

The Configuration Interaction Method: Advances in Highly Correlated Approaches

C. David Sherrill[†] and Henry F. Schaefer III

*Center for Computational Quantum Chemistry
University of Georgia
Athens, Georgia 30602*

Abstract

Highly correlated configuration interaction (CI) wavefunctions going beyond the simple singles and doubles (CISD) model space can provide very reliable potential energy surfaces, describe electronic excited states, and yield benchmark energies and molecular properties for use in calibrating more approximate methods. Unfortunately, such wavefunctions are also notoriously difficult to evaluate due to their extreme computational demands. The dimension of a full CI procedure, which represents the exact solution of the electronic Schrödinger equation for a fixed one-particle basis set, grows factorially with the number of electrons and basis functions. For very large configuration spaces, the number of CI coupling coefficients becomes prohibitively large to store on disk; these coefficients must be evaluated as needed in a so-called direct CI procedure. Work done by several groups since 1980 has focused on using Slater determinants rather than spin (\hat{S}^2) eigenfunctions because coupling coefficients are easier to compute with the former. We review the fundamentals of the configuration interaction method and discuss various determinant-based CI algorithms. Additionally, we consider some applications of highly correlated CI methods.

[†]Present address: Department of Chemistry, University of California, Berkeley, CA 94720-1460

Contents

- 1 Introduction**
- 2 The Configuration Interaction Method**
 - 2.1 Fundamentals
 - 2.2 The Variational Theorem
 - 2.2.1 The Method of Linear Variations
 - 2.2.2 The Correlation Energy
 - 2.3 Matrix Elements in Terms of One- and Two-electron Integrals
 - 2.3.1 Slater's Rules
 - 2.3.2 Second Quantization
 - 2.4 Reducing the Size of the CI Space
 - 2.4.1 Truncating by Excitation Level
 - 2.4.2 Multireference Configuration Interaction
 - 2.4.3 Other CI Selection Schemes
 - 2.4.4 The First-Order Interacting Space
 - 2.4.5 Computational Scaling
 - 2.4.6 Size Extensivity Corrections
 - 2.4.7 The Frozen Core Approximation
 - 2.5 Choice of Orbitals
 - 2.6 Excited Electronic States
- 3 Common Features of Implementations**
 - 3.1 Integral Transformation
 - 3.1.1 One-electron Integrals
 - 3.1.2 Two-electron Integrals
 - 3.2 Iterative Techniques for Solving $\mathbf{Hc} = Ec$
 - 3.2.1 Davidson's Method
 - 3.2.2 Olsen's Method
- 4 Determinant-Based Algorithms for Highly Correlated CI**
 - 4.1 Slater Determinants, CSFs, and Direct CI
 - 4.2 Alpha and Beta Strings
 - 4.3 The Vectorized Full CI Algorithm of Knowles and Handy
 - 4.4 Olsen's String-Based Full CI Algorithm
 - 4.4.1 Full CI σ Equations.
 - 4.4.2 Simplifications for $M_s = 0$
 - 4.4.3 Algorithms for Computing σ
 - 4.5 Zarrabian's Reduced Intermediate Space
 - 4.6 The Table-Based Algorithm of Bendazzoli and Evangelisti
 - 4.7 Approximate Full CI Methods

- 4.8 Restricted Active Space CI
 - 4.8.1 RAS CI σ Equations
 - 4.8.2 Algorithms for Computing σ
 - 4.8.3 Beyond RAS: More Flexible *a priori* CI Space Selection
- 4.9 Implementation of Determinant-Based Algorithms
 - 4.9.1 Graphical Representation of Alpha and Beta Strings
 - 4.9.2 Nongraphical Methods for String Addressing
 - 4.9.3 Example of CI Vector and String Addressing
 - 4.9.4 String Replacement Lists
 - 4.9.5 Algorithms for σ_2 and σ_3 Used by DETCI

5 Applications of Highly Correlated CI

- 5.1 Full CI
- 5.2 Second-Order CI
- 5.3 Restricted Active Space CI
- 5.4 CISDTQ and CISD[TQ]

1 Introduction

Most chemists picture the electronic structure of atoms or molecules by invoking orbitals. The orbital concept has its basis in Hartree-Fock theory, which determines the best wavefunction $|\Psi\rangle$ under the approximation that each electron experiences only the *average* field of the other electrons. This is also called the “one-electron,” or “independent particle” model. While the Hartree-Fock method gives very useful results in many situations, it is not always quantitatively or even qualitatively correct. When this approximation fails, it becomes necessary to include the effects of electron correlation: one must model the *instantaneous* electron-electron repulsions present in the molecular Hamiltonian.

The most broadly applicable method for describing electron correlation is configuration interaction (CI), which expresses the wavefunction as a linear combination of \hat{S}_z eigenfunctions (Slater determinants) or \hat{S}_z and \hat{S}^2 eigenfunctions (configuration state functions, or CSFs) describing the distribution of N electrons. If all possible N -electron functions are included in the CI procedure (subject to spatial and spin symmetry restrictions), then the Schrödinger equation is solved exactly within the space spanned by the one-particle basis functions. Hence, in its most general form, CI applies to difficult cases such as excited states, open-shell systems, and systems far from their equilibrium geometries. However, the dimension of this “full CI” procedure grows factorially with molecular size, so it is necessary to select only the most important N -electron functions.

A common approach is to restrict the CI space to the Hartree-Fock self-consistent-field (SCF) configuration and those configurations related to it by single and double substitutions of orbitals, in a procedure denoted CISD. In cases where the SCF method yields a good approximate wavefunction, CISD with double- ζ plus polarization (DZP) single-particle basis sets typically predicts equilibrium bond lengths of small molecules within 0.4% of experiment and harmonic vibrational frequencies within 4%.¹ Unfortunately, CISD (and most other standard CI methods short of full CI) are not size extensive, meaning that their performance degrades with increasing molecular size. Size extensive alternatives include many-body perturbation theory (MBPT) and coupled-cluster (CC) methods. The coupled-cluster singles and doubles (CCSD) method outperforms CISD with only a moderately increased computational effort (the cost of both methods scales as the sixth power of the system size) because CCSD accounts for some triple and quadruple substitutions from the SCF configuration by approximating them as products of single and double substitutions. When employed with a triple- ζ plus double polarization basis set (TZ2P), the CCSD generally predicts bond lengths within 0.2% of experiment and harmonic vibrational frequencies within 2%.²

More accurate results can be obtained by using larger one-particle basis sets and employing the CCSD(T) method, which accounts for the effects of irreducible (or connected) triple substitutions in a single non-iterative step scaling as n^7 .³⁻⁵ Furthermore, recent equation-of-motion (EOM) or linear-response coupled-cluster theories for singly-excited electronic states^{6,7} also outperform CISD in the prediction of excitation energies because CISD is biased towards the state described by the reference wavefunction.

Even the coupled-cluster methods eventually break down in cases where the SCF wavefunction is not a qualitatively correct description of the system. This can occur during bond-breaking reactions, for example, or for transition metals.^{8,9} Hence, it is necessary to make the zeroth-order wavefunction multiconfigurational. Although multireference coupled-cluster theories are very difficult to formulate, multireference CI methods have been used for many years.^{10,11} These methods typically include single and double substitutions from a set of "reference" configurations required to describe nondynamical electron correlation. Unfortunately, reference selection is not trivial, since the list of important references depends on the molecule and its geometry. This tends to make MR-CI methods unsuitable as a "model chemistry,"¹² since the quality of the wavefunction is not uniform across different molecules. One MR-CI wavefunction which is largely free from these difficulties is second-order CI (SOCI),¹³ which is a multi-reference CISD in which the references are chosen as all possible distributions of electrons within a given "active space" (unfortunately, the acronym SOCI is also sometimes used to mean spin-orbit CI). The SOCI wavefunction requires much less computational effort than a full CI, yet it produces potential energy surfaces which nearly parallel the full CI surfaces.¹⁴⁻¹⁷ If the active space is large enough, one can expect the SOCI method to provide equally good results for any small molecule at any geometry, making it a suitable model chemistry (SOCi is still not rigorously size extensive, so it may be necessary to apply size extensivity corrections for systems with eight electrons or more¹⁵). Unfortunately, the SOCI method is too computationally expensive to be generally applicable.

Hence there is a need to make SOCI more computationally efficient so that it can be used for larger chemical systems, and to develop related methods which scale better with the system size. Although the CI space can be reduced by individual selection of references or N -electron functions, for the reasons stated above it is beneficial to select the CI space in an *a priori* manner, once a minimal set of parameters, such as the active space, has been specified. For example, we have advocated a method we call CISD[TQ],¹⁶⁻¹⁸ which is a SOCI in which higher-than-quadruple substitutions have been excluded. For systems dominated by a single reference, CISD[TQ] performs nearly as well as SOCI.^{16,17}

Although extremely costly to determine, full CI results are invaluable in the calibration of such multireference CI methods, or indeed of essentially any *ab initio* electronic structure method, including many-body perturbation theory (MBPT)¹⁹⁻²³ and coupled-cluster approaches.^{4, 6, 7, 17, 19, 22, 24-37} A few full CI benchmarks^{17, 38, 39} have been carried out using the loop-driven graphical unitary group (LD-GUGA) CI approach,⁴⁰⁻⁴³ which uses a spin eigenfunction, or CSF, basis. However, a majority of the full CI calculations to date have employed Slater determinants, even though this typically makes the CI vector 2-4 times longer. In 1980, Handy demonstrated that the benefits of Slater determinants can outweigh their disadvantages, primarily because the evaluation of the required matrix elements becomes so much simpler.⁴⁴ In his 1980 article, Handy introduced the alpha and beta string notation which has commonly been used in the development of new CI algorithms. After an important reformulation of the direct CI method by Siegbahn,⁴⁵ Knowles and Handy introduced a vectorized full CI algorithm that enabled a whole series of full CI benchmark studies by Bauschlicher, Langhoff, Taylor, and others.¹⁵ In 1988, Olsen and co-workers showed how to improve the Knowles-Handy algorithm by reducing the operation count while still maintaining vectorization in the innermost loops.⁴⁶ Another important advance was the extension to certain types of truncated CI spaces in which determinants are chosen according to how many electrons they place in each of three orbital subspaces. This restricted active space (RAS) CI procedure is capable of evaluating SOCI and CISD[TQ] wavefunctions. Subsequently, other full CI algorithms involving basically the same amount of computational effort as Olsen's algorithm have been presented by Harrison and Zarrabian⁴⁷ and by Bendazzoli and Evangelisti.⁴⁸⁻⁵⁰

In this article, we provide an updated look at the configuration interaction method in general, and at highly correlated CI methods in particular. Special emphasis is given to methods which select the CI space in an *a priori* manner. After reviewing the basic theory of the CI method and the typical approximations employed, we discuss features common to all implementations: transformation of the one- and two-electron integrals from the atomic orbital to the molecular orbital basis, and iterative diagonalization methods. Next, we survey several determinant-based algorithms for full and RAS CI wavefunctions. We describe some technical issues in considerable detail and describe our experience with our own determinant-based CI program. Finally, we discuss some of the applications of highly correlated CI methods. Although crucial to the efficient determination of optimized geometries and vibrational frequencies, analytic gradients of CI wavefunctions are not discussed in this article; we refer the reader to the recent article by Shepard⁵¹ and the monograph by Yamaguchi *et al.*⁵² Furthermore, additional considerations may arise when designing a CI program to be used along with orbital optimization in the mul-

ticonfigurational (MC) or complete-active-space (CAS) SCF methods; these issues are discussed elsewhere.^{9,53-56}

2 The Configuration Interaction Method

This section presents the essential elements of the configuration interaction method and is meant to be accessible to those who are not experts in CI. The classic review by Shavitt covers the theoretical fundamentals and various formulations given prior to 1977.⁵⁷ More recent reviews have been presented by Siegbahn,⁵⁸ Karwowski,⁵⁹ and Duch.⁶⁰

2.1 Fundamentals

Configuration interaction is conceptually the simplest method for solving the time-independent electronic Schrödinger equation $\hat{H}|\Psi\rangle = E|\Psi\rangle$ under the Born-Oppenheimer approximation. The electronic wavefunction $|\Psi\rangle$ is approximated by a linear expansion of N -electron basis functions (where N is the number of electrons in the system), i.e.,

$$|\Psi\rangle = \sum_I c_I |\Phi_I\rangle. \quad (1)$$

The linear expansion coefficients c_I are the *CI coefficients*. Substituting this linear expansion into the electronic Schrödinger equation, one obtains⁶¹ a matrix form more suitable for computation:

$$\mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c}, \quad (2)$$

where the Hamiltonian operator \hat{H} has been replaced by a matrix \mathbf{H} and the CI wavefunction $|\Psi\rangle$ has been replaced by a column vector of coefficients \mathbf{c} . In principle, this “matrix mechanics” formulation is equivalent to the original electronic Schrödinger equation;⁶² hence it is said that CI constitutes an “exact theory.” In practice, however, the matrix equations are not exact because the expansion in equation (1) must be truncated to a finite number of terms. The matrix elements of the Hamiltonian are given by $H_{IJ} = \langle\Phi_I|\hat{H}|\Phi_J\rangle$ and \mathbf{S} is the overlap matrix with elements $S_{IJ} = \langle\Phi_I|\Phi_J\rangle$. If orthonormal functions $|\Phi_I\rangle$ are used for the expansion, then of course \mathbf{S} becomes the unit matrix and the equation becomes an eigenvalue equation. Since \mathbf{H} is a Hermitian matrix, the number of orthogonal eigenvectors is equal to the dimension of the matrix. The lowest-energy solution represents the electronic ground state, and higher-energy solutions represent excited electronic states.

It is generally helpful to build into the expansion functions $\{|\Phi_I\rangle\}$ the symmetry properties of the system. According to the antisymmetry principle,

a wavefunction describing a system of electrons (more generally, fermions) must be antisymmetric with respect to the interchange of spatial and spin coordinates for any pair of electrons. This requirement is very commonly satisfied by making the expansion functions Slater determinants. A Slater determinant in which the one-electron functions $\phi_i, \phi_j, \dots, \phi_k$ are occupied may be written

$$\Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_i(\mathbf{x}_1) & \phi_j(\mathbf{x}_1) & \dots & \phi_k(\mathbf{x}_1) \\ \phi_i(\mathbf{x}_2) & \phi_j(\mathbf{x}_2) & \dots & \phi_k(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \phi_i(\mathbf{x}_N) & \phi_j(\mathbf{x}_N) & \dots & \phi_k(\mathbf{x}_N) \end{vmatrix} \quad (3)$$

and abbreviated as $|\phi_i \phi_j \dots \phi_k\rangle$ or simply as $|ij \dots k\rangle$. Note that this determinant is uniquely specified (up to a phase factor) by the list of occupied orbitals. It is easy to see that such a determinant satisfies the antisymmetry principle, since the interchange of coordinates for a pair of electrons translates to the swapping of rows of the determinant, which introduces a sign change. We may also write the above determinant as

$$\Phi = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \hat{P} \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_2) \dots \phi_k(\mathbf{x}_N), \quad (4)$$

where \hat{P} is a permutation of electron coordinates with sign $(-1)^P$. In this context it can be useful to define the *antisymmetrizer* as

$$\mathcal{A} = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \hat{P}. \quad (5)$$

This operator produces a Slater determinant when applied to a simple product of spin orbitals. The antisymmetrizer is Hermitian, it commutes with \hat{H} , and its square is proportional to itself, i.e., $\mathcal{A}^2 = \sqrt{N!} \mathcal{A}$.

An electronic wavefunction can be described exactly by equation (1) if the expansion includes all possible Slater determinants formed from a complete set of one-electron functions $\{\phi\}$.⁶³ Such a procedure has been called *complete CI*.⁴⁰ Since a truly complete set of orbitals will typically be infinite, a complete CI is technically impossible to perform. However, if the one-electron basis set is truncated, then only a finite (albeit large) number of Slater determinants can be formed. Using all of these determinants in the expansion constitutes a *full CI* procedure, and the resulting eigenfunctions and eigenvalues are exact within the space spanned by the one-electron basis set. Although full CI results are extremely costly to compute, they are essential for benchmarking more approximate methods.

It is straightforward to show that if the exact wavefunction $|\Psi\rangle$ is an eigenfunction of some Hermitian operator \hat{A} , then expansion functions $|\Phi_I\rangle$ which are eigenfunctions of \hat{A} with different eigenvalues do not contribute to the CI wavefunction and can be neglected in the expansion (1). If the Slater determinants are formed from spin-orbitals which are eigenfunctions of \hat{s}_z and spatial symmetry operators, then the Slater determinants themselves will also be eigenfunctions of these spatial symmetry operators and of \hat{S}_z . However, a Slater determinant is not generally an eigenfunction of \hat{S}^2 . Hence, a common alternative to Slater determinants are *configuration state functions* (CSFs), which are simply linear combinations of Slater determinants chosen to be eigenfunctions of \hat{S}^2 . The benefit of using CSFs over determinants is that fewer N -electron functions are needed to describe the same state. The drawback is that matrix elements of the Hamiltonian are easier to compute using determinants. Of course, there are other possible choices for the N -electron basis functions. For instance, one can incorporate functions of two electrons (geminals),^{64–67} as is done in the Hylleraas treatment of the helium atom.^{68,69} Nevertheless, N -electron functions built from single-particle functions remain the most common.

Unfortunately, even with an incomplete one-electron basis, a full CI is computationally intractable for any but the smallest systems, due to the vast number of N -electron basis functions required (the size of the CI space is discussed in section 2.4.1). The CI space must be reduced, hopefully in such a way that the approximate CI wavefunction and energy are as close as possible to the exact values. By far the most common approximation is to begin with the Hartree-Fock procedure, which determines the best single-configuration approximation to the wavefunction that can be formed from a given basis set of one-electron orbitals (usually atom centered and hence called atomic orbitals, or AOs). This yields a set of molecular orbitals (MOs) which are linear combinations of the AOs:

$$\phi_i(\mathbf{x}_1) = \sum_{\mu} C_{\mu}^i \chi_{\mu}(\mathbf{x}_1), \quad (6)$$

where $\chi_{\mu}(\mathbf{x}_1)$ denotes an atomic orbital and C_{μ}^i is an *SCF coefficient*. The CI space can then be expanded according to substitution or “excitation” level relative to the SCF “reference” determinant, i.e.,

$$|\Psi\rangle = c_0|\Phi_0\rangle + \sum_{ia} c_i^a|\Phi_1^a\rangle + \sum_{a<b, i<j} c_{ij}^{ab}|\Phi_2^{ab}\rangle + \sum_{a<b<c, i<j<k} c_{ijk}^{abc}|\Phi_3^{abc}\rangle + \dots \quad (7)$$

where $|\Phi_1^a\rangle$ means the Slater determinant formed by replacing spin-orbital i in $|\Phi_0\rangle$ with spin orbital a , etc. The widely-employed CI singles and doubles (CISD) wavefunction includes only those N -electron basis functions which represent single or double substitutions relative to the reference state. Since the

Hamiltonian operator includes only one- and two-electron terms, only singly and doubly substituted configurations can interact directly with the reference, and they typically account for about 95% of the basis set correlation energy of small molecules at their equilibrium geometries,³⁸ where $|\Phi_0\rangle$ provides a good zeroth-order description. Truncation of the CI space according to excitation class is discussed more fully in section 2.4.1.

2.2 The Variational Theorem

One attractive feature of configuration interaction is that the computed lowest energy eigenvalue is always an upper bound to the exact ground state energy. This follows from the fact that the CI energy is given by the expectation value formula, or Rayleigh quotient,

$$E = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle}. \quad (8)$$

The variational theorem may be proven by expressing the approximate wavefunction $|\Phi\rangle$ as a linear combination of the exact eigenvectors $|\Psi_i\rangle$; one easily obtains

$$E - \mathcal{E}_0 = \sum_i c_i^* c_i (\mathcal{E}_i - \mathcal{E}_0), \quad (9)$$

where \mathcal{E}_i is the i th *exact* energy eigenvalue, i.e., $\hat{H}|\Psi_i\rangle = \mathcal{E}_i|\Psi_i\rangle$. Since the right-hand side of eq. (9) is necessarily non-negative, $E \geq \mathcal{E}_0$. Likewise, we can also insert an expansion over the exact eigenvectors *for a given one-electron space* to prove that the CI energy must be an upper bound to the full CI energy in the same one-electron basis set. Equation (9) demonstrates that the approximate wavefunction approaches the exact one ($c_0 \approx 1$) as the energy E is minimized (see section 2.2.1). Minimizing E is equivalent to minimizing the right-hand side of eq. (9); that is, the sum of squares of the absolute values of the coefficients of excited states is minimized with weight factors $(\mathcal{E}_i - \mathcal{E}_0)$. This means that other properties do not generally converge as quickly with CI space expansion as the energy. In fact, the error in the energy is quadratic in the wavefunction error. This can be shown by writing the energy as

$$E = \frac{\langle \Psi - \Delta\Psi | \hat{H} | \Psi - \Delta\Psi \rangle}{\langle \Psi - \Delta\Psi | \Psi - \Delta\Psi \rangle}, \quad (10)$$

with $|\Psi\rangle$ the exact wavefunction and the error $|\Delta\Psi\rangle$ chosen orthogonal to $|\Psi\rangle$. From this expression it is simple to demonstrate that all terms linear in $|\Delta\Psi\rangle$ are vanishing and that only quadratic terms remain.

It is easy to extend proofs of the variational theorem to the case of states which are the lowest roots of a given spatial and spin symmetry.⁷⁰ Since the

self-consistent-field (SCF) and multiconfigurational SCF (MCSCF) wavefunctions can be written as a linear expansion (1) containing one or a few Slater determinants, with an energy given by eq. (8), they also obey the variational theorem. Furthermore, just as the lowest CI eigenvalue is an upper bound to the exact ground-state energy, more generally any CI eigenvalue E_i is an upper bound to the corresponding exact excited state energy \mathcal{E}_i .⁷¹ Additionally, as other N -electron basis functions are added to the CI space, the eigenvalues obey the MacDonald-Hylleraas-Undheim relations^{71,72}

$$E_{i-1}^{(m)} \leq E_i^{(m+1)} \leq E_i^{(m)} \quad (11)$$

where m is the number of N -electron basis functions.

2.2.1 The Method of Linear Variations

Since the variational theorem proves that the energy of a CI wavefunction is always an upper bound to the exact energy, one might start simply from the linear expansion (1) and attempt to minimize the energy by varying the CI coefficients subject to the constraint that they remain normalized. It is easy to show⁶³ that this method of linear variations, or the Ritz method,⁷³ yields the matrix equation

$$\mathbf{Hc} = E\mathbf{Sc}. \quad (12)$$

That is, the method of linear variations is identical to the matrix formulation of the Schrödinger equation. Another way of viewing this result is that only solutions to eq. (12) are energetically stable with respect to variations in the linear expansion coefficients.

2.2.2 The Correlation Energy

Since the CI energy is always an upper bound to the exact energy, approximate CI methods can be judged according to what fraction of the correlation energy they recover. The correlation energy is defined as the difference between the energy in the Hartree-Fock limit (E_{HF}) and the exact nonrelativistic energy of a system (\mathcal{E}_0)

$$E_{\text{corr}} = \mathcal{E}_0 - E_{HF}. \quad (13)$$

This energy will always be negative because the Hartree-Fock energy is an upper bound to the exact energy. The exact nonrelativistic energy \mathcal{E}_0 could be calculated, in principle, via a full CI in a complete one-electron basis. If we have an incomplete one-electron basis set, then we can only compute the *basis set correlation energy*, which is the correlation energy for a given one-electron basis. Frequently the term correlation energy implies basis set correlation

energy. The correlation energy is the energy recovered by fully allowing the electrons to avoid each other; the Hartree-Fock method, rather than using the true instantaneous Coulomb repulsion between pairs of electrons, instead only allows each electron to experience an average potential due to all the other electrons. However, this description of *dynamical correlation* is not complete. When a molecule is pulled apart, the electrons should not need to avoid each other as much, so the magnitude of the correlation energy should decrease. In fact, the opposite is true, as shown by the basis set correlation energies given in Table 1 for H₂O at five different geometries.

Table 1: Correlation Energy in H₂O with a cc-pVDZ Basis as Both O–H Bonds are Stretched Simultaneously.

Geometry	E_{corr} (hartree) ^a
R_e	-0.217 821
$1.5 \cdot R_e$	-0.269 961
$2.0 \cdot R_e$	-0.363 954
$2.5 \cdot R_e$	-0.476 747
$3.0 \cdot R_e$	-0.567 554

^aData from Olsen *et al.*, ref 22.

All ten electrons are correlated.

The magnitude of the correlation energy increases as the O–H bonds are stretched beyond their equilibrium length because equation (13) also includes a more subtle effect called the *nondynamical* or *static* correlation energy. This part of the correlation energy reflects the inadequacy of a single reference in describing a given molecular state, and is due to nearly degenerate states or rearrangement of electrons within partially filled shells. Shavitt⁵⁷ has pointed out this deficiency in the correlation energy definition and has suggested that multiconfigurational Hartree-Fock may prove a more useful baseline than single-configuration Hartree-Fock in equation (13).

2.3 Matrix Elements in Terms of One- and Two-electron Integrals

2.3.1 Slater's Rules

The matrix elements $H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle$ can be expressed in terms of one- and two-electron integrals. If we employ Slater determinants, the matrix elements

may be evaluated using Slater's rules (also called the Slater-Condon rules)⁷⁴⁻⁷⁶ if a common set of one-electron orbitals are used for all determinants and if these orbitals are orthonormal. If nonorthogonal orbitals are employed (e.g., atomic orbitals) then the more complicated Löwdin rules⁷⁷ apply.

Slater's rules are expressed here in terms of spin-orbitals, which are functions of the spatial and spin coordinates of a single electron. The one-electron integrals are written as

$$[i|\hat{h}|j] = \int \phi_i^*(\mathbf{x}_1)\hat{h}(\mathbf{x}_1)\phi_j(\mathbf{x}_1)d\mathbf{x}_1 \quad (14)$$

and the two-electron integrals, in Mulliken notation, are written as

$$[ij|kl] = \int \phi_i^*(\mathbf{x}_1)\phi_j(\mathbf{x}_1)\frac{1}{r_{12}}\phi_k(\mathbf{x}_2)^*\phi_l(\mathbf{x}_2)d\mathbf{x}_1d\mathbf{x}_2. \quad (15)$$

Before Slater's rules can be used, the two Slater determinants must be rearranged so that they have the maximum possible number of columns in common (recalling that each column swap causes a sign change). After the determinants are in maximum coincidence, we see how many spin orbitals they differ by and employ the following rules:

1. Identical Determinants:

$$\langle \Phi_1 | \hat{H} | \Phi_1 \rangle = \sum_m^N [i|\hat{h}|i] + \sum_{i>j}^N \{ [ii|jj] - [ij|ji] \}. \quad (16)$$

2. Determinants that Differ by One Spin Orbital:

$$\begin{aligned} |\Phi_1\rangle &= |\dots i \dots\rangle \\ |\Phi_2\rangle &= |\dots j \dots\rangle \\ \langle \Phi_1 | \hat{H} | \Phi_2 \rangle &= [i|\hat{h}|j] + \sum_k^N \{ [ik|kk] - [ik|kj] \}. \end{aligned} \quad (17)$$

3. Determinants that Differ by Two Spin Orbitals:

$$\begin{aligned} |\Phi_1\rangle &= |\dots ij \dots\rangle \\ |\Phi_2\rangle &= |\dots kl \dots\rangle \\ \langle \Phi_1 | \hat{H} | \Phi_2 \rangle &= [ik|jl] - [il|jk]. \end{aligned} \quad (18)$$

Some of the terms above may vanish after integrating over spin coordinates, and a pair of determinants differing by more than two spin orbitals have a matrix element of zero. A derivation of these rules can be found in the introductory text by Szabo and Ostlund.⁶³ The rules for evaluating Hamiltonian matrix elements in a CSF basis are more complicated and are generally derived^{40,42,78} using second quantization, which we consider next.

2.3.2 Second Quantization

We are free to write the matrix elements in the more general form

$$H_{IJ} = \sum_{pq}^n \gamma_{pq}^{IJ} (p|\hat{h}|q) + \frac{1}{2} \sum_{pqrs}^n \Gamma_{pqrs}^{IJ} (pq|rs), \quad (19)$$

where the use of parentheses rather than square brackets denotes a switch to spatial orbital notation rather than spin orbital notation. Also note the factor of $1/2$ in the two-electron term. The constants γ_{pq}^{IJ} and Γ_{pqrs}^{IJ} are called the one- and two-electron coupling coefficients, respectively. The CI energy in terms of these coupling coefficients is

$$E = \sum_{IJ}^{CI} c_I \left[\sum_{pq}^n \gamma_{pq}^{IJ} (p|\hat{h}|q) + \frac{1}{2} \sum_{pqrs}^n \Gamma_{pqrs}^{IJ} (pq|rs) \right] c_J, \quad (20)$$

where we assume that the CI coefficients are real. The one- and two-electron reduced density matrices are defined as

$$\gamma_{pq} = \sum_{IJ}^{CI} c_I c_J \gamma_{pq}^{IJ}, \quad (21)$$

$$\Gamma_{pqrs} = \sum_{IJ}^{CI} c_I c_J \Gamma_{pqrs}^{IJ}, \quad (22)$$

and using these definitions the energy may be written more compactly as

$$E = \sum_{pq}^n \gamma_{pq} (p|\hat{h}|q) + \frac{1}{2} \sum_{pqrs}^n \Gamma_{pqrs} (pq|rs). \quad (23)$$

Some authors absorb the factor of $1/2$ into the definition of the two-electron coupling coefficient and reduced density matrix.

These coupling coefficients are generally derived using second quantization,^{63,79} in which the Hamiltonian is written (for a given one-particle basis set) as

$$\hat{H} = \sum_{pq}^{2n} a_p^\dagger a_q [p|\hat{h}|q] + \frac{1}{2} \sum_{pqrs}^{2n} a_p^\dagger a_r^\dagger a_s a_q [pq|rs], \quad (24)$$

where a_p^\dagger and a_p are the creation and annihilation operators, respectively, for an electron in spin orbital p . Note that the second-quantized form of the Hamiltonian is independent of the number of electrons. If the spatial parts of α and β spin orbitals are identical, it is easy to re-write the second-quantized Hamiltonian in terms of the following shift operators, which Paldus has shown⁴⁰ to be generators of the unitary group:

$$\hat{E}_{ij} = a_{i\alpha}^\dagger a_{j\alpha} + a_{i\beta}^\dagger a_{j\beta}. \quad (25)$$

Due to the anticommutation relations of creation and annihilation operators,

$$\hat{E}_{ij}^\dagger = \hat{E}_{ji} \quad (26)$$

$$[\hat{E}_{ij}, \hat{E}_{kl}] = \hat{E}_{il}\delta_{jk} - \hat{E}_{kj}\delta_{il}. \quad (27)$$

The resulting Hamiltonian in terms of these operators is

$$\hat{H} = \sum_{pq}^n h_{pq} \hat{E}_{pq} + \frac{1}{2} \sum_{pqrs}^n (pq|rs) (\hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps}), \quad (28)$$

where we have used the more compact notation $h_{pq} = (p|\hat{h}|q)$. It is clear that the one- and two-electron coupling coefficients can be written as

$$\gamma_{pq}^{IJ} = \langle \Phi_I | \hat{E}_{pq} | \Phi_J \rangle, \quad (29)$$

$$\Gamma_{pqrs}^{IJ} = \langle \Phi_I | \hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps} | \Phi_J \rangle. \quad (30)$$

Furthermore, using equations (26) and (27), one can deduce the following:

$$\gamma_{pq}^{IJ} = (\gamma_{qp}^{JI})^* \quad (31)$$

$$\Gamma_{pqrs}^{IJ} = \Gamma_{rspq}^{IJ} = (\Gamma_{srqp}^{JI})^* = (\Gamma_{qpsr}^{JI})^*. \quad (32)$$

2.4 Reducing the Size of the CI Space

This section discusses strategies for reducing the number of N -electron basis functions in the CI space (given that, in the general case, it is impossible to include all of them). We have already discussed how N -electron functions with the wrong symmetry properties (e.g., point-group symmetry, or spin symmetry) can be dismissed immediately.

2.4.1 Truncating by Excitation Level

As noted in equation (7), the CI expansion is typically truncated according to excitation level; in the vast majority of CI studies, the expansion is truncated (for computational tractability) at doubly-substituted configurations. Since the Hamiltonian contains only two-body terms, only singles and doubles can interact directly with the reference; this is a direct result of Slater's rules (cf. section 2.3.1). Furthermore, the matrix elements of singly substituted determinants (or CSFs) with the reference are zero when canonical SCF orbitals are used, according to Brillouin's theorem. Hence, one expects double excitations to make the largest contributions to the CI wavefunction after the reference

state. Indeed, this is what is observed. Even though singles, triples, etc., do not interact directly with the reference, they can still become part of the CI wavefunction (i.e., have non-zero coefficients) because they mix with the doubles, directly or indirectly. Although singles are much less important to the energy than doubles, they are generally included in CI treatments because of their relatively small number and because of their greater importance in describing one-electron properties.

After singles and doubles, the most important determinants are triples and quadruples, because only these can interact directly with the doubles. The importance of a determinant to the final CI wavefunction is expected to fall off with increasing substitution or excitation level relative to the reference, assuming that the reference is a reasonable zeroth-order description of the desired electronic state. Table 2 demonstrates the importance of various excitation classes in obtaining CI energies. Singles and doubles account for 95% of the correlation energy at the equilibrium geometries of the molecules listed. Quadruple excitations are more important than triples, at least as far as the energy is concerned. At stretched geometries, the CISD and CISDT methods become markedly poorer, yet the CISDTQ method still recovers a very high (and nearly constant) fraction of the correlation energy, suggesting that CISDTQ should give reliable results for energy differences across potential energy surfaces for small molecules so long as no more than two bonds are broken at once (simultaneously breaking three bonds would require up to sextuple substitutions).

Table 3 demonstrates that the number of N -electron basis functions increases dramatically with increasing excitation level. A DZP basis should be considered the minimum adequate basis for a meaningful benchmark study.^{15,81} While it is generally possible to perform CISD calculations on small molecules with a good one-electron basis, the CISDTQ method is limited to molecules containing very few heavy atoms, due to the extreme number of N -electron functions required. Full CI calculations are of course even more difficult to perform, so that despite their importance as benchmarks, few full CI energies using large one-electron basis sets have been obtained.

The size of the full CI space in CSFs can be calculated (including spin symmetry but ignoring spatial symmetry) by Weyl's dimension formula.⁸² If N is the number of electrons, n is the number of orbitals, and S is the total spin, then the dimension of the CI space in CSFs is given by

$$D_{nNS} = \frac{2S+1}{n+1} \binom{n+1}{N/2-S} \binom{n+1}{N/2+S+1}. \quad (33)$$

The dimension of the full CI space in determinants (again, ignoring spatial

Table 2: Percentage of correlation energy recovered by various CI excitation levels for some small molecules.

Molecule	Percent Corr. Energy ^a		
	CISD	CISDT	CISDTQ
BH	94.91	n/a	99.97
HF	95.41	96.49	99.86
H ₇ ⁺	96.36	96.87	99.96
H ₂ O(R _e)	94.48	95.85	99.85
H ₂ O(1.5 R _e)	89.36	92.05	99.48
H ₂ O(2.0 R _e)	80.21	84.59	98.40
NH ₃	94.44	95.43	99.84

^aResults are for a DZP basis and are taken from refs 38 (BH, HF), 17 (H₇⁺), 22 (H₂O), and 80 (NH₃). H₂O results correlate all ten electrons and employ the cc-pVDZ basis.

Table 3: Number of CSFs required for small molecules at several levels of CI.

Molecule	CSFs required ^a			
	CISD	CISDT	CISDTQ	FCI
BH	568	n/a	28 698	132 686
HF	552	6 712	48 963	944 348
H ₇ ⁺	1 271	24 468	248 149	2 923 933
H ₂ O	1 311	27 026	332 491	94 165 610
NH ₃	2 443	52 595	619 235	48 642 057

^aResults are for a DZP basis and are taken from refs 38 (BH, HF), 17 (H₇⁺), 22 (H₂O), and 80 (NH₃). H₂O results correlate all ten electrons and employ the cc-pVDZ basis.

Table 4: Dimension of Full CI in Determinants (CSFs in parentheses)

Orbitals	Number of electrons			
	6	8	10	12
10	14.4×10^3 (4.95×10^3)	44.1×10^3 (13.9×10^3)	63.5×10^3 (19.4×10^3)	44.1×10^3 (13.9×10^3)
20	1.30×10^6 (379×10^3)	23.5×10^6 (5.80×10^6)	240×10^6 (52.6×10^6)	1.50×10^9 (300×10^6)
30	16.5×10^6 (4.56×10^6)	751×10^6 (172×10^6)	20.3×10^9 (4.04×10^9)	353×10^9 (62.5×10^9)

symmetry) is computed simply by

$$D_{nN_\alpha N_\beta} = \binom{n}{N_\alpha} \binom{n}{N_\beta}, \quad (34)$$

or, in a form closer to equation (33),

$$D_{nNS} = \binom{n}{N/2 + S} \binom{n}{N/2 - S}. \quad (35)$$

Table 4 shows the dimension of the full CI space (neglecting spatial symmetry) in determinants and in CSFs for closed-shell systems. Current full CI algorithms are typically limited to several million determinants. Although there have been reports of larger calculations (including more than a billion determinants^{83,84}), the computational expense is currently too great for routine calculations of this size.

2.4.2 Multireference Configuration Interaction

A full CI wavefunction is invariant to orbital rotations and even to the choice of the reference function. By contrast, the simple CISD method is quite sensitive to the choice of reference and orbitals. This explains the poor performance of CISD when the bonds are stretched in H₂O (cf. Table 2): the SCF wavefunction becomes an inadequate reference at stretched geometries, and CISD is unable to overcome this inadequacy. Such difficulties can occur even at equilibrium geometries if multiple low-lying electronic states are present. For

example, the zeroth-order wavefunction for a singlet diradical often requires two electron configurations: one doubly occupies the MO formed from in-phase radical orbitals, while the other doubly occupies the out-of-phase MO. Another example is the \bar{c} state of CH₂, which also requires a two-configuration treatment: the two configurations correspond to the two choices for the lone pair of electrons, either in the molecular plane or perpendicular to it.^{85,86} More than two configurations can be critical for transition metals or when multiple bonds are broken.

