



Cite this: *Phys. Chem. Chem. Phys.*,  
2015, 17, 31491

# Tensor numerical methods in quantum chemistry: from Hartree–Fock to excitation energies

Venera Khoromskaia<sup>ab</sup> and Boris N. Khoromskij<sup>b</sup>

We resume the recent successes of the grid-based tensor numerical methods and discuss their prospects in real-space electronic structure calculations. These methods, based on the low-rank representation of the multidimensional functions and integral operators, first appeared as an accurate tensor calculus for the 3D Hartree potential using 1D complexity operations, and have evolved to entirely grid-based tensor-structured 3D Hartree–Fock eigenvalue solver. It benefits from tensor calculation of the core Hamiltonian and two-electron integrals (TEI) in  $O(n \log n)$  complexity using the rank-structured approximation of basis functions, electron densities and convolution integral operators all represented on 3D  $n \times n \times n$  Cartesian grids. The algorithm for calculating TEI tensor in a form of the Cholesky decomposition is based on multiple factorizations using algebraic 1D “density fitting” scheme, which yield an almost irreducible number of product basis functions involved in the 3D convolution integrals, depending on a threshold  $\varepsilon > 0$ . The basis functions are not restricted to separable Gaussians, since the analytical integration is substituted by high-precision tensor-structured numerical quadratures. The tensor approaches to post-Hartree–Fock calculations for the MP2 energy correction and for the Bethe–Salpeter excitation energies, based on using low-rank factorizations and the reduced basis method, were recently introduced. Another direction is towards the tensor-based Hartree–Fock numerical scheme for finite lattices, where one of the numerical challenges is the summation of electrostatic potentials of a large number of nuclei. The 3D grid-based tensor method for calculation of a potential sum on a  $L \times L \times L$  lattice manifests the linear in  $L$  computational work,  $O(L)$ , instead of the usual  $O(L^3 \log L)$  scaling by the Ewald-type approaches.

Received 28th February 2015,  
Accepted 20th May 2015

DOI: 10.1039/c5cp01215e

www.rsc.org/pccp

## 1 Introduction

The problems of numerical modeling the many-particle interactions in large molecular systems, lattice structured metallic clusters and crystals, proteins and nanomaterials are the most challenging tasks in modern computational physics and chemistry. The traditional approaches for these multidimensional problems are restricted to the concepts which have well recognized limitations for computations with higher accuracy and for larger non-periodic molecular systems, as well as for efficient calculation of excited states. Recent advances in numerical analysis for multidimensional problems and significant achievements in high performance computing suggest new creative approaches to these problems.

The Hartree–Fock (HF) equation governed by the 3D integral-differential operator<sup>29,67</sup> is one of the basic models for *ab initio* calculation of the ground state energy of molecular systems.

It is a strongly nonlinear eigenvalue problem (EVP) in a sense, that one should solve the equation in a self-consistent way when the integral part of the governing operator depends on the solution itself. Multiple strong singularities in the electron density of a molecule due to nuclear cusps impose strong requirements on the accuracy of calculations.

Commonly used numerical methods for solution of the Hartree–Fock equation are based on the analytical computation of the arising two-electron integrals (convolution type integrals in  $\mathbb{R}^3$ ) in the problem adapted naturally separable Gaussian-type bases,<sup>1,17</sup> by using erf-function expansions. This rigorous approach resulted in a number of efficient implementations which are widely used in computational quantum chemistry. The success of the analytical integration methods stems from the big amount of precomputed information based on the physical insight including the construction of problem adapted atomic orbitals basis sets and elaborate nonlinear optimization for calculation of density fitting basis. The known limitations of this approach appear due to a strong dependence of the numerical efficiency on the size and quality of the chosen Gaussian basis sets, that might be crucial for larger molecular clusters and heavier atoms.

<sup>a</sup> Max-Planck-Institute for Dynamics of Complex Systems, Magdeburg, Germany.  
E-mail: vekh@mis.mpg.de

<sup>b</sup> Max-Planck Institute for Mathematics in the Sciences, Inselstr. 22-26, Leipzig, Germany



The intention to replace or assist the analytical calculations for the Hartree–Fock problem by a data-sparse grid-based numerical schemes has a long history. First success was the grid-based numerical method for the diatomic molecules in ref. 66, though this approach was not feasible to compact (3D) molecules. The wavelet multiresolution schemes<sup>13</sup> capable for the accurate representation of nuclear cusps, have been applied to electronic structure simulations,<sup>21,28,63</sup> and recently this approach was further advanced due to achievements in high performance supercomputing.<sup>8,18,20,63</sup> However, due to extensive computational resources, the entirely wavelet-based or sparse-grid approaches<sup>24,74</sup> are limited so far only to rather small atomic systems with few electrons.<sup>2</sup> Tensor hyper-contraction decompositions in density fitting schemes have been analyzed in ref. 32 and 59.

The newly developed tensor-structured numerical methods, both the name and the concept, appeared during the work on the grid-based tensor approach to the solution of the 3D Hartree–Fock problem.<sup>38,39,42</sup> The central point is the representation of  $d$ -variate functions and integral/differential operators on large  $n^{\otimes d}$  grids and their approximation in the low-rank tensor formats, which allows numerical calculations in  $O(dn)$  complexity instead of  $O(n^d)$  by conventional methods.

In this paper we summarize the main benefits of the tensor numerical methods in electronic structure calculations, and discuss further prospects of this approach in several directions, such as

- (a) Algebraic directional density fitting and tensor factorization of the two-electron integrals in Hartree–Fock calculations;
- (b) Tensor decompositions in the MP2 energy correction, and in calculation of excitation energies based on the Bethe–Salpeter equation;
- (c) Fast tensor summation of electrostatic potentials on large 3D lattices for efficient calculation of the Fock matrix in the case of lattice-structured systems;
- (d) Fast method for calculation of the interaction energy of a large number of Coulombic potentials on a 3D lattice.

Notice that basic rank-structured tensor formats such as the canonical (PARAFAC/CANDECOMP) and Tucker tensor decompositions have been since long used in the computer science for the quantitative analysis of correlations in the experimental multidimensional data arrays in data processing and chemometrics, see ref. 12, 50 and 64 and references therein. In 2006 the exceptional properties of the Tucker decomposition for the discretized multidimensional functions have been revealed in ref. 36 and 38, where it was proven that for a class of function-related tensors the approximation error of the Tucker decomposition decays exponentially with respect to the Tucker rank. This gives the opportunity to represent the discretized multidimensional functions and integral (convolution) operators in an algebraically separable form and thus reduce the numerical treatment of the multidimensional transforms to 1D operations. It was shown that the number of vectors in such a representation depends only logarithmically on the size of the  $d$ -dimensional grid,  $O(n^d)$ , used for discretization of multivariate functions.

The above results have led to the idea to calculate the Hartree potential and the Coulomb and exchange operators

by numerical quadratures<sup>39</sup> using Gaussian-type basis functions discretized on 3D Cartesian grids and the fast tensor-product convolution.<sup>35</sup> In this way, the efficient low-rank canonical tensor representation to the Newton kernel was an essential contribution.<sup>7</sup> To reduce the initial rank of the electron density, that is quadratically proportional to the number of GTO basis functions, the canonical-to-Tucker tensor transform was invented,<sup>38,39</sup> which made computations for large tensor grids and extended molecules tractable (even in Matlab).

The initial version of tensor-structured algorithms for solving Hartree–Fock equation employed the 3D grid-based calculation of the Coulomb and exchange integral operators “on-the-fly”, thus avoiding precomputation and storage of the TEI tensor.<sup>40,42</sup> In particular, it was shown that tensor calculus allows to reduce the 3D convolution integrals to combinations of 1D convolutions, and 1D Hadamard and scalar products. Besides, these results promoted spreading of the tensor-structured methods in the community of numerical analysis,<sup>22,27,37</sup> and further development of the tree-tensor formats like tensor-train<sup>57</sup> and hierarchical Tucker representations.<sup>26</sup>

Further development of tensor methods in electronic structure calculations was due to the fast algorithm for the grid-based computation of the TEI tensor<sup>44</sup> in  $O(N_b^3)$  storage in the number of basis functions  $N_b$ . The fourth order TEI tensor is calculated in a form of the Cholesky factorization by using the algebraic 1D “density fitting” scheme, which applies to the product basis functions and the Newton convolving kernel all discretized on  $n \times n \times n$  Cartesian grid, the 3D integral transforms are calculated in  $O(n \log n)$  complexity. Given the factorized TEI, the update of the Column and exchange parts in the Fock matrix reduces to the cheap algebraic operations. Other steps are tensor calculation of the core Hamiltonian and the efficient MP2 energy correction scheme,<sup>45</sup> which all together gave rise to the black-box Hartree–Fock solver.<sup>46</sup>

Due to the grid representation of basis functions basis sets are now not restricted to Gaussian-type orbitals and allowed to be any well-separable function defined on a grid. The tensor-based Hartree–Fock solver is competitive in computational time and accuracy with the standard packages based on analytical calculation of integrals. High accuracy is attained owing to easy calculations on large 3D grids up to  $n^3 = 10^{18}$ , so that the resolution with mesh size  $h$  of the order of atomic radii,  $10^{-4}$  Å, is possible.

Motivated by tensor decompositions in the MP2 scheme, the new approach for calculation of the excited states in the framework of the Bethe–Salpeter equation was recently introduced,<sup>6</sup> that employs the reduced basis method in combination with low-rank tensor approximations.

Further developments of the tensor methods in multi-particle simulations are focused on the large lattice structures in a box and nearly periodic systems, for which the tensor approach may reduce computational costs dramatically.<sup>48</sup> One of the challenging problems in the numerical treatment of the crystalline-type molecular clusters is summation of a large number of electrostatic potentials distributed on a non-periodic (finite) lattice.



The novel grid-based method for summation of the long-range potentials in the canonical and Tucker formats<sup>47,49</sup> works on  $L \times L \times L$  lattices with the computational cost  $O(L)$  instead of  $O(L^3 \log L)$  by the traditional Ewald-type methods.<sup>19</sup> The required precision is guaranteed by employing large 3D Cartesian grids for representation of potentials. The method remains efficient for lattices with non-rectangular geometries and in the presence of multiple defects.<sup>48,49</sup>

In particular, we present fast tensor method for the summation of the long-range interaction potentials on a 3D lattice by using the assembled vectors of their canonical and Tucker tensor representations. The important application of this approach to calculation of interaction energy of the Coulombic potentials on a lattice with sub-linear cost in the lattice size is described in detail.

The rest of the paper is organized as follows. In Section 2, we overview the basic tensor formats and show why the orthogonal Tucker tensor decomposition, originating from computer science and data processing, became useful for the treatment of the multidimensional functions and operators in numerical analysis. In particular, Sections 2.3 and 2.4 discuss the matrix product states (tensor train) formats and the novel quantics tensor approximation method, respectively, while Section 2.5 addresses the basic tensor-structured multilinear algebra operations. Section 3 describes tensor calculus of the multidimensional convolution transform on examples of the Hartree potential (Section 3.1) and the TEI tensor (Section 3.2), and recalls the main building blocks in the tensor-based black-box Hartree-Fock solver. Section 4 shows benefits of the tensor approach in MP2 calculations and in calculation of the lowest part of the excitation energies in the framework of the Bethe-Salpeter equation. Section 5 describes the benefits of tensor methods in applications to lattice-type molecular systems. Appendix discusses some computational details on the Canonical-to-Tucker tensor transform, the Galerkin discretization scheme for the nonlinear Hartree-Fock equation, and the basics of the low-rank canonical tensor representation for the Newton kernel.

## 2 Rank-structured tensor representations of discretized functions and operators

In this section we discuss shortly why the rank-structured tensors, which were traditionally used in experimental data processing, appeared to be useful for the separable representation of multivariate functions and operators represented on tensor product grids. There is a numerous literature on application of the canonical and Tucker tensor decompositions in chemometrics and data processing, see ref. 12, 31, 50, 64 and 68 and references therein. In the recent decade the class of matrix product states type tensor formats became popular in the simulations of quantum spin systems and quantum molecular dynamics. We refer to ref. 37 and 62 for recent literature surveys on commonly used tensor formats.

### 2.1 Canonical and Tucker tensor formats

We consider a tensor of order  $d$ , as a real multidimensional array  $\mathbf{A} = [a_{i_1, \dots, i_d}] \in \mathbb{R}^{n_1 \times \dots \times n_d}$  numbered by a  $d$ -tuple index set,<sup>†</sup> with multi-index notation  $\mathbf{i} = (i_1, \dots, i_d)$ ,  $i_\ell \in \{1, \dots, n_\ell\}$ ,  $\ell = 1, \dots, d$ . It is an element of the linear vector space  $\mathbb{R}^{n_1 \times \dots \times n_d}$  equipped with the Euclidean scalar product,

$$\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i_1=1}^{n_1} \dots \sum_{i_d=1}^{n_d} a_{i_1, \dots, i_d} b_{i_1, \dots, i_d}.$$

Euclidean vectors and matrices are the special case of  $d$ th order tensors. For a general tensor, the required storage scales exponentially in the dimension,  $n_1 n_2 \dots n_d$ , (the so-called “curse of dimensionality”). To get rid of exponential scaling in the dimension, one can apply the rank-structured separable representations of multidimensional tensors. The simplest separable element is given by rank-1 canonical tensor,

$$\mathbf{U} = \mathbf{u}^{(1)} \otimes \dots \otimes \mathbf{u}^{(d)} \in \mathbb{R}^{n_1 \times \dots \times n_d},$$

with entries

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

requiring only  $n_1 + \dots + n_d$  numbers to store it.

A tensor in the  $R$ -term canonical format is defined by

$$\mathbf{U} = \sum_{k=1}^R c_k \mathbf{u}_k^{(1)} \otimes \dots \otimes \mathbf{u}_k^{(d)}, \quad c_k \in \mathbb{R}, \quad (1)$$

where  $\mathbf{u}_k^{(\ell)} \in \mathbb{R}^{n_\ell}$  are normalized vectors, and  $R$  is called the canonical rank of a tensor. It is convenient to introduce the so-called side matrices  $\mathbf{U}^{(\ell)} = [\mathbf{u}_1^{(\ell)} \dots \mathbf{u}_R^{(\ell)}] \in \mathbb{R}^{n_\ell \times R}$ ,  $\ell = 1, 2, 3$ , obtained by concatenation of the canonical vectors  $\mathbf{u}_k^{(\ell)}$ ,  $k = 1, \dots, R$ . Now the storage cost is bounded by  $dRn$ . For  $d \geq 3$  computation of the canonical rank of a tensor  $\mathbf{U}$ , i.e. the minimal number  $R$  in representation (1) and the respective decomposition, is an N-P hard problem. In the case  $d = 2$  the representation (1) is merely a rank- $R$  matrix.

We say that a tensor  $\mathbf{A}$  is represented in the rank- $\mathbf{r}$  orthogonal Tucker format with the rank parameter  $\mathbf{r} = (r_1, \dots, r_d)$ , if

$$\mathbf{A} = \sum_{\nu_1=1}^{r_1} \dots \sum_{\nu_d=1}^{r_d} \beta_{\nu_1, \dots, \nu_d} \mathbf{v}_{\nu_1}^{(1)} \otimes \dots \otimes \mathbf{v}_{\nu_d}^{(d)}, \quad \ell = 1, \dots, d, \quad (2)$$

where  $\{\mathbf{v}_{\nu_\ell}^{(\ell)} \in \mathbb{R}^{n_\ell}\}$  represents the set of orthonormal vectors, and  $\beta = [\beta_{\nu_1, \dots, \nu_d}] \in \mathbb{R}^{r_1 \times \dots \times r_d}$  is the Tucker core tensor. The storage cost for the Tucker tensor is bounded by  $d r n + r^d$ ,  $r = \max r_\ell$ . Often the equivalent notation for the Tucker tensor format is used,

$$\mathbf{A} = \beta \times_1 V^{(1)} \times_2 V^{(2)} \times \dots \times_d V^{(d)}, \quad (3)$$

where  $\times_\ell$  denotes the contraction along the mode  $\ell$  and orthogonal matrices  $V^{(\ell)}$  incorporate the set of vectors  $\{\mathbf{v}_{\nu_\ell}^{(\ell)} \in \mathbb{R}^{n_\ell}\}$ .

