Theor Chem Acc (2006) 116: 2–17
DOI 10.1007/s00214-005-0029-5

REGULAR ARTICLE

So Hirata

Symbolic algebra in quantum chemistry

Received: 15 February 2005 / Accepted: 16 September 2005 / Published online: 6 January 2006

Abstract Complex symbolic algebra, such as the manipulation of second-quantized operators, Slater determinants, Feynman diagrams, is inevitable in quantum chemistry. Increasingly, these operations are performed by the computerized systems that can handle higher mathematical constructs than just numbers and simple arithmetic. This article reviews these new algorithms that automate the algebraic transformation and computer implementation of many-body quantum-mechanical methods for electron correlation. They enable a whole new class of highly complex but vastly accurate methods, the manual development of which is no longer practical.

Keywords Symbolic algebra · Automated derivation and implementation · Determinant- and string-based methods · Electron correlation

1 Introduction

Computational quantum chemistry [1–7] systematically translates a given problem of chemistry to mathematical equations that are in turn subject to numerical solution by computers. In principle, there is little room for empiricism, heuristics, intuition, or any other form of nonscientific skills to enter these processes. This is particularly true in the subfield of computational quantum chemistry for electronic structure that is based on ab initio molecular orbital theory. It embodies hierarchies of methods of increasing accuracy and complexity that eventually converge, in the rigorous mathematical sense, at the exact solution of the Schrödinger equation for a given expansion basis set for wave functions. Chemical properties and transformation can be simulated or predicted, with unprecedented fidelity, by carrying through the application of these hierarchical methods to a high enough order with a large enough expansion basis set (e.g., Refs. [8–12]). Apart from the feasibility of such calculations, there is hardly any question about the faultlessness of this systematic, autonomous procedure [13]. In this sense, computational quantum chemistry may be one of the most immaculate fields of science.

It is the construction of these hierarchical methods rather than their applications that has set the pace of advancement of the field. The development of configuration-interaction, coupled-cluster, many-body perturbation theories, etc. involves *derivation*, *transformation*, and *implementation*, all of which are daunting tasks of symbolic manipulations. The derivation refers to the process of expressing the methods defined in the language of second quantization into mathematical equations (the sums of tensor products). These sum-of-product tensor expressions are transformed into a form that exposes an efficient computational sequence. Finally, this computational sequence is translated into computer programs. While these tedious processes by themselves may be inessential to chemistry or physics, they obviously dictate how complex a method we can develop and use and how quickly such progresses can be made in practice.

In the past, these processes are usually carried out by hand. However, this manual processing will soon cease to be a practical means for the symbolic manipulation in higher-order electron-correlation methods owing to its sheer volume and complexity. For instance, the coupled-cluster (CC) methods [14–22] containing through connected double, triple, or quadruple excitation operator (abbreviated as CCSD [23,24], CCSDT [25,26], and CCSDTQ [27,28]), which are defined by only three, four, and five second-quantized expressions, are expanded into the mathematical equations composed of 48, 102, and 183 antisymmetrized groups of tensor products, respectively, each of which in turn can contain up to 144 individual products. Figure 1 illustrates the complexity of one of the five tensor equations defining CCSDTQ. Each product, say, ABCDE (each letter representing a tensor) can be evaluated in a number of ways, (((AB)C)D)E, ((BC)(DE))A, ((CD)A)(BE), etc., and furthermore a sum-of-product expression, e.g., ABCD+ABEF, is subject to factorization: A(B((CD)+(EF)), B(C(AD)+A(EF)), etc. Whereas

S. Hirata
Quantum Theory Project, Department of Chemistry,
University of Florida, Gainesville, FL 32611-8435, USA
E-mail: hirata@qtp.ufl.edu

$$
\begin{aligned}
0 = {} & -P_{24}t^{p_5p_6p_7}_{h_9h_3h_4}v^{h_9p_8}_{h_1h_2} + P_{24}t^{p_9p_5p_6}_{h_2h_3h_4}v^{p_7p_8}_{h_1p_9} - P_4f^{h_9}_{h_1}t^{p_5p_6p_7p_8}_{h_9h_2h_3h_4} \\
& -P_4f^{p_8}_{p_9}t^{p_9p_5p_6p_7}_{h_1h_2h_3h_4} + \frac{1}{2}P_6t^{p_5p_6p_7p_8}_{h_9h_{10}h_3h_4}v^{h_9h_{10}}_{h_1h_2} + P_{16}t^{p_9p_5p_6p_7}_{h_{10}h_2h_3h_4}v^{h_{10}p_8}_{h_1p_9} \\
& +\frac{1}{2}P_6t^{p_9p_{10}p_5p_6}_{h_1h_2h_3h_4}v^{p_7p_8}_{p_9p_{10}} + P_{24}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_8}_{h_{10}}v^{h_9h_{10}}_{h_1h_2} - P_{48}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_{10}}_{h_2}v^{h_9p_8}_{h_1p_{10}} \\
& -P_{48}t^{p_9p_5p_6}_{h_2h_3h_4}t^{p_7}_{h_{10}}v^{h_{10}p_8}_{h_1p_9} - P_{24}t^{p_9p_5p_6}_{h_2h_3h_4}t^{p_{10}}_{h_1}v^{p_7p_8}_{p_9p_{10}} - P_4f^{h_9}_{p_{10}}t^{p_5p_6p_7p_8}_{h_9h_2h_3h_4}t^{p_{10}}_{h_1} \\
& +P_4f^{h_9}_{p_{10}}t^{p_{10}p_5p_6p_7}_{h_1h_2h_3h_4}t^{p_8}_{h_9} + \frac{1}{2}P_{12}t^{p_5p_6p_7p_8}_{h_9h_{10}h_3h_4}t^{p_{11}}_{h_2}v^{h_9h_{10}}_{h_1p_{11}} - P_{16}t^{p_9p_5p_6p_7}_{h_{10}h_2h_3h_4}t^{p_8}_{h_{11}}v^{h_{10}h_{11}}_{h_1p_9} \\
& -P_4t^{p_5p_6p_7p_8}_{h_9h_2h_3h_4}t^{p_{10}}_{h_{11}}v^{h_9h_{11}}_{h_1p_{10}} - P_{16}t^{p_9p_5p_6p_7}_{h_{10}h_2h_3h_4}t^{p_{11}}_{h_1}v^{h_{10}p_8}_{p_9p_{11}} - \frac{1}{2}P_{12}t^{p_9p_{10}p_5p_6}_{h_1h_2h_3h_4}t^{p_7}_{h_{11}}v^{h_{11}p_8}_{p_9p_{10}} \\
& +P_4t^{p_9p_5p_6p_7}_{h_1h_2h_3h_4}t^{p_{10}}_{h_{11}}v^{h_{11}p_8}_{p_9p_{10}} - P_{36}t^{p_7p_8}_{h_9h_3}t^{p_5p_6}_{h_{10}h_4}v^{h_9h_{10}}_{h_1h_2} + P_{144}t^{p_9p_7}_{h_2h_3}t^{p_5p_6}_{h_{10}h_4}v^{h_{10}p_8}_{h_1p_9} \\
& +P_{36}t^{p_9p_6}_{h_1h_2}t^{p_{10}p_5}_{h_3h_4}v^{p_7p_8}_{p_9p_{10}} - P_{24}f^{h_9}_{p_{10}}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_{10}p_8}_{h_1h_2} + P_{24}f^{h_9}_{p_{10}}t^{p_{10}p_5p_6}_{h_2h_3h_4}t^{p_7p_8}_{h_9h_1} \\
& -\frac{1}{2}P_{48}t^{p_5p_6p_7}_{h_9h_{10}h_4}t^{p_{11}p_8}_{h_2h_3}v^{h_9h_{10}}_{h_1p_{11}} + P_{72}t^{p_9p_5p_6}_{h_{10}h_3h_4}t^{p_7p_8}_{h_{11}h_2}v^{h_{10}h_{11}}_{h_1p_9} - P_{48}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_{10}p_8}_{h_{11}h_2}v^{h_9h_{11}}_{h_1p_{10}} \\
& +\frac{1}{2}P_{24}t^{p_9p_5p_6}_{h_2h_3h_4}t^{p_7p_8}_{h_{10}h_{11}}v^{h_{10}h_{11}}_{h_1p_9} - P_{72}t^{p_9p_5p_6}_{h_{10}h_3h_4}t^{p_{11}p_7}_{h_1h_2}v^{h_{10}p_8}_{p_9p_{11}} - \frac{1}{2}P_{24}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_{10}p_{11}}_{h_1h_2}v^{h_9p_8}_{p_{10}p_{11}} \\
& -\frac{1}{2}P_{48}t^{p_9p_{10}p_5}_{h_2h_3h_4}t^{p_6p_7}_{h_{11}h_1}v^{h_{11}p_8}_{p_9p_{10}} - P_{48}t^{p_9p_5p_6}_{h_2h_3h_4}t^{p_{10}p_7}_{h_{11}h_1}v^{h_{11}p_8}_{p_9p_{10}} - \frac{1}{2}P_{24}t^{p_9p_5p_6p_7}_{h_{10}h_{11}h_3h_4}t^{p_{12}p_8}_{h_1h_2}v^{h_{10}h_{11}}_{p_9p_{12}} \\
& +\frac{1}{4}P_6t^{p_5p_6p_7p_8}_{h_9h_{10}h_3h_4}t^{p_{11}p_{12}}_{h_1h_2}v^{h_9h_{10}}_{p_{11}p_{12}} + \frac{1}{2}P_{24}t^{p_9p_{10}p_5p_6}_{h_{11}h_2h_3h_4}t^{p_7p_8}_{h_{12}h_1}v^{h_{11}h_{12}}_{p_9p_{10}} - P_{16}t^{p_9p_5p_6p_7}_{h_{10}h_2h_3h_4}t^{p_{11}p_8}_{h_{12}h_1}v^{h_{10}h_{12}}_{p_9p_{11}} \\
& +\frac{1}{2}P_4t^{p_5p_6p_7p_8}_{h_9h_2h_3h_4}t^{p_{10}p_{11}}_{h_{12}h_1}v^{h_9h_{12}}_{p_{10}p_{11}} + \frac{1}{4}P_6t^{p_9p_{10}p_5p_6}_{h_1h_2h_3h_4}t^{p_7p_8}_{h_{11}h_{12}}v^{h_{11}h_{12}}_{p_9p_{10}} - \frac{1}{2}P_4t^{p_9p_5p_6p_7}_{h_1h_2h_3h_4}t^{p_{10}p_8}_{h_{11}h_{12}}v^{h_{11}h_{12}}_{p_9p_{10}} \\
& +\frac{1}{2}P_{24}t^{p_9p_7p_8}_{h_1h_2h_3}t^{p_{10}p_5p_6}_{h_{11}h_{12}h_4}v^{h_{11}h_{12}}_{p_9p_{10}} - \frac{1}{4}P_{16}t^{p_9p_{10}p_8}_{h_1h_2h_3}t^{p_5p_6p_7}_{h_{11}h_{12}h_4}v^{h_{11}h_{12}}_{p_9p_{10}} - \frac{1}{2}P_{24}t^{p_6p_7p_8}_{h_9h_1h_2}t^{p_{10}p_{11}p_5}_{h_{12}h_3h_4}v^{h_9h_{12}}_{p_{10}p_{11}} \\
& +P_{18}t^{p_9p_7p_8}_{h_{10}h_1h_2}t^{p_{11}p_5p_6}_{h_{12}h_3h_4}v^{h_{10}h_{12}}_{p_9p_{11}} - P_{48}t^{p_9}_{h_2}t^{p_8}_{h_{10}}t^{p_5p_6p_7}_{h_{11}h_3h_4}v^{h_{10}h_{11}}_{h_1p_9} + P_{24}t^{p_7}_{h_9}t^{p_8}_{h_{10}}t^{p_{11}p_5p_6}_{h_2h_3h_4}v^{h_9h_{10}}_{h_1p_{11}} \\
& -P_{24}t^{p_9}_{h_1}t^{p_{10}}_{h_2}t^{p_5p_6p_7}_{h_{11}h_3h_4}v^{h_{11}p_8}_{p_9p_{10}} - P_{48}t^{p_9}_{h_1}t^{p_7}_{h_{10}}t^{p_{11}p_5p_6}_{h_2h_3h_4}v^{h_{10}p_8}_{p_9p_{11}} + \frac{1}{2}P_6t^{p_9}_{h_1}t^{p_{10}}_{h_2}t^{p_5p_6p_7p_8}_{h_{11}h_{12}h_3h_4}v^{h_{11}h_{12}}_{p_9p_{10}} \\
& +P_{16}t^{p_9}_{h_1}t^{p_8}_{h_{10}}t^{p_{11}p_5p_6p_7}_{h_{12}h_2h_3h_4}v^{h_{10}h_{12}}_{p_9p_{11}} + P_4t^{p_9}_{h_1}t^{p_{10}}_{h_{11}}t^{p_5p_6p_7p_8}_{h_{12}h_2h_3h_4}v^{h_{11}h_{12}}_{p_9p_{10}} + \frac{1}{2}P_6t^{p_7}_{h_9}t^{p_8}_{h_{10}}t^{p_{11}p_{12}p_5p_6}_{h_1h_2h_3h_4}v^{h_9h_{10}}_{p_{11}p_{12}} \\
& -P_4t^{p_8}_{h_9}t^{p_{10}}_{h_{11}}t^{p_{12}p_5p_6p_7}_{h_1h_2h_3h_4}v^{h_9h_{11}}_{p_{10}p_{12}} - P_{72}t^{p_7p_8}_{h_9h_3}t^{p_5p_6}_{h_{10}h_4}t^{p_{11}}_{h_2}v^{h_9h_{10}}_{h_1p_{11}} - P_{144}t^{p_9p_7}_{h_2h_3}t^{p_5p_6}_{h_{10}h_4}t^{p_8}_{h_{11}}v^{h_{10}h_{11}}_{h_1p_9} \\
& -P_{144}t^{p_9p_7}_{h_2h_3}t^{p_5p_6}_{h_{10}h_4}t^{p_{11}}_{h_1}v^{h_{10}p_8}_{p_9p_{11}} - P_{72}t^{p_9p_6}_{h_1h_2}t^{p_{10}p_5}_{h_3h_4}t^{p_7}_{h_{11}}v^{h_{11}p_8}_{p_9p_{10}} + \frac{1}{2}P_{48}t^{p_5p_6p_7}_{h_9h_{10}h_4}t^{p_{11}p_8}_{h_2h_3}t^{p_{12}}_{h_1}v^{h_9h_{10}}_{p_{11}p_{12}} \\
& -P_{72}t^{p_9p_5p_6}_{h_{10}h_3h_4}t^{p_7p_8}_{h_{11}h_2}t^{p_{12}}_{h_1}v^{h_{10}h_{11}}_{p_9p_{12}} + P_{48}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_{10}p_8}_{h_{11}h_2}t^{p_{12}}_{h_1}v^{h_9h_{11}}_{p_{10}p_{12}} - \frac{1}{2}P_{24}t^{p_9p_5p_6}_{h_2h_3h_4}t^{p_7p_8}_{h_{10}h_{11}}t^{p_{12}}_{h_1}v^{h_{10}h_{11}}_{p_9p_{12}} \\
& +P_{72}t^{p_9p_5p_6}_{h_{10}h_3h_4}t^{p_{11}p_7}_{h_1h_2}t^{p_8}_{h_{12}}v^{h_{10}h_{12}}_{p_9p_{11}} + \frac{1}{2}P_{24}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_{10}p_{11}}_{h_1h_2}t^{p_8}_{h_{12}}v^{h_9h_{12}}_{p_{10}p_{11}} - P_{24}t^{p_5p_6p_7}_{h_9h_3h_4}t^{p_{10}p_8}_{h_1h_2}t^{p_{11}}_{h_{12}}v^{h_9h_{12}}_{p_{10}p_{11}} \\
& +\frac{1}{2}P_{48}t^{p_9p_{10}p_5}_{h_2h_3h_4}t^{p_6p_7}_{h_{11}h_1}t^{p_8}_{h_{12}}v^{h_{11}h_{12}}_{p_9p_{10}} + P_{48}t^{p_9p_5p_6}_{h_2h_3h_4}t^{p_{10}p_7}_{h_{11}h_1}t^{p_8}_{h_{12}}v^{h_{11}h_{12}}_{p_9p_{10}} + P_{24}t^{p_9p_5p_6}_{h_2h_3h_4}t^{p_7p_8}_{h_{10}h_1}t^{p_{11}}_{h_{12}}v^{h_{10}h_{12}}_{p_9p_{11}} \\
& +\frac{1}{2}P_{36}t^{p_9p_8}_{h_1h_2}t^{p_{10}p_7}_{h_3h_4}t^{p_5p_6}_{h_{11}h_{12}}v^{h_{11}h_{12}}_{p_9p_{10}} - P_{144}t^{p_9p_8}_{h_1h_2}t^{p_6p_7}_{h_{10}h_3}t^{p_{11}p_5}_{h_{12}h_4}v^{h_{10}h_{12}}_{p_9p_{11}} - \frac{1}{2}P_{36}t^{p_7p_8}_{h_9h_3}t^{p_5p_6}_{h_{10}h_4}t^{p_{11}p_{12}}_{h_1h_2}v^{h_9h_{10}}_{p_{11}p_{12}} \\
& -P_{24}t^{p_9}_{h_1}t^{p_{10}}_{h_2}t^{p_8}_{h_{11}}t^{p_5p_6p_7}_{h_{12}h_3h_4}v^{h_{11}h_{12}}_{p_9p_{10}} + P_{24}t^{p_9}_{h_1}t^{p_7}_{h_{10}}t^{p_8}_{h_{11}}t^{p_{12}p_5p_6}_{h_2h_3h_4}v^{h_{10}h_{11}}_{p_9p_{12}} - P_{36}t^{p_7p_8}_{h_9h_3}t^{p_5p_6}_{h_{10}h_4}t^{p_{11}}_{h_1}t^{p_{12}}_{h_2}v^{h_9h_{10}}_{p_{11}p_{12}} \\
& +P_{144}t^{p_9p_7}_{h_2h_3}t^{p_5p_6}_{h_{10}h_4}t^{p_{11}}_{h_1}t^{p_8}_{h_{12}}v^{h_{10}h_{12}}_{p_9p_{11}} + P_{36}t^{p_9p_6}_{h_1h_2}t^{p_{10}p_5}_{h_3h_4}t^{p_7}_{h_{11}}t^{p_8}_{h_{12}}v^{h_{11}h_{12}}_{p_9p_{10}}
\end{aligned}
$$

