

On tensor approximation of Green iterations for Kohn-Sham equations

Boris N. Khoromskij

Received: 15 October 2007 / Accepted: 4 February 2008 / Published online: 21 March 2008
© The Author(s) 2008

Abstract In the present paper we discuss efficient rank-structured tensor approximation methods for 3D integral transforms representing the Green iterations for the Kohn-Sham equation. We analyse the local convergence of the Newton iteration to solve the Green's function integral formulation of the Kohn-Sham model in electronic structure calculations. We prove the low-separation rank approximations for the arising discrete convolving kernels given by the Coulomb and Yukawa potentials $1/|x|$, and $e^{-\lambda|x|}/|x|$, respectively, with $x \in \mathbb{R}^d$. Complexity analysis of the nonlinear iteration with truncation to the fixed Kronecker tensor-product format is presented. Our method has linear scaling in the univariate problem size. Numerical illustrations demonstrate uniform exponential convergence of tensor approximations in the orthogonal Tucker and canonical formats.

1 Introduction

Primarily, low rank tensor decomposition methods were applied in chemometrics and independent component analysis and further spread on image processing, population modelling, mathematical biology and financial mathematics. They were generally treated as methods for statistical data processing [32, 35] with moderate accuracy requirements.

Beginning from [2, 20] the idea of tensor approximation of operators and functions has lead to the powerful com-

putational tools in large-scale problems of computational physics. Recent papers show that this technique is surprisingly efficient in multi-dimensional computations, where the traditional methods fail due to the “curse of dimensionality”. In particular, we mention results for many-particle models based on the electronic Schrödinger [3, 13, 31] and Hartree-Fock [4, 5, 22] equations in electronic structure calculations, for the Ornstein-Zernicke equation in molecular density simulations [8], and for the deterministic Boltzmann equation [25] for dilute gas.

The Hartree-Fock/Kohn-Sham equations in 3D provide a mean-field approximation for the ground state of many-electron systems initially described by the high-dimensional electronic Schrödinger equation. Methods based on the Hartree-Fock model (wavefunction methods) are preferred if the precise simulation of small size systems is required. In turn, the density functional methods based on the Kohn-Sham model can be used for numerical simulation of large molecular systems (say, proteins) as well as in the solid state physics.

In this paper, we discuss the methods of low-separation rank representations of Green iterations for the Kohn-Sham equation, which scale linearly in the univariate problem size. The techniques can be applied as well to the Hartree-Fock model. The main results presented below can be formulated as follows:

- Proof of the local convergence of the Newton iterations for the Green function formulation of the Kohn-Sham equation.
- Proof of the low-separation rank approximations for the discrete convolving kernels given by the Coulomb and Yukawa potentials $1/|x|$, and $e^{-\lambda|x|}/|x|$, respectively, with $x \in \mathbb{R}^3$.

Communicated by S. Sauter.

Dedicated to Wolfgang Hackbusch on the occasion of his 60th birthday.

B. N. Khoromskij (✉)
Max-Planck-Institute for Mathematics in the Sciences,
Inselstr. 22-26, 04103 Leipzig, Germany
e-mail: bokh@mis.mpg.de

- Complexity analysis of the nonlinear iteration with truncation to fixed Kronecker tensor-product format.
- Numerical illustrations.

Tensor approximation methods form the main building blocks of our approach. Principal advantages of the tensor-product representations of operators and functions are sublinear storage and complexity of multi-linear algebra (MLA) operations (say, matrix-vector and matrix-matrix products, matrix inversion, convolution product, functions of matrices) applied to higher-order tensors. The main difficulties of numerical MLA arise from the fact that *tensor approximation methods* are actually equivalent to *nonlinear approximations* in higher dimensions.

There are *algebraic, analytically-based and combined* strategies for computing the low tensor-rank approximations of a higher-order tensor:

- *Algebraic methods* are the most general ones. The common approach is to derive the univariate components of separable terms by straightforward minimisation of the quadratic *cost functional* (see (17) below).
- *Analytically-based approximation* methods are efficient for a special class of *function-related* tensors (see the description of sinc-methods in Sect. 3.5).
- *Combined* methods are designed to take advantage of both algebraic and analytic approaches, and, at the same time, to relax their limitations (cf. [24]). In this way, the initial *analytic approximations* can be optimised via the algebraic rank recompression.

Analytic *sinc*-quadrature based and combined methods will be in the focus of the present paper.

Efficient solution process for the Hartree-Fock/Kohn-Sham equations can be performed by approximate nonlinear iterations via the controlled tensor rank [21]. In this context, understanding of the truncated MLA operations addresses the following issues: existence of low-rank approximations, construction of a good initial guess, fast algebraic methods based on rank truncation. Notice that in the d -dimensional setting, one can apply the recent *multigrid tensor approximation method* [28], which scales linearly in all essential parameters, thus getting rid of the curse of dimensionality.

It is worth to note that the complexity of low-separation rank approximations of functions and operators scales linearly in the univariate problem size, allowing to avoid technically involved high-order approximations in the space variables. Hence, low-order Galerkin ansatz spaces, say, by piecewise constant/linear basis functions, can be applied.

The rest of the paper is organised as follows. In Sect. 2 we discuss the Green function based integral formulation of the Kohn-Sham equation. The main result of this Section is the convergence theory for the Newton iteration. Sect. 3 proves

the low tensor-rank approximations of arising convolution operator. In Sect. 4 we describe the discretisation schemes, discuss the complexity issues and present numerical illustrations.

We mention that comparative analysis of the current approach with other commonly used methods to compute invariant subspaces, say, based on the Krylov-subspace-Lanczos algorithms or the Davidson methods, will be the topic of a separate paper.

2 Green function based integral formulation

2.1 Basic equations

The Hartree-Fock equation for determination of the ground state of a molecular system consisting of M nuclei and N electrons is a self-consistent (nonlinear) eigenvalue problem in $L^2(\mathbb{R}^3)$

$$\mathcal{F}_\Phi \phi_i(x) = \lambda_i \phi_i(x), \quad \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}, \quad i, j = 1, \dots, N \quad (1)$$

with \mathcal{F}_Φ being the nonlinear Fock operator

$$\mathcal{F}_\Phi := -\frac{1}{2}\Delta - \sum_{v=1}^M \frac{Z_v}{|\cdot - a_v|} + V_H - \sum_{j=1}^N \left(\cdot \phi_j \star \frac{1}{|\cdot|} \right) \phi_j,$$

with the Hartree potential defined by

$$V_H(x) := \left(\rho \star \frac{1}{|\cdot|} \right) (x),$$

where \star denotes the convolution product in $L^2(\mathbb{R}^3)$, and Z_v, a_v ($v = 1, \dots, M$) specify charges and positions of M nuclei. The electron density $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$ is given by

$$\rho(x) = \sum_{i=1}^N |\phi_i(x)|^2, \quad x \in \mathbb{R}^3.$$

The last term in the Fock operator \mathcal{F}_Φ (the so-called exchange potential) is the most complicated part for the numerical treatment. In fact, it defines the nonlocal interaction

$$(V_x \psi)(x) = \sum_{i=1}^N \int_{\mathbb{R}^3} \frac{\phi_i(x) \phi_i(x')}{|x - x'|} \psi(x') dx',$$

that includes the convolution with the *density matrix*

$$\tau(x, x') := \sum_{i=1}^N \phi_i(x) \phi_i(x').$$

It can be proven that solutions ϕ_i to (1) are smooth functions having a cusp (singularity) at each point nucleus, and in addition, they decrease exponentially fast at infinity. The density ρ is an exponentially decreasing function at infinity, as well

(see the review paper [30]). Further regularity results for the electronic Schrödinger and Hartree-Fock equations can be found in [10,37].

In the Kohn-Sham model one deals with the system of Eq. (1), where the nonlocal operator \mathcal{F}_ϕ is substituted by a local one,

$$\mathcal{F}_\rho := -\frac{1}{2}\Delta - \sum_{v=1}^M \frac{Z_v}{|\cdot - a_v|} + V_H + v_{xc}(\rho) \tag{2}$$

with $v_{xc}(\rho)$ being a scalar function that depends only upon $\rho(x)$. One possible choice is $v_{xc}(\rho) = c\rho^{1/3}$.

A standard method to solve the system of nonlinear Eq. (1) is based on the two-level iteration which includes:

- (a) the so-called *self-consistent field* (SCF) algorithms, i.e., the iterations on the nonlinearity, which serves to solve the initial nonlinear problem,
- (b) at each SCF cycle, computation of the spectral projection to build the update of electron density/density matrix, using the current (frozen) discretisation to the Fock or Kohn-Sham operators.

Hence, the crucial point for the efficient performance of nested iterations (a), (b) would be the multiple computation of N -dimensional invariant subspace of fixed linear operators \mathcal{F}_ϕ or \mathcal{F}_ρ .

