PREPRINT 444

FEBRUARY 2016

igpm | RWTH AACHEN UNIVERSITY

Tensor Representation Techniques for Full Configuration Interaction: A Fock Space Approach Using the Canonical Product Format

Karl-Heinz Böhm[a], Mike Espig[b] and Alexander A. Auer[a] *

Institut für Geometrie und Praktische Mathematik
Templergraben 55, 52062 Aachen, Germany

*Corresponding author. Email: alexander.auer@cec.mpg.de

Tensor Representation Techniques for Full Configuration Interaction: A Fock Space Approach Using the Canonical Product Format

Karl-Heinz Böhm[a], Mike Espig[b] and Alexander A. Auer[a] *

[a] *Max-Planck-Institut for Chemical Energy Conversion*
Stiftstraße 34-36, D-45470 Mülheim a.d.R,Germany

[b] *RWTH Aachen, Institut für Geometrie und Praktische Mathematik*
Pontdriesch 14-16, D-52062 Aachen, Germany

5th February 2016

Abstract

In this proof-of-principle study we apply tensor decomposition techniques to the Full Configuration Interaction (FCI) wavefunction in order to reduce the exponential scaling of the number of wavefunction parameters and overall computational effort. For this purpose, the wavefunction ansatz is formulated in an occupation number vector representation that ensures antisymmetry. If the canonical product format tensor decomposition is then applied, the Hamiltonian and the wavefunction can be cast into a multilinear product form. As a consequence, the number of wavefunction parameters does not scale to the power of the number of particles (or orbitals) but depends on the rank of the approximation and linearly on the number of particles. The degree of approximation can be controlled by a single threshold for the rank reduction procedure required in the algorithm. We demonstrate that using this approximation, the FCI Hamiltonian matrix and the wavefunction parameters can be stored with N^5 scaling and the FCI problem can be solved with subexponential effort. The error of the approximation that is introduced is below Millihartree for a threshold of $\epsilon = 10^{-4}$ and no convergence problems are observed solving the FCI equations iteratively in the new format. While promising conceptually, all effort of the algorithm is shifted to the required rank reduction procedure after the contraction of the Hamiltonian with the coefficient tensor. At the current state, this crucial steps scales beyond N^{10}.

*Corresponding author. Email: alexander.auer@cec.mpg.de

1 Introduction

Since the early days of quantum chemistry it has been known that the full configuration interaction (FCI) method allows one to calculate the exact solution of the Schrödinger equation in principle. FCI is conceptually simple and obtaining the energies of the desired system only requires solving an eigenvalue problem. Hence one could follow Dirac's famous statement that "the fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved."
Indeed, as in FCI the wavefunction is parametrized using the weights of all possible configuration state functions (or all possible slater determinants) the number of parameters grows exponentially with the number of particles. Therefore, FCI is usually only used for benchmark calculations and in spite of the remarkable increase of computational power in recent decades solving the FCI problem for few states using Davidson-type algorithms is still only feasible for systems with much less than 20 electrons. However, FCI should not be regarded as an approach with a tiny niche in quantum chemistry - As the central part of a complete active space (CAS) treatment [1] it is essential for solving the most challenging multi-reference problems in chemistry when other methods fail.
And in spite the paradigm that solving Schrödinger's equation demands exponential effort, several attempts have been made to accelerate the method and make it feasible for larger system sizes and larger active spaces. One important improvement was achieved through the insight that the Hamiltonian matrix is data sparse, and only a few elements are required for performing FCI, which means that the Hamiltonian matrix does not need to be constructed explicitly. This, in turn, led to the development of direct CI[2] which can be applied to larger systems [3, 4, 5], but the factorial scaling behavior itself remains. In recent years, novel approaches have been developed to solve the Schrödinger equation for larger systems. One method of reducing the computational effort in FCI is to neglect certain configurations. Several schemes for selecting such configurations have been developed [6, 7, 8, 9, 10, 11]. Other approaches are not based on solving an eigenvalue problem but instead apply Monte Carlo simulations, which has resulted in the development of Quantum Monte Carlo (QMC)[12, 13] and Quantum Monte Carlo Full Configuration Interaction (QMCFCI) [14, 15, 16]. Using these methods, the size of treatable systems can be extended. In case of QMC the scaling can be reduced to polynomial scaling while in the case of QMCFCI the scaling is still exponential but with a decreased prefactor. Both methods can treat larger systems than conventional state of the art FCI algorithms.
A further approach to solve the Schrödinger equation is the density matrix renormalization group (DMRG) method where the electron density is approximated using a tensor representation [17, 18]. This Ansatz exhibits polynomial scaling, but convergence and accuracy can be difficult to control. energy depend on the ordering of the sites[19, 20, 21, 22, 23].
So, while recent years have brought about methods that, while not changing the exponential scaling of FCI, are however extremely efficient and have extended its applicability significantly. On the other hand, methods like QMC or DMRG suggest that it should even be possible to find ways to solve the FCI problem in subexponential effort.
The approach the we explore in this publication goes along the lines of the latter attempts - to find a formulation of the FCI problem that does not exhibit the exponential scaling inherently and for now not to focus in the efficiency of the final algorithm. For this purpose, we have been investigating tensor decomposition techniques from applied mathematics. In principle, these techniques are widely known in the field of quantum chemistry from approaches such as density fitting [24, 25, 26, 27, 28, 29, 30], or Cholesky Decomposition [31, 32, 33] and have also been applied in various ways to post Hartree-Fock methods. One example is the tensor hypercontraction format [34, 35] which has has been applied to Møller-Plesset perturbation theory [36] and Coupled Cluster methods [37, 38, 39]. Tensor decompositions may also be used in multiresolution analysis representations for Hartree-Fock and MP2 [40, 41], where both operators and coefficient tensors are decomposed. Furthermore, tensor decomposition techniques have also been applied to reduce the scaling of the FCI method [42].

Driven by the potential of tensor decomposition techniques to break the curse of dimensionality in quantum chemistry, our recent work has focused on the canonical product (CP) tensor format, which has been applied to post Hartree-Fock methods like MP2 and CCD [43, 44, 45], leading to a reduction of the scaling in comparison to index based algorithms. Furthermore, we have applied the Matrix Product State format to represent two electron integrals and MP2 amplitudes and compared the advantages and disadvantages to the CP format[46].
In the present work, we present a FCI algorithm based on the canonical product tensor format, which is used to represent the Fock space. In section 2.1 we introduce the basic ideas of FCI. In sections 2.2 and 2.3 it is shown

how the representation of the coefficient and Hamiltonian tensors are constructed. The next two sections focus on how the crucial steps of the denominator update and the tensor contraction of the coefficient tensor with the Hamiltonian tensor are performed. Section 2.6 gives an overview over the applied rank reduction algorithm and in Section 2.7 an outline of the current algorithm is given including timings of the crucial steps. To estimate the reliability of our algorithm, results of benchmark calculations are presented in section 3.1 and the scaling behavior with increasing system size is discussed in section 3.2.

2 Theory

2.1 Introduction of the FCI Approach

The full CI wave function $|\Phi^{(FCI)}\rangle$ is expanded as a linear combination of slater determinants.

$$|\Phi^{(FCI)}\rangle = \sum_{\mathbf{K}} C_{\mathbf{K}} |\mathbf{K}\rangle. \tag{1}$$

Each slater determinant $|\mathbf{K}\rangle$ can be represented by an occupation number vector (ONV) $|\mathbf{K}\rangle$ containing the occupation numbers k_p of spin orbitals χ_p. For a given basis of n molecular orbitals the ONV has the following form [47]:

$$|\mathbf{K}\rangle = |k_1, k_2, \ldots, k_n\rangle \text{ with } k_p = \begin{cases} 1 & \text{if } \chi_p \text{ is occupied} \\ 0 & \text{if } \chi_p \text{ is unoccupied} \end{cases} \tag{2}$$

If all $|\mathbf{K}\rangle$ constitute an orthonormal basis, the application of the variational principle in eq. 1 leads to the following eigenvalue problem:

$$\mathbf{HC} = \mathbf{EC}. \tag{3}$$

The elements of the Hamiltonian matrix $\mathbf{H}$ are given in eq. 4, where $\hat{H}$ is expressed in second quantisation:

$$\langle \mathbf{K}|\hat{H}|\mathbf{K}'\rangle = \langle \mathbf{K}| \left(\sum_{pp'} h(p,p') a_p^\dagger a_{p'} + \frac{1}{2} \sum_{pqp'q'} \langle pq||p'q'\rangle a_p^\dagger a_q^\dagger a_{p'} a_{q'} \right) |\mathbf{K}'\rangle. \tag{4}$$

Here $h(p,p')$ is the one-electron integral

$$h(p,p') = \int \chi_p^*(r_1) \left(-\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{A1}} \right) \chi_{p'}(r_1) dr_1 \tag{5}$$

and $\langle pq||p'q'\rangle = \langle pq|p'q'\rangle - \langle pq|q'p'\rangle$ corresponds to the antisymmetrised two electron integral with

$$\langle pq|p'q'\rangle = \int \chi_p^*(r_1)\chi_q^*(r_2) \frac{1}{r_{12}} \chi_{p'}(r_1)\chi_{q'}(r_2) dr_1 dr_2. \tag{6}$$

The creation and annihilation operators $a_p^\dagger$ and a_p in second quantisation are defined as follows [47]:

$$\begin{aligned} a_p^\dagger |k_1, k_2, \ldots, 0_p, \ldots, k_n\rangle \theta &= \Gamma_p^{\mathbf{g}} |k_1, k_2, \ldots, 1_p, \ldots, k_n\rangle \\ a_p^\dagger |k_1, k_2, \ldots, 1_p, \ldots, k_n\rangle &= 0 \\ a_p |k_1, k_2, \ldots, 1_p, \ldots, k_n\rangle &= \Gamma_p^{\mathbf{g}} |k_1, k_2, \ldots, 0_p, \ldots, k_n\rangle \\ a_p |k_1, k_2, \ldots, 0_p, \ldots, k_n\rangle &= 0. \end{aligned} \tag{7}$$

$$\text{with } \Gamma_p^{\mathbf{g}} = \prod_{q=1}^{p-1} (-1)^{k_q}.$$

To obtain the proper coefficients in the wave function expansion (eq. 1) and the corresponding energies, the eigenvalue problem of eq. 3 needs to be solved. In most cases only the lowest eigenvalue E, which corresponds to the ground state energy, is of interest, so that only one coefficient vector C needs to be obtained.

2.2 Canonical Tensor Representation of the Fock Space

While many approaches can be formulated to apply tensor decomposition techniques to FCI, the aim of breaking the curse of dimensionality for a many-particle approach for fermions leads to difficulties in many cases. Indeed, we have tested two further approaches. In the first one, the wave function is expanded hierarchically. Starting from a reference determinant, additional determinants are incorporated into the wave function expansion, where orbitals that are occupied in a reference determinant are replaced by virtual orbitals. In this way, the coefficients belonging to determinants where one, two, ... orbitals have been replaced, are stored in tensors according to the number of replacements. However, this results in the problem that for higher substitutions the tensor contractions get more and more complicated. Therefore, we have tested a further approach that generalises the previous one by storing all coefficients in one representation. In this case there are still problems due to unphysical elements in the Hamiltonian tensor representation, that were supposed to be zero when antisymmetry is correctly treated. In reference [48] we further discuss these two approaches.

