AN ERROR ANALYSIS OF THE MULTI-CONFIGURATION TIME-DEPENDENT HARTREE METHOD OF QUANTUM DYNAMICS

Dajana Conte[1] and Christian Lubich[2]

Abstract. This paper gives an error analysis of the multi-configuration time-dependent Hartree (MCTDH) method for the approximation of multi-particle time-dependent Schrödinger equations. The MCTDH method approximates the multivariate wave function by a linear combination of products of univariate functions and replaces the high-dimensional linear Schrödinger equation by a coupled system of ordinary differential equations and low-dimensional nonlinear partial differential equations. The main result of this paper yields an L^2 error bound of the MCTDH approximation in terms of a best-approximation error bound in a stronger norm and of lower bounds of singular values of matrix unfoldings of the coefficient tensor. This result permits us to establish convergence of the MCTDH method to the exact wave function under appropriate conditions on the approximability of the wave function, and it points to reasons for possible failure in other cases.

Highlight

INTRODUCTION

This paper provides an error analysis of the MCTDH method, which is a remarkably successful method for the approximate solution of the time-dependent multi-particle Schrödinger equation

$$\mathrm{i}\frac{\partial \psi}{\partial t} = H\psi, \tag{0.1}$$

where the wave function $\psi = \psi(x^{(1)}, \ldots, x^{(N)}, t)$ depends on the spatial coordinates $x^{(n)} \in \mathbb{R}^3$ of N particles (nuclei in a molecule), and on time t. In atomic units ($\hbar = 1$), the Hamiltonian is given by

$$H = T + V = -\frac{1}{2m_1}\Delta^{(1)} - \ldots - \frac{1}{2m_N}\Delta^{(N)} + V(x^{(1)}, \ldots, x^{(N)}). \tag{0.2}$$

In the kinetic energy operator T, the Laplacian $\Delta^{(n)}$ is taken with respect to the spatial coordinates of the nth particle of mass m_n. The real potential V, which acts as a multiplication operator, will be assumed to

[1] Dipartimento di Matematica ed Informatica, Università degli studi di Salerno, Via ponte don Melillo, 84084 Fisciano (SA), Italy. `dajconte@unisa.it`

[2] Universität Tübingen, Mathematisches Institut, Auf der Morgenstelle 10, 72076 Tübingen, Germany. `lubich@na.uni-tuebingen.de`

be bounded. The situation of primary interest here is that of a Schrödinger equation for distinguishable nuclei moving in a potential given by an (approximate) electronic energy surface, according to the time-dependent Born-Oppenheimer approximation to the full molecular Schrödinger equation. The assumption of a smooth and bounded potential is usually a reasonable modelling assumption in these applications.

For computational wavepacket propagation, the *multi-configuration time-dependent Hartree method* (MCTDH) has been put forward by Meyer *et al.* in [1,12–14] and further references therein. It has been used successfully in a variety of chemical situations such as photodissociation and reactive scattering, for problems involving 6 to 24 nuclear degrees of freedom and one or several electronic states; see, *e.g.*, [15]. In the MCTDH approach, the wave function is approximated by a linear combination of Hartree products, that is, of products of functions each depending on the coordinates of only a single particle, or of a single degree of freedom. The Dirac–Frenkel time-dependent variational principle yields equations of motion for the single-particle functions and for the coefficients in the linear combination of the Hartree products. The ordinary differential equations for the coefficients are those of a standard Galerkin method with an orthonormal basis of Hartree products, but here this basis changes in time according to the low-dimensional nonlinear partial differential equations satisfied by the single-particle functions. The MCTDH method thus replaces the high-dimensional linear Schrödinger equation by a system of ordinary differential equations and low-dimensional nonlinear partial differential equations and in this way makes the problem computationally tractable.

It is tempting to think that taking more and more linear combinations of Hartree products should give an ever better approximation to the wave function. This intuitive expectation is, however, not easily put on firm ground. Obstructions come from the difficulty to ascertain the approximation properties of the time-dependent basis of Hartree products and from the fact that the density matrices appearing in the method formulation become more and more ill-conditioned as more terms are added in the linear combination of Hartree products. These two obstructions render a standard convergence analysis illusory.

Although a smooth wave function can indeed be well approximated by linear combinations of tensor products of single-particle functions, it is not clear how this property relates to the approximation provided by the MCTDH method. The L^2 error of the MCTDH method is bounded in terms of the L^2 best-approximation error in [10], but the constants in these estimates grow without bound as the number of linear combinations increases, due to the growing ill-conditioning of the density matrices.

In the error analysis given here, we suppose that the wave function *can* be approximated by linear combinations of Hartree products in such a way that the residual in the time-dependent Schrödinger equation is small $(\leq \varepsilon)$ and the singular values of the matrix unfoldings of the coefficient tensor are not too small $(\geq \delta)$. We then obtain, in Theorem 2.1, an $O(\varepsilon)$ error bound for the MCTDH approximation on a time interval of length proportional to δ/ε. Moreover, the inverses of the density matrices of the MCTDH method are bounded by $O(\delta^{-2})$ on such a time interval. It cannot be expected that δ is substantially larger than ε, but the bound $\delta \geq c\varepsilon$ for a positive constant c independent of the number of linear combinations appears as a reasonable assumption. Under a condition that requires smallness of the time derivative of the coefficient tensor in components that correspond to the small singular values, the error bound of Theorem 2.1 then yields the convergence of MCTDH approximations to the exact wave function over some fixed time interval, as more and more terms are included in the linear combinations. It must be noted, however, that our error analysis does *not* establish convergence of the MCTDH method in general, but it does show mechanisms that lead to small errors or to possible breakdown of the approximation.

We mention that multi-configuration Hartree–Fock methods for the Schrödinger eigenvalue problem of electronic structure theory are analysed in [2,9], where convergence of the ground state and of the energies of excited states are shown.

In Section 1 we describe the MCTDH method. The main result of our error analysis is stated and discussed in Section 2. We introduce useful notation for the proof in Section 3, and in Section 4 we reformulate the MCTDH equations of motion in this compact notation. In Sections 5 and 6 we study tangent space projections onto the MCTDH manifold as an essential tool of our error analysis. The proof of the main result is then given in Section 7.

1. THE MCTDH METHOD

In the MCTDH method [1,13], the multi-variate wave function ψ is approximated by a linear combination of products of uni-variate functions, that is, for $x = (x^{(1)}, \ldots, x^{(N)})$,

$$\psi(x,t) \approx u(x,t) = \sum_{(j_1,\ldots,j_N)} a_{j_1,\ldots,j_N}(t)\, \phi^{(1)}_{j_1}(x^{(1)},t) \ldots \phi^{(N)}_{j_N}(x^{(N)},t) \equiv \sum_J a_J(t)\, \Phi_J(x,t). \tag{1.3}$$

Here, the multi-indices $J = (j_1, \ldots, j_N)$ vary for $j_n = 1, \ldots, r_n$ with $n = 1, \ldots, N$. The $a_J(t)$ are complex coefficients depending only on t. The *single-particle functions* $\phi^{(n)}_{j_n}(x^{(n)},t)$ depend on the coordinates $x^{(n)}$ of a single particle and on time t. They appear multiplied with each other in the *Hartree products* $\Phi_J(x,t) = \prod_{n=1}^N \phi^{(n)}_{j_n}(x^{(n)},t)$.

This approximation format is analogous to the Tucker format in tensor approximation (see, *e.g.*, [4,8]). The computational scaling is exponential in the dimension N, as r^N if $r_n = r$ for all n, unless a reduced format is chosen for the coefficient tensor (a_J) as, *e.g.*, in [4,17].

The Dirac–Frenkel time-dependent variational principle yields differential equations for the coefficients a_J and the single-particle functions $\phi^{(n)}_{j_n}$. We first recall this variational approximation procedure in its abstract form and then turn to the MCTDH approximation manifold and the MCTDH equations of motion. The presentation follows [5,11].

1.1. **The Dirac–Frenkel variational approximation principle**

The abstract setting is that of the time-dependent Schrödinger equation

$$\mathrm{i}\frac{\mathrm{d}\psi}{\mathrm{d}t} = H\psi\,, \qquad \psi(0) = \psi_0, \tag{1.4}$$

where the Hamiltonian H is a self-adjoint linear operator on a complex Hilbert space $\mathcal{H}$ with inner product $\langle \cdot | \cdot \rangle$ and norm $\|\cdot\|$. Let $\mathcal{M} \subset \mathcal{H}$ be a submanifold of $\mathcal{H}$ on which an approximation to the wave function $\psi(t)$ should lie, and let $\mathcal{T}_u\mathcal{M}$ denote the tangent space at $u \in \mathcal{M}$ (*i.e.*, the closed real-linear subspace of $\mathcal{H}$ formed of the derivatives of all paths on $\mathcal{M}$ passing through u, or in physical terminology, the space of admissible variations). We assume that $\mathcal{T}_u\mathcal{M}$ is in fact complex linear, *i.e.*, with $\delta u \in \mathcal{T}_u\mathcal{M}$ also $\mathrm{i}\,\delta u \in \mathcal{T}_u\mathcal{M}$. The Dirac–Frenkel principle determines the approximate wave function $t \mapsto u(t) \in \mathcal{M}$ from the condition that the time derivative $\dot{u} = \mathrm{d}u/\mathrm{d}t$ should satisfy, at every time t,

$$\left\langle \delta u \,\middle|\, \dot{u} - \tfrac{1}{\mathrm{i}}Hu \right\rangle = 0 \qquad \text{for all} \quad \delta u \in \mathcal{T}_u\mathcal{M}. \tag{1.5}$$

With the orthogonal projection $P(u)$ onto the tangent space $\mathcal{T}_u\mathcal{M}$, this can be rephrased as a differential equation on $\mathcal{M}$,

$$\dot{u} = P(u)\tfrac{1}{\mathrm{i}}Hu. \tag{1.6}$$

From a numerical analysis viewpoint, condition (1.5) can be seen as a Galerkin condition on the state-dependent approximation space $\mathcal{T}_u\mathcal{M}$.

1.2. **The MCTDH approximation manifold**

The MCTDH method determines approximations to the wave function that, for every time t, lie in the set

$$\overline{\mathcal{M}} = \left\{ u \in L^2(\mathbb{R}^d) : u = \sum_{j_1=1}^{r_1} \cdots \sum_{j_N=1}^{r_N} a_{j_1\ldots j_N}\, \phi^{(1)}_{j_1} \otimes \cdots \otimes \phi^{(N)}_{j_N} \ \text{with} \ a_{j_1\ldots j_N} \in \mathbb{C},\ \phi^{(n)}_{j_n} \in L^2(\mathbb{R}^{d_n}) \right\}, \tag{1.7}$$

Highlight

with the total dimension $d = d_1 + \ldots + d_N$. This set $\overline{\mathcal{M}}$ is not a manifold, but $\overline{\mathcal{M}}$ contains a dense subset $\mathcal{M}$ that is a manifold and is characterised by a full-rank condition.

