ANNUAL REVIEWS

Annual Review of Physical Chemistry

Many-Body Effects in Aqueous Systems: Synergies Between Interaction Analysis Techniques and Force Field Development

Joseph P. Heindel,[1] Kristina M. Herman,[1] and Sotiris S. Xantheas[1,2]

[1]Department of Chemistry, University of Washington, Seattle, Washington, USA

[2]Advanced Computing, Mathematics and Data Division, Pacific Northwest National Laboratory, Richland, Washington, USA; email: sotiris.xantheas@pnnl.gov, xantheas@uw.edu

Annu. Rev. Phys. Chem. 2023. 74:337–60

The *Annual Review of Physical Chemistry* is online at physchem.annualreviews.org

https://doi.org/10.1146/annurev-physchem-062422-023532

Keywords

intermolecular interactions, polarizable force fields, many-body expansion, molecular dynamics

Abstract

Interaction analysis techniques, including the many-body expansion (MBE), symmetry-adapted perturbation theory, and energy decomposition analysis, allow for an intuitive understanding of complex molecular interactions. We review these methods by first providing a historical context for the study of many-body interactions and discussing how nonadditivities emerge from Hamiltonians containing strictly pairwise-additive interactions. We then elaborate on the synergy between these interaction analysis techniques and the development of advanced force fields aimed at accurately reproducing the Born–Oppenheimer potential energy surface. In particular, we focus on ab initio–based force fields that aim to explicitly reproduce many-body terms and are fitted to high-level electronic structure results. These force fields generally incorporate many-body effects through (*a*) parameterization of distributed multipoles, (*b*) explicit fitting of the MBE, (*c*) inclusion of many-atom features in a neural network, and (*d*) coarse-graining of many-body terms into an effective two-body term. We also discuss the emerging use of the MBE to improve the accuracy and speed of ab initio molecular dynamics.

1. INTRODUCTION

1.1. Modeling of Molecular Interactions Involving Water

Water, as the universal solvent and critical compound for sustaining life on Earth, has attracted a plethora of research efforts targeted toward the understanding of its numerous macroscopic properties, including its anomalous behavior (1–7) compared with that of other molecular liquids (8). At the center of these efforts lies the development of atomistic-level models that are used to simulate the macroscopic properties of aqueous solutions. While first-principles approaches (9, 10) offer the advantage of transferability among different aqueous systems (e.g., binary mixtures, ions, solutes), their computational cost and challenges in theory development for periodic systems currently prevent the development of a hierarchical approach to systematically increase their accuracy, as is the case for molecular systems (11, 12). Electron correlation is starting to be added in periodic molecular dynamics (MD) simulations (13), although at a level that cannot currently parallel the advances already in place for finite (nonperiodic) systems. An alternative way to reduce cost lies with not considering the electrons of the system at all, that is, treating the system classically. Classical interaction potentials describing interactions between water molecules were introduced as early as 1933 by Bernal & Fowler (14), but it was not until 1969 and 1971, respectively, that the first Monte Carlo (15) and MD (16) computer simulations for liquid water were performed. Since then, the field of classical potential development for water has undergone a tremendous expansion.

Initially, simpler, rigid, pairwise-additive, effective potentials were developed, and subsequently, more advanced, flexible, polarizable, many-body models were introduced. In 2002, we explicitly discussed the classification of models into the previous categories in conjunction to their transferability for various aqueous environments and advocated their fitting to ab initio electronic structure results and subsequent simulations based on nuclear statistical mechanical approaches (17). These ideas laid the foundation for the development of ab initio–based many-body nonadditive potentials both by us (18–20) and by others (21–24). An earlier review (25) presented a comprehensive account of the progress made during the first 30 years of molecular simulations for water. For an updated list of existing interaction potentials for water, see Reference 8.

1.2. Types of Intermolecular Forces

Figure 1 shows a schematic of the most common types of forces used to describe interactions between atoms and/or molecules. Interactions are shown under two main categories: multipolar interactions and quantum effects. The former arise naturally in classical electrostatics; in other words, they can be understood without the knowledge of quantum mechanics. The latter consist of exchange and charge transfer and are quantum mechanical in nature. Exchange is the energetic penalty associated with enforcing the antisymmetry of an electronic wave function, and as such, it does not have a classical analog. Charge transfer, in this context, refers to partial charge transfer resulting from the delocalization of the electronic wave function.

Multipolar interactions can be further categorized into those involving the permanent charge distribution and those involving mutual polarization. Note that any interaction that depends on the polarizability will be nonadditive in nature when more than two charge densities are involved. Electrostatic interactions are the forces originating from permanent charge distributions and are pairwise additive. These interactions scale as $R_{AB}^{-(l_A+l_B+1)}$, where l_A and l_B are the order of the two interacting multipoles.

Induction is the polarization response of one charge distribution to the presence of another, static charge distribution. **Figure 1** shows that A can induce a response in B and B can induce a response in A. These interactions depend on the polarizability of each molecule. For neutral,

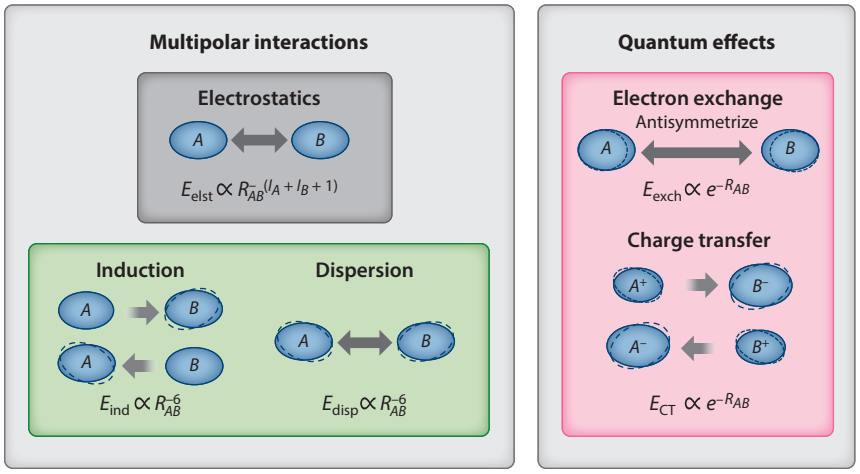

Figure 1

Common types of intermolecular forces. Ellipses represent the electron density, and dashed lines denote its deformation due to the interaction. Interactions on the left (multipolar interactions) appear naturally in classical electrostatics, whereas interactions on the right (quantum effects) are due to the electrons. Multipolar interactions involve either the permanent charge distribution (*dark gray box*) or mutual polarization (*green box*). Also shown is the nominal scaling of each interaction. In the case of electrostatics, the scaling depends on the type of interacting multipoles, so that l_A and l_B are the orders of the multipole ($l = 0$ for a point charge, $l = 1$ for a dipole, etc.).

polar molecules, the induction energy scales as R_{AB}^{-6}. In the case that either A or B is charged, the interaction would scale as R_{AB}^{-4}.

Since both A and B can induce multipoles on one another, it is natural to ask not only about the interaction of those induced multipoles with a static charge density, which we call induction, but also about the interaction of those induced multipoles with one another, which we call dispersion. The dispersion energy scales as R_{AB}^{-6}, as is well known from the London dispersion formula (Equation 6). Note that dispersion is also proportional to the product of polarizabilities, $\alpha_A \alpha_B$, which reinforces the idea that dispersion arises from mutual polarization.

Charge transfer, in the context of intermolecular forces, usually refers to partial charge transfer between two molecules A and B, arising from the delocalization of electronic wave functions. Charge transfer can be thought of as an extreme form of polarization, which is most relevant at short range. Empirically, charge transfer decreases exponentially with distance and, thus, is significant only at short range, where the charge densities overlap. In this sense, charge transfer can be regarded as a short-range correction to the polarization energy.

Finally, the exchange interaction is a purely repulsive, quantum mechanical, short-range effect, decaying exponentially with interatomic distance. Exchange is destabilizing because it originates in the requirement that the electronic wave function be antisymmetric with respect to exchanging electronic coordinates, which effectively limits the configurational freedom of the system. Even though the requirement of antisymmetry in the wave function arises from quantum mechanics, the repulsive nature of the interaction can be understood classically as an enhancement of nuclear repulsion due to the depletion of electron density between A and B (26). Therefore, just as charge transfer can be regarded as a quantum mechanical correction to the polarization energy, exchange can be considered a quantum mechanical correction to the electrostatic energy.

Note that the intermolecular interactions we have introduced here are meant to model the total interaction energies produced by solving the Schrödinger equation when allowing A and B

to interact. None of the types of interactions described above correspond to physical observables, as they are used primarily to model the physical components of the total interaction. However, we discuss various approaches of calculating these terms in ways that match physical intuition and, more importantly, guide the development of advanced force fields that can reproduce the total quantum mechanical energies.

1.3. Many-Body Expansion

We have now introduced the various ways that intermolecular forces are typically conceptualized, emphasizing the role of polarization in creating nonadditive contributions to the interaction energy. Since the partitioning of the total interaction energy among its various components is not unique, and even the components under the same name can be defined differently, an alternative path is to partition a property into its k-body components, where k is the number of bodies (i.e., subsystems) that a system can be divided into. That is, one may simply be interested in the total three-body contribution to the energy, rather than the individual three-body induction and dispersion energies separately. This partitioning of a property into k-body components can be accomplished via the many-body expansion (MBE). The MBE allows one to calculate the total pairwise and the total nonadditive contributions to any property. The earliest study of nonadditivity in the energy of water clusters is that by Hankins et al. (27), with similar studies following soon after (28, 29). Those studies calculated the nonadditivity of various configurations of $(H_2O)_3$ by subtracting the pairwise interaction energies, ΔE_{ij}, from the total energy, E_{ijk}, of the system. Later studies (e.g., 30) expanded the MBE to larger water clusters.

The MBE is a method that enables the calculation of the total k-body contribution to any molecular property. The MBE of the total energy can be written explicitly for an N-molecule system as

$$E_N^{(1)} = \sum_{i=1}^{N} E_i - E_{\text{ref}}, \qquad 1.$$

$$E_N^{(2)} = \sum_{i<j}^{N} E_{ij} - E_i - E_j, \qquad 2.$$

$$E_N^{(3)} = \sum_{i<j<k}^{N} E_{ijk} - E_{ij} - E_{ik} - E_{jk} + E_i + E_j + E_k. \qquad 3.$$

Higher many-body terms can be defined analogously. The MBE is in reality just a counting technique that keeps sets of increasing size balanced at each order, allowing it to be applied to any molecular property using the above equations. A much more compact form of the MBE, which expresses the k-body term $E_N^{(k)}$ in terms of the sum of energies of the m-mers with $m < k$, appropriately scaled by the number of times each m-mer will appear in all possible k-mers, is (31)

$$E_N^{(k+1)} = \sum_{m=0}^{k} (-1)^m \binom{N-k+m-1}{m} \sum_{n=1}^{\binom{N}{k-m+1}} E_n^{(k-m+1)}. \qquad 4.$$

This form makes the combinatorial nature of the MBE much more explicit. Although other choices are possible (32), the MBE is usually utilized in the context where each body or fragment is an entire molecule. Therefore, the one-body term can be thought of as the change of the intramolecular geometry of a molecule, due to the interaction with another body, relative to the isolated fragment. The two- through N-body terms can be thought of as containing the sum of all

the types of intermolecular interactions discussed in Section 1.2, namely electrostatics, exchange, induction, dispersion, charge transfer, and the various couplings between these interactions. In the MBE, the nonadditive contributions to the energetics can be computed without considering the types of interactions involved (as described above).

