The diene isomerization energies dataset: A difficult test for double-hybrid density functionals?

M. Wykes,[1,a)] A. J. Pérez-Jiménez,[2] C. Adamo,[3,4,5] and J. C. Sancho-García[2,b)]
[1]*Madrid Institute for Advanced Studies, IMDEA Nanoscience, Calle Faraday 9, Campus Cantoblanco, E-28049 Madrid, Spain*
[2]*Departamento de Química Física, Universidad de Alicante, E-03080 Alicante, Spain*
[3]*Institut de Recherche de Chimie Paris (IRCP), CNRS UMR-8247, Chimie ParisTech, École Nationale Superieure de Chimie de Paris, 11 rue P. et M. Curie, F-75231 Paris Cedex 05, France*
[4]*Institut Universitaire de France, 103 Boulevard Saint Michel, F-75005 Paris, France*
[5]*Istituto Italiano di Tecnologia, CompuNet, I-16163 Genova, Italy*

(Received 1 April 2015; accepted 22 May 2015; published online 10 June 2015)

We have systematically analyzed the performance of some representative double-hybrid density functionals (including PBE0-DH, PBE-QIDH, PBE0-2, XYG3, XYGJ-OS, and xDH-PBE0) for a recently introduced database of diene isomerization energies. Double-hybrid models outperform their corresponding hybrid forms (for example, PBE0-DH, PBE0-2, and PBE-QIDH are more accurate than PBE0) and the XYG3, XYGJ-OS, and xDH-PBE0 functionals perform excellently, providing root mean square deviation values within "calibration accuracy." XYGJ-OS and xDH-PBE0 also rival the best performing post-Hartree-Fock methods at a substantially lower cost. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4922058]

I. INTRODUCTION

The DIE60 database comprises a set of 60 diene isomerization energies shown in Figure 1 and has been recently added to the catalog of existing databases for benchmarking density functionals and wavefunction theories[1] against highly accurate data for π-conjugated systems.[2,3] The database includes linear (also branched) and cyclic dienes as reactants, and more specifically C_nH_{2n-2} and C_nH_{2n-4} compounds with $n = 5 - 7$, respectively, with products differing from reactants in a double-bond migration along the conjugated backbone. These reactions are also classified as hypohomodesmotic,[4] that is, having equal numbers of carbon atoms in their various possible hybridization states in both reactants and products, and are known to be highly challenging for quantum-chemical methods due to error cancellation issues, which might thus provide the right answer for the wrong reason, masking possibly insightful conclusions about the performance of theoretical methods unless systematic and hierarchical calculations are performed.

Interestingly, and very recently, reference reaction energies were obtained at the highly accurate W*n*-F12 method;[5] thus, paving the way towards the further assessment of a large number of cheaper methods in the search of a favourable trade-off (if any) between accuracy and computational cost. More specifically, density functionals belonging to all rungs of the hierarchy of existing methods (i.e., pure, hybrid, and double-hybrid expressions) have been extensively tested before.[1] For instance, selecting the BLYP, B3LYP, and B2-PLYP forms[6–10] as examples of pure, hybrid, and double-hybrid expressions and with large enough basis sets (cc-pVTZ/cc-pVQZ) to avoid basis sets incompleteness issues, the Mean Absolute Deviation (MAD) values for the DIE60 dataset are, respectively, 7.6, 6.0, and 3.9 kJ/mol, with the corresponding Root Mean Square Deviation (RMSD) values being 8.3, 6.6, and 4.2 kJ/mol. Comparing density functionals within the same rung, for example, the B3LYP, B3P86, and B3PW91 hybrid forms which differ only in the correlation functional employed, the MAD values are similar (6.0, 6.8, and 7.0 kJ/mol). However, this is not the case when the weight of the exact-like exchange (EXX) introduced into the hybrid form is varied, as it occurs for instance when comparing the B3LYP (0.2 of EXX) and BHHLYP (0.5 of EXX) forms, with MAD values of 6.0 and 3.8 kJ/mol, respectively. This general trend is preserved independently of the nature of the functionals selected: PBE and PBE0[11–13] (M06 and M06-2X[14]) were found to yield errors of 8.6 and 6.5 (5.6 and 2.5) kJ/mol, respectively. It thus seems that the weight of EXX plays a key role in determining the accuracy of results. Note also that intramolecular non-covalent interactions, successfully introduced through the D3 correction for dispersion,[15] were also ruled out as the main source of the reported discrepancies. Such corrections did however consistently improve results across all tested methods, reducing errors by approximately 1 kJ/mol or less with respect to dispersion-uncorrected values.