The representation of $u \in \overline{\mathcal{M}}$ by a coefficient tensor $\mathcal{A} = (a_J)_J$ and single-particle functions $\phi = (\phi_{j_n}^{(n)})_{j_n, n}$ is not unique. We may assume that the $\phi_{j_n}^{(n)}$ corresponding to the same particle n are orthonormal:

$$\langle \phi_{j_n}^{(n)} \,|\, \phi_{l_n}^{(n)} \rangle = \delta_{j_n, l_n}, \qquad j_n, l_n = 1, \ldots, r_n, \ n = 1, \ldots, N \tag{1.8}$$

(where δ_{j_n, l_n} is Kronecker's delta). Consider a differentiable path $t \mapsto (\mathcal{A}(t), \phi(t))$ representing a path $u(t)$ on $\overline{\mathcal{M}}$. Then, the derivative $\delta u = \dot{u}(0)$ is of the form

$$\delta u = \sum_J \delta a_J \, \Phi_J + \sum_{n=1}^{N} \sum_{j_n=1}^{r_n} \delta \phi_{j_n}^{(n)} \, \psi_{j_n}^{(n)} \tag{1.9}$$

with the Hartree products $\Phi_J = \bigotimes_{n=1}^{N} \phi_{j_n}^{(n)}$ and with the *single-hole functions*

$$\psi_{j_n}^{(n)} \quad := \quad \langle \phi_{j_n}^{(n)} \,|\, u \rangle^{(n)} = \sum_{j_1=1}^{r_1} \cdots \sum_{j_{n-1}=1}^{r_{n-1}} \sum_{j_{n+1}=1}^{r_{n+1}} \cdots \sum_{j_N=1}^{r_N} a_{j_1, \ldots, j_N} \bigotimes_{\ell \neq n} \phi_{j_\ell}^{(\ell)} \tag{1.10}$$

where the superscript (n) on the inner product indicates that the L^2 inner product is taken only with respect to the variable $x^{(n)}$, leaving a function depending on all the other variables $x^{(\ell)}$ with $\ell \neq n$. Conversely, the δa_J turn out to be uniquely determined by δu and $(\mathcal{A}, \phi)$ if we impose the *gauge constraints*

$$\langle \phi_{j_n}^{(n)} \,|\, \delta \phi_{l_n}^{(n)} \rangle = 0, \qquad j_n, l_n = 1, \ldots, r_n, \ n = 1, \ldots, N, \tag{1.11}$$

which is a stronger condition than the differentiated condition (1.8), i.e., $\langle \phi_{j_n}^{(n)} \,|\, \delta \phi_{l_n}^{(n)} \rangle + \langle \delta \phi_{j_n}^{(n)} \,|\, \phi_{l_n}^{(n)} \rangle = 0$. On taking the inner product of Φ_J with δu given by (1.9), conditions (1.8) and (1.11) together imply

$$\delta a_J = \langle \Phi_J \,|\, \delta u \rangle.$$

Taking the inner product of $\psi_{i_n}^{(n)}$ with (1.9) gives

$$\sum_{j_n=1}^{r_n} \rho_{i_n, j_n}^{(n)} \, \delta \phi_{j_n}^{(n)} = \left\langle \psi_{i_n}^{(n)} \,\Big|\, \delta u - \sum_J \delta a_J \, \Phi_J \right\rangle^{(\neg n)} \tag{1.12}$$

with the hermitian, positive semi-definite *density matrices*

$$\rho_{(n)} = \left(\rho_{i_n, j_n}^{(n)} \right)_{i_n, j_n=1}^{r_n} \qquad \text{given by} \qquad \rho_{i_n, j_n}^{(n)} := \langle \psi_{i_n}^{(n)} | \psi_{j_n}^{(n)} \rangle^{(\neg n)}. \tag{1.13}$$

The superscript $(\neg n)$ indicates that the L^2 inner product is taken over all variables except $x^{(n)}$, leaving a function depending on $x^{(n)}$ in (1.12). The orthonormality relations (1.8) allow us to express the entries of the density matrices in terms of the coefficients a_J:

$$\rho_{i_n, j_n}^{(n)} = \sum_{j_1=1}^{r_1} \cdots \sum_{j_{n-1}=1}^{r_{n-1}} \sum_{j_{n+1}=1}^{r_{n+1}} \cdots \sum_{j_N=1}^{r_N} \overline{a}_{j_1, \ldots, j_{n-1}, i_n, j_{n+1}, \ldots, j_N} \, a_{j_1, \ldots, j_N} \tag{1.14}$$

or equivalently

$$\overline{\rho}_{(n)} = A_{(n)} A_{(n)}^*, \tag{1.15}$$

where the matrix $A_{(n)} \in \mathbb{C}^{r_n \times \prod_{k \neq n} r_k}$ is the nth *unfolding* of the tensor $\mathcal{A}$, which is the matrix that in its j_nth row aligns the entries $a_{j_1,\ldots,j_N}$ with fixed j_n in lexicographical ordering.

The $\delta\phi_{j_n}^{(n)}$ are thus uniquely determined from (1.12) under the *full-rank condition* that the tensor $\mathcal{A}$ has rank $(r_1, \ldots, r_N)$, which means that its nth unfolding $A_{(n)}$ has rank r_n for each $n = 1, \ldots, N$, or equivalently,

$$\rho_{(n)} \text{ is an invertible matrix for each } n = 1, \ldots, N. \tag{1.16}$$

With the above construction of the δa_J and $\delta\phi_{j_n}^{(n)}$, we obtain local charts on

$$\mathcal{M} = \left\{ u \in L^2(\mathbb{R}^d) : u = \sum_{j_1=1}^{r_1} \cdots \sum_{j_N=1}^{r_N} a_{j_1\ldots j_N} \, \phi_{j_1}^{(1)} \otimes \cdots \otimes \phi_{j_N}^{(N)} \text{ with } a_{j_1\ldots j_N} \in \mathbb{C}, \ \phi_{j_n}^{(n)} \in L^2(\mathbb{R}^{d_n}) \right. $$
$$\left. \text{satisfying the orthonormality constraints (1.8) and the full-rank condition (1.16)} \right\}, \tag{1.17}$$

making this set an infinite-dimensional manifold, for which the tangent space at $u \in \mathcal{M}$ consists of the elements δu of the form (1.9). The MCTDH method is obtained by using the Dirac–Frenkel principle on this approximation manifold $\mathcal{M}$.

1.3. The MCTDH equations of motion

Using the Dirac–Frenkel principle on the approximation manifold $\mathcal{M}$ of (1.17) and imposing, in view of (1.11), additional orthogonality constraints on the time derivatives of the single-particle functions $\phi_{j_n}^{(n)}(x^{(n)}, t)$,

$$\left\langle \phi_{j_n}^{(n)} \,\Big|\, \frac{\partial \phi_{l_n}^{(n)}}{\partial t} \right\rangle = 0, \qquad t \geq 0, \ \ j_n, l_n = 1, \ldots, r_n, \ n = 1, \ldots, N, \tag{1.18}$$

yields a system of coupled ordinary and partial differential equations for the coefficients and single-particle functions [1,12,13]:

$$\mathrm{i}\frac{\mathrm{d}a_J}{\mathrm{d}t} = \sum_K \langle \Phi_J | H | \Phi_K \rangle \, a_K, \qquad \forall J, \tag{1.19}$$

$$\mathrm{i}\frac{\partial \phi_{j_n}^{(n)}}{\partial t} = (1 - P^{(n)}) \sum_{m_n=1}^{r_n} \sum_{l_n=1}^{r_n} (\rho_{(n)}^{-1})_{j_n, m_n} \langle \psi_{m_n}^{(n)} \,|\, H \,|\, \psi_{l_n}^{(n)} \rangle^{(\neg n)} \, \phi_{l_n}^{(n)}, \tag{1.20}$$
$$j_n = 1, \ldots, r_n, \ n = 1, \ldots, N,$$

where the Hartree products Φ_J, the single-hole functions $\psi_{l_n}^{(n)}$, and the density matrices $\rho^{(n)}$ are defined as in Section 1.2, and where $P^{(n)}$ is the orthogonal projector onto the space spanned by $\phi_1^{(n)}, \ldots, \phi_{r_n}^{(n)}$,

$$P^{(n)}\theta = \sum_{j_n=1}^{r_n} \phi_{j_n}^{(n)} \, \langle \phi_{j_n}^{(n)} \,|\, \theta \rangle^{(n)}.$$

Existence and regularity of solutions to the MCTDH equations of motion are studied in [5]. In the case of a smooth bounded potential, it is shown that the MCTDH approximation exists with the same Sobolev regularity as the initial data as long as the density matrices remain invertible.

Highlight

2. STATEMENT AND DISCUSSION OF THE MAIN RESULT

We assume that the exact wave function $\psi(t) = \psi(\cdot, t)$, which is a solution to the Schrödinger equation (0.1), can be written in the form

$$\psi(t) = v(t) + e(t), \qquad 0 \le t \le \overline{t}, \tag{2.1}$$

where $v(t) = v(\cdot, t) \in \mathcal{M}$ is *some* approximation to the wave function in the MCTDH manifold that has a small defect in the Schrödinger equation (0.1): with a small $\varepsilon > 0$, the defect is bounded in the $L^2(\mathbb{R}^d)$ norm by

$$\left\| \mathrm{i} \frac{\partial v}{\partial t}(\cdot, t) - H v(\cdot, t) \right\| \le \varepsilon. \tag{2.2}$$

We note that a sufficient condition for this inequality is given by the bound

$$\left\| \frac{\partial e}{\partial t}(\cdot, t) \right\| + \| H e(\cdot, t) \| \le \varepsilon,$$

which holds if the error $e = \psi - v$ is small in the $C^1([0, \overline{t}], L^2) \cap C([0, \overline{t}], H^2)$ norm.

We further assume that $v(\cdot, t) \in \mathcal{M}$ has a decomposition

$$v(x, t) = \sum_{j_1=1}^{r_1} \cdots \sum_{j_N=1}^{r_N} b_{j_1 \ldots j_N}(t) \, \theta_{j_1}^{(1)}(x^{(1)}, t) \ldots \theta_{j_N}^{(N)}(x^{(N)}, t) \tag{2.3}$$

with the following properties for $0 \le t \le \overline{t}$: there are a positive number $\delta > 0$ and non-negative μ, ν such that the smallest non-zero singular value $\sigma_{r_n}(B_{(n)}(t))$ of the nth unfolding $B_{(n)}(t)$ of the tensor $\mathcal{B}(t) = \left(b_{j_1 \ldots j_N}(t) \right)$ satisfies, for $n = 1, \ldots, N$,

$$\sigma_{r_n}(B_{(n)}(t)) \ge \delta, \tag{2.4}$$

and

$$\left\| B_{(n)}^\dagger(t) \, \dot{B}_{(n)}(t) \right\|_2 \le \mu, \tag{2.5}$$

where $\| \cdot \|_2$ is the spectral norm, the matrix $B_{(n)}^\dagger = B_{(n)}^* \left(B_{(n)} B_{(n)}^* \right)^{-1}$ is the pseudo-inverse of $B_{(n)}$, and $\dot{} = \mathrm{d}/\mathrm{d}t$ denotes the time derivative. Note that condition (2.4) is equivalent to the bound $\| B_{(n)}^\dagger(t) \|_2 \le \delta^{-1}$. Condition (2.5) with $\mu \ll \delta^{-1}$ requires smallness of the time derivative of the coefficient tensor in components that correspond to the small singular values.

The basis functions $\theta_{j_n}^{(n)}$ are assumed to be orthogonal to each other for each n and to satisfy the bounds

$$\sum_{j_n=1}^{r_n} \left\| \mathrm{i} \frac{\partial \theta_{j_n}^{(n)}}{\partial t}(\cdot, t) - T^{(n)} \theta_{j_n}^{(n)}(\cdot, t) \right\|^2 \le \nu^2, \tag{2.6}$$

where $T^{(n)} = -(2m_n)^{-1} \Delta^{(n)}$ is the kinetic energy operator of the nth particle, and the norm is the $L^2(\mathbb{R}^{d_n})$ norm. Moreover, we assume that the potential is bounded:

$$|V(x)| \le \beta \qquad \text{for all} \quad x = (x^{(1)}, \ldots, x^{(N)}) \in \mathbb{R}^d. \tag{2.7}$$

We then have the following error bound for the MCTDH approximation $u(t) = u(\cdot, t)$ given by (1.3) and (1.19)–(1.20).

Theorem 2.1. *Under the above conditions, the approximation error of the MCTDH method with initial value* $u(0) = v(0)$ *is bounded by*

$$\|u(t) - \psi(t)\| \leq \|e(t)\| + 2t\varepsilon \qquad for \quad t \leq \min\left(\overline{t}, \frac{c}{2}\left(\frac{\delta}{\varepsilon}\right), \frac{1}{\sqrt{\gamma}}\left(\frac{\delta}{\varepsilon}\right)^{1/2}\right),$$

where $\gamma = 2C\,(c_1\mu + c_2\nu + \beta)$ *with* $c_1 = 3N$, $c_2 = 2N^2$, $C = 8N(N+3)$, *and* $c = \frac{1}{2C}$.

