

Explicitly Correlated Double-Hybrid DFT: A Comprehensive Analysis of the Basis Set Convergence on the GMTKN55 Database

Nisha Mehta and Jan M. L. Martin*



Cite This: <https://doi.org/10.1021/acs.jctc.2c00426>



Read Online

ACCESS |



Metrics & More

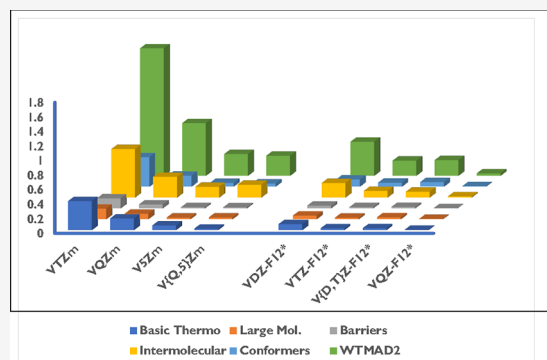


Article Recommendations



Supporting Information

ABSTRACT: Double-hybrid density functional theory (DHDFT) offers a pathway to accuracy approaching composite wavefunction approaches such as G4 theory. However, the Görling–Levy second-order perturbation theory (GLPT2) term causes them to partially inherit the slow αL^{-3} (with L the maximum angular momentum) basis set convergence of correlated wavefunction methods. This could potentially be remedied by introducing F12 explicit correlation: we investigate the basis set convergence of both DHDFT and DHDFT-F12 (where GLPT2 is replaced by GLPT2-F12) for the large and chemically diverse general main-group thermochemistry, kinetics, and noncovalent interactions (GMTKN55) benchmark suite. The B2GP-PLYP-D3(BJ) and revDSD-PBEP86-D4 DHDFs are investigated as test cases, together with orbital basis sets as large as aug-cc-pV5Z and F12 basis sets as large as cc-pVQZ-F12. We show that F12 greatly accelerates basis set convergence of DHDFs, to the point that even the modest cc-pVDZ-F12 basis set is closer to the basis set limit than cc-pV(Q+d)Z or def2-QZVPPD in orbital-based approaches, and in fact comparable in quality to cc-pV(5+d)Z. Somewhat surprisingly, aug-cc-pVDZ-F12 is not required even for the anionic subsets. In conclusion, DHDF-F12/VDZ-F12 eliminates concerns about basis set convergence in both the development and applications of double-hybrid functionals. Mass storage and I/O bottlenecks for larger systems can be circumvented by localized pair natural orbital approximations, which also exhibit much gentler system size scaling.



1. INTRODUCTION

The two most common methodologies in computational chemistry are wavefunction ab initio methods¹ and density functional theory (DFT).^{2,3} Although (correlated) wavefunction ab initio methods provide a clear road map for the convergence to the exact solution, they suffer from slow basis set convergence and hence they are only practical for small molecules. The alternative solution to the quantum many problems is given by DFT, thanks to Hohenberg–Kohn² and Kohn–Sham theorems,³ DFT currently provides the best cost-accuracy ratio for main-group thermochemistry, kinetics, and noncovalent interactions. Among various density functional theory approximations, double-hybrid density functionals (DHDFs) stand out for their general applicability, reliability, and robustness.^{4–13} In DHDFs, a portion of semilocal DFT exchange and correlation are replaced by non-local Fock exchange and second-order Görling–Levy perturbation theory¹⁴ (GLPT2) type correlation contributions, respectively. (An earlier usage^{15,16} of the term “double hybrid” referred to the combination of semilocal DFT for short-range correlation with regular MP2 correlation in a HF orbital basis for long-range correlation; see also the work of the late Angyán¹⁷ on range-separated correlation. For a detailed numerical analysis of the benefits of GLPT2 over HF-MP2 correlation, see ref 18). DHDFs offer^{8,19} a level of agreement approaching

composite wavefunction theory schemes such as G3 and G4 theories.^{20–22}

Hybrid DFT functionals (rung four on “Jacob’s Ladder”²³) exhibit basis set convergence resembling that of Hartree–Fock theory. Double hybrids (rung five on “Jacob’s Ladder”) contain a GLPT2 part, the basis set convergence of which is similar to the well-known asymptotic αL^{-3} (with L the highest angular momentum in the basis set) behavior of MP2²⁴ and of electron correlation methods more broadly.²⁵

Thus, double hybrids inherit the slow basis set convergence of MP2, although the problem is not as severe as in MP2 itself owing to the scale factors of the GLPT2 correlation (e.g., 0.25 for B2PLYP,⁴ 0.36 for B2GP-PLYP²⁶). Additionally, the computational cost can be greatly mitigated by introducing density fitting in the MP2 part,^{27,28} and two-point basis set extrapolation (e.g., refs 29–31 and references therein) can be applied.

Received: April 27, 2022

The greatest stumbling block for basis set convergence in MP2 and GLPT2 alike is the need to model the interelectronic correlation cusp, which explicitly depends on r_{12} , in terms of products of orbitals in r_1 and r_2 . In explicitly correlated approaches (see refs 32–34 for reviews), functions of r_{12} (so-called geminals) are added to the calculation to ensure that the cusp is well-described at short range, “freeing up” the orbital basis set, as it were, to cover other correlation effects.

Kutzelnigg and Morgan²⁵ showed that for two-electron model systems, the singlet-coupled pair correlation energy converges as αL^{-7} , compared to αL^{-3} for pure orbital calculations.

Initial studies (e.g., refs 35 and 36) featured a simple R12 geminal. In the last decade and a half, the F12 geminal³⁷ ($1 - \exp \gamma r_{12}$)/ γ has become the de facto standard. Meanwhile, the computational cost barrier resulting from the need for three- and four-electron integrals^{38–40} was circumvented through the introduction of auxiliary basis sets and density fitting.^{41–43}

Meanwhile, MP2-F12 and various approximations to CCSD(T)-F12 have become a mainstream tool in high-accuracy wavefunction methods: see, for example, refs 44–46 and from the Weizmann group, refs 47–49 in small-molecule thermochemistry, and refs 50–53 in noncovalent interactions.

It stands to reason that MP2-F12 in a basis of Kohn–Sham orbitals might be a way through the basis set convergence bottleneck of double-hybrid DFT. Karton and Martin⁵⁴ showed that this might be the case for a rather small set of closed-shell reactions, but to our knowledge, this has never been verified for a large and chemically diverse benchmark suite such as GMTKN55 (general main-group thermochemistry, kinetics, and noncovalent interactions,⁶ 55 problem types) or the Head-Gordon group’s even larger main-group chemistry database (MGCDB84⁵⁵). GMTKN55 has previously been used for both evaluation and parametrization of DHDFs as well as composite wavefunction methods.^{6,8–10,56–58}

We will show below that for DHDFs applied to GMTKN55, F12 accelerates basis set convergence to the point that even spd basis sets are quite close to the complete basis set limit, and that spdf basis sets effectively reach it.

2. COMPUTATIONAL DETAILS

We assess the basis set convergence of conventional and explicitly correlated DHDFs using the GMTKN55 database for general main-group thermochemistry, kinetics, and noncovalent interactions. GMTKN55 consists of 2462 total single point calculations, which are distributed over 55 subsets. The latter are divided into five categories. The first category (basic properties and reaction energies of small systems) addresses problems associated with reaction energies for small systems, total atomization energies, ionization potentials, electron affinities, and self-interaction error. The second category covers problems related to reaction energies of large systems and isomerization. The third category consists of barrier height-related problems. Intermolecular noncovalent interactions related problems are covered in the third category, while conformer equilibria (driven by intramolecular noncovalent interactions) make up the fourth category. The respective abbreviations for the five categories are “Thermo”, “Large”, “Barrier”, “Intermol”, and “Conf”. Table 1 provides a summary of GMTKN55.

We used the weighted total mean absolute deviation, type 2 (WTMAD2)—originally defined in eq 2 of ref 6—as our primary metric.

Table 1. Overview of the GMTKN55 Database and Its Five Categories: Basic Properties and Reactions of Small Systems (“Thermo”), Reaction Energies of Larger Systems and Isomerization (“Large”), Barrier Heights (“Barrier”), Intermolecular Noncovalent Interactions (“Intermol”), and Intramolecular Noncovalent Interactions (“Conf”)^a

Category	names of constituent benchmark sets	references
Thermo	W4-11, G21EA, G21IP, DIPCS10, PA26, SIE4x4, ALKBDE10, YBDE18, AL2X6, HEAVYSB11, NBPRC, ALK8, RC21, G2RC, BH76RC, FH51, TAUT15, DC13	6,59–82
Large	MB16-43, DARC, RSE43, BSR36, CDIE20, ISO34, ISOL24, C60ISO, PAreI	6,65,83–89
Barrier	BH76, BHPERI, BHDIV10, INV24, BHROT27, PX13, WCPT18	6,68,69,82,90–96
Intermol	RG18, ADIM6, S22, S66, HEAVY28, WATER27, CARBHB12, PNICO23, HALS9, AHB21, CHB6, IL16	97–106
Conf	IDISP, ICONF, ACONF, AMINO20x4, PCONF21, MCONF, SCONF, UPU23, BUT14DIOL	6,67,82,87,107–116

^aFor more details, see ref 6.

$$\text{WTMAD2} = \frac{1}{\sum_{i=1}^{55} N_i} \sum_{i=1}^{55} N_i \frac{56.85 \text{ kcal/mol}}{|\overline{\Delta E}_i|} \text{MAD}_i \quad (1)$$

where N_i represents the number of systems in each subset, $|\overline{\Delta E}_i|$ is the mean absolute value of all the reference energies for $i = 1$ to 55, and MAD_i is the mean absolute deviations of the calculated and reference energies for each subset of GMTKN55.