Fig. 1 The T_4 amplitude equation of the CCSDTQ method the result of the contraction does not depend on the order of evaluation owing to the commutative and distributive nature of tensor contraction and addition, the operation cost does depend strongly on it. Finding the most expedient order by a blanket search is out of the question without a computer and is difficult even with a computer. Once a reasonable computational sequence is found, its implementation must also honor any additional computational advantages that are brought to by symmetries of the wave function and characteristic length scales of interactions. Not a single error will be tolerated throughout the entire procedure.

Clearly, these symbolic algebraic processes must be computerized. Apart from expediting the time-consuming and error-prone derivation and implementation processes, the computerization will also facilitate the parallelization and other laborious optimization of the computer codes and enhance their maintainability, portability, and extensibility. However, most importantly, it enables a whole new class of many-body electron-correlation methods with unprecedented complexity and accuracy, the manual development of which is no longer practical or even possible.

What makes such computerization of quantum chemistry possible and effective is the systematic nature of the rules and strategies that govern the whole symbolic manipulation processes. If the rules varied considerably from one case to another, computerizing would be even more burdensome than straightforward manual processing. However, all of the *derivation*, *transformation*, and *implementation* processes are highly systematic. The vast majority of many-body electron-correlation methods are defined by the expectation values of a combination of second-quantized creation and annihilation operators for some Slater determinants. The evaluation of the expectation values can be performed by mechanical applications of either Slater's rules or the commutator rules of second quantization or by normal ordering followed by the application of Wick's theorem [29–32]. Alternatively, Goldstone or Hugenholtz diagrams may be used [31–36]. The resulting mathematical expressions are, as mentioned before, invariably the sum-of-product tensor expressions, which are again subject to computerized transformation and implementation. Systematic strategies for the latter include Kucharski and Bartlett's computational linearization and recursive factorization [37] that together suggest near-optimum computational sequences for the coupled-cluster equations, the direct product decomposition proposed by Stanton et al. [38] for point-group symmetry usage, and the index permutation symmetry logic by Kállay and Surján [39]. These systematic strategies exalt the art of derivation and implementation into the science of machine-executable algorithms.

The last few years have seen considerable advances in computerized symbolic algebra in many-body electron-correlation methods. However, as is evident from the previous paragraph, these advances are the direct consequence of the body of knowledge accumulated by the prior *manual* or *semi-computerized* derivation and implementation efforts, the latter being limited only by computer power. In fact, the most labor-intensive areas of computational quantum chemistry, e.g., that of coupled-cluster theory, many-body perturbation theory, molecular integral evaluation and multipole expansions of long-range forces, spin and spatial symmetry adaptation of wave functions, exchange-correlation functionals in density functional theory, have always embraced the computerization or graphical systematization as its predecessor (the vast field of symbolic algebra in chemistry in general has been the topic of a more comprehensive review by Barnett et al. [40]). These incessant efforts to simplify symbolic algebra were also responsible for some of the breakthroughs in the field, e.g., the linked cluster theorem and the concept of size-extensivity [33,34,41]. In this article, confining ourselves to electron correlation, we review the new paradigm of method development by virtue of computerizing the symbolic algebra.

2 Automated derivation and implementation

2.1 Automated derivation of perturbation theory

Wong and Paldus [42,43] were among the first to report an algorithm and the corresponding computer program that automated the derivation of many-body perturbation theory at any given order. This algorithm essentially simulates the diagrammatic (as opposed to algebraic) derivations of order-by-order perturbation corrections to energy. Their program generates all topologically distinct, linked diagrams of the Hugenholtz type, which can in turn be expanded to the Goldstone diagrams. Each diagram is represented by a string of integers; a similar technique was later employed by Kállay and Surján [39] in their string-based coupled-cluster method and by others. See also Csépes and Pipek [44] and Herbert and Ermler [45] (the latter is an application to molecular vibrational–rotational analyses).

2.2 Automated derivation of coupled-cluster theory

The coupled-cluster expansion was originally introduced by Coester and Kümmel [46,47] in nuclear physics and applied to quantum chemistry by Čížek and Paldus [35,48–50]. Čížek [35] detailed the second-quantization rules (including normal-ordered operators and their anti-commutation relations) and corresponding Feynman diagrams for deriving the algebraic equations of the spin-orbital and spin-adapted coupled-cluster doubles (CCD) methods and, with assistance from Paldus, performed the first CCD calculations of N_2 and benzene with a limited number of correlated orbitals. Two captivating personal accounts by Paldus [51] and by Bartlett [52] on coupled-cluster developments will appear. Their modern implementation style was shaped by the CCD development of Pople et al. [53] and of Bartlett and Purvis [54] and by the CCSD development of Purvis and Bartlett [23]. What is less known is the fact that this initial implementation of CCSD has been assisted by symbolic computer program KOMMUTE (written by Purvis) that verified the diagrammatic derivation. The coupled-cluster methods were sufficiently complex even at its lower orders to warrant an expedient diagrammatic technique or a computerized system or the combination of both. Harris [32] was among those who helped establish these techniques, although his automated derivation system based on *Maple* was not published until much later [55]. Other symbolic manipulation systems for second-quantized operators have been developed by Berente et al. [56] in their spin-restricted CCSDT development, by Bochevarov and Sherrill [57], and by Knowles and Handy [58]. Some of these groups have used general mathematical packages such as *Mathematica* and REDUCE as the basis of their symbolic manipulation programs. For the simplest reference wave functions, it is possible to derive and also solve algebraic coupled-cluster equations by using *Maple*. In this fashion, Čížek et al. [59] performed approximate CCSDT calculations (equivalent to the so-called CCSDT-1) of benzene within the Pariser–Parr–Pople model, whereas Piecuch et al. [60] studied the effect of the connected quadruple excitation operator on quasi-degenerate wave functions of hydrogen atom cluster models.

2.3 Automated implementation (Janssen and Schaefer)

An important advance was made by Janssen and Schaefer [61] who, in addition to automating the derivation, also computer synthesized the codes. Theirs was the most remarkable early attempt to explore the full potential of computerized symbolic algebra in quantum chemistry. They introduced a number of general techniques for rapid pattern matching, data representation, code generation, and optimization. The automatic derivation of the working equations for any given method defined in terms of second-quantized operators and Slater determinants is only one of the capabilities of their computerized system SQSYM. Once these sum-of-product tensor equations are derived, the SQSYM subsequently performs the strength reduction, a process that determines an optimal binary contraction order for each tensor product. After an exhaustive search, the order with the least operation cost is identified, e.g., ((AB)C)D for a tensor product ABCD, defining intermediate tensors: $I_1 = \mathrm{AB}$, $I_2 = I_1\mathrm{C}$, and $I_3 = I_2\mathrm{D}$. If two or more orders have the same minimum operation cost, the one requiring the least memory space for intermediate tensors is chosen. Furthermore, the SQSYM recognizes the occurrence of the same intermediate tensors in different tensor products. These are reusable intermediates that need not be computed more than once. The SQSYM-synthesized programs compute and store them once and reuse them as many times as needed. This is an example of a general compiler optimization technique called common subexpression elimination. The computational sequence thus determined were not translated into FORTRAN or any other general computer language, but instead directly interpreted and executed by another program CORR. It has an in-place binary tensor contraction engine at its algorithmic core, which can make use of Abelian point-group symmetry and index permutation symmetry to a limited extent; a similar contraction engine has been devised by Cole and Purvis [62] and by Kucharski and Bartlett [37]. The SQSYM has enabled the automatic derivation and implementation of a spin-adapted high-spin open-shell coupled-cluster method with single and double cluster excitation operators.

