Tensor decomposition in electronic structure calculations on 3D Cartesian grids

B.N. Khoromskij [a,*], V. Khoromskaia [a], S.R. Chinnamsetty [a], H.-J. Flad [b]

[a] Max-Planck-Institute for Mathematics in the Sciences, Inselstr. 22-26, D-04103 Leipzig, Germany
[b] TU Berlin, Sekretariat MA 3-3, Strasse des 17, Juni 136, D-10623 Berlin, Germany

ARTICLE INFO

ABSTRACT

In this paper, we investigate a novel approach based on the combination of Tucker-type and canonical tensor decomposition techniques for the efficient numerical approximation of functions and operators in electronic structure calculations. In particular, we study applicability of tensor approximations for the numerical solution of Hartree–Fock and Kohn–Sham equations on 3D Cartesian grids. We show that the orthogonal Tucker-type tensor approximation of electron density and Hartree potential of simple molecules leads to low tensor rank representations. This enables an efficient tensor-product convolution scheme for the computation of the Hartree potential using a collocation-type approximation via piecewise constant basis functions on a uniform $n \times n \times n$ grid. Combined with the Richardson extrapolation, our approach exhibits $O(h^3)$ convergence in the grid-size $h = O(n^{-1})$. Moreover, this requires $O(3rn + r^3)$ storage, where r denotes the Tucker rank of the electron density with $r = O(\log n)$, almost uniformly in n. For example, calculations of the Coulomb matrix and the Hartree–Fock energy for the CH_4 molecule, with a pseudo-potential on the C atom, achieved accuracies of the order of 10^{-6} hartree with a grid-size n of several hundreds. Since the tensor-product convolution in 3D is performed via 1D convolution transforms, our scheme markedly outperforms the 3D-FFT in both the computing time and storage requirements.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

A successful strategy to reduce the computational complexity in electronic structure calculations is based on a series of low-rank tensor-product decompositions ranging from many-electron wavefunctions and related quantities, like electron densities, up to Coulomb potentials representing inter-electronic interactions and external potentials due to atomic nuclei. In particular, we mention linear combinations of Slater determinants, *Gaussian-type orbital* (GTO) basis sets, and the commonly employed Gaussian-transform of the Coulomb potential for the evaluation of integrals involving GTO basis functions. We refer to [19,7,15] for a detailed exposition of these subjects.

Recently, the new approaches based on "optimal" tensor-product approximations (with respect to certain norms) have been discussed in the literature. Here and in the following we refer to "optimal" in the sense of the best possible separable approximation of a function

$$f(\mathbf{x}) \approx \sum_{k=1}^{R} h_{1,k}(x_1) h_{2,k}(x_2) \cdots h_{d,k}(x_d), \quad \mathbf{x} \in \mathbb{R}^d, \tag{1}$$

* Corresponding author. Tel.: +49 341 9959 766; fax: +49 341 9959 999.
  E-mail address: bokh@mis.mpg.de (B.N. Khoromskij).

for a given rank R, without any constraints imposed on the univariate components $h_{i,k}$ with $i = 1, \ldots, d$. This format has been frequently denoted as canonical decomposition in the literature. It is worth to note that the expression of a dth order tensor as a sum of decomposable, i.e., rank-1 tensors, is known in mathematics since pioneering works by Hitchcock [20,21] (see [5] for further references). It is well known long time ago that when $d > 2$, the best approximation in the form (1) might not exist. Hence, in computational practice, the approximate representation (1) is calculated up to certain termination criteria.

Initiated by the work of Beylkin and Mohlenkamp [3], "optimal" tensor-product approximations for various potentials and more general integral operators, with possible applications in quantum chemistry, have been discussed in the literature, see [18,4,12–14,27,25,28] and references therein.

We focus on the tensor-product approximation of electron densities and the computation of Hartree potentials which represents a major bottleneck for the numerical solution of *Hartree–Fock* and *Kohn–Sham* equations. Traditionally, GTO basis sets have been extensively used for this purpose in quantum chemistry. Approximation of the electron density in a so-called auxiliary GTO basis and subsequent analytic evaluation of the Hartree potential became known as density-fitting or resolution of the identity method, cf. Refs. [9,11,29], and turned out to be an essential improvement for the computational efficiency. Although such kind of approach utilizes the separability of GTO basis sets, it seems that possible generalizations concerning this issue are rarely discussed in the literature. An interesting exception is a paper by Almlöf [1] who considered possible tensor-product concepts for an efficient evaluation of the Fock matrix. Recently a density-fitting scheme based on "optimal" tensor-product approximations (1) has been studied in Ref. [6]. It turned out that significant improvements beyond GTO basis sets can be achieved, however the unconstraint optimization of univariate components represents a generically ill-posed problem and rather sophisticated optimization techniques are necessary in order to achieve the required high accuracies.

In this paper we systematically investigate a novel approach based on the combination of the orthogonal and canonical tensor decomposition techniques for the efficient numerical approximation of functions and operators in electronic structure calculations. The existence of low separation rank approximations for a class of analytic functions in $\mathbb{R}^d$ is proven in [13,24,25] (see also Section 3 below). In particular, these results apply to the classical kernels $1/\|x\|$, $e^{-\|x\|}$ and $e^{-\|x\|}/\|x\|$, $x \in \mathbb{R}^d$. The general concept of linear tensor-product operations in the orthogonal and canonical tensor formats have been first described in [27]. The main motivation in the low separation rank representation of multidimensional functions and operators is the tremendous reduction of numerical complexity from the linear in the volume size to the linear in the univariate problem size. However, the main difficulty in practical application of tensor methods in higher dimension is the arising nonlinear approximation problem. This opens a number of challenging mathematical problems to be solved in the future.

The orthogonal rank-$(r_1, \ldots, r_d)$ format (sometimes called the Tucker decomposition due to [32]) poses orthogonality constraints on the univariate components which turns the optimization into a well-posed problem. The orthogonal decomposition is traditionally applied in independent component analysis, where it was usually treated as a method with moderate accuracy requirements. The existence of such an approximation is proven in [24]. The commonly used numerical method to compute the orthogonal rank-$(r_1, \ldots, r_d)$ approximation is based on the alternating least squares (ALS) iteration [8]. The convergence theory is known for the class of rank-1 tensors, and in the case of orthogonally decomposable tensors [34]. Potential applications for large scale problems in quantum chemistry have been already shown recently in [24,27], where one can find a thorough description of tensor operations in the orthogonal rank-$(r_1, \ldots, r_d)$ format.

The main result of the paper is the efficient and accurate computation of the fully discrete representation to the Hartree potential on the Cartesian $n \times n \times n$ grid (further denoted by $n^{\otimes 3}$) via the fast tensor-product convolution [26,27] at the cost of $O(R_1 R_2 n)$ operations. Here R_1, R_2 denote the separation rank of the Coulomb potential and electron density, respectively. For a precise definition of these ranks we refer to Section 3. We want to mention that R_1, R_2 only weakly depend on the grid spacing and can be considered as almost uniformly bounded in the univariate problem size n, cf. our discussion in Section 2. It should be also noted that the input data for the discrete convolution product are usually given with an initial rank $R \gg R_1$. Applying the two-level scheme [27], we first compute its rank-r Tucker approximation and then compress the small size core tensor to the rank-R_1 canonical decomposition.

An essential new feature of this scheme is that convolutions require only 1D-FFTs (i.e. $O(n \log n)$ operations) instead of 3D-FFTs (i.e. $O(n^3 \log n)$ operations) as in traditional approaches.[1] Therefore, it is possible to perform convolutions involving tensor-product approximations with *much lower complexity* compared with the volume grid-size n^3 (see Remark 2.4). Furthermore, storage of the full 3D-grid is no longer required, which provides another potentially significant advantage with respect to the conventional methods of $O(n^3)$ complexity. The latter might become computationally infeasible just because of storage limitations for very large systems or high grid resolutions.

For simplicity we restricted ourselves to the Hartree–Fock method. We want to mention however that the techniques to be discussed in the following can be as well applied to Kohn–Sham equations in the density functional theory (DFT). The Hartree–Fock method provides a mean-field approximation for the ground state of many-electron systems. This implies the solution of a nonlinear eigenvalue problem

$$\left(-\frac{1}{2}\Delta + V_{nuc} + V_H - V_x\right)\psi_i(\mathbf{x}) = \epsilon_i \psi_i(\mathbf{x}),$$ (2)

---

[1] For example, a large variety of plane wave and wavelet based methods exhibiting linear scaling in the number $N_3 = n^3$ of grid points, see e.g. the *BigDFT* program [22,23].

for the $N/2$ lowest eigenvalues ϵ_i and spatial eigenfunctions ψ_i ($i = 1, \ldots, N/2$), in the case of a closed-shell N electron system. The Hartree–Fock equation (2) corresponds to a nonlinear single-particle Schrödinger equation in $\mathbb{R}^3$ where the potentials V_H and V_x represent a mean-field acting on a single electron which is generated by the remaining $N - 1$ electrons in the system. Furthermore, an external potential V_{nuc} contains the bare Coulomb- or pseudopotentials of the nuclei. Within the present work, we focus on the computation of the Hartree potential

$$V_H(\mathbf{x}) := \int_{\mathbb{R}^3} \frac{\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \, d\mathbf{y}, \tag{3}$$

which corresponds to a convolution of the Coulomb potential with the electron density

$$\rho(\mathbf{y}) = 2 \sum_{b=1}^{N/2} \psi_b(\mathbf{y}) \psi_b^*(\mathbf{y}). \tag{4}$$

The remaining term is the nonlocal exchange operator

$$(V_x \psi_a)(\mathbf{x}) := \sum_{b=1}^{N/2} \int_{\mathbb{R}^3} \frac{\psi_a(\mathbf{y}) \psi_b^*(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \psi_b(\mathbf{x}) \, d\mathbf{y}, \tag{5}$$

which has to be replaced by a local exchange-correlation potential in the Kohn–Sham equations.

