Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Numerical solution of the Hartree-Fock equation in multilevel tensor-structured format

(revised version: October 2010)

by

Boris N. Khoromskij, Venera Khoromskaia, and Heinz-Jürgen Flad

Preprint no.: 44 2009

Numerical solution of the Hartree-Fock equation in multilevel tensor-structured format

B.N. Khoromskij[a], V. Khoromskaia[a] and H.-J. Flad[b]
[a] Max-Planck-Institute for Mathematics in the Sciences,
Inselstr. 22-26, D-04103 Leipzig, Germany.
[b] TU Berlin, Strasse des 17.Juni 136, D-10623 Berlin, Germany.
{bokh, vekh, flad}@mis.mpg.de

Abstract

In this paper, we describe a novel method for robust and accurate iterative solution of the self-consistent Hartree-Fock equation in $\mathbb{R}^3$ based on the idea of tensor-structured computation of the electron density and the nonlinear Hartree and (nonlocal) exchange operators at all steps of the iterative process. We apply the self-consistent field (SCF) iteration to the Galerkin discretisation in a set of low separation rank basis functions that are solely specified by the respective values on a 3D Cartesian grid. The approximation error is estimated by $O(h^3)$, where $h = O(n^{-1})$ is the mesh size of $n \times n \times n$ tensor grid, while the numerical complexity to compute the Galerkin matrices scales linearly in $n \log n$. We propose the tensor-truncated version of the SCF iteration using the traditional direct inversion in the iterative subspace (DIIS) scheme enhanced by the multilevel acceleration with the grid dependent termination criteria at each discretization level. This implies that the overall computational cost scales almost linearly in the univariate problem size n. Numerical illustrations are presented for the all electron case of H_2O, and pseudopotential case of CH_4 and CH_3OH molecules. The proposed scheme is not restricted to *a priori* given rank-1 basis sets allowing analytically integrable convolution transform with the Newton kernel, that opens further perspectives for promotion of the tensor-structured methods in computational quantum chemistry.

AMS Subject Classification: 65F30, 65F50, 65N35, 65F10
Key words: Orthogonal Tucker tensor decomposition, canonical model, tensor-truncated methods, discrete convolution, Hartree-Fock equation, Coloumb and exchange matrices, multigrid accelerated SCF iteration.

1 Introduction

In recent years the concept of tensor-structured numerical methods has opened new perspectives for solving the basic equations of mathematical physics in $\mathbb{R}^d$, $d \geq 3$, in particular, the many-particle Schrödinger equation [2, 11, 3, 27], and multidimensional elliptic spectral problems [13, 21]. Tensor methods are based on the idea of separable approximation of d-variate functions and related operators arising in the discretisation and solution process

posed in a multivariate function space in $\mathbb{R}^d$. This concept appears to be particularly attractive for simplified approaches, like the *ab initio* Hartree-Fock and density functional theory methods, reducing the dimensionality of the problem to $d = 3$, though the kernels of the nonlocal operators involved are functions of six spatial dimensions [14, 6, 22, 23, 24, 19].

The traditional numerical methods in *ab initio* electronic structure calculations are based on the Galerkin approximation in the problem-dependent Gaussian-type orbitals (GTO) basis (meshless methods). Many efforts have been devoted to the development of rigorous schemes for the analytical evaluation of the so-called two-electron integrals inherent for this approach, which have yielded state-of-the-art software packages [34]. The Hartree-Fock model presupposes at least cubic (or fourfold) scaling in the number of the basis functions [26]. Hence computations using Gaussian bases for large molecules may become infeasible. The simplified models based on the appropriately adjusted pseudopotentials, can be solved by using grid-oriented methods over $n \times n \times n$ spacial grids via the traditional plane waves, wavelet or finite element discretizations, with computational cost which scales at least linearly in the volume size, $N_V = n^3$, [1, 9, 10]. In this way, the practically tractable grid-size for the calculations using these traditional approaches is limited by the value $n \approx 500$.

The principal question then arises: is it possible to solve the Hartree-Fock/Kohn-Sham models by the grid-based methods with linear scaling in the univariate grid-size n, i.e., sublinear in the volume, $O(N_V^{1/3})$? In what follows, we give the promising answer to this question by introducing a tensor-structured numerical scheme that solves the Hartree-Fock equation with $O(n \log n)$-complexity. To that end, we suggest the novel concept for the numerical solution of the Hartree-Fock equation which is based on the use of a moderate number of the problem-adapted discrete Galerkin basis functions living on 3D Cartesian grid, and represented with a low separation rank. Such a basis can be viewed as a kind of algebraic generalization/optimisation of the traditional GTO or Slater-type orbitals providing the way to $O(n \log n)$ discrete evaluation of the arising six-dimensional volume integrals.

The core of our method is the tensor-structured computation of the electron density and the Galerkin matrices of the nonlinear Hartree and (nonlocal) exchange operators at all steps of iterations on nonlinearity, based on the systematic use of the rank-truncated linear tensor-tensor operations (see Appendix and [23, 24, 17]). Within the solution process, all principal multilinear algebra operations, such as the scalar and Hadamard products, the laborious 3D convolution transform, and the rank truncation procedures are implemented with $O(n)$-complexity. Due to almost linear scaling in n of the 3D tensor-structured arithmetic, we achieve high accuracy of calculations due to accessibility of the large $n \times n \times n$ tensor grids of size up to 16384^3 at the finest approximation level. In electronic structure calculations this implies rather fine resolution with the mesh size $h \approx 10^{-4}$ $\mathring{A}$ providing possibility for arbitrary space orientation of a molecule in the computational box.

Particularly, the *self-consistent field* (SCF) iteration applies to the Galerkin discretisation of the Hartree-Fock equation with respect to certain problem adapted basis with low separation rank. Making use of piecewise linear grid representation of the Galerkin basis functions, and piecewise constant representation of the electron density, leads to the approximation error of order $O(h^2)$, in the Hartree and exchange potentials, where $h = O(n^{-1})$ is the respective mesh size. In the case of only few spacial singularities, the locally refined tensor grids can be adapted.

In this paper the traditional *direct inversion in the iterative subspace* (DIIS) SCF iteration

scheme commonly used in the physical literature [30, 15, 4] is applied. We enhance the DIIS iteration by a multilevel strategy with grid dependent termination criteria at each discretization level. It has a two-fold effect, providing a good initial guess on finer grids, and allowing improved approximation $O(h^3)$, via Richardson extrapolation over a sequence of grids. The discrete orbitals, represented by the respective coefficients vectors are updated by diagonalising the Galerkin stiffness matrix at each iteration on nonlinearity at the expense $O(N_b^3)$, where N_b is the dimension of the Galerkin subspace. In general, the convergence proof for the nonlinear DIIS iteration is still an open question [26].

We observe that in numerical practice, our multigrid accelerated DIIS iteration exhibits fast and uniform in n convergence (linear convergence rate) so that the overall computational time scales linearly in n—the tool apparently works. It is worth noting that the current version of our method still scales cubically in the size of approximating basis. Hence, any *algebraic optimisation of this basis set* within the solution process may allow fast and high accuracy *ab initio* computations for large molecules. Quadratic scaling in the size of the approximating basis might be possible in the case of iterative solution of the Galerkin spectral problem, or in the framework of direct minimization algorithms (see [26, 32] for detailed discussion on direct minimization).

We present numerical illustrations for the all electron case of H_2O, and pseudopotential cases of the CH_4 and CH_3OH molecules using the particular Galerkin basis set via discretized GTO basis functions. The GTO basis is chosen only so that convenient comparison might be made with the standard MOLPRO package which computes integrals related to this basis analytically [34]. Numerical computations confirm almost linear scaling in the grid-size, n, indicating that the proposed tensor-structured SCF iteration may provide efficient solution of *ab initio* and DFT computations for large molecules.

The rest of the paper is organized as follows. In §2 we describe the standard Galerkin scheme for the nonlinear Hartree-Fock equation, and discuss various types of commonly used sets of Galerkin basis functions. Our choice can be only constrained by requirements on the low separation rank of the individual basis functions, and possibly low dimension of the Galerkin subspace. In §2.3, we introduce nonstandard agglomerated representations of the Coulomb and Hartree-Fock exchange Galerkin matrices, which are well suited for tensor arithmetic. §2.5 describes the basic representation of the Fock operator in the rank-structured tensor format, which is the key point for efficient $O(n \log n)$-implementation of the tensor-truncated SCF iteration. In §3, we first formulate the SCF iteration that implements the unigrid tensor-truncated DIIS scheme. The (cascadic) multigrid version of the tensor-truncated DIIS iteration includes Richardson extrapolation over the final pair of sequential (refined) grids. We prove $O(N_b^3 n \log n)$-complexity of the proposed numerical method. In the particular case of GTO basis, N_b is proportional to the number of electrons in the molecule. §4 presents various numerical illustrations in the case of moderate-size molecules, which confirm the theoretical prediction of $O(n \log n)$-complexity. In §5 the main conclusions are formulated. In the appendix we describe the tensor-structured formats used in our computational scheme.

