ANALYSIS OF THE SINGLE REFERENCE COUPLED CLUSTER METHOD FOR ELECTRONIC STRUCTURE CALCULATIONS: THE DISCRETE COUPLED CLUSTER EQUATIONS

MUHAMMAD HASSAN, YVON MADAY, AND YIPENG WANG

Abstract. Coupled cluster methods are widely regarded as the gold standard of computational quantum chemistry as they are perceived to offer the best compromise between computational cost and a high-accuracy resolution of the ground state eigenvalue of the electronic Hamiltonian– an unbounded, self-adjoint operator acting on a Hilbert space of antisymmetric functions that describes electronic properties of molecular systems. The present contribution is the second in a series of two articles where we introduce a new numerical analysis of the single-reference coupled cluster method based on the invertibility of coupled cluster Fréchet derivative. In this contribution, we study discretisations of the single-reference coupled cluster equations based on a prior mean-field (Hartree-Fock) calculation. We show that under some structural assumptions on the associated discretisation spaces and assuming that the discretisation is fine enough, the discrete coupled cluster equations are locally well-posed, and we derive a priori and residual-based a posteriori error estimates for the discrete coupled cluster solutions. Preliminary numerical experiments indicate that the structural assumptions that we impose for our analysis can be expected to hold for several small molecules, and the theoretical constants that appear in our error estimates are a significant improvement over those obtained from earlier approaches.

1. Introduction

The electronic structure problem is one of the most important many-body problems in modern computational physics and chemistry. From understanding the emergent superconducting properties of twisted bilayer graphene at certain magic angles [4] to explaining the complex mechanisms underlying light-harvesting molecules [6], an incredible range of scientific phenomena are governed by many-body electronic interactions. It is therefore hardly surprising that electronic structure calculations have a wide domain of applications from drug discovery and the creation of new compounds for sustainable energy and green catalysis (see, e.g., [7, 8, 14, 21, 22]), to the design of so-called quantum materials with exotic magnetic, ferroelectric or superconducting properties (see, e.g., [1, 12, 29]).

The central problem in molecular electronic structure is the numerical computation of the ground state eigenvalue of the electronic Hamiltonian– an unbounded, self-adjoint operator acting on a Hilbert space of antisymmetric functions. Since the spatial dimension of the Hilbert space grows linearly in the number of electrons N, the computational cost of a standard Galerkin discretisation of the electronic Hamiltonian typically scales exponentially in the degrees of freedom. State-of-the-art numerical methods in molecular electronic structure theory are therefore based on a *low-rank, non-linear parametrisation* of the sought-after eigenfunction.

Coupled cluster (CC) methods are one such class of algorithms, which are based on an *exponential ansatz* for the targeted ground state of the electronic Hamiltonian. More precisely, in the so-called single reference CC method, the sought-after ground state is expressed as the action of an exponential

2020 *Mathematics Subject Classification.* 65N25, 65N30, 65Z05, 81V55, 81V70.

Key words and phrases. Electronic structure theory, coupled cluster method, numerical analysis, non-linear functions, error estimates.

cluster operator– which is the operator exponential of a linear combination of bounded maps (so-called *excitation operators* defined in Section 3.1)– on a judiciously chosen reference function. Using this ansatz, the eigenvalue problem for the ground state energy of the electronic Hamiltonian can be reformulated as a highly *non-linear* system of equations for the unknown coefficients appearing in the linear combination of excitation operators entering the operator exponential. Approximations to the ground state energy are then obtained by restricting the class of excitation operators that appear inside the exponential, which leads to a hierarchy of computationally more tractable non-linear, root-finding problems. Usually these truncations are done on the basis of the excitation orders (see Appendix B below) and one thus speaks of CCSD (single and double excitation operators), CCSDT (single, double and triple excitation operators) and so on. In particular, the so-called CCSD(T)[1] variant, which can be applied to small and medium-sized molecules at a reasonable computational cost, is widely regarded as the 'gold standard' of quantum chemistry [23].

Despite being the method of choice for computing ground states of dynamically correlated molecules in the quantum chemical community (see, e.g., [18]), the mathematical analysis of coupled cluster methods is a relatively recent phenomenon. Indeed, the numerical analysis of single-reference coupled cluster was initiated less than 15 years ago by Schneider and Rohwedder [25, 26, 28], who introduced, for the first time, the correct functional analytic setting for formulating and analysing the coupled cluster equations. These seminal contributions also pioneered the use of the *strong local monotonicity* property, which is essentially a positivity condition on the Fréchet derivative of the coupled cluster function in a neighbourhood of the sought-after CC solution $\boldsymbol{t}^*$, as a means to demonstrate the well-posedness of these equations and derive a priori error estimates. This line of reasoning based on strong local monotonicity was subsequently extended to analyse further coupled cluster variants including the so-called extended coupled cluster method [17] and the tailored coupled cluster method [10].

As we discuss in more detail in Section 3.3 below, the primary drawback of the local monotonicity approach is that a rigorous demonstration of this property for coupled cluster functions requires a rather pessimistic smallness assumption on the sought-after coupled cluster solution $\boldsymbol{t}^*$. In other words, strong local monotonicity and the local well-posedness analysis and error estimates derived from this approach, can only be shown to hold in a perturbative regime $\boldsymbol{t}^* \approx 0$. On the other hand, in many practical situations where the CC method is *known numerically* to yield accurate approximations, the sought-after root $\boldsymbol{t}^*$ is *not* in the perturbative regime (see, e.g., Table 1 below). For such problems, the existing a priori analysis yields estimates with *negative* constants, and therefore does not provide a clear path towards error certification and validation. This is particularly troublesome since coupled cluster calculations are often used for benchmark computations and the calibration of Kohn-Sham density functional theory models which are used to create data sets for large-scale machine-learning driven predictive models.

Motivated by this unsatisfactory state of affairs, the present authors have introduced a new numerical analysis of the single-reference coupled cluster method based on directly establishing the invertibility of the Fréchet derivative of the coupled cluster function using classical inf-sup arguments. Using this approach, we have shown in a recent contribution [11] that, irrespective of the smallness of $\boldsymbol{t}^*$, the continuous (infinite-dimensional) CC equations are always locally well-posed provided that the sought-after ground state is non-degenerate and the chosen reference function is non-orthogonal to the targeted ground state wave-function. Preliminary numerical experiments also indicate (see, e.g., Table 1 below) that the constants appearing in the error estimates obtained through our approach are a significant improvement over those obtained from the existing local monotonicity-based analysis. Of course, the main drawback of our new approach is that– in contrast to the strong local monotonicity methodology– local well-posedness for *discretisations* of the continuous CC equations does not immediately follow

[1]Here, the (T) emphasises the fact that triple excitation orders are not initially included in the CCSD(T) ansatz and are rather treated perturbatively through a post-processing step.

from the continuous inf-sup condition. Indeed, this is a well-known phenomenon in the numerical analysis of linear PDEs such as the Helmholtz equation where a discrete inf-sup condition must be established separately for each given discretisation (see, e.g., [27, Chapter 4]).

The aim of the current contribution is to extend the new analysis developed in [11] to discretisations of the single-reference coupled cluster equations based on a prior mean-field (Hartree-Fock) calculation. We show that under some structural assumptions on the associated discretisation spaces and assuming that the discretisation is fine enough, the discrete coupled cluster equations are locally well-posed, and we derive a priori and residual-based a posteriori error estimates for the discrete coupled cluster solutions. The structural assumptions that we impose for our analysis are one of two types. The first type corresponds to the so-called Full-CC discretisations and the second type to the so-called excitation rank-truncated CC discretisations (see, e.g., [13, Chapter 13] and Appendix B below) together with a smallness assumption on a specific operator norm involving the difference of the electronic Hamiltonian and the mean-field operator. Preliminary numerical experiments indicate that this smallness assumption can be expected to hold for several small molecules, and the resulting theoretical constants that appear in our error estimates are once again a significant improvement over those obtained from earlier approaches.

The remainder of this article is organised as follows. In Section 2, we introduce the problem formulation, i.e., the electronic Hamiltonian and the Hilbert spaces on which it acts. In Section 3.1 and 3.2, we introduce excitation operators and cluster operators respectively which are fundamental mathematical objects in the coupled cluster methodology. Next, in Section 3.3, we formulate the infinite-dimensional continuous coupled cluster equations and we briefly recall the existing results on the well-posedness of these equations, including our previous contribution [11]. Subsequently, in Section 3.4, we state the discrete coupled cluster equations whose analysis is the main subject of this contribution. We begin our analysis in Section 4 where we first establish, in Subsections 4.1 and 4.2, a number of technical lemmas. Finally, in Section 4.3, we state and prove our main result on the local well-posedness and a priori and residual-based a posteriori error estimates for the discrete coupled cluster equations. Supplementary information related to the Hartree-Fock methodology, the practicality of the assumptions needed for our analysis, as well as some auxiliary lemmas is given in Appendices A, B, and C respectively.

2. Background and Problem Setting

Computational quantum chemistry is the study of the properties of matter through modelling at the molecular scale, i.e., when matter is viewed as a collection of positively charged nuclei and negatively charged electrons. To formalise the problem setting, we assume that we are given a molecule composed of $M \in \mathbb{N}$ nuclei carrying charges $\{Z_\alpha\}_{\alpha=1}^M \subset \mathbb{R}_+$ and located at positions $\{\mathbf{x}_\alpha\}_{\alpha=1}^M \subset \mathbb{R}^3$ respectively. We further assume the presence of $N \in \mathbb{N}$ electrons whose spatial coordinates are denoted by $\{\mathbf{x}_i\}_{i=1}^N \subset \mathbb{R}^3$. Throughout this article, we will assume that the Born-Oppenheimer approximation holds, i.e., we will treat the nuclei as fixed, classical particles and we will focus purely on the quantum mechanical description of the electrons.

In order to describe the behaviour of this system of nuclei and electrons under the Born-Oppenheimer approximation, we require the notion of several functions spaces. The following construction is largely a repetition of the one found in our previous contribution [11], which itself was based on [24].

2.1. **Function Spaces and Norms.**

To begin with, we denote by $\mathrm{L}^2(\mathbb{R}^3)$ the space of real-valued square integrable functions of three variables, and we denote by $\mathrm{H}^1(\mathbb{R}^3)$ the closed subspace of $\mathrm{L}^2(\mathbb{R}^3)$ consisting of functions that additionally possess square integrable first derivatives. Both spaces are equipped with their usual inner products. Following the convention in the quantum chemical literature, we will frequently refer to $\mathrm{L}^2(\mathbb{R}^3)$ and $\mathrm{H}^1(\mathbb{R}^3)$ as infinite-dimensional *single particle* spaces.

Next, we define the tensor space

$$\mathscr{L}^2 := \bigotimes_{j=1}^{N} \mathrm{L}^2(\mathbb{R}^3),$$

which is equipped with an inner product that is constructed by defining first for all elementary tensors $f, g \in \mathscr{L}^2$ with $f = \otimes_{j=1}^N f_j$ and $g = \otimes_{j=1}^N g_j$

$$(f, g)_{\mathscr{L}^2} := \prod_{j=1}^{N} (f_j, g_j)_{\mathrm{L}^2(\mathbb{R}^3)}, \tag{1}$$

and then extending bilinearly for general tensorial elements of $\mathscr{L}^2$.

It is a consequence of Fubini's theorem that the tensor space $\mathscr{L}^2$ is isometrically isomorphic to the space $\mathrm{L}^2(\mathbb{R}^{3N})$ of real-valued square integrable functions of $3N$ variables with the associated L^2-inner product. Thanks to this result, we can define the tensor space $\mathscr{H}^1 \subset \mathscr{L}^2$ as the closure of $\mathscr{C}_0^\infty(\mathbb{R}^{3N})$ in $\mathrm{L}^2(\mathbb{R}^{3N})$ with respect to the usual gradient-gradient inner product on $\mathbb{R}^{3N}$.

In quantum mechanics, a fundamental distinction is made between so-called *bosonic* and *fermionic* particles, the latter obeying the so-called Pauli-exclusion principle and thus being described in terms of antisymmetric functions. We are therefore obligated to also define tensor spaces of antisymmetric functions. To this end, we first introduce the so-called *antisymmetric projection operator* $\mathbb{P}^{\mathrm{as}} \colon \mathscr{L}^2 \to \mathscr{L}^2$ that is defined through the action

$$\forall f \in \mathscr{L}^2 \colon \quad (\mathbb{P}^{\mathrm{as}} f)(\mathbf{x}_1, \dots, \mathbf{x}_N) := \frac{1}{\sqrt{N!}} \sum_{\pi \in \mathrm{S}(N)} (-1)^{\mathrm{sgn}(\pi)} f(\mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(N)}), \tag{2}$$

where $\mathrm{S}(N)$ denotes the permutation group of order N, and $\mathrm{sgn}(\pi)$ denotes the signature of $\pi \in \mathrm{S}(N)$.

It is easy to establish that $\mathbb{P}^{\mathrm{as}}$ is an $\mathscr{L}^2$-orthogonal projection with a closed range. We therefore define the antisymmetric tensor spaces $\widehat{\mathscr{L}}^2 \subset \mathscr{L}^2$ and $\widehat{\mathscr{H}}^1 \subset \mathscr{H}^1$ as

$$\widehat{\mathscr{L}}^2 := \bigwedge_{j=1}^{N} \mathrm{L}^2(\mathbb{R}^3) := \mathrm{ran}\ \mathbb{P}^{\mathrm{as}} \quad \text{and} \quad \widehat{\mathscr{H}}^1 := \widehat{\mathscr{L}}^2 \cap \mathscr{H}^1,$$

equipped with the $(\cdot,\cdot)_{\mathscr{L}^2}$ and $(\cdot,\cdot)_{\mathscr{H}^1}$ inner products respectively. We remark that normalised elements of $\widehat{\mathscr{L}}^2$ are known as *wave-functions*, and these are antisymmetric in the sense that for any $f \in \widehat{\mathscr{L}}^2$ we have that

$$f(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots \mathbf{x}_N) = -f(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots \mathbf{x}_N) \qquad \forall\ i, j \in \{1, \dots, N\} \text{ with } i \neq j.$$

In the sequel, we will also occasionally make use of the dual space of $\widehat{\mathscr{H}}^1$. We therefore denote $\widehat{\mathscr{H}}^{-1} := \big(\widehat{\mathscr{H}}^1\big)^*$, we equip $\widehat{\mathscr{H}}^{-1}$ with the canonical dual norm, and we write $\langle \cdot, \cdot \rangle_{\widehat{\mathscr{H}}^1, \widehat{\mathscr{H}}^{-1}}$ for the associated duality pairing. Note that higher regularity Sobolev spaces $\widehat{\mathscr{H}}^s,\ s \geq 1$ can be defined similarly to $\widehat{\mathscr{H}}^1$.

Let us remark here that, as is common for numerical analyses of the kind that we undertake in this paper, we have ignored spin variables in the construction of our function spaces (see, e.g., Remark 1 in our previous contribution [11] for a discussion of this point.)

2.2. **Governing Operators and Problem Statement.**

Throughout this article, we assume that the electronic properties of the molecule that we study are described by the action of a many-body electronic Hamiltonian given by

$$H := -\frac{1}{2}\sum_{j=1}^{N}\Delta_{\mathbf{x}_j} + \sum_{j=1}^{N}\sum_{\alpha=1}^{M}\frac{-Z_\alpha}{|\mathbf{x}_\alpha - \mathbf{x}_j|} + \sum_{j=1}^{N}\sum_{i=1}^{j-1}\frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \qquad \text{acting on } \widehat{\mathcal{L}}^2 \quad \text{with domain } \widehat{\mathscr{H}}^2. \tag{3}$$

The electronic properties of the molecule that we study are functions of the spectrum of the electronic Hamiltonian H, and we are therefore interested in its analysis and computation. It is a classical result (see, e.g., the review article [16]) that the operator H is self-adjoint on $\widehat{\mathcal{L}}^2$ with form domain $\widehat{\mathscr{H}}^1$, and under the additional assumption that $Z := \sum_{\alpha=1}^{M} Z_\alpha > N-1$, it holds that

(1) The operator H has an essential spectrum σ_{ess} of the form $\sigma_{\text{ess}} := [\Sigma, \infty)$ where $-\infty < \Sigma \leq 0$;

(2) The operator H has a bounded-below discrete spectrum that consists of a countably infinite number of eigenvalues, each with finite multiplicity, accumulating at Σ.

Consequently, under the assumption that $\sum_{\alpha=1}^{M} Z_\alpha \geq N$, the electronic Hamiltonian $\mathscr{H}$ possesses a lowest eigenvalue $\mathcal{E}^*_{\text{GS}} \in \mathbb{R}$, frequently called the ground state energy, such that

$$\mathcal{E}^*_{\text{GS}} = \min_{0\neq\Psi\in\widehat{\mathscr{H}}^1} \frac{\big\langle \Psi, H\Psi\big\rangle_{\widehat{\mathscr{H}}^1\times\widehat{\mathscr{H}}^{-1}}}{\|\Psi\|^2_{\widehat{\mathcal{L}}^2}}. \tag{4a}$$

Any function $\Psi^*_{\text{GS}} \in \widehat{\mathscr{H}}^1$ that achieves the minimum in Equation (4a) is called a ground state of H and obviously satisfies

$$H\Psi^*_{\text{GS}} = \mathcal{E}^*_{\text{GS}}\Psi^*_{\text{GS}}. \tag{4b}$$

For the purpose of this article, we will assume that indeed $Z = \sum_{\alpha=1}^{M} Z_\alpha > N-1$. Note that if the ground state eigenvalue $\mathcal{E}^*_{\text{GS}}$ is simple (which is not always the case), normalised ground states Ψ^*_{GS} (being elements of a real Hilbert space) are unique up to sign. Consequently, when $\mathcal{E}^*_{\text{GS}}$ is simple, we will simply refer to *the* ground state Ψ^*_{GS} with the convention that the sign of Ψ^*_{GS} has been fixed once and for all.

From a functional analysis point of view, the electronic Hamiltonian H possesses certain desirable properties, namely continuity and ellipticity on appropriate Sobolev spaces. More precisely (see, for instance, [31, Chapter 4]),

- The electronic Hamiltonian defined through Equation (3) is bounded as a mapping from $\widehat{\mathscr{H}}^1$ to $\widehat{\mathscr{H}}^{-1}$:

$$\forall \Phi, \Psi \in \widehat{\mathscr{H}}^1: \qquad \Big|\langle \Phi, H\Psi\rangle_{\widehat{\mathscr{H}}^1\times\widehat{\mathscr{H}}^{-1}}\Big| \leq \Big(\frac{1}{2} + 3\sqrt{N}Z\Big)\|\Phi\|_{\widehat{\mathscr{H}}^1}\|\Psi\|_{\widehat{\mathscr{H}}^1}; \tag{5}$$

- The electronic Hamiltonian defined through Equation (3) satisfies the following ellipticity condition on the Gelfand triple $\widehat{\mathscr{H}}^1 \hookrightarrow \widehat{\mathcal{L}}^2 \hookrightarrow \widehat{\mathscr{H}}^{-1}$:

$$\forall \Phi \in \widehat{\mathscr{H}}^1: \qquad \langle \Phi, H\Phi\rangle_{\widehat{\mathscr{H}}^1\times\widehat{\mathscr{H}}^{-1}} \geq \frac{1}{4}\|\Phi\|^2_{\widehat{\mathscr{H}}^1} - \Big(9NZ^2 - \frac{1}{4}\Big)\|\Phi\|^2_{\widehat{\mathcal{L}}^2}. \tag{6}$$

An important consequence of the above ellipticity estimate is that the electronic Hamiltonian, modified by any suitable shift, defines an invertible operator on a subspace of $\widehat{\mathscr{H}}^1$. This fact will be of great importance in our analysis in Sections 3 and 4.

3. Formulation of the Coupled Cluster Equations

Post-Hartree Fock numerical methods in quantum chemistry such as the single-reference coupled cluster method are based on a particular decomposition of the single-particle space $\mathrm{H}^1(\mathbb{R}^3;\mathbb{C})$ and the N-particle space $\widehat{\mathscr{H}}^1$. In order to state the single-reference coupled cluster equations, we first describe this decomposition in more detail.

Notation 1 (Occupied and Virtual Spaces)**.**

Let $\mathscr{R}$ denote an N-dimensional subspace of the single-particle space $\mathrm{H}^1(\mathbb{R}^3)$ and let $\mathscr{R}^\perp \subset \mathrm{L}^2(\mathbb{R}^3)$ denote the $\mathrm{L}^2(\mathbb{R}^3)$-orthogonal complement of $\mathscr{R}$, i.e.,

$$\mathscr{R}^\perp := \left\{\phi \in \mathrm{L}^2(\mathbb{R}^3)\colon (\phi,\psi)_{\mathrm{L}^2(\mathbb{R}^3)} = 0 \ \forall \psi \in \mathscr{R}\right\}.$$

We refer to $\mathscr{R}$ as an occupied space of $\mathrm{H}^1(\mathbb{R}^3;\mathbb{C})$, and we refer to $\mathscr{R}^\perp$ as a virtual space of $\mathrm{H}^1(\mathbb{R}^3;\mathbb{C})$ respectively.

Definition 2 (Reference Determinant for an Occupied Space)**.**

Let $\mathbb{P}^{\mathrm{as}}\colon \mathscr{L}^2 \to \mathscr{L}^2$ denote the anti-symmetric projection operator defined through Equation (2), let $\mathscr{R}$ denote an occupied space of $\mathrm{H}^1(\mathbb{R}^3)$ as introduced in Notation 1, and let $\mathscr{B}_{\mathrm{occ}} := \{\psi_j\}_{j=1}^N \subset \mathscr{R}$ denote any $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis for $\mathscr{R}$. Then we define the function $\Psi_0 \in \widehat{\mathscr{H}}^1$ as

$$\Psi_0 := \mathbb{P}^{\mathrm{as}}\left(\otimes_{i=1}^N \psi_i\right),$$

and we say that Ψ_0 is a reference determinant corresponding to the occupied space $\mathscr{R}$.

An important remark is now in order.

Remark 3 (Uniqueness up to sign of Reference Determinant)**.** *Consider Definition 2 of a reference determinant $\Psi_0 \in \widehat{\mathscr{H}}^1$ constructed from an $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis $\mathscr{B}_{\mathrm{occ}} := \{\psi_j\}_{j=1}^N$ of some occupied space $\mathscr{R}$. Given now a different $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis $\mathscr{B}'_{\mathrm{occ}} := \{\psi'_j\}_{j=1}^N$, there obviously exists a unitary matrix $\mathbf{U} = [\mathbf{U}]_{ij} \in \mathbb{R}^{N\times N}$ such that*

$$\forall j \in \{1,\ldots,N\}\colon \qquad \psi'_i := \sum_{j=1}^N [\mathbf{U}]_{ij}\psi_j.$$

It follows from the definition of the anti-symmetric projection operator $\mathbb{P}^{\mathrm{as}}$ that

$$\mathbb{P}^{\mathrm{as}}\left(\otimes_{i=1}^N \psi'_i\right) = \det(\mathbf{U})\mathbb{P}^{\mathrm{as}}\left(\otimes_{i=1}^N \psi_i\right) = \det(\mathbf{U})\Psi_0.$$

Consequently, reference determinants corresponding to a given occupied space $\mathscr{R} \subset \mathrm{H}^1(\mathbb{R}^3)$ are unique up to sign. In the sequel therefore, given an occupied space $\mathscr{R}$ of the single-particle space $\mathrm{H}^1(\mathbb{R}^3)$, we will simply refer to the reference determinant Ψ_0 corresponding to $\mathscr{R}$ with the convention that the sign of Ψ_0 has been fixed once and for all.

The motivation for introducing reference determinants is that they allow a specific complementary decomposition of the N-particle function space $\widehat{\mathscr{H}}^1$.

Definition 4 (Complementary Decomposition of N-particle Function Space)**.**

Let $\mathscr{R}$ denote an occupied space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\Psi_0 \in \widehat{\mathscr{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, and let $\{\Psi_0\}^\perp \subset \widehat{\mathscr{L}}^2$ denotes the $\widehat{\mathscr{L}}^2$-orthogonal complement of Ψ_0. Then we define the infinite-dimensional subspace $\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} \subset \widehat{\mathscr{H}}^1$ as

$$\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} := \{\Psi_0\}^\perp \cap \widehat{\mathscr{H}}^1,$$

and we introduce the complementary decomposition of $\widehat{\mathscr{H}}^1$ given by

$$\widehat{\mathscr{H}}^1 = \mathrm{span}\{\Psi_0\} \oplus \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}. \tag{7}$$

Additionally we define $\mathbb{P}_0 \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1$ and $\mathbb{P}_0^\perp := \mathrm{I} - \mathbb{P}_0 \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1$ as the $\widehat{\mathscr{L}}^2$-orthogonal projection operator onto $\mathrm{span}\,\{\Psi_0\}$ and $\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ respectively.

Another important remark is now in order.

Remark 5 (Properties of Complementary Decomposition of N-particle Function Space)**.**

Consider Definition 4 of the $\widehat{\mathscr{L}}^2$-orthogonal projection operators $\mathbb{P}_0 \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1$ and $\mathbb{P}_0^\perp \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1$.

The fact that $\mathbb{P}_0$ and $\mathbb{P}_0^\perp$ are both bounded operators with respect to the $\|\cdot\|_{\widehat{\mathscr{H}}^1}$ norm is a consequence of the fact that these operators possess a range and a kernel that are both closed in the $\widehat{\mathscr{H}}^1$ topology. Let us also point out that these projection operators allow us to equivalently express the complementary decomposition (7) as

$$\widehat{\mathscr{H}}^1 = \mathrm{Ran}\mathbb{P}_0 \oplus \widehat{\mathscr{H}}^{1,\perp} = \mathrm{Ran}\mathbb{P}_0 \oplus \mathrm{Ran}\mathbb{P}_0^\perp.$$

3.1. **Excitation Operators.**

Our next task is to introduce the notion of so-called *excitation* operators which lie at the core of the coupled cluster methodology. As the first step towards doing so, we will introduce suitable orthonormal bases for the N-particle function space $\widehat{\mathscr{H}}^1$ and its subspace $\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ introduced above.

Definition 6 (Slater Determinant Basis for N-particle function space $\widehat{\mathscr{H}}^1$)**.**

Let $\mathscr{R}$ and $\mathscr{R}^\perp$ denote an occupied and virtual space respectively of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\mathscr{B}_{\mathrm{occ}} := \{\psi_j\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ and $\mathscr{B}_{\mathrm{vir}} := \{\psi_j\}_{j=N+1}^\infty \subset \mathrm{H}^1(\mathbb{R}^3)$ denote $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal bases for $\mathscr{R}$ and $\mathscr{R}^\perp$ respectively, let $\mathcal{J}_\infty^N \subset \mathbb{N}^N$ be the index set defined as

$$\mathcal{J}_\infty^N := \Big\{\boldsymbol{\ell} = (\ell_1, \ell_2, \ldots, \ell_N) \in \mathbb{N}^N \colon \ell_1 < \ell_2 < \ldots < \ell_N\Big\},$$

and let the set $\mathscr{B}_\otimes^{\mathrm{ord}} \subset \mathscr{H}^1$ be defined as

$$\mathscr{B}_\otimes^{\mathrm{ord}} := \Big\{\Psi_{\mathbf{k}} := \psi_{k_1} \otimes \psi_{k_2} \otimes \ldots \otimes \psi_{k_N} \colon \mathbf{k} = (k_1, \ldots, k_N) \in \mathcal{J}_\infty^N\Big\}.$$

Then we define an $\widehat{\mathscr{L}}^2$-orthonormal basis for $\widehat{\mathscr{H}}^1$ as

$$\begin{aligned}\mathscr{B}_\wedge :=& \{\mathbb{P}^{\mathrm{as}}\Psi \colon \Psi \in \mathscr{B}_\otimes^{\mathrm{ord}}\} \\ =& \left\{\Psi_{\mathbf{k}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \sum_{\pi \in \mathrm{S}(N)} (-1)^{\mathrm{sgn}(\pi)} \otimes_{i=1}^N \psi_{k_i}\big(\mathbf{x}_{\pi(i)}\big) \colon \mathbf{k} = (k_1, k_2, \ldots, k_N) \in \mathcal{J}_\infty^N\right\},\end{aligned}$$

and we refer to the elements of this basis set as Slater determinants.

Two comments are now in order.

Notation 7 (Slater Determinants)**.**

Consider Definition 6 of the $\widehat{\mathcal{L}}^2$-orthonormal basis set $\mathcal{B}_\wedge$. For simplicity, given $\mathbf{k} \in \mathcal{J}_\infty^N$ and a Slater determinant $\Psi_{\mathbf{k}} \in \mathcal{B}_\wedge$ of the form

$$\Psi_{\mathbf{k}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \sum_{\pi \in \mathrm{S}(N)} (-1)^{\mathrm{sgn}(\pi)} \otimes_{i=1}^N \psi_{k_i}\big(\mathbf{x}_{\pi(i)}\big),$$

we will write the Slater determinant $\Psi_{\mathbf{k}}$ in the succinct form

$$\Psi_{\mathbf{k}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det\big(\psi_{k_i}(\mathbf{x}_j)\big)_{i,j=1}^N.$$

Definition 8 (Slater Determinant Basis for $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$)**.**

Let $\mathscr{R}$ denote an occupied space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\mathcal{B}_\wedge$ denote an $\widehat{\mathcal{L}}^2$-orthonormal basis set for the N-particle function space $\widehat{\mathcal{H}}^1$ as defined through Definition 6, let $\Psi_0 \in \widehat{\mathcal{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, and let the infinite-dimensional subspace $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0} \subset \widehat{\mathcal{H}}^1$ be defined as in Definition 4.

Then we define an $\widehat{\mathcal{L}}^2$-orthonormal basis for $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ as

$$\widetilde{\mathcal{B}_\wedge} := \mathcal{B}_\wedge \setminus \{\Psi_0\}.$$

Equipped with Definitions 6 and 8 for our N-particle basis sets, we are now ready to define the so-called excitation operators. To do so, we will first introduce the notion of excitation index sets.

Definition 9 (Excitation Index Sets)**.**

For each $j \in \{1, \ldots, N\}$ we define the index set $\mathcal{I}_j$ as

$$\mathcal{I}_j := \left\{ \binom{a_1, \ldots, a_j}{\ell_1, \ldots, \ell_j} : \ell_1 < \ldots < \ell_j \in \{1, \ldots, N\} \text{ and } a_1 < \ldots < a_j \in \{N+1, N+2, \ldots\} \right\},$$

and we say that $\mathcal{I}_j$ is the excitation index set of order j. Additionally, we define

$$\mathcal{I} := \bigcup_{j=1}^N \mathcal{I}_j,$$

and we say that $\mathcal{I}$ is the global excitation index set.

Definition 10 (Excitation Operators)**.**

Let $j \in \{1, \ldots, N\}$, let $\mu \in \mathcal{I}_j$ be of the form

$$\mu = \binom{a_1, \ldots, a_j}{\ell_1, \ldots, \ell_j} : \ell_1 < \ldots < \ell_j \in \{1, \ldots, N\} \text{ and } a_1 < \ldots < a_j \in \{N+1, N+2, \ldots\},$$

and let $\mathcal{B}_\wedge$ denote an $\widehat{\mathcal{L}}^2$-orthonormal basis set for the N-particle function space $\widehat{\mathcal{H}}^1$ as defined through Definition 6.

We define the excitation operator $\mathcal{X}_\mu \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ *through its action on the* N*-particle basis set* $\mathcal{B}_\wedge$*. For* $\Psi_\nu(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det\left(\psi_{\nu_j}(\mathbf{x}_i)\right)_{i,j=1}^N$*, we set*

$$\mathcal{X}_\mu \Psi_\nu = \begin{cases} 0 & \text{if } \{\ell_1, \ldots, \ell_j\} \not\subset \{\nu_1, \ldots, \nu_N\}, \\ 0 & \text{if } \exists a_m \in \{a_1, \ldots, a_j\} \text{ such that } a_m \in \{\nu_1, \ldots, \nu_N\}, \\ \Psi_{\nu,a} \in \mathcal{B}_\wedge & \text{otherwise}, \end{cases}$$

where the determinant $\Psi_{\nu,a}$ *is constructed from* Ψ_ν *by replacing all functions* $\psi_{\ell_1}, \ldots, \psi_{\ell_j}$ *used to construct* Ψ_ν *with functions* $\psi_{a_1}, \ldots \psi_{a_j}$ *respectively.*

Definition 11 (De-excitation Operators)**.**

Let $j \in \{1, \ldots, N\}$*, let* $\mu \in \mathcal{I}_j$ *be of the form*

$$\mu = \binom{a_1, \ldots, a_j}{\ell_1, \ldots, \ell_j} \colon \ell_1 < \ldots < \ell_j \in \{1, \ldots, N\} \text{ and } a_1 < \ldots < a_j \in \{N+1, N+2, \ldots\},$$

and let $\mathcal{B}_\wedge$ *denote an* $\widehat{\mathcal{L}}^2$*-orthonormal basis set for the* N*-particle function space* $\widehat{\mathcal{H}}^1$ *as defined through Definition 6.*

We define the de-excitation operator $\mathcal{X}_\mu^\dagger \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ *through its action on the* N*-particle basis set* $\mathcal{B}_\wedge$*. For* $\Psi_\nu(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det\left(\psi_{\nu_j}(\mathbf{x}_i)\right)_{i,j=1}^N$*, we set*

$$\mathcal{X}_\mu^\dagger \Psi_\nu = \begin{cases} 0 & \text{if } \{a_1, \ldots, a_j\} \not\subset \{\nu_1, \ldots, \nu_N\}, \\ 0 & \text{if } \exists\ \ell_m \in \{\ell_1, \ldots, \ell_j\} \text{ such that } \ell_m \in \{\nu_1, \ldots, \nu_N\}, \\ \Psi_\nu^a \in \mathcal{B}_\wedge & \text{otherwise}, \end{cases}$$

where the determinant Ψ_ν^a *is constructed from* Ψ_ν *by replacing all functions* $\psi_{a_1}, \ldots \psi_{a_j}$ *used to construct* Ψ_ν *with functions* $\psi_{\ell_1}, \ldots, \psi_{\ell_j}$ *respectively.*

Consider Definitions 10 and 11 of excitation and de-excitation operators. Two important comments are now in order.

- First, de-excitation operators reverse the action of excitation operators in a specific sense. To be precise, for any excitation index μ in the global excitation index $\mathcal{I}$ and any determinant $\Psi_\nu \in \mathcal{B}_\wedge$, if $\mathcal{X}_\mu \Psi_\nu \neq 0$, then it holds that $\mathcal{X}_\mu^\dagger \mathcal{X}_\mu \Psi_\nu = \Psi_\nu$.
- Second, we emphasise that the definition of both excitation and de-excitation operators depends a priori on the chosen basis of Slater determinants $\mathcal{B}_\wedge$ for the N-particle space $\widehat{\mathcal{H}}^1$. The N-particle Slater basis $\mathcal{B}_\wedge$ itself depends on the basis sets $\mathscr{B}_{\text{occ}}$ and $\mathscr{B}_{\text{vir}}$ for the occupied space $\mathscr{R}$ and the virtual space $\mathscr{R}^\perp$ respectively. This dependency will be the subject of further discussion in the sequel.

The following theorem summarises some basic properties of excitation and de-excitation operators.

Theorem 12 (Properties of Excitation and De-excitation Operators)**.**

Let the excitation index set $\mathcal{I}$ *be defined through Definition 9, let* $\mathcal{B}_\wedge$ *denote an* $\widehat{\mathcal{L}}^2$*-orthonormal basis of Slater determinants for* $\widehat{\mathcal{H}}^1$ *as defined through Definition 6, and let the excitation operators* $\{\mathcal{X}_\mu\}_{\mu \in \mathcal{I}}$ *and de-excitation operators* $\{\mathcal{X}_\mu^\dagger\}_{\mu \in \mathcal{I}}$ *be defined through Definitions 10 and 11 respectively. Then*

(1) For all $\mu, \nu \in \mathcal{I}$*, it holds that* $\mathcal{X}_\mu \mathcal{X}_\nu = \mathcal{X}_\nu \mathcal{X}_\mu$ *and* $\mathcal{X}_\mu^\dagger \mathcal{X}_\nu^\dagger = \mathcal{X}_\nu^\dagger \mathcal{X}_\mu^\dagger$ *.*

(2) *For all* $\mu \in \mathcal{I}$ *the de-excitation operator* $\mathcal{X}_\mu^\dagger \colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ *is the* $\widehat{\mathcal{L}}^2$*-adjoint of the excitation operator* $\mathcal{X}_\mu \colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$.

(3) *For all* $\mu \in \mathcal{I}$ *the excitation operator* $\mathcal{X}_\mu \colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ *and de-excitation operator* $\mathcal{X}_\mu^\dagger \colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ *are bounded linear maps.*

(4) *For all* $\mu \in \mathcal{I}$ *the excitation operator* $\mathcal{X}_\mu \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ *and de-excitation operator* $\mathcal{X}_\mu^\dagger \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ *are bounded linear maps.*

Proof. See [25, Section 2]. □

The primary motivation for introducing the complicated notion of excitation operators is that they provide a systematic means of indexing an N-particle Slater basis $\mathcal{B}_\wedge$ in terms of their application on the reference determinant Ψ_0. This observation is summarised more precisely in the following remark.

Remark 13 (Interpretation of N-particle Bases in terms of Excited Determinants)**.**

Let $\mathscr{R}$ *and* $\mathscr{R}^\perp$ *denote an occupied and virtual space respectively of* $\mathrm{H}^1(\mathbb{R}^3)$ *as introduced in Notation 1, let* Ψ_0 *denote the reference determinant corresponding to* $\mathscr{R}$*, let the infinite-dimensional subspace* $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0} \subset \widehat{\mathcal{H}}^1$ *be defined as in Definition 4, let* $\mathscr{B}_{\mathrm{occ}}$ *and* $\mathscr{B}_{\mathrm{vir}}$ *denote* $\mathrm{L}^2(\mathbb{R}^3)$*-orthonormal basis sets for* $\mathscr{R}$ *and* $\mathscr{R}^\perp$ *respectively, and let* $\mathcal{B}_\wedge$ *and* $\widetilde{\mathcal{B}_\wedge}$ *denote* $\widehat{\mathcal{L}}^2$*-orthonormal basis sets for* $\widehat{\mathcal{H}}^1$ *and* $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ *respectively, constructed from* $\mathscr{B}_{\mathrm{occ}}$ *and* $\mathscr{B}_{\mathrm{vir}}$*, as outlined in Definition 6.*

It is a simple exercise to show that

$$\mathcal{B}_\wedge = \{\Psi_0\} \cup \{\mathcal{X}_\mu \Psi_0 \colon \mu \in \mathcal{I}\} \qquad \text{and} \qquad \widetilde{\mathcal{B}_\wedge} = \{\mathcal{X}_\mu \Psi_0 \colon \mu \in \mathcal{I}\}\,. \tag{8}$$

3.2. **Cluster Operators.**

In view of Remark 13, we see that for any arbitrary function $\Phi \in \widehat{\mathcal{H}}^1$, there exists a constant $\boldsymbol{s}_0 \in \mathbb{R}$ and a sequence $\{\boldsymbol{s}_\mu\}_{\mu \in \mathcal{I}} \subset \mathbb{R}$ such that

$$\Phi = \boldsymbol{s}_0 \Psi_0 + \sum_{\mu \in \mathcal{I}} \boldsymbol{s}_\mu \mathcal{X}_\mu \Psi_0, \tag{9}$$

where $\{\mathcal{X}_\mu\}_{\mu \in \mathcal{I}}$ denote excitation operators defined according to Definition 10.

Equation (9) suggests that any element of the N-particle space $\widehat{\mathcal{H}}^1$ can be expressed in terms of an infinite weighted summation of excitation operators acting on the reference determinant Ψ_0. Such infinite summations of excitation operators are known as *cluster* operators in the quantum chemistry literature. Of course, while each excitation operator itself is a bounded map from $\widehat{\mathcal{L}}^2$ to $\widehat{\mathcal{L}}^2$, it does not immediately follow that an infinite summation of excitation operators is a well-defined mapping on $\widehat{\mathcal{L}}^2$. Fortunately, the following result was proven in [25, Theorem 2.7].

Proposition 14 (Cluster Operators as Bounded Maps on $\widehat{\mathcal{L}}^2$)**.**

Let the excitation index set $\mathcal{I}$ *be defined through Definition 9, let* $\mathcal{B}_\wedge$ *denote an* $\widehat{\mathcal{L}}^2$*-orthonormal basis of Slater determinants for* $\widehat{\mathcal{H}}^1$ *as defined in Definition 6, let* $\Psi_0 \in \mathcal{B}_\wedge$ *denote the reference determinant as defined through Definition 2, and let the excitation operators* $\{\mathcal{X}_\mu\}_{\mu \in \mathcal{I}}$ *and de-excitation operators* $\{\mathcal{X}_\mu^\dagger\}_{\mu \in \mathcal{I}}$ *be defined through Definitions 10 and 11 respectively. Then*

(1) For any square-summable sequence $\boldsymbol{t} = \{\boldsymbol{t}_\mu\}_{\mu\in\mathcal{J}} \in \ell^2(\mathcal{J})$*, there exists a unique bounded linear operator* $\mathcal{T}\colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$*, the so-called* cluster operator *generated by* $\boldsymbol{t}$*, such that*

$$\mathcal{T} := \sum_{\mu\in\mathcal{J}} \boldsymbol{t}_\mu \mathcal{X}_\mu$$

where the series convergence holds with respect to the $\|\cdot\|_{\widehat{\mathcal{L}}^2\to\widehat{\mathcal{L}}^2}$ *operator norm. Moreover, there exists a constant* $\beta > 0$*, depending only on* N*, such that*

$$\|\boldsymbol{t}\|_{\ell^2(\mathcal{J})} \leq \|\mathcal{T}\|_{\widehat{\mathcal{L}}^2\to\widehat{\mathcal{L}}^2} \leq \beta\|\boldsymbol{t}\|_{\ell^2(\mathcal{J})}.$$

(2) In particular, for any $\Phi := \sum_{\mu\in\mathcal{J}} \boldsymbol{t}_\mu \mathcal{X}_\mu \Psi_0 \in \{\Psi_0\}^\perp \subset \widehat{\mathcal{L}}^2$ *there exists a unique cluster operator* $\mathcal{T}(\Phi) : \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ *such that*

$$\mathcal{T}(\Phi) = \sum_{\mu\in\mathcal{J}} \boldsymbol{t}_\mu \mathcal{X}_\mu$$

where the series convergence again holds with respect to the $\|\cdot\|_{\widehat{\mathcal{L}}^2\to\widehat{\mathcal{L}}^2}$ *operator norm.*

Moreover, the mapping $\{\Psi_0\}^\perp \ni \Phi \mapsto \mathcal{T}(\Phi)$ *is linear and bounded, i.e., there exists a constant* $\beta > 0$*, depending only on* N*, such that*

$$\|\Phi\|_{\widehat{\mathcal{L}}^2} \leq \|\mathcal{T}(\Phi)\|_{\widehat{\mathcal{L}}^2\to\widehat{\mathcal{L}}^2} \leq \beta\|\Phi\|_{\widehat{\mathcal{L}}^2}.$$

(3) For any square-summable sequence $\boldsymbol{t} = \{\boldsymbol{t}_\mu\}_{\mu\in\mathcal{J}} \in \ell^2(\mathcal{J})$ *with corresponding cluster operator* $\mathcal{T}\colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$*, the* $\widehat{\mathcal{L}}^2$*-adjoint* $\mathcal{T}^\dagger\colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ *satisfies*

$$\mathcal{T}^\dagger = \sum_{\mu\in\mathcal{J}} \boldsymbol{t}_\mu \mathcal{X}_\mu^\dagger$$

where the series convergence once again holds with respect to the $\|\cdot\|_{\widehat{\mathcal{L}}^2\to\widehat{\mathcal{L}}^2}$ *operator norm.*

Consider Proposition 14 that describes the mapping properties of cluster operators and Definition 10 of the excitation operators $\{\mathcal{X}_\mu\}_{\mu\in\mathcal{J}}$. We recall that excitation operators have a priori been defined relative to an N-particle Slater basis $\mathcal{B}_\wedge$ constructed from a specific choice of occupied space $\mathscr{R}$ as well as $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis sets $\mathscr{B}_{\mathrm{occ}}$ and $\mathscr{B}_{\mathrm{vir}}$ for the occupied and virtual space $\mathscr{R}$ and $\mathscr{R}^\perp$ respectively. At first glace therefore, it might seem that the definition of a cluster operator $\mathcal{T}(\Phi)$ generated by some $\Phi \in \{\Psi_0\}^\perp$ depends not only on the function Φ but also on the basis sets $\mathscr{B}_{\mathrm{occ}}$ and $\mathscr{B}_{\mathrm{vir}}$. The following result shows that the definition of the cluster operator $\mathcal{T}(\Phi)$ is in fact *independent* of the basis sets $\mathscr{B}_{\mathrm{occ}}$ and $\mathscr{B}_{\mathrm{vir}}$, and thus depends only on the choice of occupied space $\mathscr{R}$.

