Untangling Sources of Error in the Density-Functional Many-Body Expansion

Dustin R. Broderick and John M. Herbert*
Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210 USA

Abstract

Development of meta-generalized gradient approximations (meta-GGAs) has generally led to more accurate density-functional approximations, albeit ones that have more stringent requirements for the quadrature grids that are used to evaluate the exchange-correlation energy. Here, we demonstrate that grid-induced errors are amplified when meta-GGAs are used in conjunction with a many-body expansion, which is a popular means to parameterize classical force fields using electronic structure calculations. At the same time, delocalization errors are exacerbated by the many-body expansion, leading to exaggerated estimates of nonadditive n-body interactions, as illustrated here for anion–water clusters using the meta-GGA functionals SCAN and ωB97X-V. Standard grids that are typically accurate for noncovalent interactions with meta-GGA functionals result in runaway error accumulation when used with the many-body expansion. Denser grids eliminate this problem and expose the inherent self-interaction error, which must be mitigated using other strategies that are discussed herein.

Fragment-based approximations, anchored in the many-body expansion (MBE), are an appealing means to sidestep the steep nonlinear scaling of *ab initio* quantum chemistry.[1] There is also growing interest in using the MBE to systematically decompose interaction energies in molecular liquids,[2–5] biomolecules,[6–8] and other complex systems,[9,10] as a means to generate training data for machine learning applications. A growing body of work, however, demonstrates that care must be taken to avoid error accumulation in MBE-based methods.[11–16]

The enormity of the resource requirements for data-hungry machine learning applications places a premium on low-cost electronic structure models such as density functional theory (DFT), yet we have recently demonstrated that self-interaction error (SIE) causes catastrophic failure of the DFT-based MBE.[16] In the present work, we illustrate how SIE is intermingled with quadrature grid errors, in a manner that is unique to DFT calculations based on the MBE and does not manifest in conventional electronic structure calculations. The functionals most strongly affected are meta-generalized gradient approximations (meta-GGAs), which represent many of the most accurate contemporary exchange-correlation functionals.[17] Using calculations on $F^-(H_2O)_{15}$ clusters, we demonstrate herein that grid-based error dominates higher-order MBE calculations that employ meta-GGAs. This obfuscates other sources of error, such as SIE, which must be addressed for accurate DFT-MBE calculations. Improving the grid quality brings SIE-based delocalization error to the fore.

The MBE may be written as

$$E = \sum_{I=1}^{N} \left[E_I + \sum_{J>I} \left(\Delta E_{IJ} + \sum_{K>J} \Delta E_{IJK} + \cdots \right) \right] \quad (1)$$

where the E_I are single-fragment (monomer) energies,

$$\Delta E_{IJ} = E_{IJ} - E_I - E_J \quad (2)$$

is a two-body correction, etc.[1] If eq. 1 is truncated at n-body interactions, then we call the resulting method MBE(n). In conjunction with high-quality basis sets and correlated wave function models, MBE(4) affords good accuracy for neat liquid water and monovalent ion–water interactions,[14,18–20] yet requires electronic structure calculations on systems no larger than $(H_2O)_4$ or $X^{\pm}(H_2O)_3$. However, four-body calculations manifest a crippling $\mathcal{O}(N^4)$ combinatorial prefactor, resulting in error accumulation for large systems.[12–15]

Replacing wave function methods with DFT reduces the cost but SIE becomes catastrophic, with fluctuations as large as ±200 kcal/mol in low-order MBE(n) calculations in clusters such as $F^-(H_2O)_{15}$.[16] This effect is only marginally reduced by a combination of standard hybrid functionals (such as PBE0 or B3LYP) and/or aggressive energy screening.[16] For meta-GGA functionals including SCAN[21] and ωB97X-V,[22] these error mitigation strategies are only moderately effective. Because meta-GGAs are known to have more stringent grid requirements,[23–25] we decided to revisit the quadrature grids used in the DFT-based MBE, even though previous MBE(4) calculations at the B3LYP/cc-pVDZ level suggested that the quality of the integration grid had a negligible effect on accuracy, in applications to $(H_2O)_{40}$.[13]

For water or ion–water clusters, MBE(n) should converge to same interaction energy (ΔE_{int}) as a supramolecular calculation at the same level of theory, as n increases. (We use single-monomer fragments so that no covalent bonds are severed.) As such, it makes sense to define the error in the MBE(n) approximation as

$$\text{error} = E_{\text{MBE}(n)} - E_{\text{supersystem}} \ , \quad (3)$$

where both calculations are performed using the same density functional and basis set. As a control experi-