2 Galerkin scheme and tensor approximation

2.1 Problem setting

We introduce the tensor-truncated numerical method to compute valuable quantities in the $2N$-electrons Hartree-Fock equation for pairwise L^2-orthogonal electronic orbitals $\psi_i : \mathbb{R}^3 \to \mathbb{R}$, $\psi_i \in H^1(\mathbb{R}^3)$, that reads as

$$\mathcal{F}_\Phi \psi_i(x) = \lambda_i \psi_i(x), \quad \int_{\mathbb{R}^3} \psi_i \psi_j dx = \delta_{ij}, \ i,j = 1,...,N \tag{2.1}$$

with $\mathcal{F}_\Phi$ being the nonlinear Fock operator

$$\mathcal{F}_\Phi := -\frac{1}{2}\Delta + V_c + V_H + \mathcal{K}.$$

Here we use the definitions

$$\tau(x,y) := 2\sum_{i=1}^{N} \psi_i(x)\psi_i(y), \quad \rho(x) := \tau(x,x),$$

for the density matrix $\tau(x,y)$, and electron density $\rho(x)$, and

$$V_c(x) = -\sum_{\nu=1}^{M} \frac{Z_\nu}{\|x - a_\nu\|}, \quad Z_\nu > 0, \ \ a_\nu \in \mathbb{R}^3,$$

for the nuclear potential. The Hartree potential $V_H(x)$ is given by

$$V_H(x) := \rho \star \frac{1}{\|\cdot\|} = \int_{\mathbb{R}^3} \frac{\rho(y)}{\|x-y\|}\, dy, \quad x \in \mathbb{R}^3, \tag{2.2}$$

while the nonlocal exchange operator $\mathcal{K}$ reads as

$$(\mathcal{K}\psi)(x) := -\sum_{i=1}^{N} \left(\psi\, \psi_i \star \frac{1}{\|\cdot\|} \right) \psi_i(x) = -\frac{1}{2} \int_{\mathbb{R}^3} \frac{\tau(x,y)}{\|x-y\|}\, \psi(y) dy. \tag{2.3}$$

2.2 Standard Galerkin scheme

Usually, the Hartree-Fock equation is solved by the standard Galerkin approximation of the initial problem in the form (2.1) posed in $H^1(\mathbb{R}^3)$ (see [26] for more details). For a given finite basis set $\{g_\mu\}_{1 \le \mu \le N_b}$, $g_\mu \in H^1(\mathbb{R}^3)$, the molecular orbitals ψ_i are represented (approximately) as

$$\psi_i = \sum_{\mu=1}^{N_b} C_{\mu i} g_\mu, \quad i = 1,...,N. \tag{2.4}$$

To derive the equation for the unknown coefficients matrix $C = \{C_{\mu i}\} \in \mathbb{R}^{N_b \times N}$, we first introduce the mass (overlap) matrix $S = \{S_{\mu\nu}\}_{1 \le \mu,\nu \le N_b}$, given by

$$S_{\mu\nu} = \int_{\mathbb{R}^3} g_\mu g_\nu dx,$$

the stiffness matrix $H = \{h_{\mu\nu}\}$ of the core Hamiltonian $\mathcal{H} = -\frac{1}{2}\Delta + V_c$,

$$h_{\mu\nu} = \frac{1}{2}\int_{\mathbb{R}^3} \nabla g_\mu \cdot \nabla g_\nu dx + \int_{\mathbb{R}^3} V_c(x) g_\mu g_\nu dx, \quad 1 \le \mu, \nu \le N_b,$$

and the symmetric density matrix

$$D = 2CC^* \in \mathbb{R}^{N_b \times N_b}. \tag{2.5}$$

The nonlinear terms representing the Galerkin approximation of the Hartree and exchange operators are usually constructed by using the so-called two-electron integrals, defined as

$$b_{\mu\nu,\kappa\lambda} = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{g_\mu(x) g_\nu(x) g_\kappa(y) g_\lambda(y)}{\|x-y\|} dxdy, \quad 1 \le \mu, \nu, \kappa, \lambda \le N_b.$$

Introducing the $N_b \times N_b$ matrices $J(D)$ and $K(D)$, with D defined by (2.5),

$$J(D)_{\mu\nu} = \sum_{\kappa,\lambda=1}^{N_b} b_{\mu\nu,\kappa\lambda} D_{\kappa\lambda}, \quad K(D)_{\mu\nu} = -\frac{1}{2}\sum_{\kappa,\lambda=1}^{N_b} b_{\mu\lambda,\nu\kappa} D_{\kappa\lambda},$$

and then the complete Fock matrix F,

$$F(D) = H + G(D), \quad G(D) = J(D) + K(D), \tag{2.6}$$

one obtains the respective Galerkin system of nonlinear equations for the coefficients matrix $C \in \mathbb{R}^{N_b \times N}$,

$$\begin{aligned} F(D)C &= SC\Lambda, \quad \Lambda = diag(\lambda_1, ..., \lambda_N), \\ C^*SC &= I_N, \end{aligned} \tag{2.7}$$

where the second equation represents the orthogonality constraints $\int_{\mathbb{R}^3} \psi_i \psi_j = \delta_{ij}$, with I_N being the $N \times N$ identity matrix.

In the standard implementation based on the precomputed two-electron integrals, the complexity to build up the matrix G scales as $O(N_b^4)$, that is dominated by computational cost for the exchange matrix $K(D)$. In turn, the core Hamiltonian H can be precomputed in $O(N_b^2)$ operations, hence, in the following, we will not focus on this issue.

The nonlinear system (2.7) can be solved by certain SCF iteration, where at each iterative step the respective linear eigenvalue problem has to be solved with the updated matrix $G(D)$. Given $F(D)$, using the direct diagonalization for solving the system (2.7) leads to the cost $O(N_b^3)$. The alternative approach can be based on the direct minimization of the Hartree-Fock energy functional,

$$I^{HF} = \inf\left\{ \frac{1}{2}\sum_{i=1}^{N}\int_{\mathbb{R}^3} |\nabla\psi_i|^2 + \int_{\mathbb{R}^3} \rho V_c + \frac{1}{2}\int\int_{\mathbb{R}^3} \frac{\rho(x)\rho(y) - |\tau(x,y)|^2}{\|x-y\|} dxdy \right\},$$

under the orthogonality constraints in (2.1), see [32] for more details.

2.3 Agglomerated representation of the Galerkin matrices

In our approach, fast and accurate evaluation of the Galerkin matrices $J(D)$ and $K(D)$ is based on certain reorganization of the standard computational scheme given in §2.2. Specifically, instead of precomputing the full set of two-electron integrals $b_{\mu\nu,\kappa\lambda}$ and the elements of the density matrix D, we use agglomerated representations for $J(D))$ and $K(D)$. In particular, the Galerkin representation of the Hartree operator (the Coloumb matrix) is now based on the agglomerated integrals,

$$J(D)_{\mu\nu} = \int_{\mathbb{R}^3} g_\mu(x) V_H(x) g_\nu(x) dx, \quad 1 \le \mu, \nu \le N_b, \tag{2.8}$$

including a single convolution transform in $\mathbb{R}^3$ to compute the Hartree potential in (2.2),

$$V_H = \rho * \frac{1}{\|\cdot\|},$$

where the electron density is given by

$$\rho(y) = 2 \sum_{a=1}^{N} \left(\sum_{\kappa,\lambda=1}^{N_b} C_{\kappa a} C_{\lambda a} g_\kappa(y) g_\lambda(y) \right). \tag{2.9}$$

In turn, as proposed in [17], we represent the matrix entries of $K(D)$ by the following three loops: For $a = 1, ..., N$, compute the convolution integrals,

$$W_{a\nu}(x) = \int_{\mathbb{R}^3} \frac{g_\nu(y) \sum\limits_{\kappa=1}^{N_b} C_{\kappa a} g_\kappa(y)}{\|x - y\|} dy, \quad \nu = 1, ..., N_b, \tag{2.10}$$

and then the scalar products

$$K_{\mu\nu,a} = \int_{\mathbb{R}^3} \left[\sum_{\kappa=1}^{N_b} C_{\kappa a} g_\kappa(x) \right] g_\mu(x) W_{a\nu}(x) dx, \quad \mu, \nu = 1, ..., N_b. \tag{2.11}$$

Finally, the entries of the exchange matrix are given by sums over all orbitals,

$$K(C)_{\mu\nu} = \sum_{a=1}^{N} K_{\mu\nu,a}, \quad \mu, \nu = 1, ..., N_b. \tag{2.12}$$

The advantage of above representations is due to the minimization of the number of convolution products that have to be computed by numerical quadratures. What is even more important, that we have the possibility of efficient low-rank separable approximation of the discretised density $\rho(x)$ as well as of the auxiliary potentials $W_{a\nu}(x)$ at step (2.10).

Effective realization of such a concept is based on certain unrestrictive technical assumptions on the Galerkin basis functions g_μ. First, we suppose that the initial problem is posed in the finite volume box $\Omega = [-b, b]^3 \in \mathbb{R}^3$ subject to the homogeneous Dirichlet boundary

conditions on $\partial\Omega$ (due to the exponential decay of the orbitals $\psi_i(x)$, as $\|x\| \to \infty$). For given discretization parameter $n \in \mathbb{N}$, introduce the equidistant tensor grid

$$\omega_{\mathbf{3},n} := \omega_1 \times \omega_2 \times \omega_3, \quad \omega_\ell := \{-b + (m-1)h : m = 1, ..., n+1\}, \quad \ell = 1, ..., 3, \qquad (2.13)$$

with the mesh-size $h = 2b/n$. Define the set of piecewise constant basis functions $\{\phi_{\mathbf{i}}\}$, $\mathbf{i} \in \mathcal{I} := \{1, ..., n\}^3$, associated with the respective grid-cells in $\omega_{\mathbf{3},n}$ (indicator functions), and the corresponding set $\{\chi_{\mathbf{j}}\}$, $\mathbf{j} \in \mathcal{J} := \{1, ..., n-1\}^3$, of tensor-product continuous piecewise linear (in each spacial variable) polynomials. We denote the corresponding FE spaces as

$$\mathcal{V}_n = \text{span}\{\phi_{\mathbf{i}}\}, \quad \text{and} \quad \mathcal{W}_n = \text{span}\{\chi_{\mathbf{j}}\} \in H_0^1(\Omega).$$

Now the basis set $\{g_\mu\}$ is supposed to satisfy the following properties:

(A) (Approximability). The Galerkin approximation error over the quantities in (2.7) is physically admissible.

