Perspective: *Ab initio* force field methods derived from quantum mechanics

Peng Xu, (iD) Emilie B. Guidez, Colleen Bertoni, et al.

View Online   **Export Citation**   **CrossMark**

ARTICLES YOU MAY BE INTERESTED IN

AIP Publishing

Perspective: *Ab initio* force field methods derived from quantum mechanics

Peng Xu,[1] Emilie B. Guidez,[2] Colleen Bertoni,[3] and Mark S. Gordon[1]

[1]*Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa 50011, USA*
[2]*Department of Chemistry, University of Colorado Denver, Denver, Colorado 80217, USA*
[3]*Argonne Leadership Computing Facility, Argonne, Illinois 60439, USA*

It is often desirable to accurately and efficiently model the behavior of large molecular systems in the condensed phase (thousands to tens of thousands of atoms) over long time scales (from nanoseconds to milliseconds). In these cases, *ab initio* methods are difficult due to the increasing computational cost with the number of electrons. A more computationally attractive alternative is to perform the simulations at the atomic level using a parameterized function to model the electronic energy. Many empirical force fields have been developed for this purpose. However, the functions that are used to model interatomic and intermolecular interactions contain many fitted parameters obtained from selected model systems, and such classical force fields cannot properly simulate important electronic effects. Furthermore, while such force fields are computationally affordable, they are not reliable when applied to systems that differ significantly from those used in their parameterization. They also cannot provide the information necessary to analyze the interactions that occur in the system, making the systematic improvement of the functional forms that are used difficult. *Ab initio* force field methods aim to combine the merits of both types of methods. The ideal *ab initio* force fields are built on first principles and require no fitted parameters. *Ab initio* force field methods surveyed in this perspective are based on fragmentation approaches and intermolecular perturbation theory. This perspective summarizes their theoretical foundation, key components in their formulation, and discusses key aspects of these methods such as accuracy and formal computational cost. The *ab initio* force fields considered here were developed for different targets, and this perspective also aims to provide a balanced presentation of their strengths and shortcomings. Finally, this perspective suggests some future directions for this actively developing area. *Published by AIP Publishing.* https://doi.org/10.1063/1.5009551

I. INTRODUCTION

A. Classical force fields

The accurate potential energy surface of a chemical system is usually obtained by approximately solving the Schrödinger equation at many points in the coordinate space of the component atoms. High-level *ab initio* methods now can attain results that are in reasonable agreement with experiments for small, isolated, molecular systems. However, the more likely scenarios an experimental chemist encounters are reactions occurring in condensed phases, either liquid or solid, or at the interface between two phases. The desire to model a bulk system, e.g., nano-scaled systems, with *ab initio* methods is often inhibited by the steep rise of the computational cost with respect to the system size, typically ranging from N^4–N^7 for the most popular methods, where N measures the system size.

Classical force field (FF) methods are designed to overcome the *ab initio* scaling problem by expressing the electronic energy as a parameterized function. Traditionally, the parameters in FF methods are fitted to either experiments or high-level *ab initio* calculations. The total energy is typically given as the sum of bonding and non-bonding interactions. Stretching (str), bending (bend), and torsion (tors) modes represent the bonding interactions. Electrostatic (el) and van der Waals

(vdW) terms describe the non-bonding interactions. These typically follow the form

$$E_{FF} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{el} + E_{cross}. \quad (1)$$

In Eq. (1), E_{cross} represents the coupling between the other five energy contributions. This term is often neglected or may only consider coupling between bonding interactions. In the literature, force fields are commonly classified in three categories based on their level of sophistication: Class I, Class II, and Class III. Class I refers to force fields with the simplest functional forms. Harmonic functions are typically used to describe bending and stretching. The non-bonding electrostatic and van der Waals contributions are often computed using a charge-charge Coulomb term and a Lennard-Jones potential, respectively. Cross terms are neglected. The computational cost is very low and therefore permits the modeling of very large systems of the order of tens of thousands of atoms, such as proteins, bulk liquids, and crystals. Examples include the early versions of the biological force fields CHARMM[1,2] and AMBER,[3–5] as well as OPLS.[6–8] Class II force fields use more sophisticated functions and include a number of cross terms. They are typically parameterized to reproduce certain molecular properties such as geometries and vibrational frequencies. Due to the increased complexity of the functional form, Class II FFs are computationally more

expensive. Examples include MM3[9–12] and CFF.[13,14] Class III force fields include polarization and hyperconjugation effects. Examples include polarizable versions of CHARMM,[15–17] AMBER,[18,19] and OPLS,[20] as well as AMOEBA[21,22] (for a review, see Refs. 23–26). Although the parameters in these force fields might be fit to the results of *ab initio* calculations, the functional forms themselves are not derived from quantum mechanical first principles. In particular, treating the bonding interactions with a classical bonding model is one feature that differentiates these force fields from *ab initio* force fields.

B. Why *ab initio* force fields?

One way of expressing the electronic energy is by deriving the functional forms from quantum mechanics-based first principles. In this perspective, an *ab initio* force field refers to one that is derived from quantum mechanical first principles and does not contain any empirically fitted parameters. Non-*ab initio* force fields have functional forms that are not derived from quantum mechanics or contain empirically fitted parameters. Fitting refers to either the optimization of the functional forms or parameters in the functional so that the predicted observables are accurate (agreeing well with experiments or high-level *ab initio* calculations). There is usually limited physical insight in a fitted functional form of a force field. An example of such a fitted potential is the Lennard-Jones potential: $U_{LJ} = {C_{12}}/{R^{12}} - {C_6}/{R^6}$, in which the choice of the repulsive term being proportional to R^{-12} is a computational convenience that is not based on theoretical arguments. On the other hand, the attractive R^{-6} term is based on a firm physical foundation, although the coefficient C_6 is typically fitted to experiments. However, as will be discussed later, this C_6 coefficient can also be obtained from more physically meaningful quantities such as dynamic dipole polarizability tensors. Like many concepts in chemistry, the distinction between an *ab initio* FF and a non-*ab initio* FF is not so clear-cut in practice since it is difficult to completely eliminate fitting in a force field. In this perspective, force fields that are completely or partially derived from first principles will be considered. Special attention will be given to the formulation of these first-principles derived contributions.

Ab initio force fields are generally computationally more expensive than classical force fields. If classical force fields can produce results that are comparable with experiments, why does one need a more expensive *ab initio* force field? First, having a solid physical/mathematical foundation provides a deeper understanding of the physics and chemistry of the systems to be investigated. For example, the main contribution to the binding energy of liquid argon is quite different from that of liquid water. Also, a more rigorous error analysis and a subsequent systematic improvement are easier to implement with *ab initio* force field methods because there are no empirically fitted parameters. Second, force fields with empirical parameters may not be reliable if the systems used in the parameterization process significantly differ from the system being investigated or if the data set used for parameterization is small. Most empirical force fields have well-defined parameters for atoms and combinations of atoms for elements in

the main group, where a large training set was used. However, the parameters for transition metals are typically not so well defined due to limited available experimental data and the complexity of transition metals. Transition metal complexes are much more complicated than organic compounds: they are more dynamic (ligand exchange), can form more bonds, and can have many different geometrical arrangements. Therefore, one should use extreme caution when modeling transition metal compounds with most empirical force fields. It should be emphasized that the Universal Force Field (UFF), developed by Rappe *et al.*,[27] for example, can be applied to any atom in the periodic table, but the UFF is only moderately accurate. For use in specialized cases, such as metal-organic frameworks, researchers have had to add additional parameters to the UFF to reach an adequate level of accuracy.[28] Overall, it is important to get the right answer for the right reason and inaccurate results should be able to provide insights for the improvement of the method. This is difficult to achieve when many empirical parameters are embedded in the potential.

C. Computation of intermolecular interactions

The common theoretical foundation of the *ab initio* force fields discussed in this perspective is that they all use intermolecular perturbation theory to compute intermolecular interaction energies at long range. Intermolecular interaction energies are clearly important in all phases of chemistry, ranging from diffusion processes in the gas phase to solvent effects to condensed phase phenomena.

