Supporting Information for "Density Sensitivity of Empirical Functionals"

Suhwan Song,[†] Stefan Vuckovic,[‡] Eunji Sim,[*,†] and Kieron Burke[‡]

[†]*Department of Chemistry, Yonsei University, 50 Yonsei-ro Seodaemun-gu, Seoul 03722, Korea*

[‡]*Departments of Chemistry and of Physics, University of California, Irvine, CA 92697, USA*

E-mail: esim@yonsei.ac.kr

List of Tables

List of Figures

Computational Details

All HF, DFT, HF-DFT, and MP2 calculations have been performed with the TURBOMOLE v7.0.2.[1] and PYSCF v1.7.2.[2] The following functionals have been used in DFT and HF-DFT calculations: LDA (SVWN[3,4]), GGA (PBE,[5] BLYP[6,7]), mGGA (TPSS[8]), hybrids (B3LYP,[9] PBE0,[10] M06, M06-2X,[11] B97M-V,[12] ωB97M-V,[13] B2PLYP,[14] and XYG3[15]). The scripts for performing HF-DFT energy calculations are available.[16] Unless otherwise stated, the def2-QZVPPD basis set has been used. For accelerating the self-consistent field (SCF) procedure, we adopted the resolution of the identity approximation (RI-J) with def2-QZVPPD auxiliary basis set for G08, PCONF, SCONF, and WATER27 dataset. For the KB dataset, we omit the noble gas and the S22 dataset from the original KB65. The energy and one-electron density convergence threshold have been set to 1e-8 and 1e-6 a.u., respectively. Numerical quadrature grids of size 4 have been used (grid size 4 in TURBOMOLE). For VV10 correlation, 99 radial shells with 590 angular grid points per shell are used with the SG1 prune. For all open shell calculation, unrestricted scheme is used. The parameters of Table S2 are first found in the global optimizer *shgo* and then optimized locally with the *Nelder-Mead* method in the SCIPY package with the 1e-8 convergence criterion. For revised D3, we scanned $0.00 < s_{r,6} < 2.00$ and $0.00 < s_8 < 2.00$ (0.01 grid spacing) for all 320 datapoints in 12 databases (DB) (marked by an asterisks in table S1) for all XC functionals except ωB97M-V, XYG3 and HF-DHs. The α value for HF-DHs is optimized with the *slsqp* method in the SCIPY package. All geometries and the multiplicities except for the AE6[17] have been taken from Ref.[18]

Table S1: Data sets used in this work. The asterisks indicate the data sets used for optimization of the original D3 parameters by Grimme and co-workes as described in Ref.[19] For the revised D3, additional 5 datasets (KB, S66, X40, B30, and WATER27) are used as a training set and marked by a dagger.

DB	no. points	Description
RG6*	6	rare gases
ACONF*	15	alkane conformers
S22+*	66	non-covalent interaction
CYCONF*	10	cysteine conformers
ADIM6*	6	alkane dimers
KB†	27	non-covalent interactions
G08	6	pi-interaction of small carbon complexes
PCONF*	10	peptide conformers
S66†	66	non-covalent interactions
X40†	40	halogen interactions
S22	22	noncovalently bound dimers
DARC	14	Diels–Alder reactions
CARBHB12	12	hydrogen-bonded
BHPERI	26	pericyclic reactions
G21IP	36	adiabatic ionization potentials
ALK8	8	alkaline compounds
SCONF*	17	sugar conformers
B30†	30	non-covalent interactions for halogen
AB9	9	anomalous barrier height
SIE4x4	16	self-interaction-error
BH76	76	reaction barrier height
W4	140	atomization energies
WATER27†	27	water interactions

Table S2: Optimized parameters in Fig. 1 and mean absolute error (MAE) in kcal/mol. Note that when the toy model is applied to the HF densities, we get $a = 1$ and $b = 0$ regardless of the training set, and thereby both the density-driven and functional errors are eliminated.