2.4 Automated implementation (Li and Paldus)

A more general spin-adapted open-shell coupled-cluster method has been developed by Li and Paldus [63]. The method uses an approximate wave function completely analogous to the single-reference coupled-cluster theory, but with a reference configuration that is a spin-adapted combination of determinants and with the cluster excitation operators that are orthogonal unitary group operators. It has the desirable property of the invariance with respect to the orbital rotations separately among the core, valence, and virtual spaces. The cluster excitation operators that satisfy this condition and give rise to spin-adapted states for open-shell singlet, doublet, and triplet states are sufficiently complex to warrant computerized derivation and implementation. For instance, the single and double cluster excitation operators for open-shell triplets now have thirty different types as opposed to only two for the spin-orbital case. All necessary matrix elements involving these operators are automatically generated by systematic consideration of permutation symmetry among the core, valence, and virtual spaces. Subsequently, their automated system simplifies the derived formulae, groups the terms, and searches for common loops and integrals. Furthermore, the computer-synthesized FORTRAN programs can take advantage of Abelian point-group symmetry and frozen core and virtual orbitals. A closely related work (unitary-group-based open-shell CCSD and automated derivation and implementation) was reported by Jankowski and Jeziorski [64].

2.5 Automated implementation (Crawford et al.)

A spin-adapted non-iterative, perturbation correction accounting for the effect of connected triple excitations to open-shell CCSD is also faced with greatly increased complexity in the formalisms. Crawford et al. [65] considered one such scheme, termed CCSD(zT), that is based on Z-averaged perturbation theory (ZAPT) of Lee and Jayatilaka [66] and that maintains desirable orbital invariance properties. The ZAPT divides single, double, and triple cluster excitation operators into three, six, and ten distinct component operators, respectively, according to the occupancy of doubly or singly occupied orbitals, giving rise to a multitude of terms in the (zT) energy correction. These authors devised a *Mathematica*-based symbolic algebra program that enumerates all these distinct mathematical terms in their energy expression. Furthermore, the translation of these into a computer program has also been automated. A considerable care was exercised in verifying the computer-generated programs.

2.6 Automated implementation (Nooijen and Lotrich)

Inspired by the above-mentioned work of Li and Paldus [63], Nooijen and Lotrich [67] explored the use of double similarity transformation for defining a new open-shell CCSD method. While the *computational* complexity of the method is no worse than $O\left(n^6\right)$ (n being the measure of molecular size), the complexity of development is significantly increased by the use of two exponential excitation operators and another linear excitation operator acting on the reference determinant.

They developed an automated system for formula derivation and computer implementation called APG to overcome this increasingly burdensome symbolic algebra. The APG consists of three modules. The first of the three performs the automated derivation of a set of spin-adapted or spin-orbital many-body equations from input high-level definitions. This part uses Wick's theorem for normal ordered second-quantized operators and is generally applicable to most any ansatz. Furthermore, it has additional capabilities that go outside the domain of the second-quantized language such as the multiplication of density matrices, detecting and discarding disconnected diagrammatic terms, and differentiating the equations. This step uses the rapid pattern matching techniques similar to the ones reported earlier by Janssen and Schaefer [61]. The second part performs the strength reduction followed by factorization, which were assumed to be decoupled. The third step implements the computational sequences generated by these two previous steps into computer programs that are as efficient as the equivalent, hand-written ones. This efficiency has been achieved by delegating the compute-intensive kernel of tensor contractions to highly-tuned library subroutines that are called by the computer-synthesized programs in an appropriate order.

The APG has rapidly developed new, highly involved but accurate methods for electron correlation such as extended similarity-transformed equation-of-motion (STEOM) CCSD [68], Brueckner-based generalized coupled-cluster method [69], and state-selective multireference coupled-cluster method [70]. The unique capability of the APG for differentiation is essential when the methods for analytical evaluation of energy derivatives or higher-order properties are sought. This has also been extended to a remarkable degree by M. Wladyslawski and M. Nooijen (private communication) in their symbolic manipulation program SMART and was used for the automatic derivation of analytical energy gradient formulae of STEOM-CCSD for excited, ionized, and electron-attached states and other related methods such as CCSD for the ground, excited, (one- and two-electron) ionized, and (one- and two-electron) electron-attached states. It has also been used in a cumulant expansion of reduced density matrices [71].

2.7 Automated implementation (Hirata)

Currently, one of the most advanced symbolic computing systems for electron correlation is the Tensor Contraction Engine (TCE) developed by us [72]. It automates all three basic constituents of symbolic algebra for computational quantum chemistry — *derivation*, *transformation*, and *implementation*. The input of the TCE is a second-quantized definition of any given many-electron theory that uses the following expectation value,

$$\langle\Phi_0|\left(\hat{L}^{\dagger}\hat{\Omega}\hat{R}_1\cdots\hat{R}_n\right)_{C/L}|\Phi_0\rangle\,, \tag{1}$$

where $|\Phi_0\rangle$ is a ground-state Slater determinant that can be (but is not limited to) a restricted or unrestricted Hartree–Fock reference wave function. The spin-orbital operator $\hat{\Omega}$ is a physical operator of any rank (the Hamiltonian operator, dipole moment operator, etc.), $\hat{R}_1\cdots\hat{R}_n$ is a composite operator of several excitation operators of any rank, $\hat{L}^{\dagger}$ is a deexcitation operator of any rank, and $(\cdots)_{C/L}$ means that the connectedness and linkedness among a selected group of operators can be imposed. Equation (1) may look restrictive but it covers a multitude of electron-correlation methods listed below.

The TCE derives sum-of-product tensor expressions for the ansatz of type (1) using Wick's theorem for normal-ordered second quantization. Each of the tensor products carries an appropriate index permutation operator, reflecting the anti-symmetry of spin-orbital wave functions. These mathematical expressions are then automatically transformed into computational sequences (Fig. 1). The steps are the strength reduction, the factorization, and the common subexpression elimination. Figure 2 illustrates these steps for the T_1 amplitude equation for CCSD.

The automated implementation of the computational sequences thus determined must honor domain-specific optimizations: the use of spin and spatial symmetry of the molecule in question and index permutation symmetries of tensors and tensor equations. Among these, the use of spin and spatial symmetries, which amounts to excluding tensor elements that are symmetrically zero from storage and arithmetic, is relatively straightforward. Rather, the challenge is to thoroughly incorporate the index permutation symmetry logic into automatically synthesized programs; they must distinguish unique tensor elements from symmetrically equivalent ones and exclude the latter from storage and arithmetic. To do so, the TCE must know the index permutation symmetries of not only input and output tensors but also of intermediate tensors that differ from one method to another and from one computational sequence to another.

The initial solution to this problem has been suggested by Kállay and Surján [39] (see Sect. 3.4). Currently, we have an understanding of which permutation symmetries the intermediate tensors of many-electron methods defined by Eq. (1) possess and which computational sequences give rise to the most symmetrical intermediate tensors [72–74]. For instance, each of the intermediate tensors of a coupled-cluster method of any order can be shown to have not more than four groups of permutable indices, when an appropriate computational sequence is chosen. The TCE's program synthesis algorithm is based on this critical information.

①

$$\langle 0|\{h_1^\dagger p_2\}\left[\hat{H}\exp\left(\hat{T}_1+\hat{T}_2\right)\right]_c|0\rangle$$

②

$$\begin{aligned}
&+f_{g_4}^{g_3}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4\}|0\rangle+\frac{1}{4}v_{g_5g_6}^{g_3g_4}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4^\dagger g_6g_5\}|0\rangle+f_{g_4}^{g_3}t_{h_6}^{p_5}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4\}\{p_5^\dagger h_6\}|0\rangle\\
&+\frac{1}{4}v_{g_5g_6}^{g_3g_4}t_{h_8}^{p_7}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4^\dagger g_6g_5\}\{p_7^\dagger h_8\}|0\rangle+\frac{1}{4}f_{g_4}^{g_3}t_{h_7h_8}^{p_5p_6}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4\}\{p_5^\dagger p_6^\dagger h_8h_7\}|0\rangle\\
&+\frac{1}{16}v_{g_5g_6}^{g_3g_4}t_{h_9h_{10}}^{p_7p_8}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4^\dagger g_6g_5\}\{p_7^\dagger p_8^\dagger h_{10}h_9\}|0\rangle+\frac{1}{2}f_{g_4}^{g_3}t_{h_6}^{p_5}t_{h_8}^{p_7}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4\}\{p_5^\dagger h_6\}\{p_7^\dagger h_8\}|0\rangle\\
&+\frac{1}{8}v_{g_5g_6}^{g_3g_4}t_{h_8}^{p_7}t_{h_{10}}^{p_9}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4^\dagger g_6g_5\}\{p_7^\dagger h_8\}\{p_9^\dagger h_{10}\}|0\rangle\\
&+\frac{1}{16}v_{g_5g_6}^{g_3g_4}t_{h_8}^{p_7}t_{h_{11}h_{12}}^{p_9p_{10}}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4^\dagger g_6g_5\}\{p_7^\dagger h_8\}\{p_9^\dagger p_{10}^\dagger h_{12}h_{11}\}|0\rangle\\
&+\frac{1}{24}v_{g_5g_6}^{g_3g_4}t_{h_8}^{p_7}t_{h_{10}}^{p_9}t_{h_{12}}^{p_{11}}\langle 0|\{h_1^\dagger p_2\}\{g_3^\dagger g_4^\dagger g_6g_5\}\{p_7^\dagger h_8\}\{p_9^\dagger h_{10}\}\{p_{11}^\dagger h_{12}\}|0\rangle
\end{aligned}$$

③

$$\begin{aligned}
&+f_{h_1}^{p_2}-f_{h_1}^{h_3}t_{h_3}^{p_2}+f_{p_3}^{p_2}t_{h_1}^{p_3}-t_{h_4}^{p_3}v_{h_1p_3}^{h_4p_2}+f_{p_4}^{h_3}t_{h_3h_1}^{p_4p_2}+\frac{1}{2}t_{h_4h_5}^{p_3p_2}v_{h_1p_3}^{h_4h_5}+\frac{1}{2}t_{h_5h_1}^{p_3p_4}v_{p_3p_4}^{h_5p_2}-t_{h_1}^{p_3}t_{h_4}^{p_2}f_{p_3}^{h_4}\\
&+t_{h_3}^{p_2}t_{h_5}^{p_4}v_{h_1p_4}^{h_5h_3}+t_{h_1}^{p_3}t_{h_5}^{p_4}v_{p_4p_3}^{h_5p_2}+\frac{1}{2}t_{h_4h_5}^{p_3p_2}t_{h_1}^{p_6}v_{p_6p_3}^{h_4h_5}+\frac{1}{2}t_{h_5h_1}^{p_3p_4}t_{h_6}^{p_2}v_{p_3p_4}^{h_6h_5}+t_{h_4h_1}^{p_3p_2}t_{h_6}^{p_5}v_{p_5p_3}^{h_6h_4}-t_{h_3}^{p_2}t_{h_1}^{p_4}t_{h_6}^{p_5}v_{p_5p_4}^{h_6h_3}
\end{aligned}$$

④

$$\begin{aligned}
&(\xi_8)_{h_1}^{h_4}=+f_{p_3}^{h_4}t_{h_1}^{p_3}\\
&(\xi_9)_{h_1}^{h_3}=+v_{h_1p_4}^{h_3h_5}t_{h_5}^{p_4}\\
&(\xi_{10})_{p_3}^{p_2}=+v_{p_3p_4}^{h_5p_2}t_{h_5}^{p_4}\\
&(\xi_{11})_{h_1p_3}^{h_4h_5}=+v_{p_3p_6}^{h_4h_5}t_{h_1}^{p_6}\\
&(\xi_{12})_{h_1}^{h_6}=-v_{p_3p_4}^{h_5h_6}t_{h_1h_5}^{p_3p_4}\\
&(\xi_{13})_{p_3}^{h_4}=+v_{p_3p_5}^{h_4h_6}t_{h_6}^{p_5}\\
&(\xi_{141})_{p_3}^{h_4}=+v_{p_3p_5}^{h_4h_6}t_{h_6}^{p_5}\\
&(\xi_{14})_{h_1}^{h_4}=+(\xi_{141})_{p_3}^{h_4}\,t_{h_1}^{p_3}\\
&r_{h_1}^{p_2}=+f_{h_1}^{p_2}-t_{h_3}^{p_2}f_{h_1}^{h_3}+t_{h_1}^{p_3}f_{p_3}^{p_2}-v_{h_1p_3}^{h_4p_2}t_{h_4}^{p_3}+t_{h_1h_3}^{p_2p_4}f_{p_4}^{h_3}-\frac{1}{2}v_{h_1p_3}^{h_4h_5}t_{h_4h_5}^{p_2p_3}-\frac{1}{2}v_{p_3p_4}^{h_5p_2}t_{h_1h_5}^{p_3p_4}\\
&-(\xi_8)_{h_1}^{h_4}\,t_{h_4}^{p_2}-(\xi_9)_{h_1}^{h_3}\,t_{h_3}^{p_2}-(\xi_{10})_{p_3}^{p_2}\,t_{h_1}^{p_3}+\frac{1}{2}(\xi_{11})_{h_1p_3}^{h_4h_5}\,t_{h_4h_5}^{p_2p_3}-\frac{1}{2}(\xi_{12})_{h_1}^{h_6}\,t_{h_6}^{p_2}+(\xi_{13})_{p_3}^{h_4}\,t_{h_1h_4}^{p_2p_3}-(\xi_{14})_{h_1}^{h_4}\,t_{h_4}^{p_2}
\end{aligned}$$