The paper is organized as follows. In Section 2, we introduce the orthogonal Tucker decomposition and study the corresponding approximations for all-electron densities of several simple molecules in order to demonstrate exponential convergence with respect to the tensor rank. In Section 3, using *sinc* quadrature based approximations to the Coulomb potential [13,17,24], we consider the computation of Hartree potentials within a fully discrete tensor-product collocation scheme (cf. [26]) on a 3D Cartesian grid. A numerical example is provided for the pseudodensities of the CH_4 molecule, obtained by using a pseudopotential at the C atom. Finally, in Section 4 we recover the Galerkin matrix of the approximate Hartree potential in a GTO basis and calculate the corresponding errors in the Hartree–Fock energy.

2. Tucker-type approximation of electron densities

2.1. General definitions

We introduce the linear space of real-valued dth order tensors

$$\mathcal{A} = [a_{i_1 \cdots i_d}] \in \mathbb{R}^{\mathcal{I}}, \quad \mathcal{I} = I_1 \times \cdots \times I_d, \quad I_\ell = \{1, \ldots, n_\ell\}, \quad \ell = 1, \ldots, d,$$

with the *scalar product* defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle := \sum_{(i_1, \ldots, i_d) \in \mathcal{I}} a_{i_1 \cdots i_d} b_{i_1 \cdots i_d} \quad \text{for } \mathcal{A}, \mathcal{B} \in \mathbb{R}^{\mathcal{I}}. \tag{6}$$

The corresponding Euclidean norm is given by $\|\mathcal{A}\| := \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$.

For given vectors $U^{(\ell)} \in \mathbb{R}^{I_\ell}$ with components $U^{(\ell)} = \left\{ u_{i_\ell}^{(\ell)} \right\}_{i_\ell \in I_\ell}$, we use the notion of the outer product "$\otimes$" of vectors, which form the canonical rank-1 tensor

$$\mathcal{U} = U^{(1)} \otimes \cdots \otimes U^{(d)} \in \mathbb{R}^{\mathcal{I}},$$

defined by entries

$$u_{i_1 \cdots i_d} = u_{i_1}^{(1)} \cdots u_{i_d}^{(d)}.$$

In the *orthogonal Tucker decomposition*, we approximate any tensor $\mathcal{A} \in \mathbb{R}^{\mathcal{I}}$ up to a given accuracy $\varepsilon > 0$, by a tensor $\mathcal{A}_{(\mathbf{r})}$ with $\mathbf{r} = (r_1, \ldots, r_d)$ defined by

$$\mathcal{A}_{(\mathbf{r})} = \sum_{k_1=1}^{r_1} \cdots \sum_{k_d=1}^{r_d} b_{k_1 \cdots k_d} \cdot V_{k_1}^{(1)} \otimes \cdots \otimes V_{k_d}^{(d)} \approx \mathcal{A}, \quad b_{k_1 \cdots k_d} \in \mathbb{R}, \tag{7}$$

where $V_{k_\ell}^{(\ell)} \in \mathbb{R}^{I_\ell}$ ($\ell = 1, \ldots, d$), and $\mathcal{B} = [b_{\mathbf{k}}] \in \mathbb{R}^{r_1 \times \cdots \times r_d}$ is called the core tensor. Here $\left\{ V_{k_\ell}^{(\ell)}, \ 1 \leqslant k_\ell \leqslant r_\ell \right\}$ are linear independent vectors for each fixed ℓ, hence without loss of generality, we assume that they are mutually orthogonal, i.e., $\mathbf{V}^{(\ell)} = \left[V_1^{(\ell)} V_2^{(\ell)} \cdots V_{r_\ell}^{(\ell)} \right] \in \mathbb{R}^{n_\ell \times r_\ell}$ are orthogonal matrices.

The representation (7) means that for given set of rank-parameters $r_\ell \leqslant n_\ell$ ($\ell = 1, \ldots, d$), any tensor can be decomposed approximately into $r_1 r_2 \cdots r_d$ outer products (respectively, r^d if $r_1 = \cdots = r_d = r$), which requires only $r_1 + \cdots + r_d$ canonical components $V_{k_\ell}^{(\ell)}$ of size n_ℓ and r^d coefficients $\{b_{\mathbf{k}}\}$ for its representation. The important parameter

$$r = \max_\ell \{r_\ell\}$$

is called the Tucker rank. To make the approximation numerically efficient, we have to satisfy the condition $r_\ell \ll n_\ell$ ($\ell = 1, \ldots, d$), which is usually the case in electronic structure calculations.

The *canonical decomposition* (CANDECOMP) or *parallel factors* (PARAFAC) model (shortly, the canonical model) represents a tensor by a sum of rank-1 non-orthogonal components as follows

$$\mathcal{A}_{(R)} = \sum_{k=1}^{R} c_k \cdot U_k^{(1)} \otimes \cdots \otimes U_k^{(d)}, \quad c_k \in \mathbb{R}, \tag{8}$$

with normalised components $U_k^{(\ell)} \in \mathbb{R}^{I_\ell}$ ($\ell = 1, \ldots, d$). The minimal parameter R in the above representation is called the rank (or canonical rank) of a tensor.

2.2. On theoretical validation of the tensor-product approximation

The existence of low-rank orthogonal and canonical approximations in the form (7) or (8) can be proven for the class of function related tensors [13,24,17,25]. The results directly apply to the convolving kernel defined by the Coulomb potential $\frac{1}{\|x\|}$, the Yukawa potential $\frac{e^{-\mu\|x\|}}{\|x\|}$, as well as to the functions obtained as a finite sum $\sum_{k=1}^{M} e^{-\lambda_k\|x\|}$ (represents the typical singularities of the electron densities). In electronic structure calculations the tensor decomposition of electron densities can be computed by algebraic nonlinear approximation methods.

In the following, we discuss the existence of efficient tensor decompositions for the class of function related tensors (FRT) obtained by the discretisation of multivariate function g on tensor-product grid. The approximated function is supposed to be analytic with the exception of few singularity points. For $p = 1, 2$, we define the hypercube $\Omega := \Pi^d \subset \mathbb{R}^{\mathbb{d}}$ with $\Pi := [a_0, b_0]^p$, $a_0, b_0 \in \mathbb{R}$, where $\mathbf{d} = dp$. In the case $p = 1$ we consider the functions $g : \Omega \to \mathbb{R}$ (say, density functions), while the choice $p = 2$ corresponds to $g : \Omega \times \Omega \to \mathbb{R}$ (kernel functions). To fix the idea, we consider the case $p = 2$.

Let ω_ℓ ($\ell = 1, \ldots, d$) be a tensor-product grid of uniform meshes on $\Pi := [a_0, b_0]^2$, indexed by $\mathcal{I}_\ell = I_{\ell,1} \times I_{\ell,2}$ with $\mathcal{I}_\ell$ being the product index set such that for $\mathbf{i}_\ell = (i_{\ell,1}, i_{\ell,2}) \in \mathcal{I}_\ell$, we have $I_{\ell,1}, I_{\ell,2} = I_n := \{1, \ldots, n\}$ ($\ell = 1, \ldots, d$). Furthermore, let $\omega_{\mathbf{d}} := \omega_1 \times \cdots \times \omega_d$ be the corresponding tensor-product lattice in the hypercube $\Omega := \Pi^d$. We denote by $\left\{ x_{\mathbf{i}_1}^{(1)}, \ldots, x_{\mathbf{i}_d}^{(d)} \right\}$ with $\mathbf{i}_\ell \in \mathcal{I}_\ell$ ($\ell = 1, \ldots, d$) a set of cell-centered collocation points living on the tensor-product lattice Ω_{col}.

Notice that the matrix decomposition corresponds to the choice $p = 2$. In this case we introduce the reordered index set of pairs $\mathcal{M}_\ell := \{\mathbf{m}_\ell : \mathbf{m}_\ell = (i_\ell, j_\ell), \ i_\ell, j_\ell \in I_n\}$ ($\ell = 1, \ldots, d$), so that $\mathcal{I} = \mathcal{M}_1 \times \cdots \times \mathcal{M}_d$ with $\mathcal{M}_\ell = I_n \times I_n$.

The Nyström and Galerkin approximations to function related tensors were described in [13,24]. In the following, we focus on the collocation-type schemes (see [24,17,25]), which are based on tensor-product ansatz functions

$$\psi^{\mathbf{j}}(y_1, \ldots, y_d) = \prod_{\ell=1}^{d} \psi_\ell^{j_\ell}(y_\ell), \quad \mathbf{j} = (j_1, \ldots, j_d) \in \mathcal{I}_1 \times \cdots \times \mathcal{I}_d. \tag{9}$$

Definition 2.1. Given the tensor-product basis set (9), we introduce the variable $\zeta_{i_\ell}^{(\ell)} := (x_{i_\ell}^{(\ell)}, y_\ell)$ with the collocation point $x_{i_\ell}^{(\ell)}$ and $y_\ell \in [a_0, b_0]$, the pair $\mathbf{m}_\ell := (i_\ell, j_\ell) \in \mathcal{M}_\ell$ and define the collocation-type dth order FRT by $\mathcal{A} \equiv \mathcal{A}(g) := [a_{\mathbf{m}_1 \cdots \mathbf{m}_d}] \in \mathbb{R}_1^{\mathcal{M}} \times \cdots \times \mathcal{M}_d$ with

$$a_{\mathbf{m}_1 \cdots \mathbf{m}_d} := \int_{\Omega} g(\zeta_{i_1}^{(1)}, \ldots, \zeta_{i_d}^{(d)}) \psi^{\mathbf{j}}(y_1, \ldots, y_d) \, dy, \quad \mathbf{m}_\ell = (i_\ell, j_\ell) \in \mathcal{M}_\ell. \tag{10}$$

In numerical calculations involving integral operators (e.g., arising in classical potential theory or from the Hartree–Fock, Kohn–Sham and Boltzmann equations), n may vary from several hundreds to several thousands, therefore, for $d \geqslant 2$, a naive "entry-wise" representation to the fully-populated tensor (matrix) $\mathcal{A}$ in (10) amounts to substantial computer resources, at least of the order $\mathcal{O}(n^{dp})$.