*herbert@chemistry.ohio-state.edu

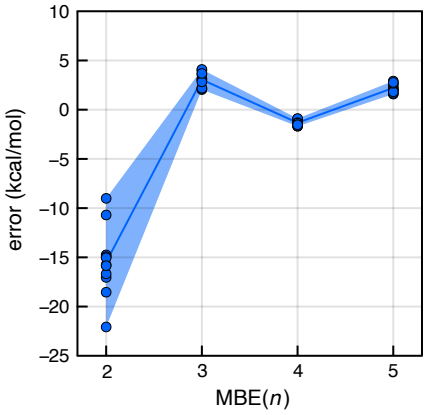

Figure 1: MBE(n) errors (as defined in eq. 3) for 11 different configurations of $F^-(H_2O)_{15}$ computed at the HF/aug-cc-pVDZ level. The solid line connects mean errors at each value of n and the shaded region highlights the span of the data.

ment, we demonstrate in Fig. 1 that convergence is indeed achieved for a set of $F^-(H_2O)_{15}$ clusters, using calculations at the Hartree-Fock (HF)/aug-cc-pVDZ level of theory. Residual errors spanning ≈ 2.5 kcal/mol, observed at the five-body level in Fig. 1, have elsewhere been shown to be artifacts of basis-set superposition error (BSSE).[18–20,26] For $Cl^-(H_2O)_9$, five body terms computed at the level of second-order Møller-Plesset perturbation theory (MP2)/aug-cc-pV5Z are negligibly small.[19]

HF-based MBE(n) calculations, for which there is no grid and no SIE, will serve as a baseline in the remaining discussion. For example, the analogous PBE/aug-cc-pVDZ data for $(H_2O)_{15}$ are shown in Fig. 2, superimposed on the span of the corresponding HF-MBE(n) errors. As demonstrated in previous work,[16] the PBE-MBE(n) errors diverge as n increases, with errors approaching ~ 150 kcal/mol at the five-body level.

In order to analyze the PBE-MBE(n) data in Fig. 2, we first review some nomenclature for DFT quadrature grids. Euler-Maclaurin-Lebedev (EML) quadrature grids have been used since the earliest days of molecular DFT,[27] and these indicated in Fig. 2 using the notation EML(N_r, N_Ω) where N_r is the number of radial shells on interval $[0, \infty)$, using an Euler-Maclaurin quadrature, and N_Ω is the number of angular points per shell (Lebedev quadrature). The choices (N_r, N_Ω) = (50,194), (75,302), and (99,590) that are examined in Fig. 2 are standard choices for a low-, medium-, and, high-quality grid, respectively.[23] The "standard grids" (SG-k) are pruned grids that reduce the number of Lebedev grid points near the nuclei (where the density is nearly spherically symmetric) and far away from the nuclei (where the density is slowly varying).[23,28] The SG-1 grid,[28] for example, starts from $N_r = 50$ and $N_\Omega = 194$ and applies a pruning procedure; this grid is typically adequate for GGA functionals.[17,23] Importantly, SG-1 uses a Gaussian quadrature,[28] and is

