Reduced-Scaling Double Hybrid Density Functional Theory with Rapid Basis Set Convergence through Localized Pair Natural Orbital F12

Nisha Mehta and Jan M. L. Martin*

ABSTRACT: Following earlier work [Mehta, N.; Martin, J. M. L. J. Chem. Theory Comput. 2022, 10.1021/acs.jctc.2c00426] that showed how the slow basis set convergence of the double hybrid density functional theory can be obviated by the use of F12 explicit correlation in the GLPT2 step (second order Görling-Levy perturbation theory), we demonstrate here for the very large and chemically diverse GMTKN55 benchmark suite that the CPU time scaling of this step can be reduced (asymptotically linearized) using the localized pair natural orbital (PNO-L) approximation at negligible cost in accuracy.

Despite the enormous success of density functional theory (DFT), the exact exchange-correlation functional continues to be elusive, and accurate density functional approximations (DFAs) are highly desirable in computational chemistry. On Perdew’s “Jacob’s Ladder” of DFT, the fifth rung corresponds to introducing dependence on unoccupied (virtual) orbitals; one special case of fifth-rung functionals are double hybrid DFAs. These combine semilocal exchange and correlation from DFT with nonlocal Fock exchange and GLPT2 (second-order Görling-Levy perturbation theory) nonlocal correlation contributions. Indeed, double hybrid density functionals (DHDFs) are known to be the most accurate DFAs, approaching composite wave function theory schemes such as G3 and G4 theories.  

Although the basis set convergence of double hybrids is faster than that of wave function ab initio (WFT) methods, they inherit the slow basis set convergence of MP2 (αL−3 with L being the highest angular momentum in the basis set) to some degree.  

WFT basis set convergence can be greatly accelerated using explicitly correlated methods in which some gemanal terms that explicitly depend on interelectronic distances are added to the one-particle basis set. For instance, Kutzelnigg and Morgan showed that, for a two-electron system, the singlet-coupled pair correlation energies converge as αL−7 for explicitly correlated methods, compared to αL−3 for orbital-only calculations.  

While early explicitly correlated WFT studies employed simple r12 geminals, the F12 geminal (1 − exp r12)/r12 has become the de facto standard for explicitly correlated WFT.  

The introduction of density fitting and auxiliary basis sets sped up the integral evaluation to the point that F12 calculations became practically feasible.  

Practical experience with MP2-F12 and various approximations to CCSD(T)-F12 has shown that, using basis sets specifically developed for F12 calculations, the basis set convergence is drastically faster than for conventional orbital calculations. Thus, F12 approaches have increasingly become a mainstay of high-accuracy WFT; see, for instance, refs 35–45.  

We have recently shown, for the large and chemically diverse GMTKN55 benchmark suite, that using MP2-F12 in a basis of Kohn–Sham orbitals accelerates basis set convergence of double hybrid DFT to the point that even spd basis sets like cc-pVDZ-F12 are quite close to the basis set limit, and the spdf cc-pVTZ-F12 effectively reaches it.  

Nevertheless, two problems remain. First of all, CPU times for the MP2-F12 step, which scale as the sixth power if F12 amplitudes are optimized in an orbital invariant manner, will become prohibitive for medium–large systems. Second, in some implementations (such as that in MOLPRO), F12 methods require a very
Table 1. Statistical Analysis (kcal/mol) of the Basis Set Convergence in Localized B2GP-PLYP-F12-D3(BJ) Calculations for the GMTKN55 Database and Its Categories

|                | WTMAD2 | THERMO | BARRIERS | LARGE | CONF | INTERMOL |
|----------------|--------|--------|----------|-------|------|----------|
| **Relative to the Ref 5 Reference Data** |        |        |          |       |      |          |
| VDZ-F12        | 3.080  | 0.590  | 0.327    | 0.662 | 0.616| 0.864    |
| VDZ-F12**      | 3.073  | 0.583  | 0.325    | 0.662 | 0.616| 0.867    |
| VTZ-F12        | 3.030  | 0.590  | 0.324    | 0.659 | 0.599| 0.858    |
| VTZ-F12**      | 2.961  | 0.587  | 0.322    | 0.659 | 0.599| 0.794    |
| V[D,T]Z-F12    | 3.025  | 0.588  | 0.327    | 0.655 | 0.593| 0.863    |
| V[D,T]Z-F12**  | 2.951  | 0.586  | 0.324    | 0.655 | 0.593| 0.793    |
| **Relative to the B2GP-PLYP-F12-D3(BJ)/V[T,Q]Z-F12 Basis Set Limits from ref 46** |        |        |          |       |      |          |
| VDZ-F12        | 0.642  | 0.086  | 0.052    | 0.093 | 0.135| 0.277    |
| VDZ-F12**      | 0.633  | 0.074  | 0.048    | 0.093 | 0.135| 0.284    |
| VTZ-F12        | 0.322  | 0.032  | 0.019    | 0.033 | 0.057| 0.182    |
| VTZ-F12**      | 0.233  | 0.023  | 0.017    | 0.033 | 0.057| 0.103    |
| V[D,T]Z-F12    | 0.300  | 0.033  | 0.023    | 0.033 | 0.048| 0.163    |
| V[D,T]Z-F12**  | 0.262  | 0.026  | 0.021    | 0.033 | 0.048| 0.134    |

