CHEMICAL REVIEWS

REVIEW

pubs.acs.org/CR

Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications†

Péter G. Szalay‡

Laboratory for Theoretical Chemistry, Institute of Chemistry, Eötvös Loránd University, P. O. Box 32, H-1518 Budapest, Hungary

Thomas Müller§

Jülich Supercomputer Centre, Institute of Advanced Simulation, Forschungszentrum Jülich, D-52425 Jülich, Germany

Gergely Gidofalvi∥

Department of Chemistry and Biochemistry, Gonzaga University, 502 East Boone Avenue, Spokane, Washington 99258-0102, United States

Hans Lischka⊥

Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, United States
Institute for Theoretical Chemistry, University of Vienna, Waehringerstrasse 17, A-1090 Vienna, Austria

Ron Shepard*

Chemical Sciences and Engineering Division, Argonne National Laboratory, Argonne, Illinois 60439, United States

CONTENTS

Special Issue: 2012 Quantum Chemistry

Received: September 21, 2011
Published: December 28, 2011

dx.doi.org/10.1021/cr200137a | *Chem. Rev.* 2012, 112, 108–181

1. INTRODUCTION

The configuration interaction (CI) method is a general procedure to compute approximate solutions to the electronic Schrödinger equation. The wave function is written as a linear expansion

$$|\Psi\rangle = \sum_{j=1}^{N_{\mathrm{dim}}} c_j |j\rangle \tag{1}$$

with expansion functions $|j\rangle$ and expansion coefficients c_j. A common feature of CI methods is that the expansion coefficients are determined variationally. A trial energy is typically an expectation value, written as the Rayleigh quotient

$$E^{\mathrm{trial}} = \rho = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{2}$$

and the CI energy is determined from the relation $\partial\rho/\partial\mathbf{c} = \mathbf{0}$. This leads to the standard hermitian eigenvalue equation

$$\mathbf{H}\mathbf{c}^k = E_k\mathbf{c}^k \tag{3}$$

with the matrix elements $H_{jl} = \langle j|H|l\rangle$. The ground state corresponds to the lowest energy solution, and excited states correspond to higher solutions. Thus the CI method is conceptually very general and simple, and therefore it has been used since the beginning of molecular quantum mechanics. In the 1950s and 1960s it was the primary source of correlated results, and these calculations showed that electron correlation is essential to understand certain properties of atoms and molecules. The review article by Shavitt[1] covers the important aspects of the CI method up through 1977, and the later reviews by Duch,[2] Shavitt,[3] Sherrill and Schaefer,[4] Čársky,[5] and Karwowski and Shavitt[6] cover important developments through 2003. In this paper, the important basic aspects of CI methods such as expansion basis choices, truncation schemes, choice of orbitals, diagonalization procedures, and size-consistency issues will be summarized. Recent developments on gradient theory, calculation of molecular properties, and nonadiabatic coupling between electronic states, as well as relativistic and spin-orbit effects will be covered in more detail.

If all possible expansion terms are included in the expansion of eq 1, then the result is the full-CI wave function. The solutions to eq 3 in this case correspond to the exact solutions to the Schrödinger equation within the given orbital basis. These are the energies and the wave functions that all approximate orbital-based electronic structure methods attempt to mimic. The dimension of the full-CI expansion grows approximately as n^N for N electrons and n molecular orbitals and for $N < n$. Consequently in practice, full-CI wave functions can be computed only for small molecular systems and for relatively small orbital basis sets, but these calculations serve as important benchmarks to assess the accuracy, reliability, and characteristics of other approximate methods.

There are N_{dim} linearly independent solutions to eq 3, and, for ordered sequences of eigenvalues, truncated expansions satisfy the bounds

$$E_k^{(\mathrm{full-CI})} \le E_k^{(N_{\mathrm{dim}})} \tag{4}$$

$$E_k^{(N_{\mathrm{dim}})} \le E_{k+1}^{(N_{\mathrm{dim}}+1)} \le E_{k+1}^{(N_{\mathrm{dim}})} \tag{5}$$

The first equation shows that the approximate energies are always higher than the exact full-CI values, and the second equation shows that convergence to those values occurs monotonically with increasing wave function flexibility. These relations allow a correspondence to be established between the exact full-CI eigenvalues and the eigenvalues from a truncated expansion, and it allows the convergence of a particular state to the exact limit to be monitored as a function of the expansion dimension. In addition, these bounds relations allow the convergence with respect to orbital basis sets to be assessed and, in some cases, extrapolated to the complete basis set (CBS) limit. In this manner, both ground and excited states can be computed, and given these wave functions, arbitrary expectation values and transition properties may be evaluated in a straightforward way. This generality and flexibility are important features of CI methods. Equations 4 and 5 are satisfied at each molecular conformation **R**, and thus these bounds relations apply not just to isolated molecular conformations but also to description of the behavior of the entire potential energy surfaces (PESs).

In most CI applications, eq 3 is solved simultaneously for both the wave function expansion coefficients $\mathbf{c}^k$ and for the energy E_k. However, consider the situation in which the coefficients are known. In this case, the energy can be evaluated by considering only a single row of eq 3.

$$(E_k - H_{jj})c_j^k = \sum_{l(\neq j)} H_{jl}c_l^k \tag{6}$$

If $c_j^k \neq 0$, then E_k may be evaluated with effort that scales only as the number of nonzero elements in the jth row of the matrix **H**. There are only, at most, $\sim N^2n^2$ nonzero elements in a row of the **H** matrix (including even the full-CI **H** matrix); thus the form of eq 6 suggests a very economical way to compute an energy, given the $\mathbf{c}^k$ from some separate computational procedure.

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

This transition-energy formula is the basis for many electronic structure methods. The coefficients associated with the nonzero elements of a single row are estimated in some fashion, and then the energy is evaluated according to eq 6. Several of these methods are discussed in section 2.

Many electronic structure methods may be characterized as either single-reference (SR) or multireference (MR) approaches. In a SR approach, a single Slater determinant is chosen as a reference function, and some procedure is used to determine the important Hamiltonian interactions with this reference function. This typically involves either empirical or a priori selection procedures that are based on perturbation theory, explicit diagonalization within small subspaces, and combinations of such approaches. The A_k and B_k procedures described in ref 1 are typical examples of methods that are used in empirical selection approaches, and excitation-limited expansions such as SR-CISD, SR-CISDT, and SR-CISDTQ, etc., are typical examples of a priori selection approaches which include respectively single + double, single + double + triple, and single + double + triple + quaduple excitations with respect to the reference function. The fundamental problem with these SR approaches is that different regions of a molecular PES are often dominated by different determinants. This means that an expansion based on the dominant function at one conformation will result in a poor description of the wave function at other conformations. SR approaches also have difficulty describing multiple states because the dominant reference function for one electronic state is often not appropriate for describing other electronic states.

In a multireference approach,[7] all of the possible important determinants, or configuratoin-state functions (CSFs), are first identified. This may involve either an a priori approach or a numerical selection procdedure. For example, in an MRCI expansion based on an a priori full optimized reaction space (FORS) or complete active space (CAS) reference, all possible determinants constructed from the set of *active* orbitals are treated as reference functions; the choice of the active orbital space alone entirely determines the expansion space. In an empirical selection approach, the important regions of the PESs are scanned, perhaps with a small orbital basis set and with some relatively cheap electronic structure method, and any important determinants would be identified and selected numerically. Once the reference space is determined, all of the individual determinants are treated equivalently to generate the MRCI expansion space. This may again be done either by empirical approach, with numerical selection procedures, or with an a priori approach (see, e.g., ref 1 for further discussions). For example, in a MR-CISD expansion, all single and double excitations from each of the reference functions are included. This additional flexibility, relative to the SR approach, allows the wave function to describe different regions of the PESs in a balanced manner and to simultaneously describe multiple electronic states. The practical difficulty with such MR expansions is that they require more effort than the analogous excitation-limited SR approach. This limits the size of the molecules and/or the size of the molecular orbital basis sets that can be studied. In some cases, even the size of the active orbital space, or the reference CSF space in general, must be restricted due to practical limitations. These considerations associated with MR approaches lead to the search for methods that require less computational effort but still share the important advantages of the MR approaches. These include, for example, internally contracted approaches, externally contracted approaches, and various fragment approaches in which the final wave function is constructed as products (or more generally sums of products) of molecular fragments. Such approaches will be discussed further in section 2.

To our knowledge, the first MRCI calculation was performed by Liu[8] on the H_3 molecule. Liu included in the expansion space all CSFs which interacted with any of the reference functions. Thus, practically speaking, it was an MR-CISD expansion. The reference functions were selected emprically.

Since the expansion functions, either Slater determinants or CSFs, depend on the choice of molecular orbitals, this choice will greatly influence the quality of the wave function. Hence, the choice of an orbital set has played an important role throughout the history of the CI method. In SR approaches, the orbitals are typically taken as the canonical SCF orbitals of the reference function. However, other choices have involved natural orbitals,[9,10] localized orbitals,[11−14] and pair natural orbitals[15−17] that are chosen separately for different fragments of the wave function, or orbitals chosen to mimic the virtual orbitals of a hypothetical molecular ion.[1,6,18] In some situations, such as reactions involving the cleavage of bonds, it is clear that SCF orbitals are inappropriate due to artifactual charge contamination of the reference wave function. In MR approaches, the orbitals are usually taken from a multiconfiguration self-consistent field (MCSCF) calculation.[19,20] In these cases, the reference space is chosen to have sufficient flexibility in order to qualitatively describe any important valence correlation effects such as bond-breaking, avoided crossings, and spin-recoupling processes. The reference space for the MRCI expansion is usually taken as the entire MCSCF expansion space, but sometimes, due primarily to computational limitations, only a selected subset of the entire MCSCF expansion space is chosen; at other times, due to some inadequacy of the MCSCF expansion, additional determinants might be added. Orbital invariance properties of the MCSCF and MRCI expansions are often further utilized to refine the molecular orbital expansion space. For example, orbital localizations might be applied to both the occupied and the virtual orbitals in the MCSCF expansion. Such procedures leave the MCSCF wave function itself unchanged, but the subsequent MRCI expansion can exploit this orbital localization through reduction in the dimension of the CSF expansion space, approximations within the Hamiltonian operator through the systematic neglect of small interactions, or both. In other situations, the MCSCF wave function expansion itself is formulated in terms of localized orbitals; examples of this approach include GVB-RCI expansions and other direct-product types of expansion spaces.[19]

Because the eigenvalues from eq 3 are defined and computed variationally, first-order response properties can be computed with the Hellmann−Feynman theorem[21] as simple expectation values $\partial E_k/\partial\lambda = \langle ci^k|\partial H/\partial\lambda|ci^k\rangle$. As discussed in section 2, this expression uses the second-quantized Hamiltonian operator which is defined in terms of the molecular orbitals. This requires much less effort than an otherwise comparable nonvariational method. One important application of this feature is the computation of analytic energy gradients, $\partial E_k/\partial\mathbf{R}$, which are used for the optimization of equilibrium molecular geometries, the computation of saddle-point geometries to determine chemical reaction barriers, for the determination of molecular forces that are used in direct-dynamics trajectory calculations, and for sophisticated PES surface fitting methods that use both energy and gradient information. For MRCI wave functions, the effort required to compute the entire gradient vector, consisting of $3N_{atom}$ elements when using Cartesian molecular coordinates,

requires typically only a small fraction of the effort to compute the energy itself.[21] Further discussion and examples are provided in sections 2 and 3.

Although the rather general form and the simple structure of the CI wave function along with the hierarchical way of truncation to practical size are very appealing properties, the energy corresponding to CI wave function truncated according to excitation level will not scale properly with the size of the system; i.e., excitation-level truncated CI is not size-extensive[22,23] or, with other terminology, not size-consistent.[24] For a general definition of these two closely related terms, refer to ref 25. Size-extensivity is more related to many-body diagramtic formulations,[26,27] while size-consistency is a property of the method related to the separability into subsystems.[24,26] Since CI is usually not formulated with many-body tools, the scaling property of CI and its corrections are mostly discussed with respect to the separation limits. Therefore, in this review the term size-consistency will be used. The lack of size-consistency hampers the investigation of molecular systems of increasing size, since an ever-smaller fraction of the correlation energy is computed, and in the limit of infinite systems no correlation energy at all is computed. On the other hand, the effect is not negligible even for middle-size molecules, so a correction is desirable. Various methods to compute this correction will be discussed in detail in section 2.

The lack of size-consistency of the truncated CI wave function is easily understood by considering two noninteracting systems by the SR-CID method. While the wave function of the supersystem treated as one entity includes only double excitations, the product of the two SR-CID wave functions of the subsystems must include quadruple excitation terms. Consequently, the energy of the supersystem will not be size-consistent as it will not be equal to the sum of the energies of the two subsystems. In fact, it can be shown that the SR-CID correlation energy grows with the square root of the system size instead of proportionally to the system size.[26] This property holds for any excitation-level-based truncation and disappears only in the case of full-CI where the excitation level is naturally exhausted. Note that in the case of CID, it is the lack of certain quadruple excitations that causes the problem; this observation is important for understanding the correction schemes discussed as follows. Also note that by increasing the excitation level included in the CI wave function, the missing terms that cause size-consistency errors will correspond to higher excitations and this typically results in a decrease of magnitude of the error. For the same reason, the size-consistency error is expected to be smaller for multireference expansions than for the single-reference expansions.

2. DISCUSSION

2.1. CI Method

2.1.1. Basic Terms and Notations. The nonrelativistic, clamped nucleus, electronic Hamiltonian operator includes the electronic kinetic energy, the electron−nuclear attraction, and the electron−electron repulsion.

$$H = \sum_j \frac{-\hbar^2}{2m_e}\nabla_j^2 + \sum_{j,a}\frac{Z_e Z_a}{|\mathbf{r}_j - \mathbf{R}_a|} + \sum_{j>k}\frac{Z_e^2}{|\mathbf{r}_j - \mathbf{r}_k|} \tag{7}$$

Although the CI method may be viewed from either the first- or second-quantization perspective, essentially all modern formulations use the latter approach. The electronic Hamiltonian operator is then written either in terms of spin-orbital creation ($a_p^\dagger$) and annihilation (a_q) operators

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2}\sum_{p,q,r,s}(pq|rs) a_p^\dagger a_r^\dagger a_s a_q \tag{8}$$

or in terms of the spatial-orbital indices using the spin-adapted generators (E_{pq}) and generator products (e_{pqrs}) of the unitary group

$$H = \sum_{p,q} h_{pq} E_{pq} + \frac{1}{2}\sum_{p,q,r,s}(pq|rs) e_{pqrs} \tag{9}$$

using what is now accepted as standard notation conventions.[6] It is sometimes convenient to treat the two-electron repulsion integrals as the elements of an array $g_{pqrs} \equiv (pq|rs)$; both notations are used herein.

The N-electron expansion functions are chosen typically to be either primitive Slater determinants of spin-orbitals (the natural bras and kets of the occupation-number representation) or configuration state functions. CSFs are linear combinations of Slater determinants that have the same spatial-orbital occupations and that are eigenfunctions of the total spin operator S^2 and the spin projection along the z-axis S_z.

$$S^2|k;S,M\rangle = S(S+1)|k;S,M\rangle \tag{10}$$

$$S_z|k;S,M\rangle = M|k;S,M\rangle;\quad M = -S, -S+1, \ldots, +S \tag{11}$$

In these equations $|k;S,M\rangle$ is the CSF indexed by k and is characterized by the spin quantum number S and the spin projection eigenvalue $M = (^1/_2)(N_\alpha - N_\beta)$. N_α and N_β are the number of occupied α and β spin-orbitals, respectively, in each of the determinants in the wave function expansion. Primitive Slater determinants are typically eigenfunctions only of the S_z operator. Determinantal expansions are usually longer (because they span several S values) than CSF expansions (which typically are chosen to span only a single S value). On the other hand, the one- and two-electron coupling coefficients in the determinantal basis, $\langle j|a_p^\dagger a_q|k\rangle$ and $\langle j|a_p^\dagger a_r^\dagger a_s a_q|k\rangle$ (see eq 8), take on the values 0 and ±1 only and are therefore somewhat simpler to evaluate than the coupling coefficients in the CSF basis, which are usually written in terms of the spin-adapted generators, $\langle j|E_{pq}|k\rangle$ and $\langle j|e_{pqrs}|k\rangle$. Consequently, various CI methods are implemented with both types of formulations. Another important difference between determinantal and CSF expansions is that the state ordering may be different (e.g., the lowest eigenvalue in a CSF basis might correspond to a higher eigenvalue in a determinantal basis), and this association of the states may change across the PESs (i.e., due to allowed crossings of states with different S values). Although determinantal CI expansion spaces are usually chosen so that there is no spin contamination (breaking of spin-symmetry) in the converged wave functions, some wave function optimization procedures are prone to introduce artificial contamination at intermediate stages.[4] CSF expansions generally avoid these spin-contamination issues.

There are two general approaches to generate the CI expansion space: through excitations and through orbital occupation restrictions. In a SR expansion, for example, a particular expansion space might be generated through excitations from the reference detrminant $|\psi_0\rangle$ as

$$\{|l\rangle; l = 1 \ldots N_{\mathrm{dim}}\} = \{|\psi_0\rangle, |\psi_i^a\rangle, |\psi_{ij}^{ab}\rangle, |\psi_{ijk}^{abc}\rangle, \ldots\} \tag{12}$$

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

with

$$|\psi_i^a\rangle = a_a^\dagger a_i|\psi_0\rangle; \quad |\psi_{ij}^{ab}\rangle \equiv a_a^\dagger a_b^\dagger a_j a_a|\psi_0\rangle; \quad |\psi_{ijk}^{abc}\rangle \equiv a_a^\dagger a_b^\dagger a_c^\dagger a_k a_j a_i|\psi_0\rangle; \quad \ldots \quad (13)$$

The spin-orbital indices $i, j, k, \ldots$ range over the occupied orbitals in the reference determinant, and $a, b, c, \ldots$ range over the unoccupied (virtual) orbitals. The determinant $|l\rangle$ in eq 12 in the wave function expansion space is therefore associated with some combination of occupied and virtual orbital indices. The expansion function $|\psi_i^a\rangle$ is a "single excitation" term in which the occupied spin-orbital φ_i in the reference determinant is replaced with virtual orbital φ_a, $|\psi_{ij}^{ab}\rangle$ is a "double excitation" in which two occupied spin-orbitals φ_i and φ_j in the reference determinant are replaced with virtual orbitals φ_a and φ_b, and so on for the other expansion terms. In most applications, the excitations are restricted to preserve the S_z eigenvalue of the reference. A wave function expanded in this space can be written as

$$|\psi\rangle = c_o|\psi_0\rangle + \sum_{i,a} c_i^a|\psi_i^a\rangle + \sum_{i>j,a>b} c_{ij}^{ab}|\psi_{ij}^{ab}\rangle + \sum_{i>j>k,a>b>c} c_{ij}^{ab}|\psi_{ijk}^{abc}\rangle + \ldots \quad (14)$$

c_0, c_i^a, c_{ij}^{ab}, ... are the expansion coefficients. It is common for SR methods to scale the wave function so that $c_0 = \langle\psi_0|\psi\rangle = 1$; this is referred to as the intermediate normalization convention, in constrast with the full normalization convention in which the norm of the wave function is unity ($\langle\psi|\psi\rangle = 1$). The full-CI expansion occurs when all possible N-rank excitations are included. Truncations of the expansion based on the overall excitation level result in SR-CIS, SR-CISD, and SR-CISDT, etc.

The expansion space can also be defined in terms of spatial orbital indices with the operators E_{ai}, $e_{bj,ai}$, and so on, but in this case the expansion terms are typically linear combinations of determinants rather than individual determinants. Further, the relations $[E_{ai},S^2] = [E_{ai},S_z] = 0$, $[e_{bj,ai},S^2] = [e_{bj,ai},S_z] = 0$, and so on ensure that the expansion space generated with these operators maintains the S^2 and S_z eigenvalues of the reference $|\psi_0\rangle$.

The alternative approach to generating the expansion space consists of imposing occupation restrictions on groups of orbitals within the set of expansion determinants. All determinants that satisfy the constraints are enumerated and retained in the expansion. In the SR example above, the occupied spin-orbital subspace would consist of $n_{occ} = N_\alpha + N_\beta$ spin-orbitals and the virtual spin-orbital space would consist of the remaining $n_{virt} = 2n - n_{occ}$ spin-orbitals. The reference wave function is the determinant that has the n_{occ} orbital space occupied by the N electrons, and the n_{virt} orbital space has zero occupation. The single-excitation determinants are those that have the n_{occ} orbitals occupied by $N - 1$ electrons and the n_{virt} orbital space occupied by one electron. The double-excitation determinants are those with the n_{occ} orbitals occupied by $N - 2$ electrons and the n_{virt} space occupied by two electrons. In multireference approaches, the orbital space is typically divided into several subsets of orbitals, and only the determinants that satisfy the occupation limits on all the subsets are included in the expansion space. Graphical methods are often used to represent and enumerate the expansion space in this form. The graphical unitary group approach (GUGA) of Shavitt[6,28] is an example of this approach for CSF expansions. Various determinant-based graphical schemes are also commonly used[29] to specify expansion spaces. (See ref 6 for a more complete discussion of graphical methods.) Most MCSCF implementations, whether or not they are based on an underlying graphical representation, use orbital subspace occupation restrictions to specify the wave function expansion space.

The practical differences in these two approaches become apparent in MR expansions. In this case, a set of reference functions $\{|m;\text{ref}\rangle; m = 1\ldots N_{ref}\}$ is used to generate the expansion space. The set of excitation operators is applied to each reference function to generate the expansion space:

$$|\psi(m)_i^a\rangle \equiv a_a^\dagger a_a|m;\text{ref}\rangle; \quad |\psi(m)_{ij}^{ab}\rangle \equiv a_a^\dagger a_b^\dagger a_j a_i|m;\text{ref}\rangle; \quad \ldots \quad \text{for } m = 1\ldots N_{ref} \quad (15)$$

Unlike the single-reference situation, some of these expansion terms may be zero (e.g., a particular spin-orbital φ_i may be unoccupied in a particular reference function $|m;\text{ref}\rangle$), and such terms must be identified and eliminated. The wave function is written in analogy to eq 14 as

$$|\psi_{MRCI}\rangle = \sum_m c_m|m\rangle + \sum_m \sum_{i,a} c(m)_i^a|\psi(m)_i^a\rangle + \sum_m \sum_{i>j,a>b} c(m)_{ij}^{ab}|\psi(m)_{ij}^{ab}\rangle + \sum_m \sum_{i>j>,a>b>c} c(m)_{ijk}^{abc}|\psi(m)_{ijk}^{abc}\rangle + \ldots \quad (16)$$

In some cases, a particular excitation from one reference expansion function can be identical to a different excitation from some other reference expansion function; the inclusion of both terms results in a linear dependency in the expansion space. Consequently some method must be adopted to identify and eliminate these redundant expansion terms from eq 16. Redundancies do not occur in an occupation-based approach; a possible determinant either succeeds or fails to satisfy the occupation restrictions, and that alone determines its expansion index $|l\rangle$. Therefore the resulting expansion space generated in this manner does not suffer from linear dependence.

Equations for the wave function parameters (the expansion coefficients) can be obtained from the variational principle by minimizing the Rayleigh quotient eq 2. Equivalently, one can insert the wave function into the Schrödinger equation and project onto the space of excited determinants. For a double excitation, for example,

$$\langle\psi_{ij}^{ab}|H - E_0|\psi_0\rangle + \sum_{k,c} c_k^c\langle\psi_{ij}^{ab}|H - E_0|\psi_k^c\rangle + \sum_{\substack{k>l,\\ c>d}} c_{kl}^{cd}\langle\psi_{ij}^{ab}|H - E_0|\psi_{kl}^{cd}\rangle + \sum_{\substack{k>l>m,\\ c>d>e}} c_{klm}^{cde}\langle\psi_{ij}^{ab}|H - E_0|\psi_{klm}^{cde}\rangle + \sum_{\substack{k>l>m>n,\\ c>d>e>f}} c_{klmn}^{cdef}\langle\psi_{ij}^{ab}|H - E_0|\psi_{klmn}^{cdef}\rangle = c_{ij}^{ab}\Delta E \quad (17)$$

For the correlation energy one obtains (in the case of SCF orbitals)

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

$$\Delta E = \sum_{i>j,a>b} \langle \psi_0|H|\psi_{ij}^{ab}\rangle c_{ij}^{ab} \equiv \sum_{i>j} \varepsilon_{ij} \tag{18}$$

where the pair energy

$$\varepsilon_{ij} = \sum_{a>b} \langle \psi_0|H|\psi_{ij}^{ab}\rangle c_{ij}^{ab} \tag{19}$$

has been introduced. Equations 17 and 18 are specific examples of the transition energy expression in eq 6. The total energy is given by $E = \Delta E + E_0$, with E_0 being the energy corresponding to the reference determinant $|\psi_0\rangle$.

Inspection of eq 17 reveals that the double excitation coefficients depend on the single, double, triple, and quadruple coefficients only, but not on the higher rank coefficients. This is because the Hamiltonian includes, at most, two-electron excitation operators. If the wave function is truncated as in CISD, the terms including triple and quadruple excited coefficients (the last two terms on the left-hand side of eq 17) will not appear.

$$\langle \psi_{ij}^{ab}|H - E_0|\psi_0\rangle + \sum_{k,c} c_k^c \langle \psi_{ij}^{ab}|H - E_0|\psi_k^c\rangle + \sum_{\substack{k>l \\ c>d}} c_{kl}^{cd} \langle \psi_{ij}^{ab}|H - E_0|\psi_{kl}^{cd}\rangle = c_{ij}^{ab}\Delta E \tag{20}$$

However, these neglected terms are comparable in size with $c_{ij}^{ab}\Delta E$, and therefore this approximation is not fully justified. To see this, examine the term including the quadruple excitations in eq 17. First, the matrix element can be simplified by using the Slater−Condon rules,[30]

$$\langle \psi_{ij}^{ab}|H - E_0|\psi_{klmn}^{cdef}\rangle = \delta_{in}\delta_{jm}\delta_{ae}\delta_{bf}\langle \psi_0|H - E_0|\psi_{kl}^{cd}\rangle \tag{21}$$

Second, the coefficients can be approximated by the leading term of coupled-cluster theory

$$c_{ijkl}^{abcd} \approx c_{ij}^{ab} c_{kl}^{cd} \tag{22}$$

With these considerations, the last term on the left-hand side of eq 17 becomes

$$\sum_{k>l,c>d}^{\neq ij,ab} c_{ij}^{ab} c_{kl}^{cd} \langle \psi_0|H - E_0|\psi_{kl}^{cd}\rangle \tag{23}$$

The restriction on the summation (that none of k, l, c, d can coincide with i, j, k, l, denoted by $\neq ij,ab$), which comes from the fact that in the quadruple excitations the same electron cannot be excited twice, ensures that the so-called exclusion principle violating (EPV) terms[23] are excluded. Note also that this term is very similar to the term on the right-hand side of eq 17 if the expression for the correlation energy (eq 18) is inserted

$$c_{ij}^{ab}\Delta E = \sum_{k>l,c>d} c_{ij}^{ab} c_{kl}^{cd} \langle \psi_0|H - E_0|\psi_{kl}^{cd}\rangle \tag{24}$$

Upon comparing eqs 23 and 24, it is seen that the two terms differ only in the restricted summation indices in the former. If the restrictions were ignored, there would be exact equivalence in the two terms. This means that instead of simply ignoring higher excitations from the CI equations, by also neglecting the energy-dependent term, a more balanced approximation can be defined. The method using this approximation goes by different names;[31] one such name is CEPA(0),[32] and it is the lowest level of the set of coupled electron pair approximation (CEPA) methods. Due to its apparent and historical relation to CI, it is discussed further below. Note that in this way, certain quadruple-excitation contributions that are neglected in CISD, namely, those responsible for the size-consistency error, are included in the energy expression. However, eqs 23 and 24 do not cancel exactly due to the restricted summation. This means that in CEPA(0) the EPV terms are not properly accounted for. Further versions of CEPA[32] address this problem by compensating for the restricted summation in eq 23. (In the literature it is often stated that CEPA methods "include EPV terms," since those terms remaining after the cancellation resemble EPV terms. Note however, that inclusion of these terms means, on the contrary, that the EPV terms have been properly handled. Since this nomenclature is very misleading, it is avoided in this review, and the terms "proper handling of EPV terms" or "compensating for EPV terms" are used instead.) The review in ref 31 discusses the proper handling of EPV terms in more detail.

Using the above approach, the CISD equations may be corrected for the size-consistency error with various degrees of partial to full cancellation. These are often referred to as the CEPA type methods. Another possibility is to leave the CI equations unchanged and to introduce a correction to the final energy expression by considering the approximate change of the CI coefficients due to the correction. For example, using CEPA-(0)-type arguments, i.e., approximating the effect of higher excitations by $c_{ij}^{ab}\Delta E$, the change of the double-excitation coefficients can be approximated as

$$\Delta c_{ij}^{ab} = \frac{\Delta E c_{ij}^{ab}}{\langle \psi_{ij}^{ab}|H|\psi_{ij}^{ab}\rangle} \tag{25}$$

which results in the change of the correlation energy

$$\Delta\Delta E = \sum_{i>j,a>b} \frac{\langle \psi_{ij}^{ab}|H|\psi_0\rangle c_{ij}^{ab}}{\langle \psi_{ij}^{ab}|H|\psi_{ij}^{ab}\rangle}\Delta E \tag{26}$$

Recognizing that

$$-\frac{\left\langle \psi_{ij}^{ab}|H|\psi_0\right\rangle}{\left\langle \psi_{ij}^{ab}|H|\psi_{ij}^{ab}\right\rangle} \approx c_{ij}^{ab}$$

is the first-order approximation to the coefficients, the correction formula may be simplified as

$$\Delta\Delta E = \sum_{i>j,a>b} (c_{ij}^{ab})^2 \Delta E \tag{27}$$

Higher order correction can also be introduced by augmenting c_{ij}^{ab} by Δc_{ij}^{ab} or by introducing a correction beyond CEPA(0). Since these methods are applied after the CI procedure, they are commonly referred to as a posteriori corrections. A variety of such approaches will be discussed in section 2.1.3.

2.1.2. Solution of the CI Eigenvalue Equation. The eigenvalue equation in eq 3 can be solved using either "direct" or "iterative" numerical methods. Direct methods are those in which the entire matrix is constructed and the individual matrix elements are modified successively in some manner until the

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

eigenvectors $\mathbf{c}^k$ and the eigenvalues E_k are eventually computed. For a matrix of dimension $N_{\text{dim.}}$, a typical direct method requires $O(N^3_{\text{dim.}})$ computational effort and $O(N^2_{\text{dim.}})$ storage and computes some or all of the eigenpairs. Direct eigenvector methods are among the most studied of linear algebra problems. Reference 33 gives a complete review of methods up through 1998.

The term "iterative" is somewhat of a misnomer because, except for some trivial situations, all eigenvector solutions are iterative at some level; even a simple 2×2 matrix involves the computation of a square root, which itself requires an iterative numerical solution. An "iterative" eigenvalue method is deemed to be one in which the underlying iterative procedure drives the overall numerical solution. Iterative methods are often based on perturbation theory or gradient searches combined with variational methods, and they focus typically on the convergence of selected eigenpairs rather than the full set of solutions. An important class of iterative methods is based on subspace expansion and includes the Lanczos and the Davidson methods (see ref 33). These methods often do not require the computation or storage of individual Hamiltonian matrix elements but instead require only the result of matrix−vector products $\mathbf{w} = \mathbf{H}\mathbf{x}$ for arbitrary expansion vectors $\mathbf{x}$. The computational effort thereby depends on the number of nonzero elements in the matrix $\mathbf{H}$ (which is $\ll N^2_{\text{dim.}}$ for a sparse matrix), and the storage requirements depend on the number of expansion vectors $\mathbf{x}$. The desired eigenvectors are eventually represented as linear combinations of these expansion vectors.

CI methods are also classified[1,34,35] as either *conventional-CI* or *direct-CI*. A conventional-CI approach is one in which the matrix elements are explicitly constructed and stored. Either a direct or an iterative diagonalization method may be used within a conventional-CI approach. A direct-CI approach is one in which the matrix−vector products within an iterative method are computed in operator form "directly" from the underlying one- and two-electron integrals. This is, of course, an unfortunate choice of terminology, but it should be clear that a direct-CI approach is always associated with an iterative diagonalization method. To further confuse this issue, there are also *AO-direct* methods, sometimes also termed *double-direct* methods; these are direct-CI methods in which the required matrix−vector products are computed using the AO two-electron repulsion integrals which are themselves recomputed on-the-fly each time they are needed during the iterative procedure.

Due to the large dimensions of MRCI wave function expansions, the eigenpairs of eq 3 are computed almost exclusively using iterative approaches based on the Davidson subspace method.[36−38] In direct-CI formulations, the contributions to $\mathbf{w}$ are accumulated on the basis of convenient organizations of the underlying two-electron repulsion integrals and the associated coupling coefficients. In well-organized codes, these underlying operations typically involve dense matrix−vector and matrix−matrix operations.[39−42] This results in efficient computational kernels that achieve near-peak performance on modern hardware. The Davidson method is a subspace method, which means that some number of expansion vectors $\{\mathbf{x}^j;\ j = 1...m\}$ are collected together to form the columns of a matrix $\mathbf{X}$, and the eigenvectors $\mathbf{c}^k$ are represented as linear combinations of these basis functions

$$\mathbf{c}^k = \mathbf{X}\tilde{\mathbf{c}}^k \tag{28}$$

The optimal coefficients within this subspace are computed from the subspace eigenvalue equation

$$\tilde{\mathbf{H}}\tilde{\mathbf{c}}^k = \tilde{\mathbf{S}}\tilde{\mathbf{c}}^k\rho_k \tag{29}$$

with the matrices of dimension m defined as $\tilde{\mathbf{H}} = \mathbf{X}^{\mathrm{T}}\mathbf{H}\mathbf{X}$ and $\tilde{\mathbf{S}} = \mathbf{X}^{\mathrm{T}}\mathbf{X}$. The subspace eigenvalues ρ_k are called the *Ritz values*, and the subspace eigenvectors $\tilde{\mathbf{c}}^k$ are called the *Ritz vectors* (see, e.g., ref 33 for a complete discussion). In some implementations the subspace expansion vectors are chosen to be orthogonal, $\tilde{\mathbf{S}} = \mathbf{1}$, while in others this constraint is not imposed, for example, in order to reduce the I/O requirements each iteration. Typically, the subspace dimension m changes throughout the iterative procedure. The Ritz values ρ_k are upper bounds to the eigenvalues E_k of the $\mathbf{H}$ matrix, and they satisfy the interlace property corresponding to eqs 4 and 5 as the subspace dimension m changes. Subspace vectors may be added, either one at a time or in blocks,[43] until convergence is achieved. In many implementations, the dimension m may be reduced periodically in order to decrease the overall storage requirements. This subspace reduction involves a contraction of the expansion vectors $\mathbf{X}^{\text{new}} \leftarrow \mathbf{X}^{\text{old}}\mathbf{T}$ where the rectangular transformation matrix $\mathbf{T}$ depends on which eigenpairs are being converged at the time of the contraction. Subspace dimensions of $m \approx 10-20$ are typical for single-state calculations, with dimensions up to $m \lesssim 100$ for multiple-state calculations. For the largest of CI calculations, computational resources may limit the maximum subspace dimension to smaller values[44] $m \approx 2$ or 3. Thus the subspace matrix diagonalization in eq 29 is relatively trivial and typically uses a direct (i.e., $O(m^3)$ effort) diagonalization approach. Although the lowest root or lowest few roots are usually desired, it is also possible to converge selected interior roots using either root-homing or vector-following approaches. In a root-homing approach, the subspace eigenvalue, or eigenvalues, closest to some target value or range are selected for improvement in the next iteration. In a vector-following approach, the approximate vector, or vectors, with the largest overlaps to some set of predefined reference vectors $\{\mathbf{c}^{(0)k};\ k = 1,m'\}$ are selected.

A crucial step of the iterative procedure is the formation of the new expansion vectors. In the Lanczos subspace method,[33] these are defined from the gradient $\partial\rho/\partial\mathbf{c}$, or equivalently the residual vector $\mathbf{r}$,

$$\mathbf{X}^{\text{new}} = \frac{1}{2}\frac{\partial\rho}{\partial\mathbf{c}} = (\mathbf{H} - \rho)\mathbf{c} = \mathbf{r} \tag{30}$$

where the normalization $\mathbf{c}\cdot\mathbf{c} = 1$ of the current approximate vector $\mathbf{c}$ is assumed for notational brevity. The Lanczos method has several desirable features which follow from the facts that the resulting subspace matrix $\tilde{\mathbf{H}}$ is tridiagonal and that (in exact arithmetic) the subspace overlap matrix $\tilde{\mathbf{S}}$ is diagonal. However, in practice this expansion vector choice exhibits slow convergence, and due to roundoff errors the diagonal nature of $\tilde{\mathbf{S}}$ cannot be assumed to be maintained. The slow convergence is a particular problem in CI calculations because each iteration requires the expensive computation of a new $\mathbf{w} = \mathbf{H}\mathbf{x}$ vector.

To overcome these problems, Davidson[36] proposed instead the expression

$$\mathbf{X}^{\text{new}} = (\mathbf{H}^0 - \rho\mathbf{1})^{-1}\mathbf{r} \tag{31}$$

where $\mathbf{H}^0$ is chosen as some easily invertible matrix. In practice, $\mathbf{H}^0$ is taken typically to be either the diagonal elements of $\mathbf{H}$ or some easily computed approximation to these diagonal elements.

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

Expressions involving orbital energies (see ref 45 for an example of a typical implementation) or averaged expressions involving the h_{pp}, g_{ppqq}, and g_{pqpq} integrals (see refs 4, 46, and 47 for further discussions) are common choices. This eliminates the need to store and retrieve the $\mathbf{H}^0$ elements during the optimization procedure. With determinantal expansions, artificial spin-contamination issues may be avoided either by choosing an $\mathbf{H}^0$ for which $[\mathbf{H}^0,S^2] = 0$ or by invoking a subsquent explicit spin-projection step (e.g., see refs 4 and 46-48 for further discussion). The expression in eq 31 may be derived either from perturbation theory or from relaxation arguments, and the matrix $(\mathbf{H}^0 - \rho\mathbf{1})^{-1}$ may be regarded as a preconditioner of the gradient.

Because the expansion vectors are created in an iterative process, it is not necessary to store the exact elements, for example, from eq 31. Any small error introduced in one expansion vector can be compensated for in subsequent expansion vectors. Lossy data compression may be applied to the vectors in order to reduce storage requirements, reduce I/O bandwidth, or to reduce communication volume in a parallel environment, and if the differences that are introduced into the individual vectors are controlled, they do not affect the overall convergence rate. Perturbation theory estimates, combined with rigorous error bounds, are typically used in the formal analysis and in the data compression process.[49−52] Typically these techniques involve deleting small expansion vector elements entirely, approximating floating point values with reduced precision, and approximating blocks of elements using linear algebra approaches (e.g., incomplete Cholesky factorizations[53−55] or truncated singular value decomposition).

Equation 31 is almost certainly the most popular expression used in iterative Davidson procedures. In practice it is usually reliable, provided some care is taken to account for small denominators in the preconditioner. These can arise in situations where a diagonal element happens to be close to the current Ritz value, particularly in excited-state calculations. Equation 31 does achieve the goal of improving the convergence rate over the Lanczos method. Typically for CI calculations, about 10−20 iterations are required to achieve convergence for each converged eigenpair. However, in contrast to the Lanczos method, the subspace matrix $\tilde{\mathbf{H}}$ is generally dense, without an explicit orthogonalization step the overlap matrix $\tilde{\mathbf{S}}$ is no longer diagonal, and all of the subspace vectors $\mathbf{X}$ and their matrix−vector products $\mathbf{W}$ need to be stored. Nonetheless, the Davidson method with diagonal preconditioners is without a doubt among the most popular and reliable approaches used within modern CI calculations.

In certain situations, however, convergence of the Davidson procedure with a diagonal preconditioner can be slow, requiring hundreds of iterations, and regardless of the convergence rate, the individual iterations with large CI expansions are expensive. This leads to the exploration for improved preconditioners in eq 31 that converge to sufficient accuracy with fewer expensive matrix−vector products. Olsen et al.[56] pointed out that in the limit $\mathbf{H}^0 \rightarrow \mathbf{H}$ in eq 31, then $\mathbf{x}^{\text{new}} \rightarrow \mathbf{c}$, the current approximate vector, and the iterative procedure makes no progress toward convergence. They replace eq 31 with the modified equation

$$\mathbf{X}^{\text{new}} = (\mathbf{H}^0 - \rho\mathbf{1})^{-1}(\mathbf{r} + \varepsilon\mathbf{c}) \tag{32}$$

with the orthogonality constraint $\mathbf{x}^{\text{new}} \cdot \mathbf{c} = 0$. The iterative subspace method based on this correction vector is called the inverse-iteration generalized Davidson (IIGD) method.[56] In the limit $\mathbf{H}^0 \rightarrow \mathbf{H}$ in eq 32, the correction vector corresponds to a Rayleigh quotient inverse iteration, which not only converges but also converges cubically[33] (i.e., the error in a particular iteration is the cube of the error of the previous iteration). This observation has little practical value (because iterating the linear equation solution of eq 32 with the exact $\mathbf{H}$ is just as expensive as iterating the eigenvalue equation), but it does show that the formulation is at least consistent in this formal limit. Sleijpen et al.[57,58] and van Dam et al.[59] arrive at a similar correction vector definition in their generalized Jacobi−Davidson (GJD) approach through approximation of the Rayleigh quotient inverse iteration equations. They also propose several ways to compute $\mathbf{x}^{\text{new}}$ with one-step rather than two-step procedures.

Finally, Shepard et al.[60] proposed the subspace projected approximate matrix (SPAM) method. This method employs a sequence of one or more approximations to the $\mathbf{H}$ matrix (denoted $\mathbf{H}^{(1)}$, $\mathbf{H}^{(2)}$, and so forth). These approximations are combined with projection operators defined with the current expansion vectors, $\mathbf{P} = \mathbf{X}\mathbf{X}^{\text{T}}$ and $\mathbf{Q} = (\mathbf{1} - \mathbf{P})$, to arrive at a recursive procedure that may be implemented as a modification of the Davidson subspace method. Some of the features of this approach are that the effective preconditioner improves each iteration because of the changing projection operators, the correction vector $\mathbf{x}^{\text{new}}$ is orthogonal not only to the current approximate eigenvector $\mathbf{c}^k$ but to the entire subspace $\mathbf{X}$, and the approximate matrices need not be easily invertible. In the formal limit $\mathbf{H}^{(1)} \rightarrow \mathbf{H}$, the SPAM approach would, in principle, converge in a single iteration, showing that the procedure is formally consistent. The usefulness of the SPAM approach rests on the difference in the computational effort in computing matrix−vector products with the approximate matrices compared to the effort for exact matrix−vector products. These approximations may consist of the neglect of off-diagonal blocks of the exact matrix (e.g., the B_k approximation[1]), approximations to the two-electron repulsion integrals (e.g., incomplete Cholesky factorizations,[53−55] RI approximations,[61] or the neglect of classes of multicenter integrals in an AO-direct implementation), density-screening methods (in either the MO or the AO basis), or representations using reduced precision arithmetic (e.g., to exploit 32-bit floating point GPU hardware).

The expansion vectors $\mathbf{X}$ and the corresponding matrix−vector products $\mathbf{W}$ are required in several steps of the Davidson subspace method, including the computation of the subspace matrices $\tilde{\mathbf{H}} = \mathbf{W}^{\text{T}}\mathbf{X}$ and $\tilde{\mathbf{S}} = \mathbf{X}^{\text{T}}\mathbf{X}$, the computation of the residual vectors $\mathbf{r}^k = (\mathbf{W}\tilde{\mathbf{c}}^k - \rho^k\mathbf{X}\tilde{\mathbf{c}}^k)$ and the corresponding residual norms $|\mathbf{r}^k|$, the orthonormalization of the expansion vectors, and contractions of the subspace in the form $\mathbf{X}^{\text{new}} \leftarrow \mathbf{X}^{\text{old}}\mathbf{T}$ and $\mathbf{W}^{\text{new}} \leftarrow \mathbf{W}^{\text{old}}\mathbf{T}$. Blocked algorithms that are designed to minimize I/O to external storage or to minimize communications requirements in parallel implementations may be used for these steps.[62] Compared to the unblocked algorithms, the overhead can be reduced from $O(m^2)$ to only $O(m)$ with this approach. These efficient blocked algorithms are particularly important for larger subspace dimensions associated with, for example, multiple-state calculations. Reference 62 contains a general discussion of these blocked algorithms along with several other features of Davidson implementations that are useful for large CI expansions.

Because each iteration of the diagonalization procedure is so expensive, convergence is seldom continued until full machine precision is achieved. Instead, the iterative process is typically terminated when the eigenvalues are converged to about 10^{-6} to 10^{-8} E_h. The convergence is typically monitored in several ways,

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

including the changes in the Ritz values and/or vectors in each iteration along with estimates of the predicted energy change for the next iteration (e.g., $\Delta\rho^k \approx \mathbf{r}^k \cdot \mathbf{x}^{\text{new}}$). In addition, there are also rigorous bounds that apply to the eigenvalue problem.[33] The most useful include the residual norm bound $|\mathbf{r}^k| \geq |E_k - \rho_k|$, the gap bound $|\mathbf{r}^k|^2/\gamma_k \geq |E_k - \rho_k|$, and the spread bound $|\mathbf{r}^1|^2/\sigma \leq |E_1 - \rho_1|$. The gap γ_k is the difference $|\rho_k - E_{k'}|$ between the Ritz value ρ_k and the nearest exact eigenvalue $E_{k'}$ for $k' \neq k$, and the spread σ is the spectral range of the matrix, $E_{N^{\text{dim.}}} - E_1$. The first two bounds apply to both interior and exterior eigenpairs, while the last bound results in an upper bound to the lowest exterior eigenvalue (e.g., the ground state) in CI calculations. A detailed discussion of the practical use of these bounds is given by Zhou et al.[63,64] The gap bound is one of several expressions that shows that the error in the eigenvalue is second-order in the error in the wave function.[1,33] Although these rigorous bounds are indeed useful in CI calculations, it has proven difficult to extend them to other types of electronic structure methods.

2.1.3. Size-Consistency Corrections to CISD. As previously discussed, the most serious formal deficiency of MRCI is the lack of size-consistency; i.e., the energy of the system does not scale properly with the system size. This shortcoming of CI led to the development of many-body methods (see, e.g., ref 26 and references therein), in particular different perturbation theory ansaetze and coupled cluster methods (for reviews see, e.g., refs 65−70), but several correction schemes to the CI energy and wave function have been suggested over the years as well. It is advantageous to group these latter into two categories: one includes a posteriori correction of the energy, the second includes corrections to the CI equations.

2.1.3.1. A Posteriori or Davidson Corrections. The first a posteriori correction of the CI energy was suggested as early as 1962 along with molecular applications by Sinanoglu:[71]

$$E_{\text{SC}} = \sum_{i>j} \varepsilon_{ij} \sum_{k>l,a>b}^{\neq i,j} (c_{kl}^{ab})^2 \tag{33}$$

where ε_{ij} is the pair energy defined in eqs 18 and 19. The second summation is restricted to k, l indices not coinciding with i and/or j, ensuring proper treatment of the EPV terms. The above form is not expensive to evaluate; therefore, it is rather surprising that it has not been used more widely by the theoretical chemistry community. Instead, the popular formulas all represent the average of the above expression where pairs are not distinguished, and thus they can be written in terms of the correlation energy (ΔE) instead of the pair energies. The first such correction has been suggested by Davidson in a book chapter[72] (see also ref 73), and therefore these are often referred to as *Davidson corrections*.

Several versions of this correction are used in the literature. The simplest one is the original suggestion by Davidson[72] and by Langhoff and Davidson[73] (cf. eq 27)

$$E_{\text{DC}} = (1 - {c_0}^2)\Delta E \tag{34}$$

i.e., the correction is proportional to the correlation energy and to the square norm of the correlation part of the wave function (see also ref 74). Note that this formula can be obtained from E_{SC} by replacing the second, restricted summation by an unrestricted one

$$\sum_{k>l,a>b}^{\neq i,j} (c_{kl}^{ab})^2 \approx \sum_{k>l,a>b} (c_{kl}^{ab})^2 \tag{35}$$

This simply means that the EPV terms that should be excluded from the correction have been included. Proper normalization of the correction has been suggested by Luken[75]

$$E_{\text{RDC}} = \frac{1 - {c_0}^2}{{c_0}^2}\Delta E \tag{36}$$

Luken's derivation closely follows Sinanoglu's analysis[71], which leads to E_{SC} (see above). This form of the correction is often referred to as the *renormalized Davidson correction* (RDC). By applying size-consistency corrections to the coefficients as well, and by using these to calculate the energy correction, a slightly different formula was derived independently by Davidson and Silver[76] and by Siegbahn.[77]

$$E_{\text{DSS}} = \frac{(1 - {c_0}^2)}{2{c_0}^2 - 1}\Delta E \tag{37}$$

Despite the different forms, all of these formulas use the CEPA(0) approximation; i.e., EPV terms[23] are incorrectly included in the energy expressions. By comparing the CC and CI methods, Paldus et al.[78] give an excellent theoretical background of these approximations. They clearly state that these formulas will overestimate the effect of higher excitations, and in this respect E_{RDC} should be preferred over E_{DC}.

The next step taken to improve the correction was the correct treatment of EPV terms. Interestingly, however, the work of Sinanoglu[71] mentioned previously was not used as the starting point in this effort. Instead, averaged forms with the correlation energy replacing the pair energies were considered. The first useful approximation was given by Pople et al.[79]

$$E_{\text{PC}} = \frac{\sqrt{N^2 + 2N\tan^2(2\theta)} - N}{2(\sec(2\theta) - 1)}\Delta E \tag{38}$$

with $\cos(\theta) = c_0$, where N is the number of correlated electrons. To show that this formula accounts for EPV terms, Meissner[80] has rewritten it into the following simpler form (assuming $c_0 \sim 1$)

$$E'_{\text{PC}} = \left(1 - \frac{2}{N}\right)\frac{1 - {c_0}^2}{{c_0}^2}\Delta E \tag{39}$$

This formula can also be derived using an averaged CEPA approximation[81] by considering noninteracting, equivalent, electron pairs. The resulting correction vanishes for two electrons, $N = 2$, which is formally correct since CISD is exact for two-electron wave functions, and there are no imposed excitation-level limitations on the wave function expansion. Note that the form E'_{PC} had been used in multireference applications (see, e.g., ref 82) prior to either ref 80 or ref 81.

A more rigorous consideration of a helium-like noninteracting system led Duch and Diercksen[83] to the following formula

$$E_{\text{DDC}} = \frac{1 - {c_0}^2}{2((N-1)/(N-2)){c_0}^2 - 1}\Delta E \tag{40}$$

which is a slightly modified version of the Davidson−Silver−Siegbahn (E_{DSS}) correction.[76,77]

Further improvement can be introduced by considering different CEPA arguments and including electron pair interaction in an averaged manner. Meissner[80] suggested the use of the formula

$$E_{\text{MC}} = \left(\frac{(N-2)(N-3)}{N(N-1)}\right)\frac{(1 - {c_0}^2)}{{c_0}^2}\Delta E \tag{41}$$

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

Upon comparison with the second form of the Pople correction in eq 39 to this expression, the above formula yields a smaller correction due to the different leading factors and therefore compensates for the well-known overestimation of most Davidson correction schemes. Furthermore, E_{MC} vanishes for three-electron wave functions. As pointed out by Meissner,[80] this is consistent with the fact that there are no contributions from quadruply excited configurations in a three-electron system.

Duch and Diercksen[83] carefully compared the above-mentioned formulas and concluded that the Duch−Diercksen and Pople corrections clearly outperform the original Davidson correction and its renormalized variants. Shortly thereafter, Meissner[84] suggested a new correction which may be regarded as an approximation of the coupled-cluster (CC) energy using the CI coefficients. The calculation of this correction is computationally more demanding than the usual Davidson-type corrections since, in addition to the coefficients, some additional Hamiltonian-matrix elements are needed as well.[84] These matrix elements are not readily available in direct-CI programs, which explains why this correction, despite its accuracy, has not been widely implemented or used in applications.

For the sake of completeness, note that a size-consistency correction for the CIS method has been suggested by Zalesny et al.[85]

As previously noted, Davidson-type corrections are largely obsolete for single-reference CI; coupled-cluster methods can, and arguably should, be used instead. In contrast to single-reference approaches, the generalization of the above formulas for the MR case is, however, of significant interest since no MR-CC method yet satisfies all theoretical and practical requirements.[86−88,140,179−181]

The first applications of Davidson-type correction for the multireference case were due to Peyerimhoff and co-workers (see, e.g., refs 89−91) who used the Davidson correction E_{DC} to correct the MR-CISD energy with

$$E_{BPB} = (1 - \sum_{p \in P} {c_p}^2)(E_{MR\text{-}CISD} - E_0) \quad (42)$$

A more precise description of this formula is given in a book chapter by Buenker et al.[92] According to this definition, c_p are the coefficients in the reference space (P) and E_0 is the reference energy. Prime et al.[93] derived the renormalized (E_{RDC}) form of the correction using quasidegenerate perturbation theory arguments, while Simons[94] reached the same conclusion by comparing the MRCI and MR-CC wave functions. In later applications, Burton et al.[82,95,96] used this along with the simplified form of the Pople correction (E'_{PC}) in multireference situations. It has been concluded (see in particular ref 96) that the latter E'_{PC} correction gives more accurate results due to a lesser extent of overestimation. Note that in all these applications by the group Peyerimhoff et al.[82,91,92,96] the correction formulas were directed toward extrapolation of the truncated CI energy to the full-CI limit rather than to correct specifically for size-extensivity effects. The E_{BPB} formula was later used by Bauschlicher,[97] Schwenke and Truhlar,[98] and Ackermann and Hogreve.[99] Shavitt et al.[100] give a nice overview of all these efforts and compare the different corrections. Jankowski et al.[101] presented a generalization of Siegbahn's derivation[77] of E_{DSS} for the quasidegenerate case. The multireference version of the Meissner correction (E_{MC}) was already proposed in the original publication,[80] and it has been pointed out that it outperforms other variants.

The multireference correction formulas can be obtained from the single-reference ones by replacing c_0 and the reference energy by their multireference counterparts. Prime et al.[93] used quasidegenerate perturbation theory arguments to derive the renormalized form of the correction (E_{RDC}) by employing

$${c_0}^2 \equiv \sum_{p \in P} {c_p}^2 \quad (43)$$

where c_p are the coefficients of the reference functions in the MRCI wave functions. Meissner gives a somewhat simpler justification for the use of eq 43 in the appendix of ref 80. Blomberg and Siegbahn[102] start their derivation with the "logical choice" by defining c_0 as the overlap between the reference and final wave functions, which leads to

$${c_0}^2 \equiv \langle \mathrm{ref}|ci\rangle^2 = (\sum_{p \in P} c_p^{(0)} c_p)^2 \quad (44)$$

where $c_p^{(0)}$ are the coefficients of the reference functions in the normalized zero-order (MCSCF) wave function. They found that this choice results in corrections that are too large, in particular near transition states where the ordering of states in the reference and final spaces may differ. Therefore, they replace $c_p^{(0)}$ in eq 44 by the normalized coefficients of the reference functions in the final MRCI expansion, which results in eq 43 again. There are several other arguments for the use of eq 44: Simons[94] derives eq 44 by comparing the MRCI and MR-CC wave functions; Van Lenthe and co-workers[103,104] use eq 44, citing a private communication by Ahlrichs; Shepard[21] (in particular see page 417 of ref 21) discusses the same choices in the context of analytic energy derivatives; recently, Werner et al.[105] also give this formula explicitly and state that their earlier applications used this expression without specifying it. Additionally, they also suggest a third variant[105] of c_0 which, similarly to Blomberg and Siegbahn,[102] uses the overlap of the *rotated reference function* and the final wave function in a formula resembling eq 44. The rotated reference functions are those functions which have the largest overlap with the final MRCI states. This latter choice is preferable in the vicinity of conical intersections and avoided crossings where the reference space part of the wave function changes rapidly and the ordering of the reference and final energies are not the same, but it should give similar results otherwise.

In a recent review Khait et al.[106] present numerical comparisons of eqs 43 and 44 and find that eq 43 gives slightly better results. (Note that Khait et al.[106] refer to eq 44 as the "original formula" and to eq 43, which probably has been used more often, as "modified", which might cause some confusion.) To the contrary, Werner et al.[105] found that eq 44 gives somewhat better results for the barrier heights of the F + H_2 reaction, although they also suspect some error-compensation or error-cancellation effects. Considering all this information, it appears that, both theoretically (see ref 80) and numerically (see ref 106), the choice eq 43 is preferred, although in practice the difference does not seem to be critical.

In addition to c_0, there are also several possibilities in the choice of E_0. Early papers[80,91,103,104] all used the energy corresponding to the reference space (usually the MCSCF energy); even the latest investigation by Khait et al.[106] uses the same choice. The use of the expectation value of the Hamiltonian in the reference space using the MRCI instead of the MCSCF coefficients has been discussed in ref 21. This choice has received

little attention in the literature until recently when Szalay[107] used this in a MR-AQCC variant (discussed below) to address some discontinuity issues.

A more involved correction using MR perturbation theory, introduced by Duch and Diercksen,[83] results in a much smaller size-consistency error than direct application of the single-reference corrections to the multireference case. Meissner and co-workers constructed a multireference correction by calculating higher excitation coefficients approximately.[108−110] Their method can be viewed as the noniterative version of the RMR-CCSD (reduced multireference) method of Li and Paldus.[111,112] Only a limited number of tests are available for this method (see, e.g., Meissner et al.[113]), and unfortunately no comparsions are published with the usual methods. Other multireference methods have also served as a starting point for Davidson-type corrections. Meller et al.[114] used their MR-$(SC)^2$-MRCI method (see below) in the derivation. Hubać et al.[115,116] obtained a correction by Brillouin−Wigner perturbation theory[117] which was found to outperform both the E_{DC} and E_{RDC} corrections. Further details on the theoretical comparison of the a priori corrections can be found in refs 31, 84, and 106.

In light of the above discussion, it is worth reviewing which form of the Davidson corrections are used by different popular program systems. COLUMBUS[118,119] uses eq 43 along with E_0 being the MCSCF energy and calculates the multireference versions of E_{DC}, E_{RDC}, E_{SDS}, and E'_{PC}. The correction in MOLPRO[120] is restricted to E_{RDC} but offers all three choices for c_0 and two possibilities for E_0 (MCSCF energy, or the expectation value with the *rotated* functions). GAMESS[121,122] also calculates E_{RDC} using eq 43. MOLCAS[123] uses the renormalized Davidson correction (E_{RDC}) and a slightly modifed form of the modified Pople correction (E'_{PC}) called there the "ACPF correction".

A full numerical evaluation of the different corrections is out of the scope of this review. To give the reader an overview of the performace of different methods, in Table 1 the results of calculations on the symmetric dissociation of water are compiled. The most important observations are that all corrections significantly improve upon MRCI, and corrections using the CEPA(0) approximation (E_{DC}, E_{RDC}, E_{DSS}) overshoot by giving energies well below the exact (full-CI) limit. The best NPE (nonparallelity error) can be found for the Pople (E_{PC}) and Meissner corrections (E_{MC}). For a more detailed analysis see refs 31, 80, and 83.

From the survey of the literature it seems that the majority of applications use either the original (E_{DC}) or the renormalized (E_{RDC}) correction. From the already mentioned analyses by Meissner,[80] Duch and Diercksen,[83] and recently by Szalay,[31] it is clear however that either the Meissner correction (E_{MC})[80] or the Pople correction (E_{PC})[79] are preferable due their simplicity and to the correct treatment of EPV terms. Consequently, these are the recommended correction expressions.

Note that MR-CISD calculations with a Davidson-type correction are often labeled as MR-CISD+Q, indicating that some quadruple excitation contributions have been included. Although this is a convenient shorthand, the problem with this notation is that the exact correction expression often is not specified, and consquently insufficient information is availabe to reproduce the results. In most cases it is the original Davidson correction (E_{DC}) which is used, but not always, and caution must be taken when using these results.

Brown and Truhlar[126] recognized that the external correlation energy of an MR-CISD wave function relative to an MCSCF reference ($E^{ci} - E_0$) is approximately a constant fraction of the exact external correlation energy. They proposed the scaled external correlation (SEC) energy

$$E^{sec} = E_0 + \frac{E^{ci} - E_0}{F} \tag{45}$$

as a semiemprical correction to the MR-CISD energy. The parameter F is taken to be a constant over the entire PES, and its value is adjusted to fit known bond dissociation energies, barrier heights, excitation energies, or other experimental data. This expression has the same form as the Davidson-correction expressions discussed in this section. For example, eq 34 may used to write the corrected energy as

$$E^{total} = E_0 + (2 - {c_0}^2)(E^{ci} - E_0) \tag{46}$$

and similar expressions apply to the other correction expressions. The calculation of analytic energy gradients for the nonvariational Davidson-corrected energy expressions is complicated by the need to compute the ${c_0}^2$ derivatives (i.e., from the CI wave function response equation). As discussed in section 2.5, this is unnecessary for the variational MCSCF and MRCI energy gradients. However, due to the assumption that F is independent of geometry, the SEC energy gradient can be computed directly from the underlying MCSCF and MRCI energy gradients, or alternatively from the appropriately weighted reduced density

Table 1. Performance[a] of Various Multireference Methods for H_2O

	distance[b]						
method	1 R_e	1.5 R_e	2 R_e	2.5 R_e	3 R_e	$\overline{\Delta E}$ [c]	NPE [d]
MR-CISD[160]	4.96	4.72	3.72	3.14	3.01	3.91	1.95
+ E_{DC}	−1.21	−1.14	−0.70	−0.60	−0.58	−0.85	0.63
+ E_{RDC}	−1.47	−1.37	−0.86	−0.73	−0.68	−1.02	0.79
+ E_{DSS}	−1.73	−1.63	−1.02	−0.84	−0.80	−1.20	0.93
+ E_{PC} (N = 8)	0.05	0.06	0.22	0.21	0.21	0.15	0.17
+ E_{PC} (N = 10)	−0.31	−0.28	−0.04	−0.01	0.01	−0.13	0.32
+ E'_{PC}(N = 8)	0.14	0.16	0.28	0.23	0.23	0.21	0.14
+ E'_{PC}(N = 10)	−0.17	−0.15	0.05	0.05	0.06	−0.03	0.23
+ E_{MC}(N = 8)	1.51	1.44	1.26	1.06	1.02	1.26	0.49
+ E_{DC}(N = 10)	0.96	0.93	0.86	0.73	0.70	0.84	0.26
MR-CEPA(0)[160]	−1.80	−2.00	−2.37	−0.92	−0.82	−1.58	1.55
MRCEPA[160]	−0.79	−0.57	−0.54	−0.62	−0.64	−0.63	0.25
MR-ACPF(N = 8)[160]	0.05	0.07	0.24	0.21	0.20	0.15	0.19
MR-ACPF (N = 10)	−0.29	−0.26	−0.01	−0.02	−0.01	−0.12	0.28
MR-ACPF-mc[160]	0.22	0.43	0.34	0.20	0.18	0.27	0.25
MR-ACPF-2 (N = 8)[162]	0.71	0.82	0.71	0.53	0.47	0.65	0.35
MR-AQCC(N = 8)[160]	1.52	1.47	1.28	1.07	1.03	1.27	0.49
MR-AQCC (N = 10)	0.92	0.91	0.87	0.72	0.70	0.82	0.22
MR-AQCC-mc[160]	1.56	1.69	1.45	1.23	1.18	1.42	0.51
MR-$(SC)^2$CI[124]	2.11	2.09	1.83	1.77	1.83	1.93	0.34
CD-MRCISD[177]	1.03	0.91	0.80	0.71	0.64	0.82	0.39

[a] Table entries are the errors ΔE in mE_h with respect to the all-electron full CI of Olsen et al.[125] with the cc-pVDZ basis. A 4^4 CASSCF reference was used for all calculations. For some corrections, the indicated number of electrons was used in the correction formula. [b] Symmetric dissociation relative to $R_{e(OH)}$ = 1.843 45 bohr, ∠HOH = 110.6°. [c] Mean error. [d] Nonparallelity error (NPE = max(ΔE) − min(ΔE)).

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

matrices, providing an efficient algorithm for the SEC energy gradient.