1.4. Electrostatic Origin of Many-Body Effects

While it has long been appreciated that many-body corrections to intermolecular forces exist, it may not be immediately clear how these terms arise from Hamiltonians that contain strictly pairwise-additive pieces. Consider the usual electronic potential (in atomic units) with electronic positions $\mathbf{r}_i$ and nuclear positions $\mathbf{R}_i$:

$$\hat{H}_{\text{el}} = -\frac{1}{2}\nabla^2_{\mathbf{r}_i} - \sum_i \sum_j \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}_j|} + \sum_{i<j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad 5.$$

Every term in this Hamiltonian is pairwise additive, depending only on the nuclear charges, Z_i, and the nuclear–nuclear, nuclear–electron, and electron–electron distances. How, then, do nonadditivities arise in intermolecular interactions?

One of the earliest answers to this question came from Fritz London, resulting in what we now call London dispersion forces. London (33) noticed that if one considers the composite Hamiltonian of two spherically symmetric atoms, $\hat{H}^{AB}_{\text{el}} = \hat{H}^{A}_{\text{el}} + \hat{H}^{B}_{\text{el}}$, the electronic spectrum will result in all states being doubly degenerate because $\hat{H}^{AB}_{\text{el}}$ is block diagonal. Circa 1930, it was already known that noble gases could interact weakly, resulting in non-ideal-gas behavior, as first explained by the van der Waals equation of state (34). Therefore, London tried to introduce interatomic interactions by perturbing $\hat{H}^{AB}_{\text{el}}$ with a Coulomb potential. To make progress, this perturbation was expanded in powers of $1/\mathbf{R}_{AB}$, which produces a multipole expansion. When the analysis is carried out to second order in perturbation theory, a pairwise interaction emerges between atoms A and B:

$$E^{AB}_{\text{disp}} = -\frac{3}{2}\frac{I_A I_B}{I_A + I_B}\frac{\alpha_A \alpha_B}{R^6_{AB}}. \quad 6.$$

In this equation, I_A is the first ionization energy of A, α_A is the dipole polarizability of atom A, and R_{AB} is the distance between atoms A and B. Even though this interaction is pairwise with respect to the atoms, it is actually a many-body effect with respect to the electronic degrees of freedom. That is, the electrons on atom B perturb the electrons on atom A, which means that the optimal electronic configuration on atom A will no longer be the ground-state wave function of the isolated atom A.

In 1943, Axilrod & Teller (35) took London's analysis to third order in the perturbative expansion and introduced the following formula:

$$E_{ABC} = E_0 \left(\frac{1 + 3\cos\theta_{AB}\cos\theta_{BC}\cos\theta_{AC}}{(R_{AB}R_{BC}R_{AC})^3} \right), \quad 7.$$

which depends on the distances between three atoms, R_{AB}, R_{BC}, and R_{AC}, and the angles between those atoms, θ_{AB}, θ_{BC}, and θ_{AC}. The interaction also has a scale, which is determined by the constant E_0.

The lessons learned by analyzing these equations is that atoms are intrinsically polarizable, as demonstrated by the appearance of a dipole polarizability in the London dispersion energy, which arises from a perturbation of atoms that do not have a permanent dipole. Additionally, we see that these nonadditive effects are short-ranged in nature. The pairwise London dispersion decays as R^{-6}, while the Axilrod–Teller term decays as the product of three R^{-3} terms. Therefore,

one source of nonadditivity in molecular interactions is that electron densities are intrinsically polarizable, and pairwise polarization will always have a many-body correction, as in the case of the Axilrod–Teller potential.

Dispersion has historically received a lot of attention due to its explanation of the interaction between nonpolar substances. However, in aqueous systems we are usually concerned with atoms and molecules that have nonvanishing multipole moments. Therefore, we note that there are related analyses when the atoms or molecules of interest have a permanent charge, dipole, and so forth. In short, the usual charge–dipole and dipole–dipole interactions will be enhanced by the presence of other charges and dipoles. This multipole enhancement, called induction, is another type of nonadditive interaction and is often more important than dispersion in aqueous systems. Note that induction can be nonzero even when no interfragment electron correlation is considered, whereas dispersion arises purely from interfragment electron correlation (36).

Another important type of nonadditivity arises from the fermionic nature of electrons. If we again consider the noninteracting, composite Hamiltonian, $\hat{H}_{\text{el}}^{AB}$, we can see that there is already a problem with the eigenfunctions of this Hamiltonian being given by products of the eigenfunctions of $\hat{H}_{\text{el}}^{A}$ and $\hat{H}_{\text{el}}^{B}$. That is, this total wave function will not be antisymmetric with respect to swapping electronic coordinates between monomers, thus violating the Pauli exclusion principle. Therefore, even if we ignore the Coulombic perturbation analyzed by London, there will be an interaction between nonpolar atoms due to antisymmetrization of the total system wave function. This interaction is typically referred to as exchange repulsion or Pauli repulsion because it turns out to be purely repulsive and decays exponentially with interatomic distance (37). Exchange repulsion can clearly cause many-molecule nonadditivities because if one had an antisymmetrized dimer wave function, ψ_{AB}, the presence of a third atom or molecule, C, would require the total system wave function, ψ_{ABC}, to be antisymmetrized, which in turn would modify the interaction energy.

Finally, we note that because both induction and dispersion interactions can be regarded as perturbations of the molecular Hamiltonian, which would then result in modified wave functions, there will always be an additional nonadditivity arising from the need to antisymmetrize the wave function after accounting for these interactions. Indeed, this is why one gets mixed exchange–induction and exchange–dispersion terms at second order in symmetry-adapted perturbation theory (36).

2. INTERACTION ANALYSIS TECHNIQUES

2.1. Symmetry-Adapted Perturbation Theory

Even though this review is concerned primarily with many-body effects and many-body force fields, it is still important to focus on the pairwise-additive pieces of molecular interactions. This is because intermolecular interactions, especially for hydrogen-bonded systems, are predominantly pairwise additive, even when many-body effects are large (30). To this end, it is important to understand the terms which dominate two-body interactions. Furthermore, if a many-body force field cannot accurately describe pairwise interactions, then the many-body terms must compensate by relying on a cancellation of errors. Therefore, a proper understanding of pairwise interactions is crucial for the development of any many-body method.

Above, we have introduced the various types of molecular interactions in the form of electrostatics, induction, dispersion, exchange, and charge transfer, all of which, except for electrostatics, contribute nonnegligibly to many-body effects. The emergence of these different interactions is pretty straightforward from symmetry-adapted perturbation theory (SAPT), which builds up these components from a perturbative expansion of the isolated monomer wave functions. SAPT

is a well-developed method of describing intermolecular forces. Due to the space limitations of this article, we refer the interested reader to recent reviews for a more comprehensive discussion (36, 38, 39).

Conceptually, the simplest form of two-body SAPT arises from considering the interaction between two fragments with exact electronic wave functions of the fragment Hamiltonians $\hat{H}_{\text{el}}^{A}$ and $\hat{H}_{\text{el}}^{B}$. In this case, Rayleigh–Schrödinger perturbation theory results in the following first- and second-order corrections to the energy:

$$E^{(1)} = \langle \Psi_A^{(0)} \Psi_B^{(0)} | V_{\text{el}} | \Psi_A^{(0)} \Psi_B^{(0)} \rangle, \qquad 8.$$

$$E^{(2)} = \sum_{i \neq 0} \frac{|\langle \Psi_A^{(0)} \Psi_B^{(0)} | V_{\text{el}} | \Psi_A^{(i)} \Psi_B^{(0)} \rangle|^2}{E_A^{(0)} - E_A^{(i)}} + \sum_{j \neq 0} \frac{|\langle \Psi_A^{(0)} \Psi_B^{(0)} | V_{\text{el}} | \Psi_A^{(0)} \Psi_B^{(j)} \rangle|^2}{E_B^{(0)} - E_B^{(j)}}$$

$$+ \sum_{i,j \neq 0} \frac{|\langle \Psi_A^{(0)} \Psi_B^{(0)} | V_{\text{el}} | \Psi_A^{(i)} \Psi_B^{(j)} \rangle|^2}{E_A^{(0)} + E_B^{(0)} - E_A^{(i)} - E_B^{(j)}}. \qquad 9.$$

The first-order correction to the eigenvalues of Hamiltonian $\hat{H}_{\text{el}}^{A} + \hat{H}_{\text{el}}^{B}$, as shown in Equation 8, is simply the expectation value of the potential energy piece of Equation 5 over the unperturbed product state wave function, $\Psi_A^{(0)} \Psi_B^{(0)}$. This has a very simple interpretation as the electrostatic interaction of the fragment charge densities. The first term of the second-order correction, Equation 9, can be interpreted as a polarization of the fragment A charge density due to the static charge density of fragment B, and vice versa for the second term. In other words, the first two terms in Equation 9 are the $B \rightarrow A$ and $A \rightarrow B$ induction energies. The third term in Equation 9 involves evaluations of the interaction operator, V_{el}, over all doubly excited states of $\Psi_A^{(0)} \Psi_B^{(0)}$. This term is interpreted as the dispersion interaction between fragments A and B. The fact that we have only excited-state wave functions, $\Psi_A^{(i)} \Psi_B^{(j)}$, in this expression means that electron correlation gives rise to dispersion, since electron correlation is often incorporated into the electronic energy by using the excited-state manifold of some reference state as a basis for further calculations.

Throughout our discussion of Equations 8 and 9, we have used only the direct product wave function $\Psi_A^{(0)} \Psi_B^{(0)}$ (and analogous excited states), which, as discussed above, is not properly antisymmetrized and hence cannot be an eigenfunction of Equation 5. For this reason, the energy corrections involving these product wave functions are referred to as the polarization approximation, as they primarily describe the energetic stabilization due to electronic polarization. To correct for this problem, one must antisymmetrize the wave functions, sometimes called symmetry adapting the wave functions, which gives rise to corrections at all orders arising from electronic exchange (36).

