Preprint
UCRL-JRNL-202957

Linear Scaling First-Principles Molecular Dynamics with Controlled Accuracy

J.-L. Fattebert and F. Gygi

This article was submitted to Elsevier Science for publication in "Computer Physics Communications"

March 18, 2004

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

Approved for public release; further dissemination unlimited

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

Linear Scaling First-Principles Molecular Dynamics with Controlled Accuracy

Jean-Luc Fattebert, François Gygi

Center for Applied Scientific Computing (CASC)
Lawrence Livermore National Laboratory
P.O. Box 808, L-561, Livermore, CA 94551

Abstract

In our quest for accurate linear scaling first-principles molecular dynamics methods for pseudopotential DFT calculations, we investigate the accuracy of real-space grid approaches, with finite differences and spherical localization regions. We examine how the positions of the localization centers affect the accuracy and the convergence rate of the optimization process. In particular we investigate the accuracy of the atomic forces computation compared to the standard $O(N^3)$ approach. We show the exponential decay of the error on the energy and forces with the size of the localization regions for a variety of realistic physical systems. We propose a new algorithm to automatically adapt the localization centers during the ground state computation which allows for molecular dynamics simulations with diffusion processes. The combination of algorithms proposed lead to a genuine linear scaling First-Principles Molecular Dynamics method with controlled accuracy. We illustrate our approach with examples of microcanonical molecular dynamics with localized orbitals.

Key words: First-Principles molecular dynamics, Density Functional Theory, finite differences, linear scaling
PACS: 71.15-m, 71.15.Dx, 71.15.Ap, 71.15.Pd

1 Introduction

During the last decade, many researchers have investigated new numerical algorithms for electronic structure calculations with a computational cost that scales linearly with the size of the physical system (for a review, see for example Ref. [1,2]). Many innovative ideas have been proposed, and very successful approaches have been implemented, in particular for Tight-Binding methods, the crudest quantum simulation approach. Nonetheless no linear scaling method proposed so far has been able to replace in practice the standard and powerful

$O(N^3)$ plane-wave approach to carry out accurate Density Functional Theory (DFT) first-principles molecular dynamics (FPMD).

In methods based on a linear combinations of atomic orbitals (LCAO) or a similar numerical basis set of localized orbitals, $O(N)$ algorithms have been obtained to compute the one-particle Density Matrix (DM) for example (see [2] and references therein). The size of this matrix is given by the total number of numerical basis functions in the calculation which is in general less than 20 per atom. But LCAO approaches lack flexibility and adaptivity since there is no systematic way of refining the accuracy of the numerical approximations. This can lead to poor performance in changing environments typical of molecular dynamics. On the other hand, in the case of accurate methods based on a very large number of variational parameters — like Plane-Waves (PW) or grid-based calculations — that can be systematically increased to improve accuracy, it becomes impossible to build the DM, or even a sparse DM. In that case, there are hundreds of degrees of freedom per electronic wave function. It is then more appropriate to represent the electronic structure by a set of orbitals that represent the occupied subspace. In standard approaches, the solution is usually expressed by a set of eigenfunctions that span the subspace of the occupied orbitals. Using the sparsity of the Hamiltonian matrix, iterative eigensolvers are used to compute these functions in $O(N^3)$ operations. $O(N)$ algorithms can be obtained by looking for Wannier-like functions that can be truncated beyond a certain radius and still describe the electronic structure accurately. This linearizes the most costly parts of the algorithm such as building the Hamiltonian and overlap matrices in the occupied subspace, or computing the gradient of the energy functional. Working with a set of truncated orbitals then allows to use the same $O(N)$ techniques used in LCAO approaches to obtain a global linear scaling[3]. Several authors have reported algorithms describing strictly localized orbitals in real-space using either finite differences to discretize the Laplacian [3–6], finite elements methods [7,8] or plane-waves in restricted regions [9]. In these approaches, accuracy can be systematically improved by mesh refinement and localization regions enlargement.

While $O(N)$ molecular dynamics (MD) schemes have been developed by several authors for tight-binding methods [10–15], only limited ab initio MD have been carried out with linear scaling methods. Fabricius et al.[16] simulated a liquid silicon surface at high temperature with a minimal basis set of 4 atomic orbitals per atom, using a thermostat to keep a constant temperature. Tsuchida and Tsukada[8] carried out a molecular dynamics simulation with localized orbitals and a much larger number of degrees of freedom — using a Finite Elements approach — for a 64-atoms diamond supercell with fixed localization regions centered at the equilibrium positions. They observed a negative drift of the total energy. This approach does not allow general MD simulations involving diffusion processes and changes in chemical bonds. To our knowledge no direct comparison between forces obtained with strictly lo-

calized orbitals and extended orbitals have been reported so far. Note that ab initio investigations have been limited by their computational cost which is much larger than in TB, in particular for systems of sizes relevant for $O(N)$ approaches.

In this paper, we report our recent developments towards a FPMD algorithm with linear scaling based on finite difference discretization. Some of the elements on which we base our approach were published in Ref. [5]. In this paper we limit our study to systems with a finite band gap and focus on the effects of localization constraints on the accuracy of the solution of DFT equations. After describing the general $O(N^3)$ reference approach in Sec. 2, we introduce the concept of localized orbitals in Sec. 3. Then in Sec. 4, we propose a new algorithm to automatically adapt the centers of the localization regions according to the solution of the electronic structure problem. These adaptive localization centers (ALC) are typically close to the Maximally Localized Wannier Centers (MLWC) defined by Marzari and Vanderbilt [17]. They often result in more accurate solution of the electronic structure problem than atomic centers while also leading to a faster convergence rate in the orbital optimization process. In Sec. 5, we present a study of the accuracy of the energies and forces as functions of localization radius. Finally in Sec. 6, we propose a molecular dynamics algorithm for localized orbitals based on the concept of adapting localization centers on the fly. We illustrate our approach with two examples of microcanonical molecular dynamics with strictly localized orbitals.