One of the simplest ways to obtain intermolecular interaction energies is by the supermolecule approach, in which the intermolecular interaction energy is calculated as the difference between the energy of the total system computed at a given level of theory and the sum of the monomer energies computed at the same level. A drawback of this approach is that analyzing the interaction energy in terms of meaningful contributions is not straightforward. Several schemes have been developed for this purpose[29] such as the Kitaura-Morokuma energy decomposition analysis (EDA),[30] the block-localized wave function energy decomposition (BLW-ED) approach,[31] the absolutely localized molecular orbital EDA,[32] the natural energy decomposition analysis (NEDA),[33] the constrained space orbital variation (CSOV) for ligand-metal interactions,[34] and the reduced variational space (RVS) procedure.[35] Su and Li have also developed a basis set independent energy decomposition analysis scheme based on HF and density functional theory (DFT) methods.[36] However, the components in these energy decomposition schemes are not entirely well defined.

An alternate way to think about intermolecular interactions is to consider the fact that these interactions are often much smaller than typical chemical bonding interactions. For instance, the interaction energy between two methane molecules is ~0.5 kcal/mol,[37] whereas the C–H bond energy in methane is ~104 kcal/mol.

Such differences in orders of magnitude can have profound implications. For example, very small interaction energies usually correspond to large separations between subsystems, where exchange interactions are small. Hence, the total wave function can be approximated as a Hartree product of

subsystem wave functions, the so-called long-range approximation. A perturbative approach can then be used to compute intermolecular interactions. With the perturbative approach, it is mostly straightforward to decompose the interaction energy into meaningful contributions (for instance, electrostatic, induction, and dispersion) that can provide insights into the nature of the interaction.

D. Building *ab initio* force fields

The force fields discussed in this perspective are derived from first principles and are theoretically applicable to any system of arbitrary size and composition. Depending on the types of systems, different approaches can be employed to generate parameters. Generally, a common theme of these methods is a divide-and-conquer or fragmentation approach:[38–40] the system is fragmented into subsystems that can easily be modeled with an *ab initio* method. The properties of these subsystems such as multipole moments and static and dynamic polarizabilities can then be used as input for an *ab initio* FF. Note that in the literature, subsystems, monomers, and fragments are all used. In this perspective, the terms monomer, fragment, and subsystem are used interchangeably.

Ab initio FF methods typically fall into one of the two categories. The first category is constructed in the framework of intermolecular perturbation theory, in which the perturbation is the electrostatic interaction operator expressed as a multipole expansion. In this category, generally, the intermolecular interaction energy is calculated, not the total energy. Examples of this category are the Effective Fragment Potential (EFP) method[39,41] and the Sum of Interaction Between Fragments *Ab initio* computed (SIBFA).[42,43] Both methods compute separate components of the total interaction energy. In the latest development of these methods, the interaction energy components are Coulomb (Coul), polarization (pol), dispersion (disp), exchange-repulsion (ex-rep), and charge transfer (CT). The total interaction energy is thus given by

$$E = E^{Coul} + E^{pol} + E^{disp} + E^{ex-rep} + E^{CT}. \quad (2)$$

The first three terms can be derived from long-range perturbation theory, whereas considering the wave function overlap typically is necessary to derive the latter two terms.

In the second category of *ab initio* FF methods considered in this perspective, the total energy is expressed in a many-body expansion,

$$
\begin{aligned}
E_{total} &= \sum_I E_I + \sum_{I>J} (E_{IJ} - E_I - E_J) \\
&\quad + \left\{ \sum_{IJK} (E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J) \right. \\
&\quad \left. - (E_{KI} - E_K - E_I) - (E_{JK} - E_J - E_K) \right\} \cdots \\
&= \sum_I E_I + \sum_{I>J} \Delta E_{IJ} + \sum_{I>J>K} \Delta E_{IJK} + \cdots .
\end{aligned} \quad (3)
$$

In Eq. (3), the total energy is split into terms containing one-body, two-body, three-body, and so on contributions. Some of the interaction energy terms may be computed with intermolecular perturbation theory instead of *ab initio* methods.

For instance, the 2-body interaction energy between fragments that are farther apart than a threshold separation can be computed with perturbation theory, while the 2-body interaction energy between fragments that are closer than the threshold separation is typically computed directly with standard *ab initio* methods. A key feature of this second category is to include the non-negligible many-body polarization interactions since polarization is the dominant contribution to the many-body interaction. Examples of force fields in this category are the effective fragment molecular orbital (EFMO) method[44] and the hybrid many-body interaction (HMBI) model.[45–49] The computational saving with these approaches stems from having to treat only small subsystems and the potential for extreme parallelism of fragmentation methods.

This perspective is organized as follows: the basic features of some representative *ab initio* FF methods are described in Secs. II and III. In Sec. IV, some key aspects of *ab initio* FF methods are discussed. In Sec. V, the applications and the accuracy of these *ab initio* FF methods are summarized. Formal computational costs are briefly discussed in Sec. VI. Summaries and outlooks are given in Sec. VII.

II. *AB INITIO* FORCE FIELDS DERIVED FROM INTERMOLECULAR PERTURBATION THEORY

At sufficiently long separations, two subsystems A and B can be regarded as two fluctuating electron densities embedding two groups of positive nuclei. Thus, interactions between the two subsystems are of electrostatic origin. The electrostatic interaction between the two subsystems can be expressed with the perturbation $\hat{H}'$ to the Hamiltonians $\hat{H}^A$ and $\hat{H}^B$ of the non-interacting subsystems,

$$\hat{H} = \hat{H}^A + \hat{H}^B + \hat{H}'. \quad (4)$$

In Eq. (4), $\hat{H}'(\boldsymbol{R}) = \int \frac{\hat{\rho}^A(\boldsymbol{r})\hat{\rho}^B(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3\boldsymbol{r} d^3\boldsymbol{r}'$, where $\hat{\rho}^A(\boldsymbol{r})$ and $\hat{\rho}^B(\boldsymbol{r}')$ are the charge density operators of subsystems A and B, respectively. In the Rayleigh-Schrödinger perturbation theory framework, the first order perturbation energy between the two subsystems representing the Coulomb term of the interaction energy is given by

$$E^{(1)} = \left\langle \Psi^{(0)} \left| \hat{H}' \right| \Psi^{(0)} \right\rangle = \iint \frac{\rho^A(\boldsymbol{r}) \rho^B(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}'. \quad (5)$$

In Eq. (5), $\Psi^{(0)}$ is the ground state wave function of the unperturbed Hamiltonian ($\hat{H} = \hat{H}^A + \hat{H}^B$). $\Psi^{(0)}$ is the Hartree product of the unperturbed wave functions of subsystems A and B ($\Psi = \Psi^A \Psi^B$). $\rho^A(\boldsymbol{r})$ and $\rho^B(\boldsymbol{r})$ represent the charge distributions of subsystems A and B, respectively. A first approximation to evaluate $E^{(1)}$ would be to consider the interaction between the point charges in the respective fragments at their centers of mass, $q^A q^B / r_{AB}$. Of course, this approximation is valid only for small (~spherical) molecules at large distances[50] and does not describe the anisotropic charge distribution. A better description of the anisotropic charge distribution can be realized by a Taylor expansion of the electrostatic potential, in which familiar interactions such as charge-dipole, charge-quadrupole, dipole-dipole,

dipole-quadrupole, etc., arise naturally,

$$
\begin{aligned}
E_{AB}^{Coul} = q^A q^B T^{AB} - \sum_{\alpha}^{x,y,z} q^A \mu_\alpha^B T_\alpha^{AB} + \frac{1}{3} \sum_{\alpha,\beta}^{x,y,z} q^A \Theta_{\alpha\beta}^B T_{\alpha\beta}^{AB} \\
+ \sum_{\alpha}^{x,y,z} \mu_\alpha^A q^B T_\alpha^{AB} - \sum_{\alpha,\beta}^{x,y,z} \mu_\alpha^A \mu_\beta^B T_{\alpha\beta}^{AB} \\
+ \frac{1}{3} \sum_{\alpha,\beta,\gamma}^{x,y,z} \mu_\alpha^A \Theta_{\beta\gamma}^B T_{\alpha\beta\gamma}^{AB} + \cdots .
\end{aligned}
\tag{6}
$$

In Eq. (6), q^A and q^B represent the point charges of fragments A and B, respectively, μ_α^A and μ_α^B represent the α component of the dipoles of fragments A and B, respectively, and $\Theta_{\alpha\beta}^A$ and $\Theta_{\alpha\beta}^B$ represent the components of the quadrupoles of fragments A and B, respectively. The electrostatic tensor T is defined as

$$
T_{\alpha,\beta\ldots\nu}^{AB} = \nabla_\alpha \nabla_\beta \ldots \nabla_\nu \frac{1}{R},
\tag{7}
$$

where R is the distance between expansion sites on fragments A and B.