2.1.3.2. CEPA-Type Multireference Methods. In this section those methods which include the extensivity correction in the equations rather than as an energy correction are summarized. In these methods the expansion coefficients are corrected along with the energy. A feature of these methods is that even though the expansion coefficients are avialable, the actual corresponding wave function often cannot be defined in a straightforward manner. In this sense these all follow the CEPA (coupled electron pair approximation) scheme. The original CEPA idea originated from Kelly[127,128] in the 1960s and was revived[129−135] in the 1970s after coupled-cluster theory (originally called CP-MET, coupled pair many electron theory[136]) provided a solid theoretical foundation. The development of CEPA methods has been largely motivated by the need for accurate, large-scale calculations, logically building upon CI technology. These methods include the variational versions CEPA-var by Pulay[137] and the CPF (coupled pair functional) method by Ahlrichs et al.,[138] which are of particular importance. Single-reference CEPA methods appear to be obsolete with the development of coupled-cluster methods,[65,66] since the latter are theoretically more rigorous yet computationally comparable in effort, although this view has been challenged recently.[139] However, in the multireference domain, CEPA-type methods have never lost importance due to the lack of widely accepted MR variant of CC theory despite recent effort (see, e.g., refs 86−88,140, and 178−180). This section focuses on the MR variants of the CEPA-type methods—for details of the single-reference CEPA methods, the interested reader is referred to the excellent review by Koch and Kutzelnigg.[141]

Surprisingly a large number of different variants of multireference CEPA-type methods has been proposed and used in the literature. Readers of the resulting papers, and the potential users of these methods, have a hard time choosing the best method to use in chemical applications. The methods are closely related, although the connections might not be obvious by just reading the derivations in the respective papers. These relations remain hidden, often even from the authors (and perhaps for reviewers); for this reason, some of the methods are completely equivalent, appearing under different names and from different derivations. A more formal comparison of all these methods can be found elsewhere.[31,142,143] The present review discusses these methods briefly in chronological order. The comparisons will concentrate on the following important properties of the methods: (i) treatment of EPV effects, (ii) redundancy contributions in the equations, (iii) whether the calculation of the energy gradient is possible via an energy functional, and (iv) applicability to excited states.

Although the motivation for introducing these methods was to account for the size-consistency error of MR-CISD, most of the methods given here are not size-consistent in the general sense. Some of them fulfill other important theoretical requirements instead, such as the proper description of the limiting case of noninteracting pairs or the proper behavior for certain numbers of electrons. Although not rigorously size-consistent, all the methods seek the goal that any remaining error is small and does not bias the application to molecular systems (see, e.g., refs 142 and 143).

Perhaps the first suggestion for a CEPA-type approach in the multireference domain is from Prime et al.[93] who made use of the cancellation of quadruple excitation contributions to quasidegenerate perturbation theory. By accounting also for the EPV terms, they developed a multireference version of the linearized CP-MET.[136] No numerical applications were given in ref 93, and no implementation results have been reported since.

The first practical multireference variant was proposed by Bartlett and co-workers[144,145] in the form of their multireference linearized coupled-cluster method (MR-LCCM). The derivation is based on linearization of the MR-CC equations as proposed by Paldus;[146] i.e., the method uses the CEPA(0) approximation. To avoid problems arising from the noncommutative nature of the excitation operators, the orthogonal complement of the reference space was excluded from the MR-LCCM wave function. On the basis of this feature, the authors of ref 144 point out the improved convergence behavior of the resulting equations. Possible inclusion of the orthogonal complement in the MR-LCCM method was considered in ref 145.

Three new variants then appeared in 1988−1989: the method called MR-CEPA(0) by Gdanitz and Ahlrichs[81] was a byproduct of MR-ACPF, the unitary CEPA (UCEPA) of Hoffmann and Simons[147] was based on their multireference unitary CC ansatz,[148] and the variational perturbation theory (VPT) method of Cave and Davidson[149] used perturbation theory arguments. In all three, the basic assumption is the CEPA(0) approximation, and redundancy is not considered. It was determined (see, e.g., ref 142) that these methods are indeed completely equivalent, and they differ from the MR-LCCM method in that they include the orthogonal complement of the reference space in the wave function. This difference does not really affect performance in most cases, but they are clearly preferable for situations where the reference wave function changes rapidly (e.g., avoided crossings, cusps, and crossing seams). The analytic gradient can easily be defined for all these methods,[142] and the calculation of excited states is possible via the diagonal shift formulation, but this option was not considered in the original papers (except for MR-CEPA(0) which is a special case of MR-ACPF[81]). In the following discussion, these three methods are all denoted MR-CEPA(0).

To address the convergence issue in cases of quasidegeneracy in the reference space, particularly when exclusion of the orthogonal complement is not acceptable, Cave and Davidson[150] proposed a variant called the quasidegenerate variational perturbation theory (QDVPT) method, which was based on an effective Hamiltonian. It has been shown[142] that it also uses the CEPA(0) approximation, and it is not expected to give substantially different results than MR-CEPA(0) (or VPT) except in the targeted quasidegenerate situations. Numerical results also support these expectations.[150] Note, however, that the CEPA(0) approximation is not valid if quasidegeneracy is present in the reference space, and therefore the use of either variant is not fully justified. Another drawback of the QDVPT method is that the analytic gradient cannot be easily formulated due to the lack of an energy functional. By construction, QDVPT can be applied to excited states.

An important development was proposed by Gdanitz and Ahlrichs[81] in the same year. On the basis of the CPF method,[138] they proposed an averaged formula in their multireference averaged coupled pair functional (MR-ACPF) method to account for EPV terms for the first time. Furthermore, the formulation was pioneering since the equations were formulated as the derivative of an energy functional, allowing immediate recognition of its relation to CI methods and providing a path for the calculation of analytic gradients similar to MRCI. The approximation used for the EPV terms appeared to be equivalent

to the modified form of Pople correction[79] eq 39 and can be explained by dividing the system into noninteracting electron pairs. Although the method is not strictly size-consistent, this formulation ensures that (i) the method is exact for noninteracting electron pairs and (ii) it is size-consistent for identical subsystems using a single function in the reference space. This method is among the most popular variants, and it has been implemented within several popular program packages.

Departing briefly from the chronological discussion, note that a variant of QDVPT (which is a MR-CEPA(0) method) using an ACPF-style correction for the EPV terms was subsequently proposed by Murray et al.[151] under the name QDVPT-APC (QDVPT with averaged pair correction). Since quasidegeneracy in the reference space is much less of a problem for MR-ACPF than for MR-CEPA(0), this method offers no apparent advantages over MR-ACPF. On the contrary, the formulation of energy derivatives is hindered by the lack of an energy functional. This method has been only rarely applied in chemical applications.

Another way to improve the MR-CEPA(0) method is to consider redundancy effects. The redundancy problem can easily be understood by realizing that configurations (or determinants) in the higher excitation space may be reached from more than one reference function. In contrast to the single-reference case, this means that correction for some higher excitations will be counted more than once. This redundancy must be considered in order to avoid overestimation of the correction. The first suggestion came from Ruttink et al.[103] under the name MRCEPA, who defined excitation classes which were characterized by the holes in the inactive orbitals and by particles on the virtual ones. These classes were considered separately when calculating the size-consistency correction. On the other hand, the EPV effects were neglected in MRCEPA; thus it remains a CEPA(0)-type method. The coefficients of the reference functions are relaxed in this procedure; i.e., the orthogonal component of the reference function was considered. An analytic gradient expression is not available for this method, but excited states can be described.

A more rigorous adaptation of the single-reference CEPA(n) series to the multireference problem was presented by Fulde and Stoll[152] in 1992. These methods were designated by the acronym MR-CEPA-n, with n = 0, 1, 2. In all variants the reference function is a prior MCSCF function that is not relaxed subsequently. The first version MR-CEPA-0 is equivalent to MR-LCCM as acknowledged by Fulde and Stoll.[152] MR-CEPA-1 and MR-CEPA-2 correspond to the single-reference CEPA(1) and CEPA(2), respectively. Redundancy effects are also considered, but only through the energy expression rather than the wave function equations. Fulde and Stoll[152] do not prove size-consistency explicitly, but they argue that the derivation through cumulants assures this property. Implementation into CI programs would be easy in a diagonal shift form, and no substantial additional computational effort compared to CI would be needed. There is no report on the implementation of any variants. Calculation of the analytic gradient would, however, be difficult since a functional of the energy cannot be associated with the equations. Furthermore, the method lacks invariance with respect to transformation of occupied orbitals, as do the single-reference counterparts. Treatment of excited states might be difficult and was not discussed in the original publication.[152]

In 1993, the method known as MC-CEPA (multiconfigurational reference CEPA) was introduced by Fink and Staemmler.[153] The equations defining the method reduce to Kelly's CEPA formula[127,128] in the single-reference case. Nonorthogonality of functions produced by products of operators are considered by the norms. Redundancies are not considered within this approximation. Implementation in the form of diagonal shift is easy; the additional computational effort is due to the calculation of the norms. Since no averaging of the pair energies is invoked, the method might be superior to MR-ACPF in situations in which the averaged pair approximation is not appropriate. No functional can be defined for the method, so an analytic gradient formulation is challenging. MC-CEPA has been implemented within the internally contracted framework (see section 2.1.5.1) and uses the PNO approximation[15,16] which allow applications for rather large molecular systems (see section 3.1).

Starting in 1989, another series of methods were developed by Tanaka and co-workers[154−157] under the name multireference coupled pair approximation (MRCPA). These methods use an effective Hamiltonian formalism similar to the QDVPT method. There are two levels of approximations: (i) MRCPA(2) (formerly known as MRCPA(0) [154,155]) uses a CEPA(0) basis, and, as such, it is equivalent to QDVPT [157] *ii*) MRCPA(4)[156] (which is a slight modification of the variant formerly known as MRCPA(2)[154,155]) and considers redundancies, but the EPV terms are not handled (see eq 68 in ref 157). The method simplifies to a CEPA(0) in the limit of a single-reference function, and therefore it is no surprise that MRCPA(4) overestimates the effect of higher excitations considerably (for more details see the comparisons in ref 31). The method is size-consistent for noninteracting electron pairs,[155] it can be applied to excited states, but analytic gradient calculations are not available.

Despite the successes of the MR-ACPF method, it was found to overestimate the effect of the higher excitations.[158] This led Szalay and Bartlett[158] in 1993 to suggest a modified version of the method; the multireference averaged quadratic CC (MR-AQCC) method[142,158] should be viewed as the CEPA version of the Meissner correction[80] just as MR-ACPF is related to the Pople correction. The functional form that was used can be justified by distributing the correlation energy among all possible electron pairs. This is in contrast to MR-ACPF where the correlation energy is distributed among the noninteracting electron pairs. Note that MR-AQCC retains all the attractive features of MR-ACPF, including an energy functional, which leads to an analytic energy derivative formulation,[81] and applicability to excited states. To enable the calculation of transition moments, a linear response version of the method (MR-AQCC-LRT) is available.[159] This means simply that the correction uses the ground-state instead of the excited-state correlation energy. There were some attempts to also include redundancy effects in the MR-AQCC-mc variant,[160] where the correction is done in the spirit of the MRCEPA method.[103] Note that ref 160 also suggests a procedure to include redundancy effects in MR-ACPF. In test calculations and in application to Be_2, Füsti-Molnar and Szalay[160,161] found that these methods performed excellently. MR-AQCC-mc received little attention (see however CD-MRCISD in subsequent text), most probably due to complicated structure of the method and to the lack of analytical gradients. A recent version of the MR-AQCC method by Szalay[107] solves some discontinuity issues by using a modified reference energy.

In 2001 Gdanitz[162] suggested a modified version of the MR-ACPF method, called MR-ACPF-2, which is essentially a combination of the original MR-ACPF and the MR-AQCC parametrization. While for the double excitation space the original MR-ACPF parametrizations[81] were used, the limiting value of

Table 2. Properties and Implementation of Different Multireference CEPA-Type Methods

method	compensation for EPV terms	handling of redundancy	reference space	availability of gradients	implementation
MR-LCCM[144,145]	no (CEPA(0))	no	unrelaxed	yes	COLUMBUS[118,119]
MR-CEPA(0)[81,147,149]	no (CEPA(0))	no	relaxed	yes	COLUMBUS[118,119]
QDVPT[150]	no (CEPA(0))	no	eff Hamiltonian	no	MELDF[282,283]
MRCEPA[103]	no (CEPA(0))	averaged	relaxed	no	GAMESS UK[121,122,182]
MRCPA(4)[156]	no (CEPA(0))	yes	eff Hamiltonian	no	local (ALCHEMY[183])
SS-MRCEPA(0)[168]	no (CEPA(0))	yes	eff Hamiltonian	no	MRCC[179,184]
MC-CEPA[153]	Kelly's CEPA	no	unrelaxed	no	Bochum code[153,185−187]
MR-ACPF[81]	averaged CEPA(2)	no	relaxed	yes	COLUMBUS,[118,119] MOLPRO[120]
QDVPT-APC[151]	averaged CEPA(2)	no	relaxed	no	MELDF[282,283]
MR-AQCC[142,158]	averaged CEPA(1)	no	relaxed	yes	COLUMBUS,[118,119] MOLPRO[120]
MR-AQCC-mc[160]	averaged CEPA(1)	averaged	relaxed	no	local[160] (COLUMBUS[118])
MR-$(SC)^2$-CI[163]	exact CEPA	exact	relaxed	no	local[163]
CD-MRSDCI[177]	averaged CEPA(1)	averaged	relaxed	no	local[177]
MR-CEPA1[176]	CEPA(1)	averaged	relaxed	no	GAMESS UK[182]
SS-MRCEPA(2)[169]	CEPA(2)	exact	eff. Hamiltonian	no	local[169]
SS-MRCEPA(I)[167]	exact CEPA	exact	eff Hamiltonian	no	local[167]

the MR-AQCC parametrization[158] was applied for the single excitation block of the Hamiltonian. In this way the notorious overestimation of higher excitations within MR-ACPF was addressed, and results similar to MR-AQCC were achieved.[162]

In 1994 Malrieu et al.[163] derived the method called multireference size-consistent self-consistent CI (MR-$(SC)^2$-CI) which is based on a single-reference variant[164] and can be viewed as an exact CEPA. It accounts for both EPV and redundancy terms exactly, and therefore it is rigorously size-consistent.[114] The method can be implemented in a diagonal shift (or dressing) form as first suggested by Heully and Malrieu.[165] Due to the storage requirement of a vector of length n_{ref} times the number of Q space functions, an extra cost that scales with the fourth power of the number of orbitals,[114] the method is much more expensive than other variants discussed previously. However, this additional cost is smaller than that of the underlying MRCI calculations. The importance of this method cannot be overstated—since it is the most rigorous among CEPA-type methods, it is often used as benchmark. It has been shown, however, that it is not significantly more accurate than MR-AQCC or MR-ACPF (see refs 142 and 166). Similarly, as in the single-reference case, the exact CEPA is not significantly better than the approximate variants.[141] Since no functional is associated with the MR-$(SC)^2$-CI method, no analytic gradient is available. There was an attempt to construct a functional form under the name MR-FCPF (multireference full coupled pair functional) by Meller et al.,[166] but it was necessary to introduce a geometry-independent diagonal shift which is formally problematic.

Most recently, Mukherjee and co-workers have proposed another family of multireference CEPA methods[167] termed SS-MRCEPA (state-specific multireference CEPA). These methods have been derived from the SS-MRCC approximation[168−172] (also known as Mk-MRCC) which is a realization of the Hilbert-space-type multireferenc CC approach.[173] As such, it is based on an effective Hamiltonian, but state-specificity has been obtained by explicitly including the reference space coefficients into the amplitude equations. For more detail, see ref 172. Four basic versions have been suggested. SS-MRCEPA(0) uses a CEPA(0) approximation (all terms are linearized).[168,179] SS-MRCEPA(2) uses approximations in the spirit of CEPA(2).[169] SS-MRCEPA-(D) includes only the diagonal terms of the dressed operators[174] and therefore is not orbital-invariant. Finally, SS-MRCEPA(I) is an "exact CEPA" in the sense that it considers EPV terms correctly.[167] Redundancy is inherently considered in these approximations since this is an essential ingredient of the parent SS-MRCC method.[172] Detailed comparison of these variants can be found in ref 175.

Recently, Ruttink et al.[176] and Ben Amor et al.[177] revisited the problem of accounting for both EPV and redundancy terms in CEPA-type approaches. Their new methods are a reconsideration of Ruttink's idea used in their MRCEPA method[103] and its extension by Füsti-Molnár and Szalay in the MR-AQCC-mc method.[160] Ruttink et al.[176] uses CEPA(1) arguments to extend MR-CEPA, and therefore the method is called MR-CEPA1.[176] The method of Ben Amor et al.[177] is called class-dressed (CD) MR-CISD and uses CEPA corrections depending on the excitation class. The new methods show some improvement over MRCEPA and MR-AQCC-mc, but they also share their feature that the analytic energy gradient cannot be formulated easily.

Table 2 summarizes the facts of this subsection by listing most CEPA-type methods and their properties. The theoretically most advanced methods are MR-$(SC)^2$-CI by Malrieu et al.[163] and SS-MRCEPA(I) by Mukherjee et al.[171] since these are exact CEPA methods. In addition, MR-AQCC-mc,[160] MR-CEPA(2),[169] MR-CEPA1,[176] and CD-MR-CISD[177] are methods which account for both EPV and redundancy effects. Still, in actual calculations, MR-AQCC[142,158] and MR-ACPF[81] perform better, competing with MRCC formulations[178−180] unless higher excitations are also considered in the latter.[180,181] From a pragmatic point of view, MR-AQCC and MR-ACPF have the advantages that analytic gradients are available,[118] that they can be readily used to calculate properties, and that they can be applied to excited states. CEPA methods should be preferred over the Davidson-type correction since the correction is introduced in the equations. However, this property can introduce intruder-state problems, in particular if proper orbitals cannot be constructed.

Table 1 also includes the results of some CEPA-type methods on the water example discussed previously with regard to the

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

Davidson-type corrections. Again, CEPA(0)-type methods overestimate the correlation effect and result in energies that are too low. Very similar NPE values can be observed for all other methods. There are several detailed comparsions of these methods in the literature, including refs 31, 142, 162, 176, 177, 179, and 180. Note that in the last two papers MR-AQCC and MR-ACPF are also compared to various MR-CC methods.

2.1.4. Inclusion of Connected Triple, Quadruple, and Higher Excitations. The size-consistency corrections discussed above address primarily disconnected higher excitation effects, those which are needed to restore size-consistency. Connected higher excitation effects, on the other hand, represent "real" higher excitations which appear also in coupled-cluster treatments. This has the apparent consequence that these methods cannot easily be classified as either CI or CC approaches. Since this review focuses on MRCI methods, the present discussion is limited to those which share some similarity to CI methods and can therefore be implemented within CI codes.

Using the formalism of dressed Hamiltonians developed for the $(SC)^2$-CI method,[164] Malrieu et al.[188,189] introduced higher excitations into a single-reference CI treatment. Meissner[190,191] used the coupled-cluster equations to derive corrections for connected triple excitations. Nooijen and Le Roy[192] also included some triple corrections in the single-reference pXCISD approach along with the size-consistency corrections (see above). Sherrill and Schaefer[193] included higher excitation effects by partitioning the (natural) orbital space according to the importance of the orbitals. Sychrovsky and Čarsky[194] used the B_k approximation[1] in the triple and quadruple excited space with respect to SR-CISD.

A very interesting novel approach to including higher excitations has been developed by Bytautas and Ruedenberg[195−198] who discovered that certain linear relationships exist between the incremental correlation energy contributions arising from different excitation levels when these increments, in turn, are considered as functions of increasing numbers of virtual natural orbitals, added in order of importance (i.e., occupations). As a result, the full correlation effects of quadruples, quintuples, and sextuples, etc., can be obtained, using a limited number of orbitals, by linear extrapolations from the doubles and triples contributions. Notably the method, called correlation energy extrapolation by intrinsic scaling (CEEIS), is effective for systematically approaching the full correlation energy of SRCI[195] as well as MRCI[196] expansions, so that bond breaking is described accurately. By additionally including extrapolations to the complete basis set limit[196] (as well as relativistic corrections), these authors were able to calculate full diatomic potential energy curves to extremely high accuracy.[199−203]

Recently, Khait et al.[204] developed the new hybrid variational-perturbational MRCISD(TQ) approach, which builds a noniterative correction to the MR-CISD energies in order to approximate the effects of triple and quadruple excitations. Several numerical studies suggest that MRCISD(TQ) recovers the dynamic electron correlation in a balanced way, i.e., not strongly dependent on the particular electronic state, and has significant promise as a computationally tenable ultrahigh-precision approximation. Most recently, Kaith et al.[106] introduced a further variant (nR-MRCISD(TQ)), which is related to the B_k method[1] and can be used to calculate the energy of several states at the same time (multiroot). Its noniterative variant is close to the original MRCISD(TQ) of ref 204. Test calculations have shown[106] that the nR-MRCISD(TQ) method provides a very high accuracy, even when there are strongly quasidegenerate states and the nonparallelity errors are typically improved by an order of magnitude relative to MR-CISD.

2.1.5. Approximate CI Methods. This section discusses several methods in which the expansion spaces are approximated relative to the MRCI or full-CI expansion space, or the Hamiltonian operator is approximated, or both.

2.1.5.1. Contracted MRCI Methods. There are two basic contraction schemes used in CI calculations: internal and external. Both are based on grouping together certain primitive expansion functions. In an internally contracted MRCI (ic-MRCI), the CI expansion space is generated by applying excitation operators to the multiconfigurational reference wave function. In analogy to the SR case eqs 12 and 13

$$\{|m\rangle; m = 1...N_{\mathrm{dim}}\} = \{|\tilde{\Psi}_0\rangle, |\tilde{\Psi}_i^a\rangle, |\tilde{\Psi}_{ij}^{ab}\rangle, |\tilde{\Psi}_{ijk}^{abc}\rangle, ...\} \tag{47}$$

with

$$|\tilde{\Psi}_0\rangle = \sum_j^{N_{\mathrm{ref}}} c_j^{\mathrm{ref}}|j; \mathrm{ref}\rangle; \qquad |\tilde{\Psi}_i^a\rangle \equiv a_a^\dagger a_i|\tilde{\Psi}_0\rangle;$$
$$|\tilde{\Psi}_{ij}^{ab}\rangle \equiv a_a^\dagger a_b^\dagger a_j a_i|\tilde{\Psi}_0\rangle; \qquad |\tilde{\Psi}_{ijk}^{abc}\rangle \equiv a_a^\dagger a_b^\dagger a_c^\dagger a_k a_j a_i|\tilde{\Psi}_0\rangle; \ ... \tag{48}$$

This approach was proposed[205−207] in the late 1970s and early 1980s. The ic-MRCI wave function is expanded as (cf. eq 16, the form of the uncontracted MRCI wave function)

$$|\psi\rangle = \tilde{c}_o|\tilde{\Psi}_0\rangle + \sum_{i,a}\tilde{c}_i^a|\tilde{\Psi}_i^a\rangle + \sum_{i>j,a>b}\tilde{c}_{ij}^{ab}|\tilde{\Psi}_{ij}^{ab}\rangle + \sum_{i>j>k,a>b>c}\tilde{c}_{ijk}^{abc}|\tilde{\Psi}_{ijk}^{abc}\rangle + ... \tag{49}$$

The coefficients of the underlying determinant expansion space are seen to be given by products of the reference coefficients and CI expansion coefficients. For example, the double-excitation terms may be written

$$\sum_{i>j,a>b}\tilde{c}_{ij}^{ab}|\tilde{\Psi}_{ij}^{ab}\rangle = \sum_{i>j,a>b}\sum_m^{N_{\mathrm{ref}}}(c_m^{\mathrm{ref}}\tilde{c}_{ij}^{ab})|\psi(m)_{ij}^{ab}\rangle \tag{50}$$

where $|\psi(m)_{ij}^{ab}\rangle$ (see eq 15) is a determinant in the uncontracted expansion space. The reference coefficients $\mathbf{c}^{\mathrm{ref}}$ are fixed by the MCSCF calculation, so the number of variational parameters, the $\tilde{c}_i^a$, $\tilde{c}_{ij}^{ab}$, $\tilde{c}_{ijk}^{abc}$, ... coefficients, is comparable to a SR expansion. Because the product $(c_m^{\mathrm{ref}}\tilde{c}_{ij}^{ab})$ for fixed $c_m{}^{\mathrm{ref}}$ does not have the full flexibility of the uncontracted coefficients $c(m)_{ij}^{ab}$ of eq 16, the contracted expansion space is a subspace of the full uncontracted MRCI expansion space and the computed eigenvalues are variationally bounded from below by the uncontracted MRCI energies. See Shavitt[208] for further discussion of the consequences of the internal contraction approximation. The Hamiltonian matrix element contributions may be computed with reduced density matrices within the occupied orbital space, and once these density matrices are available, the computational effort of each iteration of the optimization procedure is largely independent of the number of reference functions. The main advantage of the ic-MRCI scheme is that it requires less computational effort and thereby allows the use of much larger reference spaces than the traditional (uncontracted) MRCI method. The most successful implementation of this approach was reported by Werner and Knowles.[209,210] In this implementation, the expansion space is limited to single and double excitations,

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

but only the double excitations are contracted; the full reference space and the full space of uncontracted single-excitation configurations are included in the expansion space. CEPA-type versions (MR-ACPF, QDVPT, MR-CEPA(0)) have been introduced and tested,[211] and also MR-AQCC is available.[120] The appropriate version for excited states has also been proposed.[212] A parallel version has been implemented, and its efficiency has been demonstrated.[213] The ic-MRCI scheme, through its implementation in MOLPRO,[120] is very popular, and there is a very extensive list of applications (over 1500) in the literature that use this method.

Unfortunately, in many applications the ic-MRCI method is not considered as an approximate variant, and the resulting contraction error (or uncertainty bound) is usually not discussed. As should be clear from the discussion above, this is a rather severe approximation that is usually valid, but there is no systematic way to estimate the contraction error nor are there systematic studies in the literature to assess these errors. For partial studies on the accuracy of ic-MRCI see refs 214−216.

The MC-CEPA method of Fink and Steammler[153] also uses the internally contracted approximation.

Note in passing that the popular CASPT2 method[217−221] may be considered an approximation of the ic-MRCI approach. This is a "diagonalize-then-perturb" approach (see section 2.1.5.7) in which the reference function is first determined with the diagonalization of **H** within the reference space, and then the parameters $\tilde{c}_i^a$, $\tilde{c}_{ij}^{ab}$, $\tilde{c}_{ijk}^{abc}$, ... are determined with perturbation theory. In the CASPT2 method, the expansion is limited up to only contracted double excitations, and the second-order energy is determined from the first-order corrections to the wave function. This energy is not bounded from below by either the uncontracted MRCI energy or the full-CI energy.

The externally contracted MRCI concept was introduced by Siegbahn[222] on the basis of the single-reference variant.[223] This is a "perturb-then-diagonalize" approach (see section 2.1.5.7). By grouping together configurations with the same internal parts and freezing their relative weights (e.g., for double excitations),

$$|m(ij);ci\rangle = \sum_{a>b} c(m)_{ij}^{ab} |\psi(m)_{ij}^{ab}\rangle \tag{51}$$

with the contraction coefficients determined from first-order perturbation theory

$$c(m)_{ij}^{ab} = \frac{\langle\psi_0|H|\psi(m)_{ij}^{ab}\rangle}{E_0 - \langle\psi(m)_{ij}^{ab}|H|\psi(m)_{ij}^{ab}\rangle} \tag{52}$$

the number of variational parameters is drastically reduced. A variant in which the denominator is computed with orbital energy differences (i.e., Møller−Plesset PT rather than Epstein−Nesbet PT) was also discussed.[222] Although the contraction coefficients are defined with low-order PT, the final energies are computed from the eigenpairs of the **H** matrix in the space $\{|m;\mathrm{ref}\rangle, |m(i);ci\rangle, |m(ij);ci\rangle;\ m = 1...N_{\mathrm{ref}}\}$. In a GUGA implementation, which addresses the problems associated with enumerating the individual terms in this method, the external (i.e., virtual) orbitals may be placed at the bottom of the graph and the dimension of the contracted expansion space is then the number of internal walks in the Shavitt graph.[222] This expansion space, which has a larger dimension than the reference space alone, is a subspace of the full uncontracted space, and therefore the computed eigenvalues are variationally bounded from below by the uncontracted MRCI energies. The method is best suited for ground and low-lying valence excited states, but less so for higher lying excited states with significant external orbital occupations. It was shown that the loss of correlation energy is usually less than 2%. The effort required is comparable to a single iteration of an uncontracted MRCI calculation. It was implented as a conventional CI method (the matrix elements were explicitly constructed and stored) in which the eigenpairs were solved iteratively.[222] Unlike the internally contracted method, the effort depends directly on the dimension of the reference space (or, more specifically, on the number of internal walks). Consequently, it is not as popular as the ic-MRCI and CASPT2 approaches which are much more efficient for larger reference expansions. An analytic gradient has been formulated by Lee,[224] but, to our knowledge, it was never implemented into computer code. Wang et al.[225] have recently suggested an improved procedure for the perturbational determination of the external contraction coefficients.

The simultaneous use of both internal and external contraction was suggested by Wang et al.[226] This particular implementation is based on an improved hole−particle formalism.[227] Although the initial results were very promising,[226] only a few applications have been reported.

2.1.5.2. Graphically Contracted Function Method. In the graphically contracted function (GCF) method,[228−237] the wave function is represented using the graphical unitary group approach[6,28,238,239] in which the expansion CSFs of the unitary group approach[240−242] are represented graphically. The wave function is expanded as a linear combination of GCFs

$$|\psi\rangle = \sum_{P=1}^{N_{\mathrm{GCF}}} c_P |P\rangle \tag{53}$$

where the basis functions $|P\rangle$ in turn are contractions over the CSF basis of dimension N_{CSF}

$$|P\rangle = \sum_{m=1}^{N_{\mathrm{CSF}}} x_m^P |m\rangle \tag{54}$$

The contraction coefficients x_m^P are products of *arc factors* associated with the arcs of the Shavitt graph

$$x_m^P = \prod_{u=1}^{n} \alpha_{\mu(u,m)}^P \tag{55}$$

where $\mu(u,m)$ denotes the arc associated with orbital u in CSF m. In contrast to the internally and externally contracted CI approaches, the arc factors are contraction coefficients over the full orbital range. Each contracted basis function $|P\rangle$ corresponds to a particular set of arc factors $\boldsymbol{\alpha}^P$. Consequently, the wave function depends on the linear coefficients c_P and on the nonlinear arc factor parameters $\boldsymbol{\alpha}^P$. Initial calculations show that N_{GCF} dimensions typically in the range of 10−20 are sufficient to achieve chemical accuracy[231] for single-state calculations. For full-CI Shavitt graphs, the total number of variational parameters grows only as $O(N_{\mathrm{GCF}}N^2n)$ rather than exponentially n^N as in the full-CI expansion.[230,232]

An energy expectation value requires the computation of Hamiltonian

$$H_{PQ} = \langle P|H|Q\rangle = \sum_{p,q} h_{pq}\langle P|E_{pq}|Q\rangle + \frac{1}{2}\sum_{p,q,r,s} g_{pqrs}\langle P|e_{pqrs}|Q\rangle \tag{56}$$

and overlap $S_{PQ} = \langle P|Q\rangle$ (metric) matrix elements in the contracted basis. The wave functions are optimized to minimize the energy with respect to both the linear coefficients **c** and the full set of nonlinear parameters $\boldsymbol{\alpha}$. The variational energy

corresponds to a solution of the generalized symmetric eigenvalue equation

$$\mathbf{Hc} = \mathbf{Sc}E \tag{57}$$

The H_{PQ} and S_{PQ} matrix elements depend on the arc factors $\boldsymbol{\alpha}^P$ and $\boldsymbol{\alpha}^Q$, and thus the linear expansion coefficients $\mathbf{c} \equiv \mathbf{c}(\boldsymbol{\alpha})$ and the energy $E \equiv E(\boldsymbol{\alpha})$ depend on the full set of arc factors. An overlap matrix element S_{PQ} may be computed with a recursive algorithm that requires $O(N^2n)$ effort for full-CI Shavitt graphs.[228] The full overlap matrix thereby requires $O(N_{\mathrm{GCF}}{}^2N^2n)$ effort. The scaling for other expansions is discussed in ref 228. A Hamiltonian matrix element H_{PQ} may be computed with a recursive algorithm[229] that requires $O(N^2n^4)$ effort for full-CI Shavitt graphs, and the full matrix $\mathbf{H}$ thereby requires $O(N_{\mathrm{GCF}}{}^2N^2n^4)$ effort. Given a set of arc factors, a single H_{PQ} evaluation requires about the same effort as a single iteration of an SCF method. The optimization of wave functions requires the gradients of the energy with respect to the nonlinear arc factor parameters. A recent gradient algorithm has been implemented that requires $O(N_{\mathrm{GCF}}{}^2N^2n^4)$ effort for full-CI Shavitt graphs;[230] in fact the gradient costs about three times that of the energy itself. Hamiltonian matrix and gradient timings have been reported[237] for full-CI Shavitt graphs as large as $n = N = 180$ and $N_{\mathrm{CSF}} \approx 10^{104}$. The overlap $\langle D|\psi\rangle$ with an arbitrary Slater determinant $|D\rangle$ may be computed with effort that scales as $O(N_{\mathrm{GCF}}Nn)$; timings on a laptop computer of a few milliseconds per determinant have been reported for full-CI Shavitt graphs[233] as large as $n = N = 260$ and $N_{\mathrm{CSF}} \approx 10^{152}$. The spin-density matrix, which may be used to compute M-dependent expectation value properties (see eq 11), may be computed with effort that scales as $O(N_{\mathrm{GCF}}{}^2N^2n^2)$ for full-CI Shavitt graphs; timings have been reported[234] for systems as large as $n = N = 360$ and $N_{\mathrm{CSF}} \approx 10^{212}$. The method has been extended to include spin-orbit interaction using multiheaded Shavitt graphs.[232,235]

The computational effort for these quantities does not depend directly on the CSF expansion length N_{CSF}; thus, this method allows wave function expansions with N_{CSF} values that are many orders of magnitude larger than can be accommodated by traditional electronic structure methods. For all of these matrix elements and properties, the graphical representation of the underlying CSF expansion space along with the orbital-by-orbital contractions of the basis functions provided by eq 55 allows for the development of fully recursive algorithms, thereby eliminating from practical consideration any direct dependence on the large values of N_{CSF}.

The GCF method is characterized by several important features. Because the method is formulated in terms of spin eigenfunctions using GUGA, it does not suffer from spin contamination or spin instability. Open-shell spin eigenfunctions are included in the wave function expansions. This allows significant flexibility in the individual GCF basis functions to describe radicals and other open-shell electronic states. For example, a single expansion term, $N_{\mathrm{GCF}} = 1$, is sufficiently flexible to correctly dissociate the triple bond of N_2 to the high-spin ^{4}S ground-state atomic fragments. There are no artificial excitation-level or occupation restrictions with respect to a reference function or reference space. Because the wave function is expanded as a linear combination of N_{GCF} basis functions, the method can be used for both ground and excited electronic states, the increased wave function flexibility leads to more accurate wave functions, and this expansion allows the straightforward computation of transition moments, nonadiabatic coupling, and other properties that at present can only be computed reliably with MCSCF and MRCI approaches. In analogy to the subspace equation of the Lanczos or Davidson methods eq 29, the eigenvalues of eq 57 satisfy the subspace bounds relations of eqs 4 and 5. State averaging allows the arc factors to be optimized for a weighted average of states rather than for an individual state.[243]

The GCF method is still relatively immature, and only a few chemical applications have been reported. These include[230,231,243] the dissociation of the ground $1^1\Sigma^+_g$ state of N_2, the symmetric dissociation of the ground 1^1A_1 state of H_2O, the reaction path curves for the $1-2^1A_1$ states of Be + $H_2 \rightarrow$ BeH_2, and the dissociation of the $X^1\Sigma^+_g$, $B^1\Delta_g$, and $B'1\Sigma^+_g$ states of C_2. To date, the main difficulty with general application of the method is the nonlinear optimization of the arc factor parameters.[231]

2.1.5.3. Density Matrix Renormalization Group. The exponential scaling n^N of the number of CSFs (or determinants) with system size limits the practical applicability of the CASSCF approach to active spaces with approximately 16 electrons and orbitals. One approach that offers the possibility to eliminate this scaling is the density matrix renormalization group (DMRG) method originally developed by White in the context of condensed matter.[244,245] Although the DMRG method is relatively recent in the field of ab initio quantum chemistry, it has already proven useful in addressing questions that are outside of the realm of traditional quantum chemistry approaches.[244−273]

Within the language of DMRG, orbitals occupy *sites* of a one-dimensional *lattice*. In the two-site DMRG algorithm, the lattice is divided into three blocks: (i) the *system* block, (ii) the *environment* block, and (iii) two sites in between (note that some authors in the literature use *left* and *right* to denote the system and environment blocks, respectively). Each spatial orbital has four possible Fock states $|\sigma\rangle = \{|00\rangle,|10\rangle,|01\rangle,|11\rangle\}$, using the spin-orbital occupation number representation. Hence, the total number of many-electron states (determinants) for each of the three blocks is 4^{n_s}, $4^2 = 16$, and 4^{n-n_s-2}, where n_s is the number of sites (orbitals) in the system block, and n is the total number of correlated orbitals. The wave function of the total system (also called the *superblock*) is given by the tensor product space of the many-particle states of the blocks

$$\begin{aligned}|\psi\rangle &= \sum_{a_S\sigma_S\sigma_E a_E} \psi_{a_S\sigma_S\sigma_E a_E}|a_S\rangle \otimes |\sigma_S\rangle \otimes |\sigma_E\rangle \otimes |a_E\rangle \\ &= \sum_{i_S i_E} \psi_{i_S i_E}|i_S\rangle \otimes |i_E\rangle \end{aligned} \tag{58}$$

where the middle block is split into the Fock states $|\sigma_s\rangle$ for the orbital next to the system block and $|\sigma_E\rangle$ for the orbital next to the environment block. To ensure that the wave function has the proper symmetries (i.e., it has the right number of electrons and spin projection S_z), many terms are excluded from eq 58. Even with the exclusion of these terms, the number of parameters in eq 58 still scales exponentially with system size. To overcome this exponential scaling, DMRG restricts the number of many-electron states that describe the system and environment blocks. The central question of DMRG then becomes: Given a predefined threshold (usually denoted M) for the maximum number of states, how does one obtain the optimal many-electron states?

The DMRG wave function is optimized via a sweep algorithm, where sites are traversed sequentially, and each step in the sweep

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

consists of three parts:[248] *blocking*, *diagonalization*, and *decimation*. At the start of the sweep, n_s is chosen to be small enough such that the number of basis states in the system block does not exceed M, and, hence, all basis states $\{|a_s\rangle\}$ may be computed explicitly. The Hamiltonian operator and elementary creation and annihilation operators, as well as other operators that are needed to compute the Hamiltonian for the superblock are evaluated in this basis.[246] In addition, a suitable basis with dimension M for the environment block must also be available (from either a warm-up procedure[248,250,262,264,267] or the previous sweep).

In the blocking step of the algorithm, the system and environment blocks are enlarged by one site as in eq 58, and the Hamiltonian matrix for each new block is computed as a direct product of the Hamiltonian matrices and elementary operators in the old bases.[246] The dimension of the new many-electron spaces (and the associated operators) is $4M$, and the most expensive part of the blocking step scales[250] as $O(M^2n^3)$.

During the diagonalization step, the lowest few eigenpairs of the superblock Hamiltonian are computed using an iterative subspace diagonalization method such as the Davidson algorithm (see section 2.1.2). These methods require the repeated evaluation of matrix vector products, $\mathbf{w}_{S,E} = \mathbf{H}_{S,E}\mathbf{x}_{S,E}$, where $\mathbf{H}_{S,E}$ is the superblock Hamiltonian (never explicitly computed or stored) with dimension $16M^2$. It may seem that evaluation of these matrix−vector products requires a computational effort proportional to $O(M^4)$. However, utilizing the fact that the superblock Hamiltonian is a sum of products of operators that act on the system and environment blocks

$$[H_{S,E}]_{i_S k_E; j_S l_E} = \sum_{\alpha} H^{\alpha}_{i_S j_S} H^{\alpha}_{k_E l_E} \tag{59}$$

reduces the effort[250] to $O(M^3n^2)$. In eq 59, indices with an S or E subscript denote many-electron states of the system or environment block, respectively, and the sum is over all terms of the Hamiltonian. Note that the efficiency of the subspace method also depends on the quality of the initial guess for the eigenvector and the preconditioner used for computing vector updates.

In the decimation step, the many-electron reduced density matrix for the system block is computed as

$$\rho_{i_S j_S} = \sum_{k_E} |\psi_{i_S k_E}\rangle\langle\psi_{j_S k_E}| = \sum_{i=1}^{4M} w_i |\theta_i\rangle\langle\theta_i| \tag{60}$$

where the eigenvalues in the spectral representation obey[244−246] $\sum_i w_i = 1$. Note that this procedure is similar to the construction of natural orbitals; however, the eigenvectors $|\theta_i\rangle$ of the density matrix in eq 60 are many-electron functions rather than one-electron orbitals. It can be shown that the eigenvectors of the density matrix minimize the distance in the quadratic norm[244,245,272,273] $\||\Psi\rangle - |\tilde{\Psi}\rangle\|_2$, where the approximate wave function is obtained by retaining the M eigenvectors with the highest eigenvalues[244−246,248,256] in eq 60.

Using these eigenvectors, the relevant operators for the system and environment blocks are transformed according to

$$\tilde{\mathbf{A}} = \mathbf{O}\mathbf{A}\mathbf{O}^{\mathrm{T}} \tag{61}$$

Since one such transformation requires a computational effort that is proportional to the cube of the number of states retained, and there are on the order of n^2 operators that need to be transformed, the overall cost of this step scales[250] as $O(M^3n^2)$. The DMRG sweep is then continued at the blocking step.

A DMRG sweep is complete when one end of the lattice is reached. At this point the system and environment are interchanged, and the sites are traversed in reverse order. The DMRG sweeps continue until the energy is converged with respect to sweeping. The computational effort for a single sweep of the DMRG algorithm scales polynomially $O(M^3n^3) + O(M^2n^4)$ with the number of active orbitals.[250]

As can be readily seen from eq 60, the accuracy of the DMRG method depends on the number of states retained during the decimation step. Indeed, it has been pointed out that DMRG is expected to perform best for one-dimensional problems (such as the one-dimensional Hubbard model and other one-dimensional lattice models) and the treatment of higher dimensional problems (such as the two-dimensional lattice models or general three-dimensional molecular systems) should present a greater challenge since the decay of the eigenvalue spectrum of the density matrix slows exponentially with inverse system size.[272] As discussed in more detail in ref 248, the leading term in the error in the DMRG energy (δE) may be approximated as

$$\ln|\delta E| \cong -\kappa(\ln M)^2 \tag{62}$$

where κ is a model-specific constant related to the correlation length. In addition to errors associated with truncation, the ordering of the orbitals introduces an "artificial lattice correlation"[248] that can affect the convergence of DMRG. Chan and Head-Gordon[248] used a reverse Cuthill−McKee reordering of the orbitals to make the one-electron integral matrix close to band-diagonal. Mitrushenkov et al.[263] suggest ordering the orbitals on the basis of the diagonal two-electron integrals and orbital energies. Moritz et al.[268] employ several criteria on the basis of the one- and two-electron integrals as well as a genetic algorithm to examine the optimal orbital ordering in DMRG calculations involving the chromium dimer. Legeza et al.[266] find that the reverse Cuthill−McKee ordering of the orbitals, in addition to reducing the error in the energy for a given number of states, also reduces the number of sweeps for converging the DMRG calculations.

Early applications of DMRG to ab initio quantum chemistry focused on assessing the applicability of the method for small molecules. Mitrushenkov et al.[263] find that the full-CI spectroscopic constants for Be_2, HF, and N_2 are accurately reproduced with 500 states. Chan and Head-Gordon find that less than a 1000 states are sufficient to yield DMRG energies with mE_h or better accuracy for water in a DZP and TZ2P basis set.[248,249] For the potential energy curve of the nitrogen dimer in a cc-pVDZ basis set with the 1s electrons on nitrogen frozen, Chan et al.[251] find that DMRG with a 1000 states yields results with mE_h accuracy and outperforms coupled-cluster theory with up to hextuple excitations. In addition, their results indicate that the inclusion of the 1s core into the DMRG calculations does not significantly affect the number of states required to achieve sub-mE_h accuracy. Low-lying excited states of LiF,[266] CH_2,[267] HNCO,[259] acenes as large as pentacene (minimal basis and only the π orbitals and electrons are correlated),[258] CsH,[269] and CoH[270,271] have been studied with DMRG. More recently, Kurashige and Yanai[264] have applied their parallel DMRG code to assess the accuracy of DMRG for applications to transition metal complexes. Their results indicate that the number of states required to achieve sub-mE_h accuracy for these complexes requires a larger number of states. For the chromium dimer with a $\mathbf{30}^{24}$ active space (24 electrons distributed in 30 orbitals; see

section 2.2.3 for a discussion of wave function expansion spaces), the M = 6400 results are approximately 1 mE_h above the extrapolated $M = \infty$ DMRG energy, and the error is reduced by approximately a factor of 2 for M = 10 000. These calculations are the largest reported DMRG expansions; on four Intel Xeon 2.66 GHz Quad Core CPUs (using four threads per CPU) a single sweep requires 6.2 h or more of wall time.[264] Due to the number of sweeps needed to converge the energy, DMRG is more expensive than single-reference methods such as CCSD(T).[269] Nonetheless, one has to consider that the number of sweeps can be reduced by optimal ordering of the orbitals and that DMRG is applicable to systems with strong static correlations. In addition, the polynomial scaling of the DMRG algorithm allows the essentially exact treatment of much larger Hilbert spaces than are practical with traditional methods such as full-CI.

More recently, Hachmann et al.[253] have developed the local DMRG method (LDMRG) that utilizes integral screening and localized orbitals to reduce the computational cost of the LDMRG method to $O(M^3n^2) + O(M^2n^2)$ per sweep. Due to the local nature of electron correlation, for molecules that are extended along one of their directions, the accuracy of LDMRG does not depend on the number of states, and the effort scales quadratically with the number of orbitals.[253] Applications of LDMRG to all-trans polyenes (maximum active space of $\mathbf{48}^{48}$) and linear hydrogen chains (as large as $(H_2)_{50}$ with an active space of $\mathbf{100}^{100}$) demonstrate that LDMRG with M = 250 yields energies with μE_h accuracy. For the largest hydrogen chain with M = 250, a single sweep took 73 min on 18 2.0 GHz Opteron processors.[253] LDMRG has since been applied to study other strongly correlated π-systems such as (i) the radical character of acenes[254] (as large as 12-acene with an active space of $\mathbf{50}^{50}$), (ii) the excited states of the acenes[258] (as large as 5-acene with an active space of $\mathbf{22}^{22}$), (iii) excitation energies and oscillator strengths in β-carotene[256] ($\mathbf{22}^{22}$ active space and orbital optimization), and (iv) spin gaps in the poly(m-phenylenecarbenes)[265] ($\mathbf{46}^{46}$ active space).

2.1.5.4. Individual Selection. The individually selected CI method with extrapolation was introduced by Buenker and Peyerimhoff[274−276] in the 1970s (for a review up to 2000, see ref 277). The idea is to partition the CI expansion space into two subspaces: one contains the most important configurations, which are treated explicitly, while the other ones are either completely neglected or their contributions are approximated by perturbation expressions. The energies are computed with extrapolation methods to estimate the full-CI values. The method of Peyerimhoff and Buenker has been implemented into the MRD-CI program package,[278] which was widely used by various groups to study problems in several fields of chemistry including spectroscopy, reaction mechanisms, and so on. The idea of the individually selected MRCI method was also used by Whitten and Hackmeyer,[279] in the CIPSI program by Malrieu and co-workers,[280,281] and in the MELDF program of Davidson and co-workers.[282] While older programs using individual selection could not use the direct-CI algorithm, more recently several programs have appeared which address this limitation including MELDF,[283] a code by Harrison,[284] CIPSI,[285,286] and MRD-CI[287] itself. Perhaps the most advanced algorithm has been presented by Hanrath and Engels[288] in their DIESEL-MR-CI program. A massively parallel implementation of selected CI has been implemented by Stampfusset et al.[289,290]

A general selection procedure has been developed by Bytautas and Ruedenberg[291] that focuses on identifying independently all important configurations (rather than "configurational deadwood") among the quadruple, quintuple, and sextuple excitations. This "configurational livewood" is deduced in advance, on the basis of information extracted from double and triple excitations so that all important terms can be included in constructing the wave function. The procedure is notably effective for the a priori accounting of all configurations required for a specified accuracy in a reference space, as illustrated for the CISDTQ56 wave functions of the molecules HNO, N_2, and NCCN.[291]

2.1.5.5. Local Approaches. Local approaches for CI and CEPA were first applied by Saebo and Pulay,[292] following the same approach as their local perturbation theory. Walter and Carter[293] used the same techniques to first define a local MR-CISD which eliminates simultaneous excitations from widely separated internal orbitals. Shortly later they extended the method to use the locality of the virtual orbitals as well.[294] An ACPF version was also constructed and tested.[295] Similar procedures have been put forward by Bories et al.,[296] where the localized orbitals have been obtained from the LCASSCF (local CASSCF) approach of Maynau et al.[297] Recently, a linear scaling version of the local MR-CISD method has been given by Chwee and Carter[298] and more recently further improved by Cholesky decomposition (CD-LMRCISD for Cholesky-decomposition local MR-CISD)[299] and density fitting approach.[300] Applicability to excited states has been also demonstrated.[301]

Reinhardt et al.[302,303] have also developed a local contracted CI method. In this approach localized bonds are used to define contracted double excitation functions. The Hamiltonian of the system is built from the small Hamiltonians of the localized fragments. CEPA versions and inclusion of higher excitations are also discussed.[303]

2.1.5.6. Pseudospectral Methods. Pseudospectral methods were introduced in quantum chemistry by Friesner.[304,305] The details of this approximation have been reviewed by Martinez and Carter.[306] A pseudospectral version of full-CI was introduced by Martinez et al.[307] and of double-excitation CI (CID) by Martinez and Carter.[308] The first MRCI application was reported by Murphy et al.[309] and shortly afterward by Martinez and Carter.[310,311] Although the pseudospectral CI methods are very precise, i.e., only small error of a few tenths of a mE_h is introduced with respect to the traditional version, the savings in computer time is minimal. Carter and Walter[312] report speedups of ∼3.7 for SR-CID and even smaller speedups of ∼2.2 for MR-CISD. Reynolds et al.[313] combined the pseudospectral methods with local treatment of correlation but reported "only meager computational savings".[312] An improved implementation was reported by Reynolds and Carter[314] with overall gains in computer time in the range of 3−5, but also with considerable reductions of disk usage. A later version of mixed local and pseudospectral treatment was published by Walter et al.[315] with speedup of over a factor of 7.

For the sake of completeness, note that the term "reduced scaling MR-CI" has been used by Carter et al. (e.g., see ref 312) to cover both the localized approach and/or pseudospectral method.

2.1.5.7. Multireference Perturbational Approaches. A detailed discussion of multireference perturbation theory (PT) approaches is out of the scope of this review for two reasons: (i) special techniques different from MRCI are used, and (ii) the theory involved is very diversified and therefore a detailed discussion would be very lengthy. However, MRCI and multireference PT approaches are often used together in applications, and a short summary of different versions is perhaps of interest for the reader of the present MRCI review.

There are two basic ways to generalize perturbation theory for multireference situations. One is the quasidegenerate perturbation theory (QDPT) first introduced by Brandow[316] and later developed by others.[317−321] In Brandow's original formulation the reference space was chosen as a CAS, and the many-electron space, as usual in perturbation theory, was represented by determinants. Since the theory closely follows many-body perturbation theory, the acronym MR-MBPT was often used for these methods. Because of the QDPT background, these methods lead to the diagonalization of an effective Hamiltonian (often called "*perturb-then-diagonalize*") which leads to a simultaneous description of several states. However, this approach results often in the appearance of intruder states, expansion terms that are outside of the reference space with low energies, sometimes even lower than some of the reference space energies, and whose contributions are not described well with PT expressions. To avoid intruder states, a formulation with a restricted reference space was introduced by Hose and Kaldor[318] and later by Meissner and Bartlett,[319] while a different partition was used by Kozlowski and Davidson[322] and by Nakano.[323] Malrieu et al.[321] suggested the use of so-called intermediate Hamiltonians, which are built from only a few functions of the CAS model space (see also below). We note in passing that we have previously discussed the closely related quasidegenerate variational perturbation theory methods by Cave and Davidson[149,150] and Murray et al.[151] among CEPA-type MR methods (section 2.1.3.2).

The other possibility is the use of "*diagonalize-then-perturb*" type methods, which usually start with a particular state obtained from an MCSCF calculation and then follow similar lines as in Møller−Plesset PT (therefore the name MR-MP is often used). This *state-specific* formulation has clear practical advantages, which accompany the theoretical disadvantage of lack of rigorous size-extensivity. The closest analogy to the SR case can be obtained with an internally contracted formulation.[205] The GVB-MP2 of Wolinsky and Pulay[324] was perhaps the first successful application of this technique (see also the discussion of earlier methods in ref 324). Roos and co-workers[217−221] introduced their own version under the name CASPT2 which has became the most popular implementation (see also ref 123). Versions not using internal contraction have been proposed by Murphy and Messmer,[325,326] Kozlowski and Davidson,[327] and Hirao[328,329] (closely related to Nakano's MC-QDPT method[323]). The latter is implemented in GAMESS.[121,122] CASPT2 and related methods have been reviewed recently by Pulay.[330]

Dyall[331] showed that a proper zero-order MR Hamiltonian should include also two-electron terms. The NEVPT2 (*N*-electron valence state perturbation theory) method by Malrieu and co-workers[332,333] is based on this proposition. This method uses excitation classes, an idea similar to the consideration of redundancies in CEPA-type methods.[103] The drawback of this method is its formal complexity and dependence of the results on orbitals labels.[330] A quasidegenerate version has also been proposed recently.[334]

The disadvantage of the diagonalize-then-perturb methods is that they do not take into account the relaxation of the reference functions due to the correlation introduced by perturbation. This relaxation may be included by using the above-mentioned intermediate Hamiltonian idea by Malrieu et al.,[321] i.e., including a few of these reference functions in the perturbation ansatz, while avoiding the disadvantages of full QDPT theory. Such methods have been proposed by Shavitt[208] (see also Stahlberg[335]) and by Hoffman and co-workers[336−340] under the name GVVPT (generalized Van Vleck perturbation theory). The latter method has proven to yield quite accurate results in some complicated situations,[336] and also analytic gradients are available.[340]

Finally we mention that interesting new ideas have been proposed by Surján et al.[341−343] in their MC-PT (multiconfiguration PT) versions, by Rolik and Szabados[344] in multipartition multireference many-body perturbation theory, and by Mukherjee and co-workers in their state-specific (SS) MR-MBPT method.[345,346]

2.1.5.8. Semiempirical Approaches. Semiempirical MRCI approaches could also be considered as approximations to MRCI. These are not covered in this review since the main approximation is the formation of the underlying semiempirical Hamiltonian and is therefore outside of the scope of the present study. These methods are, however, important within certain fields of chemistry; therefore the interested readers are referred to recent applications (e.g., see refs 347−350). A new implementation of MRCI in the semiempirical framework has been reported recently by Lei at al.[351]

2.1.6. Transition Moments. The calculation of transition moments between different electronic states is required, for example, for the simulation of absorption and emission spectra. This calculation is straightforward if in the CI expansion the same orbital basis is used for all states. This is not always the case, considering especially, but not only, the CASSCF method where it may be desirable to perform independent calculations for different states. To take such cases into account, efficient methods have been developed on the basis of biorthogonal orbitals for full-CI by Moshinsky and Seligman.[352] On the basis of this work, Malmquist[353] and Malmquist and Roos,[354] within the framework of the restricted active space SCF (RASSCF) approach, have developed the CAS and RAS state interaction (CASSI and RASSI) methods, respectively, which have been applied successfully in many cases (see, e.g., refs 355 and 356). An extension to the internally contracted CI method has been reported by Mitrushchenkov and Werner.[357] Calculation of the transition moments at the MR-AQCC and MR-ACPF levels is also possible.[159]

2.2. MCSCF Method

The MCSCF method corresponds formally to a CI expansion for the wave function in which both the orbitals and the configuration expansion coefficients are optimized. However, practically there are also numerous other important distinctions in the methods. First the formalism for the MCSCF method is briefly reviewed. More complete reviews of MCSCF methodology, implementations, and applications are given in refs 19, 20, 214, 217− 219, 358, and 359.

2.2.1. MCSCF Wave Function Parameterization. The optimization of the orbital variations and the configuration expansion coefficients first requires a definition of the variational parameters. This is complicated by the orthonormalization constraints on the orbitals and the configuration expansion coefficients. The configurations may be chosen to be either determinants or CSFs. A CSF expansion is assumed in this discussion, but the expressions in terms of primitive Slater determinants follow in an analogous manner. One of the common parametrizations for the orbital variations is based on the fact that an orthogonal matrix **U** may be parametrized in terms of the elements of the skew-symmetric matrix **K**.

$$\mathbf{U} = \exp(\mathbf{K}) = 1 + \mathbf{K} + \frac{1}{2}\mathbf{K}^2 + \quad ... \\ + \frac{1}{m!}\mathbf{K}^m + \quad ... \tag{63}$$

with $K_{pq} = -K_{qp}$. The unique elements of **K** are unconstrained, a feature that simplifies the formulation of the iterative wave

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

function optimization and also the subsequent formulation of analytic gradients which is discussed in section 2.5. The formalism in this section focuses on real orbital and configuration coefficients. In many practical calculations, the molecular orbitals are chosen to transform as the irreps of the molecular point group. In these situations, the orbitals within each irrep may be grouped together and the transformation matrix $\mathbf{U}$ and the parameters $\mathbf{K}$ assume a block-diagonal form. Given the orbital matrix $\mathbf{K}$ there is a corresponding one-electron operator

$$K = \sum_{p,q} K_{pq}E_{pq} = \sum_{p>q} k_{(pq)}(E_{pq} - E_{qp}) \tag{64}$$

It is convenient to use the unique elements of the vector $\mathbf{k}$ (of length $n(n-1)/2$) in formulating the orbital variations. Given an initial reference orthonormal orbital basis $\boldsymbol{\varphi}^0$, and an arbitrary trial orbital basis defined as $\boldsymbol{\varphi} = \boldsymbol{\varphi}^0\mathbf{U}$, an arbitrary determinant or CSF that is written in terms of these two bases satisfies the relation

$$|j;\varphi\rangle = \exp(K)|j;\varphi^0\rangle \tag{65}$$

Because this relation is satisfied for a single configuration, it is also satisfied for an arbitrary linear combination of determinants or CSFs.

Similarly, the CSF coefficient variations are formulated in terms of the current reference wave function $|mc\rangle$ defined with the CSF coefficients $\mathbf{c}$ and some arbitrary, unnormalized, orthogonal wave function

$$|p\rangle = \sum_{m=1}^{N_{mc}-1} p_m^{\perp}|m_{\perp}\rangle = \sum_{m=1}^{N_{mc}-1} p_m|\tilde{m}\rangle \tag{66}$$

The basis $|m_{\perp}\rangle$ is some explicit representation of the orthogonal complement to $|mc\rangle$. This basis is useful for formal derivations, but the CSF basis is more useful for computer implementations. Several choices for this orthogonal complement basis are discussed in ref 19. In the second expression in terms of the primitive CSFs $|\tilde{m}\rangle$, the vector $\mathbf{p}$ satisfies $\mathbf{c}\cdot\mathbf{p} = 0$. The CSF coefficient variations may then be written with the operator

$$\exp(P) = 1 + (|mc\rangle, |\mathbf{p}|^{-1}|p\rangle)\begin{pmatrix}\cos|\mathbf{p}| - 1 & -\sin|\mathbf{p}| \\ \sin|\mathbf{p}| & \cos|\mathbf{p}| - 1\end{pmatrix}\begin{pmatrix}\langle mc| \\ |\mathbf{p}|^{-1}\langle p|\end{pmatrix} \tag{67}$$

with

$$P = |p\rangle\langle mc| - |mc\rangle\langle p| = \sum_m p_m P_m \tag{68}$$

and $P_m = |m\rangle\langle mc| - |mc\rangle\langle m|$. In this form, it is clear that the operator $\exp(P)$ defines a plane rotation between the two normalized basis vectors $|mc\rangle$ and $|\mathbf{p}|^{-1}|p\rangle$ within the CSF expansion space, and in particular

$$\exp(P)|mc\rangle = \cos|\mathbf{p}||mc\rangle + \sin|\mathbf{p}|(|\mathbf{p}|^{-1}|p\rangle) \tag{69}$$

This shows how an arbitrary normalized vector within the CSF expansion space depends on the direction $\mathbf{p}$, which defines the plane, and on the magnitude $|\mathbf{p}|$, which defines the angle of rotation within this plane. This expression also shows how, given a reference wave function $\mathbf{c}$ and a set of variational parameters $\mathbf{p}$, a new trial wave function is constructed, with no further numerical approximation, for the subsequent iterations. These two operators allow an arbitrary trial wave function, with arbitrary orthonormal orbitals and arbitrary normalized CSF coefficients, to be written

$$|\psi^{\text{trial}}\rangle = \exp(K)\exp(P)|mc\rangle \tag{70}$$

2.2.2. MCSCF Optimization Methods. The expectation value for the trial function in eq 70 may be written as

$$\begin{aligned} E^{\text{trial}} &= \langle\psi^{\text{trial}}|H|\psi^{\text{trial}}\rangle \\ &= \langle mc|\exp(-P)\exp(-K)H\exp(K)\exp(P)|mc\rangle \end{aligned} \tag{71}$$

The commutator expansion for the exponential operators allows the expansion of the trial energy in terms of the parameters $\boldsymbol{\lambda}^{\text{T}} = (\mathbf{k}^{\text{T}},\mathbf{p}^{\text{T}})$.

$$\begin{aligned} E^{\text{trial}}(\mathbf{k},\mathbf{p}) &= \langle mc|H + [H,K] + [H,P] + [[H,K],P] \\ &\quad + \frac{1}{2}[[H,K],K] + \frac{1}{2}[[H,P],P] + \ldots|mc\rangle \\ &= E^{mc} + \begin{pmatrix}\mathbf{k}\\ \mathbf{p}\end{pmatrix}^{\text{T}}\begin{pmatrix}\mathbf{f}^{mc}_{\text{orb}}\\ \mathbf{f}^{mc}_{\text{csf}}\end{pmatrix} \\ &\quad + \frac{1}{2}\begin{pmatrix}\mathbf{k}\\ \mathbf{p}\end{pmatrix}^{\text{T}}\begin{pmatrix}\mathbf{G}^{mc}_{\text{orb, orb}} & \mathbf{G}^{mc}_{\text{orb, csf}}\\ \mathbf{G}^{mc}_{\text{csf, orb}} & \mathbf{G}^{mc}_{\text{csf, csf}}\end{pmatrix}\begin{pmatrix}\mathbf{k}\\ \mathbf{p}\end{pmatrix} + \ldots \end{aligned} \tag{72}$$

The elements of the wave function optimization gradient $\mathbf{f}^{mc}$ and symmetric Hessian $\mathbf{G}^{mc}$ are

$$f^{mc}_{pq} = \langle mc|[H, E_{pq} - E_{qp}]|mc\rangle \tag{73}$$

$$f^{mc}_{j} = \langle mc|[H, P_j]|mc\rangle \tag{74}$$

$$\begin{aligned} G^{mc}_{pq,rs} &= \frac{1}{2}\langle mc|[[H, E_{pq} - E_{qp}], E_{rs} - E_{sr}] \\ &\quad + [[H, E_{rs} - E_{sr}], E_{pq} - E_{qp}]|mc\rangle \end{aligned} \tag{75}$$

$$G^{mc}_{pq,rs} = \langle mc|[[H, E_{pq} - E_{qp}], P_j]|mc\rangle \tag{76}$$

$$G^{mc}_{j,k} = \langle mc|[[H, P_j], P_k]|mc\rangle \tag{77}$$

Applying the variational condition to the trial energy gives

$$\begin{pmatrix}\partial E^{\text{trial}}(\mathbf{k},\mathbf{p})/\partial\mathbf{k}\\ \partial E^{\text{trial}}(\mathbf{k},\mathbf{p})/\partial\mathbf{p}\end{pmatrix} = \begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix} = \begin{pmatrix}\mathbf{f}^{mc}_{\text{orb}}\\ \mathbf{f}^{mc}_{\text{csf}}\end{pmatrix} + \begin{pmatrix}\mathbf{G}^{mc}_{\text{orb, orb}} & \mathbf{G}^{mc}_{\text{orb, csf}}\\ \mathbf{G}^{mc}_{\text{csf, orb}} & \mathbf{G}^{mc}_{\text{csf, csf}}\end{pmatrix}\begin{pmatrix}\mathbf{k}\\ \mathbf{p}\end{pmatrix} + \ldots \tag{78}$$

This infinite-order expression has no closed-form solution, so numerical iterative methods must be used to solve for the $\mathbf{k}$ and $\mathbf{p}$ parameters. Various iterative optimization methods involve truncation of the optimization equation (usually at first or second order, combined also with approximations to some of the matrix elements) along with a replacement of the reference wave function $|mc\rangle$ by $|\psi^{\text{trial}}\rangle$ using the approximate $\mathbf{k}$ and $\mathbf{p}$ parameters according to eq 70 for the subsequent iteration. When convergence is achieved, the gradient $\mathbf{f}^{mc}$ is zero and a solution to the nonlinear equation is obviously given by $\boldsymbol{\lambda} = \mathbf{0}$. The simple truncation at second order, resulting in a Newton−Raphson iterative procedure,

shows notoriously poor global convergence behavior, and consequently various stabilization methods are imposed in order to overcome these numerical difficulties.[19,219,358,360,214] In some situations, it is desirable to parametrize the trial energy in terms of the orbital variations only, along with the assumption that the CSF coefficients are always optimal. The appropriate trial energy expression in this case is written

$$E^{\text{trial}}(\mathbf{k}) = E^{mc} + \mathbf{k}^{\mathrm{T}}\mathbf{f}^{mc}_{\text{orb}} + \frac{1}{2}\mathbf{k}^{\mathrm{T}}(\mathbf{G}^{mc}_{\text{orb, orb}} - \mathbf{G}^{mc}_{\text{orb, csf}}(\mathbf{G}^{mc}_{\text{csf, csf}})^{-1}\mathbf{G}^{mc}_{\text{csf, orb}})\mathbf{k} + \ldots \tag{79}$$

This equation, based on the *partitioned orbital Hessian matrix*, also allows for the analysis of the eigenvalue spectrum of the wave function Hessian matrix for both ground and excited states.[19,361] For example, it is clear from the form of eq 79 that a variationally minimized energy, whether for a ground state or an excited electronic state, must correspond to a positive-definite partitioned orbital Hessian matrix.