All electronic structure calculations were performed using the MOLPRO2022 package¹¹⁷ on the ChemFarm HPC cluster of the Faculty of Chemistry at the Weizmann Institute of Science. The B2GP-PLYP²⁶-D3(BJ)¹¹⁸ and revDSD-PBEP86-D4⁸ double hybrids were investigated as test cases. The dispersion model for B2GP-PLYP considered here was DFT-D3 of Grimme et al.⁹⁷ with the Becke–Johnson damping function.¹¹⁹ We used the B2GP-PLYP-D3(BJ)²⁶ dispersion parametrization $s_6 = 0.560$, $s_8 = 0.2597$, $a_1 = 0.000$, and $a_2 = 6.3332$ from ref 118. For revDSD-PBEP86,⁸ we used the DFT-D4 dispersion correction of Grimme et al.^{120,121} with the parameters $s_6 = 0.5132$, $s_8 = 0.000$, $a_1 = 0.4400$, $a_2 = 3.60$, and $s_9 = 0.5132$ from ref 8. As per the DFT-D4 defaults, we used electronegativity equalization¹²² partial charges and the 3-body Axilrod–Teller–Muto correction term. DFT-D3 and DFT-D4 type dispersion corrections were obtained with the respective standalone programs by Grimme and co-workers.^{123,124}

Whenever possible, all of the KS, MP2, and MP2-F12 steps were carried out with density fitting (DF-KS, DF-MP2, and DF-MP2-F12 approximations). We employed the OptRI auxiliary basis set¹²⁵ within the complementary auxiliary basis set approach,¹²⁶ the JKFIT basis sets of Weigend¹²⁷ for the DF-KS calculations, and the MP2FIT set of Hättig and co-workers^{128,129} for the DF-MP2/DF-MP2-F12 steps. Throughout the manuscript, DHDF-F12 refers to the double-hybrid calculations with the MP2-F12 (or DF-MP2-F12) method, whereas DHDF refers to the orbital-only (i.e., non-F12) double-hybrid calculations. In all of the DHDF-F12 calculations, the default fixed-amplitude “3C(FIX)” approximation was employed. All self-consistent-field energies were corrected with the complementary auxiliary basis set (CABS) singles correction. Energy convergence criteria for the KS calculations

Table 2. Statistical Analysis of the Basis Set Convergence in Conventional and Explicitly Correlated B2GP-PLYP-D3(BJ) Calculations for the GMTKN55 Database and Its Categories, Relative to the Reference 6 Reference Data^a

	B2GP-PLYP-D3(BJ)						B2GP-PLYP-F12-D3(BJ)					
	WTMAD2	THERMO	BARRIERS	LARGE	CONF	INTERMOL	WTMAD2	THERMO	BARRIERS	LARGE	CONF	INTERMOL
VDZ	11.904	2.205	0.964	1.049	4.160	3.526	3.011	0.581	0.333	0.680	0.623	0.793
VDZ*	9.661	1.323	0.627	1.049	4.160	2.503	2.953	0.585	0.334	0.660	0.619	0.756
VDZ ^m	6.332	1.323	0.627	1.002	1.498	1.883	2.939	0.580	0.334	0.660	0.619	0.747
VTZ	5.649	1.062	0.553	0.698	1.405	1.930	2.979	0.584	0.331	0.652	0.587	0.825
VTZ*	4.495	0.646	0.389	0.698	1.405	1.356	2.969	0.582	0.331	0.652	0.587	0.817
VTZ ^m	3.427	0.646	0.389	0.694	0.634	1.064	3.005	0.582	0.334	0.645	0.585	0.860
VQZ	3.978	0.761	0.445	0.639	0.760	1.374	2.993	0.581	0.333	0.645	0.585	0.849
VQZ*	3.417	0.558	0.348	0.639	0.760	1.113	3.007	0.591	0.330	0.667	0.585	0.833
VQZ ^m	3.131	0.558	0.348	0.646	0.590	0.990	3.004	0.589	0.327	0.666	0.584	0.838
V{T,Q}Z	3.955	0.738	0.448	0.625	0.673	1.472	3.015	0.592	0.330	0.669	0.585	0.839
V{T,Q}Z*	3.521	0.593	0.353	0.625	0.673	1.277	3.016	0.591	0.327	0.668	0.583	0.847
V{T,Q}Z ^m	3.351	0.593	0.353	0.629	0.597	1.179						
VSZ*	3.054	0.573	0.328	0.660	0.609	0.885						
VSZ ^m	3.020	0.573	0.328	0.661	0.584	0.874						
V{Q,S}Z*	3.105	0.589	0.326	0.668	0.593	0.930						
V{Q,S}Z ^m	3.115	0.589	0.326	0.666	0.597	0.937						
def2-TZVPP	3.966	0.834	0.443	0.633	0.890	1.166						
def2-TZVPP*	3.412	0.660	0.342	0.633	0.890	0.888						
def2-TZVPP ^m	3.162	0.660	0.342	0.639	0.689	0.833						
def2-TZVPPD	3.157	0.657	0.333	0.582	0.685	0.900						
def2-QZVPP	3.267	0.653	0.364	0.643	0.624	0.984						
def2-QZVPP*	3.007	0.591	0.322	0.643	0.624	0.828						
def2-QZVPP ^m	2.953	0.591	0.322	0.645	0.592	0.803						
def2-QZVPPD	2.965	0.595	0.318	0.630	0.584	0.837						
def2-{T,Q}ZVPP	3.326	0.651	0.371	0.657	0.598	1.050						
def2-{T,Q}ZVPP*	3.177	0.623	0.329	0.657	0.598	0.969						
def2-{T,Q}ZVPP ^m	3.187	0.623	0.329	0.657	0.612	0.965						
def2-{T,Q}ZVPPD	3.111	0.625	0.321	0.670	0.600	0.894						
VDZ-F12	5.883	0.880	0.323	0.916	1.539	2.225						

^aVnZ*: AVnZ was employed for RG18, AHB21, G21EA, IL16, WATER27, BH76, and BH76RC. In the "VnZ^m" variant, we additionally treated the BUT14DIOL, S22, S66, SCONEF, PNICO23, PCONF21, PArel, MCONF, and AMINO20x4 test sets with the hAVnZ basis set. VnZ-F12*: AVnZ-F12 was employed for RG18, AHB21, G21EA, IL16, WATER27, BH76, and BH76RC. In the "VnZ-F12^m" variant, we additionally treated the BUT14DIOL, S22, S66, SCONEF, PNICO23, PCONF21, PArel, MCONF, and AMINO20x4 test sets with the hAVnZ-F12 basis set.

were set to $10^{-9}E_h$ throughout, with MOLPRO's default integration grids for this accuracy and the basis set at hand.

We considered different families of basis sets. The first category are the correlation consistent basis sets of Dunning,^{130–132} which were developed with orbital-based correlated wavefunction calculations in mind (optimized for CISD valence correlation energies of atoms). The notation VnZ, in this paper, is shorthand for the combination of regular cc-pVnZ on first-row elements, cc-pV(n+d)Z on second-row elements, and cc-pVnZ-PP for the heavy p-block elements, where PP stands for pseudopotential. Finally, we employed ad hoc modifications: for RG18⁶ and the anion-containing subsets AHB21,¹⁰⁶ G21EA,^{60,82} IL16,¹⁰⁶ WATER27,^{101,102} BH76,^{6,68,70,82} and BH76RC,^{6,82} we employed aug-cc-pVnZ (“VnZ*”). In the “VnZ^m” variant, we additionally treated the BUT14DIOL,^{6,116} S22,^{98,99} S66,¹⁰⁰ SCONF,^{6,82,114} PNICO23,^{6,103} PCONF21,^{6,111,112} PArel,⁶ MCONF,^{6,133} and AMINO20x4^{6,110} test sets with the hAVnZ basis set (cc-pVnZ on hydrogen, aug-cc-pVnZ on first-row elements, aug-cc-pV(n+d)Z on second-row elements, and aug-cc-pVnZ-PP for the heavy p-block elements).

The second class of basis sets considered are the cc-pVnZ-F12 (abbreviated VnZ-F12 in this manuscript) of Peterson and co-workers,¹³⁴ or their anionic-friendly variants aug-cc-pVnZ-F12 (AVnZ-F12).¹³⁵ These basis sets were explicitly developed with F12 calculations in mind. In fact, non-F12 basis sets in explicitly correlated calculations lead to non-monotonous convergence because of elevated and erratic basis set superposition errors (BSSEs).⁴⁹ VnZ-F12* indicates that the VnZ-F12 basis set was used for all subsets of GMTKN55 except WATER27, IL16, G21EA, BH76, BH76RC, AHB21, and RG18, where we used AVnZ-F12. Again, we employed cc-pVnZ-F12-PP for the heavy p-block elements. The geminal Slater exponent (β) values of 0.9, 1.0, and 1.0 were used for the (A-)VDZ-F12, (A-)VTZ-F12, and (A-)VQZ-F12, respectively.

Finally, we also considered the Weigend–Ahlich/Karlsruhe def2 family,¹³⁶ namely def2-TZVPP and def2-QZVPP, and their diffuse function-augmented variants def2-TZVPPD and def2-QZVPPD.¹³⁷ def2-nZVPP* and def2-nZVPP^m variants are defined analogously to the above.

The geometries, charge/multiplicity information and reference energies were obtained from ref 6 and used verbatim throughout. The most computationally demanding subset isomerization energies of fullerene C₆₀ structures (C60ISO)⁸⁹ might just barely have been feasible with the VDZ-F12 basis set with available computational resources, but near-singularity in the overlap matrix (smallest eigenvalue 3×10^{-11}) effectively made the KS calculations impossible to converge. This subset's omission does not significantly affect WTMAD2 because of its small weight in the WTMAD2 formula. For explicitly correlated DHDF calculations on the UPU23 subset,¹¹⁵ we settled for (A-)VDZ-F12 basis to reduce computational cost.