*The VnZ-F12 (where n = D,T) basis set was used for Ne-containing systems in RG18 due to numerical problems. Only for canonical VTZ-F12, computational resource limitations forced substitution of VDZ-F12 for the UPU23 subset.

The performance metric used here, as in ref 6 and subsequent studies, e.g., refs 7–10, is the weighted total mean absolute deviation (WTMAD2) as defined in ref 6:

\[
WTMAD2 = \frac{1}{\sum_{i=1}^{N} N_i} \sum_{i=1}^{N} \sum_{j=1}^{55} \frac{56.85 \text{ kcal/mol} \cdot |\Delta E_j|}{MAD_i}
\]

where \(N\) denotes the number of systems in each test set, \(|\Delta E_j|\) is the mean absolute value of all reference energies from \(i = 1\) to \(55\), and \(MAD_i\) is the mean absolute deviation of the calculated and reference energies. We also consider the decomposition of WTMAD2 into GMTKN55’s five top-level subcategories.

All electronic structure calculations were performed using the MOLPRO2022 package on the Faculty of Chemistry’s HPC cluster “ChemFarm” at the Weizmann Institute of Science. The B2GP-PLYP-D3(BJ) simple double hybrid methods were considered as a “proof of principle”. Computational details in this work largely follow those in our previous study (ref 46). All of the KS and PNO-LMP2-F12 steps were performed using density fitting (DF), and the default PNO settings were applied throughout. We considered here the cc-pVnZ-F12 (VnZ-F12 in short) basis sets, augmented versions thereof,\(^{33}\) and cc-pVnZ-PP-F12 basis sets for the heavy p-block, where \(n = D,T\). Throughout the manuscript PNO-DHF-F12 refers to DHDF-F12 calculations with the DF-PNO-LMP2-F12. For the CABS (complementary auxiliary basis set),\(^{35}\) we used the cc-pVnZ-F12/OptRI auxiliary basis sets,\(^{36,50}\) for Coulomb-exchange (“JC”) fitting, those of Weigend;\(^{37}\) finally, for RI-MP2, the MP2FIT sets of Hättig and co-workers.\(^{38,39}\) The self-consistent-field (SCF) calculations were carried out with a convergence criterion of \(10^{-9}E_{\text{C}}\). All SCF calculations were conducted with MOLPRO’s default integration grid combinations but with gridthr tightened to \(10^{-9}\). The fixed-
amplitude “3C(FIX)” approximation was employed throughout. In both ref 46 and the present work, one subset, C60ISO (isomers of C60),60 presented insurmountable near-linear dependence problems (overlap matrix elements below 10−11) owing to the p diffuse functions and was eliminated: anyhow, it has a small weight in WTMAD2.

We performed PNO-DHDF-F12 calculations with the VnZ-F12 (where n = D,T) basis sets. For six anion-containing subsets, AH2B1, G21EA,65,66 IL16, WATER27,64,65 BH76,63,65,66,67 and BH76RC,6,63 as well as for the RG18 rare-gas clusters, we also considered aug-cc-pVnZ-F12 (n = D,T) or AvnZ-F12 for short; the combination of the latter with VnZ-F12 for the rest of the GMTKN55 suite is denoted VnZ-F12 as in ref 46.

