



Equation Generator for Equation-of-Motion Coupled Cluster Assisted by Computer Algebra System

Raúl Quintero-Monsebaiz, Pierre-François Loos

► To cite this version:

Raúl Quintero-Monsebaiz, Pierre-François Loos. Equation Generator for Equation-of-Motion Coupled Cluster Assisted by Computer Algebra System. AIP Advances, 2023, 13 (8), pp.085035. 10.1063/5.0163846 . hal-04116693

HAL Id: hal-04116693

<https://hal.science/hal-04116693>

Submitted on 5 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Equation Generator for Equation-of-Motion Coupled Cluster Assisted by Computer Algebra System

Raúl Quintero-Monsebaiz^{1, a)} and Pierre-François Loos^{1, b)}

Laboratoire de Chimie et Physique Quantiques (UMR 5626), Université de Toulouse, CNRS, UPS, France

We present an equation generator algorithm that utilizes second-quantized operators in normal order with respect to a correlated or non-correlated reference and the corresponding Wick theorem. The algorithm proposed here, written with MATHEMATICA, enables the generation of non-redundant strings of second-quantized operators that, after classification, are directly assigned to intermediate quantities used to construct the coupled-cluster (CC) effective Hamiltonian. We demonstrate the capabilities of the algorithm by computing the CC amplitude equations and various blocks of the equation-of-motion CC effective Hamiltonian. A comprehensive description of this four-step algorithm is provided alongside concrete examples.

I. INTRODUCTION

The increase in computational capabilities has led us to explore more complex wave function ansätze. As these ansätze become more complicated, deriving the corresponding algebraic equations also becomes more tedious, time-consuming, and, more importantly, subject to human errors. An important feature of many-body quantum chemistry methods is that most of them can be written in the second quantization formalism based on established commutation or anti-commutation relations between second-quantized operators.¹ In the second quantization formalism, first introduced in quantum field theory, the wave function and the operators are expressed as products of creation and annihilation operators, also known as strings. These strings can be evaluated using either commutation rules, diagrammatic techniques,² or Wick's theorem,³ the latter being more efficient. These rules can then be programmed with the aid of symbolic algebra software, allowing us to write chains of products of these operators as tensors that can be efficiently evaluated by modern tensor contraction tools.^{4–10}

In the context of coupled-cluster (CC) methods,^{2,11–14} various equation generators have emerged to automate the process of deriving and implementing the corresponding equations, replacing the manual derivation and error-prone implementation in electronic structure software. In their pioneering work, Janssen and Schaefer generated and automatically implemented the open-shell CC equations by utilizing second-quantized strings.¹⁵ In a similar vein, Li and Paldus automated the implementation of spin-adapted open-shell CC equations, with the added benefit of utilizing the unitary group formalism that allows for the efficient handling of low-spin states.¹⁶ In 2001, Kállay and Surján proposed a general-order CC code by combining diagrammatic many-body perturbation theory and string-based configuration interaction.¹⁷ In such a way, the CC equations were written in terms of diagrams and stored as strings. This technique was then generalized to state-specific multireference CC,¹⁸ excited states computed within the linear response formalism,¹⁹ and approximate treatment of higher excitations.²⁰

Adopting the same design philosophy as Janssen and Schaefer, Hirata implemented the Tensor Contraction Engine (TCE)^{21–24} that performs the manipulation of second-quantized operators and the generation of the computer code. The main distinction is that TCE takes advantage of spin, spatial, and index permutation symmetries at every stage of the calculations, reducing the computational cost and storage requirement. Later on, Hirata also developed an equation generator for equation-of-motion (EOM) CC^{25–31} for neutral excitations (EE-EOM-CC),^{25–28} ionization potentials (IP-EOM-CC),^{32–36} and electron affinities (EA-EOM-CC).^{11,37–39} Hanrath *et al.* proposed an improved version of TCE by implementing the matrix-matrix multiplication-based antisymmetric tensor contraction for general CC.^{40,41} Meanwhile, Kong *et al.* developed the EOM version of state-specific multireference CC together with the automated implementation of such complicated equations.⁴² This was later generalized to arbitrary order.^{43–46} Likewise, Shiozaki implemented explicitly-correlated versions of CC in a similar way.^{47,48}

More recently appeared (i) the Symbolic Manipulation Interpreter for Theoretical cHemistry (SMITH3)⁴⁹ for complete active space methods that implies partial contractions of second-quantized operators, (ii) $\mathbf{p}^\dagger\mathbf{q}$ developed by Rubin and DePrince III⁵⁰ that combines C++ and python for proof-of-concept implementation of many-body quantum chemistry methods, and (iii) WICK&D⁵¹ that presents a strategy to evaluate Mukherjee's⁵² and Kutzelnigg's⁵³ version of Wick's theorem in the case of an arbitrary number of orbital subspaces. Each of the above programs is a clear example of the progress that has been achieved in the last three decades with respect to equation generators in quantum chemistry.

^{a)}Electronic mail: rquintero@irsamc.ups-tlse.fr

^{b)}Electronic mail: loos@irsamc.ups-tlse.fr

The aim of the present paper is to describe a four-step algorithm based on Wick's theorem to obtain the working equations of EE-EOM-CC, IP-EOM-CC, EA-EOM-CC, as well as other quantities of interest. In particular, the direct way of obtaining intermediates facilitates the comparison between many-body perturbation theory and CC methods.^{54–58} The present algorithm is implemented with the computer algebra system, MATHEMATICA,⁵⁹ in order to be able to generate and manipulate the equations in a user-friendly way.

This article is organized as follows. Section II gathers all theoretical details. In particular, Sec. II A describes Wick's theorem and the concept of normal ordering. Then, in Secs. II B and II C, we report the main theoretical details behind the CC and EOM-CC equations, respectively. Finally, in Sec. III, we provide a detailed description of our algorithm. Our conclusions are drawn in Sec. IV.

II. THEORY

A. Normal ordering and Wick's theorem

Our goal is to derive automatically the EOM-CC working equations which take the form of a matrix eigenvalue problem. Each element of the EOM-CC effective Hamiltonian is an expectation value formed by products of operators. (We shall discuss this point in more detail later in this section). An efficient way to evaluate these products of operators is via their second-quantized form.^{60–62}

As a starting point, let us introduce the reference state which is represented as a single determinant

$$|0\rangle = |ijk\dots\rangle = \hat{i}^\dagger \hat{j}^\dagger \hat{k}^\dagger \dots |\rangle. \quad (1)$$

obtained by acting on the physical (or true) vacuum $|\rangle$ with creation operators $\hat{i}^\dagger, \hat{j}^\dagger, \hat{k}^\dagger, \dots$. This reference state $|0\rangle$, typically taken as the Hartree-Fock (HF) ground-state determinant, can be used as an alternative vacuum, named Fermi vacuum.^{2,14} One property of the Fermi vacuum is that one can redefine the creation operators relative to it as particles ($\hat{a}^\dagger, \hat{b}^\dagger, \hat{c}^\dagger, \dots$) and holes ($\hat{i}, \hat{j}, \hat{k}, \dots$).^{63,64} The particle creation operators create particle states above the Fermi level while the hole creation operators remove particle states below the Fermi level (particle-hole formalism). Because one cannot annihilate a hole or remove a particle in the Fermi vacuum, we have $\hat{i}^\dagger |0\rangle = 0$ and $\hat{a} |0\rangle = 0$.

It is also possible to define “neutral” excited determinants by operating the same number of particle and hole operators onto the Fermi vacuum:

$$\hat{a}^\dagger \hat{b}^\dagger \hat{c}^\dagger \dots \hat{k} \hat{j} \hat{i} |0\rangle = |\Phi_{ijk\dots}^{abc\dots}\rangle. \quad (2)$$

In such a way, one can also define “charged” excited determinants where the number of particles and holes are different. For example, ionized and electron-attached determinants can be represented, respectively, as follows

$$\hat{a}^\dagger \hat{j} \hat{i} |0\rangle = |\Phi_{ij}^a\rangle, \quad \hat{a}^\dagger \hat{b}^\dagger \hat{i} |0\rangle = |\Phi_i^{ab}\rangle. \quad (3)$$

To preserve the antisymmetry of the electronic wave function, the second-quantized fermionic operators fulfill anti-commutation rules

$$\hat{p}\hat{q} + \hat{q}\hat{p} = 0, \quad \hat{p}^\dagger \hat{q}^\dagger + \hat{q}^\dagger \hat{p}^\dagger = 0, \quad \hat{p}^\dagger \hat{q} + \hat{q}^\dagger \hat{p} = \delta_{pq}, \quad (4)$$

where $\hat{p}, \hat{q}, \hat{r}, \dots$ are arbitrary operators that can be either hole or particle operators and δ_{pq} is a Kronecker delta.