3. RESULTS AND DISCUSSION

3.1. Basis Set Convergence for Conventional Double Hybrids. Let us first consider the basis set convergence with the orbital basis sets in conventional double-hybrid calculations, that is, B2GP-PLYP-D3(BJ)/VnZ, where $n = D, T, Q,$ and 5 (Table 2). The PT2 component slows down basis set convergence, albeit mitigated (compared to MP2 in a Hartree–Fock basis set) by the PT2 coefficients in the double hybrid (typically in the 0.1–0.5 range). Although DHDFs

converge faster than ab initio methods, their PT2 part acquires a slower basis set convergence. The VDZ basis set yields an unacceptably large WTMAD2 of 11.904 kcal/mol for the entire GMTKN55 database. This goes down to 9.661 kcal/mol when substituting the AVDZ basis set for the rare gas clusters RG18 and the six anion-containing subsets WATER27, BH76, BH76RC, AHB21, G21EA, and IL16. A further reduction to 6.332 kcal/mol was achieved for the VDZ^m variant, where the haVDZ basis set additionally was applied to BUT14DIOL, S22, S66, SCONF, PNICO23, PCONF21, PArel, MCONF, and AMINO20x4. Therefore, we will mostly discuss our statistics of conventional double-hybrid calculations with the VnZ^m variant. The VTZ^m basis set nearly halves WTMAD2 to 3.427 kcal/mol. In order to surpass this level of accuracy, VQZ^m has to be employed, yielding a WTMAD2 of 3.131 kcal/mol. For still better basis set convergence, we employed VSZ^m, which slightly further lowers WTMAD2 to 3.020 kcal/mol. As the orbital-only B2GP-PLYP-D3(BJ) complete basis set limit estimate, we extrapolate VQZ^m and VSZ^m reaction energies using the two-point extrapolation formula ($A + B/L^\alpha$, $L =$ highest angular momentum present in the basis set) where $\alpha = 8.7042$ for KS and 2.7399 for PT2 (as recommended in refs 138 and 139) components, respectively. The B2GP-PLYP-D3(BJ)/V{Q,S}Z^m level of theory results in a WTMAD2 of 3.115 kcal/mol for the entire GMTKN55 database. B2GP-PLYP-D3(BJ)/V{T,Q}Z^m ($\alpha = 7.6070$ for KS and 2.5313 for PT2) yields WTMAD2 of 3.351 kcal/mol.

A breakdown into the five top-level subdivisions of GMTKN55 (Table 2) showed that all five of them smoothly approach the basis set limit at the B2GP-PLYP-D3(BJ)/V{Q,S}Z^m level. More detailed scrutiny of the individual subsets revealed that HEAVY28^{6,97} is the major contributor to the difference between VSZ^m and V{Q,S}Z^m, with Δ WTMAD2 increased by 0.048 kcal/mol. Because of the way HEAVY28, RG18, and HAL59^{6,104,105} are weighted in WTMAD2, a small change in those subsets has an outsize contribution.

3.2. Effect of Introducing F12 Terms. Next, we investigate the basis set convergence in the explicitly correlated double-hybrid calculations. These calculations need to be done with the cc-pVnZ-F12 basis sets of Peterson et al.¹³⁴ Table 2 presents a statistical analysis of B2GP-PLYP-F12-D3(BJ) calculations. The B2GP-PLYP-F12-D3(BJ)/VDZ-F12 level of theory results in a WTMAD2 of only 2.953 kcal/mol. We would like to emphasize that this is *not* just a matter of the basis set: for illustration, we also evaluated orbital-only B2GP-PLYP-D3(BJ)/VDZ-F12 and found that WTMAD2 shot up to 5.883 kcal/mol. The difference is entirely owing to the presence versus absence of geminal F12 terms in the GLPT2 evaluation. Somewhat surprisingly, WTMAD2 with VDZ-F12* basis (AVDZ-F12 basis for the rare gas clusters RG18 and six anion containing subsets WATER27, BH76, BH76RC, AHB21, G21EA, and IL16) only was reduced to 2.939 kcal/mol, indicating that not even for anionic subsets is AVDZ-F12 required. (We do note that, unlike the VDZ orbital basis set, the VDZ-F12 already includes one diffuse function each of s and p symmetries for p-block elements, and one diffuse s function for hydrogen.) In explicitly correlated B2GP-PLYP-F12-D3(BJ), the energy differences that make up the GMTKN55 benchmark converge markedly, one might even say dramatically, faster with respect to the basis set size. For example, VDZ-F12*, VTZ-F12*, and VQZ-F12* provide WTMAD2 which are 2.939, 2.969, and 3.004 kcal/mol

Table 3. A Comparison of Total WTMAD2 of GMTKN55 Data Set (i.e., WTMAD2 (all)) and WTMAD2 after Excluding RG18, HEAVY28, and HAL59 from the Statistics (i.e., WTMAD2 (mod))^a

	WTMAD2 (all)	WTMAD2 (mod.)	WTMAD2 (all)	WTMAD2 (mod.)
	GMTKN55 as reference		CBS limit as reference	
		B2GP-PLYP-F12-D3(BJ)		
AVDZ-F12	3.011	2.706	0.418	0.320
VDZ-F12	2.953	2.686	0.499	0.391
VDZ-F12*	2.939	2.671	0.467	0.358
VTZ-F12	2.979	2.632	0.220	0.191
VTZ-F12*	2.969	2.626	0.207	0.172
V{D,T}Z-F12	3.005	2.625	0.232	0.205
V{D,T}Z-F12*	2.993	2.621	0.215	0.191
VQZ-F12	3.007	2.657	0.065	0.042
VQZ-F12*	3.004	2.649	0.032	0.024
V{T,Q}Z-F12	3.015	2.662	0.038	0.018
V{T,Q}Z-F12*	3.016	2.653	0.002	0.001
		B2GP-PLYP-D3(BJ)		
VDZ	11.904	11.295	11.303	10.825
VDZ*	9.661	9.120	9.014	8.582
VDZ ^m	6.332	5.540	5.602	4.911
VTZ	5.649	5.185	4.317	3.942
VTZ*	4.495	4.132	3.020	2.779
VTZ ^m	3.427	2.984	1.752	1.414
VQZ	3.978	3.517	1.913	1.717
VQZ*	3.417	2.969	1.172	1.065
VQZ ^m	3.131	2.662	0.723	0.582
V{T,Q}Z	3.955	3.334	1.578	1.264
V{T,Q}Z*	3.521	2.892	0.977	0.733
V{T,Q}Z ^m	3.351	2.710	0.596	0.323
VSZ*	3.054	2.678	0.372	0.278
VSZ ^m	3.020	2.641	0.299	0.199
V{Q,S}Z*	3.105	2.660	0.302	0.167
V{Q,S}Z ^m	3.115	2.671	0.275	0.138
def2-TZVPP	3.966	3.701	2.534	2.369
def2-TZVPP*	3.412	3.158	1.883	1.750
def2-TZVPP ^m	3.162	2.889	1.633	1.480
def2-TZVPPD	3.157	2.781	1.530	1.405
def2-QZVPP	3.267	2.878	1.045	0.899
def2-QZVPP*	3.007	2.671	0.748	0.636
def2-QZVPP ^m	2.953	2.613	0.750	0.638
def2-QZVPPD	2.965	2.577	0.743	0.625
def2-{T,Q}ZVPP	3.326	2.856	0.743	0.560
def2-{T,Q}ZVPP*	3.177	2.719	0.573	0.387
def2-{T,Q}ZVPP ^m	3.187	2.730	0.519	0.329
def2-{T,Q}ZVPPD	3.111	2.724	0.456	0.289

^aAll values are reported in kcal/mol.

above the reference values, respectively. Small discrepancies between the three basis sets are mostly because of rare-gas clusters RG18 with their outsize weights, which contribute 0.042 and 0.016 kcal/mol, respectively, toward the increase in WTMAD2 for VDZ-F12* to VTZ-F12* and VTZ-F12* to VQZ-F12*. The other test sets that contribute toward deviations between VDZ-F12* and VTZ-F12* are HEAVY28 (0.021 kcal/mol) and HAL59 (0.008 kcal/mol).

A reviewer inquired whether the mildly non-monotonic basis set convergence observed in Table 2 could be attributable to the excessive weights given to the three subsets RG18, HEAVY28, and HAL59 in eq 1. Table 3 compares convergence behavior for GMTKN55 including versus excluding these three subsets from the summations in the numerator and denominator of eq 1. Although the "exclusive"

WTMAD2 values are naturally considerably smaller, the mild non-monotonicity persists and likely needs to be attributed to subtle error compensations in the individual subsets between basis set incompleteness and intrinsic functional error. Consistent with this conjecture, we observe that WTMAD2 values relative to the complete basis set limit of the functional (right-hand pane of Table 3) do converge monotonically both with and without the said three subsets.

WTMAD2 obtained with V{D,T}Z-F12* (2.993 kcal/mol) and V{T,Q}Z-F12* (3.016 kcal/mol) pairs can essentially be regarded as the basis set limit. We used the two-point extrapolation formula ($A + B/L^\alpha$), L = highest angular momentum present in the basis set) for the PT2 components with $\alpha = 3.0878$ for the V{D,T}Z-F12 pair and $\alpha = 4.3548$ for the V{T,Q}Z-F12 pair.¹³⁸ The extrapolation of the KS

component essentially provides the same WTMAD2 as obtained with just the highest angular momentum present in the basis set and CABS. Switching off the CABS correction only increases the WTMAD2 value for $V\{D,T\}Z$ -F12 from 3.005 to 3.013 kcal/mol.

In order to explore whether MP2-F12 extrapolation exponents can be safely used for the PT2-F12 component in DHDF-F12, we performed a sensitivity analysis of B2GP-PLYP-F12-D3(BJ)/ $V\{D,T\}Z$ -F12 extrapolation by calculating rmsd differences [rmsd(extrapolation exponent α)—rmsd(∞)] for the atomization energies of the W4-11 set calculated relative to B2GP-PLYP-F12-D3(BJ)/ $V\{T,Q\}Z$ -F12. Figure 1 shows a minimum near $\alpha = 3.4$. $\alpha = 3.0878$ taken

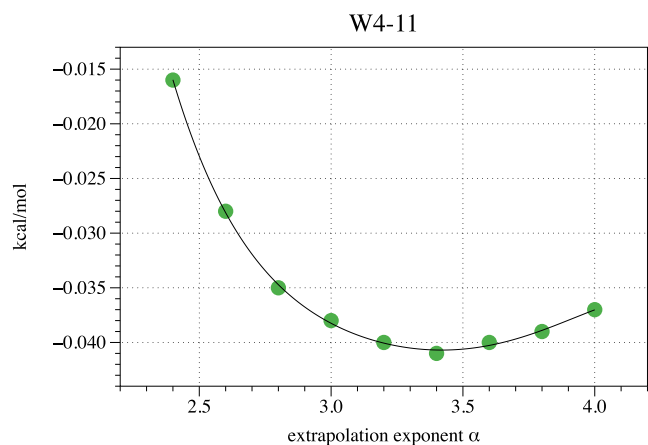


Figure 1. Sensitivity analysis of the B2GP-PLYP-F12-D3(BJ)/ $V\{D,T\}Z$ extrapolation. Root-mean-square deviation (rmsd) differences [rmsd(extrapolation exponent α)—rmsd(∞)] for the atomization energies of the W4-11⁵⁹ set calculated relative to B2GP-PLYP-F12-D3(BJ)/ $V\{T,Q\}Z$ -F12.

from ref 138 yields rmsd(α)—rmsd(∞) = -0.040 kcal/mol instead of -0.041 kcal/mol for $\alpha = 3.4$, which is a negligible difference in the larger scheme of things. For different double hybrids, the minimum of this shallow curve might vary slightly around $\alpha = 3.4$, without significantly affecting rmsd. Hence, we elected to retain the MP2-F12 extrapolation exponent.

