Hierarchical Matrices in Computations of Electron Dynamics[*]

Othmar Koch[†]   Christopher Ede   Gerald Jordan

Armin Scrinzi[‡]

July 20, 2009

Abstract

We discuss the approximation of the meanfield terms appearing in computations of the multi-configuration time-dependent Hartree–Fock method for the solution of the time-dependent multi-particle (electronic) Schrödinger equation by hierarchical matrices. We give theoretical error bounds for the *cross approximation* defined by low rank approximations of admissible matrix sub-blocks, and illustrate the gain in performance by numerical experiments.

Keywords: Multi-configuration time-dependent Hartree–Fock method, time-dependent multi-particle Schrödinger equation, hierarchical matrices.

MSC: 65M20, 65F30, 65Z05.

1   INTRODUCTION

In this paper we discuss numerical aspects of the multi-configuration time-dependent Hartree–Fock method (MCTDHF) for the approximate solution of the time-dependent Schrödinger equation

$$\mathrm{i}\frac{\partial \psi}{\partial t} = H\psi, \tag{1.1}$$

where the complex-valued *wave function* $\psi = \psi(x^{(1)}, \ldots, x^{(f)}, t)$ explicitly depends on time t and, in the case considered here, the positions

1

$x^{(1)}, \ldots, x^{(f)} \in \mathbb{R}^3$ of electrons in an atom or molecule. The Hamiltonian H is time-dependent and has the form

$$H(t) := \sum_{k=1}^{f} \left(\frac{1}{2} \left(-\mathrm{i}\nabla^{(k)} + A(t) \right)^2 + U(x^{(k)}) + \sum_{l<k} V(x^{(k)}, x^{(l)}) \right), \quad (1.2)$$

where

$$U(x) := -\frac{Z}{|x|}, \quad Z \in \mathbb{N}, \tag{1.3}$$

$$V(x,y) := \frac{1}{|x-y|}, \tag{1.4}$$

$$A(t) := (a_1(t), a_2(t), a_3(t)). \tag{1.5}$$

$A(t)$ is a smooth (vector-valued) function of t modeling an ultrafast laser pulse, and $\nabla^{(k)}$ is the nabla operator w. r. t. $x^{(k)}$ only. We will subsequently briefly denote

$$T(t) := \sum_{k=1}^{f} \left(\frac{1}{2} \left(-\mathrm{i}\nabla^{(k)} + A(t) \right)^2 + U(x^{(k)}) \right).$$

2   THE MCTDHF METHOD

In MCTDHF as put forward in [6, 15, 16], the multi-electron wave function ψ from (1.1) is approximated by a function satisfying the ansatz

$$u = \sum_{(j_1, \ldots, j_f)} a_{j_1, \ldots, j_f}(t) \phi_{j_1}(x^{(1)}, t) \cdots \phi_{j_f}(x^{(f)}, t) =: \sum_J a_J(t) \Phi_J(x, t). \tag{2.1}$$

Using (2.1) for the *electronic* Schrödinger equation, the Pauli principle implies that only solutions u are considered which are antisymmetric under exchange of any pair of arguments $x^{(j)}$, $x^{(k)}$. This assumption is particular to the MCTDHF approach, as compared to the multi-configuration time-dependent Hartree method (MCTDH) proposed in [3, 14] for quantum molecular dynamics. Antisymmetry reduces the number of equations considerably. Particularly, the assumption implies antisymmetry in the coefficients a_J. Formally, multi-indices $J = (j_1, \ldots, j_f)$ vary for $j_k = 1, \ldots, N$, $k = 1, \ldots, f$. Due to the simplifications resulting from the antisymmetry assumption, only $\binom{N}{f}$ equations for a_J have to be solved in the actual computations, however.

The Dirac–Frenkel variational principle [7, 8] is used to derive differential equations for the coefficients a_J and the *single-particle functions* ϕ_j.