Proposition 15 (Basis-Independence of Cluster Operators)**.**

Let $\mathscr{R}$ *and* $\mathscr{R}^\perp$ *denote an occupied and virtual space of* $\mathrm{H}^1(\mathbb{R}^3)$ *as stated in Notation 1, let* $\Psi_0 \in \widehat{\mathscr{H}}^1$ *denote the reference determinant corresponding to* $\mathscr{R}$ *as defined through Definition 2, let* $\mathscr{B}^1_{\mathrm{occ}} := \{\psi^1_j\}_{j=1}^N$ *and* $\mathscr{B}^2_{\mathrm{occ}} := \{\psi^2_j\}_{j=1}^N$ *denote two different* $\mathrm{L}^2(\mathbb{R}^3)$*-orthonormal bases for the occupied space* $\mathscr{R}$*, let* $\mathscr{B}^1_{\mathrm{vir}} := \{\psi^1_j\}_{j=N+1}^\infty$ *and* $\mathscr{B}^2_{\mathrm{vir}} := \{\psi^2_j\}_{j=N+1}^\infty$ *denote two different* $\mathrm{L}^2(\mathbb{R}^3)$*-orthonormal bases for the virtual space* $\mathscr{R}^\perp$*, let* $\mathcal{B}^1_\wedge$ *and* $\mathcal{B}^2_\wedge$ *denote the* N*-particle Slater bases for* $\widehat{\mathscr{H}}^1$ *constructed using the sets*
mathscrB$^1_{\mathrm{occ}} \cup \mathscr{B}^1_{\mathrm{vir}}$ *and* $\mathscr{B}^2_{\mathrm{pcc}} \cup \mathscr{B}^2_{\mathrm{vir}}$ *respectively as defined through Definition 6, let the excitation index set* $\mathcal{J}$ *be defined through Definition 9, let the excitation operators* $\{\mathcal{X}^1_\mu\}_{\mu\in\mathcal{J}}$ *and* $\{\mathcal{X}^2_\mu\}_{\mu\in\mathcal{J}}$ *corresponding*

to the N-particle Slater bases $\mathcal{B}_\wedge^1$ and $\mathcal{B}_\wedge^2$ respectively be defined through Definition 10, and let $\Phi \in \{\Psi_0\}^\perp \subset \widehat{\mathcal{L}}^2$ be given by

$$\Phi = \sum_{\mu \in \mathcal{J}} \boldsymbol{t}_\mu^1 \mathcal{X}_\mu^1 \Psi_0 = \sum_{\mu \in \mathcal{J}} \boldsymbol{t}_\mu^2 \mathcal{X}_\mu^2 \Psi_0.$$

Then it holds that

$$\mathcal{T}(\Phi) = \sum_{\mu \in \mathcal{J}} \boldsymbol{t}_\mu^1 \mathcal{X}_\mu^1 = \sum_{\mu \in \mathcal{J}} \boldsymbol{t}_\mu^2 \mathcal{X}_\mu^2.$$

Proof. See [25, Theorem 2.7 (ii)] □

Proposition 15 implies that once the reference determinant Ψ_0 constructed from an occupied space $\mathscr{R}$ of $\mathrm{H}^1_{\mathrm{as}}(\mathbb{R}^3)$ is fixed, each element $\Phi \in \{\Psi_0\}^\perp \subset \widehat{\mathcal{L}}^2$ uniquely defines a bounded cluster operator $\mathcal{T}(\Phi) : \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ that does not depend on the specific bases $\mathscr{B}_{\mathrm{vir}}$ and $\mathscr{B}_{\mathrm{vir}}$ for the occupied space $\mathscr{R}$ and virtual space $\mathscr{R}^\perp$ respectively. Of course, in order to have an explicit representation of the cluster operator $\mathcal{T}(\Phi)$, we require a Slater basis $\mathcal{B}_\wedge$ on $\widehat{\mathcal{H}}^1$, which in turn requires a specific choice of bases $\mathscr{B}_{\mathrm{occ}}, \mathscr{B}_{\mathrm{vir}}$ for the occupied and virtual spaces but the above proposition shows that any basis ultimately leads to the same operator. This implies, in particular, that we can define cluster operators by referencing only the occupied space $\mathscr{R}$ and the reference determinant Ψ_0 generated by this occupied space. This observation will be crucial for the analysis we present in Section 4.

Returning to the mapping properties of cluster operators, we notice that our results thus far pertain to $\widehat{\mathcal{L}}^2$, i.e., the space of anti-symmetric, square-integrable functions of $3N$ variables. On the other hand, the N-particle function space for our problem is the Sobolev space $\widehat{\mathcal{H}}^1$. For elements $\Phi \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0} = \{\Psi_0\}^\perp \cap \widehat{\mathcal{H}}^1$, stronger results on the mapping properties of the cluster operator $\mathcal{T}(\Phi)$ can be deduced. The following result is the main technical achievement of the seminal contribution [25].

Theorem 16 (Cluster Operators as Bounded Maps on $\widehat{\mathcal{H}}^1$ and $\widehat{\mathcal{H}}^{-1}$)**.**

Let $\mathscr{R}$ denote an occupied space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\Psi_0 \in \widehat{\mathcal{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, let the N-particle function space $\widehat{\mathcal{H}}^1$ be decomposed as $\widehat{\mathcal{H}}^1 = \mathrm{span}\{\Psi_0\} \oplus \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ as described in Definition 4, let $\Phi \in \widehat{\mathcal{H}}^{1,\perp}$, and let $\mathcal{T}(\Phi) : \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ denote the cluster operator generated by Φ as stated in Proposition 14. Then

(1) The cluster operator $\mathcal{T}(\Phi) : \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ is a bounded linear map, and there exists a constant $\beta_{\mathcal{H}} > 0$, depending only on N and $\|\Psi_0\|_{\widehat{\mathcal{H}}^1}$, such that

$$\|\Phi\|_{\widehat{\mathcal{H}}^1} \leq \|\mathcal{T}(\Phi)\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1} \leq \beta_{\mathcal{H}} \|\Phi\|_{\widehat{\mathcal{H}}^1}.$$

(2) The $\widehat{\mathcal{L}}^2$-adjoint cluster operator $\mathcal{T}(\Phi)^\dagger : \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ is also a bounded linear map and there exists a constant $\beta^\dagger_{\mathcal{H}} > 0$, depending only on N and $\|\Psi_0\|_{\widehat{\mathcal{H}}^1}$, such that

$$\|\mathcal{T}(\Phi)^\dagger\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1} \leq \beta^\dagger_{\mathcal{H}} \|\Phi\|_{\widehat{\mathcal{H}}^1}.$$

(3) The cluster operator $\mathcal{T}(\Phi) : \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ can be extended to a bounded map from $\widehat{\mathcal{H}}^{-1}$ to $\widehat{\mathcal{H}}^{-1}$ and it holds that

$$\|\mathcal{T}(\Phi)\|_{\widehat{\mathcal{H}}^{-1} \to \widehat{\mathcal{H}}^{-1}} = \|\mathcal{T}(\Phi)^\dagger\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1}.$$

Proof. See [25, Theorem 4.1, Lemma 5.1]. □

Theorem 17 (Cluster Operators Form an Algebra)**.**

Let $\mathscr{R}$ denote an occupied space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\Psi_0 \in \widehat{\mathcal{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, let the N-particle function space $\widehat{\mathcal{H}}^1$ be decomposed as $\widehat{\mathcal{H}}^1 = \mathrm{span}\{\Psi_0\} \oplus \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ as described in Definition 4, and define the set of operators

$$\mathfrak{L} := \left\{ t_0 \mathrm{I} + \mathcal{T}(\Phi) : \quad t_0 \in \mathbb{R},\ \Phi \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0} \ \text{ and } \ \mathcal{T}(\Phi) \ \text{ is the cluster operator generated by } \Phi \right\}.$$

Then the following hold:

- *The set $\mathfrak{L}$ forms a closed commutative subalgebra in the algebra of bounded linear operators acting from $\widehat{\mathcal{H}}^1$ to $\widehat{\mathcal{H}}^1$ (and also from $\widehat{\mathcal{H}}^{-1}$ to $\widehat{\mathcal{H}}^{-1}$).*
- *The subalgebra $\mathfrak{L}$ is closed under inversion and the spectrum of any $\mathcal{L}(\Phi) = t_0\mathrm{I} + \mathcal{T}(\Phi) \in \mathfrak{L}$ is exactly $\sigma(\mathcal{L}) = \{t_0\}$.*
- *Any element in $\mathfrak{L}$ of the form $\mathcal{T}(\Phi) = \sum_{\mu\in\mathcal{G}} \boldsymbol{t}_\mu \mathcal{X}_\mu$ is nilpotent: it holds that $\mathcal{T}^{N+1} \equiv 0$.*
- *The exponential function is a locally $\mathscr{C}^\infty$ map on $\mathfrak{L}$, and is also a bijection from the sub-algebra*

  $$\left\{ \mathcal{T}(\Phi) : \qquad \Phi \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0} \ \text{ and } \ \mathcal{T}(\Phi) \ \text{ is the cluster operator generated by } \Phi \right\}.$$

  to the sub-algebra

  $$\left\{ \mathrm{I} + \mathcal{T}(\Phi) : \quad \Phi \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0} \ \text{ and } \ \mathcal{T}(\Phi) \ \text{ is the cluster operator generated by } \Phi \right\}.$$

Proof. See [25, Lemma 5.2]. □

3.3. **The Continuous Coupled Cluster Equations and their Analysis.**

Theorems 16 and 17 are the foundation stones of the continuous (infinite-dimensional) coupled cluster equations. Indeed, as a consequence of Theorems 16 and 17, it becomes possible to prove that for a given reference determinant Ψ_0, any intermediately normalised element of the N-particle function space $\widehat{\mathcal{H}}^1$, i.e., any $\Phi \in \widehat{\mathcal{H}}^1$ such that $(\Psi_0, \Phi)_{\widehat{\mathcal{L}}^2} = 1$, can be parameterised through the action of an exponential cluster operator acting on Ψ_0.

More precisely, let $\mathscr{R}$ and $\mathscr{R}^\perp$ denote an occupied and virtual space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\Psi_0 \in \widehat{\mathcal{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, and let the N-particle function space $\widehat{\mathcal{H}}^1$ be decomposed as $\widehat{\mathcal{H}}^1 = \mathrm{span}\{\Psi_0\} \oplus \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ as described in Definition 4. Then for any $\Phi \in \widehat{\mathcal{H}}^1$ such that $(\Phi, \Psi_0)_{\widehat{\mathcal{L}}^2} = 1$, there exists a cluster operator $\mathcal{T}(\Theta)$ generated by some $\Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ such that

$$\Phi = e^{\mathcal{T}(\Theta)} \Psi_0. \tag{10}$$

Equation (10) implies in particular that if the sought-after ground state wave-function $\Psi^*_{\mathrm{GS}} \in \widehat{\mathcal{H}}^1$ that solves the minimisation problem (4a) is intermediately normalised, then it can be written in the form

$$\Psi^*_{\mathrm{GS}} = e^{\mathcal{T}(\Theta^*)} \Psi_0,$$

for some cluster operator $\mathcal{T}(\Theta^*)$ generated by some element $\Theta^* \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$. In other words, the minimisation problem (4a) can be replaced by an equivalent problem which consists of finding the appropriate cluster operator $\mathcal{T}(\Theta^*)$ that appears in the exponential parametrisation of Ψ^*_{GS}. This observation leads directly to the so-called continuous coupled cluster equations.

Continuous Coupled Cluster Equations:

Let $\mathscr{R}$ denote an occupied space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\Psi_0 \in \widehat{\mathscr{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, and let the N-particle function space $\widehat{\mathscr{H}}^1$ be decomposed as $\widehat{\mathscr{H}}^1 = \mathrm{span}\{\Psi_0\} \oplus \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ as described in Definition 4. We seek a cluster operator $\mathcal{T}(\Theta^*)$ generated by some element $\Theta^* \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ according to Proposition 14 such that for all $\Phi \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ it holds that

$$\left\langle \Phi, e^{-\mathcal{T}(\Theta^*)} H e^{\mathcal{T}(\Theta^*)} \Psi_0 \right\rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}} = 0. \tag{11}$$

Once Equation (11) has been solved, the associated coupled cluster energy $\mathcal{E}^*_{\mathrm{CC}}$ is given by

$$\mathcal{E}^*_{\mathrm{CC}} := \left\langle \Psi_0, e^{-\mathcal{T}(\Theta^*)} H e^{\mathcal{T}(\Theta^*)} \Psi_0 \right\rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}}. \tag{12}$$

The continuous coupled cluster equations (11) are thus an infinite-dimensional system of non-linear equations in which the unknown is the cluster operator that appears in the exponential parametrisation of the sought-after ground-state eigenfunction Ψ^*_{GS} of the electronic Hamiltonian. In practice, these equations are approximated by considering suitable Galerkin discretisations, which shall be the subject of an extensive discussion in Section 4. For the remainder of the current section however, we restrict ourselves to a discussion of the continuous coupled cluster equations themselves.

From the perspective of numerical analysis, the natural question that arises is whether or not the continuous coupled cluster equations are well-posed. To discuss the answer to this question, it is useful to first introduce the infinite-dimensional coupled cluster function.

Definition 18 (Continuous Coupled Cluster function)**.**

Let $\mathscr{R}$ denote an occupied space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\Psi_0 \in \widehat{\mathscr{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, let the N-particle function space $\widehat{\mathscr{H}}^1$ be decomposed as $\widehat{\mathscr{H}}^1 = \mathrm{span}\{\Psi_0\} \oplus \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ as described in Definition 4, and for any $\Theta \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$, let $\mathcal{T}(\Theta)$ denote the cluster operator generated by Θ as described in Proposition 14. We define the continuous coupled cluster function $f\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\right)^*$ as the mapping with the property that for all $\Theta, \Phi \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ it holds that

$$\left\langle \Phi, f(\Theta) \right\rangle_{\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} \times \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\right)^*} := \left\langle \Phi, e^{-\mathcal{T}(\Theta)} H e^{\mathcal{T}(\Theta)} \Psi_0 \right\rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}}.$$

Consider the continuous coupled cluster equations (11) and Definition 18 of the continuous coupled cluster function. It can readily be checked that if a cluster operator $\mathcal{T}(\Theta^*)$ generated by some element $\Theta^* \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ is a solution to the CC equations (11), then Θ^* is a zero of the coupled cluster function $f\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\right)^*$. The converse is also true of course, and this justifies the decision to directly study the zeros of the coupled cluster function f using the usual tools of non-linear numerical analysis.

It is now interesting to note that every intermediately normalised eigenfunction $\Psi^* \in \widehat{\mathscr{H}}^1$ of the electronic Hamiltonian generates a zero of the coupled cluster function, i.e., for any such Ψ^*, there exists a $\Phi^* \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ such that $\Psi^* = e^{\mathcal{T}(\Phi^*)}$ and Φ^* is a zero of the coupled cluster function. Even more remarkable is that the converse is true, i.e., every zero of the coupled cluster function gives rise to an intermediately normalised eigenfunction of the electronic Hamiltonian. More precisely, if $\Phi^* \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ is a zero of f, then $\Psi^* = e^{\mathcal{T}(\Phi^*)}$ is an intermediately normalised eigenfunction of the electronic Hamiltonian (see [25, Theorem 5.3]).

The upshot of these observations is that we have a complete classification of the zeros of the coupled cluster function f, namely, that they correspond precisely to intermediately normalised eigenfunctions of the electronic Hamiltonian, and the remaining task is to determine conditions under which these zeros are non-degenerate.

A sufficient condition for non-degeneracy was first given in the articles [25, 26, 28] wherein the authors proposed the use of the following *local, strong monotonicity criterion*: let $\Theta^*_{\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ denote the zero of the coupled cluster function corresponding to the intermediately normalised ground-state eigenfunction Ψ^*_{GS} of the electronic Hamiltonian $H\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$, i.e., let $\Psi^*_{\mathrm{GS}} = e^{\mathcal{T}(\Theta^*_{\mathrm{GS}})}$, and suppose there exist two constants $\delta > 0$ and $\gamma > 0$ such that

$$\forall\, \Phi, \Upsilon \in \mathbb{B}_\delta(\Theta^*_{\mathrm{GS}}) \cap \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\colon \quad \big\langle \Upsilon - \Phi, f(\Upsilon) - f(\Phi) \big\rangle_{\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} \times \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\right)^*} \geq \gamma \|\Upsilon - \Phi\|^2_{\widehat{\mathscr{H}}^1}. \tag{13}$$

Then Θ^*_{GS} is a non-degenerate zero of the coupled cluster function (see [26, Theorem 4.1]).

An important advantage of this approach is that the strong local monotonicity property is directly inherited by *Galerkin discretisations* of the continuous coupled cluster equations, i.e., by the discrete coupled cluster equations that are solved in practice. Consequently, there is a straightforward path to a priori error estimates for practical discretisations of the continuous coupled cluster equations (see [26, Theorem 4.1]).

On the other hand, an important shortcoming of the local monotonicity approach is that the theoretical local monotonicity constant that one can actually obtain (see [26, Theorem 3.4] or [28, Theorem 5.7]) is of the form

$$\begin{aligned} \gamma \approx \Lambda_0 - \|H - \mathcal{E}^*_{\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}} \|\mathcal{T}(\Theta^*_{\mathrm{GS}}) - \mathcal{T}(\Theta^*_{\mathrm{GS}})^\dagger\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \\ - \mathcal{O}\left(\|\mathcal{T}(\Theta^*_{\mathrm{GS}})\|^2_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} + \|\mathcal{T}(\Theta^*_{\mathrm{GS}})^\dagger\|^2_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \right), \end{aligned} \tag{14}$$

where $\Lambda_0 > 0$ is the coercivity constant of the shifted electronic Hamiltonian $H - \mathcal{E}^*_{\mathrm{GS}}$ on $\{\Psi^*_{\mathrm{GS}}\}^\perp \cap \widehat{\mathscr{H}}^1$, and this theoretical constant is often negative, even for well-behaved molecular systems (see, for instance, Table 1 below).

In order to improve upon this pessimistic state of affairs, the present authors proposed a new analysis based on establishing the invertibility of the Fréchet derivative of the coupled cluster function using inf-sup arguments. This analysis was the subject of our previous contribution [11], where we were able to establish, in particular, the following result.

Theorem 19 (Invertibility of the Coupled Cluster Fréchet Derivative)**.**

Let the coupled cluster function $f\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\right)^*$ be defined through Definition 18 for some choice of occupied space $\mathscr{R}$, let $\mathrm{D}f(\Theta)\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\right)^*$ denote the Fréchet derivative of f at any $\Theta \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$, and let $\Theta^* \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}$ denote a zero of the coupled cluster function corresponding to an intermediately normalised eigenfunction $\Psi^* \in \widehat{\mathscr{H}}^1$ of the electronic Hamiltonian $H\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$ with a simple, isolated eigenvalue $\mathcal{E}^*$. Then, the Fréchet derivative $\mathrm{D}f(\Theta^*)$ is an isomorphism and it holds that

$$\|\mathrm{D}f(\Theta^*)^{-1}\|_{\left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}\right)^* \to \widehat{\mathscr{H}}^{1,\perp}_{\Psi_0}} \leq \frac{\beta}{\Lambda^*},$$

where $\Lambda^* > 0$ is the inf-sup constant of the shifted electronic Hamiltonian $H - \mathcal{E}^*$ on $\{\Psi^*\}^\perp \cap \widehat{\mathscr{H}}^1$ and $\beta := \|\mathbb{P}_0^\perp e^{-\mathcal{T}(\Theta^*)}\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \|e^{\mathcal{T}(\Theta^*)^\dagger}\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}$. In particular, the zero Θ^* of the coupled cluster function f is non-degenerate.

Proof. See [11, Theorems 31 and 33, Corollary 32]. □

Molecule	Monotonicity constant Γ from Eq. (14)	$\|\mathrm{D}f(\Theta^*)^{-1}\|^{-1}_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}\right)^* \to \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}}$	Inf-Sup constant Λ^*/β
BeH_2	0.0363	0.3379	0.2568
BH_3	-0.0950	0.3060	0.2081
HF	-0.0083	0.2995	0.2529
H_2O	0.0249	0.4113	0.2784
LiH	-0.0065	0.2628	0.2164
NH_3	-0.0325	0.3576	0.2789

TABLE 1. Examples of numerically computed constants for a collection of small molecules at equilibrium geometries. The calculations were performed in STO-6G basis sets with the exception of the HF and LiH molecules for which 6-31G basis sets were used. To simplify calculations, the canonical $\widehat{\mathcal{H}}^1$ norm was replaced with an equivalent norm induced by the mean-field Hartree-Fock operator (see, e.g., [28]). The numerical results shown here are taken from the previous contribution [11].

In comparison to the local monotonicity approach, Theorem 19 provides a sharper estimate of the operator norm of the inverse coupled cluster Fréchet derivative (see also Table 1). Unfortunately, as is well known from the classical numerical analysis theory for linear PDEs, continuous inf-sup conditions are not, in general, inherited by Galerkin discretisations of the infinite-dimensional problem. The natural question to ask therefore, is if an analysis in the spirit of Theorem 19 can be carried out for discretisations of the continuous coupled cluster equations (11). This shall be the main subject of the remainder of this article.

3.4. **The Discrete Coupled Cluster Equations.**

We begin with the generic form of the discrete coupled cluster equations.

General form of the Discrete Coupled Cluster Equations:

Let $\mathscr{R}$ denote an occupied and virtual space of $\mathrm{H}^1(\mathbb{R}^3)$ as stated in Notation 1, let $\Psi_0 \in \widehat{\mathcal{H}}^1$ denote the reference determinant corresponding to $\mathscr{R}$ as defined through Definition 2, let the N-particle function space $\widehat{\mathcal{H}}^1$ be decomposed as $\widehat{\mathcal{H}}^1 = \mathrm{span}\{\Psi_0\} \oplus \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ as described in Definition 4, and let $\widetilde{\mathcal{V}}_K \subset \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ denote a finite-dimensional approximation space. We seek a cluster operator $\mathcal{T}(\Theta^*_K)$ generated by some element $\Theta^*_K \in \widetilde{\mathcal{V}}_K$ such that for all $\Phi_K \in \widetilde{\mathcal{V}}_K$ it holds that

$$\left\langle \Phi_K, e^{-\mathcal{T}(\Theta^*_K)} H e^{\mathcal{T}(\Theta^*_K)} \Psi_0 \right\rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}} = 0. \tag{15}$$

Once the discrete coupled cluster equations (11) have been solved, the associated coupled cluster energy $\mathcal{E}^*_K$ is given by

$$\mathcal{E}^*_K := \left\langle \Psi_0, e^{-\mathcal{T}(\Theta^*_K)} H e^{\mathcal{T}(\Theta^*_K)} \Psi_0 \right\rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}}. \tag{16}$$

Remark 20 (Solving the Discrete Coupled Cluster equations)**.**

Consider the discrete coupled cluster equations (15)*. The practical resolution of these equations is done by first introducing a Slater basis* $\widetilde{\mathcal{B}_{\wedge,h}}$ *for the finite-dimensional approximation space* $\widetilde{\mathcal{V}}_h$*. Using the*

notion of excitation index sets and excitation operators introduced in Section 3.1, we can then obtain a parametrisation of these basis elements, i.e.,

$$\exists \mathcal{I}_h \subset \mathcal{I}\colon \quad \widetilde{\mathcal{B}_{\wedge,h}} = \{\mathcal{X}_\mu \Psi_0 \colon \mu \in \mathcal{I}_h\}.$$

It then follows that the sought-after cluster operator $\mathcal{T}(\Theta_h^*)$ which is generated by $\Theta_h^* \in \widetilde{\mathcal{V}}_h$ can be expressed in the form

$$\mathcal{T}(\Theta_h^*) = \sum_{\mu \in \mathcal{I}_h} \boldsymbol{t}_\mu^h \mathcal{X}_\mu,$$

for the unknown expansion coefficients $\{\boldsymbol{t}_\mu^h\}_{\mu\in\mathcal{I}_h}$ of $\Theta_h^* \in \widetilde{\mathcal{V}}_h$.

Solving the discrete coupled cluster equations (15) *then consists of determining the unknown coefficients $\{\boldsymbol{t}_\mu^h\}_{\mu\in\mathcal{I}_h}$ such that for all $\mu \in \mathcal{I}_h$ it holds that*

$$\left\langle \mathcal{X}_\mu \Psi_0, e^{-\mathcal{T}(\Theta_h^*)} H e^{\mathcal{T}(\Theta_h^*)} \Psi_0 \right\rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}} = 0, \qquad \text{where } \mathcal{T}(\Theta_h^*) = \sum_{\mu \in \mathcal{I}} \boldsymbol{t}_\mu^h \mathcal{X}_\mu. \tag{17}$$

Note that Equation (17) *is a finite-dimensional system of coupled, non-linear equations which can be readily solved using iterative methods such as a quasi-Newton scheme.*

Consider the discrete coupled cluster equations (15). The approximation space $\widetilde{\mathcal{V}}_h \subset \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ is typically chosen to be the span of *some finite collection* of Slater determinants. We recall that a Slater determinant is simply an anti-symmetric tensor product of some L^2-orthonormal functions in $\mathrm{H}^1(\mathbb{R}^3)$. Broadly speaking therefore, there are two discretisation parameters that can be tuned to produce different discrete coupled cluster equations.

- We can modify the pool of L^2-orthonormal functions in $\mathrm{H}^1(\mathbb{R}^3)$ that are used to construct the Slater determinants in $\widetilde{\mathcal{V}}_h$. Typically, one starts with a collection of not necessarily orthogonal functions in $\mathrm{H}^1(\mathbb{R}^3)$– known as an atomic basis in the quantum chemistry literature– and then produces an orthonormal basis for the set, which is used to construct Slater determinants.
- We can modify the number or types of Slater determinants that appear in the approximation space $\widetilde{\mathcal{V}}_h$. In the quantum chemistry literature for instance, Slater determinants are characterised according to their excitation rank. The excitation rank of a Slater determinant $\Psi_\mu \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0}$ is the natural number $j \in \{1, \ldots, N\}$ such that the excitation index $\mu \in \mathcal{I}_j$. With this classification, we can, for example, include only singly and doubly excited Slater determinants in $\widetilde{\mathcal{V}}_h$ (this is known as CCSD) or, at the other extreme, include Slater determinants of all possible excitation ranks (this is known as Full-CC).

A graphical description of these discretisation parameters and the resulting discrete coupled cluster equations is given in Figure 1. Note that the continuous coupled cluster equations correspond to the case of taking an L^2-orthonormal complete basis of $\mathrm{H}^1(\mathbb{R}^3)$ as the atomic basis and including Slater determinants of all possible excitation ranks in the approximation space.

The task that we now confront is to develop a local well-posedness theory and a priori error estimates for the discrete coupled cluster equations (15). As is common in the non-linear numerical analysis literature, we will attempt to show that the discrete coupled cluster equations (15) are locally well-posed provided that the associated approximation space $\widetilde{\mathcal{V}}_h$ is rich enough, which, in the context of the two discretisation parameters introduced above means that both the size of the atomic basis and the maximal excitation rank of admissible Slater determinants is sufficiently large. Note that since the coupled cluster methodology is targeted at obtaining the *ground state* energy of the electronic

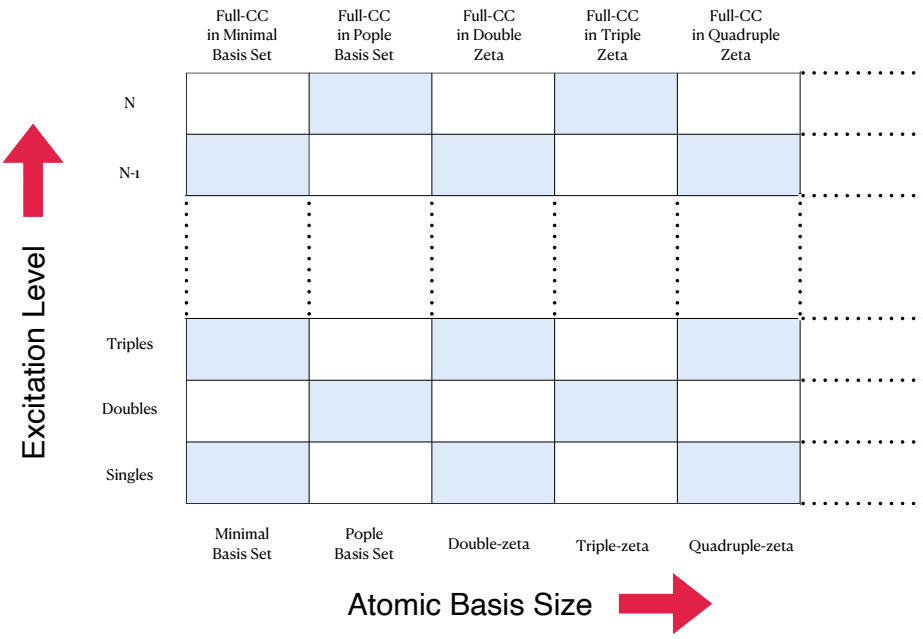

FIGURE 1. Graphical depiction of the coupled cluster discretisation parameters. Informally, the continuous coupled cluster equations should appear at the upper right corner of this graphic.

Hamiltonian, we will focus only on solutions of the discrete coupled cluster equations corresponding to a ground state energy approximation.

An asymptotic well-posedness analysis of this nature presents two principle difficulties:

Difficulty One: Due to the non-linearity of the coupled cluster function, the primary tool that we have at our disposal for the analysis of the discrete problem (15) is the inverse function theorem for Banach spaces. As we show in the subsequent Section 4, in order to apply the inverse function theorem in the current setting, we must establish a **discrete** inf-sup condition for the Fréchet derivative of the coupled cluster function on the approximation space $\widetilde{\mathcal{V}}_h$. At first glance, we can simply attempt to replicate the proof for the continuous inf-sup condition (see [11]). Unfortunately however, for an arbitrary test function $\Phi_h \in \widetilde{\mathcal{V}}_h$ and a general cluster operator $\mathcal{T}(\Theta_h)$ constructed from some element $\Theta_h \in \widetilde{\mathcal{V}}_h$, we have that

$$e^{\mathcal{T}(\Theta_h)}\Phi_h \notin \widetilde{\mathcal{V}}_h,$$

and this non-inclusion causes the continuous inf-sup argument to break down. Moreover, the 'extent' of this non-inclusion is only decreased if Slater determinants of all possible excitation ranks are included in $\widetilde{\mathcal{V}}_h$.

Difficulty Two: In the numerical practice, the reference determinant $\Psi_0 \in \widehat{\mathcal{H}}^1$ which is at the heart of the mathematical formulation of the coupled cluster equations is chosen as the anti-symmetric tensor product of the first N eigenfunctions of some discrete mean-field operator (so-called canonical orbitals). In other words, given an atomic basis, one first performs a mean-field Hartree-Fock (or Kohn-Sham) calculation *in this atomic basis*, and then uses the resulting eigenfunctions to construct the reference determinant. This means however, that as the size of the atomic basis changes, the results of the mean-field calculation and therefore the underlying

reference determinant also change. In particular, since cluster operators are defined relative to a fixed reference determinant, it is not possible to compare two different cluster operators defined with respect to two different reference determinants.

Partial solutions to **Difficulty One** shall be the subject of extensive discussion in the forthcoming Section 4. In the remainder of the current section, we introduce a formalism which, under some assumptions, will allow us to address **Difficulty Two**.

Notation 21 (Sequence of Occupied Spaces and Reference Determinants)**.**

We denote by $\{\mathscr{R}_K\}_{K\geq N}$ *a sequence of occupied spaces of* $\mathrm{H}^1(\mathbb{R}^3)$ *as stated in Notation 1, and for each* $K \geq N$*, we denote by* $\Psi_{0,K} \in \widehat{\mathscr{H}}^1$ *the reference determinant corresponding to* $\mathscr{R}_K$ *as defined through Definition 2.*

Notation 22 (Sequence of Finite-Dimensional Approximation Spaces)**.**

Let the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ *be constructed as in Notation 21 and for each* $K \geq N$*, let* $\{\Psi_{0,K}\}^\perp$ *be the* $\widehat{\mathcal{L}}^2$ *orthogonal complement of* $\Psi_{0,K}$ *in* $\widehat{\mathcal{L}}^2$*. We denote by* $\{\widetilde{\mathcal{V}}_K\}_{K\geq N}$ *a sequence of finite-dimensional subspaces of* $\widehat{\mathscr{H}}^1$ *with the following properties:*

$$\forall K \geq N\colon \quad \widetilde{\mathcal{V}}_K \subset \{\Psi_{K,0}\}^\perp, \qquad \widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\} \subset \widetilde{\mathcal{V}}_{K+1} \oplus \mathrm{span}\{\Psi_{0,K+1}\}, \qquad \textit{and}$$

$$\overline{\bigcup_{K\geq N} \widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}}^{\|\cdot\|_{\widehat{\mathscr{H}}^1}} = \widehat{\mathscr{H}}^1.$$

Additionally, for notational convenience, we define $\mathcal{V}_K := \widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}$ *for all* $K \geq N$*.*

Consider Notations 21 and 22. A valid question that arises is why we define the sequence of reference determinants and approximation spaces only for indices $K \geq N$. This choice is, in fact, motivated by the observation that in the numerical practice, the approximation spaces $\{\mathcal{V}_K\}_{K\geq N}$ are typically defined as the span of some Slater determinants constructed from $K \geq N$ single-particle functions in $\mathrm{H}^1(\mathbb{R}^3)$. Indeed, such constructions are the subject of detailed discussions in Appendix B. To ensure a uniform notation therefore, we define the sequence of reference determinants and approximation spaces only for indices $K \geq N$. Of course, this choice has no bearing on the analysis.

Equipped with Notations (21) and (22), we can now state a *sequence* of discrete coupled cluster equations whose asymptotic well-posedness will be the main object of study in the sequel.

Sequence of Discrete Coupled Cluster Equations:

Let the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ be constructed as in Notation 21 and let the sequence of finite-dimensional subspaces $\{\widetilde{\mathcal{V}}_K\}_{K\geq N}$ be constructed as in Notation 22. For each $K \geq N$, we seek a cluster operator $\mathcal{T}(\Theta_K^*)$ generated by some element $\Theta_K^* \in \widetilde{\mathcal{V}}_K$ such that

$$\forall \Phi_K \in \widetilde{\mathcal{V}}_K\colon \qquad \left\langle \Phi_K, e^{-\mathcal{T}(\Theta_K^*)} H e^{\mathcal{T}(\Theta_K^*)} \Psi_{0,K} \right\rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}} = 0. \tag{18}$$

4. Analysis of the Discrete Coupled Cluster Equations

Throughout this section, we assume the settings of Sections 2-3. As mentioned at the end of Section 3, we are now primarily concerned with the well-posedness and error analysis of the sequence of discrete coupled cluster equations (18). To do so, we will make use of the classical machinery of non-linear numerical analysis, and, in particular, the abstract framework for analysing such problems introduced by Caloz and Rappaz [3]. Of particular importance is the following fundamental result which is essentially the inverse function theorem in Banach spaces, adapted for the purpose of non-linear numerical analysis.

Theorem 23 (Inverse Function Theorem in Banach Spaces)**.**

Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ be two Banach spaces, let $\mathcal{G}\colon \mathcal{X} \to \mathcal{Z}$ be a $\mathscr{C}^1$ mapping, let $\mathbf{u} \in \mathcal{X}$ be such that the Fréchet derivative $\mathrm{D}\mathcal{G}(\mathbf{u})\colon \mathcal{X} \to \mathcal{Z}$ is an isomorphism and define the following quantities:

$$\begin{aligned} \epsilon :=& \|\mathcal{G}(\mathbf{u})\|_{\mathcal{Z}}, \\ \gamma_{\mathcal{G}} :=& \|\mathrm{D}\mathcal{G}(\mathbf{u})^{-1}\|_{\mathcal{Z}\to\mathcal{X}}, \\ \forall \alpha \in \mathbb{R}_+\colon\ \mathrm{L}(\alpha) :=& \sup_{\mathbf{v}\in\overline{\mathbb{B}_{\alpha}(\mathbf{u})}} \|\mathrm{D}\mathcal{G}(\mathbf{u}) - \mathrm{D}\mathcal{G}(\mathbf{v})\|_{\mathcal{X}\to\mathcal{Z}}. \end{aligned}$$

Under the assumption that $2\gamma_{\mathcal{G}}\mathrm{L}(2\gamma_{\mathcal{G}}\epsilon) < 1$, the closed ball $\overline{\mathbb{B}_{2\gamma_{\mathcal{G}}\epsilon}(\mathbf{u})} \subset \mathcal{X}$ contains a unique solution $\mathbf{u}^*$ to the equation

$$\mathcal{G}(\mathbf{v}) = 0,$$

the Fréchet derivative $\mathrm{D}\mathcal{G}(\mathbf{u}^*)\colon \mathcal{X} \to \mathcal{Z}$ is an isomorphism with

$$\|\mathrm{D}\mathcal{G}(\mathbf{u}^*)^{-1}\|_{\mathcal{Z}\to\mathcal{X}} \leq 2\gamma_{\mathcal{G}},$$

and for all $\mathbf{v} \in \overline{\mathbb{B}_{2\gamma_{\mathcal{G}}\epsilon}(\mathbf{u})}$ we have the error estimate

$$\|\mathbf{u}^* - \mathbf{v}\|_{\mathcal{X}} \leq 2\gamma_{\mathcal{G}}\|\mathcal{G}(\mathbf{v})\|_{\mathcal{Z}}.$$

Proof. See [3, Theorem 2.1]. □

The main focus of our subsequent analysis will be to demonstrate how, under suitable assumptions, the sequence of discrete coupled cluster equations (18) satisfies the hypotheses of Theorem 23. For clarity of exposition, we will proceed in the following steps:

(1) First, in Section 4.1, we will prove certain technical lemmas which are related to the *consistency* of the discrete coupled cluster equations (18). This step will require imposing certain assumptions on the sequence of reference determinants $\{\Psi_{0,K}\}_{K\in\mathbb{N}}$ constructed as in Notation 21.

(2) Second, in Section 4.2, we will identify certain classes of approximation spaces $\{\widetilde{\mathrm{V}}_K\}_{K\geq N}$ for which the stability of the discrete coupled cluster equations (18) can be expected to hold, at least asymptotically.

(3) In Section 4.3, we will combine the results of the previous two subsections to complete the sought-after demonstration.

4.1. **Technical Lemmas Pertaining to Consistency.**

For the purpose of the analysis in this section, we require the following assumptions.

Assumption A.I: Uniform Overlap. We assume that the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ constructed as in Notation 21 have uniformly bounded below overlaps with the exact ground state eigenfunction $\Psi^*_{\mathrm{GS}} \in \widehat{\mathscr{H}}^1$ of the electronic Hamiltonian H, i.e.,

$$\exists C > 0 \quad \text{s.t.} \quad \forall K \geq N\colon \qquad |\,(\Psi_{0,K}, \Psi^*_{\mathrm{GS}})\,| > C.$$

Assumption A.II: Convergence. We assume that the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ constructed as in Notation 21 have a well-defined limit $\Psi^*_0 \in \widehat{\mathscr{H}}^1$, i.e.,

$$\exists \Psi^*_0 \in \widehat{\mathscr{H}}^1\colon \lim_{K\to\infty} \Psi_{0,K} = \Psi^*_0 \quad \text{in the } \|\cdot\|_{\widehat{\mathscr{H}}^1} \text{ norm.}$$

We will refer to Ψ^*_0 as the *limiting reference determinant.*

In addition to the above **Assumptions A.I** and **A.II**, we of course also require that the continuous coupled cluster equations (11), stated using the limiting reference determinant $\Psi^*_0 \in \widehat{\mathscr{H}}^1$, be locally well-posed. More precisely, given the continuous coupled cluster equations (11) formulated with respect to the limiting reference determinant Ψ^*_0, we require that the solution cluster operator $\mathcal{T}^*_{\mathrm{GS}}$ corresponding to the ground state eigenfunction $\Psi^*_{\mathrm{GS}} \in \widehat{\mathscr{H}}^1$ of the electronic Hamiltonian be locally unique. In view of Theorem 19, a sufficient condition to achieve this local uniqueness is the following.

Assumption A.III: Existence of a Spectral Gap. We assume that the ground state eigenvalue $\mathcal{E}^*_{\mathrm{GS}}$ of the electronic Hamiltonian is simple.

The core idea of these assumptions is that they allow us to interpret the sequence of discrete coupled cluster equations (18) as non-conforming Galerkin discretisations of the continuous coupled cluster equations (11) constructed with respect to the limiting reference determinant Ψ^*_0, and this will allow us to deduce the consistency of the discrete equations (18).

Equipped with **Assumptions A.I-A.III**, we now prove two important lemmas.

Lemma 24 (Convergence of Zeros of Continuous Coupled Cluster Function)**.**

*Let the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ be constructed as in Notation 21 and assume that **Assumptions A.I-A.III** hold.*

Additionally, for each $K \geq N$, let $\{\Psi_{0,K}\}^\perp$ be the $\widehat{\mathscr{L}}^2$ orthogonal complement of $\Psi_{0,K}$ in $\widehat{\mathscr{L}}^2$, let the infinite-dimensional subspace $\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ be defined as $\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} := \{\Psi_{0,K}\}^\perp \cap \widehat{\mathscr{H}}^1$, let the coupled cluster function $f_K\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ be defined through Definition 18, where we adopt the convention of using the subscript K to signify the dependency of this coupled cluster function on $\Psi_{0,K}$, and let $\Theta^*_{K,\mathrm{GS}}$ denote the zero of the coupled cluster function f_K corresponding to the $\widehat{\mathscr{L}}^2$-normalised ground state eigenfunction Ψ^*_{GS} of the electronic Hamiltonian, i.e.,

$$e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K} = \frac{1}{(\Psi^*_{\mathrm{GS}}, \Psi_{0,K})_{\widehat{\mathscr{L}}^2}}\Psi^*_{\mathrm{GS}}.$$

*Finally, recalling that $\Psi_0^* \in \widehat{\mathcal{H}}^1$ denotes the limiting reference determinant defined in **Assumption A.II**, let $\{\Psi_0^*\}^\perp$ be the $\widehat{\mathcal{L}}^2$ orthogonal complement of Ψ_0^* in $\widehat{\mathcal{L}}^2$, let the infinite-dimensional subspace $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*}$ be defined as $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*} := \{\Psi_0^*\}^\perp \cap \widehat{\mathcal{H}}^1$, let the coupled cluster function $f\colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*}\right)^*$ be defined through Definition 18, and let Θ^*_{GS} denote the zero of the coupled cluster function f corresponding to the ground state eigenfunction of the electronic Hamiltonian.*

Then it holds that

$$\lim_{K\to\infty} \|\Theta^*_{K,\mathrm{GS}} - \Theta^*_{\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} = 0.$$

Proof. Since the proof is rather technical and long, we divide it into several steps.