The multipole expansion in Eq. (6) is a classical approximation to the interaction between quantum mechanical densities. Strictly speaking, the multipole expansion will only converge if the two charge densities do not overlap, which is theoretically impossible since electron densities extend throughout the entire physical space. Even with finite-sized basis sets, this stringent criterion is difficult to achieve for molecular systems described by many Gaussian basis functions including polarization and diffuse functions. A solution to this problem is to divide the molecule into regions, each of which is described by its own multipole moments defined with respect to its own origin (expansion site). Such a distributed multipole expansion mitigates the convergence issue of a single-origin multipole expansion. The choices of origins are arbitrary though atomic centers are a natural choice. As will be seen in Secs. II A–II C and III, a distributed multipole expansion is the common feature in most of the *ab initio* FF methods that are considered here. The distinctions among these methods are how a molecule is divided into regions, where to place expansion origins, and at which order the multipole expansion is truncated. It is well known that the values of multipole moments and polarizability tensors may depend on the origin of their coordinate systems. Hence, when using a distributed multipole expansion, it is important to carefully consider where to place the expansion sites. The application of the distributed multipole expansion for the interaction operator leads to distributed polarizability tensors to compute the polarization and dispersion energies.

For interactions that are electrostatic in origin (e.g., Coulomb, polarization, and dispersion, as will be discussed below), a singularity at R = 0 when using the (distributed) multipole expansion can be suppressed with damping functions. In *ab initio* FF methods, damping functions can be a source of empiricism, as discussed below.

A. The effective fragment potential (EFP) method

As will be seen below, there are very few empirically determined parameters in the EFP, even in the formulation of the damping function where the R-dependent inter-fragment overlap is employed. Since many concepts in the EFP are also used in the other *ab initio* FF methods, the EFP terms are described in some detail here. The original purpose of the EFP method was to model solvent-solute interactions. Parameters for the EFP were developed for water.[51] In this original implementation called EFP1, the total interaction energy is given by

$$
E = E^{Coul} + E^{pol} + E^{remainder}.
\tag{8}
$$

In Eq. (8), E^{Coul}, E^{pol}, and $E^{remainder}$ are the Coulomb energy, the polarization energy, and a remainder term. The parameters for the Coulomb and polarization term are generated from first principles and are explained in detail below. The $E^{remainder}$ term is fitted to a water dimer potential that was determined by either Hartree-Fock[51] or density functional theory (DFT)/B3LYP calculations.[52] So, EFP1 is not an entirely *ab initio* force field in the present sense. Subsequently, a generally applicable EFP method was developed, in which the fitted term is replaced by interaction energy contributions that are derived from first principles.[53,54]

The general EFP method, which is an *ab initio* force field that can be generated for any molecular system, is called EFP2 or general EFP. In this perspective, the general EFP will be referred to just as EFP, for simplicity.

In the EFP method, the intermolecular interaction between two fragments is partitioned into five components: Coulomb, polarization, dispersion, exchange repulsion, and charge transfer (CT),

$$
E = E^{Coul} + E^{pol} + E^{disp} + E^{ex-rep} + E^{CT}.
\tag{9}
$$

In Eq. (9), the Coulomb and polarization terms are calculated in the same way as in EFP1. The Coulomb interaction is calculated using the Stone distributed multipole analysis (DMA).[50] The distributed multipole expression for the Coulomb energy is similar to Eq. (6). Each of the subsystems A and B is divided into regions, and a multipole expansion is carried out in each region about its own expansion site. In the EFP method, nuclear centers and bond midpoints have been shown to be good choices of the expansion sites.[51] With different choices of expansion centers, the rate of convergence varies, and the multipole expansions need to be truncated accordingly. In the EFP, truncation at the charge-octopole term has shown satisfactory results,[55,56]

$$
E_{AB}^{Coul} = \sum_I^A \sum_J^B \left[q^J q^I T^{IJ} - \sum_\alpha^{x,y,z} T_\alpha^{IJ} \left(q^J \mu_\alpha^I - \mu_\alpha^J q^I \right) + \sum_{\alpha,\beta}^{x,y,z} T_{\alpha\beta}^{IJ} \left(\frac{1}{3} q^J \Theta_{\alpha\beta}^I - \mu_\alpha^J \mu_\beta^I + \frac{1}{3} \Theta_{\alpha\beta}^J q^I \right) \right.
$$
$$
\left. + \sum_{\alpha,\beta,\gamma}^{x,y,z} T_{\alpha\beta\gamma}^{IJ} \left(-\frac{1}{15} q^J \Omega_{\alpha\beta\gamma}^I + \frac{1}{3} \mu_\alpha^J \Theta_{\beta\gamma}^I - \frac{1}{3} \Theta_{\beta\gamma}^J \mu_\alpha^I + \frac{1}{15} \Omega_{\alpha\beta\gamma}^J q^I \right) + \frac{1}{9} \sum_{\alpha,\beta,\gamma,\delta}^{x,y,z} \Theta_{\alpha\beta}^J \Theta_{\gamma\delta}^I T_{\alpha\beta\gamma\delta}^{IJ} \right].
\tag{10}
$$

In Eq. (10), I and J are the expansion sites on subsystems A and B, respectively. $\Omega_{\alpha\beta\gamma}$ represents the components of the octopoles. All other terms are defined similarly as in Eq. (6).

The polarization/induction interaction in the EFP is modeled with anisotropic localized molecular orbital (LMO) dipole polarizability tensors located at LMO centroids. Induced dipoles, p^l, at the LMO centroids are self-consistently generated through the interaction of the LMO dipole polarizability tensors α^l with the static multipole and induced dipole fields of all of the other fragments in the system (F^l and F^l_{ind}, respectively), as shown in Eq. (11) for LMO l.[51,57] The self-consistent nature of this approach captures the many-body polarization effects,

$$p^l_\beta = \sum_\gamma^{\{x,y,z\}} \alpha^l_{\beta\gamma} \left(F^l_\gamma + F^l_{\gamma,ind} \right). \tag{11}$$

