Density Functional Theory of Water with the Machine-Learned DM21 Functional

Etienne Palos,[1, a)] Eleftherios Lambros,[1, b)] Saswata Dasgupta,[1, c)] and Francesco Paesani[1, 2, 3, d)]

[1)] *Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, United States*

[2)] *Materials Science and Engineering, University of California San Diego, La Jolla, California 92093, United States*

[3)] *San Diego Supercomputer Center, University of California San Diego, La Jolla, California 92093, United States*

The delicate interplay between functional-driven and density-driven errors in density functional theory (DFT) has hindered traditional density functional approximations (DFAs) from providing an accurate description of water for over 30 years. Recently, the deep-learned DeepMind 21 (DM21) functional has been shown to overcome the limitations of traditional DFAs as it is free of delocalization error. To determine if DM21 can enable a molecular-level description of the physical properties of aqueous systems within Kohn-Sham DFT, we assess the accuracy of the DM21 functional for neutral, protonated, and deprotonated water clusters. We find that the ability of DM21 to accurately predict the energetics of aqueous clusters varies significantly with cluster size. Additionally, we introduce the many-body MB-DM21 potential derived from DM21 data within the many-body expansion of the energy and use it in simulations of liquid water as a function of temperature at ambient pressure. We find that size-dependent functional-driven errors identified in the analysis of the energetics of small clusters calculated with the DM21 functional result in the MB-DM21 potential systematically overestimating the hydrogen-bond strength and, consequently, predicting a more ice-like local structure of water at room temperature.

[a)]These authors contributed equally; Electronic mail: epalos@ucsd.edu

[b)]These authors contributed equally; Electronic mail: elambros@ucsd.edu

[c)]These authors contributed equally; Electronic mail: s1dasgupta@ucsd.edu

[d)]Electronic mail: fpaesani@ucsd.edu

Since the dawn of computer simulations, there has been a continuous effort in the physical sciences to accomplish a unified, molecular-level understanding of the properties of water and aqueous systems, from small gas-phase clusters to the thermodynamic limit.[1,2] Density functional theory (DFT) provides, in principle, an exact treatment of the electronic ground-state potential energy surface of a given molecular system by solving the Kohn-Sham equations. Although more than a half-century has passed since the inception of DFT,[3,4] the exact density functional remains unknown, and so the quest for a density functional approximation (DFA) capable of accurately predicting the properties of water and aqueous systems across the entire phase diagram continues.[5–28]

One of the main obstacles that prevents DFT from achieving full reliability in molecular simulations is the error introduced by the density produced by a given DFA, *i.e* the delocalization error,[29,30] which results from the deviation of the approximate functional from the piecewise-linear behavior of the exact Kohn-Sham functional as a function of fractional electron number.[29] Recently, the strongly constrained and appropriately normed (SCAN) functional,[31] which satisfies the 17 known exact constraints imposed on a meta-GGA functional, has been the subject of numerous studies focused on modeling water, ranging from systematic analyses of the energetics of gas-phase clusters to molecular dynamics (MD) simulations of liquid water carried out *ab initio* or using SCAN-based machine-learned and data-driven potentials.[28,32–36]. Despite satisfying the 17 exact constraints known for a meta-GGA functional, the SCAN functional still suffers from density-driven errors that limit its predictive ability in MD simulations of water.[33,37] Recently, the limitations of SCAN resulting from density-driven errors have been overcome by replacing the self-consistent SCAN density with the Hartree-Fock density,[37] which is free of self-interaction error.[38,39] The development of density-corrected SCAN (DC-SCAN) and associated data-driven many-body MB-SCAN(DC) potential energy function (PEF) effectively provided the first DFT-based description of the properties of water across its phase diagram at the accuracy of "gold standard" coupled cluster theory.[37,40]

Alternatively, the deep-learned local hybrid DM21 functional recently developed by DeepMind has been shown to improve the accuracy of the DFT ansatz by reducing the delocalization error through the learning of fractional charge and spin constraints, along with the training of molecular data.[41] Unlike previous hybrid functionals, the fraction of Hartree-Fock exchange in DM21 is dynamic, meaning that it varies in space according to local conditions,

virtually removing the delocalization error. The exchange-correlation energy of DM21 is given by

$$E_{\mathrm{xc}}^{\mathrm{DM21}}[\rho(\mathbf{r})] = E_{\mathrm{xc}}^{\mathrm{MLP}}[\rho(\mathbf{r})] + E_{\mathrm{D3(BJ)}} \tag{1}$$

where $E_{\mathrm{xc}}^{\mathrm{MLP}}[\rho(\mathbf{r})]$ is a machine-learned potential trained on extensive sets of atomic and molecular data, exact (integer-charge) mathematical constraints, as well as fractional-charge and fractional-spin constraints. In Eq. 1, $E_{\mathrm{D3(BJ)}}$ is the two-body dispersion-correction to the total energy expresses in terms of the D3(BJ) parameters optimized for the B3LYP functional.[41] The DM21 functional is virtually free of density-driven error and, consequently, reproduces the piecewise-linear behavior of the exact Kohn-Sham functional for a system with fractional charge electrons.[41] Notably, DM21 outperforms traditional DFAs against the GMTKN55 database, which includes datasets in main group thermochemistry, kinetics, and noncovalent interactions.[41] Recent work also suggests that DM21 is robust against the "poison" benchmark sets that comprised systems that are notoriously difficult to model.[42] While DM21 is promising for several applications in chemistry, its reliability and predictive power in the condensed phase have not yet been investigated. In this Communication, we build upon recent efforts aimed at developing a robust DFT-based framework for MD simulations of water[33,37,40,43,44] to investigate the accuracy of the DM21 functional for water and aqueous ions.

