THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

Density Sensitivity of Empirical Functionals

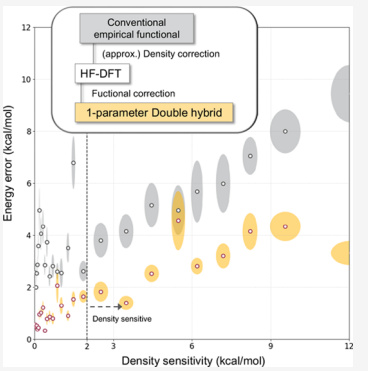
Suhwan Song, Stefan Vuckovic, Eunji Sim,* and Kieron Burke

Cite This: *J. Phys. Chem. Lett.* 2021, 12, 800−807

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Empirical fitting of parameters in approximate density functionals is common. Such fits conflate errors in the self-consistent density with errors in the energy functional, but density-corrected DFT (DC-DFT) separates these two. We illustrate with catastrophic failures of a toy functional applied to H_2^+ at varying bond lengths, where the standard fitting procedure misses the exact functional; Grimme's D3 fit to noncovalent interactions, which can be contaminated by large density errors such as in the WATER27 and B30 data sets; and double-hybrids trained on self-consistent densities, which can perform poorly on systems with density-driven errors. In these cases, more accurate results are found at no additional cost by using Hartree−Fock (HF) densities instead of self-consistent densities. For binding energies of small water clusters, errors are greatly reduced. Range-separated hybrids with 100% HF at large distances suffer much less from this effect.



Density functional theory has become a standard tool for computing electronic structure in chemistry. Early functionals, such as the local density approximation, were derived from physical intuition.[1,2] For the last quarter century, fitting of empirical parameters in approximate exchange−correlation functionals has been popular. The early successes of Becke88 exchange,[3] Lee−Yang−Parr correlation,[4] and the global hybrid ideas of Becke,[5] ultimately led to the hugely successful B3LYP.[6] Since then, the number of functionals and the number of parameters have proliferated,[7,8] and often dozens of parameters are fitted to dozens of databases, with thousands of benchmark data. Some of the most recent empirical functionals[9] achieve remarkable accuracy for molecular systems.

There are many pitfalls to such fitting, but we focus on just one. This danger is unambiguous, has nothing to do with choices of parameters or data sets, and is entirely avoidable. Almost all such fittings consist of running one or more self-consistent DFT calculations, evaluating an energy difference, and comparing it with a (presumably accurate) energy from the database. (In the case of bond lengths, the difference is an infinitesimal, determining where an energy derivative vanishes.) The accuracy of self-consistent densities was recently highlighted,[10] as was how errors in the density can be related to errors in the energy.[11−23]

The density-driven error is the (typically small) contribution to the energy error due to the error in the self-consistent density. So long as density-driven errors were small compared to the functional errors (as was the case in the halcyon days of B3LYP), they were irrelevant. But in the modern era of vast databases that include weak interactions, stretched bonds, *etc.*, these errors are sometimes as big as (or larger than) the functional errors.[20,23] However, the common practice of direct comparison with accurate energies conflates both errors and cannot distinguish the two. Recent advances in machine learning of density functionals target the density as well as the energy and likely succeed because both errors are simultaneously minimized.[24]

The cure for this difficulty is simple: where relevant, empirical schemes should be trained on purely functional errors; that is, the functional error of a parametrized approximation to the energy should be optimized against accurate energy databases, rather than the self-consistent error. For calculations that are not density-sensitive, the differences are so small as to make this irrelevant. But for those that are, this procedure isolates the self-consistency error and so avoids the corruption of the optimization process, allowing density-sensitive cases to be included even in training.

The current Letter highlights the consequences of ignoring this distinction when optimizing parameters in empirical functionals. We first create a totally artificial problem to emphasize the difficulties, especially when one uses a semilocal approximation for the self-consistent density but a more accurate form for the energy. In this case, we show how the exact functional is missed by the standard procedure. Next, we take the D3 correction of Grimme and co-workers[25] and show how, if complexes with large density-driven errors are naively included, the results become noticeably worse. On the other

Received: November 30, 2020
Accepted: January 3, 2021

https://dx.doi.org/10.1021/acs.jpclett.0c03545
J. Phys. Chem. Lett. 2021, 12, 800−807

hand, the use of DC-DFT allows previous good results to be retained and the more difficult complexes to be included. We also apply our method to double-hybrids (DHs), producing a combination that competes with similar functionals but still works when the density sensitivity is large. Finally, we find that empirical range-separated hybrid functionals suffer less from density-driven errors than their conventional global counterparts.

Background. The theory of density-corrected DFT (DC-DFT) has been developed over the past decade.[15,26] Whenever a self-consistent (SC) DFT calculation is run, there are two distinct sources of error. The total error of such calculations is $\Delta E = \tilde{E}[\tilde{n}] - E[n]$, where E and n are the exact energy functional and density, and $\tilde{E}$ and $\tilde{n}$ are their approximate counterparts. We decompose ΔE as[17,22,27]

$$\Delta E = \underbrace{\tilde{E}[\tilde{n}] - \tilde{E}[n]}_{\Delta E_D} + \underbrace{\tilde{E}[n] - E[n]}_{\Delta E_F} \quad (1)$$

where ΔE_F is the functional error, defined as the error that would be found if the exact density were used, while ΔE_D is the (usually much smaller) contribution to the energy error due to the error in the self-consistent density.

