Accurate and efficient CCSD(T)/CBS up to 1000 atoms: optimization of the local natural orbital approach and large-scale applications

Péter R. Nagy,[1] László Gyevi-Nagy[1], Bernát Szabó[1], Balázs Lőrincz[1], József Csóka[1]

[1] *Molecular Quantum Simulation Group, Budapest University Of Technology and Economics, Hungary*
e-mail: nagy.peter@vbk.bme.hu

Our recent development efforts[1,2] and large-scale applications[2,3] are presented illustrating the accuracy and efficiency of the local natural orbital (LNO) coupled-cluster (CC) family of methods. We implemented[1,2] highly-optimized and parallel closed-shell and restricted open-shell LNO-CCSD(T), general-order LNO-CC, and LNO-ADC(2) approach for excited states in the MRCC suite of quantum chemistry programs [www.mrcc.hu, 2b].

The accuracy of our local approximations can be systematically improved via a composite threshold set and we also introduced an extrapolation approach accelerating the convergence toward canonical CCSD(T). The complete basis set [CBS(X/X+1)] limit is reliably approached by extrapolating with (augmented) X-zeta basis sets to CBS(T/Q) or even CBS(Q/5). Most recently, we developed a general floating Gaussian (GFG) basis approach which can decrease the cardinal number by one step required for the same level of convergence of interaction energies. These systematically convergent hierarchies enable the formulation of robust uncertainty estimates for both the local and basis set errors[2,3].

In all existing independent[4] and our own comparisons[2,3] of popular local CCSD(T) variants, LNO-CCSD(T) provided the most accurate reaction and interaction energies, e.g., for organic thermochemistry[2], non-covalent complexes[2,3,4c], as well as for organometallic-[4a,4d] or extended π-systems [3,4b] even with minor strong correlation.

Using LNO-CCSD(T) we obtained unprecedented level of convergence for challenging intermolecular interactions[2,3] and were able to exclude the local and basis set errors as the primary source of large, up to 10±3 kcal/mol deviations between CCSD(T) and fixed-node DMC[3]. We will present new pieces to this puzzle about the disagreement of CC and DMC from the perspectives of higher-order CC, local approximation free CC, and GFG basis sets.

The enabled capabilities are demonstrated, e.g,, by very Tight LNO-CCSD(T)/CBS(Q,5) obtained for a buckyball-nanoring complex of 132 atoms[3], while Tight LNO-CCSD(T)/ CBS(T,Q) is feasible for proteins up to 1000 atoms. Embedding an LNO-CC part into a DFT environment, e.g., via the Huzinaga approach can further increase the efficiency.[2]

Since LNO-CCSD(T)/CBS(T,Q) can take about an order of magnitude more time than DF-HF/CBS(T,Q), it is an affordable choice also in computational protocols with DFT methods above rung three. This is especially useful for the accurate description of extended systems sensitive to dispersion and/or density errors of DFT methods, as demonstrated by applications even on organocatalytic, biochemical, and surface reactions.

[1] *J. Chem. Theory Comput.* **14** (2018), 4193 ; *J. Chem. Phys*: **146** (2017), 214106 ;
J. Chem. Theory Comput. **17** (2021), *860* ; *J. Chem. Theory Comput.* **17** (2021), 2886

[2] a) *J. Chem. Theory Comput.* **15** (2019), 5275 ; b) *J. Chem. Phys.* **152** (2020), 074107

[3] *Nature Communications* **12** (2021), 3927

[4] a) *J. Phys. Chem. A*, **125** (2021), 8987 ; b) *J. Chem. Theory Comput.* **16** (2020), 3641
c) *J. Phys. Chem. A*, **125** (2021), 9838 ; d) *J. Chem. Theory Comput.* **18** (2022), 883