For comparison, we also carried out conventional B2G-PLYP calculations with the commonly used Weigend-Ahlich basis sets de2-TZVPP and de2-QZVPP as well as with their diffuse-function augmented equivalents69 de2-TZVPPD and de2-QZVPPD. Additionally, we considered basis sets of the correlation consistent family:70,71 the shorthand haVnZ (heavy-averaged valence n-tuple zeta, where n = D,T,Q,S) stands here for the combination of cc-pVnZ on hydrogen,70 aug-cc-pV(n+d)Z on second-row elements,72 aug-cc-pVnZ on first-row71 and third-row73 elements, and aug-cc-pVnZ-PP on the fourth- and fifth-row p-block elements.74–77 The shorthand VnZ stands for the variant without diffuse functions of this same combo.

The largest basis set for which we were previously able to obtain fully canonical B2G-PLYP-F12-D3(B) answers, permitting a direct comparison, was VQZ-F12*, permitting extrapolation from VTZ-F12* and VQZ-F12* or for short V(T,Q)Z-F12*.

The top section of Table 1 presents the WTMAD2 for localized PNO-B2G-PLYP-F12-D3(B) for GMTKN55 and its breakdown into the five top-level categories. With the VDZ-F12 basis set, we obtained a WTMAD2 of 3.080 kcal/mol for the GMTKN55 data set, which goes down insignificantly to 3.073 kcal/mol for VDZ-F12*. Furthermore, VTZ-F12 yields a WTMAD2 of 3.030 kcal/mol, which goes down to 2.961 kcal/mol for the VTZ-F12* variant. We extrapolated VDZ-F12* and VTZ-F12* reaction energies using the two-point extrapolation formula $A + B/L^n$, where $L = \text{highest angular momentum present in the basis set}$ where $\alpha = 3.088$ for the PT2 components (Table 9 in ref 78, first row of lower pane, second column); for the KS component, we just used the highest angular momentum present in the basis set plus the CABS correction. The PNO-B2G-PLYP-F12-D3(B)/V(T,Q)Z-F12* level of theory results in WTMAD2 of 2.951 kcal/mol for the entire GMTKN55 database. The WTMAD2 results obtained with canonical B2G-PLYP-F12-D3(B) are 2.939, 2.969, and 2.993 kcal/mol, respectively, for VDZ-F12*, VTZ-F12*, and V(D,T)Z-F12* (see ref 46).

Next (Table 1, middle section), we explored the basis set convergence of localized B2G-PLYP-F12-D3(B) relative to energies calculated at the canonical B2G-PLYP-F12-D3(B)/V(T,Q)Z-F12* level of theory, which can essentially be regarded as the complete basis set limit. Relative to that, PNO-B2G-PLYP-F12-D3(B) calculations in conjunction with VDZ-F12 and VDZ-F12* provide WTMAD2CBS values of 0.642 and 0.633 kcal/mol, respectively. Increasing the basis set size to VTZ-F12 and VTZ-F12* reduces these deviations to 0.322 and 0.233 kcal/mol, respectively. PNO-B2G-PLYP-F12-D3(B) in conjunction with V(D,T)Z-F12 yields a WTMAD2CBS value of 0.300 kcal/mol. Adding diffuse functions to RG18, WATER27, BH76, BH76RC, AH2B1, G21EA, and IL16 (i.e., V(D,T)Z-F12*) slightly lowers the WTMAD2CBS value to 0.262 kcal/mol. WTMAD2CBS values obtained for canonical B2G-PLYP-F12-D3(B) with VTZ-F12*, VTZ-F12*, and V(D,T)Z-F12* basis sets are 0.467, 0.207, and 0.215 kcal/mol, respectively (see ref 46).

Another angle is offered by considering the WTMAD2 between PNO-B2G-PLYP-F12-D3(B) and canonical B2G-PLYP-F12-D3(B) energies in the same basis set. These can be found in the bottom section of Table 1. For VDZ-F12, this “$\Delta$WTMAD2$_{PNO}$” is 0.513 kcal/mol, most of which is accounted for by the intermolecular and conformational subsets (0.258 and 0.113 kcal/mol, respectively). RG18 alone contributes 0.154 kcal/mol, which together with 0.037 (HEAVY28), 0.024 (S66), and 0.018 (HAL59) accounts for nearly the entire INTERMOL component; CONFOR is mostly due to the four subsets PCONF21 (0.026), BUT14DIOL (0.025), Amino20x4 (0.024), and MCONF (0.019 kcal/mol). Essentially the same picture emerges for VDZ-F12+. For VTZ-F12, the “$\Delta$WTMAD2” is 0.333 kcal/mol, the largest two contributor sets being RG18 and
HEAVY28 at 0.094 and 0.038 kcal/mol, respectively; in the VTZ-F12* variant, the RG18 contribution drops to 0.029 kcal/mol. A similar accuracy is found for V{D,T}Z-F12 and V{D,T}Z-F12* extrapolations; in each case, RG18 is again the largest contributor.