To assess the accuracy of DM21 in representing molecular interactions for neutral, protonated, and deprotonated water, we calculate the interaction and binding energies of various gas-phase clusters. The neutral water clusters are obtained from the water subset of the BEGDB dataset,[46] whereas the protonated and deprotonated clusters are taken from the WATER27 dataset.[47] The neutral water clusters of the BEGDB dataset represent the low-energy structures of $(H_2O)_n$ clusters with $n = 2 - 10$, which were optimized at the RI-MP2/aug-cc-pVDZ level of theory.[48] The protonated and deprotonated clusters correspond to $H_3O^+(H_2O)_n$ and $OH^-(H_2O)_n$ clusters with $n = 2 - 6$, which were originally optimized at the B3LYP/6-311++G(2d,2p) level of theory.[47] The accuracy of the DM21 density functional for these aqueous clusters is compared with that of the ωB97M-V and DC-SCAN functionals, both of which have been shown to provide high accuracy relative to coupled cluster calculations carried out including single, double, and perturbative triple excitations, i.e., CCSD(T), in the complete basis set (CBS) limit.[37,49] In the following analyses, the interaction energy of a cluster containing n monomers (i.e., H_2O, H_3O^+, and OH^- molecules)

is calculated as the difference between the total energy of the cluster and the sum of the energies of the n monomers in the same distorted geometries as in the cluster. The corresponding binding energies are calculated taking the monomer relaxation into account as the difference between the total energy of the cluster and the sum of the energies of the n monomers in their optimized geometries.

Fig. 1a shows the absolute errors in the interaction energies calculated with DM21, DC-SCAN, and ωB97M-V for the neutral water clusters relative to the corresponding CCSD(T) reference values.[45] Due to the minimization of density-driven errors, all three density functionals provide significantly smaller error in interaction energy compared to

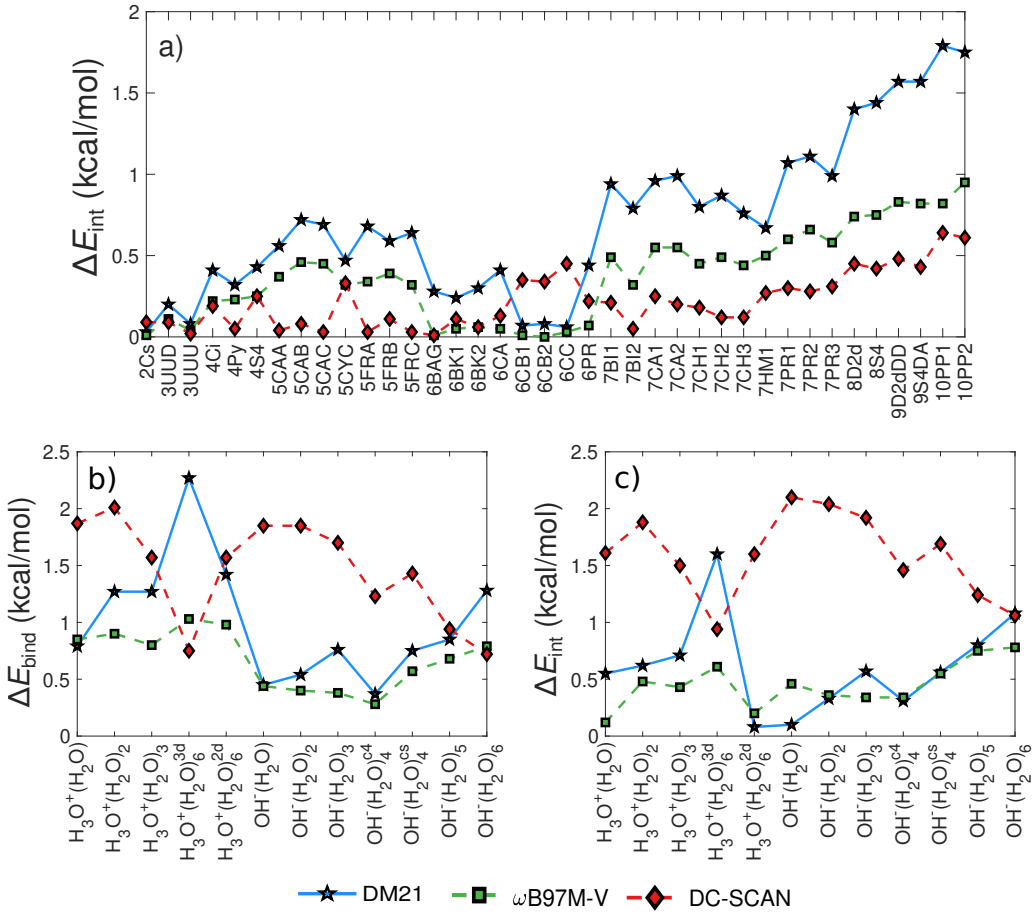

FIG. 1. Absolute errors (in kcal/mol) relative to the CCSD(T) reference values[45] for a) the interaction energies of the $(H_2O)_{n=2-10}$ clusters of the BEGDB dataset,[46] b) the binding energies of the $H_3O^+(H_2O)_n$ and $OH^-(H_2O)_n$ clusters of the WATER27 dataset, and c) the interaction energies of the $H_3O^+(H_2O)_n$ and $OH^-(H_2O)_n$ clusters of the WATER27 dataset.