For the purposes of this Letter, the general form of a 4-parameter double-hybrid functional (DH 4p) can be written as

$$E_{XC}^{DH4p} = E_X^{Slater} + \alpha(E_X^{HF} - E_X^{Slater}) + \beta(\tilde{E}_X^{GGA} - E_X^{Slater}) + \gamma\tilde{E}_C^{GGA} + \delta E_C^{ab\ initio} \quad (2)$$

where E_X^{Slater} is the local density approximation for exchange and E_X^{HF} is the HF exchange; $\tilde{E}_X^{GGA}$ and $\tilde{E}_C^{GGA}$ denote the approximate GGA exchange and correlation energy, respectively; and $E_C^{ab\ initio}$ is the correlation energy from an *ab initio* calculation such as MP2. There are numerous sets of optimal parameters depending on the choice and number of parameters. For example, the one-parameter DH functional suggested by Sharkas et al. has the form[28]

$$E_{XC}^{DH1p} = E_{XC}^{GGA} + \alpha(E_X^{HF} - E_X^{GGA}) + \alpha^2(E_C^{ab\ initio} - E_C^{GGA}) \quad (3)$$

which is based on *adiabatic connection* arguments.[28] (This is eq 2 where $\beta = 1 - \alpha$ and $\delta = 1 - \gamma = \alpha^2$.) The standard procedure then is to run self-consistent calculations of eq 2 without the *ab initio* correlation, but evaluate energies with the full DH expression on the orbitals.[28−30] The parameters are then chosen to minimize errors for specific molecular data sets. As we show, this assumes that density-driven differences between this and doing the entire procedure self-consistently are negligible.

Often, highly accurate densities required in eq 1 are too expensive to calculate. A practical measure of density sensitivity is given by[22,23,27]

$$\tilde{S} = |\tilde{E}[n_{LDA}] - \tilde{E}[n_{HF}]| \quad (4)$$

where tilde indicates a given functional approximation. Given the HF tendency to overlocalize, the LDA tendency to delocalize, and that both are nonempirical, $\tilde{S}$ is a practical guide to the density sensitivity of a given reaction and approximate functional. For small molecules, $\tilde{S} > 2$ kcal/mol implies density sensitivity and suggests DC-DFT will improve a functional's performance.[22] In such cases, usually the HF density is sufficient to produce improved energies (HF-DFT).

Illustration: Missing the Exact Solution for One Electron. In this section, we illustrate the dangers of ignoring the distinction between density-driven and functional errors in a simple, toy model: A simplified hybrid applied to the elementary case of H_2^+ as a function of bond length, which is a paradigm of self-interaction error, or more generally, delocalization error.[31,32] Standard semilocal approximations yield long-recognized catastrophic errors as the bond is stretched, missing entirely the dissociation limit (see Figure 1).[32] An HF calculation trivially gets this exactly right, because it is exact for (fully spin-polarized) one-electron systems.

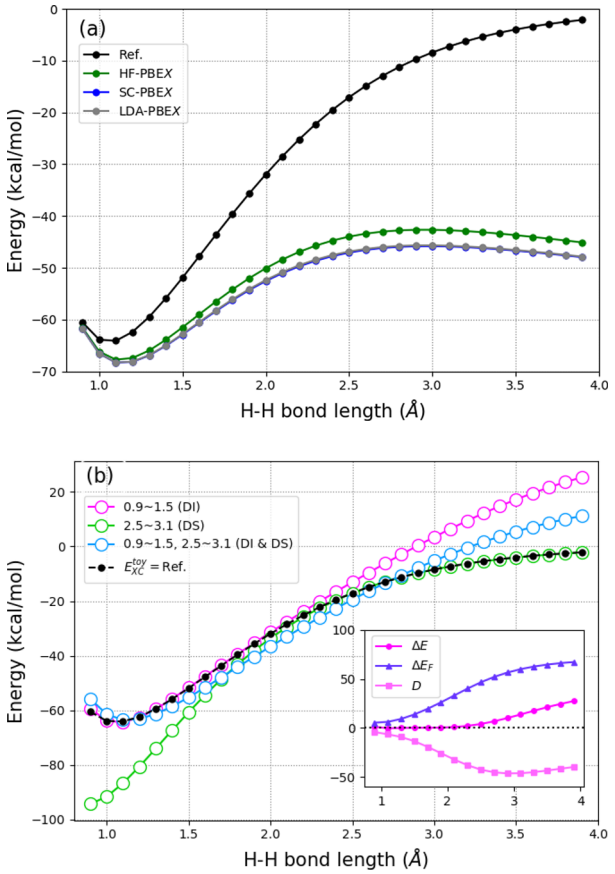


Figure 1. Potential energy surface (PES) of H_2^+ from (a) exactly (black), self-consistent PBE*X* (blue) and PBE*X* on the exact (HF) density (green) and on the LDA density (gray); (b) the toy functional of eq 2 with $\gamma = \delta = 0$ and no HF in the self-consistent density, with the *a* and *b* parameters optimized in different regions: (magenta) the density-insensitive (DI) region (0.9−1.5 Å), (green) the density-sensitive (DS) region (2.5−3.1 Å), (blue) combination of both DS and DI regions. The inset shows ΔE decomposition for the toy functional trained on the DI region. See also Figures S1 and S2 and Table S2.

Figure 1a shows the exact binding curve (black) easily found by HF, and two other curves of the PBE*X* evaluated either self-consistently (blue) or on the HF density (green). The largely irrelevant difference between blue and green curves shows that this is a true functional error, not a density-driven one. Even on the exact density, PBE*X* fails very badly as the bond is stretched. However, the difference in the two curves becomes greater than 2 kcal/mol at about 1.5 Å, showing a density sensitivity (the curve with LDA density is indistinguishable from the self-consistent curve) in this problem. (Standard HF-

https://dx.doi.org/10.1021/acs.jpclett.0c03545
J. Phys. Chem. Lett. 2021, 12, 800−807

DFT produces accurate curves for heteronuclear diatomics, not homonuclear ones.[19,20])

Now, to mimic the standard DH procedure, we perform self-consistent calculations without the HF contribution (because it yields the exact answer in this case), but we evaluate the energy with it included. We apply the DH philosophy to our H_2^+ molecule, using different separations to generate data sets. Because this is a one-electron system, we simplify the general DH form to just exchange, setting $\gamma = \delta = 0$ in eq 2, and use the PBE exchange[33] as a GGA. Figure 1b shows the results of training in the density-sensitive (stretched, DS) and density-insensitive (near equilibrium, DI) region of the binding curve. In each case, the optimal parametrization yields accurate energies on the training data but fails badly outside the training range. Even a combination of both equilibrium and stretched data does not help much.