2 General approach

We consider a physical system composed of N_a atoms in a computation domain Ω with periodic boundary conditions. We represent the electronic structure of such a system by a general non-orthogonal basis $\{\phi\}_{i=1}^N$ that spans the invariant subspace of N doubly occupied orbitals [1]. In this representation, the Kohn-Sham (KS) energy functional is given by (in atomic units)[18]

$$\begin{aligned} E_{KS}[\{\phi_i\}_{i=1}^N] = & \sum_{i,j=1}^{N} 2\left(S^{-1}\right)_{ij} \int_\Omega \phi_i(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\phi_j(\mathbf{r})d\mathbf{r} \\ & + \frac{1}{2}\int_\Omega\int_\Omega \frac{\rho_e(\mathbf{r}_1)\rho_e(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 + E_{XC}[\rho_e] \\ & + \sum_{i,j=1}^{N} 2\left(S^{-1}\right)_{ij} \int_\Omega \phi_i(\mathbf{r})(V_{ext}\phi_j)(\mathbf{r})d\mathbf{r}, \end{aligned} \tag{1}$$

[1] We use the Γ point only and thus can assume that the functions $\phi_i, i = 1, \ldots, N$ are real.

where ρ_e is the electronic density

$$\rho_e(\mathbf{r}) = 2 \sum_{i,j=1}^{N} (S^{-1})_{ij} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) \tag{2}$$

and the overlap matrix is defined by its entries

$$S_{ij} = \int_{\Omega} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}), \quad i, j = 1, \ldots, N.$$

The first two terms in the energy functional represent the kinetic energy of the electrons and the electrostatic energy of interaction between electrons. The energy E_{XC} models the exchange and correlation between electrons. In this paper, we use the local density approximation (LDA), or the first-principles exchange-correlation functional proposed by Perdew-Burke-Ernzerhof (PBE) [19] which often provides results in better agreement with experiments and is appropriate for a grid-based implementation. In the last term of Eq.(1), the potential V_{ext} represents the total potential produced by the atomic nuclei. In practice we use norm-conserving pseudopotentials in the Kleinman-Bylander form [20,21].

We discretize the energy functional (1) by finite differences (FD), using a fourth order compact scheme — Mehrstellenverfahren — for the Laplacian. The details of the FD methodology and its implementation on massively parallel computers using domain decomposition are described in Ref. [22]. We denote by M the number of grid nodes, i.e. the number of degrees of freedom for each wave function.

We denote the solution (or trial solution) $\{\vec{\phi}\}_{i=1}^{N}$ that minimizes Eq.(1) in a matrix form $\Phi = \left(\vec{\phi}_1, \ldots, \vec{\phi}_N\right)$ ($M \times N$ matrix). We assume that Φ always has full column rank. If Φ minimizes (1), $\Phi \cdot B$ is also a minimum for any nonsingular $N \times N$ matrix B. In these notations, we can write the gradient of the discretized energy functional as [5]

$$\nabla_{\Phi} E = 2 \left(H\Phi - \Phi S^{-1} H^{(\Phi)} \right) = (\nabla_{\phi_1} E, \ldots, \nabla_{\phi_N} E) \tag{3}$$

where H is the DFT Hamiltonian operator associated to (1) and $H^{(\Phi)}$ is its matrix representation in the basis Φ, i.e.

$$H^{(\Phi)} = \Phi^T H \Phi.$$

The gradient (3) is a contravariant vector expressed for the basis Φ. It is related to the gradient expressed for the basis of the Ritz vectors Ψ by the

relation

$$\nabla_\Psi E = \nabla_\Phi E \cdot C$$

where C is a $N \times N$ matrix solution of the eigenvalue problem

$$H^{(\Phi)} C = SC\Lambda$$

with Λ an $N \times N$ diagonal matrix.

This gradient is the basis for many iterative algorithms to search for the minimum of the energy functional (1). One key ingredient for a fast convergence is a good preconditioner[23]. For a real-space discretization, an appropriate technique is a multigrid preconditioner based on the Laplacian. It provides a convergence rate independent of the size of the discretization grid for a simple steepest descent algorithm [24]. This approach is particularly efficient when some unoccupied states are included in the calculation.

If no unoccupied states are needed, we found that a simple block preconditioned steepest descent algorithm with fixed shift can be sped up significantly by using the extrapolation scheme of Anderson [25]. In the nonorthogonal basis Φ, we write this extrapolation scheme as

$$\bar{\Phi}^{(l)} := \Phi^{(l)} + \sum_{j=1}^{m} \theta_j^{(l)} \left(\Phi^{(l-j)} - \Phi^{(l)} \right)$$

where $\Phi^{(l)}$ denote the trial solution Φ at step l. The coefficients $\theta_j^{(l)} \in \mathcal{R}$ are determined by the solution of the linear system

$$\sum_{j=1}^{m} \left(R^{(l)} - R^{(l-i)}, R^{(l)} - R^{(l-j)} \right) \theta_j^{(l)} = \left(R^{(l)} - R^{(l-i)}, R^{(l)} \right), \; i = 1, \ldots, m \quad (4)$$

which corresponds to minimizing the norm of a coupled iterative sequence $R^{(l)}$, extrapolated according to the same scheme

$$\bar{R} := R^{(l)} + \sum_{j=1}^{m} \theta_j^{(l)} \left(R^{(l-j)} - R^{(l)} \right).$$

Finally, the new trial solution is given by

$$\Phi^{(l+1)} = \bar{\Phi}^{(l)} + \beta \bar{R}^{(l)}.$$

In practice, we use the preconditioned steepest descent — negative gradient — for the sequence $R^{(l)}$ so that solving (4) minimizes the preconditioned residual of the eigenvalue problem. We usually choose $\beta = 1$. In the space of the nonorthogonal wave functions, the natural scalar product is

$$(U, V) = Tr(S^{-1} U^T V) \quad (5)$$

for U and V $M \times N$ matrices. This scalar product has the property of being independent of the choice of the basis Φ we choose. It can however become computationally very expensive for large problems. To reduce its cost, we drop the matrix S^{-1} in Eq.(5). To limit the effects of this change, we orthonormalize the orbitals at regular intervals (every 20-30 iterations). Numerical tests show no major difference using this approximation.

Anderson's extrapolation scheme was designed to solve nonlinear equations [25]. In the case of an eigenvalue problem like the KS equations, the residual vanishes not only for the lowest eigenvalues we are interested in, but for any set of eigenvalues. In order to avoid being trapped in undesired solutions, some caution is required in particular during the first few steps when the trial solution is far from the ground state. In practice, we avoid problems by choosing a "safety" interval for θ_j^l outside of which the extrapolation is turned off (very large absolute values) or truncated to be inside the safety interval.

