Journal of Computational Physics **149,** 75–94 (1999)
Article ID jcph.1998.6138, available online at http://www.idealibrary.com on IDEAL®

Finite Difference Schemes and Block Rayleigh Quotient Iteration for Electronic Structure Calculations on Composite Grids

Jean-Luc Fattebert[1]

Department of Mathematics, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
E-mail: fatteber@nemo.physics.ncsu.edu

Received December 10, 1997; revised August 10, 1998

We present an original numerical method to discretize the Kohn–Sham equations by a finite difference scheme in real-space when computing the electronic structure of a molecule. The singular atomic potentials are replaced by pseudopotentials and the discretization of the 3D problem is done on a composite mesh refined in part of the domain. A "Mehrstellenverfahren" finite difference scheme is used to approximate the Laplacian on the regular parts of the grid. The nonlinearity of the potential operator in the Kohn–Sham equations is treated by a fixed point algorithm. At each step an iterative scheme is applied to determine the searched solutions of the eigenvalue problem for a given fixed potential. The eigensolver is a block generalization of the Rayleigh quotient iteration which uses Petrov–Galerkin approximations. The algorithm is adapted to a multigrid resolution of the linear systems obtained in the inverse iterations. Numerical tests of the different algorithms are presented on problems coming from the electronic structure calculation of some molecules.

Key Words: Rayleigh quotient iteration; multigrid method; finite differences; mesh refinement; electronic structure calculations; Kohn–Sham equations.

1. INTRODUCTION

Over the last years real-space methods have appeared in the domain of 3D *ab initio* electronic structure calculations for numerically solving the Kohn–Sham (KS) equations [1–10]. The main reason to replace the classical plane wave expansion of the orbitals by these new discretization schemes is their local feature, allowing natural local mesh refinements [7, 10] and efficient subdomains decompositions on massively parallel supercomputers [4]. Moreover, this approach is more natural for nonperiodic systems. The computation

[1] Present address: Department of Physics, North Carolina State University, Raleigh, NC 27695-8202.

0021-9991/99 $30.00

of the kinetic energy associated to a wave function in real-space requires high-order finite difference (FD) schemes. Such schemes can lead to wrong results if they involve nodes too far from each other, in particular if they are applied to rapidly oscillating functions, compared to the mesh spaces. The "Mehrstellenverfahren" finite difference schemes [11] permit us to obtain higher orders by using nodes confined in a smaller volume than classical schemes. In the case of the 3D Laplacian, order four formulas are available that involve only the neighbor nodes confined in a cube of volume $2h \times 2h \times 2h$ (where h denotes the grid spacing) around the node considered, so allowing easier combination with local mesh refinement and Dirichlet boundary conditions.

Discretization of the KS equations by FD, or finite elements methods, leads to large-scale linear algebra eigenvalue problems requiring efficient iterative eigensolvers. In general, the involved matrices are sparse and we are only interested in a small fraction of the spectrum (the smallest eigenvalues) and the corresponding invariant subspace. In quantum calculations, Davidson's method [12, 13], a diagonal preconditioned Lanczos' method, is the most used because of the diagonal dominance of the matrices. Recently an improved version has been proposed: the Jacobi–Davidson method [14]. The idea of these methods is to generate in an adequate way a subspace of much smaller dimension than the space in which the problem is defined, and then to process diagonalization in this subspace. The dimension of these subspaces has to be larger than the number of searched eigenpairs and could become prohibitively large on very large-scale eigenvalue problems, so that these methods are not well designed for the resolution of the problems we encounter in quantum chemistry when discretizing the equations with a large number of degrees of freedom.

The multigrid method [15, 16] is also a mathematical tool that has recently appeared in different methods used for solving the KS equations [3–6, 9]. It drastically improves the convergence rates of the classical iterative linear systems solvers by using coarse grid approximations of the partial differential equation for solving. At the same time, it allows us to use iterative methods for the eigenvalue problems like the Rayleigh quotient iteration (RQI) method that could not be applied on large-scale problems without an efficient linear systems solver.

In this paper we present a generalization of the method first described in [5]. We propose a scheme to discretize and numerically solve the KS equations on a grid composed of two regular grids, the coarser covering the whole domain and the finer having a mesh space two times smaller and covering just the part of the domain where we want more precision. The ionic potentials are represented by nonlocal pseudopotentials so that we do not have to deal with singular potentials and core electrons [17]. We use the local density approximation (LDA) model for the exchange and correlation potential. The discretized KS equations are solved by a self-consistent (SC) iterative scheme in which, at every step, an eigenvalue problem with fixed potential is solved by a block generalization of the RQI. One of the main features of the method is that it avoids generating search subspaces (in which we diagonalize the operator) of dimension larger than the number of desired eigenvalues. Moreover, it treats simultaneously and in the same way all the search eigenpairs. The power of the multigrid method is used to efficiently solve the indefinite linear systems resulting from the inverse iteration equations present in the algorithm. Numerical tests of convergence of the multigrid method and of the eigenvalue solver are presented in Section 5 on examples provided in the case of the electronic structure calculation of the furan molecule.

The method described here includes the particular case of regular grids treated in [5] and extend it to composite grids. In particular the multigrid algorithm is generalized to the resolution of the linear systems obtained in the inverse iterations on the composite grids. The

applications on which the algorithms are tested and applied are less academic and show the applicability of the method to more general electronic structure calculations. In particular we treat the furan molecule (C_4OH_4) with the ionic Bachelet–Hamman–Schlueter (BHS) pseudopotentials introduced in [18]. More mathematical details can be found in [6].

2. FD SCHEME ON COMPOSITE GRID

For simplicity reasons, we choose to work in the cubic domain $\Omega = (0, L) \times (0, L) \times (0, L)$ with homogeneous Dirichlet boundary conditions. The generalization to a parallelepiped domain is straightforward. Let us first define a regular mesh on the domain Ω,

$$\Omega_\ell = \{(\xi, \eta, \zeta) \in \Omega \mid \xi = ih_\ell, \eta = jh_\ell, \zeta = kh_\ell, 1 \le i, j, k \le n_\ell\}, \tag{1}$$

defined for $\ell \in \mathbb{N}$, composed of n_ℓ^3 inner nodes, with $n_\ell = a \cdot 2^{\ell+1} - 1$, $h_\ell = L/(n_\ell + 1)$ being the distance between nodes along the three axis ξ, η, and ζ, and $a \in \mathbb{N}$ fixed. In this paper, we will always take $a = 1$. As in the multigrid literature, ℓ designates the grid level. The grid are indexed so that they contain more points as ℓ increases. Implicitly (by not introducing any nodes on $\partial\Omega$), we have imposed homogeneous Dirichlet boundary conditions, but periodic ones could be chosen without main changes in the algorithms presented below. By analogy with the notations in $\mathbb{R}^3$ we write

$$\bar{\Omega}_\ell = \{(\xi, \eta, \zeta) \in \Omega \mid \xi = ih_\ell, \eta = jh_\ell, \zeta = kh_\ell, 0 \le i, j, k \le n_\ell + 1\}.$$

We denote $\mathcal{E}_\ell$ the finite dimensional vectorial space of the real functions defined on Ω_ℓ.

Let (i_1, j_1, k_1) and (i_2, j_2, k_2) be the coordinates of two nodes of $\bar{\Omega}_\ell$ such that $0 \le i_1 < i_2 \le n_\ell + 1$, $0 \le j_1 < j_2 \le n_\ell + 1$, and $0 \le k_1 < k_2 \le n_\ell + 1$. We define the parallelepiped subdomain ω of Ω by

$$\omega = \left\{ (\xi, \eta, \zeta) \in \Omega \;\middle|\; \begin{array}{l} \left(i_1 - \frac{1}{2}\right)h_\ell < \xi < \left(i_2 + \frac{1}{2}\right)h_\ell, \\ \left(j_1 - \frac{1}{2}\right)h_\ell < \eta < \left(j_2 + \frac{1}{2}\right)h_\ell, \\ \left(k_1 - \frac{1}{2}\right)h_\ell < \zeta < \left(k_2 + \frac{1}{2}\right)h_\ell \end{array} \right\}.$$

In this subdomain, we introduce the regular grids

$$\omega_\ell = \Omega_\ell \cap \omega,$$
$$\omega_{\ell+1} = \Omega_{\ell+1} \cap \omega,$$

where $\Omega_{\ell+1}$ is defined by (1) with the index $\ell + 1$ instead of ℓ. By analogy with the notations in $\mathbb{R}^3$, we write

$$\partial\omega_{\ell+1} = \Omega_{\ell+1} \cap \partial\omega,$$
$$\bar{\omega}_{\ell+1} = \omega_{\ell+1} \cup \partial\omega_{\ell+1}.$$

The composite mesh $\hat{\Omega}_{\ell+1}$ is then defined by the union of the two regular grids,

$$\hat{\Omega}_{\ell+1} = \bar{\omega}_{\ell+1} \cup \Omega_\ell.$$

We note that, if $\Omega \subset \omega$, $\hat{\Omega}_{\ell+1} = \Omega_{\ell+1}$ is a regular grid.

