Evolution-operator method for density functional theory

E. R. Hernández,[1] S. Janecek,[2] M. Kaczmarski,[1,3] and E. Krotscheck[2]

[1]*Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus de Bellaterra, 08193 Barcelona, Spain*
[2]*Institut für Theoretische Physik, Johannes Kepler Universität, Altenbergstrasse 69, A 4040 Linz, Austria*
[3]*Institute of Physics, University of Silesia, Uniwersytecka 4, 40-007 Katowice, Poland*

We describe an implementation of density functional theory that is formulated fully in configuration space, where all wave functions, densities, and potentials are represented on a grid. Central to the method is a fourth-order factorization of the evolution operator for the Kohn-Sham Hamiltonian. Special attention is paid to nonlocal pseudopotentials of the Kleinman-Bylander type, which are necessary for a quantitatively accurate description of molecules, clusters, and solids. It is shown that the fourth-order factorization improves the computational efficiency of the method by about an order of magnitude compared with second-order schemes. We use the Ono-Hirose filtering method to reduce the resolution of the grid used for representing the wave functions. Some care is needed to maintain the fourth-order convergence using filtered projectors, and the necessary precautions are discussed. We apply the method to an isolated carbon atom as well as the carbon-monoxide molecule, the benzene molecule, and the buckminsterfullerene cluster, obtaining quantitative agreement with previous results. The convergence of the method with respect to time step, grid resolution, and filtering method is discussed in detail.

PACS number(s): 71.15.Dx, 71.15.Ap

I. INTRODUCTION

Electronic structure calculations in the framework of density functional theory (DFT),[1,2] combined with the pseudopotential approximation, have established themselves as a widely used research tool[3] in condensed matter physics, physical chemistry, materials science, and molecular physics, with numerous examples of applications in these different fields. In DFT the problem of solving the Schrödinger equation for a system of interacting electrons moving in the electrostatic field of the nuclei is reformulated in terms of a collection of independent particles moving in an effective field, the so-called Kohn-Sham potential. These independent particles are represented by Kohn-Sham orbitals, which are the solutions of an effective Schrödinger equation where the role of the potential is played by the Kohn-Sham potential. Despite its conceptual simplicity—the complicated many-body problem is effectively mapped onto a one-body problem—DFT calculations still constitute a significant computational challenge, particularly if the system to be studied is large, or if one wishes to combine DFT calculations with molecular dynamics (MD) simulations. This high computational cost significantly hinders the applicability of DFT, in spite of the dramatic increase in computational power experienced over the past few decades. Because of this, there is a continuing search in the scientific community for more efficient methods to solve the Kohn-Sham equations, as well as for better ways to represent the Kohn-Sham orbitals. Several well-established approaches have been put forward. In the context of computational solid state physics, perhaps the most frequently employed approach is the so-called momentum space formalism,[4] in which the Kohn-Sham orbitals are represented in terms of plane waves. This strategy has been implemented in a number of widely used programs, such as VASP,[5] CASTEP,[6] ABINIT,[7] and others. An alternative approach also commonly used is to employ a basis set formed by atom-centered functions resembling atomic orbitals, such as

the *fireball* functions[8] or Gaussians. This is the approach implemented in the SIESTA[9] package and in a number of similar codes.[10–12] Yet a third alternative, which is the focus of the present work, is to employ a real space grid spanning the volume of the system to represent the Kohn-Sham orbitals, the electron density, and the potential.[13–18] Real space approaches are particularly well suited for parallel computer architectures and for achieving the aim of linear scaling of the computational cost with the size of the system.[19] Closely related to real space grids are finite elements approaches,[20] discrete variable representations,[21] Lagrange meshes,[22,23] and wavelets.[24,25]

All of the above different strategies have their relative merits and drawbacks. The momentum space approach is well suited for periodic systems, but its use in the context of finite systems, such as molecules or clusters, is more troublesome. The plane-wave basis set is unbiased, orthonormal, the convergence is controlled by a single parameter (the number of plane waves), and, since it does not depend on the positions of the ions, it does not lead to Pulay terms in the forces. However, the number of plane waves per atom is usually quite large, and iterative diagonalization schemes[26] must be used to obtain the Kohn-Sham orbitals. Atomiclike basis sets, on the other hand, are usually much smaller, and work indistinctly in both periodic and finite boundary conditions. However, unlike plane waves, the convergence is not systematic, and is less obviously controllable. Since the basis functions are centered on the atoms, there are Pulay contributions to the forces. The nonorthogonality of the basis set requires the use of a generalized eigensolver, which sometimes can be ill-conditioned, and the localized nature of the basis functions may lead to basis-set superposition errors. Grid-based methods are in many respects similar to the plane-wave based approach, although one advantage is that the boundaries can indistinctly be periodic or not.

While all-electron calculations are possible,[27] they are computationally demanding, and it is far more common to

employ the pseudopotential approximation.[3,28] This approximation consists of substituting the combined effect of the relatively inert inner electrons (the so-called core electrons) and the atomic nucleus on the chemically active valence electrons by an effective (pseudo) potential. In order to construct accurate and transferable potentials that correctly mimic the properties of different atoms in varying environments, it is necessary to employ a nonlocal (angular momentum-dependent) form. For excellent reviews on pseudopotentials and their use, see Refs. 3 and 28.

In this paper, we describe an implementation of DFT combined with nonlocal pseudopotentials employing a real space grid to represent the Kohn-Sham orbitals, the density, and the potential. The novelty of our work lies in how we obtain the Kohn-Sham orbitals. Our strategy consists of employing the evolution operator in imaginary time, approximated by a fourth-order factorization. As will be demonstrated below, this approach is particularly simple, and extremely efficient.

The structure of this paper is the following. In Sec. II, we describe the theoretical and computational aspects of our method. In Sec. III, we present a series of numerical tests that illustrate the capabilities and robustness of the method. Our conclusions and a description of the future prospects of this method are discussed in Sec. IV.

II. METHODOLOGY

A. Basic strategy

In DFT calculations, the Hamiltonian is dependent on the self-consistent potential, which is unknown at the start. Typically, one begins by assuming a given potential (or equivalently, an electron density), and then solves the Kohn-Sham equations for that fixed potential. The resulting orbitals are then used to obtain a new density and potential, and the process is iterated until self-consistency (i.e., the mutual consistency of Kohn-Sham potential and orbitals) is obtained. Therefore, two different types of operations need to be performed sequentially in any such calculation, namely (a) given a trial fixed potential, and hence a Hamiltonian, obtain the eigenvalue/eigenstate pairs, and (b) update the density and potential based on the previous history of the calculation and the current Kohn-Sham orbitals as well as the structure of the Kohn-Sham energy functional in such a way as to optimize the convergence towards self-consistency.

The second step, i.e., the update of the Kohn-Sham potential or electronic density, has been the subject of numerous studies, and efficient strategies have been developed over the past few decades.[5,26,29–31] Likewise, there are well established techniques for addressing the first step, namely, the calculation of eigenpairs/eigenstates given a fixed potential. When the dimensions of the problem are large, as occurs when plane-wave or grid based methods are employed, it is normally not possible to hold simultaneously in the computer memory the entire matrix representation of the Hamiltonian. Therefore, one must resort to employing some iterative diagonalization technique, such as a conjugate gradients minimization of the total energy.[26] In this work, we advocate a different strategy for addressing this part of the calculation.

Specifically, we use the evolution operator in imaginary time to obtain the Kohn-Sham orbitals. The basic step consists of first applying the diffusion operator

$$\mathcal{T}(\epsilon) = e^{-\epsilon H}, \tag{2.1}$$

where ϵ is the time step and H is the Kohn-Sham Hamiltonian, on some set of trial wave functions, $\{\psi_i^{(k)}\}, i = 1, \ldots, n$, represented on the grid. Since the evolution operator in imaginary time is not unitary (i.e., it does not preserve the normalization nor the orthogonality), it is then necessary to orthonormalize the resulting functions after each propagation step. We will hereafter refer to acting with $e^{-\epsilon H}$ *and* the subsequent orthonormalization as a *time step*. Repeatedly applying

$$\psi_i^{(k+1)} \equiv \mathcal{T}(\epsilon)\psi_i^{(k)} \tag{2.2}$$

and orthonormalizing makes the functions evolve toward, and eventually converge to, the lowest n eigenfunctions of the Hamiltonian H. Once this has been achieved, the resulting orbitals are used to update the density, from which a new potential is calculated, and the process is repeated until self-consistency is attained.