Note that the ill-conditioning of $B_{(n)}$ as expressed by the smallness of δ does not affect the error bound on intervals of a length proportional to δ/ε.

Remark 2.2. (a) The same error bound holds, with essentially the same proof, if conditions (2.4)–(2.6) are assumed for the MCTDH approximation, *i.e.*, for $A_{(n)}$ and $\phi_{j_n}^{(n)}$ instead of $B_{(n)}$ and $\theta_{j_n}^{(n)}$. The condition $\sigma_{r_n}(A_{(n)}(t)) \geq \delta$ is equivalent to the bound $\|\rho_{(n)}^{-1}(t)\|_2 \leq \delta^{-2}$ of the inverse of the nth density matrix (1.15) which appears in (1.20). We further note that under condition (2.4) on $B_{(n)}$, we obtain $\sigma_{r_n}(A_{(n)}(t)) \geq \frac{1}{2}\delta$ and $\|\rho_{(n)}^{-1}(t)\|_2 \leq 4\delta^{-2}$ for t in the interval given in the theorem.

(b) In Theorem 2.1 it is not essential that the operator T is a mass-scaled Laplacian. It is sufficient that T is a sum of densely defined one-particle operators, which ensures the basic property that $Tu \in \mathcal{T}_u\mathcal{M}$ for all $u \in \mathcal{M} \cap D(T)$. On the other hand, the boundedness of the potential V appears substantially in our proof. It is not clear to us how to extend the result to the MCTDHF method of electron dynamics, where unbounded Coulomb potentials are present.

Relating δ and ε. We describe a situation in which the key quantities δ and ε of Theorem 2.1 are in close relationship. For ease of notation only, we consider here the case where all ranks r_n are equal, $r = r_1 = \ldots = r_N$. We are interested in the behaviour of δ and ε as r grows, and therefore write δ_r and ε_r in the following.

We expand the wave function into a basis of tensor products of orthogonal functions, *e.g.*, of shifted and scaled Hermite functions as considered below. We then have

$$\psi(x,t) = \sum_{j_1=0}^{\infty} \ldots \sum_{j_n=0}^{\infty} b_{j_1\ldots j_N}(t)\theta_{j_1}(x^{(1)},t)\ldots\theta_{j_N}(x^{(N)},t).$$

We denote by ν_r the bound (2.6) in dependence of r for these basis functions.

We now make the assumption on the wave function that with respect to this basis, the coefficients $b_{j_1\ldots j_N}(t)$ decay with growing indices, in the sense that for some $\eta_r > 0$ with $\eta_r \to 0$ for $r \to \infty$ (*e.g.*, $\eta_r = r^{-m}$ for some $m > 1$),

$$\sum_{(j_1,\ldots,j_N):\, j_n > r \text{ for some } n} |b_{j_1\ldots j_N}(t)|^2 \leq \eta_r^2 \tag{2.8}$$

$$\sum_{(j_1,\ldots,j_N):\, j_n > r \text{ for some } n} |\dot{b}_{j_1\ldots j_N}(t)|^2 \leq \alpha^2\,\eta_r^2. \tag{2.9}$$

When we take $v(t)$ as the truncated expansion with r terms in each coordinate, we have the following inequalities:

$$\delta_r \leq \eta_r, \tag{2.10}$$

and (2.2) holds with

$$\varepsilon_r = \eta_r\,(\sqrt{\alpha^2 + \nu_r^2} + \beta). \tag{2.11}$$

To prove (2.10), let $B_{(n)}^{[r]}(t)$ denote the nth unfolding matrix for the tensor with coefficients $b_{j_1\ldots j_N}(t)$ for $j_1, \ldots, j_N \leq r$. We then have from [3], p. 448, that the difference between the smallest singular value of $B_{(n)}^{[r]}(t)$

and the smallest singular value 0 of the matrix $B_{(n)}^{[r-1]}(t)$ filled up with zeros, is bounded by the Frobenius norm of the difference of these matrices, which is bounded by η_r. This yields (2.10). The bound (2.2) with (2.11) follows by combining the inequalities (2.6), (2.7), and (2.8)–(2.9).

The estimate (2.10) makes it clear that δ_r must tend to 0 if $r \to \infty$, and the faster the better the exact wave function is approximated with just a few Hartree products. With (2.10)–(2.11) we have $\delta_r \leq \text{const.}\,\varepsilon_r$. It would be of interest to have also a converse inequality. By replacing the coefficient tensor $(b_{j_1 \ldots j_N})$ with a perturbed one, where all singular values of the unfolding matrices are made greater than η_r, one obtains for this perturbation $\delta_r = \eta_r$ and the bound (2.2) with $\varepsilon_r = c_r \eta_r$. Unfortunately it does not seem possible to bound c_r independently of r in general, since the number of singular values that need to be increased to η_r may grow linearly with r. In such a situation we would only get $c_r = O(\sqrt{r})$. While this cannot be excluded, an assumption $\delta_r / \varepsilon_r \geq \text{const.} > 0$ appears realistic for many cases.

Basis functions with $\nu = 0$. As the following example shows, assuming condition (2.6) with ν independent of r is not an unreasonable assumption. Consider the kth Hermite function

$$\varphi_k(x) = \frac{1}{\pi^{1/4}} \frac{1}{\sqrt{2^k k!}} H_k(x)\, e^{-x^2/2}$$

with the Hermite polynomial H_k of degree k. These functions form an orthonormal basis of $L^2(\mathbb{R})$ (see, *e.g.*, [16]). We choose the basis functions

$$\theta_{j_n}(x^{(n)}, t) = \varphi_{j_n}\left(\frac{x^{(n)} - q^{(n)}(t)}{a^{(n)}(t)}\right) e^{i p^{(n)} x^{(n)}} e^{i S^{(n)}(t)}$$

as shifted and scaled Hermite functions multiplied with complex exponentials, and expand

$$\psi(x, t) = \sum_{j_1=0}^{\infty} \ldots \sum_{j_n=0}^{\infty} b_{j_1 \ldots j_N}(t) \theta_{j_1}(x^{(1)}, t) \ldots \theta_{j_N}(x^{(N)}, t).$$

With the choice $q^{(n)}(t) = q^{(n)}(0) + \frac{t}{m_n} p^{(n)}$, $a^{(n)}(t) = a^{(n)}(0) + \frac{it}{2m_n}$, $S^{(n)}(t) = S^{(n)}(0) + \frac{t}{2m_n}(p^{(n)})^2$ of the position, scaling and phase parameters, respectively, the functions $\theta_{j_n}(x^{(n)}, t)$ are solutions of the free Schrödinger equation, and hence (2.6) holds with $\nu = 0$.

Conditional convergence of the MCTDH method. An interesting theoretical question concerns the convergence of the MCTDH method to the solution of the Schrödinger equation as the numbers r_n of basis functions tend to infinity. It appears already obvious from the method formulation that such a result cannot be given without assuming some *a priori* bounds on $\|\rho_{(n)}^{-1}(t)\|_2$, that is, on δ^{-2} with δ of (2.4). As we have seen above, such bounds cannot reasonably be assumed independently of r. *If we have $\delta_r / \varepsilon_r \geq \text{const.} > 0$ and if the bounds (2.5) and (2.6) are uniform in r, then Theorem 2.1 yields convergence of the MCTDH approximations to the exact wave function over a fixed time interval, as $r \to \infty$.*

Road map. The remaining sections of this paper develop the proof of Theorem 2.1. The reader who wishes to get a quick impression of the ideas and procedure of the proof before entering into technical details, may proceed directly to Section 7. The proof given there uses bounds that are derived in Sections 5 and 6 using suitable notation introduced in Sections 3 and 4.

3. NOTATION AND BASIC ESTIMATES FOR TENSOR FUNCTIONS

For the proof of Theorem 2.1 it is appropriate to use more compact notation than, say, in (1.3). We collect some notation of tensors of functions, partly inspired by the tensor notation in [8]. As a general rule in the following, superscripts refer to variables of a function and subscripts to indices of a tensor.

Function spaces. For a subset Q of $\{1, ..., N\}$, we denote by $\mathcal{H}^Q$ the space of square-integrable functions $f = f(x^{(q)} \,|\, q \in Q)$ depending on the variables $x^{(q)} \in \mathbb{R}^{d_q}$ with $q \in Q$, i.e., for $Q \neq \emptyset$ we have $\mathcal{H}^Q = L^2(\mathbb{R}^{d_Q})$, where $d_Q = \sum_{q \in Q} d_q$ is the total dimension of variables in Q, and $\mathcal{H}^\emptyset = \mathbb{C}$. We write $\mathcal{H} = \mathcal{H}^{\{1,...,N\}} = L^2(\mathbb{R}^d)$.

Tensors of functions. For subsets P and Q of $\{1, ..., N\}$, we denote by $\mathcal{T}_P^Q = (\mathcal{H}^Q)^{s_1 \times ... \times s_N}$ the space of tensors of dimension $s_1 \times ... \times s_N$ where

$$s_n = \begin{cases} r_n \text{ for } n \in P \\ 1 \text{ for } n \notin P, \end{cases}$$

and whose elements are functions of the variables $x^{(q)}$ with $q \in Q$, i.e., an element $\mathcal{Y} \in \mathcal{T}_P^Q$ is of the form

$$\mathcal{Y} = \left(y_{J_P}^{(Q)} \right)_{J_P},$$

with $y_{J_P}^{(Q)} \in \mathcal{H}^Q$, where the multi-indices $J_P = (j_p \,|\, p \in P)$ vary for $j_p = 1, ..., r_p$.

We consider in $\mathcal{T}_P^Q$ the norm

$$\|\mathcal{Y}\|_{\mathcal{T}_P^Q}^2 = \langle \mathcal{Y}, \mathcal{Y} \rangle_{\mathcal{T}_P^Q} = \sum_{J_P} \left\| y_{J_P}^{(Q)} \right\|_{\mathcal{H}^Q}^2, \tag{3.1}$$

induced by the inner product

$$\langle \mathcal{X}, \mathcal{Y} \rangle_{\mathcal{T}_P^Q} := \sum_{J_P} \left\langle x_{J_P}^{(Q)}, y_{J_P}^{(Q)} \right\rangle_{\mathcal{H}^Q}, \tag{3.2}$$

where if $Q = \emptyset$ we consider, for $x, y \in \mathcal{H}^\emptyset = \mathbb{C}$, $\langle x, y \rangle_\mathbb{C} = \overline{x}y$. In particular we have $\mathcal{T}_\emptyset^{\{1,...,N\}} = L^2(\mathbb{R}^d)$, $\mathcal{T}_\emptyset^Q = \mathcal{H}^Q$, and $\mathcal{T}_{\{1,...,N\}}^\emptyset = \mathbb{C}^{r_1 \times ... \times r_N}$, equipped with the Frobenius norm.

Matrices of functions and their products. We denote by $\mathcal{M}_{\alpha \times \beta}^Q$ the space of matrices of dimension $\alpha \times \beta$ whose elements are functions of $x^{(Q)} = (x^{(q)} \,|\, q \in Q)$. For $A^Q \in \mathcal{M}_{\alpha \times \beta}^Q$, $B^R \in \mathcal{M}_{\beta \times \gamma}^R$ and $S \subseteq Q \cap R \neq \emptyset$, we define the matrix $A^Q \bullet^S B^R \in \mathcal{M}_{\alpha \times \gamma}^{(Q \cup R) \setminus S}$ having elements

$$\left(A^Q \bullet^S B^R \right)_{ij} = \sum_{k=1}^{\beta} \left\langle a_{ik}^{(Q)}, b_{kj}^{(R)} \right\rangle_{\mathcal{H}^S}, \tag{3.3}$$

where we denote by $\langle \cdot, \cdot \rangle_{\mathcal{H}^S}$ the L^2 inner product taken only over the variables $x^{(s)}$ with $s \in S$. If $Q \cap R = \emptyset$, we define the matrix $A^Q \otimes B^R \in \mathcal{M}_{\alpha \times \gamma}^{Q \cup R}$ having elements

$$\left(A^Q \otimes B^R \right)_{ij} = \sum_{k=1}^{\beta} a_{ik}^{(Q)} \otimes b_{kj}^{(R)}, \tag{3.4}$$

that is, $\left(A^Q \otimes B^R \right)_{ij} (x^{(Q)}, x^{(R)}) = \sum_{k=1}^{\beta} a_{ik}^{(Q)} (x^{(Q)}) \, b_{kj}^{(R)} (x^{(R)})$.