⑤

$$\begin{aligned}
&(\xi_{22})_{p_3}^{h_7}=+f_{p_3}^{h_7}-t_{h_6}^{p_5}v_{p_3p_5}^{h_6h_7}\\
&(\xi_2)_{h_1}^{h_7}=+f_{h_1}^{h_7}+t_{h_1}^{p_3}\,(\xi_{22})_{p_3}^{h_7}-t_{h_5}^{p_4}v_{h_1p_4}^{h_5h_7}-\frac{1}{2}t_{h_1h_5}^{p_3p_4}v_{p_3p_4}^{h_5h_7}\\
&(\xi_3)_{p_3}^{p_2}=+f_{p_3}^{p_2}-t_{h_5}^{p_4}v_{p_3p_4}^{h_5p_2}\\
&(\xi_5)_{p_7}^{h_8}=+f_{p_7}^{h_8}+t_{h_6}^{p_5}v_{p_5p_7}^{h_6h_8}\\
&(\xi_6)_{h_1p_3}^{h_4h_5}=+v_{h_1p_3}^{h_4h_5}-t_{h_1}^{p_6}v_{p_3p_6}^{h_4h_5}\\
&r_{h_1}^{p_2}=+f_{h_1}^{p_2}-t_{h_7}^{p_2}\,(\xi_2)_{h_1}^{h_7}+t_{h_1}^{p_3}\,(\xi_3)_{p_3}^{p_2}-t_{h_4}^{p_3}v_{h_1p_3}^{h_4p_2}+t_{h_1h_8}^{p_2p_7}\,(\xi_5)_{p_7}^{h_8}-\frac{1}{2}t_{h_4h_5}^{p_2p_3}\,(\xi_6)_{h_1p_3}^{h_4h_5}-\frac{1}{2}t_{h_1h_5}^{p_3p_4}v_{p_3p_4}^{h_5p_2}
\end{aligned}$$

⑥

$$\begin{aligned}
&(\Xi_1)_{p_3}^{h_7}=-t_{h_6}^{p_5}v_{p_3p_5}^{h_6h_7}\\
&(\xi_{22})_{p_3}^{h_7}=+f_{p_3}^{h_7}+(\Xi_1)_{p_3}^{h_7}\\
&(\xi_2)_{h_1}^{h_7}=+f_{h_1}^{h_7}+t_{h_1}^{p_3}\,(\xi_{22})_{p_3}^{h_7}-t_{h_5}^{p_4}v_{h_1p_4}^{h_5h_7}-\frac{1}{2}t_{h_1h_5}^{p_3p_4}v_{p_3p_4}^{h_5h_7}\\
&(\xi_3)_{p_3}^{p_2}=+f_{p_3}^{p_2}-t_{h_5}^{p_4}v_{p_3p_4}^{h_5p_2}\\
&(\xi_5)_{p_7}^{h_8}=+f_{p_7}^{h_8}+(\Xi_1)_{p_7}^{h_8}\\
&(\xi_6)_{h_1p_3}^{h_4h_5}=+v_{h_1p_3}^{h_4h_5}-t_{h_1}^{p_6}v_{p_3p_6}^{h_4h_5}\\
&r_{h_1}^{p_2}=+f_{h_1}^{p_2}-t_{h_7}^{p_2}\,(\xi_2)_{h_1}^{h_7}+t_{h_1}^{p_3}\,(\xi_3)_{p_3}^{p_2}-t_{h_4}^{p_3}v_{h_1p_3}^{h_4p_2}+t_{h_1h_8}^{p_2p_7}\,(\xi_5)_{p_7}^{h_8}-\frac{1}{2}t_{h_4h_5}^{p_2p_3}\,(\xi_6)_{h_1p_3}^{h_4h_5}-\frac{1}{2}t_{h_1h_5}^{p_3p_4}v_{p_3p_4}^{h_5p_2}
\end{aligned}$$

Fig. 2 Automated derivation and transformation of the CCSD T_1 amplitude equation. (*1*) The second-quantized definition. (*2*) Expanded normal-ordered second-quantized definition. (*3*) The sum-of-product tensor expressions obtained after the contraction of creation and annihilation operators and the deletion of disconnected terms. (*4*) The result of strength reduction and the definition of intermediate tensors as binary tensor products. (*5*) The result of factorization and the redefinition of intermediate tensors as sums of binary tensor products. (*6*) The identification of reusable intermediate tensors (common subexpression elimination)

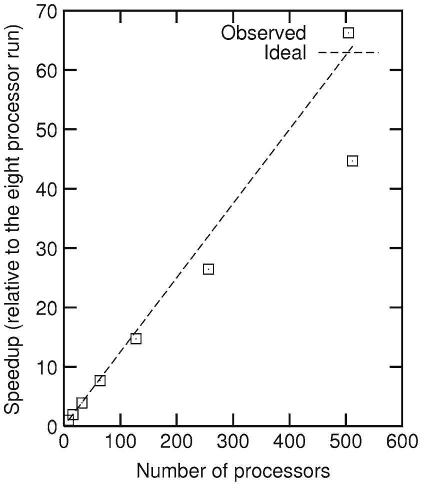

Fig. 3 The speedup (relative to the eight processor run) of a parallel execution of the perturbation correction calculation of the CCSD(T) code [75]

Another difficult problem is to find efficient storage and contraction algorithms for tensors with multiple symmetry attributes. These algorithms must also be general enough to be subject to automation. Kállay and Surján [39] resorted to the string-based algorithm (see Sect. 4.8) to address this issue effectively, whereas we employed the so-called *tiling* algorithm. The latter divides the orbital range into spin- and spatial-symmetry-adapted groups (*tiles*) and handles the tensors at the *tile* level rather than at the *element* level. This allows the TCE to incorporate the complex permutation symmetry logic as well as the spin and spatial symmetry ones in loops over *tiles* (not *elements*) in the synthesized programs (*cf.* the direct product decomposition scheme in Sect. 3.2). In this way, we can capture the essential symmetry advantages and simultaneously avoid an excessive overhead of examining symmetry properties of individual elements of tensors. Furthermore, the tile-wise tensor contractions are largely independent operations from one another and are ideally executed in parallel.

The TCE has synthesized the parallel execution programs for configuration-interaction, many-body perturbation, and coupled-cluster methods through fourth order [72], coupled-cluster Λ equation solver for ground-state properties, and equation-of-motion coupled-cluster methods for excitation energies and excited-state properties through fourth order [73]. It has also *semi* automated the development of various second-order perturbation corrections to coupled-cluster methods [75] and second- [76] through fourth-order perturbation corrections to configuration-interaction singles for excitation energies [77]. The critical importance of our ability to verify and validate the optimized computer programs was emphasized recently in the context of computational physics [78]. All of our automatically synthesized programs were tested against independent *determinant-based* implementations (see Sect. 4). The scalability of the parallel execution speed is reasonable for the perturbation correction calculation in CCSD(T) (Fig. 3) [75]. The synthesized programs are characterized by the optimal polynomial dependence on the system sizes [e.g., $O\left(n^{2k+2}\right)$ for kth order coupled-cluster methods], by the use of spin symmetry (within the spin-orbital formalisms), real Abelian point-group symmetry, and index permutation symmetry at every stage, and by the general applicability to open- and closed-shell molecules. These computerized implementations enabled highly accurate predictions of spectroscopic properties of small molecules, some containing heavy elements (Tables 1–3) [11].

2.8 Computerized optimizations

In addition to the strength reduction, factorization, and common subexpression elimination mentioned above, automated implementations can perform *coupled* strength reduction and factorization, *memory minimization* (that fuses common loops and minimizes the memory consumption without increasing the operation cost), *space-time-tradeoff* (that more aggressively minimizes the memory consumption by recomputing some large input or intermediate tensors), *data partitioning* (that maximizes the data locality and minimizes the interprocessor communications in parallel executions), etc. These are employed in conjunction with the domain-specific optimizations (symmetries, sparsity, boundary conditions, etc.). A unique accomplishment is being made along this direction by a team of quantum chemists and computer scientists led by Sadayappan [79–82]. For instance, their optimization capability can perform an exhaustive or quick search of the *coupled* strength reduction and factorization problem (see Refs. [61,67] for discussions on its impossibly vast search space). This yields an efficient computational sequence for, e.g., CCSD and CCSDT equations. When the ratio of the number of virtual orbitals (v) to the number of occupied orbitals (o) becomes high, the *decrease* in the number of $O\left(\mathrm{o}^1 v^4\right)$ contractions at the cost of the *increase* in the number of $O\left(\mathrm{o}^4 v^2\right)$ contractions leads to an overall cheaper computational sequence for the CCSD T_2 equation [80]. This runs counter to the intuition of most quantum chemists.

Lotrich et al. (private communication) developed a computer language specialized for the development of scalable parallel computer programs for coupled-cluster and other electron-correlation methods. This computer language, called the super instruction assembly language (SIAL), is placed higher in the scale of abstraction than the usual high-level computer languages such as FORTRAN or C. It describes the necessary arithmetic operations and in-core and out-of-core data storage and retrieval across different processors in an abstract syntax. This enabled Lotrich et al. to implement various alternative parallel algorithms of computing energies and analytical energy gradients at the CCSD level in a short time.

3 Systematization of symbolic algebra in quantum chemistry

3.1 Diagrammatic techniques

It must not be overlooked that, in the context of manual development of electron correlation methods, considerable efforts

Table 1 Equilibrium bond lengths (r_e in Å) and dissociation energies (D_0 in eV) of second- through fifth-row diatomic hydrides [11]

Diatomic hydrides	r_e		D_0	
	Theory	Experimental	Theory	Experimental
BH	1.233	1.232	3.53	3.54
CH	1.120	1.120	3.47	3.47
NH	1.037	1.037	3.39	3.40
OH	0.970	0.970	4.40	4.39
FH	0.917	0.917	5.85	5.87
AlH	1.649	1.648	3.07	3.06
SiH	1.522	1.520	3.07	3.04
PH	1.425	1.422	3.03	3.02
SH	1.344	1.341	3.61	3.62
ClH	1.278	1.275	4.41	4.43
GaH	1.659	1.663	2.82	<2.84
GeH	1.587	1.588	2.82	<3.3
AsH	1.523	1.523	2.75	2.77
SeH	1.465	1.464	3.19	3.2
BrH	1.415	1.414	3.74	3.76
InH	1.841	1.838	2.46	2.48
SnH	1.771	1.770	2.53	<2.73
SbH	1.712	1.711	2.47	
TeH	1.656	1.656	2.76	
IH	1.609	1.609	3.05	3.05

The theoretical estimates are based on the various combinations of CCSD, CCSDT, and CCSDTQ results and cc-pVDZ, TZ, and QZ basis sets. The corrections due to the special theory of relativity are taken into account within the third-order Douglas–Kroll approximation. For the details of complete-basis-set extrapolation and the sources of experimental data, the readers are referred to the original paper [11]

Table 2 Singlet–triplet separations (in kcal mol^{-1}) of CH_2, NH_2^+, SiH_2, PH_2^+ and AsH_2^+ [11]

	CH_2	NH_2^+	SiH_2	PH_2^+	AsH_2^+
Theory	9.1	29.3	−20.9	−18.3	−23.2
Experiment	9.0	30.1	−21.0	−18.4	−23.8

The theoretical estimates are based on the various combinations of CCSD, CCSDT, and CCSDTQ results and cc-pVDZ, TZ, QZ, and 5Z basis sets. The zero-point vibrational energy, relativistic, and core correlation (and Born–Oppenheimer for CH_2) corrections are taken into account. For the details of complete-basis-set extrapolation and the sources of experimental data, the readers are referred to the original paper [11]

Table 3 Adiabatic excitation energies of CH and vertical excitation energies of formaldehyde (in eV) [73]

CH	$a^4\Sigma^-$	$A^2\Delta$	$B^2\Sigma^-$	$C^2\Sigma^+$	
EOM-CCSDT/aug-cc-pVTZ	0.74	2.94	3.27	4.03	
Experiment	0.74	2.88	3.23	3.94	
CH_2O	$^1A_2$	$^1B_2$	$^1B_2$	$^1A_1$	$^1B_1$
EOM-CCSD/d-aug-cc-pVTZ	3.92	7.18	8.02	8.12	9.06
EOM-CCSDT/aug-cc-pVDZ	3.99	7.02	7.99	8.05	9.29
Experiment	4.0	7.08, 7.10	7.97, 7.98	8.14	9.03

have been made to systematize and generalize the strategies for efficient and accurate symbolic manipulations. These are the foundation of the above-mentioned computerized methods. It is well known that the specialized form of Feynman diagrams for many-body perturbation theory devised by Goldstone led to the linked cluster theorem and the concept of size extensivity [33,34,41]. This and related diagrammatic techniques and underlying normal ordering and Wick's theorem [29–32] can be viewed as the predecessor of the ultimate simplification of this entire process – the complete automation. It is extraordinary that coupled-cluster method including up to as high as quintuple cluster excitation operator [83, 84] and equation-of-motion coupled-cluster method including triple cluster and linear excitation operators [85–88] have been developed manually by this time-honored technique.