3.3. Aside on BSSE. A brief digression on BSSE might shed more light on basis set convergence behavior. For the intermolecular subset of GMTKN55, one has the option of applying counterpoise (CP) corrections¹⁴⁰ (for detailed

discussion and further references, see Burns et al.¹⁴¹ for WFT methods, Brauer et al.¹⁴² for F12 methods, and ref 143 for DFT and double hybrids). For the intramolecular subset, CP corrections would be rather more awkward, although geometric CP corrections do exist^{144,145} for some levels of theory. (For an alternative approach to noncovalent interactions for large systems, involving small tailored basis sets, see ref 146 and references therein). Hence, most groups that employ GMTKN55 avoid CP corrections, which of course presuppose basis sets large enough that these no longer matter (much).

One major benefit of F12 methods (with F12 basis sets) was previously found to be^{142,147} a drastic reduction in BSSE, as shown for thermochemistry¹⁴⁷ and for noncovalent interactions.^{142,143}

Table 4 presents CP corrections for the Watson–Crick and stacked uracil dimers (systems 17 and 26, respectively, in S66), as representative examples of strong hydrogen bonding and π -stacking, respectively. As seen in Table 4, B2GP-PLYP-F12/ VnZ -F12 leads to a BSSE reduction by an order of magnitude (or more) over the corresponding B2GP-PLYP/ VnZ calculation, and indeed one has to go all the way to V5Z to find a basis set with a similarly low BSSE as B2GP-PLYP-F12/VDZ-F12 (!). For haVnZ-F12 versus haVnZ, and for AVnZ-F12 versus AVnZ, one likewise sees one order of magnitude reduction in BSSE. Additionally, AVnZ-F12 further reduced BSSE by about a factor 2–3 over the already low values for VnZ-F12.

At the CBS limit, the BSSE correction should of course be zero, as raw and CP-corrected calculations should yield the same answer. The deviation from zero when extrapolating CP corrections to the CBS limit is a good proxy for the quality of the extrapolation (and its underlying basis sets). For $V\{T,Q\}Z$ -F12 and AV $\{T,Q\}Z$ -F12, this evidently works beautifully. For $V\{D,T\}Z$ -F12 and AV $\{D,T\}Z$ -F12, not much improvement over the already low BSSE of VTZ-F12 viz. AVTZ-F12 can be seen. For $V\{Q,S\}Z$, on the other hand, we find a large negative BSSE that indicates overcorrection. In fact, simple VDZ-F12 has less BSSE than $V\{Q,S\}Z$ and similar to haV $\{Q,S\}Z$.

3.4. Basis Set Convergence Relative to the Complete Basis Set Limit. Furthermore, we explored the basis set convergence of conventional and explicitly correlated double-hybrid calculations using basis set limit reference values (Table 5). For this purpose, we used energies calculated at the B2GP-PLYP-F12-D3(BJ)/ $V\{T,Q\}Z$ -F12* level of theory, as they are sufficiently converged to the basis set limit. Conventional

Table 4. B2GP-PLYP-F12 Compared to B2GP-PLYP CP Corrections (kcal/mol) for the Two Uracil Dimer Structures in S66 Using Different Basis Sets^a

	dimer 17 Watson-Crick	dimer 26 π -stacked		dimer 17 Watson-Crick	dimer 26 π -stacked		dimer 17 Watson-Crick	dimer 26 π -stacked
B2GP-PLYP-F12								
VDZ-F12	0.191	0.317	haVDZ-F12	0.107	0.128	AVDZ-F12	0.097	0.124
VTZ-F12	0.106	0.208	haVTZ-F12	0.056	0.113	AVTZ-F12	0.065	0.116
VQZ-F12	0.042	0.068	haVQZ-F12	0.014	0.026	AVQZ-F12	0.017	0.027
B2GP-PLYP								
VDZ	4.768	4.545	haVDZ	1.285	2.927	AVDZ	1.848	3.083
VTZ	1.630	2.089	haVTZ	0.650	1.120	AVTZ	0.901	1.223
VQZ	0.634	0.908	haVQZ	0.261	0.465	AVQZ	0.348	0.507
V5Z	0.210	0.289	haV5Z	0.108	0.169	AV5Z	0.160	0.188

^ahaVnZ-F12, by analogy with haVnZ, corresponds to AVnZ-F12 on nonhydrogen elements and VnZ-F12 on hydrogen.

Table 5. Statistical Analysis of the Basis Set Convergence in Conventional and Explicitly Correlated B2GP-PLYP-D3(BJ) Calculations for the GMTKN55 Database and Its Categories, Relative to the B2GP-PLYP-F12-D3(BJ)/V{T,Q}Z-F12* Reference Data^a

	B2GP-PLYP-D3(BJ)						B2GP-PLYP-F12-D3(BJ)						
	WTMAD2	THERMO	BARRIERS	LARGE	CONF	INTERMOL	WTMAD2	THERMO	BARRIERS	LARGE	CONF	INTERMOL	
VDZ	11.303	2.251	0.869	0.854	4.110	3.220	AVDZ-F12	0.418	0.087	0.036	0.061	0.083	0.152
VDZ*	9.014	1.340	0.464	0.854	4.110	2.246	VDZ-F12	0.499	0.096	0.042	0.050	0.098	0.212
VDZ ^m	5.602	1.340	0.464	0.755	1.416	1.627	VDZ-F12*	0.467	0.082	0.038	0.050	0.098	0.199
VTZ	4.317	0.853	0.373	0.174	1.325	1.592	VTZ-F12	0.220	0.035	0.019	0.021	0.055	0.090
VTZ*	3.020	0.399	0.137	0.174	1.325	0.984	VTZ-F12*	0.207	0.025	0.016	0.021	0.055	0.090
VTZ ^m	1.752	0.399	0.137	0.145	0.404	0.666	V{D,T}Z-F12	0.232	0.033	0.023	0.027	0.062	0.088
VQZ	1.913	0.407	0.201	0.082	0.431	0.792	V{D,T}Z-F12*	0.215	0.026	0.020	0.027	0.062	0.079
VQZ*	1.172	0.160	0.054	0.082	0.431	0.446	VQZ-F12	0.065	0.012	0.006	0.004	0.006	0.038
VQZ ^m	0.723	0.160	0.054	0.075	0.151	0.284	VQZ-F12*	0.032	0.005	0.002	0.004	0.006	0.016
V{T,Q}Z	1.578	0.244	0.180	0.075	0.246	0.833	V{T,Q}Z-F12	0.038	0.006	0.004	0.000	0.001	0.028
V{T,Q}Z*	0.977	0.061	0.045	0.075	0.246	0.550	V{T,Q}Z-F12*	REF	REF	REF	REF	REF	REF
V{T,Q}Z ^m	0.596	0.061	0.045	0.073	0.066	0.351							
V5Z*	0.372	0.063	0.016	0.023	0.091	0.179							
V5Z ^m	0.299	0.063	0.016	0.023	0.050	0.147							
V{Q,S}Z*	0.302	0.021	0.014	0.027	0.063	0.178							
V{Q,S}Z ^m	0.275	0.021	0.014	0.024	0.042	0.175							
def2-TZVPP	2.534	0.600	0.240	0.169	0.694	0.831							
def2-TZVPP*	1.883	0.400	0.098	0.169	0.694	0.523							
def2-TZVPP ^m	1.633	0.400	0.098	0.157	0.474	0.505							
def2-TZVPPD	1.530	0.326	0.074	0.175	0.473	0.483							
def2-QZVPP	1.045	0.253	0.106	0.082	0.202	0.402							
def2-QZVPP*	0.748	0.171	0.038	0.082	0.202	0.254							
def2-QZVPP ^m	0.750	0.171	0.038	0.082	0.201	0.258							
def2-QZVPPD	0.743	0.157	0.030	0.078	0.209	0.269							
def2-{T,Q}ZVPP	0.743	0.132	0.077	0.070	0.091	0.372							
def2-{T,Q}ZVPP*	0.573	0.089	0.025	0.070	0.091	0.298							
def2-{T,Q}ZVPP ^m	0.519	0.089	0.025	0.067	0.085	0.254							
def2-{T,Q}ZVPPD	0.456	0.090	0.020	0.033	0.088	0.225							
VDZ-F12	5.324	0.795	0.324	0.576	1.238	2.392							

^aValues are heat-mapped from red for the largest via yellow for median to green for the smallest. Note that values are heat-mapped separately for each category of GMTKN55 and the entire database.

B2GP-PLYP-D3(BJ) calculations in conjunction with the VDZ^m basis set yield a WTMAD2 value that is 5.602 kcal/mol above the basis set limit. Increasing the basis set to VTZ^m and VQZ^m reduces this deviation to 1.752 and 0.723 kcal/mol, respectively. VSZ^m yields a deviation that is only 0.299 kcal/mol above our best estimate (B2GP-PLYP-F12-D3(BJ)/V{T,Q}Z-F12*). Basis set limit reaction energies for the conventional B2GP-PLYP-D3(BJ)/V{Q,S}Z^m calculations differ by only 0.275 kcal/mol from explicitly correlated B2GP-PLYP-F12-D3(BJ)/V{T,Q}Z-F12*, of which inter- and intramolecular noncovalent interactions account for the lion's share. A closer inspection of the individual subsets revealed that HEAVY28, HALS9, and RG18 are the three largest contributors to the discrepancies, their Δ WTMAD2 of HEAVY28, RG18, and HALS9 being 0.086, 0.034, and 0.027 kcal/mol, respectively, relative to B2GP-PLYP-F12-D3(BJ)/V{T,Q}Z-F12*. As discussed above, the way these three subsets are weighted in the WTMAD2 formula, a small change in reaction energies has a disproportionate contribution to WTMAD2.