$R_{\mathrm{H-H}}$	a	b	training	test	all
0.9∼1.5	1.397	6.799	0.30	10.15	7.93
2.5∼3.1	1.237	-1.532	0.02	6.64	5.14
0.9∼1.5 & 2.5∼3.1	1.218	4.665	2.22	6.85	4.76

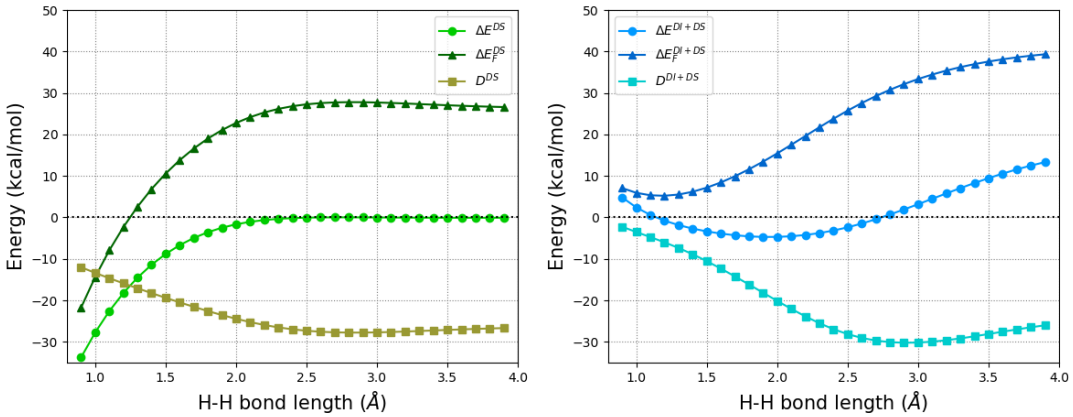

Figure S1: D and ΔE_F error decomposition for H_2^+ dissociation curve of the empirical toy functional trained on the DS region (left panel) and on the combination of DS and DI regions (right panel) when applied to H_2^+ along the dissociation curve . The same result, but for the functional trained on the DI region is shown in the inset of the bottom panel of of Fig. 1.

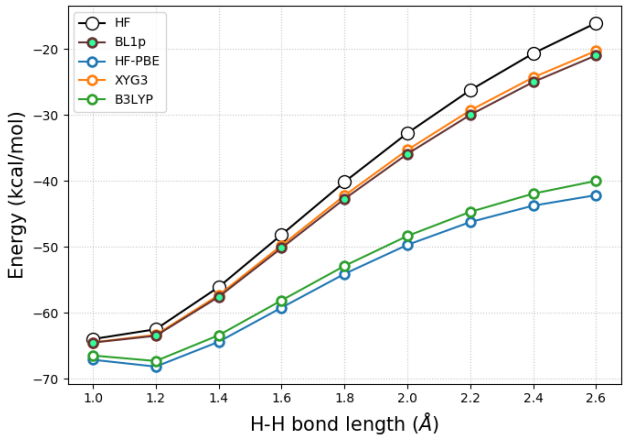

Figure S2: H_2^+ potential energy surface for the various methods.

