JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau
Article

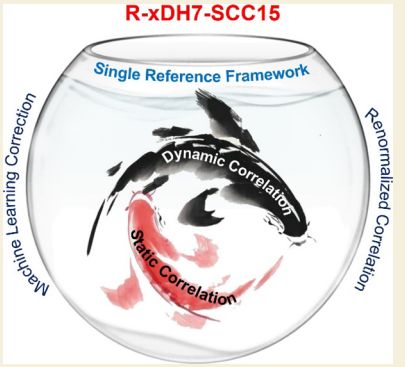
Toward Efficient and Unified Treatment of Static and Dynamic Correlations in Generalized Kohn–Sham Density Functional Theory

Yizhen Wang, Zihan Lin, Runhai Ouyang, Bin Jiang, Igor Ying Zhang,* and Xin Xu*

Cite This: JACS Au 2024, 4, 3205–3216
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Accurate description of the static correlation poses a persistent challenge in electronic structure theory, particularly when it has to be concurrently considered with the dynamic correlation. We develop here a method in the generalized Kohn–Sham density functional theory (DFT) framework, named R-xDH7-SCC15, which achieves an unprecedented accuracy in capturing the static correlation, while maintaining a good description of the dynamic correlation on par with the state-of-the-art DFT and wave function theory methods, all grounded in the same single-reference black-box methodology. Central to R-xDH7-SCC15 is a general-purpose static correlation correction (SCC) model applied to the renormalized XYG3-type doubly hybrid method (R-xDH7). The SCC model development involves a hybrid machine learning strategy that integrates symbolic regression with nonlinear parameter optimization, aiming to achieve a balance between generalization capability, numerical accuracy, and interpretability. Extensive benchmark studies confirm the robustness and broad applicability of R-xDH7-SCC15 across a diverse array of main-group chemical scenarios. Notably, it displays exceptional aptitude in accurately characterizing intricate reaction kinetics and dynamic processes in regions distant from equilibrium, where the influence of static correlation is most profound. Its capability to consistently and efficiently predict the whole energy profiles, activation barriers, and reaction pathways within a user-friendly "black-box" framework represents an important advance in the field of electronic structure theory.

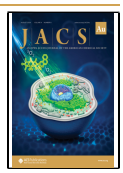


KEYWORDS: *density functional theory, XYG3-type doubly hybrid functional, renormalized PT2 correlation model, static correlation effect, machine learning*

Kohn–Sham density functional theory (KS-DFT), grounded in single-reference black-box methodology, has emerged as the preeminent electronic-structure framework for analyzing large-scale systems across diverse scientific fields.[1−6] In KS-DFT, density functional approximations (DFAs) form a development hierarchy, known as the Jacob's Ladder, stratified into five rungs reflecting increasing complexity toward chemical accuracy.[7] Doubly hybrid (DH) approximations, which stand on the top rung (the fifth rung) of the Jacob's Ladder, have demonstrated significant superiority over their popular lower-rung counterparts, particularly for properties dominated by the dynamic correlation.[8−24] Despite this significant progress, accurately capturing the static correlation remains an unsolved challenge in KS-DFT,[4,25,26] necessitating computationally demanding multideterminant wave function treatments that are nontrivial for routine applications.[1,27]

In this study, we start with the development of a renormalized XYG3-type doubly hybrid method, designated as R-xDH7, which utilizes B3LYP density and orbitals, and incorporates seven empirical parameters through an integration of a spin-distinctive random-phase approximation (RPA)-type renormalization of the second-order perturbative contribution (PT2).[5] The R-xDH7 method encapsulates a substantial portion of the static correlation alongside a broad range of the dynamic correlation toward chemical accuracy. Based on this foundation, we further devise a general-purpose term for the residual static correlation correction specially adapted to R-xDH7 with 15 parameters (SCC15), utilizing a hybrid machine learning algorithm that synergistically combines symbolic and nonlinear parameter regressions. The working scheme is illustrated in Figure 1 and implemented in the Rust-based Electronic Structure Toolkit (REST).[28] The final R-xDH7-SCC15 method achieves an unprecedented balance between accuracy in describing both dynamic and static correlations. Its robustness is further underscored by the outstanding performance across a comprehensive benchmark suite for main-group chemistry, displaying its exceptional aptitude in providing precise characterizations of complex reaction mechanisms

Received: June 10, 2024
Revised: July 26, 2024
Accepted: August 6, 2024
Published: August 12, 2024



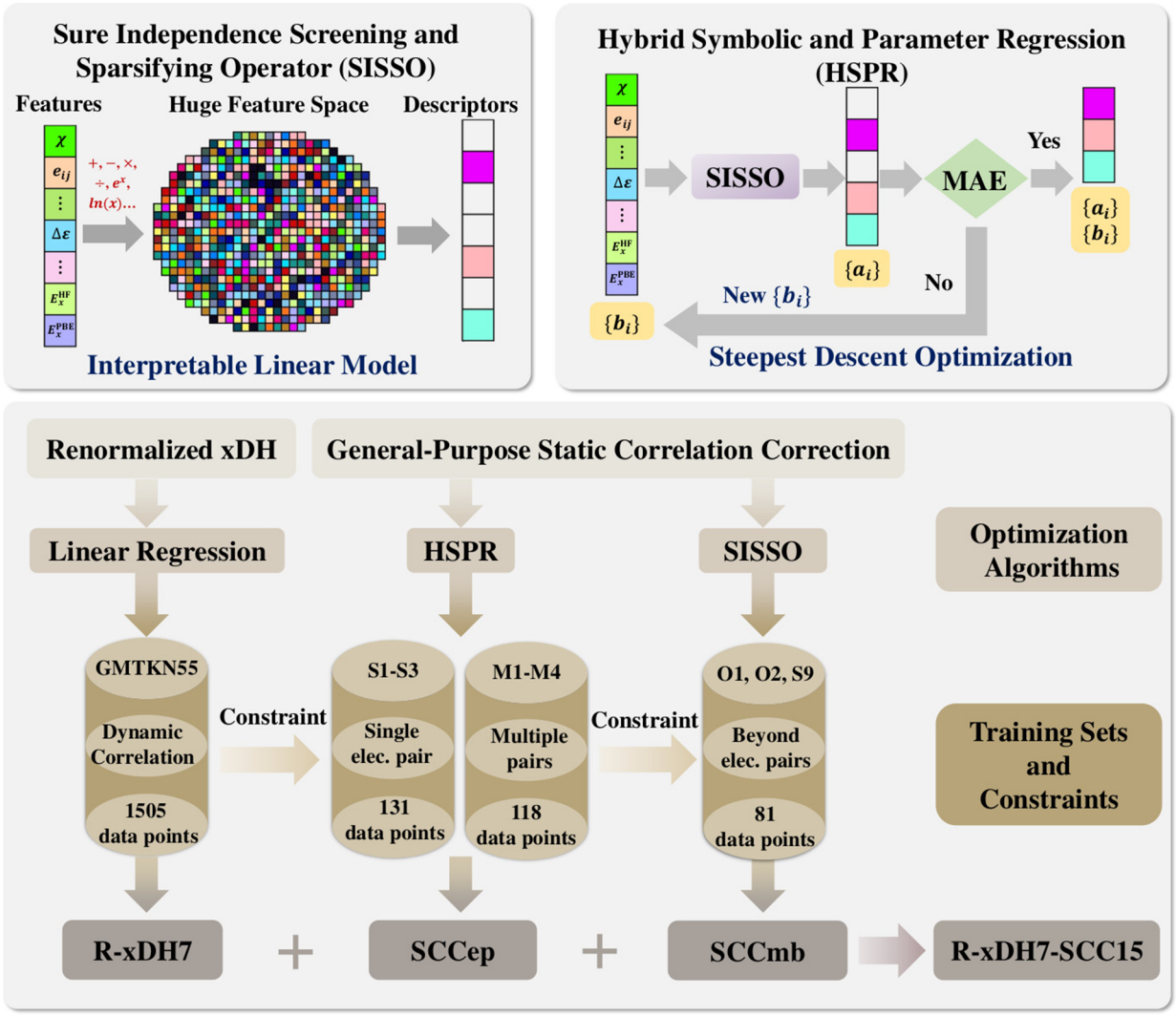


Figure 1. Development strategy for the renormalized XYG3-type doubly hybrid approximation (R-xDH7) and the machine learned static correlation correction (SCC15). Only a subset of 10 out of total 23 bond dissociation curves within the MBD23−594 (refer to Figure 2 for more details) were incorporated into the training set or utilized as constraints during the establishment of the SCC15 model.

beyond the equilibrium regions where static correlation effects are significant.

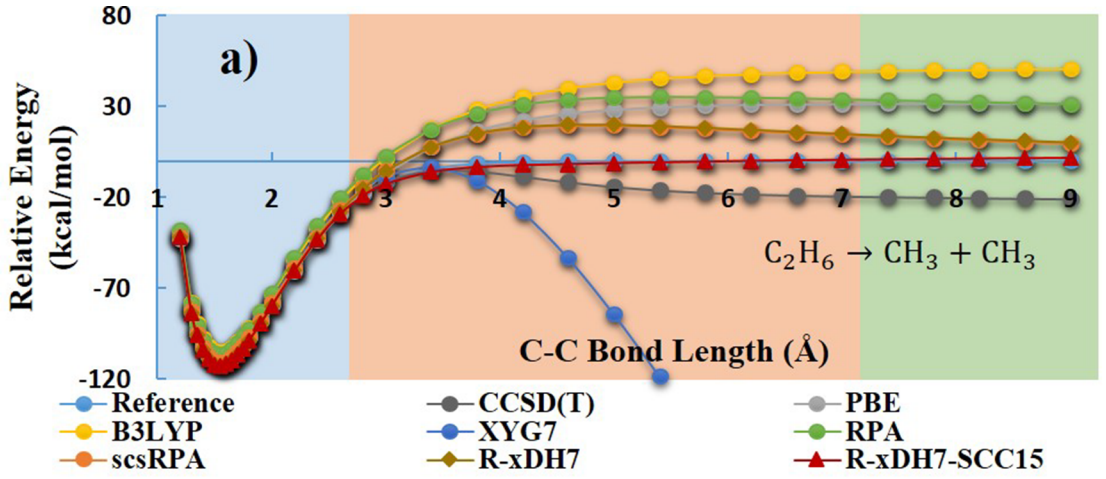
The XYG3-type doubly hybrid approximations (xDH)[12,17,24,29] have garnered widespread recognition for their ability of accurately and efficiently describing dynamic correlation effects.[30−35] This is exemplified by XYG7[36] recently developed by our group. Some other successful examples that follow the same construction strategy as xDH are ωB97M(2)[22] and xrevDSD-PBEB86-D4,[37] independently developed by the Head-Gordon and Martin research groups, respectively. The use of the PT2 term is a key that accounts for the notably good performance of all doubly hybrid functionals.[23,29] Despite this advancement, PT2 is notoriously inadequate for (near)-degenerated systems, where static correlation effects are predominant.[4] As shown in Figure 2, the XYG7 results diverge during molecular bond dissociation processes. To address this shortcoming, we have previously explored enhancements to the PT2 correlation model, which included the integration of the screened second-order Bethe-Goldstone equation (sBGE2)[6] and a spin-distinctive random-phase approximation (RPA)-type renormalization of PT2.[5] These novel correlation models have demonstrated their efficacy in resolving the inherent divergence of PT2 in (near)-degenerated systems.[4]

Building upon these advancements, we introduced here a renormalized xDH method, R-xDH7, which, akin to the standard xDH method of XYG7, utilized the B3LYP[38−41] density and orbitals, but improved the standard PT2 by the spin-distinctive RPA-type renormalization[5]:

$$E_{xc}^{R-xDH7} = a_1 E_x^{HF} + a_2 E_x^{S} + a_3 E_x^{B88} + a_4 E_c^{VWN} + a_5 E_c^{LYP} + a_6 E_c^{osRPT2} + a_7 E_c^{ssRPT2+} \quad (1)$$

Here E_x^{HF} is the contribution from the Hartree−Fock (HF)-like exact exchange, while E_x^{S} and E_x^{B88} denote the exchange contributions from the Slater-type local density approximation (LDA)[42] and the Becke88 generalized gradient approximation (GGA),[38] respectively. E_c^{VWN} is the local Vosko-Wilk-Nusair correlation.[43] The R-xDH7 method adhered to the specific VWN-III version utilized in the original B3LYP implementation to maintain consistency.[41] E_c^{LYP} is the Lee−Yang−Parr correlation approximation.[39] Specifically, E_c^{osRPT2} stands for the RPA-type renormalization of opposite-spin PT2, derived from the spin-pair distinctive adiabatic connection fluctuation−

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216



Close-shell Single-bond Dissociations (CSD9-256)

S1	$H_2 \rightarrow H + H$	S2	$C_2H_6 \rightarrow CH_3 + CH_3$	S3	$F_2 \rightarrow F + F$
S4	$CH_4 \rightarrow CH_3 + H$	S5	$NH_3 \rightarrow NH_2 + H$	S6	$H_2O \rightarrow OH + H$
S7	$HF \rightarrow H + F$	S8	$CH_3F \rightarrow CH_3 + F$	S9	$CH_3Cl \rightarrow CH_3 + Cl$

Close-shell Multiple-bond Dissociations (CMD8-204)

M1	$N_2 \rightarrow N + N$	M2	$C_2 \rightarrow C + C$
M3	$CH_4 \rightarrow C + 4H$	M4	$NH_3 \rightarrow N + 3H$
M5	$HCN \rightarrow H + CN$	M6	$PH_3 \rightarrow P + 3H$
M7	$H_2O \rightarrow O + 2H$	M8	$H_2S \rightarrow S + 2H$

Open-shell Bond Dissociations (OBD6-134)

O1	$CH_3 \rightarrow CH_2 + H$
O2	$NH_2 \rightarrow NH + H$
O3	$CH_2 \rightarrow CH + H$
O4	$NH \rightarrow N + H$
O5	$OH \rightarrow O + H$
O6	$NO \rightarrow N + O$

b) Main-group Bond Dissociations (MBD23-594)

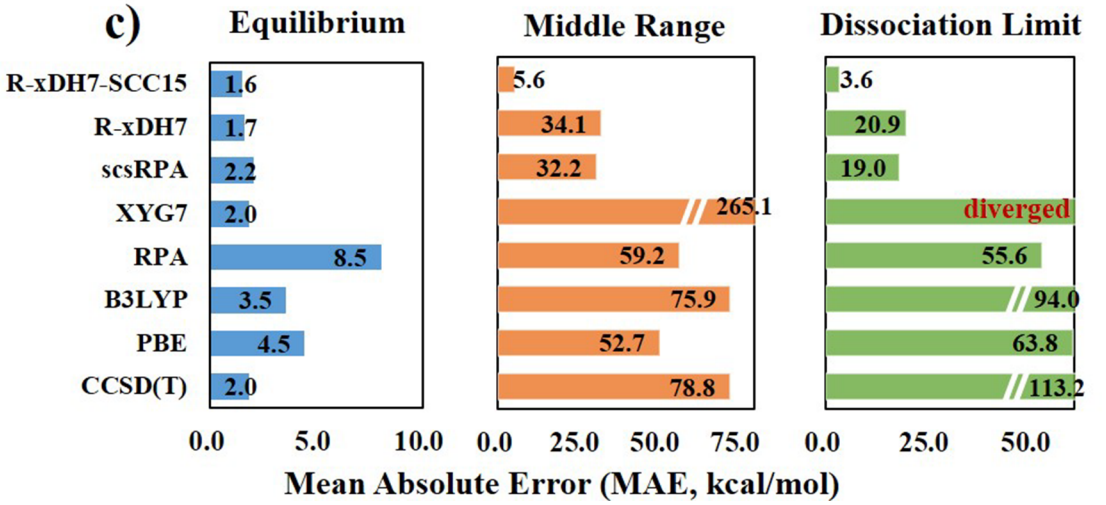


Figure 2. Assessment of various methods on the MBD23−594 data set for chemical bond dissociation. (a) Detailed depiction of calculated potential energy curves during C−C bond cleavage in ethane (C_2H_6), exemplifying method's performance. (b) 23 bond dissociation reactions encompassed within MBD23−594, which are categorized into three subsets based on the type of bond being broken. (c) Comprehensive comparison of Mean Absolute Errors (MAEs) of these methods across three critical regions: equilibrium, middle, and the dissociation limit. All calculations were performed employing the def2-QZVPP basis set. For benchmarking, reference data were prepared using the CCSD(T) method extrapolated to the complete basis set limit (CBS) for the equilibrium region and the NVEPT2 method for the middle and dissociation limit regions.

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

dissipation (ACFD) framework,[5,44,45] while $E_c^{ssRPT2+}$ presents the difference between the standard RPA and osRPT2, which includes not only the renormalization of the same-spin PT2 but also accounts for the high-order contribution beyond PT2.[5] Note that $E_c^{ssRPT2+}$ covers only the high-order direct diagrams within the same-spin channel, explicitly excluding the second-order screened exchange (SOSEX) contributions.[46] The seven parameters in R-xDH7 were optimized against the weighted total mean absolute deviation (WTMAD2) of 1505 relative energies of main-group chemistry in the GMTKN55 data set,[47] which can be found in the Supporting Information (SI). While no constraints were applied to the seven parameters during the optimization process, an interesting pattern emerges in their final values. Specifically, the first parameter, associated with the HF-like exact exchange, was found to be nearly 1, suggesting that the R-xDH7 functional has less self-interaction error (SIE) than the conventional DFAs. Furthermore, the sum of the three exchange parameters (a_1, a_2, and a_3) is approximately close to 1. Likewise, the sum of the four correlation parameters is also nearly equal to 1. We believe that this balance in the final optimized values is crucial for maintaining the accuracy and stability of the functional, as it ensures that the exchange and correlation contributions are neither overemphasized nor underrepresented.

Table 1 presents the performance of various DFAs on the GMTKN55 benchmark set. Notably, the leading performers predominantly emerge from the standard PT2-based xDH methods, including ωB97M(2)[22] and xrevDSD-PBEB86-D4[37]

Table 1. WTMAD2 Values of Different DFAs (in kcal/mol) for the Full GMTKN55 Benchmark (All), Basic Properties and Reactions of Small Systems (Sub1), Isomerizations and Reactions of Large Systems (Sub2), Barrier Heights (Sub3), Intermolecular Noncovalent Interactions (Sub4), and Intramolecular Noncovalent Interactions (Sub5)

	all	sub1	sub2	sub3	sub4	sub5
Hybrid DFAs						
PBE0-D3(BJ)[a]	6.59	4.42	8.35	9.88	6.65	6.40
B3LYP-D3(BJ)[a]	6.37	4.25	10.21	9.04	5.56	5.68
ωB97X-V[a]	3.93	3.21	6.59	4.21	3.03	3.62
ωB97M-V[b]	3.47	2.57	4.75	3.40	2.90	4.53
CF22-D[c]	3.64	2.53	4.01	3.46	4.22	4.68
Machine-learned hybrid DFA						
DM21[d]	3.97	1.99	4.64	3.63	6.56	4.16
B2PLYP-type double hybrids (bDHs)						
DSD-BLYP-D3(BJ)[b]	3.07	1.84	4.3	3.04	3.92	3.15
B2GPPLYP-D3(BJ)[b]	3.26	1.95	4.59	3.24	4.28	3.21
XYG3-type double hybrids (xDHs)						
XYG3[e]	3.75	1.72	4.69	2.23	6.04	4.88
XYGJ-OS[e]	4.04	4.04	1.85	4.42	2.94	6.01
xrevDSD-PBEB86-D4[b]	2.23	1.80	3.03	2.11	2.33	2.23
ωB97M(2)[b]	2.19	1.41	2.58	1.99	2.45	2.99
XYG7[e]	2.05	1.31	2.43	2.30	2.78	2.00
Renormalized xDHs (R-xDHs)						
scsRPA	3.54	2.66	3.13	2.79	5.58	3.68
R-xDH7	2.23	1.74	3.09	2.48	2.80	1.55
R-xDH7-SCC15	2.24	1.77	3.05	2.54	2.80	1.55

[a]Results from ref 37 using the basis set of def2-QZVPP. [b]Results from ref 22 using the basis set of def2-QZVPP. [c]Results from ref 49 using the basis set of def2-QZVPP. [d]Results from ref 50 using the basis set of def2-QZVPP. [e]Results from ref 36 using the numerical basis set of NAO-VCC-4Z.

with an impressively small WTMAD2 of 2.19 and 2.23 kcal/mol, respectively. These methods also benefit from the pairwise dispersion corrections involving several tens of empirical parameters to properly capture nonbonded interactions. Interestingly, the XYG7 method, featuring only 7 empirical parameters, achieves an outstanding WTMAD2 of 2.05 kcal/mol, although no pairwise dispersion corrections have been adopted.[36] These PT2-based xDH methods outperform the best hybrid methods such as ωB97X-V[48] (3.92 kcal/mol), CF22-D[49] (3.62 kcal/mol), and also the machine-learning local hybrid DM21[50] (3.97 kcal/mol), by approximately 1.5 to 2.0 kcal/mol.

Our R-xDH7 method attains an overall WTMAD2 of 2.23 kcal/mol for GMTKN55, placing it among the top class of computational methods in this benchmark. This performance underscores the efficacy of the renormalized PT2 in preserving the accuracy of PT2-based xDH methods across a broad range of chemical interactions, when dynamic correlations dominate. Furthermore, the R-xDH7 method demonstrates remarkable versatility in accurately modeling various types of interactions, including covalent bond interactions both in small (Sub1) and large systems (Sub2), barrier heights (Sub3), and nonbonded interactions (Sub4 and Sub5) in GMTKN55, as shown in Table 1. To extend our analysis, we followed the work of Liu and co-workers[49] and partitioned the 55 subsets of GMTKN55 into Radical7 (7 subsets) and Nonradical48 (48 subsets). This classification offered a focused lens through which to evaluate the performance of DFAs in reactions involving radical versus nonradical species. Supplementary Figure S1 in the Supporting Information suggests that R-xDH7, paralleling other state-of-the-art xDH methods, provides outstanding and balanced performance in both radical and nonradical systems. Remarkably, R-xDH7, also with 7 empirical parameters akin to XYG7, achieves this level of precision without resorting to any pairwise dispersion corrections, indicating the capabilities of renormalized PT2 correlation model for the description of nonbonded interactions.

Turning to the challenge of the static correlation, its significance grows substantially as molecules deviate from their ground state equilibrium geometries.[4,25,26] It is a notorious limitation that conventional DFAs struggle to accurately portray molecular bond dissociations, particularly when bonds are heavily stretched and the static correlation prevails due to (near)-degeneracy.[4] As demonstrated in Figure 2a, lower-rung DFAs, such as GGA of PBE[51] and the hybrid GGA of B3LYP,[38−41] exhibit significant static correlation errors. For instance, in the C−C bond dissociation of C_2H_6, these inaccuracies lead to deviations from the correct dissociation curves as large as 30 to 50 kcal/mol. PT2-based DH approximations, such as XYG7, fail badly; they tend to diverge at the dissociation limit, a direct consequence of PT2's inherent inability to handle the degenerate situations.

The proposed R-xDH7 method successfully mitigates this divergence (Figure 2a). By integrating a spin-distinctive RPA-type renormalization into the PT2 model, R-xDH7 guides the C_2H_6 dissociation process toward the correct limit. Given its excellent performance on the GMTKN55 data set, R-xDH7 manifests its proficiency in effectively capturing a good portion of dynamic correlation and a substantial amount of the static correlation. However, an intriguing observation from Figure 2a reveals the persistence of the residual static correlation error in R-xDH7, which is particularly noticeable in the middle range of

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

the dissociation curve, leading to an unphysical repulsive bump.

To further substantiate our findings, we compiled a comprehensive data set, MDB23−594, encompassing a total of 594 data points from 23 distinct bond dissociations in main-group molecules. This data set, detailed graphically in Figure 2b, includes 9 close-shell single-bond dissociations (CSD9−256), 8 multiple-bond dissociations (CMD8−204), and 6 open-shell systems (OBD6−134). Aiming to achieve an unbiased assessment, we partitioned the dissociation curves into three sections, namely equilibrium, middle and limit regions, respectively. As illustrated in Figure 2c, both XYG7 and R-xDH7 attain a good accuracy in the equilibrium region, being comparable to the expensive CCSD(T) method, which is considered the "gold standard" in quantum chemistry. However, as the static correlation effects intensifies moving from the equilibrium region to the middle and limit regions, their performances significantly degrade. The results from these 23 dissociation benchmarks confirm that R-xDH7 successfully overcomes the divergence problem observed in XYG7. However, it still exhibits mean absolute errors (MAEs) of 34.1 and 20.9 kcal/mol in the middle and limit regions, respectively. While these errors are notably reduced compared to other methods, including CCSD(T), they underscore the need for further improvement in order to accurately capture the static correlation on par with the accuracy being achieved by xDHs for the dynamic correlation.

In the realm of dispersion interactions as associated with, e.g., ωB97M(2)[22] and xrevDSD-PBEB86-D4,[37] pairwise corrections with empirical parameters have proven to be highly effective.[52−54] However, addressing static correlation errors poses a significantly greater challenge due to its complex multireference character.[55,56] Although current diagnostics, derived from both KS-DFT and wave function theory, often exhibit weak correlations with the static correlation,[57−59] machine learning has emerged as a potent tool, offering innovative avenues to decipher complex correlation patterns embedded in data.[49,50,60−63] Given that R-xDH7 has already captures a substantial portion of the static correlation, we anticipated that the residual errors, as those evident in the middle and limit regions for bond dissociations, could be rectified by devising a static correlation correction (SCC). Acknowledging that the static correlation is more pronounced yet less well understood than the dispersion interactions, we turned to machine learning. To precisely target both dynamic and static correlations for general purpose, we identified four system-independent input features at diverse theoretical levels as follows:

Electron-Pair Energy Ratios ($\{r_i\}$)

Defined as $\left\{ r_i = \frac{e_i^{\mathrm{sBGE2}}}{e_i^{\mathrm{PT2}}} \right\}$, these ratios compare sBGE2[6,64] and PT2 correlation energies for each electron pair involved in double excitations. These electron pairs are sorted in descending order according to the sum of their eigenvalues, with each pair indexed by "i" to reflect its position in this ordered sequence. This descending order ensures that the electron pair with the largest sum of eigenvalues is considered first. Due to the nature that sBGE2 was designed to mitigate the PT2 divergence,[4,6,64] $\{r_i\}$ were used to serve as an effective indicator for systems with significant static correlation. As evidenced in Supplementary Figure S2 in the Supporting Information, the distinct values of $\{r_i\}$ accurately demarcated different dissociation regimes and displayed differing convergence tendencies with respect to the number of bonds being cleaved. This characteristic behavior forms a fundamental basis of our SCC pairwise correction.

HOMO−LUMO Gap (ΔE)

Calculated as the orbital energy difference between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of the B3LYP method. As also evidenced in Supplementary Figure S2, ΔE was used to serve as a direct indicator of correlation effects across different dissociation regimes: a larger gap in the equilibrium region, a smaller gap in the middle region, and a near-degenerate gap in the limit region.

Exchange Energy Deviation (ΔE_x)

Represented as $\Delta E_x = \frac{(E_x^{\mathrm{PBE}} - E_x^{\mathrm{HF}})}{E_x^{\mathrm{HF}}} \times 100\%$. This metric was used to quantify the percentage difference between PBE and HF exchange energies. First introduced by Truhlar and colleagues[65,66] and also evidenced in Supplementary Figure S2, it is indicative of the extent of the static correlation effect within KS-DFT, with a larger difference pointing to a more important static correlation effect.

Response Functional Radii ($\{\chi_\alpha, \chi_\beta\}$)

These special radii of spin-pair distinctive response functions were used to provide insight into many-body static correlation effects. The detailed definition and implementation of $\{\chi_\alpha, \chi_\beta\}$ was provided in Supplementary Section S1. Our recent finding suggested that such effects in RPA-type correlation models are linked to the special radius of the KS noninteraction response function.[4,5] A radius exceeding unity would imply a divergence in the Dyson expansion of the RPA-like correlation, indicative of a strong static correlation effect beyond the scope of standard single-reference perturbation treatments.

Static correlation presents a unique challenge due to its elusive and ill-defined nature, as well as the scarcity of suitable data sets. To overcome these obstacles, we built here the MDB23−594 data set and devised a hybrid machine learning approach that combined symbolic and nonlinear parameter regression methods based on ($\{r_i\}$, ΔE, ΔE_x, $\{\chi_\alpha, \chi_\beta\}$). As illustrated in Figure 1, we developed a systematic approach using the Sure Independence Screening and Sparsifying Operator (SISSO)[67] method. It allowed us to thoroughly explore the descriptor space and identify the most relevant descriptors for the residual static correlation errors associated with R-xDH. Consequently, we introduced a select number of parameters in a nonlinear fashion to refine and enhance the accuracy. The training process was iterative, continuing until the desired level of accuracy was reached using descriptors that displayed well-understood behaviors.

The final correction model, coined SCC15, is composed of two distinct parts, involving 9 descriptors and 15 empirical parameters:

$$\mathrm{SCC15} = \mathrm{SCC}_{\mathrm{ep}} + \mathrm{SCC}_{\mathrm{mb}} \tag{2}$$

The detailed formulation of the SCC15 model was provided in the Supporting Information. Briefly speaking, the first segment $\mathrm{SCC}_{\mathrm{ep}} = \sum_i \mathrm{SCC}_{\mathrm{ep}}^i$ represents the pairwise correction to the static correlation error, where "i" indexes each electron pair in double excitations. The second term $\mathrm{SCC}_{\mathrm{mb}}$ addresses the residual many-body contributions beyond electron-pair excitations.

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

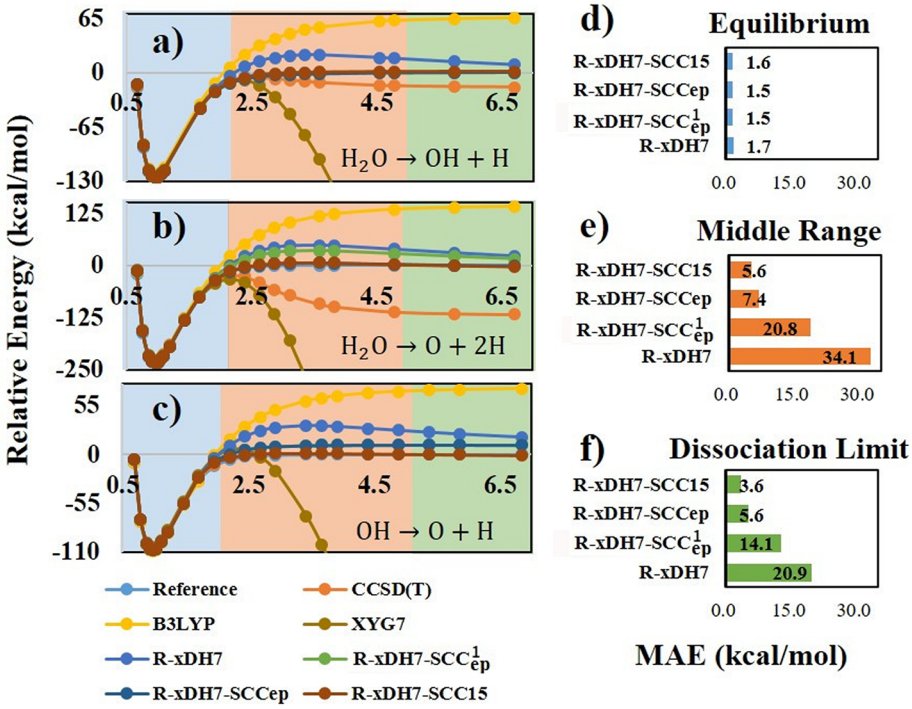


Figure 3. Effect of different components of the R-xDH7-SCC15 method on three critical bond dissociation regions for H_2O (a−c) and on the MBD23−594 data set (d−f). B3LYP and CCSD(T) results are also presented for comparison in (a−c).

We started by focusing on R-xDH7's errors in close-shell single-bond dissociations, where the relevant static correlation correction is pairwise and specific to the first electron pair: SCC^1_{ep}. We trained the SCC^1_{ep} model using the data set consisting of 131 data points from the first three reactions (S1−3) in CSD9−256, as shown in Figure 2b. Meanwhile, 1505 data points in GMTKN55 were also included, which served as constraints to preserve the good description of dynamic correlations of R-xDH7. To this stage, we arrived at the SCC^1_{ep} model with 3 descriptors $\{D^{i=1}_{1-3}\}$ and 9 empirical parameters:

$$SCC^1_{ep} = c_1 D_1^{i=1}[b_{1-5}] + c_2 D_2^{i=1}[b_{1-5}] + c_3 D_3^{i=1}[b_{1-6}] \quad (3)$$

The 3 parameters $\{c_{1-3}\}$ were concurrently determined alongside the 3 descriptors during the SISSO operation, while the remaining 6 nonlinear parameters $\{b_{1-6}\}$ underwent iterative optimization in subsequent outer loops.

In Figure 3a−c, we present the calculated curves for three critical bond dissociations in H_2O, which were not part of the training sets used in this study. The results showed that the SCC^1_{ep} model notably improved R-xDH7's performance in dissociating a single O−H bond from water ($H_2O \rightarrow OH + H$); however, it fell short when simultaneously dissociating two O−H bonds ($H_2O \rightarrow O + 2H$), as expected. For the comprehensive MBD23−594 data set, the MAEs of R-xDH7-SCC^1_{ep} in the middle region and dissociation limit were reduced to 20.8 and 14.1 kcal/mol, respectively, which marked a substantial progress compared to the uncorrected R-xDH7 method, although further improvement is needed to meet the desired accuracy comparable to that observed in the equilibrium region.

To extend the SCC^i_{ep} model for electron pairs beyond the first pair ($i > 1$), we then constructed a new training set comprising 118 data points from the first four reactions (M1−4) in CMD8−204. To maintain the success already achieved, constraints were imposed using data from GMTKN55 and also Reactions S1−3. Utilizing the hybrid regression approach led us to an interesting finding: the SCC^i_{ep} model with $i > 1$ shares two common descriptors with SCC^1_{ep}. Instead of reoptimizing the parameters established in the SCC^1_{ep} model, we introduced a novel parameter-descriptor pair $\{c_4, D_4\}$ to recalibrate the static correlation correction for higher-order electron pairs. Then the form of the SCC^i_{ep} model was updated to

$$SCC^i_{ep} = c_4 D_4 (c_1 D_1^i[b_{1-5}] + c_2 D_2^i[b_{1-5}]) \text{ for } i > 1 \quad (4)$$

where $\{c_4, D_4\}$ were optimized through a supplementary SISSO operation. Figure 3 vividly reveals that, by incorporating the static correlation correction for all electron pairs, the refined R-xDH7-SCCep approach accurately predicts the concurrent dissociation of both O−H bonds in the water molecule. It bears repeating that the three bond dissociation instances depicted in Figure 3 were neither part of the training set nor utilized as constraints for the SCC^i_{ep} model. A detailed depiction of bond dissociation curves for all 23 molecules can be found in the Supporting Information (Supplementary Figures S4−S26). A thorough analysis of these curves reveals that, R-xDH7, when augmented with the SCC_{ep} model, removes the undesired repulsive bump in the middle range of the dissociation curves.

Then we proceeded to develop the corresponding many-body static correlation correction model SCC_{mb}. We employed a training set comprising 81 data points from two reactions (O1−2) in OBD6−134 and one reaction (S9) in CSD9−256, as illustrated in Figure 2b. Consistently applying the same strategy, data from GMTKN55 and Reactions S1−3, along with M1−4 were included as constraints to minimize potential model bias or unintended influence on the outcomes. In light of the complex nature of many-body static correlation, the final SCC_{mb} model incorporates another 5 descriptors and their

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

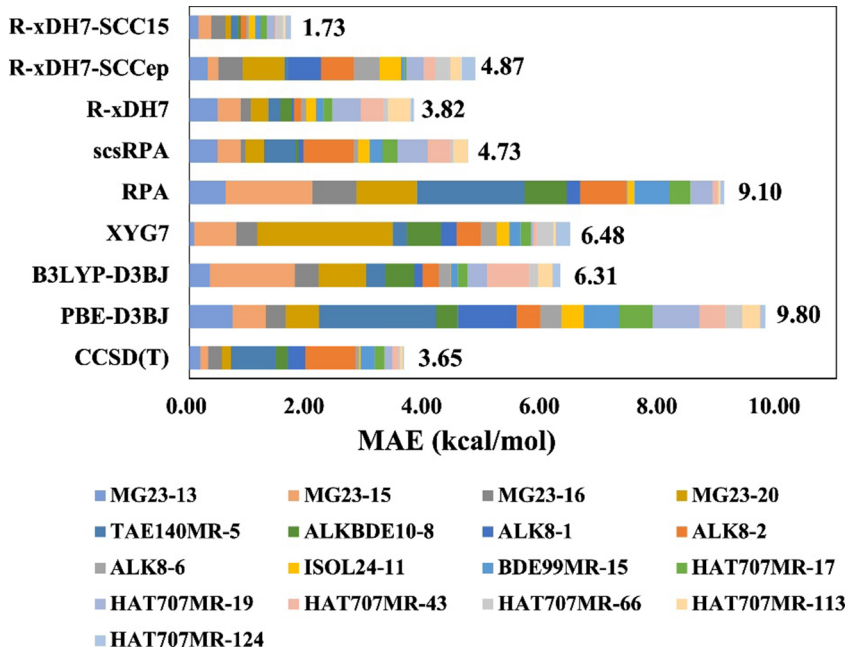


Figure 4. Performance comparison of various methods on a specialized validation set of 17 reactions where the SCC15 model exhibits a substantial improvement (>3 kcal/mol). Each reaction is filtered out from the GMTKN55, MGDC84, or MG23 databases and identified by its corresponding subset name and original index. Details of these 17 reactions are presented in Supplementary Table S2.

corresponding parameters, which were systematically identified and fine-tuned using the standard SISSO symbolic regression approach:

$$\mathrm{SCC}_{\mathrm{mb}} = \sum_{i=5}^{9} c_i D_i \tag{5}$$

To this end, we arrived at the SCC15 model step by step, which is specifically adapted to R-xDH7.

As evident in Figure 3, the resultant R-xDH7-SCC15 method achieves a remarkably accurate description for all three types of bond dissociation processes involved in H_2O. Notably, within the MBD23−594 data set, 13 out of total 23 dissociation curves (over 300 distinct data points) were deliberately excluded from the training of the SCC15 model (see Figure 1). Deeper exploration within this expended framework further highlights the advancements brought forth by the R-xDH7-SCC15 method for the static correlation are indeed significant. The overall MAEs in the middle range and dissociation limit are now 5.6 and 3.6 kcal/mol, respectively, approaching to its performance in the equilibrium region (MAE = 1.6 kcal/mol) where the dynamic correction is dominant. Therefore, R-xDH7-SCC15 strikes an unprecedented balance between accurately capturing dynamic and static correlation effects. This conclusion is further reinforced by its high accuracy on the GMTKN55 data set. As detailed in Table 1, the overall MAE of R-xDH7-SCC15 for GMTKN55 is 2.24 kcal/mol. Upon examination across five subsets of GMTKN55, it is clear that the SCC15 model introduces, on average, an influence of less than 0.1 kcal/mol. Thus, the significant advancement of R-xDH7-SCC15 in describing strongly correlated systems is achieved without compromising its accuracy in describing dynamic correlations.

The MGCDB84 data set,[68] developed by the Head-Gordon group, represents a chemically diverse and comprehensive database designed for main-group chemistry. It encompasses 84 subsets containing a total of 4986 individual data points. Similar to GMTKN55, MGCDB84 primarily targets dynamic correlations with exception to 3 subsets denoted by the "MR" suffix.[68] Given the limited overlap of only 804 data points across 24 subsets shared by GMTKN55 and MGCDB84, we utilized the MGCDB84 data set as a supplementary means of validation. Supplementary Figure S3 in the Supporting Information suggests XYG7 and R-xDH7 as the top performers with MAEs of 0.65 and 0.67 kcal/mol, respectively. Notably, the SCC15 model shows a negligible effect on MGCDB84, with R-xDH7-SCC15 maintaining the good performance with an overall MAE of 0.67 kcal/mol. This consistency highlights the robustness and transferability of the predictive power of R-xDH7-SCC15.

While MGCDB84 is predominantly governed by the dynamic correlation outside the training set, we further evaluate the performance of R-xDH7-SCC15 on the MG23 set, which comprises 23 chemical reaction energies having important static correlation effects. These testing cases involve both main-group metal and nonmetal elements as detailed in Supplementary Table S1 in the Supporting Information, meticulously collected from prior literature sources[69,70] and identified using the %[(T)] diagnostic for their static correlation characters. While the CCSD(T) method yields an MAE of 4.03 kcal/mol, standard PT2-based methods, hampered by their inherent limitations in handling static correlation effects, demonstrate less accurate results. For instance, XYG7 records an MAE of 6.23 kcal/mol on the MG23 set. In sharp contrast, R-xDH7 effectively captures a significant portion of the static correlation with its innovative renormalized design, reducing the MAE to 2.96 kcal/mol. More significantly, R-xDH7-SCC15 achieves top-tier accuracy with an MAE of 2.43 kcal/mol on the MG23 set, underlining its superior potential to tackle problems with important static correlation effects.

We would like to emphasize that the SCC15 model was devised to eliminate the residual errors adapted for R-xDH7 in

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

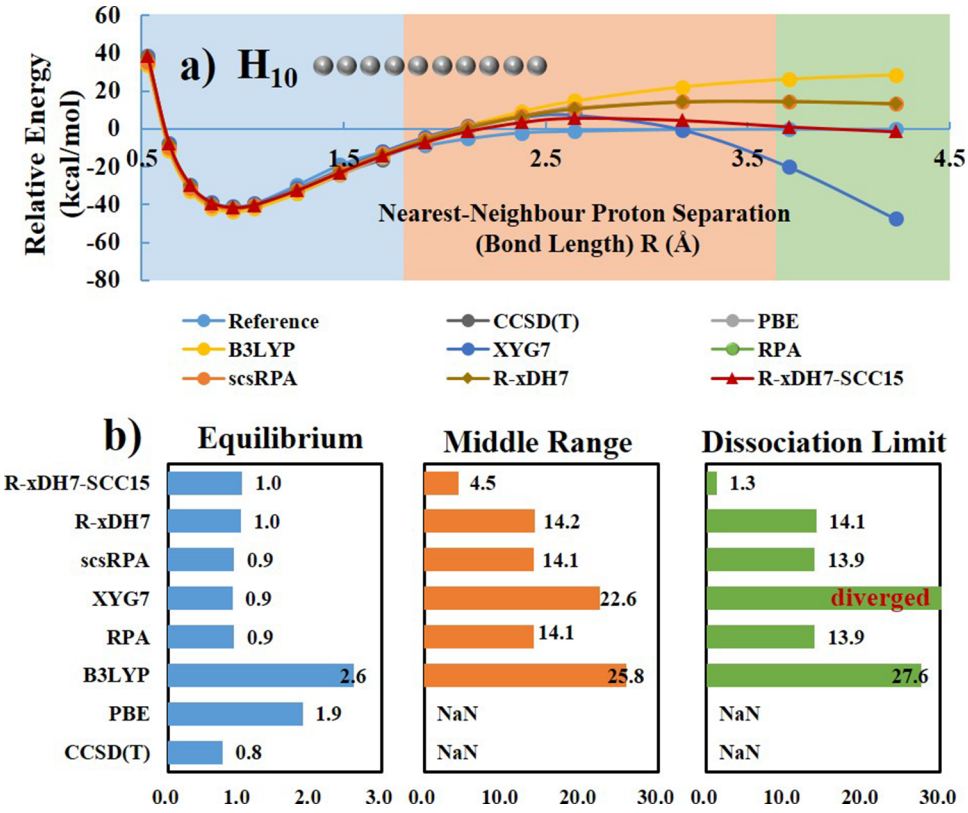


Figure 5. Relative potential energy curve of H10. Energies per atom were evaluated by $\frac{1}{10}E[H_{10}] - E[H]$ (a) and Mean Absolute Errors (MAEs) of various methods across three critical regions: equilibrium, middle, and the dissociation limit (b). All calculations were performed employing the def2-QZVPP basis set. For benchmarking, accurate reference data were taken from ref 71. 'NaN' denotes failure of SCF convergence for PBE or non-convergence of coupled-cluster iteration for CCSD(T).

the treatment of the static correlation. This point was substantiated by our findings: 4 out of 23 reactions in the MG23 test set exhibit SCC_{ep}, SCC_{mb}, or SCC15 values exceeding 3.0 kcal/mol. Extending this threshold across more than 6400 data points from GMTKN55, MGCDB84, and MG23 altogether, we identified and filtered out 17 specific chemical reactions which were detailed in Supplementary Table S2 in the Supporting Information. Figure 4 and Supplementary Table S2 show that R-xDH7 yields an MAE of 3.82 kcal/mol, which closely parallels the performance of the CCSD(T) method and significantly surpassed other DFAs. Significantly, none of these 17 reactions are part of training sets for the SCC15 model, yet R-xDH7-SCC15 achieves an outstanding performance with an MAE of only 1.73 kcal/mol, demonstrating again its remarkable robustness and transferability across a diverse spectrum of chemical scenarios. When applying only the electron-pair correction (i.e., R-xDH7-SCCep), the accuracy is inferior (MAE = 4.87 kcal/mol), highlighting the crucial interplay between electron-pair and many-body corrections in achieving optimal results. The success of the SCC15 model is rooted in its developing strategy that effectively tackled both types of static correlation errors, thereby providing strong evidence for the validity and effectiveness of our systematic approaches.

The linear hydrogen chain model holds a profound position within the realm of computational materials science,[71] as it provides a pivotal means to mimic strong electron correlation effects, prevalent in many complex materials. Unlike the Hubbard model that simplifies interactions to one-dimensional lattice sites, the linear H chain model offers a more authentic depiction of molecular bonding inherent to condensed matter systems. Figure 5a presents the energy profile that dissociates all bonds simultaneously in a finite hydrogen chain consisting of ten atoms (H10), while Supplementary Table S5 provides the calculated energies per atom at the thermodynamic limit (TDL). This specific model system presents a stringent test, as it involves multiple bond dissociations simultaneously, being a far more challenging task compared to those usually encountered in typical molecules as in MBD23−594. R-xDH7-SCC15 demonstrates its robustness by providing a balanced and accurate description for the whole dissociation curve. As showcased in Figure 5b. the overall MAEs of R-xDH7-SCC15 in the middle range and dissociation limit are 4.5 and 1.3 kcal/mol, respectively, which are notably close to its performance in the equilibrium region (MAE = 1.0 kcal/mol), where the dynamic correction is dominant.

As a practical case study, we examined here the dissociative chemisorption of methanol (CH_3OH) on the Cu(111) surface, where three primary reaction channels are involved with the cleavage of O−H, C−H, and C−O bonds, respectively. The control of the product branching ratios via vibrational mode excitation has been proposed.[72,73] A proper description of these product branching ratios necessitates accurate molecular dynamics simulations, where constructing precise potential energy surfaces is essential. Until now, however, these surfaces have primarily been constructed at the GGA level.[72]

First, we calculated the dissociation curves of methanol in the gas phase (Figure 6a,c, and e). Among the methods tested, R-xDH7-SCC15 performs significantly better with an MAE of only 1.91 kcal/mol against reference curves, outperforming

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

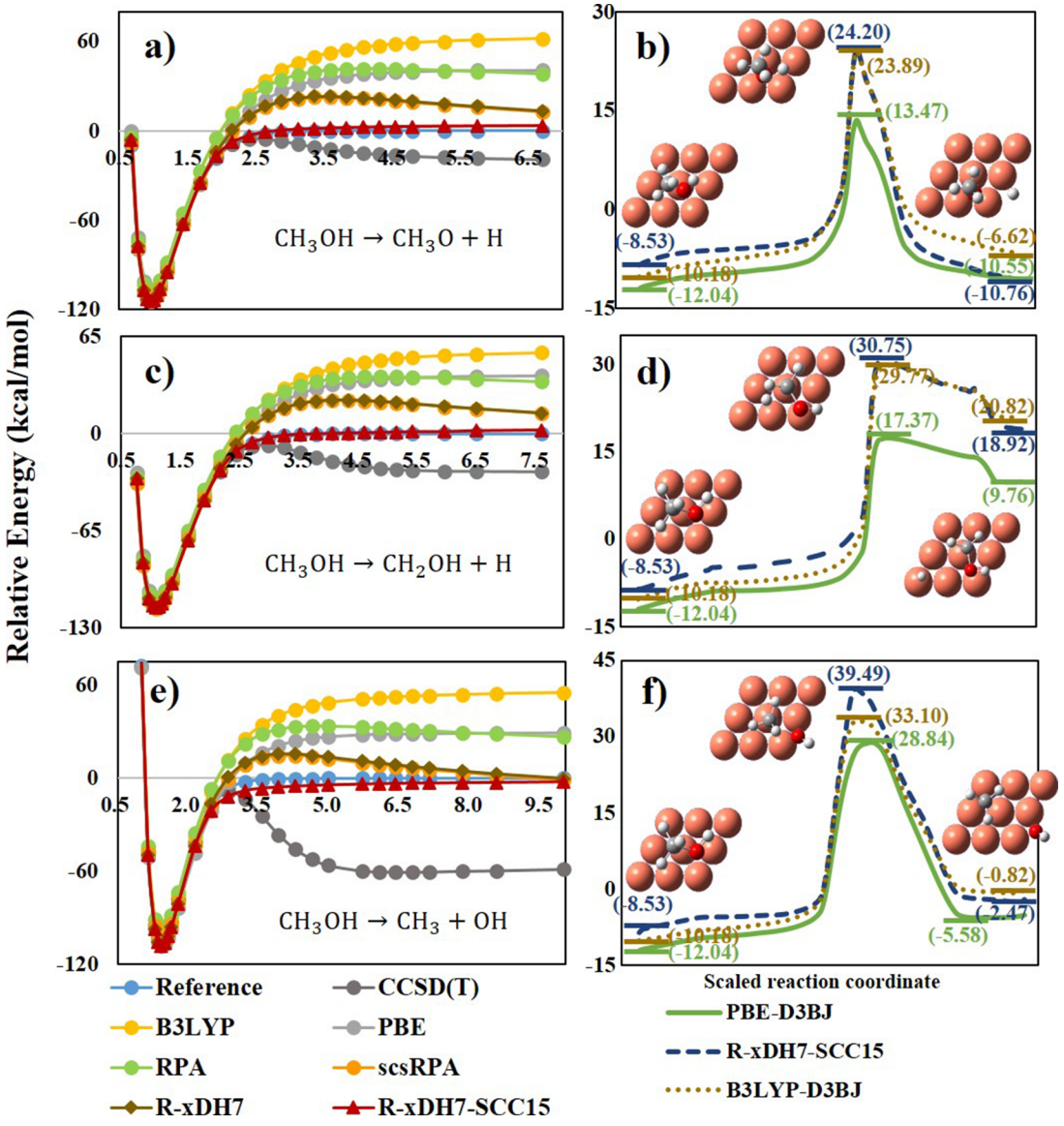


Figure 6. Comparative analysis of the performance of various methods in calculating the dissociation processes of methanol (CH_3OH) in the gas phase (a,c, and e) and on the Cu(111) surface (b,d, and f).

traditional single-reference based methods such as conventional DFAs and even CCSD(T). We then extended our analysis to the bond dissociation curves of methanol on the Cu(111) surface by employing the XO-PBC strategy recently developed (See the Supporting Information for the detail of XO-PBC).[74−76] Using R-xDH7-SCC15 results as the benchmark, we identified that the widely used GGA methods, in particular PBE-D3BJ,[52,54] have a strong tendency to overestimate the adsorption energies, while significantly underestimating reaction barriers. Figure 6b,d, and f show an improvement when utilizing the hybrid B3LYP-D3BJ method. For O−H and C−H bond dissociation pathways, the errors in B3LYP-D3BJ reaction barriers are less than 1 kcal/mol. However, the error exceeds 6 kcal/mol for C−O bond dissociation. Such discrepancies in predicted reaction barriers of different pathways are critical, as they have a direct impact on the determination of the product branching ratios, highlighting the need for more advanced computational approaches in an unbiased modeling of the complex surface reactions.

In summary, we have developed the R-xDH7 functional that stands at the apex of Jacob's Ladder of KS-DFT. This renormalized xDH method, containing seven empirical parameters, excels in capturing a significant portion of the static correlation, while preserving the generally good performance of xDH methods for a wide range of systems with the dynamic correlation. A pivotal achievement lies in the successful development of the SCC15 model through symbolic and parameter regressions to eliminate the residual static correlation errors associated with R-xDH7. The resultant R-xDH7-SCC15 method achieves an extraordinary balance in representing dynamic and static correlation effects over conventional DFAs, as well as the "gold standard" CCSD(T) in wave function theory. This superiority has been confirmed by extensive benchmarking over a diverse array of 6400 main-group chemical reactions sourced from GMTKN55, MGDCB84, and MG23 data sets, as well as 594 data points associated with 23 molecular dissociation curves. Applications of R-xDH7-SCC15 to the dissociative chemisorption of methanol on copper surfaces have showcased the inadequacies

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

of the widely used GGAs, as well as hybrid DFAs, in accurately describing complex reaction energetics and kinetics. We anticipate that the R-xDH7-SCC15 method, integrating superior accuracy, high computational efficiency and black-box-like accessibility all-in-one, will significantly boost the predictive power of computational chemistry, particularly in elucidating intricate reaction mechanisms when both static and dynamic correlation effects are crucial.

METHODS

The electronic calculations for R-xDH7-SCC15 were conducted using our in-house Rust-based Electronic Structure Toolkits (REST).[28] Unless specified otherwise, the def2-QZVPP basis set was employed. The resolution-of-identity approach was introduced in conjunction with the def2-SV(P)-JKFIT auxiliary basis set to enhance the computational efficiency. Only valence electrons from the outer shell were considered in the post self-consistent field (post-SCF) calculations. For the molecular dissociation curves, the calculations were performed without breaking spin symmetry. A detailed account of the computational parameters and methodology is provided in the SI. It is worth noting that the computational scaling of R-xDH7-SCC15 is on part with standard PT2-based doubly hybrid approximations. This is attributed to the advanced correlation models introduced, which exhibit the same or even lower computational scaling than the PT2 term. As illustrated in Table S8 in Supporting Information, the consumption cost of evaluating the PT2 correlations in XYG7 is comparable to a single B3LYP SCF iteration. This efficiency is particularly notable for large systems, as demonstrated by our calculations on systems comprising up to 80 atoms and over 3000 basis functions. Furthermore, the many-body renormalization of the PT2 evaluation and the SCC15 correction are cost-effective, similar to the standard PT2 term. This highlights the practicality of the R-xDH7-SCC15 method within the REST package for complex systems. While the REST package currently supports numerical evaluation of energy gradients for R-xDH7-SCC15, the development of analytic energy gradients is underway and will be a significant enhancement in the forthcoming versions of the package.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.4c00488.

R-xDH7 functional, SCC15 model, and computational details (PDF)

Geometry files for MBD23−594 and MG23 data sets and the surface reaction system of CH_3OH on Cu(111) (zip)

AUTHOR INFORMATION

Corresponding Authors

Igor Ying Zhang − *Shanghai Key Laboratory of Molecular Catalysis and Innovation Materials, Collaborative Innovation Centre of Chemistry for Energy Materials, MOE Laboratory for Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200438, China; Hefei National Laboratory, Hefei 230088, China; Shanghai Key Laboratory of Bioactive Small Molecules, Shanghai 200032, China;* orcid.org/0000-0002-8703-1912; Email: igor_zhangying@fudan.edu.cn

Xin Xu − *Shanghai Key Laboratory of Molecular Catalysis and Innovation Materials, Collaborative Innovation Centre of Chemistry for Energy Materials, MOE Laboratory for Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200438, China; Hefei National Laboratory, Hefei 230088, China;* orcid.org/0000-0002-5247-2937; Email: xxchem@fudan.edu.cn

Authors

Yizhen Wang − *Shanghai Key Laboratory of Molecular Catalysis and Innovation Materials, Collaborative Innovation Centre of Chemistry for Energy Materials, MOE Laboratory for Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200438, China*

Zihan Lin − *Shanghai Key Laboratory of Molecular Catalysis and Innovation Materials, Collaborative Innovation Centre of Chemistry for Energy Materials, MOE Laboratory for Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200438, China*

Runhai Ouyang − *Materials Genome Institute, Shanghai University, Shanghai 200444, China;* orcid.org/0000-0001-9020-0484

Bin Jiang − *Key Laboratory of Precision and Intelligent Chemistry, Department of Chemical Physics, University of Science and Technology of China, Hefei, Anhui 230026, China; Hefei National Laboratory, Hefei 230088, China;* orcid.org/0000-0003-2696-5436

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.4c00488

Author Contributions

CRediT: **Yizhen Wang** data curation, formal analysis, investigation, writing-original draft; **Zihan Lin** data curation, formal analysis, investigation; **Runhai Ouyang** data curation, formal analysis, funding acquisition, investigation; **Bin Jiang** data curation, formal analysis, funding acquisition, investigation; **Igor Ying Zhang** conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing-original draft, writing-review & editing; **Xin Xu** conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, writing-review & editing.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (22125301, 22233002, 22393911, 22393912, 22321003, 22325304), the Innovation Program for Quantum Science and Technology (2021ZD0303305), Innovative research team of high-level local universities in Shanghai and a key laboratory program of the Education Commission of Shanghai Municipality (ZDSYS14005).

REFERENCES

(1) Szabó, A.; Ostlund, N. S. *Modern quantum chemistry: Introduction to advanced electronic structure theory*; Dover Publications, Inc.: Mineola, New York, 1982.

(2) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136* (3B), B864−B871.

(3) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140* (4A), A1133−A1138.

(4) Zhang, I. Y.; Xu, X. On the top rung of Jacob's ladder of density functional theory: Toward resolving the dilemma of SIE and NCE. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021**, *11* (1), No. e1490.

(5) Zhang, I. Y.; Xu, X. Simultaneous attenuation of both self-interaction error and nondynamic correlation error in density functional theory: A spin-pair distinctive adiabatic-connection approximation. *J. Phys. Chem. Lett.* **2019**, *10* (10), 2617−2623.
(6) Zhang, I. Y.; Rinke, P.; Perdew, J. P.; Scheffler, M. Towards efficient orbital-dependent density functionals for weak and strong correlation. *Phys. Rev. Lett.* **2016**, *117* (13), No. 133002.
(7) Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conf. Proc.* **2001**, *577* (1), 1−20.
(8) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124* (3), No. 034108.
(9) Tarnopolsky, A.; Karton, A.; Sertchook, R.; Vuzman, D.; Martin, J. M. L. Double-hybrid functionals for thermochemical kinetics. *J. Phys. Chem. A* **2008**, *112* (1), 3−8.
(10) Karton, A.; Tarnopolsky, A.; Lamere, 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* (50), 12868−12886.
(11) Chai, J. D.; Head-Gordon, M. Long-range corrected double-hybrid density functionals. *J. Chem. Phys.* **2009**, *131* (17), No. 174105.
(12) Zhang, Y.; Xu, X.; Goddard, W. A., III Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (13), 4963−4968.
(13) Kozuch, S.; Gruzman, D.; Martin, J. M. L. DSD-BLYP: A general purpose double hybrid density functional including spin component scaling and dispersion correction. *J. Phys. Chem. C* **2010**, *114* (48), 20801−20808.
(14) 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* (2), 291−309.
(15) Kozuch, S.; Martin, J. M. L. DSD-PBEP86: in search of the best double-hybrid DFT with spin-component scaled MP2 and dispersion corrections. *Phys. Chem. Chem. Phys.* **2011**, *13* (45), 20104−20107.
(16) Sharkas, K.; Toulouse, J.; Savin, A. Double-hybrid density-functional theory made rigorous. *J. Chem. Phys.* **2011**, *134* (6), No. 064113.
(17) Zhang, I. Y.; Su, N. Q.; Brémond, É. A. G.; Adamo, C.; Xu, X. Doubly hybrid density functional xDH-PBE0 from a parameter-free global hybrid model PBE0. *J. Chem. Phys.* **2012**, *136* (17), 174103.
(18) Brémond, E.; Adamo, C. Seeking for parameter-free double-hybrid functionals: The PBE0-DH model. *J. Chem. Phys.* **2011**, *135* (2), No. 024106.
(19) Brémond, É.; Sancho-García, J. C.; Pérez-Jiménez, Á. J.; Adamo, C. Communication: Double-hybrid functionals from adiabatic-connection: The QIDH model. *J. Chem. Phys.* **2014**, *141* (3), No. 031101.
(20) Su, N. Q.; Xu, X. Construction of a parameter-free doubly hybrid density functional from adiabatic connection. *J. Chem. Phys.* **2014**, *140* (18), 18a512.
(21) Bremond, E.; Savarese, M.; Perez-Jimenez, A. J.; Sancho-Garcia, J. C.; Adamo, C. Range-separated double-hybrid functional from nonempirical constraints. *J. Chem. Theory Comput.* **2018**, *14* (8), 4052−4062.
(22) Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ωB97M(2) double hybrid density functional. *J. Chem. Phys.* **2018**, *148* (24), 241736.
(23) Zhang, I. Y.; Xu, X. Reaching a uniform accuracy for complex molecular systems: Long-range-corrected XYG3 doubly hybrid density functional. *J. Phys. Chem. Lett.* **2013**, *4* (10), 1669−1675.
(24) Zhang, I. Y.; Xu, X.; Jung, Y.; Goddard, W. A. A fast doubly hybrid density functional method close to chemical accuracy using a local opposite spin ansatz. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (50), 19896−19900.
(25) Li, C.; Yang, W. On the piecewise convex or concave nature of ground state energy as a function of fractional number of electrons for approximate density functionals. *J. Chem. Phys.* **2017**, *146* (7), No. 074107.
(26) Mori-Sanchez, P.; Cohen, A. J.; Yang, W. Localization and delocalization errors in density functional theory and implications for band-gap prediction. *Phys. Rev. Lett.* **2008**, *100* (14), No. 146401.
(27) Helgaker, T.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Jørgensen, P.; Olsen, J. Multiconfigurational Self-Consistent Field Theory. In *Molecular electronic-structure theory*, 2000, 598647.
(28) Zhang, I. Y.; Xu, X., git@github.com:igor-1982/rest.git.
(29) Zhang, I. Y.; Xu, X. A New Generation of Doubly Hybrid Density Functionals (DHDFs). In *A new-generation density functional: Towards chemical accuracy for chemistry of main group elements*, Zhang, I. Y.; Xu, X., Eds.; Springer: Berlin Heidelberg: Berlin, Heidelberg, 2014; 25−45.
(30) Wang, Y.; Li, Y.; Chen, J.; Zhang, I. Y.; Xu, X. Doubly Hybrid Functionals Close to Chemical Accuracy for Both Finite and Extended Systems: Implementation and Test of XYG3 and XYGJ-OS. *JACS Au* **2021**, *1* (5), 543−549.
(31) 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.
(32) Zhang, I. Y.; Xu, X. XYG3 and XYGJ-OS performances for noncovalent binding energies relevant to biomolecular structures. *Phys. Chem. Chem. Phys.* **2012**, *14* (36), 12554−12570.
(33) Zhang, I. Y.; Wu, J.; Luo, Y.; Xu, X. Accurate bond dissociation enthalpies by using doubly hybrid XYG3 functional. *J. Comput. Chem.* **2011**, *32* (9), 1824−1838.
(34) Su, N. Q.; Xu, X. Insights into direct methods for predictions of ionization potential and electron affinity in density functional theory. *J. Phys. Chem. Lett.* **2019**, *10* (11), 2692−2699.
(35) Zhang, I. Y.; Xu, X. Gas-phase thermodynamics as a validation of computational catalysis on surfaces: A case study of Fischer−Tropsch synthesis. *ChemPhysChem* **2012**, *13* (6), 1486−1494.
(36) Zhang, I. Y.; Xu, X. Exploring the limits of the XYG3-type doubly hybrid approximations for the main-group chemistry: The xDH@B3LYP model. *J. Phys. Chem. Lett.* **2021**, *12* (10), 2638−2644.
(37) 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* (24), 5129−5143.
(38) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38* (6), 3098−3100.
(39) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37* (2), 785−789.
(40) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648−5652.
(41) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98* (45), 11623−11627.
(42) Dirac, P. A. M. Note on exchange phenomena in the Thomas atom. *Math. Proc.* **1930**, *26* (3), 376−385.
(43) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Can. J. Phys.* **1980**, *58* (8), 1200−1211.
(44) Langreth, D. C.; Perdew, J. P. Exchange-correlation energy of a metallic surface: Wave-vector analysis. *Phys. Rev. B* **1977**, *15* (6), 2884−2901.
(45) Langreth, D. C.; Perdew, J. P. The exchange-correlation energy of a metallic surface. *Solid State Commun.* **1975**, *17* (11), 1425−1429.

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216

(46) Grüneis, A.; Marsman, M.; Harl, J.; Schimka, L.; Kresse, G. Making the random phase approximation to electronic correlation accurate. *J. Chem. Phys.* **2009**, *131* (15), 154115.
(47) 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* (48), 32184−32215.
(48) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16* (21), 9904−9924.
(49) Liu, Y.; Zhang, C.; Liu, Z.; Truhlar, D. G.; Wang, Y.; He, X. Supervised learning of a chemistry functional with damped dispersion. *Nat. Comput. Sci.* **2023**, *3* (1), 48−58.
(50) Kirkpatrick, J.; McMorrow, B.; Turban, D. H. P.; Gaunt, A. L.; Spencer, J. S.; Matthews, A.; Obika, A.; Thiry, L.; Fortunato, M.; Pfau, D.; Castellanos, L. R.; Petersen, S.; Nelson, A. W. R.; Kohli, P.; Mori-Sánchez, P.; Hassabis, D.; Cohen, A. J. Pushing the frontiers of density functionals by solving the fractional electron problem. *Science* **2021**, *374* (6573), 1385.
(51) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865−3868.
(52) 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* (15), 154104.
(53) Liu, Y.; Goddard, W. A., III A universal damping function for empirical dispersion correction on density functional theory. *Mater. Trans.* **2009**, *50* (7), 1664−1670.
(54) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32* (7), 1456−1465.
(55) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. Insights into current limitations of density functional theory. *Science* **2008**, *321* (5890), 792−794.
(56) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. Challenges for density functional theory. *Chem. Rev.* **2012**, *112* (1), 289−320.
(57) Lee, T. J.; Taylor, P. R. A diagnostic for determining the quality of single-reference electron correlation methods. *Int. J. Quantum Chem.* **1989**, *36* (S23), 199−207.
(58) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions. *J. Chem. Phys.* **2006**, *125* (14), 144108.
(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* (4), 165−178.
(60) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big data meets quantum chemistry approximations: The Δ-machine learning approach. *J. Chem. Theory Comput.* **2015**, *11* (5), 2087−2096.
(61) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine learning force fields. *Chem. Rev.* **2021**, *121* (16), 10142−10186.
(62) Snyder, J. C.; Rupp, M.; Hansen, K.; Müller, K.-R.; Burke, K. Finding density functionals with machine learning. *Phys. Rev. Lett.* **2012**, *108* (25), No. 253002.
(63) Li, L.; Snyder, J. C.; Pelaschier, I. M.; Huang, J.; Niranjan, U.-N.; Duncan, P.; Rupp, M.; Müller, K.-R.; Burke, K. Understanding machine-learned density functionals. *Int. J. Quantum Chem.* **2016**, *116* (11), 819−833.
(64) Zhang, I. Y.; Rinke, P.; Scheffler, M. Wave-function inspired density functional applied to the H2/H2+ challenge. *New J. Phys.* **2016**, *18* (7), No. 073026.
(65) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. Density functionals for inorganometallic and organometallic chemistry. *J. Phys. Chem. A* **2005**, *109* (49), 11127−11143.
(66) 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* (7), 2968−2983.
(67) Ouyang, R.; Curtarolo, S.; Ahmetcik, E.; Scheffler, M.; Ghiringhelli, L. M. SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys. Rev. Mater.* **2018**, *2* (8), No. 083802.
(68) 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* (19), 2315−2372.
(69) Peverati, R.; Truhlar, D. G. Quest for a universal density functional: The accuracy of density functionals across a broad spectrum of databases in chemistry and physics. *Philos. Trans. R. Soc. Lond. Ser. A: Math. Phys. Eng. Sci.* **2011**, *2014* (372), No. 20120476.
(70) Zhao, Y.; Tishchenko, O.; Gour, J.; Li, W.; Lutz, J.; Piecuch, P.; Truhlar, D. Thermochemical kinetics for multireference systems: Addition reactions of ozone. *J. Phys. Chem. A* **2009**, *113*, 5786−5799.
(71) Motta, M.; Ceperley, D. M.; Chan, G. K. L.; Gomez, J. A.; Gull, E.; Guo, S.; Jiménez-Hoyos, C. A.; Lan, T. N.; Li, J.; Ma, F.; Millis, A. J.; Prokof'ev, N. V.; Ray, U.; Scuseria, G. E.; Sorella, S.; Stoudenmire, E. M.; Sun, Q.; Tupitsyn, I. S.; White, S. R.; Zgid, D.; Zhang, S. Towards the solution of the many-electron problem in real materials: Equation of state of the hydrogen chain with state-of-the-art many-body methods. *Phys. Rev. X* **2017**, *7* (3), No. 031059.
(72) Chen, J.; Zhou, X.; Zhang, Y.; Jiang, B. Vibrational control of selective bond cleavage in dissociative chemisorption of methanol on Cu(111). *Nat. Commun.* **2018**, *9* (1), 4039.
(73) Hundt, P. M.; Jiang, B.; van Reijzen, M. E.; Guo, H.; Beck, R. D. Vibrationally promoted dissociation of water on Ni(111). *Science* **2014**, *344* (6183), 504−507.
(74) Chen, Z.; Liu, Z.; Xu, X. Accurate descriptions of molecule-surface interactions in electrocatalytic CO2 reduction on the copper surfaces. *Nat. Commun.* **2023**, *14* (1), 936.
(75) Chen, B.; Xu, X. XO-PBC: An accurate and efficient method for molecular crystals. *J. Chem. Theory Comput.* **2020**, *16* (7), 4271−4285.
(76) Guo, W.; Wu, A.; Zhang, I. Y.; Xu, X. XO: An extended ONIOM method for accurate and efficient modeling of large systems. *J. Comput. Chem.* **2012**, *33* (27), 2142−2160.

NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP on August 12, 2024, with errors throughout the paper. The corrected version was reposted August 14, 2024.

https://doi.org/10.1021/jacsau.4c00488
JACS Au 2024, 4, 3205−3216