Real-Space Density Functional Theory with Localized Orbitals and Multiwavelets

S. R. Jensen, J. Jusélius, A. Durdek, P. Wind, T. Flåand L. Frediani

Abstract

Real-space methods for quantum chemistry have been gaining popularity in recent years. Despite the still significant overhead with respect to more common Gaussian-Type Orbitals and plane waves, they possess very attractive properties, especially when combined with a rigorous mathematical framework such as Multiresolution Analysis and Multiwavelets. Among the distinctive features are a localized orthonormal basis with disjoint support which is ideal in combination with massively parallel computing architectures, full adaptivity for automatic local refinement around nuclei, vanishing moments for sparse representations of functions and operators, rigorous error control with respect to basis-set limit results (a "golden standard" of Quantum Chemistry). We have implemented a Self Consistent Field solver within the Multiwavelet framework for restricted and unrestricted Hartree-Fock and Density Functional Theory. Our solver is based on a preconditioned steepest descent step combined with a Krylov accelerator. Among the distinctive features are the use of localized orbitals throughout and the construction of the full Fock matrix without any reference to the kinetic energy. In our results we have shown that we are able to attain high accuracy for Density Functional Theory as well as Hartree-Fock, both for restricted closed shell cases as well as unrestricted open shell ones. Moreover we have found that the use of localized orbitals is highly beneficial for the SCF convergence of larger species, making the acceleration scheme less critical compared to canonical orbitals.

1 Introduction

Atom-centered Gaussians have traditionally been the most common and widespread choice of basis set for molecules[1]. Several strong arguments are in favor of such a choice: the compactness of the representation which is defined by a handful of coefficients, the ability to represent atomic orbitals well (Slater functions are in theory superior due to the cusp at the nuclear position and the correct asymptotic), the simplification in the computation of molecular integrals which are often obtained analytically (this is the weak point of Slater orbitals which require expensive

numerical evaluations). Their main disadvantage is the non-orthogonality of the basis which can become a severe problem especially for large bases leading to a computational bottleneck when orthonormalization is required or worse numerical instabilities due to near linear-dependency in the basis[2].

On the opposite side of the spectrum, plane waves are ideally suited for periodic systems and are orthonormal by construction. However a very large number of them needs to be employed in order to achieve good accuracy, especially if one is interested in high resolution in the nuclear-core regions[3]. A popular choice to circumvent the problem is to use pseudopotentials[4] in the core region, thereby reducing the number of electrons to be treated and at the same time removing the need for very high-frequency components. Another challenge is constituted by non-periodic systems, which can only be dealt with by using a supercell approach[5].

Quantum chemical modeling is constantly expanding its horizons: cutting edge research is focused on achieving good accuracy (either in energetics or molecular properties) on large non-periodic systems such as large biomolecules or molecular nanosystems. This progression is constantly exposing the weaknesses of the traditional approach thus rendering the use of unconventional methods, which are free from the above mentioned limitations ever more attractive. One such choice is constituted by numerical, real-space grid-based methods which are gaining popularity in quantum chemistry as a promising strategy to deal with the Self Consistent Field (SCF) problem of Hartree-Fock and Density Functional Theory.

Among real-space approaches three strategies have been commonly employed: Finite Differences, Finite Elements and Wavelets/Multiwavelets[6]. Among these methods, Multiwavelets are particularly well suited for all-electron calculations[7, 8]. The basis functions are localized (as Gaussian-type orbitals) yet orthonormal (as plane waves). One crucial property of Multiwavelets is the disjoint support (zero overlap) between basis functions in adjacent nodes[9], paving the way for adaptive refinement of the mesh, tailored to each given function. This is essential for an all-electron description where varying resolution is a prerequisite for efficiency. The price to pay is that, in order to provide a representation with a given number of vanishing moments, a corresponding number of basis functions must be employed. The most common choice of basis functions in the Multiwavelet framework is a generic orthonormal polynomial basis of order k, providing a second possibility to increase the resolution of the representation alongside the adaptive grid refinement[10]. Currently, the main drawbacks of this approach are a large memory footprint (a numerical representation of a molecular orbital is much larger in terms of number of coefficients), and a significant computational overhead[11, 12]. On the other hand, a localized orthonormal basis is an ideal match for modern massively-parallel architectures[13] and we are confident that it is only a matter of time before real-space grid methods in general and Multiwavelets in particular will become competitive with or even superior to traditional ones.

Adaptivity is an excellent strategy to achieve good accuracy, as expensive high-resolution functions on a fine grid are present locally only where necessary[14]. The challenge with such an approach is constituted by the consequent lack of a fixed basis set which instead is only present "on demand". This detail has a profound impact on the minimization strategies that can be adopted in order to solve SCF problems such as the Roothaan-Hall equations of Hartree-Fock (HF) or the Kohn-Sham (KS) equations of Density Functional Theory (DFT). In other words all such strategies which strictly depend on having a fixed basis, such as the most common atomic orbital based methods[15] are excluded. On the other hand, only the occupied molecular orbitals are strictly needed both in HF and DFT to describe the wavefunction/electronic density. Several strategies, where only the knowledge of the occupied molecular orbitals and the electronic density are needed to minimize the energy, can be employed. The most straightforward albeit not efficient choices are the Steepest Descent (SD) and the Conjugate Gradient (CG) methods [16]. A better approach is constituted by the computation of the lowest eigenpairs (eigenvalue/eigenfunction) of the Fock operator, such as in the Lanczos method[17] and the algorithm proposed by Davidson[18]. The drawback of such an approach is the focus on the canonical, delocalized orbitals, which prevent linear scaling of any algorithm.

Currently the most popular minimization scheme is constituted by the Direct Inversion of the Iterative Subspace (DIIS), originally proposed by Pulay[19] and later revised by Wood and Zunger[20] in their Residual Minimization Method. Other strategies have recently been employed such a preconditioned conjugate gradient method[21] and a Krylov method proposed by Harrison[22]. An analysis of DIIS and Krylov methods has been presented by Rohwedder and Schneider[23]. A general formalization of the problem has been laid out by Schneider *et al.* [24], who considered the HF and KS problems as a minimization problem subject to (1) orthonormality constraint and (2) invariance with respect to rotation among the occupied orbitals. By properly defining the admissible Grassmann manifold based on the two constraints above, they were able to make use of standard minimization techniques such as a preconditioned steepest descent which could then later be combined with any acceleration method such as the DIIS technique[19] or the Krylov method proposed by Harrison[22].

In the present work we have considered a preconditioned steepest descent approach, in combination with the Krylov acceleration and localized molecular orbitals. In Sec. 2 we briefly summarize the mathematical framework of Multiwavelets, whereas in Sec. 3 we introduce the formalism of the SCF method, which is then applied in Sec. 4 to develop the equations for the real-space SCF optimizer both with canonical and localized orbitals. In particular, in Sec. 4.1 we describe how the Fock matrix is computed circumventing the need to apply the kinetic energy operator at any stage and in Sec. 1 we present the details of our algorithm. In Sec. 5 we present some benchmark results for DFT, RHF and UHF, emphasizing the convergence of the algorithm, the accuracy of

the results and the scaling of the code. The final conclusions and future perspectives are presented in Sec. 6.

2 Functions and operators in the Multiwavelet framework

Multiwavelets are kind of wavelets where disjoint support between adjacent nodes is enforced[25], by allowing more than one function to be present in each interval. A common choice for the initial scaling basis is a set of Legendre polynomials as originally proposed by Alpert[10]. Successive refinements are obtained by dilation and translation:

$$\phi^1_{j,l}(x) = 2^{1/2}\phi_j(2x-l), \qquad l = 0, 1 \tag{1}$$

which yields the usual ladder of scaling spaces:

$$V^0_k \subset V^1_k \subset \ldots \subset V^n_k \subset \ldots \tag{2}$$

The wavelet functions $\psi^n_{j,l}$ constitute the basis spanning two consecutive scaling spaces:

$$W^n_k \oplus V^n_k = V^{n+1}_k \tag{3}$$

Representation of functions is done by projection either on the scaling space at the finest scale $n+1$ (reconstructed representation) or on V^0 and the full ladder of wavelet spaces from W^0 to W^n (compressed representation). The two representations are equivalent and can be interconverted through the so-called Multiwavelet *filters*.

If we indicate with P^n and Q^n the projectors onto the scaling and wavelet spaces at scale n, respectively, then, it follows from the definition of the wavelet space that $P^{n+1} = P^n + Q^n$.

$$f^{n+1} \stackrel{\text{def}}{=} P^{n+1}f = (P^n + Q^n)f = f^n + \mathrm{d}f^n \tag{4}$$

where $\mathrm{d}f^n \stackrel{\text{def}}{=} Q^n f$.

The application of operators using a Multiwavelet basis can be performed using either the Standard form or the Non-Standard form. We have chosen the Non-Standard form [26, 27] as it virtually decouples scales from each other rendering the implementation of adaptive algorithms much simpler.

As shown in the following section, for SCF algorithms within a preconditioned steepest descent framework the necessary operators are the Poisson operator for the electrostatic potential, the Coulomb electronic energy and the quantistic exact exchange and the bound-state Helmholtz operator for the SCF iteration. Their Green's kernel can be written as

$$H^\mu(r - r') = \frac{e^{-\mu\|r-r'\|}}{4\pi\|r-r'\|} \tag{5}$$

, where $\mu > 0$ yields the bound-state Helmholtz kernel, whereas $\mu = 0$ is the Poisson kernel. Their application is achieved by convolution of a function with the corresponding Green's kernel

$$g(r) = [Tf](r) = \int G(r - r')f(r')\,\mathrm{d}r' \tag{6}$$

once an approximate separated form in terms of Gaussian functions has been computed[28, 26, 29]:

$$G(r - r') \approx \sum_{i=1}^{M} a_i e^{-\alpha_i (r-r')^2} \tag{7}$$

.

The Non-Standard form of the operator T is built as a telescopic expansion of the finest scale projection $T^N = P^N T P^N$

$$T^N = T^0 + \sum_{n=0}^{N-1} (A^n + B^n + C^n) \tag{8}$$

where $A^n = Q^n T Q^n$, $B^n = Q^n T P^n$, $C^n = P^n T Q^n$. Thanks to the vanishing moments of the Multiwavelet basis, the matrix representations of A^n, B^n and C^n are sparse and diagonally dominant for the Poisson and bound-state Helmholtz kernels. Therefore all terms beyond a predetermined bandwidth can be omitted in the operator application, without a loss of accuracy. In particular, we have shown that the application of the Poisson operator for the calculation of the electrostatic potential scales linearly with the size of the system [30].