(B) (Separability). Each basis function $g_\mu(x) \in H_0^1(\Omega)$, can be represented by the R_G-term separable expansion in $x = (x_1, x_2, x_3)$, with moderate number of terms R_G,

$$g_\mu(x) = \sum_{k=1}^{R_G} g_{\mu,k}^{(1)}(x_1) g_{\mu,k}^{(2)}(x_2) g_{\mu,k}^{(3)}(x_3), \quad \mu = 1, ..., N_b. \qquad (2.14)$$

(C) (Discrete separability). Functions $g_\mu(x)$ allow the approximate representation in either basis sets $\{\phi_{\mathbf{i}}\}$ and $\{\chi_{\mathbf{j}}\}$, by the rank-R_G coefficients tensors $G_\mu = [G_{\mu,\mathbf{i}}] \in \mathbb{R}^{\mathcal{I}}$ and $X_\mu = [X_{\mu,\mathbf{j}}] \in \mathbb{R}^{\mathcal{J}}$, respectively.

(D) (Separable quadratures). The Galerkin integrals for $J(D)$ and $K(D)$ given by (2.8) - (2.12) can be accurately represented by the well separable numerical quadratures in the discrete basis sets $\{G_\mu\}$ and $\{X_\mu\}$, providing asymptotical convergence as $h \to 0$.

Notice that the basis sets $\{\phi_{\mathbf{i}}\}$ and $\{\chi_{\mathbf{j}}\}$ can be generalized to those based on the higher order piecewise polynomials over nonuniform (locally refined) tensor grids. This will only concern with some technical aspects of our approach. The particular numerical effects of such generalizations should be carefully verified on realistic data in electronic structure calculation.

2.4 On the choice of the Galerkin basis functions

The examples of *problem-independent* grid-oriented basis sets are given by plane waves, wavelets, and by the piecewise polynomial *finite element* (FE) basis functions already mentioned in the Introduction. Usually, the dimension of the respective Galerkin spaces is much larger than in the case of problem dependent basis sets (see below). The practically tractable grids (indices) of size $n \times n \times n$ are presently limited by the value $n \approx 500$.

Several efficient "meshless" basis sets $\{g_\mu\}$ are known in the literature on computational quantum chemistry. In particular, we mention the *linear combination of atomic orbitals* (LCAO) and their successors, *Slater-type orbitals* (STOs). The most popular are the so-called *Gaussian-type orbitals* (GTOs) and their more general version, *contracted Gaussian*

functions, which probably constitute the best compromise between STOs and GTOs (cf. [26] for detailed discussion). The construction of such *problem dependent* basis sets is distinctively based on the precomputed electronic orbitals for single atoms.

An alternative to the analytically given GTO-type basis functions are the so-called *fully numerical atomic orbitals* [26], that are solely specified by their numerical values on a grid. Such a choice of basis functions fits well the spirit of our tensor-structured numerical method. On the one hand, this allows to utilize the already existing problem adapted basis sets taking advantage of the important physical information, that is well known for the individual atoms. At this step, in the present approach, one has the possibility to further algebraic optimization of the Galerkin subspace (reduction of the Galerkin dimension N_b). On the other hand, the (low-rank) separable representation of functions and operators reduces the 3D calculations to fast numerical operations implemented only on the univariate grids (1D calculations) [22, 17]. In this way, the computation of the volume integrals, convolution transforms, scalar products and function-function multiplications can be simplified dramatically.

The particular requirements on the approximating basis set to be fulfilled in the framework of our tensor-structured numerical scheme are formulated in the previous section (see conditions (A)-(D) in §2.3). The systematic construction of the high-quality low tensor rank approximating basis can be established on:

- Algebraic optimization of the conventional 'meshless" GTO-type basis sets ($R_G = 1$);
- Rank reduction of the Slater-type basis ($R_G = O(\log \varepsilon^{-1})$, up to the tolerance $\varepsilon > 0$, cf. [18]);
- Using the united (agglomerated) orthogonal Tucker vectors, whose rank is supposed to be weakly dependent on the particular molecule and the grid parameters [22, 23, 24].

All these concepts still require further theoretical and numerical analysis and will be addressed elsewhere.

The main advantage of the low tensor rank approximating basis sets is the linear scaling of the resultant algorithms in n, that already allows to advent the huge $n \times n \times n$-grids in $\mathbb{R}^3$ (specifically, $n \le 2 \cdot 10^4$, in the contemporary computing practice on the base of tensor-structured methods). This could be benefitial in the FEM-DFT computations applied to large molecular clusters.

2.5 Tensor computation of the Galerkin integrals in $J(D)$, $K(D)$

The beneficial feature of our method is that functions and operators involved in the computational scheme (2.8) - (2.12) are efficiently evaluated using (approximate) low-rank tensor-product representations in the basis sets $\{G_\mu\}$ and $\{X_\mu\}$ at the expense that scales linear-logarithmic in n, $O(n \log n)$.

To that end, we introduce some interpolation/prolongation operators interconnecting the continuous functions on Ω and their discrete representation on the grid via the coefficient tensors in $\mathbb{R}^{\mathcal{I}}$ (or in $\mathbb{R}^{\mathcal{J}}$). Note that the coefficients space of 3-tensors

$$\mathbb{V}_n = \mathbb{R}^{\mathcal{I}} := V_1 \otimes V_2 \otimes V_3,$$

is the tensor-product space with $V_\ell = \mathbb{R}^n$, $(\ell = 1, 2, 3)$, cf. Appendix. Conventionally, we use the canonical isomorphism between $\mathcal{V}_n$ and $\mathbb{V}_n$,

$$\mathcal{V}_n \ni f(x) = \sum_{\mathbf{i}} f_{\mathbf{i}} \phi_{\mathbf{i}}(x) \quad \Longleftrightarrow \quad F := [f_{\mathbf{i}}]_{\mathbf{i}\in\mathcal{I}} \in \mathbb{V}_n.$$

We make use of similar entities for the pair $\mathcal{W}_n$ and $\mathbb{W}_n = \mathbb{R}^{\mathcal{J}} := W_1 \otimes W_2 \otimes W_3$, with $W_\ell = \mathbb{R}^{n-1}$, $(\ell = 1, 2, 3)$.

Now we define the collocation and L^2-projection mappings onto $\mathbb{V}_n$. For the continuous function f, we introduce the collocation "projection" operator by

$$\mathcal{P}_C : f \mapsto \sum_{\mathbf{i}} f_(y_{\mathbf{i}}) \phi_{\mathbf{i}}(x) \quad \Longleftrightarrow \quad F := [f(y_{\mathbf{i}})]_{\mathbf{i}\in\mathcal{I}} \in \mathbb{V}_n,$$

where $\{y_{\mathbf{i}}\}$ is the set of cell-centered points with respect to the grid $\omega_{\mathbf{3},n}$. Furthermore, for functions $f \in L^2(\Omega)$, we define L^2-projection by

$$\mathcal{P}_0 : f \mapsto \sum_{\mathbf{i}} \langle f, \phi_{\mathbf{i}} \rangle \phi_{\mathbf{i}}(x) \quad \Longleftrightarrow \quad F := [\langle f, \phi_{\mathbf{i}} \rangle]_{\mathbf{i}\in\mathcal{I}} \in \mathbb{V}_n.$$

Likewise, we denote the L^2-projection onto $\mathbb{W}_n$ by $\mathcal{Q}_0$.

Using the discrete representations as above, we are able to rewrite all functional and integral transforms in (2.8) - (2.12), in terms of tensor operations in $\mathbb{V}_n$. In particular, for the continuous targets, the function-times-function, and the L^2-scalar product can be dicretised by tensor operations as

$$f \cdot g \mapsto F \odot G \in \mathbb{V}_n, \quad \text{and} \quad \langle f, g \rangle \mapsto h^3 \langle F, G \rangle,$$

with

$$F = \mathcal{P}_C(f), \quad G = \mathcal{P}_C(g),$$

where the scaling constant defines the grid-cell volume h^3, and $\odot$ means the Hadamard (entrywise) product of tensors.

The convolution product is represented by

$$f * g \mapsto F *_T G \in \mathbb{V}_n, \quad \text{with} \quad F = \mathcal{P}_C(f) \in \mathbb{V}_n,\ G = \mathcal{P}_0(g) \in \mathbb{V}_n,$$

where the tensor operation $*_T$ stands for the tensor-structured convolution transform in $\mathbb{V}_n$ described in [20] (see also [24, 23] for application of fast $*_T$ transform in electronic structure calculations). Related to the *separable quadrature* assumption (cf. item (D) in §2.3), we notice that under certain assumptions on the regularity of the input functions the tensor product convolution $*_T$ can be proven to provide an approximation error of order $O(h^2)$, while the two-grid version via the Richardson extrapolation leads to the improved error bound $O(h^3)$ (cf. [20]).

Representations (2.8) - (2.9) can be now rewritten (approximately) in the discrete tensor form as follows,

$$\rho \approx \Theta := \sum_{a=1}^{N} \left(\sum_{\kappa,\lambda=1}^{N_b} C_{\kappa a} C_{\lambda a} G_\kappa \odot G_\lambda \right), \quad \text{where} \quad G_\kappa = \mathcal{P}_C(g_\kappa),$$

and then

$$V_H = \rho * g \approx \Theta *_T P_N, \quad \text{where} \quad P_N = \mathcal{P}_0(g),\ g = \frac{1}{\|\cdot\|}, \tag{2.15}$$

with $P_N \in \mathbb{V}_n$ being the collocation tensor for the Coloumb potential. This implies the tensor representation of the Coloumb matrix,

$$J(D)_{\mu\nu} \approx \langle G_\mu \odot G_\nu, \Theta *_T P_N \rangle, \quad 1 \le \mu, \nu \le N_b. \tag{2.16}$$

The separability property (A) ensures that $rank(G_\mu) \le R_G$, while tensors Θ and P_N are to be approximated by low-rank tensors. Hence, in our method, the corresponding tensor operations are implemented using fast multilinear algebra equipped with the corresponding rank optimization (tensor truncation).