<sup>†</sup> The alternative notation  $\mathbf{A} = [A(i_1, \dots, i_d)]$  can be utilized.



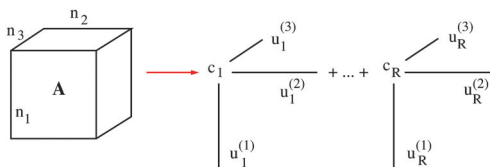


Fig. 1 The canonical tensor decomposition for  $d = 3$ .

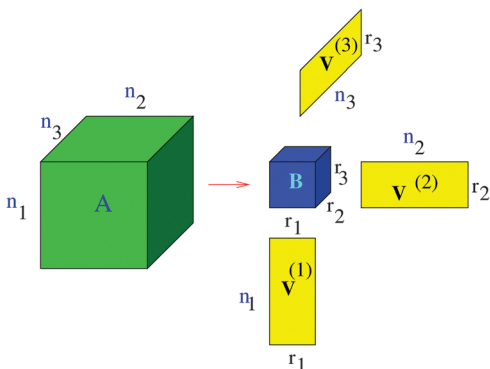


Fig. 2 Visualizing the Tucker tensor for  $d = 3$ .

In the case  $d = 2$ , the orthogonal Tucker decomposition is equivalent to the singular value decomposition (SVD) of a rectangular matrix. Fig. 1 and 2 visualizes the canonical and Tucker tensors in the case  $d = 3$ . Canonical tensor decomposition leads to an ill-posed problem and, up to our best knowledge, there are no robust algebraic algorithms for the canonical approximation of an arbitrary tensor. For some classes of analytic functions, the explicit low-rank approximation can be constructed in analytic form by using sinc-quadrature approximation to the Laplace transform. The robust methods for Tucker decomposition are based on the orthogonal projections using the higher order singular value decomposition<sup>16</sup> (HOSVD) that is the generalization of the singular value decomposition (SVD) of matrices.

The rank-structured decomposition (approximation) of multidimensional tensors provides means for the separation of variables in the discretized representation of multivariate functions and operators, and thus the possibility to substitute multidimensional algebraic transforms by univariate operations. Notice that in the computer science community these possibilities were restricted to moderate-dimensional tensors of small mode size (with rather large rank parameters) obtained from the experimental data sets.

## 2.2 Tucker decomposition for function related tensors.

### Canonical-to-Tucker approximation

It was proven and shown in numerical simulations,<sup>36,38</sup> that the rank of the (fixed accuracy) Tucker approximation to some function related tensors depends only logarithmically on the size of the discretization grid. In particular, it was shown that for tensors resulting from the grid discretization of physically relevant multidimensional functions one can find algebraically a reduced subspace in each mode of the 3D tensor, which lead

to a separable approximation of the function by a tensor product of a small number of vectors.

For a given continuous function  $f: \Omega \rightarrow \mathbb{R}$ ,  $\Omega := \prod_{\ell=1}^d [a_\ell, b_\ell] \subset \mathbb{R}^3$ , we introduced the function related 3rd order tensor, obtained by Galerkin discretization in a volume box using  $n \times n \times n$  3D Cartesian grid. In particular, the low-rank tensor approximations were calculated for functions of the Slater type  $f(x) = \exp(-\alpha||x||)$ , Newton kernel  $f(x) = \frac{1}{||x||}$ , Yukawa potential  $f(x) = \frac{e^{-\alpha||x||}}{||x||}$ , and the Helmholtz potential  $f(x) = \frac{\cos \alpha||x||}{||x||}$ ,  $x \in \mathbb{R}^3$ .

Numerical tests demonstrated that the error of the rank- $\mathbf{r}$  (with  $\mathbf{r} = (r, r, r)$ ) Tucker approximation applied to these third order tensors decays exponentially with respect to the Tucker rank  $r$ . Moreover, for fixed approximation error, the Tucker rank  $r$  scales as  $O(\log n)$  for above mentioned functions.<sup>35</sup>

Fig. 3 (bottom) shows the exponential convergence of the Tucker tensor approximation in the Tucker rank  $r$  of the Slater function in the relative Frobenius norm,  $E_{\text{FN}} = \frac{||A - A_r||}{||A||}$ . Fig. 3 and 4 demonstrate that the exponential decay of approximation error is nearly the same for both a single potential and for a sum of the same potentials distributed in nodes of a cubic lattice. This property of the Tucker decomposition will be further gainfully applied to the fast lattice summation of interaction potentials.

Motivated by these observations, we invented the canonical-to-Tucker (C2T) decomposition for function related tensors,<sup>38</sup>

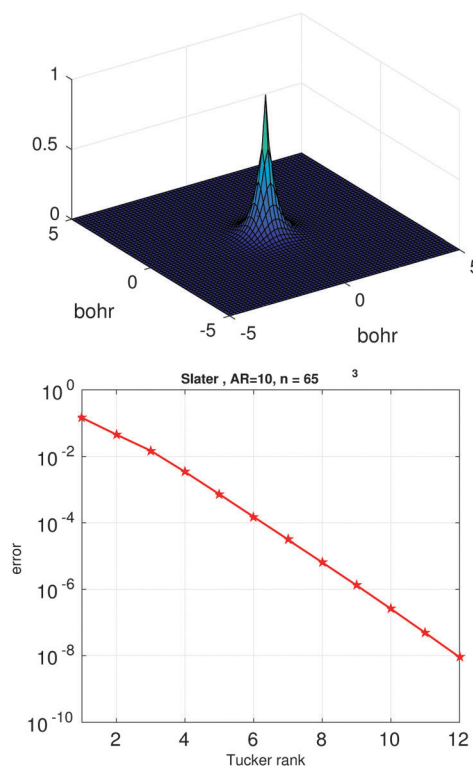


Fig. 3 The Tucker approximation error in the Frobenius norm vs. the Tucker rank  $r$  for a single Slater function.





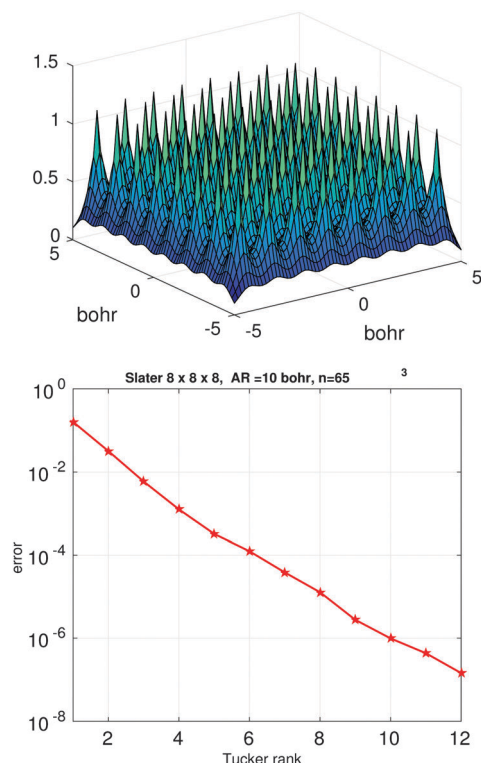


Fig. 4 The Tucker approximation error in the Frobenius norm vs. the Tucker rank  $r$  for a sum of 512 Slater potentials over the cubic  $8 \times 8 \times 8$  lattice (right).

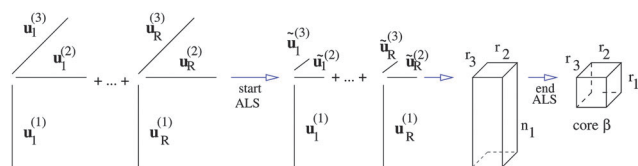


Fig. 5 Part of the canonical-to-Tucker decomposition algorithm for mode  $\ell = 1$ .

based on the reduced higher order singular value decomposition (RHOSVD) (Theorem 2.5, ref. 39). The canonical-to-Tucker algorithm (see Appendix) combined with the Tucker-to-canonical transform serves for reducing the ranks of canonical tensors with large  $R$ . Fig. 5 in its ALS part shows the algorithmic step for  $\ell = 1$ , which is repeated for every mode  $\ell = 1, 2, 3$  (see Appendix). The computational work for the multigrid tensor decomposition C2T algorithm introduced in ref. 39 exhibits linear complexity scaling with respect to all input parameters,  $O(Rnr)$ .

The multigrid Tucker decomposition algorithm applied to full format tensors<sup>39</sup> leads to the complexity scaling  $O(n^3)$ , whereas its standard version (commonly used HOSVD algorithm in computer science) scales as  $O(n^4)$ , becoming intractable even for moderate sizes of tensors.

We conclude by notice that the optimized Tucker and canonical tensor approximations can be computed by the alternating least square (ALS) iteration with initial guess obtained by HOSVD/RHOSVD approximations.

### 2.3 Matrix product states/tensor train format

The product-type representation of  $d$ th order tensors, which is called in the physical literature as the matrix product states (MPS) decomposition (or more generally, tensor network states models), was introduced and successfully applied in DMRG quantum computations,<sup>62,69,73</sup> and, independently, in quantum molecular dynamics as the multilayer (ML) MCTDH methods.<sup>54,71</sup> MPS type representations reduce the storage complexity to  $O(dr^2N)$ , where  $r$  is the maximal rank parameter.

In the recent years the various versions of the MPS-type tensor approximations were discussed and further investigated in mathematical literature including the hierarchical dimension splitting,<sup>36</sup> the tensor train (TT),<sup>14,57</sup> the quantics-TT (QTT),<sup>34</sup> as well as the hierarchical Tucker representation,<sup>26</sup> which belongs to the class of tensor network states model.

The TT format is the particular form of MPS type factorization in the case of open boundary conditions. For a given rank parameter  $\mathbf{r} = (r_0, \dots, r_d)$ , the rank- $\mathbf{r}$  TT format contains all elements  $\mathbf{A} = [a_{i_1, \dots, i_d}] \in \mathbb{R}^{n_1 \times \dots \times n_d}$ , which can be represented as the contracted products of 3-tensors, that in the index notation takes a form,

$$a_{i_1, \dots, i_d} = \sum_{\alpha_1=1}^{r_1} \dots \sum_{\alpha_d=1}^{r_d} \mathbf{a}_{\alpha_1}^{(1)}(i_1) \mathbf{a}_{\alpha_1, \alpha_2}^{(2)}(i_2) \dots \mathbf{a}_{\alpha_{d-1}}^{(d)}(i_d) \\ \equiv \mathbf{A}^{(1)}(i_1) \mathbf{A}^{(2)}(i_2) \dots \mathbf{A}^{(d)}(i_d),$$

specified by the set of column vectors,  $\mathbf{a}_{\alpha_\ell, \alpha_{\ell+1}}^{(\ell)} \in \mathbb{R}^{n_\ell}$ , ( $\ell = 1, \dots, d$ ), or equivalently by the vector-valued  $r_\ell \times r_{\ell+1}$  matrices,  $\mathbf{A}^{(\ell)} = [\mathbf{a}_{\alpha_\ell, \alpha_{\ell+1}}^{(\ell)}]$ , ( $i.e.$ , 3-tensors), cf. (2). The latter representation is written in the matrix product form, explaining the notion MPS, where  $\mathbf{A}^{(\ell)}(i_\ell)$  is  $r_{\ell-1} \times r_\ell$  matrix.

Fig. 6 illustrates the TT representation of a 5th-order tensor, where each particular entry is factorized as a product of five matrices,  $a_{i_1, i_2, \dots, i_5} = \mathbf{A}^{(1)}(i_1) \mathbf{A}^{(2)}(i_2) \dots \mathbf{A}^{(5)}(i_5)$ , where, for example,  $\mathbf{A}^{(2)}(i_2) \in \mathbb{R}^{r_1 \times r_2}$ .

The rank-structured tensor formats like canonical, Tucker and MPS/TT-type decompositions also apply to matrices. For example, the  $d$ -dimensional FFT matrix over  $N^{\otimes d}$  grid can be implemented on the rank- $k$  tensor with the linear-logarithmic cost  $O(dkN \log_2 N)$ , due to the rank-1 factorized representation

$$\mathcal{F}_N^{(d)} = (F_N^{(1)} \otimes I \dots \otimes I)(I \otimes F_N^{(2)} \dots \otimes I) \dots (I \otimes I \dots \otimes F_N^{(d)}) \\ \equiv F_N^{(1)} \otimes \dots \otimes F_N^{(d)},$$

where  $F_N^{(\ell)} \in \mathbb{R}^{N \times N}$  represents the univariate FFT matrix along mode  $\ell$ .

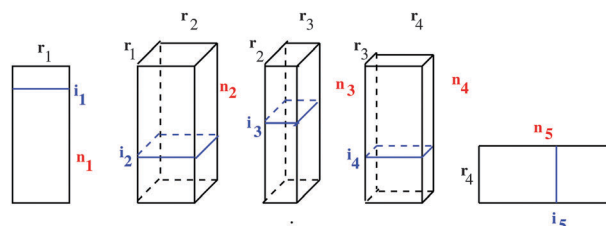


Fig. 6 Visualizing 5th-order MPS/TT tensor.

## 2.4 Quantics tensor approximation of functional vectors

In the case of large mode size, the asymptotic storage cost for a class of function related  $N$ - $d$  tensors can be reduced to  $O(d \log N)$  by using quantics-TT (QTT) tensor approximation method.<sup>34</sup> The QTT-type approximation of an  $N$ -vector with  $N = 2^L$ ,  $L \in \mathbb{N}$ , is defined as the tensor decomposition (approximation) in the canonical, TT or more general formats applied to a tensor obtained by the dyadic folding (reshaping) of the target vector to an  $L$ -dimensional  $2 \times \dots \times 2$  data array (tensor) that is thought as an element of the  $L$ -dimensional quantized tensor space.

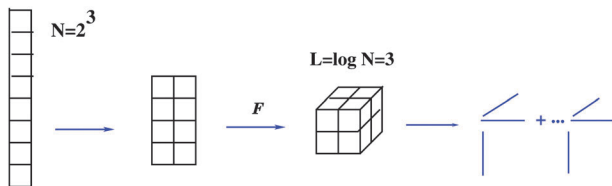
In the vector case, i.e. for  $d = 1$ , a vector  $\mathbf{x} = [x_i] \in \mathbb{R}^N$ , with  $N = 2^L$ , is reshaped to its quantics (quantized) image in  $\otimes_{j=1}^L \mathbb{R}^2$ , by dyadic folding,

$$\mathcal{F}_{2,L} : \mathbf{x} \rightarrow \mathbf{Y} = [y_{\mathbf{j}}] \in \otimes_{j=1}^L \mathbb{R}^2, \quad \mathbf{j} = (j_1, \dots, j_L), \quad j_\nu \in \{1, 2\},$$

where for fixed  $i$ , we have  $y_{\mathbf{j}} = x_i$ , with  $j_\nu = j_\nu(i) = C_{\nu-1} (\nu = 1, \dots, L)$  being defined *via* binary coding, i.e. the coefficients  $C_{\nu-1} \in \{0, 1\}$  are found from the binary representation of  $i - 1$ ,

$$i - 1 = C_0 + C_1 2^1 + \dots + C_{L-1} 2^{L-1} \equiv \sum_{\nu=1}^L (j_\nu - 1) 2^{\nu-1}.$$

Next figure visualizes the QTT approximation process.



Suppose that the quantics image for an  $N$ -vector, i.e. an element of  $L$ -dimensional quantized tensor space  $\otimes_{j=1}^L \mathbb{R}^2$  with  $L = \log_2 N$ , can be represented (approximated) by the low-rank canonical or TT tensor of order  $L$ , thus introducing the QTT approximation of an  $N$ -vector. Given rank parameters  $\{r_k\}$  ( $k = 1, \dots, L$ ), the QTT approximation of an  $N$ -vector requires a number of representation parameters estimated by

$$2r^2 \log_2 N \ll N, \quad \text{where } r_k \leq r, \quad k = 1, \dots, L,$$

providing log-volume complexity scaling in the size of initial vector,  $N$ . For  $d > 1$  the construction is similar.<sup>34</sup>

The power of QTT approximation method is explained by the theoretical substantiation of the QTT approximation properties discovered in ref. 34 and establishing the perfect rank- $r$  decomposition for the wide class of function-related tensors obtained by sampling continuous functions over uniform (or properly refined) grid:

- $r = 1$  for complex exponents,
- $r = 2$  for trigonometric functions and plain-waves,
- $r \leq m + 1$  for polynomials of degree  $m$ ,
- $r$  is a small constant for wavelet functions, Gaussians, etc.

all independently on the vector size  $N$ .

