Tensor decomposition in electronic structure calculations on 3D Cartesian grids

(revised version: November 2007)

by

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

Preprint no.: 65 2007
Tensor decomposition in electronic structure calculations on 3D Cartesian grids

S. R. Chinnamsetty*, H.-J. Flad°, V. Khoromskaia*, and B.N. Khoromskij*

* Max-Planck-Institute for Mathematics in the Sciences,
Inselstr. 22-26, D-04103 Leipzig, Germany.
° Christian-Albrechts-Universität zu Kiel,
Christian-Albrechts-Platz 4, 24098 Kiel, Germany.
Corresponding author {vekh}@mis.mpg.de (V. Khoromskaia)

Abstract

In this paper we propose 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 potential applications of tensor approximations for the numerical solution of Hartree-Fock and Kohn-Sham equations on 3D Cartesian grids.

Low-rank orthogonal Tucker-type tensor approximations are investigated for electron densities and Hartree potentials of simple molecules, where exponential convergence with respect to the Tucker rank is observed. 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 Richardson extrapolation, our approach exhibits \( O(h^3) \) convergence with \( h = O(n^{-1}) \), and requires \( O(3rn + r^3) \) storage, where \( r \) denotes the Tucker rank of the electron density with \( r \ll n \) almost uniformly in \( n \) (specifically, \( r = O(\log n) \)). For example, Hartree-Fock calculations for the \( \text{CH}_4 \) molecule, with a pseudopotential on the C atom, achieved accuracies of the order of \( 10^{-5} \) hartree with a grid-size \( n \) of several hundreds. For large 3D grids \( n \geq 128 \), the tensor-product convolution scheme markedly outperforms the 3D–FFT in both the computing time and storage requirements.

AMS Subject Classification: 65F30, 65F50, 65N35, 65F10
Key words: Tucker-type tensor decomposition, Hartree-Fock equation, discrete convolution, orthogonal adaptive tensor-product basis.

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. Let us just 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 the monographs [20, 7] for a detailed exposition of these subjects. Furthermore, separable approximations of energy denominators, following the pioneering work of Almlöf and Häser [1, 15], became a popular approach to reduce the computational effort in many-body perturbation theory. Recently, alternative 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

\[ f(x) \approx \sum_{k=1}^{R} h_{1,k}(x_1)h_{2,k}(x_2)\ldots h_{d,k}(x_d), \quad x \in \mathbb{R}^d \] (1.1)

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. Initiated by the work of Beylkin and Mohlemkamp [4, 5], “optimal” tensor-product approximations for various potentials and more general integral operators, with possible applications in quantum chemistry, have been discussed in the literature, cf. Refs. [19, 12, 13, 14] 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, 26], 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 [2] 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 (1.1) 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 propose 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. The Tucker format poses orthogonality constraints on the univariate components which turns the optimization into a well-posed problem. Primarily, the Tucker-type decomposition was applied in chemometrics and independent component analysis, and it was generally treated as a method for statistical data processing [29] with moderate accuracy requirements. Potential applications for large scale problems in quantum chemistry have been already shown recently in Ref. [25] where one can find a thorough description of tensor operations in the Tucker 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 [24, 25] 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 $R >> R_1$. Applying the two-level scheme [25], we first compute its rank-$r$ Tucker approximation and then recompress the small size core tensor by the rank-$R_1$ canonical decomposition.

An essential new feature of this scheme is that convolutions require only 1D–FFTs instead of 3D–FFTs as in traditional approaches (for example, a large variety of plane wave and wavelet methods exhibiting linear scaling in the number $N_3 = n^3$ of grid points, see e.g. the BigDFT program [21, 22]). Therefore, it is possible to perform convolutions involving tensor product approximations with much lower complexity with respect to the grid size $N_3$ (see Remark 2.1). Furthermore, storage of the full 3D–grid is no longer required which provides another potentially significant advantage with respect to 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 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

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

for the $N/2$ lowest eigenvalues $\epsilon_i$ and spatial eigenfunctions $\psi_i$ ($i = 1, ..., N/2$), in the case of a closed-shell $N$ electron system. The Hartree-Fock equation (1.2) 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 $V_{\text{nuc}}$ contains the bare Coulomb- or pseudopotentials of the nuclei. Within the present work, we focus on the computation of the Hartree potential