Step One:

Given $K \geq N$, let $\mathcal{T}(\Theta^*_{K,\mathrm{GS}})$ and $\mathcal{T}(\Theta^*_{\mathrm{GS}})$ denote the cluster operators constructed from $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ and $\Theta^*_{\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*}$ respectively. It follows from the definition of these cluster operators that

$$\mathcal{T}(\Theta^*_{K,\mathrm{GS}})\Psi_{0,K} = \Theta^*_{K,\mathrm{GS}} \quad \text{and} \quad \mathcal{T}(\Theta^*_{\mathrm{GS}})\Psi_0^* = \Theta^*_{\mathrm{GS}}.$$

As a consequence, for any $K \geq N$ it holds that

$$\begin{aligned}
\|\Theta^*_{K,\mathrm{GS}} - \Theta^*_{\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} &= \|\mathcal{T}(\Theta^*_{K,\mathrm{GS}})\Psi_{0,K} - \mathcal{T}(\Theta^*_{\mathrm{GS}})\Psi_0^*\|_{\widehat{\mathcal{H}}^1} \\
&\leq \|\mathcal{T}(\Theta^*_{K,\mathrm{GS}})\Psi_{0,K} - \mathcal{T}(\Theta^*_{\mathrm{GS}})\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} + \|\mathcal{T}(\Theta^*_{\mathrm{GS}})\Psi_{0,K} - \mathcal{T}(\Theta^*_{\mathrm{GS}})\Psi_0^*\|_{\widehat{\mathcal{H}}^1} \\
&\leq \|\mathcal{T}(\Theta^*_{K,\mathrm{GS}}) - \mathcal{T}(\Theta^*_{\mathrm{GS}})\|_{\widehat{\mathcal{H}}^1\to\widehat{\mathcal{H}}^1}\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} + \|\mathcal{T}(\Theta^*_{\mathrm{GS}})\|_{\widehat{\mathcal{H}}^1\to\widehat{\mathcal{H}}^1}\|\Psi_{0,K} - \Psi_0^*\|_{\widehat{\mathcal{H}}^1}.
\end{aligned}$$

Thanks to **Assumption A.II**, we see that the second term on the right hand side of the inequality approaches zero in the limit $K \to \infty$. Additionally, **Assumption A.II** also implies that the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ is uniformly bounded in $\widehat{\mathcal{H}}^1$, and therefore it suffices to prove that

$$\lim_{K\to\infty} \|\mathcal{T}(\Theta^*_{K,\mathrm{GS}}) - \mathcal{T}(\Theta^*_{\mathrm{GS}})\|_{\widehat{\mathcal{H}}^1\to\widehat{\mathcal{H}}^1} = 0. \tag{19}$$

To this end, recall that Ψ^*_{GS} denotes the $\widehat{\mathcal{L}}^2$-normalised ground state eigenfunction of the electronic Hamiltonian. For each $K \geq N$, we define the constant

$$c_{0,K} := \frac{1}{(\Psi_{0,K}, \Psi^*_{\mathrm{GS}})_{\widehat{\mathcal{L}}^2}} < \infty,$$

where the upper bound holds thanks to **Assumption A.I**. It follows that for every $K \geq N$

$$\widetilde{\Psi}^*_{K,\mathrm{GS}} := c_{0,K}\Psi^*_{\mathrm{GS}} - \Psi_{0,K} \ \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}.$$

Next, for every $K \geq N$, we denote by $\mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}})\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ the cluster operator generated by $\widetilde{\Psi}^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$. It follows that

$$\left(\mathrm{I} + \mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}})\right)\Psi_{0,K} = c_{0,K}\Psi^*_{\mathrm{GS}} = e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K}.$$

Using now Theorem 16 on the mapping properties of cluster operators together with the fact that the cluster operators $\mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}})$ and $\mathcal{T}(\Theta^*_{K,\mathrm{GS}})$ are both constructed from functions in $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ with respect to the same reference determinant, we conclude that for every $K \geq N$ it holds that

$$\|\mathrm{I} + \mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}}) - e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\|_{\widehat{\mathcal{H}}^1\to\widehat{\mathcal{H}}^1} \leq \beta_{\widehat{\mathcal{H}}}\left\|\left(\mathrm{I} + \mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}}) - e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\right)\Psi_{0,K}\right\|_{\widehat{\mathcal{H}}^1} = 0,$$

and therefore,

$$\mathrm{I} + \mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}}) = e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}.$$

Recalling from Theorem 17 the bijection property of the exponential mapping between certain subalgebras of cluster operators, a direct calculation involving the Taylor series expansion of the logarithm function together with the nilpotency property of cluster operators allows us to conclude that for all $K \geq N$, it holds that

$$\mathcal{T}(\Theta^*_{K,\mathrm{GS}}) = \sum_{j=1}^{N} (-1)^{j+1} \frac{\mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}})^j}{j}. \tag{20}$$

An identical line of reasoning can be carried out for the cluster operator $\mathcal{T}(\Theta^*_{\mathrm{GS}})$ to obtain

$$\mathcal{T}(\Theta^*_{\mathrm{GS}}) = \sum_{j=1}^{N} (-1)^{j+1} \frac{\mathcal{C}(\widetilde{\Psi}^*_{\mathrm{GS}})^j}{j}, \tag{21}$$

where $\mathcal{C}(\widetilde{\Psi}^*_{\mathrm{GS}})$ is the cluster operator constructed from $\widetilde{\Psi}^*_{\mathrm{GS}} := c_0 \Psi^*_{\mathrm{GS}} - \Psi^*_0 \ \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi^*_0}$ using the limiting reference determinant Ψ^*_0, and the constant c_0 is defined as $c_0^{-1} := (\Psi^*_0, \Psi^*_{\mathrm{GS}})_{\widehat{\mathcal{L}}^2} \neq 0$.

In view of Equations (20) and (21), it suffices to prove that

$$\lim_{K\to\infty} \|\mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}}) - \mathcal{C}(\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} = 0. \tag{22}$$

Note that while the estimate (22) may appear superficially similar to the prior estimate (19), they are, in fact, very different. The key observation is that the functions $\widetilde{\Psi}^*_{K,\mathrm{GS}}$ and $\widetilde{\Psi}^*_{\mathrm{GS}}$ that appear in Estimate (22) have the property that

$$\lim_{K\to\infty} \|\widetilde{\Psi}^*_{K,\mathrm{GS}} - \widetilde{\Psi}^*_{\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1} = 0. \tag{23}$$

Indeed, this is simply a consequence of **Assumption A.II**. On the other hand, a similar result for the functions $\Theta^*_{K,\mathrm{GS}}$ and Θ^*_{GS} that appear in the prior estimate (19) has not yet been established, and is, in fact, exactly what we wish to show.

Step Two:

Unfortunately, a demonstration of the estimate (22) is far from trivial since the cluster operators $\mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}})$ and $\mathcal{C}(\widetilde{\Psi}^*_{\mathrm{GS}})$ are defined with respect to *different* reference determinants. Thus, the explicit definition of these operators (in the spirit of Proposition 14) cannot directly be compared. An additional difficulty arises from the fact that the infinite series representation of cluster operators is absolutely convergent only with respect to the $\widehat{\mathcal{L}}^2$-operator norm (see Proposition 14) and *not* with respect to the $\widehat{\mathscr{H}}^1$-operator norm. Thus, a term-by-term analysis of the partial sums of these series representations does not yield the required estimate (22).

To tackle these difficulties, we will, as a first step split the operator norm appearing in Equation (22) into two terms, each of which will be handled individually. We begin by defining for each $K \geq N$, the two-dimensional subspace $\mathrm{W}_K \subset \widehat{\mathscr{H}}^1$ as

$$\mathrm{W}_K := \mathrm{span}\{\Psi_{0,K}, \Psi^*_0\}.$$

For each $K \geq N$, we also denote by $\mathrm{W}_K^\perp$ the $\widehat{\mathcal{L}}^2$-orthogonal complement of W_K and we decompose the N-particle function space $\widehat{\mathscr{H}}^1$ as

$$\widehat{\mathscr{H}}^1 = \mathrm{W}_K \oplus \widetilde{\mathrm{W}_K^\perp} \quad \text{where} \quad \widetilde{\mathrm{W}_K^\perp} := \mathrm{W}_K^\perp \cap \widehat{\mathscr{H}}^1.$$

Additionally, we denote the projection operators associated with this decomposition as $\Pi_{\mathrm{W}_K}\colon \widehat{\mathscr{H}}^1 \to \mathrm{W}_K$ and $\Pi^{\perp}_{\mathrm{W}_K}\colon \widehat{\mathscr{H}}^1 \to \widetilde{\mathrm{W}^{\perp}_K}$.

Using the fact that $\widetilde{\Psi}^*_{K,\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ and $\widetilde{\Psi}^*_{\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi^*_0}$, a direct calculation now reveals that

$$0 = \left(\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi_{0,K}\right)_{\widehat{\mathscr{L}}^2} = \left(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi_{0,K}\right)_{\widehat{\mathscr{L}}^2} + \underbrace{\left(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi_{0,K}\right)_{\widehat{\mathscr{L}}^2}}_{=0 \text{ since } \widetilde{\mathrm{W}^{\perp}_K} \subset \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}} = \left(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi_{0,K}\right)_{\widehat{\mathscr{L}}^2},$$

$$0 = \left(\widetilde{\Psi}^*_{\mathrm{GS}}, \Psi^*_0\right)_{\widehat{\mathscr{L}}^2} = \left(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi^*_0\right)_{\widehat{\mathscr{L}}^2} + \underbrace{\left(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}, \Psi^*_0\right)_{\widehat{\mathscr{L}}^2}}_{=0 \text{ since } \widetilde{\mathrm{W}^{\perp}_K} \subset \widehat{\mathscr{H}}^{1,\perp}_{\Psi^*_0}} = \left(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}, \Psi^*_0\right)_{\widehat{\mathscr{L}}^2}.$$

Consequently, it holds that

$$\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \quad \text{and} \quad \Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}},$$

and

$$\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi^*_0} \quad \text{and} \quad \Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi^*_0}.$$

Thanks to the linearity of the mappings $\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \ni \Phi_K \mapsto \mathcal{C}_K(\Phi_K)$ and $\widehat{\mathscr{H}}^{1,\perp}_{\Psi^*_0} \ni \Phi \mapsto \mathcal{C}(\Phi)$ (see Proposition 14), we can thus deduce that for all $K \geq N$ it holds that

$$\begin{aligned}\|\mathcal{C}_K(\widetilde{\Psi}^*_{K,\mathrm{GS}}) - \mathcal{C}(\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \leq{}& \underbrace{\|\mathcal{C}_K(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}) - \mathcal{C}(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}}_{:=(\mathrm{I})} \\ &+ \underbrace{\|\mathcal{C}_K(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}) - \mathcal{C}(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}}_{:=(\mathrm{II})}. \end{aligned} \tag{24}$$

Let us first consider the term (II). Since the function $\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}} \in \widetilde{\mathrm{W}^{\perp}_K} \subset \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$, we can use the linearity property of cluster operators from Proposition 14 together with the mapping properties of cluster operators from Theorem 16, to further deduce that all $K \geq N$ it holds that

$$\begin{aligned}(\mathrm{II}) &\leq \|\mathcal{C}(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}) - \mathcal{C}_K(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} + \|\mathcal{C}_K(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}} - \Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \\ &\leq \|\mathcal{C}(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}) - \mathcal{C}_K(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} + \beta_{\widehat{\mathscr{H}}}\|\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}} - \Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1} \\ &\leq \|\mathcal{C}(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}) - \mathcal{C}_K(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} + C\|\widetilde{\Psi}^*_{K,\mathrm{GS}} - \widetilde{\Psi}^*_{\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1}, \end{aligned} \tag{25}$$

where the last step also introduces the continuity constant of the projection operator $\Pi^{\perp}_{\mathrm{W}_K}\colon \widehat{\mathscr{H}}^1 \to \widetilde{\mathrm{W}^{\perp}_K}$ which is uniformly bounded in $K \geq N$ thanks to **Assumption A.II**.

Recalling now the definitions of the functions $\widetilde{\Psi}^*_{K,\mathrm{GS}}$ and $\widetilde{\Psi}^*_{\mathrm{GS}}$, we deduce that for all $K \geq N$ it holds that

$$\begin{aligned}\|\widetilde{\Psi}^*_{K,\mathrm{GS}} - \widetilde{\Psi}^*_{\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1} &= \|c_{0,K}\Psi^*_{\mathrm{GS}} - \Psi_{0,K} - c_0\Psi^*_{\mathrm{GS}} + \Psi^*_0\|_{\widehat{\mathscr{H}}^1} \\ &\leq |c_{0,K} - c_0|\|\Psi^*_{\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1} + \|\Psi^*_0 - \Psi_{0,K}\|_{\widehat{\mathscr{H}}^1}. \end{aligned} \tag{26}$$

Combining the estimates (25) and (26), making use of the definitions of the constants $c_{0,K}$ and c_0, and appealing to **Assumption A.II** we can conclude that

$$\lim_{K\to\infty} (\mathrm{II}) \leq \lim_{K\to\infty} \underbrace{\|\mathcal{C}(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}) - \mathcal{C}_K(\Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1\to\widehat{\mathscr{H}}^1}}_{:=(\mathrm{III})}, \tag{27}$$

and therefore, in order to prove our required estimate, it suffices to study the terms (I) and (III).

Step Three:

We begin with the study of the term (III) appearing in Inequality (27). For simplicity, let us fix $K \geq N$ and let us denote $\Upsilon_K := \Pi^{\perp}_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}$.

Let now $\{\psi_{j,0}\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ and $\{\psi_{j,K}\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ denote $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal, single-particle functions used to construct the reference determinants $\Psi_0^* \in \widehat{\mathscr{H}}^1$ and $\Psi_{0,K} \in \widehat{\mathscr{H}}^1$ respectively with the ordering convention that

$$\forall j \in \{1,\dots,N\}\colon \quad \lim_{K\to\infty} \|\psi_{j,K} - \psi_{j,0}\|_{\mathrm{H}^1(\mathbb{R}^3)} = 0. \tag{28}$$

The fact that such an ordering convention is indeed possible is a consequence of **Assumption A.II** and a proof is given in Appendix D.

For simplicity, we now assume that for all $K \geq N$ and all $j \in \{1,\dots,N\}$ it holds that

$$\psi_{j,K} \neq \psi_{j,0} \quad \text{so that} \quad \dim\{\psi_{j,0}\}_{j=1}^N \cup \{\psi_{j,K}\}_{j=1}^N = 2N. \tag{29}$$

If this is not the case, then the subsequent arguments must be rephrased in terms of subsequences. More precisely, we must individually consider all subsequences $\{K_\ell\}_{\ell\in\mathbb{N}} \subset \mathbb{N}$ with the property that there exist fixed indices $j_1,\dots,j_L \in \{1,\dots,N\}$ such that for all $\{K_\ell\}_{\ell\in\mathbb{N}}$ and all $i \in \{1,\dots,L\}$ it holds that

$$\psi_{j_i,K_\ell} = \psi_{j_i,0} \quad \text{so that} \quad \dim\{\psi_{j,0}\}_{j=1}^N \cup \{\psi_{j,K_\ell}\}_{j=1}^N = 2N - L.$$

However, this formalism introduces considerable notational overhead without affecting our core arguments. For clarity of exposition therefore, we will assume that the relation (29) holds.

Using now the single-particle functions introduced above, we define the following subsets of $\mathrm{H}^1(\mathbb{R}^3)$.

$$\mathscr{W} := \mathrm{span}\{\psi_{j,0}\}_{j=1}^N \cup \{\psi_{j,K}\}_{j=1}^N, \quad \text{and} \quad \mathscr{W}^{\perp} := \left\{\phi \in \mathrm{L}^2(\mathbb{R}^3)\colon (\phi,\psi)_{\mathrm{L}^2(\mathbb{R}^3)} = 0 \quad \forall \psi \in \mathscr{W}\right\}.$$

We equip $\mathscr{W}^{\perp}$ with an $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis set $\mathscr{S}^{\perp}$, and we introduce two different $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis sets $\mathscr{S}_0$ and $\mathscr{S}_K$ for the finite-dimensional space $\mathscr{W}$, which are constructed as follows.

- The basis $\mathscr{S}_0$ for $\mathscr{W}$ is obtained by taking the L^2-orthonormal functions $\{\psi_{j,0}\}_{j=1}^N$ and applying a Gram-Schmidt process on the remaining N functions $\{\psi_{j,K}\}_{j=1}^N$. We thus have

  $$\mathscr{S}_0 = \{\psi_{j,0}\}_{j=1}^{2N}$$

  where $\psi_{N+1,0}$ is obtained by orthogonalising $\psi_{1,K}$ with respect to $\{\psi_{j,0}\}_{j=1}^N$, and $\psi_{N+2,0}$ is obtained by orthogonalising $\psi_{2,K}$ with respect to $\{\psi_{j,0}\}_{j=1}^{N+1}$ and so on.

  Note that this Gram-Schmidt process is necessary because the functions $\{\psi_{j,K}\}_{j=1}^N$, while mutually L^2-orthonormal themselves are *not orthogonal* to $\{\psi_{j,0}\}_{j=1}^N$.

- Analogously, the basis $\mathscr{S}_K$ for $\mathscr{W}$ is obtained by taking the L^2-orthonormal functions $\{\psi_{j,K}\}_{j=1}^N$ and applying a Gram-Schmidt process on the remaining N functions $\{\psi_{j,0}\}_{j=1}^N$. We thus have

$$\mathscr{S}_K = \{\psi_{j,K}\}_{j=1}^{2N}$$

  where $\psi_{N+1,K}$ is obtained by orthogonalising $\psi_{1,0}$ with respect to $\{\psi_{j,K}\}_{j=1}^N$, and $\psi_{N+2,K}$ is obtained by orthogonalising $\psi_{2,0}$ with respect to $\{\psi_{j,K}\}_{j=1}^{N+1}$ and so on.

Since both $\mathscr{S}_0$ and $\mathscr{S}_K$ are L^2-orthonormal basis sets for the finite-dimensional space $\mathscr{W}$, there exists a finite-dimensional unitary mapping $\mathcal{U}\colon \mathscr{W} \to \mathscr{W}$ which transforms the basis $\mathscr{S}_0$ into $\mathscr{S}_K$. Obviously, we can view $\mathcal{U}$ as a $2N \times 2N$ matrix $\mathbf{U}$ whose entries are given by $\mathbf{U}_{i,j} = (\psi_{i,K}, \psi_{j,0})_{\mathrm{L}^2(\mathbb{R}^3)}$. It follows from Equation (28), **Assumption A.II**, and a direct, albeit very tedious Gram-Schmidt calculation that the unitary matrix $\mathbf{U}$ defined in this manner converges to a signature matrix in the limit $K \to \infty$, i.e., to a diagonal matrix with entries ± 1. Consequently, by changing the signs of some basis elements in $\mathcal{S}_K$ if necessary, we may assume that $\mathbf{U}$ converges to the $2N \times 2N$ identity matrix in the limit $K \to \infty$.

Next, notice that both $\mathscr{S}_0 \cup \mathscr{S}^\perp$ and $\mathscr{S}_K \cup \mathscr{S}^\perp$ form $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis sets for $\mathrm{H}^1(\mathbb{R}^3)$. Consequently, both bases can be used to define $\widehat{\mathcal{L}}^2$-orthonormal Slater basis sets for $\widehat{\mathcal{H}}^1$. We denote by $\mathscr{B}_{\wedge,0}$ the Slater basis constructed from $\mathscr{S}_0 \cup \mathscr{S}^\perp$ and we denote by $\mathscr{B}_{\wedge,K}$ the Slater basis constructed from $\mathscr{S}_K \cup \mathscr{S}^\perp$. Clearly, the reference determinants Ψ_0^* and $\Psi_{0,K}$ satisfy

$$\Psi_0^* \in \mathscr{B}_{\wedge,0} \quad \text{and} \quad \Psi_{0,K} \in \mathscr{B}_{\wedge,K}.$$

The key point now is that we have constructed $\widehat{\mathcal{L}}^2$-orthonormal Slater basis sets $\mathscr{B}_{\wedge,0}$ and $\mathscr{B}_{\wedge,K}$ for the N-particle sub-spaces $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*}$ and $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ respectively such that

(1) the cluster operators $\mathcal{C}(\Upsilon_K)$ and $\mathcal{C}_K(\Upsilon_K)$ can be expressed explicitly in terms of the Slater bases $\mathscr{B}_{\wedge,0}$ and $\mathscr{B}_{\wedge,K}$ respectively, where we recall that $\Upsilon_K = \Pi^\perp_{\mathrm{W}_K} \widetilde{\Psi}^*_{\mathrm{GS}}$;

(2) only a finite number, at most $\binom{2N}{N}$, of the single-particle functions used to construct these two basis sets vary from one Slater basis to another.

In the remaining part of this proof, we will make use of these two properties.

To begin with, let us consider the Slater basis $\mathscr{B}_{\wedge,0}$. Recall that $\mathscr{B}_{\wedge,0}$ is constructed from the single-particle basis set $\mathscr{S}_0 \cup \mathscr{S}^\perp$, and that $\mathscr{S}_0$ is itself a basis set for the space

$$\mathscr{W} = \mathrm{span}\{\psi_{j,0}\}_{j=1}^N \ \cup\ \{\psi_{j,K}\}_{j=1}^N \quad \text{and and we have imposed the constraint that} \{\psi_{j,0}\}_{j=1}^N \subset \mathscr{S}_0.$$

Consequently, any Slater determinant $\Psi_\nu \in \mathscr{B}_{\wedge,0}$ falls into precisely one of a finite number of categories depending on exactly *how many and which* single-particle functions from $\mathscr{S}_0$ are used to construct it. Thus, there are $\binom{2N}{1}$ categories of Slater determinants in $\mathscr{B}_{\wedge,0}$ that are constructed using exactly 1 element of $\mathscr{S}_0$ and $N-1$ elements of $\mathscr{S}^\perp$, while there $\binom{2N}{2}$ categories of Slater determinants in $\mathscr{B}_{\wedge,0}$ that are constructed using exactly 2 elements of $\mathscr{S}_0$ and $2N-2$ elements of $\mathscr{S}^\perp$ and so on. We therefore decompose the basis set $\mathscr{B}_{\wedge,0}$ as

$$\mathscr{B}_{\wedge,0} = \sum_{\ell=1}^{P} \mathscr{B}^{(\ell)}_{\wedge,0}, \tag{30}$$

where $P = \sum_{i=0}^N \binom{2N}{i}$ denotes the total number of categories highlighted above and for each $\ell \in \{1, \dots, P\}$, the set $\mathscr{B}^{(\ell)}_{\wedge,0}$ consists precisely of all Slater determinants in $\mathscr{B}_{\wedge,0}$ that fall into a particular category.

Corresponding to the decomposition (30), we can now define a complementary decomposition of the space $\widehat{\mathscr{H}}^1$ given by

$$\widehat{\mathscr{H}}^1 = \mathrm{span}\mathscr{B}^{(1)}_{\wedge,0} \cap \widehat{\mathscr{H}}^1 \ \oplus\ \mathrm{span}\mathscr{B}^{(2)}_{\wedge,0} \cap \widehat{\mathscr{H}}^1 \ \oplus\ \cdots\ \oplus\ \mathrm{span}\mathscr{B}^{(P)}_{\wedge,0} \cap \widehat{\mathscr{H}}^1.$$

We denote by $\mathbb{P}_\ell,\ \ell \in \{1,\ldots,P\}$ the corresponding projector operator onto $\mathrm{span}\mathscr{B}^{(\ell)}_{\wedge,0} \cap \widehat{\mathscr{H}}^1$. Since each subspace $\mathrm{span}\mathscr{B}^{(\ell)}_{\wedge,0} \cap \widehat{\mathscr{H}}^1,\ \ell \in \{1,\ldots,P\}$ is a closed subspace of $\widehat{\mathscr{H}}^1$, it follows (see, e.g., [2, Chapter 2.4]) that the corresponding projector $\mathbb{P}_\ell$ is a bounded linear map in the $\widehat{\mathscr{H}}^1$ topology. Moreover, since the continuity constant of each projector depends only on the projection space whose dependence on K is through the the single-particle functions $\{\psi_{j,K}\}_{j=1}^N$, we deduce from **Assumption A.II** that the continuity constants of these projectors are all uniformly bounded in K.

It therefore follows that the term (III) appearing in Inequality (27) can be bounded as

$$\text{(III)} = \|\mathcal{C}(\Upsilon_K) - \mathcal{C}_K(\Upsilon_K)\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \leq \sum_{\ell=1}^{P} \|\mathcal{C}\big(\mathbb{P}_\ell \Upsilon_K\big) - \mathcal{C}_K\big(\mathbb{P}_\ell \Upsilon_K\big)\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}. \tag{31}$$

We will now study the individual terms appearing in the summation in Inequality (31).

To this end, let us pick some $\ell \in \{1,\ldots,P\}$ and denote $\Upsilon_{K,\ell} = \mathbb{P}_\ell \Upsilon_K \in \mathrm{span}\mathscr{B}^{(\ell)}_{\wedge,0} \cap \widehat{\mathscr{H}}^1$. For concreteness, we may assume that the set $\mathcal{B}^{(\ell)}_{\wedge,0}$ consists of all Slater determinants constructed using exactly L single particle basis elements from $\mathscr{S}_0$, which we denote $\psi_{j_1,0}, \psi_{j_2,0}, \ldots, \psi_{j_L,0}$. It follows that we can express the function $\Upsilon_{K,\ell}$ as

$$\Upsilon_{K,\ell} = \sum_{\mu} \boldsymbol{c}_\mu \psi_{j_1,0} \wedge \ldots \wedge \psi_{j_L,0} \wedge \psi^\perp_{\mu_1} \wedge \ldots \wedge \psi^\perp_{\mu_{N-j_L}}. \tag{32}$$

Here, each $\boldsymbol{c}_\mu$ is some real number, the single-particle functions of the form $\{\psi^\perp_{\mu_j}\}$ belong to the basis set $\mathscr{S}^\perp$, the summation is over suitable indices involving these functions, and the so-called wedge product $\wedge$ has been introduced as a short hand notation to signify anti-symmetric tensor products.

It can now readily be seen, recalling the unitary transformation matrix we have introduced earlier, that the function $\Upsilon_{K,\ell}$ can equivalently be expressed in the second Slater basis $\mathscr{B}_{\wedge,K}$ as

$$\begin{aligned}
\Upsilon_{K,\ell} &= \sum_{\mu} \sum_{i_1=1}^{2N} \sum_{i_2=1}^{2N} \cdots \sum_{i_L=1}^{2N} \mathbf{U}_{j_1,i_1} \ldots \mathbf{U}_{j_L,i_L} \boldsymbol{c}_\mu \psi_{i_1,K} \wedge \ldots \wedge \psi_{i_L,K} \wedge \psi^\perp_{\mu_1} \wedge \ldots \wedge \psi^\perp_{\mu_{N-j_L}} \\
&= \sum_{i_1=1}^{2N} \sum_{i_2=1}^{2N} \cdots \sum_{i_L=1}^{2N} \mathbf{U}_{j_1,i_1} \ldots \mathbf{U}_{j_L,i_L} \sum_{\mu} \boldsymbol{c}_\mu \psi_{i_1,K} \wedge \ldots \wedge \psi_{i_L,K} \wedge \psi^\perp_{\mu_1} \wedge \ldots \wedge \psi^\perp_{\mu_{N-j_L}} \\
&= \mathbf{U}_{j_1,j_1} \ldots \mathbf{U}_{j_L,j_L} \sum_{\mu} \boldsymbol{c}_\mu \psi_{j_1,K} \wedge \ldots \wedge \psi_{j_L,K} \wedge \psi^\perp_{\mu_1} \wedge \ldots \wedge \psi^\perp_{\mu_{N-j_L}} \\
&\quad + \sum_{\substack{i_1=1 \\ i_1 \neq j_1}}^{2N} \sum_{\substack{i_2=1 \\ i_2 \neq j_2}}^{2N} \cdots \sum_{\substack{i_L=1 \\ i_L \neq j_L}}^{2N} \mathbf{U}_{j_1,i_1} \ldots \mathbf{U}_{j_L,i_L} \sum_{\mu} \boldsymbol{c}_\mu \psi_{i_1,K} \wedge \ldots \wedge \psi_{i_L,K} \wedge \psi^\perp_{\mu_1} \wedge \ldots \wedge \psi^\perp_{\mu_{N-j_L}},
\end{aligned} \tag{33}$$

where the interchange of the finite and infinite sums is justified by the fact that for fixed indices $i_1, i_2, \ldots i_L$, the infinite series

$$\sum_{\mu} \boldsymbol{c}_\mu \psi_{i_1,K} \wedge \ldots \wedge \psi_{i_L,K} \wedge \psi^\perp_{\mu_1} \wedge \ldots \wedge \psi^\perp_{\mu_{N-j_L}},$$

is absolutely convergent in the $\widehat{\mathcal{L}}^2$-norm. Note that the expressions in Equation (33) may appear complicated but are simply the result of replacing the single-particle functions $\psi_{j_1,0}\cdots\psi_{j_L,0} \in \mathscr{S}_0$ with single-particle functions $\psi_{1,K}\cdots\psi_{2N,K}$ from the alternate basis $\mathscr{S}_K$.

Step Four:

Let us now consider an explicit representation for the first cluster operator $\mathcal{C}(\Upsilon_{K,\ell})$ appearing in the term (III) that we wish to study. Using Equation (32) we can write (suppressing notational details)

$$\mathcal{C}(\Upsilon_{K,\ell}) = \sum_{\mu} \boldsymbol{c}_\mu \mathcal{X}_\mu. \tag{34}$$

Here, each excitation operator $\mathcal{X}_\mu$ is defined with respect to the limiting reference determinant Ψ_0^* and satisfies

$$\mathcal{X}_\mu \Psi_0^* = \psi_{j_1,0} \wedge \ldots \wedge \psi_{j_L,0} \wedge \psi_{\mu_1}^\perp \wedge \ldots \wedge \psi_{\mu_{N-j_L}}^\perp.$$

In other words, recalling Definition 10, we see that the excitation operator $\mathcal{X}_\mu$

$$\text{replaces all occupied orbitals} \quad \underset{\substack{i=1,\ldots,N \\ i\neq j_1,j_2,\ldots,j_L}}{\{\psi_{i,0}\}} \subset \mathscr{S}_0$$

$$\text{with the virtual orbitals} \quad \underset{\substack{i=N+1,\ldots,2N \\ i\in\{j_1,j_2,\ldots,j_L\}}}{\{\psi_{i,0}\}} \subset \mathscr{S}_0 \ \text{ and } \ \psi_{\mu_1}^\perp, \ldots, \psi_{\mu_{N-j_L}}^\perp.$$

Indeed, this is just a consequence of the fact that the limiting reference determinant Ψ_0^* is constructed from the occupied orbitals $\{\psi_{i,0}\}_{i=1}^N$ so that all remaining orbitals are considered virtual.

Using the alternative representation (33), a similar expression can be derived for the second cluster operator $\mathcal{C}_K(\Upsilon_{K,\ell})$ appearing in the term (III) that we wish to study. Indeed, we have

$$\mathcal{C}_K(\Upsilon_{K,\ell}) = \mathbf{U}_{j_1,j_1}\ldots\mathbf{U}_{j_L,j_L}\sum_{\mu}\boldsymbol{c}_\mu\widetilde{\mathcal{X}}_\mu^{j_1,\ldots,j_L} + \sum_{\substack{i_1=1\\ i_1\neq j_1}}^{2N}\sum_{\substack{i_2=1\\ i_2\neq j_2}}^{2N}\cdots\sum_{\substack{i_L=1\\ i_L\neq j_L}}^{2N}\mathbf{U}_{j_1,i_1}\ldots\mathbf{U}_{j_L,i_L}\sum_{\mu}\boldsymbol{c}_\mu\widetilde{\mathcal{X}}_\mu^{i_1,\ldots,i_L}. \tag{35}$$

In contrast to the simpler expression (34), each excitation operator $\widetilde{\mathcal{X}}_\mu^{i_1,\ldots,i_L}$ in Equation (37) is defined with respect to the *second* reference determinant $\Psi_{0,K}$ and is such that

$$\widetilde{\mathcal{X}}_\mu^{i_1,\ldots,i_L}\Psi_{0,K} = \psi_{i_1,K} \wedge \ldots \wedge \psi_{i_L,K} \wedge \psi_{\mu_1}^\perp \wedge \ldots \wedge \psi_{\mu_{N-j_L}}^\perp.$$

Recalling once again Definition 10, we see that the excitation operator $\widetilde{\mathcal{X}}_\mu^{i_1,\ldots,i_L}$

$$\text{replaces all occupied orbitals} \quad \underset{\substack{i=1,\ldots,N \\ i\neq i_1,i_2,\ldots,i_L}}{\{\psi_{i,K}\}} \subset \mathscr{S}_K$$

$$\text{with the virtual orbitals} \quad \underset{\substack{i=N+1,\ldots,2N \\ i\in\{i_1,i_2,\ldots,i_L\}}}{\{\psi_{i,K}\}} \subset \mathscr{S}_K \ \text{ and } \ \psi_{\mu_1}^\perp, \ldots, \psi_{\mu_{N-j_L}}^\perp.$$

The key point now is that the excitation operator $\mathcal{X}_\mu$ and the "diagonal" ($i_1 = j_1, \ldots, i_L = j_L$) excitation operator $\widetilde{\mathcal{X}}_\mu^{j_1,\ldots,j_L}$ can be directly compared. To see this, notice first that we may assume

without loss of generality[2] that the indices $j_1, \dots, j_L$ satisfy, for some $M \in \{1, \dots, N\}$

$$j_1 < \dots < j_M \in \{1, \dots, N\} \quad \text{and} \quad j_{M+1} < \dots < j_L \in \{N+1, \dots, 2N\}. \tag{36}$$

Next, let us introduce two generalised excitation operators $\widetilde{\mathcal{X}}^{j_1,\dots,j_M} : \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ and $\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} : \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ such that

$$\widetilde{\mathcal{X}}^{j_1,\dots,j_M} \quad \text{replaces occupied orbitals} \quad \underset{\substack{i=1,\dots,N \\ i\neq j_1,j_2,\dots,j_M}}{\{\psi_{i,K}\}} \subset \mathscr{S}_K \quad \text{with occupied orbitals} \quad \underset{\substack{i=1,\dots,N \\ i\neq j_1,j_2,\dots,j_M}}{\{\psi_{i,0}\}} \subset \mathscr{S}_0,$$

$$\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \quad \text{replaces virtual orbitals} \quad \underset{\substack{i=N+1,\dots,2N \\ i\in\{j_{M+1},\dots,j_L\}}}{\{\psi_{i,0}\}} \subset \mathscr{S}_0 \quad \text{with virtual orbitals} \quad \underset{\substack{i=N+1,\dots,2N \\ i\in\{j_{M+1},\dots,j_L\}}}{\{\psi_{i,K}\}} \subset \mathscr{S}_K.$$

It then follows that for every excitation index μ appearing in Equations (34) and (35), we can write

$$\widetilde{\mathcal{X}}_\mu^{j_1,\dots,j_L} = \widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \mathcal{X}_\mu \widetilde{\mathcal{X}}^{j_1,\dots,j_M}. \tag{37}$$

Applying now Equation (37) to Equation (35) which expresses the cluster operator $\mathcal{C}_K(\Upsilon_{K,\ell})$, we deduce that

$$\begin{aligned}\mathcal{C}_K(\Upsilon_{K,\ell}) =& \mathbf{U}_{j_1,j_1} \dots \mathbf{U}_{j_L,j_L} \left(\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L}\right) \left(\sum_\mu \boldsymbol{c}_\mu \mathcal{X}_\mu\right) \left(\widetilde{\mathcal{X}}^{j_1,\dots,j_M}\right) \\ &+ \sum_{\substack{i_1=1 \\ i_1\neq j_1}}^{2N} \sum_{\substack{i_2=1 \\ i_2\neq j_2}}^{2N} \dots \sum_{\substack{i_L=1 \\ i_L\neq j_L}}^{2N} \mathbf{U}_{j_1,i_1} \dots \mathbf{U}_{j_L,i_L} \sum_\mu \boldsymbol{c}_\mu \widetilde{\mathcal{X}}_\mu^{i_1,\dots,i_L}.\end{aligned}$$

Returning now to the bound given by Inequality (31) for the term (III), we see that

$$\begin{aligned}\text{(III)} =& \|\mathcal{C}\big(\mathbb{P}_\ell \Upsilon_K\big) - \mathcal{C}_K\big(\mathbb{P}_\ell \Upsilon_K\big)\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \\ =& \|\mathcal{C}\big(\Upsilon_{K,\ell}\big) - \mathcal{C}_K\big(\Upsilon_{K,\ell}\big)\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \\ \leq& \underbrace{\left\| \left(\sum_\mu \boldsymbol{c}_\mu \mathcal{X}_\mu\right) - \mathbf{U}_{j_1,j_1} \dots \mathbf{U}_{j_L,j_L} \left(\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L}\right) \left(\sum_\mu \boldsymbol{c}_\mu \mathcal{X}_\mu\right) \left(\widetilde{\mathcal{X}}^{j_1,\dots,j_M}\right) \right\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}}_{:=\text{(IV)}} \\ &+ \underbrace{\left\| \sum_{\substack{i_1=1 \\ i_1\neq j_1}}^{2N} \sum_{\substack{i_2=1 \\ i_2\neq j_2}}^{2N} \dots \sum_{\substack{i_L=1 \\ i_L\neq j_L}}^{2N} \mathbf{U}_{j_1,i_1} \dots \mathbf{U}_{j_L,i_L} \sum_\mu \boldsymbol{c}_\mu \widetilde{\mathcal{X}}_\mu^{i_1,\dots,i_L} \right\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}}_{:=\text{(V)}}\end{aligned} \tag{38}$$

To conclude our study of the term (III), we will now argue that both terms (IV) and (V) converge to zero in the limit $K \to \infty$.

Step Five:

[2]The only other possibility is that all indices $j_1 < \dots j_L \in \{N+1, \dots, 2N\}$, and the same fundamental arguments apply with a slightly different notation.

We begin by considering the term (IV) appearing in Equation (38). For clarity of exposition, let us denote $\mathcal{Y} := \sum_\mu \boldsymbol{c}_\mu \mathcal{X}_\mu$. We can then deduce that

$$\begin{aligned}
\text{(IV)} &= \left\| \mathcal{Y} - \mathbf{U}_{j_1,j_1} \dots \mathbf{U}_{j_L,j_L} \left(\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \right) \mathcal{Y} \left(\widetilde{\mathcal{X}}^{j_1,\dots,j_M} \right) \right\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1} \\
&\leq \underbrace{\left\| \left(\mathrm{I} - \mathbf{U}_{j_1,j_1} \dots \mathbf{U}_{j_L,j_L} \widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \right) \mathcal{Y} \right\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}}_{:=\text{(IVA)}} \\
&\quad + \underbrace{|\mathbf{U}_{j_1,j_1} \dots \mathbf{U}_{j_L,j_L}| \left\| \left(\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \mathcal{Y} \right) \left(\mathrm{I} - \widetilde{\mathcal{X}}^{j_1,\dots,j_M} \right) \right\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}}_{:=\text{(IVB)}}
\end{aligned} \tag{39}$$

Consider the first term (IVA). As argued previously, $\mathbf{U}_{j_1,j_1} \dots \mathbf{U}_{j_L,j_L} \to 1$ as $K \to 0$, and $\mathcal{Y} = \sum_\mu \boldsymbol{c}_\mu \mathcal{X}_\mu = \mathcal{C}(\Upsilon_{K,\ell})$ is a bounded linear map on $\widehat{\mathscr{H}}^1$ since $\Upsilon_{K,\ell} \in \widehat{\mathscr{H}}^1$ (recall Theorem 16). Consequently, it suffices to show that

$$\lim_{K \to \infty} \sup_{0 \neq \Phi \in \ \mathrm{Ran}\mathcal{Y}} \frac{\left\| \left(\mathrm{I} - \widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \right) \Phi \right\|_{\widehat{\mathscr{H}}^1}}{\|\Phi\|_{\widehat{\mathscr{H}}^1}} = 0. \tag{40}$$

Consider an arbitrary $\Phi \in \mathrm{Ran}\mathcal{Y}$. Recalling the Slater basis $\mathscr{B}_{\wedge,0}$ introduced in **Step Three** above, we see that we can express Φ in the form

$$\Phi = \sum_\mu \mathbf{d}_\mu \psi_{j_{M+1},0} \wedge \dots \wedge \psi_{j_L,0} \wedge \psi^\perp_{\mu_1} \wedge \dots \wedge \psi^\perp_{\mu_{N-L+M}}, \tag{41}$$

where each $\mathbf{d}_\mu$ is some real number, the single-particle functions of the form $\{\psi^\perp_{\mu_j}\}$ belong to the basis set $\mathscr{S}^\perp$, the summation is over suitable indices involving these functions, and we once again use the wedge product $\wedge$ as a short hand notation to signify anti-symmetric tensor products. Note that the appearance of single-particle functions $\psi_{j_{M+1},0} \wedge \dots \wedge \psi_{j_L,0}$ in the expression for $\Phi \in \mathrm{Ran}\mathcal{Y}$ is a consequence of the definition of $\mathcal{Y} = \sum_\mu \boldsymbol{c}_\mu \mathcal{X}_\mu = \mathcal{C}(\Upsilon_{K,\ell})$ (see **Step Three** and **Step Four** above) together with the ordering of the indices $j_{M+1}, \dots, j_L$ given by Equation (36).

It then follows from the definition of the generalised excitation operator $\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L}$ that

$$\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \Phi = \sum_\mu \mathbf{d}_\mu \psi_{j_{M+1},K} \wedge \dots \wedge \psi_{j_L,K} \wedge \psi^\perp_{\mu_1} \wedge \dots \wedge \psi^\perp_{\mu_{N-L+M}}.$$

Appealing now to the unitary matrix $\mathbf{U} \in \mathbb{R}^{2N \times 2N}$ introduced in **Step Three**, we deduce that

$$\begin{aligned}
&\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L} \Phi \\
&\qquad = \sum_{i_{M+1}=1}^{2N} \dots \sum_{i_L=1}^{2N} \mathbf{U}_{i_{M+1},j_{M+1}} \dots \mathbf{U}_{i_L,j_L} \sum_\mu \mathbf{d}_\mu \psi_{i_{M+1},0} \wedge \dots \wedge \psi_{i_L,0} \wedge \psi^\perp_{\mu_1} \wedge \dots \wedge \psi^\perp_{\mu_{N-L+M}},
\end{aligned} \tag{42}$$

where the interchange of the finite and infinite sums is once again justified by the fact that for fixed indices $i_1, i_2, \dots i_L$, the infinite series above is absolutely convergent in the $\widehat{\mathcal{L}}^2$-norm.