3 Self Consistent Field methods

In the SCF approximations of an N-electron molecular system, the $3N$-dimensional wavefunction is expressed in terms of N three-dimensional one-electron spinorbitals $\{\phi_i\}_{i=1}^N$, in the form of a Slater determinant $\Phi(\langle\phi_1 \dots \phi_N\rangle)$. The spinorbitals are a solution of the HF or KS equations, if they a minimize the HF or KS energy functional $\mathcal{I}^{\mathrm{HF/KS}}(\Phi)$. Since Slater determinants are invariant with respect to unitary transformations among the spinorbitals, this redundancy is traditionally exploited to write the problem in the so-called canonical form

$$\hat{F}|\phi_i\rangle = \epsilon_i|\phi_i\rangle \tag{9}$$

in which the Fock matrix is diagonal

$$F_{ij} = \langle\phi_i|\hat{F}|\phi_j\rangle = \delta_{ij}\epsilon_i \tag{10}$$

where $\hat{F}$ is the Fock operator. Solving these equations will give the natural, delocalized molecular orbitals that are eigenfunctions of the given Fock operator.

As the SCF energy is invariant among rotations in the occupied orbital space, the more general set of explicitly coupled equations can be written as:

$$\hat{F}|\phi_i\rangle = \left[\sum_j |\phi_j\rangle\langle\phi_j|\right] \hat{F}\,|\phi_i\rangle = \sum_j F_{ji}\,|\phi_j\rangle \tag{11}$$

This invariance can be exploited in several ways, e.g. to achieve orbital localization in space, thus leading to more compact real-space representations, faster convergence for larger systems and prospects of low-scaling algorithms.

3.1 Kohn-Sham equations

In Kohn-Sham DFT, the electron density is given in terms of the occupied orbitals, assuming a closed-shell system

$$\rho(r) = 2\sum_{i=1}^{N/2} |\phi_i(r)|^2 . \tag{12}$$

Minimizing the DFT energy functional with respect to orbital variations (under the constraints that the orbitals remain orthonormal and the density integrates to the number of electrons) is equivalent to finding the $N/2$ lowest energy eigenfunctions of the Kohn-Sham operator

$$\hat{F} = -\frac{1}{2}\nabla^2 + V_{eff}(r), \tag{13}$$

where the effective potential have three contributions $V_{eff} = V_{nuc} + V_{el} + V_{xc}$. The electron-nuclear attraction is described by the nuclear potential

$$V_{nuc}(r) = -\sum_I \frac{Z_I}{|r - R_I|}, \tag{14}$$

the electron-electron repulsion is given through the electronic potential

$$V_{el}(r) = \int \frac{\rho(r')}{|r - r'|}\,\mathrm{d}r', \tag{15}$$

while all non-classical effects are included in the exchange-correlation potential V_{xc}, which is commonly approximated as a scalar function that depends only upon the density (LDA) or upon the density and its gradient (GGA) and higher order derivatives (meta-GGA).

The non-canonical Kohn-Sham equations are thus given as

$$\Big[-\frac{1}{2}\nabla^2 + V_{nuc}(r) + V_{el}(r) + V_{xc}(r)\Big]|\phi_i\rangle = \sum_j F_{ji}|\phi_j\rangle \tag{16}$$

and as both the electronic and exchange-correlation potentials depend upon the orbitals through the electron density, we have a set of coupled, non-linear equations.

3.2 Hartree-Fock equations

In Hartree-Fock theory the multiplicative exchange-correlation potential of Kohn-Sham theory is replaced by the non-multiplicative exchange operator $\hat{K}$, and the Fock operator is given as

$$\hat{F} = -\frac{1}{2}\nabla^2 + V_{nuc}(r) + V_{el}(r) - \hat{K}, \tag{17}$$

where the nuclear and electronic potentials are the same as in Eqs. (14) and (15), respectively, for the Kohn-Sham case. The exchange operator is defined through its effect on an orbital and can be written (for a closed-shell system) as[7]

$$\hat{K}\phi_i(r) = \sum_{j=1}^{N/2} \phi_j(r) \int \frac{\phi_i(r')\phi_j(r')}{|r-r'|}\,\mathrm{d}r' \tag{18}$$

and we get the non-canonical Hartree-Fock equations

$$\Big[-\frac{1}{2}\nabla^2 + V_{nuc}(r) + V_{el}(r) - \hat{K}\Big]|\phi_i\rangle = \sum_j F_{ji}|\phi_j\rangle \tag{19}$$

which again are non-linear through the electronic potential and exchange operator.

3.3 Integral formulation

The similarity between the Hartree-Fock and Kohn-Sham equations allows for the same solution algorithms to be used in both cases, and to get a unified description of we introduce the potential operator $\hat{V}$, defined as

$$\hat{V}^{HF} = V_{nuc}(r) + V_{el}(r) - \hat{K} \tag{20}$$

$$\hat{V}^{KS} = V_{nuc}(r) + V_{el}(r) + V_{xc}(r) \tag{21}$$

in Hartree-Fock and Kohn-Sham theory, respectively. Following Kalos[31] and Harrison *et al.* [7], we can rewrite the non-canonical equations in terms of the Helmholtz integral operator of Eq. (5) in the following way

$$\Big[-\frac{1}{2}\nabla^2 + \hat{V}\Big]\phi_i(\boldsymbol{r}) = \sum_j F_{ji}\phi_j(\boldsymbol{r}) \tag{22}$$

$$\big[-\nabla^2 - 2\lambda\big]\phi_i(\boldsymbol{r}) = -2\Big[\hat{V}\phi_i(\boldsymbol{r}) + \sum_j \big(\lambda\delta_{ij} - F_{ji}\big)\phi_j(\boldsymbol{r})\Big]. \tag{23}$$

By imposing that $\mu = \sqrt{-2\lambda}$ and by making use of the Green's function definition $(-\nabla^2 - 2\lambda)H^\mu(r-r') = \delta(r-r')$:

$$\phi_i(r) = -2\int H^\mu(r-r')\Big[\hat{V}\phi_i(r') + \sum_j \big(\lambda\delta_{ij} - F_{ji}\big)\phi_j(r')\Big]\,\mathrm{d}r' \tag{24}$$

$$\phi_i = -2\hat{H}^\mu\Big[\hat{V}\phi_i + \sum_j \big(\lambda\delta_{ij} - F_{ji}\big)\phi_j\Big] \tag{25}$$

This general expression can be simplified. In the canonical orbital basis, the Fock matrix becomes diagonal $F_{ji} = \epsilon_i\delta_{ij}$ and the orbital equations become formally decoupled:

$$\phi_i = -2\hat{H}^\mu\Big[\hat{V}\phi_i + \big(\lambda - \epsilon_i\big)\phi_i\Big] \tag{26}$$

By choosing $\lambda = \epsilon_i$, each orbital equation is further simplified:

$$\phi_i = -2\hat{H}^{\mu_i}\Big[\hat{V}\phi_i\Big] \tag{27}$$

with $\mu_i = \sqrt{-2\epsilon_i}$. We underline that, as for the standard differential formulations, the equations are still implicitly coupled through the potential operator $\hat{V}$, which is defined in terms of the orbitals: the solution must therefore be sought self-consistently by iterative methods. Finally, it is important to notice that both the orbitals ϕ_i and their corresponding energy ϵ_i are unknowns in the equations, and must be determined simultaneously.

The choice of orbital basis (canonical or not) as well as the λ parameter that appears in the Helmholtz operator will affect the numerical stability and the rate of convergence of the iterative solution algorithms that are presented in the following section.

4 Iterative solution algorithm

For simple one-orbital systems, the integral equation (27) can easily be brought to convergence up to arbitrary accuracy by straightforward iteration

$$\tilde{\phi}^{n+1} = -2\hat{H}^{\mu^n}\left[\hat{V}\phi^n\right] \tag{28}$$

$$\phi^{n+1} = \frac{\tilde{\phi}^{n+1}}{\|\tilde{\phi}^{n+1}\|} \tag{29}$$

where the Helmholtz operator is reconstructed in each iteration using the latest approximation of the orbital energy $\mu^n = \sqrt{-2\epsilon^n}$. The energy update can be approximated to first order at each iteration[7]

$$\epsilon^{n+1} = \epsilon^n - \frac{\langle \hat{V}\phi^n | \phi^n - \tilde{\phi}^{n+1} \rangle}{\|\tilde{\phi}^{n+1}\|^2} \tag{30}$$

This energy update is calculated very efficiently since the product $\hat{V}\phi^n$ is a byproduct of the iteration in Eq. (28), and inner products are efficiently evaluated in the orthonormal Multiwavelet basis.

For many-electron systems it is necessary to enforce orthonormality between the occupied orbitals in order to arrive at a true Aufbau solution of the HF/KS equations, as a straightforward iteration of Eq. (27) would bring all orbitals to the lowest eigenfunction.

The simplest approach to keep orthonormality would be to apply Eq. (27) for each orbital followed by a Gram-Schmidt orthogonalization in order of increasing energy. This would however lead to very slow convergence, especially for valence orbitals, as the convergence of each orbital is restrained by the level of convergence of lower-lying orbitals, whose convergence is again dependent on the accuracy of *all* orbitals through the potential energy operator $\hat{V}$.

Harrison *et al.* [7] described how to use deflation to extract multiple eigenpairs from the Fock operator by recasting the equation for each orbital as a ground state problem. Another approach, which was also suggested by Harrison *et al.* , is to diagonalize the Fock matrix at each iteration. By comparing a simple iteration of Eq. (28) with the general expression in Eq. (25), the off-diagonal

coupling terms are missing unless the Fock matrix is diagonal, and neglecting these will severely hamper the convergence for systems with many orbitals.

To keep a satisfactory convergence one can either include the off-diagonal elements and iterate Eq. (25) instead (using $\lambda_i^n = F_{ii}^n$)

$$\tilde{\phi}_i^{n+1} = -2\hat{H}^{\mu_i^n}\Big[\hat{V}^n\phi_i^n - \sum_{j\neq i} F_{ji}^n\phi_j^n\Big], \tag{31}$$

or one can diagonalize the Fock matrix in each iteration. In the latter case Eq. (28) becomes exact. In the following we will describe both approaches, where specifically we have the choice in each iteration to either diagonalize the Fock matrix [1] and obtain the canonical molecular orbitals (CMO), or to obtain localized molecular orbitals (LMO), in which case the Fock matrix is not diagonal. In either case we need to calculate the Fock matrix, and in the following we show how this can be done with no approximations without applying the kinetic energy operator.