Notice that the notion quantics (or quantized) tensor approximation (with a shorthand QTT) originally introduced in 2009, see ref. 34, is reminiscent of the entity “quantum of

information”, that mimics the minimal possible mode size ( $n = 2$ ) of the quantized image.

Concerning the matrix case, it was first found in ref. 56 by numerical tests that in some cases the dyadic reshaping of an  $N \times N$  matrix with  $N = 2^L$  may lead to a small TT-rank of the resultant matrix rearranged to the tensor form. The efficient low-rank QTT representation for a class of discrete multidimensional elliptic operators (matrices) and their inverse was proven in ref. 33. Moreover, based on the QTT approximation, the important algebraic matrix operations like FFT, convolution and wavelet transforms can be implemented by superfast algorithms in  $O(\log_2 N)$ -complexity, see survey paper<sup>37</sup> and references therein.

## 2.5 Rank-structured tensor operations in 1D complexity

The rank-structured tensor representation provides 1D complexity of multilinear operations with multidimensional tensors. Rank-structured tensor representation provides fast multi-linear algebra with linear complexity scaling in the dimension  $d$ .

For given canonical tensors  $\mathbf{A}$  and  $\mathbf{B}$  as in (1), with ranks  $R_a$  and  $R_b$ , respectively, their Euclidean scalar product can be computed by univariate operations

$$\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i=1}^{R_a} \sum_{j=1}^{R_b} c_i c_j \prod_{\ell=1}^d \langle \mathbf{a}_i^{(\ell)}, \mathbf{b}_j^{(\ell)} \rangle, \quad (4)$$

at the expense  $O(dnR_aR_b)$ .

The Hadamard (entrywise) product of tensors  $\mathbf{A}, \mathbf{B}$  is defined by  $\mathbf{Y} = [\mathbf{y}_{i_1 \dots i_d}] = \mathbf{A} \odot \mathbf{B}$ , where  $y_{i_1 \dots i_d} = a_{i_1 \dots i_d} b_{i_1 \dots i_d}$ . For canonical tensors  $\mathbf{A}$  and  $\mathbf{B}$  given in form (1), the Hadamard product is calculated in  $O(dnR_aR_b)$  operations by 1D entrywise products of vectors,

$$\mathbf{A} \odot \mathbf{B} = \sum_{i=1}^{R_a} \sum_{j=1}^{R_b} c_i c_j \left( \mathbf{a}_i^{(1)} \odot \mathbf{b}_j^{(1)} \right) \otimes \dots \otimes \left( \mathbf{a}_i^{(d)} \odot \mathbf{b}_j^{(d)} \right). \quad (5)$$

Summation of two tensors in the canonical format  $\mathbf{C} = \mathbf{A} + \mathbf{B}$  is performed by a simple concatenation of their factor matrices,  $\mathbf{A}^{(\ell)} = [\mathbf{a}_1^{(\ell)}, \dots, \mathbf{a}_{R_a}^{(\ell)}]$  and  $\mathbf{B}^{(\ell)} = [\mathbf{b}_1^{(\ell)}, \dots, \mathbf{b}_{R_b}^{(\ell)}]$ ,

$$\mathbf{C}^{(\ell)} = [\mathbf{a}_1^{(\ell)}, \dots, \mathbf{a}_{R_a}^{(\ell)}, \mathbf{b}_1^{(\ell)}, \dots, \mathbf{b}_{R_b}^{(\ell)}] \in \mathbb{R}^{n_\ell \times (R_a + R_b)}. \quad (6)$$

The rank of the resulting canonical tensor increases up to  $R_c = R_a + R_b$ .

In electronic structure calculations, the 3D convolution transform with the Newton kernel,  $\frac{1}{\|x - y\|}$ , is the most computationally expensive operation. The tensor method to compute convolution over large  $n \times n \times n$  Cartesian grids in  $O(n \log n)$  complexity was introduced in ref. 35. Given canonical tensors  $\mathbf{A}, \mathbf{B}$ , their convolution product is represented by the sum of tensor products of 1D convolutions,

$$\mathbf{A} \times \mathbf{B} = \sum_{i=1}^{R_a} \sum_{j=1}^{R_b} c_i c_j \left( \mathbf{a}_i^{(1)} * \mathbf{b}_j^{(1)} \right) \otimes \left( \mathbf{a}_i^{(2)} * \mathbf{b}_j^{(2)} \right) \otimes \left( \mathbf{a}_i^{(3)} * \mathbf{b}_j^{(3)} \right), \quad (7)$$



where  $\mathbf{a}_k^{(\ell)} * \mathbf{b}_m^{(\ell)}$  denotes the univariate convolution product of  $n$ -vectors. The cost of tensor convolution in both storage and time is estimated by  $O(R_a R_b n \log n)$ . The resulting algorithm considerably outperforms the conventional 3D FFT-based approaches of complexity  $O(n^3 \log n)$ , see numerics in ref. 39.

The sequences of rank-structured operations on matrices and vectors normally lead to the increase of tensor ranks, usually being multiplied or added after each operation. The necessary rank reduction in the Tucker and MPS type formats can be implemented by stable algorithms based on the higher order SVD. In the physical community the HOSVD-type algorithms are known since longer as the Schmidt decomposition.<sup>62,69,70</sup> In the case of canonical tensors the rank reduction can be performed by the RHOSVD algorithm based on the canonical-to-Tucker and then Tucker-to-canonical transforms described and analyzed in ref. 38, 39 and 49 (see Section 2.2 and Appendix), which demonstrated the stable behavior for most of examples in the Hartree-Fock calculations we considered so far. The stability conditions for the RHOSVD have been discussed in ref. 39 and 49.

### 3 Tensor calculus for the Hartree-Fock equation

#### 3.1 Calculation of multi-dimensional integrals

Tensor-structured calculation of the multidimensional convolution integral operators with the Newton kernel have been introduced in ref. 39, 40 and 42, where on the examples of the Hartree and exchange operators in the Hartree-Fock equation, it was shown that calculation of the 3D and 6D convolution integrals can be reduced to a combination of 1D Hadamard products, 1D convolutions and 1D scalar products.

The molecule is embedded in a certain fixed computational box  $\Omega = [-b, b]^3 \in \mathbb{R}^3$ , as in Fig. 7.† For a given discretization parameter  $n \in \mathbb{N}$ , we use the equidistant  $n \times n \times n$  tensor grid  $\omega_{3,n} = \{x_i\}$ ,  $\mathbf{i} \in \mathcal{I} := \{1, \dots, n\}^3$ , with the mesh-size  $h = 2b/(n+1)$ . In calculations of integral terms, the Gaussian basis functions  $g_k(x)$ ,  $x \in \mathbb{R}^3$ , are approximated by sampling their values at the centers of discretization intervals, as in Fig. 8, using one-dimensional piecewise constant basis functions  $\bar{g}_k(x) \approx \prod_{\ell=1}^3 \bar{g}_k^{(\ell)}(x_\ell)$ ,  $\ell = 1, 2, 3$ , yielding their rank-1 tensor representation,

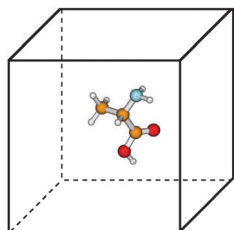


Fig. 7 Glycine amino acid in a computational box.

† Usually for small to moderate size molecules we use the computational box of size  $40^3$  bohr<sup>3</sup>.

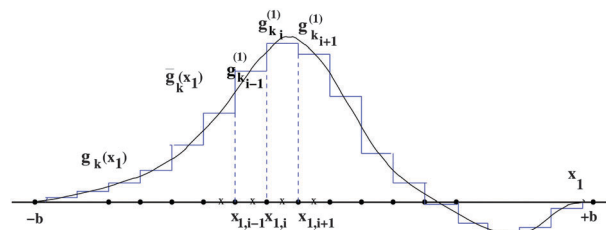


Fig. 8 Approximation of the Gaussian-type basis function by a piecewise constant function.

$$\mathbf{G}_k = \mathbf{g}_k^{(1)} \otimes \mathbf{g}_k^{(2)} \otimes \mathbf{g}_k^{(3)} \in \mathbb{R}^{n \times n \times n}, \quad k = 1, \dots, N_b. \quad (8)$$

Let us consider the tensor calculation of the Hartree potential

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

and the corresponding Coulomb matrix,

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

where the electron density,  $\rho(x) = 2 \sum_{a=1}^{N_{\text{orb}}} (\varphi_a)^2$ , is represented in terms of molecular orbitals  $\varphi_a(x) = \sum_{k=1}^{N_b} c_{a,k} g_k(x)$ . Given the discrete tensor representation of basis functions (8), the electron density is approximated using 1D Hadamard products of rank-1 tensors (instead of product of Gaussians),

$$\begin{aligned} \rho \approx \Theta &= 2 \sum_{a=1}^{N_{\text{orb}}} \sum_{k=1}^{N_b} \sum_{m=1}^{N_b} c_{a,k} c_{a,m} \left( \mathbf{g}_k^{(1)} \odot \mathbf{g}_m^{(1)} \right) \\ &\otimes \left( \mathbf{g}_k^{(2)} \odot \mathbf{g}_m^{(2)} \right) \otimes \left( \mathbf{g}_k^{(3)} \odot \mathbf{g}_m^{(3)} \right) \in \mathbb{R}^{n \times n \times n}. \end{aligned}$$

Further, the representation of the Newton kernel  $\frac{1}{\|x - y\|}$  by a canonical rank- $R_N$  tensor<sup>7</sup> is used (see Appendix for details),

$$\mathbf{P}_R = \sum_{q=1}^{R_N} \mathbf{p}_q^{(1)} \otimes \mathbf{p}_q^{(2)} \otimes \mathbf{p}_q^{(3)} \in \mathbb{R}^{n \times n \times n}. \quad (9)$$

Since large ranks make tensor operations inefficient, the multi-grid canonical-to-Tucker and Tucker-to-canonical algorithms should be applied to reduce the initial rank of  $\Theta \mapsto \Theta'$  by several orders of magnitude, from  $N_b^2/2$  to  $R_\rho \ll N_b^2/2$ . Then the 3D tensor representation of the Hartree potential is calculated by using the 3D tensor product convolution, which is a sum of tensor products of 1D convolutions,

$$\begin{aligned} V_H \approx \mathbf{V}_H &= \Theta' * \mathbf{P}_R \\ &= \sum_{m=1}^{R_\rho} \sum_{q=1}^{R_N} c_m \left( \mathbf{u}_m^{(1)} * \mathbf{p}_q^{(1)} \right) \otimes \left( \mathbf{u}_m^{(2)} * \mathbf{p}_q^{(2)} \right) \otimes \left( \mathbf{u}_m^{(3)} * \mathbf{p}_q^{(3)} \right). \end{aligned}$$

The Coulomb matrix entries  $J_{km}$  are obtained by 1D scalar products of  $\mathbf{V}_H$  with the Galerkin basis consisting of rank-1 tensors,

$$J_{km} \approx \langle \mathbf{G}_k \odot \mathbf{G}_m, \mathbf{V}_H \rangle, \quad k, m = 1, \dots, N_b.$$



**Table 1** Times (s) for the 3D tensor product convolution vs. 3D FFT convolution

$n^3$	1024 <sup>3</sup>	2048 <sup>3</sup>	4096 <sup>3</sup>	8192 <sup>3</sup>	16384 <sup>3</sup>
FFT <sub>3</sub>	~6000	—	—	—	~2 years
$C \times C$	8.8	20.0	61.0	157.5	299.2
C2T	6.9	10.9	20.0	37.9	86.0

The cost of 3D tensor product convolution is  $O(n \log n)$  instead of  $O(n^3 \log n)$  for the standard benchmark 3D convolution using the 3D FFT. Table 1 shows CPU times (s) for the Matlab computation of  $V_H$  for H<sub>2</sub>O molecule<sup>39</sup> on a SUN station using 8 Opteron Dual-Core/2600 processors (times for 3D FFT for  $n \geq 1024$  are obtained by extrapolation). C2T shows the time for the canonical-to-Tucker rank reduction. In a similar way, the algorithm for 3D grid-based tensor-structured calculation of 6D integrals in the exchange potential operator was introduced in ref. 40,  $K_{km} = \sum_{a=1}^{N_{\text{orb}}} K_{km,a}$  with

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

the contribution from the  $a$ -th orbital are approximated by tensor anzats,

$$K_{km,a} := \left\langle \mathbf{G}_k \odot \left[ \sum_{\mu=1}^{N_b} c_{\mu a} \mathbf{G}_{\mu} \right], \left[ \mathbf{G}_m \odot \sum_{\nu=1}^{N_b} c_{\nu a} \mathbf{G}_{\nu} \right] * \mathbf{P}_R \right\rangle.$$

Here, the tensor product convolution is first calculated for each  $a$ th orbital, and then scalar products in canonical format yield the contributions to entries of the exchange Galerkin matrix from the  $a$ -th orbital.

These algorithms were employed in the first tensor-structured solver using 3D grid-based evaluation of the Coulomb and exchange matrices in 1D complexity at every step of SCF EVP iteration.<sup>41,42</sup> A sequence of dyadically refined 3D Cartesian grids was used for reducing time in first iterations, with an  $\varepsilon$  convergence criterion for switching to larger grids. This is a nonstandard computational scheme avoiding calculation of the two-electron integrals. The accuracy for small molecules like H<sub>2</sub>O and CH<sub>4</sub> was of the order of  $10^{-4}$  Hartree. Though time performance of this solver was not compatible with the standard Hartree-Fock packages it was the first proof of concept for the tensor numerical methods.

### 3.2 3D grid-based calculation of the two-electron integrals

The basic tensor-structured Hartree-Fock solver employs the factorized 3D grid-based calculation of the two-electron integrals tensor,  $\mathbf{B} = [b_{\mu\nu\kappa\lambda}]$ , in the form

$$b_{\mu\nu\kappa\lambda} = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{g_{\mu}(x) g_{\nu}(x) g_{\kappa}(y) g_{\lambda}(y)}{\|x-y\|} dx dy \quad (10)$$

$$= \langle \mathbf{G}_{\mu} \odot \mathbf{G}_{\nu}, \mathbf{P}_R * (\mathbf{G}_{\kappa} \odot \mathbf{G}_{\lambda}) \rangle_{n^{\otimes 3}},$$

by using univariate tensor operations. Introduce the side matrices  $G^{(\ell)}$  representing on the grid the full set of canonical

vectors composing the products of the Gaussian basis functions,  $\{\mathbf{G}_{\mu} \odot \mathbf{G}_{\nu}\}$ ,

$$G^{(\ell)} = \left[ \mathbf{g}_{\mu}^{(\ell)} \odot \mathbf{g}_{\nu}^{(\ell)} \right]_{1 \leq \mu, \nu \leq N_b} \in \mathbb{R}^{n \times N_b^2} \quad \ell = 1, 2, 3,$$

where in most of our Hartree-Fock calculations grids of size  $n^3 = 32 \times 10^3$  or  $n^3 = 64 \times 10^3$  have been used. It was found that the large matrices  $G^{(\ell)}$  of size  $n \times N_b^2$  (e.g.  $N_b^2 = 40\,000$  for Alanine amino acid) can be approximated with high accuracy by low rank matrices with the rank parameter bounded by  $R_{\ell} \leq N_b$ .<sup>44,45</sup> The corresponding low-rank factorizations (“1D density fitting”) in the form (for  $\ell = 1, 2, 3$ )

$$G^{(\ell)} \cong U^{(\ell)} V^{(\ell)T}, \quad U^{(\ell)} \in \mathbb{R}^{n \times R_{\ell}}, \quad V^{(\ell)} \in \mathbb{R}^{N_b^2 \times R_{\ell}}, \quad (11)$$

is computed by the truncated Cholesky decomposition of the symmetric, positive definite  $G^{(\ell)} G^{(\ell)T}$  (see ref. 44 and 45 for more details).