We consider a class of multi-variate functions $g : \mathbb{R}^d \to \mathbb{R}$ parametrised by $g(\zeta) = G(\rho(\zeta)) \equiv G(\rho)$ with $\rho \equiv \rho(\zeta) = \rho_0(\zeta^{(1)}) + \cdots + \rho_0(\zeta^{(d)}) > 0$, $\rho_0 : \mathbb{R}^2 \to \mathbb{R}_+$, corresponding to $p = 2$, $\zeta^{(\ell)} = (x_\ell, y_\ell)$ (cf. Definition 2.1). Here the univariate function $G : \mathbb{R}_+ \to \mathbb{R}$ can be represented via the Laplace transform

$$G(\rho) = \int_{\mathbb{R}_+} \mathcal{G}(\tau) e^{-\rho\tau} \, d\tau.$$

Without loss of generality, we introduce one and the same scaling function

$$\phi^i(\cdot) = \phi(\cdot + (i-1)h), \quad i \in I_n, \tag{11}$$

for all spatial dimensions $\ell = 1, \ldots, d$, where $h > 0$ is the mesh parameter.

For $i \in I_n$, let $\{\bar{x}_i\} \in \Omega_{col}$ be the set of cell-centred collocation points on $[a, b]$. For each $(i, j) \in I_n \times I_n$, we introduce the parameter dependent integral

$$\Psi_{i,j}(\tau) := \int_{\mathbb{R}^2} e^{-\rho_0(\bar{x}_i, y)\tau} \phi(y + (j-1)h) \, dy, \quad \tau \geqslant 0. \tag{12}$$

Theorem 2.2 ([24,17], Canonical approximation). *Assume that:*

(a) $\mathcal{G}(\tau)$ *has an analytic extension* $\mathcal{G}(w)$, $w \in \Omega_{\mathcal{G}}$, *into a certain domain* $\Omega_{\mathcal{G}} \subset \mathbb{C}$ *which can be mapped conformally onto the strip* $D_\delta := \{z \in \mathbb{C} : |\mathfrak{Im}z| < \delta\}$, *such that* $w = \varphi(z)$, $z \in D_\delta$ *and* $\varphi^{-1} : \Omega_{\mathcal{G}} \to D_\delta$;
(b) *for all* $(\mathbf{i},\mathbf{j}) \in \mathcal{I} \times \mathcal{J}$ *the transformed integrand*

$$f(z) := \varphi'(z)\mathcal{G}(\varphi(z)) \prod_{\ell=1}^{d} \Psi_{i_\ell j_\ell}(\varphi(z)), \tag{13}$$

belongs to the Hardy space $H^1(D_\delta)$;
(c) *the function* $f(t)$, $t \in \mathbb{R}$, *in* (13) *has either exponential (c1) or hyper-exponential (c2) decay as* $t \to \pm\infty$.

Then, for each $M \in \mathbb{N}_+$, *the FGT(C)*, $\mathcal{A}(g)$, *defined on* $[a,b]^d$ *allows an exponentially convergent symmetric*[2] *rank-r canonical decomposition* $\mathcal{A}_{(r)}$, *such that we have*

$$\|\mathcal{A}(g) - \mathcal{A}_{(r)}\|_\infty \leqslant Ce^{-\alpha M^\nu} \quad \text{with } r = 2M + 1, \tag{14}$$

where $\nu = \frac{1}{2}$, $\alpha = \sqrt{2\pi\delta b}$ *in the case (c1) and with* $\nu = 1$, $\alpha = \frac{2\pi\delta b}{\log(2\pi aM/b)}$ *in the case (c2).*

Let x, $y \in \mathbb{R}^d$, $p = 2$, and define $\rho = \|x - y\|^2 = \zeta_1^2 + \cdots + \zeta_d^2$ with $\zeta_\ell = x_\ell - y_\ell : \mathbb{R}^2 \to \mathbb{R}$. Theorem 2.2 can be applied to the family of functions

$$g(x,y) \equiv g(\zeta) := 1/\rho^\lambda \quad \text{with } \lambda \in \mathbb{R}_{>0}, \quad \zeta \in \mathbb{R}^d \times \mathbb{R}^d,$$

which arises in potential theory, in quantum chemistry and in computational gas dynamics. The choice $\lambda = 1/2$ corresponds to the classical Newton potential, while $\lambda = -1/2$ refers to the Euclidean distance function. Low separation rank decomposition to the multi-variate functions $1/\rho$, $1/\sqrt{\rho}$ and to the related Galerkin approximations were discussed in [13,14,24], while the kernel function $e^{-\mu\sqrt{\rho}}/\sqrt{\rho}$, $\mu \in \mathbb{R}$, was analysed in [25].

Let us take a closer look to the collocation-type FRT corresponding to the Newton potential $1/\sqrt{\rho}$ in the hypercube $[-R,R]^3 \in \mathbb{R}^3$. As a basic example, we consider piecewise constant finite elements on the uniform grid with step-size $h > 0$, defined by scaling functions $\phi(x) = \psi(x)$ associated with a tensor-product grid. Again, we let $\{\bar{x}_i\}$ be the set of cell-centred collocation points.

In our case, for the function in (12) we have $\rho_0(x,y) = (x - y)^2$ $(x,y \in \mathbb{R})$, hence, setting $d = 3$ and making use of the Gaussian-transform

$$G(\rho) \equiv \frac{1}{\sqrt{\rho}} = \frac{2}{\sqrt{\pi}} \int_{\mathbb{R}_+} e^{-\rho\tau^2}\,d\tau,$$

we obtain

$$\Psi_{i,j}(\tau) = \Psi_{|i-j|}(\tau) := \int_{\mathbb{R}^2} e^{-\tau^2(\bar{x}_i-y)^2}\phi^j(y)\,dy, \quad \tau \geqslant 0, \quad i,j \in I_n$$

(see (12) and (11) for the definition of $\Psi_{i,j}$, ϕ^j). Theorem 2.2 now implies the following result.

Lemma 2.3. *The FRT for the Newton potential* $1/\sqrt{\rho}$ *for* $d = 3$, *allows a canonical approximation in any finite hypercube* $[-A,A]^3 \times [-A,A]^3 \in \mathbb{R}^3 \times \mathbb{R}^3$ *with exponential convergence rate as in* (14), *where* $\nu = 1/2$.

2.3. Rank structured approximation of the electron densities

Below we consider approximating properties of the orthogonal Tucker-type representations (7) in $\mathbb{R}^3$ for electron densities of some simple molecules. We apply the ALS iterative scheme described in [8,27,28] to compute the low-rank tri-linear approximations for electron densities of the H atom, LiH, CH_4, C_2H_6 and H_2O molecules.

Remark 2.4. The Tucker model applied to a fully populated tensor of size $n^{\otimes 3}$ has the complexity of order $O(n^4)$ (storage and computational time). If the input tensor is already presented in the rank-R canonical format then the corresponding canonical-to-Tucker ALS method with the Tucker rank r has the complexity in the range $O(rRn) \div O(r^2Rn)$, depending on the problem setting (see Section 2.4.2 in [27,28]). Notice that the so-called cross approximation method of complexity $O(r^3n)$ was recently proposed in [30].

In Sections 3 and 4 both the Tucker-type and canonical decompositions are used in the fast computation of the convolution product in $\mathbb{R}^3$ appearing as the main ingredient in the Fock operator.

---

[2] A d th order tensor is called symmetric if it is invariant under arbitrary permutations of indices in $\{1,\ldots,d\}$.

The reference orbitals and electron densities of the mentioned molecules are given in the form of an expansion by "Cartesian Gaussian" basis functions, i.e.,

$$\psi(\mathbf{x}) = \sum_{k=1}^{R_0} \mu_k (x_1 - A_{1,k})^{\ell_k} (x_2 - A_{2,k})^{m_k} (x_3 - A_{3,k})^{n_k} \exp(-\alpha_k (\mathbf{x} - \mathbf{A}_k)^2), \tag{15}$$

with

$$(\mathbf{x} - \mathbf{A}_k)^2 = (x_1 - A_{2,k})^2 + (x_2 - A_{2,k})^2 + (x_3 - A_{3,k})^2, \quad \mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3,$$

where the parameters of the representation are taken from the standard quantum chemistry package MOLPRO [33] (cf. Ref [10]).

The electron densities of the considered molecules are represented in the form

$$\rho(\mathbf{x}) := \sum_{a=1}^{K} \psi_a^2(\mathbf{x}) = \sum_{a=1}^{K} \left(\sum_{k=1}^{R_0} P_{k,a}(\mathbf{x}) e^{-\alpha_k (\mathbf{x} - \mathbf{A}_k)^2} \right)^2, \tag{16}$$

where K is the number of electron pairs and R_0 is the number of GTO basis functions given by expansion (15) for orbitals $\psi_a(\mathbf{x})$.

Recall that in the case of the H atom the orbital and electron density are given explicitly by

$$\phi(\mathbf{x}) = C_\alpha \exp(-\alpha |\mathbf{x}|) \quad \text{with } \mathbf{x} \in \mathbb{R}^3,$$

with $\alpha = 1$ for the orbital and $\alpha = 2$ for the electron density.