The EFP dispersion interaction is currently obtained at the 2nd-order perturbation theory level, and only two-body pairwise dispersion interactions are included. By truncating the interaction potential at the quadrupole term, EFP dispersion includes not only the usual R^{-6} term[58] but also the R^{-7} contribution.[59,60] An analytic expression for the R^{-8} term has been derived and is in the process of being implemented. The imaginary frequency-dependent polarizability tensors used for computing the EFP dispersion term arise from decomposing the total imaginary frequency-dependent polarizability tensor into LMO contributions, with appropriate origin-shifting transformations.[59] It is worth noting that, for entities such as atoms or centrosymmetric molecules, the dipole-quadrupole polarizability is zero. Hence a distributed polarizability formulation that is based on distributed atomic centers will not have an R^{-7} contribution. However, this is not the case if the distributed polarizability formulation is based on LMOs, as it is in the EFP method. The R^{-6} contribution to the dispersion energy can be computed using the isotropic and spherical approximations in the EFP, as shown in Eqs. (12a) and (12b), respectively. The expression for the R^{-7} term is given in Eq. (13),

$$E_6^{iso} = -\frac{1}{2\pi} \sum_{k\in A}^{LMO} \sum_{j\in B}^{LMO} \sum_{a,b}^{x,y,z} T_{ab}^{kj} T_{ab}^{kj} \int_0^\infty \alpha_{aa}^k(i\omega)\, \alpha_{bb}^j(i\omega)\, d\omega, \tag{12a}$$

$$E_6^{spher} = -\frac{3\hbar}{\pi} \sum_{k\in A}^{LMO} \sum_{j\in B}^{LMO} \frac{1}{R_{kj}^6} \int_0^\infty \bar{\alpha}^k(i\omega)\, \bar{\alpha}^j(i\omega)\, d\omega, \tag{12b}$$

$$E_7 = -\frac{\hbar}{3\pi} \sum_{k\in A} \sum_{j\in B} \sum_{\alpha\beta\gamma\sigma\kappa}^{x,y,z} T_{\alpha\beta}^{kj} T_{\gamma\sigma\kappa}^{kj} \int_0^\infty d\omega$$
$$\times \left[\alpha_{\alpha\gamma}^k(i\omega) A_{\beta,\sigma\kappa}^j(i\omega) - \alpha_{\beta\kappa}^j(i\omega) A_{\alpha,\gamma\sigma}^k(i\omega) \right]. \tag{13}$$

In Eq. (12a), the isotropic approximation was used, assuming the off-diagonal elements of the dipole polarizability matrix are zero. α^l and A^l in Eqs. (12a) and (12b) and Eq. (13) represent the dynamic dipole-dipole polarizability tensor and the dynamic dipole-quadrupole polarizability tensor, respectively, at LMO I. In Eq. (12b), $\bar{\alpha}^l$ is the average of the trace of the dynamic dipole-dipole polarizability tensor. R_{kj} is the distance between centroids of LMOs k and j.

The Coulomb, polarization, and dispersion interactions are "long-range interactions," and each formulation starts from a (distributed) multipole expansion. At short range, the R^{-n} functional form cannot be correct because of its singularity at R = 0. Without sufficient damping/screening, "polarization collapse" can occur. To suppress the R^{-n} singularity, the three types of interactions mentioned above require damping functions. The damping functions should ideally asymptotically tend to 1 as R → ∞ and to 0 as R → 0. Thus, the various forms of damping functions either contain exponentially decaying functions or Gaussian functions. To reduce the number of empirical parameters, damping functions based on the intermolecular overlap integrals have been devised and employed in the EFP and EFMO methods. The explicit forms of damping functions are discussed in Sec. IV D.

As mentioned in Sec. II, constructing the total wave function as a Hartree product of subsystem wave functions allows the interaction energy to be separated into terms that only depend on individual fragments, such as distributed polarizabilities. But, at short range, this Hartree product wave function is not adequate. Using an antisymmetrized wave function gives rise to the exchange repulsion and charge transfer interactions.

Both exchange repulsion and charge transfer interactions are considered to be short-range. These two types of interactions fall off exponentially with respect to subsystem separation. Thus, they will be small at 5 Å separation. But at intermediate separations, these interactions are not negligible, especially considering the accumulation from hundreds or thousands of molecules, as one might have in a typical force field molecular dynamics (MD) simulation. Since both of these interactions arise from the overlap of subsystem wave functions, a correct zeroth order wave function for the perturbation treatment should be the antisymmetrized product of wave functions |A⟩ and |B⟩, |AB⟩. With such a wave function, exchange repulsion emerges as part of the first-order perturbation energy and charge transfer emerges as part of the second-order perturbation energy.

Exchange repulsion, a purely quantum mechanical effect, is a direct consequence of the Pauli exclusion principle. The exchange repulsion in the EFP is derived from a power series expansion in terms of intermolecular overlap integrals between LMOs.[61–63] The parameters for calculating exchange repulsion include the intermolecular overlap, kinetic energy integrals, and Fock matrices of the subsystems in the LMO basis.[54,64,65]

Because of the same 2nd-order perturbation formulation, charge transfer may be viewed as part of the polarization energy that arises from states in which one of the subsystems is excited. Some of these states have one electron from subsystem A excited into a virtual orbital localized on subsystem B; such states contribute to the charge transfer interaction in the EFP framework. The parameters in the EFP for calculating charge transfer are canonical occupied molecular orbitals and the so-called "valence virtual orbitals" (VVOs[66,67]) as well as the orbital energies of those occupied MOs.[66] The VVOs, obtained through a single-valued decomposition (SVD) algorithm, can be regarded as the "chemically relevant" virtual orbitals. There are two major advantages of using these VVOs:

(1) the number of occupied MO + VVOs equals the number of minimal basis functions, which is constant for a given system, and (2) the VVOs are essentially basis-set independent. The original implementation of the EFP charge transfer interaction energy uses canonical virtual orbitals, whose number will typically be much larger than the number of occupied orbitals with a contracted Gaussian basis including polarization and diffuse functions. The VVO approach for the charge transfer interaction appears to be a very good approximation.[66]

For the EFP method, a periodic boundary condition using the minimum image convention (MIC-PBC) is implemented with appropriate switching functions in order to conserve the energy and the masses of molecular clusters during molecular dynamics (MD) simulations.[56] This implementation is not intended for infinite crystalline systems.

B. Sum of interaction between fragments *ab initio* computed (SIBFA)

Like EFP, SIBFA computes the intermolecular interaction energy as a sum of the electrostatic multipolar (E_{MTP}), short-range repulsion (E_{rep}), polarization (E_{pol}), dispersion (E_{disp}), and charge transfer (E_{ct}) terms,[68–70]

$$\Delta E_{\text{int}} = E_{MTP} + E_{rep} + E_{pol} + E_{disp} + E_{ct}. \quad (14)$$

The electrostatic interaction is computed with multipoles distributed at the atoms and bond barycenters up to quadrupoles using the Vigné-Maeder and Claverie scheme.[71] This scheme is similar to the Stone DMA method used in the EFP. The main difference is that a charge density originally located at a center **P** (resulting from the product of two Gaussian functions) is divided between the two closest expansion sites (atomic centers or bond midpoints) based on their relative distances from **P**. On the other hand, in the DMA, the charge density is fully transferred to the closest expansion site unless **P** is equidistant to two expansion sites. In that case, the charge density is divided equally between the two sites. Charge penetration effects were later included in SIBFA by expanding the three most important contributions to the electrostatic energy: monopole-monopole, monopole-dipole,[72] and monopole-quadrupole.[43] The monopole-monopole energy contribution between two interacting centers I and J is expanded into core-core repulsion, core-electron attraction, and electron-electron repulsion,

$$E_{mono-mono} = \begin{bmatrix} Z_I Z_J - Z_I \left(Z_J - q_J\right)\left(1 - \exp\left(-\alpha_J r_{IJ}\right)\right) \\ -Z_J \left(Z_I - q_I\right)\left(1 - \exp\left(-\alpha_I r_{IJ}\right)\right) \\ +\left(Z_I - q_I\right)\left(Z_J - q_J\right)\left(1 - \exp\left(-\beta_I r_{IJ}\right)\right)\left(1 - \exp\left(-\beta_J r_{IJ}\right)\right) \end{bmatrix} \left(1/r_{IJ}\right). \quad (15)$$

In Eq. (15), I and J are the interacting centers (i.e., atomic centers or bond midpoints). Z_I and Z_J represent the number of valence electrons of atoms I and J for monopoles q_I and q_J located at atomic centers. For monopoles located at bond midpoints, Z is 0. α_I and β_I are fitted parameters that are inversely proportional to the van der Waals (vdW) radii of the atom I if I is located at an atomic center. If I is a bond midpoint, the vdW radius is the average between the two radii of the atoms forming the bond. Similarly, the monopole-dipole energy is given by

$$E_{mono-dip} = -\mu_J \left[Z_I - \left(Z_I - q_I\right)\left(1 - \exp\left(-\eta r_{IJ}\right)\right)\right] r_{IJ}/r_{IJ}^3, \quad (16)$$

where $\eta = \chi \big/ \left(\frac{r_{vdw,i} + r_{vdw,j}}{2}\right)$, with χ also being a fitted parameter.