Based on factorization (11), the number of convolutions in (10) is reduced dramatically from  $N_b^2$  to  $N_b$  at most (say from 40 000 to 200). In fact, using canonical factors from the rank- $R$  canonical tensor  $\mathbf{P}_R$  representing the Newton kernel (9) (see (31) in Appendix) we, first, precompute the set of “convolution” matrices for every space variable  $\ell$ ,  $\ell = 1, 2, 3$ ,

$$M_k^{(\ell)} = U^{(\ell)T} \left( \mathbf{p}_k^{(\ell)} *_{\mathbf{n}} U^{(\ell)} \right) \in \mathbb{R}^{R_{\ell} \times R_{\ell}}, \quad k = 1, \dots, R, \quad (12)$$

which includes the convolution products with  $R_{\ell} \leq N_b$  column vectors of the matrix  $U^{(\ell)} \in \mathbb{R}^{n \times R_{\ell}}$  instead of  $N_b^2/2$  convolutions in the initial formulation.

Given matrices  $M_k^{(\ell)}$ , then the resulting 4-th order TEI tensor is represented in a matrix form

$$B = \text{mat}(\mathbf{B}) \cong B_e := \sum_{k=1}^R \odot_{\ell=1}^3 V^{(\ell)} M_k^{(\ell)} V^{(\ell)T} \in \mathbb{R}^{N_b^2 \times N_b^2}.$$

The above nonstandard factorization of the TEI matrix  $B$  allows to reduce dramatically the computational cost of the standard Cholesky factorization schemes<sup>3,5,30</sup> applied to the reduced-rank symmetric positive definite matrix  $B$ . In fact, the low-rank Cholesky decomposition of  $B$  is calculated in the following sequence. First, the diagonal elements of  $B$  are calculated as

$$B(i, i) = \sum_{k=1}^R \odot_{\ell=1}^3 V^{(\ell)}(i, :) M_k^{(\ell)} V^{(\ell)}(:, i)^T.$$

Then the selected columns of the matrix  $B$ , required for the rank-truncated Cholesky factorization scheme, are computed by the following fast tensor operations

$$B(:, j^*) = \sum_{k=1}^R \odot_{\ell=1}^3 V^{(\ell)} M_k^{(\ell)} V^{(\ell)}(:, j^*)^T,$$

leading to representation of the matrix  $B$  in the form of rank- $R_B$  approximate decomposition<sup>44,45</sup>

$$B := [b_{\mu\nu\kappa\lambda}] \approx LL^T, \quad \text{with } L \in \mathbb{R}^{N_b^2 \times R_B}, \quad R_B \sim N_b.$$

This algorithm is much faster than the direct Cholesky decomposition of the matrix  $B$  with on-the-fly computation of the required column vectors.





Table 2 represents times (s) for 3D grid-based calculation of the directional density fitting and the TEI tensor (electron repulsion integrals) for H<sub>2</sub>O molecule in a box  $[-20,20]^3$  bohr<sup>3</sup>, performed in Matlab on a 2-Intel Xeon Hexa-Core/2677. Time for convolution integrals in (12) scales almost linearly in the 1D grid-size  $n$  as expected by theory.

Note that the algorithm for fast tensor computation of TEI is described in detail for the case of rank-1 GTO basis functions in form (8). However, this approach can be also applied to basis sets with low separation rank, say, to Slater-type basis functions or some kind of agglomerated basis. The advantage of use the “more compressed” basis is the essential reduction in the number of basis functions,  $N_b$ , that enters to the size of TEI tensor in the fourth power. The payoff for such a reduction in the tensor size is the additional costs for calculation of the particular entries of TEI tensor in (10), since in that case the number of terms in tensor representation of the scalar products in (10) increases quadratically in the separation rank of basis functions. The resulting effect from implementation of the reduced basis sets (with larger separation ranks) depends on subsequent (post) Hartree–Fock calculations and it will be studied in detail in the forthcoming paper.

### 3.3 Core Hamiltonian

The Galerkin representation of the 3D Laplace operator in the nonlocal Gaussian basis  $\{g_k(x)\}_{1 \leq k \leq N_b}$ ,  $x \in \mathbb{R}^3$ , leads to the fully populated matrix  $A_g = [a_{km}] \in \mathbb{R}^{N_b \times N_b}$ . Tensor calculation of the matrix entries  $a_{km}$  for the discrete Laplacian  $A_g$  in the separable Gaussian basis is reduced to 1D matrix operations<sup>46</sup> involving the FEM Laplacian  $\Delta_3$ , defined on  $n \times n \times n$  grid,

$$\Delta_3 = \Delta_1^{(1)} \otimes I^{(2)} \otimes I^{(3)} + I^{(1)} \otimes \Delta_1^{(2)} \otimes I^{(3)} + I^{(1)} \otimes I^{(2)} \otimes \Delta_1^{(3)},$$

where  $\Delta_1 = \frac{1}{h} \text{tridiag}\{-1, 2, -1\}$ . Specifically, we have

$$a_{km} = \langle \Delta_3 \mathbf{G}_k, \mathbf{G}_m \rangle,$$

where  $\mathbf{G}_k$  is the tensor representation of Gaussian basis functions using the piecewise linear finite elements (see Fig. 9).

In the case of large  $n \times n \times n$  grids, this calculation can be implemented with logarithmic cost in  $n$  by using the low-rank QTT representation of the large matrix  $\Delta_3$ , see ref. 33, 43 and 46.

For tensor calculation of the nuclear potential operator

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

we apply the rank-1 windowing operator,  $\mathcal{W}_\alpha = \mathcal{W}_\alpha^{(1)} \otimes \mathcal{W}_\alpha^{(2)} \otimes \mathcal{W}_\alpha^{(3)}$ , for shifting the reference Newton kernel  $\tilde{\mathbf{P}}_R \in \mathbb{R}^{2n \times 2n \times 2n}$

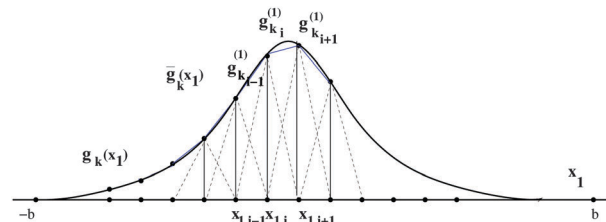


Fig. 9 Discretization of a Gaussian by piecewise linear finite elements.

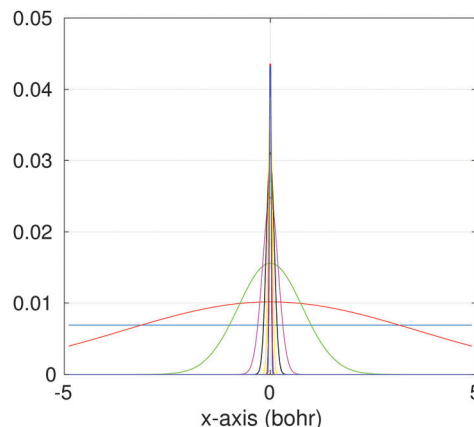


Fig. 10 Several vectors of the canonical representation of the Newton kernel along one of variables.

according to the coordinates of nuclei in a molecule (see Section 5 and Appendix). Then the resulting nuclear potential,  $\mathbf{P}_c \in \mathbb{R}^{n \times n \times n}$ , is obtained as a direct tensor sum of shifted potentials,<sup>46</sup>

$$\begin{aligned} \mathbf{P}_c &= \sum_{\alpha=1}^M Z_\alpha \mathcal{W}_\alpha \tilde{\mathbf{P}}_R \\ &= \sum_{\alpha=1}^M Z_\alpha \sum_{q=1}^R \mathcal{W}_\alpha^{(1)} \tilde{\mathbf{P}}_q^{(1)} \otimes \mathcal{W}_\alpha^{(2)} \tilde{\mathbf{P}}_q^{(2)} \otimes \mathcal{W}_\alpha^{(3)} \tilde{\mathbf{P}}_q^{(3)}. \end{aligned}$$

This leads to the following representation of the Galerkin matrix,  $V_c = [\bar{v}_{km}]$ , by tensor operations

$$\bar{v}_{km} = \int_{\mathbb{R}^3} V_c(x) \bar{g}_k(x) \bar{g}_m(x) dx \approx \langle \mathbf{G}_k \odot \mathbf{G}_m, \mathbf{P}_c \rangle, \quad 1 \leq k, m \leq N_b.$$

Fig. 10, shows several vectors of the canonical representation of the Coulomb kernel along one of variables. Fig. 11, represents the cross-section of the resulting nuclear potential  $\mathbf{P}_c$  for C<sub>2</sub>H<sub>5</sub>OH molecule.

### 3.4 Black-box tensor solver

The tensor-structured Hartree–Fock solver<sup>46</sup> based on factorized calculation of the two-electron integrals<sup>44</sup> includes efficient tensor implementation of the MP2 energy correction<sup>45</sup> scheme. Though it is yet implemented in Matlab, its performance in time and accuracy is compatible with the standard packages based on analytical evaluation of the two-electron integrals. Due to 1D complexity of all calculations, it enables 3D grids of the size  $10^{15}$ ,

Table 2 Times (s) for 3D grid-based calculation of the directional density fitting and the TEI for H<sub>2</sub>O molecule

$n^3$	16 384 <sup>3</sup>	32 768 <sup>3</sup>	65 536 <sup>3</sup>	131 072 <sup>3</sup>
Mesh (bohr)	0.0024	0.0012	$6 \times 10^{-4}$	$3 \times 10^{-4}$
Had. prod.	1.6	3.4	12	19
Density fit.	0.5	1.0	2.0	4.0
3D conv. time	69	151	698	496
Chol. time	2.2	2.2	2.2	2.2



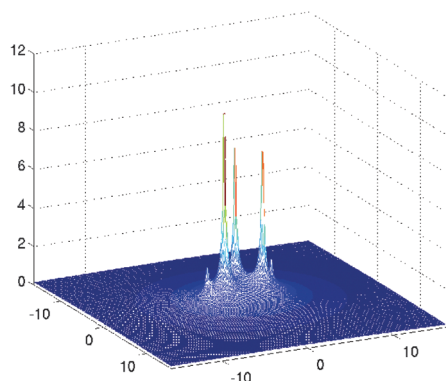


Fig. 11 A sum of the nuclear potentials for ethanol molecule. Size of the computational box is given in bohr.

yielding mesh size of the order of atomic radii,  $10^{-4}$  Å. That ensures high accuracy of calculations, which is controlled by the  $\varepsilon$ -ranks of tensor truncation.

The solver works in a black-box way: input the grid-based basis-functions and coordinates of nuclei in a molecule and start the program. Calculation of TEI for H<sub>2</sub>O on grids  $32\,768^3$  takes two minutes on a laptop. The time for TEI with  $n^3 = 131\,072^3$  for Alanine amino acid takes approximately one hour in Matlab, including incorporated density fitting.

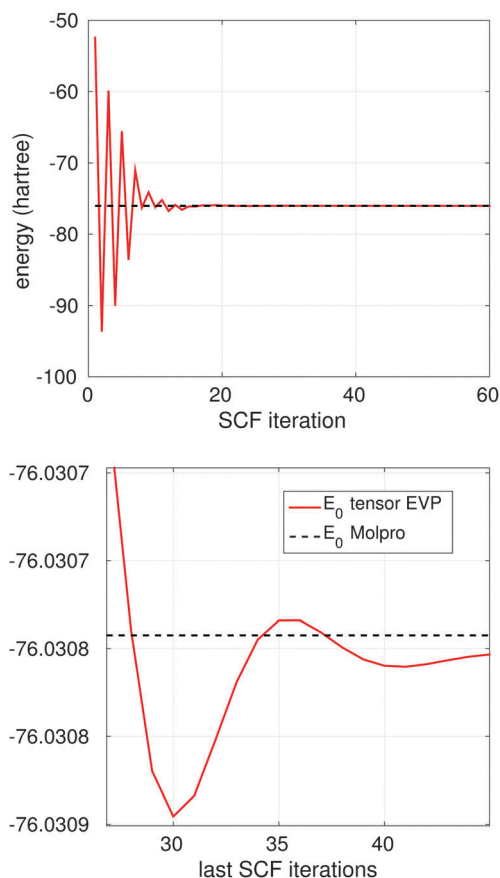


Fig. 12 Top: convergence of the ground state energy to  $E_0$  in *ab initio* SCF Hartree–Fock iterations for H<sub>2</sub>O; bottom: zooming of final 20 iterations for H<sub>2</sub>O.

Next examples compare the results from benchmark Molpro program with calculations by the tensor-based solver by using the same Gaussian basis, but now discretized on 3D Cartesian grid. For all molecules we use the “cc-pVDZ” Gaussian basis set. The core Hamiltonian part in these calculations is taken from Molpro. Note that since in the tensor solver the density fitting is included in “blind” calculation of TEI, it is not easy to compare the CPU times of our calculations with those in “*ab initio*” procedure in the standard programs, because the density fitting step there is usually considered as off-line pre-computing.

Fig. 12, top, shows convergence to the ground state energy  $E_0$  in *ab initio* iterations for H<sub>2</sub>O molecule (cc-pVDZ-41), where TEI is computed on the 3D grid of size  $1\,31\,072^3$ . Fig. 12, bottom presents the zoom of the last 20 iterations. The ground state energy from Molpro ( $E_0 = 76.308$  hartree) corresponding to the basis set cc-pVDZ-41 with 41 Gaussian-type basis functions is shown by the black dashed line.

Fig. 13 (top) shows the SCF EVP iteration history for Glycine amino acid (cc-pVDZ-170), where dashed line indicates convergence of the residual and the red solid line shows convergence of the error  $\Delta E_{0,g}$  in ground state energy compared with the result from MOLPRO package<sup>72</sup> with the same basis set,

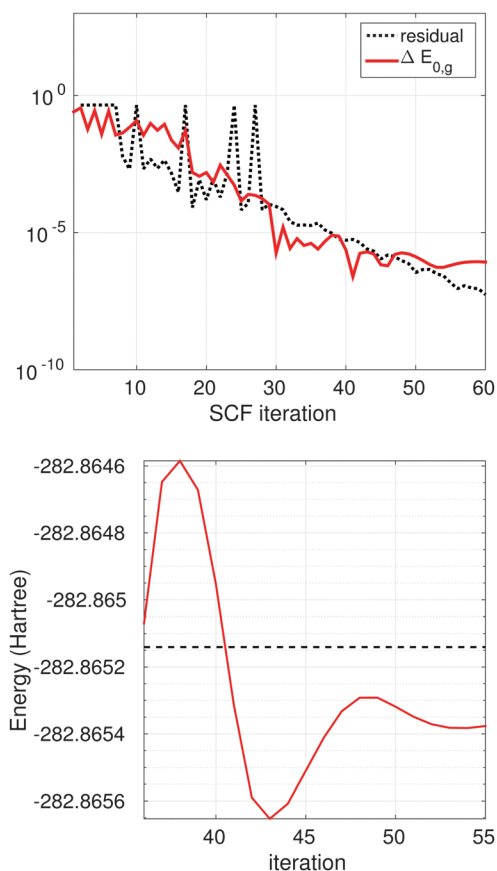


Fig. 13 Top: convergence of the residual in *ab initio* Hartree–Fock iteration for glycine (C<sub>2</sub>H<sub>5</sub>NO<sub>2</sub>) molecule. Red line shows the difference with  $E_0$  from molpro at every iteration. Bottom: ground state energy  $E_{0,g}$  (red solid line) at final 20 iterations for glycine molecule, dashed line shows  $E_0$  from molpro.



**Table 3** Time (s) for one *ab initio* SCF EVP iteration and accuracy of *ab initio* solution with respect to grid-size in TEI calculations. Matlab on an Intel Xeon X5650

	H <sub>2</sub> O	H <sub>2</sub> O <sub>2</sub>	C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub>
$N_{\text{orb}}, N_{\text{b}}$	5; 41	9; 68	20; 170
Time	0.35	0.55	6.0
$\Delta E_{0,g}, (65\ 536^3)$	$3.0 \times 10^{-7}$	$8.0 \times 10^{-8}$	$9.1 \times 10^{-7}$
$\Delta E_{0,g}, (131\ 072^3)$	$1.4 \times 10^{-7}$	$3.9 \times 10^{-8}$	$8.0 \times 10^{-7}$

$$\Delta E_{0,g} = |E_{0,g} - E_0|/|E_0|. \quad (13)$$

Fig. 13 (bottom) illustrates the convergence of the ground state energy at final 20 iterations for Glycine, dashed line is the energy from Molpro.