Upon a closer inspection of the results provided by all the tested double-hybrid forms, we can see unexpectedly large variations going from very small to large errors; the MAD obtained by B2GP-PLYP,[16] for example, is only 1.2 kJ/mol, and is thus close to the "calibration accuracy" of ±1 kJ/mol of error with respect to reference results, while the B2K-PLYP model[17] yields a MAD value of 8.8 kJ/mol. The remaining double-hybrids are within a 1–4 kJ/mol range of values, and thus are within the "chemical accuracy" range defined as ±1 kcal/mol of error with respect to reference results. For the

a)E-mail: michael.wykes@imdea.org
b)E-mail: jc.sancho@ua.es

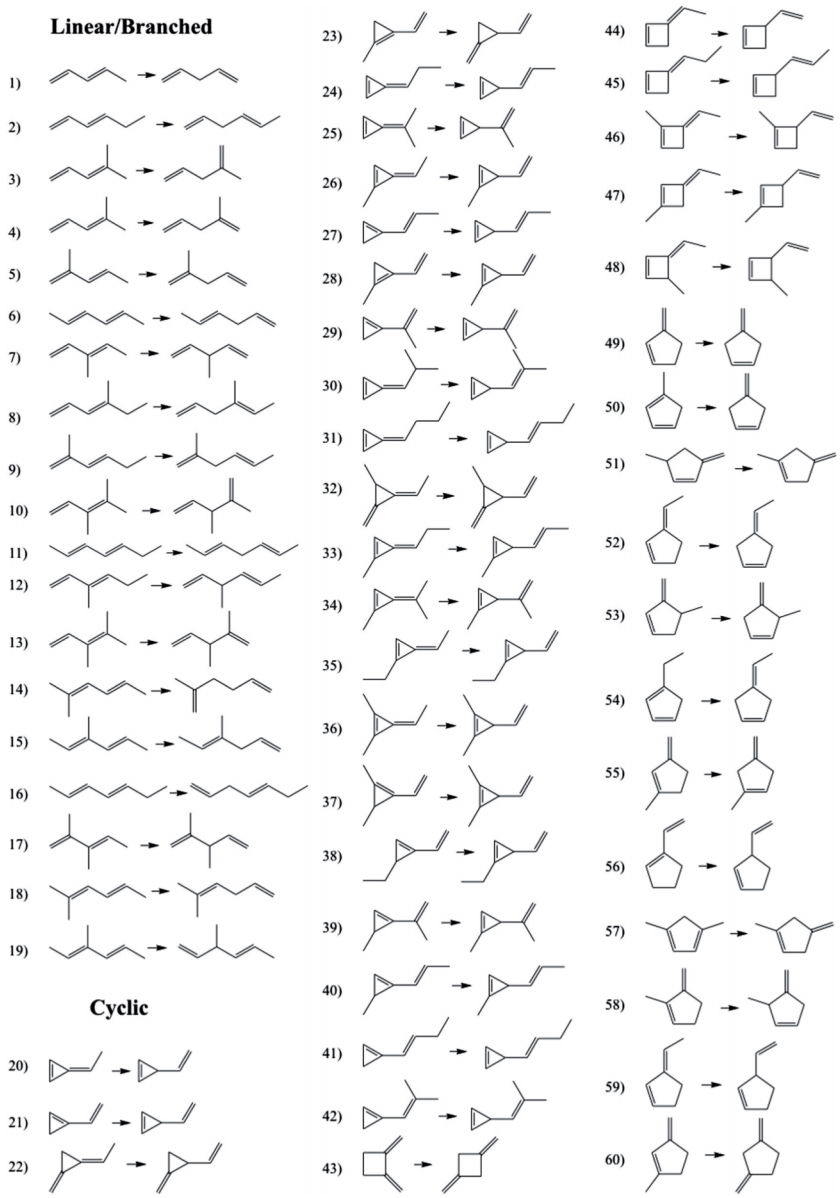

FIG. 1. Chemical structures of reactants and products used for the DIE60 isomerization reactions. The hydrogen atoms and corresponding C–H bonds are omitted for clarity. Reprinted with permission from L.-J. Yu and A. Karton, Chem. Phys. **441**, 166–177 (2015). Copyright 2015 by the Elsevier.

set of functionals analyzed thus far, not only the EXX weight but also the weight and form of the perturbative part entering into the expression is obtained by fitting against some training set of accurate data.[18] The particularly poor performance of B2K-PLYP was attributed to its specific parameterization against kinetic, and thus not to thermochemical data. However, double-hybrid forms which allow separate coefficients for same-spin and opposite-spin correlations, like DSD-BLYP[19] or DSD-PBEP86,[20] provided accurate values.

On the other hand, the families of parameter-free double-hybrid functionals (obviously less affected by parameterization issues) or X-based functionals (using a set of orbitals (density) arising from a standard B3LYP or PBE0 calculation) have not been tested yet on this challenging reactions set. Among them, we will select here some of the most modern expressions developed so far: PBE0-DH,[21] PBE0-2,[22] PBE-QIDH and TPSS-QIDH,[23] together with the XYG3,[24,25] XYGJ-OS,[26] and xDH-PBE0[27] models. In doing so, we aim to systematically investigate for these functionals: (i) the role played by intra-molecular interactions (by comparing the results of dispersion-corrected PBE0-DH-D3(BJ)[28] with those of uncorrected PBE0-DH); (ii) the role played by the nature of the exchange-correlation density functionals used (by studying PBE-QIDH and TPSS-QIDH models, the latter built with the TPSS exchange-correlation functional[29] instead of PBE); (iii) the role played by the underlying orbitals (density) used to evaluate the double hybrid energy (by analyzing the performance of the XYG3, XYGJ-OS, and xHD-PBE0 models, which use the converged orbitals of a B3LYP or PBE0 calculation for evaluating all the terms, including the EXX and perturbative energies[30]); (iv) the role played by the specific weight given to the various energy terms in the final expression, which might be related to the self-interaction error of common functionals;[31] and (v) the role played by different spin-scaling approaches applied to the perturbative term (e.g., by comparing the XYG3 and the XYGJ-OS models, the latter of which neglects same-spin correlation). In such a way, we hope to systematically disentangle the effect(s)

driving the final accuracy of the results, thus providing some robust guidelines for the future improvement of these two families of theoretical expressions.