All the results presented in this paper were obtained using the block Anderson extrapolation scheme described above with multigrid preconditioning. No unoccupied states were included in the calculations. In the iterative process, the self-consistent potential is updated at each step, *i.e.* after updating all the wave functions only once. Our numerical experiments show little improvements when using $m > 1$ so that all the results presented here were obtained with $m = 1$.

3 Localization and partial linear scaling

Since we are only interested in finding the subspace that minimizes (1), we can choose to represent this subspace by a basis of *maximally localized* functions in some sense. If each localized function can be represented by a finite number of parameters that does not grow with the size of the physical system, it naturally leads to a better scaling. To determine such a solution we explicitly define localization regions (LR) that restrict the support of each orbital. These LR should not be too small so that the problem with constraints admits a solution that approximates well the solution of the problem without constraints.

We choose to use spherical LR. Localization is imposed by forcing the orbitals solution of the electronic structure problem to be strictly zero outside of a spherical region of radius R_c. The problem is then to minimize the energy functional (1) under localization constraints. This is done in practice by truncating the gradients (3) used in the iterative solver, applying a projector P_i to $\nabla_{\phi_i} E, i = 1, \ldots, N$. Before applying a multigrid preconditioner, a smooth mask is applied to $P_i(\nabla_{\phi_i} E)$ to remove the sharp discontinuities at the boundaries of the LRs. The preconditioned gradient leads to corrections at each step

of the iterative optimization procedure that keep each orbital localized [5]. Orbitals truncation linearizes the computational cost of S, $H^{(\Phi)}$, as well as $\nabla_\Phi E$ in Eq. (3) and ρ in Eq. (2), the most expensive parts of the algorithm — which would otherwise scale like $O(N^2M)$ where M is the size of the discretization grid. The inverse of the overlap matrix S^{-1} is computed exactly in $O(N^3)$ operations. This leads to a global algorithm with a remaining part with cubic scaling. However this part is not a bottleneck for a large range of problems since N is usually much smaller than M. Also, keeping this term exact allows us to focus on the effect of localization constraints on the orbitals only.

With localization constraints, the minimization process can be carried out using Anderson's extrapolation scheme as described above. Such a scheme is particularly well suited since it does not delocalize the trial solution. It only determines an appropriate weight for the extrapolation and correction contributions. As opposed to the case of extended orbitals, no orthonormalization can be applied to reduce S to the identity in the scalar product (5). Instead we only normalize the functions ϕ to keep the relative weights of all the functions equal in the process. Numerical experiments indicate that this approach does not slow down the iterative process compared to using the scalar product with S^{-1}.

4 Adaptive localization centers

In previous approaches, LR have been centered either on atoms [5,8] or on bonds [6], based on a priori knowledge of the chemistry. In many cases, such a choice of localization centers is not straightforward — like in disordered systems for example. Moreover, these centers may not be optimal from the accuracy point of view. Also, attaching localization centers at bonds or atoms may not be appropriate for general molecular dynamics. Recent development in Maximally Localized Wannier Functions (MLWF) theory [17] suggest that centering localized orbitals at Maximally Localized Wannier Centers (MLWC) may be a much better approach. However, finding the exact Wannier centers requires in general $O(N^3)$ operations[26]. Here we propose a simple approach which consists in centering the localization region of each orbital at the centroid of charge of the orbital itself. We use an iterative approach in which the localization regions adapt themselves and "follow" the orbitals during the minimization procedure. The algorithm can be summarized as follows:

Until convergence, do:

- *Optimize localized non-orthogonal orbital for fixed LR (about 20-40 iterations)*

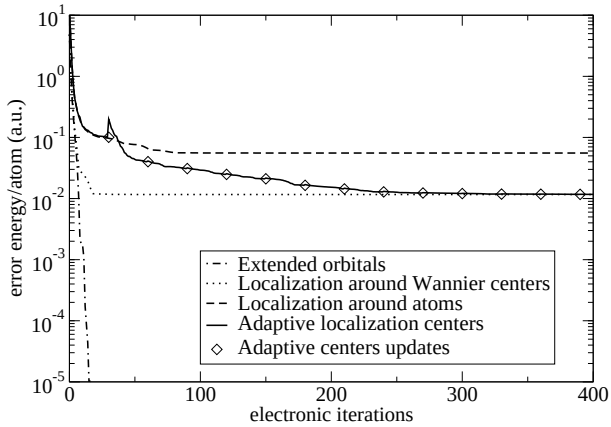

Fig. 1. Convergence and error in the computed KS energy for silicon bulk (64 atoms in a cubic cell) using various localization schemes.

- *Compute the centroids of charge:* $\mathbf{Q}_i =< \phi_i|\mathbf{X}|\phi_i >, i = 1, \ldots, N$
- *Move localization centers to* $\mathbf{Q}_i$ *to define new LR.*

Here $\mathbf{X}$ denotes the position operator. For periodic boundary conditions, the position operator for extended systems proposed by Resta [27] should be used. We call Adaptive Localization Centers (ALC) the centers obtained by this procedure. Note that in the particular case of MLWF (without any localization constraint), the centroids of charge are exactly the MLWC.

We now illustrate the behavior of this algorithm in some applications. For the first test we use a perfect silicon bulk cell with 64 atoms. In this case the MLWC are exactly located in the middle of the bonds. In Fig.1, we compare the results of various approaches by measuring the residual error on the energy due to the localization constraint on the orbitals, as well as the convergence rate. The reference energy is the value obtained by the standard $O(N^3)$ algorithm without any localization constraints. First, we observe that localization on the bonds (equivalent to MLWC in this case) reduces the error by a factor 5 compared to localization on the atoms. Secondly, we can see that in the adaptive procedure — which starts with centers on the atoms — the error converges to the same value as the one obtained with centers on the bonds. Actually the localization centers themselves converge to the MLWC. Note that this is not a general result but a consequence of the high symmetry of the problem. Usually we observe that the energy obtained with ALC is slightly higher than with MLWC, but this difference is very small in comparison with centering on the atoms.