We shall now briefly compare the computational cost of localized (denoted PNO-F12 for short in the tables) and canonical B2GP-PLYP-F12 calculations. This is perhaps best illustrated by considering timings for the linear n-alkane dimers as a function of n; we considered first (ethane)$_2$ through (n-heptane)$_2$, which make up the ADIM6$_{6,79}$ subset of GMTKN55. Then, we extended this series for n = 8 through n = 24 by manually inserting CH$_2$ groups (since we were only interested in timings, optimized structures were not deemed necessary). All calculations in this comparison are run on identical hardware, namely, Intel(R) Haswell 2.40 GHz, 16 cores, 256 GB RAM, and 3.6TB of fast striped SSDs. Detailed single-core and 16-core timings for both F12 and orbital-only calculations with various basis sets can be found in the Supporting Information; in Table 2, we report values relative to PNO-F12/VDZ-F12 = 1.00. As we have seen above, localized PNO-B2GP-PLYP-F12/VDZ-F12 yields results of comparable quality to canonical B2GP-PLYP-F12/VDZ-F12. The localized PNO-B2GP-PLYP-F12/VDZ-F12 results are comparable in cost to their canonical counterparts for n-butane dimer, but as the chain length grows, the gap opens up to the point where for (n-hexadecane)$_2$, the PNO calculation is almost 18 times faster than its canonical counterpart.

For mass storage requirements, the ratios are even more lopsided (see Table S1): the PNO-F12 calculation on the n-butane dimer requires only one-sixth the scratch space of its canonical counterpart, and for n-hexadecane, this drops to one-sixteenth. These ratios appear pretty much independent of the basis set. What this means in concrete terms: for the (n-hexadecane)$_2$ dimer with the VDZ-F12 basis set, the canonical calculation requires almost 3.5 TB of scratch space versus about 0.2 TB for the PNO calculation. On our cluster, this makes the difference between having to run the calculation on a dedicated “heavyio” node with a large local scratch SSD and being able to run it on any available general-purpose node.

Note that, for the conventional B2GP-PLYP-F12-D3(BJ)/cc-pVDZ-F12 calculation, the PT2-F12 step dominates the CPU time to such an extent that the “F12 total” line is obscured by the crimson “PT2-F12” line. KS is the time spent in the Kohn−Sham iterations; CABS(KS) denotes that for the evaluation of the CABS correction. PT2-F12 and PNO-PT2-F12 refer to the canonical and PNO perturbation theory steps, respectively; “F12 total” is the time for the entire canonical calculation, and “PNO-F12 total” refers to the entire localized calculation.

However, for larger systems, CPU time scaling of the PNO-GLPT2-F12 step in PNO-B2GP-PLYP-F12-D3(BJ) becomes even gentler. Figure 1 illustrates this for dimers of parallel chains of n-alkanes through n = 24 (tetraicosane dimer) with the chains grow longer and, for def2-{T,Q}ZVPPD, from 2 to 5 times faster. The B2GP-PLYP/haV{Q,S}Z calculations range from about 9 times longer than PNO-B2GP-PLYP-F12/VDZ-F12 for the butane dimer to about 27 times as long for (n-hexadecane)$_2$. In fact, a power law fit to the total computational time for n-heptane through n-hexadecane dimers, (C$_n$H$_{2n+2})_2$ ($n = 7$−10, 12, 14, 16), reveals $\propto n^{4.88}$ scaling (nearly the expected $\propto n^5$) for canonical B2GP-PLYP-F12-D3(BJ) but approximately $\propto n^{2.58}$ scaling for PNO-B2GP-PLYP-F12-D3(BJ).
the cc-pVDZ-F12 and cc-pVTZ-F12 basis sets. The canonical calculations were only feasible through \( n = 16 \) (for which the PT2-F12 step required all available scratch space): a power law fit of CPU times for the PT2-F12 step reveals an almost perfect \( \propto n^2 \) dependence \((R^2 = 0.9999)\) as will the total CPU time that is completely dominated by this step. For PNO-B2GP-PLYP-F12-D3(BJ)/cc-pVDZ-F12, a power law fit for \( n = 8 \) through \( n = 24 \) reveals a much gentler \( \propto n^{0.6} \) scaling \((R^2 = 0.997)\): a component breakdown of the times reveals an approximately \( \propto n^2 \) scaling for the Kohn–Sham step \((R^2 = 0.998)\) paired with a scaling for the PNO-PT2-F12 step that follows a roughly \( \propto n^{1.6} \) power law \((R^2 = 0.997)\) but hews more closely \((R^2 = 0.997)\) to a quadratic fit. Broadly speaking, the same trends are found for cc-pVTZ-F12 (see the Excel sheet in the Supporting Information).