Using Equation (42), we can thus write:

$$\begin{aligned}
&\left\|\left(\mathrm{I}-\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L}\right)\Phi\right\|_{\widehat{\mathscr{H}}^1}\\
&\leq \left|1-\mathbf{U}_{j_{M+1},j_{M+1}}\dots\mathbf{U}_{j_L,j_L}\right|\underbrace{\left\|\sum_\mu \mathbf{d}_\mu\psi_{j_{M+1},0}\wedge\ldots\wedge\psi_{j_L,0}\wedge\psi_{\mu_1}^\perp\wedge\ldots\wedge\psi_{\mu_{N-L+M}}^\perp\right\|_{\widehat{\mathscr{H}}^1}}_{=\Phi}\\
&+\left\|\sum_{\substack{i_{M+1}=1\\ i_{M+1}\neq j_{M+1}}}^{2N}\dots\sum_{\substack{i_L=1\\ i_L\neq j_L}}^{2N}\mathbf{U}_{i_{M+1},j_{M+1}}\dots\mathbf{U}_{i_L,j_L}\sum_\mu \mathbf{d}_\mu\psi_{i_{M+1},0}\wedge\ldots\wedge\psi_{i_L,0}\wedge\psi_{\mu_1}^\perp\wedge\ldots\wedge\psi_{\mu_{N-L+M}}^\perp\right\|_{\widehat{\mathscr{H}}^1}.
\end{aligned}$$

As described in **Step Three**, the definition of the unitary matrix $\mathbf{U}$ implies that $\mathbf{U}_{ij}=\delta_{ij}$ for all $i,j\in\{1,\dots,2N\}$ in the limit $K\to\infty$. Consequently, in order to establish the required Equation (40), we need only show that for any choice of indices $i_{M+1},\dots,i_L\in\{1,\dots,2N\}$, there exists a constant $C_{i_{M+1},\dots,i_L}>0$ such that

$$\begin{aligned}
&\left\|\sum_\mu \mathbf{d}_\mu\psi_{i_{M+1},0}\wedge\ldots\wedge\psi_{i_L,0}\wedge\psi_{\mu_1}^\perp\wedge\ldots\wedge\psi_{\mu_{N-L+M}}^\perp\right\|_{\widehat{\mathscr{H}}^1}\\
&\leq C_{i_{M+1},\dots,i_L}\|\Phi\|_{\widehat{\mathscr{H}}^1}\\
&= C_{i_{M+1},\dots,i_L}\left\|\sum_\mu \mathbf{d}_\mu\psi_{j_{M+1},0}\wedge\ldots\wedge\psi_{j_L,0}\wedge\psi_{\mu_1}^\perp\wedge\ldots\wedge\psi_{\mu_{N-L+M}}^\perp\right\|_{\widehat{\mathscr{H}}^1}.
\end{aligned}$$

In other words, we must show that, for any choice of indices $i_{M+1},\dots,i_L\in\{1,\dots,2N\}$, the so-called *generalised* excitation operator $\mathcal{X}^{i_{M+1},\dots i_L}_{j_{M+1},\dots j_L}:\widehat{\mathcal{L}}^2\to\widehat{\mathcal{L}}^2$ which

$$\begin{aligned}
&\text{replaces the orbitals } \psi_{j_{M+1},0},\dots,\psi_{j_L,0}\subset\mathscr{S}_0\\
&\text{with the orbitals } \psi_{i_{M+1},0},\dots,\psi_{i_L,0}\subset\mathscr{S}_0,
\end{aligned}\tag{43}$$

is, in fact, a bounded linear map from $\widehat{\mathscr{H}}^1\to\widehat{\mathscr{H}}^1$. Notice that this result does *not* follow directly from Theorem 16 since the excitation operators considered therein always replace occupied orbitals (in this case $\psi_{1,0},\dots,\psi_{N,0}$) with virtual orbitals $\{\psi_{j,0}\}_{j>N+1}$. In the current setting however, if we consider, for instance the case, $j_{M+1}=N+1, i_{M+1}=N+2$, then the excitation operator $\mathcal{X}^{i_{M+1},\dots i_L}_{j_{M+1},\dots j_L}$ can be seen to excite a virtual orbital to another virtual orbital, and it is not a priori clear that such an excitation operator is $\widehat{\mathscr{H}}^1$-bounded.

A rigorous demonstration that the generalised operator $\mathcal{X}^{i_{M+1},\dots i_L}_{j_{M+1},\dots j_L}:\widehat{\mathscr{H}}^1\to\widehat{\mathscr{H}}^1$ is indeed bounded can be derived from the general framework introduced by Rohwedder [25] for analysing the continuity properties of the usual occupied-to-virtual cluster operators but this derivation is not trivial. In order to avoid distracting from the core results of this paper, we postpone this demonstration to Appendix C. For the moment, we continue with the remainder of the current proof and show that the convergence of some remaining terms in our estimates (namely, the convergence of the terms (IVB) and (V)) can be reduced to establishing similar boundedness results for generalised occupied-to-occupied and virtual-to-virtual excitation operators. Note that the core part of these demonstrations is rather similar to the arguments for the term (IVA) and very tedious to write in full detail so we will only provide a sketch.

To start with, we consider the term (IVB) appearing in Inequality (40). Thanks to our analysis of the term (IVA) and, in particular, the generalised excitation operator $\widetilde{\mathcal{X}}_{j_{M+1},\dots,j_L}$, and recalling the convergence properties of the unitary matrix $\mathbf{U}$, we see that it suffices to show that

$$\lim_{K\to\infty} \sup_{0\neq\Phi\in\ \widehat{\mathscr{H}}^1} \frac{\left\| \mathcal{Y}\left(\mathrm{I} - \widetilde{\mathcal{X}}^{j_1,\dots,j_M}\right)\Phi\right\|_{\widehat{\mathscr{H}}^1}}{\|\Phi\|_{\widehat{\mathscr{H}}^1}} = 0. \tag{44}$$

Using an argument very similar– albeit more tedious– to the one carried out to establish Equation (40), it can be shown that the limit (44) indeed holds provided we can show that for any choice of indices $i_1,\dots,i_M \in \{1,\dots,2N\}$, the generalised excitation operator $\mathcal{X}^{i_1,\dots i_M}_{j_1,\dots j_M}\colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ which

$$\text{replaces the orbitals } \ \psi_{j_1,0},\dots,\psi_{j_M,0} \subset \mathscr{S}_0 \tag{45}$$

$$\text{with the orbitals } \ \psi_{i_1,0},\dots,\psi_{i_M,0} \subset \mathscr{S}_0,$$

is, in fact, a bounded linear map from $\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1$.

In other words, similar to the case of the term (IVA), the required convergence of the term (IVB) in the limit $K \to \infty$ can be established by demonstrating the $\widehat{\mathscr{H}}^1$-boundedness of occupied-to-occupied generalised excitation operators. As before, a demonstration of this continuity result is given in Appendix C.

Having completed our study of the terms (IVA) and (IVB) and thus the term (IV) appearing in Inequality (38), we can now proceed to an analysis of the term (V) from Inequality (38). In order to demonstrate that the term (V) also converges to zero in the limit $K \to \infty$, we recall the convergence properties of the unitary matrix $\mathbf{U}$ and deduce that it suffices to prove that for any choice of indices $i_1,\dots,i_L \in \{1,\dots,2N\}$, the cluster operator

$$\sum_\mu \boldsymbol{c}_\mu \widetilde{\mathcal{X}}^{i_1,\dots,i_L}_\mu \quad \text{is a bounded linear map from } \ \widehat{\mathscr{H}}^1 \text{ to } \widehat{\mathscr{H}}^1. \tag{46}$$

It can now be seen that for any choice of indices $i_1,\dots,i_L \in \{1,\dots,2N\}$, there exists corresponding generalised occupied-to-occupied excitation operator $\widetilde{\mathcal{X}}^{j_1,\dots,j_L}_{i_1,\dots,i_L}\colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ and a generalised virtual-to-virtual excitation operator $\widetilde{\mathcal{X}}^{i_1,\dots,i_L}_{j_1,\dots,j_L}\colon \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ such that

$$\sum_\mu \boldsymbol{c}_\mu \widetilde{\mathcal{X}}^{i_1,\dots,i_L}_\mu = \widetilde{\mathcal{X}}^{i_1,\dots,i_L}_{j_1,\dots,j_L}\left(\sum_\mu \boldsymbol{c}_\mu \widetilde{\mathcal{X}}^{j_1,\dots,j_L}_\mu\right)\widetilde{\mathcal{X}}^{j_1,\dots,j_L}_{i_1,\dots,i_L}. \tag{47}$$

Here, the generalised occupied-to-occupied excitation operator $\widetilde{\mathcal{X}}^{j_1,\dots,j_L}_{i_1,\dots,i_L}$

$$\text{replaces all occupied orbitals } \ \underset{\substack{i=1,\dots,N\\ i\neq i_1,i_2,\dots,i_L}}{\{\psi_{i,K}\}} \subset \mathscr{S}_K \qquad \text{with the occupied orbitals } \ \underset{\substack{j=1,\dots,N\\ j\neq j_1,j_2,\dots,j_L}}{\{\psi_{j,K}\}} \subset \mathscr{S}_K,$$

while the generalised virtual-to-virtual excitation operator $\widetilde{\mathcal{X}}^{i_1,\dots,i_L}_{j_1,\dots,j_L}$

$$\text{replaces all virtual orbitals } \ \underset{\substack{j=N+1,\dots,2N\\ j\in\{j_1,j_2,\dots,j_L\}}}{\{\psi_{j,K}\}} \subset \mathscr{S}_K \qquad \text{with the virtual orbitals } \ \underset{\substack{i=N+1,\dots,2N\\ i\in\{i_1,i_2,\dots,i_L\}}}{\{\psi_{i,K}\}} \subset \mathscr{S}_K,$$

As a consequence of our analysis of the term (IV), we can already conclude that the cluster operator

$$\sum_\mu \boldsymbol{c}_\mu \widetilde{\mathcal{X}}^{j_1,\dots,j_L}_\mu \quad \text{is a bounded linear map from } \ \widehat{\mathscr{H}}^1 \text{ to } \widehat{\mathscr{H}}^1.$$

Consequently, as in the proof of convergence of the terms (IVA) and (IVB), we deduce from (46) and (47) that the term (V) also converges to zero in the limit $K \to \infty$ provided we can establish the $\widehat{\mathscr{H}}^1$ -boundedness of generalised occupied-to-occupied and virtual-to-virtual excitation operators. Once again we refer the readers to Appendix C for a demonstration of this result.

Step Six:

To conclude the proof, we return finally to the term (I) from Inequality (24). We must now show that

$$\lim_{K\to\infty} \|\mathcal{C}_K(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}) - \mathcal{C}(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1\to\widehat{\mathscr{H}}^1} = 0.$$

To this end, let us first fix $K \geq N$. Notice that $\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}},\ \Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}} \in \mathrm{W}_K = \mathrm{span}\{\Psi_{0,K}, \Psi^*_0\}$. Recall also from **Step Three**, that we have introduced Slater bases

$$\mathscr{B}_{\wedge,0} \text{ and } \mathscr{B}_{\wedge,K} \quad \text{constructed from single-particle bases} \quad \mathscr{S}_0 \cup \mathscr{S}^\perp \text{ and } \mathscr{S}_K \cup \mathscr{S}^\perp \text{ respectively.}$$

Let us now introduce the following subsets of the Slater bases $\mathscr{B}_{\wedge,0}$ and $\mathscr{B}_{\wedge,K}$:

$$\mathscr{B}^{\Pi}_{\wedge,0} = \big\{\psi_{i_1,0} \wedge \ldots \wedge \psi_{i_N,0} \colon \quad \psi_{i_j,0} \in \mathscr{S}_0\big\} \subset \mathscr{B}_{\wedge,0},$$
$$\mathscr{B}^{\Pi}_{\wedge,K} = \big\{\psi_{i_1,K} \wedge \ldots \wedge \psi_{i_N,K} \colon \quad \psi_{i_j,K} \in \mathscr{S}_K\big\} \subset \mathscr{B}_{\wedge,K}.$$

We claim that $\mathrm{W}_K \subset \mathrm{span}\mathscr{B}^{\Pi}_{\wedge,0}$ and $\mathrm{W}_K \subset \mathrm{span}\mathscr{B}^{\Pi}_{\wedge,K}$. Indeed, by the definition of the single-particle basis sets $\mathscr{S}_0$ and $\mathscr{S}_K$, we obviously have $\Psi^*_0 \in \mathscr{B}^{\Pi}_{\wedge,0}$ and $\Psi_K \in \mathscr{B}^{\Pi}_{\wedge,K}$. Since, the unitary matrix $\mathbf{U}$ provides a transformation between the bases $\mathscr{S}_0$ and $\mathscr{S}_K$, the claim readily follows.

Next, let us introduce the $\binom{2N}{N}$-dimensional index set $\mathcal{K} \subset \{1,\ldots,2N\}^N$ as

$$\mathcal{K} = \Big\{\mu = (\mu_1,\ldots,\mu_N) \colon 1 < \mu_1 < \ldots < \mu_N < 2N\Big\}.$$

It follows that $\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}$ and $\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}$ have *finite-dimensional* representations of the form

$$\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}} = \sum_{\mu\in\mathcal{K}} \mathbf{a}_{\mu,0}\Psi_{\mu,0} \qquad \text{where} \quad \Psi_{\mu,0} = \psi_{\mu_1,0} \wedge \ldots \wedge \psi_{\mu_N,0}, \tag{48}$$
$$\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}} = \sum_{\mu\in\mathcal{K}} \mathbf{a}_{\mu,K}\Psi_{\mu,K} \quad \text{where} \quad \Psi_{\mu,K} = \psi_{\mu_1,K} \wedge \ldots \wedge \psi_{\mu_N,K}.$$

We now claim that for all $\mu \in \mathcal{K}$ it holds that

$$\lim_{K\to\infty} a_{\mu,K} = a_{\mu,0}. \tag{49}$$

Indeed, a direct calculation shows that

$$\begin{aligned} |a_{\mu,0} - a_{\mu,K}| &= \Big(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}, \Psi_{\mu,0}\Big)_{\widehat{\mathcal{L}}^2} - \Big(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi_{\mu,K}\Big)_{\widehat{\mathcal{L}}^2} \\ &= \Big(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}} - \Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi_{\mu,0}\Big)_{\widehat{\mathcal{L}}^2} + \Big(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}, \Psi_{\mu,0} - \Psi_{\mu,K}\Big)_{\widehat{\mathcal{L}}^2}, \end{aligned}$$

Since additionally, **Assumption A.II** yields that $\Psi_{0,K} \to \Psi^*_0$ in the $\widehat{\mathcal{L}}^2$-sense as $K \to \infty$, and we have by definition that $\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}} \to \Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}$ in the $\widehat{\mathcal{L}}^2$-sense as $K \to \infty$ (recall Equation (23)), the claim now readily follows.

Using now Equation (48), we can express the cluster operators $\mathcal{C}_K(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}})$ and $\mathcal{C}(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}})$ as

$$\mathcal{C}_K(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{K,\mathrm{GS}}) = \sum_{\mu\in\mathcal{K}} \mathbf{a}_{\mu,0}\mathcal{X}_{\mu,0} \quad \text{and} \quad \mathcal{C}(\Pi_{\mathrm{W}_K}\widetilde{\Psi}^*_{\mathrm{GS}}) = \sum_{\mu\in\mathcal{K}} \mathbf{a}_{\mu,K}\mathcal{X}_{\mu,K},$$

where, for each $\mu \in \mathcal{K}$, the excitation operator $\mathcal{X}_{\mu,0}$ is generated by the determinant $\Psi_{\mu,0}$ and is defined with respect to the reference Ψ_0^* while the excitation operator $\mathcal{X}_{\mu,K}$ is generated by the determinant $\Psi_{\mu,K}$ and is defined with respect to the reference $\Psi_{0,K}$.

In view of Equation (49), and since the cluster operators under consideration are *finite sums* of excitation operators, we conclude that it suffices to show that

$$\forall \mu \in \mathcal{K}: \qquad \lim_{K\to\infty} \|\mathcal{X}_{\mu,0} - \mathcal{X}_{\mu,K}\|_{\widehat{\mathcal{H}}^1\to\widehat{\mathcal{H}}^1} = 0. \tag{50}$$

Consider an arbitrary index $\mu = (\mu_1, \ldots, \mu_N)$ of the form[3]

$$1 < \mu_1 < \ldots \mu_M \leq N \quad \text{and} \quad N < \mu_{M+1} < \ldots < \mu_N < 2N \qquad \text{where } M \in \{1, \ldots, N\} \text{ is arbitrary.}$$

It follows that the excitation operator $\mathcal{X}_{\mu,0}$

$$\text{replaces occupied orbitals } \{\psi_{i,0}\}_{\substack{i=1,\ldots,N\\ i\neq\mu_1,\ldots\mu_M}} \qquad \text{with virtual orbitals } \{\psi_{i,0}\}_{\substack{i=N+1,\ldots,2N\\ i=\mu_{M+1},\ldots\mu_N}}.$$

Similarly the excitation operator $\mathcal{X}_{\mu,K}$

$$\text{replaces occupied orbitals } \{\psi_{i,K}\}_{\substack{i=1,\ldots,N\\ i\neq\mu_1,\ldots\mu_M}} \qquad \text{with virtual orbitals } \{\psi_{i,K}\}_{\substack{i=N+1,\ldots,2N\\ i=\mu_{M+1},\ldots\mu_N}}.$$

Consequently, we may introduce generalised excitation operators $\widetilde{\mathcal{X}}_{\mu_1,\ldots,\mu_M} : \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ and $\widetilde{\mathcal{X}}^{\mu_{M+1},\ldots,\mu_N} : \widehat{\mathcal{L}}^2 \to \widehat{\mathcal{L}}^2$ such that

$$\widetilde{\mathcal{X}}_{\mu_1,\ldots,\mu_M} \text{ replaces the orbitals } \{\psi_{i,K}\}_{\substack{i=1,\ldots,N\\ i\neq\mu_1,\ldots\mu_M}} \text{ with the orbitals } \{\psi_{i,0}\}_{\substack{i=1,\ldots,N\\ i\neq\mu_1,\ldots\mu_M}} \text{ and}$$

$$\widetilde{\mathcal{X}}^{\mu_{M+1},\ldots,\mu_N} \text{ replaces the orbitals } \{\psi_{i,0}\}_{\substack{i=N+1,\ldots,2N\\ i\in\{\mu_{M+1},\ldots,\mu_N\}}} \text{ with the orbitals } \{\psi_{i,K}\}_{\substack{i=N+1,\ldots,2N\\ i\in\{\mu_{M+1},\ldots,\mu_N\}}}.$$

It immediately follows that

$$\mathcal{X}_{\mu,K} = \widetilde{\mathcal{X}}^{\mu_{M+1},\ldots,\mu_N}\mathcal{X}_{\mu,0}\widetilde{\mathcal{X}}_{\mu_1,\ldots,\mu_M}$$

and thus, in order to establish the convergence result (50), we need only prove that

$$\forall \mu \in \mathcal{K}: \qquad \lim_{K\to\infty} \underbrace{\|\mathcal{X}_{\mu,0} - \widetilde{\mathcal{X}}^{\mu_{M+1},\ldots,\mu_N}\mathcal{X}_{\mu,0}\widetilde{\mathcal{X}}_{\mu_1,\ldots,\mu_M}\|_{\widehat{\mathcal{H}}^1\to\widehat{\mathcal{H}}^1}}_{:=(\mathrm{I}*)}. \tag{51}$$

It can now readily be seen that the term (I*) above has the same basic form (albeit much simpler) as the term (IV) appearing in Inequality (38). In particular, the same analysis used to study the term (IV) in **Step Five** can be carried out to prove the convergence result (51). As before, this analysis yields that the convergence result (51) does indeed hold, provided we can establish $\widehat{\mathcal{H}}^1$-boundedness of generalised occupied-to-occupied and virtual-to-virtual excitation operators. Appealing to the demonstration of this continuity result in Appendix C allows us to finally conclude the proof.

□

[3]The remaining case where all indices $\{\mu_i\}_{i=1}^N \subset \{N+1, \ldots 2N\}$ can be dealt with using similar arguments but slightly different notation

An important consequence of Lemma 24 is the following lemma, which will be used to to establish the consistency of the sequence discrete coupled cluster equations (18) in the forthcoming Section 4.3.

Lemma 25 (Approximability of Zeros of Continuous Coupled Cluster Function)**.**

*Let the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ and the sequence of approximation spaces $\{\widetilde{\mathcal{V}}_K\}_{K\geq N}$ be constructed as in Notation 21 and Notation 22 respectively, and assume that **Assumptions A.I-A.III** hold.*

Additionally, for each $K \geq N$, let $\{\Psi_{0,K}\}^\perp$ be the $\widehat{\mathcal{L}}^2$ orthogonal complement of $\Psi_{0,K}$ in $\widehat{\mathcal{L}}^2$, let the infinite-dimensional subspace $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ be defined as $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} := \{\Psi_{0,K}\}^\perp \cap \widehat{\mathcal{H}}^1$, let the coupled cluster function $f_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ be defined through Definition 18, where we adopt the convention of using the subscript K to signify the dependency of this coupled cluster function on $\Psi_{0,K}$, and let $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ denote the ground state zero of the coupled cluster function f_K. Then for every $\epsilon > 0$ there exists $K_0 \geq N$ such that for each $K \geq K_0$ there exists $\Upsilon_K \in \widetilde{\mathrm{V}}_K$ with the property that

$$\|\Upsilon_K - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} < \epsilon.$$

Proof. The proof is essentially a direct consequence of Lemma 24. To see this, recall that $\Psi^*_0 \in \widehat{\mathcal{H}}^1$ denotes the limiting reference determinant defined in **Assumption A.II**. Let $\{\Psi^*_0\}^\perp$ be the $\widehat{\mathcal{L}}^2$ orthogonal complement of Ψ^*_0 in $\widehat{\mathcal{L}}^2$, let the infinite-dimensional subspace $\widehat{\mathcal{H}}^{1,\perp}_{\Psi^*_0}$ be defined as $\widehat{\mathcal{H}}^{1,\perp}_{\Psi^*_0} := \{\Psi^*_0\}^\perp \cap \widehat{\mathcal{H}}^1$, let the coupled cluster function $f \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi^*_0} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi^*_0}\right)^*$ be defined through Definition 18, and let Θ^*_{GS} denote the zero of the coupled cluster function f corresponding to the ground state eigenfunction of the electronic Hamiltonian. Lemma 24 then implies that

$$\lim_{K\to\infty} \|\Theta^*_{K,\mathrm{GS}} - \Theta^*_{\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} = 0. \tag{52}$$

Since $\Theta^*_{\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi^*_0} \subset \widehat{\mathcal{H}}^1$, the definition of the sequence of finite-dimensional approximation spaces $\{\widetilde{\mathcal{V}}_K\}_{K\geq N}$ given in Notation 22 implies the existence of a sequence of functions $\{\overline{\Upsilon}_K\}_{K\geq N}$ with each $\overline{\Upsilon}_K \in \widehat{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}$ such that

$$\lim_{K\to\infty} \|\Theta^*_{\mathrm{GS}} - \overline{\Upsilon}_K\|_{\widehat{\mathcal{H}}^1} = 0. \tag{53}$$

Defining now for every $K \geq N$, the function $\Upsilon_K \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ as $\Upsilon_K := \overline{\Upsilon}_K - (\overline{\Upsilon}_K, \Psi_{0,K})_{\widehat{\mathcal{L}}^2}\Psi_{0,K}$, and using the fact that $\Theta^*_{\mathrm{GS}} \in \{\Psi^*_0\}^\perp$, we deduce that

$$\begin{aligned}
\lim_{K\to\infty} \|\Theta^*_{\mathrm{GS}} - \Upsilon_K\|_{\widehat{\mathcal{H}}^1} &\leq \lim_{K\to\infty} \|\Theta^*_{\mathrm{GS}} - \overline{\Upsilon}_K\|_{\widehat{\mathcal{H}}^1} + \lim_{K\to\infty} \|(\overline{\Upsilon}_K, \Psi_{0,K})_{\widehat{\mathcal{L}}^2}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \\
&= \lim_{K\to\infty} \|(\overline{\Upsilon}_K, \Psi_{0,K})_{\widehat{\mathcal{L}}^2}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \\
&\leq \lim_{K\to\infty} \|(\overline{\Upsilon}_K - \Theta^*_{\mathrm{GS}}, \Psi_{0,K})_{\widehat{\mathcal{L}}^2}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} + \lim_{K\to\infty} \|(\Theta^*_{\mathrm{GS}}, \Psi_{0,K})_{\widehat{\mathcal{L}}^2}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \\
&= \lim_{K\to\infty} \|(\overline{\Upsilon}_K - \Theta^*_{\mathrm{GS}}, \Psi_{0,K})_{\widehat{\mathcal{L}}^2}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} + \lim_{K\to\infty} \|(\Theta^*_{\mathrm{GS}}, \Psi_{0,K} - \Psi^*_0)_{\widehat{\mathcal{L}}^2}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \\
&\leq \lim_{K\to\infty} \|\overline{\Upsilon}_K - \Theta^*_{\mathrm{GS}}\|_{\widehat{\mathcal{L}}^2}\|\Psi_{0,K}\|_{\widehat{\mathcal{L}}^2}\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \\
&\quad + \lim_{K\to\infty} \|\Theta^*_{\mathrm{GS}}\|_{\widehat{\mathcal{L}}^2}\|\Psi_{0,K} - \Psi^*_0\|_{\widehat{\mathcal{L}}^2}\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \\
&= 0,
\end{aligned}$$

where the last step follows from Equation (53) and **Assumption A.II**. Appealing now to Equation (52) completes the proof. □

4.2. **Technical Lemmas Pertaining to Stability.**

Throughout this subsection, we assume the setting of Section 4.1. Our goal now is to identify certain classes of finite-dimensional approximation spaces $\{\widetilde{\mathcal{V}}_K\}_{K\geq N}$, defined as in Notation 22, for which we can reasonably establish a specific stability property of the discrete coupled cluster equations (18). Essentially, we wish to show that the Fréchet derivative $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})$ at the ground state zero $\Theta^*_{K,\mathrm{GS}}$ of the coupled cluster function $f_K\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}})^*$ satisfies a discrete inf-sup condition on $\widetilde{\mathcal{V}}_K$, i.e.,

$$\exists \gamma > 0, \forall K \geq N\colon \inf_{\Upsilon_K \in \widetilde{\mathcal{V}}_K} \sup_{\Phi_K \in \widetilde{\mathcal{V}}_K} \frac{\big\langle \Upsilon_K, f_K(\Theta^*_{K,\mathrm{GS}})\Phi_K \big\rangle_{\widehat{\mathscr{H}}^1 \times \mathscr{H}^{-1}}}{\|\Upsilon_K\|_{\widehat{\mathscr{H}}^1}\|\Phi_K\|_{\widehat{\mathscr{H}}^1}} > \gamma. \tag{54}$$

Unfortunately, establishing a discrete inf-sup condition of the form (54) for arbitrary choices of approximation spaces $\{\widetilde{\mathcal{V}}_K\}_{K\geq N}$ and all $K \geq N$ seems out of reach; indeed, we strongly suspect that the condition (54) does not hold in complete generality without further assumptions. On the other hand, if we impose more structure on the choice of the approximation spaces $\{\widetilde{V}_K\}_{K\geq N}$, and we restrict ourselves to the asymptotic regime, i.e., when K is sufficiently large (meaning that $\widetilde{\mathcal{V}}_K$ is sufficiently rich), then it becomes possible to establish certain results.

In the current article, we will focus on two classes of finite-dimensional approximation spaces, each with a different additional structure (beyond the definition given in Notation 22 and **Assumptions A.I-A.III**). Roughly speaking, the two classes that we consider correspond to the so-called Full-CC approximation spaces and more general discretised CC equations arising from an initial mean-field calculation, and the additional structure that we impose are natural abstractions of the core properties of these two types of coupled cluster approximation spaces (see Appendix B for a detailed discussion of this point).

The remainder of this subsection is organised as follows. We will first specify the two types of additional structure that we impose on the finite-dimensional approximation spaces $\{\widetilde{V}_K\}_{K\geq N}$ and all $K \geq N$; we will refer to these conditions as **Structure Assumption B.I** and **Structure Assumption B.II**. We will then state precisely the discrete inf-sup condition that we require for our analysis and we will show how **Structure Assumption B.I** or **Structure Assumption B.II** allows us to prove this result. A detailed discussion pertaining to our earlier claim that **Structure Assumption B.I** and **B.II** are satisfied by Full-CC approximation spaces and more general discretised CC equations arising from an initial mean-field calculation respectively, is postponed to Appendix B although we briefly address this point following our proofs.

Structure Assumption B.I: We assume that the sequence of approximation spaces $\{\widetilde{\mathcal{V}}_K\}_{K\in\mathbb{N}}$ constructed as in Notation 22are both **excitation and de-excitation complete**, i.e., for any $K \geq N$ and any function $\Upsilon_K \in \widetilde{\mathcal{V}}_K$, the cluster operator $\mathcal{T}(\Upsilon_K)$ generated by $\Upsilon_K \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ has the property that

$$\forall \Phi_K \in \widetilde{\mathcal{V}}_K\colon \qquad \mathcal{T}(\Upsilon_K)\Phi_K \in \widetilde{\mathcal{V}}_K \quad \text{and} \quad \mathcal{T}(\Upsilon_K)^\dagger \Phi_K \in \widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}. \tag{55}$$

Structure Assumption B.II: We assume that the sequence of approximation spaces $\{\widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}\}_{K\in\mathbb{N}}$ constructed as in Notation 22 are **weakly invariant subspaces** of a sequence of symmetric, uniformly bounded linear operators $\{\mathcal{F}_K\}_{K\in\mathbb{N}}$ that satisfy a uniform coercivity condition on a specific subspace. More precisely, we assume that for each $K \geq N$, there exists a symmetric, bounded linear operator $\mathcal{F}_K \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ and a finite-dimensional subspace $\mathcal{W}_K \subset \widehat{\mathcal{H}}^1$ such that

(1) It holds that $\widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\} \subset \mathcal{W}_K$ and for any function $\Upsilon_K \in \widetilde{\mathcal{V}}_K$, the cluster operator $\mathcal{T}(\Upsilon_K)$ generated by $\Upsilon_K \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ has the property that

$$\forall \Phi_K \in \widetilde{\mathcal{V}}_K \colon \qquad \mathcal{T}(\Upsilon_K)\Phi_K, \in \mathcal{W}_K \quad \text{and} \quad \mathcal{T}(\Upsilon_K)^\dagger \Phi_K \in \widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}. \tag{56}$$

(2) The approximation space $\widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}$ is an invariant subspace of the restricted operator $\widetilde{\mathcal{F}}_K \colon \mathcal{W}_K \to \mathcal{W}_K^*$ defined as

$$\forall \Phi_K, \Upsilon_K \in \mathcal{W}_K \colon \quad \langle \Upsilon_K, \widetilde{\mathcal{F}}_K \Phi_K \rangle_{\mathcal{W}_K \times \mathcal{W}_K^*} := \langle \Upsilon_K, \mathcal{F}_K \Phi_K \rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}}. \tag{57}$$

(3) The operator $\mathcal{F}_K$ has continuity constant uniform in K and is coercive on $\mathcal{W}_K$ with coercivity constant independent of K, i.e., $\exists \widetilde{\gamma} > 0$ independent of K such that

$$\forall \Phi_K \in \mathcal{W}_K \colon \quad \left\langle \Phi_K, \mathcal{F}_K \Phi_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}} \geq \widetilde{\gamma} \|\Phi_K\|^2_{\widehat{\mathcal{H}}^1}.$$

(4) For each $K \geq N$ the operator $\mathcal{U}_K := H - \mathcal{F}_K$ is a bounded linear mapping from $\widehat{\mathcal{H}}^1$ to $\widehat{\mathcal{L}}^2$ with continuity constant uniformly bounded in K and small in the following sense: denoting by $(\cdot,\cdot)_{\mathcal{F}_K}$ and $\|\cdot\|_{\mathcal{F}_K}$ the inner product and norm generated by the restricted operator $\widetilde{\mathcal{F}_K} \colon \mathcal{W}_K \to \mathcal{W}_K^*$, by $\mathbb{P}_{\mathcal{F}_K} \colon \mathcal{W}_K \to \mathcal{W}_K$ the $(\cdot,\cdot)_{\mathcal{F}_K}$-orthogonal projection operator onto $\widetilde{\mathcal{V}}_K \oplus \mathrm{span}\{\Psi_{0,K}\}$, by $\widetilde{\mathcal{U}}_K \colon \mathcal{W}_K \to \mathcal{W}_K$ the restriction of the operator $\mathcal{U}_K$, by $\Psi^*_{\mathrm{GS},K} \in \mathcal{W}_K$ the normalised ground state eigenfunction of H in $\mathcal{W}_K$, and by $\Theta^{\Pi}_{K,\mathrm{GS}} \in \widetilde{\mathcal{V}}_K$, the best approximation with respect to the $\widehat{\mathcal{H}}^1$ inner product of the ground state coupled cluster zero $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$, we impose the condition that

$$\begin{aligned}
\left\| (\mathrm{I} - \mathbb{P}_{\mathcal{F}_K})\, \widetilde{\mathcal{U}}_K \mathbb{P}_{\mathcal{F}_K} \right\|_{\mathcal{F}_K \to \widehat{\mathcal{L}}^2} < & \min_{\substack{\Phi_K \in \mathcal{W}_K \\ \Phi_K \in \mathcal{V}_K^\perp}} \frac{|\langle \Phi_K, \mathcal{F}_K \Phi_K \rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}}|^{\frac{1}{2}}}{\|\Phi_K\|_{\widehat{\mathcal{L}}^2}} \\
& \times \min_{\Phi_K \in \{\Psi^*_{\mathrm{GS},K}\}^\perp \cap \mathcal{W}_K} \frac{\langle \Phi_K, (H - \mathcal{E}_{\mathrm{GS}^*})\Phi_K \rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}}}{\|\Phi_K\|^2_{\mathcal{F}_K}} \\
& \times \left\| (\mathrm{I} - \mathbb{P}_{\mathcal{F}_K})\, e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} \mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} \mathbb{P}_{\mathcal{F}_K} \right\|^{-1}_{\mathcal{F}_K \to \mathcal{F}_K}.
\end{aligned}$$

A few comments are now in order. It can readily be seen that **Structure Assumption B.I** is a strong constraint that is not satisfied by the excitation rank-truncated coupled cluster schemes such as CCSD or CCSDT etc. Indeed, in the numerical practice, the only class of coupled cluster approximation spaces that satisfy such a constraint are the Full-CC equations in a finite basis. This shortcoming is addressed by the weaker **Structure Assumption B.II**. We will show later that this alternative assumption is in fact a natural abstraction of combined mean-field excitation-rank truncated coupled cluster equations together with an additional smallness assumption on the operator norm of the mean-field operator. While the smallness assumption is problem-dependent and unlikely to be universally satisfied, preliminary numerical tests (shown in Appendix B) indicate that it is indeed satisfied by small molecules in several cases. We remark that such a splitting of the electronic Hamiltonian using

a mean-field operator together with a smallness assumption has already been considered in earlier contributions [26, 28] but our analysis uses this splitting in a somewhat different manner.

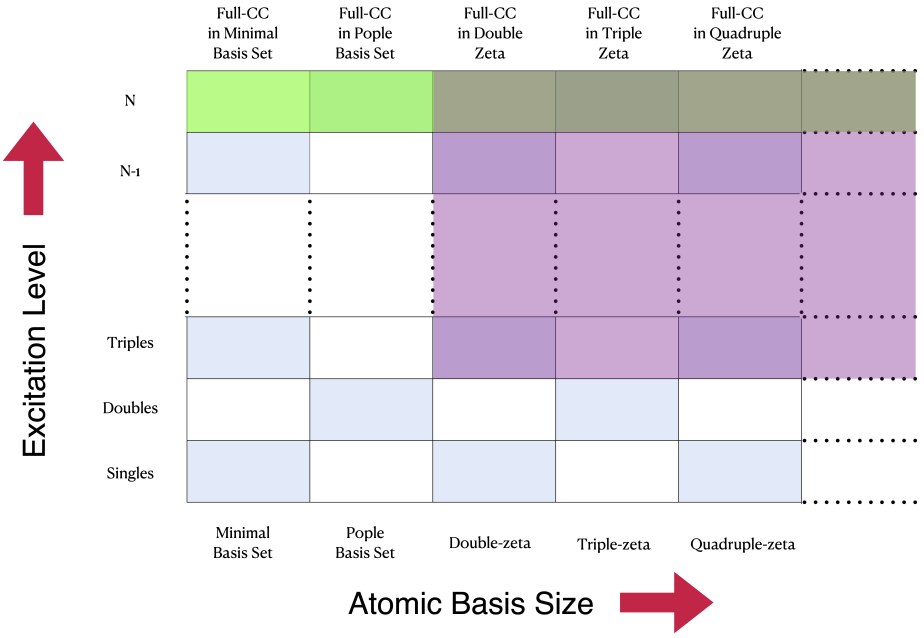

FIGURE 2. Graphical depiction of region of validity of **Structure Assumption B.I** (in green) and **Structure Assumption B.II** (in magenta). Note that the exact size of the magenta region of validity, i.e., whether it begins at the double-zeta level or quadruple-zeta level etc., or whether it includes quadruples or triples etc., depends on the properties of the mean-field operator (c.f., the proof of Lemma 26).

Equipped with **Structure Assumption B.I** and **Structure Assumption B.II**, we are now ready to state the precise discrete inf-sup condition that we wish to prove.

Lemma 26 (Discrete inf-sup condition on Fréchet Derivative of Coupled Cluster Function)**.**

Let the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ *and the sequence of approximation spaces* $\{\widetilde{\mathcal{V}}_K\}_{K\geq N}$ *be constructed as in Notation 21 and Notation 22 respectively, and assume that* ***Assumptions A.I-A.III*** *hold.*

Additionally, for each $K \geq N$*, let* $\{\Psi_{0,K}\}^{\perp}$ *be the* $\widehat{\mathcal{L}}^2$ *orthogonal complement of* $\Psi_{0,K}$ *in* $\widehat{\mathcal{L}}^2$*, let the infinite-dimensional subspace* $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ *be defined as* $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} := \{\Psi_{0,K}\}^{\perp} \cap \widehat{\mathcal{H}}^1$*, let the coupled cluster function* $f_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ *be defined through Definition 18, where we adopt the convention of using the subscript* K *to signify the dependency of this coupled cluster function on* $\Psi_{0,K}$*, let* $\Theta^*_{K,\mathrm{GS}}$ *denote the zero of the coupled cluster function* f_K *corresponding to the* $\widehat{\mathcal{L}}^2$*-normalised ground state eigenfunction* Ψ^*_{GS} *of the electronic Hamiltonian, i.e.,*

$$e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K} = \frac{1}{(\Psi^*_{\mathrm{GS}}, \Psi_{0,K})_{\widehat{\mathcal{L}}^2}}\Psi^*_{\mathrm{GS}},$$

and let $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ *denote the Fréchet derivative of the coupled cluster function* f_K *evaluated at* $\Theta^*_{K,\mathrm{GS}}$.

Under the constraint that either ***Structure Assumption B.I*** *or* ***Structure Assumption B.II*** *holds, we have the existence of a* $K_0 \geq N$ *and a constant* $\gamma > 0$ *such that*

$$\forall K \geq K_0\colon \inf_{\Upsilon_K \in \widetilde{\mathcal{V}}_K} \sup_{\Phi_K \in \widetilde{\mathcal{V}}_K} \frac{\left\langle \Upsilon_K, \mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\Phi_K \right\rangle_{\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}}{\|\Upsilon_K\|_{\widehat{\mathscr{H}}^1}\|\Phi_K\|_{\widehat{\mathscr{H}}^1}} > \gamma. \tag{58}$$

We will provide separate proofs for each of the two structural assumptions. Before proceeding to these proofs however, let us first demonstrate a simple, auxiliary result which will considerably simplify our arguments.

Lemma 27 (Auxiliary result to simplify the proof of Lemma 26)**.**

Consider the setting and statement of Lemma 26, let $\mathcal{E}^*_{\mathrm{GS}}$ *denote the ground state energy of the electronic Hamiltonian* $H\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$*, and for all* $K \geq N$ *let* $\Theta^{\Pi}_{K,\mathrm{GS}} \in \widetilde{\mathcal{V}}_K$ *denote the best approximation with respect to the* $\widehat{\mathscr{H}}^1$ *inner product of the exact ground state zero* $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ *of the coupled cluster function* $f_K\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$*. Then the discrete inf-sup condition* (59) *holds if there exists* $\widehat{K}_0 \geq N$ *and a constant* $\widehat{\gamma} > 0$ *such that*

$$\forall K \geq \widehat{K}_0\colon \inf_{\Upsilon_K \in \widetilde{\mathcal{V}}_K} \sup_{\Phi_K \in \widetilde{\mathcal{V}}_K} \frac{\left\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H - \mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\Phi_K \right\rangle_{\widehat{\mathscr{H}}^1 \times \mathscr{H}^{-1}}}{\|\Upsilon_K\|_{\widehat{\mathscr{H}}^1}\|\Phi_K\|_{\widehat{\mathscr{H}}^1}} > \widehat{\gamma}. \tag{59}$$

Proof. Let $K \geq N$ be fixed. It has been demonstrated in detail in [11, Corollary 27] that the Fréchet derivative $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ of the coupled cluster function $f_K\colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ evaluated at $\Theta^*_{K,\mathrm{GS}}$ has the expression

$$\forall \Upsilon, \Phi \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\colon \quad \left\langle \Upsilon, \mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\Phi \right\rangle_{\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} = \langle \Upsilon, e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}(H - \mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Phi \rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}}.$$

On the other hand, Lemma 24 implies that the best approximation $\Theta^{\Pi}_{K,\mathrm{GS}} \in \widetilde{\mathcal{V}}_K$ of the exact ground state zero $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ satisfies

$$\lim_{K \to \infty} \|\Theta^{\Pi}_{K,\mathrm{GS}} - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1} = 0.$$

Recalling therefore from Theorem 16 that the mapping $\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \ni \Theta \mapsto \mathcal{T}(\Theta)$ is bounded with continuity constant depending only on N and $\|\Psi_{0,K}\|_{\widehat{\mathscr{H}}^1}$, and using the fact that the exponential is a $\mathscr{C}^\infty$ mapping on the algebra of cluster operators (see Theorem 17) yields the required result. □

In view of Lemma 27, it suffices to establish the validity of the estimate (59) in order to prove the discrete inf-sup Lemma 26. The primary motivation for proceeding in this fashion is that the underlying exponential cluster operators appearing in Estimate (59) are constructed from functions in the finite-dimensional approximation spaces $\{\widetilde{\mathcal{V}}_K\}_{K \geq N}$. Consequently, we can immediately take advantage of the relations (55) and (56) appearing in **Structure Assumption B.I** and **Structure Assumption B.II** respectively.

We begin with the simpler case of **Structure Assumption B.I**. Since this proof closely resembles that of [11, Theorem 42], we shall be succinct in our arguments.

*Proof of Lemma 26 for **Structure Assumption B.I**.*

In view of Lemma 27, it suffices to establish the validity of the estimate (59). To this end, let $K \geq N$ be fixed, let $\mathbb{P}_{0,K} \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ denote the $\widehat{\mathcal{L}}^2$-orthogonal projector onto $\operatorname{span}\{\Psi_{0,K}\}$ and let $\mathbb{P}_{0,K}^{\perp} \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ denote its complement, i.e., $\mathbb{P}_{0,K}^{\perp} := \mathrm{I} - \mathbb{P}_{0,K}$.

Then for any arbitrary $\Upsilon_K \in \widetilde{\mathcal{V}}_K \subset \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$, we may define $\Phi_K \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ as

$$\Phi_K := \mathbb{P}_{0,K}^{\perp} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K, \tag{60}$$

where, as in Lemma 27, $\Theta^{\Pi}_{K,\mathrm{GS}} \in \widetilde{\mathcal{V}}_K$ denotes the best approximation with respect to the $\widehat{\mathcal{H}}^1$ inner product of the exact ground state zero $\Theta^{*}_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ of the coupled cluster function $f_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$.

We claim that in fact $\Phi_K \in \widetilde{\mathcal{V}}_K$. Indeed, it follows directly from **Structure Assumption B.I** that $e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K \in \widetilde{\mathcal{V}}_K \oplus \operatorname{span}\{\Psi_{0,K}\}$. Using therefore the fact that $\widetilde{\mathcal{V}}_K$ is $\widehat{\mathcal{L}}^2$-orthogonal to $\operatorname{span}\{\Psi_{0,K}\}$ by construction (recall Notation 22) proves the claim.

We can thus deduce that (c.f., Estimate (59))

$$\begin{aligned}
&\left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H - \mathcal{E}^{*}_{\mathrm{GS}}) e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} \Phi_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \mathcal{H}^{-1}} \\
&= \left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H - \mathcal{E}^{*}_{\mathrm{GS}}) e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} \mathbb{P}^{\perp}_{0,K} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \mathcal{H}^{-1}} \\
&= \underbrace{\left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H - \mathcal{E}^{*}_{\mathrm{GS}}) e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \mathcal{H}^{-1}}}_{:=(\mathrm{I})} \\
&\underbrace{- \left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H - \mathcal{E}^{*}_{\mathrm{GS}}) e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} \mathbb{P}_{0,K} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \mathcal{H}^{-1}}}_{:=(\mathrm{II})},
\end{aligned} \tag{61}$$

where we remind the reader that, as in Lemma 27, $\mathcal{E}^{*}_{\mathrm{GS}}$ denotes the ground state eigenvalue of the electronic Hamiltonian $H \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$.