popular GGA and meta-GGA density functionals analyzed in Ref. 40. As shown in previous studies,[37,40] DC-SCAN describes the energetics of neutral water clusters with similar accuracy as CCSD(T)/CBS. In the case of DM21, the error per monomer varies from -0.02 to -0.18 kcal/mol with the cluster size (Fig. S1). Similar trend is found for ωB97M-V, although the increase in the error with cluster size is a less pronounced, resulting in error per monomer that varies from 0.01 to -0.10 kcal/mol. Moving from smaller to larger clusters does not significantly affect the error per monomer for DC-SCAN, which is found to lie within 0.06 kcal/mol of the CCSD(T)/CBS reference values from the dimer to decamer. Fig. S2 shows the mean absolute error per monomer for the interaction energies of the neutral water clusters. Minimization of the density-driven errors leads to accurate interaction energies as demonstrated by the smaller mean absolute errors per monomer associated with DM21, ωB97M-V, and DC-SCAN, with the latter providing the closest agreement with the CCSD(T)/CBS reference values. For comparison, Fig. S2 also shows that, due to large density-driven errors displayed by semi-local functionals, the mean absolute error per monomer associated with SCAN calculations increases rapidly with the cluster size.

Figs. 1b and 1c show the errors in binding and interaction energies relative to the CCSD(T) reference values[45] for protonated and deprotonated water clusters, respectively. In this case, both ωB97M-V and DM21 are in closer agreement with the reference values than DC-SCAN. Interestingly, the errors appears to systematically increase as the cluster size increases for both DM21 and ωB97M-V. An opposite trend is observed for DC-SCAN, with the errors becoming smaller as n increases. It should be noted that increasing the number of water molecules in an ion–water cluster results in a smaller charge-per-monomer, which implies that the properties of ion–water clusters effectively tend to those of neutral water clusters as n increases. Since the (small) errors associated with DC-SCAN calculations of neutral water clusters are effectively independent of the cluster size, the interaction energies of ion–water clusters predicted by DC-SCAN become relatively more accurate as the clusters become larger. On the contrary, the errors associated with DM21 and ωB97M-V calculations of neutral water clusters are dependent on the cluster size which explain the increase in the errors for both interaction and binding energies predicted by these two functionals for larger protonated and deprotonated water clusters. Fig. S3 shows that DM21, ωB97M-V, and DC-SCAN are significantly more accurate than the semi-local SCAN functional for both protonated and deprotonated water clusters, which is primarily due to the removal of

density-driven errors that affect SCAN.[37]

To gain further insight into how molecular interactions in aqueous systems are represented by the DM21 functional, we analyze individual n-body contributions to the interaction energies of small clusters. Many-body decomposition analyses have been shown to provide a quantitative assessment of the overall accuracy of different molecular models of aqueous systems.[50–52] Figs. 2a-c show the errors ($\Delta\varepsilon_{\mathrm{nB}}$) associated with each n-body contribution to the interaction energy of the lowest-energy isomer (Fig. S4) of the neutral, $(H_2O)_6$,[53] protonated, $H_3O^+(H_2O)_5$,[52] and deprotonated, $OH^-(H_2O)_5$,[51] water hexamer clusters calculated with DM21, ωB97M-V, and DC-SCAN. DM21 predicts the 2-body (2B) energy of the neutral water hexamer with CCSD(T) accuracy, while the 3-body (3B) error is appreciably larger ($\sim$ 0.83 kcal/mol) and similar to that associated with ωB97M-V. Analyses of interaction and many-body energies of the first eight low-energy isomers of the neutral water hexamer shown in Figs. S5-S7 indicate that the 3B error associated with DM21 reduces for the planar isomers.

DM21 also provides 2B energies with CCSD(T) accuracy for the protonated water cluster (Fig. 2b), outperforming both ωB97M-V and DC-SCAN, which instead display errors of $\sim$0.6 kcal/mol and $\sim$2.0 kcal/mol, respectively. Similar 3B errors are instead found for all three functionals, with DM21 and DC-SCAN displaying the largest ($\sim$0.8 kcal/mol) and smallest ($\sim$0.5 kcal/mol) errors, respectively. As for the neutral water hexamer, DM21 predicts relatively smaller 3B errors for the planar isomers of $H_3O^+(H_2O)_5$ (Figs. S8a and S8b).