There are several common approximations used in MCSCF methods. The evaluation of the commutators in eqs 73−77 allows the elements to be written directly in terms of the current Hamiltonian integrals and the current reduced density and transition matrix elements.

$$E^{\text{ref}} = \langle mc|H|mc\rangle = \sum_{pq} h_{pq}D_{pq} + \sum_{pq}(pq|rs)d_{pqrs} \tag{80}$$

$$f^{mc}_{pq} = 2(F_{pq} - F_{qp}) \tag{81}$$

$$F_{pq} = \sum_{t} h_{pt}D_{tq} + \sum_{tuv}(pt|uv)d_{qtuv} \tag{82}$$

$$f^{mc}_{j} = 2\langle j|H|mc\rangle = \sum_{pq} h_{pq}D^{j:mc}_{pq} + \sum_{pqrs}(pq|rs)d^{j:mc}_{pqrs} \tag{83}$$

$$G^{mc}_{pq,rs} = (1-P_{pq})(1-P_{rs})\{(F_{ps} + F_{sp})\delta_{qr} - 2h_{ps}D_{qr} + \sum_{uv} 4(pu|rv)d_{qusv} + 2(pr|uv)d_{qsuv}\} \tag{84}$$

$$G^{mc}_{pq,j} = 4(1-P_{pq})\{\sum_{t} h_{pt}D^{j:mc}_{qt} + \sum_{tuv}(pt|uv)d^{j:mc}_{qtuv}\} \tag{85}$$

$$G^{mc}_{j,k} = 2\langle j|H - E^{mc}|k\rangle \tag{86}$$

(For notational brevity, the above expressions are written with the index-permutation operators P_{pq} and P_{rs}.) Most MCSCF calculations partition the orbitals into three disjoint subsets: the *inactive* orbitals which are doubly occupied in each CSF in the reference expansion space, the *active* orbitals which have arbitrary occupations (0, 1, or 2) in the various CSFs, and the *virtual* orbitals which are unoccupied in each of the reference CSFs. The inactive and active orbitals together form the *occupied* orbital subset. Orbital rotations between pairs of inactive orbitals leave each individual CSF unchanged, the corresponding gradient element is always zero, and this set of orbital rotation variables k_{ij} is trivially redundant. Similarly, rotations k_{ab} between two virtual orbitals always leaves each CSF unchanged and are also trivially redundant. In most MCSCF implementations, these rotation variables are simply ignored (set to zero) during the orbital optimization procedure. A particular active−active rotation may or may not be redundant; some examples are discussed below. The density matrix elements in the above general expressions simplify with this type of orbital partitioning. If any density matrix subscript indexes a virtual orbital, then that element is zero; this places restrictions on the summation indices in the above expressions to range over only the occupied orbitals. Furthermore, if a subscript of a density matrix indexes an inactive orbital, then that density matrix element may be simplified with the identities

$$D_{it} = 2\delta_{it} \tag{87}$$

$$d_{ituv} = D_{it}D_{uv} - \frac{1}{4}D_{iu}D_{tv} - \frac{1}{4}D_{iv}D_{tu} \tag{88}$$

with i inactive and t,u,v general. This effectively means that only density matrix elements with active orbital indices need to be computed explicitly. This partitioning of the orbitals results in four different types of orbital gradient terms (with orbital indices ip, ia, pq, and pa) and in the corresponding 10 different types of orbital−orbital Hessian terms. Examination of the E^{ref}, the CSF gradient f^{mc}_{j}, and the CSF Hessian $G_{jk}{}^{mc}$ expressions show that only the subset of integrals with all occupied orbital indices are required for these elements. If there are o occupied orbitals and n total orbitals, then this two-electron integral subset requires about $(({}^{1}/_{2})on^4 + ({}^{1}/_{4})o^2n^3 + ({}^{1}/_{12})o^4n)$ effort to compute using the normal four-step integral transformation algorithm compared to about $4n^5$, $({}^{29}/_{24})n^5$, or $({}^{25}/_{24})n^5$ effort for the full transformation, depending on the treatment of the orbital index permutation symmetry.[362] Only integrals with at most one virtual index are required for the orbital gradient f^{mc}_{pq} and the orbital-CSF Hessian elements $G^{mc}_{pq,j}$; these integrals require an additional $(({}^{1}/_{6})o^3n^2 + ({}^{5}/_{12})o^4n)$ effort. Finally, integrals with at most two virtual indices are required for the orbital−orbital Hessian elements $G^{mc}_{pq,rs}$; these integrals require about $(on^4 + ({}^{1}/_{2})o^2n^3 - ({}^{1}/_{6})o^4n)$ effort, which is about a factor of 2 larger than that required for the energy and gradient terms. This shows that substantial effort can be eliminated by restricting the two-electron integral transformation to compute only the required subset at any time. There is a substantial history within the computational chemistry community related to efficient computation of various subsets of the two-electron integrals, and it is noted here only that similar orbital index subset restrictions apply to many other electronic structure methods (e.g., the B_k approximation and various second-order PT methods).

The orbital−orbital Hessian elements $G^{mc}_{pq,rs}$ in eq 84 are seen to consist of terms computed from the Fock matrix elements **F**, which in turn require only the zero- and one-virtual integral subsets, and only the terms in the last summation require the two-virtual integral subset. As discussed above, the computation of these latter integrals requires about twice the computational effort of the smaller subset. If the orbital−orbital Hessian matrix is approximated with only the terms involving the **F** matrix elements, and if that approximation is sufficient to allow convergence with less than twice the number of iterations of the full second-order method, then the overall effort for the integral transformation steps would be less than that for the full second-order procedure. This approximation, along with the total neglect of the orbital−CSF and CSF−CSF blocks of the Hessian matrices, is used in the CASSCF implementation in MOLCAS.[123] These approximations make it problematic to converge to excited states—the orbital variations tend to cause variational

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

collapse of the excited states, resulting in oscillatory convergence and other difficulties. However, other levels of approximation, which include these missing Hessian matrix blocks while still using only the zero- and one-virtual integral subsets, could be implemented for these kinds of excited-state calculations. Chaban et al.[363] demonstrated the convergence characteristics of several MCSCF approaches using a diagonal Hessian approximation. Other approximations used in MCSCF calculations involve approximations to the molecular integrals. These have been done with pseudospectral methods,[304,306] RI methods,[61] and the use of incomplete Cholesky factorizations.[53−55]

In many situations, the MCSCF energies and wave functions themselves are only of secondary importance; the primary goal is the computation of the orbitals which are then used in a subsequent MRCI calculation. These orbitals are used to describe simultaneously several electronic states, sometimes with the same spin values S, and sometimes with different S values (see eq 10). However, the optimization of the MCSCF orbitals for a particular state sometimes causes the subsequent MRCI calculation to be biased. For example, excitation energies from lower states might be systematically underestimated while excitation energies to higher states are overestimated, or other properties, such as dipole moments or $\langle r^2 \rangle$ expectation values might be biased by the orbital choice. To avoid this bias, it is typical to optimize the MCSCF orbitals to minimize a weighted average of the states of interest.[364,214,219,365]

$$\bar{E} = \sum_{k=1}^{N_{av}} w_k E_k \tag{89}$$

The optimization of the orbitals is done through the minimization.

$$\frac{\partial \bar{E}}{\partial \boldsymbol{\lambda}} = \mathbf{0} = \sum_{k=1}^{N_{av}} w_k \frac{\partial E_k}{\partial \boldsymbol{\lambda}} \tag{90}$$

with $\Sigma_k w_k = 1$. The individual states in the averaging procedure are described with optimal CSF coefficients $\mathbf{c}^k$, with each vector k satisfying the eigenvalue equation. It is only the variation of the orbitals, which are shared by all of the states in the averaging procedure, that is described with a single set of parameters $\mathbf{k}$. This state-average optimization also addresses the problematic convergence of first-order convergent MCSCF optimizations for excited states; if all lower states are included in the averaging procedure, then variational collapse during the excited-state orbital optimization process can be avoided.[219,364,214] Unlike the single-state optimization case, the individual gradients $\partial E_k/\partial \mathbf{k}$ are generally nonzero; it is only the weighted average that is zero at convergence of the optimization procedure. If the individually optimized orbitals for the individual states are similar, then the individual gradients from the state-averaged procedure are small, but if there is strong competition among the states for the orbitals to have very different character, then the individual gradients from the state-averaged procedure are large. If the individual gradients are too large, then the state-averaging procedure is inappropriate. In this case, additional CSFs should be added to the expansion space, or the active orbital space should be increased, in order for the wave function to have sufficient flexibility. This situation can be detected by monitoring the individual gradients during the optimization procedure and for the final converged orbitals.

2.2.3. MCSCF Wave Function Expansions. One of the important differences between MCSCF and MRCI methods is the choice of expansion space. MCSCF expansion spaces are discussed in detail in ref 19, and only some of the important features are mentioned here. The MCSCF expansion space is chosen typically to describe the important valence correlation effects. In a PES calculation, for example, the relevant bond-breaking and spin-recoupling effects should be described well by the MCSCF expansion. In a FORS/CAS expansion,[219,364,366] for example, the important valence orbitals and electrons would be identified and included in the active orbital list, and then all possible CSFs (i.e., all possible occupations and spin couplings) with those active orbitals and electrons would be included in the MCSCF expansion. There are several important features of this expansion form. One is that all active−active orbital rotations are redundant, which means that all orbital parameters of the type k_{pq} for active orbital indices p and q can be ignored during the optimization. This means that one block of the orbital gradient vector and four blocks of the orbital−orbital Hessian matrix can be ignored, which simplifies the implementation considerably. Another desirable feature is that such expansions are size-consistent. Separate calculations on molecular fragments A and B are consistent with the calculation on the combined molecule AB with noninteracting fragments provided the AB active orbital space is the union of the A and B active orbital spaces; in such cases, the wave functions satisfy the multiplication property $|\psi^{AB}\rangle = |\psi^{A} \otimes \psi^{B}\rangle$, and the energies satisfy the sum property $E_{AB} = E_A + E_B$. Furthermore, if the active spaces are chosen appropriately, these relations hold even, for example, for closed-shell molecules that dissociate to open-shell fragments and also for other types of spin-recoupling processes.

For larger molecules, not all valence orbitals and electrons can be included in FORS/CAS expansions due to the large expansion dimensions, and a selection process is required. Ignoring any simplifications due to point group symmetry, the number of expansion terms for such an expansion is given by[1]

$$N_{\text{det}}^{\text{full-CI}} = \binom{n}{N_\alpha}\binom{n}{N_\beta} \tag{91}$$

$$N_{\text{csf}}^{\text{full-CI}} = \frac{2S+1}{n+1}\binom{n+1}{\frac{1}{2}N - S}\binom{n+1}{\frac{1}{2}N + S + 1} \tag{92}$$

for expansions in terms of Slater determinants and CSFs, respectively. In these expressions, N and n refer only to the active electrons and orbitals, respectively, not to those for the full molecule. It is convenient to denote such an expansion space as $(\mathbf{n}^N)$, in analogy to the familiar atomic shell notation, meaning N electrons distributed in n orbitals. Another common notation for such an expansion is CAS(N,n); the more compact notation will be used in this discussion. Practically speaking, such FORS/CAS expansions are limited to about $N \approx 16$ electrons and $n \approx 16$, for which $N_{\text{det}}(\mathbf{16}^{16}) = 1.66 \times 10^8$ and $N_{\text{csf}}(\mathbf{16}^{16}) = 3.48 \times 10^7$ for $M = 0$ and $S = 0$, respectively, according to eqs 10 and 11. Although even larger MCSCF calculations are possible, it is then extremely difficult to do any kind of subsequent MRCI calculation based on these larger reference expansion spaces. Consequently, additional orbital occupation restrictions are typically imposed in order to reduce the expansion length. Two of the more commonly encountered types of occupation restrictions

will be discussed here. A wide range of other restricted expansion spaces are discussed in ref 19.

One type of occupation restriction that is imposed on MCSCF expansions is termed a direct-product expansion.[19] With this approach, the active orbital space is divided into an arbitrary number of disjoint subsets, and each of these orbital subspaces is associated with a fixed number of electrons. Within each subspace, all possible orbital occupations are allowed, and the final expansion space is then the direct product of all combinations of all these occupations. In the occupation restricted multiple active space (ORMAS) approach of Ivanic[367,368] this is termed an ORMAS0 expansion. There are two important features of direct-product expansions. The first is that the energies are rigorously size-consistent provided the orbital subsets of the fragments are taken also as the orbital subsets of the molecule. This follows from the product nature of the expansion space. The second important feature is that the orbital rotations within each orbital subset are redundant. The orbitals within a subset form an *invariant subspace*. This means practically that these rotations should be ignored in setting up the orbital-rotation vector $\mathbf{k}$, and it also means that any arbitrary choice of orbital rotations within each of the product spaces may be chosen for the final optimized orbitals. This last step is called *orbital resolution*, and it may be imposed to facilitate analysis of the wave function or to simplify some aspect of a subsequent MRCI calculation. Orbital resolutions based on spatial localization (both with and without orthogonalization constraints), natural orbitals, or canonical orbitals based on the diagonalization of Fock matrices are common. In all of these situations, a particular orbital rotation can be classified as either essential (it is required in order to minimize the energy) or redundant (it has no effect on the wave function or energy, provided the CSF coefficients are allowed to adjust accordingly). This essential/redundant structure of the orbital rotation parameters is discussed in more detail in ref 21.

For an example of a direct-product expansion, suppose that 16 active orbitals are divided into eight pairs and two electrons are associated with each orbital pair. Each individual pair would then have three possible occupations: φ_1^2, $\varphi_1\varphi_2$, and φ_2^2 which can be denoted by the full-CI designation $\mathbf{2}^2$. The direct product of the eight subspaces, denoted $\mathbf{2}^2\mathbf{2}^2\mathbf{2}^2\mathbf{2}^2\mathbf{2}^2\mathbf{2}^2\mathbf{2}^2\mathbf{2}^2$ would then consist of $3^8 = 6561$ possible orbital occupations. Each of the open-shell occupations in a direct-product expansion can have multiple spin couplings, and in this particular example the 6561 orbital occupations result in 71 398 singlet CSFs, which is several orders of magnitude smaller than the $\mathbf{16}^{16}$ expansion discussed above for this same active orbital space. Note that although there are severe occupation restrictions imposed on this expansion space, the overall excitation level relative to a single configuration is not artificially limited; in this particular example, there are 16-fold excitations in the expansion space, the same as for the $\mathbf{16}^{16}$ full-CI expansion. This particular type of direct-product expansion, in which pairs of electrons are constrained to occupy pairs of orbitals, is called the GVB-RCI expansion, and a general expression for the CSF expansion dimension for singlets is

$$N_{\mathrm{csf}}^{\mathrm{GVB-RCI}} = \sum_{k=0}^{n/2} \frac{2^{n/2-k}}{k+1} \binom{2k}{k} \binom{n/2}{k} \tag{93}$$

The expansion dimension in terms of determinants is the same as eq 93 except that the $(k + 1)$ factor in the denominator is dropped.[19] The GVB-RCI expansion is appealing because the optimal orbitals tend to localize into chemically intuitive bonding and antibonding pairs. Given an arbitrary molecule, it is usually straightforward to construct the GVB-RCI expansion along with a reasonable set of initial orbitals from the Lewis dot structure for the molecule, and this expansion is capable of breaking single and multiple bonds and dissociating correctly in many situations to the correct high-spin or low-spin molecular fragments. Due to the direct-product construction of the expansion space, the resulting wave functions and energies are rigorously size-consistent. Another useful feature of GVB-RCI expansions, and direct-product expansions in general, is that the two-particle density matrix is relatively sparse; the nonzero elements d_{pqrs} have orbital indices in which either all correspond to a single invariant subspace or two indices belong to one subspace and the other two indices belong to another subspace. In the GVB-RCI case, there are only about $(^{19}/_{2})n^2$ unique nonzero elements compared to about $(^{1}/_{8})n^4$ for the general full-CI expansion. Because of these features, the GVB-RCI expansion often results in an excellent reference expansion space for subsequent MRCI calculations. However, excited states and sometimes even ground states are not described adequately by the GVB-RCI expansion. Three representative examples of failures are the benzene molecule for which the GVB-RCI expansion computes a symmetry-broken wave function with localized double bonds rather than the delocalized aromatic structure, the CO_2 molecule for which a symmetry-broken D_{2d} wave function is computed rather than the correct $D_{\infty h}$ wave function, and the O_3 molecule for which important interpair correlations between the lone-pair π orbital and the two open-shell π orbitals are poorly described. In all of these cases, a generalization of the occupations within the direct-product expansion form solves these issues but at the expense of larger expansion spaces. In the benzene case the three π orbital products $\mathbf{2}^2\mathbf{2}^2\mathbf{2}^2$ can be replaced with the larger $\mathbf{6}^6$ expansion, in the CO_2 case the $1\pi_x^2(2\pi_x3\pi_x)^2(1\pi_y2\pi_y)^2 3\pi_y^2$ direct product can be replaced with the $(1\pi_x2\pi_x3\pi_x)^4$-$(1\pi_y2\pi_y3\pi_y)^4$ direct product, and in the O_3 case the $1\pi^2(2\pi3\pi)^2$ product can be replaced with the $(1\pi2\pi3\pi)^4$ subspace. For a more complete discussion of these issues with application to several molecules, see ref 369.

Another type of occupation restriction is the restricted active space (RAS).[370] In this approach, the active orbitals are divided into subsets, labeled I, II, and III. A minimum total occupation is imposed on subspace I, a maximum occupation is imposed on subspace III, and subspace II is allowed any occupation consistent with the total number of electrons. In general, this expansion space is equivalent to the union of several direct-product expansion spaces, each with the same orbital subset partitioning but with different numbers of subspace electrons. If $\mathbf{n}_X$ corresponds to the orbitals in each of the three subsets, $X = \{\mathrm{I,II,III}\}$ and if $N_X^{\min}$ and $N_X^{\max}$ are the corresponding minimum and maximum occupations for $X = \{\mathrm{I,III}\}$, then the RAS expansion space may be written symbolically as

$$\{RAS\} = \bigcup_{j=N_I^{\min}}^{N_I^{\max}} \bigcup_{l=N_{III}^{\min}}^{N_{III}^{\max}} \mathbf{n}_{\mathrm{I}}^{j}\mathbf{n}_{\mathrm{II}}^{k}\mathbf{n}_{\mathrm{III}}^{l}; \quad j+k+l=N \text{ and } k \geq 0 \tag{94}$$

The orbital subspace invariance properties for the RAS expansion follow from those of the component direct-product expansions. Specifically, each subspace $X = \{\mathrm{I,II,III}\}$ is invariant because it is invariant in each of the individual direct-product expansions in

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

eq 94. For example, for the $N = n = 16$ expansion used in the previous discussions, suppose the first four orbitals, subspace I, are required to have at least four electrons (and no more than eight), the last four orbitals, subspace III, are allowed to have no more than four electrons, and the second group of eight orbitals, subspace II, is occupied by the complement of electrons so that the total in each CSF is 16. The RAS expansion then consists of the union of 25 direct-product expansions of the form $\mathbf{4}^j\mathbf{8}^k\mathbf{4}^l$ for the five values of $j = 4...8$, the five values of $l = 0...4$, and with $k = 16 - j - l$. This expansion consists of 16 462 550 CSFs or 79 342 342 determinants, about half of that for the $\mathbf{16}^{16}$ expansion. For this RAS expansion, the six orbital rotations within subspace I are redundant, the 28 rotations within subspace II are redundant, and the six rotations within subspace III are redundant. All other orbital rotations are essential. The ORMAS expansion implemented by Ivanic[367] may be regarded as a generalization of the RAS expansion approach to allow an arbitrary number of orbital subsets, with each subset associated with its own minimum and maximum occupation. This approach is also termed the generalized active space (GAS) expansion.[56,635] The ORMAS/GAS expansion is also a union of multiple direct-product expansions, and therefore the orbital rotations within each orbital subspace are redundant. One of the practical features of RAS expansions is that the coupling coefficients can be generated as products of factors that are computed separately for the three orbital subspaces.[370] Unlike the sparse nature of the two-particle density matrix for direct-product expansions, the RAS/ORMAS/GAS expansion form generally results in a dense 2-RDM.

Size-consistency can be achieved with the RAS/ORMAS/GAS expansion approach in certain special situations, but in general this expansion form is not size-consistent. To demonstrate this, consider a simple example with RAS expansions for fragments A and B compared to the combined RAS-AB expansion for the molecule. Suppose fragment A has four electrons and has subspaces I_A, with two orbitals $\{\varphi_{1A}\varphi_{2A}\}$ occupied by at least two electrons, and III_A, with two orbitals $\{\varphi_{3A}\varphi_{4A}\}$ occupied by no more than two electrons. This RAS expansion is equivalent to an SRCI-SD expansion from the configuration $|\varphi^2_{1A}\varphi^2_{2A}\rangle$. Similarly, let fragment B with four electrons have subspaces I_B, with two orbitals $\{\varphi_{1B}\varphi_{2B}\}$ occupied by at least two electrons, and III_B, with two orbitals $\{\varphi_{3B}\varphi_{4B}\}$ occupied by no more than two electrons. To include all possible product configurations in the RAS-AB expansion, the I_{AB} subspace must consist of four orbitals $\{\varphi_{1A}\varphi_{2A}\varphi_{1B}\varphi_{2B}\}$ occupied by at least four electrons, and the III_{AB} subspace must consist of four orbitals $\{\varphi_{3A}\varphi_{4A}\varphi_{3B}\varphi_{4B}\}$ occupied by no more than four electrons. This corresponds to the SRCI-SDTQ expansion from the configuration $|\varphi^2_{1A}\varphi^2_{2A}\varphi^2_{1B}\varphi^2_{2B}\rangle$. The inconsistency arises by noting that a configuration such as $|\varphi^2_{1A}\varphi^2_{2A}\varphi^2_{3B}\varphi^2_{4B}\rangle$, which corresponds to a quadruple excitation on fragment B, is included in this RAS-AB expansion space, but it does not appear in the product of the fragment RAS expansions. Thus this RAS-AB expansion space is more flexible than it should be, the AB wave function does not satisfy the product requirement, and $E_{AB} < E_A + E_B$. On the other hand, if the I_{AB} subspace is required to have at least six electrons, and the III_{AB} subspace is restricted to have no more than two electrons, then there would be configurations in the product expansion that would not be included in this RAS-AB expansion, this expansion space would be insufficiently flexible, and $E_{AB} > E_A + E_B$. This is an example of the usual size-consistency problem for SRCI-SD expansions. There is no choice of occupation limits for the I_{AB} and III_{AB} orbital subspaces that would result in exact size-consistency. The lack of size-consistency for RAS/ORMAS/GAS expansions has been discussed and demonstrated by Ivanic for the $2NO_2 \rightarrow N_2O_4$ reaction.[368]

2.2.4. Computing the Matrix Exponential. Given the skew-symmetric matrix $\mathbf{K}$ from an iteration of the MCSCF procedure, it is necessary to compute the orthogonal matrix $\mathbf{U} = \exp(\mathbf{K})$. There are several approaches to compute the matrix exponential.[371,372] Errors may be tolerated in some situations in the computation of $\mathbf{U}$ provided they are sufficiently small (e.g., $\lesssim O(|\mathbf{K}|^3)$ for a second-order convergent wave function optimization procedure), but it is essential that the final $\mathbf{U}$ be orthogonal to within a small factor of the machine precision. In other situations, the elements of $\mathbf{U}$ must be accurate to within some tolerance regardless of the size of the matrix elements $\mathbf{K}$. Therefore the goal is an efficient procedure that satisfies both kinds of accuracy criteria. One obvious approach to compute $\mathbf{U}$ is to truncate the Taylor expansion in eq 63 at some value m. This may be done recursively with the sequence,

$$\left.\begin{aligned} \mathbf{U}^{(0)} &= 1; \quad \mathbf{X}^{(1)} = \mathbf{K} \\ \mathbf{U}^{(j)} &= \mathbf{U}^{(j-1)} + \mathbf{X}^{(j)} \\ \mathbf{X}^{(j+1)} &= \frac{1}{(j+1)}\mathbf{K}\mathbf{X}^{(j)} \end{aligned}\right\} j = 1 \ \ldots \ m \tag{95}$$

Either eq 95 may be implemented with a predetermined truncation order m or m may be determined dynamically by monitoring $||\mathbf{X}^{(j)}||$. This is a relatively expensive algorithm since each step requires about $2n^3$ arithmetic operations for the matrix product, ignoring any sparsity in the matrix $\mathbf{K}$. With termination after m steps, the matrix $\mathbf{U} = \mathbf{U}^{(m)}$ is not strictly orthogonal due to the truncation error, so a subsequent orthonormalization step requires an additional $2n^3$ operation.

The identity $\exp(\mathbf{K}) = (\exp(\mathbf{K}/k))^k$ may be used to compute the matrix exponential with the scaling and squaring method.[373] The matrix $\mathbf{Z}$ is initialized as $\mathbf{Z} = \mathbf{K}/2^p$ for some suitably large value of p for which $\exp(\mathbf{Z})$ may be computed accurately with a small value of m in eq 95. $\mathbf{U}$ is then computed with p recursive squaring steps. For a specified error tolerance, this approach is usually cheaper than the straightforward truncated Taylor expansion. Either the truncated $\exp(\mathbf{Z})$ must be orthogonalized or the resulting matrix $\mathbf{U}$ must be orthogonalized afterward because the truncation error accumulates with each recursive squaring step.[374]

A closely related approach is based on the rational function approximation

$$\mathbf{U}^{(k/m)} = q_{km}(\mathbf{K})^{-1}p_{km}(\mathbf{K}) \tag{96}$$

Here $p_{km}(\mathbf{K})$ and $q_{km}(\mathbf{K})$ are (commuting) polynomials of degree k and m, respectively. In addition to the recursive effort for the matrix powers for the two polynomials, the linear equation solutions in eq 96 for $\mathbf{U}^{(k/m)}$ requires about the same effort as two matrix multiplications. If the expansion coefficients in the two polynomials are chosen to reproduce the first $k + m$ Taylor expansion coefficients of the exponential, this is a Padé approximant and the polynomial coefficients have known closed-form expressions.[373] A simple example is $\mathbf{U}^{(1/1)} = (\mathbf{1} - (1/2)\mathbf{K})^{-1}(\mathbf{1} + (1/2)\mathbf{K})$ which is correct through second order. Although $\mathbf{U}^{(1/1)}$ is also only an approximation and suffers from truncation error, it is orthogonal in exact arithmetic. Thus no subsequent orthogonalization step is required unless the

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

condition number of the matrix $(\mathbf{1} - (^1/_2)\mathbf{K})$ is exceedingly large. The condition number of a matrix is the ratio of the largest and smallest singular values, and large values can result in large numerical errors in linear and eigenvector equations when using finite-precision floating point arithmetic. Other more general $p_{mm}(\mathbf{K})$ and $q_{mm}(\mathbf{K})$ diagonal polynomials may also be chosen to satisfy this orthogonal relation. This rational function approximation can be combined with recursive squaring.[373] For a given error tolerance, lower order k and m may be used, and this can result in overall less effort. This initial scaling also reduces the impact of error growth due to the condition number of the $q_{km}(\mathbf{K})$ matrix. In those situations for which $\mathbf{U}^{(k/m)}$ in the initial squaring step is orthogonal, the final computed matrix is also orthogonal. Otherwise, an explicit orthogonalization either before or after the recursive squaring may be necessary.

The matrix $\mathbf{K}$ is normal (it commutes with its transpose), so it may be diagonalized by a complex unitary matrix $\mathbf{V}$. It has purely imaginary eigenvalues, the nonzero elements of which occur in complex conjugate pairs.

$$\mathbf{KV} = \mathbf{V}i\boldsymbol{\lambda} \tag{97}$$

This allows the matrix exponential to be computed as

$$\mathbf{U} = \exp(\mathbf{K}) = \mathbf{V}\exp(i\boldsymbol{\lambda})\mathbf{V}^\dagger = \mathbf{V}(\cos(\boldsymbol{\lambda}) + i\sin(\boldsymbol{\lambda}))\mathbf{V}^\dagger \tag{98}$$

The result should be purely real, but in finite precision arithmetic there is always some small imaginary error that must be discarded. Furthermore, the complex diagonalization costs about four times that of a comparable real symmetric diagonalization. In a diagonalization-based approach, the full eigenpair spectrum is required, so direct methods which require $O(n^3)$ effort are typically employed.

The negative semidefinite symmetric matrix $\mathbf{K}^2$ may be diagonalized with a real orthogonal matrix $\mathbf{X}$

$$\mathbf{K}^2\mathbf{X} = \mathbf{X}\boldsymbol{\lambda} \tag{99}$$

Using the purely real diagonal matrix $\mathbf{d} = \mathrm{Sqrt}(-\boldsymbol{\lambda})$, the matrix exponential can be computed[358,375] as

$$\mathbf{U} = \exp(\mathbf{K}) = \mathbf{X}\cos(\mathbf{d})\mathbf{X}^{\mathrm{T}} + \mathbf{K}\mathbf{X}\mathbf{d}^{-1}\sin(\mathbf{d})\mathbf{X}^{\mathrm{T}} \tag{100}$$

Zero diagonal elements are treated with $\lim_{x\to 0}\sin(x)/x = 1$. This expression has the advantage that the entire operation involves only real arithmetic, and although it has no truncation error, it suffers from numerical error due to the larger condition number of the matrix $\mathbf{K}^2$ relative to that of the matrix $\mathbf{K}$. Consequently a subsequent orthonormalization step is sometimes required. In addition to the effort for the real symmetric diagonalization, effort for three matrix products is also required.

The final approach relies on the fact that the matrix $\mathbf{K}$ may be factored in the form

$$\mathbf{K} = \mathbf{WDW}^{\mathrm{T}} \tag{101}$$

with real orthogonal $\mathbf{W}$ and a skew-symmetric block diagonal matrix $\mathbf{D}$. The diagonal subblocks are either 2×2 or 1×1 with all diagonal elements zero. The factorization in eq 101 is described in TOMS Algorithm 530 by Ward and Gray,[376] and the corresponding software is available from netlib.[377] The matrix exponential may then be computed as $\exp(\mathbf{K}) = \mathbf{W}\exp(\mathbf{D})\mathbf{W}^{\mathrm{T}}$, where the 2×2 subblocks of the exponential matrix are given as

$$\exp\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{102}$$

The entire procedure consists of sequences of products of orthogonal transformations, so there is no growth of roundoff errors. This is generally the recommended approach for computing the accurate matrix exponential for the orbital transformation. It requires only real arithmetic, it only requires effort comparable to a single real symmetric matrix diagonalization plus a single matrix product, the algorithm has no truncation error, and the resulting matrix $\mathbf{U}$ is orthogonal to within a small factor of the machine precision.

2.3. New Multireference Approaches

Several new methods that strive to address the shortcomings of MRCI have been developed recently. Two such methods are the canonical transformation (CT) theory developed by Yanai and Chan[378,379] and the anti-Hermitian contracted Schrödinger equation (ACSE) developed by Mazziotti.[380−383]

2.3.1. Density-Based Approach to Dynamical Correlation. Although other density-based methods related to the ACSE have been developed,[384−388] the treatment of dynamic correlation within the context of the ACSE is discussed here. Consider the density-matrix formulation of the Schrödinger equation

$$\hat{H}^N D = E^N D \tag{103}$$

where $\hat{H}$ is the Hamiltonian, ${}^N D = \Psi\Psi^*$ is the N-electron density, and E is the energy. Contraction of eq 103 onto the space of two particles yields the contracted Schrödinger equation (CSE), which in second-quantized notation takes the form

$$\langle\Psi|{}^2\hat{\Gamma}^{i,j}_{k,l}\hat{H}|\Psi\rangle = E^2 D^{i,j}_{k,l} \tag{104}$$

In eq 104, ${}^2\hat{\Gamma}$ is the two-electron reduced density operator (RDO), and the elements of the two-particle reduced density matrix (2-RDM)

$${}^2D^{i,j}_{k,l} = \frac{1}{2}\langle\Psi|{}^2\hat{\Gamma}^{i,j}_{k,l}|\Psi\rangle = \frac{1}{2}\langle\Psi|\hat{a}^\dagger_i\hat{a}^\dagger_j\hat{a}_l\hat{a}_k|\Psi\rangle \tag{105}$$

are defined here such that the trace of the 2-RDM is $N(N-1)/2$. Because the Hamiltonian contains at most two-electron interactions, it can be verified that the left-hand side of the CSE in eq 104 depends not only on the 2-RDM but also on the 3- and 4-RDMs. As a result, until recently work in the CSE community has focused on developing accurate approximations to the 3- and 4-RDMs in terms of the 1- and 2-RDMs.

An alternative approach is obtained when the CSE in eq 104 is expressed as a sum of its Hermitian and anti-Hermitian components

$$\langle\Psi|[{}^2\hat{\Gamma}^{i,j}_{k,l}, (\hat{H}-E)]_+|\Psi\rangle + \langle\Psi|[{}^2\hat{\Gamma}^{i,j}_{k,l}, \hat{H}]|\Psi\rangle = 0 \tag{106}$$

in which $[\ldots]$ and $[\ldots]_+$ denote the commutator and anticommutator, respectively. For the CSE to be satisfied, both terms in eq 106 must equal zero separately, defining the ACSE[380−382,389−395] (also called the k-particle Brillouin conditions)

$$\langle\Psi|[{}^2\hat{\Gamma}^{i,j}_{k,l}, \hat{H}]|\Psi\rangle = 0 \tag{107}$$

In contrast to the CSE, the ACSE only depends on the 1-, 2-, and 3-RDMs since the commutator of two tensors with rank m and n, respectively, is a tensor of rank $(m + n - 1)$.

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

2.3.2. Solution of the ACSE Equations. To solve the ACSE for the 2-RDO and hence the energy, Mazziotti[380−382] considers a sequence of unitary two-body transformations applied to a reference wave function $\Psi(\lambda)$, where the unitary transformations are ordered according to a continuous time-like variable λ

$$|\Psi(\lambda + \epsilon)\rangle = e^{\epsilon\hat{S}(\lambda)}|\Psi(\lambda)\rangle \tag{108}$$

where the operator

$$\hat{S}(\lambda) = \sum_{i,j,k,l} {}^2S^{i,j}_{k,l}(\lambda)\cdot{}^2\hat{\Gamma}^{i,j}_{k,l} \tag{109}$$

is restricted to be antihermitian ($\hat{S}^\dagger = -\hat{S}$) in order to ensure the unitarity of the sequence of transformations. Inserting eq 108 into the definition of the ACSE in eq 107 and expanding the exponentials up to first order in ϵ yields

$$E(\lambda + \epsilon) = E(\lambda) + \epsilon\langle\Psi(\lambda)|[\hat{H}, \hat{S}(\lambda)]|\Psi(\lambda)\rangle + \ldots \tag{110}$$

Taking the $\epsilon \rightarrow 0$ limit yields the following differential equation for the energy

$$\frac{dE}{d\lambda} = \langle\Psi(\lambda)|[\hat{H}, \hat{S}(\lambda)]|\Psi(\lambda)\rangle \tag{111}$$

and a similar equation for the evolution of the 2-RDM is obtained when the Hamiltonian in the ACSE is replaced with the 2-RDO

$$\frac{d^2D^{i,j}_{k,l}}{d\lambda} = \langle\Psi(\lambda)|[{}^2\hat{\Gamma}^{i,j}_{k,l}, \hat{S}(\lambda)]|\Psi(\lambda)\rangle \tag{112}$$

Inserting the definition of $\hat{S}(\lambda)$ in eq 109 into eq 110 and differentiating, the elements of ${}^2\mathbf{S}$ are chosen in order to minimize the energy along the gradient

$${}^2S^{i,j}_{k,l}(\lambda) = -\frac{1}{\epsilon}\frac{\partial E(\lambda + \epsilon)}{\partial({}^2S^{i,j}_{k,l}(\lambda))}\bigg|_{{}^2S=0} = \langle\Psi(\lambda)|[{}^2\hat{\Gamma}^{i,j}_{k,l}, \hat{H}]|\Psi(\lambda)\rangle \tag{113}$$

Although the differential eqs 111−113 formally employ the wave function, they can be expressed in terms of the 1-, 2-, and 3-RDMs (because they all involve the commutator of two rank-2 tensors). Provided there is a suitable initial guess for the density matrices at $\lambda = 0$, the differential equations are propagated until either (i) the energy or (ii) the least-squares error of the ACSE increases. Technically, the ACSE (the right-hand side of eq 113) should vanish upon convergence; however, due to the approximations related to the reconstruction of the 3-RDM (see below), the ACSE equations are evolved until either the energy or the least-squares error norm of the ACSE increases.

Note that eqs 111−113, in addition to the 2-RDM, also depend on the 3-RDM. As such, the ACSE equations are indeterminate. To remove this indeterminacy, the 3-RDM is approximated with its cumulant expansion

$$\begin{aligned}{}^3D^{i,j,k}_{l,m,n} &= {}^1D^i_l \wedge {}^1D^j_m \wedge {}^1D^k_n \\ &+ 3({}^2D^{i,j}_{l,m} - {}^1D^i_l \wedge {}^1D^j_m) \wedge {}^1D^k_n + {}^3\Delta^{i,j,k}_{l,m,n}\end{aligned} \tag{114}$$

where $\wedge$ denotes the antisymmetrized tensor product, and the cumulant (connected) part of the 3-RDM (${}^3\mathbf{\Delta}$) vanishes if all three electrons are statistically independent. There are several approximations to the connected part of the 3-RDM. The simplest approximation, also known as the Valdemoro reconstruction,[396] is to simply neglect the three-electron cumulant (${}^3\mathbf{\Delta} = \mathbf{0}$). As argued in ref 383, this simple approximation yields energies with sufficient accuracy. More elaborate reconstruction functionals, such as the Nakatsuji−Yasuda[397] and the Mazziotti functional,[398,399] can lead to numerical difficulties when solving the ACSE since the elements of the connected 3-RDM can become large when all six orbitals are in the active space.[383] Because of these considerations, the ACSE neglects ${}^3\mathbf{\Delta}$ when treating multireference problems.[383]

In applications to multireference problems, the initial guess for the ACSE is generated from a CASSCF wave function. For this reason, these methods may be considered multireference methods. This limits the applicability of the ACSE to applications with small active spaces, but in principle this could be circumvented by using one of the new methods (e.g., GCF, DMRG, or the active-space variational RDM[400−404] method) to compute the initial density. Nonetheless, the working assumptions in applying the ACSE to multireference problems are that (i) the reference (CASSCF) 2-RDM captures the effects of static correlations, (ii) that unitary transformations among the active orbitals can be neglected, and (iii) the two-body operator $\hat{S}(\lambda)$ neglects the terms with more than two virtual orbitals. These assumptions correspond to those of other multireference methods. Some recent applications of ACSE include geometry optimizations,[405,406] reaction barriers,[407,408] sigmatropic shifts,[409] excited states,[410] open-shell systems,[411] and conical intersections.[412]

2.3.3. Canonical Transformation Theory. To incorporate the effects of dynamic correlation on top of a multiconfigurational reference wave function $|\psi_0\rangle$, the canonical transformation theory of Yanai and Chan[378,379] expresses the wave function based on the unitary exponential ansatz

$$|\Psi_{CT}\rangle = e^{\hat{A}}|\psi_0\rangle \tag{115}$$

When the antisymmetric excitation operator $\hat{A}$ contains at most single- and double-replacement operators

$$\begin{aligned}\hat{A} = &\sum_{i_1,s_1} A^{s_1}_{i_1}(\hat{a}^\dagger_{s_1}\hat{a}_{i_1} - \hat{a}^\dagger_{i_1}\hat{a}_{s_1}) \\ &+ \sum_{i_1,i_1,s_1,s_2} A^{s_1,s_2}_{i_1,i_2}(\hat{a}^\dagger_{s_1}\hat{a}^\dagger_{s_2}\hat{a}_{i_2}\hat{a}_{i_1} - \hat{a}^\dagger_{i_1}\hat{a}^\dagger_{i_2}\hat{a}_{s_2}\hat{a}_{s_1})\end{aligned} \tag{116}$$

that only rotate between the active−external and external−external spaces,[378] the method is denoted CTSD. The orbital indices i and s in eq 116 denote arbitrary orbitals within the active and external spaces. As long as the reference wave function is size-consistent, the exponential parametrization of the wave function guarantees that CT is rigorously size-consistent.[378] Application of the unitary operator to the Hamiltonian

$$\bar{\hat{H}} = e^{-\hat{A}}\hat{H}e^{\hat{A}} = \hat{H} + [\hat{H}, \hat{A}] + \frac{1}{2}[[\hat{H}, \hat{A}], \hat{A}] + \ldots \tag{117}$$

generates the effective CT Hamiltonian $\bar{\hat{H}}$. The excitation operator amplitudes ($A^{s_1}_{i_1}$ and $A^{s_1,s_2}_{i_1,i_2}$) in eq 116 are determined by solving the generalized Brillouin conditions.[390] The energy is then computed as the expectation value of this effective Hamiltonian with

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

the reference wave function,

$$E = \langle\psi_0|\hat{\overline{H}}|\psi_0\rangle \tag{118}$$

In the current form, the computation in eq 118 is just as challenging as the original Schrödinger equation since, as can be seen from the commutator expansion in eq 117, the CT Hamiltonian contains, in addition to one- and two-electron operators, also three-electron and higher rank operators. However, as noted by Yanai and Chan,[378] this complexity may be eliminated by decomposing the three-body operators after each commutator as in linearized CTSD (LCTSD)

$$\hat{\overline{H}} = e^{-\hat{A}}\hat{H}e^{\hat{A}} = \hat{H} + [\hat{H},\hat{A}]_{1,2} + \frac{1}{2}[[\hat{H},\hat{A}]_{1,2},\hat{A}]_{1,2} + \ldots \tag{119}$$

When employing the generalized normal ordering of Mukherjee and Kutzelnigg,[413,414] the terms $[\hat{H},\hat{A}]_{1,2}$ only depend on one- and two-body operators and the 1-, 2-, and 3-RDMs of the reference wave function.[379] Unlike the energy defined in eq 118, this approximation results in an energy that is no longer bounded from below by the exact full-CI energy. As in the density-based approaches, the dependence on the 3-RDM may be eliminated via the cumulant theory, and the resulting approximate CT Hamiltonian only depends on one- and two-electron terms. In principle, the expansion still requires an infinite number of terms to be computed; however, 8−10 terms are sufficient[415] for a precision of 10^{-9} E_h.

As discussed in more detail in ref 379, the performance of single-reference LCTSD is expected to "perform intermediate between linearized CCSD and the full CCSD theories" at a computational scaling that is similar to SR-CCSD (n^6, where n is the number of orbitals). Note, however, that LCTSD yields potential energy surfaces that are similar in accuracy to those from other multireference approaches. The accuracy of CTSD may be improved, while retaining the computational cost of CCSD, by delaying the operator decompositions for as long as possible.[416] The resulting quadratic CTSD (QCTSD) Hamiltonian

$$\hat{\overline{H}} = e^{-\hat{A}}\hat{H}e^{\hat{A}} = \hat{H} + [\hat{H},\hat{A}]_{1,2} + \frac{1}{2!}[[\hat{H},\hat{A}],\hat{A}]_{1,2} + \frac{1}{3!}[[[\hat{H},\hat{A}],\hat{A}]_{1,2},\hat{A}]_{1,2} + \frac{1}{4!}[[[[\hat{H},\hat{A}],\hat{A}]_{1,2},\hat{A}],\hat{A}]_{1,2} + \ldots \tag{120}$$

is more complex as it requires the decomposition of four-body operators ($[[\hat{H},\hat{A}],\hat{A}]_{1,2}$ versus $[\hat{H},\hat{A}]_{1,2}$ in LCTSD). Nonetheless, the overall scaling of QCTSD with system size is the same as that of LCTSD, and QCTSD is accurate to the same order in perturbation theory as CCSD.[416]

Although the initial applications[378,379,416] have focused on assessing the accuracy of LCTSD and QCTSD relative to multireference approaches such as CASPT2, MR-CISD, MR-ACPF, and MR-AQCC, more recent work toward developing strongly contracted CTSD for eliminating intruder states,[417] in conjunction with a reference wave function from DMRG,[418] shows great promise. Since the computational cost of CTSD does not depend on the number of determinants in the reference wave function, the DMRG-CT method has been applied to a variety of problems that are out of the reach of standard multireference approaches: (i) the evaluation of total correlation energies for conjugated polyenes ($\mathbf{24}^{24}$ active space),[417] (ii) the isomerization of $[Cu_2O_2]^{2+}$ ($\mathbf{32}^{28}$ active space),[417] and (iii) singlet−triplet gap of free base porphirin ($\mathbf{24}^{26}$ active space).[415]

2.4. Basis Extrapolation and R12 Methods

The quality and flexibility of the orbital basis set affects the quality of the computed wave functions and molecular properties. Thus any inherent limitations in a particular orbital basis set are reflected in the quality of the computed properties. There are two common approaches to address orbital basis set limitations: methods that are based on extrapolations and methods that include explicit interelectronic interactions into the wave function.

One of the important features of modern, generally contracted, basis sets (e.g., the ANO[467,468] and the correlation-consistent[466] basis sets) is that they display systematic convergence of the energy and other molecular properties. These sequences of basis sets are often used to extrapolate a given property to the complete basis set (CBS) limit for a given method. These CBS limits may be used to obtain high-quality results, to measure the inherent error in a given electronic structure method, and also to assess the quality of the individual calculations that are used in the extrapolation procedure.

The basis set extrapolation can be justified by the observation that the correlation energy converges with the inverse third power of the angular momentum in the helium atom[419] and in other N-electron atoms.[420] On the other hand, it is also known that the uncorrelated energy (SCF) shows an exponential convergence pattern;[421,422] i.e., it converges much faster. Extrapolation schemes developed for SR theories use these facts and extrapolate the SCF and correlated energies separately.[423−425,436] For the SCF energy an exponential equation

$$E(X) = E(\mathrm{CBS}) + A\exp(-BX) \tag{121}$$

may be used, where X is the cardinal index (e.g., $X = 2$ for DZ, $X = 3$ for TZ, and so on). This requires energy values with three different basis sets to determine the three parameters. For the correlation energy an inverse cubic formula requiring two energy values can be used[423]

$$E(X) = E(\mathrm{CBS}) + AX^{-3} \tag{122}$$

Other formulas are also used, e.g., the mixed Gaussian expression by Peterson et al.[426]

$$E(X) = E(\mathrm{CBS}) + A\exp(-(X-1)) + B\exp(-(n-1)^2) \tag{123}$$

In the preceding expressions $E(\mathrm{CBS})$, A, B, and n are the adjustable parameters that are fitted to the $E(X)$ data points.

The apparent problem in MR calculations is that uncorrelated and correlated energy components cannot easily be distinguished. First, MR calculations are often used in situations when the one-determinant approximation fails; i.e., separating the uncorrelated part is not possible. However, a separation of the static (nondynamic) and the dynamic parts of the correlation might be attempted. It is expected that the static part will converge similarly to the SCF case, and the remaining part can be extrapolated with a "dynamic electron correlation" expression. In the case of excited states, even this distinction is problematic since the "uncorrelated" excited state is often not defined. The fundamental idea of separation of static and dynamic correlation is therefore nontrivial. A practical separation can be defined as the difference between the total and the reference energy. In case of a

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

large CAS reference, however, particularly one that includes extra-valence orbitals in the active space, the reference energy would include already some portion of the dynamic correlation energy, and the convergence with respect to the basis set size might not be described well by the exponential formula. Therefore there is no unique and unambiguous basis set extrapolation recipe for MR calculations.

Several suggestions can be found in the literature to address this problem. Petersson et al.[427] observed that the convergence pattern of the CAS energy is very similar to that of the UHF energy, and they employ the UHF CBS limit to get the extrapolated CAS energy. The advantage of this procedure is that the larger basis set calculations are only performed at the UHF level resulting in some saving in computer time. In a subsequent paper[428] a similar procedure was used to extrapolate the MR-CISD energy; here the UHF-CCSD energy calculated with larger basis sets are used to extrapolate the MR-CISD correlation energy. Unfortunately if UHF fails to describe the system (e.g., due to strong static correlation, or spin contamination, or for excited-state wave functions), this approach is inappropriate. Nevertheless, the procedure was successfully used to describe several states of N_2.[428] Jiang and Wilson[429] tested several combinations of extrapolation formulas, including exponential type extrapolations of the CAS energy and subsequent extrapolations (again using different formulas) for the dynamic correlation. They also considered the possibility of extrapolating the total energy according to the eq 123 above. Earlier Müller et al.[655] successfully used the inverse cubic formula eq 122 to extrapolate the total MR-CISD and MR-AQCC energy in a systematic study of the excited states of diatomic molecules. Extrapolation of the total energy was also performed successfully in several studies of the ozone molecule[661−663] and to obtain a high-quality potential energy surface for the F + H_2 reaction.[105] In a study on the vibrational states of LiH molecule, Holka et al.[650] defined the uncorrelated energy corresponding to $\mathbf{2}^2$ CAS and calculated it with large basis sets, essentially obtaining the basis set limit. The correlation energy was defined as the total MR-CISD energy (using a large CAS reference wave function) less the $\mathbf{2}^2$ CAS uncorrelated energy, and this difference was extrapolated with the cubic formula. Very accurate results could be obtained this way (1−2 cm^{-1} accuracy for vibrational levels up to dissociation), but large basis sets (up to 6Z) were required. The drawback of this method is that it is difficult to generalize to other systems in which the "static" correlation is not as well-defined as in the case of a single-bond breaking.

For a given orbital basis, a sequence of progressively more accurate electronic structure methods can be used to extrapolate to the full-CI limit. With CI methods, such sequences typically rely on orbital basis truncations that are based on natural orbital occupations, on overall excitation level (CISD, CISDT, and CISDTQ, etc.), or on numerical selection methods. These two different extrapolations, orbital basis and method, can be used together to estimate the combined CBS and full-CI limit. This is the complete-CI limit which, after accounting for relativistic and nonadiabatic effects, may be compared to experimental values. The Gaussian Gx methods,[430] the focal-point method,[431] the CEEIS method,[195,196] the Wx methods,[432] and the HEAT methods[433,434] are popular examples of this combined extrapolation approach, although these, except CEEIS, involve typically single-reference CC and PT calculations. Recently a new multireference extrapolation procedure has also been introduced by Jiang and Wilson[429] under the name MR-ccCA (multireference analog of the correlation consistent composite approach).

Three difficulties can arise with these extrapolations. The first occurs with the CBS limit for a given electronic structure method. If the wave function expansion or the energy expression is insufficiently flexible to describe the property, then erratic or unreliable results can be computed for the basis set sequence. Typically this occurs when there are two or more qualitatively different components of the property, and the different components converge at different rates with basis set expansion. As one or the other of these competing effects dominates within the basis function sequence, erratic convergence to the CBS limit is observed, making reliable extrapolations difficult. Examples of this were observed in the bond length extrapolations in ref 369 for several of the single-reference methods; it was argued in these cases that the additional wave function flexibility from the increasing basis set size improves the description of the electron correlation, which tends to shorten the computed bond lengths, but it also reduces the artificial charge contamination of the SCF reference function, which tends to lengthen the computed bond lengths. MR expansions are inherently more flexible than SR expansions, and such erratic convergence is less likely to occur, or to be smaller in magnitude, for a MR sequence than for an otherwise comparable SR sequence. In the bond length calculation example,[369] the MCSCF reference space eliminates the spurious charge contamination, leaving only the more well-behaved and smoothly convergent dynamical correlation effects to be described by the MRCI extrapolation sequence. However, even with MR methods, extrapolations that begin with small basis sets can be unreliable.

Another difficulty with extrapolation methods is that a particular sequence of electronic structure methods may not be convergent for all basis sets. An example of this is the SR-MP*n* sequence of energy calculations. For a small cc-pVDZ basis expansion, the SR-MP*n* sequence might converge smoothly to a value, but for larger cc-pVTZ and cc-pVQZ basis sets, the SR-MP*n* sequence can diverge. Olsen et al.[435] and Helgaker et al.[436] argue this divergence is often due to singularities for negative values of the perturbation parameter which arise from the diffuse functions of the larger basis sets. Negative values of the perturbation parameter correspond to the nonphysical situation in which electron interactions are attractive rather than repulsive, but these nonphysical singularities adversely affect the mathematical convergence properties of the SR-MP*n* sequence. (These are called *back-door intruder states* in Chapter 14 of ref 436.) Variational MR methods are typically less susceptible to these convergence issues than PT methods, but they can arise nonetheless. An example is the MR-AQCC calculation on O_3 with a small $\mathbf{2}^2$ reference space. cc-pVDZ calculations result in a reasonably accurate PES near the equilibrium geometry, but larger cc-pVTZ and cc-pVQZ basis sets result in inaccurate PESs and in very slow convergence. Both effects are due to basis-dependent intruder-state issues. These problems all disappear with larger, more flexible, reference spaces.[369]

Finally the general difficulty with all basis extrapolation methods is that the effort for each basis scales with the basis size as $O(n^X)$, where X ranges from 4 for low-level methods (e.g., SCF on small molecules), to 5 or 6 for higher accuracy electronic structure methods (e.g., MCSCF, MRCI), up to the number of electrons N for full-CI methods. The memory, storage, I/O, and communication requirements for various methods also increase with increasing basis size n. Thus the calculations required for the

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

larger basis sets in the extrapolation sequence can become expensive or, due to practical limitations, impossible.

The other general approach to addressing the orbital basis set convergence issue is through the incorporation of explicit interelectron coordinates into the wave function or energy expression. This approach dates from the work of Hylleraas[437] in 1928. For the present discussion, it is the general formulation of Kutzelnigg[438] and Klopper[439,440] in terms of standard MO basis sets that is most applicable. The details of the method are discussed in the recent review by Klopper et al.[441] These approaches are called R12 or F12 methods depending on the specific form of the interelectronic interactions. This approach was first applied to CI by Röhse et al.[442] for benchmark full-CI calculations on H_2 and H_3^+. The formulation for general MRCI expansions was given by Gdanitz[443,444] in 1993 using the COLUMBUS codes. In both cases, the R12 terms were limited only to the reference space, the remaining expansion terms were treated in the normal CI or ACPF approach. The newest version of this method is available within the AMICA program,[445] and some applications are presented in refs 446−449. Te-No[450] introduced an F12 extension to MR-MP2 method with internal contraction that is applicable to larger molecular systems. Varganov and Martinez[451] use geminal augmentation with the MCSCF method, but their implementation is limited to two-electron systems. A novel approach was introduced by Torheyden and Valeev[452] which allows the calculation of the R12 second-order correction for any reference state for which 1-RDM and 2-RDM are available. The formalism uses internal contractions to obtain the geminal replacements, but it can be applied equally well to uncontracted MR-CISD as was demonstrated in ref 452. A spin-free version has recently been proposed by Kong and Valeev.[453] This is a very promising development since it would allow application of R12 methods into arbitrary electronic structure methods such as the GCF, DMRG, and CT approaches discussed in sections 2.1.5 and 2.3. Recently an ic-MR-CISD-F12 approach has been presented by Shiozaki et al.[454,455] In general, the R12 terms in these methods are designed primarily to address the description of the Kato cusp associated with dynamical correlation,[456] while the underlying MCSCF and MRCI expansions describe the valence and other strong correlation effects, e.g., due to near-degeneracies and curve crossings. Thus the fundamental flexibility and general advantages of MR methods are retained.

These R12 methods are incorporated within a standard MO basis set expansion, and therefore they still have artifacts associated with the finite basis set truncation. However, when combined with basis extrapolation approaches, the convergence to the CBS limit occurs much faster, and therefore explicit calculations only with the smaller basis sets[441] are required. Thus within the basis extrapolation sequence, any increase in effort due to the R12 methodology is offset by the use of smaller orbital basis sets.

2.5. Analytic Gradients of Multireference Methods

In this section an overview of some of the general features of MCSCF and MRCI analytic energy gradients is presented. A more detailed and complete discussion is given in the review article of ref 21. The gradient formalism for these MR methods may be compared to those of various SR methods discussed in the review article of Pulay.[457] The CSF expansion coefficients for MCSCF and MRCI wave functions are variationally determined, and this allows the Hellmann−Feynman theorem, using second-quantized conventions for the Hamiltonian operator, to be exploited. The choice of orbitals for the MRCI expansion is somewhat arbitrary because of redundant orbital rotation variables associated with the MCSCF expansion space, and in certain situations the resolution of these rotations must be accounted for when computing the MRCI energy gradient. One of the challenges of analytic gradient methods for MR expansions has been the development and implementation of general efficient procedures that match the flexibility and capability of the wave functions themselves.

A formal approach that meets this challenge of generality and efficiency is based on a sequence of successive geometry-dependent orbital transformations in which the effects of individual constraints or conditions imposed on the orbitals may be considered individually.[21,458,459] In the straightforward case, there would be four orbital basis sets.

$$\boldsymbol{\varphi}^{[C]}(\mathbf{R}) = \boldsymbol{\chi}(\mathbf{R})\,\mathbf{C}(0) \tag{124}$$

$$\boldsymbol{\varphi}^{[S]}(\mathbf{R}) = \boldsymbol{\varphi}^{[C]}(\mathbf{R})\,\mathbf{S}^{[C]}(\mathbf{R})^{-1/2} \tag{125}$$

$$\boldsymbol{\varphi}^{[K]}(\mathbf{R}) = \boldsymbol{\varphi}^{[S]}(\mathbf{R})\,\exp(\mathbf{K}) \tag{126}$$

$$\boldsymbol{\varphi}^{[Z]}(\mathbf{R}) = \boldsymbol{\varphi}^{[K]}(\mathbf{R})\,\exp(\mathbf{Z}) \tag{127}$$

$\mathbf{R}$ denotes the coordinates of the atom centers, or more generally the basis function centers, within the molecule. The actual coordinates that are used in a calculation may be, for example, the Cartesian coordinates of the atom centers or some choice of internal coordinates. The basis $\boldsymbol{\chi}(\mathbf{R})$ is the atom-centered AO basis; as the atom centers move with the molecular geometry, the associated basis functions move along with them. This AO basis may be symmetry-adapted to the point group of the molecule. The $\mathbf{C}(0)$ matrix contains the fully optimized and resolved orbital coefficients at the reference geometry denoted, for notational convenience, $\mathbf{R} = 0$. It is the analytic gradient at this reference geometry that is of interest. The basis $\boldsymbol{\varphi}^{[C]}(\mathbf{R})$ is a geometry-dependent basis that generally is orthonormal only at $\mathbf{R} = 0$. The symmetric positive-definite matrix $\mathbf{S}^{[C]}(\mathbf{R})$ is the orbital overlap matrix in the $\boldsymbol{\varphi}^{[C]}(\mathbf{R})$ basis, and it is used to define[460] the basis $\boldsymbol{\varphi}^{[S]}(\mathbf{R})$ which is orthonormal at all $\mathbf{R}$. The basis $\boldsymbol{\varphi}^{[K]}(\mathbf{R})$ is the energy-optimized orthonormal orbital basis defined in terms of the skew-symmetric matrix $\mathbf{K}$ whose nonzero elements correspond to the essential MCSCF orbital rotation parameters. Finally, $\boldsymbol{\varphi}^{[Z]}(\mathbf{R})$ is the fully resolved orbital basis defined in terms of the skew-symmetric matrix $\mathbf{Z}$ whose nonzero elements correspond to the redundant MCSCF orbital rotation parameters. The two sets of orbital rotation parameters, essential and redundant, are disjointed in the sense that a nonzero K_{pq} element implies a zero Z_{pq} element, and a nonzero Z_{pq} element implies a zero K_{pq} element. With an appropriate ordering of the MO basis functions, the $\mathbf{Z}$ matrix assumes a block-diagonal form and the $\mathbf{K}$ matrix assumes a complementary block-off-diagonal form (see Figure 1 in ref 21). It is this final orbital basis $\boldsymbol{\varphi}^{[Z]}(\mathbf{R})$ that is used to define the MRCI wave function at the reference geometry. With this sequence of orbital transformations it is seen that $\boldsymbol{\varphi}^{[C]}(0) = \boldsymbol{\varphi}^{[S]}(0) = \boldsymbol{\varphi}^{[K]}(0) = \boldsymbol{\varphi}^{[Z]}(0)$, but these orbital bases are generally different at arbitrary $\mathbf{R} \neq 0$. In this formulation, the orbital bases $\boldsymbol{\varphi}^{[K]}(\mathbf{R})$ and $\boldsymbol{\varphi}^{[Z]}(\mathbf{R})$ are orthonormal because the transformation matrices $\exp(\mathbf{K})$ and $\exp(\mathbf{Z})$ are intrinsically orthogonal; thus no additional constraints need to be satisfied, and no additional optimization variables, particularly in the form of Lagrange multipliers, are introduced. In more general situations, there might be several orbital optimization steps involved in computing the orbitals for the final MRCI energy and wave function. This successive orbital transformation

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

approach may be generalized in a straightforward way to account for any arbitrary sequence of orbital optimization steps.[21]

The general approach to computing a particular analytic energy gradient will be similar for all of the electronic structure methods discussed in this section. The energy as a function of $\mathbf{R}$ will be written in terms of the geometry-dependent one- and two-electron integrals in the most appropriate orbital basis. This will also involve the geometry-dependent density matrices, transition density matrices, and various other combinations of these quantities such as geometry-dependent Fock matrices. Expansion techniques will be used to determine the first-order dependence of the energy on the various geometry-dependent quantities at the reference geometry $\mathbf{R} = 0$. Finally, the transformation properties of these quantities, consistent with eqs 124−127, will be used as necessary to simplify the expressions, to isolate the geometry-dependent factors from the geometry-independent factors, and eventually to express the analytic energy gradient as simple summations of the derivatives of the overlap integrals and the one- and two-electron Hamiltonian integrals in the AO basis.

2.5.1. Single-State MCSCF Gradient. A summary of some of the important features of the MCSCF gradient for a single electronic state follows. These features will then form the foundation for the discussion of the other wave functions. A trial MCSCF wave function may be written in the $\boldsymbol{\varphi}^{[S]}(\mathbf{R})$ orbital basis as a generalization of eq 70.

$$|\psi^{\mathrm{trial}}(\mathbf{R})\rangle = \exp(K(\mathbf{R}))\exp(P(\mathbf{R}))|ref(\mathbf{R});[S]\rangle \tag{128}$$

The reference wave function $|ref(\mathbf{R})\rangle$ may be chosen to be either a ground or an excited electronic state; a ground state would correspond to the lowest Hamiltonian matrix eigenpair at $\mathbf{R} = 0$, whereas an excited state would correspond to a higher eigenpair. The ordering of the states may, of course, change at various values of $\mathbf{R}$, so it is important for the formalism to remain general in this respect. It is convenient to take $|ref(\mathbf{R})\rangle$ to be the optimized MCSCF wave function at $\mathbf{R} = 0$, and at displaced geometries to be the wave function with the same normalized CSF expansion coefficients, $\|\mathbf{c}^{mc}(0)\|_2 = 1$, but represented in the corresponding orthonormal $\boldsymbol{\varphi}^{[S]}(\mathbf{R})$ orbital basis. This reference wave function will be denoted $|mc(0);[S]\rangle$, and there is a corresponding reference PES associated with this reference wave function defined as

$$\begin{aligned} E^{\mathrm{ref}}(\mathbf{R}) &= \langle mc(0);[S]|H^{[S]}(\mathbf{R})|mc(0);[S]\rangle \\ &= \sum_{r,s} h^{[S]}_{rs}(\mathbf{R})\langle mc(0);[S]|E_{rs}|mc(0);[S]\rangle \\ &\quad + \frac{1}{2}\sum_{p,q,r,s} g^{[S]}_{pqrs}(\mathbf{R})\langle mc(0);[S]|e_{pqrs}|mc(0);[S]\rangle \\ &= \mathrm{Tr}(\mathbf{h}^{[S]}(\mathbf{R})\mathbf{D}^{mc[s]}(0)) + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[S]}(\mathbf{R})\mathbf{d}^{mc[S]}(0)) \end{aligned} \tag{129}$$

The last expression in particular shows that all of the geometry dependence of this reference energy surface derives from the geometry dependence of the one- and two-electron integrals that define the second-quantized Hamiltonian operator. The reduced density matrices $\mathbf{D}^{mc[S]}(0)$ and $\mathbf{d}^{mc[S]}(0)$ are geometry-independent because (i) they depend on the geometry-independent coupling coefficients and (ii) they depend on the fixed, reference-geometry, CSF expansion coefficients $\mathbf{c}^{mc}(0)$. The above reference PES is never actually computed at arbitrary $\mathbf{R}$; it is rather a formal construct that is used to reveal how the various quantities depend on the molecular displacements $\mathbf{R}$.