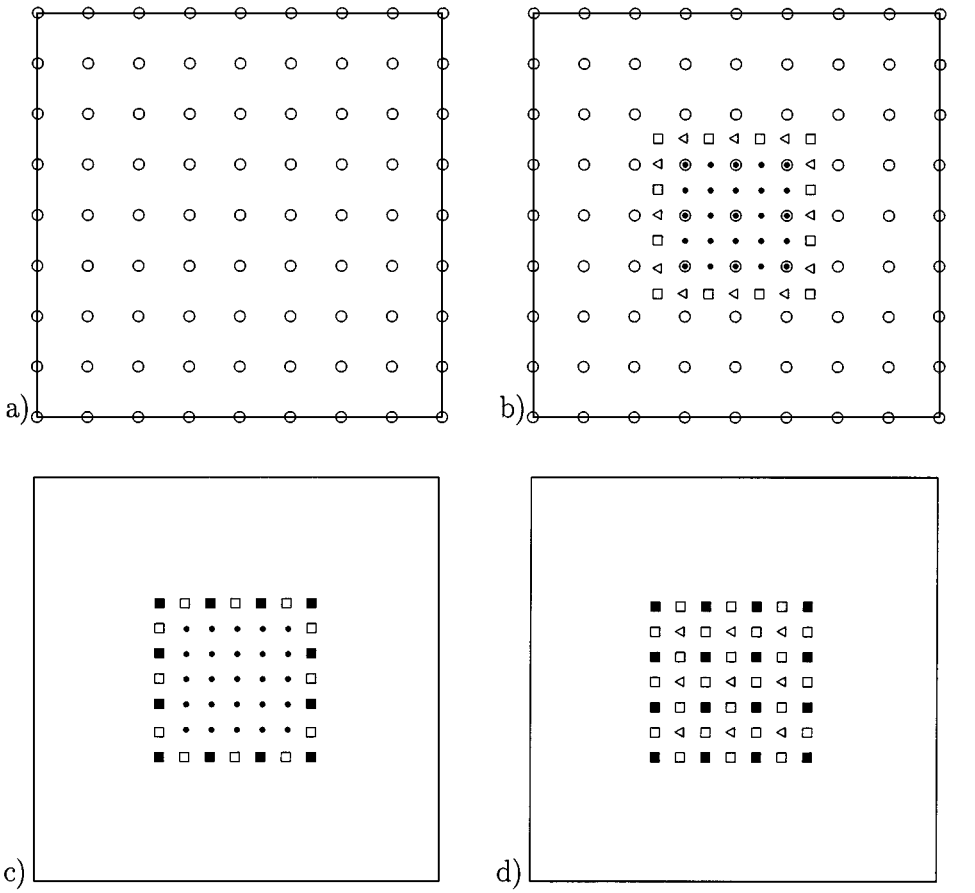

FIG. 1. Composite grid with $\ell = 2$, $i_1 = j_1 = k_1 = 3$, $i_2 = j_2 = k_2 = 5$: (a) $\zeta = kh_\ell$, $k \le k_1$ or $k \ge k_2$; (b) $\zeta = kh_\ell$, $k_1 \le k \le k_2$; (c) $\zeta = (k + \frac{1}{2})h_\ell$, $k_1 \le k < k_2$; (d) $\zeta = (k_1 - \frac{1}{2})h_\ell$ or $\zeta = (k_2 + \frac{1}{2})h_\ell$. Nodes: $\circ : \bar{\Omega}_\ell \backslash \omega_{\ell+1}$, $\odot : \omega_{\ell+1} \cap \Omega_\ell$, $\bullet : \omega_{\ell+1} \backslash \Omega_\ell$, $\triangleleft : \partial\omega^{(1)}_{\ell+1}$, $\square : \partial\omega^{(2)}_{\ell+1}$, $\blacksquare : \partial\omega^{(3)}_{\ell+1}$.

Instead of introducing slave points on the interface between the refined ($\bar{\omega}_{\ell+1}$) and non-refined ($\Omega_\ell \backslash \omega_\ell$) parts of the grid, we consider the composite mesh as a nonuniform grid near the interface. So we are led to introduce nonuniform finite difference stencils at the interface grid points contained in $\partial\omega_{\ell+1}$ [19]. Note here that such an approach would be much more difficult to put in practice with high order finite difference schemes involving values on more distant nodes. We decompose $\partial\omega_{\ell+1}$ in three disjoint subsets of nodes,

$$\partial\omega_{\ell+1} = \partial\omega^{(1)}_{\ell+1} \cup \partial\omega^{(2)}_{\ell+1} \cup \partial\omega^{(3)}_{\ell+1},$$

to which we associate three specific stencils. These subsets are defined (see Fig. 1) as

$$\begin{aligned}
\partial\omega^{(1)}_{\ell+1} = \{(\xi,\eta,\zeta) \in \partial\omega_{\ell+1} |\ & (\xi \pm h_{\ell+1}, \eta, \zeta) \in \hat{\Omega}_{\ell+1}, \\
& (\xi, \eta \pm h_{\ell+1}, \zeta) \in \hat{\Omega}_{\ell+1}, \\
& (\xi, \eta, \zeta \pm h_{\ell+1}) \in \hat{\Omega}_{\ell+1}\}, \\
\partial\omega^{(3)}_{\ell+1} = \{(\xi,\eta,\zeta) \in \partial\omega_{\ell+1} |\ & (\xi \pm h_{\ell+1}, \eta \pm h_{\ell+1}, \zeta \pm h_{\ell+1}) \in \hat{\Omega}_{\ell+1}\},
\end{aligned}$$

and

$$\partial\omega_{\ell+1}^{(2)} = \partial\omega_{\ell+1} \big\backslash \big\{ \partial\omega_{\ell+1}^{(3)} \cup \partial\omega_{\ell+1}^{(1)} \big\}.$$

For each subset, we will define a finite difference operator for the Laplacian that is adapted to the neighbor nodes available. Figure 1 shows views of the composite mesh $\hat{\Omega}_{\ell+1}$ on planes $\zeta = \text{const}$ for an example with $\ell = 2$. We associate to $\hat{\Omega}_{\ell+1}$ the finite-dimensional functional space $\hat{\mathcal{E}}_{\ell+1}$ defined by

$$\begin{aligned}\hat{\mathcal{E}}_{\ell+1} = \big\{ \hat{v}_{\ell+1} = \big(\hat{v}_{\ell+1}^{(1)}, \hat{v}_{\ell+1}^{(2)}, \hat{v}_{\ell+1}^{(3)}, \hat{v}_{\ell+1}^{(4)}, \hat{v}_{\ell+1}^{(5)}\big) \;\big|\; & \hat{v}_{\ell+1}^{(1)} : \Omega_\ell \backslash \omega_\ell \to \mathbb{R}, \\ & \hat{v}_{\ell+1}^{(2)} : \omega_{\ell+1} \to \mathbb{R}, \\ & \hat{v}_{\ell+1}^{(3)} : \partial\omega_{\ell+1}^{(1)} \to \mathbb{R}, \\ & \hat{v}_{\ell+1}^{(4)} : \partial\omega_{\ell+1}^{(2)} \to \mathbb{R}, \\ & \hat{v}_{\ell+1}^{(5)} : \partial\omega_{\ell+1}^{(3)} \to \mathbb{R} \big\}.\end{aligned}$$

It is easy to see that it is isomorphic to the finite-dimensional vectorial space of the real functions defined on $\hat{\Omega}_{\ell+1}$. In the following, we will not distinguish between them. We will use the scalar product $(.,.)_{\widehat{\ell+1}}$ between two functions $\hat{u}_{\ell+1}$ and $\hat{v}_{\ell+1} \in \hat{\mathcal{E}}_{\ell+1}$, defined by

$$(\hat{u}_{\ell+1}, \hat{v}_{\ell+1})_{\widehat{\ell+1}} = h_\ell^3 \sum_{x \in (\Omega_\ell \backslash \omega_\ell)} \hat{u}_{\ell+1}(x)\hat{v}_{\ell+1}(x) + h_{\ell+1}^3 \sum_{x \in \overline{\omega}_{\ell+1}} \hat{u}_{\ell+1}(x)\hat{v}_{\ell+1}(x), \qquad (2)$$

which is a consistent approximation of the integral $\int_\Omega u(x)v(x)\,dx$ if $\hat{u}_{\ell+1}(x) = u(x)$ and $\hat{v}_{\ell+1}(x) = v(x)$ for $x \in \hat{\Omega}_{\ell+1}$.

The tools are now ready to introduce a FD scheme for the Laplacian on the composite grid $\hat{\Omega}_{\ell+1}$. It is based on a "Mehrstellenverfahren" scheme of order 4 in the regular parts of the grid ($\Omega_\ell \backslash \omega_\ell$ and $\omega_{\ell+1}$) and on specific schemes of order 2 on the nodes $\partial\omega_{\ell+1}$. In this purpose, we define the discrete operators $\hat{C}_{\ell+1}$ and $\hat{B}_{\ell+1}$ by their action on $\hat{u}_{\ell+1} \in \hat{\mathcal{E}}_{\ell+1}$,

$$\hat{C}_{\ell+1}\hat{u}_{\ell+1} = \big(\hat{c}_{\ell+1}^{(1)}, \hat{c}_{\ell+1}^{(2)}, \hat{c}_{\ell+1}^{(3)}, \hat{c}_{\ell+1}^{(4)}, \hat{c}_{\ell+1}^{(5)}\big), \qquad (3)$$

where

$$\hat{c}_{\ell+1}^{(1)}(x_0) = \frac{1}{6h_\ell^2} \left\{ 24\hat{u}_{\ell+1}(x_0) - 2 \sum_{\substack{x \in \Omega_\ell \\ \|x - x_0\| = h_\ell}} \hat{u}_{\ell+1}(x) - \sum_{\substack{x \in \Omega_\ell \\ \|x - x_0\| = \sqrt{2}h_\ell}} \hat{u}_{\ell+1}(x) \right\}$$

for $x_0 \in (\Omega_\ell \backslash \omega_\ell)$,

$$\hat{c}_{\ell+1}^{(2)}(x_0) = \frac{1}{6h_{\ell+1}^2} \left\{ 24\hat{u}_{\ell+1}(x_0) - 2 \sum_{\substack{x \in \overline{\omega}_{\ell+1} \\ \|x - x_0\| = h_{\ell+1}}} \hat{u}_{\ell+1}(x) - \sum_{\substack{x \in \overline{\omega}_{\ell+1} \\ \|x - x_0\| = \sqrt{2}h_{\ell+1}}} \hat{u}_{\ell+1}(x) \right\}$$

for $x_0 \in \omega_{\ell+1}$,

$$\hat{c}_{\ell+1}^{(3)}(x_0) = \frac{1}{h_{\ell+1}^2} \left\{ 6\hat{u}_{\ell+1}(x_0) - \sum_{\substack{x \in \hat{\Omega}_{\ell+1} \\ \|x - x_0\| = h_{\ell+1}}} \hat{u}_{\ell+1}(x) \right\}$$

for $x_0 \in \partial\omega_{\ell+1}^{(1)}$,

$$\hat{c}_{\ell+1}^{(4)}(x_0) = \frac{1}{2h_{\ell+1}^2} \left\{ 8\hat{u}_{\ell+1}(x_0) - 2 \sum_{\substack{x \in \partial\omega_{\ell+1}^{(3)} \\ \|x-x_0\|=h_{\ell+1}}} \hat{u}_{\ell+1}(x) - \sum_{\substack{x \in \Omega_\ell \\ \|x-x_0\|=\sqrt{2}h_{\ell+1}}} \hat{u}_{\ell+1}(x) \right\}$$