Likewise, tensor representations (2.10) - (2.12) realized in [17], now look as follows,

$$W_{a\nu} \approx \Upsilon_{a\nu} := \left[G_\nu \odot \sum_{\kappa=1}^{N_b} C_{\kappa a} \odot G_\kappa \right] *_T P_N, \quad \nu = 1, ..., N_b, \tag{2.17}$$

with the tensor $P_N \in \mathbb{V}_n$ defined by (2.15),

$$K_{\mu\nu,a} \approx \chi_{\mu\nu,a} := \langle \left[\sum_{\kappa=1}^{N_b} C_{\kappa a} G_\kappa \right] \odot G_\mu, \Upsilon_{a\nu} \rangle, \quad \mu, \nu = 1, ..., N_b, \tag{2.18}$$

finally providing the entries of the exchange matrix,

$$K(D)_{\mu\nu} = \sum_{a=1}^{N} \chi_{\mu\nu,a}, \quad \mu, \nu = 1, ..., N_b. \tag{2.19}$$

Again, the auxiliary tensors and respective algebraic operations have to implemented with the truncation to low-rank tensor formats.

Notice that the core Hamiltonian $H = \{h_{\mu\nu}\}$ can be computed by the respective tensor operations in $\mathbb{W}_n$, and $\mathbb{V}_n$,

$$h_{\mu\nu} \approx \frac{1}{2} \langle \nabla_T \overline{G}_\mu, \nabla_T \overline{G}_\nu \rangle_{(\mathbb{W}_n)^3} + \langle V_0, G_\mu \odot G_\nu \rangle_{\mathbb{V}_n}, \quad 1 \le \mu, \nu \le N_b, \tag{2.20}$$

where $V_0 = \mathcal{P}_0(V_c) \in \mathbb{V}_n$, and where the rank-$R_G$ tensors $\overline{G}_\mu$ ($\mu = 1, ..., N_b$) represent the Galerkin basis functions g_μ in $\mathbb{W}_n$ by $\overline{G}_\mu = \mathcal{Q}_0(g_\mu)$. Furthermore, the operator

$$\nabla_T : \mathbb{W}_n \to (\mathbb{W}_n)^3 := \{ \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} : w_k \in \mathbb{W}_n,\ k = 1, 2, 3 \},$$

denotes the discrete gradient map by 3-way central differences at $y_{\mathbf{j}}$, $\mathbf{j} \in \mathcal{J}$, where $(\mathbb{W}_n)^3$ is the "vector space" of 3-tensors.

With the particular requirements on the rank, $rank(\overline{G}_\mu) \le R_G$, the operator ∇_T applies to each individual rank-1 canonical 3-tensor in $\mathbb{W}_n$ by

$$\nabla_T(x_1 \otimes x_2 \otimes x_3) := \begin{pmatrix} \nabla_1 x_1 \otimes x_2 \otimes x_3 \\ x_1 \otimes \nabla_2 x_1 \otimes x_3 \\ x_1 \otimes x_2 \otimes \nabla_3 x_3 \end{pmatrix} \in (\mathbb{W}_n)^3,$$

where the univariate discrete gradient matrix ∇_k $(k = 1, 2, 3)$ is defined conventionally by central differences on the vectors in $\mathbb{R}^{n-1}$ with zero extension, imposed by the homogeneous Dirichlet boundary conditions (recall that $\mathcal{W}_n \in H_0^1(\Omega)$).

3 Multilevel tensor-truncated SCF iteration via DIIS

3.1 General SCF iteration

The standard SCF algorithm can be formulated as the following "fixed-point" iteration ([26]): Starting from initial guess C_0, perform iterations of the form

$$\begin{aligned} \overline{F}_k C_{k+1} &= S C_{k+1} \Lambda_{k+1}, \quad \Lambda_{k+1} = \mathrm{diag}(\lambda_1^{k+1}, ..., \lambda_{N_{orb}}^{k+1}) \\ C_{k+1}^T S C_{k+1} &= I_{N_{orb}}, \end{aligned} \tag{3.1}$$

where current Fock matrix $\overline{F}_k = \Phi(C_k, C_{k-1}, \ldots, C_0)$, $k = 0, 1, ...$, is specified by the particular relaxation scheme. For example, for the simplest approach, called the *Roothaan algorithm*, one has $\overline{F}_k = F(C_k)$. In practically interesting situations this algorithm usually leads to "flip-flop" stagnation [26].

Recall that here, $\lambda_1^{k+1} \le \lambda_2^{k+1} \le ... \le \lambda_{N_{orb}}^{k+1}$ are N_{orb} negative eigenvalues of the linear generalized eigenvalue problem

$$\overline{F}_k U = \lambda S U, \tag{3.2}$$

and $R_0 \times N_{orb}$ matrices C_{k+1} contain the respective N_{orb} orthonormal eigenvectors $U_1, ..., U_{N_{orb}}$. We denote by $\overline{C}_{k+1} \in \mathbb{R}^{R_0 \times R_0}$ a matrix representing the full set of orthogonal eigenvectors in (3.2).

We use the particular choice of $\overline{F}_k$, $k = 0, 1, ...$, via the DIIS-algorithm, (cf. [30]), with the starting value $\overline{F}_0 = F(C_0) = H$.

We propose the modification to the standard DIIS iteration, by carrying out the iteration on a sequence of successively refined grids with the grid-dependent stopping criteria. The multilevel implementation provides robust convergence from the zero initial guess for the Hartree and exchange operators. The coarse-to-fine grids iteration, in turn, accelerates the solution process dramatically due to low cost of the coarse grid calculations.

The principal feature of our tensor-truncated iteration is revealed on the fast update of the Fock matrix $F(C)$ by using tensor-product multilinear algebra of 3-tensors accomplished with the rank truncation described above. Another important point is that the multilevel implementation provides simple and robust scheme for construction good initial guess on the fine grid-levels.

3.2 SCF iteration using DIIS scheme

Recall that in the case of orthonormal basis, i.e., the overlap matrix equals the identity, the DIIS iteration is substituted by the commutation property, $[F(D), D] = 0$, on the exact solution. In the general case, the DIIS algorithm is based on the fact that the equation

$$F(D)DS - SDF(D) = 0$$

is the equivalent formulation of the initial Hartree-Fock Galerkin equation (2.7), see [26].

In the present paper, we use the original version of DIIS scheme (cf. [15]), defined by the following choice of the *residual error vectors* (matrices),

$$E_i := [\overline{C}_{i+1}^T F(C_i)\overline{C}_{i+1}]_{|\{1\le\mu\le N_{orb};N_{orb}+1\le\nu\le R_0\}} \in \mathbb{R}^{N_{orb}\times(R_0-N_{orb})}, \tag{3.3}$$

for iteration number $i = 0, 1, ..., k$, that should vanish on the exact solutions of the Hartree-Fock Galerkin equation due to the orthogonality property. Hence, some stopping criteria applies to residual error vector E_i for $i = 0, 1, 2, ...$

The minimizing coefficients vector $\overline{c} := (c_0, ..., c_k)^T \in \mathbb{R}^{k+1}$ is computed by solving the constrained quadratic minimization problem for the respective cost functional (the averaged residual error vector over previous iterands),

$$f(\overline{c}) := \frac{1}{2}\left\|\sum_{i=0}^{k} c_i E_i\right\|_F^2 \equiv \frac{1}{2}\langle B\overline{c},\overline{c}\rangle \to \min, \quad \text{provided that} \quad \sum_{i=0}^{k} c_i = 1,$$

where

$$B = \{B_{ij}\}_{i,j=0}^{k} \quad \text{with} \quad B_{ij} = \langle E_i, E_j\rangle,$$

and with E_i defined by (3.3). Introducing the Lagrange multiplier $\xi \in \mathbb{R}$, the problem is reduced to minimization of the Lagrangian functional

$$L(\overline{c},\xi) = f(\overline{c}) - \xi(\langle \mathbf{1},\overline{c}\rangle - 1),$$

where $\mathbf{1} = (1, ..., 1)^T \in \mathbb{R}^{k+1}$, that leads to the linear augmented system of equations

$$\begin{aligned} B\overline{c} - \xi\mathbf{1} &= 0, \\ \langle \mathbf{1},\overline{c}\rangle &= 1. \end{aligned} \tag{3.4}$$

Finally, the updated Fock operator $\overline{F}_k$ is built up by

$$\overline{F}_k = \sum_{i=0}^{k-1} c_i^{opt}\overline{F}_i + c_k^{opt}F(C_k), \quad k = 0, 1, 2, ..., \tag{3.5}$$

where the minimizing coefficients $c_i^{opt} = \overline{c}_i$ $(i = 0, 1, ..., k)$ solve the linear system (3.4). For $k = 0$ the first sum in (3.5) is assumed to be zero, hence providing $c_0^{opt} = 1$, and $\overline{F}_0 = F(C_0)$.

Recall that if the stopping criteria on C_k, $k = 1, ...,$ is not satisfied, one updates $\overline{F}_k$ by (3.5) and solves the eigenvalue problem (3.1) for C_{k+1}.

Note that in practice one can use the averaged residual vector only on a reduced subsequence of iterands, $E_k, E_{k-1}, ..., E_{k-k_0}$, with small $k_0 < k$. In our numerical examples below, we usually set $k_0 = 4$.

3.3 Tensor-truncated multilevel DIIS iteration

In this section, we describe the resultant numerical algorithm. Recall that the discrete nonlinear Fock operator is specified by a matrix

$$F(D) = H + J(D) + K(D), \tag{3.6}$$

where H corresponds to the core Hamiltonian (fixed in our scheme) and the discrete Hartree and exchange operators are given by tensor representations (2.16) and (2.19), respectively.

First, we describe the unigrid tensor-truncated DIIS scheme.