If a CISD procedure includes all the important N -electron functions from the zeroth-order wavefunction (the “references”) and also the single and double substitutions for *each* of these references, then the resulting method is referred to as multireference (MR) CISD. The MR-CISD wavefunction may be written

$$|\Phi_{MRCI}\rangle = \sum_R c(R)|\Phi(R)\rangle + \sum_R \sum_{ix} c_i^x(R)|\Phi_i^x(R)\rangle + \sum_R \sum_{ijxy} c_{ij}^{xy}(R)|\Phi_{ij}^{xy}(R)\rangle, \quad (36)$$

where R denotes a reference function, and i, j (x, y) run over orbitals which are occupied (unoccupied) for a given R . Clearly, a determinant or CSF which is generated as a single or double substitution from one reference state might also be generated as a single or double from a different reference; only unique N -electron functions are included in the MR-CISD procedure. If a sufficient number of references are included, then a MR-CISD can provide results nearly as good as the full CI^{14,15,87} at a dramatically reduced computational expense. In the MR-CI method, the set of orbitals which are occupied in any of the references constitutes the *internal space*, and all other orbitals are in the *external space*. Sometimes a further distinction is made among the internal space orbitals: those whose occupancy is constant for all references are called *inactive* (even though their electrons may be excited in the final wavefunction), and the rest are called *active*. In the direct CI method⁸⁸ (see section 4), it is more convenient to rewrite (36) in an equivalent form which emphasizes the number of external orbitals:

$$|\Phi_{MRCI}\rangle = \sum_I c_I |\Phi_I\rangle + \sum_S \sum_a c_S^a |\Phi_S^a\rangle + \sum_P \sum_{ab} c_P^{ab} |\Phi_P^{ab}\rangle, \quad (37)$$

where a and b are external orbitals, and I , S , and P denote internal states (including spin coupling) with N , $N - 1$, and $N - 2$ electrons, respectively.

One very straightforward, *a priori* selection scheme is to make references of all those N -electron functions which can be obtained by distributing electrons in all possible ways in a subset of the most important orbitals (the “active space”). This results in the second-order CI (SOC),¹³ which is known^{16,17,89} to provide high-quality potential energy surfaces nearly parallel to those from a full CI. Unfortunately, this prescription typically produces too many references

and the final CI space is too large to be computationally tractable. One strategy which has received relatively little attention^{17,46,90} is to approximate a SOCI by restricting the references according to their excitation level. It is often reasonable to assume that the most important references are single and double substitutions from a dominant single reference.^{16,17,91} Making this restriction leads to a MR-CISD which includes those triples and quadruples which have no more than two electrons outside the active space. This wavefunction has been designated CISD[TQ] to emphasize the variational treatment of limited triples and quadruples, and it has been shown to closely match SOCI when a single reference configuration dominates.^{16,17} Although the CISD[TQ] expansion is much smaller than SOCI, it remains intractable for systems with more than two or three heavy atoms. Further strategies to reduce the cost of a CISD[TQ] are discussed later.

A much more common procedure for reference selection is to accept references whose estimated importance is greater than some given threshold; this can involve perturbative estimates of a function's energetic contribution or its coefficient in some preliminary wavefunction. These approaches are more successful at obtaining the best wavefunction at the lowest expense, but they sacrifice the simplicity of the excitation class selection and can become more difficult to implement and to use. One complication is that potential energy surfaces determined using such methods may not be smooth; to alleviate this, one may need to determine the important references at each geometry and use the union of these sets at every point.

Discarding some of the single and double excitations is another way to reduce the CI space. As with reference selection, the most common approaches involve estimates of a function's energetic contribution or coefficient. The CIPSI method of Malrieu and co-workers selects determinants based on perturbation theory estimates of their coefficients in the first-order wavefunction.^{92,93} Alternatively, Buenker and Peyerimhoff¹⁰ select spatial orbital configurations on the basis of each configuration's energetic contribution to a small CI consisting only of the references and the CSFs formed from that configuration. Obviously this involves solving a very large number of small CI problems. Alternatively, one can estimate the importance of all configurations simultaneously via a procedure such as Gershgorin and Shavitt's B_k method.^{94,95} Shavitt's 1977 review article⁵⁷ surveys these and related alternatives.

Finally, Siegbahn has suggested two procedures for reducing the number of variational parameters in a MR-CISD wavefunction. The first method, *externally contracted* MR-CISD,⁸⁸ expresses eq. (37) as

$$|\Phi_{EC-MRCI}\rangle = \sum_I c_I |\Phi_I\rangle + \sum_S c_S \sum_a \bar{c}_S^a |\Phi_S^a\rangle + \sum_P c_P \sum_{ab} \bar{c}_P^{ab} |\Phi_P^{ab}\rangle$$

$$= \sum_I c_I |\Phi_I\rangle + \sum_S c_S |\Phi_S\rangle + \sum_P c_P |\Phi_P\rangle, \quad (38)$$

where the external contraction coefficients \tilde{c}_S^a and \tilde{c}_P^{ab} are determined perturbatively and the coefficients c_I , c_S , and c_P are determined variationally; hence, this is a type of variational perturbation theory.^{96–98} Note that the total number of variational coefficients is now drastically reduced compared to eq. (37). Siegbahn states⁵⁸ that the error in the correlation energy due to external contraction is roughly 1-3%.

An alternative contraction scheme which has received more attention is *internally contracted* multireference CISD (usually denoted simply CMRCI), which was first discussed by Meyer⁹⁹ and Siegbahn.¹⁰⁰ This method applies the single and double excitation operators to a single multiconfigurational reference wavefunction *as a whole*, including the reference coefficients. Thus, if the reference wavefunction is

$$|\Phi_0\rangle = \sum_R c_R |\Phi_R\rangle, \quad (39)$$

then there are at least three other classes of expansion functions—singly external, semi-internal, and doubly-external:⁵⁸

$$|\Phi_i^a\rangle = \hat{E}_{ai} |\Phi_0\rangle = \sum_S d_S |\Phi_S^a\rangle \quad (40)$$

$$|\Phi_{ij}^{ak}\rangle = \hat{E}_{ai} \hat{E}_{kj} |\Phi_0\rangle = \sum_S d_S |\Phi_S^a\rangle \quad (41)$$

$$|\Phi_{ijp}^{ab}\rangle = (\hat{E}_{ai} \hat{E}_{bj} + p \hat{E}_{aj} \hat{E}_{bi}) |\Phi_0\rangle = \sum_P d_P |\Phi_P^{ab}\rangle, \quad (42)$$

where p is +1 (-1) for singlet (triplet) coupling of a and b . The coefficients d are not variational parameters, but fixed linear combinations of the reference coefficients c_R . The final wavefunction is then

$$|\Phi_{IC-MRCI}\rangle = c_0 |\Phi_0\rangle + \sum_{ia} c_i^a |\Phi_i^a\rangle + \sum_{ijk a} c_{ij}^{ak} |\Phi_{ij}^{ak}\rangle + \sum_{ijp ab} c_{ijp}^{ab} |\Phi_{ijp}^{ab}\rangle. \quad (43)$$

Once again, the contraction has dramatically reduced the number of variational parameters. One difficulty with the internally contracted multireference CI method is that the relevant coupling coefficients become considerably more difficult to calculate. Werner and Knowles¹⁰¹ alleviate this problem by leaving internal and singly-external configurations uncontracted, i.e.,

$$|\Phi_{IC-MRCI}\rangle = \sum_I c_I |\Phi_I\rangle + \sum_{Sa} c_S^a |\Phi_S^a\rangle + \sum_{ijp ab} c_{ijp}^{ab} |\Phi_{ijp}^{ab}\rangle. \quad (44)$$

The remaining coupling coefficients still require elements of the third and fourth order reduced density matrices, which can now be evaluated each time

they are required due to advances by Werner and Knowles.^{54,102} The contraction error in the internally contracted MR-CISD method is generally only 0.1-0.2%.⁵⁸

2.4.3 Other CI Selection Schemes

In 1988, Olsen and co-workers⁴⁶ presented the restricted active space (RAS) CI, which specifies the CI space in an *a priori* manner reminiscent of the second-order CI (SOC) and its derivatives. Olsen partitions the orbitals into three subspaces, labeled RAS I, RAS II, and RAS III. Typically, RAS I contains occupied and possibly very important virtual orbitals, RAS II contains the most important virtuals, and RAS III contains the less important virtuals. The CI space includes all determinants with a minimum of p electrons in RAS I and a maximum of q electrons in RAS III. There is no restriction on RAS II, which is akin to the complete active space. Using this simple procedure, it is possible to formulate any CI space truncated according to excitation level (e.g., CISD, CISDT, etc.) as well as excitation class selected MR-CI spaces, such as SOC and CISD[TQ]. The RAS CI method is discussed more fully in section 4.8.

There are of course many other possible ways to select the CI space. For example, it is possible to generalize the RAS scheme to allow for more flexible CI spaces; work along these lines is presented later. In contrast, it is also possible to make the CI selection process essentially random. A recent paper by Greer discusses the unusual strategy of selecting CSFs using a Monte Carlo algorithm.¹⁰³

2.4.4 The First-Order Interacting Space

Another way of limiting the size of the CI space is to include only those N -electron functions which contribute to the first-order wavefunction in Rayleigh-Schrödinger perturbation theory. This is the motivation behind the *interacting space* classification: the zeroth-order interacting space consists of the set of references, and the first order interacting space includes all those N -electron functions which have a nonzero Hamiltonian matrix element with some member of the zeroth-order interacting space.¹⁰⁴⁻¹⁰⁶ Similarly, it is possible to define n -th order interacting spaces as those functions having nonzero Hamiltonian matrix elements with some member of the $(n - 1)$ th order interacting space.¹⁰⁶ Various methods for obtaining the first-order interacting space have been presented by Bunge,¹⁰⁴ Schaefer and co-workers,^{43,105} and McLean and Liu.¹⁰⁶

The first-order interacting space restriction is generally used to reduce the number of double substitutions included in single- or multi-reference CISD

wavefunctions. Although single substitutions from a Hartree-Fock reference determinant may be noninteracting due to Brillouin's theorem, they are often included nevertheless because of their strong interaction with the doubles and their importance in describing one-electron properties. For closed shell systems, the first-order interacting space criterion is inconsequential because Slater's rules dictate that all double substitutions are interacting. However, for open-shell systems there can be CSFs whose *spatial* orbital configuration differs from the reference by two electrons but which are noninteracting because of their spin coupling. In a basis of Slater determinants, one can enforce the first-order interacting space restriction simply by ensuring that all determinants differ from the reference by at most two *spin orbitals*. Aside from being more compact, wavefunctions limited to the first-order interacting space can exhibit certain orbital invariance properties.¹⁰⁷

2.4.5 Computational Scaling

Depending on the relative sizes of the number of electrons, the number of orbitals, and the excitation level, one can derive several different simple estimates of the computational cost of a configuration interaction procedure. Obviously that cost relates to the number of N -electron functions in the linear expansion of the wavefunction, and the size of the CI space for various methods has already been discussed in section 2.4.1.

For present purposes, it is sufficient to work with spin orbitals. Typically, the dimension of the CI space is dominated by determinants with the highest excitation level, m . Thus

$$N_{det} \sim \binom{N}{m} \binom{n_v}{m} \sim \frac{1}{(m!)^2} N^m n_v^m, \quad (45)$$

with n_v spin orbitals unoccupied in the reference. Most CI procedures solve only for the lowest or lowest few eigenvectors, via an iterative procedure (cf. section 3.2). In such situations, the scaling is much less than the $\mathcal{O}(N_{det}^3)$ typical of standard matrix diagonalization methods. The most expensive step in iterative procedures such as the Davidson method¹⁰⁸ is the construction of the so-called σ vectors,

$$\sigma_i = \mathbf{H}\mathbf{b}_i, \quad (46)$$

where \mathbf{b}_i belongs to a set of trial vectors which is expanded each iteration until convergence is reached. If the Hamiltonian matrix \mathbf{H} were formed directly, this procedure would require $\mathcal{O}(N_{det}^2)$ operations. This is never actually done because the storage requirements would be too great, and such an approach ignores the fact that the Hamiltonian contains only two-body terms, so that the majority of the matrix elements are zero.

Each element of a trial vector \mathbf{b}_i need only be multiplied by the nonzero elements of \mathbf{H} . The Hamiltonian will connect a maximally excited determinant with other maximally excited determinants and with other determinants having excitation level $m \geq m' \geq m - 2$. The number of interacting determinants is roughly

$$\binom{m}{2} \binom{n_v + m}{2} + \binom{m}{2} \binom{N - m}{2} + m^2 n_v (N - m), \quad (47)$$

which we further approximate as

$$\frac{1}{4} m^2 (n_v + m)^2 + \frac{1}{4} m^2 (N - m)^2 + m^2 n_v (N - m). \quad (48)$$

Each element in \mathbf{b}_i must be multiplied by the relevant nonzero matrix elements, leading to an overall operation count on the order of

$$\mathcal{O}(N^m n_v^m \{m^2 n_v^2 + m^2 N^2 + m^2 N n_v\}). \quad (49)$$

Except for full CI, we typically expect $N, n_v \gg m$. Furthermore, we almost always have $n_v > N$, so that the leading term becomes

$$\mathcal{O}(N^m n_v^{2+m}). \quad (50)$$

Thus the number of operations for a CISD procedure has a sixth power dependence on the total number of orbitals, while CISDTQ scales as the tenth power. For a given system, the number of occupied orbitals is fixed, and the cost of increasing the basis set size scales as $\mathcal{O}(n_v^{2+m})$; for CISD and CISDTQ, this scaling becomes $\mathcal{O}(n_v^4)$ and $\mathcal{O}(n_v^6)$, respectively.

The scaling of a multireference CI procedure can be estimated by multiplying the single-reference scaling by the number of references. The scaling of the CISD[TQ] method, for example, is roughly $\mathcal{O}(N^4 n_v^4)$, since the number of references is roughly N^2 if the active space is small relative to the external orbital space. For very high levels of excitation, including full CI, the number of interacting matrix elements for a given determinant becomes approximately $N^2 n^2$, so that the computational cost becomes roughly

$$\mathcal{O}^{FCI}(N_{det} N^2 n^2), \quad (51)$$

where we have replaced the term $N^m n_v^m$ with the actual number of determinants, N_{det} . For comparison, for n spatial orbitals the determinant full CI algorithm of Knowles and Handy¹⁰⁹ scales as $\mathcal{O}(N_{det} n^4)$, while the algorithm of Olsen *et al.*⁴⁶ and similar approaches scale as $\mathcal{O}(N_{det} N^2 (n - N/2)^2)$ for a closed-shell system. Although the exponents appearing in (51) are smaller than those in (50), it is important to remember that N_{det} contains a factorial dependence on N and n (see section 2.4.1); hence, a full CI procedure is extremely demanding computationally.

2.4.6 Size Extensivity Corrections

If we truncate the CI (either in the one-electron or N -electron space), we no longer have an exact theory. Of course either of these truncations will introduce an error in the wavefunction, which will cause errors in the energy and all other properties. One particularly unwelcome result of truncating the N -electron basis is that CI energies are no longer size extensive or size consistent.

These two terms—size extensive and size consistent—are used somewhat loosely in the literature. Of the two, size extensivity is the most well-defined. A method is said to be size extensive if the energy calculated thereby scales linearly with the number of particles N ; the word “extensive” is used in the same sense as in thermodynamics. A method is called size consistent if it gives an energy $E_A + E_B$ for two well separated subsystems A and B . While the definition of size extensivity applies at any geometry, the concept of size consistency applies only in the limiting case of infinite separation. In addition, size consistency *usually* also implies correct dissociation into fragments; this is the source of much of the confusion arising from this term. Thus restricted Hartree-Fock (RHF) is size extensive, but it is not necessarily size consistent, since it cannot properly describe dissociation into open-shell fragments. It can be shown that many-body perturbation theory (MBPT) and coupled-cluster (CC) methods are size extensive, but they will be size consistent only if they are based on reference wavefunction which dissociates properly.

As previously stated, truncated CI's are neither size extensive nor size consistent. A simple (and often used) example makes this clear. Consider two noninteracting hydrogen molecules. If the CISD method is used, then the energy of the two molecules at large separation will not be the same as the sum of their energies when calculated separately. For this to be the case, one would have to include *quadruple* excitations in the supermolecule calculation, since local double excitations could happen simultaneously on A and B .

Clearly the fraction of the correlation energy recovered by a truncated CI will diminish as the size of the system increases, making it a progressively less accurate method. There have been many attempts to correct the CI energy to make it size extensive. The most widely-used (and simplest) of these methods is referred to as the Davidson correction,^{110,111} which is

$$\Delta E_{DC} = (1 - c_0^2)(E_{CISD} - E_{SCF}). \quad (52)$$

This correction approximately accounts for the effects of “unlinked quadruple” excitations (i.e. simultaneous pairs of double excitations), and there are many similar expressions in use. For instance, the “renormalized” Davidson correction¹¹² is

$$\Delta E_{RDC} = \frac{1 - c_0^2}{c_0^2}(E_{CISD} - E_{SCF}). \quad (53)$$

Note that, when $c_0^2 \approx 1$, the two versions are nearly equivalent. A number of other variations exist,^{113,114} some of which force the correction to vanish for two-electron systems.

A multireference version of Davidson's correction is given by¹¹⁵

$$\Delta E_{MR-DC} = \left(1 - \sum_{i \in \text{Ref}} |c_i|^2 \right) (E_{MRCI} - E_{MR}), \quad (54)$$

where E_{MRCI} is the multireference CI energy and E_{MR} is the energy obtained from a CI in the space spanned by the references. We have simply replaced the CISD correlation energy in eq. (52) with the analogous multireference correlation energy, and we have replaced c_0^2 with the analogous sum of squares of all the reference coefficients. If the sum of the squares of reference coefficients is not near unity, better results may be obtained by using the renormalized version of this equation:

$$\Delta E_{MR-RDC} = \frac{1 - \sum_{i \in \text{Ref}} |c_i|^2}{\sum_{i \in \text{Ref}} |c_i|^2} (E_{MRCI} - E_{MR}). \quad (55)$$

It should be noted, however, that for a fixed system size, increasing the number of references decreases size extensivity errors. Indeed, for very highly correlated MR-CI wavefunctions, applying corrections such as (54) and (55) can sometimes lead to less reliable results.

There are a number of other *a posteriori* size extensivity corrections, most of which are computationally trivial once the wavefunction has been obtained. Duch¹¹⁴ compares several of the more common corrections. Of course it is also possible to allow coupling between the wavefunction and the size extensivity correction. This leads to such methods as the coupled electron pair approximation (CEPA),¹¹⁶ and the coupled pair functional (CPF) approaches.¹¹⁷ This is also the motivation behind the quadratic configuration interaction method of Pople, Head-Gordon, and Raghavachari.¹¹⁸ These authors determine the correlation energy and CI coefficients for quadratic CI with singles and doubles (QCISD) by the following set of projection equations, in spin-orbital notation:

$$\langle \Phi_0 | \hat{H} | \hat{C}_2 \Phi_0 \rangle = E_{corr} \quad (56)$$

$$\langle \Phi_i^a | \hat{H} - E_{scf} | (\hat{C}_1 + \hat{C}_2 + \hat{C}_1 \hat{C}_2) \Phi_0 \rangle = c_i^a E_{corr} \quad (57)$$

$$\langle \Phi_{ij}^{ab} | \hat{H} - E_{scf} | (1 + \hat{C}_1 + \hat{C}_2 + \frac{1}{2} \hat{C}_2^2) \Phi_0 \rangle = c_{ij}^{ab} E_{corr}, \quad (58)$$

where intermediate normalization ($\langle \Phi_0 | \Phi \rangle = 1$) has been employed, the Brillouin condition has been assumed ($\langle \Phi_i^a | \hat{H} | \Phi_0 \rangle = 0$), and \hat{C}_1 and \hat{C}_2 are the

standard single and double substitution operators,

$$\hat{C}_1 = \sum_{ia} c_i^a a_a^\dagger a_i \quad (59)$$

$$\hat{C}_2 = \frac{1}{4} \sum_{ijab} c_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i. \quad (60)$$

The QCISD projection equations differ from the equivalent CISD equations only in the addition of the quadratic terms $\hat{C}_1 \hat{C}_2$ and $\frac{1}{2} \hat{C}_2^2$, which lead to size-extensive energies. Alternatively, the QCISD equations may be considered an approximation to CCSD in which certain terms have been neglected. Pople and co-workers show how to extend this approach to include triples fully (QCISDT) or perturbatively [QCISD(T)].¹¹⁸

A multireference method building approximate size extensivity into the wavefunction is the Averaged Coupled Pair Functional (ACPF) method of Gdanitz and Ahlrichs,¹¹⁹ which introduces an electron number dependence into the denominator of the MR-CISD energy functional. A similar method has been presented by Szalay and Bartlett under the name multireference averaged quadratic coupled-cluster (MR-AQCC).^{120,121} Also noteworthy is the work of Malrieu and co-workers, who have presented a state-specific self-consistent dressing of the MR-CISD Hamiltonian matrix which gives size extensive results.¹²²⁻¹²⁴

2.4.7 The Frozen Core Approximation

It is quite common in correlated methods (including many-body perturbation theory, coupled-cluster, etc., as well as configuration interaction) to invoke the frozen core approximation, whereby the lowest-lying molecular orbitals, occupied by the inner-shell electrons, are constrained to remain doubly-occupied in all configurations. The frozen core for atoms lithium to neon typically consists of the 1s atomic orbital, while that for atoms sodium to argon consists of the atomic orbitals 1s, 2s, 2p_x, 2p_y and 2p_z. The frozen molecular orbitals are those made primarily from these inner-shell atomic orbitals.

A justification for this approximation is that the inner-shell electrons of an atom are less sensitive to their environment than the valence electrons. Thus the error introduced by freezing the core orbitals is nearly constant for molecules containing the same types of atoms. In fact, it is often preferable to employ the frozen core approximation as a general rule because most of the basis sets commonly used in *ab initio* quantum chemistry do not provide sufficient flexibility in the core region to accurately describe the correlation of the core electrons. Recently, Woon and Dunning have attempted to alleviate this problem by publishing correlation consistent core-valence basis sets.¹²⁵

Not only does the frozen core approximation reduce the number of configurations, but it also reduces the computational effort required to evaluate matrix elements between the configurations which remain. Assuming that all frozen core orbitals are doubly occupied and orthogonal to all other molecular orbitals, it can be shown¹²⁶ that

$$\langle \bar{\Phi}_I | \hat{H} | \bar{\Phi}_J \rangle = \langle \bar{\Phi}_I | \hat{H}_0 | \bar{\Phi}_J \rangle, \quad (61)$$

where $\bar{\Phi}_I$ and $\bar{\Phi}_J$ are identical to Φ_I and Φ_J , respectively, except that the core orbitals have been deleted from $\bar{\Phi}_I$ and $\bar{\Phi}_J$, and \hat{H} has been replaced by \hat{H}_0 defined by

$$\hat{H}_0 = E_c + \sum_{i=1}^{N-N_c} \hat{h}_c(i) + \sum_{i>j}^{N-N_c} \frac{1}{r_{ij}}, \quad (62)$$

where N is the number of electrons and N_c is the number of core electrons. E_c is the so-called "frozen-core energy," which is the expectation value of the determinant formed from only the N_c core electrons doubly occupying the $n_c = N_c/2$ core orbitals

$$E_c = 2 \sum_i^{n_c} h_{ii} + \sum_{ij}^{n_c} \{2(ii|jj) - (ij|ji)\}. \quad (63)$$

Finally, $\hat{h}_c(i)$ is the one-electron Hamiltonian operator for electron i in the average field produced by the N_c core electrons,

$$\hat{h}_c(i) = \hat{h}(i) + \sum_{j=1}^{n_c} \{2\hat{J}_j(i) - \hat{K}_j(i)\}, \quad (64)$$

with $\hat{J}_j(i)$ and $\hat{K}_j(i)$ representing the standard Coulomb and exchange operators, respectively. Note that, although we have written the frozen core energy E_c and frozen core operator \hat{h}_c in terms of molecular orbitals, it is not necessary to explicitly transform the one- and two-electron integrals involving core orbitals. Assuming real orbitals, we can define a frozen core density matrix¹²⁷ in atomic (or symmetry adapted) orbitals as

$$P_{\rho\sigma}^c = \sum_i^{n_c} C_\rho^i C_\sigma^i, \quad (65)$$

where C_ρ^i is the contribution of atomic orbital ρ to molecular orbital i . Now the frozen core operator in atomic orbitals becomes

$$h_{\mu\nu}^c = h_{\mu\nu} + 2 \sum_{\rho\sigma} (\rho\sigma|\mu\nu) P_{\rho\sigma}^c - \sum_{\rho\sigma} (\rho\mu|\nu\sigma) P_{\rho\sigma}^c, \quad (66)$$

and the frozen core operator in molecular orbitals h_{ij}^c can be obtained simply by transforming $h_{\mu\nu}^c$. Similarly the frozen core energy can be evaluated as

$$\begin{aligned} E_c &= \sum_{\mu\nu} P_{\mu\nu}^c (h_{\mu\nu} + h_{\mu\nu}^c) \\ &= Tr(P^c h) + Tr(P^c h^c). \end{aligned} \quad (67)$$

An analogous approximation is the *deleted virtual* approximation, whereby a few of the highest-lying virtual (unoccupied) molecular orbitals are constrained to remain unoccupied in all configurations. Since these orbitals can never be occupied, they can be removed from the CI procedure entirely. The rationalization for this procedure is that it is unlikely that electrons will choose to partially populate high-energy orbitals in their attempt to avoid other electrons. However, this conclusion is generally true only for very high-lying virtual orbitals (such as those formed by antisymmetric combinations of core orbitals for a given atom). For all other virtual orbitals, such simplistic reasoning is insufficient. Indeed, Davidson points out that those high energy SCF virtual orbitals which result from the antisymmetric combination of the two basis functions describing each valence atomic orbital in a double- ζ basis set (such as the 3p-like orbital formed from the minus combination of the larger and smaller 2p atomic orbitals on oxygen) often make the largest contribution to the correlation energy in Møller-Plesset (MPn) wavefunctions.¹²⁸

2.5 Choice of Orbitals

The results of any configuration interaction procedure depend on the choice of the atomic orbital (AO) basis. However, for a fixed AO basis, certain choices of molecular orbitals give equivalent CI wavefunctions. CI wavefunctions which are based on a single closed-shell reference and are truncated at a given excitation level are invariant to nonsingular linear transformations which mix doubly occupied orbitals with each other or unoccupied orbitals with each other. The invariance properties of CI wavefunctions based on open-shell references are more complicated, and the energy is generally not invariant to the rotation of open-shell orbitals unless certain extra references are added or the spin couplings are restricted to the first-order interacting space.¹⁰⁷ A full CI is invariant to all nonsingular linear transformations among the orbitals, even those that mix occupied and unoccupied orbitals; hence, the choice of the “reference” is irrelevant for a full CI procedure.

Some of the more elaborate CI spaces also exhibit invariance properties. Shavitt has defined the *full class CI* as one which partitions the orbitals into an arbitrary number of orthogonal subsets and includes all or none of the

N -electron functions which have a given partitioning of electrons among the subspaces.⁵⁷ The RAS CI wavefunctions are of this type, as are the RAS formulations of SOCI and CISD[TQ]. Such CI wavefunctions are invariant to separate, nonsingular linear transformations within any of the orbital subspaces. This property is relevant to the formulation of analytic gradients.¹²⁹

For bond breaking processes, the restricted Hartree-Fock approximation will not yield a good reference. This can be remedied by employing a generalized valence bond (GVB)¹³⁰ reference or an unrestricted Hartree-Fock reference. However, the latter entails spin contamination in the CI wavefunction by states of higher spin multiplicity. Another alternative is to use a multiconfigurational (MC)⁵³ or complete-active-space (CAS) SCF^{55,131} reference, which can be constructed to behave properly at all locations on the potential energy surface.

For multireference CI's such as SOCI and CISD[TQ], or any RAS CI which uses the RAS II orbital subspace, it is important that the orbitals of the active space be good correlating orbitals (i.e., they should be localized in the same region of space as the occupied orbitals). This is equally important for selected CI procedures, in that the number of configurations needed to achieve a given accuracy will be reduced. This localization criterion is not generally satisfied by canonical SCF virtual orbitals, whose construction is not physically motivated because they are based on an N -electron potential rather than an $(N - 1)$ electron potential. One possible solution is to determine the virtual orbitals using a different (and more suitable) effective Hamiltonian than that used for the occupied orbitals, and to orthogonalize the resulting orbitals against the occupied orbitals. This is the procedure in the improved virtual orbital (IVO) method of Hunt and Goddard.¹³² IVO's look like excited state orbitals and are more contracted than canonical SCF orbitals. Nevertheless, they remain somewhat too diffuse for making the CI expansion as small as possible. A related and improved method is the modified virtual orbital (MVO) approach of Bauschlicher,¹³³ who obtains virtual orbitals by diagonalizing the virtual subspace of a Fock matrix constructed for the core electrons only (eq. 64). Another possibility for obtaining compact virtual orbitals is Davidson's K -orbital approach.¹³⁴

More commonly, good correlating orbitals are obtained with the MCSCF⁵³ or CASSCF⁵⁵ methods. Yet another possibility are the natural orbitals from a CISD wavefunction. Natural orbitals (NOs)^{77,135} are defined as the eigenfunctions of the one-particle density matrix; the eigenvalues are called the occupation numbers of the NOs. One drawback of NOs is that the Hamiltonian is no longer diagonally-dominant,¹³⁶ and this can decrease the efficiency of iterative diagonalization methods (section 3.2).

Grev and Schaefer have shown for a number of small molecules that the

SOCI method performs just as well when based on CISD NOs as when based on CASSCF orbitals.¹⁶ Furthermore, they demonstrated that this is not merely due to the highly accurate treatment of correlation in the SOCI method, because a SOCI wavefunction based on canonical SCF orbitals performs notably worse. One advantage of CISD natural orbitals is that they can be easier to compute than CASSCF orbitals. Another is that their importance to the CI wavefunction falls off very rapidly with occupation number. This means that one can delete several of the most weakly-occupied NOs from the MR-CISD procedure with little loss in the correlation energy recovered; such considerations do not necessarily hold for high-lying MCSCF or CASSCF orbitals. Additionally, Parisel and Ellinger have investigated the use of CI natural orbitals in variation-perturbation methods which employ a CI wavefunction as the zeroth-order solution in a subsequent second-order perturbation treatment,¹³⁷ and Blomberg and Liu have considered the use of CI natural orbitals in SOCI transition moment calculations.¹³⁸ Balasubramanian uses SOCI natural orbitals in his relativistic CI procedure.¹³⁹

Finally, it has long been recognized that spatially localized orbitals should allow an efficient truncation of the CI space (see, for example, the PCILO method of Malrieu and co-workers^{140,141}). SCF orbitals can be localized according to the Boys procedure¹⁴² or various other methods. In most cases, the savings realized should outweigh any extra effort due to the loss of point-group symmetry. In the 1980s, Saebo and Pulay developed various localized correlation methods (including CISD) which can achieve computational savings in two distinct ways: first, the pair correlation energy for distant pairs can be neglected (or estimated), and second, the set of virtual orbitals used as correlating orbitals can be restricted to the atomic orbitals in the vicinity of the orbital to be correlated (with components of occupied orbitals projected out).^{143,144} Since standard CISD scales as the sixth power of the system size, some type of localized correlation treatment is inevitable as quantum chemists seek to apply correlated methods to large molecules.

2.6 Excited Electronic States

Here we will briefly discuss configuration interaction descriptions of excited electronic states. As previously mentioned, excited states are described by higher-energy eigenvectors of the Hamiltonian. However, since one can apply spin and spatial symmetry restrictions to the N -electron basis functions, solving for excited states which are energetically the lowest of a given symmetry species proceeds just as for the ground state. In this way, one can use orbitals which are optimal for each state.

Much more challenging is the case when several states of the same symme-

try species are required. Here, all but the lowest state are described by higher roots of the CI secular equations. Better zeroth-order descriptions are obtained if molecular orbitals are optimized separately for each state. However, this means that the resulting CI wavefunctions are interacting and nonorthogonal (complicating, for example, the evaluation of transition moments). The states can be made noninteracting and orthogonal by carrying out a non-orthogonal CI procedure,^{77,145-148} which requires the matrix **S** of overlaps between *N*-electron functions and a more complex procedure for evaluating matrix elements of the Hamiltonian (the Slater-Condon rules no longer apply because they assume a single set of orthonormal orbitals). Using orbitals optimized separately for each excited state should allow one to use smaller CI expansions to achieve a given level of accuracy.^{145,149}

However, optimizing excited state orbitals can be difficult because variational optimization always finds the lowest solution of a given symmetry species; this problem is generally called "variational collapse".¹⁵⁰ One solution is to first obtain the SCF ground state solution, and then obtain the first excited state solution by requiring it to remain orthogonal to the ground state;¹⁵¹ this process could in principle be repeated for higher-lying excited states. Another solution is to optimize the orbitals by following a higher root of the MCSCF secular equations. An early application of this idea was presented by Bauschlicher and Yarkony,⁸⁵ who optimized orbitals for the 2^1A_1 state of methylene by following the second root in a two-configuration SCF procedure. A correlated treatment of this state was obtained by solving for the second root of a two-reference CISD. This same procedure, in conjunction with more highly correlated CI methods, was recently used to re-examine the 2^1A_1 state of methylene.⁸⁶ In 1987, Allen and Schaefer presented analytic gradients for this type of TCSCF-CI procedure and used them to study the 2^1A_1 state of formaldehyde and ketene.^{150,152}

Unfortunately, the MCSCF optimization generally worsens the description of the ground state while it improves the description of the excited state. Frequently it happens that the energetic ordering of the two states will become swapped, in a process called "root flipping," and further optimization will yield orbitals describing the ground state.⁵³ One way around these difficulties is to use a single set of orbitals for all the states of a given symmetry. The improved virtual orbitals (IVO) and modified virtual orbital (MVO) methods described in the preceding section may be useful in this respect. A more typical approach is to modify the MCSCF method to yield a set of compromise orbitals; these can be obtained by the "state-averaged" procedure, which optimizes an averaged MCSCF energy obtained from averaged one- and two-electron reduced density matrices.⁵⁴ A related possibility is to use averaged natural orbitals (NOs).¹⁵³ Finally, one might simply use ground state orbitals

in conjunction with a CI method including a sufficiently complete treatment of electron correlation that the choice of orbitals becomes less important.

The most commonly used CI procedure, CISD based on the ground state configuration and using SCF orbitals, yields excitation energies which are substantially too large. One reason for this is that the ground state is correlated by all singles and double substitutions, whereas singly excited states are correlated only by singles and those doubles that involve replacement of the singly excited electron. Hence, the correlation treatment is imbalanced in favor of the ground state. This has been considered in more detail by Head-Gordon and Lee, who have analyzed the performance of CISD for excited states in the context of perturbation theory for electronic excitation energies; they find that CISD is not even correct through second order.¹⁵⁴ Another problem is that the SCF orbitals themselves bias results towards the ground state.

Alternatively, one might specifically design modified CI methods for excited states; the symmetry-adapted cluster (SAC) CI approach of Nakatsuji is such a method,^{32, 155, 156} although it also contains elements from coupled-cluster theory.⁶ Another alternative is the CASPT2 method of Andersson, Malmqvist, Roos, and co-workers,¹⁵⁷⁻¹⁵⁹ which is a second-order perturbation theory based on a CASSCF reference function. Rather than employ more complex CI approaches, Pople, Head-Gordon, and others have advocated the use of configuration interaction with only singles (CIS) as a qualitative excited state theory and as a starting point for more advanced treatments.^{154, 160} Clearly CIS offers no improvement for the ground state (Brillouin's theorem), but higher roots represent excited states with an accuracy in the excitation energies of around 1 eV (23 kcal mol⁻¹). CIS has the unusual property of being both size extensive and variational; no other truncated CI method is size extensive. Its low computational cost and size extensivity make CIS applicable to large systems. Head-Gordon and co-workers have introduced a perturbative doubles correction for CIS which they denote CIS(D);¹⁶¹ this method tends to improve excitation energies, but it does not necessarily improve geometries or other properties.¹⁶²

The performance of simple CIS for open-shell systems such as radicals is not as good as for closed-shell systems, regardless of whether an unrestricted Hartree-Fock (UHF) or restricted open-shell Hartree-Fock (ROHF) reference is used.¹⁶³ Maurice and Head-Gordon find improved results for these systems by using a spin-pure CI wavefunction, denoted XCIS, which adds to the singles those doubly substituted determinants in which the excited electron has its spin flipped and one of the open-shell electrons is also spin flipped to conserve S_z .¹⁶⁴ It is interesting to note that these limited double substitutions are actually single substitutions from the point of view of spatial orbital configurations; this problem of the non-transferability of the substitution level (or

“excitation level”) definition between determinants and CSFs has occasionally been mentioned in the literature.^{35,165} Size extensivity is maintained in the XCIS method by uncoupling the SCF solution from the excited states. More flexible approaches which still scale favorably with system size would provide a useful alternative to the more expensive EOM-CCSD⁶ and CASPT2¹⁵⁹ methods and are eagerly anticipated.

3 Common Features of Implementations

This section briefly discusses two elements common to all configuration interaction programs: transformation of integrals, and iterative subspace diagonalization of the Hamiltonian.