for $x_0 \in \partial\omega_{\ell+1}^{(2)}$,

$$\hat{c}_{\ell+1}^{(5)}(x_0) = \frac{1}{4h_{\ell+1}^2} \left\{ 8\hat{u}_{\ell+1}(x_0) - \sum_{\substack{x \in \Omega_\ell \\ \|x-x_0\|=\sqrt{3}h_{\ell+1}}} \hat{u}_{\ell+1}(x) \right\}$$

for $x_0 \in \partial\omega_{\ell+1}^{(3)}$, while

$$\hat{B}_{\ell+1}\hat{u}_{\ell+1} = \left(\hat{b}_{\ell+1}^{(1)}, \hat{b}_{\ell+1}^{(2)}, \hat{b}_{\ell+1}^{(3)}, \hat{b}_{\ell+1}^{(4)}, \hat{b}_{\ell+1}^{(5)}\right), \tag{4}$$

where

$$\hat{b}_{\ell+1}^{(1)}(x_0) = \frac{1}{72} \left\{ 48\hat{u}_{\ell+1}(x_0) + 2 \sum_{\substack{x \in \Omega_\ell \\ \|x-x_0\|=h_\ell}} \hat{u}_{\ell+1}(x) + \sum_{\substack{x \in \Omega_\ell \\ \|x-x_0\|=\sqrt{2}h_\ell}} \hat{u}_{\ell+1}(x) \right\}$$

for $x_0 \in (\Omega_\ell \backslash \omega_\ell)$,

$$\hat{b}_{\ell+1}^{(2)}(x_0) = \frac{1}{72} \left\{ 48\hat{u}_{\ell+1}(x_0) + 2 \sum_{\substack{x \in \bar{\omega}_{\ell+1} \\ \|x-x_0\|=h_{\ell+1}}} \hat{u}_{\ell+1}(x) + \sum_{\substack{x \in \bar{\omega}_{\ell+1} \\ \|x-x_0\|=\sqrt{2}h_{\ell+1}}} \hat{u}_{\ell+1}(x) \right\}$$

for $x_0 \in \omega_{\ell+1}$,

$$\begin{aligned} \hat{b}_{\ell+1}^{(3)}(x_0) &= \hat{u}_{\ell+1}(x_0) \quad \text{for } x_0 \in \partial\omega_{\ell+1}^{(1)}, \\ \hat{b}_{\ell+1}^{(4)}(x_0) &= \hat{u}_{\ell+1}(x_0) \quad \text{for } x_0 \in \partial\omega_{\ell+1}^{(2)}, \\ \hat{b}_{\ell+1}^{(5)}(x_0) &= \hat{u}_{\ell+1}(x_0) \quad \text{for } x_0 \in \partial\omega_{\ell+1}^{(3)}. \end{aligned}$$

If $\hat{\Omega}_{\ell+1} \neq \Omega_{\ell+1}$, the operators $\hat{C}_{\ell+1}$ and $\hat{B}_{\ell+1}$ are not symmetric, but nevertheless $\hat{B}_{\ell+1}$ is well conditioned, in contrast to the less expensive "Mehrstellenverfahren" finite difference scheme used in [4], even for periodic boundary conditions. Moreover, both operators can be represented by sparse matrices. The operators $\hat{C}_{\ell+1}$ and $\hat{B}_{\ell+1}$ have been designed in order that we obtain, for $u \in C^6(\bar{\Omega})$, $u = 0$ on $\partial\Omega$ [6],

$$(\hat{C}_{\ell+1}u)(x_0) = (\hat{B}_{\ell+1}(-\Delta u))(x_0) + O\left(h_\ell^\kappa\right) \tag{5}$$

with $\kappa = 4$ for $x_0 \in (\Omega_\ell \cup \omega_{\ell+1})$ and $\kappa = 2$ for $x_0 \in \partial\omega_{\ell+1}$. It means that on the nodes of $\partial\omega_{\ell+1}$, the finite difference schemes give the order 2, while on the other nodes the

"Mehrstellenverfahren" schemes using all the nodes x such that $\|x - x_0\| \leq \sqrt{2}h_{\ell+1}$ give the order 4.

Let us begin by applying the discretization scheme described above to the Poisson problem:

$$\begin{cases} (-\Delta v)(x) = 4\pi\rho(x), & x \in \Omega \\ v(x) = 0, & x \in \partial\Omega. \end{cases} \tag{6}$$

We discretize this equation on the composite mesh $\hat{\Omega}_{\ell+1}$ using the finite difference scheme (5). So we have to look for a solution $\hat{v}_{\ell+1} \in \hat{\mathcal{E}}_{\ell+1}$ of the equation

$$\hat{C}_{\ell+1}\hat{v}_{\ell+1} = 4\pi\,\hat{B}_{\ell+1}\hat{\rho}_{\ell+1}, \tag{7}$$

where $\hat{\rho}_{\ell+1} \in \hat{\mathcal{E}}_{\ell+1}$ is defined by $\hat{\rho}_{\ell+1}(x) = \rho(x)$ for $x \in \hat{\Omega}_{\ell+1}$. Using standard finite difference theory (see [20] for example), we easily obtain a convergence result for this scheme. If we assume that Eq. (6) has a unique solution u whose regularity is $C^{2+\kappa}(\Omega)$, $\kappa = 1, 2$, then the maximum of the error between the discrete solution $\hat{u}_{\ell+1}$ of (7) and the exact solution u is $O(h_\ell^\kappa)$.

We note that the number of nodes on the boundary $\partial\omega_{\ell+1}$ is small ($O(n_\ell^2)$), compared to the total number of nodes ($O(n_\ell^3)$) on the whole mesh. We therefore expect to have globally a convergence of order larger than 2. But more important is the fact that the solutions computed on the grid $\hat{\Omega}_{\ell+1}$ will exhibit an error of same magnitude as the solutions computed on $\Omega_{\ell+1}$ (see Section 5). This method will be used in the SC resolution of the KS equations to define the Hartree potential, solution of a Poisson problem.

Our purpose is to discretize the KS equations [21] (in Rydberg units)

$$H^{KS}\Psi^{(j)} = -\Delta\Psi^{(j)} + V^{KS}\Psi^{(j)} = \epsilon^{(j)}\Psi^{(j)}, \quad j = 1, \ldots, p,$$

where Δ designates the Laplacian operator in $\mathbb{R}^3$, and $\epsilon^{(1)} \leq \cdots \leq \epsilon^{(p)} < \cdots$ are the smallest eigenvalues. V^{KS} is the potential operator composed of the Hartree (Coulomb interactions), exchange and correlation, and pseudopotential terms. It is nonlinear because of the dependence of the two first terms on the electronic density

$$\rho(x) = \sum_{j=1}^{p} 2\left|\Psi^{(j)}(x)\right|^2.$$

The factor 2 accounts for the electronic occupation numbers assumed to be all equal. We work with the exchange and correlation potential given by the local density approximation (LDA) model and describe the ions by the BHS nonlocal pseudopotentials given in [18].

If we assume that a regular grid $\Omega_{\ell+1}$ is fine enough on the whole domain to obtain a good approximation of the solutions of the KS problem, we will replace this one by a composite grid $\hat{\Omega}_{\ell+1}$ in order to decrease the number of nodes outside of ω, where the solutions are very smooth and regular. If ℓ and ω are well chosen, the solutions obtained should be almost the same using $\Omega_{\ell+1}$ or $\hat{\Omega}_{\ell+1}$ (see Section 5), but much less expensive to compute on the composite mesh.

The discretized problem on $\hat{\Omega}_{\ell+1}$ consists in looking for the p smallest eigenvalues $\hat{\epsilon}_{\ell+1}^{(j)}$, $\hat{\epsilon}_{\ell+1}^{(1)} \leq \hat{\epsilon}_{\ell+1}^{(2)} \leq \cdots \leq \hat{\epsilon}_{\ell+1}^{(p)}$ and the p corresponding eigenfunctions $\hat{\Psi}_{\ell+1}^{(j)}$, $j = 1, 2, \ldots, p$,

solutions of the eigenvalue problem

$$\hat{B}_{\ell+1}^{-1}\hat{H}_{\ell+1}^{KS}[\hat{\rho}_{\ell+1}]\hat{\Psi}_{\ell+1}^{(j)} = \hat{\epsilon}_{\ell+1}^{(j)}\hat{\Psi}_{\ell+1}^{(j)}, \tag{8i}$$

$$\hat{\rho}_{\ell+1}(x) = \sum_{j=1}^{p} 2\left|\hat{\Psi}_{\ell+1}^{(j)}(x)\right|^2, \quad x \in \hat{\Omega}_{\ell+1}, \tag{8ii}$$

$$\left(\hat{\Psi}_{\ell+1}^{(j)}, \hat{\Psi}_{\ell+1}^{(j)}\right)_{\widehat{\ell+1}} = \delta_{ij}, \qquad i = 1, \ldots, p \tag{8iii}$$

where $\hat{H}_{\ell+1}^{KS}[\hat{\rho}_{\ell+1}] = \hat{C}_{\ell+1} + \hat{B}_{\ell+1}\hat{V}_{\ell+1}[\hat{\rho}_{\ell+1}]$. Using the BHS nonlocal pseudopotentials [18], $\hat{V}_{\ell+1}$ can be defined by its action on a function $\hat{u}_{\ell+1} \in \hat{\mathcal{E}}_{\ell+1}$ which can be written in the form

$$(\hat{V}_{\ell+1}[\hat{\rho}_{\ell+1}]\hat{u}_{\ell+1})(x) = v[\hat{\rho}_{\ell+1}](x)\hat{u}_{\ell+1}(x) + \sum_{i=1}^{\bar{s}} \left(\hat{\phi}_{\ell+1}^{(i)}, \hat{u}_{\ell+1}\right)_{\widehat{\ell+1}} \hat{\phi}_{\ell+1}^{(i)}(x) \tag{9}$$

for $x \in \hat{\Omega}_{\ell+1}$, where v is the local part of the potential and $\hat{\phi}_{\ell+1}^{(i)}$, $i = 1, \ldots, \bar{s}$, are radial functions, each one centered on one atom and decreasing very rapidly to zero far from this atom.