We consider then a more interesting disordered system: hydrogen under pressure. In this case, not only hydrogen dimers are found, but also individual atoms, trimers, and so on. This makes it difficult to center a priori localization regions on bonds. For a test cell of 992 atoms, we show in Fig. 2 how the adaptive algorithm finds a good solution, starting from 496 initial centers

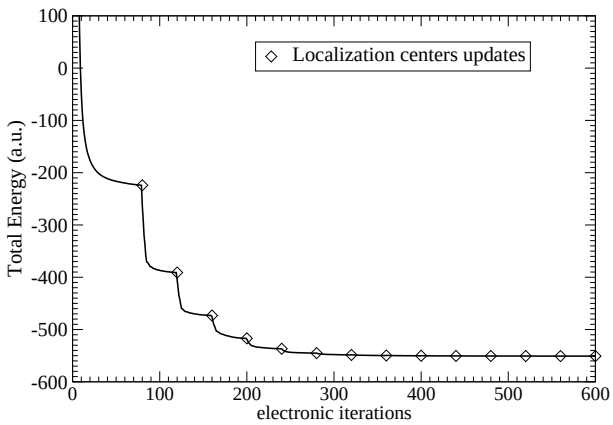

Fig. 2. Convergence of the Adaptive Localization Centers algorithm for a sample of disordered hydrogen (992 atoms in a cubic cell of 28.26 a.u. side length). The initial 446 localizations centers are randomly distributed at 446 different atomic positions.

distributed on randomly chosen atoms. Figure 2 shows the convergence of the total energy with adaptation of the localization regions.

Note that in both tests described above, we start from a configuration of centers very far from the optimal solution. This explains why a large number of electronic iterations is required to converge.

On the other hand, we also notice that convergence towards the ground state is often much faster for localization regions centered at MLWC or ALC, than for centers at atomic positions. To understand why such phenomenon may happen, we consider a bulk system like silicon. We realize that if we put 2 localized orbitals on each atom, a global *arrangement* is required to find a set of localized orbitals actually located at mid-distance between pairs of atoms, with each orbital associated by localization region to one of the two atoms. The iterative solver has to associate 2 bonds to each atom to center the 2 orbitals associated to this atom. This can slow down convergence significantly or even drive the iterative process into a local minimum where 2 orbitals try to settle on the same bond and no orbital occupies another bond elsewhere in the system. If localizations regions are positioned on the bonds and *own* one orbital each, this problem is avoided. Another way of avoiding such local minimas would be to compute more orbitals than needed, including unoccupied states in the calculation as proposed in Ref. [5]. This gives more possibilities to quickly *fill* all the bonds. In the approach proposed by Kim et al[15], these additional states become linearly dependent at convergence, helping the convergence process but without contributing to the final solution.

5 Effect of localization constraints on atomic forces

The ability to run molecular dynamics with localized orbitals of accuracy comparable to standard $O(N^3)$ simulations depends critically on the accuracy of the atomic forces. This error is a more significant measure of global accuracy than error on the total energy. As in standard approaches, the force on an ion is obtained by evaluating a derivative of the energy with respect to the ionic position. Expressed for the nonorthogonal solution Φ, the component α of the force on atom k is given by

$$\frac{dE}{dR_{k,\alpha}} = \frac{\partial E}{\partial R_{k,\alpha}} + \sum_{j=1}^{N} \frac{\delta E}{\delta \phi_j} \frac{d\phi_j}{dR_{k,\alpha}} + \sum_{k=1}^{N} \sum_{\beta=1}^{3} \frac{\delta E}{\delta Q_{k,\beta}} \frac{dQ_{k,\beta}}{dR_{k,\alpha}}. \tag{6}$$

The second term on the right hand side vanishes at the ground state for extended orbitals. If we consider that the functions ϕ_i are restricted to be zero outside of the LRs, this term vanishes too if we solve exactly

$$P_j \left(\frac{\delta E}{\delta \phi_j} \right) = 0, \; j = 1, \ldots, N, \tag{7}$$

where P_j is the localization projector discussed in Section 3. This is impractical since using the left hand side of (7) — a tensorially incorrect gradient — leads to very slow convergence[6,28]. On the other hand, using the gradient (3), we actually solve

$$P_j \left(\sum_{i=1}^{N} S_{ij} \frac{\delta E}{\delta \phi_i} \right) = 0, \; j = 1, \ldots, N. \tag{8}$$

while (7) is not satisfied. However we know that the second term on the right hand side of Eq.(6) goes to zero for large localization radius. Since this term is difficult to estimate, we consider it as part of the error introduced by localization constraints and drop it.

The third term in Eq.(6) corresponds to the Pulay forces described in numerical methods using a limited basis set attached to the atomic positions. The situation is different here since the orbitals are not explicitly attached to the atoms. Our claim is that $\delta E / \delta Q_{k,\beta}$ is usually very small and can be neglected. This term would be exactly zero if the localization centers were minimizing E. This is not the case for ALC and appears to be difficult to satisfy in practice. As will be shown in our numerical results below, we claim that the choice of localization centers does not have much influence on the total energy *as long as they are close enough to the configuration that minimizes* E, i.e. the energy

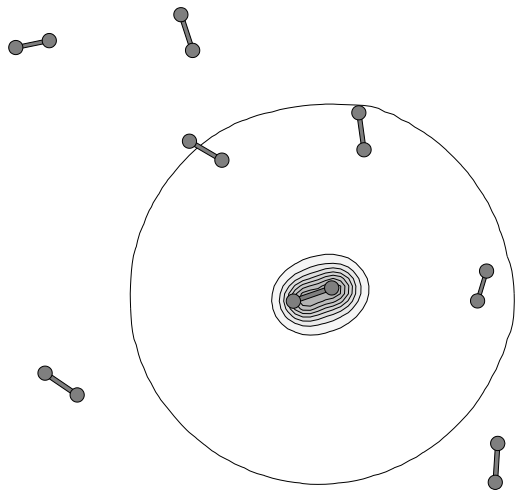

Fig. 3. Contour plot of a localized orbital in a slice through a deuterium sample at 1000 K with $R_c = 7a.u.$. The projections of all the atoms located within a distance of 2 a.u. from the slicing plane are also plotted.

surface is very flat near this minimum. This can be explained by the observation that moving a localization region only changes boundary conditions for an orbital, which is a minor perturbation for a function that goes smoothly to zero at these boundaries (see Fig.3 for example). This means that the first variation of the energy with respect to the localization centers is in general non-zero but very small.

Since we neglect the third term in Eq.(6) and consider the second term as part of the localization error, the forces are evaluated as usual by simply computing the partial derivative of the energy with respect to the ionic positions (Hellmann-Feynman theorem). In practice, these derivatives are evaluated by computing changes in the total energy for small virtual displacements of the ions, without the need for full self-consistent calculations. This computation is local for each atom and thus also scales linearly with the number of atoms.