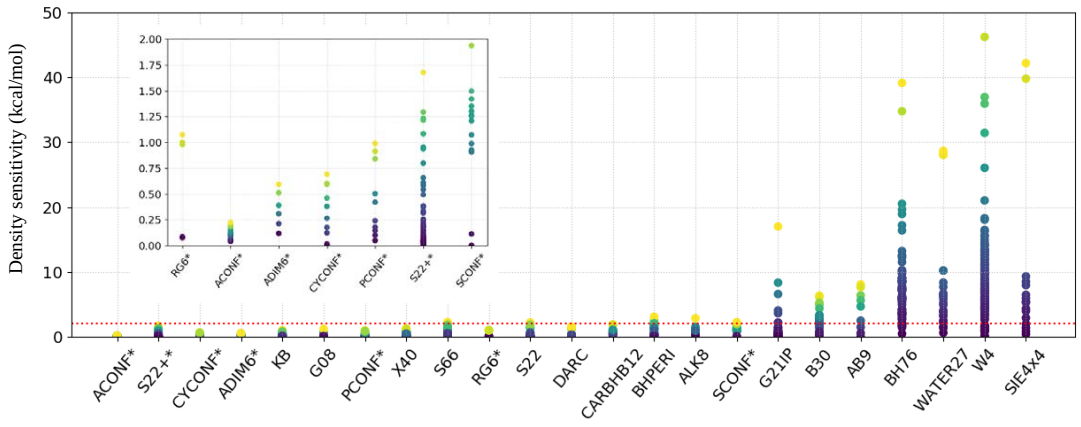

Figure S3: Density sensitivity (S^{PBE}) of PBE for the 23 databases in Table S1. The red dotted line denotes 2 kcal/mol. If S^{PBE} is greater than 2 kcal/mol, it is considered density sensitive.[20] As we go from the left to the right, the averaged S^{PBE}) of the databases increases. The databases used for the training of the Grimme's original D3 parameters are marked with an asterisk. The same databases are also shown in the inset.

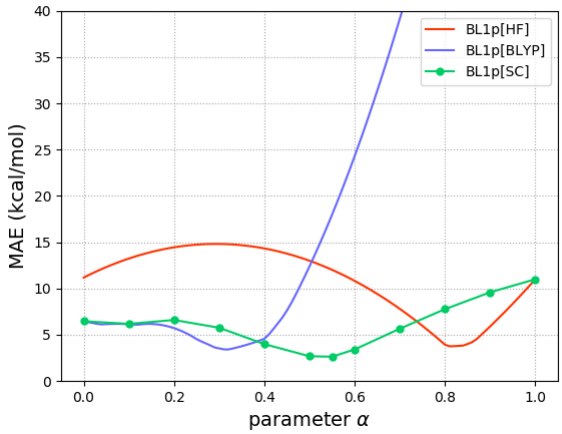

Figure S4: BL1p MAE for the AE6 dataset in Ref.[21] HF, BLYP and SC density is used. (SC denotes the self-consistent density from $E_{XC} = \alpha E_X^{HF} + (1-\alpha) E_X^{B88} + (1-\alpha^2) E_C^{LYP}$.) Reference values are from Ref.[22]

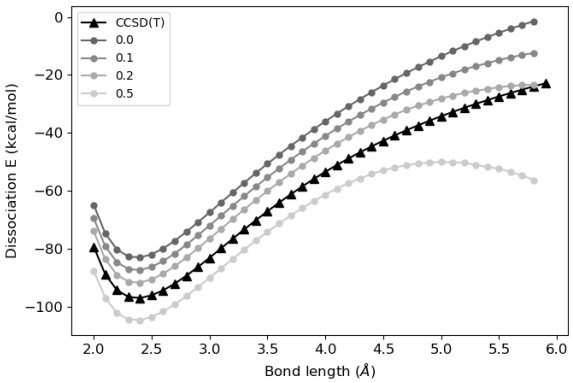

Figure S6: Na-Cl dissociation curve for the XYG3 with different MP2 admixture. The larger the MP2 portion, the quicker the XYG3 DH bends down.

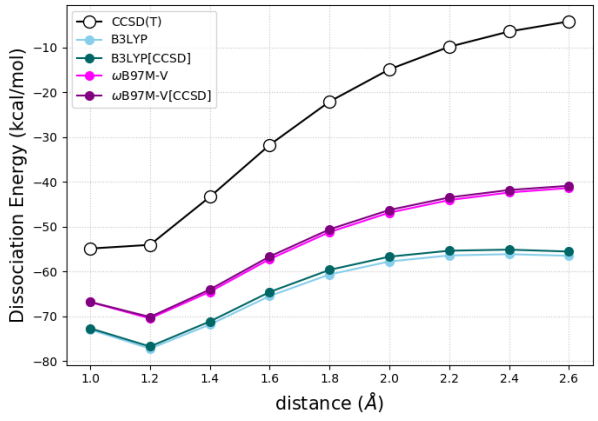

Figure S5: Dissociation curve of He_2^+. Note that CCSD density is used within the WY KS-inversion method to obtain accurate KS orbitals. Detailed information about KS-inversion can be found in the Ref.[23]