If $\Omega \subset \omega$, the eigenvalue problem (8) is symmetric so that the eigenvalues are real and the eigenfunctions can be chosen real and orthogonal. If $\hat{\epsilon}_{\ell+1}^{(p+1)} > \hat{\epsilon}_{\ell+1}^{(p)}$, the problem (8) is then mathematically well defined. But if the refinement does not cover the whole domain, the operator $\hat{B}_{\ell+1}^{-1}\hat{H}_{\ell+1}^{KS}$ is not symmetric in general. Nevertheless, if the composite grid is fine enough to well describe the searched eigenfunctions, what should be the case in practice by the choice of ω, the computed solutions are almost the same on $\hat{\Omega}_{\ell+1}$ or $\Omega_{\ell+1}$ (see Section 5) so that we can expect the eigenvalues $\hat{\epsilon}_{\ell+1}^{(j)}$ being real and the eigenfunctions being in $\hat{\mathcal{E}}_{\ell+1}$. In the following we will assume it is like this and the orthogonality will be imposed only for the eigenfunctions associated to multiple eigenvalues. Actually the departure from this assumed ideal case is only due to the discretization error, the nondiscretized operator being symmetric.

3. BGII ALGORITHM FOR SOLVING THE EIGENVALUE PROBLEM

We propose to treat the nonlinear problem (8) by a classical relaxed fixed point algorithm also called self-consistent (SC) iteration. One step consists in solving the eigenvalue problem (8) for a fixed operator $\hat{B}_{\ell+1}^{-1}\hat{H}_{\ell+1}^{KS}$, that is a fixed potential $\hat{V}_{\ell+1}^{\text{old}}$. The solution of this problem defines a new density $\hat{\rho}_{\ell+1}^{\text{new}}$. Then the operator $\hat{H}_{\ell+1}^{KS}$ is updated, replacing the potential $\hat{V}_{\ell+1}^{\text{old}}$ by a linear mixing of $\hat{V}_{\ell+1}^{\text{old}}$ and $\hat{V}_{\ell+1}[\hat{\rho}_{\ell+1}^{\text{new}}]$. Iterations are performed until self-consistency is achieved. That is until the difference between $\hat{V}_{\ell+1}^{\text{old}}$ and $\hat{V}_{\ell+1}[\hat{\rho}_{\ell+1}^{\text{new}}]$ is smaller than a given tolerance. The purpose of the block Galerkin inverse iteration (BGII) algorithm we describe below is to solve efficiently the eigenvalue problem with linear operator (fixed potential) at each step of the SC scheme.

Let us assume that we know a good approximation $\hat{\mathcal{V}}_{\ell+1}$ of the p-dimensional subspace $\hat{\mathcal{U}}_{\ell+1}$ spanned by the p searched eigenfunctions $\hat{\Psi}_{\ell+1}^{(j)}$ on the refined grid $\hat{\Omega}_{\ell+1}$ for a given potential $\hat{V}_{\ell+1}$. We want to improve $\hat{\mathcal{V}}_{\ell+1}$ by a generalization of the RQI method. First of all we compute good initial eigenpairs for the process by performing a subspace diagonalization of $\hat{B}_{\ell+1}^{-1}\hat{H}_{\ell+1}^{KS}$ in $\hat{\mathcal{V}}_{\ell+1}$. In order not to have to invert $\hat{B}_{\ell+1}$, this is done by applying the Petrov–Galerkin method (see, for example, [22]) with $\hat{\mathcal{V}}_{\ell+1}$ as the right subspace and $\hat{B}_{\ell+1}^{*}\hat{\mathcal{V}}_{\ell+1}$

as the left subspace, where the exponent $*$ indicates we take the adjoint of the operator. It means that if the approximate subspace is spanned by p linear independent functions,

$$\hat{\mathcal{V}}_{\ell+1} = \operatorname{span}\{\hat{v}_{\ell+1}^{(1)}, \ldots, \hat{v}_{\ell+1}^{(p)}\},$$

we have to solve the p-dimensional matrix eigenvalue problems

$$G\mathbf{w}^{(j)} = \theta^{(j)} M\mathbf{w}^{(j)}, \tag{10}$$

where the matrix elements are given by

$$\begin{aligned} G_{ik} &= \big(\hat{B}_{\ell+1}^{*}\hat{v}_{\ell+1}^{(i)}, \hat{B}_{\ell+1}^{-1}\hat{H}_{\ell+1}^{KS}\hat{v}_{\ell+1}^{(k)}\big)_{\widehat{\ell+1}} = \big(\hat{v}_{\ell+1}^{(i)}, \hat{H}_{\ell+1}^{KS}\hat{v}_{\ell+1}^{(k)}\big)_{\widehat{\ell+1}}, \\ M_{ik} &= \big(\hat{B}_{\ell+1}^{*}\hat{v}_{\ell+1}^{(i)}, \hat{v}_{\ell+1}^{(k)}\big)_{\widehat{\ell+1}} = \big(\hat{v}_{\ell+1}^{(i)}, \hat{B}_{\ell+1}\hat{v}_{\ell+1}^{(k)}\big)_{\widehat{\ell+1}}. \end{aligned}$$

In general, if the subspace $\hat{\mathcal{V}}_{\ell+1}$ is not invariant, the matrix $M^{-1}G$ is not symmetric so that the eigenvalues $\theta^{(j)}$ can be complex. In order to only work with real numbers, we have to take it into account. If $\theta^{(j)}$ is real we simply choose

$$\tilde{\epsilon}_{\ell+1}^{(j)} = \theta^{(j)}$$

and define the new approximation of the jth eigenfunction as

$$\tilde{\Psi}_{\ell+1}^{(j)} = \sum_{k=1}^{p} \big(\mathbf{w}^{(j)}\big)_k \hat{v}_{\ell+1}^{(k)}, \tag{11}$$

where $(\mathbf{w}^{(j)})_k$ is the kth component of $\mathbf{w}^{(j)}$. If the eigenvalues $\theta^{(j)}$ and $\theta^{(j+1)}$ are complex conjugate, i.e. $\theta^{(j)} = \overline{\theta^{(j+1)}}$, we note $\mathbf{w}^{(j)} \pm i\mathbf{w}^{(j+1)}$ the associated eigenvectors—as returned by the LAPACK routine DGEEV [23] used to solve (10). In that case we treat the two vectors $\mathbf{w}^{(j)}$ and $\mathbf{w}^{(j+1)}$ as two basis vectors of a two-dimensional eigensubspace associated to a double eigenvalue $\operatorname{Re}(\theta^{(j)})$.

To improve the trial subspace $\hat{\mathcal{V}}_{\ell+1}$, we use the generalized inverse iteration reported in [5]. For the problem (8), a classical RQI iteration [24] can be written for the eigenpair j,

$$\big(\hat{B}_{\ell+1}^{-1}\hat{H}_{\ell+1}^{KS} - \tilde{\epsilon}_{\ell+1}^{(j)}\big)\frac{1}{\xi^{(j)}}\big(\tilde{\Psi}_{\ell+1}^{(j)} + \delta\tilde{\Psi}_{\ell+1}^{(j)}\big) = \tilde{\Psi}_{\ell+1}^{(j)}, \tag{12}$$

where $\delta\hat{\Psi}_{\ell+1}^{(j)}$, the correction to add to $\tilde{\Psi}_{\ell+1}^{(j)}$, is chosen to be orthogonal to $\tilde{\Psi}_{\ell+1}^{(j)}$ for the scalar product $(\cdot, \hat{B}_{\ell+1}\cdot)_{\widehat{\ell+1}}$, and $\xi^{(j)}$ is a real coefficient. Let us define the subspace

$$\hat{\mathcal{V}}_{\ell+1}^{(j)} = \operatorname{span}\big\{\tilde{\Psi}_{\ell+1}^{(j-\alpha_1(j))}, \ldots, \tilde{\Psi}_{\ell+1}^{(j+\alpha_2(j))}\big\} \subset \hat{\mathcal{E}}_{\ell+1}$$

for $\alpha_1(j)$ and $\alpha_2(j)$ integers given by

$$\begin{aligned} \alpha_1(j) &= \max_{1\le i\le j,\, \tilde{\epsilon}_{\ell+1}^{(j)} - \tilde{\epsilon}_{\ell+1}^{(i)} \le \delta} j - i, \\ \alpha_2(j) &= \max_{j\le i\le p,\, \tilde{\epsilon}_{\ell+1}^{(i)} - \tilde{\epsilon}_{\ell+1}^{(j)} \le \delta} i - j \end{aligned} \tag{13}$$

for a chosen parameter $\delta \in \mathbb{R}$. If we multiply (12) by $\hat{B}_{\ell+1}$, then project both sides of the equation onto the subspace $(\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$ and restrict its solution to be in the same subspace $(\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$, we obtain a new equation for $\delta\hat{\Psi}_{\ell+1}^{(j)} \in (\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$,

$$\tilde{Q}_{\ell+1}^{(j)}\big(\hat{H}_{\ell+1}^{KS} - \tilde{\epsilon}_{\ell+1}^{(j)}\hat{B}_{\ell+1}\big)\delta\hat{\Psi}_{\ell+1}^{(j)} = -\tilde{Q}_{\ell+1}^{(j)}\big(\hat{H}_{\ell+1}^{KS} - \tilde{\epsilon}_{\ell+1}^{(j)}\hat{B}_{\ell+1}\big)\tilde{\Psi}_{\ell+1}^{(j)}, \tag{14}$$

where $\tilde{Q}_{\ell+1}^{(j)}$ is the orthogonal projector onto $(\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$. Equations (12) and (14) are equivalent if $\alpha_1(j) = \alpha_2(j) = 0$. In this case the iterative scheme defined by $\tilde{\Psi}_{\ell+1}^{(j)} \leftarrow \tilde{\Psi}_{\ell+1}^{(j)} + \delta\hat{\Psi}_{\ell+1}^{(j)}$, $j = 1, \ldots, p$, with $\delta\hat{\Psi}_{\ell+1}^{(j)}$ the solution of (14), is the classical RQI iteration. But the linear systems can be singular or very ill-conditioned if we have multiple or close eigenvalues. As pointed out by Hackbusch [16], it is important to have a well-posed problem for the multigrid resolution we have in view (see Section 4). Within our formulation, the larger δ is chosen, the more the linear system defined by (14) is well-conditioned.

Combining the PG method and the generalization (14) of the inverse iteration, we define the following algorithm.