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

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

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

The remaining term is the nonlocal exchange operator

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

3
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 Tucker 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, 18, 23], we consider the computation of Hartree potentials within a fully discrete tensor-product collocation scheme (cf. [24]) on a 3D Cartesian grid. A numerical example is provided for the CH_4 molecule 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

We introduce the linear space of real-valued d-th order tensors

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

with the scalar product defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle := \sum_{(i_1,...,i_d) \in I} a_{i_1...i_d} b_{i_1...i_d} \quad \text{for} \ \mathcal{A}, \mathcal{B} \in \mathbb{R}^I. \quad (2.1)$$

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

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

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

defined by entries

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

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

$$\mathcal{A}(\mathbf{r}) = \sum_{k_1=1}^{r_1} \cdots \sum_{k_d=1}^{r_d} b_{k_1...k_d} \cdot V_{k_1}^{(1)} \otimes \cdots \otimes V_{k_d}^{(d)} \approx \mathcal{A}, \quad b_{k_1...k_d} \in \mathbb{R}, \quad (2.2)$$

where $$V_{k_\ell}^{(\ell)} \in \mathbb{R}^{n_\ell} \ (\ell = 1,...,d)$$, and $$\mathcal{B} = \{b_k\} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$$ is called the core tensor. Here $$\{V_{k_\ell}^{(\ell)}, \ 1 \leq k_\ell \leq r_\ell\}$$ are linear independent vectors for each fixed $$\ell$$, hence without loss of generality, we assume that they are mutually orthogonal, i.e., $$V^{(\ell)} = \left[V_{1}^{(\ell)} V_{2}^{(\ell)} \cdots V_{r_\ell}^{(\ell)}\right] \in \mathbb{R}^{n_\ell \times r_\ell}$$ are orthogonal matrices.

The representation (2.2) means that for given set of rank parameters $$r_\ell \leq n_\ell \ (\ell = 1,...,d)$$, any tensor can be decomposed approximately into $$r_1r_2\ldots r_d$$ outer products (respectively, $$r_d$$
if \( r_1 = \ldots = r_d = r \), which requires only \( r_1 + \cdots + r_d \) canonical components \( V_{k_1}^{(1)} \) of size \( n_1 \) and \( r^d \) coefficients \( \{b_k\} \) 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, CP model) represents a tensor by a sum of rank-1 non-orthogonal components as follows

\[
\mathbf{A}(R) = \sum_{k=1}^R c_k \cdot \mathbf{U}_k^{(1)} \otimes \ldots \otimes \mathbf{U}_k^{(d)}, \quad c_k \in \mathbb{R},
\]

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

Below we consider approximating properties of the orthogonal Tucker-type representations (2.2) in \( \mathbb{R}^3 \) for electron densities of some simple molecules. We apply the alternating least-squares (ALS) iterative scheme described in [8, 25] to compute the low-rank tri-linear approximations for electron densities of the H atom, LiH, CH$_4$, C$_2$H$_6$ and H$_2$O molecules.

**Remark 2.1** The Tucker model applied to a fully populated tensor of size \( n^3 \) has the complexity of order \( O(n^4) \) (storage and computational time). If the input tensor is already presented in the rank-R CP format then the corresponding canonical-to-Tucker ALS method with the Tucker rank \( r \) has the complexity in the range \( O(r R n) : O(r^2 R n) \), depending on the problem setting (see §2.4.2 in [25]). Notice that the so-called adaptive cross approximation method of complexity \( O(r^3 n) \) was recently proposed in [27].

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

As the first step, the approximate reference orbitals and electron densities of the mentioned molecules are computed by the standard quantum chemistry package MOLPRO [30] in the form of an expansion by “Cartesian Gaussian” basis functions, i.e.,

\[
g(\mathbf{x}) = \sum_{k=1}^{R_0} \mu_k (x_1 - A_k)^{\ell_k} (x_2 - B_k)^{m_k} (x_3 - C_k)^{n_k} \exp(-\alpha_k \zeta_k^2)
\]

with

\[
\zeta_k^2 = (x_1 - A_k)^2 + (x_2 - B_k)^2 + (x_3 - C_k)^2, \quad \mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3,
\]

where the exponents \( \alpha_k \) have been taken from standard VDZ basis sets [10].

