The coupled-cluster formalism – a mathematical perspective

A. Laestadius & F. M. Faulstich

Published online: 09 Jan 2019.

Submit your article to this journal ⌗

View related articles ⌗

View Crossmark data ⌗

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS   Check for updates

The coupled-cluster formalism – a mathematical perspective

A. Laestadius [ORCID] and F. M. Faulstich

Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Oslo, Blindern, Oslo, Norway

ABSTRACT
The Coupled-Cluster (CC) theory is one of the most successful high precision methods used to solve the stationary Schrödinger equation. In this article, we address the mathematical foundation of this theory with focus on the advances made in the past decade. Rather than solely relying on spectral gap assumptions (non-degeneracy of the ground state), we highlight the importance of coercivity assumptions – Gårding type inequalities – for the local uniqueness of the CC solution. Based on local strong monotonicity, different *sufficient* conditions for a local unique solution are suggested. One of the criteria assumes the relative smallness of the total cluster amplitudes (after possibly removing the single amplitudes) compared to the Gårding constants. In the extended CC theory the Lagrange multipliers are wave function parameters and, by means of the bivariational principle, we here derive a connection between the exact cluster amplitudes and the Lagrange multipliers. This relation might prove useful when determining the quality of a CC solution. Furthermore, the use of an Aubin–Nitsche duality type method in different CC approaches is discussed and contrasted with the bivariational principle.

$$\begin{cases} \mathcal{E}(t) = \langle \phi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \phi_0 \rangle \,, \\ (f(t))_\mu = \langle \phi_\mu | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \phi_0 \rangle = 0 \end{cases}$$

1. Introduction

One of the most successful high accuracy *ab initio* computational schemes is the Coupled-Cluster (CC) approach [1]. It goes back to Coester [2], who in 1958 suggested using an exponential parametrisation of the wave function. This parametrisation was derived independently by Hubbard [3] and Hugenholtz [4] in 1957 as an alternative to summing many-body perturbation theory (MBPT) contributions order by order. At that time, Coester was not able to come up with working equations that one might try to solve. Those were presented by Čížek [5] after the relevant concepts had been introduced in the context of quantum chemistry. In this work, Čížek mentioned the projective approach of the equations, which is exploited in all conventional CC methods until today. Firstly, in [5] the working amplitudes and energy equations were derived when the cluster operator is approximated by merely double excitations (CCD). Secondly, the CC theory was compared with

MBPT, configuration interaction (CI), and the pair cluster expansions of Sinanoğlu [6]. Thirdly, the first ever CCD and linearised CCD computations were reported for nitrogen and a model of benzene. For a more detailed description of the CC history, we refer to reviews by pioneers of the theory. For example, Kümmel [7] and Čížek [8] wrote such articles within the workshop 'Coupled Cluster Theory of Electron Correlation'. Furthermore, see the articles by Bartlett [9], Paldus [10], Arponen [11] and Bishop [12].

Unlike the CI method, the CC formalism does not arise from the Rayleigh–Ritz variational principle and is therefore said to be non-variational in that sense. This yields the well-known fact that the CC energy is in general not equal to the expectation value of the Hamiltonian and in general not an upper bound to the ground-state energy. The reliability of quantum chemical methods is in most cases based on benchmarking, and the results' physical and chemical consistency with existing theory.

The *gold standard of quantum chemistry* – the CCSD(T) method [13,14] – is no exception of this. It is the importance of sharp statements of an *ab initio* method's reliability that is the motivation of this work. Here, we build on a local analysis [15] of the CC theory that also holds in the exact, so-called *continuous*, formulation with infinitely many one-particle basis functions [16,17].

There is a rich history of mathematical investigations addressing CC methods prior to the local analyses in [15–17]. To give a complete historical account is beyond the scope of this article. We therefore limit ourselves and mention only a few important results. As a system of polynomial equations, the CC equations can have real or if the cluster operator is truncated, complex solutions. Furthermore, using quasi-Newton–Raphson methods to compute solutions of non-linear equations can lead to divergence since the approximated Jacobian may become singular. This is, in particular, the case when strongly correlated systems are considered. These and other related aspects of the CC theory have been addressed by Živković and Monkhorst [18,19] and Piecuch *et al.* [20]. Significant advances in the understanding of the nature of multiple solutions of single-reference CC have been made by Živković and Monkhorst [19], Kowalski and Jankowski [21], and by Piecuch and Kowalski [22]. An interesting attempt to address the existence of a cluster operator and cluster expansion in the open-shell case was done by Jeziorski and Paldus [23]. We would also like to mention the coupled-electron pair approximation (CEPA) [24–27]. This approach was introduced as a size-consistent alternative to the CISD method that was achieved by modifying (through topological factors [28]) the CI equations to account for higher excitations. This makes CEPA non-variational (for an adapted variational formulation of CEPA see [29]). CEPA can be regarded as an approximation of the CC method and does not form a truncation hierarchy that converges to the full-CI limit [30].

Mathematical analysis is a well-established part of many natural sciences. Plenty examples show how various fields benefit from mathematical rigor and that mathematical analysis can define a framework of the method's applicability. This work takes off from recent developments of local analyses of CC methods, including the single-reference CC, the extended CC, the tailored CC (TCC) and its special case the CC method tailored by tensor network states (TNS–TCC) [15–17,31,32]. In the spirit of Robert Parr's fundamental approach to quantum chemistry, which was honored during the 58*th* Sanibel Symposium, we here present some mathematical concepts used to analyse CC methods in a functional analytic framework. These yield rigorous analytical results that are independent of benchmarks and interpretations but rather based on mathematical assumptions. Adapting these assumptions to cover the computations performed in practice remains a challenge and is subject of future work. The local analysis puts as a sufficient – but not necessary – condition that the cluster amplitudes are small relative to other constants. We discuss a possible way out of this restriction motivated by the fact that CC calculations are known to work for large (single) amplitudes as well. We furthermore address the t_1-diagnostic [33] and mathematically derive a more sophisticated strategy that includes all cluster amplitudes and offers a sufficient condition of a locally unique and quasi-optimal solution (after possibly rotating out the single amplitudes) rather than rejection based on just large single amplitudes. We furthermore complement the literature by a detailed discussion on spectral gap assumptions. In this context, spectrum refers to the point spectrum, i.e. the eigenvalues of relevant operators. Although a gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital (HOMO–LUMO gap), or a spectral gap of the exact Hamiltonian $\hat{H}$ (non-degenerate ground state), is crucial for the analysis, we highlight the importance of coercivity conditions, either for $\hat{H}$ or the Fock operator $\hat{F}$. Additionally, we derive an optimal constant in the monotonicity proof of the CC function for the finite dimensional case, i.e. the projected CC theory. Comparing the CC Lagrangian with the extended CC formulation [31], we propose by means of the bivariational principle an alternative to measure the quality of the Lagrange multipliers, here interpreted as wave function parameters.

This article is structured as follows: In Section 2, a brief summary of the CC theory is presented. We introduce the set of admissible wave functions and moreover define cluster operators, the CC function, and the CC energy (for a full scope treatment of the mathematical formulation of CC theory presented here we refer to [16,17]). In Section 3, we discuss the use of local analysis within different CC methods. Key concepts here are (see Section 3.1 for definitions) *local strong monotonicity* and *local Lipschitz continuity* of the CC function *f*, which – if fulfilled – are sufficient conditions for a locally unique solution of $f = 0$ by Zarantonello's theorem. In particular, the importance of so-called Gårding inequalities is demonstrated. This is done both for the Hamiltonian, Section 3.2.1, and for the Fock operator, Section 3.2.2. We conclude in Section 3.3 with an overview of the Aubin-Nitsche method and the bivariational principle as they are used in CC methods for estimating the truncation error of the energy.

2. Wave functions on an exponential manifold

The aim of electronic many-body methods, such as the CC approach, is to solve the electronic Schrödinger equation (SE) $\hat{H}\psi = E_0\psi$ of an N-electron system. Here, E_0 is the ground-state energy and $\hat{H}$ the self-adjoint Coulomb Hamiltonian. In this work, we restrict our attention to real Hamiltonians and wave functions. We emphasise that the mathematical framework of Hermitian operators is not sufficient to support the necessary spectral theory for quantum mechanics. Thirring exemplified this with the radial momentum operator $\hat{P}_r = -i\hbar(\partial/\partial r)$ on $\mathcal{D}(\hat{P}_r) = \{\psi \in L^2((0,\infty)) : \psi(r = 0) = 0 \text{ and } \hat{P}_r\psi \in L^2((0,\infty))\}$ [34].