ALGORITHM BGII.

(a) Compute $(\tilde{\epsilon}_{\ell+1}^{(j)}, \tilde{\Psi}_{\ell+1}^{(j)}) \in \mathbb{R} \times \hat{\mathcal{E}}_{\ell+1}$, $j = 1, \ldots, p$, by resolution of the PG problem (10) in $\hat{\mathcal{V}}_{\ell+1}$.

(b) Compute the solution $\delta\hat{\Psi}_{\ell+1}^{(j)}$ of Eq. (14) for $j = 1, \ldots, p$.

(c) Compute $\hat{v}_{\ell+1}^{(j)} = \tilde{\Psi}_{\ell+1}^{(j)} + \delta\hat{\Psi}_{\ell+1}^{(j)}$, $j = 1, \ldots, p$.

(d) Compute $(\tilde{\epsilon}_{\ell+1}^{(j)}, \tilde{\Psi}_{\ell+1}^{(j)}) \in \mathbb{R} \times \hat{\mathcal{E}}_{\ell+1}$, $j = 1, \ldots, p$, by resolution of the PG problem (10) in $\hat{\mathcal{V}}_{\ell+1}$.

(e) Test for convergence. Goto (b) if not satisfied.

At step (e), we check for the convergence by evaluating the norm of the residuals

$$\hat{r}_{\ell+1}^{(j)} = \big(\hat{A}_{\ell+1} - \tilde{\epsilon}_{\ell+1}^{(j)}\hat{B}_{\ell+1}\big)\tilde{\Psi}_{\ell+1}^{(j)}, \quad j = 1, \ldots, p. \tag{15}$$

It can be shown in the Galerkin case, i.e. when the left subspace is chosen to be $\hat{\mathcal{V}}_{\ell+1}^{(j)}$ instead of $\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)}$ and when $\hat{B}_{\ell+1}^{-1}\hat{H}_{\ell+1}^{KS}$ is symmetric, that the convergence of that algorithm is locally quadratic, provided that the starting subspace is sufficiently accurate [25].

4. MULTIGRID METHOD

At step (b) of the BGII algorithm, we need an efficient tool to solve the linear systems. In this purpose we propose a multigrid method adapted to the composite grids. Let us first describe the idea of the multigrid algorithm for the resolution of a Poisson problem discretized on $\hat{\Omega}_{\ell+1}$. Assume we know some approximation $\tilde{v}_{\ell+1}$ of the solution $\hat{v}_{\ell+1}$ of Eq. (7). Let us define the defect

$$\hat{d}_{\ell+1} = \hat{C}_{\ell+1}\tilde{v}_{\ell+1} - 4\pi\hat{B}_{\ell+1}\hat{\rho}_{\ell+1} \in \hat{\mathcal{E}}_{\ell+1}. \tag{16}$$

To describe a multigrid iteration, we introduce the grid transfer operators $\hat{R}_{\ell+1} : \hat{\mathcal{E}}_{\ell+1} \to \mathcal{E}_{\ell}$, an operator of weighted restriction, and $\hat{P}_{\ell} : \mathcal{E}_{\ell} \to \hat{\mathcal{E}}_{\ell+1}$, an operator of prolongation. Instead of solving Eq. (7) directly, we look for an iterative solution of the equation with the defect

as a right-hand side,

$$\hat{C}_{\ell+1}\hat{z}_{\ell+1} = \hat{d}_{\ell+1}, \tag{17}$$

starting from $\hat{z}_{\ell+1} \equiv 0$ and updating $\hat{v}_{\ell+1} = \hat{v}_{\ell+1} - \hat{z}_{\ell+1}$ and $\hat{d}_{\ell+1}$ after each correction.

To do this we write an algorithm based on the strategy proposed in [16, Section 15.2]. The main difference comes from the way to deal with the nodes on the interface between the refined and nonrefined parts of $\hat{\Omega}_{\ell+1}$. A heuristic justification of the method is based on a decomposition of $\hat{z}_{\ell+1}$ in components of "high frequency" and "low frequency" (relative to the grid Ω_ℓ). The "low frequency" components are those we can represent on Ω_ℓ, while the "high frequency" components are functions in $\hat{\mathcal{E}}_{\ell+1}$ which need a finer grid to be well represented and whose nonzero values are localized in ω by definition of $\hat{\mathcal{E}}_{\ell+1}$. The former will be reduced by solving a coarse grid approximation of Eq. (17) corresponding to a problem without refinement,

$$C_\ell z_\ell = \hat{R}_{\ell+1}\hat{d}_{\ell+1}. \tag{18}$$

The latter will be reduced by a smoothing process applied to Eq. (17) restricted to the subspace $\mathcal{E}^{\omega}_{\ell+1} \subset \hat{\mathcal{E}}_{\ell+1}$ of the functions whose values are zero outside of $\bar{\omega}_{\ell+1}$,

$$\begin{aligned} \Pi^{\omega}_{\ell+1}\hat{C}_{\ell+1}z^{(\omega)}_{\ell+1} &= \Pi^{\omega}_{\ell+1}\hat{d}_{\ell+1}, \\ z^{(\omega)}_{\ell+1} &\in \mathcal{E}^{\omega}_{\ell+1}. \end{aligned} \tag{19}$$

Here $\Pi^{\omega}_{\ell+1} : \hat{\mathcal{E}}_{\ell+1} \to \mathcal{E}^{\omega}_{\ell+1}$ denotes the orthogonal projector from $\hat{\mathcal{E}}_{\ell+1}$ onto $\mathcal{E}^{\omega}_{\ell+1}$.

By analogy with a standard multigrid V-cycle, we write the algorithm in the following form

TWO-GRID ALGORITHM FOR THE POISSON PROBLEM

(a) Perform ν_1 presmoothing iterations on Eq. (19).
(b) Update $\tilde{v}_{\ell+1} \leftarrow \tilde{v}_{\ell+1} - z^{(\omega)}_{\ell+1}$ and $\hat{d}_{\ell+1}$.
(c) Solve the coarse grid equation (18).
(d) Update $\tilde{v}_{\ell+1} \leftarrow \tilde{v}_{\ell+1} - \hat{P}_\ell z_\ell$ and $\hat{d}_{\ell+1}$.
(e) Perform ν_2 postsmoothing iterations on Eq. (19).
(f) Update $\tilde{v}_{\ell+1} \leftarrow \tilde{v}_{\ell+1} - z^{(\omega)}_{\ell+1}$ and $\hat{d}_{\ell+1}$.

This algorithm can be generalized to more than two levels by solving the coarse grid problem at step (c) with a classic multigrid V-cycle, i.e. the same algorithm for the particular case $\Omega \subset \omega$. Numerical results show that the rate of convergence is the same if the finer grid is composite or regular (i.e. refined everywhere) [6].

Remark 4.1. Since the values on all the nodes are defined by the unique solution of the Poisson's equation discretized on the composite grid (and not by some extension of the solutions of a Poisson's equation on the coarse grid), we do not need to ensure that the equation on the coarse grid to be conservative as in [26]. Here the use of a global coarse grid and a local fine grid are just tools to solve efficiently the problem on the composite grid.

We will solve iteratively the generalized inverse iteration equation (14) for a given j by an algorithm similar to the one presented for the Poisson problem, i.e. by combining

corrections coming from "local smoothing" and "coarse grid corrections." Starting from an approximation $\delta\tilde{\Psi}_{\ell+1}^{(j)} \in (\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$ of $\delta\hat{\Psi}_{\ell+1}^{(j)}$, the search solution of (14), let us define the defect

$$\hat{d}_{\ell+1}^{(j)} = \big(\hat{H}_{\ell+1}^{KS} - \tilde{\epsilon}_{\ell+1}^{(j)}\hat{B}_{\ell+1}\big)\big(\tilde{\Psi}_{\ell+1}^{(j)} + \delta\tilde{\Psi}_{\ell+1}^{(j)}\big). \tag{20}$$

We look for $\hat{z}_{\ell+1}^{(j)} \in (\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$ such that

$$\delta\hat{\Psi}_{\ell+1}^{(j)} = \hat{z}_{\ell+1}^{(j)} + \delta\tilde{\Psi}_{\ell+1}^{(j)}, \tag{21}$$

starting from the approximation $\hat{z}_{\ell+1}^{(j)} \equiv 0$. We define $Q_{\ell}^{(j,R)} : \mathcal{E}_{\ell} \to (\hat{R}_{\ell+1}\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$ the orthogonal projector onto $(\hat{R}_{\ell+1}\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$. The function $z_{\ell}^{(j)} \in (\hat{R}_{\ell+1}\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)})^{\perp}$, solution of the equation

$$Q_{\ell}^{(j,R)}\big(H_{\ell+1}^{KS} - \tilde{\epsilon}_{\ell+1}^{(j)}B_{\ell}\big)z_{\ell}^{(j)} = -Q_{\ell}^{(j,R)}\hat{R}_{\ell+1}\hat{d}_{\ell+1}^{(j)} \tag{22}$$

can be considered as a coarse grid correction for a multigrid resolution of (14) by the update $\delta\tilde{\Psi}_{\ell+1}^{(j)} \leftarrow \tilde{Q}_{\ell+1}^{(j)}\hat{P}_{\ell}z_{\ell}^{(j)} + \delta\tilde{\Psi}_{\ell+1}^{(j)}$. As before, "high frequency" components are supposed to be localized in ω. By first restricting the RQI equation (12) multiplied by $\hat{B}_{\ell+1}$ to $\mathcal{E}_{\ell+1}^{\omega}$ and then introducing an orthogonal projector $\tilde{Q}_{\ell+1}^{(j,\omega)} : \mathcal{E}_{\ell+1}^{\omega} \to \tilde{\mathcal{W}}_{\ell+1}^{(j,\omega)}$ onto the subspace

$$\tilde{\mathcal{W}}_{\ell+1}^{(j,\omega)} = \mathcal{E}_{\ell+1}^{\omega} \cap \big(\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)}\big)^{\perp}, \tag{23}$$

in the same way as for the problem on the whole grid, we obtain an equation for $z_{\ell+1}^{(j,\omega)} \in \tilde{\mathcal{W}}_{\ell+1}^{(j,\omega)}$:

$$\tilde{Q}_{\ell+1}^{(j,\omega)}\Pi_{\ell+1}^{\omega}\big(\hat{H}_{\ell+1}^{KS} - \tilde{\epsilon}_{\ell+1}^{(j)}\hat{B}_{\ell+1}\big)z_{\ell+1}^{(j,\omega)} = -\tilde{Q}_{\ell+1}^{(j,\omega)}\Pi_{\ell+1}^{\omega}\hat{d}_{\ell+1}^{(j)}. \tag{24}$$