Now that we have established how SAPT describes two-body interactions, we turn our attention to many-body effects described within the SAPT framework. The theory and development of three-body SAPT have been led primarily by the Szalewicz group, and these calculations can be carried out with their SAPT2020 package (38). Many of the earliest three-body SAPT calculations performed by this group focused on rare gas trimers, such as He_3 (40), Ar_3 (41), and Ar_2HF (42). Due to the computational expense of ordinary three-body SAPT, recent developments have focused on three-body SAPT(DFT) (43). The SAPT(DFT) method describes the intramonomer electron correlation at the DFT level, whereas all intermonomer interactions are handled by a perturbative expansion. One rather impressive use of three-body SAPT(DFT) was the calculation of the nonadditive interactions in a benzene trimer (44). The three-body SAPT(DFT) is a promising technique for studying nonadditive interactions in great detail, and

more applications of this method to aqueous systems are expected in the future. The highly detailed nature of the results could be valuable in guiding the development of aqueous force fields, especially those involving ions.

2.2. Many-Body Expansion

We have already introduced the MBE as a method for calculating the total *k*-body contribution to any molecular property. In this section, we present some recent examples of its usefulness in analyzing molecular interactions. In general, the MBE provides a coarser view of many-body effects than SAPT (36), which explicitly calculates all of these types of molecular interactions through perturbation theory. However, the MBE is much easier to calculate for three-body and higher terms because it requires only the sums and differences of independent electronic structure calculations. The MBE can also be applied to obtain the many-body contributions to properties such as the dipole moment, or forces (45–47).

An advantage of three-body SAPT is that it casts the nonadditive interactions into physically intuitive and interpretable terms such as induction, dispersion, and exchange, albeit at a great computational cost (38). The MBE falls on the opposite end of the spectrum. The various terms in the MBE can be evaluated using any electronic structure method that yields the energies of the various "bodies" (dimers, trimers, tetramers, etc.). This flexibility allows for nonadditive contributions to molecular properties to be quantified with relative ease. The many-body effects converge with respect to the level of electron correlation and basis set much more quickly than pairwise interactions, as long as a correction for basis set superposition error is taken into account, as recently shown for water clusters (48), monatomic ion–water clusters (49), and polyatomic ion–water clusters including ions in the Hofmeister series (50). Therefore, although the MBE provides less-detailed information about the nature of nonadditive interactions than, for example, the three-body SAPT, it tends to be much simpler to apply, a fact that allows for a higher throughput of chemical systems and the elucidation of complex trends in an efficient way.

As an example, **Figure 2*a*,*b*** shows the magnitude of the two-, three-, and four-body energies of various clusters of Hofmeister ions of $Z^{+/-}(H_2O)_9$ either internally or externally solvated by nine water molecules (for full details, see Reference 50). For charge-dense ions, such as $Ca^{2+}(H_2O)_9$, there can be significant many-body contributions even at the four-body level (denoted I-W-W-W). In contrast, chaotropic (structure-breaking) ions display smaller many-body contributions to their total interaction energies than kosmotropic (structure-making) ones. Calculating four-body contributions to binding energies is currently infeasible using SAPT, and likely will be for the foreseeable future. On the other hand, while we are able to note the potential importance of four-body effects via the MBE, we cannot discern what type of interaction results in these high-order nonadditivities.

An emerging trend shows that three-body I-W-W interactions are generally repulsive, whereas **Figure 2*c*,*d*** suggests the existence of a surprisingly linear, uniform anticorrelation between the two-body I-W and the three-body I-W-W terms; in other words, the former is a descriptor of the latter. These correlations emphasize the delicate balance between the attractive and repulsive components of the total binding energy that determine the stability of a particular configuration. These balances become more delicate when considering the dynamics of ion–water systems, demonstrating the difficulty of constructing general and accurate many-body ion–water force fields.

The ease with which one can explore many-body effects via the MBE makes it a valuable tool in guiding the development of many-body force fields. Indeed, the MBE is the most common technique for studying three-body effects and, to the best of our knowledge, the only approach to explicitly calculate four-body and higher terms.

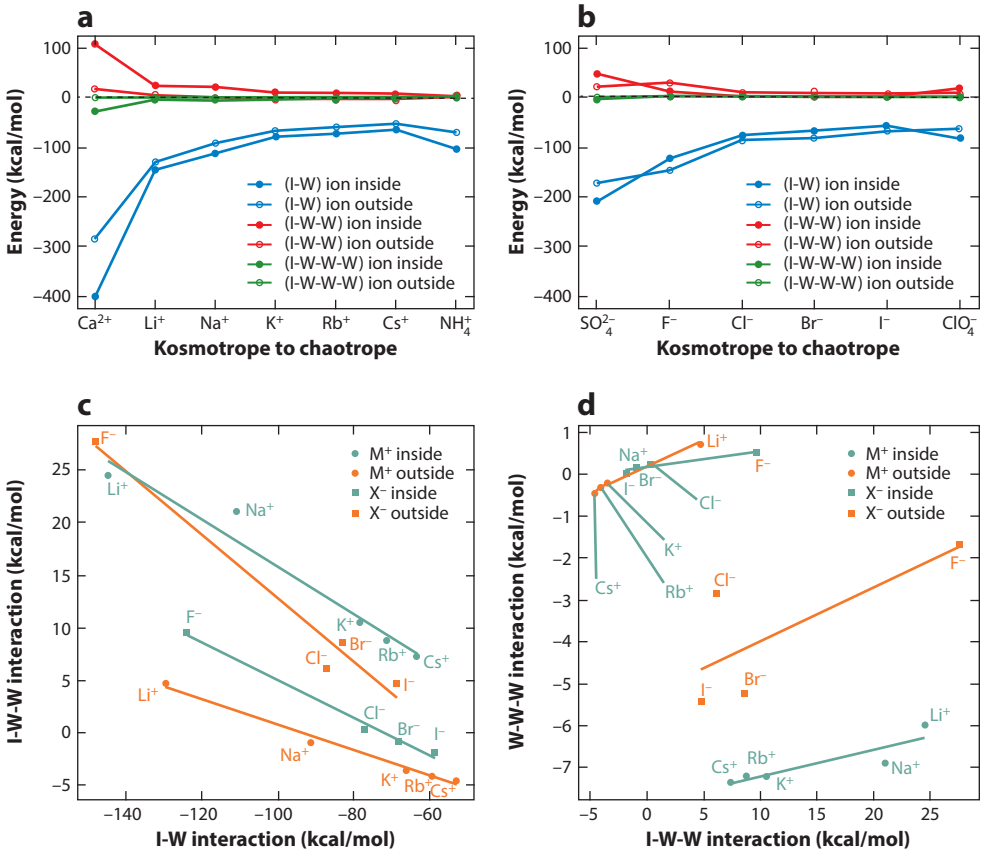

Figure 2

Two- through four-body I-W many-body terms for various (*a*) positively charged and (*b*) negatively charged aqueous Hofmeister ions. (*c*,*d*) Correlation between various many-body terms in $X^-(H_2O)_9$ and $M^+(H_2O)_9$ I-W clusters, where $X^- = F^-$, Cl^-, Br^-, or I^- and $M^+ = Li^+$, Na^+, K^+, Rb^+, or Cs^+. Abbreviation: I-W, ion–water. Panels *a* and *b* adapted from Reference 50, and panels *c* and *d* adapted from Reference 49.

2.3. Energy Decomposition Analysis

Thus far, we have described two paradigms for analyzing intermolecular interactions. SAPT builds up the various interaction components from perturbative corrections to the isolated fragments, whereas the MBE provides a mechanism for separating the total energy of a collection of fragments into the *k*-body contributions through a series of calculations of all possible subsystems. Energy decomposition analysis (EDA) attempts to combine the merits of SAPT and MBE, namely by using a well-established framework of electronic structure calculations (like the MBE does), and still produce a breakdown of the energy in terms of components similar to those in **Figure 1** (as SAPT does). Many notable EDA techniques exist, but we only have space here to focus on the newest ones, which have been used specifically in the development of many-body force fields. For a more complete background on EDAs, we direct the reader to a recent review (51).

An EDA scheme was recently introduced and utilized by the Head-Gordon groups (51). It originates from the definition of absolutely localized molecular orbitals (ALMOs) (51, 52). Some desirable properties of the scheme are as follows. It provides an estimate of electrostatics, dispersion, polarization, Pauli repulsion, and charge transfer. Also, the full calculation is variational, is

amenable to use with correlated wave functions from either DFT (53) or MP2 (54, 55), and obeys various other constraints that try to minimize the arbitrariness of any EDA method (56).

3. APPROACHES FOR DEVELOPING MANY-BODY FORCE FIELDS

Having introduced techniques for studying intermolecular interactions, we now focus on the construction of many-body force fields, bearing in mind that our goal is to describe how these analysis techniques can guide the construction of advanced many-body force fields.

The goal of fitting a many-body force field can be stated as follows: Find a fast-to-evaluate and accurate approximation of the electronic Schrödinger equation for a particular system or set of systems. To circumvent the fact that the task of explicitly constructing electronic wave functions is relatively slow, the electronic degrees of freedom must be neglected, and we must either resort to a classical approximation of the interactions or develop a protocol for fitting the potential energy surface (PES). To that end, we provide a brief background of the most common approaches used for developing many-body force fields.

Before we begin our discussion, we establish a framework that broadly classifies many-body force fields into two classes, namely implicitly and explicitly many-body force fields. Below, we contrast these two with effective two-body force fields. Explicitly many-body (e-MB) force fields, in one way or another, directly fit the terms of the MBE as calculated by some level of electronic structure. The data used to fit the many-body terms must come from theoretical calculations because many-body terms are not experimentally measurable quantities. In contrast, implicitly many-body (i-MB) force fields form a much wider class of force fields, partly because they are historically older than e-MB models and partly because they are derived from the numerous classical models of polarization. i-MB force fields are typically constructed using some kind of classical electrostatic model, which attempts to reproduce the relevant quantum mechanical interactions. Many different electrostatic models can be used to develop a force field, but they all must model polarization in some way. We discuss some of the most popular techniques for creating i-MB force fields, but because of the wide and disparate array of models that have been proposed over the years, this cannot be considered an exhaustive list. Rather, we skew toward approaches that have been the most successful and are actively being pursued.

3.1. Implicit Many-Body Potentials

i-MB potentials are those that leverage classical models of polarization to describe the nonadditive interactions induced by a molecular environment. Usually, this means i-MB models use a multipole expansion, in which the multipoles are able to change in response to the local electric field. For reasons of simplicity and computational speed, i-MB models are typically restricted to either polarizable charges or polarizable dipoles.