3.1 Integral Transformation

As discussed in section 2.3, the Hamiltonian matrix elements are generally written in terms of one- and two-electron integrals in the molecular orbital (MO) basis. However, these integrals are originally calculated in the atomic orbital (AO) basis, or perhaps the symmetry-adapted orbital (SO) basis. Therefore it is necessary to transform the AO or SO integrals into the MO basis, according to

$$h_{ij} = \sum_{\mu\nu} C_{\mu}^i C_{\nu}^j h_{\mu\nu}, \quad (68)$$

$$(ij|kl) = \sum_{\mu\nu\rho\sigma} C_{\mu}^i C_{\nu}^j C_{\rho}^k C_{\sigma}^l (\mu\nu|\rho\sigma), \quad (69)$$

where C_{μ}^i is the coefficient for the contribution of atomic orbital μ to molecular orbital i , and real orbitals have been assumed. Although the coefficients C_{μ}^i are generally the SCF coefficients, they might instead be the coefficients of the CI natural orbitals in the atomic orbital basis, etc.

3.1.1 One-electron Integrals

The transformation of the one-electron integrals is computationally inexpensive and easily accomplished: without point group symmetry, this transformation can be performed as two half-transformations, each of which requires a multiplication of the one-electron integral matrix by the SCF coefficient matrix, for a total of $2n^3$ multiplications. Spatial symmetry reduces this cost because the one-electron integral and SCF coefficient matrices are block diagonal according to irreducible representation (irrep), and the transformation can be carried out an irrep at a time (cf. Figure 1). Note that it would also be

Figure 1: Transformation of one-electron integrals.

```

Loop over irreps  $\Gamma$ 
   $h_{\mu j}^{\Gamma} = \sum_{\nu} C_{\nu j}^{\Gamma} h_{\mu \nu}^{\Gamma}$  (matrix mult)
   $h_{i j}^{\Gamma} = \sum_{\mu} C_{\mu i}^{\Gamma} h_{\mu j}^{\Gamma}$  (matrix mult)
end loop over  $\Gamma$ 

```

possible to utilize the permutational symmetry $h_{ij} = h_{ji}$, but this would typically reduce efficiency because the transformation could no longer be written in terms of matrix multiplications, which are very fast on vector supercomputers (e.g., CRAY C90) or pipelined workstations (e.g., IBM RS/6000).¹⁶⁶ Note that if orbitals are frozen in the correlated procedure (cf. section 2.4.7), $h_{\mu\nu}$ is replaced by the frozen core operator $h_{\mu\nu}^c$ (eq. 66).

3.1.2 Two-electron Integrals

Transforming the two-electron integrals is considerably more time-consuming. Equation (69) implies that this transformation is an n^4 process for each of n^4 integrals ($ij|kl$) (or n^8 overall), but of course it can be carried out as four separate quarter-transformations analogous to the two half-transformations required for the one-electron integrals; this strategy requires $4n^5$ multiplications if symmetry is neglected and it constitutes a fairly demanding computational procedure if n is larger than 100 or so. Fortunately, the full transformation is not necessary for the simple CIS method because only MO integrals with two internal and two external indices are relevant. Matrix elements for most other CI wavefunctions are expressed in terms of the full set of MO integrals, but by performing some compensating work, one can avoid the full transformation for CISD,^{167,168} internally-contracted MR-CISD,^{101,168} and even uncontracted MR-CISD.¹⁶⁹ For the latter, however, Saunders and van Lenthe argue that the extra steps required to avoid the full transformation may cost more than the transformation itself unless the AO integral list exhibits considerable sparsity.¹²⁷ In the general case, the full set of integrals is required. Therefore, various methods for employing spatial and permutational symmetry have been proposed to reduce the operation count. In this context, the “permutational symmetry” refers to the eight-fold redundancy in the two-electron integrals for real orbitals.

Wilson¹⁷⁰ provides a very clear and helpful survey of four-index transformation methods published before 1987. More recent work has focused attention on the sparsity of quantities in the AO basis. For example, Häser, Almlöf, and Feyereisen have presented an integral-direct transformation algorithm which

Figure 2: Pre-sorting the two-electron AO integrals in TRANSQT.

```

Form frozen core density matrix, eq. (65)
Initialize Yoshimine structure for sorting AO two-el ints
Read two-el ints from disk; form frozen core operator, eq. (66)
    and write integrals to Yoshimine buffers
Complete Yoshimine sort, ensuring  $p \geq q$ ,  $r \geq s$ , but not  $pq \geq rs$ 
Free Yoshimine pre-sort structure
Evaluate frozen core energy, eq. (67)

```

employs integral pre-screening techniques and can even exploit non-abelian point-group symmetry.¹⁷¹ However, our attention here is focused on integral transformation routines for highly correlated CI wavefunctions, which typically means that one can consider only small molecules for which there is less benefit in exploiting sparsity in the AO basis. Of the conventional approaches discussed by Wilson, one of the most promising is the Saunders-van Lenthe algorithm,¹²⁷ which has an operation count of $\sim 25n^5/24$ (the operation count is somewhat less if the number of transformed orbitals is less than the number of AO's). Saunders and van Lenthe present an explicit algorithm for the case of no spatial symmetry.¹²⁷ However, in our experience it is not entirely straightforward to symmetry adapt their algorithm and simultaneously maintain a high degree of vectorization. On the other hand, we find it straightforward to symmetry adapt a simpler series of quarter transformations in which some of the permutational symmetry of the integrals is ignored. This simpler code remains efficient because it calls optimized matrix multiplication subroutines. This new program, TRANSQT, developed by Daniel Crawford, Justin Fermann, and the present authors, runs faster than previous transformation programs produced by this group which take more advantage of permutational symmetries.

The algorithm consists of three major parts: a pre-sort of the two-electron SO integrals, the first half-transform, and the second half-transform (Figures 2-4). To keep track of spatial symmetry, the loops over orbitals are broken up into loops over irreps of the point group and over orbitals within those irreps. The atomic orbitals are numbered consecutively within each irrep, which allows the use of relative indices (denoted by capital letters in the figures) for numbering orbitals with a fixed irrep. Molecular orbitals are also numbered this way until they are written out at the end of the transformation, when they are renumbered according to whatever order is used by subsequent programs. As seen in the figures, all multiplications which give zero by symmetry are

avoided. The use of matrix multiplications means that the algorithm takes only partial advantage of permutational symmetry. The pre-sorted integrals P use the permutational symmetries $(pq|rs) = (qp|rs) = (pq|sr) = (qp|sr)$, but they do not allow the swapping of the first pair of indices with the last pair [e.g., $(pq|rs) = (rs|pq)$]. These same permutational symmetries are employed during multiplication except that if $rsym = ssym$, then the symmetry $(pq|rs) = (pq|sr)$ is not utilized. Similar considerations apply to the second half-transform: the half-transformed integrals J are stored similarly to P , with $(pq|kl) = (pq|lk) = (qp|kl) = (qp|lk)$.

The TRANSQT algorithm employs canonical indices⁵⁷ for pairs of orbitals, such as $pq = \text{ioff}[p] + q$. These indices are useful for computing the address of an element in a symmetric matrix which is stored by writing only the lower triangle to a linear array. The array $\text{ioff}[p]$ contains the address of the first element in row p , and it is assumed that $p \geq q$. If orbital numbering starts from zero, then $\text{ioff}[0] = 0$ and $\text{ioff}[p] = p + \text{ioff}[p-1]$. The memory requirements are for two matrices (A and B) with dimension equal to the number of atomic orbitals, a matrix of SCF coefficients for each irrep, two blocks which hold all two-electron integrals $(pq|rs)$ with a fixed pair of first indices pq , and various buffers associated with Yoshimine sorting. This sorting method, first described by Yoshimine¹⁷² in 1969, is needed to sort the integrals so that they can be read sequentially in the required order. The pre-sort is necessary because the first half-transform requires all $(pq|rs)$ for a given pq , but the integrals are not stored this way in the disk file produced by our integrals program, where they possess the full eight-fold permutational symmetry. For instance, the first half-transform will require integrals such as $(11|43)$, but this integral is only stored as $(43|11)$ on disk. The pre-sort adds the redundant integrals $(kl|ij) = (ij|kl)$ and places them all in the correct order for reading. The second Yoshimine sort involves the half-transformed integrals: these integrals are formed in the order $(pq|kl)$, where pq is fixed. In the second half-transform, however, the program needs to read all $(kl|pq)$ for a fixed kl .^{*} Since the integrals were not written in this order, they must be sorted so they can be read this way.

3.2 Iterative Techniques for Solving $Hc = Ec$

Standard numerical methods exist for diagonalizing real symmetric matrices such as the Hamiltonian H .[†] However, such methods usually require the stor-

^{*}The convention used here is to write the fixed orbital pair first in the two-electron integral. However, one must exercise caution because the half-transformed integrals do not possess the symmetry $(pq|kl) = (kl|pq)$ since a distinction must be drawn between the AO and MO pairs.

[†]Once again, real orbitals have been assumed, along with a nonrelativistic Hamiltonian.

Figure 3: First half-transform in TRANSQT.

```

Initialize tmp matrices A and B and buffers Pblock and Jblock
Loop psym over irreps
  Loop p over orbitals in irrep psym
    Loop qsym over irreps w/ qsym ≤ psym
      Loop q over orbitals in irrep qsym w/ q ≤ p
        pq = ioff[p] + q
        Read (pq|rs) for all rs given pq into Pblock
        Loop rsym over irreps
          Compute ssym from psym, qsym, rsym
          Loop r over orbitals in irrep rsym (relative idx R)
            Loop s over orbs in ssym w/ s ≤ r (rel idx S)
              rs = ioff[r] + s
              A[R][S] = Pblock[rs]
              if rsym = ssym, A[S][R] = Pblock[rs]
            end loop over s
          end loop over r
          matrix multiply:
          loop rel idx R over orbs in rsym
            loop rel idx L over active orbs in ssym
              loop rel idx S over orbs in ssym
                 $B[R][L] = B[R][L] + A[R][S] * C^{ssym}[S][L]$ 
              end loops over S, L, R
            matrix multiply:
            loop rel idx K over active orbs in rsym
              loop rel idx L over active orbs in ssym
                loop rel idx R over orbs in rsym
                   $A[K][L] = A[K][L] + (C^{rsym})^T[K][R] * B[R][L]$ 
                end loops over R, L, K
              loop k over active orbitals in rsym (rel idx K)
                loop l over active orbs in ssym (rel idx L), l ≤ k
                  kl = ioff[k] + l
                  Jblock[kl] = A[K][L]
                end loops over l, k
              Write Jblock to Yoshimine buffers
            end loop over rsym
          end loop over q
        end loop over qsym
      end loop over p
    end loop over psym
  flush, close I/O files
  free Pblock
  Yoshimine sort half-transformed integrals J, free Yoshimine struct

```

Figure 4: Second half-transform in TRANSQT.

```

Loop ksym over irreps
  Loop k over active orbitals in irrep ksym
    Loop lsym over irreps w/ lsym ≤ ksym
      Loop l over active orbitals in irrep lsym w/ l ≤ k
         $kl = \text{iof}[k] + l$ 
        zero Jblock
        Read all (kl|pq) for given kl into Jblock
        Loop psym over irreps
          Compute qsym from ksym, lsym, and psym
          if (qsym > psym) next psym
            Loop qsym over irreps, with qsym ≤ psym
              Loop p over orbs in psym (rel idx P)
                Loop q over orbs in qsym (rel idx Q)
                   $pq = \text{iof}[p] + q$  if  $p \geq q$ , else  $\text{iof}[q] + p$ 
                   $A[P][Q] = \text{Jblock}[pq]$ 
                end loop over q
              end loop over p
              matrix multiply:
              Loop rel idx P over orbs in psym
                Loop rel idx J over active orbs in qsym
                  Loop rel idx Q over orbs in qsym
                     $B[P][J] = B[P][J] + A[P][Q] * C^{qsym}[Q][J]$ 
                  End loops over Q, J, P
                matrix multiply:
                Loop rel idx I over active orbs in psym
                  Loop rel idx J over active orbs in qsym
                    Loop rel idx P over orbs in psym
                       $A[I][J] = A[I][J] + (C^{psym})^T[I][P] * B[P][J]$ 
                    end loops over P, J, I
                  Write matrix A to buffer
                End loop over qsym
              End loop over psym
            End loop over l
          End loop over lsym
        End loop over k
      End loop over ksym
    free Jblock
  flush and close I/O buffers

```

age of \mathbf{H} in core memory (if eigenvectors are also computed, then one actually needs memory to store *two* matrices of this size). If the CI includes a mere 10,000 determinants (certainly a small CI space), storing the full matrix \mathbf{H} would require 800 megabytes. As of 1997, this represents a large amount of core memory (although disk storage would not be a problem). It is little consolation that the symmetry of \mathbf{H} could be used to cut this requirement approximately in half. Another very important difficulty is the time required to diagonalize matrices this size or larger. Most diagonalization routines scale as $O(n^3)$, which is certainly problematic for $n \geq 10^4$. Only for smaller matrices do the standard methods become practical.

In typical applications, only the ground electronic state or perhaps a few of the low lying excited states are of interest. Hence methods which obtain only the lowest few roots of the CI matrix are greatly preferred over methods which compute the entire spectrum. Furthermore, storage requirements are greatly reduced if \mathbf{H} is not stored at all; *direct CI* methods, discussed in section 4, form products $\mathbf{H}\mathbf{c} = \sigma$ directly from the MO integrals.

Most techniques for solving large eigenvalue problems fall under the category of *subspace iteration* methods, which iteratively solve the eigenvalue problem in a linear vector subspace spanned by only a few vectors. Malmqvist¹⁷³ provides a concise review of the subspace iteration methods most commonly found in quantum chemistry. Here we will outline some of these methods and note recent advances.

3.2.1 Davidson's Method

Davidson's method for the iterative solution of the lowest few eigenvalues and eigenvectors of large real, symmetric matrices¹⁰⁸ is undoubtedly the most widely used technique for solving the CI secular equations. In this method, one applies standard diagonalization methods to a small Hamiltonian matrix formed in a subspace $\{\mathbf{b}_i\}$ of L orthonormal expansion vectors, where L increases from iteration to iteration but is typically very much smaller than the dimension of \mathbf{H} (the subspace generally includes less than a dozen vectors per root). At each iteration, the Davidson algorithm estimates a correction vector for each root currently under consideration and adds it to the set $\{\mathbf{b}_i\}$ after Schmidt orthogonalization.

Davidson used perturbation theory to argue¹⁰⁸ that the best correction vector δ to the current iteration's guess vector \mathbf{c} satisfies

$$(\mathbf{H} - \lambda\mathbf{I})\delta = -(\mathbf{H} - \lambda\mathbf{I})\mathbf{c}. \quad (70)$$

In the Davidson method, one approximates λ by the current iteration's eigenvalue, and \mathbf{H} is assumed to be diagonally dominant so that δ can be approxi-

mated by

$$\delta = -(\mathbf{H}_d - \lambda\mathbf{I})^{-1}(\mathbf{H} - \lambda\mathbf{I})\mathbf{c}, \quad (71)$$

where \mathbf{H}_d is the diagonal of \mathbf{H} and the denominator is referred to as the *preconditioner*.

Liu showed how to extend Davidson's method to solve for several roots simultaneously,¹⁷⁴ leading to what is called the Simultaneous Expansion Method, the Davidson-Liu method, or the block Davidson method. The detailed Davidson-Liu algorithm, adapted from ref. 174, is presented in Figure 5.

At each iteration, the current approximations to the eigenvalues of \mathbf{H} are given by the eigenvalues of the small matrix \mathbf{G} , which is the Hamiltonian in the subspace spanned by the expansion vectors $\{\mathbf{b}_i\}$, with matrix elements $G_{ij} = (\mathbf{b}_i, \mathbf{H}\mathbf{b}_j)$. Likewise, the current approximate eigenvectors are linear combinations of the subspace vectors with coefficients given by the eigenvectors α of \mathbf{G} :

$$\mathbf{c}^k = \sum_{i=1}^L \alpha_i^k \mathbf{b}_i. \quad (76)$$

The convergence of the k -th root can be checked by the sum of squares of the last m components of α^k in step 2 or by the norm of the residue vector \mathbf{r} in step 3.

Unless very tight convergence criteria are specified, it is possible for the Davidson-Liu method to converge on the wrong eigenvector if the initial guess vectors are poor. Although this will not happen for the ground state unless a completely inappropriate guess is provided, it occasionally happens when several roots are sought. Davidson and co-workers recommend initial loose convergence of more roots than are actually needed, and then tighter convergence on the desired roots.¹⁷⁵ Possible choices for the initial vectors include unit vectors (chosen according to the diagonal elements of \mathbf{H} with the largest magnitudes) or eigenvectors of some small block of \mathbf{H} .

Equations (73) and (75) show that after a matrix-vector product $\sigma_i = \mathbf{H}\mathbf{c}_i$ is computed it is needed again in subsequent iterations. Since the construction of the vectors $\{\sigma_i\}$ is the most time consuming step in the iterative diagonalization of \mathbf{H} , they are stored on disk along with the expansion vectors $\{\mathbf{b}_i\}$. The original Davidson method converges one root at a time and requires storage for two (segments of) vectors at once in core memory. If more core memory is available, then the Davidson-Liu method can reduce computational and I/O requirements. For example, one pass through the subspace expansion vectors is sufficient to construct several correction vectors δ simultaneously. Likewise, a single construction of the matrix elements of \mathbf{H} , several σ vectors can be formed simultaneously.

Note that the preconditioner in eq. (74) requires the diagonal elements of

Figure 5: The Davidson-Liu Iterative Method for the Lowest Few Eigenvectors and Eigenvalues of Real, Symmetric Matrices (Ref. 174)

1. Select a set of L orthonormal guess vectors, at least one for each root desired, and place in the set $\{\mathbf{b}_i\}$.
2. Use a standard diagonalization method to solve the $L \times L$ eigenvalue problem

$$\mathbf{G}\alpha^k = \lambda^k \alpha^k, k = 1, 2, \dots, M \quad (72)$$

where

$$G_{ij} = (\mathbf{b}_i, \mathbf{H}\mathbf{b}_j) = (\mathbf{b}_i, \sigma_j), 1 \leq i, j \leq L \quad (73)$$

and M is the number of roots of interest.

3. Form the correction vectors $\{\delta^k\}, k = 1, 2, \dots, M$, defined as

$$\delta_I^k = (\lambda^k - H_{II})^{-1} r_I^k, I = 1, 2, \dots, N \quad (74)$$

where

$$\mathbf{r}^k = \sum_{i=1}^L \alpha_i^k (\mathbf{H} - \lambda^k) \mathbf{b}_i \quad (75)$$

and N is the number of determinants or CSFs.

4. Normalize $\{\delta^k\}$.
5. Schmidt orthonormalize δ^1 against the set $\{\mathbf{b}_i\}$ and append the result to $\{\mathbf{b}_i\}$. Repeat this process for each of the other $M - 1$ correction vectors, neglecting those whose Schmidt orthonormalized norm is less than some threshold $T \sim 10^{-3}$. This results in the addition of m new \mathbf{b} vectors, with $1 \leq m \leq M$.
6. Increase L by m and return to step 2.

the Hamiltonian. These can be precomputed and stored on disk, or they can be computed on-the-fly. Alternatively, they can be approximated in some cases using orbital energies. In a determinantal basis, Davidson's preconditioner can actually cause the CI vector to break spin symmetry. Indeed, Knowles and Handy noted this difficulty in their pioneering 1984 paper on determinant based configuration interaction.¹⁰⁹ They found that this problem can be avoided if the diagonal elements of the Hamiltonian H_{II} are replaced by an average \bar{H}_{II} over all determinants which have the same spatial orbital configuration as determinant I but differ in the distribution of spins.

If several roots are sought, or convergence takes many iterations, then the number of vectors stored on disk can become large, leading to I/O delays during the construction of \mathbf{G} or \mathbf{r} . Furthermore, disk storage can become a problem if the vectors themselves are very large. One solution is to apply compression algorithms to the subspace vectors and σ vectors.^{176,177} Additionally, the Davidson-Liu method can be restarted with only M expansion vectors by using the current approximation to each eigenvector, eq. (76), as the new starting guess vectors. Clearly this procedure hinders the rate of convergence because information is lost after the vector subspace is collapsed: subsequent diagonalizations have less variational freedom because of the reduced dimension of the subspace. However, in 1990 van Lenthe and Pulay presented the remarkable conclusion¹⁷⁸ that when only a single root is sought, collapsing the subspace does not substantially degrade the rate of convergence if the subspace is collapsed down to two vectors instead of just one. This procedure, which may be justified by the theory of conjugate gradients, was later generalized to multiple roots by Murray, Racine, and Davidson.¹⁷⁵ The collapsed vector subspace contains the current guess vector for each root, as before, and also the guess vectors from the previous iteration (after they have been Schmidt orthogonalized against the other vectors in the collapsed subspace).

Other work has focused on improving the correction vector. As noted by Olsen,⁸³ Saad,¹⁷⁹ Sleijpen and van der Vorst,¹⁸⁰ and others,^{181,182} Davidson's equation (70) seems to imply that the optimal update vector δ is just the negative of the current guess CI vector \mathbf{c} . Clearly, this would not allow for the expansion of the vector subspace. Sleijpen *et al.*¹⁸⁰ have pointed out that Davidson assumed that δ is orthogonal to \mathbf{c} in deriving eq. (70). However, Davidson's method only enforces this orthogonality after δ has already been determined. Hence, a more effective preconditioner may result from explicitly enforcing this orthogonality while δ is being constructed, and several authors have recently proposed such preconditioners.^{83,181,182}

Another improvement suggested by these authors and others^{54,56,173} is to lift the assumption of strict diagonal dominance of the Hamiltonian in the preconditioner. One selects a subspace of the most important N -electron functions

in the CI space and, for the purposes of the preconditioner, approximates the Hamiltonian as

$$\mathbf{H} \approx \begin{bmatrix} \mathbf{H}' & 0 \\ 0 & \Lambda'' \end{bmatrix}, \quad (77)$$

where primes denote the small selected subspace and double primes denote the complement subspace. Although relatively few (up to several hundred) determinants might be included in the selected space, it is important to take complete sets of determinants which are capable of forming spin eigenfunctions so that the spin symmetry of the CI vector can be maintained during the iterative procedure. Note that the Hamiltonian is assumed diagonal in the complement subspace and coupling is ignored. This leads to two equations for the correction vector:

$$(\delta^k)' = -(\mathbf{H}' - \lambda^k \mathbf{I}')^{-1} (\mathbf{r}^k)' \quad (78)$$

$$(\delta^k)'' = -(\Lambda'' - \lambda^k \mathbf{I}'')^{-1} (\mathbf{r}^k)'' \quad (79)$$

The second equation of course becomes

$$(\delta_I^k)'' = \frac{(r_I^k)''}{\lambda_k - H_{II}}, \quad (80)$$

and is analogous to eq. (74). The first equation can be written in terms of the eigenvalues μ^l and eigenvectors \mathbf{u}^l of the small matrix \mathbf{H}' :

$$(\delta_I^k)' = \sum_l \frac{u_I^l (\mathbf{u}^l \cdot (\mathbf{r}^k)')}{\lambda^k - \mu^l} \quad (81)$$

3.2.2 Olsen's Method

Realizing the difficulties of storing several \mathbf{b} and σ vectors for very large CI spaces, Olsen proposed that each correction vector be added directly to the current CI vector, and that the resulting (renormalized) vector be used as the next iteration's guess vector. Of course for this scheme to work well, the correction vector must be as good as possible. Olsen therefore introduced an improved method for generating the correction vector, using some of the ideas just discussed above. If the current (normalized) CI vector is denoted \mathbf{c} , then its energy is

$$E = (\mathbf{c}, \mathbf{H}\mathbf{c}). \quad (82)$$

The Hamiltonian is then divided into two terms,

$$\mathbf{H} = \mathbf{H}^{(0)} + \mathbf{H}^{(1)}, \quad (83)$$

and the CI eigenvalue equation can be written

$$(\mathbf{H}^{(0)} + \mathbf{H}^{(1)})(\mathbf{c} + \delta\mathbf{c}) = (E + \delta E)(\mathbf{c} + \delta\mathbf{c}), \quad (84)$$

where $\delta\mathbf{c}$ and δE are the corrections to the current CI vector and energy. If $\delta\mathbf{c}$ is required to be orthogonal to \mathbf{c} then (neglecting quadratic terms)

$$\delta\mathbf{c} = -(\mathbf{H}^{(0)} - E)^{-1} [(\mathbf{H} - E)\mathbf{c} - \mathbf{c}\delta E], \quad (85)$$

where

$$\delta E = \frac{(\mathbf{c}, (\mathbf{H}^{(0)} - E)^{-1}(\mathbf{H} - E)\mathbf{c})}{(\mathbf{c}, (\mathbf{H}^{(0)} - E)^{-1}\mathbf{c})} \quad (86)$$

The correction vector $\delta\mathbf{c}$ is superior to that used in the standard Davidson method since it remains rigorously orthogonal to \mathbf{c} and therefore retains the ability to introduce new character into the CI vector even near convergence.⁸³ This correction vector was also derived by Bofill and Anglada from other considerations.¹⁸¹ In 1990, Olsen, Jørgensen, and Simons used this method to perform three iterations of the first one-billion determinant CI calculation.⁸³ The zeroth-order part of the Hamiltonian $\mathbf{H}^{(0)}$ was defined as a 400×400 block of determinants formed from the lowest diagonal elements of \mathbf{H} , and as the diagonal of \mathbf{H} outside this block. This procedure requires the storage of four vectors on disk (three if the diagonal elements of \mathbf{H} are computed as needed).

Although the Olsen method can be very helpful when disk space is limited, its convergence characteristics are not always very good. Indeed, as first pointed out by Mitrushenkov,¹⁸³ the Olsen method does not guarantee that the energy decreases every iteration. However, it is of course possible to use Olsen's preconditioner in conjunction with iterative methods which keep more than one CI vector and σ vector. Mitrushenkov¹⁸³ advocates diagonalizing the Hamiltonian in the space of the current and previous CI vectors:

$$H_{ii} = (\mathbf{c}^{(i)}, \sigma^{(i)}) \quad (87)$$

$$H_{i,i-1} = H_{i-1,i} = (\sigma^{(i)}, \mathbf{c}^{(i-1)}). \quad (88)$$

Explicitly, the nonorthogonal pseudo-eigenvalue equation is

$$\begin{pmatrix} H_{i-1,i-1} & H_{i-1,i} \\ H_{i,i-1} & H_{i,i} \end{pmatrix} \begin{pmatrix} \alpha_{i-1} \\ \alpha_i \end{pmatrix} = E_i \begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix} \begin{pmatrix} \alpha_{i-1} \\ \alpha_i \end{pmatrix}, \quad (89)$$

where s is the overlap between the current and previous CI vectors, $(\mathbf{c}^{(i)}, \mathbf{c}^{(i-1)})$. At each iteration, the current CI vector is recomputed as $\mathbf{c}^{(i)} = \alpha_{i-1}\mathbf{c}^{(i-1)} + \alpha_i\mathbf{c}^{(i)}$, and $\sigma^{(i)} = \alpha_{i-1}\sigma^{(i-1)} + \alpha_i\sigma^{(i)}$. H_{ii} is set to E_i in eq. (89), and then the new vector $\mathbf{c}^{(i+1)}$ is computed using Olsen's method. This process is repeated until convergence is reached. In our experience Mitrushenkov's method

improves convergence substantially in the first few iterations compared to the single-vector version of Olsen's method. Unfortunately, as CI vector approaches convergence, eq. (89) becomes ill-conditioned because $H_{i-1,i}$ approaches H_{ii} and s approaches unity. This difficulty can be avoided by reverting to the single-vector Olsen method near convergence.

4 Determinant-Based Algorithms for Highly Correlated CI

This section describes several determinant-based CI algorithms. The alpha and beta string formalism of Handy⁴⁴ is introduced, and the equations for $\sigma = \mathbf{Hc}$ within this formalism are derived for the full CI and RAS CI cases.⁴⁶ Practical considerations for implementation are also discussed.

4.1 Slater Determinants, CSFs, and Direct CI

Slater determinants are eigenfunctions of \hat{S}_z ; therefore, the CI space includes only those determinants with a given value of M_s , unless a spin-dependent Hamiltonian is used.¹⁸⁴ However, Slater determinants are not eigenfunctions of \hat{S}^2 as are configuration state functions (CSFs), and dimension of the CI space in Slater determinants is typically about 2-4 times larger than in CSFs (for more about CSFs, see the books by Pauncz^{185,186}). Thus it would seem that CSFs are preferable to Slater determinants for use as CI expansion functions. However, determinants offer certain advantages in the context of "direct CI" methods that can outweigh the disadvantage of a larger CI space.

In computational quantum chemistry the term "direct" has come to mean that certain quantities, which are too large to hold in core memory, are computed on-the-fly instead of being stored on disk and read as needed. A *direct SCF* implies that the two-electron integrals are computed on-the-fly. For a *direct CI*,⁸⁸ the integrals are still held on disk, but the Hamiltonian matrix itself is not explicitly constructed or stored. Instead, the vector $\sigma = \mathbf{Hc}$, which is required in iterative subspace methods for diagonalizing the Hamiltonian (cf. section 3.2), is computed directly from the one- and two-electron integrals and the CI vector; the construction of σ is the time-consuming step in the direct CI method. The coefficients for multiplying the CI vector by the integrals have already been introduced (cf. section 2.3.2) as the one- and two-electron coupling coefficients; these may be written to disk in a file traditionally called the "formula tape." Unfortunately, this procedure is still unsuitable for a direct CI, since the coupling coefficients will require as much storage space as the Hamiltonian matrix itself, leading to long I/O delays. Hence, in a direct

CI, the coupling coefficients should be calculated as needed (or they should be built into the program). A next step would be to eliminate storage of the CI vector itself; efforts along these lines have been described as *superdirect CI*.^{187,188} Alternatively, Carter and co-workers have considered pseudospectral approaches which eliminate the need to construct two-electron integrals as separate intermediates.¹⁸⁹⁻¹⁹²

The direct CI method was first introduced in 1972 by Roos for the case of CISD from a closed-shell reference function.¹⁹³ However, generalization to more complex CI spaces, such as MR-CISD, proved exceedingly difficult due to the large number of special cases. The next breakthrough did not occur until Shavitt^{41,42} cast the work of Paldus on the unitary group approach (UGA)^{40,78} into a graphical formalism which represents CSFs in the Gelfand-Tsetlin canonical basis as walks on a directed graph. Not only did this graphical representation make the UGA more accessible to chemists, but it also provided a convenient formalism for carrying out computations. Any pair of walks (CSFs) forms a loop, and matrix elements are evaluated based on the shapes of these loops, with only certain loop types giving nonzero matrix elements.^{41,42} For example, one-electron coupling coefficients are expressed as

$$\gamma_{ij}^{IJ} = \prod_{k=i}^j W(T_k, b_k), \quad (90)$$

where T_k identifies the shape of the loop formed by walks I and J at level k on the Shavitt graph, and b_k is the “b-value” of the vertex on walk J at level k (see ref. 42 for more details). The coupling coefficient vanishes unless walks I and J coincide everywhere below level $i - 1$ and above level j on the graph (assuming $i < j$). This graphical unitary group approach (GUGA) was developed with a philosophy similar to that of the direct CI. The idea was to use each coupling coefficient, specified by a loop on the Shavitt graph, for a whole series of Hamiltonian matrix elements differing in their common upper and lower walks. The first multireference CI method based on the ideas of Paldus and Shavitt was developed by Brooks and Schaefer.⁴³ Particularly notable was their computation on the ${}^1B_{1u}$ state of ethylene involving all single and double excitations relative to three open-shell singlet reference configurations.⁴³ However, a detailed analysis of the “loop-driven” GUGA CI program of Brooks and Schaefer indicates that, in practice, few loops contribute to very many matrix elements, and it remains more efficient to write the coupling coefficients to the formula tape rather than to recompute them as needed.

Nevertheless, the graphical approach afforded new insight into the structure of the Hamiltonian. In particular, for CI spaces which allow only one or two electrons in the external space, the graphical representation of the external space becomes very regular, and the external portion of a loop can only have

a few possible shapes.¹⁹⁴ Siegbahn made the crucial observation that the one- and two-electron coupling coefficients can be factored into contributions from the internal orbitals and from the external orbitals,

$$\gamma_{ij}^{IJ} = \text{int}\gamma_{ij}^{IJ} \times \text{ext}\gamma_{ij}^{IJ} \quad (91)$$

$$\Gamma_{ijkl}^{IJ} = \text{int}\Gamma_{ijkl}^{IJ} \times \text{ext}\Gamma_{ijkl}^{IJ}, \quad (92)$$

and that the calculation can be "direct" in the external space when the external factors are very simple.

In 1979, Siegbahn showed¹⁹⁵ that for the case of single replacements from a reference wavefunction which is a full CI in the active space (i.e., first-order CI¹³), the external factors for the coupling coefficients are all simply +1; hence, only the internal space coupling coefficients must be precomputed and stored on the formula tape. This results in a substantial savings because the number of internal coupling coefficients will be much smaller than the total number of coupling coefficients. In 1980, Siegbahn extended these ideas to the general case of all single and double substitutions for an arbitrary set of references (i.e., MR-CISD).¹⁹⁴ Once again, the external coupling coefficient factors are simple (± 1 , $\pm\sqrt{2}$, and 2) and can be dealt with in a direct fashion. The shape-driven GUGA program of Saxe, Fox, Schaefer, and Handy⁹¹ is based in part on Siegbahn's approach, as is the program of Saunders and van Lenthe¹²⁷ and the COLUMBUS program of Shavitt, Lischka, Shepard, and co-workers.¹⁹⁶⁻¹⁹⁸

Unfortunately, these simplifications are not directly applicable⁹¹ when more than two electrons are allowed into the external space (e.g., CISDT, CISDTQ, and full CI). Furthermore, even for MR-CISD wavefunctions, large active spaces can lead to a large number of internal coupling coefficients, which can become difficult to deal with.¹⁹⁷ The next advance was once again due to Siegbahn, who suggested a factorization of the two-electron coupling coefficients by inserting the resolution of the identity:

$$\Gamma_{ijkl}^{IJ} = \sum_K \gamma_{ij}^{IK} \gamma_{kl}^{KJ} - \gamma_{il}^{IJ} \delta_{jk}. \quad (93)$$

Although the resolution of the identity requires an infinite sum in principle, in this case only a finite number of states $|\Phi_K\rangle$ are relevant. For fixed i, j, k, l, I , and J , the product term will vanish unless $|\Phi_K\rangle$ is obtained from $|\Phi_J\rangle$ by a single substitution from orbital l to orbital k and from $|\Phi_I\rangle$ by a single substitution from orbital i to orbital j . For configuration state functions, this completely specifies the orbital configuration of $|\Phi_K\rangle$, and only a few spin couplings must be summed over. This approach led Knowles and Handy¹⁰⁹ to present a vectorized full CI algorithm based on Slater determinants rather than CSFs. For Slater determinants, the one-electron coupling coefficients

appearing in (93) and elsewhere are very easy to calculate on-the-fly, allowing a fully direct CI procedure. For Slater determinants $|I\rangle$ and $|J\rangle$, $\gamma_{ij}^{IJ} = \langle I|\hat{E}_{ij}|J\rangle$ is 0 unless determinant $|J\rangle$ becomes determinant $|I\rangle$ (within a sign) when an alpha or beta electron is moved from orbital j to orbital i , in which case γ_{ij}^{IJ} becomes ± 1 . For the special case $i = j, I = J$, γ_{ii}^{II} counts the number of electrons in orbital i for determinant I , yielding 0, 1, or 2.

Because of this simplicity and the ability to carry out computations in a fully direct fashion, many of the full and restricted CI algorithms developed over the last ten years have employed Slater determinants, and in this section we will focus our attention on these determinant-based methods. However, we should note that further progress has also been made in CSF-based approaches. Much of the recent work in direct CI methods has been based on the symmetric group approach (SGA)^{199–203} rather than the related unitary group approach (UGA).^{40–42,78} Given the factorization (93), the problem of formulating a fully direct CI procedure can be turned into the problem of determining the one-electron coupling coefficients on-the-fly in the desired order. Knowles and Werner presented a way of doing this in 1988.¹⁰² They use the identity

$$\hat{E}_{ij} = \hat{E}_{ia}\hat{E}_{aj}, \quad (94)$$

which holds for any orbital ϕ_a which is always unoccupied. This hypothetical orbital, referred to as a “ghost” orbital, does not actually appear in any of the integrals. The one-electron coupling coefficient becomes

$$\gamma_{ij}^{IJ} = \sum_{K_s} \langle I|\hat{E}_{ia}|K_s\rangle \langle K_s|\hat{E}_{aj}|J\rangle, \quad (95)$$

where the sum is over all spin couplings of the uniquely-specified orbital configuration K . By fixing one of the two orbital indices, it becomes feasible to store the intermediates needed to evaluate the one-electron coupling coefficients efficiently in the desired order. This ghost-orbital technique was used by Werner and Knowles in their implementation of internally-contracted multireference CI.¹⁰¹ That method requires third- and fourth-order reduced density matrices, which can be evaluated by approaches analogous to (93). Another possibility along the lines of Siegbahn’s internal/external factorization has been suggested by Malmqvist, Rendell, and Roos in their implementation of the RAS SCF method.⁵⁶ They modify the GUGA method to split all walks into upper and lower portions and calculate coupling coefficients as products of upper and lower factors. Although the upper factors are not necessarily very simple for a RAS case, the storage requirements are substantially reduced in this approach.

Finally, we note that many other important advances have been made in CSF-based approaches, even outside the context of direct CI. Much of this effort in recent years has focused on extending the unitary and symmetric group

approaches to the spin-dependent Hamiltonians needed to account for relativistic effects.^{204–210} Examples of other work include specialized unitary group approaches for MRCI wavefunctions based on CAS references²¹¹ and application of the unitary group approach to CI calculations on atoms using Hylleraas coordinates⁶⁵ and to spin-adapted open-shell coupled-cluster theory.²¹² However, we now turn our attention to determinant-based formulations of direct CI.

4.2 Alpha and Beta Strings

A 1980 paper by Handy⁴⁴ represented a major advance in determinant-based CI, even though the paper was more concerned with how integrals and CI coefficients are stored than with the computational advantages of determinants over CSFs. Handy realized that if determinants are used as N -electron basis functions, and particularly if these determinants are expressed as “alpha strings” and “beta strings,” then the vector σ can be evaluated very efficiently.