The original monopole-quadrupole energy is given by the sum of the interaction energies between a monopole and two axial quadrupoles,

$$E_{mono-quad} = E_{mono-quad}^1 + E_{mono-quad}^2, \quad (17a)$$

$$E_{mono-quad}^n = q \left(\frac{Q_{a,n}}{2r^3}\right) \left[3\left(\frac{\mathbf{a} \cdot \mathbf{r}}{r}\right)^2 - 1\right]. \quad (17b)$$

In Eq. (17b), **a** is a unit vector, **r** is the vector from monopole to the *n*th axial quadrupole, and $Q_{a,n}$ is the quadrupole magnitude for the *n*th axial quadrupole in direction *a*. The superscript *n* is either 1 or 2.

In the new SIBFA formulation that includes charge penetration, the monopole q in Eq. (17b) is replaced by an exponential function of r, similar to the monopole-dipole case,[43]

$$E_{mono-quad} = \left[Z_I - \left(Z_I - q_I\right)\left(1 - \exp\left(-\varphi r_{IJ}\right)\right)\right] \times \left(3(\mathbf{a} \cdot \mathbf{u})^2 - 1\right)\left(\frac{Q_a}{2r^3}\right), \quad (18)$$

where φ is a parameter that depends on the vdW radii of the atoms I and J, similar to η in Eq. (16).

Similar to the EFP, the SIBFA many-body polarization is obtained by iterating the induced dipoles $\Delta\mu_P$ at points P on each fragment A in a self-consistent fashion,[73]

$$\Delta\mu_P (i) = \sum_j^{x,y,z} \left(\alpha_P (i,j) E_P(j)\right). \quad (19)$$

In Eq. (19), $\alpha_P (i,j)$ are the induced anisotropic polarizabilities at point P on fragment A and $E_P(j)$ is the electrostatic field generated by all surrounding fragments at point P. The electrostatic field generated by every point Q of each fragment B on fragment A is damped by the following Gaussian function:[73,74]

$$1 - q_Q E \exp\left(\frac{-F\left(R_{PQ}^2\right)}{V_P + V_Q}\right). \quad (20)$$

In Eq. (20), q_Q is the partial charge of point Q on fragment A and R_{PQ} is the distance between point P on fragment A and point Q on fragment B. V_P and V_Q are the effective radii of

points P and Q. E and F are scaling factors. The effective radii V are defined to be proportional to the absolute value of the induced dipole moment and are iterated self-consistently. The polarizability tensors for obtaining the induced dipoles are distributed on the LMOs. The details of the remaining three terms in Eq. (14) differ substantially from the EFP. Since

there are parameters in the remaining terms that require fitting/calibration, this method is not entirely an *ab initio* force field.

The SIBFA repulsion interaction between fragment 1 and fragment 2 is the sum of bond-bond, bond-lone pair, and lone pair-lone pair interactions, given by[43]

$$E_{rep}^{12} = C_I \left(\sum_{AB}^{\substack{\text{bonds} \\ \in 1}} \sum_{CD}^{\substack{\text{bonds} \\ \in 2}} rep(AB, CD) + \sum_{AB}^{\substack{\text{bonds} \\ \in 1}} \sum_{L\gamma}^{\substack{\text{lone pairs} \\ \in 2}} rep(AB, L\gamma) \right. \\ \left. + \sum_{L\alpha}^{\substack{\text{lone pairs} \\ \in 1}} \sum_{CD}^{\substack{\text{bonds} \\ \in 2}} rep(L\alpha, CD) + \sum_{L\alpha}^{\substack{\text{lone pairs} \\ \in 1}} \sum_{L\gamma}^{\substack{\text{lone pairs} \\ \in 2}} rep(L\alpha, L\gamma) \right). \tag{21}$$

In Eq. (21), the first term represents the repulsion between bond AB in fragment 1 and bond CD in fragment 2. The second term represents the repulsion interaction between the bond AB in fragment 1 and the lone pairs $L\gamma$ in fragment 2. The third term represents the repulsive interaction between bond CD in fragment 2 and lone pairs $L\alpha$ in fragment 1. The last term represents the interaction between lone pairs in fragment 1 with lone pairs in fragment 2. C_I is an empirically defined parameter. Each term in Eq. (21) has the form[43,75]

$$rep(AB, CD) = N_{occ}(AB) N_{occ}(CD) \\ \times \left[\frac{C_1 S^{**2}(AB, CD)}{D_{AB,CD}} + \frac{C_2 S^{**2}(AB, CD)}{D_{AB,CD}^2} \right], \tag{22}$$

where N_{occ} is the occupation number of the bonds. $S^{**2}(AB,CD)$ is a functional of the overlap between localized molecular orbitals of the interacting atom pairs (AC, AD, BC, BD).[73,76,77] $D_{AB,CD}$ is the distance between the barycenters of bonds AB and CD. C_1 and C_2 are fitted parameters. The overlap between interacting localized molecular orbitals on atoms A and C is given by the expression[77]

$$S_{AC} = M_{AC} \exp(-\alpha \rho_{AC}), \tag{23}$$

where $\rho_{AC} = \frac{r_{AC}}{4\sqrt{W_A W_C}}$ and $M_{AC} = \sqrt{K_{AC}\left(1 - \frac{q_A}{Z^A}\right)\left(1 - \frac{q_C}{Z^C}\right)}$, W_A and W_C are the effective radii of atoms A and C, respectively. q_A and q_C are the net charges of atoms A and C. K_{AC} is a parameter that depends on the atomic numbers of A and C.

Similar to the EFP, the SIBFA dispersion energy E_{disp} is a pairwise additive quantity defined as[77,78]

$$E_{disp} = \sum_{AB} \left[\frac{C_6^{AB}}{\mathbf{Z}_{AB}^6} + \frac{C_8^{AB}}{\mathbf{Z}_{AB}^8} + \frac{C_{10}^{AB}}{\mathbf{Z}_{AB}^{10}} \right]. \tag{24}$$

In Eq. (24), $\mathbf{Z}_{AB}$ is a function of the distance R_{AB} between atoms A and B, given by

$$\mathbf{Z}_{AB} = \frac{R_{AB}}{\sqrt{W_A + W_B}}. \tag{25}$$

The C_n ($n = 6, 8, 10$) coefficients are empirically fitted parameters. Each energy term in Eq. (24) is damped according to the

following expression:

$$E_{damp(n)} = \left(1/\mathbf{Z}_{AB}^n\right) L_{AB} \exp\left(-a_n \left(\frac{(W_A + W_B) b_n}{R_{AB}} - 1\right)\right), \tag{26}$$

where a_n and b_n are empirically fitted coefficients. L_{AB} is a parameter that depends on the atoms A and B. In addition, an exchange-dispersion term is added to the dispersion energy for each atom pair,

$$E_{ex-disp} = L_{AB} \left(1 - \frac{q_A}{N_{val}(A)}\right) \left(1 - \frac{q_B}{N_{val}(B)}\right) \\ \times C_{exch} \exp(-\beta_{exch} \mathbf{Z}_{AB}), \tag{27}$$

where q_A and q_B are the net charges of atoms A and B and $N_{val}(A)$ and $N_{val}(B)$ represent the number of valence electrons of atoms A and B, respectively. C_{exch} and β_{exch} are empirical parameters. The parameterization was performed so that the dispersion energy fits Symmetry-Adapted Perturbation Theory (SAPT) results obtained for a set of hydrogen-bonded dimers.