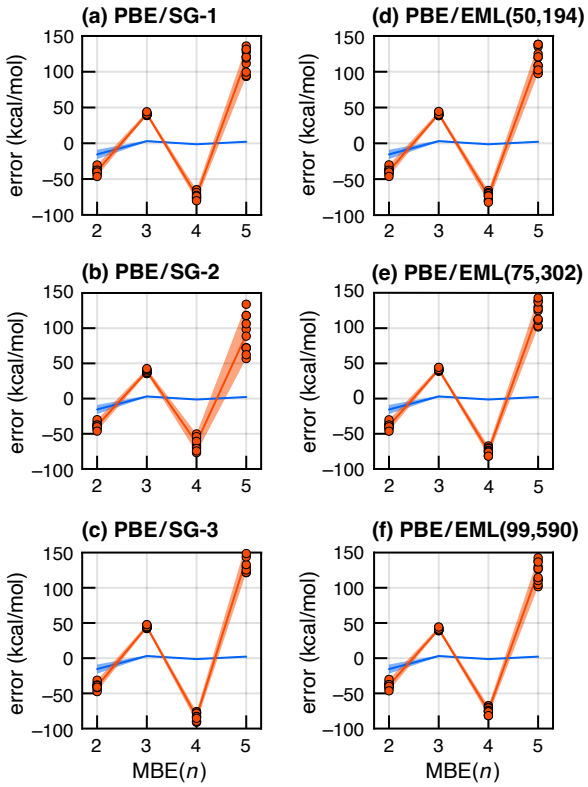

Figure 2: MBE(n) errors (circles) for 11 configurations of $F^-(H_2O)_{15}$, computed at the PBE/aug-cc-pVDZ level using various quadrature grids: (a) SG-1, (b) SG-2, (c) SG-3, (d) EML(50,194), (e) EML(75,302), (f) EML(99,590). The orange shaded region and solid line connect the range of the MBE(n) errors and their mean, for each value of n. In blue are the mean errors and their range, computed for the same set of clusters at the HF/aug-cc-pVDZ level. (The latter are the same data as in Fig. 1 but without the data points themselves.)

not a pruned version of EML(50,194). Similarly, SG-2 and SG-3 are pruned versions of (75,302) and (99,590) grids but both use a double-exponential quadrature,[29,30] so are not simply pruned versions of EML(75,302) and EML(99,590).[23] The SG-2 grid is adequate for many meta-GGA functionals,[23] whereas EML(75,302) had been recommended for meta-GGAs,[17] prior to the development of SG-2 and SG-3.

Regardless of quadrature grid, PBE-MBE(n) calculations for $F^-(H_2O)_{15}$ exhibit oscillations > 200 kcal/mol between $n = 4$ and $n = 5$. As such, we conclude that these oscillations are driven by SIE, as reported previously.[16]

Meta-GGA functionals depend on the Laplacian of the electron density and/or the kinetic energy density, both of which are more oscillatory than the density gradients, and thus more challenging for numerical integration. Oscillations in potential energy surfaces for noncovalent dimers have been noted for low-quality integration grids.[23,31,32] The SCAN functional[21] is especially sen-

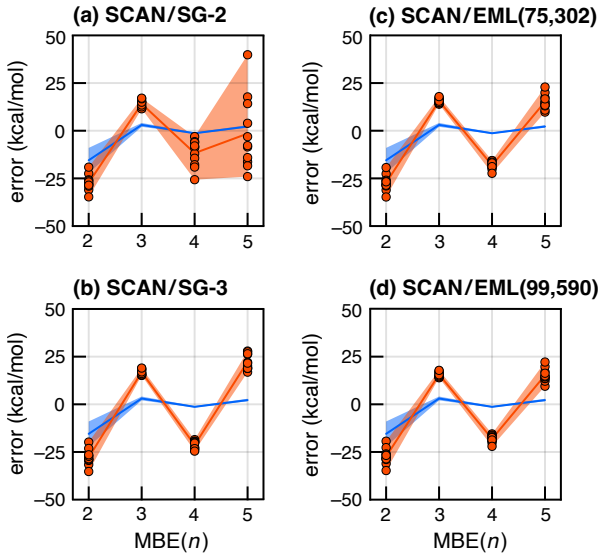

Figure 3: MBE(n) errors (circles) for 11 configurations of $F^-(H_2O)_{15}$, computed at the SCAN/aug-cc-pVDZ level using various quadrature grids: (a) SG-1, (b) SG-2, (c) SG-3, (d) EML(50,194), (e) EML(75,302), (f) EML(99,590). The orange shaded region and solid line connect the range of the errors and their mean, for each value of n. Quantities in blue indicate the HF/aug-cc-pVDZ error distribution from Fig. 1.

sitive to the choice of quadrature grid,[24,25] with direct consequences for SCAN-MBE(n) calculations that are on display for $F^-(H_2O)_{15}$ clusters in Fig. 3. Although the SG-2 grid works well for many meta-GGA functionals,[23] including SCAN in some instances,[33] in the present context it affords MBE(n) errors that fluctuate over a range of 75 kcal/mol at the $n = 5$ level.

Error distributions for SCAN-MBE(n) calculations are qualitatively different than those computed using PBE-MBE(n), for the same set of $F^-(H_2O)_{15}$ clusters, even when the basis set and quadrature grid are the same. In particular, the PBE-MBE(n) errors have strictly alternating signs as a function of n, regardless of the choice of quadrature grid (Fig. 2), which is not true for SCAN calculations using SG-2 (Fig. 3a). The difference lies in the fact that the PBE errors are relatively tightly clustered across the ensemble of structures that we consider, whereas SCAN errors for the same ensemble are more spread out at a given n, to the point where the errors span both sides of zero for $n = 5$. The SCAN errors become much more tightly clustered, and alternating signs are recovered for SCAN-MBE(n) calculations, when the SG-2 grid is replaced by the EML(75,302); see Fig. 3b. This behavior is not unique to the SCAN functional and is seen also for ωB97X-V; see Fig. S1.