4.1 Calculation of Fock matrix

The starting point is a set of orthonormal orbitals $\{\phi_i^n\}$ and an initial guess for the corresponding Fock matrix $F_{ij}^n \approx \langle\phi_i^n|\hat{F}|\phi_j^n\rangle$. We emphasize that such a guess need not to be the exact Fock matrix for the given orbital set. The new and now exact Fock matrix $\tilde{F}_{ij}^{n+1} = \langle\tilde{\phi}_i^{n+1}|\hat{F}|\tilde{\phi}_j^{n+1}\rangle$ in the non-orthonormal basis obtained by applying Eq. (31) (λ can formally be chosen arbitrarily) is computed without any reference to the kinetic energy operator.

As will be shown shortly, for the definition of the new Fock matrix to be consistent, the potential operator needs to be constructed *after* orthonormalization of the new orbitals. This requires a temporary set of orbitals $\{\bar{\phi}_i\}$, constructed e.g. through a Gram-Schmidt process, so that $\langle\bar{\phi}_i|\bar{\phi}_j\rangle = \delta_{ij}$. In this basis we calculate the potentials V_{el} and V_{xc} (in the case of DFT) through the density

$$\rho^{n+1}(r) = 2\sum_i^{N/2} |\bar{\phi}_i^{n+1}(r)|^2, \tag{32}$$

and exchange (in the case of Hartree-Fock) as

$$\hat{K}^{n+1}\tilde{\phi}_i^{n+1} = \sum_j^{N/2} \bar{\phi}_j^{n+1}(r) \int \frac{\tilde{\phi}_i^{n+1}(r')\bar{\phi}_j^{n+1}(r')}{|r-r'|}\,\mathrm{d}r'. \tag{33}$$

In order to avoid the application of the kinetic energy operator we exploit the formal definition of the bound-state Helmholtz operator as the inverse of the level-shifted Laplacian:

$$\left(\hat{T} - \lambda_i^n\right)^{-1} = 2\hat{H}^{\mu_i^n}. \tag{34}$$

[1] Since we are working only with occupied orbitals, the matrix diagonalization is considerably cheaper than the corresponding diagonalization using high-quality Gaussian basis sets, where the number of virtual orbitals can be much larger than the number of occupied orbitals.

The application of $(\hat{T} - \lambda_i^n)$ to the new orbital will return the argument from the Helmholtz operator (provided that $\mu_i^n = \sqrt{-2\lambda_i^n}$ was used in this operator):

$$\big(\hat{T} - \lambda_i^n\big)\tilde{\phi}_i^{n+1} = -\Big[\hat{V}\phi_i^n + \sum_j \big(\Lambda_{ji}^n - F_{ji}^n\big)\phi_j^n\Big], \tag{35}$$

where we have defined the diagonal matrix $\Lambda_{ji}^n = \lambda_i^n \delta_{ji}$. Before we proceed we also define the updates in the orbitals and in the potential

$$\Delta\tilde{\phi}_i^n = \tilde{\phi}_i^{n+1} - \phi_i^n, \tag{36}$$

$$\Delta\hat{V}^n = \hat{V}^{n+1} - \hat{V}^n. \tag{37}$$

We can now use the above observations to eliminate the kinetic operator from the calculation of the Fock matrix.

$$\begin{aligned}
\tilde{F}_{ij}^{n+1} &= \langle\tilde{\phi}_i^{n+1}|\hat{T} + \hat{V}^{n+1}|\tilde{\phi}_j^{n+1}\rangle \\
&= \langle\tilde{\phi}_i^{n+1}|\hat{T} - \lambda_j^n|\tilde{\phi}_j^{n+1}\rangle + \langle\tilde{\phi}_i^{n+1}|\hat{V}^{n+1} + \lambda_j^n|\tilde{\phi}_j^{n+1}\rangle \\
&= \langle\tilde{\phi}_i^{n+1}| - \Big[\hat{V}^n\phi_j^n + \sum_k \big(\Lambda_{kj}^n - F_{kj}^n\big)\phi_k^n\Big]\rangle + \langle\tilde{\phi}_i^{n+1}|\hat{V}^{n+1} + \lambda_j^n|\tilde{\phi}_j^{n+1}\rangle \\
&= -\langle\tilde{\phi}_i^{n+1}|\hat{V}^n|\phi_j^n\rangle - \sum_k \langle\tilde{\phi}_i^{n+1}|\phi_k^n\rangle\big(\Lambda_{kj}^n - F_{kj}^n\big) + \langle\tilde{\phi}_i^{n+1}|\hat{V}^{n+1}|\tilde{\phi}_j^{n+1}\rangle + \langle\tilde{\phi}_i^{n+1}|\tilde{\phi}_j^{n+1}\rangle\lambda_j^n \\
&= \langle\tilde{\phi}_i^{n+1}|\Delta\hat{V}^n|\phi_j^n\rangle + \langle\tilde{\phi}_i^{n+1}|\hat{V}^n|\Delta\tilde{\phi}_j^n\rangle + \sum_k \langle\tilde{\phi}_i^{n+1}|\phi_k^n\rangle\big(F_{kj}^n - \Lambda_{kj}^n\big) + \langle\tilde{\phi}_i^{n+1}|\tilde{\phi}_j^{n+1}\rangle\lambda_j^n \\
&= \langle\tilde{\phi}_i^{n+1}|\Delta\hat{V}^n|\phi_j^n\rangle + \langle\tilde{\phi}_i^{n+1}|\hat{V}^n|\Delta\tilde{\phi}_j^n\rangle + \sum_k \langle\phi_i^n + \Delta\tilde{\phi}_i^n|\phi_k^n\rangle F_{kj}^n + \langle\tilde{\phi}_i^{n+1}|\Delta\tilde{\phi}_j^n\rangle\lambda_j^n \\
&= \langle\tilde{\phi}_i^{n+1}|\Delta\hat{V}^n|\phi_j^n\rangle + \langle\tilde{\phi}_i^{n+1}|\hat{V}^n|\Delta\tilde{\phi}_j^n\rangle + \big(S^n F^n\big)_{ij} + \big(\Delta\tilde{S}_1^n F^n\big)_{ij} + \big(\Delta\tilde{S}_2^n \Lambda^n\big)_{ij}
\end{aligned}$$

where S^n is the overlap matrix (assumed identity). Finally, we see that we can calculate the new Fock matrix with no approximations by evaluating three update terms

$$\tilde{F}^{n+1} = F^n + \Delta\tilde{S}_1^n F^n + \Delta\tilde{S}_2^n \Lambda^n + \Delta\tilde{F}_{pot}^n \tag{38}$$

where we have defined two updates involving the overlap matrix

$$\big(\Delta\tilde{S}_1^n\big)_{ij} = \langle\Delta\tilde{\phi}_i^n|\phi_j^n\rangle \tag{39}$$

$$\big(\Delta\tilde{S}_2^n\big)_{ij} = \langle\tilde{\phi}_i^{n+1}|\Delta\tilde{\phi}_j^n\rangle \tag{40}$$

and one update involving the potential operator

$$\big(\tilde{F}_{pot}^n\big)_{ij} = \langle\tilde{\phi}_i^{n+1}|\hat{V}^n|\Delta\tilde{\phi}_j^n\rangle + \langle\tilde{\phi}_i^{n+1}|\Delta\hat{V}^n|\phi_j^n\rangle \tag{41}$$

The above expression is exact, but can be approximated and simplified while keeping track of the approximation order (e.g. the two overlap updates are equal to first order). As already mentioned,

the new potential should ideally be evaluated in a temporary orthonormal set of orbitals, but to avoid a costly intermediate orthogonalization one can choose to only normalize the orbitals and calculate the new potential operator in a non-orthogonal basis. While this is usually a fair approximation, we have observed a significant loss of convergence for certain problematic cases (in particular for some unrestricted open-shell Hartree-Fock calculations presented in Sec. 5).

We want to emphasize that these expressions *require* that the orbitals of the new set $\{\tilde{\phi}_i^{n+1}\}$ are related to the orbitals of the old set $\{\phi_i^n\}$ exactly through the application of the Helmholtz operator in Eq. (31). Otherwise the application of the kinetic energy operator cannot be avoided to obtain the Fock matrix.

4.2 Orbital orthogonalization

As already mentioned, the basic iteration of the integral operators in Eq. (31) to all orbitals $\{\phi_i\}$ does not preserve the orthonormality of the orbitals. This could be achieved by making use of an exponential parameterization[1, 24], but we have not pursued this strategy yet. At each iteration, orthonormalization needs to be restored explicitly. This is done in combination with a unitary transformation that either brings the equations to canonical form, or that localizes the orbitals in space.

Canonical orbitals

The orbital orthonormalization and Fock matrix diagonalization can be collected into a single orbital transformation. With the usual definition of the overlap matrix

$$\tilde{S}_{ij} = \langle \tilde{\phi}_i | \tilde{\phi}_j \rangle \tag{42}$$

the orbitals can be orthonormalized through the transformation

$$\bar{\phi}_i = \sum_{j=1}^{N} \tilde{S}_{ij}^{-1/2} \tilde{\phi}_j, \qquad \langle \bar{\phi}_i | \bar{\phi}_j \rangle = \delta_{ij}. \tag{43}$$

The Fock matrix in the non-orthonormal basis reads: $\tilde{F}_{ij} = \langle \tilde{\phi}_i | T + \hat{V} | \tilde{\phi}_j \rangle$. Its transformation to the orthonormal basis is achieved as $\bar{F} = \tilde{S}^{-1/2} \tilde{F} \tilde{S}^{1/2}$. By calling M_X the unitary transformation to the requested basis ($X = C, L$ for canonical or localized basis respectively) the overall transformation matrix becomes $U = M^T \tilde{S}^{-1/2}$.

Localized orbitals

Instead of diagonalizing the Fock matrix, it is beneficial for bigger systems to work with localized molecular orbitals (LMOs), both in terms of the familiar prospects of low-scaling algorithms of

conventional SCF implementations[32], but also to reduce the large storage requirements of real-space methods.