Let us first consider the term (II). A straightforward calculation reveals that

$$\begin{aligned}
(\mathrm{II}) &= \left(\Psi_{0,K}, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K \right)_{\widehat{\mathcal{L}}^2} \left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H - \mathcal{E}^{*}_{\mathrm{GS}}) e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} \Psi_{0,K} \right\rangle_{\widehat{\mathcal{H}}^1 \times \mathcal{H}^{-1}} \\
&\leq \|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \|e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1} \|e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K\|^2_{\widehat{\mathcal{H}}^1} \|(H - \mathcal{E}^{*}_{\mathrm{GS}}) e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} \Psi_{0,K}\|_{\widehat{\mathcal{H}}^{-1}}.
\end{aligned} \tag{62}$$

Note that since $\Theta^{*}_{K,\mathrm{GS}}$ is by definition the ground state zero of the coupled cluster function $f_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$, it follows that $H e^{\mathcal{J}(\Theta^{*}_{K,\mathrm{GS}})} \Psi_{0,K} = \mathcal{E}^{*}_{\mathrm{GS}} e^{\mathcal{J}(\Theta^{*}_{K,\mathrm{GS}})} \Psi_{0,K}$ (see, e.g., [25, Theorem 5.3]) and we

therefore have

$$\begin{aligned}\|(H-\mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^{-1}} \leq & \|(H-\mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K}\|_{\widehat{\mathcal{H}}^{-1}} \\ & + \|(H-\mathcal{E}^*_{\mathrm{GS}})\left(e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}-e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\right)\Psi_{0,K}\|_{\widehat{\mathcal{H}}^{-1}} \\ = & \|(H-\mathcal{E}^*_{\mathrm{GS}})\left(e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}-e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\right)\Psi_{0,K}\|_{\widehat{\mathcal{H}}^{-1}}. \end{aligned} \tag{63}$$

On the one hand, Lemma 24 implies that the best approximation $\Theta^{\Pi}_{K,\mathrm{GS}} \in \widetilde{\mathcal{V}}_K$ of the exact ground state zero $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ satisfies

$$\lim_{K\to\infty} \|\Theta^{\Pi}_{K,\mathrm{GS}} - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} = 0.$$

On the other hand, we can recall from Theorem 16 that the mapping $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \ni \Theta \mapsto \mathcal{T}(\Theta)$ is bounded with continuity constant depending only on N and $\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}$. Since the exponential is a $\mathscr{C}^\infty$ mapping on the algebra of cluster operators (see Theorem 17), we can take advantage of **Assumption A.II** and thus deduce from Inequalities (62) and (63) that there exists a constant $\epsilon^{(\mathrm{II})}_K \geq 0$ such that

$$(\mathrm{II}) \leq \epsilon^{(\mathrm{II})}_K \|e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\|^2_{\widehat{\mathcal{H}}^1} \quad \text{and} \quad \lim_{K\to\infty}\epsilon^{(\mathrm{II})}_K = 0. \tag{64}$$

It remains to estimate the term (I). To do so, let us first write

$$(\mathrm{I}) = \left\langle e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K, (H-\mathcal{E}^*_{\mathrm{GS}})e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}}.$$

Notice that thanks to **Assumption A.III**, the shifted Hamiltonian $H-\mathcal{E}^*_{\mathrm{GS}}\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ is coercive on $\{\Psi^*_{\mathrm{GS}}\}^\perp \cap \widehat{\mathcal{H}}^1$ with a coercivity constant that we denote by $\gamma^*_{\mathrm{GS}} > 0$, where we remind the reader that $\Psi^*_{\mathrm{GS}} \in \widehat{\mathcal{H}}^1$ denotes the $\widehat{\mathcal{L}}^2$-normalised ground state eigenfunction of the electronic Hamiltonian H.

In order to take advantage of this coercivity result, let us first observe that the function $e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})^\dagger}\Upsilon_K \in \widetilde{\mathcal{V}}_K$ is $\widehat{\mathcal{L}}^2$-orthogonal to the ground state Ψ^*_{GS}. Indeed, since the ground state zero $\Theta^*_{K,\mathrm{GS}}$ of the coupled cluster function f_K satisfies by definition that $e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K} = \frac{1}{(\Psi_{0,K},\Psi^*_{\mathrm{GS}})_{\widehat{\mathcal{L}}^2}}\Psi^*_{\mathrm{GS}}$, we deduce that

$$\frac{1}{(\Psi_{0,K},\Psi^*_{\mathrm{GS}})_{\widehat{\mathcal{L}}^2}}\left(e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})^\dagger}\Upsilon_K, \Psi^*_{\mathrm{GS}}\right)_{\widehat{\mathcal{L}}^2} = \left(e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})^\dagger}\Upsilon_K, e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K}\right)_{\widehat{\mathcal{L}}^2} = (\Upsilon_K,\Psi_{0,K})_{\widehat{\mathcal{L}}^2} = 0,$$

where the last step follows from the fact that $\Upsilon_K \in \widetilde{\mathcal{V}}_K$, which is by definition $\widehat{\mathcal{L}}^2$-orthogonal to $\Psi_{0,K}$ (see Notation 22).

Consequently, using again the fact that $\lim_{K\to\infty}\|\Theta^{\Pi}_{K,\mathrm{GS}} - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} = 0$ (see Lemma 24), we deduce the existence of a constant $\gamma^{(K)}_{\mathrm{GS}} > 0$ such that

$$(\mathrm{I}) \geq \gamma^{(K)}_{\mathrm{GS}} \|e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\|^2_{\widehat{\mathcal{H}}^1} \quad \text{and} \quad \lim_{K\to\infty}\gamma^{(K)}_{\mathrm{GS}} = \gamma^*_{\mathrm{GS}}. \tag{65}$$

Combining now Inequalities (64) and (65) and recalling our choice of test function Φ_K given by Equation (60), we see that we have shown the existence of constant $\widetilde{\gamma} > 0$ and $\widehat{K}_0$ such that for all

$K \geq \widehat{K}_0$ and all $\Upsilon_K \in \widetilde{\mathcal{V}}_K$ it holds that

$$\begin{aligned}
&\sup_{\Phi_K \in \widetilde{\mathcal{V}}_K} \frac{\left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H - \mathcal{E}^*_{\mathrm{GS}}) e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} \Phi_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \mathcal{H}^{-1}}}{\|\Upsilon_K\|_{\widehat{\mathcal{H}}^1} \|\Phi_K\|_{\widehat{\mathcal{H}}^1}} \\
&\qquad \geq \widetilde{\gamma} \frac{\|e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger} \Upsilon_K\|^2_{\widehat{\mathcal{H}}^1}}{\|\Upsilon_K\|_{\widehat{\mathcal{H}}^1} \|\mathbb{P}^{\perp}_{0,K} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger} \Upsilon_K\|_{\widehat{\mathcal{H}}^1}} \\
&\qquad \geq \widetilde{\gamma} \frac{1}{\|e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1} \|\mathbb{P}^{\perp}_{0,K} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1}} \frac{\|e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger} \Upsilon_K\|_{\widehat{\mathcal{H}}^1} \|\Upsilon_K\|_{\widehat{\mathcal{H}}^1}}{\|\Upsilon_K\|_{\widehat{\mathcal{H}}^1} \|e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger} \Upsilon_K\|_{\widehat{\mathcal{H}}^1}} \\
&\qquad = \widetilde{\gamma} \frac{1}{\|e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1} \|\mathbb{P}^{\perp}_{0,K} e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\|_{\widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1}}.
\end{aligned}$$

Taking now the infimum over all $\Upsilon_K \in \widetilde{\mathcal{V}}_K$ therefore completes the proof.

□

Remark 28 (Method of proof of Lemma 26 for **Structure Assumption B.I**).

A few comments on the above proof are now in order. First, notice that the main reason for introducing ***Structure Assumption B.I*** *was to allow us to construct an appropriate test function belonging to the finite-dimensional approximation space $\widetilde{\mathcal{V}}_K$. Second, many of our subsequent arguments relied on the approximability of the exact ground state zero $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ of the coupled cluster function in the approximation space $\widetilde{\mathcal{V}}_K$; indeed, the same approximability property was used in the proof of Lemma 27. Finally, we used the coercivity of the shifted electronic Hamiltonian $H - \mathcal{E}^*_{\mathrm{FCI}}$ on $\{\Psi^*_{\mathrm{GS}}\}^\perp \cap \widehat{\mathcal{H}}^1$ where Ψ^*_{GS} denotes the $\widehat{\mathcal{L}}^2$-normalised ground state eigenfunction of $H \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ with corresponding eigenvalue $\mathcal{E}^*_{\mathrm{GS}}$. In particular, the question of whether the discrete inf-sup condition* (59) *holds for a particular choice of finite-dimensional approximation space $\widetilde{\mathcal{V}}_K$ depends on how well the exact ground state coupled cluster zero $\Theta^*_{K,\mathrm{GS}}$ is approximated in $\widetilde{\mathcal{V}}_K$. Finally, let us point out that the numerator $\widetilde{\gamma} > 0$ in the discrete inf-sup constant that we derived has the property that $\lim_{K\to\infty} \widetilde{\gamma} = \gamma^*_{\mathrm{GS}}$, where γ^*_{GS} denotes the coercivity constant of the shifted Hamiltonian $H - \mathcal{E}^*_{\mathrm{GS}} \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ on $\{\Psi^*_{\mathrm{GS}}\}^\perp \cap \widehat{\mathcal{H}}^1$.*

As mentioned previously, the so-called Full-CC approximation spaces satisfy **Structure Assumption B.I**. This is the subject of further discussion in Appendix B.

We next turn our attention to the proof of Lemma 26 in the case of our second structural assumption. This proof is considerably more technical.

Proof of Lemma 26 for ***Structure Assumption B.II****.*

As before, we will make use of Lemma 27 and establish the validity of the estimate (59). To this end, for any $K \geq N$ we denote by $\mathbb{P}_{0,K} \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ the $\widehat{\mathcal{L}}^2$-orthogonal projector onto $\mathrm{span}\{\Psi_{0,K}\}$ and by $\mathbb{P}^{\perp}_{0,K} \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^1$ its complement, i.e., $\mathbb{P}^{\perp}_{0,K} := \mathrm{I} - \mathbb{P}_{0,K}$. Additionally, for each $K \geq N$, let us denote $\mathcal{V}_K := \widetilde{V}_K \oplus \mathrm{span}\{\Psi_{0,K}\}$. This notation will be useful for clarity of exposition in the forthcoming proof.

We begin our proof by demonstrating two auxiliary results that follow directly from **Structure Assumption B.II**. The first of these minor results is the observation that for each $K \geq N$, the mean-field operator $\mathcal{F}_K \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ induces an inner-product, denoted $(\cdot,\cdot)_{\mathcal{F}_K}$, on the finite-dimensional space $\mathcal{W}_K \subset \widehat{\mathcal{H}}^1$ through the relation

$$\forall K \geq N,\ \forall \Upsilon_K, \Phi_K \in \mathcal{W}_K \colon \qquad (\Upsilon_K, \Phi_K)_{\mathcal{F}_K} := \left\langle \Upsilon_K, \mathcal{F}_K \Phi_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}}. \tag{66}$$

Indeed, this is simply a consequence of the symmetry of $\mathcal{F}_K$ and its coercivity on $\mathcal{W}_K$. More importantly however, the fact that both the continuity constant and coercivity constant of $\mathcal{F}_K$ are uniformly bounded in K immediately implies that the norm $\|\cdot\|_{\mathcal{F}_K}$ induced by this new inner-product $(\cdot,\cdot)_{\mathcal{F}_K}$ on the finite-dimensional space $\mathcal{W}_K$ is equivalent to the usual $\widehat{\mathcal{H}}^1$ norm with equivalence constants independent of K.

Consider now, for any $K \geq N$, the inner-product $(\cdot,\cdot)_{\mathcal{F}_K} \colon \mathcal{W}_K \times \mathcal{W}_K \to \mathbb{R}$ and recall that the approximation space $\mathcal{V}_K \subset \mathcal{W}_K$ by assumption. Let us denote by $\mathbb{P}_{\mathcal{F}_K} \colon \mathcal{W}_K \to \mathcal{W}_K$, the $(\cdot,\cdot)_{\mathcal{F}_K}$-orthogonal projection operator onto $\mathcal{V}_K$ and let us denote by $\mathbb{P}^{\perp}_{\mathcal{F}_K} \colon \mathcal{W}_K \to \mathcal{W}_K$, its complement, i.e., $\mathbb{P}^{\perp}_{\mathcal{F}_K} = \mathrm{I} - \mathbb{P}_{\mathcal{F}_K}$. We now claim that these projection operators $\mathbb{P}_{\mathcal{F}_K}$ and $\mathbb{P}^{\perp}_{\mathcal{F}_K}$ are, in fact, $\widehat{\mathcal{L}}^2$-orthogonal.

To see this, recall once again from **Structural Assumption B.II** that the approximation space $\mathcal{V}_K$ is an invariant subspace of the restricted operator $\widetilde{\mathcal{F}}_K \colon \mathcal{W}_K \to \mathcal{W}_K^*$ defined through Equation (57). Since this restricted operator $\widetilde{\mathcal{F}}_K$ is by definition symmetric, it follows that $\mathcal{W}_K$ has an $\widehat{\mathcal{L}}^2$-orthonormal basis consisting of eigenvectors of $\widetilde{\mathcal{F}}_K$ and that $\mathcal{V}_K$ has an $\widehat{\mathcal{L}}^2$-orthonormal basis consisting of *a subset* of these eigenvectors. The claim now readily follows.

Equipped with the above two auxiliary results, let us now return to our goal of showing the validity of Estimate (59) from Lemma 27. For any arbitrary $\Upsilon_K \in \widetilde{\mathcal{V}}_K \subset \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$, let us now choose the test function $\Phi_K \in \widetilde{\mathcal{V}}_K$ as

$$\Phi_K := \mathbb{P}^{\perp}_{0,K} \mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K. \tag{67}$$

where, as in Lemma 27, $\Theta^{\Pi}_{K,\mathrm{GS}} \in \widetilde{\mathcal{V}}_K$ denotes the best approximation with respect to the $\widehat{\mathcal{H}}^1$ inner product of the exact ground state zero $\Theta^{*}_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ of the coupled cluster function $f_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$.

Note that $\Phi_K \in \widetilde{\mathcal{V}}_K$ is well-defined since by **Structural Assumption B.II**, the function

$$e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}} \Upsilon_K \in \mathcal{W}_K,$$

and applying the projection operator $\mathbb{P}_{\mathcal{F}_K}$ yields an element of $\mathcal{V}_K$.

It follows that for any arbitrary $\Upsilon_K \in \widetilde{\mathcal{V}}_K$, with the above choice of test function $\Phi_K \in \widetilde{\mathcal{V}}_K$, we have that (c.f., Estimate (59))

$$
\begin{aligned}
&\left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\Phi_K \right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}} \\
&= \left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}^{\perp}_{0,K}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}}\Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}} \\
&= \left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}}\Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}} \\
&- \left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{0,K}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}}\Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}} \\
&= \underbrace{\left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}}\Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}}}_{:=(\mathrm{I}*)} \\
&+ \underbrace{\left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})\mathbb{P}^{\perp}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}}\Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}}}_{:=(\mathrm{II}*)} \\
&- \underbrace{\left\langle \Upsilon_K, e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{0,K}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}}\Upsilon_K \right\rangle_{\widehat{\mathcal{H}}^1\times\mathcal{H}^{-1}}}_{:=(\mathrm{III}*)},
\end{aligned}
\tag{68}
$$

where we remind the reader that, as in Lemma 27, $\mathcal{E}^*_{\mathrm{GS}}$ denotes the ground state eigenvalue of the electronic Hamiltonian $H\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$.

Intuitively, the key idea of the remainder of this proof is to argue that the term (I*) is positive while the terms (II*) and (III*) are 'small'– at least asymptotically. We proceed one term at a time.

Estimation of the term (I*)

We begin by considering the term (I*). As the first step, we claim that

$$
\begin{aligned}
&\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}\colon \mathcal{V}_K \to \mathcal{V}_K \ \text{ is an invertible mapping with} \\
&\left(\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}\right)^{-1} = \mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}.
\end{aligned}
\tag{69}
$$

Indeed, a direct calculation shows that

$$
\begin{aligned}
\left(\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}\right)\left(\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}\right) &= \mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} \\
&= \mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} \\
&- \underbrace{\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}^{\perp}_{\mathcal{F}_K}e^{-\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}}_{:=(\mathrm{IA}*)}.
\end{aligned}
\tag{70}
$$

We claim that the term (IA*) is zero. Indeed, suppose there exists $\widehat{\Phi}_K \in \mathcal{V}_K$ such that

$$
\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}^{\perp}_{\mathcal{F}_K}e^{\mathcal{J}(-\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}\widehat{\Phi}_K = \mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}^{\perp}_{\mathcal{F}_K}e^{\mathcal{J}(-\Theta^{\Pi}_{K,\mathrm{GS}})}\widehat{\Phi}_K \neq 0.
$$

It follows that there must exist some $\widehat{\Upsilon}_K \in \mathcal{V}_K$ such that

$$
0 \neq \left(\widehat{\Upsilon}_K, \mathbb{P}_{\mathcal{F}_K}e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}^{\perp}_{\mathcal{F}_K}e^{\mathcal{J}(-\Theta^{\Pi}_{K,\mathrm{GS}})}\widehat{\Phi}_K\right)_{\widehat{\mathcal{L}}^2} = \left(e^{\mathcal{J}(\Theta^{\Pi}_{K,\mathrm{GS}})^{\dagger}}\widehat{\Upsilon}_K, \mathbb{P}^{\perp}_{\mathcal{F}_K}e^{\mathcal{J}(-\Theta^{\Pi}_{K,\mathrm{GS}})}\widehat{\Phi}_K\right)_{\widehat{\mathcal{L}}^2}.
$$

But thanks to **Structure Assumption B.II**, it holds that $e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\widehat{\Upsilon}_K \in \mathcal{V}_K$. And since, the projection operator $\mathbb{P}^{\perp}_{\mathcal{F}_K}$ is $\widehat{\mathcal{L}}^2$-orthogonal, the inner product on the right must be zero. Recalling Equation (70), we see that our claim concerning the invertibility of the projected exponential cluster operator readily follows.

Let us now return to our earlier aim of estimating the term (I*) appearing in Equation (68). Using the invertibility result (69) and taking advantage of the fact that $e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K \in \mathcal{V}_K$, which is again a consequence of **Structure Assumption B.II**, we deduce that

$$\begin{aligned}
(\mathrm{I}*) &= \Big\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\Big\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}} \\
&= \Big\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})\mathbb{P}_{\mathcal{F}_K}e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\Big\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}} \\
&= \Big\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})\mathbb{P}_{\mathcal{F}_K}e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\Big\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}} \\
&= \Big\langle e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K, (H-\mathcal{E}^*_{\mathrm{GS}})e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\Big\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}}.
\end{aligned}$$

Next, let $(\mathcal{E}^*_{\mathrm{GS},K}, \Psi^*_{\mathrm{GS},K})$ denote a ground state (normalised) eigenpair of the electronic Hamiltonian in the finite-dimensional approximation space $\mathcal{W}_K$. Recall that the approximation spaces $\{\mathcal{V}_K\}_{K\geq N}$ have a dense union in $\widehat{\mathscr{H}}^1$ by definition (see Notation 22), and that **Structure Assumption B.II** imposes that for each $K \geq N$, we have $\mathcal{V}_K \subset \mathcal{W}_K$. We can therefore conclude that the sequence of finite dimensional spaces $\{\mathcal{W}_K\}_{K\geq N}$ also posses a dense union in $\widehat{\mathscr{H}}^1$. Consequently, by well-known approximability results for linear eigenvalue problems (see, e.g., [5, Example 5.9]), the eigenvalue $\mathcal{E}^*_{\mathrm{GS},K}$ must be simple for sufficiently large $K \geq N$, and with the correct choice of normalisation, $\Psi^*_{\mathrm{GS},K} \to \Psi^*_{\mathrm{GS}}$ as $K \to \infty$ in the $\widehat{\mathscr{H}}^1$ sense. This implies in particular that the shifted Hamiltonian $H - \mathcal{E}^*_{\mathrm{GS}}$ is coercive on $\{\Psi^*_{\mathrm{GS},K}\}^\perp \cap \mathcal{W}_K$ for all K sufficiently large with a coercivity constant uniformly bounded in K. Furthermore, recalling the convergence result from Lemma 24, we also have that

$$\lim_{K\to\infty}\left\| \frac{1}{\big(\Psi^*_{\mathrm{GS},K}, \Psi_{0,K}\big)_{\widehat{\mathcal{L}}}}\Psi^*_{\mathrm{GS},K} - e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\Psi_{0,K}\right\|_{\widehat{\mathcal{L}}^2} = 0.$$

Using now the fact that the test function $e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K$ appearing in the simplification of the term (I*) is $\widehat{\mathcal{L}}^2$-orthogonal to the function $e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\Psi_{0,K}$, and denoting by $\Gamma^*_{\mathrm{GS}} > 0$ the uniformly bounded coercivity constant $H - \mathcal{E}^*_{\mathrm{GS}}$ with respect to the equivalent $\|\cdot\|_{\mathcal{F}_K}$ norm, we deduce for all $K \geq N$ sufficiently large, the existence of a constant $\Gamma^{\mathcal{F}_K}_{\mathrm{GS}} > 0$ such that

$$(\mathrm{I}*) \geq \Gamma^{\mathcal{F}_K}_{\mathrm{GS}}\|e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\|^2_{\mathcal{F}_K} \quad \text{and} \quad \lim_{K\to\infty}\Gamma^{\mathcal{F}_K}_{\mathrm{GS}} = \Gamma^*_{\mathrm{GS}}. \tag{71}$$

Estimation of the term (II*)

We now turn our attention to the term (II*). We begin by using once again **Structure Assumption B.II** and employing the splitting of the electronic Hamiltonian H given by

$$\forall K \geq N\colon \quad H = \mathcal{F}_K + \mathcal{U}_K \quad \text{where } \mathcal{U}_K := H - \mathcal{F}_K\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathcal{L}}^2 \text{ is a bounded linear operator.}$$

Employing this splitting allows us to write the term (II*) as

$$\begin{aligned}(\mathrm{II}*) &= \left\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}(H-\mathcal{E}^*_{\mathrm{GS}})\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K \right\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}} \\ &= \left\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}(\mathcal{F}_K-\mathcal{E}^*_{\mathrm{GS}})\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K \right\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}} \\ &+ \left\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathcal{U}_K\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K \right\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}} \\ &= \left\langle \Upsilon_K, e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathcal{U}_K\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K \right\rangle_{\widehat{\mathscr{H}}^1\times\mathscr{H}^{-1}}.\end{aligned}$$

Here, the last step follows from the fact that by **Structure Assumption B.II**, $e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K \in \mathcal{V}_K$, and $\mathcal{V}_K$ is an invariant subspace of the restricted mean-field operator $\widetilde{\mathcal{F}}_K\colon \mathcal{W}_K \to \mathcal{W}_K^*$ while $\mathbb{P}_{\mathcal{F}_K}$ is the $(\cdot,\cdot)_{\mathcal{F}_K}$-orthogonal projection operator onto $\mathcal{V}_K$ which also happens to be $\widehat{\mathcal{L}}^2$-orthogonal.

A direct calculation now yields

$$\begin{aligned}(\mathrm{II}*) &\geq -\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\mathcal{U}_K e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\right\|_{\widehat{\mathcal{L}}^2}\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\right\|_{\widehat{\mathcal{L}}^2} \\ &\geq -\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\widetilde{\mathcal{U}}_K\mathbb{P}_{\mathcal{F}_K}\right\|_{\mathcal{F}_K\to\widehat{\mathcal{L}}^2}\left\|e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\right\|_{\mathcal{F}_K}\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\right\|_{\widehat{\mathcal{L}}^2},\end{aligned}$$

where $\widetilde{\mathcal{U}}_K\colon \mathcal{W}_K \to \mathcal{W}_K$ denotes the restriction of the formally infinite-dimensional operator $\mathcal{U}_K\colon \widehat{\mathscr{H}}_1 \to \widehat{\mathcal{L}}_2$ to the finite-dimensional space $\mathcal{W}_K$, and we have denoted by $\|\cdot\|_{\mathcal{F}_K\to\widehat{\mathcal{L}}^2}$, the operator norm on the space of bounded linear operators acting from $\mathcal{W}_K$, equipped with the $\widehat{\mathscr{H}}^1$-equivalent $\|\cdot\|_{\mathcal{F}_K}$ norm to the space $\mathcal{W}_K$ equipped with the usual $\|\cdot\|_{\widehat{\mathcal{L}}^2}$ norm. Notice that this is permitted since the function $e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K \in \mathcal{W}_K$ for all $K \geq N$.

To proceed, let us consider an arbitrary element $\widehat{\Phi}_K \in \mathcal{W}_K$ for any $K \geq N$. It is easy to see that we can write

$$\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\widehat{\Phi}_K\right\|^2_{\widehat{\mathcal{L}}^2} \leq \frac{1}{\Lambda^{\mathcal{F}_K}_{\min}}\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\widehat{\Phi}_K\right\|^2_{\mathcal{F}_K}, \tag{72}$$

where $\Lambda^{\mathcal{F}_K}_{\min} > 0$ denotes the minimum of the spectrum of the restricted mean-field operator $\widetilde{\mathcal{F}}_K$ in the subspace $\operatorname{ran}\mathbb{P}^{\perp}_{\mathcal{F}_K}$, which can be expressed as

$$\Lambda^{\mathcal{F}_K}_{\min} = \min_{\substack{\Phi_K\in\mathcal{W}_K\\ \Phi_K\in\mathcal{V}_K^{\perp}}} \frac{\langle\Phi_K, \mathcal{F}_K\Phi_K\rangle_{\widehat{\mathscr{H}}^1\times\widehat{\mathscr{H}}^{-1}}}{\|\Phi_K\|^2_{\widehat{\mathcal{L}}^2}}, \quad \text{and we have that} \quad \mathcal{W}_K = \underbrace{\operatorname{ran}\mathbb{P}_{\mathcal{F}_K}}_{:=\mathcal{V}_K} \oplus \operatorname{ran}\mathbb{P}^{\perp}_{\mathcal{F}_K}.$$

Returning now to the term (II*) and using Inequality (72) together with some elementary simplifications, we deduce that for all $K \geq N$ it holds that

$$(\mathrm{II}*) \geq -\frac{\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\widetilde{\mathcal{U}}_K\mathbb{P}_{\mathcal{F}_K}\right\|_{\mathcal{F}_K\to\widehat{\mathcal{L}}^2}\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}\right\|_{\mathcal{F}_K\to\mathcal{F}_K}}{\sqrt{\Lambda^{\mathcal{F}_K}_{\min}}}\left\|e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})^\dagger}\Upsilon_K\right\|^2_{\mathcal{F}_K}. \tag{73}$$

Estimation of the term (III*)

It remains to estimate the final term (III*). To do so, we will recall the estimation of the very similar term (II) appearing in Equation (61) from the earlier proof of Lemma 26 under **Structure**

Assumption B.I. Following exactly the same arguments as before (which we do not repeat for the sake of brevity), we deduce for every $K \geq N$, the existence of a constant $\epsilon_K^{(\mathrm{III})}$ such that

$$(\mathrm{III}*) \geq -\epsilon_K^{(\mathrm{III})} \left\| e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K \right\|_{\mathcal{F}_K}^2 \quad \text{and} \quad \lim_{K\to\infty} \epsilon_K^{(\mathrm{III})} = 0. \tag{74}$$

Combining now Equation (68) with Inequalities (71), (73), and (74), recalling our choice of test function Φ_K given by Equation (67), and making use of the smallness assumption on the continuity constant of $\mathcal{U}_K$ from **Structure Assumption B.II**, we see that we have shown the existence of a constant $\widetilde{\gamma} > 0$ and $\widehat{K}_0$ such that for all $K \geq \widehat{K}_0 \geq N$ and all $\Upsilon_K \in \widetilde{\mathcal{V}}_K$ it holds that

$$\begin{aligned}
&\sup_{\Phi_K \in \widetilde{\mathcal{V}}_K} \frac{\left\langle \Upsilon_K, e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})}(H - \mathcal{E}_{\mathrm{GS}}^*) e^{\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})} \Phi_K \right\rangle_{\widehat{\mathcal{H}}^1 \times \mathcal{H}^{-1}}}{\|\Upsilon_K\|_{\mathcal{F}_K} \|\Phi_K\|_{\mathcal{F}_K}} \\
&\quad \geq \widetilde{\gamma} \frac{\|e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K\|_{\mathcal{F}_K}^2}{\|\Upsilon_K\|_{\mathcal{F}_K} \|\mathbb{P}_{0,K}^{\perp} \mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K\|_{\mathcal{F}_K}} \\
&\quad = \widetilde{\gamma} \frac{\|\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K\|_{\mathcal{F}_K}^2}{\|\Upsilon_K\|_{\mathcal{F}_K} \|\mathbb{P}_{0,K}^{\perp} \mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K\|_{\mathcal{F}_K}} \\
&\quad \geq \widetilde{\gamma} \frac{1}{\|\mathbb{P}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \mathbb{P}_{\mathcal{F}_K}\|_{\mathcal{F}_K \to \mathcal{F}_K} \|\mathbb{P}_{0,K}^{\perp} \mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})} \mathbb{P}_{\mathcal{F}_K}\|_{\mathcal{F}_K \to \mathcal{F}_K}} \frac{\|e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K\|_{\mathcal{F}_K} \|\Upsilon_K\|_{\mathcal{F}_K}}{\|\Upsilon_K\|_{\mathcal{F}_K} \|e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K\|_{\mathcal{F}_K}} \\
&\quad = \widetilde{\gamma} \frac{1}{\|\mathbb{P}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \mathbb{P}_{\mathcal{F}_K}\|_{\mathcal{W}_K \to \mathcal{W}_K} \|\mathbb{P}_{0,K}^{\perp} \mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})} \mathbb{P}_{\mathcal{F}_K}\|_{\mathcal{F}_K \to \mathcal{F}_K}}.
\end{aligned}$$

Here, the second to last step uses the invertibility of the projected exponential cluster operator, i.e., $\mathbb{P}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \mathbb{P}_{\mathcal{F}_K}$ as a mapping from the approximation space $\mathcal{V}_K$ to $\mathcal{V}_K$. This invertibility is a consequence of the fact that $e^{\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Phi_K \in \mathcal{V}_K$ for any $\Phi_K \in \mathcal{V}_K$, which itself follows from **Structure Assumption B.II**.

Taking now the infimum over all $\Upsilon_K \in \widetilde{\mathcal{V}}_K$ completes the proof.

□

A number of remarks on the above proof of Lemma 26 for **Structure Assumption B.II** are in order.

Remark 29 (The role of **Structure Assumption B.II** in the proof of Lemma 26)**.**

Consider the above proof of Lemma 26. It is instructive to briefly discuss exactly where each condition appearing in ***Structure Assumption B.II*** *is used in our proof.*

- *The first point in* ***Structure Assumption B.II****, pertaining to the closedness of the approximation space $\mathcal{V}_K$ under the action of excitation and de-excitation operators is used to ensure that the chosen test function Φ_K, which is constructed by applying certain exponential cluster operators to an element $\Upsilon_K \in \mathcal{V}_K$ and then taking a projection, is indeed an element of the approximation space $\mathcal{V}_K$. Additionally, the so-called de-excitation completeness condition, i.e., $e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})^\dagger} \Upsilon_K \in \mathcal{V}_K$ for all Υ_K is used to argue the invertibility of the exponential cluster operator in the approximation space, which is a crucial step in our proof.*

- *The fact that the approximation space $\mathcal{V}_K$ is an invariant subspace of the restricted mean-field operator $\widetilde{\mathcal{F}}_K \colon \mathcal{W}_K \to \mathcal{W}_K$ has two uses. On the one hand, it implies that the projection operator $\mathbb{P}_{\mathcal{F}_K}$ onto $\mathcal{V}_K$, which is a priori defined with respect to the inner product generated by the mean-field operator $\mathcal{F}_K$, is simultaneously $\widehat{\mathcal{L}}^2$-orthogonal. On the other hand, this invariance property plays a key role in our estimate of the term* (II*) *as it allows us to (eventually) conclude that the term* (II*) *is small.*
- *The fact that the operator $\mathcal{F}_K$ is symmetric and coercive on the finite-dimensional space $\mathcal{W}_K$ with coercivity and continuity constants independent of K allows us to introduce a new inner-product $(\cdot,\cdot)_{\mathcal{F}_K}$ on $\mathcal{W}_K$ which is equivalent to the usual $\widehat{\mathcal{H}}^1$ inner product. This allows us (see above) to construct a projection operator $\mathbb{P}_{\mathcal{F}_K}$ which is simultaneously $\widehat{\mathcal{L}}^2$-orthogonal and $(\cdot,\cdot)_{\mathcal{F}_K}$ orthogonal. These simultaneous orthogonalities aid us in estimating the term* (II*) *using, what basically amount to, spectral arguments.*
- *Finally, the splitting of the electronic Hamiltonian H in terms of the mean-field operator $\mathcal{F}_K$ and $\mathcal{U}_K \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{L}}^2$ is used crucially in the estimate of term* (II*). *The smallness assumption on the continuity constant of $\mathcal{U}_K$ is used to argue that our final bound on the term* (II*) *is indeed small.*

Remark 30 (The role of other assumptions in the proof of Lemma 26)**.**

Consider again the above proof of Lemma 26. In addition to ***Structure Assumption B.II****, we also use* ***Assumptions A.I-A.III*** *and of course the density of the approximation spaces $\{\mathcal{V}_K\}_{K\geq N}$ in $\widehat{\mathcal{H}}^1$. These results are used to make asymptotic arguments pertaining to the approximability of the exact ground state zero $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ of the coupled cluster function in the approximation space $\widetilde{\mathcal{V}}_K$. More precisely, we use these assumptions to argue that the term* (I*) *appearing in the above proof is indeed positive and that the term* (III*) *goes to zero asymptotically as $K \to \infty$.*

From a close study of the method of our proof, it can be deduced that the the density of the approximation spaces is not a necessary condition. Indeed, for a given choice of approximation space $\mathcal{V}_K$, the question of whether the discrete inf-sup condition (59) *holds depends (in addition to* ***Structure Assumption B.II****) on how well the exact ground state coupled cluster zero $\Theta^*_{K,\mathrm{GS}}$ is approximated in $\widetilde{\mathcal{V}}_K$.*

Having completed our second proof of Lemma 26, let us finally turn to the question of how restrictive **Structure Assumption B.II** really is.

Remark 31 (Restrictiveness of **Structure Assumption B.II**)**.**

As we discuss in more detail in Appendix B, the first two points of this assumption are satisfied by standard coupled cluster discretisations based on an initial Hartree-Fock computation in a finite basis with a suitably shifted so-called N-particle Hartree-Fock operator (defined below in Appendix A) playing the role of the mean-field operator $\mathcal{F}_K \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$. We emphasise that this is true not just for coupled cluster methods based on the canonical Hartree-Fock orbitals but even for rotated orbitals as long as the rotations maintain the L^2-orthogonality of occupied and virtual spaces.

The third point is also satisfied by such Hartree-Fock discretisations, provided that we assume convergence of the discrete Hartree-Fock ground state towards some limiting Slater determinant in the complete basis set limit. As far as we are aware, this complete basis set convergence of the Hartree-Fock method has not been rigorously established in the literature although it seems reasonable to assume.

The fourth condition in ***Structure Assumption B.II*** *is by far the most restrictive. Indeed, while it can be shown (again, see Appendix A below) that the N-particle Hartree-Fock operator $\mathcal{F}_K$ can be*

used to split the electronic Hamiltonian H as $H = \mathcal{F}_K + \mathcal{U}_K$ with $\mathcal{U}_K \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{L}}^2$, the smallness of the continuity constant of $\mathcal{U}_K$ is problem-dependent and not a universal property. To be more precise, recalling the above proof of Lemma 26 for ***Structure Assumption B.II****, we see that the precise condition that must be satisfied is (c.f., Point 4) in* ***Structure Assumption B.II****)*

$$\left\| \mathbb{P}_{\mathcal{F}_K}^{\perp} \widetilde{\mathcal{U}}_K \mathbb{P}_{\mathcal{F}_K} \right\|_{\mathcal{F}_K \to \widehat{\mathcal{L}}^2} < \frac{\sqrt{\Lambda_{\min}^{\mathcal{F}_K}}\, \Gamma_{\mathrm{GS}}^*}{\left\| \mathbb{P}_{\mathcal{F}_K}^{\perp} e^{\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})} \mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta_{K,\mathrm{GS}}^{\Pi})} \mathbb{P}_{\mathcal{F}_K} \right\|_{\mathcal{F}_K \to \mathcal{F}_K}}, \tag{75}$$

where we remind the reader that the constant Γ_{GS}^* is the coercivity constant of the shifted electronic Hamiltonian $H - \mathcal{E}_{\mathrm{GS}}^*$ on the finite-dimensional space $\{\Psi_{\mathrm{GS},K}^*\}^{\perp} \cap \mathcal{W}_K$, while the constant $\Lambda_{\min}^{\mathcal{F}_K} > 0$ denotes the minimum of the spectrum of the restricted mean-field operator $\widetilde{\mathcal{F}}_K$ in the range of the projection operator $\mathbb{P}_{\mathcal{F}_K}^{\perp} \colon \mathcal{W}_K \to \mathcal{W}_K$, i.e.,

$$\Lambda_{\min}^{\mathcal{F}_K} = \min_{\substack{\Phi_K \in \mathcal{W}_K \\ \Phi_K \in \mathcal{V}_K^{\perp}}} \frac{\langle \Phi_K, \mathcal{F}_K \Phi_K \rangle_{\widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^{-1}}}{\|\Phi_K\|_{\widehat{\mathcal{L}}^2}^2}.$$

The coercivity constant Γ_{GS}^* depends on the spectral gap of the specific molecular system being studied while the operator norm appearing in Inequality (75) depends on the norm of the exact ground state coupled cluster zero $\Theta_{K,\mathrm{GS}}$ which is also molecule-dependent. Moreover, even though we have assumed the density of the approximation spaces $\{\mathcal{V}_K\}_{K \geq N}$ in $\widehat{\mathcal{H}}^1$, it is <u>not the case</u> that the constant $\Lambda_{\min}^{\mathcal{F}_K} \to \infty$ as $K \to \infty$. Essentially, this is a consequence of the fact that the Laplacian, considered as an operator on the unbounded domain $\mathbb{R}^3$, has an essential spectrum consisting of $[0, \infty)$. Thus any mean-field operator $\mathcal{F}_K \colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ constructed according to the stipulations of ***Structure Assumption B.II*** *(and, in particular, the N-particle Hartree-Fock operator) will also posses an essential spectrum that contains the interval $[0, \infty)$. Having said this, for the specific case of the N-particle Hartree-Fock operator and the so-called truncated coupled cluster equations which are based on excitation-rank truncations (such as CCSD, CCSDT etc.), lower bounds can be derived for $\Lambda_{\min}^{\mathcal{F}_K}$ in terms of the so-called Hartree-Fock excitation energies. In particular, thanks to the density of the approximation spaces $\{\mathcal{V}_K\}_{K \geq N}$ in $\widehat{\mathcal{H}}^1$, we can deduce that for K sufficiently large, a simple upper bound for $\Lambda_{\min}^{\mathcal{F}_K}$ is given by the sum of the first N excitation energies of the N-particle Hartree-Fock operator $\mathcal{F}_K$, which is typically a large quantity (see Appendix B).*

The upshot of the above discussion is that, for an arbitrary molecular system, we are, unfortunately, unable to guarantee that the continuity constant of the operator $\mathcal{U}_K$ is sufficiently small, even in the asymptotic limit $K \to \infty$. We are therefore also unable to guarantee unconditional validity of the discrete inf-sup estimate (58) in the asymptotic limit $K \to \infty$ for arbitrary molecules. Nevertheless, as we discuss in the sequel, our numerical results indicate that the smallness condition (75) is often met for small molecules, particularly if the approximation space $\mathcal{V}_K$ is sufficiently rich.

4.3. **Main Results on the Well-posedness of the Discrete Coupled Cluster Equations.**

Throughout this subsection, we assume the setting of Sections 4.1 and 4.2. Equipped with the technical lemmas that we have proved in these section, we are now ready to state and prove the main result of the present article.

Theorem 32 (Local Well-posedness of the Discrete Coupled Cluster Equations)**.**

Let the sequence of reference determinants $\{\Psi_{0,K}\}_{K \geq N}$ and the sequence of approximation spaces $\{\widetilde{\mathcal{V}}_K\}_{K \geq N}$ be constructed as in Notation 21 and Notation 22 respectively, and assume that ***Assumptions A.I-A.III*** *and either* ***Structure Assumption B.I*** *or* ***Structure Assumption B.II*** *holds.*

Additionally, for each $K \geq N$, let $\{\Psi_{0,K}\}^{\perp}$ be the $\widehat{\mathcal{L}}^2$ orthogonal complement of $\Psi_{0,K}$ in $\widehat{\mathcal{L}}^2$, let the infinite-dimensional subspace $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ be defined as $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} := \{\Psi_{0,K}\}^{\perp} \cap \widehat{\mathcal{H}}^1$, let the coupled cluster function $f_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ be defined through Definition 18, and let $\Theta^*_{K,\mathrm{GS}}$ denote the zero of the coupled cluster function f_K corresponding to the ground state eigenfunction of the electronic Hamiltonian.

Consider now for each $K \geq N$, the discrete coupled cluster problem which consists of seeking $\Theta_K \in \widetilde{\mathcal{V}}_K$ such that (c.f., Equation (18))

$$\forall \Phi_K \in \widetilde{\mathcal{V}}_K \colon \qquad \langle \Phi_K, f_K(\Theta_K)\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} = 0. \tag{76}$$

There exists $\delta > 0$ and $\overline{K_0^*} \geq N$ such that for all $K \geq \overline{K_0^*}$, there exists a unique $\Theta_K \in \widetilde{\mathcal{V}}_K \cap \mathbb{B}_\delta(\Theta^*_{K,\mathrm{GS}})$ that solves the discrete coupled cluster equation (76), and we have the quasi-optimality result

$$\exists C > 0,\ \forall K \geq \overline{K_0^*} \colon \qquad \|\Theta_K - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} \leq C \inf_{\Phi_K \in \widetilde{\mathcal{V}}_K} \|\Phi_K - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1}, \tag{77}$$

as well as the residual-based error estimate

$$\forall K \geq \overline{K_0^*} \colon \qquad \|\Theta_K - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} \leq 2\|\mathrm{D}f_K\left(\Theta_K\right)^{-1}\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^* \to \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}} \|f_K\left(\Theta_K\right)\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}. \tag{78}$$

Proof. The proof that we construct is based on the arguments developed by Caloz and Rappaz [3, Theorems 6.1, 6.2, and 7.1] for conforming Galerkin discretisations of non-linear equations but adapted to the present non-conforming setting. Throughout this proof, we will assume the setting of, and use results from, Sections 4.1 and 4.2. It will be useful, in particular, to recall that $\Psi_0^* \in \widehat{\mathcal{H}}^1$ denotes the limiting reference determinant defined in **Assumption A.II**. Moreover, denoting by $\{\Psi_0^*\}^{\perp}$, the $\widehat{\mathcal{L}}^2$ orthogonal complement of Ψ_0^* in $\widehat{\mathcal{L}}^2$ and defining the infinite-dimensional subspace $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*}$ as $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*} := \{\Psi_0^*\}^{\perp} \cap \widehat{\mathcal{H}}^1$, we can introduce the coupled cluster function $f \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_0^*}\right)^*$ through Definition 18. In this context, we denote by Θ^*_{GS} the zero of the coupled cluster function f corresponding to the ground state eigenfunction of the electronic Hamiltonian.

We begin by defining for every $K \geq N$, the bilinear form $\mathcal{b}_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \mathbb{R}$ given by

$$\forall \Phi, \Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \colon \qquad \mathcal{b}_K(\Phi, \Upsilon) = \langle \Upsilon, \mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\Phi\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}, \tag{79}$$

where we remind the reader that $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}}) \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ is the Fréchet derivative of the coupled cluster function f_k at $\Theta^*_{K,\mathrm{GS}}$.