How can this be happening? Obviously, if we set $\alpha = 1$ and $\beta = 0$ in eq 2, we get HF and so produce the exact answer. But, because the self-consistent calculation uses only a GGA form, which has an unbalanced self-interaction error as the bond is stretched, the exact result is never found. To quantify, we define

$$D[n'] = \tilde{E}[n'] - \tilde{E}[n] \tag{5}$$

generalizing[26] ΔE_D to arbitrary densities ($D[\tilde{n}] = \Delta E_D$, and $D[n] = 0$). We decompose the error for the functional trained near equilibrium, showing ΔE_F and D in the inset of Figure 1b. The optimal parameters (which are nonsensical, see Table S2 of the Supporting Information) keep the total error to a minimum in the training region where ΔE_F and D cancel each other by being about equal and opposite. Outside the training region of our H_2^+ curve, this artificial cancellation of errors fails badly. Obviously, we trivially solve this toy problem if we always train on the HF (exact, in this case) density instead of the self-consistent GGA density.

DFT-D3 for Weak Interactions. The D3 empirical correction of Grimme and co-workers has become a standard technique for improving the accuracy of DFT approximations when applied to noncovalent interactions.[25,34] While most such calculations are density-insensitive, DFT calculations of specific types of noncovalent interactions, such as halogen bonds, are plagued by density errors, which can be larger than the D3 correction itself.[23]

HF-DFT, as a simple form of DC-DFT, fixes this problem by replacing the SC density and orbitals with those of HF, for which semilocal functionals yield more accurate energies in such cases.[17,22,23,26] The differences between HF-DFT and exact DC-DFT are typically negligible relative to the improvements of HF-DFT over self-consistent DFT when density-driven errors are significant.[27] This does not imply that the pointwise accuracy of the underlying HF density is better than that of SC-DFT densities, even in cases of large density-driven errors.[22]

The example of ref 23 was an extreme case. Here we study the effects of density sensitivity on SC-DFT-D3 calculations of weak interactions when they are more subtle. We use 12 data sets (7 from the original D3 parametrization[25]) of noncovalent interactions (320 data points in total, see Table S1 of the Supporting Information).[35] The data points are classified as DS or DI based on their PBE sensitivity, S^{PBE} (see eq 4 and Figure S3). Only 46 are DS, and these are mostly from B30[36] and WATER27,[35] with only one such data point present in the data set used for the training of the original D3 parameters.

In Table 1, we demonstrate the importance of accounting for the density sensitivity when optimizing parameters for D3

Table 1. Mean Absolute Errors (kcal/mol) of PBE and Modifications on Density-Insensitive (DI) and Density-Sensitive (DS) Test Cases (Columns) versus Optimization on Various Databases (Rows), with Self-Consistent (SC) Densities on Left and HF Densities on Right[a]

	[SC]		[HF]	
opt. data set	DI	DS	DI	DS
without opt.	1.53	2.90	1.89	4.95
$D3_{orig}$	0.43	6.74	0.42	1.20
12DB	0.48	5.66	0.31	0.98
DS-12DB	1.47	2.96	0.38	0.87
DI-12DB	0.42	6.53	0.31	1.01

[a] $D3_{orig}$ denotes the original Grimme dataset; 12DB is our large (320 values) mixed dataset; DI-12DB are its 274 DI cases, and DS-12DB its 46 DS cases.

corrections. The first two numbers in the second column show the dramatic reduction in error in the PBE functional when the original D3 correction is made on the density-insensitive cases. The next entry shows that when we optimize over our much expanded database, the errors for DI cases are only slightly worse. But if we optimize specifically over our DS cases (fourth entry), this greatly worsens results on our DI test cases.

Moving over one column, we find results when tested on the DS cases. Now the original D3 parametrization yields a large (greater than 6 kcal/mol) error, demonstrating that density-sensitivity creates large errors. Even when optimized for DS cases, the error remains about 3 kcal/mol.

In the next column, we report the DI test results, but using HF densities instead of SC densities. In all cases of interest, the errors are slightly reduced once D3 with any of the parameters is turned on. The errors fall by more than a factor of 6 if the D3 is trained on the DI cases (from 1.89 to 0.31 kcal/mol). Furthermore, the differences between the optimal D3 parameters for DS and DI cases are much smaller when HF densities are used. Figure 2 shows the variation of the error with parameters. Figure 2a shows the usual case (SC densities on DI cases). Figure 2b is SC densities on DS cases, showing a totally different landscape. A green circle lying at the minimum of case (a) is denoted in all three panels. Figure 2c is HF densities on DS cases, showing about the same landscape as panel (a).

Finally, the fourth column of Table 1 shows results on the DS cases using HF densities. While overall these are much less accurate than the DI cases (by about a factor of 3), they are much better than those of column 2, which uses SC densities.

From these findings we can also see the effects of including DS cases in the training set. Their naive inclusion without the density correction via HF-DFT gives some improvements for DS cases at the cost of deteriorated accuracy for DI cases resulting from the abrupt changes in the optimal parameters. On the other hand, after the density correction is applied, the inclusion of DS cases in the training set improves their accuracy without the side effects for DI cases (Table 1) and without abrupt changes in the parameter landscape (Figure 2).

Most of the DS noncovalent complexes used in the training set in Table 1 belong to the B30 and WATER27 data sets. In Figure 3, we compare errors of SC-PBE and HF-PBE, with and without the (revised) D3 correction for binding energies of

https://dx.doi.org/10.1021/acs.jpclett.0c03545
J. Phys. Chem. Lett. 2021, 12, 800−807

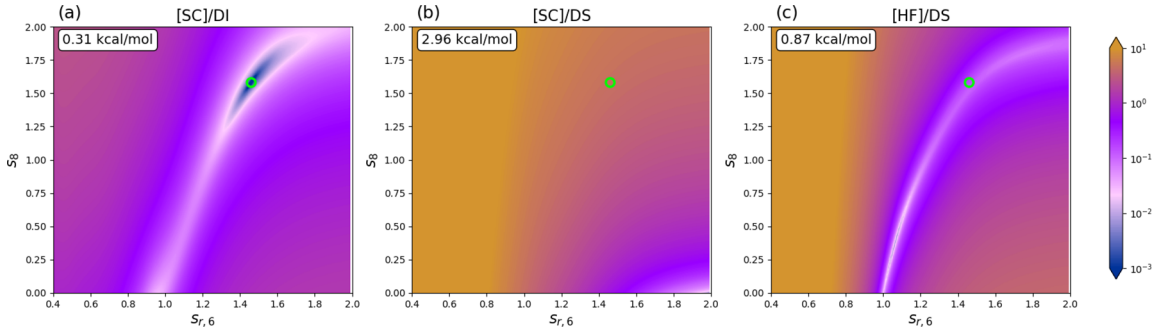


Figure 2. Mean absolute error (MAE) of PBE-D3 as a function of dispersion parameters,[25] for various densities and test sets: (a) self-consistent (SC) density on density-insensitive (DI) cases, (b) SC density for density-sensitive (DS) cases, and (c) HF density for DS cases. Contours are shifted by the minimum value (upper left corner) for clarity. The green circle is at the position of the global minimum of panel (a).

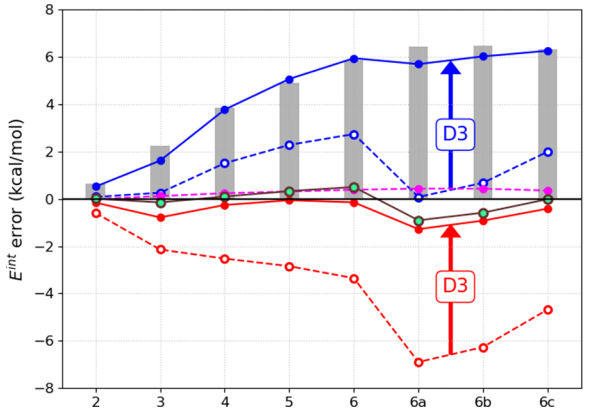


Figure 3. PBE binding energy error for small water clusters, $E^{int} = nE_{H_2O} - E_{(H_2O)_n}$ (n = 2–6), in WATER27 data set. Blue denotes self-consistent (PBE), while red is for the HF density (HF-PBE); dashed is without dispersion correction, while solid denotes with D3 (revised is similar to original). The gray bars show the density-sensitivity of eq 4. For comparison, we also show ωB97M-V (magenta) and BL1p (green, defined in text) results.

small water clusters of the WATER27 data set. The standard DFT calculations of these binding energies are highly DS, as shown by the large values for S^{PBE} shown in Figure 3.

We see that HF-DFT corrections are larger than D3 here and that D3 on self-consistent densities actually corrects in the wrong direction. HF-DFT-D3 reduces errors for the largest clusters from about 6 kcal/mol to less than 1 kcal/mol and thus delivers the performance comparable to ωB97M-V,[9] which includes nonlocal correlation,[37] and BL1p (a DH that will be introduced later).

Double-Hybrids. The energy functional of widely popular DHs (*e.g.*, eq 2) is typically evaluated on the hybrid density and orbitals found in a self-consistent calculation that neglects the $E_C^{ab\ initio}$ term.[28,29] In contrast, we find that HF-DHs obtained by applying a DH energy expression to the HF density and orbitals yield an overall accuracy competitive with their standard counterparts but remain accurate for cases where the standard DHs fail because of density sensitivity. We test the HF-DH idea with only one empirical parameter. On the basis of eq 3, we use here a combination of B88 exchange,[3] semilocal LYP correlation,[4] and MP2 correlation for $E_C^{ab\ initio}$.[38] We call this functional BL1p. Also, see Figure S4 to compare 1DH-BLYP (BL1p[SC]) of ref 28 and BL1p[HF]. Here we do not aim at reaching the accuracy limit of the HF-DH approach. This is already prohibited by a functional form of eq 3, which contains only one empirical parameter. Our goal is to show that this approach delivers an overall performance comparable to the standard DHs while not being plagued by large density-driven errors. Thus, we perform the optimization of α of eq 3 in an old-fashioned way, by training BL1p[HF] on the AE6 data set, containing atomization energies of 6 molecules.[39] The results of the training are shown in Figure 4. At α = 0, our

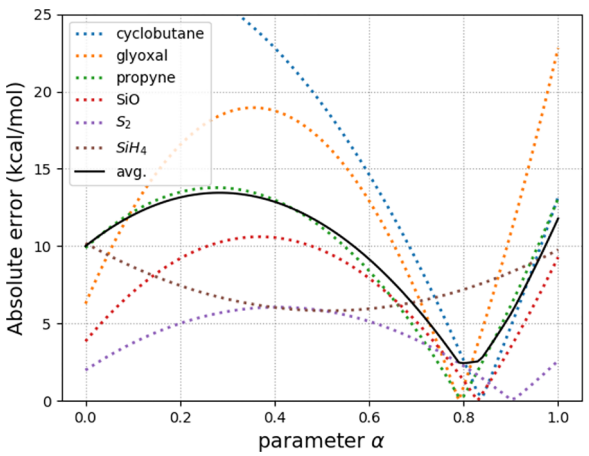


Figure 4. Absolute errors for the AE6 data set of BL1p as a function of α (see eq 3) for individual molecules (dashed lines). In the black solid curve, where the averaged errors are shown, the minimum is achieved at α = 0.82.

BL1p reduces to HF-BLYP, whereas at α = 1, it reduces to MP2. The optimal BL1p that minimizes MAE for AE6 has α = 0.82, which varies little between molecules, except for SiH_4 whose minimum is much shallower. Also, the MAE of optimal BL1p is about 7.5 kcal/mol smaller than the α = 0 case (HF-BLYP) and about 9 kcal/mol smaller than the α = 1 case (MP2).

In Figure 5, we compare the performance of BL1p with the standard DHs (B2PLYP[29] and XYG3[40]), hybrids (B3LYP, M06, M06-2X), and also the range-separated functional (ωB97M-V[9]), which we detail in the Supporting Information. This figure shows that the one-parameter BL1p, trained for only 6 atomization energies, yields an accuracy that is competitive with the standard DHs for all databases and works for noncovalent interactions, without using Grimme's empirical correction. For example, BL1p does well for the π–π stacking interactions where MP2 is known to severely overbind.[41,42] BL1p also contains MP2 correlation, but it has much less overbinding than MP2 and reduces errors by 1/4

https://dx.doi.org/10.1021/acs.jpclett.0c03545
J. Phys. Chem. Lett. 2021, 12, 800–807

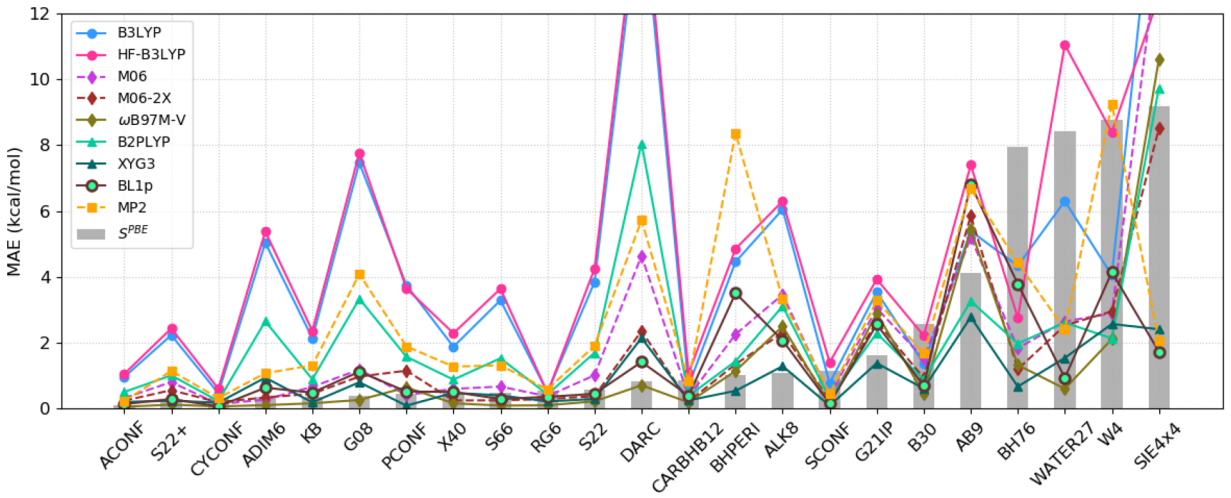


Figure 5. MAEs for several methods on many databases: BL1p, other double-hybrids (B2PLYP and XYG3), hybrids (B3LYP,M06, M06-2X), range-separated meta-GGA hybrid (ωB97M-V), and MP2.

(see Figure S7 in the Supporting Information). Usually, we recommend against using the HF density when it suffers from spin-contamination.[17] Nevertheless, for all data in this section, we include the spin-contaminated cases for fair comparison. Performance without spin-contaminated cases is shown in the Supporting Information (see Table S3).

Returning to our starting point, stretched NaCl is a prototypical case where self-consistent hybrids and GGAs are contaminated by large density errors.[19] These errors are typical of semilocal functionals for dissociating heterodimers.[43,44] HF densities fix this problem, and HF-DFT is able to dissociate heterodimers correctly.[19] From Figure 6, in contrast to

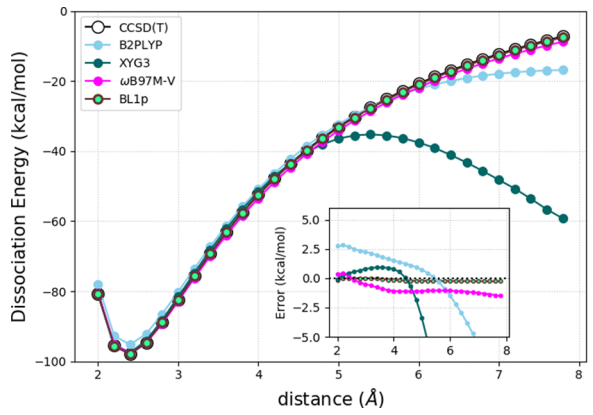


Figure 6. Dissociation curves of NaCl obtained from various approaches. For stretched bond lengths, standard double-hybrid functionals fail because of the density-driven errors (see ref 19). Note that only BL1p uses HF densities.

standard DHs (B2PLYP and XYG3 shown here) that fail at large bond lengths, our BL1p, as a representative of HF-DH, dissociates NaCl correctly (see also Figure S6).

Another case where BL1p outperforms other methods is the SIE4×4 data set, containing four positively charged dimers at four different separations, where standard DFT methods have large self-interaction error.[35] Figure 7 shows the dissociation curve of He_2^+, as a representative of this data set. First, the errors of the standard DFT methods for He_2^+ are almost entirely functional errors (see Figure S5), because they differ

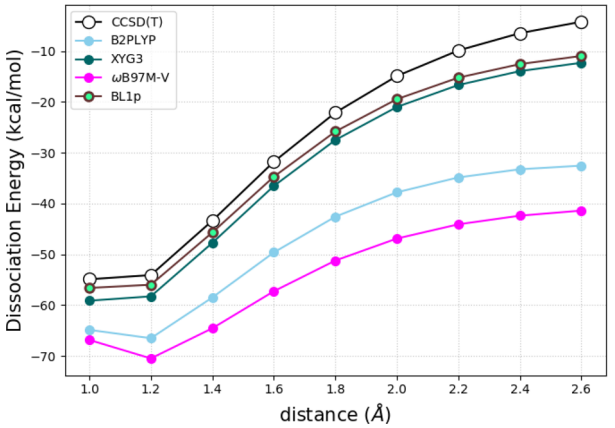


Figure 7. Dissociation curve of He_2^+ obtained from various functionals. See also Figure S5, showing that the errors of standard functionals for He_2^+ are mostly functional errors, because self-consistent results are almost identical to those when the functionals are applied to accurate densities (obtained from the Kohn−Sham inversion scheme from the CCSD wave function[27]).

little between accurate and self-consistent densities. The accurate densities are obtained by Kohn−Sham inversion from CCSD densities.[27] In this way, the source of error of the standard DFT for He_2^+ is very different from that of stretched NaCl. Figure 7 shows that, even though these are not density-driven errors, the error of BL1p and XYG3 for He_2^+ is much smaller than that of other approaches because of the high amount of exact exchange.