B2GP-PLYP-D3(BJ) in conjunction with def2-TZVPP, def2-QZVPP, and def2-{T,Q}ZVPP ($\alpha = 7.6070$ for KS and 2.5313 for PT2) basis sets provides WTMAD2 which are 2.534, 1.045, and 0.743 kcal/mol above our best estimate, respectively. Adding diffuse functions to RG18, AHB21, BH76, BH76RC, IL16, G21EA, and WATER27 (i.e., def2-nZVPP* basis set) lowers the WTMAD2 values to 1.883, 0.748, and 0.573 kcal/mol, respectively, for TZ, QZ, and {T,Q}Z basis. On the other hand, def2-TZVPPD, def2-QZVPPD, and def2-{T,Q}ZVPPD provide WTMAD2 which are 1.530, 0.743, and 0.456 kcal/mol, respectively.

Turning our attention to explicitly correlated B2GP-PLYP-F12-D3(BJ) calculations with VnZ-F12 type basis sets, we note that VDZ-F12* already yields an acceptable WTMAD2 which

is only 0.467 kcal/mol from the F12 basis set limit. Moving on to AVDZ-F12 provides a WTMAD2 which is just 0.050 kcal/mol below VDZ-F12*. The WTMAD2 component breakdown revealed that S66, HEAVY28, and AMINO20x4 together account for 0.043 kcal/mol of the total improvement in Δ WTMAD2 of AVDZ-F12 in comparison to VDZ-F12*. Increasing the basis set size to VTZ-F12* yields a WTMAD2 of 0.207 kcal/mol. WTMAD2 obtained with the VQZ-F12* basis set (0.032 kcal/mol) can essentially be regarded as the basis set limit.

3.5. Note on Systematic Errors. Thus far, we have only discussed WTMAD2. It would be of interest to compare, as a measure of systematic error, the basis set convergence of conventional and explicitly correlated double hybrids in terms of weighted total mean signed deviations (WTMSDs), where MAD in eq 1 is replaced by the MSD. (For clarity, the way the reference data's sign conventions work, a positive WTMSD2 indicates overbinding, and a negative one underbinding). Tables S1 and S2 in Supporting Information present the WTMSD2 of conventional and explicitly correlated B2GP-PLYP-D3(BJ) relative to the reference data⁶ and to the complete basis set limit as obtained at the B2GP-PLYP-F12-D3(BJ)/V{T,Q}Z-F12* level. WTMSD2 obtained for B2GP-PLYP-F12-D3(BJ) with VDZ-F12* (0.081 kcal/mol), VTZ-F12* (0.071 kcal/mol), and V{D,T}Z-F12* (0.087 kcal/mol) basis sets are all close to zero and indicate that there is little systematic error relative to the complete basis set limit (Table S2). Turning our attention to conventional B2GP-PLYP-D3(BJ), we obtained WTMSD2 of 2.693, 1.077, 0.501, 0.076, and 0.042 kcal/mol for the VDZ*, VTZ*, VQZ*, VSZ*, and V{Q,S}Z* basis sets, respectively. These results indicate that even VQZ* still overbinds significantly due to BSSE: a breakdown into the five top-level subsets (Table S2) reveals

Table 6. Relative CPU Timings for the B2GPPLYP-D3(BJ) and B2GPPLY-F12-D3(BJ) Calculations for $(C_nH_{n+2})_2^a$

	B2GP-PLYP-F12			B2GP-PLYP					
	PNO-LPT2		canonical	{T,Q}ZVPP	{T,Q}ZVPPD	V{T,Q}Z	haV{T,Q}Z	haV{Q,5}Z	AV{T,Q}Z
	PNO-F12	VDZ-F12	VDZ-F12						
(ethane) ₂	1.08	1.00	1.00	1.38	1.92	1.74	2.02	7.20	3.69
(propane) ₂	1.05	1.00	1.00	1.22	1.84	1.53	2.19	8.33	4.15
(butane) ₂	0.83	1.00	1.00	0.91	1.47	1.58	1.84	7.15	3.47
(pentane) ₂	0.63	1.00	1.00	0.70	1.16	0.91	1.49	5.83	2.78
(hexane) ₂	0.49	1.00	1.00	0.57	0.95	0.73	1.23	4.93	2.30
(heptane) ₂	0.38	1.00	1.00	0.48	0.81	0.62	1.05	4.24	1.96
(n-octane) ₂	0.28	1.00	1.00	0.38	0.66	0.49	0.85	3.43	1.58
(n-nonane) ₂	0.22	1.00	1.00	0.32	0.56	0.41	0.72	2.96	1.34
(n-decane) ₂	0.18	1.00	1.00	0.28	0.49	0.36	0.63	2.60	1.16
(n-dodecane) ₂	0.11	1.00	1.00	0.22	0.39	0.28	0.49	2.09	0.92

^aTiming is shown relative to B2GP-PLYP-F12-D3(BJ)/VDZ-F12; white = 1, blue = faster, and red = slower.

that intermolecular interactions account for about two-thirds, and conformers for almost all the remainder. Turning to the VnZ^m variants, VDZ^m yields a deceptively low WTMSD2 = 0.345 kcal/mol owing to compensation between sizable systematic underestimates of basic thermochemistry, barrier heights, a large molecule reaction, systematic overestimates of conformer energies, and especially intermolecular interactions. Although VTZ^m sees WTMSD2 “only” reduced to 0.235 kcal/mol, the constituent values for the five top-level subsets are actually reduced by factors of 3–4. (This, incidentally, illustrates the dangers of blindly relying on a single global metric). The 0.182 kcal/mol for VQZ^m is mostly driven by the noncovalent interactions (0.213 kcal/mol), slightly compensated by basic thermochemistry and barrier heights. Finally, $V5Z^m$ has only a mild systematic error, again mostly from noncovalent interactions. $V\{T,Q\}Z^m$ extrapolation yields a WTMSD2 of 0.337 kcal/mol, which at first sight seems inferior even to VDZ^m ; however, closer inspection reveals that essentially all of that number comes from overbinding in the intermolecular interactions due to BSSE and that the remaining four top-level subsets are nearly free of systematic bias. For $V\{Q,5\}Z^m$, WTMSD2 is down to a paltry 0.048 kcal/mol, essentially all of it again from intermolecular interactions. Interestingly, def2-QZVPP (−0.228 kcal/mol) and especially def2-QZVPP* (−0.024 kcal/mol) and def2-QZVPPD (−0.016 kcal/mol) have much smaller WTMSD2 values, also for the top-level subsets: the negative signs reflect mostly underestimates for small-molecule thermochemistry. The different signs of the def2- $\{T,Q\}ZVPPD^*$ and def2- $\{T,Q\}ZVPPD$ WTMSD2s essentially reflect systematic overbinding versus underbinding of intermolecular interactions, where the former lacks diffuse functions on such subsets as S22, S66, and the like.

3.6. Computational Cost Considerations. It is of interest to compare the relative computational cost of conventional B2GP-PLYP-D3(BJ) and B2GP-PLYP-F12-D3(BJ) procedures. Each of these timing evaluation jobs was run on otherwise empty nodes with identical hardware (Intel Haswell 2.4 GHz with 256 GB RAM and a 3.6TB SSD RAID array). These jobs were run serially, on a single core, in order to eliminate differences in parallelization as a confounding factor. Timing data relative to VDZ-F12 are reported in Table 6 for the six linear n -alkane dimers in ADIM6,^{6,97} (ethane)₂ through (n-heptane)₂, plus additionally (n-octane)₂, (n-nonane)₂, (n-decane)₂, and (n-dodecane)₂ with the structures

obtained by manually inserting additional CH₂ groups (because we needed them only for timing purposes). As we have seen above, even the VDZ-F12 basis set yields results of a quality comparable to $V\{Q,5\}Z^m$. In view of the fact that no CP correction is being applied, one would definitely want to use $haV\{Q,5\}Z$ for this subset rather than $V\{Q,5\}Z$. The sum of timings for both calculations involved in the extrapolation ranges from 8 times longer than VDZ-F12 for propane dimer to twice as long for n -dodecane dimer, with the ratio increasingly less favorable to VDZ-F12 as the chain lengths increase. For $AV\{T,Q\}Z$, one goes from about 4 times to about the same time, and for def2- $\{T,Q\}ZVPPD$ that falls from 2 times slower to over twice as fast. In addition, the canonical VDZ-F12 calculations will be increasingly more demanding in mass storage requirements, at least for MOLPRO.

However, in a very recent communication,¹⁴⁸ we have shown that the F12 step can be drastically accelerated by evaluating it in terms of localized pair natural orbitals (i.e., PNO-DHDF-F12) without materially sacrificing accuracy. In addition, it is shown there that the scratch storage requirements are an order of magnitude smaller and that parallelization is more efficient than canonical F12 as well. By way of illustration of what this approximation enables, we offer a PNO-B2GP-PLYP-LMP2/VDZ-F12 calculation on C₆₀ (no symmetry): after deleting a diffuse p function that causes insurmountable near-linear dependence issues, it took just 51 min wall clock time on 16 cores of an Ice Lake 2.2 GHz node.

For the discussion at hand here, single-core PNO-DHDF-F12 timing data can be found in the first column of Table 6. For the smallest systems, as expected, the localized approach offers no benefit, but as the chain length increases ever-better speedup is realized, to reach an order of magnitude for n -dodecane dimer. By way of illustration, we applied a power fit to the canonical DHDF-F12 and PNO-DHDF-F12 wall clock times for n -pentane through n -dodecane dimers and found very good fits with $R^2 = 0.99917$ and $R^2 = 0.99996$, respectively: the scaling exponents are 4.57 and 2.60, respectively, showing the scaling advantage of the PNO-F12 approach. Indeed, in ref 148 we show that for still longer chains through n -tetracosane dimer (i.e., $n = 24$), scaling with n decreases further toward linearity.

It is clear that PNO-DHDF-F12/VDZ-F12 offers a more economical alternative than any of the basis set extrapolations that would yield comparable, or even somewhat inferior, accuracy. We therefore do believe that the DH-F12 approach,

especially in its localized orbital form, compares favorably in both accuracy and efficiency with large basis set B2GP-PLYP-D3(BJ).

3.7. Spin-component-scaled Double Hybrids. We will now evaluate GMTKN55 performance for the more recent and accurate revDSD-PBEP86-D3(BJ) functional⁸ with and without explicit correlation. Table 7 presents statistical analysis for conventional revDSD-PBEP86-D4 and explicitly correlated revDSD-PBEP86-F12-D4 calculations. Using the VnZ^m basis set in conjunction with conventional revDSD-PBEP86-D4 results in WTMAD2 values of 2.236 and 2.104 kcal/mol, respectively, for VQZ^m and VSZ^m basis sets. Leaving out diffuse functions altogether—including in the anionic subsets such as the G21EA electron affinities and the hydroxide clusters in WATER27—WTMAD2 unacceptably increases by 0.6 kcal/mol from VQZ^* to VQZ and by 0.4 kcal/mol from VSZ^* to VSZ . G21EA alone accounts for 0.176 (VQZ) and 0.077 (VSZ) kcal/mol, respectively.