Following Yanai *et al.* [8], we use the Foster-Boys[33, 34] algorithm for orbital localization, where the unitary matrix is calculated based on the one-electron dipole integrals $\langle\phi_i|\boldsymbol{r}|\phi_j\rangle$. More specifically, the dipole integrals are calculated in the non-orthogonal orbitals $\{\tilde{\phi}_i\}$ and then orthonormalized using the $\tilde{S}^{-1/2}$ matrix before the unitary matrix M_L that localize the orbitals is calculated using an iterative non-linear optimizer[35].

With the formulation of the HF/KS equations given in Eq. (31) it is possible to work exclusively in the LMO basis throughout the SCF optimization. This is in contrast to the algorithm proposed by Yanai *et al.* [8], which relies on CMOs for the application of the Helmholtz operator, and low-scaling calculation of HF exchange is obtained by first transforming the CMOs into LMOs, then apply the exchange operator, before transforming the result back to the CMO basis.

It is in general not necessary to localize the orbitals in every iteration, as the new orbitals will only be slightly perturbed from the old ones, and a simple orthonormalization using the $\tilde{S}^{-1/2}$ matrix will to a large extent keep the localization of the old orbitals. However, in this case the calculation of the M_L matrix is very efficient requiring only a handful of iterations, keeping the computational overhead of the extra localization low, compared to the actual orthonormalization and rotation of the orbitals, which is indeed rather inefficient in real-space bases (although it should become increasingly diagonally dominant as the SCF iteration proceeds).

4.3 Krylov subspace accelerated inexact Newton method

By collecting the orbital vector $\Phi = (\phi_0, \ldots, \phi_N)^T$ and the Fock matrix F into a new vector $\boldsymbol{x} = (\Phi, F)$, the SCF problem in Eq. (25) can be viewed as finding the roots of the following residual function

$$f(\boldsymbol{x}) = -2\hat{H}^{\mu}\big[\hat{V}\Phi + \big(\Lambda - F\big)\Phi\big] - \Phi. \tag{44}$$

At a given iteration n, we have the current approximation $\boldsymbol{x}^n = (\Phi^n, F^n)$ and the corresponding residual $f(\boldsymbol{x}^n) = (\Delta\Phi^n, \Delta F^n)$. In the Krylov subspace accelerated inexact Newton (KAIN)[22] method the new update $\delta\boldsymbol{x}^n$ is calculated in terms of the m latest iterations

$$\delta\boldsymbol{x}^n = f(\boldsymbol{x}^n) + \sum_{j=(n-m)}^{n-1} c_j\big[\,\big(\boldsymbol{x}^j - \boldsymbol{x}^n\big) + \big(f(\boldsymbol{x}^j) - f(\boldsymbol{x}^n)\big)\,\big], \tag{45}$$

where the coefficients c_j are obtained by solving the linear system $Ac = b$

$$A_{ij} = \langle \boldsymbol{x}^n - \boldsymbol{x}^i | f(\boldsymbol{x}^n) - f(\boldsymbol{x}^j)\rangle, \tag{46}$$

$$b_i = \langle \boldsymbol{x}^n - \boldsymbol{x}^i | f(\boldsymbol{x}^n)\rangle. \tag{47}$$

The Frobenius inner product is employed for the Fock matrix. The size m of the Krylov subspace is without constraints.

We conclude this section with a few important aspects to consider with the KAIN accelerator with more than one orbital:

1. As the orbital energies are updated in each iteration, it might occur that orbitals switch place during the optimization process, if ordered by increasing energy. This means that one must make sure that each orbital is linked to the correct orbital in the iterative history.

2. In the case of degeneracies, the matrix diagonalization will not result in a unique set of orbitals, as any basis for the degenerate subspace can be obtained. It is thus important to keep a consistent definition of all orbitals as the iteration proceeds. This and the previous issue can be treated simultaneously by always applying the same orbital rotation M_D that diagonalized the Fock matrix to the entire iterative history.

3. With localization there is a consistency in the definition of the orbital history unless there is an infinite rotational symmetry axis in the molecule, which makes the localization minimum non-unique. Otherwise the algorithm will always fall back to the same minimum, as long as the Helmholtz step does not bring the orbitals too far from equilibrium.

4. The KAIN updates do not conserve the orthogonality between the orbitals, and a second orthogonalization process can be added at the end of the cycle, although this is not strictly necessary to achieve convergence.

4.4 Algorithm

Finally, we put all the pieces together to a general algorithm for the SCF optimization for many-electron systems. Starting from an arbitrary initial guess for the orbitals and the Fock matrix, we calculate the electron density and corresponding potentials. The Helmholtz operator is applied once to each orbital, and we judge the convergence by the norm of the orbital update at this point. The potential update is calculated using an intermediate orthonormalized orbital set, and the new Fock matrix is calculated as described in Sec. 4.1. The appropriate transformation matrix U for Fock matrix diagonalization or orbital localization is obtained and applied to the orbitals and Fock matrix. The orbital and Fock matrix residuals are calculated *after* the orbital orthonormalization and rotation described in Sec. 4.2

$$\Delta\Phi^n = \Phi^{n+1} - \Phi^n = U\tilde{\Phi}^{n+1} - \Phi^n, \tag{48}$$

$$\Delta F^n = F^{n+1} - F^n = U\tilde{F}^{n+1}U^{-1} - F^n. \tag{49}$$

and then added to the KAIN history

$$\boldsymbol{x}^n = (\Phi^n, F^n) \qquad f(\boldsymbol{x}^n) = (\Delta\Phi^n, \Delta F^n)\,. \tag{50}$$

If the length of the history exceeds some modest number the oldest vector is discarded. New updates are then calculated based on Eq. (45)

$$\delta \boldsymbol{x}^n = (\delta\Phi^n, \delta F^n) \tag{51}$$

and subsequently added to the previous guess. The orbitals are then (optionally) orthonormalized. This sequence is iterated until the maximum norm among the orbital updates (after the Helmholtz operator application) is below some predefined threshold. The whole procedure is summarized in Algorithm 1.

Algorithm 1 Iterative SCF optimization of many-electron systems.

Given initial orbitals ϕ_i^0 and Fock matrix F^0
while $\varepsilon > \varepsilon_r$ **do**
  Calculate electron density through Eq.(12)
  Calculate potential operator $\hat{V}^n$ through Eq.(20) or Eq. (21)
  for each orbital i **do**
    if F_{ii}^n differs significantly from λ_i^{n-1} **then**
      Set $\lambda_i^n = F_{ii}^n$ and reconstruct Helmholtz operator
    else
      Set $\lambda_i^n = \lambda_i^{n-1}$ and reuse Helmholtz operator
    end if
    Calculate Helmholtz argument $\varphi_i^n = \hat{V}^n\phi_i^n + \sum_j \left(\lambda_i^n\delta_{ji} - F_{ji}^n\right)\phi_j^n$
    Apply Helmholtz operator $\tilde{\phi}_i^{n+1} = -2\hat{H}^{\mu_i^n}\left[\varphi_i^n\right]$
    Calculate orbital update $\Delta\tilde{\phi}_i^n = \tilde{\phi}_i^{n+1} - \phi_i^n$
  end for
  Determine maximum orbital error $\varepsilon = \max_i \|\Delta\tilde{\phi}_i^n\|$
  Calculate potential operator update $\Delta\hat{V}^n = \hat{V}^{n+1} - \hat{V}^n$
  Calculate Fock matrix through Eq. (38)
  Calculate rotation matrix U to diagonalize or localize
  Calculate KAIN residual $f(\boldsymbol{x}^n)$ through Eqs. (48) and (49)
  Calculate KAIN updates $\delta\boldsymbol{x}^n$ through Eq. (45)
  Update orbitals $\phi_i^{n+1} = \phi_i^n + \delta\phi_i^n$
  Update Fock matrix $F^{n+1} = F^n + \delta F^n$
  Orthonormalize orbitals
end while

5 Results

The performance of the code has been tested focusing in particular on convergence rate and accuracy for small- and medium-sized molecular systems. Two accuracy parameters ε_r and ε_{tot} have been considered: the former denotes the convergence threshold in the iterative algorithm, and the latter is the overall accuracy of the underlying function representations and mathematical operations.

All DFT calculations are performed using the standard Local Density Approximation (LDA), consisting of the Slater-Dirac[36, 37] exchange and the VWN5[38] correlation potentials. The exchange-correlation potentials are provided by the XCFun library[39], and we use the same smoothed nuclear potential as described by Yanai *et al.* [7].

For all calculations, the initial guess is obtained from conventional Gaussian basis calculations using the LSDalton[32, 40] program with a 3-21G[41, 42] basis set. All calculations have been performed on a single compute node containing 2×8 cores Intel Xeon E5-2670 processors with 16 GB memory (128 GB for the most demanding calculations).

5.1 Convergence

The convergence rate of the algorithms presented in Sec. 4 for one- and many-electron systems has been investigated with respect to three parameters: the size of the iterative subspace in the KAIN method, the choice of orbital transformation (canonical vs. localized), and the choice of λ in Eq. (31).

5.1.1 Hydrogen atom

The Hydrogen atom is a simple one-electron system, where the Schrödinger equation can be solved by straightforward power iteration of Eqs. (28)-(30). The potential operator $\hat{V}$ contains only the nuclear potential and does not depend on the wavefunction.

As can be seen from Fig.1, the convergence of the wavefunction and the corresponding energy of the power iteration is remarkably uniform: the norm of the wavefunction update is almost exactly halved between each iteration, while the energy update is divided by 4, showing that the error in the energy is quadratic in the error of the wavefunction. The overall accuracy is kept to $\varepsilon_{tot} = 10^{-10}$ throughout the iterations.

The convergence of the Hydrogen atom is significantly improved when Krylov subspace acceleration ($m = 4$) is included, with a threefold reduction in the number of iterations required to achieve a given accuracy in the wavefunction. The error in the energy is however no longer quadratic with respect to the wavefunction error, since both the wavefunction and the energy are included in the KAIN subspace vector, leading to a comparable accuracy for both.