We assume that any particular molecule is embedded in a certain fixed computational box $[-A, A]^3$ with a suitable $A > 0$. Let $\omega_3 \subset [-A, A]^3$ be a uniform $n \times n \times n$ tensor-product grid indexed by $\mathcal{I} = I_1 \times I_2 \times I_3$. For each particular molecule we use the following physically relevant parameters: $A = 10$ bohr , $R_0 = 10$ for H atom, $A = 7$ bohr, $R_0 = 34$ for LiH, $A = 5$ bohr, $R_0 = 55$ for CH_4, $A = 5.8$ bohr, $R_0 = 96$ for C_2H_6 and $A = 10$ bohr, $R_0 = 41$ for H_2O.

For a given continuous function $g : [-A, A]^3 \to \mathbb{R}$, we introduce the collocation-type function related tensor of order 3 by

$$\mathcal{A}_0 \equiv \mathcal{A}(g) := [a_{i_1 i_2 i_3}]_{(i_1, i_2, i_3) \in \mathcal{I}} \in \mathbb{R}^{n \times n \times n} \quad \text{with } a_{i_1 i_2 i_3} := g\left(x_{i_1}^{(1)}, x_{i_2}^{(2)}, x_{i_3}^{(3)} \right),$$

where $\left(x_{i_1}^{(1)}, x_{i_2}^{(2)}, x_{i_3}^{(3)} \right) \in \omega_3$ are the grid collocation points. In our applications, the function g is defined by the electron density ρ as in (16). Then tensor $\mathcal{A}_0$ is approximated by a rank $\mathbf{r} = (r, r, r)$ Tucker-type representation for a sequence of rank-parameters $r = 1, 2, \ldots, r_0$. The orthogonal vectors and the corresponding core tensor of the size $r^{\otimes 3}$ are then used for the reconstruction of the approximating tensor $\mathcal{A}_{(\mathbf{r})} \approx \mathcal{A}_0$. All numerical multi-linear approximations are performed in MATLAB, release 7.3.

For every rank-$\mathbf{r}$ Tucker-type approximation, we compute the relative error with respect to the Euclidean norm, as well as the relative maximum error:

$$E_{FN}^{(r)} = \frac{\|\mathcal{A}_0 - \mathcal{A}_{(\mathbf{r})}\|}{\|\mathcal{A}_0\|}, \quad E_C^{(r)} := \frac{\max\limits_{i \in \mathcal{I}} |a_{0,i} - a_{r,i}|}{\max\limits_{i \in \mathcal{I}} |a_{0,i}|}. \tag{17}$$

The relative difference of the norms denoted by

$$E_{FE}^{(r)} = \frac{\|\mathcal{A}_0\| - \|\mathcal{A}_{(\mathbf{r})}\|}{\|\mathcal{A}_0\|} \leqslant \frac{\|\mathcal{A}_0 - \mathcal{A}_{(\mathbf{r})}\|}{\|\mathcal{A}_0\|},$$

can be used as a stopping criteria (see [27] for the proof). The ALS iteration is terminated if the value of $E_{FE}^{(r)}$ reaches the machine precision.

Fig. 1 gives the approximation error with respect to the Tucker rank r for the LiH electron densities computed by (4) and (15) for different grid sizes $n^{\otimes 3}$. We see that for a fixed approximation error the Tucker rank remains to be almost independent on the univariate problem size n. The approximation errors shown in Fig. 2 verify exponential convergence of the orthogonal Tucker approximation (in the rank-parameter r) of electron densities reaching the relative accuracy $\sim 10^{-5}$ for CH_4, H_2O and C_2H_6 with $r = 16$.

The examples of orthogonal vectors $V_k^{(\ell)}$ ($k = 1, \ldots, 4$, $\ell = 1, 3$), of the tensor-product decomposition for the H atom, LiH and C_2H_6 molecules shown in Figs. 3 and 4 resemble the shape of the decomposed electron density along the corresponding spatial axis. Due to orthogonality of the decomposition, the Tucker model appears to be suitable for constructing low-dimensional problem-dependent orthogonal basis. Entries of the core tensor presented in Fig. 4 are the weights $b_{k_1 k_2 k_3}$ of the corresponding outer products of orthogonal vectors $V_{k_1}^{(1)} \otimes V_{k_2}^{(2)} \otimes V_{k_3}^{(3)}$, which compose the summands of $\mathcal{A}_{(\mathbf{r})}$ in (7) for $k_\ell \leqslant 6$.

Numerical experiments demonstrate high approximation power of the problem-dependent orthogonal tensor-product basis. For example, let $\mathcal{B}_{10}$ be the core tensor of the rank $r = 10$ Tucker-type approximation of the electron density for the CH_4 molecule, while $\widetilde{\mathcal{B}}_p$ represents the principal $p \times p \times p$ subtensor of $\mathcal{B}_{10}$. We observe that the contribution of the subspace $span\left\{ V_k^{(\ell)} \right\}_{k=1}^{p}$ ($\ell = 1, 2, 3$) to the relative Euclidean norm $\frac{\|\mathcal{B}_{10} - \mathcal{B}_p\|}{\|\mathcal{B}_{10}\|}$, is significant even for small p. In fact, already with $p = 3$, the
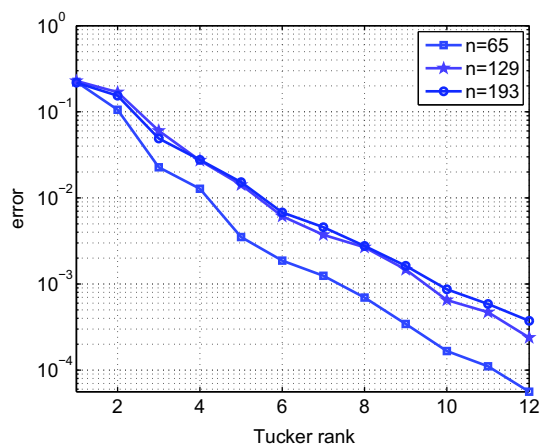
Fig. 1. Convergence of the Tucker approximation for the electron density of LiH with respect to the E_{FN} norm for different grid sizes $n = 65, 129$ and 193 and fixed box size $A = 7$ bohr.
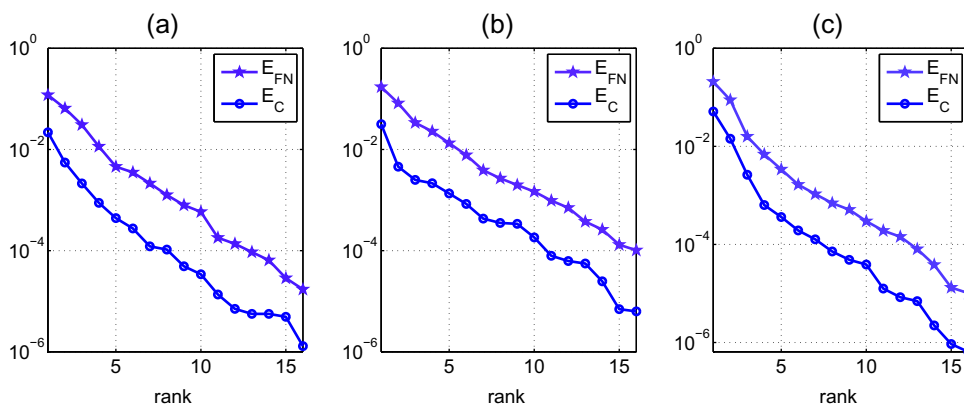

Fig. 2. Approximation error in E_{FN} and E_C norms versus Tucker rank for the electron densities of (a) CH_4, (b) C_2H_6 and (c) H_2O with $n = 65$.
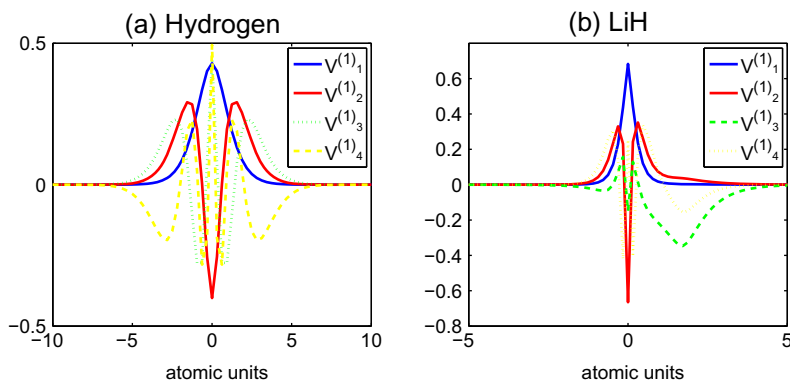

Fig. 3. Orthogonal vectors $V_k^{(1)}$, $k = 1, \ldots, 4$, for the rank-10 orthogonal Tucker approximation of the electron density for (a) the H atom and (b) for the LiH molecule.