II. THEORETICAL DETAILS

The starting point in the development of the parameter-free family of double-hybrid density functionals is the Adiabatic Connection Model (ACM), which defines the exchange-correlation contribution, E_{xc}, to the total Kohn-Sham (KS) energy as

$$E_{xc}[\rho] = \int_0^1 \mathcal{W}_{xc,\lambda}[\rho]d\lambda, \quad (1)$$

with

$$\mathcal{W}_{xc,\lambda}[\rho] = \langle\Psi_\lambda|\hat{V}_{ee}|\Psi_\lambda\rangle - \frac{1}{2}\int\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'d\mathbf{r}, \quad (2)$$

being $\hat{V}_{ee}$ the electron-electron interaction operator with associated mean value $\langle\Psi_\lambda|\hat{V}_{ee}|\Psi_\lambda\rangle$. The meaning of Ψ_λ is the wavefunction minimizing $\langle\Psi_\lambda|\hat{T}+\lambda\hat{V}_{ee}|\Psi_\lambda\rangle$. Note how the variable λ connects the non-interacting particle system (ideal system for which $\lambda = 0$) with the interacting one (real system for which $\lambda = 1$), assuming also that the external potential is adjusted to hold the electron density ρ fixed at its value for $\lambda = 1$. The mathematical form of $\mathcal{W}_{xc,\lambda}[\rho]$ needs to be conveniently fixed, and some examples already existed in the literature.[32–34] One of the latter is the recently developed quadratic function $\mathcal{W}_{xc,\lambda}[\rho] = a[\rho] + b[\rho]\lambda + c[\rho]\lambda^2$, which after imposing some known and realistic conditions leads to the final form of the PBE-QIDH model,

$$E_{xc}[\rho] = \lambda_x E_x[\phi]^{\mathrm{EXX}} + (1-\lambda_x)E_x[\rho] + \lambda_c E_c[\phi,\phi']^{\mathrm{PT2}} + (1-\lambda_c)E_c[\rho], \quad (3)$$

with the values found for λ_x and λ_c given in Table I. $E_x[\phi]^{\mathrm{EXX}}$ and $E_c[\phi,\phi']^{\mathrm{PT2}}$ are, respectively, the EXX and the perturbation correlation correction up to second-order (PT2), both built with the set of occupied (ϕ) and unoccupied (ϕ') orbitals of the corresponding Kohn-Sham solution. The PBE0-DH and PBE0-2 can be viewed as a simplification of the above function for $\mathcal{W}_{xc,\lambda}[\rho]$ together with a somehow different philosophy to obtain the final values for λ_i ($i = x, c$).[35]

TABLE I. Composition of the functionals used in this study (in chronological order following the year of its publication).

Name	λ_x	λ_c	Exchange	Correlation	Year
XYG3[a]	0.80	0.32	S+B88[b]	LYP	2009
XYGJ-OS[a]	0.77	0.44[c]	S	VWN+LYP[d]	2011
PBE0-DH	1/2	1/8	PBE	PBE	2011
PBE0-2	$2^{-1/3}$	0.50	PBE	PBE	2012
xDH-PBE0[e]	0.83	0.54[c]	PBE	PBE[f]	2012
PBE-QIDH	$3^{-1/3}$	1/3	PBE	PBE	2014
TPSS-QIDH	$3^{-1/3}$	1/3	TPSS	TPSS	2014

[a]Uses B3LYP orbitals.
[b]Uses 1-λ_x S and 0.2107 B88.
[c]Introduces perturbative correlation only between opposite-spin particles.
[d]Uses 0.2309 VWN and 0.2754 LYP.
[e]Uses PBE0 orbitals.
[f]Uses 0.5292 PBE correlation.

On the other hand, the X-based functionals XYG3, XYGJ-OS, and xDH-PBE0 models mainly differ in the set of orbitals used for the calculation of all the energy terms of Eq. (3). Whereas in the other PBE-based cases, one uses the orbitals obtained self-consistently using a truncated Kohn-Sham potential, i.e., discarding the perturbative part of Eq. (3) and thus building the corresponding exchange-correlation potential as $v_{xc}[\rho] = \lambda_x\frac{\delta E_x^{\mathrm{EXX}}}{\delta\rho} + (1-\lambda_x)\frac{\delta E_x[\rho]}{\delta\rho} + (1-\lambda_c)\frac{\delta E_c[\rho]}{\delta\rho}$; the X-based functionals employ orbitals arising from a complete (and fully converged) B3LYP solution (XYG3 and XYGJ-OS) or PBE0 solution (xDH-PBE0). A further distinction is that XYG3, XYGJ-OS, and xDH-PBE0 are partially fitted to reproduce accurate benchmark datasets. Table I summarizes the double-hybrid functionals selected for this study. Note also how orbital-optimized double-hybrid functionals have been recently developed,[36] which can be particularly interesting for open-shell systems, but are however not expected to significantly influence the conclusions reached here.