When running in a more realistic fashion on 16 cores, we find speedups by about a factor of 8 for B2GP-PLYP-F12-D3(BJ) and of 12 for PNO-B2GP-PLYP-F12-D3(BJ). More detailed scrutiny reveals that, while in the former case, the scaling is determined by the PT2-F12 step (where I/O bandwidth limitations place a practical limit on parallelism), in the latter case PNO-PT2-F12 actually parallelizes close to ideally and the Kohn–Sham step is the parallelization efficiency-limiting factor. At the end of the day, parallelization additionally favors PNO-B2GP-PLYP-F12-D3(BJ) and makes it an even more attractive method.

Now, would we be able to realize similar gains from PNO-LMP2 in orbital-only (i.e., non-F12) calculations? We have considered this for haVTZ and haVQZ basis sets along the same n-alkane dimer series for \( n = 8–16 \), parallel on 16 cores; the timing data and their breakdown can again be found in the Supporting Information. For low \( n \), the PNO approach is actually slightly costlier, but for \( n = 16 \), we can see a reduction in total CPU time by about 25%. Needless to say, this does not even come close to the order-of-magnitude or more that can be saved in PNO-F12 vs canonical F12 double hybrids. (We note that, in the orbital-only calculations, due to the larger basis sets, the KS step accounts for the lion’s share of the total time, although this will eventually be reversed as chains grow still longer.)

The principal conclusion of this work is that double hybrid F12 calculations, which largely eliminate the slow basis set convergence of double hybrids, can be carried out without significant loss of accuracy using localized pair natural orbitals in the F12 step. Thus, CPU and mass storage requirements scale much more gently with the system size, making the method amenable also to, and promising for, larger systems.

- **ASSOCIATED CONTENT**

- **Supporting Information**

  The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcl.2c02620.

  A Microsoft Excel workbook containing statistical results of all assessed methods as well as a sample MOLPRO input file for PNO-B2GP-PLYP-F12 and 1-core and 16-core timing data for the \( n \)-alkane dimer series using different approaches and broken down by calculation step (XLSX).

- **AUTHOR INFORMATION**

  **Corresponding Author**

  Jan M. L. Martin  —  Department of Molecular Chemistry and Materials Science, Weizmann Institute of Science, Rehovot 7610001, Israel; orcid.org/0000-0002-0005-5074; Phone: +972-8-9342533; Email: gershom@weizmann.ac.il; Fax: +972-8-9343029

  **Author**

  Nisha Mehta  —  Department of Molecular Chemistry and Materials Science, Weizmann Institute of Science, Rehovot 7610001, Israel; orcid.org/0000-0001-7222-4108

  Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcl.2c02620

  **Notes**

  The authors declare no competing financial interest.

- **ACKNOWLEDGMENTS**

  Work on this paper was supported by the Israel Science Foundation (grant 1969/20), by the Minerva Foundation (grant 2020/05), and by a research grant from the Artificial Intelligence and Smart Materials Research Fund (in memory of Dr. Uriel Arnon), Israel. The authors would like to thank Golokesh Santra for helpful comments on the draft manuscript and Emmanouil Semidalas for proofreading the revised manuscript.