contribution from the first three orthogonal components, e.g., for $1 \leqslant k \leqslant 3$, represents the relative ℓ^2-error with an accuracy of about 0.01%. This means that the adaptive orthogonal basis obtained via the Tucker-type approximation with moderate rank already represents the important physical quantities with satisfactory accuracy.
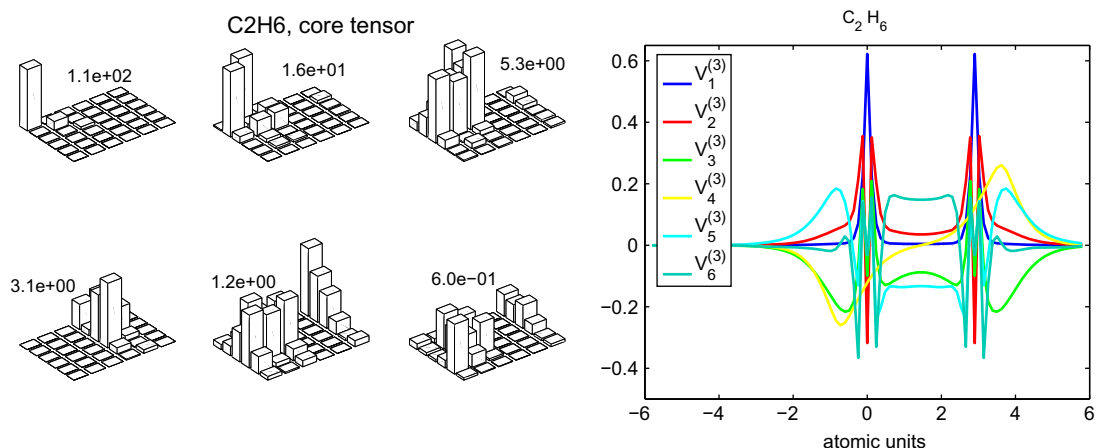
Fig. 4. Entries of the core tensor $6^{\otimes 3}$ and the orthogonal vectors $V_k^{(3)}$, $k = 1, \ldots, 6$ of the rank-6 Tucker-type decomposition of the electron density for the C_2H_6 molecule. Numbers given for every slice of core-tensor entries correspond to the maximum of $|b_\mathbf{k}|$ for the corresponding slice.

The electron densities of CH_4 and C_2H_6 are considered in the cubes $[-7, 7]^3$ and $[-5.8, 5.8]^3$, respectively, on a uniform $n^{\otimes 3}$ grid of size $n = 65$. Fig. 5(a) and (c) visualise the electron density of CH_4 in a plane containing the C atom and of C_2H_6 in a plane containing two C and two H atoms, correspondingly. Fig. 5(b) and (d) visualise the absolute approximation error for the electron densities ρ of these molecules in the corresponding planes for $r = 16$. In spite of large values of ρ at the cusp regions (~ 60 units) we observe a rather uniform distribution of the absolute value of the approximation error of the order $\sim 10^{-4} \div 10^{-5}$ a.u. in the computational domain. For electron density ρ of H_2O, which has even larger cusp (~ 148 units) at the origin, Fig. 2(c) shows that the tensor approximation in the volume $[-10, 10]^3$ gives the relative accuracy $\sim 10^{-5}$ with the rank $r = 16$. Finally, we present the convergence behaviour of the Tucker-type approximation applied to the Hartree potentials of CH_4, C_2H_6 and H_2O molecules (see Fig. 6).

Numerical results demonstrate the efficiency of the low-rank orthogonal approximations to the electron densities of the considered molecules. They show exponential convergence upon the Tucker rank r with a convergence rate which is almost independent of the resolution level depending on the grid-size n (theoretical results indicate $r = O(\log n)$). The low-rank tensor approximations of orbitals and electron densities can be further utilised for low-complexity calculations of functionals and operators of physical relevance, e.g., the electron–electron and electron–ion contributions to the total energy, the Hartree and exchange potentials, as well as to local exchange-correlation potentials in the Kohn–Sham equation. As an example, in the next section we consider the computation of the low-rank tensor approximation of the Hartree potential using tensor-product approximations of all functions and operators involved.
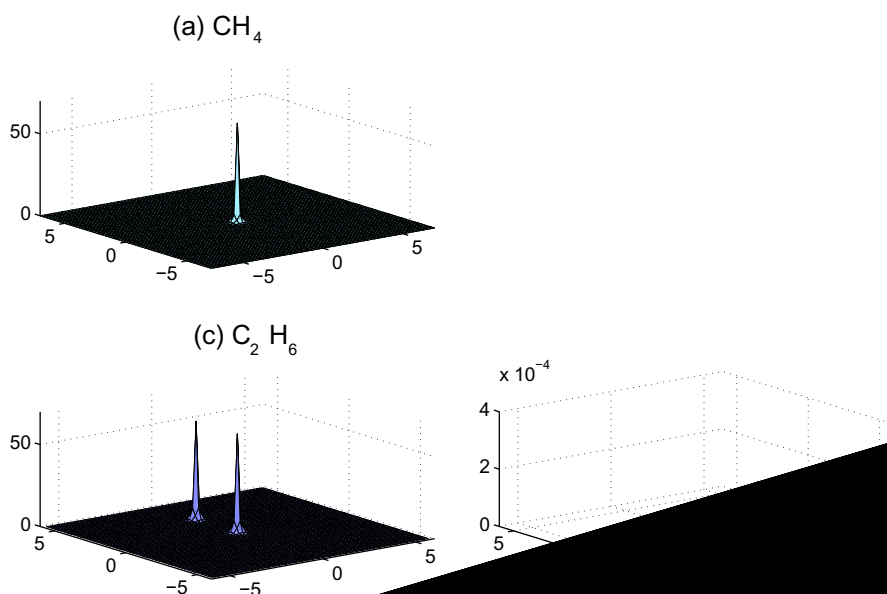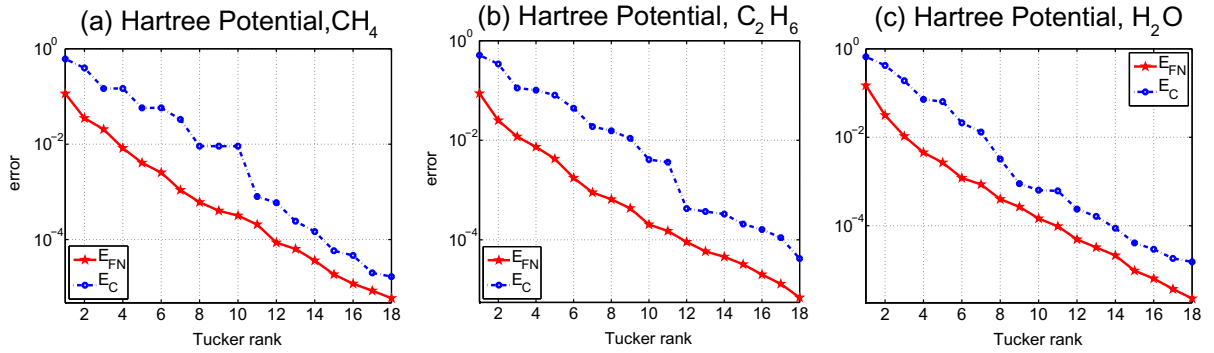
Fig. 6. Approximation error in E_{FN} and E_C norms versus Tucker rank for the Hartree potentials of (a) CH_4, (b) C_2H_6 and (c) H_2O.

3. Computing the Hartree potential by tensor-product convolution

3.1. General description

We consider the computation of the Hartree potential

$$V_H(\mathbf{x}) = \int_{\mathbb{R}^3} \frac{\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \, d\mathbf{y}, \quad \mathbf{x} \in \mathbb{R}^3, \tag{18}$$

by the discrete multi-dimensional tensor-product convolution on uniform grids in $\mathbb{R}^3$, described in Ref. [26]. The convolution product is defined by

$$w(\mathbf{x}) := (f \star g)(\mathbf{x}) := \int_{\mathbb{R}^3} f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) \, d\mathbf{y} \quad f, \ g \in L^2(\mathbb{R}^3).$$

Our particular choice is $g(\mathbf{x}) = \frac{1}{|\mathbf{x}|}$, $f(\mathbf{x}) = \rho(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$. Due to the physical prerequisites (orbitals and electron density have exponential decay) the function $f \star g$ is computed in some fixed box $\Omega = [-A, A]^3$ and f has its support in Ω.

We apply the standard collocation scheme to discretise the convolution product. First we introduce the equidistant tensor grid $\omega_{3,n} := \omega_1 \times \omega_2 \times \omega_3$ of collocation points $\{x_{\mathbf{m}}\}$ in Ω, $\mathbf{m} \in \mathcal{M} := \{1, \ldots, n+1\}^3$. Here $\omega_\ell := \{-A + (m-1)h : m = 1, \ldots, n+1\}$ $(\ell = 1, \ldots, 3)$ with mesh-size $h = 2A/n$.

For given piecewise constant basis functions $\{\phi_{\mathbf{i}}\}$, $\mathbf{i} \in \mathcal{I} := \{1, \ldots, n\}^3$, associated with $\omega_{3,n}$, let $f_{\mathbf{i}} = f(\mathbf{y}_{\mathbf{i}})$ be the representation coefficients of f in $\{\phi_{\mathbf{i}}\}$,

$$f(\mathbf{y}) \approx \sum_{\mathbf{i} \in \mathcal{I}} f_{\mathbf{i}} \phi_{\mathbf{i}}(\mathbf{y}),$$

where $\mathbf{y}_{\mathbf{i}}$ is the midpoint of the grid-cell numbered by $\mathbf{i} \in \mathcal{I}$. Now the *projection-collocation scheme* reads as

$$f \star g \approx [w_{\mathbf{m}}]_{\mathbf{m} \in \mathcal{M}}, \quad w_{\mathbf{m}} := \sum_{\mathbf{i} \in \mathcal{I}} f_{\mathbf{i}} \int_{\mathbb{R}^3} \phi_{\mathbf{i}}(\mathbf{y}) g(\mathbf{x}_{\mathbf{m}} - \mathbf{y}) \, d\mathbf{y}, \quad \mathbf{x}_{\mathbf{m}} \in \omega_3.$$

As a first step, we precompute the projection coefficients

$$g_{\mathbf{i}} = \int_{\mathbb{R}^3} \frac{\phi_{\mathbf{i}}(\mathbf{y})}{|\mathbf{y}|} \, d\mathbf{y}, \quad \mathbf{i} \in \mathcal{I}.$$

Following [27], we compute $w_{\mathbf{m}}$ by copying the corresponding portion of the *discrete convolution* in $\mathbb{R}^3$

$$\mathcal{F} \star \mathcal{G} := [z_{\mathbf{j}}], \quad z_{\mathbf{j}} := \sum_{\mathbf{i} \in \mathcal{I}} f_{\mathbf{i}} g_{\mathbf{j} - \mathbf{i} + 1}, \quad \mathbf{j} \in \mathcal{J} := \{1, \ldots, 2n - 1\}^3, \tag{19}$$

centred at $\mathbf{j} = \mathbf{n}$. Here the sum is over all $\mathbf{i} \in \mathcal{I}$, which lead to legal subscripts for $g_{\mathbf{j} - \mathbf{i} + 1}$, i.e., $\mathbf{j} - \mathbf{i} + 1 \in \mathcal{I}$. Based on 3D-FFT this computation can be performed with $O(n^3 \log n)$ arithmetic operations.