3.2 Recursive intermediate factorization and direct product decomposition

Others include the strength reduction (computational linearization) of coupled-cluster equations [32,37]. Kucharski and Bartlett [37] proposed a simple algorithm (recursive intermediate factorization) for factorizing these computationally linearized coupled-cluster equations, which is suitable for automation. This may not yield an optimally efficient computational sequence for any given coupled-cluster equation, but is reasonable and has been adopted by Nooijen and Lotrich [67] and by us [72] in their automated systems. The direct product decomposition scheme of Stanton et al. [38] is a systematic and general algorithm of incorporating Abelian point-group and other symmetries (e.g., double group symmetry in relativistic quantum chemistry) in virtually any electron-correlation calculations that are in essence the evaluation of sum-of-product tensor expressions. Hence, this offers a universal algorithm for computerized program synthesis. Stanton et al. [38] observed that their algorithm exposes an adequate granularity of parallelism, which has later been demonstrated to be the case by us [72].

3.3 Tensor formulation of many-electron theories

Another important generalization of electron-correlation theories was brought forth by Head-Gordon et al. [89]. They have shown that, when recast in the most general form that accommodates any reference orbitals, many-body electron-correlation methods are invariably in the sum-of-product tensor expressions. For instance, the usual Møller–Plesset perturbation energy expressions that involve orbital energies (which are vectors) can be brought to a more general tensor algebraic form that no longer involves orbital energies. This general tensor formalism, unlike the mixed tensor-vector formalism, works with noncanonical or nonorthogonal reference orbitals and is considered more fundamental. The formalisms of disparate groups of theories (coupled-cluster, configuration-interaction, many-body perturbation theories, and various combinations thereof) are subject to the unified representation that is best suited to computerized derivation and implementation.

3.4 Index permutation symmetry of intermediate tensors

Electron-correlation methods are defined either in spin-adapted forms or in spin-orbital forms. In the latter, the thorough use of index permutation symmetry in the tensor storage and contractions is crucial [39]. For instance, a tensor representing the triple excitation operator $T^{p_1p_2p_3}_{h_4h_5h_6}$ (p and h denote particle and hole indices, respectively) has the 36-fold index permutation symmetry:

$$\begin{aligned}
T^{p_1p_2p_3}_{h_4h_5h_6} &= -T^{p_1p_3p_2}_{h_4h_5h_6} = -T^{p_2p_1p_3}_{h_4h_5h_6} = T^{p_2p_3p_1}_{h_4h_5h_6} = T^{p_3p_1p_2}_{h_4h_5h_6} \\
&= -T^{p_3p_2p_1}_{h_4h_5h_6} = -T^{p_1p_2p_3}_{h_4h_6h_5} = T^{p_1p_3p_2}_{h_4h_6h_5} = T^{p_2p_1p_3}_{h_4h_6h_5} \\
&= -T^{p_2p_3p_1}_{h_4h_6h_5} = -T^{p_3p_1p_2}_{h_4h_6h_5} = T^{p_3p_2p_1}_{h_4h_6h_5} = -T^{p_1p_2p_3}_{h_5h_4h_6} \\
&= T^{p_1p_3p_2}_{h_5h_4h_6} = T^{p_2p_1p_3}_{h_5h_4h_6} = -T^{p_2p_3p_1}_{h_5h_4h_6} = -T^{p_3p_1p_2}_{h_5h_4h_6} \\
&= T^{p_3p_2p_1}_{h_5h_4h_6} = T^{p_1p_2p_3}_{h_5h_6h_4} = -T^{p_1p_3p_2}_{h_5h_6h_4} = -T^{p_2p_1p_3}_{h_5h_6h_4} \\
&= T^{p_2p_3p_1}_{h_5h_6h_4} = T^{p_3p_1p_2}_{h_5h_6h_4} = -T^{p_3p_2p_1}_{h_5h_6h_4} = T^{p_1p_2p_3}_{h_6h_4h_5} \\
&= -T^{p_1p_3p_2}_{h_6h_4h_5} = -T^{p_2p_1p_3}_{h_6h_4h_5} = T^{p_2p_3p_1}_{h_6h_4h_5} = T^{p_3p_1p_2}_{h_6h_4h_5} \\
&= -T^{p_3p_2p_1}_{h_6h_4h_5} = -T^{p_1p_2p_3}_{h_6h_5h_4} = T^{p_1p_3p_2}_{h_6h_5h_4} = T^{p_2p_1p_3}_{h_6h_5h_4} \\
&= -T^{p_2p_3p_1}_{h_6h_5h_4} = -T^{p_3p_1p_2}_{h_6h_5h_4} = T^{p_3p_2p_1}_{h_6h_5h_4}.
\end{aligned} \tag{2}$$

Any implementation that does not take advantage of this symmetry will be consuming too much memory and involving too many redundant arithmetic operations to be useful. The prize will be progressively greater with the increasing order of electron-correlation methods. The index permutation symmetry for $T^{p_1p_2p_3p_4}_{h_5h_6h_7h_8}$ and $T^{p_1p_2p_3p_4p_5}_{h_6h_7h_8h_9h_{10}}$ is 576 and 14,400 fold, respectively. This ultimately reflects the antisymmetry of wave functions and has more profound physical significance than just a performance optimization aspect. In fact, spin-orbital coupled-cluster programs that do not pay due attention to the index permutation symmetry can yield a non-physical wave function that satisfies neither Fermi–Dirac nor Bose–Einstein statistics [72].

Nonetheless, the index permutation symmetry has largely been ignored in automated implementation until recently. This is partly because the index permutation symmetry of *intermediate* tensors – the tensors which are neither input nor output tensors and that result from the strength reduction and factorization – was not even known a priori. However, in an important paper (mentioned in Sect. 4.8 in a greater detail) by Kállay and Surján [39], they showed that, for a certain class of electron-correlation methods (including coupled-cluster methods for the ground state), the intermediate tensors arising from reasonably optimal computational sequences indeed have a priori known index permutation symmetry, i.e.,

$$I^{e_1<\cdots<e_a,\, i_1<\cdots<i_c}_{e_{a+1}<\cdots<e_b,\, i_{c+1}<\cdots<i_d}, \tag{3}$$

where e and i represent external indices (the indices of output tensors) and internal (summation) indices, respectively. Each group of indices that are arranged in an ascending order is permutable; each and any intermediate tensor has (at maximum) four groups of permutable indices in contrast to two or four groups in the input and output tensors.

This observation is the cornerstone of their string-based coupled-cluster methods and our automated system TCE that achieve the optimal theoretical efficiency of $O\left(n^{2k+2}\right)$ for kth order approximation and take advantage of permutation symmetry for storage and contraction of input, output, and intermediate tensors. The logic for the index permutation symmetry for intermediate tensors has been extended to coupled-cluster analytical gradients [90], to coupled-cluster methods for excited states and excited-state properties [73, 90], etc.

4 Determinant- and string-based methods

4.1 General-order perturbation method

Avoiding the tedious manual symbolic algebra can also be achieved by schemes alternative to complete automation. One such scheme is the *determinant-based algorithm* that enables general-order coupled-cluster and many-body perturbation methods without any explicit consideration of the methods at individual orders. This, just as the automated systems, manipulates higher mathematical constructs such as Slater determinants and operators, rather than just numbers and simple arithmetic. Therefore, the difference between the two approaches is merely in whether the computerized symbolic manipulations occur at run time (in the determinant-based algorithm) or at compile time (in automation). Highly efficient algorithms of handling determinants (and spin-adapted configurations) [90] have been devised for truncated or full configuration-interaction methods [92–98].

In this algorithm, a wave function is represented *computationally* (not just *conceptually*) as a linear combination of Slater determinants and each determinant is in turn expressed by a pair of α and β strings of bits specifying the α and β orbital occupancies. Knowles and Handy [92] developed an ingenious algorithm to map these strings to consecutive, lexical addresses and performed rapid second-quantized operations on the determinants represented by these strings. This and related algorithms therefore permit the facile implementation of the whole hierarchical methods in a single algorithmic framework. For instance, with the highest level of generality, the Rayleigh–Schrödinger perturbation expansions of exact energy and wave function can be defined in a recursive fashion as

$$E^{(i)} = \left\langle \Psi^{(0)} \right| \hat{V} \left| \Psi^{(i-1)} \right\rangle, \tag{4}$$

$$\left(E^{(0)} - \hat{H}_0\right) \left|\Psi^{(i)}\right\rangle = \hat{V} \left|\Psi^{(i-1)}\right\rangle - \sum_{j=1}^{i} E^{(j)} \left|\Psi^{(i-j)}\right\rangle, \tag{5}$$

where $E^{(i)}$ and $\left|\Psi^{(i)}\right\rangle$ are the ith order corrections to energy and wave function and the Hamiltonian operator is partitioned as $\hat{H} = \hat{H}_0 + \hat{V}$. Representing second-quantized operators and wave functions as computational objects that can be explicitly and directly handled by the algorithm, one can translate these equations *literally* into a single computer code and perform the perturbation expansions to any arbitrary high order. In this fashion, Knowles et al. [99] carried out Møller–Plesset perturbation calculations ($\hat{H}_0$ is the sum of one-particle Fock operators) through tenth order. Essentially the same method has been proposed independently by Laidig et al. [100], who studied the convergence of the Møller–Plesset perturbation expansion through fifth order. An application to molecular vibrations has also been reported by Christiansen [101].

The generality and ease of implementation brought to by the determinant-based algorithm come at the cost of considerably increased number of operations. The calculation of the wave function and energy at each perturbation order involves a factorial number of determinants and its cost scaling is essentially that of full configuration-interaction method. When a priori information about the structure of wave functions (e.g., the knowledge of the highest rank of the constituent excited determinants) is used, the cost scaling can be reduced, but to a level that is still much higher than the optimal one. Therefore, the applicability of this algorithm for perturbation theory is severely limited, but its generality makes it indispensable for initial assessments of the performance of the whole hierarchical methods and for obtaining reference computer programs against which more optimized programs can be verified. For instance, Knowles and Handy [58] used this algorithm to verify their open-shell second-order Møller–Plesset perturbation method.

4.2 Divergence of perturbation series

It was with the aid of the determinant-based algorithm that Olsen and coworkers [102–104] made a striking observation about the divergence of the Møller–Plesset perturbation theory. While occasional divergent behavior in perturbation theory is well known (electron gas, hydrogenic stark effect, and harmonic oscillator) [105,106], they found that the Møller–Plesset perturbation expansion of a wave function dominated by a single determinant can frequently become divergent by the presence of diffuse basis functions [102]. The seemingly sensible use of the aug-cc-pVDZ basis set [107] for the neon atom, the fluorine anion, and hydrogen fluoride all results in oscillatory divergent behavior manifesting itself at the 16th, 5th, and 14th orders, respectively. The cause of this is now identified as the intruder states [108] – spurious low-lying excited states – that make the single-determinant reference appear less dominant in the wave function and perturbation treatment of electron correlation less adequate (the most thorough analysis on this subject appeared recently [109]). On this basis, Olsen et al. [102] and others [110] questioned the fundamental value of the Møller–Plesset perturbation theory as a convergent many-electron theory, while others argue in favor of the theory that the divergent series contains sufficient information to recover the same exact limit in a number of ways, e.g., variation perturbation theory [111], Padé approximant [109,111,112], Feenberg scaling [113] (see Table 4).