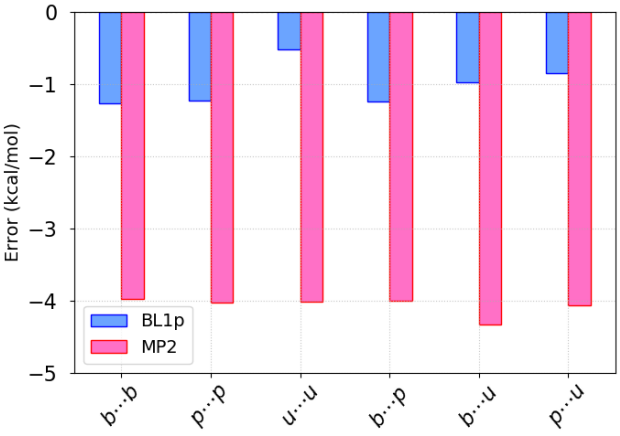

Figure S7: $\pi - \pi$ interaction energy error of BL1p and MP2 for the dimer formation of benzene (b), pyridine (p), and uracil (u). Negative signs denote overbinding.

	ACONF	S22+	CYCONF	ADIM6	KB	G08	PCONF	X40	S66	RG6	S22	DARC
BL1p	0.156	0.263	0.067	0.724	0.457	1.061	0.456	0.487	0.279	0.350	0.393	1.299
BL1p*	0.156	0.263	0.067	0.724	0.457	1.061	0.456	0.487	0.279	0.350	0.393	1.299
XYG3	0.205	0.237	0.179	0.937	0.193	0.801	0.090	0.476	0.397	0.223	0.293	2.146
wB97M-V	0.056	0.121	0.065	0.106	0.161	0.263	0.644	0.161	0.092	0.101	0.220	0.709
B3LYP	0.959	2.231	0.469	5.034	2.137	7.467	3.750	1.878	3.291	0.494	3.842	15.762
B3LYP[HF]	1.052	2.435	0.606	5.375	2.364	7.771	3.648	2.288	3.639	0.482	4.248	16.380
M06	0.318	0.807	0.135	0.275	0.661	1.186	0.411	0.601	0.669	0.375	1.013	4.642
M06-2X	0.240	0.561	0.193	0.341	0.452	0.973	1.152	0.249	0.253	0.288	0.377	2.358
B2PLYP	0.505	1.008	0.230	2.654	0.889	3.337	1.575	0.880	1.529	0.432	1.694	8.059
MP2	0.228	1.145	0.320	1.076	1.315	4.087	1.888	1.276	1.315	0.582	1.908	5.720

	CARBHB12	BHPERI	ALK8	SCONF	G21IP	B30	AB9	BH76	WATER27	W4	SIE4x4
BL1p	0.379	3.289	2.043	0.168	2.571	0.654	6.830	3.751	1.003	4.183	1.860
BL1p*	0.379	3.289	2.043	0.168	1.991	0.654	5.851	2.492	1.003	3.077	2.044
XYG3	0.248	0.543	1.295	0.092	1.382	0.611	2.785	0.674	1.541	2.572	2.411
wB97M-V	0.189	1.145	2.501	0.129	2.901	0.450	5.438	1.327	0.603	2.113	10.608
B3LYP	0.700	4.459	6.051	0.771	3.541	1.369	5.391	4.336	6.306	4.046	17.475
B3LYP[HF]	1.076	4.855	6.303	1.400	3.928	2.216	7.403	2.759	11.051	8.391	12.512
M06	0.363	2.258	3.464	0.381	3.052	1.522	5.199	1.831	2.647	2.894	14.146
M06-2X	0.212	1.351	2.300	0.254	2.645	0.931	5.855	1.196	2.527	2.955	8.507
B2PLYP	0.385	1.440	3.113	0.305	2.294	0.886	3.265	1.975	2.613	2.132	9.737
MP2	0.838	8.369	3.318	0.458	3.283	1.675	6.678	4.432	2.399	9.242	2.064