3.1.1. Models based on the distributed multipole expansion. Within classical electrostatics, one might first consider approximating intermolecular interactions via a multipole expansion of the molecular charge density. Unfortunately, this approach fails because it lacks the flexibility to capture short-range variations in the charge density and often predicts the wrong structure for molecular complexes due to a lack of accounting for anisotropy and the short-range variation in the interactions (57, 58). Additional flexibility is introduced via a separate set of multipoles that are often placed on each atom in a molecule, resulting in a so-called distributed multipole expansion (57). If we consider two molecules A and B, with multipolar interactions sites i and j containing charges q_i, dipoles $\boldsymbol{\mu}_i$, and quadrupoles $\boldsymbol{\Theta}_{ij}$, the interaction between these molecules (to second

order in the multipole expansion) is

$$V = \sum_{i \in A} \sum_{j \in B} \left[\boldsymbol{T}_{ij} q_i q_j + \boldsymbol{T}^{\alpha}_{ij} (q_i \boldsymbol{\mu}_j - \boldsymbol{\mu}_i q_j) + \boldsymbol{T}^{\alpha\beta}_{ij} \left(\frac{1}{3} q_i \boldsymbol{\Theta}^{b}_{ij} - \boldsymbol{\mu}_i \boldsymbol{\mu}_j + \frac{1}{3} \boldsymbol{\Theta}^{a}_{ij} q_j \right) \right]. \quad 10.$$

In Equation 10, $\boldsymbol{T}_{ij}$, $\boldsymbol{T}^{\alpha}_{ij}$, and $\boldsymbol{T}^{\alpha\beta}_{ij}$ are the interaction tensors formed by expanding an $\boldsymbol{R}^{-1}$ potential in a Taylor series. Stone (57) has described the explicit form of these tensors, with $\boldsymbol{\mu}_i$ and $\boldsymbol{\Theta}^{a}_{ij}$ being the dipole and quadrupole moments at site i (on atom a) in molecule A, respectively. For a thorough discussion of distributed multipole analyses, including consideration of higher-order multipoles, we refer the reader to Stone's authoritative book (57).

An important feature of Equation 10 is that the multipoles, q_i, $\boldsymbol{\mu}_i$, and $\boldsymbol{\Theta}^{a}_{ij}$, need not be static. If the multipoles were static, then the interaction potential would be strictly pairwise additive. To recover the many-body character of these interactions, we must allow the multipoles to be modulated by polarizabilities and hyperpolarizabilities, explicitly accounting for the induction energy. This gives rise to the implicit description of the many-body terms within Equation 10. Force fields employing distributed multipoles are perhaps the most common form of many-body force fields. These potentials typically include electrostatic, induction, exchange, dispersion, and sometimes charge-transfer components. For a detailed explanation of how these intermolecular forces are implemented, see the **Supplemental Material**.

Supplemental Material >

Some of the earliest data-driven models for describing pure water are TTM2-F (18), TTM2.1-F (19), and TTM3-F (20). The TTM models utilize a spectroscopically accurate one-body PES (59) and ab initio dimer data to fit the two-body PES while including all higher-order terms by an induction scheme. The development of H_2O–solute interaction potentials based on distributed multipoles is an active area of research. The i-TTM force fields (60, 61) were constructed to provide the long-range representation of the e-MB ion–water potentials (discussed in the next section).

The MB-UCB potential is one of the best examples to date of the synergy between interaction analysis techniques and force field construction (62). It is based on the ALMO-EDA method (63–65) of decomposing interactions in small water clusters [up to $(H_2O)_5$] and explicitly matching the interaction terms by using appropriate multipolar descriptions for each term. The MB-UCB separately predicts electrostatics, polarization, charge transfer, and van der Waals energies. This is important because it allows the force field to be parameterized against a physically meaningful set of interactions, thereby mitigating a reliance on cancellation of errors. In this manner, the MB-UCB potential is more transferable than other polarizable potentials (66). An EDA-guided ion–water force field following the same approach as for MB-UCB was recently published (67). The overall success of this and similar models remains to be seen, but the methodology of using EDA to guide the parameterization of force fields seems promising, especially because strong ion–water interactions can result in highly nonadditive interactions (50).

3.1.2. Drude and fluctuating charge models. There are two other common types of polarizable force field, known as Drude (68) and fluctuating charge models (69, 70). We have left the discussion of these models until now because they are most commonly used for large-scale biomolecular simulations and they typically do not aim to reproduce the underlying Born–Oppenheimer PES. These models allow for an approximate treatment of polarization, making the parameterization of the model transferable to a wider range of systems and areas of their phase diagram.

Drude models use an empirical form of polarization, in which a charge is allowed to fluctuate on a spring attached to a nucleus. The position of these Drude oscillators is then optimized in the presence of other Drude oscillators. A large displacement of the Drude oscillator results in a

larger induced dipole on that atom. Fluctuating charge models are similar to fixed charge models in that at any point in the simulation, every atom has a specified charge and the interaction energy is simply the Coulomb interaction. The difference, however, is that the charges are treated as dynamical variables of the system using an extended Lagrangian approach (71) subject to an electroneutrality constraint (the total system charge cannot fluctuate).

3.2. Explicit Many-Body Potentials

e-MB potentials are those that aim to achieve an accurate description of short-range molecular interactions by fitting the terms of the MBE using a set of electronic structure calculations. Their development is made possible by the many advances in correlated electronic structure over the last few decades.

3.2.1. Fitting the many-body expansion with permutationally invariant polynomials. One major difficulty with explicitly fitting the terms of the MBE is the need for the fit to be consistent with the relevant physical symmetries. Most notably, the potential must be invariant with respect to permutations of identical atoms and translation of atoms, and it must be rotationally equivariant. One of the most successful ways to capture these molecular symmetries in a fitted potential is the use of permutationally invariant polynomials (PIPs) (72, 73). PIPs are formed from a collection of symmetrized monomials built from a transformation of the internuclear distances which decay appropriately at long range (for details, see the recent review in Reference 74).

3.2.2. Water potentials fitted to the many-body expansion. The WHBB (Wang–Huang–Braams–Bowman) potential from the Bowman group (75) is the first potential for water based on fitting PIPs to large quantities of ab initio data meant to reproduce the two- and three-body terms of the MBE. The one-body term, as first introduced in the TTM potentials (19, 20), utilizes the spectroscopically accurate Partridge–Schwenke PES and dipole moment surface (DMS) (59). The two-body term is fitted to roughly 30,000 CCSD(T)/aug-cc-pVTZ dimer configurations, while the three-body term is fitted to roughly 30,000 MP2/aug-cc-pVTZ trimer configurations. The DMS is a sum of one- and two-body MP2/aug-cc-pVTZ DMSs, which should be very nearly converged with respect to the full DMS, since the electrostatic properties of water are known to converge more quickly than the energy (45). Note that the two-body and three-body fitted potentials for the WHBB model provide a short-range correction to another polarizable potential, namely TTM3-F (20). As discussed above, it can be difficult to reproduce the many-body quantum mechanical interactions, which dominate at short range, by using distributed multipoles. These potentials, however, are guaranteed to be asymptotically correct, so the two-body PES in WHBB switches to the TTM3-F PES beyond some cutoff. Additionally, all terms beyond three-body are provided by the TTM3-F potential.

After the introduction of the WHBB potential, a very similar protocol was used to construct the MB-Pol potential. The one-body potential is again the Partridge–Schwenke PES, originally introduced in the TTM potentials. The two-body PES is fitted with PIPs to 42,508 CCSD(T)/CBS water dimer energies, whereas the three-body PES is a fit to 12,000 CCSD(T)/aug-cc-pVTZ energies (76). Once again, each of the terms switches to a long-range polarizable potential, namely the TTM4-F potential (77). The TTM4-F potential is also used to capture the four-body and higher terms in MB-Pol.

This approach to fitting a water potential has proven to be quite successful in reproducing water cluster binding energies and zero-point energies (78), as well as condensed-phase properties (79, 80). By explicitly fitting the terms of the MBE, one manages to circumvent the many difficulties associated with reproducing the various intermolecular interactions, which appear in a complex nonadditive manner at short range. These potentials are quite accurate and have set the benchmark

for how accurate a fitted potential can be (79). Note, however, that constructing force fields in this way requires a very large quantity of data and results in potentials that are typically between two and ten times slower than distributed multipole potentials.

3.2.3. Solute–water potentials fitted to the many-body expansion. The careful consideration of many-body terms has resulted in several excellent pure water potentials, so the next frontier to consider is the interaction of water with various solutes, such as small molecules and positive/negative ions. If one is building on an e-MB potential, then the additional solvent–solute and solute–solute many-body terms must be fitted. To be more explicit, consider the case of some molecule or ion M interacting with H_2O where the e-MB potential for H_2O is already in place. This means we need to fit the one-body term, E_{1B}^{M}, the two-body term, $E_{2B}^{M,W}$, and the three-body term, $E_{3B}^{M,W,W}$, where W is shorthand for H_2O. Notice that with these terms, only a single solute M interacting with H_2O can be simulated. To allow for multiple solutes or ions, the intersolute interaction terms $E_{2B}^{M,M}$, $E_{3B}^{M,M,W}$, and $E_{3B}^{M,M,M}$ must also be fitted. For certain systems or at low concentrations, some of these intersolute terms may be negligible, but this is not known a priori. Therefore, to study a system containing three different molecules, such as water solvating both Na^+ and Cl^-, the number of terms in the MBE gets increasingly large. This fitting process has already been implemented for water interacting with H_2 (81, 82), CH_4 (83), CO_2 (84), halide ions (60, 85), and alkali metal ions (61, 86). The typical process involves sampling the dimer and trimer PESs, performing hundreds of thousands of electronic structure calculations, and fitting these terms using PIPs. This process, however, provides only the short-range part of the potential, so one must also construct a distributed multipole potential, called i-TTM for alkali metal and halide ions (60, 61), to provide the long-range part of the potential. The recently released MB-Fit software can aid this rather involved process (87). Further advances in the use of gradients for fitting (88), Δ-ML techniques (89), and active learning (90) can lower the barrier to constructing e-MB force fields.

Efforts to construct a highly accurate PIP-based force field for protonated water systems have recently been reported (73, 91). Due to the reactive nature of $H^+(H_2O)_n$, namely the conversion between Eigen and Zundel ions (92), this type of heterogeneous e-MB force field is more difficult to construct than the corresponding homogeneous ones, as for pure water. However, they are quite accurate when compared with high-level electronic structure calculations of clusters (91).

One can also construct e-MB potentials by building up a description of intermolecular interactions with SAPT (38). Indeed, the Szalewicz group (93) has used this approach to construct the SAPT-3B force field, which includes explicitly fitted two- and three-body terms from the corresponding perturbation expansions and layers on a polarizable model for the four- to *N*-body terms.

3.3. Machine-Learned Potentials (Implicit/Explicit)

Another promising method for developing potentials for aqueous systems is using machine learning (ML). There are too many ways of fitting PESs using ML to describe in detail here, so we refer the interested reader to various recent reviews and focus only on neural network (NN) potentials (94–98). As with PIPs, NN potentials have been used to reproduce ab initio reference data for protonated water clusters (99). An important feature of NN potentials that has not been thoroughly explored is the transferability of a model from the gas phase to clusters to the condensed phase. There has been some success in going from smaller to medium-sized clusters (100), but whether these potentials can describe the condensed phase remains an open question. NNs must also respect all the physical symmetries of the potential. Models like DPMD (101), SchNet (102), and NewtonNet (103) succeed in doing so and, as a result, are more data efficient and accurate than earlier networks.

