe an attempt to remedy these difficulties by performing Gram-Schmidt orthonormalization and Ritz projection on predetermined overlapping eigenvalue clusters on the fine level. These clusters need not necessarily correspond to degenerate subspaces. The fine-scale cluster Ritz process is performed along with GRBR on coarser grids. Since the Gram-Schmidt orthonormalization and Ritz projection are performed on eigenvalue-eigenfunction clusters, it is computationally less expensive than the regular Gram-Schmidt orthonormalization and Ritz projection performed on the full occupied subspace on the finest level. Further, we successfully test this method with the simultaneous multigrid technique, where the potential is updated on coarse levels.

The calculations were all performed at the Kohn-Sham local density approximation (LDA) level using a 12th-order finite-difference representation of the Laplacian. The method presented here differs from the alternative approach in Ref. 33 in several respects. First, with some modifications, we provide a direct implementation and test of the methods proposed in Refs. 31 and 32; in this approach, the FAS multigrid cycles are employed to update the eigenfunctions on the fine scale, not the charge density as in Ref. 33. In our method, the charge density is updated once the eigenfunctions are corrected. Second, we utilize a high-order finite-difference representation to estimate more accurately the kinetic energy operator, while in Ref. 33, a second-order form was employed. Finally, our work is aimed at reducing the $q^2N_g^l$ (or N_e^3, where N_e is the number of electrons) scaling of the algorithm. Wang *et al.* performed the orthonormalization and Ritz projection operations for the full occupied subspace on the fine scale, which maintains the N_e^3 scaling.

The structure of the paper is as follows. First, we describe the fine-level Ritz projection multigrid method.[30] Then the GRBR algorithm and the fine-level Ritz projection on preselected clusters are discussed. Next the simultaneous update of the potential on coarse levels is described. This is followed by discussions of the iterative relaxation scheme and the application of the relativistic separable dual-space Gaussian pseudopotentials. In the section concerning computational implementation, a detailed description of the technical details of the numerical calculations is given. The comparative convergence results for three molecular test cases are then presented, followed by results of calculations on larger chemical species. Finally, we discuss the algorithms and computational results and present our conclusions.

II. MULTIGRID CYCLE WITH FINE-LEVEL RITZ PROJECTION: FLR

Let us denote the fine-level eigenvalue problem with q wave functions by

$$H^l U^l = U^l \Lambda. \tag{1}$$

Here H^l is the fine-level Hamiltonian, which is a highly banded $N_g^l \times N_g^l$ square matrix. The matrix U^l is an $N_g^l \times q$ matrix, and its columns are the wave function values on the grid. The matrix Λ is a $q \times q$ diagonal matrix with the eigenvalues along the diagonal. Since Λ is the same on all levels at convergence, we do not attach a grid label to this matrix (see below). We denote the exact grid solution U^l and the current approximation u^l. We obtain an initial fine-scale approximation u^l via a full multigrid cycle as discussed in Sec. VII

Once an initial approximation is obtained on the finest level l and a few relaxation steps are performed there, the problem is passed to the next coarser level k. On the coarse level, the eigenvalue problem takes the form

$$H^k u^k = u^k \Lambda + \tau^k, \tag{2}$$

where τ^k is the defect correction defined as

$$\tau^k = I_{k+1}^k \tau^{k+1} + H^k (I_{k+1}^k u^{k+1}) - I_{k+1}^k (H^{k+1} u^{k+1}). \tag{3}$$

The operator I_{k+1}^k is the restriction operator, which involves a local trapezoid-rule average of the function. The defect cor-

rection is zero on the finest level. The first term on the right-hand side (RHS) in Eq. (3) is a contribution from the previous level's defect correction on grids at least two levels removed from the finest scale. If the exact solution from the fine grid is inserted into the coarse-grid eigenvalue equation, it is easily seen that an identity is obtained, where the restricted fine-grid eigenfunction solves the coarse-grid equation. This formulation thus satisfies the important condition of zero correction at convergence.

We also note that the same FAS formulation is used to solve the Poisson equation during the self-consistent solution. For the Poisson problem the Hamiltonian in Eq. (2) is replaced by the Laplacian operator, and the first term on the RHS of Eq. (2) is replaced by $-4\pi\rho$ where ρ is the charge density. The coarse-grid density is obtained by restricting the fine-grid density.

On the coarse level, iterative relaxation is performed subject to constraints designed to maintain eigenfunction orthonormality on the finest grid.[30] We also explored use of Kaczmarz relaxation on the coarse levels as in Ref. 31; no clear advantage was observed, however. A correction is then made back on the next higher level $k+1$ ($k+1=l$ for the finest level):

$$u^{k+1} = u^{k+1} + I_k^{k+1}(u^k - I_{k+1}^k u^{k+1}), \tag{4}$$

where I_k^{k+1} is the interpolation operator. Linear interpolation was used throughout our work. The sum of the orders of the transfer operations (restriction and interpolation) should equal the order of the differential equation (here second order), not necessarily the order of the Laplacian representation.[27,28]

Once back on the finest grid, the eigenvalues can be updated using the Rayleigh quotient:

$$\lambda = \langle u^l|H^l|u^l\rangle/\langle u^l|u^l\rangle. \tag{5}$$

Following relaxation steps on the fine level, Gram-Schmidt orthonormalization and Ritz projection are performed to improve the occupied subspace. The complete process of moving to a sequence of coarser levels followed by correction and relaxation steps on succeedingly finer levels (until the finest level is reached) is termed a V cycle. The Ritz projection step is discussed in the next section [see Eqs. (8) and (11) for the fine-scale version]. The above discussion can be extended to any number of levels. For the Poisson problem, the multigrid cycle is continued all the way to the coarsest level, which has only a single interior point. For the eigenvalue problem, the coarsest level must possess enough structure to (at least approximately) sample the oscillations in the eigenfunctions. For the eigenvalue problem, we utilized three grids levels (17^3, 33^3, and 65^3) in all of the calculations reported here.