B. Operator factorizations for nonlocal pseudopotentials

Let us now consider how to evaluate the action of the complex-time evolution operator on some trial function ψ. Without previous knowledge of the eigenstates of H, it is not possible to do this exactly, but different kinds of approximations have been proposed in the literature. For local Hamiltonians, one can use Trotter factorizations of the form

$$\begin{aligned} e^{-\epsilon H} &= e^{-\epsilon(T+V)} \\ &\approx e^{-a_i\epsilon A}e^{-b_i\epsilon B}\cdots e^{-a_1\epsilon A}e^{-b_1\epsilon B}e^{-c_0\epsilon A} \\ &\times e^{-b_1\epsilon B}e^{-a_1\epsilon A}\cdots e^{-b_i\epsilon B}e^{-a_i\epsilon A}, \end{aligned} \tag{2.3}$$

where $A=T(B=V)$ or $A=V(B=T)$, T being the kinetic energy operator, V the potential operator, and the numerical coefficients a_i, b_i, c_0 are chosen so as to ensure the desired order of the factorization. For example, two possible second-order factorizations of $e^{-\epsilon H}$ are

$$\mathcal{T}(\epsilon) = e^{-(\epsilon/2)A}e^{-\epsilon B}e^{-(\epsilon/2)A} + O(\epsilon^3) \equiv \mathcal{T}^{(2)}(\epsilon) + O(\epsilon^3). \tag{2.4}$$

This second-order factorization is useful, and indeed it has been employed in applications of time-dependent DFT, or quantum Monte Carlo calculations. Replacing the exact evolution operator by a factorization is equivalent to approximating the Hamiltonian; hence the converged eigenvalue/eigenfunction pairs become functions of the time step ϵ, the exact result being obtained in the limit $\epsilon \to 0+$. The smaller the time step ϵ, the more accurate the approximation for the eigenvalue/eigenfunction pairs is. On the other hand, the imaginary time evolution converges faster when the time step is large. Since we would like to converge to the eigenstates of H reasonably quickly, it is desirable to use approximations of the diffusion operator that are both stable and accurate for large time steps. The most obvious strategy

would be to go to factorizations of order higher than 2.

For factorizations of the form (2.3) to be useful for imaginary time-step propagation, the coefficients a_i, b_i, and c_0 must be positive, otherwise some of the $e^{-a_i \epsilon T}$ and/or $e^{-b_i \epsilon V}$ will be unbound and the method diverges. It has been shown,[32] however, that beyond second order in ϵ *no* factorizations of the form (2.3) exist having purely positive coefficients. A way out has been proven by Suzuki[33] and Chin:[34] it is possible to obtain factorizations with positive coefficients if one introduces an additional operator having the form $[V,[T,V]]$. For example, a possible fourth-order factorization is

$$e^{-\epsilon H} = e^{-(\epsilon/6)V}e^{-(\epsilon/2)T}e^{-(2\epsilon/3)\widetilde{V}}e^{-(\epsilon/2)T}e^{-(\epsilon/6)V} + O(\epsilon^5)$$
$$\equiv \mathcal{T}^{(4)}(\epsilon) + O(\epsilon^5), \qquad (2.5)$$

where

$$\widetilde{V} = V + \frac{\epsilon^2}{48}[V,[T,V]] = V + \frac{\epsilon^2 \hbar^2}{48m}|\nabla V|^2. \qquad (2.6)$$

The power of the method lies in the fact that the correction term $[V,[T,V]]$ is again just a local repulsive potential, therefore dealing with $\widetilde{V}$ is numerically identical to dealing with V.

We have implemented this method in the past for local potentials in two and three dimensions[31] as well as in magnetic fields[35,36] and have demonstrated that the fourth-order factorization (2.5) accelerates the convergence rate of the method by one to two orders of magnitude, depending on the accuracy required.

Unfortunately, quantitative accuracy of electronic structure calculations for realistic systems can normally not be obtained by describing the ion cores by local potentials. The purpose of this paper is therefore to generalize the above evolution operator scheme for DFT calculations combined with nonlocal pseudopotentials. The Hamiltonian then consists of three noncommuting terms,

$$H = T + V_{\text{loc}} + V_{\text{nl}}, \qquad (2.7)$$

where V_{loc} gathers all the local contributions to the Kohn-Sham potential (Hartree, exchange-correlation, and local part of the pseudopotentials), while V_{nl} represents the nonlocal part of the pseudopotentials used to describe the interaction of the valence electrons with the ion cores. Several fourth-order factorization schemes have been derived for the case of three noncommuting operators;[37] a simple method to obtain such factorizations is to use Eqs. (2.4) or (2.5) recursively. There are many ways to do this; we have verified that the factorization utilized here is computationally the most advantageous one because it contains the smallest number of nonlocal operators,

$$e^{-\epsilon H} = e^{-(\epsilon/6)V_{\text{nl}}}e^{-(\epsilon/2)H_{\text{loc}}}e^{-(2\epsilon/3)\widetilde{V}_{\text{nl}}}e^{-(\epsilon/2)H_{\text{loc}}}e^{-(\epsilon/6)V_{\text{nl}}} + O(\epsilon^5), \qquad (2.8)$$

where $H_{\text{loc}} = T + V_{\text{loc}}$, and $\widetilde{V}_{\text{nl}}$ is defined analogously to Eq. (2.6). The factors containing H_{loc} in the exponent are in turn factorized in terms of T and V_{loc} as indicated by Eq. (2.5).

Let us now consider how each term in the factorization of $e^{-\epsilon H}$ affects a trial wave function, when the Hamiltonian is given by Eq. (2.7). The simplest case is that of terms involving the local potential: $\phi = e^{-\epsilon V_{\text{loc}}}\psi$. In a real space representation, a grid spanning the volume of the system is employed, and objects depending on $\mathbf{r}$ are represented on this grid by their numerical values at each grid point $\mathbf{r}_n$, where n is an index labeling the grid points. Thus, the effect of terms such as $e^{-\epsilon V_{\text{loc}}}$ on a trial function ψ is directly evaluated locally at each grid point,

$$\phi(\mathbf{r}_n) = e^{-\epsilon V_{\text{loc}}(\mathbf{r}_n)}\psi(\mathbf{r}_n). \qquad (2.9)$$

Next, consider terms containing the kinetic energy operator in the exponent, $e^{-\epsilon T}\psi$. One way to evaluate these terms is to numerically Fourier-transform the trial function $\psi(\mathbf{r}_n)$ to momentum space, where $e^{-\epsilon T}$ is diagonal. Then we can easily calculate

$$\phi(\mathbf{p}_n) = e^{-\epsilon t(\mathbf{p}_n)}\psi(\mathbf{p}_n), \qquad (2.10)$$

and Fourier transform $\phi(\mathbf{p}_n)$ back to real space. Here, $t(p)$ is the eigenvalue of the kinetic energy operator for wave vector $\mathbf{p}$; one can use the eigenvalues of any finite-difference formula[18,31] or the continuum limit $t(\mathbf{p}) = p^2/2m$. Since very efficient schemes exist to perform numerical Fourier transforms,[38] the operations required to proceed along this way can be carried out rather effectively. We note that there is also a simple alternative for the propagation with the kinetic energy operator, which avoids the need to transform between coordinate and momentum representations, an operation that can only be carried out efficiently on a uniform Cartesian grid. The idea is to directly represent $e^{-\epsilon T}$ in coordinate space. In the limit of an infinitely dense real space grid (infinite basis limit), the real space representation of $e^{-\epsilon T}$ is

$$\lim_{h \to 0} \langle \mathbf{r}|e^{-\epsilon T}|\mathbf{r}' \rangle = \left(\frac{m}{2\pi\epsilon}\right)^{3/2} e^{-m|\mathbf{r} - \mathbf{r}'|^2/2\epsilon}, \qquad (2.11)$$

where h is the grid spacing, $|\mathbf{r}\rangle$ are eigenvectors of the position operator; in this limit these are Dirac δ functions. Of course $\langle \mathbf{r}|e^{-\epsilon T}|\mathbf{r}'\rangle$ is not diagonal, but it is clearly diagonally dominant, the more so the smaller the time step ϵ. Thus, the action of $e^{-\epsilon T}$ on the trial function ψ, which formally is given by

$$\langle \mathbf{r}|\phi \rangle = \int d\mathbf{r}' \langle \mathbf{r}|e^{-\epsilon T}|\mathbf{r}'\rangle\langle \mathbf{r}'|\psi \rangle, \qquad (2.12)$$

does not require to carry out an integral over all space, but only over the neighborhood $\mathbf{r}'$ of $\mathbf{r}$ where $\langle \mathbf{r}|e^{-\epsilon T}|\mathbf{r}'\rangle$ is sufficiently different from zero. In practice, the infinite basis limit is not attainable; rather, we work with a truncated basis, i.e., the real space grid, with a finite grid spacing. In this case, Eq. (2.11) is not strictly valid any longer, but it is still true that $\langle \mathbf{r}|e^{-\epsilon T}|\mathbf{r}'\rangle$ is diagonally dominant. Its actual form can be obtained numerically by performing a discrete Fourier transform of $e^{-\epsilon p^2/2m}$ from momentum to coordinate space, which is done once and for all at the beginning of the calculation. On a Cartesian grid, the factorization

$$e^{-\epsilon p^2/2m} = e^{-\epsilon p_x^2/2m} e^{-\epsilon p_y^2/2m} e^{-\epsilon p_z^2/2m} \qquad (2.13)$$

holds, so we only need to store a one-dimensional array containing $\langle x|e^{-\epsilon p_x^2}|x'\rangle$ as a function of $|x'-x|$. The advantage of the method is that one is not committed to a regular Cartesian grid, which opens the possibility to implement the imaginary time-step method on adaptive grids.