We consider in $\mathcal{M}_{\alpha \times \beta}^Q$ the Frobenius-type inner product

$$\left\langle A^Q, B^Q \right\rangle_{\mathcal{M}_{\alpha \times \beta}^Q} = \sum_{i=1}^{\alpha} \sum_{j=1}^{\beta} \left\langle a_{ij}^{(Q)}, b_{ij}^{(Q)} \right\rangle_{\mathcal{H}^Q}$$

Highlight

which induces the norm

$$\left\|A^Q\right\|^2_{\mathcal{M}^Q_{\alpha\times\beta}} = \left\langle A^Q, A^Q\right\rangle_{\mathcal{M}^Q_{\alpha\times\beta}} = \sum_{i=1}^{\alpha}\sum_{j=1}^{\beta}\left\|a_{ij}^{(Q)}\right\|^2_{\mathcal{H}^Q}. \tag{3.5}$$

Vectors of functions. We denote by $\mathcal{V}^Q_\alpha = \mathcal{M}^Q_{\alpha\times 1}$ the space of column vectors of dimension α with entries depending on $x^{(Q)}$. The norm of a vector $\Psi^Q = [\psi_1^{(Q)}, ..., \psi_\alpha^{(Q)}]^T \in \mathcal{V}^Q_\alpha$ is given by

$$\left\|\Psi^Q\right\|^2_{\mathcal{V}^Q_\alpha} = \sum_{i=1}^{\alpha}\left\|\psi_i^{(Q)}\right\|^2_{\mathcal{H}^Q} = (\Psi^Q)^T \bullet^Q \Psi^Q. \tag{3.6}$$

Lemma 3.1. *Let* $\Psi^n \in \mathcal{V}^{\{n\}}_\alpha$, $A^Q \in \mathcal{M}^Q_{\beta\times\alpha}$, $\Phi^Q \in \mathcal{V}^Q_\beta$. *Then,*
 (1) *For* $n \notin Q$,
$$\left\|A^Q \otimes \Psi^n\right\|_{\mathcal{V}^{Q\cup\{n\}}_\beta} \leq \left\|A^Q\right\|_{\mathcal{M}^Q_{\beta\times\alpha}}\left\|\Psi^n\right\|_{\mathcal{V}^{\{n\}}_\alpha}. \tag{3.7}$$
 (2) *For* $n \in Q$,
$$\left\|\Psi^n \bullet^n (\Phi^Q)^T\right\|_{\mathcal{M}^{Q\setminus\{n\}}_{\alpha\times\beta}} \leq \left\|\Psi^n\right\|_{\mathcal{V}^{\{n\}}_\alpha}\left\|\Phi^Q\right\|_{\mathcal{V}^Q_\beta}. \tag{3.8}$$

Proof. (1) It follows from (3.6), from the definitions (3.3)–(3.4) and from the fact that the matrix $(A^Q)^T \bullet^Q A^Q$ is Hermitian, that

$$\begin{aligned}
\left\|A^Q \otimes \Psi^n\right\|^2_{\mathcal{V}^{Q\cup\{n\}}_\beta} &= \left((\Psi^n)^T \otimes (A^Q)^T\right) \bullet^{Q\cup\{n\}} \left(A^Q \otimes \Psi^n\right)\\
&= \sum_{i=1}^{\beta}\left\langle \sum_{j=1}^{\alpha}\psi_j^{(n)} \otimes a_{ij}^{(Q)}, \sum_{k=1}^{\alpha}a_{ik}^{(Q)} \otimes \psi_k^{(n)}\right\rangle_{\mathcal{H}^{Q\cup\{n\}}}\\
&= \sum_{j=1}^{\alpha}\sum_{k=1}^{\alpha}\left(\sum_{i=1}^{\beta}\left\langle a_{ij}^{(Q)}, a_{ik}^{(Q)}\right\rangle_{\mathcal{H}^Q}\right)\left\langle \psi_j^{(n)}, \psi_k^{(n)}\right\rangle_{\mathcal{H}^{\{n\}}}\\
&= \sum_{j=1}^{\alpha}\sum_{k=1}^{\alpha}\left((A^Q)^T \bullet^Q A^Q\right)_{jk}\left(\Psi^n \bullet^n (\Psi^n)^T\right)_{jk}\\
&= tr\left(\left((A^Q)^T \bullet^Q A^Q\right)^T\left(\Psi^n \bullet^n (\Psi^n)^T\right)\right)\\
&= \left\langle\left((A^Q)^T \bullet^Q A^Q\right)^T, \Psi^n \bullet^n (\Psi^n)^T\right\rangle_F \leq \left\|(A^Q)^T \bullet^Q A^Q\right\|_F\left\|\Psi^n \bullet^n (\Psi^n)^T\right\|_F,
\end{aligned}$$

where $\langle\cdot,\cdot\rangle_F$ denotes the Frobenius inner product. By observing that the matrix $(A^Q)^T \bullet^Q A^Q$ is positive semi-definite, we obtain

$$\begin{aligned}
\left\|(A^Q)^T \bullet^Q A^Q\right\|_F &= \sqrt{\sum_{i,j=1}^{\alpha}\left|((A^Q)^T \bullet^Q A^Q)_{ij}\right|^2} \leq \sqrt{\sum_{i,j=1}^{\alpha}\left(((A^Q)^T \bullet^Q A^Q)_{ii}((A^Q)^T \bullet^Q A^Q)_{jj}\right)}\\
&= tr((A^Q)^T \bullet^Q A^Q) = \left\|A^Q\right\|^2_{\mathcal{M}^Q_{\beta\times\alpha}},
\end{aligned}$$

and, analogously,

$$\left\|\Psi^n \bullet^n (\Psi^n)^T\right\|_F \leq \left\|\Psi^n\right\|^2_{\mathcal{V}^{\{n\}}_\alpha}.$$

From the above inequalities it follows that

$$\left\|A^Q \otimes \Psi^n\right\|^2_{\mathcal{V}^{Q\cup\{n\}}_\beta} \leq \left\|A^Q \bullet^Q (A^Q)^T\right\|_F\left\|\Psi^n \bullet^n (\Psi^n)^T\right\|_F \leq \left\|A^Q\right\|^2_{\mathcal{M}^Q_{\alpha\times\beta}}\left\|\Psi^n\right\|^2_{\mathcal{V}^{\{n\}}_\alpha},$$

which completes the proof of the first inequality.

(2) Let $\Psi^n = [\psi_1^{(n)}, ..., \psi_\alpha^{(n)}]^T$, $\Phi^Q = [\phi_1^{(Q)}, ..., \phi_\beta^{(Q)}]^T$. We have from the definition (3.3),

$$
\begin{aligned}
\left\| \Psi^n \bullet^n (\Phi^Q)^T \right\|_{\mathcal{M}_{\alpha \times \beta}^{Q \setminus \{n\}}}^2 &= \sum_{i=1}^{\alpha} \sum_{j=1}^{\beta} \left\| \left(\Psi^n \bullet^n (\Phi^Q)^T \right)_{ij} \right\|_{\mathcal{H}^{Q \setminus \{n\}}}^2 \\
&= \sum_{i=1}^{\alpha} \sum_{j=1}^{\beta} \left\| \left\langle \psi_i^{(n)}, \phi_j^{(Q)} \right\rangle_{\mathcal{H}^{\{n\}}} \right\|_{\mathcal{H}^{Q \setminus \{n\}}}^2 \\
&\leq \sum_{i=1}^{\alpha} \sum_{j=1}^{\beta} \left\| \left\| \psi_i^{(n)} \right\|_{\mathcal{H}^{\{n\}}} \left\| \phi_j^{(Q)} \right\|_{\mathcal{H}^Q} \right\|_{\mathcal{H}^{Q \setminus \{n\}}}^2 \\
&= \sum_{i=1}^{\alpha} \sum_{j=1}^{\beta} \left\| \psi_i^{(n)} \right\|_{\mathcal{H}^{\{n\}}}^2 \left\| \phi_j^{(Q)} \right\|_{\mathcal{H}^Q}^2 = \left\| \Psi^n \right\|_{\mathcal{V}_\alpha^{\{n\}}}^2 \left\| \Phi^Q \right\|_{\mathcal{V}_\beta^Q}^2 . \qquad \square
\end{aligned}
$$

Unfolding a tensor. Let $\mathcal{Y} \in \mathcal{T}_P^Q$, $\alpha := \prod_{n=1}^{N} s_n$, $\beta_n := \prod_{k \neq n} s_k$. We denote by $\mathrm{vec}(\mathcal{Y}) \in \mathcal{V}_\alpha^Q$ the vector that carries the entries $y_{J_P}^{(Q)}$ of $\mathcal{Y}$ in lexicographical order, and by

$$
Y_{(n)} = [\mathcal{Y}]_{(n)} \in \mathcal{M}_{s_n \times \beta_n}^Q
$$

the nth unfolding matrix of the tensor $\mathcal{Y}$, *i.e.*, the matrix that aligns the entries $y_{J_P}^{(Q)}$ of $\mathcal{Y}$ with fixed j_n in the j_nth row of the matrix, ordered lexicographically. Clearly the tensor $\mathcal{Y}$ can be reshaped from its unfolding $Y_{(n)}$. Note that if $n \notin P$, then $s_n = 1$ and hence $Y_{(n)} = \mathrm{vec}(\mathcal{Y})^T$. We further note

$$
\|\mathcal{Y}\|_{\mathcal{T}_P^Q} = \|\mathrm{vec}(\mathcal{Y})\|_{\mathcal{V}_\alpha^Q} , \tag{3.9}
$$

and, for each $n = 1, ..., N$,

$$
\|\mathcal{Y}\|_{\mathcal{T}_P^Q} = \left\| Y_{(n)} \right\|_{\mathcal{M}_{s_n \times \beta_n}^Q} = \left\| Y_{(n)}^T \right\|_{\mathcal{M}_{\beta_n \times s_n}^Q} . \tag{3.10}
$$

Tensor products. Let $\mathcal{Y} \in \mathcal{T}_P^Q$, $\Psi^n \in \mathcal{V}_{r_n}^{\{n\}}$. For $n \in P \setminus Q$ we define the nth-mode raising product $\mathcal{Y} \times_n \Psi^n \in \mathcal{T}_{P \setminus \{n\}}^{Q \cup \{n\}}$ by the relation

$$
[\mathcal{Y} \times_n \Psi^n]_{(n)} := (\Psi^n)^T \otimes Y_{(n)} , \tag{3.11}
$$

that is, $\mathcal{Z} = \mathcal{Y} \times_n \Psi^n$ is the tensor of functions given by summing over the nth index *via*

$$
z_{j_1 ... j_{n-1} j_{n+1} ... j_N}(x^{(Q)}, x^{(n)}) = \sum_{\ell_n=1}^{r_n} \psi_{\ell_n}^n(x^{(n)}) \, y_{j_1 ... j_{n-1} \ell_n j_{n+1} ... j_N}(x^{(Q)}).
$$