The trial energy expectation value may then be written

$$\begin{aligned} E^{\mathrm{trial}}(\mathbf{K},\mathbf{p};\mathbf{R}) &= \langle\psi^{\mathrm{trial}}(\mathbf{R})|H(\mathbf{R})|\psi^{\mathrm{trial}}(\mathbf{R})\rangle \\ &= \langle mc(0);[S]|\exp(-P(\mathbf{R}))\exp(-K(\mathbf{R}))H(\mathbf{R}) \\ &\quad \times\exp(K(\mathbf{R}))\exp(P(\mathbf{R}))\rangle mc(0);[S]\rangle \\ &= E^{\mathrm{ref}}(\mathbf{R}) + \begin{pmatrix}\mathbf{k}(\mathbf{R})^{\mathrm{T}} & \mathbf{p}(\mathbf{R})^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\mathbf{f}_{\mathrm{orb}}(\mathbf{R})\\ \mathbf{f}_{\mathrm{csf}}(\mathbf{R})\end{pmatrix} \\ &\quad + \frac{1}{2}\begin{pmatrix}\mathbf{k}(\mathbf{R})^{\mathrm{T}} & \mathbf{p}(\mathbf{R})^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\mathbf{G}_{\mathrm{orb,orb}}(\mathbf{R}) & \mathbf{G}_{\mathrm{orb,csf}}(\mathbf{R})\\ \mathbf{G}_{\mathrm{csf,orb}}(\mathbf{R}) & \mathbf{G}_{\mathrm{csf,csf}}(\mathbf{R})\end{pmatrix}\begin{pmatrix}\mathbf{k}(\mathbf{R})\\ \mathbf{p}(\mathbf{R})\end{pmatrix} + \ldots \\ &= E^{\mathrm{ref}}(\mathbf{R}) + \boldsymbol{\lambda}(\mathbf{R})\cdot\mathbf{f}(\mathbf{R}) + \frac{1}{2}\boldsymbol{\lambda}(\mathbf{R})^{\mathrm{T}}\mathbf{G}(\mathbf{R})\boldsymbol{\lambda}(\mathbf{R}) + \ldots \end{aligned} \tag{130}$$

which corresponds to eq 72 with the $\mathbf{R}$-dependence denoted explicitly. At any arbitrary $\mathbf{R}$, the MCSCF wave function parameters $\boldsymbol{\lambda}^{mc}(\mathbf{R})$ are those that satisfy the variational conditions $\partial E^{\mathrm{trial}}(\boldsymbol{\lambda}(\mathbf{R});\mathbf{R})/\partial\boldsymbol{\lambda} = \mathbf{0}$. This results in a coupled set of nonlinear equations

$$\mathbf{0} = \mathbf{f}^{mc}(\mathbf{R}) + \mathbf{G}^{mc}(\mathbf{R})\boldsymbol{\lambda}^{mc}(\mathbf{R}) + O(\boldsymbol{\lambda}^{mc}(\mathbf{R})^2)\ldots \tag{131}$$

that must be satisfied by the parameters $\boldsymbol{\lambda}^{mc}(\mathbf{R})$ at arbitrary $\mathbf{R}$. There is no closed-form solution to this equation. However, this equation is sufficient to determine the corresponding Taylor expansion of the geometry-dependent parameters $\boldsymbol{\lambda}(\mathbf{R})$ relative to the reference geometry $\mathbf{R} = 0$ values. Differentiating eq 131 with respect to a displacement of a representative atomic center coordinate denoted x, evaluation at the reference geometry, and using the relation $\boldsymbol{\lambda}^{mc}(0) = \mathbf{0}$ give

$$\boldsymbol{\lambda}^{mc}(0)^x = -\mathbf{G}^{mc}(0)^{-1}\mathbf{f}^{mc}(0)^x \tag{132}$$

The superscript x denotes differentiation, and it is used to identify quantities that depend on some displacement of the molecular geometry. This gives the first-order change in the MCSCF orbitals and CSF expansion coefficients at the reference geometry to the displacement along the coordinate direction labeled by x. The MCSCF energy at arbitrary $\mathbf{R}$ is given by eq 130 with the specific $\mathbf{K}^{mc}(\mathbf{R})$ and $\mathbf{p}^{mc}(\mathbf{R})$ parameters determined from eq 131. Differentiation of this energy expression with respect to a geometry displacement and evaluation at $\mathbf{R} = 0$ gives an element of the MCSCF analytic energy gradient

$$E^{mc}(0)^x = E^{\mathrm{ref}}(0)^x + \boldsymbol{\lambda}(0)\cdot\mathbf{f}(0)^x + \boldsymbol{\lambda}(0)^x\cdot\mathbf{f}(0) + \ldots \tag{133}$$

$$= E^{\mathrm{ref}}(0)^x \tag{134}$$

The truncation follows from the relations $\mathbf{f}(0) = \mathbf{0}$ and $\boldsymbol{\lambda}(0) = \mathbf{0}$. This is an example of the Hellmann−Feynman theorem using second-quantized conventions for the definition of the Hamiltonian operator: the first-order wave function does not contribute to the single-state MCSCF energy gradient. The MCSCF energy gradient may be written

$$\begin{aligned} E^{mc}(0)^x = E^{\mathrm{ref}}(0)^x &= \langle mc(0);[S]|\hat{H}^{[s]}(0)^x|mc(0);[S]\rangle \\ &= \mathbf{c}^{mc}(0)^{\mathrm{T}}\hat{\mathbf{H}}^{[s]}(0)^x\mathbf{c}^{mc}(0) \\ &= \mathrm{Tr}(\mathbf{h}^{[S]}(0)^x\mathbf{D}^{mc[S]}(0)) + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[S]}(0)^x\mathbf{d}^{mc[S]}(0)) \end{aligned} \tag{135}$$

This last expression shows that the density matrices contain the displacement-independent factors of the energy gradient

elements, which are shared by all possible displacements x, and the derivative integrals contain all of the displacement-dependent factors for each displacement direction x. If the energy gradient were evaluated using this expression, the entire set of derivative terms $\mathbf{h}^{[S]}(0)^x$ and $\mathbf{g}^{[S]}(0)^x$ for up to $3N_{\mathrm{atom}}$ possible displacement directions x would need to be computed. This would require effort proportional to $3N_{\mathrm{atom}}$. A more efficient approach is to transform the gradient expression back to the original AO basis. Using the sequence of orbital transformations eqs 124 and 125, the gradient component may be written as[21]

$$E^{mc}(0)^x = \mathrm{Tr}(\mathbf{h}^{[\chi]}(0)^x\, \mathbf{D}^{mc[\chi]}(0)) + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[\chi]}(0)^x\, \mathbf{d}^{mc[\chi]}(0)) - \mathrm{Tr}(\mathbf{S}^{[\chi]}(0)^x\, \mathbf{F}^{mc[\chi]}(0)) \tag{136}$$

The AO density and Fock matrices are computed as

$$\begin{aligned} F^{mc[\chi]}_{\mu\nu}(0) &= \sum_{pq} C(0)_{\mu p}\, C(0)_{\nu q}\, F^{mc[C]}_{pq}(0) \\ D^{mc[\chi]}_{\mu\nu}(0) &= \sum_{pq} C(0)_{\mu p}\, C(0)_{\nu q}\, D^{mc[C]}_{pq}(0) \\ d^{mc[\chi]}_{\mu\nu\lambda\sigma}(0) &= \sum_{pqrs} C(0)_{\mu p}\, C(0)_{\nu q}\, C(0)_{\lambda r}\, C(0)_{\sigma s}\, d^{mc[C]}_{pqrs}(0) \end{aligned} \tag{137}$$

The transformation of these arrays is similar to the one- and two-electron integral transformation operation. The actual operation counts differ because, in the typical situation, arrays are transformed from the smaller occupied orbital basis to the larger AO basis.

The final expression eq 136 is important because the two-electron Hamiltonian integrals are very sparse in the atom-centered AO basis. A particular two-electron repulsion integral depends on, at most, only four atom centers, or 12 Cartesian displacements, out of the $3N_{\mathrm{atom}}$ total possible displacements. Consequently, there are only about 12 times as many nonzero AO derivative integrals as undifferentiated AO integrals. Furthermore, the computation of the AO derivatives by shells allows reuse of various intermediate quantities, resulting in an even more efficient overall procedure.[461] By exploiting these features, the trace operation may be computed in the AO basis with effort that is formally independent of N_{atom}; i.e., $O(N^0_{\mathrm{atom}}) = O(1)$. This simplification affects the number of arithmetic operations required to evaluate the energy gradient and the total amount of memory and external storage space that is required for the computation. Thus the analytic gradient procedure described above is both more efficient and more accurate than a finite-difference approach, and it has similar advantages when these gradients are used to fit molecular potential energy surfaces,[462] to optimize molecular geometries, or when they are used directly to compute classical dynamical trajectories.[365] In ref 369 molecular geometry optimizations were performed for 20 molecules using MCSCF wave functions and with a variety of orbital basis sets and a wide range of CSF expansion spaces. The effort for the MCSCF gradient evaluations for these molecules required between 8.0 and 84.4% of the total computational effort (including integral evaluation and wave function optimization), with a mean of 58.1%. These timings demonstrate that the computational procedure presented above is very efficient, that it is independent of N_{atom}, and that it may be applied to any molecule for which the MCSCF wave function optimization itself is practical.

2.5.2. State-Averaged MCSCF Gradient. For state-averaged calculations, the trial wave function parametrization of eqs 63−70 is generalized as

$$P(\mathbf{R}) = \sum_j^{N_{\mathrm{av}}} P^j(\mathbf{R}) \tag{138}$$

$$P^j(\mathbf{R}) = |p^j(\mathbf{R})\rangle\langle mc^j(0)| - |mc^j(0)\rangle\langle p^j(\mathbf{R})| \tag{139}$$

$$|p^j(\mathbf{R})\rangle = \sum_m^{N_{\mathrm{csf}}} p^j_m(\mathbf{R})|m;[S]\rangle \tag{140}$$

In analogy to the single-state case, the reference states $|mc^j(0)\rangle$ are defined with the $\boldsymbol{\varphi}^{[S]}(\mathbf{R})$ orbitals and with the fixed CSF coefficients $\mathbf{c}^j(0)$ corresponding to the jth eigenpair at $\mathbf{R} = 0$. This generalization allows an averaged trial energy to be written

$$\begin{aligned} \bar{E}^{\mathrm{trial}}(\mathbf{k}, \mathbf{p}^{1:N_{\mathrm{av}}}; \mathbf{R}) &= \sum_j^{N_{\mathrm{av}}} w_j E^{\mathrm{trial}}_j(\mathbf{k}, \mathbf{p}^{1:N_{\mathrm{av}}}; \mathbf{R}) \\ &= \sum_j^{N_{\mathrm{av}}} w_j \langle mc^j(0);[S]|\exp(-P)\exp(-K)H\exp(K)(P)|mc^j(0);[S]\rangle \\ &= \bar{E}^{\mathrm{ref}}(\mathbf{R}) + \begin{pmatrix} \mathbf{k}^{\mathrm{T}} & \mathbf{p}^{1:N_{\mathrm{av}},\mathrm{T}} \end{pmatrix} \begin{pmatrix} \bar{\mathbf{f}}_{\mathrm{orb}} \\ \mathbf{f}^{1:N_{\mathrm{av}}}_{\mathrm{csf}} \end{pmatrix} \\ &\quad + \frac{1}{2}\begin{pmatrix} \mathbf{k}^{\mathrm{T}} & \mathbf{p}^{1:N_{\mathrm{av}},\mathrm{T}} \end{pmatrix} \begin{pmatrix} \bar{\mathbf{G}}_{\mathrm{orb,orb}} & \mathbf{G}^{1:N_{\mathrm{av}}}_{\mathrm{orb,csf}} \\ \mathbf{G}^{1:N_{\mathrm{av}}}_{\mathrm{csf,orb}} & \mathbf{G}^{1:N_{\mathrm{av}}}_{\mathrm{csf,csf}} \end{pmatrix} \begin{pmatrix} \mathbf{k} \\ \mathbf{p}^{1:N_{\mathrm{av}}} \end{pmatrix} + \cdots \end{aligned} \tag{141}$$

The vector $\mathbf{p}^{1:N_{\mathrm{av}}}(\mathbf{R})$ corresponds to the concatenation of the individual $\mathbf{p}^j(\mathbf{R})$ vectors for each of the states included in the state average. The state-averaged quantities

$$\bar{E}^{\mathrm{ref}}(\mathbf{R}) = \sum_j^{N_{\mathrm{av}}} w_j E^{\mathrm{ref}}_j(\mathbf{R}) \tag{142}$$

$$\bar{\mathbf{f}}_{\mathrm{orb}}(\mathbf{R}) = \sum_j^{N_{\mathrm{av}}} w_j \mathbf{f}^j_{\mathrm{orb}}(\mathbf{R}) \tag{143}$$

$$\bar{\mathbf{G}}_{\mathrm{orb,orb}}(\mathbf{R}) = \sum_j^{N_{\mathrm{av}}} w_j \mathbf{G}^j_{\mathrm{orb,orb}} \tag{144}$$

are defined in terms of their state-specific components, but they are most efficiently computed using state-averaged density matrices.[365] The first-order equation for the orbital and CSF response in terms of these augmented gradient and Hessian matrices is

$$\begin{pmatrix} \mathbf{k}(0)^x \\ \mathbf{p}^{1:N_{\mathrm{av}}}(0)^x \end{pmatrix} = - \begin{pmatrix} \bar{\mathbf{G}}_{\mathrm{orb,orb}}(0) & \mathbf{G}^{1:N_{\mathrm{av}}}_{\mathrm{orb,csf}}(0) \\ \mathbf{G}^{1:N_{\mathrm{av}}}_{\mathrm{csf,orb}}(0) & \mathbf{G}^{1:N_{\mathrm{av}}}_{\mathrm{csf,csf}}(0) \end{pmatrix}^{-1} \begin{pmatrix} \bar{\mathbf{f}}_{\mathrm{orb}}(0)^x \\ \mathbf{f}^{1:N_{\mathrm{av}}}_{\mathrm{csf}}(0)^x \end{pmatrix} \tag{145}$$

Differentiation of the state-averaged trial energy with the optimal wave function parameters with respect to a displacement, and evaluation at $\mathbf{R} = 0$, gives, in principle, the state-averaged energy gradient $\bar{E}^{mc}(0)^x$ analogous to eq 135, in terms of the state-averaged density matrices. However, it is not the gradient of the state-averaged $\bar{E}^{mc}(\mathbf{R})$ that is of interest; it is rather

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

the gradients of the individual states that are important. It is the energy gradients of the individual states that determine, for example, the classical trajectories on these PESs. This gradient is given by substituting the $\mathbf{k}(0)^x$ and $\mathbf{p}^j(0)^x$ for a particular state into the state-specific energy expression, eq 130. After truncation, this gives

$$E_j^{mc}(0)^x = E_j^{\mathrm{ref}}(0)^x + \mathbf{k}(0)^x \cdot \mathbf{f}_{\mathrm{orb}}^j(0) \tag{146}$$

In contrast to eq 134, the generally nonzero state-specific orbital gradient terms $\mathbf{f}_{\mathrm{orb}}^j(0)$ are seen to contribute to the energy gradient expression through the first-order change in the orbitals. The $\mathbf{f}_{\mathrm{csf}}^j(0)$ terms are zero, and the $\mathbf{p}^j(0)^x$ response terms do not contribute in eq 146 because of the Hellmann−Feynman relation. The next step is to express this contribution to the gradient in a form that allows for efficient evaluation. To this end, eq 146 is written in the slightly modified form

$$E_j^{mc}(0)^x = E_j^{\mathrm{ref}}(0)^x + (\mathbf{k}(0)^{x,\mathrm{T}}, \mathbf{p}^{1:N_{\mathrm{av}}}(0)^{x,\mathrm{T}}) \begin{pmatrix} \mathbf{f}_{\mathrm{orb}}^j(0) \\ \mathbf{0}^{1:N_{\mathrm{av}}} \end{pmatrix} \tag{147}$$

and eq 145 is used to give the following sequence of identities[21,458,459,463,464]

$$\begin{aligned} E_j^{mc}(0)^x &= E_j^{\mathrm{ref}}(0)^x - (\bar{\mathbf{f}}_{\mathrm{orb}}(0)^{x,\mathrm{T}}, \mathbf{f}_{\mathrm{csf}}^{1:N_{\mathrm{av}}}(0)^{x,\mathrm{T}}) \\ &\quad \times \begin{pmatrix} \bar{\mathbf{G}}_{\mathrm{orb,orb}}(0) & \mathbf{G}_{\mathrm{orb,csf}}^{1:N_{\mathrm{av}}}(0) \\ \mathbf{G}_{\mathrm{csf,orb}}^{1:N_{\mathrm{av}}}(0) & \mathbf{G}_{\mathrm{csf,csf}}^{1:N_{\mathrm{av}}}(0) \end{pmatrix} \begin{pmatrix} \mathbf{f}_{\mathrm{orb}}^j(0) \\ \mathbf{0}^{1:N_{\mathrm{av}}} \end{pmatrix} \\ &= E_j^{\mathrm{ref}}(0)^x + (\bar{\mathbf{f}}_{\mathrm{orb}}(0)^{x,\mathrm{T}}, \mathbf{f}_{\mathrm{csf}}^{1:N_{\mathrm{av}}}(0)^{x,\mathrm{T}}) \begin{pmatrix} \boldsymbol{\lambda}_{\mathrm{orb}}^j(0) \\ \boldsymbol{\lambda}_{\mathrm{csf}}^{1:N_{\mathrm{av}}}(0) \end{pmatrix} \end{aligned} \tag{148}$$

$$= E_j^{\mathrm{ref}}(0)^x + \bar{\mathbf{f}}_{\mathrm{orb}}(0)^x \cdot \boldsymbol{\lambda}_{\mathrm{orb}}^j(0) + \sum_k^{N_{\mathrm{av}}} \mathbf{f}_{csf}^k(0)^x \cdot \boldsymbol{\lambda}_{\mathrm{csf}}^{k,j}(0) \tag{149}$$

As written, this expression for the energy gradient would require the computation of the $\bar{\mathbf{f}}_{\mathrm{orb}}(0)^x$ and $\bar{\mathbf{f}}_{\mathrm{csf}}^I(0)^x$ vectors for all $3N_{\mathrm{atom}}$ possible displacements. However, the density matrices[21,459]

$$\begin{aligned} \bar{D}_{pq}^{j;\lambda} &= \sum_m^{N_{\mathrm{csf}}} \sum_k^{N_{\mathrm{av}}} \lambda_{\mathrm{csf}}^{k,j}(0)_m \langle m|E_{pq} + E_{pq}|mc^k(0)\rangle \\ \bar{d}_{pqrs}^{j;\lambda} &= \frac{1}{2} \sum_m^{N_{\mathrm{csf}}} \sum_k^{N_{\mathrm{av}}} \lambda_{\mathrm{csf}}^{k,j}(0)_m \langle m|e_{pqrs} + e_{qprs} + e_{pqsr} + e_{qpsr}|mc^k(0)\rangle \\ \bar{\mathbf{D}}^{\Lambda_j} &= -\{\mathbf{D}^{mc}(0); \Lambda_j\} \qquad \bar{\mathbf{d}}^{\Lambda_j} = -\{\mathbf{d}^{mc}(0); \Lambda_j\} \end{aligned} \tag{150}$$

may be used to write the gradient in the same general form as the single-state expression

$$\begin{aligned} E_j^{mc}(0)^x &= \mathrm{Tr}(\mathbf{h}^{[S]}(0)^x(\mathbf{D}^j + \bar{\mathbf{D}}^{\Lambda_j} + \bar{\mathbf{D}}^{j;\lambda})) \\ &\quad + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[S]}(0)^x(\mathbf{d}^j + \bar{\mathbf{d}}^{\Lambda_j} + \bar{\mathbf{d}}^{j;\lambda})) \\ &= \mathrm{Tr}(\mathbf{h}^{[S]}(0)^x \mathbf{D}^{j,\mathrm{total}}) + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[S]}(0)^x \mathbf{d}^{j,\mathrm{total}}) \end{aligned} \tag{151}$$

All of the displacement dependence occurs within the Hamiltonian integrals, and the effective density matrices $\mathbf{D}^{j,\mathrm{total}}$ and $\mathbf{d}^{j,\mathrm{total}}$ are computed only from displacement-independent quantities. The analytic energy gradient can be computed in the atom-centered AO basis using the same sequence of transformations as in the single-state case.

$$\begin{aligned} E_j^{mc}(0)^x &= \mathrm{Tr}(\mathbf{h}^{[\chi]}(0)^x \mathbf{D}^{j,\mathrm{total}[\chi]}(0)) \\ &\quad + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[\chi]}(0)^x \mathbf{d}^{j,\mathrm{total}[\chi]}(0)) - \mathrm{Tr}(\mathbf{S}^{[\chi]}(0)^x \mathbf{F}^{j,\mathrm{total}[\chi]}(0)) \end{aligned} \tag{152}$$

As with the single-state wave function optimization case, this allows the energy gradient for each state j within the state-averaging procedure to be computed with effort that is formally independent of N_{atom}. The additional effort corresponding to eqs 148−151 is comparable to that of a single iteration of the state-averaged MCSCF energy optimization procedure. Thus, if the state-averaged wave functions and energy can be computed, then it is also practical to compute the energy gradients for the states of interest.

2.5.3. MRCI Gradient. For the MRCI energy, the CSF expansion coefficients are variationally optimized, which means that the eigenvalue equation

$$H^{[Z]}(\mathbf{R})\, \mathbf{c}^j(\mathbf{R}) = E_j^{ci}(\mathbf{R})\, \mathbf{c}^j(\mathbf{R}) \tag{153}$$

is satisfied at all $\mathbf{R}$. Differentiating this expression with respect to an atomic center displacement and evaluation at $\mathbf{R} = 0$ results in

$$\begin{aligned} E_j^{ci}(0)^x &= \mathbf{c}^j(0)^{\mathrm{T}}\, \mathbf{H}^{[Z]}(0)^x\, \mathbf{c}^j(0) \\ &= \langle ci^j(0); [Z]|H^{[Z]}(0)^x|ci^j(0); [Z]\rangle \\ &= \sum_{m,n}^{N_{\mathrm{csf}}} c_m^j(0)\, c_n^j(0) \langle m; [Z]|H^{[Z]}(0)^x|n; [Z]\rangle \\ &= \mathrm{Tr}(\mathbf{h}^{[Z]}(0)^x \mathbf{D}^{j,ci[z]}(0)) + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[Z]}(0)^x \mathbf{d}^{j,ci[Z]}(0)) \end{aligned} \tag{154}$$

with the wave function normalization $\|\mathbf{c}^j(0)\|_2 = 1$. With regard to the MCSCF energy gradient expression in eq 135, the first-order CI wave function response $\mathbf{c}^j(0)^x$ does not contribute to the energy gradient, and the Hellmann−Feynman theorem is seen to be satisfied for the CI energy gradient. In order to avoid the effort of constructing the derivative integrals in the $\boldsymbol{\varphi}^{[Z]}(\mathbf{R})$ basis, the orbital transformation sequence in eqs 124−127 and the commutator expansion of the Hamiltonian operator allows the CI energy gradient to be written in the $\boldsymbol{\varphi}^{[S]}(\mathbf{R})$ basis.

$$\begin{aligned} E_j^{ci}(0)^x &= \langle ci^j(0); [S]|H^{[S]}(0)^x + [H^{[S]}(0), K(0)^x] \\ &\quad + [H^{[S]}(0), Z(0)^x]|ci^j(0); [S]\rangle \end{aligned} \tag{155}$$

$$\begin{aligned} &= \langle ci^j(0); [S]|H^{[S]}(0)^x|ci^j(0); [S]\rangle \\ &\quad + \mathbf{k}^{mc}(0)^x \cdot \mathbf{f}_{\mathrm{orb}}^{j,ci}(0) + \mathbf{z}^{mc}(0)^x \cdot \mathbf{f}_{\mathrm{orb}}^{j,ci}(0) \end{aligned} \tag{156}$$

The first two terms of eq 156 are of the type previously considered in eqs 146−152 for the state-averaged MCSCF energy gradient, except here, the CI density matrices and CI orbital rotation gradient vector elements are used rather than MCSCF density and gradient elements. In general, the $\mathbf{f}^{j,ci}(0)$ orbital rotation gradient vector is nonzero because the orbitals

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

are optimized for the MCSCF expansion, not for the CI expansion.

The last term is a new type of gradient contribution that derives from the orbital resolution of the MCSCF wave function. The last term contributes to the gradient only when the combination $z^{mc}(0)^x_{pq}$ and $f^{j,ci}(0)_{pq}$ are both nonzero for some particular orbital rotation pair indexed by (pq). As discussed in section 2.2, the $\mathbf{Z}^{mc}$ matrix assumes a block-diagonal form with the appropriate choice of orbital ordering, and it has nonzero elements only for orbital rotations for which the MCSCF energy is invariant. Thus, a particular $f^{j,ci}(0)_{pq}$ element can contribute to only one of the last two terms in eq 156; it cannot contribute to both. Due to the variational determination of the CI expansion coefficients, however, the elements of the CI orbital gradient $f^{j,ci}(0)_{pq}$ are zero for orbital rotations that are redundant in the CI wave function.[21] This eliminates such gradient contributions for large classes of CI expansion spaces. For example, if the CI expansion space has exactly the same invariant orbital subspaces as the MCSCF reference expansion, which is a common occurrence, then the last term in eq 156 cannot contribute to the energy gradient, and it may be ignored. Thus, the only remaining contributions to the last term in eq 156 are the orbital rotations that are redundant in the MCSCF expansion (nonzero Z^{mc}_{pq}, zero K^{mc}_{pq}) but are essential in the CI expansion space (nonzero $f^{j,ci}(0)_{pq}$). There are three common situations for which this occurs. The first, and probably most common, is the *frozen core* situation which occurs when a proper subset of the MCSCF inactive orbitals is constrained to be doubly occupied in the CI expansion CSFs. The CI wave function, energy, and therefore the gradient then depend on exactly which of those orbitals are frozen from this set and which are correlated. The typical orbital resolution consists of diagonalizing the diagonal subblock of an MCSCF Fock matrix within the MCSCF inactive orbital space. The orbitals associated with the lowest eigenvalues are frozen in the CI expansion, and the remaining orbitals are allowed to have various occupations. Some orbital basis sets are designed to describe well the core orbitals but are not designed to describe well the dynamical correlation of those core electrons, and wave functions expanded with these types of AO basis sets should have their core orbitals frozen in any post-SCF type of calculation.[461]

A second common situation occurs with *frozen virtual* orbitals in which there are very high-lying MCSCF virtual orbitals that would contribute only negligibly to the CI wave function. It is typical in this case to diagonalize the diagonal subblock of an MCSCF Fock matrix and delete the MOs associated with the highest eigenvalues. This is formally equivalent to constraining those orbitals to be unoccupied in all CSFs within the CI expansion, but by removing the orbitals entirely from the MO basis, the overall cost of the CI calculation is reduced. The existence of such high-lying virtual orbitals also depends on the choice of AO basis sets. Most modern, generally contracted,[465] AO basis sets, e.g., the correlation-consistent basis sets of Dunning[466] or the ANO approach of Almlöf and Taylor,[467,468] are designed so that these types of molecular orbitals do not appear. It is primarily the older-style AO basis sets defined with segmented contractions that suffer from this artifact, or also certain benchmark calculations that are sometimes done in the primitive uncontracted Gaussian basis.

A third, perhaps less common, situation occurs when an active orbital in the MCSCF expansion becomes nearly doubly occupied or when an active MCSCF orbital occupation becomes almost zero. In these cases, the subsequent CI expansion is sometimes constructed that treats this orbital as if it were an MCSCF inactive orbital or an MCSCF virtual orbital. This might be done in order to reduce the CI expansion length, but it might also be done, for example, to eliminate some kind of convergence problem associated with intruder states in the expansion (e.g., caused by configurations with occupied MCSCF virtual orbitals that have higher occupations or lower energies than reference configurations involving MCSCF active orbitals). In these situations, the orbitals are typically resolved by diagonalizing either the diagonal subblock of an active-orbital Fock matrix or the diagonal subblock of the MCSCF 1-RDM (i.e., natural orbital resolution). The formalism has been developed generally to allow various other orbital resolution approaches, but, practically speaking, these other possibilities are less common. For example, the resolution based on successive diagonalizations of the CI 1-RDM, the iterative natural orbital approach,[1] which was once a fairly common technique, is now used rarely. The orbital resolutions discussed above may be combined in a very flexible and general manner. Different invariant orbital subspaces may be treated separately and resolved in different ways, and the corresponding effective operators and density matrices are computed accordingly. A more complete discussion of the orbital resolution effects for various combinations of reference and MRCI expansion spaces may be found in the review article in ref 21.

Ultimately, the last term in eq 154, involving $\mathbf{z}^{mc}(0)^x$, may be cast into an expression similar to eq 147. From there it is straightforward to separate the displacement-dependent terms (which carry the coordinate x superscript in these equations) from the displacement-independent terms and to arrive at the final equation for the gradient in terms of the sparse derivative integrals in the AO basis.

$$E^{mc}_j(0)^x = \mathrm{Tr}(\mathbf{h}^{[\chi]}(0)^x\,\mathbf{D}^{j,\mathrm{total}[\chi]}(0)) + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[\chi]}(0)^x\,\mathbf{d}^{j,\mathrm{total}[\chi]}(0)) - \mathrm{Tr}(\mathbf{S}^{[\chi]}(0)^x\,\mathbf{F}^{j,\mathrm{total}[\chi]}(0)) \quad (157)$$

The total effort to construct the gradient vector is dominated, particularly for large CI expansions, by the effort to construct the CI reduced density matrices $\mathbf{D}^{j,ci}(0)$ and $\mathbf{d}^{j,ci}(0)$. This requires roughly the same effort as that of a single Hamiltonian matrix–vector product operation during the iterative solution of the eigenvalue equation at the reference geometry. A consequence of this is that, unlike most other electronic structure methods, the CI energy gradient requires typically less effort than the computation of the CI wave function and energy itself. The above gradient formulation also applies to MR-ACPF and MR-AQCC energy expressions.[21] For the 20 molecules studied in ref 369, the effort for the MRCI and MR-AQCC gradient evaluations required between 2.5 and 52.2% of the total computational effort, with a mean of 9.9% over the entire set of molecules and basis sets. This demonstrates that the computational procedure outlined above is very efficient, that it is independent of N_{atom}, and that it may be applied to any molecule for which the CI wave function optimization itself is practical.

2.6. Nonadiabatic Coupling for Multireference Wave Functions

The formalism and application of nonadiabatic coupling using multireference wave functions have been reviewed recently by Barbatti et al.[365] Some of the important features of this overall approach are summarized here. The basic problem in dynamics simulations of molecules is to solve the time-dependent

Schrödinger equation for the molecular system

$$\left(i\hbar\frac{\partial}{\partial t} - H(\mathbf{r},\mathbf{R})\right)\Psi(\mathbf{r},\mathbf{R},t) = 0 \tag{158}$$

This molecular Hamiltonian operator includes both nuclear **R** and electronic **r** coordinates, $H(\mathbf{r},\mathbf{R}) = T^{\mathrm{nuc}}(\mathbf{R}) + H^{\mathrm{el}}(\mathbf{r};\mathbf{R})$. $T^{\mathrm{nuc}}(\mathbf{R})$ is the nuclear kinetic energy, and $H^{\mathrm{el}}(\mathbf{r};\mathbf{R})$ includes (cf. eq 7) the electronic kinetic energy, the electron−nuclear attraction, the electron−electron repulsion, and, by convention, also the nuclear−nuclear repulsion. The total wave function can be written as a Born expansion (see, e.g., ref 469)

$$\Psi(\mathbf{r},\mathbf{R},t) = \sum_j \Psi_j^{\mathrm{nuc}}(\mathbf{R},t)\,\Psi_j^{\mathrm{el}}(\mathbf{r};\mathbf{R}) \tag{159}$$

In this expansion, $\psi_j^{\mathrm{el}}(\mathbf{r};\mathbf{R})$ is the usual clamped-nucleus electronic wave function for electronic state j, evaluated at the nuclear configuration **R**. Equation 159 itself is not an approximation, but truncation of the summation to a limited number of interacting states is a practical necessity. Substituting eq 159 into eq 158, multiplying from the left by $\psi_k^{\mathrm{el}}(\mathbf{r};\mathbf{R})^*$, and integrating over the electronic coordinates gives

$$\left(i\hbar\frac{\partial}{\partial t} - T^{\mathrm{nuc}}\right)\Psi_k^{\mathrm{nuc}} + \sum_j (-H_{jk} + i\hbar\mathbf{F}_{jk}\cdot\mathbf{v} + G_{jk})\Psi_j^{\mathrm{nuc}} = 0 \tag{160}$$

with

$$H_{jk}(\mathbf{R}) = \langle\psi_j^{\mathrm{el}}|H^{\mathrm{el}}|\psi_k^{\mathrm{el}}\rangle_{\mathbf{r}} \tag{161}$$

$$G_{jk}(\mathbf{R}) = \langle\psi_j^{\mathrm{el}}|T^{\mathrm{nuc}}|\psi_k^{\mathrm{el}}\rangle_{\mathbf{r}} \tag{162}$$

$$\mathbf{F}_{jk}^m(\mathbf{R}) = \langle\psi_j^{\mathrm{el}}|\nabla_m|\psi_k^{\mathrm{el}}\rangle_{\mathbf{r}} \tag{163}$$

$$\mathbf{v}_m \equiv \frac{-i\hbar}{M_m}\nabla_m \tag{164}$$

The index m ranges over the nuclear centers, each with mass M_m. Reference 365 discusses various methods in which eq 160 is approximated and solved in order to determine chemical reaction dynamics. The quantities H_{jk}, G_{jk}, and $\mathbf{F}_{jk}^m$ together determine how, for example, a quantum wave packet transfers its amplitude among the various electronic states during a chemical reaction. In the adiabatic representation $H_{jk}(\mathbf{R}) = V_k(\mathbf{R})\delta_{jk}$, the electronic energy $V_k(\mathbf{R})$ assumes the role of a potential energy for the nuclear motion, and the kinetic energy and velocity-dependent terms determine the nonadiabatic coupling among the various electronic states. Alternatively, in the diabatic representation the electronic states are transformed among themselves in order to eliminate the kinetic energy and velocity-dependent coupling in eq 160, and it is the off-diagonal H_{jk} elements in this diabatic basis that determine the nonadiabatic coupling. In either case, the important electronic states are first computed in order to form the basis states for eq 159, and the coupling elements in this basis must be evaluated according to eqs 160−164. This section focuses on the nonadiabatic coupling vector elements $\mathbf{F}_{jk}(\mathbf{R})$. This is a vector of length $3N_{\mathrm{atom}}$, and an individual element of this vector

$$\mathbf{F}_{jk}(\mathbf{R})^x = \left\langle \psi_j^{\mathrm{el}}(\mathbf{r};\mathbf{R})\left|\frac{\partial}{\partial x}\right|\psi_k^{\mathrm{el}}(\mathbf{r};\mathbf{R})\right\rangle_{\mathbf{r}} \tag{165}$$

determines the coupling between electronic states j and k at the molecular geometry **R**. Expansion of the electronic wave functions in a CSF basis in the fully optimized and resolved orbital basis $\boldsymbol{\varphi}^{[Z]}(\mathbf{R})$ in eq 127 allows the wave function derivative to be written

$$\begin{aligned}\frac{\partial}{\partial x}|\psi_k^{\mathrm{el}}\rangle &= \frac{\partial}{\partial x}\sum_m^{N_{\mathrm{csf}}} c_m^k(\mathbf{R})|m(\mathbf{R};[Z])\rangle \\ &= \sum_m^{N_{\mathrm{csf}}}\left(\frac{\partial}{\partial x}c_m^k(\mathbf{R})\right)|m(\mathbf{R};[Z])\rangle + \sum_m^{N_{\mathrm{csf}}} c_m^k(\mathbf{R})\frac{\partial}{\partial x}|m(\mathbf{R};[Z])\rangle\end{aligned} \tag{166}$$

The CSF expansion in eq 166 can be either an MCSCF expansion or a general MRCI expansion, with the orbitals optimized for either single-state MCSCF or state-averaged MCSCF energies. The following equations are the same in any case. Equation 166 allows the nonadiabatic coupling element to be written as two contributions

$$f_{jk}(\mathbf{R})^x = f_{jk}^{ci}(\mathbf{R})^x + f_{jk}^{\mathrm{csf}}(\mathbf{R})^x \tag{167}$$

with

$$\begin{aligned}f_{jk}^{ci}(\mathbf{R})^x &= \sum_{m,n}^{N_{\mathrm{csf}}} c_m^j(\mathbf{R})\left(\frac{\partial}{\partial x}c_n^k(\mathbf{R})\right)\langle m(\mathbf{R};[Z])|n(\mathbf{R};[Z])\rangle \\ &= \mathbf{c}^j(\mathbf{R})\cdot\mathbf{c}^k(\mathbf{R})^x\end{aligned} \tag{168}$$

$$f_{jk}^{\mathrm{csf}}(\mathbf{R})^x = \sum_{m,n}^{N_{\mathrm{csf}}} c_m^j(\mathbf{R})\,c_n^k(\mathbf{R})\langle m(\mathbf{R};[Z])|\hat{X}|n(\mathbf{R};[Z])\rangle\rangle \tag{169}$$

with

$$\begin{aligned}\hat{X}^{[Z]}(\mathbf{R}) &= \sum_{p,q} X_{pq}^{[Z]}(\mathbf{R})E_{pq} \\ X_{pq}^{[Z]}(\mathbf{R}) &= \int \varphi_p^{[z]}(\tau;\mathbf{R})\frac{\partial}{\partial x}\varphi_q^{[z]}(\tau;\mathbf{R})\,\mathrm{d}\tau\end{aligned} \tag{170}$$

These two contributions to the nonadiabatic coupling element are examined separately. Differentiating eq 153 with respect to a coordinate and evaluation at **R** = 0 gives an expression for the first-order response of the CSF expansion coefficients to a perturbation

$$\begin{aligned}&(\mathbf{H}^{[Z]}(0) - E_k^{ci}(0)\mathbf{1})\mathbf{c}^k(0)^x \\ &\quad = -(\mathbf{H}^{[Z]}(0)^x - E_k^{ci}(0)^x\mathbf{1})\mathbf{c}^k(0)\end{aligned} \tag{171}$$

Multiplication from the left by $\mathbf{c}^j(0)^{\mathrm{T}}$ results in

$$\begin{aligned}f_{JI}^{ci}(0)^x &= (E_k^{ci}(0) - E_j^{ci}(0))^{-1}\mathbf{c}^j(0)^{\mathrm{T}}\,\mathbf{H}^{[Z]}(0)^x\,\mathbf{c}^k(0) \\ &= (E_k^{ci}(0) - E_j^{ci}(0))^{-1}\left(\mathrm{Tr}(\mathbf{h}^{[Z]}(0)^x\,\mathbf{D}^{jk,ci}(0)) + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[Z]}(0)^x\,\mathbf{d}^{jk,ci}(0))\right)\end{aligned} \tag{172}$$

$$\begin{aligned}&= (E_k^{ci}(0) - E_j^{ci}(0))^{-1}\Big(\mathrm{Tr}(\mathbf{h}^{[S]}(0)^x\,\mathbf{D}^{jk,ci}(0)) \\ &\quad + \frac{1}{2}\mathrm{Tr}(\mathbf{g}^{[S]}(0)^x\,\mathbf{d}^{jk,ci}(0)) + \mathbf{k}^{mc}(0)^x\cdot\mathbf{f}^{jk,ci}(0) + \mathbf{z}^{mc}(0)^x\cdot\mathbf{f}^{jk,ci}(0)\Big)\end{aligned} \tag{173}$$

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

Using the variational nature of $\mathbf{c}^j(\mathbf{R})$ in this manner is analogous to the Hellmann−Feynmann theorem for expectation values in that it avoids the explicit computation of $\mathbf{c}^k(0)^x$ in eq 171; if computed, this would be relatively expensive, and also the effort would be proportional to N_{atom}. Equation 173 is analogous to eq 156, but it uses the symmetrized transition density matrices $\mathbf{D}^{jk,ci}(0)$ and $\mathbf{d}^{jk,ci}(0)$ in place of the state-specific CI density matrices in the trace expressions and in the effective orbital gradient vectors $\mathbf{f}^{jk,ci}(0)$. As with the MCSCF and CI gradients discussed in the previous section, these density matrices are the displacement-independent factors, whereas the derivative integrals are the displacement-dependent factors in this expression. The expression in eq 173 can be transformed to the AO basis using the same sequence of steps as in the MCSCF and CI gradients, but with the effective orbital gradient, Fock matrices, and all other quantities written in terms of the symmetric transition density matrices rather than the state-specific density matrices in the various orbital basis sets. Before considering this transformation further, the other contribution to the nonadiabatic coupling element is examined.

The $f_{jk}^{\text{csf}}(\mathbf{R})^x$ coupling of eq 169 may be written in the $\boldsymbol{\varphi}^{[Z]}(\mathbf{R})$ orbital basis at $\mathbf{R} = 0$ as

$$f_{jk}^{\text{csf}}(0)^x = \sum_{p,q} X_{pq}^{[Z]} \langle ci(0)|E_{pq}|ci^k(0)\rangle \tag{174}$$

The relation $0 = S_{pq}^{[Z]}(0)^x = X_{pq}^{[Z]}(0) + X_{qp}^{[Z]}(0)$ shows that the orbital matrix $\mathbf{X}^{[Z](0)}$ is skew-symmetric. This allows the $f_{jk}{}^{\text{csf}}(0)^x$ coupling to be written in terms of the skew-symmetric CI one-particle density

$$\begin{aligned} f_{jk}^{\text{csf}}(0)^x &= \sum_{p,q} \frac{1}{2} X_{pq}^{[Z]}(0) \langle ci^j(0)|\hat{E}_{pq} - \hat{E}_{qp}|ci^k(0)\rangle \\ &= \sum_{p,q} X_{pq}^{[Z]}(0)\, D_{qp}^{(-)jk}(0) \\ &= \text{Tr}(\mathbf{X}^{[Z]}(0)\, \mathbf{D}^{(-)jk}(0)) \end{aligned} \tag{175}$$

Using eqs 124−127 and evaluation at $\mathbf{R} = 0$ give

$$\begin{aligned} X_{pq}^{[Z]}(0) &= X_{pq}^{[C]}(0) - \frac{1}{2} S_{pq}^{[C]}(0)^x + K^{mc}(0)_{pq}^x + Z^{mc}(0)_{pq}^x \\ &= \frac{1}{2}(X_{pq}^{[C]}(0) - X_{qp}^{[C]}(0)) + K^{mc}(0)_{pq}^x + Z^{mc}(0)_{pq}^x \end{aligned} \tag{176}$$

Because of the disjoint partitioning of the MCSCF essential and redundant orbital rotation elements, only one of the last two terms in eq 176 can be nonzero for a given orbital index pair (pq). Equation 175 can then be written as

$$\begin{aligned} f_{jk}^{\text{csf}}(0)^x &= \text{Tr}(\mathbf{X}^{[C]}(0)\, \mathbf{D}^{(-)jk}(0)) \\ &+ \text{Tr}(\mathbf{K}^{mc}(0)^x\, \mathbf{D}^{(-)jk}(0)) + \text{Tr}(\mathbf{Z}^{mc}(0)^x\, \mathbf{D}^{(-)jk}(0)) \\ &= \text{Tr}(\mathbf{X}^{[C]}(0)\, \mathbf{D}^{(-)jk}(0)) + \mathbf{k}^{mc}(0)^x \cdot \mathbf{f}_{\text{orb}}^{jk,D}(0) \\ &+ \mathbf{z}^{mc}(0)^x \cdot \mathbf{f}_{\text{orb}}^{jk,D}(0) \end{aligned} \tag{177}$$

with $f_{\text{orb}}^{jk,D}(0)_{(pq)} = 2D_{qp}^{(-)jk}(0)$. Upon comparing eq 177 with eq 173, the common factors can be combined to give

$$\begin{aligned} f_{jk}(0)^x &= f_{jk}^{ci}(0)^x + f_{jk}^{\text{csf}}(0)^x = \text{Tr}(\mathbf{X}^{[C]}(0)\, \mathbf{D}^{(-)jk}(0)) \\ &+ (E_k^{ci}(0) - E_j^{ci}(0))^{-1} \Big(\text{Tr}(\mathbf{h}^{[S]}(0)^x\, \mathbf{D}^{jk,ci[Z]}(0)) \\ &+ \frac{1}{2}\text{Tr}(\mathbf{g}^{[S]}(0)^x\, \mathbf{d}^{jk,ci[Z]}(0)) \Big) + (\mathbf{k}^{mc}(0)^x \\ &+ \mathbf{z}^{mc}(0)^x) \cdot ((E_k^{ci}(0) - E_j^{ci}(0))^{-1} \mathbf{f}^{jk,ci}(0) + \mathbf{f}^{jk,D}(0)) \\ &= \text{Tr}(\mathbf{X}^{[C]}(0)\, \mathbf{D}^{(-)jk}(0)) + \text{Tr}(\mathbf{h}^{[S]}(0)^x\, \mathbf{D}^{jk,\text{eff}}(0)) \\ &+ \frac{1}{2}\text{Tr}(\mathbf{g}^{[S]}(0)^x\, \mathbf{d}^{jk,\text{eff}}(0)) + \mathbf{k}^{mc}(0)^x \cdot \mathbf{f}_{\text{orb}}^{jk,\text{eff}}(0) \\ &+ \mathbf{z}^{mc}(0)^x \cdot \mathbf{f}_{\text{orb}}^{jk,\text{eff}}(0) \end{aligned} \tag{178}$$

where the effective density matrices and orbital gradient vector include the CI energy-difference factors as appropriate. The transformation steps of eqs 146−152 may be applied to transform this expression to the AO basis in which the derivative integral sparseness may be exploited.

$$\begin{aligned} f_{jk}(0)^x &= \text{Tr}(\mathbf{h}^{[\chi]}(0)^x\, \mathbf{D}^{jk,\text{total}[\chi]}(0)) + \frac{1}{2}\text{Tr}(\mathbf{g}^{[\chi]}(0)^x\, \mathbf{d}^{jk,\text{total}[\chi]}(0)) \\ &- \text{Tr}(\mathbf{S}^{[\chi]}(0)^x\, \mathbf{F}^{jk,\text{total}[\chi]}(0)) + \text{Tr}(\mathbf{X}^{[\chi]}(0)^x\, \mathbf{D}^{(-)jk[\chi]}(0)) \end{aligned} \tag{179}$$

Given the symmetric transition density matrices and effective orbital gradient vectors, the steps of the analytic energy gradient procedure may be applied, and the contributions from the first three terms in eq 179 may be computed in a straightforward manner. The last term in eq 179 is unique to the nonadiabatic coupling element. However, it involves only the skew-symmetric component of the one-particle transition CI density matrix, which requires an insignificant additional effort to compute along with the symmetric component which is used in the first terms. The above procedure may be applied to both MCSCF and general MRCI wave functions. As a practical matter, only the states with the largest coupling are included in the electronic basis. These states are the ones that are nearly degenerate and for which the energy difference factors in eq 178 have the largest magnitudes.

The effort required for all $3N_{\text{atom}}$ components x of the nonadiabatic coupling vector for MRCI wave functions is almost exactly the same as that required for the computation of a single energy gradient vector, which in turn is typically only a small fraction of the effort required for the energy and wave function optimization steps. Thus, for all practical purposes, if the wave functions and energies can be optimized for the states of interest, then the energy gradient vectors and the nonadiabatic coupling vectors can also be computed.

2.7. Relativistic Effects and Spin-Orbit Interaction

The accurate quantum-chemical treatment of molecules containing heavy atoms must account for relativistic effects to describe their properties. See, e.g., refs 470−472 for a general discussion of relativistic effects. Relativistic electrons whose energy is a sizable fraction of the rest mass $(m_e c^2)$ will have less kinetic energy than their nonrelativistic counterparts $(p^2/2m_e)$. Thus, the strong Coulomb interaction with the nuclei results in a contraction of the (nonrelativistic) orbitals close to the nuclei, which primarily affects the s and p orbitals, and, to lesser extent, the higher angular momentum orbitals. The more efficient

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

screening of the nuclear charge due to the contraction of the core orbitals typically causes the outer valence d and f orbitals to expand.

Spin-orbit coupling lifts the degeneracy of atomic orbitals belonging to the same l quantum number, thereby splitting the energy levels of the nonrelativistic atomic states. This introduces irregularities of the atomic properties in the periodic table. A good example is the anomalies of gold.[473] Especially for atoms with a multitude of low-lying (degenerate and quasidegenerate) electronic states such as actinides, spin-orbit coupling results in a high density of states in the vicinity of the electronic ground state. In fact, due to the coupling between electron correlation and relativistic effects, it is frequently difficult to unambiguously assign the ground state on theoretical grounds as illustrated by the dimers of actinides.[474] Hence, the coupling of electron-correlation and relativistic effects is in general best described using multireference methods. Since spin-orbit splitting may be large in free atoms but is usually quenched in the molecule, bond energies are strongly affected. The classical example is Tl_2 where spin-orbit coupling reduces the dissociation energy from 1.05 to 0.43 eV.[475] The importance of relativistic effects for molecular properties was pioneered by Pyykkö and Desclaux[476−478] and subsequently explored by many research groups (see, e.g., refs 471, 479 480, and 566).

2.7.1. Relativistic Hamiltonians. The starting point for almost all relativistic Hamiltonians for quantum chemistry is the Dirac equation, the relativistic quantum mechanical description of a one-electron system in an external scalar potential V. The Dirac equation with the Hamiltonian H_D, shifted by $-m_ec^2$, in position space representation is given by

$$H_D\Psi = (c\boldsymbol{\alpha}\cdot\mathbf{p} + (\boldsymbol{\beta} - 1)m_ec^2 + V]\Psi = E\Psi \qquad (180)$$

The eigenfunction Ψ is a four-component vector containing two "large" and two "small" components Ψ^L and Ψ^S, respectively.

$$\Psi = \begin{pmatrix} \Psi^L \\ \Psi^S \end{pmatrix} = \begin{pmatrix} \Psi^L_1 \\ \Psi^L_2 \\ \Psi^S_1 \\ \Psi^S_2 \end{pmatrix} \qquad (181)$$

The large and small components, respectively, originate from the electronic and positronic degrees of freedom. In the nonrelativistic limit $c\rightarrow\infty$ the small components vanish while the large components correspond to the nonrelativistic wave functions for α and β electron spin. $\mathbf{p}$ is the familiar momentum vector and $\boldsymbol{\alpha}$ is a three-dimensional vector. Together with $\boldsymbol{\beta}$ its components $\boldsymbol{\alpha}_x$, $\boldsymbol{\alpha}_y$, and $\boldsymbol{\alpha}_z$ constitute the four Dirac matrices

$$\boldsymbol{\alpha}_x = \begin{pmatrix} 0 & 0 & 0 & +1 \\ 0 & 0 & +1 & 0 \\ 0 & +1 & 0 & 0 \\ +1 & 0 & 0 & 0 \end{pmatrix} \quad \boldsymbol{\alpha}_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & +i & 0 \\ 0 & -i & 0 & 0 \\ +i & 0 & 0 & 0 \end{pmatrix} \qquad (182)$$

$$\boldsymbol{\alpha}_z = \begin{pmatrix} 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & -1 \\ +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad \boldsymbol{\beta} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad (183)$$

The spectrum of the Dirac equation contains a negative $(-\infty,-2m_ec^2)$ and a positive $(0,+\infty)$ continuum of scattering type solutions along with the discrete spectrum of eigenstates in between. The off-diagonal term in eq 180, $c\boldsymbol{\alpha}\cdot\mathbf{p}$, introduces a coupling between the small and the large components. Ψ^L and Ψ^S are related to each other by the (exact) kinetic balance condition.[481,482] This causes the basis set of the small component to be about twice the size of that for the large component. In the physical vacuum the negative continuum is completely filled by electrons while all other states are unoccupied.

There is no unique derivation of a molecular many-electron analogue to the Dirac equation. Most commonly, the Dirac−Coulomb, H_{DC}, and the Dirac−Coulomb−Breit Hamiltonian, H_{DCB}, are used in practice, where the Breit term, derived by perturbation theory within the framework of QED, is taken in the frequency-independent form[483−485] (in atomic units)

$$H_{DCB} = H_{DC} - \sum_{j>k}\left[\frac{\boldsymbol{\alpha}_j\cdot\boldsymbol{\alpha}_k}{2r_{jk}} + \frac{(\boldsymbol{\alpha}_j\cdot\mathbf{r}_{jk})(\boldsymbol{\alpha}_k\cdot\mathbf{r}_{jk})}{2r_{jk}^3}\right] \qquad (185)$$

Within the algebraic approximation the presence of a negative energy continuum results in the tendency of variational solutions to "collapse". This continuum dissolution problem is characteristic for a Hamiltonian whose spectrum is not bounded from below. With finite GTO basis set expansions, however, there is not necessarily sufficient flexibility for the collapse into the negative continuum to be observed. Yet lacking a bounded Hamiltonian, computed energies and wave functions do not necessarily correspond to the desired bounded fermion states. To circumvent this problem, it is necessary to construct a basis that distinguishes between positron and electron states. Thus, the first step is the construction of a four-spinor basis by computing the eigenvalues of a simplified independent particle Hamiltonian H_0 and classifying them into positronic and electronic spinors. This information can subsequently be used to constrain the wave function optimization to the interaction of electron states only, since positronic excitations and pair-creation processes are usually of no interest. The no-pair Hamiltonian depends on the choice of H_0 used to define the four-spinor basis because it is not invariant to rotations between positronic and electronic spinors (vacuum polarization).

Within the fully relativistic four-component formalism, Dirac[486] and Kramers restricted Dirac−Hartree−Fock SCF calculations furnish the necessary four-spinor basis. As already mentioned, the kinetic balance condition defines a relation between the basis sets for the small and large components. While primitive basis sets are best suited to ensure the kinetic balance condition, contracted basis sets reduce the computational costs substantially. Thus, to generate contracted basis sets, the atomic balance procedure has been proposed.[487] Here the contraction coefficients are derived from the spinor coefficients of atomic DHF calculations with kinetically balanced primitive basis sets, thereby assuming that core regions are not affected by bond formation. To retain sufficient flexibility to describe the different radial character of the $l \pm 1/2$ spinors, the size of the contracted basis essentially doubles. Similar to nonrelativistic Hartree−Fock, the energy of a single determinant is minimized with respect to rotations between occupied and virtual electron spinors plus the constraint that the energy should be maximized for rotations between electron and positron spinors.[488] The formalism can be extended to MCSCF.[489] The optimized four-spinor bases are

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

subsequently used for more accurate electron correlation treatments. Due to the integration over four spinors, the computational resource and memory requirements are substantially higher. There are 16 times as many complex two-electron integrals compared to nonrelativistic calculations, in addition to the need for large primitive basis sets for a sufficiently flexible description of the relativistic wave function. Hence, much effort has been devoted to develop a hierarchy of approximations to the four-component DCB Hamiltonian.

Two-component approaches can be derived by removing the small component of the four spinors. Elimination techniques exploit the fact that, for electronic solutions, the Ψ^S is suppressed by a factor of $1/(2m_ec)$ with respect to Ψ^L. Zero- and first-order regular approximations[490] (ZORA, FORA) and the method of normalized elimination of the small component[491] (NESC) belong to this group of approaches.

Alternatively, transformation techniques aim at a unitary transformation U of the Dirac Hamiltonian to block-diagonal form such that Ψ^S and Ψ^L (in an eigenspinor basis of H_D) are decoupled

$$\tilde{H} = UH_DU^\dagger = \begin{pmatrix} h_+ & 0 \\ 0 & h_- \end{pmatrix} \tag{186}$$

$$\tilde{\Psi} = U\Psi = \begin{pmatrix} \widetilde{\Psi^L} \\ \widetilde{\Psi^S} \end{pmatrix} \tag{187}$$

For electronic solutions, $\tilde{\Psi}^S = 0$, the upper-left 2×2 block (h_+) only is retained. The generalized Douglas−Kroll transformation[492] employs a sequence of unitary transformations successively eliminating the off-diagonal 2×2 block in orders of the external potential.

$$\tilde{H} = ...U_2U_1U_0H_DU_0^\dagger U_1^\dagger U_2^\dagger ... = \sum_{k=0}^{\infty} V_k \tag{188}$$

$$= \sum_{k=0}^{\infty} \begin{pmatrix} V_{k+}^{sf} + V_{k+}^{sd} & 0 \\ 0 & V_{k-}^{sf} + V_{k-}^{sd} \end{pmatrix} \tag{189}$$

The Douglas−Kroll method yields variationally stable, well-defined expressions and allows for systematic improvement of the approximation.[493] $\tilde{H}$ is well-defined only in momentum space, and the operators contain only even powers of momentum. The (scalar) DKH approximation is applied to the one-electron operators only, with spin-dependent terms V_{k+}^{sd} discarded. Most conveniently, the modified kinetic energy and electron−nuclear attraction integrals are evaluated in the basis of the eigenvalues of the kinetic energy matrix and subsequently transformed to position space.[492] In an incomplete basis, this amounts to an additional approximation. From the computational point of view, this method is very attractive, as it combines well with the entire existing machinery of nonrelativistic quantum chemistry. Restricting the scalar DKH transformation to the one-electron part neglects a renormalized two-electron Darwin term for which the integrals should be small.[491]

The Breit−Pauli Hamiltonian[484] incorporates spin-orbit coupling in the reduced basis of the large component, and the spin-orbit component is given by

$$H^{BPSO} = \frac{1}{\alpha^2}\left[\sum_{jA} \frac{Z_A}{r_{jA}^3}(\mathbf{l}_j \cdot \mathbf{s}_j)\right] - \sum_{j \neq k} \frac{1}{r_{jk}^3}(\mathbf{r}_{jk} \times \mathbf{p}_j) \cdot (\mathbf{s}_j + 2\mathbf{s}_k) \tag{190}$$

$$= \frac{1}{\alpha^2}\left[\sum_{jA} \frac{Z_A}{r_{jA}^3}(\mathbf{l}_j \cdot \mathbf{s}_j)\right] - \sum_{j \neq k} \frac{1}{r_{jk}^3} \mathbf{l}_{jk} \cdot (\mathbf{s}_j + 2\mathbf{s}_k) \tag{191}$$

$$= \sum_j H^{SO}(j) + \sum_{jk} H^{SOO}(j,k) \tag{192}$$

where α is the fine-structure constant. The first term is the *spin-orbit* interaction, while the second term is denoted the *spin-other-orbit* interaction. Including the Breit interaction in the DKH transformation up to second order in V, a relativistically corrected no-pair Hamiltonian can be derived whose spin-dependent part is of the same form as H^{BPSO} amended by kinematic bracketing factors.[494] The nonvanishing matrix elements of the spin-orbit operator for a spin-orbital excitation are given by

$$\begin{aligned} H_{ja}^{SO} = {} & \langle \varphi_j(1)|H^{SO}|\varphi_a(1)\rangle \\ & + \frac{1}{2}\sum_k n_k[\langle \varphi_j(1)\,\varphi_k(2)|H^{SOO}(1,2)|\varphi_a(1)\,\varphi_k(2)\rangle \\ & - \langle \varphi_j(1)\,\varphi_k(2)|H^{SOO}(1,2)|\varphi_k(1)\,\varphi_a(2)\rangle \\ & - \langle \varphi_k(1)\,\varphi_j(2)|H^{SOO}(1,2)|\varphi_a(1)\,\varphi_k(2)\rangle] \end{aligned} \tag{193}$$

where n_k denotes the occupancy of orbital k. To reduce the computational effort, the two-electron part contributes through fixed predefined average atomic densities, and, due to the short-range property of the spin-orbit two-electron operator, only one-center terms are retained.[495] More recently, also the one-electron integrals are restricted to one-center integrals.[496] The dependence on the choice of the atomic densities and the neglect of the multicenter spin-orbit integrals appears to be small.

In the above, all-electron approximations to the fully relativistic treatment have been discussed. Further reduction of the computational effort and the complexity of the approach can be achieved by use of effective valence electron Hamiltonians which include relativistic effects solely by means of suitably parametrized core potentials. Both one- and two-component effective core potentials (ECPs) use a nonrelativistic model Hamiltonian, and relativistic contributions arise solely from the parametrization of the core−valence potential.

The ab initio model potential (AIMP) method[497] aims at reproducing atomic frozen-core calculations by replacing the Fock operator of a valence electron in the field of the core electrons

$$\left[\frac{Z'}{r_j} + \sum_c 2J_c^{core}(j)\right] - \sum_c^{core} K_c(j) = V_c(j) + V_x(j) \tag{194}$$

by a local expansion for the Coulomb part

$$V_c(j) \approx \frac{1}{r_j}\sum_k A_k e^{-\alpha_k r_j^2} \tag{195}$$

where the atom-specific parameters $\{\alpha_k, A_k\}$ are adjusted in the least-squares sense to the all-electron Coulomb potential. A spectral representation is used for the exchange contribution

$$K_c(j) \approx \sum_{pq} |\chi_p(j)\rangle A_{pq} \langle \chi_q(j)| \tag{196}$$

Due to the short-range property of V_x a moderate basis $\{\chi\}$ should suffice and the one-center approximation is expected to be good. To prevent the valence orbitals from collapsing into the core

region, the operator $P(j) = -\sum_c^{\text{core}} 2\varepsilon_c|\varphi_c(j)\rangle\langle\varphi_c(j)|$ shifts the core orbitals well above the valence orbitals to retain (approximate) orthogonality of the core and valence orbitals. The molecular AIMP Hamiltonian is defined by superposition of the atomic AIMP Hamiltonians plus the core−core repulsion approximated as point−charge interactions.

$$H^{\text{AIMP}} = \sum_j^{n_v}\left[-\frac{1}{2}\nabla_j^2 + V_c(j) + V_x(j) + P_c(j)\right] + \sum_{j>k}^{n_v}\frac{1}{r_{jk}} + \sum_{A>B}\frac{Z'_A Z'_B}{R_{AB}} \quad (197)$$

Keeping the AIMP parameters fixed, the valence basis can now be optimized in the standard procedure. The molecular Hamiltonian neglects many-center core−valence exchange, while retaining the nodal structure of the valence orbitals, provided the valence basis set contains also sufficiently steep core functions. This makes AIMP calculations more expensive, while opening the possibility to explicitly include relativistic effects for the valence electrons in the framework of the DKH no-pair Hamiltonian[498] by use of the DKH transformed kinetic energy and electron−nuclear attraction integrals. This may be extended to spin-orbit coupling within the framework of the AMFI approach.[499] For computational reasons, using the smaller scalar relativistic ECPs in combination with AMFI for spin-orbit interaction is attractive. A simple procedure based on the equivalence of AMFI integrals from all-electron and ECP calculations using suitably matched basis sets has been proposed.[500,501]

The model core potential (MCP) method[502,503] differs from the AIMP approach by omitting the spectral representation of the exchange operator V_x and by expanding the local potential V_c in terms of radial Gaussian functions. The parametrization of V_c also implicitly includes the exchange term. Since MCPs share the same basic features of AIMPs, scalar-relativistic contributions and spin-orbit coupling have been recently incorporated in a spirit similar to the afore-mentioned AMFI approach.[504] Promising results for hydrides and cationic dimers for *p*-block elements up to Rn at the SO complete active space CI (SO-CASCI)[505] and the SO multiconfigurational quasidegenerate perturbation theory (SO-MCQDPT)[506] level of theory have been reported.[504,507]

An alternative form of an effective valence-electron Hamiltonian is the use of pseudopotentials, where the atomic Hamiltonian is given by

$$H = -\frac{1}{2}\sum_j \nabla_j^2 + \sum_j V_{PP}(\mathbf{r}_j) + \sum_{j>k}\frac{1}{r_{jk}} \quad (198)$$

$$V_{PP}(\mathbf{r}) = -\frac{Z'}{r} + \sum_{ljk} B_{lj}^k \mathrm{e}^{-\beta_{lj}^k r^2} P_{lj} = V_{PP}^{\text{av}} + V_{PP}^{\text{SO}} \quad (199)$$

$$P_{lj} = \sum_{m=-j}^{j} |ljm\rangle\langle ljm| \quad (200)$$

The kinetic energy and electron−electron interaction are nonrelativistic; i.e., all relativistic effects are included in the pseudopotential. The long-range behavior is determined by the core charge; the short-range behavior is modeled in a semilocal ansatz with the projection operator P_{lj} inducing different radial potentials for different $j = l \pm 1/2$.