Finally, but most importantly, we must address the treatment of the nonlocal potential. The specifics depend, of course, on the type of nonlocality. In this work, we adopt the Kleinman-Bylander[39] separable form of the ionic pseudopotentials. The nonlocal part of the Hamiltonian has the form

$$V_{nl}(\mathbf{r},\mathbf{r}') = \sum_{i=1}^{N} v_{nl}^{(i)}(\mathbf{r}-\mathbf{R}_i,\mathbf{r}'-\mathbf{R}_i)$$
$$= \sum_{i=1}^{N}\sum_{\ell m} A_\ell^{(i)} \langle \mathbf{r}-\mathbf{R}_i|P_{\ell m}^{(i)}\rangle\langle P_{\ell m}^{(i)}|\mathbf{r}'-\mathbf{R}_i\rangle. \qquad (2.14)$$

In the above equation, $A_\ell^{(i)}$ are numerical constants characterizing the pseudopotential, and the $P_{\ell m}^{(i)}$ are projectors defined as $P_{\ell m}^{(i)} = R_\ell^{(i)}(r)Y_{\ell m}$, where $R_\ell(r)$ is a radial function, and $Y_{\ell m}$ are spherical harmonics; the $\mathbf{R}_i$ are the positions of the ions, and the superscript (i) indicates that the $A_\ell^{(i)}$, $P_{\ell m}^{(i)}$, and $R_\ell^{(i)}$ depend on the chemical species of atom i. Because the range of the nonlocal pseudopotential extends only up to a sphere of radius $r_c^{(i)}$ around each atomic position $\mathbf{R}_i$, and normally the core radii $r_c^{(i)}$ are chosen so that there is no overlap between the different core spheres, the pseudopotentials centered on different atoms commute. We can therefore write

$$e^{-\epsilon V_{nl}} = \prod_{i=1}^{N} e^{-\epsilon v_{nl}^{(i)}}. \qquad (2.15)$$

Expanding $e^{-\epsilon v_{nl}^{(i)}}$ in a power series, it is easy to see that its effect on a trial wave function ψ is given by

$$\langle \mathbf{r}|e^{-\epsilon v_{nl}^{(i)}}|\psi\rangle = \langle \mathbf{r}|\left[1 + \sum_{\ell m}(e^{-\epsilon A_\ell^{(i)}}-1)|P_{\ell m}^{(i)}\rangle\langle P_{\ell m}^{(i)}|\right]|\psi\rangle, \qquad (2.16)$$

i.e., acting with $e^{-\epsilon V_{nl}}$ on a wave function costs the same numerical effort as acting with V_{nl}.

The double commutators require further attention. There are two such terms, namely the one arising in the factorization of $e^{-\epsilon H_{loc}}$ in terms of exponentials of the kinetic energy and local potential operator, i.e., $[V_{loc},[T,V_{loc}]]$, and that arising from the factorization of $e^{-\epsilon H}$ in terms of exponentials of the local Hamiltonian, H_{loc}, and the nonlocal part of the potential, V_{nl}. The double commutator $[V_{loc},[T,V_{loc}]]$ $= \frac{\hbar^2}{m}|\nabla V_{loc}|^2$ has been dealt with in previous work,[31] and with minimal attention to numerical consistency, it does not pose any problems. It is straightforward to verify that the double commutator $[v_{nl}^{(i)},[H_{loc},v_{nl}^{(i)}]]$ has the form

$$[v_{nl}^{(i)},[H_{loc},v_{nl}^{(i)}]] = -\sum_{\ell m} B_\ell^{(i)}[|Q_{\ell m}\rangle\langle P_{\ell m}| + |P_{\ell m}\rangle\langle Q_{\ell m}|], \qquad (2.17)$$

where $|Q_{\ell m}\rangle$ is defined as

$$|Q_{\ell m}\rangle = \frac{1}{N_\ell}[H_{loc}|P_{\ell m}\rangle - \langle P_{\ell m}|H_{loc}|P_{\ell m}\rangle|P_{\ell m}\rangle], \qquad (2.18)$$

where the N_ℓ are the normalization constants, and $B_\ell^{(i)}$ $=(A_\ell^{(i)})^2 N_\ell$. Note in particular that $\langle Q_{\ell m}|P_{\ell m}\rangle=0$. We must pay attention to the fact that the operation of the local Hamiltonian H_{loc} on the projectors $|P_{\ell m}^{(i)}\rangle$ mixes all angular momenta; this additional angular dependence stems from the angular dependence of the local potential $V_{loc}(\mathbf{r})$ in the vicinity of the ion where $v_{nl}^{(i)}$ is defined. In principle, one should expand the product $V_{loc}(\mathbf{r}-\mathbf{R}_i)|P_{\ell m}\rangle$ in a new set of spherical harmonics and radial functions. In the practical applications we discuss below, we have approximated the $V_{loc}(\mathbf{r}-\mathbf{R}_i)$ by their spherical average, which greatly simplifies the implementation. This approximation implies that only the diagonal matrix elements $\langle P_{\ell m}|H_{loc}|P_{\ell m}\rangle$ appear in the commutator. Note that this approximation can only change the convergence rate of the eigenvalue/eigenfunction pairs as a function of the time step ϵ, but does not affect the final result. We illustrate below (see Sec. III) that this approximation indeed affects the convergence rate, but only for very small time steps, while allowing a significant simplification of the methodology.

We see that $[v_{nl}^{(i)},[H_{loc},v_{nl}^{(i)}]]$ retains a form very similar to that of the nonlocal pseudopotential itself. Since H_{loc} will change at each self-consistent cycle, it is necessary to update $[v_{nl}^{(i)},[H_{loc},v_{nl}^{(i)}]]$ every time H_{loc} changes. This simply requires the recalculation of the $|Q_{\ell m}\rangle$ for each atom, which is not an expensive calculation.

The full nonlocal operator $\widetilde{V}_{nl}$ defined in Eq. (2.8) is then

$$\widetilde{V}_{nl}(\mathbf{r},\mathbf{r}') = \sum_i \widetilde{v}_{nl}^{(i)}(\mathbf{r}-\mathbf{R}_i,\mathbf{r}'-\mathbf{R}_i),$$

$$\widetilde{v}_{nl}^{(i)}(\mathbf{r},\mathbf{r}') = \sum_{\ell m}\left[A_\ell^{(i)}|P_{\ell m}\rangle\langle P_{\ell m}| - \frac{\epsilon^2}{48}B_\ell^{(i)}[|Q_{\ell m}\rangle\langle P_{\ell m}| + |P_{\ell m}\rangle \right.$$
$$\left. \times\langle Q_{\ell m}|]\right]. \qquad (2.19)$$

The action of $e^{-\epsilon\widetilde{v}_{nl}^{(i)}}$ on $|\psi\rangle$ can be evaluated by transforming Eq. (2.19) to a representation where it is diagonal in the space spanned by the two projectors $|P_{\ell m}\rangle$ and $|Q_{\ell m}\rangle$, i.e., by diagonalizing the matrix

$$\begin{pmatrix} A_\ell^{(i)} & -\dfrac{\epsilon^2}{48}B_\ell^{(i)} \\[3mm] -\dfrac{\epsilon^2}{48}B_\ell^{(i)} & 0 \end{pmatrix}. \qquad (2.20)$$

C. Numerical implementation

There are a number of issues worth paying attention to when implementing the fourth-order algorithm Eq. (2.8) numerically on a discrete mesh. Formally speaking, the method is rigorously fourth order in the time step, provided there are no discretization errors. The presence of the latter in any numerical implementation may spoil the fourth-order convergence unless care is taken to treat the different operators appearing in Eq. (2.8) in a consistent manner.[31]

A key consideration to retain the fourth-order convergence while employing a discrete mesh is to treat the kinetic energy operator in the double commutator in exactly the same manner as the kinetic energy operator in the exponentials $e^{-1/2\epsilon T}$. The problem is that the double commutator is local only in the limit of an infinitely fine mesh; in any finite discrete representation of the kinetic energy operator the double-commutator is nonlocal. A possible way to proceed is to expand the exponential

$$\exp\left\{-\frac{2}{3}\frac{\epsilon^3}{48}[V,[T,V]]\right\} = 1 - \frac{\epsilon^3}{72}[2VTV - V^2T - TV^2]$$

(2.21)

and act with this operator on the wave functions. Then, the algorithm is rigorously fourth order, independent of the discretization, but the approach is computationally costly and reduces the convergence rate significantly compared to keeping the double commutator in the exponential. We have shown in Ref. 31 that the optimal way to proceed is to use the form

$$\frac{\hbar^2}{m}|\nabla V|^2 = -(TV^2) + 2V(TV)$$

(2.22)

for the double commutator, the round brackets indicating that T acts on V or V^2 alone and not on a wave function behind it, and keep this form in the exponential. One can still add a correction term to establish rigorous fourth-order convergence, but we found that this term is very small and rather cumbersome to implement for discretizations of the kinetic energy operator other than a three-point formula.

The same consideration applies for the calculation of the double commutator $[H_{\mathrm{loc}}[V_{\mathrm{nl}},H_{\mathrm{loc}}]]$: Since the kinetic energy in $e^{-(\epsilon/2)T}$ is calculated on a Cartesian grid, the *same* form should be used for the operation of the kinetic energy in that double commutator.