Note that we have labeled ML potentials as having either an implicit or an explicit description of many-body effects. The vast majority of popular NN force fields rely on the use of many-body features of the atomic coordinates to describe the local environment of each atom, which means that many-body effects are included implicitly (for a discussion of many popular types of features and the relationships between them, see Reference 104). On the other hand, an e-MB force field can be produced by replacing the PIPs (as described above) with any type of NN (105). Either approach can be very successful, so the best may turn out to be the one that is easiest to produce and is most transferable. We anticipate that NN potentials will continue to become more popular, especially with the aid of well-developed packages for training NNs that automatically allow the potential to be evaluated on a GPU.

3.4. Effective Two-Body Force Fields

Our main focus in this review is the discussion of polarizable (many-body) force fields, which aim to reproduce the quantum mechanical PES via model representations of polarization and the quantum effects discussed in Section 1.2. However, focusing only on many-body force fields ignores the large class of effective two-body potentials. These force fields incorporate only pairwise-additive interactions, meaning that the potential energy function is cast as

$$E_{\text{tot}} = \sum_i E_i + \sum_{i<j} \Delta E_{ij}^{\text{eff}}. \qquad 11.$$

In Equation 11, the one-body term incorporates the deformation energy for each fragment, i, and the effective two-body potential, $\Delta E_{ij}^{\text{eff}}$, describes all intermolecular interactions.

Historically, effective two-body force fields have prioritized speed and simplicity over quantitative accuracy. For this reason, the form of $\Delta E_{ij}^{\text{eff}}$ is often some combination of Lennard–Jones interactions and Coulomb interactions between fixed charges. Additionally, many models are rigid, meaning that the one-body term is zero. Common models of this type include the TIP4P (106) and SPC/E (107) potentials. Almost every effective two-body force field is parameterized to reproduce some experimental property and possibly reference quantum mechanical energies. The use of experimental data means that polarization and zero-point energy are implicitly included in the model in some average way. In other words, the potential energies with these models reproduce not the Born–Oppenheimer PES but rather an effective free energy surface. This results in effective two-body potentials having inconsistent accuracy across the phase diagram of a system. These assumptions are quite severe, and it is virtually impossible to know a priori how large an effect they will have in any particular simulation.

An interesting approach is the development of effective two-body models that explicitly incorporate higher-body terms of the MBE with the goal of reproducing the Born–Oppenheimer PES. Two different methods of calculating an effective two-body potential from ab initio calculations have recently been introduced (108). The aim of these approaches is to guarantee that the effective two-body potential exactly incorporates interactions up to some order k of the MBE ($k > 2$). It is already known that effective two-body models of water wrap many-body effects into the two-body term by artificially enhancing the dipole moment. Therefore, formalizing this process seems a natural step to take. In particular, Reference 108 describes how one can construct an effective two-body potential that exactly incorporates the energy up to the k-body term by appropriately scaling the energy of an environment surrounding a dimer. As far as we are aware, this is the only known approach that allows pairwise-additive potentials to reproduce the quantum mechanical PES rather than some effective free energy surface.

COMPARISON OF MANY-BODY FORCE FIELDS			
	Two-body	Implicit	Explicit
Accuracy compared with QM	Inconsistent	Good	Very good
Speed of evaluation	Fastest	Fast	Slow
PES	Effective	Born–Oppenheimer	Born–Oppenheimer
Dipole surface	Explicitly included	Explicitly included	Included manually
Simulation methodology	Classical	Classical Quantum	Classical Quantum
Ease of construction	Easiest	Moderate	Difficult
Transferability across chemical environments	Restricted by parameterization	Transferable	Transferable

Figure 3

A qualitative comparison between the three classes of force fields discussed in this review, focusing on various desirable properties. Abbreviations: PES, potential energy surface; QM, quantum mechanics.

3.5. Comparison of the Different Classes of Force Fields

There are inevitable drawbacks to each of the three classes of force field described above. **Figure 3** depicts some important properties of a force field and describes how each class represents those properties. This figure represents a visual overview of the discussion (17) we initiated in 2002, at the beginning of the development of the TTM potentials. We rely on this figure to facilitate a comparison between the three force field classes.

3.5.1. Accuracy and speed. Arguably, the two most important properties of any force field are accuracy and speed. The order of importance may depend on the use case. Aiming for accuracy, e-MB- and PIP-based potentials fit the short-range part of the two- and three-body potentials while using a polarizable potential to capture all of the long-range and higher-order terms. Because accurate quantum mechanical references are included in the potential, an e-MB potential can be very accurate. Additionally, because e-MB force fields usually use a distributed multipole potential to provide the long-range part of the potential, they are inevitably slower than other polarizable force fields. Furthermore, in the context of MD, distributed multipole potentials can reuse the induced dipoles from a previous step, greatly decreasing the number of iterations to converge new induced dipoles, whereas explicit force fields have to redo all calculations for every new structure. Both i-MB and e-MB force fields prioritize accuracy over speed, although the former add a bit more weight on speed than the latter. Effective two-body potentials take a very different approach and prioritize speed over accuracy. Generally speaking, two-body force fields operate under the assumption that many types of errors will be averaged out during a simulation. Therefore, by prioritizing speed, one can study large-scale emergent phenomena and/or very slow processes.

3.5.2. Potential energy surface and dipole moment surface. Both i-MB and e-MB force fields aim to accurately reproduce the Born–Oppenheimer PES. Effective two-body force fields, on the other hand, propagate dynamics not on the Born–Oppenheimer PES but instead on some kind of effective free energy surface that includes zero-point energy, temperature effects, and other phenomena. We have noted that there are new types of two-body force fields that aim to reproduce

the Born–Oppenheimer PES (108), but even these potentials will do so only for a specific region of the system's phase diagram.

Since a common use of a force field is to simulate spectroscopic experiments, the DMS associated with a force field is also important. Both two-body and i-MB force fields come packaged with a DMS. That is, the energy of a given configuration depends on the assignment of charges on the atoms, which determine the molecular dipole moment. e-MB force fields, however, do not have a notion of atomic charges associated with the short-range part of the PES, so the DMS must be manually supplied. For instance, WHBB utilizes a DMS constructed from MP2/aug-cc-pVTZ calculations, which, incidentally, are not the same level of theory used to construct the PES (75). Additionally, after the MB-Pol potential was introduced, an ab initio DMS was constructed to calculate the intensities of the infrared spectra (109). Although this does not represent a drawback, it does pose an issue that the PES and DMS in e-MB models are not internally consistent. Note that because e-MB force fields typically use distributed multipole models for the long-range part of the potential, they could use the DMS generated by this potential. However, this DMS is typically ignored because it is expected to be less accurate than the PES used. Similarly, to generate Raman spectra, a polarizability surface is necessary. Typically, force fields include not a polarizability surface (dependent on intramolecular geometry) but rather a single set of polarizabilities (independent of intramolecular geometry). However, note that polarizability surfaces dependent on bond length and bond angle have been established (110, 111).

3.5.3. Classical versus nuclear quantum simulations. A force field that is fit to electronic structure calculations will attempt to reproduce the Born–Oppenheimer PES. However, classical MD simulations do not account for the quantum nature of nuclei. Since both the i-MB and the e-MB types of force fields aim to reproduce the Born–Oppenheimer PES, they should be combined with path integral (112) rather than classical (Newtonian) simulations. On the other hand, using an effective two-body body force field with quantized nuclei would almost certainly result in double-counting of zero-point effects (see **Figure 3** and the discussion in Reference 17).

3.5.4. Parameterization and transferability. While opinions may vary on this point, perhaps the most useful feature of effective two-body force fields is that they can be used for any system for which atom-specific parameters have been defined (113, 114). However, due to the empirical nature of this parameterization, effective two-body force fields may provide physically meaningful results only near the points on the phase diagram that they were parameterized to reproduce. In this sense, two-body force fields are often transferable across chemical space while being restricted in their transferability in chemical environments.

At the other extreme, e-MB models often have to be reconstructed from the ground up for any system of interest. Or, in the best-case scenario, new one-, two-, and three-body terms need to be fitted to facilitate the modeling of heterogeneous systems. This is the case, for example, for the CH_4–H_2O (83), CO_2–H_2O (84), and H_2–H_2O (81, 82) e-MB potentials introduced above. These models are designed to be applicable to the entire phase diagram of a certain chemical system, but they are not transferable across chemical space at all. That is, even though all of the element types are fixed, one could not easily take an e-MB model of H_2O and make an e-MB model of H_2O_2.

Since i-MB models typically have atom-specific parameters, they are easier to reparameterize than e-MB models and hence are more transferable across chemical space. Additionally, because i-MB models also aim to reproduce the Born–Oppenheimer PES, they are applicable to any region of the phase diagram. In this sense, i-MB models strike a balance between transferability across chemical environments and transferability across chemical space.

4. FRAGMENT-BASED AB INITIO MOLECULAR DYNAMICS

Thus far, we have discussed approaches for reproducing the many-body quantum mechanical interactions between molecules classically, that is, without having to describe actual electronic interactions. We have presented classical techniques via distributed multipoles as well as data-driven techniques utilizing PIPs or NNs. It is natural to wonder, however, if there is a way to preserve the transferability and accuracy of electronic structure without resorting to fitting a force field. One such idea is to leverage the MBE in the same manner as e-MB force fields with the exception of calculating the many-body terms as needed. This approach is frequently referred to as fragment-based quantum mechanics. There are several methods of fragmenting a system, doing calculations on these fragments, then approximately reconstructing properties of the total system from the ones of the individual pieces. Because these methods have been reviewed elsewhere (115), here we only summarize new approaches that use fragment-based quantum calculations to propagate ab initio molecular dynamics (AIMD) simulations.

4.1. Background and Theory

The implementation of fragment-based quantum mechanical approaches is similar to that of Born–Oppenheimer MD (116). **Figure 4** shows a simplified schematic of the main steps in a fragment-based MD simulation, which we refer to as MBE-MD. In ordinary MD, the main loop consists of calculating the forces on all atoms, then integrating the relevant equations of motion forward in time. MBE-MD modifies how the forces are evaluated by splitting the total system into collections of one-body (monomers), two-body (dimers), three-body (trimers), and higher

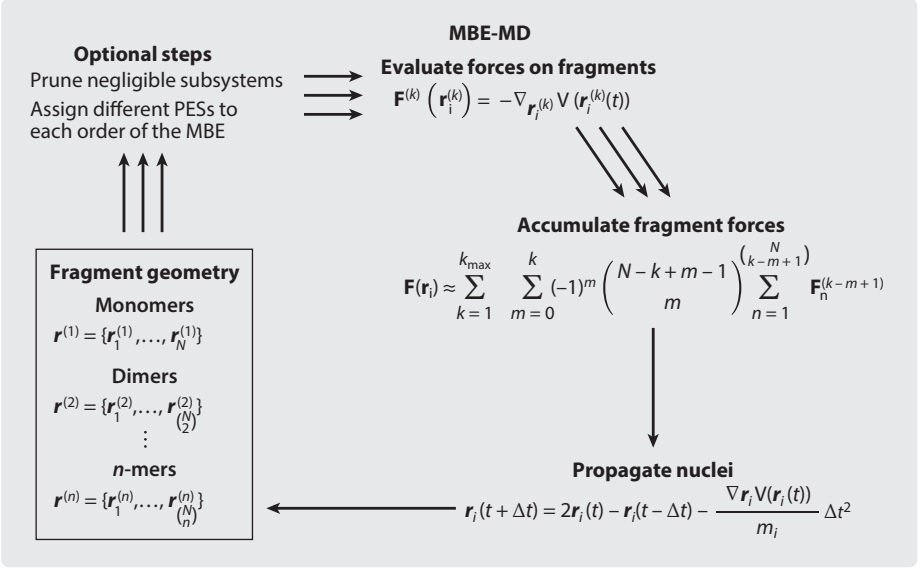

Figure 4

A schematic showing the main steps in a fragment-based MBE-MD simulation. Fragmenting the system into the relevant subsystems is typically quite simple because the fragments are chosen ahead of time. Calculations are performed on all fragments, and the results are recombined to approximate the total atomic forces. Integration of the nuclei is carried out in exactly the same way as in ordinary MD. One can also optionally prune certain subsystems or choose to use a less expensive force field on weakly interacting fragments. Abbreviations: MBE, many-body expansion; MD, molecular dynamics; PES, potential energy surface.

sets of fragments. The forces on these collections of fragments are then calculated and combined to approximate the total forces on each atom. Integration of the equations of motion is the same in MBE-MD as in ordinary MD. One could even use an extended Lagrangian approach as in Car–Parrinello MD without much difficulty (117).