4.3 General-order coupled-cluster method

Evidently, the determinant-based algorithm is applicable to any method that possesses a wave function that is a linear combination of Slater determinants. Such methods include coupled-cluster methods and various combinations of coupled-cluster and many-body perturbation or configuration-interaction methods. The initial applications of the algorithm to the entire hierarchy of standard (i.e., projection-type) coupled-cluster methods were reported independently by us [111], by Kállay and Surján [114], and by Olsen [115]. The algorithm proposed by us [111] consisted of generating the coupled-cluster wave function explicitly as a linear combination of Slater determinants,

$$|\Psi\rangle = \exp\left(\hat{T}\right)|\Phi_0\rangle = \left(1 + \hat{T} + \frac{\hat{T}^2}{2!} + \cdots + \frac{\hat{T}^m}{m!}\right)|\Phi_0\rangle\,, \tag{6}$$

and then having the wave function satisfy the projection equations

$$\langle\Phi_0|\,\hat{H}\,|\Psi\rangle = E, \tag{7}$$

$$\left\langle\Phi_{h_2}^{p_1}\right|\hat{H}\,|\Psi\rangle = E\left\langle\Phi_{h_2}^{p_1}\,\middle|\,\Psi\right\rangle, \tag{8}$$

$$\left\langle\Phi_{h_3h_4}^{p_1p_2}\right|\hat{H}\,|\Psi\rangle = E\left\langle\Phi_{h_3h_4}^{p_1p_2}\,\middle|\,\Psi\right\rangle, \tag{9}$$

Table 4 Convergence of configuration-interaction, many-body perturbation, coupled-cluster methods for the fluorine anion (with the 6-31+G basis set) studied by the determinant-based algorithm [111]

Order	CI	MBPT	VPCI	Padé	CC
1	0.151798	0.151798	0.151798		0.151798
2	0.009614	−0.000415	0.016774	0.011727	0.003662
3	0.007507	0.012780	0.003572	0.005326	0.000290
4	0.000442	−0.004349	0.000509	0.002144	−0.000040
5	0.000132	0.006661	0.000094	−0.006088	−0.000005
6	0.000007	−0.006350	0.000011	0.000158	0.000000
7	0.000001	0.007430	0.000002	−0.000087	0.000000
8	0	−0.008365	0	0.000010	0
9		0.010318		0.000000	
10		−0.012676		0.000000	
11		0.015885		0.000000	
12		−0.020201		0.000000	
13		0.026007		−0.000001	
14		−0.033808		0.000000	
15		0.044301		0.000000	
16		−0.058422		0.000000	

The values shown are the deviations (in hartree) from the exact (full configuration-interaction) results. The divergent perturbation series are subject to resummation in a number of ways (variation perturbation configuration interaction, Padé approximant, and coupled-cluster theory) all converging to the exact limit

etc., by adjusting the cluster excitation operator $\hat{T}$. In the above equations, $|\Phi_0\rangle$ is the reference wave function (or determinant), $\left|\Phi_{h_2}^{p_1}\right\rangle$ and $\left|\Phi_{h_3h_4}^{p_1p_2}\right\rangle$ are singly and doubly substituted determinants, E is the total electronic energy, and m is the number of electrons. The (repeated) actions of the cluster excitation operator $\hat{T}$ and the Hamiltonian operator $\hat{H}$ on a wave function are straightforward by virtue of the determinant-based algorithm, regardless of the highest excitation rank of $\hat{T}$. It should be noted that the Taylor expansion of the exponential operator in Eq. (6) terminates after a finite number of terms and hence the operation is computationally well defined.

The determinant-based algorithm can compute the energies and wave functions of the coupled-cluster method recursively through any arbitrary high order. Again, the trade-off is between the generality or ease of implementation and the optimal computational cost. The most naïve determinant implementation will incur $O(n!)$ operations even at the lowest order of coupled-cluster theory [111], but the operation cost can be easily reduced to $O\left(n^{2k+4}\right)$ for kth order coupled-cluster method by inspecting the structure of projection equations [114,115]. Furthermore, Kállay and Surján [114] devised a scheme which, on the basis of the identify,

$$\exp\left(\hat{T}\right)|\Phi_0\rangle = \left[\prod_{k=1}^{m}\exp\left(\hat{T}_k\right)\right]|\Phi_0\rangle = \left[\prod_{k=1}^{m}\lim_{p\to\infty}\left(1+\frac{\hat{T}_k}{p}\right)^p\right]|\Phi_0\rangle\,, \tag{10}$$

updates the exponential wave function by acting the cluster excitation operator once in every iteration instead of m times ($\hat{T}_k$ is k-fold cluster excitation operator). The computational cost is however still significantly greater than the optimum dependence of $O\left(n^{2k+2}\right)$. Nonetheless, these higher-order coupled-cluster results (Table 4) have become an invaluable benchmark for optimized manual [83,84] or automated [72] implementations. Christiansen [116,117] extended this to molecular vibrations recently.

Van Voorhis and Head-Gordon [118] also invoked the algorithm in the facile implementation of the variational coupled-cluster method [119] truncated at the double excitation level. Unlike the standard projection-type ansatz, this method uses an alternative energy definition of

$$E = \frac{\langle\Psi|\,\hat{H}\,|\Psi\rangle}{\langle\Psi\mid\Psi\rangle} \tag{11}$$

and variationally optimizes the cluster excitation operator $\hat{T}$ in the wave function of Eq. (6). The determinant-based algorithm is particularly well suited for this sort of performance assessment as Eq. (7) is as easily evaluated in this algorithm just as any other energy expressions. For a related analysis, see Refs. [120–122].

4.4 Equation-of-motion coupled-cluster method

We have also extended the algorithm to general-order coupled-cluster methods for excited states [123] (see also Ref. [114]) and ionized and electron-attached states [124]. These are based on a unified formalism [125–130] that is called the equation-of-motion coupled-cluster theory and it uses the wave function parameterized as

$$|\Psi_k\rangle = \exp\left(\hat{T}\right)\hat{R}_k\,|\Phi_0\rangle\,, \tag{12}$$

for kth excited (or ionized or electron-attached) state. This is also equivalent to time-dependent linear response coupled-cluster theory when just excitation energies are considered [131–138] (see also the next section). The cluster excitation operator $\hat{T}$ is held fixed at the one determined for the ground state (hence $\hat{R}_0 = 1$) and only the linear excitation (or ionization or electron-attachment) operator $\hat{R}_k$ will be varied such that the wave function satisfies

$$\langle \Phi_0 | \hat{H} | \Psi_k \rangle = E_k, \tag{13}$$

$$\left\langle \Phi_{h_2}^{p_1} \right| \hat{H} | \Psi_k \rangle = E_k \left\langle \Phi_{h_2}^{p_1} \middle| \Psi_k \right\rangle, \tag{14}$$

$$\left\langle \Phi_{h_3 h_4}^{p_1 p_2} \right| \hat{H} | \Psi_k \rangle = E_k \left\langle \Phi_{h_3 h_4}^{p_1 p_2} \middle| \Psi_k \right\rangle, \tag{15}$$

etc., in complete analogy to Eqs. (7)–(9). Hence, the equation-of-motion coupled-cluster theory combines the cluster expansion that is suitable for electron-correlation in the ground state and the configuration-interaction expansion that is the most natural to describe an excitation (or ionization or electron-attachment) process (and the differential correlation thereof) [130]. With this, the excited-state coupled-cluster calculations including up to the sextuple [123] or decuple [114] cluster and linear excitation operators have been performed and the results were later used to verify manual [83, 84] and automatic [72] implementations. We have also studied all possible combinations of the truncation ranks of the cluster and linear excitation operators and concluded that the ones naturally suggested by the linear response derivation (the equal highest rank for cluster and linear excitation operators) are also the best performing.

4.5 Symmetry-adapted-cluster configuration-interaction method

The use of the determinant-based algorithm in coupled-cluster methods, however, predates the above-mentioned work. Nakatsuji and Hirao [139,140] and Nakatsuji [141–143] reported a coupled-cluster method for closed- and open-shell excited, ionized, and electron-attached states, years prior to the modern implementations of above-mentioned excited-state coupled-cluster methods (see, e.g., Refs. [144,145] for the connection among these methods). They call the method SAC-CI for symmetry-adapted-cluster configuration-interaction. The SAC-CI shares the same physical principle which is the division of electron-correlation into dynamical correlation in the ground state described by the cluster expansion and differential correlation between the ground and excited states accounted for by the linear expansion. The uniqueness of their approach is the insistence upon the space and spin symmetry adaptation of the wave function at various stages. Hence, the ground- and excited-state wave functions are,

$$|\Psi_0\rangle = \hat{O}_0 \exp\left(\hat{T}\right) |\Phi_0\rangle, \tag{16}$$

$$|\Psi_k\rangle = (1 - |\Psi_0\rangle \langle \Psi_0|) \hat{O}_k \hat{R}_k \hat{O}_0 \exp\left(\hat{T}\right) |\Phi_0\rangle, \tag{17}$$

where the cluster excitation operator $\hat{T}$ is a symmetry-adapted linear-combination of spin-orbital excitation operator and symmetry projectors $\hat{O}_0$ and $\hat{O}_k$ further remove any symmetry contamination that may arise from the disconnected clusters. The adjustable parameters are determined either variationally or by the projection-type equations, defining *variational* or *non-variational* SAC-CI, respectively.

Nakatsuji [146] resorted to a determinant-based algorithm (which they call *excitator* method) to implement these methods, not just for the purpose of initial performance assessment but to arrive at the production-level optimized programs. They form explicitly the exponential wave function as a linear combination of Slater determinants by repeated actions of cluster excitation operator on a reference wave function. Once the exponential wave function is formed, the subsequent procedure is the symmetry-adapted configuration-interaction that is well suited to the determinant-based algorithm. Nevertheless, such an implementation, as mentioned before, will have at least $O\left(n^{2k+4}\right)$ complexity for kth order coupled-cluster method that is much greater than the optimal dependence of $O\left(n^{2k+2}\right)$. They managed to overcome this increased cost dependence by aggressively reducing the effective number of determinants according to a criterion guided by a perturbation theory argument [147]. Remarkably, the resulting computer program can perform SAC-CI calculations of the excited states of porphin [148, 149] (see also Refs. [150,151]) or even larger molecules [152–154]. Clearly, the space and spin symmetry adaptation, the perturbation selection, and the so-called general- R extension [155–157] (the use of a higher-ranked linear excitation operator than that of cluster operator) are all facilitated by the determinant-based algorithm.

4.6 Combined coupled-cluster and perturbation methods

Another example that illustrates the strength of the determinant-based algorithm was reported by us [158,159]. We implemented the entire, new hierarchy of combined coupled-cluster and many-body perturbation method for ground and excited states. This was inspired by the vast success [160] of second-order perturbation corrections to coupled-cluster methods, such as CCSD+T(CCSD) [161], CCSD(T) [162], CCSD(TQ) [163,164], CCSD(2) [75,158,159,165–168], CR-CCSD(T) [169–171], CCSDT(2) [75,158,159]. They use a three- or four-tiered approach: variation (Hartree–Fock, e.g.) for the reference wave function, cluster expansion (CCSD, e.g.) for the majority of correlation, perturbation for the tiny residual correlation, and perhaps configuration-interaction for differential correlation when an electronic transition is involved. Our formalism is based on a transparent application of Rayleigh–Schrödinger perturbation theory to the equation-of-motion coupled-cluster reference wave function that satisfies

$$\bar{H} \hat{R}_k |\Phi_0\rangle = E_k \hat{R}_k |\Phi_0\rangle, \tag{18}$$

where $\bar{H} = \exp\left(-\hat{T}\right) \hat{H} \exp\left(\hat{T}\right)$. Partitioning this effective (similarity-transformed) Hamiltonian into the zeroth order part and perturbation, $\bar{H} = \hat{H}_0 + \hat{V}$, in an appropriate proportion, we arrive at recursive equations for order-by-order

Table 5 Convergence of combined coupled-cluster and many-body perturbation theories for hydrogen fluoride and water (in their respective ground states) with the 6-31G basis set (frozen core) [158,159]

Theory	FH (0.917 Å)	H_2O (0.967 Å and 107.6°)
CCS = MP1 = HF	0.131398	0.136671
CCS(2) = MP2	0.003723	0.008215
CCS(3) = MP3	0.004932	0.006577
CCSD	0.001032	0.001545
CCSD(2)	0.000004	0.000102
CCSD(3)	0.000034	0.000077
CCSDT	0.000350	0.000449
CCSDT(2)	0.000031	0.000049
CCSDT(3)	0.000008	0.000013
CCSDTQ	0.000008	0.000012
CCSDTQ(2)	0.000001	0.000002
CCSDTQ(3)	0.000000	0.000000
CC5	0.000001	0.000003
CC6	0.000000	0.000000
CC7	0.000000	0.000000
CC8 = FCI	0	0
CCSDT-1a	0.000354	0.000492
CCSDT-2	0.000403	0.000587
CCSDT-3	0.000409	0.000590
CCSD+T(CCSD)	0.000287	0.000435
CCSD(T)	0.000414	0.000534
$\mathrm{CCSD(TQ}_f)_\Lambda$	0.000014	0.000075

The values shown are the deviations (in hartree) from the exact (full configuration-interaction) results

perturbation corrections to energy and wave functions:

$$E_k^{(i)} = \langle \Phi_0 | \hat{L}_k^\dagger \hat{V} \hat{R}_k^{(i-1)} | \Phi_0 \rangle , \tag{19}$$

$$\left(E_k^{(0)} - \hat{H}_0\right)\hat{R}_k^{(i)} |\Phi_0\rangle = \hat{V}\hat{R}_k^{(i-1)} |\Phi_0\rangle - \sum_{j=1}^{i} E_k^{(j)} \hat{R}_k^{(i-j)} |\Phi_0\rangle, \tag{20}$$

where $\hat{L}_k^\dagger$ is the deexcitation operator corresponding to the left-hand side eigenvector of $\bar{H}$ matrix representation for the kth electronic state. Each order of perturbation theory has a well-defined wave function and energy and every operation in Eqs. (18)–(20) is a combination of a finite number of second-quantized operations. With the determinant-based algorithm, we have obtained the benchmark results for the perturbation series to any arbitrary high order starting from all orders of coupled-cluster theory for the ground and excited states through the equivalents to the full configuration-interaction method (see Table 5). An important conclusion about the (lack of) size-extensivity of the methods has been drawn on this basis. The perturbation series included the second-order corrections to coupled-cluster methods for the ground state [CCSD(2), CCSDT(2)] [75] and second- [76] and higher-order corrections to configuration-interaction singles [CIS(2), CIS(3), CIS(4)] [77], which were later implemented into optimized programs by the TCE (see Sect. 2.8).