This yields equations of motion [3] for the coefficients and single-particle functions in (2.1)

$$\mathrm{i}\frac{da_J}{dt} = \sum_K \langle \Phi_J \,|V|\, \Phi_K \rangle\, a_K, \quad \forall J, \tag{2.2}$$

$$\mathrm{i}\frac{\partial \phi_j}{\partial t} = T(t)\phi_j + (I - P)\sum_{k=1}^N \sum_{l=1}^N \rho_{j,l}^{-1}\overline{V}_{l,k}\phi_k, \quad j = 1, \ldots, N, \tag{2.3}$$

where

$$\psi_j := \langle \phi_j | u \rangle, \quad j = 1, \ldots, N, \tag{2.4}$$
$$\rho_{j,l} := \langle \psi_j | \psi_l \rangle, \quad j, l = 1, \ldots, N, \tag{2.5}$$
$$\overline{V}_{j,l} := \langle \psi_j \,|V|\, \psi_l \rangle, \quad j, l = 1, \ldots, N, \tag{2.6}$$

and P is the orthogonal projector onto the space spanned by the functions ϕ_j. Integrals defining the inner products $\langle \cdot | \cdot \rangle$ are with respect to all variables present in both arguments.

3   EFFICIENT COMPUTATION

To make the numerical solution of the equations of motion (2.2), (2.3) computationally tractable, some care is required in the evaluation of the right-hand side of the differential equations. The computationally most demanding part is represented by the evaluation of the meanfield operators (2.6). This becomes clear when we realize that the computations involve the evaluation of integrals of the form[1]

$$\langle \phi_1(x) \,|V(x,y)|\, \phi_2(x) \rangle_{L^2(x)}\, \phi_3(y), \tag{3.1}$$

$$\left\langle \phi_1(x)\tilde{\phi}_1(y) \,|V(x,y)|\, \phi_2(x)\tilde{\phi}_2(y) \right\rangle_{L^2(x,y)}, \tag{3.2}$$

where ϕ_i, $\tilde{\phi}_i$ are any single-particle functions from (2.1). To reduce the computational effort necessary for the evaluation of these terms, a procedure based on discretization and low rank approximation is proposed in [6] to evaluate the integrals. In [12], bounds for the error introduced by this approximation are given. In fact, the analysis deals with terms of the form

$$\langle \phi_1(x)|V(x,y)|\phi_2(y)\rangle_{L^2(x,y)} = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \overline{\phi_1(x)}V(x,y)\phi_2(y)\,dy\,dx, \tag{3.3}$$

---

[1]Subscripts of inner products $\langle \cdot | \cdot \rangle$ refer to the respective integration variables, and likewise for norms $\| \cdot \|$. We will drop the subscripts where the arguments are clear.

which are however representative for the actually appearing quantities, see [13] for an exhaustive treatment of extensions of our estimates.

In the present paper, we are going to extend the analysis of [12] to the case where the meanfield operators are alternatively approximated by $\mathcal{H}$-matrices. These were first introduced in [10, 11], and compress the data represented by a matrix of a particular structure by using *cluster trees*. Storage requirements and complexity of arithmetic operations involving $\mathcal{H}$-matrices scale *almost linearly* with the dimension (that is, linearly with logarithmic terms) [9]. Appendix A gives a short description of the most important features of $\mathcal{H}$-matrices we are going to use, see also [13].

Following [6], we choose a discretization in terms of a set of basis functions $\mathcal{B} = \{|i\rangle = i(x) : i = 1, \ldots, L\}$ and approximate V by

$$V \approx V_{\text{app}} = RVR, \tag{3.4}$$

where R is the orthogonal projection onto the subspace B spanned by $\mathcal{B}$. V here is short for the operator on L^2 with integral kernel $V(x, y)$. As basis functions we use real-valued functions with finite support defined on the spatial grid used to solve (2.2), (2.3). In general, it is natural to use polynomial finite elements on a suitable subdivision of the spatial domain. The details may vary from case to case. Here, we want to give error bounds under the assumption that the space partition is sensible and the choice of finite elements corresponds with the task at hand. The notions used in the subsequent presentation are explained for example in [1, 4, 5].