From a mathematical viewpoint the Coulomb Hamiltonian, like most differential operators, is studied in its *weak form* to allow a larger variety of solutions. Set $\Omega = \mathbb{R}^3 \times \{\pm\frac{1}{2}\}$ (or any other appropriate region in space and number of spin states) and let $\int_{\Omega^N} d\tau$ denote both integration and summation over spatial and spin degrees of freedom. Multiplying the SE on both sides with a smooth and compactly supported function $\phi \in \mathcal{C}_c^\infty(\Omega^N)$, a so-called *test function*, and integrating by parts yields ($\nabla = (\nabla_{r_1}, \ldots, \nabla_{r_N})$)

$$\frac{1}{2}\int_{\Omega^N} \nabla\psi \cdot \nabla\phi \, d\tau + \int_{\Omega^N} \psi\,\hat{V}_C\phi \, d\tau = E_0 \int_{\Omega^N} \psi\phi \, d\tau, \tag{1}$$

where $\hat{V}_C$ denotes the Coulomb operator (containing both the Coulomb attraction and repulsion) and ψ a solution of the SE. It follows immediately that the l.h.s. of Equation (1) defines a bilinear form $a(\cdot,\cdot) : \mathcal{C}_c^\infty \times \mathcal{C}_c^\infty \to \mathbb{K}$ with $\mathbb{K}$ being the underlying algebraic field. Boundedness and ellipticity of this bilinear form, however, are non-trivial consequences that go back to Hardy–Rellich inequalities proving that $\hat{V}_C : \mathcal{C}_c^\infty \to L^2$ is bounded (for a general introduction see [35]). Note that this treatment of the SE extends the set of admissible wave functions to the set of antisymmetric L^2-functions ψ of finite kinetic energy $\mathcal{K}(\psi)$, i.e.

$$\|\psi\|_2^2 := \int_{\Omega^N} |\psi|^2 \, d\tau < +\infty$$

and

$$\mathcal{K}(\psi) := \frac{1}{2}\sum_{i=1}^{N} \int_{\Omega^N} |\nabla_{r_i}\psi|^2 \, d\tau < +\infty.$$

We denote this space $H^1(\Omega^N)$ and impose the norm

$$\|\cdot\| : H^1 \to \mathbb{R}; \psi \mapsto \sqrt{\|\psi\|_2^2 + 2\mathcal{K}(\psi)}.$$

In this topology $\mathcal{C}_c^\infty(\Omega^N) \subseteq H^1(\Omega^N)$ is dense. Hence, the bilinear form $a(\cdot,\cdot)$ is continuously extendable to $H^1(\Omega^N)$. We define the operator

$$\hat{H}_w : H^1 \to (H^1)'; \psi \mapsto \hat{H}_w\psi = a(\psi,\cdot),$$

where $(H^1)'$ is dual space of H^1, which we shall denote H^{-1} from now on. Note that $\hat{H}_w$ maps indeed into H^{-1} since boundedness and ellipticity are preserved under continuous extensions. Furthermore, the r.h.s. of Equation (1) can be generalised to the *dual pairing* allowing to reformulate the SE as an operator equation: Find $\psi \in H^1$ such that $\hat{H}_w\psi = E_0\psi'$, with ψ' being the *Riesz representation* of ψ. This general approach to the SE was to the best of our knowledge not considered in the mathematical analyses of CC theory prior to the work of Schneider and Rohwedder [15–17]. Subsequently, we consider this weak formulation and for simplicity write $\hat{H}_w = \hat{H}$.

Different parameterisations of ψ lead to different approximation schemes, subject of this article is the CC scheme, i.e. we parameterise ψ on an exponential manifold. We assume that the solution ψ_* can be written $\psi_* = \phi_0 + \psi_\perp$, where ϕ_0 is a reference determinant of N one-electron functions and $\psi_\perp$ is an element of $\{\phi_0\}^\perp$, the L^2-orthogonal complement of ϕ_0. We denote the L^2-inner product by $\langle\cdot|\cdot\rangle$ and follow the quantum chemistry notation for expectation values of operators, i.e. $\langle\psi|\hat{A}|\psi\rangle$. In particular, assuming that $\hat{H}$ supports a ground state, which is always the case for Coulomb systems [36], the Rayleigh–Ritz variational principle reads

$$E_0 = \min_{\psi \neq 0} \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} =: \min_{\psi \neq 0} \mathcal{R}(\psi),$$

with $\psi \in H^1$. Note that although we assume ψ to be normalisable (L^2-summable), we do not impose $\|\psi\|_2 = 1$, but rather $\|\phi_0\|_2 = 1$. Furthermore, by construction of the solution ψ_*, we assume intermediate normalisation $\langle\phi_0|\psi\rangle = 1$.

Next, let $\{\chi_k\} \subset H^1(\Omega)$ be an $L^2(\Omega)$-orthonormal one-electron basis of the space of admissible one-electron wave functions. Unless we explicitly write $\{\chi_k\}_{k=1}^K$ we refer to the infinite dimensional setting. We construct from this set an $L^2(\Omega^N)$-orthonormal *Slater basis* in the

usual fashion denoted $\{\phi_\mu\}$. Note that the N-particle basis functions $\{\phi_\mu\}$ span the infinite dimensional space of all possible excitations with respect to the reference determinant ϕ_0. In this notation we have $\psi = \phi_0 + \psi_\perp = \phi_0 + \sum_\mu s_\mu \phi_\mu$, where $\{s_\mu\}$ are the L^2-weights of ψ in the given Slater basis, i.e. $s_\mu = \langle \phi_\mu | \psi \rangle$. We formally define the cluster operators by $\hat{S} = \sum_\mu s_\mu \hat{X}_\mu$, where $\hat{X}_\mu$ excites the reference state ϕ_0 to the state ϕ_μ. We obtain $\psi = (\hat{I} + \hat{S})\phi_0$ with $\hat{I}$ denoting the identity operator. The coefficients s_μ are called *cluster amplitudes* and we say that $s = \{s_\mu\}$ is a set of admissible cluster amplitudes if $\hat{S}\phi_0 \in L^2$ and $\mathcal{K}(\hat{S}\phi_0) < +\infty$. Due to the one-to-one relationship between cluster amplitudes and linearly parametrised wave functions, a natural choice for a norm on the space of admissible cluster amplitudes is the corresponding wave function norm of $\hat{S}\phi_0$ [15,16], i.e.

$$\|s\|^2 = \|\hat{S}\phi_0\|^2 = \|\hat{S}\phi_0\|_2^2 + \mathcal{K}(\hat{S}\phi_0).$$

2.1. The exponential ansatz

The CC theory is based on an exponential parametrisation of wave functions. This is an alternative and, assuming full *excitation rank* (explained below) of the cluster operators, equivalent description of the full CI (FCI) wave function. Since its introduction by Hubbard [3] and, independently, Hugenholtz [4], the unique parametrisation of a wave function ψ by the exponential $\psi = e^{\hat{T}}\phi_0$ was assumed to be true and motivated from formal manipulations. However, the unique representation of functions in a Hilbert space is by nature a mathematical problem and was rigorously proven for the exponential parametrisation in the infinite dimensional case by Rohwedder [16].

A key element in deriving the exponential parameterisation from the mathematical viewpoint is the well-definedness of the exponential of $\hat{T}$ (or equivalently the logarithm of $\hat{I} + \hat{S}$), which is subject of functional calculus. We emphasise that the applicability of functional calculus depends strongly on the operator's domain since different domains may imply different properties of the operator, e.g. boundedness, essential self-adjointness, sectorial spectrum, etc. By the fact that Rohwedder [16] showed the H^1-continuity of cluster operators in a continuous setting, the functional calculus for bounded operators was proven to be applicable.

In the finite dimensional case this result was known in the quantum chemistry community, see, e.g. Živković and Monkhorst [19]. However, this result was revisited by Schneider [15] using the *Cauchy–Dunford* calculus. To the best of our knowledge, the subtleties addressed in [15,16] have not been part of previous considerations

in mathematical analysis of CC theory. These important results demonstrate how quantum chemistry benefits from mathematics on a very fundamental level. The continuous CC theory amounts to the exact formulation where the set $\{\chi_k\}$ forms a basis (in the strict mathematical sense) of the one particle space $H^1(\Omega)$. In a for this article appropriate form, we recall Rohwedder's result [16]:

(i) *Let ϕ_0 denote a reference determinant, e.g. the Hartree–Fock solution. Given a wave function $\psi_\perp \in \{\phi_0\}^\perp \cap L^2$, i.e. $\langle \psi_\perp | \phi_0 \rangle = 0$, set $S = S_{\psi_\perp}$ where $S_{\psi_\perp}\phi_0 = \psi_\perp$ and note that $S \in \mathcal{B}(L^2, L^2)$, i.e. a bounded linear operator from L^2 into L^2. Then, $\psi_\perp \in H^1$ if and only if $S \in \mathcal{B}(H^1, H^1)$. Furthermore, there exists a constant C independent of $\psi_\perp$ such that*

$$\|\psi_\perp\| \leq \|S\| \leq C\|\psi_\perp\|.$$

An equivalent statement holds for the L^2-adjoint of S.