3.2. Tensor-product convolution

The complexity can be reduced to $O(n \log^q n)$ based on the idea of low tensor rank approximation combined with the multi-linear algebra [27,26]. To that end, the coefficient tensor $\mathcal{G} = [g_{\mathbf{i}}] \in \mathbb{R}^{\mathcal{I}}$ for the Coulomb potential $\frac{1}{|\mathbf{x} - \mathbf{y}|}$ is approximated in the rank-R canonical tensor format (see Section 2) using optimised *sinc*-quadratures [13,17,26], where the rank-parameter $R = O(|\log \varepsilon| \log n)$ depends logarithmically on both the required accuracy and the problem-size n. In all computations

presented below it was enough to choose the tensor rank in the range $R \in [10, 20]$ to provide the desired accuracy. The third order tensor $\mathcal{F} = [f_{\mathbf{i}}] \in \mathbb{R}^{\mathcal{I}}$ is approximated either in the rank $\mathbf{r} = (r, r, r)$ Tucker format or via the canonical model with tensor rank R_1.

Let $\mathcal{F}$ be represented in the Tucker format (7). Then (19) can be implemented in the tensor-product form

$$\mathcal{F} \star \mathcal{G} = \sum_{k=1}^{R} \sum_{\mathbf{m}=\mathbf{1}}^{\mathbf{r}} c_k b_{m_1 m_1 m_3} \left(U_k^{(1)} \star V_{m_1}^{(1)} \right) \otimes \left(U_k^{(2)} \star V_{m_2}^{(2)} \right) \otimes \left(U_k^{(3)} \star V_{m_3}^{(3)} \right).$$

The 1D convolutions $U_k^{(\ell)} \star V_{m_\ell}^{(\ell)} \in \mathbb{R}^{2n-1}$ can be computed on an equidistant grid by FFT in $O(n \log n)$ operations. This leads to the overall complexity

$$\mathcal{N}_{C \star T} = O(3rRn \log n + Rr^3). \tag{20}$$

Likewise, approximating $\mathcal{F}$ in the rank-R_1 canonical format (see (8)) enables us to compute $\mathcal{F} \star \mathcal{G}$ in the form

$$\mathcal{F} \star \mathcal{G} = \sum_{k=1}^{R} \sum_{m=1}^{R_1} c_k b_m \left(U_k^{(1)} \star V_m^{(1)} \right) \otimes \left(U_k^{(2)} \star V_m^{(2)} \right) \otimes \left(U_k^{(3)} \star V_m^{(3)} \right),$$

which leads to the numerical cost

$$\mathcal{N}_{C \star C} = O(RR_1 n \log n).$$

3.3. Numerical illustrations on the fast tensor-product convolution

As an example we consider the CH_4 molecule. In order to test the performance of our approach for the high accuracy demands of quantum chemical calculations, it becomes necessary to use a pseudopotential on the C atom. We have used a semilocal energy-adjusted pseudopotential and corresponding basis set of Dolg and co-workers [2]. This enables us to achieve the desired accuracies on grids $\omega_{\mathbf{3},n}$ already for values of n that can be easily handled by our MATLAB code.Tensor approximations of full potentials will require the fast convolution schemes applicable to locally refined grids, see e.g. Refs. [16,26,27].

The pseudo-density for the CH_4 molecule obtained from a Hartree–Fock calculation is given by (16), where $K = 4$ is the number of pseudo-orbitals, $R_0 = 50$ and polynomial degree$(P_{k,i}) \leqslant 2$. In the following we represent the convolving tensor $\rho(\mathbf{x})$ in the canonical format. In this way, the products of two Gaussians can be written in the form of single Gaussians by

$$e^{-\lambda(x-a)^2} \cdot e^{-\beta(x-b)^2} = e^{\sigma} \cdot e^{-\gamma(x-c)^2}, \quad \sigma = \frac{\lambda \beta (a-b)^2}{\lambda + \beta}, \quad \gamma = \lambda + \beta, \quad c = \frac{a\lambda + b\beta}{\lambda + \beta},$$

which leads to the following bound on the initial rank of the input tensor

$$R_2 \leqslant \frac{R_0(R_0 + 1)}{2} = 1275. \tag{21}$$

We approximate $V_H(\mathbf{x})$, $\mathbf{x} \in \omega_{\mathbf{3},n}$, in (18) via a tensor-product collocation convolution scheme by $V_H^{(n)}$ on the $n^{\otimes 3}$ tensor-grid $\omega_{\mathbf{3},n}$ in the reference cube $[-A, A]^3$, $A = 10.6$ bohr. The left part of Fig. 7 gives the approximation error $V_H - V_H^{(n)}$ for the sequence of grids with $n = 112, 224, 448$. It indicates that the discrete convolution has an accuracy of order $O(h^2)$ in the relevant physical domain corresponding to the essential support of $\rho(\mathbf{x})$ (see [26] for detailed error estimates). Following [26], we then apply the Richardson extrapolation technique to obtain higher accuracy approximations of order $O(h^3)$ without extra computational cost. The numerical gain of using an extrapolated solution is obtained due to the fact that the approximation error $O(h^3)$ on the single grid would require the univariate grid-size $N = n^{3/2} \gg n$. The corresponding Richardson extrapolant $V_{H,Rich}^{(n)}$ approximating $V_H(\mathbf{x})$ over a pair of nested grids $\omega_{\mathbf{3},n}$ and $\omega_{\mathbf{3},2n}$, and defined on the "coarse" $n^{\otimes 3}$-grid, is given by

$$V_{H,Rich}^{(n)} = \left(4 \cdot V_H^{(2n)} - V_H^{(n)} \right) \Big/ 3 \quad \text{in the grid-points on } \omega_{\mathbf{3},n}.$$

The right part of Fig. 7 shows the effect of the Richardson extrapolation on a sequence of grids $\omega_{\mathbf{3},n}$ with $n = 112, 224, 448$.

Next we demonstrate the computational efficiency of fully discrete tensor-product convolution to compute the approximate Hartree potential $V_H^{(n)}$ on a sequence of grids $\omega_{\mathbf{3},n}$ with $n = 32, 64, 128, \ldots, 512$. We use the same input data as above (corresponding to the CH_4 molecule) with initial rank $R_2 = 1275$. To reduce the canonical rank, we apply the two-level scheme as described in [27]: first, we compute the rank-r Tucker decomposition of the input tensor and build the corresponding rank-r^2 canonical representation (with $r = 15$); second, we recompress it to the reduced rank $R_2' = 120 < r^2$ using the slicewise "truncated" SVD of small $r \times r$ matrices. This is a heuristic procedure that allows to reduce the canonical rank of the convolving tensor. Quantitative estimates on the behaviour of the corresponding singular values seem to be the difficult problem. In the following we present the related numerical results.

Table 1 compares the elapsed time to compute the discrete convolution on $n^{\otimes 3}$ grid by using the 3D-FFT (asymptotic complexity $O(8n^3 \log n)$) and the fast tensor-product scheme of complexity $\mathcal{N}_{C \star C} = O(R_2' R_1 n \log n)$ described above. We recall that $R_1 \in [10, 20]$ denotes the tensor rank of the Coulomb coefficient tensor in $\mathbb{R}^3$.
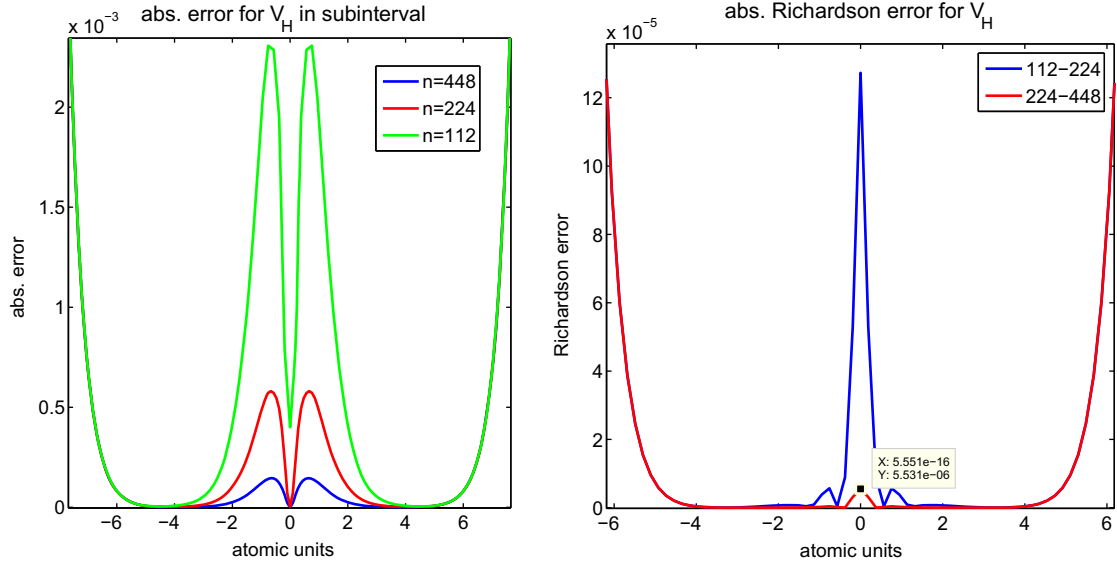
Fig. 7. Absolute approximation error in the Hartree potential V_H for the pseudo-density of CH_4 in the subinterval $\Omega = [-7,7] \times \{0\} \times \{0\}$ (left) and the reduced error by Richardson extrapolation involving two grids (right).