The SIBFA charge transfer term only considers electron donation from lone pair orbital $L\alpha$ in molecule A (electron donor) to antibonding orbitals β^* linking a hydrogen atom with a heavy atom in the acceptor molecule B. The charge transfer energy is computed with the formula[77]

$$E_{ct} = -2C \sum_{L\alpha} N_{occ}(\alpha) \left(\frac{\left(T_{\alpha\beta^*}\right)^2}{\Delta E_{\alpha\beta^*}} \right), \tag{28}$$

where C is a fitted parameter and N_{occ} is the occupation number of the lone pair. $\Delta E_{\alpha\beta^*}$ is the energy involved in the electron transfer between the electron donor and acceptor. It is expressed as a function of the ionization potential $I_{L\alpha}$ of the electron donor, the electron affinity of the acceptor molecule A_{β^*}, and the electron potential V exerted by all the molecules C in the system,[77]

$$\Delta E_{\alpha\beta^*} = \left(I_{L\alpha} + \sum_C V_{C \to A} \right) - \left(A_{\beta^*} + \sum_C V_{C \to B} \right). \tag{29}$$

$T_{\alpha\beta^*}$ is given by the integral

$$T_{\alpha\beta^*} = \int \rho_{\alpha\beta^*}(r)V(r)\,d\nu. \tag{30}$$

In Eq. (30), $V(r)$ is the potential generated by the acceptor molecule and $\rho_{\alpha\beta^*}(r)$ is the transition density between orbitals α and β^*. $\rho_{\alpha\beta^*}$ is given by[76]

$$\rho_{\alpha\beta^*} = -\left(\alpha\beta^* - \alpha^2 S_{\alpha\beta^*}\right). \tag{31}$$

In Eq. (31), $S_{\alpha\beta^*}$ is the overlap between the orbitals α and β^*.

C. Related *ab initio* force field methods

Several force fields use DMA to compute electrostatic interactions as it is computationally efficient.[21,22,79] However, in order to describe other intermolecular interactions, they use fitted functional forms that are not derived from first-principles and/or contain empirical parameters. There are several other force field methods, in which the interaction energy decomposition is similar to that of the EFP, but the details of each energy component differ substantially.

For example, Piquemal *et al.* developed a variation of SIBFA, which computes the Coulomb energy from the fitted electron density,[80–82]

$$E_{Coulomb} = \sum_i \sum_j \frac{Z_{A_i}Z_{B_j}}{r_{A_iB_j}} - \sum_i \int \frac{Z_{A_i}\tilde{\rho}^B(r_2)}{r_{2A_i}}dr_2$$
$$- \sum_j \int \frac{Z_{B_j}\tilde{\rho}^A(r_1)}{r_{1B_j}}dr_1 + \iint \frac{\tilde{\rho}^A(r_1)\tilde{\rho}^B(r_2)}{r_{12}}dr_1 dr_2. \tag{32}$$

In Eq. (32), $\tilde{\rho}^A(r_1)$ is the fitted electron density of fragment A, $\tilde{\rho}^B(r_2)$ is the fitted electron density of fragment B, and $r_{A_iB_j}$ is the distance between nucleus i in monomer A and nucleus j in monomer B. Z_{A_i} and Z_{B_j} are the charges of the nuclei i and j in fragments A and B, respectively. r_{2A_i} represents the distance between electrons in fragment B and nucleus i in fragment A. r_{1B_j} represents the distance between electrons in fragment A and nucleus j in fragment B. r_{12} is the distance between electrons in fragments A and B. In the original GEM-0 method,[82] the electron density $\tilde{\rho}$ is a linear combination of s-type ($l = 0$) Gaussian auxiliary basis functions, where the coefficients are fitted to minimize the self-interaction error between the fitted density and *ab initio* density (obtained with DFT or CCSD, for example). The main advantage of this method over DMA is that short-range penetration effects are included, therefore avoiding the use of an arbitrary damping function. Another advantage that has been exploited by the developers is the use of the fitted electron density to compute the wave function overlap S between the spherical Gaussian functions on the different fragments and then estimate the exchange-repulsion energy using the function,

$$E_{ex-rep} \approx KS = K\int \tilde{\rho}^A(r)\tilde{\rho}^B(r)dr. \tag{33}$$

In Eq. (33), K is a parameter obtained from the slope of a linear regression of the *ab initio* exchange repulsion vs the density overlap. One disadvantage of the GEM-0 method is that the evaluation of the Coulomb energy requires com-

puting many integrals, therefore making the method more expensive than DMA. Another disadvantage is that the use of $l = 0$ Gaussian functions as an auxiliary basis set (ABS) does not properly account for the anisotropy of the electron density. As a result, several expansion sites in addition to atomic centers and bond midpoints are required (a total of 9 for water) to properly model the electron density. GEM-0 was later improved to include higher angular Gaussian-Hermite basis functions.[81] The permanent electric field and electrostatic potential may be generated from the density fitting procedure to compute the polarization and charge transfer terms using the original SIBFA formalism.[82] Alternatively, the other energy terms may be computed with the AMOEBA force field (GEM*).[83] More recently, Chaudret *et al.* developed the S/G-1 force field, which uses GEM densities to model metallic cations and compute the polarization and charge transfer energies between the cation and ligands, while all other energy contributions (Coulomb, dispersion, and exchange repulsion) are modeled with SIBFA.[74] It was shown that the electric fields generated with GEM yield improved polarization energies through a better description of quantum effects at short range.

Another FF method that has features that are similar to the EFP is explicit polarization (X-Pol). Similarly to other *ab initio* Force Fields, X-Pol partitions a total system into fragments that can be individual molecules or, in the case of a polypeptide chain, a peptide unit where the boundary C atom shares four generalized hybrid orbitals (GHO) between neighboring fragments.[84] A buffering scheme was developed for faster convergence.[85] As for the EFP, the total wave function is approximated as a Hartree product of individual fragment antisymmetric wave functions. The effective Hamiltonian of the total system is given by

$$\hat{H} = \sum_A^N \hat{H}_A^0 + \frac{1}{2}\sum_A^N \sum_{B\neq A}^N \left(\hat{H}_A^{int}\left[V_E^B\right] + \Delta E_{AB}^{XCD}\right). \tag{34}$$

The first term in Eq. (34) is the sum over the Hamiltonians of all isolated fragments. The second term accounts for pairwise interactions among all the fragments. The $\hat{H}_A^{int}\left[V_E^B\right]$ term represents the electrostatic interaction between fragments A and B. ΔE_{AB}^{XCD} is the exchange repulsion (X), charge delocalization (C), and dispersion (D) interactions between fragments A and B. For the electrostatic interaction, the potential due to fragment B is represented by partial atomic charges (truncated at the monopole terms in a distributed multipole expansion).[84] The exchange repulsion and dispersion pairwise interaction can be represented as a Lennard-Jones[84] or Buckingham potential.[86]

The Lennard-Jones potential has the form

$$E_{AB}^{XD} = \sum_I^{N_A} \sum_J^{N_B} 4\varepsilon_{IJ}\left[\left(\frac{\sigma_{IJ}}{R_{IJ}}\right)^{12} - \left(\frac{\sigma_{IJ}}{R_{IJ}}\right)^6\right], \tag{35}$$

where $\sigma_{IJ} = (\sigma_I\sigma_J)^{1/2}$, $\varepsilon_{IJ} = (\varepsilon_I\varepsilon_J)^{1/2}$, N_A is the number of atoms in fragment A, and σ_I and ε_I are taken from the CHARMM protein force field.[84]

The Buckingham potential has the following expression:

$$E_{AB}^{XD} = \sum_{I}^{N_A} \sum_{J}^{N_B} A_{IJ} e^{-B_{IJ}R_{IJ}} - \frac{C_{IJ}}{R_{IJ}}, \qquad (36)$$

where the parameters A_{IJ}, B_{IJ}, and C_{IJ} are defined as $A_{IJ} = (A_I A_J)^{1/2}$, $B_{Ij} = (B_I + B_J)/2$, and $C_{IJ} = (C_I C_J)^{1/2}$.

To include polarization effects, the double self-consistent field (DSCF)[84] approach is used, in which each individual fragment wave function is iteratively converged, followed by one iterative cycle of the total system SCF. This process is repeated until the convergence of the total energy is achieved. The Fock matrix of fragment A can be written as

$$F_{\mu\nu}^A = F_{\mu\nu}^0 - \frac{1}{2} \sum_{B \neq A} \sum_{b \in B} q_b^B \left(I_b^B\right)_{\mu\nu}^A + \frac{1}{2} \sum_{a \in A} X_a^A \left(\Lambda_a^A\right)_{\mu\nu}. \quad (37)$$

In Eq. (37), $F_{\mu\nu}^0$ is the Fock matrix of the isolated fragment A, q_b^B is the point charge of atom b in fragment B, and I_b^B is the one-electron matrix of the embedding potential due to B. The last term is the response of the charge density of B due to variational optimization of fragment A.