(ii) *The exponential map $\hat{T} \mapsto e^{\hat{T}}$ is a C^∞ isomorphism between $\mathcal{C} := \{\hat{T} : \hat{T} \in \mathcal{B}(H^1, H^1)\}$ and $\mathcal{I} + \mathcal{C} := \{\hat{I} + \hat{T} : \hat{T} \in \mathcal{B}(H^1, H^1)\}$. In particular, for any $\psi \in H^1$ with $\langle \phi_0 | \psi \rangle = 1$ there exists a unique $\hat{T}$ such that $\psi = e^{\hat{T}}\phi_0$.*

Note that this result holds for any orthonormal set of N-particle basis functions spanning the space of selected excitations with respect to the reference determinant ϕ_0. However, it is required that the excitation rank of the cluster operators remains untruncated, i.e. $\hat{T} = \sum_{k=1}^N \hat{T}_k$, where $\hat{T}_1$ corresponds to single excitations, $\hat{T}_2$ to double excitations, ..., $\hat{T}_N$ to N-fold excitations. Consequently, we have

$$\psi = \exp(\hat{T}_1 + \cdots + \hat{T}_N)\phi_0 \qquad (2)$$

in the case of full excitation rank.

The usual identification between the linear and exponential parametrisation holds [37]: Write $\hat{S} = \hat{S}_1 + \cdots + \hat{S}_N$ and suppose that the linear parametrisation is given by

$$\psi = (\hat{I} + \hat{S}_1 + \cdots + \hat{S}_N)\phi_0. \qquad (3)$$

Expanding the exponential in Equation (2), and comparing with Equation (3), then yields

$$\hat{T}_1 = \hat{S}_1, \quad \hat{T}_2 = \hat{S}_2 - \frac{1}{2}\hat{S}_1^2, \quad \cdots$$

and for the amplitudes

$$t_i^a = c_i^a/c_0, \quad t_{ij}^{ab} = c_{ij}^{ab}/c_0 - (c_i^a c_j^b - c_j^a c_i^b)/c_0^2, \quad \cdots \quad ,$$

where c_0 is the FCI coefficient of the reference determinant (here $c_0 = 1$). This shows a one-to-one relation for untruncated linear and exponential parameterisations. Restricting the parametrisation on the sub-manifold of

excitation rank $k < N$, this one-to-one relationship is in general not true (see Remark 2 in [31]): Consider CCSD for $N > 2$ particles, i.e. $\psi = e^{\hat{T}_1 + \hat{T}_2}\phi_0$. Expanding the exponential yields

$$\hat{T}_1 + \hat{T}_2 + \frac{(\hat{T}_1 + \hat{T}_2)^2}{2} + \cdots + \frac{(\hat{T}_1 + \hat{T}_2)^N}{N!} = \hat{S},$$

which is not a CISD parametrisation, unless for the trivial case $\hat{T}_1 = \hat{T}_2 = 0$.

2.2. The CC energy

Being able to express any wave function in H^1 on an exponential manifold, it is straightforward to derive the linked CC equations [37]:

$$\mathcal{E}(t) = \langle\phi_0| e^{-\hat{T}}\hat{H} e^{\hat{T}}|\phi_0\rangle,$$

$$(f(t))_\mu = \langle\phi_\mu| e^{-\hat{T}}\hat{H} e^{\hat{T}}|\phi_0\rangle = 0, \quad \text{for all } \phi_\mu. \quad (4)$$

Here, ϕ_0 and all the (visavi ϕ_0) excited determinants ϕ_μ are assumed to form a basis of the anti-symmetric part of H^1. Note that the above equation defines the CC function f and the CC energy function $\mathcal{E}$. Theorem 5.3 from [16] demonstrates that the CC theory provides a wave function that satisfies $\mathcal{R}(\psi_*) = E_0 = \mathcal{E}(t_*)$:

The continuous (and with full excitation rank) CC amplitudes t_* solve $f(t_*) = 0$ fulfilling $\mathcal{E}(t_*) = E_0$ if and only if the corresponding function $\psi_* = e^{\hat{T}_*}\phi_0$ solves the SE $\hat{H}\psi_* = E_0\psi_*$.

By this fact and with $E_0 = \mathcal{E}(t_*)$, if t_* solves $f(t_*) = 0$ the SE yields

$$\langle\psi_*|\hat{H}|\psi_*\rangle = E_0\langle\psi_*|\psi_*\rangle.$$

Hence, the CC amplitudes describe a function ψ_* that provides the system's energy in the usual quantum mechanical setting, i.e. $\mathcal{R}(\psi_*) = \mathcal{E}(t_*)$.

In practice, computations are carried out using a finite basis $\{\chi_k\}_{k=1}^K$ and furthermore with a truncated excitation rank $\hat{T}^{(n)} = \hat{T}_1 + \cdots \hat{T}_n, n < N$. The total truncation level can then be denoted by $d = (K, n)$, and where we solve $f = 0$ on $\mathcal{V}^{(d)}$ to obtain $f(t_d) = 0$, $E_d = \mathcal{E}(t_d)$. We note the following from the literature:

(i) Given a finite one-electron basis $\{\chi_k\}_{k=1}^K$, we denote the span of the corresponding Slater basis by H_K^1. With full excitation rank $(n = N)$ Proposition 4.7 in [15] gives: $f(t_d) = 0$ and $E_K = \mathcal{E}(t_d)$ if $\psi_d = e^{\hat{T}_d}\phi_0$ solves the SE on H_K^1, i.e. $\hat{H} e^{\hat{T}_d}\phi_0 = E_K e^{\hat{T}_d}\phi_0$. By the argument of Monkhorst in [38] we can establish the reverse: Assume $f(t_d) = 0$ and set $E_K = \mathcal{E}(t_d)$, then since $\hat{I}_K = |\phi_0\rangle\langle\phi_0| + \sum_\mu |\phi_\mu\rangle\langle\phi_\mu|$ we obtain (Equation (38) in [38])

$$\langle\phi_0| e^{\hat{T}_d^\dagger} e^{\hat{T}_d}|\phi_0\rangle \mathcal{R}(e^{\hat{T}_d}\phi_0)$$

$$= \langle\phi_0| e^{\hat{T}_d^\dagger} e^{\hat{T}_d}\hat{I}_K e^{-\hat{T}_d}\hat{H} e^{\hat{T}_d}|\phi_0\rangle$$

$$= \langle\phi_0| e^{\hat{T}_d^\dagger} e^{\hat{T}_d}|\phi_0\rangle\mathcal{E}(t_d) + \sum_\mu \langle\phi_0| e^{\hat{T}_d^\dagger} e^{\hat{T}_d}|\phi_\mu\rangle f(t_d).$$

From this we can conclude $\mathcal{R}(e^{\hat{T}_d}\phi_0) = \mathcal{E}(t_d) = E_K$, i.e. the CC wave function gives the energy when inserted into the Rayleigh–Ritz quotient. Furthermore, we have (where $\mathcal{C}_d$ denotes the truncated version of $\mathcal{C}$)

$$\inf\{\mathcal{R}(\psi) : \psi = e^{\hat{T}}\phi_0, \hat{T} \in \mathcal{C}_d\}$$

$$= \inf\{\mathcal{R}(\psi) : \psi$$

$$= (\hat{I} + \hat{S})\phi_0, \hat{S} \in \mathcal{C}_d\} = E_K,$$

by the equivalence between linear and exponential parametrisation as long as full excitation rank is kept. Consequently, $\psi_d = e^{\hat{T}_d}\phi_0$ solves the SE on H_K^1, which establishes the reversed implication in Proposition 4.7 in [15].

(ii) However, for $n < N$ we have in general (see for instance Remark 4.9 in [15])

$$E_d^{\mathrm{var}} := \inf\{\mathcal{R}(\psi) : \psi = e^{\hat{T}}\phi_0, \hat{T} \in \mathcal{C}_d\} \neq E_d, \quad n < N,$$

which gives the well-known result that the computed E_d is not an upper bound to E_K. Hence, $E_d \neq \mathcal{R}(e^{\hat{T}_d}\phi_0)$ where $f(t_d) = 0$ and $E_d = \mathcal{E}(t_d)$.