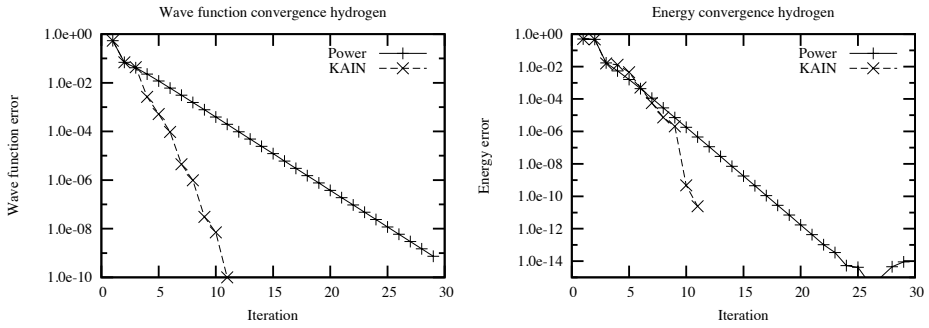

Figure 1: Convergence (relative error) of the Hydrogen wavefunction and total energy for a simple power iteration and using Krylov subspace acceleration with $m = 4$. The overall accuracy is kept at $\varepsilon_{tot} = 10^{-10}$.

5.1.2 Many-electron systems

Using canonical orbitals, we iterate the separated orbital equations

$$\tilde{\phi}_i^{n+1} = -2\hat{H}^{\mu_i^n}\left[\hat{V}^n\phi_i^n\right]. \tag{52}$$

provided that the Fock matrix is diagonalized in each iteration and the Helmholtz operator is updated using the latest orbital energy $\lambda_i^n = \epsilon_i^n$. This corresponds to the algorithm proposed by Harrison *et al.* [7].

The convergence of the canonical orbitals and the total energy of methane is presented in Fig. 2, where the overall accuracy is kept at $\varepsilon_{tot} = 10^{-8}$ throughout the iterations. Again we observe linear convergence, albeit somewhat slower than the Hydrogen atom: the error is each time roughly reduced by a factor of 1.6 instead of 2. The total energy shows similar convergence, and the error in the energy lies consistently below the corresponding error in the orbitals by about two orders of magnitude. With the KAIN method ($m = 4$) a significant improvement in the orbital convergence is achieved and the error in the energy is consistent with the error in the orbitals.

5.1.3 Convergence acceleration and orbital localization

The size of the iterative subspace m used in the KAIN algorithm affects the convergence for many-electron systems. In Tab. 1 the number of iterations required to reach convergence ($\varepsilon_r \leq 10^{-4}$ for all orbitals) for some of the smallest linear alkanes (C_nH_{2n+2}, $n = 1, 24, 6, 8, 10$) is reported as a function of m ($m = 0$ reduces to the power method) at the LDA level of theory. The first series of data refers to calculations performed with canonical orbitals and the second one to localized orbitals as described in Sec. 4.2.

For canonical orbitals, the straightforward iteration of Eq. (52) followed by a Fock matrix diagonalization displays poor convergence as the size of the system increases. However, a modest

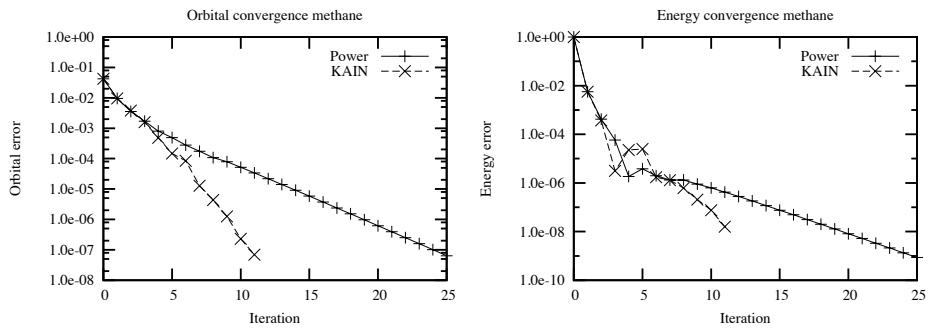

Figure 2: Convergence (relative error) of the maximum orbital residual and total energy of the methane molecule by power iteration of the canonical orbitals, and using Krylov subspace acceleration with $m = 4$. The overall accuracy is kept at $\varepsilon_{tot} = 10^{-8}$.

number ($m = 2 - 4$) of the iterative history improves convergence significantly. A larger iterative subspace $m > 4$ has practically no effect on the convergence.

With localized orbitals, the more general Eq. (31) is employed, followed by an orbital localization. The final energies are observed to be consistent with the canonical energies within the truncation threshold. Without acceleration ($m = 0$) the smaller molecules display a rate of convergence similar to the canonical orbitals. However, the number of iterations does not increase significantly for larger systems. When acceleration is included we observe an improvement until $m = 2$. Overall, using localized orbitals with $m = 2$ is slightly better than using canonical orbitals with $m = 4$.

In conclusion, for atoms and small molecules using canonical or localized orbitals is almost equivalent when KAIN is switched on: canonical orbitals converge in fewer iterations but require a larger subspace ($m = 4$) whereas localized orbitals require 2-3 additional iterations but a smaller subspace $m = 2$. For larger systems, orbital localization yields faster convergence even with a smaller subspace $m = 2$.

5.1.4 Fixed λ

As Eq. (52) can be viewed as a preconditioned steepest descent iteration, it is interesting to investigate how the convergence is affected by the choice of the parameter λ entering the preconditioner $\hat{H}^{\mu}$.

We have studied the convergence when a *fixed* λ value is employed in the preconditioner throughout the iterations. The canonical orbital equations then take the form:

$$\tilde{\phi}_i^{n+1} = -2\hat{H}^{\mu_i^n}\Big[\hat{V}^n\phi_i^n + \big(\lambda - \epsilon_i^n\big)\phi_i^n\Big]. \tag{53}$$

For the Hydrogen atom the exact energy is known ($E = 1/2$ Hartrees). This value will then be used as a reference, expressing λ relative to it. The number of iterations necessary to converge

Table 1: Number of iterations required to bring the orbital residual to $\epsilon_r \leq 10^{-4}$ for different lengths of KAIN history ($m = 0$ corresponding to regular power iteration) using canonical and localized orbitals.

		Size m of KAIN history							
Molecule	N orbitals	0	1	2	3	4	5	6	7
		Canonical orbitals							
C_1H_4	5	8	8	7	6	6	6	6	6
C_2H_6	9	9	8	7	7	7	7	7	7
C_4H_{10}	17	17	15	10	9	10	10	10	10
C_6H_{14}	25	36	23	13	12	11	11	11	11
		Localized orbitals							
C_1H_4	5	9	9	7	7	7	7	7	7
C_2H_6	9	10	10	8	8	8	8	8	8
C_4H_{10}	17	13	12	9	9	9	9	9	9
C_6H_{14}	25	14	11	9	10	9	9	9	9
C_8H_{18}	33	11	10	8	9	8	8	8	8
$C_{10}H_{22}$	41	11	11	9	9	8	8	8	8

Eq. (53) up to a residual norm of $\varepsilon_r \leq 10^{-7}$ without any subspace acceleration, is presented in Fig. 3 as a function of λ/E. We underline that $\lambda = E$ does not mean that the last term in Eq. (53) vanishes in all iterations; however this term approaches zero as the eigenvalue ϵ_0^n approaches the true energy. The value corresponding to the "dynamic" algorithm (the Helmholtz operator is updated in each iteration as in Eq. (52)) is denoted in the figure by a dashed line.

Surprisingly, the fastest convergence is achieved by choosing λ slightly smaller (in absolute value) than the true energy. The dynamic update is indeed equally fast as setting $\lambda = E$. The value of λ cannot however be chosen arbitrarily, as the rate of convergence deteriorates rapidly away from $\lambda = E$, ultimately leading to divergence. It is therefore important to choose λ close to the final energy, but updating the operator in each iteration can safely be avoided, especially close to convergence.

The same analysis has been repeated for methane. Since the exact orbital energies ϵ_i are not available, the converged results from a previous calculation have been employed: for each occupied orbital i λ_i is expressed relative to the converged ϵ_i. The fixed-λ equations (53) are then iterated to a maximum residual orbital norm of $\varepsilon_r \leq 10^{-5}$ without subspace acceleration, and the Fock matrix is diagonalized in each iteration. The number of iterations required with different choices of λ_i/ϵ_i is presented in figure 3.

In contrast to the Hydrogen atom, the best choice is now $\lambda_i/\epsilon_i = 1$ (exact eigenvalues). The dynamic strategy (dashed line) is in this case equally good. The conclusion is again that λ should always be chosen close to the current orbital energy, but not necessarily equal to ϵ_i^n in each

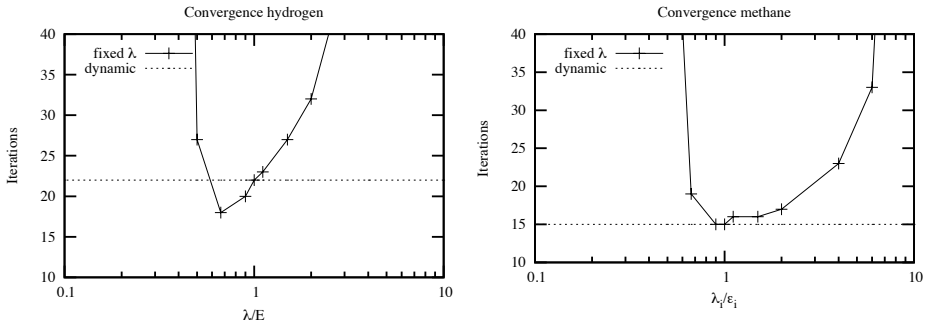

Figure 3: The effect of using a fixed λ in the power iteration. The number of iterations required to bring the residual below $\varepsilon_r \leq 10^{-7}$ for Hydrogen and $\varepsilon_r \leq 10^{-5}$ for methane for different choices of λ. The number of iterations when using a dynamic λ is given for comparison.

iteration.

Based on these observations the Helmholtz operator is recomputed with $\lambda_i = \epsilon_i^n$ whenever the new energy differs from the old λ_i by more than 1%. Our tests show that this choice is close to optimal and has minimal effect on the overall convergence: at most one extra iteration was required on a wide variety of molecules and to different choices of final precision. The significant advantage is that the Helmholtz operator associated with a given orbital is in practice fixed once the orbital is converged within $\varepsilon_r \leq 10^{-2}$. This strategy leads to considerable savings in computing time, especially for large systems and high accuracy.

5.2 Accuracy

As the approach followed is unconventional (real-space minimization with Multiwavelets) and presents also novel aspects compared to previous work (localized orbitals and full Fock matrix computation) it is important to investigate the overall accuracy that can be attained.