It follows directly from Lemma 26 that there exists $K_0 \geq N$ such that for all $K \geq K_0$, the bilinear form $\mathcal{b}_K$ satisfies a discrete inf-sup condition on the approximation space $\widetilde{\mathcal{V}}_K$.

Consequently, for any $K \geq K_0$ we can introduce projection operators $\Pi^{\mathrm{L}}_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \widetilde{\mathcal{V}}_K$ and $\Pi^{\mathrm{R}}_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \widetilde{\mathcal{V}}_K$ through the relations

$$\text{For any } \Phi \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}},\ \forall \Upsilon_K \in \widetilde{\mathcal{V}}_K \colon \qquad \mathcal{b}_K\left(\Phi - \Pi^{\mathrm{L}}_K\Phi, \Upsilon_K\right) = 0, \tag{80}$$

$$\text{For any } \Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}},\ \forall \Phi_K \in \widetilde{\mathcal{V}}_K \colon \qquad \mathcal{b}_K\left(\Phi_K, \Upsilon - \Pi^{\mathrm{R}}_K\Upsilon\right) = 0. \tag{81}$$

Note that both projection operators are well-defined thanks to the validity of the discrete inf-sup condition.

Equipped with the projections operators $\Pi_K^{\mathrm{L}}, \Pi_K^{\mathrm{R}}$, we next define for each $K \geq K_0$, the non-linear mapping $\mathcal{G}_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ given by

(82)
$$\forall \Phi, \Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \colon \quad \langle \Upsilon, \mathcal{G}_K(\Phi) \rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} := \left\langle \Pi_K^{\mathrm{R}} \Upsilon, \mathit{f}_K(\Phi) \right\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} + \mathcal{b}_K\left(\Phi, \Upsilon - \Pi_K^{\mathrm{R}} \Upsilon\right).$$

We now claim that for any $K \geq K_0$, the function $\Phi_K \in \widetilde{\mathcal{V}}_K$ solves the discrete coupled cluster equations (76) if and only if Φ_K solves the infinite-dimensional non-linear equation

$$\forall \Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \colon \qquad \langle \Upsilon, \mathcal{G}_K(\Phi_K) \rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} = 0. \tag{83}$$

Indeed, assume $\Phi_K \in \widetilde{\mathcal{V}}_K$ solves the discrete coupled cluster equations (76). Then it follows directly from the definition of the projection operator Π_K^{R} given by Equation (81) that Φ_K also solves Equation (83). Conversely, let $\Phi_K \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ be a solution to Equation (83). We must show that $\Phi_K \in \widetilde{\mathcal{V}}_K$.

To this end, notice that Equation (82) implies that for all $\Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$0 = \left\langle \Upsilon - \Pi_K^{\mathrm{R}} \Upsilon, \mathcal{G}_K(\Phi_K) \right\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} = \mathcal{b}_K(\Phi_K, \Upsilon - \Pi_K^{\mathrm{R}} \Upsilon).$$

Thanks to Definitions (80) and (81) of the projection operators Π_K^{L} and Π_K^{R} respectively, we can then deduce that for all $\Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$0 = \mathcal{b}_K(\Phi_K, \Upsilon - \Pi_K^{\mathrm{R}} \Upsilon) = \mathcal{b}_K(\Phi_K - \Pi_K^{\mathrm{L}} \Phi_K, \Upsilon).$$

Recalling now the definition of bilinear form $\mathcal{b}_K \colon : \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \mathbb{R}$ given by Equation (79) and using the fact that the Fréchet derivative $\mathrm{D}\mathit{f}_K(\Theta^*_{K,\mathrm{GS}}) \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ is an isomorphism (see Theorem 19), we conclude that $\Phi_K = \Pi_K^{\mathrm{L}} \Phi$, or equivalently, $\Phi_K \in \widetilde{\mathcal{V}}_K$ as claimed.

We will now analyse the (local) well-posedness of the equivalent Equation (83). To do so, we first prove four properties of the sequence of non-linear functions $\{\mathcal{G}_K\}_{K \geq K_0}$.

Claim One: The mapping $\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \ni \Theta \mapsto \mathcal{G}_K(\Theta)$ is of class $\mathscr{C}^1$ for all $K \geq K_0$.

To demonstrate this first claim, we use the fact (see, e.g., [11, Proposition 26]) that for any $K \geq N$ the coupled cluster function $\mathit{f}_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ is a $\mathscr{C}^\infty$ mapping. It then follows directly from Equation (82) that $\mathcal{G}_K$ is also a $\mathscr{C}^\infty$ mapping for any $K \geq K_0$. In particular, for any $\Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ the Fréchet derivative $\mathrm{D}\mathcal{G}_K(\Theta) \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ is given by

$$\begin{aligned} \forall \Phi, \Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \colon \quad \langle \Upsilon, \mathrm{D}\mathcal{G}_K(\Theta)\Phi \rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} &:= \left\langle \Pi_K^{\mathrm{R}} \Upsilon, \mathrm{D}\mathit{f}_K(\Theta)\Phi \right\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} \\ &\quad + \mathcal{b}_K\left(\Phi, \Upsilon - \Pi_K^{\mathrm{R}} \Upsilon\right). \end{aligned} \tag{84}$$

Claim Two: There exists $\delta_0 > 0, L_0 > 0$ and $\widetilde{K_0} \geq K_0$ such that for all $K \geq \widetilde{K_0}$ it holds that

$$\forall \Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \cap \mathbb{B}_{\delta_0}(\Theta^*_{K,\mathrm{GS}}) \colon \quad \|\mathrm{D}\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}\mathcal{G}_K(\Theta)\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} \leq L_0 \|\Theta^*_{K,\mathrm{GS}} - \Theta\|_{\widehat{\mathcal{H}}^1}.$$

To demonstrate the second claim, we begin by using Equation (84) to deduce that for all $K \geq K_0$ and all $\Theta, \Upsilon, \Phi \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$\begin{aligned}\left\langle \Upsilon, \left(\mathrm{D}\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}\mathcal{G}_K(\Theta)\right)\Phi\right\rangle_{\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} &= \left\langle \Pi^{\mathrm{R}}_K \Upsilon, \left(\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}f_K(\Theta)\right)\Phi\right\rangle_{\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} \\ &\leq \|\Pi^{\mathrm{R}}_K \Upsilon\|_{\widehat{\mathscr{H}}^1} \left\| \left(\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}f_K(\Theta)\right)\Phi\right\|_{\left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}. \end{aligned} \tag{85}$$

To proceed, we must bound the term $\|\Pi^{\mathrm{R}}_K \Upsilon\|_{\widehat{\mathscr{H}}^1}$. To this end, we observe that thanks to Equation (80), for every $K \geq K_0$ and any $\Upsilon \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$, the function $\Pi^{\mathrm{R}}_K \Upsilon$ is a solution to the following adjoint problem:

$$\forall \Phi_K \in \widetilde{\mathcal{V}}_K : \quad \mathcal{b}_K\left(\Phi_K, \Pi^{\mathrm{R}}_K \Upsilon\right) = \mathcal{b}_K\left(\Phi_K, \Upsilon\right).$$

Recalling the definition of the bilinear form $\mathcal{b}_K$ given by Equation (79) and using Lemma 26, which implies that the Fréchet derivative $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})$ satisfies a discrete inf-sup condition on $\widetilde{\mathcal{V}}_K$ with constant $\gamma > 0$, a simple calculation yields that

$$\forall K \geq K_0,\ \forall \Upsilon \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} : \qquad \|\Pi^{\mathrm{R}}_K \Upsilon\|_{\widehat{\mathscr{H}}^1} \leq \frac{\|\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}}{\gamma}.$$

Using now the expression for the Fréchet derivative $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}}) \colon \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ found in [11, Proposition 26], we further deduce that for all $K \geq K_0$ and all $\Upsilon \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$\|\Pi^{\mathrm{R}}_K \Upsilon\|_{\widehat{\mathscr{H}}^1} \leq \frac{\|e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\|_{\widehat{\mathscr{H}}^{-1} \to \widehat{\mathscr{H}}^{-1}} \|H - \mathcal{E}^*_{\mathrm{GS}}\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}} \|e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\|_{\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1}}{\gamma}.$$

Finally, recalling the continuity properties of cluster operators given by Theorem 16, together with the series expansion of the exponential mapping, and Lemma 24 which implies that $\Theta^*_{K,\mathrm{GS}} \to \Theta^*_{\mathrm{GS}}$ as $K \to \infty$ in the $\widehat{\mathscr{H}}^1$ norm, we finally deduce the existence of a constant $\breve{\beta} > 0$ such that for all $K \geq K_0$ and all $\Upsilon \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$\|\Pi^{\mathrm{R}}_K \Upsilon\|_{\widehat{\mathscr{H}}^1} \leq \frac{\breve{\beta}}{\gamma}, \tag{86}$$

where we emphasise that both $\breve{\beta}, \gamma > 0$ are uniformly bounded in K.

In view of Inequalities (85) and (86), it remains to prove that for every $K \geq K_0$ the Fréchet derivative $\mathrm{D}f_K$ of the coupled cluster function f_K is locally Lipschitz continuous in a neighbourhood around $\Theta^*_{K,\mathrm{GS}}$ with Lipschitz constant uniformly bounded in K. While this local Lipschitz continuity has been established in prior results (see, e.g., [11, Proposition 26]), the Lipschitz constants obtained therein have not been shown to be uniform in K. We now state a brief argument that bridges this gap.

To do so, we will make use of the expression for the Fréchet derivative of the coupled cluster function found in our previous contribution [11, Proposition 26]. Using this expression, we deduce that for all $K \geq K_0$ and all $\Theta, \Phi \in \widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$\begin{aligned} &\left\| \left(\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}f_K(\Theta)\right)\Phi\right\|_{\left(\widehat{\mathscr{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} \\ &\leq \left\| e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Big[H, \mathcal{T}(\Phi)\Big] e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K} - e^{-\mathcal{T}(\Theta)}\Big[H, \mathcal{T}(\Phi)\Big] e^{\mathcal{T}(\Theta)}\Psi_{0,K}\right\|_{\widehat{\mathscr{H}}^{-1}}, \end{aligned}$$

where $\mathcal{T}(\Phi), \mathcal{T}(\Theta)$ denote, as usual the cluster operators generated by $\Phi, \Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ respectively, and $[\cdot,\cdot]$ denotes the usual commutator.

Adding and subtracting suitable terms now yields

$$\| \left(\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}f_K(\Theta)\right)\Phi\|_{\widehat{\mathcal{H}}^{-1}} \leq \underbrace{\left\| \left(e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})} - e^{-\mathcal{T}(\Theta)}\right)\left[H, \mathcal{T}(\Phi)\right] e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K}\right\|_{\widehat{\mathcal{H}}^{-1}}}_{:=(\mathrm{I})} + \underbrace{\left\| e^{-\mathcal{T}(\Theta)}\left[H, \mathcal{T}(\Phi)\right]\left(e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})} - e^{-\mathcal{T}(\Theta)}\right)\Psi_{0,K}\right\|_{\widehat{\mathcal{H}}^{-1}}}_{:=(\mathrm{II})}. \tag{87}$$

Let us first consider the term (I). We immediately observe that

$$(\mathrm{I}) \leq \underbrace{\left\|\left(e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})} - e^{-\mathcal{T}(\Theta)}\right)\right\|_{\widehat{\mathcal{H}}^{-1}\to\widehat{\mathcal{H}}^{-1}}}_{:=(\mathrm{IA})} \underbrace{\left\|\left[H, \mathcal{T}(\Phi)\right]\right\|_{\widehat{\mathcal{H}}^{1}\to\widehat{\mathcal{H}}^{-1}}}_{:=(\mathrm{IB})} \underbrace{\left\| e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K}\right\|_{\widehat{\mathcal{H}}^{1}}}_{:=(\mathrm{IC})}. \tag{88}$$

We proceed term-by-term beginning with (IC). We first notice that thanks to Theorem 16 concerning the continuity properties of cluster operators, there exists a constant $\widetilde{\beta} > 0$ depending only on N and $\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}$ such that

$$(\mathrm{IC}) = \left\| e^{\mathcal{T}(\Theta^*_{K,\mathrm{GS}})}\Psi_{0,K}\right\|_{\widehat{\mathcal{H}}^1} \leq e^{\widetilde{\beta}\|\Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1}}\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}. \tag{89}$$

Using again Theorem 16 and appealing to the continuity of the electronic Hamiltonian $H\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$, we further deduce the existence of a constant $\widehat{\beta} > 0$ depending only on N, $\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}$ and the continuity constant of H such that

$$(\mathrm{IB}) = \left\|\left[H, \mathcal{T}(\Phi)\right]\right\|_{\widehat{\mathcal{H}}^{1}\to\widehat{\mathcal{H}}^{-1}} \leq \widehat{\beta}\|\Phi\|_{\widehat{\mathcal{H}}^1}. \tag{90}$$

Finally, relying once again on the $\widehat{\mathcal{H}}^{-1}$ continuity of cluster operators given by Theorem 16 together with the series expansion of the exponential, we deduce the existence of a constant $\overline{\beta} > 0$, again depending only on N and $\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}$, such that for all $\Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ that satisfy

$$\|\mathcal{T}(\Theta) - \mathcal{T}(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathcal{H}}^{-1}\to\widehat{\mathcal{H}}^{-1}} < 1, \tag{91}$$

it holds that

$$(\mathrm{IA}) = \left\|\left(e^{-\mathcal{T}(\Theta^*_{K,\mathrm{GS}})} - e^{-\mathcal{T}(\Theta)}\right)\right\|_{\widehat{\mathcal{H}}^{-1}\to\widehat{\mathcal{H}}^{-1}} \leq \overline{\beta}\|\Theta - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1}. \tag{92}$$

On the other hand, the linearity and continuity properties of cluster operators given by Theorem 16 imply the existence of a constant $\widetilde{\delta} > 0$ depending only on N and $\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}$ such that for any $\Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$\|\mathcal{T}(\Theta) - \mathcal{T}(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathcal{H}}^{-1}\to\widehat{\mathcal{H}}^{-1}} \leq \widetilde{\delta}\|\Theta - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1}. \tag{93}$$

In view of Inequalities (88)-(93), we deduce that for any $K \geq K_0$ and all $\Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \cap \mathbb{B}_{1/\widetilde{\delta}}(\Theta^*_{K,\mathrm{GS}})$ it holds that

$$(\mathrm{I}) \leq \overline{\beta}\widehat{\beta}\, e^{\widetilde{\beta}\|\Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1}}\ \|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}\|\Phi\|_{\widehat{\mathcal{H}}^1}\|\Theta - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1}$$

A similar argument can be used to bound the term (II) appearing in Inequality (87), and we thus conclude that for any $K \geq K_0$ and all $\Theta \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \cap \mathbb{B}_{1/\widetilde{\delta}}(\Theta^*_{K,\mathrm{GS}})$ it holds that

$$\| \left(\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}f_K(\Theta) \right) \Phi \|_{\widehat{\mathcal{H}}^{-1}} \leq \overline{\beta}\widehat{\beta} \left(1 + e^{\frac{\widetilde{\beta}}{\widetilde{\delta}}} \right) e^{\widetilde{\beta}\|\Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1}} \|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1} \|\Phi\|_{\widehat{\mathcal{H}}^1} \|\Theta - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1},$$

where we emphasise that the constants $\overline{\beta}, \widehat{\beta}, \widetilde{\beta}, \widetilde{\delta} > 0$ all depend only on N and $\|\Psi_{0,K}\|_{\widehat{\mathcal{H}}^1}$. Consequently, appealing to Lemma 24 which implies the convergence of $\Theta^*_{K,\mathrm{GS}}$ to Θ^*_{GS} as $K \to \infty$, together with **Assumption A.II** which implies the convergence of the reference determinant $\Psi_{0,K}$ to $\Psi^*_0 \in \widehat{\mathcal{H}}^1$ as $K \to \infty$, we deduce the required local Lipschitz continuity of the Fréchet derivative $\mathrm{D}f_K$ at $\Theta^*_{K,\mathrm{GS}}$ with Lipschitz constant uniformly bounded in K.

Claim Three: The sequence of non-linear maps $\{\mathcal{G}_K\}_{K \geq K_0}$ defined through Equation (82) satisfy

$$\lim_{K \to \infty} \|\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}})\|_{(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*} = 0.$$

To demonstrate this claim, we will use the definition of the non-linear mapping $\mathcal{G}_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ given by Equation (82) together with the fact that $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ is by definition, a zero of the coupled cluster function $f_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$. Combining these two facts yields that for all $K \geq K_0$ and all $\Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ we have that

$$\begin{aligned} \left\langle \Upsilon, \mathcal{G}_K(\Theta^*_{K,\mathrm{GS}}) \right\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} &= \mathcal{b}_K \left(\Theta^*_{K,\mathrm{GS}}, \Upsilon - \Pi^{\mathrm{R}}_K \Upsilon \right) \\ &= \mathcal{b}_K \left(\Theta^*_{K,\mathrm{GS}} - \Pi^{\mathrm{L}}_K \Theta^*_{K,\mathrm{GS}}, \Upsilon \right) \\ &\leq \|\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} \|\Theta^*_{K,\mathrm{GS}} - \Pi^{\mathrm{L}}_K \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} \|\Upsilon\|_{\widehat{\mathcal{H}}^1}. \end{aligned} \tag{94}$$

where the second step follows from the definition of the projection operators Π^L_K and Π^R_K given by Equations (80) and (81) respectively, and the third step follows from the definition of the bilinear form $\mathcal{b}_K \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \mathbb{R}$ given by Equation (79).

We first focus on bounding the term involving the projection operator Π^{L}_K. To this end, we recall that by assumption, the Fréchet derivative $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})$– and hence the bilinear form $\mathcal{b}_K$– satisfies a discrete inf-sup condition on $\widetilde{\mathcal{V}}_K$. Consequently, recalling further the definition of the projection operator Π^{L}_K given by Equation (80), we can use the well-known Céa's lemma to deduce that for all $K \geq K_0$ and any $\Phi \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ it holds that

$$\|\Theta^*_{K,\mathrm{GS}} - \Pi^{\mathrm{L}}_K \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} \leq \left(1 + \frac{\|\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}}{\gamma} \right) \inf_{\widetilde{\Phi}_K \in \widetilde{\mathcal{V}}_K} \|\Theta^*_{K,\mathrm{GS}} - \widetilde{\Phi}_K\|_{\widehat{\mathcal{H}}^1}. \tag{95}$$

Consider now Inequalities (94) and (95). We have already argued (see the proof of **Claim Two** above) that the operator norm $\|\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}$ is uniformly bounded above in K. Recalling therefore Lemma 25 which implies that the $\widehat{\mathcal{H}}^1$-best approximation error of $\Theta^*_{K,\mathrm{GS}}$ in the approximation space $\widetilde{\mathcal{V}}_K$ goes to zero as $K \to \infty$ now yields the required result.

Claim Four: For every $K \geq K_0$ it holds that the Fréchet derivative $\mathrm{D}\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}}) \colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ is an isomorphism with continuity and inf-sup constants uniformly bounded in K.

To demonstrate this final claim, we make use of Equation (84) which implies that for all $K \geq K_0$ it holds that

$$\begin{aligned}\forall \Phi, \Upsilon \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}: \quad \left\langle \Upsilon, \mathrm{D}\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}})\Phi \right\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} &= \left\langle \Pi^{\mathrm{R}}_K \Upsilon, \mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\Phi \right\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} \\ &\quad + \mathcal{b}_K\left(\Phi, \Upsilon - \Pi^{\mathrm{R}}_K \Upsilon\right) \\ &= \left\langle \Upsilon, \mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\Phi \right\rangle_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \times \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*},\end{aligned}$$

where the second step follows from the definition of the bilinear form $\mathcal{b}_K$ given by Equation (79). Thus, we have that $\mathrm{D}\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}}) = \mathrm{D}f_k(\Theta^*_{K,\mathrm{GS}})$ and the isomorphism property immediately follows from Theorem 19.

To demonstrate the remaining portion of **Claim Four**, note that we have already shown in the proof of **Claim Two** that the continuity constant of the mapping $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})\colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ is uniformly bounded in $K \geq K_0$. Similar arguments allow us to conclude that the inf-sup constant of $\mathrm{D}f_K(\Theta^*_{K,\mathrm{GS}})$ (given by Theorem 19) is also uniformly bounded in $K \geq K_0$. For the sake of brevity and to avoid repeating tedious arguments, we omit these details.

Equipped with **Claims One** through **Four**, we are now ready to conclude our proof of the present theorem. We seek to apply the inverse function theorem for Banach spaces, namely, Theorem 23 to the non-linear mapping $\mathcal{G}_K\colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ for K sufficiently large. To this end, let us define for each $K \geq \widetilde{K_0} \geq K_0$ the quantities

$$\epsilon_K := \|\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}})\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*},$$

$$\gamma^*_{\mathcal{G}} := \|\mathrm{D}\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}})^{-1}\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^* \to \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}},$$

$$\forall \alpha \in [0, \delta_0): \qquad \mathrm{L}(\alpha) := \sup_{\Theta \in \mathbb{B}_\alpha(\Theta^*_{K,\mathrm{GS}})} \|\mathrm{D}\mathcal{G}_K(\Theta^*_{K,\mathrm{GS}}) - \mathrm{D}\mathcal{G}_K(\Theta)\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}.$$

We now observe that

- Thanks to **Claim One**, for all $K \geq \widetilde{K_0}$, the mapping $\mathcal{G}_K\colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ is of regularity class $\mathscr{C}^1$.
- Thanks to **Claim Two**, for all $K \geq \widetilde{K_0}$ and $\alpha \in [0, \delta_0)$ with $\delta_0 > 0$ independent of K, the constant $\mathrm{L}(\alpha)$ is uniformly bounded above in K.
- Thanks to **Claim Three**, the constant ϵ_K defined above satisfies $\lim_{K\to\infty} \epsilon_K = 0$.
- Thanks to **Claim Four**, for all $K \geq \widetilde{K_0}$, the Fréchet derivative $\mathrm{D}\mathcal{G}_K\colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ is an isomorphism and the constant $\gamma^*_{\mathcal{G}}$ defined above is uniformly bounded above in K.

Consequently, there exists $\overline{K_0} \geq N$ such that for all $K \geq \overline{K_0}$ it holds that

$$2\gamma^*_{\mathcal{G}} \mathrm{L}(2\gamma^*_{\mathcal{G}}\epsilon_K) < 1,$$

and the conclusions of Theorem 23 apply: for all $K \geq \overline{K_0}$ the closed ball $\overline{\mathbb{B}_{2\gamma^*_{\mathcal{G}}\epsilon_K}\left(\Theta^*_{K,\mathrm{GS}}\right)} \subset \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ contains a unique solution Θ_K to the equation

$$\mathcal{G}_K(\Theta_K) = 0,$$

the Fréchet derivative $\mathrm{D}\mathcal{G}_K(\Theta_K)\colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*$ is an isomorphism with

$$\|\mathrm{D}\mathcal{G}_K(\Theta_K)^{-1}\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^* \to \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}} \leq 2\gamma^*_{\mathcal{G}}, \tag{96}$$

and for all $\Theta \in \overline{\mathbb{B}_{2\gamma^*_{\mathcal{G}}\epsilon_K}(\Theta^*_K)} \cap \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$, we have the error estimate

$$\|\Theta_K - \Theta\|_{\widehat{\mathcal{H}}^1} \leq 2\gamma^*_{\mathcal{G}} \|\mathcal{G}_K(\Theta)\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}. \tag{97}$$

Taking now Inequality (97) and applying the earlier Inequalities (94) and (95) immediately yields the quasi-optimality result (77).

It remains to establish the residual-based error estimate (78). To this end, let us define for each $K \geq \overline{K_0}$ the quantities

$$\widetilde{\epsilon_K} := \|f_K(\Theta_K)\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*},$$

$$\gamma^*_f := \|\mathrm{D}f_K(\Theta_K)^{-1}\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^* \to \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}},$$

$$\forall \alpha \in [0,\delta_0)\colon \qquad \widetilde{\mathrm{L}(\alpha)} := \sup_{\Theta \in \overline{\mathbb{B}_\alpha(\Theta_K)}} \|\mathrm{D}f_K(\Theta_K) - \mathrm{D}f_K(\Theta)\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to \left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*}.$$

To proceed, let us first observe that the quasi-optimality result (77) in combination with the best-approximation result given by Lemma 25 together imply that

$$\lim_{K\to\infty} \|\Theta_K - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^1} = 0. \tag{98}$$

Next, we claim that we also have the consistency result

$$\lim_{K\to\infty} \widetilde{\epsilon_K} = \lim_{K\to\infty} \|f_K(\Theta_K)\|_{\left(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}\right)^*} = 0. \tag{99}$$

Indeed, recalling on the one hand that $\Theta^*_{K,\mathrm{GS}} \in \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ is a zero of the coupled cluster function f_K, and on the other hand that f_K is a $\mathscr{C}^\infty$ mapping, we may appeal to the mean-value inequality for Banach spaces (see, e.g., [9, Theorem 3.2.7]) to deduce that for all $K \geq \overline{K_0}$ it holds that

$$\begin{aligned}
\|f_K(\Theta_K) - f_K(\Theta^*_{K,\mathrm{GS}})\|_{(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*} &\leq \sup_{t\in[0,1]} \|\mathrm{D}f_K\left(\Theta^*_{K,\mathrm{GS}} + t(\Theta_K - \Theta^*_{K,\mathrm{GS}})\right)\left(\Theta_K - \Theta^*_{K,\mathrm{GS}}\right)\|_{(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*} \\
&\leq \sup_{t\in[0,1]} \|\mathrm{D}f_K\left(\Theta^*_{K,\mathrm{GS}} + t(\Theta_K - \Theta^*_{K,\mathrm{GS}})\right)\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*} \\
&\quad \cdot \|\Theta_K - \Theta^*_{K,\mathrm{GS}}\|_{\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}}.
\end{aligned}$$

Notice now that for any $t \in [0,1]$ and all $K \geq \overline{K_0}$, the function $\Theta^*_{K,\mathrm{GS}} + t(\Theta_K - \Theta^*_{K,\mathrm{GS}}) \in \overline{\mathbb{B}_\alpha(\Theta^*_{K,\mathrm{GS}})}$ for any $\alpha \in [0,\delta_0)$. We therefore deduce from **Claim Two** that the Fréchet derivative $\mathrm{D}f_K$ is locally Lipschitz continuous in a closed ball containing the set $\{\Theta^*_{K,\mathrm{GS}} + t(\Theta_K - \Theta^*_{K,\mathrm{GS}})\colon t \in [0,1]\}$ with Lipschitz constant uniformly bounded in K. The consistency result (99) then immediately follows from Equation (98).

We now observe that

- For all $K \geq \overline{K_0}$, the mapping $f_K\colon \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}} \to (\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*$ is of regularity class $\mathscr{C}^1$ (see, e.g., [11, Proposition 26].

- Thanks to **Claim Two** and the $\widehat{\mathcal{H}}^1$-convergence of Θ_K to $\Theta^*_{K,\mathrm{GS}}$, there exists $\widetilde{\widetilde{K_0}} \geq \overline{K_0}$ and $\widetilde{\delta_0} > 0$ sufficiently small and independent of K, such that for all $\alpha \in [0, \delta_0)$ the constant $\widetilde{\mathrm{L}(\alpha)}$ is uniformly bounded above in K.
- Thanks to Equation (99), the constant $\widetilde{\epsilon_K}$ defined above satisfies $\lim_{K\to\infty} \epsilon_K = 0$.
- Thanks to Inequality (96), the constant γ_f^* defined above is uniformly bounded above in K.

We can therefore conclude that there exists $\widehat{K_0} \geq N$ such that for all $K \geq \widehat{K_0}$ it holds that

$$2\gamma_f^* \mathrm{L}(2\gamma_f^* \widetilde{\epsilon_K}) < 1.$$

As a consequence, the conclusions of Theorem 23 are once again applicable. In particular, for all $K \geq \widehat{K_0}$ the closed ball $\overline{\mathbb{B}_{2\gamma_f^*\widetilde{\epsilon_K}}(\Theta_K)} \subset \widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}}$ contains a unique solution $\widetilde{\Theta_K}$ to the equation

$$f_K(\widetilde{\Theta_K}) = 0,$$

and we have the error estimate

$$\|\widetilde{\Theta_K} - \Theta_K\|_{\widehat{\mathcal{H}}^1} \leq 2\gamma_f^* \|f_K(\Theta_K)\|_{(\widehat{\mathcal{H}}^{1,\perp}_{\Psi_{0,K}})^*}.$$

Recalling that $\Theta^*_{K,\mathrm{GS}}$ is the locally unique zero of the coupled cluster function f_K (see Theorem 19), we deduce that in fact $\widetilde{\Theta_K} = \Theta^*_{K,\mathrm{GS}}$. The required residual-based error estimate thus follows.

□

Acknowledgements

The first author warmly thanks Antoine Levitt for several helpful discussions. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 810367). The majority of this work was done while the first author was employed as a post-doctoral researcher at the Laboratoire Jacques-Louis Lions (LJLL) at Sorbonne Université.

References

[1] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan. Quantum algorithms for quantum chemistry and quantum materials science. *Chemical Reviews*, 120(22):12685–12717, 2020.
[2] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer-Verlag New York, 2011.
[3] G. Caloz and J. Rappaz. Numerical analysis for nonlinear and bifurcation problems. In *Handbook of numerical analysis, Vol. V*. North-Holland, Amsterdam, 1997.
[4] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero. Unconventional superconductivity in magic-angle graphene superlattices. *Nature*, 556(7699):43–50, 2018.
[5] F. Chatelin. *Spectral approximation of linear operators*. Academic Press Inc., New York, 1983.
[6] C. Curutchet and B. Mennucci. Quantum chemical studies of light harvesting. *Chemical reviews*, 117(2):294–343, 2017.
[7] P. Deglmann, A. Schäfer, and C. Lennartz. Application of quantum calculations in the chemical industry—an overview. *International Journal of Quantum Chemistry*, 115(3):107–136, 2015.
[8] J. M. Dieterich and E. A. Carter. Opinion: Quantum solutions for a sustainable energy future. *Nature Reviews Chemistry*, 1(4):1–7, 2017.
[9] P. Drábek and J. Milota. *Methods of nonlinear analysis*. Birkhäuser Verlag, Basel, 2007.
[10] F. M. Faulstich, A. Laestadius, O. Legeza, R. Schneider, and S. Kvaal. Analysis of the tailored coupled-cluster method in quantum chemistry. *SIAM Journal on Numerical Analysis*, 57(6):2579–2607, 2019.
[11] M. Hassan, Y. Maday, and Y. Wang. Analysis of the single reference coupled cluster method for electronic structure calculations: the full-coupled cluster equations. *Numerische Mathematik*, 155(1):121–173, 2023.
[12] K. Head-Marsden, J. Flick, C. J. Ciccarino, and P. Narang. Quantum information and algorithms for correlated quantum matter. *Chemical Reviews*, 121(5):3061–3120, 2020.
[13] T. Helgaker, P. Jorgensen, and J. Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2014.

[14] A. Hillisch, N. Heinrich, and H. Wild. Computational chemistry in the pharmaceutical industry: from childhood to adolescence. *ChemMedChem*, 10(12):1958–1962, 2015.
[15] P. D. Hislop and I. M. Sigal. *Introduction to spectral theory*. Springer-Verlag, New York, 1996. With applications to Schrödinger operators.
[16] W. Hunziker and I. M. Sigal. The quantum N-body problem. *Journal of Mathematical Physics*, 41(6):3448–3510, 2000.
[17] A. Laestadius and S. Kvaal. Analysis of the extended coupled-cluster method in quantum chemistry. *SIAM Journal on Numerical Analysis*, 56(2):660–683, 2018.
[18] T. J. Lee and G. E. Scuseria. Achieving chemical accuracy with coupled-cluster theory. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, pages 47–108. Springer, 1995.
[19] E. H. Lieb and B. Simon. The Hartree-Fock theory for Coulomb systems. *Communications in Mathematical Physics*, 53(3):185–194, 1977.
[20] P.-L. Lions. Solutions of Hartree-Fock equations for Coulomb systems. *Communications in Mathematical Physics*, 109(1):33–97, 1987.
[21] S. L. Lovelock, R. Crawshaw, S. Basler, C. Levy, D. Baker, D. Hilvert, and A. P. Green. The road to fully programmable protein catalysis. *Nature*, 606(7912):49–58, 2022.
[22] C. J. Manly, S. Louise-May, and J. D. Hammer. The impact of informatics and computational chemistry on synthesis and screening. *Drug discovery today*, 6(21):1101–1110, 2001.
[23] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters*, 157(6):479–483, 1989.
[24] T. Rohwedder. *An analysis for some methods and algorithms of Quantum Chemistry*. PhD thesis, Technische Universität Berlin, 2010.
[25] T. Rohwedder. The continuous coupled cluster formulation for the electronic Schrödinger equation. *ESAIM. Mathematical Modelling and Numerical Analysis*, 47(2):421–447, 2013.
[26] T. Rohwedder and R. Schneider. Error estimates for the coupled cluster method. *ESAIM. Mathematical Modelling and Numerical Analysis*, 47(6):1553–1582, 2013.
[27] S. Sauter and C. Schwab. *Boundary element methods*. Springer-Verlag Berlin Heidelberg, 2011.
[28] R. Schneider. Analysis of the projected coupled cluster method in electronic structure calculation. *Numerische Mathematik*, 113(3):433–471, 2009.
[29] L. Si and K. Held. Electronic structure of the putative room-temperature superconductor pb 9 cu (po 4) 6 o. *Physical Review B*, 108(12):L121110, 2023.
[30] H. Yserentant. On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numerische Mathematik*, 98(4):731–759, 2004.
[31] H. Yserentant. *Regularity and approximability of electronic wave functions*. Springer-Verlag, Berlin, 2010.

Appendix A. The Hartree-Fock Method

The Hartree-Fock method is a basic numerical algorithm for approximating the ground state of the electronic Hamiltonian $H\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$ defined through Equation (3). The essential idea of the Hartree-Fock approximation is to replace the infinite-dimensional tensor space $\widehat{\mathscr{H}}^1$ of antisymmetric functions– over which the minimum expectation value of the electronic Hamiltonian is computed– with a more computationally tractable subset, in this case the set of *single Slater determinants*. As in the main text of this article, we assume throughout this section that the total nuclear charge $Z = \sum_{\alpha=1}^{M} Z_\alpha > N-1$.

Definition 33 (Hartree-Fock Minimisation Set)**.**

We define the Hartree-Fock minimisation set S_{HF} *as the subset of* $\widehat{\mathscr{H}}^1$ *given by*

$$\mathrm{S}_{\mathrm{HF}} := \Big\{ \Phi \in \widehat{\mathscr{H}}^1 \colon \text{ there exist } \mathrm{L}^2\text{-orthonormal } \{\phi_i\}_{i=1}^N \subset \mathrm{H}^1(\mathbb{R}^3) \\ \text{with } \Phi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det\big(\phi_i(\mathbf{x}_j)\big)_{i,j=1}^N \Big\}.$$

Hartree-Fock Approximation of Minimisation Problem (4a)

Let the Hartree-Fock minimisation set S_{HF} be defined through Definition 33. We seek the solution(s) $\mathcal{E}^*_{\mathrm{HF}} \in \mathbb{R}$ to the following minimisation problem

$$\mathcal{E}^*_{\mathrm{HF}} := \min_{\Psi \in \mathrm{S}_{\mathrm{HF}}} \langle \Psi, H\Psi \rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}}. \tag{100a}$$

Several remarks are in order.

First, in Equation (100a), the normalisation constraints have been removed since each element of the set S_{HF} is normalised by construction. Second, any function $\Psi^*_{\mathrm{HF}} \in \mathrm{S}_{\mathrm{HF}}$ that achieves the minimum in Equation (100a) is called a Hartree-Fock ground state and obviously satisfies

$$\forall \Psi \in \mathrm{S}_{\mathrm{HF}} \colon \quad \langle \Psi, H\Psi^*_{\mathrm{HF}} \rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}} = \mathcal{E}^*_{\mathrm{HF}} \langle \Psi, \Psi^*_{\mathrm{HF}} \rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}}. \tag{101}$$

Third, concerning the well-posedness of the minimisation problem (100a), the existence of an infimum of the Hartree-Fock energy $\mathcal{E}^*_{\mathrm{HF}}$ is relatively straightforward to establish but the existence of a minimising Slater determinant is a difficult problem that was finally solved by Lieb and Simon [19]. This result was generalised in a celebrated paper by P.L. Lions [20]. The *uniqueness* of such minimisers is still an open question however.

Fourth, a useful simplification occurs in the expression of the Hamiltonian H when it acts on an element of S_{HF} and is projected against another Slater determinant. Indeed, given $\Phi := \frac{1}{\sqrt{N!}} \det\big(\phi_i(\mathbf{x}_j)\big)_{i,j=1}^N \in \mathrm{S}_{\mathrm{HF}}$ we have

$$\begin{aligned} \langle \Phi, H\Phi \rangle_{\widehat{\mathscr{H}}^1 \times \widehat{\mathscr{H}}^{-1}} = &-\frac{1}{2} \sum_{j=1}^{N} \int_{\mathbb{R}^3} |\nabla \phi_j(\mathbf{x})|^2 \, d\mathbf{x} + \sum_{j=1}^{N} \sum_{\alpha=1}^{M} \int_{\mathbb{R}^3} \frac{-Z_\alpha}{|\mathbf{x}_\alpha - \mathbf{x}|} |\phi_j(\mathbf{x})|^2 \, d\mathbf{x} \\ &+ \frac{1}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{x})|^2 |\phi_i(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|} \, d\mathbf{x} d\mathbf{y} \\ &- \frac{1}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\phi_i(\mathbf{x}) \phi_j(\mathbf{x}) \phi_i(\mathbf{y}) \phi_j(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \, d\mathbf{x} d\mathbf{y}. \end{aligned} \tag{102}$$

Equation (102) suggests that we can introduce, in a natural manner, the so-called single-particle Hartree-Fock operator $\mathrm{F}_\Phi \colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R}^3)$.

Definition 34 (Single-Particle Hartree-Fock Operator)**.**

For any $\Phi = \frac{1}{\sqrt{N!}}\det\left(\phi_i(\mathbf{x}_j)\right)_{i,j=1}^N \in \mathrm{S}_{\mathrm{HF}} \subset \widehat{\mathscr{H}}^1$*, we define the single-particle Hartree-Fock (or mean-field) operator* $\mathrm{F}_\Phi \colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R}^3)$ *as the mapping with the property that*

$$\forall \psi \in \mathrm{H}^1(\mathbb{R}^3)\colon \quad \mathrm{F}_\Phi\psi(\mathbf{x}) := -\frac{1}{2}\Delta\psi(\mathbf{x}) + \sum_{\alpha=1}^M \frac{-Z_\alpha}{|\mathbf{x}_\alpha - \mathbf{x}|}\psi(\mathbf{x}) + \sum_{j=1}^N \int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|}\, d\mathbf{y}\psi(\mathbf{x})$$
$$- \sum_{j=1}^N \int_{\mathbb{R}^3} \frac{\phi_j(\mathbf{x})\phi_j(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\psi(\mathbf{y})\, d\mathbf{y} \quad \text{for a.e. } \mathbf{x} \in \mathbb{R}^3.$$

It is readily verified from Kato-Rellich theory (see, e.g., [15, Chapter 13]) that the single-particle Hartree-Fock operator (constructed from any determinant $\Phi \in \mathrm{S}_{\mathrm{HF}}$) is self-adjoint on $\mathrm{L}^2(\mathbb{R}^3)$ with domain $\mathrm{H}^2(\mathbb{R}^3)$ and form domain $\mathrm{H}^1(\mathbb{R}^3)$. Additionally, since the minimum of the Hartree-Fock energy exists, the single-particle Fock operator constructed from the Hartree-Fock determinant $\Psi_{\mathrm{HF}} \in \mathrm{S}_{\mathrm{HF}}$ and modified with a suitable shift is a coercive operator[4].

The Euler-Lagrange equations for the Hartree-Fock minimisation problem (102), i.e., the first-order optimality conditions then read as follows:

Infinite-Dimensional Hartree-Fock Eigenvalue Problem

For any determinant $\Phi \in \mathrm{S}_{\mathrm{HF}}$, let the single-particle Hartree-Fock operator $\mathrm{F}_\Phi \colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R}^3)$ be defined through Definition 34. We seek $\Psi = \frac{1}{\sqrt{N!}}\det\left(\psi_i(\mathbf{x}_j)\right)_{i,j=1}^N \in \mathrm{S}_{\mathrm{HF}} \subset \widehat{\mathscr{H}}^1$ and $\{\lambda_i\}_{i=1}^N \subset \mathbb{R}$ with the property that for all $i = 1, \ldots, N$ it holds that

$$\mathrm{F}_\Psi\psi_i = \lambda_i\psi_i. \tag{103}$$

Equation (103) is a non-linear eigenvalue problem, the Hartree-Fock operator being dependent on the sought-after eigenfunctions $\{\psi_i\}_{i=1}^N$. Partial results about the analysis of this eigenvalue problem are available, and in particular, it is known that any minimiser of the Hartree-Fock minimisation problem (102) satisfies Equation (103) with all $\lambda_i < 0$ (see, e.g., [20]).

Convention 35 (Ordering of Eigenvalues)**.**

Throughout the remainder of this article, we adopt the convention that the eigenvalues of any self-adjoint operator– including, in particular, any single-particle Hartree-Fock operator F_Ψ*,* $\Psi \in \mathrm{S}_{\mathrm{HF}}$ *or discretisations thereof– are ordered ascendingly. Thus, in the case of the eigenvalue problem* (103)*, we have* $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots$.

Before proceeding further, let us point out that the single-particle Hartree-Fock operator $\mathrm{F}_\Phi \colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R}^3)$ can be used to build an N-particle Hartree-Fock operator that acts on the tensor space $\widehat{\mathscr{H}}^1$. Indeed, we have the following definition.

[4]Similar to the electronic Hamiltonian H, the single-particle operator $\mathrm{F}_\Phi, \Phi \in \mathrm{S}_{\mathrm{HF}}$ can be shown to satisfy an ellipticity estimate. This is simply a consequence of the Hardy inequality.

Definition 36 (N-Particle Hartree-Fock Operator)**.**

For any $\Phi = \frac{1}{\sqrt{N!}}\det\left(\phi_i(\mathbf{x}_j)\right)_{i,j=1}^N \in \mathrm{S_{HF}} \subset \widehat{\mathscr{H}}^1$*, we define the* N*-particle Fock operator* $\mathscr{F}_\Phi \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$ *as the mapping with the property that for any* $\Psi = \frac{1}{\sqrt{N!}}\det\left(\psi_{k_i}(\mathbf{x}_j)\right)_{i,j=1}^N \in \mathrm{S_{HF}}$ *it holds that*

$$\left(\mathscr{F}_\Phi\Psi\right)(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\sum_{j=1}^N \sum_{\pi\in\mathrm{S}(N)} (-1)^{\mathrm{sgn}(\pi)}\psi_{\pi(k_1)}(\mathbf{x}_1)\otimes\psi_{\pi(k_2)}(\mathbf{x}_2)\otimes\ldots \\ \ldots\otimes \mathrm{F}_\Phi\psi_{\pi(k_j)}(\mathbf{x}_j)\otimes\ldots\otimes\psi_{\pi(k_N)}(\mathbf{x}_N),$$

and whose action on arbitrary elements of $\widehat{\mathscr{H}}^1$ *is defined through linearity using the fact that any* $\widehat{\mathscr{L}}^2$*-orthonormal Slater basis* $\mathscr{B}_\wedge$ *of* $\widehat{\mathscr{H}}^1$ *is a subset of* $\mathrm{S_{HF}}$*.*

As expected, the N-particle Fock operator inherits many properties from the single-particle Fock operator. In particular, for any determinant $\Phi \in \mathrm{S_{HF}}$, the operator $\mathscr{F}_\Phi$ is self-adjoint on $\widehat{\mathscr{L}}^2$ with domain $\widehat{\mathscr{H}}^2$ and form domain $\widehat{\mathscr{H}}^1$. Moreover, the N-particle Fock operator constructed from the Hartree-Fock determinant $\Psi_{\mathrm{HF}} \in \mathrm{S_{HF}}$ and modified with a suitable shift is also a coercive operator.