(iii) By (ii), strictly speaking, CC methods do not compute wave functions, as ψ_d does not provide the system's energy and therewith does not fulfill the Copenhagen interpretation's first principle [39]. However, as mathematical analyses in [15–17,31,32] have demonstrated, CC methods do provide approximate wave functions that converge to the solution of the SE (as $K \to \infty$, $n \to N$). The Copenhagen interpretation is formulated for full systems, which correspond to the continuous CC formulation, and does not contain any statement about approximative solutions. This raises the fundamental question of what properties should be demanded for approximative solutions.

(iv) To contrast with the next section, we would also like to point out the work [19] where, for a finite basis, the CC equations were analysed in a perturbational setting. Writing

$$e^{\hat{T}^{(n)}} = \hat{I} + \hat{T}_1 + \cdots + \hat{T}_n + \lambda(\hat{T}_{n+1}(n) + \hat{T}_{n+2}(n)),$$

where we followed the notation in [19] (see Equations (A9) and (A10)), the CI equations are obtained at

$\lambda = 0$ and $\lambda = 1$ corresponds to the CC case. From this and under the assumption of a finite one-electron basis, both the reality and multiplicity of the CC solutions were investigated with respect to pole and branch cut singularities in the complex plane. The emergence of multiple solutions is certainly interesting and worth pursuing, however, the local analysis studied here instead deals with the establishment of a *locally* unique solution under certain assumptions. Note that the local behaviour of a solution is important for the applicability and convergence of Newton–Rhapson and quasi-Newton methods.

3. Local analysis in CC theory

The CC equations (linked and unlinked) can be formulated as a *non-linear Galerkin scheme*, which is a well-established framework in numerical analysis to convert the continuous Schrödinger equation to a discrete problem. Instead of solving the full problem, Galerkin methods solve the CC equations in a finite dimensional subspace $H_d \subseteq H^1$. Note that the CC equations remain the same, only the space spanned by the considered $\{\phi_\mu\}$ has changed. Reducing the problem to a finite-dimensional vector subspace allows to numerically compute an approximate solution via Newton–Rhapson or quasi-Newton methods. Galerkin methods allow a local analysis, which is useful for CC theory due to the manifold of solutions [18–23] and the use of quasi-Newton methods that require certain local behaviour of the solutions. Local analysis furthermore allows reliable statements about the existence and local uniqueness of Galerkin solutions as well as quantitative statements on the basis-truncation error. Its backbone is formed by a local version of Zarantonello's theorem [40]:

Let $f : X \to X'$ be a map between a Hilbert space $(X, \langle \cdot, \cdot \rangle, \| \cdot \|)$ and its dual X', and let $x_* \in B_\delta$ be a root, $f(x_*) = 0$, where B_δ is an open ball of radius δ around x_*. Assume that f is Lipschitz continuous and strongly monotone in B_δ with constants $L > 0$ and $\gamma > 0$, respectively. Then the root x_* is unique in B_δ. Indeed, there is a ball $C_\varepsilon \subset X'$ with $0 \in C_\varepsilon$ such that the solution map $f^{-1} : C_\varepsilon \to X$ exists and is Lipschitz continuous, implying that the equation $f(x_* + x) = y$ has a unique solution $x = f^{-1}(y) - x_*$, depending continuously on y, with norm $\|x\| \leq \delta$. Moreover, let $X^{(d)} \subset X$ be a closed subspace such that x_* can be approximated sufficiently well, i.e. the distance $d(x_*, X^{(d)})$ is small. Then, the projected problem $f_d(x_d) = 0$ has a unique solution $x_d \in X^{(d)} \cap B_\delta$ and

$$\|x_* - x_d\| \leq \frac{L}{\gamma} d(x_*, X^{(d)}),$$

i.e. x_d is a quasi-optimal solution.

The concept of quasi optimality was introduced in Jean Céa's dissertation [41] in 1964 for linear Galerkin schemes and got extended over the years to the non-linear case. It ensures that the Galerkin solution in a fixed approximative space is, up to a multiplicative constant, the closest element to the exact solution. For obvious reasons this is a desired property for CC schemes. The different CC methods vary, however, in more than just minor details, which makes this property a conceptual different and challenging task to establish for each method.

3.1. Local unique solutions and quasi-optimality

We start by elaborating on the assumptions of Zarantonello's theorem in a more demonstrative way. Here, the notation $\langle s, t \rangle = \sum_\mu s_\mu t_\mu$ is used for sequences $s = \{s_\mu\}$ and $t = \{t_\mu\}$. In the context of the CC theory, the CC function f from Equation (4) is said to be *strongly monotone* if for sets of cluster amplitudes $t = \{t_\mu\}$ and $t' = \{t'_\mu\}$ there exists a $\gamma > 0$ such that

$$\langle f(t) - f(t'), t - t' \rangle \geq \gamma \|t - t'\|^2. \tag{5}$$

If this inequality is true for all $t, t' \in B_\delta(t_*)$ then f is said to be locally strongly monotone. The CC function f is further said to be *Lipschitz continuous* if there exists a constant $L > 0$ such that

$$\|f(t) - f(t')\| \leq L \|t - t'\|. \tag{6}$$

In direct analogy with local strong monotonicity, we define local Lipschitz continuity if Equation (6) is fulfilled for all cluster amplitudes t, t' inside some ball.

To exemplify these concepts in a simple way we consider a smooth function $f : \mathbb{R} \to \mathbb{R}$. By the Cauchy–Schwarz inequality, the strong monotonicity implies that the derivative $f'(t) \geq \gamma$, i.e. f is a strictly monotonically increasing function. Note that strictly monotone functions are injective (one-to-one), which implies local invertibility. Hence, this already ensures local uniqueness of the function's root t_*, if supported. Lipschitz continuity on the other hand implies that $-L \leq f'(t) \leq L$. Hence, the assumptions in Zarantonello's theorem are restrictions to the function's slope, namely

$$0 < \gamma \leq f'(t) \leq L.$$

By introducing normed operator spaces, these restrictions can be generalised to vector valued and even infinite dimensional functions f.

Returning to the general case, the Lipschitz continuity is key to derive the quasi-optimality in case of Galerkin solutions. We assume that $X^{(d)} \subsetneq X$ is the considered approximation space supporting the Galerkin solution t_d,

i.e. $\langle f(t_d), s \rangle = 0$ for all $s \in X^{(d)}$. Then, $f(t_*) - f(t_d) \in (X^{(d)})^{\perp}$, i.e. $\langle f(t_*) - f(t_d), u \rangle = 0$ for all $u \in X^{(d)}$, in particular for $u = t_d$. Starting from the strong monotonicity, we deduce for any $u \in X^{(d)}$ that

$$\begin{aligned} \gamma \| t_* - t_d \|^2 &\leq \langle f(t_*) - f(t_d), t_* - t_d \rangle \\ &= \langle f(t_*) - f(t_d), t_* - u \rangle \\ &\leq L \| t_* - t_d \| \| t_* - u \|. \end{aligned}$$

Because $u \in X^{(d)}$ was chosen arbitrarily, the above estimate holds for all u, which implies the quasi optimality:

$$\| t_* - t_d \| \leq L/\gamma \min_{u \in X^{(d)}} \| t_* - u \|. \qquad (7)$$

To apply Zarantonello's theorem to CC methods, the main challenge is to demonstrate a strictly positive γ in Equation (5) such that strong monotonicity holds locally around the solution that corresponds to the ground state. The original idea in [15] to obtain such a result in the finite-dimensional projected CC theory assumed the existence of an HOMO–LUMO gap. Further, more technical assumptions on the Fock operator $\hat{F}$ (see Gårding inequality below) were needed to achieve a generalisation to the continuous CC setting [17], which also has a counterpart for $\hat{H}$. We refer the reader to [15–17,31,32] for the detailed proofs and made assumptions, not only within the traditional CC formalism, but also for the TCC and extended CC methods. However, we remark that these assumptions are sufficient conditions but not necessary. One example is given by metals: Despite their typically small or negligible HOMO–LUMO gaps, the single-reference CC method can compute metallic effects often quite well. This suggests that the HOMO–LUMO gap assumption, which limits the results' applicability, can be lifted in the case of non-multi-configuration systems [32]. See also [23] for a CC theory that considers open-shell systems where no HOMO–LUMO gap exists.