Finally, the $V\{Q_5\}Z^m$ pair yields a WTMAD2 of 2.233 kcal/mol. Clearly, in the F12 calculations, WTMAD2 converges spectacularly faster with respect to the basis set size, with even VDZ-F12* reaching statistics comparable to VSZ^* in the non-F12 approach. VDZ-F12* and VTZ-F12* yield WTMAD2 values which are 2.233 and 2.218 kcal/mol above the reference values; the latter is close to the “basis set limit” goal as WTMAD2 of $V\{D,T\}Z$ -F12* is only 0.006 kcal/mol below VTZ-F12*.

At a reviewer’s request, we further explored the basis set convergence of the same spin and opposite spin components of the PT2 term in a double hybrid. It is well established (see, e.g., Kutzelnigg and Morgan²⁵) that in the large- L limit, MP2 same-spin correlation energies converge as L^{-5} and their opposite-spin counterparts as L^{-3} . Hence, for sufficiently large basis sets, opposite-spin will dominate the convergence behavior and same-spin will be effectively saturated.

Although it stands to reason that this would also be the case for GLPT2 correlation, by way of illustration we show in Figure S1 that this is indeed the case for the same-spin (E_{2ss}) and opposite-spin (E_{2ab}) components of the B2GP-PLYP atomic correlation energy of neon atom along the $nZaP$ basis set sequence ($n = 3-8$, with maximum angular momentum $L = n$) of Petersson.^{149,150} Hence, for sufficiently large L , same-spin and opposite-spin contributions in B2GP-PLYP converge as L^{-5} and L^{-3} , respectively, and the latter will completely dominate convergence of the overall correlation energy.

4. CONCLUSIONS

We have investigated the basis set convergence of double hybrids in conjunction with explicitly correlated (F12) on a large and chemically diverse GMTKN55 database. We chose B2GP-PLYP-D3(BJ) and revDSD-PBEP86-D3(BJ) as test cases. Two families of basis sets were considered: orbital basis sets as large as aug-cc-pV(5+d)Z and F12 basis sets as large as cc-pVQZ-F12. We found that explicitly correlated double-hybrid calculations with F12 basis converge markedly faster than the conventional double-hybrid calculations with orbital (aug-)cc-pV(5+d)Z or def2 basis sets. In fact, DHDF-F12 calculations with just a cc-pVDZ-F12 basis set are closer to the basis set limit than DHDF/cc-pV(Q+d)Z or def2-QZVPPD and approach DHDF/cc-pV(5+d)Z in quality at about one-third the cost. One significant benefit of DHDF-F12 is reducing BSSE by an order of magnitude over orbital-only DHDF in a similar-sized basis set: this particularly benefits the

Table 7. Statistical Analysis of the Basis Set Convergence in Conventional and Explicitly Correlated revDSD-PBEP86-D4 Calculations for the GMTKN55 Database and Its Categories, Relative to the Reference 6 Reference Data

	revDSD-PBEP86-D4					revDSD-PBEP86-F12-D4						
	WTMAD2	THERMO	BARRIERS	LARGE	CONF	INTERMOL	WTMAD2	THERMO	BARRIERS	LARGE	CONF	INTERMOL
VQZ	3.087	0.767	0.355	0.489	0.539	0.937	2.247	0.518	0.310	0.545	0.412	0.463
VQZ*	2.494	0.547	0.244	0.489	0.539	0.675	2.233	0.513	0.306	0.545	0.412	0.458
VQZ ^m	2.236	0.547	0.244	0.487	0.399	0.559	2.216	0.520	0.308	0.530	0.397	0.462
VSZ	2.502	0.628	0.318	0.501	0.386	0.670	2.218	0.521	0.305	0.530	0.397	0.466
VSZ*	2.101	0.496	0.226	0.501	0.386	0.492	2.213	0.517	0.310	0.526	0.389	0.471
VSZ ^m	2.104	0.496	0.226	0.499	0.402	0.480	2.213	0.520	0.307	0.526	0.389	0.471
V{Q ₅ }Z	2.563	0.600	0.313	0.503	0.451	0.696						
V{Q ₅ }Z*	2.235	0.516	0.227	0.503	0.451	0.537						
V{Q ₅ }Z ^m	2.233	0.516	0.227	0.502	0.433	0.555						

noncovalent interaction subsets (both intermolecular and conformer). We also found that even for anionic systems, the anion-friendly aug-cc-pVDZ-F12 basis set proved unnecessary and cc-pVDZ-F12 was adequate. Finally, although the application of DH-F12 to larger molecules will eventually face mass storage and I/O bandwidth challenges in a disk-based algorithm, these can be circumvented through localized pair natural orbital approaches,¹⁴⁸ which also reduce CPU time scaling by (in practice) about n^2 .

Summing up, explicitly corrected double-hybrid calculations are an economical and accurate alternative if (near-)basis set limit results are required, for example, for benchmarking or parametrizing double-hybrid DFT methods. Implementation in other electronic structure systems of MP2-F12 in a basis of Kohn–Sham orbitals would be a very worthwhile endeavor, especially if said implementation is parsimonious in I/O requirements.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.2c00426>.

Statistical results of all assessed methods, as well as sample MOLPRO input files for B2GP-PLYP-F12-D3(BJ) and revDSD-PBEP86-F12-D4 and Figure S1 (PDF)

GMTKN55 statistical summary files for various basis set - functional combinations (ZIP)

Excel spreadsheet with additional timing information (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Jan M. L. Martin – Department of Molecular Chemistry and Materials Science, Weizmann Institute of Science, 7610001 Rehovot, 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, 7610001 Rehovot, Israel; orcid.org/0000-0001-7222-4108

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.jctc.2c00426>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Work on this paper was supported in part 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.

■ REFERENCES

- (1) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics*; Cambridge University Press: Cambridge, 2009.
- (2) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.
- (3) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- (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-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129–5143.
- (9) Mehta, N.; Casanova-Páez, 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 ω B97M(2) double hybrid density functional. *J. Chem. Phys.* **2018**, *148*, 241736.
- (12) Brémond, E.; Ciofini, I.; Sancho-García, J. C.; Adamo, C. Nonempirical Double-Hybrid Functionals: An Effective Tool for Chemists. *Acc. Chem. Res.* **2016**, *49*, 1503–1513.
- (13) Zhang, I. 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) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Doubly Hybrid Meta DFT: New Multi-Coefficient Correlation and Density Functional Methods for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A* **2004**, *108*, 4786–4791.
- (16) 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.
- (17) Ángyán, J. G.; Gerber, I. C.; Savin, A.; Toulouse, J. Van Der Waals Forces in Density Functional Theory: Perturbational Long-Range Electron-Interaction Corrections. *Phys. Rev. A* **2005**, *72*, 012510.
- (18) Santra, G.; Martin, J. M. L. Does GLPT2 Offer Any Actual Benefit Over Conventional HF-MP2 in the Context of Double-Hybrid Density Functionals? **2021**. AIP Conf. Proc., **2022**, in press. <http://arxiv.org/abs/2111.01880>
- (19) Semidalas, E.; Martin, J. M. L. Canonical and DLPNO-Based G4(MP2)XK-Inspired Composite Wave Function Methods Parametrized Against Large and Chemically Diverse Training Sets: Are They More Accurate and/or Robust Than Double-Hybrid DFT? *J. Chem. Theory Comput.* **2020**, *16*, 4238–4255.
- (20) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.
- (21) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 Theory. *J. Chem. Phys.* **2007**, *126*, 084108.
- (22) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *Gn* theory. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 810–825.
- (23) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conf. Proc.* **2001**, *577*, 1–20.