Other theoretical details are presented next: (i) the cc-pVTZ and cc-pVQZ basis sets were used to avoid basis set incompleteness issues mainly due to the $E_c[\phi,\phi']^{\mathrm{PT2}}$ term—the calculations can be thus safely considered as nearly converged (*vide infra*) at the cc-pVQZ level; (ii) the G09 program suite[37] was used for the PBE0-DH, PBE0-DH-D3(BJ), PBE0-2, PBE-QIDH, and TPSS-QIDH calculations, while the Firefly program[38] (which is partially based on the GAMESS-US source code[39]) was used for all XYG3, XYGJ-OS, xDH-PBE0, and all other X-based calculations; (iii) the geometries of all reactants and products were taken from the original study and rigidly used herein; (iv) we refer in the following to electronic-only energies and compare with the all-electron, vibrationless, and non-relativistic reference energies obtained at the W*n*-F12 level; and (v) the frozen core approximation was utilized for calculating PT2 energies. Gaussian calculations used ultrafine grids. All self-consistent energies (SCE) employed convergence criterion of 10^{-8} a.u. Results of new functionals differing from existing functionals only by the weighting of same- and opposite-spin PT2 energies were obtained by hand by reweighting the PT2 terms computed in either Firefly or Gaussian.

III. RESULTS AND DISCUSSION

The performance of all the selected double-hybrid density functionals is gathered in Table II,[40] together with the results obtained in previous studies at the Hartree-Fock (HF), second-order Møller-Plesset perturbation theory (MP2), and its spin-component-scaled (SCS) version (SCS-MP2).[41] Note that most of the methods we will discuss formally scale as $O(N^5)$ (though xDH-PBE0 and XYGJ-OS can be made to scale as $O(N^4)$ and $O(N^3)$, respectively), where N is related to the system size, selectively discarding other more costly methods. We choose as a threshold (Δ_e) for recommending a functional: Δ_e = RMSD ≤ 2.1 kJ/mol (half the "chemical accuracy" threshold of 4.2 kJ/mol). Note that the canonical MP2 method has a RMSD of only 2.7 kJ/mol, and thus is close to that threshold at roughly the same formal computational cost

TABLE II. Average errors (kJ/mol) for the set of double-hybrid density functionals initially considered. The cc-pVTZ and cc-pVQZ basis sets were used.

	Name	RMSD[a]	MAD[a]	MSD[a]
cc-pVTZ	HF[b]	5.0	3.9	−0.7
	MP2[b]	2.8	2.3	2.3
	SCS-MP2[b]	1.1	0.8	0.3
	PBE0-DH	5.67	4.93	4.82
	PBE0-DH-D3(BJ)	5.21	4.50	4.44
	PBE0-2	3.65	3.29	3.28
	PBE-QIDH	4.35	3.84	3.78
	TPSS-QIDH	4.38	3.86	3.82
	XYG3	2.28	2.17	2.17
	XYGJ-OS	1.05	0.79	0.41
	xDH-PBE0	1.17	0.90	0.37
cc-pVQZ	HF[b]	5.1	3.8	−0.9
	MP2[b]	2.7	2.2	2.2
	SCS-MP2[b]	1.0	0.6	0.3
	PBE0-DH	5.54	4.78	4.64
	PBE0-DH-D3(BJ)	5.08	4.35	4.26
	PBE0-2	3.61	3.23	3.19
	PBE-QIDH	4.29	3.77	3.67
	TPSS-QIDH	4.29	3.76	3.69
	XYG3	2.20	2.08	2.08
	XYGJ-OS	1.00	0.72	0.35
	xDH-PBE0	1.13	0.83	0.35

[a]RMSD means "root mean square deviation," MAD means "mean absolute deviation," and MSD means "mean signed deviation.".
[b]Results taken from Ref. 1.

as any density functional including a perturbative correction, with perhaps the exception of the xDH-PBE0 and XYGJ-OS models due to the algorithm advantage of computing only the opposite-spin contribution to the correlation energy.[26,42] This accuracy (Δ_e) will be thus considered as the envisioned target for the double-hybrid functionals studied here. In order to get a handle on the impact of the different effects analyzed, we will dub any effect "very marked," "marked," or "small" if the corresponding RMSD value is reduced by at least $\frac{1}{2}\Delta_e \approx 1.0$ kJ/mol, $\frac{1}{4}\Delta_e \approx 0.5$ kJ/mol, or less than the latter value, respectively.