The accuracy in the computation of the total energy is related to the overall accuracy of the calculation ε_{tot}. It is however important to assess how accurate the orbital representation has to be in order to guarantee the demanded accuracy on the energy. Fig. 4 shows the convergence of the maximum orbital residual and the total LDA energy for small closed-shell atoms. The overall accuracy of the calculations was $\varepsilon_{tot} = 10^{-6}$, and the canonical orbital equations were iterated without Krylov acceleration in order to determine the maximum accuracy attainable. The total energies were compared to the basis set limit taken from the National Institute of Standards and Technology (NIST)[43].

The error in the total energy is brought below the threshold within ten iterations for all atoms, and it is afterwards stabilized at about an order of magnitude below ε_{tot}. A similar behavior is

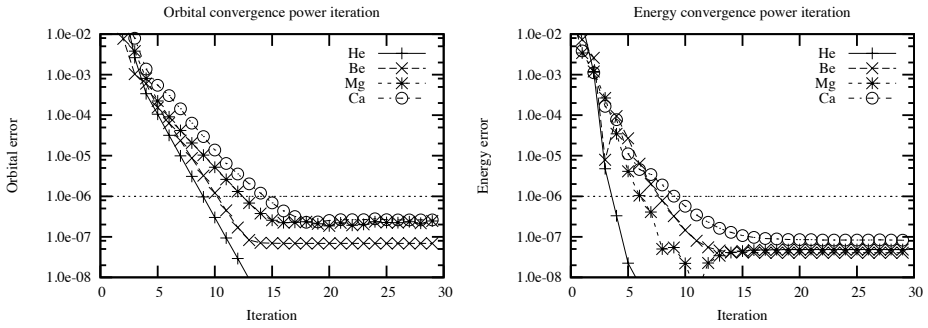

Figure 4: Orbital and energy convergence for power iteration without subspace acceleration. Overall accuracy is kept at $\varepsilon_{tot} = 10^{-6}$, and the equations are iterated to maximum accuracy. The error in the total energy is taken with respect to the NIST standard as a reference, whereas the error in the orbital is the maximum orbital residual defined as the norm of the largest orbital update at each iteration.

observed for the orbital error. However, at any given iteration the error in the energy is at least an order of magnitude below the error in the orbitals. Therefore, in order to achieve an accuracy ε in the total energy, it is sufficient to set $\varepsilon_{tot} = \varepsilon$, and $\varepsilon_r \leq 10\varepsilon$.

The corresponding results using Krylov subspace acceleration are reported in Fig. 5. Although the convergence is less regular, the same trend is observed as for the straightforward power iteration. The accuracy obtained is the same, but it is reached with 30-40% fewer iterations. The relationship between the error in the orbitals and in the total energy remains at one order of magnitude.

A further confirmation of the correctness of this choice can be found by inspecting the results obtained for the electronic structure of selected atoms (see Tab. 2 for closed shell systems and Tab. 3 for open shell ones) using the Local Spin-Density Approximation (LSDA). All calculations were iterated to a maximum orbital residual of $\varepsilon_r \leq 10\varepsilon_{tot}$.

In Tables 4 to 6 results for atoms and small molecules at the HF level are reported. The test molecules are taken from the HEAT[44, 45, 46] project which is aiming at high-accuracy *ab initio* thermochemistry. The molecular geometries as well as the presented energies for estimated Hartree-Fock limit and using high-quality Gaussian basis sets can be found in the references above. Robust and rapid convergence is achieved for all investigated systems with accuacy comparable to the LDA calculations. An important point to achieve fast and accurate convergence is the calculation of the electronic density in Eq. (32) and the Hartree-Fock exchange in Eq. (33): in an attempt to improve performance the orbitals used therein were only normalized (orthogonalization is naturally achieved close to convergence) but such an approximation affected convergence severely, especially for some open-shell systems (e.g. NO, CH, CCH, HO_2).

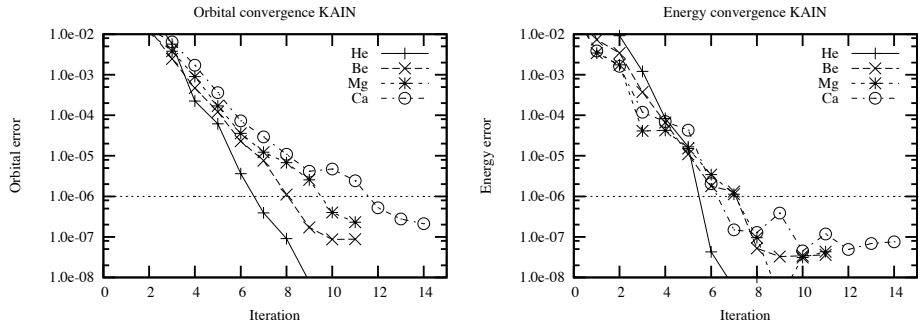

Figure 5: Orbital and energy convergence for power iteration including subspace acceleration $m = 4$. Overall accuracy is kept at $\varepsilon_{tot} = 10^{-6}$, and the equations are iterated to maximum accuracy. The error in the total energy is taken with respect to the NIST standard as a reference, whereas the error in the orbital is the maximum orbital residual defined as the norm of the largest orbital update at each iteration.

Table 2: LDA energies for closed-shell atoms. NIST represents basis set limit[43]. Orbitals are converged to a maximum residual of $\varepsilon_r \leq 10\varepsilon_{tot}$.

Requested	Energy (Hartree)				
precision	E_{kin}	E_{ee}	E_{en}	E_{xc}	E_{tot}
			Helium		
$\varepsilon_{tot} \leq 10^{-3}$	2.77205228	1.99665754	-6.63046074	-0.97373563	-2.83548654
$\varepsilon_{tot} \leq 10^{-5}$	2.76793649	1.99612381	-6.62558036	-0.97331623	-2.83483629
$\varepsilon_{tot} \leq 10^{-7}$	2.76792245	1.99611976	-6.62556386	-0.97331398	-2.83483563
NIST	2.767922	1.996120	-6.625564	-0.973314	-2.834836
			Beryllium		
$\varepsilon_{tot} \leq 10^{-3}$	14.32209505	7.11589404	-33.37086518	-2.51552236	-14.44839844
$\varepsilon_{tot} \leq 10^{-5}$	14.30973819	7.11528422	-33.35734165	-2.51487298	-14.44719222
$\varepsilon_{tot} \leq 10^{-7}$	14.30942392	7.11525737	-33.35703462	-2.51485614	-14.44720948
NIST	14.309424	7.115257	-33.357034	-2.514856	-14.447209
			Magnesium		
$\varepsilon_{tot} \leq 10^{-3}$	197.83334926	95.71507014	-477.14806651	-15.44282502	-199.04247213
$\varepsilon_{tot} \leq 10^{-5}$	198.53996893	95.67281328	-477.89751361	-15.45500303	-199.13973443
$\varepsilon_{tot} \leq 10^{-7}$	198.54150414	95.67328823	-477.89914697	-15.45505141	-199.13940601
NIST	198.541505	95.673290	-477.899149	-15.455051	-199.139406
			Calcium		
$\varepsilon_{tot} \leq 10^{-3}$	674.19516760	285.15968984	-1600.92618633	-34.13913904	-675.71046793
$\varepsilon_{tot} \leq 10^{-5}$	674.66137975	285.20533903	-1601.46680475	-34.14256738	-675.74265335
$\varepsilon_{tot} \leq 10^{-7}$	674.65729500	285.20613876	-1601.46319769	-34.14253866	-675.74230258
NIST	674.657334	285.206130	-1601.463209	-34.142538	-675.742283

Table 3: LSDA energies for open-shell atoms. NIST represents basis set limit[43]. Orbitals are converged to a maximum residual of $\varepsilon_r \leq 10\varepsilon_{tot}$.

Requested	Energy (Hartree)				
precision	E_{kin}	E_{ee}	E_{en}	E_{xc}	E_{tot}
			Hydrogen		
$\varepsilon_{tot} \leq 10^{-3}$	0.46408449	0.29795528	-0.96306035	-0.27764807	-0.47866865
$\varepsilon_{tot} \leq 10^{-5}$	0.46664520	0.29837713	-0.96562129	-0.27807194	-0.47867090
$\varepsilon_{tot} \leq 10^{-7}$	0.46664311	0.29837681	-0.96561918	-0.27807151	-0.47867076
NIST	0.466643	0.298377	-0.965619	-0.278072	-0.478671
			Lithium		
$\varepsilon_{tot} \leq 10^{-3}$	7.25182279	4.00976028	-16.93974456	-1.66542991	-7.34359140
$\varepsilon_{tot} \leq 10^{-5}$	7.24980921	4.00923212	-16.93781276	-1.66517957	-7.34395099
$\varepsilon_{tot} \leq 10^{-7}$	7.24989420	4.00925979	-16.93792111	-1.66518948	-7.34395660
NIST	7.249892	4.009258	-16.937918	-1.665189	-7.343957
			Sodium		
$\varepsilon_{tot} \leq 10^{-3}$	160.14833493	79.70030511	-387.75498135	-13.51431838	-161.42065969
$\varepsilon_{tot} \leq 10^{-5}$	160.90590302	79.78423947	-388.60422402	-13.53369303	-161.44777456
$\varepsilon_{tot} \leq 10^{-7}$	160.90724741	79.78490710	-388.60601781	-13.53376279	-161.44762609
NIST	160.907255	79.784904	-388.606022	-13.533763	-161.447625
			Potassium		
$\varepsilon_{tot} \leq 10^{-3}$	597.01717215	257.40676800	-1421.00811117	-31.65933886	-598.24350988
$\varepsilon_{tot} \leq 10^{-5}$	597.17434653	257.43558611	-1421.15953204	-31.65796306	-598.20756246
$\varepsilon_{tot} \leq 10^{-7}$	597.17998931	257.43704851	-1421.16500554	-31.65806931	-598.20603703
NIST	597.179968	257.437000	-1421.164934	-31.658067	-598.206032

Table 4: RHF total energies for closed-shell molecules. Estimated HF limit, Gaussian basis energies and geometries are taken from Refs.[44, 45, 46].