In this section, we follow the ansatz by Legeza et al. and Szalay et al. [49, 50] that will be discussed in this work as it does not lead to aforementioned difficulties.

As shown in eq. 2, the ONV possesses n entries k_p, which can each take one of two values. In order to address the proper coefficients, the corresponding coefficient tensor can be arranged in a way that follows a similar logic as is used for the ONVs. As a consequence, the coefficients are elements of an n dimensional tensor $\mathfrak{C}$. Each dimension μ of this tensor can be assigned to one orbital $|\chi_\mu\rangle$ and there exist two entries in each dimension corresponding to an occupied ($k_\mu = 1$) or an unoccupied orbital ($k_\mu = 0$).

The coefficient tensor can be represented in the canonical product tensor format as shown in eqs. 8 and 9 where for each of the J ranks there exist n representing vectors $c_j^{(\mu)}$ and each representing vector holds two elements $c_j^{(\mu)}(0)$ and $c_j^{(\mu)}(1)$.

$$\mathfrak{C} = \sum_{j=1}^{J} c_j^{(1)} \otimes c_j^{(2)} \otimes \ldots \otimes c_j^{(n)} \tag{8}$$

$$= \sum_{j=1}^{J} \bigotimes_{\mu=1}^{n} c_j^{(\mu)}. \tag{9}$$

Below, we illustrate the ansatz using a four-orbital ($n = 4$), two-electron ($N = 2$) example like H_2, which is also used later in this article. In order to use a transparent nomenclature, the coefficients $c_{k_i,k_a,k_{\bar{i}},k_{\bar{a}}}$ are named in the same way as the ONVs, where k_i, k_a, $k_{\bar{i}}$ and $k_{\bar{a}}$ are the corresponding occupation numbers of orbitals χ_i, χ_a, $\chi_{\bar{i}}$ and $\chi_{\bar{a}}$. As shown in eq. 1, each determinant can be assigned one coefficient leading in the present example to a total number of $\binom{n}{N} = \frac{4\cdot3}{2\cdot1} = 6$ coefficients, which are stored in a vector $\vec{c}$ in conventional FCI as shown below:

$$\vec{c} = \begin{pmatrix} c_{0,1,0,1} \\ c_{1,0,0,1} \\ c_{0,1,1,0} \\ c_{1,1,0,0} \\ c_{0,0,1,1} \\ c_{1,0,1,0} \end{pmatrix} \tag{10}$$

The same configuration space can be spanned in a CP format representation of an n dimensional tensor as follows:

$$\mathfrak{C} = \sum_{j=1}^{J} c_j^{(i)} \otimes c_j^{(a)} \otimes c_j^{(\bar{i})} \otimes c_j^{(\bar{a})} \tag{11}$$

where each coefficient $c(k_i, k_{\bar{i}}, k_a, k_{\bar{a}})$ can be obtained from

$$c(k_i, k_a, k_{\bar{i}}, k_{\bar{a}}) = \sum_{j=1}^{J} c_j^{(i)}(k_i) \cdot c_j^{(a)}(k_a) \cdot c_j^{(\bar{i})}(k_{\bar{i}}) \cdot c_j^{(\bar{a})}(k_{\bar{a}}). \tag{12}$$

Suppose that only one coefficient $c(1,0,1,0)$ that belongs to the Hartree-Fock determinant is equal to one while all the other coefficients are zero as, for example, in the initial guess of an FCI procedure. In this case the vector $\vec{c}$ contains six elements

$$\vec{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{13}$$

while in the representation of the coefficient tensor it takes the form of a rank one expansion:

$$\mathfrak{C} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{14}$$

Note that such a tensor representation covers the whole Fock space instead of an N particle subspace of the Fock space. For example, elements belonging to the vacuum state $|0,0,0,0\rangle$ are accessible in the representation as:

$$c(0,0,0,0) = \sum_{j=1}^{J} c_j^{(i)}(0) \cdot c_j^{(a)}(0) \cdot c_j^{(\bar{i})}(0) \cdot c_j^{(\bar{a})}(0). \tag{15}$$

There are many instances of coefficient tensors that can occur for a four-orbital, two-electron system. Some examples of representations of the coefficient tensors are given in Fig. 1.

The wave function is obtained by inserting the coefficient tensor representation 8 into eq. 1:

$$|\Phi^{(FCI)}\rangle = \sum_k C(k_1, k_2, \ldots, k_n)|k_1, k_2, \ldots, k_n\rangle \tag{16}$$

$$= \sum_k \sum_{j=1}^{J} c_j^{(1)}(k_1) \cdot c_j^{(2)}(k_2) \cdot \ldots \cdot c_j^{(n)}(k_n)|k_1, k_2, \ldots, k_n\rangle. \tag{17}$$

Using the tensor representation of the coefficient tensor for the aforementioned four-orbital, two-electron example would yield the explicit form of the wave function

$$|\Phi^{(FCI)}\rangle = \sum_k \sum_{j=1}^{J} c_j^{(1)}(k_1) \cdot c_j^{(2)}(k_2) \cdot c_j^{(3)}(k_3) \cdot c_j^{(4)}(k_4)|k_1, k_2, k_3, k_4\rangle. \tag{18}$$

2.3 Decomposed Representation of the Hamiltonian Matrix Using Canonical Products

As introduced in section 2.1, to obtain the appropriate coefficients, an eigenvalue problem of a matrix $\mathbf{H}$ needs to be solved. To find a tensor representation of the Hamiltonian matrix, the representation needs to be constructed in an analogous way as the matrix elements in eq. 4. In a first step tensor representations are set up that are analogues of the annihilation and creation operators $a_p^\dagger$ and a_p. Therefore, matrices $\mathbf{A}^T$, $\mathbf{A}$, $\mathbf{1}$ and $\mathbf{G}$ are introduced [49, 50]:

$$\mathbf{A}^T = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{19}$$

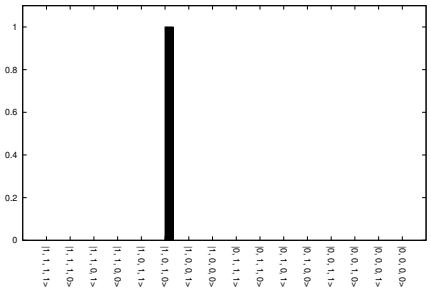
$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{20}$$

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{21}$$

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{22}$$

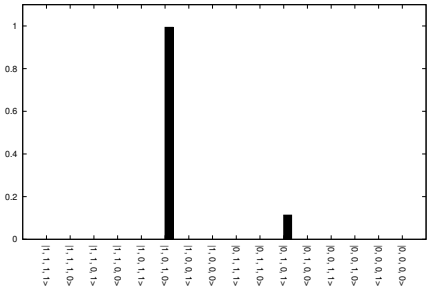
HF-Solution of H_2

$$
\begin{aligned}
|\Psi\rangle &= 1 \cdot |1,0,1,0\rangle \\
\mathfrak{C} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\end{aligned}
$$



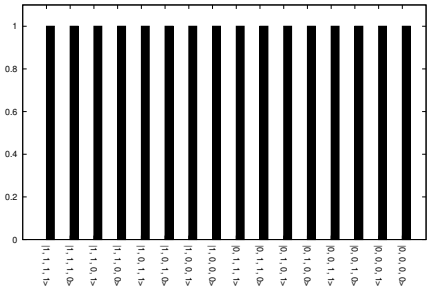
FCI-Solution of H_2

$$
\begin{aligned}
|\Psi\rangle &= 0.99 \cdot |1,0,1,0\rangle - 0.11 \cdot |0,1,0,1\rangle \\
\mathfrak{C} &= \begin{pmatrix} 0 \\ 0.99 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&+ \begin{pmatrix} -0.11 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}
$$



Equal Coefficients for all Determinants in the Fock-Space

$$
\begin{aligned}
|\Psi\rangle &= \sum_{k_i} \sum_{k_a} \sum_{k_{\bar{i}}} \sum_{k_{\bar{a}}} 1 \cdot |k_i, k_a, k_{\bar{i}}, k_{\bar{a}}\rangle \\
\mathfrak{C} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix}
\end{aligned}
$$



Equal Coefficients for all two-electron Determinants

$$
\begin{aligned}
|\Psi\rangle &= 1 \cdot |1,0,1,0\rangle + 1 \cdot |0,1,0,1\rangle + 1 \cdot |0,1,1,0\rangle \\
&= 1 \cdot |1,0,0,1\rangle + 1 \cdot |1,1,0,0\rangle + 1 \cdot |0,0,1,1\rangle \\
\mathfrak{C} &= \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \\
&+ \begin{pmatrix} 0.58 \\ -1.22 \end{pmatrix} \otimes \begin{pmatrix} -0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ -1.22 \end{pmatrix} \otimes \begin{pmatrix} -0.58 \\ 1.22 \end{pmatrix} \\
&+ \begin{pmatrix} 0 \\ 1.45 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ -1.45 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ -1.45 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ -1.45 \end{pmatrix} \\
&+ \begin{pmatrix} 0.69 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} -0.69 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} -0.69 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} -0.69 \\ -0 \end{pmatrix}
\end{aligned}
$$

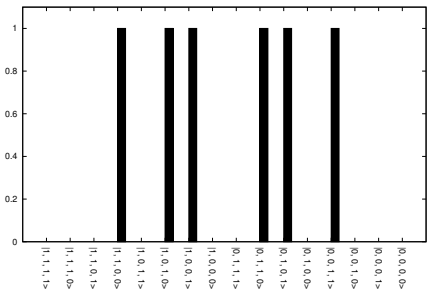


Figure 1: Examples of coefficient tensor representations of the four-orbital, two-electron example. Shown are an explicit form of the wave function using ONVs, the representation of the coefficient tensor, and a bar plot of the absolute values of the coefficients.