We would first like to bracket the effect of intra-molecular dispersion interactions, by comparing the PBE0-DH and PBE0-DH-D3(BJ) RMSD values, estimated here to be the order of 0.5 kJ/mol independently of basis sets size, see Table II, and thus classified to be in the limit of being marked. By inspecting the difference between cc-pVTZ and cc-pVQZ results, it can be concluded that the latter basis set reduces errors by approximately 0.1 kJ/mol, and hence, that basis-set extension beyond cc-pVQZ is not needed. Unless otherwise mentioned, all the discussion in the following refers to cc-pVQZ results. Furthermore, comparison of PBE-QIDH and TPSS-QIDH error metrics shows that variation of the underlying parameter-free exchange-correlation functional (i.e., PBE or TPSS) has a very limited impact. Thus, effects related to basis sets and the underlying exchange-correlation functionals are considered small in the following and will not be discussed further.

We report next in Figure 2 how the RMSD values for the DIE60 dataset evolve as a function of the weight of the EXX introduced into the exchange-correlation functional, i.e., the value of λ_x in Eq. (3) for the double-hybrid forms belonging to the PBE-based family of functionals. PBE, a pure (non-hybrid) functional, for which $\lambda_x = 0$ by definition, provides the largest error of 9.4 kJ/mol; PBE0 is the well-known hybrid functional with $\lambda_x = 0.25$ for which the error is very markedly reduced (relative to PBE) to 7.3 kJ/mol. This behavior prompted us to study this effect further: PBE0-DH ($\lambda_x = 0.50$) leads to a further marked decrease with respect to PBE0 (from 7.3 to 5.5 kJ/mol). Furthermore, the remaining PBE-based double-hybrid forms included in this study (see Table I for λ_x values) also perfectly fit into this trend (see Figure 2). Linear regression yields the result RMSD = $9.26 - 7.23\lambda_x$ ($r^2 = 0.9986$) for which extrapolation to $\lambda_x = 1$ predicts a RMSD value of only 2.0 kJ/mol, and thus substantially lower than that

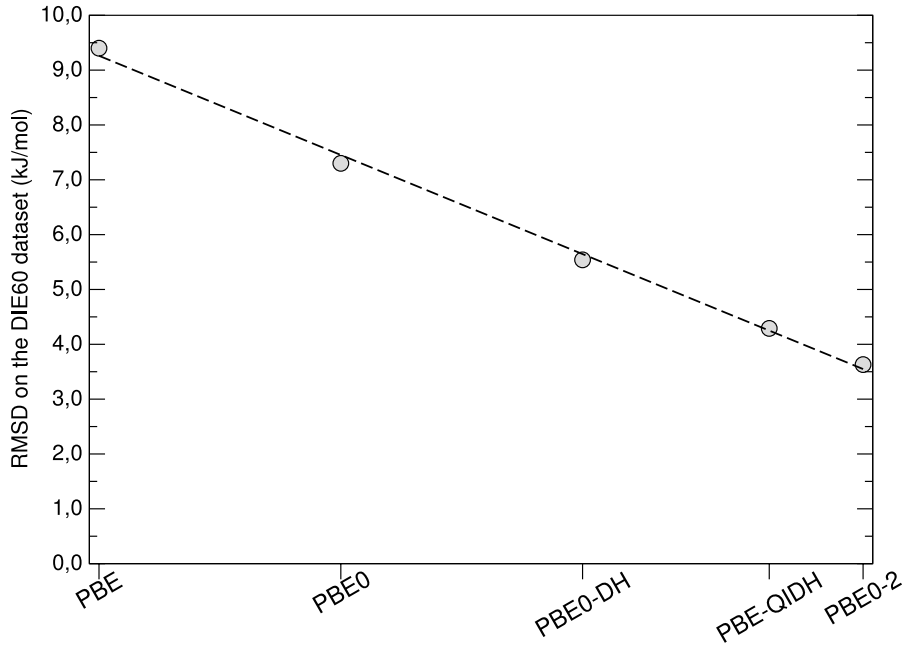

FIG. 2. Evolution of the RMSD for the DIE dataset as a function of the EXX weight of the PBE-based family of functionals: PBE ($\lambda_x = 0$), PBE0 ($\lambda_x = 0.25$), PBE0-DH ($\lambda_x = 0.50$), PBE-QIDH ($\lambda_x = 0.693$), and PBE0-2 ($\lambda_x = 0.79$). The dashed straight line is the result of the linear regression performed (see main text for coefficients).