Although Handy was the first to use alpha and beta strings, we will employ the subsequent notation of Olsen *et al.*⁴⁶ An alpha string is defined as an ordered product of creation operators for spin orbitals with alpha spin. If I_α contains a list $\{i, j, \dots k\}$ of the N_α occupied spin orbitals with alpha spin in determinant $|I\rangle$, then the alpha string $\alpha(I_\alpha)$ is $a_{i\alpha}^\dagger a_{j\alpha}^\dagger \dots a_{k\alpha}^\dagger$. A beta string is defined similarly. Thus a Slater determinant $|I\rangle$ in terms of alpha and beta strings is

$$|I\rangle = |\alpha(I_\alpha)\beta(I_\beta)\rangle = \alpha(I_\alpha)\beta(I_\beta)|\rangle. \quad (96)$$

For example, consider the Slater determinant $|I\rangle = |\phi_{1\alpha}\phi_{2\alpha}\phi_{3\alpha}\phi_{1\beta}\phi_{2\beta}\phi_{4\beta}\rangle$. Then the alpha string $\alpha(I_\alpha)$ is given by

$$\alpha(I_\alpha) = a_{1\alpha}^\dagger a_{2\alpha}^\dagger a_{3\alpha}^\dagger, \quad (97)$$

and the beta string is given by

$$\beta(I_\beta) = a_{1\beta}^\dagger a_{2\beta}^\dagger a_{4\beta}^\dagger. \quad (98)$$

Note that the order of the creation operators matters; if we swap the order of two creation operators within the alpha string (or within the beta string), then we introduce a sign change due to the anticommutation relation of creation operators. Also, applying the alpha string to the vacuum first, rather than the beta string, may introduce a minus sign, depending on the number of alpha and beta electrons. Typically, the beta string will be placed to the right of the alpha string in equations like (96). Further, within each string, orbitals are listed in strictly increasing order.

Since a determinant is now specified by an ordered pair of indices representing its alpha and beta components, the I th element of the CI vector becomes $c(I_\alpha, I_\beta)$. Note that this vector can also be considered as a matrix with coordinates I_α and I_β . Both vector and matrix addressing schemes are computationally useful. The σ vector,

$$\sigma_I = \sum_J H_{IJ} c_J, \quad (99)$$

can be written in the new notation as

$$\sigma(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \langle \beta(J_\beta) \alpha(J_\alpha) | \hat{H} | \alpha(I_\alpha) \beta(I_\beta) \rangle c(J_\alpha, J_\beta), \quad (100)$$

where we have used $H_{IJ} = H_{JI}^*$ and assumed that \mathbf{H} is a real matrix to obtain the form most commonly seen in the literature when this notation is used. Handy realized the following advantages to alpha and beta strings:

1. Direct CI methods often require an index vector which points to a list of all allowed excitations from a given N -electron basis function. Using alpha and beta strings, the index vector need not be the length of the CI vector—its size is dictated by the number of alpha or beta strings, which (for a full CI) is approximately the square root of the number of determinants. This results from the fact that in determinant-based CI, electrons in alpha spin-orbitals can be excited only to other alpha spin-orbitals, and electrons in beta spin-orbitals can be excited only to other beta spin-orbitals (because of the restriction to a single value of M_s).
2. To form $\sigma(I_\alpha, I_\beta)$ in equation (100), all functions $|\alpha(J_\alpha)\beta(J_\beta)\rangle$ which have non-zero matrix elements with $|\alpha(I_\alpha)\beta(I_\beta)\rangle$ are generated, one at a time, with the appropriate integral being looked up and multiplied by the appropriate CI coefficient. No time is wasted considering determinants which are noninteracting, and the coefficients of the integrals are easy to evaluate.
3. Efficiency is increased by realizing that all integrals which enter the expression $\langle \alpha(I_\alpha)\beta(I_\beta) | \hat{H} | \alpha(J_\alpha)\beta(I_\beta) \rangle$ (equation 100), where $\alpha(J_\alpha)$ differs from $\alpha(I_\alpha)$ by two orbitals, are independent of $\beta(I_\beta)$.

This approach allowed several benchmark full CI computations, including the first CI procedure (1981) to include more than one million determinants.^{38, 213}

4.3 The Vectorized Full CI Algorithm of Knowles and Handy

In 1984, Knowles and Handy introduced a new direct CI algorithm for full CI wavefunctions.^{109,214} As Siegbahn had pointed out,⁴⁵ the efficiency of direct CI algorithms is increased if, for a given $|K\rangle$, all one-electron coupling coefficients γ_{kl}^{KJ} are available together. Examination of Siegbahn's expression for σ elucidates this observation:⁴⁵

$$\sigma_I = \sum_J H_{IJ} c_J \quad (101)$$

$$= \sum_{ij} \sum_J \gamma_{ij}^{IJ} h_{ij} c_J + \frac{1}{2} \sum_{ijkl} (ij|kl) \sum_J \Gamma_{ijkl}^{IJ} c_J \quad (102)$$

$$= \sum_{ij} \{h_{ij} - \frac{1}{2} \sum_l (il|lj)\} \sum_J \gamma_{ij}^{IJ} c_J \\ + \frac{1}{2} \sum_{ijkl} (ij|kl) \sum_K \gamma_{ij}^{IK} \sum_J \gamma_{kl}^{KJ} c_J, \quad (103)$$

where the resolution of the identity has been used to turn the two-electron coupling coefficients into products of one-electron coupling coefficients (eq. 93). Notice that part of the two-electron contribution has been folded into the one-electron term. The remaining two-electron term is the time consuming part in the evaluation of σ_I , and it is most efficiently written as

$$\sigma_I^{(2)} = \frac{1}{2} \sum_K \sum_{ij} \gamma_{ij}^{IK} \sum_{kl} (ij|kl) \sum_J \gamma_{kl}^{KJ} c_J. \quad (104)$$

Thus this part of σ_I can be evaluated by the following set of operations:

$$D_{kl}^K = \sum_J \gamma_{kl}^{KJ} c_J \quad (105)$$

$$E_{ij}^K = \sum_{kl} (ij|kl) D_{kl}^K \quad (106)$$

$$\sigma_I^{(2)} = \frac{1}{2} \sum_K \sum_{ij} \gamma_{ij}^{IK} E_{ij}^K. \quad (107)$$

The one-electron coupling coefficients would ordinarily be stored on disk, making the evaluation of the D and σ quantities I/O intensive and thus inefficient. However, Knowles and Handy noted that in a basis of determinants the one-electron coupling coefficients can be evaluated on-the-fly (direct CI) even in the general case. A given determinant $|K\rangle$ can interact with at most two

other determinants $|I\rangle$, and their contributions can be separated by rewriting the shift operator in $\gamma_{ij}^{IK} = \langle I|\hat{E}_{ij}|K\rangle$ as

$$\hat{E}_{ij} = \hat{E}_{ij}^{\alpha} + \hat{E}_{ij}^{\beta}, \quad (108)$$

where \hat{E}_{ij}^{α} replaces an α -spin electron in orbital j with an α -spin electron in orbital i (cf. section 2.3.2). Note that this approach requires a sum over a complete set of intermediate determinants $|K\rangle$ (or at least all determinants which interact with the allowed determinants through \hat{E}_{ij}), including those determinants with the wrong spatial symmetry. This means that the Knowles and Handy algorithm would be considerably less efficient for restricted CI.

Given equation (108), it is possible to write the one-electron coupling coefficients in terms of alpha and beta strings:

$$\begin{aligned} \gamma_{ij}^{IJ} &= \langle \alpha(I_{\alpha})\beta(I_{\beta})|\hat{E}_{ij}^{\alpha}|\alpha(J_{\alpha})\beta(J_{\beta})\rangle\delta(I_{\beta}, J_{\beta}) \\ &+ \langle \alpha(I_{\alpha})\beta(I_{\beta})|\hat{E}_{ij}^{\beta}|\alpha(J_{\alpha})\beta(J_{\beta})\rangle\delta(I_{\alpha}, J_{\alpha}), \end{aligned} \quad (109)$$

where $|\alpha(I_{\alpha})\rangle$ is related to $|\alpha(J_{\alpha})\rangle$ by a single excitation (and likewise for $|\beta(I_{\beta})\rangle$ and $|\beta(J_{\beta})\rangle$). Thus the one-electron coupling coefficients are generated from lists of strings related by single excitations. For each alpha (beta) string, one stores a list of all allowed single replacements to other alpha (beta) strings; for closed-shell systems, the two lists will be identical and only one must be stored. Each list contains the address of the new string, the orbital index ij , and a phase factor, denoted by $\text{sgn}(ij)$, which is ± 1 . The sign can be determined as $(-1)^p$, where p is the number of transpositions of creation operators needed to bring an excited string to its canonical form. String addresses were computed by table lookups using a canonical addressing scheme explained in section 4.9.2. One can take advantage of the permutational symmetries $(ij|kl) = (ji|kl) = (ij|lk) = (ji|lk)$ of the two-electron integrals by requiring $i \geq j$, $k \geq l$. This entails replacing γ_{kl}^{KJ} in (105) by $(\gamma_{kl}^{KJ} + \gamma_{lk}^{KJ})$, an analogous change for (107), and a modification of the integrals to avoid double counting when $i = j$ or $k = l$.

The algorithm of Knowles and Handy is described as “vectorized” because each of the three major operations (105)-(107) may be written as an operation performed on an entire vector at once. This is very beneficial for vector supercomputers, which actually perform such operations a vector at a time and give substantial increases in speed. To illustrate, consider Fig. 6, which shows the Knowles-Handy algorithm for the formation of D , eq. (105). Due to memory limitations, operations are performed for a block of strings at a time. In the first half of Fig. 6, the operations in the innermost loop are identical but independent of each other for different K_{β} . In the second half of the algorithm, the same applies to K_{α} ; hence, this operation can be performed for a

Figure 6: Knowles-Handy Vectorized Formation of D (Refs. 109,214).

```

loop alpha strings  $K_\alpha$  in block
  loop over excitations  $|\alpha(J_\alpha)\rangle = \text{sgn}(ij)\hat{E}_{ij}^\alpha|\alpha(K_\alpha)\rangle$ 
    loop over  $K_\beta$  in block
       $|K\rangle = |\alpha(K_\alpha)\beta(K_\beta)\rangle, |J\rangle = |\alpha(J_\alpha)\beta(K_\beta)\rangle$ 
       $D(K_\alpha, K_\beta, ij) = D(K_\alpha, K_\beta, ij) + \text{sgn}(ij)c(J_\alpha, K_\beta)$ 

loop alpha strings  $K_\beta$  in block
  loop over excitations  $|\beta(J_\beta)\rangle = \text{sgn}(ij)\hat{E}_{ij}^\beta|\beta(K_\beta)\rangle$ 
    loop over  $K_\alpha$  in block
       $|K\rangle = |\alpha(K_\alpha)\beta(K_\beta)\rangle, |J\rangle = |\alpha(K_\alpha)\beta(J_\beta)\rangle$ 
       $D(K_\alpha, K_\beta, ij) = D(K_\alpha, K_\beta, ij) + \text{sgn}(ij)c(K_\alpha, J_\beta)$ 

```

whole range of K_β (K_α) values simultaneously with a vector processor. These same considerations apply to the analogous eq. (107). The remaining (and most time consuming) step, (106), can be performed as a matrix multiplication when one uses compound indices ij and kl , and of course this is also a vectorized operation.

These concerns about vectorization remain relevant even though quantum chemists now perform a substantial fraction of their computations on workstation machines which lack vector processors. Nevertheless, workstations (and now even personal computers, or PCs) feature *pipelined* processors. Pipelines allow machine instructions to overlap to some extent, giving the processor a limited ability to perform several tasks at once.¹⁶⁶ The superscalar IBM RS/6000 POWER2 workstation processor has two floating point pipelines, each of which can hold up to twelve instructions. If the processor can keep a steady stream of independent instructions coming down the pipeline, then overall performance will be increased substantially. However, if one instruction depends on results from another, then the pipeline can become stalled and performance is degraded. Since vectorizable code implies many similar but independent operations, as a general rule, vectorizable code becomes good pipelined code.

The Knowles-Handy approach was expected to be very efficient on vector supercomputers, and indeed it enabled many important full CI benchmark calculations.¹⁵ Nevertheless, one can see that this algorithm does more work than is strictly necessary. Equation (106) demonstrates that the operation count for the time-consuming step is approximately $\frac{1}{4}\tilde{N}_{det}n^4$, where n is the number of orbitals and \tilde{N}_{det} is the number of interacting intermediate states,

which may be larger than the number of determinants in the full CI space because the intermediate states are not subject to spatial symmetry restrictions. The factor of $\frac{1}{4}$ arises from the permutational symmetries of the integrals. As discussed previously in section 2.4.5, the computational cost of a full CI procedure should actually scale as $\mathcal{O}(N_{det}N^2n^2)$. Thus, the Knowles-Handy algorithm replaces N_{det} with the larger \tilde{N}_{det} , and N^2 with the larger n^2 . Some of the extra work is due to the fact that the intermediate matrix D can contain a substantial number of zeros. For example, D_{kl}^K will be zero when K has orbital k unoccupied or orbital l doubly occupied (D becomes less sparse under the condition $k \geq l$). Even though matrix multiplications are ideal for vector computers, Olsen and co-workers⁴⁶ realized that abandoning the matrix formulation (105)-(107) might still lead to a faster algorithm due to the substantially reduced operation count.

4.4 Olsen's String-Based Full CI Algorithm

In order to avoid the unnecessarily large operation counts in the full CI algorithm of Knowles and Handy, Olsen *et al.* abandoned the explicit use of a complete set of intermediate states and returned to some of Handy's original (1980) ideas⁴⁴ concerning string-driven full CI approaches.

4.4.1 Full CI σ Equations

We begin by describing Olsen's expressions for the σ vector. In second quantized-form (cf. section 2.3.2), \hat{H} becomes

$$\hat{H} = \sum_{kl}^n h_{kl} \hat{E}_{kl} + \frac{1}{2} \sum_{ijkl}^n (ij|kl) (\hat{E}_{ij} \hat{E}_{kl} - \delta_{jk} \hat{E}_{il}). \quad (110)$$

Inserting this expression into that for σ , eq. (100), yields

$$\begin{aligned} \sigma(I_\alpha, I_\beta) &= \sum_{J_\alpha, J_\beta} \langle \beta(J_\beta) \alpha(J_\alpha) | \sum_{kl}^n h_{kl} \hat{E}_{kl} \\ &+ \frac{1}{2} \sum_{ijkl}^n (ij|kl) (\hat{E}_{ij} \hat{E}_{kl} - \delta_{jk} \hat{E}_{il}) | \alpha(I_\alpha) \beta(I_\beta) \rangle c(J_\alpha, J_\beta) \end{aligned} \quad (111)$$

Now expanding the shift operators into their two spin components, $\hat{E}_{kl} = \hat{E}_{kl}^\alpha + \hat{E}_{kl}^\beta$, we write σ as a sum of three terms:⁴⁶

$$\sigma(I_\alpha, I_\beta) = \sigma_1(I_\alpha, I_\beta) + \sigma_2(I_\alpha, I_\beta) + \sigma_3(I_\alpha, I_\beta), \quad (112)$$

where

$$\begin{aligned} \sigma_1(I_\alpha, I_\beta) &= \sum_{J_\beta} \sum_{kl}^n \langle \beta(J_\beta) | \hat{E}_{kl}^\beta | \beta(I_\beta) \rangle \left[h_{kl} - \frac{1}{2} \sum_j^n (kj|jl) \right] c(I_\alpha, J_\beta) \\ &+ \frac{1}{2} \sum_{J_\beta} \sum_{ijkl}^n \langle \beta(J_\beta) | \hat{E}_{ij}^\beta \hat{E}_{kl}^\beta | \beta(I_\beta) \rangle (ij|kl) c(I_\alpha, J_\beta), \end{aligned} \quad (113)$$

$$\begin{aligned} \sigma_2(I_\alpha, I_\beta) &= \sum_{J_\alpha} \sum_{kl}^n \langle \alpha(J_\alpha) | \hat{E}_{kl}^\alpha | \alpha(I_\alpha) \rangle \left[h_{kl} - \frac{1}{2} \sum_j^n (kj|jl) \right] c(J_\alpha, I_\beta) \\ &+ \frac{1}{2} \sum_{J_\alpha} \sum_{ijkl}^n \langle \alpha(J_\alpha) | \hat{E}_{ij}^\alpha \hat{E}_{kl}^\alpha | \alpha(I_\alpha) \rangle (ij|kl) c(J_\alpha, I_\beta), \end{aligned} \quad (114)$$

$$\sigma_3(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \sum_{ijkl}^n \langle \beta(J_\beta) | \hat{E}_{ij}^\beta | \beta(I_\beta) \rangle \langle \alpha(J_\alpha) | \hat{E}_{kl}^\alpha | \alpha(I_\alpha) \rangle (ij|kl) c(J_\alpha, J_\beta). \quad (115)$$

For efficient implementation, it is convenient to precompute the quantities

$$h'_{kl} = h_{kl} - \frac{1}{2} \sum_j^n (kj|jl). \quad (116)$$

Note that the first term (σ_1) involves only beta shift operators, the second (σ_2) involves only alpha shift operators, and the third (σ_3) involves both alpha and beta shift operators. These terms are also called the $\beta\beta$, $\alpha\alpha$, and $\alpha\beta$ terms.^{48,49} Several determinant-based CI algorithms presented over the last few years^{83,183,215,216} have been based on this set of σ equations or the analogous equations for restricted CI (see section 4.8.1).

4.4.2 Simplifications for $M_s = 0$

Certain simplifications arise if the $M_s = 0$ component of an electronic state is used. The first of these is the time-reversal symmetry of the CI vector, which may be expressed as

$$c(I_\alpha, I_\beta) = (-1)^S c(I_\beta, I_\alpha), \quad (117)$$

where S is the spin quantum number. Olsen *et al.* use this fact to show how the σ_2 contribution can be determined entirely from the σ_1 contribution when $M_s = 0$.⁴⁶ The remarkably simple result is

$$\sigma_2(I_\alpha, I_\beta) = (-1)^S \sigma_1(I_\beta, I_\alpha). \quad (118)$$

Likewise, it is also possible to show that the $ijkl$ -th component of σ_3 satisfies the relation

$$\sigma_3^{ijkl}(I_\alpha, I_\beta) = (-1)^S \sigma_3^{klij}(I_\beta, I_\alpha). \quad (119)$$

This equation may be used to eliminate contributions from $I_\beta > I_\alpha$ or $(kl) > (ij)$, where (ij) and (kl) are compound indices. Olsen argues that the restriction $I_\alpha \geq I_\beta$ is to be preferred where this can be used to eliminate entire blocks of the σ_3 matrix. If all alpha/beta strings with the same irreducible representation are grouped together, then states which are not totally symmetric in their molecular point group will have off-diagonal blocks which can be eliminated using this restriction. On the other hand, when applied to totally symmetric states, this restriction eliminates the upper half of each symmetry block of σ_3 . Since this reduces the average vector length in the vectorized algorithm, Olsen recommends using the alternative restriction $(ij) \geq (kl)$ in these cases. This may be accomplished by rewriting σ_3 as

$$\begin{aligned} \sigma_3(I_\alpha, I_\beta) &= \sum_{(ij) \geq (kl)} \sigma_3^{ijkl}(I_\alpha, I_\beta) + \sum_{(ij) < (kl)} (-1)^S \sigma_3^{klij}(I_\beta, I_\alpha) \\ &= \sigma'_3(I_\alpha, I_\beta) + (-1)^S \sigma'_3(I_\beta, I_\alpha), \end{aligned} \quad (120)$$

where

$$\sigma'_3(I_\alpha, I_\beta) = \sum_{(ij) \geq (kl)} \sigma_3^{ijkl}(I_\alpha, I_\beta) (1 + \delta_{(ij), (kl)})^{-1}. \quad (121)$$

Hence the total σ vector can be evaluated as

$$\sigma(I_\alpha, I_\beta) = \sigma_1(I_\alpha, I_\beta) + \sigma'_3(I_\alpha, I_\beta) + (-1)^S [\sigma_1(I_\beta, I_\alpha) + \sigma'_3(I_\beta, I_\alpha)]. \quad (122)$$

The $M_s = 0$ simplifications therefore reduce computational expense by roughly a factor of two.

One further observation must be made about the loss of spin symmetry in the CI vector in the iterative diagonalization of the Hamiltonian. Even very slight deviations from (117), such as might occur from roundoff errors, become magnified in subsequent iterations and cause the iteration procedure to become numerically unstable because precise adherence to (117) is *assumed* if any of the $M_s = 0$ simplifications just described. If necessary, these difficulties can be avoided by explicitly enforcing the spin symmetry of any new vector in the subspace expansion. In this respect it is important to modify the diagonal elements of the Hamiltonian in the preconditioner for the subspace iteration method, as already discussed in section 3.2.

4.4.3 Algorithms for Computing σ

From (113), one can see that the mathematical operations required to form $\sigma_1(I_\alpha, I_\beta)$ are identical but independent of each other for different I_α . That is,

Figure 7: Olsen's Vectorized Algorithm for σ_1 (Ref. 46).

```

loop over beta strings  $I_\beta$ 
  Zero array  $F$ 
  Loop over excitations  $\hat{E}_{kl}^\beta$  from  $|\beta(I_\beta)\rangle$ 
     $|\beta(K_\beta)\rangle = \text{sgn}(kl)\hat{E}_{kl}^\beta|\beta(I_\beta)\rangle$ 
     $F(K_\beta) = F(K_\beta) + \text{sgn}(kl)h'_{kl}$ 
    Loop over excitations  $\hat{E}_{ij}^\beta$  from  $|\beta(K_\beta)\rangle$ 
       $|\beta(J_\beta)\rangle = \text{sgn}(ij)\hat{E}_{ij}^\beta|\beta(K_\beta)\rangle$ 
       $F(J_\beta) = F(J_\beta) + (1/2)\text{sgn}(kl)\text{sgn}(ij)(ij|kl)$ 
    end loop over  $\hat{E}_{ij}^\beta$ 
  end loop over  $\hat{E}_{kl}^\beta$ 
  loop over beta strings  $J_\beta$  and alpha strings  $I_\alpha$ 
     $\sigma_1(I_\alpha, I_\beta) = \sigma_1(I_\alpha, I_\beta) + F(J_\beta)c(I_\alpha, J_\beta)$ ; vect'd
  end loop over  $I_\alpha, J_\beta$ 
end loop over  $I_\beta$ 

```

column I_β of σ_1 can be constructed by two multiplications of scalars by columns (J_β) of c . Hence the construction of σ_1 is vectorizable over I_α . The vectorized algorithm for the evaluation of σ_1 , adapted from Olsen *et al.*,⁴⁶ appears in Figure 7. An analogous algorithm can be used to obtain σ_2 . However, one can also obtain σ_2 for $M_s = 0$ cases by (118). These algorithms require the same string replacement lists used by Knowles and Handy¹⁰⁹ (sec. 4.3). Note that the vector F is sparse, and multiplication of F by c should only take place for nonzero values of F .

Unfortunately, the construction of σ_3 (115) is harder to vectorize. A simple, non-vectorized algorithm for σ_3 is presented in Fig. 8. One can see that this does not appear as a simple set of arithmetic operations on vectors. For example, the contributions of the beta strings are not identical for different alpha strings because each alpha string connects to a different set of excited alpha strings with different indices k and l . Olsen *et al.* remedy this by operating a fixed kl at a time;⁴⁶ this makes their algorithm vectorizable in the innermost loop. Their algorithm, adapted and expanded from Ref. 46, is presented in Fig. 9. Note that this algorithm also employs scatter/gather (i.e., data rearrangement) operations to ensure that all of the data relevant to the multiplication step $V = Fc'$ are contiguous. This avoids "indirect addressing," which could substantially degrade performance due to long waits for data to be fetched from scattered memory locations.¹⁶⁶ For $M_s = 0$, an improvement to the σ_3 algorithm can be made by utilizing equations discussed in section 4.4.2.

Figure 8: Simple Algorithm for σ_3 .

```

loop over  $I_\alpha$ 
  loop over  $|\alpha(J_\alpha)\rangle = \text{sgn}(kl)\hat{E}_{kl}^\alpha|\alpha(I_\alpha)\rangle$ 
    loop over  $I_\beta$ 
      loop over  $|\beta(J_\beta)\rangle = \text{sgn}(ij)\hat{E}_{ij}^\beta|\beta(I_\beta)\rangle$ 
         $\sigma_3(I_\alpha, I_\beta) = \sigma_3(I_\alpha, I_\beta) + \text{sgn}(ij)\text{sgn}(kl)(ij|kl)c(J_\alpha, J_\beta)$ 
      end loop over  $J_\beta$ 
    end loop over  $I_\beta$ 
  end loop over  $J_\alpha$ 
end loop over  $I_\alpha$ 
    
```

Furthermore, if the integrals possess the full eightfold permutational symmetry, then \hat{E}_{kl}^α can be replaced by $(\hat{E}_{kl}^\alpha + \hat{E}_{lk}^\alpha)(1 + \delta_{kl})^{-1}$ in order to increase the average vector length in the formation of V . Note once again that F is sparse.

Clearly this algorithm takes less advantage of vector processors than the Knowles-Handy algorithm, since it involves some overhead (setup of the L and R arrays, and the scatter and gather) and uses smaller vector lengths. Nevertheless, one would expect this algorithm to be faster in many cases due to the substantially reduced number of mathematical operations performed. Counting only multiplications, the operation counts for each part of σ are approximately⁴⁶

$$N_1 \approx \frac{1}{4}N_{det}N_\beta^2(n - N_\beta)^2 \quad (123)$$

$$N_2 \approx \frac{1}{4}N_{det}N_\alpha^2(n - N_\alpha)^2 \quad (124)$$

$$N_3 \approx N_{det}N_\alpha N_\beta(n - N_\alpha)(n - N_\beta). \quad (125)$$

When $N_\alpha = N_\beta$, the overall operation count is thus approximately

$$N_{op} \approx \frac{3}{2}N_{det}N_\alpha^2(n - N_\alpha)^2. \quad (126)$$

Recall that this operation count can be cut approximately in half for $M_s = 0$ cases. Knowles and Handy are also able to take advantage of time reversal symmetry for singlet states, by employing the combinations $2^{-1/2}(\alpha(I_\alpha)\beta(I_\beta) + \alpha(I_\beta)\beta(I_\alpha))$. Recalling that the operation count for the Knowles-Handy algorithm is approximately $\frac{1}{4}\tilde{N}_{det}n^4$, we might expect the greatest savings for Olsen's algorithm when n/N is large.

Figure 9: Olsen's Vectorized Algorithm for σ_3 (Ref. 46).

```

loop over  $kl$ 
  set up lists  $L(I)$ ,  $R(I)$ , and  $\text{sgn}(I)$ , such that
     $|\alpha[L(I)]\rangle = \text{sgn}(I)\hat{E}_{kl}^\alpha|\alpha[R(I)]\rangle$ 
  loop over list entries  $I$  and beta strings  $J_\beta$ 
     $c'(I, J_\beta) = c(L(I), J_\beta)\text{sgn}(I)$ ; vect'd gather
  end loop over  $I$  and  $J_\beta$ 
  loop over  $I_\beta$ 
    zero array  $F$ 
    loop over excitations  $\hat{E}_{ij}^\beta$  from  $|\beta(I_\beta)\rangle$ 
       $|\beta(J_\beta)\rangle = \text{sgn}(ij)\hat{E}_{ij}^\beta|\beta(I_\beta)\rangle$ 
       $F(J_\beta) = F(J_\beta) + \text{sgn}(ij)(ij|kl)$ 
    end loop over  $\hat{E}_{ij}^\beta$ 
    loop over beta strings  $J_\beta$  and list entries  $I$ 
       $V(I) = F(J_\beta)c'(I, J_\beta)$ ; vect'd over  $I$ 
    end loop over  $J_\beta, I$ 
    loop over list entries  $I$ 
       $\sigma_3(R(I), I_\beta) = \sigma_3(R(I), I_\beta) + V(I)$ ; vect'd scatter
    end loop over  $I$ 
  end loop over  $I_\beta$ 
end loop over  $kl$ 

```

4.5 Zarrabian's Reduced Intermediate Space

Shortly after the publication of the 1988 paper by Olsen *et al.*, Zarrabian, Sarma, and Paldus presented²¹⁷ an alternative approach to avoid the unnecessarily large scaling of the Knowles-Handy full CI algorithm. These workers employed an $(N - 2)$ -electron intermediate space for the two-electron contributions rather than an N -electron intermediate space. Their expressions for σ were originally derived using generators of the group $SO(4)$, but to avoid introducing new notation we will consider the later derivation of Harrison and Zarrabian,⁴⁷ which uses only the standard spin-orbital creation and annihilation operators.

We begin by rewriting (24) over spatial orbitals, as

$$\hat{H} = \sum_{ij}^n h_{ij} \sum_{\sigma=\alpha,\beta} a_{i\sigma}^\dagger a_{j\sigma} + \frac{1}{2} \sum_{ijkl}^n (ij|kl) \sum_{\lambda,\sigma=\alpha,\beta} a_{i\sigma}^\dagger a_{k\lambda}^\dagger a_{l\lambda} a_{j\sigma}. \quad (127)$$

Next, insert the resolution of the identity between the pairs of creation and annihilation operators in the two-electron term. Clearly, the sum must run over $(N - 2)$ -electron states. The expression for σ becomes

$$\begin{aligned} \sigma_I &= \sum_J H_{IJ} c_J \\ &= \sum_J \sum_{ij}^n h_{ij} \sum_{\sigma=\alpha,\beta} \langle I^{(N)} | a_{i\sigma}^\dagger a_{j\sigma} | J^{(N)} \rangle c_J \\ &\quad + \frac{1}{2} \sum_{JK} \sum_{ijkl}^n (ij|kl) \sum_{\lambda,\sigma=\alpha,\beta} \langle I^{(N)} | a_{i\sigma}^\dagger a_{k\lambda}^\dagger | K^{(N-2)} \rangle \langle K^{(N-2)} | a_{l\lambda} a_{j\sigma} | J^{(N)} \rangle c_J, \end{aligned} \quad (128)$$

where the superscripts (N) and $(N - 2)$ denote the number of electrons for each state. The one-electron terms are exactly the same as before. The two-electron contributions to σ , denoted $\sigma^{(2)}$, may be written in terms of separate contributions from each possible spin case:

$$\begin{aligned} \sigma_{\alpha\alpha}^{(2)} &= \sum_{JK} \sum_{i>k,j>l}^n [(ij|kl) - (il|jk)] \langle I^{(N)} | a_{i\alpha}^\dagger a_{k\alpha}^\dagger | K^{(N-2)} \rangle \langle K^{(N-2)} | a_{l\alpha} a_{j\alpha} | J^{(N)} \rangle c_J \\ \sigma_{\beta\beta}^{(2)} &= \sum_{JK} \sum_{i>k,j>l}^n [(ij|kl) - (il|jk)] \langle I^{(N)} | a_{i\beta}^\dagger a_{k\beta}^\dagger | K^{(N-2)} \rangle \langle K^{(N-2)} | a_{l\beta} a_{j\beta} | J^{(N)} \rangle c_J \\ \sigma_{\alpha\beta}^{(2)} &= \sum_{JK} \sum_{ijkl}^n (ij|kl) \langle I^{(N)} | a_{i\beta}^\dagger a_{k\alpha}^\dagger | K^{(N-2)} \rangle \langle K^{(N-2)} | a_{l\alpha} a_{j\beta} | J^{(N)} \rangle c_J, \end{aligned} \quad (129)$$

where the restrictions over the orbital indices in the $\sigma_{\alpha\alpha}^{(2)}$ and $\sigma_{\beta\beta}^{(2)}$ terms is made possible by the permutational symmetry of the integrals and the anticommutation relations of the creation and annihilation operators. Likewise, the two

Figure 10: Harrison and Zarrabian's Vectorized Algorithm for $\sigma_{\alpha\alpha}^{(2)}$ (Ref. 47).

```

loop over alpha strings  $I_\alpha$ 
  loop over orbital pairs  $i > k$  (creation op.)
    define  $(N_\alpha - 2)$ -electron string  $K_\alpha$ 
     $|\alpha(K_\alpha)\rangle = \text{sgn}(ik)a_{k\alpha}a_{i\alpha}|\alpha(I_\alpha)\rangle$ 
    loop over orbital pairs  $j > l$  (annihilation op.)
      define new  $N_\alpha$ -electron string  $J_\alpha$ 
       $|\alpha(J_\alpha)\rangle = \text{sgn}(jl)a_{j\alpha}^\dagger a_{l\alpha}^\dagger |\alpha(K_\alpha)\rangle$ 
       $V = \text{sgn}(ik)\text{sgn}(jl)[(ij|kl) - (il|kj)]$ 
      loop over beta strings  $I_\beta$ 
         $\sigma_{\alpha\alpha}^{(2)}(I_\beta, I_\alpha) = \sigma_{\alpha\alpha}^{(2)}(I_\beta, I_\alpha) + V * c(I_\beta, J_\alpha)$ 

```

mixed contributions $\alpha\beta$ and $\beta\alpha$ have been combined in $\sigma_{\alpha\beta}^{(2)}$, eliminating the coefficient of $\frac{1}{2}$.

The algorithm for constructing $\sigma_{\alpha\alpha}^{(2)}$, adapted from Harrison and Zarrabian,⁴⁷ is given in Fig. 10. The algorithm for $\sigma_{\beta\beta}^{(2)}$ is of course analogous. It is easy to show that the number of floating-point multiplications involved in the construction of $\sigma_{\beta\beta}^{(2)}$ and $\sigma_{\alpha\alpha}^{(2)}$ with this algorithm are

$$N_1 = \frac{1}{4}N_{det}N_\beta(N_\beta - 1)(n - N_\beta + 2)(n - N_\beta + 1) \quad (130)$$

$$N_2 = \frac{1}{4}N_{det}N_\alpha(N_\alpha - 1)(n - N_\alpha + 2)(n - N_\alpha + 1), \quad (131)$$

which are basically the same as the approximate operation counts (123)-(124) for Olsen's algorithm.⁴⁶ Harrison and Zarrabian point out that this algorithm can be parallelized over the outermost loop. Note that they address the CI vector with the beta string as the row index instead of the alpha string; the earlier paper by Zarrabian *et al.* used the alternative convention. This choice can have some relevance for $\sigma_{\alpha\alpha}^{(2)}$ and $\sigma_{\beta\beta}^{(2)}$ when only one of the terms is explicitly constructed (i.e., when $M_s = 0$). In that case, it is best to access the data in C sequentially (i.e., with "unit stride").¹⁶⁶

For $\sigma_{\alpha\beta}^{(2)}$, one can use a similar loop structure to that in Fig. 9 or Fig. 10.²¹⁷ This yields an operation count⁴⁷ of

$$N_3 = N_{det}N_\alpha N_\beta (n - N_\alpha + 1)(n - N_\beta + 1). \quad (132)$$

However, Harrison and Zarrabian suggest that for parallel-vector machines, it is better to revert to a matrix multiplication such as that used by Knowles and Handy.¹⁰⁹ This algorithm is produced in Fig. 11. These loops are run

Figure 11: Harrison and Zarrabian's Vectorized Algorithm for $\sigma_{\alpha\beta}^{(2)}$ (Ref. 47).

```

precompute info for adding orbs to  $(N_\alpha - 1)$ -elec.  $\alpha$  string
precompute info for adding orbs to  $(N_\beta - 1)$ -elec.  $\beta$  string
zero  $D$ 
loop over orbitals  $l$  to be added to  $(N_\alpha - 1)$ -elec. string  $K_\alpha$ 
  loop over orbitals  $j$  to be added to  $(N_\beta - 1)$ -elec. string  $K_\beta$ 
    loop over  $(N_\alpha - 1)$ -electron strings  $K_\alpha$ 
      define  $N_\alpha$ -electron string  $J_\alpha$ 
       $|\alpha(J_\alpha)\rangle = \text{sgn}(l)a_{l\alpha}^\dagger|\alpha(K_\alpha)\rangle$ 
      loop over  $(N_\beta - 1)$ -electron strings  $K_\beta$ 
        define  $N_\beta$ -electron string  $J_\beta$ 
         $|\beta(J_\beta)\rangle = \text{sgn}(j)a_{j\beta}^\dagger|\beta(K_\beta)\rangle$ 
         $D(K_\beta, K_\alpha, jl) = D(K_\beta, K_\alpha, jl) + \text{sgn}(k)\text{sgn}(j)c(J_\beta, J_\alpha)$ 
      end loop over  $K_\beta$ 
    end loop over  $K_\alpha$ 
  end loop over  $j$ 
end loop over  $l$ 

call optimized matrix multiply for  $E_{K,ik} = D_{K,jl}\langle jl|ik\rangle$ 

loop over orbitals  $k$  to be added to  $(N_\alpha - 1)$ -elec. string  $K_\alpha$ 
  loop over orbitals  $i$  to be added to  $(N_\beta - 1)$ -elec. string  $K_\beta$ 
    loop over  $(N_\alpha - 1)$ -electron strings  $K_\alpha$ 
      define  $N_\alpha$ -electron string  $I_\alpha$ 
       $|\alpha(I_\alpha)\rangle = \text{sgn}(k)a_{k\alpha}^\dagger|\alpha(K_\alpha)\rangle$ 
      loop over  $(N_\beta - 1)$ -electron strings  $K_\beta$ 
         $|\alpha(I_\alpha)\rangle = \text{sgn}(k)a_{k\alpha}^\dagger|\alpha(K_\alpha)\rangle$ 
        define  $N_\beta$ -electron string  $I_\beta$ 
         $|\beta(I_\beta)\rangle = \text{sgn}(i)a_{i\beta}^\dagger|\beta(K_\beta)\rangle$ 
         $\sigma_{\alpha\beta}^{(2)}(I_\beta, I_\alpha) = \sigma_{\alpha\beta}^{(2)}(I_\beta, I_\alpha) + \text{sgn}(i)\text{sgn}(k)E(K_\beta, K_\alpha, ik)$ 
      end loop over  $K_\beta$ 
    end loop over  $K_\alpha$ 
  end loop over  $i$ 
end loop over  $k$ 

```

for blocks of several intermediate states K at a time, and additional loops account for spatial symmetry. Note the use of two-electron integrals in Dirac notation rather than Mulliken notation; i.e., $\langle jl|ik\rangle = (ij|kl)$ for real integrals. Furthermore, the integrals are stored without any permutational symmetry. This algorithm has an operation count of

$$N_3 = N_{det} N_\alpha N_\beta n^4 / (n - N_\alpha + 1)(n - N_\beta + 1), \quad (133)$$

which can be obtained⁴⁷ by using $N_\alpha/(n - N_\alpha + 1)$ as the ratio of the number of $(N_\alpha - 1)$ -electron strings to the number of N -electron strings (and by ignoring spatial symmetry). Note that this operation count is not too much greater than that of the non-matrix version (132) when $m \gg N_\alpha, N_\beta$. In such cases, and given $N_\alpha = N_\beta$, the overall number of multiplications $N_1 + N_2 + N_3$ is about $\frac{3}{2} N_{det} N_\alpha^2 n^2$, compared to $\frac{1}{4} \tilde{N}_{det} n^4$ for the Knowles-Handy algorithm.