D. Filtering of the nonlocal projector functions

The calculation of the action (2.16) of the projection operators $|P_{\ell m}^{(i)}\rangle\langle P_{\ell m}^{(i)}|$ on the wave functions requires integration of a rapidly varying function—the projector function $|P_{\ell m}^{(i)}\rangle$—with the more slowly varying wave function $|\psi\rangle$. Several schemes have been proposed for treating these situations, e.g., using more sophisticated discretizations[40] or filtering of the pseudopotentials.[41] In particular, we have chosen to implement the scheme of Ono and Hirose,[42] which allows to do this accurately even for rather coarse grid spacings. The basic idea is that the projectors are rapidly varying

functions and should be represented on a fine mesh, whereas the wave functions are relatively slowly varying and can be discretized on a coarse grid. To calculate integrals of the form $\langle P_{\ell m}^{(i)}|\psi\rangle$ on a fine mesh appropriate for the projectors, it would be sufficient to interpolate the wave function on the fine grid. The Ono-Hirose[42] method permits the calculation of integrals as if they were carried out by summing on the fine mesh, but by summing only over the coarse grid. In the following, points on the coarse grid are denoted by lower case indices (so ψ_i is the value of ψ on the ith point of the coarse grid), whereas fine grid points are denoted by upper case indices (i.e., P_I is the value of the projector on the Ith point of the fine grid).

Given some interpolation scheme to calculate the value of the wave function on a fine grid point (ψ_I), one can approximate the scalar product by a sum over the fine mesh,

$$\langle P|\psi\rangle = \int d\mathbf{r}\,P(\mathbf{r})\psi(\mathbf{r}) \approx h_f\sum_I P_I\psi_I,$$

(2.23)

where h_f is the volume of one fine-grid cell. For interpolation schemes where the interpolated function values can be written in the form

$$\psi_I = \sum_j t_{Ij}\psi_j$$

(2.24)

with constant coefficients t_{Ij}, one can transform the summation in Eq. (2.23) into a sum over the coarse mesh only,

$$\langle P|\psi\rangle = h_f\sum_j \underbrace{\sum_I t_{Ij}P_I}_{=:P'_j}\psi_j = h_f\sum_j P'_j\psi_j,$$

(2.25)

where the filtered projectors P' need to be calculated only once at the beginning of the program. Interpolation methods that fulfill Eq. (2.24) are, e.g., Fourier interpolation and Lagrange interpolating polynomials of any order. Spline interpolation, on the other hand, cannot be used, as the coefficients t_{Ij} would depend on the coarse-grid function values ψ_j, and thus would have to be recalculated every time ψ changes.

However, we not only need to calculate scalar products, but also the action of the exponential of the nonlocal potential operator onto a wave function $|\psi\rangle$. Summing up the exponential series (2.16) is only possible if the nonlocal operators

$$\mathcal{P}_{\ell m}^{(i)} = |P_{\ell m}^{(i)}\rangle\langle P_{\ell m}^{(i)}|$$

(2.26)

commute for different values of ℓm and i, and if they are projection operators, so that $(\mathcal{P}_{\ell m}^{(i)})^n = \mathcal{P}_{\ell m}^{(i)}$. This poses the additional constraint on the projector functions that they must be orthonormalized,

$$\langle P_{\ell m}^{(i)}|P_{rs}^{(j)}\rangle = \delta_{\ell r}\delta_{ms}\delta_{ij}.$$

(2.27)

Filtering, on the other hand, changes the norm of the projectors, because it modifies the projector functions such that scalar products with slowly varying functions give the result one would get on a finer grid. The norm of the filtered projector, being the scalar product with itself, is then not a

TABLE I. Operations needed to do one propagation step using the second and fourth order factorizations for non-local pseudopotentials.

Operation type		Comp. time	Operation count	
			Second order	Fourth order
FFT		T_{FFT}	2	8
vector-vector multiplication	$e^{-\epsilon V},\ e^{-\epsilon T},\ e^{-\epsilon \tilde{V}}$	T_{mul}	3	10
nonlocal propagation	$e^{-\epsilon V_{\text{nl}}}$	T_{nl}	2	2
nonlocal double commutator	$e^{-\epsilon \tilde{V}_{\text{nl}}}$	T_{nldc}	0	1

meaningful quantity anymore, because the projectors are not slowly varying functions. This does not pose any problem, however, since one can always renormalize the projectors by changing the coefficients $A_\ell^{(i)}$ in Eq. (2.14) accordingly.

Due to the discretization of the functions on a finite grid, there are also more subtle issues with regard to the orthonormalization condition (2.27). When represented on a finite mesh, the projector functions are of course not strictly orthogonal anymore, they may even have slightly different norms for different m channels. Filtering cannot cure this problem; it does not even guarantee that the filtered functions are of the form $R_\ell(r)Y_{\ell m}$ at all. Although these are very small effects for reasonable grid spacings, it turns out that they can deteriorate fourth order convergence at small time steps. To maintain fourth-order behavior even down to small time steps, we therefore orthogonalize the projectors after filtering, then normalize all projectors using $m-$ dependent coefficients $A_{\ell m}^{(i)}$ in Eq. (2.14). We use Fourier interpolation for filtering the projector functions from a fine grid having a grid spacing four times finer than the grid used to represent the orbitals. We have checked that this is sufficient, and finer grids do not result in any improvement, at least in the cases considered below. At the beginning of the program, the projector functions are calculated on a fine grid and transformed to Fourier space. Then they are cut off so that the transformation back to real space yields them on the coarse grid we use to describe the wave functions. Afterwards, the filtered projectors are orthonormalized using a Gram-Schmidt procedure and the coefficients $A_{\ell m}^{(i)}$ are changed accordingly.

To propagate the double commutator term (2.17), we also need to calculate the functions $|Q_{\ell m}\rangle$ defined in Eq. (2.18). Notice that $H_{\text{loc}} = T + V_{\text{loc}}$, so that the kinetic energy part $T|P_{\ell m}\rangle$ can also be computed at the beginning of the program. This is done by differentiating the filtered and normalized projectors in Fourier space, in exactly the same way as the action of the kinetic energy operator is calculated. The other terms contain the Kohn-Sham potential and have to be recomputed before each self-consistent iteration.

III. NUMERICAL TESTS AND APPLICATIONS

The purpose of this section is to demonstrate the functionality of the fourth-order factorization for real-space electronic structure calculations involving nonlocal pseudopotentials. We have chosen four different cases, namely, an isolated carbon atom, the diatomic molecule CO, the medium-size C_6H_6 benzene molecule, and the relatively large C_{60} cluster. These different cases allow us to probe different aspects of the method, as discussed below. We focus on the following issues: the efficiency of the eigensolver (i.e., the ability of the imaginary time evolution method to rapidly locate the eigenstates and eigenenergies for a fixed Hamiltonian) and the influence of the numerical grid on the eigenvalues and total energy.

A. Convergence of the eigenvalue solver

Here we illustrate the efficiency and robustness of the fourth-order factorization, comparing it with the second-order method. In our implementation of the second-order method, we have factorized the diffusion operator as

$$\mathcal{T}^{(2A)}(\epsilon) = e^{-(\epsilon/2)V_{\text{nl}}}e^{-\epsilon H_{\text{loc}}}e^{-(\epsilon/2)V_{\text{nl}}} = \mathcal{T}(\epsilon) + O(\epsilon^3) \quad (3.1)$$

and then factorize $e^{-\epsilon H_{\text{loc}}}$ according to Eq. (2.4) with $A = V$ and $B = T$. One time step in the second-order factorization requires to compute twice the action of $e^{-(\epsilon/2)V_{\text{nl}}}$ and one propagation with a second-order approximation of $e^{-\epsilon H_{\text{loc}}}$, which amounts to twice acting with $e^{-(\epsilon/2)V_{\text{loc}}}$ [cf. Eq. (2.4)] and once acting with $e^{-\epsilon T}$. Alternatively one could use

$$\mathcal{T}^{(2B)}(\epsilon) = e^{-(\epsilon/2)H_{\text{loc}}}e^{-\epsilon V_{\text{nl}}}e^{-(\epsilon/2)H_{\text{loc}}} = \mathcal{T}(\epsilon) + O(\epsilon^3). \quad (3.2)$$

which would use only one propagation with the nonlocal potential but twice propagating with the local Hamiltonian. Compared to this, the fourth-order factorization requires twice the action of $e^{-(\epsilon/6)V_{\text{nl}}}$ and once the action of $e^{-2\epsilon/3 \tilde{V}_{\text{nl}}}$ which is about twice as expensive as the action of $e^{-\epsilon/6 V_{\text{nl}}}$; recall that the double commutator generates two projectors. It also requires twice the fourth-order propagation with $e^{-(\epsilon/2)H_{\text{loc}}}$, each of which is about twice as expensive as the second-order propagation. The operations necessary to do one propagation step using the second- and fourth-order factorization schemes are summarized in Table I. The computation times for the vector-vector multiplications (such as $e^{-\epsilon V}\psi$) and the Fourier transforms roughly scale the same way with system size, and we have found that $T_{\text{FFT}} \approx 4T_{\text{mul}}$, for more precise timings see Ref. 31. The nonlocal propagation steps also scale the same way; we found $T_{\text{nldc}} \approx 1.5 T_{\text{nl}}$. Thus the ratio of computation times needed for one propagation step in the second- and fourth-order schemes is

$$\frac{T_{4\text{th}}}{T_{2\text{nd}}} \approx \frac{50T_{\text{mul}} + 3.5T_{\text{nl}}}{13T_{\text{mul}} + 2T_{\text{nl}}}, \tag{3.3}$$

which lies between $T_{4\text{th}} = 1.75T_{2\text{nd}}$ and $T_{4\text{th}} = 3.8T_{2\text{nd}}$.