Using these matrices, tensors $\mathfrak{A}_p^\dagger$ and $\mathfrak{A}_p$ representing the creation and annihilation operators can be constructed as follows:

$$\mathfrak{A}_p^\dagger = \mathbf{G}^{(1)} \otimes \mathbf{G}^{(2)} \otimes \ldots \otimes \mathbf{G}^{(p-1)} \otimes \mathbf{A}^{(p)T} \otimes \mathbf{1}^{(p+1)} \otimes \ldots \otimes \mathbf{1}^{(n)} \tag{23}$$

$$\mathfrak{A}_p = \mathbf{G}^{(1)} \otimes \mathbf{G}^{(2)} \otimes \ldots \otimes \mathbf{G}^{(p-1)} \otimes \mathbf{A}^{(p)} \otimes \mathbf{1}^{(p+1)} \otimes \ldots \otimes \mathbf{1}^{(n)}. \tag{24}$$

In analogy to eqs. 7, these tensors can be contracted with tensor representations of the occupation number vectors. Again, consider the four-orbital, two-electron example. Application of a creation operator $a_a^\dagger$ on the ONV $|1,0,1,0\rangle$ gives:

$$a_a^\dagger |1,0,1,0\rangle = -1|1,1,1,0\rangle \tag{25}$$

If a tensor representation $\mathfrak{E}$ of this occupation number vector

$$\mathfrak{E} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{26}$$

is used, the elements of the well known form of the ONV are obtained by setting the occupation numbers k_p equal to the elements $e^{(p)}(1)$ of the representing vectors. In analogy to the operation shown in eq. 25 a corresponding tensor contraction $\mathfrak{A}_a^\dagger \mathfrak{E}$ is done as follows:

$$\mathfrak{A}_a^\dagger \mathfrak{E} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{27}$$

$$= \begin{pmatrix} 0 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{28}$$

Note that the resulting representation of an ONV is exactly the same as the well known form of the ONV, where the prefactor is multiplied with the representing vector of the first dimension. In order to set up the Hamiltonian tensor, the contributions of all one- and two-electron integrals from all ONVs $\langle \mathbf{K}|$ and $|\mathbf{K}'\rangle$ need to be included. For simplicity, just consider the one-electron part $\sum_{pp'} h(p,p') a_p^\dagger a_{p'}$ of eq. 4. In analogy to the term, $h(p,p')a_p^\dagger a_{p'}$, a rank one tensor representation $\mathfrak{H}^{pp'}$ can be set up that contains only elements $\langle \mathbf{K}|h(p,p')a_p^\dagger a_{p'}|\mathbf{K}'\rangle$ as in eq. 29. The contraction is done as shown in eq. 30 where the representations of $\mathfrak{A}_p^\dagger$ and $\mathfrak{A}_{p'}$ from eqs. 23 and 24 are applied, and the representing matrices are multiplied with each other in each dimension:

$$\mathfrak{H}^{pp'} = h(p,p')\mathfrak{A}_p^\dagger \mathfrak{A}_{p'} \tag{29}$$

$$\begin{aligned} = {} & h(p,p') \cdot \mathbf{G}^{(1)}\mathbf{G}^{(1)} \otimes \ldots \otimes \mathbf{G}^{(p'-1)}\mathbf{G}^{(p'-1)} \otimes \mathbf{G}^{(p')}\mathbf{A}^{(p')} \otimes \\ & \mathbf{G}^{(p'+1)}\mathbf{1}^{(p'+1)} \otimes \ldots \otimes \mathbf{G}^{(p-1)}\mathbf{1}^{(p-1)} \otimes \mathbf{A}^{(p)T}\mathbf{1}^{(p)} \otimes \mathbf{1}^{(p+1)}\mathbf{1}^{(p+1)} \otimes \\ & \ldots \otimes \mathbf{1}^{(n)}\mathbf{1}^{(n)}. \end{aligned} \tag{30}$$

The contributions of the two electron integrals to all elements of the Hamiltonian tensor can be obtained in the same way as for the one-electron integrals. Therefore, for each one- and two-electron integral, a rank one tensor can be constructed in the same manner.

The four-orbital, two-electron example introduced above yields tensor representations with four representing matrices for each one- and two-electron integral. Two examples $\mathfrak{H}^{ia}$ and $\mathfrak{H}^{i\bar{i}i\bar{i}}$ of such tensor representations are given below. These representations $\mathfrak{H}^{ia}$ and $\mathfrak{H}^{i\bar{i}i\bar{i}}$ constitute tensors that possess only elements equal to $\langle \mathbf{K}|h(i,a)a_i^\dagger a_a|\mathbf{K}'\rangle$ or $\langle \mathbf{K}|\langle i\bar{i}||i\bar{i}\rangle a_i^\dagger a_{\bar{i}}^\dagger a_{\bar{i}} a_i|\mathbf{K}'\rangle$, respectively, where $\langle \mathbf{K}|$ and $|\mathbf{K}'\rangle$ denote arbitrary ONVs:

$$\mathfrak{H}^{ia} = \underbrace{h(i,a) \cdot \mathbf{A}^{(i)T}\mathbf{G}^{(i)}}_{\mathbf{h}_{ia}^{(i)}} \otimes \underbrace{\mathbf{1}^{(a)}\mathbf{A}^{(a)}}_{\mathbf{h}_{ia}^{(a)}} \otimes \underbrace{\mathbf{1}^{(\bar{i})}\mathbf{1}^{(\bar{i})}}_{\mathbf{h}_{ia}^{(\bar{i})}} \otimes \underbrace{\mathbf{1}^{(\bar{a})}\mathbf{1}^{(\bar{a})}}_{\mathbf{h}_{ia}^{(\bar{a})}} \tag{31}$$

$$= \begin{pmatrix} 0 & 0 \\ h(i,a) & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{32}$$

$$\mathfrak{H}^{i\bar{i}i\bar{i}} = \underbrace{\langle i\bar{i}||i\bar{i}\rangle \cdot \mathbf{A}^{(i)T}\mathbf{G}^{(i)}\mathbf{G}^{(i)}\mathbf{A}^{(i)}}_{\mathbf{h}_{i\bar{i}i\bar{i}}^{(i)}} \otimes \underbrace{\mathbf{1}^{(a)}\mathbf{G}^{(a)}\mathbf{G}^{(a)}\mathbf{1}}_{\mathbf{h}_{i\bar{i}i\bar{i}}^{(a)}} \otimes \underbrace{\mathbf{1}^{(\bar{i})}\mathbf{A}^{(\bar{i})T}\mathbf{A}^{(\bar{i})}\mathbf{1}^{(\bar{i})}}_{\mathbf{h}_{i\bar{i}i\bar{i}}^{(\bar{i})}} \otimes \underbrace{\mathbf{1}^{(\bar{a})}\mathbf{1}^{(\bar{a})}\mathbf{1}^{(\bar{a})}\mathbf{1}^{(\bar{a})}}_{\mathbf{h}_{i\bar{i}i\bar{i}}^{(\bar{a})}} \tag{33}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & \langle i\bar{i}||i\bar{i}\rangle \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{34}$$

As mentioned above, the contraction of the representing matrices in each dimension leads to new matrices that are named $\mathbf{h}_{ia}^{(i)}$,$\mathbf{h}_{ia}^{(a)}$, $\mathbf{h}_{ia}^{(\bar{i})}$ and $\mathbf{h}_{ia}^{(\bar{a})}$ for the representation of the one-electron integral $h(i,a)$. The same is done for the two-electron integral contribution.

Now it becomes clear that for all integrals $h(p,p')$ and $\langle pq||p'q'\rangle$ that occur in an N-electron, n-orbital system, a rank one tensor is set up with n representing matrices. By summing up all rank one representing tensors of the one- and two- electron contributions, a tensor representation $\mathfrak{H}$ is built. This representation becomes the *Hamiltonian tensor* given in eqs. 35 to 38 and corresponds to a $2n$ dimensional tensor representation of the Hamiltonian matrix. This representation of matrices by decomposed tensors is similar to the decomposition shown by Beylkin [51]. Note again the similarity between eq. 4 and 35.

$$\begin{aligned}
\mathfrak{H} &= \sum_{p=1}^{n}\sum_{p'=1}^{n}\mathfrak{H}^{pp'} + \sum_{p=1}^{n}\sum_{p'=1}^{n}\sum_{q=1}^{p}\sum_{q'=1}^{p'}\mathfrak{H}^{pqp'q'} && (35)\\
&= \sum_{p=1}^{n}\sum_{p'=1}^{n}\mathbf{h}_{pp'}^{(1)}\otimes\mathbf{h}_{pp'}^{(2)}\otimes\ldots\otimes\mathbf{h}_{pp'}^{(n)} + \\
&\quad \sum_{p=1}^{n}\sum_{p'=1}^{n}\sum_{q=1}^{p}\sum_{q'=1}^{p'}\mathbf{h}_{pqp'q'}^{(1)}\otimes\mathbf{h}_{pqp'q'}^{(2)}\otimes\ldots\otimes\mathbf{h}_{pqp'q'}^{(n)} && (36)\\
&= \sum_{k=1}^{K}\mathbf{h}_{k}^{(1)}\otimes\mathbf{h}_{k}^{(2)}\otimes\ldots\otimes\mathbf{h}_{k}^{(n)} && (37)\\
&= \sum_{k=1}^{K}\bigotimes_{\mu=1}^{n}\mathbf{h}_{k}^{(\mu)}. && (38)
\end{aligned}$$

Because all one- and two-electron integral contributions are taken into account, this representation is exact. The rank K of this representation is dominated by the two electron contributions, as their number rises with n^4 (if no sparsity in the integrals is assumed). Because of the n^4 dependence of the rank and the n 2×2 representing matrices, the overall scaling of this representation with the system size is $\sim\mathcal{O}(N^5)$ under the assumption that the the number of molecular orbitals n is proportional to the number of electrons N. This is a significant reduction from a factorial scaling of a $\binom{n}{N}\times\binom{n}{N}$ matrix to a $\sim\mathcal{O}(N^5)$ scaling of an n dimensional tensor.

For the four-orbital, two-electron example each element $\langle k_i,k_a,k_{\bar{i}},k_{\bar{a}}|\hat{H}|k_i',k_a',k_{\bar{i}}',k_{\bar{a}}'\rangle$ of the Hamiltonian matrix can be obtained via the corresponding ONVs:

$$\langle k_i,k_a,k_{\bar{i}},k_{\bar{a}}|\hat{H}|k_i',k_a',k_{\bar{i}}',k_{\bar{a}}'\rangle = \sum_{k=1}^{K}h_k^{(i)}(k_i,k_i')\cdot h_k^{(a)}(k_a,k_a')\cdot h_k^{(\bar{i})}(k_{\bar{i}},k_{\bar{i}}')\cdot k_k^{(\bar{a})}(k_{\bar{a}},k_{\bar{a}}'). \tag{39}$$

Again, it is important to mention, that the present representation covers the whole Fock space of a given basis and not just an N electron subspace. Thus, for the aforementioned example also elements that even belong to the vacuum state $|0,0,0,0\rangle$ ($N=0$) are accessible and are not necessarily equal to zero as shown below.

$$\langle 0,0,0,0|\hat{H}|0,0,0,0\rangle = \sum_{k=1}^{K}h_k^{(i)}(0,0)\cdot h_k^{(a)}(0,0)\cdot h_k^{(\bar{i})}(0,0)\cdot h_k^{(\bar{a})}(0,0) \tag{40}$$

2.4 Tensor Contraction

As known from conventional algorithms, one of the crucial steps in FCI algorithms is the contraction $\sigma=\mathbf{H}C$ [47]. This means that every element $\sigma_{\mathbf{K}}$ is obtained by contracting the Hamiltonian matrix with the coefficient vector.

$$\sigma_{\mathbf{K}} = \sum_{\mathbf{K}'}\langle\mathbf{K}|\hat{H}|\mathbf{K}'\rangle C_{\mathbf{K}'} \tag{41}$$

However, in the present approach this step is considerably simplified because the contraction needs to be done only for the representing matrices and vectors rather than of the whole tensors. Furthermore, the coefficient tensor retains the same structure as an ONV in the tensor representation so that each coefficient can be addressed in the same way. In order to illustrate the contraction, the four-orbital, two-electron example is used again. Here the ONVs are given explicitly instead of $\mathbf{K}'$ and $\mathbf{K}$. By inserting the explicit elements of the Hamiltonian and the coefficient tensors (eqs. 39 and 12) into eq. 41 and rearranging the summations, the contraction in the chosen tensor format is carried out by performing simple matrix-vector multiplications in each dimension.