4.7 Multireference coupled-cluster and perturbation methods

The CI-like algorithmic kernel of the determinant-based scheme permits flexible multi-determinant reference for coupled-cluster and perturbation theories. Zarrabian and coworkers studied the convergence of multireference perturbation expansions of the ground-state correlation energies and excitation energies with and without Padé approximant convergence acceleration [172–174]. They employed a straightforward generalization of Rayleigh–Schrödinger perturbation theory to multireference wave function [175] which led to an order-by-order expansion of the Bloch equation. The perturbation series can be computed recursively through any arbitrarily high order by the determinant-based algorithm. Olsen's initial implementation of a general-order determinant-based coupled-cluster method also addressed the use of multireference wave function [115] (see also Ref. [176] and the next section). Olsen introduced a general active space, which can be more compact than complete active space and more flexible than restricted active space [93] and which allows almost arbitrary divisions of orbital space and flexible conditions imposed on the excitation operators permitted in each orbital space. For instance, one can use only those higher-order cluster excitation operators ($\hat{T}_3$, $\hat{T}_4$, etc.) that involve a certain number of selected important orbitals (active orbitals). This is closely related to the state-specific multireference coupled-cluster methods of Oliphant, Adamowicz, Piecuch, and others [85–87,177–180]. In fact, Adamowicz's group also recently developed an automated system that generated the diagrammatic equations and computer programs (which however did not define intermediate tensors as binary tensor products but evaluated each of the multiple tensor products at once) for this class of methods including the complete-active-space coupled-cluster method [181].

4.8 String-based coupled-cluster method

While evolving out of the above-mentioned determinant-based algorithm [114], Kállay and Surján's [39] string-based, general-order coupled-cluster method is distinguished from its predecessor in a number of important respects. It does not expand a wave function explicitly as a linear combination of Slater determinants, which would make the algorithm inevitably much more expensive than an optimal implementation. Rather, it carries out the operations that are essentially equivalent to evaluating the conventional sum-of-product tensor expressions, which are also automatically derived. The *strings* therefore are no longer associated with Slater determinants but are employed to address tensors such as the molecular integrals, cluster excitation amplitudes, and intermediates that possess index permutation symmetries. Kállay and Surján [39] developed an algorithm to take advantage of the complex index permutation symmetries of intermediate tensors when storing them or contracting them with other tensors (Sect. 3.4). Hence, this remarkable paper [39] introduces at least three innovative elements in coupled-cluster implementation: (1) the automatic diagrammatic derivation of coupled-cluster equations; (2) the structure of intermediate tensors; (3) the string-based mapping of tensors with index permutation symmetries. Together, they accomplish the $O\left(n^{2k+2}\right)$ algorithm that also incorporates the index permutation and other symmetries for kth order coupled-cluster method.

The string-based algorithm is complementary in many ways to the automatic derivation and implementations described in the previous section, with the primary difference being in whether the symbolic manipulations occur at run time or compile time. However, a relative merit of the string-based algorithm may be its compact program kernel that is well suited to extensions to higher orders and multireference wave functions. By virtue of this algorithm, Kállay and coworkers have rapidly implemented analytical first [90] and second [182] derivatives for coupled-cluster methods at any arbitrary high order and excited-state [74] and the state-selective multireference [176] coupled-cluster methods also at any order. These implementations, which would be unthinkable without computerized symbolic algebra, are enabling highly accurate chemical predictions [12,183–185].

5 Future prospects

With the computing power increasing steadily, a greater proportion of intellectual work and technical decisions will be delegated to computers. In quantum chemistry, this manifests itself as the automation of the development of highly complex computational methods that exceeds normal human ability. With the determinant-based algorithm, it seems entirely possible to develop an interpreter that performs numerical calculations of most any given many-body electron-correlation method specified only by its highest-level definition in terms of Slater determinants and operators. Some methods that prove the most promising can then be translated into production-level computer programs that take advantage of domain-specific optimizations (spin, spatial, index permutation symmetries, sparsity and locality, etc.) and mathematical and computer-science optimizations (strength reduction, factorization, loop fusion, data locality, parallelization, etc.) by an automated system such as the TCE or by the string-based algorithm. A new generation of quantum chemists may be able to develop new electron-correlation methods without having to know the inner workings of the symbolic algebra involved, although the contrary opinion also exists among the experts that the development always requires the in-depth knowledge of derivation and computer implementation processes.

The applications of these computational methods may also be automated to the extent that an intelligent computer system will decide what particular method and basis set (or a combination thereof) should be used for some desired chemical property of any given molecule. This is in line with the effort by Pople and co-workers [186–188] to provide an ultimate blackbox computational method (Gaussian-*n* theory) or Bartlett's [189] convergent sequence: HF $<$ MP2 $\approx$ MP3 $<$ CCD $<$ CCSD $\approx$ MP4 $<$ CCSD(T) $<$ CCSDT $<$ full CI. However, the computers may enable more flexible and on-demand adjustments of the optimal combination of methods and basis sets for each case.

Acknowledgements I thank Prof. Kimihiko Hirao for his continuous support to our work and Profs. Christopher J. Cramer and Donald G. Truhlar for the invitation to contribute to this special issue. I am indebted to Prof. Frank E. Harris, Dr. Mihály Kállay, Dr. Curtis L. Janssen, Prof. Marcel Nooijen, and Prof. Josef Paldus for a critical reading of the manuscript prior to publication. I also thank Profs. Hiroshi Nakatsuji and Josef Paldus for their assistance with literature. I am privileged to have been introduced to coupled-cluster and many-body perturbation theories by Prof. Rodney J. Bartlett. This work is sponsored by the U.S. Department of Energy, Office of Basic Energy Sciences through Grant No. DE-FG02-04ER15621.

References

1. Szabo A, Ostlund NS (1989) Modern quantum chemistry: Introduction to advanced electronic structure theory. McGraw-Hill, New York
2. McWeeny R (1992) Methods of molecular quantum mechanics. Academic Press, San Diego
3. Roos BO (1992) Lecture notes in quantum chemistry. Springer, Berlin Heidelberg New York
4. Roos BO (1994) Lecture notes in quantum chemistry II. Springer, Berlin Heidelberg New York
5. Schleyer PvR, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer III HF, Schreiner PR (1998) Encyclopedia of computational chemistry. Wiley, Chichester, New York
6. Helgaker T, Jørgensen P, Olsen J (2000) Molecular electronic-structure theory. Wiley, Chichester, New York
7. Shavitt I, Bartlett RJ (2005) Many-body methods in quantum chemistry. Cambridge University Press, Cambridge

8. Ruscic B, Feller D, Dixon DA, Peterson KA, Harding LB, Asher RL, Wagner AF (2001) J Phys Chem A 105:1
9. Ruscic B, Wagner AF, Harding LB, Asher RL, Feller D, Dixon DA, Peterson KA, Song Y, Qian X, Ng CY, Liu J, Chen W, Schwenke DW (2002) J Phys Chem A 106:2727
10. Császár AG, Leininger ML, Szalay V (2003) J Chem Phys 118:10631
11. Hirata S, Yanai T, de Jong WA, Nakajima T, Hirao K (2004) J Chem Phys 120:3297
12. Tajti A, Szalay PG, Császár AG, Kállay M, Gauss J, Valeev EF, Flowers BA, Vázquez J, Stanton JF (2004) J Chem Phys 121:11599
13. Schaefer III HF (1986) Science 231:1100
14. Bartlett RJ (1991) Theor Chim Acta 80:71
15. Paldus J (1992) In: Wilson S, Diercksen GHF (ed) Methods in computational molecular physics (NATO ASI Series B: Physics), vol 293 Plenum, New York, p 99
16. Paldus J (1994) In: Malli GL (ed) Relativistic and correlation effects in molecules and solids (NATO ASI Series B: Physics), vol 318. Plenum, New York, p 207
17. Bartlett RJ, Stanton JF (1994) In: Lipkowitz KB, Boyd DB (ed) Reviews in computational chemistry, vol 5. VCH, New York, p 65
18. Bartlett RJ (1995) In: Yarkony DR (ed) Modern electronic structure theory, Part II. World Scientific, Singapore, p 1047
19. Lee TJ, Scuseria GE (1995) In: Langhoff SR (ed) Quantum mechanical electronic structure calculations with chemical accuracy. Kluwer, Dordrecht, p 47
20. Bartlett RJ (1997) Recent advances in coupled-cluster methods. World Scientific, Singapore
21. Crawford TD, Schaefer III HF (2000) Rev Comp Chem 14:33
22. Paldus J (2003) In: Wilson S (ed) Handbook of molecular physics and quantum chemistry, vol 2. Wiley, Chichester, New York, p 272
23. Purvis III GD, Bartlett RJ (1982) J Chem Phys 76:1910
24. Scuseria GE, Scheiner AC, Lee TJ, Rice JE, Schaefer III HF (1987) J Chem Phys 86:2881
25. Noga J, Bartlett RJ (1987) J Chem Phys 86:7041
26. Scuseria GE, Schaefer III HF (1988) Chem Phys Lett 152:382
27. Oliphant N, Adamowicz L (1991) J Chem Phys 95:6645
28. Kucharski SA, Bartlett RJ (1992) J Chem Phys 97:4282
29. Wick GC (1950) Phys Rev 80:268
30. Paldus J, Čížek J (1975) Adv Quantum Chem 9:105
31. Surján PR (1989) Second quantized approach to quantum chemistry: an elementary introduction with 11 figures. Springer, Berlin Heidelberg New York
32. Harris FE, Monkhorst HJ, Freeman DL (1992) Algebraic and diagrammatic methods in many-fermion theory. Oxford University Press, New York
33. Goldstone J (1957) Proc R Soc Lon Ser-A 239:267
34. Hugenholtz NM (1957) Physica 23:481
35. Čížek J (1966) J Chem Phys 45:4256
36. Kucharski SA, Bartlett RJ (1986) Adv Quantum Chem 18:281
37. Kucharski SA, Bartlett RJ (1991) Theor Chim Acta 80:387
38. Stanton JF, Gauss J, Watts JD, Bartlett RJ (1991) J Chem Phys 94:4334
39. Kállay M, Surján PR (2001) J Chem Phys 115:2945
40. Barnett MP, Capitani JF, von zur Gathen J, Gerhard J (2004) Int J Quantum Chem 100:80
41. Brueckner KA (1955) Phys Rev 100:36
42. Paldus J, Wong HC (1973) Comput Phys Commun 6:1
43. Wong HC, Paldus J (1973) Comput Phys Commun 6:9
44. Csépes Z, Pipek J (1988) J Comput Phys 77:1
45. Herbert JM, Ermler WC (1998) Comput Chem 22:169
46. Coester F (1958) Nucl Phys 7:421
47. Coester F, Kümmel H (1960) Nucl Phys 17:477
48. Čížek J (1969) Adv Chem Phys 14:35
49. Paldus J, Čížek J, Shavitt I (1972) Phys Rev A 5:50
50. Čížek J, Paldus J (1980) Phys Scripta 21:251
51. Paldus J (2005) In: Dykstra CF, Franking G, Kim KS, Scuseria GE (ed) Theory and applications of computational chemistry: the first 40 years, Elsevier, Amsterdam
52. Bartlett RJ (2005) In: Dykstra CF, Franking G, Kim KS, Scuseria GE (ed) Theory and applications of computational chemistry: the first 40 years, Elsevier Science, Amsterdam
53. Pople JA, Krishnan R, Schlegel HB, Binkley JS (1978) Int J Quantum Chem 14:545
54. Bartlett RJ, Purvis GD (1978) Int J Quantum Chem 14:561
55. Harris FE (1999) Int J Quantum Chem 75:593
56. Berente I, Szalay PG, Gauss J (2002) J Chem Phys 117:7872
57. Bochevarov AD, Sherrill CD (2004) J Chem Phys 121:3374
58. Knowles PJ, Handy NC (1988) J Chem Phys 88:6991
59. Čížek J, Vinette F, Paldus J (1990) Int J Quantum Chem 38:831
60. Piecuch P, Toboła R, Paldus J (1996) Phys Rev A 54:1210
61. Janssen CL, Schaefer III HF (1991) Theor Chim Acta 79:1
62. Cole SJ, Purvis GD (1986) Int J Quantum Chem Symp 20:665
63. Li X, Paldus J (1994) J Chem Phys 101:8812
64. Jankowski P, Jeziorski B (1999) J Chem Phys 111:1857
65. Crawford TD, Lee TJ, Schaefer III HF (1997) J Chem Phys 107:7943
66. Lee TJ, Jayatilaka D (1993) Chem Phys Lett 201:1
67. Nooijen M, Lotrich V (2001) J Mol Struc-Theochem 547:253
68. Nooijen M, Lotrich V (2000) J Chem Phys 113:494
69. Nooijen M, Lotrich V (2000) J Chem Phys 113:4549
70. Nooijen M (2002) Int J Mol Sci 3:656
71. Nooijen M, Wladyslawski M, Hazra A (2003) J Chem Phys 118:4832
72. Hirata S (2003) J Phys Chem A 107:9887
73. Hirata S (2004) J Chem Phys 121:51
74. Kállay M, Gauss J (2004) J Chem Phys 121:9257
75. Hirata S, Fan PD, Auer AA, Nooijen M, Piecuch P (2004) J Chem Phys 121:12197
76. Head-Gordon M, Rico RJ, Oumi M, Lee TJ (1994) Chem Phys Lett 219:21
77. Hirata S (2005) J Chem Phys 122:094105
78. Post DE, Votta LG (2005) Physics Today January:35
79. Cociorva D, Baumgartner G, Lam C, Sadayappan P, Ramanujam J, Nooijen M, Bernholdt D, Harrison R (2002) Acm Sigplan Notices 37:177
80. Krishnan S, Krishnamoorthy S, Baumgartner G, Cociorva D, Lam C, Sadayappan P, Ramanujam J, Bernholdt D, Choppella V (2003) High Perform Comput - HIPC 2913:406
81. Bibireata A, Krishnan S, Baumgartner G, Cociorva D, Lam C, Sadayappan P, Ramanujam J, Bernholdt D, Choppella V (2004) Lect Notes Comput Sci 2958:93
82. Baumgartner G, Auer AA, Bernholdt DE, Bibireata A, Choppella V, Cociorva D, Gao X, Harrison RJ, Hirata S, Krishnamoorthy S, Krishnan S, Lam CC, Lu Q, Nooijen M, Pitzer RM, Ramanujam J, Sadayappan P, Sibiryakov A (2005) Proc IEEE 93:276
83. Musiał M, Kucharski SA, Bartlett RJ (2000) Chem Phys Lett 320:542
84. Musiał M, Kucharski SA, Bartlett RJ (2002) J Chem Phys 116:4382
85. Kowalski K, Piecuch P (2000) J Chem Phys 113:8490
86. Kowalski K, Piecuch P (2001) J Chem Phys 115:643
87. Kowalski K, Piecuch P (2001) Chem Phys Lett 347:237
88. Kucharski SA, Włoch M, Musiał M, Bartlett RJ (2001) J Chem Phys 115:8263
89. Head-Gordon M, Maslen PE, White CA (1998) J Chem Phys 108:616
90. Kállay M, Gauss J, Szalay PG (2003) J Chem Phys 119:2991
91. Duch W (1985) J Phys A-Math Gen 18:3283
92. Knowles PJ, Handy NC (1984) Chem Phys Lett 111:315
93. Olsen J, Roos BO, Jørgensen P, Jensen HJA (1988) J Chem Phys 89:2185
94. Knowles PJ, Handy NC (1989) Comput Phys Commun 54:75
95. Harrison RJ, Zarrabian S (1989) Chem Phys Lett 158:393
96. Zarrabian S, Sarma CR, Paldus J (1989) Chem Phys Lett 155:183
97. Olsen J, Jørgensen P, Simons J (1990) Chem Phys Lett 169:463
98. Sherrill CD, Schaefer III HF (1999) Adv Quantum Chem 34: 143