For both SCAN and ωB97X-V, tight clustering and alternating signs for the errors persist in higher-quality grids including SG-3 and EML(99,590). In contrast, when SG-2 is used errors at the MBE(5) level appear to be almost random for different cluster geometries, suggesting MBE(n) results computed using meta-GGA functionals are dominated by grid error when SG-2 is used. Notably, grid-dependent of this magnitude are not observed in conventional supramolecular calculations. For example, noncovalent interaction energies (ΔE_{int}) for the S66 data set,[34] computed at the ωB97X-V/aug-cc-pVTZ level, differ by an average of only 0.03 kcal/mol for SG-2 as compared to its unpruned ($N_r = 75, N_\Omega = 302$) counterpart.[23] Thus, SG-2 may be considered effectively converged to the grid limit for conventional, monolithic DFT calculations but the same grid is insufficient for DFT-MBE(n), where $\sim 10^4$ subsystem calculations (or more) might be required to evaluate a single-point energy.

These noncovalent benchmarks help to explain why the error patterns for PBE (Fig. 2) and SCAN (Fig. 3) are qualitatively different. SIE drives delocalization that stabilizes clusters relative to their constituent monomers, and stabilizes larger clusters relative to smaller ones. The *lack* of SIE in the smaller fragments is overrepresented in the MBE(n) as system size and expansion order increases, which is more readily seen if the n-body corrections are written in closed form,[13]

$$E_{\text{MBE}(n)} = \sum_{k=1}^{n} (-1)^{n-k} \binom{N-k-1}{n-k} \sum_{\alpha=1}^{\binom{N}{k}} E_\alpha^{(k)} \qquad (4)$$

where α indexes the subsystems consisting of k fragments, whose individual energies are $E_\alpha^{(k)}$. In this form, each fragment energy is scaled by a signed combinatoric coefficient. These coefficients are listed in Table S1 for $N = 16$, as in $F^-(H_2O)_{15}$.

For $k = n$, the coefficient in eq. 4 is unity regardless of N. For $k < n$, the signs alternate with n for fixed k and increase in magnitude with n (see Table S1). Because the SIE inherent in $E_\alpha^{(k)}$ is consistently stabilizing, the alternating signs of the lower order terms will overwhelm the always-positive $n = k$ terms, leading to the pattern of errors that is observed in the PBE-MBE(n) calculations. Note that these errors need not result in a divergent MBE(n), as demonstrated by the HF/aug-cc-pVDZ data, where BSSE is strictly stabilizing for larger subsystems, yet the errors grow smaller as n increases. This suggests that there exists some expansion-depended error threshold, related to the combinatorial coefficients in eq. 4, beyond which MBE(n) calculations will diverge.

In contrast, grid-based errors need not be strictly stabilizing so the error pattern is more scattered in cases where the errors are dominated by grid artifacts rather than SIE. This is the case for SCAN-based MBE(5) calculations using SG-2 (Fig. 3a). Only upon saturating the grid does SIE take over as the dominant source of error, leading to oscillatory results for MBE(n) with meta-GGA functionals, as in Fig. 3b. However, fluctuations of almost 50 kcal/mol persist between MBE(4) and MBE(5) results, even when the high-quality EML(99,590) grid is used.

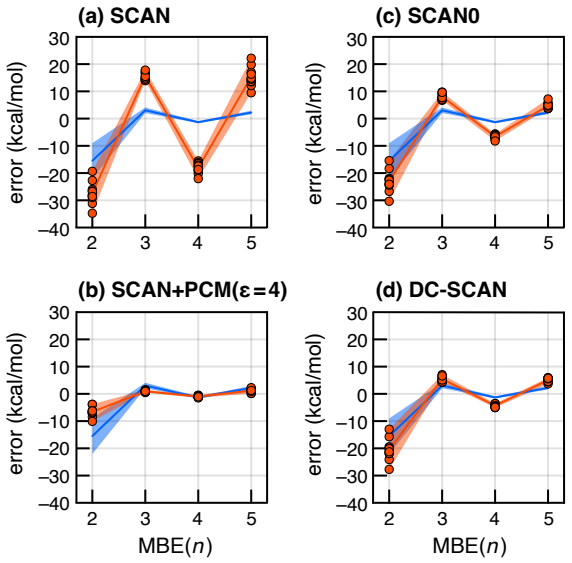

Figure 4: MBE(n) errors for 11 different $F^-(H_2O)_{15}$ clusters, computed using (a) SCAN, (b) SCAN + PCM($\varepsilon = 4$), and (c) SCAN0 with 25% exact exchange, and (d) DC-SCAN, starting from a converged HF density. All calculations used the aug-cc-pVDZ basis set and an unpruned EML(99,590) grid. Orange lines connect the mean errors at each value of n and the orange shaded region highlights the span of the data. The blue line and shaded region represent the HF/aug-cc-pVDZ data from Fig. 1.