Although the work done by this algorithm is basically equivalent to that done in Olsen's algorithm, Zarrabian *et al.* suggest^{47,217} that their approach would be better suited for the evaluation of three- and four-electron reduced density matrices, which are important in the context of internally contracted MR-CISD.¹⁰¹ They also note that it should be possible to adapt their algorithm to restricted CI spaces,^{47,217} and some work along these general lines has been presented by Duch.²⁰³

4.6 The Table-Based Algorithm of Bendazzoli and Evangelisti

Using Handy's alpha and beta string formalism,⁴⁴ along with some of the notation of Olsen,⁴⁶ Bendazzoli and Evangelisti have presented a full CI algorithm^{48,49} which uses tables to represent the excitation operators \hat{E}_{ij}^β rather than the string replacement lists of Knowles and Handy. The operation count of their method is essentially the same as that of Olsen *et al.*⁴⁶ and of Zarrabian *et al.*,^{47,217} but the data are organized differently and the authors note that their loop structure is more easily parallelized than that of Olsen *et al.*⁴⁶ The algorithm of Bendazzoli and Evangelisti⁴⁸ for σ_1 (which they call the $\beta\beta$ term), is presented in Figure 12. When $M_s = 0$, σ_2 can be obtained from (118) just as in Olsen's approach.⁴⁶

The tables $OOVV$ represent the shift operator products $\hat{E}_{il}^\beta \hat{E}_{jk}^\beta$; for a given set of orbitals (i, j, k, l) , $OOVV(i, j, l, k)$ gives a list of all beta strings with orbitals i, j occupied and l, k unoccupied. This is the list of all strings which can be acted on to the left by the shift operator product. Similarly, $OOVV(l, k, i, j)$ gives a list of all strings which can be acted on to the right by this same product. The clever aspect of this approach is that the I th entry of $OOVV(i, j, l, k)$ (denoted I_1) is the *same* as the string produced by applying $\hat{E}_{il}^\beta \hat{E}_{jk}^\beta$ to the I th

Figure 12: Bendazzoli and Evangelisti's Algorithm for σ_1 (Ref. 48).

```

loop over  $i > j, k > l$ 
   $V = (ik|jl) - (il|jk)$ 
  loop over  $I = 1$ , length of list  $OOVV(i, j, k, l)$ 
     $I_1 = I$ th entry of list  $OOVV(i, j, l, k)$ 
     $S_1 =$  sign associated with  $I_1$ 
     $VS = V * S_1$ 
     $I_2 = I$ th entry of list  $OOVV(l, k, i, j)$ 
    loop over  $J = 1$ , number of alpha strings
       $\sigma(J, I_2) = \sigma(J, I_2) + c(J, I_1) * VS$ 
    end loop over  $J$ 
  end loop over  $I$ 
end loop over  $i, j, k, l$ 

```

element of $OOVV(l, k, i, j)$ (denoted I_2). Bendazzoli and Evangelisti have so far limited their attention to full CI; for restricted CI, the size of the lists OV and $OOVV$ will rapidly become large relative to the size of the CI vector (sec. 4.9.4), so that these lists are probably appropriate only for full CI.

The σ_3 algorithm is presented in Figure 13. Note the same scatter/gather structure as in Figures 9 and 22. Like our own version (cf. section 4.9.5), this algorithm eliminates the F array and uses a DAXPY operation¹⁶⁶ in the innermost loop. Compared to the algorithm in Figure 22, our initial attempts to implement this algorithm for σ_3 yielded a program running roughly 50% slower on the IBM RS/6000 POWER2 model 3CT workstation.

More recently, Evangelisti, Bendazzoli, and co-workers have developed a parallel implementation of their algorithm for the Cray T3D, a distributed memory machine.^{50,84} The newest out-of-core version of their program allows the CI and σ vectors to be processed one symmetry block at a time. To avoid storage of the diagonal of the Hamiltonian, they approximate it using orbital eigenvalues. Following Olsen⁸³ (sec. 3.2.2), they minimize storage space by using only one CI vector and one σ vector in their iterative diagonalization method, although the details of their iterative procedure differ somewhat from those of Olsen and co-workers. In 1996, this parallelized version was used on a 64-processor Cray T3D to obtain⁸⁴ the full CI wavefunction for Be_2 , with all electrons correlated and using a 9s2p1d basis (derived from a 4s2p1d ANO basis by uncontracting the primitive Gaussians corresponding to the five largest coefficients in the first ANO orbital). This represents the first converged CI calculation requiring more than one billion Slater determinants

Figure 13: Bendazzoli and Evangelisti's Algorithm for σ_3 (Ref. 48).

```

loop over  $k, l$ 

  loop over  $I = 1$ , length of list  $OV(l, k)$ 
     $I_2 = I$ th entry of  $OV(l, k)$ 
     $S_2 =$  sign associated with  $I_2$ 
    loop over  $J = 1$ , number of beta strings
       $c'(I, J) = c(I_2, J) * S_2$ 
    end loop over  $J$ 
  end loop over  $I$ 

  loop over  $i, j$ 
     $V = (ij|kl)$ 
    loop over  $J = 1$ , length of list  $OV(i, j)$ 
       $J_1 = J$ th entry of  $OV(i, j)$ 
       $J_2 = J$ th entry of  $OV(j, i)$ 
       $S_2 =$  sign associated with  $J_2$ 
       $VS = V * S_2$ 
      loop over  $I = 1$ , length of  $OV(l, k)$ 
         $\sigma'(I, J_2) = \sigma'(I, J_2) + c'(I, J_1) * VS$ 
      end loop over  $I$ 
    end loop over  $J$ 
  end loop over  $i, j$ 

  loop  $I = 1$ , length of list  $OV(l, k)$ 
     $I_1 = I$ th entry of  $OV(l, k)$ 
    loop  $J = 1$ , number of beta strings
       $\sigma(I_1, J) = \sigma'(I, J)$ 
    end loop over  $J$ 
  end loop over  $I$ 

end loop over  $k, l$ 

```

(an unconverged calculation on the Mg atom involving more than a billion determinants was reported in 1990 by Olsen, Jørgensen, and Simons⁸³).

4.7 Approximate Full CI Methods

In 1989, Knowles introduced⁸¹ a modified full CI procedure which exploits the sparsity of the Hamiltonian matrix and affords approximate full CI results at a dramatically reduced computational cost. Employing the Davidson method¹⁰⁸ (cf. section 3.2.1), the correction to the current CI vector is given by

$$\Delta c_I = \frac{r_I}{(E - H_{II})}, \quad (134)$$

where r_I is the residual $\sigma_I - E c_I$. Knowles estimates the importance of these corrections using the following simple expression inspired by second-order perturbation theory:

$$\Delta E_I = r_I \Delta c_I. \quad (135)$$

If $|\Delta E_I|$ is less than some threshold, Δc_I is neglected. Thus far fewer determinants are actually included in the correction vector, which is stored on disk in a packed format. One problem with this approach is that neglected corrections Δc_I can reappear during the standard Schmidt orthogonalization against previous subspace vectors (cf. section 3.2). Knowles thus avoids the Schmidt orthogonalization step and employs a non-orthogonal space of expansion vectors. This allows for tight control over the size of the expansion space vectors.

A potential difficulty of this approach is that the σ vectors (which must also be stored) are not necessarily sparse. Knowles notes⁸¹ that even when \mathbf{c} is only 1% populated, typically 50% of σ will be nonzero. Nevertheless, in order to obtain variationally correct energies, the full σ vector must be formed in core memory and its dot product taken with all expansion vectors \mathbf{c} . However, once this is done, the only further use of σ is in the construction of new subspace vectors; hence, Knowles only writes to disk those elements of σ greater than some threshold. According to (134)-(135), these neglected elements of σ would only contribute to elements of $\Delta \mathbf{c}$ which make very small energy contributions.

Given the Knowles-Handy full CI algorithm of section 4.3, it is clear that the matrix formulation no longer applies with a sparse CI vector \mathbf{c} . Instead, the formation of σ is driven from the list of nonzero elements in \mathbf{c} , employing scatter and gather operations to obtain some vectorization in the innermost loops; this approach is therefore similar to the original string-driven approach of Handy⁴⁴ or the subsequent algorithm of Olsen *et al.*⁴⁶ To avoid core storage problems, the exact σ can be formed one symmetry block at a time (where a symmetry block of σ contains all elements σ_I having the same alpha string

symmetry). Memory requirements can be further reduced, with some loss in efficiency, by processing σ in smaller batches of arbitrary size.⁸¹

Knowles and Handy demonstrated the power of this approach by estimating the full CI energy of NH_3 in an atomic natural orbital (ANO) basis set of DZP quality.²¹⁸ The full CI expansion contains more than 209 million determinants, yet Knowles and Handy were able to obtain an apparently reliable variational energy of -56.4235 hartree using a CI vector with only 665,247 nonzero elements (0.3% of the full CI vector). Employing perturbation theory to estimate the remaining energy error (presumably via equation 135), Knowles and Handy arrived at a final estimate of -56.4236 ± 0.0001 hartree.²¹⁸

Using perturbation theory to estimate the importance of determinants in configuration interaction is a very old idea (see Ref. 57 for a detailed review). Indeed, it is perturbation theory which provides the justification for truncating the CI space at only singles and doubles from one or several references (i.e., the CISD and MR-CISD methods). The CIPSI method (1973) of Huron, Malrieu, and Rancurel⁹² diagonalizes the Hamiltonian in some subspace of selected determinants and uses the resulting eigenvector as the zeroth-order wavefunction in a subsequent perturbation theory treatment. Determinants having a contribution to the first-order wavefunction greater than some threshold η are added to the selected CI space, and this process is repeated until the selection threshold is considered acceptably small or until the selected CI space becomes too large to handle. The effect of unselected determinants is evaluated by second-order perturbation theory. The procedure of Knowles^{81,218} is similar to this, but differs in two important respects: first, Knowles selects determinants based on a perturbative estimate of their contribution to the energy rather than to the first-order wavefunction, and second, Knowles applies the selection *during* the Davidson procedure, whereas CIPSI solves the CI problem exactly for each selected CI space. A more recent version of the CIPSI method⁹³ is somewhat more flexible and introduces a third class of determinants of intermediate importance; interacting determinants with an estimated CI coefficient less than η but greater than a second threshold τ can be treated by higher-order perturbation theory or variationally, while those with contributions less than τ are treated by second-order perturbation theory as before. The CIPSI scheme should yield wavefunctions approaching the full CI limit, and indeed it has been benchmarked against full CI.^{93,122,123,215,219} The most recent studies have added a self-consistent dressing of the Hamiltonian matrix to ensure size consistency.^{122,123}

Another long-established approach to approximating full CI is to employ successively larger MR-CISD spaces. Since the size of the CI space grows very rapidly as the number of references is increased, Buenker and Peyerimhoff (1974-5)^{10,11} suggested retaining only the most important singly and doubly

substituted configurations and treating discarded configurations by Brillouin-Wigner perturbation theory and extrapolation procedures; they call their procedure MRD-CI. Their strategy implies that the most compact wavefunctions are obtained by truncating the singles and doubles space rather than the reference space, and indeed CIPSI studies support this idea.^{122,219} Unlike the CIPSI method, Buenker and Peyerimhoff do not use perturbation theory in the configuration selection; rather, orbital configurations are accepted or rejected on the basis of the energy lowering they cause when added to the reference space. A separate small CI procedure is required for each possible spatial orbital configuration. Although this may require somewhat more effort than the perturbational estimates of CIPSI, the energy lowerings can be reused in the extrapolations to zero threshold.^{10,11} Alternative approaches to making MR-CISD more computationally tractable are the internal and external contraction schemes discussed in section 2.4.2.

Knowles' 1989 program^{81,218} was able to approach the full CI limit more closely than selected CI methods such as MRD-CI and CIPSI because it was efficient enough to treat a much larger number of determinants variationally. Subsequently in 1992, Povill, Rubio, and Illas noted²¹⁵ that the principal difficulty with the standard CIPSI program was its need to store the Hamiltonian matrix \mathbf{H} , allowing it to handle no more than 50,000 determinants variationally. Hence, they presented²¹⁵ the direct selected configuration interaction using strings (DISCIUS) algorithm employing the alpha and beta string formalism of Handy,⁴⁴ the notation, $(\alpha\alpha, \beta\beta, \alpha\beta)$ spin decomposition, and $M_s = 0$ simplifications of Olsen *et al.*,⁴⁶ and the reduced intermediate space of Zarrabian *et al.*^{47,217} Special ordering and addressing schemes, which make use of large index arrays, allow for some degree of vectorization despite the lack of a well-defined structure in the CI space.²¹⁵ Nevertheless, a more recent (1995) version of this algorithm by Povill and Rubio²²⁰ largely abandons the vectorization of σ_3 , noting that the average vector length for selected CI spaces is generally too small for effective vectorization. These authors also found that too much time is spent checking to see if doubly excited strings in the construction of σ_1 or σ_2 belong to the selected space; hence, they consider every pair of allowed strings and determine all single and double excitations connecting them. The DISCIUS algorithm is capable of treating selected CI spaces with more than one million determinants.²²⁰

A related algorithm, which has also been coupled to the CIPSI approach, was presented by Caballol and Malrieu²²¹ in 1992. Their approach is also direct and determinant-based, but the strings are written as particle-hole excitations from a single reference state; the program is named SCIEL, for selected CI with excitation labeling. For a determinant with excitation level m , the particle-hole labeling lists m holes and m particles. This is inefficient for full CI,²²¹ since

it would require the listing of $2N_\alpha$ orbitals for a maximally-excited ($m = N_\alpha$) alpha string, rather than only N_α orbitals in the standard approach. However, for CI spaces dominated by determinants with a relatively low excitation level, this formalism could offer some benefits. Povill *et al.* have commented that the DISCIUS and SCIEL programs seem to have similar efficiencies.¹²²

Similar improvements have been made to the MRD-CI program of Buenker and Peyerimhoff,^{10,11} which was previously limited to about 50,000 configurations.²²² In 1995, Krebs and Buenker presented²²² a new table-direct CI algorithm for use in the MRD-CI selection scheme which is capable of handling variational spaces including at least several hundred thousand determinants.

Knowles' 1989 sparse CI method has been the subject of additional study in the last few years. In 1994, Mitrushenkov presented a very similar method¹⁸³ which differs primarily in that it selects components of the CI vector based on their magnitude (134) and not on their expected energy lowering (135). This choice was motivated by the belief that it would yield more physical CI vectors less likely to give errors for properties other than the total energy.¹⁸³ Mitrushenkov described how to adapt Olsen's full CI algorithm to implement his approach, which he has called dynamic CI. Of particular interest is his technique for avoiding core storage of the entire σ vector: he calculates $\sigma(I_\alpha, I_\beta)$ for a fixed I_β (i.e., the algorithm is driven by σ rather than by nonzero elements of c). The exact σ values are used to update the Hamiltonian in the small Davidson subspace, and then components larger than a given cutoff are written to disk. Like Knowles, Mitrushenkov uses a nonorthogonal Davidson subspace; however, he uses only two vectors and employs the improved preconditioner of Olsen *et al.*⁸³ (cf. section 3.2.2). Mitrushenkov reported results for NH₃, H₂O, and Mg test cases,¹⁸³ but unfortunately no results were presented for systems where the exact full CI result was known (DZP NH₃ full CI results have subsequently been reported,^{50,80} see below).

In 1991, Harrison emphasized the use of second-order perturbation theory to approach the full CI results more rapidly.²²³ In Harrison's method, denoted CI+PT, one chooses an initial reference space (perhaps a single determinant), and an initial selection threshold η . A CI is performed in the reference space, yielding eigenvectors for all roots of interest. Unlike most of the other algorithms discussed in this section, Harrison's program employs CSFs rather than determinants; two-electron coupling coefficients are evaluated as products of one-electron coupling coefficients, which in turn are evaluated by the method of Knowles and Werner¹⁰² (see section 4.1). For every configuration I interacting with (but not included in) the reference space, the second-order perturbation theory energy contribution is determined for each of the desired roots k , using

the selected-space eigenvectors \mathbf{c}^k as zeroth-order solutions:

$$\Delta E_I^k = \frac{|\langle I | \hat{H} | \mathbf{c}^k \rangle|^2}{E^k - \langle I | \hat{H} | I \rangle}. \quad (136)$$

The set of all configurations I with $|\Delta E_I^k| > \eta$ is added to the reference space, and the new Hamiltonian is diagonalized; this process is repeated until no new important configurations are found. The reference selection is thus self-consistent, and Harrison notes²²³ that typically only two iterations are required. The sum of the estimated energy lowerings (136) provides a perturbative correction for excluded configurations, and the total energy (variational energy plus perturbative correction) is an approximation to the full CI energy. An increasingly accurate sequence of wavefunctions is generated by repeating this whole process for a series of thresholds η with decreasing values. Note, however, that those configurations which do not interact directly with any reference function are completely neglected. Although the individual contributions of such configurations should be very small, there are a large number of them. Harrison finds that the CI+PT method approaches the full CI energy from above, suggesting that the error due to the neglect of noninteracting configurations is greater than the error due to the perturbative estimates of configurations which are interacting but not included in the reference space (the latter correction, being nonvariational, could conceivably lead to energies below the full CI limit). Harrison's approach is similar to the original two-class CIPSI algorithm,⁹² but it selects references based on their contributions to the energy rather than to the first-order wavefunction; Harrison notes that for properties other than the energy, this might not be the optimal choice.²²³ An advantage of his program was that it could handle larger variational spaces than the versions of the CIPSI and MRD-CI programs available at that time (recent improvements in these programs are described above). Results for H_2O , O , and O^- were compared to full CI and indicate that the perturbation theory energy correction rapidly accelerates convergence to the full CI energy. For example, the full CI energy for H_2O at three different geometries was obtained within $0.1 \text{ kcal mol}^{-1}$ with a variational reference space and perturbative interacting space spanning less than 0.23% and 25%, respectively, of the full CI space (these results correspond to a selection threshold of $\eta = 4 \times 10^{-7}$).

The wavefunction operator (WFO) approach of Luzanov, Wulfow, and Krouglov (1992)²²⁴ seems to be the same as Harrison's CI+PT approach, although it is formulated differently and implemented using determinants. This method appears to involve the same amount of work as other determinant-based sparse CI methods,^{215,220,221} but it uses rather different intermediate arrays. Explicit algorithms for the WFO method were presented in 1996 by Wulfow,²²⁵ who obtained results for HF dimer in a 4s3p1d/2s1p basis using a

personal computer (PC) equipped with only 16 megabytes (MB) of RAM and 50 MB of disk space. The largest computation, with threshold $\eta = 10^{-7}$, considered 65,751 determinants variationally and treated an unreported number of interacting determinants perturbatively.

Perturbation theory corrections to variational energies have also been considered recently by Mitrushenkov and Dmitriev (1995),²¹⁶ who express the second-order energy correction as

$$\begin{aligned}\Delta E &= \sum_{I_\alpha I_\beta} \frac{\sigma_0(I_\alpha, I_\beta)[\sigma_0(I_\alpha, I_\beta) - E_0 c_0(I_\alpha, I_\beta)]}{E_0 - H_{sa}^D(I_\alpha, I_\beta)} \\ &= (\mathbf{c}_0, \hat{H} \Delta \mathbf{c}),\end{aligned}\quad (137)$$

where \mathbf{c}_0 is the current CI vector and $\Delta \mathbf{c}_0$ is the correction vector. The term $H_{sa}^D(I_\alpha, I_\beta)$ is the diagonal element of the Hamiltonian for determinant $|I_\alpha I_\beta\rangle$, and the subscript *sa* indicates an average over all determinants with the same spatial orbital configuration (cf. section 3.2.1). This correction is essentially the same used by Harrison²²³ and others. Rather than employ this expression exactly as it is, Mitrushenkov and Dmitriev note²¹⁶ that for a converged full CI vector $\mathbf{c} = \mathbf{c}_0 + \Delta \mathbf{c}$, the full CI energy can be expressed as:

$$\begin{aligned}E &= \frac{(\mathbf{c}_0, \hat{H} \mathbf{c})}{(\mathbf{c}_0, \mathbf{c})} \\ &= \frac{E_0 + (\mathbf{c}_0, \hat{H} \Delta \mathbf{c})}{1 + (\mathbf{c}_0, \Delta \mathbf{c})}.\end{aligned}\quad (138)$$

Mitrushenkov and Dmitriev designate this the norm-consistent zero threshold full CI estimate. When $\Delta \mathbf{c}$ is orthogonal to \mathbf{c} , this is the second-order perturbation theory estimate (137). However, since selection in the dynamic CI occurs during the subspace iteration process, $\Delta \mathbf{c}$ and \mathbf{c} are not orthogonal. Of course the exact σ_0 and $\Delta \mathbf{c}$ must be used in the evaluation of (138); these are evaluated for fixed I_β as described above. Mitrushenkov and Dmitriev demonstrated this zero threshold energy (ZTE) estimate for Ne, NH₃, Mg, and H₂O; for Ne, comparisons with full CI indicate that the ZTE estimates approach the correct energy much faster than the variational energies with respect to decreasing threshold, but the approach is not monotonic and can occur from below. Very recently, Mitrushenkov has extended this approach to perform the dynamic CI+ZTE in the active space and to treat the external space using only second-order perturbation theory.²²⁶

In order to compare some of these selected CI methods, we present results for DZP NH₃ in Table 5. These results were obtained in point group C_s with the 1a' core orbital frozen (i.e., constrained to remain doubly occupied). The

Table 5: Selected and Full Configuration Interaction Benchmarks for NH_3 with an ANO DZP Basis Set.^a

Threshold	N_{CI}	E_{CI}	$\frac{E_{CI}-E_{FCI}}$	$\frac{N_{PT}}{\times 10^6}$	E_{CI+PT}	$\frac{E_{CI+PT}-E_{FCI}}$
Knowles and Handy, Ref. 218: ^b						
2.0×10^{-4}	171 867	-56.4219	0.0021	n/a	-56.4220	0.0020
1.0×10^{-4}	393 666	-56.4229	0.0011	n/a	-56.4230	0.0010
5.0×10^{-5}	450 763	-56.4232	0.0008	n/a	-56.4234	0.0006
4.0×10^{-5}	665 247	-56.4235	0.0005	n/a	-56.4236	0.0004
Harrison, Ref. 223: ^c						
1.0×10^{-4}	786	-56.390631	0.033376	0.61	-56.422854	0.001153
1.3×10^{-5}	1 889	-56.411996	0.012011	1.04	-56.423766	0.000241
1.6×10^{-6}	5 814	-56.417497	0.006510	2.93	-56.423748	0.000259
3.9×10^{-7}	18 921	-56.420203	0.003804	5.92	-56.423719	0.000288
2.0×10^{-7}	32 288	-56.421211	0.002796	7.68	-56.423737	0.000270
Povill <i>et al.</i> , Ref. 123: ^d						
6.0×10^{-6}	1.00×10^6	-56.423659	0.000348	2.49	-56.423681	0.000326
0.4×10^{-6}	1.17×10^6	-56.423785	0.000222	6.55	-56.423824	0.000183
0.1×10^{-6}	1.25×10^6	-56.423825	0.000182	15.10	-56.423875	0.000132
Mitrushenkov and Dmitriev, Refs. 183, 216: ^{b,e}						
1.0×10^{-3}	2 000	-56.4085	0.0155	n/a	-56.42397	0.00004
1.0×10^{-4}	34 000	-56.4195	0.0045	n/a	-56.42392	0.00009
1.0×10^{-5}	590 000	-56.4235	0.0005	n/a	-56.42400	0.00001

^aBasis set and geometry of Ref. 218. Only valence electrons are correlated. N_{CI} denotes the size of the variational space, and N_{PT} denotes the size of the interacting space treated by second-order perturbation theory. All energies are given in hartree.

^bNumber of nonzero elements in final CI vector.

^cDimensions are given in CSFs instead of determinants.

^dThe CIPSI algorithm uses two thresholds, η and τ . η is given, and $\tau = 10^{-10}$.

^eMitrushenkov reports the number of nonzero elements in the CI vector as a percentage of the full CI space (209 626 425 determinants); the N_{CI} values given are thus only approximate.

first highly accurate NH_3 benchmark using a basis set this large was that of Knowles and Handy in their 1989 demonstration of their new selected CI method.²¹⁸ Hence, most subsequent selected CI benchmarks on this system have used their geometry and atomic natural orbital basis set. An exception is the wavefunction operator (WFO) benchmark,²²⁴ which employed a different basis set; that method is therefore excluded from the table. The exact full CI energy of NH_3 with this basis set was unavailable when these selected CI benchmarks were published, and the Knowles-Handy²¹⁸ extrapolated selected CI estimate of -56.4236 hartree has sometimes been used in place of the full CI value. The apparent convergence of this estimate led Knowles and Handy to propose error bars of ± 0.0001 hartree. However, in 1994 Povill *et al.*¹²³ used the 3-class CIPSI method⁹³ to select the most important 1.25 million determinants and obtained a variational energy of -56.423825 hartree using the DISCIUS algorithm.²¹⁵ This energy is lower than the lowest estimate of Knowles and Handy, including the error bar. This difficulty was cleared up in 1995 by Evangelisti *et al.*,⁵⁰ who used the full CI program of Bendazzoli and Evangelisti^{48,49} to obtain an energy of -56.424007 hartree (this same value was obtained independently by Olsen⁸⁰). This demonstrates that the energy of Knowles and Handy was not converged as tightly as expected, and that it is easy to underestimate the importance of a large number of neglected determinants.

The data in Table 5 show that the perturbation theory corrected energies are better approximations to the full CI energy than the purely variational results, and that the full CI energy is always approached from above; these conclusions are in general agreement with previous benchmarks for smaller systems.^{93,216,223,224} However, we note that the perturbation theory corrections become less effective for large variational spaces, and conversely, the CI+PT energies are slowly convergent (and not monotonic) with the relatively small CI spaces used in Harrison's study²²³ (note, however, that Harrison's N_{CI} values are in CSFs rather than determinants, making his variational space look smaller than it actually is). The norm-consistent zero threshold energies of Mitrushenkov and Dmitriev²¹⁶ appear particularly effective, although they do not approach the full CI energy monotonically.

Table 5 demonstrates that it is difficult to establish the convergence of the energy for selected CI methods. Another problem which has received relatively little attention is the convergence of properties other than the energy. It would be expected that other properties should not converge as quickly with respect to the size of the CI space as the energy (cf. section 2.2), particularly for selected CI methods which use an energy selection criterion rather than a coefficient criterion. However, a 1992 study by Cave, Xantheas, and Feller²²⁷ used a selected CI method which is similar to the two-class CIPSI⁹² method but

uses energy-based selection. These authors came to the remarkable conclusion that most one-electron properties considered in their study (including isotropic hyperfine values and dipole and quadrupole moments) converged even more rapidly than the energy.²²⁷ Additionally, Wulfov²²⁸ has recently considered the convergence of the equilibrium geometries and harmonic vibrational frequencies of several diatomic molecules as a function of the selection threshold value.^{224,225} The convergence of these properties could not be firmly established due to the lack of corresponding exact full CI values; however, a recent full CI benchmark study by our group²²⁹ on C_2 and CN^+ finds only very small errors in Wulfov's best CI+PT geometries and frequencies (less than 0.001 Å and 4.0 cm^{-1} , respectively).

4.8 Restricted Active Space CI

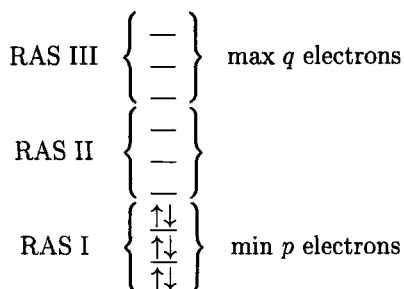
Rather than select individual determinants based on computational estimates of their importance, one might instead select entire classes of determinants which are expected *a priori* to be important based on their partitioning of electrons among various orbital subspaces. This is the motivation behind the truncation of the CI space according to "excitation level" (how many electrons are placed in the virtual subspace) and the second-order CI (SOCI),¹³ which includes all determinants with at most two electrons in the external subspace. Such CI selection schemes were described in general terms by Shavitt,⁵⁷ who defined the *full class CI* as one which partitions the orbitals into an arbitrary number of orthogonal subsets and includes all or none of the N -electron functions which have a given partitioning of electrons among the subspaces. As Shavitt points out, a full-class CI wavefunction is invariant to separate, nonsingular linear transformations within any of the orbital subspaces.

A benefit of such class selection schemes is that the CI space exhibits a regular structure which can be used to advantage in computational implementations. Additionally, it appears to be easier to gauge the general reliability of wavefunctions obtained using class selection schemes as opposed to individual selection. However, class selection methods will invariably include some less important determinants and therefore cannot yield wavefunctions as compact as those from an individual selection method.

A specialization of the full class CI which uses only three orbital subspaces is the Restricted Active Space (RAS) CI approach introduced by Olsen *et al.*⁴⁶ in 1988 along with the string-based full CI algorithm already discussed (section 4.4). The three subspaces are labeled I, II, and III, and the CI space is limited by requiring a *minimum of p electrons in RAS I* and a *maximum of q electrons in RAS III* (cf. Figure 14). There are no restrictions on the number of electrons in RAS II, and thus it is analogous to the complete active space

(CAS). There may be an additional frozen core subspace in which each orbital is constrained to remain doubly occupied; these core electrons and orbitals need not be treated explicitly in the RAS procedure (cf. section 2.4.7). The full CI space may be obtained as the maximum limit of the RAS space. The focus of Olsen's paper was on the utility of the RAS method in limiting the size of CI calculations, but thus far its maximum impact has been on the development of determinant-based full CI algorithms.^{83,183}

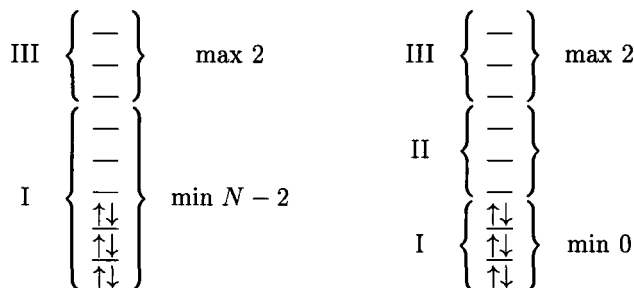
Figure 14: Orbital partitioning and configuration selection in the Restricted Active Space Configuration Interaction method. The CI space includes all determinants in which at least p electrons are in RAS I and at most q electrons are in RAS III.



Any CI space truncated according to excitation level may be formulated within the RAS CI framework: the occupied orbitals are placed in RAS I, and the unoccupied orbitals are placed in RAS III, and the RAS II subspace is absent. The maximum number of electrons in RAS III is set equal to the maximum excitation level, and the minimum number of electrons in RAS I is simply the total number of electrons N minus the maximum excitation level. A full CI can be obtained by applying trivial restrictions, such as a minimum of zero electrons in RAS I and a maximum of N electrons in RAS III.

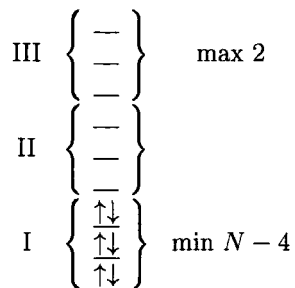
One may also formulate excitation class selected MR-CI spaces within the RAS framework. A SOCI can be obtained by setting RAS I equal to the active space, deleting RAS II, allowing a maximum of two electrons in RAS III, and requiring a minimum of $N - 2$ electrons in RAS I. Alternatively, this same CI space may be constructed by placing the virtual orbitals of the active space in RAS II, allowing a maximum of two electrons in RAS III, and requiring zero electrons in RAS I. To obtain the CISD[TQ] wavefunction mentioned in section 2.4.2, one places occupied orbitals in RAS I and virtual active space

Figure 15: Two formulations of the SOCI wavefunction within the RAS CI method.



orbitals in RAS II. A maximum of two electrons are allowed in RAS III, and a minimum of $N - 4$ electrons are required in RAS I (cf. Figure 16). Some possibilities of extending the RAS selection scheme are discussed in section 4.8.3.

Figure 16: Formulation of the CISD[TQ] wavefunction within the RAS method.



The RAS CI algorithm of Olsen *et al.*⁴⁶ relies on Handy's separation of determinants into alpha and beta strings (cf. section 4.2). As in other determinant-based CI methods, the determinants are restricted to those having a given value of M_s . Since the number of electrons N is also fixed, this means that the alpha and beta strings always have constant lengths of N_α and N_β , respectively. For a full CI, one forms all possible alpha and beta strings for a given N_α and N_β ,

and the determinants employed are all possible combinations of these alpha and beta strings. In a RAS CI wavefunction, the CI space is restricted in two ways: first, not all alpha and beta strings are allowed, and secondly, not all combinations of alpha and beta strings to form determinants are accepted. This is best understood from an example: consider the case of 6 orbitals, with $N_\alpha = N_\beta = 3$. If orbitals 4, 5, and 6 constitute RAS III, with a maximum of 2 electrons allowed, then clearly alpha strings such as $a_{4\alpha}^\dagger a_{5\alpha}^\dagger a_{6\alpha}^\dagger$ are not allowed. Similarly, even though $a_{1\alpha}^\dagger a_{4\alpha}^\dagger a_{5\alpha}^\dagger$ and $a_{1\beta}^\dagger a_{4\beta}^\dagger a_{5\beta}^\dagger$ are allowed alpha and beta strings, these strings cannot be combined with each other because the resulting determinant would place four electrons in RAS III.

If the CI coefficient vector is viewed as a matrix, $c(I_\alpha, I_\beta)$, then these restrictions, as well as those due to point-group symmetry, can be implemented by allowing only certain blocks of c to be nonzero. If strings with the same irreducible representation are numbered consecutively, then only those blocks of the CI coefficient matrix with the correct overall symmetry (determined as the direct product of the alpha and beta string irreducible representations) are allowed to be nonzero. For the RAS restrictions, it is convenient to assign a particular code to each alpha and beta string which represents the distribution of electrons among the RAS orbital subspaces. This means that the allowed combinations of strings may be determined directly from their respective codes. In Olsen's nomenclature,⁴⁶ these codes correspond to different string graphs; string representation and addressing is discussed further in section 4.9. If strings with a given code (or within a given graph) are numbered consecutively, then the allowed combinations of strings become allowed subblocks of the symmetry blocks in the CI coefficient matrix. These are referred to henceforth as the RAS subblocks. For a full CI, one can group all strings belonging to the same irreducible representation in the same graph, so that the RAS subblocks are the same as the symmetry blocks.

4.8.1 RAS CI σ Equations

The products of shift operators in equations (113) and (114) mean that some alpha/beta strings can contribute to the σ vector even if they are not necessary to describe the CI space. Therefore the string lists must include all valid strings and all the singly substituted strings derived from them. For a full CI, all possible strings for a given number of alpha/beta electrons are allowed, so all singly substituted strings are automatically present. For a restricted CI, however, including all singly substituted strings is an inefficient procedure.

Olsen *et al.* showed⁴⁶ how to eliminate the contributions from these "out-of-space" strings for the case of a RAS CI wavefunction. Since the equation for σ_3 (115) contains no products of shift operators, only σ_1 and σ_2 can contain

contributions from out-of-space strings. Moreover, since σ_2 is analogous to σ_1 (and can be derived from it for $M_s = 0$ cases), it is sufficient to focus on σ_1 . Olsen's approach is to change the sum over i, j, k, l into a restricted sum, $(ij) \geq (kl)$, where (ij) and (kl) are canonical indices defined as

$$(ij) = in + j, \quad (139)$$

if there are n orbitals and the numbering starts from zero. Consider the term

$$\sum_{J_\beta} \sum_{(ij) \geq (kl)} \langle \beta(J_\beta) | \hat{E}_{ij}^\beta \hat{E}_{kl}^\beta | \beta(I_\beta) \rangle (ij|kl) c(I_\alpha, J_\beta). \quad (140)$$

An out-of-space string K_β can only contribute if it is produced by $\hat{E}_{kl}^\beta | \beta(I_\beta) \rangle$, and if \hat{E}_{ij}^β transforms it back into an allowed string (J_β) . But if the orbitals are numbered consecutively within each RAS space, and if equation (139) is used to define the canonical index (ij) , such contributions are impossible. For example, consider the case where a maximum of two electrons are allowed in RAS III. If string I_β already contains two electrons in RAS III, then a single replacement $\hat{E}_{kl}^\beta | \beta(I_\beta) \rangle$ which promotes a third electron to RAS III will yield an out-of-space string. The shift operator \hat{E}_{ij}^β could produce an allowed string again if it moves any of the three electrons in RAS III back down into RAS I or RAS II. However, if the summation over orbitals is restricted to $(ij) \geq (kl)$, this is impossible, since $k > i$.