The computational cost of acting with the nonlocal potential relative to acting with the local potential depends on the number of angular momentum components retained and on the fraction of the simulation box that is covered by nonlocal projectors. For the systems studied in this publication, the factor $T_{\text{nl}}/T_{\text{mul}}$ turned out to be between 3 and 6. Hence we conclude that the fourth-order form is about two to three times as expensive as the second-order form. We shall see shortly that this is easily compensated by the faster convergence of the fourth order method.

The calculations of the eigenvalues are then carried out as follows: We start with a set of wave functions $\{\psi_i^{(0)}\}$ and iterate the αth order algorithm

$$\psi_i^{(k+1)}(\epsilon) = \mathcal{T}^{(\alpha)}(\epsilon)\psi_i^{(k)}(\epsilon) \tag{3.4}$$

followed by orthonormalization for fixed value of ϵ to convergence; the iterations were normally terminated if the change in the eigenvalues $e_i(\epsilon)$ was less than 10^{-8}. This provides a set of approximate eigenvalue/eigenfunction pairs $\{e_i(\epsilon), \psi_i(\epsilon)\}$ for each version of the algorithm and each value of the time step ϵ. The "exact" eigenvalue is then obtained by extrapolating to $\epsilon \to 0$.

As noted above, we have performed calculations for four different systems, namely, the isolated carbon atom, the carbon-monoxide molecule, the benzene molecule, and the buckminsterfullerene cluster. Tests of the efficiency of the second- and fourth-order eigensolvers have been performed as follows: We have first self-consistently solved the Kohn-Sham equations, taking the Perdew-Wang density functional[43] and employing Troullier-Martins[44] pseudopotentials generated by the program FHI98PP.[45] We used projector functions for the $\ell = 0, 1, 2$ angular momentum channels, where the $\ell = 2$ projector was taken as the local potential. The core radii were $1.37a_0$ for the oxygen ion, $1.46a_0$ for carbon, and $1.28a_0$ for hydrogen. We have checked that the projectors generated in this manner do not lead to "ghost states."

The orthogonalization and the density update were performed as described in Ref. 31. We calculate the overlap matrix of propagated vectors,

$$M_{ij}^{(k)}(\epsilon) = \langle \psi_i^{(k)}(\epsilon)\mathcal{T}^{(\alpha)}(\epsilon)|\mathcal{T}^{(\alpha)}(\epsilon)\psi_j^{(k)}(\epsilon)\rangle, \tag{3.5}$$

and solve the eigenvalue problem

$$\sum_j M_{ij}^{(k)}(\epsilon)c_j^{(n)} = m_n^{(k)}(\epsilon)c_i^{(n)}. \tag{3.6}$$

The linear combinations

$$|\psi_j^{(k+1)}(\epsilon)\rangle \equiv \frac{1}{\sqrt{m_j^{(k)}}}\sum_i c_i^{(j)}\mathcal{T}^{(\alpha)}(\epsilon)|\psi_i^{(k)}(\epsilon)\rangle \tag{3.7}$$

form the new orthonormal set of wave functions, and the eigenvalues $m_j^{(k)}(\epsilon)$ are related to the *overlap energies* $e_j(\epsilon)$ through

$$m_j^{(k)}(\epsilon) = \exp[-2\epsilon e_j(\epsilon)]; \tag{3.8}$$

they converge toward the exact eigenvalues as $\epsilon \to 0$.

We have then taken the electron density and the corresponding Kohn-Sham potential as a fixed local field and have solved again the eigenvalue problem, using as initial guess for the evolution the wave functions of a particle in a box, and repeating the process for a sequence of time steps ϵ.

Although in principle the fourth-order power-law convergence is guaranteed by the fourth-order factorization, in practice this may be prevented by the numerical approximations incurred in the actual calculation (see Sec. II C), such as the use of a real space grid to represent orbitals, density, and potential. We have checked the sensitivity of the factorization formula against discretization errors in previous work[31] and determined an "optimal" way to deal with the double commutator that minimizes the discretization errors. A second potentially detrimental feature is that, in calculating the double commutator (2.17), we have assumed that H_{loc} is spherically symmetric inside the core radius, where $v_{\text{nl}}^{(i)}$ is defined. Evidently, keeping only the spherically symmetric part of $V_{\text{loc}}(\mathbf{r} - \mathbf{R}_i)$ in the calculation of the double commutator $[V_{\text{nl}}, [H_{\text{loc}}, V_{\text{nl}}]]$ is an approximation that greatly simplifies the calculation, but which destroys rigorous fourth-order convergence. Any deviation from fourth-order behavior is then, among other things, an indication as to whether this approximation is a valid one. It should be stressed, however, that the approximation of spherical symmetry of H_{loc} affects only the evaluation of the double commutator (2.17), and possibly the rate of convergence, but it does not affect the final result. This is because the double commutators appearing in the factorization of the evolution operator have the effect of guaranteeing the fourth-order convergence with respect to the time step, but they do not have any effect on the converged eigenpairs that result at the end of the evolution.

The two smallest systems that we have considered here are extreme cases for the spherical approximation of the commutators Eq. (2.17): An isolated carbon atom has, of course, a spherically symmetric local Hamiltonian, and therefore any deviation from rigorous fourth-order convergence is a measure of the discretization errors alone. On the other hand, one would expect a maximum amount of anisotropy (i.e., a maximum deviation from spherical symmetry) in a diatomic molecule such as CO. Hence a comparison between these two examples will provide estimates for both the discretization error and the error induced by the assumption of isotropy inside the core regions. Figures 1–4 show different aspects of the convergence of these examples. All figures show the time-step error $\Delta e_1(\epsilon)/|e_1(0)| \equiv (e_1(\epsilon) - e_1(0))/|e_1(0)|$ as a function of the imaginary time step ϵ, where e_1 is the lowest eigenvalue. The two representations provide different pieces of information: The linear scale shows in a direct visual way the convergence features and the range of validity of the power law, while the double-logarithmic scale gives immediate information on the order of the convergence, the range of time steps for which truncation errors spoil the fourth-order convergence, and the relative computational effort of the calculation of the eigenvalues to a given accuracy.
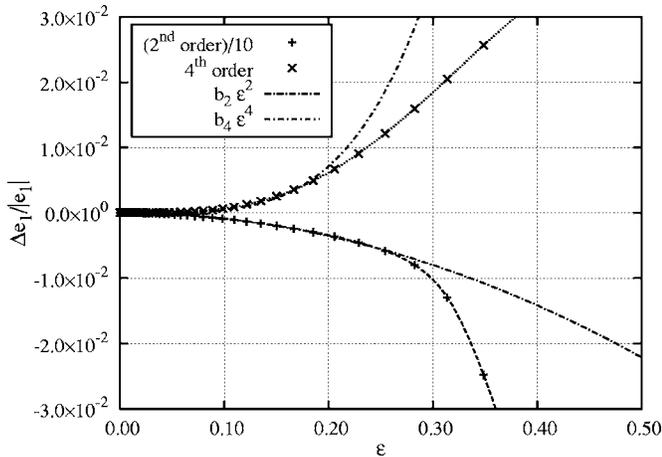
FIG. 1. Relative error $\Delta e_1(\epsilon)/|e_1|$ of the lowest eigenvalue of electrons in a single C atom on a linear scale for the second- and the fourth-order algorithm. Also shown are the fits to the eigenvalues $b_2\epsilon^2$ and $b_4\epsilon^4$. A cubic grid of 48^3 mesh points has been used for the calculation, with a resolution of $h=0.3a_0$.