All three functionals display similar trends for $\Delta\varepsilon_{\mathrm{nB}}$ calculated for the deprotonated water

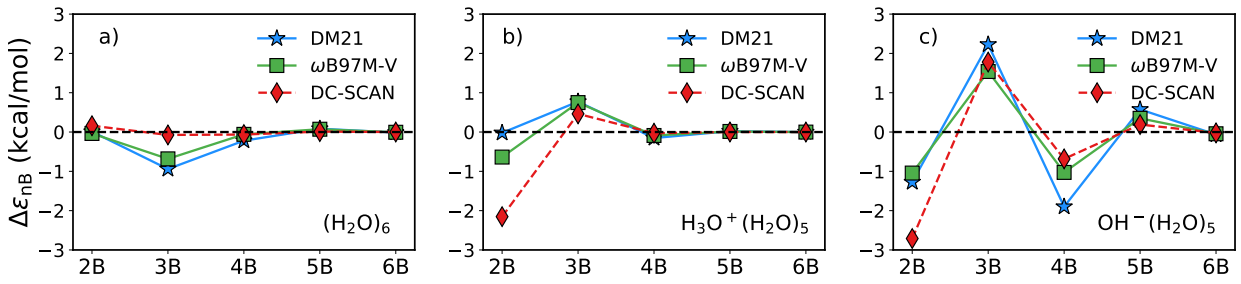

FIG. 2. Errors (in kcal/mol) associated with individual n-body contributions to the interaction energies of the lowest-energy isomers (Fig. S4) of the $(H_2O)_6$ (panel a), $H_3O^+(H_2O)_5$ (panel b), and $OH^-(H_2O)_5$ (panel c) clusters relative to the corresponding CCSD(T) reference values.

cluster (Fig. 2c), with DM21 providing the smallest 2B error but relatively larger errors than ωB97 and DC-SCAN for the 3B, 4B, and 5B energies. The more pronounced "oscillatory" trend of $\Delta\varepsilon_{\mathrm{nB}}$ associated with DM21, which suggests relatively larger error compensation among individual n-body contributions, becomes even more evident for deprotonated water clusters with different geometries and sizes (Figs. S8b and S8c). These comparisons indicate that all three functionals are less accurate for the deprotonated clusters than the corresponding neutral and protonated clusters. This suggests that all three functionals overestimate, to a different extent, the energy of the highest occupied molecular orbital (HOMO) of the deprotonated water clusters, which results from incorrect features of the exchange-correlation functional in the asymptotic region.[51,54]

To asses the accuracy of DM21 for liquid water, we introduce MB-DM21, a data-driven many-body PEF derived from the many-body expansion (MBE) of the total energy. Following Ref. 55, MB-DM21 approximates the MBE to the sum of explicit 1-body (1B), 2-body (2B), and 3-body (3B) terms, along with an implicit many-body term based on classical polarization (V_{pol}) representing all n-body interactions with $n > 3$:

$$E_N^{\mathrm{MB-DM21}}(\mathbf{r}_1, .., \mathbf{r}_N) = \sum_{i=1}^{N} \varepsilon_{1\mathrm{B}}(\mathbf{r}_i) + \sum_{i>j}^{N} \varepsilon_{2\mathrm{B}}(\mathbf{r}_i, \mathbf{r}_j) + \\ + \sum_{i>j>k}^{N} \varepsilon_{3\mathrm{B}}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + V_{\mathrm{pol}} \tag{2}$$

Following the MB-DFT formalism,[55] the 2B and 3B terms of Eq. 2 are fitted to reproduce 2-body and 3-body energies calculated with the DM21 functional. It has been shown that the MB-DFT PEFs closely reproduce the results of *ab initio* MD simulations carried out with the corresponding functionals, especially for functionals that display small density-driven errors.[37,55]

It has been recently demonstrated that the interplay between functional-driven and density-driven errors effectively determines the ability of a given functional to correctly predict the properties of liquid water.[37,40] Since the DM21 functional was trained on chemical data as well as fractional charge and fractional spin constraints, and was shown to predict highly accurate electronic densities, 41 it can be concluded that the errors associated with DM21 calculations for the energetics of small water clusters discussed above are primarily functional-driven.

To investigate the effect of functional-driven errors in DM21 for liquid water, we carry

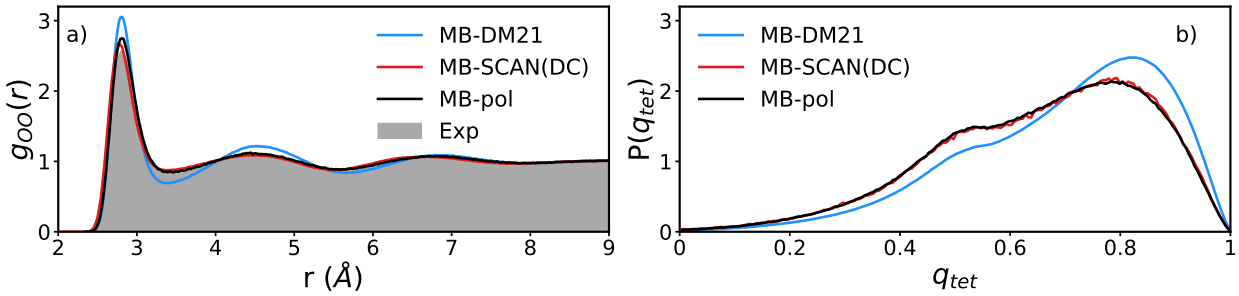

FIG. 3. Oxygen-oxygen radial distribution function (panel a), and tetrahedral order parameter distribution (panel b) calculated from MD simulations carried out with MB-DM21 at 298 K and 1 atm.

out MD simulations in the isothermal-isobaric (NPT) ensemble at 1 atm using the MB-DM21 PEF. The oxygen-oxygen radial distribution function ($g_{\mathrm{oo}}(r)$) and tetrahedral order parameter distribution (q_{tet}) calculated at 298 K and 1 atm shown in Fig. 3 indicate that MB-DM21 predicts a slightly over-structured and more tetrahedral liquid phase than both experiment and previous simulations carried out with analogous data-driven many-body PEFs derived from CCSD(T) and DC-SCAN data, i.e., MB-pol[53,56–58] and DC-SCAN(DC),[37] respectively.