We consider for $\mathcal{B}$ the *nodal basis* for globally continuous finite elements consisting of piecewise polynomials of degree $\leq m - 1$. The resulting finite elements are *conforming*, i.e., $B \subseteq H^1$, where we denote by H^m the Sobolev space of functions which are square integrable together with their first m weak derivatives. For the theoretical analysis we will always assume that our spatial subdivision is *quasi-uniform*, where the diameters of the elements are proportional to $h = 1/L$, also referred to as *mesh width*. In our numerical experiments in Section 4, we use underlying cylindrical coordinates, where this assumption does not hold uniformly, but nonetheless the results are along the line of our theoretical analysis.

Now, the projection R used in (3.4) is characterized by

$$R = \sum_{i,j=1}^{L} |i\rangle [Q^{-1}]_{i,j} \langle j|, \tag{3.5}$$

with the *mass matrix* Q given as

$$Q_{i,j} = \langle i|j\rangle. \tag{3.6}$$

Obviously, Q is real, symmetric and nonsingular, and thus Q^{-1} is also symmetric. Moreover, we have $\|Q^{-1}\|_2 = O(1/h)$ for the spectral norm $\|\cdot\|_2$ of the matrix Q^{-1}.

In [12], a bound

$$|\langle\phi_1(x)|V(x,y)-V_{\text{app}}(x,y)|\phi_2(y)\rangle_{L^2(x,y)}| = \begin{cases} O(h^m) & \text{for } V \text{ bounded,} \\ O(h^{m-1}) & \text{for } V \text{ Coulomb} \end{cases}$$
(3.7)

is shown for the discretization error if $\phi_1,\ \phi_2 \in H^m$.

Subsequently, the discrete operator V_{app} is approximated by a suitable operator whose application is computationally cheap. The global approximation put forward in [6] is analyzed in [12]. Here, we propose an approximation by a hierarchical matrix instead.

First, we rewrite (3.4) as

$$V_{\text{app}} = \sum_{i,j=1}^{L}\sum_{i',j'=1}^{L} |i\rangle[Q^{-1}]_{i,i'}\tilde{V}_{i',j'}[Q^{-1}]_{j',j}\langle j| = \sum_{i,j=1}^{L} |i\rangle[Q^{-1}\tilde{V}Q^{-1}]_{i,j}\langle j|$$

with

$$\tilde{V}_{i,j} = \int\int i(x)V(x,y)j(y)\,dx\,dy.$$

Obviously, the matrix $\tilde{V}$ is symmetric. We now rewrite $\tilde{V}$ as follows: define a matrix S by

$$S_{i,j} := \int i(x)g(x)j(x)\,dx$$

with a real function $g > 0$ which is large in regions which shall be emphasized, usually near the core. Obviously, S is symmetric, and it is easy to see that S is positive definite. Note that $\|\tilde{V}\|_2$, $\|S\|_2 = O(h)$, since $\|\tilde{V}\|_2 \le \sqrt{\|\tilde{V}\|_1\|\tilde{V}\|_\infty}$ and the row sums and the column sums are $O(h)$.

In [6, 12], the Cholesky decomposition

$$S = C^T C,$$

is used to compute a low rank approximation of $\tilde{V}$, where the cut-off parameter can be chosen independently of the discretization parameter h, see [12]. This is realized by resorting to the transformed standard eigenvalue problem

$$C^{-T}\tilde{V}C^{-1}\tilde{u} = \lambda\tilde{u}, \tag{3.8}$$

with $\tilde{u} := Cu$. Discarding eigenvalues smaller than a tolerance ε, finally leads to an approximation V_{low} satisfying an error bound

$$|\langle\phi_1(x)|V_{\text{app}}(x,y) - V_{\text{low}}(x,y)|\phi_2(y)\rangle_{L^2(x,y)}| = O(\varepsilon). \tag{3.9}$$

This corresponds to a (global) low rank approximation of V_{app}. By computing this low rank approximation via (3.8), we obtained an approximation error which is independent of the mesh width.