In previous work,[16] we tested several strategies to mitigate accumulation of SIE, including hybrid functionals, dielectric boundary conditions, and density-corrected (DC-)DFT. Although these procedures *reduced* the errors in DFT-MBE(n) calculations, none was sufficient to restore convergence, at least not for standard hybrid functionals with 20–25% exact exchange. (Functionals with 50% exact exchange did afford convergent results, indicating that SIE plays a prominent role in divergent behavior for semilocal functionals and for hybrid functionals such as B3LYP, PBE0, and SCAN0.) Because the calculations in Ref. 16 used the SG-2 quadrature grid for meta-GGA functionals, SIE-driven errors were intertwined with MBE-induced grid errors. We next revisit these calculations using the unpruned EML(99,590) grid.

MBE(n) errors for $F^-(H_2O)_{15}$ clusters are plotted in Fig. 4 for several SCAN-based calculations, all using the EML(99,590) grid. For reference, the SCAN-MBE(n) results in Fig. 4a are the same as those in Fig. 3d, and they are compared in Fig. 4 alongside MBE(n) calculations using SCAN + PCM, SCAN0, and DC-SCAN, all with the same high-quality grid.

In MBE(n) calculations on proteins,[35–37] we demonstrated that low-dielectric boundary conditions are necessary to obtain converged results in the presence of ionic residues.[35] Dielectric boundaries can be implemented by means of a polarizable continuum model (PCM),[38] using a dielectric constant $\varepsilon = 4$ as in previous work.[35–37] Results in Fig. 4b show that MBE(n) errors are significantly reduced for SCAN + PCM($\varepsilon = 4$). For $n \geq 3$, the errors are comparable to (and in some cases smaller than) the corresponding HF-MBE(n) errors.

Alternatively, adding 25% exact exchange (to obtain the SCAN0 functional[39]) results in convergent behavior for MBE(n) as shown in Fig. 4c. This was not the case when SG-2 was used.[16] Instead, SCAN0-MBE(5) errors for this same set of $F^-(H_2O)_{15}$ clusters, computed using the SG-2 grid, span a range of almost 250 kcal/mol and lack the telltale alternating signs of SIE-dominated errors.[16] Nevertheless, oscillations in SCAN0-MBE(n) errors obtain using the EML(99,590) grid remain larger than the corresponding HF-MBE(n) values, as shown in Fig. 4c. We attribute this to residual delocalization error in SCAN0, since the smaller HF-MBE(5) errors should provide an estimate of MBE(n) errors due to finite-basis effects (specifically, mismatch between subsystem and supersystem BSSE),[14] and the SCAN0-MBE(5) errors are larger than any such effects.

In the DC-DFT approach,[40] also known as "DFT@HF",[41] an exchange-correlation functional is evaluated in a one-shot, non-self-consistent fashion using a converged HF density. This approach can significantly mitigate SIE-driven errors,[40–44] and DC-SCAN has been suggested as a workhorse for development of MBE-based force fields.[45,46] Results in Fig. 4d show errors on the order of 5 kcal/mol for MBE(4) and MBE(5) calculations based on DC-SCAN, which is a bit larger than the corresponding HF-MBE(n) errors but slightly smaller than SCAN0-MBE(n) errors. Residual five-body errors using DC-SCAN are more surprising than the corresponding SCAN0 errors, since DC-SCAN uses a SIE-free density and the errors cannot be blamed on delocalization error. Enhanced fluctuations MBE(n) errors using DC-SCAN may arise from electron correlation effects and a more appropriate comparison for DC-DFT might be MP2, which is SIE-free but includes correlation. Indeed, MP2-MBE(n) fluctuations for $F^-(H_2O)_{15}$ clusters are larger than those observed for HF-MBE(n) calculations.[20]