For $n \in Q \setminus P$ we define the nth-mode lowering product $\mathcal{Y} \circ^n \Psi^n \in \mathcal{T}_{P \cup \{n\}}^{Q \setminus \{n\}}$ by the relation

$$
[\mathcal{Y} \circ^n \Psi^n]_{(n)} := \Psi^n \bullet^n Y_{(n)} , \tag{3.12}
$$

that is, $\mathcal{W} = \mathcal{Y} \circ^n \Psi^n$ is the tensor of functions obtained by integrating out over $x^{(n)}$ *via*

$$
w_{j_1 ... j_N}(x^{(Q \setminus \{n\})}) = \int_{\mathbb{R}^{d_n}} \overline{\psi_{j_n}^n(x^{(n)})} \, y_{j_1 ... j_{n-1} j_{n+1} ... j_N}(x^{(Q \setminus \{n\})}, x^{(n)}) \, \mathrm{d}x^{(n)}.
$$

Highlight

Lemma 3.2. *In the situation of the previous definition, we have*

$$\left\| \mathcal{Y} \times_n \Psi^n \right\|_{\mathcal{T}_{P \setminus \{n\}}^{Q \cup \{n\}}} \leq \left\| \Psi^n \right\|_{\mathcal{V}_{r_n}^{\{n\}}} \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^Q}$$

$$\left\| \mathcal{Y} \circ^n \Psi^n \right\|_{\mathcal{T}_{P \cup \{n\}}^{Q \setminus \{n\}}} \leq \left\| \Psi^n \right\|_{\mathcal{V}_{r_n}^{\{n\}}} \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^Q} .$$

Proof. By the definition (3.11) and the properties (3.10), (3.7) we obtain

$$\left\| \mathcal{Y} \times_n \Psi^n \right\|_{\mathcal{T}_{P \setminus \{n\}}^{Q \cup \{n\}}} = \left\| Y_{(n)}^T \otimes \Psi^n \right\|_{\mathcal{V}_{\beta_n}^{Q \cup \{n\}}} \leq \left\| \Psi^n \right\|_{\mathcal{V}_{r_n}^{\{n\}}} \left\| Y_{(n)}^T \right\|_{M_{r_n \times \beta_n}^Q} = \left\| \Psi^n \right\|_{\mathcal{V}_{r_n}^{\{n\}}} \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^Q} ,$$

where $\beta_n = \prod_{k \neq n} s_k$, and thus the first inequality is proved.

By the definition (3.12) and the properties (3.9), (3.8) we obtain

$$\left\| \mathcal{Y} \circ^n \Psi^n \right\|_{\mathcal{T}_{P \cup \{n\}}^{Q \setminus \{n\}}} = \left\| \Psi^n \bullet^n \mathrm{vec}(\mathcal{Y})^T \right\|_{M_{r_n \times \beta_n}^{Q \setminus \{n\}}} \leq \left\| \Psi^n \right\|_{\mathcal{V}_{r_n}^{\{n\}}} \left\| \mathrm{vec}(\mathcal{Y}) \right\|_{\mathcal{V}_{\beta_n}^{\{n\}}} = \left\| \Psi^n \right\|_{\mathcal{V}_{r_n}^{\{n\}}} \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^Q} ,$$

which concludes the proof.   □

For a subset $K = \{k_1, \ldots, k_m\}$ of P with $K \cap Q = \emptyset$ we denote

$$\mathcal{Y} \mathop{\mathsf{X}}_{k \in K} \Psi^k = \mathcal{Y} \times_{k_1} \Psi^{k_1} \ldots \times_{k_m} \Psi^{k_m} \in \mathcal{T}_{P \setminus K}^{Q \cup K} ,$$

and similarly, for a subset $L = \{\ell_1, \ldots, \ell_n\}$ of Q with $L \cap P = \emptyset$ we write

$$\mathcal{Y} \mathop{\bigcirc}_{\ell \in L} \Psi^\ell = \mathcal{Y} \circ^{\ell_1} \Psi^{\ell_1} \ldots \circ^{\ell_n} \Psi^{\ell_n} \in \mathcal{T}_{P \cup L}^{Q \setminus L} .$$

Tensor products with vectors of orthogonal functions. In the following, let $\Phi^n \in \mathcal{V}_{r_n}^{\{n\}}$ be such that the components are mutually orthogonal functions:

$$\Phi^n \bullet^n (\Phi^n)^T = I_{r_n}, \quad n = 1, \ldots, N, \tag{3.13}$$

where I_{r_n} is the identity matrix of dimension r_n. For $\mathcal{Y} \in \mathcal{T}_P^{\{n\}}$ with $n \notin P$, we consider the projections

$$\mathcal{P}^n(\mathcal{Y}) = (\mathcal{Y} \circ^n \Phi^n) \times_n \Phi^n, \quad \mathcal{P}_\perp^n(\mathcal{Y}) = \mathcal{Y} - \mathcal{P}^n(\mathcal{Y}). \tag{3.14}$$

Lemma 3.3. *Let $\mathcal{Y} \in \mathcal{T}_P^Q$. Then,*

(1) *For $K \subseteq P$, $K \cap Q = \emptyset$,*

$$\left\| \mathcal{Y} \mathop{\mathsf{X}}_{k \in K} \Phi^k \right\|_{\mathcal{T}_{P \setminus K}^{Q \cup K}} = \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^Q} .$$

(2) *For $L \subseteq Q$, $P \cap L = \emptyset$,*

$$\left\| \mathcal{Y} \mathop{\bigcirc}_{\ell \in L} \Phi^\ell \right\|_{\mathcal{T}_{P \cup L}^{Q \setminus L}} \leq \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^Q} .$$

(3) *For $Q = \{n\}$, $n \notin P$,*

$$\left\| \mathcal{P}^n(\mathcal{Y}) \right\|_{\mathcal{T}_P^{\{n\}}} \leq \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^{\{n\}}}, \quad \left\| \mathcal{P}_\perp^n(\mathcal{Y}) \right\|_{\mathcal{T}_P^{\{n\}}} \leq \left\| \mathcal{Y} \right\|_{\mathcal{T}_P^{\{n\}}} .$$

Proof. (1) By the definition of norm (3.1) and by the orthonormality conditions (3.13) we obtain

$$\left\| \mathcal{Y} \underset{k\in K}{\mathsf{X}} \Phi^k \right\|^2_{\mathcal{T}^{Q\cup K}_{P\setminus K}} = \sum_{J_{P\setminus K}} \left\| \sum_{J_K} y^{(Q)}_{J_P} \bigotimes_{k\in K} \phi^{(k)}_{j_k} \right\|^2_{\mathcal{H}^{Q\cup K}} = \sum_{J_{P\setminus K}} \sum_{J_K} \left\| y^{(Q)}_{J_P} \right\|^2_{\mathcal{H}^Q} = \|\mathcal{Y}\|^2_{\mathcal{T}^Q_P}$$

(2) Let $\mathcal{X} \in \mathcal{T}^Q_P$ defined by having elements

$$x^{(Q)}_{J_P} = \sum_{J_L} \left\langle \bigotimes_{\ell\in L} \phi^{(k)}_{j_\ell}, y^{(Q)}_{J_P} \right\rangle_{\mathcal{H}^L} \bigotimes_{\ell\in L} \phi^{(k)}_{j_\ell} \in \mathcal{H}^Q.$$

From the orthonormality conditions (3.13) it follows

$$\left\| x^{(Q)}_{J_P} \right\|^2_{\mathcal{H}^Q} = \sum_{J_L} \left\| \left\langle \bigotimes_{\ell\in L} \phi^{(k)}_{j_\ell}, y^{(Q)}_{J_P} \right\rangle_{\mathcal{H}^L} \right\|^2_{\mathcal{H}^{Q\setminus L}},$$

i.e.,

$$\|\mathcal{X}\|_{\mathcal{T}^Q_P} = \left\| \mathcal{Y} \underset{\ell\in L}{\bigcirc} \Phi^\ell \right\|_{\mathcal{T}^{Q\setminus L}_{P\cup L}},$$

and

$$\langle \mathcal{X}, \mathcal{Y} \rangle_{\mathcal{T}^Q_P} = \langle \mathcal{Y}, \mathcal{X} \rangle_{\mathcal{T}^Q_P} = \left\| \mathcal{Y} \underset{\ell\in L}{\bigcirc} \Phi^\ell \right\|^2_{\mathcal{T}^{Q\setminus L}_{P\cup L}}.$$

Thus it follows

$$\left\| \mathcal{Y} \underset{\ell\in L}{\bigcirc} \Phi^\ell \right\|^2_{\mathcal{T}^{Q\setminus L}_{P\cup L}} = \|\mathcal{Y}\|^2_{\mathcal{T}^Q_P} - \|\mathcal{Y} - \mathcal{X}\|^2_{\mathcal{T}^Q_P} \leq \|\mathcal{Y}\|^2_{\mathcal{T}^Q_P}.$$

(3) It follows from the definition (3.14) of $\mathcal{P}^n_\perp$ that the elements of $\mathcal{P}^n_\perp(\mathcal{Y}) \in \mathcal{T}^{\{n\}}_P$ have norm

$$\left\| (\mathcal{P}^n_\perp(\mathcal{Y}))_{J_P} \right\|^2_{\mathcal{H}^{\{n\}}} = \left\| y^{(n)}_{J_P} \right\|^2_{\mathcal{H}^{\{n\}}} - \sum_{j_n} \left| \left\langle \phi^{(n)}_{j_n}, y^{(n)}_{J_P} \right\rangle_{\mathcal{H}^{\{n\}}} \right|^2$$

$$\|\mathcal{P}^n_\perp(\mathcal{Y})\|^2_{\mathcal{T}^{\{n\}}_P} = \sum_{J_P} \left\| (\mathcal{P}^n_\perp(\mathcal{Y}))_{J_P} \right\|^2_{\mathcal{H}^{\{n\}}} = \|\mathcal{Y}\|^2_{\mathcal{T}^{\{n\}}_P} - \|\mathcal{Y} \circ^n \Phi^n\|^2_{\mathcal{T}^\emptyset_{P\cup\{n\}}} \leq \|\mathcal{Y}\|^2_{\mathcal{T}^{\{n\}}_P}. \qquad \square$$

4. REFORMULATION OF THE MCTDH METHOD

The MCTDH manifold and its tangent space. With the notation of the previous section, we can rewrite the MCTDH manifold $\mathcal{M}$ defined by (1.17) as

$$\mathcal{M} = \Big\{ u \in L^2(\mathbb{R}^d) : u = \mathcal{A} \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n, \quad \Phi^n \in \mathcal{V}^{\{n\}}_{r_n}, \quad \Phi^n \bullet^n (\Phi^n)^T = I_{r_n}, \quad n = 1, ..., N, \tag{4.1}$$
$$\mathcal{A} \in \mathbb{C}^{r_1 \times \cdots \times r_N} \text{ has full rank } (r_1, \ldots, r_N) \Big\}.$$

We recall that $\mathcal{V}^{\{n\}}_{r_n}$ denotes the space of column vectors of length r_n of functions of the variable $x^{(n)}$. In the definition of the manifold $\mathcal{M}$ we may as well just require that for each n, the components of $\Phi^n = (\phi^{(n)}_{j_n})$ are linearly independent functions, since we can then always do an orthogonalisation without changing $u \in \mathcal{M}$.