Here, we extend the results in [15–17,31,32] by optimising the strong monotonicity constant γ, which yields lesser restrictions on the solution's cluster amplitudes $t_* = \{(t_*)_\mu\}$. Further investigations need to be undertaken before the presented analysis can lead to practical results of the reliability of the CC approach. However, we suggest an estimate on the CC amplitudes that is sufficient to guarantee the existence of a locally unique CC solution (see Equation (13)) and contrast it with the single amplitudes diagnostic of [33].

3.2. Local strong monotonicity of the CC function

In the literature there are two different proofs that the infinite dimensional (continuous) CC function f is locally strongly monotone [17] (see also [31] for the extended CC function). Even though spectral-gap assumptions enter the arguments, it is the so-called Gårding constants that give a sufficient condition for the local strong monotonicity, as will be demonstrated below. This fact emerges from the analysis in [17] but was noted and elaborated within the analysis of the extended CC method in [31]. We here furthermore improve the existing analysis by optimising the constants. We start by defining the Gårding inequality that will be used extensively in the sequel:

An operator $\hat{A}$ fulfills a Gårding inequality if there exists a real constant e such that $\hat{A} + e$ is coercive, i.e. there exists a constant $c > 0$ that depends on e (we denote this dependence by $c(e)$) such that

$$\langle \psi | \hat{A} + e | \psi \rangle \geq c(e) \| \psi \|^2.$$

The coercivity above describes a particular growth behaviour of $\hat{A} + e$ as the lower bound becomes large when the wave function is at the extreme of the space, e.g. wave functions with a large kinetic energy. Subsequently, we denote the l.h.s. of Equation (5) by Δ, i.e. for two sets of CC amplitudes $t = \{t_\mu\}$ and $t' = \{t'_\mu\}$ we have

$$\Delta = \langle f(t) - f(t'), t - t' \rangle.$$

We further set $\Delta \hat{T} = \hat{T} - \hat{T}'$, which yields by the CC equations in Equation (4) the equality

$$\Delta = \langle \Delta \hat{T} \phi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} - e^{-\hat{T}'} \hat{H} e^{\hat{T}'} | \phi_0 \rangle. \qquad (8)$$

Next, we elaborate on Gårding inequalities for two different operators that imply local strong monotonicity of the CC function, by bounding the r.h.s. of Equation (8). Interestingly, for the finite-dimensional (projected) CC method, only the latter approach has a counterpart (using the particular structure of the Fock operator $\hat{F}$).

3.2.1. A Gårding inequality for the hamiltonian

We here assume a spectral gap γ_* of $\hat{H}$, i.e. for all ψ that are L^2-orthogonal to the ground state ψ_* we have $\mathcal{R}(\psi) - E_0 \geq \gamma_*$, for some $\gamma_* > 0$, i.e. we assume a non-degenerate ground state. We also assume that ϕ_0 is a good approximation of the exact wave function, i.e. $\varepsilon = \| \psi_* - \phi_0 \|_2$ is small. It then holds (see Lemma 11 in [31])

$$\langle \hat{T} \phi_0 | \hat{H} - E_0 | \hat{T} \phi_0 \rangle \geq \gamma_*(\varepsilon) \| \hat{T} \phi_0 \|_2^2, \qquad (9)$$

with $\gamma_*(\varepsilon) = \gamma_*(1 - 4\varepsilon + \mathcal{O}(\varepsilon^2))$. Thus, $\gamma_*(\varepsilon)$ is close to γ_* and strictly positive, if ε is sufficiently close to zero. Using the argument in [17,31] (see proof of Theorem 3.4 in [17], and also Equation (16) with $\hat{\Lambda}_* = 0$ together with

the proof of Theorem 16 in [31]), we obtain

$$\Delta \geq \langle \Delta\hat{T}\phi_0 | \hat{H} - E_0 | \Delta\hat{T}\phi_0 \rangle$$
$$- (\| e^{-\hat{T}_*^\dagger} - \hat{I} \| + \| e^{-\hat{T}_*^\dagger} \| \| e^{\hat{T}_*} - \hat{I} \|) \| \Delta\hat{T}\phi_0 \|^2. \tag{10}$$

In [17], the first term of Equation (10) was bounded by a constant times $\| \Delta\hat{T}\phi_0 \|^2$, achieved by combining the Gårding inequality with Equation (9).

From Lemma 11 in [31], it follows that

$$\langle \Delta\hat{T}\phi_0 | \hat{H} - E_0 | \Delta\hat{T}\phi_0 \rangle \geq \frac{\gamma_*(\varepsilon)}{\gamma_*(\varepsilon) + e + E_0} c(e) \| \Delta\hat{T}\phi_0 \|^2.$$

However, this can be further strengthened to

$$\langle \Delta\hat{T}\phi_0 | \hat{H} - E_0 | \Delta\hat{T}\phi_0 \rangle \geq \eta_{\text{opt}}(\varepsilon) \| \Delta\hat{T}\phi_0 \|^2,$$

with the optimal constant

$$\eta_{\text{opt}}(\varepsilon) := \max_{e > 0} \frac{\gamma_*(\varepsilon)}{\gamma_*(\varepsilon) + e + E_0} c(e).$$

From this we conclude

$$\Delta \geq \left(\eta_{\text{opt}}(\varepsilon) - \| e^{-\hat{T}_*^\dagger} - \hat{I} \| \right.$$
$$\left. - \| e^{-\hat{T}_*^\dagger} \| \| e^{\hat{T}_*} - \hat{I} \| \right) \| t - t' \|^2, \tag{11}$$

which yields the following *sufficient condition* for the local strong monotonicity of f, namely

$$\eta_{\text{opt}}(\varepsilon) > \| e^{-\hat{T}_*^\dagger} - \hat{I} \| + \| e^{-\hat{T}_*^\dagger} \| \| e^{\hat{T}_*} - \hat{I} \|. \tag{12}$$

Given $\gamma_* > 0$, we observe that a sufficiently small ε and t_*, such that $\| \hat{T}_* \|$ is small enough relative to $\eta_{\text{opt}}(\varepsilon)$, guarantees that Equation (12) is fulfilled. (Recall that $\| t \|$ and $\| \hat{T} \|$ are equivalent, see Section 2.1.)

To finalise this section, we offer the following interpretation of Equation (11), providing a more descriptive approach to Equation (12). We see as e tends to $-E_0$ from above, the quotient $\gamma_*(\varepsilon)/(\gamma_*(\varepsilon) + e + E_0)$ goes to one from below. Furthermore, assume that $c(e)$ goes to zero from above as e approaches $-E_0$ from above. This suggest an optimal value of $e_{\text{opt}} > -E_0$. For instance, choosing $e_n = -E_0 + \gamma_*(\epsilon)/n$ implies

$$\frac{\gamma_*(\varepsilon)}{\gamma_*(\epsilon) + e + E_0} c(e_n) = \frac{1}{1 + 1/n} c(e_n)$$

such that $\gamma_*(\varepsilon)$ is eliminated from the expression. Assuming further that e_{opt} corresponds to an $n_{\text{opt}} \gg 1$ yields $\eta_{\text{opt}} \approx c(e_{\text{opt}})$. In conclusion, as long as $\gamma_*(\varepsilon) > 0$, the Gårding constant $c(e_{\text{opt}})$ offers a direct estimate of the monotonicity constant $\gamma \approx c(e_{\text{opt}}) - 2\| \hat{T}_* \| +$

$\mathcal{O}(\| \hat{T}_* \|^2)$. We therefore obtain the following (approximate) sufficient condition for local strong monotonicity

$$c(e_{\text{opt}}) > 2\| T_* \|. \tag{13}$$

Note that $\| \hat{T} \| \geq K\| t \|$, for some constant K. However, a sharp estimate for this constant is object of current research. Thus, for Zarantonello's theorem to guarantee a locally unique solution, the exact amplitudes $t_* = \{(t_*)_\mu\}$ cannot be too large relative to $c(e_{\text{opt}})$. We remark that by an appropriate choice of the reference determinant ϕ_0, the single amplitudes $t_1 = \{(t_1)_\mu\}$ do not contribute to (the overall) $\| t \|$. Thus, if $\| t \|$ is too large then this is a consequence of $t_2, t_3, \ldots$ (doubles, triples, etc.). Numerical investigations are left for future work but we can already compare this mathematically derived sufficient condition for locally unique CC solutions with the t_1-diagnostics of [33]. Given the truncation level n of the excitation rank, here the proposed diagnostic uses all cluster amplitudes $t_1, t_2, \ldots, t_n$ and not just the single amplitudes t_1. This is a clear advantage since, as mentioned above, orbital rotations can be used to rotate out the single amplitudes. However, our diagnostic offers only a sufficient and not a necessary criterion for a local unique solution, i.e. for large $t_2, t_3, \ldots$ the current diagnostic is agnostic about local uniqueness and only states that local strong monotonicity cannot be inferred from this particular analysis. We hope that future work will clarify the situation further.