Scalar relativistic effects are approximated by the *j*-averaged potential V_{PP}^{av} while the spin-orbit part is the difference from the averaged potential. Both are usually expanded in Gaussians.

$$V_{PP}^{\text{av}}(\mathbf{r}) = \frac{Z}{r} + \sum_{lk}\left[\frac{l}{2l+1}B_{l,l-1/2}^k \mathrm{e}^{(-\beta_{l,l-1/2}^k r^2)} + \frac{l+1}{2l+1}B_{l,l+1/2}^k \mathrm{e}^{(-\beta_{l,l+1/2}^k r^2)}\right]P_l \quad (201)$$

$$V_{PP}^{\text{SO}}(\mathbf{r}) = \sum_{lk}\frac{2}{2l+1}[B_{l,l+1/2}^k \mathrm{e}^{(-\beta_{l,l+1/2}^k r^2)} - B_{l,l-1/2}^k \mathrm{e}^{(-\beta_{l,l-1/2}^k r^2)}]P_l \mathbf{l}\cdot\mathbf{s} P_l \quad (202)$$

Shape-consistent pseudopotentials[508,509] replace the valence orbitals of the relativistic all-electron calculation by nodeless pseudoorbitals which retain the shape of the valence orbital beyond a matching radius r_c separating spatially the valence and core region, while inside this radius the pseudoorbital is described by a smooth nodeless polynomial expansion. The pseudopotential and basis set to describe the pseudoorbitals are tightly coupled. Ambiguities arise from the choice of the reference data. Some of the popular sets include the following: (i) Christiansen and co-workers[510−519] generated shape-consistent pseudopotentials including spin-orbit potentials derived from DCHF all-electron calculations; (ii) Hay and Wadt[520−524] derived another popular set based on scalar-relativistic Cowan−Griffin all-electron calculations; (iii) Stevens and co-workers[525−527] compiled a more compact representation also based on DCHF all-electron calculations.

Energy-consistent pseudopotentials[528−530] fit the adjustable parameters on the basis of least-squares deviations of atomic energy levels with respect to relativistic all-electron calculations. This approach has the advantage that it does not rely on reproducing quantities in the one-particle picture, and the formalism can be used to generate pseudopotentials at any level of relativity approximation. Given that the optimization of the pseudopotential parameters is carried out close to the basis set limit or using accurate numerical atomic calculations, the subsequent basis set optimization with fixed parameters is straightforward. Parameters and corresponding valence basis sets are available for almost all elements in the periodic table both for small and large cores.[528,529,531−546]

2.7.2. Relativistic Implementations. Currently, not only are nonrelativistic methods in quantum chemistry much more developed than their two- and four-component relativistic counterparts, but they are also computationally much less expensive. This is to some extent related to requirements for the construction of the one- and many-particle basis, but it is also apparent that the underlying machinery is quite different when comparing relativistic and nonrelativistic methods. Since this is closely related to the different symmetry properties of the spin-orbit term in the following paragraph, specific features are discussed in order to clarify the more detailed comparison of the various implementations.

Although the Hamiltonian is scalar and thus rotationally invariant, the spin-orbit coupling term contains the scalar product of two vector operators. While nonrelativistic and scalar relativistic Hamiltonians preserve the symmetry of space and spin coordinates separately, the $\mathbf{l}\cdot\mathbf{s}$ term commutes only with symmetry operations applied simultaneously to both spatial and spin coordinates. Thus the symmetry properties are no longer described by the regular (single) point groups. To account for spin 1/2 transformation properties, the double groups are

introduced. Half-integer spin functions transform according to the $D^{1/2}$ irrep of the full rotation group such that a rotation by 2π leads to a sign change while a 4π rotation is the unit operation. The double group may be derived from the regular (single) group by adding the extra symmetry element $\tilde{E}$ representing a rotation by 2π. Thus the number of symmetry operations is doubled, and in addition to the regular (bosonic, even number of electrons) irreps describing the transformation properties of integer angular momentum functions, fermionic (odd number of electrons) irreps describe the transformation properties of half-integer angular momentum functions.[547,548]

In the absence of an external magnetic field, the Hamiltonian also commutes with the time-reversal symmetry operator (Kramers operator). By virtue of Kramers theorem, wave functions with integer spin momentum may be chosen to be real. Half-integer (fermion) functions, however, are 2-fold degenerate.[549] The degenerate components are related by Kramers operator (K) forming a Kramers pair (ϕ, $\bar{\phi}$).

$$K\phi = \bar{\phi}; \quad K^2\phi = -\phi \tag{203}$$

The group-theoretical properties of the Kramers pairs are connected to the structure of the fermionic irreps of the double group. For the odd-electron case, three different cases are distinguished depending upon the transformation properties of the Kramers pair functions: (i) in real groups they span the basis of a two-dimensional fermion irrep (e.g., D_{2h}^*, C_{2v}^*, D_2^*); (ii) in complex groups the components transform as two different fermion irreps (C_{2h}^*, C_s^* C_2^*); (iii) in quaternionic groups the components transform as the same one-dimensional irrep (C_1^*, C_i^*). Thus, given a hermitian one-electron operator h symmetric under time-reversal and spatial symmetry, in a Kramers basis the off-diagonal block $h_{p\bar{q}}$ vanishes for real and complex groups, whereas a quaternionic transformation is necessary to have case iii block diagonal. For real groups, the Hamiltonian and the wave function may be chosen to be real,[550−552] whereas they are complex in cases ii and iii. Thus time-reversal symmetry may be exploited reducing the overall effort, as well as to classify the algebra (real, complex, quaternionic) in the general case. Hence, to incorporate double group and time-reversal symmetry, one may work in a Kramers basis of one-electron spinors and rewrite the Hamiltonian in terms of Kramers single- and double-replacement operators.[489] This results already in a considerable reduction in the number of two-electron integrals. This carries over to the classification of determinants according to the number of occupied Kramers pair functions, $N(p)$ and $N(\bar{p})$, and the projection $M_k = (^1/_2)(N(p) - N(\bar{p}))$. The Hamiltonian matrix in a basis of these determinants displays a block structure due to the selection rule $H_{IJ} = 0$ for $|\Delta M_k| = |M^I{}_k - M^J{}_k| > 2$ and the algebraic classification (real, complex, and quaternionic).[527] Within this formalism, the nature of the spin function in the one-electron spinor is not restricted to either α or β spin but is a general linear combination thereof. Since the spin-orbit term couples determinants of different spin projections M_k, the underlying configuration space is usually limited to some maximum range of $|M_k - M_{k'}|$ with respect to a reference value $M_{k'}$ (see also ref 553).

An alternative procedure sets out with nonrelativistic spin- and spatial-symmetry-adapted many-electron functions, i.e., the union of all $(2S + 1)$ degenerate components for multiple spin multiplicities of all spatial symmetry irreps ($|j;S,M,\Gamma\rangle$). The nonrelativistic Hamiltonian matrix is block diagonal with respect to S, M, and Γ (cf. eqs 10 and 11). Adding (effective) one-electron spin-orbit coupling terms introduces off-diagonal blocks in the Hamiltonian matrix since nonzero matrix elements occur for $|\Delta S| \leq 1$ and $|\Delta M| \leq 1$, so that the Hamiltonian matrix has a complicated block structure.[549] The matrix elements of the spin-orbit coupling term can be evaluated by means of the Wigner-Eckart theorem, reducing the overall effort. However, this technique is limited to real orbitals, and in general the Hamiltonian matrix is complex, although the above group theoretical restrictions and simplifications apply.

Spin-orbit coupling affects the selection of the one-electron basis. A large difference between the radii $\langle r^2 \rangle$ of the pairs of spinors $l \pm 1/2$ from DHF calculations[554] suggests the need for an optimized two-spinor basis. Main group elements primarily fall into this category, while transition metals, lanthanides, and actinides may well work with a scalar-relativistically optimized (real) MO basis. Otherwise, an inadequate MO basis must be compensated for by a larger N-electron basis, i.e., including single excitations to (highly) excited states, in order to account for spin polarization of the average orbitals toward the $l \pm 1/2$ spinors. On the basis of experiences with two-component Kramers-restricted HF calculations using spin-orbit pseudopotentials to generate the j-specific spinors for subsequent two-component electron correlation calculations, it has been argued that the configuration space expansion is more compact and converges more rapidly.[555]

2.7.3. Scalar Relativistic Effects. In this section, some general aspects related to the extension of existing nonrelativistic MCSCF and MRCI methods to approximate two-component relativistic calculations are discussed.

Scalar relativistic effects approximated in terms of pseudopotentials, AIMPs, or using the DKH-no-pair Hamiltonians are completely transparent to any conventional MCSCF or MRCI code since only the one-electron integrals are affected. Separation of spatial and spin coordinates, and thus the related symmetry considerations, are unchanged as compared to nonrelativistic calculations. Whereas effective valence electron approaches (pseudopotentials, AIMPs) rely on the proper choice of the core region and the parametrization, all-electron treatments are free of these limitations. The DKH-based description of scalar relativistic effects requires basis sets capable of describing the core region, which leads to increased computational costs. Similar considerations apply to the AIMPs, as the nodal structure of the valence orbitals in the core region must be adequately reproduced. All-electron calculations can be systematically improved, a feature that does not apply to pseudopotential approaches, even with multiple parametrizations for different core sizes, since the transferability to the molecular case can be limited. Given the cost/accuracy ratio, pseudopotentials and, to lesser extent, AIMPs are very efficient.

2.7.4. Two-Component Extensions of MRCI. There are some technical and conceptual obstacles to incorporate spin-orbit coupling into traditional MRCI codes while retaining the simplifications due to the use of one-component (real) orbitals. This is equivalent to a Kramers basis with collinear spin (α or β). For any direct-CI method, it is important to evaluate the Hamiltonian matrix elements very efficiently on-the-fly, so that spin-averaged, real, molecular orbitals, a real Hamiltonian, and the use of existing nonrelativistic machinery for the compute-intense matrix−vector product formation are all important assets.

The evaluation of the spin-orbit matrix elements is split into the product of an orbital-dependent contribution and a spin-dependent contribution evaluated on-the-fly using partially tabulated data. The one-electron integrals have been implemented for more than 2 decades.[556] The rather numerous two-electron

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

spin-orbit integrals (with different symmetry properties) are either completely neglected with pseudopotentials or approximated within an effective one-electron operator (AMFI approximation[495]). However, in a real spin-adapted or determinant basis, the spin-orbit matrix elements are imaginary.

The dimension of the configuration space including spin-orbit coupling may be an order of magnitude larger than for the nonrelativistic counterpart. Since the efficient evaluation of Hamiltonian matrix elements is closely tied to the enumeration of the many-electron basis, the seamless integration of this superset into the enumeration scheme is quite important. GUGA-based implementations, for example, employ a multiheaded Shavitt graph in which each head represents a separate spin multiplicity.[557] Both CSF- and determinant-based expansions need to deal with multiple M values and spatial symmetries. To gain insight into the structure of the determinant space, graphical representations are helpful.[29]

For a complex Hermitian matrix $\mathbf{H} = \mathbf{P} + i\mathbf{Q}$, where $\mathbf{P}$ and $\mathbf{Q}$ are real-symmetric and skew-symmetric matrices, respectively, the corresponding eigenvalue equation is $\mathbf{Hz} = \mathbf{z}E$ for real E and complex $\mathbf{z} = \mathbf{u} + i\mathbf{v}$ for real components $\mathbf{u}$ and $\mathbf{v}$. If $\mathbf{z}$ is an eigenvector, then $e^{i\theta}\mathbf{z} = (\cos(\theta)\mathbf{u} - \sin(\theta)\mathbf{v}) + i(\sin(\theta)\mathbf{u} + \cos(\theta)\mathbf{v})$ is also an eigenvector for arbitrary phase θ. Thus the splitting between the real and imaginary components is somewhat arbitrary, and θ may be chosen freely to simplify the wave function analysis (e.g., to minimize the complex component norm or to maximize the component overlaps with some reference vector). The complex eigenvalue equation of dimension m may be expanded as

$$\mathbf{H}'\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix} = \begin{pmatrix}\mathbf{P} & -\mathbf{Q}\\ \mathbf{Q} & \mathbf{P}\end{pmatrix}\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix} = \begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix}E \qquad (204)$$

where $\mathbf{H}'$ is real symmetric with dimension $2m$. The vector $\begin{pmatrix}-\mathbf{v}\\ \mathbf{u}\end{pmatrix}$ is also an eigenvector of $\mathbf{H}'$ with the same eigenvalue E. In general, each of the eigenvalues of the complex $\mathbf{H}$ are replicated[558] in the eigenvalue spectrum of the expanded $\mathbf{H}'$. An arbitrary plane rotation of the two degenerate eigenvectors is also an eigenvector, $\cos(\theta)\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix} + \sin(\theta)\begin{pmatrix}-\mathbf{v}\\ \mathbf{u}\end{pmatrix}$, which is seen to be simply another expression for the arbitrary complex phase of the original complex eigenvector. Thus, for the pair of degenerate eigenvectors of the expanded $\mathbf{H}'$, there exists only a single linearly independent complex eigenvector $\mathbf{z}$. For small dimensions m, the choice between these two representations is rather arbitrary, but for larger dimensions in which a direct diagonalization is used, the complex representation is about a factor of 2 more efficient ($\sim(4m^3)$ operations for $\mathbf{H}$ compared to $\sim(2m)^3$ operations for $\mathbf{H}'$) in addition to the larger storage requirements for the expanded $\mathbf{H}'$. For iterative subspace methods for only a few eigenpairs, the computational effort is comparable for the two approaches. The complex Hermitian generalization of the Davidson subspace method is straightforward. In the expanded representation, the matrix−vector products $\mathbf{Pu}$, $\mathbf{Pv}$, $\mathbf{Qu}$, and $\mathbf{Qv}$ are required within each iteration for trial vector components $\mathbf{u}$ and $\mathbf{v}$. These vectors can then be combined to increase the subspace dimension by two each iteration, equivalent to adding both expansion terms $\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix}$ and $\begin{pmatrix}-\mathbf{v}\\ \mathbf{u}\end{pmatrix}$; this requires no substantial additional effort, and the resulting subspace Hamiltonian has replicated Ritz values, mimicking the true eigenvalue spectrum of the expanded form $\mathbf{H}'$. In the case of additional degeneracies, e.g., the Π states in linear molecules, the four associated eigenvectors of the expanded $\mathbf{H}'$ may mix arbitrarily, and when combined into complex form, only two linearly independent complex $\mathbf{z}$ vectors can be constructed. The choice of representation of these two orthogonal complex eigenvectors can affect the continuity of the wave function representation and the computation of transition properties.[565]

The complex matrix $\mathbf{H}$ of dimension m can take a partitioned form $H = \begin{pmatrix}\mathbf{A} & -i\mathbf{C}^{\mathrm{T}}\\ i\mathbf{C} & \mathbf{B}\end{pmatrix}$ for real symmetric square $\mathbf{A}$ and $\mathbf{B}$ and for arbitrary real rectangular $\mathbf{C}$. In this situation, a unitary transformation of the form $\mathbf{H} \leftarrow \mathbf{U}^{\dagger}\mathbf{H}\mathbf{U}$ with $\mathbf{U} = \mathrm{diag}(\mathbf{1}, i\mathbf{1})$ will result in a real, symmetric $\mathbf{H} = \begin{pmatrix}\mathbf{A} & \mathbf{C}^{\mathrm{T}}\\ \mathbf{C} & \mathbf{B}\end{pmatrix}$ of dimension m that has the same eigenvalues as the original matrix. The eigenpairs may be determined with a real symmetric diagonalization in the straightforward way without complications due to replicated eigenvalues. The eigenvectors in the original basis representation are given by $\mathbf{U}^{\dagger}\mathbf{c}$. This property can generalize to certain matrices with multiple partitions. By using symmetry arguments, it can be shown[559] for the real groups (D_{2h}^{*}, C_{2v}^{*}, D_2^{*}) in a double-group adapted many-electron basis either that the complex Hamiltonian matrix can be brought to this real form or alternatively that the phases of the many-electron basis may be modified in order to make the corresponding Hamiltonian representation real. All computations may then be done with real arithmetic only. This applies to both even and odd electron cases, as expected for real groups from the discussion above. For even electrons, it can be shown that the spin functions in terms of real spherical tensors (linear combinations of $|k;S,M\rangle$ and $|k;S,-M\rangle$) form a symmetry-adapted basis for the double group, and, with the appropriate phase convention, all nonvanishing spin-orbit coupling matrix elements are real.[557] The odd-electron case can be adapted to the even-electron formalism by formally adding an additional, fictitious, noninteracting electron; this results in a doubling of the dimension of the Hamiltonian matrix, and in replicating the eigenvalues, since for complex and quaternionic groups the Hamiltonian is generally complex.

When dealing with highly degenerate open-shell cases such as actinides and lanthanide compounds, the sheer size of the potential configuration space raises the fundamental question, to what extent it is admissible to approximately decouple electron correlation and spin-orbit coupling effects in order to reduce the computational effort? Due to the substantially reduced complexity, as well as computational effort, two-step approaches have been popular from the very beginning.[560] The simplest approach amounts to computing a range of electronic states of different spin multiplicity or spatial symmetry including scalar relativistic effects.

There are two types of general approaches based upon modifications of the nonrelativistic MRCI procedure, (i) the simultaneous treatment of electron correlation and spin-orbit coupling on the same footing (one-step procedure) and (ii) the model space or effective Hamiltonian approach (two-step procedure). Treating spin-orbit coupling and electron correlation on the same footing, the one-step approach, is thus quite expensive and requires highly efficient programs. There are three major codes capable of such calculations. (i) The conventional CIDBG code by Pitzer and co-workers[556,559,561,562] which supports double groups is, due to the storage requirements for the CI matrix elements, limited to $\sim 10^5$ CSFs when used in combination with spin-orbit pseudopotentials; closely related is a variant of this approach using an additional configuration selection scheme (selected intermediate coupling CI).[563,564] (ii) The GUGA-based direct-CI code supporting double groups

by Yabushita et al.,[557] the parallel version of this code,[118] is capable of dealing with CSF expansions up to $\sim 10^9$. (iii) The individually selected direct-CI code SPOCK.CI[565] which uses a basis of spin-adapted CSFs without double group support but optionally in combination with the DFT/MRCI method.[642] In this last (empirical) method, dynamical electron correlation (and size-extensivity) is approximately accounted for by DFT while MRCI is responsible for static electron correlation. DFT MOs furnish the one-electron basis. To avoid double counting of electron correlation, the off-diagonal CI matrix elements are scaled by an empirical energy-dependent term while the diagonal elements are modified by Kohn−Sham matrix elements. The selection of near-degenerate configurations relies on an energy gap criterion. The DFT/MRCI Hamiltonian combination works very well with an individually selected CI code because individual Hamiltonian matrix elements may be easily modified in the direct-CI scheme, and there is an immediate computational gain with the substantial reduction of the MRCI expansion dimension. The parallel version of this method, however, seems to be not well parallelizable and, due to random memory access, suffers from cache misses.[641] A priori selected CI codes are not generally suitable because they define a particular expansion form for the wave function and exploit this to efficiently compute the CI matrix elements on-the-fly. Since the contributions of individual integrals are additive, CI matrix elements are constructed incrementally to exploit vectorization and retain only a minimum of logic. For example, the GUGA-based MR-CISD implementation in COLUMBUS can utilize DFT/MRCI efficiently, only if the entire CSF space is defined within the internal orbital space. Although this is a natural choice, the size of the internal orbital space may be large and some structural selection scheme must be devised to avoid the exponential scaling of the underlying full-CI space. Thus, the large combined CI space dimension in one-step spin-orbit CI is reduced, and calculations including spin-orbit coupling on larger molecules are possible.

For individually selected MRCI codes, the automated selection of configurations for variational treatment is critical in the case of spin-orbit CI because the usual selection criteria based on perturbation theory estimates of the correlation energy contributions would leave out the important single excitations because, due to Brillouin's theorem, their contribution to the correlation energy is insignificant. However, many of these single excitations are indispensable in order to compensate for the choice of spin-averaged real molecular orbitals which are biased as compared to the respective optimized two-component spinors. These single excitations allow for orbital relaxation effects relative to the spin-averaged MO basis. Thus, all singles, or a major subset of all singles, with respect to the already selected CSF space (e.g., by symmetry criteria or perturbation theory estimates) must be added in order to obtain reliable zero-field splitting. Although SPOCK.CI simultaneously treats electron correlation and spin-orbit coupling on the same footing, the initial contracted, effective, complex Hamiltonian is set up and diagonalized. This serves to select the CSF space, to generate starting vectors for the final, fully coupled, MR-SOCI step, and to obtain a quasidegenerate perturbation theory (QDPT) estimate of the spin-orbit effects (equivalent to the two-step approaches below).

Two-step approaches are characterized by some separation of spin-orbit coupling and electron correlation effects. Among the simplest approaches is to compute various $|\psi_k;S,M\rangle$ and to evaluate spin-orbit coupling by first- and second-order perturbation theory from the spin-orbit coupling matrix elements between pairs of these electronic states.[566] The set $\{|\psi_k;S,M\rangle\}$ may also include the spatial and spin-degenerate components of the scalar-relativistic electronic wave functions. Thus, this may be described as a two-step approach using contracted (scalar-relativistic) wave functions while ignoring any relaxation due to spin-orbit coupling.

An improved approach employs QDPT:[26] a model Hamiltonian in the contracted basis (of up to a few hundred) $\{|\psi_k;S,M\rangle\}$ is constructed, all off-diagonal spin-orbital coupling matrix elements $\langle\psi_j;S,M|H^{SO}|\psi_k;S',M'\rangle$ are evaluated, and this small, complex-hermitian, model Hamiltonian is diagonalized. The major limitation of this approach is that spin-orbit coupling allows for relaxation of the wave functions only within the small model space. Since an effective Hamiltonian is constructed, this allows for further variations. The SO-RASPT2 approach[475,567] assumes the separability of dynamical and static correlation; the contracted basis $\{|\psi_k;S,M\rangle\}$ is computed at the RASSCF level of theory, and the off-diagonal spin-orbit matrix elements are computed within the AMFI approximation. The matrix elements of the model Hamiltonian are shifted to account for state-specific dynamic electron correlation effects (e.g., derived from scalar-relativistic CASPT2 or MRCI calculations), and the eigenvalues and eigenvectors are computed. The implementations of SO-CASCI[505] and SO-MCQDPT[506] limited to CAS reference spaces follow a similar strategy.

Individually selected MRCI codes typically suffer from inherent inefficiencies in evaluating the Hamiltonian matrix elements on-the-fly. They primarily gain from a drastic reduction of the size of the variationally treated configuration space; a posteriori perturbational corrections for the contributions of the omitted configurations are indispensable. In a contracted variant, initially the individual states $\{|\psi_k;S,M\rangle\}$ are computed with a large CSF expansion space, supplying the diagonal elements of the model Hamiltonian. The contracted $\{|\psi_k;S,M\rangle\}$ states are projected onto a reduced basis of determinants $\{|\psi_k;S,M\rangle\}_{red.}$ which still gives a qualitatively correct description of the $\{|\psi_k;S,M\rangle\}$ while offering a more economical basis to evaluate the off-diagonal elements $\langle\psi_j;S,M|H^{SO}|\psi_k;S',M'\rangle_{red.}$ of the model Hamiltonian. The small complex model Hamiltonian is diagonalized.[568] Similar two-step effective Hamiltonian approaches have been proposed by Hess et al.,[560] Rakowitz and Marian,[569] and Buenker et al.[570] As before, orbital relaxation due to spin-orbit interaction (spin-polarization) is difficult to account for in a contracted approach. This problem is addressed in the EPCISO method.[571] Here the model Hamiltonian is computed from the uncontracted reduced determinant basis, enhanced by singly excited determinants to incorporate spin-polarization, and includes corrections for electron correlation effects of the individual $\{|\psi_k;S,M\rangle\}$ states in the large basis. The size of the full complex model Hamiltonian is limited to $\sim 10^5$ with conventional CI methodology. Hence, this approach is fairly closely related to the SPOCK.CI implementation but considerably more restricted in terms of configuration spaces.

Another two-step approach has been proposed by DiLabio.[564] Electron correlation effects for the multiple $\{|\psi_k;S,M\rangle\}$ states are computed in extended scalar-relativistic MR-CISD calculations. The spin-orbit splitting (E^{SO}) is estimated from the energy difference between the scalar-relativistic and two-component MR-CIS calculations; i.e., the configuration space includes static electron correlation plus spin-polarization effects using a version of COLUMBUS CIDBG with configuration selection.[563]

2.7.5. Two- and Four-Component MCSCF and MRCI. The relativistic MOLFDIR/DIRRCI program package[572] supports open-shell DHF calculations for the optimization and selection of the molecular spinor basis for subsequent CI calculations. Double-group and time-reversal symmetry is supported. The direct-CI code operates in a determinant basis following a generalization of the RASSCF approach[370] and using a (multiheaded)[29] graphical representation of the CI space. An improved version of the DIRRCI code is part of the DIRAC10 program package.[573]

The two- and four-component relativistic direct-CI code LUCIAREL[574−576,635,] employs the DKH transformation for the one-electron integrals in combination with the AMFI approximation using a Kramers-pair spinor basis. It was pointed out that the DKH-transformed two-electron integrals lead only to minor corrections of the total energies of the Ag and Au atoms.[494] Both collinear and noncollinear spin functions are supported, though the latter requires on optimized two- or four-component spinor basis optimized at the SO-MCSCF level of theory with, in general, complex orbitals. Since the AMFI approximation includes an approximate treatment of the Breit term, this kind of relativistic approach might be considered superior to four-component Dirac−Coulomb CI calculations. Abelian and quaternionic double groups (C_1^*, C_i^*, C_2^*, C_s^*, C_{2h}^*) are supported with work toward the subgroups of D_{2h}^* in progress. Many implementations of two- and four-component MRCI codes are limited to real double groups or cannot necessarily exploit double group and time-reversal symmetry. GAS-type configuration spaces (multiple orbital subspaces with individual minimum and maximum occupation number constraints) are supported. Because this code is initially based on a nonrelativistic full-CI implementation using a string-based determinant formalism, there are no hard excitation-level limits. Since the structure of the Hamiltonian does not change, the code supports both two- and four-component calculations provided the appropriate integrals are available. A parallel version has been implemented, and calculations on the BiH ground state with up to 428 million determinants showed modest scaling for 32 Linux quad-core nodes with a 1 Gb/s ethernet network.

For heavy elements, in particular actinides and lanthanides, the near degeneracy of s, p, d, and f shells, already apparent with nonrelativistic and scalar-relativistic treatments, suggests the need for a multiconfigurational orbital optimization incorporating static electron correlation effects. Including spin-orbit coupling, the splitting of otherwise degenerate states increases the density of low-lying electronic states so that a relativistic MCSCF treatment appears even more important. Atomic MCSCF calculations show that orbital optimizations with static electron correlation only must include spin-orbit coupling in order to arrive at reasonable spin-orbit splitting. From two-component MRCI calculations, it is well-known that spin-polarization must be accounted for with a spin-averaged orbital basis; in fact with individually selected CI the necessary single excitations may constitute the major part of the CSF space. Hence, more compact relativistic CI expansions may be expected from relativistic MCSCF orbital optimization. A rather general formalism for two- and four-component molecular MCSCF, based on one-electron spinors in the Kramers-pair basis, has been proposed by Jensen et al.[489] and Fleig et al.[577] and has been recently implemented.[578] Another somewhat less general two-component MCSCF implementation, based on spin-orbit pseudopotentials, has been reported by Kim and Lee[579]

2.8. Parallel Computing

Quantum chemistry has historically been one of the leading fields in the use of parallel computing.[580−582] In view of the resource demands of many MCSCF and MRCI implementations, both in terms of CPU time and memory consumption, it is reasonable to efficiently exploit the inherent potential of today's computational resources. However, the performance available from state-of-the-art computer systems is almost exclusively due to massively parallel execution. Current supercomputer hardware offers PFLOP/s (10^{15} floating point operations per second) peak performance to those who manage to exploit it. Until recently, PFLOP/s computer systems relied on an increasingly larger number of compute nodes ($\geq 10^4$). Graphical processor units (GPUs) and accelerator cards are emerging as tools for computational science. These use hardware-accelerated densely integrated parallelization and vectorization to achieve peak performance in the TFLOP/s (10^{12} floating point operations per second) range at significantly lower cost and energy consumption. High-end multicore CPUs, GPUs, and accelerator devices are all currently available on the general market. Thus, to take advantage of the tremendous computational power available with current and upcoming hardware developments, parallelization is essential. However, quantum chemists must modify and adjust their algorithms to utilize this inherent computational power.

The first part of this subsection gives a brief overview of current parallel computer architectures by presenting some details about the computational units, memory, and network architecture. The next subsection sketches some general parallel programming paradigms. The last subsection includes details of parallel multireference methods (MCSCF and MRCI) and the corresponding analytic energy derivatives.

2.8.1. Parallel Computer Architectures. Parallel computer systems can be classified as distributed or shared memory computers. While all CPUs in shared memory systems have direct access to the same address space, distributed memory systems need to exchange data between the nodes by explicit message-passing through a communication network. Most modern parallel computers are a combination of shared and distributed memory architectures since each individual node hosts multiple (multicore) CPUs sharing the same address space. Currently GPUs and accelerator devices operate on dedicated high-speed memory, physically separated from the main memory of the compute node, and therefore they resemble the characteristics of distributed memory systems with a high-speed interconnect.

2.8.1.1. Computational Units. Current parallel computers tend to be built from widely available standard components of high-end personal computers (PCs), and, albeit equipped with special hardware to connect to the communication network, the technical development of the individual compute nodes parallels those available with PCs to the general public. Supercomputer systems combine a large number of processors with a high-speed network and aim at massively parallel codes. Considerable effort, both in terms of hardware and software maintenance, must be invested to keep them fully operational. Standard parallel computer clusters combine a smaller number of nodes, but for this same reason are easier to operate, and they provide a powerful general-purpose resource for many applications.

Initially vectorization (or pipelining) took advantage of performing multiple identical instructions in several overlapping stages, followed by CPUs supporting multiple instructions to

begin executing each clock cycle (multiple issue), combined with increasing clock speeds allowed for an increase in the serial performance by about 4 orders of magnitude since 1980. Since about 2005 the performance gain is almost exclusively due to multicore CPUs; i.e., multiple general-purpose processors are combined on a single chip. However this development does not reduce the execution time of a serial code; it is only through parallelization that performance gains are possible from the hardware.

On GPUs, or accelerator cards, many more, albeit much simpler, cores are combined on a single chip to provide a tremendous performance (~100−1000 GFLOP/s peak) for the single-instruction multiple data (SIMD) mode. Recently GPUs, initially developed to accelerate graphics operations, have reached advanced levels of sophistication. Most importantly, more user-friendly programming models and tools (e.g., CUDA,[583] OpenCL[584]) for these devices have led to growing popularity of GPU programming in a variety of scientific disciplines. Accelerator devices (e.g., ClearSpeed e710[585]) along with CSXL programming tools are fairly similar from the application point of view. They are characterized by multiple SIMD array processors that support tightly coupled vectorization and parallelization combined with a very high internal memory bandwidth (>150 GB/s) and are well-suited for data-parallel applications.

2.8.1.2. Memory Architecture. Current computers have a hierarchical memory structure with multiple data storage devices, each with widely different characteristics.[586] Registers, level 1 (L1), and level 2 (L2) cache operate at very high speed and are private to each core. The L3 cache is substantially larger but shared among all cores of a CPU. The next level in the hierarchy is the computer's main memory, which is usually shared. On a node with multiple multicore CPUs, the access to the node's shared memory is not uniform, and the available memory bandwidth may be optimized by preferentially accessing memory banks physically close to the core. Access to remote memory on a distributed-memory machine is slower by about 1−2 orders of magnitude due to the use of the interconnect.

As a consequence of the ratio of FLOP rate to memory bandwidth, the increasing number of cores residing on a single node causes the observable performance to appear increasingly memory bandwidth limited. The efficiency of the hardware-controlled cache management is very sensitive to the memory access pattern of the application.[586] The compiler primarily manages efficient register allocation, optimum order of instructions, and use of available features of the CPU. Hence, it is important for the application code to avoid random access to vast portions of the memory. Since paging may drastically deteriorate performance, some supercomputer architectures, such as the IBM BlueGene/P, do not permit the virtual memory space to exceed the available real (physical) memory.

2.8.1.3. Network Architecture. The communication network connects the nodes and consists of communication links and switching elements connecting these links. It is characterized by the network topology, bandwidth, and latency. The network topology describes how the nodes are physically connected (e.g., ring, 2D mesh, 2D torus, 3D torus, or crossbar) and each topology may display very different performance characteristics. The performance of a network correlates with the maximum number of hops required to pass a message between two arbitrary nodes in the network. The economical cost of a network rather correlates with the number of switches and the number of links per switch. Hence, a crossbar connecting all computing elements with a single switch and a single hop would be perfect for performance but economically and technically unfeasible for a large parallel computer system. Thus, mesh and torus network topologies are commonly encountered as compromises between costs and performance. In addition, the time needed to pass a message between two nodes is given by $t = t_{lat} + V/bw$, that is, the latency of the initial startup time (t_{lat}) and the data volume (V) divided by the bandwidth bw. For (nonfarming type) applications, remote data access is a major bottleneck, and efforts devoted to optimize the algorithm should address the following: (i) sending a few large messages instead of many small ones, (ii) passing messages asynchronously to hide latency, (iii) passing messages between nodes such as to fit the network topology and to minimize the number of hops, and (iv) avoiding unnecessary collective all-to-all data transfers.

2.8.1.4. File Systems and Disk I/O. The slowest, but also the largest, medium on which to store data in routine parallel computations is the hard disk. Even on the consumer market, multi-TB hard disks are available at moderate prices. For compute servers, it is not uncommon to have 10 TB file systems, and supercomputer systems provide petabyte (PB) storage systems. To avoid I/O contention, the accumulated I/O bandwidth should ideally scale linearly with the number of compute nodes. The simplest way to achieve this is by adding local RAID-based file systems. The disadvantage of this is that the files are only locally visible and thus primarily used for local scratch files. In addition, with $\gtrsim 10^4$ hard disks distributed over the entire machine, the overall mean time between failure (MTBF) may cause machine instability and frequent data loss, which requires additional considerations of fault tolerance and redundancy. The alternative is a common parallel file system (e.g., PVFS,[587] GPFS,[588] and LUSTER[589]) which resides separately on a cluster of dedicated file servers, with files visible to all compute nodes, and hiding all the storage details. The maximum aggregate I/O bandwidth is available only when the I/O characteristics of the application match the performance characteristics of the parallel file system. Special parallel I/O libraries (e.g., SIONlib[590] and MPI-I/O[595]) may be invoked by the application code to transparently ensure optimum usage of the file system. Despite all these advances in I/O technology, it should be kept in mind that today's compute nodes easily saturate the bandwidth of any file system. To avoid dramatic deterioration of the I/O bandwidth, the data access pattern is significant: while sequential (continuous) access to a data file allows the file system to efficiently perform read/write operations (with either fixed or variable record lengths), this may be impossible for random data access patterns. For small files fitting completely into the available file cache, usually no performance degradation occurs.

2.8.2. Parallel Programming Techniques. Parallelism offers great opportunities to extend the range of feasible applications both in terms of the molecular system size and turnaround time. Additionally, as already pointed out, parallelism is essential in order to exploit modern computer architecture. The nature of algorithms in quantum chemistry tends to be rather complex, compute intensive, and frequently very data intensive. Parallel algorithms add an additional layer of complexity and must be sufficiently flexible to account for current and future hardware developments. Using high-level tools to isolate the user code from the details of the bookkeeping and hardware specifics is recommended. The easiest way to approach parallelization is by using community-supported open-source or vendor-supplied

parallel libraries for major parts of the work (e.g., linear algebra, sorting, and other common tasks). Libraries alone are frequently not a sufficient solution because the code rarely spends all of the time in these library calls.

The next programming technique is automatic parallel code generation by the compiler (e.g., OpenMP[591] and HPF[592]); with the possible exception of single-node parallelization, this often results in disappointing performance because, without further guidance by the programmer, the compiler can recognize data parallelism only at very low levels.

Hence, in many cases it is necessary to manually parallelize a program. The major programming paradigms are thread- and process-based parallelization. Using processes amounts to running multiple copies of the application in disjoint address spaces; data exchange between them must be programmed explicitly—all data are private by default. The popular single-program multiple-data (SPMD) programming model is an example of process-based parallelization. In the thread-based model, code sections are executed simultaneously by multiple cores sharing a common address space; all data are shared unless specified explicitly otherwise. Consequently, it is limited primarily to shared-memory address machines. The hybrid model aims at combining the strength of the process- and thread-based programming paradigms. Specifically, this may mean spawning one process per node and additionally creating up to n_c threads per process on each node where n_c should not exceed the number of cores per node. Parallel programming paradigms are constantly adapting to new challenges posed by hardware development and method development in the respective fields of research.

2.8.2.1. Internode Communication. The fundamental problem for parallel threads of execution running on different nodes is how to exchange data between nodes since the address space is physically disjoint. The first approach was the message passing model with both partners actively participating in the data transmission. This was implemented in terms of libraries such as TCGMSG,[593] PVM[594] and MPI,[595] where MPI is today's de facto standard. The disadvantage of this paradigm is a vulnerability to load-balancing problems if tasks are not uniform; this can be addressed by dynamic task assignment at the cost of additional complexity. The active pairwise participation in the data transfer sometimes leads to awkward code and complicated bookkeeping of the distributed data. One-sided memory access, accessing remote data without explicit coordination of the remote processes (e.g., ARMCI,[596] DDI,[597] PPIDD,[598] and MPI-2[595]), can simplify the application code programming considerably. By extending this paradigm to globally distributed data structures (global arrays[596]), much of the bookkeeping in the application code can be avoided. Yet, due to efficiency or compatibility with the underlying algorithms, it is not always sensible to use such global data structures. While decomposing the dominant task in a serial code into independent tasks, with either static or dynamic load balancing, the structure of the code remains basically intact. Carrying out data-parallel operations on distributed data or using parallel linear algebra (via libraries, e.g., ScaLAPACK[599]) is straightforward. To take full advantage of distributed data (e.g., because the available memory per node is insufficient), the basic (serial) kernels must be modified in order to match operations and data access patterns under constraints arising from communication overhead and bottlenecks, load balance, and local memory requirements.

2.8.2.2. Intranode Communication. Processes run on the same node and work within the same physical memory but with distinct virtual address spaces. Thus, the message-passing paradigm can also be used for process communication within a node. This can simplify the programming demands because it allows the same programming model to apply to both the distributed- and shared-memory levels of the hardware. Alternatively, the processes may set up a common memory segment to be shared by all, which can be much faster than transferring data via a message-passing interface. However, the application must ensure that no two processes write simultaneously to the same memory address since the result would be unpredictable for any nonatomic operation (race condition). This can be achieved either by appropriately modifying the application algorithm to allow only disjoint access or by using semaphores to define critical code regions that can be entered by only a single process at a time. The latter approach involves some overhead and implies partial serialization of the code, which can result in performance bottlenecks. Usually, processes are created for the lifetime of the code, though this is not mandatory.

The paradigm of multithreading parallelism advocates the use of threads, lightweight processes that require low overhead for creation and destruction, making feasible threads with short lifetimes dedicated to specific tasks. In addition, threads inherit and share all data of the parent processes plus additionally having private local data (possibly copies of certain data of the parent process). Thus, a thread-safe code must avoid modifying shared data in an unpredictable manner (e.g., simultaneous access by multiple threads). Threads can be manipulated either explicitly through calls to the pthread library[600] or implicitly by use of compiler directives (e.g., OpenMP[591]) that signal the compiler to automatically add the supporting code. Although multithreading is used for fine-grain parallelization, it should not be taken to extremes. With all data residing in cache, a thread lifetime of the order of at least about a millisecond (or $\sim 10^7$ operations) greatly exceeds the thread administration overhead ($1-10\,\mu$s on current architectures) and should yield good parallel performance.

Process-based parallelism (either message-passing or intranode shared-memory) can be combined with multithreading in a hybrid programming model.

2.8.2.3. GPU and Accelerator Usage: Recent Developments. GPUs, as well as accelerators, take advantage of tightly coupled vectorization and parallelization by distributing data-parallel SIMD-like tasks across multiple stream processors. Technical specifications quickly become outdated due to the rapid development in this field, so the following performance values serve only as a guide. The internal memory bandwidth is very high ($\sim$150 GB/s), while the interface to the computer's main memory is relatively slow ($\sim$4 GB/s) given the peak performance between 100 and 1000 GFLOP/s. Multiple GPUs or accelerator cards, equipped with several GB memory each, can be attached to a single host. However, in order to extract the computational power, the GPU kernel must contain little or no logic and have a large ratio of floating point operations to data items; the matrix−matrix multiply operation is the typical benchmark. Due to the slow connection to the computers main memory, especially with multiple GPUs, invoking GPU kernels is often bandwidth-limited. Ideally, the compute-intensive, data-parallel kernels should remain as long as possible on the GPU while generating only a modest amount of output data. The actual hardware-dependent code is usually generated by compilers, guided by compiler directives that define the parallel sections and the partitioning of data and that initiate data transfer between host and GPU. Optimization of algorithms is

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

challenging because the actual run-time behavior is difficult to analyze and to connect to the structure of the algorithm. Simple data-parallel algorithms are advantageous in this task. GPU or accelerator card usage has been reported in context with integral evaluation within the framework of Hartree−Fock[601−603] and density functional theory.[604,605]

2.8.3. Parallel Multireference Methods. *2.8.3.1. Four-Index Transformations.* Because the electron correlation energy converges very slowly with basis set size, highly accurate calculations even on rather small molecules can require very large basis sets. On the other hand, for more qualitative investigations of excited states of larger molecules (20−30 atoms), the AO basis set requirements are somewhat relaxed, yet the total basis set size easily reaches 300−500 basis functions and more. While the one- and two-electron integrals are best evaluated in the atom-centered basis (possibly symmetry-adapted), the wave function optimization and manipulation is best carried out in an orthogonal basis of molecular orbitals. Hence, a four-index transformation is necessary to switch between AO- and MO-based representations. Since four-index transformations play a central role in the wave function optimization and gradient evaluation steps, some variants of parallel implementations are discussed in detail.

The full four-index transformation

$$(ij|kl) = \sum_{\mu\nu\kappa\lambda}^{n} C_{\mu i} C_{\nu j} C_{\kappa k} C_{\lambda l} (\mu\nu|\kappa\lambda) \tag{205}$$

evaluated straightforwardly with eight nested loops requires $O(n^8)$ multiplications. Decomposing the transformation into four quarter-transformations and storing the intermediates reduces the effort to $O(4n^5)$

$$(ij|kl) = \sum_{\mu}^{n} C_{\mu i} [\sum_{\nu}^{n} C_{\nu j} [\sum_{\kappa}^{n} C_{\kappa k} [\sum_{\lambda}^{n} C_{\lambda l} (\mu\nu|\kappa\lambda)]]] \tag{206}$$

More efficiently, the AO integrals are sorted into distributions such that all integrals with the common index pair $\mu\nu$ are collected into a matrix $I^{\mu\nu}_{\kappa\lambda}$ with indices $\kappa\lambda$. The transformation proceeds by the first half-transformation of each distribution $\mathbf{I}^{\mu\nu}$

$$I^{\mu\nu}_{lk} = \sum_{\kappa\lambda} C_{\kappa k} I^{\mu\nu}_{\kappa\lambda} C_{\lambda l} \tag{207}$$

In matrix notation, $\mathbf{I}^{\mu\nu} = \mathbf{C}^{\mathrm{T}} \mathbf{I}^{\mu\nu} \mathbf{C}$. This localizes the memory access for each distribution within the procedure. This is followed by a disk-based (or, in a parallel environment, distributed-memory-based) transposition[606] of the half-transformed integrals $I^{\mu\nu}_{kl} \rightarrow I^{kl}_{\mu\nu}$ to prepare for the second half-transformation

$$I^{kl}_{ij} = \sum_{\mu\nu} C_{\mu i} I^{kl}_{\mu\nu} C_{\nu j} = (ij|kl) \tag{208}$$

Substantial savings can be achieved by exploiting the 8-fold permutational index symmetry of the two-electron integrals.[362,364] This conventional procedure requires the intermediate storage of $n^4/4$ half-transformed integrals, and along with the storage of the $n^4/8$ AO integrals and $n^4/8$ MO integrals (ignoring any numerical sparsity in these arrays) amounts to 250 GB and 1.27 TB for 500 and 750 basis functions, respectively. Abelian point group symmetry adds a further reduction factor of approximately the order of the point group. Even with TB hard disks, the initial sorting and the transposition step requires at least one, relatively slow, random access I/O step.

In many cases, it is necessary to perform the four-index transformation for only a subset of all MO integrals with up to n_x external MO indices required. If o denotes the number of internal orbitals occupied in the MCSCF or reference wave function, the number of required MO integrals is $O(o^{4-n_x} n^{n_x})$ and the number of the half-transformed integrals is $O(o^2 n^2)$ for $n_x = \{0,1\}$, $O(on^3)$ for $n_x = \{2,3\}$, and $O(n^4)$ for $n_x = 4$. The number of integrals determines whether a particular subset must be stored in external storage (requiring I/O), distributed in memory across the entire machine (requiring internode communications), or replicated on each of the nodes (requiring local memory).

Integral-direct four-index transformations skip the separate AO integral evaluation step, and evaluate and process the AO integrals directly as needed. Thus the first half-transformation in the AO-driven procedure is replaced by four nested loops over shell-blocks, thereby computing and processing batches of shell quadruples. The memory requirement of $O(s^2 n^2)$, where s denotes the number of basis functions per shell block, can be further reduced to $O(s^3 n)$ by partially discarding the permutational index symmetry[607,608] and computing certain AO integrals more than once. This allows the first half-transformation to be split into two quarter-transformations, so that the latter step can operate on subsets of the first quarter-transformed integrals, at the expense of a 4-fold redundant integral evaluation. The number of half-transformed integrals remains unaffected.

A variant of the four-index transformation, which is better suited to prescreening techniques, has been proposed by Taylor.[609] Coulomb and exchange operator matrices are directly assembled from the AO integrals and subsequently transformed to the MO basis to yield the final MO integrals.

$$J^{kl}_{\mu\nu} = \sum_{\kappa\lambda} C_{\kappa k} C_{\lambda l} (\mu\nu|\kappa\lambda) = \sum_{\kappa\lambda} D^{kl}_{\kappa\lambda} (\mu\nu|\kappa\lambda) \tag{209}$$

$$K^{jk}_{\mu\lambda} = \sum_{\nu\kappa} C_{\nu j} C_{\kappa k} (\mu\nu|\kappa\lambda) = \sum_{\mu\nu} D^{jk}_{\nu\kappa} (\mu\nu|\kappa\lambda) \tag{210}$$

The formal scaling of $O(n^6)$ appears very unfavorable compared to the standard algorithm. However, density matrices are more localized than MO coefficients. By using efficient prescreening techniques based upon the maximum density matrix element D', along with precomputed estimates of the AO integrals per shell block using the Cauchy−Schwartz inequality

$$|(\mu\nu|\kappa\lambda)| \leq |(\mu\nu|\mu\nu)|^{1/2} |(\kappa\lambda|\kappa\lambda)|^{1/2} \tag{211}$$

the number of integrals may be reduced asymptotically to $O(n^2)$ for extended molecules, and hence the overall scaling reduces to $O(n^4)$ for the **J** and **K** matrix construction. A similar screening, based on MO coefficients, is also applicable to AO-direct implementations of the standard scheme.[608]

Another approach is based on use of the resolution of the identity (RI) or density fitting methods,[610−612] i.e., approximating one- and two-center orbital products by a one-center auxiliary basis χ_P

$$(\mu\nu|\kappa\lambda) \approx \sum_{PQ} (\mu\nu|P)(P|Q)^{-1}(Q|\kappa\lambda) \tag{212}$$

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

For partially transformed two-external integrals

$$(ai|bj) \approx \sum_P B_{P,ai} B_{P,bj} \quad (213)$$

$$(ab|ij) \approx \sum_P B_{P,ab} B_{P,ij} \quad (214)$$

$$B_{P,pq} = \sum_Q [(P|Q)^{-1/2}]_{PQ} \sum_\mu C_{\mu p} \sum_\nu C_{\nu,q} (Q|\mu\nu) \quad (215)$$

the construction of the intermediate quantities $B_{P,ab}$ requires $O(n^3 N_{aux})$ operations and $O(n^2 N_{aux})$ storage. The final step to transform the intermediates to the final MO integrals requires $O(o^2 n^2 N_{aux})$ operations. Auxiliary (RI) basis sets for different purposes have been optimized which scale linearly with the basis set size; typically $N_{aux} \approx 3n$.

A closely related approach initially proposed by Beebe and Linderberg[53] is the Cholesky decomposition (CD). The positive semidefinite two-electron integral array $\mathbf{V}$ is written as a matrix product of the lower triangular supermatrix $\mathbf{L}$

$$(\mu\nu|\lambda\sigma) = \mathbf{V} = \mathbf{L}\mathbf{L}^T = \sum_K L_{\mu\nu,K} L_{\lambda\sigma,K} \quad (216)$$

The expression is exact if the $\mathbf{L}$ matrix (or Cholesky factor) contains the full $n(n+1)/2$ columns. This would result in an effort of $O(n^6)$ for the construction of $\mathbf{L}$ followed by another $O(n^6)$ step to obtain the final integrals. The error of an expansion up to M columns is given in terms of the residual matrix elements $D^M{}_{\mu\nu,\lambda\sigma}$

$$\begin{aligned} D^M_{\mu\nu,\lambda\sigma} &\equiv |(\mu\nu|\lambda\sigma) - \sum_{K=1}^{M} L_{\mu\nu,K} L_{\lambda\sigma,K}| \\ &\leq \sqrt{D^M_{\mu\nu,\mu\nu} D^M_{\lambda\sigma,\lambda\sigma}} \leq \max_{\mu'\nu'} (D^M_{\mu'\nu',\mu'\nu'}) \leq \delta \end{aligned} \quad (217)$$

Since CD is a recursive procedure, it offers a way to truncate the expansion when the acceptable error is below a given limit; $\delta = 10^{-4}$ approximately corresponds to the accuracy with optimized basis sets. In practice, the number of columns M that need to be included has been found to grow linearly with the basis set size ($M \approx 3n$ to $10n$). Under this assumption, the computational effort is of the same order as that for the RI methods. The CD method can be viewed as a scheme to eliminate the redundancy (or linear dependency) of the one-electron basis function product space ($\mu\nu$). Therefore, further approximations such as retaining only one-center product terms (1C-CD)[613] or generating atom-specific (acCD)[614] Cholesky decompositions for the product densities have been proposed. These, however, are no longer exact even in the limit of a complete Cholesky decomposition, but they do offer computational advantages. See ref 55 for a discussion on the relationship between the RI and CD methods. A critical comparison in terms of accuracy and performance can be found in ref 612. Analytical gradients using this approach are available.[615]

The significantly reduced storage requirements for the three-index intermediates results in this approach being particularly beneficial when the resulting MO integrals are immediately processed. The RI approximation has been successfully applied to MP2[610] and CC2[616] methods, including parallel implementations,[617] where the auxiliary basis was tuned to reproduce occupied-virtual orbital products. CD-based variants of CASSCF,[618] CASPT2,[619] and local MR-CISD[299] as well as parallel, integral-direct, CD schemes[620−622] have been reported recently.

2.8.3.2. Integral-Direct MCSCF and MRCI. As discussed in section 2.2, first-order convergent MCSCF methods require only the 0- and 1-virtual MO subset integrals, of which there are only $O(o^3 n)$ in number. The small number of integrals, the small memory requirements for intermediate quantities, as well as additional simplifications for the inactive orbitals makes these first-order convergent approaches attractive for large basis sets despite the slow convergence of the orbital optimization and the associated difficulties related to excited-state optimizations. Second-order methods require, in addition, the 2-virtual MO subset integrals as well as one- and two-particle transition density matrices. The orbital optimization equations may be solved using a variety of approaches. One approach is to explicitly construct the gradient and Hessian elements in the MO basis using eqs 80−86. This requires the appropriate 0-, 1-, and 2-virtual MO subset of integrals which, in turn, can be computed either from the stored integrals, from the AO-direct methods discussed above, or from the Cholesky or RI approximations. Alternatively, the optimization equations may be solved using iterative subspace methods, in which matrix−vector products of the form $\mathbf{G}^{mc}_{orb,orb}\mathbf{k}$, $\mathbf{G}^{mc}_{csf,orb}\mathbf{k}$, $\mathbf{G}^{mc}_{orb,csf}\mathbf{p}$, and $\mathbf{G}^{mc}_{csf,csf}\mathbf{p}$ are constructed for arbitrary expansion vectors $\mathbf{k}$ and $\mathbf{p}$. These matrix−vector products may be computed in operator form using symmetrized one-index transformed integrals and transition density matrices,[19,358] which in turn may be computed either from explicitly stored integrals or with AO-direct methods.

In MR-CISD, the contributions of all three- and four-external MO subset integrals to the $\mathbf{w} = \mathbf{H}\mathbf{v}$ vector can be formulated in terms of exchange operator matrices, i.e., the contraction of densities with $(ab|cd)$ and $(ab|ci)$. These integrals typically constitute the vast majority of all two-electron integrals.[39−42,207,209]

$$\mathbf{K}^{ab,p} = \sum_{cd} D^p_{cd} (ac|db) = \mathbf{C}^T [\sum_{\nu\kappa} [\mathbf{C}\mathbf{D}^p\mathbf{C}^T]_{\nu\kappa} (\mu\nu|\kappa\lambda)] \mathbf{C} \quad (218)$$

The index p represents a particular occupation and spin coupling pattern of the internal orbitals (i.e., an *internal walk* in a GUGA implementation), and D^p_{cd}, which is composed of the elements of a subblock of the current trial vector and the coupling coefficients, assumes the role of a density matrix element. All density matrix elements can be grouped into as many matrices as there are internal walks. Rather than transforming the AO integrals once to the MO basis and storing and accessing them repeatedly within the matrix−vector product step (the Hamiltonian blocks are never explicitly constructed), the density matrices are transformed to the AO basis, contracted with the on-the-fly computed AO integrals, and finally the resulting exchange operator matrices are transformed from the AO to the MO basis as indicated in eq 218. The number of $\mathbf{D}^p$ matrices for ic-MRCI is much smaller than for uncontracted MRCI, a consequence of the different number of variationally optimized parameters. This AO-based approach eliminates the need to store the three- and four-external subset MO integrals, which then simplifies the transformation and storage steps for the remaining two-electron integrals.

2.8.3.3. Parallel MCSCF. Although there are several parallel MCSCF codes available,[219,123,120−122,623] they do not necessarily perform well in terms of scalability or resource requirements. The MCSCF procedure is composed of three major steps per macroiteration: (i) partial AO−MO transformation, (ii) CI

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

eigenvalue problem, and (iii) solving the wave function correction equations. Of these, the AO−MO transformation (possibly integral-direct) is most frequently implemented in parallel. This is particularly beneficial when the intermediates are completely kept in (distributed) memory. For large wave function expansions, the CI wave function optimization step may consume a substantial fraction of the total CPU time and can benefit from parallelization. For large active spaces, it is imperative to either recompute the coupling coefficients within each Davidson iteration cycle (e.g., GUGA[28] and SGUGA[624]) or to make use of the string-based algorithms which operate in a basis of determinants. By use of the resolution of the identity,

$$H_{jk} = \sum_{pq} h_{pq}\langle j|E_{pq}|k\rangle + \frac{1}{2}\sum_{pq}(pq|rs)\langle j|E_{pq}E_{rs} - \delta_{qr}E_{ps}|k\rangle \quad (219)$$

$$= \sum_{pq} h_{pq}\langle j|E_{pq}|k\rangle + \frac{1}{2}\sum_{pqrs}(pq|rs)[\sum_{k'}\langle j|E_{pq}|k'\rangle\langle k'|E_{rs}|k\rangle - \delta_{qr}\langle j|E_{ps}|k\rangle] \quad (220)$$

the two-electron coupling coefficients may be expanded as products of one-electron coupling coefficients. Several variants of the original scheme[46] eq 220, which wrote the determinants in terms of α and β strings ($|K\rangle = |K_\alpha K_\beta\rangle$), have been proposed and implemented.[370,625] An important feature of this approach is that the contraction of the integral arrays with the one-particle transition density matrix elements can be cast as a dense matrix−matrix product, yielding a very efficient computational kernel that parallelizes readily. A related concept for determinant-based full-CI uses α, β strings and replaces the resolution of the identity in favor of a more efficient identification of determinant pairs, yielding nonzero coupling coefficients in terms of reduced lists.[626] Further development of string-based CI has been directed toward the efficient handling of multiple active spaces with occupation restrictions,[4,359,367,574] retaining flexible wave function expansions while reducing the expansion space dimensions compared to traditional complete active spaces.

While the AO−MO transformations dominates for large basis sets and moderate CI expansions, the iterative optimization steps become more costly and more resource-consuming with larger CI expansions. For a recent parallel determinant-based MCSCF implementation[627] performance data were reported for an $\mathbf{11}^{12}$ expansion with 451 basis functions: steps i to iii consume wall clock time with the ratio 90:1:10. Only the orbital Hessian block $\mathbf{G}^{mc}_{\mathrm{orb,orb}}$ was included in step iii. In case of a dominant AO−MO transformation, a fast and reliable second-order wave function optimization with a minimum number of expensive four-index transformations such as the model Hamiltonian approach proposed by Werner[214] is advantageous, although no parallel version exists to date. The main focus in parallel MCSCF codes has been on the first two major steps. The combination of large basis set sizes and large CI expansions gives rise to huge dimensions of the orbital optimization problem, within which each iteration requires a pass through integrals, densities, and transition densities. With increasing size of the equations, the iterative solution tends to take not only more time per iteration but it also often converges more slowly. With state-averaged MCSCF, this technical problem can be even more pronounced.[628]

2.8.3.4. Parallel MRCI. The matrix−vector product formation $\mathbf{w} = \mathbf{H}\mathbf{v}$ in the direct-CI approach accounts for almost the entire computer time of a CI calculation, and the parallelization effort is usually directed to this step. Some other steps of the procedure, such as the operations involving the subspace $\mathbf{W}$ and $\mathbf{X}$ arrays, computation of the subspace elements $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{S}}$, and computation of the residual vectors are trivial to parallelize on the basis of the distributed storage of the subspace arrays. With the appropriate partitioning of the two-electron integrals and of the vectors $\mathbf{w}$ and $\mathbf{v}$, the entire workload can be split into an almost arbitrary number of tasks.[638,213,631,636,629] The computational efficiency of MRCI implementations derives from the ability (i) to efficiently evaluate the coupling coefficients and (ii) to contract them with the corresponding integrals and trial vector coefficients in terms of efficient vectorizable matrix−vector and matrix−matrix operations. There are three major variants of such schemes: the graphical unitary group approach (GUGA),[6,28] the symmetric group graphical approach (SGGA),[6,630] and string-based CI.[46,625,370,436] Both GUGA and SGGA operate in a basis of spin-adapted configurations while string-based CI employs a basis of Slater determinants. GUGA and SGGA are typically limited to MR-CISD expansions since they exploit the simplicity of the coupling coefficients in the two-electron external orbital space while providing efficient means to evaluate the internal coupling coefficients.[6] Unlike GUGA, the closely related SGGA method does not explicitly incorporate spin-coupling information into the graphical representation of the configuration space. String-based CI is particularly well suited for full-CI because the insertion of the resolution of the identity does not introduce additional intermediate determinants (i.e., $k' \in \{j,k\}$ in eq 220). Also a priori configuration selection schemes such as ORMAS and GASCI, which closely resemble direct products of full-CI subspaces, are implemented in terms of string-based CI, although in this case the intermediate states k' extend the underlying determinant space. As discussed in more detail below, relativistic two- and four-component MCSCF and MRCI methods are more frequently represented in a basis of determinants than in terms of CSFs, although the choice is of a technical nature. Implementations of parallel MRCI code 10 years and older are currently only of historical or conceptual interest because the computational hardware architecture has changed dramatically, particularly at the higher end supercomputer level. Nevertheless it is safe to state that there are several very promising determinant-based implementations of full-CI[56,631−634,636] and GASCI[578,635] with up to 60 billion determinants for calculations on diatomics.[636,637] For unitary and symmetric group based implementations COLUMBUS[118] and MOLPRO[120] are among the most popular general parallel uncontracted and internally contracted MR-CISD codes, respectively. Both codes have been applied to the notoriously difficult chromium dimer potential curve. The largest uncontracted expansion reported is about 2.8 billion CSFs,[638] and the largest contracted expansion reported is about 147 million variational parameters and corresponds to an underlying uncontracted expansion space of about 10.2 billion[639] CSFs. Another GUGA-CI code with less favorable parallel scaling properties has been recently described.[640] As was demonstrated in the application of DFT/MRCI on β-carotenes,[641] individually selected MRCI implementations tend to scale considerably worse than the MR-CISD expansions because of the unfavorable memory access patterns, cache misses, and poor load balancing. This method[642] involves the approximate treatment of dynamical electron correlation with DFT by empirically modifying selected CI matrix elements, while near-degeneracy effects were treated by configuration interaction. Since individually selected MRCI

implementations compute the CI matrix elements independently, this merges well with the DFT/MRCI approach, while the computational effort is significantly reduced due to the drastically decreased size of the variational CSF space. Parallel scaling of some full-CI[631] and GASCI[635] implementations was reported to be sensitive to I/O for large expansions. This is because the subspace expansion vectors were stored on disk, and, although computationally inexpensive, the construction of the subspace representations, along with the subspace contractions and other subspace operations, are I/O bound.

2.8.3.5. Parallel MCSCF Derivatives, MR-CISD Gradients, and Nonadiabatic Coupling. For the single-state MCSCF gradient, the Fock matrix, 1-RDM, and 2-RDM in the MO basis are computed as part of the normal iterative procedure. These arrays from the final MCSCF iteration are then back-transformed to the AO basis according to eq 137. This may be done in a parallel implementation, in analogy to the two-electron repulsion integral transformation, with two main differences. One is the trivial observation that the roles of the AO and MO indices are interchanged, which is easily accounted for by transposing the orbital coefficient arrays within the procedure. The other is that the MO density matrix involves only occupied orbital indices, so the total array storage is usually small enough to be replicated as necessary on multiple nodes without bandwidth or total storage concerns. This allows the larger AO arrays to be computed in independent blocks on various nodes without a prior sorting step. These AO arrays are then used as input to the final gradient computation step that implements the derivative contractions of eq 136. In a parallel implementation, this involves sorting the AO arrays into shell blocks, distribution of those shell blocks to the compute nodes, computation of the set of derivative integrals for those shell blocks, and finally the accumulation of the gradient contributions. The partial gradient contributions from the various compute nodes are then globally summed after all shell blocks have been computed.

The state-averaged MCSCF gradients additionally involve the solution of the linear equations of eq 148 and the effective two-particle density construction in eq 151. These steps are required for each state for which the gradient is being computed. After the effective density matrices are constructed, the remaining steps are identical to the single-state MCSCF case. In the back-transformation step, if the density matrix distributions for all of the states are considered together, then there are either a larger number of tasks than in the single-state case, or if the distributions are grouped together, the tasks are larger than in the single-state case; both situations are beneficial to parallel efficiency. In the AO contraction step, there is the further choice of computing the derivative integrals redundantly for each state (resulting in a larger number of tasks with less local memory requirements), or of grouping together the shell blocks of density matrices for all the states of interest and performing the contractions of eq 152 with the unique set of derivative integrals.

The evaluation of the analytic gradients for MRCI wave functions consists of the following major steps. First, after the CI energy and wave function are computed, the CI 1-RDM and 2-RDM are constructed. In a parallel implementation, this step is very similar to a single iteration of the wave function optimization procedure and requires a comparable amount of effort. The CSF expansion coefficients are combined with the coupling coefficients to produce the density matrix elements. The density elements are generally stored either on external disk or in distributed memory. Next, the CI 1-RDM and 2-RDM are combined with the integrals to compute the MO Fock matrix according to eq 82. Those elements are then used to compute the CI orbital rotation gradient. This may be accomplished by sorting simultaneously the integrals and density matrices by distributions, and the Fock matrix elements are computed as the summation of the matrix−matrix products across all of the nodes

$$F_{pq}^{ci,j} = \sum_{\mathrm{st}} \sum_{r} g_{pr}^{\mathrm{st}} d_{qr}^{\mathrm{st}(ci,j)} = \sum_{\mathrm{st}} \mathbf{g}^{\mathrm{st}} \mathbf{d}^{\mathrm{st}(ci,j)} \quad (221)$$

This results in good memory locality while retaining an efficient computational kernel. In multiple state calculations, these computations can be distributed independently of the state index j, resulting in a larger overall number of tasks, or all of the (st) distributions with all states of interest can be treated together with the same number of tasks but with more computation within each task. Once the CI Fock matrices are available, the remaining steps are analogous to those of the state-averaged gradient procedure discussed above.

The parallel nonadiabatic coupling computation is almost entirely analogous to the corresponding analytic gradient computation steps. After the energy denominators are factored into the effective density matrices, the computational equations are almost identical to those of the analytic gradient procedure (i.e., either the state-averaged MCSCF or multiple-state MRCI). The only new quantity required in eq 179 is the skew-symmetric one-particle transition density matrix, the construction and transformation of which is a trivial operation with no significant parallel consequences.