Highlight

The tangent space at $u \in \mathcal{M}$ consists of the elements $\dot{u}$ of the form

$$\dot{u} = \dot{\mathcal{A}} \mathop{\mathsf{X}}_{n=1}^{N} \Phi^n + \sum_{n=1}^{N} \mathcal{A} \times_n \dot{\Phi}^n \mathop{\mathsf{X}}_{k \neq n} \Phi^k, \tag{4.2}$$

with arbitrary $\dot{\mathcal{A}} \in \mathbb{C}^{r_1 \times \cdots \times r_N}$ and $\dot{\Phi}^n \in \mathcal{V}_{r_n}^{\{n\}}$. We obtain that $\dot{\mathcal{A}}$ and $\dot{\Phi}^n$ are uniquely determined by $\dot{u}$ and the chosen $\mathcal{A}$ and Φ^n in the representation of u if we impose the gauge conditions (1.18), viz.,

$$\Phi^n \bullet^n (\dot{\Phi}^n)^T = 0. \tag{4.3}$$

In fact we obtain, by multiplying (4.2) by $\mathop{\bigcirc}\limits_{n=1}^{N} \Phi^n$ and by $\mathop{\bigcirc}\limits_{k \neq n} \Phi^k$,

$$\begin{cases} \dot{\mathcal{A}} &= \dot{u} \mathop{\bigcirc}\limits_{n=1}^{N} \Phi^n \\ (\dot{\Phi}^n)^T &= \left[\mathcal{P}_\perp^n \left(\dot{u} \mathop{\bigcirc}\limits_{k \neq n} \Phi^k \right) \right]_{(n)} A_{(n)}^\dagger, \end{cases} \tag{4.4}$$

where the projection $\mathcal{P}_\perp^n$ is defined by (3.14) with the above Φ^n, and $A_{(n)}^\dagger = A_{(n)}^* (A_{(n)} A_{(n)}^*)^{-1}$ is the pseudo-inverse of the nth unfolding matrix $A_{(n)}$ of $\mathcal{A}$. Note that the second equation in (4.4) is equivalent to

$$\mathcal{A} \times_n \dot{\Phi}^n = \mathcal{P}_\perp^n \left(\dot{u} \mathop{\bigcirc}\limits_{k \neq n} \Phi^k \right),$$

since $[\mathcal{A} \times_n \dot{\Phi}^n]_{(n)} = (\dot{\Phi}^n)^T A_{(n)}$.

The MCTDH equations of motion. The variational approximation (1.5) on the MCTDH manifold (4.1) is given by

$$u(t) = \mathcal{A}(t) \mathop{\mathsf{X}}_{n=1}^{N} \Phi^n(t),$$

where the coefficients and the single-particle functions are solutions of the system of differential equations (1.19)–(1.20). In the present notation these equations turn out to simplify to

$$\begin{cases} \dot{\mathcal{A}} &= \frac{1}{\mathrm{i}} H u \mathop{\bigcirc}\limits_{n=1}^{N} \Phi^n \\ (\dot{\Phi}^n)^T &= \left[\mathcal{P}_\perp^n \left(\frac{1}{\mathrm{i}} H u \mathop{\bigcirc}\limits_{k \neq n} \Phi^k \right) \right]_{(n)} A_{(n)}^\dagger, \end{cases} \tag{4.5}$$

which are formally obtained by equations (4.4) with $\dot{u}$ replaced by $\frac{1}{\mathrm{i}} H u$.

Rederivation of the MCTDH equations. For the convenience of the reader, we give the derivation of equations (4.5) directly from the Dirac-Frenkel principle (1.5). Choosing in (1.5) $\delta u = \mathsf{S} \mathop{\mathsf{X}}\limits_{n=1}^{N} \Phi^n \in \mathcal{T}_u \mathcal{M}$ with arbitrary $\mathsf{S} \in \mathbb{C}^{r_1 \times \cdots \times r_N}$, and using (4.4), we have

$$\left\langle \mathsf{S} \mathop{\mathsf{X}}_{n=1}^{N} \Phi^n \,\middle|\, \dot{u} \right\rangle = \left\langle \mathsf{S}, \dot{u} \mathop{\bigcirc}\limits_{n=1}^{N} \Phi^n \right\rangle = \langle \mathsf{S}, \dot{\mathcal{A}} \rangle,$$

$$\left\langle \mathsf{S} \mathop{\mathsf{X}}_{n=1}^{N} \Phi^n \,\middle|\, \frac{1}{\mathrm{i}} H u \right\rangle = \left\langle \mathsf{S}, \frac{1}{\mathrm{i}} H u \mathop{\bigcirc}\limits_{n=1}^{N} \Phi^n \right\rangle.$$

Since this holds for every $\mathbf{S} \in \mathbb{C}^{r_1 \times \cdots \times r_N}$, the first equation of (4.5) follows. We now choose $\delta u = \mathcal{A} \times_n \Lambda^n \underset{k \neq n}{\mathsf{X}} \Phi^k$ in $\mathcal{T}_u \mathcal{M}$, where $\Lambda^n \in \mathcal{V}_{r_n}^{\{n\}}$ is arbitrary. We then have

$$\left\langle \mathcal{A} \times_n \Lambda^n \underset{k \neq n}{\mathsf{X}} \Phi^k \,\middle|\, \dot{u} \right\rangle = \left\langle \mathcal{A} \times_n \Lambda^n , \dot{u} \underset{k \neq n}{\bigcirc} \Phi^k \right\rangle_{\mathcal{T}_{\{\neg n\}}^{\{n\}}} = \left\langle \mathcal{A} \times_n \Lambda^n , \dot{\mathcal{A}} \times_n \Phi^n + \mathcal{A} \times_n \dot{\Phi}^n \right\rangle_{\mathcal{T}_{\{\neg n\}}^{\{n\}}} , \qquad (4.6)$$

where we have used (4.2) and (4.3) in the second equality, and

$$\left\langle \mathcal{A} \times_n \Lambda^n \underset{k \neq n}{\mathsf{X}} \Phi^k \,\middle|\, \tfrac{1}{\mathrm{i}} H u \right\rangle = \left\langle \mathcal{A} \times_n \Lambda^n , \tfrac{1}{\mathrm{i}} H u \underset{k \neq n}{\bigcirc} \Phi^k \right\rangle_{\mathcal{T}_{\{\neg n\}}^{\{n\}}} . \qquad (4.7)$$

Then, by subtracting the two equations (4.6) and (4.7), using the first equation of (4.5) and recalling the definition (3.14) of the projection $\mathcal{P}_\perp^n$, we obtain

$$\left\langle \mathcal{A} \times_n \Lambda^n , \mathcal{P}_\perp^n \left(\tfrac{1}{\mathrm{i}} H u \underset{k \neq n}{\bigcirc} \Phi^k \right) - \mathcal{A} \times_n \dot{\Phi}^n \right\rangle_{\mathcal{T}_{\{\neg n\}}^{\{n\}}} = 0.$$

Unfolding the tensors in $\mathcal{T}_{\{\neg n\}}^{\{n\}}$ to vectors in $\mathcal{V}_{\beta_n}^{\{n\}}$ with $\beta_n = \prod_{k \neq n} r_k$, this is the same as the inner product

$$\left\langle (\Lambda^n)^T A_{(n)}, \left[\mathcal{P}_\perp^n \left(\tfrac{1}{\mathrm{i}} H u \underset{k \neq n}{\bigcirc} \Phi^k \right) - \mathcal{A} \times_n \dot{\Phi}^n \right]_{(n)} \right\rangle_{\mathcal{V}_{\beta_n}^{\{n\}}} = 0,$$

from which

$$\left\langle (\Lambda^n)^T , \left[\mathcal{P}_\perp^n \left(\tfrac{1}{\mathrm{i}} H u \underset{k \neq n}{\bigcirc} \Phi^k \right) \right]_{(n)} A_{(n)}^* - (\dot{\Phi}^n)^T A_{(n)} A_{(n)}^* \right\rangle_{\mathcal{V}_{r_n}^{\{n\}}} = 0,$$

with arbitrary $\Lambda^n \in \mathcal{V}_{r_n}^{\{n\}}$, and then the second equation of (4.5) follows. This completes the derivation of the MCTDH equations (4.5).

Reformulation of (2.3) and (2.6). In (2.3), we have the decomposition

$$v(t) = \mathcal{B}(t) \underset{n=1}{\overset{N}{\mathsf{X}}} \Theta^n(t), \qquad 0 \leq t \leq \overline{t}, \qquad (4.8)$$

with $\Theta^n(t) \in \mathcal{V}_{r_n}^{\{n\}}$ satisfying the orthogonality relations $\Theta^n \bullet^n (\Theta^n)^T = I_{r_n}$. Condition (2.6) can be rephrased as

$$\left\| \mathrm{i} \dot{\Theta}^n(t) - T^{(n)} \Theta^n(t) \right\|^{(n)} \leq \nu.$$

5. TANGENT SPACE PROJECTION

The orthogonal projection onto the tangent space of the MCTDH manifold $\mathcal{M}$ at u is given by the following explicit formula.

Lemma 5.1. *Let $u = \mathcal{A} \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n \in \mathcal{M}$. Then, the orthogonal projection $P(u)$ onto the tangent space $\mathcal{T}_u \mathcal{M}$ of $\varphi \in \mathcal{H}$ is given by*

$$P(u)\varphi = \left(\varphi \underset{n=1}{\overset{N}{\bigcirc}} \Phi^n \right) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n + \sum_{n=1}^{N} \mathcal{Q}^n \underset{k \neq n}{\mathsf{X}} \Phi^k,$$

where $\mathcal{Q}^n \in \mathcal{T}_{\{\neg n\}}^{\{n\}}$ is given by its nth unfolding matrix $Q_n = [\mathcal{Q}^n]_{(n)}$ as

$$Q_n = \left[\mathcal{P}_\perp^n \left(\varphi \underset{k \neq n}{\bigcirc} \Phi^k \right) \right]_{(n)} A_{(n)}^\dagger A_{(n)}.$$

(sidebar) **Highlight**

Proof. The equations of motion (4.5) give an expression of $\dot{u}$ determined by $\dot{u} = P(u)\frac{1}{i}Hu$. With φ in place of $\frac{1}{i}Hu$, by noting that the second equation of (4.5) is equivalent to

$$\left[\mathcal{A} \times_n \dot{\Phi}^n\right]_{(n)} = \left[\mathcal{P}^n_\perp\left(\varphi \bigcirc_{k\neq n} \Phi^k\right)\right]_{(n)} A^\dagger_{(n)} A_{(n)},$$

the result immediately follows. $\square$

We denote by $P^\perp(u) = \mathrm{Id} - P(u)$ the complementary projection. The following lemma is analogous to bounds of tangent space projections of manifolds of low-rank matrices and tensors given in [6,7].

Lemma 5.2. *Let* $v = \mathcal{B} \overset{N}{\underset{n=1}{\mathsf{X}}} \Theta^n \in \mathcal{M}$ *with smallest singular value* $\sigma_{r_n}(B_{(n)}) \geq \delta > 0$ *for* $n = 1, ..., N$. *Let* $u \in \mathcal{M}$ *such that* $\|u - v\| \leq c\delta$ *with* $c = \frac{1}{2C}$, $C = 8N(N+3)$. *Then, for all* $\varphi \in \mathcal{H}$, *we have*

$$\left\|\left(P(u) - P(v)\right)\varphi\right\| \leq C\delta^{-1}\|u - v\|\|\varphi\| \tag{5.1}$$
$$\left\|P^\perp(v)\left(u - v\right)\right\| \leq C\delta^{-1}\|u - v\|^2. \tag{5.2}$$

We note that (5.1)–(5.2) further yield the bound

$$\left\|P^\perp(u)\left(u - v\right)\right\| \leq 2C\delta^{-1}\|u - v\|^2. \tag{5.3}$$

Proof. We decompose the functions on the straight line connecting v and u as

$$v + \tau\left(u - v\right) = y(\tau) + z(\tau) \tag{5.4}$$

with $y(\tau) \in \mathcal{M}$ and $z(\tau) \perp \mathcal{T}_v\mathcal{M}$. This decomposition exists at least for small τ but the argument below shows that in fact it exists for $0 \leq \tau \leq 1$. Let us denote

$$w = P(v)(u - v) \in \mathcal{T}_v\mathcal{M}.$$

Applying $P(v)$ to (5.4), we obtain

$$\tau w = P(v)(y(\tau) - v),$$

and then

$$\frac{\partial y}{\partial \tau}(\tau) = P(y(\tau))\frac{\partial y}{\partial \tau}(\tau) = (P(y(\tau)) - P(v))\frac{\partial y}{\partial \tau}(\tau) + P(v)\frac{\partial y}{\partial \tau}(\tau) = (P(y(\tau)) - P(v))\frac{\partial y}{\partial \tau}(\tau) + w.$$