- (24) Schwartz, C. Importance of Angular Correlations Between Atomic Electrons. *Phys. Rev.* **1962**, *126*, 1015–1019.
- (25) Kutzelnigg, W.; Morgan, J. D. Rates of convergence of the partial-wave expansions of atomic correlation energies. *J. Chem. Phys.* **1992**, *96*, 4484–4508.
- (26) 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.
- (27) Kendall, R. A.; Früchtl, H. A. The Impact of the Resolution of the Identity Approximate Integral Method on Modern Ab Initio Algorithm Development. *Theor. Chem. Acc.* **1997**, *97*, 158–163.
- (28) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: Optimized Auxiliary Basis Sets and Demonstration of Efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152.
- (29) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N₂, and H₂O. *Chem. Phys. Lett.* **1998**, *286*, 243–252.
- (30) Ranasinghe, D. S.; Petersson, G. A. CCSD(T)/CBS Atomic and Molecular Benchmarks for H Through Ar. *J. Chem. Phys.* **2013**, *138*, 144104.
- (31) Martin, J. M. L. A Simple ‘Range Extender’ for Basis Set Extrapolation Methods for MP2 and Coupled Cluster Correlation Energies. *AIP Conf. Proc.* **2018**, *2040*, 020008.
- (32) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75–107.
- (33) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly Correlated Electrons in Molecules. *Chem. Rev.* **2012**, *112*, 4–74.
- (34) Ten-No, S.; Noga, J. Explicitly Correlated Electronic Structure Theory From R12/F12 Ansätze. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 114–125.
- (35) Klopper, W.; Kutzelnigg, W. Møller-plešset calculations taking care of the correlation CUSP. *Chem. Phys. Lett.* **1987**, *134*, 17–22.
- (36) Kutzelnigg, W.; Klopper, 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.
- (37) Ten-no, S. Initiation of Explicitly Correlated Slater-Type Geminal Theory. *Chem. Phys. Lett.* **2004**, *398*, 56–61.
- (38) Persson, B. J.; Taylor, P. R. Molecular Integrals Over Gaussian-Type Geminal Basis Functions. *Theor. Chem. Acc.* **1997**, *97*, 240–250.
- (39) Helgaker, T.; Klopper, W.; Tew, D. P. Quantitative Quantum Chemistry. *Mol. Phys.* **2008**, *106*, 2107–2143.
- (40) Klopper, W.; Manby, F. R.; Ten-No, S.; Valeev, E. F. R12 Methods in Explicitly Correlated Molecular Electronic Structure Theory. *Int. Rev. Phys. Chem.* **2006**, *25*, 427–468.
- (41) Klopper, W.; Samson, C. C. M. Explicitly correlated second-order Møller-Plesset methods with auxiliary basis sets. *J. Chem. Phys.* **2002**, *116*, 6397–6410.
- (42) Manby, F. R. Density fitting in second-order linear-r12 Møller-Plesset perturbation theory. *J. Chem. Phys.* **2003**, *119*, 4607–4613.
- (43) 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.
- (44) 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.
- (45) Peterson, K. A.; Dixon, D. A.; Stoll, H. The Use of Explicitly Correlated Methods on XeF₆ Predicts a C_{3v} Minimum with a Sterically Active, Free Valence Electron Pair on Xe. *J. Phys. Chem. A* **2012**, *116*, 9777–9782.
- (46) 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.
- (47) Karton, A.; Martin, J. M. L. Explicitly correlated Wn theory: W1-F12 and W2-F12. *J. Chem. Phys.* **2012**, *136*, 124114.
- (48) Peterson, K. A.; Kesharwani, M. K.; Martin, J. M. L. The cc-pV5Z-F12 Basis Set: Reaching the Basis Set Limit in Explicitly Correlated Calculations. *Mol. Phys.* **2015**, *113*, 1551–1558.
- (49) Sylvetsky, N.; Peterson, K. A.; Karton, A.; 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.
- (50) Fogueri, 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*, 2269–2277.
- (51) Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. L. The S66x8 Benchmark for Noncovalent Interactions Revisited: Explicitly Correlated Ab Initio Methods and Density Functional Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20905–20925.
- (52) Manna, D.; Kesharwani, M. K.; Sylvetsky, N.; Martin, J. M. L. Conventional and Explicitly Correlated Ab Initio Benchmark Study on Water Clusters: Revision of the BEGDB and WATER27 Data Sets. *J. Chem. Theory Comput.* **2017**, *13*, 3136–3152.
- (53) Kesharwani, M. K.; Manna, D.; Sylvetsky, N.; Martin, J. M. L. The X40×10 Halogen Bonding Benchmark Revisited: Surprising Importance of (n-1)d Subvalence Correlation. *J. Phys. Chem. A* **2018**, *122*, 2184–2197.
- (54) Karton, A.; Martin, J. M. L. Basis Set Convergence of Explicitly Correlated Double-Hybrid Density Functional Theory Calculations. *J. Chem. Phys.* **2011**, *135*, 144119.
- (55) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.
- (56) Semidalas, E.; Martin, J. M. L. Canonical and DLPNO-Based Composite Wavefunction Methods Parametrized against Large and Chemically Diverse Training Sets. 2: Correlation-Consistent Basis Sets, Core-Valence Correlation, and F12 Alternatives. *J. Chem. Theory Comput.* **2020**, *16*, 7507–7524.
- (57) Santra, G.; Semidalas, E.; Martin, J. M. L. Surprisingly Good Performance of XYG3 Family Functionals Using a Scaled KS-MP3 Correlation. *J. Phys. Chem. Lett.* **2021**, *12*, 9368–9376.
- (58) Najibi, A.; Casanova-Páez, M.; Goerigk, L. Analysis of Recent BLYP- and PBE-Based Range-Separated Double-Hybrid Density Functional Approximations for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Phys. Chem. A* **2021**, *125*, 4026–4035.
- (59) Karton, A.; Daon, S.; Martin, J. M. L. W4-11: A High-Confidence Benchmark Dataset for Computational Thermochemistry Derived From First-Principles W4 Data. *Chem. Phys. Lett.* **2011**, *510*, 165–178.
- (60) 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.
- (61) Parthiban, S.; Martin, J. M. L. Assessment of W1 and W2 Theories for the Computation of Electron Affinities, Ionization Potentials, Heats of Formation, and Proton Affinities. *J. Chem. Phys.* **2001**, *114*, 6014–6029.
- (62) Zhao, Y.; Truhlar, D. G. Assessment of Density Functionals for π Systems: Energy Differences between Cumulenes and Polyynes; Proton Affinities, Bond Length Alternation, and Torsional Potentials of Conjugated Polyenes; and Proton Affinities of Conjugated Shiff Bases. *J. Phys. Chem. A* **2006**, *110*, 10478–10486.
- (63) Yu, H.; Truhlar, D. G. Components of the Bond Energy in Polar Diatomic Molecules, Radicals, and Ions Formed by Group-1 and Group-2 Metal Atoms. *J. Chem. Theory Comput.* **2015**, *11*, 2968–2983.
- (64) Zhao, Y.; Ng, H. T.; Peverati, R.; Truhlar, D. G. Benchmark Database for Ylidic Bond Dissociation Energies and Its Use for

Assessments of Electronic Structure Methods. *J. Chem. Theory Comput.* **2012**, *8*, 2824–2834.

(65) Johnson, E. R.; Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Delocalization Errors in Density Functionals and Implications for Main-Group Thermochemistry. *J. Chem. Phys.* **2008**, *129*, 204112.

(66) Grimme, S.; Kruse, H.; Goerigk, L.; Erker, G. The Mechanism of Dihydrogen Activation by Frustrated Lewis Pairs Revisited. *Angew. Chem., Int. Ed.* **2010**, *49*, 1402–1405.

(67) Goerigk, L.; Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals-Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.

(68) 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.

(69) Zhao, Y.; González-García, N.; Truhlar, D. G. Benchmark Database of Barrier Heights for Heavy Atom Transfer, Nucleophilic Substitution, Association, and Unimolecular Reactions and Its Use to Test Theoretical Methods. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.

(70) Friedrich, J.; Hänchen, J. Incremental CCSD(T)(F12*)/MP2: A Black Box Method To Obtain Highly Accurate Reaction Energies. *J. Chem. Theory Comput.* **2013**, *9*, 5381–5394.

(71) Friedrich, J. Efficient Calculation of Accurate Reaction Energies-Assessment of Different Models in Electronic Structure Theory. *J. Chem. Theory Comput.* **2015**, *11*, 3596–3609.

(72) Grimme, S.; Mück-Lichtenfeld, C.; Würthwein, E.-U.; Ehlers, A. W.; Goumans, T. P. M.; Lammertsma, K. Consistent Theoretical Description of 1,3-Dipolar Cycloaddition Reactions. *J. Phys. Chem. A* **2006**, *110*, 2583–2586.

(73) Piacenza, M.; Grimme, S. Systematic Quantum Chemical Study of DNA-Base Tautomers. *J. Comput. Chem.* **2004**, *25*, 83–99.

(74) Woodcock, H. L.; Schaefer, H. F.; Schreiner, P. R. Problematic Energy Differences Between Cumulenes and Poly-Ynes: Does This Point to a Systematic Improvement of Density Functional Theory? *J. Phys. Chem. A* **2002**, *106*, 11923–11931.

(75) Schreiner, P. R.; Fokin, A. A.; Pascal, R. A.; de Meijere, A. Many Density Functional Theory Approaches Fail to Give Reliable Large Hydrocarbon Isomer Energy Differences. *Org. Lett.* **2006**, *8*, 3635–3638.

(76) Lepetit, C.; Chermette, H.; Gicquel, M.; Heully, J.-L.; Chauvin, R. Description of Carbo-oxocarbons and Assessment of Exchange-Correlation Functionals for the DFT Description of Carbomers. *J. Phys. Chem. A* **2007**, *111*, 136–149.

(77) Lee, J. S. Accurate Ab Initio Binding Energies of Alkaline Earth Metal Clusters. *J. Phys. Chem. A* **2005**, *109*, 11927–11932.

(78) Karton, A.; Martin, J. M. L. Explicitly correlated benchmark calculations on C₈H₈ isomer energy separations: how accurate are DFT, double-hybrid, and composite ab initio procedures? *Mol. Phys.* **2012**, *110*, 2477–2491.

(79) Zhao, Y.; Tishchenko, O.; Gour, J. R.; Li, W.; Lutz, J. J.; Piecuch, P.; Truhlar, D. G. Thermochemical Kinetics for Multi-reference Systems: Addition Reactions of Ozone. *J. Phys. Chem. A* **2009**, *113*, 5786.

(80) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(81) Manna, D.; Martin, J. M. L. What Are the Ground State Structures of C₂₀ and C₂₄? An Explicitly Correlated Ab Initio Approach. *J. Phys. Chem. A* **2016**, *120*, 153–160.

(82) 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.

(83) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of Orbital-Optimized, Spin-Component Scaled Second-

Order Many-Body Perturbation Theory for Thermochemistry and Kinetics. *J. Chem. Theory Comput.* **2009**, *5*, 3060–3073.

(84) Steinmann, S. N.; Csonka, G.; Corminboeuf, C. Unified Inter- and Intramolecular Dispersion Correction Formula for Generalized Gradient Approximation Density Functional Theory. *J. Chem. Theory Comput.* **2009**, *5*, 2950–2958.

(85) Krieg, H.; Grimme, S. Thermochemical benchmarking of hydrocarbon bond separation reaction energies: Jacob's ladder is not reversed! *Mol. Phys.* **2010**, *108*, 2655–2666.

(86) Yu, L.-J.; Karton, A. Assessment of Theoretical Procedures for a Diverse Set of Isomerization Reactions Involving Double-Bond Migration in Conjugated Dienes. *Chem. Phys.* **2014**, *441*, 166–177.

(87) Grimme, S.; Steinmetz, M.; Korth, M. How to Compute Isomerization Energies of Organic Molecules With Quantum Chemical Methods. *J. Org. Chem.* **2007**, *72*, 2118–2126.

(88) Huenerbein, R.; Schirmer, B.; Moellmann, J.; Grimme, S. Effects of London Dispersion on the Isomerization Reactions of Large Organic Molecules: A Density Functional Benchmark Study. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6940.

(89) Sure, R.; Hansen, A.; Schwerdtfeger, P.; Grimme, S. Comprehensive theoretical study of all 1812 C₆₀ isomers. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14296–14305.

(90) Guner, V.; Khuong, K. S.; Leach, A. G.; Lee, P. S.; Bartberger, M. D.; Houk, K. N. A Standard Set of Pericyclic Reactions of Hydrocarbons for the Benchmarking of Computational Methods: The Performance of ab Initio, Density Functional, CASSCF, CASPT2, and CBS-QB3 Methods for the Prediction of Activation Barriers, Reaction Energetics, and Transition State Geometries. *J. Phys. Chem. A* **2003**, *107*, 11445–11459.