4.2. Making Many-Body Expansion–Molecular Dynamics Simulations Practical: Algorithmic and Software Improvements

While the MBE is known to improve on the scaling of highly correlated electronic structure methods (47), many algorithmic and software improvements are required to make MBE-MD practical. Algorithmically, one would like to take advantage of the fact that many fragments will make a negligible contribution to the total energy, especially trimers and tetramers with far-separated monomers, which contribute primarily through short-range polarization interactions. Incorporation of protocols for pruning fragments with negligible contributions will need to be incorporated into MBE-MD. Both distance-based and energy-based pruning schemes have been proposed for use in the MBE, the latter of which seems particularly promising (118, 119). Another technique which could be useful in MBE-MD is to embed fragments in an environment with the goal of speeding up the convergence of the MBE. Among the many ways of incorporating embedding in the MBE, perhaps the most straightforward and popular one is to use electrostatically embedded charges (120). This technique effectively folds some of the higher-order terms into the lower-order ones, speeding up the convergence of the MBE. Electrostatically embedded AIMD simulations have been carried out for liquid water with the CCD electronic structure method (121), although this method includes only interactions up to the two-body term. Another path that could be pursued is to take advantage of the composability of the MBE. That is, because three-body and higher terms have very small contributions from electron correlation (48, 49), one can use Hartree–Fock or even a many-body force field (if a suitable force field exists) to save tremendous amounts of time on the calculation of higher-body terms. We believe that it is possible for MBE-MD simulations to become a common form of simulation utilizing high-scaling electronic structure methods, such as MP2 and even CCSD(T), just as Car–Parrinello MD made DFT-based AIMD practical. In order for that to happen, however, all of these algorithmic improvements will likely have to be utilized. Each of the improvements described above introduces a well-controlled and well-understood approximation. Utilizing all of these techniques simultaneously, while also managing a very large number of electronic structure calculations, makes for a complex distributed computing problem, which will require the development of new software tailored to tackle this exact problem.

4.3. Making Many-Body Expansion–Molecular Dynamics Simulations Practical: Software Improvements

Currently, most electronic structure packages aim to provide the best performance possible for a single molecular system in such a way that the code scales with the system size. A software package that is well suited for the MBE, however, does not need to scale with the size of a single system but rather needs to efficiently manage resources to do many small calculations simultaneously. Additional complexities arise from the fact that each fragment of a particular size may not take the same amount of time to finish, so the computational resources need to be managed dynamically. Not many electronic structure packages have the capability to dynamically swap threads in and out of calculations, let alone distribute a calculation on hundreds or even thousands of systems at the same time. The development of efficient scheduling and scripting language software that makes calls to an electronic structure package can address this problem by controlling the resources to maximize efficiency and minimize undesirable overheads. In our opinion, the ideal

computing environment for the development of MBE-MD software would include a mix of CPUs and GPUs with the capability of dynamically assigning resources depending on the load, such as in the TeraChem (122, 123) software.

5. CONCLUSIONS AND OUTLOOK

We have described the main techniques for analyzing intermolecular interactions and constructing many-body force fields, emphasizing the role of these techniques in constructing highly accurate force fields. In particular, e-MB force fields which are fitted to results of electronic structure calculations have proven to be extremely successful. However, new force fields based on a distributed multipole expansion rival the accuracy of e-MB force fields while providing the efficiency and ease of construction of i-MB ones. These new distributed multipole potentials will need to rely heavily on mimicking the terms from interaction analysis techniques to achieve quantitatively accurate PESs.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The writing of this review was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, Chemical Sciences, Geosciences and Biosciences Division at Pacific Northwest National Laboratory (PNNL). It was jointly supported by the Center for Scalable and Predictive Methods for Excitations and Correlated Phenomena, as part of the Computational Chemical Sciences program, and by the Condensed Phase and Interfacial Molecular Science program at PNNL. Battelle operates PNNL for the US Department of Energy. This research used computer resources provided by the National Energy Research Scientific Computing Center, which is supported by the US Department of Energy, Office of Science, under contract DE-AC02-05CH11231.

LITERATURE CITED

1. Nilsson A, Pettersson LGM. 2015. The structural origin of anomalous properties of liquid water. *Nat. Commun.* 6:8998
2. Russo J, Akahane K, Tanaka H. 2018. Water-like anomalies as a function of tetrahedrality. *PNAS* 115:E3333–41
3. Poole PH, Sciortino F, Essmann U, Stanley HE. 1992. Phase behaviour of metastable water. *Nature* 360:324–28
4. Sellberg JA, Huang C, McQueen TA, Loh ND, Laksmono H, et al. 2014. Ultrafast X-ray probing of water structure below the homogeneous ice nucleation temperature. *Nature* 510:381–84
5. Kim KH, Spah A, Pathak H, Perakis F, Mariedahl D, et al. 2017. Maxima in the thermodynamic response and correlation functions of deeply supercooled water. *Science* 358:1589–93
6. Niskanen J, Fondell M, Sahle CJ, Eckert S, Jay RM, et al. 2019. Compatibility of quantitative X-ray spectroscopy with continuous distribution models of water at ambient conditions. *PNAS* 116:4058–63
7. Smolentsev N, Smit WJ, Bakker HJ, Roke S. 2017. The interfacial structure of water droplets in a hydrophobic liquid. *Nat. Commun.* 8:15548
8. Chaplin M. 2020. *Anomalous properties of water*. Work. Pap., London South Bank Univ., London. **https://water.lsbu.ac.uk/water/water_anomalies.html**
9. Marx D, Hutter J. 2009. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge, UK: Cambridge Univ. Press

10. Iftimie R, Minary P, Tuckerman ME. 2005. Ab initio molecular dynamics: concepts, recent developments, and future trends. *PNAS* 102:6654–59
11. Hehre WJ, Radom L, Schleyer PVR, Pople JA. 1986. *Ab Initio Molecular Orbital Theory*. New York: Wiley
12. Crittenden DL. 2013. A hierarchy of static correlation models. *J. Phys. Chem. A* 117:3852–60
13. Del Ben M, Hutter J, VandeVondele J. 2012. Second-order Møller-Plesset perturbation theory in the condensed phase: an efficient and massively parallel Gaussian and plane waves approach. *J. Chem. Theory Comput.* 8:4177–88
14. Bernal JD, Fowler RH. 1933. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J. Chem. Phys.* 1:515–48
15. Barker JA, Watts RO. 1969. Structure of water: a Monte Carlo calculation. *Chem. Phys. Lett.* 3:144–45
16. Rahman A, Stillinger FH. 1971. Molecular dynamics study of liquid water. *J. Chem. Phys.* 55:3336–59
17. Burnham CJ, Xantheas SS. 2002. Development of transferable interaction models for water. I. Prominent features of the water dimer potential energy surface. *J. Chem. Phys.* 116:1479–92
18. Burnham CJ, Xantheas SS. 2002. Development of transferable interaction models for water. IV. A flexible, all-atom polarizable potential TTM2-F based on geometry dependent charges derived from an *ab initio* monomer dipole moment surface. *J. Chem. Phys.* 116:5115–24
19. Fanourgakis GS, Xantheas SS. 2006. The flexible, polarizable, Thole-type interaction potential for water (TTM2-F) revisited. *J. Phys. Chem. A* 110:4100–6
20. Fanourgakis GS, Xantheas SS. 2008. Development of transferable interaction potentials for water. V. Extension of the flexible, polarizable, Thole-type model potential (TTM3-F, v. 3.0) to describe the vibrational spectra of water clusters and liquid water. *J. Chem. Phys.* 128:074506
21. Bowman JM, Braams BJ, Carter S, Chen C, Czakó G, et al. 2010. Ab-initio-based potential energy surfaces for complex molecules and molecular complexes. *J. Phys. Chem. Lett.* 1:1866–74
22. Bowman JM, Czakó G, Fu B. 2011. High-dimensional ab initio potential energy surfaces for reaction dynamics calculations. *Phys. Chem. Chem. Phys.* 13:8094–111
23. Paesani F, Xantheas SS, Voth GA. 2009. Infrared spectroscopy and hydrogen-bond dynamics of liquid water from centroid molecular dynamics with an ab initio–based force field. *J. Phys. Chem. B* 113:13118–30
24. Paesani F, Iuchi S, Voth GA. 2007. Quantum effects in liquid water from an ab initio–based polarizable force field. *J. Chem. Phys.* 127:074506
25. Guillot B. 2002. A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liq.* 101:219–60
26. Salem L. 1961. The forces between polyatomic molecules. II. Short-range repulsive forces. *Proc. R. Soc. A* 264:379–91
27. Hankins D, Moskowitz JW, Stillinger FH. 1970. Water molecule interactions. *J. Chem. Phys.* 53:4544–54
28. Del Bene JE, Pople JA. 1973. Theory of molecular interactions. III. A comparison of studies of H_2O polymers using different molecular-orbital basis sets. *J. Chem. Phys.* 58:3605–8
29. Clementi E, Kołos W, Lie GC, Ranghino G. 1980. Nonadditivity of interaction in water trimers. *Int. J. Quantum Chem.* 17:377–98
30. Xantheas SS. 1994. Ab initio studies of cyclic water clusters $(H_2O)_n$, $n = 1$–6. II. Analysis of many-body interactions. *J. Chem. Phys.* 100:7523–34
31. Richard RM, Lao KU, Herbert JM. 2014. Understanding the many-body expansion for large systems. I. Precision considerations. *J. Chem. Phys.* 141:014108
32. Richard RM, Herbert JM. 2012. A generalized many-body expansion and a unified view of fragment-based methods in electronic structure theory. *J. Chem. Phys.* 137:064113
33. London F. 1930. Zur Theorie und Systematik der Molekularkräfte. *Z. Phys.* 63:245–79
34. van der Waals JD. 1873. *Over de Continuiteit van den Gas- en Vloeistoftoestand*, Vol. 1. Leiden, Neth.: Sijthoff
35. Axilrod BM, Teller E. 1943. Interaction of the van der Waals type between three atoms. *J. Chem. Phys.* 11:299–300
36. Patkowski K. 2020. Recent developments in symmetry-adapted perturbation theory. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 10:e1452