Range-Separated versus Conventional Hybrids. We have shown a number of examples where large density-driven errors of conventional (global) hybrid functionals are substantially reduced when they are evaluated on the HF instead of SC densities. Range-separated hybrids (RSHs) often use 100% of the HF exchange in the long range (lr)[9,45,46] and so should suffer less from density-driven errors.[47] To test this, we use ωB97M-V as a representative of RSHs,[9] given its remarkable performance for many of the databases in Figure 5. We will compare ωB97M-V with B97M-V, its conventional analogue.[48] The density-driven errors of ωB97M-V and B97M-V are shown in Figure 8 for our two standard cases, with PBE and B3LYP also shown for comparison. For H_2^+, the HF density is

https://dx.doi.org/10.1021/acs.jpclett.0c03545
J. Phys. Chem. Lett. 2021, 12, 800−807

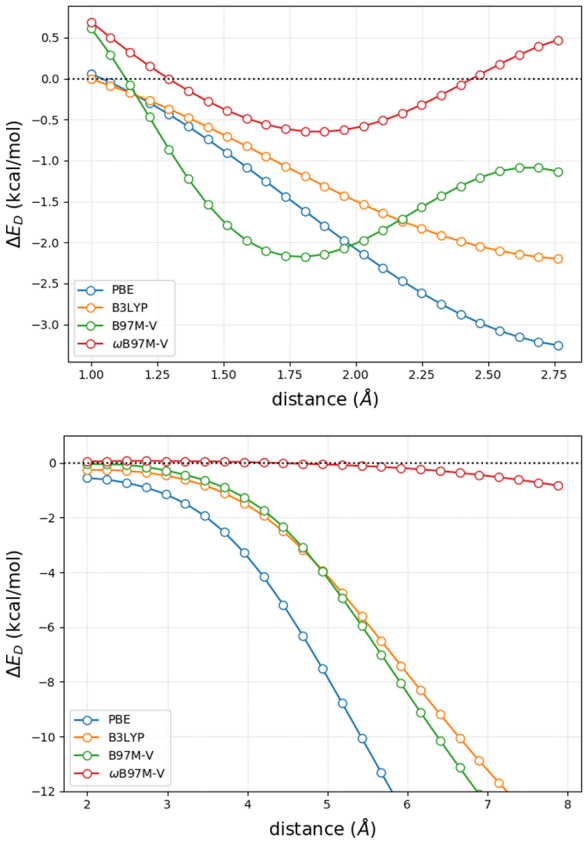


Figure 8. Density-driven errors (see eq 1) of selected functionals along the dissociation curves of:e H_2^+ (top panel) and NaCl (bottom panel). For H_2^+, the (exact) HF density is used to extract the density-driven errors. For NaCl, we use CCSD as a reference in tandem with the Kohn−Sham inversion scheme described in ref 27 to obtain the "exact" density and orbitals needed to isolate density-driven errors.

exact, while for NaCl, we invert the accurate Kohn−Sham density from CCSD.[27] In each case, the density-driven error of ωB97M-V is much smaller than that of the other functionals. It does not vanish, because of the semilocal part of the functional. We see similar behavior for larger systems where the error of conventional hybrid functionals is contaminated by the densities and is much smaller in ωB97M-V. Sensitivity plots are used as a diagnostic tool for density-driven errors, and in Figure 9 we show that the sensitivity of ωB97M-V for the WATER27 complexes is a fraction of that of B3LYP and B97M-V.

We have shown the dangers of ignoring density errors in the construction of empirical approximations. In our simple H_2^+ example, a parametrized semilocal functional trained on a limited region of the H_2^+ binding curve fails in all other regions. Even high accuracy in the training region results from an enforced error cancellation between the density and functional error (eq 1), which fails outside this region. We found that the standard DFT with empirical D3 corrections breaks down in *density-sensitive* calculations of noncovalent systems but is fixed by using the HF density.

We also found that resilience to density-driven errors could be achieved with simple 1-parameter double-hybrids, once they are trained and applied to HF densities. As always, our use of HF densities does not imply that they are point-wise more accurate than self-consistent densities, but simply that they yield more accurate energetics when a reaction is density-sensitive. Our BL1p is trained only on atomization energies of only 6 molecules, but its accuracy is comparable to the standard doubled hybrids tested here. Moreover, ωB97M-V outperforms BL1p for most of the data sets considered in Figure 5, except for the SIE4×4 data set, where BL1p does much better. BL1p would also be beaten by ωB97(2), a very recent highly accurate DH designed to improve over ωB97M-V.[49] Given its excellent performance,[49,50] we expect it to beat BL1p on most of the data sets, but not SIE4×4.

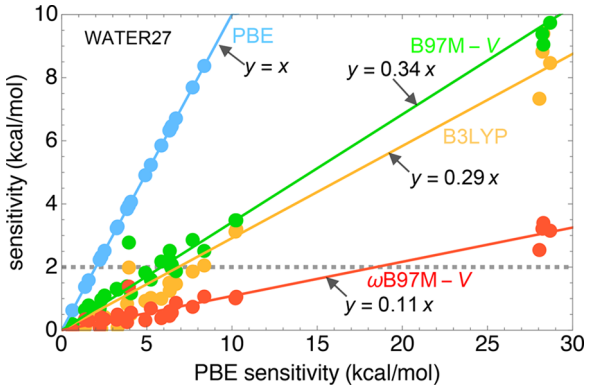


Figure 9. Sensitivity (see eq 4) of selected functionals vs PBE sensitivities for binding energies of the WATER27 clusters.

Our goal here is not the introduction of a new empirical XC functional, but to illustrate contamination due to density errors in fitting procedures and to show how minimizing the functional error can improve the performance of empirical functionals. Thus, our primitively optimized BL1p does not reach the accuracy limit of the HF-DH class of functionals. Technical advances in optimization and larger parameter spaces could further improve its accuracy. Furthermore, to improve HF-DHs, one may also use the new insights into functionals that explicitly depend on the HF density obtained from the adiabatic connection that has the MP2 theory as its weak-interaction expansion.[51,52] Finally, we have found that using 100% of HF exchange in range-separated hybrids means they suffer much less from density-driven errors than their conventional counterparts.