Smoothing on this equation will provide a fine grid correction $z_{\ell+1}^{(j,\omega)}$ to add to $\hat{z}_{\ell+1}^{(j)}$. In Eq. (24), the operator is not positive definite and can be singular, unlike the operator on the whole composite grid. Nevertheless it will not be a problem if we only apply smoothing iterations on it with GMRES [27]. We verify that the action of the projector $\tilde{Q}_{\ell+1}^{(j,\omega)}\Pi_{\ell+1}^{\omega}$ on a function $\hat{u}_{\ell+1}$ in $\hat{\mathcal{E}}_{\ell+1}$ consists in restricting $\hat{u}_{\ell+1}$ to $\bar{\omega}_{\ell+1}$, then subtracting its projection onto the subspace spanned by the functions of $\hat{B}_{\ell+1}\hat{\mathcal{V}}_{\ell+1}^{(j)}$ restricted to $\bar{\omega}_{\ell+1}$, and finally extending the function obtained to $\hat{\Omega}_{\ell+1}$ with value zero outside of $\bar{\omega}_{\ell+1}$. The coarse and fine grid corrections described above are combined in the following algorithm:

TWO-GRID ALGORITHM FOR THE KS EQUATION

(a) Perform ν_1 GMRES iterations on equation (24) starting from $z_{\ell+1}^{(j,\omega)} = 0$.
(b) Update $\delta\tilde{\Psi}_{\ell+1}^{(j)} \leftarrow \delta\tilde{\Psi}_{\ell+1}^{(j)} + z_{\ell+1}^{(j,\omega)}$ and $\hat{d}_{\ell+1}^{(j)}$.
(c) Solve the coarse grid equation (22).
(d) Update $\delta\tilde{\Psi}_{\ell+1}^{(j)} \leftarrow \delta\tilde{\Psi}_{\ell+1}^{(j)} + \tilde{Q}_{\ell+1}^{(j)}\hat{P}_{\ell}z_{\ell}^{(j)}$ and $\hat{d}_{\ell+1}^{(j)}$.
(e) Perform ν_2 GMRES iterations on equation (24) starting from $z_{\ell+1}^{(j,\omega)} = 0$.
(f) Update $\delta\tilde{\Psi}_{\ell+1}^{(j)} \leftarrow \delta\tilde{\Psi}_{\ell+1}^{(j)} + z_{\ell+1}^{(j,\omega)}$ and $\hat{d}_{\ell+1}^{(j)}$.

This two-grid algorithm can be generalized to a multigrid algorithm by applying recursively the same method for solving the coarse grid problem (case $\Omega \subset \omega$). In this case, the

projector $\Pi^{\omega}_{\ell+1}$ is equal to the identity, while the shifts $\tilde{\epsilon}^{(j)}_{\ell+1}$ and the projectors $\tilde{Q}^{(j)}_{\ell+1}$ are defined by the eigenpairs solutions of the stationary Schrödinger problem on the corresponding grid with the potential defined by a restriction of the one used on the finest grid. We note that if ω is a cube whose side length is $L/2$ (L being the length of the side of Ω) the cost of one local smoothing iteration on the grid $\hat{\Omega}_{\ell+1}$ is almost the same as the cost of one smoothing iteration on Ω_{ℓ}. But it is about two times cheaper that one iteration on the whole composite grid $\hat{\Omega}_{\ell+1}$ and eight times cheaper that one iteration on a grid $\Omega_{\ell+1}$ (i.e., refined everywhere).

5. NUMERICAL TESTS OF CONVERGENCE

The whole resolution of the problem (8) by the proposed algorithm involves three levels of iterations: SC, BGII, and the multigrid method for the resolution of the linear systems in the inverse iterations. Here we are primarily concerned with the BGII and multigrid parts of the method, the original ones. A local quadratic convergence rate can be proved for the BGII algorithm when the linear systems are solved exactly and when the Petrov–Galerkin process is replaced by a Galerkin one (i.e., using $\hat{\mathcal{V}}_{\ell+1}$ as the left subspace) for a symmetric operator $\hat{B}^{-1}_{\ell+1}\hat{H}^{KS}_{\ell+1}$, but theoretical convergence rates are difficult to obtain for the multigrid resolution of the linear systems such as Eq. (14). Nevertheless the efficiency of these two algorithms in practical problems can be assessed by isolating the different parts of the method and measuring their convergence numerically. This has be done on problems appearing in the electronic structure calculation of molecules with the proposed method.

The first example we present is the furan molecule (C_4OH_4) discretized on a grid $\hat{\Omega}_6$, where Ω is a cell of side 15 Bohr radii and ω is a cell of side 8 Bohr centered in Ω. (One Bohr radius is approximately 0.529 Å) This problem involves the self-consistent search of 13 eigenfunctions. The whole calculation of the electronic structure of the molecule for given atomic positions (as given in [28]) is made by using the coarser grids to initiate the functions and potentials. The initial approximations of the eigenfunctions and of the nonlinear potential on a given grid are obtained by extension of the ones obtained on the previous coarser grid, after applying a first PG process that also furnishes the initial eigenvalues (or shifts). If these approximate eigenfunctions are too far from the solutions corresponding to the starting potential, a block variant of the Jacobi–Davidson method is applied for the first SC step. This is avoided in the next SC steps by controlling the variation of the potential (keeping it relatively small) and using the last computed solutions as starting vectors. Figure 2 shows the convergence of the total energy $E[\{\Psi^{(j)}\}]$ of the molecule, during the computation on the successive grids Ω_4, Ω_5, and $\hat{\Omega}_6$. The convergence criterion on each grid is the relative L_2 norm of the variation of the nonlinear part of the potential (Hartree and exchange-correlation parts). In the present example, it has been fixed to 10^{-3}. Here, only one BGII iteration is applied in a SC step. In Table I, we give the eigenvalues and some energies obtained on the grids Ω_4, Ω_5, and $\hat{\Omega}_6$, the ones we use in the multigrid algorithm. They are compared to numerical values obtained for the same model with a different method. E^{GS} denotes the ground state energy of the system and E^{cin} the kinetic energy of the electrons.

In order to investigate its convergence rate, the two-grid algorithm for the KS equation defined in Section 4 is applied for the resolution of linear problems like (14). Numerical

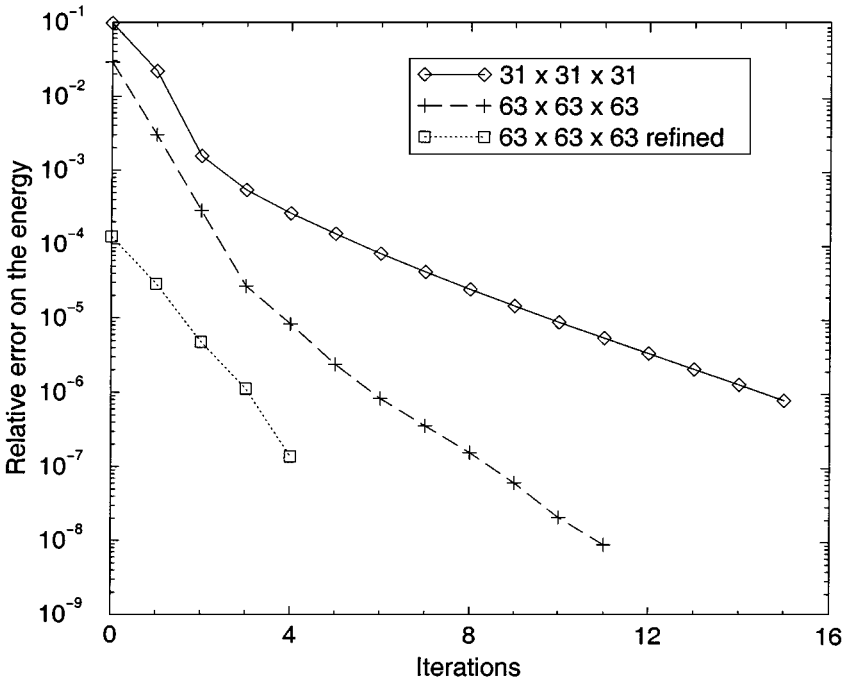

FIG. 2. Relative error on the energy $E[\{\Psi^{(j)}\}]$ as a function of SC iterations during the ground state search on the three finest grids used in the calculation.

experiments are performed using operators obtained in the electronic structure calculation of the furan molecule when using BGII, between two SC iterations. The solutions of the linear problems are chosen a priori with values chosen randomly on $\hat{\Omega}_6$ and define the right-hand side for the tests. The coarse grid problems on Ω_5 are approximately solved by the same two-grid algorithm. For the tests, we use very accurate solutions for the coarse

TABLE I
Eigenvalues and Energies of the Furan Molecule as Obtained from Our Computation on Different Grids

	Ω_4	Ω_5	$\hat{\Omega}_6$	[29]
$\epsilon^{(1)}$	−2.644	−2.028	−2.055	−2.041
$\epsilon^{(2)}$	−1.801	−1.444	−1.452	−1.439
$\epsilon^{(3)}$	−1.281	−1.351	−1.356	−1.342
$\epsilon^{(4)}$	−1.157	−1.086	−1.089	−1.075
$\epsilon^{(5)}$	−1.001	−1.016	−1.019	−1.005
$\epsilon^{(6)}$	−0.942	−0.978	−0.982	−0.969
$\epsilon^{(7)}$	−0.926	−0.829	−0.828	−0.814
$\epsilon^{(8)}$	−0.784	−0.795	−0.795	−0.782
$\epsilon^{(9)}$	−0.762	−0.722	−0.723	−0.710
$\epsilon^{(10)}$	−0.687	−0.712	−0.712	−0.699
$\epsilon^{(11)}$	−0.578	−0.657	−0.660	−0.646
$\epsilon^{(12)}$	−0.370	−0.518	−0.516	−0.502
$\epsilon^{(13)}$	−0.367	−0.423	−0.423	−0.411
E^{cin}	107.31	65.231	64.022	63.970
E^{GS}	−94.042	−82.500	−82.251	−82.268