The calculation of the eigenvalue/eigenfunction pairs at each time step ϵ requires iterating algorithms (2.4) and (2.5) until convergence to a desired accuracy has been reached. It is useful to note that the ultimate accuracy of eigenvalues $e_i(\epsilon)$ is approximately given by the total evolved time $\tau = N\epsilon$, where N is the number of iterations. The necessary evolution time τ to reach a given accuracy is characteristic of the system, but depends only weakly on the particular algorithm used and on the value of ϵ. This estimate of a constant evolution time assumes, however, that one starts with the same wave function. But normally one will calculate the wave functions corresponding to smaller time steps from those of larger time steps. Since the fourth-order method provides more accurate wave functions for larger time steps, one starts the iterations with better initial functions, which improves the convergence rate for large time steps.
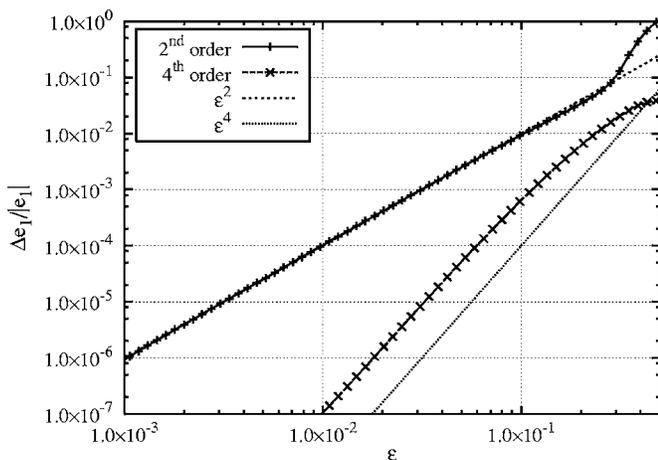
FIG. 2. Relative error of the lowest eigenvalue of electrons in a single C atom on a double-logarithmic scale for the second- and the fourth-order algorithm. Also shown are the functions ϵ^2 and ϵ^4 to verify the power-law convergence. A cubic grid of 48^3 mesh points has been used for the calculation, with a resolution of $h=0.3a_0$.
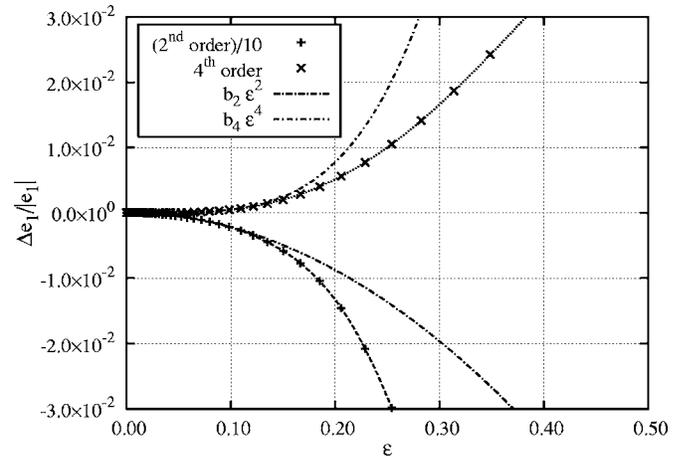
FIG. 3. Same as Fig. 1 for a CO molecule.

We can determine from the double-logarithmic figures an estimate of the time step needed for a desired accuracy for each algorithm, and by the ratio of these time steps determine the speed advantage of the fourth-order algorithm. From Fig. 2, a 10^{-5} accuracy of the eigenvalues is obtained with the fourth-order algorithm at a time step $\epsilon \approx 3 \times 10^{-2}$ a.u., whereas the second-order algorithm needs a time step of $\epsilon \approx 3 \times 10^{-3}$ a.u. This translates into a factor of 10 in total integration time or, for the system at hand, a factor of at least 3 in speed advantage. Requiring an accuracy of 10^{-6} increases the saving to almost an order of magnitude.

The second piece of information to be drawn from Fig. 2 is that the algorithm is perfectly fourth-order down to accuracies of better than 10^{-7}. There are small deviations from that power law for smaller time steps due to discretization errors, see Sec. II C; we have shown in Ref. 31 how to overcome this problem. Our calculations here are consistent with those for local pseudopotentials, and the additional effort does not appear to be warranted, particularly in view of the fact that other approximations have a more visible detrimental effect on the fourth-order convergence.

As pointed out above, the case of CO is perhaps the most sensitive test on the assumption of spherical symmetry of the local Hamiltonian in the evaluation of the nonlocal double
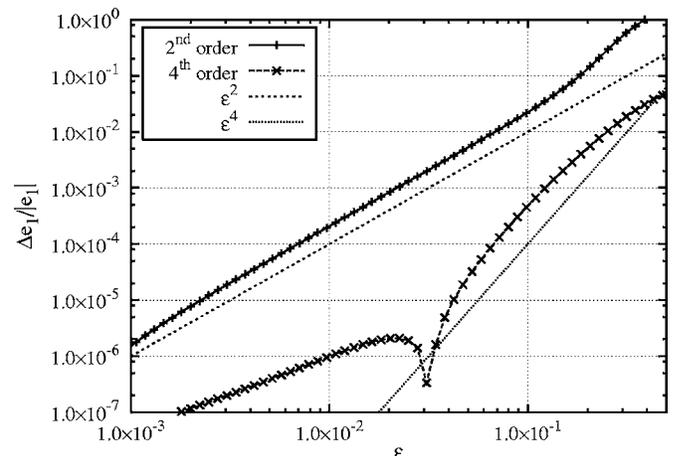
FIG. 4. Same as Fig. 2 for a CO molecule.
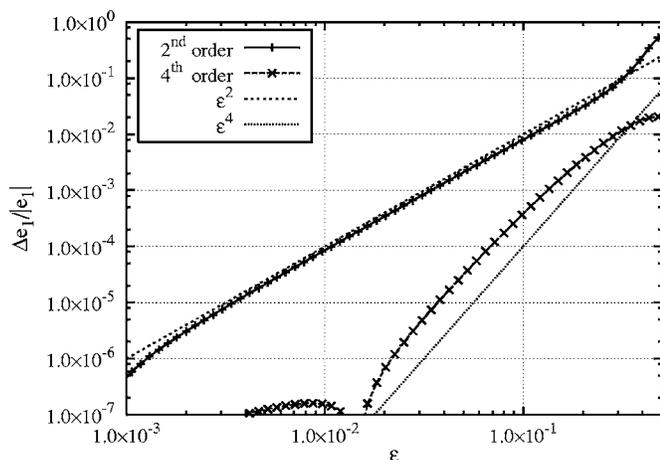
FIG. 5. Same as Fig. 2 for a C_6H_6 molecule.

commutator. Figure 4 shows the convergence of the operator factorization for that case. Clearly, the convergence is close to fourth-order for time steps $\epsilon > 4 \times 10^{-2}$ but turns second order for smaller time steps. As pointed out above, we attribute this to the deviation of H_{loc} from spherical symmetry and we expect that the power-law behavior can be improved by including higher components. From a more pragmatic point of view, we see that we obtain again 10^{-5} accuracy in second order at a time step of $\epsilon \approx 2 \times 10^{-3}$, whereas the same accuracy is reached in the fourth-order algorithm at $\epsilon \approx 4 \times 10^{-2}$, indicating a speed advantage of an order of magnitude.

Figures 5 and 6 give further support for the conclusions drawn from the two first cases: The relative performance of the second- and the fourth-order algorithms is roughly the same. The deviation from a ϵ^4 power law starts at different step sizes for different systems, which we can again take as an indication for the deviation of the local Hamiltonian from spherical symmetry.

B. Iteration strategy

Figures 1–4 show the accuracy of the results obtained with the second- and fourth-order algorithms as a function of
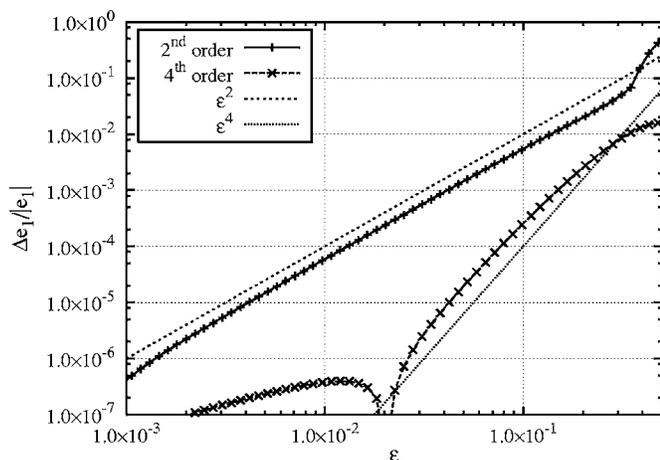
FIG. 6. Same as Fig. 2 for a C_{60} molecule. A cubic grid of 64^3 mesh points has been used for the calculation, with a resolution of $h = 0.4a_0$.

the time step ϵ. An estimate of the required computer time can be obtained from these for the simplified situation that one starts, for each time step, from the *same* set of initial wave functions. Such a strategy is clearly not optimal: In reality, one will start with some set of wave functions $\{\psi_i^{(0)}\}$ (in our case, simple "particle in a box" states) and a large time step, carry out a number of iterations, then reduce the time step and iterate the wave functions obtained at the larger time step again. The process is repeated until the desired accuracy has been reached.

Hence, there are a number of adjustable and model-dependent settings, which are as follows: (i) The initial time step, (ii) a condition for terminating the diffusion iterations for a given time step, (iii) a prescription for reducing the time step, and (iv) a condition for terminating the whole procedure.

A very good estimate for the absolute convergence of the iterations is provided by comparing the "overlap" energies $e_j^{(k)}(\epsilon)$ [Eq. (3.8)] with the "variational" energies

$$h_j^{(k)}(\epsilon) = \langle \psi_i^{(k)}(\epsilon) | H | \psi_j^{(k)}(\epsilon) \rangle. \tag{3.9}$$

These estimates for the eigenvalues converge as $O(\epsilon^{2\alpha})$ if the error in the wave function is $O(\epsilon^{\alpha})$. Looking at the overlap energies serves, therefore, two purposes: For one, it provides a very accurate prediction of the exact eigenvalue as $\epsilon \to 0$, and second, it serves as a verification that not only the eigenvalues, but also the wave functions, have converged to order ϵ^{α}.