2.2 Green’s function formulation

In the present notes we focus on the solution of the spectral problem at step (b) of SCF iteration in the set of low-separation rank functions. We will consider the eigenvalue problem of the form

$$H\phi := [-\frac{1}{2}\Delta + V]\phi = \lambda\phi, \tag{3}$$

posed in $L^2(\mathbb{R}^3)$, where the “interaction” potential V ensures the existence of the discrete spectrum that belongs to $(-\infty, 0)$, and also it allows the low-separation rank representation in univariate variables in the sense which will be specified later on. To avoid unessential technicalities, below, we consider the Kohn-Sham model, that is represented by the local multiplication operator with the potential

$$V := -\sum_{v=1}^M \frac{Z_v}{|\cdot - a_v|} + \left(\rho \star \frac{1}{|x|}\right) + v_{xc}(\rho). \tag{4}$$

However, the Hartree-Fock equation can be treated with minor modifications.

For the ease of presentation, we consider the computation of the minimal eigenvalue

$$\lambda^* = \min(\sigma(-\frac{1}{2}\Delta + V))$$

and the corresponding eigenfunction ϕ^* of the problem (3). Introducing the elliptic resolvent (the Green’s function operator)

$$R_z = (-\Delta + z^2 I)^{-1}$$

with the kernel function defined by the Yukawa potential

$$G_z(x) := \frac{e^{-z|x|}}{4\pi|x|}, \quad z \in (0, \infty), \tag{5}$$

and setting $z = \sqrt{-2\lambda}$, we obtain the equivalent formulation to the eigenvalue problem (3)

$$\phi = \mathcal{G}_z\phi \quad \text{with} \quad \mathcal{G}_z := -2(V \cdot) \star G_z \equiv -2R_z V \tag{6}$$

with the compact operator \mathcal{G}_z . This formulation was first introduced in [23]. The important feature of this approach is that any eigenvalue-eigenfunction pair, (λ, ϕ) , (λ in the discrete spectrum), for the operator H generates a fixed point solution of the problem (6) (cf. [22]).

Numerical algorithms for solving the integral Eq. (6) are based on considering the eigenvalue problem for the Lippmann-Schwinger type parametric integral operator

$$\mu_z\phi_z = \mathcal{G}_z\phi_z \quad \text{for} \quad z \in (0, \infty), \tag{7}$$

where both $\mu_z = \mu(z)$ and ϕ_z depend on the parameter z .

Lemma 1 *Suppose that the “exchange part” $v_{xc}(\rho)$ in the Kohn-Sham potential (4) satisfies assumption*

$$v_{xc}(\rho)(x) = V_1(x) + V_2(x)$$

with

$$V_1 \in L^2(\mathbb{R}^3), \quad V_2 \in L^\infty(\mathbb{R}^3),$$

where V_2 can be taken to be arbitrarily small in the L^∞ sense.

Then $\mu_{\lambda^*} = 1$ is the largest eigenvalue of \mathcal{G}_{λ^*} , and the corresponding single eigenfunction $\phi_{\lambda^*} = \phi_*$ is the desired ground-state eigenfunction of (3).

Proof The proof is a slight modification of respective arguments in the proof of Theorem 1.1, [31]. This includes the relation $\|\mathcal{G}_{\lambda^*}\| = 1$ in Proposition 3.1, [31], which can be similarly justified in our case. □

2.3 Analysis of the Newton iteration

Lemma 1 implies that finding the eigenvalue λ^* can be reduced to iterative solution of the scalar nonlinear equation

$$\mu(z) = 1, \quad z \in \mathbb{R}_+ \tag{8}$$

with an initial guess $z_0 \in \mathbb{R}_+$ that belongs to the attraction basin of λ^* (we never start electronic structure calculations from scratch).

Remark 1 Other eigenvalues/eigenfunctions may be obtained by deflation which is used to recast the integral equation for each orbital as a ground-state problem (see [22] for more details). Let \mathcal{P}_m be the orthoprojection onto the space of eigenfunctions of lower energy than orbital m . Then the m th occupied orbital ψ_m will be the lowest energy solution of

$$(I - \mathcal{P}_m)H(I - \mathcal{P}_m)\psi_m = \lambda_m\psi_m,$$

which leads to the modified integral formulation

$$\psi_m = -2R_z(V + \mathcal{P}_m H(I - \mathcal{P}_m))\psi_m.$$

Several iteration schemes were considered in the literature to solve the nonlinear problem (6). In particular, in the case of Schrödinger equation, the convergence of the power method in the form

$$\phi_{n+1} = \frac{\mathcal{G}_{z_n}\phi_n}{\|\mathcal{G}_{z_n}\phi_n\|}, \quad z_{n+1} = \langle H\phi_{n+1}, \phi_{n+1} \rangle$$

was analysed in [31]. Under the corresponding assumptions, the power method can be applied to the Kohn-Sham equation as well (see Theorem 1). Newton’s type iteration in the form

$$\begin{aligned} \tilde{\phi}_n &= \mathcal{G}_{z_n}\phi_n, \quad \phi_{n+1} = \tilde{\phi}_n / \|\tilde{\phi}_n\|, \\ z_{n+1} &= z_n - \langle V\phi_n, \phi_n - \tilde{\phi}_n \rangle / \|\tilde{\phi}_n\|^2, \end{aligned}$$

was applied in [22] to the Kohn-Sham equation defined by the potential (4).

The following result provides sufficient conditions for the quadratic (local) convergence of the Newton iteration applied directly to Eq. (8),

$$\text{given } z_0 : \quad z_{n+1} = z_n - \frac{\mu(z_n) - 1}{\frac{\partial \mu}{\partial z}(z_n)}, \quad n = 0, 1, \dots \quad (9)$$

Particular realization of the iteration (9) is determined by approximations chosen for $\mu(z_n)$ and $\frac{\partial \mu}{\partial z}(z_n)$.

Theorem 1 (I) *Under assumptions in Lemma 1 the power method converges geometrically.*

(II) *Let the exact eigenvalue-eigenfunction pair of (3), (λ_*, ϕ_*) , satisfy*

$$\langle (V\phi_*) \star e^{-z_*|x|}, \phi_* \rangle \neq 0, \quad z_* = \sqrt{-2\lambda_*}. \quad (10)$$

Then (a) the Newton iteration (9) converges (locally) quadratically, and (b) the quasi-Newton iteration defined by the approximation

$$\frac{\partial \mu}{\partial z}(z_n) \approx \langle (V\phi_n) \star e^{-z_n|x|}, \phi_n \rangle,$$

converges (locally) quadratically as well.

Proof First, we note that G_z is an analytic family in z , hence $(V \cdot) \star G_z$ is an analytic family and thus its eigenvalue μ_z and

eigenfunction ϕ_z depend analytically on z . Differentiating Eq. (7) at $z = z_*$, we obtain

$$\frac{\partial \mu_z}{\partial z} \phi_* + \frac{\partial \phi_z}{\partial z} = \frac{\partial \mathcal{G}_z}{\partial z} \phi_* + \mathcal{G}_{z_*} \frac{\partial \phi_z}{\partial z}.$$

Scalar multiplication of this equation with ϕ_* and taking into account the symmetry of \mathcal{G}_{z_*} and relations

$$\mathcal{G}_{\lambda_*} \phi_* = \phi_*, \quad \langle \phi_*, \phi_* \rangle = 1,$$

leads to

$$\frac{\partial \mu_z}{\partial z} \Big|_{z=z_*} = \left\langle \frac{\partial \mathcal{G}_z}{\partial z} \Big|_{z=z_*} \phi_*, \phi_* \right\rangle. \quad (11)$$

Moreover, the direct computation shows that

$$\frac{\partial \mathcal{G}_z}{\partial z} \Big|_{z=z_*} = (V \cdot) \star e^{-z_*|x|}. \quad (12)$$

This proves (a).

Furthermore, due to analyticity of G_z in z , and in view of (11) and (12), we conclude that the approximation of $\frac{\partial \mu_z}{\partial z}$ (in general noncomputable) given by

$$\frac{\partial \mu_z}{\partial z}(z_n) \approx \frac{\partial \mu_z}{\partial z}(z_*) = \langle (V\phi_*) \star e^{-z_*|x|}, \phi_* \rangle$$

ensures the local quadratic convergence. The desired computable approximation can be proven by perturbation argument. Indeed, the functional $f : \mathbb{R}_+ \times L^2(\mathbb{R}^3) \rightarrow \mathbb{R}$,

$$f(z, \phi) := 1 / \langle (V\phi) \star e^{-z|x|}, \phi \rangle, \quad z > 0,$$

is continuously differentiable at (z_*, ϕ_*) , hence, in the small vicinity of (z_*, ϕ_*) , we have

$$f(z_n, \phi_n) = 1 / \langle (V\phi_*) \star e^{-z_*|x|}, \phi_* \rangle + \zeta_n, \quad \zeta_n \in \mathbb{R},$$

where

$$|\zeta_n| \leq C(|z_* - z_n| + \|\phi_* - \phi_n\|).$$

This completes the proof of part (b). □

In this paper, we focus on the efficient implementation of the Lippmann-Schwinger type operator \mathcal{G}_z for $z > 0$, which constitutes the computational kern of the Green iterations in both power method and Newton’s scheme.

Example 1 The non-degeneracy condition (10) can be justified in the case of Hydrogen atom/ion with $V(x) = C/|x|$ and $\phi_*(x) = e^{-\mu_*|x|}$ with $\mu_* > 0$ (say, $\mu_* = 1/2$). In fact, using representation in spherical coordinates, it is easy to check that

$$\left\langle \frac{e^{-\mu_*|x|}}{|x|} \star e^{-\mu_*|x|}, e^{-\mu_*|x|} \right\rangle \neq 0.$$

3 Low tensor-rank representation of operators

3.1 Main motivations

We are interested in the low tensor-rank approximation of the Green’s function operator \mathcal{G}_z in Eq. (6), which can be presented (up to the scaling factor) in one of the following forms

$$\mathcal{G}_z = (V \cdot) \star G_z \equiv R_z V. \tag{13}$$

The first one is well suited for the integral representation of Eq. (6), while the second form, that contains the elliptic resolvent, can be applied to the FE/FD discretisations of the initial Eq. (3) posed on the bounded domain. We refer to [12, 17, 19] concerning the low-rank tensor approximation methods for the elliptic resolvent operators.

To fix the point, in the following we discuss the method of fast tensor-product convolution of the Yukawa kernel and the action of interaction potential, $(V \phi) \star G_z$, involved in the integral representation of Eq. (6). Our approach is based on the ideas from [15, 26] applied to the collocation convolution schemes. The main advantages of the integral representation $\mathcal{G}_z = (V \cdot) \star G_z$ are the following:

- Compactness of the operator \mathcal{G}_z .
- Applicability of simple collocation schemes with discontinuous basis functions (L^2 setting).
- Possibility to compute the total energy using only integral operators (avoid the application of Δ that requires at least H^1 basis functions).
- Existence of low-separation rank approximations to the operators V and G_z .
- Nonlinear (say, Newton’s) quadratically convergent iteration, which is well suited for the numerical multi-linear algebra via truncation to fixed tensor format.

Notice that for the sake of generality, most constructions in the next section will be described in the d -dimensional case with the arbitrary dimension parameter d , while the basic equations in electronic structure calculations are formulated for the case $d = 3$.

3.2 Tucker-type and canonical models

We consider linear space \mathcal{H} of real-valued d -th order tensors

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

equipped with the Euclidean inner product

$$\langle \mathcal{A}, \mathcal{B} \rangle := \sum_{(i_1, \dots, i_d) \in \mathcal{I}} a_{i_1 \dots i_d} b_{i_1 \dots i_d},$$

inducing the Frobenius norm $\|\mathcal{A}\| := \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$. One needs $\prod_{\ell=1}^d n_\ell$ units to store the multi-dimensional array \mathcal{A} . The outer product of vectors forms the canonical rank-1 tensor

$$\mathcal{U} \equiv [u_i]_{i \in \mathcal{I}} = U^{(1)} \otimes \dots \otimes U^{(d)} \in \mathbb{R}^{\mathcal{I}}$$

with the entries

$$u_{i_1 \dots i_d} = u_{i_1}^{(1)} \cdots u_{i_d}^{(d)}, \tag{14}$$

where

$$U^{(\ell)} \equiv [u_{i_\ell}^{(\ell)}]_{i_\ell \in I_\ell} \in \mathbb{R}^{I_\ell} \quad \ell = 1, \dots, d.$$

The storage requirements are reduced to $\sum_{\ell=1}^d n_\ell$. In the case $d = 2$, the outer product of vectors represents a rank-1 matrix.

Rank- (r_1, \dots, r_d) Tucker-type decomposition approximates the tensor \mathcal{A} by a sum of rank-1 tensors

$$\mathcal{A}_{(\mathbf{r})} = \sum_{k_1=1}^{r_1} \cdots \sum_{k_d=1}^{r_d} b_{k_1 \dots k_d} \cdot U_{k_1}^{(1)} \otimes \dots \otimes U_{k_d}^{(d)}, \tag{15}$$

where $\mathbf{r} = (r_1, \dots, r_d)$, so that $r = \max_{\ell} \{r_\ell\}$ is called the Tucker rank. The coefficients tensor $\mathcal{B} = [b_{k_1 \dots k_d}] \in \mathbb{R}^{r_1 \times \dots \times r_d}$ is usually called the core tensor. Without loss of generality we conventionally assume that the Tucker components $\mathbf{U}^{(\ell)} = [U_1^{(\ell)} U_2^{(\ell)} \dots U_{r_\ell}^{(\ell)}]$ are orthogonal matrices of the respective size $n_\ell \times r_\ell$, that is

$$\mathbf{U}^{(\ell)T} \mathbf{U}^{(\ell)} = I_{r_\ell \times r_\ell} \quad (\ell = 1, \dots, d).$$

We denote by $\mathcal{T}_{\mathbf{r}}$ the set of tensors represented in the Tucker-type format (15).

Figure 1 represents the orthogonal components $U_k^{(3)}$ ($k = 1, \dots, 5$) from the rank-5 Tucker approximation of the Yukawa potential on the $n \times n \times n$, $n = 65$, grid (see

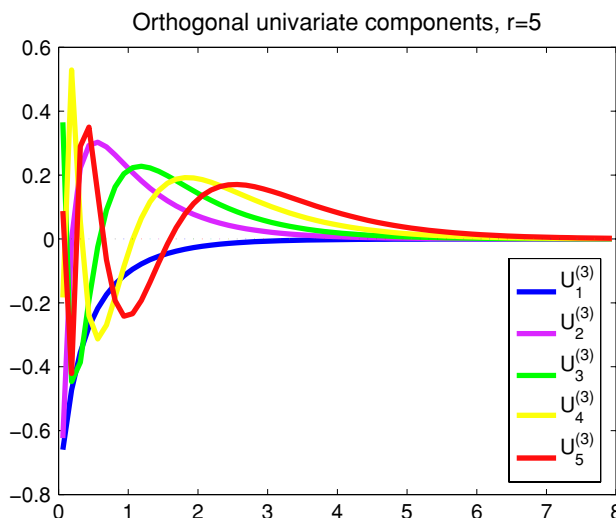


Fig. 1 Orthogonal univariate components $U_k^{(3)}$ ($k = 1, \dots, 5$) for Tucker approximation of the Yukawa potential

comments in Sects. 3.6, 4.1). This components resolve both the singularity at the origin and exponential decay at the infinity. By the construction, the set of functions $U_k^{(\ell)}$ ($\ell = 1, \dots, d$) represents the best problem adapted orthogonal basis, computed by the nonlinear approximation.

The canonical (CP) decomposition of a tensor is defined by

$$\mathcal{A}_{(R)} = \sum_{k=1}^R b_k \cdot U_k^{(1)} \otimes \dots \otimes U_k^{(d)}, \quad b_k \in \mathbb{R} \tag{16}$$

with normalised components $U_k^{(\ell)}$ ($\ell = 1, \dots, d$). The minimal parameter R in the above representation is called the rank (or canonical rank) of a tensor. The set of tensors which can be represented in the CP format (16) will be denoted by \mathcal{C}_R .

We denote by $\#\mathcal{S}(\mathcal{A})$ the number of nonzero elements corresponding to the sparsity pattern $\mathcal{S}(\mathcal{A})$ of \mathcal{A} . In particular, for $\mathcal{A}_{(R)} \in \mathcal{C}_R$ we have for the corresponding core tensor $\mathcal{B} = \text{diag}\{b_1, \dots, b_R\}$, $\#\mathcal{S}(\mathcal{B}) = R$, while in general for $\mathcal{A}_{(r)} \in \mathcal{T}_r$ one obtains $\#\mathcal{S}(\mathcal{B}) = \prod_{\ell=1}^d r_\ell$.

3.3 Numerical multi-linear approximations

Since both \mathcal{T}_r and \mathcal{C}_R are not linear spaces, we obtain a severe nonlinear approximation problem as soon as we want to estimate

$$\sigma(\mathcal{A}, \mathcal{S}) := \inf_{s \in \mathcal{S}} \|\mathcal{A} - s\|, \tag{17}$$

for $\mathcal{A} \in \mathcal{H}$, where either $\mathcal{S} = \mathcal{M}_r$ or $\mathcal{S} = \mathcal{C}_R$.

As it was already mentioned, there are algebraic, analytically-based and combined strategies for computing a Kronecker tensor-product decomposition of a higher-order tensor.

The most commonly used methods to compute the Tucker decomposition are based on direct minimisation in (17) via the alternating least-squares (ALS) iteration. We refer to [6] concerning the orthogonal rank- (r_1, \dots, r_d) Tucker decomposition. Special versions of the ALS iteration adapted to the format of input data presented either by the Tucker or canonical models are discussed in [27].

For $\mathcal{S} = \mathcal{C}_R$, the approximation problem (17) can be considered in the framework of best r -term approximation with regard to a redundant dictionary (cf. [34]). The difficulties in the rigorous analysis and efficient implementation of the minimisation process in \mathcal{C}_R are due to

- (a) multiple local minima of the cost functional,
- (b) degeneracy of the minimising sequence (see examples in [24]),
- (c) high-dimensional nonlinear optimisation.

In the case of canonical decomposition one can find local minima in (17) via the ALS algorithm as follows: with b_k from (16) define $B = \text{diag}\{b_1, \dots, b_d\}$ and for any fixed $m \in \{1, \dots, d\}$ assume that all matrices $\mathbf{V}^{(\ell)}$, $\ell \neq m$, are fixed. Then (17) is a quadratic expression in the components of the matrix $\mathbf{V}^{(m)} \cdot B$, hence we obtain a classical least-squares problem. The ALS iteration repeats this procedure for all $m = 1, \dots, d$ until convergence (or termination). The components of $B = \text{diag}\{b_1, \dots, b_d\}$ are obtained by normalisation of the columns $V_k^{(m)}$ ($k = 1, \dots, R$).

A second approach is based on the Newton-type algorithms applied to the Lagrange equation corresponding to the unconstrained minimisation problem: Find $\mathcal{A} \in \mathcal{C}_R$ and the Lagrange multipliers $\lambda^{(k,\ell)} \in \mathbb{R}$ such that

$$\langle \mathcal{A} - \mathcal{A}_0, \mathcal{A} - \mathcal{A}_0 \rangle + \sum_{k=1}^R \sum_{\ell=1}^d \lambda^{(k,\ell)} \left(\|V_k^{(\ell)}\|^2 - 1 \right) \rightarrow \min.$$

An alternative method to implement the CP model is based on a simultaneous generalised Schur decomposition, see [7].

The two-level methods to compute a rank- R CP approximation start from the rank- r Tucker approximation with $r = (r_1, \dots, r_d)$ and then proceed with the rank- R CP decomposition of the small-size core tensor $\mathcal{B} \in \mathbb{R}^{r_1 \times \dots \times r_d}$, which results in the \mathcal{C}_R -representation of the initial tensor (see [7,27] for more details).

3.4 Collocation discretisation of the convolution in \mathbb{R}^d

The multi-dimensional convolution transform of functions $f, g \in L^2(\mathbb{R}^d)$ is given by

$$w(x) := (f * g)(x) := \int_{\mathbb{R}^d} f(y)g(x - y)dy, \quad x \in \mathbb{R}^d.$$

We are interested in approximate computation of $f * g$ in some fixed box $\Omega = [-A, A]^d$. We suppose that the convolving function f has a support in $\Omega' = [-B, B]^d \subset \Omega$ ($B < A$), i.e., $\text{supp } f \subset \Omega'$ (usually, f has the exponential decay). In our applications we have two particular choices

$$g(x) = \frac{1}{|x|} \quad \text{and} \quad f(x) = \rho(x),$$

as well as

$$g(x) = e^{-z|x|}/|x| \quad \text{and} \quad f(x) = V(x)\phi(x),$$

where ρ, ϕ are exponentially decaying functions.

Note that for both the Hartree-Fock and Kohn-Sham equations the interaction potential V already contains the Hartree potential $V_H = \rho \star \frac{1}{|x|}$ (ρ is the electron density), which can be efficiently approximated by the low tensor rank convolution algorithms, to be described below. For this case, the theoretical analysis and numerical results can be found in [5,26].

There are three commonly used discretisation methods of the integral operators: the so-called Nyström, collocation and Galerkin type schemes. The *Galerkin scheme* is known as the most convenient for theoretical analysis. However, compared with the collocation method, it has higher implementational cost since the presence of double integration. In the following, we focus on the collocation methods.

Introduce the equi-distant tensor-product lattice $\omega_{\mathbf{d}} := \omega_1 \times \dots \times \omega_d$ of size $h = 2A/n$ by setting $\omega_\ell = \omega := \{-A + (i - 1)h : i = 1, \dots, n + 1\}$, $\ell = 1, \dots, d$,

where for the sake of convenience $n = 2p$, $p \in \mathbb{N}$. Hence $\Omega = \cup_{\mathbf{i} \in \mathcal{I}} \Omega_{\mathbf{i}}$ becomes the union of closed boxes $\Omega_{\mathbf{i}} = \bigotimes_{\ell=1}^d \Omega_{i_\ell}$ with intervals

$$\Omega_{i_\ell} := \{x_\ell : x_\ell \in [-A + (i_\ell - 1)h, -A + i_\ell h]\} \subset \mathbb{R}$$

for $\ell = 1, \dots, d$.

We consider *collocation discretisation* in the case of tensor-product piecewise constant basis functions $\{\phi_{\mathbf{i}}\}$ associated with $\omega_{\mathbf{d}}$, so that $\phi_{\mathbf{i}} = \chi_{\Omega_{\mathbf{i}}}$ is the characteristic function of $\Omega_{\mathbf{i}}$,

$$\phi_{\mathbf{i}}(x) = \prod_{\ell=1}^d \phi_{i_\ell}(x_\ell), \quad \text{where } \phi_{i_\ell} = \chi_{\Omega_{i_\ell}}. \tag{18}$$

Let $x_{\mathbf{m}} \in \omega_{\mathbf{d}}$ be the set of collocation points with $\mathbf{m} \in \mathcal{M} := \{1, \dots, n + 1\}^d$, and let $f_{\mathbf{i}}$ be the representation coefficients of f in $\{\phi_{\mathbf{i}}\}$,

$$f(y) \approx \tilde{f}(y) := \sum_{\mathbf{i} \in \mathcal{I}} f_{\mathbf{i}} \phi_{\mathbf{i}}(y).$$

In the following, we specify the coefficients as $f_{\mathbf{i}} = f(y_{\mathbf{i}})$, where $y_{\mathbf{i}}$ is the midpoint of $\Omega_{\mathbf{i}}$, $\mathbf{i} \in \mathcal{I}$ (the alternative choice would be the exact L^2 -projection to the set of basis functions $\{\phi_{\mathbf{i}}\}$). We consider the following *discrete collocation scheme*

$$f * g \approx \{w_{\mathbf{m}}\}, \quad w_{\mathbf{m}} := \sum_{\mathbf{i} \in \mathcal{I}} f_{\mathbf{i}} \int_{\mathbb{R}^d} \phi_{\mathbf{i}}(y) g(x_{\mathbf{m}} - y) dy \tag{19}$$

with $x_{\mathbf{m}} \in \omega_{\mathbf{d}}$, $\mathbf{m} \in \mathcal{M}$. In our application this scheme approximates the part in the Fock operator defined by the Hartree potential. Pointwise evaluation of this scheme has exponential in d complexity, $O(n^{2d})$ (curse of dimensionality). In the case of uniform grids the computational complexity can be reduced to $O(n^d \log n)$ by applying the multi-dimensional FFT.

Here we present the convolution algorithm of almost linear complexity $O(n \log^q n)$ based on the tensor decomposition of convolving functions. Precompute the collocation coefficients

$$g_{\mathbf{i}} = \int_{\mathbb{R}^d} \phi_{\mathbf{i}}(y) g(-y) dy, \quad \mathbf{i} \in \mathcal{I}, \tag{20}$$

and define the d -th order tensors

$$\mathcal{F} = [f_{\mathbf{i}}], \quad \mathcal{G} = [g_{\mathbf{i}}] \in \mathbb{R}^{\mathcal{I}}.$$

Let us introduce the d -dimensional *discrete convolution* $\mathcal{F} * \mathcal{G} := \{z_{\mathbf{j}}\}$, where

$$z_{\mathbf{j}} := \sum_{\mathbf{i}} f_{\mathbf{i}} g_{\mathbf{j}-\mathbf{i}+1}, \quad \mathbf{j} \in \mathcal{J} := \{1, \dots, 2n - 1\}^d, \tag{21}$$

where the sum is over all $\mathbf{i} \in \mathcal{I}$ which lead to legal subscripts for $g_{\mathbf{j}-\mathbf{i}+1}$, i.e., $\mathbf{j} - \mathbf{i} + 1 \in \mathcal{I}$ for $\mathbf{j} \in \mathcal{J}$, then the discrete collocation scheme $\{w_{\mathbf{m}}\}$, $\mathbf{m} \in \mathcal{M}$, is obtained by copying the corresponding portion of $\{z_{\mathbf{j}}\}$ from (21), centred at $\mathbf{j} = \mathbf{n} = n^{\otimes d}$ (cf. [26]),

$$\{w_{\mathbf{m}}\} = \{z_{\mathbf{j}}\}_{\mathbf{j}=\mathbf{j}_0+\mathbf{m}}, \quad \mathbf{m} \in \mathcal{M}, \quad \mathbf{j}_0 = \mathbf{n}/2.$$

For the collocation method with piecewise constant basis functions and for the input data $f \in C^2$, we are able to prove the $O(h^2)$ -error bound (superconvergence). We recall that continuous Fourier transform in \mathbb{R}^d is given by

$$\mathcal{F}(f)(\kappa) := \int_{\mathbb{R}^d} f(x) e^{-i(\kappa, x)} dx, \quad \kappa \in \mathbb{R}^d.$$

Lemma 2 ([26]) *Let $f \in C^2(\Omega)$ and let $g \in L^1(\Omega)$, where $\Omega \in \mathbb{R}^d$ is a tensor-product domain. Furthermore, we assume that there exist $\mu \geq 1$ and $\beta > 0$, such that*

$$|\mathcal{F}(g)| \leq C/|\kappa|^\mu \text{ as } |\kappa| \rightarrow \infty, \quad \kappa \in \mathbb{R}^d \tag{22}$$

and

$$|\nabla_y g(x - y)| \leq C/|x - y|^\beta \text{ for } x, y \in \Omega, \quad x \neq y.$$

Then there is a constant $C > 0$ independent of h such that

$$|w(x_{\mathbf{m}}) - w_{\mathbf{m}}| \leq Ch^2, \quad \mathbf{m} \in \mathcal{M}. \tag{23}$$

Example 2 To illustrate Lemma 2 we notice that the fundamental solution of the Laplace operator in \mathbb{R}^d for $d \geq 3$, is given by $G(x) = c(d)/|x|^{d-2}$, and $\mathcal{F}(G) = C/|\kappa|^2$. Hence Lemma 2 applies with $\beta = d - 1$, $\mu = 2$.

More refined error analysis justifies the Richardson extrapolation method on a sequence of diadically refined grids providing the better approximation error $O(h^3)$.

Lemma 3 ([26]) *Let $f \in C^3(\Omega)$, then there exists a function $c_1 \in C(\Omega)$ which is independent of h , such that for $\mathbf{m} \in \mathcal{M}$ we have*

$$w(x_{\mathbf{m}}) = w_{\mathbf{m}} + c_1(x_{\mathbf{m}})h^2 + \eta_{\mathbf{m},h}, \quad \eta_{\mathbf{m},h} \in \mathbb{R} \tag{24}$$

with $|\eta_{\mathbf{m},h}| \leq Ch^3$.

The Richardson extrapolation allows substantial improvement of the approximation accuracy using the simple basis set and without an extra cost. Moreover, computation on a sequence of grids optimises the numerical cost of tensor approximations of arising functions.

3.5 Tensor-product convolution on uniform grids

We notice that the convolution product appears to be one of the most computationally elaborate MLA operations since in general one might have for the corresponding core tensors $\#\mathcal{S}(\mathcal{B}) \cdot \#\mathcal{S}(\mathcal{C}) = O(r^{2d})$, where by $\mathcal{S}(\mathcal{B})$ we denote the sparsity pattern of a tensor. Significant complexity reduction is observed if at least one of the convolving tensors can be represented by the CP model.

Theorem 2 *Let $\mathcal{F} \in \mathcal{T}_r, \mathcal{G} \in \mathcal{C}_R$, then the corresponding discrete convolution $\mathcal{F} * \mathcal{G}$ can be computed in the linear in n complexity*

$$\mathcal{N}_{T * C} = O(drRn \log n + Rr^d).$$

Furthermore, if $\mathcal{F} \in \mathcal{C}_{R_1}$, and $\mathcal{G} \in \mathcal{C}_{R_2}$ with the core tensors $\mathcal{B}_{\mathcal{F}} = \text{diag}\{b_1, \dots, b_{R_1}\}$ and $\mathcal{B}_{\mathcal{G}} = \text{diag}\{c_1, \dots, c_{R_2}\}$, respectively, then $\mathcal{F} * \mathcal{G}$ can be computed in lower cost, which scales linearly in both the dimensionality parameter d and the grid-size n (up to a log-factor),

$$\mathcal{N}_{C * C} = O(dR_1R_2n \log n).$$

Proof We tensorize the convolution product as follows (cf. [27])

$$\mathcal{F} * \mathcal{G} = \sum_{k=1}^r \sum_{m=1}^R b_k c_m \left(U_{k_1}^{(1)} * V_m^{(1)} \right) \otimes \dots \otimes \left(U_{k_d}^{(d)} * V_m^{(d)} \right),$$

which contains $R\#\mathcal{S}(\mathcal{B}) = Rr^d$ terms. Computing the 1D convolution $U_{k_\ell}^{(\ell)} * V_m^{(\ell)} \in \mathbb{R}^{2n-1}$ in $O(n \log n)$ operations leads to the desired bound. In the second case, we obtain

$$\mathcal{F} * \mathcal{G} = \sum_{k=1}^{R_1} \sum_{m=1}^{R_2} b_k c_m \left(U_k^{(1)} * V_m^{(1)} \right) \otimes \dots \otimes \left(U_k^{(d)} * V_m^{(d)} \right), \tag{25}$$

leading to the reduced cost since the number of terms is equal to R_1R_2 . \square

The resultant convolution product $\mathcal{F} * \mathcal{G}$ can be again represented (approximately) in either low-rank Tucker or CP formats, depending on further MLA operations applied to this tensor. Fast tensor-product convolution clearly outperforms the traditional FFT based algorithm that scales as $O(n^d \log n)$ (see [5] for more details). The fast tensor-product convolution can be implemented on a nonuniform grid as well (cf. [15, 26, 27]), however, this topic is beyond the scope of our paper.

3.6 Low-rank approximation of \mathcal{F} and \mathcal{G}

In our applications, the function related collocation coefficients tensor $\mathcal{F} = [f_i]_{i \in \mathcal{I}}$ is generated by either density function $\rho(x)$ or by the product $V(x)\psi(x)$. In this way we make *a priori* assumption on the existence of low rank approximation to both tensors. This assumption is not easy to analyse, however, it works well in practice.

Example 3 In the case of hydrogen atom we have

$$\rho(x) = e^{-2|x|}, \quad \text{and} \quad V(x)\psi(x) = \frac{e^{-|x|}}{|x|},$$

hence, the corresponding low-rank tensor approximations can be proven along the line of Lemma 4.3 [24] and Theorem 3 below.

To construct low-rank approximation of tensor \mathcal{G} , we consider a class of multi-variate spherically-symmetric convolving kernels $g : \mathbb{R}^d \rightarrow \mathbb{R}$ parametrised by

$$g(y) = G(\rho(y)) \equiv G(\rho) \quad \text{with} \quad \rho \equiv \rho(y) = y_1^2 + \dots + y_d^2,$$

where the univariate function $G : \mathbb{R}_+ \rightarrow \mathbb{R}$ can be represented via the generalised Laplace transform

$$G(\rho) = \int_{\mathbb{R}_+} \widehat{G}(\tau^2) e^{-\rho\tau^2} d\tau. \tag{26}$$

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

$$\phi_i(\cdot) = \phi(\cdot + (i - 1)h), \quad i \in I_n,$$

for all spatial dimensions $\ell = 1, \dots, d$, where $h > 0$ is the mesh parameter, so that the corresponding tensor-product basis function ϕ_i is defined by (18).

Using *sinc*-quadrature methods [33], we approximate the collocation coefficients tensor $\mathcal{G} = [g_i]_{i \in \mathcal{I}}$ in (20) via rank- $(2M + 1)$ canonical decomposition

$$\mathcal{G} \approx \sum_{k=-M}^M w_k \mathcal{E}(\tau_k) \quad \text{with} \quad \mathcal{E} = [e_i], \quad \mathbf{i} \in \mathcal{I},$$

with suitably chosen coefficients $w_k \in \mathbb{R}$ and $\tau_k \in \mathbb{R}_+$, and with the rank-1 components $\mathcal{E}(\tau_k) \in \mathbb{R}^{\mathcal{I}}$ given by

$$e_i(\tau_k) = \widehat{G}(\tau_k^2) \prod_{\ell=1}^d \int_{\mathbb{R}} e^{-y_\ell^2 \tau_k^2} \phi_{i_\ell}(y_\ell) dy_\ell. \tag{27}$$

Following the standard description of the *sinc*-methods, we introduce the Hardy space $H^1(D_\delta)$ as the set of all complex-valued functions f , which are analytic in the strip

$$D_\delta := \{z \in \mathbb{C} : |\Im z| < \delta\},$$

such that

$$N(f, D_\delta) := \int_{\mathbb{R}} (|f(x + i\delta)| + |f(x - i\delta)|) dx < \infty.$$

Given $f \in H^1(D_\delta)$, $h > 0$, and $M \in \mathbb{N}_0$, the corresponding *sinc*-quadrature reads as

$$T_M(f, h) := h \sum_{k=-M}^M f(kh) \approx \int_{\mathbb{R}} f(\xi) d\xi. \tag{28}$$

Proposition 1 *Let $f \in H^1(D_\delta)$, $\mathfrak{h} > 0$, and $M \in \mathbb{N}_0$ be given. If f possesses the hyper-exponential decay*

$$|f(\xi)| \leq C \exp(-be^{a|\xi|}) \text{ for all } \xi \in \mathbb{R} \tag{29}$$

with $a, b, C > 0$, then the choice $\mathfrak{h} = \log(\frac{2\pi aM}{b})/(aM)$ leads to (cf. [11])

$$\left| \int_{\mathbb{R}} f(\xi) d\xi - T_M(f, \mathfrak{h}) \right| \leq C N(f, D_\delta) e^{-\frac{2\pi \delta aM}{\log(2\pi aM/b)}}.$$

For a class of analytic functions the exponential convergence of above quadrature in M can be proven (see [16]). In our particular case of the Yukawa potential for $\kappa \in [0, \infty)$, we apply the Gauss transform (cf. (26))

$$G(\rho) = \frac{e^{-\kappa\sqrt{\rho}}}{\sqrt{\rho}} = \frac{2}{\sqrt{\pi}} \int_{\mathbb{R}_+} \exp(-\rho\tau^2 - \kappa^2/\tau^2) d\tau, \tag{30}$$

corresponding to the choice

$$\widehat{G}(\tau^2) = \frac{2}{\sqrt{\pi}} e^{-\kappa^2/\tau^2}.$$

Theorem 3 *For given $G(\rho)$ in (30) with fixed $\kappa > 0$, we set*

$$w_k = \mathfrak{h}_M \widehat{G}(\tau_k^2) \text{ and } \tau_k = e^{t_k},$$

where $t_k = k\mathfrak{h}_M$ with $\mathfrak{h}_M = C_0 \log(M)/M$ for some $C_0 > 0$. Then we have the exponentially convergent CP approximation of the rank $(2M + 1)$ for the collocation coefficients tensor $\mathcal{G} = [g_i]_{i \in \mathcal{I}}$ in (20) corresponding to the Yukawa potential $G(\rho)$,

$$\left\| \mathcal{G} - \sum_{k=-M}^M w_k \mathcal{E}(\tau_k) \right\| \leq C e^{-\pi^2 M / (C + \log(M))}, \tag{31}$$

where $\mathcal{E} = [e_i]$ with (27).

Proof We apply Proposition 1. Following [33], we choose the analyticity domain for the integrand in (26) as a sector $S_\delta := \{w \in \mathbb{C} : |\arg(w)| < \delta\}$ with apex angle $0 < 2\delta < \pi/2$, and then make use the conformal map

$$\varphi^{-1} : S_\delta \rightarrow D_\delta \text{ with } w = \varphi(z) = e^z, \varphi^{-1}(w) = \log(w).$$

Hence, we apply the change of variables $\tau = e^t$ to obtain

$$G(\rho) = \int_{\mathbb{R}} f(t; \rho) dt \text{ with } f(t; \rho) = \mathcal{Q}(t) e^{-\rho e^{2t}},$$

where $\mathcal{Q}(t) = \mathcal{Q}(t; \kappa) = \frac{2}{\sqrt{\pi}} e^{t-\kappa^2 e^{-2t}}$. By the way, f can be analytically extended into the strip D_δ . By definition (20) we have

$$g_i = \langle G(\rho), \phi_i \rangle = \int_{\mathbb{R}} \langle f(t; \rho), \phi_i \rangle dt \equiv \int_{\mathbb{R}} p_i(t) dt,$$

where

$$p_i(t) = \mathcal{Q}(t) \prod_{\ell=1}^d \int_{\mathbb{R}} e^{-y_\ell^2 e^{2t}} \phi_{i_\ell}(y_\ell) dy_\ell.$$

Furthermore, the function $p_i : \mathbb{R} \rightarrow \mathbb{R}$, can be analytically extended into the strip D_δ with $0 < \delta < \pi/4$, and this extension belongs to the Hardy space $H^1(D_\delta)$. In fact, introducing the error function $\text{erf} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$\text{erf}(t) := \frac{2}{\sqrt{\pi}} \int_0^t e^{-\tau^2} d\tau,$$

we calculate the explicit representation

$$\int_{\mathbb{R}} e^{-y^2 e^{2t}} \phi_i(y) dy = \frac{1}{2t} \{ \text{erf}(t i h) - \text{erf}(t(i-1)h) \}, \tag{32}$$

with $h = 2A/n$ (uniform grid spacing) for $i = 1, \dots, n$. Since $\text{erf}(z)/z$ is an entire function it proves the required analyticity of p_i .

Now we estimate the constant $N(f, D_\delta)$ applying arguments similar to those in Lemma 4.7, [24]. To that end, we let $H_i = h(i_1 - 1, \dots, i_d - 1) \in \mathbb{R}^d$ to obtain

$$\int_{\mathbb{R}^d} e^{-w^2 |y|^2} \phi(y + H_i) dy = \int_{\mathbb{R}^d} e^{-w^2 |v - H_i|^2} \phi(v) dv,$$

taking into account that ϕ has compact support $[-h, h]^d$.

We notice that $|\mathcal{Q}(\zeta \exp(i\delta))| \leq C_0 < \infty$ for $\zeta \in [0, \infty)$. The following argument is slight modification of those in Lemma 4.3, [24], applied there to the Galerkin approximation of the function $e^{-|x|}$,

$$\begin{aligned} N(f, D_\delta) &= \int_{\partial S_\delta} |f(w)| |dw| \\ &= \int_{\partial S_\delta} \int_{\mathbb{R}^d} |\mathcal{Q}(w)| \left| e^{-w^2 |y|^2} \phi(y + H_i) dy \right| |dw| \\ &\leq 2 \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} |\mathcal{Q}(\zeta e^{i\delta})| \left| e^{-\zeta^2 \exp(2i\delta) |u - H_i|^2} \phi(u) du \right| d\zeta \\ &\leq 2C_0 \int_{\mathbb{R}^d} \int_{\mathbb{R}_+} \left| e^{-\zeta^2 \exp(2i\delta) |u + H_i|^2} \right| d\zeta |\phi(u)| du \\ &= 2C_0 \int_{\mathbb{R}^d} \int_{\mathbb{R}_+} e^{-\zeta^2 \cos(2\delta) |u + H_i|^2} d\zeta |\phi(u)| du \\ &= \frac{2C_0}{\sqrt{\cos(2\delta)}} \int_{\mathbb{R}^d} \frac{|\phi(u)|}{|u - H_i|^2} du. \end{aligned}$$

The latter bound is uniform in H_i .

The decay of the integrand $p(t) := p(t; \kappa)$ on the real axis is

$$p(t; \kappa) \approx e^{t-\kappa^2 e^{2t}} \text{ as } t \rightarrow \infty,$$

$$p(t; \kappa) \approx e^{t-\kappa^2 e^{2|t|}} \text{ as } t \rightarrow -\infty,$$

corresponding to $a = 2, b = \kappa^2$ and $C = 1$ in (29).

Finally, we apply Proposition 1, providing the exponential convergence of the quadrature approximation

$$g_i = \int_{\mathbb{R}} p(t; \kappa) dt \approx h_M \sum_{k=-M}^M p(t_k; \kappa),$$

which proves (31). □

Theorem 3 applies to the case with fixed $\kappa > 0$. To construct the approximation, which is uniform for $\kappa \in [0, \infty)$ (i.e., including the case of Coulomb potential corresponding to $\kappa = 0$) we modify the above quadrature using a variable transformation $t = \sinh(u)$ to obtain the quadrature

$$\int_{\mathbb{R}} p_i(t) dt = \int_{\mathbb{R}_+} 2 \cosh(u) p_i(\sinh(u)) du \approx \sum_{k=0}^M w_k p_i(t_k)$$

with $t_k := \sinh(kh_M)$ and

$$w_k := \begin{cases} h_M & \text{for } k = 0 \\ 2 h_M \cosh(kh_M) & \text{for } k > 0 \end{cases} \quad (33)$$

with the choice $h_M = C_0 \frac{\ln(M)}{M}$ for some $C_0 > 0$. The analysis is similar to those in Lemma 4.3, [16], taking into account the symmetry of the integrand.

4 Discretisation, complexity issues, numerics

4.1 Discretisation of the Green function operator

Our primal interest is concerned with integral formulation in L^2 -setting. Hence, we consider the collocation-type approximation of the operator $\mathcal{G}_z = (V \cdot) \star G_z$ with respect to the certain ansatz space $W = \text{span}\{\phi_i\} \subset L^2(\mathbb{R}^3)$ of discontinuous functions (see, e.g., (18)). Letting

$$\psi = \sum_{i \in \mathcal{I}} a_i \phi_i, \quad \mathbb{V} = \{(V \phi_i, \phi_j)\}_{i, j \in \mathcal{I}},$$

we calculate the L^2 -projection of $V\psi$ onto W ,

$$\mathcal{P}_W(V\psi) = \sum_{j \in \mathcal{I}} \sum_{i \in \mathcal{I}} a_i (V \phi_i, \phi_j) \phi_j = \sum_{j \in \mathcal{I}} b_j \phi_j$$

with $b_j = (\mathbb{V}\mathcal{A})_j$, $\mathcal{A} = [a_i]$, and introduce the coefficients tensor

$$\mathcal{F} = [b_j]_{j \in \mathcal{I}} \in \mathbb{R}^{n \times n \times n}, \quad \text{i.e., } \mathcal{F} = \mathbb{V}\mathcal{A}.$$

The discretisation of $\mathcal{G}_z \psi$ is then defined by

$$\mathcal{G}_z \psi \approx \mathcal{G} \star \mathcal{F},$$

where tensor \mathcal{G} , corresponding to the convolving kernel $e^{-z|x|}/|x|$, was described in Sect. 3.6. In the case of piecewise constant basis functions (18) the stiffness matrix \mathbb{V} becomes diagonal, $\mathbb{V} = \text{diag}\{(V \phi_i, \phi_i)\}_{i \in \mathcal{I}}$.

Finite element or finite difference approximations result into the representation $\mathcal{G}_z \psi = R_z V \psi$, where R_z is the discrete elliptic resolvent and $V \psi$ is the action of the discrete interaction potential. In this case, we first approximate the initial operator H using the ansatz space of continuous functions in $H^1(\mathbb{R}^3)$, and then compute the discrete elliptic inverse $(-\Delta + z^2 I)^{-1}$.

For example, the finite difference Helmholtz operator on the uniform $n \times n \times n$ tensor-product grid (subject to homogeneous Dirichlet boundary conditions) is represented by a matrix $\mathbb{A} + z^2 h^2 \mathbb{I} \in \mathbb{R}^{n \times n \times n}$ with

$$\mathbb{A} := V^{(1)} \otimes I \otimes I + I \otimes V^{(2)} \otimes I + I \otimes I \otimes V^{(3)},$$

and with $V^{(\ell)}, I \in \mathbb{R}^{n \times n}$, where I is the identity matrix and $V^{(\ell)} = \text{tridiag}\{-1, 2, -1\}$, $\ell = 1, 2, 3$. Using sinc-methods, we choose coefficients $t_k, c_k \in \mathbb{R}$ and then construct the rank- $(2M + 1)$ CP approximation in the form

$$B_{(M)} = \sum_{k=-M}^M c_k \bigotimes_{\ell=1}^3 \exp(-t_k (V^{(\ell)} + \frac{z^2 h^2}{d} I)) \approx (\mathbb{A} + z^2 h^2 \mathbb{I})^{-1},$$

providing exponential convergence in $r = 2M + 1$,

$$\|(\mathbb{A} + z^2 h^2 \mathbb{I})^{-1} - B_{(M)}\| \leq C e^{-sM/\log(M)}.$$

The particular choice of coefficients t_k, c_k is described in [16].

4.2 Complexity issues

We assume that at each step m of the Newton iteration, we have rank- R_G and rank- R_F canonical representations of tensors $\mathcal{G}_n \in \mathbb{R}^{n^3}$ and $\mathcal{F}_n \in \mathbb{R}^{n^3}$ at our disposal, respectively, as well as the good initial guess for the orthogonal components of the Tucker approximation for the next iterant $\psi_{m+1} = \mathcal{G}_m \star \mathcal{F}_m$. We apply the two-level approximation method [27]. Since the tensor ψ_{m+1} has maximal canonical rank $R_G R_F$, its rank- r Tucker approximation has the cost

$$Q_{CT} = O(R_G R_F n r^2 + n r^2 \min\{r^2, n\}).$$

Notice that in the d -dimensional setting, one can apply the recent *multigrid tensor approximation method* [28], which scales linearly in all essential parameters (getting rid of the curse of dimensionality),

$$Q_{CT} = O(d R_G R_F n r + d n r^2).$$

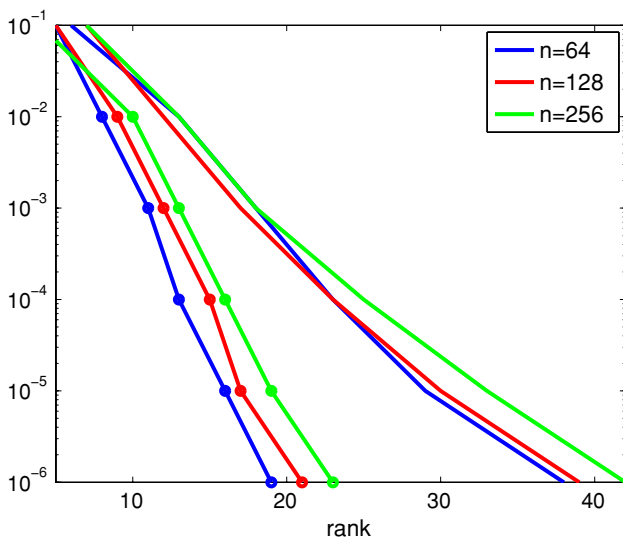


Fig. 2 CP approximation of the Coulomb potentials via *sinc* quadratures (solid lines). Algebraically recompressed approximations (marked solid lines)

Computing the rank- R_F CP approximation to the corresponding core tensor of size $r \times r \times r$ (by ALS iteration) with the cost that does not depend on n ,

$$Q_{FC} = O(N_{it} R_F r^2),$$

where N_{it} is the number of ALS iteration loops, we return the current iterant ψ_{m+1} to the required format $\psi_{m+1} \in \mathcal{C}_{R_F}$. The overall numerical cost is estimated by $Q_{TC} + Q_{FC}$, which scales linearly in d and in the univariate problem size n .

Similar complexity bounds can be derived for the FE/FD discretizations.

4.3 Numerics

Computations were performed in MATLAB, Release 7.3. General purpose MATLAB subroutines to compute tensor approximations by ALS iteration can be found in [1,32]. Some special algorithms for fast numerical multi-linear algebra are described in [24,26,27].

The collocation coefficients tensor $\mathcal{G} \in \mathbb{R}^{\mathcal{I}}$, corresponding to the Coulomb potential $g(x) = 1/|x|$, is approximated with accuracy 10^{-6} in the rank- R CP format with $R \in [10, 20]$ depending logarithmically on n (see [18] for more details). Convergence results corresponding to the *sinc*-quadrature method are shown in Fig. 2 (solid lines). Marked solid lines correspond to algebraically recompressed *sinc* approximations described in [26].

Figure 3 illustrates the *sinc*-approximation method for the discrete elliptic inverse defined on $n \times n \times n$ grid with $n = 128$. Here M defines the number of terms in the *sinc*-quadrature, so that the total rank is $r = 2M + 1$.

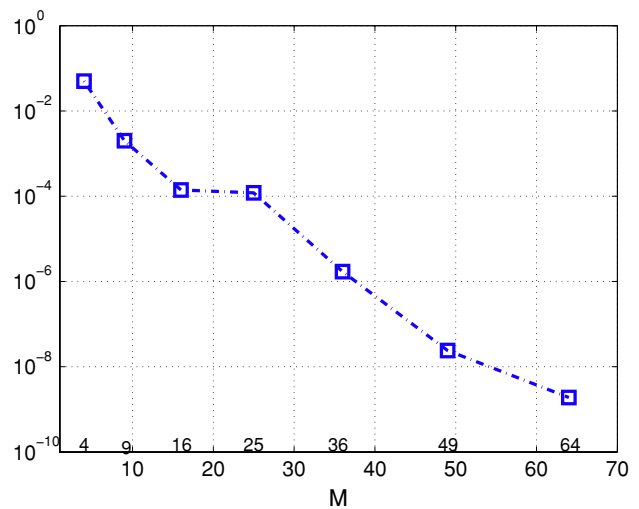


Fig. 3 CP approximation for the discrete elliptic inverse by *sinc*-quadratures

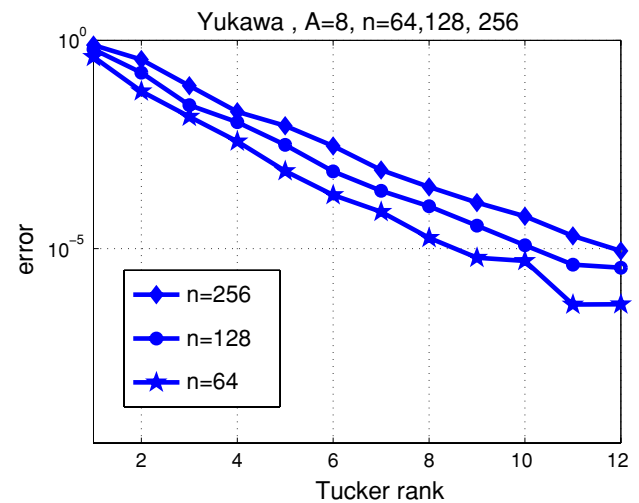


Fig. 4 Convergence of the Tucker approximation for the Yukawa potential

Figure 4 presents the convergence history for the Tucker approximation of the Yukawa potential defined on $[0, 8]^3$ with different number of the grid points $n = 64, 128, 256$. We visualise the relative error in the Frobenius norm.

Next example demonstrates the convergence of the two-level approximation scheme, where the core tensor \mathcal{B} of the Tucker representation is further approximated by the CP model. Figure 5 presents the convergence behaviour of the CP ALS iteration applied to the $12 \times 12 \times 12$ core tensor \mathcal{B} from the previous approximation with $n = 64$.

The last Fig. 6 presents the pointwise absolute error of the rank-12 Tucker approximation for the Yukawa potential corresponding to Fig. 4 with $n = 64$. Note that the maximal value of the approximated function $e^{-|x|}/|x|$ at the respective collocation points equals to ≈ 8.3 .

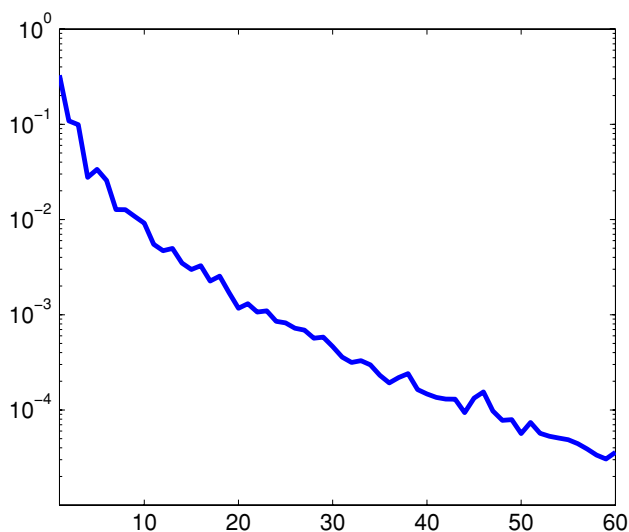


Fig. 5 CP approximation for the Yukawa potential by ALS iteration

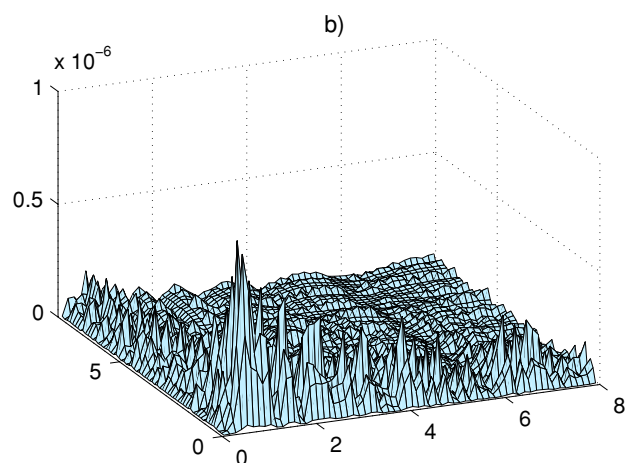


Fig. 6 Pointwise absolute error of the rank-12 Tucker approximation for the Yukawa potential

Numerical experiments on the rank-structured solution of the eigenvalue problems will be presented in the forthcoming papers.

Acknowledgments The author is appreciative to Prof. M. Mohlenkamp (Ohio University, OH) for providing the stimulating paper [31]. The author acknowledges Dipl.-Phys. V. Khoromskaia (MPI MIS, Leipzig) for assistance with numerical computations.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Bader, B.W., Kolda, T.G.: MATLAB Tensor Toolbox Version 2.2. SANDIA Report, Sandia National Laboratories (2007)
- Beylkin, G., Mohlenkamp, M.J.: Numerical operator calculus in higher dimensions. *Proc. Natl. Acad. Sci. USA* **99**, 10246–10251 (2002)
- Beylkin, G., Mohlenkamp, M.J., Pérez, F.: Approximating a wavefunction as an unconstrained sum of Slater determinants. University of Colorado at Boulder, 2007 (to appear)
- Chinnamsetty, S.R., Espig, M., Hackbusch, W., Khoromskij, B.N., Flad, H.-J.: Optimal Kronecker tensor-product approximation in quantum chemistry. *J. Chem. Phys.* **127**, 084110 (2007)
- Chinnamsetty, S.R., Flad, H.-J., Khoromskaia, V., Khoromskij, B.N.: Tensor decomposition in electronic structure calculations on Cartesian grids. *MPI MIS* 65, Leipzig 2007 (submitted)
- De Lathauwer, L., De Moor, B., Vandewalle, J.: On the best rank-1 and rank- (R_1, \dots, R_N) approximation of higher-order tensors. *SIAM J. Matrix Anal. Appl.* **21**, 1324–1342 (2000)
- De Lathauwer, L., De Moor, B., Vandewalle, J.: Computation of the canonical decomposition by means of a simultaneous generalized Schur decomposition. *SIAM J. Matrix Anal. Appl.* **26**, 295–327 (2004)
- Fedorov, M.V., Flad, H.-J., Chuev, G.N., Grasedyck, L., Khoromskij, B.N.: A structured low-rank wavelet solver for the Ornstein-Zernike integral equation. *Computing* **80**(1), 47–73 (2007)
- Flad, H.-J., Hackbusch, W., Khoromskij, B.N., Schneider, R.: Concept of data-sparse tensor-product approximation in many-particle modeling. *MPI MIS* 3, Leipzig 2008 (submitted)
- Flad, H.-J., Hackbusch, W., Schneider, R.: Best N-term approximation in electronic structure calculation: I. One-electron reduced density matrix. *ESAIM:M2AN* **40**, 49–61 (2006)
- Gavrilyuk, I.P., Hackbusch, W., Khoromskij, B.N.: Data-sparse approximation to a class of operator-valued functions. *Math. Comp.* **74**, 681–708 (2005)
- Gavrilyuk, I.P., Hackbusch, W., Khoromskij, B.N.: Tensor-product approximation to elliptic and parabolic solution operators in higher dimensions. *Computing* **74**, 131–157 (2005)
- Griebel, M., Hamaekers, J.: Sparse grids for the Schrödinger equation. *ESAIM: M2AN* **41**, 215–247 (2007)
- Hackbusch, W.: Fast and exact projected convolution for non-equidistant grids. Preprint 102 MPI MIS, Leipzig (2006)
- Hackbusch, W.: Fast and exact projected multi-dimensional convolution for functions defined in locally refined grids. *MPI MIS, Leipzig*, in progress (2007)
- Hackbusch, W., Khoromskij, B.N.: Low-rank Kronecker product approximation to multi-dimensional nonlocal operators. Part I. Separable approximation of multi-variate functions. *Computing* **76**, 177–202 (2006)
- Hackbusch, W., Khoromskij, B.N.: Low-rank Kronecker product approximation to multi-dimensional nonlocal operators. Part II. HKT representations of certain operators. *Computing* **76**, 203–225 (2006)
- Hackbusch, W., Khoromskij, B.N.: Tensor-product approximation to operators and functions in high dimension. *J. Complex.* **23**, 697–714 (2007). doi:10.1016/j.jco.2007.03.007
- Hackbusch, W., Khoromskij, B.N.: Tensor-product approximation to multi-dimensional integral operators and Green's functions. Preprint 38, MPI MIS, Leipzig 2006, *SIAM J. Matr. Anal. Appl.* (to appear) (2007)
- Hackbusch, W., Khoromskij, B.N., Tyrtshnikov, E.: Hierarchical Kronecker tensor-product approximation. *J. Numer. Math.* **13**, 119–156 (2005)
- Hackbusch, W., Khoromskij, B.N., Tyrtshnikov, E.: Approximate iteration for structured matrices. *Numer. Math.* (2008). doi:10.1007/s00211-008-0143-0
- Harrison, R.J., Fann, G.I., Yanai, T., Gan, Z., Beylkin, G.: Multiresolution quantum chemistry: basic theory and initial applications. *J. Chem. Phys.* **121**(23), 11587–11598 (2004)

23. Kalos, M.H.: Monte Carlo calculations of the ground state of three- and four-body nuclei. *Phys. Rev.* **2**(128), 1791–1795 (1962)
24. Khoromskij, B.N.: Structured rank- (r_1, \dots, r_d) decomposition of function-related tensors in \mathbb{R}^d . *Comp. Meth. Appl. Math.* **6**(2), 194–220 (2006)
25. Khoromskij, B.N.: Structured data-sparse approximation to high-order tensors arising from the deterministic Boltzmann equation. *Math. Comp.* **76**(259), 1291–1315 (2007)
26. Khoromskij, B.N.: Tensor approximation of multivariate convolution with linear scaling in the dimension. MPI MIS, Leipzig, 2008 (to be submitted)
27. Khoromskij, B.N., Khoromskaia, V.: Low rank Tucker tensor approximation to the classical potentials. *Cent. Eur. J. Math.* **5**(3), 1–28 (2007)
28. Khoromskij, B.N., Khoromskaia, V.: Two-level tensor approximation with multigrid acceleration. MPI MIS, Leipzig, 2008 (to be submitted)
29. Le Bris, C.: Computational chemistry from the perspectives of numerical analysis. *Acta Numer.* **14**, 363–444 (2005)
30. Le Bris, C., Lions, P.-L.: From atoms to crystals: A mathematical journey. *Bull. Am. Math. Soc.* **42**(3), 291–363 (2005)
31. Mohlenkamp, M.J., Young, T.: Convergence of Green iterations for Schrödinger equations (to appear) (2007)
32. Smilde, A., Bro, R., Geladi, P.: *Multi-way Analysis*. Wiley, New York (2004)
33. Stenger, F.: *Numerical Methods Based on Sinc and Analytic Functions*. Springer, Berlin (1993)
34. Temlyakov, V.N.: Greedy algorithms and M -Term approximation with regard to redundant dictionaries. *J. Approx. Theory* **98**, 117–145 (1999)
35. Tucker, L.R.: Some mathematical notes on three-mode factor analysis. *Psychometrika* **31**, 279–311 (1966)
36. Tyrtshnikov, E.E.: Kronecker-product approximations for some function-related matrices. *Linear Algebra Appl.* **379**, 423–437 (2004)
37. Yserentant, H.: On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numer. Math.* **98**, 731–759 (2004)