In order to apply this concept also in the context of hierarchical matrices, we adapt the idea of (3.8) for the approximation of the respective matrix blocks $\tilde{V}_{\mathcal{I}\times\mathcal{J}}$ ($\mathcal{I}$ and $\mathcal{J}$ are suitable subsets of the index set, and $\tilde{V}_{\mathcal{I}\times\mathcal{J}}$ the corresponding submatrix of $\tilde{V}$) to obtain an $\mathcal{H}$-matrix approximation $\tilde{V}_{\mathcal{H},\varepsilon}$. Thus, for an *admissible* block [9] $\tilde{V}_{\mathcal{I}\times\mathcal{J}}$ with corresponding diagonal blocks $S_{\mathcal{I}\times\mathcal{I}}$ and $S_{\mathcal{J}\times\mathcal{J}}$ of S, we compute the singular value decomposition of

$$C_{\mathcal{I}\times\mathcal{I}}^{-T}\tilde{V}_{\mathcal{I}\times\mathcal{J}}C_{\mathcal{J}\times\mathcal{J}}^{-1}.$$

Using a cutoff parameter ε to define the low rank approximation of each block, it is found that the overall error committed in the approximation of $\tilde{V}$ results in

$$\|\tilde{V}-\tilde{V}_{\mathcal{H},\varepsilon}\|_2 = |\ln(h)|O(\varepsilon),$$

see [2, Lemma 2.1, p.409]. Consequently, it is possible to write[2]

$$\tilde{V}-\tilde{V}_{\mathcal{H},\varepsilon} = \hat{U}\Lambda\hat{U}^T,$$

with a unitary matrix $\hat{U}$ and $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_L)$ with $\lambda_j = |\ln(h)|O(\varepsilon h)$. Thus, the error analysis proceeds analogously to [12],

$$|\langle\phi_1(x)|V_{\mathrm{app}}(x,y) - V_{\mathcal{H},\varepsilon}(x,y)|\phi_2(y)\rangle_{L^2(x,y)}| =$$

$$\left|\sum_{i,j=1}^{L}\langle\phi_1|i\rangle\langle j|\phi_2\rangle\sum_{\mu=1}^{L}[Q^{-1}\hat{U}]_{i,\mu}\lambda_\mu[\hat{U}^TQ^{-1}]_{\mu,j}\right|$$

$$\leq \|\phi_1\|\|\phi_2\|\sum_{i,j=1}^{L}\|i\|\|j\|\left|\sum_{\mu=1}^{L}[Q^{-1}\hat{U}]_{i,\mu}\lambda_\mu[\hat{U}^TQ^{-1}]_{\mu,j}\right|$$

$$\leq \mathrm{const.}\,\varepsilon\ln(h)\,h^3\sum_{i,j=1}^{L}\sum_{\mu=1}^{L}|[Q^{-1}\hat{U}]_{i,\mu}||[\hat{U}^TQ^{-1}]_{\mu,j}|$$

$$\leq \mathrm{const.}\,\varepsilon|\ln(h)|h^3\sum_{i,j=1}^{L}[|Q^{-1}\hat{U}||\hat{U}^TQ^{-1}|]_{i,j}$$

$$\leq \mathrm{const.}\,\varepsilon|\ln(h)|h^3L\|Q^{-1}\hat{U}\|_2^2 \leq \mathrm{const.}\,\varepsilon|\ln(h)|h^3L\|Q^{-1}\|_2^2\|\hat{U}\|_2^2$$

$$\leq \mathrm{const.}\,\varepsilon|\ln(h)|,$$

where the absolute value of a matrix, $|A|$, is meant entry-wise. For this estimate, we have used a simple property proven in [12, Lemma 2.1]:

Let $X \in \mathbb{R}^{L\times L}$ be a nonsingular matrix. Then,

$$\sum_{i,j=1}^{L}[|X^T||X|]_{i,j} \leq L\|X\|_2^2.$$

---

[2]Note that by construction, $\tilde{V}_{\mathcal{H},\varepsilon}$ is symmetric, and consequently, $\tilde{V}-\tilde{V}_{\mathcal{H},\varepsilon}$ has a symmetric singular value decomposition.