The first row in Table 3 shows the number of orbitals and basis functions used in Hartree–Fock calculations for H<sub>2</sub>O (cc-pVDZ-41), H<sub>2</sub>O<sub>2</sub> (cc-pVDZ-68), and C<sub>2</sub>H<sub>5</sub>NO<sub>2</sub> (cc-pVDZ-170) molecules. Second row represents times (s) for one step of SCF iteration in *ab initio* solution of the Hartree Fock EVP. Next two rows show the relative difference in energy (13), for different grid sizes used in TEI calculations. This numerics demonstrates that in the case of fine enough spacial  $n \times n \times n$ -grids the accuracy in 7–8 digits (*i.e.* relative accuracy about  $10^{-7}$ – $10^{-8}$ ) can be achieved for moderate size molecules up to small amino acids. All calculations are performed in the computational box of size  $[-20, 20]^3$  bohr<sup>3</sup>. This tensor-based solver can be considered as the computational tool for trying the alternatives to Gaussian-type basis sets.

## 4 From MP2 energy correction to excited states

### 4.1 MP2 correction scheme by using tensor formats

Given the set of Hartree–Fock molecular orbitals  $\{C_p\}$  and the corresponding energies  $\{\epsilon_p\}$ ,  $p = 1, 2, \dots, N_{\text{b}}$ , where  $\{C_i\}$  and  $\{C_a\}$  denote the occupied and virtual orbitals, respectively. First, one has to transform the TEI matrix  $B = [b_{\mu\nu, \lambda\sigma}]$ , corresponding to the initial AO basis set, to those represented in the molecular orbital (MO) basis,

$$V = [v_{ia,jb}] : v_{iajb} = \sum_{\mu, \nu, \lambda, \sigma=1}^{N_{\text{b}}} C_{\mu i} C_{\nu a} C_{\lambda j} C_{\sigma b} b_{\mu\nu, \lambda\sigma}, \quad (14)$$

where  $a, b \in \mathcal{I}_{\text{v}}$ ,  $i, j \in \mathcal{I}_{\text{o}}$ , and  $\mathcal{I}_{\text{o}} := \{1, \dots, N_{\text{orb}}\}$ ,  $\mathcal{I}_{\text{v}} := \{N_{\text{orb}} + 1, \dots, N_{\text{b}}\}$ , with  $N_{\text{orb}}$  denoting the number of occupied orbitals. In the following, we shall use the notation

$$N_{\text{v}} = N_{\text{b}} - N_{\text{orb}}, \quad N_{\text{ov}} = N_{\text{orb}} N_{\text{v}}.$$

The straightforward computation of the matrix  $V$  by above representation makes the dominating impact to the overall numerical cost of order  $O(N_{\text{b}}^5)$ . The method of complexity  $O(N_{\text{b}}^4)$  based on the low-rank tensor decomposition of the matrix  $V$  was introduced in ref. 45. Indeed, it can be shown that the rank  $R_{\text{B}} = O(N_{\text{b}})$  approximation to the TEI matrix  $B \approx LL^T$ , with the  $N \times R_{\text{B}}$  Cholesky factor  $L$ , allows to introduce the low-rank representation of the matrix  $V$ , (see ref. 6 and 45)

$$V = L_V L_V^T, \quad L_V \in \mathbb{R}^{N_{\text{ov}} \times R_{\text{B}}},$$

and then reduce the asymptotic complexity of calculations to  $O(N_{\text{b}}^4)$ .

Given the tensor  $V = [v_{iajb}]$ , the second order MP2 perturbation to the HF energy is calculated by

$$E_{\text{MP2}} = - \sum_{a,b \in I_{\text{vir}}} \sum_{i,j \in I_{\text{occ}}} \frac{v_{iajb} (2v_{iajb} - v_{ibja})}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}, \quad (15)$$

where the real numbers  $\epsilon_k$ ,  $k = 1, \dots, N_{\text{b}}$ , represent the HF eigenvalues.

Introducing the so-called doubles amplitude tensor  $\mathbf{T}$ ,

$$\mathbf{T} = [t_{iajb}] : t_{iajb} = \frac{(2v_{iajb} - v_{ibja})}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}, \quad a, b \in I_{\text{vir}}; i, j \in I_{\text{occ}},$$

the MP2 perturbation takes the form of a simple scalar product of tensors,

$$E_{\text{MP2}} = -\langle \mathbf{V}, \mathbf{T} \rangle = -\langle \mathbf{V} \odot \mathbf{T}, \mathbf{1} \rangle,$$

where  $\mathbf{1}$  denotes the rank-1 all-ones tensor. Introducing the low  $\epsilon$ -rank reciprocal “energy” tensor

$$\mathbf{E} = [e_{abij}] := \left[ \frac{1}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \right], \quad a, b \in I_{\text{vir}}; i, j \in I_{\text{occ}}, \quad (16)$$

and the partly transposed tensor (transposition in indices  $a$  and  $b$ )

$$\mathbf{V}' = [v'_{iajb}] := [v_{ibja}],$$

allows to decompose the doubles amplitude tensor  $\mathbf{T}$  as follows

$$\mathbf{T} = \mathbf{T}^{(1)} + \mathbf{T}^{(2)} = 2\mathbf{V} \odot \mathbf{E} - \mathbf{V}' \odot \mathbf{E}. \quad (17)$$

Notice that the denominator in (15) remains strongly positive if  $\epsilon_a > 0$  for  $a \in I_{\text{vir}}$  and  $\epsilon_i < 0$  for  $i \in I_{\text{occ}}$ . The latter condition (nonzero homo lumo gap) allows to prove the low  $\epsilon$ -rank decomposition of the tensor  $\mathbf{E}$ .<sup>44,45</sup>

Each term in the right-hand side in (17) can be treated separately by using rank-structured tensor decompositions of  $\mathbf{V}$  and  $\mathbf{E}$ , possibly combined with various symmetries and data sparsity. Numerical tests illustrating the tensor approach to the MP2 energy correction are presented in ref. 45.

### 4.2 Toward low-rank approximation of the Bethe–Salpeter equation for calculation of excitation energies

One of the commonly used approaches for calculation of the excited states in molecules and solids, along with the time-dependent DFT, is based on the solution of the Bethe–Salpeter equation (BSE), see for example.<sup>11,61</sup> The BSE approach leads to the challenging computational task on the solution of the eigenvalue problem for determining the excitation energies  $\omega_n$ , governed by a large fully populated matrix of size  $O(N_{\text{ov}}^2) \approx O(N_{\text{b}}^2)$ ,

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} \mathbf{x}_n \\ \mathbf{y}_n \end{pmatrix} = \omega_n \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \mathbf{x}_n \\ \mathbf{y}_n \end{pmatrix}, \quad (18)$$

so that the computation of the entire spectrum is prohibitively expensive. Here the large matrix blocks of size  $N_{\text{ov}} \times N_{\text{ov}}$  take a form

$$A = \Delta\epsilon + V - \bar{W}, \quad B = V - \tilde{W},$$



where the diagonal “energy” matrix is defined by

$$\Delta\varepsilon = [\Delta\varepsilon_{ia,jb}] \in \mathbb{R}^{N_{ov} \times N_{ov}}: \quad \Delta\varepsilon_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab},$$

while the matrices  $\tilde{W} = [\tilde{w}_{ia,jb}]$  and  $\tilde{W} = [\tilde{w}_{ia,jb}]$  are determined by permutation of the so-called static screened interaction matrix  $W = [w_{ia,jb}]$ , via  $\tilde{W}_{ia,jb} = w_{ji,ab}$ , and  $[\tilde{w}_{ia,jb}] = [w_{ib,aq}]$ , respectively. In turn, the fourth order tensor  $\mathbf{W} = [w_{ia,jb}]$  is constructed by certain linear transformations of the tensor  $\mathbf{V} = [v_{ia,jb}]$ , see ref. 11 and 61.

A number of numerical methods for structured eigenvalue problems have been discussed in the literature.<sup>4,15,51,55</sup>

The tensor approach to the solution of the partial BSE eigenvalue problem for eqn (18) proposed in ref. 6 suggests to compute the reduced basis set by solving the simplified eigenvalue problem via the low-rank plus diagonal approximation to the matrix blocks  $A$  and  $B$ , and then solve spectral problem for the subsequent Galerkin projection of the initial system (18) to this reduced basis. This procedure relies entirely on multiplication of the simplified BSE matrix with vectors.

It was demonstrated on the examples of moderate size molecules<sup>6</sup> that a small reduced basis set, obtained by separable approximation with the rank parameters of about several tens, allows to reveal several lowest excitation energies and respective excited states with the accuracy about 0.1–0.02 eV.

Fig. 14 illustrates the BSE energy spectrum of the  $\text{NH}_3$  molecule (based on HF calculations with cc-pDVZ-48 GTO

basis) for the lowest  $N_{\text{red}} = 30$  eigenvalues vs. the rank truncation parameter  $\varepsilon = 0.6$  and  $0.1$ , where the ranks of  $V$  and the BSE matrix block  $W$  are 4, 5 and 28, 30, respectively. For the choice  $\varepsilon = 0.6$  and  $\varepsilon = 0.1$ , the error in the 1st (lowest) eigenvalue for the solution of the problem in reduced basis is about 0.11 eV and 0.025 eV, correspondingly. The CPU time in the laptop Matlab implementation of each example is about 5 s.

## 5 Tensor approach to simulation of large crystalline clusters

In this section, we briefly discuss the generalization of the tensor-based Hartree–Fock solver to the case of large lattice structured and periodic systems<sup>47,48</sup> arising in the numerical modeling of crystalline, metallic and polymer type compounds.

### 5.1 Fast tensor calculation of a lattice sum of interaction potentials

One of the challenges in the numerical treatment of large molecular systems is the summation of long-range potentials allocated on large 3D lattices.<sup>23,52,53,58</sup> The conventional Ewald summation techniques based on a separate evaluation of contributions from the short- and long-range parts of the interaction potential exhibit  $O(L^3 \log L)$  complexity scaling for a cubic  $L \times L \times L$  3D lattice.

In the contrary, the main idea of the novel tensor summation method introduced in ref. 47 and 49 suggests to benefit from the low-rank tensor decomposition of the generating kernel approximated on the fine  $n \times n \times n$  representation grid in the 3D computational box  $\Omega_L$ . This allows to completely decouple the 3D sum into the three independent 1D summations, thus reducing drastically the numerical expenses. The resultant potential sum, which now requires only  $O(n)$  storage and  $O(nL)$  computational demands, is represented by a few assembled canonical/Tucker  $n$ -vectors of complicated shape (see Fig. 15), where  $n$  is the univariate grid size for a cubic 3D lattice.

For ease of exposition, we consider the electrostatic nuclear potential of a single hydrogen atom,  $V_c(x) = \frac{Z}{\|x\|}$ . Define the scaled unit cell,  $\Omega_0 = [0, b]^3$ , of size  $b \times b \times b$  and consider a sum of a finite number of interaction potentials in a symmetric computational box

$$\Omega_L = B \times B \times B, \quad \text{with} \quad B = b \left[ -\frac{L}{2}, \frac{L}{2} \right], \quad L = 2L_0 \in \mathbb{N},$$

consisting of a union of  $L \times L \times L$  unit cells  $\Omega_{\mathbf{k}}$ , obtained from  $\Omega_0$  by a shift along the lattice vector  $b\mathbf{k}$ , where  $\mathbf{k} = (k_1, k_2, k_3) \in \mathbb{Z}^3$ , such that  $k_\ell \in \mathcal{K} := \mathcal{K}_- \cup \mathcal{K}_+$  for  $\ell = 1, 2, 3$  with  $\mathcal{K}_- := \left\{ -1, \dots, -\frac{L}{2} \right\}$  and  $\mathcal{K}_+ := \left\{ 0, 1, \dots, \frac{L}{2} - 1 \right\}$ . Hence, we have  $\Omega_L = \cup_{k_1, k_2, k_3 \in \mathcal{K}} \Omega_{\mathbf{k}} \in \mathbb{R}^3$ .

Recall that  $b = nh$ , where  $h > 0$  is the fine mesh size that is the same for all spatial variables, and  $n$  is the number of grid points for each variable. We also define the accompanying domain  $\tilde{\Omega}_L$  obtained by scaling of  $\Omega_L$  with the factor of 2,

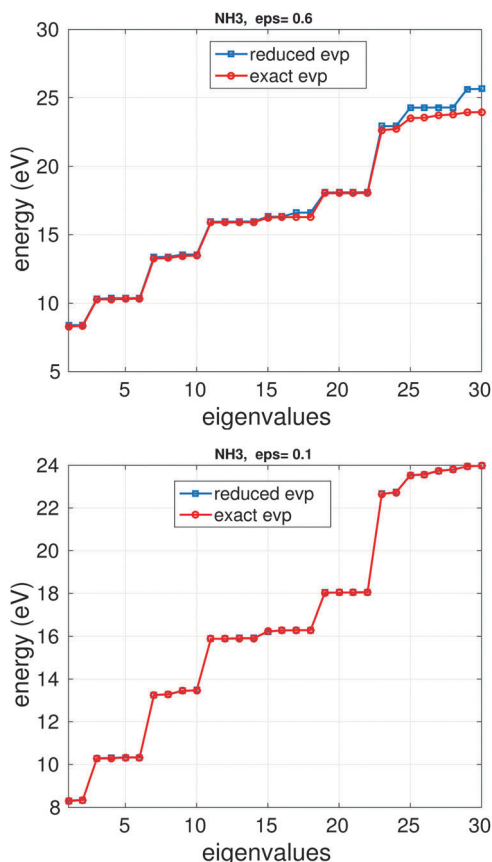


Fig. 14 Comparison of  $m_0 = 30$  lower eigenvalues for the reduced and exact BSE systems for  $\text{NH}_3$  molecule:  $\varepsilon = 0.6$ , top;  $\varepsilon = 0.1$ , bottom.





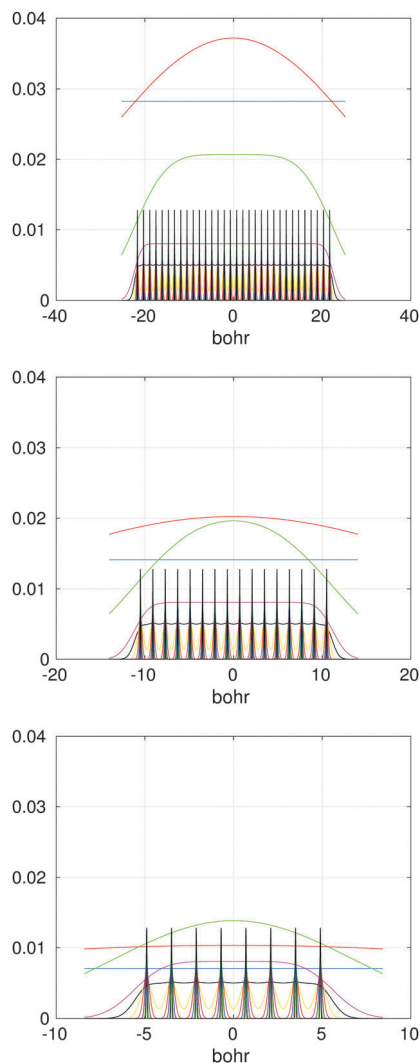


Fig. 15 Assembled x-, y- and z-axis canonical vectors for a cluster of  $32 \times 16 \times 8$  hydrogen atoms.

$\tilde{\Omega}_L = 2\Omega_L$ , and, similar to (31), introduce the respective rank- $R$  reference (master) tensor

$$\tilde{\mathbf{P}} = \sum_{q=1}^R \tilde{\mathbf{p}}_q^{(1)} \otimes \tilde{\mathbf{p}}_q^{(2)} \otimes \tilde{\mathbf{p}}_q^{(3)} \in \mathbb{R}^{2n \times 2n \times 2n}, \quad (19)$$

approximating  $\frac{1}{\|x\|}$  in  $\tilde{\Omega}_L$  on a  $2n \times 2n \times 2n$  representation grid with mesh size  $h$ .