So, as long as the operator norm of $P(y(\tau)) - P(v)$ is bounded by $1/2$, we have, with $\omega = \|w\|$,

$$\left\|\frac{\partial y}{\partial \tau}(\tau)\right\| \leq 2\omega. \tag{5.5}$$

Let $y(\tau) \in \mathcal{M}$ have the factorization

$$y(\tau) = \mathcal{A}(\tau) \overset{N}{\underset{n=1}{\mathsf{X}}} \Phi^n(\tau)$$

with $\Phi^n \bullet^n \left(\frac{\partial \Phi^n}{\partial \tau}\right)^T = 0$. Then, by (4.4), we have

$$\begin{cases} \dfrac{\mathrm{d}\mathcal{A}}{\mathrm{d}\tau} = \dfrac{\partial y}{\partial \tau} \overset{N}{\underset{n=1}{\bigcirc}} \Phi^n(\tau) \\[2ex] \mathcal{A}(\tau) \times_n \dfrac{\partial \Phi^n}{\partial \tau} = \mathcal{P}^n_\perp\left(\dfrac{\partial y}{\partial \tau} \underset{k\neq n}{\bigcirc} \Phi^k(\tau)\right). \end{cases}$$

Hence, by Lemma 3.3 and (5.5), we obtain

$$\left\|\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}\tau}\right\| \le 2\omega$$

$$\left\|\mathcal{A}(\tau)\times_n\frac{\partial\Phi^n}{\partial\tau}\right\|^{(n)} \le 2\omega.$$

Moreover, since

$$\|\mathcal{A}(\tau) - \mathcal{B}\| = \|\mathcal{A}(\tau) - \mathcal{A}(0)\| \le 2\omega \le 2c\delta \le \tfrac{1}{2}\delta, \qquad 0 \le \tau \le 1,$$

it follows from Lemma 5.3 below that $\|A^\dagger_{(n)}(\tau)\|_2 \le 2\delta^{-1}$, and then, using (3.10),

$$\left\|\frac{\partial\Phi^n}{\partial\tau}\right\|^{(n)} \le \left\|\frac{\partial\Phi^n}{\partial\tau}^T A_{(n)}\right\|^{(n)}\left\|A^\dagger_{(n)}\right\|_2 \le 4\omega\,\delta^{-1}. \tag{5.6}$$

By Lemma 5.1 and writing $\|(P(y(\tau)) - P(v))\,\varphi\| = \left\|\int_0^\tau\left(\frac{\partial}{\partial\sigma}P\left(y(\sigma)\right)\varphi\right)\mathrm{d}\sigma\right\|$, we obtain

$$\|(P(y(\tau)) - P(v))\,\varphi\| \le \int_0^\tau\left\|\frac{\partial}{\partial\sigma}\left[\left(\varphi\mathop{\bigcirc}_{n=1}^N\Phi^n(\sigma)\right)\mathop{\mathsf{X}}_{n=1}^N\Phi^n(\sigma)\right]\right\|\mathrm{d}\sigma$$

$$+ \sum_{n=1}^N\int_0^\tau\left\|\frac{\partial\mathcal{Q}^n}{\partial\sigma}\mathop{\mathsf{X}}_{k\neq n}\Phi^k(\sigma) + \sum_{\substack{j=1\\j\neq n}}^N\left[\mathcal{Q}^n(\sigma)\times_j\frac{\partial\Phi^j}{\partial\sigma}\mathop{\mathsf{X}}_{k\neq n,j}\Phi^k(\sigma)\right]\right\|\mathrm{d}\sigma, \tag{5.7}$$

where

$$[\mathcal{Q}^n(\sigma)]_{(n)} = \left[\mathcal{P}^n_\perp\left(\varphi\mathop{\bigcirc}_{k\neq n}\Phi^k(\sigma)\right)\right]_{(n)}A^\dagger_{(n)}(\sigma)A_{(n)}(\sigma).$$

By Lemmas 3.2–3.3 and by (5.6) we obtain

$$\left\|\frac{\partial}{\partial\sigma}\left[\left(\varphi\mathop{\bigcirc}_{n=1}^N\Phi^n(\sigma)\right)\mathop{\mathsf{X}}_{n=1}^N\Phi^n(\sigma)\right]\right\| \le \sum_{n=1}^N\left\|\left(\varphi\circ^n\frac{\partial\Phi^n}{\partial\sigma}\mathop{\bigcirc}_{k\neq n}\Phi^k\right)\mathop{\mathsf{X}}_{k=1}^N\Phi^k(\sigma)\right\|$$

$$+ \sum_{n=1}^N\left\|\left(\varphi\mathop{\bigcirc}_{k=1}^N\Phi^k(\sigma)\right)\times_n\frac{\partial\Phi^n}{\partial\sigma}\mathop{\mathsf{X}}_{k\neq n}\Phi^k(\sigma)\right\|$$

$$\le 8N\omega\,\delta^{-1}\,\|\varphi\|,$$

and analogously we obtain

$$\left\|\frac{\partial}{\partial\sigma}\mathcal{P}^n_\perp\left(\varphi\mathop{\bigcirc}_{k\neq n}\Phi^k(\sigma)\right)\right\| \le 4(N+1)\omega\,\delta^{-1}\,\|\varphi\|.$$

Using the product rule for $\frac{\partial}{\partial\sigma}\left(A^*_{(n)}(A_{(n)}A^*_{(n)})^{-1}A_{(n)}\right)$ and the estimates of $\left\|\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}\tau}\right\|$ and $\left\|A^\dagger_{(n)}\right\|$ by 2ω and $2\delta^{-1}$ respectively, we find that

$$\left\|\frac{\partial}{\partial\sigma}\left(A^\dagger_{(n)}(\sigma)A_{(n)}(\sigma)\right)\right\| \le 16\omega\delta^{-1}.$$

Hence we obtain

$$\left\|\frac{\partial\mathcal{Q}^n}{\partial\sigma}\right\| = \left\|\left[\frac{\partial\mathcal{Q}^n}{\partial\sigma}\right]_{(n)}\right\| \le \left\|\frac{\partial}{\partial\sigma}\mathcal{P}^n_\perp\left(\varphi\mathop{\bigcirc}_{k\neq n}\Phi^k(\sigma)\right)\right\| + \|\varphi\|\left\|\frac{\partial}{\partial\sigma}\left(A^\dagger_{(n)}(\sigma)A_{(n)}(\sigma)\right)\right\|$$

$$\le 4(N+5)\omega\,\delta^{-1}\,\|\varphi\|.$$

Substituting in (5.7) the above bounds it follows

$$\|(P(y(\tau)) - P(v))\,\varphi\| \leq 8N(N+3)\omega\,\delta^{-1}\tau\,\|\varphi\| = C\omega\delta^{-1}\tau\,\|\varphi\|. \tag{5.8}$$

Thus, as we have $C\omega\delta^{-1} \leq Cc = 1/2$, the norm of the operator $(P(y(\tau)) - P(v))$ does not exceed $1/2$ for $0 \leq \tau \leq 1$.

By considering (5.8) at $\tau = 1$ we obtain (5.1). Using (5.8) and (5.5) we get

$$
\begin{aligned}
P^\perp(v)\,(u-v) &= P^\perp(v)\int_0^1 \frac{\partial y}{\partial \tau}\mathrm{d}\tau = \int_0^1 P^\perp(v)\frac{\partial y}{\partial \tau}\mathrm{d}\tau \\
&= \int_0^1 \left(P^\perp(v) - P^\perp(y(\tau))\right)\frac{\partial y}{\partial \tau}\mathrm{d}\tau = \int_0^1 \left(P(y(\tau)) - P(v)\right)\frac{\partial y}{\partial \tau}\mathrm{d}\tau,
\end{aligned}
$$

and with (5.5) and (5.8) we then obtain

$$\left\|P^\perp(v)\,(u-v)\right\| \leq 2C\omega^2\delta^{-1}\int_0^1 \tau\mathrm{d}\tau = C\delta^{-1}\,\|u-v\|^2,$$

which is the desired bound (5.2). □

It remains to give the linear algebra lemma to which we referred in the above proof.

Lemma 5.3. *Let $\mathcal{A}$, $\mathcal{B} \in \mathbb{C}^{r_1 \times \dots \times r_N}$ be such that the nth unfolding matrix $B_{(n)}$ has smallest singular value $\sigma_{r_n}(B_{(n)}) \geq \delta > 0$ for $n = 1, \dots, N$ and $\|\mathcal{A} - \mathcal{B}\| \leq \frac{1}{2}\delta$. Then, the pseudo-inverse $A^\dagger_{(n)}$ of the nth unfolding matrix of $\mathcal{A}$ is bounded in the spectral norm by $\|A^\dagger_{(n)}\|_2 \leq 2\delta^{-1}$.*

Proof. From the chain of inequalities

$$\left|\sigma_{r_n}(A_{(n)}) - \sigma_{r_n}(B_{(n)})\right| \leq \left\|A_{(n)} - B_{(n)}\right\|_2 \leq \left\|A_{(n)} - B_{(n)}\right\|_F = \|\mathcal{A} - \mathcal{B}\| \leq \tfrac{1}{2}\delta,$$

where the first inequality holds by [3], p. 448, we obtain

$$\sigma_{r_n}(A_{(n)}) \geq \sigma_{r_n}(B_{(n)}) - \left|\sigma_{r_n}(A_{(n)}) - \sigma_{r_n}(B_{(n)})\right| \geq \tfrac{1}{2}\delta,$$

and hence $\|A^\dagger_{(n)}\|_2 = 1/\sigma_{r_n}(A_{(n)}) \leq 2/\delta$. □

6. A FURTHER PROJECTION BOUND

The following lemma, which takes up the arguments of the proof of Lemma 5.2, provides a key estimate for the proof of Theorem 2.1. It bounds the expression

$$P^\perp(u)\big(\dot{v} + \mathrm{i}Tv\big) = \big(P^\perp(u) - P^\perp(v)\big)\big(\dot{v} + \mathrm{i}Tv\big) = -\big(P(u) - P(v)\big)\big(\dot{v} + \mathrm{i}Tv\big),$$

where the first equality results from the fact that both $\dot{v}$ and Tv are in the tangent space $\mathcal{T}_v\mathcal{M}$. Note that in contrast to a direct application of Lemma 5.2, no factor δ^{-1} appears in the following bound.

Lemma 6.1. *Under the assumptions of Theorem 2.1 and as long as $\|u(t) - v(t)\| \leq c\delta$ with c of Lemma 5.2, we have*

$$\left\|P^\perp(u(t))\big(\dot{v}(t) + \mathrm{i}Tv(t)\big)\right\| \leq (c_1\mu + c_2\nu)\,\|u(t) - v(t)\| \tag{6.9}$$

with $c_1 = 3N$ and $c_2 = 2N^2$.