In summary, DFT energy errors can be separated into functional and density-driven using DC-DFT. To avoid inaccuracies, empirical functionals can be trained on functional errors only, where practical. In cases of large density sensitivity, HF densities (unless flawed by, *e.g.*, spin-contamination) are typically more useful than self-consistent semilocal densities. With 100% exchange at large distances, range-separated functionals are relatively density-insensitive and suffer much less from these issues.

COMPUTATIONAL DETAILS

All HF, DFT, HF-DFT, and MP2 calculations have been performed with the TURBOMOLE v7.0.2.[53] and PYSCF v1.7.2.[54] The following functionals have been used in DFT and HF-DFT calculations: LDA (SVWN[1,55]), GGA (PBE[33] and BLYP[3,4]), mGGA (TPSS[56]), hybrids (B3LYP,[6] PBE0,[57] M06, M06-2X,[58] B97M-V,[48] ωB97M-V,[9] B2PLYP,[29] and XYG3[40]). The scripts for performing HF-DFT energy calculations are available.[59] Unless otherwise stated, the def2-QZVPPD basis set has been used. All geometries and the multiplicities except for the AE6[39] have been taken from ref 35. Further

https://dx.doi.org/10.1021/acs.jpclett.0c03545
J. Phys. Chem. Lett. 2021, 12, 800−807

computational details can be found in the Supporting Information.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.0c03545.

Data set description, optimized parameters and mean absolute error for H_2^+, and mean absolute error value of Figure 5 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Eunji Sim − *Department of Chemistry, Yonsei University, Seoul 03722, Korea;* orcid.org/0000-0002-4139-0960; Email: esim@yonsei.ac.kr

Authors

Suhwan Song − *Department of Chemistry, Yonsei University, Seoul 03722, Korea;* orcid.org/0000-0002-7768-6181

Stefan Vuckovic − *Departments of Chemistry and of Physics, University of California, Irvine, California 92697, United States;* orcid.org/0000-0002-0768-9176

Kieron Burke − *Departments of Chemistry and of Physics, University of California, Irvine, California 92697, United States;* orcid.org/0000-0002-6159-0054

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.0c03545

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work at Yonsei University was supported by a grant from the Korean Research Foundation (NRF-2020R1A2C2007468 and NRF-2020R1A4A1017737). K.B. acknowledges funding from NSF (CHEM 1856165). S.V. acknowledges funding from the Rubicon project (019.181EN.026), which is financed by The Netherlands Organisation for Scientific Research (NWO). We thank Professor Martin Head-Gordon and Dr. Narbe Mardirossian for stimulating discussions.

■ REFERENCES

(1) Dirac, P. A. Note on Exchange Phenomena in the Thomas Atom. *Math. Proc. Cambridge Philos. Soc.* **1930**, *26*, 376−385.

(2) Kohn, W.; Sham, L. J. Self-consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133.

(3) Becke, A. D. Density-functional Exchange-energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.

(4) 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: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.

(5) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(6) 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*, 11623−11627.

(7) Yu, H. S.; Li, S. L.; Truhlar, D. G. Perspective: Kohn-Sham Density Functional Theory Descending a Staircase. *J. Chem. Phys.* **2016**, *145*, 130901.

(8) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115*, 2315−2372.

(9) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A Combinatorially Optimized, Range-separated Hybrid, meta-GGA Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(10) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density Functional Theory is Straying from the Path Toward the Exact Functional. *Science* **2017**, *355*, 49−52.

(11) Clementi, E.; Chakravorty, S. J. A Comparative Study of Density Functional Models to Estimate Molecular Atomization Energies. *J. Chem. Phys.* **1990**, *93*, 2591−2602.

(12) Oliphant, N.; Bartlett, R. J. A Systematic Comparison of Molecular Properties Obtained Using Hartree-Fock, A Hybrid Hartree-Fock Density-Functional-Theory, and Coupled-Cluster Methods. *J. Chem. Phys.* **1994**, *100*, 6550−6561.

(13) Janesko, B. G.; Scuseria, G. E. Hartree-Fock Orbitals Significantly Improve the Reaction Barrier Heights Predicted by Semilocal Density Functionals. *J. Chem. Phys.* **2008**, *128*, 244112.

(14) Brittain, D. R.; Lin, C. Y.; Gilbert, A. T.; Izgorodina, E. I.; Gill, P. M.; Coote, M. L. The Role of Exchange in Systematic DFT Errors for Some Organic Reactions. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1138−1142.

(15) Kim, M.-C.; Sim, E.; Burke, K. Communication: Avoiding Unbound Anions in Density Functional Calculations. *J. Chem. Phys.* **2011**, *134*, 171103.

(16) Verma, P.; Perera, A.; Bartlett, R. J. Increasing the Applicability of DFT I: Non-variational Correlation Corrections from Hartree-Fock DFT for Predicting Transition States. *Chem. Phys. Lett.* **2012**, *524*, 10−15.

(17) Kim, M.-C.; Sim, E.; Burke, K. Understanding and Reducing Errors in Density Functional Calculations. *Phys. Rev. Lett.* **2013**, *111*, 073003.

(18) Kim, M.-C.; Sim, E.; Burke, K. Ions in Solution: Density Corrected Density Functional Theory (DC-DFT). *J. Chem. Phys.* **2014**, *140*, 18A528.

(19) Kim, M.-C.; Park, H.; Son, S.; Sim, E.; Burke, K. Improved DFT Potential Energy Surfaces via Improved Densities. The. *J. Phys. Chem. Lett.* **2015**, *6*, 3802−3807.

(20) Wasserman, A.; Nafziger, J.; Jiang, K.; Kim, M.-C.; Sim, E.; Burke, K. The Importance of Being Self-Consistent. *Annu. Rev. Phys. Chem.* **2017**, *68*, 555−581.

(21) Song, S.; Kim, M.-C.; Sim, E.; Benali, A.; Heinonen, O.; Burke, K. Benchmarks and Reliable DFT Results for Spin Gaps of Small Ligand Fe (II) Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2304−2311.