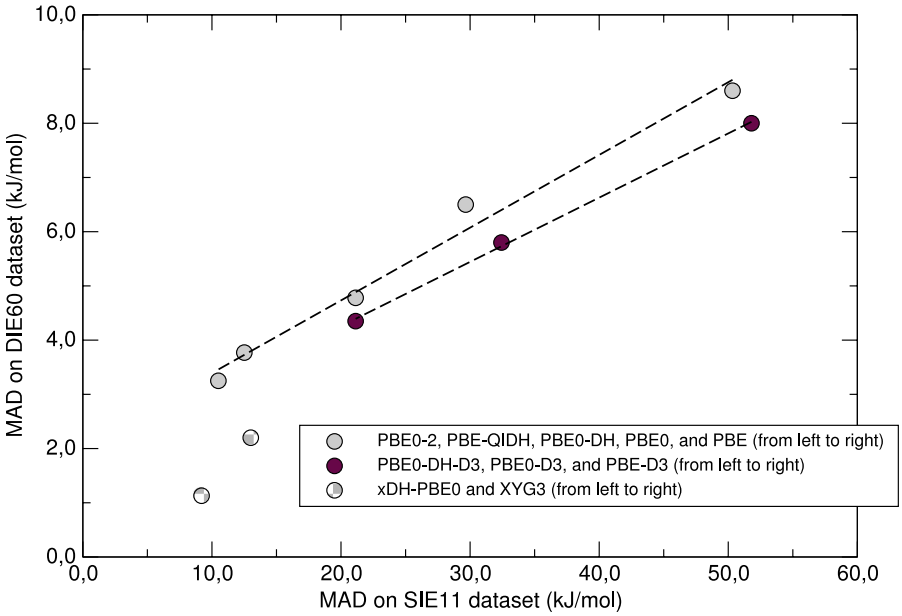

FIG. 3. Correlation between the MAD for the DIE60 and SIE11 datasets. The dashed straight line is the result of the linear regression performed for the PBE-based set of functionals.

provided by the HF method alone (5.1 kJ/mol) indicating that the correlation correction also plays a leading role.

Note that for values of λ_x corresponding to XYG3 ($\lambda_x = 0.8033$), XYGJ-OS ($\lambda_x = 0.7731$), and xDH-PBE0 ($\lambda_x = 0.8335$), the previous fit would predict RMSD values between 3.2 and 3.6 kJ/mol, and thus larger than those finally obtained (2.2, 1.0, and 1.1 kJ/mol, respectively, see Table II) indicating that in addition to large EXX weights, other factors must also contribute to the improved performance of these models, again highlighting the importance of PT2 correlation in determining the achieved accuracy. In fact, these three models (XYG3, XYGJ-OS, and xDH-PBE0) outperform the accuracy of MP2, with the latter two rivalling the best performing $O(N^5)$ scaling method, SCS-MP2 (RMSD of just 1.0 kJ/mol) and even outperforming more costly methods scaling as $O(N^6)$ such as MP3 and CCSD.[1] Comparing results of XYG3 and XYGJ-OS (both of which employ B3LYP orbitals in their PT2 and large weights of EXX) highlights the large impact of different spin-scaling approaches; neglecting same-spin correlation and scaling only the opposite-spin contribution in XYGJ-OS have a very marked influence (in the right direction) on final accuracy. Additional systematic investigations of the impact of PT2 spin-scaling within PBE-based functionals support this conclusion (*vide infra*).

The qualitative correlation of RMSD with λ_x values resembles that found in previous studies analyzing the behavior of double-hybrid functionals with respect to the underlying self-interaction error (SIE).[43–45] To confirm this point, in Figure 3, we present MAD values for the DIE60 and the SIE11 datasets, the latter of which was specifically designed to test self-interaction errors.[46] Indeed, we find a large correlation between the performance of density functionals for both datasets, independently of dispersion-corrections and/or basis sets' issues, with the XYG3 and xDH-PBE0 functionals behaving as outliers of trends found for the PBE-based family of functionals.

To investigate this behavior further, we will systematically compare PBE0-2 and xDH-PBE0 models, which have rather similar weights of λ_x and λ_c values (see Table I) but different types of MP2 term: PBE0-2 uses MP2 canonical expression with orbitals arising from the underlying (truncated) exchange-correlation potential, while xDH-PBE0 uses SOS-MP2 scaling, neglecting same-spin correlation, together with PBE0 self-consistent orbitals. Starting from these models, we can modify them to have: (i) a DH functional based on xDH-PBE0 but with slightly modified self-consistent orbitals, arising from the PBE0-1/3 hybrid functional ($\lambda_x = 1/3$, see Ref. 12) instead of the original PBE0 ($\lambda_x = 1/4$) one, dubbed as xDH-PBE0-1/3; (ii) a DH functional based on xDH-PBE0, but employing self-consistent orbitals of the truncated xDH-PBE0 v_{xc} in the PT2 term, dubbed SC-xDH-PBE0; (iii) a DH functional based on xDH-PBE0 but applying now the λ_c scaling to both opposite- and same-spin contributions, which will be dubbed as MP2-xDH-PBE0; (iv) a DH functional based

TABLE III. Composition of the double-hybrid density functionals newly composed.

Name	λ_x	$\lambda_c^{\text{opposite-spin}}$	$\lambda_c^{\text{same-spin}}$	Exchange	Correlation	Orbitals
xDH-PBE0	0.83	0.54	...	PBE	PBE	PBE0
xDH-PBE0-1/3	0.83	0.54	...	PBE	PBE	PBE0-1/3
SC-xDH-PBE0	0.83	0.54	...	PBE	PBE	Truncated v_{xc}
MP2-xDH-PBE0	0.83	0.54	0.54	PBE	PBE	PBE0
PBE0-2	0.79	0.50	0.50	PBE	PBE	Truncated v_{xc}
SOS-PBE0-2	0.79	0.50	...	PBE	PBE	Truncated v_{xc}
SOS-xDH-PBE0-2	0.79	0.50	...	PBE	PBE	PBE0

TABLE IV. Statistical analysis (kJ/mol) for the set of newly composed double-hybrid density functionals. The cc-pVQZ basis set was used.