Table 1
Elapsed CPU time (s) to compute the discrete convolution on $n^{\otimes 3}$ grids for the tensor-product scheme $Conv_{CC}$ and 3D-FFT on a Sun Fire X4600 computer with 2.6 GHz processor.

n^3	32^3	64^3	128^3	256^3	512^3
3D-FFT	0.06	0.41	4.3	55.4	582.8
$Conv_{CC}$	0.51	0.73	1.26	3.05	10.01

Note that in both 3D-FFT based and tensor-product convolution schemes on the $n^{\otimes 3}$ grid one requires the double-size FFT with grid-size $2n - 1$. Since the tensor-product convolution requires only 1D-FFT (negligible cost in the range $n \leqslant 10^4$), the double problem size $2n - 1$ does not lead to severe restrictions on CPU-time and on the storage space in the case of fast tensor-product scheme.

Our results indicate that the tensor-product convolution outperforms the conventional 3D-FFT for large problems ($n \geqslant 128$). But even more important, it has much less restrictive memory consumptions of order $O(RR_1 n \log n)$ compared with $O(n^3 \log n)$. In particular, the 3D-FFT already runs out of memory on the grid $\omega_{3,512}$, hence, in the table above, we give the extrapolated CPU-time ~ 500.0 s in this case.

4. Coulomb matrix and Hartree–Fock energy

In the following, we apply the tensor-product scheme [27] to compute the Coulomb matrix of the Hartree potential for the CH_4 molecule in a GTO basis set. The remaining parts of the Fock operator are computed by the MOLPRO package in a conventional way. The Coulomb matrix for V_H with respect to the set of normalised "Cartesian Gaussians" $\{g_k\}$ is given by

$$J_{km} := \int g_k(\mathbf{x}) g_m(\mathbf{x}) V_H(\mathbf{x}) \, d\mathbf{x}, \quad k, \, m = 1, \ldots R_0. \tag{22}$$

Fig. 8(a) visualises the exact Coulomb matrix elements J_{km}. The error E_n between J_{km} and its approximation obtained by the discrete collocation scheme for calculating the Hartree potential V_H is given in Fig. 8(b). Results on high accuracy calculations based on Richardson extrapolation are illustrated in Fig. 9.

Let $\mathcal{F} = \{\mathcal{F}_{km}\} = \{(g_k, Fg_m)_{L^2}\}$ be the Galerkin matrix of the Fock operator with respect to the normalised Gaussian basis $\{g_k\}$ and $\mathcal{M} = \{\mathcal{M}_{km}\} = \{(g_k, g_m)_{L^2}\}$ $(k, m = 1, \ldots, R_0)$ be the corresponding overlap matrix. We substitute the Galerkin matrix $\mathcal{J} = \{J_{km}\}_{k,m=1}^{R_0}$ from (22), which represents the approximate Hartree potential $V_H^{(n)}$ in the basis $\{g_k\}$, into the Fock matrix

$$\mathcal{F} = \mathcal{F}_{lin} + \mathcal{J} + \mathcal{K},$$

where $\mathcal{F}_{lin} = \{(g_k, (-\frac{1}{2}\Delta + V_N)g_m)_{L^2}\}_{k,m=1}^{R_0}$ and $\mathcal{K} = \{(g_k, V_x g_m)_{L^2}\}_{k,m=1}^{R_0}$ represent the linear part of the Fock operator and the exchange potential, respectively.
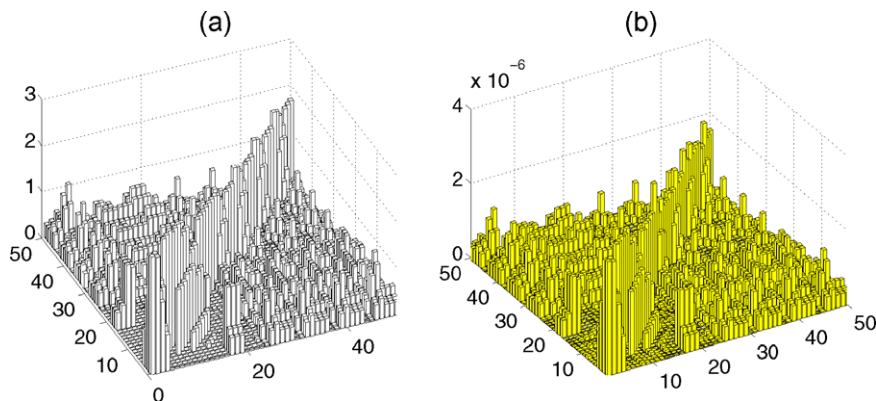
Fig. 8. (a) Coulomb matrix J_{km} for CH_4 computed for the GTO basis set k, $m = 1, \ldots R_0$. (b) Absolute error of the discrete approximation using Richardson extrapolation involving the grids with $n = 400$ and 800.
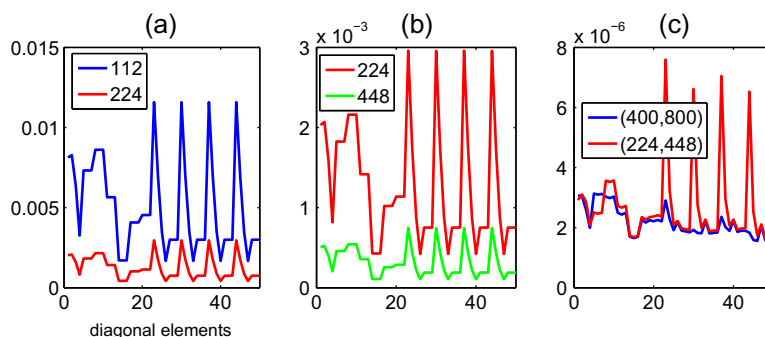

Fig. 9. Absolute error in the diagonal Coulomb matrix elements for (a) $n = 112, 224$ and (b) $n = 224, 448$. (c) Absolute error in the diagonal Coulomb matrix elements computed by Richardson extrapolation on the pairs of grids $(224, 448)$ and $(400, 800)$.

Now the total energy of the system is computed by solving the eigenvalue problem

$$\mathcal{F}\Psi_a = \lambda_a \mathcal{M}\Psi_a, \quad a = 1, \ldots, K,$$

for the first $K = N/2$ smallest eigenvalues λ_a (for the CH_4 molecule with pseudopotential $K = 4$). For this computation we apply the MOLPRO package. The total energy is calculated by

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

with $\widetilde{J}_a = (\psi_a, V_H \psi_a)_{L^2}$ and $\widetilde{K}_a = (\psi_a, \mathcal{V}_{ex}\psi_a)_{L^2}$ $(a = 1, \ldots, N/2)$ being the so-called Coulomb and exchange integrals, respectively, computed with respect to the orbitals ψ_a $(a = 1, \ldots, N/2)$ (compare with (22)). The corresponding errors in the Hartree–Fock energy E_{HF} and of individual eigenvalues for different grids are given in Table 2.

Let $E_n \in \mathbb{R}^{R_0 \times R_0}$ be the error in the Coulomb matrix $\{J_{km}\}$. Notice that both matrices $\{J_{km}\}$ and E_n are symmetric up to round-off errors. We consider the approximation error in the eigenvalues λ_a, $a = 1, \ldots, K$. In this way, we introduce the error estimator $\mathcal{E}_1(E_n)$ to the energy error based on the $\| \cdot \|_1$-evaluation of the error matrix E_n calculated by

$$\mathcal{E}_1(E_n) = \max_{j=1,\ldots,R_0} \sum_{i=1}^{R_0} |(E_n)_{ij}|.$$

Table 2
Error of the Hartree–Fock energy E_{HF} (hartree) and of individual eigenvalues λ_i computed on different grids, denoted by δ_n, for the CH_4 molecule. The values for $\delta_{(n,2n)}$ represent the Richardson extrapolation error for pairs of grids of size n and $2n$, where $n = 112, 224$ and $n = 224, 448$, respectively.