III. GENERALIZED RITZ PROJECTION AND BACKROTATION: GRBR

The fine-scale eigenvalue problem can be rewritten using the relation

$$H^l v^l = v^l \Lambda, \tag{6}$$

where

$$v^l = u^l E^l. \tag{7}$$

Thus

$$H^l u^l E^l = u^l E^l \Lambda. \tag{8}$$

Here v^l and u^l are $N_g^l \times q$ matrices and E^l is a $q \times q$ matrix to be found. When Eq. (8) is transferred to a coarse level using an FAS transfer, it takes the form

$$H^k u^k E^k = u^k E^k \Lambda + \tau^k E^k, \tag{9}$$

where τ^k is the defect correction vector. Multiplying Eq. (9) on the left by $[u^k]^T$, the generalized eigenvalue problem results:

$$[u^k]^T (H^k u^k - \tau^k) E^k = ([u^k]^T u^k) E^k \Lambda. \tag{10}$$

This equation can be solved using standard linear algebra packages. The solution gives E^k and a new set of eigenvalues $\{\lambda_i\}$. New wave functions and defect corrections τ^k are obtained as linear combinations of the previous values:

$$v^k = u^{k,new} = u^{k,old} E^k, \tag{11}$$

$$\tau^{k,new} = \tau^{k,old} E^k. \tag{12}$$

On the fine grid where the defect correction τ^l is zero, the generalized Ritz projection reduces to the standard Ritz projection. Notice that the eigenvalues are the same on all levels at convergence by construction.

In the generalized Ritz projection method, rotations of solutions in subspaces of close eigenvalues, rescaling, sign changes, and permutations of the solutions may occur. Therefore the back-rotation algorithm was introduced by Costiner and Ta'asan[31] to deal with these problems. The back-rotation algorithm proceeds as follows. Once E^k and Λ are found by solving Eq. (10), the eigenvalues in Λ are sorted and the columns of the E^k matrix are permuted accordingly. The clusters of degenerate and nondegenerate clusters of Λ are identified. For each diagonal block of E^k (nondegenerate clusters) the dominant elements of the block are brought to the diagonal by permuting the columns of E^k and the diagonal of Λ. Define F^k as a block-diagonal matrix having only the diagonal blocks of E^k. Each nondegenerate diagonal block of F^k is replaced with the identity matrix. Set $E^k = E^k[F^k]^{-1}$ and change the signs of the columns of E^k such that the diagonal elements are all positive. The columns of E^k are normalized and the wave functions and the defect corrections are updated as per Eqs. (11) and (12). Extensive details of the back-rotation steps are given in Ref. 31.

GRBR performed on a coarse level is intended to bring separation of eigenstates on the fine level. However, this worked with limited success in our previous large-molecule test case (benzenedithiol).[24]

IV. FINE-SCALE RITZ PROJECTION PERFORMED ON CLUSTERS ALONG WITH COARSE-SCALE GRBR: CR-GRBR

As discussed above, the original GRBR algorithm stalled for large systems with ambiguous eigenvalue cluster struc-

tures. We found that incorporation of the following algorithm in addition to GRBR is a first step toward alleviating those difficulties. The general idea is to perform fine-scale Ritz projection on previously selected eigenvalue clusters in addition to the GRBR processing on coarse levels. Cluster structures are chosen at the beginning of the multigrid V cycles on the fine level, as described below.

Consider a system with q wave functions. The eigenvalues along with the wave functions are sorted in increasing order as $\lambda_0, \ldots, \lambda_{q-1}$, and we preselect m clusters of eigenvalues along with the wave functions. We denote the ith cluster of the eigenvalues by (a_i, b_i) where a_i and b_i are the first and last elements of the cluster, and i goes from 1 to m. Fine-level orthonormalization and Ritz projection are performed sequentially on the first cluster (a_1, b_1), followed by the rest of the clusters $(a_i - s, b_i)$ $(i=2, \ldots, m)$, where s is the number of overlap elements. The first cluster (a_1, b_1) has no overlap. The inclusion of the overlap elements (s) during the fine-level separation was found to be crucial. The number of required overlap elements depends on the system. For the test molecules benzene, glycine, and benzenedithiol, it was sufficient to choose $s=1$. For larger systems such as the C_{20} molecule, however, a larger number of overlap elements ($s=3$) was necessary. The coarse-level (middle level in all our calculations) GRBR process is performed using the clusters with the same overlap as on the fine scale. This addition to the algorithm of Costiner and Ta'asan[31,32] resulted in separation of the wave functions on the fine level sufficient to achieve full convergence for the cases where GRBR alone stalled. The preselected clusters are treated as degenerate clusters in the BR process.

This fine-level separation technique is relatively inexpensive compared to performing orthonormalization and Ritz projection on the full occupied subspace. If m clusters each consisting of p wave functions are chosen, the cost of orthonormalization and Ritz projection on clusters scales as $m(p+s)^2 N_g^l$. The Ritz projection method scales as $q^2 N_g^l$, where $q=mp$. Thus the ratio of the scalings of the CR-GRBR to the FLR algorithms is

$$R = \frac{1}{m}\left(\frac{p+s}{p}\right)^2. \tag{13}$$

Therefore, there is nearly an m-fold decrease in computational time on the fine grid for large systems. If smaller cluster sizes p are chosen, m will be larger and the savings of computational time will be higher. In our experience, however, it is found that, if the clusters are too small, the algorithm tends to stall. The size of the clusters can be gradually increased until a suitable size (typically $p=10$) is found. Using this approach, we observed convergence rates comparable to the full fine-level Ritz projection in our comparative test studies (benzene, glycine, and benzenedithiol). We further illustrate the new method by applying it to larger systems (glycine dimer, glutamine, tyrosine, phenylalanine, and the C_{20} molecule).

V. SIMULTANEOUS MULTIGRID METHOD: S-FLR AND S-CR-GRBR

Due to the nonlinear FAS formulation of the self-consistent problem, it is also possible to update the coarse-grid effective potential simultaneously with the eigenfunctions.[32,33] This contrasts with the sequential algorithm in which the fixed coarse-grid potential is restricted from the fine level and kept fixed during the multigrid cycle for the eigenfunctions.

The simultaneous update of the potential on the coarse grids requires some care. The density must be updated on the coarse grids, from which the coarse-level effective potential is obtained. A key advantage of the FAS approach is that the coarse-level density and potential terms can be constructed with defect correction terms such that, at convergence, the coarse-level values are precisely those restricted from the fine grid. The fine-grid eigenfunctions are normalized so as to preserve charge conservation throughout the various grids at convergence. This yields the optimal representation of the coarse-grid potential. In order to effect these points, new defect corrections are required for the charge density and the exchange-correlation potential.

We define $\bar{\rho}^k(u^k)$ for coarse level k as

$$\bar{\rho}^k(u^k) = 2\sum_{i=1}^{q} |u_i^k|^2. \tag{14}$$

It is important to note that $\bar{\rho}^k(u^k)$ is not a true charge density on coarse levels since, even at convergence, the restricted eigenfunctions are neither normalized nor orthogonal on coarse grids. Once the wave functions are updated on level k, the quantity $\bar{\rho}^k$ is generated using Eq. (14). We then construct the coarse-level density via

$$\rho^k = \bar{\rho}^k(u^k) + I_{k+1}^k \rho^{k+1} - \bar{\rho}^k(I_{k+1}^k u^{k+1}). \tag{15}$$

It is clear that, at convergence, the first and third terms cancel and we are left with the charge density restricted from the finer level. Therefore, charge conservation is maintained on all levels since the restriction operation preserves the total charge. The last two terms can be defined as a defect correction for ρ^k. A different method of updating the charge density on coarse levels can be found in Ref. 33. Once ρ^k is updated, the Coulomb potential can be found by solving the coarse-grid Poisson problem using again the FAS technique.[9] Using this FAS approach with the above-defined charge density, the coarse-level Coulomb potential is precisely the restricted fine-level potential at convergence.

On the coarse level k the exchange-correlation potential v_{xc}^k is updated as soon as ρ^k is obtained. We define a form for the coarse grid v_{xc}^k which also yields the restricted fine level v_{xc}^l at convergence:

$$v_{xc}^k = v_{xc}^k(\rho^k) + I_{k+1}^k v_{xc}^{k+1} - v_{xc}^k(I_{k+1}^k \rho^{k+1}). \tag{16}$$

The last two terms of Eq. (16) can be considered a defect correction for v_{xc}^k. For v_{xc} we have used either the Vosko-Wilk-Nussair (VWN) form[35] or the formula given in Ref. 36. The pseudopotentials (discussed below) do not depend on the valence charge density, so they do not require modification on coarse levels.

VI. RELAXATION WITH SHIFT PARAMETER

Consider a fine-level eigenvalue equation for a single state (where λ and u^l are the current approximations):

$$H^l u^l = \lambda u^l,$$

$$H^l = L^l + D^l + U^l. \tag{17}$$

Here the Hamiltonian matrix H^l is written using the lower-triangular part L^l, the diagonal part D^l, and the upper-triangular part U^l.

On any level k, Gauss-Seidel relaxation for the eigenvalue problem (with a shift parameter η^k) can then be written as[30]

$$u^{k,n+1} = u^{k,n} + (D^k - \eta^k I)^{-1}[\tau^k - (D^k u^{k,n} + L^k u^{k,n+1} + U^k u^{k,n} - \lambda I u^{k,n})], \tag{18}$$

where

$$\eta^k = \mu^k + \lambda + v^k_{Coul}(\mathbf{r}) + v^k_{xc}(\mathbf{r}). \tag{19}$$

Here μ^k is a parameter (units of energy), λ is the eigenvalue, and $v^k_{Coul}(\mathbf{r})$ includes the Coulomb potential from the nuclei and the electrons (no nonlocal parts of the pseudopotential). We found that addition of the eigenvalue and the potential led to smoother convergence. In the iteration equation $u^{k,n+1}$ is the updated k-level vector and $u^{k,n}$ is the old vector.

We also extended the iterative scheme to successive over-relaxation (SOR) with a shift parameter and found improved convergence rates in our case studies

$$u^{k,n+1} = u^{k,n} + (D^k - \eta^k I)^{-1}[\tau^k - (D^k u^{k,n} + L^k u^{k,n+1} + U^k u^{k,n} - \lambda I u^{k,n})]\omega^k, \tag{20}$$

where ω^k is the (unitless) overrelaxation parameter and $\omega^k = 1$ yields the usual Gauss-Seidel relaxation [Eq. (18)]. On the finest level, $k=l$ and $\tau^l=0$.

Choosing the appropriate parameters ω and μ for the relaxation scheme is important; otherwise, consistent convergence may not be obtained. We determined optimal convergence parameters for the relaxation by numerical experimentation during the test calculations. Values used for ω and μ are given for the test cases below.

When nonlocal pseudopotentials are present, the Hamiltonian includes an additional term, which is the nonlocal pseudopotential operator. Relativistic separable dual-space Gaussian pseudopotentials[34] are used in this work. The total pseudopotential, neglecting spin-orbit effects,[34] is given by

$$v(\mathbf{r},\mathbf{r}') = v_{loc}(r)\delta(\mathbf{r}-\mathbf{r}') + \sum_l v^l_{nl}(\mathbf{r},\mathbf{r}'). \tag{21}$$

Denoting the last term of Eq. (21) by $v_{nl}(\mathbf{r},\mathbf{r}')$, the action of the nonlocal pseudopotential operator ($\mathcal{V}_{nl}$) on a wave function $u(\mathbf{r})$ is expressed as

$$\mathcal{V}_{nl}\{u(\mathbf{r})\} = \int v_{nl}(\mathbf{r},\mathbf{r}')u(\mathbf{r}')d\mathbf{r}'. \tag{22}$$

The separable nature of the pseudopotenial is utilized to obtain an efficient real-space integration (see below).[37] The decay properties of the projectors allow restriction of the domain to a small region around each nucleus (specified below). Spin-orbit effects were neglected for the calculations presented here. On coarse levels, the pseudopotential is applied in the same form as on the fine level over the same physical domain (but sampling fewer grid points). We found that adding $\int v_{nl}(\mathbf{r},\mathbf{r}')d\mathbf{r}'$ to the diagonal term D^k in the denominator of Eqs. (18) and (20) brought improved convergence rates; this is an approximate way to include the effect of the nonlocal pseudopotential in the relaxation scheme. The quantity $\int v_{nl}(\mathbf{r},\mathbf{r}')d\mathbf{r}'$ can be evaluated and stored at the beginning of the program since it does not involve the wave functions.

VII. COMPUTATIONAL IMPLEMENTATION

In this section, details of the computational implementation for self-consistent Kohn-Sham eigenvalue problems are presented. All the calculations are performed on a three-dimensional real-space grid using the finite-difference method with 12th-order accuracy for the Laplacian. The FAS approach allows for different (typically lower) approximation orders on coarse levels; in the calculations presented here, we maintained the same Laplacian order on all levels for computational simplicity. We have tested lower-order representations on coarse levels and observe similar convergence rates to those presented here. The trapezoid rule is used for integration formulas. Three grid levels are utilized, comprising 17^3, 33^3, and 65^3 grid points. Interpolation and restriction operations are performed using linear interpolation and full weighted averaging, respectively. Atomic units are used throughout. The calculation parameters, sequence of operations, numbers of iteration, etc., were optimized during extensive numerical experimentation in the course of algorithm development.

A. FLR

Here, we describe the implementation for the case where the Ritz projection is performed on the full occupied fine-level subspace (FLR). Initially on the coarse grid (A in Fig. 1), random numbers are assigned to all wave function values on the grid, except the boundary points; for these finite systems, the boundary points are set to zero and fixed throughout the calculations. Next, the wave functions are orthonormalized and the effective potential is found. This is followed by the evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$ [Eq. (22)]. All these operations prepare the Hamiltonian matrix, and in the next step, the Ritz projection uses this matrix. The operation $\mathcal{V}_{nl}\{u(\mathbf{r})\}$ is evaluated once again, and the eigenvalues are updated using the Rayleigh quotient. The wave functions are relaxed once using the Gauss-Seidel method and normalized. The last four operations are repeated twice, and all the operations are repeated 50 times. At this point, the residuals of the eigenvalue equations are around 10^{-5}, and further iterations have almost no effect in reducing the residuals.

Next, the wave functions are interpolated to the next higher level (B in Fig. 1) and the same operations which are performed at A are repeated at B. An additional evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$ is done since this is needed in the evaluation of the defect correction for the next coarser level (C in Fig. 1). The wave functions are restricted to the coarse level, and the defect correction is found. On the coarse grid C, the evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, relaxation, and imposition of constraints are

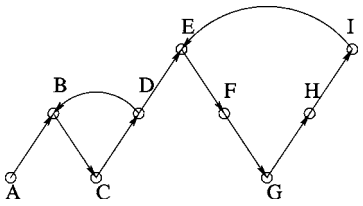

FIG. 1. Three-level full multigrid process (FMG). The bottom of the figure is the coarsest level.

repeated 5 times. The solution is then updated by an FAS correction step at *D* in Fig. 1. The same operations which are performed at *B* are repeated at *D*. This completes the two-level multigrid V cycle. The two-level V cycle is repeated roughly 15 times, and the solution is then interpolated to the finest level (*E* in Fig. 1).

The interpolated wave functions provide a good initial approximation on the fine level (*E*), which is the main advantage of using the full multigrid (FMG) scheme. On the fine level (*E*), orthonormalization of wave functions, update of the effective potential, the evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, and Ritz projection are performed once. This is followed by the evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, updating the eigenvalues using the Rayleigh quotient, relaxation (SOR) and normalization twice, evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, and relaxation and normalization twice again. An additional application of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$ is performed since this is required for calculating the defect correction on the middle level (*F* in Fig. 1).

Next, the solution is restricted to the middle level (*F*), where the defect correction is first computed from the fine-scale wave functions. The evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, relaxation, and imposition of constraints[30] are repeated 7 times at *F*, and another evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$ is performed since this is required for the defect correction on the next coarser level. Then the solution is restricted to the coarsest level (*G* in Fig. 1), and the defect correction is found. The operations which are performed at *C* are performed 20 times at *G*. The current approximation on the next finer level (*H* in Fig. 1) is then updated by an FAS correction step. The same operations which are done at *F* are performed at *H*, and the finest-level (*I* in Fig. 1) wave functions are updated with an FAS correction step. The same operations which are performed at *E* are performed at *I*. This completes the three-level V cycle. The three-level V cycle can be repeated until a desired convergence is reached. On the fine level (65^3), the Gram-Schmidt orthonormalization and the Ritz projection are performed only twice (once at each end of the V cycle), and the wave functions are relaxed altogether 8 times.

The convergence rate can be improved by performing more relaxation steps on the fine grid, although the optimum number of relaxation steps depends to some extent on the system. For our comparative test cases, we found eight relaxations per V cycle, four at each end of the V cycle, to be optimal. However, for larger sytems such as the C_{20} molecule, four relaxations were performed on the fine grid, two at each end of the V cycle; further increase in the number of relaxation steps did not improve the convergence rate for that case. The Poisson problem is solved once at each end of the V cycle.

B. CR-GRBR

Here we describe the modified algorithm in which fine-level Ritz projection is performed only on preselected overlapping eigenvalue clusters. The CR-GRBR process is performed on the three-level V cycle (*E-F-G-H-I* in Fig. 1). The standard Ritz projection method as described above is used prior to this point. Therefore, only the three-level V cycle is described.

Upon entering the three-level V cycle (*E* in Fig. 1), the wave functions are normalized and the effective potential is updated. Next, Ritz projection is performed on clusters once upon entering *E* to stabilize the wave functions initially. In the process of performing Ritz projection on clusters, $\mathcal{V}_{nl}\{u(\mathbf{r})\}$ is always evaluated right after the orthonormalization within a cluster so that an updated function is used in the Ritz projection process.

The Ritz projection on clusters is followed by the evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, relaxation, and normalization; these three steps are performed twice. Next, during the first three V cycles, we perform Ritz projection on clusters 3 times consecutively; for the subsequent V cycles, it is performed only once. The additional Ritz projections on clusters help to stabilize the wave finctions in their initial stages. This was especially necessary for larger systems such as the C_{20} molecule. These steps are followed by the evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, relaxation, and normalization twice and an additional evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, which is required for the defect correction on the next coarser level.

Then the solution is restricted to the middle level (*F* in Fig. 1), and the defect correction is evaluated. This is followed by the evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, relaxation, evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, and GRBR. These operations are repeated 7 times. An additional evaluation of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$ is made since it is required for the defect correction on the next coarser level.

The solution is restricted to the coarsest grid (*G* in Fig. 1), and the defect correction is found. Application of $\mathcal{V}_{nl}\{u(\mathbf{r})\}$, relaxation, and constraint imposition are performed 20 times at *G*. The next finer level (*H* in Fig. 1) solution is updated with an FAS correction step. The same operations which are performed at *F* are repeated at *H*, followed by updates of the fine-level (*I* in Fig. 1) functions in an FAS correction step. The operations which are performed at *E* are repeated at *I*. This completes the three-level V cycle. It can be repeated until a desired convergence is reached.

For the comparitive test cases considered here (benzene, glycine, and benzenedithiol), the following clusters are used for the CR-GRBR method. For benzene two clusters of (0, 7) and (7, 14) are chosen, and for glycine two clusters of (0, 6) and (6, 14) are used. Benzenedithiol has 21 wave functions, and two clusters of (0, 9) and (9, 20) are chosen; an overlap *s* of one element was sufficient in all three cases. The values of the parameters μ and ω, which are used for the relaxation schemes, are as follows. On the coarse level (at point *A* in Fig. 1), values of μ and ω are -30 and 1, respectively; for the two-level V cycle (*B*,*C*,*D* in Fig. 1), on both the coarse and middle levels, $\mu=-20$ and $\omega=1$; for the three-level V cycle (*E-F-G-H-I* in Fig. 1), on the coarse and middle levels, $\mu=-20$ and $\omega=1$, while on the finest level $\mu=0$ and ω

=1.75. Slight adjustments of the ω values on the finest level around the 1.75 value can improve the convergence for some cases. For the simultaneous algorithm, the finest-level ω value was taken as 1.5. These values were obtained through numerical optimization.

The cluster sizes are not limited to two as in these modest-sized test cases. Up to four clusters were used with an overlap of $s=3$ elements for the larger glycine dimer, glutamine, phenylalanine, tyrosine, and C_{20} illustrative cases, which are used to further display this new method. We found it necessary to include more than one overlap element for the larger systems. Also, for these larger systems we employed SOR relaxation on coarse levels in addition to the fine level, which brought improved convergence rates. As an example, for the C_{20} test case, the SOR ω values were taken as 1.8, 1.5, and 1.88 for the coarse, middle, and fine levels, respectively, while the μ values were −15, −20, and 0 for the three levels.

C. S-FLR and S-CR-GRBR

The simultaneous technique is similar to the method described above, except the charge density and effective potential are updated on the coarse grids during the FAS processing. The appropriate defect corrections for the charge density and the exchange-correlation potential described above must be included during these updates. Those defect corrections ensure that the charge density and the exchange correlation potential are simply the values obtained by restricting the fine-scale quantities at convergence. Updating the potential along with the eigenfunctions on coarse levels is intended to bring improved convergence rates. However, in our calculations we did not find any major improvements in the convergence rates compared to the sequential methods. Slightly different values for μ and ω are used in the relaxation as discussed above.

D. Poisson problem, exchange correlation potential, and density mixing

In all calculations except the simultaneous update methods (S-FLR and S-CR-GRBR), the Poisson equation is solved using an FAS multigrid solver given the total charge density on the fine grid. We solve the Poisson equation for the whole system (electrons plus nuclei),

$$\nabla^2\phi = -4\pi\rho, \tag{23}$$

and then the nuclear contribution is replaced by the smooth local component of the pseudopotential.

The boundary values for the Poisson problem are obtained using a multipole expansion up to the quadrupole term. On coarse levels, restricted values of the Coulomb potential are used except in the simultaneous techniques, where the potential is repeatedly updated during the coupled FAS iterations. Since the nuclear charge distribution for a large molecule cannot in general be placed on grid points, the charges are distributed to the neighboring grid points. The charge assignment to the nearest eight grid points is performed by splitting the charge according to the volume fraction within the cube.

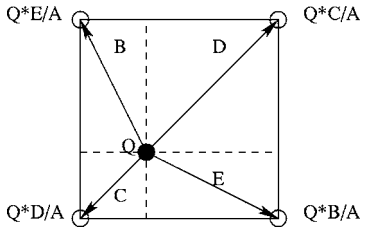

FIG. 2. Illustration of how the charge Q is distributed to neighboring grid points in two dimensions. B, C, D, and E are the partitioning areas, and A is the total area of the square. The same procedure can be performed with volumes in three dimensions.

Figure 2 displays how the partitioning is enacted in two dimensions, and this partitioning is easily generalized to three dimensions.

The exchange-correlation potential is computed using the formula given in Refs. 34 and 36. The VWN formula[35] for the exchange-correlation potential can also be used if desired, and it has been implemented in our algorithm. On coarse levels, the restricted exchange-correlation potential from the fine level is employed, except in the simultaneous technique, where it is repeatedly updated. Concerning mixing of old and new charge densities during self-consistency iterations, we found no mixing was required in the FLR algorithm for any of the test cases. In the CR-GRBR method and S-CR-GRBR algorithms, however, a mixing parameter of 0.5 was utilized (on the fine level at the end of the V cycle) to stabilize the self-consistent process.

E. Nonlocal pseudopotential

The application of the nonlocal component of the pseudopotential, Eq. (22), is performed only over a small cube around each nucleus. This is possible because the radial projectors tend to zero outside the covalent radius of an atom. The domain sizes used in our calculations are: 11 grid points (per side) on the fine grid, 9 grid points on the mid level, and 5 grid points for the coarse level. Equation (22) is implemented on all grid levels without any modification.

VIII. RESULTS

The benzene, glycine, and benzenedithiol molecules are used in our comparitive numerical studies. Utilizing pseudopotentials to remove the core electrons, benzene and glycine each have 15 wave functions and benzenedithiol has 21 wave functions. These molecules are chosen due to their different structures and resulting degeneracies. The benzene molecule has clear symmetry and thus degeneracies in the eigenvalues, while glycine possesses no clear symmetry and exhibits almost no eigenvalue degeneracy. Benzenedithiol is intermediate between these two extremes and has a larger number of wavefunctions.

The convergence rates for each of the algorithms are analyzed by plotting $\log_{10}|E-E_{con}|$ versus the number of V cycles (self-consistent iterations). E is the total energy for each iteration given by the formula in Ref. 38 (Eq. 7.2.10), and E_{con} is the fully converged numerical result. The total

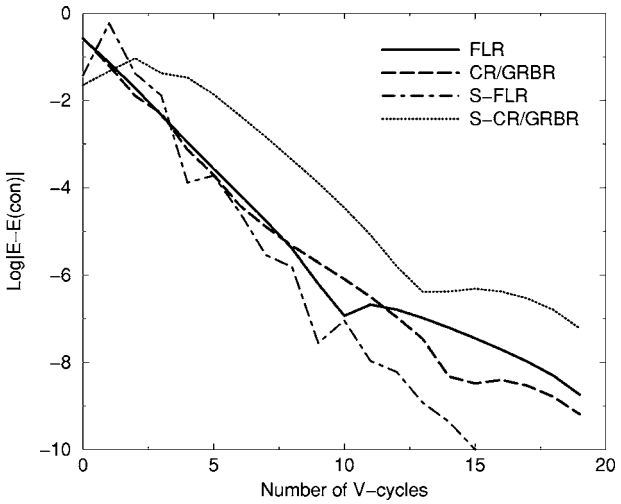

FIG. 3. Convergence rates for benzene. The $\log_{10}$ of the difference between the current total energy E (a.u.) and the fully converged value E_{con} is plotted against the number of V cycles (self-consistent iterations). The plots are for the FLR, CR-GRBR, S-FLR, and S-CR-GRBR methods. The fine grid spacing is $h=0.3$ a.u.

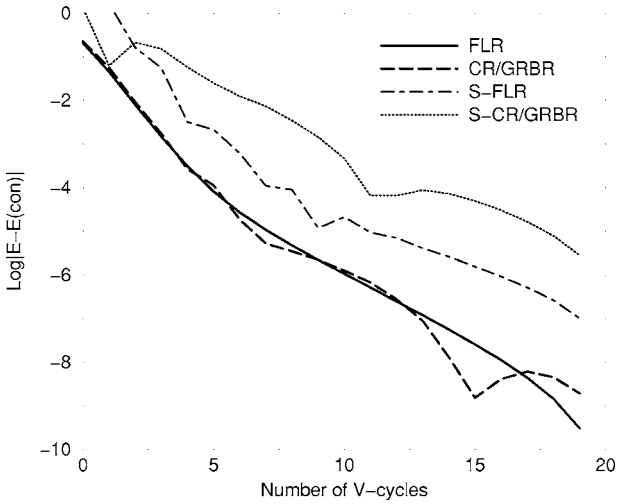

FIG. 4. Convergence rates for glycine. The $\log_{10}$ of the difference between the current total energy E (a.u.) and the fully converged value E_{con} is plotted against the number of V cycles (self-consistent iterations). The plots are for the FLR, CR-GRBR, S-FLR, and S-CR-GRBR methods. The fine grid spacing is $h=0.3$ a.u.

energy E is evaluated and stored at the end of each three-level V cycle.

The FLR, CR-GRBR, and S-FLR methods exhibit similar convergence rates for all test cases as shown in Figs. 3–5. In all cases except the S-FLR result for glycine, $\log_{10}|E-E_{con}|$ drops to −6 in about ten V cycles. The S-CR-GRBR method displays the slowest convergence rate in all three cases. A key result in this study is the excellent convergence rates shown by the CR-GRBR algorithm.

We also present results for the CR-GRBR convergence rates obtained for the larger illustrative cases glutamine (29 orbitals, 3 clusters), glycine dimer (26 orbitals, 3 clusters), tyrosine (35 orbitals, 4 clusters), phenylalanine (32 orbitals, 4 clusters), and the C_{20} molecule (40 orbitals, 4 clusters) in Fig. 6. Three overlap elements ($s=3$) were used for these cases. The larger required overlap is likely due to the smaller eigenvalue separation in the denser spectrum.

The convergence rates for the larger systems are robust but somewhat slower than for the smaller test cases. This behavior can be linked to three factors. First, it has been shown that ill-conditioning due to a wide separation in energy scales between the lowest and highest eigenvalues can lead to slowing of the convergence rates.[39] At first glance, the multigrid method should address this issue directly. For simple cases such as Poisson problems, this is certainly the case. The eigenvalue problems addressed here possess additional difficulties, however: the coarsest grid employed in the solver must be fine enough to maintain some resolution of the oscillations in the eigenfunctions. Therefore, to some extent there is incomplete decimation of the longer-wavelength features of the error. For eigenfunctions that possess both rapid oscillations and longer-wavelength features, this can be expected to lead to partial slowing down. It is worth exploring to what extent alternative methods for handling the ill-conditioning might help in this regard.

Second, Kerker preconditioning has been found to remove long-wavelength changes in the charge density upon mixing.[19] We have not addressed kinetic-energy ill-conditioning or charge preconditioning in this work. Third, the C_{20} physical system contains electrons which are more delocalized, so charge sloshing can be expected to be a problem. To some extent, the FAS multigrid process addresses these issues; limited charge mixing was utilized in our work, which illustrates the potential of the coarse-grid processing itself to stabilize the charge density. Future incorporation of these improvements can be expected to accelerate the convergence rates for larger systems. The CR-GRBR method presented here performs well in comparison with other multigrid[22,25] and plane-wave[40] codes, the advantage being

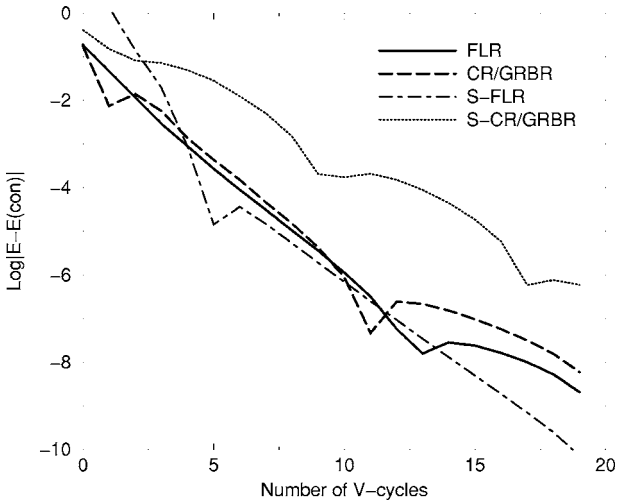

FIG. 5. Convergence rates for benzenedithiol. The $\log_{10}$ of the difference between the current total energy E (a.u.) and the fully converged value E_{con} is plotted against the number of V cycles (self-consistent iterations). The plots are for the FLR, CR-GRBR, S-FLR, and S-CR-GRBR methods. The fine grid spacing is $h=0.3$ a.u.

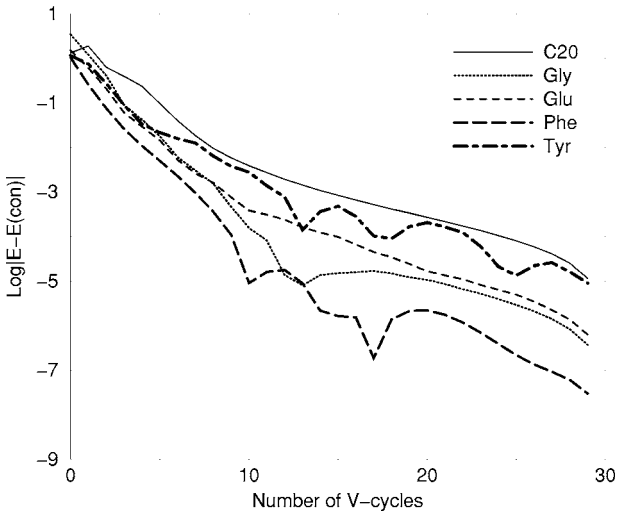

FIG. 6. Convergence rates for glycine dimer (Gly), glutamine (Glu), tyrosine (Tyr), phenylalanine (Phe), and the C_{20} molecule (C20). The CR-GRBR method is used. The $\log_{10}$ of the difference between the current total energy E (a.u.) and the fully converged value E_{con} is plotted against the number of V cycles (self-consistent iterations). The fine grid spacing for the first three molecules is $h=0.32$ a.u., and for the last two molecules $h=0.35$ a.u. is used.

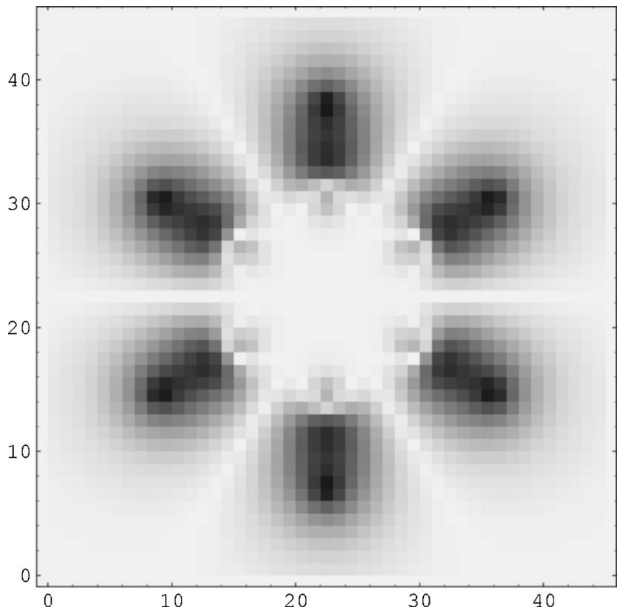

FIG. 7. This figure displays the wave function magnitude across the x-y plane for the seventh-highest occupied state of benzene. The intensity is proportional to the wave function magnitude. The benzene molecule lies on the x-y plane at the center of the domain. The axis labels are grid indices, and the grid spacing is 0.3 a.u.

the reduced fine-level cost of maintaining eigenfunction orthogonality in the algorithm.

In the CR-GRBR method, a measure of the overall convergence is the degree of orthogonality of the wave functions. The orthogonality of the wave functions between different clusters improves with increasing numbers of V cycles. Initially the degree of orthogonality starts around 10^{-2} (for $\langle u_i^l | u_j^l \rangle$), and it improves to 10^{-14} at full convergence. Finally, as an example of the physical output of the calculations, we present a density plot of the wave function magnitude for one of the states of the benzene molecule (Fig. 7).

To give an indication of the computational requirements for the results presented here, the calculations were performed on a 3.0 GHz CPU with 2 GB of physical memory. For the C_{20} case (the largest considered here), 20 three-level multigrid V-cycle iterations were performed with the CR-GRBR method. The fine grid consisted of 65^3 points. That calculation required roughly 1.5 h of CPU time.

IX. DISCUSSION AND CONCLUSIONS

We have successfully tested four nonlinear FAS multigrid techniques for solving the self-consistent Kohn-Sham equations of density functional theory. The four methods are the usual fine-level Ritz projection (FLR), Ritz projection performed on clusters in conjunction with GRBR (CR-GRBR), the simultaneous technique with fine-level Ritz projection (S-FLR), and the simultaneous technique with Ritz projection performed on clusters with GRBR on coarse levels (S-CR-GRBR).

Earlier we showed that the coarse-level GRBR method converged for small molecules and atoms with a small number of wave functions;[24] the convergence rate, however, was slower than the usual fine-level Ritz projection. Convergence rates could be restored by performing occasional orthonormalization on the fine grid for these small molecules. For large molecules such as benzenedithiol, however, the GRBR method stalled, and the GRBR method with occasional fine-level orthonormalization either stalled (benzenedithiol) or converged with a much slower convergence rate (benzene). The cause of the stalling is likely the relatively large number of wave functions these molecules posses and the ambiguous eigenvalue cluster structures presented. In cases with ambiguous eigenvalue cluster structure, it is difficult to automatically determine the relevant clusters for processing in the GRBR method. Also, the strongly varying effective potentials in the Kohn-Sham problem present a greater challenge than the relatively smooth model potentials examined by Costiner and Ta'asan.[31,32]

We have shown in this paper that typical multigrid convergence rates can be restored by including fine-scale Ritz projection on eigenvalue clusters in the CR-GRBR algorithm. In this method, cluster structures are chosen initially, and then the size of the cluster structures can be increased until a suitable set of clusters is reached (typically clusters of roughly ten eigenvalues). The Ritz projection is performed on those overlapping clusters on the fine grid in addition to GRBR on a coarse level (the middle level in our cases). This procedure removes the stalling observed in the original form of the GRBR algorithm, and good convergence rates are then observed. The nominal scaling of this algorithm for modest-sized systems is qN_g^l (or N_e^2, where N_e is the number of electrons), as opposed to $q^2N_g^l$ when the full fine-scale Ritz procedure is implemented.

The orthonormalization and the Ritz projection performed on m overlapping clusters having p wave functions in each

cluster scale as $m(p+s)^2N_g^l$, where s is the number of overlap elements and s is a small number. Therefore, nearly an m-fold savings in computational time can be obtained relative to the regular Ritz projection method. In addition to the main test cases examined here, we have successfully applied the CR-GRBR method to larger systems such as the glycine dimer, glutamine, tyrosine, phenylalanine, and the C_{20} molecule. While the scaling for the modest-sized systems examined here is quadratic, it remains to be seen how the method performs for very large systems in which the eigenvalue spectrum becomes more dense.

The CR-GRBR algorithm is a clustering algorithm in energy space. It is clear from our convergence results that a proper separation of the eigenvalues can be enforced by the fine-scale clustering process; the algorithm can achieve full convergence to machine precision in the residuals. The energy-clustering method can be viewed as a kind of perturbation theory, where communication between the exactly diagonalized cases within each cluster is communicated through the overlap elements. In addition, widely separated eigenvalue-eigenfunction pairs tend to be inherently orthogonal due to their differing length scales of oscillation.

The CR-GRBR method can be contrasted with real-space localization techniques,[3,17,19] which have been utilized to achieve near-linear scaling DFT algorithms. Those methods generally rely on a band gap, leading to relatively rapid decay of the density matrix in real space. The localization then yields a density matrix and Hamiltonian (in the basis of occupied states), which are banded matrices in real space. The localization, however, may come at a price (which decreases with increasing localization radius);[4] the total energy convergence may stall at an energy scale dictated by lost information from the density matrix or orbital truncation. The energy-clustering CR-GRBR algorithm can be viewed as an alternative algorithm, which has the drawback of steeper scaling, but the advantage of full numerical convergence given a chosen real-space approximation to the Kohn-Sham equations.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the National Science Foundation Grant No. (CHE-0112322), the DOD MURI (Army) program, and the Army Research Office for this research. We also thank Achi Brandt, Shlomo Ta'asan, Frank Pinski, and Duane Johnson for many helpful discussions.

*Author to whom correspondence should be addressed. Electronic address: becktl@email.uc.edu

[1] M. Payne, M. Teter, D. Allan, T. Arias, and J. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[2] D. P. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler, Int. J. Quantum Chem. **65**, 453 (1997).
[3] J.-L. Fattebert and J. Bernholc, Phys. Rev. B **62**, 1713 (2000).
[4] J.-L. Fattebert and F. Gygi, Comput. Phys. Commun. **162**, 24 (2004).
[5] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, J. Chem. Phys. **122**, 084119 (2005).
[6] M. Challacombe, Comput. Phys. Commun. **128**, 93 (2000).
[7] J. R. Chelikowsky, N. Troullier, K. Wu, and Y. Saad, Phys. Rev. B **50**, 11355 (1994).
[8] C. M. Goringe, E. Hernández, M. J. Gillan, and I. J. Bush, Comput. Phys. Commun. **102**, 1 (1997).
[9] T. L. Beck, Rev. Mod. Phys. **72**, 1041 (2000).
[10] K. Hirose, T. Ono, Y. Fujimoto, and S. Tsukamoto, *First-Principles Calculations in Real-Space Formalism: Electronic Configurations and Transport Properties of Nanostructures* (Imperial College Press, London, 2005).
[11] N. A. Modine, G. Zumbach, and E. Kaxiras, Phys. Rev. B **55**, 10289 (1997).
[12] T. Hoshi and T. Fujiwara, J. Phys. Soc. Jpn. **66**, 3710 (1997).
[13] L. Kronik *et al.*, Phys. Status Solidi B **243**, 1063 (2006).
[14] I.-H. Lee, Y.-H. Kim, and R. M. Martin, Phys. Rev. B **61**, 4397 (2000).
[15] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[16] D. R. Bowler, I. J. Bush, and M. J. Gillan, Int. J. Quantum Chem. **77**, 831 (2000).
[17] G. Feng and T. L. Beck, Phys. Status Solidi B **243**, 1054 (2006).
[18] J. E. Pask and P. A. Sterne, Modell. Simul. Mater. Sci. Eng. **13**, R71 (2005).
[19] D. R. Bowler, R. Choudhury, M. J. Gillan, and T. Miyazaki, Phys. Status Solidi B **243**, 989 (2006).
[20] T. Torsti *et al.*, Phys. Status Solidi B **243**, 1016 (2006).
[21] M. Heiskanen, T. Torsti, M. J. Puska, and R. M. Nieminen, Phys. Rev. B **63**, 245106 (2001).
[22] E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **54**, 14362 (1996).
[23] J. Wang and T. L. Beck, J. Chem. Phys. **112**, 9223 (2000).
[24] N. Wijesekera, G. Feng, and T. L. Beck, J. Theor. Comput. Chem. **2**, 553 (2003).
[25] F. Ancilotto, P. Blandin, and F. Toigo, Phys. Rev. B **59**, 7868 (1999).
[26] J.-L. Fattebert, J. Comput. Phys. **149**, 75 (1999).
[27] A. Brandt, Math. Comput. **31**, 333 (1977).
[28] A. Brandt, *Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics* (GMD-Studien, 1984).
[29] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial*, 2nd ed. (Society for Industrial and Applied Mathematics, Philadelphia, 2000).
[30] A. Brandt, S. F. McCormick, and J. Ruge, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **4**, 244 (1983).
[31] S. Costiner and S. Ta'asan, Phys. Rev. E **51**, 3704 (1995).
[32] S. Costiner and S. Ta'asan, Phys. Rev. E **52**, 1181 (1995).
[33] J. Wang, Y. Wang, S. Yu, and D. Kolb, J. Phys.: Condens. Matter **17**, 3701 (2005).
[34] C. Hartwigsen, S. Goedecker, and J. Hutter, Phys. Rev. B **58**, 3641 (1998).
[35] S. H. Vosko, L. Wilk, and M. Nussair, Can. J. Phys. **58**, 1200 (1980).

[36] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).
[37] R. D. King-Smith, M. C. Payne, and J. S. Lin, Phys. Rev. B **44**, 13063 (1991).
[38] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989), p. 275.
[39] A. A. Mostofi, P. D. Haynes, C.-K. Skylaris, and M. C. Payne, J. Chem. Phys. **119**, 8842 (2003).
[40] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).