To estimate the error on the forces introduced by localization constraints, we have carried out simulations of various systems with localization regions of various sizes and compared the forces obtained with those calculated with standard extended orbitals. The ALC were obtained using the algorithm described in Section 4, while the MLWC were computed according to Ref. [26].

Our first test system is a sample of 72 deuterium atoms in a cubic cell of size 21.24 Bohr from a snapshot of a molecular dynamics simulation at 1000K [2]. Looking at the numerical results obtained for this system (Fig.4), we observe

[2] This corresponds to a pressure of 1.5 GPa.

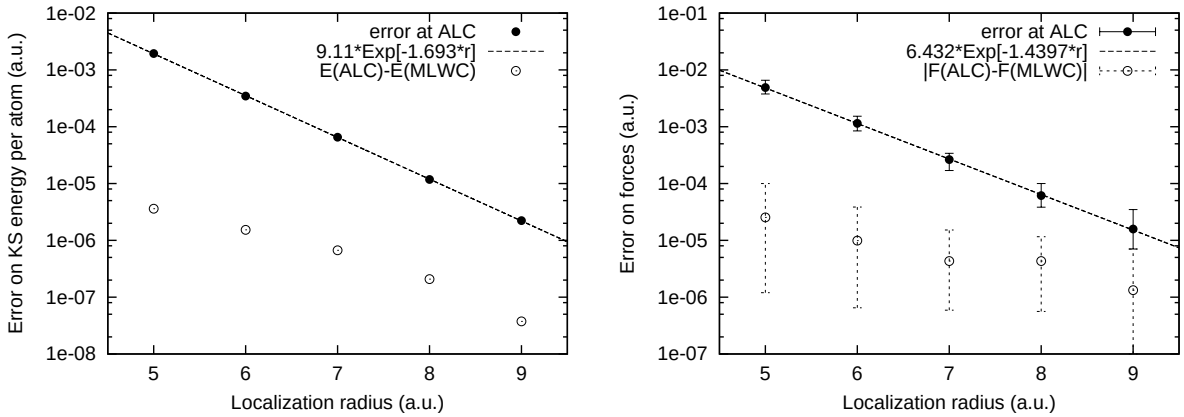

Fig. 4. Error in energy/atom and in forces for a snapshot of deuterium (72 atoms) as a function of the truncation radius, centering localization regions at ALC or MLWC. The computational cell is cubic of side length 21.24 a.u.

that the differences in energies and forces obtained by using either ALC or MLWC are very small in comparison with the error introduced by localization. We measured the average norm of the error on the forces compared to the calculation with extended orbitals. The error bars show the minimum and maximum errors measured in the sample. We also observe that both the error on the energy and the average error on the forces decay exponentially with the size of the localization radius. A localized orbital for this system is shown in Fig. 3.

To investigate the general validity of the observations on the deuterium sample, we measured the errors introduced by localization constraints for various other systems. We considered a sample of liquid water with 54 molecules at ambient conditions, a peptide chain in solution (representing the solvent by a continuum model[29]) composed of 14 amino acids, a sample of liquid lithium hydride, as well as a silicon bulk cell. For the silicon cell, atoms were randomly displaced from their equilibrium positions to have non-zero forces. The results plotted in Fig.5, 6, 8 and 7 confirm the exponential decay law for the error on the energies and forces. Results on perfect bulk Silicon showing an exponential decay of the error on the energy with respect to the size of the localization regions have also been reported elsewhere [6].

We observe that for all these systems except the silicon crystal, we can find localization radii of reasonable sizes (7-9 a.u.) which generate errors on the forces of the order $10^{-4} a.u.$, a tolerance often used in geometry optimization.

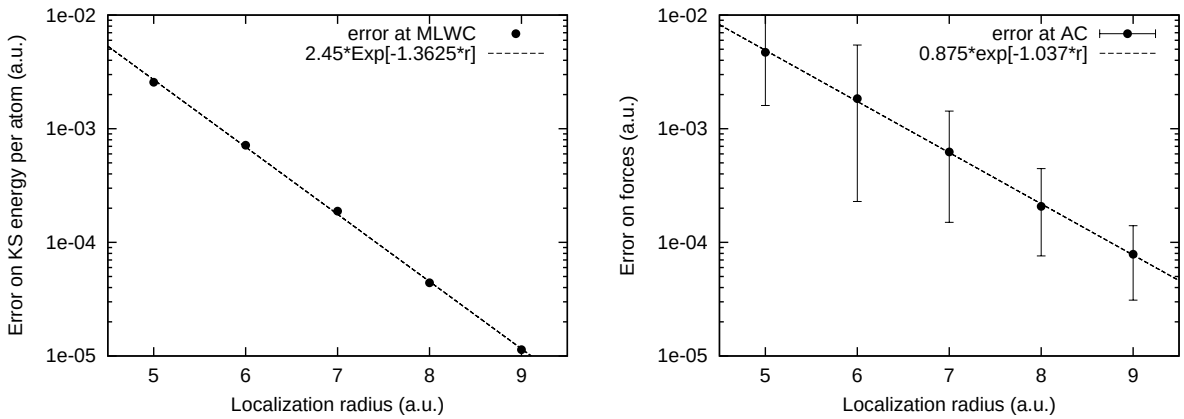

Fig. 5. Error in energy/atom and in forces for a snapshot of liquid water (54 molecules) as a function of the truncation radius. The computational cell is cubic of side length 22.19 a.u.

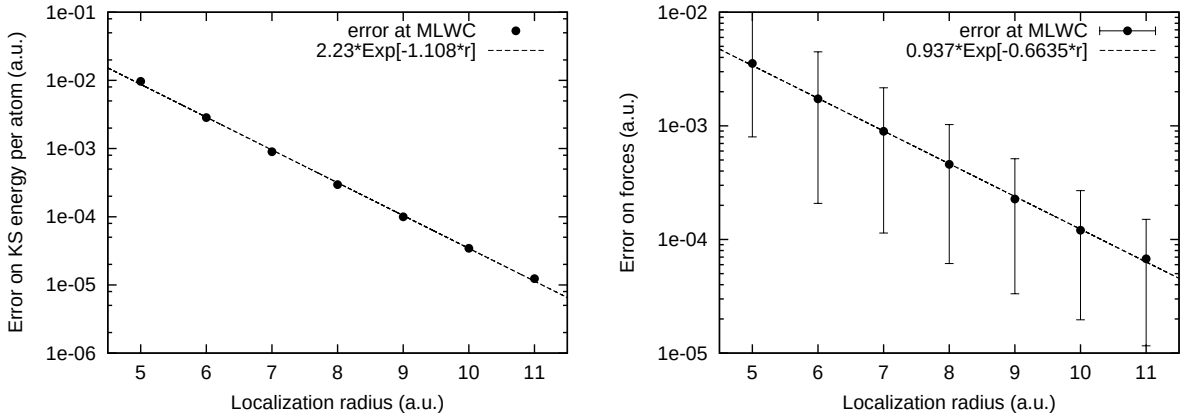

Fig. 6. Error in energy/atom and in forces for a snapshot of liquid LiH (216 atoms) as a function of the truncation radius. The computational cell is cubic of side length 26.8 a.u.