Besides the anti-commutation rules defined in Eq. (4), there is another powerful “bookkeeping system” called normal ordering,⁶⁵ that consists of placing all the creation operators to the left and all annihilation operators to the right. Therefore, applying the Fermi vacuum on a normal-ordered string yield zero. At this stage, it is convenient to introduce a more compact notation to take full advantage of this bookkeeping system. To do so, we define the contraction of arbitrary operators

$$\overline{\hat{p}^\dagger \hat{q}} = \hat{p}^\dagger \hat{q} - \{\hat{p}^\dagger \hat{q}\} = \hat{p}^\dagger \hat{q} + \hat{q} \hat{p}^\dagger = \delta_{pq}, \quad (5)$$

where $\{\hat{p}^\dagger \hat{q}\}$ means that the product $\hat{p}^\dagger \hat{q}$ is normal ordered with respect to the Fermi vacuum. By definition, the only non-zero contractions in Eq. (5) are

$$\overline{\hat{a} \hat{b}^\dagger} = \delta_{ab}, \quad \overline{\hat{i}^\dagger \hat{j}} = \delta_{ij}. \quad (6)$$

When a product of second-quantized operators is normal-ordered, we name it a normal product.

We are now in a position to introduce the generalized Wick theorem stating that a product of creation and annihilation operators is equal to their normal product plus the sum of all possible contractions.² For example, applying this theorem to a given string of normal-ordered operators (which is typically what one has to do to derive the EOM-CC equations), one gets

$$\{\hat{A}\}\{\hat{B}\}\{\hat{C}\}\cdots = \{\hat{A}\hat{B}\hat{C}\cdots\} + \sum_{\text{singles}} \overbrace{\{\hat{A}\hat{B}\hat{C}\cdots\}} + \sum_{\text{doubles}} \overbrace{\{\hat{A}\hat{B}\hat{C}\cdots\}} + \cdots + \sum_{\text{fully contracted}} \overbrace{\{\hat{A}\hat{B}\hat{C}\cdots\}}, \quad (7)$$

where the first term on the right-hand side is the normal product, the second term contains all the single contractions, the third contains all double contractions, and the last sum gathers the so-called fully contracted terms. The key feature of this theorem is that, if we evaluate the Fermi vacuum expectation value associated with Eq. (7), the only non-zero terms correspond to the fully contracted ones, *i.e.*,

$$\langle 0 | \{\hat{A}\}\{\hat{B}\}\{\hat{C}\}\cdots | 0 \rangle = \sum_{\text{fully contracted}} \langle 0 | \overbrace{\{\hat{A}\hat{B}\hat{C}\cdots\}} | 0 \rangle \quad (8)$$

To have a consistent nomenclature, the products of second-quantized operators $\{\hat{A}\}\{\hat{B}\}\{\hat{C}\}$ are called strings, and each operator $\{\hat{A}\}$ is called a sub-string.

B. Coupled-cluster equations

To derive the CC equations, one first starts with the definition of the CC wave function which is an exponential parametrization applied to a reference determinant^{66,67}

$$|\Psi^{(0)}\rangle = e^{\hat{T}} |0\rangle, \quad (9)$$

where $\Psi^{(0)}$ is the CC ground-state wave function and

$$\hat{T} = \sum_{n=1}^N \hat{T}_n \quad (10)$$

is the cluster operator with

$$\hat{T}_n = \frac{1}{(n!)^2} \sum_{ijk\cdots} \sum_{abc\cdots} t_{ijk\cdots}^{abc\cdots} \{\hat{a}^\dagger \hat{b}^\dagger \hat{c}^\dagger \hat{k} \cdots\}, \quad (11)$$

where $t_{ijk\cdots}^{abc\cdots}$ are the antisymmetric cluster amplitudes and the indices i, j, k, \dots indicate occupied spinorbitals in the reference configuration, while a, b, c, \dots are unoccupied spinorbitals. In general, \hat{T}_n produces n -fold excited N -electron determinants of the form $|\Phi_{ijk\cdots}^{abc\cdots}\rangle$ (with $n \leq N$). It is worth mentioning here that the excitation operators do commute with each other, *i.e.*, $[\hat{T}_n, \hat{T}_m] = 0$.

The electronic Schrödinger equation associated with the CC wave function is

$$\hat{H} |\Psi^{(0)}\rangle = E^{(0)} |\Psi^{(0)}\rangle, \quad (12)$$

where $E^{(0)}$ is the CC ground-state electronic energy, and \hat{H} is the electronic Hamiltonian, which can be expressed in second-quantized form as follows:

$$\hat{H} = \sum_{pq} h_q^p \hat{p}^\dagger \hat{q} + \frac{1}{4} \sum_{pqrs} v_{rs}^{pq} \hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r}, \quad (13)$$

where the indices p, q, r , and s indicate arbitrary (*i.e.*, occupied or virtual) spinorbitals. The matrix element h_q^p is the sum of the kinetic and nuclear attraction components. The electronic repulsion is accounted for by the (antisymmetrized) two-electron integrals v_{rs}^{pq} . Because one is usually interested in the correlated part of the Hamiltonian, it is common practice to divide it as

$$\hat{H} = \hat{H}_N + \langle 0 | \hat{H} | 0 \rangle, \quad (14)$$

where the first term

$$\hat{H}_N = \hat{F}_N + \hat{V}_N = \sum_{pq} f_q^p \{\hat{p}^\dagger \hat{q}\} + \frac{1}{4} \sum_{pqrs} v_{rs}^{pq} \{\hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r}\} \quad (15)$$

is the normal-ordered Hamiltonian (where f_q^p is an element of the Fock matrix) that corresponds to fluctuations (*i.e.*, correlation) with respect to the second term that represents the reference energy $E_0 = \langle 0 | \hat{H} | 0 \rangle$. It is worth emphasizing that \hat{H}_N do not commute with \hat{T}_n , *i.e.*, $[\hat{T}_n, \hat{H}_N] \neq 0$.

Thanks to the exponential ansatz of the CC wave function, Eq. (9), and the introduction of the following CC effective Hamiltonian via a similarity transformation

$$\bar{H}^{\text{st}} = e^{-\hat{T}} \hat{H}_N e^{\hat{T}}, \quad (16)$$

where the superindex “st” stands for “similarity-transformed”. Then, one can recast the Schrödinger equation as

$$\bar{H}^{\text{st}} |0\rangle = \Delta E^{(0)} |0\rangle. \quad (17)$$

where $\Delta E^{(0)} = E^{(0)} - E_0$ is the CC ground-state correlation energy obtained via projection

$$\langle 0 | \bar{H}^{\text{st}} | 0 \rangle = \Delta E^{(0)}, \quad (18)$$

while

$$\langle \Phi_{ijk\dots}^{abc\dots} | \bar{H} | 0 \rangle = 0 \quad (19)$$

are the so-called amplitude equations, a set of non-linear equations where the unknowns are the cluster amplitudes $t_{ijk\dots}^{abc\dots}$.

To evaluate efficiently Eq. (19), one usually relies on the Baker-Campbell-Hausdorff expansion^{68,69} of the similarity-transformed Hamiltonian,

$$\bar{H}^{\text{st}} = \hat{H}_N + [\hat{H}_N, \hat{T}] + \frac{1}{2!} [[\hat{H}_N, \hat{T}], \hat{T}] + \frac{1}{3!} [[[\hat{H}_N, \hat{T}], \hat{T}], \hat{T}] + \frac{1}{4!} [[[[\hat{H}_N, \hat{T}], \hat{T}], \hat{T}], \hat{T}], \quad (20)$$

a series of nested commutators between \hat{H}_N and \hat{T} that naturally terminates at the four-fold commutator thanks to the two-body nature of the electronic Hamiltonian.

Then, using Wick’s theorem [see Eq. (7)], one can establish that the non-zero terms in Eq. (20) involve \hat{H}_N as the left-most operator contracted with at most four \hat{T} operators, *i.e.*,

$$\bar{H}^{\text{st}} = \hat{H}_N + \overbrace{\hat{H}_N \hat{T}} + \frac{1}{2} \overbrace{\hat{H}_N \hat{T} \hat{T}} + \frac{1}{3!} \overbrace{\hat{H}_N \hat{T} \hat{T} \hat{T}} + \frac{1}{4!} \overbrace{\hat{H}_N \hat{T} \hat{T} \hat{T} \hat{T}}, \quad (21)$$

where the multi-leg symbol indicates that \hat{H}_N must be contracted at least once with each \hat{T} . Hence, \hat{H}_N is said to be connected as it does not include disconnected terms of the form $\overbrace{\hat{H}_N \hat{T} \hat{T} \hat{T}}$.

If one introduces the right-hand side of Eq. (21) in Eqs. (18) and (19), one gets

$$\langle 0 | \left(\hat{H}_N e^{\hat{T}} \right)_C | 0 \rangle = \Delta E^{(0)}, \quad (22a)$$

$$\langle \Phi_{ijk\dots}^{abc\dots} | \left(\hat{H}_N e^{\hat{T}} \right)_C | 0 \rangle = 0. \quad (22b)$$

where the subindex C stands for connected. Since one only deals with the connected terms, the algebraic form of Eqs. (22a) and (22b) is simpler than Eqs. (18) and (19).

C. Equation-of-motion coupled-cluster equations

Having explained how to derive the CC equations, next, we shift our attention to the EOM-CC equations.^{25,29–31} As a starting point, let us consider the Schrödinger equation for a target excited state $\Psi^{(k)}$, *i.e.*,

$$\hat{H}_N |\Psi^{(k)}\rangle = \Delta E^{(k)} |\Psi^{(k)}\rangle, \quad (23)$$

with $\Delta E^{(k)} = E^{(k)} - E_0$. One way to access this target state is by transforming the initial state described by Eq. (17) via an excitation operator of the form

$$\hat{\mathcal{R}}^{(k)} |\Psi^{(0)}\rangle = |\Psi^{(k)}\rangle \quad (24)$$

If one wants to access neutral excited states, one must rely on the excitation energy (EE) operator

$$\hat{\mathcal{R}}^{(k)} = \hat{1} + \hat{\mathcal{R}}_1^{(k)} + \hat{\mathcal{R}}_2^{(k)} + \dots \quad (25)$$

with

$$\hat{\mathcal{R}}_1^{(k)} = \sum_i \sum_a r_i^a \{\hat{a}^\dagger \hat{i}\}, \quad (26a)$$

$$\hat{\mathcal{R}}_2^{(k)} = \frac{1}{4} \sum_{ij} \sum_{ab} r_{ij}^{ab} \{\hat{a}^\dagger \hat{i} \hat{b}^\dagger \hat{j}\}, \quad (26b)$$

\vdots

where $\hat{\mathcal{R}}_n^{(k)}$ is an excitation operator of degree n for the state k associated with the EOM-CC amplitudes $r_{ij\dots}^{ab\dots}$. (Here, for the sake of simplicity, we have omitted the k -dependence of the amplitudes.)

After several transformations, the Schrödinger equation for the target excited state [see Eq. (23)] can be recast as

$$\left[\bar{H}^{\text{st}}, \hat{\mathcal{R}}^{(k)} \right] |0\rangle = \Omega_k \hat{\mathcal{R}}^{(k)} |0\rangle. \quad (27)$$

where $\Omega_k = E^{(k)} - E^{(0)}$ is the excitation energy associated with the k th excited state. Defining the normal-ordered similarity-transformed Hamiltonian, $\bar{H}_N^{\text{st}} = \bar{H} - E^{(0)}$, the commutator in Eq. (27) can be further simplified as

$$\left(\bar{H}_N^{\text{st}} \hat{\mathcal{R}}^{(k)} \right)_C |0\rangle = \Omega_k \hat{\mathcal{R}}^{(k)} |0\rangle. \quad (28)$$

The excitation energies are then directly computed by projecting Eq. (28) in a determinant basis truncated at a given excitation degree, which is equivalent to truncate $\hat{\mathcal{R}}^{(k)}$.

Introducing the simplified notation, $|\Phi_i^a\rangle \equiv |S\rangle$, $|\Phi_{ij}^{ab}\rangle \equiv |D\rangle$, ..., the EOM-CC linear eigenvalue problem has the following form:

$$\bar{\mathcal{H}} \mathbf{r}_k = \Omega_k \mathbf{r}_k \Rightarrow \begin{pmatrix} \langle S | \bar{H}_N^{\text{st}} | S \rangle & \langle S | \bar{H}_N^{\text{st}} | D \rangle & \dots \\ \langle D | \bar{H}_N^{\text{st}} | S \rangle & \langle D | \bar{H}_N^{\text{st}} | D \rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} r_i^a \\ r_{ij}^{ab} \\ \vdots \end{pmatrix} = \Omega_k \begin{pmatrix} r_i^a \\ r_{ij}^{ab} \\ \vdots \end{pmatrix}. \quad (29)$$

Due to the non-hermitian nature of this eigenvalue problem, each block of the matrix has to be computed.

Following the same procedure, one can also obtain the EOM-CC equations for the ionized and electron-attached states, just by changing the definition of the operator $\hat{\mathcal{R}}^{(k)}$ in Eq. (26). For example, removing one or two particles, one gets IP-EOM-CC and DIP-EOM-CC,^{70–74} respectively, while removing one or two holes yields EA-EOM-CC and DEA-EOM-CC,^{73,74} respectively.

D. Many-body Hamiltonian

When one computes the expectation value associated with a given block of the EOM-CC matrix $\bar{\mathcal{H}}$ [see Eq. (29)], redundant terms are generated. One way to avoid redundancies is to rely on the many-body representation of the EOM-CC effective Hamiltonian^{75–78}

$$\bar{H}_N^{\text{mb}} = \sum_{pq} \chi_q^p \{\hat{p}^\dagger \hat{q}\} + \sum_{pqrs} \chi_{rs}^{pq} \{\hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r}\} + \chi_{stu}^{pqr} \{\hat{p}^\dagger \hat{q}^\dagger \hat{r}^\dagger \hat{u} \hat{t} \hat{s}\} + \dots, \quad (30)$$

where χ_q^p , χ_{rs}^{pq} , χ_{stu}^{pqr} are one-, two-, and three-body terms respectively. The superindex “mb” stands for “many-body”. Note that \bar{H}_N^{mb} and \bar{H}_N^{st} are exactly the same quantity; the labels “mb” and “st” are here to indicate the type of expansion.

The many-body terms $\{\chi_{stu...}^{pqr...}\}$ are usually represented through diagrammatic techniques.² In the present work though, we are interested in describing them within the second quantization formalism. To achieve this, we take as an example the block $\langle D | \bar{H}_N^{\text{mb}} | S \rangle$ and the many-body Hamiltonian defined in Eq. (30) to obtain

$$\begin{aligned} \langle \Phi_{ij}^{ab} | \bar{H}_N^{\text{mb}} | \Phi_k^c \rangle = & \sum_{pq} \chi_q^p \langle 0 | \{\hat{i}^\dagger \hat{j}^\dagger \hat{b} \hat{a}\} \{\hat{p}^\dagger \hat{q}\} \{\hat{c}^\dagger \hat{k}\} | 0 \rangle + \sum_{pqrs} \chi_{rs}^{pq} \langle 0 | \{\hat{i}^\dagger \hat{j}^\dagger \hat{b} \hat{a}\} \{\hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r}\} \{\hat{c}^\dagger \hat{k}\} | 0 \rangle \\ & + \sum_{pqrstu} \chi_{stuv}^{pqr} \langle 0 | \{\hat{i}^\dagger \hat{j}^\dagger \hat{b} \hat{a}\} \{\hat{p}^\dagger \hat{q}^\dagger \hat{r}^\dagger \hat{u} \hat{t} \hat{s}\} \{\hat{c}^\dagger \hat{k}\} | 0 \rangle. \end{aligned} \quad (31)$$

Equation (31) effectively terminates at the three-body level since higher-order terms produce partial contractions which are zero with respect to the Fermi vacuum.

By applying Wick's theorem to Eq. (31) and taking advantage of the antisymmetric permutation of the many-body terms, one gets

$$\langle \Phi_{ij}^{ab} | \bar{H}_N^{\text{mb}} | \Phi_k^c \rangle = \delta_{jk} \delta_{bc} \chi_i^a - \delta_{ik} \delta_{bc} \chi_j^a - \delta_{jk} \delta_{ac} \chi_i^b + \delta_{jk} \delta_{ac} \chi_j^b - \delta_{bc} \chi_{ij}^{ak} - \delta_{jk} \chi_{ci}^{ab} + \delta_{ik} \chi_{cj}^{ab} + \delta_{ac} \chi_{ij}^{bk} + \chi_{ijc}^{abk}. \quad (32)$$

Since $\bar{H}_N^{\text{mb}} = \bar{H}_N^{\text{st}}$, the block $\langle D | \bar{H}_N^{\text{mb}} | S \rangle$ can be naturally written in terms of \hat{T} , \hat{F}_N , and \hat{V}_N . To illustrate this, we consider the general many-body term

$$\chi_{ij...cd...}^{ab...kl...} = \langle \Phi_{ij...}^{ab...} | \bar{H}_N^{\text{st}} | \Phi_{kl...}^{cd...} \rangle_{iC}, \quad (33)$$

where the subscript iC represents “internal contractions”, *i.e.*, contractions occurring exclusively between $\langle \Phi_{ij...}^{ab...} |$ and \bar{H}_N , as well as \bar{H}_N and $|\Phi_{kl...}^{cd...}\rangle$. Notably, contractions between $\langle \Phi_{ij...}^{ab...} |$ and $|\Phi_{kl...}^{cd...}\rangle$ are not allowed.

Taking into account Eq. (33) and the non-zero matrix elements of χ_{ij}^{ak} , one gets

$$\chi_{ij}^{ak} = \left\langle \Phi_{ij}^a \left| \left(\hat{V}_N + \hat{T}_1 \hat{V}_N + \hat{F}_N \hat{T}_2 + \hat{V}_N \hat{T}_2 + \frac{1}{2!} \hat{V}_N \hat{T}_1^2 + \hat{V}_N \hat{T}_1 \hat{T}_2 + \frac{1}{3!} \hat{V}_N \hat{T}_1^3 \right) \right| \Phi_k^c \right\rangle_{iC}. \quad (34)$$

Applying the same procedure for χ_{ci}^{ab} and χ_{ijc}^{abk} , we obtain

$$\chi_{ic}^{ab} = \left\langle \Phi_i^{ab} \left| \left(\hat{V}_N + \hat{T}_1 \hat{V}_N + \hat{F}_N \hat{T}_2 + \hat{V}_N \hat{T}_2 + \frac{1}{2!} \hat{V}_N \hat{T}_1^2 + \hat{V}_N \hat{T}_1 \hat{T}_2 + \frac{1}{3!} \hat{V}_N \hat{T}_1^3 \right) \right| \Phi_c^c \right\rangle_{iC}, \quad (35)$$

$$\chi_{ijc}^{abk} = \left\langle \Phi_{ij}^{ab} \left| \left(\hat{V}_N \hat{T}_2 \right) \right| \Phi_k^c \right\rangle_{iC}. \quad (36)$$

The one-body terms in Eq. (32) are defined as follows

$$\chi_i^a = \left\langle \Phi_i^a \left| \left[\hat{H}_N \left(1 + \hat{T}_1 + \hat{T}_2 + \frac{\hat{T}_1^2}{2} + \hat{T}_1 \hat{T}_2 + \frac{\hat{T}_1^3}{3!} \right) \right] \right| 0 \right\rangle_C, \quad (37)$$

which is identical to the CCS amplitude equations. Consequently, all the one-body terms are zero in Eq. (32). This is also true for the many-body terms χ_{ij}^{ab} for CCSD, χ_{ijk}^{abc} for CCSDT, and so on.

Setting the one-body terms equal to zero in Eq. (37), we have

$$\langle \Phi_{ij}^{ab} | \bar{H}_N^{\text{mb}} | \Phi_k^c \rangle = -\delta_{bc} \chi_{ij}^{ak} - \delta_{jk} \chi_{ci}^{ab} + \delta_{ik} \chi_{cj}^{ab} + \delta_{ac} \chi_{ij}^{bk} + \chi_{ijc}^{abk}. \quad (38)$$

It is possible to obtain the rest of the many-body terms in Eq. (38) by substituting the indices of χ_{ij}^{ab} , χ_{ci}^{ab} , and χ_{ijc}^{abk} . These replacements are restricted to indices of the same category: hole-creation (i, j), particle-annihilation (a, b), hole-annihilation (k), and particle-creation (c).

To obtain the block $\langle \Phi_{ij}^{ab} | \bar{H}_N^{\text{mb}} | \Phi_k^c \rangle$ in a compact and non-redundant form, we rely on Eqs. (34), (35), and (36) to define the many-body terms and Eq. (38) to define the blocks in terms of these many-body terms. Finally, to generate the entire eigenvalue equations for EE-EOM-CCSD, one must repeat the same procedure for the blocks $\langle \Phi_i^a | \bar{H}_N^{\text{st}} | \Phi_k^c \rangle$, $\langle \Phi_{ij}^a | \bar{H}_N^{\text{st}} | \Phi_{kl}^{cd} \rangle$, $\langle \Phi_{ij}^{ab} | \bar{H}_N^{\text{st}} | \Phi_{kl}^{cd} \rangle$, and then remove the redundant many-body terms.

TABLE I. List of variables used in the EOM-CC equation generator program.

Description	Mathematical symbol	MATHEMATICA notation
Elements of the Fock matrix	f_q^{pp}	<code>F[[p,q]]</code>
Two-electron repulsion integrals	v_{rs}^{pq}	<code>ERI[[p,q,r,s]]</code>
Amplitudes of single excitations	t_i^a	<code>t1[[i,a]]</code>
Amplitudes of double excitations	t_{ij}^{ab}	<code>t2[[i,j,a,b]]</code>
Amplitudes of triple excitations	t_{ijk}^{abc}	<code>t3[[i,j,k,a,b,c]]</code>
Amplitudes of quadruple excitations	t_{ijkl}^{abcd}	<code>t4[[i,j,k,l,a,b,c,d]]</code>
One-body terms	χ_q^p	<code>χ1[[p,q]]</code>
Two-body terms	χ_{rs}^{pq}	<code>χ2[[p,q,r,s]]</code>
Three-body terms	χ_{stu}^{pqr}	<code>χ3[[p,q,r,s,t,u]]</code>
Four-body terms	χ_{tuvw}^{pqrs}	<code>χ4[[p,q,r,s,t,u,v,w]]</code>

III. EOM-CC EQUATION GENERATOR

In this section, we describe the algorithm that we have written with MATHEMATICA to automatically derive the working equations of EOM-CC. Although the program is not optimized for efficiency, our goal is to obtain the working equations in terms of the many-body terms by using straightforward input quantities. First of all, we adapt the index notations in order to have more suitable and general notations to obtain the EOM-CC equations up to fourth order. For the operators belonging to the bra and the ket, instead of using $\hat{a}^\dagger, \hat{b}^\dagger, \dots, \hat{i}, \hat{j}, \dots$, we use $\hat{p}_1^\dagger, \hat{p}_2^\dagger, \dots, \hat{h}_1, \hat{h}_2, \dots$. For the particles/holes that play the role of dummy indices in the cluster operator \hat{T} , instead of using $\hat{e}^\dagger, \hat{f}^\dagger, \dots, \hat{m}, \hat{n}, \dots$, we use $\hat{o}_1^\dagger, \hat{o}_2^\dagger, \dots, \hat{v}_1, \hat{v}_2, \dots$, where the notation of o and v refers to occupied and virtual, respectively. Finally, for the arbitrary indexes that can be either particle or hole [like the Hamiltonian in Eq. (14)], we switch from $\hat{p}^\dagger, \hat{q}^\dagger, \hat{r}, \hat{s}$ to $\hat{q}_1^\dagger, \hat{q}_2^\dagger, \hat{q}_3, \hat{q}_4$.

Thanks to this change in notation, we can now describe the four-step algorithm that we use to derive the EOM-CC equations. Each step of the algorithm is written as a MATHEMATICA module. These modules are gathered within a main module called `eomccgen`. Variables tailored-made for MATHEMATICA are summarized in Table I alongside their description and corresponding mathematical expression.

The input quantities for `eomccgen` are the list of (neutral) excitation operators, named `ClusterOperator` (that determine the similarity-transform Hamiltonian) and the list of EOM operators, named `EOMOperator`. For example, the input to obtain the EE-EOM-CCSDT equations is

$$\begin{aligned} \text{ClusterOperator} &= \{ \{ "1h1p" \}, \{ "2h2p" \}, \{ "3h3p" \} \} \\ \text{EOMOperator} &= \{ \{ "1h1p" \}, \{ "2h2p" \}, \{ "3h3p" \} \} \end{aligned}$$

If one wants to get IP-EOM-CCSD, the input is

$$\begin{aligned} \text{ClusterOperator} &= \{ \{ "1h1p" \}, \{ "2h2p" \} \} \\ \text{EOMOperator} &= \{ \{ "1h0p" \}, \{ "2h1p" \} \} \end{aligned}$$

To obtain DEA-EOM-CCSD, the second line has to be changed as

$$\text{EOMOperator} = \{ \{ "0h2p" \}, \{ "1h3p" \} \}$$

It is also possible to have different numbers of operators in `EOMOperator` and `ClusterOperator`. For example, the EA-EOM-CCSD(3h2p) equations can be produced as

$$\text{EOMOperator} = \{ \{ "0h1p" \}, \{ "1h2p" \}, \{ "3h2p" \} \}$$

In practice, it is possible to carry out all possible combinations of excitation operators up to 4h4p, and it is also possible to calculate the CC amplitude equations. For example, to generate the CCSD amplitude equations, the following input must be entered

$$\begin{aligned} \text{ClusterOperator} &= \{ \{ "1h1p" \}, \{ "2h2p" \} \} \\ \text{EOMOperator} &= \{ \{ "0h0p" \} \} \end{aligned}$$

We shall now discuss in detail the algorithm summarized in Fig. 1.

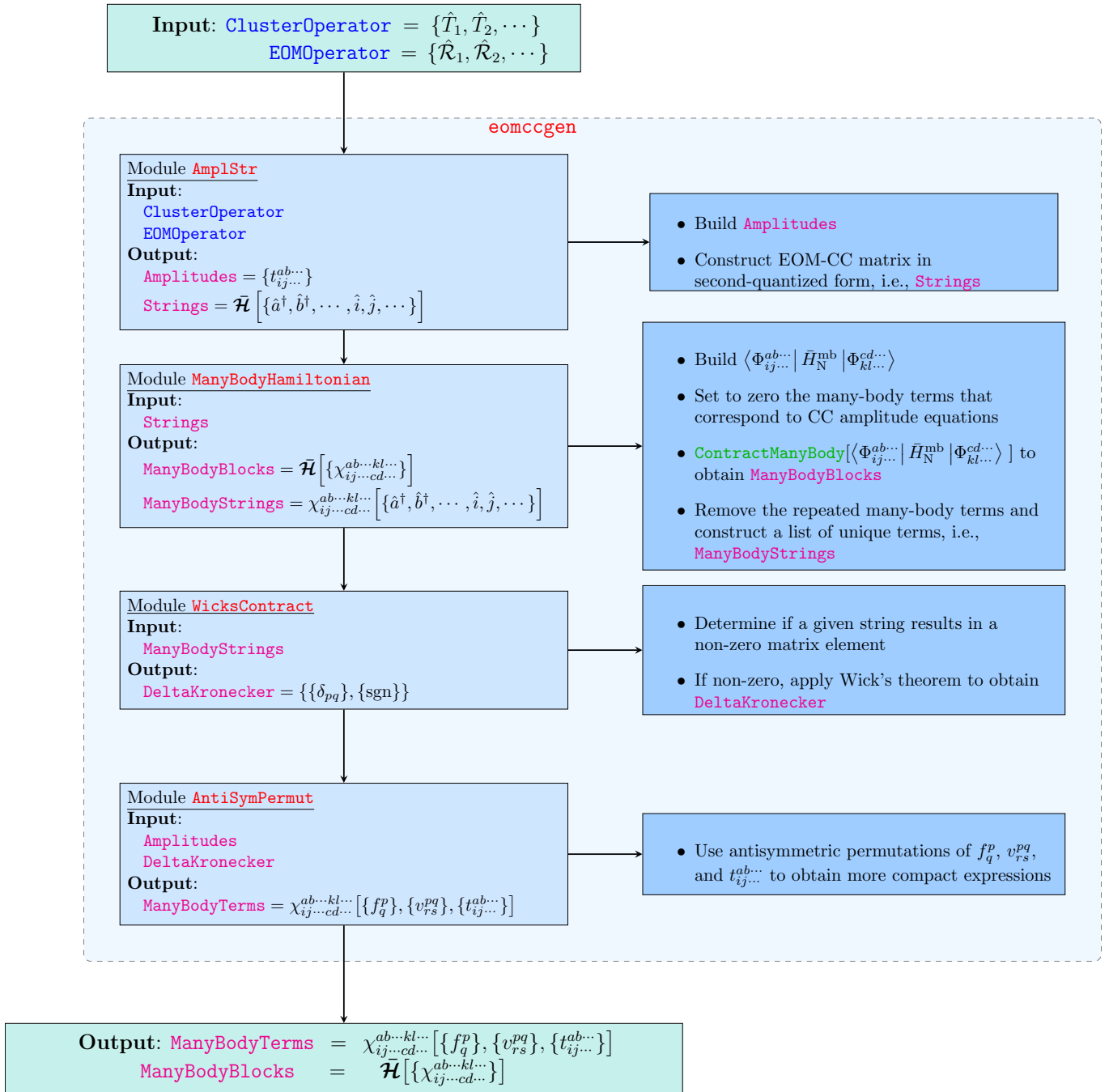


FIG. 1. Flowchart of the entire set of modules constituting the EOM-CC equation generator.

A. Step 1

The first step is carried out through the module **AmplStr**. The inputs of this module are the **ClusterOperator** and **EOMOperator**:

$$\{\text{Amplitudes}, \text{Strings}\} = \text{AmplStr}[\text{ClusterOperator}, \text{EOMOperator}]$$

It returns the list of CC amplitudes (stemming from the truncated Taylor expansion of $e^{\hat{T}}$) in **Amplitudes**. For example, in the case of EE-EOM-CCSD, we have

$$\text{Amplitudes} = \{1, t_{o_1}^{v_1}, t_{o_1}^{v_1} t_{o_2}^{v_2}, t_{o_1}^{v_1} t_{o_2}^{v_2} t_{o_3}^{v_3}, t_{o_1}^{v_1} v_{o_2}^{v_2}, t_{o_1}^{v_1} v_{o_2}^{v_2} t_{o_3}^{v_3}\}$$

while, in MATHEMATICA, it takes the following form:

```
1 Amplitudes={{1},{t1[o1,v1]},{t1[o1,v1]t1[o2,v2]},{t1[o1,v1]t1[o2,v2]t1[o3,v3]},{t2[o1,o2,v1,v2]},{t1[o1,v1]t2[o2,o3,v2,v3]}}
```

The set of excitation operators gathered in `EOMOperator` determines the basis in which $\bar{\mathcal{H}}$ [see Eq. (29)] is diagonalized. With this basis together with the similarity-transformed Hamiltonian, the EOM-CC matrix is written in second-quantized form, $\bar{\mathcal{H}}[\{\hat{p}_1^\dagger, \hat{p}_2^\dagger, \dots, \hat{h}_1, \hat{h}_2, \dots\}]$, and stored in the variable `Strings`, which is made of two substrings, i.e., `Strings` = {`StringF`, `StringV`} (F and V stand for \hat{F}_N and \hat{V}_N). For example, the element `StringV`[[2,1,6]] corresponds to

$$\langle D | \hat{V}_N \hat{T}_1 \hat{T}_2 | S \rangle = \{\{h1^\dagger, h2^\dagger, p2, p1\}, \{q1^\dagger, q2^\dagger, q4, q3\}, \{\{v1^\dagger, o1\}, \{v2^\dagger, o2, v3^\dagger, o3\}\}, \{p3^\dagger, h3\}\}$$

The values 2 and 1 in `StringV`[[2,1,6]] correspond to the indices of the row and column of the EOM-CC matrix, respectively, while 6 represents the sixth amplitude, i.e., `Amplitudes`[[6]]. For the sake of simplicity, the summation and the Fermi vacuum of `StringF` and `StringV` are implicit.

B. Step 2

In Fig. 1, the module that corresponds to the second step is named `ManyBodyHamiltonian` and requires the input quantity `Strings` generated in `AmplStr` (see Subsec. III A):

$$\{\text{ManyBodyBlocks}, \text{ManyBodyStrings}\} = \text{ManyBodyHamiltonian}[\text{Strings}]$$

where `ManyBodyBlocks` corresponds to the EOM-CC matrix defined in Eq. (29) written in terms of the many-body terms, i.e., $\bar{\mathcal{H}}[\{\chi_{h_1 h_2 \dots p_3 p_4 \dots}^{p_1 p_2 \dots h_3 h_4 \dots}\}]$, and `ManyBodyStrings` is a list that consists of all unique many-body terms in second-quantized form, i.e., $\chi_{h_1 h_2 \dots p_3 p_4 \dots}^{p_1 p_2 \dots h_3 h_4 \dots}[\{\hat{p}_1^\dagger, \hat{p}_2^\dagger, \dots, \hat{h}_1, \hat{h}_2, \dots\}]$.

More precisely, the algorithm inside `ManyBodyHamiltonian` that generates `ManyBodyBlocks` and `ManyBodyStrings` performs the following steps:

1. Extract the EOM basis from `Strings` and evaluate the many-body Hamiltonian represented in Eq. (30) to form matrix elements, such as $\langle \Phi_{h_1 h_2 \dots}^{p_1 p_2 \dots} | \bar{H}_N^{\text{mb}} | \Phi_{h_3 h_4 \dots}^{p_3 p_4 \dots} \rangle$.
2. Set to zero the many-body terms that correspond to the CC amplitude equations, such as the ones defined in Eq. (37).
3. Contract blocks made of second-quantized strings of the form $\langle \Phi_{h_1 h_2 \dots}^{p_1 p_2 \dots} | \bar{H}_N^{\text{mb}} | \Phi_{h_3 h_4 \dots}^{p_3 p_4 \dots} \rangle$ with the help of the `ContractManyBody` function. For example, the contraction of the block $\langle \Phi_{h_1}^{p_1} | \bar{H}_N^{\text{mb}} | \Phi_{h_3}^{p_3} \rangle$ generates

$$\{-\delta_{p_1 p_3} \chi_{h_1}^{h_3}, \delta_{h_1 h_3} \chi_{p_3}^{p_1}, -\chi_{h_1 p_3}^{p_1 h_3}\} = \text{ContractManyBody}[\langle \Phi_{h_1}^{p_1} | \bar{H}_N^{\text{mb}} | \Phi_{h_3}^{p_3} \rangle]$$

Applying `ContractManyBody` to each block of $\bar{\mathcal{H}}$ allows us to obtain the EOM-CC matrix in terms of many-body terms, i.e., $\bar{\mathcal{H}}[\{\chi_{h_1 h_2 \dots p_3 p_4 \dots}^{p_1 p_2 \dots h_3 h_4 \dots}\}]$. These are stored in `ManyBodyBlocks`. In MATHEMATICA, it takes the following form:

```
1 ManyBodyBlocks={-delta[p1,p3]chi[h3,h1]+delta[h1,h3]chi[p1,p3]-chi2[h3,p1,h1,p3]}
```

4. For each block of `ManyBodyBlocks`, we remove any redundant many-body term to create a list of unique terms, $\chi_{h_1 h_2 \dots p_3 p_4 \dots}^{p_1 p_2 \dots h_3 h_4 \dots}[\{\hat{p}_1^\dagger, \hat{p}_2^\dagger, \dots, \hat{h}_1, \hat{h}_2, \dots\}]$, stored in `ManyBodyStrings`. This is done by counting the number of hole-annihilation, hole-creation, particle-annihilation, and particle-creation operators in each term. For example, in the case of IP-EOM-CCSD, we have

$$\text{ManyBodyStrings} = \{\chi_{h_1}^{h_3}, \chi_{p_3}^{h_4}, \chi_{p_3}^{p_1}, \chi_{h_1 h_2}^{h_3 p_1}, \chi_{h_2 p_3}^{h_4 p_1}, \chi_{h_1 h_2 p_3}^{h_3 h_4 p_1}\}$$

Next, we express each many-body term as in Eq. (33) to obtain their second-quantized form. For example,

$$\chi_{h_1 h_2 p_3}^{h_3 h_4 p_1} = \langle \Phi_{h_1 h_2}^{p_1} | \bar{H}_N^{\text{st}} | \Phi_{h_3 h_4}^{p_3} \rangle_{iC} \quad (39)$$

while, in MATHEMATICA, this three-body term corresponds to the sixth position of `ManyBodyStrings`:

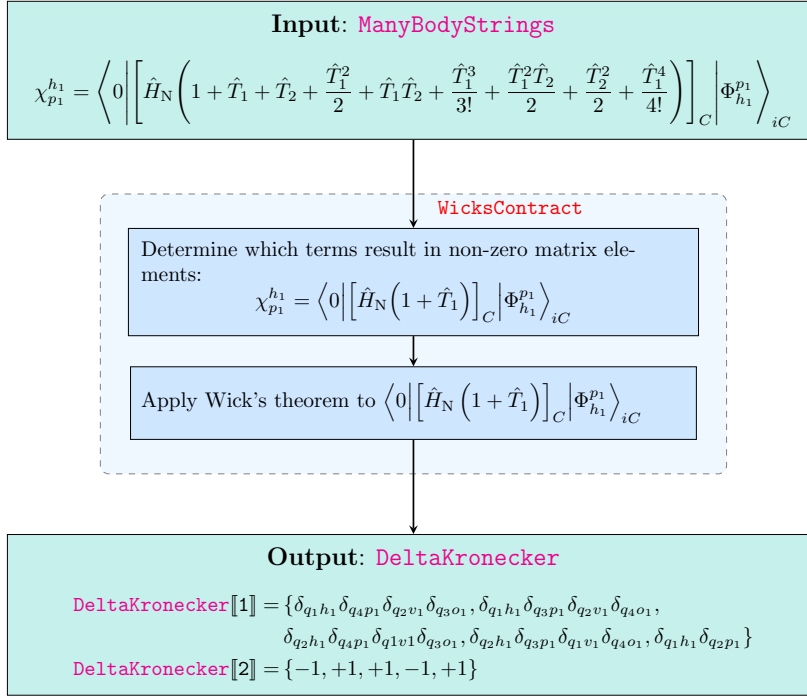


FIG. 2. Flowchart of the module **WicksContract**. The module takes as input $\chi_{p_1}^{h_1} = \langle 0 | \hat{H}_N | \Phi_{h_1}^{p_1} \rangle_{iC}$ stored in the variable **ManyBodyStrings** produced from the module **ManyBodyHamiltonian** (see Fig. 1).

```

1 ManyBodyBlocks[6]={{{SuperDagger[h1],SuperDagger[h2],p1},{SuperDagger[q1],SuperDagger[q2],q4,
q3},{},{SuperDagger[p3],h4,h3}},{SuperDagger[h1],SuperDagger[h2],p1},{SuperDagger[q1],
SuperDagger[q2],q4,q3},{SuperDagger[v1],o1},{SuperDagger[p3],h4,h3}},{SuperDagger[h1],
SuperDagger[h2],p1},{SuperDagger[q1],SuperDagger[q2],q4,q3},{SuperDagger[v1],o1},{
SuperDagger[v2],o2},{SuperDagger[p3],h4,h3}},{SuperDagger[h1],SuperDagger[h2],p1},{
SuperDagger[q1],SuperDagger[q2],q4,q3},{SuperDagger[v1],o1},{SuperDagger[v2],o2},{
SuperDagger[v3],o3},{SuperDagger[p3],h4,h3}},{SuperDagger[h1],SuperDagger[h2],p1},{
SuperDagger[q1],SuperDagger[q2],q4,q3},{SuperDagger[v1],o1,SuperDagger[v2],o2},{
SuperDagger[p3],h4,h3}},{SuperDagger[h1],SuperDagger[h2],p1},{SuperDagger[q1],SuperDagger[
q2],q4,q3},{SuperDagger[v1],o1},{SuperDagger[v2],o2,SuperDagger[v3],o3}},{SuperDagger[p3],
h4,h3}}}
```

C. Step 3

The third step is carried out by the module named **WicksContract**, which performs Wick's contractions [see Eq. (8)] of the second-quantized strings in **ManyBodyStrings** obtained in the module **ManyBodyHamiltonian**, as follows:

$$\text{DeltaKronecker} = \text{WicksContract}[\text{ManyBodyStrings}]$$

The output **DeltaKronecker** contains two quantities, the Kronecker deltas and the signs of each contraction.

Figure 2 illustrates the structure of the module **WicksContract**. Its input is the one-body term $\chi_{p_1}^{h_1}$ which is stored in the variable **ManyBodyStrings**[3] produced by the module **ManyBodyHamiltonian** discussed in Subsec. III B. The first step consists of replacing the similarity-transformed Hamiltonian with the terms yielding non-zero matrix elements and then applying Wick's theorem to the remaining terms. Finally, the Kronecker deltas and signs obtained from Wick's theorem are stored in **DeltaKronecker**.

D. Step 4

The module **AntiSymPermut** requires three inputs: **DeltaKronecker**, **ManyBodyStrings**, and **Amplitudes**.

$$\text{ManyBodyTerms} = \text{AntiSymPermut}[\text{DeltaKronecker}, \text{ManyBodyStrings}, \text{Amplitudes}]$$

This module applies the Kronecker deltas obtained in **WicksContract** (see Subsec. III C) to the **Amplitudes** (generated in **AmplStr** and discussed in Subsec. III A), the Fock matrix elements, $\{f_q^p\}$, and the two-electron repulsion integrals, $\{v_{rs}^{pq}\}$. Next, it takes advantage of the antisymmetric permutations of the amplitudes and integrals to obtain more compact expressions.

Finally, as output, **ManyBodyBlocks** obtained in the module **ManyBodyHamiltonian** (see Subsec. III B) and **ManyBodyTerms**, $\chi_{h_1 h_2 \dots p_3 p_4 \dots}^{p_1 p_2 \dots h_3 h_4 \dots} [\{f_q^p\}, \{v_{rs}^{pq}\}, \{t_{h_1 h_2 \dots}^{p_1 p_2 \dots}\}]$, from **AntiSymPermut** are printed. From these, one can easily generate the EOM-CC equations.

E. Examples

The software has also the advantage of being able to generate various outputs, including CC energy and amplitude equations, second-quantized strings in normal order, many-body terms of the similarity-transformed Hamiltonian, and blocks of the EOM-CC Hamiltonian matrix in terms of many-body terms.

For example, the input code to obtain IP-EOM-CCSD is

```
1 (* Input for IP-EOM-CCSD *)
2
3 ClusterOperator={{ "1h1p"}, {"2h2p"} };
4 EOMOperator={{ "1h0p"}, {"2h1p"} };
5 eomccgen [ClusterOperator, EOMOperator]
6
```

while the output consists of a list of many-body terms and the EOM-CC matrix in terms of these quantities:

```
1 (* Many-body terms *)
2
3  $\chi_1[h_3, h_1] = -F[h_3, h_1] - F[h_3, v_1] t_1[h_1, v_1] + t_1[o_1, v_1] \text{ERI}[h_3, o_1, v_1, h_1] +$ 
4  $t_1[h_1, v_2] t_1[o_1, v_1] \text{ERI}[h_3, o_1, v_1, v_2] + 1/2 t_2[o_1, h_1, v_1, v_2] \text{ERI}[h_3, o_1, v_1, v_2]$ 
5  $\chi_1[h_4, p_3] = F[h_4, p_3] - t_1[o_1, v_1] \text{ERI}[h_4, o_1, v_1, p_3]$ 
6  $\chi_2[h_4, h_3, h_1, p_3] = \text{ERI}[h_4, h_3, h_1, p_3] + t_1[h_1, v_1] \text{ERI}[h_4, h_3, v_1, p_3]$ 
7  $\chi_2[h_3, p_1, h_1, h_2] = t_1[o_1, p_1] \text{ERI}[h_3, o_1, h_1, h_2] - t_1[h_2, v_1] t_1[o_1, p_1] \text{ERI}[h_3, o_1, v_1, h_1] +$ 
8  $t_1[h_1, v_1] t_1[o_1, p_1] \text{ERI}[h_3, o_1, v_1, h_2] + t_1[h_1, v_1] t_1[h_2, v_2] t_1[o_1, p_1] \text{ERI}[h_3, o_1, v_1, v_2] -$ 
9  $\text{ERI}[h_3, p_1, h_1, h_2] + t_1[h_2, v_1] \text{ERI}[h_3, p_1, v_1, h_1] - t_1[h_1, v_1] \text{ERI}[h_3, p_1, v_1, h_2] -$ 
10  $t_1[h_1, v_1] t_1[h_2, v_2] \text{ERI}[h_3, p_1, v_1, v_2] - F[h_3, v_1] t_2[h_1, h_2, v_1, p_1] +$ 
11  $1/2 t_1[o_1, p_1] t_2[h_1, h_2, v_1, v_2] \text{ERI}[h_3, o_1, v_1, v_2] - 1/2 \text{ERI}[h_3, p_1, v_1, v_2] t_2[h_1, h_2, v_1, v_2] +$ 
12  $t_1[o_1, v_1] t_2[h_1, h_2, v_2, p_1] \text{ERI}[h_3, o_1, v_1, v_2] - t_2[o_1, h_1, v_1, p_1] \text{ERI}[h_3, o_1, v_1, h_2]$ 
13  $+ t_1[h_2, v_1] t_2[o_1, h_1, v_2, p_1] \text{ERI}[h_3, o_1, v_1, v_2] + t_2[o_1, h_2, v_1, p_1] \text{ERI}[h_3, o_1, v_1, h_1]$ 
14  $- t_1[h_1, v_1] t_2[o_1, h_2, v_2, p_1] \text{ERI}[h_3, o_1, v_1, v_2]$ 
15  $\chi_1[p_1, p_3] = F[p_1, p_3] - F[o_1, p_3] t_1[o_1, p_1] - t_1[o_2, p_1] \text{ERI}[o_1, o_2, v_1, p_3] + t_1[o_1, v_1] \text{ERI}[o_1, p_1, v_1, p_3]$ 
16  $- 1/2 \text{ERI}[o_1, o_2, v_1, p_3] t_2[o_1, o_2, v_1, p_1]$ 
17  $\chi_2[h_4, h_3, h_1, h_2] = -\text{ERI}[h_4, h_3, h_1, h_2] + t_1[h_2, v_1] \text{ERI}[h_4, h_3, v_1, h_1] - t_1[h_1, v_1] \text{ERI}[h_4, h_3, v_1, h_2] - t_1[h_1, v_1]$ 
18  $t_1[h_2, v_2] \text{ERI}[h_4, h_3, v_1, v_2] - 1/2 \text{ERI}[h_4, h_3, v_1, v_2] t_2[h_1, h_2, v_1, v_2]$ 
19  $\chi_2[h_4, p_1, h_2, p_3] = t_1[o_1, p_1] \text{ERI}[h_4, o_1, h_2, p_3] + t_1[h_2, v_1] t_1[o_1, p_1] \text{ERI}[h_4, o_1, v_1, p_3] - \text{ERI}[h_4, p_1, h_2, p_3]$ 
20  $- t_1[h_2, v_1] \text{ERI}[h_4, p_1, v_1, p_3] - \text{ERI}[h_4, o_1, v_1, p_3] t_2[o_1, h_2, v_1, p_1]$ 
21  $\chi_3[h_4, h_3, p_1, h_1, h_2, p_3] = \text{ERI}[h_4, h_3, v_1, p_3] t_2[h_1, h_2, v_1, p_1]$ 
22
23 (* EOM-CC blocks *)
24
25  $\{-\chi_1[h_3, h_1]\},$ 
26  $\{\chi_2[h_4, h_3, h_1, p_3] + \delta_{h_1, h_3} \chi_1[h_4, p_3] - \delta_{h_1, h_4} \chi_1[h_3, p_3]\},$ 
27  $\{-\chi_2[h_3, p_1, h_2, h_1]\},$ 
28  $\{-\chi_1[h_4, h_2] \delta_{h_1, h_3} \delta_{p_1, p_3} + \chi_1[h_3, h_2] \delta_{h_1, h_4} \delta_{p_1, p_3} + \chi_1[h_4, h_1] \delta_{h_2, h_3} \delta_{p_1, p_3} - \chi_1[h_3, h_1] \delta_{h_2, h_4} \delta_{p_1, p_3} - \chi_1[p_1, p_3] \delta_{h_1, h_4} \delta_{h_2, h_3}$ 
29  $+ \chi_1[p_1, p_3] \delta_{h_1, h_3} \delta_{h_2, h_4} - \chi_2[h_4, p_1, h_2, p_3] \delta_{h_1, h_3} + \chi_2[h_3, p_1, h_2, p_3] \delta_{h_1, h_4} + \chi_2[h_4, p_1, h_1, p_3] \delta_{h_2, h_3} - \chi_2[h_3, p_1, h_1, p_3]$ 
30  $\delta_{h_2, h_4} +$ 
31  $\chi_2[h_4, h_3, h_2, h_1] \delta_{p_1, p_3} + \chi_3[h_4, h_3, p_1, h_2, h_1, p_3]\}$ 
```

The input code to generate the CCD equations is

```
1 (* Input for CCD *)
2
3 ClusterOperator={{ "2h2p"} };
4 EOMOperator={{ "0h0p"} };
5 eomccgen [ClusterOperator, EOMOperator]
6
```

and it produces the following output:

```
1 (* CCD energy *)
```

```

2      {1/4 ERI[o1,o2,v1,v2] t2[o1,o2,v1,v2]}
3
4      (* T2 equations *)
5
6      {-ERI[p2,p1,h1,h2]
7      -F[p2,v1]t2[h1,h2,v1,p1]+F[p1,v1]t2[h1,h2,v1,p2]
8      -F[o1,h2]t2[o1,h1,p2,p1]+F[o1,h1]t2[o1,h2,p2,p1]
9      -1/2 ERI[p2,p1,v1,v2]t2[h1,h2,v1,v2] +ERI[o1,p2,v1,h2]t2[o1,h1,v1,p1]
10     -ERI[o1,p1,v1,h2]t2[o1,h1,v1,p2]-ERI[o1,p2,v1,h1]t2[o1,h2,v1,p1]
11     +ERI[o1,p1,v1,h1]t2[o1,h2,v1,p2]-1/2ERI[o1,o2,h1,h2]t2[o1,o2,p2,p1]
12     -1/4ERI[o1,o2,v1,v2]t2[h1,h2,v1,v2]t2[o1,o2,p2,p1]
13     -1/2ERI[[o1,o2,v1,v2]t2[h1,h2,v2,p2]t2[o1,o2,v1,p1]
14     +1/2ERI[o1,o2,v1,v2]t2[h1,h2,v2,p1]t2[o1,o2,v1,p2]
15     -1/2ERI[o1,o2,v1,v2]t2[o1,h2,v1,v2]t2[o2,h1,p2,p1]
16     +1/2ERI[o1,o2,v1,v2]t2[o1,h1,v1,v2]t2[o2,h2,p2,p1]
17     +ERI[o1,o2,v1,v2]t2[o1,h2,v1,p2]t2[o2,h1,v2,p1]
18     -ERI[o1,o2,v1,v2]t2[o1,h1,v1,p2]t2[o2,h2,v2,p1]}
19

```

In the following example, we compute the block {2,1} of the IP-EOM-CCSD matrix. The corresponding input is

```

1      (* Input for IP-EOM-CCSD *)
2
3      ClusterOperator={{ "1h0p"}, {"2h1p"} };
4      EOMOperator={{ "1h1p"}, {"2h2p"} };
5      EOMBlock={2,1};
6      eomccgen[ClusterOperator,EOMOperator,EOMBlock]
7

```

and the output is

```

1      (* Many body terms *)
2
3      χ1[h4,p3]==F[h4,p3]-ERI[h4,o1,v1,p3]t1[o1,v1]
4      χ2[h4,h3,h1,p3]==ERI[h4,h3,h1,p3]-ERI[h4,h3,v1,p3]t1[h1,v1]
5
6      (* EOM-CC blocks *)
7
8      {χ1[h4,p3]δh1,h3-χ1[h3,p3]δh1,h4+χ2[h4,h3,v1,p3]}

```

The MATHEMATICA code described in this paper is publicly available in a dedicated repository on GITHUB with the name **eomccgen**.⁷⁹ In particular, the repository contains examples of scripts and data files that can be easily adapted for particular research projects. Additionally, the repository includes another notebook, called **eomccnum**, where the equations obtained in **eomccgen** can be implemented and numerically tested for small atoms and molecules. **eomccnum** comes with predefined examples that can serve as a guide for implementing the equations obtained in **eomccgen**. An example of the water molecule in the STO-3G basis can be found in the repository.

IV. CONCLUSION

The present paper discusses the development of a code generator for equation-of-motion coupled-cluster (EOM-CC) methods, a class of many-body quantum chemistry methods known for their accuracy in predicting excitation energies, ionization potentials, and electron affinities in molecular systems. Because their implementation can be complex and time-consuming, we develop an easy-to-use code generator that automates the process of deriving and implementing the EOM-CC equations, reducing the potential for human error.

We begin by discussing the second-quantization formalism, a practical and modern mathematical language used to describe many-body quantum systems in terms of creation and annihilation operators. We then discuss the development of our code generator for EOM-CC methods. Our approach builds on previous work by using second-quantized strings to automate the derivation and implementation of EOM-CC equations. However, we introduce several new features that make the code generator more efficient and flexible. For example, we rely on a symbolic algebra software package, MATHEMATICA, to generate these equations that can be easily read by humans and machines. The paper provides a detailed description of each step of the algorithm used in the code generator. We also describe how the code generator can be used to calculate excitation energies, ionization potentials, and electron affinities, by providing several concrete examples.

Although many improvements are still needed to generate a production-level code, especially in the definition of adequate intermediates to ensure the right computational scaling of the methods, we hope that the present work illustrates nicely the capabilities of MATHEMATICA-based code generator in the context of CC theory.

ACKNOWLEDGMENTS

The authors thank Yann Damour and Antoine Marie for useful comments on the manuscript. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 863481).