In contrast to the evaluation of analytical gradients, the evaluation of analytical geometric second derivatives even in the favorable case of a variational CASSCF wave function requires the solution of $\sim 3N_{\mathrm{atom}}$ coupled-perturbed MCSCF equations. Integral-direct approaches are preferred to prevent a memory storage or disk I/O bottleneck.[643] Recently, Dudley et al. reported an integral-direct parallel implementation of analytical geometric second derivatives of CASSCF wave functions with good parallel scaling.[644]

3. APPLICATIONS

In this section, some applications of the methods in this review are discussed. A complete overview of the several hundred applications is certainly out of the scope of the present review; our aim is rather to show the potential and applicability of the different methods and techniques and, by discussing some representative calculations, to help the reader to select the appropriate approach for particular molecular systems.

Multireference methods are generally necessary for the description of potential energy surfaces, in particular when a wide range of the coordinate space is considered. Typical examples are the calculation of vibrational spectra including higher vibrational states, chemical reactions, and photochemistry. Additionally, excited states very often require the flexibility of a multireference description, even near equilibrium geometries. Furthermore, multireference methods are certainly needed to describe weak bonds and transition metal complexes characterized by nearby electronic states.

The applications in this section will be discussed from two perspectives. First in section 3.1, a survey is presented of typical applications for the methods described in section 2. Applications using the size-consistency corrected methods, approximate CI methods, and relativistic methods are discussed. Then in sections

3.2 and 3.3 a molecule-oriented overview is given. Specific applications are described, and the strengths and weaknesses of the multireference methods are discussed. Finally, in section 3.4 some general aspects of the choice of molecular orbitals are discussed.

3.1. Overview of the Application Fields of Various MR Methods

3.1.1. Size-Consistency Corrected Methods. Without size-consistency corrections, MRCI results can be biased due to the change of the correlation energy across a PES. Therefore, since the beginning of MRCI calculations, some kind of Davidson-type correction has been applied. Examples include the works by the groups of Buenker and Peyerimhoff (see, e.g., refs 82, 89−91, and 96), by Bauschlicher and co-workers (see, e.g., refs 97 and 645−647), by Schwenke and Truhlar,[98] and by Hogreve[99,690−693] or the newer results by Werner et al. (see, e.g., refs 105, 648, and 649). A very recent, high-accuracy calculation on the vibrational levels of LiH by Holka et al.[650] also shows the importance of size-consistency corrections.

In a certain sense, CEPA type methods are preferable over the Davidson-type corrections. Indeed, the size-consistency corrected methods are widely used in several fields of chemistry. These include calculation of potential energy surfaces, associated vibrational and rotational spectra, excited states and associated spectroscopy, reaction mechanism, properties of transition metal compounds, molecules with heavy elements, noble gas complexes, and reaction control. These methods often serve as a source for benchmarks in studies where lower level methods, such as DFT or MP2, are used as the main method. They are also used to benchmark high-level single-reference calculations, such as CCSD(T), in cases where eventual multireference effects are significant. The most popular methods are MR-ACPF and MR-AQCC, followed by MCCEPA. There is no clear preference in the literature of applications between MR-ACPF and MR-AQCC, and often they are used together. QDVPT was used for several studies shortly after it appeared (e.g., for conjugated polyenes[651,652]), but much less recently. Different versions of MC-CPA have been used in calculations mostly on diatomic molecules of d- and p-shell filled metals, such as GaH or TiCl, concentrating on both ground and low-lying excited states.[653,654] Interestingly, two advanced versions, MR-CEPA and MR-$(SC)^2$-CI, have received little attention in applications (one example is discussed as follows).

The popularity of the MR-ACPF and MR-AQCC methods is clearly due to their simplicity, the close relation to MRCI, the availability of analytic gradients, and, perhaps most importantly, their availability in popular program systems (COLUMBUS,[118,119] MOLPRO,[120] and MOLCAS[123]). MC-CEPA has been implemented into the Bochum suite of codes[153,185−187] with the PNO approximation,[15,16] and this enables its application to larger molecules.

To obtain accurate potential energy surfaces, it is apparent that the MRCI method without the correction for size-consistency error will not provide the necessary accuracy. Therefore, the CEPA-type methods are used in most cases. A large number of applications on diatomic molecules, including the study of their excited states, can be found in the literature. This includes a systematic study of the homonuclear diatomic molecules by Müller et al.,[655] the N_2 molecule by Gdanitz,[656] and the Be_2 dimer by Füsti-Molnár and Szalay,[161] Gdanitz,[657] and Martin.[658] All of these studies used MR-ACPF and/or MR-AQCC methods. Van de Bovenkamp and van Duijneveldt[659] used the MC-CEPA method in an often-cited, systematic study of He_2. Müller[638] recently used the MR-AQCC method to study the long-standing Cr_2 problem.

Ozone has been investigated extensively by multireference methods. Müller et al.[660] calculated the energy difference between the open and the ring isomers of ozone, as well as the dissociation energy, with MR-AQCC using up to quintuple-zeta basis sets and basis extrapolation. MR-AQCC has been also used to obtain accurate three-dimensional potential energy surface by the group of Schinke.[661,662] The complete three-dimensional global surface has been constructed using basis sets up to quadruple-zeta quality and basis set extrapolation. Vibrational levels, and, in particular, the dissociation channel were studied. Recently, Holka et al.[663] extended these calculations, and some parts of the ozone surface were calculated with even larger basis sets.

A very accurate PES has been constructed by Barletta et al.[664] for the water molecule along the bending coordinate up to dissociation. The basic surface is obtained with the ic-MR-AQCC method, and corrections for core-correlation, relativistic effects, infinite basis set have been applied.

He_2 has already been discussed in connection with diatomic molecules. Other rare gas molecules were also treated by MR-ACPF (Ar−CO and Ar_2 by Jansen,[665] neutral XeF by Schröder et al.[666]) and by MR-AQCC (HArF and HKrF by Chaban et al.[667]). Both MR-ACPF and MR-AQCC (in connection with the ic implementation in MOLPRO) was suggested for "the theoretical study of neutral rare-gas compounds for organic chemists".[668]

Electron collision of Cl_2,[669] HBr, and DBr[670] were studied using the MR-AQCC method. These surfaces were found to describe accurately the low-energy electron−HBr collision dynamics.[670]

MR-AQCC has been used for method calibration in regard of the PuO_2^{2+} by Ismail et al.[671] This work extends the application range of the CEPA-type methods even to the f-shell metals (actinides).

To simulate the HCN-HNC isomerization in a laser-controlled pump−dump scheme, Jakubetz and Lan[672] calculated the three-dimensional ground-state potential energy and dipole surfaces with the MR-AQCC method. They showed that it is possible to prepare a pulse that brings the molecule from the HCN ground vibrational state to HNC excited bending states.

In cases where DFT or low-level ab initio methods are often sufficient to describe preparative chemistry, MR-AQCC has been used to describe intermediates (see, e.g., the work of Creamer et al.[673] on dioxirane and its substituted counterparts) or to characterize unstable species (see, e.g., Pasinszki et al.[674]).

Other applications include PESs for excited-state dynamics, but these will be discussed separately in section 3.2.

Finally, some applications of MC-CEPA are given since these are somewhat differerent from those of MR-ACPF and MR-AQCC, due, in particular, to the use of PNO formalism which allows application to larger systems. It was mostly used to describe oxides (e.g., ZnO) and their surfaces and processes thereon (see, e.g., ref 675). There are spectroscopy applications in the literature,[676] as well.

Applications with explicit higher excitation corrections to MRCI are rare because of higher costs and also because of the lack of availability within popular program systems. An exception is the accurate diatomic potential energy curves obtained with the CEEIS method.[199−201,203]

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

3.1.2. Approximate CI Methods. Due to their cost-effective nature, approximate CI methods are indeed very popular in applications. Most common are the ic-MRCI method of Werner and Knowles[209,210] and the MRD-CI method (and program) by the group of Buenker and Peyerimhoff.[274−276]

Over 1500 application papers cite the papers of Werner and Knowles[209,210] describing the ic-MRCI method. Some of these use size-consistency-corrected MR-AQCC or MR-ACPF and have previously been discussed. Typical applications mostly include determination of potential energy surfaces for spectroscopy or the study of reaction mechanism. The most cited applications are the water surface by Partridge and Schwenke,[677] the transition state for the H_2 + F and the Cl + HD reactions by Manolopoulos et al.,[678,679] properties of lanthanide compounds by Dolg et al.,[532] OH + Ar as well as CN + He potential energy surface and quantum scattering thereon,[680,681] the structure of alkaline-earth dihalides,[682] the correlation energy of gold,[683] the accurate enthalpy of formation of OH by Ruscic et al.,[684] HSO by Xantheas and Dunning,[685] and the low-lying states of ClO_2.[686] ic-MRCI was also used in photodynamics (see, e.g., ref 687), the singlet−triplet splitting in benzyne,[688] and many more. Concerning the "very accurate" regime, the magnitude of the error arising from the internal contraction remains unclear.

Individually selected CI methods, in particular the MRD-CI method of Buenker and Peyerimhoff,[274−276] have also been used mainly in the field of excited states (spectroscopy,[733,785] including nonadiabatic couplings,[689] photodissociation,[778] etc.), but they have also been used for the determination of molecular structure, and reaction mechanisms. Hogreve successfully used this procedure to calculate the structure and properties of molecular ions including the ion of helium dimer[99] and trimer,[690] the dication of carbon trimer[691] and carbon dioxide,[692] as well as highly charged atoms.[693] The new version in DIESEL-CI opened the way for organic chemistry applications.[694−696]

3.1.3. Relativistic Calculations. The most popular approach to deal with the inclusion of scalar-relativistic effects has been the use of relativistic pseudopotentials in otherwise nonrelativistic calculations, serving both the purpose of reducing the computational effort by freezing the core−electrons and retaining the well-developed nonrelativistic machinery. For light elements, a posteriori perturbational corrections for the mass-velocity and Darwin terms based on the nonrelativistic electron density are frequently applied. These applications are primarily concerned with molecular structure and energetics of compounds containing one or more d-transition metal, lanthanide or actinide elements, and, to lesser extent, the main group post-f elements, where the relativistic contraction has substantial impact on the chemical properties. This work has triggered the development of systematic basis set series for relativistic pseudopotential calculations in the spirit of the correlation-consistent basis sets.[555,697] Many conventional MRCI studies of transition metal compounds followed this approach.[698,699]

For the spectroscopic properties of the heavier elements, the inclusion of spin-orbit coupling is imperative. Until recently, the primary method of choice has been RECPs with spin-orbit potentials in combination with two-component spin-orbit CI based on nonrelativistically optimized (scalar) molecular orbitals. The spectroscopy of actinyl ions (AcO_2^{n+}, Ac = U, Np, Pu, Am) including intensities was of particular interest in recent years,[700−703] while early spin-orbit CI calculations focused on main group dimers and hydrides.[704,705,562,706] The analysis of the X−A electronic spectrum of Ag_3 and the presence of Jahn−Teller effects has also been addressed recently.[707] On the basis of spin-orbit RECPs or ECPs combined with AMFI spin-orbit integrals, two-step approaches assuming some degree of separability of electron correlation and spin-orbit coupling have been applied to the study of the spectroscopy of iodine,[708] structure optimizations and reaction energies of actinyl ions in aqueous solution,[709] or the spectroscopy of[710] Ag/Ag^+. The AMFI/DKH approach integrates well with two-component spin-orbit CI and offers a route to all-electron spin-orbit coupling treatments to both one- and two-step approaches. Yet, CI-based applications are limited to small molecules and atoms.[711] The two-step variants of the DFT/MRCI method have been found to face difficulties in selecting the appropriate CSF space to give a balanced treatment of spin-polarization. The recently implemented one-step approach appears more reliable, and application to large molecular systems, such as models of β-carotenes[641] or porphyrins,[712] is possible.

Relativistic two-component approaches have been quite successful approximations to the four-component methods, provided the innermost core orbitals of the heavy atoms are kept inactive. In general, excitation energies, bond lengths, and vibrational frequencies are obtained with high accuracy at substantially lower cost.[711]

Due to their simpler structure, fully relativistic four-component MRCI codes have been implemented for some time. Yet, the combination of large primitive basis sets and invariably huge configuration spaces has been a tremendous obstacle, and early applications have been of limited nature (UF_6,[713] PtH,[714] and HX, X = F, Cl, Br, I, At [715]). With the development of efficient, parallel CI codes, new possibilities exist to treat small systems more accurately (RbYb,[716] BiH,[635] and I_3^- [717]).

For the computation of dissociation energies, harmonic vibrational frequencies, equilibrium distances for ground and excited states (independent of the choice of a high-quality SO-RECPS, AMFI/DKH based SO-MRCI, or four-component MRCI treatments), it is important to include sufficient valence electron correlation and, hence, to supply a valence basis with sufficiently high angular momentum functions. Many calculations are far from converged with respect to basis set and correlation treatment and thus rely on error cancellation effects to achieve agreement with experimental data. Only recently have sufficient theoretical data been reported[576] from four-component MCSCF/MRCI calculations on UO_2 that furnish a solid basis for comparison with the more approximate data from one-step SO-MRCI/RECPs[718] (14 valence electrons correlated) and two-step SO-CASPT2[719] (14 valence electrons plus 10 semicore electrons correlated) methods. Initially, the SO-MRCI and the four-component MRCI calculations are in excellent agreement for valence electron correlation, and they find a ground state of ungerade symmetry ($\Omega = 2u$). The SO-CASPT2 calculations on the other hand disagree by about 1500 cm^{-1} for excited states of ungerade symmetry. Upon including an additional 10 semicore electrons in the four-component-MRCI correlation treatment, and improving the basis set with *g* and *h* functions, the difference reduces to about 900 cm^{-1}. Four-component CC calculations are in agreement with these data.[720]

These findings suggest that (i) the additivity assumption of two-step approaches does not generally hold and may introduce error of the order of up to several thousand cm^{-1} and (ii) valence electron correlation alone is insufficient to obtain converged results.

Molecules containing second and third row transition metals are particularly difficult to treat since spin-orbit coupling and

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

electron correlation induce coupling of atomic states with different d-shell orbital occupation. Large differential electron correlation effects are typical, ruling out the practical use of any single-reference electron correlation method. Molecular compounds MeX (Me = transition metal, X = F, Cl, Br, I, O, H) are frequently studied as prototype models for the M−X chemical bonds.[566,721] Large spin-orbit coupling, and mixing of ionic and covalent states upon bond elongation, gives rise to a delicate balance between spin-orbit coupling and electron correlation effects; the computed results are sensitive to both the choice of the molecular orbitals and to the parametrization details of the RECPs.

The f-elements (lanthanides and actinides) may contain a significant number of unpaired f electrons, causing narrow bands of excited states (for a review, see ref 722). For CeO, about 16 excited states within 4500 cm^{-1} are reported. For the lanthanides, the 4f electrons are spatially well separated from the partially filled valence shell so that large-core pseudopotentials (including the 4f electrons) may produce good results.[723,724] In cases where the excited states are not well separated, one-step SO-CI or uncontracted two-step SO-CI methods are adequate. For the early actinides, the 5f orbitals contribute to the chemical bond, and thus they cannot be neglected. On the basis of SO-CASPT2 calculations,[474] the highest bond order in the periodic system, a quintuple designation, has been assigned to U_2. The actinyl compounds (UO_n^{2+}, NpO_n^{2+}, PuO_n^{2+}) containing one to two unpaired f electrons have been studied by several methods. They are characterized by a ground-state multiplet of f states, starting with the neptonyl ion. They display low-lying charge-transfer states, and the density of states is considerably lower than for the corresponding lanthanide compounds. The position of the charge-transfer state does not seem to be related to the number of open f shells. The low-lying excited states corresponding to different f-electron distributions are consistently produced by different methods—differences clearly occur for higher excited and especially CT states.

As has been shown by Vallet et al.,[709] spin-orbit coupling effects can change the thermodynamics of a set of redox reactions for the early actinides. The reduction of actinyl ions from oxidation state VI to state IV in aqueous solution follows a two-step mechanism

$$(MO_2^{2+})(VI) + \frac{1}{2}H_2O \rightarrow HOMO^{2+}(V) + \frac{1}{4}O_2$$

$$HOMO^{2+}(V) + \frac{1}{2}H_2O \rightarrow M(OH_2^{2+})(IV) + \frac{1}{4}O_2$$

The total reaction is endothermic for M = U, while increasingly exothermic for M = Pu, Np, and Am. For UO_2^{2+} the spin-orbit effects reduce the reaction energy by 59 kJ/mol. spin-orbit effects for the self-exchange electron transfer in binuclear Np(VI)/Np(V) complexes in solution has been studied in a modified two-step SO-CI procedure.[725]

Another field of interest is the modeling of the spectroscopic properties of f-element impurities in some crystal environments. One approach to model the environment uses a variant of the model potential technique without periodic boundary conditions, the AIMP embedding potential method developed by Seijo et al.,[726,727] in order to incorporate a polarizable environment. As AIMP, it integrates well with various electron correlation methods and it is usually used in combination with MS-CASPT2 and MRCI. Separate calculations for the excited states of Pa^{4+}, with and without its first coordination shell, allow the different effects of crystal and ligand fields to be discriminated. The crystal field lifts the atomic degeneracies and quenches SO coupling, while the total effect yields a significantly more pronounced SO splitting.

Similar studies on the uranyl and neptunyl ion in the crystalline environment of $Cs_2UO_2Cl_4$ have been carried out by Matsika and Pitzer[728] using the one-step SO-CI method. The UO_2Cl_4 unit was treated ab initio, the first nearest neighbor shell was described by all-electron pseudopotentials, and the remaining shells were represented by point charges. The size of the cluster representing the environment was chosen sufficiently large to converge the Madelung potential for the central $Cs_2UO_2Cl_4$ unit. Only valence electron correlation was computed.

3.2. Applications in Detail: Energy Gradients, Excited States, and Nonadiabatic Coupling

The availability of analytic energy gradients for uncontracted MRCI and MR-AQCC allows the systematic and consistent treatment at the same computational level for geometry optimization and for single-point calculations. The applications described in this section focus on "difficult" situations which cannot be described by standard single-reference methods. These cases contain radical and biradical structures, excited-state minima, and conical intersection between different electronic states. These examples will show not only results but also the choices for how typical reference spaces for the MR-CISD and MR-AQCC calculations can be constructed. The scope of analytic energy gradients does not end with the static description of energy surfaces. Challenging applications can be found in dynamics simulations where, in classical[729] or surface hopping dynamics,[730] analytic energy gradients and nonadiabatic coupling vectors are most useful to describe the dynamics "on-the-fly".[731] The large variety of available applications requires also a significant selection in the presentation. Thus the present discussion will focus on conjugated π systems, starting with a short discussion of vertical electronic excitations, and will continue with a presentation of radical or biradical structures, nonadiabatic effects, and photodynamics.

3.2.1. Vertical Excitations in Ethylene and Butadiene.

The calculation of the lowest singlet valence excitation in ethylene, the V($\pi-\pi^*$) state, has a long history and is full of difficulties. Without going back in detail into the early days of ethylene calculations, the major problem observed in SCF and CISD calculations was the fact that the π^* orbital is too diffuse in such calculations, resembling more a Rydberg orbital.[732] The deeper reason for this problem lies in the ionic character of the V state as described in terms of valence bond theory. Several successful approaches were chosen to cope with this problem. One early approach consisted of extensive selected CI at the MRD-CI level with extrapolation to zero threshold.[733] A second approach followed the general observation that $\sigma-\pi$ correlation was important to account for the different relaxation of the σ core in the fields of ionic and covalent valence bond structures.[734] The "all σ single excitations from all π configurations" (ASSEFAPC)[732] expansion produced remarkably stable and good results that were practically independent of the type of orbitals used in the CI. Subsequently, a two-step procedure was developed[735] in which the differential effects of $\sigma-\pi$ correlation was used to construct improved molecular orbitals in the first step, and in the second step MR-CISD and MR-AQCC calculations were used to

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

compute global electron correlation energies for the ground and V states. In this work, the importance of size-consistency contributions as computed at the MR-AQCC level was also noted. This topic has been taken up recently again in detailed work analyzing the effect of $\sigma-\pi$ correlation at the different levels of theory.[736]

MRCI calculation of the V(1^1B_{1u}) state of butadiene poses problems similar to those of ethylene. CASSCF calculations with active space spanned by the π orbitals result in too diffuse character of the V state (see, e.g., refs 737 and 738). Additionally, a second excited valence state of A_g symmetry is close in energy. It has a pronounced multireference character and is dipole-forbidden. The relative order of the two states was, and to some extent remains, controversial. Theoretical calculations show similar energies for vertical excitation to the two states. Two-step MR-CISD and MR-AQCC calculations with orbitals optimized to include $\sigma-\pi$ electron correlation[738] give $2^1A_g/1^1B_u$ excitation energies of 6.55/6.18 eV. For comparison, CASPT2 calculations[739] give 6.27/6.23 eV, EOM-CCSD(T)[740] gives 6.76/6.13 eV, and recent MRCI calculations[741] give 6.07/6.29 eV. A systematic study of excited states of butadiene and *trans*-2-propeniminium cation has been reported by Lehtonen et al.,[742] showing good results in terms of excitation energies at several coupled cluster levels, except for the two multireference ^1A_g states due to their multireference character as already discussed above. It is certainly necessary to go beyond selected single-point calculations and to continue the effort in simulation of optical spectra similar to the work of Krawczyk et al.[743] in order to obtain a better account of the elusive 2^1A_g state of butadiene.

3.2.2. *p*-Benzyne. 1,4-Didehydrobenzene, or *p*-benzyne, is a biradical intermediate formed by the Bergman cyclization.[744] There have been numerous studies of the *p*-benzyne singlet ground state (1^1A_g);[745,746] however, there have been very few reports of the characterization of the excited states other than the low-lying triplet. Extensive MR-CISD, MR-AQCC, and MR-AQCC-LRT calculations were performed for vertical electronic excitations, including valence and Rydberg states by Wang et al.[747] The minimal space for describing the biradical is the CAS $(\sigma\sigma^*)^2$ expansion consisting of the two in-plane singly occupied orbitals. The active space was extended in the work of Wang et al.[747] to a CAS $\mathbf{8}^8$ by including the six π orbitals and six electrons into the active subspace as well. Rydberg states were considered by adding one auxiliary active orbital to represent the Rydberg 3s orbital. Only single excitations were allowed from the CAS into the auxiliary space. In total, 32 states were included in the state-averaged MCSCF calculations. The same space was used as the reference space for the aforementioned MR-CISD and MR-AQCC computations. A high density of electronic states was observed in this biradical system due to the fact that there are more than 17 states within 7 eV of the ground state, including two 3s Rydberg states. All excitations, except the 2^1A_g state, consist primarily of excitations from the π system into the (σ,σ^*) biradical orbitals. Of the 32 states characterized, 15 were significantly multiconfigurational, including the ground 1^1A_g state, providing further evidence for the necessity of a multireference approach for *p*-benzyne.

3.2.3. Stability of the Allyl Wave Function. Symmetry breaking in radical systems and Hartree−Fock instability is an interesting problem and has been investigated over a long period of time.[748] The allyl radical is an outstanding example in this respect; for a review on ab initio results see ref 749. Stable wave functions have been constructed on the basis of a CASSCF($\mathbf{3}^3$) calculation[750] with the active space consisting of the π orbitals. Extensive CASSCF and MR-ACPF calculations have been performed by Szalay et al.[751] to compute an accurate allyl geometry and vibrational force field. The stability properties of the wave functions employed have been tested by performing geometry optimizations starting from a symmetry-broken structure of 2-propenyl-type. Use of the CAS($\mathbf{3}^3$) space in CASSCF and MR-ACPF resulted in a stable wave function with equivalent CC bonds and C_{2v} symmetry. Increasing the CAS($\mathbf{3}^3$) to CAS($\mathbf{4}^3$) and CAS($\mathbf{5}^3$) in the π space of the allyl radical led to interesting results displaying instability and symmetry-breaking for some expansions. Optimizing the allyl radical at CASSCF($\mathbf{4}^3$) showed symmetry-breaking, whereas the CASSCF($\mathbf{5}^3$) approach gave again the correct symmetric geometry. This fact was explained in terms of a near degeneracy of the $2a_2$ and $3b_1$ π orbitals used for the extension of the CAS($\mathbf{3}^3$) to CAS($\mathbf{5}^3$). Both orbitals are simultaneously needed, and omission of one of them leads to symmetry-breaking of the CASSCF wave function. The MR-CISD and MR-ACPF wave functions are always stable, even with symmetry-broken SCF orbitals.

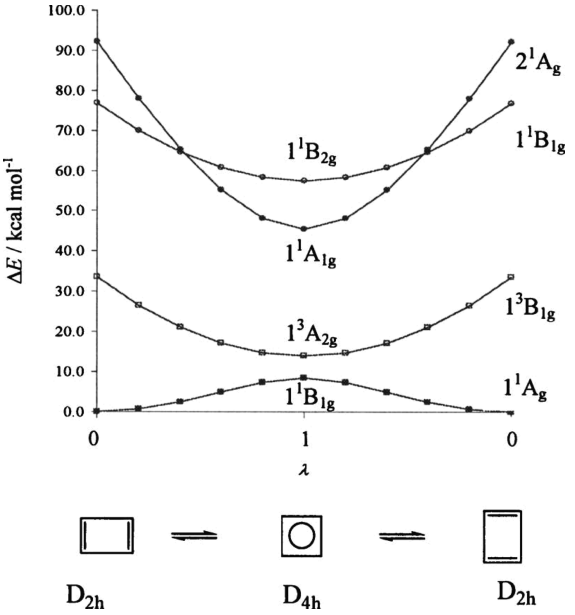


Figure 1. Energy variation along the automerization path of cyclobutadiene for the ground state and the three lowest excited states. Reprinted with permission from ref 754. Copyright 2006, American Institute of Physics.

3.2.4. Automerization in Cyclobutadiene. Cyclobutadiene is an interesting molecular system since it is the smallest neutral organic compound that shows the effect of antiaromaticity. Another characteristic feature is its strong angular strain. The D_{2h} ground-state structure is of closed-shell character, implying that it is well-described by single-reference methods. However, the transition-state structure has a square geometry (D_{4h} symmetry) and is an open-shell system, for which a multireference approach is required. The crucial factor determining the automerization rate is the barrier height for the process leading from one D_{2h} structure to the equivalent one via the D_{4h} saddle point (see Figure 1).

Experimental values for the activation energy vary considerably within the range of 1.6−12 kcal/mol.[752] Multireference coupled cluster calculations including single and double

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

excitations augmented by noniterative triplet excitations (MR-CCSD(T)) based on a two-determinant reference gave a barrier height of 6.6 kcal/mol.[753] Inclusion of zero-point vibrational energy (ZPVE) lowered this value to 4.0 kcal/mol. CASSCF(4^4) calculations result in a barrier of only 2.5 kcal/mol when ZPVE corrections are applied. MR-AQCC calculations were performed[754] in order to determine the automerization barrier height and to investigate the properties of the lowest excited states. Several wave functions were used. At the MCSCF level, a CAS(4^4) within the π space was initially chosen. Additionally, a restricted direct product (RDP) space[19] consisting of the four ($\sigma_{CC}\sigma_{CC}^*$) and four ($\sigma_{CH}\sigma_{CH}^*$) subspaces was used with each subspace restricted to singlet spin-coupling of two electrons. Subsequent MR-AQCC calculations used either a CAS(4^4) as reference space or a space complemented by all configurations generated by single excitations from the manifold formed by the four π and π^* orbitals into the eight σ^* orbitals (CC and CH), in addition to all configurations generated by single excitations from the eight σ orbitals into the two π, two π^*, and eight σ^* orbitals. Basis sets up to cc-pVQZ quality[466] were used in single-point calculations, and basis set extrapolations were also performed. A final barrier height of 6.3 kcal/mol including ZPE corrections was obtained. Concerning investigations of excited states, Figure 1 shows that the 1^3A_{2g} (see also ref 755) and the 1^1B_{2g} states have square-planar equilibrium structures. The first triplet state is of particular interest because of its alleged aromatic character. In a recent investigation[756] two conical intersections between S_1/S_0 were located, one with ionic character and the other an open-shell tetra-radical.

3.2.5. Diels−Alder Reaction of Ethylene and 1,3-Butadiene. The Diels−Alder (DA) reaction of ethylene and butadiene is the prototype of a Woodward−Hoffmann-allowed 4s + 2s cycloaddition (for review see, e.g., ref 757). Both concerted and nonconcerted mechanisms have been discussed. The concerted case involves an aromatic boatlike transition structure, whereas the nonconcerted case involves a biradical intermediate. A balanced description of biradical and nonradical structures is a difficult task. Unrestricted DFT calculations suffer from spin contamination, and CASSCF calculations lack dynamical electron correlation.[758] The multireference Møller−Plesset calculations to second order (MRMP2) performed in ref 758 are expected to give better balanced results. MR-AQCC calculations allow for the integration of multireference effects and dynamic electron correlation, and the geometry optimization capabilities have been used to study both the present DA reaction[759] and the Cope rearrangement of 1,5-hexadiene.[760] In the case of the DA reaction a CAS(6^6) was used in CASSCF calculations and as reference space in the MR-AQCC expansion. Basis sets ranged from 6-31G* [761] to 6-311G**.[762] In summary, the best estimate for the concerted barrier of 22.2 kcal/mol for the forward reaction, ethylene + butadiene to cyclohexene, is in good agreement with the 21.85 kcal/mol deduced by Huang et al.[763] from fits to experimental rate constants using variational transition state theory. Various stationary points on the biradical region of the PES were investigated. The energy difference of 6.5 kcal/mol between the concerted transition state and the biradical fragmentation transition state for the anticonformer is in line with experimental estimates.

3.2.6. Excited States: Energy Surfaces and Conical Intersections. Vertical electronic excitations starting in the ground-state minimum and characterized by single excitations are usually well-described by single-reference methods such as EOM-CC techniques.[764] These methods can also be well-suited for the description of the neighboring Franck−Condon region, including excited-state minima. In many cases, however, sufficient energy is available from the vertical electronic excitation to also reach nearby energy barriers; this allows access to regions of the PES for which significant multireference character occurs and where the electronic wave function is significantly more complicated. The chemical intuition guiding many theoretical calculations for the ground state is of little or no help in this situation, and flexible, unbiased, general multireference methods are required. For smaller molecules such as the prototypical ethylene molecule, extended MR-CISD calculations are feasible and provide accurate results, but with increasing molecular size these calculations become increasingly more costly. Because of the pressing need for information on excited-state surfaces for significantly larger molecules, less flexible methods must be used such as MRCI with single excitations (MR-CIS) or state-averaged MCSCF. These methods are significantly faster than MR-CISD, and enable interesting applications in many areas. However, it must be noted that the missing dynamical electron correlation and the ambiguities of selecting the proper active orbital space may lead to nonnegligible errors.

3.2.6.1. Ethylene: Energy Surfaces and Conical Intersections. The ethylene molecule plays a fundamental role in the understanding of photoisomerization processes and, in particular, the ultrafast energy conversion through nonadiabatic transitions. Therefore, quantum chemical calculations on vertical excitations and of important sections of the PESs have a long history.[765−767] The vertical excitation to the V state of ethylene has already been discussed previously. The primary interest in the excited-state energy surface concentrated on the torsion around the CC bond; this torsion leads to a degeneracy between the V and Z valence excited states at 90°[768] but not to an intersection with the electronic ground state. The first global analysis of modes that lead excited-state ethylene to intersections with the ground state was made by Ohmine[769] and Freund and Klessinger.[770] In those investigations, the importance of hydrogen migration and of CH_2 pyramidalization for reaching the S_1/S_0 intersection was demonstrated. Later, Ben-Nun and Martínez[771] performed extensive studies on the structures and stabilities of the important conical intersections: the ethylidene, twisted/pyramidalized, and twisted/H-migration intersections. They used SA-CASSCF(7^4) and single-point ic-MR-CISD calculations using the active space of the CASSCF calculation as reference wave function. State-averaging was performed in the CASSCF calculations over two and three states, respectively. The aug-cc-pVDZ basis set[466,772] was used, allowing also the description of Rydberg states along additional one-dimensional potential energy searches. Optimization of the ethylidene and twisted/pyramidalized intersections led to true minima on the intersection seam (MXSs), whereas for the twisted/H-migration structure only a representative structure on the seam was given. The first two structures were found to be similar in energy, whereas the latter one was significantly higher in energy. Comparison of the energy difference between S_1 and S_0 states computed at the CASSCF level (at which the MXS optimization was performed) and at MRCI levels suggest possible inconsistencies. The degeneracy of the S_1/S_0 states is fulfilled at the CASSCF level by virtue of the MXS optimization performed at this level. The MR-CISD single-point calculations using the CASSCF MXS structures show small splittings for the ethylidene and twisted/pyramidalized MXSs, but larger splittings

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

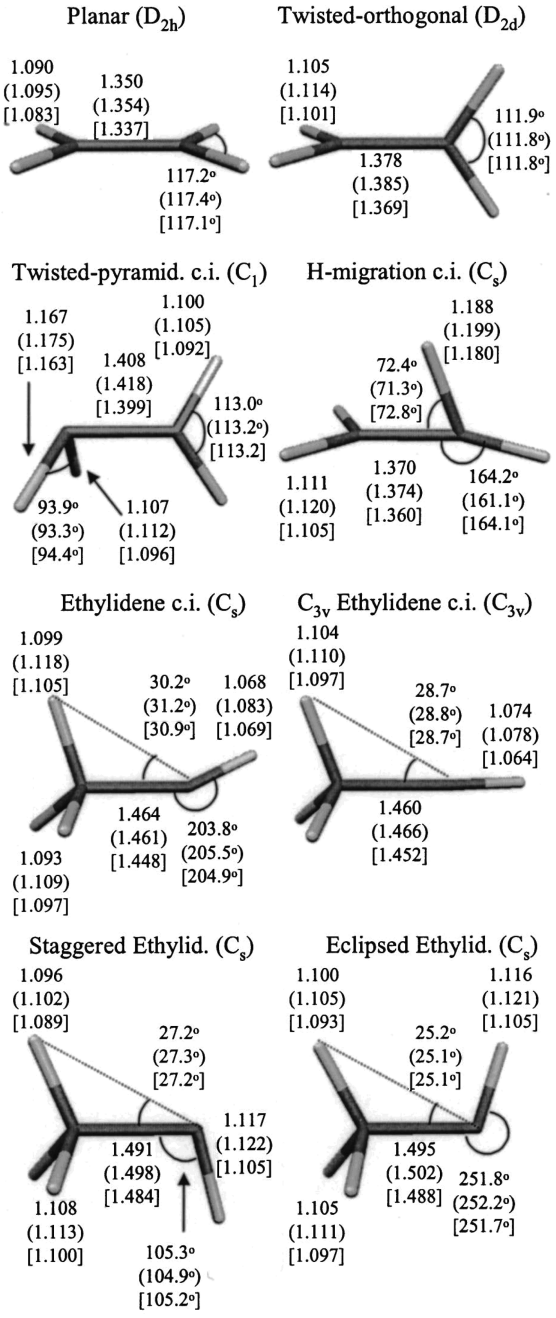


Figure 2. Selected geometrical parameters for the main C_2H_4 structures studied in the present work optimized at the MR-CISD/SA-3-CAS(2^2)/aug-cc-pVDZ level. Values in parentheses and in square brackets were obtained at the MR-CISD/SA-3-RDP/aug-cc-pVDZ and MR-CISD/SA-3-RDP/aug-cc-pVTZ levels, respectively. For the twisted-pyramidalized MXS, the pyramidalization angle b is 104.7° (103.5°) [104.4°]. Distances are given in angstroms and angles in degrees. c.i. stands for conical intersection space. Reprinted with permission from ref 773. Copyright 2004, American Institute of Physics.

in the case of the twisted/H-migration conical intersection structure.

MR-CISD calculations have been used by Barbatti et al.[773] to optimize MXSs for ethylene consistently at the correlated level. These investigations went beyond standard CAS reference spaces which are limited, for practical reasons, to small active orbital spaces. In addition to the $(\pi\pi^*)^2$ CAS space, a RDP space was constructed for the σ orbitals. The RDP space was composed in this case of 10 orbitals grouped in five subspaces, one for each σ bond, i.e., four $[(\sigma\sigma^*)_{CH}]$ pairs and one $[(\sigma\sigma^*)_{CC}]$ pair. Each $(\sigma\sigma^*)$ subspace was restricted to singlet pairing. The MCSCF calculation based on the RDP wave function resulted in localized orbitals very similar to those obtained in generalized valence bond (GVB) calculations.[774] The same RDP expansion space was used as the reference space in the subsequent MR-CISD calculations. Basis sets used range from aug-cc-pVDZ to aug′-cc-pVTZ quality, where the prime indicates that the augmented f functions on the carbon atoms and the augmented d functions on the hydrogen atoms were omitted. Comparison of optimized geometries computed with different methods and basis sets (see Figure 2) showed that the $(\pi\pi^*)^2$ CAS reference space and the aug-cc-pVDZ basis represented a good compromise between accuracy and efficiency for calculations of the valence regions relevant for the photodynamics within the S_0, S_1, and S_2 states. The topology of the seam was investigated as well. It has a complex structure in which all conical intersections presently known for ethylene are connected.

3.2.6.2. Heterosubstitution in Ethylene. As shown by Michl and Bonačić-Koutecký[775] using two-electron two-orbital model calculations, the gap between the S_1 and S_0 states in the orthogonal structure narrows as the electronegativity difference between the two central atoms diminishes. This behavior has important consequences on the photochemical deactivation mechanism to the ground state because the pyramidalization and H-migration mechanisms should lose importance relative to ethylene.

Substitution of one carbon atom in ethylene by silicon leads to silaethylene. SA-CASSCF and MR-CISD calculations were performed[776] to investigate vertical excitations, conical intersections, and potential energy curves for selected coordinates. To take valence and Rydberg states into account, a $(\pi\pi^*)^2$ CAS was augmented with an auxiliary space describing the 3s and 3p Rydberg orbitals. Only single excitations were allowed into the auxiliary space. This expansion space was used in SA-MCSCF calculations and as the reference space for the MR-CISD expansion. At the MCSCF level, the Rydberg states were lower in energy than the $\pi-\pi^*$ state, whereas at the MR-CISD and MR-CISD+Davidson levels the $\pi-\pi^*$ state was the lowest excited singlet state. Optimization of conical intersections showed—as expected—the twisted orthogonal structure to be a MXS. Twisting of silaethylene around the C−Si bond leads directly to the MXS.

Isoelectronic substitution of a carbon atom in ethylene with a nitrogen cation leads to the methaniminium cation $CH_2NH_2^+$. This is an interesting molecule since it is the first one in the series of protonated Schiff bases that are used as models for retinal, the chromophore of the opsin visual protein.[777] Bonačić-Koutecký et al.[778] applied the two-electron two-orbital model to the methaniminium cation and verified the orthogonal MXS structure by means of direct-CI[779] and MRD-CI calculations. The calculations showed that at least two excited singlet states are involved, a $\sigma\pi^*$ and a $\pi\pi^*$. In the vertical excitation, the $\sigma\pi^*$ state is energetically lower than the $\pi\pi^*$ state. Torsion around the CN bond leads directly to a crossing with the ground state. This intersection has also been investigated in a qualitative way[780] using the Longuet−Higgins phase change theorem. The vertical excitations in the methaniminium cation were also investigated

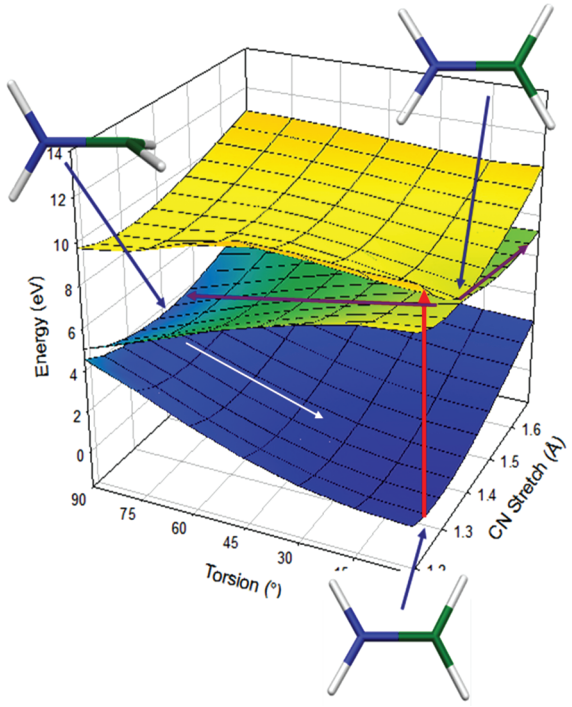


Figure 3. Potential energy surfaces for the methaniminium cation in terms of the torsional and CN stretching coordinates. The red arrow indicates the Franck−Condon excitation.

by Du et al.[781] with a selected MR-CISD approach, confirming the energetic ordering of the above-mentioned $\sigma\pi^*$ and $\pi\pi$ states. CASSCF and MR-CISD calculations performed by Barbatti et al.[782] identified the CN stretching coordinate as another important degree of freedom in addition to the CN torsion. Calculations were performed at the SA-CASSCF(4^6) level that included three states into the state-averaging procedure. The same 4^6 expansion space was used as reference in MR-CISD calculations with cc-pVDZ and cc-pVTZ basis sets. Starting in the Franck−Condon point (see Figure 3), stretching of the CN bond leads to a crossing between the $S_2(\pi-\pi^*)$ and $S_1(\sigma-\pi^*)$ states at planar geometries. Torsion around the CN bond results in the familiar intersection between S_1 and S_0 states. As discussed in what follows, dynamics calculations are necessary to obtain an adequate picture of the true photodynamical processes.

Substitution of one of the hydrogen atoms of ethylene by a heteroatom also creates a polar π bond. For example, MR-CISD calculations show[783] that twisting around the CC bond in fluorethylene leads to a conical intersection at the twisted orthogonal structure, as was the case for silaethylene and the methaniminium cation. Analogous calculations have been performed on the chiral (4-methylcyclohexylidene) fluoromethane.[784]

3.2.6.3. Formaldehyde. In addition to Rydberg transitions, at least three valence excited states ($n-\pi^*$, $\sigma-\pi^*$, and $\pi-\pi^*$) are necessary for the characterization of the vertical electronic excitation spectrum of formaldehyde. The interactions between the $\pi-\pi^*$ and Rydberg states have been investigated in great detail in the benchmark MRD-CI calculations of Hachey et al.[785] Because of these strong interactions, the $\pi-\pi^*$ state appears to be elusive and difficult to characterize experimentally. The work described in ref 786 shows how a larger set of electronic states can be computed simultaneously at the MR-CISD, MR-CISD+Q, and MR-AQCC levels. Five valence states (ground state, $n-\pi^*$, $\sigma-\pi^*$, $\pi-\pi^*$, and $n^2-\pi^*$) and 10 Rydberg states, n−(3s3p3d) and $\pi-3s$, were computed using an MRCI reference space in which the Rydberg states were represented by an auxiliary orbital subspace into which only single excitations were allowed. Standard MR-CISD and MR-AQCC calculations were performed, and the interactions of valence and Rydberg states with the CO bond stretch were analyzed. On the basis of these results, EOM-CCSD investigations were subsequently performed,[787] computing one-dimensional potential curves for all vibrational normal modes, two-dimensional (involving the CO stretch and HCH bend modes), and three-dimensional (including additionally the out-of-plane mode) energy surfaces. Using these surfaces, wave packet dynamics simulations were performed to compute the UV spectrum. The computed spectrum reproduces well the experimental data in the 7−10 eV region, including highly irregular features due to strong interactions between the $\pi-\pi^*$ state and the nearby Rydberg states.

The structure of the $\pi-\pi^*$ state is also of interest for other reasons. The MRD-CI [785] and CASPT2 [788] methods predict a planar structure for this state. However, it was later shown[789] that the $\sigma-\pi^*$ and $\pi-\pi^*$ states cross and, in combination with a CH_2 out-of-plane bend, lead to a conical intersection. As a consequence, the planar $\pi-\pi^*$ structure is a saddle point rather than a true minimum. Full geometry optimization of this state leads to strong deviation from planarity and strong mixing of the $\sigma-\pi^*$ and $\pi-\pi^*$ states.

In contrast to the $\pi-\pi^*$ state, the $S_1(n-\pi^*)$ state is spectroscopically well characterized.[790] The photodissociation dynamics of formaldehyde with respect to the S_1 state has been investigated intensively (see, e.g., refs 731 and 791−798). There are two photodissociation products: H_2 + CO and H + HCO. The molecular photoproducts are exclusively obtained from dissociation on S_0, whereas the radical fragments can be derived from T_1 and S_0. The dynamics simulations of the molecular channel have been performed on the S_0 surface starting at the corresponding transition state. Beyond the standard dissociation process that was observed in earlier investigations, an interesting additional "roaming" hydrogen atom process was found.[793] Along this pathway the molecular transition state is avoided and radical formation is initiated. However, this does not lead to full dissociation because of insufficient energy. Instead, the two fragments orbit around each other, eventually resulting in hydrogen abstraction and leading to H_2 + CO.

Explanation of the dynamics of the radical products is significantly more difficult since it involves the three electronic states S_0, S_1, and T_1. Beyond this fact, long simulation times are required to describe the singlet−triplet transitions. Straightforward on-the-fly dynamics is, therefore, precluded in spite of the small size of the formaldehyde molecule. Instead, global fits to the potential energy surfaces have been developed. Some examples include the investigations on the S_1/S_0 conical intersection[795] using density functional theory and CASSCF calculations, and the calculations on the S_1/T_1 intersection[796,797] based on the EOM-CCSD and MRCI methods. Quasiclassical trajectory calculations have been performed that exclude the S_1/S_0 intersection because of energetic reasons[797] and focus instead on the S_1/T_1 intersection. Wavepacket simulations involving both the S_1/T_1 and S_1/S_0 crossings using the CASSCF method have been reported.[798] Despite the fact that existing investigations have led to important insight into the formaldehyde S_1 photodissociation,

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

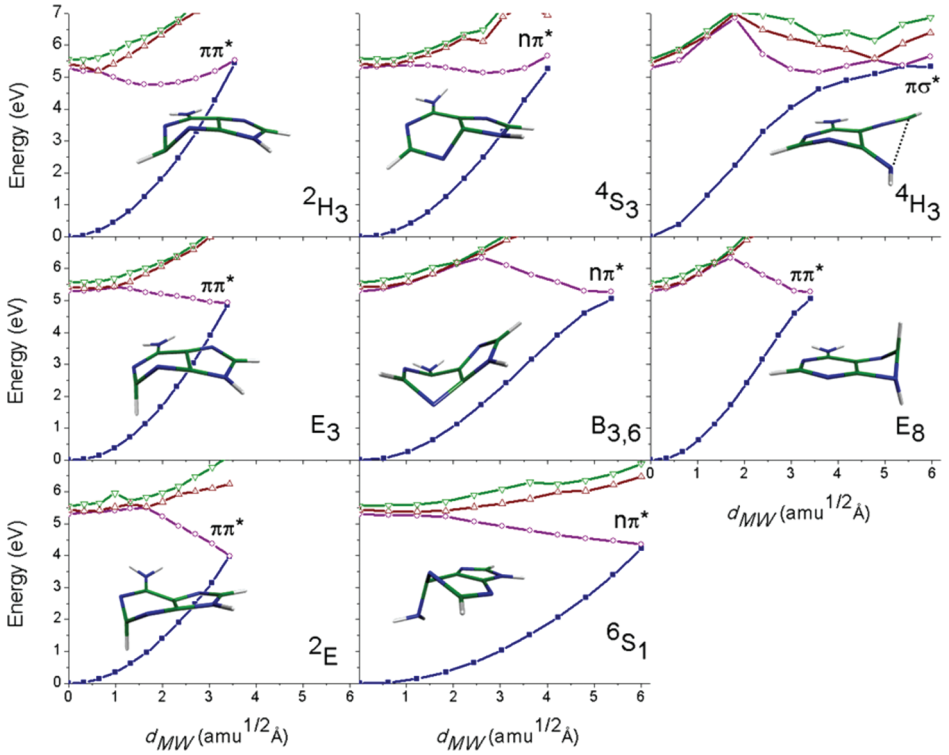


Figure 4. Paths connecting the ground-state minimum geometry to the puckered MXSs of adenine. The MXS geometry is shown in each case. The coordinate *d* describes the mass-weighted Cartesian distance between a given geometry and the ground-state minimum. Reprinted from ref 803. Copyright 2008 American Chemical Society.

a comprehensive understanding based on accurate dynamics simulations has yet to be proposed.

3.2.6.4. Excited States of DNA Bases. The study of the UV spectra, conical intersections, and photophysical deactivation paths for the DNA and RNA bases leads into the fascinating field of the photostability of DNA. The lowest electronic excitations are characterized as valence $n-\pi^*$ and $\pi-\pi^*$ states with varying energetic order depending on the particular nucleobase investigated, but also depending on the computational method selected. Because of the increased molecular size of the nucleobases in comparison to the substituted ethylenes discussed above, MRCI calculations of comparable accuracy are difficult to perform because of the drastically increased computer times. The quantum chemical calculations on excited states presented in the literature are, therefore, dominated by SA-CASSCF and CASPT2 calculations. MR-CISD calculations suffer from overshooting certain $\pi-\pi^*$ states as already observed for the V state of ethylene. These errors are significantly reduced for the strongly distorted MXS structures, and therefore, MR-CISD has been used mostly for verification purposes of such structures. MR-CIS (MR-CI plus single excitations) has been applied with the goal of compensating for inadequacies of the state-averaged orbital-optimization procedure, and in selected cases this resulted in significant improvements of excitation energies.

UV spectra, energy minima in excited states, conical intersections, and photodeactivation paths have been investigated in detail by several groups.[799−804] Energy minima have been located in both the S_1 and S_2 states with usually small energy barriers of a few tenths of an electronvolt. Reaction paths crossing these barriers lead, in most cases, to conical intersections at strongly distorted, ring-puckered structures. As an example, Figure 4 shows several reaction paths and types of MXS structures that have been computed for adenine in ref 803. The calculations were performed at the SA-3-CASSCF($\mathbf{10}^{12}$) and MR-CIS($\mathbf{5}^6$) levels where the CAS($\mathbf{5}^6$) reference space was obtained by moving for all stationary points and MXSs all orbitals with natural occupations larger than 0.9 to the doubly occupied space and smaller than 0.1 to the virtual space. The Cremer−Pople parameters[805] have been used in Figure 4 to describe systematically the ring-puckered conformations. This figure illustrates nicely the multitude of different pathways. Assuming that the photodynamics starts in the lowest $\pi-\pi^*$ state, the most likely reaction paths lead to the ^2E and $^6S_1$ MXSs.

Other examples of successful application of MRCI methods can be found in the calculations performed by Matsika on uracil[801] and Kistler and Matsika on cytosine.[806,807] It was shown that MRCI calculations on molecules of the size of nucleobases are possible. The calculations are based on a SA-CASSCF($\mathbf{9}^{12}$) calculation providing the orbitals for the subsequent MRCI calculation. The nine orbitals are composed of seven π, a lone pair on one N atom (n_N), and a lone pair on the O atom (n_O). Three different MRCI expansions were constructed in ref 806. The first (MRCI1) included only single excitation CSFs generated from the CAS orbitals. This low-level expansion was used for MXS searches. The next two expansion sets incorporated dynamical correlation of the σ electrons with the active π and nonbonded electrons. This type of excitation has been shown to be important for the description of excited states of organic molecules.[734,735] In the MRCI$\sigma\pi$1 expansion, only single excitations were included. The third method (MRCI$\sigma\pi$2) includes single excitations from the σ orbitals and the oxygen lone pair, plus single and double excitations from the CAS into the virtual

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

Table 3. S_1 to S_3 Excitation Energies[a] of Cytosine Computed at Three MRCI Levels

method	$S_1(\pi-\pi^*)$	$S_2(n_N-\pi^*)$	$S_3(n_O-\pi^*)$
MRCI1 [b]	5.101 (0.067)	5.394 (0.002)	5.888 (0.001)
MRCI$\sigma\pi$1	4.941	5.131	5.625
MRCI$\sigma\pi$2	5.136	5.289	5.927
expt [808]	4.66		

[a] Data are from ref 806. Values are in electronvolts. [b] Oscillator strengths are given in parentheses.

orbitals. The cc-pVDZ basis set was used. These calculations demonstrate the flexibility in choosing MRCI expansons, and how this can be largely decoupled from the MCSCF step. Table 3 collects vertical excitation energies computed in ref 806 for the three expansions. For comparison, an experimental value available for aqueous solution[808] is given as well. The results show that the bright $\pi-\pi^*$ state is lowest, followed by two $n-\pi^*$ states. See ref 806 for details of the geometries and of the energetic stability of the S_1 minima and conical intersections. Using these MRCI methods, three-state conical intersections could also be located by a three-state seam search.[807]

Cytosine MXS structures have also been optimized at SA-4-CASSCF($\mathbf{10}^{14}$) and MR-CISD($\mathbf{5}^{6}$) level[809] using the 6-31G* basis. The CASSCF and MR-CISD optimized geometries are found to be rather similar.

3.2.7. Radical−Radical Reactions. Radical−radical reactions present a special challenge to electronic structure theory.[810] These reactions typically have no barriers, and hence the dynamical bottlenecks for these reactions usually occur at large, 2−4 Å, separations between the two radical centers. It is well-known that Hartree−Fock wave functions are poorly suited to this task; a restricted Hartree−Fock wave function introduces spurious ionic character at large radical−radical separations, and an unrestricted Hartree−Fock wave function introduces significant spin contamination. Multireference methods, both MRCI and CASPT2, have been shown to yield accurate results for a wide variety of radical−radical reactions including combination reactions,[811−814] disproportionation reactions,[815,816] and roaming radical reactions.[793,817−820]

3.2.8. Bond Length Comparisons. Experimental R_e values were compared with computed R_e values for 20 molecules using three multireference electronic structure methods, MCSCF, MR-CISD, and MR-AQCC by Shepard at al.[369] Three correlation-consistent orbital basis sets were used, along with CBS extrapolations, for all of the molecules. These data complement those computed previously by Helgaker et al.[821] and Bak et al.[822] with SR methods. The MCSCF wave function expansions were all of the direct-product form, including GVB-RCI and CASSCF expansions. Several trends were observed. The SCF R_e values tend to be shorter than the experimental values, and the MCSCF values tend to be longer than the experimental values. These trends were attributed to the ionic contamination of the SCF wave function and to the corresponding systematic distortion of the potential energy curve. Upon orbital basis improvement, the SCF values tend to shorten even further from the experimental values, while the MCSCF values shorten toward the experimental values. For the individual bonds, the MR-CISD R_e values tend to be shorter than the MR-AQCC values, which in turn tend to be shorter than the MCSCF values. Compared to the previous SR results, the MCSCF values were roughly comparable to the MP4 and CCSD methods, which is more accurate than might be expected due to the fact that these MCSCF wave functions include no extra-valence electron correlation effects. This suggests that static valence correlation effects, such as near degeneracies and the ability to dissociate correctly to neutral fragments, play an important role in determining the shape of the potential energy surface, even near equilibrium structures. The MR-CISD and MR-AQCC methods predict R_e values with an accuracy comparable to, or better than, the best SR methods (MP4, CCSD, and CCSD-(T)), despite the fact that triple and higher excitations into the extra-valence orbital space are included in the SR methods but are absent in the multireference wave functions. The computed R_e values using the multireference methods tend to be smooth and monotonic with basis set improvement.

Full-CI wave function R_e values were compared for some of the molecules in this study. This allows some direct comparisons among the various SR and MR methods without interference from experimental uncertainties. Overall, the MR-CISD and MR-AQCC R_e values agree very well with the corresponding full-CI R_e values, and the mean errors and standard deviations for these methods are better than for any of the SR methods. The MR-AQCC method is seen to have the smallest mean error and smallest standard deviation of all of the methods. The MCSCF statistics are seen to be surprisingly good, supporting the premise that valence correlation effects play a larger role than expected in determining the shape of potential energy surfaces, even near equilibrium structures.

The variational nature of the MR-SDCI and MR-AQCC energies allows the Hellmann−Feynman theorem to be exploited in the analytic energy gradient computations associated with the molecular structure optimizations. This is seen to result in a very efficient analytic energy gradient computation for a wide range of reference wave function expansion dimensions and for a wide range of orbital basis sets. The timings for these calculations show the practical advantage of using variational wave functions for which the Hellmann−Feynman theorem can be exploited.

3.3. Applications in Detail: Nonadiabatic Dynamics

The availability of analytic energy gradients and nonadiabatic coupling vectors opens the way to mixed quantum-classical dynamics simulations (see, e.g., ref 365), for which Tully surface-hopping is probably the most popular form.[730] For the purpose of this review it is of interest to note that these calculations are performed on-the-fly,[731] which means that at each time step of the classical dynamics a full quantum chemical calculation is performed, including energy gradient and coupling elements. The advantage of this approach is that all internal degrees of freedom can be taken into account without any restrictions since the analytic energy methods automatically provide the complete derivative vectors along all Cartesian coordinates. Thus it is not required to preselect any active internal coordinates as is necessary when energy grids have to be calculated. The drawback of the on-the-fly method is the large computational demand, which is even more pronounced by the fact that a batch of trajectories must be computed in order to obtain statistically useful results. Thus, when ab initio methods are used for the quantum chemical part of the calculation, simulation times are usually restricted to ultrafast processes in the order of a few picoseconds. Ab initio multiple spawning (AIMS)[771] is an interesting alternative method that combines the rigor of quantum dynamics with computational efficiency comparable to surface-hopping.

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

3.3.1. Ethylene Photodynamics. The study of the photochemistry of ethylene gives insights into basic chemical processes and at the same time serves as a prototype for studying properties of larger molecules. The structure of the conical S_1/S_0 intersections, the topology of the intersections seam, and the modes leading to the seam have been discussed previously in section 3.2.6.1. Extensive photodynamical investigations have been performed by several groups. AIMS calculations carried out at the CASSCF(2^2) level[771,823] nicely demonstrate the combined torsional and pyramidalization character of the deactivation paths leading from the Franck−Condon region to the conical intersection. A lifetime of 180 fs has been deduced from these simulations. Wavepacket dynamics calculations based on a fitted potential described in six coordinates (CC torsion, CC stretch, CH_2 scissoring, and pyramidalization)[824] show a strong coupling between the torsion and the CC stretching mode, along with a significant bottleneck in the slow activation of the pyramidalization, leading to an overall slow population transfer to the ground state. Although these simulations illustrate the proper deactivation paths qualitatively, they predict lifetimes that are much too long compared to the 20−40 fs values obtained in femtosecond pump−probe experiments.[825,826]

Surface-hopping dynamics using semiempirical CI methods are an interesting alternative to the above simulations based on ab initio methods. Granucci et al.[349] used the MINDO/3 method[827] and Fabiano et al.[828] used the orthogonalization-corrected Hamiltonian (OM2) method.[829] These investigations confirm the torsional and pyramidalization modes as the main deactivation mechanism, similar to the observations made in the ab initio simulations described before. Interestingly, much shorter lifetimes are reported in these investigations: 50 fs in the MINDO/3 study and 70.8 fs in the OM2 calculations. Semiempirical dynamics performed afterwards at the AM1 level[830] with reparameterization of the original AM1 parameters based on MRCI calculations led again to larger lifetimes of 105−139 fs.[831] Closer analysis of these dynamics indicates[832] also the importance of the H-migration and ethylidene regions of the S_1/S_0 intersection seam for the photodeactivation. A possible reason for the discrepancies between theoretical and experimental predictions was proposed:[771,831] the energy of the probe pulse for ionization was not sufficient for the whole course of the dynamics, and the experiments were actually giving only the time to leave the observation window rather than that to return to the ground state. Multiphoton experiments,[826,833] however, have ruled out this explanation and reinforced the previous experimental conclusions.

Recent AIMS calculations performed for ethylene, using analytic multistate perturbation theory to second-order (MSPT2) energy gradients[834] and numerical finite differences for the nonadiabatic coupling vector,[835] result in a lifetime of 89 fs. This value is a significant improvement over the previous ab initio results and shows the sensitivity of the lifetime with respect to the computational method. Even more accurate calculations are certainly necessary to resolve completely the puzzle of the ethylene lifetime.

3.3.2. Ethylene Heterosubstitution: Effect on Photodynamics. The fact that heterosubstitution in ethylene promotes the S_1/S_0 crossing at the twisted orthogonal structure (see discussion above) may lead to the expectation that the corresponding photodynamical processes should be simpler than those observed for ethylene. The central torsional mode should lead the dynamics directly to the conical intersection and thus to the ground state. However, the situation is not really that simple.[836] The first reason follows from the fact that not only the torsional mode but also the central stretching coordinate defines the intersection space, and a specific combination of these two modes is required to reach this region. Normally, this will not be the case at the beginning, and the dynamics will lead initially to a finite gap (avoided crossing region) rather that to a crossing (intersection). Usually, several torsional cycles will be required to reach the intersection seam. E.g., in silaethylene the 90° structure is reached in 10−20 fs. However, the lifetime computed from surface-hopping dynamics using a MR-CISD(2^2 reference) wave function is found to be 124 fs.[837] Since the molecule remains in the excited state longer, sufficient time is available to activate additional modes. Consequently, other regions of the crossing seam, different from the ones belonging to twisted configurations, may actually be accessed during the return to the ground state. This situation is not unique to the combination of torsional and stretching modes, but rather occurs in many other situations such as the ring puckering found important for the deactivation of DNA nucleobases.

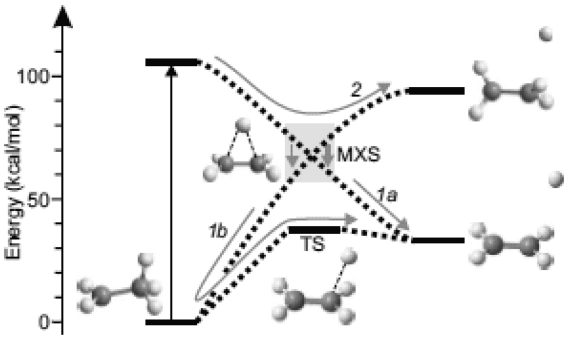


Figure 5. Schematic energy level diagram for the dissociation of ethyl radical $C_2H_5 \rightarrow C_2H_4 + H$ following excitation to the $\tilde{A}$-state. Reprinted with permission from ref 838. Copyright 2009, American Institute of Physics.

The second reason that deviations from the simple rotor model occur is competition with other processes. Such cases were observed in silaethylene[837] and methaniminium cation.[782] While one group of the trajectories follows the torsional paths, another undergoes strong stretches of the central bond connected with a simultaneous pyramidalization of both terminal groups.

3.3.3. Adiabatic and Nonadiabatic Dissociation of the Ethyl Radical. The dynamics of the dissociation of the ethyl radical following excitation to the $\tilde{A}$ state has been studied by Hostettler et al.[838] As Figure 5 shows, adiabatic dissociation (2) produces excited-state ethylene and H in competition with nonadiabatic dissociation (1a). Alternatively, hot ground-state radicals can be generated followed by unimolecular dissociation (1b). For the nonadiabatic dynamics calculations, the SA-2-CASSCF(4^3) and MR-CISD(4^3 reference) methods have been used. To correctly describe the dissociation of the ethyl radical into ethylene and hydrogen atom, the four active orbitals contain three valence orbitals (two 2p orbitals on the carbon atoms and the 1s orbital on the departing hydrogen atom, forming the nonbonding singly occupied molecular orbital and the bonding and antibonding σ orbitals) and the 3s Rydberg orbital. The 6-31++G(d,p) basis set was used. A set of 4956 trajectories was run at the SA-CASSCF level for a maximum of 7 ps, and 245 trajectories were run at the MR-CISD level for a maximum of 1 ps. When comparing the dynamics results obtained with the two

dx.doi.org/10.1021/cr200137a |*Chem. Rev.* 2012, 112, 108–181

methods, the changes in the branching ratio between adiabatic and nonadiabatic dissociation are illustrative in showing the effect of the computational level. At the SA-CASSCF level, only about half of the trajectories dissociate nonadiabatically; the majority of the other trajectories follow path 2. At the MRCI level almost 70% of the trajectories follow the nonadiabatic path, and the remainder is equally split between paths 1b and 2. A possible explanation for the dramatic difference in the branching ratio between the two levels of theory comes from the computed dissociation energy for the $C_2H_4(\tilde{A})$ + H channel that is too low at the SA-CASSCF level but agrees well with experiment at the MRCI level.

3.3.4. Photostability of DNA/RNA Nucleobases. All nucleobases show ultrafast radiationless decay to the ground state within a few picoseconds. These ultrafast deactivation mechanisms and their related photostability are of great interest since they are considered to contribute to a natural chemical defense of the genetic code against damaging photochemically induced processes in reactive excited states. The structure and energetic location of the conical intersections responsible for these ultrafast processes have been discussed previously. Together with ultrafast time-resolved spectroscopy,[839,840] these methods have provided a high degree of detail in the characterization of internal conversion processes, suggesting ring puckering as one major structural theme for the photodeactivation of nucleobases. Photodynamical simulations have been performed at the ab initio level using surface-hopping[802,841−844,809,845] and AIMS[846,847] techniques for investigating the individual bases. A summarizing survey was presented with the goal of extracting the common pattern of the deactivation dynamics.[848] Analysis of the dynamics simulations shows that the purine bases, adenine and guanine, follow a direct decay proceeding diabatically on the $\pi-\pi^*$ energy surface. These two bases possess the shortest gas-phase lifetimes among all nucleobases. In contrast, the pyrimidine bases display much richer deactivation characteristics where also the $n-\pi^*$ state is involved and trapping in S_1 minima may occur.