Further insight into the ability of MB-DM21 to describe the properties of liquid water are gained from the analysis of the temperature dependence of the enthalpy of vaporization (ΔH_{vap}), density (ρ), and isothermal compressibility (κ_{T}) shown in Fig. 4. Consistent with the more ice-like structure inferred from the analysis of $g_{\mathrm{oo}}(r)$ and q_{tet}, MB-DM21 predicts a density (ρ) of 0.977 g/cm^3 for liquid water at 298 K, which is lower than the experimental value (0.997 g/cm^3) and the values predicted by MB-pol and MB-SCAN(DC). Overall, the density isobar predicted by MB-DM21 is shifted down by $\sim$0.03 g/cm^3 for temperature below 300 K and by $\sim$ 0.01 g/cm^3 for higher temperatures. By fitting a 3rd degree polynomial to the data points of the density isobar, the density maximum for MB-DM21 is determined to be 0.978 g/cm^3 at 296 K, which is 19 K higher than the experimental value of 277 K. Fig. 4a shows that ΔH_{vap} calculated from MB-DM21 simulations is systematically up-shifted and displays a more pronounced temperature dependence than the experimental curve, with a mean absolute error of 0.67 kcal/mol. It should be noted that the neglect of nuclear quantum effects in classical MD simulations may be responsible, at least in part, for the relatively larger deviations from the experimental values at the lower temperatures.[53,58] The

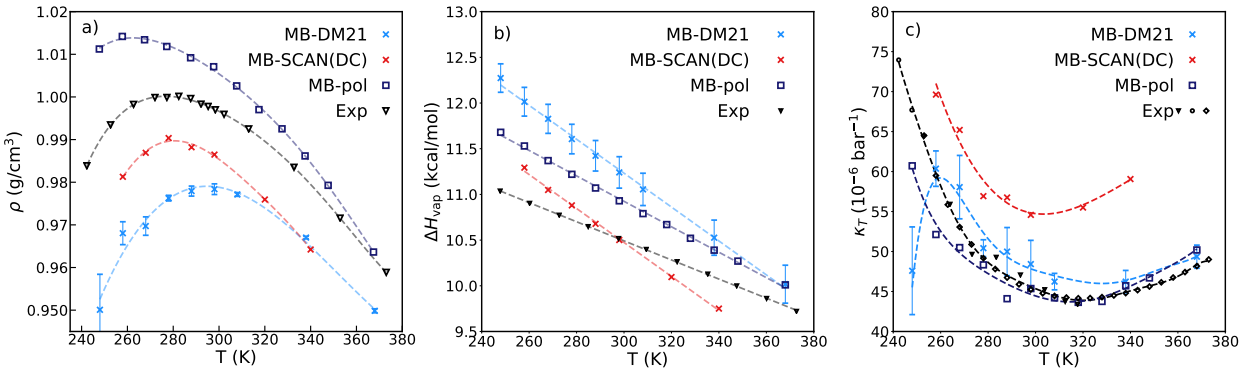

FIG. 4. Temperature dependence of the density (panel a), enthalpy of vaporization (panel b), and isothermal compressibility (panel c) calculated from MD simulations carried out with MB-DM21 at 1 atm. Also shown for comparison are the corresponding values reported in the literature for the MB-pol[53] and MB-SCAN(DC)[37] PEFs. The experimental data for the density are from the NIST Chemistry WebBook,[60], from Ref. 61 for the heat of vaporization, and the values for the isothermal compressibility are from Ref. 62 (triangles), Ref. 59 (squares), and Ref. 63 (circles).

overestimation of $\Delta H_{\rm vap}$ indicates that MB-DM21 predicts stronger hydrogen bonds between water molecules in the liquid phase, which is consistent with analysis of the DM21 interaction energies of neutral water clusters (Fig. S1) as well as with the more tetrahedral structure and lower density obtained from the MD simulations with MB-DM21. The isothermal compressibility of liquid water is notoriously difficult to predict by common force fields and DFT-based models. As shown in Fig. 4c, $\kappa_{\rm T}$ calculated from the MB-DM21 simulations is in agreement with the corresponding experimental values at higher temperatures but displays a maximum of $\sim 60\times10^{-6}$ bar^{-1} at $\sim$260 K, which is significantly different from the value of $\sim 105\times10^{-6}$ bar^{-1} at $\sim$229 K determined by recent ultrafast x-ray spectroscopic measurements.[59]

In conclusion, we have reported a systematic analysis of the DM21 functional applied to water. The analyses of interaction energies, binding energies, and individual many-body energies of neutral, protonated, and deprotonated water clusters demonstrate that DM21 is quite robust for the description of gas-phase clusters, providing similar accuracy as the range-separated, hybrid, meta-GGA ωB97M-V density functional. However, the error per monomer for the interaction energies predicted by DM21 increases with cluster size, which suggests that DM21 suffers from functional-driven errors that can be emphasized in ex-