Algorithm TT_DIIS *(Unigrid tensor-truncated DIIS iteration).*

1. *Given the core Hamiltonian matrix H, the grid parameter n, and the termination parameter $\varepsilon > 0$.*

2. *Set $C_0 = 0$ (i.e. $J(C_0) = 0$, $K(C_0) = 0$), and $\overline{F}_0 = H$.*

3. *For $k = 0, 1, ...,$ perform*

a) Solve the full linear eigenvalue problem of size $R_0 \times R_0$, given by (3.2), and define C_{k+1} as a matrix containing the N_{orb} eigenvectors corresponding to N_{orb} minimal eigenvalues.

b) Terminate iteration by checking the stopping criteria

$$\|C_{k+1} - C_k\|_F \leq \varepsilon.$$

c) If $\|C_{k+1} - C_k\|_F > \varepsilon$, compute the Fock matrix

$$F(C_{k+1}) = H + J(C_{k+1}) + K(C_{k+1})$$

by the tensor-structured calculations of $J(C_{k+1})$ and $K(C_{k+1})$, using grid-based basis functions with expansion coefficients C_{k+1}, update the Fock matrix $\overline{F}_{k+1}$ by (3.5), and switch to Step a).

4. *Returns: Eigenvalues $\lambda_1, ..., \lambda_{N_{orb}}$ and eigenvectors $C \in \mathbb{R}^{R_0 \times N_{orb}}$.*

Figure 4.3 shows the convergence of Algorithm TT_DIIS for solution of the Hartree-Fock equation in the pseudopotential case of CH_4. It demonstrates that convergence history is almost independent on the grid size on the examples with $n = 64$ and $n = 256$, correspondingly.

To enhance the unigrid DIIS iteration, we propose the multilevel version of Algorithm TT_DIIS defined on a sequence of discrete Hartree-Fock equations specified by a sequence of grid parameters $n_p = n_0, 2n_0, \ldots, 2^M n_0$, with $p = 0, ..., M$, corresponding to the succession of dyadically refined spacial grids. To that end, for ease of exposition, we also introduce the incomplete version of Algorithm TT_DIIS, further called Algorithm TT_DIIS($\overline{k}$), which represents only its part starting from iteration number $k = \overline{k} \geq 1$. The input data for Algorithm TT_DIIS($\overline{k}$) include the current approximation $C_{\overline{k}}$ and a sequence of all already precomputed Fock matrices, $\overline{F}_0, \overline{F}_1, ..., \overline{F}_{\overline{k}-1}$.

We sketch this algorithm as follows.

Algorithm TT_DIIS($\overline{k}$) *(Incomplete unigrid tensor-truncated DIIS iteration).*

1. *Given the core Hamiltonian matrix H, the grid parameter n, the termination parameter $\varepsilon > 0$, $C_{\overline{k}}$, and a sequence of Fock matrices $\overline{F}_0, \overline{F}_1, ..., \overline{F}_{\overline{k}-1}$.*

2. *Compute $J(C_{\overline{k}})$, $K(C_{\overline{k}})$, $F(C_{\overline{k}}) = H + J(C_{\overline{k}}) + K(C_{\overline{k}})$, and $\overline{F}_{\overline{k}}$ by (3.5).*

3. *For $k = \overline{k}, \overline{k}+1, ...,$ perform steps a) - c) in Algorithm MTT_DIIS.*

We further assume that the core Hamiltonian H is precomputed beforehand.

Algorithm MTT_DIIS *(Multilevel tensor-truncated DIIS scheme).*
1. *Given the core Hamiltonian matrix H, the coarsest grid parameter n_0, the termination parameter $\varepsilon_0 > 0$, and the number of grid refinements M.*
2. *For $p = 0$, apply unigrid Algorithm TT_DIIS with $n = n_0$, $\varepsilon_p = \varepsilon_0$, and return the number of iterations k_0, matrix C_{k_0+1}, and a sequence of Fock matrices $\overline{F}_0, \overline{F}_1, ..., \overline{F}_{k_0}$.*
3. *For $p = 1, ..., M$, apply successively Algorithm TT_DIIS($k_{p-1} + 1$), with the input parameters $n_p := 2^p n_0$, $\varepsilon_p := \varepsilon_0 2^{-2p}$, $C_{k_{p-1}+1}$. Keep continuous numbering of the DIIS iterations through all levels, such that the maximal iteration number at level p is given by*

$$k_p = \sum_{p=0}^{p} m_p,$$

with m_p being the number of iterative steps at level p.
4. *Returns: k_M, C_{k_M+1}, and a sequence of Fock matrices $\overline{F}_0, \overline{F}_1, ..., \overline{F}_{k_M}$.*

In the present numerical examples, we start calculations on $n_0 \times n_0 \times n_0$ 3D Cartesian grid, with $n_0 = 64$, and end up with maximum $n_M = 8192$, for all electron case computations, or $n_M = 1024$, for the pseudopotential case. Further, in Section 4, we show by numerical examples that in large scale computations the multilevel Algorithm MTT_DIIS allows us to perform most of iterative steps on coarse grids thus reducing dramatically the computational cost and, at the same time providing a good initial guess for DIIS iteration on nonlinearity at each consequent approximation level.

3.4 Complexity estimate in terms of R_G, N_b and n

The rest of this section addresses the complexity estimate of the multilevel tensor-truncated iteration in terms of R_N, R_0, n and other governing parameters of the algorithm. For the ease of discussion we suppose that $rank(G_\mu) = 1$, $\mu = 1, ..., R_0$.

Lemma 3.1 *Let $rank(G_\mu) = 1$, $\mu = 1, \dots R_0$ and $rank(P_N) \le R_N \le CN_{orb}$. Suppose that the rank reduction procedure applied to the convolution products $\Upsilon_{a\nu}$ in (2.10) provides the rank estimate $rank(\Upsilon_{a\nu}) \le r_0$. Then the numerical cost of one iterative step in Algorithm MTT_DIIS at level p, can be bounded by*

$$W_p = O(R_0 R_N n_p \log n_p + R_0^3 r_0 N_{orb} n_p).$$

Assume that the number of multigrid DIIS iterations at each level is bounded by the unique constant I_0, then the total cost of Algorithm MTT_DIIS does not exceed the double cost at the finest level $n = n_M$, $2W_M = O(I_0 R_0^3 r_0 N_{orb} n)$.

Proof. The rank bound $rank(G_k) = 1$ implies $rank(\sum_{m=1}^{R_0} c_{ma} G_m) \le R_0$. Hence, the numerical cost to compute tensor-product convolution $\Upsilon_{a\nu}$ in (2.10) amounts to

$$W(\Upsilon_{a\nu}) = O(R_0 R_N n_p \log n_p).$$

Since the initial canonical rank of $\Upsilon_{a\nu}$ is estimated by $rank(\Upsilon_{a\nu}) \le R_0 R_N$, the multigrid rank reduction algorithm, having linear scaling in $rank(\Upsilon_{a\nu})$, see [23], provides the complexity

bound $O(r_0 R_0 R_N n_p)$. Hence the total cost to compute scalar products in $\chi_{\mu\nu,a}$ (see (2.11)) can be estimated by

$$W(\chi_{\mu\nu,a}) = O(R_0^3 r_0 N_{orb} n_p),$$

which completes the first part of our proof. The second assertion is due to linear scaling of the unigrid algorithm in n_p, that implies the following bound

$$n_0 + 2n_0 + ... + 2^p n_0 \leq 2^{p+1} n_0 = 2n_M,$$

hence, completing the prove. ■

Remark 3.2 *In the case of large molecules and $R_G = rank(G_\mu) \geq 1$, further optimization of the algorithm up to $O(R_N R_0^2 n_p)$-complexity may be possible on the base of rank reduction applied to the rank-$R_G R_0$ orbitals and by using an iterative eigenvalue solver instead of currently employed direct solver via matrix diagonalization. Alternatively, one can apply direct minimization schemes [32].*

4 Numerical illustrations

4.1 General discussion

Our algorithm for the *ab initio* solution of the Hartree-Fock equation in *tensor-structured format* is examined numerically on some moderate size molecules. In particular, we consider the all electron case of H_2O, and the pseudopotential case of CH_4 and CH_3OH molecules. In the present numerical examples the discrete GTO basis functions are used for the reasons of convenient comparison with the standard MOLPRO package based on the analytical evaluation of the integral operators in the GTO basis.

The size of the computational box introduced in §2.3 varies from $2b = 11.21$ $\mathring{A}$ and $2b = 15.45$ $\mathring{A}$ for H_2O, and CH_4 molecules, respectively. The smallest step-size of the grid $h = 0.0027$ $\mathring{A}$ is reached in the SCF iterations for H_2O molecule, using the finest grid level with $n = 4096$, while the average step size for the computations using the pseudopotentials of the moderate size molecules is about $h = 0.015$ $\mathring{A}$, corresponding to the grid size $n = 1024$.

Using the equidistant $n \times n \times n$ tensor grid given by (2.13), the functions in $\mathcal{V}_n$ of the type (2.14) are discretized in the intervals $[-b, b]$. In this way, we obtain the canonical representation of the electron densities and orbitals on 3D Cartesian grid, further applied for the tensor-structured calculation of the Hartree (2.2) and the nonlocal exchange (2.3) potentials. Notice, that the Galerkin representation of the exchange operator leads to the evaluation of the exchange integral term over $\mathbb{R}^3 \times \mathbb{R}^3$. Our tensor-structured techniques provide means for the discrete evaluation of this integral using the fast tensor-product convolution and other tensor-product operations (see Appendix), which are of almost linear complexity with respect to the one-dimension grid size n. The convolving Poisson kernel is effectively represented in the rank-R_N canonical format with the rank parameter in the range $20 \leq R_N \leq 30$, depending logarithmically on the univariate grid size n, $R_N = O(\log n)$.

4.2 Tensor computation of the Coulomb and exchange matrices

A detailed description of the tensor-product numerical schemes for the Coulomb and exchange matrices in the case of closed shell systems is presented in [23, 17]. In this section, we verify the performance of the tensor-structured computations of the Coulomb and exchange matrices by using the *solutions for the orbitals* of molecules, taking the expansion coefficients $C_{\mu i}$ for the GTO basis in (2.4) from the MOLPRO package.