For starting the iterations from "minimal knowledge," it is preferable to start with a relatively large time step; for the calculations to be discussed here, an initial time step $\epsilon = 1$ Ry^{-1} has turned out to be adequate. We have then propagated the wave functions for fixed ϵ until the change in the variational energies between subsequent iterations was less than 0.1 times the change in the variational energies from the first to the second iteration. An alternative method is the one proposed in Ref. 31, where we compared $e_j(\epsilon) - h_j(\epsilon)$ directly. When the desired accuracy was reached, the time step was reduced by a factor r and the procedure was repeated. In the cases studied here, a constant reduction factor $r = 0.5$ was taken, but the results up to $r = 0.8$ are similar.

Figures 7 and 8 show the normalized differences $|e_j^{(k)}(\epsilon) - h_j^{(k)}(\epsilon)| / |h_j^{(k)}(\epsilon)|$ as a function of iteration for the two cases CO and C_6H_6 for both the second- and the fourth-order algorithm. The discontinuities in the curves show the places where the time step has been changed. It is seen in all cases that the change in the eigenvalues is largest during the first two iterations, and then levels off. This does not, however, mean that one can naively terminate the iterations much earlier; the reason for this is that the wave functions need a few more iterations for convergence; when iterations for large time steps are terminated too early, the iterations for smaller time steps will take much longer to converge.

In all cases, the second-order method needs, on average, 20–50 times more iterations for convergence to a relative accuracy between 10^{-5} and 10^{-6}. In a realistic DFT calculation, one will, of course, include a density update typically
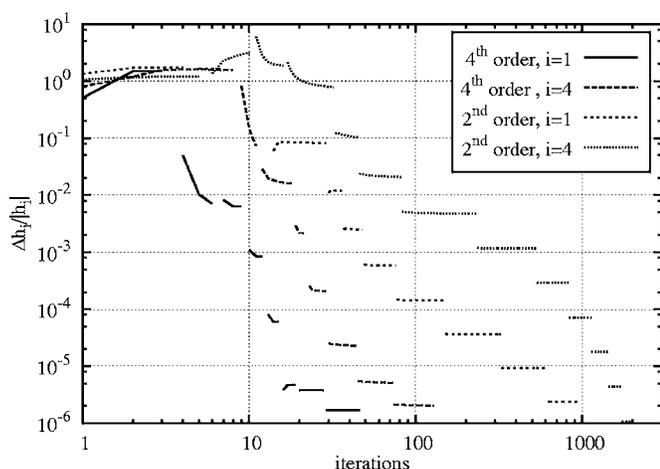
FIG. 7. Relative error $\Delta h_j(\epsilon)/|h_j|$ of the lowest and the fourth eigenvalue of electrons in a CO molecule as a number of iteration for both the fourth- and the second-order algorithm. Discontinuities of the curves occur whenever the time step is changed.

whenever the time step is reduced. That means that one needs initially five to ten iterations of the diffusion algorithm, close to the converged density, and when the minimum time step has been reached, one or two iterations are often sufficient.

C. Convergence with respect to grid spacing

The purpose of the filtering algorithm described in Sec. II D is, of course, to improve the accuracy of integrals with rapidly varying functions such that the calculation can be done on a relatively coarse grid. There are several indicators for both convergence and the efficiency of the filtering method: Working on a rectangular grid breaks, of course, both translational and rotational invariance; the energy of the system will depend slightly on where the individual ions are located relative to the grid points. We can therefore estimate the discretization errors by moving or rotating the system under consideration relative to the grid.
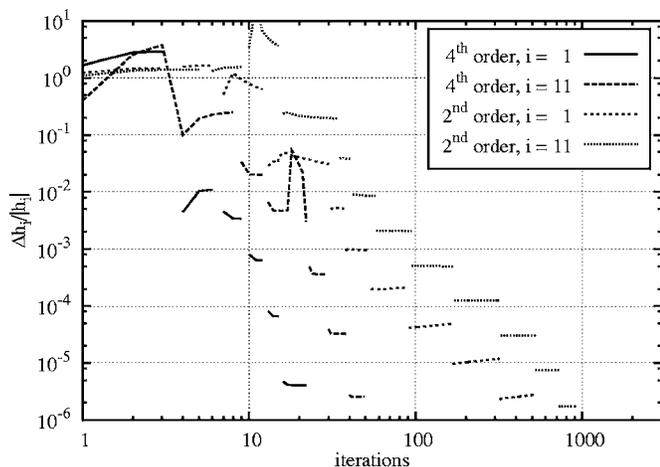
FIG. 8. Same as Fig. 7 for the lowest and the 10th eigenvalue in a C_6H_6 molecule.
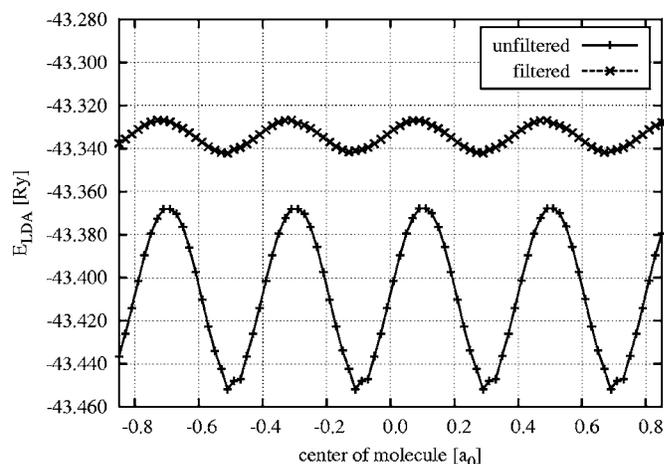
FIG. 9. The figure shows sensitivity of the LDA ground-state energy of a CO molecule to the location of the two atoms relative to the fixed grid. The calculation shown here is for the equilibrium distance of $2.2a_0$, a cubic grid of 48^3 points, and a resolution of $h=0.4a_0$.

Figures 9 and 10 show examples for a CO molecule. Figure 9 shows the resulting energy for the discretization on a mesh of 48^3 points with a mesh resolution $h=0.4a_0$, and Fig. 10 for the discretization on a mesh of 64^3 points with a mesh resolution $h=0.3a_0$. On the fine mesh, the estimated discretization error is about 5×10^{-3} Ry without filtering, which is reduced by filtering to about 7×10^{-4} Ry. The picture is very similar on the coarse mesh; the error is reduced from 8×10^{-2} to 1.5×10^{-2}, in other words, filtering improves the accuracy in this range of discretization by about a factor of 5–7. Notice that the influence of the grid spacing can be clearly discerned from that of any time-step related error, which, as seen from Figs. 2 and 4–6, is easily made several orders of magnitude smaller than the amplitude of grid-induced energy oscillations.

A very similar convergence estimate can be obtained by rotating the molecule. The fluctuation range of the LDA ground-state energy is identical to that obtained from translation and shown in Figs. 9 and 10, therefore we refrain from showing these results.
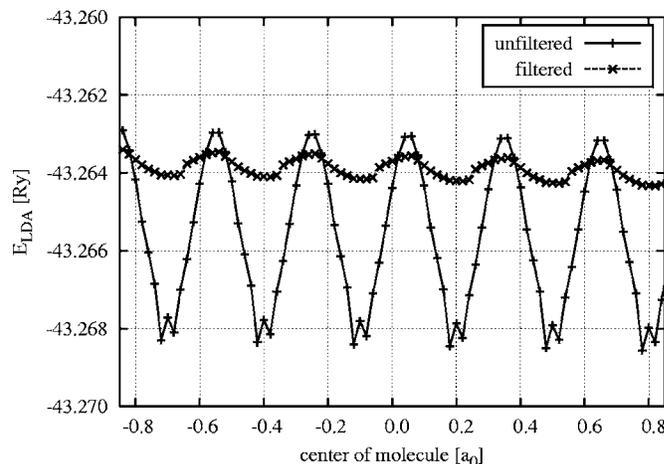
FIG. 10. Same as Fig. 9 for a cubic grid of 64^3 points and a resolution of $h=0.3a_0$.
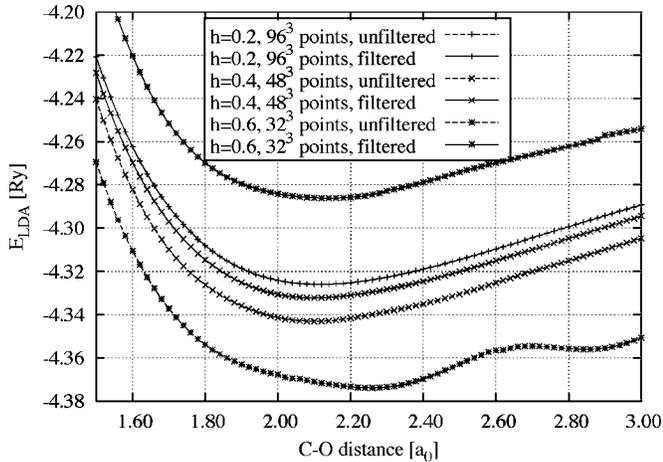
FIG. 11. The figure shows the LDA energy of a CO molecule as a function of the distance between the carbon and the oxygen atom for different discretizations. Solid lines reflect results using filtering, whereas dashed lines show unfiltered results. The grid sizes are depicted by markers (+ signs for 96^3 point mesh, crosses for a 48^3 point mesh, and stars for a 32^3 point mesh).