	Exact value	δ_{112}	δ_{224}	δ_{448}	$\delta_{(112,224)}$	$\delta_{(224,448)}$
λ_1	−0.95072154	0.0148	3.73×10^{-3}	9.37×10^{-4}	1.53×10^{-5}	6.05×10^{-6}
$\lambda_2, \lambda_3, \lambda_4$	−0.54457593	0.0121	3.03×10^{-3}	7.62×10^{-4}	1.57×10^{-5}	5.21×10^{-6}
E_{HF}	−7.84226746	0.0113	2.86×10^{-2}	7.19×10^{-3}	1.08×10^{-4}	3.53×10^{-5}

Table 3
Comparison of error estimators $\mathcal{E}_1$ and $\mathcal{E}_2$ for Richardson extrapolated Coulomb matrix elements.

n	(112,224)	(224,448)
$\mathcal{E}_1(E_n)$	1.1212×10^{-4}	3.742×10^{-5}
$\mathcal{E}_2(E_n)$	6.9×10^{-5}	2.40×10^{-5}

Our numerical results presented in Table 3 indicates that the error estimator $\mathcal{E}_1(E_n)$ gives a quite accurate upper bound to the approximate total energy. The second row of Table 3 illustrates the validity of the error estimator $\mathcal{E}_2(E_n)$ based on the discrete $\|\cdot\|_2$-norm evaluation,

$$\mathcal{E}_2(E_n) = \|E_n\|_2.$$

Analysis of both estimators can be found, for example, in [31]. In particular, for each individual eigenvalue we have

$$\left|\lambda_a - \lambda_a^{(n)}\right| \leqslant \mathcal{E}_2(E_n), \quad \left|\lambda_a - \lambda_a^{(n)}\right| \leqslant cond(P)\mathcal{E}_1(E_n),$$

where the matrix P is defined from the diagonalisation procedure $\mathcal{M}^{-1/2}\mathcal{F}\mathcal{M}^{-1/2} = PDP^T$ with the diagonal matrix D. In our case we have $cond(P) \approx 3.46$. In Table 2, we present the approximation errors δ_n and the corresponding extrapolated errors $\delta_{(n,2n)}$ for the single eigenvalue λ_1 and the three degenerate eigenvalues $\lambda_2 = \lambda_3 = \lambda_4$.

From Table 3, we observe that the error estimators $\mathcal{E}_2(E_n)$ and $\mathcal{E}_1(E_n)$ can be utilised for an a priori error estimation to the total energy based only on precomputed error bounds to the Fock matrix.

5. Conclusions

We have investigated adaptive orthogonal tensor-product approximations for some quantities of interest in Hartree–Fock and density functional theory. In particular, we studied the convergence rate upon the tensor rank for electron densities, Coulomb and Hartree potentials as well as for the Hartree–Fock energy for a number of simple molecules.

For the considered molecules, the ALS iterative scheme to compute the orthogonal Tucker-type decomposition has demonstrated robust behaviour providing exponentially fast convergence in the tensor rank r. It is a remarkable fact that the orthogonal tensor-product ansatz typically provides rather "uniformly distributed" max-norm errors, so that the tensor approximant fits the initial electron density equally well at the near- and far-field regions of the nuclear cusps as shown in Fig. 5.

The orthogonal univariate components of the Tucker-type decomposition resemble the shape of the decomposed electron density along the corresponding axis and appear to be suitable for the construction of a low-dimensional problem-adapted set of orthogonal basis functions. We observe that due to the fast decay in the entries of the core tensor (see Fig. 4) the Tucker-type approximation with *moderate rank* already represents the physical quantity with satisfactory accuracy.

The convolution integral representing the Hartree potential is computed by the fully discrete tensor-product collocation scheme combined with the Richardson extrapolation on a sequence of grids. Our approach exhibits $O(h^3)$ convergence rate with $h = O(n^{-1})$. It requires $O(3rn + r^3)$ storage, where r denotes the Tucker rank of the electron density with $r \ll n$, specifically, $r = O(\log n)$. This method ensures, in particular for the CH_4-molecule, the relative accuracy of the order of 10^{-6}, on the fine $n^{\otimes 3}$ grid with $n \sim 500$. It requires low computational resources since the corresponding three-dimensional discrete convolution transform is performed via 1D tensor-product FFT. This leads to sublinear complexity $O(n\log^3 n)$ instead of the cost of the conventional 3D-FFT, $O(n^3 \log n)$. In this way, the presented approach offers a principal step toward the numerical solution of the Hartree–Fock/Kohn–Sham equations in the fully discrete *tensor-product format* with linear scaling in the univariate grid-size n.

Acknowledgments

References

[1] J. Almlöf, Direct methods in electronic structure theory, in: D.R. Yarkony (Ed.), Modern Electronic Structure Theory, vol. II, World Scientific, Singapore, 1995, pp. 110–151.
[2] A. Bergner, M. Dolg, W. Küchle, H. Stoll, H. Preuss, Ab initio energy-adjusted pseudopotentials for elements of groups 13–17, Mol. Phys. 80 (1993) 1431–1441.
[3] G. Beylkin, M.J. Mohlenkamp, Numerical operator calculus in higher dimensions, Proc. Natl. Acad. Sci. USA 99 (2002) 10246–10251.
[4] G. Beylkin, M.J. Mohlenkamp, Algorithms for numerical analysis in high dimension, SIAM J. Sci. Comput. 26 (6) (2005) 2133–2159.
[5] P. Burgisser, M. Clausen, M.A. Shokrollahi, Algebraic Complexity Theory, Springer, 1997.
[6] S.R. Chinnamsetty, M. Espig, W. Hackbusch, B.N. Khoromskij, H.-J. Flad, Tensor product approximation with optimal rank in quantum chemistry, J. Chem. Phys. 127 (2007) 084110.

[7] P.G. Ciarlet, C. Le Bris (Eds.), Handbook of Numerical Analysis, vol. X, Computational Chemistry, Elsevier, Amsterdam, 2003.
[8] L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-$(R_1, \ldots, R_N)$ approximation of higher-order tensors, SIAM J. Matrix Anal. Appl. 21 (2000) 1324–1342.
[9] B.I. Dunlap, Robust and variational fitting, Phys. Chem. Chem. Phys. 2 (2000) 2113–2116.
[10] T.H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. 90 (1989) 1007–1023.
[11] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, Auxiliary basis sets to approximate Coulomb potentials, Chem. Phys. Lett. 240 (1995) 283–290.
[12] H.-J. Flad, W. Hackbusch, B.N. Khoromskij, R. Schneider, Concept of data-sparse tensor-product approximation in many-particle modeling, Matrix Methods: Theory, Algorithms, Applications, World Scientific Publ., Singapore, 2010, pp. 313–347.
[13] W. Hackbusch, B.N. Khoromskij, Low-rank Kronecker product approximation to multi-dimensional nonlocal operators. Part I. Separable approximation of multi-variate functions, Computing 76 (2006) 177–202.
[14] W. Hackbusch, B.N. Khoromskij, E.E. Tyrtyshnikov, Hierarchical Kronecker tensor-product approximations, J. Numer. Math. 13 (2005) 119–156.
[15] M. Häser, J. Almlöf, Laplace transform techniques in Møller–Plesset perturbation theory, J. Chem. Phys. 96 (1992) 489–494.
[16] W. Hackbusch, Efficient Convolution with the Newton Potential in d Dimensions, Preprint 81, MPI MIS, Leipzig, 2007.
[17] W. Hackbusch, B.N. Khoromskij, Tensor-product approximation to operators and functions in high dimension, J. Complex. 23 (2007) 697–714.
[18] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan, G. Beylkin, Multiresolution quantum chemistry: basic theory and initial applications, J. Chem. Phys. 121 (2004) 11587–11598.
[19] T. Helgaker, P. Jørgensen, J. Olsen, Molecular Electronic-Structure Theory, Wiley, New York, 1999.
[20] F.L. Hitchcock, The expression of a tensor or a polyadic as a sum of products, J. Math. Phys. 6 (1927) 164–189.
[21] F.L. Hitchcock, Multiple invariants and generalized rank of a p-way matrix or tensor, J. Math. Phys. 7 (1927) 39–79.
[22] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, G. Beylkin, Efficient Solution of Poisson's Equation with Free Boundary Conditions, J. Chem. Phys. 125 (2006) 074105.
[23] S.A. Ghasemi, A. Neelov, S. Goedecker, A Particle-Particle, Particle-Density (P^3D) Algorithm for the Calculation of Electrostatic Interactions of Particles with Slab-like Geometry, J. Chem. Phys. 127 (2007) 224102.
[24] B.N. Khoromskij, Structured rank-$(r_1, \ldots, r_d)$ decomposition of function-related tensors in $\mathbb{R}^d$, Comput. Method Appl. Math. 6 (2) (2006) 194–220.
[25] B.N. Khoromskij, On tensor approximation of Green iterations for Kohn–Sham equations, Comput. Visual. Sci. 11 (2008) 259–271.
[26] B.N. Khoromskij, Fast and accurate tensor approximation of multivariate convolution with linear scaling in dimension. Preprint MPI MIS 36, Leipzig 2008. J. Comp. Appl. Math., accepted for publication.
[27] B.N. Khoromskij, V. Khoromskaia, Low rank Tucker tensor approximation to the classical potentials, Central Eur. J. Math. 5 (3) (2007) 1–28.
[28] B.N. Khoromskij, V. Khoromskaia, Multigrid Accelerated Tensor Approximation of Function Related Multi-dimensional Arrays, Preprint MPI MIS, 40/2008, Leipzig, 2008 (SIAM J. Sci. Comput. 2009, to appear).
[29] F.R. Manby, P.J. Knowles, A.W. Lloyd, The Poisson equation in density fitting for the Kohn–Sham Coulomb problem, J. Chem. Phys. 115 (2001) 9144–9148.
[30] I.V. Oseledets, D.V. Savostianov, E.E. Tyrtyshnikov, Tucker dimensionality reduction of three-dimensional arrays in linear time, SIAM J. Matrix Anal. Appl. 30 (3) (2008) 939–956.
[31] J. Stoer, R. Bulirsch, Numerische Mathematik, vol. 2, Springer, Berlin, 1990.
[32] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (1966) 279–311.
[33] H.-J. Werner, P.J. Knowles, et al., MOLPRO, Version 2002.10, A Package of Ab Initio Programs for Electronic Structure Calculations.
[34] T. Zhang, G.H. Golub, Rank-one approximation to high order tensors, SIAM J. Matrix Anal. Appl. 23 (2001) 534–550.