tended aqueous systems. In addition, many-body analyses of the DM21 interaction energies of neutral and protonated water clusters shows that DM21 predicts accurate 2B energies but is affected by relatively large 3B errors. All higher-order many-body terms calculated with DM21 for neutral and protonated water display neglible errors. In the case of deprotonated water clusters, DM21 displays significant error compensation among individual n-body energy terms. MD simulations carried out with the data-driven many-body MB-DM21 PEF derived from DM21 many-body energies predict a slightly overstructured and more tetrahedral liquid at ambient conditions. Consistently, the MD simulations with MB-DM21 also predict a lower density and a higher heat of vaporization for liquid water at 1 atm in the temperature range from 240 to 370 K. In addition, the MB-DM21 simulations fail to qualitatively reproduce the temperature dependence of the isothermal compressibility of liquid water at 1 atm, predicting a significantly lower maximum at a relatively higher temperature compared to experiment. These results suggest that functional-driven errors in DM21 are not negligible and affect the overall accuracy of DM21 when applied to various aqueous systems, especially in the liquid phase. While our study is centered on the DM21 functional, the analyses presented here suggest that improving the functional form and physical content of machine-learned density functionals should allow for further reducing functional-driven errors and, in turn, enable quantitative DFT-based representations of aqueous systems from gas-phase clusters to the liquid phase.

SUPPLEMENTARY MATERIAL

Computational details, and additional analyses of interaction and binding energies of neutral, protonated, and deprotonated water clusters of different sizes.

ACKNOWLEDGEMENTS

This research was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Science, through grant no. DE-SC0019490. This research used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231, the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported

by the National Science Foundation grant number ACI-1548562, and the Triton Shared Computing Cluster (TSCC) at the San Diego Supercomputer Center (SDSC). E.P. acknowledges support from the Alfred P. Sloan Foundation Graduate Fellowship Program.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY STATEMENT

Any data generated and analyzed for this study that are not included in this Communication and its Supplementary Information are available from the authors upon request.

REFERENCES

[1]J. A. Barker and R. Watts, "Structure of water; a Monte Carlo calculation," Chem. Phys. Lett. **3**, 144–145 (1969).

[2]A. Rahman and F. H. Stillinger, "Molecular dynamics study of liquid water," J. Chem. Phys. **55**, 3336–3359 (1971).

[3]P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," Phys. Rev. **136**, B864–B871 (1964).

[4]W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," Phys. Rev. **140**, A1133 (1965).

[5]D. Prendergast and G. Galli, "X-ray absorption spectra of water from first principles calculations," Phys. Rev. Lett. **96**, 215502 (2006).

[6]J. A. Morrone and R. Car, "Nuclear quantum effects in water," Phys. Rev. Lett. **101**, 017801 (2008).

[7]P. L. Silvestrelli, M. Bernasconi, and M. Parrinello, "Ab initio infrared spectrum of liquid water," Chem. Phys. Lett. **277**, 478–482 (1997).

[8]P. L. Silvestrelli and M. Parrinello, "Water molecule dipole in the gas and in the liquid phase," Phys. Rev. Lett. **82**, 3308–3311 (1999).

[9]K. Laasonen, M. Sprik, M. Parrinello, and R. Car, ""Ab initio" liquid water," J. Chem. Phys. **99**, 9080–9089 (1993).

[10]B. L. Trout and M. Parrinello, "The dissociation mechanism of H_2O in water studied by first-principles molecular dynamics," Chem. Phys. Lett. **288**, 343–347 (1998).

[11]B. L. Trout and M. Parrinello, "Analysis of the dissociation of H_2O in water using first-principles molecular dynamics," J. Phys. Chem. B **103**, 7340–7345 (1999).

[12]T. D. Kühne, M. Krack, and M. Parrinello, "Static and dynamical properties of liquid water from first principles by a novel Car-Parrinello-like approach," J. Chem. Theory Comput. **5**, 235–241 (2009).

[13]I.-F. W. Kuo, C. J. Mundy, M. J. McGrath, J. I. Siepmann, J. VandeVondele, M. Sprik, J. Hutter, B. Chen, M. L. Klein, F. Mohamed, M. Krack, and M. Parrinello, "Liquid water from first principles: Investigation of different sampling approaches," J. Phys. Chem. B **108**, 12990–12998 (2004).

[14]M. Boero, K. Terakura, T. Ikeshoji, C. C. Liew, and M. Parrinello, "Hydrogen bonding and dipole moment of water at supercritical conditions: A first-principles molecular dynamics study," Phys. Rev. Lett. **85**, 3245–3248 (2000).

[15]M. Boero, K. Terakura, T. Ikeshoji, C. C. Liew, and M. Parrinello, "Water at supercritical conditions: A first principles study," J. Chem. Phys. **115**, 2219–2227 (2001).

[16]M. Boero, M. Parrinello, K. Terakura, T. Ikeshoji, and C. C. Liew, "First-principles molecular-dynamics simulations of a hydrated electron in normal and supercritical water," Phys. Rev. Lett. **90**, 226403 (2003).

[17]G. Hura, D. Russo, R. M. Glaeser, T. Head-Gordon, M. Krack, and M. Parrinello, "Water structure as a function of temperature x-ray scattering experiments and ab initio molecular dynamics," Phys. Chem. Chem. Phys. **5**, 1981–1991 (2003).

[18]D. Marx, M. E. Tuckerman, and M. Parrinello, "Solvated excess protons in water: Quantum effects on the hydration structure," J. Condens. Matter Phys. **12**, A153–A159 (2000).

[19]M. Sprik, J. Hutter, and M. Parrinello, "Ab initio molecular dynamics simulation of liquid water: Comparison of three gradient-corrected density functionals," J. Chem. Phys. **105**, 1142–1152 (1996).