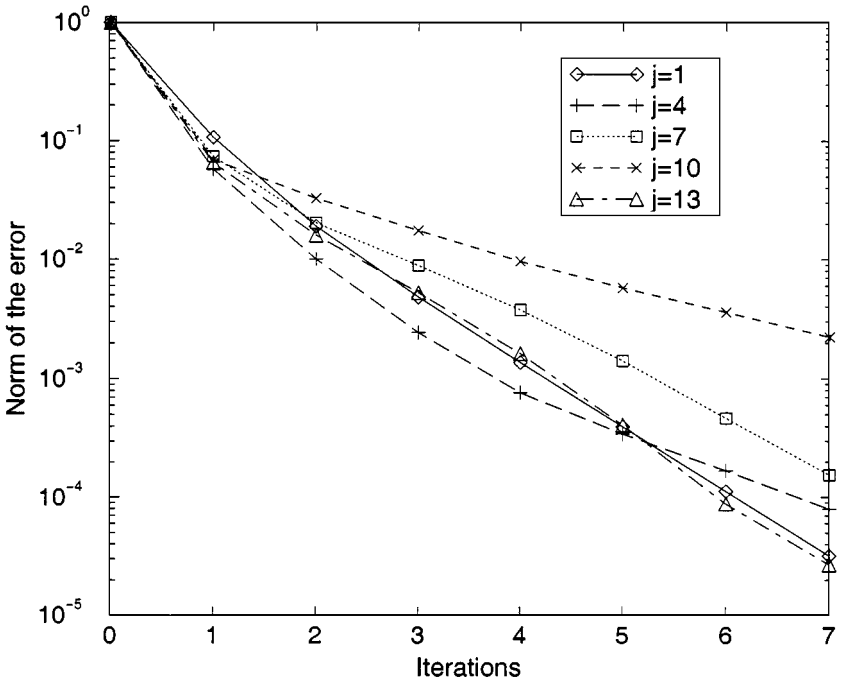

FIG. 3. Convergence of the error for the multigrid resolution of linear systems appearing in the electronic structure calculation of the furan molecule by BGII, for different eigenpairs of indices j. The norm of the error is plotted as a function of the number of cycles of the algorithm.

grid problem on Ω_4. But in practice, the resolution of the problem on the coarsest grid ($31 \times 31 \times 31$) is not trivial. However, an approximate solution obtained after a limited number of GMRES iterations ($\sim$20) is in general sufficient. We so apply a tri-grid algorithm. Indeed we observed that the employ of coarser grids for the potentials we use damage the convergence rate. Figure 3 shows the norm of the error after each cycle. The index j indicates to which eigenvalue are related the shift and projector used in the operators. For the BGII algorithm we report in Fig. 4 the minimum, maximum, and mean values of the norm of the residuals, as defined in Eq. (15), measured on the 13 simultaneously computed eigenpairs, for initial approximations and potential obtained between two SC iterations of the whole calculation. In this test, the linear systems obtained in the inverse iteration part of the algorithm are only approximately solved by two cycles of the two-grid algorithm (with $\nu_1 = \nu_2 = 2$). Note that all the results for the furan molecule have been obtained using a coefficient $\delta = 0.1$[Ry] in (13), leading to subspaces $\hat{\mathcal{V}}_{\ell+1}^{(j)}$ of dimension between 1 and 4 (see Table I).

A good way to check the accuracy of the discretization scheme proposed in this paper is to compute the electronic structure of a single atom and to compare the numerical results with a calculation using the same model (LDA) and pseudopotentials (BHS), but solving the radial equations in the spherical approximation (for an equidistribution of the electrons in the highest occupied orbitals). That way, we have a very precise estimation of the results we try to reproduce in our 3D calculation. Actually, the only differences between the two approaches for fully converged results are the boundary conditions. In Table II, we present the results obtained for the oxygen atom, an element which requires a very fine discretization. We can observe the good agreement between the results obtained on a grid $127 \times 127 \times 127(\Omega_6)$ and a grid $63 \times 63 \times 63$ refined in the center ($\hat{\Omega}_6$).

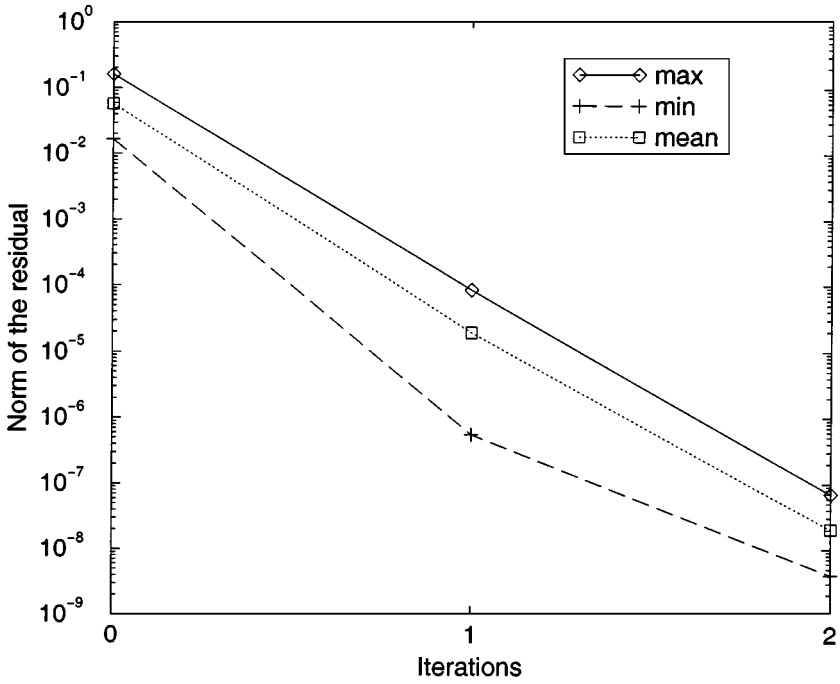

FIG. 4. Illustration of the convergence of the residuals for the BGII algorithm. The convergence is evaluated for an eigenvalue problem with fixed potential between two SC iterations in the electronic structure calculation of the furan molecule. In the inverse iterations, the linear systems are approximately solved by two cycles of the multigrid algorithm with $\nu_1 = \nu_2 = 2$.

As an example of the agreement between the eigenfunctions computed on regular and refined grids, we show in Table III the results obtained for the silicon atom without interaction between electrons. Actually, we look at the first eigenpair of a stationary Schrödinger problem with the silicon atom pseudopotential. In this case, the single eigenfunction computed on different grids can be compared. $I_{\ell+1} : \mathcal{E}_{\ell+1} \to \mathcal{E}_\ell$, $\ell = 4, 5$, denotes the restriction operator defined by $(I_{\ell+1} v_{\ell+1})(x) = v_{\ell+1}(x)$, $x \in \Omega_\ell$, $v_{\ell+1} \in \mathcal{E}_{\ell+1}$, and $\hat{I}_6 : \mathcal{E}_6 \to \hat{\mathcal{E}}_6$, denotes the restriction operator defined by $(\hat{I}_6 v_6)(x) = v_6(x)$, $x \in \hat{\Omega}_6$, $v_6 \in \mathcal{E}_6$.

To show the importance of the grid refinement when looking for the equilibrium positions of the atoms in a molecule, we present and compare in Fig. 5 the ground-state energy of the

TABLE II

Numerical Results for the Oxygen Atom (in Rydberg), Centered in a Cubic Cell Ω of Dimensions (12 Bohr)^3, for Different Discretization Grids, Compared to a Result Obtained for the Radial Problem (1D) for the Same Model (LDA + Oxygen BHS Pseudopotential). The Refinement Domain is a Cube of Dimensions (6 Bohr)^3 Centered on the Ion

	Ω_4	Ω_5	$\hat{\Omega}_6$	Ω_6	Radial
$\epsilon^{(1)}$	−1.87774	−1.72484	−1.74379	−1.74345	−1.74645
$\epsilon^{(2)}$	−0.62944	−0.67738	−0.67299	−0.67266	−0.67409
$\epsilon^{(3)}$	−0.62944	−0.67738	−0.67299	−0.67266	−0.67409
$\epsilon^{(4)}$	−0.62944	−0.67738	−0.67299	−0.67266	−0.67409
E^{GS}	−31.0485	−31.6008	−31.4994	−31.4989	−31.4936
E^{cin}	21.0788	24.0349	23.5281	23.5286	23.4697

TABLE III

First Eigenvalue Obtained for the Pseudopotential of a Silicon Atom without Any Interaction Between Electrons, and Comparison of the First Eigenfunction (Normalized) Obtained on the Grids Ω_4, Ω_5, $\hat{\Omega}_6$, Ω_6

$\epsilon_4^{(1)}$	$\epsilon_5^{(1)}$	$\epsilon_6^{(1)}$	$\hat{\epsilon}_6^{(1)}$	$\epsilon^{(1)}$(1D)
−3.32311963	−3.31844813	−3.31823856	−3.31823816	−3.318226
	$\|u_4^{(1)} - I_5 u_5^{(1)}\|_4$	$\|u_5^{(1)} - I_6 u_6^{(1)}\|_5$	$\|\hat{u}_6^{(1)} - \hat{I}_6 u_6^{(1)}\|_5$	
	4.8e-03	2.3e-04	4.1e-06	

CO molecule for different interatomic distances obtained on a regular grid and on a refined grid. The computation has been performed in a cubic cell of side 16 Bohr, using a regular grid of $63 \times 63 \times 63$ nodes. The refinement has been done in a cube of side 8 Bohr, centered in the cell. For the computation without refinement, we observe the presence of two local minimas, separated by a distance close to the grid spacing $h_{63} = 0.25$ Bohr, manifesting a lack of translational invariance. The same phenomena has been observed in [7] for the CO_2 molecule. The introduction of the mesh refinement corrects this problem and the regular curve obtained has a single minima we can estimate by a quadratic fit to be at 2.129 Bohr. This value is very close to the value obtained in [7] (2.132 Bohr) with the same model (LDA and BHS pseudopotentials). The results clearly show the improvement in the determination of the ground state of the molecule by using the refined grid, at a much lower cost that on a regular grid of $127 \times 127 \times 127$ nodes.