37. Rackers JA, Ponder JW. 2019. Classical Pauli repulsion: an anisotropic, atomic multipole model. *J. Chem. Phys.* 150:084104
38. Garcia J, Podeszwa R, Szalewicz K. 2020. SAPT codes for calculations of intermolecular interaction energies. *J. Chem. Phys.* 152:184109
39. Szalewicz K, Bukowski R, Jeziorski B. 2005. On the importance of many-body forces in clusters and condensed phase. In *Theory and Applications of Computational Chemistry*, ed. C Dykstra, G Frenking, K Kim, G Scuseria, pp. 919–62. Amsterdam: Elsevier
40. Lotrich VF, Szalewicz K. 2000. Perturbation theory of three-body exchange nonadditivity and application to helium trimer. *J. Chem. Phys.* 112:112–21
41. Lotrich VF, Szalewicz K. 1997. Three-body contribution to binding energy of solid argon and analysis of crystal structure. *Phys. Rev. Lett.* 79:1301–4
42. Lotrich VF, Jankowski P, Szalewicz K. 1998. Symmetry-adapted perturbation theory of three-body nonadditivity in the Ar_2HF trimer. *J. Chem. Phys.* 108:4725–38
43. Podeszwa R, Szalewicz K. 2007. Three-body symmetry-adapted perturbation theory based on Kohn-Sham description of the monomers. *J. Chem. Phys.* 126:194101
44. Podeszwa R, Rice BM, Szalewicz K. 2008. Predicting structure of molecular crystals from first principles. *Phys. Rev. Lett.* 101:115503
45. Medders GR, Paesani F. 2013. Many-body convergence of the electrostatic properties of water. *J. Chem. Theory Comput.* 9:4844–52
46. Demerdash O, Head-Gordon T. 2016. Convergence of the many-body expansion for energy and forces for classical polarizable models in the condensed phase. *J. Chem. Theory Comput.* 12:3884–93
47. Heindel JP, Xantheas SS. 2021. Molecular dynamics driven by the many-body expansion (MBE-MD). *J. Chem. Theory Comput.* 17:7341–52
48. Heindel JP, Xantheas SS. 2020. The many-body expansion for aqueous systems revisited. I. Water–water interactions. *J. Chem. Theory Comput.* 16:6843–55
49. Heindel JP, Xantheas SS. 2021. The many-body expansion for aqueous systems revisited. II. Alkali metal and halide ion–water interactions. *J. Chem. Theory Comput.* 17:2200–16
50. Herman KM, Heindel JP, Xantheas SS. 2021. The many-body expansion for aqueous systems revisited. III. Hofmeister ion–water interactions. *Phys. Chem. Chem. Phys.* 23:11196–210
51. Mao Y, Loipersberger M, Horn PR, Das A, Demerdash O, et al. 2021. From intermolecular interaction energies and observable shifts to component contributions and back again: a tale of variational energy decomposition analysis. *Annu. Rev. Phys. Chem.* 72:641–66
52. Khaliullin RZ, Cobar EA, Lochan RC, Bell AT, Head-Gordon M. 2007. Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals. *J. Phys. Chem. A* 111:8753–65
53. Horn PR, Mao Y, Head-Gordon M. 2016. Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals. *Phys. Chem. Chem. Phys.* 18:23067–79
54. Thirman J, Head-Gordon M. 2015. An energy decomposition analysis for second-order Møller–Plesset perturbation theory based on absolutely localized molecular orbitals. *J. Chem. Phys.* 143:084124
55. Loipersberger M, Lee J, Mao Y, Das AK, Ikeda K, et al. 2019. Energy decomposition analysis for interactions of radicals: theory and implementation at the MP2 level with application to hydration of halogenated benzene cations and complexes between CO_2^- and pyridine and imidazole. *J. Phys. Chem. A* 123:9621–33
56. Andrés J, Ayers PW, Boto RA, Carbó-Dorca R, Chermette H, et al. 2019. Nine questions on energy decomposition analysis. *J. Comput. Chem.* 40:2248–83
57. Stone A. 2013. *The Theory of Intermolecular Forces*. Oxford, UK: Oxford Univ. Press
58. Herbert JM. 2021. Neat, simple, and wrong: debunking electrostatic fallacies regarding noncovalent interactions. *J. Phys. Chem. A* 125:7125–37
59. Partridge H, Schwenke DW. 1997. The determination of an accurate isotope dependent potential energy surface for water from extensive ab initio calculations and experimental data. *J. Chem. Phys.* 106:4618–39
60. Arismendi-Arrieta DJ, Riera M, Bajaj P, Prosmiti R, Paesani F. 2016. i-TTM model for ab initio–based ion–water interaction potentials. 1. Halide–water potential energy functions. *J. Phys. Chem. B* 120:1822–32

61. Riera M, Götz AW, Paesani F. 2016. The i-TTM model for ab initio–based ion–water interaction potentials. II. Alkali metal ion–water potential energy functions. *Phys. Chem. Chem. Phys.* 18:30334–43
62. Das AK, Urban L, Leven I, Loipersberger M, Aldossary A, et al. 2019. Development of an advanced force field for water using variational energy decomposition analysis. *J. Chem. Theory Comput.* 15:5001–13
63. Das AK, Demerdash ON, Head-Gordon T. 2018. Improvements to the AMOEBA force field by introducing anisotropic atomic polarizability of the water molecule. *J. Chem. Theory Comput.* 14:6722–33
64. Mao Y, Demerdash O, Head-Gordon M, Head-Gordon T. 2016. Assessing ion–water interactions in the AMOEBA force field using energy decomposition analysis of electronic structure calculations. *J. Chem. Theory Comput.* 12:5422–37
65. Demerdash O, Mao Y, Liu T, Head-Gordon M, Head-Gordon T. 2017. Assessing many-body contributions to intermolecular interactions of the AMOEBA force field using energy decomposition analysis of electronic structure calculations. *J. Chem. Phys.* 147:161721
66. Mobley DL, Wymer KL, Lim NM, Guthrie JP. 2014. Blind prediction of solvation free energies from the SAMPL4 challenge. *J. Comput. Aided Mol.* 28:135–50
67. Das AK, Liu M, Head-Gordon T. 2022. Development of a many-body force field for aqueous alkali metal and halogen ions: an energy decomposition analysis guided approach. *J. Chem. Theory Comput.* 18:953–67
68. Lemkul JA, Huang J, Roux B, MacKerell AD Jr. 2016. An empirical polarizable force field based on the classical Drude oscillator model: development history and recent applications. *Chem. Rev.* 116:4983–5013
69. Baker CM. 2015. Polarizable force fields for molecular dynamics simulations of biomolecules. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 5:241–54
70. Patel S, Brooks CL III. 2006. Fluctuating charge force fields: recent developments and applications from small molecules to macromolecular biological systems. *Mol. Simul.* 32:231–49
71. Rick SW, Stuart SJ, Berne BJ. 1994. Dynamical fluctuating charge force fields: application to liquid water. *J. Chem. Phys.* 101:6141–56
72. Braams BJ, Bowman JM. 2009. Permutationally invariant potential energy surfaces in high dimensionality. *Int. Rev. Phys. Chem.* 28:577–606
73. Xie Z, Bowman JM. 2010. Permutationally invariant polynomial basis for molecular energy surface fitting via monomial symmetrization. *J. Chem. Theory Comput.* 6:26–34
74. Qu C, Yu Q, Bowman JM. 2018. Permutationally invariant potential energy surfaces. *Annu. Rev. Phys. Chem.* 69:151–75
75. Wang Y, Huang X, Shepler BC, Braams BJ, Bowman JM. 2011. Flexible, ab initio potential, and dipole moment surfaces for water. I. Tests and applications for clusters up to the 22-mer. *J. Chem. Phys.* 134:094509
76. Babin V, Leforestier C, Paesani F. 2013. Development of a "first principles" water potential with flexible monomers: dimer potential energy surface, VRT spectrum, and second virial coefficient. *J. Chem. Theory Comput.* 9:5395–403
77. Burnham CJ, Anick DJ, Mankoo PK, Reiter GF. 2008. The vibrational proton potential in bulk liquid water and ice. *J. Chem. Phys.* 128:154519
78. Mallory JD, Mandelshtam VA. 2016. Diffusion Monte Carlo studies of MB-Pol $(H_2O)_{2-6}$ and $(D_2O)_{2-6}$ clusters: structures and binding energies. *J. Chem. Phys.* 145:064308
79. Reddy SK, Straight SC, Bajaj P, Huy Pham C, Riera M, et al. 2016. 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
80. Reddy SK, Moberg DR, Straight SC, Paesani F. 2017. Temperature-dependent vibrational spectra and structure of liquid water from classical and quantum simulations with the MB-Pol potential energy function. *J. Chem. Phys.* 147:244504
81. Homayoon Z, Conte R, Qu C, Bowman JM. 2015. Full-dimensional, high-level ab initio potential energy surfaces for $H_2(H_2O)$ and $H_2(H_2O)_2$ with application to hydrogen clathrate hydrates. *J. Chem. Phys.* 143:084302