X-Pol incorporates charge delocalization by expanding the localization space from a single fragment (monomer) block to pairwise blocks (dimers). As mentioned above, the total wave function is the Hartree product of individual antisymmetric fragment wave functions. So the antisymmetrized total wave function, Ψ_x^A, can be expressed as

$$\Psi_x^A = R_x^A \hat{A} \left(\prod \Phi_a\right), \qquad (38)$$

where $\hat{A}$ is an antisymmetrizing operator and R_x^A is the normalization constant.

A charge transfer state between fragments A and B is defined as

$$\Psi_{(ab)} = R_{(ab)} \hat{A}(\Phi_1) \ldots \hat{A}(\Phi_{ab}) \ldots \hat{A}(\Phi_M), \qquad (39)$$

where Φ_{ab} is expanded over the basis functions of the dimer space. The analogous fully antisymmetrized total wave function is

$$\Psi_{(ab)}^A = R_{(ab)}^A \hat{A}(\Phi_1 \ldots \Phi_{ab} \ldots \Phi_M). \qquad (40)$$

Then the energy difference between the two states is the X-Pol charge transfer energy between the two fragments,

$$\Delta E_{(ab)}^{CT} = \left\langle \Psi_{(ab)}^A \middle| H \middle| \Psi_{(ab)}^A \right\rangle - \left\langle \Psi_x^A \middle| H \middle| \Psi_x^A \right\rangle. \qquad (41)$$

Since 2-body delocalization can be generalized to grouping multiple fragment blocks into one, this implementation is referred to as the generalized X-Pol (GX-Pol).

In an effort to reduce the dependence on empirical parameters in the dispersion and exchange-repulsion terms in X-Pol, Jacobson and Herbert[87] adapted X-Pol to add in SAPT interaction energy terms, in a method called XSAPT.[88]

In addition to XSAPT, other force field methods also use SAPT to avoid depending on empirical parameters. In these methods, many of the parameters are derived from monomer or atom properties, but some part of the potential is typically fit to the SAPT energy or energy decomposition of a dimer. Schmidt and co-workers[89–91] have leveraged the energy decomposition provided by SAPT to fit parameters to simple functional forms for the different SAPT interaction energy terms (Coulomb,

exchange repulsion, etc.). The parameters produced can be used in current standard MD simulation packages. Also based on SAPT, Stone, Misquitta, and co-workers[92–94] developed potentials focused on rigorously deriving distributed properties such as multipole moments and polarizabilities used in the potentials, as well as including anisotropic terms.

The above examples demonstrate that some *ab initio* force fields often differ in the details of the distributed multipole approach for the Coulomb interactions, the use of an iterative procedure to obtain many-body polarization, and the decision regarding whether and how to include charge transfer. Empiricism is often used in the treatment of exchange repulsion and dispersion interactions. In addition, other force fields use parameters that are derived from electronic structure theory as input data, but the functional forms of some energy components are not quantum mechanically derived.[95,96] Discussing all of those methods in detail is beyond the scope of this perspective, and the use of empirical parameters does not fully conform to the concept of *ab initio* force fields.

III. *AB INITIO* FORCE FIELDS THAT COMBINE MANY-BODY EXPANSION AND INTERMOLECULAR PERTURBATION THEORY

The methods discussed in this section combine a many-body expansion with intermolecular perturbation theory. There are many methods that calculate the total energy using a many-body expansion approach alone, such as the fragment molecular orbital (FMO) method,[97,98] the systematic molecular fragmentation (SMF) method,[99–101] molecular fractionation with conjugate caps (MFCC),[102,103] and the molecular tailoring approach (MTA).[104,105] Taking the FMO method as an example, the total energy of a system is partitioned through a many-body expansion, as described in Eq. (3). The n-body interactions can be obtained either as an isolated n-mer (denoted as ΔE_x^0) or in the presence of the rest of the fragments (ΔE_x). The FMO method self-consistently converges the monomer energies in the electrostatic potential (ESP), which includes the nuclei-electron attraction and electron-electron repulsion, of all the other monomers. The dimer and trimer energies may also be computed in the ESP of all of the remaining fragments (by excluding the interacting dimer or trimer) but are not iterated to self-consistency.

Fragmentation methods like FMO compute the energies of monomer, dimer, trimer, etc., using standard *ab initio* methods and therefore are not force fields in a traditional sense. The accuracy of fragmentation methods using the many-body expansion alone will be limited to the level of *ab initio* theory employed as well as the level of n-body truncation. Many-body interactions can have non-negligible or even substantial contributions in some systems; it is therefore important to include three-body terms to obtain accurate results.[106] The number of dimer and trimer calculations increases rapidly as the number of fragments increases. Moreover, in large systems, many dimers and trimers are comprised of pairs of distant monomers. The interactions between such distant monomers can be evaluated using one of the force field methods mentioned in Sec. II. Among the various types of interactions,

polarization is usually the dominant many-body interaction. The many-body polarization can be obtained in an economical way through the use of some force fields, as exemplified in the EFP method.

A. The effective fragment molecular orbital (EFMO) method

The total EFMO energy of a system is evaluated as[44,107]

$$E^{EFMO} = \sum_I E_I^0 + \sum_{I>J}^{R_{IJ} \leq R_{cut}} \left(\Delta E_{IJ}^0 - E_{IJ}^{pol} \right)$$
$$+ \sum_{I>J}^{R_{IJ} > R_{cut}} \left(E_{IJ}^{Coul} + E_{IJ}^{disp} + E_{IJ}^{exrep} + E_{IJ}^{ct} \right) + E_{tot}^{pol}. \quad (42)$$

The superscript 0 means the energies are calculated from isolated gas-phase monomers and dimers. All of the other terms in Eq. (42) are obtained from the EFP method. In the FMO method, the monomers are iterated until self-consistency in an electrostatic potential due to all the other monomers. In the EFMO method, the electrostatic potential is replaced by the EFP if the fragments are far enough apart. This is determined by the use of a unitless "distance" cutoff R_{cut} defined as

$$R_{IJ} = \min_{i \in I, j \in J} \left\{ \frac{|\vec{r}_i - \vec{r}_j|}{r_i^{vdw} + r_j^{vdw}} \right\}. \quad (43)$$

In Eq. (43), R_{IJ} is the minimal ratio of interatomic distance and the sum of van der Waals radii of atom i in fragment I and atom j in fragment J.

The EFP is generated for every fragment in the system, and the energy of each fragment is computed as an isolated system using an *ab initio* method of choice. For dimers that are closer to each other than R_{cut}, dimer interaction energies are calculated with the chosen *ab initio* method, and the pairwise EFP polarization energy is subtracted to avoid double counting the dimer polarization effect. For dimers separated by a distance larger than R_{cut}, the dimer interaction, except pairwise polarization, is calculated using the EFP. Finally, the total EFP many-body polarization is added after the induced dipoles achieve self-consistency, as is done in the EFP method.

Compared to the FMO method truncated at two-body terms (FMO2), the EFMO method shows noticeable time saving and typically smaller errors.[107] Molecular dynamics simulations can be carried out with the EFMO method. For simulations in which all dimers are considered to be EFP dimers, periodic boundary conditions with the minimum image convention (MIC-PBC) are implemented.