Consider once again the non-linear eigenvalue problem (103). The practical resolution of this problem requires the introduction of a finite-dimensional subspace of $\mathrm{H}^1(\mathbb{R}^3)$. Indeed, denoting by X_K any K-dimensional subspace of $\mathrm{H}^1(\mathbb{R}^3)$, we can define:

Definition 37 (Hartree-Fock Finite-Dimensional Minimisation Set)**.**

We define the finite-dimensional set $\mathrm{S}_{\mathrm{HF}}^K$ *as the subset of* $\widehat{\mathscr{H}}^1$ *given by*

$$\mathrm{S}_{\mathrm{HF}}^K := \Big\{\Phi \in \widehat{\mathscr{H}}^1 \colon \text{ there exist } \mathrm{L}^2\text{-orthonormal } \{\phi_i\}_{i=1}^N \subset \mathrm{X}_K \\ \text{with } \Phi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\det\left(\phi_i(\mathbf{x}_j)\right)_{i,j=1}^N\Big\}.$$

A finite-dimensional approximation of the non-linear eigenvalue problem (103) is then given as follows.

Finite-Dimensional Hartree-Fock Eigenvalue Problem

For any determinant $\Phi \in \mathrm{S_{HF}}$, let the single-particle Fock operator $\mathrm{F}_\Phi \colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R}^3)$ be defined through Definition 34. We seek $\Psi = \frac{1}{\sqrt{N!}}\det\left(\psi_i(\mathbf{x}_j)\right)_{i,j=1}^N \in \mathrm{S}_{\mathrm{HF}}^K \subset \widehat{\mathscr{H}}^1$ and $\{\lambda_i\}_{i=1}^N \subset \mathbb{R}$ with the property that for all $\phi \in \mathrm{X}_K \subset \mathrm{H}^1(\mathbb{R}^3)$ and all $i = 1,\ldots,N$ it holds that

$$\langle \mathrm{F}_\Psi\psi_i, \phi\rangle_{\mathrm{H}^{-1}(\mathbb{R}^3)\times\mathrm{H}^1(\mathbb{R}^3)} = \lambda_i\left(\psi_i,\phi\right)_{\mathrm{L}^2(\mathbb{R}^3)}. \tag{104}$$

Let us remark here that since the single-particle Fock operator is self-adjoint on $\mathrm{X}_K \subset \mathrm{H}^1(\mathbb{R}^3)$, it will in fact have K orthonormal eigenfunctions $\{\Psi_i\}_{i=1}^K \subset \mathrm{X}_K$. By convention, the N solution eigenfunctions of Equation (104) are chosen to be the ones with the lowest associated eigenvalues (the so-called Aufbau principle). Thus, the diagonalisation of the single-particle Fock-operator provides an L^2-orthonormal basis for X_K, which can be denoted as $\{\psi_i\}_{i=1}^N \cup \{\psi_i\}_{i=N+1}^K$. A standard approach in electronic structure calculation is to take $\mathscr{R} = \mathrm{span}\{\psi_i\}_{i=1}^N$ as an occupied space (recall Definition 1). We will adopt precisely this view in the forthcoming Appendix B.

Next, let us point out a fact is often exploited in computational quantum chemistry (see, e.g., Møller Plesset perturbation theory [13, Chapter 14]), namely that the N-particle Fock operator defined through Definition 36 induces a decomposition of the electronic Hamiltonian $H \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$. Indeed, we have the following lemma.

Lemma 38 (Electronic Hamiltonian Decomposition and the Fluctuation Potential)**.**

For any $\Phi = \frac{1}{\sqrt{N!}}\det\left(\phi_i(\mathbf{x}_j)\right)_{i,j=1}^N \in \mathrm{S_{HF}} \subset \widehat{\mathscr{H}}^1$*, let the* N*-particle Fock operator* $\mathscr{F}_\Phi\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$ *be defined as in Definition 36 and let the electronic Hamiltonian* $H\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^{-1}$ *be defined through Equation* (3)*. Then we have the decomposition*

$$\mathscr{H} = \mathscr{F}_\Phi + \mathscr{U}_\Phi,$$

where $\mathscr{U}_\Phi := \mathscr{H} - \mathscr{F}_\Phi$ *is known as the Fluctuation potential, and is a bounded linear mapping from* $\widehat{\mathscr{H}}^1$ *to* $\widehat{\mathscr{L}}^2$ *with a continuity constant that depends only on* Φ *and* N*.*

Proof. The fact that $\mathscr{U}_\Phi := \mathscr{H} - \mathscr{F}_\Phi$ is a bounded linear mapping from $\widehat{\mathscr{H}}^1$ to $\widehat{\mathscr{L}}^2$ is essentially a consequence of the following Hardy-type inequalities whose proof can be found in [30, Lemma 1]: For all $u \in \mathscr{C}^\infty_{\mathrm{comp}}(\mathbb{R}^3)$ and $v \in \mathscr{C}^\infty_{\mathrm{comp}}(\mathbb{R}^6)$ it holds that

$$\int_{\mathbb{R}^3} \frac{u^2(\mathbf{x})}{|\mathbf{x}|^2}\, d\mathbf{x} \leq 4 \int_{\mathbb{R}^3} |\nabla u(\mathbf{x})|^2\, d\mathbf{x}, \tag{105}$$

$$\int_{\mathbb{R}^6} \frac{v^2(\mathbf{x},\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^2}\, d\mathbf{x}d\mathbf{y} \leq 2 \int_{\mathbb{R}^6} |\nabla_{\mathbf{x}} v(\mathbf{x},\mathbf{y})|^2 + |\nabla_{\mathbf{y}} v(\mathbf{x},\mathbf{y})|^2\, d\mathbf{x}d\mathbf{y}. \tag{106}$$

Equipped with the Hardy-type inequalities (105) and (106), we will now show that there exists a constant $C_{\Phi,N} > 0$ such that for all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$ it holds that

$$(\mathscr{U}_\Phi \Psi_H, \Psi_L)_{\widehat{\mathscr{L}}^2} \leq C_{\Phi,N}\, \|\Psi_H\|_{\widehat{\mathscr{H}}^1}\, \|\Psi_L\|_{\widehat{\mathscr{L}}^2} \tag{107}$$

To this end, let us recall Equation (3) and Definitions 34 and 36 to deduce that for any $\Psi_H \in \widehat{\mathscr{H}}^1$, we have that

$$\begin{aligned}
\mathscr{U}_\Phi \Psi_H =& \mathscr{H}\Psi_H - \mathscr{F}_\Phi \Psi_H \\
=& \underbrace{\sum_{i=1}^N \sum_{j=1}^{i-1} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \Psi_H(\mathbf{x}_1,\ldots,\mathbf{x}_N)}_{:=\mathscr{U}_1\Psi_H} - \underbrace{\sum_{i=1}^N \sum_{j=1}^N \int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{y})|^2}{|\mathbf{x}_i - \mathbf{y}|}\, d\mathbf{y} \Psi_H(\mathbf{x}_1,\ldots,\mathbf{x}_N)}_{:=\mathscr{U}_2\Psi_H} \\
&+ \underbrace{\sum_{i=1}^N \sum_{j=1}^N \int_{\mathbb{R}^3} \frac{\phi_j(\mathbf{x}_i)\phi_j(\mathbf{y})}{|\mathbf{x}_i - \mathbf{y}|} \Psi_H(\mathbf{x}_1,\cdots,\mathbf{x}_{i-1},\mathbf{y},\mathbf{x}_{i+1},\cdots,\mathbf{x}_N)\, d\mathbf{y}}_{:=\mathscr{U}_3\Psi_H} \quad \text{for a.e. } (\mathbf{x}_1,\ldots,\mathbf{x}_N) \in \mathbb{R}^{3N}.
\end{aligned} \tag{108}$$

Thus, the fluctuation potential $\mathscr{U}_\Phi$ can be written as the sum of three terms, which we denote $\mathscr{U}_\ell,\ \ell \in \{1,2,3\}$. We will now prove that for each $\ell \in \{1,2,3\}$, there exists a constant $K^\ell_{N,\Phi} > 0$ such that for all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$ it holds that

$$(\mathscr{U}_\ell \Psi_H, \Psi_L)_{\widehat{\mathscr{L}}^2} \leq K^\ell_{N,\Phi}\, \|\Psi_H\|_{\widehat{\mathscr{H}}^1}\, \|\Psi_L\|_{\widehat{\mathscr{L}}^2}. \tag{109}$$

Let us begin by considering the operator $\mathcal{U}_1$. Observe that for any $1 \leq j < i \leq N$, all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$, and a.e. $(\mathbf{x}_1, \dots, \mathbf{x}_N) \in \mathbb{R}^{3N}$ it holds that

$$\begin{aligned}
&\int_{\mathbb{R}^6} \frac{\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)}{|\mathbf{x}_i - \mathbf{x}_j|} \Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)\, d\mathbf{x}_j d\mathbf{x}_i \\
\leq &\Big(\int_{\mathbb{R}^6} \frac{|\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2}{|\mathbf{x}_i - \mathbf{x}_j|^2}\, d\mathbf{x}_j d\mathbf{x}_i \Big)^{\frac{1}{2}} \Big(\int_{\mathbb{R}^6} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_j d\mathbf{x}_i \Big)^{\frac{1}{2}} \\
\leq &\sqrt{2} \Big(\int_{\mathbb{R}^6} |\nabla_{\mathbf{x}_j} \Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2 + |\nabla_{\mathbf{x}_i} \Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_j \mathbf{x}_i \Big)^{\frac{1}{2}} \Big(\int_{\mathbb{R}^6} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_j \mathbf{x}_i \Big)^{\frac{1}{2}},
\end{aligned}$$

where the last step follows from the Hardy Inequality (106). Using now Fubini's theorem and the Cauchy-Schwarz inequality together with the shorthand notation

$$d\mathbf{x}_r^{ij} := d\mathbf{x}_1 \cdots d\mathbf{x}_{j-1} d\mathbf{x}_{j+1} \cdots d\mathbf{x}_{i-1} d\mathbf{x}_{i+1} \cdots d\mathbf{x}_N,$$

we can further deduce that for any $1 \leq j < i \leq N$, all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$ it holds that

$$\begin{aligned}
&\int_{\mathbb{R}^{3N-6}} \int_{\mathbb{R}^6} \frac{\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)}{|\mathbf{x}_i - \mathbf{x}_j|} \Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)\, d\mathbf{x}_j d\mathbf{x}_i d\mathbf{x}_r^{ij} \\
\leq &\sqrt{2} \int_{\mathbb{R}^{3N-6}} \left(\int_{\mathbb{R}^6} |\nabla_{\mathbf{x}_j} \Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2 + |\nabla_{\mathbf{x}_i} \Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_j d\mathbf{x}_i \right)^{\frac{1}{2}} \\
&\qquad \cdot \left(\int_{\mathbb{R}^6} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_j d\mathbf{x}_i \right)^{\frac{1}{2}} d\mathbf{x}_r^{ij} \\
\leq &\sqrt{2} \left(\|\nabla_{\mathbf{x}_j} \Psi_H\|_{\widehat{\mathscr{L}}^2}^2 + \|\nabla_{\mathbf{x}_i} \Psi_H\|_{\widehat{\mathscr{L}}^2}^2 \right)^{\frac{1}{2}} \|\Psi_L\|_{\widehat{\mathscr{L}}^2},
\end{aligned}$$

with the last step following from a Cauchy-Schwarz inequality applied to the outer integral.

Recalling now the definition of operator $\mathcal{U}_1$, we can conclude from a basic combinatorial argument that for all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$ it holds that

$$\begin{aligned}
(\mathcal{U}_1 \Psi_H, \Psi_L)_{\widehat{\mathscr{L}}^2} \leq &\sqrt{2} \|\Psi_L\|_{\widehat{\mathscr{L}}^2} \sum_{i=1}^{N} \sum_{j=1}^{i-1} \left(\|\nabla_{\mathbf{x}_j} \Psi_H\|_{\widehat{\mathscr{L}}^2}^2 + \|\nabla_{\mathbf{x}_i} \Psi_H\|_{\widehat{\mathscr{L}}^2}^2 \right)^{\frac{1}{2}} \\
\leq &\sqrt{N}(N-1) \|\nabla \Psi_H\|_{\widehat{\mathscr{L}}^2} \|\Psi_L\|_{\widehat{\mathscr{L}}^2},
\end{aligned} \tag{110}$$

which shows that the bound (109) indeed holds for the operator $\mathcal{U}_1$.

Next, we focus on the term $\mathcal{U}_2$. Once again, we observe that for any $i \leq j \leq N$ and a.e. $\mathbf{x}_i \in \mathbb{R}^3$ it holds that

$$\begin{aligned}
\int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{y})|^2}{|\mathbf{x}_i - \mathbf{y}|}\, d\mathbf{y} \leq \left(\int_{\mathbb{R}^3} |\phi_j(\mathbf{y})|^2\, d\mathbf{y} \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{y})|^2}{|\mathbf{x}_i - \mathbf{y}|^2}\, d\mathbf{y} \right)^{\frac{1}{2}} &\leq \left(\int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{y})|^2}{|\mathbf{x}_i - \mathbf{y}|^2}\, d\mathbf{y} \right)^{\frac{1}{2}} \\
&\leq 2 \left(\int_{\mathbb{R}^3} |\nabla \phi_j(\mathbf{y})|^2\, d\mathbf{y} \right)^{\frac{1}{2}},
\end{aligned}$$

where the second inequality results from the fact that $\|\phi_j\|_{\mathrm{L}^2} = 1$ for all $j \in \{1, \dots, N\}$ by assumption, and the third inequality follows from the Hardy inequality (105). We therefore deduce that

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{y})|^2}{|\mathbf{x}_i - \mathbf{y}|}\, d\mathbf{y} \leq 2N \sum_{j=1}^{N} \|\nabla \phi_j\|_{\mathrm{L}^2},$$

and consequently, for all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$ it holds that

$$\begin{aligned}(\mathcal{U}_2\Psi_H, \Psi_L)_{\widehat{\mathscr{L}}^2} \leq & 2N \sum_{j=1}^{N} \|\nabla\phi_j\|_{\mathrm{L}^2} \int_{\mathbb{R}^{3N}} \left|\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)\right| \, d\mathbf{x}_1, \cdots, d\mathbf{x}_N \\ \leq & 2N \sum_{j=1}^{N} \|\nabla\phi_j\|_{\mathrm{L}^2} \|\Psi_H\|_{\widehat{\mathscr{L}}^2} \|\Psi_L\|_{\widehat{\mathscr{L}}^2}.\end{aligned} \tag{111}$$

We have thus shown that the bound (109) also holds for the operator $\mathcal{U}_2$.

It remains to consider the final term $\mathcal{U}_3$ that appears in the decomposition of the fluctuation potential. To this end, we once again observe that for any $1 \leq i, j \leq N$, all $\Psi_H \in \widehat{\mathscr{H}}^1$, and a.e. $(\mathbf{x}_1, \ldots, \mathbf{x}_N) \in \mathbb{R}^{3N}$ it holds that

$$\begin{aligned}&\int_{\mathbb{R}^3} \frac{\phi_j(\mathbf{x}_i)\phi_j(\mathbf{y})}{|\mathbf{x}_i - \mathbf{y}|} \Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N) \, d\mathbf{y} \\ \leq & |\phi_j(\mathbf{x}_i)| \left(\int_{\mathbb{R}^3} \frac{|\phi_j(\mathbf{y})|^2}{|\mathbf{x}_i - \mathbf{y}|^2} \, d\mathbf{y}\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N)|^2 \, d\mathbf{y}\right)^{\frac{1}{2}} \\ \leq & 2|\phi_j(\mathbf{x}_i)| \, \|\nabla\phi_j\|_{\mathrm{L}^2} \left(\int_{\mathbb{R}^3} |\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N)|^2 \, d\mathbf{y}\right)^{\frac{1}{2}},\end{aligned}$$

where the last step follows again from the Hardy inequality (105).

Consequently, for any $1 \leq i, j \leq N$, all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$, and a.e. $(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_N) \in \mathbb{R}^{3N-3}$, we can use Fubini's theorem and the Cauchy-Schwarz inequality to deduce that

$$\begin{aligned}&\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\phi_j(\mathbf{x}_i)\phi_j(\mathbf{y})}{|\mathbf{x}_i - \mathbf{y}|} \Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N) \, d\mathbf{y} \Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N) \, d\mathbf{x}_i \\ \leq & \int_{\mathbb{R}^3} 2|\phi_j(\mathbf{x}_i)| \, \|\nabla\phi_j\|_{\mathrm{L}^2} \left(\int_{\mathbb{R}^3} |\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N)|^2 \, d\mathbf{y}\right)^{\frac{1}{2}} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)| \, d\mathbf{x}_i \\ \leq & 2\|\nabla\phi_j\|_{\mathrm{L}^2} \left(\int_{\mathbb{R}^3} |\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N)|^2 \, d\mathbf{y}\right)^{\frac{1}{2}} \\ & \qquad \cdot \left(\int_{\mathbb{R}^3} |\phi_j(\mathbf{x}_i)|^2 \, d\mathbf{x}_i\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2 \, d\mathbf{x}_i\right)^{\frac{1}{2}} \\ = & 2\|\nabla\phi_j\|_{\mathrm{L}^2} \left(\int_{\mathbb{R}^3} |\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N)|^2 \, d\mathbf{y}\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2 \, d\mathbf{x}_i\right)^{\frac{1}{2}},\end{aligned}$$

where the last simplification results from the fact that $\|\phi_j\|_{\mathrm{L}^2} = 1$ for all $j \in \{1, \ldots, N\}$ by assumption. Using again the shorthand notation

$$d\mathbf{x}_r^i := d\mathbf{x}_1 \cdots d\mathbf{x}_{i-1} d\mathbf{x}_{i+1} \cdots d\mathbf{x}_N,$$

together with Fubini's theorem and the Cauchy-Schwarz inequality, we can conclude that for any $1 \leq j < i \leq N$, all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$ it holds that

$$\begin{aligned}
&\int_{\mathbb{R}^{3(N-1)}} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\phi_j(\mathbf{x}_i)\phi_j(\mathbf{y})}{|\mathbf{x}_i - \mathbf{y}|} \Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N)\, d\mathbf{y} \Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)\, d\mathbf{x}_i d\mathbf{x}_r^i \\
\leq & 2\|\nabla\phi_j\|_{\mathrm{L}^2} \int_{\mathbb{R}^{3(N-1)}} \left(\int_{\mathbb{R}^3} |\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_N)|^2\, d\mathbf{y} \right)^{\frac{1}{2}} \\
& \cdot \qquad \left(\int_{\mathbb{R}^3} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_i \right)^{\frac{1}{2}} d\mathbf{x}_r^i \\
\leq & 2\|\nabla\phi_j\|_{\mathrm{L}^2} \left(\int_{\mathbb{R}^{3(N-1)}} \int_{\mathbb{R}^3} |\Psi_H(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_i d\mathbf{x}_r^i \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^{3(N-1)}} \int_{\mathbb{R}^3} |\Psi_L(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2\, d\mathbf{x}_i d\mathbf{x}_r^i \right)^{\frac{1}{2}} \\
= & 2\|\nabla\phi_j\|_{\mathrm{L}^2}\, \|\Psi_H\|_{\widehat{\mathscr{L}}^2}\, \|\Psi_L\|_{\widehat{\mathscr{L}}^2}.
\end{aligned}$$

Recalling now the definition of operator $\mathcal{U}_3$, we see that for all $\Psi_H \in \widehat{\mathscr{H}}^1$ and $\Psi_L \in \widehat{\mathscr{L}}^2$ it holds that

$$(\mathcal{U}_3\Psi_H, \Psi_L)_{\widehat{\mathscr{L}}^2} \leq 2N \sum_{j=1}^{N} \|\nabla\phi_j\|_{\mathrm{L}^2}\, \|\Psi_H\|_{\widehat{\mathscr{L}}^2}\, \|\Psi_L\|_{\widehat{\mathscr{L}}^2}, \tag{112}$$

which demonstrates that the bound (109) also holds for the operator $\mathcal{U}_3$.

Combining now Inequalities (110), (111) and (112) yields the continuity of the fluctuation potential $\mathcal{U}_\Phi$ as a mapping from $\widehat{\mathscr{H}}^1$ to $\widehat{\mathscr{L}}^2$ with a continuity constant that depends only on N and $\sum_{j=1}^N \|\nabla\phi_j\|_{\mathrm{L}^2}$. Let us remark here that since $\Phi = \frac{1}{\sqrt{N!}}\det\left(\phi_i(\mathbf{x}_j)\right)_{i,j=1}^N$, it is easy to show that $\sum_{j=1}^N \|\nabla\phi_j\|_{\mathrm{L}^2} \lesssim \|\Phi\|_{\widehat{\mathscr{H}}^1}$. □

Appendix B. Approximation Spaces Satisfying Structure Assumptions B.I and B.II

The goal of this section is to state some prototypical examples of coupled cluster approximation spaces that satisfy **Structure Assumptions B.I and B.II**. Throughout this section, we assume the settings of Sections 3 and 4 and we recall the Hartree-Fock method introduced in Appendix A.

We begin by taking an $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis of $\mathrm{H}^1(\mathbb{R}^3)$ which we denote as $\{\phi_j\}_{j\in\mathbb{N}}$, and we define, for every $K \geq N$ the set $\mathcal{B}_K := \{\phi_j\}_{j=1}^K$ and the space $\mathrm{X}_K := \operatorname{span} \mathcal{B}_K$.

Equipped with the sequence of spaces $\{\mathrm{X}_K\}_{K\geq N}$ we define, for each $K \geq N$, the Hartree-Fock finite-dimensional minimisation set $\mathrm{S}_{\mathrm{HF}}^K$ according to Definition 37, and we consider the following finite-dimensional Hartree-Fock eigenvalue problem (c.f., Equation (104)):

Find $\Psi_K = \frac{1}{\sqrt{N!}}\det\left(\psi_i^K(\mathbf{x}_j)\right)_{i,j=1}^N \in \mathrm{S}_{\mathrm{HF}}^K \subset \widehat{\mathscr{H}}^1$ and $\{\lambda_i^K\}_{i=1}^N \subset \mathbb{R}$ with the property that for all $\phi \in \mathrm{X}_K$ and all $j = 1, \ldots, N$ it holds that

$$\left\langle \mathrm{F}_{\Psi_K}\psi_j^K, \phi \right\rangle_{\mathrm{H}^{-1}(\mathbb{R}^3)\times\mathrm{H}^1(\mathbb{R}^3)} = \lambda_j^K \left(\psi_j^K, \phi\right)_{\mathrm{L}^2(\mathbb{R}^3)}, \tag{113}$$

where $\mathrm{F}_{\Psi_K} \colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R})$ denotes the single-particle Fock operator defined through Definition 34.

We emphasise that, as described in Appendix A, for each $K \geq N$, the eigenfunctions $\{\psi_j^K\}_{j=1}^K$ of the restricted single-particle Fock operator $\mathrm{F}_{\Psi_K} \colon \mathrm{X}_K \to \mathrm{X}_K^*$ form an L^2-orthonormal basis of X_K, and

the N solution eigenfunctions of Equation (113) are chosen according to the Aufbau principle, i.e., as the ones with the lowest eigenvalues.

Next, we introduce, for every $K \geq N$, the index set $\mathcal{I}_K^N \subset \{1, \ldots, K\}^N$ given by

$$\mathcal{I}_K^N := \Big\{ \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N) \in \{1, \ldots, K\}^N \colon \alpha_1 < \alpha_2 < \ldots < \alpha_N \Big\}.$$

We now have the following definition.

Definition 39 (Finite Dimensional N-Particle Basis)**.**

We define the $\widehat{\mathcal{L}}^2$-orthonormal, $\binom{K}{N}$-dimensional N-particle basis $\mathcal{B}_K^N \subset \widehat{\mathcal{H}}^1$ as

$$\mathcal{B}_K^N := \left\{ \Psi_\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det\big(\psi_{\alpha_i}^K(\mathbf{x}_j)\big)_{i,j=1}^N \colon \ \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N) \in \mathcal{I}_K^N \right\}. \tag{114}$$

Additionally, we define the subspace spanned by this basis set as $\mathcal{Z}_K := \operatorname{span} \mathcal{B}_K^N$.

The so-called Full-CC discretisations based on canonical Hartree-Fock orbitals now consist of defining, for each $K \geq N$,

$$\Psi_{0,K} := \frac{1}{\sqrt{N!}} \det\left(\psi_j^K(\mathbf{x}_i)\right)_{i,j=1}^N \quad \text{with } \{\psi_j^K\}_{j=1}^N \text{ denoting the Aufbau solutions to Equation (113),}$$

$$\widetilde{\mathcal{V}}_K := \{\Psi_{0,K}\}^\perp \cap \mathcal{Z}_K, \qquad \text{and} \qquad \mathcal{V}_K := \widetilde{\mathcal{V}}_K \oplus \operatorname{span}\{\Psi_{0,K}\}. \tag{115}$$

A direct calculation shows (c.f., [11, Remark 36]) that the Full-CC approximation spaces $\{\mathcal{V}_K\}_{K \geq N}$ defined in the above manner indeed satisfy **Structure Assumption B.I**. Let us remark here that the above construction of Full-CC approximation spaces satisfies **Structure Assumption B.I**, even if the canonical Hartree-Fock orbitals $\{\psi_j^K\}_{j=1}^K$ are rotated through a unitary transformation– as long as the unitary transformation preserves the $\mathrm{L}^2(\mathbb{R}^3)$ orthogonality of the set of so-called occupied orbitals $\{\psi_j^K\}_{j=1}^N$ and the set of so-called virtual orbitals $\{\psi_j^K\}_{j=N=1}^K$.

We now turn our attention to coupled cluster approximation spaces that satisfy **Structure Assumption B.II**. To this end, we recall again the sequence of N-particle basis sets $\{\mathcal{B}_K^N\}_{K \geq N}$ and the sequence of finite-dimensional spaces $\{\mathcal{Z}_K\}_{K \geq N}$ defined through Definition 39. Next, we define a so-called excitation index set and appropriate excitation operators to characterise the Slater determinants in basis set $\mathcal{B}_K^N$

Definition 40 (Excitation Index Sets For Finite Bases)**.**

For each $K \geq N$ and each $Q \in \{1, \ldots, N\}$ we define the index set $\mathcal{I}_Q^K$ as

$$\mathcal{I}_Q^K := \left\{ \binom{a_1, \ldots, a_Q}{\ell_1, \ldots, \ell_Q} \colon \ell_1 < \ldots < \ell_Q \in \{1, \ldots, N\} \text{ and } a_1 < \ldots < a_j \in \{N+1, \ldots, K\} \right\}.$$

Definition 41 (Excitation Operators for Finite Bases)**.**

Let $K \geq N$, let $Q \in \{1, \ldots, N\}$, let $\mu \in \mathcal{I}_Q^K$ be of the form

$$\mu = \binom{a_1, \ldots, a_Q}{\ell_1, \ldots, \ell_Q} \colon \ell_1 < \ldots < \ell_Q \in \{1, \ldots, N\} \text{ and } a_1 < \ldots < a_Q \in \{N+1, \ldots, K\},$$

and let the N-particle basis set $\mathcal{B}_K^N$ and finite-dimensional space $\mathcal{Z}_K$ be defined as in Definition 39.

We define the excitation operator $\mathcal{X}_{K,\mu}\colon \mathcal{Z}_K \to \mathcal{Z}_K$ *through its action on the* N*-particle basis set* $\mathcal{B}_K^N$*: For* $\Psi_\nu(\mathbf{x}_1,\dots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\det\left(\psi_{\nu_j}^K(\mathbf{x}_i)\right)_{i,j=1}^N \in \mathcal{B}_K^N$*, we set*

$$\mathcal{X}_{K,\mu}\Psi_\nu = \begin{cases} 0 & \text{if } \{\ell_1,\dots,\ell_Q\} \not\subset \{\nu_1,\dots,\nu_N\}, \\ 0 & \text{if } \exists a_m \in \{a_1,\dots,a_Q\} \text{ such that } a_m \in \{\nu_1,\dots,\nu_N\}, \\ \Psi_{\nu,a} \in \mathcal{B}_\wedge & \text{otherwise,} \end{cases}$$

where the determinant $\Psi_{\nu,a}$ *is constructed from* Ψ_ν *by replacing all functions* $\psi_{\ell_1}^K,\dots,\psi_{\ell_Q}^K$ *used to construct* Ψ_ν *with functions* $\psi_{a_1}^K,\dots\psi_{a_Q}^K$ *respectively.*

Equipped with Definitions 40 and 41 , we are now ready to define the class of coupled cluster equations that can reasonably be expected to satisfy **Structure Assumption B.II**: The so-called truncated-CC discretisations of excitation rank Q, based on canonical Hartree-Fock orbitals consist of defining, for each $K \geq N$ and some given $Q \in \{1,\dots,N\}$

$$\Psi_{0,K} := \frac{1}{\sqrt{N!}}\det\left(\psi_j^K(\mathbf{x}_i)\right)_{i,j=1}^N \quad \text{with } \{\psi_j^K\}_{j=1}^N \text{ denoting the Aufbau solutions to Equation (113),}$$

(116)

$$\widetilde{\mathcal{V}}_K := \text{span}\left\{\Psi_\nu \in \mathcal{B}_K^N \colon \exists j \in \{1,\dots,Q\},\ \mu \in \mathcal{I}_j^K \text{ such that } \mathcal{X}_{K,\mu}\Psi_{0,K} = \Psi_\nu\right\},$$

$$\mathcal{V}_K := \widetilde{\mathcal{V}}_K \oplus \text{span}\{\Psi_{0,K}\}, \qquad \mathcal{W}_K := \mathcal{Z}_K, \quad \text{and} \quad \mathcal{F}_K := \begin{cases} \mathcal{F}_{\Psi_{0,K}} - \sum_{j=1}^N \lambda_j^K & \text{on } \mathcal{H}^{1,\perp}_{\Psi_{0,K}} \\ \mathrm{I} & \text{on } \text{span}\{\Psi_{0,K}\} \end{cases},$$

where $\mathcal{F}_{\Psi_{0,K}}\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ is the N-particle Hartree-Fock operator constructed from the reference determinant $\Psi_{0,K}$ as defined in Definition 36 and $\{\lambda_j^K\}_{j=1}^N$ denote the N lowest eigenvalues of the finite-dimensional Hartree-Fock eigenvalue problem (113) in the space X_K. Let us remark, by way of examples, that the classical CCSD method corresponds to taking $Q = 2$, the CCSDT method corresponds to taking $Q = 3$ and so on.

As before, a direct calculation confirms that the truncated-CC approximation spaces $\{\mathcal{V}_K\}_{K\geq N}$ defined in the above manner satisfy the first two conditions of **Structure Assumption B.II**, and this is the case even if the canonical Hartree-Fock orbitals $\{\psi_j^K\}_{j=1}^K$ are rotated through a unitary transformation– as long as the unitary transformation preserves the $\mathrm{L}^2(\mathbb{R}^3)$ orthogonality of the set of so-called occupied orbitals $\{\psi_j^K\}_{j=1}^N$ and the set of so-called virtual orbitals $\{\psi_j^K\}_{j=N=1}^K$.

Coming now to the third condition of **Structure Assumption B.II**, we see that for every $K \geq N$, the mean-field operator $\mathcal{F}_K$ is defined as $\mathcal{F}_K := \mathcal{F}_{\Psi_{0,K}} - \sum_{j=1}^N \lambda_j^K$ on $\mathcal{H}^{1,\perp}_{\Psi_{0,K}}$. Recall now that we have the decomposition of the electronic Hamiltonian $H := \mathcal{F}_{\Psi_{0,K}} + \mathcal{U}_{\Psi_{0,K}}$, and we have shown in Lemma 38 that $\mathcal{U}_{\Psi_{0,K}}\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{L}}^2$ is a bounded operator with continuity constant depending only on N and $\Psi_{0,K}$. Consequently, under the assumption that the N discrete Aufbau eigenfunctions and corresponding Aufbau eigenvalues of the sequence of finite-dimensional Hartree-Fock eigenvalue problems (113) converge in the complete basis set limit $K \to \infty$, we have uniform boundedness of the continuity and coercivity constants of the sequence of mean-field operators $\{\mathcal{F}_K\}_{K\geq N}$. Thus the third condition of **Structure Assumption B.II** will be satisfied.

Let us incidentally remark that while the convergence of the discrete Aufbau Hartree-Fock eigenpairs in the complete basis set limit is a reasonable assumption, to the best of our knowledge, no formal proof of this fact or precise conditions under which it holds, have been stated in the literature. Note that this convergence assumption will also guarantee that the prior **Assumption A.II** concerning the convergence of the reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ to a limiting reference determinant Ψ_0^* holds.

Finally, let us consider the fourth and final condition of **Structure Assumption B.II**. We have already demonstrated in Lemma 38 that $\mathcal{U}_{\Psi_{0,K}}\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{L}}^2$ is a bounded operator with continuity constant depending only on N and $\Psi_{0,K}$, and thus, under the same assumption that the discrete Aufbau Hartree-Fock eigenpairs converge in the complete basis set limit $K \to \infty$ (c.f., **Assumption A.II**), we have uniform boundedness of the continuity constants of the sequence of operators $\{\mathcal{U}_K\}_{K\geq N}$ with each $\mathcal{U}_K := \mathcal{U}_{\Psi_{0,K}}$. It remains to study the additional smallness assumption that we impose on the continuity constant of each $\mathcal{U}_K$, namely (see also, Inequality (75)),

$$\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\widetilde{\mathcal{U}}_K\mathbb{P}_{\mathcal{F}_K}\right\|_{\mathcal{F}_K\to\widehat{\mathscr{L}}^2} < \frac{\sqrt{\Lambda^{\mathcal{F}_K}_{\min}}\,\Gamma^*_{\mathrm{GS}}}{\beta_K(\Theta^*_{K,\mathrm{GS}})}, \tag{117}$$

where

$$\Lambda^{\mathcal{F}_K}_{\min} := \min_{\substack{\Phi_K\in\mathcal{W}_K\\ \Phi_K\in\mathcal{V}_K^{\perp}}} \frac{|\langle\Phi_K, \mathcal{F}_K\Phi_K\rangle_{\widehat{\mathscr{H}}^1\times\widehat{\mathscr{H}}^{-1}}|^{\frac{1}{2}}}{\|\Phi_K\|_{\widehat{\mathscr{L}}^2}},$$

$$\Gamma^*_{\mathrm{GS}} := \min_{\Phi_K\in\{\Psi^*_{\mathrm{GS},K}\}^{\perp}\cap\mathcal{W}_K} \frac{\langle\Phi_K, (H-\mathcal{E}_{\mathrm{GS}^*})\Phi_K\rangle_{\widehat{\mathscr{H}}^1\times\widehat{\mathscr{H}}^{-1}}}{\|\Phi_K\|^2_{\mathcal{F}_K}}, \text{and}$$

$$\beta_K(\Theta^*_{K,\mathrm{GS}}) := \left\|\mathbb{P}^{\perp}_{\mathcal{F}_K} e^{\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K} e^{-\mathcal{T}(\Theta^{\Pi}_{K,\mathrm{GS}})}\mathbb{P}_{\mathcal{F}_K}\right\|_{\mathcal{F}_K\to\mathcal{F}_K}.$$

As stated earlier in Remark 31, the validity of Inequality (117) is, in general, problem dependent and not a universal property that holds for arbitrary molecules. Indeed, it can readily be seen that the constant Γ^*_{GS} depends on the spectral gap of the electronic Hamiltonian H that describes the molecule under study. Similarly, the constant $\beta_K(\Theta^*_{K,\mathrm{GS}})$ can be shown to be $\mathcal{O}\big(\|\mathcal{T}(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1\to\widehat{\mathscr{H}}^1}\big)$ for $\|\mathcal{T}(\Theta^*_{K,\mathrm{GS}})\|_{\widehat{\mathscr{H}}^1\to\widehat{\mathscr{H}}^1} < 1$, and therefore reflects how well the reference determinant $\Psi_{0,K}$ approximates the sought-after ground state eigenfunction $\Psi^*_{\mathrm{GS},K}$ of the electronic Hamiltonian H. Finally, as is well known, the single particle Hartree-Fock operator defined through Definition 34 possesses an essential spectrum consisting of $[0,\infty)$ and therefore the constant $\Lambda^{\mathcal{F}_K}_{\min}$, will <u>not</u> diverge to $+\infty$ in the complete basis set limit $K\to\infty$.

Having said this, in the specific case of the truncated-CC approximation spaces constructed as in Equation (116), the constant $\Lambda^{\mathcal{F}_K}_{\min}$ can reasonably be expected to be fairly large in comparison to the operator norm $\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\widetilde{\mathcal{U}}_K\mathbb{P}_{\mathcal{F}_K}\right\|_{\mathcal{F}_K\to\widehat{\mathscr{L}}^2}$, at least for large excitation truncation ranks Q. Indeed, it follows directly from the definition of $\Lambda^{\mathcal{F}_K}_{\min}$, the mean-field operator $\mathcal{F}_K$ and the excitation rank-truncated approximation space $\widetilde{\mathcal{V}}_K$ that for any $K \geq N+Q+1$, it holds that

$$\Lambda^{\mathcal{F}_K}_{\min} := \min_{\substack{\Phi_K\in\mathcal{W}_K\\ \Phi_K\in\mathcal{V}_K^{\perp}}} \frac{|\langle\Phi_K, \mathcal{F}_K\Phi_K\rangle_{\widehat{\mathscr{H}}^1\times\widehat{\mathscr{H}}^{-1}}|^{\frac{1}{2}}}{\|\Phi_K\|_{\widehat{\mathscr{L}}^2}} = \sum_{j=1}^{N-Q-1}\lambda^K_j + \sum_{j=N+1}^{N+Q+1}\lambda^K_j - \sum_{j=1}^{N}\lambda^K_j, \tag{118}$$

where $\{\lambda^K_j\}_{j=1}^{N+Q+1}$ denote the $N+Q+1$ lowest eigenvalues of the finite-dimensional Hartree-Fock eigenvalue problem (113) in the space X_K.

Equation (118) is simply a consequence of the fact the space $\mathcal{W}_K\cap\mathcal{V}_K^{\perp}$ is spanned by Slater determinants from $\mathcal{B}^N_K$ that can be written as excitations of the reference determinant $\Psi_{0,K}$ of rank at least $Q+1$. Since the Slater determinants in $\mathcal{B}^N_K$ are, additionally, discrete eigenfunctions of the N-particle Hartree-Fock operator $\mathcal{F}_{\Psi_{0,K}}$, and we have adopted the convention of ordering all eigenvalues in ascending order, it is relatively simple to see that the Slater determinant that achieves the minimum in the above

equation must be given by

$$\Psi_{K,\min} = \psi_1^K \wedge \psi_2^K \wedge \ldots \psi_{N-Q-1}^K \wedge \ldots \psi_{N+1}^K \wedge \psi_{N+2}^K \wedge \ldots \wedge \psi_{N+Q+1}^K,$$

which immediately leads to Equation (118). Notice that we can equivalently write Equation (118) as

$$\Lambda_{\min}^{\mathcal{F}_K} = \sum_{j=N+1}^{N+Q+1} \lambda_j^K - \sum_{j=N-Q}^{N} \lambda_j^K = \sum_{j=1}^{Q+1} \varepsilon_j^K,$$

where each $\varepsilon_j^K := \lambda_{N+j}^K - \lambda_{N-j+1}^K$ is the j^{th} discrete excitation energy associated with the N-particle Hartree-Fock operator $\mathcal{F}_{\Psi_{0,K}}$. Note that if the discrete Hartree-Fock eigenvalues $\{\lambda_j^K\}_{j=1}^{2N}$ converge to some limiting eigenvalues $\{\lambda_j^0\}_{j=1}^{2N}$, then we can, of course, replace Estimate (118) with a K-independent bound.

We end this section with some numerical tests that explore the validity of Inequality (117) in different regimes. We begin by considering the most common excitation rank-truncated coupled cluster discretisations, namely, the CCSD (truncation at rank two) and CCSDT (truncation at rank three) discretisations with canonical Hartree-Fock orbitals. Our results for a collection of small molecules at equilibrium geometries using the minimal STO-6G basis sets are displayed in Tables 2 and 3. For reference, we include the theoretical continuous inf-sup constants corresponding to Theorem 19. We also include the numerically computed estimate of the discrete inf-sup constant given by Equation (59), as well as the corresponding theoretical estimate resulting from our analysis (see the proof of Lemma 26), which is given by

$$\gamma_{\text{inf-sup}} := \frac{\Gamma_{\text{GS}}^{\mathcal{F}_K} - \left\| \mathbb{P}_{\mathcal{F}_K}^{\perp} \widetilde{u}_K \mathbb{P}_{\mathcal{F}_K} \right\|_{\mathcal{F}_K \to \widehat{\mathcal{L}}^2} \frac{\beta_K(\Theta_{K,\text{GS}}^*)}{\sqrt{\Lambda_{\min}^{\mathcal{F}_K}}} - \| \left(H - \mathcal{E}_{\text{GS}}^*\right) \Psi_{0,K} \|_{\mathcal{F}_K^{-1}}}{\left\| \mathbb{P}_{\mathcal{F}_K}^{\perp} e^{-\mathcal{T}(\Theta_{K,\text{GS}}^{\Pi})} \right\|_{\mathcal{F}_K \to \mathcal{F}_K} \left\| e^{\mathcal{T}(\Theta_{K,\text{GS}}^{\Pi})^{\dagger}} \right\|_{\mathcal{F}_K \to \mathcal{F}_K}}.$$

Molecule	$\left\| \mathbb{P}_{\mathcal{F}_K}^{\perp} \widetilde{u}_K \mathbb{P}_{\mathcal{F}_K} \right\|_{\mathcal{F}_K \to \widehat{\mathcal{L}}^2}$	$\frac{\sqrt{\Lambda_{\min}^{\mathcal{F}_K}}\, \Gamma_{\text{GS}}^*}{\beta_K(\Theta_{K,\text{GS}}^*)}$	Continuous inf-sup constant Λ^*/β	Discrete inf-sup constant $\gamma_{\text{inf-sup}}$	$\|\mathrm{D}f_K^{-1}(\Theta_{K,\text{GS}}^{\Pi})\|_{\mathcal{F}_K^{-1} \to \mathcal{F}_K}^{-1}$ with $\mathrm{D}f_K(\Theta_{K,\text{GS}}^{\Pi})$ viewed as a mapping from $\widetilde{\mathcal{V}}_K$ to $\widetilde{\mathcal{V}}_K^*$
BeH_2	0.2508	1.9807	0.2568	0.2532	0.3592
BH_3	0.3056	1.5447	0.2081	0.2064	0.3254
HF	0.5010	2.4073	0.2529	0.2016	0.2993
H_2O	0.3067	2.2724	0.2789	0.2652	0.3646
LiH	0.1878	2.4044	0.2164	0.1953	0.2630
NH_3	0.3721	2.0420	0.2784	0.2732	0.4302

TABLE 2. Numerical results for the **CCSD** scheme (excitation rank-truncation of order **two**). The calculations were performed in STO-6G basis sets with the exception of the HF and LiH molecules for which 6-31G basis sets were used. Since the exact coupled cluster ground state zero $\Theta_{K,\text{GS}}^* \in \widehat{\mathcal{H}}_{\Psi_{0,K}}^{1,\perp}$ is unavailable, the Full-CC ground state zero in the corresponding basis set (STO-6G or 6-31G) was taken as a reference.