82. Qu C, Bowman JM. 2019. Assessing the importance of the H_2–H_2O–H_2O three-body interaction on the vibrational frequency shift of H_2 in the sII clathrate hydrate and comparison with experiment. *J. Phys. Chem. A* 123:329–35
83. Riera M, Hirales A, Ghosh R, Paesani F. 2020. Data-driven many-body models with chemical accuracy for CH_4/H_2O mixtures. *J. Phys. Chem. B* 124:11207–21
84. Riera M, Yeh EP, Paesani F. 2020. Data-driven many-body models for molecular fluids: CO_2/H_2O mixtures as a case study. *J. Chem. Theory Comput.* 16:2246–57
85. Bajaj P, Götz AW, Paesani F. 2016. Toward chemical accuracy in the description of ion–water interactions through many-body representations. I. Halide–water dimer potential energy surfaces. *J. Chem. Theory Comput.* 12:2698–705
86. Riera M, Mardirossian N, Bajaj P, Götz AW, Paesani F. 2017. 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
87. Bull-Vulpe EF, Riera M, Götz AW, Paesani F. 2021. MB-Fit: software infrastructure for data-driven many-body potential energy functions. *J. Chem. Phys.* 155:124801
88. Nandi A, Qu C, Bowman JM. 2019. Using gradients in permutationally invariant polynomial potential fitting: a demonstration for CH_4 using as few as 100 configurations. *J. Chem. Theory Comput.* 15:2826–35
89. Nandi A, Qu C, Houston PL, Conte R, Bowman JM. 2021. Δ-machine learning for potential energy surfaces: a PIP approach to bring a DFT-based PES to CCSD(T) level of theory. *J. Chem. Phys.* 154:051102
90. Zhai Y, Caruso A, Gao S, Paesani F. 2020. Active learning of many-body configuration space: application to the Cs^+–water MB-nrg potential energy function as a case study. *J. Chem. Phys.* 152:144103
91. Heindel JP, Yu Q, Bowman JM, Xantheas SS. 2018. Benchmark electronic structure calculations for $H_3O^+(H_2O)_n$, $n = 0$–5, clusters and tests of an existing 1,2,3-body potential energy surface with a new 4-body correction. *J. Chem. Theory Comput.* 14:4553–66
92. Headrick JM, Diken EG, Walters RS, Hammer NI, Christie RA, et al. 2005. Spectral signatures of hydrated proton vibrations in water clusters. *Science* 308:1765–69
93. Góra U, Cencek W, Podeszwa R, van der Avoird A, Szalewicz K. 2014. Predictions for water clusters from a first-principles two- and three-body force field. *J. Chem. Phys.* 140:194101
94. Behler J. 2016. Perspective: machine learning potentials for atomistic simulations. *J. Chem. Phys.* 145:170901
95. Noé F, Tkatchenko A, Müller KR, Clementi C. 2020. Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.* 71:361–90
96. Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. 2018. Machine learning for molecular and materials science. *Nature* 559:547–55
97. Behler J. 2021. Four generations of high-dimensional neural network potentials. *Chem. Rev.* 121:10037–72
98. Manzhos S, Carrington T Jr. 2020. Neural network potential energy surfaces for small molecules and reactions. *Chem. Rev.* 121:10187–217
99. Schran C, Behler J, Marx D. 2019. Automated fitting of neural network potentials at coupled cluster accuracy: protonated water clusters as testing ground. *J. Chem. Theory Comput.* 16:88–99
100. Schran C, Brieuc F, Marx D. 2021. Transferability of machine learning potentials: protonated water neural network potential applied to the protonated water hexamer. *J. Chem. Phys.* 154:051101
101. Zhang L, Han J, Wang H, Car R, E W. 2018. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* 120:143001
102. Schütt KT, Sauceda HE, Kindermans PJ, Tkatchenko A, Müller KR. 2018. SchNet—a deep learning architecture for molecules and materials. *J. Chem. Phys.* 148:241722
103. Haghighatlari M, Li J, Guan X, Zhang O, Das A, et al. 2022. NewtonNet: a Newtonian message passing network for deep learning of interatomic potentials and forces. *Digit. Discov.* 1:333–43
104. Goscinski A, Fraux G, Imbalzano G, Ceriotti M. 2021. The role of feature space in atomistic learning. *Mach. Learn. Sci. Technol.* 2:025028

105. Nguyen TT, Székely E, Imbalzano G, Behler J, Csányi G, et al. 2018. Comparison of permutationally invariant polynomials, neural networks, and Gaussian approximation potentials in representing water interactions through many-body expansions. *J. Chem. Phys.* 148:241725
106. Abascal JL, Vega C. 2005. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* 123:234505
107. Berendsen HJC, Grigera JR, Straatsma TP. 1987. The missing term in effective pair potentials. *J. Phys. Chem.* 91:6269–71
108. Mackie C, Zech A, Head-Gordon M. 2021. Effective two-body interactions. *J. Phys. Chem. A* 125:7750–58
109. Medders GR, Paesani F. 2015. Infrared and Raman spectroscopy of liquid water through "first-principles" many-body molecular dynamics. *J. Chem. Theory Comput.* 11:1145–54
110. Loboda O, Ingrosso F, Ruiz-López MF, Reis H, Millot C. 2016. Dipole and quadrupole polarizabilities of the water molecule as a function of geometry. *J. Comput. Chem.* 37:2125–32
111. Lao KU, Jia J, Maitra R, Distasio RA Jr. 2018. On the geometric dependence of the molecular dipole polarizability in water: a benchmark study of higher-order electron correlation, basis set incompleteness error, core electron effects, and zero-point vibrational contributions. *J. Chem. Phys.* 149:204303
112. Feynman RP, Hibbs AR, Styer DF. 2010. *Quantum Mechanics and Path Integrals*. Mineola, NY: Dover
113. Damm W, Frontera A, Tirado-Rives J, Jorgensen WL. 1997. OPLS all-atom force field for carbohydrates. *J. Comput. Chem.* 18:1955–70
114. Rappe AK, Casewit CJ, Colwell KS, Goddard WA, Skiff WM. 1992. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* 114:10024–35
115. Herbert JM. 2019. Fantasy versus reality in fragment-based quantum chemistry. *J. Chem. Phys.* 151:170901
116. Barnett RN, Landman U. 1993. Born-Oppenheimer molecular-dynamics simulations of finite systems: structure and dynamics of $(H_2O)_2$. *Phys. Rev. B* 48:2081–97
117. Car R, Parrinello M. 1985. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* 55:2471–74
118. Cui J, Liu H, Jordan KD. 2006. Theoretical characterization of the $(H_2O)_{21}$ cluster: application of an *n*-body decomposition procedure. *J. Phys. Chem. B* 110:18872–78
119. Liu KY, Herbert JM. 2019. Energy-screened many-body expansion: a practical yet accurate fragmentation method for quantum chemistry. *J. Chem. Theory Comput.* 16:475–87
120. Dahlke EE, Truhlar DG. 2007. Electrostatically embedded many-body expansion for large systems, with applications to water clusters. *J. Chem. Theory Comput.* 3:46–53
121. Liu J, He X, Zhang JZ, Qi LW. 2018. Hydrogen-bond structure dynamics in bulk water: insights from ab initio simulations with coupled cluster theory. *Chem. Sci.* 9:2065–73
122. Seritan S, Thompson K, Martínez TJ. 2020. TeraChem cloud: a high-performance computing service for scalable distributed GPU-accelerated electronic structure calculations. *J. Chem. Inf. Model.* 60:2126–37
123. Seritan S, Bannwarth C, Fales BS, Hohenstein EG, Kokkila-Schumacher SIL, et al. 2020. TeraChem: accelerating electronic structure and ab initio molecular dynamics with graphical processing units. *J. Chem. Phys.* 152:224110

Annual Review of Physical Chemistry
Volume 74, 2023

Contents

Remembering the Work of Phillip L. Geissler: A Coda to His Scientific Trajectory
Gregory R. Bowman, Stephen J. Cox, Christoph Dellago, Kateri H. DuBay, Joel D. Eaves, Daniel A. Fletcher, Layne B. Frechette, Michael Grünwald, Katherine Klymko, JiYeon Ku, Ahmad K. Omar, Eran Rabani, David R. Reichman, Julia R. Rogers, Andreana M. Rosnik, Grant M. Rotskoff, Anna R. Schneider, Nadine Schwierz, David A. Sivak, Suriyanarayanan Vaikuntanathan, Stephen Whitelam, and Asaph Widmer-Cooper 1

Gas-Phase Computational Spectroscopy: The Challenge of the Molecular Bricks of Life
Vincenzo Barone and Cristina Puzzarini 29

Magneto-Optical Properties of Noble Metal Nanostructures
Juniper Foxley and Kenneth L. Knappenberger Jr. 53

Ultrafast X-Ray Probes of Elementary Molecular Events
Daniel Keefer, Stefano M. Cavaletto, Jérémy R. Rouxel, Marco Garavelli, Haiwang Yong, and Shaul Mukamel 73

Spectroscopic Studies of Clusters of Atmospheric Relevance
Nicoline C. Frederiks, Annapoorani Hariharan, and Christopher J. Johnson 99

Photoacid Dynamics in the Green Fluorescent Protein
Jasper J. van Thor and Paul M. Champion 123

Photochemical Upconversion
Jiale Feng, Jessica Alves, Damon M. de Clercq, and Timothy W. Schmidt 145

Adsorption at Nanoconfined Solid–Water Interfaces
Anastasia G. Ilgen, Kevin Leung, Louise J. Criscenti, and Jeffery A. Greathouse 169

The Predictive Power of Exact Constraints and Appropriate Norms in Density Functional Theory
Aaron D. Kaplan, Mel Levy, and John P. Perdew 193

Modeling Anharmonic Effects in the Vibrational Spectra of High-Frequency Modes
Edwin L. Sibert III 219

Studies of Local DNA Backbone Conformation and Conformational Disorder Using Site-Specific Exciton-Coupled Dimer Probe Spectroscopy
Andrew H. Marcus, Dylan Heussman, Jack Maurer, Claire S. Albrecht, Patrick Herbert, and Peter H. von Hippel 245

In Situ Measurement of Evolving Excited-State Dynamics During Deposition and Processing of Organic Films by Single-Shot Transient Absorption
Zachary S. Walbrun and Cathy Y. Wong 267

Toward Ab Initio Reaction Discovery Using the Artificial Force Induced Reaction Method
Satoshi Maeda, Yu Harabuchi, Hiroki Hayashi, and Tsuyoshi Mita 287

Interactive Quantum Chemistry Enabled by Machine Learning, Graphical Processing Units, and Cloud Computing
Umberto Raucci, Hayley Weir, Sukolsak Sakshuwong, Stefan Seritan, Colton B. Hicks, Fabio Vannucci, Francesco Rea, and Todd J. Martínez 313

Many-Body Effects in Aqueous Systems: Synergies Between Interaction Analysis Techniques and Force Field Development
Joseph P. Heindel, Kristina M. Herman, and Sotiris S. Xantheas 337

Surface-Mediated Formation of Stable Glasses
Peng Luo and Zahra Fakhraai 361

3D Super-Resolution Fluorescence Imaging of Microgels
Oleksii Nevskyi and Dominik Wöll 391

Photodarkening, Photobrightening, and the Role of Color Centers in Emerging Applications of Lanthanide-Based Upconverting Nanomaterials
Changhwan Lee and P. James Schuck 415

Isotope Effects and the Atmosphere
Julia M. Carlstad and Kristie A. Boering 439

The Optical Signatures of Stochastic Processes in Many-Body Exciton Scattering
Hao Li, S.A. Shah, Ajay Ram Srimath Kandada, Carlos Silva, Andrei Piryatinski, and Eric R. Bittner 467

Ultrafast Dynamics of Photosynthetic Light Harvesting: Strategies for Acclimation Across Organisms
Olivia C. Fiebig, Dvir Harris, Dihao Wang, Madeline P. Hoffmann, and Gabriela S. Schlau-Cohen 493

Mechanisms of Photothermalization in Plasmonic Nanostructures: Insights into the Steady State
Shengxiang Wu and Matthew Sheldon 521

Modeling Excited States of Molecular Organic Aggregates for Optoelectronics
Federico J. Hernández and Rachel Crespo-Otero 547

Errata

An online log of corrections to *Annual Review of Physical Chemistry* articles may be found at http://www.annualreviews.org/errata/physchem