We assume that any particular molecule is imbedded in a certain fixed computational box \([-A, A]^3\) with a suitable \( A > 0 \). Let \( \mathcal{I} \subseteq [-A, A]^3 \) be a uniform \( n \times n \times n \) tensor-product grid (shortly \( n^{\otimes 3}\)) indexed by \( \mathcal{I} = I_1 \times I_2 \times I_3 \). For a given continuous function \( g : [-A, A]^3 \to \mathbb{R} \), we introduce the collocation-type function related tensor of order 3 by
Figure 2.1: 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.

$$A_0 \equiv A(g) := [a_{i_1 i_2 i_3}]_{(i_1, i_2, i_3) \in \mathcal{I}} \in \mathbb{R}^{n \times n \times n} \text{ with } a_{i_1 i_2 i_3} := g(x_{i_1}^{(1)}, x_{i_2}^{(2)}, x_{i_3}^{(3)}),$$

where $(x_{i_1}^{(1)}, x_{i_2}^{(2)}, x_{i_3}^{(3)}) \in \omega_3$ are the grid collocation points. In our applications, the function $g$ is defined by (2.4) using the set of separable “Cartesian Gaussian” functions. The reconstructed 3-rd order tensor $A_0$ is approximated by a rank $r = (r, r, r)$ Tucker-type representation for a sequence of rank-parameters $r = 1, 2, \ldots, r_0$. The orthogonal components and the corresponding core tensor of the size $r \otimes 3$ are then used for the reconstruction of the approximating tensor $A_{(r)} \approx A_0$. All numerical multi-linear approximations are performed in MATLAB, release 7.3.

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

$$E_{FN}^{(r)} = \frac{\|A_0 - A_{(r)}\|}{\|A_0\|}, \quad E_{C}^{(r)} := \frac{\max_{i \in \mathcal{I}} |a_{0,i} - a_{r,i}|}{\max_{i \in \mathcal{I}} |a_{0,i}|}.$$  \hspace{1cm} (2.5)

The relative difference of the norms denoted by

$$E_{FE}^{(r)} = \frac{\|A_0\| - \|A_{(r)}\|}{\|A_0\|} \leq \frac{\|A_0 - A_{(r)}\|^2}{\|A_0\|^2},$$

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

The electron densities of the considered molecules are first computed in the form

$$\rho(x) := \sum_{a=1}^{K} \psi_a^2(x) = \sum_{a=1}^{K} \left( \sum_{k=1}^{R_0} P_{k,a}(x)e^{\lambda_k(x-x_k)^2} \right)^2,$$  \hspace{1cm} (2.6)

where $K$ is the number of electron pairs and $R_0$ is the number of GTO basis functions given by expansion (2.4) for orbitals $\psi_a(x)$. For each particular molecule we use the following
Figure 2.2: Approximation error in $E_{FN}$ and $E_C$ norms versus Tucker rank for the electron densities of a) CH$_4$, b) C$_2$H$_6$ and c) H$_2$O with $n = 65$.

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$_2$H$_6$ and $A = 10$ bohr, $R_0 = 41$ for H$_2$O.

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

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

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

Figure 2.3: Orthogonal vectors $V^{(1)}_k$, $k = 1, \ldots, 4$, for the rank-10 orthogonal Tucker approximation of the electron density for a) the H atom and b) for the LiH molecule.

We approximate the function (2.6) using the Tucker model. Figure 2.1 gives the approximation error with respect to the Tucker rank $r$ for the LiH electron densities computed by (1.4) and (2.4) for different grid sizes $n^{\circ 3}$. 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 Figure 2.2 verify exponential convergence of the orthogonal
Figure 2.4: 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 $|b_k|$ for the corresponding slice.

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 Figures 2.3 and 2.4 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 Figure 2.4 are the weights $b_{k_1k_2k_3}$ 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 $A_{[k]}$ in (2.2) for $k_t \leq 6$.

Numerical experiments demonstrate high approximation power of the problem-dependent orthogonal tensor-product basis. For example, let $B_{10}$ be the core tensor of the rank $r = 10$ Tucker-type approximation of the electron density for the $CH_4$ molecule, while $B_p$ represents the principal $p \times p \times p$ subtensor of $B_{10}$. We observe that the contribution of the subspace $\text{span}\{V_k^{(\ell)}\}_{k=1}^p$, ($\ell = 1, 2, 3$) to the relative Euclidean norm $\frac{\|B_{10} - B_p\|}{\|B_{10}\|}$, is significant even for small $p$. In fact, already with $p = 3$, the contribution from the first three orthogonal components, e.g., for $1 \leq k \leq 3$, represents the relative $\ell^2$-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.

The electron densities of $CH_4$ and $C_2H_6$ are considered in the cubes $[-7, 7]^3$ and $[-5.8, 5.8]^3$, correspondingly, on a uniform $n^{\otimes 3}$ grid of size $n = 65$. Figures 2.5 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 plane containing two C and two H atoms, correspondingly. Figures 2.5 b) and d) visualise the absolute approximation error for the electron densities $\rho$ of these molecules in the corresponding planes for $r = 16$. In spite of large values of $\rho$ at the cusp regions ($\sim 60$ 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. This is a typical feature
Figure 2.5: Electron densities and absolute errors of rank $r = 16$ Tucker-type approximations for CH$_4$ and C$_2$H$_6$ molecules.

of the orthogonal Tucker decomposition. For H$_2$O with even larger cusp ($\sim 148$ units) at the origin, we see in Figure 2.2 c) that the convergence of $\rho$ for this molecule approximated in the volume $[-10, 10]^3$ gives the relative accuracy $\sim 10^{-5}$ with the rank $r = 16$. Finally, we

Figure 2.6: Approximation error in $E_{FN}$ and $E_C$ norms versus Tucker rank for the Hartree potentials of a) CH$_4$, b) C$_2$H$_6$ and c) H$_2$O.

present the convergence behaviour of the Tucker-type approximation applied to the Hartree potentials of CH$_4$, C$_2$H$_6$ and H$_2$O molecules (see Figure 2.6) which indicates exponential convergence in the Tucker rank $r$.

Our numerical results demonstrate the efficiency of low-rank orthogonal Tucker 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 decompositions of all functions and operators involved.

3 Computing the Hartree potential by tensor-product convolution

We consider the computation of the Hartree potential

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

(3.1)

by the discrete multi-dimensional tensor-product convolution on uniform grids in $\mathbb{R}^3$, described in Ref. [24]. The convolution product is defined by

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

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

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 $\{x_m\}$ in $\Omega$, $m \in \mathcal{M} := \{1, \ldots, n+1\}^3$. 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 $\{\phi_i\}$, $i \in \mathcal{I} := \{1, \ldots, n\}^3$, associated with $\omega_{3,n}$, let $f_i = f(y_i)$ be the representation coefficients of $f$ in $\{\phi_i\}$,

$$f(y) \approx \sum_{i \in \mathcal{I}} f_i \phi_i(y),$$

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

$$f \ast g \approx \{w_m\}_{m \in \mathcal{M}}, \quad w_m := \sum_{i \in \mathcal{I}} f_i \int_{\mathbb{R}^3} \phi_i(y)g(x_m - y)dy, \quad x_m \in \omega_3.$$

As a first step, we precompute the coefficients

$$g_i = \int_{\mathbb{R}^3} \frac{\phi_i(y)}{|y|} dy, \quad i \in \mathcal{I}.$$
The coefficient tensor \( \mathcal{G} = \{ g_i \} \in \mathbb{R}^T \) for the Coulomb potential \( \frac{1}{r-y} \) is approximated in the rank-\( R \) CP tensor format (see Section 2) using optimised sinc-quadratures [13, 18, 24], 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 3-th order tensor \( \mathcal{F} = \{ f_i \} \in \mathbb{R}^T \) is approximated either in the rank \( r = (r, r, r) \) Tucker format or via the CP model with tensor rank \( R_2 \).

Following [24, 25], we compute \( w_m \) by copying the corresponding portion of the discrete convolution in \( \mathbb{R}^3 \)

\[
\mathcal{F} * \mathcal{G} := \{ z_j \}, \quad z_j := \sum_{i \in I} f_i g_{i-j+1}, \quad j \in J := \{1, \ldots, 2n-1\}^3,
\]

centred at \( j = n \).

Let \( \mathcal{F} \) be represented in the Tucker format (2.2). Then (3.2) is represented in the tensor-product form

\[
\mathcal{F} * \mathcal{G} = \sum_{k=1}^{R} \sum_{m=1}^{r} c_k b_{m_1 \ldots m_3} \left( U_k^{(1)} * V_k^{(1)} \right) \otimes \ldots \otimes \left( U_k^{(3)} * V_k^{(3)} \right).
\]

![Figure 3.1: 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).](image)

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

\[
N_{CT} = O(3r Rn \log n + R r^3).
\]
Likewise, approximating $\mathcal{F}$ in the rank-$R_1$ CP format (see (2.3)) enables us to compute $\mathcal{F} * \mathcal{G}$ in the form

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

which leads to the cost

$$
N_{CaC} = O(RR_1n \log n).
$$

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 et al. [3]. This enables us to achieve the desired accuracies on grids $\omega_{3,n}$ 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, 17, 24, 25].

The pseudo-density for the CH$_4$ molecule obtained from a Hartree-Fock calculation is given by (2.6), where $K = 4$ is the number of pseudo-orbitals, $R_0 = 50$ and polynomial degree $P_{k,i} \leq 2$. In the following we represent the convolving tensor $\rho(x)$ 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

$$
R_2 \leq \frac{R_0 (R_0 + 1)}{2} = 1275. \quad (3.4)
$$

We approximate $V_H(x)$, $x \in \omega_{3,n}$, in (3.1) via a tensor-product collocation convolution scheme by $V_H^{(n)}$ on the $n^{\otimes 3}$ tensor-grid $\omega_{3,n}$ in the reference cube $[-A,A]^3$, $A = 10.6$ bohr. The left part of Figure 3.1 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 $\rho(x)$ (see [24] for detailed error estimates). Following [24], we then apply the Richardson extrapolation technique to obtain higher accuracy approximations of order $O(h^3)$ without extra computational cost. The corresponding Richardson extrapolant $V_{H, Rich}^{(n)}$ approximating $V_H(x)$ over a pair of nested grids $\omega_{3,n}$ and $\omega_{3,2n}$, and defined on the “coarse” $n^{\otimes 3}$-grid, is given by

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

The right part of figure 3.1 shows the effect of the Richardson extrapolation on a sequence of grids $\omega_{3,n}$ 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 $\omega_{3,n}$ 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 [25]: first, we compute the rank-$r$ Tucker decomposition of the input tensor and build the corresponding rank-$r^2$ CP representation (with $r = 15$); second, we recompress it to the reduced rank $R'_2 = 120 < r^2$ using the slicewise SVD of small $r \times r$ matrices. Table 3.1 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 $N_{CC} = O(R'_2R_1n\log n)$ described above. We recall that $R_1 \in [10, 20]$ denotes the tensor rank of the Coulomb coefficient tensor in $\mathbb{R}^3$.

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 $n \leq 10^4$), 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 ($n \geq 128$). But even more important, it has much less restrictive memory consumptions of order $O(RR_1n\log n)$ 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 $\sim 500.0$ sec in this case.

## 4 Coulomb matrix and Hartree-Fock energy computation

In the following, we apply the tensor-product scheme [25] 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 normalized “Cartesian Gaussians” $\{g_k\}$ is given by

$$J_{km} := \int g_k(\mathbf{x})g_m(\mathbf{x})V_H(\mathbf{x})d\mathbf{x}, \quad k, m = 1, \ldots, R_0. \quad (4.1)$$

Figure 4.1 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 Figure 4.1 b). Results on high accuracy calculations based on Richardson extrapolation are illustrated in Figure 4.2.

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 $\mathcal{M} = \{\mathcal{M}_{km}\} = \{(g_k, g_m)_{L^2}\}$ ($k, m = 1, \ldots, R_0$)
Figure 4.1: 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.

be the corresponding overlap matrix. We substitute the Galerkin matrix $\mathcal{J} = \{J_{km}\}_{k,m=1}^{R_0}$ from (4.1), 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.

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

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

for the first $K = N/2$ smallest eigenvalues $\lambda_a$ (for the CH$_4$ molecule with pseudopotential $K = 4$). For this computation we apply the MOLPRO package. The total energy is calculated by

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

with $\tilde{J}_a = (\psi_a, V_H \psi_a)_{L^2}$ and $\tilde{R}_a = (\psi_a, V_{ex} \psi_a)_{L^2}$ ($a = 1, \ldots, N/2$) being the so-called Coulomb and exchange integrals, respectively, computed with respect to the orbitals $\psi_a$ ($a = 1, \ldots, N/2$) (compare with (4.1)). The corresponding errors in the Hartree-Fock energy $E_{HF}$ and of individual eigenvalues for different grids are given in Table 4.1.