$$\begin{aligned}
\sigma_{k_i,k_a,k_{\bar{i}},k_{\bar{a}}} &= \sum_{k'_i}\sum_{k'_a}\sum_{k'_{\bar{i}}}\sum_{k'_{\bar{a}}}\langle k_i,k_a,k_{\bar{i}},k_{\bar{a}}|\hat{H}|k'_i,k'_a,k'_{\bar{i}},k'_{\bar{a}}\rangle C_{k'_i,k'_a,k'_{\bar{i}},k'_{\bar{a}}} && (42)\\
&= \sum_{k'_i}\sum_{k'_a}\sum_{k'_{\bar{i}}}\sum_{k'_{\bar{a}}}\sum_{k=1}^{K} h_k^{(i)}(k_i,k'_i)\cdot h_k^{(a)}(k_a,k'_a)\cdot h_k^{(\bar{i})}(k_{\bar{i}},k'_{\bar{i}})\cdot h_k^{(\bar{a})}(k_{\bar{a}},k'_{\bar{a}})\\
&\quad \sum_{j=1}^{J} c_j^{(i)}(k'_i)\cdot c_j^{(a)}(k'_a)\cdot c_j^{(\bar{i})}(k'_{\bar{i}})\cdot c_j^{(\bar{a})}(k'_{\bar{a}}) && (43)\\
&= \sum_{k=1}^{K}\sum_{j=1}^{J}\sum_{k'_i} h_k^{(i)}(k_i,k'_i)c_j^{(i)}(k'_i)\cdot\sum_{k'_a} h_k^{(a)}(k_a,k'_a)\cdot c_j^{(a)}(k'_a)\cdot\sum_{k'_{\bar{i}}} h_k^{(\bar{i})}(k_{\bar{i}},k'_{\bar{i}})c_j^{(\bar{i})}(k'_{\bar{i}})\\
&\quad \cdot\sum_{k'_{\bar{a}}} h_k^{(\bar{a})}(k_{\bar{a}},k'_{\bar{a}})c_j^{(\bar{a})}(k'_{\bar{a}}) && (44)\\
&= \sum_{m=1}^{M=J\cdot K} s_m^{(i)}(k_i)\cdot s_m^{(a)}(k_a)\cdot s_m^{(\bar{i})}(k'_{\bar{i}})\cdot s_m^{(\bar{a})}(k'_{\bar{a}}) && (45)
\end{aligned}$$

A more general form of the contraction is shown in eq. 46. For each dimension μ, the contraction $\mathbf{h}_k^{(\mu)}c_j^{(\mu)}$ must be carried out for all ranks of the Hamiltonian tensor K and for all ranks of the coefficient tensor J. Hence, the rank M of the resulting tensor $\mathfrak{S}$ is increased and needs to be reduced afterwards.

$$\begin{aligned}
\mathfrak{S} &= \sum_{k=1}^{K}\sum_{j=1}^{J}\bigotimes_{\mu=1}^{n}\left(\mathbf{h}_k^{(\mu)}c_j^{(\mu)}\right) && (46)\\
&= \sum_{m=1}^{M=J\cdot K}\bigotimes_{\mu=1}^{n} s_m^{(\mu)} && (47)
\end{aligned}$$

The scaling of this contraction will be $K\cdot J\cdot n$. Note that there is no further logic required for the ordering of coefficients or Hamiltonian tensor elements.

Beside this contraction in a decomposed format, further mathematical operations in a decomposed format are required for an eigenvalue solver. For further operations such as analogons to the scalar product or summations, interested readers are referred to the corresponding literature [52, 53].

2.5 Denominator Update

The Davidson algorithm [54] requires the construction of a matrix $(\mathbf{H}-\lambda Id)^{-1}$. To set up a tensor representation of this matrix, some assumptions are made because there exists no trivial solution. Because a representation of such a tensor in n dimensions should possess a low rank, the representation is approximated by exponential sums [55]. Therefore, the diagonal dominance of Hamiltonian matrix is exploited by considering only the elements $\langle\Psi_K|\hat{H}|\Psi_K\rangle$ in the representation. All other elements are assumed to be zero. For example, one element $\left(H_{i,\bar{i}}-\lambda Id\right)^{-1}$ for the four-orbital, two-electron case is given below:

$$\Big(\langle 1,0,1,0|\hat{H}|1,0,1,0\rangle - \lambda Id\Big)^{-1} = \frac{1}{\epsilon_i + 0 + \epsilon_{\bar{i}} + 0 - \lambda} \tag{48}$$

$$= \int_0^\infty e^{-(\epsilon_i+0+\epsilon_{\bar{i}}+0-\lambda)t} dt \tag{49}$$

$$\approx \sum_s \omega_s \cdot e^{(\alpha_s\lambda)} \cdot e^{(-\alpha_s\cdot\epsilon_i)} \cdot e^{(-\alpha_s\cdot 0)} \cdot e^{(-\alpha_s\cdot\epsilon_{\bar{i}})} \cdot e^{(-\alpha_s\cdot 0)} \tag{50}$$

$$\approx \sum_s \omega_s \cdot e^{(\alpha_s\lambda)} \cdot e^{(-\alpha_s\cdot\epsilon_i)} \cdot 1 \cdot e^{(-\alpha_s\cdot\epsilon_{\bar{i}})} \cdot 1. \tag{51}$$

More generally, the tensor approximation for the inverse is:

$$(\mathfrak{H} - \lambda\mathfrak{Id})^{-1} \approx \sum_{s=1}^{S} \bigotimes_{\mu=1}^{n} i_s^{(\mu)}. \tag{52}$$

The representing vectors consist of the following elements:

$$i_s^{(x\neq 1)} = \begin{pmatrix} e^{(-\alpha_s\cdot\epsilon_x)} \\ e^0 \end{pmatrix} \text{ and } i_s^{(1)} = \begin{pmatrix} \omega_s \cdot e^{(\alpha_s\cdot\lambda)} \cdot e^{(-\alpha_s\cdot\epsilon_1)} \\ \omega_s e^{(\alpha_s\cdot\lambda)} \end{pmatrix}. \tag{53}$$

The use of exponential sums allows the representation of a sum in the denominator by a sum of products where each summand carries one index. However, the diagonal elements of the Hamiltonian matrix contain a contribution of the two electron integrals that still depend on two indices $\langle pq||pq\rangle$, meaning a connection of two dimensions. Because of this, the exponential sums currently do not allow the inclusion of the two electron integrals. Therefore, the orbital energies ϵ_i are used as a compromise to include the contribution of electron-electron interactions.

2.6 Rank Reduction - Pivotised Alternating Steepest Descent Procedure

The efficient approximation of quantities by low-rank tensor representations is a crucial step in our approach. Usually, the numerical solution of these approximation problem needs most of the computational time of the whole simulation (See Figure 3). While in previous work, we have applied an alternating least square scheme, in the following we outline a more efficient alternative.

Let $\mathfrak{S} = \sum_{i=1}^{R} \bigotimes_{\mu=1}^{n} s_i^{(\mu)}$ and $r < R$ be given. We are looking for $\hat{\xi} = \sum_{j=1}^{r} \bigotimes_{\mu=1}^{d} \hat{\xi}_j^{(\mu)}$ such that

$$\|\mathfrak{S} - \hat{\xi}\| = \min_{\xi\in\mathcal{T}_r} \|\mathfrak{S} - \xi\|, \tag{54}$$

under the constraints

$$\|\hat{\xi}_j^{(\mu)}\| = \|\hat{\xi}_j^{(\mu)}\|, \quad \text{for all } \mu,\nu \in \{1,\ldots,d\},\ j \in \{1,\ldots,r\}, \tag{55}$$

where $\mathcal{T}_r$ denotes the set of tensors which can be represented with r bounded terms.[56] In our applications it is a priori not obvious how to choose the size or the representation rank r. Rather, a desired approximation accuracy ε is of importance. Hence, the following extended approximation problem is stated. For a given accuracy $\varepsilon > 0$ we have to find minimal $r_\varepsilon \leq R$ and $\hat{\xi}_\varepsilon \in \mathcal{T}_{r_\varepsilon}$ such that

$$\|\mathfrak{S} - \hat{\xi}_\varepsilon\| \leq \varepsilon, \tag{56}$$

$$\|\mathfrak{S} - \hat{\xi}_\varepsilon\| = \min_{\xi\in\mathcal{T}_{r_\varepsilon}} \|\mathfrak{S} - \xi\|, \tag{57}$$

under the constraints from Eq. (55). The solution of the extended approximation problem is closely related to the choice of the initial guess for an efficient optimisation method. Here, we are following a scheme which was introduced in [57, 58], where the extended optimisation problem was solved by successive use of a low-rank

approximation method. With the specific definition of the initial guess described in [57, 58], we ensure that the approximation error will not increase during the low-rank approximation of $\mathfrak{S} = \sum_{i=1}^{R} \bigotimes_{\mu=1}^{n} s_i^{(\mu)}$. For a fixed target rank r, the approximation problem from Eq. (54) is solved by the pivotised alternating steepest descent (PASD) method [59]. Since the low-rank approximation in the canonical tensor format is not well defined, we have to consider the following stabilised target function:

$$f_\lambda(\xi_1^{(1)}, \ldots, \xi_r^{(d)}) = \frac{1}{2}\|\xi\|^2 - \langle \xi, b \rangle + \frac{\lambda}{2} \sum_{j=1}^{r} \prod_{\mu=1}^{d} \left\| \xi_j^{(\mu)} \right\|^2, \tag{58}$$

where λ is a positive real number. In Fig. 2 we briefly outline the PASD method, for details and convergence analysis we refer the the reader to [59].

Figure 2: Pivotised Alternating Steepest Descent Algorithm (PASD) method

1: Choose the initial guess as described in [57, 58] and define $k := 1$.
2: **while** Stop Condition **do**
3: $\mu := \text{argmax}_{1 \le \nu \le d} \left\| \frac{\partial f_\lambda}{\partial \xi_\nu}(\xi_{k,\nu}) \right\|_\infty$
4:

$$\begin{aligned}
\xi_{k,\mu} &:= (\xi_1^{k+1}, \ldots, \xi_{\mu-1}^{k+1}, \xi_{\mu+1}^{k}, \ldots, \xi_d^k) \\
d_\mu^k &:= \frac{\partial f_\lambda}{\partial \xi_\nu}(\xi_{k,\nu}) \\
\lambda_\mu^k &:= \frac{\langle d_\mu^k, d_\mu^k \rangle}{\langle G_{k,\mu} d_\mu^k, d_\mu^k \rangle} \\
\xi_\mu^{k+1} &:= \xi_\mu^k - \lambda_\mu^k d_\mu^k \\
\xi_{k,\mu+1} &:= (\xi_1^{k+1}, \ldots, \xi_{\mu}^{k+1}, \xi_{\mu+1}^{k}, \ldots, \xi_d^k)
\end{aligned}$$

5: balance the representation system $\xi_{k,\mu+1}$ such that Eq. (55) is satisfied
6: $k \mapsto k + 1$.
7: **end while**

2.7 Application of the Davidson Algorithm

The Davidson algorithm is used to solve the eigenvalue problem posed in FCI. This algorithm constitutes a very efficient algorithm, since it was developed for sparse and diagonal dominant matrices. The eigenvector is expanded as a linear combination in an orthonormal basis. In each iteration the basis is extended by adding an additional vector, which is derived from the residual vector of the eigenvalue problem. An overview over the algorithm is given in Fig. 3. For further details, see reference [54].