(22) Sim, E.; Song, S.; Burke, K. Quantifying Density Errors in DFT. *J. Phys. Chem. Lett.* **2018**, *9*, 6385−6392.

(23) Kim, Y.; Song, S.; Sim, E.; Burke, K. Halogen and Chalcogen Binding Dominated by Density-Driven Errors. *J. Phys. Chem. Lett.* **2019**, *10*, 295−301.

(24) Nagai, R.; Akashi, R.; Sugino, O. Completing Density Functional Theory by Machine Learning Hidden Messages from Molecules. *npj Computational Materials* **2020**, *6*, 43.

(25) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(26) Vuckovic, S.; Song, S.; Kozlowski, J.; Sim, E.; Burke, K. Density Functional Analysis: The Theory of Density-corrected DFT. *J. Chem. Theory Comput.* **2019**, *15*, 6636−6646.

(27) Nam, S.; Song, S.; Sim, E.; Burke, K. Measuring Density-Driven Errors Using Kohn−Sham Inversion. *J. Chem. Theory Comput.* **2020**, *16*, 5014−5023.

(28) Sharkas, K.; Toulouse, J.; Savin, A. Double-hybrid Density-functional Theory Made Rigorous. *J. Chem. Phys.* **2011**, *134*, 064113.

(29) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(30) Martin, J. M.; Santra, G. Empirical Double-Hybrid Density Functional Theory: A 'Third Way'in Between WFT and DFT. *Isr. J. Chem.* **2020**, *60*, 787.
(31) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into Current Limitations of Density Functional Theory. *Science* **2008**, *321*, 792−794.
(32) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289−320.
(33) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(34) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-charge Dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150*, 154122.
(35) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A Look at the Density Functional Theory Zoo with the Advanced GMTKN55 Database for General Main Group Thermochemistry, Kinetics and Noncovalent Interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184−32215.
(36) Bauza, A.; Alkorta, I.; Frontera, A.; Elguero, J. On the Reliability of Pure and Hybrid DFT Methods for the Evaluation of Halogen, Chalcogen, and Pnicogen Bonds Involving Anionic and Neutral Electron Donors. *J. Chem. Theory Comput.* **2013**, *9*, 5201−5210.
(37) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals Density Functional: The Simpler the Better. *J. Chem. Phys.* **2010**, *133*, 244103.
(38) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618−622.
(39) Lynch, B. J.; Truhlar, D. G. Small Representative Benchmarks for Thermochemical Calculations. *J. Phys. Chem. A* **2003**, *107*, 8996−8999.
(40) Zhang, Y.; Xu, X.; Goddard, W. A. Doubly Hybrid Density Functional for Accurate Descriptions of Nonbond Interactions, Thermochemistry, and Thermochemical Kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4963−4968.
(41) Cybulski, S. M.; Lytle, M. L. The Origin of Deficiency of the Supermolecule Second-order Møller-Plesset Approach for Evaluating Interaction Energies. *J. Chem. Phys.* **2007**, *127*, 141102.
(42) Vuckovic, S.; Fabiano, E.; Gori-Giorgi, P.; Burke, K. MAP: An MP2 Accuracy Predictor for Weak Interactions from Adiabatic Connection Theory. *J. Chem. Theory Comput.* **2020**, *16*, 4141−4149.
(43) Gritsenko, O. V.; Baerends, E. J. Effect of Molecular Dissociation on the Exchange-correlation Kohn-Sham Potential. *Phys. Rev. A: At., Mol., Opt. Phys.* **1996**, *54*, 1957.
(44) Giarrusso, S.; Vuckovic, S.; Gori-Giorgi, P. Response Potential in the Strong-interaction Limit of DFT: Analysis and Comparison with the Coupling-constant Average. *J. Chem. Theory Comput.* **2018**, *14*, 4151.
(45) Vydrov, O. A.; Scuseria, G. E. Assessment of a Long-range Corrected Hybrid Functional. *J. Chem. Phys.* **2006**, *125*, 234109.
(46) 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*, 9904.
(47) Johnson, E. R.; Otero-de-la Roza, A.; Dale, S. G. Extreme Density-driven Delocalization Error for a Model Solvated-electron System. *J. Chem. Phys.* **2013**, *139*, 184116.
(48) Mardirossian, N.; Head-Gordon, M. Mapping the Genome of meta-Generalized Gradient Approximation Density Functionals: The Search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.
(49) Mardirossian, N.; Head-Gordon, M. Survival of the Most Transferable at the Top of Jacob's Ladder: Defining and Testing the ω B97M(2) Double Hybrid Density Functional. *J. Chem. Phys.* **2018**, *148*, 241736.
(50) Santra, G.; Sylvetsky, N.; Martin, J. M. L. Minimally Empirical Double-Hybrid Functionals Trained against the GMTKN55 Database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129−5143.
(51) Seidl, M.; Giarrusso, S.; Vuckovic, S.; Fabiano, E.; Gori-Giorgi, P. Communication: Strong-interaction Limit of an Adiabatic Connection in Hartree-Fock Theory. *J. Chem. Phys.* **2018**, *149*, 241101.
(52) Daas, T. J.; Grossi, J.; Vuckovic, S.; Musslimani, Z. H.; Kooi, D. P.; Seidl, M.; Giesbertz, K. J.; Gori-Giorgi, P. Large Coupling-strength Expansion of the Møller-Plesset Adiabatic Connection: From Paradigmatic Cases to Variational Expressions for the Leading Terms. *J. Chem. Phys.* **2020**, *153*, 214112.
(53) TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007; TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
(54) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; et al. PySCF: the Python-based Simulations of Chemistry Framework. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1340.
(55) 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*, 1200−1211.
(56) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
(57) Burke, K.; Ernzerhof, M.; Perdew, J. P. The Adiabatic Connection Method: A Non-empirical Hybrid. *Chem. Phys. Lett.* **1997**, *265*, 115−120.
(58) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(59) http://tccl.yonsei.ac.kr/mediawiki/index.php/DC-DFT.

https://dx.doi.org/10.1021/acs.jpclett.0c03545
J. Phys. Chem. Lett. 2021, 12, 800−807