- **REFERENCES**

  1. Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. Phys. Rev. 1964, 136, B864.
  2. Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. Phys. Rev. 1965, 140, A1133–A1138.
  3. Perdew, J. P.; Schmidt, K. Jacob’s ladder of density functional approximations for the exchange-correlation energy. AIP Conf. Proc. 2001, 577, 1–20.
  4. Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. J. Chem. Phys. 2006, 124, 034108.
  5. Goerigk, L.; Grimme, S. Double-Hybrid Density Functionals. Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2014, 4, 576–600.
  6. Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. Phys. Chem. Chem. Phys. 2017, 19, 32184–32215.
  7. Goerigk, L.; Mehta, N. A trip to the density functional theory zoo: warnings and recommendations for the user. Aust. J. Chem. 2019, 72, 563–573.
  8. Santra, G.; Sylvetsky, N.; Martin, J. M. L. Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBE-P86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. J. Phys. Chem. A 2019, 123, 5129–5143.
  9. Mehta, N.; Casanova-Paez, M.; Goerigk, L. Semi-empirical or non-empirical double-hybrid density functionals: which are more robust? Phys. Chem. Chem. Phys. 2018, 20, 23175–23194.
  10. Martin, J. M. L.; Santra, G. Empirical double-hybrid density functional theory: A ‘third way’ in between WFT and DFT. Isr. J. Chem. 2020, 60, 787–804.
  11. Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob’s ladder: defining and testing the obB97M(2) double hybrid density functional. J. Chem. Phys. 2018, 148, 241736.
  12. Brémond, E.; Ciolfini, I.; Sancho-Garcia, J. C.; Adamo, C. Nonempirical Double-Hybrid Functionals: An Effective Tool for Chemists. Acc. Chem. Res. 2016, 49, 1503–1513.
(13) Zhang, L. Y.; Xu, X. Doubly hybrid density functional for accurate description of thermochemistry, thermochemical kinetics and nonbonded interactions. Int. Rev. Phys. Chem. 2011, 30, 115–160.
(14) Görling, A.; Levy, M. Exact Kohn-Sham scheme based on perturbation theory. Phys. Rev. A 1994, 50, 196–204.
(15) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolow, V.; Pople, J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. J. Chem. Phys. 1998, 109, 7764–7776.
(16) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 theory. J. Chem. Phys. 2007, 126, 084108.
(17) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gnu theory. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2011, 1, 810–825.
(18) Hättig, C.; Klopfer, W.; Köhn, A.; Tew, D. P. Explicitly correlated electrons in molecules. Chem. Rev. 2012, 112, 4–74.
(19) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly correlated R12/F12 methods for electronic structure. Chem. Rev. 2012, 112, 75–107.
(20) Ten-no, S.; Noga, J. Explicitly correlated electronic structure theory from R12/F12 ansätze. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2012, 2, 114–125.
(21) Kutzelnigg, W.; Morgan, J. D., III Rates of convergence of the partial-wave expansions of atomic correlation energies. J. Chem. Phys. 1992, 96, 4484–4508.
(22) Kloepper, W.; Kutzelnigg, W. Moller-Plesset calculations taking care of the correlation cusp. Chem. Phys. Lett. 1987, 134, 17–22.
(23) Kutzelnigg, W.; Klopfer, W. Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. I. General theory. J. Chem. Phys. 1991, 94, 1985–2001.
(24) Ten-no, S. Initiation of explicitly correlated second-order Møller–Plesset methods with auxiliary basis sets. J. Chem. Phys. 2002, 116, 6397–6410.
(25) Manby, F. R. Density fitting in second-order linear-r12 Møller–Plesset perturbation theory. J. Chem. Phys. 2003, 119, 4607–4613.
(26) May, A. J.; Valeev, E.; Polly, R.; Manby, F. R. Analysis of the errors in explicitly correlated electronic structure theory. Phys. Chem. Chem. Phys. 2005, 7, 2710–2713.
(27) Werner, H. J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. J. Chem. Phys. 2007, 126, 164102.
(28) Valeev, E. F. Coupled-cluster methods with perturbative inclusion of explicitly correlated terms: A preliminary investigation. Phys. Chem. Chem. Phys. 2008, 10, 106–113.
(29) Valeev, E. F.; Crawford, T. D. Simple coupled-cluster singles and doubles method with perturbative inclusion of triples and explicitly correlated geminals: The CCSD(T)R12 model. J. Chem. Phys. 2008, 128, 244113.