6 Molecular Dynamics with Adaptive localization centers

Our new adaptive algorithm to define localization regions provides a general tool to compute the electronic structure under localization constraints defined by a localization radius only. This procedure does not depend explicitly on atomic positions. We now propose an algorithm based on this adaptive scheme to carry out general Born-Oppenheimer molecular dynamics with strictly localized orbitals. The basic idea is to adapt the localization centers during the dynamics according to the algorithm proposed in Section 4, and use a simple extrapolation scheme to define a new localization region when moving the atoms. Specifically, we use the following algorithm:

For each ionic step k, do:

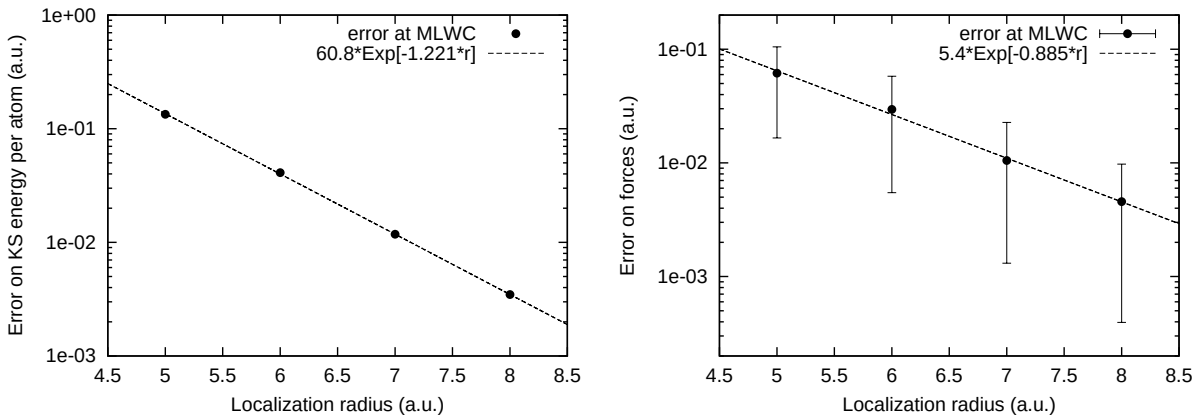

Fig. 7. Error in energy/atom and in forces for bulk silicon (216 atoms) as a function of the truncation radius. The computational cell is cubic of side length 30.78 a.u. The atoms are randomly displaced from their equilibrium position according to a Gaussian distribution of width 0.1 a.u.

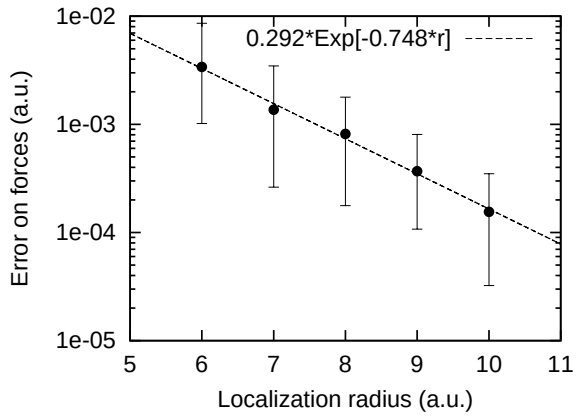

Fig. 8. Error in forces for a peptide chain in solution as a function of the truncation radius.

- *Initialize trial orbitals by extrapolation from 2 previous steps:*

$$\phi_i^{(k)} = 2\phi_i^{(k-1)} - \phi_i^{(k-2)},\ i = 1, \ldots, N$$

- *Initialize localization centers* $\mathbf{Q}_i$ *by extrapolation from 2 previous steps:*

$$\mathbf{Q}_i^{(k)} = 2\mathbf{Q}_i^{(k-1)} - \mathbf{Q}_i^{(k-2)},\ i = 1, \ldots, N$$

- *Optimize localized non-orthogonal orbital* $\phi_i^{(k)}$ *for LR centered at* $\mathbf{Q}_i^{(k)}, i = 1, \ldots, N$ *until convergence criterion satisfied.*
- *Adapt localization centers:*

$$\mathbf{Q}_i^{(k)} =< \phi_i|\mathbf{X}|\phi_i >,\ i = 1, \ldots, N$$

- *Move atoms according to Verlet scheme:*

$$\mathbf{R}_I^{(k+1)} = 2\mathbf{R}_I^{(k)} - \mathbf{R}_I^{(k-1)} + \mathbf{F}_I^{(k)}(\triangle t)^2/M_I,\ I = 1, \ldots, N_a$$

where $\triangle t$ is the time step.

Note that localization regions are adapted only once per ionic step. More adaptations could be used if needed. Note also that in the orbitals extrapolation procedure, one should take into account the new localization regions and truncate the functions ϕ at nodes that are outside of the new localization regions.

We tested our MD algorithm for the sample of deuterium introduced in Section 5. We carried out a molecular dynamics at 1000 K. Starting from an equilibrated sample, we ran the simulation twice, once with a localization radius of 7 a.u. and once with extended orbitals, to study the effects of localization. We used a time step of 15 a.u. For the run with localized orbitals a convergence criterion of 10^{-6} on the total energy change — to be satisfied in two consecutive iterations — was used. This criterion was reached at each ionic step after an average of 31 iterations, *i.e.* 31 updates for each wave function. A tighter tolerance of 10^{-8} was required to conserve energy in the extended orbitals calculation, reached in average in 17 iterations. Figure 9 illustrates the excellent conservation of the total energy during a 1 ps run. This system exhibits considerable diffusion as illustrated in Fig. 10. The success of this simulation shows that the adaptive centers follow closely the dimer molecules during the diffusion process. As expected we also observe that the atoms follow qualitatively the same trajectories for some time in the 2 simulations — the first 0.25 ps — until a bifurcation occurs.