Name	RMSD[a]	MAD[a]	MSD[a]
xDH-PBE0	1.13	0.83	0.35
xDH-PBE0-1/3	1.46	1.02	0.59
SC-xDH-PBE0	2.59	1.76	1.11
MP2-xDH-PBE0	4.21	3.77	3.75
PBE0-2	3.61	3.23	3.19
SOS-PBE0-2	2.72	1.91	1.41
SOS-xDH-PBE0-2	1.38	0.99	0.66

[a]RMSD means "root mean square deviation," MAD means "mean absolute deviation," and MSD means "mean signed deviation.".

on PBE0-2 but neglecting the same-spin correlation, keeping thus only the opposite-spin scaling but without reweighting it, dubbed as SOS-PBE0-2; and (v) a DH functional based on SOS-PBE0-2 but with the PBE0 self-consistent orbitals, dubbed as SOS-xDH-PBE0-2. Table III summarizes the new set of double-hybrid functionals used for this part of the study, with the statistical metrics for their performance shown in Table IV.

Perusing Table IV, we can observe the large interplay found between opposite-spin scaling and the set of orbitals used. In other words, one can imagine that any double-hybrid expression might have potentially two distinct sources of error: the error due to the energy approximation itself and the error due to the corresponding approximate orbitals (density) used therein, as it has been recently emphasized.[47] Indeed, comparing PBE0-2 and SOS-PBE0-2 data, the former effect can be classified as very marked, reducing the RMSD from 3.63 to 2.72 kJ/mol. This effect is further confirmed when going from xDH-PBE0 to MP2-xDH-PBE0, which undoes the opposite-spin scaling and dramatically increases the RMSD from 1.13 to 4.21 kJ/mol. It seems that standard double-hybrid models might suffer from some double-counting of same-spin correlation, and thus neglecting that contribution to the perturbative energy term would improve accuracy. On the other hand, when PBE0 orbitals feed the SOS-PBE0-2 model, becoming thus SOS-xDH-PBE0-2 according to our notation, the RMSD still decreases passing from 2.72 to 1.38 kJ/mol, with this effect being very marked again. Finally, using PBE0-1/3 orbitals rather than those obtained from PBE0 for the xDH-PBE0 model slightly increases the error bars relative to xDH-PBE0.

IV. CONCLUSIONS

We have benchmarked a set of double-hybrid density functionals from the recent literature: PBE0-DH, PBE-QIDH, TPSS-QIDH, PBE0-2, XYG3, XYGJ-OS, and xDH-PBE0, with respect to isomerization energies of a large set of diene reactions, involving some double-bond migration and thus constituting a database for π-conjugation effects. Our results show that, on this dataset, any of the double-hybrid functionals tested, including the parameter-free PBE0-DH and PBE-QIDH examples, systematically improves accuracy relative to corresponding hybrid methods, i.e., the gold-standard PBE0 functional, consistently providing lower RMSD and MAD errors and rivalling the accuracy of other more empirical double-hybrid methods.

Particularly, striking is the good performance of the XYGJ-OS and xDH-PBE0 methods, reaching the so-called "calibration accuracy" (an error deviation of around ±1 kJ/mol with respect to reference values) and rivalling the accuracy of (more computationally expensive) wavefunction (post-HF) methods. We have also systematically investigated the origin of this accuracy, trying to disentangle the effects of the various ingredients (orbitals, functionals, correlation contributions, weights of every term, etc.) entering into the xDH-PBE0 formulation. We conclude that not only the orbitals used to calculate the $E_c[\phi,\phi']^{PT2}$ part but also the neglect of its same-spin correlation term (which reduces its formal scaling from $O(N^5)$ to $O(N^4)$) is key to its excellent accuracy. The threshold (Δ_e) imposed at the beginning of the work, to investigate if any DH density functional belonging to the parameter-free or the X-based families might perform better than the MP2 method itself (as measured by a RMSD below half the "chemical accuracy" value of 2.1 kJ/mol), is fulfilled by the original xDH-PBE0 model, but also by some variants (e.g., xDH-PBE0-1/3 and SOS-xDH-PBE0-2) specifically devised here to analyze the concurring effects, as well as by the closely related XYGJ-OS model, with the XYG3 lying at the limit of it.

It thus seems that research on double-hybrid density functionals will continue to pave the way towards more accurate and less costly methods, which combine the best of both wavefunction and density functional worlds, and will stimulate (thanks to improved accuracy at low-cost) more applied studies on π-conjugated molecules, especially once algorithmic advantages of neglecting same-spin correlation are fully exploited in widely available software.

ACKNOWLEDGMENTS

The work in Alicante is supported by the "Ministerio de Economía y Competitividad" of Spain and the "European Regional Development Fund" through Project No. CTQ2014-55073P. The work at IMDEA was supported by the Campus of International Excellence (CEI) UAM+CSIC. M.W. thanks the European Commission for his Marie Curie Fellowship (Grant No. FP7-PEOPLE-2012-IEF-331795).