(30) Hättig, C.; Tew, D. P.; Köhn, A. Communications: Accurate and efficient approximations to explicitly correlated coupled-cluster singles and doubles, CCSD-F12. J. Chem. Phys. 2010, 132, 231012.
(31) Peterson, K. A.; Adler, T. B.; Werner, H. J. Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B–Ne, and Al–Ar. J. Chem. Phys. 2008, 128, 084102.
(32) Sylvestre, N.; Kesharwani, M. K.; Martin, J. M. L. The aug-cc-pVnZ-F12 basis set family: Correlation consistent basis sets for explicitly correlated benchmark calculations on anions and noncovalent complexes. J. Chem. Phys. 2017, 147, 134106.
(33) Hill, J. G.; Peterson, K. A. Correlation consistent basis sets for explicitly correlated wavefunctions: Pseudopotential-based basis sets for the post-d main group elements Ga–Rn. J. Chem. Phys. 2014, 141, 094106.
(34) Marchetti, O.; Werner, H.-J. Accurate calculations of intermolecular interaction energies using explicitly correlated coupled cluster wave functions and a dispersion-weighted MP2 method. J. Phys. Chem. A 2009, 113, 11580–11585.
(35) Peterson, K. A.; Dixon, D. A.; Stoll, H. The Use of Explicitly Correlated Methods on XeF2 Predicts a C3e Minimum with a Sterically Active, Free Valence Electron Pair on Xe. J. Phys. Chem. A 2012, 116, 9777–9782.
(36) Pavošević, F.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. SparseMaps—A systematic infrastructure for reduced-scaling electronic structure methods. IV. Linear-scaling second-order explicitly correlated energy with pair natural orbitals. J. Chem. Phys. 2016, 144, 144109.
(37) Klopfer, W.; Samson, C. C. M. Explicitly correlated second-order Møller–Plesset perturbation theory. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2011, 1, 810–825.
(38) Hättig, C.; Klopfer, W.; Köhn, A.; Tew, D. P. Explicitly correlated electrons in molecules. Chem. Rev. 2012, 112, 4–74.
(39) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly correlated R12/F12 methods for electronic structure. Chem. Rev. 2012, 112, 75–107.
(40) Fougier, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. The melatonin conformer space: benchmark and assessment of wave function and DFT methods for a paradigmatic biological and pharmacological molecule. J. Phys. Chem. A 2013, 117, 2269–77.
(41) Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. L. Toward a W4-F12 approach: Can explicitly correlated and orbital-based ab initio CCSD(T) limits be reconciled? J. Chem. Phys. 2016, 144, 214101.
(42) Brauer, B.; Kesharwani, M. K.; Martin, J. M. L. Some observations on counterpoise corrections for explicitly correlated calculations on noncovalent interactions. J. Chem. Theory Comput. 2014, 10, 3791–3799.
(43) Curtiss, L. A.; Raghavachari, K. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. J. Chem. Phys. 1992, 96, 7764–7776.
(44) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 theory. J. Chem. Phys. 2007, 126, 084108.
(45) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gnu theory. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2011, 1, 810–825.
(46) Hättig, C.; Klopfer, W.; Köhn, A.; Tew, D. P. Explicitly correlated R12/F12 methods for electronic structure. Chem. Rev. 2012, 112, 75–107.
(47) Manby, F. R. Density fitting in second-order linear-r12 Møller–Plesset perturbation theory. J. Chem. Phys. 2003, 119, 4607–4613.
(48) May, A. J.; Valeev, E.; Polly, R.; Manby, F. R. Analysis of the errors in explicitly correlated electronic structure theory. Phys. Chem. Chem. Phys. 2005, 7, 2710–2713.
(49) Werner, H. J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. J. Chem. Phys. 2007, 126, 164102.
(50) Werner, H.-J.; Doll, K.; Heßelmann, A.; Kats, D.; Köhn, A.; Korona, T.; Kreplin, D. A.; et al. The Molpro quantum chemistry package. J. Chem. Theory Comput. 2017, 13, 3136–3152.
(51) Nagy, P. R.; Kállay, M. Approaching the Basis Set Limit of CCSD(T) Energies for Large Molecules with Local Natural Orbital Coupled-Cluster Methods. J. Chem. Theory Comput. 2019, 15, 5275–5298.
(52) Ma, Q.; Werner, H.-J. Scalable Electron Correlation Methods. 2. Parallel PNO-LMP2-F12 with Near Linear Scaling in the Molecular Size. J. Chem. Theory Comput. 2015, 11, 5291–5304.
(53) Karton, A.; Tarnopolsky, A.; Lamère, J.-F.; Schatz, G. C.; Martin, J. M. L. Highly accurate first-principles benchmark data sets for the parametrization and validation of density functional and other approximate methods. Derivation of a robust, generally applicable, double-hybrid functional for thermochemistry and thermochemical kinetics. J. Phys. Chem. A 2008, 112, 12868–12886.