Using the commutation relationship

$$[\hat{E}_{ij}^\beta, \hat{E}_{kl}^\beta] = \delta_{kj} \hat{E}_{il}^\beta - \delta_{il} \hat{E}_{kj}^\beta, \quad (141)$$

one can rewrite (113) to employ this restriction. The result⁴⁶ is:

$$\begin{aligned} \sigma_1(I_\alpha, I_\beta) &= \sum_{J_\beta} \sum_{kl} \langle \beta(J_\beta) | \hat{E}_{kl}^\beta | \beta(I_\beta) \rangle g_{kl} c(I_\alpha, J_\beta) \\ &+ \sum_{J_\beta} \sum_{(ij) \geq (kl)} \langle \beta(J_\beta) | \hat{E}_{ij}^\beta \hat{E}_{kl}^\beta | \beta(I_\beta) \rangle (ij|kl) c(I_\alpha, J_\beta) (1 + \delta_{(ij), (kl)})^{-1}, \end{aligned} \quad (142)$$

where g is an asymmetric matrix defined as

$$g_{kl} = \begin{cases} h_{kl} - \sum_{j < k} (kj|jl) - (kk|kl)(1 + \delta_{kl})^{-1} & k \geq l \\ h_{kl} - \sum_{j < k} (kj|jl) & k < l. \end{cases} \quad (143)$$

4.8.2 Algorithms for Computing σ

The full CI algorithms must be modified to treat the more general RAS case. Most significantly, the σ_1 and σ_2 equations for RAS CI spaces differ from

the full CI equations: (142) requires that the summation over orbitals obey $(ij) \geq (kl)$ (recall that introducing this restriction into the full CI algorithm is one way to reduce effort when $M_s = 0$). The other required modification is that the innermost loops must sum over allowed combinations of alpha and beta strings only. For instance, the loop over I_α near the end of the σ_1 algorithm in Figure 7 must be split into two loops: one over alpha string graphs, and another over strings within each graph. This allows for the sum over J_β to run over only those beta strings which are allowed to combine with strings from the current alpha string graph. Similar modifications must be made to the σ_2 and σ_3 algorithms.

It is perhaps not entirely obvious what is the most efficient way to adapt Olsen's algorithms to account for point group symmetry and RAS restrictions. Although it is a relatively simple matter to introduce a loop over graphs in the multiplication of F by c or c' , it is also necessary to introduce loops over graphs in earlier parts of the algorithm; otherwise, F can contain irrelevant entries. When the CI vector is processed a symmetry block at a time or a RAS subblock at a time, it seems best to place these loops over graphs within the σ_1 , σ_2 , and σ_3 routines. However, when the CI vector is processed a RAS subblock at a time, these loops over graphs may be placed outside the σ subroutines. Further details are presented in section 4.9.

4.8.3 Beyond RAS: More Flexible *a priori* CI Space Selection

Although many useful CI spaces can be obtained in the RAS CI method, it may nevertheless be beneficial to employ more complex CI spaces. Olsen has begun to investigate CI spaces formed as the union of two RAS spaces,⁸⁰ while we have considered the addition of another orbital subspace.¹⁸ As first pointed out by Grev and Schaefer,¹⁶ the most weakly occupied CISD natural orbitals contribute almost exclusively to singly and doubly substituted configurations, rather than to triples, quadruples, etc. This suggests the utility of extending the RAS method to include another orbital subspace, formed from the most weakly occupied natural orbitals. Labeling this new orbital set as "RAS IV" may be somewhat misleading, in the sense that the orbital index restriction $(ij) \geq (kl)$ is no longer sufficient to remove out-of-space contributions from (142): we have alternatively referred to this new orbital set as the "tertiary virtual subspace".¹⁸

Let us assume that the occupied orbitals are collected in RAS I, and that if an electron occupies RAS IV, the determinant must represent a single or double substitution of the reference determinant. Out-of-space contributions arise if one electron occupies RAS IV and another electron occupies RAS II or RAS III: such strings are allowed, but the promotion of another electron

from RAS I results in a disallowed string $|\beta(K_\beta)\rangle = \hat{E}_{kl}^\beta |\beta(I_\beta)\rangle$. Application of the other shift operator in the σ_1 equation, \hat{E}_{ij}^β , can result in an allowed string $|\beta(J_\beta)\rangle$ with $(ij) \geq (kl)$ if the electron in RAS IV is moved into one of the lower RAS subspaces, as long as it occupies an orbital $i > k$. Thus it is necessary to include in the string space all allowed strings, plus the disallowed strings which have one electron in RAS IV and two electrons in (RAS II + RAS III). Once these strings are included, the RAS σ equations can be used for σ_1 and σ_2 . As in the standard RAS method, no out-of-space strings can contribute to σ_3 . By definition, the out-of-space strings are not allowed to combine with other strings to form RAS subblocks of the c or σ matrices. Alternatively, in some cases these out-of-space contributions might be dealt with by employing the $(N - 2)$ -electron reduced space of Zarrabian *et al.*^{47,217} (cf. section 4.5).

4.9 Implementation of Determinant-Based Algorithms

In this section we discuss some of the practical issues relevant to the actual implementation of the determinant-based CI algorithms. We also describe our experience with our own fully direct CI program, DETCI, which is capable of evaluating any CI wavefunction which can be formulated as a RAS CI, subject to memory and disk limitations. This program is based in part on the alpha and beta string formalism of Handy⁴⁴ and the algorithms of Olsen *et al.*^{46,83} It has been modified to allow more complex CI spaces, as described in sections 4.8.3 and 5.4.

Our program requires at least two memory buffers for CI vectors, where a buffer can be either the length of the entire vector, or a spatial symmetry block, or a RAS subblock. Using the fastest algorithm, the program also requires a smaller memory buffer to hold a portion of the CI coefficients in a given RAS subblock. An additional buffer the size of the largest RAS subblock may also be required for taking transposes of the c subblocks if $M_s = 0$ symmetry is employed (this is determined by the core memory option, and if possible the same buffer is used for transposes and gathered CI coefficients). For diagonalizing the Hamiltonian, we have implemented many of the iterative methods described in section 3.2.

4.9.1 Graphical Representation of Alpha and Beta Strings

As discussed in the previous section, it is necessary to have a method for numbering the alpha and beta strings and a reasonable way of grouping these strings together so that allowed combinations of strings can be determined a group at a time. Olsen *et al.* use a graphical method to compute string addresses, and they group strings together by placing them on the same graph.⁴⁶

We have employed a similar method.

The present approach is based on the work of Duch, who has described²³⁰ the graphical representation of CI spaces in considerable detail. First, we consider the simple two-slope directed graphs (“digraphs”) which represent alpha or beta strings without consideration of point group symmetry. Figure 17 presents a digraph representing all strings with five electrons in seven orbitals. Each string is represented by a “walk” on the graph, from the head (at $e = o = 0$) to the tail (at $e = N_{\alpha/\beta}$, $o = n$). Moving straight down from vertex (e, o) to vertex $(e, o + 1)$ indicates that orbital $o + 1$ is unoccupied in the current string, while moving down diagonally from vertex (e, o) to vertex $(e + 1, o + 1)$ indicates that orbital $o + 1$ is occupied. Each vertex on the graph is assigned a weight $x(e, o)$, and each arc connecting two vertices is assigned an arc weight $Y(e, o)$ for the arc leaving vertex (e, o) . Since, in general, two different arcs can leave a given vertex, we write $Y_0(e, o)$ for the arc originating from vertex (e, o) which leaves orbital $o + 1$ unoccupied, and $Y_1(e, o)$ for the arc which occupies orbital $o + 1$.[‡] The index or address of a string or walk is obtained by adding weights for each arc contained in the walk, i.e.,

$$I_{\alpha}(L^{\alpha}) = X(L^{\alpha}) + \sum_{i=0}^n Y_{L_i}(e_i, i), \quad (144)$$

where L_i is the occupation (0 or 1) of the i th arc, and (e_i, i) are the coordinates of the vertices crossed by L^{α} . The term $X(L^{\alpha})$ gives the offset of a given graph, if more than one graph is employed. The relative index for a determinant in a block may be given by $I(L^{\alpha}, L^{\beta}) = I_{\alpha}(L^{\alpha})S_{\beta} + I_{\beta}(L^{\beta})$, where S_{β} is the number of beta strings in the block.

There are several different methods for assigning the arc weights by which one evaluates the index of a string according to equation (144). Under the *lexical ordering* scheme, the tail (N_{α}, n) of an alpha string graph is assigned a weight $x = 1$. Other vertex weights are computed according to the recursive formula

$$x(e, o) = x(e + 1, o + 1) + x(e, o + 1). \quad (145)$$

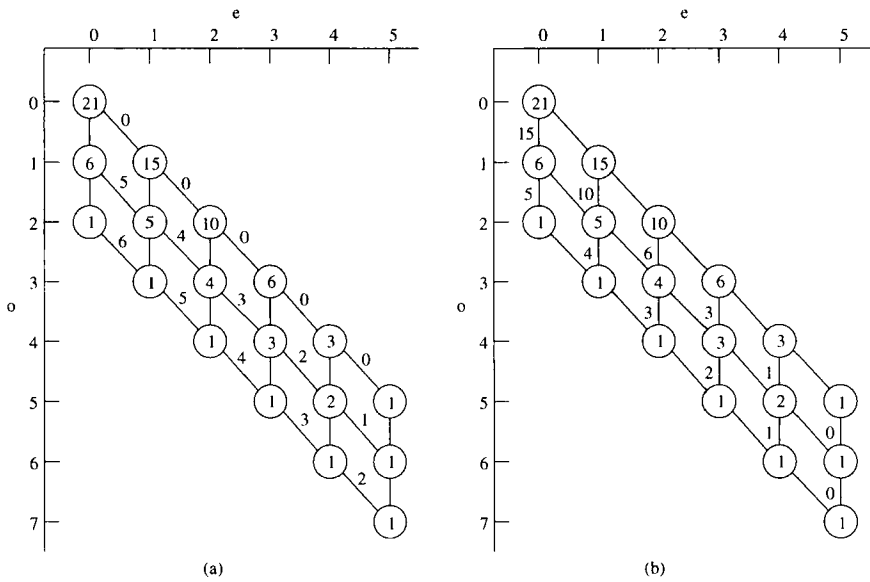
Using lexical ordering, typically all arc weights $Y_0(e, o)$ are set equal to zero, and the arc weights $Y_1(e, o)$ are determined according to

$$Y_1(e, o) = x(e + 1, o + 1) + x(e + 1, o) + \cdots + x(e + 1, e + 1). \quad (146)$$

Figure 17a features vertex and arc weights computed in this manner. A result of the lexical ordering scheme is that paths with a fixed upper part and an

[‡]This differs somewhat from Duch,²³⁰ who sometimes uses $\bar{Y}(e, o)$ to denote the arc entering vertex (e, o) in reverse-lexical addressing.

Figure 17: Alpha string graph for $n_\alpha = 5, n = 7$. Vertex weights are determined according to lexical ordering, and arc weights are given so that the rightmost path has index zero. (a) All unoccupied arc weights $Y_0(e, o)$ are zero. (b) All occupied arc weights $Y_1(e, o)$ are zero.



arbitrary lower part have consecutive indices. The particular choice of Y values above is appropriate if the rightmost path is to have an index of zero. The same effects can be achieved using

$$Y_1(e, o) = 0 \quad (147)$$

$$Y_0(e, o) = x(e + 1, o + 1), \quad (148)$$

as illustrated in Figure 17b. Any walk has the same index in Figures 17a and 17b. For instance, the walk $a_{2\alpha}^\dagger a_{3\alpha}^\dagger a_{4\alpha}^\dagger a_{5\alpha}^\dagger a_{7\alpha}^\dagger$ has an index of $5+4+3+2+2 = 16$ (equation 144) from Figure 17a, and an index of $15 + 1 = 16$ from Figure 17b.

In the so-called "reverse-lexical" ordering scheme, all upper paths for a fixed lower path have consecutive indices. Vertex weights are now determined as

$$\bar{x}(e, o) = \bar{x}(e, o - 1) + \bar{x}(e - 1, o - 1), \quad (149)$$

where the overbar indicates reversed-lexical ordering. Figure 18a depicts a reversed-lexical graph with all non-occupied orbital arcs set to zero. The occupied orbital arcs are computed as

$$\bar{Y}_1(e, o) = \bar{x}(e + 1, o). \quad (150)$$

Figure 18b is the same except that now all occupied arcs have weights of zero. The non-occupied arc weights are

$$\bar{Y}_0(e, o) = \bar{x}(e, o) + \bar{x}(e + 1, o + 1) + \cdots + \bar{x}(N - 1, o + N - e - 1). \quad (151)$$

Note that string indices for reverse-lexical ordering are not necessarily the same as indices for lexical ordering. For the string $a_{2\alpha}^\dagger a_{3\alpha}^\dagger a_{4\alpha}^\dagger a_{5\alpha}^\dagger a_{7\alpha}^\dagger$ considered previously, the index is calculated as $1 + 1 + 1 + 1 + 6 = 10$ from Figure 18a, or as $5 + 5 = 10$ from Figure 18b.

The arc weights given in Figures 17 and 18 cause the rightmost path to have an index $I(R_m) = 0$. If we change the arc weights so that the leftmost path has index $I(L_m) = 0$, we obtain four more addressing schemes. The two simplest schemes for $I(L_m) = 0$ are

$$Y_0(e, o) = 0 \quad Y_1(e, o) = x(e, o + 1) \quad (152)$$

$$\bar{Y}_1(e, o) = 0 \quad \bar{Y}_0(e, o) = \bar{x}(e - 1, o) \quad (153)$$

where the overbars indicate that reversed-lexical vertex weights have been used. Alpha strings for 5 electrons in 7 orbitals employing these addressing schemes are depicted in Figure 19.

If we add another coordinate Γ to each vertex, we can extend these simple digraphs to include point group symmetry. However, this procedure is not

Figure 18: Alpha string graph for $n_\alpha = 5, n = 7$. Vertex weights are determined according to reverse-lexical ordering, and arc weights are given so that the rightmost path has index zero. (a) All unoccupied arc weights $Y_0(e, o)$ are zero. (b) Occupied arc weights $Y_1(e, o)$ are set to zero.

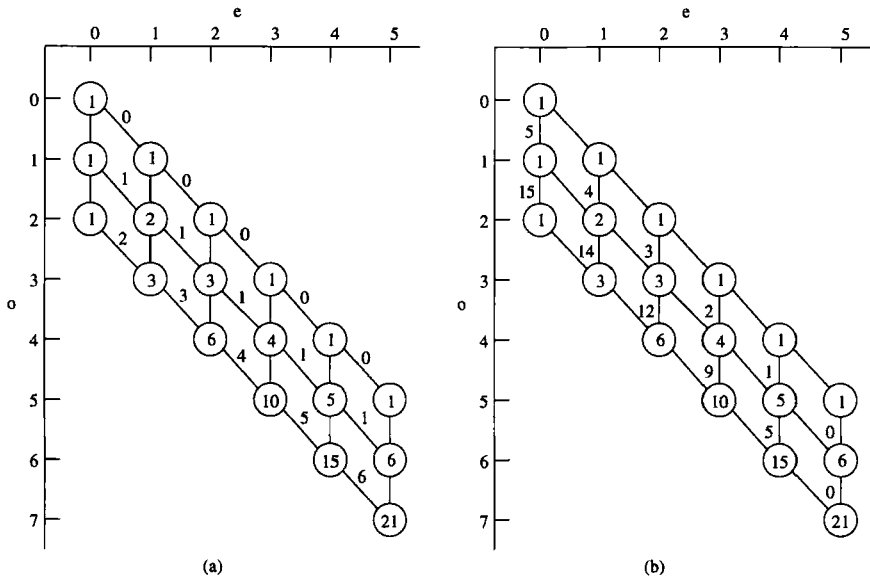
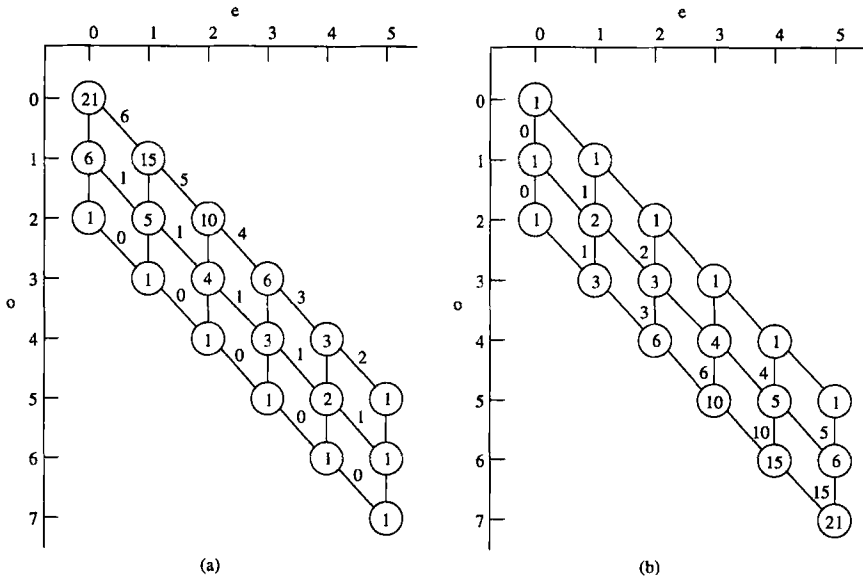


Figure 19: Alpha string graph for $n_\alpha = 5, n = 7$, with arc weights determined so that the leftmost path has index zero. (a) Vertex weights for lexical ordering, and arc weights according to $Y_0(e, o) = 0, Y_1(e, o) = x(e, o + 1)$. (b) Vertex weights according to reverse-lexical ordering, and arc weights according to $\bar{Y}_1(e, o) = 0, \bar{Y}_0(e, o) = \bar{x}(e - 1, o)$.



really necessary as long as strings with different irreps are placed on different graphs. These two-slope digraphs are actually simpler versions of the four-slope digraphs first used by Shavitt to compute the addresses of configuration state functions (CSFs) in the graphical unitary group approach.⁴¹ The indexing scheme used by DETCI is the reverse-lexical ordering with all unoccupied arc weights set to zero, as depicted in Figure 18a. Olsen *et al.*⁴⁶ use the lexical ordering of Figure 17a.

In order to make the CI coefficient matrix block diagonal according to irreducible representation, it is necessary to number the strings consecutively within each irrep. This is accomplished by grouping strings with the same irrep into the same graph. For RAS CI spaces, it is also useful to number the strings such that only certain subblocks of each symmetry block are nonzero. This is accomplished by forming a different string graph for each different distribution of electrons among the RAS subspaces; for example, all strings with irrep a_1 having four electrons in RAS I and two electrons in RAS III are grouped together, whereas strings with irrep a_1 but five electrons in RAS I and one electron in RAS III are grouped together in a separate graph. In this way, the allowed combinations of strings become allowed combinations of graphs, and each allowed pair of alpha and beta graphs becomes a RAS subblock.

4.9.2 Nongraphical Methods for String Addressing

In their full CI program, Bendazzoli and Evangelisti^{48,49} dispense with the graphs entirely and compute string addresses directly from their bit patterns, with one bit for each orbital; set bits (ones) represent occupied orbitals. Strings of N_α ones and $n - N_\alpha$ zeroes are a common representation of the combinations of n objects taken N_α at a time, and there exist standard numerical methods of computing lexical addresses for such bit patterns.²³¹

Knowles and Handy present an explicit formula for computing a string's address from a list of its occupied orbitals.¹⁰⁹ They employ an auxiliary array defined according to

$$Z(k, l) = \sum_{m=n-l}^{n-k-1} \left[\binom{m}{N-k-1} - \binom{m-1}{N-k-2} \right]$$

$$Z(k, l) = l + 1 - N \quad n \geq l \geq N; k = N - 1 \quad (154)$$

where k labels electrons, l labels orbitals, n is the number of orbitals, and N is the number of electrons (we have modified the equation so that electron and orbital numbering starts from zero). The address of a string is then computed

according to:

$$I_\alpha = \sum_{i=0}^{N_\alpha-1} Z(i, \phi_i). \quad (155)$$

Unfortunately, generalizing this formula to assign consecutive indices to strings with the same irrep becomes complicated. Bendazzoli *et al.* note that, since their strings are generated in the desired order, they can write the string's address to intermediate arrays and thus obviate the need to compute string addresses on-the-fly. Such considerations are also true of Olsen's program and of ours. However, storing all strings with their single replacement information can require a very large amount of memory, as discussed in section 4.9.4. In this case, it becomes useful to compute string addresses on-the-fly as rapidly as possible.

If only occupied orbitals are assigned arc weights in the graphical procedure, it is possible to obtain an equation similar to (155), but using the graphical numbering scheme instead of the Knowles-Handy numbering. This is easily seen by comparing equation (155) to equation (144) when only occupied arcs have nonzero weights:

$$I_\alpha(L^\alpha) = X(L^\alpha) + \sum_{i=0}^{N_\alpha-1} Y(i, \phi_i). \quad (156)$$

This method of evaluation can be very efficient if the matrix Y (with dimension $N_\alpha \times n$) is stored for each of the string graphs. The memory requirement for this approach will generally be manageable, and far preferable to storing the string replacement lists or Bendazzoli and Evangelisti's *OV* and *OOVV* lists.^{48,49} In the worst case, each excited string would then require N_α matrix lookups and $(N_\alpha - 1)$ additions to determine its address. However, note that a single excitation changes the occupancy of perhaps one and at most two RAS subspaces. It can be seen from eq. (156) that the contribution of unchanged RAS spaces to the string address are *constant* and need only be computed once. Graphically, this means that certain RAS spaces are traversed by the same segment of a walk. Finally, unless fast access to the arc weights is required, it is preferable to keep the arc weights and other graph information in a compact form which stores only allowed vertices.

4.9.3 Example of CI Vector and String Addressing

This section applies the CI vector and string addressing methods just discussed to the specific case of a CISD for H_2O in a minimal basis. Assume that the calculation is performed in C_{2v} symmetry, and that the core orbital has been frozen. The orbitals are then ordered according to Figure 20. Note that the

Figure 20: Orbital ordering for minimum basis CISD H₂O in C_{2v} symmetry.

$$\begin{array}{r}
 \text{RAS III} \\
 \\
 \text{RAS I} \\
 \\
 \text{Frozen core}
 \end{array}
 \left\{ \begin{array}{l}
 \text{---} \\
 \text{---} \\
 \uparrow\downarrow \\
 \uparrow\downarrow \\
 \uparrow\downarrow \\
 \uparrow\downarrow \\
 \uparrow\downarrow
 \end{array} \right\}
 \begin{array}{l}
 5 (2b_2) \\
 4 (4a_1) \\
 3 (1b_2) \\
 2 (1b_1) \\
 1 (3a_1) \\
 0 (2a_1) \\
 (1a_1)
 \end{array}$$

Table 6: Allowed strings for minimum basis CISD H₂O in C_{2v} symmetry.

Graph	Γ	N_{III}	Strings
0	a_1	1	(0 1 3 5)
1	a_1	2	(0 3 4 5), (1 3 4 5)
2	a_2	0	(0 1 2 3)
3	a_2	1	(0 2 3 4), (1 2 3 4), (0 1 2 5)
4	a_2	2	(0 2 4 5), (1 2 4 5)
5	b_1	1	(0 1 2 4), (0 2 3 5), (1 2 3 5)
6	b_1	2	(2 3 4 5)
7	b_2	1	(0 1 3 4)
8	b_2	2	(0 1 4 5)

frozen core orbital is not assigned a number by DETCI because the frozen core electrons are treated implicitly (cf. section 2.4.7).

The distribution of electrons among RAS subspaces can be determined simply from the number of electrons in RAS III, since there are only two RAS subspaces in this particular example. For a given string, there can be 0, 1, or 2 electrons in RAS III, and since there are four irreps in C_{2v} , there can be up to twelve string graphs. For this closed-shell case, the same set of graphs can be used to represent both alpha and beta strings. Table 6 lists all 15 allowed strings. Note that some strings transform as a_2 even though there are no a_2 orbitals; as a general rule for larger cases, strings are almost evenly distributed among irreps. The strings in Table 6 are listed in increasing index order. For the few graphs containing more than one string, it is straightforward to verify using the techniques of section 4.9.1 that the relative addresses within the graph are in the order shown if reverse-lexical ordering is used.

Table 7: Single replacement list for the first string in minimum basis CISD H_2O . J^g is the graph of the target string and J is the target string's relative index.

i	j	J^g	J	Sgn
0	0	0	0	+
1	1	0	0	+
3	3	0	0	+
5	5	0	0	+
4	0	1	1	+
4	1	1	0	-
2	5	2	0	-
2	3	3	2	+
2	0	5	2	-
2	1	5	1	+
4	5	7	0	+
4	3	8	0	+

As discussed later in section 4.9.4, it is necessary to compute lists of all singly excitations from each string. These excitations can be written in the form $|\alpha(J_\alpha)\rangle = \text{sgn}(ij)\hat{E}_{ij}^\alpha|\alpha(I_\alpha)\rangle$, and for each string I_α , the lists need to contain the address of J_α , the sign, and the pair of orbitals ij . The string replacement lists for string 0 are given in Table 7.

Finally, it is helpful to show how this string ordering determines the addressing of the CI vector. There are two restrictions on the pairing of alpha strings with beta strings: first, the direct product of the two string irreps must be the irrep of the electronic state of interest (a_1); second, the total number of electrons in RAS III must be two or less. Due to the arrangement of the graphs, these restrictions can be satisfied for pairs of graphs rather than for pairs of strings. Each allowed pair of graphs becomes a RAS subblock. Table 8 lists the allowed RAS subblocks. An interesting feature of this unusually small CI space is that not all allowed strings contribute to allowed determinants: there are no beta strings which can be combined with alpha strings from graphs 1, 6, or 8 which give allowed determinants.

4.9.4 String Replacement Lists

One of the first steps in implementing Olsen's algorithm is the construction of string lists which hold the information needed to generate all single replace-

Table 8: Allowed RAS subblocks for minimum basis CISD H₂O.

Block	α Graph	β Graph	Dets
0	0	0	1
1	2	2	1
2	2	3	3
3	2	4	2
4	3	2	3
5	3	3	9
6	4	2	2
7	5	5	9
8	7	7	1

ments for each string; for example, $|\beta(J_\beta)\rangle = \text{sgn}(ij)\hat{E}_{ij}^\beta|\beta(I_\beta)\rangle$ in Figure 9. For each excited string, one needs i , j , (or a composite index ij), the string address J_β , and $\text{sgn}(ij)$, which tells whether \hat{E}_{ij}^β sends $\beta(I_\beta)$ into plus or minus $\beta(J_\beta)$. The sign is most easily derived using the rules of second quantization, recalling that the phase convention is defined by always listing alpha/beta strings according to increasing orbital number (cf. section 4.2). Note that the equations require the inclusion of operators \hat{E}_{ii} , which might not normally be considered “single substitutions.”

In the innermost loop of Olsen’s σ_3 algorithm, $V(I) = \sum_{J_\beta} F(J_\beta)c'(I, J_\beta)$, the summation should be restricted to only allowed pairs of strings (I, J_β) . If loops over graphs are introduced, then one loops over J_α belonging to a given alpha string graph (J_α^g) and J_β belonging to a given beta string graph (J_β^g), where (J_α^g, J_β^g) is an allowed combination of graphs. Now the problem of summing over all substituted strings becomes one of summing over all substituted strings belonging to a given graph. Thus DETCI divides the list of single replacements into multiple lists, one for each target graph. A counter array is used to store how many singly substituted strings there are for each graph. For a large number of graphs (there can be several hundred), this method of storage can be very memory intensive, since the number of pointers for each string is proportional to the number of graphs; however, this method also provides very fast access to the relevant replacement information in the innermost loops.

It is important to point out that the string lists can become very large if there are more than a few thousand strings. For a full CI, the number of strings is approximately equal to the square root of the number of determinants, so quite large full CI spaces can be described using only a small number of strings.

Table 9: Number of strings and determinants for selected CI problems.

CI problem	Fzc	Strings	Dets	Str/Det
cc-pVDZ+ Ne full CI ^a	0	8 568	9 185 280	9.3×10^{-4}
DZ H ₂ O full CI	0	2 002	1 002 708	2.0×10^{-3}
DZ H ₂ O full CI	1	715	128 829	5.5×10^{-3}
DZP H ₂ O CISDTQ	1	10 626	558 823	1.9×10^{-2}
DZP H ₂ O CISD[TQ] ^b	1	4 326	78 895	5.5×10^{-2}
DZP H ₂ O CISDT	1	5 781	50 187	1.2×10^{-1}
DZP H ₂ O CISD	1	1 221	2 349	5.2×10^{-1}

^aUsing the basis set of Koch *et al.*⁷

^bThe RAS II space consists of the 2b₂,4a₁,2b₁, and 5a₁ orbitals.

However, it is perhaps not well appreciated that for RAS CI spaces, the number of strings can grow much faster with the number of determinants: this is illustrated in Table 9. Such considerations indicate that any method of storing the string replacement lists whose memory requirements are proportional to the number of strings (with a large coefficient, no less) is unmanageable for large-scale RAS CI procedures.

If the number of determinants remains much larger than the number of strings for a given case, one might consider storing the string lists to disk and loading them as needed. Indeed, such an approach would probably work in conjunction with the σ_3 algorithm. However, this strategy is not amenable to the σ_1 algorithm as implemented by Olsen *et al.*,⁴⁶ because one must consider single replacements from *all* graphs which can be reached from the graph containing the strings I_β . It seems preferable to form the string replacement information on-the-fly. A prototype method for doing this exists in DETCI, but it is inefficient: walking down graphs to add up arc weights is rather slow with our current storage scheme for the graphs. However, if the arc weights are stored in a slightly different format, it is possible to compute string addresses much more quickly; indeed, it is possible to compute addresses a RAS space at a time, meaning that arc weight contributions from RAS spaces with fixed orbital occupancies are constant. This strategy was discussed in section 4.9.2 and is currently being implemented in DETCI. Another option is to use an entirely different implementation for restricted CI where storage of string replacement information becomes a problem. Malmqvist *et al.*⁵⁶ have implemented RAS CI using a split-graph unitary group approach, and this algorithm may be more efficient for CI spaces which are not close to the full CI.

4.9.5 Algorithms for σ_2 and σ_3 Used by DETCI

Assuming that the string replacement lists are available in-core, the algorithms used by our program for computing σ_2 and σ_3 (simplified somewhat for clarity) are presented in Figures 21 and 22 for the case in which σ is computed a RAS subblock at a time, where the loops over combinations of strings allowed by symmetry and CI space restrictions are placed outside the σ_2 and σ_3 subroutines. The symbol γ in Fig. 22 represents an irreducible representation of the molecular point group. These are the same algorithms used by our program for full CI; if enough memory is available to hold a symmetry block of c and σ in core, then the RAS subblocks are the same as the symmetry blocks for a full CI. Otherwise, one can use a larger number of string graphs and smaller subblocks. For $M_s = 0$ cases, it is possible to compute only the lower (or upper) triangle of σ , according to (119). Thus only one of each pair of off-diagonal subblocks is determined explicitly. For diagonal blocks, (119) is used to impose the restriction $(ij) \geq (kl)$ in the evaluation of σ_3 (cf. section 4.4.2). In FORTRAN, presumably it is best to eliminate redundant subblocks from the upper triangle, since blocks with only a few rows will typically be found in the upper triangle (assuming that larger indices are assigned to strings with more electrons in RAS III): these blocks do not vectorize well due to short vector lengths. In C, exactly the opposite holds: it is presumably most efficient to eliminate the lower triangle subblocks. However, it is also true that longer vector lengths require more time to set up the lists L , R , and Sgn .

The signs are actually taken care of by very fast bitwise logic operations. σ_2 is computed instead of σ_1 because the preferred direction of vectorization is reversed in c as compared to FORTRAN. In the present context, this means that the step $[\sigma_2(I_\alpha, I_\beta) = \sigma_2(I_\alpha, I_\beta) + F(J_\alpha) * c(J_\alpha, I_\beta)]$ is performed with unit-stride access to σ_2 and c . Note that this is in fact a DAXPY operation, which could be performed by calling the `daxpy` function from the BLAS library. For the IBM RS/6000 POWER2 implementation, calling the BLAS library from the σ_2 or σ_3 routines resulted in no real savings. In fact, this slowed down the σ_3 subroutine, presumably due to the overhead of placing the function call within nested loops.

Note that the σ_2 routine is actually quite close to Olsen's σ_1 routine (swapping alphas for betas, of course), except that the adaptation to point group symmetry and RAS restrictions is now explicit. The σ_3 routine is also similar to Olsen's (with alphas swapped for betas), although the intermediate vector F has been eliminated (see below). Note that this particular version of the σ_3 routine takes advantage of the $(ij) \geq (kl)$ simplification possible for $M_s = 0$ cases; this restriction cannot be used for off-diagonal RAS subblocks if only unique subblocks of σ are computed.

Figure 21: DETCI algorithm for σ_2 .

```

Enter subroutine for block  $(I_\alpha^g, I_\beta^g)$  of  $\sigma_2$  and block  $(J_\alpha^g, I_\beta^g)$  of  $c$ 
loop  $I_\alpha = 1$ , number of strings in  $I_\alpha$  graph  $I_\alpha^g$ 
  zero  $F(J_\beta)$ 
  loop over alpha string graphs  $K_\alpha^g$ 
     $Kcnt = cnt(I_\alpha, K_\alpha^g)$ 
     $Kidx = idx(I_\alpha, K_\alpha^g)$ 
     $Kij = ij(I_\alpha, K_\alpha^g)$ 
    loop  $K = 1, Kcnt$ 
       $kl = Kij(K)$ 
       $K_\alpha = Kidx(K)$ 
       $S_1 = \text{sign associated with } K_\alpha$ 
      if  $(K_\alpha^g = J_\alpha^g)$ 
         $F(K_\alpha) = F(K_\alpha) + S_1 * h(kl)$ 
       $Jcnt = cnt(K_\alpha, J_\alpha^g)$ 
       $Jidx = idx(K_\alpha, J_\alpha^g)$ 
       $Jij = ij(K_\alpha, J_\alpha^g)$ 
      loop  $J = 1, Jcnt$ 
         $J_\alpha = Jidx(J)$ 
         $S_2 = (\text{sign associated with } J_\alpha) * S_1$ 
         $ij = Jij(J)$ 
         $F(J_\alpha) = F(J_\alpha) + 0.5 * (ij|kl) * S_2$ 
      end loop over  $J$ 
    end loop over  $K$ 
  end loop over  $K_\alpha^g$ 
  loop  $J_\alpha = 1$ , number of alpha strings in  $J_\alpha^g$ 
    if  $(F(J_\alpha) = 0)$  skip to next  $J_\alpha$ 
    loop  $I_\beta = 1$ , number of beta strings in  $I_\beta$  graph  $I_\beta^g$ 
       $\sigma_2(I_\alpha, I_\beta) = \sigma_2(I_\alpha, I_\beta) + F(J_\alpha) * c(J_\alpha, I_\beta)$ 
    end loop over  $I_\beta$ 
  end loop over  $J_\alpha$ 
end loop over  $I_\alpha$ 

```

Figure 22: DETCI algorithm for σ_3 .

```

Enter subroutine for block  $(I_\alpha^g, I_\beta^g)$  of  $\sigma_3$  and block  $(J_\alpha^g, J_\beta^g)$  of  $c$ 
loop over  $ij$ 
  if  $(\gamma(i) \otimes \gamma(j) \otimes \gamma(J_\beta^g) \otimes \gamma(I_\beta^g) \neq 0)$  skip to next  $ij$ 
   $jlen = \text{form\_list}(I_\beta^g, J_\beta^g, ij, L, R, Sgn)$ 
  if  $(jlen = 0)$  skip to next  $ij$ 
  loop  $I_\alpha = 1$ , number of alpha strings in  $I_\alpha^g$ 
    loop  $J = 1, jlen$ 
       $c'(I_\alpha, J) = c(I_\alpha, L(J)) * Sgn(J)$ 
    end loop over  $J$ 
  end loop over  $I_\alpha$ 

  loop  $I_\alpha = 1$ , number of strings in graph  $I_\alpha^g$ 
     $Jacnt = cnt(I_\alpha, J_\alpha^g)$ 
     $Jaidx = idx(I_\alpha, J_\alpha^g)$ 
     $Jaij = ij(I_\alpha, J_\alpha^g)$ 
    zero  $V$ 
    loop  $Ja = 1, Jacnt$  AND  $((kl = Jaij(Ja)) \leq ij)$ 
       $J_\alpha = Jaidx(Ja)$ 
       $S_1 = \text{sign associated with } J_\alpha$ 
       $VS = 0.5 * \delta_{ij,kl} * (ij|kl) * S_1$ 
      loop  $Jb = 1, jlen$ 
         $V(Jb) = V(Jb) + VS * c'(J_\alpha, Jb)$ 
      end loop over  $Jb$ 
    end loop over  $Ja$ 

    loop  $Jb = 1, jlen$ 
       $\sigma_3(I_\alpha, R(Jb)) = \sigma_3(I_\alpha, R(Jb)) + V(Jb)$ 
    end loop over  $Jb$ 
  end loop over  $I_\alpha$ 
end loop over  $ij$ 

```

The figures clearly indicate that an efficient algorithm should take advantage of the sparsity of F in both the σ_2 and σ_3 routines. Although it is possible to ignore this sparsity in order to obtain more highly vectorized algorithms, this results in a slower code due to the large number of multiplications by zero. Attempts to formulate the σ_3 algorithm as a standard matrix multiplication resulted in a program which was slower by a factor of at least four (and usually more) on the IBM RS/6000 model 3CT. The reader may wonder whether further reductions in the overall operation count might increase efficiency: after all, formulating the innermost loop as a DAXPY requires a certain amount of overhead work (namely, the gather and scatter operations). Indeed, a very simple algorithm which performs the minimum number of operations has already been presented in Fig. 8. However, there are a number of reasons to think that this algorithm should be less efficient than that in Figure 22 for workstation computers. One reason is that indirect addressing in the innermost loop causes cache misses and thus longer waits for elements of c . Another is that the simple algorithm takes no advantage of the pipelining features of current workstations, discussed in section 4.3. For DZ H₂O full CI (1 million determinants), using the algorithm in Fig. 22 instead of the simple algorithm in Fig. 8 results in a speedup by a factor of five on the IBM 3CT. However, it is also important to point out that the smaller block sizes in a RAS CI (as opposed to a full CI) mean that the payoff for vectorization is less.