This first set of molecules constitutes an example of 'well-behaved' molecules, i.e., molecules for which the basic CCSD scheme work well, attaining in all cases the so-called chemical accuracy threshold of an error less than 10^{-3} Hartree. As a more stringent test case, we consider the nitrogen dimer and the

Molecule	$\left\|\mathbb{P}^{\perp}_{\mathcal{F}_K}\widetilde{u}_K\mathbb{P}_{\mathcal{F}_K}\right\|_{\mathcal{F}_K\to\widehat{\mathcal{L}}^2}$	$\frac{\sqrt{\Lambda^{\mathcal{F}_K}_{\min}}\,\Gamma^*_{\mathrm{GS}}}{\beta_K(\Theta^*_{K,\mathrm{GS}})}$	Continuous inf-sup constant Λ^*/β	Discrete inf-sup constant $\gamma_{\mathrm{inf-sup}}$	$\|\mathrm{D}f_K^{-1}(\Theta^{\Pi}_{K,\mathrm{GS}})\|^{-1}_{\mathcal{F}_K^{-1}\to\mathcal{F}_K}$ with $\mathrm{D}f_K(\Theta^{\Pi}_{K,\mathrm{GS}})$ viewed as a mapping from $\widetilde{\mathcal{V}}_K$ to $\widetilde{\mathcal{V}}_K^*$
BeH_2	0.1835	2.2865	0.2568	0.2321	0.3403
BH_3	0.2581	2.1659	0.2081	0.1752	0.3081
HF	0.4903	3.5897	0.2529	0.2187	0.2995
H_2O	0.2431	3.0995	0.2789	0.2504	0.3592
LiH	0.1629	3.3790	0.2164	0.2038	0.2628
NH_3	0.3038	2.6085	0.2784	0.2338	0.4147

TABLE 3. Numerical results for the **CCSDT** scheme (excitation rank-truncation of order **three**). As before, the calculations were performed in STO-6G basis sets with the exception of the HF and LiH molecules for which 6-31G basis sets were used, and the Full-CC ground state zero in the corresponding basis set (STO-6G or 6-31G) was taken as a reference.

carbon monoxide molecule at equilibrium geometry, these being examples of molecules for which the CCSD scheme is unable to achieve a chemically accurate solution[5]. The corresponding results, using once again minimal STO-6G basis sets, are displayed in Table 4. In the case of the carbon monoxide molecule, we see that Inequality (117) is not satisfied for the CCSD and CCSDT discretisations. On the other hand, the higher excitation rank-truncation discretisations indeed satisfy Inequality (117).

Molecule	N_2	CO
CCSD Energy Error (milli-Hartree)	4.0	8.2
Local monotonicity constant Γ according to Equation (14)	-0.8613	-1.2467
Discrete inf-sup constant $\gamma_{\mathrm{inf-sup}}$ at excitation truncation-rank two	0.0004	-0.1074
Discrete inf-sup constant $\gamma_{\mathrm{inf-sup}}$ at excitation truncation-rank three	0.0402	-0.0426
Discrete inf-sup constant $\gamma_{\mathrm{inf-sup}}$ at excitation truncation-rank four	0.0908	0.0225
Discrete inf-sup constant $\gamma_{\mathrm{inf-sup}}$ at excitation truncation-rank five	0.1364	0.0666
Continuous inf-sup constant Λ^*/β	0.1614	0.1255

TABLE 4. Numerical results for the nitrogen dimer and the carbon monoxide molecule at equilibrium geometries. The calculations were performed in STO-6G basis sets, and as before, the Full-CC ground state zero in the STO-6G basis set was taken as a reference. For the sake of comparison, we also include the local monotonicity constant Γ for these molecules, computed according to Equation (14).

[5]In fact, the CCSD scheme applied to these two molecules is an order of magnitude less accurate compared to the first set of molecules.

Appendix C. Continuity Properties of Generalised Excitation Operators

The goal of this section is to complete a missing argument in the proof of Lemma 24 by demonstrating that the so-called generalised excitation operators are bounded operators from $\widehat{\mathscr{H}}^1$ to $\widehat{\mathscr{H}}^1$. Recall that generalised excitation operators– in contrast to the usual excitation operators defined through Definition 10, function by replacing occupied orbitals with occupied or virtual orbitals with virtual orbitals.

The remainder of this section is organised as follows. We begin by stating the precise definition of a generalised excitation operator. We then state a brief summary of the mathematical framework introduced by Rohwedder [25] for establishing the continuity properties of cluster operators constructed from ordinary excitation operators. Finally, we adapt the arguments in [25, Theorem 4.1] to prove that generalised excitation operators are indeed bounded operators from $\widehat{\mathscr{H}}^1$ to $\widehat{\mathscr{H}}^1$. Let us emphasise here that our proof is essentially an adaptation of the arguments of Rohwedder, and we will therefore make heavy use of the notation and definitions introduced in [25].

Throughout this section, we will use the setting and notation developed in Section 3. In particular, we assume that we are given an N-dimensional occupied space $\mathscr{R} \subset \mathrm{H}^1(\mathbb{R}^3)$, a corresponding virtual space $\mathscr{R}^\perp$, $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis sets $\mathscr{B}_{\mathrm{occ}} := \{\psi_j\}_{j=1}^N$ and $\mathscr{B}_{\mathrm{vir}} := \{\psi_j\}_{j=N+1}^\infty$ for $\mathscr{R}$ and $\mathscr{R}^\perp$ respectively, a reference determinant Ψ_0 constructed from $\mathscr{B}_{\mathrm{occ}}$, and an N-particle Slater basis $\mathcal{B}_\wedge$ for $\widehat{\mathscr{H}}^1$ constructed from $\mathscr{B}_{\mathrm{occ}} \cup \mathscr{B}_{\mathrm{vir}}$ (see Definition 6).

Definition 42 (Occupied-to-Occupied Generalised Excitation Operator)**.**

Let $j \in \{1, \dots, N\}$, let $\widetilde{\mu}$ be an index of the form

$$\mu = \binom{i_1, \dots, i_j}{\ell_1, \dots, \ell_j} : i_1 < \ldots < i_j \in \{1, \dots, N\} \text{ and } \ell_1 < \ldots < \ell_j \in \{1, \dots, N\}.$$

We define the occupied-to-occupied generalised excitation operator $\mathcal{X}_{\widetilde{\mu}} \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{L}}^2$ through its action on the N-particle basis set $\mathcal{B}_\wedge$: For $\Psi_\nu(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det\left(\psi_{\nu_j}(\mathbf{x}_i)\right)_{i,j=1}^N$, we set

$$\mathcal{X}_\mu \Psi_\nu = \begin{cases} 0 & \text{if } \{\ell_1, \dots, \ell_j\} \not\subset \{\nu_1, \dots, \nu_N\}, \\ 0 & \text{if } \exists i_m \in \{i_1, \dots, i_j\} \text{ such that } i_m \in \{\nu_1, \dots, \nu_N\}, \\ \Psi_{\nu,i} \in \mathcal{B}_\wedge & \text{otherwise}, \end{cases}$$

where the determinant $\Psi_{\nu,i}$ is constructed from Ψ_ν by replacing all functions $\psi_{\ell_1}, \dots \psi_{\ell_j}$ used to construct Ψ_ν with functions $\psi_{i_1}, \dots, \psi_{i_j}$ respectively.

Definition 43 (Virtual to Virtual Generalised Excitation Operator)**.**

Let $j \in \{1, \dots, N\}$, let $\widetilde{\mu}$ be an index of the form

$$\mu = \binom{a_1, \dots, a_j}{b_1, \dots, b_j} : a_1 < \ldots < a_j \in \{N+1, N+2, \dots\} \text{ and } b_1 < \ldots < b_j \in \{N+1, N+2, \dots\}.$$

We define the virtual-to-virtual generalised excitation operator $\mathcal{X}_{\widetilde{\mu}} \colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{L}}^2$ through its action on the N-particle basis set $\mathcal{B}_\wedge$: For $\Psi_\nu(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det\left(\psi_{\nu_j}(\mathbf{x}_i)\right)_{i,j=1}^N$, we set

$$\mathcal{X}_\mu \Psi_\nu = \begin{cases} 0 & \text{if } \{b_1, \dots, b_j\} \not\subset \{\nu_1, \dots, \nu_N\}, \\ 0 & \text{if } \exists a_m \in \{a_1, \dots, a_j\} \text{ such that } a_m \in \{\nu_1, \dots, \nu_N\}, \\ \Psi_{\nu,a} \in \mathcal{B}_\wedge & \text{otherwise}, \end{cases}$$

where the determinant $\Psi_{\nu,a}$ is constructed from Ψ_ν by replacing all functions $\psi_{b_1}, \dots \psi_{b_j}$ used to construct Ψ_ν with functions $\psi_{a_1}, \dots, \psi_{a_j}$ respectively.

Our goal now is prove the following basic result.

Lemma 44 (Continuity of Generalised Excitation Operators)**.**

Let $\mathcal{X}_{\widetilde{\mu}}\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{L}}^2$ be an occupied-to-occupied or virtual-to-virtual generalised excitation operator defined according to Definition 42 or Definition 43 respectively. Then it holds that $\mathcal{X}_{\widetilde{\mu}}$ is a bounded linear mapping from $\widehat{\mathcal{H}}^1$ to $\widehat{\mathcal{H}}^1$.

As mentioned previously, we will prove Lemma 44 by making heavy usage of the basic mathematical framework introduced by Rohwedder [25] for the analysis of cluster operators constructed from the usual (occupied-to-virtual) excitation operators defined according to Definition 10. We therefore begin by recalling some essential notations and definitions from [25, Section 3].

Definition 45 (An Equivalent Norm on $\widehat{\mathcal{H}}^1$)**.**

Recalling that the occupied space $\mathscr{R}$ is an N-dimensional subspace of $\mathrm{H}^1(\mathbb{R}^3)$, let $\mathbb{P}_{\mathscr{R}}\colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^1(\mathbb{R}^3)$ denote the $\mathrm{L}^2(\mathbb{R}^3)$-orthogonal projection operator onto $\mathscr{R}$, and let $\mathbb{P}_{\mathscr{R}}^{\perp} := \mathrm{I} - \mathbb{P}_{\mathscr{R}}$. Then

- *We introduce the operator $\mathscr{A}\colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R}^3)$ as*
$$\mathscr{A} := \mathbb{P}_{\mathscr{R}}^{*}(\mathrm{I}-\Delta)\mathbb{P}_{\mathscr{R}} + (\mathbb{P}_{\mathscr{R}}^{\perp})^{*}(\mathrm{I}-\Delta)\mathbb{P}_{\mathscr{R}}^{\perp},$$
where $\mathbb{P}_{\mathscr{R}}^{*}$ and $(\mathbb{P}_{\mathscr{R}}^{\perp})^{*}$ denote the adjoint operators of $\mathbb{P}_{\mathscr{R}}$ and $\mathbb{P}_{\mathscr{R}}^{\perp}$ respectively.
- *We denote by $(\cdot,\cdot)_{\mathscr{R}}\colon \mathrm{H}^1(\mathbb{R}^3)\times\mathrm{H}^1(\mathbb{R}^3) \to \mathbb{R}$ the inner product induced by the operator $\mathscr{A}\colon \mathrm{H}^1(\mathbb{R}^3) \to \mathrm{H}^{-1}(\mathbb{R}^3)$, i.e.,*
$$\begin{aligned}\forall \phi,\psi \in \mathrm{H}^1(\mathbb{R}^3)\colon \qquad (\phi,\psi)_{\mathscr{R}} :=& \langle \phi, \mathscr{A}\psi\rangle_{\mathrm{H}^1(\mathbb{R}^3)\times\mathrm{H}^{-1}(\mathbb{R}^3)}\\ =& \langle \mathbb{P}_{\mathscr{R}}\phi, (\mathrm{I}-\Delta)\mathbb{P}_{\mathscr{R}}\psi\rangle_{\mathrm{H}^1(\mathbb{R}^3)\times\mathrm{H}^{-1}(\mathbb{R}^3)}\\ &+\langle \mathbb{P}_{\mathscr{R}}^{\perp}\phi, (\mathrm{I}-\Delta)\mathbb{P}_{\mathscr{R}}^{\perp}\psi\rangle_{\mathrm{H}^1(\mathbb{R}^3)\times\mathrm{H}^{-1}(\mathbb{R}^3)},\end{aligned}$$
and we note that the norm $\|\cdot\|_{\mathscr{R}}$ induced by the inner product $(\cdot,\cdot)_{\mathscr{R}}$ is equivalent to the usual $\mathrm{H}^1(\mathbb{R}^3)$ norm.
- *Finally, we introduce the operator $\mathscr{A}_N\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$ as*
$$\mathscr{A}_N := \sum_{j=1}^{N} \underbrace{\mathrm{I}\otimes\ldots\otimes\ldots\otimes\mathrm{I}}_{j-1 \text{ times}} \otimes \mathscr{A} \otimes \underbrace{\mathrm{I}\otimes\ldots\otimes\mathrm{I}}_{N-j\text{times}},$$
we denote by $(\cdot,\cdot)_{\mathscr{R}^N}\colon \widehat{\mathcal{H}}^1 \times \widehat{\mathcal{H}}^1 \to \mathbb{R}$ the inner product induced by $\mathscr{A}_N\colon \widehat{\mathcal{H}}^1 \to \widehat{\mathcal{H}}^{-1}$, i.e.,
$$\forall \Phi,\Psi \in \widehat{\mathcal{H}}^1\colon \qquad (\Phi,\Psi)_{\mathscr{R}^N} := \langle \Phi, \mathscr{A}_N\Psi\rangle_{\widehat{\mathcal{H}}^1\times\widehat{\mathcal{H}}^{-1}},$$
and we note that the norm $\|\cdot\|_{\mathscr{R}^N}$ induced by the inner product $(\cdot,\cdot)_{\mathscr{R}^N}$ is equivalent to the usual $\widehat{\mathcal{H}}^1$ norm.

Consider Definition 45. The fact that the new norms introduced through this definition are indeed equivalent to the usual $\mathrm{H}^1(\mathbb{R}^3)$ and $\widehat{\mathcal{H}}^1$ norms has been demonstrated in [25, Lemma 3.1].

Notation 46. *We denote by $\{\widetilde{\phi_j}\}_{j\in\mathbb{N}}$ an $(\cdot,\cdot)_{\mathscr{R}}$-orthonormal basis for $\mathrm{H}^1(\mathbb{R}^3)$ with the property that*
$$\mathrm{span}\{\widetilde{\phi_j}\}_{j=1}^{N} = \mathscr{R} \qquad \overline{\mathrm{span}\{\widetilde{\phi_j}\}_{j=N+1}^{\infty}} = \mathscr{R}^{\perp}.$$
Additionally, we denote by $\mathscr{N} \subset \mathbb{N}^{N-1}$, the set of multi-indices defined as
$$\mathscr{N} := \{\underline{\nu} := (\nu_1,\nu_2,\ldots,\nu_{N-1})\colon \quad \nu_1 < \nu_2 < \ldots < \nu_{N-1} \in \mathbb{N}\}.$$

Equipped with Definition 45 and Notation 46, we now have the following important lemma.

Lemma 47 ([25, Lemma 3.3, Proposition 3.4])**.**

Let the inner products $(\cdot,\cdot)_{\mathscr{R}}\colon \mathrm{H}^1(\mathbb{R}^3)\times\mathrm{H}^1(\mathbb{R}^3)\to\mathbb{R}$ and $(\cdot,\cdot)_{\mathscr{R}^N}\colon \widehat{\mathcal{H}}^1\times\widehat{\mathcal{H}}^1\to\mathbb{R}$ be defined according to Definition 45, recall the definition of the index set $\mathcal{I}_\infty^N$ given by Definition 6, and consider Notation 46. Then the following holds:

(1) For any antisymmetric function $\Phi\in\widehat{\mathcal{H}}^1$ it holds that

$$\|\Phi\|_{\mathscr{R}^N}^2 = N(\Phi,\Phi)_{\mathscr{R}_1} := N\langle\Phi,\mathscr{A}_1\Phi\rangle_{\widehat{\mathcal{H}}^1\times\widehat{\mathcal{H}}^{-1}},$$

where the operator $\mathscr{A}_1\colon\widehat{\mathcal{H}}^1\to\widehat{\mathcal{H}}^{-1}$ is defined as

$$\mathscr{A}_1 := \mathscr{A}\otimes\underbrace{\mathrm{I}\otimes\ldots\otimes\mathrm{I}}_{N-1\text{times}}.$$

(2) The set $\mathbb{B}_{N-1}$ defined as

$$\mathbb{B}_{N-1} := \{\Psi_{\underline{\nu}} := \psi_{\nu_1}\wedge\psi_{\nu_2}\wedge\ldots\wedge\psi_{\nu_{N-1}}\colon \quad \underline{\nu}=(\nu_1,\nu_2,\ldots,\nu_{N-1})\in\mathscr{N}\},$$

is an orthonormal basis for the antisymmetric tensor product space

$$\widehat{\mathcal{L}}_{N-1}^2 := \bigwedge_{j=1}^{N-1}\mathrm{L}^2(\mathbb{R}^3),$$

where we recall that $\{\psi_j\}_{j=1}^\infty = \mathscr{B}_{\mathrm{occ}}\cup\mathscr{B}_{\mathrm{vir}}$.

(3) The set $\overline{\mathbb{B}}$ defined as

$$\overline{\mathbb{B}} := \{\Phi_{p,\underline{\nu}} := \widetilde{\phi}_p\otimes\Psi_{\underline{\nu}}\colon \quad p\in\mathbb{N},\ \Psi_{\underline{\nu}}\in\mathbb{B}_{N-1}\},$$

is an $(\cdot,\cdot)_{\mathscr{R}_1}$-orthonormal system that satisfies the property that $\widehat{\mathcal{H}}^1\subset\overline{\mathrm{span}\overline{\mathbb{B}}}$.

(4) For any antisymmetric function $\Phi\in\widehat{\mathcal{H}}^1$ of the form $\Phi = d_{(1,\ldots,N)}\Psi_0 + \sum_{\mu\in\mathcal{I}_\infty^N} d_\mu\Psi_\mu$, it holds that

$$\|\Phi\|_{\mathscr{R}^N}^2 = N\sum_{\Phi_{p,\underline{\nu}}\in\overline{\mathbb{B}}}(\Phi_{p,\underline{\nu}},\Phi)_{\mathscr{R}_1}^2 = \sum_{i=1}^N\sum_{\underline{\nu}\in\mathscr{N}}\left|\sum_{j=1}^N(-1)^{|j,\underline{\nu}|-1}d_{[j,\underline{\nu}]}\gamma_{i,j}\right|^2 + \sum_{a=N+1}^\infty\sum_{\underline{\nu}\in\mathscr{N}}\left|\sum_{b=N+1}^\infty(-1)^{|b,\underline{\nu}|-1}d_{[b,\underline{\nu}]}\gamma_{a,b}\right|^2, \tag{119}$$

where we have introduced the following notation:

- *For any $p\in\mathbb{N}$ and any $\underline{\nu}\in\mathscr{N}$, we denote by $[p,\underline{\nu}]$ the unique element of the index set $\mathcal{I}_\infty^N$ formed by combining the index p and the $N-1$ indices from $\underline{\nu}$.*
- *For any $p\in\mathbb{N}$ and any $\underline{\nu}\in\mathscr{N}$, if $p\in\underline{\nu}$, then we set $d_{[p,\underline{\nu}]}=0$.*
- *For any $p\in\mathbb{N}$ and any $\underline{\nu}\in\mathscr{N}$, we denote by $|p,\underline{\nu}|$ the position of the index p in the multi-index $[p,\underline{\nu}]\in\mathcal{I}_\infty^N$.*
- *For any $p,q\in\mathbb{N}$, we define $\gamma_{p,q} := (\widetilde{\phi_p},\psi_q)_{\mathscr{R}}$, where we remind the reader that $\{\psi_q\}_{q=1}^\infty\subset\mathrm{H}^1(\mathbb{R}^3)$ denotes the original $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis for $\mathrm{H}^1(\mathbb{R}^3)$, which was used to construct the basis set $\mathcal{B}_\wedge$ for the N-particle function space $\widehat{\mathcal{H}}^1$ (see Definition 6).*

For a proof of this lemma, see [25, Lemma 3.3, Proposition 3.4].

We are now ready to state the proof of Lemma 44.

Proof of Lemma 44. We first consider the case where $\mathcal{X}_{\widetilde{\mu}}$ is an occupied to occupied generalised excitation operator. Notice that we may assume without loss of generality that $\mathcal{X}_{\widetilde{\mu}}$ is a first-order excitation operator, i.e., the index $\widetilde{\mu}$ is of the form

$$\mu = \begin{pmatrix} i_1 \\ \ell_1 \end{pmatrix} : i_1, \ell_1 \in \{1, \dots, N\},$$

Indeed, higher-order excitation operators $\mathcal{X}_{\widetilde{\nu}}$, i.e., excitation operators for which the index $\widetilde{\nu}$ is of length $j \in \{2, \dots, N\}$ can be written as products of first-order excitation operators so that the continuity properties of such excitation operators will follow once we have proved the continuity of an arbitrary first-order excitation operator $\mathcal{X}_{\widetilde{\mu}}$ as a mapping from $\widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1$.

To proceed, let $\Phi = c_{(1,\dots,N)}\Psi_0 + \sum_{\mu \in \mathcal{I}_\infty^N} c_\mu \Psi_\mu \in \widehat{\mathscr{H}}^1$ be arbitrary. We define $\Upsilon := \mathcal{X}_{\widetilde{\mu}}\Phi$. Clearly, $\Upsilon \in \widehat{\mathcal{L}}^2 \cap \{\Psi_0\}^\perp$, and we can therefore express Υ in the form

$$\Upsilon = \mathcal{X}_{\widetilde{\mu}} \left(c_{(1,\dots,N)}\Psi_0 + \sum_{\mu \in \mathcal{I}_\infty^N} c_\mu \Psi_\mu \right) = \sum_{\mu \in \mathcal{I}_\infty^N} d_\mu \Psi_\mu, \tag{120}$$

Using now Lemma 47, we can deduce that

(121)

$$\|\Upsilon\|_{\mathscr{H}^N}^2 = \|\mathcal{X}_{\widetilde{\mu}}\Phi\|_{\mathscr{H}^N}^2 = \underbrace{\sum_{i=1}^N \sum_{\underline{\nu} \in \mathcal{N}} \left| \sum_{j=1}^N (-1)^{|j,\underline{\nu}|-1} d_{[j,\underline{\nu}]} \gamma_{i,j} \right|^2}_{:=(\mathrm{I})} + \underbrace{\sum_{a=N+1}^\infty \sum_{\underline{\nu} \in \mathcal{N}} \left| \sum_{b=N+1}^\infty (-1)^{|b,\underline{\nu}|-1} d_{[b,\underline{\nu}]} \gamma_{a,b} \right|^2}_{:=(\mathrm{II})}.$$

Let us first consider the term (I). A simple application of the Cauchy-Schwarz inequality yields that

$$\begin{aligned} (\mathrm{I}) = \sum_{i=1}^N \sum_{\underline{\nu} \in \mathcal{N}} \left| \sum_{j=1}^N (-1)^{|j,\underline{\nu}|-1} d_{[j,\underline{\nu}]} \gamma_{i,j} \right|^2 &\leq \sum_{i=1}^N \sum_{\underline{\nu} \in \mathcal{N}} \left(\sum_{j=1}^N d_{[j,\underline{\nu}]}^2 \sum_{j=1}^N \gamma_{i,j}^2 \right) \\ &= \left(\sum_{i=1}^N \sum_{j=1}^N \gamma_{i,j}^2 \right) \left(\sum_{\underline{\nu} \in \mathcal{N}} \sum_{j=1}^N d_{[j,\underline{\nu}]}^2 \right). \end{aligned}$$

On the one hand, recalling that each orbital $\psi_i,\ i \in \{1, \dots, N\}$ belongs to $\mathscr{R} = \mathrm{span}\{\widetilde{\phi}_j\}_{j=1}^N$, we have

$$\sum_{i=1}^N \sum_{j=1}^N \gamma_{i,j}^2 = \sum_{i=1}^N \|\psi_i\|_{\mathscr{R}}^2 = \|\Psi_0\|_{\mathscr{R}^N}^2.$$

On the other hand, we easily deduce that

$$\sum_{\underline{\nu} \in \mathcal{N}} \sum_{j=1}^N d_{[j,\underline{\nu}]}^2 = \sum_{j=1}^N \sum_{\underline{\nu} \in \mathcal{N}} d_{[j,\underline{\nu}]}^2 \leq N \sum_{\mu \in \mathcal{I}_\infty^N} d_\mu^2 = N \|\Upsilon\|_{\widehat{\mathcal{L}}^2}^2.$$

It therefore follows that

$$(\mathrm{I}) \leq N \|\Psi_0\|_{\mathscr{H}^N}^2 \|\Upsilon\|_{\widehat{\mathcal{L}}^2}^2, \tag{122}$$

and it remains only to consider the term (II) in Equation (121).

We first attempt to express the coefficients $d_{[b,\underline{\nu}]},\ b \in \{N+1,\dots,\},\underline{\nu} \in \mathscr{N}$ that appear in the term (II) in terms of the original expansion coefficients $\{c_\mu\}_{\mu\in\mathcal{I}_\infty^N}$ of the input function $\Phi \in \widehat{\mathscr{H}}^1$. To do so, fix some $b \in \{N+1,\dots,\},\underline{\nu} \in \mathscr{N}$ such that the coefficient $d_{[b,\underline{\nu}]} \neq 0$, and recall the expression of the function Υ given by Equation (120). It can readily be seen that there must exist a unique index $\delta \in \mathcal{I}_\infty^N$ such that

$$\mathcal{X}_{\widetilde{\mu}}\Psi_\delta = \Psi_{[b,\underline{\nu}]}. \tag{123}$$

Moreover, since $b \in \{N+1,\dots\}$ and $\mathcal{X}_{\widetilde{\mu}}$ is an *occupied-to-occupied* generalised excitation operator, the index $\delta \in \mathcal{I}_\infty^N$ must be of the form

$$\delta = [b,\underline{\beta}] \quad \text{for some } \underline{\beta} = \underline{\beta}(\widetilde{\mu},\underline{\nu}) \in \mathscr{N}, \tag{124}$$

where, we have highlighted the dependency of the multi-index $\underline{\beta}$ on $\widetilde{\mu}$ and $\underline{\nu}$.

Consequently, each non-zero coefficient $d_{[b,\underline{\nu}]},\ b \in \{N+1,\dots,\},\underline{\nu} \in \mathscr{N}$ appearing in Equation (121) can be written in the form

$$(-1)^{|b,\underline{\nu}|-1}d_{[b,\underline{\nu}]} = (-1)^{|b,\underline{\beta}(\widetilde{\mu},\underline{\nu})|-1}\ \delta\big(\underline{\nu},\underline{\beta}(\widetilde{\mu},\underline{\nu})\big)\ c_{[b,\underline{\beta}(\widetilde{\mu},\underline{\nu})]},$$

where $\delta\big(\underline{\nu},\underline{\beta}(\widetilde{\mu},\underline{\nu})\big) \in \{\pm 1\}$ is a sign distribution. The fact that the sign distribution δ depends only on $\underline{\nu}$ and $\underline{\beta}$, and is, in particular, independent of b can be deduced from the argument given in [25, Lemma 4.3].

It now follows that the term (II) appearing in Equation (121) can be written as

$$\begin{aligned}
\text{(II)} &= \sum_{a=N+1}^{\infty}\sum_{\underline{\nu}\in\mathscr{N}}\left|\sum_{b=N+1}^{\infty}(-1)^{|b,\underline{\nu}|-1}d_{[b,\underline{\nu}]}\gamma_{a,b}\right|^2 \\
&= \sum_{a=N+1}^{\infty}\sum_{\underline{\nu}\in\mathscr{N}}\left|\sum_{b=N+1}^{\infty}(-1)^{|b,\underline{\beta}(\widetilde{\mu},\underline{\nu})|-1}\ \delta\big(\underline{\nu},\underline{\beta}(\widetilde{\mu},\underline{\nu})\big)\ c_{[b,\underline{\beta}(\widetilde{\mu},\underline{\nu})]}\gamma_{a,b}\right|^2 \\
&= \sum_{a=N+1}^{\infty}\sum_{\underline{\nu}\in\mathscr{N}}\left|\sum_{b=N+1}^{\infty}(-1)^{|b,\underline{\beta}(\widetilde{\mu},\underline{\nu})|-1}\ c_{[b,\underline{\beta}(\widetilde{\mu},\underline{\nu})]}\gamma_{a,b}\right|^2. \qquad (125)
\end{aligned}$$

In order to proceed, let us consider two multi-indices $\underline{\nu}_1,\underline{\nu}_2 \in \mathscr{N}$ such that $\underline{\nu}_1 \neq \underline{\nu}_2$. It is easy to see that for any $b \in \{N+1,\dots,\}$ with the property that $d_{[b,\underline{\nu}_1]}, d_{[b,\underline{\nu}_2]} \neq 0$, the corresponding multi-indices $\underline{\beta}(\widetilde{\mu},\underline{\nu}_1)$ and $\underline{\beta}(\widetilde{\mu},\underline{\nu}_2)$ defined according to Equations (123) and (124) must satisfy

$$\underline{\beta}(\widetilde{\mu},\underline{\nu}_1) \neq \underline{\beta}(\widetilde{\mu},\underline{\nu}_2).$$

Returning to Equation (125), we conclude that

$$\text{(II)} = \sum_{a=N+1}^{\infty}\sum_{\underline{\nu}\in\mathscr{N}}\left|\sum_{b=N+1}^{\infty}(-1)^{|b,\underline{\beta}(\widetilde{\mu},\underline{\nu})|-1}\ c_{[b,\underline{\beta}(\widetilde{\mu},\underline{\nu})]}\gamma_{a,b}\right|^2 \leq \sum_{a=N+1}^{\infty}\sum_{\underline{\alpha}\in\mathscr{N}}\left|\sum_{b=N+1}^{\infty}(-1)^{|b,\underline{\alpha}|-1}\ c_{[b,\underline{\alpha}]}\gamma_{a,b}\right|^2.$$

Comparing the above expression with Equation (119) for the $\|\cdot\|_{\mathscr{R}^N}$ norm of an arbitrary element of $\widehat{\mathscr{H}}^1$, we conclude that

$$\text{(II)} \leq \sum_{a=N+1}^{\infty}\sum_{\underline{\alpha}\in\mathscr{N}}\left|\sum_{b=N+1}^{\infty}(-1)^{|b,\underline{\alpha}|-1}\ c_{[b,\underline{\alpha}]}\gamma_{a,b}\right|^2 \leq \|\Phi\|^2_{\mathscr{R}^N}. \tag{126}$$

Combining now Inequalities (122) and (126), and using the fact that the $\|\cdot\|_{\mathscr{R}^N}$ norm is equivalent to the usual $\|\cdot\|_{\widehat{\mathscr{H}}^1}$ on $\widehat{\mathscr{H}}^1$ yields the required continuity of the occupied-to-occupied generalised excitation operator $\mathcal{X}_{\widetilde{\mu}}\colon \widehat{\mathscr{H}}^1 \to \widehat{\mathscr{H}}^1$.

It remains to consider the case when $\mathcal{X}_{\widetilde{\mu}}$ is a virtual-to-virtual generalised excitation operator. The argument in this case, while more tedious, is very similar to the proof we have stated above. Indeed, the only difference arises in Equation (124), where we must consider two different cases, namely, one in which $\delta = [b, \underline{\beta}]$ and the other in which the excitation index $\widetilde{\mu}$ itself involves the virtual orbital b. For the sake of brevity, we skip this result and instead refer interested readers to the proof of [25, Theorem 4.1] which is the basic template that we used to construct the present occupied-to-occupied proof and can also be adapted to prove the virtual-to-virtual case. □

Appendix D. Some Auxiliary Results

The goal of this final section is to provide a demonstration of an earlier claim from **Step Three** in the proof of Lemma 24. Let us first state precisely the claim.

Lemma 48. *Let the sequence of reference determinants $\{\Psi_{0,K}\}_{K\geq N}$ be defined as in Notation 21 and let the limiting reference determinant $\Psi_0^* \in \widehat{\mathscr{H}}^1$ be given in accordance with **Assumption A.II**. Then for each K large enough, there exist L^2-orthonormal, single-particle functions $\{\psi_{j,0}\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ and $\{\psi_{j,K}\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ such that*

$$\Psi_0^*(\mathbf{x}_1,\dots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\det\big(\psi_{j,0}(\mathbf{x}_i)\big)_{i,j=1}^N \quad \textit{for a.e. } (\mathbf{x}_1,\dots\mathbf{x}_N)\in\mathbb{R}^{3N},$$
$$\Psi_{0,K}(\mathbf{x}_1,\dots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\det\big(\psi_{j,K}(\mathbf{x}_i)\big)_{i,j=1}^N \quad \textit{for a.e. } (\mathbf{x}_1,\dots\mathbf{x}_N)\in\mathbb{R}^{3N},$$

and

$$\forall j\in\{1,\dots,N\}\colon \quad \lim_{K\to\infty}\|\psi_{j,K}-\psi_{j,0}\|_{\mathrm{H}^1(\mathbb{R}^3)} = 0.$$

Proof. Let $K \geq N$ be sufficiently large and fixed. By the definition of a Slater determinant, there exist some L^2-orthonormal, single-particle functions $\{\psi_{j,0}\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ and $\{\rho_{j,K}\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ such that

$$\Psi_0^*(\mathbf{x}_1,\dots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\det\big(\psi_{j,0}(\mathbf{x}_i)\big)_{i,j=1}^N \quad \text{for a.e. } (\mathbf{x}_1,\dots\mathbf{x}_N)\in\mathbb{R}^{3N},$$
$$\Psi_{0,K}(\mathbf{x}_1,\dots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\det\big(\rho_{j,K}(\mathbf{x}_i)\big)_{i,j=1}^N \quad \text{for a.e. } (\mathbf{x}_1,\dots\mathbf{x}_N)\in\mathbb{R}^{3N}.$$

Let $\mathscr{R}_K := \operatorname{span}\{\rho_{j,K}\}_{j=1}^N \subset \mathrm{H}^1(\mathbb{R}^3)$ be an N-dimensional subspace. We first claim that there exists a (not necessarily orthonormal) basis $\{\widetilde{\rho}_{j,K}\}_{j=1}^N$ of $\mathscr{R}_K$ such that

$$\Psi_{0,K}(\mathbf{x}_1,\dots,\mathbf{x}_N) = \alpha_K\det\big(\widetilde{\rho}_{j,K}(\mathbf{x}_i)\big)_{i,j=1}^N \quad \text{for a.e. } (\mathbf{x}_1,\dots\mathbf{x}_N)\in\mathbb{R}^{3N}, \qquad \text{and} \tag{127}$$
$$\forall i,j\in\{1,\dots,N\}\colon \quad (\widetilde{\rho}_{j,K},\psi_{i,0})_{\mathrm{L}^2(\mathbb{R}^3)} = \delta_{ij}.$$

To see this, let us first notice that there must exist $j\in\{1,\dots,N\}$ such that $(\rho_{j,K},\psi_{1,0})_{\mathrm{L}^2(\mathbb{R}^3)} \neq 0$. Indeed, if this were not the case, then we would clearly have for a.e. $(\mathbf{x}_2,\dots,\mathbf{x}_N)\in\mathbb{R}^{3(N-1)}$ that

$$\int_{\mathbb{R}^3}\psi_{1,0}(\mathbf{x}_1)\Psi_{0,K}(\mathbf{x}_1,\dots,\mathbf{x}_N)\,d\mathbf{x}_1 = \frac{1}{\sqrt{N!}}\int_{\mathbb{R}^3}\psi_{1,0}(\mathbf{x}_1)\det\big(\rho_{j,K}(\mathbf{x}_i)\big)_{i,j=1}^N\,d\mathbf{x}_1 \equiv 0. \tag{128}$$

On the other hand, **Assumption A.II** yields that $\|\Psi_{0,K} - \Psi_0^*\|_{\widehat{\mathcal{L}}^2} \to 0$ as $K \to \infty$ and therefore Equation (128) contradicts the fact that for a.e. $(\mathbf{x}_2, \dots, \mathbf{x}_N) \in \mathbb{R}^{3(N-1)}$

$$\begin{aligned}\int_{\mathbb{R}^3} \psi_{1,0}(\mathbf{x}_1)\Psi_0^*(\mathbf{x}_1, \dots, \mathbf{x}_N)\, d\mathbf{x}_1 =& \frac{1}{\sqrt{N!}} \int_{\mathbb{R}^3} \psi_{1,0}(\mathbf{x}_1) \det\big(\psi_{j,0}(\mathbf{x}_i)\big)_{i,j=1}^N \, d\mathbf{x}_1 \\ =& \frac{1}{\sqrt{N!}} \det\big(\psi_{j,0}(\mathbf{x}_i)\big)_{i,j=2}^N \\ \not\equiv& \, 0,\end{aligned}$$

As a consequence, we can define the function

$$\widehat{\rho}_{1,K} := \frac{1}{(\psi_{1,0}, \rho_{j,K})_{\mathrm{L}^2(\mathbb{R}^3)}} \rho_{j,K}.$$

Denoting now by $\Pi_1^j \colon \{1, \dots, N\} \to \{1, \dots, N\}$ the permutation defined as

$$\forall i \in \{1, \dots, N\} \colon \quad \Pi_1^j(i) = \begin{cases} i & \text{if } \ i \notin \{1, j\}, \\ 1 & \text{if } \ i = j, \\ j & \text{if } \ i = 1, \end{cases}$$

we can introduce the functions $\{\widehat{\rho}_{i,K}\}_{i=2}^N$ as

$$\widehat{\rho}_{i,K} := \rho_{\Pi_1^j(i),K} - \Big(\rho_{\Pi_1^j(i),K}, \psi_{1,0}\Big)_{\mathrm{L}^2(\mathbb{R}^3)} \widehat{\rho}_{1,K}.$$

A direct calculation now confirms that

$$\forall i \in \{1, \dots, N\} \colon \quad (\psi_{1,0}, \widehat{\rho}_{i,K}) = \delta_{1i},$$

and that the set $\{\widehat{\rho}_{i,K}\}_{i=1}^N$ forms a basis for the subspace $\mathscr{R}_K$. Note that this implies in particular that $\Psi_{0,K}$ can be constructed as a determinant from the N functions $\{\widehat{\rho}_{i,K}\}_{i=1}^N$ and a scaling factor.

We now repeat the above construction starting with the new set $\{\widehat{\rho}_{i,K}\}_{i=1}^N$ and the function $\psi_{2,0} \in \mathrm{H}^1(\mathbb{R}^3)$. Arguing exactly as before and using the fact that that $(\widehat{\rho}_{1,K}, \psi_{1,0})_{\widehat{\mathcal{L}}^2} = 1$ while $\psi_{1,0} \perp \psi_{2,0}$, we first deduce that there must exist $\ell \in \{2, \dots, N\}$ such that $(\widehat{\rho}_{\ell,K}, \psi_{2,0})_{\mathrm{L}^2(\mathbb{R}^3)} \neq 0$.

As a consequence, we can define the function

$$\widehat{\widehat{\rho}}_{2,K} := \frac{1}{(\psi_{2,0}, \widehat{\rho}_{\ell,K})_{\mathrm{L}^2(\mathbb{R}^3)}} \widehat{\rho}_{\ell,K},$$

Once again we denote by $\Pi_2^\ell \colon \{2, \dots, N\} \to \{2, \dots, N\}$ the permutation defined as

$$\forall i \in \{1, \dots, N\} \colon \quad \Pi_2^\ell(i) = \begin{cases} i & \text{if } \ i \notin \{2, \ell\}, \\ 2 & \text{if } \ i = \ell, \\ \ell & \text{if } \ i = 2, \end{cases}$$

and we introduce the functions $\{\widehat{\widehat{\rho}}_{i,K}\}_{\substack{i=1 \\ i \neq 2}}^N$ as

$$\widehat{\widehat{\rho}}_{i,K} := \widehat{\rho}_{\Pi_2^\ell(i),K} - \Big(\widehat{\rho}_{\Pi_2^\ell(i),K}, \psi_{2,0}\Big)_{\mathrm{L}^2(\mathbb{R}^3)} \widehat{\widehat{\rho}}_{2,K}.$$

A direct calculation now confirms that

$$\forall i \in \{1,\dots,N\}: \quad \left(\psi_{1,0}, \widehat{\widetilde{\rho}}_{i,K}\right) = \delta_{1i}, \quad \text{and}$$
$$\forall i \in \{1,\dots,N\}: \quad \left(\psi_{2,0}, \widehat{\widetilde{\rho}}_{i,K}\right) = \delta_{2i},$$

and that the set $\{\widehat{\widetilde{\rho}}_{i,K}\}_{i=1}^N$ still forms a basis for the subspace $\mathscr{R}_K$, which implies that $\Psi_{0,K}$ can be constructed as a determinant from the N functions $\{\widehat{\widetilde{\rho}}_{i,K}\}_{i=1}^N$ and a scaling factor. Continuing in this fashion allows us to produce the required basis $\{\widetilde{\rho}_{j,K}\}_{j=1}^N$ of $\mathscr{R}_K$.

Equipped with the basis $\{\widetilde{\rho}_{j,K}\}_{j=1}^N$ which satisfies the relations (127), we can now deduce that for any $j \in \{1,\dots,N\}$ and a.e. $\mathbf{x}_j \in \mathbb{R}^3$ it holds that

$$\alpha_K \widetilde{\rho}_{j,K}(\mathbf{x}_j) = \int_{\mathbb{R}^{3(N-1)}} \Psi_{0,K}(\mathbf{x}_1,\dots,\mathbf{x}_N)\psi_{1,0}(\mathbf{x}_1)\dots\psi_{j-1,0}(\mathbf{x}_{j-1}) \\ \psi_{j+1,0}(\mathbf{x}_{j+1})\dots\psi_{N,0}(\mathbf{x}_N)\, d\mathbf{x}_1\dots d\mathbf{x}_{j-1}d\mathbf{x}_{j+1}\dots d\mathbf{x}_N.$$

Similarly, for any $j \in \{1,\dots,N\}$ and a.e. $\mathbf{x}_j \in \mathbb{R}^3$ it holds that

$$\frac{1}{\sqrt{N!}}\psi_{j,0}(\mathbf{x}_j) = \int_{\mathbb{R}^{3(N-1)}} \Psi_0^*(\mathbf{x}_1,\dots,\mathbf{x}_N)\psi_{1,0}(\mathbf{x}_1)\dots\psi_{j-1,0}(\mathbf{x}_{j-1}) \\ \psi_{j+1,0}(\mathbf{x}_{j+1})\dots\psi_{N,0}(\mathbf{x}_N)\, d\mathbf{x}_1\dots d\mathbf{x}_{j-1}d\mathbf{x}_{j+1}\dots d\mathbf{x}_N.$$

Appealing now to **Assumption A.II** which yields the convergence of $\Psi_{0,K}$ to the limiting determinant Ψ_0^* in the $\widehat{\mathscr{H}}^1$ norm as $K \to \infty$, we readily deduce that for any $j \in \{1,\dots,N\}$

$$\lim_{K\to\infty} \left\| \alpha_K \widetilde{\rho}_{j,K} - \frac{1}{\sqrt{N!}}\psi_{j,0} \right\|_{\mathrm{H}^1(\mathbb{R}^3)} = 0.$$

Using the fact that, by construction, $(\widetilde{\rho}_{1,K}, \psi_{1,0})_{\mathrm{L}^2(\mathbb{R}^3)} = 1$, we can easily deduce from the above convergence result that $\lim_{K\to\infty} \alpha_K = \frac{1}{\sqrt{N!}}$, and therefore

$$\forall j \in \{1,\dots,N\}: \quad \lim_{K\to\infty} \|\widetilde{\rho}_{j,K} - \psi_{j,0}\|_{\mathrm{H}^1(\mathbb{R}^3)} = 0.$$

In order to conclude the proof (recall that the basis $\{\widetilde{\rho}_{j,K}\}_{j=1}^N$ is not $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal), we can perform a Gram-Schmidt procedure on the basis $\{\widetilde{\rho}_{j,K}\}_{j=1}^N$, and it is a simple exercise to show that the resulting $\mathrm{L}^2(\mathbb{R}^3)$-orthonormal basis $\{\Psi_{j,K}\}_{j=1}^N$ satisfies

$$\forall j \in \{1,\dots,N\}: \quad \lim_{K\to\infty} \|\psi_{j,K} - \psi_{j,0}\|_{\mathrm{H}^1(\mathbb{R}^3)} = 0.$$

□

Sorbonne Université, CNRS, Université Paris Cité, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France and Institute of Mathematics, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland.

(Y. Maday and Y. Wang) Sorbonne Université, CNRS, Université Paris Cité, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris.