

# Generalized nonorthogonal matrix elements: Unifying Wick's theorem and the Slater–Condon rules

Cite as: J. Chem. Phys. **154**, 144109 (2021); <https://doi.org/10.1063/5.0045442>

Submitted: 26 January 2021 • Accepted: 16 March 2021 • Published Online: 13 April 2021

 Hugh G. A. Burton



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

### Electronic structure software

The Journal of Chemical Physics **153**, 070401 (2020); <https://doi.org/10.1063/5.0023185>

### Spin contamination in MP2 and CC2, a surprising issue

The Journal of Chemical Physics **154**, 131101 (2021); <https://doi.org/10.1063/5.0044362>

### From orbitals to observables and back

The Journal of Chemical Physics **153**, 080901 (2020); <https://doi.org/10.1063/5.0018597>



Chemical Physics Reviews

First Articles Now Online!

READ NOW >>>



# Generalized nonorthogonal matrix elements: Unifying Wick's theorem and the Slater–Condon rules

Cite as: J. Chem. Phys. 154, 144109 (2021); doi: 10.1063/5.0045442

Submitted: 26 January 2021 • Accepted: 16 March 2021 •

Published Online: 13 April 2021



Hugh G. A. Burton<sup>a)</sup>

## AFFILIATIONS

Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

<sup>a)</sup> Author to whom correspondence should be addressed: [hugh.burton@chem.ox.ac.uk](mailto:hugh.burton@chem.ox.ac.uk)

## ABSTRACT

Matrix elements between nonorthogonal Slater determinants represent an essential component of many emerging electronic structure methods. However, evaluating nonorthogonal matrix elements is conceptually and computationally harder than their orthogonal counterparts. While several different approaches have been developed, these are predominantly derived from the first-quantized generalized Slater–Condon rules and usually require biorthogonal occupied orbitals to be computed for each matrix element. For coupling terms between nonorthogonal excited configurations, a second-quantized approach such as the nonorthogonal Wick's theorem is more desirable, but this fails when the two reference determinants have a zero many-body overlap. In this contribution, we derive an entirely generalized extension to the nonorthogonal Wick's theorem that is applicable to all pairs of determinants with nonorthogonal orbitals. Our approach creates a universal methodology for evaluating any nonorthogonal matrix element and allows Wick's theorem and the generalized Slater–Condon rules to be unified for the first time. Furthermore, we present a simple well-defined protocol for deriving arbitrary coupling terms between nonorthogonal excited configurations. In the case of overlap and one-body operators, this protocol recovers efficient formulas with reduced scaling, promising significant computational acceleration for methods that rely on such terms.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0045442>

## I. INTRODUCTION

Matrix elements between nonorthogonal Slater determinants are increasingly common in emerging electronic structure methods. For example, capturing strong correlation using a linear combination of nonorthogonal Slater determinants is a relatively old idea<sup>1–4</sup> that has seen a renaissance in the past decade.<sup>5–14</sup> Similarly, nonorthogonal matrix elements arise in projected Hartree–Fock methods,<sup>15,16</sup> while the combination of geminal-based nonorthogonal functions is an area of ongoing research.<sup>17</sup> In each method, the nonorthogonality of different determinants can capture strong static correlation effects by breaking and restoring symmetries of the Hamiltonian,<sup>15</sup> or it can provide a quasi-diabatic representation of dominant electronic configurations.<sup>5,9</sup> Alternatively, multiple wave functions built from different orbitals arise in

orbital-optimized excited states identified through methods such as  $\Delta$ SCF,<sup>18–21</sup> excited-state mean-field theory,<sup>22,23</sup> or the complete active space self-consistent field (SCF).<sup>24,25</sup> In these cases, orbital optimization can significantly improve predictions of charge transfer excitations, but nonorthogonal matrix elements are required for inter-state coupling terms such as oscillator strengths.

A variety of different approaches have been developed for the efficient evaluation of nonorthogonal matrix elements,<sup>26–30</sup> which are predominantly derived from Löwdin's general formula.<sup>31</sup> The most popular framework in quantum chemistry is the generalized Slater–Condon rules,<sup>27,32</sup> where biorthogonal occupied orbitals are constructed<sup>33,34</sup> and a modified form of the Slater–Condon rules<sup>35</sup> is applied depending on the number of zero-overlap orbital pairs in the biorthogonal basis. This approach is applicable to any

pair of determinants but requires the diagonalization of the occupied orbital overlap matrix each time. In contrast, the development of many-body correlation methods using orthogonal determinants has greatly benefited from the second-quantized Wick's theorem.<sup>36</sup> While a nonorthogonal variant of Wick's theorem exists, it is limited to determinants that have a non-zero many-body overlap and is not applicable if there are zero-overlap orbital pairs in the biorthogonal basis.<sup>37,38</sup> This limitation arises because the Thouless theorem,<sup>39</sup> used to relate two nonorthogonal determinants via an exponential transformation, breaks down when the two determinants have a zero many-body overlap. As a result, the nonorthogonal Wick's theorem has seen only limited use in quantum chemistry.<sup>16,40,41</sup>

Computationally efficient nonorthogonal matrix elements become increasingly important in methods that use orthogonally excited configurations from nonorthogonal reference determinants. For example, including post-NOCI dynamic correlation in methods such as perturbative NOCI-MP2<sup>42–44</sup> and NOCI-PT2,<sup>14</sup> or nonorthogonal multireference CI,<sup>3,16,41,45</sup> requires overlap, one-body, or two-body coupling terms between excitations from nonorthogonal determinants. The number of nonorthogonal matrix elements therefore grows rapidly, and repeated biorthogonalization of the occupied orbitals becomes prohibitively expensive. In principle, the nonorthogonal Wick's theorem could allow these matrix elements to be evaluated using only biorthogonal reference orbitals, but until now, this requires the reference determinants to have a strictly non-zero overlap.

In this contribution, we derive an entirely generalized nonorthogonal form of Wick's theorem that applies to any pair of determinants with nonorthogonal orbitals, even if the overall determinants have a zero overlap. This new framework, which we call the “extended nonorthogonal Wick's theorem,” provides the most general approach for deriving matrix elements using second-quantization, allowing Wick's theorem and the generalized Slater–Condon rules to be unified for the first time.

One particular advantage of our approach is that it allows all matrix elements between excited configurations from a pair of nonorthogonal determinants to be derived using a single well-defined protocol. While some of the resulting expressions have previously been derived for various bespoke applications (see, e.g., Refs. 6, 42, and 45), these have often relied on the properties of matrix determinants to account for orbital excitations. Instead, we present a unifying theory that can recover all of these results and is automatically applicable in cases where the reference determinants have a zero overlap. Furthermore, we show how evaluating intermediates for a given pair of determinants can reduce the scaling of overlap and one-body coupling terms between excited configurations to  $\mathcal{O}(1)$ . These particular non-orthogonal matrix elements then become almost as straightforward as the orthogonal Slater–Condon rules or Wick's theorem, promising considerable acceleration for methods that rely on such terms.

To derive our generalized nonorthogonal matrix elements, we first define our notation in Sec. II. In Sec. III, we extend Thouless theorem<sup>39</sup> to the case where the two determinants have nonorthogonal orbitals and a zero many-body overlap. Section IV combines this extended Thouless transformation with Wick's theorem to create a generalized protocol for evaluating nonorthogonal matrix elements

using second-quantization. We then illustrate the application of this approach by re-deriving the generalized Slater–Condon rules for one- and two-body operators in Sec. V. Finally, in Sec. VI, we extend our framework to the matrix elements between excited configurations and show how  $\mathcal{O}(1)$  scaling can be achieved for overlap and one-body operators.

## II. NOTATION

We will consider matrix elements between the two determinants  $|^w\Phi\rangle$  and  $|^x\Phi\rangle$ . Each determinant is constructed from a bespoke set of molecular orbitals (MOs), represented in terms of the atomic spin-orbital basis functions  $|\chi_\mu\rangle$  as

$$|^w\phi_p\rangle = \sum_\mu |\chi_\mu\rangle {}^wC_p^\mu. \quad (1)$$

Here, we employ the nonorthogonal tensor notation of Head-Gordon *et al.*<sup>46</sup> to explicitly keep track of any required overlap matrices. Occupied MOs are indexed as  $i, j, k$ , etc., virtual MOs are indexed as  $a, b, c$ , etc., and any general MO is indexed as  $p, q, r$ , etc. We emphasize that the MOs are orthogonal *within* each Slater determinant but are nonorthogonal *between* the different determinants.

In second-quantization, the  $N$ -electron determinant  $|^w\Phi\rangle$  is defined as

$$|^w\Phi\rangle = \prod_{i=1}^N {}^wb_i^\dagger |-\rangle, \quad (2)$$

where  $|-\rangle$  is the physical vacuum and the molecular orbital creation operators  ${}^wb_i^\dagger$  satisfy the standard fermionic anticommutation rules.<sup>47</sup> Using the expansion [Eq. (1)], the MO creation and annihilation operators can be represented in terms of the covariant atomic spin-orbital creation  $a_\mu^\dagger$  and annihilation  $a_\mu$  operators as

$${}^wb_p^\dagger = \sum_\mu a_\mu^\dagger {}^wC_p^\mu \quad \text{and} \quad {}^wb_p = \sum_\mu ({}^wC^*)_{p,\mu}^\mu a_\mu. \quad (3)$$

The covariant atomic spin-orbital operators have only one non-zero anticommutator,<sup>47</sup>

$$[a_\nu^\dagger, a_\mu]_+ = g_{\mu\nu}, \quad (4)$$

where  $g_{\mu\nu} = \langle\chi_\mu|\chi_\nu\rangle$  defines the corresponding covariant metric tensor (overlap matrix).<sup>46</sup>

Throughout this paper, we will need to express the atomic spin-orbital creation and annihilation operators in terms of the MO creation and annihilation operators using the inverse of Eq. (3),

$$a_\mu^\dagger = \sum_{p\sigma} {}^wb_p^\dagger ({}^wC^*)_{p,\mu}^\sigma g_{\sigma\mu}, \quad (5a)$$

$$a_\mu = \sum_{p\sigma} g_{\mu\sigma} ({}^wC)_{p,\mu}^\sigma {}^wb_p. \quad (5b)$$

To avoid the introduction of overlap matrices throughout our expressions, we will often use the contravariant atomic spin-orbital operators  $(a^\mu)^\dagger$  and  $a^\mu$  defined as

$$(a^\nu)^\dagger = \sum_\mu a_\mu^\dagger g^{\mu\nu} \quad \text{and} \quad a^\nu = \sum_\mu g^{\nu\mu} a_\mu, \quad (6)$$

where  $g^{\mu\nu}$  is the contravariant metric tensor corresponding to the inverse covariant overlap matrix,<sup>46</sup> i.e.,

$$g^{\mu\nu} = (g^{-1})_{\mu\nu}. \quad (7)$$

If the AO basis is overcomplete, this contravariant metric tensor becomes the pseudo-inverse of the covariant overlap matrix. Note that the anticommutator of these contravariant atomic spin-orbital operators is

$$[(a^\nu)^\dagger, a^\mu]_+ = g^{\mu\nu}. \quad (8)$$

### III. EXTENDED THOULESS TRANSFORMATION

#### A. Conventional Thouless transformation

The conventional form of Thouless theorem allows two nonorthogonal determinants to be related by an exponential operator of single excitations as<sup>39</sup>

$$|^w\Phi\rangle = \exp(\mathcal{Z})|^x\Phi\rangle \langle^x\Phi|^w\Phi\rangle. \quad (9)$$

To derive the single excitation operator  $\mathcal{Z}$ , the occupied orbitals can be transformed to a biorthogonal basis using Löwdin pairing<sup>33,34</sup> such that

$$\sum_{\mu\nu} ({}^x\tilde{C}^*)_{i\cdot}^\mu g_{\mu\nu} {}^w\tilde{C}_{\cdot j}^{\nu\cdot} = {}^{xw}\tilde{S}_{ij}. \quad (10)$$

The virtual–occupied and virtual–virtual blocks of the biorthogonalized overlap matrix become

$$\sum_{\mu\nu} ({}^x\tilde{C}^*)_{a\cdot}^\mu g_{\mu\nu} {}^w\tilde{C}_{\cdot p}^{\nu\cdot} = {}^{xw}\tilde{S}_{ap}, \quad (11)$$

and the transformed molecular orbital creation and annihilation operators are given as

$${}^w\tilde{b}_p^\dagger = \sum_\mu a_\mu^\dagger {}^w\tilde{C}_{\cdot p}^{\mu\cdot} \quad \text{and} \quad {}^w\tilde{b}_p = \sum_\mu ({}^w\tilde{C}^*)_{p\cdot}^\mu a_\mu. \quad (12)$$

The single excitation operator in Eq. (9) is then defined as

$$\mathcal{Z} = \sum_{ia} {}^{xw}Z_{ai} {}^x\tilde{b}_a^\dagger {}^x\tilde{b}_i, \quad (13)$$

with the  ${}^{xw}Z_{ai}$  matrix elements given by

$${}^{xw}Z_{ai} = \sum_{\mu\nu} ({}^x\tilde{C}^*)_{a\cdot}^\mu g_{\mu\nu} ({}^w\tilde{C})_{\cdot i}^{\nu\cdot} \frac{1}{{}^{xw}\tilde{S}_{ii}}. \quad (14)$$

A brief derivation of this result can be found in [Appendix A](#).

Unfortunately, this exponential representation relies on the strict nonorthogonality of the two determinants  $\langle^x\Phi|^w\Phi\rangle \neq 0$ ; in other words, it is not applicable to a pair of determinants that are orthogonal but contain mutually nonorthogonal orbitals. Our first step is therefore a generalization of the Thouless transformation to the case where  $\langle^x\Phi|^w\Phi\rangle = 0$ .

#### B. Introducing zero-overlap orbitals

We begin in the biorthogonal basis identified through Löwdin pairing, with orbital coefficients satisfying Eq. (10). For a general pair of nonorthogonal orbitals, it is possible for orbital pairs to have a zero-overlap in the biorthogonal basis, where  ${}^{xw}\tilde{S}_i = 0$ . Taking the case with  $m$  zero overlaps between orbitals  $k_1, \dots, k_m$ , we construct “reduced” determinants by removing the electrons in these zero-overlap orbitals to give

$$|^x\Phi_{k_1\dots k_m}\rangle = {}^x\tilde{b}_{k_m}\dots{}^x\tilde{b}_{k_1}|^x\Phi\rangle, \quad (15a)$$

$$|^w\Phi_{k_1\dots k_m}\rangle = {}^w\tilde{b}_{k_m}\dots{}^w\tilde{b}_{k_1}|^w\Phi\rangle. \quad (15b)$$

These reduced determinants are strictly nonorthogonal with the non-zero reduced overlap defined as

$${}^{xw}\tilde{S} = \langle^x\Phi_{k_1\dots k_m}|^w\Phi_{k_1\dots k_m}\rangle = \prod_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} {}^{xw}\tilde{S}_i. \quad (16)$$

Therefore, the Thouless transformation can now be applied to these reduced determinants to give

$$|^w\Phi_{k_1\dots k_m}\rangle = \exp(\tilde{\mathcal{Z}})|^x\Phi_{k_1\dots k_m}\rangle {}^{xw}\tilde{S}. \quad (17)$$

Here, we have introduced the reduced single excitation operator  $\tilde{\mathcal{Z}}$  that only contains excitations from occupied orbitals with a non-zero overlap as

$$\tilde{\mathcal{Z}} = \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} \sum_a {}^{xw}Z_{ai} {}^x\tilde{b}_a^\dagger {}^x\tilde{b}_i. \quad (18)$$

The full  $N$ -electron determinants are then related through second-quantization as

$$|^w\Phi\rangle = {}^w\tilde{b}_{k_1}^\dagger \dots {}^w\tilde{b}_{k_m}^\dagger \exp(\tilde{\mathcal{Z}}) {}^x\tilde{b}_{k_m} \dots {}^x\tilde{b}_{k_1} |^x\Phi\rangle {}^{xw}\tilde{S}. \quad (19)$$

Equation (19) can be further simplified by exploiting the commutativity relation  $[\tilde{\mathcal{Z}}, {}^x\tilde{b}_k] = 0$  to shift the  $\exp(\tilde{\mathcal{Z}})$  operator to the far right-hand side, giving

$$|^w\Phi\rangle = {}^w\tilde{b}_{k_1}^\dagger \dots {}^w\tilde{b}_{k_m}^\dagger {}^x\tilde{b}_{k_m} \dots {}^x\tilde{b}_{k_1} \exp(\tilde{\mathcal{Z}}) |^x\Phi\rangle {}^{xw}\tilde{S}. \quad (20)$$

We can then introduce single-electron excitation operators  $\hat{z}_k$  for the zero-overlap orbitals as

$$\begin{aligned} \hat{z}_k &= {}^w\tilde{b}_k^\dagger {}^x\tilde{b}_k = \sum_p \left( \sum_{\mu\nu} ({}^x\tilde{C}^*)_{p\cdot}^\mu g_{\mu\nu} ({}^w\tilde{C})_{\cdot k}^{\nu\cdot} \right) {}^x\tilde{b}_p^\dagger {}^x\tilde{b}_k \\ &= \sum_a {}^{xw}\tilde{S}_{ak} {}^x\tilde{b}_a^\dagger {}^x\tilde{b}_k, \end{aligned} \quad (21)$$

where we have exploited the biorthogonality and zero-overlap of the occupied orbitals such that  ${}^{xw}\tilde{S}_{ik} = 0$  for all  $i$ . The commutativity of these single excitation operators with the  ${}^x\tilde{b}_k$  annihilation operators leads to the simplified relationship

$$|{}^w\Phi\rangle = \prod_{\{k|{}^{xw}\tilde{S}_k=0\}} \hat{z}_k \exp(\tilde{Z})|{}^x\Phi\rangle^{xw} \tilde{S}. \quad (22)$$

Introducing the relationship  $\hat{z}_k = \exp(\hat{z}_k) - 1$  allows Eq. (22) to be expanded as

$$|{}^w\Phi\rangle = \prod_{\{k|{}^{xw}\tilde{S}_k=0\}} (\exp(\hat{z}_k) - 1) \exp(\tilde{Z})|{}^x\Phi\rangle^{xw} \tilde{S}. \quad (23)$$

Expanding the product of  $(\exp(\hat{z}_k) - 1)$  terms then leads to a sum of exponential transformations where every combination of the zero-overlap single excitation operators  $\hat{z}_k$  is either included or excluded with an appropriate phase factor, giving

$$|{}^w\Phi\rangle = \sum_{n=0}^m (-1)^{(m-n)} \sum_{C_n} \exp(\tilde{Z}^{C_n})|{}^x\Phi\rangle^{xw} \tilde{S}. \quad (24)$$

Here, we have introduced the compound index  $C_n$  to denote a particular combination of  $n$  out of  $m$  zero-overlap orbitals, while the superscript notation  $\tilde{Z}^{C_n}$  indicates which particular zero-overlap excitations are included in the corresponding operator, i.e.,

$$\tilde{Z}^{C_n} = \tilde{Z} + \sum_{k \in C_n} \hat{z}_k. \quad (25)$$

We refer to this transformation in Eq. (24) as the “extended Thouless transformation.” To explicitly illustrate its application, the case of two zero-overlaps in orbital pairs  $k_1$  and  $k_2$  leads to

$$|{}^w\Phi\rangle = {}^{xw}\tilde{S} \left( \exp \tilde{Z}^{(k_1, k_2)} - \exp \tilde{Z}^{(k_1)} - \exp \tilde{Z}^{(k_2)} + \exp \tilde{Z} \right) |{}^x\Phi\rangle, \quad (26)$$

where  $\tilde{Z}^{(k_1, k_2)} = \tilde{Z} + \hat{z}_{k_1} + \hat{z}_{k_2}$  and  $\tilde{Z}^{(k_1)} = \tilde{Z} + \hat{z}_{k_1}$ .

#### IV. EXTENDED NONORTHOGONAL WICK'S THEOREM

##### A. Conventional Wick's theorem

Efficiently deriving matrix elements using the conventional Wick's theorem requires the introduction of contractions, defined for two creation or annihilation operators  ${}^x b_p$  and  ${}^x b_q$  as

$$\overline{{}^x b_p {}^x b_q} = {}^x b_p {}^x b_q - \{ {}^x b_p {}^x b_q \}, \quad (27)$$

where  $\{ {}^x b_p {}^x b_q \}$  represents a normal-ordered operator string with respect to the reference Fermi vacuum  $\langle {}^x\Phi | \dots | {}^x\Phi \rangle$ .<sup>36</sup> The only non-zero contractions between creation and annihilation operators with respect to this symmetric Fermi vacuum are

$$\overline{{}^x b_a {}^x b_b^\dagger} = \delta_{ab} \quad \text{and} \quad \overline{{}^x b_i^\dagger {}^x b_j} = \delta_{ij}. \quad (28)$$

Through Wick's theorem, the Fermi vacuum expectation of an operator product is given by the sum over all fully contracted products of operators, e.g.,

$$\langle {}^x\Phi | {}^a_i {}^x\Phi | {}^b_j \rangle = \langle {}^x\Phi | \overline{{}^x b_i^\dagger {}^x b_a} \overline{{}^x b_b^\dagger {}^x b_j} | {}^x\Phi \rangle = \delta_{ij} \delta_{ab}. \quad (29)$$

##### B. Zero-overlap transformed operators

Using the extended Thouless transformation, we can now extend the nonorthogonal Wick's theorem<sup>37,38,48</sup> to derive matrix elements between *any* pair of determinants with mutually nonorthogonal orbitals. In what follows, we will consider the contravariant atomic spin-orbital operators (see Sec. II) to avoid large numbers of overlap matrices in our expressions. The matrix elements for general operators expressed in the atomic spin-orbital basis requires the evaluation of terms containing a string of creation and annihilation operators, such as  $\langle {}^x\Phi | (a^\mu)^\dagger (a^\nu)^\dagger \dots a^\sigma a^\tau | {}^w\Phi \rangle$ . Applying the extended Thouless transformation leads to the linear combination

$$\begin{aligned} & \langle {}^x\Phi | (a^\mu)^\dagger (a^\nu)^\dagger \dots a^\sigma a^\tau | {}^w\Phi \rangle \\ &= {}^{xw}\tilde{S} \sum_{n=0}^m (-1)^{(m-n)} \sum_{C_n} \langle {}^x\Phi | (a^\mu)^\dagger (a^\nu)^\dagger \dots a^\sigma a^\tau \exp(\tilde{Z}^{C_n}) | {}^x\Phi \rangle. \end{aligned} \quad (30)$$

To evaluate each constituent matrix element for the combinations  $C_n$ , we follow the approach described in Refs. 38 and 40 and introduce a similarity-transformed set of spin-orbital creation and annihilation operators as

$$(d[C_n]^\mu)^\dagger = \exp(-\tilde{Z}^{C_n}) (a^\mu)^\dagger \exp(\tilde{Z}^{C_n}), \quad (31a)$$

$$d[C_n]^\mu = \exp(-\tilde{Z}^{C_n}) a^\mu \exp(\tilde{Z}^{C_n}). \quad (31b)$$

These operators clearly depend on the particular combination  $C_n$  of included zero-overlap single excitation operators. Expanding the similarity transformation as

$$\exp(-\tilde{Z}^{C_n}) a^\mu \exp(\tilde{Z}^{C_n}) = a^\mu - [\tilde{Z}^{C_n}, a^\mu], \quad (32)$$

and similarly for  $(a^\mu)^\dagger$ , leads to the explicit forms

$$\begin{aligned} (d[C_n]^\mu)^\dagger &= \sum_i {}^x \tilde{b}_i^\dagger ({}^x \tilde{C}^*)_{i.}^\mu + \sum_a {}^x \tilde{b}_a^\dagger \left[ ({}^x \tilde{C}^*)_{a.}^\mu \right. \\ &\quad \left. - \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} {}^{xw} Z_{ai} ({}^x \tilde{C}^*)_{i.}^\mu - \sum_{k \in C_n} {}^{xw} \tilde{S}_{ak} ({}^x \tilde{C}^*)_{k.}^\mu \right], \end{aligned} \quad (33a)$$

$$\begin{aligned} d[C_n]^\mu &= \sum_i ({}^x \tilde{C})_{.i}^{\mu \cdot x} \tilde{b}_i + \sum_a \left[ ({}^x \tilde{C})_{.a}^{\mu \cdot x} {}^{xw} Z_{ai} {}^x \tilde{b}_i \right. \\ &\quad \left. + \sum_{k \in C_n} ({}^x \tilde{C})_{.a}^{\mu \cdot x} \tilde{S}_{ak} {}^x \tilde{b}_k + ({}^x \tilde{C})_{.a}^{\mu \cdot x} \tilde{b}_a \right]. \end{aligned} \quad (33b)$$

An explicit derivation of these relationships can be found in Appendix B. Exploiting the relationship

$$\langle {}^x\Phi | \exp(-\tilde{Z}^{\mathcal{C}_n}) = \langle {}^x\Phi | \quad (34)$$

and the resolution of the identity

$$\exp(\tilde{Z}^{\mathcal{C}_n}) \exp(-\tilde{Z}^{\mathcal{C}_n}) = \mathcal{I} \quad (35)$$

then allows the constituent matrix elements within Eq. (30) to be expressed as

$$\begin{aligned} & \langle {}^x\Phi | (a^\mu)^\dagger (a^\nu)^\dagger \cdots a^\sigma a^\tau \exp(\tilde{Z}^{\mathcal{C}_n}) | {}^x\Phi \rangle \\ &= \langle {}^x\Phi | (d[\mathcal{C}_n]^\mu)^\dagger (d[\mathcal{C}_n]^\nu)^\dagger \cdots d[\mathcal{C}_n]^\sigma d[\mathcal{C}_n]^\tau | {}^x\Phi \rangle. \end{aligned} \quad (36)$$

### C. The fundamental contraction

The extended Thouless transformation essentially converts the nonorthogonal matrix element with an asymmetric Fermi vacuum  $\langle {}^x\Phi | \cdots | {}^w\Phi \rangle$  to a transformed matrix element with respect to symmetric Fermi vacuum  $\langle {}^x\Phi | \cdots | {}^x\Phi \rangle$ . Since the transformed operators  $d[\mathcal{C}_n]^\mu$  and  $(d[\mathcal{C}_n]^\mu)^\dagger$  are expressed purely in terms of the  ${}^x\tilde{b}_p$  creation and  ${}^x\tilde{b}_p^\dagger$  annihilation operators, with respect to the  $\langle {}^x\Phi | \cdots | {}^x\Phi \rangle$  vacuum, their non-zero contractions with respect to  $\langle {}^x\Phi | \cdots | {}^x\Phi \rangle$  can be derived by combining Eqs. (28) and (33) to give

$$\overline{(d[\mathcal{C}_n]^\mu)^\dagger (d[\mathcal{C}_n]^\nu)} = {}^{xw}M^{\nu\mu} - \sum_{k \notin \mathcal{C}_n} {}^{xw}P_k^{\nu\mu}, \quad (37a)$$

$$\overline{(d[\mathcal{C}_n]^\mu) (d[\mathcal{C}_n]^\nu)^\dagger} = g^{\mu\nu} - {}^{xw}M^{\mu\nu} + \sum_{k \notin \mathcal{C}_n} {}^{xw}P_k^{\mu\nu}, \quad (37b)$$

where we have introduced the general notation

$${}^{xw}P_k^{\nu\mu} = ({}^w\tilde{C})_{\cdot k}^\nu ({}^x\tilde{C}^*)_{\cdot k}^\mu, \quad (38a)$$

$${}^{xw}P^{\nu\mu} = \sum_{\{k | {}^{xw}\tilde{S}_k = 0\}} {}^{xw}P_k^{\nu\mu}, \quad (38b)$$

$${}^{xw}W^{\nu\mu} = \sum_{\{i | {}^{xw}\tilde{S}_i \neq 0\}} ({}^w\tilde{C})_{\cdot i}^\nu \frac{1}{{}^{xw}\tilde{S}_i} ({}^x\tilde{C}^*)_{\cdot i}^\mu, \quad (38c)$$

$${}^{xw}M^{\nu\mu} = {}^{xx}P^{\nu\mu} + {}^{xw}P^{\nu\mu} + {}^{xw}W^{\nu\mu}. \quad (38d)$$

These expressions closely resemble the co-density matrices in Ref. 5, and their derivation can be found in Appendix C.

The overall matrix element  $\langle {}^x\Phi | (a^\mu)^\dagger (a^\nu)^\dagger \cdots a^\sigma a^\tau | {}^w\Phi \rangle$  requires the derivation of contractions between the  $a^\mu$  and  $(a^\mu)^\dagger$  atomic spin-orbital operators rather than the transformed  $d[\mathcal{C}_n]^\mu$  and  $(d[\mathcal{C}_n]^\mu)^\dagger$  operators. To show how a general string of operators can be evaluated using the contractions defined in Eqs. (37a) and (37b), we first demonstrate the derivation of the one-body co-density matrix element

$${}^{xw}\Gamma_1^{\nu\mu} = \langle {}^x\Phi | (a^\mu)^\dagger a^\nu | {}^w\Phi \rangle. \quad (39)$$

Assuming that some form of Wick's theorem can be derived, we expect this matrix element to be represented by the single contraction

$${}^{xw}\Gamma_1^{\nu\mu} = \langle {}^x\Phi | \overline{(a^\mu)^\dagger a^\nu} | {}^w\Phi \rangle. \quad (40)$$

Taking the most general case with  $m$  zero-overlap orbitals, the extended Thouless transformation leads to

$$\begin{aligned} \langle {}^x\Phi | \overline{(a^\mu)^\dagger a^\nu} | {}^w\Phi \rangle &= {}^{xw}\tilde{S} \sum_{n=0}^m (-1)^{(m-n)} \sum_{\mathcal{C}_n} \langle {}^x\Phi | \overline{(d[\mathcal{C}_n]^\mu)^\dagger d[\mathcal{C}_n]^\nu} | {}^x\Phi \rangle \\ &= {}^{xw}\tilde{S} \sum_{n=0}^m (-1)^{(m-n)} \sum_{\mathcal{C}_n} \left( {}^{xw}M^{\nu\mu} - \sum_{k \notin \mathcal{C}_n} {}^{xw}P_k^{\nu\mu} \right). \end{aligned} \quad (41)$$

Reversing the order of summation over  $k$  and  $\mathcal{C}_n$  yields

$$\langle {}^x\Phi | \overline{(a^\mu)^\dagger a^\nu} | {}^w\Phi \rangle = {}^{xw}\tilde{S} \sum_{n=0}^m (-1)^{(m-n)} \left( {}^{xw}M^{\nu\mu} \left[ \sum_{\mathcal{C}_n} 1 \right] - \sum_k {}^{xw}P_k^{\nu\mu} \left[ \sum_{\mathcal{C}_n \not\ni k} 1 \right] \right). \quad (42)$$

The first term in square brackets  $\left[ \sum_{\mathcal{C}_n} 1 \right]$  can be recognized as the total number of ways to pick  $n$  orbitals from the  $m$  zero-overlap orbitals, given by  $\binom{m}{n}$ . Similarly, the second term in square brackets  $\left[ \sum_{\mathcal{C}_n \not\ni k} 1 \right]$  is the total number of ways to pick  $n$  orbitals from the  $m-1$  zero-overlap orbitals that remain when orbital  $k$  is excluded, given by  $\binom{m-1}{n}$ . These combinatorial expansions allow Eq. (42) to be expressed as

$$\langle {}^x\Phi | \overline{(a^\mu)^\dagger a^\nu} | {}^w\Phi \rangle = {}^{xw}\tilde{S} \left[ \sum_{n=0}^m (-1)^{(m-n)} \binom{m}{n} {}^{xw}M^{\nu\mu} + \sum_{n=0}^{m-1} (-1)^{(m-1-n)} \binom{m-1}{n} {}^{xw}P^{\nu\mu} \right]. \quad (43)$$



Here, we note that there are no ways to exclude an orbital  $k$  when all zero-orbital overlaps are included in the complete combination  $C_m$ . Exploiting the binomial expansion

$$\sum_{n=0}^y (-1)^{(y-n)} \binom{y}{n} = \delta_{0,y} \quad (44)$$

then leads to the closed-form expression

$$\langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger | {}^w\Phi \rangle = {}^{xw}\tilde{S} [\delta_{0,m} {}^{xw}M^{\nu\mu} + \delta_{0,m-1} {}^{xw}P^{\nu\mu}]. \quad (45)$$

The reduced overlap  ${}^{xw}\tilde{S}$  will be a prefactor for every matrix element between these nonorthogonal determinants. The remaining terms can then be used to define “fundamental contractions” for second-quantization operators with respect to the asymmetric Fermi vacuum  $\langle {}^x\Phi | \dots | {}^w\Phi \rangle$ . The form of these contractions depends on the number of zero-overlap orbitals  $m$ , and we define the first fundamental contraction as

$$(\overline{a^\mu})^\dagger \overline{a^\nu} = \begin{cases} {}^{xw}M^{\nu\mu}, & m = 0 \\ {}^{xw}P^{\nu\mu}, & m = 1 \\ 0, & m > 1. \end{cases} \quad (46)$$

Similarly, the second fundamental contraction can be identified as

$$\begin{aligned} \langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau} | {}^w\Phi \rangle &= {}^{xw}\tilde{S} \sum_{n=0}^m (-1)^{(m-n)} \sum_{c_n} \langle {}^x\Phi | (\overline{d[c_n]^\mu})^\dagger (\overline{d[c_n]^\nu})^\dagger \overline{d[c_n]^\sigma} \overline{d[c_n]^\tau} | {}^x\Phi \rangle. \\ &= {}^{xw}\tilde{S} \sum_{n=0}^m (-1)^{(m-n)} \sum_{c_n} \left( {}^{xw}M^{\tau\mu} - \sum_{k_1 \notin c_n} {}^{xw}P_{k_1}^{\tau\mu} \right) \left( {}^{xw}M^{\sigma\nu} - \sum_{k_2 \notin c_n} {}^{xw}P_{k_2}^{\sigma\nu} \right). \end{aligned} \quad (50)$$

Once again, the reduced overlap  ${}^{xw}\tilde{S}$  appears as an overall prefactor. The order of summation over the  $k_1, k_2$  indices and  $C_n$  can then be swapped to give

$$\begin{aligned} \langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau} | {}^w\Phi \rangle &= {}^{xw}\tilde{S} \sum_{n=0}^m (-1)^{(m-n)} \left( {}^{xw}M^{\tau\mu} {}^{xw}M^{\sigma\nu} \left[ \sum_{c_n} 1 \right] \right. \\ &\quad \left. - \sum_k ({}^{xw}M^{\tau\mu} {}^{xw}P_k^{\sigma\nu} + {}^{xw}P_k^{\tau\mu} {}^{xw}M^{\sigma\nu}) \left[ \sum_{c_n \not\ni k} 1 \right] \right. \\ &\quad \left. + \sum_{k_1, k_2} {}^{xw}P_{k_1}^{\tau\mu} {}^{xw}P_{k_2}^{\sigma\nu} \left[ \sum_{c_n \not\ni k_1, k_2} 1 \right] \right). \end{aligned} \quad (51)$$

The third term in square brackets  $\left[ \sum_{c_n \not\ni k_1, k_2} 1 \right]$  is simply the number of ways to pick  $n$  orbitals from the  $m-2$  zero-overlap orbitals that remain when orbitals  $k_1$  and  $k_2$  are removed, given by  $\binom{m-2}{n}$ . Applying the binomial expansion Eq. (44) in an analogous way to the single contraction leads to the closed form

$$\langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau} | {}^w\Phi \rangle = {}^{xw}\tilde{S} [\delta_{0,m} ({}^{xw}M^{\tau\mu} {}^{xw}M^{\sigma\nu}) + \delta_{0,m-1} ({}^{xw}M^{\tau\mu} {}^{xw}P^{\sigma\nu} + {}^{xw}P^{\tau\mu} {}^{xw}M^{\sigma\nu}) + \delta_{0,m-2} ({}^{xw}P^{\tau\mu} {}^{xw}P^{\sigma\nu})]. \quad (52)$$

A similar expression can be derived for the second contraction as

$$\langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau} | {}^w\Phi \rangle = {}^{xw}\tilde{S} [\delta_{0,m} ({}^{xw}M^{\sigma\mu} {}^{xw}M^{\tau\nu}) + \delta_{0,m-1} ({}^{xw}M^{\sigma\mu} {}^{xw}P^{\tau\nu} + {}^{xw}P^{\sigma\mu} {}^{xw}M^{\tau\nu}) + \delta_{0,m-2} ({}^{xw}P^{\sigma\mu} {}^{xw}P^{\tau\nu})], \quad (53)$$

$$\overline{a^\mu} (\overline{a^\nu})^\dagger = \begin{cases} g^{\mu\nu} - {}^{xw}M^{\mu\nu}, & m = 0 \\ -{}^{xw}P^{\mu\nu}, & m = 1 \\ 0, & m > 1. \end{cases} \quad (47)$$

Crucially, we emphasize that these fundamental contractions are defined with respect to the asymmetric Fermi vacuum  $\langle {}^x\Phi | \dots | {}^w\Phi \rangle$ .

## D. Combining several contractions

Next, we show how the fundamental contractions can be combined to derive matrix elements for longer products of creation and annihilation operators. As an example, consider the two-body reduced co-density matrix element, defined as

$${}^{xw}\Gamma_2^{\tau\mu\sigma\nu} = \langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau} | {}^w\Phi \rangle. \quad (48)$$

Applying Wick's theorem, this matrix element should be given by the sum of the two contractions,

$$\langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau} | {}^w\Phi \rangle + \langle {}^x\Phi | (\overline{a^\mu})^\dagger (\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau} | {}^w\Phi \rangle. \quad (49)$$

Note that the second term in this expression carries a phase of  $-1$  from the intersection of the contraction lines, representing the fundamental parity of fermionic operators.<sup>36</sup> Taking the first contraction as an example, we apply the extended Thouless transformation and the transformed contractions defined in Eqs. (37a) and (37b) to give

which, analogously with the orthogonal case, will carry a  $-1$  phase factor from the intersection of the contraction lines. Combining Eqs. (52) and (53), with their associated phase factors, yields the full expression for the two-body reduced co-density matrix elements with  $m$  zero-overlap orbitals as

$$\langle {}^x\Phi|(a^\mu)^\dagger(a^\nu)^\dagger a^\sigma a^\tau|{}^w\Phi\rangle = \begin{cases} {}^{xw}\tilde{S}({}^{xw}M^{\tau\mu}{}^{xw}M^{\sigma\nu} - {}^{xw}M^{\sigma\mu}{}^{wx}M^{\tau\nu}), & m = 0 \\ {}^{xw}\tilde{S}({}^{xw}P^{\tau\mu}{}^{xw}M^{\sigma\nu} + {}^{xw}M^{\tau\mu}{}^{xw}P^{\sigma\nu} - {}^{xw}P^{\sigma\mu}{}^{xw}M^{\tau\nu} - {}^{xw}M^{\sigma\mu}{}^{xw}P^{\tau\nu}), & m = 1 \\ {}^{xw}\tilde{S}({}^{xw}P^{\tau\mu}{}^{xw}P^{\sigma\nu} - {}^{xw}P^{\sigma\mu}{}^{xw}P^{\tau\nu}), & m = 2 \\ 0, & m \geq 3. \end{cases} \quad (54)$$

## E. General rules for constructing matrix elements

To simplify the derivation of even longer operator strings, we note that the two-body matrix elements in Eq. (52) can be factorized into the product of two fundamental contractions with individual  $m_1$  and  $m_2$  values under the constraint  $m_1 + m_2 = m$ , giving

$$\langle {}^x\Phi|(\overline{a^\mu})^\dagger(\overline{a^\nu})^\dagger \overline{a^\sigma} \overline{a^\tau}|{}^w\Phi\rangle = {}^{xw}\tilde{S} \sum_{\substack{m_1, m_2 \\ m_1 + m_2 = m}} (\delta_{0, m_1} {}^{xw}M^{\tau\mu} + \delta_{0, m_1 - 1} {}^{xw}P^{\tau\mu})(\delta_{0, m_2} {}^{xw}M^{\sigma\nu} + \delta_{0, m_2 - 1} {}^{xw}P^{\sigma\nu}). \quad (55)$$

This factorization can be extended for a general product of contractions with each  $m_i$  restricted to the values 0 or 1 for the overall product to be non-zero. We can therefore define an intuitive approach for extending Wick's theorem to generalized nonorthogonal matrix elements as follows:

1. Construct all fully contracted combinations of the operator product with the associated phase factors.
2. For each term, sum every possible way to distribute  $m$  zeros among the contractions such that  $\sum_i m_i = m$ .
3. For every set of  $\{m_i\}$  in each term, construct the relevant contribution as a product of fundamental contractions depending on whether each contraction has  $m_i = 0$  or 1.
4. Multiply the final combined expression with the reduced overlap  ${}^{xw}\tilde{S}$ .

These rules and contractions therefore allow any matrix element to be evaluated with respect to the asymmetric Fermi vacuum  $\langle {}^x\Phi|\dots|{}^w\Phi\rangle$  for nonorthogonal orbitals, regardless of whether the many-body determinants are orthogonal or not. As a result, our formulation is the most flexible form of Wick's theorem and reduces to the previous nonorthogonal<sup>37,38,48</sup> or orthogonal<sup>36</sup> variants under suitable restrictions on the MO coefficients. We refer to this approach as the "extended nonorthogonal Wick's theorem." In Secs. V and VI, we will show how these steps can be applied to recover the generalized Slater–Condon rules<sup>32</sup> and to derive matrix elements between excited configurations with respect to the nonorthogonal reference determinants.

## V. GENERALIZED SLATER–CONDON RULES

The generalized Slater–Condon rules provide a first-quantized approach for evaluating matrix elements between nonorthogonal determinants. A detailed description of these rules, their derivation, and their application can be found in Ref. 32. However, to the best of our knowledge, the generalized Slater–Condon rules have never

previously been derived though a fully second-quantized framework. In this section, we will show how these rules can be recovered using the extended nonorthogonal Wick's theorem.

### A. One-body operators

Consider first the one-body operator

$$\hat{f} = \sum_{\mu\nu} f_{\mu\nu} (a^\mu)^\dagger a^\nu, \quad (56)$$

with the corresponding matrix element

$$\langle {}^x\Phi|\hat{f}|{}^w\Phi\rangle = \sum_{\mu\nu} f_{\mu\nu} \langle {}^x\Phi|(a^\mu)^\dagger a^\nu|{}^w\Phi\rangle. \quad (57)$$

Applying the extended nonorthogonal Wick's theorem yields only one non-zero contraction to give

$$\langle {}^x\Phi|\hat{f}|{}^w\Phi\rangle = \sum_{\mu\nu} f_{\mu\nu} \langle {}^x\Phi|(\overline{a^\mu})^\dagger \overline{a^\nu}|{}^w\Phi\rangle. \quad (58)$$

Substituting the fundamental contraction [Eq. (46)] and considering the possible values of  $m$  immediately yield the one-body generalized Slater–Condon rules<sup>32</sup> in the form presented in Ref. 5 as

$$\langle {}^x\Phi|\hat{f}|{}^w\Phi\rangle = \begin{cases} {}^{xw}\tilde{S} \sum_{\mu\nu} f_{\mu\nu} ({}^{xw}W)^{\nu\mu} & \text{no zeros} \\ {}^{xw}\tilde{S} \sum_{\mu\nu} f_{\mu\nu} ({}^{xw}P_k)^{\nu\mu} & \text{one zero } (k) \\ 0 & \text{otherwise.} \end{cases} \quad (59)$$

Here, we note that  ${}^{xw}M^{\nu\mu} = {}^{xw}W^{\nu\mu}$  when there are no zero-overlap orbitals, while  ${}^{xw}P^{\nu\mu} = {}^{xw}P_k^{\nu\mu}$  for one-zero overlap in orbital  $k$ . The matrices  ${}^{xw}M$  and  ${}^{xw}P_k$  correspond, respectively, to the weighted and unweighted co-density matrices discussed in Ref. 5.

### B. Two-body operators

Next, consider a general two-body operator  $\hat{v}$  defined in second-quantization as



$$\hat{v} = \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} (a^\mu)^\dagger (a^\nu)^\dagger a^\sigma a^\tau, \quad (60)$$

with the two-electron integrals in the spin-orbital basis defined as

$$v_{\mu\nu\sigma\tau} = \langle \chi_\mu \chi_\nu | \hat{v} | \chi_\sigma \chi_\tau \rangle. \quad (61)$$

The corresponding matrix element is given by

$$\langle {}^x\Phi | \hat{v} | {}^w\Phi \rangle = \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} \langle {}^x\Phi | (a^\mu)^\dagger (a^\nu)^\dagger a^\sigma a^\tau | {}^w\Phi \rangle. \quad (62)$$

Recognizing the  $\langle {}^x\Phi | (a^\mu)^\dagger (a^\nu)^\dagger a^\sigma a^\tau | {}^w\Phi \rangle$  term as the two-body reduced co-density matrix derived in Eq. (54), we immediately recover

$$\langle {}^x\Phi | \hat{v} | {}^w\Phi \rangle = \begin{cases} {}^{xw}\tilde{S} \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} ({}^{xw}M^{\tau\mu} {}^{xw}M^{\sigma\nu} - {}^{xw}M^{\sigma\mu} {}^{xw}M^{\tau\nu}), & m = 0 \\ {}^{xw}\tilde{S} \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} ({}^{xw}P^{\tau\mu} {}^{xw}M^{\sigma\nu} + {}^{xw}M^{\tau\mu} {}^{xw}P^{\sigma\nu} - {}^{xw}P^{\sigma\mu} {}^{xw}M^{\tau\nu} - {}^{xw}M^{\sigma\mu} {}^{xw}P^{\tau\nu}), & m = 1 \\ {}^{xw}\tilde{S} \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} ({}^{xw}P^{\tau\mu} {}^{xw}P^{\sigma\nu} - {}^{xw}P^{\sigma\mu} {}^{xw}P^{\tau\nu}), & m = 2 \\ 0, & m \geq 3. \end{cases} \quad (63)$$

Here, note that the  $m = 1$  terms each contain two terms with the single zero-overlap in either the first or second contraction. Exploiting the symmetry  $v_{\mu\nu\sigma\tau} = v_{\nu\mu\sigma\tau}$  and the identity

$${}^{xw}P^{\tau\mu} {}^{xw}P^{\sigma\nu} - {}^{xw}P^{\sigma\mu} {}^{xw}P^{\tau\nu} = \sum_{k_1 \neq k_2} \left( {}^{xw}P^{\tau\mu}_{k_1} {}^{xw}P^{\sigma\nu}_{k_2} - {}^{xw}P^{\sigma\mu}_{k_1} {}^{xw}P^{\tau\nu}_{k_2} \right) \quad (64)$$

then allows the two-body generalized Slater–Condon rules<sup>32</sup> to be recovered as

$$\langle {}^x\Phi | \hat{v} | {}^w\Phi \rangle = \begin{cases} {}^{xw}\tilde{S} \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} ({}^{xw}W^{\tau\mu} {}^{xw}W^{\sigma\nu} - {}^{xw}W^{\sigma\mu} {}^{xw}W^{\tau\nu}), & \text{no zeros} \\ 2 {}^{xw}\tilde{S} \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} ({}^{xw}P^{\tau\mu}_k {}^{xw}W^{\sigma\nu} - {}^{xw}P^{\sigma\mu}_k {}^{xw}W^{\tau\nu}), & \text{one zero } (k) \\ 2 {}^{xw}\tilde{S} \sum_{\mu\nu\sigma\tau} v_{\mu\nu\sigma\tau} ({}^{xw}P^{\tau\mu}_{k_1} {}^{xw}P^{\sigma\nu}_{k_2} - {}^{xw}P^{\sigma\mu}_{k_1} {}^{xw}P^{\tau\nu}_{k_2}), & \text{two zeros } (k_1, k_2) \\ 0, & \text{otherwise.} \end{cases} \quad (65)$$

Note that for the  $m = 1$  case with one zero-overlap in orbital  $k$ , we have exploited the identities  ${}^{xw}P^{\tau\mu}_k {}^{xx}P^{\sigma\nu}_k = {}^{xw}P^{\sigma\mu}_k {}^{xx}P^{\tau\nu}_k$  and  ${}^{xw}P^{\tau\mu}_k {}^{xw}P^{\sigma\nu}_k = {}^{xw}P^{\sigma\mu}_k {}^{xw}P^{\tau\nu}_k$  to introduce the simplification  ${}^{xw}M \rightarrow {}^{xw}W$ .

## VI. MATRIX ELEMENTS FOR EXCITED CONFIGURATIONS

While re-deriving the generalized Slater–Condon rules provides an important verification of the extended nonorthogonal Wick’s theorem, it does not provide any computational advantage over the original framework. On the contrary, the primary focus of our new framework involves deriving matrix elements between excited configurations from a pair of nonorthogonal determinants, e.g.,  $\langle {}^x\Phi_{ij\dots}^{ab\dots} | \dots | {}^w\Phi_{kl\dots}^{cd\dots} \rangle$ . Terms of this form arise in perturbative corrections to NOCI,<sup>14,42–44</sup> the NOCI–CIS approach for core excitations,<sup>8,49</sup> and the evaluation of  $\langle S^2 \rangle$  coupling terms in NOCI expansions.<sup>6</sup> Furthermore, these nonorthogonal matrix elements will be required to evaluate inter-state coupling elements between orbital-optimized excited-state wave functions identified using excited-state mean-field theory<sup>22</sup> or state-specific complete active space SCF.<sup>24</sup>

Until now, evaluating these matrix elements has required the direct application of the generalized Slater–Condon rules to each pair of excitations. This approach leads to significant computational costs associated with the biorthogonalization of the excited

determinants, which scales as  $\mathcal{O}(N^3)$  each time. In this section, we show how the extended nonorthogonal Wick’s theorem allows these matrix elements to be evaluated using only biorthogonalized reference determinants. When large numbers of coupling elements are required, avoiding the biorthogonalization of these excited configurations significantly reduces the overall computational cost. Furthermore, the cost of certain nonorthogonal matrix elements (including all one-body operators) becomes independent of the number of electrons or basis functions if additional intermediate matrix elements are computed for each pair of reference determinants.

In this section, we describe how the extended nonorthogonal Wick’s theorem can be applied to matrix elements between excited configurations. First, we discuss how the fundamental contractions described in Sec. IV C can be modified to an MO-based form that makes excited configurations easier to handle. We then illustrate the derivation of certain overlap, one-body, and two-body matrix elements between nonorthogonal excited configurations. The resulting expressions are entirely generalized for any pair of nonorthogonal reference determinants and can significantly reduce the computational scaling compared to a naïve application of the generalized Slater–Condon rules.

### A. Asymmetric representation

Evaluating matrix elements between two excited determinants will require the evaluation of the asymmetric contractions  ${}^x\overline{b_p^\dagger w} b_q$

with respect to the asymmetric Fermi vacuum  $\langle {}^x\Phi | \cdots | {}^w\Phi \rangle$ . Expanding the molecular orbital creation and annihilation operators using Eq. (3) yields

$${}^x\overline{b_p^\dagger} {}^w b_q = \sum_{\mu\nu\sigma\tau} ({}^w C^*)_{q,\mu} g_{\mu\sigma} ({}^{\overline{a^\tau}})^\dagger ({}^{\overline{a^\sigma}}) g_{\tau\nu} ({}^x C)_{p,\nu}^\vee, \quad (66a)$$

$${}^x\overline{b_p} {}^w b_q^\dagger = \sum_{\mu\nu\sigma\tau} ({}^x C^*)_{p,\mu}^\vee g_{\mu\sigma} ({}^{\overline{a^\sigma}}) ({}^{\overline{a^\tau}})^\dagger g_{\tau\nu} ({}^w C)_{q,\nu}^\vee. \quad (66b)$$

Introducing the fundamental contractions for  $({}^{\overline{a^\mu}})^\dagger a^\nu$  and  $({}^{\overline{a^\mu}})^\dagger a^\nu$  defined in Eqs. (46) and (47), respectively, then reduces these expressions to different forms depending on the number of zero-overlaps  $m_k$  associated with the contraction,

$${}^x\overline{b_p^\dagger} {}^w b_q = \begin{cases} {}^{wx} X_{qp}, & m_k = 0 \\ {}^{wx} \tilde{X}_{qp}, & m_k = 1 \\ 0, & m_k > 1, \end{cases} \quad (67a)$$

$${}^x\overline{b_p} {}^w b_q^\dagger = \begin{cases} {}^{xw} Y_{pq}, & m_k = 0 \\ {}^{xw} \tilde{Y}_{pq}, & m_k = 1 \\ 0, & m_k > 1. \end{cases} \quad (67b)$$

Here, we have defined the “screened” overlap terms

$${}^{wx} X_{qp} = \sum_{\mu\nu\sigma\tau} ({}^w C^*)_{q,\mu}^\vee g_{\mu\sigma} ({}^{xw} M)^{\sigma\tau} g_{\tau\nu} ({}^x C)_{p,\nu}^\vee, \quad (68a)$$

$${}^{wx} \tilde{X}_{qp} = \sum_{\mu\nu\sigma\tau} ({}^w C^*)_{q,\mu}^\vee g_{\mu\sigma} ({}^{xw} P)^{\sigma\tau} g_{\tau\nu} ({}^x C)_{p,\nu}^\vee, \quad (68b)$$

and

$${}^{xw} Y_{pq} = -({}^{xw} X_{pq} - {}^{xw} S_{pq}), \quad (69a)$$

$${}^{xw} \tilde{Y}_{pq} = -{}^{xw} \tilde{X}_{pq}, \quad (69b)$$

with  ${}^{xw} P$  and  ${}^{xw} M$  defined in Eqs. (38b) and (38d), respectively, and the overlap element

$${}^{wx} S_{pq} = \sum_{\mu\nu} ({}^w C^*)_{p,\mu}^\vee g_{\mu\nu} ({}^x C)_{q,\nu}^\vee. \quad (70)$$

Although the indexing notation used in Eq. (68a) may seem counter-intuitive, we find that it helps to keep track of the bra and ket orbital coefficients in the screened overlap terms.

Crucially, the orbital coefficients used to evaluate these contractions do not need to be the same as those used to evaluate the  ${}^{xw} M$  and  ${}^{xw} P$  matrices. This feature is particularly advantageous as the excited configurations can be represented in terms of the original orbital basis while the  ${}^{xw} M$  and  ${}^{xw} P$  matrices are evaluated in the biorthogonal basis. As a result, only the reference determinants need to be biorthogonalized, and the remaining matrix elements are evaluated in terms of these screened overlap terms. Furthermore, the screened overlap elements are themselves one-body matrix elements that can be computed once for a given pair of determinants and stored, before being combined to evaluate more complicated matrix

elements. With  $N_{\text{ref}}$  determinants, the total cost of computing these intermediates therefore scales as  $\mathcal{O}(N_{\text{ref}}^2 n^3)$ .

To take full advantage of these asymmetric contractions, the one- and two-body operators can also be represented in terms of one set of molecular orbitals as

$$\hat{f} = \sum_{pq} {}^x f_{pq} {}^x b_p^\dagger {}^x b_q, \quad (71a)$$

$$\hat{v} = \sum_{pqrs} {}^x v_{pqrs} {}^x b_p^\dagger {}^x b_q^\dagger {}^x b_s {}^x b_r, \quad (71b)$$

where we have defined the transformed matrix elements

$${}^x f_{pq} = \sum_{\mu\nu} ({}^x C^*)_{p,\mu}^\vee f_{\mu\nu} ({}^x C)_{q,\nu}^\vee \quad (72)$$

and

$${}^x v_{pqrs} = \sum_{\mu\nu\sigma\tau} ({}^x C^*)_{p,\mu}^\vee ({}^x C^*)_{q,\nu}^\vee v_{\mu\nu\sigma\tau} ({}^x C)_{r,\sigma}^\vee ({}^x C)_{s,\tau}^\vee. \quad (73)$$

Evaluating nonorthogonal matrix elements through the extended nonorthogonal Wick's theorem then proceeds using similar steps outlined in Sec. IV E:

1. Assemble all fully contracted combinations of the asymmetric operator and excitation operator product and compute the corresponding phase factors.
2. For each term, sum every possible way to distribute  $m$  zeros among the contractions such that  $\sum_i m_i = m$ .
3. For every set of  $\{m_i\}$  in each term, construct the relevant contribution as a product of fundamental contractions defined in Eqs. (67a) and (67b) depending on whether each contraction has  $m_i = 0$  or 1.
4. Multiply the combined expression with the reduced overlap  ${}^{xw} \tilde{S}$  of the reference determinants.

Note that the number of zero-overlap orbitals in these expressions corresponds to the biorthogonalized reference determinants, not the excited configurations.

In Secs. IV B–IV E, we will illustrate this process through a series of typical nonorthogonal matrix elements. As we shall see below, evaluating the sum of every combination of zero-overlap indices assigned to each contraction quickly leads to complicated equations. However, we can derive the general structure of a matrix element for the case with no zero-overlap orbitals  $m = 0$  and then recover the forms for different values of  $m$  by distributing the zero-overlap indices over each fundamental contraction. We will therefore focus on the parent equation with  $m = 0$ , which we refer to as the “canonical form” and denote using the notation  $\langle \cdots \rangle_0$ . Furthermore, any matrix element with more zero-overlap orbitals than the total number of contractions must be strictly zero.

## B. Overlap terms

First, we consider the overlap element between excited determinants. The simplest overlap matrix element involves only a single excitation  $\langle {}^x\Phi | {}^w \Phi_i^a \rangle$ . For  $m$  zero-overlap orbitals in the reference determinants, this single excitation matrix element can be identified using one contraction as

$$\langle {}^x\Phi | {}^w\Phi_i^a \rangle = \langle {}^x\Phi | {}^w\overline{b_a^\dagger b_i} | {}^w\Phi \rangle = \begin{cases} {}^{xw}\tilde{S} {}^{ww}X_{ia}, & m = 0 \\ {}^{xw}\tilde{S} {}^{ww}\bar{X}_{ia}, & m = 1 \\ 0, & m > 1. \end{cases} \quad (74)$$

Here, the canonical form is simply

$$\langle {}^x\Phi | {}^w\Phi_i^a \rangle_0 = {}^{xw}\tilde{S} {}^{ww}X_{ia}, \quad (75)$$

and there is only one way to assign one zero-overlap orbital to the single contraction. Note that the reduced overlap between the *reference* determinants remains a prefactor for the overall matrix elements.

Next, the overlap of two single excitations  $\langle {}^x\Phi_i^a | {}^w\Phi_j^b \rangle$  can be evaluated as

$$\begin{aligned} \langle {}^x\Phi_i^a | {}^w\Phi_j^b \rangle &= \langle {}^x\Phi | {}^w\overline{b_i^\dagger b_a} {}^w\overline{b_j^\dagger b_j} | {}^w\Phi \rangle + \langle {}^x\Phi | {}^w\overline{b_i^\dagger b_a} {}^w\overline{b_b^\dagger b_j} | {}^w\Phi \rangle \\ &= \begin{cases} {}^{xw}\tilde{S}({}^{ww}X_{jb} {}^{xx}X_{ai} + {}^{wx}X_{ji} {}^{xw}Y_{ab}), & m = 0 \\ {}^{xw}\tilde{S}({}^{ww}X_{jb} {}^{xx}\bar{X}_{ai} + {}^{wx}X_{ji} {}^{xw}\bar{Y}_{ab} + {}^{ww}\bar{X}_{jb} {}^{xx}X_{ai} + {}^{wx}\bar{X}_{ji} {}^{xw}Y_{ab}), & m = 1 \\ {}^{xw}\tilde{S}({}^{ww}\bar{X}_{jb} {}^{xx}\bar{X}_{ai} + {}^{wx}\bar{X}_{ji} {}^{xw}\bar{Y}_{ab}), & m = 2 \\ 0, & m > 2. \end{cases} \end{aligned} \quad (76)$$

The canonical form for this element is the  $m = 0$  term

$$\langle {}^x\Phi_i^a | {}^w\Phi_j^b \rangle_0 = {}^{xw}\tilde{S}({}^{ww}X_{jb} {}^{xx}X_{ai} + {}^{wx}X_{ji} {}^{xw}Y_{ab}). \quad (77)$$

The  $m = 1$  term is recovered by taking the sum of the two different ways to distribute one zero-overlap orbital to the two contractions, while there is only one way to distribute two zero-overlap orbitals for the  $m = 2$  term.

As a third example, consider the double excitation overlap  $\langle {}^x\Phi | {}^w\Phi_{ij}^{ab} \rangle$ , which can be evaluated as

$$\begin{aligned} \langle {}^x\Phi | {}^w\Phi_{ij}^{ab} \rangle &= \langle {}^x\Phi | {}^w\overline{b_a^\dagger b_i} {}^w\overline{b_b^\dagger b_j} | {}^w\Phi \rangle + \langle {}^x\Phi | {}^w\overline{b_a^\dagger b_i} {}^w\overline{b_j^\dagger b_b} | {}^w\Phi \rangle \\ &= \begin{cases} {}^{xw}\tilde{S}({}^{ww}X_{ia} {}^{ww}X_{jb} - {}^{ww}X_{ib} {}^{ww}X_{ja}), & m = 0 \\ {}^{xw}\tilde{S}({}^{ww}X_{ia} {}^{ww}\bar{X}_{jb} - {}^{ww}X_{ib} {}^{ww}\bar{X}_{ja} + {}^{ww}\bar{X}_{ia} {}^{ww}X_{jb} - {}^{ww}\bar{X}_{ib} {}^{ww}X_{ja}), & m = 1 \\ {}^{xw}\tilde{S}({}^{ww}\bar{X}_{ia} {}^{ww}\bar{X}_{jb} - {}^{ww}\bar{X}_{ib} {}^{ww}\bar{X}_{ja}), & m = 2 \\ 0, & m > 2. \end{cases} \end{aligned} \quad (78)$$

In this case, the minus sign arises from the intrinsic  $-1$  phase that results from the intersection of the contraction lines in  $\langle {}^x\Phi | {}^w\overline{b_a^\dagger b_i} {}^w\overline{b_b^\dagger b_j} | {}^w\Phi \rangle$ . The relevant canonical form of this matrix element is

$$\langle {}^x\Phi | {}^w\Phi_{ij}^{ab} \rangle_0 = {}^{xw}\tilde{S}({}^{ww}X_{ia} {}^{ww}X_{jb} - {}^{ww}X_{ib} {}^{ww}X_{ja}), \quad (79)$$

from which the terms for  $m = 1$  and  $m = 2$  can be derived. Notably, this form of  $\langle {}^x\Phi | {}^w\Phi_{ij}^{ab} \rangle$  has previously been derived for  $m = 0$  in Refs. 42 and 45, but our derivation generalizes for any number of zero-overlap orbitals.

### C. One-body operators

We now consider matrix elements for one-body operators of the form given in Eq. (71a). To further simplify the subsequent expressions, we can introduce intermediate matrices that account for partial contraction with the one-body operator. In particular, we introduce the partially contracted intermediate terms

$$F_0 = \sum_{pq} {}^x f_{pq} {}^{xx}X_{qp}, \quad (80a)$$

$${}^{yw}[XFX]_{rs} = \sum_{pq} {}^{yx}X_{rp} {}^x f_{pq} {}^{xw}X_{qs}, \quad (80b)$$

$${}^{yw}[XFY]_{rs} = \sum_{pq} {}^{yx}X_{rp} {}^x f_{pq} {}^{xw}Y_{qs}, \quad (80c)$$

$${}^{yw}[YFX]_{rs} = \sum_{pq} {}^{yx}Y_{rp} {}^x f_{pq} {}^{xw}X_{qs}, \quad (80d)$$

$${}^{yw}[YFY]_{rs} = \sum_{pq} {}^{yx}Y_{rp} {}^x f_{pq} {}^{xw}Y_{qs}. \quad (80e)$$

Note that the individual contraction in the  $F_0$  term, indicated by the  $X$  matrix, may correspond to a zero-overlap orbital pair, as indicated

by the notation  $\tilde{F}_0$ . The  ${}^{yw}[XFX]_{rs}$  (and similar) terms correspond to two contractions and can thus be assigned two zero overlap orbitals. The different possibilities of assigning these zero-overlap contractions is denoted as  ${}^{yw}[\tilde{X}FX]_{rs}$  or  ${}^{yw}[X\tilde{F}X]_{rs}$  and similarly for terms involving the  $Y$  contraction. Crucially, these intermediate values also correspond to orbital pairs and can be precomputed once for each pair of reference determinants, leading to a one-off computational cost that scales as  $\mathcal{O}(N_{\text{ref}}^2 n^3)$ . As a result, the summation over the  $p, q$  indices is avoided for the subsequent evaluation of matrix elements between excited determinants, and the computational scaling of these one-body terms is the same as the overlap matrix elements.

To illustrate the application of this approach, we take the simplest one-body matrix element  $\langle {}^x\Phi|\hat{f}|{}^w\Phi_i^a\rangle$ . For  $m$  zero-overlap orbitals in the reference determinants, this matrix element can be expanded as

$$\langle {}^x\Phi|\hat{f}|{}^w\Phi_i^a\rangle = \sum_{pq} x f_{pq} \left( \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_i} |{}^w\Phi\rangle + \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_i} |{}^w\Phi\rangle \right). \quad (81)$$

Using the contractions defined in Eqs. (67a) and (67b), and the intermediate terms defined in Eq. (80), the canonical form for  $m = 0$  is given as

$$\langle {}^x\Phi|\hat{f}|{}^w\Phi_i^a\rangle_0 = {}^{xw}\tilde{S}(F_0 {}^{ww}X_{ia} + {}^{ww}[XFY]_{ia}). \quad (82)$$

If  $F_0$ ,  ${}^{ww}X_{ia}$ , and  ${}^{ww}[XFY]_{ia}$  are precomputed and stored once for the reference determinants, the subsequent cost of evaluating  $\langle {}^x\Phi|\hat{f}|{}^w\Phi_i^a\rangle$  is independent of the number of electrons or basis functions and scales as  $\mathcal{O}(1)$ . From this canonical form, expressions for  $m = 1$  and 2 (or higher) can be identified as

$$\langle {}^x\Phi|\hat{f}|{}^w\Phi_i^a\rangle = \begin{cases} {}^{xw}\tilde{S}(F_0 {}^{ww}X_{ia} + {}^{ww}[XFY]_{ia}), & m = 0 \\ {}^{xw}\tilde{S}(\tilde{F}_0 {}^{ww}X_{ia} + {}^{ww}[\tilde{X}FY]_{ia} + F_0 {}^{ww}\tilde{X}_{ia} + {}^{ww}[X\tilde{F}Y]_{ia}), & m = 1 \\ {}^{xw}\tilde{S}(\tilde{F}_0 {}^{ww}\tilde{X}_{ia} + {}^{ww}[\tilde{X}\tilde{F}Y]_{ia}), & m = 2 \\ 0, & m > 2. \end{cases} \quad (83)$$

Next, consider the coupling of two single excitations  $\langle {}^x\Phi_i^a|\hat{f}|{}^w\Phi_j^b\rangle$ , corresponding to the fully contracted terms

$$\begin{aligned} \langle {}^x\Phi_i^a|\hat{f}|{}^w\Phi_j^b\rangle = \sum_{pq} x f_{pq} & \left( \langle {}^x\Phi|\overline{b_i^\dagger x b_a} \overline{w b_p^\dagger w b_q} \overline{w b_b^\dagger w b_j} |{}^w\Phi\rangle + \langle {}^x\Phi|\overline{b_i^\dagger x b_a} \overline{w b_p^\dagger w b_q} \overline{w b_b^\dagger w b_j} |{}^w\Phi\rangle \right) \\ & + \langle {}^x\Phi|\overline{b_i^\dagger x b_a} \overline{w b_p^\dagger w b_q} \overline{w b_b^\dagger w b_j} |{}^w\Phi\rangle + \langle {}^x\Phi|\overline{b_i^\dagger x b_a} \overline{w b_p^\dagger w b_q} \overline{w b_b^\dagger w b_j} |{}^w\Phi\rangle \\ & + \langle {}^x\Phi|\overline{b_i^\dagger x b_a} \overline{w b_p^\dagger w b_q} \overline{w b_b^\dagger w b_j} |{}^w\Phi\rangle + \langle {}^x\Phi|\overline{b_i^\dagger x b_a} \overline{w b_p^\dagger w b_q} \overline{w b_b^\dagger w b_j} |{}^w\Phi\rangle. \end{aligned} \quad (84)$$

In this case, the canonical form is given by

$$\langle {}^x\Phi_i^a|\hat{f}|{}^w\Phi_j^b\rangle_0 = {}^{xw}\tilde{S}(F_0 {}^{xx}X_{ai} {}^{ww}X_{jb} + F_0 {}^{wx}X_{ji} {}^{xw}Y_{ab} + {}^{ww}[XFY]_{jb} {}^{xx}X_{ai} + {}^{xx}[YFX]_{ai} {}^{ww}X_{jb} + {}^{xw}[YFY]_{ab} {}^{wx}X_{ji} - {}^{wx}[XFX]_{ji} {}^{xw}Y_{ab}). \quad (85)$$

Again, the form for  $m$  zero-overlap orbitals can be recovered as the sum of every way to distribute the zero-overlap contractions over the  $X$ ,  $Y$ , or  $F_0$  terms in each product. We omit the explicit form of these  $m > 0$  expressions to maintain brevity.

Finally, consider the one-body coupling of a reference determinant and a double excitation

$$\begin{aligned} \langle {}^x\Phi|\hat{f}|{}^w\Phi_{ij}^{ab}\rangle = \sum_{pq} x f_{pq} & \left( \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_b^\dagger w b_j} \overline{w b_i} |{}^w\Phi\rangle + \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_b^\dagger w b_j} \overline{w b_i} |{}^w\Phi\rangle \right) \\ & + \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_b^\dagger w b_j} \overline{w b_i} |{}^w\Phi\rangle + \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_b^\dagger w b_j} \overline{w b_i} |{}^w\Phi\rangle \\ & + \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_b^\dagger w b_j} \overline{w b_i} |{}^w\Phi\rangle + \langle {}^x\Phi|\overline{b_p^\dagger x b_q} \overline{w b_a^\dagger w b_b^\dagger w b_j} \overline{w b_i} |{}^w\Phi\rangle. \end{aligned} \quad (86)$$

The corresponding canonical form for  $m = 0$  is

$$\langle {}^x\Phi|\hat{f}|{}^w\Phi_{ij}^{ab}\rangle_0 = {}^{xw}\tilde{S}(F_0 {}^{ww}X_{ia} {}^{ww}X_{jb} - F_0 {}^{ww}X_{ib} {}^{ww}X_{ja} + {}^{ww}[XFY]_{ia} {}^{ww}X_{jb} + {}^{ww}[XFY]_{jb} {}^{ww}X_{ia} - {}^{ww}[XFY]_{ja} {}^{ww}X_{ib} - {}^{ww}[XFY]_{ib} {}^{ww}X_{ja}), \quad (87)$$

from which expressions for  $m > 0$  can be obtained as described above.

## D. Two-body operators

Finally, we consider matrix elements for two-body operators, with the general form given in Eq. (71b). Again, we can define intermediate matrices in the orbital basis that can be pre-computed for a given pair of nonorthogonal determinants. First, we introduce analogs of the Coulomb and exchange matrices, respectively, defined as

$$\begin{aligned} {}^x J_{pr} &= \sum_{qs} {}^x v_{pqrs} {}^{xx} X_{sq}, \\ {}^x K_{ps} &= \sum_{qr} {}^x v_{pqrs} {}^{xx} X_{rq}. \end{aligned} \quad (88)$$

Partially contracted intermediate matrices can then be defined as, e.g.,

$${}^{yw} [X(J-K)X]_{rs} = \sum_{pq} {}^{yx} X_{rp} ({}^x J - {}^x K)_{pq} {}^{xw} X_{qs}, \quad (89a)$$

$${}^{yw} [Y(J-K)X]_{rs} = \sum_{pq} {}^{yx} Y_{rp} ({}^x J - {}^x K)_{pq} {}^{xw} X_{qs}, \quad (89b)$$

$${}^{yw} [X(J-K)Y]_{rs} = \sum_{pq} {}^{yx} X_{rp} ({}^x J - {}^x K)_{pq} {}^{xw} Y_{qs}, \quad (89c)$$

$${}^{yw} [Y(J-K)Y]_{rs} = \sum_{pq} {}^{yx} Y_{rp} ({}^x J - {}^x K)_{pq} {}^{xw} Y_{qs}, \quad (89d)$$

with the constant value

$$V_0 = \sum_{pqrs} {}^x v_{pqrs} ({}^{xx} X_{sq} {}^{xx} X_{rp} - {}^{xx} X_{rq} {}^{xx} X_{sp}). \quad (90)$$

The one-off computational cost of evaluating each of these intermediate matrices is dominated by the cost of evaluating Eqs. (88)

and (90) for each reference determinant, and so the overall scaling is  $\mathcal{O}(N_{\text{ref}} n^4)$ .

Like the one-body intermediate matrices, when the zero-overlap orbitals are distributed among the contractions, these zeros may be independently assigned to any of the  $X$ ,  $Y$ , or  $(J-K)$  terms, as each of these contain one contraction. For example, with  $m = 1$ , one must consider the sum of three terms given as, e.g.,

$${}^{yw} [\tilde{X}(J-K)X]_{rs} + {}^{yw} [X(\tilde{J}-\tilde{K})X]_{rs} + {}^{yw} [X(J-K)\tilde{X}]_{rs}, \quad (91)$$

where

$${}^x \tilde{J}_{pr} = \sum_{qs} {}^x v_{pqrs} {}^{xx} \tilde{X}_{sq} \quad (92a)$$

$${}^x \tilde{K}_{ps} = \sum_{qr} {}^x v_{pqrs} {}^{xx} \tilde{X}_{rq}. \quad (92b)$$

On the contrary, the  $V_0$  constant term corresponds to two contractions and can be assigned two zero-overlap orbitals. Noting the symmetry  ${}^x v_{pqrs} = {}^x v_{qpsr}$ , we denote these possibilities as

$$\tilde{V}_0 = 2 \sum_{pqrs} {}^x v_{pqrs} ({}^{xx} \tilde{X}_{sq} {}^{xx} X_{rp} - {}^{xx} \tilde{X}_{rq} {}^{xx} X_{sp}), \quad (93a)$$

$$\tilde{\tilde{V}}_0 = \sum_{pqrs} {}^x v_{pqrs} ({}^{xx} \tilde{X}_{sq} {}^{xx} \tilde{X}_{rp} - {}^{xx} \tilde{X}_{rq} {}^{xx} \tilde{X}_{sp}). \quad (93b)$$

Here, the factor of two in Eq. (93a) arises because the zero-overlap orbital can be assigned to either the first or second contraction to give the same result.

To illustrate the application of this approach for excited configurations, we first consider the two-body matrix element

$$\begin{aligned} \langle {}^x \Phi | \hat{v} | {}^w \Phi_i^a \rangle &= \sum_{pqrs} {}^x v_{pqrs} \left( \langle {}^x \Phi | \overbrace{b_p^\dagger b_q^\dagger b_s^\dagger b_r}^{x} \overbrace{b_a^\dagger b_i}^{w} | {}^w \Phi \rangle + \langle {}^x \Phi | \overbrace{b_p^\dagger b_q^\dagger b_s^\dagger b_r}^{x} \overbrace{b_a^\dagger b_i}^{w} | {}^w \Phi \rangle \right. \\ &\quad + \langle {}^x \Phi | \overbrace{b_p^\dagger b_q^\dagger b_s^\dagger b_r}^{x} \overbrace{b_a^\dagger b_i}^{w} | {}^w \Phi \rangle + \langle {}^x \Phi | \overbrace{b_p^\dagger b_q^\dagger b_s^\dagger b_r}^{x} \overbrace{b_a^\dagger b_i}^{w} | {}^w \Phi \rangle \\ &\quad \left. + \langle {}^x \Phi | \overbrace{b_p^\dagger b_q^\dagger b_s^\dagger b_r}^{x} \overbrace{b_a^\dagger b_i}^{w} | {}^w \Phi \rangle + \langle {}^x \Phi | \overbrace{b_p^\dagger b_q^\dagger b_s^\dagger b_r}^{x} \overbrace{b_a^\dagger b_i}^{w} | {}^w \Phi \rangle \right). \end{aligned} \quad (94)$$

Combining each contraction and exploiting the intermediates in Eqs. (89) and (90) yields the canonical form ( $m = 0$ ) for this matrix element as

$$\langle {}^x \Phi | \hat{v} | {}^w \Phi_i^a \rangle_0 = {}^{xw} \tilde{S} (V_0 {}^{ww} X_{ia} + 2 {}^{ww} [X(J-K)Y]_{ia}). \quad (95)$$

Distributing the zero-overlap orbitals among each contraction then leads to explicit expressions for all  $m$  values as

$$\langle {}^x \Phi | \hat{v} | {}^w \Phi_i^a \rangle = \begin{cases} {}^{xw} \tilde{S} (V_0 {}^{ww} X_{ia} + 2 {}^{ww} [X(J-K)Y]_{ia}), & m = 0 \\ {}^{xw} \tilde{S} (V_0 {}^{ww} \tilde{X}_{ia} + 2 {}^{ww} [\tilde{X}(J-K)Y]_{ia} + \tilde{V}_0 {}^{ww} X_{ia} + 2 {}^{ww} [X(\tilde{J}-\tilde{K})Y]_{ia} + 2 {}^{ww} [X(J-K)\tilde{Y}]_{ia}), & m = 1 \\ {}^{xw} \tilde{S} (\tilde{\tilde{V}}_0 {}^{ww} X_{ia} + 2 {}^{ww} [\tilde{X}(\tilde{J}-\tilde{K})Y]_{ia} + \tilde{V}_0 {}^{ww} \tilde{X}_{ia} + 2 {}^{ww} [\tilde{X}(J-K)\tilde{Y}]_{ia} + 2 {}^{ww} [X(\tilde{J}-\tilde{K})\tilde{Y}]_{ia}), & m = 2 \\ {}^{xw} \tilde{S} (\tilde{\tilde{V}}_0 {}^{ww} \tilde{X}_{ia} + 2 {}^{ww} [\tilde{X}(\tilde{J}-\tilde{K})\tilde{Y}]_{ia}), & m = 3 \\ 0 & m > 3. \end{cases} \quad (96)$$

Next, we consider the two-body coupling of two singly excited determinants as

$$\langle {}^x\Phi_i^a|\hat{v}|{}^w\Phi_j^b\rangle = \sum_{pqrs} {}^xv_{pqrs} \langle {}^x\Phi|{}^xb_i^\dagger b_a^\dagger b_p^\dagger b_q^\dagger b_s^x b_r^w b_b^\dagger b_j|{}^w\Phi\rangle. \quad (97)$$

With a total of 24 possible ways to fully contract the corresponding matrix element, we maintain brevity by directly presenting the canonical form for  $m = 0$  as

$$\begin{aligned} \langle {}^x\Phi_i^a|\hat{v}|{}^w\Phi_j^b\rangle_0 = {}^{xw}\tilde{S} & \left[ V_0({}^{xx}X_{ai}{}^{ww}X_{jb} + {}^{wx}X_{ji}{}^{xw}Y_{ab}) + 2({}^{xx}[Y(J-K)X]_{ai}{}^{ww}X_{jb} + {}^{ww}[X(J-K)Y]_{jb}{}^{xx}X_{ai}) \right. \\ & \left. + 2({}^{xw}[Y(J-K)Y]_{ab}{}^{wx}X_{ji} - {}^{wx}[X(J-K)X]_{ji}{}^{xw}X_{ab}) + 2\sum_{pqrs} {}^xv_{pqrs}{}^{xx}Y_{ap}{}^{wx}X_{jq}({}^{xw}Y_{sb}{}^{xx}X_{ri} - {}^{xw}Y_{rb}{}^{xx}X_{si}) \right]. \quad (98) \end{aligned}$$

From here, expressions for the cases with  $m > 0$  can be recovered by distributing the zero-overlap orbitals over the contractions associated with the  $V$ ,  $X$ ,  $(J - K)$ , or  $Y$  matrices. While partially contracted intermediate expressions could also be evaluated to avoid the nested summation on the last line in Eq. (98), this will generally incur an unacceptably large storage cost.

Finally, we consider the two-body coupling of a reference determinant and a double excitation

$$\langle {}^x\Phi|\hat{v}|{}^w\Phi_{ij}^{ab}\rangle = \sum_{pqrs} {}^xv_{pqrs} \langle {}^x\Phi|{}^xb_p^\dagger b_q^\dagger b_s^x b_r^w b_a^\dagger b_b^\dagger b_j^w b_i|{}^w\Phi\rangle. \quad (99)$$

Again, there are a total of 24 possible ways to fully contract the corresponding matrix element, so we advance directly to the canonical form for  $m = 0$ , given as

$$\begin{aligned} \langle {}^x\Phi|\hat{v}|{}^w\Phi_{ij}^{ab}\rangle_0 = {}^{xw}\tilde{S} & \times \left[ V_0({}^{ww}X_{ia}{}^{ww}X_{jb} - {}^{ww}X_{ib}{}^{ww}X_{ja}) + 2({}^{ww}[X(J-K)Y]_{ia}{}^{ww}X_{jb} - {}^{ww}[X(J-K)Y]_{ib}{}^{ww}X_{ja}) \right. \\ & \left. + 2({}^{ww}[X(J-K)Y]_{jb}{}^{ww}X_{ia} - {}^{ww}[X(J-K)Y]_{ja}{}^{ww}X_{ib}) + 2\sum_{pqrs} {}^xv_{pqrs}{}^{wx}X_{ip}{}^{wx}X_{jq}({}^{xw}Y_{sb}{}^{xw}Y_{ra} - {}^{xw}Y_{sa}{}^{xw}Y_{rb}) \right]. \quad (100) \end{aligned}$$

We note that this formula has previously been identified in Ref. 45 for  $m = 0$ , but our approach now allows this result to be extended for  $m > 0$ .

## E. Illustration of scaling

Through the use of intermediate matrices in Secs. VI A–VI C, the cost of evaluating overlap and one-body matrix elements between excited configurations can be made independent of the number of electrons  $N$  or basis functions  $n$ . In both cases, the computational cost of evaluating the intermediate matrices scales as  $\mathcal{O}(N_{\text{ref}}^2 n^3)$ , which is determined by the cost of evaluating the screened overlap terms in Eqs. (68a) and (68b). The subsequent cost of evaluating matrix elements between excited configurations then scales as  $\mathcal{O}(1)$ . In contrast, the cost of applying the generalized Slater–Condon rules for overlap or one-body operators is dominated by the evaluation of the occupied orbital overlap matrix, giving a scaling of  $\mathcal{O}(n^2 N)$  for every excited coupling term.

For two-body operators, a similar application of two-body intermediates requires transformations of the two-electron integrals into an asymmetric MO representation for each pair of determinants

and for every possible combination of four contractions in the four possible forms given in Eq. (67a) or (67b). These two-body intermediates therefore carry a storage overhead that scales as  $\mathcal{O}(4^4 N_{\text{ref}}^2 n^4)$ , which will generally be unacceptably large. We imagine that there may be certain applications that require only a subset of these two-electron intermediates, which could then be stored to achieve the subsequent  $\mathcal{O}(1)$  scaling. In general, however, the  $n^4$  scaling for two-body matrix elements cannot be overcome, although the overall cost is still reduced by avoiding biorthogonalization of the occupied orbitals in the excited configurations.

To explicitly illustrate the computational speed-up that might be expected using the extended nonorthogonal Wick's theorem, consider the evaluation of matrix elements between two configuration interaction singles (CIS) wave functions built from different reference Slater determinants,

$$|{}^x\Psi_{\text{CIS}}\rangle = \sum_{ia} {}^xt_{ia} |{}^x\Phi_{ia}\rangle, \quad (101a)$$

$$|{}^w\Psi_{\text{CIS}}\rangle = \sum_{jb} {}^wt_{jb} |{}^w\Phi_{jb}\rangle. \quad (101b)$$



**TABLE I.** Computational scaling for evaluating the overlap, one-body, and two-body coupling between two CIS wave functions with different reference determinants using the generalized Slater–Condon rules (Slat.–Con.) or extended non-orthogonal Wick’s theorem.

Coupling term	Intermediate		Subsequent cost	
	Slat.–Con.	Wick’s	Slat.–Con.	Wick’s
$\langle {}^x\Psi_{\text{CIS}}   {}^w\Psi_{\text{CIS}} \rangle$	N/A	$n^3$	$N^3 n^4$	$N^2 n^2$
$\langle {}^x\Psi_{\text{CIS}}   \hat{f}   {}^w\Psi_{\text{CIS}} \rangle$	N/A	$n^3$	$N^3 n^4$	$N^2 n^2$
$\langle {}^x\Psi_{\text{CIS}}   \hat{v}   {}^w\Psi_{\text{CIS}} \rangle$	N/A	$n^3$	$N^2 n^6$	$N^2 n^6$

In particular, consider the evaluation of the overlap

$$\langle {}^x\Psi_{\text{CIS}} | {}^w\Psi_{\text{CIS}} \rangle = \sum_{ia,jb} {}^x t_{ia}^* {}^w t_{jb} \langle {}^x\Phi_{ia} | {}^w\Phi_{jb} \rangle, \quad (102)$$

a one-body operator  $\hat{f}$ , e.g., the transition dipole moment,

$$\langle {}^x\Psi_{\text{CIS}} | \hat{f} | {}^w\Psi_{\text{CIS}} \rangle = \sum_{ia,jb} {}^x t_{ia}^* {}^w t_{jb} \langle {}^x\Phi_{ia} | \hat{f} | {}^w\Phi_{jb} \rangle, \quad (103)$$

and a two-body operator  $\hat{v}$  such as the two-electron repulsion

$$\langle {}^x\Psi_{\text{CIS}} | \hat{v} | {}^w\Psi_{\text{CIS}} \rangle = \sum_{ia,jb} {}^x t_{ia}^* {}^w t_{jb} \langle {}^x\Phi_{ia} | \hat{v} | {}^w\Phi_{jb} \rangle. \quad (104)$$

The relevant overlap, one-body, and two-body nonorthogonal matrix elements are given in Eqs. (76), (84), and (98), respectively. In each case, only one set of intermediates need to be evaluated using the biorthogonalized reference determinants, followed by a total of  $N^2 n^2$  nonorthogonal matrix elements between singly excited configurations (assuming  $n \gg N$ ). Here, we only consider one pairing of the reference determinants.

The computational scaling associated with first evaluating the intermediate terms from the reference determinants and then evaluating all the nonorthogonal coupling terms between the excited configurations is summarized in Table I. For the overlap and one-body terms, the computational scaling is dominated by the evaluation of the intermediates, which scales as  $\mathcal{O}(n^3)$ . In comparison, the total cost of biorthogonalizing the occupied orbitals for every pair of excited configurations scales as  $\mathcal{O}(N^3 n^4)$ . For two-body operators, the cost for both approaches is dominated by the contraction with the two-body integrals for each pair of excited configurations. This leads to a total scaling  $\mathcal{O}(N^2 n^6)$  for both cases, but the extended nonorthogonal Wick’s theorem may still be faster as the prefactor is reduced for each element.

## VII. CONCLUDING REMARKS

Intuitive derivations and efficient implementations of nonorthogonal matrix elements are becoming increasingly important for the development of nonorthogonal configuration interaction methods and inter-state coupling terms for state-specific excited state wave functions. However, until now, the evaluation of these matrix elements has relied on the generalized Slater–Condon rules,

while the second-quantized nonorthogonal Wick’s theorem fails when there are any zero-overlap orbital pairs between the reference determinants. In this work, we have extended the nonorthogonal Wick’s theorem to the case where two determinants have nonorthogonal orbitals but have a zero many-electron overlap. This new theory, which we call the extended nonorthogonal Wick’s theorem, provides the most generalized framework for evaluating matrix elements between two nonorthogonal determinants using second quantization.

Among the primary advantages of our new approach is the ease of deriving and evaluating matrix elements between excited configurations from nonorthogonal reference determinants. To illustrate this feature, we have derived a series of overlap, one-body, and two-body coupling terms between nonorthogonal excited configurations. For overlap terms and one-body operators, these excited nonorthogonal matrix elements can be expressed in terms of one-body intermediate terms that can be precomputed and stored for a given pair of reference determinants. As a result, the subsequent cost of evaluating the coupling of excited configurations scales as  $\mathcal{O}(1)$ , which is the same as the conventional Slater–Condon rules or Wick’s theorem for orthogonal reference determinants. However, in the current form of the theory, the cost of evaluating two-body coupling terms scales as  $\mathcal{O}(n^4)$ , which is the same as applying the generalized Slater–Condon rules for each pair of nonorthogonal excited configurations. From a computational perspective, the extended nonorthogonal Wick’s theorem will therefore provide the greatest acceleration in tasks that require at most one-body coupling terms, such as the evaluation of transition dipole moments for orbital-optimized excited states,<sup>22–25</sup> or the coupling elements of a one-body reference Hamiltonian in nonorthogonal perturbation theories.<sup>14,42–44</sup>

Looking forward, we hope that this work will encourage and accelerate future developments in nonorthogonal electronic structure theory by creating a unifying theory for deriving challenging nonorthogonal matrix elements. As part of this vision, we are working on an open-source C++ library for evaluating typical nonorthogonal matrix elements, and we intend to report on this project soon.

## ACKNOWLEDGMENTS

H.G.A.B. was supported by New College, Oxford, through the Astor Junior Research Fellowship. The author is grateful to Rebecca Lloyd for careful proof-reading.

## APPENDIX A: THOULESS THEOREM

Thouless’ theorem<sup>39</sup> allows any Slater determinant to be represented using only single excitations from another determinant, i.e.,

$$|{}^w\Phi\rangle = \exp(\mathcal{Z}) |{}^x\Phi\rangle. \quad (A1)$$

Here, we follow Ref. 40 and provide a brief derivation of this theorem. First, the two sets of second-quantization operators can be related as

$${}^w b_i^\dagger = \sum_j {}^x b_j^\dagger {}^{xw} A_{ji} + \sum_b {}^x b_b^\dagger {}^{xw} B_{bi}, \quad (A2a)$$

$${}^w b_a^\dagger = \sum_j {}^x b_j^\dagger {}^{xw} \tilde{B}_{ja} + \sum_b {}^x b_b^\dagger {}^{xw} D_{ba}, \quad (A2b)$$

where

$${}^{xw}A_{ji} = \sum_{\mu\nu} ({}^xC^*)_{j\cdot}^{\nu} g_{\nu\mu} ({}^wC)_{i\cdot}^{\mu}, \quad (\text{A3a})$$

$${}^{xw}B_{bi} = \sum_{\mu\nu} ({}^xC^*)_{b\cdot}^{\nu} g_{\nu\mu} ({}^wC)_{i\cdot}^{\mu}, \quad (\text{A3b})$$

$${}^{xw}\tilde{B}_{ja} = \sum_{\mu\nu} ({}^xC^*)_{j\cdot}^{\nu} g_{\nu\mu} ({}^wC)_{a\cdot}^{\mu}, \quad (\text{A3c})$$

$${}^{xw}D_{ba} = \sum_{\mu\nu} ({}^xC^*)_{b\cdot}^{\nu} g_{\nu\mu} ({}^wC)_{a\cdot}^{\mu} \quad (\text{A3d})$$

are sub-blocks of the orbital overlap matrix. Taking a suitable transformation among the orbitals of  $|{}^w\Phi\rangle$  allows these to be further reduced to

$${}^w\tilde{b}_i^{\dagger} = {}^xb_i^{\dagger} + \sum_{ja} {}^xb_a^{\dagger} {}^{xw}B_{aj} ({}^{xw}A^{-1})_{ji}, \quad (\text{A4a})$$

$${}^w\tilde{b}_a^{\dagger} = {}^xb_a^{\dagger} + \sum_{jb} {}^xb_j^{\dagger} {}^{xw}\tilde{B}_{jb} ({}^{xw}D^{-1})_{ba}. \quad (\text{A4b})$$

The occupied orbitals can then be represented by the transformation

$${}^w\tilde{b}_i^{\dagger} = {}^xb_i^{\dagger} + \sum_a {}^xb_a^{\dagger} {}^{xw}Z_{ai}, \quad (\text{A5})$$

where now

$${}^{xw}Z_{ai} = \sum_j {}^{xw}B_{aj} ({}^{xw}A^{-1})_{ji}. \quad (\text{A6})$$

The overall transformation in Eq. (A1) is then given by

$$\tilde{Z} = {}^{xw}Z_{ai} {}^xb_a^{\dagger} b_i. \quad (\text{A7})$$

Finally, if the occupied orbitals are represented in a biorthogonal basis satisfying Eq. (10), then the  $Z$  matrix elements are given simply as

$${}^{xw}Z_{ai} = \sum_{\mu\nu} ({}^xC^*)_{a\cdot}^{\mu} g_{\mu\nu} ({}^w\tilde{C})_{i\cdot}^{\nu} \frac{1}{{}^{xw}\tilde{S}_i}. \quad (\text{A8})$$

## APPENDIX B: SIMILARITY TRANSFORMED OPERATORS

To evaluate common matrix elements using the single excitation operators defined in Eq. (25), we follow Ref. 40 and consider the similarity transformed operators

$$(d[C_n]^{\mu})^{\dagger} = \exp(-\tilde{Z}^{C_n})(a^{\mu})^{\dagger} \exp(\tilde{Z}^{C_n}), \quad (\text{B1a})$$

$$d[C_n]^{\mu} = \exp(-\tilde{Z}^{C_n})a^{\mu} \exp(\tilde{Z}^{C_n}). \quad (\text{B1b})$$

First, consider the expansion of  $(d[C_n]^{\mu})^{\dagger}$  as

$$\begin{aligned} (d[C_n]^{\mu})^{\dagger} &= \exp(-\tilde{Z}^{C_n})(a^{\mu})^{\dagger} \exp(\tilde{Z}^{C_n}) \\ &= (a^{\mu})^{\dagger} - [\tilde{Z}^{C_n}, (a^{\mu})^{\dagger}]. \end{aligned} \quad (\text{B2})$$

Expanding  $(a^{\mu})^{\dagger}$  using Eq. (6) and inserting the definition of  $\tilde{Z}^{C_n}$  from Eq. (25) lead to

$$\begin{aligned} (d[C_n]^{\mu})^{\dagger} &= \sum_p {}^xb_p^{\dagger} ({}^xC^*)_{p\cdot}^{\mu} - \sum_{ap} \left( \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} {}^{xw}Z_{ai} [{}^xb_a^{\dagger} {}^xb_i, {}^xb_p^{\dagger}] \right. \\ &\quad \left. + \sum_{k \in C_n} {}^{xw}\tilde{S}_{ak} [{}^xb_a^{\dagger} {}^xb_k, {}^xb_p^{\dagger}] \right) ({}^xC^*)_{p\cdot}^{\mu}. \end{aligned} \quad (\text{B3})$$

Noting the commutation relation  $[{}^xb_a^{\dagger} {}^xb_i, {}^xb_p^{\dagger}] = {}^xb_a^{\dagger} \delta_{ip}$  then allows this expression to be simplified as

$$\begin{aligned} (d[C_n]^{\mu})^{\dagger} &= \sum_p {}^xb_p^{\dagger} ({}^xC^*)_{p\cdot}^{\mu} - \sum_a {}^xb_a^{\dagger} \left( \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} {}^{xw}Z_{ai} ({}^xC^*)_{i\cdot}^{\mu} \right. \\ &\quad \left. + \sum_{k \in C_n} {}^{xw}\tilde{S}_{ak} ({}^xC^*)_{k\cdot}^{\mu} \right). \end{aligned} \quad (\text{B4})$$

Finally, separating the summation over  $p$  into a summation over the occupied orbitals  $i$  and the virtual orbitals  $a$  recovers the form given in Eq. (33a). Applying a similar approach for the  $(d[C_n]^{\mu})$  operators then recovers the expression in Eq. (33b).