More important than an estimate of the absolute accuracy of the calculation is an assessment of the accuracy of the prediction of physical observables, which are often energy differences. An obvious quantity to look at is the equilibrium distance and the oscillation frequencies. Figure 11 shows these results for the CO molecule for different discretizations, with and without filtering.

Evidently the 96^3 point mesh is sufficiently accurate such that there is no advantage anymore in filtering; the results are practically identical and we can take them as reference. The only discretization that has an unacceptable accuracy is the unfiltered calculation on a 32^3 point mesh. Table II collects the equilibrium distances d_0 and harmonic vibrational frequencies ν for the CO molecule for different discretizations. The equilibrium distances for all acceptable calculations agree within better than 0.5%. The uncertainty of the vibrational frequency is somewhat larger, the $h=0.6a_0$ result is 4% below the best value, whereas the filtered $h=0.4a_0$ result is 2% below the best value. We conclude that even a rather coarse mesh can lead, when used with sufficient care, to very good results—note that the calculation for the $h=0.6a_0$ reso-

TABLE II. CO equilibrium distances d_0 and harmonic vibrational frequencies ν obtained from quadratic fits in the range [2.06:2.20] to the curves in Fig. 11. The entry labeled with ABINIT was obtained with the ABINIT code.[7]

	d_0 (a_0)	ν (cm^{-1})
$h=0.2$, unfiltered	2.12185	2117
$h=0.2$, filtered	2.12180	2119
$h=0.4$, unfiltered	2.10555	2299
$h=0.4$, filtered	2.10726	2153
$h=0.6$, unfiltered		
$h=0.6$, filtered	2.13422	2036
ABINIT	2.113	2149

lution is about 30 times faster than the best calculation.

Chelikowsky *et al.*[46] obtained a value of $2.135a_0$ for the equilibrium bond distance of the CO molecule, and a vibrational frequency of 2000 cm^{-1}, using a finite difference real space grid approach. Using an adaptive grid method, Gygi and Galli[47] obtained values of $2.132a_0$ and 2151 cm^{-1}. With the same pseudopotential as employed in our calculations and the same calculation parameters (grid spacing, box size, etc.), the ABINIT[7] code provides values of $2.113a_0$ and 2149 cm^{-1}, respectively. It is therefore seen that our approach leads to results in good agreement with previous calculations, both those carried out employing similar techniques to those used here, as well as plane-wave codes.

IV. CONCLUSIONS

In this paper, we have presented a full implementation of DFT employing a real space grid to represent Kohn-Sham orbitals, electron density, and potential. The eigenvalue/eigenstate pairs for a fixed potential are determined using a fourth-order factorization of the evolution operator in imaginary time. We have generalized the method here to nonlocal pseudopotentials of the Kleinman-Bylander type, which allow for quantitative simulations of most s, p, and d electron elements and complexes within the LDA. The high order of the factorization allows the method to efficiently locate the eigenpairs, with a computational cost that is typically one order of magnitude less than that required when employing the simpler second-order factorization. The rigorous fourth-order convergence is slightly compromised by numerical sacrifices that were made in the treatment of the nonlocal components of the potential, but this deviation from fourth order has no *practical* consequences.

In conventional iterative diagonalization schemes, one needs to calculate the action of the Hamiltonian on the trial wave functions. This can be costly if the representation used results in a dense matrix representation of the Hamiltonian. In the coordinate space representation, the Hamiltonian is sparse, and the action of the kinetic energy part of the diffusion operator can be done very efficiently using FFT, or entirely in real space, which allows for great speed. In other words, in a real-space representation, the action of the diffusion operator on a wave function is not much more costly than the action of the Hamiltonian on a wave function.

In order to reduce the computational effort, we have also implemented the filtering method by Ono and Hirose.[42] Some care must be taken to maintain (or reestablish) the normalization and orthogonality of the projection operators of the nonlocal potentials. The method then allows some reduction of the mesh resolution, which can lead to significant savings in computer time, in particular for larger systems like C_{60}. An alternative approach to filtering would be to use the Lagrange mesh method,[22] which has been recently illustrated in the context of electronic structure calculations.[23] We will explore this possibility in our subsequent work.

We have implemented the method in the context of DFT combined with the pseudopotential approximation, and illustrated its capabilities for a number of test systems, including

carbon monoxide, benzene, and C_{60}. These calculations allowed us to illustrate the capabilities of the method in terms of efficiency. Although all our tests were conducted for finite systems (molecules and clusters), the methodology is not restricted to these geometries, and can work equally well in periodic boundary conditions with few modifications. Therefore, this method should also find applicability in electronic structure calculations of solids, surfaces, etc. Throughout the paper we have assumed that there is only one projection operator in each angular momentum channel. It is quite feasible to generalize the effort to more than one projector, which is needed, for example, for the pseudopotentials of Ref. 48. The essential change needed is that Eq. (2.16) becomes a matrix equation that can be dealt with similarly to the treatment of the double commutator.

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[3] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).

[4] J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).

[5] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[6] M. D. Segall, P. J. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Condens. Matter **14**, 2717 (2002).

[7] X. Gonze *et al.*, Comput. Mater. Sci. **25**, 478 (2002).

[8] O. F. Sankey and D. J. Niklewski, Phys. Rev. B **40**, 3979 (1989).

[9] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquer, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).

[10] A. P. Horsfield, Phys. Rev. B **56**, 6594 (1997).

[11] T. Ozaki and H. Kino, Phys. Rev. B **72**, 045121 (2005).

[12] J. M. Frisch and *et al.*, GAUSSIAN 98 (2001).

[13] J. R. Chelikowsky, N. Troullier, and Y. Saad, Phys. Rev. Lett. **72**, 1240 (1994).

[14] E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **54**, 14362 (1996).

[15] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, Phys. Rev. B **71**, 035109 (2005).

[16] M. Heiskanen, T. Torsti, M. J. Puska, and R. M. Nieminen, Phys. Rev. B **63**, 245106 (2001).

[17] R. Schmid, M. Tafipolsky, P. H. Konig, and H. Kostler, Phys. Status Solidi B **243**, 1001 (2006).

[18] T. L. Beck, Rev. Mod. Phys. **72**, 1041 (2000).

[19] J.-L. Fattebert and F. Gygi, Phys. Rev. B **73**, 115124 (2006).

[20] J. E. Pask and P. A. Sterne, Modell. Simul. Mater. Sci. Eng. **13**, 71 (2005).

[21] Y. Liu, D. A. Yarne, and M. E. Tuckerman, Phys. Rev. B **68**, 125110 (2003).

[22] D. Baye, Phys. Status Solidi B **243**, 1095 (2006).

[23] K. Varga and S. T. Pantelides, Phys. Status Solidi B **243**, 1110 (2006); K. Varga, Z. Zhang and S. T. Pantelides, Phys. Rev. Lett. **93**, 176403 (2004).

[24] T. A. Arias, Rev. Mod. Phys. **71**, 267 (1999).

[25] R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan, and G. Beylkin, J. Chem. Phys. **121**, 11587 (2004).

[26] M. Payne, M. Teter, D. Allen, T. Arias, and J. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[27] K. Schwarz and P. Blaha, Comput. Mater. Sci. **28**, 259 (2003).

[28] W. E. Pickett, Comput. Phys. Rep. **9**, 115 (1989).

[29] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[30] V. Eyert, J. Comput. Phys. **124**, 271 (1996).

[31] M. Aichinger and E. Krotscheck, Comput. Mater. Sci. **34**, 188 (2005).

[32] M. Suzuki, J. Math. Phys. **32**, 400 (1991).

[33] M. Suzuki, in *Computer Simulation Studies in Condensed Matter Physics*, edited by D. P. Landau, K. K. Mon, and H.-B. Schüttler (Springer, Berlin, 1996), Vol. VIII, pp. 1–6.

[34] S. A. Chin, Phys. Lett. A **226**, 344 (1997).

[35] M. Aichinger, S. A. Chin, and E. Krotscheck, Comput. Phys. Commun. **71**, 197 (2005).

[36] M. Aichinger, S. A. Chin, E. Krotscheck, and E. Räsänen, Phys. Rev. B **73**, 195310 (2006).

[37] H. A. Forbert and S. A. Chin, Phys. Rev. B **63**, 144518 (2001).

[38] www.fftw.org.

[39] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[40] E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **52**, R5471 (1995).

[41] M. Tafipolsky and R. Schmid, J. Chem. Phys. **124**, 174102 (2006).

[42] T. Ono and K. Hirose, Phys. Rev. Lett. **82**, 5016 (1999).

[43] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[44] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[45] M. Fuchs and M. Scheffler, Comput. Phys. Commun. **119**, 67 (1999).

[46] J. R. Chelikowsky, N. Troullier, K. Wu, and Y. Saad, Phys. Rev. B **50**, 11355 (1994).

[47] F. Gygi and G. Galli, Phys. Rev. B **52**, R2229 (1995).

[48] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).