(91) Ess, D. H.; Houk, K. Activation Energies of Pericyclic Reactions: Performance of DFT, MP2, and CBS-QB3 Methods for the Prediction of Activation Barriers and Reaction Energetics of 1,3-Dipolar Cycloadditions, and Revised Activation Enthalpies for a Standard Set of Hydrocarbon Pericyclic Reactions. *J. Phys. Chem. A* **2005**, *109*, 9542–9553.

(92) Dinadayalane, T. C.; Vijaya, R.; Smitha, A.; Sastry, G. N. Diels–Alder Reactivity of Butadiene and Cyclic Five-Membered Dienes ((CH)₄X, X = CH₂, SiH₂, O, NH, PH, and S) with Ethylene: A Benchmark Study. *J. Phys. Chem. A* **2002**, *106*, 1627–1633.

(93) Karton, A.; Goerigk, L. Accurate Reaction Barrier Heights of Pericyclic Reactions: Surprisingly Large Deviations for the CBS-QB3 Composite Method and Their Consequences in DFT Benchmark Studies. *J. Comput. Chem.* **2015**, *36*, 622–632.

(94) Goerigk, L.; Sharma, R. the INV24 Test Set: How Well Do Quantum-Chemical Methods Describe Inversion and Racemization Barriers? *Can. J. Chem.* **2016**, *94*, 1133–1143.

(95) Karton, A.; O'Reilly, R. J.; Chan, B.; Radom, L. Determination of Barrier Heights for Proton Exchange in Small Water, Ammonia, and Hydrogen Fluoride Clusters with G4(MP2)-Type, MP_n, and SCS-MP_n Procedures-A Caveat. *J. Chem. Theory Comput.* **2012**, *8*, 3128–3136.

(96) Karton, A.; O'Reilly, R. J.; Radom, L. Assessment of Theoretical Procedures for Calculating Barrier Heights for a Diverse Set of Water-Catalyzed Proton-Transfer Reactions. *J. Phys. Chem. A* **2012**, *116*, 4211–4221.

(97) 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.

(98) Jurečka, P.; Šponer, J.; Cerný, J.; Hobza, P. Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(99) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. Basis set convergence of the coupled-cluster correction, δ MP2CCSD(T): Best practices for benchmarking non-covalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases. *J. Chem. Phys.* **2011**, *135*, 194102.

- (100) Řezáč, J.; Riley, K. E.; Hobza, P. S66: A Well-Balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.
- (101) 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.
- (102) Manna, D.; Kesharwani, M. K.; Sylvetsky, N.; Martin, J. M. L. Conventional and Explicitly Correlated Ab Initio Benchmark Study on Water Clusters: Revision of the BEGDB and WATER27 Data Sets. *J. Chem. Theory Comput.* **2017**, *13*, 3136–3152.
- (103) Setiawan, D.; Kraka, E.; Cremer, D. Strength of the Pnictogen Bond in Complexes Involving Group Va Elements N, P, and As. *J. Phys. Chem. A* **2015**, *119*, 1642–1656.
- (104) Kozuch, S.; Martin, J. M. L. Halogen Bonds: Benchmarks and Theoretical Analysis. *J. Chem. Theory Comput.* **2013**, *9*, 1918–1931.
- (105) Řezáč, J.; Riley, K. E.; Hobza, P. Benchmark Calculations of Noncovalent Interactions of Halogenated Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 4285–4292.
- (106) Lao, K. U.; Schäffer, R.; Jansen, G.; Herbert, J. M. Accurate Description of Intermolecular Interactions Involving Ions Using Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2015**, *11*, 2473–2486.
- (107) Schwabe, T.; Grimme, S. Double-Hybrid Density Functionals With Long-Range Dispersion Corrections: Higher Accuracy and Extended Applicability. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- (108) Grimme, S. Seemingly Simple Stereoelectronic Effects in Alkane Isomers and the Implications for Kohn-Sham Density Functional Theory. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460–4464.
- (109) Gruzman, D.; Karton, A.; Martin, J. M. L. Performance of Ab Initio and Density Functional Methods for Conformational Equilibria of C_nH_{2n+2} Alkane Isomers ($n = 4–8$). *J. Phys. Chem. A* **2009**, *113*, 11974–11983.
- (110) Kesharwani, M. K.; Karton, A.; Martin, J. M. L. Benchmark Ab Initio Conformational Energies for the Proteinogenic Amino Acids Through Explicitly Correlated Methods. Assessment of Density Functional Methods. *J. Chem. Theory Comput.* **2016**, *12*, 444–454.
- (111) Řeha, D.; Valdés, H.; Vondrášek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; De Vries, M. S. Structure and IR Spectrum of Phenylalanyl-Glycyl-Glycine Tripeptide in the Gas-Phase: IR/UV Experiments, Ab Initio Quantum Chemical Calculations, and Molecular Dynamic Simulations. *Chem.—Eur. J.* **2005**, *11*, 6803–6817.
- (112) Goerigk, L.; Karton, A.; Martin, J. M. L.; Radom, L. Accurate Quantum Chemical Energies for Tetrapeptide Conformations: Why MP2 Data With an Insufficient Basis Set Should Be Handled With Caution. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7028–7031.
- (113) Fogueri, 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*, 2269–2277.
- (114) Csonka, G. I.; French, A. D.; Johnson, G. P.; Stortz, C. A. Evaluation of Density Functionals and Basis Sets for Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 679–692.
- (115) Kruse, H.; Mladek, A.; Gkionis, K.; Hansen, A.; Grimme, S.; Spöner, J. Quantum Chemical Benchmark Study on 46 RNA Backbone Families Using a Dinucleotide Unit. *J. Chem. Theory Comput.* **2015**, *11*, 4972–4991.
- (116) Kozuch, S.; Bachrach, S. M.; Martin, J. M. L. Conformational Equilibria in Butane-1,4-Diol: A Benchmark of a Prototypical System With Strong Intramolecular H-Bonds. *J. Phys. Chem. A* **2014**, *118*, 293–303.
- (117) Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Black, J. A.; Doll, K.; Heßelmann, A.; Kats, D.; Köhn, A.; Korona, T.; Kreplin, D. A.; et al. The Molpro Quantum Chemistry Package. *J. Chem. Phys.* **2020**, *152*, 144107.
- (118) 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.
- (119) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (120) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 Dispersion Coefficient Model. *J. Chem. Phys.* **2017**, *147*, 034112.
- (121) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-Charge Dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150*, 154122.
- (122) Ghasemi, S. A.; Hofstetter, A.; Saha, S.; Goedecker, S. Interatomic Potentials for Ionic Systems With Density Functional Accuracy Based on Charge Densities Obtained by a Neural Network. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2015**, *92*, 045131.
- (123) Grimme, S. *DFT-D3 V3.1*; University of Bonn, 2014, <https://www.chemie.uni-bonn.de/pctc/mulliken-center/software/dft-d3/Dft-d3>.
- (124) Caldeweyher, E.; Ehlert, S.; Grimme, S. *DFT-D4 V2.5.0*; University of Bonn, 2019, <https://www.chemie.uni-bonn.de/pctc/mulliken-center/software/dftd4>.
- (125) 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.
- (126) Valeev, E. F. Improving on the Resolution of the Identity in Linear R12 Ab Initio Theories. *Chem. Phys. Lett.* **2004**, *395*, 190–195.
- (127) Weigend, F. A. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.
- (128) 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.
- (129) Hättig, C. Optimization of Auxiliary Basis Sets for RI-MP2 and RI-CC2 Calculations: Core–Valence and Quintuple- ζ Basis Sets for H to Ar and QZVPP Basis Sets for Li to Kr. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.
- (130) 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.
- (131) 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.
- (132) 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.
- (133) Fogueri, 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*, 2269–2277.
- (134) 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.
- (135) Sylvetsky, 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.
- (136) 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.
- (137) Rappoport, D.; Furche, F. Property-Optimized Gaussian Basis Sets for Molecular Response Calculations. *J. Chem. Phys.* **2010**, *133*, 134105.
- (138) 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.

(139) Karton, A.; Martin, J. M. L. Comment on: "Estimating the Hartree-Fock limit from finite basis set calculations" [Jensen F (2005) *Theor Chem Acc* 113:267]. *Theor. Chem. Acc.* **2006**, *115*, 330–333.

(140) Boys, S.; Bernardi, F. the Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures With Reduced Errors. *Mol. Phys.* **1970**, *19*, 553–566.

(141) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing Counterpoise-Corrected, Uncorrected, and Averaged Binding Energies for Benchmarking Noncovalent Interactions. *J. Chem. Theory Comput.* **2014**, *10*, 49–57.

(142) 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.

(143) Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. L. The S66x8 Benchmark for Noncovalent Interactions Revisited: Explicitly Correlated Ab Initio Methods and Density Functional Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20905–20925.

(144) Kruse, H.; Grimme, S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *J. Chem. Phys.* **2012**, *136*, 154101.

(145) Witte, J.; Neaton, J. B.; Head-Gordon, M. Effective Empirical Corrections for Basis Set Superposition Error in the def2-SVPD Basis: gCP and DFT-C. *J. Chem. Phys.* **2017**, *146*, 234105.

(146) Brémond, E.; Li, H.; Sancho-García, J. C.; Adamo, C. Double Hybrids and Noncovalent Interactions: How Far Can We Go? *J. Phys. Chem. A* **2022**, *126*, 2590–2599.

(147) Sylvetsky, N.; Peterson, K. A.; Karton, A.; 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.

(148) Mehta, N.; Martin, J. M. L. Reduced-Scaling Double Hybrid DFT With Rapid Basis Set Convergence Through Localized Pair Natural Orbital F12. **2022**, Preprint arxiv.org/abs/2206.05072.

(149) Zhong, S.; Barnes, E. C.; Petersson, G. A. Uniformly convergent n -tuple- ζ augmented polarized (n ZaP) basis sets for complete basis set extrapolations. I. Self-consistent field energies. *J. Chem. Phys.* **2008**, *129*, 184116.

(150) Barnes, E. C.; Petersson, G. A.; Feller, D.; Peterson, K. A. The CCSD(T) Complete Basis Set Limit for Ne Revisited. *J. Chem. Phys.* **2008**, *129*, 194115.