Table S3: **Mean absolute error (MAE, kcal/mol) of 23 databases for methods in Fig. 5. Note that BL1p* omits spin contamination cases for G21IP, AB9, BH76, W4, and SIE4x4.**

References

(1) 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`.

(2) 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. 2017; `https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1340`.

(3) Dirac, P. A. Note on Exchange Phenomena in the Thomas Atom. Mathematical Proceedings of the Cambridge Philosophical Society. 1930; pp 376–385.

(4) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Canadian Journal of Physics* **1980**, *58*, 1200–1211.

(5) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865.

(6) Becke, A. D. Density-functional Exchange-energy Approximation with Correct Asymptotic Behavior. *Physical Review A* **1988**, *38*, 3098–3100.

(7) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-energy Formula into a Functional of the Electron Density. *Physical Review B* **1988**, *37*, 785–789.

(8) 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. *Physical Review Letters* **2003**, *91*, 146401.

(9) 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. *The Journal of Physical Chemistry* **1994**, *98*, 11623–11627.

(10) Burke, K.; Ernzerhof, M.; Perdew, J. P. The Adiabatic Connection Method: A Non-empirical Hybrid. *Chemical Physics Letters* **1997**, *265*, 115–120.

(11) 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. *Theoretical Chemistry Accounts* **2008**, *120*, 215–241.

(12) Mardirossian, N.; Head-Gordon, M. Mapping the Genome of meta-Generalized Gradient Approximation Density Functionals: The Search for B97M-V. *The Journal of Chemical Physics* **2015**, *142*, 074111.

(13) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A Combinatorially Optimized, Range-separated Hybrid, meta-GGA Density Functional with VV10 Nonlocal Correlation. *The Journal of Chemical Physics* **2016**, *144*, 214110.

(14) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-order Correlation. *The Journal of Chemical Physics* **2006**, *124*, 034108.

(15) Zhang, Y.; Xu, X.; Goddard, W. A. Doubly Hybrid Density Functional for Accurate Descriptions of Nonbond Interactions, Thermochemistry, and Thermochemical Kinetics. *Proceedings of the National Academy of Sciences* **2009**, *106*, 4963–4968.

(16) `http://tccl.yonsei.ac.kr/mediawiki/index.php/DC-DFT`.

(17) Lynch, B. J.; Truhlar, D. G. Small Representative Benchmarks for Thermochemical Calculations. *The Journal of Physical Chemistry A* **2003**, *107*, 8996–8999.

(18) 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. *Physical Chemistry Chemical Physics* **2017**, *19*, 32184–32215.

(19) 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. *The Journal of Chemical Physics* **2010**, *132*, 154104.

(20) Sim, E.; Song, S.; Burke, K. Quantifying Density Errors in DFT. *The Journal of Physical Chemistry Letters* **2018**, *9*, 6385–6392.

(21) Sharkas, K.; Toulouse, J.; Savin, A. Double-hybrid Density-functional Theory Made Rigorous. *The Journal of Chemical Physics* **2011**, *134*, 064113.

(22) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid meta-GGA Density Functionals by Range Separation. *The Journal of Physical Chemistry Letters* **2011**, *2*, 2810–2817.

(23) Nam, S.; Song, S.; Sim, E.; Burke, K. Measuring Density-Driven Errors Using Kohn–Sham Inversion. *Journal of Chemical Theory and Computation* **2020**, *16*, 5014–5023.