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 $\lambda_a$, $a = 1, \ldots, K$. 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

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

Our numerical results presented in Table 4.2 indicate 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
Figure 4.2: 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 $(400, 800)$.

Table 4.1: Error of the Hartree-Fock energy $E_{HF}$ (hartree) and of individual eigenvalues $\lambda_i$ computed on different grids, denoted by $\delta_n$, 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.

<table>
<thead>
<tr>
<th></th>
<th>Exact value</th>
<th>$\delta_{112}$</th>
<th>$\delta_{224}$</th>
<th>$\delta_{448}$</th>
<th>$\delta_{112,224}$</th>
<th>$\delta_{224,448}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>-0.95072154</td>
<td>0.0148</td>
<td>3.73 x 10^{-4}</td>
<td>9.37 x 10^{-4}</td>
<td>1.53 x 10^{-5}</td>
<td>6.05 x 10^{-6}</td>
</tr>
<tr>
<td>$\lambda_2, \lambda_3, \lambda_4$</td>
<td>-0.54457593</td>
<td>0.0121</td>
<td>3.03 x 10^{-3}</td>
<td>7.62 x 10^{-4}</td>
<td>1.57 x 10^{-5}</td>
<td>5.21 x 10^{-6}</td>
</tr>
<tr>
<td>$\bar{E}_{HF}$</td>
<td>-7.84226746</td>
<td>0.0113</td>
<td>2.86 x 10^{-2}</td>
<td>7.19 x 10^{-3}</td>
<td>1.08 x 10^{-4}</td>
<td>3.53 x 10^{-5}</td>
</tr>
</tbody>
</table>

4.2 illustrates the validity of the error estimator $E_2(E_n)$ based on the discrete $\| \cdot \|_2$-norm evaluation,

$$E_2(E_n) = \| E_n \|_2.$$  

Analysis of both estimators can be found, for example, in [28]. In particular, for each individual eigenvalue we have

$$|\lambda_a - \lambda_a^{(n)}| \leq E_2(E_n), \quad |\lambda_a - \lambda_a^{(n)}| \leq \text{cond}(P)E_1(E_n),$$

where the matrix $P$ is defined from the diagonalisation procedure $\mathcal{M}^{-1/2}\mathcal{F}\mathcal{M}^{-1/2} = PDPT^T$ with the diagonal matrix $D$. In our case we have $\text{cond}(P) \approx 3.46$. In Table 4.1 we present the approximation errors $\delta_n$ and the corresponding extrapolated errors $\delta_{(n,2n)}$ for the single eigenvalue $\lambda_1$ and the three degenerate eigenvalues $\lambda_2 = \lambda_3 = \lambda_4$.

From Table 4.2 we observe that the error estimators $E_2(E_n)$ and $E_1(E_n)$ can be utilised for an a priori error estimation to the total energy based only on precomputed error bounds to the Fock matrix.
Table 4.2: Comparison of error estimators $\mathcal{E}_1$ and $\mathcal{E}_2$ for Richardson extrapolated Coulomb matrix elements.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$(112,224)$</th>
<th>$(224,448)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{E}_1(E_n)$</td>
<td>$1.12\cdot 10^{-4}$</td>
<td>$3.74\cdot 10^{-5}$</td>
</tr>
<tr>
<td>$\mathcal{E}_2(E_n)$</td>
<td>$6.9\cdot 10^{-5}$</td>
<td>$2.40\cdot 10^{-5}$</td>
</tr>
</tbody>
</table>

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 approximand fits the initial electron density equally well at the near- and far-field regions of the nuclear cusps as shown in Figures 25. Contrary to the orthogonal decomposition, the accuracy of the canonical approximation usually noticeably deteriorates near the nuclear cusps [6]. 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 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 with $h = O(n^{-1})$, and requires $O(3rn + r^3)$ storage, where $r$ denotes the Tucker rank of the electron density with $r \ll n$ almost uniformly in $n$ (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 $n^{3\times3}$ 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 towards 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 $n$ of a 3D Cartesian grid.

Acknowledgments. The authors gratefully acknowledge Prof. W. Hackbusch (Leipzig) and Prof. R. Schneider (Berlin) for useful discussions and valuable comments. This work was partially supported by the DFG (SPP 1145).
References


17