Requested precision	Total energy (Hartree)			
	H_2O	H_2O_2	CO	CO_2
$\varepsilon_{tot} = 10^{-5}$	-76.067611455	-150.85253297	-112.79087294	-187.72538886
$\varepsilon_{tot} = 10^{-6}$	-76.067556696	-150.85249254	-112.79069389	-187.72541991
$\varepsilon_{tot} = 10^{-7}$	-76.067535613	-150.85246986	-112.79081263	-187.72538522
$\varepsilon_{tot} = 10^{-8}$	-76.067535431	-150.85247037	-112.79081269	-187.72538560
Est. HF limit	-76.0675	-150.8525	-112.7908	-187.7254
aug-cc-pCV5Z	-76.067379371	-150.85218780	-112.79063514	-187.72508317
aug-cc-pCVQZ	-76.066140457	-150.84985235	-112.78919290	-187.72260431
	H_2	C_2H_2	HF	F_2
$\varepsilon_{tot} = 10^{-5}$	-1.1336231555	-76.855426687	-100.07104518	-198.77361614
$\varepsilon_{tot} = 10^{-6}$	-1.1336185308	-76.855583643	-100.07090391	-198.77356693
$\varepsilon_{tot} = 10^{-7}$	-1.1336185408	-76.855589419	-100.07088606	-198.77354557
$\varepsilon_{tot} = 10^{-8}$	-1.1336185528	-76.855588322	-100.07088501	-198.77354427
Est. HF limit	-1.1336	-76.8556	-100.0709	-198.7735
aug-cc-pCV5Z	-1.1335996653	-76.855469571	-100.07066449	-198.77312544
aug-cc-pCVQZ	-1.1334622625	-76.854638249	-100.06873826	-198.76928779
	N_2	NH_3	HCN	HNO
$\varepsilon_{tot} = 10^{-5}$	-108.99285921	-56.225111832	-92.915931357	-129.84997375
$\varepsilon_{tot} = 10^{-6}$	-108.99312380	-56.225050995	-92.915781860	-129.84996270
$\varepsilon_{tot} = 10^{-7}$	-108.99309384	-56.225052223	-92.915802570	-129.84998798
$\varepsilon_{tot} = 10^{-8}$	-108.99309024	-56.225051873	-92.915801013	-129.84998863
Est. HF limit	-108.9931	-56.2250	-92.9158	-129.8500
aug-cc-pCV5Z	-108.99293047	-56.224938182	-92.915668251	-129.84977143
aug-cc-pCVQZ	-108.99167619	-56.224119853	-92.914667092	-129.84803908

Table 5: UHF total energies for open-shell molecules. Estimated HF limit, Gaussian basis energies and geometries are taken from Refs.[44, 45, 46].

Requested precision	Total energy (Hartree)			
	H	OH	HO_2	O_2
$\varepsilon_{tot} = 10^{-5}$	-0.5000005418	-75.428181578	-150.25211601	-149.69178275
$\varepsilon_{tot} = 10^{-6}$	-0.5000000246	-75.428114257	-150.25269142	-149.69155112
$\varepsilon_{tot} = 10^{-7}$	-0.5000000111	-75.428109595	-150.25267450	-149.69154144
$\varepsilon_{tot} = 10^{-8}$	-0.5000000001	-75.428109170	-150.25267364	-149.69154036
Est. HF limit	-0.5000	-75.4281	-150.2527	-149.6915
aug-cc-pCV5Z	-0.4999947846	-75.427965654	-150.25239658	-149.69126719
aug-cc-pCVQZ	-0.4999483215	-75.426782123	-150.25009561	-149.68900724
	CH	CH_2	CH_3	CCH
$\varepsilon_{tot} = 10^{-5}$	-38.284525066	-38.941000192	-39.581248206	-76.183662196
$\varepsilon_{tot} = 10^{-6}$	-38.284539890	-38.940974280	-39.581213883	-76.183555283
$\varepsilon_{tot} = 10^{-7}$	-38.284505692	-38.940972985	-39.581211226	-76.183559436
$\varepsilon_{tot} = 10^{-8}$	-38.284505511	-38.940972753	-39.581211023	-76.183559250
Est. HF limit	-38.2845	-38.9410	-39.5812	-76.1835
aug-cc-pCV5Z	-38.284449323	-38.940902183	-39.581129872	-76.183449583
aug-cc-pCVQZ	-38.284067801	-38.940408627	-39.580563543	-76.182657470
	NH	NH_2	NO	NO_2
$\varepsilon_{tot} = 10^{-5}$	-54.986447671	-55.592360153	-129.30950551	-204.13091462
$\varepsilon_{tot} = 10^{-6}$	-54.986416818	-55.592337083	-129.30951408	-204.13070528
$\varepsilon_{tot} = 10^{-7}$	-54.986414045	-55.592333120	-129.30953328	-204.13072391
$\varepsilon_{tot} = 10^{-8}$	-54.986413719	-55.592332706	-129.30953318	-204.13072160
Est. HF limit	-54.9864	-55.5923	-129.3095	
aug-cc-pCV5Z	-54.986325573	-55.592232278	-129.30932180	-204.13038130
aug-cc-pCVQZ	-54.985662610	-55.591500695	-129.30760011	-204.12761715

Table 6: UHF total energies for open-shell molecules. Estimated HF limit, Gaussian basis energies and geometries are taken from Refs.[44, 45, 46].

Requested precision	Total energy (Hartree)			
	C	N	O	F
$\varepsilon_{tot} = 10^{-5}$	-37.693753180	-54.404705282	-74.819062105	-99.416366868
$\varepsilon_{tot} = 10^{-6}$	-37.693741232	-54.404546399	-74.818985544	-99.416319097
$\varepsilon_{tot} = 10^{-7}$	-37.693740484	-54.404548263	-74.818980697	-99.416306995
$\varepsilon_{tot} = 10^{-8}$	-37.693740393	-54.404548319	-74.818980196	-99.416306085
Est. HF limit	-37.6937	-54.4045	-74.8190	-99.4163
aug-cc-pCV5Z	-37.693694337	-54.404470064	-74.818844591	-99.416094067
aug-cc-pCVQZ	-37.693383688	-54.403857000	-74.817689844	-99.414170791
	CN	CF	HCO	OF
$\varepsilon_{tot} = 10^{-5}$	-92.242768032	-137.23900869	-113.30415122	-174.21099973
$\varepsilon_{tot} = 10^{-6}$	-92.242824581	-137.23895560	-113.30393123	-174.21091105
$\varepsilon_{tot} = 10^{-7}$	-92.242825135	-137.23898474	-113.30400749	-174.21091140
$\varepsilon_{tot} = 10^{-8}$	-92.242824460	-137.23898575	-113.30401610	-174.21091074
Est. HF limit	-92.2428	-137.2390	-113.3040	-174.2109
aug-cc-pCV5Z	-92.242699028	-137.23872998	-113.30382984	-174.21056660
aug-cc-pCVQZ	-92.241728449	-137.23646629	-113.30232555	-174.20751193

Table 7: Computation time for LDA calculation on small alkane systems. The residual norm is converged to $\varepsilon_r \leq 10^{-4}$ using localized orbitals. N_e is the number of electrons and n_{it} is the number of iterations. The overall accuracy of the calculations were $\varepsilon_{tot} = 10^{-5}$, which should reflect the relative accuracy of the presented energies.

Molecule	N_e	n_{it}	Time (sec)	Wall time 16 CPUs	Energy (Hartree)
C_1H_4	10	7	2300	$\sim$ 2 min	-40.12193
C_2H_6	18	8	5800	$\sim$ 5 min	-79.07702
C_4H_{10}	34	9	15500	$\sim$ 15 min	-156.99457
C_6H_{14}	50	9	30000	$\sim$ 30 min	-234.90188
C_8H_{18}	66	8	52000	$\sim$ 60 min	-312.82352
$C_{10}H_{22}$	82	8	82000	$\sim$ 90 min	-390.72429

5.3 Computing time and scaling

A brief illustration of the performance and scaling of the code is presented in Tab. 7 for LDA calculations of small alkane molecules using localized orbitals. As the code is still a prototype in a development phase, no explicit attempt has been made exploiting the locality of the molecular orbitals to reduce the scaling with respect to system size. Therefore the code displays a formal quadratic scaling due to the sum appearing in the non-canonical orbital equations (31), in the calculation of the Fock matrix and in the orbital orthogonalizations. The calculation of the Coulomb potential has been demonstrated to scale linearly in a previous study[30]. In the case of Hartree-Fock there is also quadratic scaling in the calculation of the exchange potential, assuming that the application of the Coulomb operator scales linearly with the system size.

In the future, it should be possible to approach an asymptotic linear scaling of both Kohn-Sham and Hartree-Fock calculations, by taking full advantage of the exponential fall-off of the localized orbitals. Similar screening techniques to the ones already adopted in traditional Gaussian based SCF implementations [32, 47] could be employed in our Multiwavelet implementation.

6 Conclusions

We have presented a new implementation of a Multiwavelet-based SCF solver for HF and DFT, which is able to deal both with closed-shell as well as open-shell systems. Our solver is based on a preconditioned steepest descent step[24], accelerated through the KAIN method[22]. Our implementation is parallel (both OMP and MPI parallelizations are present) and able to deal exclusively with localized orbitals, without any reference to the delocalized canonical orbitals. In order to do that we have shown how it is possible to compute the Fock matrix in the localized orbital basis exploiting the formal relation between the level-shifted Laplacian and the bound-state

Helmholtz kernel, thus avoiding any reference to the kinetic energy operator. We have shown that we are able to obtain high accuracy results (basis-set limit within an arbitrary, predefined threshold) both with the canonical basis and the localized one. We have empirically found that keeping the parameter λ which defines the Helmholtz operator close to the current orbital energy value for each orbital is a near-optimal choice, thus limiting the need to update H^μ to the initial steps of the SCF cycle. Another important finding of our study is the significant improvement of the preconditioned steepest descent convergence without acceleration when a localized basis is employed. This result seems to indicate that localization improves significantly the quality of the preconditioner, thus making the non-accelerated convergence simpler. In fact when Krylov acceleration is added on top only a few iterations are gained indicating that each preconditioned step is already a very good guess. This observation deserves further attention and we will assess whether this is achieved also for other molecular systems, with more complicated electronic structure. The use of localized orbitals is also important for practical but important reasons: (1) the development of low-scaling (asymptotically linear) algorithms; (2) the reduction of the memory footprint of the orbital representation; (3) the exploitation of modern massive parallel architectures.

7 Acknowledgments

This work has been supported by the Research Council of Norway through a Centre of Excellence Grant (Grant No. 179568/V30) and from the Norwegian Supercomputing Program (NOTUR) through a grant of computer time (Grant No. NN4654K).