3.2.2. A Gårding inequality for the fock operator

On the other hand, assume an HOMO–LUMO gap $\gamma_0 > 0$ of the Fock operator $\hat{F}$ and that ϕ_0 is the Hartree–Fock solution, i.e. $\hat{F}\phi_0 = \Lambda_0\phi_0$ with

$$\langle \psi | \hat{F} - \Lambda_0 | \psi \rangle \geq \gamma_0 \| \psi \|_2^2, \quad \text{for all } \psi \perp \phi_0.$$

The HOMO–LUMO gap thus corresponds to a spectral gap of the Fock operator and we regard Λ_0 as the ground-state energy of $\hat{F}$. Let $\hat{F} = \sum_{i=1}^N \hat{f}(r_i)$ and choose $\{\chi_k\}$ as eigenbasis of $\hat{f}$, i.e. $\hat{f}\chi_k = \lambda_k\chi_k$ for all k. We observe that $\Lambda_0 = \sum_{i=1}^N \lambda_i$, $\gamma_0 = \lambda_{N+1} - \lambda_N > 0$ and $\hat{F}\phi_\mu = (\Lambda_0 + \varepsilon_\mu)\phi_\mu$ with $\varepsilon_\mu = \sum_{l \leq |\mu|} \lambda_{a_l} - \lambda_{i_l}$. The argument proving that the CC function f is locally strongly monotone can then be outlined as follows.

The considered Fock operator is assumed to fulfill a Gårding inequality. Thus there exists a constant e such that $\hat{F} + e$ is coercive, i.e.

$$\langle \psi | \hat{F} + e | \psi \rangle \geq c(e) \| \psi \|^2.$$

For the sake of simplicity we use the same symbols for the Gårding constants of $\hat{F}$ as for the Hamiltonian. In

complete analogy with $\hat{H}$, the argument in [15,31] shows that

$$\langle\psi|\hat{F}-\Lambda_0|\psi\rangle \geq \max_{e>0}\frac{\gamma_0}{\gamma_0+e+\Lambda_0}c(e)\|\psi\|^2 \quad (14)$$

and we moreover define

$$\eta_{\text{opt}}^{(0)} := \max_{e>0}\frac{\gamma_0}{\gamma_0+e+\Lambda_0}c(e). \quad (15)$$

Following [17], for a fixed ϕ_0 we define the map from the space of cluster amplitudes into the space of wave functions $O_{\phi_0} : t \mapsto \hat{O}(t)\phi_0$, with $\hat{O} : t \mapsto [[\hat{F},\hat{T}],\hat{T}] + e^{-\hat{T}}\hat{W}e^{\hat{T}}$. Hence,

$$e^{-\hat{T}}\hat{H}\,e^{\hat{T}}\phi_0 = e^{-\hat{T}}(\hat{F}+\hat{W})\,e^{\hat{T}}\phi_0$$
$$= (\hat{F}+[\hat{F},\hat{T}])\phi_0 + \hat{O}(t)\phi_0, \quad (16)$$

where $\hat{H} = \hat{F} + \hat{W}$, and assume that for some $L > 0$ (not too large)

$$\langle\Delta\hat{T}\phi_0|\hat{O}(t)-\hat{O}(t')|\phi_0\rangle \geq -L\|t-t'\|^2. \quad (17)$$

As a technical remark, the assumption in [17] is the stronger requirement that $t \mapsto \hat{O}(t)\phi_0$ is Lipschitz continuous as a map from the space of cluster amplitudes to H^{-1}. However, we here note that Equation (17) is sufficient to derive the CC function's local strong monotonicity, as will be evident shortly. Inserting the identity (a consequence of Equation (16) and $\hat{F}\phi_0 = \Lambda_0\phi_0$)

$$e^{-\hat{T}}\hat{H}\,e^{\hat{T}}\phi_0 = (\hat{F}+(\hat{F}-\Lambda_0)\hat{T})\phi_0 + \hat{O}(t)\phi_0$$

into Equation (8), as well as using Equations (14) and (17), we obtain

$$\Delta = \langle\Delta\hat{T}\phi_0|\hat{F}-\Lambda_0|\Delta\hat{T}\phi_0\rangle$$
$$+ \langle\Delta\hat{T}\phi_0|\hat{O}(t)-\hat{O}(t')|\phi_0\rangle$$
$$\geq (\eta_{\text{opt}}^{(0)}-L)\|t-t'\|^2. \quad (18)$$

Consequently, local strong monotonicity holds if $\eta_{\text{opt}}^{(0)} > L$. Repeating the argument presented in the previous section, with the obvious adaptations, we obtain

$$c(e_{\text{opt}}) > L \quad (19)$$

as a sufficient condition for f to be locally strongly monotone. Here, no explicit assumption on $\|t_*\|$ enters. The main drawback of the assumption in Equation (19) is that the constant L of the inequality in Equation (17) has to be determined. Further analysis of this constant is postponed for later work.

Before we conclude this section we exemplify how the Gårding constant c can be chosen in the finite dimensional setting. In this case the commutator $[\hat{F},\hat{T}]$ is an excitation operator (which implies $[[\hat{F},\hat{T}],\hat{T}] = 0$) and $\hat{O}(t)$ is simply the similarity transformation of the fluctuation potential $\hat{W}$. This offers the following insight into the optimal constant $\eta_{\text{opt}}^{(0)}$ in Equation (14) for the truncated case. As in [15], we define the norm on $\{\phi_0\}^\perp$ by

$$\|\hat{T}\phi_0\|_F^2 = \sum_\mu \varepsilon_\mu t_\mu^2 = \|t\|_F^2.$$

It follows that

$$\langle\Delta\hat{T}\phi_0|\hat{F}-\Lambda_0|\Delta\hat{T}\phi_0\rangle = \sum_\mu \varepsilon_\mu(\Delta t)_\mu^2 = \|\Delta\hat{T}\phi_0\|_F^2.$$

Using $\|\hat{T}\phi_0\|_F$ instead of $\|\hat{T}\phi_0\|$ and making the assumption in Equation (17) also for the truncated theory (denoting the Lipschitz constant in this new topology by L'), we obtain

$$\Delta = \sum_\mu \varepsilon_\mu(\Delta t)_\mu^2 + \langle\Delta\hat{T}\phi_0|\hat{O}(t)-\hat{O}(t')|\phi_0\rangle$$
$$= \|t-t'\|_F^2 + \langle\Delta\hat{T}\phi_0|\hat{O}(t)-\hat{O}(t')|\phi_0\rangle$$
$$\geq (1-L')\|t-t'\|_F^2. \quad (20)$$

Comparing the local strong monotonicity estimates Equations (18) and (20) suggests that the finite-dimensional version of $\eta_{\text{opt}}^{(0)}$ equals one. Furthermore, at first glance it appears that the estimate in Equation (20) is obtained without imposing a Gårding inequality. A key observation here is that the choice of the norm makes $\hat{F}$ on $\{\phi_0\}^\perp$ fulfill a Gårding inequality with $e_{\text{opt}} = -\Lambda_0$ and $c(e_{\text{opt}}) = 1$. Indeed, the inequality is saturated, meaning that equality holds. It follows then immediately from Equation (15) that $\eta_{\text{opt}}^{(0)} = c(e_{\text{opt}}) = 1$. Thus, in agreement with Equation (19) we have obtained the condition $\eta_{\text{opt}}^{(0)} - L' > 0$.

To conclude this section, we note that we have formulated an alternative to the diagnostic in Equation (13): Assume a finite one-electron basis and suppose that $\hat{O}(t)$ satisfies Equation (17) with the norm $\|\cdot\|_F$ and $L' < 1$ locally around the solution amplitudes. Then local strong monotonicity implies a locally unique CC solution. Whether Equation (17) with $L' < 1$ holds without the assumption of a small $\|t_*\|$ is an interesting and still open question. Furthermore, the above analysis can be generalised to any single particle operator fulfilling certain properties (see [15,32]).

3.3. The CC method's numerical analysis

As computational schemes, the convergence behaviour of CC methods is one of the main objects of study. This

covers whether or not the method converges towards the exact solution as well as how fast it converges. We note that the quasi optimality as given in Equation (7) yields $t_d \rightarrow t_*$ as $d \rightarrow \infty$ (for increasing approximation spaces $X^{(d)}$). Furthermore, in the case of the CC method one studies the CC-energy residual

$$|\mathcal{E}(t_*) - \mathcal{E}(t)|.$$

A major difference between the CI and CC method is that the CC formalism is not variational in the Rayleigh–Ritz sense. Consequently, it is not evident that the CC energy error decays quadratically with respect to the error of the wave function or cluster amplitudes. In the sequel we present two approaches that were used in previous mathematical analyses of different CC methods to derive such quadratic error bounds [17,31,32].