(54) Goerigk, L.; Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. Phys. Chem. Chem. Phys. 2011, 13, 6670–6688.
(55) Valeev, E. F. Improving on the resolution of the identity in linear R12 ab initio theories. Chem. Phys. Lett. 2004, 395, 190–195.
(56) Yousaf, K. E.; Peterson, K. A. Optimized complementary auxiliary basis sets for explicitly correlated methods: aug-cc-pVnZ orbital basis sets. Chem. Phys. Lett. 2009, 476, 303–307.
(57) Weigend, F. A fully direct RI-HF algorithm: Implementation, optimized auxiliary basis sets, demonstration of accuracy and efficiency. Phys. Chem. Chem. Phys. 2002, 4, 4285–4291.
(58) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. J. Chem. Phys. 2002, 116, 3175–3183.
(59) Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core—valence and quintuple-$\zeta$ basis sets for H to Ar and QZVPP basis sets for Li to Kr. Phys. Chem. Chem. Phys. 2005, 7, 59–66.
(60) Sure, R.; Hansen, A.; Schwerdtfeger, P.; Grimme, S. Comprehensive theoretical study of all 1812 C$_{60}$ isomers. Phys. Chem. Chem. Phys. 2017, 19, 14296–14305.
(61) Lao, K. U.; Schäffer, R.; Jansen, G.; Herbert, J. M. Accurate Description of Intramolecular Interactions Using Symmetry-Adapted Perturbation Theory. J. Chem. Theory Comput. 2015, 11, 2473–2486.
(62) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 theory for molecular energies of first-and second-row compounds. J. Chem. Phys. 1991, 94, 7221–7230.
(63) Goerigk, L.; Grimme, S. A General Database for Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. Assessment of Common and Reparameterized (Meta)-GGA Density Functionals. J. Chem. Theory Comput. 2010, 6, 107–126.
(64) Bryantsev, V.; Diallo, M.; van Duin, A.; Goddard, W. Evaluation of B3LYP, X3LYP, and M06-Class Density Functionals for Predicting the Binding Energies of Neutral, Protonated, and Deprotonated Water Clusters. J. Chem. Theory Comput. 2009, 5, 1016–1026.
(65) Manna, D.; Kesharwani, M. K.; Sylvestsky, N.; Martin, J. M. L. Conventional andExplicitly Correlated ab Initio Benchmark Study on Water Clusters: Revision of the BEGDB and WATER27 Data Sets. J. Chem. Theory Comput. 2017, 13, 3136–3152.
(66) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. Phys. Chem. Chem. Phys. 2005, 7, 43–52.
(67) Friedrich, J.; Hanchen, J. Incremental CCSD(T) (F12$^*$) MP2: A Black Box Method To Obtain Highly Accurate Reaction Energies. J. Chem. Theory Comput. 2013, 9, 5381–5394.
(68) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. Phys. Chem. Chem. Phys. 2005, 7, 3297–3305.
(69) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. J. Chem. Phys. 2010, 133, 134105.
(70) Dunning, T. H., Jr Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys. 1989, 90, 1007–1023.
(71) Kendall, R. A.; Dunning, T. H., Jr; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. J. Chem. Phys. 1992, 96, 6796–6806.
(72) Dunning, T. H., Jr; Peterson, K. A.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. J. Chem. Phys. 2001, 114, 9244–9253.
(73) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IX. The atoms gallium through krypton. J. Chem. Phys. 1999, 110, 7667–7676.
(74) Metz, B.; Stoll, H.; Dolg, M. Small-core multiconfiguration-Diarc—Hartree—Fock-adjusted pseudopotentials for post-d main group elements: Application to PbH and PbO. J. Chem. Phys. 2000, 113, 2563–2569.
(75) Peterson, K. A. Systematically convergent basis sets with relativistic pseudopotentials. I. Correlation consistent basis sets for the post-d group 13–15 elements. J. Chem. Phys. 2003, 119, 11099–11112.
(76) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16–18 elements. J. Chem. Phys. 2003, 119, 11113–11123.
(77) Peterson, K. A.; Shepler, B. C.; Figgen, D.; Stoll, H. On the Spectroscopic and Thermochemical Properties of ClO, BrO, IO, and Their Anions. J. Phys. Chem. A 2006, 110, 13877–13883.
(78) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. Extrapolating MP2 and CCSD explicitly correlated correlation energies to the complete basis set limit with first and second row correlation consistent basis sets. J. Chem. Phys. 2009, 131, 194105.
(79) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys. 2010, 132, 154104.