References

[1] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular Electronic-Structure Theory*. Wiley-Blackwell, 2008.

[2] David Moncrieff and S Wilson. Computational linear dependence in molecular electronic structure calculations using universal basis sets. *Int J Quantum Chem*, 101(4):363–371, 2005.

[3] G Kresse and J Furthmuller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B*, 54(16):11169–11186, 1996.

[4] D Vanderbilt. Soft Self-Consistent Pseudopotentials in a Generalized Eigenvalue Formalism. *Phys. Rev., B Condens. Matter*, 41(11):7892–7895, 1990.

[5] G Y Sun, J Kurti, P Rajczy, M Kertesz, J Hafner, and G Kresse. Performance of the Vienna ab initio simulation package (VASP) in chemical applications. *J Mol Struc-Theochem*, 624:37–45, 2003.

[6] T Torsti, T Eirola, J Enkovaara, T Hakala, P Havu, V Havu, T Hoynalanmaa, J Ignatius, M Lyly, I Makkonen, T T Rantala, J Ruokolainen, K Ruotsalainen, E Rasanen, H Saarikoski, and M J Puska. Three real-space discretization techniques in electronic structure calculations. *Physica Status Solidi B-Basic Solid State Physics*, 243(5):1016–1053, April 2006.

[7] RJ Harrison, GI Fann, T Yanai, Z Gan, and G Beylkin. Multiresolution quantum chemistry: Basic theory and initial applications. *J Chem Phys*, 121:11587, 2004.

[8] T Yanai, GI Fann, Z Gan, RJ Harrison, and G Beylkin. Multiresolution quantum chemistry in multiwavelet bases: Hartree–Fock exchange. *J Chem Phys*, 121:6680, 2004.

[9] B Alpert, G Beylkin, D Gines, and L Vozovoi. Adaptive solution of partial differential equations in multiwavelet bases. *J Comput Phys*, 1999.

[10] Bradley K Alpert. A Class of Bases in L^2 for the Sparse Representation of Integral Operators. *SIAM Journal on Mathematical Analysis*, 24(1):246–262, January 1999.

[11] Florian A Bischoff, Robert J Harrison, and Edward F Valeev. Computing many-body wave functions with guaranteed precision: The first-order Moller-Plesset wave function for the ground state of helium atom. *J Chem Phys*, 137(10):104103, September 2012.

[12] A. Durdek, S. R. Jensen, J. Jusélius, P. Wind, T. Flå, and L. Frediani. Adaptive order polynomial algorithm in a multi-wavelet representation scheme. Submitted, 2013.

[13] G Beylkin, R J Harrison, and K E Jordan. Singular operators in multiwavelet bases. *IBM Journal of Research and Development*, 48(2):161–171, 2004.

[14] G Beylkin. Fast adaptive algorithms in the non-standard form for multidimensional problems. *Appl Comput Harmon A*, 24(3):354–377, 2008.

[15] Stinne Høst, Jeppe Olsen, Branislav Jansik, Lea Thøgersen, Poul Jorgensen, and Trygve Helgaker. The augmented Roothaan-Hall method for optimizing Hartree-Fock and Kohn-Sham density matrices. *J Chem Phys*, 129(12):124106, September 2008.

[16] Yousef Saad. *Numerical Methods for Large Eigenvalue Problems*. Revised Edition. SIAM, 2011.

[17] Lin-Wang Wang and Alex Zunger. Large scale electronic structure calculations using the Lanczos method. *Computational Materials Science*, 2(2):326–340, 1994.

[18] Ernest R Davidson. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J Comput Phys*, 17(1):87–94, 1975.

[19] Peter Pulay. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chem Phys Lett*, 73(2):393–398, 1980.

[20] D M Wood and Alex Zunger. A new method for diagonalising large matrices. *J. Phys. A: Math. Gen.*, 18(9):1343–1359, 1985.

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

[22] RJ Harrison. Krylov subspace accelerated inexact Newton method for linear and nonlinear equations. *J Comput Chem*, 25(3):328–334, 2004.

[23] Thorsten Rohwedder and Reinhold Schneider. An analysis for the DIIS acceleration method used in quantum chemistry calculations. *J Math Chem*, 49(9):1889–1914, October 2011.

[24] Reinhold Schneider, Thorsten Rohwedder, Alexej Neelov, and Johannes Blauert. Direct minimization for calculating invariant subspaces in density functional computations of the electronic structure. *J Comput Math*, 27(2-3):360–387, 2008.

[25] Fritz Keinert. *Wavelets and Multiwavelets*, volume 42 of *Studies in advanced mathematics*. Chapman and Hall, CRC Press, Boca Raton, FL, USA, 2003.

[26] Gregory Beylkin, Vani Cheruvu, and Fernando Perez. Fast adaptive algorithms in the non-standard form for multidimensional problems. *Appl Comput Harmon A*, 24(3):354–377, 2008.

[27] D Gines, G Beylkin, and J Dunn. LU Factorization of Non-standard Forms and Direct Multiresolution Solvers* 1. *Appl Comput Harmon A*, 5(2):156–201, 1998.

[28] Luca Frediani, Eirik Fossgaard, Tor Flå, and Kenneth Ruud. Fully adaptive algorithms for multivariate integral equations using the non-standard form and multiwavelets with applications to the Poisson and bound-state Helmholtz kernels in three dimensions. *Mol Phys*, 111(9-11):1143–1160, July 2013.

[29] G Beylkin and MJ Mohlenkamp. Numerical operator calculus in higher dimensions. *P Natl Acad Sci Usa*, 99(16):10246, 2002.

[30] S. R. Jensen, J. Jusélius, A. Durdek, T. Flå, P. Wind, and L. Frediani. Linear scaling coulomb interaction in the multiwavelet basis, a parallel implementation. Submitted, 2013.

[31] M H Kalos. Monte Carlo calculations of the ground state of three-and four-body nuclei. *Physical Review*, 128(4):1791, 1962.

[32] Lsdalton, a linear scaling molecular electronic structure program, release dalton2013 (2013), see http://daltonprogram.org.

[33] S. F. Boys. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.*, 32:296–299, 1960.

[34] J. M. Foster and S. F. Boys. Canonical configurational interaction procedure. *Rev. Mod. Phys.*, 32:300–302, 1960.

[35] Helgaker T. Optimization of minima and saddle points. In BjrnO. Roos, editor, *Lecture Notes in Quantum Chemistry*, volume 58 of *Lecture Notes in Chemistry*, pages 295–324. Springer Berlin Heidelberg, 1992.

[36] John C Slater. A simplification of the Hartree-Fock method. *Physical Review*, 81(3):385, 1951.

[37] Paul Adrien Maurice Dirac. Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London A*, 123(729):714–733, 1929.

[38] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.

[39] Ulf Ekström, Lucas Visscher, Radovan Bast, Andreas J Thorvaldsen, and Kenneth Ruud. Arbitrary-Order Density Functional Response Theory from Automatic Differentiation. *J Chem Theory Comput*, 6(7):1971–1980, July 2010.

[40] Kestutis Aidas, Celestino Angeli, Keld L Bak, Vebjørn Bakken, Radovan Bast, Linus Boman, Ove Christiansen, Renzo Cimiraglia, Sonia Coriani, Pål Dahle, Erik K Dalskov, Ulf Ekström, Thomas Enevoldsen, Janus J Eriksen, Patrick Ettenhuber, Berta Fernández, Lara Ferrighi, Heike Fliegl, Luca Frediani, Kasper Hald, Asger Halkier, Christof Hättig, Hanne Heiberg, Trygve Helgaker, Alf Christian Hennum, Hinne Hettema, Eirik Hjertenaes, Stinne Høst, Ida-Marie Høyvik, Maria Francesca Iozzi, Branislav Jansik, Hans Jørgen Aa Jensen, Dan Jonsson, Poul Jorgensen, Joanna Kauczor, Sheela Kirpekar, Thomas Kjaergaard, Wim Klopper, Stefan Knecht, Rika Kobayashi, Henrik Koch, Jacob Kongsted, Andreas Krapp, Kasper Kristensen, Andrea Ligabue, Ola B Lutnaes, Juan I Melo, Kurt V Mikkelsen, Rolf H Myhre, Christian Neiss, Christian B Nielsen, Patrick Norman, Jeppe Olsen, Jógvan Magnus H Olsen, Anders Osted, Martin J Packer, Filip Pawlowski, Thomas B Pedersen, Patricio F Provasi, Simen Reine, Zilvinas Rinkevicius, Torgeir A Ruden, Kenneth Ruud, Vladimir V Rybkin, Pawel Sałek, Claire C M Samson, Alfredo Sánchez de Merás, Trond Saue, Stephan P A Sauer, Bernd Schimmelpfennig, Kristian Sneskov, Arnfinn H Steindal, Kristian O Sylvester-Hvid,

Peter R Taylor, Andrew M Teale, Erik I Tellgren, David P Tew, Andreas J Thorvaldsen, Lea Thøgersen, Olav Vahtras, Mark A Watson, David J D Wilson, Marcin Ziolkowski, and Hans Agren. The Dalton quantum chemistry program system. *WIREs Comput Mol Sci*, pages n/a–n/a, September 2013.

[41] J Stephen Binkley, John A Pople, and Warren J Hehre. Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J Am Chem Soc*, 102(3):939–947, 1980.

[42] Mark S Gordon, J Stephen Binkley, John A Pople, William J Pietro, and Warren J Hehre. Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. *J Am Chem Soc*, 104(10):2797–2803, May 1982.

[43] http://physics.nist.gov/PhysRefData/DFTdata/Tables/ptable.html.

[44] A. Tajti, P. G. Szalay, A. G. Csszr, M. Kllay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vzquez, and J. F. Stanton. Heat: High accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.*, 121(23):11599–11613, 2004.

[45] Y. J. Bomble, J. Vzquez, M. Kllay, C. Michauk, P. G. Szalay, A. G. Csszr, J. Gauss, and J. F. Stanton. High-accuracy extrapolated ab initio thermochemistry. ii. minor improvements to the protocol and a vital simplification. *J. Chem. Phys.*, 125(6), 2006.

[46] M. E. Harding, J. Vzquez, B. Ruscic, A. K. Wilson, J. Gauss, and J. F. Stanton. High-accuracy extrapolated ab initio thermochemistry. iii. additional improvements and overview. *J. Chem. Phys.*, 128(11), 2008.

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