[20]E. Fois, M. Sprik, and M. Parrinello, "Properties of supercritical water: An ab initio simulation," Chem. Phys. Lett. **223**, 411–415 (1994).

[21]J. VandeVondele, F. Mohamed, M. Krack, J. Hutter, M. Sprik, and M. Parrinello, "The influence of temperature and density functional models in ab initio molecular dynamics simulation of liquid water," J. Chem. Phys. **122**, 014515 (2005).

[22]A. K. Soper, "The quest for the structure of water and aqueous solutions," J. Condens. Matter Phys. **9**, 2717–2730 (1997).

[23]S. Izvekov and G. A. Voth, "Car–parrinello molecular dynamics simulation of liquid water: New results," J. Chem. Phys. **116**, 10372–10376 (2002).

[24]J. Ortega, J. P. Lewis, and O. F. Sankey, "First principles simulations of fluid water: The radial distribution functions," J. Chem. Phys. **106**, 3696–3702 (1997).

[25]I.-C. Lin, A. P. Seitsonen, I. Tavernelli, and U. Rothlisberger, "Structure and dynamics of liquid water from ab initio molecular dynamics—comparison of BLYP, PBE, and revPBE density functionals with and without van der waals corrections," J. Chem. Theory Comput. **8**, 3902–3910 (2012).

[26]L. Ruiz Pestana, N. Mardirossian, M. Head-Gordon, and T. Head-Gordon, "Ab initio molecular dynamics simulations of liquid water using high quality meta-gga functionals," Chem. Sci. **8**, 3554–3565 (2017).

[27]L. Ruiz Pestana, O. Marsalek, T. E. Markland, and T. Head-Gordon, "The quest for accurate liquid water properties from first principles," J. Phys. Chem. Lett. **9**, 5009–5016 (2018).

[28]M. Chen, H.-Y. Ko, R. C. Remsing, M. F. Calegari Andrade, B. Santra, Z. Sun, A. Selloni, R. Car, M. L. Klein, J. P. Perdew, and X. Wu, "Ab initio theory and modeling of water," Proc. Natl. Acad. Sci. U.S.A. **114**, 10846–10851 (2017).

[29]J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz Jr, "Density-functional theory for fractional particle number: Derivative discontinuities of the energy," Phys. Rev. Lett. **49**, 1691 (1982).

[30]Y. Zhang and W. Yang, "A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons," J. Chem. Phys. **109**, 2604–2608 (1998).

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

[32]K. Sharkas, K. Wagle, B. Santra, S. Akter, R. R. Zope, T. Baruah, K. A. Jackson, J. P. Perdew, and J. E. Peralta, "Self-interaction error overbinds water clusters but cancels in

structural energy differences," Proc. Natl. Acad. Sci. U.S.A. **117**, 11283–11288 (2020).

[33]E. Lambros, J. Hu, and F. Paesani, "Assessing the accuracy of the SCAN functional for water through a many-body analysis of the adiabatic connection formula," J. Chem. Theory Comput. **17**, 3739–3749 (2021).

[34]P. M. Piaggi, A. Z. Panagiotopoulos, P. G. Debenedetti, and R. Car, "Phase equilibrium of water with hexagonal and cubic ice using the SCAN functional," J. Chem. Theory Comput. **17**, 3065–3077 (2021).

[35]L. Zhang, H. Wang, R. Car, and W. E, "Phase diagram of a deep potential water model," Phys. Rev. Lett. **126**, 236001 (2021).

[36]C. Zhang, F. Tang, M. Chen, J. Xu, L. Zhang, D. Y. Qiu, J. P. Perdew, M. L. Klein, and X. Wu, "Modeling liquid water by climbing up jacob's ladder in density functional theory facilitated by using deep neural network potentials," J. Phys. Chem. B **125**, 11444–11456 (2021).

[37]S. Dasgupta, E. Lambros, J. Perdew, and F. Paesani, "Elevating density functional theory to chemical accuracy for water simulations through a density-corrected many-body formalism," Nat. Commun. **12**, 1–12 (2021).

[38]G. E. Scuseria, "Comparison of coupled-cluster results with a hybrid of hartree–fock and density functional theory," J. Chem. Phys. **97**, 7528–7530 (1992).

[39]M.-C. Kim, E. Sim, and K. Burke, "Understanding and reducing errors in density functional calculations," Phys. Rev. Lett. **111**, 073003 (2013).

[40]E. Palos, E. Lambros, S. Swee, J. Hu, S. Dasgupta, and F. Paesani, "Assessing the interplay between functional-driven and density-driven errors in dft models of water," (2022).

[41]J. Kirkpatrick, B. McMorrow, D. H. P. Turban, A. L. Gaunt, J. S. Spencer, A. G. D. G. Matthews, A. Obika, L. Thiry, M. Fortunato, D. Pfau, L. R. Castellanos, S. Petersen, A. W. R. Nelson, P. Kohli, P. Mori-Sánchez, D. Hassabis, and A. J. Cohen, "Pushing the frontiers of density functionals by solving the fractional electron problem," Science **374**, 1385–1389 (2021).

[42]T. Gould and S. Dale, "Poisoning density functional theory with benchmark sets of difficult systems," Phys. Chem. Chem. Phys. , DOI: 10.1039/D2CP00268J (2022).