99. Knowles PJ, Somasundram K, Handy NC, Hirao K (1985) Chem Phys Lett 113:8
100. Laidig WD, Fitzgerald G, Bartlett RJ (1985) Chem Phys Lett 113:151
101. Christiansen O (2003) J Chem Phys 119:5773
102. Olsen J, Christiansen O, Koch H, Jørgensen P (1996) J Chem Phys 105:5082
103. Halkier A, Larsen H, Olsen J, Jørgensen P (1999) J Chem Phys 110:7127
104. Larsen H, Halkier A, Olsen J, Jørgensen P (2000) J Chem Phys 112:1107
105. Simon B (1982) Int J Quantum Chem 21:3
106. Silverstone HJ, Harris JG, Čížek J, Paldus J (1985) Phys Rev A 32:1965
107. Kendall RA, Dunning Jr TH, Harrison RJ (1992) J Chem Phys 96:6796
108. Christiansen O, Olsen J, Jørgensen P, Koch H, Malmqvist PÅ (1996) Chem Phys Lett 261:369
109. Sergeev AV, Goodson DZ, Wheeler SE, Allen WD (2005) J Chem Phys 123:064105
110. Leininger ML, Allen WD, Schaefer III HF, Sherrill CD (2000) J Chem Phys 112:9213
111. Hirata S, Bartlett RJ (2000) Chem Phys Lett 321:216
112. Laidig WD, Fitzgerald G, Bartlett R (1985) Chem Phys Lett 113:151
113. Schmidt C, Warken M, Handy NC (1993) Chem Phys Lett 211: 272
114. Kállay M, Surján PR (2000) J Chem Phys 113:1359
115. Olsen J (2000) J Chem Phys 113:7140
116. Christiansen O (2004) J Chem Phys 120:2140
117. Christiansen O (2004) J Chem Phys 120:2149
118. Van Voorhis T, Head-Gordon M (2000) J Chem Phys 113:8873
119. Szalay PG, Nooijen M, Bartlett RJ (1995) J Chem Phys 103:281
120. Van Voorhis T, Head-Gordon M (2001) J Chem Phys 115:5033
121. Piecuch P, Kowalski K, Fan PD, Jedziniak K (2003) Phys Rev Lett 90:113001
122. Meissner L, Hirata S, Bartlett RJ (2005) Theor Chem Acc (to appear)
123. Hirata S, Nooijen M, Bartlett RJ (2000) Chem Phys Lett 326:255
124. Hirata S, Nooijen M, Bartlett RJ (2000) Chem Phys Lett 328:459
125. Emrich K (1981) Nucl Phys A351:379
126. Emrich K (1981) Nucl Phys A351:397
127. Sekino H, Bartlett RJ (1984) Int J Quantum Chem Symp 18:255
128. Geertsen J, Rittby M, Bartlett RJ (1989) Chem Phys Lett 164:57
129. Comeau DC, Bartlett RJ (1993) Chem Phys Lett 207:414
130. Stanton JF, Bartlett RJ (1993) J Chem Phys 98:7029
131. Monkhorst HJ (1977) Int J Quantum Chem Symp 11:421
132. Mukherjee D, Mukherjee PK (1979) Chem Phys 39:325
133. Ghosh S, Mukherjee D, Bhattacharyya S (1981) Mol Phys 43: 173
134. Dalgaard E, Monkhorst HJ (1983) Phys Rev A 28:1217
135. Takahashi M, Paldus J (1986) J Chem Phys 85:1486
136. Koch H, Jørgensen P (1990) J Chem Phys 93:3333
137. Koch H, Jensen HJA, Jørgensen P, Helgaker T (1990) J Chem Phys 93:3345
138. Rico RJ, Head-Gordon M (1993) Chem Phys Lett 213:224
139. Nakatsuji H, Hirao K (1977) Chem Phys Lett 47:569
140. Nakatsuji H, Hirao K (1978) J Chem Phys 68:2053
141. Nakatsuji H (1978) Chem Phys Lett 59:362
142. Nakatsuji H (1979) Chem Phys Lett 67:329
143. Nakatsuji H (1979) Chem Phys Lett 67:334
144. Hirao K (1983) J Chem Phys 79:5000
145. Ehara M, Ishida M, Toyota K, Nakatsuji H (2003) In: Sen KD (ed) Reviews in Modern Quantum Chemistry (A Tribute to Professor Robert G. Parr), World Scientific, Singapore, p 293
146. Nakatsuji H (1991) J Chem Phys 94:6716
147. Nakatsuji H (1983) Chem Phys 75:425
148. Toyota K, Hasegawa J, Nakatsuji H (1996) Chem Phys Lett 250:437
149. Tokita Y, Hasegawa J, Nakatsuji H (1998) J Phys Chem A 102: 1843
150. Nooijen M, Bartlett RJ (1997) J Chem Phys 106:6449
151. Gwaltney SR, Bartlett RJ (1998) J Chem Phys 108:6790
152. Hasegawa J, Ozeki Y, Ohkawa K, Hada M, Nakatsuji H (1998) J Phys Chem B 102:1320
153. Miyahara T, Tokita Y, Nakatsuji H (2001) J Phys Chem B 105:7341
154. Miyahara T, Nakatsuji H, Hasegawa J, Osuka A, Aratani N, Tsuda A (2002) J Chem Phys 117:11196
155. Nakatsuji H (1985) J Chem Phys 83:713
156. Nakatsuji H (1985) J Chem Phys 83:5743
157. Nakatsuji H (1991) Chem Phys Lett 177:331
158. Hirata S, Nooijen M, Grabowski I, Bartlett RJ (2001) J Chem Phys 114:3919
159. Hirata S, Nooijen M, Grabowski I, Bartlett RJ (2001) J Chem Phys 115:3967
160. Stanton JF (1997) Chem Phys Lett 281:130
161. Urban M, Noga J, Cole SJ, Bartlett RJ (1985) J Chem Phys 83:4041
162. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M (1989) Chem Phys Lett 157:479
163. Kucharski SA, Bartlett RJ (1998) J Chem Phys 108:5243
164. Kucharski SA, Bartlett RJ (1998) J Chem Phys 108:9221
165. Gwaltney SR, Head-Gordon M (2000) Chem Phys Lett 323:21
166. Gwaltney SR, Sherrill CD, Head-Gordon M, Krylov AI (2000) J Chem Phys 113:3548
167. Gwaltney SR, Head-Gordon M (2001) J Chem Phys 115:2014
168. Gwaltney SR, Byrd EFC, Van Voorhis T, Head-Gordon M (2002) Chem Phys Lett 353:359
169. Kowalski K, Piecuch P (2000) J Chem Phys 113:5644
170. Piecuch P, Kowalski K, Pimienta ISO, McGuire MJ (2002) Int Rev Phys Chem 21:527
171. Piecuch P, Kowalski K, Pimienta ISO, Fan PD, Lodriguito M, McGuire MJ, Kucharski SA, Kus T, Musiał M (2004) Theor Chem Acc 112:349
172. Zarrabian S, Bartlett RJ (1988) Chem Phys Lett 153:133
173. Zarrabian S, Laidig WD, Bartlett RJ (1990) Phys Rev A 41:4711
174. Zarrabian S, Paldus J (1990) Int J Quantum Chem 38:761
175. Lindgren I (1978) Int J Quantum Chem Symp 12:33
176. Kállay M, Szalay PG, Surján PR (2002) J Chem Phys 117:980
177. Oliphant N, Adamowicz L (1991) J Chem Phys 94:1229
178. Oliphant N, Adamowicz L (1992) J Chem Phys 96:3739
179. Piecuch P, Oliphant N, Adamowicz L (1993) J Chem Phys 99: 1875
180. Piecuch P, Kucharski SA, Bartlett RJ (1999) J Chem Phys 110:6103
181. Lyakh DI, Ivanov VV, Adamowicz L (2005) J Chem Phys 122: 024108
182. Kállay M, Gauss J (2004) J Chem Phys 120:6841
183. Boese AD, Oren M, Atasoylu O, Martin JML, Kállay M, Gauss J (2004) J Chem Phys 120:4129
184. Chan GKL, Kállay M, Gauss J (2004) J Chem Phys 121: 6110
185. Rizzo A, Kállay M, Gauss J, Pawlowski F, Jørgensen P, Hattig C (2004) J Chem Phys 121:9461
186. Pople JA, Head-Gordon M, Fox DJ, Raghavachari K, Curtiss LA (1989) J Chem Phys 90:5622
187. Curtiss LA, Raghavachari K, Trucks GW, Pople JA (1991) J Chem Phys 94:7221
188. Curtiss LA, Raghavachari K, Redfern PC, Rassolov V, Pople JA (1998) J Chem Phys 109:7764
189. Bartlett RJ (2000) In: Keinan E, Schechter I (ed) Chemistry for the 21st century, Wiley-VCH, Weinheim