A few remaining comments should be made about our σ_3 algorithm in Fig. 22. First, note that the intermediate vector F has been eliminated. Storage to and retrieval from this array requires several extra operations, yet the efficiency of the DAXPY operation is not enhanced by the use of F , since F is sparse (the only work saved is for excitations which map a string into itself). Second, writing the innermost loop of this routine (and also of the σ_1 routine) as a DAXPY is particularly efficient on the RS/6000 architecture, which performs each pair of floating point multiplications and additions in the *same* machine instruction, the floating point multiply-add (FMA).

In order to gauge the efficiency of our new determinant-based CI program, several timing comparisons have been made between DETCI and the standard Schaefer group CI program GUGACI. Table 10 presents averaged CPU and total times for several full CI test cases. Both programs used the standard Davidson method, keeping on disk the c and σ vectors for every iteration; this lead to noticeable I/O delays for some of the larger test cases. Our new program is also capable of using iteration methods which require fewer vectors on disk (cf. section 3.2). Although we do not have access to Olsen's code, he has reported that his program takes about 40s per iteration for the 1 million determinant DZ H₂O full CI calculation,²³² whereas DETCI requires 30s per iteration on the same model workstation (IBM RS/6000 model 590).

Table 10: CPU and total times per iteration to evaluate several benchmark full CI wavefunctions using the IBM RS/6000 model 3CT. Fzc/Fzv denote the number of frozen core/frozen virtual orbitals.

Molecule	Fzc/Fzv	CSFs	Dets	Time/Iter (seconds)			
				GUGACI		DETCI	
				CPU	Tot	CPU	Tot
DZ H ₂ O	1/0	37 353	128 829	77	79	3	3
DZ H ₂ O	0/0	256 473	1 002 708	1430	1465	28	33
cc-pVDZ+ Ne ^a	0/0	2 083 968	9 185 280	13788	14007	282	532
DZP C ₂	2/2	6 571 116	27 944 940	32489	34831	1329	3160

^aUsing the basis set of Koch *et al.*⁷

5 Applications of Highly Correlated CI

Most studies limit the CI space to single and double substitutions from a single reference (CISD). Occasionally, when one reference does not provide a sufficient zeroth-order wavefunction, the multi-reference CISD method will be employed. The applications of such methods are too numerous to discuss here; their general performance has already been described in section 2. We consider studies which go beyond CISD for one or a few references; such highly correlated wavefunctions are useful when very accurate results are desired or when several electron configurations are needed for a qualitatively correct reference wavefunction (such as when multiple bonds are broken). We will limit our attention to methods which select the CI space in an *a priori* fashion based on the distribution of electrons among various orbital subspaces.

5.1 Full CI

The most highly correlated configuration interaction method is of course full CI, which solves the Schrödinger equation exactly within the space spanned by the single-particle basis set (section 2.1). Unfortunately, as explained in section 2.4.1, the full CI space grows factorially with the number of electrons or single-particle basis functions; currently, full CI is limited to very small molecules (one or two heavy atoms with a few hydrogen atoms) described by a basis set of double- ζ or perhaps triple- ζ quality. Furthermore, errors due to the incompleteness of the single-particle basis set are generally more severe than those introduced by the neglect of triple, quadruple, etc., substitutions in the treatment of electron correlation. Hence full CI is useful primarily as

a benchmark method for evaluating approximate treatments of correlation. Given their extreme computational requirements, it is perhaps surprising how many full CI benchmarks have been reported. Here we will focus primarily on the more noteworthy or more recent benchmarks, and on systems containing more than four electrons.

The early (1980) full CI algorithm of Handy⁴⁴ enabled Saxe, Schaefer, and Handy to obtain the exact variational solution for the ground state of H₂O within a modest double- ζ (DZ) basis set; this represented the first CI wavefunction to include more than one million determinants.²¹³ In 1983, Harrison and Handy used this same algorithm, along with the loop-driven graphical unitary group approach CI (LD-GUGACI) program of Brooks and Schaefer,⁴³ to report full CI results for H₂O and NH₃ with a DZ basis and for BH and HF in a double- ζ plus polarization (DZP) basis.³⁸ These two studies, which also gave results for the CISD, CISDT, and CISDTQ methods, clearly demonstrated that triply and quadruply substituted configurations account for nearly all of the error in the CISD correlation energy (cf. Table 2). Moreover, by comparing to Harrison and Handy's results for H₂O, Bartlett, Sekino, and Purvis demonstrated that fourth-order many-body perturbation theory [MBPT(4)] performs poorly when both O-H bonds are stretched,¹⁹ even though it partially accounts for quadruple excitations. This occurs because the restricted Hartree-Fock determinant provides an inadequate zeroth-order wavefunction as the bonds are stretched too far from their equilibrium lengths, and it emphasizes the need for multireference approaches.

The subsequent (1984) vectorized full CI algorithm¹⁰⁹ of Knowles and Handy allowed Bauschlicher, Taylor, Langhoff, and others to carry out a series of important benchmark calculations. In 1986, these authors, along with Partridge, presented full CI results for the Ne atom using a triple- ζ plus double polarization (TZ2P) basis set.²³³ Once again, the CISDTQ wavefunction yielded more than 99% of the basis set correlation energy; furthermore, the contribution of quintuple and higher substitutions decreased with increasing basis set size. Subsequent benchmark results using a DZP basis were presented for HF and NH₂;²³⁴ H₂O, F, and F⁻;²³⁵ the ¹A₁-³B₁ separation in CH₂;²³⁶ the barrier height to the reaction F + H₂ → HF + H;²³⁷ the ¹A₁, ¹B₁, and ³B₁ states of SiH₂;²³⁸ the 2 ¹A₁ states of CH₂ and SiH₂ and the ²A₁ and ²B₂ states of CH₂⁺;²³⁹ the low-lying states of C₂;²⁴⁰ N₂, NO, and O₂;²⁴¹ the CH₃ radical;²⁴² and Be₃.²⁴³ Bauschlicher and co-workers have also investigated transition metals. They estimated the ⁵D-⁵F energy separation for the Fe atom using a 5s4p2d1f basis set,²⁴⁴ excitation energies and oscillator strengths for the ²D Rydberg series in the Al atom with a 7s6p4d3f ANO basis,²⁴⁵ the ³F-⁵F separation for the Ti atom, and the ⁴Φ-²Δ separation for TiH in a 5s4p3d1f/2s basis set.²⁴⁶ Illas, Rubio, Ricart, and Bagus used the

Knowles-Handy program to obtain full CI energies for the first row atoms and their hydrides in 4s3p1d/2s1p basis sets; these benchmarks were used to evaluate the CIPSI method²¹⁹ and the performance of ANO basis sets versus more traditional segmented basis sets.²⁴⁷ Casanovas, Rubio, and Illas have also performed full CI studies on the interaction of H atom with Cu₅ and Ag₅ cluster models to investigate the transferability of the correlation contribution to the chemisorption bond for different pseudopotentials.²⁴⁸

Several of these studies considered molecules at three geometries, with equivalent bonds simultaneously stretched to 1.0, 1.5, and 2.0 times their equilibrium lengths; hence, approximate methods could be judged not only by what fraction of the correlation energy they recovered, but also by how well they paralleled the full CI potential energy surface. The MR-CISD method was found to parallel full CI very well, particularly when CASSCF orbitals are used and all configurations present in the CASSCF are used as references. More recently, these authors have used large basis set full CI benchmarks to examine core-core and core-valence correlation effects²⁴⁹ and to calibrate more approximate MR-CISD approaches for the dissociation energy of BH.²⁵⁰ Many of these results are summarized in a 1990 review article by Bauschlicher, Langhoff, and Taylor.¹⁵

Most of the full CI studies just discussed involved CI spaces spanning tens of millions of determinants. More recent full CI algorithms, which follow Olsen *et al.* in sacrificing some degree of vectorization for reduced operation counts,^{46,48,49} have allowed for CI spaces including hundreds of millions of determinants. Indeed, Olsen, Jørgensen, and Simons reported a full CI calculation on the Mg atom using a 5s3p2d1f ANO basis and requiring more than one billion determinants; unfortunately, it was not possible to fully converge the wavefunction due to the extreme amount of CPU time required.⁸³ The first *converged* full CI benchmark requiring more than one billion determinants was reported recently by Evangelisti, Bendazzoli, Ansaloni, Durí, and Rossi, who presented an out-of-core adaptation of their full CI program for distributed-memory parallel computers and used it to obtain the full CI energy of Be₂, with all electrons correlated, in a 4s2p1d ANO basis set partially uncontracted to 9s2p1d.⁸⁴

In 1995, these workers used an in-core version of this algorithm to resolve an uncertainty concerning the full CI energy of NH₃ with an ANO DZP basis set.⁵⁰ Knowles and Handy had in 1989 presented an energy of -56.4236 ± 0.0001 hartree using their approximate full CI method which truncates the σ vector; 665,247 of 209,626,425 determinants were treated variationally, and the convergence limit was estimated perturbatively.²¹⁸ This value underestimates the importance of the large number of neglected determinants, as was first demonstrated by Povill, Rubio, Caballol, and Malrieu.¹²³ The true full CI

energy was shown to be -56.424007 hartree.^{50,80} The program of Bendazzoli and Evangelisti⁴⁸ has also been applied to DZ H_3 and N_2 ⁴⁹ and to $C_{18}H_{18}$ using a simple Pariser-Parr-Pople (PPP) Hamiltonian.⁴⁸

In 1996, Olsen, Jørgensen, Koch, Balková, and Bartlett presented DZP full CI results for H_2O at three geometries, where all ten electrons were correlated. Although the basis set was not designed to describe core correlation, these results are valuable in that they are fully invariant to orbital rotations, whereas they would not have been if the $1s$ -like orbital on oxygen were frozen.²² As shown by Handy and co-workers,²⁰ it is fairly straightforward (and valuable) to generalize a full CI program to produce many-body perturbation theory (MBPT) energies order by order. Hence, Olsen *et al.* considered MBPT through 15th order and examined the convergence of the series as the two OH bonds are simultaneously stretched, causing the zeroth-order wavefunction to become progressively worse. In another 1996 study, Olsen, Christiansen, Koch, and Jørgensen examined the convergence of MBPT for several small molecules using DZP and larger basis sets. These authors concluded that perturbation theory corrections grow with increasing basis size and that, remarkably, the inclusion of diffuse functions can cause the perturbation series to diverge even for well-behaved molecules such as HF,^{23,251} Olsen's algorithm has also been used to provide benchmark full CI excitation energies for CH^+ ,²⁵² BH, CH_2 , and Ne atom;⁷ H_2O , N_2 , and C_2 ;³⁷ and H_2O^+ .²²

Most full CI studies have focused solely on energies. A few papers, however, have presented full CI results for other molecular properties. In 1987 Bauschlicher and Taylor presented full CI dipole moments and polarizabilities for HF, CH_2 , SiH_2 , and F^- with a DZP basis.²⁵³ Moreover, Bauschlicher and Langhoff gave full CI equilibrium geometries, dissociation energies, and harmonic vibrational frequencies for CH, NH, and OH using flexible ANO basis sets.²⁵⁴ The following year, these authors presented full CI and SOCI transition moments for two transitions in CH_2 and selected dipole and quadrupole transitions in BeO,²⁵⁵ and Bauschlicher presented geometries, frequencies, and the dipole moment for the $^4\Phi$ and $^2\Delta$ states of TiH.²⁴⁶ Later, Bauschlicher and Taylor also presented full CI transition moments for H_2 and BH in a comparison of the length and velocity representations for the transition moment.²⁵⁶ Full CI transition moments and polarizabilities for CH^+ have been presented by Olsen and co-workers.²⁵² Koch and Bauschlicher have presented a method for computing analytically the frequency dependent linear and quadratic response functions for full CI wavefunctions; they have considered Be atom in a $9s9p5d$ basis set and have reported transition energies and dipole moments for several states, the first polarizability at real and imaginary frequencies, and the static second hyperpolarizability.²⁵⁷ Bauschlicher *et al.* have determined the full CI isotropic hyperfine coupling constant for the nitrogen atom using

basis sets as large as 8s4p2d.²⁵⁸ Spin-orbit coupling has been investigated at the full CI level for CH₂ by Vahtras *et al.*²⁵⁹ and for LiBe by Marino *et al.*²⁶⁰

Researchers have even obtained exact full CI equilibrium geometries for a few polyatomic systems with a DZP basis: the linear transition state for the reaction F + H₂ → HF + H,²³⁷ the systems H₆, H₇⁺ and H₅⁺, He,^{17,261} the \tilde{X}^2B_1 and \tilde{A}^2A_1 states of NH₂,²⁶² BH₃,^{39,262} H₅⁺,²⁶³ and the four lowest-lying states of methylene^{262,264} and NH₂⁺.²⁶⁵ For the latter four molecules, full CI dipole moments and harmonic vibrational frequencies have been reported at the full CI equilibrium geometries.^{39,263–265} The DZP full CI studies of methylene clearly demonstrate the need to use larger basis sets: predictions for the singlet-triplet energy gap^{236,262,264} are at least 2.5 kcal mol⁻¹ too large compared to experiment.²⁶⁴ We have recently completed a considerably more challenging full CI benchmark study of the four lowest states of methylene using the more reliable TZ2P basis set.²⁶⁶

In the past few years, full CI benchmarks have been used to calibrate methods for systems featuring weak interactions. Woon has obtained the full CI well-depth, equilibrium separation, and harmonic vibrational frequency for He₂ with basis sets as large as augmented correlation-consistent polarized quadruple- ζ (aug-cc-pVQZ).²⁶⁷ In a subsequent study, van Mourik and van Lenthe used the full CI program of Harrison and Zarrabian⁴⁷ to obtain dimer energies at two separate geometries using large basis sets including *h*-type polarization functions and bond functions.³⁶ One interesting conclusion of this study, which could only be determined using highly correlated wavefunctions, was that the usually-reliable CCSD method is unsuitable for obtaining accurate potential energy curves for He₂. Other recent full CI studies have also focused on helium dimer,^{268–271} as well as the dimer of two H₂ molecules^{268–270} and the He–H₂ system.^{268,269} Such benchmarks have been helpful in examining the problem of basis set superposition error.²⁷⁰

5.2 Second-Order CI

One important conclusion from the full CI benchmark studies of Bauschlicher, Taylor, Langhoff, and others in the 1980's is that the MR-CISD method based on CASSCF orbitals provides potential energy surfaces which accurately parallel the full CI surfaces.^{14, 15, 234, 238–240, 242, 254} For example, the CASSCF MR-CISD method predicts singlet-triplet energy separations in CH₂ and SiH₂ within 0.01 kcal mol⁻¹ and 0.03 kcal mol⁻¹, respectively, of the full CI results.^{236,238} The best results are obtained when no threshold is used for reference selection: that is, when all CSFs in the CASSCF wavefunction are used as references. This CAS-ref MR-CISD procedure is intimately related¹⁰⁷ to second-order configuration interaction (SOC), which distributes electrons in

all possible ways as long as no more than two electrons are allowed in external orbitals at once (of course, spatial symmetry and spin symmetry may also be imposed on the final N -electron basis functions). For closed-shell systems these procedures are identical, but they can differ for open-shell systems. In cases where all occupied orbitals which are correlated in the SOCI are included in the active space, one can guarantee that the SOCI space is generated by using as references all CSFs arising from CASSCF wavefunctions of all possible spatial symmetries.¹⁰⁷ SOCI wavefunctions are invariant to orbital rotations within the active space, whereas this is not necessarily the case for CAS-ref MR-CISD when the references are symmetry-restricted. In cases where there are "inactive" orbitals (occupied orbitals which are correlated in the SOCI but not included in the CASSCF), Bauschlicher has recommended that the SOCI be defined according to the RAS CI scheme, such that one places inactive orbitals in RAS I, active orbitals in RAS II, and external orbitals in RAS III;¹⁵ this maintains the desired orbital invariance properties. This SOCI can alternatively be generated an MR-CISD in which the references are now all CSFs arising from all CASSCF wavefunctions of every spatial and spin symmetry.

The reliability of SOCI, coupled with its essentially *a priori* selection of the CI space, makes it an attractive alternative to full CI. Unfortunately, for reasonable active spaces the dimension of the SOCI grows very rapidly with system size and thus the method is applicable only to small molecules. Nevertheless, for quite a few molecules it is possible to use the SOCI method in conjunction with large one-particle basis sets and hence to obtain wavefunctions very close to the exact nonrelativistic Born-Oppenheimer limit. Below, we will attempt to give the reader a sense of the types of problems to which the SOCI method has been applied. Bauschlicher, Langhoff, and Taylor have already given¹⁵ an excellent review of the related CASSCF MR-CISD method, and we refer the reader to their article for a discussion of additional important studies.

The SOCI method has been applied primarily to diatomics and triatomics. Two of the early applications of SOCI were studies of Be_2 by Blomberg, Siegbahn, and Roos²⁷² (1980) and by Lengsfeld, McLean, Yoshimine, and Liu⁸⁹ (1983). With only four active electrons, a SOCI for this system is identical to the CISD[TQ] method discussed in the following section. The Be_2 molecule is challenging to theory because of the near degeneracy of the $1s^2 2s^2$ and $1s^2 2p^2$ configurations in the the Be atom and because the bonding in the dimer is dominated by dispersion forces. The latter study⁸⁹ used a $6s4p3d1f$ Slater basis set to yield a SOCI dissociation energy of $D_e = 1.87 \text{ kcal mol}^{-1}$ ($2.04 \pm 0.21 \text{ kcal mol}^{-1}$, including estimates of core correlation and basis set errors), in contradiction to previous coupled-cluster studies²⁷³ giving $D_e \leq 0.2 \text{ kcal mol}^{-1}$; the experimental value²⁷⁴ is $2.26 \pm 0.09 \text{ kcal mol}^{-1}$, signaling a success

of the SOCI method and a failure of coupled-cluster models which neglect connected triple substitutions. A SOCI study of the related Mg_2 molecule using large ANO basis sets was reported²⁷⁵ in 1990 by Partridge *et al.* Other SOCI studies on the ground states of diatomic molecules include an investigation of the potential energy curves of NeN^{2+} and NeN^+ by Koch, Liu, and Frenking,²⁷⁶ and of N_2 and O_2 by Langhoff, Bauschlicher, and Taylor.²⁷⁷ The latter study used a 5s4p3d2f1g basis and reported equilibrium bond lengths, harmonic vibrational frequencies, and dissociation energies of (1.101 Å, 2343 cm^{-1} , 9.723 eV) for N_2 and (1.209 Å, 1561 cm^{-1} , 5.139 eV) for O_2 , in excellent agreement with experimental values of (1.0977 Å, 2358 cm^{-1} , 9.905 eV) and (1.2075 Å, 1580 cm^{-1} , 5.214 eV), respectively. The multi-reference Davidson correction for size extensivity (section 2.4.6) improves the SOCI dissociation energies but worsens the predicted bond lengths and vibrational frequencies. A subsequent study by Almlöf *et al.*²⁷⁸ considered even larger basis sets (including *i* polarization functions), core-valence correlation, and basis set superposition effects in an investigation of remaining sources of error in the dissociation energy of N_2 ; their best theoretical estimate was within about 2 kcal mol⁻¹, or 1% of the experimental dissociation energy.

Several studies have used the SOCI method to describe excited electronic states of diatomic molecules. In 1988, Partridge *et al.*²⁷⁹ reported SOCI potential energy curves for the $A' \ ^5\Sigma_g^+$ and $C'' \ ^5\Pi_u$ states of N_2 ; these results had important implications for theories of the N_2 afterglow and for the first time allowed an assignment of the Hermann infrared system. Also in 1988, Bauschlicher *et al.*²⁸⁰ reported SOCI spectroscopic constants (r_e , T_e , and ω_e) for several low-lying quartet states of AlC; this study indicated that the SOCI results based on CASSCF orbitals optimized separately for each state were nearly identical to SOCI using a common set of state-averaged CASSCF orbitals. Balasubramanian has used the SOCI method along with relativistic pseudopotentials to study diatomics incorporating elements below the second row of the periodic table. His work has included studies of the low-lying electronic states of InSb;²⁸¹ Ga_2 , Ga_2^- , and Ga_2^+ ;²⁸² GeH and GeH^+ ;²⁸³ WH;²⁸⁴ GaH ;²⁸⁵ TiH and InH;²⁸⁶ and GeCl.²⁸⁷ Another study by Balasubramanian considers the low-lying states of the transition metal hydrides YH-CdH.²⁸⁸

The SOCI method has also been used for a number of triatomics, including the much-studied²⁶⁴ methylene molecule (CH_2). An early application of the SOCI method to methylene was presented by Saxe, Schaefer, and Handy,²⁸⁹ who used Handy's 1980 string-based determinant CI program⁴⁴ to obtain an estimate of $T_e = 10.5$ kcal mol⁻¹ for the singlet-triplet energy gap using a 8s5p3d/4s1p basis set. More recently, Bauschlicher, Taylor, and Langhoff have used the SOCI method in conjunction with a 5s4p3d2f1g/4s3p2d ANO basis set to predict²⁹⁰ optimized geometries for the $\tilde{X} \ ^3B_1$ and $\tilde{a} \ ^1A_1$ states of

CH_2 of (1.079 Å, 133.6°) and (1.110 Å, 102.0°), which compare to experimentally derived estimates of (1.0753 Å, 133.93°) and (1.116 Å, 101.8°), respectively. McLean *et al.*²⁹¹ used the SOCI method with a 4s3p2d1f/3s2p basis to obtain 37 energy points which were fit with a nonrigid bender Hamiltonian model to yield $\nu_1 = 2985 \pm 20 \text{ cm}^{-1}$ and $\nu_3 = 3205 \pm 20 \text{ cm}^{-1}$ for the fundamental stretching frequencies, compared to experimentally derived values of 2992 and 3213 cm^{-1} . For the singlet-triplet energy splitting (T_e), the 5s4p3d2f1g/4s3p2d ANO SOCI estimate of 9.13 kcal mol⁻¹ by Bauschlicher *et al.*²⁹⁰ and the TZ3P(2f,2d)+2diff SOCI estimate of 9.02 kcal mol⁻¹ by Yamaguchi *et al.*⁸⁶ compare very favorably with the best nonrelativistic Born-Oppenheimer-corrected experimental estimate of 9.372 kcal mol⁻¹.^{264,292-294} Yamaguchi *et al.* also report⁸⁶ SOCI excitation energies for the \tilde{b}^1B_1 and \tilde{c}^1A_1 states which will hopefully guide further experimental efforts.

Among the numerous other studies of polyatomic molecules using the SOCI method, Balasubramanian has reported energies and optimized geometries for low-lying states of SnH_2 ,^{139,295} PbH_2 ,¹³⁹ GaH_2 , GaH_2^+ , GaH_3 , and GaH_3^+ ,²⁸⁵ YH_2^+ and ZrH_2^+ ,²⁹⁶ AsH_2 , AsH_2^+ , SbH_2^+ , and BiH_2^+ ,²⁹⁷ GeH_2 ,^{139,283} HfH_2 ,²⁹⁸ TlH_2 , TlH_2^+ , InH_2 , and InH_2^+ ,²⁸⁶ and PH_2 , PH_2^+ , and PH_2^- .²⁹⁹ Yarkony *et al.* have used SOCI wavefunctions as zeroth-order solutions in first-order perturbation theory treatments of the full Breit-Pauli spin-orbit Hamiltonian; in this fashion, these workers have studied the spin-forbidden decay of systems including $\tilde{X}^2\Pi \text{HS}^{2+}$,³⁰⁰ $\tilde{a}^3\Sigma_1^+ \text{NO}^+$,³⁰¹ and $\tilde{a}^4\Sigma^- \text{CH}$.³⁰² SOCI has also been used to assess the quality of the CCSD(T) method for electron affinities,³⁰³ to obtain accurate barrier heights for the termolecular reaction of 3H_2 ,²⁶¹ and to yield transition moments.^{138,255}

Finally, we note that it may be advantageous to use natural orbitals (NOs) rather than the CASSCF orbitals usually employed in SOCI studies. Grev and Schaefer¹⁶ have shown for a number of small molecules (NH_2 , CH_3 , SiH_2 , N_2) that a SOCI procedure based on CISD natural orbitals yields energies which are very close to the CASSCF SOCI energies, even when several bonds are simultaneously stretched to twice their equilibrium length. Blomberg and Liu have also observed similar performance of MCSCF orbitals and SOCI natural orbitals for energies and transition moments of CH and CH^+ ,¹³⁸ although we note that CISD NOs are much less expensive to obtain than SOCI NOs. CISD NOs are also easier to obtain than CASSCF orbitals, and they are better suited to orbital truncation; that is, the energy lost by deleting a few of the most weakly occupied NOs will typically be smaller than that lost by deleting a few of the highest-lying CASSCF orbitals. This allows for an effective reduction in the number of determinants included in a SOCI wavefunction with minimal loss in accuracy. Moreover, the energy lost by deleting these weakly occupied NOs is primarily due to neglected singly and doubly substituted determinants

which occupy these orbitals, and not to neglected triples, quadruples, etc.^{16,17} This suggests a strategy of using a smaller NO basis to treat higher-than-double substitutions, and this idea has been implemented in the CISD[TQ] method described in section 5.4.

5.3 Restricted Active Space CI

As already discussed, many commonly encountered CI spaces can be formulated within the RAS CI scheme. However, at the moment our attention is focused on CI calculations using the RAS CI program of Olsen and co-workers.⁴⁶ Examples of full CI calculations using Olsen's program have already been discussed above, so here we limit our attention to truncated CI wavefunctions.

The RAS CI method can be very valuable when used in conjunction with the multiconfigurational linear response (MCLR) method, which allows calculation of excitation energies, transition moments, and second-order properties such as polarizability. In 1989, Olsen and co-workers compared MCLR and full CI results for CH^+ and found that for highly accurate results, it is necessary to provide more extensive treatments of electron correlation than the valence CAS. Furthermore, these authors found that the RAS method provides an accurate means of reducing the size of the CI space in these MCLR studies. Jensen *et al.* were able to show that the RAS MCLR calculations which include dynamical correlation give very reliable frequency-dependent polarizabilities for the nitrogen molecule.³⁰⁴ Sanchez de Merás *et al.* found that the polarizabilities of H_2O and CO_2 obtained using the RAS MCLR method with polarized basis sets were within 5% of experiment.³⁰⁵ More recently, Sundholm and Olsen have used the RAS approach as part of a finite element multiconfigurational Hartree-Fock method for determining the atomic quadrupole moment of Ca ($3d4s; {}^1D$) and the electron affinity of the 1S ground state.^{306,307}

5.4 CISDTQ and CISD[TQ]

Few studies have employed configuration interaction with all singles, doubles, triples, and quadruples (CISDTQ) because the number of triple and quadruple substitutions grows very rapidly with the number of electrons and basis functions (cf. Table 3). CISDTQ results are most commonly reported in benchmark full CI studies to indicate the fraction of the basis set correlation energy recovered by triples and quadruples.^{17,22,39,80,234,241,264} Nevertheless, the CISDTQ method has occasionally been used for benchmarking in cases where the full CI was not technically feasible, because the CISDTQ results are expected to be very close to full CI for small molecules. For systems with eight electrons or less at their equilibrium geometries, the CISDTQ method recovers more

than 99.8% of the correlation energy for a DZP basis set (cf. Table 2). One benchmark study by Lee *et al.* examined the effects of triple and quadruple excitations on equilibrium geometries, harmonic vibrational frequencies, and infrared intensities of several small molecules.³⁰⁸ A similar study by Scuseria, Hamilton, and Schaefer³⁰⁹ used CISDTQ equilibrium geometries and harmonic vibrational frequencies to evaluate the performance of the CCSDT method for several diatomic molecules.⁴

Chemical applications of the CISDTQ method include the 1988 study by Scuseria and Schaefer³¹⁰ on the barrier height for the $F + H_2 \rightarrow FH + H$ reaction, which is very sensitive to the level of theory employed. By truncating the CI space at quadruples, these authors were able to increase the basis set from 28 functions (in Bauschlicher and Taylor's full CI study²³⁷) to 47 basis functions and enabled them to consider the effects of correlating the fluorine 2s orbital. Additionally, Tanaka and Nishimoto have used CISDTQ to examine the reaction mechanism for 1,3 hydrogen transfer in excited states of formamide,³¹¹ and Du, Hrovat, and Borden³¹² have used CISDTQ as part of a study on singlet-triplet gaps in diradicals.

Multireference CISD methods generally offer a more economical treatment of the dominant effects of triple and quadruple substitutions and allow the use of larger one-particle basis sets. The primary disadvantage of MR-CISD compared to CISDTQ is that the choice of references (and truncation, if any, of the generated singles and doubles space) must be performed carefully so as not to bias the results. An *a priori* selection scheme for MR-CISD which has been investigated in our laboratory is the CISD[TQ] method,¹⁶ which selects as references all single and double substitutions in the active space. This is equivalent to a CISDTQ in which no more than two electrons are allowed into external orbitals, or to a second-order CI (SOCi) in which greater-than-quadruple substitutions have been eliminated. This method was used by Saxe *et al.* in a 1982 study of ethylene using a DZP basis and the shape-driven graphical unitary group approach (SD-GUGA) CI program.⁹¹ The resulting wavefunction (spanning more than 1 million CSFs) was considered a benchmark result. In 1992 study of NH_2 , CH_3 , SiH_2 , C_2 , and N_2 , Grev and Schaefer¹⁶ found that the CISD[TQ] method provides results which are very close to SOCi when a single reference function dominates. A subsequent study of several other small molecules by Fermann *et al.*¹⁷ reinforced these conclusions. Some results from these studies which indicate the ability of CISD[TQ] to parallel the SOCi and full CI surfaces are presented in Table 11. The CISD[TQ] and SOCi wavefunctions were based on CISD natural orbitals, which perform as well as CASSCF orbitals for SOCi.¹⁶ Although the savings in the number of CSFs for CISD[TQ] compared to SOCi in Table 11 is relatively modest, this savings increases very rapidly with basis set and number of electrons: for

Table 11: Errors in Total Energies (millihartree) Relative to Full CI for Several Molecules Using a DZP Basis.^a

Method	No. CSFs	$E(r_e)$	$E(1.5 \cdot r_e)$	$E(2.0 \cdot r_e)$
² B ₁ NH ₂ (Ref. 16):				
CISD	898	9.003	23.475	69.168
CISD[TQ]	18 396	2.897	2.630	4.957
SOCI	21 687	2.853	2.107	1.703
Full CI	2 435 160	0.000	0.000	0.000
² A ₂ ' CH ₃ (Ref. 16):				
CISD	1 385	8.384	23.216	70.646
CISD[TQ]	51 818	2.156	2.065	4.910
SOCI	76 660	2.090	1.254	0.889
Full CI	9 591 312	0.000	0.000	0.000
¹ A ₁ H ₂ O (Ref. 17):				
CISD	926	12.851	30.421	75.644
CISDTQ	151 248	0.397	1.547	6.280
CISD[TQ]	32 361	1.630	2.537	6.867
SOCI	76 660	1.276	1.058	1.020
Full CI	6 740 280	0.000	0.000	0.000

^aCISD[TQ] and SOCI methods employed CISD natural orbitals.

DZP ethylene, a CISD[TQ] description requires over 1 million CSFs, whereas a SOCI requires over 42 million. In addition to these results, Grev and Schaefer¹⁶ also demonstrate that the CISD[TQ] method yields reliable dissociation energies for the difficult N₂ and C₂ molecules.

More recent work has explored the quality of CISD[TQ] equilibrium geometries and harmonic vibrational frequencies. King *et al.*³¹³ have found that for H₂O with a TZ2P basis, the equilibrium geometry predicted by CISD[TQ] differs from that of CISDTQ by less than 0.0001 Å in the bond length and 0.2° in the bond angle. Furthermore, the CISD[TQ] harmonic vibrational frequencies differ from those of the complete CISDTQ by an average of only 5 cm⁻¹. Such agreement is outstanding, particularly in view of the fact that the CISD[TQ] wavefunction contains 45 times fewer CSFs than CISDTQ. In related work, Hoffman *et al.*³¹⁴ have found similar results for H₂S. Another study by Leininger and Schaefer³¹⁵ considers the ozone molecule, which is challenging to theory because of the unusual importance of triple and quadruple substitutions. It is not currently feasible to obtain the complete CISDTQ wavefunction with a DZP basis, but the DZP CISD[TQ] equilibrium geometry differs from experiment by only 0.009 Å, and the harmonic vibrational frequencies differ by an average of 2.4%, with the treacherous antisymmetric stretching frequency predicted within 4.5%.

Finally, even though the computational cost of a CISD[TQ] procedure is substantially reduced compared to CISDTQ, the scaling with system size (cf. section 2.4.5) remains unfavorable. Hence, it is necessary to seek further reductions in the CI expansion with a minimal loss in the quality of the wavefunction. One promising strategy was suggested by Grev and Schaefer's 1992 study¹⁶ on the use of CISD natural orbitals (NOs) in the CISD[TQ] and SOCI methods. This work indicated not only that CISD NOs are as effective as CASSCF orbitals at providing good correlating orbitals in the active space, but also that the most weakly occupied NOs contribute almost negligibly to the energy. Furthermore, the errors in the energy caused by deleting weakly occupied orbitals are almost entirely due to the neglect of singles and doubles occupying these orbitals, and not to triples or quadruples. This suggests a more compact CISD[TQ] wavefunction¹⁷ which splits the external orbital space into two sets. If the unoccupied (virtual) orbitals of the active space are labeled the primary virtual subspace, then the external orbitals are divided into secondary and tertiary virtual subspaces, where the tertiary subspace comprises the set of the most weakly occupied NOs. One may then modify the CISD[TQ] method to eliminate those triple and quadruple substitutions which place an electron in one of the most weakly occupied orbitals.^{17,18,90} In effect, one uses a larger basis set for the singles and doubles than for the triples and quadruples.

This "split-virtual" CISD[TQ] method has been implemented in our de-

Table 12: Correlation Energy Recovered (Relative to CISDTQ) by the Split-Virtual CISD[TQ] Method for H₂O in a cc-pVTZ Basis Set.^a

Method/ Secondary Space	NO Cutoff	Number of Dets	%SDTQ Energy
CISD		15 939	93.1
CISDT		938 679	95.8
CISDTQ		28 085 271	100.
CISD[TQ] w/ primary space 2b ₂ 4a ₁ 2b ₁ 5a ₁			
(18 7 9 15)	none	984 789	98.8
(9 4 6 8)	10 ⁻⁴	320 531	98.6
(3 1 1 2)	10 ⁻³	41 261	96.5
(0 0 0 0)		16 713	94.7

^aCISD[TQ] methods employed CISD natural orbitals (NOs). The secondary orbital space selects all virtual orbitals with NO occupation numbers greater than the given cutoff and is identified according to how many orbitals of each irrep of C_{2v} it contains, in the order (a₁, a₂, b₁, b₂). Only valence electrons are correlated.

terminant based CI program (cf. section 4.8.3), and we have presented some preliminary results¹⁸ for the neon atom and for H₂O. Table 12 gives some of our results for H₂O with a cc-pVTZ basis set. We find the NO-based selection scheme to be effective in obtaining the dominant effects of triple and quadruple substitutions while using fewer CSFs, and we believe this promising strategy should be even more effective in the coupled-cluster approach, where the connected triples (\hat{T}_3 operator) could be evaluated using a smaller NO space.

Acknowledgments

We thank Dr. Jeppe Olsen, Dr. Wesley Allen, and Matthew Leininger for helpful discussions, and Dr. Daniel Crawford and Dr. Justin Fermann for their work on the two-electron integral transformation algorithm. This work was sponsored by the U.S. National Science Foundation, grant no. CHE-9527468, and by NSF graduate and postdoctoral fellowships (grant no. CHE-9626094) to C.D.S.

References

- [1] J. R. Thomas, B. J. DeLeeuw, G. Vacek, and H. F. Schaefer, *J. Chem. Phys.* **98**, 1336 (1993).
- [2] J. R. Thomas, B. J. DeLeeuw, G. Vacek, T. D. Crawford, Y. Yamaguchi, and H. F. Schaefer, *J. Chem. Phys.* **99**, 403 (1993).
- [3] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- [4] G. E. Scuseria and T. J. Lee, *J. Chem. Phys.* **93**, 5851 (1990).
- [5] G. E. Scuseria, *Chem. Phys. Lett.* **94**, 27 (1991).
- [6] J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.* **98**, 7029 (1993).
- [7] H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, *Chem. Phys. Lett.* **244**, 75 (1995).
- [8] B. O. Roos, The multiconfigurational (MC) SCF method, in *Methods in Computational Molecular Physics*, edited by G. H. F. Diercksen and S. Wilson, pages 161–187. D. Reidel, Dordrecht, 1983.
- [9] B. O. Roos, The multiconfigurational (MC) self-consistent field (SCF) theory, in *Lecture Notes in Quantum Chemistry: European Summer*

- School in Quantum Chemistry*, edited by B. O. Roos, volume 58 of *Lecture Notes in Chemistry*, pages 177–254. Springer-Verlag, New York, 1992.
- [10] R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta* **35**, 33 (1974).
- [11] R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta* **39**, 217 (1975).
- [12] J. A. Pople, in *Energy, Structure and Reactivity: Proceedings of the 1972 Boulder Summer Research Conference on Theoretical Chemistry*, edited by D. W. Smith and W. B. McRae, page 51. Wiley, New York, 1973.
- [13] H. F. Schaefer, *Configuration Interaction Wave Functions and the Properties of Atoms and Diatomic Molecules*, PhD thesis, Stanford University, Stanford, CA, 1969.
- [14] F. B. Brown, I. Shavitt, and R. Shepard, *Chem. Phys. Lett.* **105**, 363 (1984).
- [15] C. W. Bauschlicher, S. R. Langhoff, and P. R. Taylor, Accurate quantum chemical calculations, in *Advances in Chemical Physics*, edited by I. Prigogine and S. A. Rice, volume 77, pages 103–161. Wiley, New York, 1990.
- [16] R. S. Grev and H. F. Schaefer, *J. Chem. Phys.* **96**, 6850 (1992).
- [17] J. T. Fermann, C. D. Sherrill, T. D. Crawford, and H. F. Schaefer, *J. Chem. Phys.* **100**, 8132 (1994).
- [18] C. D. Sherrill and H. F. Schaefer, *J. Phys. Chem.* **100**, 6069 (1996).
- [19] R. J. Bartlett, H. Sekino, and G. D. Purvis, *Chem. Phys. Lett.* **98**, 66 (1983).
- [20] N. C. Handy, P. J. Knowles, and K. Somasundram, *Theor. Chim. Acta* **68**, 87 (1985).
- [21] S. A. Kucharski, J. Noga, and R. J. Bartlett, *J. Chem. Phys.* **90**, 7282 (1989).
- [22] J. Olsen, P. Jørgensen, H. Koch, A. Balková, and R. J. Bartlett, *J. Chem. Phys.* **104**, 8007 (1996).
- [23] J. Olsen, O. Christiansen, H. Koch, and P. Jørgensen, *J. Chem. Phys.* **105**, 5082 (1996).