Thus, the error of our low rank approximation is bounded in terms of the error margin ε, (almost) uniformly in the diameters of the spatial partition. We formulate this fact in the following theorem:

Theorem 3.1 For ϕ_1, $\phi_2 \in L^2$, and an underlying spatial partition at L points distributed quasi-uniformly, the error introduced by the approximation by $\mathcal{H}$-matrices satisfies

$$|\langle \phi_1(x)|V_{\text{app}}(x,y) - V_{\mathcal{H},\varepsilon}(x,y)|\phi_2(y)\rangle_{L^2(x,y)}| = O(\varepsilon \ln(L)), \qquad (3.10)$$

where ε is an error margin which can be freely chosen.

4 NUMERICAL EXPERIMENTS

First, we give a numerical illustration of Theorem 3.1. For the spatial discretization with underlying cylindrical coordinates, we were using polynomial finite elements characterized by quadruples $(n_1, p_1; n_2, p_2)$, with the first pair n_1, p_1 giving the number n_1 of polynomials of degree p_1 for the discretization of the z-coordinate and likewise n_2 and p_2 for the radial coordinate. The resulting matrix size equals $n_1 n_2 \times n_1 n_2$. To prevent too small subdivision of the matrix, a minimal blocksize of 16 was enforced. In Table 1 we give the values of the inner products (3.3) with underlying discretization characterized by $(n_1, p_1; n_2, p_2) = (48, 2; 96, 2)$, approximated by using hierarchical matrices with cut-off parameters ε, where $\phi_1 = \phi_2$ are the eigenfunctions associated with the ground state energy E_0 and the first excited state E_1 of the He$^+$ ion, respectively. The rows give the corresponding values for the "exact" computation, where no low rank approximation of the meanfield integrals is used, and the approximation by hierarchical matrices with cut-off parameter equal to $\varepsilon \in \{10^{-4}, 10^{-5}, 10^{-6}\}$. The number of digits which coincide with the "exact" computations is roughly proportional to the cut-off parameter ε.

ε	E_0	E_1
0	3.692434	3.774842
1e−6	3.692433	3.774841
1e−5	3.692498	3.774953
1e−4	3.693201	3.776187

Table 1: Meanfield integrals (3.3) for the He$^+$ ion.

Next, we demonstrate the gain in efficiency when the MCTDHF code [16] uses model reduction by $\mathcal{H}$-matrices as compared to the global approach [6, 12].

We compute a model for the helium atom under the influence of a laser pulse. Consequently, in (1.2) we set $Z = 2$, $a_1 \equiv 0$, $a_2 \equiv 0$, $a_3(t) = A_0 \cos\left(\frac{\pi t}{2\sigma}\right)$, $t \in [-\sigma, \sigma]$, with $A_0 = 2.964$, $\sigma = 110.32$. The time propagation took place for $t \in [-1, -0.999]$ with 10 steps of a Runge-Kutta-Fehlberg 4(5) integrator on that interval. For MCTDHF, we set $N = 4$, $f = 2$.

Table 2 gives the results of our test runs. For the cut-off parameter $\varepsilon \in \{10^{-4}, 10^{-5}, 10^{-6}\}$, the computation times per evaluation (evaluations of the meanfield terms in both the imaginary time propagation for the computation of the initial state and the 10 steps performed for the time propagation) and numerically obtained values for the energy at the end of the time integration are given for the $\mathcal{H}$-matrix approximations as compared to the global low rank approximation. The top half of Table 2 gives the results for a spatial discretization with $(n_1, p_1; n_2, p_2) = (24, 2; 48, 2)$, while the lower half shows $(n_1, p_1; n_2, p_2) = (48, 2; 96, 2)$. Note that this results in matrix dimensions $N = 1152$ and $N = 4608$, respectively.

Obviously, the computation time is significantly decreased by the use of hierarchical matrices. The scaling appears to be roughly linear with respect to the matrix size for the $\mathcal{H}$-matrix approximation, commensurate with the theory summed up in Appendix A, and scales less favorably in the case of the global low rank approximation by singular value decomposition. Moreover, the approximation quality is indeed controlled by the cut-off parameter ε, the difference to the most accurate of the approximations is related to ε. The dependence is not as clear as in Table 1, which is caused by other algorithmic components influencing the accuracy during time integration.