Finally, we comment on how denser quadrature grids impacts computational cost. The cost to evaluate the exchange-correlation quadrature is proportional to the number of grid points and ultimately $\mathcal{O}(N)$ with molecular size, although it can be a significant fraction of the computational time for semilocal functionals.[23] The cost of SCAN-MBE(4) calculations on $F^-(H_2O)_{15}$ is listed in Table 1 for various quadrature grids. The SG-2 grid is deficient for such calculations but is included as a baseline, since this is a typical grid used for meta-GGAs in conventional quantum chemistry.[23] SG-3 is satisfactory for MBE(n) calculations, and its cost is functionally equivalent to that of EML(75,302), and both exhibit similar errors. Further testing is required to fully differentiate the performance of these grids for DFT-MBE(n) applications, but the present work serves as a cautionary note that supramolecular tests may not reveal problems en-

Table 1: Cost of MBE(4) Calculations on $F^-(H_2O)_{15}$ at the SCAN/aug-cc-pVDZ Level.[a]

Grid	Time (sec)[b]
SG-2	59,176 ± 660
SG-3	96,282 ± 1205
EML(75,302)	99,744 ± 1450
EML(99,590)	202,371 ± 2282

[a]Averaged over 11 cluster geometries. [b]Uncertainty represents one standard deviation.

gendered by MBE(n) calculations. This is analogous to the manner in which MBE(n) is more sensitive to numerical thresholds as compared to supersystem calculations at the same level of theory.[12–14]

In summary, while meta-GGA functionals have distinct advantages compared to functionals situated at lower rungs on Jacob's ladder, their more stringent quadrature grid requirements are significantly amplified when such functionals are used in conjunction with the MBE. In particular, the SG-2 grid (which is otherwise recommended for meta-GGA functionals such as SCAN and ωB97X-V),[23] results in cumulative errors that can mask other artifacts, including SIE. To eliminate the grid errors, we recommend the SG-3 grid or (in case of doubt) an unpruned grid such as EML(75,302).

This does not eliminate SIE-driven delocalization errors, and MBE(n) calculations with meta-GGA functionals remain divergent even if the order-by-order fluctuations are greatly reduced through the use of high-quality integration grids. However, by saturating the grid and removing it as a source of error, other strategies for mitigating SIE can help, which were unsuccessful in calculations based on SG-2.[16] These include the use of hybrid functionals such as SCAN0, low-dielectric boundary conditions (based on a PCM with $\varepsilon = 4$), and DC-DFT (DFT@HF), based on a self-consistent HF density.

Methods

Cluster structures were taken from Ref. 16 where they were obtained from a molecular dynamics simulation of F^- in bulk water. All calculations were performed using the FRAGME∩T code,[47] interfaced to Q-CHEM v. 6.2.[48] Timing data were obtained using compute nodes with two Intel Xeon CPU Max 9470 processors (26 cores each) with 128 Gb of memory. Each fragment calculation was provisioned to use 4 cores and 10 Gb of memory.

Supporting Information

Additional calculations and data (PDF)
Coordinates for $F^-(H_2O)_{15}$ clusters (TXT)

Conflicts of interest

J.M.H. is part owner of Q-Chem Inc. and serves on its board of directors.

Acknowledgements

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under Award No. DE-SC0008550. Calculations were performed at the Ohio Supercomputer Center.[49]