Note that in LCAO methods, energy discontinuities occur when atoms move in or out of a localization region[12]. In the approach presented here, this problem is avoided since localization regions are not defined by atoms located within a given radius, but by a region of the grid independently of the atomic positions.

We also tested our new algorithm on a sample of liquid water at ambient conditions. We started from an equilibrated sample composed of 64 molecules at 300 K. Once again we compare runs with and without localization constraints (Fig.11) over 0.14 ps. We used a time step of 10 a.u. For the run with localized orbitals we chose a radius of 8 a.u. and a convergence criterion of $2 \cdot 10^{-6} a.u.$ on the total KS energy at each ionic step. An average of 22 updates per wave function was required to reach this tolerance. In comparison, we used a tolerance of $10^{-7} a.u.$ reached in an average of 15 iterations for the run with extended orbitals. A small negative drift of 26 mHa/ps on the total energy was observed for the run with localized orbitals. The KS energy for the run with localized orbitals follows closely the KS energy curve obtained without localization, showing that the two runs generate very similar trajectories.

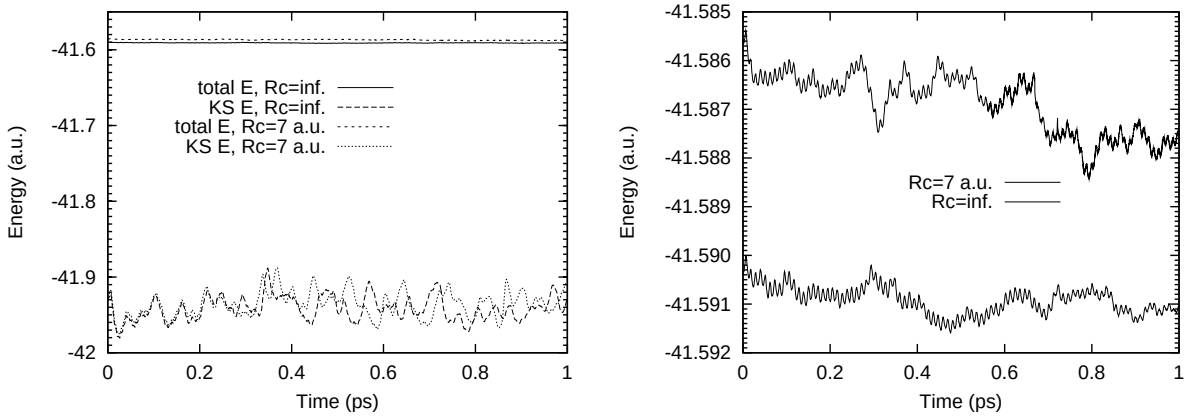

Fig. 9. Total energy and KS energy during a molecular dynamics simulation of deuterium at 1000 K without localization and with $R_c = 7a.u.$ respectively. Zooming in on the total energy we observe a very small drift of about 2 [mHa/ps] for the run with localized orbitals.

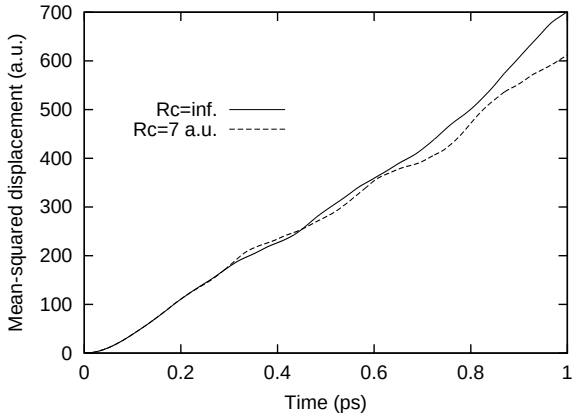

Fig. 10. Diffusion during a molecular dynamics simulation of deuterium at 1000 K without localization and with $R_c = 7a.u.$ respectively .

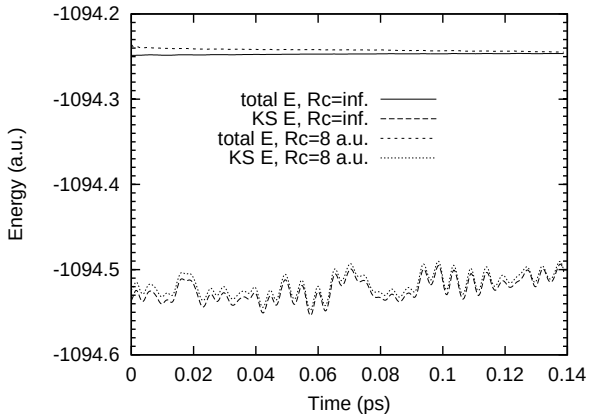

Fig. 11. Total energy and KS energy during a molecular dynamics simulation of water at 300 K without localization and with $R_c = 8a.u.$ respectively.

7 Concluding remarks

As noted by several other authors it is sometimes difficult to reach perfectly well converged energies with localization constraints. While $O(N^3)$ iterative solvers can usually reach total energies converged to 8 digits or more in less than 100 iterations — at least for systems with a finite band gap and no charge sloshing — $O(N)$ algorithms may take much longer to converge to the same accuracy. We observe however that slow convergence appears only after reaching a tolerance of the order of the error introduced by the localization constraints. This means that pushing convergence further reduces the error by only a small fraction compared to the calculation with extended orbitals. This is also true for ionic forces. For molecular dynamics simulations, this is not a problem as long as this error is not systematic and does not lead to a drift of the total energy.

One early motivation of real-space methods was their inherent parallelism. With our real-space code MGmol[30], we were able to show an excellent scaling up to 1600 CPUs — on the Lawrence Livermore National laboratory Linux cluster MCR — for large problems with extended orbitals ($O(N^3)$). With the introduction of localized orbitals, this advantage remains, at least for systems with an homogeneous spatial distribution of the electrons. In our current implementation, localized orbitals are grouped together in sets of non-overlapping orbitals. Each set is stored in one global array distributed over all the processors, like an extended orbital. This way all the functions in one set are computed in parallel. Beside issues related to implementation and the small $O(N^3)$ part of the algorithm, a global linear scaling can be observed.