[1]L.-J. Yu and A. Karton, Chem. Phys. **441**, 166 (2014).
[2]L.-J. Yu, F. Sarrami, A. Karton, and R. J. O'Reilly, Mol. Phys. **113**, 1284 (2015).
[3]Y. Zhao and D. G. Truhlar, J. Phys. Chem. A **110**, 10478 (2006).
[4]S. E. Wheeler, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 204 (2012).
[5]A. Karton and J. M. L. Martin, J. Chem. Phys. **136**, 124114 (2012).
[6]A. D. Becke, Phys. Rev. A **38**, 3098 (1998).
[7]C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[8]A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[9]V. Barone and C. Adamo, Chem. Phys. Lett. **224**, 432 (1994).
[10]S. Grimme, J. Chem. Phys. **124**, 034108 (2006).
[11]J. P. Perdew, M. Ernzenhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[12]C. A. Guido, E. Brémond, C. Adamo, and P. Cortona, J. Chem. Phys. **138**, 021104 (2013).
[13]C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).
[14]Y. Zhao and D. G. Truhlar, Acc. Chem. Res. **41**, 157 (2008).
[15]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[16]A. Karton, A. Tarnopolsky, J.-F. Lamère, G. C. Schatz, and J. M. L. Martin, J. Phys. Chem. A **112**, 12868 (2008).

[17]A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman, and J. M. L. Martin, J. Phys. Chem. A **112**, 3 (2008).
[18]L. Goerigk and S. Grimme, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 576 (2014).
[19]S. Kozuch, D. Gruzman, and J. M. L. Martin, J. Phys. Chem. C **114**, 20801 (2010).
[20]S. Kozuch and J. M. L. Martin, Phys. Chem. Chem. Phys. **13**, 20104 (2011).
[21]E. Brémond and C. Adamo, J. Chem. Phys. **135**, 024106 (2011).
[22]J.-D. Chai and S.-P. Mao, Chem. Phys. Lett. **538**, 121 (2012).
[23]E. Brémond, J. C. Sancho-García, A. J. Pérez-Jiménez, and C. Adamo, J. Chem. Phys. **141**, 031101 (2014).
[24]Y. Zhang, X. Xu, and W. A. Goddard III, Proc. Natl. Acad. Sci. U. S. A. **106**, 4963 (2009).
[25]Y. Zhang, J. Wu, and X. Xu, Chem. Commun. **46**, 3057 (2010).
[26]Y. Zhang, X. Xu, Y. Jung, and W. A. Goddard III, Proc. Natl. Acad. Sci. U. S. A. **108**, 19896 (2011).
[27]I. Y. Zhang, N. Q. Su, E. Brémond, C. Adamo, and X. Xu, J. Chem. Phys. **136**, 174103 (2012).
[28]D. Bousquet, E. Brémond, J. C. Sancho-García, I. Ciofini, and C. Adamo, Theor. Chem. Acc. **134**, 1602 (2015).
[29]J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
[30]I. Y. Zhang and X. Xu, Int. Rev. Phys. Chem. **30**, 115 (2011).
[31]P. Mori-Sánchez, A. J. Cohen, and W. Yang, J. Chem. Phys. **125**, 201102 (2006).
[32]M. J. G. Peach, A. M. Teale, and D. J. Tozer, J. Chem. Phys. **126**, 244104 (2007).
[33]A. J. Cohen, P. Mori-Sánchez, and W. Yang, J. Chem. Phys. **127**, 034101 (2007).
[34]M. J. G. Peach, A. M. Miller, A. M. Teale, and D. J. Tozer, J. Chem. Phys. **129**, 064105 (2008).
[35]J. C. Sancho-García and C. Adamo, Phys. Chem. Chem. Phys. **15**, 14581 (2013).
[36]R. Peverati and M. Head-Gordon, J. Chem. Phys. **139**, 024110 (2013).
[37]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, GAUSSIAN 09, Revision D.01, Gaussian, Inc., Wallingford, CT, 2009.
[38]A. A. Granovsky, Firefly version 8, http://classic.chem.msu.su/gran/firefly/index.html.
[39]M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[40]See supplementary material at http://dx.doi.org/10.1063/1.4922058 for tabulated data and associated errors for all the methods employed.
[41]S. Grimme, J. Chem. Phys. **118**, 9095 (2003).
[42]Y. Jung, R. C. Lochan, A. D. Dutoi, and M. Head-Gordon, J. Chem. Phys. **121**, 9793 (2004).
[43]J. C. Sancho-García, J. Chem. Phys. **134**, 234102 (2011).
[44]M. Wykes, N. Q. Su, X. Xu, C. Adamo, and J. C. Sancho-García, J. Chem. Theory Comput. **11**, 832 (2015).
[45]N. Q. Su and X. Xu, J. Phys. Chem. A **119**, 1590 (2015).
[46]L. Goerigk and S. Grimme, J. Chem. Theory Comput. **6**, 107 (2010).
[47]M.-C. Kim, E. Sim, and K. Burke, J. Chem. Phys. **140**, 18A528 (2014).