## APPENDIX C: SIMILARITY TRANSFORMED CONTRACTIONS

To evaluate the contractions  $\overline{(d[c_n]^{\mu})^{\dagger}(d[c_n]^{\nu})}$  and  $\overline{(d[c_n]^{\mu})(d[c_n]^{\nu})^{\dagger}}$  with respect to the Fermi vacuum  $\langle {}^x\Phi | \dots | {}^x\Phi \rangle$ , we first expand the transformed operators in terms of the  ${}^xb_p$  and  ${}^xb_p^{\dagger}$  operators. The only non-zero contractions between these MO operators are  $\overline{{}^xb_a^{\dagger} {}^xb_b} = \delta_{ab}$  and  $\overline{{}^xb_i^{\dagger} {}^xb_j} = \delta_{ij}$ . We can therefore expand the contracted product  $\overline{(d[c_n]^{\mu})^{\dagger}(d[c_n]^{\nu})}$  as

$$\overline{(d[c_n]^{\mu})^{\dagger}(d[c_n]^{\nu})} = \sum_i ({}^xC^*)_{i\cdot}^{\nu} ({}^xC^*)_{i\cdot}^{\mu} + \sum_a \left( \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} ({}^xC^*)_{i\cdot}^{\nu} {}^{xw}Z_{ai} ({}^xC^*)_{i\cdot}^{\mu} + \sum_{k \in C_n} ({}^xC^*)_{i\cdot}^{\nu} {}^{xw}\tilde{S}_{ak} ({}^xC^*)_{k\cdot}^{\mu} \right). \quad (\text{C1})$$

Inserting the definition of  ${}^{xw}Z_{ai}$  and  ${}^{xw}\tilde{S}_{ak}$  from Eqs. (14) and (11), respectively, then leads to the form

$$\overline{(d[c_n]^\mu)^\dagger(d[c_n]^\nu)} = \sum_i ({}^x\tilde{C})_{\cdot i}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot i}^{\mu\cdot} + \sum_a \sum_{\sigma\tau} ({}^x\tilde{C})_{\cdot a}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot a}^{\sigma\cdot} g_{\sigma\tau} \left( \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} ({}^w\tilde{C})_{\cdot i}^{\tau\cdot} \frac{1}{{}^{xw}\tilde{S}_i} ({}^x\tilde{C}^*)_{\cdot i}^{\mu\cdot} + \sum_{k \in c_n} ({}^w\tilde{C})_{\cdot k}^{\tau\cdot} ({}^x\tilde{C}^*)_{\cdot k}^{\mu\cdot} \right). \quad (\text{C2})$$

Resolving the identity allows us to obtain the additional relationship

$$\sum_a ({}^x\tilde{C})_{\cdot a}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot a}^{\sigma\cdot} = g^{\nu\sigma} - \sum_j ({}^x\tilde{C})_{\cdot j}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot j}^{\sigma\cdot}, \quad (\text{C3})$$

which, when inserted into Eq. (C2), leads to

$$\overline{(d[c_n]^\mu)^\dagger(d[c_n]^\nu)} = \sum_{\{k|{}^{xw}\tilde{S}_k=0\}} ({}^x\tilde{C})_{\cdot k}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot k}^{\mu\cdot} + \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} ({}^w\tilde{C})_{\cdot i}^{\nu\cdot} \frac{1}{{}^{xw}\tilde{S}_i} ({}^x\tilde{C}^*)_{\cdot i}^{\mu\cdot} + \sum_{k \in c_n} ({}^w\tilde{C})_{\cdot k}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot k}^{\mu\cdot}. \quad (\text{C4})$$

To obtain Eq. (C4) from Eq. (C2), we have exploited the relationships

$$\sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} \sum_j \sum_{\sigma\tau} ({}^x\tilde{C})_{\cdot j}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot j}^{\sigma\cdot} g_{\sigma\tau} ({}^w\tilde{C})_{\cdot i}^{\tau\cdot} \frac{1}{{}^{xw}\tilde{S}_i} ({}^x\tilde{C}^*)_{\cdot i}^{\mu\cdot} = \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} ({}^x\tilde{C})_{\cdot i}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot i}^{\mu\cdot} \quad (\text{C5})$$

and

$$\sum_{k \in c_n} \sum_j \sum_{\sigma\tau} ({}^x\tilde{C})_{\cdot j}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot j}^{\sigma\cdot} g_{\sigma\tau} ({}^w\tilde{C})_{\cdot k}^{\tau\cdot} ({}^x\tilde{C}^*)_{\cdot k}^{\mu\cdot} = 0. \quad (\text{C6})$$

Finally, Eq. (37a) is recovered by modifying the separation of summation indices in Eq. (C4) to give the form

$$\overline{(d[c_n]^\mu)^\dagger(d[c_n]^\nu)} = \sum_{\{k|{}^{xw}\tilde{S}_k=0\}} ({}^x\tilde{C})_{\cdot k}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot k}^{\mu\cdot} + \sum_{\{k|{}^{xw}\tilde{S}_k=0\}} ({}^w\tilde{C})_{\cdot k}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot k}^{\mu\cdot} + \sum_{\{i|{}^{xw}\tilde{S}_i \neq 0\}} ({}^w\tilde{C})_{\cdot i}^{\nu\cdot} \frac{1}{{}^{xw}\tilde{S}_i} ({}^x\tilde{C}^*)_{\cdot i}^{\mu\cdot} - \sum_{k \notin c_n} ({}^w\tilde{C})_{\cdot k}^{\nu\cdot} ({}^x\tilde{C}^*)_{\cdot k}^{\mu\cdot}. \quad (\text{C7})$$

and introducing the matrices defined in Eqs. (38a)–(38d). The second contraction [Eq. (37b)] then follows from

$$\overline{(d[c_n]^\mu)^\dagger(d[c_n]^\nu)} = g^{\mu\nu} - \overline{(d[c_n]^\nu)^\dagger(d[c_n]^\mu)}. \quad (\text{C8})$$

## DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## REFERENCES

- <sup>1</sup>H. Fukutome, *Prog. Theor. Phys.* **80**, 417 (1988).
- <sup>2</sup>H. Koch and E. Dalgaard, *Chem. Phys. Lett.* **212**, 193 (1993).
- <sup>3</sup>S. Ten-no, *Theor. Chem. Acc.* **98**, 182 (1997).
- <sup>4</sup>P. Y. Ayala and H. B. Schlegel, *J. Chem. Phys.* **108**, 7560 (1998).
- <sup>5</sup>A. J. W. Thom and M. Head-Gordon, *J. Chem. Phys.* **131**, 124113 (2009).
- <sup>6</sup>E. J. Sundstrom and M. Head-Gordon, *J. Chem. Phys.* **140**, 114103 (2014).
- <sup>7</sup>N. J. Mayhall, P. R. Horn, E. J. Sundstrom, and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **16**, 22694 (2014).
- <sup>8</sup>K. J. Oosterbaan, A. F. White, and M. Head-Gordon, *J. Chem. Phys.* **149**, 044116 (2018).
- <sup>9</sup>K. T. Jensen, R. L. Benson, S. Cardamone, and A. J. W. Thom, *J. Chem. Theory Comput.* **14**, 4629 (2018).
- <sup>10</sup>H. G. A. Burton and A. J. W. Thom, *J. Chem. Theory Comput.* **15**, 4851 (2019).
- <sup>11</sup>B. C. Huynh and A. J. W. Thom, *J. Chem. Theory Comput.* **16**, 904 (2020).
- <sup>12</sup>J. Nite and C. A. Jiménez-Hoyos, *J. Chem. Theory Comput.* **15**, 5343 (2019).
- <sup>13</sup>R. K. Kathir, C. de Graaf, R. Broer, and R. W. A. Havenith, *J. Chem. Theory Comput.* **16**, 2941 (2020).
- <sup>14</sup>H. G. A. Burton and A. J. W. Thom, *J. Chem. Theory Comput.* **16**, 5586 (2020).
- <sup>15</sup>G. E. Scuseria, C. A. Jiménez-Hoyos, T. M. Henderson, K. Samanta, and J. K. Ellis, *J. Chem. Phys.* **135**, 124108 (2011).
- <sup>16</sup>T. Tsuchimochi and S. Ten-no, *J. Chem. Phys.* **144**, 011101 (2016).
- <sup>17</sup>R. Dutta, G. P. Chen, T. M. Henderson, and G. E. Scuseria, *J. Chem. Phys.* **154**, 114112 (2021).
- <sup>18</sup>A. T. B. Gilbert, N. A. Besley, and P. M. W. Gill, *J. Phys. Chem. A* **112**, 13164 (2008).
- <sup>19</sup>D. Hait and M. Head-Gordon, *J. Chem. Theory Comput.* **16**, 1699 (2020).
- <sup>20</sup>K. Carter-Fenk and J. M. Herbert, *J. Chem. Theory Comput.* **16**, 5067 (2020).
- <sup>21</sup>G. Levi, A. V. Ivanov, and H. Jónsson, *Faraday Discuss.* **224**, 448 (2020).
- <sup>22</sup>J. A. R. Shea and E. Neuscamman, *J. Chem. Phys.* **149**, 081101 (2018).
- <sup>23</sup>T. S. Hardikar and E. Neuscamman, *J. Chem. Phys.* **153**, 164108 (2020).
- <sup>24</sup>L. N. Tran, J. A. R. Shea, and E. Neuscamman, *J. Chem. Theory Comput.* **15**, 4790 (2019).
- <sup>25</sup>L. N. Tran and E. Neuscamman, *J. Phys. Chem. A* **124**, 8273 (2020).
- <sup>26</sup>S. C. Leasure and G. G. Balint-Kurti, *Phys. Rev. A* **31**, 2107 (1985).
- <sup>27</sup>J. Verbeek and J. H. Van Lenthe, *Int. J. Quantum Chem.* **40**, 201 (1991).
- <sup>28</sup>A. Igawa, *Int. J. Quantum Chem.* **54**, 235 (1995).
- <sup>29</sup>Y. Utsuno, N. Shimizu, T. Otsuka, and T. Abe, *Comput. Phys. Commun.* **184**, 102 (2013).

- <sup>30</sup>J. Rodríguez-Laguna, L. M. Robledo, and J. Dukelsky, *Phys. Rev. A* **101**, 012105 (2020).
- <sup>31</sup>P.-O. Löwdin, *Phys. Rev.* **97**, 1490 (1955).
- <sup>32</sup>I. Mayer, *Simple Theorems, Proofs, and Derivations in Quantum Chemistry* (Springer, 2003).
- <sup>33</sup>A. T. Amos and G. G. Hall, *Proc. R. Soc. London, Ser. A* **263**, 483 (1961).
- <sup>34</sup>G. G. Hall, *Proc. R. Soc. A* **205**, 541 (1951).
- <sup>35</sup>A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover Publications, Inc., 1989).
- <sup>36</sup>I. Shavitt and R. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, 2009).
- <sup>37</sup>J. Hendeković, M. Pavlović, and F. Sokolić, *Chem. Phys. Lett.* **77**, 382 (1981).
- <sup>38</sup>P. Ring and P. Schuck, *The Nuclear Many-Body Problem* (Springer-Verlag, 1980).
- <sup>39</sup>D. J. Thouless, *Nucl. Phys.* **21**, 225 (1960).
- <sup>40</sup>C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, and G. E. Scuseria, *Phys. Rev. A* **86**, 052102 (2012).
- <sup>41</sup>T. Tsuchimochi and S. Ten-no, *J. Chem. Theory Comput.* **12**, 1741 (2016).
- <sup>42</sup>S. R. Yost, T. Kowalczyk, and T. Van Voorhis, *J. Chem. Phys.* **139**, 174104 (2013).
- <sup>43</sup>S. R. Yost and M. Head-Gordon, *J. Chem. Phys.* **145**, 054105 (2016).
- <sup>44</sup>S. R. Yost and M. Head-Gordon, *J. Chem. Theory Comput.* **14**, 4791 (2019).
- <sup>45</sup>J. Nite and C. A. Jiménez-Hoyos, "Efficient multi-configurational wavefunction method with dynamical correlation using non-orthogonal configuration interaction singles and doubles (NOCISD)," *chemRxiv* (2019).
- <sup>46</sup>M. Head-Gordon, P. E. Maslen, and C. A. White, *J. Chem. Phys.* **108**, 616 (1998).
- <sup>47</sup>T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, 2000).
- <sup>48</sup>R. Balian and E. Brezin, *Nuovo Cimento B* **64**, 37 (1969).
- <sup>49</sup>K. J. Oosterbaan, A. F. White, and M. Head-Gordon, *J. Chem. Theory Comput.* **15**, 2966 (2019).