[43]M. Riera, E. Lambros, T. T. Nguyen, A. W. Götz, and F. Paesani, "Low-order many-body interactions determine the local structure of liquid water," Chem. Sci. **10**, 8211–8218 (2019).

[44]T. T. Duignan, S. M. Kathmann, G. K. Schenter, and C. J. Mundy, "Toward a first-principles framework for predicting collective properties of electrolytes," Acc. Chem. Res. **54**, 2833–2843 (2021).

[45]D. Manna, M. K. Kesharwani, N. Sylvetsky, and J. M. Martin, "Conventional and explicitly correlated ab initio benchmark study on water clusters: Revision of the BEGDB and WATER27 data sets," J. Chem. Theory Comput. **13**, 3136–3152 (2017).

[46]J. Řezáč, P. Jurečka, K. E. Riley, J. Černỳ, H. Valdes, K. Pluháčková, K. Berka, T. Řezáč, M. Pitoňák, J. Vondrášek, and P. Hobza, "Quantum chemical benchmark energy and geometry database for molecular clusters and complex molecular systems (www.begdb.com): A users manual and examples," Collect. Czechoslov. Chem. Commun. **73**, 1261–1270 (2008).

[47]V. S. Bryantsev, M. S. Diallo, A. C. Van Duin, and W. A. Goddard III, "Evaluation of B3LYP, X3LYP, and M06-class density functionals for predicting the binding energies of neutral, protonated, and deprotonated water clusters," J. Chem. Theory Comput. **5**, 1016–1026 (2009).

[48]B. Temelso, K. A. Archer, and G. C. Shields, "Benchmark structures and binding energies of small water clusters with anharmonicity corrections," J. Phys. Chem. A **115**, 12034–12046 (2011).

[49]M. Riera, E. Lambros, T. T. Nguyen, A. W. Götz, and F. Paesani, "Low-order many-body interactions determine the local structure of liquid water," Chem. Sci. **10**, 8211–8218 (2019).

[50]G. A. Cisneros, K. T. Wikfeldt, L. Ojamäe, J. Lu, Y. Xu, H. Torabifard, A. P. Bartók, G. Csányi, V. Molinero, and F. Paesani, "Modeling molecular interactions in water: From pairwise to many-body potential energy functions," Chem. Rev. **116**, 7501–7528 (2016).

[51]C. K. Egan and F. Paesani, "Assessing many-body effects of water self-ions. I: $OH^{-}(H2O)_n$ clusters," J. Chem. Theory Comput. **14**, 1982–1997 (2018).

[52]C. K. Egan and F. Paesani, "Assessing many-body effects of water self-ions. II: $H_3O^{+}(H_2O)_n$ clusters," J. Chem. Theory Comput. **15**, 4816–4833 (2019).

[53]S. K. Reddy, S. C. Straight, P. Bajaj, C. Huy Pham, M. Riera, D. R. Moberg, M. A. Morales, C. Knight, A. W. Götz, and F. Paesani, "On the accuracy of the MB-pol many-body potential for water: Interaction energies vibrational frequencies and classical thermodynamic and dynamical properties from clusters to liquid water and ice," J. Chem. Phys. **145**, 194504 (2016).

[54]D. J. Tozer and N. C. Handy, "The importance of the asymptotic exchange-correlation potential in density functional theory," Molecular Physics **101**, 2669–2675 (2003).

[55]M. Riera, N. Mardirossian, P. Bajaj, A. W. Götz, and F. Paesani, "Toward chemical accuracy in the description of ion–water interactions through many-body representations. alkali–water dimer potential energy surfaces," J. Chem. Phys. **147**, 161715 (2017).

[56]V. Babin, C. Leforestier, and F. Paesani, "Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface vrt spectrum and second virial coefficient," J. Chem. Theory Comput. **9**, 5395–5403 (2013).

[57]V. Babin, G. R. Medders, and F. Paesani, "Development of a "first principles" water potential with flexible monomers. ii: Trimer potential energy surface third virial coefficient and small clusters," J. Chem. Theory Comput. **10**, 1599–1607 (2014).

[58]G. R. Medders, V. Babin, and F. Paesani, "Development of a "first principles" water potential with flexible monomers. III. liquid phase properties," J. Chem. Theory Comput. **10**, 2906–2910 (2014).

[59]K. H. Kim, A. Späh, H. Pathak, F. Perakis, D. Mariedahl, K. Amann-Winkel, J. A. Sellberg, J. H. Lee, S. Kim, J. Park, K. H. Nam, T. Katayama, and A. Nilsson, "Maxima in the thermodynamic response and correlation functions of deeply supercooled water," Science **358**, 1589–1593 (2017).

[60]E. W. Lemmon, M. O. McLinden, and D. G. Friend, "Thermophysical properties of fluid systems," in *NIST Chemistry WebBook*, edited by P. Linstrom and W. Mallard (National Institute of Standards and Technology, Gaithersburg (MD)).

[61]W. Wagner and A. Pruß, "The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use," J. Phys. Chem. Ref. Data **31**, 387–535 (2002).

[62]R. Speedy and C. Angell, "Isothermal compressibility of supercooled water and evidence for a thermodynamic singularity at -45 °C," J. Chem. Phys. **65**, 851–858 (1976).

[63]G. S. Kell, "Isothermal compressibility of liquid water at 1 atm." J. Chem. Eng. Data **15**, 119–122 (1970).