- ¹P. A. M. Dirac and N. H. D. Bohr, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **114**, 243 (1927).
- ²I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*, Cambridge Molecular Science (Cambridge University Press, 2009).
- ³G. C. Wick, *Phys. Rev.* **80**, 268 (1950).
- ⁴D. A. Matthews, *SIAM J. Sci. Comput.* **40**, C1 (2018).
- ⁵A. Abdelfattah, M. Baboulin, V. Dobrev, J. Dongarra, C. Earl, J. Falcou, A. Haidar, I. Karlin, T. Kolev, I. Masliah, and S. Tomov, *Procedia Computer Science* **80**, 108 (2016), international Conference on Computational Science 2016, ICCS 2016, 6-8 June 2016, San Diego, California, USA.
- ⁶P. Springer, T. Su, and P. Bientinesi, in *Proceedings of the 4th ACM SIGPLAN International Workshop on Libraries, Languages, and Compilers for Array Programming*, ARRAY 2017 (Association for Computing Machinery, New York, NY, USA, 2017) pp. 56–62.
- ⁷J. A. Calvin, C. A. Lewis, and E. F. Valeev, in *Proceedings of the 5th Workshop on Irregular Applications: Architectures and Algorithms*, IA3 ’15 (Association for Computing Machinery, New York, NY, USA, 2015).
- ⁸E. Solomonik, D. Matthews, J. F. Hammond, Jeff R and Stanton, and J. Demmel, *Journal of Parallel and Distributed Computing* **74**, 3176 (2014).
- ⁹T. Herault, Y. Robert, G. Bosilca, R. J. Harrison, C. A. Lewis, E. F. Valeev, and J. J. Dongarra, “Distributed-memory multi-GPU block-sparse tensor contraction for electronic structure (revised version),” Research Report RR-9365 (Inria - Research Centre Grenoble – Rhône-Alpes, 2020).
- ¹⁰C. Psarras, L. Karlsson, J. Li, and P. Bientinesi, “The landscape of software for tensor computations,” (2021).
- ¹¹F. Coester, *Nuc. Phys.* **7**, 421 (1958).
- ¹²F. Coester and H. Kümmel, *Nuc. Phys.* **17**, 477 (1960).
- ¹³J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966).
- ¹⁴T. D. Crawford and H. F. Schaefer III, “An introduction to coupled cluster theory for computational chemists,” in *Reviews in Computational Chemistry* (John Wiley and Sons, Ltd, 2000) pp. 33–136.
- ¹⁵C. L. Janssen and H. F. Schaefer, *Theor. Chim. Acta* **79**, 1 (1991).
- ¹⁶X. Li and J. Paldus, *J. Chem. Phys.* **101**, 8812 (1994).
- ¹⁷M. Kállay and P. R. Surján, *J. Chem. Phys.* **115**, 2945 (2001).
- ¹⁸S. Das, D. Mukherjee, and M. Kállay, *J. Chem. Phys.* **132**, 074103 (2010).
- ¹⁹M. Kállay and J. Gauss, *J. Chem. Phys.* **121**, 9257 (2004).
- ²⁰M. Kállay and J. Gauss, *J. Chem. Phys.* **123**, 214105 (2005).
- ²¹S. Hirata, *J. Phys. Chem. A* **107**, 9887 (2003).
- ²²S. Hirata, *J. Chem. Phys.* **121**, 51 (2004).
- ²³A. A. Auer, G. Baumgartner, D. E. Bernholdt, A. Bibireata, V. Choppella, D. Cociorva, X. Gao, R. Harrison, S. Krishnamoorthy, S. Krishnan, C.-C. Lam, Q. Lu, M. Nooijen, R. Pitzer, J. Ramanujam, P. Sadayappan, and A. Sibiryakov, *Mol. Phys.* **104**, 211 (2006).
- ²⁴M. Kamiya and S. Hirata, *J. Chem. Phys.* **125**, 074111 (2006).
- ²⁵J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.* **98**, 7029 (1993).
- ²⁶R. J. Bartlett, *WIREs Comput. Mol. Sci.* **2**, 126 (2012).
- ²⁷A. I. Krylov, *Annu. Rev. Phys. Chem.* **59**, 433 (2008), pMID: 18173379.
- ²⁸M. Musiał, “Equation-of-motion coupled-cluster models,” in *Quantum Chemistry and Dynamics of Excited States* (John Wiley and Sons, Ltd, 2020) Chap. 4, pp. 77–108.
- ²⁹K. Emrich, *Nuc. Phys. A* **351**, 379 (1981).
- ³⁰H. Sekino and R. J. Bartlett, *Int. J. Quantum Chem.* **26**, 255 (1984).
- ³¹D. C. Comeau and R. J. Bartlett, *Chem. Phys. Lett.* **207**, 414 (1993).
- ³²J. F. Stanton and J. Gauss, *J. Chem. Phys.* **101**, 8938 (1994).
- ³³R. J. Bartlett, J. E. Del Bene, S. Perera, and R. Mattie, *J. of Mol.Struc.: THEOCHEM* **400**, 157 (1997), ab Initio Benchmark Studies.
- ³⁴J. F. Stanton and J. Gauss, *J. Chem. Phys.* **111**, 8785 (1999).
- ³⁵M. Kamiya and S. Hirata, *J. Chem. Phys.* **125**, 074111 (2006).
- ³⁶Y. J. Bomble, J. C. Saeh, J. F. Stanton, P. G. Szalay, M. Kállay, and J. Gauss, *J. Chem. Phys.* **122**, 154107 (2005).
- ³⁷M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **102**, 6735 (1995).
- ³⁸M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **102**, 3629 (1995).
- ³⁹M. Kamiya and S. Hirata, *J. Chem. Phys.* **126**, 134112 (2007).
- ⁴⁰M. Hanrath and A. Engels-Putzka, *J. Chem. Phys.* **133**, 064108 (2010).
- ⁴¹A. Engels-Putzka and M. Hanrath, *J. Chem. Phys.* **134**, 124106 (2011).
- ⁴²L. Kong, K. R. Shamasundar, O. Demel, and M. Nooijen, *J. Chem. Phys.* **130**, 114101 (2009).
- ⁴³S. Das, M. Kállay, and D. Mukherjee, *J. Chem. Phys.* **133**, 234110 (2010).
- ⁴⁴F. A. Evangelista and J. Gauss, *J. Chem. Phys.* **134**, 114102 (2011).
- ⁴⁵M. Hanauer and A. Köhn, *J. Chem. Phys.* **134**, 204111 (2011).
- ⁴⁶M. Hanauer and A. Köhn, *J. Chem. Phys.* **137**, 131103 (2012).
- ⁴⁷T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3358 (2008).
- ⁴⁸A. Köhn, G. W. Richings, and D. P. Tew, *J. Chem. Phys.* **129**, 201103 (2008).
- ⁴⁹M. K. MacLeod and T. Shiozaki, *J. Chem. Phys.* **142**, 051103 (2015).
- ⁵⁰N. C. Rubin and A. E. D. III, *Mol. Phys.* **119**, e1954709 (2021).
- ⁵¹F. A. Evangelista, *J. Chem. Phys.* **157**, 064111 (2022).

- ⁵²D. Mukherjee, *Chem. Phys. Lett.* **274**, 561 (1997).
- ⁵³W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **107**, 432 (1997).
- ⁵⁴E. Monino and P.-F. Loos, *J. Chem. Phys.* **156**, 231101 (2022).
- ⁵⁵T. C. Berkelbach, *J. Chem. Phys.* **149**, 041103 (2018).
- ⁵⁶G. E. Scuseria, T. M. Henderson, and D. C. Sorensen, *J. Chem. Phys.* **129**, 231101 (2008).
- ⁵⁷G. E. Scuseria, T. M. Henderson, and I. W. Bulik, *J. Chem. Phys.* **139**, 104113 (2013).
- ⁵⁸M. F. Lange and T. C. Berkelbach, *J. Chem. Theory Comput.* **14**, 4224 (2018), pMID: 30028614.
- ⁵⁹W. R. Inc., “*Mathematica, Version 13.2*,” Champaign, IL, 2022.
- ⁶⁰T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, LTD, Chichester, 2000).
- ⁶¹P. R. Surján, *Second quantized approach to quantum chemistry: an elementary introduction* (Springer Science & Business Media, 2012).
- ⁶²A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Courier Corporation, 2012).
- ⁶³N. N. Bogoliubov, D. V. Shirkov, and E. M. Henley, *Introduction to the Theory of Quantized Fields* (1960).
- ⁶⁴J. Paldus and J. Čížek (Academic Press, 1975) pp. 105–197.
- ⁶⁵E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, 1998).
- ⁶⁶J. Hubbard and R. E. Peierls, *Proc. R. Soc. London.* **240**, 539 (1957).
- ⁶⁷J. Hubbard and R. E. Peierls, *Proc. R. Soc. London.* **243**, 336 (1958).
- ⁶⁸F. Jensen, *Introduction to Computational Chemistry* (John Wiley and Sons, 2017).
- ⁶⁹K. Kumar, *J. Math. Phys.* **6**, 1928 (1965).
- ⁷⁰K. W. Sattelmeyer, H. F. Schaefer III, and J. F. Stanton, *Chem. Phys. Lett.* **378**, 42 (2003).
- ⁷¹O. Demel, K. R. Shamasundar, L. Kong, and M. Nooijen, *J. Phys. Chem. A* **112**, 11895 (2008).
- ⁷²M. Musiał, A. Perera, and R. J. Bartlett, *J. Chem. Phys.* **134**, 114108 (2011).
- ⁷³T. Kuś and A. I. Krylov, *J. Chem. Phys.* **135**, 084109 (2011).
- ⁷⁴J. Shen and P. Piecuch, *J. Chem. Phys.* **138**, 194102 (2013).
- ⁷⁵S. A. Kucharski and R. J. Bartlett, *Theoret. Chim. Acta* **80**, 387 (1991).
- ⁷⁶S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **97**, 4282 (1992).
- ⁷⁷M. Musiał, S. Kucharski, and R. Bartlett, *J. Chem. Phys.* **116**, 4382 (2002).
- ⁷⁸M. Musiał, S. Kucharski, and R. Bartlett, *Mol. Phys.* **100**, 1867 (2002).
- ⁷⁹R. Quintero-Monsebaiz and P.-F. Loos, “*eomccgen*,” (2023).