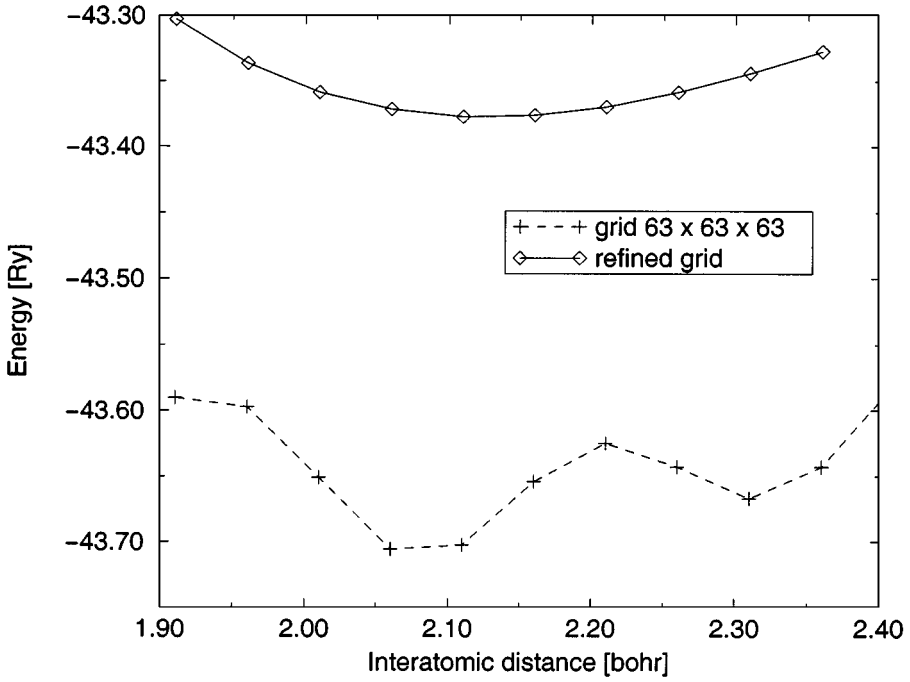

FIG. 5. Ground state energy of the CO molecule computed on grids with and without refinement, as a function of the interatomic space.

6. CONCLUDING REMARKS

In this paper we have shown the applicability of a collection of numerical techniques to electronic structure calculations problems. In particular we have shown that algorithms based on the inverse iteration, using multigrid techniques for the resolution of the linear systems, can be very efficient. This has been done by applying the algorithm BGII as described above. Nevertheless some modifications can be considered in the method in order to reduce the computational cost for very large systems.

The eigensolver presented here scales as $p^2 \times N$ (where N denotes the number of nodes on the grid and p is the number of searched eigenpairs) like most of the methods currently used in this field to treat problems involving nonlocalized orbitals, i.e. systems with wavefunctions that are nonzero in the whole domain of computation. The bottleneck is in general the Galerkin part (or orthogonalization) which requires the computation of matrix elements associated to every pair of eigenvectors followed by an update of the different approximations as a linear combination of the p trial eigenvectors. But since the Rayleigh quotient-like iterations drive naturally the eigenvector approximations to vectors contained in eigensubspaces, we can consider a future implementation of the BGII algorithm that would require only the matrix elements between vectors associated to close or multiple eigenvalues (at least the ones associated to the vectors pairs contained in $\hat{\mathcal{V}}_{\ell+1}^{(j)}$, the other ones converging naturally to zero) and that would build the new vectors as linear combinations of a small number of vectors (see Eq. (11)). Such an approach could reduce drastically the cost of the method on systems much larger than the ones treated in this paper. Unfortunately, a reduction of the asymptotic scaling of the method, when the number of eigenpairs becomes very large, is difficult to obtain. Indeed the density of the eigenvalues in the part of the spectrum in which we are interested is likely to increase in this case and thus require us to enlarge the dimension of the subspaces $\hat{\mathcal{V}}_{\ell+1}^{(j)}$ in order to keep the same conditioning of the linear systems.

However, the present algorithm BGII has the advantage over the conjugate gradient (CG) like algorithms with simultaneous update of all the eigenfunctions [30] that it requires much less memory storage. In our implementation, when computing the correction for the vector j, we chose to store the vectors $\hat{B}_{\ell+1}\tilde{\Psi}_{\ell+1}^{(j-\alpha_1(j))}, \ldots, \hat{B}_{\ell+1}\tilde{\Psi}_{\ell+1}^{(j+\alpha_2(j))}$ defining the projector $\tilde{Q}_{\ell+1}^{(j)}$. In this case, the maximal number of vectors to store (for the eigenfunctions and their corrections) at the same time is only: p eigenfunction approximations $+ (\max_j(\alpha_1(j) + \alpha_2(j)) + 1)$ functions defining the projectors $+ 1$ correction currently computed. Moreover, the optimal coefficient multiplying a correction before adding it to the corresponding eigenvector (often called the time-step) is always 1 in BGII, while it can be expensive to compute in a CG approach or difficult to estimate (and very small) in a steepest descent algorithm.

Finally, we mention that the BGII method described in this paper has also been applied with success to the computation of the electronic structure of the C_{60} molecule [31], using a regular grid of $79 \times 79 \times 79$ nodes, showing its applicability to problems involving more than 100 eigenvalues.

ACKNOWLEDGMENTS

The author is very indebted to Professor J. Descloux and Dr. F. Gygi for numerous valuable discussions about this work.

REFERENCES

1. J. Bernholc, J.-Y. Yi, and D. J. Sullivan, Structural transitions in metal clusters, *Faraday Soc. Discuss. Chem. Soc.* **92**, 217 (1991).
2. J. R. Chelikowsky, N. Trouiller, and Y. Saad, Finite-difference-pseudopotential method: Electronic structure calculations without a basis, *Phys. Rev. Lett.* **72**(8), 1240 (1994).
3. E. L. Briggs, D. J. Sullivan, and J. Bernholc, Large scale electronic structure calculations with multigrid acceleration, *Phys. Rev. B* **52**(8), 5471 (1995).
4. E. L. Briggs, D. J. Sullivan, and J. Bernholc, Real-space multigrid-based approach to large-scale electronic structure calculations, *Phys. Rev. B* **54**(20), 14362 (1996).
5. J.-L. Fattebert, An inverse iteration method using multigrid for quantum chemistry, *BIT* **36**(3), 509 (1996).
6. J.-L. Fattebert, *Une méthode numérique pour la résolution des problèmes aux valeurs propres liés au calcul de structure électronique moléculaire*, Ph.D. thesis, Thèse No. 1640, Ecole Polytechnique Fédérale de Lausanne, 1997.
7. F. Gygi and G. Galli, Real-space adaptive-coordinate electronic-structure calculations, *Phys. Rev. B* **52**(4), 2229 (1995).
8. E. Tsuchida and M. Tsukada, Real space approach to electronic-structure calculations, *Solid State Com.* **94**(1), 5 (1995).
9. T. L. Beck, K. A. Iver, and M. P. Merrick, Multigrid methods in density functional theory, *Int. J. Quant. Chem.* **61**, 341 (1997).
10. N. A. Modine, G. Zumbach, and E. Kaxiras, Adaptive-coordinate real-space electronic-structure calculations for atoms, molecules, and solids, *Phys. Rev. B* **55**(16), 10289 (1997).
11. L. Collatz, *The Numerical Treatment of Differential Equations* (Springer, Berlin, 1966).
12. E. R. Davidson, The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices, *J. Comput. Phys.* **17**, 87 (1975).
13. M. Crouzeix, B. Philippe, and M. Sadkane, The Davidson method, *SIAM J. Sci. Comput* **15**(1), 62 (1994).
14. G. L. G. Sleijpen and H. A. Van Der Vorst, A generalized Jacobi–Davidson iteration method for linear eigenvalue problem, *SIAM Matrix Anal. Appl.* **17**(2), 401 (1996).
15. A. Brandt, Multilevel adaptative solutions to boundary-value problems, *Math. Comp.* **31**(138), 333 (1977).
16. W. Hackbusch, *Multi-grid Methods and Applications* (Springer, Berlin, 1985).
17. W. E. Pickett, Pseudopotential methods in condensed matter applications, *Comput. Phys. Rep.* **9**, 115 (1989).
18. G. B. Bachelet, D. R. Hamman, and M. Schlueter, Pseudopotentials that work: From H to Pu, *Phys. Rev. B* **26**, 4199 (1982).
19. A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations* (Wiley, New York, 1980).
20. E. Isaacson and H. B. Keller, *Analysis of Numerical Methods* (Wiley, New York, 1966).
21. R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer–Verlag, Berlin, 1990).
22. Y. Saad, *Numerical Methods for Large Eigenvalue Problems* (Manchester Univ. Press, Manchester, 1992).
23. E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen, *LAPACK User's Guide, Release 1.0* (SIAM, Philadelphia, 1992).
24. B. N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, NJ, 1980).
25. J.-L. Fattebert, A block Rayleigh quotient iteration with local quadratic convergence, *Elec. Trans. Num. Anal.*, in press.
26. D. Bai and A. Brandt, Local mesh refinement multilevel techniques, *SIAM J. Sci. Stat. Comput.* **8**(2), 109 (1987).
27. Y. Saad and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Stat. Comput.* **7**(3), 856 (1986).
28. F. Gygi, Ab-initio molecular dynamics of organic compounds on a massively parallel computer, in *Proceedings of the 1995 Fall Meeting of the Materials Research Society: Materials Theory, Simulations, and Parallel Algorithms, of MRS Symposium Proceedings Series*, Vol. 408 (Materials Res. Soc., Pittsburgh, PA, 1996).

29. F. Gygi, personal communication. [Results obtained with a basis of 64^3 plane waves, with periodic boundary conditions and adaptive coordinates (according to [32]) corresponding to a mesh refinement of a factor 2 in the regions where the functions vary rapidly]
30. I. Stich, R. Car, M. Parrinello, and S. Baroni, Conjugate gradient minimization of the energy functional: A new method for electronic structure calculation. *Phys. Rev. B* **39**(8), 4997 (1989).
31. J. Descloux, J.-L. Fattebert, and F. Gygi, RQI (Rayleigh quotient iteration), an old recipe for solving modern large scale eigenvalue problems, *Comput. Phys.* **12**(1), 22 (1998).
32. F. Gygi, Electronic-structure calculations in adaptive coordinates, *Phys. Rev. B* **48**(16), 11692 (1993).