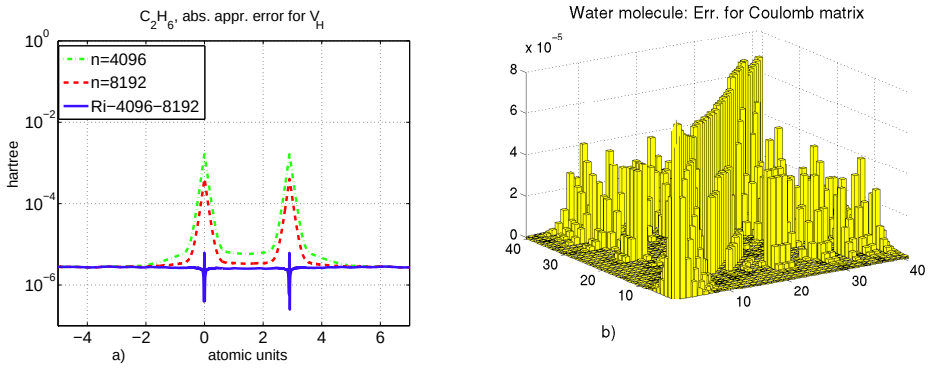

Figure 4.1: Left: Absolute approximation error (blue line: $\approx 10^{-6}$ a.u.) in the tensor-product computation of the Hartree potential of C_2H_6, measured in the grid line $\Omega = [-5, 7] \times \{0\} \times \{0\}$. Right: the absolute error for the Coulomb matrix of H_2O, ($\approx 10^{-5}$ a.u.).

The Coulomb (Galerkin) matrix of the Hartree potential, V_H, is computed by tensor inner products in $\{g_k\}$,

$$J_{km} := \int_{\mathbb{R}^3} g_k(x) V_H(x) g_m(x) dx, \quad k, m = 1, \dots N_b, \quad x \in \mathbb{R}^3.$$

Figure 4.1,left shows the absolute approximation error (blue line: $\approx 10^{-6}$ a.u.) in the tensor-product representation of the Hartree potential of C_2H_6 molecule, measured in the subinterval $\Omega = [-5, 7] \times \{0\} \times \{0\}$. Figure 4.1,right presents the absolute accuracy for the Coloumb matrix of H_2O, computed by tensor-structured techniques on the large grid with the one-dimension size $n = 8192$, providing the absolute accuracy $\approx 10^{-5}$. This corresponds to the approximation error $O(h^3)$ achieved on the grid with high resolution, $h \approx 0.0008$ $\overset{\circ}{A}$.

The exchange matrix in the Galerkin GTO basis is given by

$$K_{k,m} := -\frac{1}{2} \int_{\mathbb{R}^3} g_k(x) \frac{\tau(x, y)}{|x - y|} g_m(y) dx dy, \quad k, m = 1, \dots N_b.$$

Figure 4.2,right shows the L^∞-error in the matrix elements of K for the pseudodensity of CH_3OH computed on the grid with $n = 1024$. Figures 4.2,left illustrate the high accuracy achieved in the computation of the exchange matrix of H_2O molecule on the grids $n = 8192$, that corresponds to the fine step-size $h \approx 0.0016$ $\overset{\circ}{A}$, and the asymptotic approximation error $O(h^3)$, $h = 1/n$.

4.3 Multilevel tensor-truncated SCF iteration

The tensor-structured algorithms for calculation of the Coulomb and exchange parts of the Fock operator enable numerical solution of the *ab initio* Hartree-Fock equation, by using Algorithms TT_DIIS and MTT_DIIS in §3.3. Starting with the zero initial guess for matrices $J(D) = 0$ and $K(D) = 0$ in the Galerkin Fock matrix (2.6), we solve the eigenvalue problem at the first iterative step ($p = 0$) using only the H part of the Fock matrix in (2.6), that does not depend on the solution.

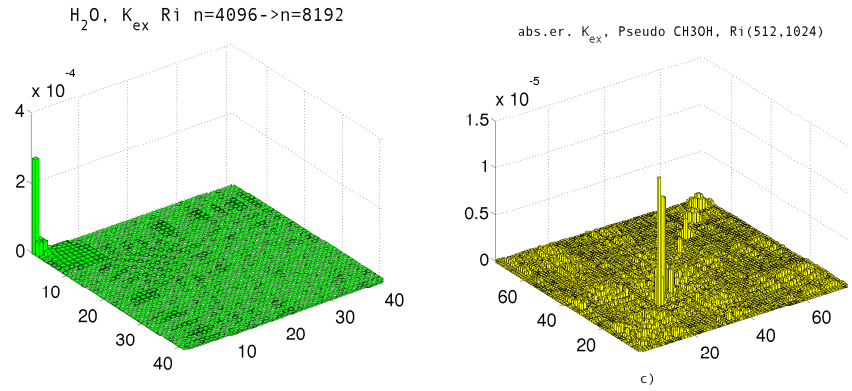

Figure 4.2: L^∞-error in $K_{ex} = K$ for the density of H_2 and pseudodensity of CH_3OH.

Thus, SCF iterations start with the expansion coefficients $C_{\mu i}$ for orbitals in the GTO basis, computed using only the core Hamiltonian $\mathcal{H}$. At every iteration step, the Hartree and exchange potentials and the corresponding Galerkin matrices, are computed using the updated expansions coefficients $C_{\mu i}$. The L^2-error in the correction term to the coefficient matrix C_k, is used for the convergence control. Figure 4.3 shows the convergence of the single-grid scheme for the solution of the Hartree-Fock equation in the pseudopotential case of CH_4.

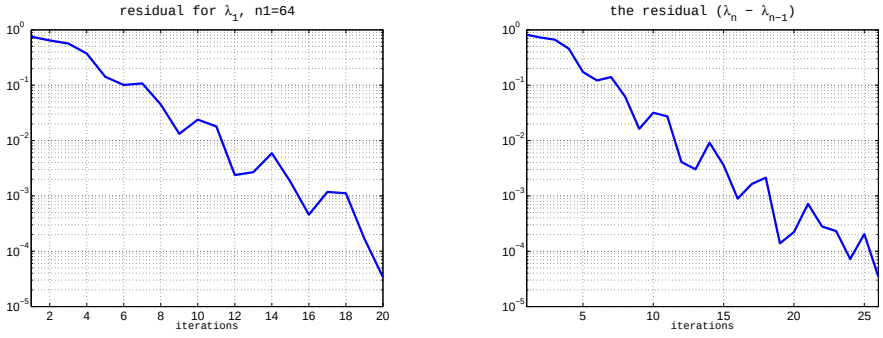

Figure 4.3: Convergence of the unigrid SCF iteration (Algorithm TT_DIIS) for the pseudopotentail case of the CH_4 molecule on the grid sizes $n = 64$ (left) and $n = 256$ (right).

The multilevel solution of the nonlinear eigenvalue problem (2.7) is realized via SCF iteration on a sequence of uniform grids, beginning from the initial coarse grid, say, with

$n_0 = 64$, and proceeding on the diadically refined grids, $n_p = n_0 2^p$, $p = 1, ..., M$. We use the grid dependent termination criteria $\varepsilon_{n_p} := \varepsilon_0 2^{-2p}$, keeping the continuous numbering of the iterations. Figure 4.4 (left) shows the convergence of the iterative scheme in the case of pseudopotential of CH_4. Convergence in the total Hartree-Fock energy reaching the absolute error $9 \cdot 10^{-6}$ a.u. on the grid size $n = 1024$ is shown in Figure 4.4 (right). The total energy is calculated by

$$E_{HF} = 2 \sum_{a=1}^{N} \lambda_a - \sum_{a=1}^{N} \left(\widetilde{J}_a - \widetilde{K}_a \right)$$

with $\widetilde{J}_a = (\psi_a, V_H \psi_a)_{L^2}$, and $\widetilde{K}_a = (\psi_a, \mathcal{V}_{ex} \psi_a)_{L^2}$, being the so-called Coulomb and exchange integrals, respectively, computed in the orbital basis ψ_a $(a = 1, ..., N)$.

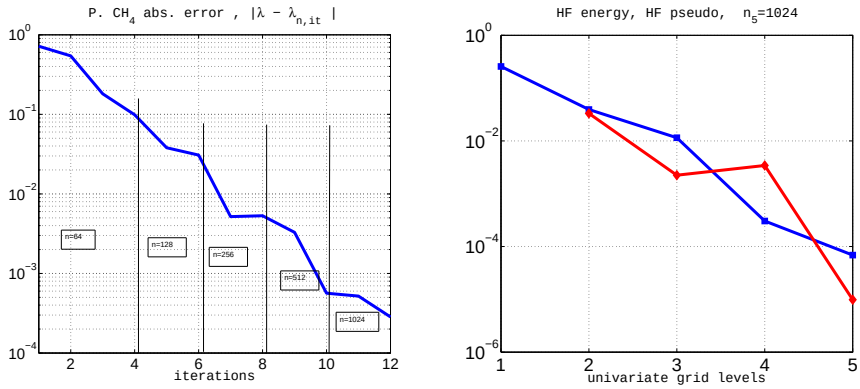

Figure 4.4: Multilevel convergence (left) for the pseudopotential of CH_4 molecule, and convergence of the HF energy in the grid levels (right). The initial level 1 corresponds to $n = 64$, while the final level 5 corresponds to $n = 1024$.

Figure 4.5 (left) shows the linear scaling in n, corresponding to CPU time at one iteration. Figure 4.5 (right) shows the number of "effective" iterations counted by rescaling the total computational time to the iteration time-unit observed at each iterative step at the finest grid-level.

Figure 4.6 represents the convergence history of the nonlinear iteration for CH_3OH measured by the iteration residual error (left), and by the energy error (right), respectively.

Figure 4.7 (left) shows convergence of the SCF iteration for all electron case of H_2O. This challenging problem is solved efficiently due to usage of large 3D Cartesian grids up to the volume size $N = 8192^3$. Figure 4.7 (right) shows convergence of the HF energy for the corresponding grid levels.