In the present implementation the algorithm proceeds as follows: in a first step a guess tensor C_1 is set up which contains only one coefficient equal to one. This coefficient is assigned to the Hartree-Fock-Slater determinant. All other coefficients are equal to zero (an example for such a tensor is given in eq. 14 and in Fig. 1). Next, the coefficient tensor is contracted with the Hamiltonian tensor as discussed in section 2.4. Subsequently, the rank of the resulting tensor $\tilde{\sigma_1}$ is reduced to σ_1. Note that this rank reduction is the most time consuming step in all iterations. The scalar product $\langle \sigma_1, C_1 \rangle$ provides the first approximation λ_1 to the eigenvalue and using this value the residual r_1 will be calculated. If the norm of the residual is larger than the threshold ϵ_{res}, a tensor δ is constructed. This tensor is orthogonalised to all coefficient tensors of the previous iterations. After normalisation the tensor is added to the basis in which the coefficient tensor is expanded. Now, the contraction with the Hamiltonian tensor and the rank reduction is done in the same way as shown for the first iteration. A matrix $\mathbf{A}$ is constructed whose elements are obtained by the scalar products of all coefficient tensors with all σ-tensors. The eigenvalue λ_{K+1} constitutes the next approximation of the desired eigenvalue while the elements of the eigenvectors are used as the expansion coefficients for the desired coefficient tensors in a tensor representation. Using these coefficients, a new residual is constructed. In case that the residual res_K doesn't fulfill the convergence criterion $\|res_K\| > \epsilon_{res}$, a new iteration is initiated.

Create C_1 with $\|C_1\| = 1$, $K = 1$	rank	time in s[60]
$\tilde{\sigma}_1 = \mathbf{H}C_1$, $\sigma_1 = \text{Approx}(\tilde{\sigma}_1)$	rank($\mathbf{H}$) = 918	
$\lambda_1 = \langle \sigma_1, C_1 \rangle$, $r_1 = (\sigma_1 - \lambda_1 C_i)$		
while $(\|res_K\| > \epsilon_{res})$		
$\delta = -\left(H - \lambda^K Id\right)^{-1} res_K$		0.30
$\tilde{C}_{K+1} = \prod_{i=1}^{K} (\mathbf{1} - C_i C_i^T)\delta$	rank($\tilde{C}_{K+1}$) = 594	
$C_{K+1} = \text{Approx}(\tilde{C}_{K+1})$	rank(C_{K+1}) = 44	2.62
$C_{K+1}/ = \|C_{K+1}\|$		
$\tilde{\sigma}_{K+1} = \mathbf{H}C_{K+1}$	rank($\tilde{\sigma}_{K+1}$) = 5120	0.99
$\sigma_{K+1} = \text{Approx}(\tilde{\sigma}_{K+1})$	rank(σ_{K+1}) = 109	219.00
$\lambda_{K+1} = \langle \sigma_{K+1}, C_{K+1} \rangle$, $r_{K+1} = (\sigma_{K+1} - \lambda_{K+1} C_{K+1})$		
$K+ = 1$		
for i=1 ... K		
for j=1 ... K		
$A_{i,j} = \langle \sigma_j, C_i \rangle$,		0.01
Diagonalise $\mathbf{A}$, store the lowest eigenvalue λ_K		
and the corresponding eigenvector α		
$res_K = \sum_{i=1}^{K} (\alpha_i \sigma_i - \alpha_i \lambda_K C_i)$		

Figure 3: Workflow of the Davidson algorithm currently used as FCI eigenvalue solver (left) with the ranks (middle) and the timings in seconds (right) of the corresponding steps for the example of $(H_2)_3$ given for iteration 2.

3 Results and Discussion

3.1 Benchmarks

Below we present benchmark calculations of a pilot implementation and investigate the accuracy, the convergence behavior and the scaling of the tensor decomposed FCI algorithm.

The goal of the present algorithm is to achieve high accuracy in the calculation of the energy of the systems – the so called "chemical accuracy". This means that the deviations from the exact energies should lie in the range of $10^{-3}\,\text{E}_h$. In order to set the convergence criterion ϵ_{res} (see Fig. 3) appropriately, the correlation between the accuracy of the energy and $\|res_K\|$ needs to be investigated.

Using a test set of molecules it is confirmed that $\|res_K\|$ is about hundred times larger than the deviation of the energy from the exact value (in E_h) (see Table 1). Therefore, to obtain results in the range of the "chemical accuracy", the iterations should continue until $\|res_K\|$ is smaller than 10^{-1}.

As the parameter ϵ is the only parameter that controls the accuracy in the energy, the influence of this parameter needs to be investigated. For different accuracies in the rank reduction algorithm of $\epsilon = 10^{-3}$ (left) and $\epsilon = 10^{-4}$ (right), the convergence of the energy using the Davidson algorithm is shown in Fig. 4. Note that for this estimation the algorithm was not stopped when the convergence criterion was fulfilled because the maximal achievable accuracy for a given ϵ was to be reached. In both plots of Fig. 4 a rapid convergence towards the lowest eigenvalue in a decomposed tensor format is obtained for all examples that are shown. This is in agreement to the conventional Davidson algorithm. However, in the present algorithm the eigenvalues are obtained only to a certain accuracy. Additional iterations do not improve the accuracy any further. For the examples shown here, the maximal accuracy in the energy is better than the desired "chemical accuracy" of $10^{-3}\,\text{E}_h$. This holds for both criteria $\epsilon = 10^{-3}$ and $\epsilon = 10^{-4}$. A closer look shows that for these examples the largest deviation (in E_h) lies in the range of ϵ, which is applied for the rank reduction. This means that the expected error in the energy for $\epsilon = 10^{-3}$ can lie in the range of $10^{-3}\,\text{E}_h$ and for $\epsilon = 10^{-4}$ in the range of $10^{-4}\,\text{E}_h$. As shown above, the deviation in the energy is about two orders of magnitude larger than the criterion $\|res_K\|$. According to this and the data shown above, a convergence criterion of $\|res_K\| \leq 10^{-1}$ can be expected to be fulfilled in case of $\epsilon = 10^{-3}$ and $\|res_K\| \leq 10^{-2}$ will be fulfilled for $\epsilon = 10^{-4}$.

example	ΔE in E_h at $\|res_K\| \leq 10^{-1}$	ΔE in E_h at $\|res_K\| \leq 10^{-2}$
H_2	0.000001	0.000001
He	0.000584	0.000019
$(H_2)_2$	0.000532	0.000004
He_2	0.000165	0.000007
$(H_2)_3$	0.001535	0.000010
He_3	0.000440	0.000018
LiH	0.001574	0.000010
Be	0.000667	0.000014
H_2O	0.000304	0.000012
H_4	0.000288	0.000004
BeH_2	0.001331	0.000044
HF	0.000532	0.000420

Table 1: Deviation of the energy from the exact eigenvalues ΔE for various convergence thresholds ϵ_{res} for various examples[61] that were calculated without any tensor approximation using the STO-3G basis set [62].

Because the deviation in the energy is mostly influenced by the threshold ϵ in the rank reductions, the errors that can be expected for various values of ϵ are estimated by calculating the energies for a set of molecules. Both, the energies and the deviations to the desired eigenvalues for the test set are given in table 2[11]. All molecules were calculated using the STO-3G basis set[62], and the iterations were stopped when $\|res_K\|$ was smaller than the corresponding threshold estimated above. For $\epsilon = 10^{-3}$, the mean absolute deviation is $1.95 \cdot 10^{-3}\,\mathrm{E}_h$, and the mean square deviation is $2.95 \cdot 10^{-3}\,\mathrm{E}_h$. For $\epsilon = 10^{-4}$, the mean absolute error amounts to $1.73 \cdot 10^{-4}\,\mathrm{E}_h$, and the mean square error amounts to $3.98 \cdot 10^{-4}\,\mathrm{E}_h$. In order to obtain reasonable results, the accuracy in the rank reductions should be set to $\epsilon = 10^{-3}$. However, for smaller deviations, especially smaller than the chemical accuracy, a threshold of $\epsilon = 10^{-4}$ is recommended.

3.2 Scaling Behavior

A decisive variable for the scaling of the algorithm is the rank of the coefficient tensors. Because for each iteration a new coefficient tensor is constructed, each of them possesses its own rank. These ranks are shown in Fig. 5 for the example of a H_2 trimer. Starting from an initial guess of rank one, the ranks of the coefficient tensors seem to increase linearly over the iterations.The ranks become larger as the accuracy of the rank reductions ϵ increases.

To estimate the scaling with the system size of the whole algorithm, the most time consuming step has to be

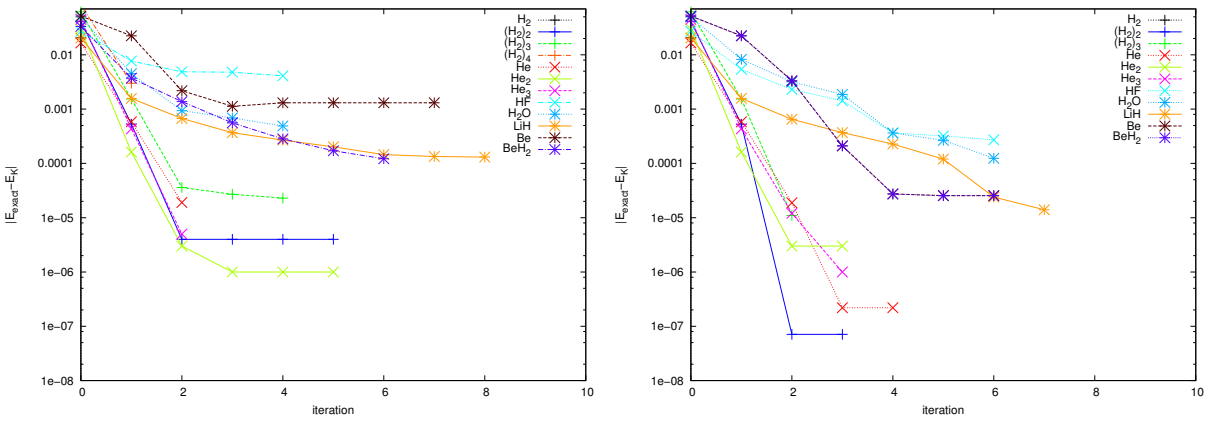


Figure 4: Deviation in the energy from the exact values for various molecules. The accuracy in the rank reduction algorithm is $\epsilon = 10^{-3}$ (left) and $\epsilon = 10^{-4}$ (right).