3.3.1. The Aubin–Nitsche duality method

The Aubin–Nitsche duality method is a standard tool for deriving *a priori* error estimates for finite element methods. It was introduced independently by Aubin [42], Nitsche [43] and Oganesyan–Ruchovets [44]. We here elaborate the Aubin–Nitsche duality type method used in [15,17,32] to derive a quadratic error bound for the CC method and the closely related TNS-TCC method, a special case of the tailored CC method [45]. This approach exploits the mathematical framework introduced by Bangerth and Rannacher [46]. The untruncated Euler–Lagrange method gives the Lagrangian $\mathcal{L}(t, s) = \mathcal{E}(t) - \langle f(t), s \rangle$ with f and $\mathcal{E}$ from Equation (4). The corresponding Gâteaux derivative in direction (u, v) is denoted $\mathcal{L}'(\cdot, \cdot)(u, v)$ and we study (t_*, s_*) fulfilling

$$\mathcal{L}'(t_*, s_*)(u, v) = \begin{Bmatrix} \mathcal{E}'(t_*)u - \langle f'(t_*)u, s_* \rangle \\ -\langle f(t_*), v \rangle \end{Bmatrix} = 0, \quad (21)$$

for all pairs of CC amplitude vectors (u, v). Under the assumption that f is locally strongly monotone inside a ball around t_*, there exists a unique s_* determined by t_* such that (t_*, s_*) solves Equation (21). Note, that the assumptions imposed to ensure local strong monotonicity are different for the single-reference CC method [15,17] and the TNS-TCC method [32]. Moreover, there exists a solution s_d to the corresponding discretisation of the problem that approximates s_* quasi optimally [17,32]. Equipped with these so called *dual solutions*, the energy-error characterisation given by Bangerth and Rannacher [46] yields

$$2(\mathcal{E}(t_*) - \mathcal{E}(t_d)) = \mathcal{R}_d^{(3)} + \rho(t_d)(s_* - \upsilon_d) \\ + \rho^*(t_d, s_d)(t_* - w_d),$$

with arbitrarily chosen discrete CC amplitudes υ_d, w_d. The given remainder term $\mathcal{R}_d^{(3)}$ is cubic in the *primal* and

dual error, i.e. $e = t_* - t_d$ and $e^* = s_* - s_d$. Using this energy-error characterisation, a quadratic energy-error bound for the single-reference CC method [15,17] and the TNS-TCC method [32] follows.

3.3.2. The bivariational approach

The extended version of the CC method rests on Arponen's bivariational approach [47,48]. This unconventional formulation of the CC method parametrises two independent wave functions and thus makes use of two sets of cluster amplitudes $t = \{t_\mu\}$ and $\lambda = \{\lambda_\mu\}$. It gained recent attention in the study [31] and has a major advantage as far as the error analysis is concerned, namely, the energy itself is stationary, i.e. the solution (t_*, λ_*) is a critical point of the *bivariational energy*, see Equation (22). Consequently, when the corresponding Galerkin solution (t_d, λ_d) is close to the exact solution, a quadratic error estimate is guaranteed. Subsequently, we elaborate on this further.

Consider the Rayleigh–Ritz quotient, we write

$$E_0 = \mathcal{R}(\psi_*) = \min_{\psi \neq 0} \mathcal{R}(\psi).$$

Hence, ψ_* is a stationary point of $\mathcal{R}$, i.e. $\mathcal{R}'(\psi_*) = 0$. By Taylor expanding $\mathcal{R}$ around ψ_* we obtain the quadratic error estimation for the Rayleigh–Ritz quotient

$$|\mathcal{R}(\psi) - \mathcal{R}(\psi_*)| \leq \frac{1}{2} \|\mathcal{R}''(\psi_*)\| \, \|\psi - \psi_*\|^2 \\ + \mathcal{O}(\|\psi - \psi_*\|^3).$$

As mentioned before, the CC formalism does not arise from the Rayleigh–Ritz variational principle. However, it can be described by Arponen's bivariational approach, as follows. Let the bivariate quotient be

$$\mathcal{B}(\psi, \tilde{\psi}) = \frac{\langle \tilde{\psi} | \hat{H} | \psi \rangle}{\langle \tilde{\psi} | \psi \rangle}. \quad (22)$$

Equation (22) can be seen as a generalisation of the Rayleigh–Ritz quotient where a stationary point $(\psi_*, \tilde{\psi}_*)$ is given by a left and right eigenvector of $\hat{H}$ with corresponding eigenvalue $E = \mathcal{B}(\psi, \tilde{\psi})$. Note that $\mathcal{B}$ is no longer a below bounded functional, hence critical points do not necessarily correspond to extremal points as they do for $\mathcal{R}$. In the extended CC theory, the bivariational quotient is studied indirectly by means of the so-called *flipped gradient* [31]. Following [47], we assume $\langle \phi_0 | \psi \rangle = \langle \psi | \tilde{\psi} \rangle = 1$ and note that there exists $\hat{T}$ such that $\psi = e^{\hat{T}} \phi_0$ (cf. Section 2.1). Then $1 = \langle \psi | \tilde{\psi} \rangle = \langle \phi_0 | e^{\hat{T}^\dagger} | \tilde{\psi} \rangle$ and consequently there exists a cluster operator $\hat{\Lambda}$ so that $e^{\hat{T}^\dagger} \tilde{\psi} = e^{\hat{\Lambda}} \phi_0$. This defines a smooth coordinate map Φ from cluster amplitudes (t, λ) to wave

functions $(\psi, \tilde{\psi})$. The flipped gradient is then given by $\mathcal{F}(t, \lambda) := \hat{R} \nabla \mathcal{B}(\Phi(t, \lambda))$, where we introduced the flipping map

$$\hat{R} = \begin{pmatrix} 0 & \hat{I} \\ \hat{I} & 0 \end{pmatrix}.$$

Under certain assumptions, $\mathcal{F}$ is locally strongly monotone [31]. By the extended CC approach [31], $\tilde{\psi}_* = e^{-\hat{T}_*^\dagger} e^{\hat{\Lambda}_*} \phi_0$ and $\psi_* = e^{\hat{T}_*} \phi_0$ solve the SE if and only if $\mathcal{F}(t_*, \lambda_*) = 0$. Note that $\mathcal{F}(t_*, \lambda_*) = 0$ implies

$$\nabla \mathcal{B}(\Phi(t_*, \lambda_*)) = 0$$

and therewith a quadratic energy error.

Furthermore, by identifying $e^{\hat{\Lambda}} = \hat{I} + \hat{S}$ we obtain from Equation (22) the CC Lagrangian, i.e.

$$\mathcal{B}(e^{\hat{T}} \phi_0, e^{-\hat{T}^\dagger} e^{\hat{\Lambda}} \phi_0) = \langle \phi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \phi_0 \rangle$$
$$+ \sum_\mu s_\mu \langle \phi_\mu | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \phi_0 \rangle =: \mathcal{L}(t, s). \quad (23)$$

Introducing the Lagrangian is a general method for optimisation with constraints. In the special case of CC theory with fixed orbitals, as in this article, Equation (23) demonstrates the equivalence to Arponen's bivariational method [47]. In the context of obtaining an efficient evaluation of CC energy gradient, the derivative of the variational functional was obtained by Bartlett [49]. The functional itself (Equation (23)) was first used in quantum chemistry by Helgaker and Jørgensen [50] to derive CC energy derivatives. We would also like to mention the related extended CC work of Piecuch and Bartlett [51]. Note that their assumption that the reference determinant ϕ_0 is both a left- and right eigenvector of the doubly similarity transformed $\hat{H}$ can be rigorously proven in the continuous case (see Lemma 13 in [31]).

Denoting the dual solution $s_* = \{(s_*)_\mu\}$ as in Section 3.3.1, it can then be seen that s_* also describes cluster amplitudes parameterising the wave function $\tilde{\psi}_*$. Indeed, using the relation $e^{\hat{\Lambda}_*} = \hat{I} + \hat{S}_*$, we obtain that $\tilde{\psi}_* = e^{-\hat{T}_*^\dagger}(\hat{I} + \hat{S}_*) \phi_0$ together with $\psi_* = e^{\hat{T}_*} \phi_0$ solve the SE corresponding to the same energy $\mathcal{B}(\psi_*, \tilde{\psi}_*)$. Assuming non-degeneracy and using the constraint $\langle \tilde{\psi}_* | \psi_* \rangle = 1$, we arrive at the condition

$$e^{-\hat{T}_*^\dagger}(\hat{I} + \hat{S}_*) \phi_0 = \frac{1}{\| e^{\hat{T}_*} \phi_0 \|^2} e^{\hat{T}_*} \phi_0$$

for the primal and dual solutions t_* and s_*. Thus, from the extended CC theory we have obtained a constraint relating s_* to t_* for the traditional CC method.