5 Conclusions and further perspectives

We present the grid-based tensor-truncated numerical method for the robust and accurate iterative solution of the self-consistent Hartree-Fock equation at the cost $O(n \log n)$, discretised over the 3D $n \times n \times n$ Cartesian grid. The computational scheme is based on the

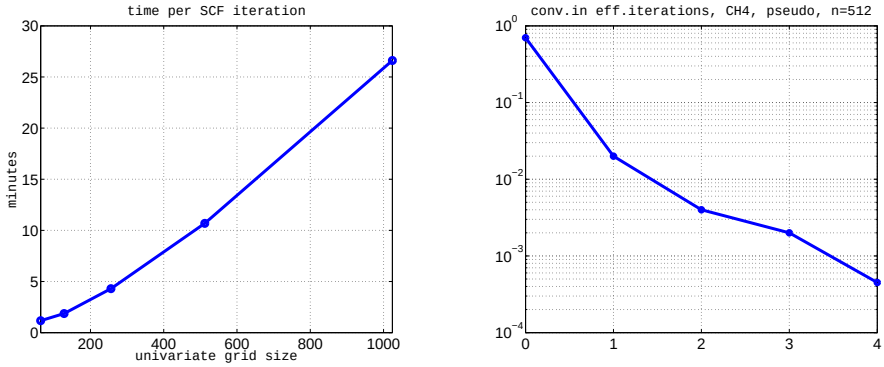

Figure 4.5: Linear scaling in n (left) and convergence in the effective iterations (right).

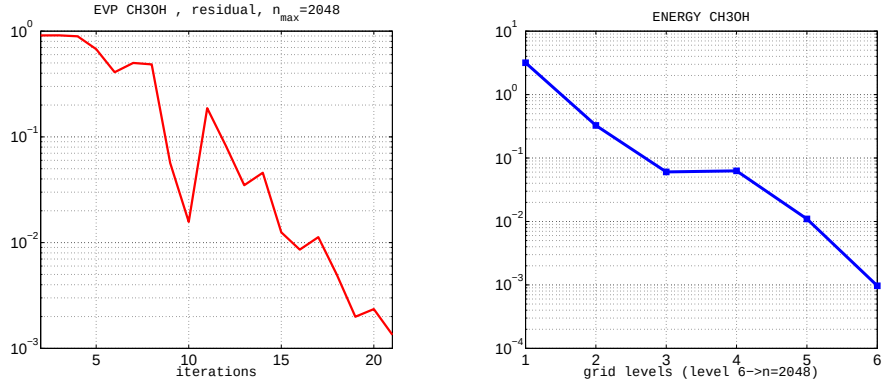

Figure 4.6: Residual (left) and the energy (right) iteration history for CH_3OH molecule.

discrete tensor representation of the Fock operator at each step of the multilevel SCF iteration applied to the nonlinear 3D eigenvalue problem. Numerical tests confirm the theoretical complexity estimates.

This scheme is neither restricted to the analytically separable basis functions like GTO orbitals nor to the traditional plane waves approximations. The Galerkin basis can be modified by adapting to the particular problem in the framework of the tensor-structured solution process.

The main computational blocks of the numerical scheme allow the natural parallelization on the level of matrix elements computation, rank decompositions, and the multilinear tensor operations.

6 Appendix: Description of tensor-structured formats

Let $\mathbb{H} = H_1 \otimes ... \otimes H_d$ be a tensor-product Hilbert space (see [31]), where H_ℓ ($\ell = 1, ..., d$) is a real, separable Hilbert space of functions of the continuous or discrete argument (say, $H_\ell = L^2([a, b])$ or $H_\ell = \mathbb{R}^n$). Each $w \in \mathbb{H}$ can be written as a sum of rank-1 (elementary,

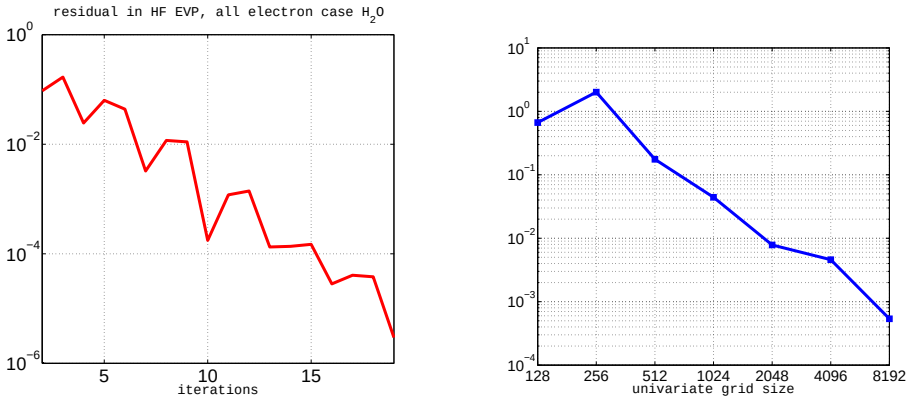

Figure 4.7: Multilevel convergence of the SCF iteration applied to the all electron case of H_2O (left), and convergence in the energy in n (right).

separable) tensors

$$w = \sum_k w_k^{(1)} \otimes w_k^{(2)} \otimes \ldots \otimes w_k^{(d)} \quad (w_k^{(\ell)} \in H_\ell).$$

The scalar product of rank-1 tensors in $\mathbb{H}$ is defined by

$$\langle w^{(1)} \otimes \ldots \otimes w^{(d)}, h^{(1)} \otimes \ldots \otimes h^{(d)} \rangle = \prod_{\ell=1}^{d} \langle w^{(\ell)}, h^{(\ell)} \rangle.$$

A dth order tensor is a function of d discrete arguments, $f : \mathbb{R}^{I_1 \times \ldots \times I_d} \to \mathbb{R}$, specified by a multi-dimensional array over $\mathcal{I} = I_1 \times ... \times I_d$, with $I_\ell = \{1, ..., n_\ell\}$, $\ell = 1, ..., d$. We write

$$V = [v_{i_1,\ldots,i_d} : i_\ell \in I_\ell] \in \mathbb{R}^{\mathcal{I}}, \quad \ell = 1, ..., d,$$

to denote a real-valued dth order tensor, that is an element of the tensor-product Hilbert space $\mathbb{H} := \mathbb{V}_{\mathbf{n}} = \otimes_{\ell=1}^{d} \mathbb{V}_\ell$, with $\mathbb{V}_\ell = \mathbb{R}^{I_\ell}$, and $\mathbf{n}$ being the d-tuple $(n_1, ..., n_d)$. $\mathbb{V}_{\mathbf{n}}$ is equipped with the Euclidean *inner product* $\langle \cdot, \cdot \rangle : \mathbb{V}_{\mathbf{n}} \times \mathbb{V}_{\mathbf{n}} \to \mathbb{R}$. Tensor $V \in \mathbb{V}_{\mathbf{n}}$ requires $\prod_{\ell=1}^{d} n_\ell$ reals for storage.

The concept of tensor methods is based on approximation of multivariate functions with low separation rank. In particular, we are interested in decomposition of a tensor $f \in \mathbb{V}_{\mathbf{n}}$, in the set of separable tensors, i.e., in some classes $\mathcal{S} \subset \mathbb{V}_{\mathbf{n}}$ of "rank structured" elements based on sums of rank-1 tensors. In this way, the canonical rank-1 tensor is represented by *outer product* of vectors $t^{(\ell)} = \{t_{i_\ell}^{(\ell)}\}_{i_\ell \in I_\ell} \in \mathbb{V}_\ell$ ($\ell = 1, ..., d$),

$$T \equiv [t_{\mathbf{i}}]_{\mathbf{i} \in \mathcal{I}} = t^{(1)} \otimes ... \otimes t^{(d)} \in \mathbb{V}_{\mathbf{n}} \quad \text{with entries} \quad t_{\mathbf{i}} = t_{i_1}^{(1)} \cdots t_{i_d}^{(d)},$$

requiring only $\sum_{\ell=1}^{d} n_\ell \ll \prod_{\ell=1}^{d} n_\ell$ reals to store it (now linear scaling in the dimension d).

In the present paper, we apply data sparse representation to high order tensors based on the Tucker and canonical models.

The rank-$(r_1, \ldots, r_d)$ Tucker approximation [5] is based on subspaces $\mathbb{T}_{\mathbf{n}} := \otimes_{\ell=1}^{d} \mathbb{T}_\ell$ of $\mathbb{V}_{\mathbf{n}}$ for certain $\mathbb{T}_\ell \subset \mathbb{V}_\ell$ with $r_\ell := \dim \mathbb{T}_\ell \ll n_\ell$. Given the rank parameter $\mathbf{r} = (r_1, ..., r_d)$,

we denote by $\boldsymbol{\mathcal{T}}_{\mathbf{r},\mathbf{n}}$ (or simply $\boldsymbol{\mathcal{T}}_{\mathbf{r}}$) the subset of tensors in $\mathbb{V}_{\mathbf{n}}$ represented in the so-called Tucker format

$$V_{(\mathbf{r})} = \sum\nolimits_{\nu_1=1}^{r_1} \dots \sum\nolimits_{\nu_d=1}^{r_d} \beta_{\nu_1,\dots,\nu_d}\, t^{(1)}_{\nu_1} \otimes \dots \otimes t^{(d)}_{\nu_d}, \tag{6.1}$$

with some vectors $t^{(\ell)}_{\nu_\ell} \in \mathbb{V}_\ell = \mathbb{R}^{I_\ell}$ $(1 \le \nu_\ell \le r_\ell)$, which form the orthonormal basis of $\mathbb{T}_\ell = \mathrm{span}\{t^{(\ell)}_\nu\}_{\nu=1}^{r_\ell}$ $(\ell = 1, ..., d)$. The coefficients tensor $\boldsymbol{\beta} = [\beta_{\nu_1,\dots,\nu_d}]$, that is an element of a dual tensor space $\mathbb{V}_{\mathbf{r}} = \mathbb{R}^{r_1 \times \dots \times r_d}$, is called the *core tensor*. The parameter $r = \max\limits_{\ell}\{r_\ell\}$ is called the maximal Tucker rank. In our applications, we normally have $r \ll n = \max n_\ell$, say $r = O(\log n)$.

