CC2 excitation energy calculations on large molecules using the resolution of the identity approximation

Christof Hättig and Florian Weigend
Universit"at Karlsruhe (TH), Institut für Physikalische Chemie, Lehrstuhl für Theoretische Chemie, Kaiserstraße 12, D-76128 Karlsruhe, Germany

(Received 10 April 2000; accepted 7 July 2000)

A new implementation of the approximate coupled cluster singles and doubles method CC2 is reported, which is suitable for large scale integral-direct calculations. It employs the resolution of the identity (RI) approximation for two-electron integrals to reduce the CPU time needed for calculation and I/O of these integrals. We use a partitioned form of the CC2 equations which eliminates the need to store double excitation cluster amplitudes. In combination with the RI approximation this formulation of the CC2 equations leads to a reduced scaling of memory and disk space requirements with the number of correlated electrons \( n \) and basis functions \( (N) \) to, respectively, \( O(N^2) \) and \( O(nN^2) \), compared to \( O(n^2N^2) \) in previous implementations. The reduced CPU, memory and disk space requirements make it possible to perform CC2 calculations with accurate basis sets on large molecules, which would not be accessible with conventional implementations of the CC2 method. We present an application to vertical excitation energies of alkenes \( C_nH_{2n+2} \), for \( n = 1–12 \), and report results for the lowest lying dipole-allowed transitions for the TZVPP basis sets, which for \( n = 12 \) contain 1108 basis functions. Comparison with conventional CC2 results for the smaller alkenes show that for CC2 ground state energies and for excitation energies of valence states, the error due to the RI approximation is negligible compared to the usual basis set error, if auxiliary basis sets are used, which have been optimized for MP2 energy calculations. © 2000 American Institute of Physics. [S0021-9606(00)31237-5]

I. INTRODUCTION

Since the first implementations of excitation energies within the framework of coupled cluster response theory were reported\(^1\) between 1984 and 1990 for the coupled cluster singles and doubles (CCSD) model, the coupled cluster response approach has gained increasing popularity and has been applied to many different properties. Today it can be considered one of the most important standard approaches for the calculation of accurate excitation energies and frequency-dependent molecular properties for small and medium sized molecules. Implementations for excitation energies have now been reported for a broad spectrum of methods covering coupled cluster singles (CCS, for excitation energies also known as configuration interaction singles, CIS, or Tamm–Dancoff approximation, TDA), the approximate coupled cluster singles and doubles method CC2, the CCSD model, and numerous methods with iterative or noniterative approximations for the contributions of connected triples (CCSDT-1a,\(^6\) CCSDT-1b,\(^8\) CCSDT-3,\(^9\) EOM-CCSD(T),\(^10\) CC3,\(^11,12\) CCSDR(3),\(^13\) CCSDR(1a),\(^13\) CCSDR(1b)\(^13\) \ldots.). With increasing complexity and computational costs, these methods give increasingly accurate results for excitation energies. In particular the sequence CCS, CC2, CCSD and CC3 has been shown to form a systematic hierarchy for excitation energy calculations with average and maximum errors decreasing approximately by a factor of 3 at each new level.\(^14\)

Recently, integral-direct implementations have been reported\(^5,16\) for some of these methods which extended in particular the applicability of the CC2 and the CCSD model to large scale calculations on molecules with up to 10–15 atoms and 400–600 basis functions, provided that molecular point group symmetry can be employed to reduce computational efforts. Applications to larger systems are hampered for the CCSD model by the steep increase of the computational costs (in terms of CPU time), which for the most expensive contributions scale as \( n^2N^4 \) and \( n^3N^3 \), where \( n \) is the number of correlated electrons and \( N \) the size of the basis set. For integral-direct CC2 calculations\(^16\) the most severe bottlenecks are, beside the computational costs for calculation \( (\approx N^4) \), I/O \( (\approx N^4) \) and transformation of the two-electron integrals \( (\approx nN^4) \), the large memory and disk space demands scaling as \( n^2N^2 \). In particular the latter demands restrict today the applicability of the implementation to systems with less than \( 4 \times 10^8 \) double excitation amplitudes (e.g., \( \approx 40 \) occupied orbitals and \( \approx 500 \) basis functions in \( C_1 \) symmetry).

A number of different approaches have been proposed in the last years to reduce the computational costs of electron correlated methods. For example the local correlation ansatz developed by Pulay and Sæbo\(^17,18\) and Schütz and Werner,\(^19\) or the Laplace transformation proposed by Almlöf and Häser\(^20–22\) for MP2 and the perturbative triples correction included in CCSD(T) or the resolution of the identity (RI) approximation\(^23,24\) implemented, e.g., for MP2\(^25–28\) and CCSD.\(^29\) Despite their unquestionable advantages all of these approaches still have some weaknesses. The local correlation methods neglect about 1%–5% of the correlation energy and...