- [24] G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- [25] W. D. Laidig and R. J. Bartlett, *Chem. Phys. Lett.* **104**, 424 (1984).
- [26] Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **81**, 5906 (1984).
- [27] Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **82**, 5761(E) (1985).
- [28] J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987).
- [29] J. Noga and R. J. Bartlett, *J. Chem. Phys.* **89**, 3401(E) (1988).
- [30] J. D. Watts, G. W. Trucks, and R. J. Bartlett, *Chem. Phys. Lett.* **157**, 359 (1989).
- [31] J. D. Watts and R. J. Bartlett, *J. Chem. Phys.* **93**, 6104 (1990).
- [32] H. Nakatsuji and M. Ehara, *J. Chem. Phys.* **98**, 7179 (1993).
- [33] X. Li, P. Piecuch, and J. Paldus, *Chem. Phys. Lett.* **224**, 267 (1994).
- [34] A. Balková and R. J. Bartlett, *J. Chem. Phys.* **102**, 7116 (1995).
- [35] X. Li and J. Paldus, *J. Chem. Phys.* **102**, 8059 (1995).
- [36] T. van Mourik and J. H. van Lenthe, *J. Chem. Phys.* **102**, 7479 (1995).
- [37] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, *Chem. Phys. Lett.* **256**, 185 (1996).
- [38] R. J. Harrison and N. C. Handy, *Chem. Phys. Lett.* **95**, 386 (1983).
- [39] J. M. Galbraith, G. Vacek, and H. F. Schaefer, *J. Mol. Structure* **300**, 281 (1993).
- [40] J. Paldus, *J. Chem. Phys.* **61**, 5321 (1974).
- [41] I. Shavitt, *Int. J. Quantum Chem. Symp.* **11**, 131 (1977).
- [42] I. Shavitt, *Int. J. Quantum Chem. Symp.* **12**, 5 (1978).
- [43] B. R. Brooks and H. F. Schaefer, *J. Chem. Phys.* **70**, 5092 (1979).
- [44] N. C. Handy, *Chem. Phys. Lett.* **74**, 280 (1980).
- [45] P. E. M. Siegbahn, *Chem. Phys. Lett.* **109**, 417 (1984).

- [46] J. Olsen, B. O. Roos, P. Jørgensen, and H. J. Aa. Jensen, *J. Chem. Phys.* **89**, 2185 (1988).
- [47] R. J. Harrison and S. Zarrabian, *Chem. Phys. Lett.* **158**, 393 (1989).
- [48] G. L. Bendazzoli and S. Evangelisti, *J. Chem. Phys.* **98**, 3141 (1993).
- [49] G. L. Bendazzoli and S. Evangelisti, *Int. J. Quantum Chem. Symp.* **27**, 287 (1993).
- [50] S. Evangelisti, G. L. Bendazzoli, R. Ansalmi, and E. Rossi, *Chem. Phys. Lett.* **233**, 353 (1995).
- [51] R. Shepard, The analytic gradient method for configuration interaction wave functions, in *Modern Electronic Structure Theory*, edited by D. R. Yarkony, volume 2 of *Advanced Series in Physical Chemistry*, pages 345–458. World Scientific, Singapore, 1995.
- [52] Y. Yamaguchi, Y. Osamura, J. D. Goddard, and H. F. Schaefer, *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab initio Molecular Electronic Structure Theory*. Oxford, New York, 1994.
- [53] R. Shepard, The multiconfiguration self-consistent field method, in *Ab Initio Methods in Quantum Chemistry-II*, edited by K. P. Lawley, volume 69 of *Advances in Chemical Physics*, pages 63–200. Wiley, New York, 1987.
- [54] H.-J. Werner, Matrix-formulated direct multiconfiguration self-consistent field and multiconfiguration reference configuration-interaction methods, in *Ab Initio Methods in Quantum Chemistry-II*, edited by K. P. Lawley, volume 69 of *Advances in Chemical Physics*, pages 1–62. Wiley, New York, 1987.
- [55] B. O. Roos, The complete active space self-consistent field method and its applications in electronic structure calculations, in *Ab Initio Methods in Quantum Chemistry-II*, edited by K. P. Lawley, volume 69 of *Advances in Chemical Physics*, pages 399–445. Wiley, New York, 1987.
- [56] P.-Å. Malmqvist, A. Rendell, and B. O. Roos, *J. Phys. Chem.* **94**, 5477 (1990).
- [57] I. Shavitt, The method of configuration interaction, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer, pages 189–275. Plenum Press, New York, 1977.

- [58] P. E. M. Siegbahn, The configuration interaction method, in *Lecture Notes in Quantum Chemistry: European Summer School in Quantum Chemistry*, edited by B. O. Roos, volume 58 of *Lecture Notes in Chemistry*, pages 255–293. Springer-Verlag, New York, 1992.
- [59] J. Karwowski, The configuration interaction approach to electron correlation, in *Methods in Computational Molecular Physics*, edited by S. Wilson and G. H. F. Diercksen, pages 65–98. Plenum Press, New York, 1992.
- [60] W. Duch, *J. Mol. Struct. (Theochem)* **80**, 27 (1991).
- [61] R. McWeeny, *Methods of Molecular Quantum Mechanics*. Academic Press, New York, 2nd edition, 1992.
- [62] E. Schrödinger, *Ann. d. Phys.* **79**, 734 (1926).
- [63] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. McGraw-Hill, New York, 1989.
- [64] W. Klopper, R. Röhse, and W. Kutzelnigg, *Chem. Phys. Lett.* **178**, 455 (1991).
- [65] R. D. Kent and M. Schlesinger, *Phys. Rev. A* **46**, 6881 (1992).
- [66] W. Cencek and J. Rychlewski, *J. Chem. Phys.* **102**, 2533 (1995).
- [67] B. J. Persson and P. R. Taylor, *J. Chem. Phys.* **105**, 5915 (1996).
- [68] E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929).
- [69] E. A. Hylleraas, The Schrödinger two-electron atomic problem, in *Advances in Quantum Chemistry*, edited by P.-O. Löwdin, volume 1, pages 1–33. Academic Press, New York, 1964.
- [70] L. Pauling and E. B. Wilson, Jr., *Introduction to Quantum Mechanics with Applications to Chemistry*. McGraw-Hill, New York, 1935.
- [71] J. K. L. MacDonald, *Phys. Rev.* **43**, 830 (1933).
- [72] E. A. Hylleraas and B. Undheim, *Z. Phys.* **65**, 759 (1930).
- [73] W. Ritz, *J. Reine Angew. Math.* **135**, 1 (1909).
- [74] J. C. Slater, *Phys. Rev.* **34**, 1293 (1929).
- [75] E. U. Condon, *Phys. Rev.* **36**, 1121 (1930).

- [76] J. C. Slater, *Phys. Rev.* **38**, 1109 (1931).
- [77] P. O. Löwdin, *Phys. Rev.* **97**, 1474 (1955).
- [78] J. Paldus, *Int. J. Quantum Chem. Symp.* **9**, 165 (1975).
- [79] J. Olsen, The method of second quantization, in *Lecture Notes in Quantum Chemistry: European Summer School in Quantum Chemistry*, edited by B. O. Roos, volume 58 of *Lecture Notes in Chemistry*, pages 37–87. Springer-Verlag, New York, 1992.
- [80] J. Olsen, unpublished.
- [81] P. J. Knowles, *Chem. Phys. Lett.* **155**, 513 (1989).
- [82] H. Weyl, *The Classical Groups, Their Invariants and Representations*. Princeton University Press, Princeton, 1946.
- [83] J. Olsen, P. Jørgensen, and J. Simons, *Chem. Phys. Lett.* **169**, 463 (1990).
- [84] S. Evangelisti, G. L. Bendazzoli, R. Ansaloni, F. Durí, and E. Rossi, *Chem. Phys. Lett.* **252**, 437 (1996).
- [85] C. W. Bauschlicher and D. R. Yarkony, *J. Chem. Phys.* **69**, 3875 (1978).
- [86] Y. Yamaguchi, C. D. Sherrill, and H. F. Schaefer, *J. Phys. Chem.* **100**, 7911 (1996).
- [87] P. G. Burton and P. D. Gray, *Chem. Phys. Lett.* **96**, 453 (1983).
- [88] P. E. M. Siegbahn, The direct CI method, in *Methods in Computational Molecular Physics*, edited by G. H. F. Diercksen and S. Wilson, pages 189–207. D. Reidel, Dordrecht, 1983.
- [89] B. H. Lengsfeld, A. D. McLean, M. Yoshimine, and B. Liu, *J. Chem. Phys.* **79**, 1891 (1983).
- [90] M. M. Steiner, W. Wenzel, K. G. Wilson, and J. W. Wilkins, *Chem. Phys. Lett.* **231**, 263 (1994).
- [91] P. Saxe, D. J. Fox, H. F. Schaefer, and N. C. Handy, *J. Chem. Phys.* **77**, 5584 (1982).
- [92] B. Huron, J. P. Malrieu, and P. Rancurel, *J. Chem. Phys.* **58**, 5745 (1973).

- [93] S. Evangelisti, J. P. Daudey, and J. P. Malrieu, *Chem. Phys.* **75**, 91 (1983).
- [94] Z. Gershgorin and I. Shavitt, *Int. J. Quantum Chem.* **2**, 751 (1968).
- [95] V. Sychrovský and P. Čársky, *Mol. Phys.* **88**, 1137 (1996).
- [96] O. Goscinski and E. Brändas, *Chem. Phys. Lett.* **2**, 299 (1968).
- [97] J. Cullen and M. C. Zerner, *Int. J. Quantum Chem. Symp.* **9**, 343 (1975).
- [98] J. M. Cullen and M. C. Zerner, *Int. J. Quantum Chem.* **33**, 373 (1988).
- [99] W. Meyer, Configuration expansion by means of pseudonatural orbitals, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer, pages 413–446. Plenum Press, New York, 1977.
- [100] P. E. M. Siegbahn, *Int. J. Quantum Chem.* **18**, 1229 (1980).
- [101] H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **88**, 5803 (1988).
- [102] P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.* **145**, 514 (1988).
- [103] J. C. Greer, *J. Chem. Phys.* **103**, 1821 (1995).
- [104] A. Bunge, *J. Chem. Phys.* **53**, 20 (1970).
- [105] C. F. Bender and H. F. Schaefer, *J. Chem. Phys.* **55**, 4798 (1971).
- [106] A. D. McLean and B. Liu, *J. Chem. Phys.* **58**, 1066 (1973).
- [107] C. W. Bauschlicher and H. Partridge, *Theor. Chim. Acta* **85**, 255 (1993).
- [108] E. R. Davidson, *J. Comp. Phys.* **17**, 87 (1975).
- [109] P. J. Knowles and N. C. Handy, *Chem. Phys. Lett.* **111**, 315 (1984).
- [110] E. R. Davidson, Configuration interaction description of electron correlation, in *The World of Quantum Chemistry*, edited by R. Daudel and B. Pullman. D. Reidel, Dordrecht, 1974.
- [111] S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.* **8**, 61 (1974).
- [112] E. R. Davidson and D. W. Silver, *Chem. Phys. Lett.* **52**, 403 (1977).
- [113] L. Meissner, *Chem. Phys. Lett.* **146**, 204 (1988).
- [114] W. Duch and G. H. F. Diercksen, *J. Chem. Phys.* **101**, 3018 (1994).

- [115] P. J. Bruna, S. D. Peyerimhoff, and R. J. Buenker, *Chem. Phys. Lett.* **72**, 278 (1980).
- [116] W. Kutzelnigg, Pair correlation theories, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer, pages 129–188. Plenum Press, New York, 1977.
- [117] R. Ahlrichs, P. Scharf, and C. Ehrhardt, *J. Chem. Phys.* **82**, 890 (1985).
- [118] J. A. Pople, M. Head-Gordon, and K. Raghavachari, *J. Chem. Phys.* **87**, 5968 (1987).
- [119] R. J. Gdanitz and R. Ahlrichs, *Chem. Phys. Lett.* **143**, 413 (1988).
- [120] P. G. Szalay and R. J. Bartlett, *J. Chem. Phys.* **103**, 3600 (1995).
- [121] L. Füst-Molnár and P. G. Szalay, *J. Phys. Chem.* **100**, 6288 (1996).
- [122] A. Povill, R. Caballol, J. Rubio, and J. P. Malrieu, *Chem. Phys. Lett.* **209**, 126 (1993).
- [123] A. Povill, J. Rubio, R. Caballol, and J. P. Malrieu, *Chem. Phys. Lett.* **218**, 283 (1994).
- [124] J. P. Malrieu, J. P. Daudey, and R. Caballol, *J. Chem. Phys.* **101**, 8908 (1994).
- [125] D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **103**, 4572 (1995).
- [126] R. P. Hosteny, T. H. Dunning, R. R. Gilman, A. Pipano, and I. Shavitt, *J. Chem. Phys.* **62**, 4764 (1975).
- [127] V. R. Saunders and J. H. van Lenthe, *Mol. Phys.* **48**, 923 (1983).
- [128] E. R. Davidson, Perspectives on *ab initio* calculations, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd. VCH Publishers, New York, 1990.
- [129] M. Page, P. Saxe, G. F. Adams, and B. H. Lengsfeld, *J. Chem. Phys.* **81**, 434 (1984).
- [130] F. W. Bobrowicz and W. A. Goddard, The self-consistent field equations for generalized valence bond and open-shell Hartree-Fock wave functions, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer, pages 79–127. Plenum Press, New York, 1977.

- [131] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, *Chem. Phys.* **48**, 157 (1980).
- [132] W. J. Hunt and W. A. Goddard, *Chem. Phys. Lett.* **3**, 414 (1969).
- [133] C. W. Bauschlicher, *J. Chem. Phys.* **72**, 880 (1980).
- [134] D. Feller and E. R. Davidson, *J. Chem. Phys.* **74**, 3977 (1981).
- [135] E. R. Davidson, *Reduced Density Matrices in Quantum Chemistry*. Academic Press, New York, 1976.
- [136] I. Shavitt, B. J. Rosenberg, and S. Palalikit, *Int. J. Quantum Chem. Symp.* **10**, 33 (1976).
- [137] O. Parisel and Y. Ellinger, *Chem. Phys.* **189**, 1 (1994).
- [138] M. R. A. Blomberg and B. Liu, *J. Chem. Phys.* **89**, 6870 (1988).
- [139] K. Balasubramanian, *J. Chem. Phys.* **89**, 5731 (1988).
- [140] S. Diner, J. P. Malrieu, P. Claverie, and F. Jordan, *Chem. Phys. Lett.* **2**, 319 (1968).
- [141] J. P. Malrieu, The pcilo method, in *Semiempirical Methods of Electronic Structure Calculation*, edited by G. A. Segal, volume 7 of *Modern Theoretical Chemistry*, pages 69–103. Plenum Press, New York, 1977.
- [142] S. F. Boys, Localized orbitals and localized adjustment functions, in *Quantum Theory of Atoms, Molecules, and the Solid State*, edited by P.-O. Löwdin, pages 253–262. Academic Press, New York, 1966.
- [143] S. Saebø and P. Pulay, *Chem. Phys. Lett.* **113**, 13 (1985).
- [144] S. Saebø and P. Pulay, *J. Chem. Phys.* **88**, 1884 (1988).
- [145] I. D. Petsalakis, G. Theodorakopoulos, C. A. Nicolaides, and R. Buenker, *J. Chem. Phys.* **81**, 3161 (1984).
- [146] P.-Å. Malmqvist, *Int. J. Quantum Chem.* **30**, 479 (1986).
- [147] R. Broer and W. C. Nieuwpoort, *Theor. Chim. Acta* **73**, 405 (1988).
- [148] P.-Å. Malmqvist and B. O. Roos, *Chem. Phys. Lett.* **155**, 189 (1989).
- [149] R. J. Buenker and S. D. Peyerimhoff, *J. Chem. Phys.* **53**, 1368 (1970).

- [150] W. D. Allen and H. F. Schaefer, *J. Chem. Phys.* **87**, 7076 (1987).
- [151] G. Fitzgerald and H. F. Schaefer, *J. Chem. Phys.* **83**, 1162 (1985).
- [152] T. J. Lee, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **87**, 7062 (1987).
- [153] S. A. Houlden and I. G. Csizmadia, *Theor. Chim. Acta* **30**, 209 (1973).
- [154] M. Head-Gordon and T. J. Lee, Single reference coupled cluster and perturbation theories of electronic excitation energies, in *Modern Ideas in Coupled Cluster Theory*, edited by R. J. Bartlett. World Scientific, Singapore, 1997.
- [155] H. Nakatsuji, *Chem. Phys. Lett.* **67**, 329 (1979).
- [156] H. Nakatsuji, *Chem. Phys. Lett.* **67**, 334 (1979).
- [157] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *J. Chem. Phys.* **96**, 5483 (1990).
- [158] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96**, 1218 (1992).
- [159] K. Andersson and B. O. Roos, Multiconfigurational second-order perturbation theory, in *Modern Electronic Structure Theory*, edited by D. R. Yarkony, volume 2 of *Advanced Series in Physical Chemistry*, pages 55–109. World Scientific, Singapore, 1995.
- [160] J. B. Foresman, M. Head-Gordon, J. A. Pople, and M. J. Frisch, *J. Phys. Chem.* **96**, 135 (1992).
- [161] M. Head-Gordon, R. J. Rico, M. Oumi, and T. J. Lee, *Chem. Phys. Lett.* **219**, 21 (1994).
- [162] J. F. Stanton, J. Gauss, N. Ishikawa, and M. Head-Gordon, *J. Chem. Phys.* **103**, 4160 (1995).
- [163] D. Maurice and M. Head-Gordon, *Int. J. Quantum Chem. Symp.* **29**, 361 (1995).
- [164] D. Maurice and M. Head-Gordon, *J. Phys. Chem.* **100**, 6131 (1996).
- [165] J. Wasilewski, *Int. J. Quantum Chem.* **36**, 503 (1989).
- [166] K. Dowd, *High Performance Computing*. O'Reilly, Sebastopol, CA, 1993.

- [167] J. A. Pople, R. Seeger, and R. Krishnan, *Int. J. Quantum Chem. Symp.* **11**, 149 (1977).
- [168] H.-J. Werner and E.-A. Reinsch, *J. Chem. Phys.* **76**, 3144 (1982).
- [169] R. Ahlrichs, in *Proceedings of the 5th Seminar on Computational Methods in Quantum Chemistry, Groningen*, edited by P. T. van Duijn and W. C. Nieuwpoort. Max Planck Institut, Garching, Munchen, 1981.
- [170] S. Wilson, Four-index transformations, in *Electron Correlation in Atoms and Molecules*, edited by S. Wilson, volume 1 of *Methods in Computational Chemistry*. Plenum, New York, 1987.
- [171] M. Häser, J. Almlöf, and M. W. Feyereisen, *Theor. Chim. Acta* **79**, 115 (1991).
- [172] M. Yoshimine, The use of direct access devices in problems requiring the reordering of long lists of data, Technical Report RJ-555, IBM Research Laboratory, San Jose, California, 1969.
- [173] P.-Å. Malmqvist, Mathematical tools in quantum chemistry, in *Lecture Notes in Quantum Chemistry: European Summer School in Quantum Chemistry*, edited by B. O. Roos, volume 58 of *Lecture Notes in Chemistry*, pages 1–35. Springer-Verlag, New York, 1992.
- [174] B. Liu, The simultaneous expansion method for the iterative solution of several of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices, Technical Report LBL-8158, Lawrence Berkeley Laboratory, University of California, Berkeley, 1978.
- [175] C. W. Murray, S. C. Racine, and E. R. Davidson, *J. Comput. Phys.* **103**, 382 (1992).
- [176] R. Shepard, *J. Comput. Chem.* **11**, 45 (1990).
- [177] H. Dachsels and H. Lischka, *Theor. Chim. Acta* **92**, 339 (1995).
- [178] J. H. van Lenthe and P. Pulay, *J. Comput. Chem.* **11**, 1164 (1990).
- [179] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*. Manchester University Press, Manchester, 1992.
- [180] G. L. G. Sleijpen and H. A. van der Vorst, A generalized Jacobi-Davidson iteration method for linear eigenvalue problems, Technical Report Preprint 856, Dept. Math., Utrecht University, 1994.

- [181] J. M. Bofill and J. M. Anglada, *Chem. Phys.* **183**, 19 (1994).
- [182] H. J. J. van Dam, J. H. van Lenthe, G. L. G. Sleijpen, and H. A. van der Vorst, *J. Comput. Chem.* **17**, 267 (1996).
- [183] A. O. Mitrushenkov, *Chem. Phys. Lett.* **217**, 559 (1994).
- [184] C. Ribbing and C. Daniel, *J. Chem. Phys.* **100**, 6591 (1994).
- [185] R. Pauncz, *Spin Eigenfunctions: Construction and Use*. Plenum, New York, 1979.
- [186] R. Pauncz, *The Symmetric Group in Quantum Chemistry*. CRC Press, Boca Raton, 1995.
- [187] W. Duch, *Chem. Phys. Lett.* **162**, 56 (1989).
- [188] W. Duch and J. Meller, *Int. J. Quantum Chem.* **50**, 243 (1994).
- [189] T. J. Martinez, A. Mehta, and E. A. Carter, *J. Chem. Phys.* **97**, 1876 (1992).
- [190] T. J. Martinez, A. Mehta, and E. A. Carter, *J. Chem. Phys.* **99**, 4238(E) (1993).
- [191] T. J. Martinez and E. A. Carter, *J. Chem. Phys.* **102**, 7564 (1995).
- [192] G. Reynolds, T. J. Martinez, and E. A. Carter, *J. Chem. Phys.* **105**, 6455 (1996).
- [193] B. Roos, *Chem. Phys. Lett.* **15**, 153 (1972).
- [194] P. E. M. Siegbahn, *J. Chem. Phys.* **72**, 1647 (1980).
- [195] P. E. M. Siegbahn, *J. Chem. Phys.* **70**, 5391 (1979).
- [196] H. Lischka, R. Shepard, F. B. Brown, and I. Shavitt, *Int. J. Quantum Chem. Symp.* **15**, 91 (1981).
- [197] M. Schüler, T. Kovar, H. Lischka, R. Shepard, and R. J. Harrison, *Theor. Chim. Acta* **84**, 489 (1993).
- [198] H. Dachsel, H. Lischka, R. Shepard, J. Nieplocha, and R. J. Harrison, *J. Comput. Chem.* **18**, 430 (1997).
- [199] W. Duch and J. Karwowski, *Computer Phys. Rept.* **2**, 93 (1985).

- [200] W. Duch, *Chem. Phys. Lett.* **124**, 442 (1986).
- [201] A. Guldberg, S. Rettrup, G. L. Bendazzoli, and P. Palmieri, *Int. J. Quantum Chem. Symp.* **21**, 513 (1987).
- [202] W. Duch and J. Karwowski, *Theor. Chim. Acta* **71**, 187 (1987).
- [203] W. Duch, *Int. J. Quantum Chem. Symp.* **234**, 683 (1990).
- [204] N. Flocke, *Int. J. Quantum Chem.* **61**, 747 (1997).
- [205] M. W. Lucht, *Phys. Rev. A* **54**, 266 (1996).
- [206] R. D. Kent and M. Schlesinger, *Phys. Rev. A* **50**, 186 (1994).
- [207] Z. Wen, Y. Wang, and H. Lin, *Chem. Phys. Lett.* **230**, 41 (1994).
- [208] M. D. Gould and J. S. Battle, *J. Chem. Phys.* **99**, 5961 (1993).
- [209] M. D. Gould and J. S. Battle, *J. Chem. Phys.* **99**, 5983 (1993).
- [210] R. D. Kent, M. Schlesinger, and I. Shavitt, *Int. J. Quantum Chem.* **41**, 89 (1992).
- [211] P. J. Burton and M. D. Gould, *J. Chem. Phys.* **96**, 5261 (1992).
- [212] B. Jeziorski, J. Paldus, and P. Jankowski, *Int. J. Quantum Chem.* **56**, 129 (1995).
- [213] P. Saxe, H. F. Schaefer, and N. C. Handy, *Chem. Phys. Lett.* **79**, 202 (1981).
- [214] P. J. Knowles and N. C. Handy, *Computer Phys. Commun.* **54**, 75 (1989).
- [215] A. Povill, J. Rubio, and F. Illas, *Theor. Chim. Acta* **82**, 229 (1992).
- [216] A. O. Mitrushenkov and Y. Y. Dmitriev, *Chem. Phys. Lett.* **235**, 410 (1995).
- [217] S. Zarrabian, C. R. Sarma, and J. Paldus, *Chem. Phys. Lett.* **155**, 183 (1989).
- [218] P. J. Knowles and N. C. Handy, *J. Chem. Phys.* **91**, 2396 (1989).
- [219] F. Illas, J. Rubio, J. M. Ricart, and P. S. Bagus, *J. Chem. Phys.* **95**, 1877 (1991).

- [220] A. Povill and J. Rubio, *Theor. Chim. Acta* **92**, 305 (1995).
- [221] R. Caballol and J. P. Malrieu, *Chem. Phys. Lett.* **188**, 543 (1992).
- [222] S. Krebs and R. J. Buenker, *J. Chem. Phys.* **103**, 5613 (1995).
- [223] R. J. Harrison, *J. Chem. Phys.* **94**, 5021 (1991).
- [224] A. V. Luzanov, A. L. Wulfov, and V. O. Krouglov, *Chem. Phys. Lett.* **197**, 614 (1992).
- [225] A. L. Wulfov, *Chem. Phys. Lett.* **255**, 300 (1996).
- [226] A. O. Mitrushenkov, *J. Chem. Phys.* **105**, 10487 (1996).
- [227] R. J. Cave, S. S. Xantheas, and D. Feller, *Theor. Chim. Acta* **83**, 31 (1992).
- [228] A. L. Wulfov, *Chem. Phys. Lett.* **263**, 79 (1996).
- [229] M. L. Leininger, C. D. Sherrill, W. D. Allen, and H. F. Schaefer, manuscript in preparation.
- [230] W. Duch, *GRMS or Graphical Representation of Model Spaces*, volume 1. Springer-Verlag, Berlin, 1986.
- [231] D. H. Lehmer, in *Applied Combinatorial Mathematics*, edited by E. F. Beckenbach. Wiley, New York, 1964.
- [232] J. Olsen, Private communication, 1995.
- [233] C. W. Bauschlicher, S. R. Langhoff, P. R. Taylor, and H. Partridge, *Chem. Phys. Lett.* **126**, 436 (1986).
- [234] C. W. Bauschlicher, S. R. Langhoff, P. R. Taylor, N. C. Handy, and P. J. Knowles, *J. Chem. Phys.* **85**, 1469 (1986).
- [235] C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.* **85**, 2779 (1986).
- [236] C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.* **85**, 6510 (1986).
- [237] C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.* **86**, 858 (1987).
- [238] C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.* **86**, 1420 (1987).
- [239] C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.* **87**, 2844 (1987).
- [240] C. W. Bauschlicher and S. R. Langhoff, *J. Chem. Phys.* **87**, 2919 (1987).

- [241] C. W. Bauschlicher and S. R. Langhoff, *J. Chem. Phys.* **86**, 5595 (1987).
- [242] C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.* **86**, 5600 (1987).
- [243] J. D. Watts, I. Cernusak, J. Noga, R. J. Bartlett, C. W. Bauschlicher, T. J. Lee, A. P. Rendell, and P. R. Taylor, *J. Chem. Phys.* **93**, 8875 (1990).
- [244] C. W. Bauschlicher, *J. Chem. Phys.* **86**, 5591 (1987).
- [245] P. R. Taylor, C. W. Bauschlicher, and S. R. Langhoff, *J. Phys. B* **21**, L333 (1988).
- [246] C. W. Bauschlicher, *J. Phys. Chem.* **92**, 3020 (1988).
- [247] F. Illas, J. M. Ricart, J. Rubio, and P. S. Bagus, *J. Chem. Phys.* **93**, 4982 (1990).
- [248] J. Casanovas, J. Rubio, and F. Illas, *Chem. Phys. Lett.* **180**, 578 (1991).
- [249] C. W. Bauschlicher, S. R. Langhoff, and P. R. Taylor, *J. Chem. Phys.* **88**, 2540 (1988).
- [250] C. W. Bauschlicher, S. R. Langhoff, and P. R. Taylor, *J. Chem. Phys.* **93**, 502 (1990).
- [251] O. Christiansen, J. Olsen, P. Jørgensen, H. Koch, and P.-Å. Malmqvist, *Chem. Phys. Lett.* **261**, 369 (1996).
- [252] J. Olsen, A. M. Sánchez de Meras, H. J. Aa. Jensen, and P. Jørgensen, *Chem. Phys. Lett.* **154**, 380 (1989).
- [253] C. W. Bauschlicher and P. R. Taylor, *Theor. Chim. Acta* **71**, 263 (1987).
- [254] C. W. Bauschlicher and S. R. Langhoff, *Chem. Phys. Lett.* **135**, 67 (1987).
- [255] C. W. Bauschlicher and S. R. Langhoff, *Theor. Chim. Acta* **73**, 43 (1988).
- [256] C. W. Bauschlicher and S. R. Langhoff, *Theor. Chim. Acta* **79**, 93 (1991).
- [257] H. Koch and R. J. Harrison, *J. Chem. Phys.* **95**, 7479 (1991).
- [258] C. W. Bauschlicher, S. R. Langhoff, H. Partridge, and D. P. Chong, *J. Chem. Phys.* **89**, 2985 (1988).

- [259] O. Vahtras, H. Ågren, P. Jørgensen, H. J. A. Jensen, T. Helgaker, and J. Olsen, *J. Chem. Phys.* **96**, 2118 (1992).
- [260] M. M. Marino, W. C. Ermler, C. W. Kern, and V. E. Bondybey, *J. Chem. Phys.* **96**, 3756 (1992).
- [261] P. R. Taylor, A. Komornicki, and D. A. Dixon, *J. Am. Chem. Soc.* **111**, 1259 (1988).
- [262] J. M. Anglada and J. M. Bofill, *Theor. Chim. Acta* **92**, 369 (1995).
- [263] Y. Yamaguchi, J. F. Gaw, R. B. Remington, and H. F. Schaefer, *J. Chem. Phys.* **86**, 5072 (1987).
- [264] C. D. Sherrill, T. J. Van Huis, Y. Yamaguchi, and H. F. Schaefer, *J. Mol. Struct. (Theochem)* **400**, 139 (1997).
- [265] T. J. Van Huis, M. L. Leininger, C. D. Sherrill, and H. F. Schaefer, manuscript in preparation.
- [266] C. D. Sherrill, M. L. Leininger, T. J. Van Huis, and H. F. Schaefer, *J. Chem. Phys.*, submitted.
- [267] D. E. Woon, *J. Chem. Phys.* **100**, 2838 (1994).
- [268] R. Moszynski, B. Jeziorski, A. Ratkiewicz, and S. Rybak, *J. Chem. Phys.* **99**, 8856 (1993).
- [269] R. Moszynski, B. Jeziorski, and K. Szalewicz, *J. Chem. Phys.* **100**, 1312 (1994).
- [270] I. Mayer, A. Vibok, and P. Valiron, *Chem. Phys. Lett.* **224**, 166 (1994).
- [271] J. V. Burda, R. Zahradnik, P. Hobza, and M. Urban, *Mol. Phys.* **89**, 425 (1996).
- [272] M. R. A. Blomberg, P. E. M. Siegbahn, and B. O. Roos, *Int. J. Quantum Chem. Symp.* **14**, 229 (1980).
- [273] R. A. Chiles and C. E. Dykstra, *J. Chem. Phys.* **74**, 4544 (1981).
- [274] V. E. Bondybey, *Chem. Phys. Lett.* **109**, 436 (1984).
- [275] H. Partridge, C. W. Bauschlicher, L. G. M. Pettersson, A. D. McLean, B. Liu, M. Yoshimine, and A. Komornicki, *J. Chem. Phys.* **92**, 5377 (1990).

- [276] W. Koch, B. Liu, and G. Frenking, *J. Chem. Phys.* **92**, 2464 (1990).
- [277] S. R. Langhoff, C. W. Bauschlicher, and P. R. Taylor, *Chem. Phys. Lett.* **135**, 543 (1987).
- [278] J. Almlöf, B. J. DeLeeuw, P. R. Taylor, C. W. Bauschlicher, and P. Siegbahn, *Int. J. Quantum Chem. Symp.* **23**, 345 (1989).
- [279] H. Partridge, S. R. Langhoff, C. W. Bauschlicher, and D. W. Schwenke, *J. Chem. Phys.* **88**, 3174 (1988).
- [280] C. W. Bauschlicher, S. R. Langhoff, and L. G. M. Pettersson, *J. Chem. Phys.* **89**, 5747 (1988).
- [281] K. Balasubramanian, *J. Chem. Phys.* **93**, 507 (1990).
- [282] K. Balasubramanian, *J. Phys. Chem.* **94**, 7764 (1990).
- [283] K. K. Das and K. Balasubramanian, *J. Chem. Phys.* **93**, 5883 (1990).
- [284] Z. Ma and K. Balasubramanian, *Chem. Phys. Lett.* **181**, 467 (1991).
- [285] K. Balasubramanian, *Chem. Phys. Lett.* **164**, 231 (1989).
- [286] K. Balasubramanian and J. X. Tao, *J. Chem. Phys.* **94**, 3000 (1991).
- [287] D.-W. Liao and K. Balasubramanian, *Chem. Phys. Lett.* **213**, 174 (1993).
- [288] K. Balasubramanian, *J. Chem. Phys.* **93**, 8061 (1990).
- [289] P. Saxe, H. F. Schaefer, and N. C. Handy, *J. Phys. Chem.* **85**, 745 (1981).
- [290] C. W. Bauschlicher, S. R. Langhoff, and P. R. Taylor, *J. Chem. Phys.* **87**, 387 (1987).
- [291] A. D. McLean, P. R. Bunker, R. M. Escibano, and P. Jensen, *J. Chem. Phys.* **87**, 2166 (1987).
- [292] P. Jensen and P. R. Bunker, *J. Chem. Phys.* **89**, 1327 (1988).
- [293] E. R. Davidson, D. Feller, and P. Phillips, *Chem. Phys. Lett.* **76**, 416 (1980).
- [294] N. C. Handy, Y. Yamaguchi, and H. F. Schaefer, *J. Chem. Phys.* **84**, 4481 (1986).

- [295] K. Balasubramanian, *Chem. Phys. Lett.* **127**, 585 (1986).
- [296] K. K. Das and K. Balasubramanian, *J. Chem. Phys.* **91**, 2433 (1989).
- [297] K. Balasubramanian, *J. Chem. Phys.* **91**, 2443 (1989).
- [298] K. K. Das and K. Balasubramanian, *J. Phys. Chem.* **95**, 3979 (1991).
- [299] K. Balasubramanian, *Chem. Phys. Lett.* **204**, 601 (1993).
- [300] G. Parlant, J. Senekowitsch, S. V. O'Neil, and D. R. Yarkony, *J. Chem. Phys.* **94**, 7208 (1991).
- [301] M. R. Manaa and D. R. Yarkony, *J. Chem. Phys.* **95**, 6562 (1991).
- [302] H. Hettrema and D. R. Yarkony, *J. Chem. Phys.* **100**, 8991 (1994).
- [303] D. L. Strout and G. E. Scuseria, *J. Chem. Phys.* **96**, 9025 (1992).
- [304] H. J. Aa. Jensen, P. Jørgensen, T. Helgaker, and J. Olsen, *Chem. Phys. Lett.* **162**, 355 (1989).
- [305] A. M. Sánchez de Meras, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen, *Chem. Phys. Lett.* **186**, 379 (1991).
- [306] D. Sundholm and J. Olsen, *Chem. Phys. Lett.* **198**, 526 (1992).
- [307] D. Sundholm and J. Olsen, *Chem. Phys. Lett.* **217**, 451 (1994).
- [308] T. J. Lee, R. B. Remington, Y. Yamaguchi, and H. F. Schaefer, *J. Chem. Phys.* **89**, 408 (1988).
- [309] G. E. Scuseria, T. P. Hamilton, and H. F. Schaefer, *J. Chem. Phys.* **92**, 568 (1990).
- [310] G. E. Scuseria and H. F. Schaefer, *J. Chem. Phys.* **88**, 7024 (1988).
- [311] H. Tanaka and K. Nishimoto, *J. Mol. Struct. (Theochem)* **50**, 297 (1988).
- [312] P. Du, D. A. Hrovat, and W. T. Borden, *J. Am. Chem. Soc.* **108**, 8086 (1986).
- [313] R. A. King, C. D. Sherrill, and H. F. Schaefer, *Spectrochim. Acta, Part A* **53**, 1163 (1997).
- [314] B. C. Hoffman, C. D. Sherrill, and H. F. Schaefer, *Chem. Phys. Lett.* , submitted.
- [315] M. L. Leininger and H. F. Schaefer, *J. Chem. Phys.*, in press.