Let us consider a sum of single Coulomb potentials on a  $L \times L \times L$  lattice,

$$V_{c_L}(x) = \sum_{k_1, k_2, k_3 \in \mathcal{K}} \frac{Z_1}{\|x - a_1(k_1, k_2, k_3)\|}, \quad x \in \Omega_L \in \mathbb{R}^3. \quad (20)$$

The assembled tensor approach applies to the potentials defined on  $n \times n \times n$  3D Cartesian grid. It reduces the sum over a rectangular 3D lattice,  $\mathbf{P}_{c_L} \in \mathbb{R}^{n \times n \times n}$ ,

$$\mathbf{P}_{c_L} = \sum_{\mathbf{k} \in \mathcal{K}^3} \mathcal{W}_{\nu(\mathbf{k})} \tilde{\mathbf{P}} = \sum_{k_1, k_2, k_3 \in \mathcal{K}} \sum_{q=1}^R \mathcal{W}_{\nu(\mathbf{k})} \left( \tilde{\mathbf{p}}_q^{(1)} \otimes \tilde{\mathbf{p}}_q^{(2)} \otimes \tilde{\mathbf{p}}_q^{(3)} \right),$$

to the summation of directional vectors for the canonical decomposition of shifted single Newton kernels,<sup>47</sup>

$$\mathbf{P}_{c_L} = \sum_{q=1}^R \left( \sum_{k_1 \in \mathcal{K}} \mathcal{W}_{(k_1)} \tilde{\mathbf{p}}_q^{(1)} \right) \otimes \left( \sum_{k_2 \in \mathcal{K}} \mathcal{W}_{(k_2)} \tilde{\mathbf{p}}_q^{(2)} \right) \otimes \left( \sum_{k_3 \in \mathcal{K}} \mathcal{W}_{(k_3)} \tilde{\mathbf{p}}_q^{(3)} \right), \quad (21)$$

where  $\mathcal{W}_{\nu(\mathbf{k})} = \mathcal{W}_{k_1} \otimes \mathcal{W}_{k_2} \otimes \mathcal{W}_{k_3}$  is the shift-and-windowing (onto  $\Omega_L$ ) separable transform along the  $\mathbf{k}$ -grid. Remarkably that the rank of the resulting sum is the same as for the  $R$ -term canonical reference tensor (19) representing the single Newton kernel.

The numerical cost and storage size are bounded by  $O(RLn)$  and  $O(Rn)$ , respectively, where  $n = n_0L$ , and  $n_0$  is the grid size in the unit cell.

Fig. 15 presents the assembled canonical vectors for the cluster of  $32 \times 16 \times 8$  Hydrogen nuclei, with a distance 1.4 bohr between nuclei. The resulting potential is shown in Fig. 16. Size of the computational box is  $50.4 \times 28.0 \times 16.8$  bohr<sup>3</sup>. The corresponding 3D Cartesian grid size is  $9216 \times 5120 \times 3072$ . Here the empty interval between the lattice and the boundary of the computational box equals to 3 bohr.

The next table represents CPU times for the lattice summation of the Newton kernels over  $L \times L \times L$  cubic box, with very fine  $n \times n \times n$  representation grid.

$L^3$	4096	32 768	262 144	2 097 152
Time (s)	1.8	0.8	3.1	15.8
3D grid size, $n^3$	5632 <sup>3</sup>	9728 <sup>3</sup>	17 920 <sup>3</sup>	34 304 <sup>3</sup>

The summation method in the canonical format was extended to the Tucker tensors which allows the principal generalization of this techniques to the case of rather complicated lattices with defects (Theorem 3.2, ref. 49), so that the resulting sum takes a form

$$\mathbf{T}_{c_L} = \sum_{m_1=1}^{r_1} \sum_{m_2=1}^{r_2} \sum_{m_3=1}^{r_3} b_{m_1, m_2, m_3} \left( \sum_{k_1 \in \mathcal{K}} \mathcal{W}_{(k_1)} \tilde{\mathbf{t}}_{m_1}^{(1)} \right) \otimes \left( \sum_{k_2 \in \mathcal{K}} \mathcal{W}_{(k_2)} \tilde{\mathbf{t}}_{m_2}^{(2)} \right) \otimes \left( \sum_{k_3 \in \mathcal{K}} \mathcal{W}_{(k_3)} \tilde{\mathbf{t}}_{m_3}^{(3)} \right)$$

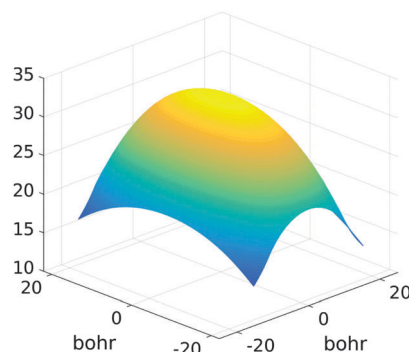


Fig. 16 Assembled canonical sum of the Newton potentials for the cluster of  $32 \times 16 \times 8$  hydrogen atoms.



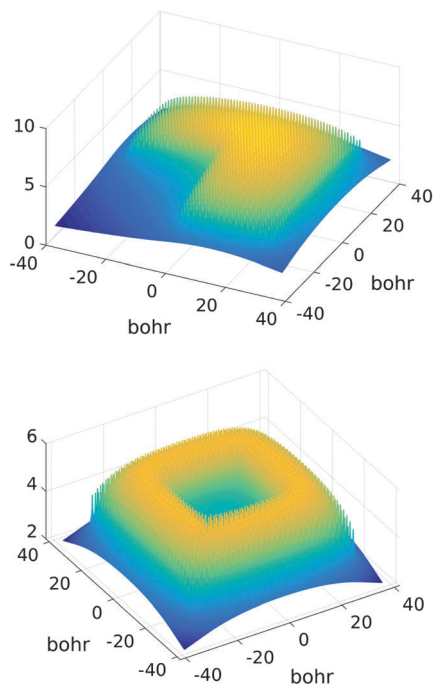


Fig. 17 Assembled canonical sum of the Coulomb potentials on the L-shaped (top) and O-shaped (bottom) sub-lattices of a  $32 \times 32 \times 1$  lattice.

where  $\tilde{\mathbf{t}}_{m_\ell}^{(\ell)}$ ,  $\ell = 1, 2, 3$ , represents the Tucker vectors of the rank- $r$  master tensor approximating the Newton kernel  $\frac{1}{\|x\|}$  in  $\tilde{\Omega}_L$  on a  $2n \times 2n \times 2n$  representation grid.

In the case of defected lattices the increasing rank of the final sum of Tucker tensors can be reduced by the stable ALS based algorithms applicable to the Tucker tensors (see ref. 48). The particular examples of the lattice geometries suitable for our approach are presented in Fig. 17, see ref. 48 for more detailed discussion.

For the reduced Hartree–Fock equation, where the Fock operator is confined to the core Hamiltonian, the tensor-structured block-circulant representation of the Fock matrix was introduced<sup>48</sup> that allows the special low-rank approximation of the matrix blocks. This opens the way for the numerical treatment of large eigenvalue problems with structured matrices arising in the solution of the Hartree–Fock equation for large crystalline systems with defects.

## 5.2 Interaction energy of the Coulomb lattice sums

Given the nuclear charges  $\{Z_{\mathbf{k}}\}$ , centered at points  $\mathbf{x}_{\mathbf{k}}$ ,  $\mathbf{k} \in \mathcal{K}^3$ , located on a finite  $L \times L \times L$  lattice  $\mathcal{L}_L = \{\mathbf{x}_{\mathbf{k}}\}$  with the step-size  $b$ , the interaction energy of the total electrostatic potential of these charges is defined by the lattice sum

$$E_L = \frac{1}{2} \sum_{\mathbf{k}, \mathbf{j} \in \mathcal{K}, \mathbf{k} \neq \mathbf{j}} \frac{Z_{\mathbf{k}} Z_{\mathbf{j}}}{\|\mathbf{x}_{\mathbf{j}} - \mathbf{x}_{\mathbf{k}}\|}, \quad \text{i.e. for } \|\mathbf{x}_{\mathbf{j}} - \mathbf{x}_{\mathbf{k}}\| \geq b. \quad (22)$$

The fast and accurate computation of electrostatic interaction energy (as well as the related forces and stresses) is one of the difficult tasks in computer modeling of macromolecular structures like finite crystals, and biological systems.

The tensor summation scheme (21) can be directly applied to this computational problem. In what follows we show that (22) can be treated as a particular case of the previous scheme served for calculation of (20) on a fine spacial grid. For this discussion, we assume that all charges are equal,  $Z_{\mathbf{k}} = Z$ .

First, notice that the rank- $R$  reference tensor  $h^{-3} \tilde{\mathbf{P}}$  defined in (19) approximates with high accuracy  $O(h^2)$  the (and its shifted version) Coulomb potential  $\frac{1}{\|x\|}$  in  $\tilde{\Omega}_L$  (for  $\|x\| \geq b$  that is required for the energy expression) on the fine  $2n \times 2n \times 2n$  representation grid with mesh size  $h$ . Likewise, the tensor  $h^{-3} \mathbf{P}_{c_L}$  approximates the potential sum  $V_{c_L}(x)$  on the same fine representation grid including the lattice points  $\mathbf{x}_{\mathbf{k}}$ .

We propose to evaluate the energy expression (22) by using tensor sums as in (21), but now applied to a small sub-tensor of the rank- $R$  canonical reference tensor  $\tilde{\mathbf{P}}$ , that is  $\tilde{\mathbf{P}}_L := [\tilde{\mathbf{P}}_{|\mathbf{x}_{\mathbf{k}}}] \in \mathbb{R}^{2L \times 2L \times 2L}$ , obtained by tracing of  $\tilde{\mathbf{P}}$  at the accompanying lattice of the double size  $2L \times 2L \times 2L$ , i.e.  $\tilde{\mathcal{L}}_L = \{\mathbf{x}_{\mathbf{k}}\} \cup \{\mathbf{x}_{\mathbf{k}'}\} \in \tilde{\Omega}_L$ . Here  $\tilde{\mathbf{P}}_{|\mathbf{x}_{\mathbf{k}}}$  denotes the tensor entry corresponding to the  $\mathbf{k}$ -th lattice point designating the atomic center  $\mathbf{x}_{\mathbf{k}}$ . We are interested in the computation of the rank- $R$  tensor  $\hat{\mathbf{P}}_{c_L} = [\mathbf{P}_{c_L|\mathbf{x}_{\mathbf{k}}}]_{\mathbf{k} \in \mathcal{K}} \in \mathbb{R}^{L \times L \times L}$ , where  $\mathbf{P}_{c_L|\mathbf{x}_{\mathbf{k}}}$  denotes the tensor entry corresponding to the  $\mathbf{k}$ -th lattice point on  $\mathcal{L}_L$ . The tensor  $\hat{\mathbf{P}}_{c_L}$  can be computed at the expense  $O(L^2)$  by

$$\hat{\mathbf{P}}_{c_L} = \sum_{q=1}^R \left( \sum_{\mathbf{k}_1 \in \mathcal{K}} \mathcal{W}_{(\mathbf{k}_1)} \tilde{\mathbf{p}}_{L,q}^{(1)} \otimes \sum_{\mathbf{k}_2 \in \mathcal{K}} \mathcal{W}_{(\mathbf{k}_2)} \tilde{\mathbf{p}}_{L,q}^{(2)} \otimes \sum_{\mathbf{k}_3 \in \mathcal{K}} \mathcal{W}_{(\mathbf{k}_3)} \tilde{\mathbf{p}}_{L,q}^{(3)} \right).$$

This leads to the representation of the energy sum (22) with accuracy  $O(h^2)$  in a form

$$E_{L,T} = \frac{Z^2 h^{-3}}{2} \left( \langle \hat{\mathbf{P}}_{c_L}, \mathbf{1} \rangle - \sum_{\mathbf{k} \in \mathcal{K}} \mathbf{P}_{|\mathbf{x}_{\mathbf{k}}=0} \right),$$

where the first term in brackets represents the full canonical tensor lattice sum restricted to the  $\mathbf{k}$ -grid composing the lattice  $\mathcal{L}_L$ , while the second term introduces the correction at singular points  $\mathbf{x}_{\mathbf{j}} - \mathbf{x}_{\mathbf{k}} = 0$ . Here  $\mathbf{1} \in \mathbb{R}^{L \times L \times L}$  is the all-ones tensor. By using the rank-1 tensor  $\mathbf{P}_{0L} = \mathbf{P}_{|\mathbf{x}_{\mathbf{k}}=0} \mathbf{1}$ , the correction term can be represented by a simple tensor operation

$$\sum_{\mathbf{k} \in \mathcal{K}} \mathbf{P}_{|\mathbf{x}_{\mathbf{k}}=0} = \langle \mathbf{P}_{0L}, \mathbf{1} \rangle.$$

Finally, the interaction energy  $E_L$  allows the approximate representation

$$E_L \approx E_{L,T} = \frac{Z^2 h^{-3}}{2} \left( \langle \hat{\mathbf{P}}_{c_L}, \mathbf{1} \rangle - \langle \mathbf{P}_{0L}, \mathbf{1} \rangle \right), \quad (23)$$

that can be implemented in  $O(L^2) \ll L^3 \log L$  complexity by tensor operations with the rank- $R$  canonical tensors in  $\mathbb{R}^{L \times L \times L}$ .

Table 4 illustrates the performance of the algorithm described above. We compare the exact value computed by (22) with the approximate tensor representation in (23) computed on the fine representation grid with  $n = n_0 L$ ,  $n_0 = 128$ . We consider the lattice systems composed of Hydrogen atoms with interatomic distance 2.0 bohr. The geometric size of the



**Table 4** Comparison of times for the full ( $T_{\text{full}}$ ) and tensor-based ( $T_{\text{tens}}$ ) calculation of the interaction energy sum for the lattice electrostatic potentials. Matlab on a 2 Intel Xeon Hexa-Core/2677

$L^3$	$T_{\text{full}}, (O(L^3))$	$T_{\text{tens}}$	$E_{L,T}$	abs. err.
$24^3$	37, (6.1)	1.2	$3.7 \times 10^6$	$2 \times 10^{-8}$
$32^3$	250, (15.8)	1.5	$1.5 \times 10^7$	$1.5 \times 10^{-9}$
$48^3$	3374, (58.8)	2.8	$1.12 \times 10^8$	0
$64^3$	—	5.7	$5.0 \times 10^8$	—
$128^3$	—	13.5	$1.6 \times 10^{10}$	—
$256^3$	—	68.2	$5.2 \times 10^{11}$	—

largest 3D lattice with  $256^3$  potentials is of the order of  $(2.0 \cdot 256 + 6.0)^3$  bohr<sup>3</sup>,§ which is more than  $20^3$  cubic nanometers. Here  $T_{\text{full}}$  means CPU time (sec) for direct summation of by (22) of the order of  $L^6$ , where the number in brackets shows  $\sqrt{T_{\text{full}}}$  of the order of  $O(L^3)$ .

The presented approach for fast calculation of the interaction energy can be extended to the case of non-uniform rectangular lattices and, under certain assumptions, to the case of non-equal nuclear charges  $Z_k$ . Moreover, it applies to many other types of spherically symmetric interaction potentials, for example, to shielded Coulomb interaction or van der Waals attraction sums corresponding to the distance function  $||x||^{-2}$  and  $||x||^{-6}$ , respectively.

## 6 Conclusions

The goal of this paper is to attract interest of the specialists in computational quantum chemistry to recent results and open questions of the grid-based tensor approach in electronic structure calculations. Here we focus mostly on the description of main mathematical and algorithmic aspects of the tensor decomposition schemes and demonstrate their benefits in some applications.

The scope of applications which can be regarded as consistent, ranges from the Hartree–Fock energy estimates for moderate size molecules, including calculation of TEI tensor with incorporated algebraic density fitting, to calculation of the excitation energies for molecules, and up to a unique superfast method for calculating the lattice potential sums and the interaction energy of long range potentials on a lattice in a finite volume. Tensor approach allows to treat above problems using moderate computational facilities. All numerics given in the paper presents implementations in Matlab.

The described numerical tools are not restricted to the applications presented here, but can be applied to various hard computational problems in (post) Hartree–Fock calculations related to accurate evaluation of multidimensional integrals, and efficient storage and manipulations with large multivariate data arrays.

The presented method for summation of long-range potentials with sub-linear computational cost does not have analogues in what is used so far in computational quantum chemistry and can have a good future. For example, it can be useful in modeling of

large finite molecular structures like nanostructures or quantum dots, where the periodic approach may be inconsistent. Calculating a sum of several millions of lattice potentials on fine 3D grids takes only several seconds in Matlab on a laptop.<sup>47</sup> This method gives also the unique possibility to present the summed 3D lattice potential in the whole computational region with very high accuracy. Integration and differentiation of this 3D potential can be easily performed on representation grid due to 1D computational costs.

Tensor approach is now being evolved for modeling the electronic structure of finite crystalline-type molecular systems. We hope that tensor numerical methods will have a good future in solving challenging multidimensional problems of computational quantum chemistry.

## Appendix

### Canonical-to-Tucker transform

The Canonical-to-Tucker tensor transform combined with the Tucker-to-Canonical scheme introduced in ref. 39 usually applies for the rank reduction of the function related canonical tensors with the large initial rank. Here we sketch Algorithm Canonical-to-Tucker which includes the following basic steps:

Input data: side matrices  $U^{(\ell)} = [\mathbf{u}_1^{(\ell)} \dots \mathbf{u}_R^{(\ell)}] \in \mathbb{R}^{n_\ell \times R}$ ,  $\ell = 1, 2, 3$ , composed of vectors  $\mathbf{u}_k^{(\ell)} \in \mathbb{R}^{n_\ell}$ ,  $k = 1, \dots, R$ , see (1); maximal Tucker-rank parameter  $r$ ; maximal number of the alternating least square (ALS) iterations  $m_{\text{max}}$  (usually a small number).

(A) Compute the singular value decomposition (SVD) of side matrices

$$U^{(\ell)} = Z^{(\ell)} S^{(\ell)} V^{(\ell)}, \quad \ell = 1, 2, 3.$$

Discard the singular vectors in  $Z^{(\ell)}$  and the respective singular values up to given rank threshold, yielding the small orthogonal matrices  $Z_{r_\ell}^{(\ell)} \in \mathbb{R}^{n_\ell \times r_\ell}$ ,  $\ell = 1, 2, 3$ .

(B) Project side matrices  $U^{(\ell)}$  onto the orthogonal basis set defined by  $Z_{r_\ell}^{(\ell)}$

$$U^{(\ell)} \mapsto \tilde{U}^{(\ell)} = \left( Z_{r_\ell}^{(\ell)} \right)^T U^{(\ell)}, \quad \tilde{U}^{(\ell)} \in \mathbb{R}^{r_\ell \times R}, \quad \ell = 1, 2, 3. \quad (24)$$

(C) (Find dominating subspaces). Implement the following ALS iteration  $m_{\text{max}}$  times at most. For  $\ell = 1, 2, 3$  implement the following ALS iteration  $m_{\text{max}}$  times at most.

(D) Start ALS iteration for  $\ell = 1, 2, 3$ :

— For  $\ell = 1$ : construct partially projected image of the full tensor,

$$\mathbf{U} \mapsto \tilde{\mathbf{U}}_1 = \sum_{k=1}^R c_k \mathbf{u}_k^{(1)} \otimes \tilde{\mathbf{u}}_k^{(2)} \otimes \tilde{\mathbf{u}}_k^{(3)}, \quad c_k \in \mathbb{R}. \quad (25)$$

Here  $\mathbf{u}_k^{(1)} \in \mathbb{R}^{n_1}$  are in physical space for mode  $\ell = 1$ , while  $\tilde{\mathbf{u}}_k^{(2)} \in \mathbb{R}^{r_2}$  and  $\tilde{\mathbf{u}}_k^{(3)} \in \mathbb{R}^{r_3}$ , the column vectors of  $\tilde{U}^{(2)}$  and  $\tilde{U}^{(3)}$ , respectively, belong to the coefficients space by means of projection.

— Reshape the tensor  $\tilde{\mathbf{U}}_1 \in \mathbb{R}^{n_1 \times r_2 \times r_3}$  into a matrix  $M_{U_1} \in \mathbb{R}^{n_1 \times (r_2 r_3)}$ , representing the span of the optimized subset of mode-1 columns in partially projected tensor  $\tilde{\mathbf{U}}_1$ . Compute the SVD of the matrix  $M_{U_1}$ :

$$M_{U_1} = Z^{(1)} S^{(1)} V^{(1)},$$

§ Here 6 bohr is the chosen dummy distance.



and truncate the set of singular vectors in  $Z^{(1)} \mapsto \tilde{Z}^{(1)} \in \mathbb{R}^{n_1 \times r_1}$ , according to the restriction on the mode-1 Tucker rank,  $r_1$ .

— Update the current approximation to the mode-1 dominating subspace  $Z_{r_1}^{(1)} \mapsto \tilde{Z}^{(1)}$ .

— Implement the single loop of ALS iteration for mode  $\ell = 2$  and for  $\ell = 3$ .

— End of the single ALS iteration step.

— Repeat the complete ALS iteration  $m_{\max}$  times to obtain the optimized Tucker orthogonal side matrices  $\tilde{Z}^{(1)}, \tilde{Z}^{(2)}, \tilde{Z}^{(3)}$ , and final projected image  $\tilde{\mathbf{U}}_3$ .

(E) Project the final iterated tensor  $\tilde{\mathbf{U}}_3$  in (25) using the resultant basis set in  $\tilde{Z}^{(3)}$  to obtain the core tensor,  $\beta \in \mathbb{R}^{r_1 \times r_2 \times r_3}$ .

Output data: the Tucker core tensor  $\beta$  and the Tucker orthogonal side matrices  $\tilde{Z}^{(\ell)}$ ,  $\ell = 1, 2, 3$ .

The Canonical-to-Tucker algorithm can be easily modified to the  $\varepsilon$ -truncation stopping criteria. Notice that in the case of equal Tucker ranks,  $r_\ell = r$ , maximal canonical rank of the core tensor  $\beta$  does not exceed  $r^2$ , see ref. 39, which completes the Tucker-to-Canonical part of the total algorithm.¶

### The Hartree-Fock equation in AO basis set

The  $2N$ -electrons Hartree-Fock equation for pairwise  $L^2$ -orthogonal electronic orbitals,  $\psi_i: \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $\psi_i \in H^1(\mathbb{R}^3)$ , reads as

$$\mathcal{F}\psi_i(x) = \lambda_i\psi_i(x), \quad \int_{\mathbb{R}^3} \psi_i\psi_j dx = \delta_{ij}, \quad i, j = 1, \dots, N_{\text{orb}} \quad (26)$$

where the nonlinear Fock operator  $\mathcal{F}$  is given by

$$\mathcal{F} := -\frac{1}{2}\Delta + V_c(\cdot) + V_H(\cdot) + \mathcal{K}.$$

Here the nuclear potential takes the form  $V_c(x) = -\sum_{\nu=1}^M \frac{Z_\nu}{\|x - a_\nu\|}$ ,  $Z_\nu > 0$ ,  $a_\nu \in \mathbb{R}^3$ , while the Hartree potential  $V_H(x)$  and the nonlocal exchange operator  $\mathcal{K}$  read as

$$V_H(x) := \rho \star \frac{1}{\|\cdot\|} = \int_{\mathbb{R}^3} \frac{\rho(y)}{\|x - y\|} dy, \quad x \in \mathbb{R}^3, \quad (27)$$

and

$$\begin{aligned} (\mathcal{K}\psi)(x) &:= -\frac{1}{2} \sum_{i=1}^{N_{\text{orb}}} \left( \psi\psi_i \star \frac{1}{\|\cdot\|} \right) \psi_i(x) \\ &= -\frac{1}{2} \int_{\mathbb{R}^3} \frac{\tau(x, y)}{\|x - y\|} \psi(y) dy, \end{aligned} \quad (28)$$

respectively. Conventionally, we use the definitions

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

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

Usually, the Hartree-Fock equation is approximated by the standard Galerkin projection of the initial problem (26) by using the physically justified reduced basis sets (say, GTO type

orbitals). For a given finite Galerkin basis set  $\{g_\mu\}_{1 \leq \mu \leq N_b}$ ,  $g_\mu \in H^1(\mathbb{R}^3)$ , the occupied molecular orbitals  $\psi_i$  are represented (approximately) as  $\psi_i = \sum_{\mu=1}^{N_b} C_{\mu i} g_\mu$ ,  $i = 1, \dots, N_{\text{orb}}$ . To derive an equation for the unknown coefficients matrix  $C = \{C_{\mu i}\} \in \mathbb{R}^{N_b \times N_{\text{orb}}}$ , first, we introduce the mass (overlap) matrix  $S = \{S_{\mu\nu}\}_{1 \leq \mu, \nu \leq N_b}$ , given by  $S_{\mu\nu} = \int_{\mathbb{R}^3} g_\mu g_\nu dx$ , and the stiffness matrix  $H = \{h_{\mu\nu}\}$  of the core Hamiltonian  $\mathcal{H} = -\frac{1}{2}\Delta + V_c$  (the single-electron integrals),

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

The core Hamiltonian matrix  $H$  can be precomputed in  $O(N_b^2)$  operations via grid-based approach.

Given the finite basis set  $\{g_\mu\}_{1 \leq \mu \leq N_b}$ ,  $g_\mu \in H^1(\mathbb{R}^3)$ , the associated fourth order two-electron integrals (TEI) tensor,  $\mathbf{B} = [b_{\mu\nu\lambda\sigma}]$ , is defined entrywise by (10), where  $\mu, \nu, \lambda, \sigma \in \{1, \dots, N_b\} =: \mathcal{I}_b$ . In computational quantum chemistry the nonlinear terms representing the Galerkin approximation to the Hartree and exchange operators are calculated traditionally by using the low-rank Cholesky decomposition of a matrix associated with the TEI tensor  $\mathbf{B} = [b_{\mu\nu\kappa\lambda}]$  defined in (10), that initially has the computational and storage complexity of order  $O(N_b^4)$ .

Introducing the  $N_b \times N_b$  matrices  $J(D)$  and  $K(D)$ ,

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

where  $D = 2CC^T \in \mathbb{R}^{N_b \times N_b}$  is the rank- $N_{\text{orb}}$  symmetric density matrix, one then represents the complete Fock matrix  $F$  by

$$F(D) = H + J(D) + K(D).$$

The resultant Galerkin system of nonlinear equations for the coefficients matrix  $C \in \mathbb{R}^{N_b \times N_{\text{orb}}}$ , and the respective eigenvalues  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{N_{\text{orb}}})$ , reads as

$$F(D)C = SCA, \quad C^T SC = I_N,$$

where the second equation represents the orthogonality constraints  $\int_{\mathbb{R}^3} \psi_i \psi_j dx = \delta_{ij}$ , and  $I_N$  denotes the  $N_b \times N_b$  identity matrix, and is usually solved by self-consistent field (SCF) DIIS iteration.<sup>60</sup>

### Tensor approximation to the Newton kernel in 3D

Methods of separable approximation to multivariate spherically symmetric functions by using the Gaussian sums have been addressed in the chemical and mathematical literature since ref. 9, 10 and 65, respectively.

In this section, we discuss for the readers convenience the grid-based method for the low-rank canonical and Tucker tensor representations of a spherically symmetric functions  $p(\|x\|)$ ,  $x \in \mathbb{R}^d$  in the particular case of the 3D Newton kernel  $p(\|x\|) = \frac{1}{\|x\|}$ ,  $x \in \mathbb{R}^3$  by its projection onto the set of piecewise constant basis functions, see ref. 7 for more details.

¶ Further optimization of the canonical rank in the small-size core tensor  $\beta$  can be implemented by applying the ALS iterative scheme in the canonical format, see e.g. ref. 50.





In the computational domain  $\Omega = [-b/2, b/2]^3$ , let us introduce the uniform  $n \times n \times n$  rectangular Cartesian grid  $\Omega_n$  with the mesh size  $h = b/n$  (usually,  $n = 2k$ ). Let  $\{\psi_{\mathbf{i}}\}$  be a set of tensor-product piecewise constant basis functions,  $\psi_{\mathbf{i}}(\mathbf{x}) = \prod_{\ell=1}^d \psi_{i_\ell}^{(\ell)}(x_\ell)$ , for the 3-tuple index  $\mathbf{i} = (i_1, i_2, i_3)$ ,  $i_\ell \in \{1, \dots, n\}$ ,  $\ell = 1, 2, 3$ . The kernel  $p(\|\mathbf{x}\|)$  can be discretized by its projection onto the basis set  $\{\psi_{\mathbf{i}}\}$  in the form of a third order tensor of size  $n \times n \times n$ , defined entrywise as

$$\mathbf{P} := [p_{\mathbf{i}}] \in \mathbb{R}^{n \times n \times n}, \quad p_{\mathbf{i}} = \int_{\mathbb{R}^3} \psi_{\mathbf{i}}(\mathbf{x}) p(\|\mathbf{x}\|) d\mathbf{x}. \quad (29)$$

Given  $M$ , the low-rank canonical decomposition of the 3rd order tensor  $\mathbf{P}$  is based on using exponentially convergent sinc-quadratures for approximation of the Laplace-Gauss transform to the analytic function  $p(z) = 1/z$  as follows,

$$\begin{aligned} 1/\|\mathbf{x}\| &= \frac{2}{\sqrt{\pi}} \int_{\mathbb{R}_+} e^{-t^2 \|\mathbf{x}\|^2} dt \\ &\approx \sum_{k=-M}^M a_k e^{-t_k^2 \|\mathbf{x}\|^2} = \sum_{k=-M}^M a_k \prod_{\ell=1}^3 e^{-t_k^2 x_\ell^2}, \end{aligned} \quad (30)$$

where the quadrature points and weights are given by

$$t_k = k h_M, \quad a_k = \frac{2}{\sqrt{\pi}} h_M, \quad h_M = C_0 \log(M)/M, \quad C_0 > 0.$$

Under the assumption  $0 < a \leq \|\mathbf{x}\| < \infty$ ,  $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ , this quadrature can be proven to provide the exponential convergence rate in  $M$  for a class of analytic functions including  $p(z) = 1/z$ , see ref. 10, 25 and 65,

$$\left| 1/\|\mathbf{x}\| - \sum_{k=-M}^M a_k e^{-t_k^2 \|\mathbf{x}\|^2} \right| \leq \frac{C}{a} e^{-\beta \sqrt{M}}, \quad \text{with some } C, \beta > 0.$$

Combining (29) and (30), and taking into account the separability of the Gaussian basis functions, we arrive at the low-rank approximation to each entry of the tensor  $\mathbf{P}$ ,

$$\begin{aligned} p_{\mathbf{i}} &\approx \sum_{k=-M}^M a_k \int_{\mathbb{R}^3} \psi_{\mathbf{i}}(\mathbf{x}) e^{-t_k^2 \|\mathbf{x}\|^2} d\mathbf{x} \\ &= \sum_{k=-M}^M a_k \prod_{\ell=1}^3 \int_{\mathbb{R}} \psi_{i_\ell}^{(\ell)}(x_\ell) e^{-t_k^2 x_\ell^2} dx_\ell. \end{aligned}$$

Define the vector (recall that  $a_k > 0$ )

$$\mathbf{p}_k^{(\ell)} = a_k^{1/3} \mathbf{b}^{(\ell)}(t_k) \equiv a_k^{1/3} \left[ b_{i_\ell}^{(\ell)}(t_k) \right]_{i_\ell=1}^n \in \mathbb{R}^n$$

with

$$b_{i_\ell}^{(\ell)}(t_k) = \int_{\mathbb{R}} \psi_{i_\ell}^{(\ell)}(x_\ell) e^{-t_k^2 x_\ell^2} dx_\ell,$$

then the 3rd order tensor  $\mathbf{P}$  can be approximated by the  $R$ -term canonical representation

$$\mathbf{P} \approx \mathbf{P}_R = \sum_{k=-M}^M a_k \overset{3}{\otimes}_{\ell=1} \mathbf{b}^{(\ell)}(t_k) = \sum_{q=1}^R \mathbf{p}_q^{(1)} \otimes \mathbf{p}_q^{(2)} \otimes \mathbf{p}_q^{(3)} \in \mathbb{R}^{n \times n \times n}, \quad (31)$$

where  $R = 2M + 1$ , and the canonical vectors are renumbered by  $k \rightarrow q = k + M + 1$ ,  $\mathbf{p}_q^{(\ell)} = \mathbf{p}_k^{(\ell)} \in \mathbb{R}^n$ ,  $\ell = 1, 2, 3$ . For the given threshold  $\varepsilon > 0$ ,  $M = O(|\log \varepsilon|^2)$  is chosen as the minimal number, such that in the max-norm we have

$$\|\mathbf{P} - \mathbf{P}_R\| \leq \varepsilon \|\mathbf{P}\|.$$

The symmetric canonical tensor  $\mathbf{P}_R$  in (31) approximates the discretized 3D symmetric kernel function  $p(\|\mathbf{x}\|) = 1/\|\mathbf{x}\|$  ( $\mathbf{x} \in \Omega$ ), centered at the origin, such that  $\mathbf{p}_q^{(1)} = \mathbf{p}_q^{(2)} = \mathbf{p}_q^{(3)}$  ( $q = 1, \dots, R$ ).