Given a rank parameter $R \in \mathbb{N}$, we denote by $\boldsymbol{\mathcal{C}}_{R,\mathbf{n}} = \boldsymbol{\mathcal{C}}_R \subset \mathbb{V}_{\mathbf{n}}$ a set of tensors which can be represented in *the canonical format*

$$V_{(R)} = \sum\nolimits_{\nu=1}^{R} \mu_\nu u^{(1)}_\nu \otimes \dots \otimes u^{(d)}_\nu, \quad \mu_\nu \in \mathbb{R}, \tag{6.2}$$

with normalized vectors $u^{(\ell)}_\nu \in \mathbb{V}_\ell$ $(\ell = 1, ..., d)$. The minimal parameter R in (6.2) is called the rank (or canonical rank) of a tensor $V_{(R)}$.

The storage requirement for the Tucker (resp. canonical) decomposition is bounded by $r^d + drn$ (resp. $R + dRn$), where usually $r \ll R$.

The above defined classes of rank-structured tensors are being applied in our tensor-product approximation schemes.

It is worth to note that linear transforms of elements in tensor-structured representation are reduced to 1D-operations, that can be accomplished with the rank truncation. In particular, for tensors A_1, A_2 in the canonical format

$$A_1 = \sum_{k=1}^{R_1} c_k u^{(1)}_k \otimes \dots \otimes u^{(d)}_k, \quad A_2 = \sum_{m=1}^{R_2} b_m v^{(1)}_m \otimes \dots \otimes v^{(d)}_m,$$

we make use of the following operations:

1. Euclidean inner product (complexity $O(dR_1R_2n) \ll n^d$),

$$\langle A_1, A_2 \rangle := \sum_{k=1}^{R_1}\sum_{m=1}^{R_2} c_k b_m \prod_{\ell=1}^{d} \left\langle u^{(\ell)}_k, v^{(\ell)}_m \right\rangle.$$

2. Hadamard product (complexity $O(dR_1R_2n) \ll n^d$),

$$A_1 \odot A_2 := \sum_{k=1}^{R_1}\sum_{m=1}^{R_2} c_k b_m \left(u^{(1)}_k \odot v^{(1)}_m\right) \otimes \dots \otimes \left(u^{(d)}_k \odot v^{(d)}_m\right) \in \boldsymbol{\mathcal{C}}_{R_1R_2}.$$

3. Convolution of two $3rd$ order tensors A_1, A_2 (same for high order tensors),

$$A_1 * A_2 = \sum_{k=1}^{R_1}\sum_{m=1}^{R_2} c_k b_m (u^{(1)}_m * v^{(1)}_k) \otimes (u^{(2)}_m * v^{(2)}_k) \otimes (u^{(3)}_m * v^{(3)}_k) \in \boldsymbol{\mathcal{C}}_{R_1R_2},$$

with linear scaling in n, $O(R_1R_2n \log n) \ll n^3 \log n$ (corresponds to 3D FFT).

These basic properties lead to the linear scaling in n of tensor-structured multilinear algebra applied in the framework of tensor-truncated iteration.

Acknowledgements. The authors acknowledge Prof. Dr. R. Schneider (TU Berlin), and Prof. E. Tyrtyshnikov (INM RAS Moscow) for stimulating discussions on the topic.

References

[1] M. Barrault, E. Cancés, W. Hager and Le Bris, *Multilevel domain decomposition for electronic structure calculations.* J. Comput. Phys. **222**, 2007, 86-109.

[2] G. Beylkin and M.J. Mohlenkamp, *Algorithms for numerical analysis in high dimension.* SIAM J. Sci. Comput. **26** (6) (2005) 2133-2159.

[3] G. Beylkin, M.J. Mohlenkamp and F. Pérez, *Approximating a wavefunction as an unconstrained sum of Slater determinants.* Journal of Math. Phys., **49** 032107 (2008).

[4] E. Cancés and Le Bris, *On the convergence of SCF algorithms for the Hartree-Fock equations.* ESAIM: M2AN, vol. 34/4, 2000, 749-774.

[5] L. De Lathauwer, B. De Moor, J. Vandewalle, *On the best rank-*1 *and rank-*$(R_1, ..., R_N)$ *approximation of higher-order tensors.* SIAM J. Matrix Anal. Appl., 21 (2000) 1324-1342.

[6] H.-J. Flad, W. Hackbusch, B.N. Khoromskij, and R. Schneider, *Concept of data-sparse tensor-product approximation in many-particle modeling.* In: "Matrix Methods: Theory, Algorithms, Applications", V. Olshevsky and E. Tyrtyshnikov, eds., World Scientific Publishers, Singapoure, 2010 (313-347).

[7] H.-J. Flad, B.N. Khoromskij, D.V. Savostyanov, and E.E. Tyrtyshnikov: *Verification of the cross 3d algorithm on quantum chemistry data.* J. Numer. Anal. and Math. Modelling, 4 (2008), 1-16.

[8] I.P. Gavrilyuk, W. Hackbusch, and B.N. Khoromskij, *Tensor-product approximation to elliptic and parabolic solution operators in higher dimensions.* Computing **74** (2005), 131-157.

[9] L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman, R. Schneider, Daubechies wavelets as a basis set for density functional pseudopotential calculations, J. Chem. Phys. **129** (2008) 014109.

[10] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, D. C. Allan, First-principles computation of material properties: the ABINIT software project, Comput. Mater. Sci. **25** (2002) 478.

[11] M. Griebel and J. Hamaekers, *Sparse grids for the Schrödinger equation.* M2AN, 41 (2007), pp. 215-247.

[12] W. Hackbusch, B.N. Khoromskij and E. Tyrtyshnikov. *Approximate iteration for structured matrices.* Numer. Math., **109** (2008), 365-383.

[13] W. Hackbusch, B.N. Khoromskij, S. Sauter and E. Tyrtyshnikov. *Use of Tensor Formats in Elliptic Eigenvalue Problems.* Preprint 78, MPI MiS, Leipzig 2008. NLAA submitted.

[14] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan, and G. Beylkin: *Multiresolution quantum chemistry: Basic theory and initial applications.* J. of Chemical Physics, 121 (23): 11587-11598, 2004.

[15] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory.* Wiley, New York, 1999.

[16] F.L. Hitchcock, *The expression of a tensor or a polyadic as a sum of products.* J. Math. Phys., 6 (1927), 164-189.

[17] V. Khoromskaia, *Computation of the Hartree-Fock Exchange in the Tensor-structured Format.* Computational Methods in Applied Mathematics, Vol. 10(2010), No 2, pp.204-218.

[18] B.N. Khoromskij, *Structured Rank-$(r_1, ..., r_d)$ Decomposition of Function-related Tensors in $\mathbb{R}^d$.* Comp. Meth. in Applied Math., **6** (2006), 2, 194-220.

[19] B. N. Khoromskij, *On tensor approximation of Green iterations for Kohn-Sham equations.* Comp. and Visualization in Sci., **11** (2008) 259-271.

[20] B.N. Khoromskij, *Fast and Accurate Tensor Approximation of a Multivariate Convolution with Linear Scaling in Dimension.* J. of Comp. Appl. Math., **234** (2010) 3122-3139.

[21] B.N. Khoromskij, *Tensor-Structured Preconditioners and Approximate Inverse of Elliptic Operators in $\mathbb{R}^d$.* J. Constr. Approx. **30** (2009) 599-620.

[22] B.N. Khoromskij and V. Khoromskaia, *Low Rank Tucker Tensor Approximation to the Classical Potentials.* Central European J. of Math., **5**(3) 2007, 1-28.

[23] B.N. Khoromskij and V. Khoromskaia: *Multigrid Tensor Approximation of Function Related Arrays.* SIAM J. on Sci. Comp., **31**(4), 3002-3026 (2009).

[24] B.N. Khoromskij, V. Khoromskaia, S. R. Chinnamsetty, H.-J. Flad, *Tensor Decomposition in Electronic Structure Calculations on 3D Cartesian Grids.* J. of Comput. Phys. 228 (2009) 5749-5762.

[25] T. G. Kolda and B. W. Bader, *Tensor Decompositions and Applications. SIAM Review*, **51/3**, 2009 455-500.

[26] C. Le Bris, *Computational chemistry from the perspective of numerical analysis.* Acta Numerica (2005), 363 - 444.

[27] Ch. Lubich, *From quantum to calssical molecular dynamics: reduced models and numerical analysis.* Zurich Lectures in Advanced Mathematics, EMS, 2008.

[28] I.V. Oseledets, D.V. Savostyanov, and E.E. Tyrtyshnikov, *Tucker dimensionality reduction of three-dimensional arrays in linear time.* SIMAX, 30(3), 939-956 (2008).

[29] I.V. Oseledets, D.V. Savostyanov, and E.E. Tyrtyshnikov, *Cross approximation in electron density computations.* Numer. Lin. Alg. with Appl., 2009, submitted.

[30] P. Pulay, *Improved SCF convergence acceleration.* J. Comput. Chem. **3**, 556-560 (1982).

[31] M. Reed, and B. Simon: Functional analysis, Academic Press, 1972.

[32] R. Schneider, Th. Rohwedder, J. Blauert, A. Neelov, *Direct minimization for calculating invariant subspaces in density functional computations of the electronic structure*, Journal of Comp. Math. **27** (2009), no. 2-3, 360-387.

[33] L.R. Tucker, *Some mathematical notes on three-mode factor analysis.* Psychometrika 31 (1966) 279-311.

[34] H.-J. Werner, P.J. Knowles, et al., MOLPRO, Version 2002.10, A Package of Ab Initio Programs for Electronic Structure Calculations.

[35] H. Yserentant, *The hyperbolic cross space approximation of electronic wavefunctions.* Numer. Math., 105 (2007) 659-690.