4. Conclusion

In this article, we have introduced the reader to a local analysis of the CC method and its variations. In particular, we have demonstrated that the Gårding inequalities for $\hat{F}$ and $\hat{H}$ are key as far as a better understanding of the sufficient conditions for a locally unique and quasi-optimal solution of the CC equations is concerned. Moreover, these investigations are geared towards an *a posteriori* criterion of assessing the CC amplitudes from a given computation. This is a mathematical approach that is alternative to the controversial diagnostic suggested in [33]. Indeed, the mathematically derived criteria in Equations (12) and (13) use the total $\|t\|$ and not just the single amplitudes t_1. Since the single amplitudes could be removed by an appropriate choice of the reference determinant (i.e. an ideal choice of the basis functions), the sufficient condition for a locally unique solution given by Equation (13) puts constraints on the remaining amplitudes $(t_2, t_3, \ldots)$. However, it is not yet a rejection criterion since it only *implies* locally unique and quasi-optimal solutions under certain conditions. As outlined, the upper bound in Equation (13) is fundamentally different from previous heuristic and potentially misleading diagnostics [33] since the former is derived in a rigorous mathematical framework, where not just the singles amplitudes are taken into consideration. We have also shown that the condition on the two particle operator in Equation (17) implies a locally unique CC solution. Here, the condition does not explicitly depend on the amplitude norm and might offer a broader understanding of the reliability of a CC solution. Moreover, the derived condition is independent of the chosen single particle operator. In connection with the extended CC formalism, we have set up a constraint for the exact CC Lagrange multipliers $s_* = \{(s_*)_\mu\}$, relating them to the exact CC amplitudes $t_* = \{(t_*)_\mu\}$. Numerical investigations are left for future work.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

ORCID

A. Laestadius 🄳 http://orcid.org/0000-0001-7391-0396

References

[1] R.J. Bartlett and M. Musiał, Rev. Mod. Phys. **79** (1), 291 (2007).

[2] F. Coester, Nucl. Phys. **7**, 421–424 (1958).

[3] J. Hubbard, Proc. R. Soc. Lond. A **240** (1223), 539–560 (1957).

[4] N. Hugenholtz, Physica **23** (1–5), 533–545 (1957).

[5] J. Čížek, J. Chem. Phys. **45** (11), 4256–4266 (1966).

[6] O. Sinanoğlu, J. Chem. Phys. **36** (3), 706–717 (1962).

[7] H. Kümmel, Theor. Chim. Acta. **80** (2–3), 81–89 (1991).

[8] J. Čížek, Theor. Chim. Acta. **80** (2–3), 91–94 (1991).

[9] R. Bartlett, in *Theory and Applications of Computational Chemistry: The First Forty years*, edited by C.E. Dykstra, G. Frenking, K.S. Kim, and G.E. Scuseria (2005), pp. 1191–1221.

[10] J. Paldus, *Theory and Applications of Computational Chemistry* (Elsevier, 2005), pp 115–147.

[11] J.S. Arponen, Theor. Chim. Acta. **80** (2–3), 149–179 (1991).

[12] R. Bishop, Theor. Chim. Acta. **80** (2–3), 95–148 (1991).

[13] K. Raghavachari, G.W. Trucks, J.A. Pople and M. Head-Gordon, Chem. Phys. Lett. **157** (6), 479–483 (1989).

[14] R.J. Bartlett, J. Watts, S. Kucharski and J. Noga, Chem. Phys. Lett. **165** (6), 513–522 (1990).

[15] R. Schneider, Numerische Mathematik **113** (3), 433–471 (2009).

[16] T. Rohwedder, ESAIM Math. Model. Numer. Anal. **47** (2), 421–447 (2013).

[17] T. Rohwedder and R. Schneider, ESAIM: Math. Model. Numer. Anal. **47** (6), 1553–1582 (2013).

[18] T.P. Živković, Int. J. Quantum. Chem. **12** (S11), 413–420 (1977).

[19] T.P. Živković and H.J. Monkhorst, J. Math. Phys. **19** (5), 1007–1022 (1978).

[20] P. Piecuch, S. Zarrabian, J. Paldus and J. Čížek, Phys. Rev. B **42** (6), 3351 (1990).

[21] K. Kowalski and K. Jankowski, Phys. Rev. Lett. **81** (6), 1195 (1998).

[22] P. Piecuch and K. Kowalski, in *Computational Chemistry: Reviews of Current Trends*, edited by J. Leszczynski (World Scientific, Singapore, 2000), Vol. 5.

[23] B. Jeziorski and J. Paldus, J. Chem. Phys. **90** (5), 2714–2731 (1989).

[24] W. Meyer, Int. J. Quantum. Chem. **5** (S5), 341–348 (1971).

[25] W. Meyer, J. Chem. Phys. **58** (3), 1017–1035 (1973).

[26] R. Ahlrichs, F. Driessler, H. Lischka, V. Staemmler and W. Kutzelnigg, J. Chem. Phys. **62** (4), 1235–1247 (1975).

[27] W. Kutzelnigg, *Methods of Electronic Structure Theory* (Plenum Press, New York, 1977), p.129.

[28] R. Ahlrichs, P. Scharf and C. Ehrhardt, J. Chem. Phys. **82** (2), 890–898 (1985).

[29] C. Kollmar and F. Neese, Mol. Phys. **108** (19–20), 2449–2458 (2010).

[30] F. Wennmohs and F. Neese, Chem. Phys. **343** (2–3), 217–230 (2008).

[31] A. Laestadius and S. Kvaal, SIAM. J. Numer. Anal. **56** (2), 660–683 (2018).

[32] F.M. Faulstich, A. Laestadius, S. Kvaal, Ö. Legeza and R. Schneider, arXiv preprint arXiv:1802.05699 (2018).

[33] T.J. Lee and P.R. Taylor, Int. J. Quantum. Chem. **36** (S23), 199–207 (1989).

[34] C.W. Kilmister and E. Schrödinger, *Schrödinger: Centenary Celebration of a Polymath* (Cambridge University Press, Cambridge, 1987).

[35] D. Yafaev, J. Funct. Anal. **168** (1), 121–144 (1999).

[36] H. Yserentant, *Regularity and Approximability of Electronic Wave Functions* (Springer, Berlin, 2010).

[37] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, New York, 2014).

[38] H.J. Monkhorst, Int. J. Quantum. Chem. **12** (S11), 421–432 (1977).

[39] W. Heisenberg, *Die Kopenhager Deutung der Quantentheorie* (Battenberg, Stuttgart, 1963).

[40] E. Zaidler, *Nonlinear Functional Analysis and Its Applications* (Springer, New York, 1990).

[41] J. Céa, Ann. Inst. Fourier (Grenoble) **14** (fasc. 2), 345–444 (1964).

[42] J.P. Aubin, Annali della Scuola Normale Superiore di Pisa-Classe di Scienze **21** (4), 599–637 (1967).

[43] J. Nitsche, Numerische Mathematik **11** (4), 346–348 (1968).

[44] L.A. Oganesyan and L.A. Rukhovets, USSR Comput. Math. Math. Phys. **9** (5), 158–183 (1969).

[45] T. Kinoshita, O. Hino and R.J. Bartlett, J. Chem. Phys. **123** (7), 074106 (2005).

[46] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations* (Birkhäuser, Basel, 2013).

[47] J. Arponen, Ann. Phys. **151** (2), 311–382 (1983).

[48] P.O. Löwdin, J. Math. Phys. **24** (1), 70–87 (1983).

[49] R. Bartlett, in *Geometrical Derivatives of Energy Surfaces and Molecular Properties*, edited by P. Jorgensen and J. Simons (Reidel, Dordrecht, 1986).

[50] T. Helgaker and P. Jørgensen, in *Advances in Quantum Chemistry*, edited by Per-Olov Löwdin (Academic Press, Cambridge, MA, 1988), Vol. 19, pp. 183–245.

[51] P. Piecuch and R.J. Bartlett, in *Advances in Quantum Chemistry* (Academic Press, Cambridge, MA, 1999), Vol. 34, pp. 295–380.