## References

- 1 J. Almlöf, in *Direct methods in electronic structure theory*, ed. D. R. Yarkony, Modern Electronic Structure Theory, World Scientific, Singapore, 1995, vol. II, pp. 110–151.
- 2 M. Bachmayr, *Adaptive low-rank wavelet methods and applications to two-electron Schrödinger equations*, PhD dissertation, RWTH Aachen, 2012.
- 3 P. Baudin, J. Marin, I. G. Cuesta and A. M. S. de Meras, Calculation of excitation energies from the CC2 linear response theory using Cholesky decomposition, *J. Chem. Phys.*, 2014, **140**, 104111.
- 4 P. Benner, H. Faßbender and M. Stoll, Solving Large-Scale Quadratic Eigenvalue Problems with Hamiltonian Eigenstructure using a Structure-Preserving Krylov Subspace Method, *Electronic Transactions on Numerical Analysis*, 2008, **29**, 212–229.
- 5 P. Benner and T. Mach, The LR Cholesky Algorithm for Symmetric Hierarchical Matrices, *Linear Algebra and its Applications*, 2013, **439**(4), 1150–1166.
- 6 P. Benner, V. Khoromskaia and B. N. Khoromskij, A reduced basis approach for calculation of the Bethe–Salpeter excitation energies using low-rank tensor factorizations, 2015, Preprint arXiv:1505.02696.
- 7 C. Bertoglio and B. N. Khoromskij, Low-rank quadrature-based tensor approximation of the Galerkin projected Newton/Yukawa kernels, *Comput. Phys. Commun.*, 2012, **183**(4), 904–912.
- 8 F. A. Bischoff, R. J. Harrison and E. F. Valeev, Computing many-body wave functions with guaranteed precision: The first-order Møller-Plesset wave function for the ground state of helium atom, *J. Chem. Phys.*, 2012, **137**, 104103.
- 9 S. F. Boys, G. B. Cook, C. M. Reeves and I. Shavitt, Automatic Fundamental Calculations of Molecular Structure, *Nature*, 1956, **178**, 1207–1209.
- 10 D. Braess, Asymptotics for the Approximation of Wave Functions by Exponential-Sums, *J. Approx. Theory*, 1995, **83**, 93–103.



- 11 M. E. Casida, in *Recent Advances in Density Functional Methods, Part I*, ed. D. P. Chong, World Scientific, Singapore, 1995, p. 155.
- 12 A. Cichocki and Sh. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, Wiley, 2002.
- 13 W. Dahmen, S. Proessdorf and R. Schneider, *Multiscale methods for pseudo-differential equations on manifolds*, Wavelet analysis and its applications, 5, Academic Press, 1995.
- 14 S. Dolgov, *Tensor product methods in numerical simulation of high-dimensional dynamical problems*, PhD thesis, University of Leipzig, 2014, <http://nbn-resolving.de/urn:nbn:de:bsz:15-qucosa-151129>.
- 15 S. Dolgov, B. N. Khoromskij, D. Savostyanov and I. Oseledets, Computation of extreme eigenvalues in higher dimensions using block tensor train format, *Comput. Phys. Commun.*, 2014, **185**(4), 1207–1216.
- 16 L. De Lathauwer, B. De Moor and J. Vandewalle, A multilinear singular value decomposition, *SIAM J. Matrix Anal. Appl.*, 2000, **21**, 1253–1278.
- 17 T. H. Dunning Jr, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, *J. Chem. Phys.*, 1989, **90**, 1007–1023.
- 18 A. Durdek, S. R. Jensen, J. Juselius, P. Wind, T. Flå and L. Frediani, Adaptive order polynomial algorithm in a multi-wavelet representation scheme, *Appl. Num. Math.*, 2014, **92**, 40–53.
- 19 P. P. Ewald, Die Berechnung optische und elektrostatischer Gitterpotentiale, *Ann. Phys.*, 1921, **64**, 253.
- 20 L. Frediani, E. Fossgaard, T. Flå and K. Ruud, Fully adaptive algorithms for multivariate integral equations using the non-standard form and multiwavelets with applications to the Poisson and bound-state Helmholtz kernels in three dimensions, *Mol. Phys.*, 2013, **111**(9–11), 1143–1160.
- 21 L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman and R. Schneider, Daubechies wavelets as a basis set for density functional pseudopotential calculations, *J. Chem. Phys.*, 2008, **129**, 014109.
- 22 L. Grasedyck, D. Kressner and C. Tobler, *A literature survey of low-rank tensor approximation techniques*, 2013, arXiv:1302.7121v1.
- 23 L. Greengard and V. Rokhlin, A fast algorithm for particle simulations, *J. Comput. Phys.*, 1987, **73**, 325.
- 24 M. Griebel and J. Hamaekers, Sparse grids for the Schroedinger equation, *M2AN*, 2007, **41**, 215–247.
- 25 W. Hackbusch and B. N. Khoromskij, Low-rank Kronecker product approximation to multi-dimensional nonlocal operators, Part I. Separable approximation of multi-variate functions, *Computing*, 2006, **76**, 177–202.
- 26 W. Hackbusch and S. Kühn, A new scheme for the tensor representation, *J. Fourier Anal. Appl.*, 2009, **15**, 706–722.
- 27 W. Hackbusch and R. Schneider, in *Tensor Spaces and Hierarchical Tensor Representations*, ed. S. Dahlke and W. Dahmen *et al.*, Lecture Notes in Computer Science and Engineering, 102, Springer, 2014.
- 28 R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan and G. Beylkin, Multiresolution quantum chemistry: Basic theory and initial applications, *J. Chem. Phys.*, 2004, **121**(23), 11587–11598.
- 29 T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, New York, 1999.
- 30 N. Higham, in *Analysis of the Cholesky decomposition of a semi-definite matrix*, ed. M. G. Cox and S. J. Hammarling, Reliable Numerical Computations, Oxford University Press, Oxford, 1990, pp. 161–185.
- 31 F. L. Hitchcock, The expression of a tensor or a polyadic as a sum of products, *J. Math. Phys.*, 1927, **6**, 164–189.
- 32 E. G. Hohenstein, R. M. Parrish and T. J. Martinez, Tensor hypercontraction density fitting, Quartic scaling second- and third-order Møller-Plesset perturbation theory, *J. Chem. Phys.*, 2012, **137**, 044103.
- 33 V. Kazeev and B. N. Khoromskij, Explicit low-rank QTT representation of Laplace operator and its inverse, *SIAM J. Matrix Anal. Appl.*, 2012, **33**(3), 742–758.
- 34 B. N. Khoromskij,  $O(d \log N)$ -Quantics Approximation of N-d Tensors in High-Dimensional Numerical Modeling, *J. Const. Approx.*, 2011, **34**(2), 257–289, Preprint 55/2009, MPI MiS, Leipzig 2009, <http://www.mis.mpg.de/publications/preprints/2009/prepr2009-55.html>.
- 35 B. N. Khoromskij, Fast and Accurate Tensor Approximation of Multivariate Convolution with Linear Scaling in Dimension, *J. Comput. Appl. Math.*, 2010, **234**, 3122–3139.
- 36 B. N. Khoromskij, Structured Rank- $(r_1, \dots, r_d)$  Decomposition of Function-related Tensors in  $\mathbb{R}^d$ , *Comput. Method. Appl. Math.*, 2006, **6**(2), 194–220.
- 37 B. N. Khoromskij, in *Tensor Numerical Methods for High-dimensional PDEs: Basic Theory and Initial Applications*, ESAIM: Proceedings and Surveys, ed. N. Champagnat, T. Lelièvre and A. Nouy, 2015, vol. 48, pp. 1–28, DOI: 10.1051/proc/201448001.
- 38 B. N. Khoromskij and V. Khoromskaia, Low Rank Tucker-Type Tensor Approximation to Classical Potentials, *Cent. Eur. J. Math.*, 2007, **5**(3), 523–550.
- 39 B. N. Khoromskij and V. Khoromskaia, Multigrid Accelerated Tensor Approximation of Function Related Multi-dimensional Arrays, *SIAM J. Sci. Comput.*, 2009, **31**(4), 3002–3026.
- 40 V. Khoromskaia, Computation of the Hartree–Fock Exchange in the Tensor-structured Format, *Comput. Method. Appl. Math.*, 2010, **10**(2), 1–16.
- 41 V. Khoromskaia, *Numerical Solution of the Hartree–Fock Equation by Multilevel Tensor-structured methods*, PhD dissertation, TU Berlin, 2010, <http://opus.kobv.de/tuberlin/volltexte/2011/2948/>.
- 42 B. N. Khoromskij, V. Khoromskaia and H.-J. Flad, Numerical Solution of the Hartree–Fock Equation in Multilevel Tensor-structured Format, *SIAM J. Sci. Comput.*, 2011, **33**(1), 45–65.
- 43 V. Khoromskaia, D. Andrae and B. N. Khoromskij, Fast and accurate 3D tensor calculation of the Fock operator in a general basis, *Comput. Phys. Commun.*, 2012, **183**, 2392–2404.
- 44 V. Khoromskaia, B. N. Khoromskij and R. Schneider, Tensor-structured factorized calculation of two-electron integrals in a general basis, *SIAM J. Sci. Comput.*, 2013, **35**(2), A987–A1010.



- 45 V. Khoromskaia and B. N. Khoromskij, Møller-Plesset (MP2) Energy Correction Using Tensor Factorizations of the Grid-based Two-electron Integrals, *Comput. Phys. Commun.*, 2014, **185**, 2–10.
- 46 V. Khoromskaia, Black Box Hartree–Fock Solver by the Tensor Numerical Methods, *Comput. Method. Appl. Math.*, 2014, **14**(1), 89–111.
- 47 V. Khoromskaia and B. N. Khoromskij, Grid-based lattice summation of electrostatic potentials by assembled canonical tensors, *Comput. Phys. Commun.*, 2014, **185**, 3162–3174.
- 48 V. Khoromskaia and B. N. Khoromskij, *Tensor Approach to Linearized Hartree–Fock Equation for Lattice-type and Periodic Systems*, 2014, Preprint 62/2014, MIS MPI, arXiv:1408.3839, submitted.
- 49 V. Khoromskaia and B. N. Khoromskij, Tucker tensor method for fast grid-based summation of long-range potentials on 3D lattices with defects, 2014, Preprint arXiv:14.1994.
- 50 T. G. Kolda and B. W. Bader, Tensor Decompositions and Applications, *SIAM Rev.*, 2009, **51**(3), 455–500.
- 51 D. Kressner, *Numerical methods and software for general and structured eigenvalue problems*, PhD dissertation, TU Berlin, 2004.
- 52 K. N. Kudin and G. E. Scuseria, Revisiting infinite lattice sums with the periodic Fast Multipole Method, *J. Chem. Phys.*, 2004, **121**, 2886–2890.
- 53 S. A. Losilla, D. Sundholm and J. Juselius, The direct approach to gravitation and electrostatics method for periodic systems, *J. Chem. Phys.*, 2010, **132**(2), 024102.
- 54 Ch. Lubich, On Variational Approximations in Quantum Molecular Dynamics, *Math. Comput.*, 2005, **74**, 765–779.
- 55 L. Lin, Y. Saad and C. Yang, *Approximating spectral densities of large matrices*, 2014, Preprint arXiv:1308.5467v2.
- 56 I. V. Oseledets, Approximation of  $2^d \times 2^d$  matrices using tensor decomposition, *SIAM J. Matrix Anal. Appl.*, 2010, **31**(4), 2130–2145.
- 57 I. V. Oseledets and E. E. Tyrtshnikov, Breaking the Curse of Dimensionality, or How to Use SVD in Many Dimensions, *SIAM J. Sci. Comput.*, 2009, **31**(5), 3744–3759.
- 58 C. Pisani, M. Schütz, S. Casassa, D. Usvyat, L. Maschio, M. Lorenz and A. Erba, CRYSCOR: a program for the post-Hartree–Fock treatment of periodic systems, *Phys. Chem. Chem. Phys.*, 2012, **14**, 7615–7628.
- 59 R. Parrish, E. G. Hohenstein, T. J. Martinez and C. D. Sherrill, Discrete variable representation in electronic structure theory: Quadrature grids for least-squares tensor hypercontraction, *J. Chem. Phys.*, 2013, **138**, 194107.
- 60 P. Pulay, Improved SCF convergence acceleration, *J. Comput. Chem.*, 1982, **3**, 556–560.
- 61 E. Ribolini, J. Toulouse and A. Savin, in *Electronic excitation energies of molecular systems from the Bethe–Salpeter equation: Example of H<sub>2</sub> molecule*, ed. S. Ghosh and P. Chattaraj, *Concepts and Methods in Modern Theoretical Chemistry*, Electronic Structure and Reactivity, 2013, vol. 1, p. 367.
- 62 U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.*, 2011, **326**(1), 96–192.
- 63 H. Sekino, Y. Maeda, T. Yanai and R. J. Harrison, Basis set limit Hartree Fock and density functional theory response property evaluation by multiresolution multiwavelet basis, *J. Chem. Phys.*, 2008, **129**, 034111.
- 64 A. Smilde, R. Bro and P. Geladi, *Multi-way Analysis*, Wiley, 2004.
- 65 F. Stenger, *Numerical methods based on Sinc and analytic functions*, Springer-Verlag, Heidelberg, 1993.
- 66 D. Sundholm, P. Pykkö and L. Laaksonen, Two-dimensional Fully Numerical Molecular Calculations. X. Hartree–Fock Results for He<sub>2</sub>, Li<sub>12</sub>, Be<sub>2</sub>, HF, OH<sup>−</sup>, N<sub>2</sub>, CO, BF, NO<sup>+</sup>, and CN<sup>−</sup>, *Mol. Phys.*, 1985, **56**, 1411–1418.
- 67 A. Szabo and N. Ostlund, *Modern Quantum Chemistry*, Dover Publication, New York, 1996.
- 68 L. R. Tucker, Some mathematical notes on three-mode factor analysis, *Psychometrika*, 1966, **31**, 279–311.
- 69 F. Verstraete, D. Porras and J. I. Cirac, DMRG and periodic boundary conditions: A quantum information perspective, *Phys. Rev. Lett.*, 2004, **93**(22), 227205.
- 70 G. Vidal, Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.*, 2003, **91**(14), 147902.
- 71 H. Wang and M. Thoss, Multilayer formulation of the multi-configuration time-dependent Hartree theory, *J. Chem. Phys.*, 2003, **119**, 1289–1299.
- 72 H.-J. Werner and P. J. Knowles *et al.*, MOLPRO, version 2002.10, a package of ab initio programs for electronic structure calculations.
- 73 S. R. White, Density-matrix algorithms for quantum renormalization groups, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1993, **48**(14), 10345–10356.
- 74 H. Yserentant, *Regularity and Approximability of Electronic Wave Functions*, Lecture Notes in Mathematics series, Springer-Verlag, 2010.