References

[1] Herbert, J. M. Fantasy versus reality in fragment-based quantum chemistry. *J. Chem. Phys.* **2019,** *151,* 170901.
[2] Riera, M.; Knight, C.; Bull-Vulpe, E. F.; Zhu, X.; Agnew, H.; Smith, D. G. A.; Simmonett, A. C.; Paesani, F. MBX: A many-body energy and force calculator for data-driven many-body simulations. *J. Chem. Phys.* **2023,** *159,* 054802.
[3] Yao, K.; Herr, J. E.; Parkhill, J. The many-body expansion combined with neural networks. *J. Chem. Phys.* **2017,** *146,* 014106.
[4] Bull-Vulpe, E. F.; Riera, M.; Götz, A. W.; Paesani, F. MB-Fit: Software infrastructure for data-driven many-body potential energy functions. *J. Chem. Phys.* **2021,** *155,* 124801.
[5] Maldonado, A. M.; Poltavsky, I.; Vassilev-Galindo, V.; Tkatchenko, A.; Keith, J. A. Modeling molecular ensembles with gradient-domain machine learning force fields. *Digital Discov.* **2023,** *2,* 871–880.
[6] Li, W.; Ma, H.; Li, S.; Ma, J. Computational and data driven molecular material design assisted by low scaling quantum mechanics calculations and machine learning. *Chem. Sci.* **2021,** *12,* 14987–15006.
[7] Cheng, Z.; Du, J.; Zhang, L.; Ma, J.; Li, W.; Li, S. Building quantum mechanics quality force fields of proteins with the generalized energy-based fragmentation approach and machine learning. *Phys. Chem. Chem. Phys.* **2021,** *24,* 1326–1337.
[8] Liao, K.; Dong, S.; Cheng, Z.; Li, W.; Li, S. Combined fragment-based machine learning force field with classical force field and its application in the NMR calculations of macromolecules in solutions. *Phys. Chem. Chem. Phys.* **2022,** *24,* 18559–18567.
[9] Ricard, T. C.; Zhu, X.; Iyengar, S. S. Capturing weak interactions in surface adsorbate systems at coupled cluster accuracy: A graph-theoretic molecular fragmentation aapproach improved through machine learning. *J. Chem. Theory Comput.* **2023,** *19,* 8541–8556.
[10] Iyengar, S. S.; Ricard, T. C.; Zhu, X. Reformulation of all ONIOM-type molecular fragmentation approaches and many-body theories using graph-theory-based projection operators: Applications to dynamics, molecular potential surfaces, machine learning, and quantum computing. *J. Phys. Chem. A* **2024,** *128,* 466–478.
[11] Ouyang, J. F.; Cvitkovic, M. W.; Bettens, R. P. A. Trouble with the many-body expansion. *J. Chem. Theory Comput.* **2014,** *10,* 3699–3707.
[12] Richard, R. M.; Lao, K. U.; Herbert, J. M. Aiming for benchmark accuracy with the many-body expansion. *Acc.*

Chem. Res. **2014,** *47,* 2828–2836.

[13] Richard, R. M.; Lao, K. U.; Herbert, J. M. Understanding the many-body expansion for large systems. I. Precision considerations. *J. Chem. Phys.* **2014,** *141,* 014108.

[14] Lao, K. U.; Liu, K.-Y.; Richard, R. M.; Herbert, J. M. Understanding the many-body expansion for large systems. II. Accuracy considerations. *J. Chem. Phys.* **2016,** *144,* 164105.

[15] Liu, K.-Y.; Herbert, J. M. Understanding the many-body expansion for large systems. III. Critical role of four-body terms, counterpoise corrections, and cutoffs. *J. Chem. Phys.* **2017,** *147,* 161729.

[16] Broderick, D. R.; Herbert, J. M. Delocalization error poisons the density-functional many-body expansion. *Chem. Sci.* **2024,** *15,* 19893–19906.

[17] 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.

[18] Heindel, J. P.; Xantheas, S. S. The many-body expansion for aqueous systems revisited: I. Water-water interactions. *J. Chem. Theory Comput.* **2020,** *16,* 6843–6855.

[19] Heindel, J. P.; Xantheas, S. S. The many-body expansion for aqueous systems revisited: II. Alkali metal and halide ion–water interactions. *J. Chem. Theory Comput.* **2021,** *17,* 2200–2216.

[20] Broderick, D. R.; Herbert, J. M. Scalable generalized screening for high-order terms in the many-body expansion: Algorithm, open-source implementation, and demonstration. *J. Chem. Phys.* **2023,** *159,* 174801.

[21] Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **2015,** *115,* 036402.

[22] 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,* 9904–9924.

[23] Dasgupta, S.; Herbert, J. M. Standard grids for high-precision integration of modern density functionals: SG-2 and SG-3. *J. Comput. Chem.* **2017,** *38,* 869–882.

[24] Bartók, A. P.; Yates, J. R. Regularized SCAN functional. *J. Chem. Phys.* **2019,** *150,* 161101.

[25] Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and numerically efficient r^{2}SCAN meta-generalized gradient approximation. *J. Phys. Chem. Lett.* **2020,** *11,* 8208–8215.

[26] Richard, R. M.; Lao, K. U.; Herbert, J. M. Approaching the complete-basis limit with a truncated many-body expansion. *J. Chem. Phys.* **2013,** *139,* 224102.

[27] Murray, C. W.; Handy, N. C.; Laming, G. J. Quadrature schemes for integrals of density functional theory. *Mol. Phys.* **1993,** *78,* 997–1014.

[28] Gill, P. M. W.; Johnson, B. G.; Pople, J. A. A standard grid for density-functional calculations. *Chem. Phys. Lett.* **1993,** *209,* 506–512.