molecule		$\epsilon = 10^{-3}$, $\|res_K\| \leq 10^{-1}$		$\epsilon = 10^{-4}$, $\|res_K\| \leq 10^{-2}$	
	E	E_{calc}	$\|\Delta E\|$	E_{calc}	$\|\Delta E\|$
LiH	-8.87561	-8.87465	$9.58{\cdot}10^{-4}$	-8.87533	$2.80{\cdot}10^{-4}$
BeH	-16.51306	-16.51093	$2.13{\cdot}10^{-3}$	-16.51297	$8.74{\cdot}10^{-5}$
BH	-26.94869	-26.94624	$2.45{\cdot}10^{-3}$	-26.94836	$6.62{\cdot}10^{-5}$
CH	-40.62149	-40.61768	$3.81{\cdot}10^{-3}$	-40.62147	$2.18{\cdot}10^{-5}$
NH	-57.83139	-57.82793	$3.47{\cdot}10^{-3}$	-57.83097	$4.29{\cdot}10^{-4}$
OH	-78.69955	-78.69414	$5.41{\cdot}10^{-3}$	-78.69915	$3.96{\cdot}10^{-4}$
FH	-103.72787	-103.72024	$7.63{\cdot}10^{-3}$	-103.72647	$1.39{\cdot}10^{-3}$
H_2	-1.84993	-1.84993	$9.87{\cdot}10^{-7}$	-1.84993	$9.87{\cdot}10^{-7}$
$(H_2)_2$	-4.27697	-4.27644	$5.32{\cdot}10^{-4}$	-4.27697	$4.04{\cdot}10^{-6}$
He_2	-6.44588	-6.44571	$1.65{\cdot}10^{-4}$	-6.44587	$6.64{\cdot}10^{-6}$
$(H_2)_3$	-6.98821	-6.98667	$1.53{\cdot}10^{-3}$	-6.98820	$1.02{\cdot}10^{-5}$
He_3	-10.37441	-10.37397	$4.40{\cdot}10^{-4}$	-10.37440	$1.84{\cdot}10^{-5}$
Be	-14.40366	-14.40299	$6.67{\cdot}10^{-4}$	-14.40367	$1.42{\cdot}10^{-5}$
H_2O	-84.20745	-84.20714	$3.04{\cdot}10^{-4}$	-84.20746	$1.24{\cdot}10^{-5}$
H_4	-4.64653	-4.64624	$2.88{\cdot}10^{-4}$	-4.64652	$4.11{\cdot}10^{-6}$
BeH_2	-19.08012	-19.07879	$1.33{\cdot}10^{-3}$	-19.08008	$4.38{\cdot}10^{-5}$

Table 2: Energies E_{calc} and deviations $|\Delta E|$ to the desired energies E for various molecules [63] calculated using the STO-3G basis set [62].

considered. This step is the rank reduction of $\tilde{\sigma}$ after the contraction $\tilde{\sigma} = \mathbf{H}c$ (eqs. 46 and 47, Fig. 3). As mentioned in section 2.6, the rank reduction algorithm scales with $\sim \mathcal{O}(r_{in} \cdot r_{out}^2 \cdot n)$. The initial rank r_{in} results from the contraction $\tilde{\sigma} = \mathbf{H}C_K$ and scales as $r_{in} = K \cdot J$. K is the rank of the Hamiltonian tensor which scales with $\sim \mathcal{O}(N^4)$, as can be seen in eq. 35, and N is the number of electrons, which is assumed to be proportional to the number of basis functions n.

As the rank J of the coefficient tensor is not known in advance, calculations of a chain of H_2 molecules with a bond distance of 0.7425 Å that are separated by a distance of 3 Å are performed, for an optimistic estimate. The threshold for the rank reduction was set to $\epsilon = 10^{-3}$ and the convergence criterion was set to $\epsilon_{res} = 10^{-1}$ to reach the chemical accuracy. The length of the chain increases up to 4 molecules. Because there are several tensors used to represent the coefficient tensors that each possess a different rank, the overall rank needs to be determined for each system. This is done by setting up the linear expansion of the coefficient tensor $\tilde{C}$ that is obtained from the Davidson algorithm and performing a successive rank reduction to obtain C, which corresponds to the decomposed expression of the FCI solution (eqs. 3 and 9). The rank of C is used to investigate the scaling behavior with the system size.

The ranks of the tensors C with increasing chain length are shown in Fig. 6. The data points are fitted by a curve of the function $f(x) = ax^b$ leading to the values $a = 0.055 \pm 0.007$ and $b = 3.75 \pm 0.06$. Here, $f(x)$ corresponds to rank(C) and x corresponds to N. This leads to a scaling of the rank of the C tensor of $\sim \mathcal{O}(N^4)$. Therefore, the FCI wave function can be represented using N^5 parameters. Hence, the initial rank of the most time consuming rank reduction step scales $\sim \mathcal{O}(N^4 \cdot N^4) = \mathcal{O}(N^8)$, which corresponds to the scaling of the rank of the $\tilde{\sigma}$ tensor. Furthermore, the rank of the σ tensor, which corresponds to r_{out}, has the same order of magnitude as the ranks of the C tensor. This leads to an overall scaling of the FCI algorithm of $\sim \mathcal{O}(N^8 \cdot (N^4)^2 \cdot N) \approx \mathcal{O}(N^{17})$.

Although the required storage shows a scaling of only $\sim \mathcal{O}(N^5)$ for both the coefficient and Hamiltonian tensors, the overall scaling of the algorithm is much steeper but still subexponential. As an example, the largest system calculated with the present algorithm is $(H_2)_4$, where rank(C_1)=134 and rank($\mathbf{H}$)=2912. While in conventional FCI-Algorithms the coefficient vector consists of 12,870 elements, there are only 4288 elements to be stored in the CP format. Nevertheless, much larger ranks occur during the iterations, which need to be reduced further, e.g. rank($\tilde{\sigma}_1$)=42,552 (Note that without a prescreening the rank would be 390,208.). At the current state of the development of the reduction algorithm, the rank reduction of tensors with such large ranks is very time consuming. Therefore, this approach is only feasible for small systems and further improvements of the rank reduction algorithm are required to reduce the total scaling. In comparison to conventional algorithms, the effort of the contraction $\sigma = \mathbf{H}C$ is shifted to the rank reduction of the σ tensor in the present algorithm.

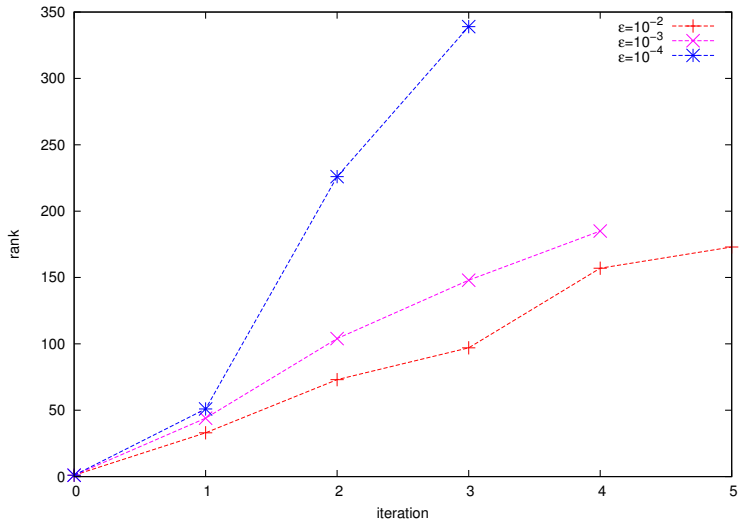


Figure 5: Ranks of the coefficient tensors for the example of a H_2 trimer that was calculated using the STO-3G basis set [62].

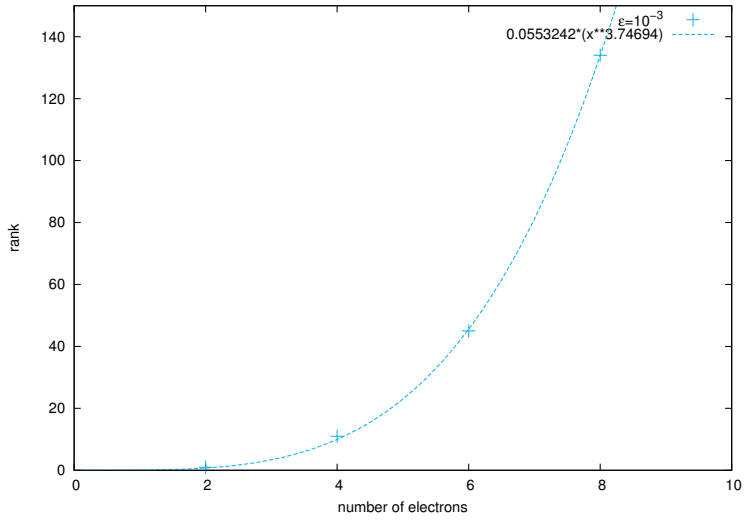


Figure 6: Ranks of the coefficient tensors of chains of $n = 1, 2, 3, 4$ H_2 molecules calculated using the STO-3G basis set and $\epsilon = 10^{-3}$.

4 Conclusions

In the present proof of principle study, a tensor representation of the Fock space is used to approximate the FCI problem in subexponential effort. The coefficient tensor and the Hamiltonian tensor are represented in a canonical product based on an occupation number vector formulation. A modified Davidson algorithm is used to solve the FCI problem in the decomposed format. Both, the scaling behavior and the robustness of the algorithm are tested in a series of benchmark calculations. The number of elements in the representation of the Hamiltonian tensor and the coefficient tensor scale subexponentially ($\sim \mathcal{O}(N^5)$). Note that in the present work, the representation of the Hamiltonian tensor is exact. Benchmark calculations show that the error in the energy in E_h lies in the range of ϵ and hence, the desired "chemical accuracy" can be achieved by choosing an value of $\epsilon \simeq 10^{-4}$. The algorithm converges rapidly and in most of our test cases in less than five iterations. Because of these properties, the algorithm is a robust approximation, and its accuracy is controllable via the single parameter ϵ.

Although the memory required for the Hamiltonian and coefficient tensors scales with $\sim \mathcal{O}(N^5)$, the overall scaling of the algorithm is much steeper. This is because the rank increases after certain operations, especially after the contraction $\mathbf{H}C$ and the increased rank needs to be reduced afterwards. As the rank reduction algorithm scales with the rank to be reduced and the rank after reduction, the actual scaling of the whole algorithm is much steeper ($\sim \mathcal{O}(N^{17})$) than the scaling of the Hamiltonian / coefficient tensor representations. As a consequence, still only small systems can be computed using the present algorithm.

In ongoing work in our group, we intend to make the present algorithm applicable to lager systems by investigating two possible approaches: Either the efficiency of the rank reduction algorithm is improved upon, or the rank reduction is avoided whenever possible. The second approach could include using an initial guess to define an initial tensor space so that fewer iterations are required, or it would optimise the parameters within the tensor representation rather than the wave function parameters. This way, the $\mathbf{H}C$ contraction and the subsequent rank reduction can be avoided, but an accurate guess for C would be required. For such an algorithm, the results of computationally less demanding methods like CCSD, CISD, CIPSI [64, 65] or ICE-CI [11] could provide an initial guess for the coefficient tensors. This way, a hybrid method could be constructed in analogy to SQMC that is a hybrid for QMCFCI[66].

5 Acknowledgements

We like to thank R. Schneider for fruitful discussions and B. van Kuiken and U. Böhm for some helpful comments on the manuscript. A.A.A. and K.-H. Böhm acknowledge financial support by the Max Planck Society and MaxNet Energy research initiative.

References

[1] B. O. ROOS, P. R. TAYLOR, and P. E. SIEGBAHN, *Chem. Phys.* **48**, 157 (1980).

[2] B. ROOS, *Chem. Phys. Lett.* **15**, 153 (1972).

[3] E. ROSSI, G. L. BENDAZZOLI, S. EVANGELISTI, and D. MAYNAU, *Chem. Phys. Lett.* **310**, 530 (1999).

[4] J. OLSEN, P. JØRGENSEN, and J. SIMONS, *Chem. Phys. Lett.* **169**, 463 (1990).

[5] C. W. BAUSCHLICHER and P. R. TAYLOR, *J. Chem. Phys.* **85**, 2779 (1986).

[6] J. C. GREER, *J. Chem. Phys.* **103**, 1821 (1995).

[7] A. O. MITRUSHENKOV, *Chem. Phys. Lett.* **217**, 559 (1994).

[8] D. R. ALCOBA, A. TORRE, L. LAIN, O. B. OÑA, P. CAPUZZI, M. VAN RAEMDONCK, P. BULTINCK, and D. VAN NECK, *J. Chem. Phys.* **141**, (2014).

[9] M. C. TROPAREVSKY and A. FRANCESCHETTI, *J. Phys.: Condens. Matter* **20**, 055211 (2008).

[10] M. SAMBATARO, D. GAMBACURTA, and L. LO MONACO, *Phys. Rev. B* **83**, 045102 (2011).

[11] F. NEESE, Iterative Configuration Expansion Full CI, To be published.

[12] M. D. TOWLER, *physica status solidi (b)* **243**, 2573 (2006).

[13] B. M. AUSTIN, D. Y. ZUBAREV, and J. WILLIAM A. LESTER, *Chem. Rev.* **112**, 263 (2012), PMID: 22196085.

[14] G. H. BOOTH, A. J. W. THOM, and A. ALAVI, *J. Chem. Phys.* **131** (2009).

[15] G. H. BOOTH, D. CLELAND, A. J. W. THOM, and A. ALAVI, *J. Chem. Phys.* **135**, (2011).

[16] J. J. SHEPHERD, G. E. SCUSERIA, and J. S. SPENCER, *Phys. Rev. B* **90**, 155130 (2014).

[17] S. R. WHITE, *Phys. Rev. B* **48**, 10345 (1993).

[18] G. K. L. CHAN, J. J. DORANDO, D. GHOSH, J. HACHMANN, E. NEUSCAMMAN, H. WANG, and T. YANAI, *Frontiers In Quantum Systems In Chemistry and Physics* **18**, 49 (2008).

[19] O. LEGEZA, J. RÖDER, and B. A. HESS, *Phys. Eev. B* **67**, 125114 (2003).

[20] G. MORITZ, B. A. HESS, and M. REIHER, *J. Chem. Phys.* **122**, (2005).

[21] G. K.-L. CHAN and M. HEAD-GORDON, *J. Chem. Phys.* **116**, 4462 (2002).

[22] V. MURG, F. VERSTRAETE, O. LEGEZA, and R. M. NOACK, *Phys. Rev. B* **82**, 205105 (2010).

[23] V. MURG, F. VERSTRAETE, R. SCHNEIDER, P. R. NAGY, and O. LEGEZA, *J. Chem. Theory Comput.* **11**, 1027 (2015), PMID: 25844072.

[24] F. R. MANBY, *J. Chem. Phys.* **119**, 4607 (2003).

[25] M. SCHÜTZ and F. R. MANBY, *Phys. Chem. Chem. Phys.* **5**, 3349 (2003).

[26] H. J. WERNER, F. R. MANBY, and P. J. KNOWLES, *J. Chem. Phys.* **118**, 8149 (2003).

[27] R. AHLRICHS, *Phys. Chem. Chem. Phys.* **6**, 5119 (2004).

[28] M. FEYEREISEN, G. FITZGERALD, and A. KOMORNICKI, *Chem. Phys. Lett.* **208**, 359 (1993).

[29] R. KENDALL and H. FRÜCHTL, *Theor. Chem. Acc.* **97**, 158 (1997).

[30] O. VAHTRAS, J. ALMLÖF, and M. FEYEREISEN, *Chem. Phys. Lett.* **213**, 514 (1993).

[31] N. H. F. BEEBE and J. LINDERBERG, *Int. J. Quantum Chem.* **12**, 683 (1977).

[32] H. KOCH, A. S. DE MERÁS, and T. B. PEDERSEN, *J. Chem. Phys.* **118**, 9481 (2003).

[33] I. RØEGGEN and E. WISLØFF-NILSSEN, *Chem. Phys. Lett.* **132**, 154 (1986).

[34] J. CARROLL and J.-J. CHANG, *Psychometrika* **35**, 283 (1970).

[35] R. A. HARSHMAN and M. E. LUNDY, *Comput. Stat. Data Anal.* **18**, 39 (1994).

[36] E. G. HOHENSTEIN, R. M. PARRISH, and T. J. MARTÍNEZ, *J. Chem. Phys.* **137**, 044103 (2012).

[37] E. G. HOHENSTEIN, R. M. PARRISH, C. D. SHERRILL, and T. J. MARTÍNEZ, *J. Chem. Phys.* **137**, 221101 (2012).

[38] E. G. HOHENSTEIN, S. I. L. KOKKILA, R. M. PARRISH, and T. J. MARTÍNEZ, *J. Chem. Phys.* **138**, 124111 (2013).

[39] R. M. PARRISH, C. D. SHERRILL, E. G. HOHENSTEIN, S. I. L. KOKKILA, and T. J. MARTÍNEZ, *J. Chem. Phys.* **140**, (2014).

[40] F. A. BISCHOFF and E. F. VALEEV, *J. Chem. Phys.* **134**, 104104 (2011).

[41] F. A. BISCHOFF and E. F. VALEEV, *J. Chem. Phys.* **139**, (2013).

[42] W. UEMURA and O. SUGINO, *Phys. Rev. Lett.* **109** (2012).

[43] U. BENEDIKT, A. A. AUER, M. ESPIG, and W. HACKBUSCH, *J. Chem. Phys.* **134**, 054118 (2011).

[44] U. BENEDIKT, K.-H. BÖHM, and A. A. AUER, *J. Chem. Phys.* **139**, (2013).

[45] U. BENEDIKT, *Low-Rank Tensor Approximation in post Hartree-Fock Methods*, PhD thesis, TU Chemnitz, 2013.

[46] U. BENEDIKT, H. AUER, W. HACKBUSCH, and A. AUER, *Mol. Phys.* , 1 (2013).

[47] T. HELGAKER, P. JØRGENSON, and J. OLSEN, *Molecular Electronic Structure Theory*, Wiley and Sons, 1. edition, 2000.

[48] K.-H. BÖHM, PhD thesis, TU Chemnitz, 2016.

[49] O. LEGEZA, T. ROHWEDDER, R. SCHNEIDER, and S. SZALAY, Tensor Product Approximation (DMRG) and Coupled Cluster Method in Quantum Chemistry, in *Many-Electron Approaches in Physics, Chemistry and Mathematics*, edited by V. BACH and L. DELLE SITE, Mathematical Physics Studies, pp. 53–76, Springer International Publishing, 2014.

[50] S. SZALAY, M. PFEFFER, V. MURG, G. BARCZA, F. VERSTRAETE, R. SCHNEIDER, and O. LEGEZA, *Int. J. Quantum Chem* **115**, 1342 (2015).

[51] G. BEYLKIN and M. J. MOHLENKAMP, *P. Natl. Acad. Sci. USA* **99**, 10246 (2002).

[52] G. BEYLKIN and M. J. MOHLENKAMP, *SIAM J. Sci. Comput.* **26**, 2133 (2005).

[53] T. G. KOLDA and B. W. BADER, *Siam Review* **51**, 455 (2009).

[54] E. R. DAVIDSON, *J. Comp. Phys.* **17**, 87 (1975).

[55] D. BRAESS and W. HACKBUSCH, *IMA J. Numer. Anal.* **25**, 685 (2005).

[56] Especially in high dimensions it is a good advise to balance the representation system of every tensor which is stored in real computer implementations. Otherwise, one cannot avoid a number overflow in some entries of the representation system. For example consider the rank-one tensor $v := \bigotimes_{\mu=1}^{n} v_\mu \in \bigotimes_{\mu=1}^{n} \mathbb{R}^t$ with $(v_\mu)_l = 1$ for all $l \in \{1, \dots, t\}$. If one introduces the following constrains for the representation system: $\|v_\mu\| = 1$ for all $\mu \in \{2, \dots, d\}$, we have $\|v_1\| = t^n$. If n and t are large enough, we will produce a number overflow on every computer systems. With (55) we have a balanced and solid representation system. Furthermore, with (55) we avoid unnecessary scaling influences in our objective function and consequently their second derivative with respect to the representation system.

[57] M. ESPIG, *Effiziente Bestapproximation mittels Summen von Elementartensoren in hohen Dimensionen*, PhD thesis, Universität Leipzig, 2008.

[58] M. ESPIG and W. HACKBUSCH, *Numer. Math.* **122**, 489 (2012).

[59] M. ESPIG, *in preparation* (2016).

[60] Calculation done on one node with an Intel® Xenon® CPU X5650 @ 2,67 GHz.

[61] The geometries are: H_2 r_{H-H}=0.7425 Å, LiH r_{Li-H}=1.5950 Å, H_2O r_{O-H}=0.9572 Å, θ_{HOH}=104.52°, H_4 square geometry r_{H-H}=1.0584 Å, BeH_2 r_{H-H}=1.2905 Å θ_{HBeH}=180.00°, HF r_{F-H}=0.9327 Å. Intermolecular distance is always 3.0 Å.

[62] W. J. HEHRE, R. F. STEWART, and J. A. POPLE, *J. Chem. Phys.* **51**, 2657 (1969).

[63] The geometries are: LiH r_{Li-H}=1.5983 Å, BeH r_{Be-H}=1.3601 Å, BH r_{B-H}=1.2372 Å, CH r_{C-H}=1.1301 Å, NH r_{N-H}=1.0447 Å, OH r_{O-H}=0.9821 Å, FH r_{F-H}=0.9285 Å, H_2 r_{H-H}=0.7425 Å, H_2O r_{O-H}=0.9572 Å, θ_{HOH}=104.52°, H_4 square geometry r_{H-H}=1.0584 Å, BeH_2 r_{Be-H}=1.2905 Å θ_{HBeH}=180.00°. Intermolecular distance is always 3.0 Å.

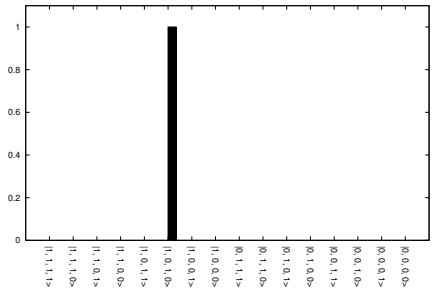
[64] B. HURON, J. MALRIEU, and P. RANCUREL, *J. Chem. Phys.* **58**, 5745 (1973).

[65] S. EVANGELISTI, J. DAUDEY, and J. MALRIEU, *Chem. Phys.* **75**, 91 (1983).

[66] F. R. PETRUZIELO, A. A. HOLMES, H. J. CHANGLANI, M. P. NIGHTINGALE, and C. J. UMRIGAR, *Phys. Rev. Lett.* **109**, 230201 (2012).

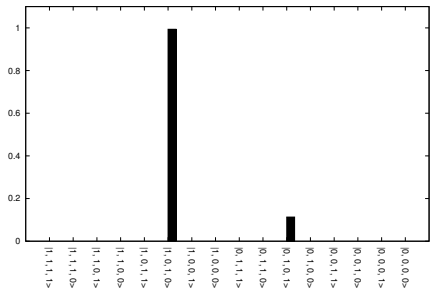
HF-Solution of H_2

$$
\begin{aligned}
|\Psi\rangle &= 1 \cdot |1,0,1,0\rangle \\
\mathfrak{C} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\end{aligned}
$$



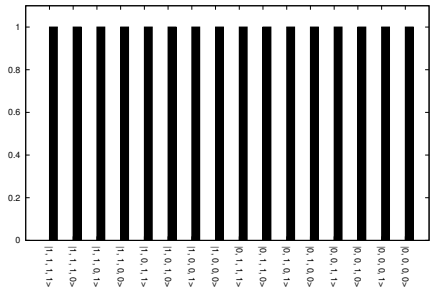
FCI-Solution of H_2

$$
\begin{aligned}
|\Psi\rangle &= 0.99 \cdot |1,0,1,0\rangle - 0.11 \cdot |0,1,0,1\rangle \\
\mathfrak{C} &= \begin{pmatrix} 0 \\ 0.99 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&+ \begin{pmatrix} -0.11 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}
$$



Equal Coefficients for all Determinants in the Fock-Space

$$
\begin{aligned}
|\Psi\rangle &= \sum_{k_i} \sum_{k_a} \sum_{k_{\bar{i}}} \sum_{k_{\bar{a}}} 1 \cdot |k_i, k_a, k_{\bar{i}}, k_{\bar{a}}\rangle \\
\mathfrak{C} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix}
\end{aligned}
$$



Equal Coefficients for all two-electron Determinants

$$
\begin{aligned}
|\Psi\rangle &= 1 \cdot |1,0,1,0\rangle + 1 \cdot |0,1,0,1\rangle + 1 \cdot |0,1,1,0\rangle \\
&= 1 \cdot |1,0,0,1\rangle + 1 \cdot |1,1,0,0\rangle + 1 \cdot |0,0,1,1\rangle \\
\mathfrak{C} &= \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ 1.22 \end{pmatrix} \\
&+ \begin{pmatrix} 0.58 \\ -1.22 \end{pmatrix} \otimes \begin{pmatrix} -0.58 \\ 1.22 \end{pmatrix} \otimes \begin{pmatrix} 0.58 \\ -1.22 \end{pmatrix} \otimes \begin{pmatrix} -0.58 \\ 1.22 \end{pmatrix} \\
&+ \begin{pmatrix} 0 \\ 1.45 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ -1.45 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ -1.45 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ -1.45 \end{pmatrix} \\
&+ \begin{pmatrix} 0.69 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} -0.69 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} -0.69 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} -0.69 \\ -0 \end{pmatrix}
\end{aligned}
$$

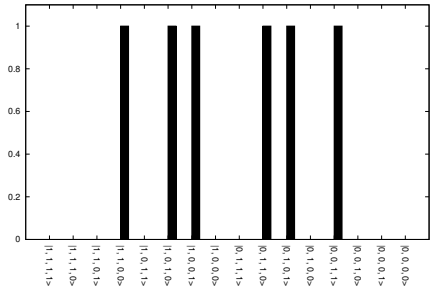


Figure 1: Examples of coefficient tensor representations of the four orbital, two electron example. Shown are an explicit form of the wave function using ONVs, the representation of the coefficient tensor, and a bar plot of the absolute values of the coefficients.

1: Choose the initial guess as described in [57, 58] and define $k := 1$.
2: **while** Stop Condition **do**
3: $\mu := \text{argmax}_{1\leq\nu\leq d} \left\| \frac{\partial f_\lambda}{\partial \xi_\nu}(\xi_{k,\nu}) \right\|_\infty$
4:

$$\begin{aligned}
\xi_{k,\mu} &:= (\xi_1^{k+1}, \ldots, \xi_{\mu-1}^{k+1}, \xi_{\mu+1}^{k}, \ldots, \xi_d^k) \\
d_\mu^k &:= \frac{\partial f_\lambda}{\partial \xi_\nu}(\xi_{k,\nu}) \\
\lambda_\mu^k &:= \frac{\langle d_\mu^k, d_\mu^k \rangle}{\langle G_{k,\mu} d_\mu^k, d_\mu^k \rangle} \\
\xi_\mu^{k+1} &:= \xi_\mu^k - \lambda_\mu^k d_\mu^k \\
\xi_{k,\mu+1} &:= (\xi_1^{k+1}, \ldots, \xi_{\mu}^{k+1}, \xi_{\mu+1}^{k}, \ldots, \xi_d^k)
\end{aligned}$$

5: balance the representation system $\xi_{k,\mu+1}$ such that Eq. (55) is satisfied
6: $k \mapsto k+1$.
7: **end while**

Figure 2: Pivotised Alternating Steepest Descent Algorithm (PASD) method.

Create C_1 with $\|C_1\| = 1$, $K = 1$		rank	time in s[60]
$\tilde{\sigma}_1 = \mathbf{H}C_1$, $\sigma_1 = \text{Approx}(\tilde{\sigma}_1)$		rank($\mathbf{H}$) = 918	
$\lambda_1 = \langle \sigma_1, C_1 \rangle$, $r_1 = (\sigma_1 - \lambda_1 C_i)$			
while	$(\|res_K\| > \epsilon_{res})$		
	$\delta = -\left(H - \lambda^K Id\right)^{-1} res_K$		0.30
	$\tilde{C}_{K+1} = \prod_{i=1}^{K} (\mathbf{1} - C_i C_i^T)\delta$	rank($\tilde{C}_{K+1}$) = 594	
	$C_{K+1} = \text{Approx}(\tilde{C}_{K+1})$	rank(C_{K+1}) = 44	2.62
	$C_{K+1}/ = \|C_{K+1}\|$		
	$\tilde{\sigma}_{K+1} = \mathbf{H}C_{K+1}$	rank($\tilde{\sigma}_{K+1}$) = 5120	0.99
	$\sigma_{K+1} = \text{Approx}(\tilde{\sigma}_{K+1})$	rank(σ_{K+1}) = 109	219.00
	$\lambda_{K+1} = \langle \sigma_{K+1}, C_{K+1} \rangle \; r_{K+1} = (\sigma_{K+1} - \lambda_{K+1} C_{K+1})$		
	$K+ = 1$		
	for i=1 ... K		
	for j=1 ... K		
	$A_{i,j} = \langle \sigma_j, C_i \rangle$,		0.01
	Diagonalise $\mathbf{A}$, store the lowest eigenvalue λ_K		
	and the corresponding eigenvector α		
	$res_K = \sum_{i=1}^{K} (\alpha_i \sigma_i - \alpha_i \lambda_K C_i)$		

Figure 3: Workflow of the Davidson algorithm currently used as FCI eigenvalue solver (left) with the ranks (middle) and the timings in seconds (right) of the corresponding steps for the example of $(H_2)_3$ given for iteration 2.

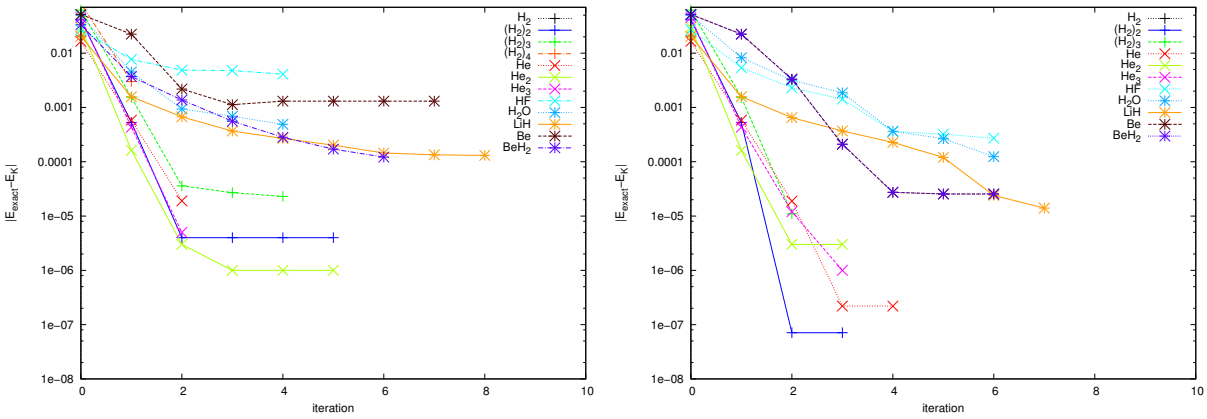


Figure 4: Deviation in the energy from the exact values for various molecules. The accuracy in the rank reduction algorithm is $\epsilon = 10^{-3}$ (left) and $\epsilon = 10^{-4}$ (right).

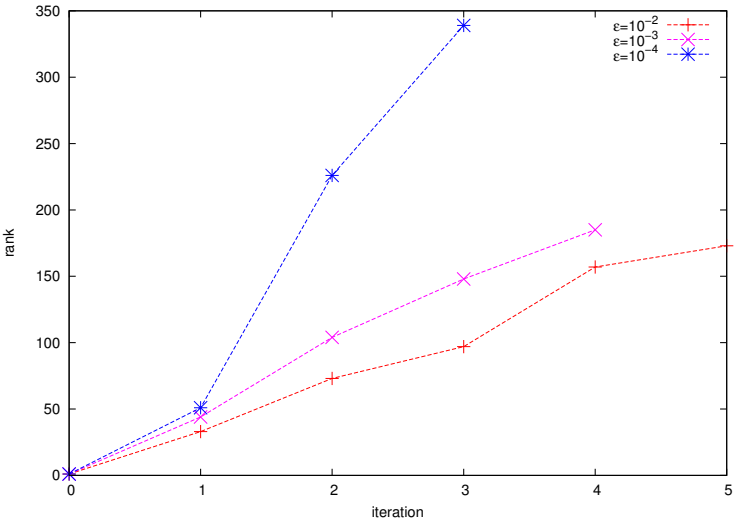


Figure 5: Ranks of the coefficient tensors for the example of a H_2 trimer that was calculated using the STO-3G basis set [62].

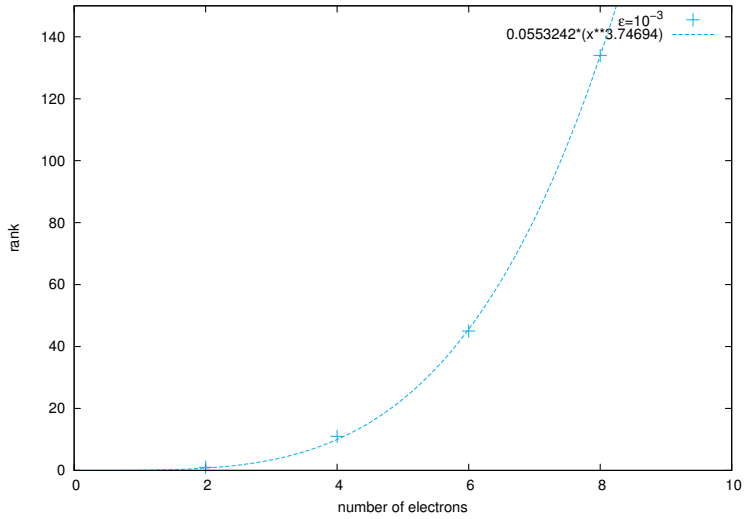


Figure 6: Ranks of the coefficient tensors of chains of $n = 1, 2, 3, 4$ H_2 molecules calculated using the STO-3G basis set and $\epsilon = 10^{-3}$.