In summary, we have shown that a real-space discretization approach with localized orbitals allows for general first-principles molecular dynamics simulations. We have presented the first systematic comparison of such an approach with the equivalent $O(N^3)$ approach. In our numerical experiments, we observe that the total energy shows only a very small drift and thus would require a very limited coupling to a thermostat. Global accuracy can be systematically increased by varying both the localization radius and the grid spacing. The error on forces and energies decay exponentially with the size of the localization regions, while the accuracy of the discretization scheme is $O(h^4)$ for the Mehrstellen finite difference stencil we use on a mesh of grid spacing h. This makes the approach presented in this paper a genuine $O(N)$ ab initio molecular dynamics method with controlled accuracy.

Acknowledgements

We wish to thank S. Bonev, V. Couch, T. Ogitsu, J.C. Grossman and E. Schwegler for providing snapshots of their molecular dynamics simulations. The numerical simulations were carried out on the Lawrence Livermore National Laboratory Linux cluster MCR. This work was performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

References

[1] G. Galli, Linear scaling methods for electronic structure calculations and quantum molecular dynamics simulations, Curr. Opin. Solid State Mater. Sci. 1 (6) (1996) 864–874.

[2] S. Goedecker, Linear scaling electronic structure methods, Rev. Mod. Phys. 71 (4) (1999) 1085–1123.

[3] E. Hernandez, M. J. Gillan, C. M. Goringe, Linear-scaling density-functional-theory technique: The density-matrix approach, Phys. Rev. B 53 (11) (1996) 7147–7157.

[4] T. Hoshi, T. Fujiwara, Fully-selfconsistent electronic-structure calculation using nonorthogonal localized orbitals within a finite-difference real-space scheme and ultrasoft pseudopotential, J. Phys. Soc. Jpn. 66 (12) (1997) 3710–3713.

[5] J.-L. Fattebert, J. Bernholc, Towards grid-based O(N) density-functional theory methods: optimized non-orthogonal orbitals and multigrid acceleration, Phys. Rev. B 62 (3) (2000) 1713–1722.

[6] J. Mortensen, M. Parrinello, Localized orbitals in silicon, J. Phys.: Condens. Matter 13 (2001) 5731–5741.

[7] E. Hernandez, M. J. Gillan, C. M. Goringe, Basis functions for linear-scaling first-principles calculations, Phys. Rev. B 55 (20) (1997) 13485.

[8] E. Tsuchida, M. Tsukada, Large-scale electronic-structure calculations based on the adaptive finite-element method, J. Phys. Soc. Jap. 67 (11) (1998) 3844–3858.

[9] A. Mostofi, C. Skylaris, P. Haynes, M. Payne, Total-energy calculations on a real space grid with localized functions and a plane-wave basis, Comp. Phys. Comm. 147 (3) (2002) 788–802.

[10] F. Mauri, G. Galli, R. Car, Orbital formulation for electronic-structure calculations with linear system-size scaling, Phys. Rev. B 47 (15) (1993) R9973–9976.

[11] S. Goedecker, L. Colombo, Efficient linear scaling algorithm for tight-binding molecular dynamics, Phys. Rev. Lett. 73 (1994) 122–125.

[12] F. Mauri, G. Galli, Electronic-structure calculations and molecular-dynamics simulations with linear system-size scaling, Phys. Rev. B 50 (7) (1994) 4316–4326.

[13] P. Ordejon, D. A. Drabold, R. M. Martin, M. P. Grumbach, Linear system size scaling methods for electronic structure calculations, Phys. Rev. B 51 (3) (1995) 1456–1476.

[14] S. Goedecker, M. Teter, Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals, Phys. Rev. B 51 (15) (1995) 9455–9464.

[15] J. Kim, F. Mauri, G. Galli, Total-energy global optimizations using nonorthogonal localized orbitals, Phys. Rev. B 52 (3) (1995) 1640–1648.

[16] G. Fabricius, E. Artacho, D. Sanchez-Portal, P. Ordejon, D. Drabold, J. Soler, Atomic layering at the liquid silicon surface: A first-priciples simulation, Phys. Rev. B 60 (24) (1999) R16283.

[17] N. Marzari, D. Vanderbilt, Maximally localized generalized Wannier functions for composite energy bands, Phys. Rev. B 56 (20) (1997) 12847.

[18] I. Stich, R. Car, M. Parrinello, S. Baroni, Conjugate gradient minimization of the energy functional: A new method for electronic structure calculation, Phys. Rev. B 39 (8) (1989) 4997–5004.

[19] J. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (18) (1996) 3865–3868.

[20] D. R. Hamann, Generalized norm-conserving pseudopotentials, Phys. Rev. B 40 (5) (1989) 2980–2987.

[21] N. Troullier, J. Martins, Efficient pseudopotentials for plane-wave calculations, Phys. Rev. B 43 (3) (1991) 1993–2006.

[22] E. L. Briggs, D. J. Sullivan, J. Bernholc, Real-space multigrid-based approach to large-scale electronic structure calculations, Phys. Rev. B 54 (20) (1996) 14362–14375.

[23] M. C. Payne, M. P. Teter, D. C. Allan, T. Arias, J. Joannopoulos, Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients, Rev. Mod. Phys. 64 (4) (1992) 1045–1097.

[24] J.-L. Fattebert, M. Buongiorno Nardelli, Finite difference methods for ab initio electronic structure and quantum transport calculations of nanostructures, in: P. Ciarlet, C. Le Bris (Eds.), Handbook of Numerical Analysis, Vol. X, Elsevier Science, 2003.

[25] D. Anderson, Iterative procedures for nonlinear integral equations, J. Ass. Comput. Mach. 12 (4) (1965) 547–560.

[26] F. Gygi, J.-L. Fattebert, E. Schwegler, Computation of maximally localized Wannier functions using a simultaneous diagonalization algorithm, Comput. Phys. Comm. 155 (2003) 1–6.

[27] R. Resta, Quantum-mechanical position operator in extended systems, Phys. Rev. Lett. 80 (9) (1998) 1800–1803.

[28] U. Stephan, Comparison of the convergence properties of linear-scaling electronic-structure schemes for nonorthogonal bases, Phys. Rev. B 62 (24) (2000) 16412–16424.

[29] J.-L. Fattebert, F. Gygi, Density functional theory for efficient ab initio molecular dynamics simulations in solution, J. Comput. Chem. 23 (2002) 662–666.

[30] J.-L. Fattebert, MGmol, a real-space multigrid first-principles molecular dynamics code, CASC, Lawrence Livermore National Laboratory, 2001-2004.