In addition to the study of the five nucleobases, the dynamics of related compounds which serve as nucleobase models, such as 4-aminopyrimidine,[849] and 2,4-diaminopyrimidine,[850] has been performed.

This collection of examples should demonstrate the range of applications available. Advances in the parallel performance of quantum chemical codes and implementation of quantum mechanical/molecular mechanics (QM/MM) methods, in combination with MCSCF and MRCI methods (see, e.g., refs 851 and 852) will lead to significantly enhanced possibilities allowing more realistic applications in many important fields of research where multireference theory is the method of choice.

3.4. Role of the Molecular Orbital Basis

The use of advanced SR and MR electron correlation methods is invariably tied to the optimization of the one-particle basis in terms of molecular orbitals separate from the optimization of the N-electron basis expansion coefficients (Slater determinants, CSFs). With the exception of full-CI, which by construction treats all MOs on the same footing, this introduces an implicit dependence on the choice of molecular orbitals of the results from the electron correlation methods. In the area effectively covered by SR methods, due to the formalism there is usually no choice of the MO basis, which has unfortunately led to the misconception that a suitable set of MOs can be taken for granted if they satisfy the basic requirement to cope with possible near-degeneracy effects.

For about 40 years, it has been well-known that natural orbitals derived from some electron correlation treatment lead to a more compact expansion of the CI wave function.[1,6,9,10] However, the iterative improvement of the NO basis does not always converge,[853] so that this procedure, although quite helpful, does not lead to a uniquely defined orbital basis. Thus, the effort to compute good approximations to the full-CI limit depends on the choice of the molecular orbitals, and, in the early days of quantum chemistry, natural orbitals (or related quantities such as PNOs[15,16]) have been used to reduce the size of the N-particle space while keeping the incompleteness error small (e.g., the PNO−CI method[205]).

Of conceptual importance is the fact that HF wave functions tend to overestimate the ionic character of a chemical bond. This leads, e.g., to systematically overestimated harmonic vibrational frequencies that are typically corrected by empirical methods and basis set dependent scaling factors.[854] More advanced scaling techniques employ multiple scaling factors depending upon the internal coordinates (e.g., ref 855). The increased ionic character is also reflected in the shape of the occupied canonical HF orbitals, so that subsequent electron correlation treatments must correct for this orbital bias. In a recent study it has been pointed out that even for typical single-reference cases, high-level electron-correlation methods such as CCSD(T) cannot completely compensate for this defect if the results are extrapolated to the CBS limit.[369] For the same set of molecules, MCSCF/MRCI or MCSCF/MR-AQCC did not suffer appreciably from ionic contamination.

The notoriously problematic $\pi-\pi^*$ singlet states in polyenes, carbonyl compounds, and other conjugated or aromatic systems exhibit ionic character, which recently has been shown to be systematically overestimated even at the MCSCF/MRCI level of theory when the active space is not flexible enough to incorporate $\sigma-\pi$ polarization effects.[734]

In a basis of orthonormal MOs, a $(\pi\pi^*)^2$ CASSCF calculation can represent three singlet and one triplet states of A_g and B_{1u} symmetry (using D_{2h} irrep labels). If c_π and c_{π^*} denote arbitrary coefficients subject to the orthonormalization constraint $(c_\pi)^2 + (c_{\pi^*})^2 = 1$, these states are (ignoring the core electrons)

$$1\Psi^{1}A_g(N) = c_\pi|\pi\overline{\pi}| + c_{\pi^*}|\pi^*\overline{\pi}^*| \quad (222)$$

$$2\Psi^{1}A_g = c_{\pi^*}|\pi\overline{\pi}| - c_\pi|\pi^*\overline{\pi}^*| \quad (223)$$

$$1\Psi^{1}B_{1u}(V) = \frac{1}{\sqrt{2}}(|\pi\overline{\pi}^*| + |\pi^*\overline{\pi}|) \quad (224)$$

$$1\Psi^{3}B_{1u}(T) = \frac{1}{\sqrt{2}}(|\pi\overline{\pi}^*| - |\pi^*\overline{\pi}|) \quad (225)$$

In a basis of orthonormalized atomic p orbitals centered on the two carbon atoms (p_A,p_B), the same wave functions can be written as

$$1\Psi^{1}A_g(N) = \frac{1}{\sqrt{2}}(c_\pi + c_{\pi^*})(|p_A\overline{p}_A| + |p_B\overline{p}_B|) + \frac{1}{\sqrt{2}}(c_\pi - c_{\pi^*})(|p_A\overline{p}_B| + |\overline{p}_A p_B|) \quad (226)$$

$$2\Psi^{1}A_g = \frac{1}{\sqrt{2}}(c_{\pi^*} - c_\pi)(|p_A\overline{p}_A| + |p_B\overline{p}_B|) + \frac{1}{\sqrt{2}}(c_\pi + c_{\pi^*})(|p_A\overline{p}_B| + |\overline{p}_A p_B|) \quad (227)$$

$$1\Psi^{1}B_{1u}(V) = \frac{1}{\sqrt{2}}(|p_A\overline{p}_A| - |p_B\overline{p}_B|) \quad (228)$$

$$1\Psi^{3}B_{1u}(T) = \frac{1}{\sqrt{2}}(|p_B\overline{p}_A| - |p_A\overline{p}_B|) \quad (229)$$

dx.doi.org/10.1021/cr200137a |Chem. Rev. 2012, 112, 108–181

The two singlets of A_g symmetry are complementary mixtures of ionic and covalent VB structures, while the $^1B_{1u}$ state and its $^3B_{1u}$ counterpart are represented by purely ionic and covalent VB structures, respectively. Hence, state-specific CASSCF calculations with the minimum active space yield widely different π and π^* MOs for the ionic V state as opposed to the predominantly covalent N and T states. Constraining the occupied σ core orbitals to be inactive in the CASSCF optimization prevents them from being properly polarized for the ionic configurations as both components share the same core. To allow for σ polarization, σ to σ^* single excitations relative to the $|(core)\pi\bar{\pi}|$ and $|(core)\pi^*\bar{\pi}^*|$ determinants are required. The neglect of this polarization results in π and π^* MOs that are too diffuse, and the ionic components are destabilized.[736]

The vertical excitation energy of the V state has been systematically overestimated by ~0.25 eV and a too diffuse character is predicted by a wide variety of methods focusing on extending the *N*-particle expansions (CI,[853,733,651,856] EOM-CCSD,[740] and CASPT2 [739]), which is unfortunately slowly convergent. The observation of better results with more compact basis sets[853] is a direct consequence of preventing too diffuse orbitals by basis set constraints. Focusing instead on the MO basis leads to the aforementioned inclusion of $\sigma\pi$ polarization, either as furnishing the final results[857] or as a MO basis optimization step followed by large-scale MRCI for quantitative predictions.[735] A third possibility involves QDPT or intermediate Hamiltonian theories,[858] constructing a model space from multiple valence states (V state and states representing $\sigma\pi$ polarization effects) such as MS-CASPT2[859] or QD-PC-NEVPT2,[334,736] which also corresponds to an extension of the *N*-particle space. Results that are in good agreement with ref 735 have been presented with eight reference states using the QD-PC-NEVPT2 approach if limited to valence excited states.

For the standard $\mathbf{2}^2$ CASSCF calculations for ethylene, in addition to the neglect of $\sigma-\pi$ polarization effects, the now energetically too high ionic V state mixes with a near-degenerate low-lying Rydberg state of the same symmetry, which further increases the diffuse character of the V state. The degree of mixing with the Rydberg state depends very much on the choice of the active space, since it is primarily an artifactual near-degeneracy effect (experimental data indicate a valencelike V state). It is important to include $\sigma-\pi$ polarization and dynamic $\pi-\pi^*$ correlation during the MO optimization step in order to suppress both sources of error. QD-PC-NEVPT2 fails in the presence of Rydberg valence state mixing because a very large number of reference states would be required to correct the large diffuse character of the V state. Optimizing the orbitals with large RASSCF expansions including sufficient $\sigma-\pi$ and $\pi\pi^*$ correlation leads also to qualitatively correct orbitals for the V state, and subsequent single-state PC-NEVPT2 results are in good agreement with extended MR-AQCC and MR-CISD results. These findings reemphasize the importance to closely inspect and analyze the optimized molecular orbitals. The longer polyenes are less problematic with respect to $\pi-\pi^*$ excitations, and a similar procedure leads to good agreement with the experimental data.[738]

The simultaneous treatment of Rydberg and valence excited states is of considerable importance in dealing with the photophysics of many aromatic heterocycles and carbonyl compounds. In particular, formaldehyde has been thoroughly investigated[785−789] (cf. section 3.2.6.3); upon elongating the C═O double bond, the $\pi-\pi^*$ state crosses four different Rydberg states. This has a profound impact on the observed electronic spectrum and is an almost ideal example for intensity borrowing.[787] In terms of efficiency, it is desirable to construct an MO basis in which each MO describes either a diffuse Rydberg or a compact valence orbital.[786] Since Rydberg states are essentially singly charged molecular cations plus a single electron in a diffuse orbital, it is sufficient to choose reference spaces with at most a single electron in a Rydberg orbital. Although such a CSF space is not much larger than the corresponding CSF space for the valence-state correlation treatment, it is flexible enough for arbitrary valence/Rydberg state mixing within the *N*-particle basis. Even with sophisticated CSF spaces at the MCSCF level, it will not be possible to always disentangle Rydberg/valence orbital mixing at the orbital optimization level; the consequence is the requirement of extremely large CSF spaces in order to suppress artifacts. To this end, response methods display the desirable feature that it is sufficient to optimize the MO basis for the reference state only. Since the ground state is well-separated from any Rydberg states, it is easier to suppress the undesirable valence/Rydberg mixing at the orbital optimization level.

4. SUMMARY AND CONCLUSIONS

We have discussed the relevant aspects of the MCSCF and the MRCI methods which are generally applicable procedures to compute approximations to the electronic Schrödinger equation. These multireference methods share the important asset that they are not inherently tied to some restricted reference state, and thus they are applicable to arbitrary electronic states and molecular geometries. The variational nature of these methods greatly simplifies the formulation and the implementation of analytical gradients and nonadiabatic coupling vectors. While MCSCF is primarily used to optimize the one-electron basis functions and to describe static electron correlation effects due to nearly degenerate electronic states, the subsequent MRCI method aims at quantitative treatment of dynamical electron correlation. The most prominent restrictions of traditional CI procedures are the lack of size-consistency and the exponential growth of the *N*-particle expansion space. The latter initiated the development of various approximate CI variants achieving reductions of the number of variational parameters by CSF selection, contraction, or reparametrization schemes of various forms, and also multireference methods that follow completely different approaches. The size-consistency, or in this context size-extensivity, error is related to the presence of unlinked clusters in the truncated CI equations which can be dealt with by inclusion of (approximate) higher excitation contribution to the energy expression or equations. In the former case the correction takes the form of a posteriori corrections or extrapolations of CI results, while in the latter case a new method is defined. These new CEPA-type methods do not necessarily deal consistently with the EPV terms, causing overestimation of the correction. Also, the applied modifications often ruin the structure of the CI equation, destroying the possibility of straightforward calculation of analytic gradients and properties; the central tasks of computational chemistry, such as the evaluation of critical points on the PES, is thereby restricted to relatively expensive numerical gradient evaluation.

The last two decades have witnessed a remarkable development of relativistic MCSCF and MRCI methods, and calculations on general molecules of chemical interest with high accuracy are becoming feasible. The computational effort is demanding for both nonrelativistic and, in particular, full four-component

relativistic multireference methods, and the use of advanced parallel (super-) computers is inevitable. Hence, the general current status of parallel computer architectures and the implications of implementations of multireference methods on (massively) parallel were discussed. Replacing disk storage by distributed memory, trading computer time for reduced data storage requirements, and data-parallel algorithms are important and common underlying concepts. In particular, schemes that reduce the exponential growth of the number of variational parameters of the FCI expansion to low-order polynomial dependence, while still maintaining small errors, are of great current interest.

A variety of applications from a selection of different fields including approximate CI and size-consistency-corrected methods, applications of multireference methods to excited states, gradients, and nonadiabatic surface-hopping dynamics, and relativistic quantum chemistry have been discussed. Several illustrative applications with references to the original works have been included.

Finally, the dependencies and intricacies that can arise from the separate optimization of the one- and *N*-particle bases were discussed. Most notable is the fact that it is difficult for the electron correlation method to compensate for an inappropriate one-particle MO basis except at the full-CI limit.

Perspectives on the future of quantum chemical multireference methods are not easy to assess. As the results collected in this review show, multireference methods are indispensible in many very important application areas. It is also clear that dynamical electron correlation is necessary to obtain the required accuracy; i.e., lower level valence-only multireference methods such as MCSCF are not always satisfactory. In the single-reference realm, coupled-cluster theory brought about substantial improvement of the quality of the results, and a similar impact is expected when an all-purpose MRCC method is developed. Despite large efforts, no such MRCC method exists at the time of writing this review (see also the review by Lyakh and Bartlett[860] in this issue); thus MRCI, and in particular extensivity corrected versions such as MR-ACPF and MR-AQCC, are the primary choice for applications. Note also that the MR-CISD (including MR-ACPF and MR-AQCC) are expected to give closer results to the hypothetical MRCC level than does SR-CI compared to SR-CC since a much larger amount of correlation is included with the MR wave functions. From the technical point of view, the availability of ever growing computational resources, in terms of computer power, memory, and storage capabilities, will certainly continue, albeit the characteristics of the computational hardware are likely to change dramatically; this will have a profound impact on algorithm choices as the software strives to match the inherent capabilities of the hardware. Relatively simple data-parallel and data-local compute kernels, preferentially combined with on-the-fly computation of large amounts of internal data (such as integrals and coupling coefficients) followed by immediate contraction to compact representations (e.g., Fock matrices), are likely to benefit from the upcoming hardware. More conventional techniques relying on the processing of large amounts of precomputed data, stored either in distributed memory (or much worse on external disks), will likely encounter difficulties in this new environment. Generally, methods that allow for the reduction of the volume of data to be stored and manipulated have the most promising potential. The current revival of interest in multireference methods is due to their many inherent advantages in chemical applications and also, in part, to the success in adapting and developing these methods to take advantage of these new hardware characteristics.

AUTHOR INFORMATION

Corresponding Author

*E-mail: shepard@tcg.anl.gov.
‡E-mail: szalay@chem.elte.hu.
§E-mail: th.mueller@fz-juelich.de.
∥E-mail: gidofalvi@gonzaga.edu.
⊥E-mail: hans.lischka@univie.ac.at.

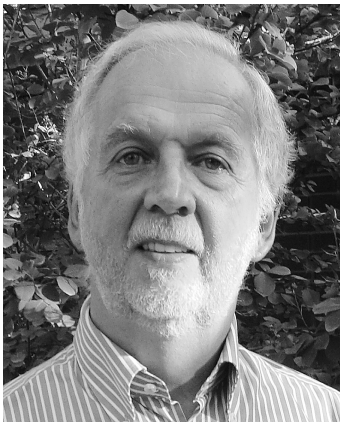
BIOGRAPHIES

Péter G. Szalay was born in Szentes, Hungary, in 1962. He received his M.Sc. degree in 1986 at Eötvös Loránd University (ELTE) in Budapest under the supervision of Géza Fogarasi and in 1989 his Ph.D. at the University of Vienna under the supervision of Hans Lischka. The latter work consisted of MR-CI calculations on the ground and excited states of conjugated molecules. At the same time he joined the COLUMBUS community and worked on improvements of the MR-CI code, including the first analytic derivative code for this type of wave function. Between 1991 and 1993 he was postdoc with Rod Bartlett at the Quantum Theory Project at the University of Florida, where he worked on the development of coupled-cluster methods, including the MR-AQCC method. Since 2003 he has been a Full Professor at ELTE and served as chairman of the Institute of Chemistry between 2005 and 2008. His reasearch, besides method development, concentrates on excited states and corresponding spectroscopy.

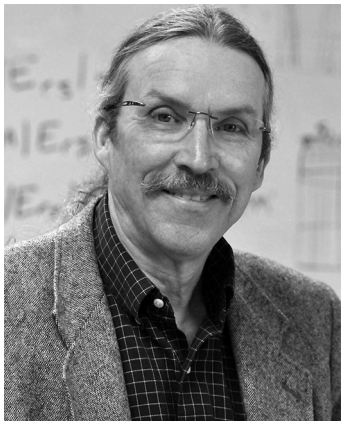
Thomas Müller was born near Verden, Germany, in 1966. He received his diploma degree in 1992 and his Ph.D. in 1995 at the University of Marburg. In the years 1996 to 2001 he worked as a

postdoctoral researcher with Hans Lischka at the University of Vienna and joined the development activities centered on the COLUMBUS quantum chemistry codes. Since 2001 he has been Senior Scientist at the Jülich Supercomputer Centre. His research interests primarily focus on the development of efficient parallel electronic structure methods with emphasis on multi-reference techniques and accurate calculations on ground and excited states of small molecules.

Gergely Gidofalvi was born in Budapest, Hungary, in 1977. He received his B.S. degree in 2002 from San Diego State University, his M.S. degree from The University of Chicago in 2003, and in 2006 his Ph.D. from The University of Chicago under the supervision of David A. Mazziotti. The latter work consisted of the development and application of variational reduced-density-matrix theory to strongly correlated systems. Between 2007 and 2010, he worked with Dr. Ron Shepard on the graphically contracted function method as a Director's Postdoctoral Fellow at Argonne National Laboratory. Since 2010 he has been an Assistant Professor at Gonzaga University. His research focuses on extending the graphically contracted function method to the computation of excited-state potential energy surfaces as well as the implementation of geometery optimization.



Hans Lischka was born in Vienna, Austria, in 1943 and is currently Research Professor of Chemistry at the Texas Tech University. He received his Ph.D. degree at the University of Vienna in 1969. In the years 1972 to 1973 he worked as a postdoctoral researcher with Professor Kutzelnigg at the University of Karlsruhe. After finishing his habilitation in 1976 he became Professor of Theoretical Chemistry at the University of Vienna in 1980. In the same year, during his stay as Visiting Professor at the Ohio State University, the development of the COLUMBUS program system was started in cooperation with Professor I. Shavitt and Dr. Ron Shepard. Research interests focus on simulaitons of nonadiabatic photodynamical processes and on solvent effects with particular emphasis on hydrogen bonded interactions.



Ron Shepard was born in Malvern, Arkansas, USA, in 1952. He was a Trinity Foundation Scholar at the University of Central Arkansas, where he received a B.S. with majors in Mathematics and Chemistry in 1975 under the supervision of Prof. J. M. Manion. He received a Ph.D. in Physical Chemistry from the University of Utah in 1980 under the supervision of Prof. J. Simons where he developed and applied electronic structure methods. He was a Battelle Memorial Postdoctoral Fellow under the supervision of Profs. I. Shavitt and R. J. Bartlett at Battelle Columbus Laboratories in Columbus, Ohio. His research involved the development of the COLUMBUS Program System in collaboration also with Prof. Hans Lischka who was a visiting faculty member, Prof. R. M. Pitzer, and the Ph.D. student F. B. Brown at The Ohio State University. He joined the Theoretical Chemistry Group at Argonne National Laboratory (ANL) in 1981 under the supervision of Dr. T. H. Dunning, Jr. He is now a Senior Scientist in the Gas Phase Dynamics Group at ANL, where he develops and applies electronic structure methods to address combustion chemistry problems.

ACKNOWLEDGMENT

The authors gratefully acknowledge many discussions with our senior collaborators, Profs. Isaiah Shavitt and Russell M. Pitzer, who have inspired and stimulated our interests in the topics discussed in this review. We also thank the reviewers for their careful reading of this long paper and for many useful suggestions. R.S. was supported by the Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences, U.S. Department of Energy under Contract DE-AC02-06CH11357. P.G.S. acknowledges financial support by the Hungarian American Enterprise Scholarship Foundation (HAESF) during his sabbatical stay at the University of Florida and by Orszagos Tudomanyos Kutatasi Alap (OTKA; Grant No. F72423). Contribution from TAMOP, supported by the European Union and cofinanced by the European Social Fund (Grant Agreement No. TAMOP 4.2.1/B-09/1/KMR-2010-0003) is also acknowledged. H.L. was supported by the Austrian Science Fund within the framework of the Special Research Program F41 (ViCoM); this work was performed as part of

research supported by the National Science Foundation Partnership in International Research and Education (PIRE) Grant No. OISE-0730114; support was also provided by the Robert A. Welch Foundation under Grant No. D-0005. G.G. was supported by an award from Research Corporation for Science Advancement and a grant to Gonzaga University from the Howard Hughes Medical Institute through the Undergraduate Science Education Program. T.M. acknowledges support by the John-von-Neumann Centre for Computing.

DEDICATION

[†]We dedicate this paper to Professor Isaiah Shavitt, our mentor, collaborator, and friend.

REFERENCES

(1) Shavitt, I. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977; pp 189−275.
(2) Duch, W. *J. Mol. Struct. (THEOCHEM)* **1991**, *80*, 27.
(3) Shavitt, I. *Mol. Phys.* **1998**, *94*, 3.
(4) Sherrill, C. D.; Schaefer, H. F. *Adv. Quantum Chem.* **1999**, *34*, 142.
(5) Čársky, P. Configuration Interaction. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; Wiley: New York, 2002.
(6) Karwowski, J. A.; Shavitt, I. Configuration Interaction. In *Handbook of Molecular Physics and Quantum Chemistry*, Vol. 2, Part 3;Wilson, S., Ed.; Wiley: Chichester, U.K., 2003; pp 227−271.
(7) See for example: Hirao, K., Ed. *Recent Advances in Multireference Methods*; World Scientific: London, 1999.
(8) Liu, B *J. Chem. Phys.* **1973**, *58*, 1925.
(9) Löwdin, P. O. *Phys. Rev.* **1955**, *97*, 1474.
(10) Davidson, E. R. *Reduced Density Matrices in Quantum Chemistry*; Academic Press: New York, 1976.
(11) Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300.
(12) Edmiston, C.; Ruedenberg, K. *Rev. Mod. Phys.* **1963**, *35*, 457.
(13) Weinstein, H.; Pauncz, R.; Cohen, M. *Adv. At. Mol. Phys.* **1971**, *7*, 97.
(14) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1988**, *90*, 4916.
(15) Kutzelnigg, W. *Theor. Chim. Acta* **1963**, *1*, 327.
(16) Ahlrichs, R.; Driessler, F. *Theor. Chim. Acta* **1975**, *36*, 275.
(17) Taylor, P. R. *J. Chem. Phys.* **1981**, *74*, 1256.
(18) Hunt, W. J.; Goddard, W. A., III. *Chem. Phys. Lett.* **1969**, *3*, 414.
(19) Shepard, R. *Adv. Chem. Phys.* **1987**, *69*, 63.
(20) McWeeny, R.; Sutcliffe, B. T. *Comput. Phys. Rep.* **1985**, *5*, 217.
(21) Shepard, R. In *Modern Electronic Structure Theory*, Part I; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; pp 345−458.
(22) Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* **1978**, *14*, 561.
(23) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359.
(24) Pople, J. A. *Int. J. Quantum Chem. Symp.* **1976**, *10*, 1.
(25) Keywords "*Size-consistency*" and "*Size-extensivity*": *Encyclopedia Computational Chemistry* (on line version); Schleyer, P. v. R., Alinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Eds.; Wiley: New York, 2005; DOI: 10.1002/0470845015.
(26) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics*; Cambridge University Press: Cambridge, U.K., 2009.
(27) Nooijen, M.; Shamasundar, K. R.; Mukherjee, D. *Mol. Phys.* **2005**, *103*, 2277.
(28) Shavitt, I. In *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements*, Lecture Notes in Chemistry 22; Hinze, J., Ed.; Springer-Verlag: Berlin, 1981; pp 51−99.
(29) Duch, W. *Lect. Notes Chem.* **1986**, *42*, 189.
(30) Condon, E. U.; Shortley, G. H. *The Theory of Atomic Spectra*; Cambridge University Press: New York, 1957.
(31) Szalay, P. G. Configuration Interaction: Corrections for Size-Consistency. In *Encyclopedia of Computational Chemistry* (on line version); Schleyer, P. v. R., Alinger, N. L., Clark T., Gasteiger J., Kollman, P. A., Schaefer H. F., Eds.; 2005; DOI: 10.1002/0470845015.cn0066.
(32) Kutzelnigg, W. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977; pp 129−188.
(33) Parlett, B. N. *The Symmetric Eigevalue Problem*; SIAM: Philadelphia, PA, USA, 1998.
(34) Roos, B. *Chem. Phys. Lett.* **1972**, *15*, 153.
(35) Roos, B.; Siegbahn, P. E. M. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977; pp 277−318.
(36) Davidson, E. R. *J. Comput. Phys.* **1975**, *17*, 87.
(37) Davidson, E. R. *Comput. Phys. Commun.* **1989**, *53*, 49.
(38) Davidson, E. R. *Comput. Phys.* **1993**, *7*, 519.
(39) Ahlrichs, R. In *Proceedings of the 5th Seminar on Computational Quantum Chemistry*; van Duijen, P. Th., Nieuwpoort, W. C., Eds.; Max Planck Institute: Munich, Germany, 1981; pp 142−153.
(40) Ahlrichs, R. In *Methods in Computational Molecular Physics*; Diercksen, G. H. F., Wilson, S., Eds.; Reidel: Dordrecht, The Netherlands, 1983; pp 209−226.
(41) Ahlrichs, R.; Böhm, H.-J.; Ehrhardt, C.; Scharf, P.; Schiffer, H.; Lischka, H.; Schindler, M. *J. Comput. Chem.* **1985**, *6*, 200.
(42) Saunders, V. R.; van Lenthe, J. H. *Mol. Phys.* **1983**, *48*, 923.
(43) Liu, B. In *Numerical Algorithms in Chemistry: Algebraic Methods*; Moler, C., Shavitt, I., Eds.; Lawrence Berkeley Laboratory: Berkeley, CA, USA, 1978, pp 49−53.
(44) van Lenthe, J. H.; Pulay, P. *J. Comput. Chem.* **1990**, *11*, 1164.
(45) Evangelisti, S.; Bendazzoli, G. L.; Ansaloni, R.; Duri, F.; Rossi, E. *Chem. Phys. Lett.* **1996**, *252*, 437.
(46) Knowles, P. J.; Handy, N. C. *Chem. Phys. Lett.* **1984**, *111*, 315.
(47) Knowles, P. J.; Handy, N. C. *Comput. Phys. Commun.* **1989**, *54*, 75.
(48) Rolik, Z.; Szabados, A.; Surjan, P. R. *J. Chem. Phys.* **2008**, *128*, 144101.
(49) Shepard, R. *J. Comput. Chem.* **1990**, *11*, 45.
(50) Dachsel, H.; Lischka, H. *Chem. Phys. Lett.* **1995**, *92*, 339.
(51) Knowles, P. J. *Chem. Phys. Lett.* **1989**, *155*, 513.
(52) Gansterer, W. N.; Kreuzer, W.; Lischka, H. *Lect. Notes Comput. Sci.* **2006**, *3743*, 564.
(53) Beebe, N. H. F.; Linderberg, J. *Int. J. Quantum Chem.* **1977**, *12*, 683.
(54) O'Neal, D. W.; Simons, J. *Int. J. Quantum Chem.* **1989**, *36*, 673.
(55) Pedersen, T. B.; Aquilante, F.; Lindh, R. *Theor. Chem. Acc.* **2009**, *124*, 1.
(56) Olsen, J.; Jørgensen, P.; Simons, J. *Chem. Phys. Lett.* **1990**, *169*, 463.
(57) Sleijpen, G. L. G.; Booten, A. G. L.; Fokkema, D. R.; van der Vorst, H. A. *BIT Num. Math.* **1996**, *36*, 595.
(58) Sleijpen, G. L. G.; van der Vorst, H. A. *J. Matrix Anal. Appl.* **1996**, *17*, 401.
(59) van Dam, H. J. J.; van Lenthe, J. H.; Sleijpen, G. L. G.; van der Vorst, H. A. *J. Comput. Chem.* **1996**, *17*, 267.
(60) Shepard, R.; Wagner, A. F.; Tilson, J. L.; Minkoff, M. *J. Comput. Phys.* **2001**, *172*, 472.
(61) Fulsher, M. P.; Widmark, P. O. *J. Comput. Chem.* **1993**, *14*, 8.
(62) Shepard, R.; Shavitt, I.; Lischka, H. *J. Comput. Chem.* **2002**, *23*, 1121.
(63) Zhou, Y.; Shepard, R.; Minkoff, M. *Comput. Phys. Commun.* **2005**, *167*, 90.
(64) Shepard, R.; Minkoff, M.; Zhou, Y. *Comput. Phys. Commun.* **2005**, *170*, 109.
(65) Bartlett, R. J.; Musial, M. *Rev. Mod. Phys.* **2007**, *79*, 291.
(66) Gauss, J. In *Encyclopedia of Comutational Chemistry*; Schleyer, P. v. R., Ed.; Wiley: New York, **1998**; Vol. *I*, pp 615−636.
(67) Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697.
(68) Bartlett, R. J. In *Modern Electronic Structure Theory*, Part I; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; Vol. 2, p 1047.
(69) Bartlett, R. J.; Stanton, J. F. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH: New York, 1994; Vol. 5, p 65.

(70) Lee, T. J.; Scuseria, G. E. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 1995.
(71) Sinanoglu, O. *J. Chem. Phys.* **1962**, *36*, 706.
(72) Davidson, E. R. In *The World of Quantum Chemistry*; Daudel, R., Pullman, B., Eds.; Reidel: Dodrecht, The Netherlands, 1974.
(73) Langhoff, S. R.; Davidson, E. R. *Int. J. Quantum Chem.* **1974**, *8*, 61.
(74) Bartlett, R. J.; Shavitt, I. *Int. J. Quantum Chem.* **1977**, *S11*, 165. Erratum:Bartlett, R. J.; Shavitt, I. *Int. J. Quantum Chem.* **1978**, *S12*, 543.
(75) Luken, W. L. *Chem. Phys. Lett.* **1978**, *58*, 421.
(76) Davidson, E. R.; Silver, D. W. *Chem. Phys. Lett.* **1977**, *52*, 403.
(77) Siegbahn, P E. M. *Chem. Phys. Lett.* **1978**, *55*, 386.
(78) Paldus, J.; Wormer, P. E. S.; Visser, F.; Vanderavoird, A. *J. Chem. Phys.* **1982**, *76*, 2458.
(79) Pople, J. A.; Seeger, R.; Krishnan, R. *Int. J. Quantum Chem.* **1977**, 149.
(80) Meissner, L. *Chem. Phys. Lett.* **1988**, *146*, 204.
(81) Gdanitz, R.; Ahlrichs, R. *Chem. Phys. Lett.* **1988**, *143*, 413.
(82) Burton, P.; Peyerimhoff, S. D.; Buenker, R. J. *Chem. Phys.* **1982**, *73*, 83.
(83) Duch, W.; Diercksen, G. H. F. *J. Chem. Phys.* **1994**, *101*, 3018.
(84) Meissner, L. *Chem. Phys. Lett.* **1996**, *263*, 351.
(85) Zalesny, R.; Sadlej, A. J.; Leszczynski, J. *Struct. Chem.* **2004**, *15*, 379.
(86) Bartlett, R. J. *Int. J. Mol. Sci.* **2002**, *3*, 579.
(87) Hanrath, M. *Chem. Phys. Lett.* **2006**, *420*, 426.
(88) Evangelista, F. A.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2007**, *127*, 024102.
(89) Butscher, W.; Shih, S.; Buenker, R. J.; Peyerimhoff, S. D. *Chem. Phys.* **1977**, *52*, 457.
(90) Hirsch, G.; Bruna, P. J.; Peyerimhoff, S. D.; Buenker, R. J. *Chem. Phys. Lett.* **1977**, *52*, 442.
(91) Bruna, P. J.; Peyerimhoff, S. D.; Buenker, R. J. *Chem. Phys. Lett.* **1980**, *72*, 278.
(92) Buenker, R. J.; Peyerimhoff, S. D.; Bruna, P. In *Computational Theoretical Organic Chemistry*; Csizmadia, I. G., Daudel, R., Eds.; Reidel: Dordrecht, The Netherlands, 1981.
(93) Prime, S.; Rees, C.; Robb, M. A. *Mol. Phys.* **1981**, *44*, 173.
(94) Simons, J. *J. Phys. Chem.* **1989**, *93*, 626.
(95) Burton, P. G. *Int. J. Quantum Chem.* **1983**, *23*, 613.
(96) Burton, P. G.; Buenker, R. J.; Bruna, P. J.; Peyerimhoff, S. D. *Chem. Phys. Lett.* **1983**, *95*, 379.
(97) Bauschlicher, C. W. *J. Chem. Phys.* **1988**, *88*, 1743.
(98) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1987**, *86*, 2443.
(99) Ackermann, J.; Hogreve, H. *Chem. Phys.* **1991**, *157*, 75.
(100) Shavitt, I.; Brown, F.; Burton, P. G. *Int. J. Quantum Chem.* **1987**, *31*, 507.
(101) Jankowski, K.; Meissner, L.; Wasilewski, J. *Int. J. Quantum Chem.* **1985**, *28*, 931.
(102) Blomberg, M.; Siegbahn, P. E. M. *J. Chem. Phys.* **1983**, *78*, 5682.
(103) Ruttink, P.; Vanlenthe, J.; Zwaans, R.; Groenenboom, G. *J. Chem. Phys.* **1991**, *94*, 7212.
(104) Verbeek, J.; Vanlenthe, J.; Timmermans, P.; Mackor, A.; Budzelaar, P. *J. Org. Chem.* **1987**, *52*, 2955.
(105) Werner, H.-J.; Kallay, M.; Gauss, J. *J. Chem. Phys.* **2008**, *128*, 034305.
(106) Khait, Y. G.; Jiang, W.; Hoffmann, M. R. *Chem. Phys. Lett.* **2010**, *493*, 1.
(107) Szalay, P. G. *Chem. Phys.* **2008**, *349*, 121.
(108) Meissner, L.; Grabowski, I. *Chem. Phys. Lett.* **1999**, *300*, 53.
(109) Meissner, L.; Nooijen, M. *Chem. Phys. Lett.* **2000**, *316*, 501.
(110) Meissner, L. *J. Mol. Struct. (THEOCHEM)* **2006**, *768*, 63.
(111) Li, X.; Paldus, J. *J. Chem. Phys.* **1997**, *107*, 6257.
(112) Li, X.; Paldus, J. *J. Chem. Phys.* **1998**, *108*, 637.
(113) Meissner, L.; Gryniakow, J.; Hubac, I. *Mol. Phys.* **2005**, *103*, 2173.
(114) Meller, J.; Malrieu, J.; Heully, J. *Chem. Phys. Lett.* **1995**, *244*, 440.
(115) Hubac, I.; Mach, P.; Wilson, S. *Int. J. Quantum Chem.* **2002**, *89*, 198.
(116) Hubac, I.; Mach, P.; Wilson, S. *Mol. Phys.* **2003**, *101*, 3493.
(117) Masik, J.; Hubac, I. *Adv. Quantum Chem.* **1981**, *31*, 75.
(118) Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Muller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664.
(119) Lischka, H.; Müller, T.; Szalay, P. G.; Shavitt, I.; Pitzer, R. M.; Shepard, R. *WIREs: Comput. Mol. Sci.* **2011**, *1*, 191.
(120) Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Schütz, M.; Celani, P.; Knizia, G.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R. M.; Reiher, M.; Shiozaka, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. MOLPRO, version 2010.1, a package of ab initio programs. See: http://www.molpro.net.
(121) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Jensen, J. H.; Koseki, S.; Gordon, M. S.; Nguyen, K. A.; Windus, T. L.; Elbert, S. T. *QCPE Bull.* **1990**, *10*, 52–54.
(122) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
(123) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Neogrády, P.; Pedersen, T. B.; Pitonak, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224.
(124) Ben Amor, N.; Maynau, D. *Chem. Phys. Lett.* **1998**, *286*, 211.
(125) Olsen, J.; Jørgensen, P.; Koch, H.; Balkova, A.; Bartlett, R. J. *J. Chem. Phys.* **1996**, *104*, 8007.
(126) Brown, F. B.; Truhlar, D. G. *Chem. Phys. Lett.* **1985**, *117*, 307.
(127) Kelly, H.; Sessler, A. *Phys. Rev.* **1963**, *132*, 2091.
(128) Kelly, H. *Phys. Rev.* **1964**, *134*, 1450.
(129) Meyer, W. *Int. J. Quantum Chem.* **1971**, *5*, 341.
(130) Meyer, W. *J. Chem. Phys.* **1973**, *58*, 1017.
(131) Meyer, W. *Theor. Chim. Acta* **1974**, *35*, 277.
(132) Ahlrichs, R.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *62*, 1225.
(133) Ahlrichs, R.; Driessler, F.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *62*, 1235.
(134) Ahlrichs, R.; Keil, F.; Lischka, H.; Kutzelnigg, W.; Staemmler, V. *J. Chem. Phys.* **1975**, *63*, 455.
(135) Ahlrichs, R.; Lischka, H.; Zurawski, B.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *63*, 4685.
(136) Cizek, J. *J. Chem. Phys.* **1966**, *45*, 4256.
(137) Pulay, P. *Int. J. Quantum Chem.* **1983**, *24*, 257.
(138) Ahlrichs, R.; Scharf, P.; Ehrhardt, C. *J. Chem. Phys.* **1985**, *82*, 890.
(139) Neese, F.; Hansen, A.; Wennmohs, F.; Grimme, S. *Acc. Chem. Res.* **2009**, *42*, 641.
(140) Szalay, P. G. *Mol. Phys.* **2010**, *108*, 3055.
(141) Koch, S.; Kutzelnigg, W. *Theor. Chim. Acta* **1981**, *59*, 387.
(142) Szalay, P. G.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *103*, 3600.
(143) Szalay, P. G. Towards state-specific formulation of multi-reference coupled-cluster theory: Coupled electron pair approximations (CEPA) leading to multireference configuration interaction (MR-CI) type equations. In *Recent Advances in Coupled-Cluster Methods*; Bartlett, R. J., Ed.; World Scientific: Singapure, 1997; p 81.
(144) Laidig, W.; Bartlett, R. *Chem. Phys. Lett.* **1984**, *104*, 424.
(145) Laidig, W.; Saxe, P.; Bartlett, R. *J. Chem. Phys.* **1987**, *86*, 887.
(146) Paldus, J. In *New Horizons of Quantum Chemistry*; Löwdin, P.-O., Pullman, B., Eds.; Reidel: Dordrecht, The Netherlands, 1983; p 31.
(147) Hoffmann, M. R.; Simons, J. *J. Chem. Phys.* **1989**, *90*, 3671.

(148) Hoffmann, M. R.; Simons, J. *J. Chem. Phys.* **1988**, *88*, 993.
(149) Cave, R. J.; Davidson, E. R. *J. Chem. Phys.* **1988**, *88*, 5770.
(150) Cave, R. J.; Davidson, E. R. *J. Chem. Phys.* **1988**, *89*, 6798.
(151) Murray, C.; Racine, S. C.; Davidson, E. R. *Int. J. Quantum Chem.* **1992**, *42*, 273.
(152) Fulde, P.; Stoll, H. *J. Chem. Phys.* **1992**, *97*, 4185.
(153) Fink, R.; Staemmler, V. *Theor. Chim. Acta* **1993**, *87*, 129.
(154) Tanaka, K.; Sakai, T.; Terashima, H. *Theor. Chim. Acta* **1989**, *76*, 213.
(155) Sakai, T.; Tanaka, K. *Theor. Chim. Acta* **1993**, *85*, 451.
(156) Tanaka, K.; Mochizuki, Y. *Theor. Chem. Acc.* **1997**, *98*, 165.
(157) Tanaka, K.; Ghosh, T.; Sakai, T. *Int. J. Quantum Chem.* **1999**, *74*, 661.
(158) Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481.
(159) Szalay, P. G.; Müller, T.; Lischka, H. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2067.
(160) Füsti-Molnár, L.; Szalay, P. G. *J. Phys. Chem.* **1996**, *100*, 6288.
(161) Füsti-Molnár, L.; Szalay, P. G. *Chem. Phys. Lett.* **1996**, *258*, 400.
(162) Gdanitz, R. J. *Int. J. Quantum Chem.* **2001**, *85*, 281.
(163) Malrieu, J.-P.; Daudey, J.-P.; Caballol, R. *J. Chem. Phys.* **1994**, *101*, 8908.
(164) Daudey, J.-P.; Heully, J.-L.; Malrieu, J.-P. *J. Chem. Phys.* **1993**, *99*, 1240.
(165) Heully, J.; Malrieu, J. *Chem. Phys. Lett.* **1992**, *199*, 545.
(166) Meller, J.; Malrieu, J.-P.; Heully, J.-L. *Mol. Phys.* **2003**, *101*, 2029.
(167) Pahari, D.; Chattopadhyay, S.; Deb, A.; Mukherjee, D. *Chem. Phys. Lett.* **2004**, *386*, 307.
(168) Chattopadhyay, S.; Mahapatra, U.; Datta, B.; Mukherjee, D. *Chem. Phys. Lett.* **2002**, *357*, 426.
(169) Chattopadhyay, S.; Pahari, D.; Mukherjee, D.; Mahapatra, U. *J. Chem. Phys.* **2004**, *120*, 5968.
(170) Mahapatra, U. S.; Datta, B.; Mukherjee, D. *Mol. Phys.* **1998**, *94*, 157.
(171) Mahapatra, U. S.; Datta, B.; Bandyopadhyay, B.; Mukherjee, D. *Adv. Quantum Chem.* **1998**, *30*, 163.
(172) Mahapatra, U.; Datta, B.; Mukherjee, D. *J. Chem. Phys.* **1999**, *110*, 6171.
(173) Jeziorski, B.; Monkhorst, H. *Phys. Rev. A* **1981**, *24*, 1668.
(174) Chattopadhyay, S.; Mahapatra, U. *J. Phys. Chem. A* **2004**, *108*, 11664.
(175) Pahari, D.; Ghosh, P.; Mukherjee, D.; Chattopadhyay, S. *Theor. Chem. Acc.* **2006**, *116*, 621.
(176) Ruttink, P. J. A.; van Lenthe, J. H.; Todorov, P. *Mol. Phys.* **2005**, *103*, 2497.
(177) Ben-Amor, N.; Maynau, D.; Malrieu, J.-P.; Monari, A. *J. Chem. Phys.* **2008**, *129*, 064112.
(178) Kállay, M.; Szalay, P. G.; Surján, P. R. *J. Chem. Phys.* **2002**, *117*, 980.
(179) Das, S.; Mukherjee, D.; Kállay, M. *J. Chem. Phys.* **2010**, *132*, 074103.
(180) Das, S.; Kállay, M.; Mukherjee, D. *J. Chem. Phys.* **2010**, *133*, 234110.
(181) Evangelista, F. A.; Gauss, J. *J. Chem. Phys.* **2010**, *133*, 044101.
(182) Guest, M. F.; van Lenthe, J. H.; Kendrick, J.; Schöffel, K.; Sherwood, P.; Harrison, R. J. *GAMESS-UK*, a package of ab initio programs; Computing for Science: 2002 (with contributions from: Amos, R. D.; Buenker, R. J.; Dupuis, M.; Handy, N. C.; Hillier, I. H.; Knowles, P. J.; Bonačić-Koutecký, V.; von Niessen, W.; Saunders V. R.; Stone. A. J.; it is derived from the original GAMESS code due to Dupuis, M.; Spangler, D.; Wendolowski, J., *NRCC Software Catalog*, Vol. *1*, Program No. QG01 (GAMESS), 1980).
(183) Lengsfield, B. H. *J. Chem. Phys.* **1980**, *73*, 382.Lengsfield, B. H.; Liu, B. *J. Chem. Phys.* **1981**, *75*, 475, Liu, B.; Yoshimine, M. *J. Chem. Phys.* **1981**, *74*, 612.
(184) Kállay, M. *MRCC*, a string-based quantum chemical program suite, http://www.mrcc.hu.
(185) Staemmler, V. *Theor. Chim. Acta* **1977**, *45*, 89.
(186) Wasilewski, J. *Int. J. Quantum Chem.* **1989**, *36*, 503.
(187) Meier, U.; Staemmler, V. *Theor. Chim. Acta* **1989**, *76*, 95.
(188) Sanchezmarin, J.; Nebot-Gil, I.; Maynau, D.; Malrieu, J.-P. *Theor. Chim. Acta* **1995**, *92*, 241.
(189) Nebot-Gil, I.; Sanchezmarin, J.; Malrieu, J.-P.; Heully, J.-L.; Maynau, D. *J. Chem. Phys.* **1995**, *103*, 2576.
(190) Meissner, L. *Int. J. Quantum Chem.* **2008**, *108*, 2199.
(191) Meissner, L. *Mol. Phys.* **2010**, *108*, 2961.
(192) Nooijen, M.; Roy, R. J. L. *J. Mol. Struct. (THEOCHEM)* **2006**, *768*, 25.
(193) Sherrill, C.; Schaefer, H. *J. Phys. Chem.* **1996**, *100*, 6069.
(194) Sychrovsky, V.; Čársky, P. *Mol. Phys.* **1996**, *88*, 1137.
(195) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *121*, 10905.
(196) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *121*, 10919.
(197) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2005**, *122*, 154110.
(198) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2006**, *124*, 174304.
(199) Bytautas, L.; Nagata, T.; Gordon, M. S.; Ruedenberg, K. *J. Chem. Phys.* **2007**, *127*, 164317.
(200) Bytautas, L.; Matsunaga, N.; Nagata, T.; Gordon, M. S.; Ruedenberg, K. *J. Chem. Phys.* **2007**, *127*, 204301.
(201) Bytautas, L.; Matsunaga, N.; Nagata, T.; Gordon, M. S.; Ruedenberg, K. *J. Chem. Phys.* **2007**, *127*, 204313.
(202) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2010**, *132*, 074109.
(203) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2010**, *132*, 074307.
(204) Khait, Y.; Song, H.; Hoffmann, M. *Chem. Phys. Lett.* **2003**, *372*, 674.
(205) Meyer, W. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977; pp 413−446.
(206) Siegbahn, P. *Int. J. Quantum Chem.* **1980**, *18*, 1229.
(207) Werner, H.-J.; Reinsch, E.-A. *J. Chern. Phys.* **1982**, *76*, 3144.
(208) Shavitt, I. *Int. J. Mol. Sci.* **2002**, *3*, 639.
(209) Werner, H.-J.; Knowles, P. *J. Chem. Phys.* **1988**, *89*, 5803.
(210) Knowles, P.; Werner, H. *Chem. Phys. Lett.* **1988**, *145*, 514.
(211) Werner, H.-J.; Knowles, P. J. *Theor. Chim. Acta* **1990**, *78*, 175.
(212) Knowles, P. J.; Werner, H.-J. *Theor. Chim. Acta* **1992**, *84*, 95.
(213) Dobbyn, A.; Knowles, P. J.; Harrison, R. J. *J. Comput. Chem.* **1998**, *19*, 1215.
(214) Werner, H.-J. *Adv. Chem. Phys.* **1987**, *69*, 1.
(215) Peterson, K. A.; Kendall, R. A.; Dunning, T. H. *J. Chem. Phys.* **1993**, *99*, 1930.
(216) Peterson, K. A.; Kendall, R. A.; Dunning, T. H. *J. Chem. Phys.* **1993**, *99*, 9790.
(217) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-A.; Serrano-Andres, L. *Adv. Chem. Phys.* **1996**, *93*, 219.
(218) Roos, B. O.; Fülscher, M. P., Malmqvist, P.-A.; Merchan, M.; Serrano-Andres, L. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer: Dordrecht, The Netherlands, 1995.
(219) Roos, B. *Adv. Chem. Phys.* **1987**, *69*, 399.
(220) Andersson, K.; Malmqvist, P.-A.; Roos, B. O.; Sadlej, A. J.; Wolinski., K. *J. Phys. Chem.* **1990**, *94*, 5483.
(221) Andersson, K.; Malmqvist, P.-A.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
(222) Siegbahn, P. E. M. *Int. J. Quantum Chem.* **1983**, *23*, 1869.
(223) Siegbahn, P. E. M. *Chem. Phys.* **1977**, *25*, 197.
(224) Lee, T. *J. Chem. Phys.* **1987**, *87*, 2825.
(225) Wang, Y.; Gan, Z.; Su, K.; Wen, Z. *Chem. Phys. Lett.* **1999**, *312*, 277.
(226) Wang, Y.; Suo, B.; Zhai, G.; Wen, Z. *Chem. Phys. Lett.* **2004**, *389*, 315.
(227) Wang, Y.; Zhai, G.; Suo, B.; Gan, Z.; Wen, Z. *Chem. Phys. Lett.* **2003**, *375*, 134.
(228) Shepard, R. *J. Phys. Chem. A* **2005**, *109*, 11629.
(229) Shepard, R. *J. Phys. Chem. A* **2006**, *110*, 8880.
(230) Shepard, R.; Minkoff, M. *Int. J. Quantum Chem.* **2006**, *106*, 3190.
(231) Shepard, R.; Minkoff, M.; Brozell, S. R. *Int. J. Quantum Chem.* **2007**, *107*, 3203.

(232) Brozell, S. R.; Shepard, R.; Zhang, Z. *Int. J. Quantum Chem.* **2007**, *107*, 3191.
(233) Gidofalvi, G.; Shepard, R. *J. Comput. Chem.* **2009**, *30*, 2414.
(234) Gidofalvi, G.; Shepard, R. *Int. J. Quantum Chem.* **2009**, *109*, 3552.
(235) Brozell, S. R.; Shepard, R. *J. Phys. Chem. A* **2009**, *113*, 12741.
(236) Shepard, R.; Gidofalvi, G.; Hovland, P. D. *Int. J. Quantum Chem.* **2010**, *110*, 2938.
(237) Gidofalvi, G.; Shepard, R. *Mol. Phys.* **2010**, *108*, 2717.
(238) Shavitt, I. *Int. J. Quantum Chem.* **1977**, *S11*, 131.
(239) Shavitt, I. In *Mathematical Frontiers in Computational Physics*; Truhlar, D. G., Ed.; Springer-Verlag: New York, 1988; p 300.
(240) Paldus, J. *J. Chem. Phys.* **1974**, *61*, 5321.
(241) Paldus, J., In *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements*; Lecture Notes in Chemistry 22; Hinze, J., Ed.; Springer-Verlag: Berlin, 1981; pp 1−50.
(242) Paldus, J. In *Mathematical Frontiers in Computational Physics*; Truhlar, D. G., Ed.; Springer-Verlag: New York, 1988; p 262.
(243) Gidofalvi, G.; Shepard, R. Abstracts of Papers, 238th ACS National Meeting, Washington, DC, United States, Aug. 16−20, 2009, PHYS-198.
(244) White, S. R.; Martin, R. L. *Phys. Rev. Lett.* **1992**, *69*, 2863.
(245) White, S. R.; Martin, R. L. *Phys. Rev. B* **1993**, *48*, 10345.
(246) White, S. R.; Martin, R. L. *J. Chem. Phys.* **1999**, *110*, 4127.
(247) Daul, S.; Ciofini, I.; Daul, C.; White, S. R. *Int. J. Quantum Chem.* **2000**, *79*, 331.
(248) Chan, G. K.-L.; Head-Gordon, M. *J. Chem. Phys.* **2002**, *116*, 4462.
(249) Chan, G. K.-L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *118*, 8551.
(250) Chan, G. K.-L. *J. Chem. Phys.* **2004**, *120*, 3172.
(251) Chan, G. K.-L.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *121*, 6110.
(252) Chan, G. K.-L.; Van Voorhis, T. *J. Chem. Phys.* **2004**, *122*, 204101.
(253) Hachmann, J.; Cardoen, W.; Chan, G. K.-L. *J. Chem. Phys.* **2006**, *125*, 144101.
(254) Hachmann, J.; Dorando, J. J.; Avilés, M.; Chan, G. K.-L. *J. Chem. Phys.* **2007**, *127*, 134309.
(255) Dorando, J. J.; Hachmann, J.; Chan, G. K.-L. *J. Chem. Phys.* **2007**, *127*, 084109.
(256) Ghosh, D.; Hachmann, J.; Yanai, T.; Chan, G. K.-L. *J. Chem. Phys.* **2008**, *128*, 144117.
(257) Chan, G. K.-L. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3454–3459.
(258) Dorando, J. J.; Hachmann, J.; Chan, G. K.-L. *J. Chem. Phys.* **2009**, *130*, 184111.
(259) Zgid., D.; Nooijen, M. *J. Chem. Phys.* **2008**, *128*, 014107.
(260) Zgid., D.; Nooijen, M. *J. Chem. Phys.* **2008**, *128*, 144115.
(261) Zgid., D.; Nooijen, M. *J. Chem. Phys.* **2008**, *128*, 144116.
(262) Mitrushenkov, A. O.; Fano, G.; Ortolani, F.; Linguerri, R.; Palmieri, P. *J. Chem. Phys.* **2001**, *115*, 6815.
(263) Mitrushenkov, A. O.; Linguerri, R.; Palmieri, P. *J. Chem. Phys.* **2003**, *119*, 4148.
(264) Kurashige, Y.; Yanai, T. *J. Chem. Phys.* **2009**, *130*, 234114.
(265) Mizukami, W.; Kurashige, Y.; Yanai, T. *J. Chem. Phys.* **2010**, *133*, 091101.
(266) Legeza, Ö.; Röder, J.; Hess, B. A. *Mol. Phys.* **2003**, *101*, 2019.
(267) Legeza, Ö.; Röder, J.; Hess, B. A. *Phys. Rev. B* **2003**, *67*, 125114.
(268) Moritz, G.; Hess, B. A.; Reiher, M. *J. Chem. Phys.* **2005**, *122*, 024107.
(269) Moritz, G.; Wolf, A.; Reiher, M. *J. Chem. Phys.* **2005**, *123*, 184105.
(270) Marti, K. H.; Ondík, I. M.; Moritz, G.; Reiher, M. *J. Chem. Phys.* **2008**, *128*, 014104.
(271) Marti, K. H.; Reiher, M. *Mol. Phys.* **2010**, *101*, 501.
(272) Schollwöck, U. *Rev. Mod. Phys.* **2005**, *77*, 259.
(273) Schollwöck, U. *Ann. Phys.* **2011**, *326*, 96.
(274) Buenker, R. J.; Peyerimhoff, S. D. *Theor. Chim. Acta* **1974**, *35*, 33.
(275) Buenker, R. J.; Peyerimhoff, S. D. *Theor. Chim. Acta* **1975**, *39*, 217.
(276) Buenker, R. J.; Peyerimhoff, S. D.; Butscher, W. *Mol. Phys.* **1978**, *35*, 771.
(277) Engels, B.; Hanrath, M.; Lennartz, C. *Comput. Chem.* **2001**, *25*, 15.
(278) Buenker, R. J., Peyerimhoff, S. D. In *New Horizon in Quantum Chemistry*; Löwdin, P. O., Pullmann, B., Eds.; Reidel: Dordrecht, The Netherlands, 1983.
(279) Whitten, J.; Hackmeyer, M. *J. Chem. Phys.* **1969**, *51*, 5584.
(280) Huron, B.; Malrieu, J.; Rancurel, P. *J. Chem. Phys.* **1973**, *58*, 5745.
(281) Evangelisti, S.; Daudey, J.-P; Malrieu, J.-P. *Chem. Phys.* **1983**, *75*, 91.
(282) McMurchie, L.; Elbert, S. T.; Langhoff, S. R.; Davidson, E. R.; Feller, D. *MELDF Program Package*; Indiana University: Bloomington, IN, USA, 1989.
(283) McMurchie, L.; Elbert, S.; Langhoff, S.; Davidson, E. R. Thesis, MELDF, The University of Washington, Indiana University, and the Environmental Molecular Sciences Laboratory, 1997. The MELDF suite of programs was substantially modified by Feller, D.; Cave, R.; Rawlings, D.;Frey, R.; Daasch, R.; Nitchie, L.; Phillips, P.; Iberle, K.; Jackels, C.; Davidson, E. R. 1997.
(284) Harrison, R. J. *J. Chem. Phys.* **1991**, *94*, 5021.
(285) Caballol, R.; Malrieu, J.-P. *Chem. Phys. Lett.* **1992**, *188*, 543.
(286) Povill, A.; Rubio, J.; Illas, F. *Theor. Chim. Acta* **1992**, *82*, 229.
(287) Krebs, S.; Buenker, R. *J. Chem. Phys.* **1995**, *103*, 5613.
(288) Hanrath, M.; Engels, B. *Chem. Phys.* **1997**, *225*, 197.
(289) Stampfuss, P; Wenzel, W; Keiter, H. *J. Comput. Chem.* **1999**, *20*, 1559.
(290) Stampfuss, P; Hamcher, K.; Wenzel, W. *J. Mol. Struct. (THEOCHEM)* **2000**, *506*, 99.
(291) Bytautas, L; Ruedenberg, K. *Chem. Phys.* **2009**, *356*, 64.
(292) Saebo, S.; Pulay, P. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213.
(293) Walter, D.; Carter, E. A. *Chem. Phys. Lett.* **2001**, *346*, 177.
(294) Walter, D.; Venkatnathan, A.; Carter, E. A. *J. Chem. Phys.* **2003**, *118*, 8127.
(295) Venkatnathan, A.; Szilva, A.; Walter, D.; Gdanitz, R. J.; Carter, E. A. *J. Chem. Phys.* **2004**, *120*, 1693.
(296) Bories, B.; Maynau, D.; Bonnet, M.-L. *J. Comput. Chem.* **2007**, *28*, 632.
(297) Maynau, D.; Evangelisti, S.; Guihéry, N.; Calzado, C. J.; Malrieu, J.-P. *J. Chem. Phys.* **2002**, *116*, 10060.
(298) Chwee, T. S.; Carter, E. A. *J. Chem. Phys.* **2008**, *128*, 224106.
(299) Chwee, T. S.; Carter, E. A. *J. Chem. Phys.* **2010**, *132*, 074104.
(300) Chwee, T. S.; Carter, E. A. *Mol. Phys.* **2010**, *108*, 2519.
(301) Chwee, T. S.; Carter, E. A. *J. Chem. Theory Comput.* **2011**, *7*, 103.
(302) Reinhardt, P.; Zhang, H.; Ma, J.; Malrieu, J. P. *J. Chem. Phys.* **2008**, *129*, 164106.
(303) Zhang, H.; Malrieu, J.-P.; Reinhardt, P.; Ma, J. *J. Chem. Phys.* **2010**, *132*, 034108.
(304) Friesner, R. *Annu. Rev. Phys. Chem.* **1991**, *42*, 341.
(305) *SCHRÖDINGER*, http://www.schrodinger.com.
(306) Martinez, T. J.; Carter, E. A. In *Modern Electronic Structure Theory*, Vol. II; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; p 1132.
(307) Martinez, T. J.; Mehta, A.; Carter, E. A. *J. Chem. Phys.* **1992**, *97*, 1876. **1993**, *99*, 4238 (Erratum).
(308) Martinez, T. J.; Carter, E. A. *J. Chem. Phys.* **1993**, *98*, 7081.
(309) Murphy, R.; Friesner, R.; Ringnalda, M.; Goddard, W. *J. Chem. Phys.* **1994**, *101*, 2986.
(310) Martinez, T. J.; Carter, E. A. *J. Chem. Phys.* **1995**, *102*, 7564.
(311) Martinez, T. J.; Carter, E. A. *Chem. Phys. Lett.* **1995**, *241*, 490.
(312) Carter, E. A.; Walter, D. Reduced Scaling Electron Correlation Methods. In *Encyclopedia Computational Chemistry* (on line ed.);

Schleyer, P. v. R., Alinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Eds.; Wiley: Chichester, U.K., 2004; DOI: 10.1002/0470845015.cu0024.
(313) Reynolds, G.; Martinez, T. J.; Carter, E. A. *J. Chem. Phys.* **1996**, *105*, 6455.
(314) Reynolds, G.; Carter, E. A. *Chem. Phys. Lett.* **1997**, *265*, 660.
(315) Walter, D.; Szilva, A.; Niedfeldt, K.; Carter, E. A. *J. Chem. Phys.* **2002**, *117*, 1982.
(316) Brandow, B. H. *Rev. Mod. Phys.* **1967**, *39*, 771.
(317) Lindgren, I. *J. Phys. B: At. Mol. Phys.* **1974**, *7*, 2241.
(318) Hose, G.; Kaldor., U. *J. Phys. B: At. Mol. Phys.* **1979**, *12*, 3827.
(319) Meissner, L.; Bartlett., R. J. *J. Chem. Phys.* **1989**, *91*, 4800.
(320) Finley, J. P.; Freed, K. F. *J. Chem. Phys.* **1995**, *102*, 1306.
(321) Malrieu, J.-P.; Heully, J.-L.; Zaitsevskii, A. *J. Phys. A: Math. Gen.* **1985**, *18*, 809.
(322) Kozlowski, P. M.; Davidson, E. R. *Chem. Phys. Lett.* **1994**, *222*, 615.
(323) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.
(324) Wolinski, K.; Pulay, P. *J. Chem. Phys.* **1989**, *90*, 3647.
(325) Murphy, R. B.; Messmer, R. P. *Chem. Phys. Lett.* **1991**, *183*, 443.
(326) Murphy, R. B.; Messmer, R. P. *J. Chem. Phys.* **1992**, *97*, 4170.
(327) Kozlowski, P. M.; Davidson, E. R. *J. Chem. Phys.* **1992**, *100*, 3672.
(328) Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.
(329) Hirao, K. *Chem. Phys. Lett.* **1993**, *201*, 59.
(330) Pulay, P. *Int. J. Quantum Chem.* **2011**, *111*, 3273.
(331) Dyall, K. G. *J. Chem. Phys.* **1995**, *102*, 4909.
(332) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. *J. Chem. Phys.* **2001**, *114*, 10252.
(333) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *J. Chem. Phys.* **2002**, *117*, 297.
(334) Angeli, C.; Borini, S.; Cestari, M.; Cimiraglia, R. *J. Chem. Phys.* **2004**, *121*, 4043.
(335) Stahlberg, E. A. Application of Multireference Based Correlation Methods to the Study of Weak Bonding Interactions. Ph.D. Dissertation, Ohio State University, Columbus, OH, 1991.
(336) Khait, Y. G.; Song, J.; Hoffmann, M. R. *J. Chem. Phys.* **2002**, *117*, 4133.
(337) Hoffmann, M. R. *J. Phys. Chem.* **1996**, *100*, 6125.
(338) Hoffmann, M. R. *Chem. Phys. Lett.* **1992**, *195*, 127.
(339) Hoffmann, M. R. *Chem. Phys. Lett.* **1993**, *210*, 193.
(340) Theis, D.; Khait, Y. G.; Hoffmann, M. R. *J. Chem. Phys.* **2011**, *135*, 044117.
(341) Rolik, Z.; Szabados, Á.; Surján, P. R. *J. Chem. Phys.* **2003**, *119*, 1922.
(342) Surján, P. R.; Rolik, Z.; Szabados, Á.; Kőhalmi, D. *Ann. Phys. (Leipzig)* **2004**, *13*, 223.
(343) Szabados, Á; Tóth, G.; Rolik, Z.; Surján., P. R. *J. Chem. Phys.* **2005**, *122*, 114104.
(344) Rolik, Z.; Szabados, Á. *Int. J. Quantum Chem.* **2009**, *109*, 2554.
(345) Mahapatra, U. S.; Datta, B.; Mukherjee, D. *J. Phys. Chem. A* **1999**, *103*, 1822.
(346) Ghosh, P.; Chattopadhyay, S.; Jana, D.; Mukherjee, D. *Int. J. Mol. Sci.* **2002**, *3*, 733.
(347) Koslowski, A.; Beck, M.; Thiel, W. *J. Comput. Chem.* **2003**, *24*, 714.
(348) Granucci, G.; Toniolo, A. *Chem. Phys. Lett.* **2000**, *325*, 79.
(349) Granucci, G.; Persico, M.; Toniolo, A. *J. Chem. Phys.* **2001**, *114*, 10608.
(350) Toniolo, A.; Ben-Nun, M.; Martinez, T. *J. Phys. Chem. A* **2002**, *106*, 4679.
(351) Lei, Y.; Suo, B.; Dou, Y.; Wang, Y.; Wen, Z. *J. Comput. Chem.* **2010**, *31*, 1752.
(352) Moshinsky, M.; Seligman, T. H. *Ann. Phys.* **1971**, *66*, 311.
(353) Malmqvist, P.-A. *Int. J. Quantum Chem.* **1986**, *30*, 479.
(354) Malmqvist, P.-A.; Roos, B. *Chem. Phys. Lett.* **1989**, *155*, 189.
(355) Olsen, J.; Godefroid, M. R.; Jonsson, P.; Malmqvist, P.-A.; Fischer, C. F. *Phys. Rev. E* **1995**, *52*, 4499.
(356) Lorentzon, J.; Fuelscher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 9265.
(357) Mitrushchenkov, A.; Werner, H.-J. *Mol. Phys.* **2007**, *105*, 1239.
(358) Olsen, J.; Yeager, D. L.; Jørgensen, P. *Adv. Chem. Phys.* **1983**, *54*, 1.
(359) Schmidt, M. W.; Gordon, M. S. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233.
(360) Shepard, R. *Theor. Chim. Acta* **1992**, *84*, 55.
(361) Shepard, R.; Shavitt, I.; Simons, J. *J. Chem. Phys.* **1982**, *76*, 543.
(362) Elbert, S. T. In *Numerical Algorithms in Chemistry: Algebraic Methods*; Moler, C.; Shavitt, I., Eds.; Lawrence Berkeley Laboratory: Berkeley, CA. 1978; pp 129−141.
(363) Chaban, G.; Schmidt, M. W.; Gordon, M. S. *Theor. Chem. Acc.* **1997**, *97*, 88.
(364) Ruedenberg, K.; Cheung, L. M.; Elbert, S. T. *Int. J. Quantum Chem.* **1979**, *16*, 1069.
(365) Barbatti, M.; Shepard, R.; Lischka, H. Computational and Methodological Elements for Nonadiabatic Trajectory Dynamics Simulations of Molecules. In *Conical Intersections: Theory, Computation and Experiment*; Domcke, W., Yarkony, D. R., Köppel, H., Eds.; Advanced Series in Physical Chemistry, Vol. 17; World Scientific: Singapore, in press.
(366) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert., S. T. *Chem. Phys.* **1982**, *71*, 41.
(367) Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364.
(368) Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9377.
(369) Shepard, R.; Kedziora, G. S.; Lischka, H.; Shavitt, I.; Müller, T.; Szalay, P. G.; Kállay, M.; Seth, M. *Chem. Phys.* **2008**, *349*, 37.
(370) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1988**, *89*, 2185.
(371) Moler, C.; Van Loan, C. F. *SIAM Rev.* **1978**, *20*, 801.
(372) Moler, C.; Van Loan, C. F. *SIAM Rev.* **2003**, *45*, 3.
(373) Higham, N. J. *SIAM Rev.* **2009**, *51*, 747.
(374) Higham, N. J. *SIAM J. Matrix Anal. Appl.* **2005**, *26*, 1179.
(375) Dalgaard, E.; Jørgensen, P. *J. Chem. Phys.* **1978**, *69*, 3833.
(376) Ward, R. C.; Gray, L. J. *ACM Trans. Math. Software* **1978**, *4*, 278.
(377) Netlib Repository at UTK and ORNL, http://www.netlib.org.
(378) Yanai, T; Chan, G. K.-L. *J. Chem. Phys.* **2006**, *124*, 194106.
(379) Yanai, T; Chan, G. K.-L. *J. Chem. Phys.* **2007**, *127*, 104107.
(380) Mazziotti, D. A. *Phys. Rev. Lett.* **2006**, *97*, 143002.
(381) Mazziotti, D. A. *J. Chem. Phys.* **2007**, *126*, 184101.
(382) Mazziotti, D. A. *Phys. Rev. A* **2007**, *75*, 022505.
(383) Mazziotti, D. A. *Phys. Rev. A* **2007**, *76*, 052502.
(384) Valdemoro, C.; Tel, L. M.; Pérez-Romero; Alcoba, D. R. *Int. qJ. Quantum Chem.* **2008**, *108*, 1090.
(385) Valdemoro, C.; Alcoba, D. R.; Tel, L. M.; Pérez-Romero *Int. J. Quantum Chem.* **2009**, *109*, 2622.
(386) Alcoba, D. R.; Valdemoro, C.; Tel, L. M.; Pérez-Romero *Int. J. Quantum Chem.* **2009**, *109*, 3178.
(387) Valdemoro, C.; Alcoba, D. R.; Tel, L. M.; Pérez-Romero *Int. J. Quantum Chem.* **2011**, *111*, 245.
(388) Alcoba, D. R.; Tel, L. M.; Pérez-Romero; Valdemoro, C. *Int. J. Quantum Chem.* **2011**, *111*, 937.
(389) Harriman, J. E. *Phys. Rev. A* **1979**, *19*, 1893.
(390) Kutzelnigg, W. *Chem. Phys. Lett.* **1979**, *64*, 383.
(391) Kutzelnigg, W. *Int. J. Quantum Chem.* **1980**, *18*, 3.
(392) Mukherjee, D.; Kutzelnigg, W. *J. Chem. Phys.* **2001**, *114*, 2047.
(393) Kutzelnigg, W.; Mukherjee, D. *J. Chem. Phys.* **2002**, *116*, 4787.
(394) Kutzelnigg, W.; Mukherjee, D. *J. Chem. Phys.* **2004**, *120*, 7340.
(395) Kutzelnigg, W.; Mukherjee, D. *J. Chem. Phys.* **2004**, *120*, 7350.
(396) Colmenero, F.; Valdemoro, C. *Phys. Rev. A* **1993**, *47*, 979.
(397) Nakatsuji, H.; Yasuda, K. *Phys. Rev. Lett.* **1996**, *76*, 1039.
(398) Mazziotti, D. A. *Phys. Rev. A* **1998**, *57*, 4219.
(399) Mazziotti, D. A. *Chem. Phys. Lett.* **1998**, *289*, 419.
(400) Mazziotti, D. A. *Phys. Rev. Lett.* **2011**, *106*, 083001.
(401) Sinitskiy, A.; Greenman, L.; Mazziotti, D. A. *J. Chem. Phys.* **2010**, *133*, 014104.