Proof. From the proof of Lemma 5.2 we observe that there exists an homotopy

$$y(t,\tau) = \mathcal{A}(t,\tau) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi_n(t,\tau) \qquad \text{with } y(t,0) = v(t), \ y(t,1) = u(t),$$

such that the orthogonality relations $\Phi^n(t,\tau) \bullet^n \Phi^n(t,\tau)^T = I_{r_n}$ hold for all (t,τ), and

$$\left\| \frac{\partial \mathcal{A}}{\partial \tau}(t,\tau) \right\| \le 2 \left\| u(t) - v(t) \right\|$$

$$\left\| \mathcal{A}(t,\tau) \times_n \frac{\partial \Phi^n}{\partial \tau}(t,\tau) \right\|^{(n)} \le 2 \left\| u(t) - v(t) \right\| \tag{6.10}$$

$$\left\| \frac{\partial \Phi^n}{\partial \tau}(t,\tau) \right\|^{(n)} \le 4 \left\| u(t) - v(t) \right\| \delta^{-1} \le 4c.$$

We thus obtain, with $\Lambda^n(t,\tau) = \dot{\Phi}^n(t,\tau) + iT^{(n)}\Phi^n(t,\tau)$ (where $\dot{} = d/dt$),

$$P^\perp(u(t))\big(\dot{v}(t) + iTv(t)\big) = P^\perp(u(t)) \left(\dot{\mathcal{A}}(t,0) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n(t,0) + \sum_{n=1}^{N} \mathcal{A}(t,0) \times_n \Lambda^n(t,0) \underset{k \ne n}{\mathsf{X}} \Phi^k(t,0) \right.$$

$$\left. - \dot{\mathcal{A}}(t,0) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n(t,1) - \sum_{n=1}^{N} \mathcal{A}(t,1) \times_n \Lambda^n(t,0) \underset{k \ne n}{\mathsf{X}} \Phi^k(t,1) \right),$$

where we could add the last line because it belongs to the tangent space $\mathcal{T}_{u(t)}\mathcal{M}$, see (4.2). We note $\mathcal{A}(t,0) = \mathcal{B}(t)$, and with Lemma 3.3 we obtain the bound

$$\left\| \dot{\mathcal{B}}(t) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n(t,0) - \dot{\mathcal{B}}(t) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n(t,1) \right\| = \left\| \int_0^1 \frac{\partial}{\partial \tau} \left(\dot{\mathcal{B}}(t) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n(t,\tau) \right) d\tau \right\|$$

$$\le \sum_{n=1}^{N} \int_0^1 \left\| \dot{\mathcal{B}}(t) \times_n \frac{\partial \Phi^n}{\partial \tau}(t,\tau) \underset{k \ne n}{\mathsf{X}} \Phi^k(t,\tau) \right\| d\tau \le \sum_{n=1}^{N} \int_0^1 \left\| \dot{\mathcal{B}}(t) \times_n \frac{\partial \Phi^n}{\partial \tau}(t,\tau) \right\|^{(n)} d\tau.$$

Using (3.10) and (3.11), we estimate

$$\left\| \dot{\mathcal{B}} \times_n \frac{\partial \Phi^n}{\partial \tau} \right\|^{(n)} = \left\| \left[\dot{\mathcal{B}} \times_n \frac{\partial \Phi^n}{\partial \tau} \right]_{(n)} \right\|^{(n)} = \left\| \left(\frac{\partial \Phi^n}{\partial \tau} \right)^T \dot{B}_{(n)} \right\|^{(n)}$$

$$\le \left\| \left(\frac{\partial \Phi^n}{\partial \tau} \right)^T B_{(n)} \right\|^{(n)} \| B_{(n)}^\dagger \dot{B}_{(n)} \|_2 = \left\| \mathcal{B} \times_n \frac{\partial \Phi^n}{\partial \tau} \right\|^{(n)} \| B_{(n)}^\dagger \dot{B}_{(n)} \|_2.$$

Writing $\mathcal{B}(t) = \mathcal{A}(t,0)$ and

$$\mathcal{A}(t,0) \times_n \frac{\partial \Phi^n}{\partial \tau}(t,\tau) = \mathcal{A}(t,\tau) \times_n \frac{\partial \Phi^n}{\partial \tau}(t,\tau) - \int_0^\tau \left(\frac{\partial \mathcal{A}}{\partial \sigma}(t,\sigma) \times_n \frac{\partial \Phi^n}{\partial \tau}(t,\tau) \right) d\sigma$$

we have, using Lemma 3.2 and the bounds (6.10),

$$\left\| \mathcal{B}(t) \times_n \frac{\partial \Phi^n}{\partial \tau}(t,\tau) \right\|^{(n)} \le (2 + 8c) \left\| u(t) - v(t) \right\| \le 3 \left\| u(t) - v(t) \right\|.$$

Highlight

Combining the above estimates and using condition (2.5), we thus obtain, with $c_1 = 3N$,

$$\left\| \dot{\mathcal{B}}(t) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n(t,0) - \dot{\mathcal{B}}(t) \underset{n=1}{\overset{N}{\mathsf{X}}} \Phi^n(t,1) \right\| \le c_1 \mu \left\| u(t) - v(t) \right\|.$$

We have analogously, using Lemmas 3.2–3.3 and condition (2.6),

$$\left\| \sum_{n=1}^{N} \left(\mathcal{A}(t,0) \times_n \Lambda^n(t,0) \underset{k \ne n}{\mathsf{X}} \Phi^k(t,0) - \mathcal{A}(t,1) \times_n \Lambda^n(t,0) \underset{k \ne n}{\mathsf{X}} \Phi^k(t,1) \right) \right\|$$

$$\le \sum_{n=1}^{N} \int_0^1 \left\| \frac{\partial}{\partial \tau} \left(\mathcal{A}(t,\tau) \times_n \Lambda^n(t,0) \underset{k \ne n}{\mathsf{X}} \Phi^k(t,\tau) \right) \right\| d\tau \le c_2 \nu \left\| u(t) - v(t) \right\|,$$

with $c_2 = 2N^2$. These bounds yield (6.9).   □

7. Proof of Theorem 2.1

With the estimates of the preceding sections, we are finally in the position to prove Theorem 2.1. We start from (1.6) (omitting the argument t here and in the following),

$$\dot{u} = -P(u)\mathrm{i}Hu.$$

We write, decomposing $\mathrm{Id} = P(u) + P^\perp(u)$ and $H = T + V$,

$$\dot{u} - \dot{v} = -P(u)\mathrm{i}H(u-v) - P(u)\mathrm{i}Hv - P(u)\dot{v} - P^\perp(u)\dot{v}$$
$$= -\mathrm{i}H(u-v) + P^\perp(u)\mathrm{i}T(u-v) + P^\perp(u)\mathrm{i}V(u-v) - P(u)\left(\dot{v} + \mathrm{i}Hv\right) - P^\perp(u)\dot{v}.$$

Since $Tu \in \mathcal{T}_u\mathcal{M}$, we have $P^\perp(u)\mathrm{i}T(u-v) = -P^\perp(u)\mathrm{i}Tv$, and hence we obtain

$$\dot{u} - \dot{v} = -\mathrm{i}H(u-v) + P^\perp(u)\mathrm{i}V(u-v) - P^\perp(u)\left(\dot{v} + \mathrm{i}Tv\right) - P(u)\left(\dot{v} + \mathrm{i}Hv\right).$$

Taking the inner product of $\dot{u} - \dot{v}$ with $u - v$ and considering the real part, we then have

$$\|u - v\| \frac{\mathrm{d}}{\mathrm{d}t} \|u - v\| = \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|u - v\|^2 = \mathrm{Re}\,\langle u - v, \dot{u} - \dot{v} \rangle$$

and

$$\mathrm{Re}\,\langle u - v, \dot{u} - \dot{v} \rangle = \mathrm{Re}\,\left\langle P^\perp(u)\left(u - v\right), \mathrm{i}V(u-v) \right\rangle - \mathrm{Re}\,\left\langle P^\perp(u)\left(u - v\right), P^\perp(u)\left(\dot{v} + \mathrm{i}Tv\right) \right\rangle$$
$$- \mathrm{Re}\,\langle u - v, P(u)\left(\dot{v} + \mathrm{i}Hv\right) \rangle.$$

We now bound the expressions on the right-hand side. Using Lemma 5.2 and the bound (2.7) of the potential V after applying the Cauchy–Schwarz inequality, we obtain

$$\left| \left\langle P^\perp(u)\left(u - v\right), \mathrm{i}V(u-v) \right\rangle \right| \le 2C\delta^{-1}\beta \left\| u - v \right\|^3.$$

Using Lemmas 5.2 and 6.1, we estimate

$$\left| \left\langle P^{\perp}(u)\,(u - v)\,, P^{\perp}(u)(\dot{v} + \mathrm{i}Tv) \right\rangle \right| \leq 2C\delta^{-1}\left(c_1\mu + c_2\nu\right)\left\|u - v\right\|^3.$$

The defect bound (2.2) yields

$$\left| \left\langle u - v, P(u)\,(\dot{v} + \mathrm{i}Hv) \right\rangle \right| \leq \varepsilon\left\|u - v\right\|.$$

In total we thus obtain

$$\operatorname{Re}\left\langle u - v, \dot{u} - \dot{v} \right\rangle \leq 2C\delta^{-1}\left(\beta + c_1\mu + c_2\nu\right)\left\|u - v\right\|^3 + \varepsilon\left\|u - v\right\|,$$

and hence, with $\gamma = 2C\left(\beta + c_1\mu + c_2\nu\right)$, we arrive at the *quadratic* differential inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|u - v\right\| \leq \gamma\delta^{-1}\left\|u - v\right\|^2 + \varepsilon.$$

Therefore, $\left\|u - v\right\|$ is majorized by the solution of the initial-value problem

$$\dot{y} = \gamma\delta^{-1}y^2 + \varepsilon, \qquad y(0) = 0,$$

which equals $y(t) = \sqrt{\frac{\varepsilon}{\gamma\delta^{-1}}}\tan\left(t\sqrt{\gamma\delta^{-1}\varepsilon}\right)$ and is bounded by $2t\varepsilon$ for $t\sqrt{\gamma\delta^{-1}\varepsilon} \leq 1$, that is, for $t \leq \frac{1}{\sqrt{\gamma}}\left(\frac{\delta}{\varepsilon}\right)^{1/2}$. So we obtain

$$\left\|u(t) - v(t)\right\| \leq 2t\varepsilon$$

as long as Lemma 5.2 remains applicable, *i.e.*, as long as $\left\|u(t) - v(t)\right\| \leq c\delta$, which is satisfied for $t \leq \frac{c}{2}\left(\frac{\delta}{\varepsilon}\right)$. The triangle inequality for $u - \psi = (u - v) + (v - \psi)$ finally yields the result.

Highlight

REFERENCES

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

[2] G. Friesecke, The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions. *Arch. Ration. Mech. Anal.* **169** (2003) 35–71.

[3] R.A. Horn and C.R. Johnson, *Matrix Analysis*. Cambridge Univ. Press, UK (1985).

[4] B.N. Khoromskij, Structured rank-$(R_1, ..., R_d)$ decomposition of function-related tensors in $\mathbb{R}^d$. *Comput. Meth. Appl. Math.* **6** (2006) 194–220.

[5] O. Koch and C. Lubich, Regularity of the multi-configuration time-dependent Hartree approximation in quantum molecular dynamics. *ESAIM: M2AN* **41** (2007) 315–331.

[6] O. Koch and C. Lubich, Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.* **29** (2007) 434–454.

[7] O. Koch and C. Lubich, *Dynamical tensor approximation*. Preprint (2009).

[8] T.G. Kolda and B.W. Bader, Tensor decompositions and applications. *SIAM Rev.* **51** (2009) 455–500.

[9] M. Lewin, Solutions of the multiconfiguration equations in quantum chemistry. *Arch. Ration. Mech. Anal.* **171** (2004) 83–114.

[10] C. Lubich, On variational approximations in quantum molecular dynamics. *Math. Comp.* **74** (2005) 765–799.

[11] C. Lubich, *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. Europ. Math. Soc., Zurich, Switzerland (2008).

[12] H.-D. Meyer, F. Gatti and G.A. Worth, Eds., Multidimensional Quantum Dynamics: MCTDH Theory and Applications. Wiley, New York, USA (2009).

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

[14] H.-D. Meyer and G.A. Worth, Quantum molecular dynamics: propagating wavepackets and density operators using the multi-configuration time-dependent Hartree (MCTDH) method. *Theo. Chem. Acc.* **109** (2003) 251–267.

[15] A. Raab, G.A. Worth, H.-D. Meyer and L.S. Cederbaum, Molecular dynamics of pyrazine after excitation to the S_2 electronic state using a realistic 24-mode model Hamiltonian. *J. Chem. Phys.* **110** (1999) 936–946.

[16] B. Thaller, *Visual Quantum Mechanics*. Springer, New York, USA (2000).

[17] H. Wang and M. Thoss, Multilayer formulation of the multiconfiguration time-dependent Hartree theory. *J. Chem. Phys.* **119** (2003) 1289–1299.

Highlight