ε	time global	time $\mathcal{H}$	energy global	energy $\mathcal{H}$
$(n_1, p_1; n_2, p_2) = (24, 2; 48, 2)$				
1e−4	0.3622 sec.	0.0766 sec.	−2.8113067472	−2.8110777972
1e−5	0.3924 sec.	0.1017 sec.	−2.8113067471	−2.8113045538
1e−6	0.3887 sec.	0.1216 sec.	−2.8113067471	−2.8113067403
$(n_1, p_1; n_2, p_2) = (48, 2; 96, 2)$				
1e−4	2.5951 sec.	0.2938 sec.	−2.8612096380	−2.8607521876
1e−5	4.1204 sec.	0.4682 sec.	−2.8610407552	−2.8610453291
1e−6	5.1268 sec.	0.6321 sec.	−2.8610407551	−2.8610406239

Table 2: Timings for global vs. hierarchical meanfield approximations.

A   Hierarchical Matrices

Let $\mathcal{S} = \{1, \ldots, L\}$ be an index set. The elements of $\mathcal{S}$ are called *degrees of freedom*. Associate with each index i a point $x_i \in \mathbb{R}^L$. Non-empty subsets $\sigma \subseteq \mathcal{S}$ are referred to as *clusters* and associated with a bounded domain $\cup \sigma$. The latter is assumed to be contained in an axis oriented *bounding box* B_σ of minimal size.

Definition A.1 $\mathcal{T}$ is a *cluster tree* for the index set $\mathcal{S}$, if for a parameter C there holds

1. $\mathcal{S}$ is the root of $\mathcal{T}$.

2. Each node $\sigma \in \mathcal{T}$ is a subset of $\mathcal{S}$.

3. If $\sigma \in \mathcal{T}$ is a leaf, then $|\sigma| \leq C$.

4. If $\sigma \in \mathcal{T}$ is not a leaf, then there are two unique non-empty clusters σ', σ'' satisfying $\sigma' \cap \sigma'' = \emptyset$, $\sigma' \cup \sigma'' = \sigma$, called *sons*.

A cluster tree can be constructed for instance from purely geometrical information as follows: Starting with the index set $\mathcal{S}$, in each step the index set is recursively split in two by halving the associated bounding box along the longest edge and partitioning the index set according to the resulting new bounding boxes. This is repeated until the size of the remaining index sets is smaller than the constant C.

For the elements of a cluster tree, we define the *level* recursively by associating with the index set $\mathcal{S}$ the level 0 and defining the level of σ' as the level of σ plus one, if σ' is a son of σ. The *depth* of a cluster tree is the maximal level of its members. Note that excluding pathological cases, the depth of $\mathcal{T}$ is proportional to $\log(L)$.

As a next step, a hierarchical partition of $\mathcal{S} \times \mathcal{S}$ is derived. If σ, τ are clusters with respective bounding boxes B_σ, B_τ, then the *block* (σ, τ) is *admissible* as long as

$$\max(\operatorname{diam}(B_\sigma), \operatorname{diam}(B_\tau)) < \eta \operatorname{dist}(B_\sigma, B_\tau), \qquad \text{(A.1)}$$

where diam denotes the diameter of a box (in Euclidean distance) and dist is the distance between the boxes. Otherwise, the block is *inadmissible*.

The construction described above provides *almost linear scaling* of both storage requirements for and arithmetic operations with hierarchical matrices. More precisely, for a matrix $A \in \mathbb{R}^{L \times L}$, both the storage requirement for the $\mathcal{H}$-matrix approximation and number of floating point operations required for matrix-vector multiplication when the matrix is in $\mathcal{H}$-format is bounded by $\mathcal{C} \ln(L)$ [9].