(402) Greenman, L.; Mazziotti, D. A. *J. Chem. Phys.* **2010**, *130*, 164110.
(403) Greenman, L.; Mazziotti, D. A. *J. Chem. Phys.* **2009**, *130*, 184101.
(404) Gidofalvi, G.; Mazziotti, D. A. *J. Chem. Phys.* **2008**, *129*, 134108.
(405) Mazziotti, D. A. *J. Phys. Chem. A* **2007**, *111*, 12635.
(406) Foley, J. J., IV; Mazziotti, D. A. *Mol. Phys.* **2010**, *108*, 2543.
(407) Mazziotti, D. A. *J. Phys. Chem. A* **2008**, *112*, 13684.
(408) Greenman, L.; Mazziotti, D. A. *J. Phys. Chem. A* **2010**, *114*, 583.
(409) Foley, J. J., IV; Rothman, A. E.; Mazziotti, D. A. *J. Chem. Phys.* **2009**, *130*, 22410.
(410) Gidofalvi, G.; Mazziotti, D. A. *Phys. Rev. A* **2009**, *80*, 022507.
(411) Rothman, A. E.; Foley, J. J., IV; Mazziotti, D. A. *Phys. Rev. A* **2009**, *80*, 052508.
(412) Snyder, J. W.; Rothman, A. E.; Foley, J. J., IV; Mazziotti, D. A. *J. Chem. Phys.* **2010**, *132*, 154109.
(413) Mukherjee, D. *Chem. Phys. Lett.* **1997**, *274*, 561.
(414) Kutzelnigg, W.; Mukherjee, D. *J. Chem. Phys.* **1997**, *107*, 432.
(415) Neuscamman, E; Yanai, T; Chan, G. K.-L. *Int. Rev. Phys. Chem.* **2010**, *29*, 231.
(416) Neuscamman, E; Yanai, T; Chan, G. K.-L. *J. Chem. Phys.* **2009**, *130*, 124102.
(417) Neuscamman, E; Yanai, T; Chan, G. K.-L. *J. Chem. Phys.* **2010**, *132*, 024106.
(418) Yanai, T; Kurashige, Y; Chan, G. K.-L. *J. Chem. Phys.* **2010**, *132*, 024105.
(419) Schwartz, C. *Phys. Rev.* **1962**, *126*, 1015.
(420) Kutzelnigg, W.; Morgan, J. D., III *J. Chem. Phys.* **1992**, *96*, 4484. *erratum* **1992**, *97*, 8821.
(421) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.
(422) Feller, D. *J. Chem. Phys.* **1993**, *98*, 7059.
(423) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.
(424) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.
(425) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437.
(426) Peterson, K. A.; Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *100*, 7410.
(427) Petersson, G. A.; Malick, D. K.; Frisch, M. J.; Braunstein, M. *J. Chem. Phys.* **2005**, *123*, 074111.
(428) Petersson, G. A.; Malick, D. K.; Frisch, M. J.; Braunstein, M. *J. Chem. Phys.* **2006**, *125*, 044107.
(429) Jiang, W.; Wilson, A. K. *J. Chem. Phys.* **2011**, *134*, 034101.
(430) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *126*, 144108.
(431) Schuurman, M. S.; Muir, S. R.; Allen, W. D.; Schaefer, H. F., III *J. Chem. Phys.* **2004**, *120*, 11586.
(432) Karton, A.; Rabinovich, E.; Martin, J. M. L. *J. Chem. Phys.* **2006**, *125*, 11586.
(433) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599.
(434) Bomble, Y. J.; Vázquez, K.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108.
(435) Olsen, J.; Christiansen, O.; Koch, H.; Jørgensen, P. *J. Chem. Phys.* **1996**, *105*, 5082.
(436) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley: New York, 2000.
(437) Hylleraas, E. A. *Z. Phys.* **1928**, *48*, 469.
(438) Kutzelnigg, W. *Theor. Chim. Acta* **1985**, *68*, 445.
(439) Klopper, W.; Kutzelnigg, W. *Chem. Phys. Lett.* **1987**, *134*, 17.
(440) Kutzelnigg, W.; Klopper, W. *J. Chem. Phys.* **1991**, *94*, 1985.
(441) Klopper, W.; Manby, F. R.; Ten-No, S.; Valeev, E. F. *Int. Rev. Phys. Chem.* **2006**, *25*, 427.
(442) Röhse, R.; Klopper, W.; Kutzelnigg, W. *J. Chem. Phys.* **1993**, *99*, 8830.
(443) Gdanitz, R. J. *Chem. Phys. Lett.* **1993**, *210*, 253.
(444) Gdanitz, R. J.; Röhse, R. *Int. J. Quantum Chem.* **1995**, *55*, 147. **1996**, *59*, 505. (Erratum)
(445) Gdanitz, R. J.; Black, G. D.; Lansing, C. S.; Palmer, B. J.; Schuchardt, K. L. *J. Comput. Chem.* **2005**, *26*, 214.
(446) Cardoen, W.; Gdanitz, R. J. *J. Chem. Phys.* **2005**, *123*, 024304.
(447) Flores, J. R.; Gdanitz, R. J. *J. Chem. Phys.* **2005**, *123*, 144316.
(448) Cardoen, W.; Simons, J.; Gdanitz, R. J. *Int. J. Quantum Chem.* **2006**, *106*, 1516.
(449) Cardoen, W.; Gdanitz, R. J.; Simons, J. *J. Phys. Chem. A* **2006**, *110*, 564.
(450) Ten-No, S. *Chem. Phys. Lett.* **2007**, *447*, 175.
(451) Varganov, S. A.; Martinez, T. J. *J. Chem. Phys.* **2010**, *132*, 054103.
(452) Torheyden, M.; Valeev, E. F. *J. Chem. Phys.* **2009**, *131*, 171103.
(453) Kong, L.; Valeev, E. F. *Studies of excited states of polyenes with multireference explicitly correlated methods*. Poster presented at the 241st ACS National Meeting, Anheim, CA, 2011.
(454) Shiozaki, T.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2011**, *134*, 034113.
(455) Shiozaki, T.; Werner, H.-J. *J. Chem. Phys.* **2011**, *134*, 184104.
(456) Kato, T. *Commun. Pure Appl. Math.* **1957**, *10*, 151.
(457) Pulay, P. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; pp 1191−1240.
(458) Shepard, R. *Int. J. Quantum Chem.* **1987**, *31*, 33.
(459) Shepard, R.; Lischka, H.; Szalay, P. G.; Kovar, T.; Ernzerhof, M. *J. Chem. Phys.* **1992**, *96*, 2085.
(460) Helgaker, T. U.; Almlöf, J. *Int. J. Quantum Chem.* **1984**, *26*, 275.
(461) Helgaker, T.; Taylor In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.;World Scientific: Singapore, 1995; pp 725−856.
(462) Schlegel, H. B. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; pp 459−500.
(463) Handy, N. C.; Schaefer, H. F., III *J. Chem. Phys.* **1984**, *81*, 5031.
(464) Rice, J. E.; Amos, R. D. *Chem. Phys. Lett.* **1985**, *122*, 585.
(465) Raffenetti, R. C. *J. Chem. Phys.* **1973**, *58*, 4452.
(466) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(467) Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1987**, *86*, 4070.
(468) Almlöf, J.; Taylor, P. R. *Adv. Quantum Chem.* **1992**, *22*, 301.
(469) Baer, M. In *Beyond Born-Oppenheimer: Electronic Nonadiabatic Coupling Terms and Conical Intersections*; Wiley: Hoboken, NJ, USA, 2006.
(470) Balasubramanian, K. *Relativistic Effects in Chemistry Part A Theory and Techniques*; Wiley: New York, 1997.
(471) Balasubramanian, K. *Relativistic Effects in Chemistry Part B Applications*; Wiley: New York, 1997.
(472) Schwerdtfeger, P., Ed. *Relativistic Electronic Structure Theory: Part 1. Theory*; Elsevier: Amsterdam, 2004.
(473) Pyykkö, P. *Angew. Chem., Int. Ed.* **2004**, *43*, 4412.
(474) Roos, B. O.; Malmqvist, P.-Å.; Gagliardi, L. *J. Am. Chem. Soc.* **2006**, *128*, 17000.
(475) Roos, B. O.; Malmqvist., P.-Å. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2919.
(476) Pyykkö, P.; Desclaux, J. P. *Acc. Chem. Res.* **1979**, *12*, 276.
(477) Pyykkö, P. *Chem. Rev.* **1988**, *88*, 563.
(478) Pyykkö, P. *Chem. Rev.* **1997**, *97*, 597.
(479) Schwerdtfeger, P., Ed. *Relativistic Electronic Structure Theory: Part 2. Applications*; Elsevier: Amsterdam, 2004.
(480) Pitzer, K. S. *Acc. Chem. Res.* **1979**, *12*, 271.
(481) Kutzelnigg, W. *Chem. Phys.* **1997**, *225*, 203.
(482) Stanton, R. E.; Havriliak, S. *J. Chem. Phys.* **1984**, *81*, 1910.
(483) Breit, G. *Phys. Rev.* **1929**, *34*, 553.
(484) Moss, R. E. *Advanced Molecular Quantum Mechanics*; Chapman and Hall: London, 1973.
(485) Strange, P. *Relativistic Quantum Mechanics: With Applications in Condensed Matter and Atomic Physics*; Cambridge University Press: Cambridge, U.K., 1998.

(486) Mark, F.; Lischka, H.; Rosicky, F. *Chem. Phys. Lett.* **1980**, *71*, 507.
(487) Visscher, L.; Aerts, P. J. C.; Visser, O.; Nieuwpoort, W. C. *Int. J. Quantum Chem.* **1991**, *S25*, 131.
(488) Talman, J. D. *Phys. Rev. Lett.* **1986**, *57*, 1091.
(489) Jensen, J. J. Aa.; Dyall, K. G.; Saue, T.; Faegri, K., Jr. *J. Chem. Phys.* **1996**, *104*, 4083.
(490) van Lenthe, E.; van Leeuwen, R.; Baerends, E. J.; Snijders, J. G. *Int. J. Quantum Chem.* **1996**, *57*, 281.
(491) Dyall, K. G. *J. Comput. Chem.* **2002**, *23*, 786.
(492) Wolf., A.; Reiher, M.; Hess, B. A. *J. Chem. Phys.* **2002**, *117*, 9215.
(493) Wolf, A.; Reiher, M.; Hess, B. A. In *Relativistic Quantum Chemistry—Theory*. Schwerdtfeger, P., Ed; Elsevier: New York, 2002.
(494) Samzow, R.; Hess, B. A. *Chem. Phys. Lett.* **1991**, *184*, 491.
(495) Hess, B. A.; Marian, C. M.; Wahlgren, U.; Gropen, O. *Chem. Phys. Lett.* **1996**, *251*, 365.
(496) Christiansen, O.; Gauss, J.; Schimmelpfennig, B. *Phys. Chem. Chem. Phys.* **2000**, *2*, 965.
(497) Huzinaga, S.; Seijo, L.; Barandiaran, Z.; Klobukowski, M. *J. Chem. Phys.* **1987**, *86*, 2132.
(498) Rakowitz, F.; Marian, C. M.; Wahlgren, U.; Sejo, L. *J. Chem. Phys.* **1999**, *110*, 3678.
(499) Rakowitz, F.; Marian, C. M.; Schimmelpfennig, B. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2481.
(500) Schimmelpfennig, B.; Maron, L.; Wahlgren, U.; Teichteil, C.; Fagerli, H.; Gropen, O. *Chem. Phys. Lett.* **1998**, *286*, 267.
(501) Vallet, V.; Schimmelpfennig, B.; Maron, L.; Teichteil, C.; Leininger, T.; Gropen, O.; Grenthe, I.; Wahlgren, U. *Chem. Phys.* **1999**, *244*, 185.
(502) Klobukowski, M.; Huzinaga, S.; Sakai, Y. In *Computational Chemistry, Review of Current Trends*; Leszczynski, J., Ed.; World Scientific: Singapore, 1999; Vol. 3, p 49.
(503) Zeng, T.; Mori, H.; Miyoshi, E.; Klobukowski, M. *Int. J. Quantum Chem.* **2009**, *109*, 3235.
(504) Zeng, T.; Fedorov, D. G.; Klobukowski, M. *J. Chem. Phys.* **2010**, *132*, 074102.
(505) Fedorov, D. G.; Gordon, M. S. *J. Chem. Phys.* **2000**, *112*, 5611.
(506) Fedorov, D. G.; Finley, J. P. *Phys. Rev. A* **2001**, *64*, 042502.
(507) Zeng, T.; Fedorov, D. G.; Klobukowski, M. *J. Chem. Phys.* **2010**, *133*, 114107.
(508) Ermler, W. C.; Ross, R. B; Christiansen, P. A. *Adv. Quantum Chem.* **1988**, *19*, 139.
(509) Christiansen, P. A.; Lee, Y. S.; Pitzer, K. S. *J. Chem. Phys.* **1979**, *71*, 4445.
(510) Pacios, L. F.; Christiansen, P. A. *J. Chem. Phys.* **1985**, *82*, 2664.
(511) Hurley, M. M.; Pacios, L. F.; Christiansen, P. A.; Ross, R. B.; Ermler, W. C. *J. Chem. Phys.* **1986**, *84*, 2812.
(512) LaJohn, L. A.; Christiansen, P. A.; Ross, R. B.; Atashroo, T.; Ermler, W. C. *J. Chem. Phys.* **1987**, *87*, 2812.
(513) Ross, R. B.; Power, J. M.; Atashroo, T.; Ermler, W. C.; LaJohn, L. A.; Christiansen, P. A. *J. Chem. Phys.* **1990**, *93*, 6654.
(514) Ermler, W. C.; Ross, R. B.; Christiansen, P. A. *Int. J. Quantum Chem.* **1991**, *40*, 829.
(515) Wallace, N. M.; Blaudeau, J.-P.; Pitzer, R. M. *Int. J. Quantum Chem.* **1991**, *40*, 789.
(516) Ross, R. B.; Gayen, S.; Ermler, W. C. *J. Chem. Phys.* **1994**, *100*, 8145.
(517) Nash, C. S.; Bursten, B. E.; Ermler, W. C. *J. Chem. Phys.* **1997**, *106*, 5133.
(518) Wildman, S. A.; DiLabio, G. A.; Christiansen, P. A. *J. Chem. Phys.* **1997**, *107*, 9975.
(519) Blaudeau, J.-P.; Curtiss, L. A. *Int. J. Quantum Chem.* **1997**, *61*, 943.
(520) Hay, P. J. *J. Chem. Phys.* **1983**, *79*, 5469.
(521) Hay, P. J; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270.
(522) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284.
(523) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.
(524) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1998**, *109*, 3875.
(525) Stevens, W. J.; Basch, H.; Krauss, M. *J. Chem. Phys.* **1984**, *81*, 6026.
(526) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612.
(527) Cundari, T. R.; Stevens, W. J. *J. Chem. Phys.* **1993**, *98*, 5555.
(528) Küchle, W.; Dolg, M.; Stoll, H.; Preuss, H. *Mol. Phys.* **1991**, *74*, 1245.
(529) Dolg, M.; Stoll, H.; Preuss, H.; Pitzer, R. M. *J. Phys. Chem.* **1993**, *97*, 5852.
(530) Dolg, M. *Relativistic Effective Core Potential in Relativistic Electronic Structure Theory, Part 1: Fundamentals*, Theoretical and Computational Chemistry, Vol. 11; Elsevier Science B.V: Amsterdam, 2002.
(531) Dolg, M.; Stoll, H.; Savin, A.; Preuss, H. *Theor. Chim. Acta* **1989**, *75*, 173.
(532) Dolg., M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1993**, *85*, 441.
(533) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86*, 866.
(534) Schwerdtfeger, P.; Dolg, M.; Schwarz, W. H. E.; Bowmaker, G. A.; Boyd, P. D. W. *J. Chem. Phys.* **1989**, *91*, 1762.
(535) Dolg, M.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1989**, *90*, 1730.
(536) Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; Preuß, H. *Theor. Chim. Acta* **1990**, *77*, 123.
(537) Häussermann, U.; Dolg, M.; Stoll, H.; Preuss, H.; Schwerdtfeger, P.; Pitzer, R. M. *Mol. Phys.* **1993**, *78*, 1211.
(538) Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuss, H. *Mol. Phys.* **1993**, *80*, 1431.
(539) Küchle, W.; Dolg, M.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1994**, *100*, 7535.
(540) Metz, B.; Schweizer, M.; Stoll, H.; Dolg, M.; Liu, W. *Theor. Chem. Acc.* **2000**, *104*, 22.
(541) Seth, M.; Schwerdtfeger, P.; Dolg, M. *J. Chem. Phys.* **1997**, *106*, 3623.
(542) Metz, B.; Stoll, H.; Dolg, M. J. *Chem. Phys.* **2000**, *113*, 2563.
(543) Cao, X.; Dolg, M. *J. Chem. Phys.* **2001**, *115*, 7348.
(544) Stoll, H.; Metz, B.; Dolg, M. *J. Comput. Chem.* **2002**, *23*, 767.
(545) Cao, X.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2003**, *118*, 487.
(546) Cao, X.; Dolg, M. *J. Mol. Struct. (THEOCHEM)* **2002**, *581*, 139.
(547) Tinkham, M. *Group Theory and Quantum Mechanics*; McGraw-Hill: New York, 1964.
(548) Heine, V. *Group Theory in Quantum Mechanics*; Pergamon Press: New York, 1960.
(549) Messiah, A. *Quanten Mechanik 2*; Walter de Gruyter: Berlin, 1985. Messiah, A. *Quantum Mechanics*; Dover: Minneola, 1999.
(550) Yanai, T.; Harrison, R. J.; Nakajima, T.; Ishikawa, Y.; Hirao, K. *Int. J. Quantum Chem.* **2007**, *107*, 1382.
(551) Visscher, L. *Chem. Phys. Lett.* **1996**, *253*, 20.
(552) Saue, T.; Jensen, H. J. A. *J. Chem. Phys.* **1999**, *111*, 6211.
(553) Sjovoll, M.; Gropen, O.; Olsen, J. *Theor. Chim. Acc.* **1997**, *97*, 301.
(554) Desclaux, J. P. *At. Data Nucl. Data Tables* **1973**, *12*, 311.
(555) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. *J. Chem. Phys.* **2003**, *119*, 11113.
(556) Pitzer, R. M.; Winter, N. W. *Int. J. Quantum Chem.* **1991**, *40*, 773.
(557) Yabushita, S.; Zhang, Z.; Pitzer, R. M. *J. Phys. Chem.* **1999**, *103*, 5791.
(558) Wilkensen, J. H. *The Algebraic Eigenvalue Problem*; Clarendon Press: Oxford, U.K., 1965.
(559) Pitzer, R. M.; Winter, N. W. *J. Phys. Chem.* **1988**, *92*, 3061.
(560) Hess, B. A.; Buenker, R. J.; Marian, C. M.; Peyerimhoff, S. D. *Chem. Phys. Lett.* **1982**, *89*, 459.
(561) Tilson, J. L.; Ermler, W. C.; Pitzer, R. M. *Comput. Phys. Commun.* **2000**, *128*, 128.
(562) Christiansen, P. A.; Balasubramanian, K.; Pitzer, K. S. *J. Chem. Phys.* **1982**, *76*, 5087.

(563) DiLabio, G. A.; Christiansen, P. A. *Chem. Phys. Lett.* **1997**, *277*, 473.
(564) DiLabio, G. A.; Christiansen, P. A. *J. Chem. Phys.* **1998**, *108*, 7527.
(565) Kleinschmidt, M.; Tatchen, J.; Marian, C. M. *J. Chem. Phys.* **2006**, *124*, 124101.
(566) Hess, B. A.; Marian, C. M.; Peyerimhoff, S. D. Ab Initio Calculation of Spin-Orbit Effects in Molecules Including Electron Correlation. In *Modern Electronic Structure Theory*, Vol. I; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995.
(567) Malmqvist, P.-Å.; Roos, B. O.; Schimmelpfennig, B. *Chem. Phys. Lett.* **2002**, *357*, 230.
(568) Teichteil, C.; Spiegelmann, F. *Chem. Phys.* **1983**, *81*, 283.
(569) Rakowitz, F.; Marian, C. M. *Chem. Phys.* **1997**, *225*, 233.
(570) Buenker, R. J.; Alekseyev, A. B.; Liebermann, H.-P.; Lingott, R.; Hirsch, G. *J. Chem. Phys.* **1998**, *108*, 3400.
(571) Vallet, V.; Maron, L.; Teichteil, C.; Flament, J.-P. *J. Chem. Phys.* **2000**, *113*, 1391.
(572) Visscher, L.; Visser, O.; Aerts, P. J. C.; Merenga, H.; Nieuwpoort, W. C. *Comput. Phys. Commun.* **1994**, *81*, 120.
(573) Saue, T.; Visscher, L.; Jensen, H. J. Aa., with new contributions from: Bast, R.; Dyall, K. G.; Ekstrm, U.; Eliav, E.; Enevoldsen, T.; Fleig, T.; Gomes, A. S. P.; Henrikson, J.; Ilias, M.; Jacob, Ch. R., Knecht, S.; Nataraj, H. S.; Norman, P.; Olsen, J.; Pernpointner, M.; Ruud, K.; Schimmelpfennig, B.; Sikkema, J.; Thorvaldsen, A.; Thyssen, J.; Villaume, S.; Yamamoto, S. *DIRAC*, a relativistic ab-initio electronic structure program, release DIRAC10, 2010, http://dirac.chem.vu.nl.
(574) Fleig, T.; Olsen, J.; Marian, C. M. *J. Chem. Phys.* **2001**, *114*, 4775.
(575) Fleig, T.; Olsen, J.; Visscher, L. *J. Chem. Phys.* **2003**, *119*, 2963.
(576) Fleig, T.; Jensen, H. J. Aa.; Olsen, J.; Visscher, L. *J. Chem. Phys.* **2006**, *124*, 104106.
(577) Fleig, T.; Marian, C. M.; Olsen, J. *Theor. Chem. Acc.* **1997**, *97*, 125.
(578) Knecht, S.; Jensen, H. J. Aa.; Fleig, T. *J. Chem. Phys.* **2008**, *129*, 034109.
(579) Kim, Y. S.; Lee, Y. S. *J. Chem. Phys.* **2003**, *119*, 12169.
(580) Harrison, R. J.; Shepard, R. *Annu. Rev. Phys. Chem.* **1994**, *45*, 623.
(581) Kendall, R. A.; Harrison, R. J.; Littlefield, R. J.; Guest, M. F. In *Reviews in Computational Chemistry*, Vol. 6; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley: Hoboken, NJ, USA, 2007; pp 209−316.
(582) de Jong, W. A.; Bylaska, E.; Govind, N.; Janssen, C. L.; Kowalski, K.; Müller, T.; Nielsen, I. M. B.; van Dam, H. J. J.; Veryazov, V.; Lind, R. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6896.
(583) *NVIDIA CUDA Compute Unified Device Architecture Programming Guide*, Version 3.2, developer.download.nvidia.com/compute/cuda/3_2_prod/toolkit/docs/CUDA_C_Programming_Guide.pdf.
(584) *OpenCL. Open Computing Language*, www.khronos.org/opencl.
(585) *Advance e710 product brief*, www.clearspeed.com/products/documents/-e710_product_brief.pdf. *CSXL User Guide*, www.clearspeed.com/resources/documentation/-CSXL_User_Guide_3.1_Rev1.C.pdf.
(586) Drepper, U. "What every programmer should know about memory," 2007, http://people.redhat.com/drepper/cpumemory.pdf.
(587) *Parallel Virtual File System*, http://www.pvfs.org.
(588) *General Parallel File System*, http://www-03.ibm.com/systems/software/gpfs.
(589) http://www.lustre.org.
(590) http://www2.fz-juelich.de/jsc/sionlib.
(591) *OpenMP, specification: OpenMP Application Program Interface*, Version 3.0 www.openmp.org
(592) *High Performance Fortran*, http://hpff.rice.edu.
(593) Harrison, R. J. *Int. J. Quantum Chem.* **1991**, *40*, 847.
(594) Sunderam, V. S. *Concurr. Comput. Pract. Exper.* **1990**, *2*, 315.
(595) *MPI: A Message-Passing Interface Standard. Message Passing Interface Forum*, www.mpi-forum.org. Gropp, W.; Lusk, E.; Thakur, R. *Using MPI-2—Advanced Features of the Message-Passing Interface*; MIT Press: New Haven, CT, USA, 1999.
(596) Nieplocha, J.; Palmer, B.; Tipparaju, V.; Krishnan, M.; Trease, H.; Apra, E. *Int. J. High Perform. Comput. Appl.* **2006**, *20*, 203.
(597) Fedorov, D. M.; Olsen, R. M.; Kitaura, K.; Gordon, M. S.; Koseki, S. *J. Comput. Chem.* **2004**, *25*, 872.
(598) Wang, M.; May, A. J.; Knowles, P. *J. Comput. Phys. Commun.* **2009**, *180*, 2673.
(599) *ScaLAPACK, Scalable Linear Algebra Package Users Guide*, www.netlib.org/scalapack/slug.
(600) Lewis, B.; Berg, D. J. *Multithreaded Programming with pthreads*, Prentice-Hall: Upper Saddle River, NJ, USA, 1998.
(601) Ufimisev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 1004.
(602) Ufimisev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 2619.
(603) Wilkinson, K. A.; Sherwood, P.; Guest, M. F.; Naidoo, K. J *J. Comput. Chem.* **2011**, *32*, 2313.
(604) Brown, P.; Woods, C.; McIntosh-Smith, S.; Manby, F. R. *J. Chem. Theor. Comput.* **2008**, *4*, 1620.
(605) Yasuda, K. *J. Chem. Theory Comput.* **2008**, *4*, 1230.
(606) Yoshimine, M. *J. Comput. Phys.* **1973**, *11*, 449.
(607) Schütz, M.; Lindh, R. *Theor. Chem. Acc.* **1997**, *95*, 13.
(608) Schütz, M.; Lindh, R.; Werner, H.-J. *Mol. Phys.* **1999**, *96*, 719.
(609) Taylor, P. *Int. J. Quantum Chem.* **1987**, *31*, 521.
(610) Vahtras, O.; Almlöf, J. E.; Feyereisen, M. W. *Chem. Phys. Lett.* **1996**, *261*, 105.
(611) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143.
(612) Weigend, F.; Kattannek, M.; Ahlrichs, R. *J. Chem. Phys.* **2009**, *130*, 164106.
(613) Aquilante, F.; Lindh, R.; Pedersen, T. B. *J. Chem. Phys.* **2007**, *127*, 114107.
(614) Aquilante, F.; Gagliardi, L.; Pedersen, T. B.; Lindh, R. *J. Chem. Phys.* **2009**, *130*, 154107.
(615) Aquilante, F.; Lindh, R.; Pedersen, T. B. *J. Chem. Phys.* **2008**, *129*, 034106.
(616) Hättig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154.
(617) Hättig, C.; Hellweg, A.; Köhn, A. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1159.
(618) Aquilante, F.; Pedersen, T. B.; Lindh, R.; Roos, B. O.; Sanchez de Meras, A.; Koch, H. *J. Chem. Phys.* **2008**, *129*, 024113.
(619) Aquilante, F.; Malmqvist, P.-A.; Pedersen, T. B.; Ghosh, A.; Roos, B. O. *J. Chem. Theory Comput.* **2008**, *4*, 694.
(620) Koch, H.; Sanchez de Meras, A.; Pedersen, T. B. *J. Chem. Phys.* **2003**, *118*, 9481.
(621) Pierloot, K.; Vancoillie, S. *J. Chem. Phys.* **2008**, *128*, 034104.
(622) Roeggen, I.; Johansen, T. *J. Chem. Phys.* **2008**, *128*, 194107.
(623) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. *Comput. Phys. Commun.* **2010**, *181*, 1477.
(624) Malmqvist, P.-Å.; Rendell, A.; Roos, B. O. *J. Phys. Chem.* **1990**, *94*, 5477.
(625) Zarrabian, S.; Sarma, C. R.; Paldus, J. *Chem. Phys. Lett.* **1989**, *155*, 183.
(626) Klene, M.; Robb, M. A.; Frisch, M. J.; Celani, P. *J. Chem. Phys.* **2000**, *113*, 5654.
(627) Fletcher, G. D. *Mol. Phys.* **2007**, *105*, 2971.
(628) Lischka, H.; Dallos, M.; Shepard, R. *Mol. Phys.* **2002**, *100*, 1647.
(629) Dachsel, H.; Lischka, H.; Shepard, R.; Nieplocha, J.; Harrison, R. J. *J. Comput. Chem.* **1995**, *18*, 430.
(630) Duch, W.; Karwoski, J. *Int. J. Quantum Chem.* **1982**, *22*, 783.
(631) Ansaloni, R.; Bendazzoli, G. L.; Evangelisti, S.; Rossi, E. *Comput. Phys. Commun.* **2000**, *128*, 496.
(632) Ben-Amor, N.; Evangelisti, S.; Maynau, D.; Rossi, E. *Chem. Phys. Lett.* **1998**, *288*, 348.
(633) Gan, Z.; Alexeev, Y.; Gordon, M. S.; Kendall, R. A. *J. Chem. Phys.* **2003**, *119*, 47.

(634) Tanaka, K.; Mochizuki, Y.; Ishikawa, T.; Terashima, H.; Tokiwa, H. *Theor. Chem. Acc.* **2007**, *117*, 397.
(635) Knecht, S.; Jensen, H. J. Aa.; Fleig, T. *J. Chem. Phys.* **2010**, *132*, 014108.
(636) Gan, Z.; Harrison, R. J. *J. Proc. 2005 ACM/IEEE SC05 Conf.* **2005**, *22*; DOI: //10.1109/SC.2005.17.
(637) Gan, Z.; Grant, D. G.; Harrison, R. J.; Dixon, D. A. *J. Chem. Phys.* **2006**, *125*, 124311.
(638) Müller, T. *J. Phys. Chem. A* **2009**, *113*, 12729.
(639) Celani, P.; Stoll, H.; Werner, H.-J.; Knowles, P. J. *Mol. Phys.* **2004**, *102*, 2369.
(640) Suo, B.; Zhai, G.; Wang, Y.; Wen, Z.; Hu, X.; Li, L. *J. Comput. Chem.* **2005**, *26*, 88.
(641) Kleinschmidt, M.; Marian, C. M.; Waletzke, M.; Grimme, S. *J. Chem. Phys.* **2009**, *130*, 044708.
(642) Grimme, S.; Waletzke, M. *J. Chem. Phys.* **1999**, *111*, 5645.
(643) Bernhardsson, A.; Lindh, R.; Olsen, J.; Fülscher, M. *Mol. Phys.* **1999**, *96*, 617.
(644) Dudley, T. J.; Olson, R. M.; Schmidt, M. W.; Gordon, M. S. *J. Comput. Chem.* **2006**, *27*, 352.
(645) Siegbahn, P. M. E.; Blomberg, M. R. A.; Bauschlicher, C. W. *J. Chem. Phys.* **1984**, *81*, 1373.
(646) Pettersson, L. G. M.; Bauschlicher, C. W.; Halicioglu, T. *J. Chem. Phys.* **1987**, *87*, 2205.
(647) Bauschlicher, C. W.; Langhoff, S. R.; Taylor, P. R. *Astrophys. J.* **1988**, *332*, 531.
(648) Stark, K.; Werner, H.-J. *J. Chem. Phys.* **1996**, *104*, 6515.
(649) Alexander, M. H.; Werner, H.-J.; Dagdigian, P. J. *J. Chem. Phys.* **1988**, *89*, 1388.
(650) Holka, F.; Szalay, P. G.; Fremont, J.; Rey, M; Peterson, K. A.; Tyuterev, Vl. G. *J. Chem. Phys.* **2011**, *134*, 094306.
(651) Cave, R. J. *J. Chem. Phys.* **1990**, *92*, 2450.
(652) Lappe, J.; Cave, R. J. *J. Phys. Chem. A* **2000**, *104*, 2294.
(653) Mochizuki, Y.; Tanaka, K. *Theor. Chem. Acc.* **1998**, *99*, 88.
(654) Sakai, Y.; Mogi, K.; Miyoshi, E. *J. Chem. Phys.* **1999**, *111*, 3989.
(655) Müller, T.; Dallos, M.; Lischka, H.; Dubrovay, Zs; Szalay, P. G. *Theor. Chem. Acc.* **2001**, *105*, 227.
(656) Gdanitz, R. J. *Chem. Phys. Lett.* **2001**, *348*, 67.
(657) Gdanitz, R. J. *Chem. Phys. Lett.* **1999**, *312*, 578.
(658) Martin, J. M. L. *Chem. Phys. Lett.* **1999**, *303*, 399.
(659) van de Bovenkamp, J.; van Duijneveldt, F. B. *J. Chem. Phys.* **1999**, *110*, 11141.
(660) Müller, T.; Xantheas, S. S.; Dachsel, H.; Harrison, R. J.; Nieplocha, J.; Shepard, R.; Kedziora, G. S.; Lischka, H. *Chem. Phys. Lett.* **1998**, *293*, 72.
(661) Schinke, R.; Grebenshchikov, S. Y.; Ivanov, M. V.; Fleurat-Lessard, P. *Annu. Rev. Phys. Chem.* **2006**, *57*, 625.
(662) Siebert, R.; Fleurat-Lessard, P.; Schinke, R.; Bittererova, M.; Farantos, S. C. *J. Chem. Phys.* **2002**, *116*, 9749.
(663) Holka, F.; Szalay, P. G.; Müller, T; Tyuterev, Vl. G. *J. Phys. Chem. A* **2010**, *114*, 9927.
(664) Barletta, P.; Shirin, S. V.; Zobov, N. F.; Polyansky, O. L.; Tennyson, J.; Valeev, E. F.; Császár, A. G. *J. Chem. Phys.* **2006**, *125*, 204307.
(665) Jansen, G *Chem. Phys. Lett.* **1994**, *223*, 377.
(666) Schröder, D; Harvey, J. N.; Aschi, M; Schwarz, H *J. Chem. Phys.* **1998**, *108*, 8446.
(667) Chaban, G. M.; Lundell, J.; Gerber, R. B. *Chem. Phys. Lett.* **2002**, *364*, 628.
(668) McDowell, S. A. C. *Curr. Org. Chem.* **2006**, *10*, 791.
(669) Leininger, T; Galdea, X *J. Phys. B: At., Mol. Opt. Phys.* **2000**, *33*, 735.
(670) Cizek, M; Horacek, J.; Sergenton, A. C.; Popovic, D. B.; Allan, M.; Domcke, W.; Leininger, T.; Gadea, F. X. *Phys. Rev. A* **2001**, *63*, 062710.
(671) Ismail, N.; Heully, J. L.; Saue, T.; Daudey, J.-P.; Marsden, C. J. *Chem. Phys. Lett.* **1999**, *300*, 296.
(672) Jakubetz, W.; Lan, B. L. *Chem. Phys.* **1997**, *217*, 375.
(673) Cremer, D.; Kraka, E.; Szalay, P. G. *Chem. Phys. Lett.* **1998**, *292*, 97.
(674) Pasinszki, T.; Hajgató, B; Havasi, B.; Westwood, N. P. C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 5263.
(675) Kurtz, M.; Strunk, J.; Hinrichsen, O.; Muhler, M.; Fink, K.; Meyer, B.; Wöll, C. *Angew. Chem., Int Ed.* **2005**, *44*, 2790.
(676) Cossart-Magos, C.; Jungen, M.; Xu, R.; Launay, F. *J. Chem. Phys.* **2003**, *119*, 3219.
(677) Partridge, H.; Schwenke, D. W. *J. Chem. Phys.* **1997**, *106*, 4618.
(678) Manolopoulos, D. E.; Stark, K.; Werner, H. J.; Arnold, D. W.; Bradforth, S,E.; Neumark, D. M. *Science* **1993**, *262*, 1852.
(679) Skouteris, D.; Manolopoulos, D. E.; Bian, W. S.; Werner, H. J.; Lai, L. H.; Liu, K. P. *Science* **1999**, *286*, 713.
(680) Egli Esposti, A.; Werner, H. J. *J. Chem. Phys.* **1990**, *93*, 3351.
(681) Werner, H. J.; Follmeg, B.; Alexander, M. H. *J. Chem. Phys.* **1988**, *89*, 3139.
(682) Kaupp, M.; Schleyer, P. v. R.; Stoll, H.; Preuss, H. *J. Am. Chem. Soc.* **1991**, *113*, 6012.
(683) Stoll, H. *Phys. Rev. B* **1992**, *46*, 6700.
(684) Ruscic, B.; Feller, D.; Dixon, D. A.; Peterson, K. A.; Harding, L. B.; Asher, R. L.; Wagner, A. F. *J. Phys. Chem. A* **2001**, *105*, 1.
(685) Xantheas, S. S.; Dunning, T. H. *J. Phys. Chem.* **1993**, *97*, 18.
(686) Peterson, K. A.; Werner, H. J. *J. Chem. Phys.* **1992**, *96*, 8948.
(687) Ben-Nun, M.; Martinez, T. J. *Adv. Chem. Phys.* **2002**, *121*, 439.
(688) Wierschke, S. G.; Nash., J. J.; Squires, R. R. *J. Am. Chem. Soc.* **1993**, *115*, 1958.
(689) Hirsch, G.; Bruna, P. J.; Buenker, R. J.; Peyerimhoff, S. D. *Chem. Phys.* **1980**, *45*, 335.
(690) Hogreve, H. *Chem. Phys. Lett.* **1993**, *215*, 72.
(691) Hogreve, H. *J. Chem. Phys.* **1995**, *102*, 3281.
(692) Hogreve, H. *J. Phys. B: At., Mol. Opt. Phys.* **1995**, *28*, L263.
(693) Hogreve, H. *J. Phys. B: At., Mol. Opt. Phys.* **1998**, *31*, L439.
(694) Schreiner, P. R.; Navarro-Vazquez, A.; Prall, M. *Acc. Chem. Res.* **2005**, *38*, 29.
(695) Engels, B.; Hanrath, M. *J. Am. Chem. Soc.* **1998**, *120*, 6356.
(696) Schmittel, M.; Steffen, J. P.; Angel, M. A. W.; Engels, B.; Lennartz, C.; Hanrath, M. *Angew. Chem., Int. Ed.* **1998**, *37*, 1562.
(697) Peterson, K. A.; Puzzarini, C. *Theor. Chem. Acc.* **2005**, *114*, 283.
(698) Langhoff, S. R.; Bauschlicher, C. W., Jr. *Annu. Rev. Phys. Chem.* **1988**, *39*, 181.
(699) Bauschlicher, C. W., Jr.; Langhoff, S. R.; Patridge, H. the Application of Ab Initio Electronic Structure Calculations to Molecules Containing Transition Metal Atoms. In *Modern Electronic Structure Theory*, Vol. 2; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; pp 1280−1374.
(700) Matsika, S.; Pitzer, R. M. *J. Phys. Chem. A* **2000**, *104*, 4064.
(701) Matsika, S.; Zhang, Z.; Brozell, S. R.; Blaudeau, J.-P.; Wang, Q.; Pitzer, R. M. *J. Phys. Chem. A* **2001**, *105*, 3825.
(702) Zhang, Z.; Pitzer, R. M. *J. Phys. Chem. A* **1999**, *103*, 6880.
(703) Matsika, S.; Pitzer, R. M. *J. Phys. Chem. A* **2000**, *104*, 11983.
(704) Christiansen, P. A. *Chem. Phys. Lett.* **1984**, *109*, 145.
(705) Christiansen, P. A. *J. Chem. Phys.* **1983**, *79*, 2928.
(706) Balasubramanian, K.; Pitzer, K. S. *J. Chem. Phys.* **1983**, *78*, 321.
(707) Sioutis, I.; Stakhursky, V. L.; Pitzer, R. M.; Miller, T. A. *J. Chem. Phys.* **2007**, *126*, 124308.
(708) Teichteil, C.; Pelissier, M. *Chem. Phys.* **1994**, *180*, 1.
(709) Vallet, V.; Schimmelpfennig, B; Leininger, T.; Teichteil, C.; Gropen, O.; Grenthe, I.; Wahlgren, U. *J. Phys. Chem. A* **1999**, *103*, 9285.
(710) Ramirez-Solis, A.; Vallet, V.; Teichteil, C.; Leininger, T.; Daudey, J. P. *J. Chem. Phys.* **2001**, *115*, 3201.
(711) Hess, B. A.; Marian, C. M.; Relativistic Effects in the Calculation of Electronic Energies. In *Computational Molecular Spectroscopy*; Jensen, P.; Bunker, P. R., Eds.; Wiley: New York, 2000.
(712) Kleinschmidt, M.; Marian, C. M. *Chem. Phys. Lett.* **2008**, *458*, 190.
(713) de Jong, W. A.; Nieuwpoort, W. C. *Int. J. Quantum Chem.* **1996**, *58*, 203.

(714) Visscher, L.; Saue, T.; Nieuwpoort, W. C.; Faegri, K.; Gropen, O. *J. Chem. Phys.* **1993**, *99*, 6704.
(715) Visscher, L.; Styszynski, J.; Nieuwpoort, W. C. *J. Chem. Phys.* **1996**, *105*, 1987.
(716) Sorensen, L. K.; Knecht, S.; Fleig, T.; Marian, C. M. *J. Phys. Chem. A* **2009**, *113*, 12607.
(717) Gomes, A. S. P.; Visscher, L.; Bolvin, H.; Saue, T.; Knecht, S.; Fleig, T.; Eliav, E. *J. Chem. Phys.* **2010**, *133*, 064305.
(718) Chang, Q. M.S. Thesis, Graduate School, The Ohio State University, Columbus, OH, USA, 2002.
(719) Gagliardi, L.; Heaven, M. C.; Krogh, J. W.; Roos, B. O. *J. Am. Chem. Soc.* **2005**, *127*, 86.
(720) Infante, I.; Eliav, E.; Vilkas, M. J.; Ishikawa, Y.; Kaldor, U.; Visscher, L. *J. Chem. Phys.* **2007**, *127*, 124308.
(721) Guichemerre, M.; Chambaud, G.; Stoll, H. *Chem. Phys.* **2002**, *280*, 71.
(722) Dolg, M. Lanthanides and Actinides. In *Encyclopedia of Computational Chemistry*, Vol. 2; Schleyer, P. v. R., et al., Eds.; Wiley: New York, 1998; pp 1478−1486.
(723) Dolg, M.; Stoll, H.; Preuss, H. *J. Mol. Struct.* **1991**, *77*, 243.
(724) Dolg, M.; Stoll, H.; Flad, H.-J.; Preuss, H. *J. Chem. Phys.* **1992**, *97*, 1162.
(725) Fromager, E.; Vallet, V.; Schimmelpfennig, B.; Macak, P.; Privalov, T.; Wahlgren, U. *J. Phys. Chem. A* **2005**, *109*, 4957.
(726) Seijo, L.; Barandiaran, Z. Relativistic Electronic Structure Theory, Part 2: Applications. In *Theoretical and Computational Chemistry*, Vol. 14; Schwerdtfeger, P., Ed.; Elsevier: Amsterdam, 2004.
(727) Pascual, J. L.; Barros, N.; Barandiaran, Z.; Seijo, L *J. Phys. Chem. A* **2009**, *113*, 12454.
(728) Matsika, S.; Pitzer, R. M. *J. Phys. Chem. A* **2001**, *105*, 637.
(729) Sun, L.; Hase, W. L. *Rev. Comput. Chem.* **2003**, *19*, 79.
(730) Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061.
(731) Chen, W.; Hase, W. L.; Schlegel, H. B. *Chem. Phys. Lett.* **1994**, *228*, 436.
(732) Davidson, E. R. *J. Phys. Chem.* **1996**, *100*, 6161.
(733) Krebs, S.; Buenker, R. J. *J. Chem. Phys.* **1997**, *106*, 7208.
(734) Borden, W. T.; Davidson, E. R. *Acc. Chem. Res.* **1996**, *29*, 67.
(735) Muller, T.; Dallos, M.; Lischka, H. *J. Chem. Phys.* **1999**, *110*, 7176.
(736) Angeli, C. *J. Comput. Chem.* **2009**, *30*, 1319.
(737) Szalay, P. G.; Karpfen, A.; Lischka, H. *Chem. Phys.* **1989**, *130*, 219.
(738) Dallos, M.; Lischka, H. *Theor. Chem. Acc.* **2004**, *112*, 16.
(739) Serrano-Andres, L.; Merchan, M.; Nebot-Gil, I.; Lindh, R.; Roos, B. O. *J. Chem. Phys.* **1993**, *98*, 3151.
(740) Watts, J. D.; Gwaltney, S. R.; Bartlett, R. J. *J. Chem. Phys.* **1996**, *105*, 6979.
(741) Palmer, M. H.; Walker, I. C. *Chem. Phys.* **2010**, *373*, 159.
(742) Lehtonen, O.; Sundholm, D.; Send, R.; Johansson, M. P. *J. Chem. Phys.* **2009**, *131*, 024301.
(743) Krawczyk, R. P.; Malsch, K.; Hohlneicher, G.; Gillen, R. C.; Domcke, W. *Chem. Phys. Lett.* **2000**, *320*, 535.
(744) Jones, R. R.; Bergman, R. G. *J. Am. Chem. Soc.* **1972**, *94*, 660.
(745) Crawford, T. D.; Kraka, E.; Stanton, J. F.; Cremer, D. *J. Chem. Phys.* **2001**, *114*, 10638.
(746) Slipchenko, L. V.; Krylov, A. I. *J. Chem. Phys.* **2002**, *117*, 4694.
(747) Wang, E. B.; Parish, C. A.; Lischka, H. *J. Chem. Phys.* **2008**, *129*, 044306.
(748) Cizek, J.; Paldus, J. *J. Chem. Phys.* **1967**, *47*, 3976.
(749) Paldus, J.; Veillard, A. *Mol. Phys.* **1978**, *35*, 445.
(750) Takada, T.; Dupuis, M. *J. Am. Chem. Soc.* **1983**, *105*, 1713.
(751) Szalay, P. G.; Császar, A. G.; Fogarasi, G.; Karpfen, A.; Lischka, H. *J. Chem. Phys.* **1990**, *93*, 1246.
(752) Whitman, D. W.; Carpenter, B. K. *J. Am. Chem. Soc.* **1982**, *104*, 6473.
(753) Balkova, A.; Bartlett, R. J. *J. Chem. Phys.* **1994**, *101*, 8972.
(754) Eckert-Maksić, M.; Vazdar, M.; Barbatti, M.; Lischka, H.; Maksić, Z. B. *J. Chem. Phys.* **2006**, *125*, 064310.
(755) Borden, W. T.; Davidson, E. R.; Hart, P. *J. Am. Chem. Soc.* **1978**, *100*, 388.
(756) Sumita, M.; Saito, K. *Chem. Phys.* **2010**, *371*, 30.
(757) Houk, K. N.; Gonzalez, J.; Li, Y. *Acc. Chem. Res.* **1995**, *28*, 81.
(758) Isobe, H.; Takano, Y.; Kitagawa, Y.; Kawakami, T.; Yamanaka, S.; Yamaguchi, K.; Houk, K. N. *Mol. Phys.* **2002**, *100*, 717.
(759) Lischka, H.; Ventura, E.; Dallos, M. *Chem. Phys. Chem.* **2004**, *5*, 1365.
(760) Ventura, E.; do Monte, S. A.; Dallos, M.; Lischka, H. *J. Phys. Chem. A* **2003**, *107*, 1175.
(761) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(762) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
(763) Huang, C. H.; Tsai, L. C.; Hu, W. P. *J. Phys. Chem. A* **2001**, *105*, 9945.
(764) Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029.
(765) Buenker, R. J.; Bonačíć-Koutecký, V.; Pogliani, L. *J. Chem. Phys.* **1980**, *73*, 1836.
(766) Petrongolo, C.; Buenker, R. J.; Peyerimhoff, S. D. *J. Chem. Phys.* **1982**, *76*, 3655.
(767) Persico, M.; Bonačíć-Koutecký, V. *J. Chem. Phys.* **1982**, *76*, 6018.
(768) Merer, A. J.; Mulliken, R. S. *Chem. Rev.* **1969**, *69*, 639.
(769) Ohmine, I. *J. Chem. Phys.* **1985**, *83*, 2348.
(770) Freund, L.; Klessinger, M. *Int. J. Quantum Chem.* **1998**, *70*, 1023.
(771) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J. Phys. Chem. A* **2000**, *104*, 5161.
(772) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(773) Barbatti, M.; Paier, J.; Lischka, H. *J. Chem. Phys.* **2004**, *121*, 11614.
(774) Bobrowicz, F. W.; Goddard, W. A., III. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977; p 79−127.
(775) Michl, J.; Bonačíć-Koutecký, V. *Electronic Aspects of Organic Photochemistry*; Wiley-Interscience: New York, 1990.
(776) Pitonak, M.; Lischka, H. *Mol. Phys.* **2005**, *103*, 855.
(777) Wald, G. *Science* **1968**, *162*, 230.
(778) Bonačíć-Koutecký, V.; Schoffel, K.; Michl, J. *Theor. Chim. Acta* **1987**, *72*, 459.
(779) Saunders, V. R.; Guest, M. F. *Comput. Phys. Commun.* **1982**, *26*, 389.
(780) Zilberg, Sh.; Haas, Y. *Photochem. Photobiol. Sci.* **2003**, *2*, 1256.
(781) Du, P.; Racine, S. C.; Davidson, E. R. *J. Phys. Chem.* **1990**, *94*, 3944.
(782) Barbatti, M.; Aquino, A. J. A.; Lischka, H. *Mol. Phys.* **2006**, *104*, 1053.
(783) Barbatti, M.; Aquino, A. J. A.; Lischka, H. *J. Phys. Chem. A* **2005**, *109*, 5168.
(784) Schreiber, M.; Barbatti, M.; Zilberg, S.; Lischka, H.; Gonzalez, L. *J. Phys. Chem. A* **2007**, *111*, 238.
(785) Hachey, M. R. J.; Bruna, P. J.; Grein, F. *J. Phys. Chem.* **1995**, *99*, 8050.
(786) Müller, Th.; Lischka, H. *Theor. Chem. Acc.* **2001**, *106*, 369.
(787) Gómez-Carrasco, S.; Müller, T.; Köppel, H. *J. Phys. Chem. A* **2010**, *114*, 11436.
(788) Merchán, M.; Roos, B. O. *Theor. Chim. Acta* **1995**, *92*, 227.
(789) Dallos, M.; Müller, Th.; Lischka, H. *J. Chem. Phys.* **2001**, *114*, 746.
(790) Moule, D. C.; Walsh, A. D. *Chem. Rev.* **1975**, *75*, 67.
(791) Green, W. H.; Moore, C. B.; Polik, W. F. *Annu. Rev. Phys. Chem.* **1992**, *43*, 591.
(792) Hopkins, W. S.; Loock, H.-P.; Cronin, B.; Nix, M. G. D.; Devine, A. L.; Dixon, R. N.; Ashfold, M. N. R.; Yin, H.-M.; Rowling, St. J.; Büll, A.; Kable, S. H. *J. Phys. Chem. A* **2008**, *112*, 9283.
(793) Townsend, D.; Lahankar, S. A.; Lee, S. K.; Chambreau, S. D.; Suits, A. G.; Zhang, X.; Rheinecker, J.; Harding, L. B.; Bowman, J. M. *Science* **2004**, *306*, 1158.

(794) Bowman, J. M.; Shepler, B. C. *Annu. Rev. Phys. Chem.* **2011**, *62*, 53.
(795) Simonsen, J. B.; Rusteika, N.; Johnson, M. S.; Sølling, T. I. *Phys. Chem. Chem. Phys.* **2008**, *10*, 674.
(796) Zhang, P.; Maeda, S.; Morokuma, K.; Braams, B. J. *J. Chem. Phys.* **2009**, *130*, 114304.
(797) Shepler, B. C.; Epifanovsky, E.; Zhang, P.; Bowman, J. M.; Krylov, A. I.; Morokuma, K. *J. Phys. Chem. A Lett.* **2008**, *112*, 13267.
(798) Araujo, M.; Lasorne, B.; Bearpark, M. J.; Magalhaes, A. L.; Worth, G. A.; Bearpark, M. J.; Robb, M. A. *J. Chem. Phys.* **2009**, *131*, 144301.
(799) Marian, C. M. *J. Chem. Phys.* **2005**, *122*, 104314.
(800) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.
(801) Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584.
(802) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, Luis; Rodríguez, E.; Reguero, M.; Peláez, D. *J. Phys. Chem. A* **2006**, *110*, 26471.
(803) Barbatti, M.; Lischka, H. *J. Am. Chem. Soc.* **2008**, *130*, 6831.
(804) Barbatti, M.; Aquino, A. J. A.; Lischka, H. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4959.
(805) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354.
(806) Kistler, K. A.; Matsika, S. *J. Phys. Chem. A* **2007**, *111*, 2650.
(807) Kistler, K. A.; Matsika, S. *J. Chem. Phys.* **2008**, *128*, 215102.
(808) Zaloudek, F.; Novros, J. S.; Clark, L. B. *J. Am. Chem. Soc.* **1985**, *107*, 7344.
(809) Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallová, D.; Lischka, H. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6145.
(810) Harding, L. B.; Klippenstein, S. J.; Jasper, A. W. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4055.
(811) Harding, L. B.; Georgievskii, Y.; Klippenstein, S. J. *J. Phys. Chem. A* **2005**, *109*, 4646.
(812) Klippenstein, S. J.; Georgievskii, Y.; Harding, L. B. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1133.
(813) Harding, L. B.; Klippenstein, S. J.; Georgievskii, Y. *J. Phys. Chem. A* **2007**, *111*, 3789.
(814) Klippenstein, S. J.; Harding, L. B. *J. Phys. Chem. A* **1999**, *103*, 9388.
(815) Klippenstein, S. J.; Harding, L. B. *Phys. Chem. Chem. Phys.* **1999**, *1*, 989.
(816) Harding, L. B.; Klippenstein, S. J. *Symp. (Int.) Combust.* **1998**, *27*, 151.
(817) Harding, L. B.; Georgievskii, Y.; Klippenstein, S. J. *J. Phys. Chem. A* **2010**, *114*, 765.
(818) Harding, L. B.; Klippenstein, S. J. *J. Phys. Chem. Lett.* **2010**, *1*, 3016.
(819) Sivaramakrishnan, R.; Michael, J. V.; Wagner, A. F.; Dawes, R.; Jasper, A. W.; Harding, L. B.; Georgievskii, Y.; Klippenstein, S. *J. Combust. Flame* **2011**, *158*, 618.
(820) Sivaramakrishnan, R.; Su, M.-C.; Michael, J. V.; Harding, L. B.; Klippenstein, S. J. *J. Phys. Chem. A*, in press.
(821) Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *J. Chem. Phys.* **1997**, *106*, 6430.
(822) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. *J. Chem. Phys.* **2001**, *114*, 6548.
(823) Ben-Nun, M.; Martínez, T. J. *Chem. Phys. Lett.* **1998**, *298*, 57.
(824) Viel, A.; Krawczyk, R. P.; Manthe, U.; Domcke, W. *J. Chem. Phys.* **2004**, *120*, 11000.
(825) Farmanara, P.; Stert, V.; Radloff, W. *Chem. Phys. Lett.* **1998**, *288*, 518.
(826) Kosma, K.; Trushin, S. A.; Fuss, W.; Schmid, W. E. *J. Phys. Chem. A* **2008**, *112*, 7514.
(827) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975**, *97*, 1285.
(828) Fabiano, E.; Keal, T. W.; Thiel, W. *Chem. Phys.* **2008**, *349*, 334.
(829) Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495.
(830) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(831) Barbatti, M.; Granucci, G.; Persico, M.; Lischka, H. *Chem. Phys. Lett.* **2005**, *401*, 276.
(832) Barbatti, M.; Ruckenbauer, M.; Lischka, H. *J. Chem. Phys.* **2005**, *122*, 174307.
(833) Stert, V.; Lippert, H.; Ritze, H. H.; Radloff, W. *Chem. Phys. Lett.* **2004**, *388*, 144.
(834) Celani, P.; Werner, H.-J. *J. Chem. Phys.* **2003**, *119*, 5044.
(835) Tao, H. L.; Levine, B. G.; Martinez, T. J. *J. Phys. Chem. A* **2009**, *113*, 13656.
(836) Barbatti, M.; Ruckenbauer, M.; Szymczak, J. J.; Aquino, A. J. A.; Lischka, H. *Phys. Chem. Chem. Phys.* **2008**, *10*, 482.
(837) Zechmann, G.; Barbatti, M.; Lischka, H.; Pittner, J.; Bonačić-Koutecký, V. *Chem. Phys. Lett.* **2006**, *418*, 377.
(838) Hostettler, J. M.; Bach, A.; Chen, P. *J. Chem. Phys.* **2009**, *130*, 034303.
(839) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2796.
(840) Canuel, C.; Mons, M.; Piuzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 074316.
(841) Szymczak, J. J.; Barbatti, M.; Hoo, J. T. S.; Adkins, J. A.; Windus, T. L.; Nachtigallova, D.; Lischka, H. *J. Phys. Chem. A* **2009**, *113*, 12686.
(842) Asturiol, D.; Lasorne, B.; Robb, M. A.; Blancafort, L. *J. Phys. Chem. A* **2009**, *113*, 10211.
(843) Gonzaléz-Vázquez, J.; Gonzaléz, L. *Chem. Phys. Chem.* **2010**, *11*, 3617.
(844) Barbatti, M.; Szymczak, J. J.; Aquino, A. J. A.; Nachtigallova, D.; Lischka, H. *J. Chem. Phys.* **2011**, *134*, 014304.
(845) Nachtigallová, D.; Aquino, A. J. A.; Szymczak, J. J.; Barbatti, M.; Hobza, P.; Lischka, H. *J. Phys. Chem. A* **2011**, *115*, 5247.
(846) Hudock, H. R.; Levine, B. G.; Thompson, A. L.; Satzger, H.; Townsend, D.; Gador, N.; Ullrich, S.; Stolow, A.; Martinez, T. J. *J. Phys. Chem. A* **2007**, *111*, 8500.
(847) Hudock, H. R.; Martinez, T. J. *Chem. Phys. Chem.* **2008**, *9*, 2486.
(848) Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallová, D.; Hobza, P.; Lischka, H. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 21453.
(849) Barbatti, M.; Lischka, H. *J. Phys. Chem. A* **2007**, *111*, 2852.
(850) Nachtigallova, D.; Barbatti, M.; Szymczak, J. J.; Hobza, P.; Lischka, H. *Chem. Phys. Lett.* **2010**, *497*, 129.
(851) Groenhof, G.; Schafer, L. V.; Boggio-Pasqua, M.; Goette, M.; Grubmuller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 6812.
(852) Ruckenbauer, M.; Barbatti, M.; Sellner, B.; Muller, T.; Lischka, H. *J. Phys. Chem. A* **2010**, *114*, 12585.
(853) Lindh, R.; Roos, B. O. *Int. J. Quantum Chem.* **1989**, *35*, 813.
(854) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 96502.
(855) Pulay, P.; Fogarasi, G.; Pongor, G.; Boggs, J. E.; Vargha, A. *J. Am. Chem. Soc.* **1983**, *105*, 7037.
(856) Wiberg, K. B.; Hadad, C. M.; Foresman, J.; Chupka, W. A. *J. Phys. Chem.* **1992**, *96*, 10756.
(857) McMurchie, L. E.; Davidson, E. R. *J. Chem. Phys.* **1977**, *66*, 2959.
(858) Zaitsevskii, A.; Malrieu, J.-P. *Int. J. Quantum Chem.* **1995**, *55*, 117.
(859) Finley, J.; Malmqvist, P.-A.; Roos, B. O.; Serrano-Andres, L. *Chem. Phys. Lett.* **1998**, *288*, 299.
(860) Lyakh, D. I.; Musial, M.; Lotrich, V. F.; Bartlett, R. J. *Chem. Rev.* **2012**, *112*; DOI: 10.1021/cr2001417.