[29] Mitani, M. An application of double exponential formula to radial quadrature grid in density functional calculation. *Theor. Chem. Acc.* **2011,** *130,* 645–669.

[30] Mitani, M.; Yoshioka, Y. Numerical integration of atomic electron density with double exponential formula for density functional calculation. *Theor. Chem. Acc.* **2012,** *131,* 1169.

[31] Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. Application of 25 density functionals to dispersion-bound homomolecular dimers. *Chem. Phys. Lett.* **2004,** *394,* 334–338.

[32] Johnson, E. R.; Becke, A. D.; Sherrill, C. D.; DiLabio, G. A. Oscillations in meta-generalized-gradient approximation potential energy surfaces for dispersion-bound complexes. *J. Chem. Phys.* **2009,** *131,* 034111.

[33] Jana, S.; Herbert, J. M. Slater transition methods for core-level electron binding energies. *J. Chem. Phys.* **2023,** *158,* 094111.

[34] Řezác, J.; Riley, K. E.; Hobza, P. S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures. *J. Chem. Theory Comput.* **2011,** *7,* 2427–2438.

[35] Bowling, P. E.; Broderick, D. R.; Herbert, J. M. Fragment-based calculations of enzymatic thermochemistry require dielectric boundary conditions. *J. Phys. Chem. Lett.* **2023,** *14,* 3826–3834.

[36] Bowling, P. E.; Broderick, D. R.; Herbert, J. M. Convergent protocols for protein–ligand interaction energies using fragment-based quantum chemistry. *J. Chem. Theory Comput.* **2025,** (in press; preprint available at DOI: 10.26434/chemrxiv-2024-7v7pv).

[37] Bowling, P. E.; Broderick, D. R.; Herbert, J. M. Quick-and-easy validation of protein–ligand binding models using fragment-based semi-empirical quantum chemistry. *ChemRxiv* **2024** (DOI: 10.26434/chemrxiv-2024-lcgb6).

[38] Herbert, J. M. Dielectric continuum methods for quantum chemistry. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021,** *11,* e1519.

[39] Hui, K.; Chai, J.-D. SCAN-based hybrid and double-hybrid density functionals from models without fitted parameters. *J. Chem. Phys.* **2016,** *144,* 044114.

[40] Sim, E.; Song, S.; Vuckovic, S.; Burke, K. Improving results by improving densities: Density-corrected density functional theory. *J. Am. Chem. Soc.* **2022,** *144,* 6625–6639.

[41] Kaplan, A. D.; Shahi, C.; Sah, R. K.; Bhetwal, P.; Kanungo, B.; Gavini, V.; Perdew, J. P. How does HF-DFT achieve chemical accuracy for water clusters? *J. Chem. Theory Comput.* **2024,** *20,* 5517–5527.

[42] Song, S.; Vuckovic, S.; Sim, E.; Burke, K. Density-corrected DFT explained: Questions and answers. *J. Chem. Theory Comput.* **2022,** *18,* 817–827.

[43] Rana, B.; Coons, M. P.; Herbert, J. M. Detection and correction of delocalization errors for electron and hole polarons using density-corrected DFT. *J. Phys. Chem. Lett.* **2022,** *13,* 5275–5284.

[44] Rana, B.; Beran, G. J. O.; Herbert, J. M. Correcting π-delocalisation errors in conformational energies using density-corrected DFT, with application to crystal polymorphs. *Mol. Phys.* **2023,** *121,* e2138789.

[45] Dasgupta, S.; Lambros, E.; Perdew, J. P.; Paesani, F. Elevating density functional theory to chemical accuracy for water simulations through a density-corrected many-body formalism. *Nat. Commun.* **2021,** *12,* 6359.

[46] Dasgupta, S.; Shahi, C.; Bhetwal, P.; Perdew, J. P.; Paesani, F. How good is the density-corrected SCAN functional for neutral and ionic aqueous systems, and what is so right about the Hartree–Fock density? *J. Chem. Theory Comput.* **2022,** *18,* 4745–4761.

[47] Broderick, D. R.; Bowling, P. E.; Shockey, J.; Higley, J.; Dickerson, H.; Ahmed, S.; Herbert, J. M. FRAGME∩T, a framework for fragment-based quantum chemistry calculations (https://gitlab.com/fragment-qc/fragment).

[48] Epifanovsky, E. *et al.* Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021,** *155*, 084801.
[49] Ohio Supercomputer Center, http://osc.edu/ark:/19495/f5s1ph73 (accessed 2024-12-15).

TOC Graphic