The algebraic structure just derived is now applied for the approximation of a matrix $A \in \mathbb{R}^{L \times L}$, where each block in the sense of the definition given above corresponds with a submatrix $A_{\mathcal{I} \times \mathcal{J}}$ of A, where $\mathcal{I}$, $\mathcal{J}$ denote subsets of the index set $\mathcal{S}$. In the method of *cross approximation*, a low rank approximation of each admissible block is computed by some suitable method. In our case we can simply use a truncated singular value decomposition, as the computational cost of setting up the $\mathcal{H}$-matrix approximation is negligible as compared to the application of this matrix in the course of the time integration of the MCTDHF equations.

In our analysis in §3, we make use of the following error bound resulting from cross approximation. A proof of this result is given in [2].

Theorem A.2 Let a matrix $A \in \mathbb{R}^{L \times L}$ be approximated by a hierarchical matrix $A_{\mathcal{H}}$, where on each admissible block $A_{\mathcal{I} \times \mathcal{J}}$,

$$\|A_{\mathcal{I} \times \mathcal{J}} - A_{\mathcal{H}, \mathcal{I} \times \mathcal{J}}\|_2 \leq \varepsilon$$

holds with an error margin ε which can be freely chosen. Then,

$$\|A - A_{\mathcal{H}}\|_2 \leq \text{const. } \ln(L)\varepsilon.$$

References

[1] O. Axelsson, V. Barker, Finite Element Solution of Boundary Value Problems: Theory and Computation, Academic Press, Orlando, Fa., 1984.

[2] M. Bebendorf, W. Hackbusch, Stabilized rounded addition of hierarchical matrices, Numer. Linear Algebra Appl. 14 (2007) 407–423.

[3] M. H. Beck, A. Jäckle, G. A. Worth, H.-D. Meyer, The multiconfiguration time-dependent Hartree (MCTDH) method: A highly efficient algorithm for propagating wavepackets, Phys. Rep. 324 (2000) 1–105.

[4] D. Braess, Finite Elements, 2nd ed., Cambridge University Press, Cambridge, U.K., 2001.

[5] S. Brenner, L. Scott, The Mathematical Theory of Finite Element Methods, 2nd ed., Springer-Verlag, New York, 2002.

[6] J. Caillat, J. Zanghellini, M. Kitzler, W. Kreuzer, O. Koch, A. Scrinzi, Correlated multielectron systems in strong laser pulses — an MCTDHF approach, Phys. Rev. A 71 (2005) 012712.

[7] P. Dirac, Note on exchange phenomena in the Thomas atom, Proc. Cambridge Phil. Soc. 26 (1930) 376–385.

[8] J. Frenkel, Wave Mechanics, Advanced General Theory, Clarendon Press, Oxford, 1934.

[9] L. Grasedyck, W. Hackbusch, Construction and arithmetic of $\mathcal{H}$-matrices, Computing 70 (2003) 295–334.

[10] W. Hackbusch, A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices, Computing 62 (1999) 89–108.

[11] W. Hackbusch, B. Khoromskij, A sparse $\mathcal{H}$-matrix arithmetic. Part II: Application to multi-dimensional problems, Computing 64 (1999) 21–47.

[12] O. Koch, Efficient computation of the MCTDHF approximation to the time-dependent Schrödinger equation, Opuscula Math. 26 (2006) 473–487.

[13] O. Koch, Approximation of meanfield terms in MCTDHF computations by $\mathcal{H}$-matrices, ASC Report 5/2008, Inst. for Anal. and Sci. Comput., Vienna Univ. of Technology (2008).

[14] H.-D. Meyer, U. Manthe, L. S. Cederbaum, The multi-configurational time-dependent Hartree approach, Chem. Phys. Lett. 165 (1990) 73–78.

[15] J. Zanghellini, M. Kitzler, T. Brabec, A. Scrinzi, Testing the multi-configuration time-dependent Hartree–Fock method, J. Phys. B: At. Mol. Phys. 37 (2004) 763–773.

[16] J. Zanghellini, M. Kitzler, C. Fabian, T. Brabec, A. Scrinzi, An MCTDHF approach to multi-electron dynamics in laser fields, Laser Physics 13 (8) (2003) 1064–1068.