B. Hybrid many-body interaction (HMBI) model

The hybrid many-body interaction (HMBI) model combines a many-body expansion with an approximate many-body polarization calculated using an *ab initio* force field.[47] Short-range pairwise interactions are modeled with quantum mechanics (QM), whereas long range and many-body interactions are modeled with molecular mechanics (MM). The total

energy is given in the following equation:[48]

$$E_{total}^{HMBI} = \sum_i^{monomer} E_i^{QM} + \sum_{ij}^{dimer} \Delta^2 E_{ij}^{QM} + \left[\sum_{ijk}^{trimer} \Delta^3 E_{ijk}^{MM} + \cdots \right]$$
$$= \sum_i^{monomer} E_i^{QM} + \sum_{ij}^{dimer} \Delta^2 E_{ij}^{QM}$$
$$+ \left[E_{total}^{MM} - \sum_i^{monomer} E_i^{MM} - \sum_{ij}^{dimer} \Delta^2 E_{ij}^{MM} \right]$$
$$= E_{total}^{MM} + \sum_i^{monomer} \left(E_i^{QM} - E_i^{MM} \right)$$
$$+ \sum_{ij}^{dimer} \left(\Delta^2 E_{ij}^{QM} - \Delta^2 E_{ij}^{MM} \right). \quad (44)$$

In Eq. (44), $\Delta^2 E_{ij} = E_{ij} - E_i - E_j$. E_i, E_{ij}, and E_{ijk} represent the monomer, dimer, and trimer energies.

In the original HMBI implementation, the classical AMOEBA force field was used for the MM part of the simulation.[47] However, the empirical nature of AMOEBA yields high mean signed errors and total error ranges, as illustrated by water and formamide clusters.[47] In addition, structures that are far away from equilibrium are poorly described. In order to overcome this problem, an *ab initio* force field was developed in which parameters are computed on the fly for each geometry.

The MM part of the HMBI model includes the two-body Coulomb interaction, many-body polarization, two-body dispersion, and recently also the three-body Axilrod-Teller-Muto dispersion.[108,109] The Coulomb interaction between two fragments is evaluated using a distributed multipole representation,

$$E_{Coul}^{AB} = \sum_{a \in A} \sum_{b \in B} \sum_{tu} Q_t^a T_{tu}^{ab} Q_u^b, \quad (45)$$

where Q_t^a represents the t-th order multipole moment on atom a in fragment A and Q_u^b represents the u-th order multipole moment on atom b in fragment B. T_{tu}^{ab} is the *tu* component of the electrostatic T tensor that contains the distance between atoms a and b [see Eq. (7)]. Anisotropic atomic multipole moments up to rank 4 (hexadecapole) on heavy atoms and rank 1 (dipole) on hydrogen atoms are used. The Coulomb energy of the system is given by the sum of all pairwise interactions, which is evaluated using a multipolar Ewald summation.

HMBI many-body polarization is evaluated by iterating induced multipole moments to self-consistency,

$$\Delta Q_t^a = - \sum_{B \neq A} \sum_b \sum_{a'} \sum_{t'u} \alpha_{tt'}^{aa'} f_n(R, \beta) T_{t'u}^{a'b} \left(Q_u^b + \Delta Q_u^b \right). \quad (46)$$

The Tang-Toennies formula[110] is used as a damping function $f_n(R, \beta)$, where β is a fitted parameter. $\alpha_{tt'}^{aa'}$ is the atomic static polarizability tensor, which contains terms up to rank 2 (quadrupole-quadrupole polarizability for heavy atoms and dipole-dipole polarizability for hydrogen atoms). T tensor, $T_{t'u}^{a'b}$, and multipole moment, Q_u^b, are the same as in Eq. (45). ΔQ_u^b represents the induced multipole on atom b due to atom a.

An attractive feature of HMBI is the implementation of periodic boundary conditions, which allows simulations of periodic crystals at a reasonable cost. In periodic HMBI, short-range interactions are treated using quantum mechanics, whereas long-range interactions are treated with molecular mechanics. In order to have a smooth potential energy surface, a transition between these two regimes is defined, in which intermolecular interactions are computed as a linear combination of QM and classical interactions. Another recent addition to HMBI is the development of an algorithm to determine and exploit space-group symmetry.[111]

IV. SOME KEY ASPECTS OF *AB INITIO* FF METHODS

A. Flexible fragments

Both the EFMO and HMBI methods generate the force field parameters "on the fly" to account for small but potentially important geometrical variations within each fragment. Clearly, this will impact the speed of the calculation significantly if the total system is large. Moreover, sensitivity to geometry changes is property-dependent and system-dependent.[48] During the recalculation of the force field, the definition of the fragments does not change. This implies that bond-breaking/formation across the fragment border cannot occur unless a dynamic fragmentation scheme has been implemented. Inspired by the idea of the Multistate Empirical Valence Bond (MS-EVB) force field, Lange and Voth carried out pioneer work on dynamic fragmentation using a simplified version of FMO2. However, their method, Fragment Molecular Orbital Multistate Reactive Molecular Dynamics (FMO-MS-RMD), is not yet general.[112]

B. QM/MM interface

Ab initio force fields can be interfaced with full *ab initio* quantum mechanical calculations. The EFP method was originally created to explicitly model aqueous solvation. In a QM/EFP calculation, the interaction between a quantum mechanical solute and EFP fragments as well as between EFP fragments is computed. Currently the Coulomb, polarization, and exchange repulsion interactions are accounted for by adding these terms to the *ab initio* Hamiltonian as a one-electron operator.[51] The dispersion is added as a post-SCF correction.[113] Recently an alternative implementation of the QM/EFP dispersion interaction has been developed.[114] The QM/EFP charge transfer energy has not yet been developed. The development of analytic gradients for the components of the QM/EFP interactions is in progress. Multiple levels of fragmentation can be used if the solute molecule is very large (for instance, a protein). In fact, the solute can be modeled with the FMO or EFMO method, while the solvent is modeled with the EFP method.

C. Interface with continuum solvation models

Force fields are often used to model solvation. The solvent molecules surrounding a solute can have a substantial influence on the properties of the solute molecules. For example, protic solvent molecules can hydrogen bond with solute molecules, and even aprotic solvent molecules can interact with the solute via dispersive interactions. Therefore, solvents are best modeled as discreet molecules when practical. A disadvantage of modeling solvation with explicit solvent molecules is that the need for extensive configurational sampling can be computationally demanding. As the distance between the solvent and solute molecules increases (second solvation shell outwards), the electrostatic field is the dominant interaction, so these interactions can be reliably modeled by a continuum solvation model,[115] e.g., the polarizable continuum model (PCM),[116–119] without loss of accuracy. This type of approach has the advantage that it can significantly reduce the number of explicit solvent molecules in the simulation, thereby reducing the computational cost. However, if enough explicit solvent molecules are included, a continuum solvent is not necessary. The original effective fragment potential (EFP1) method was interfaced with the Onsager[120] and PCM models.[121,122] The SIBFA method was interfaced with the Langlet-Claverie continuum model.[123–125]

D. Damping/screening and penetration energy

The damping functions for some of the methods discussed in this perspective are summarized in Table I.

For the Coulomb term in the EFP method, there are multiple ways to account for the short-range electrostatic screening and charge penetration effect. One way is to use damping functions to screen the multipole moment interactions. For this option, an exponential damping function is applied to the charge-charge interaction term in the multipole moment expansion. Damping of the higher multipole moment interactions is an option in the EFP method as well. The damping parameter is determined by fitting the damped multipole potential to the QM potential on grid points. Another way to account for screening is to include a separate term that approximates the charge penetration energy and does not modify the terms in the multipole moment expansion. In that approach, when the charge penetration term is used, the damped Coulomb energy is the sum of the undamped multipole moment expansion and the penetration energy. In the EFP method, the charge penetration term is computed using overlap integrals between the LMOs of fragments.

For the EFP polarization interaction, screening is accounted for with a damping function that multiplies the interaction tensors. The damping function can have a Gaussian or exponential form. The damping parameters are set to 0.6 and 1.5 in Gaussian damping function and exponential damping function, respectively.[126] The EFP dispersion interaction is also damped with a multiplicative damping factor. The damping factor can be Tang-Toennies-like[110] or based on an expansion of LMO overlap integrals. The Tang-Toennies damping for EFP dispersion interaction employs a single-value (1.5) parameter β, based on a study of small dimers.[126] The overlap damping is preferred, especially since no arbitrary or empirically fitted parameter is required. Similar to the EFP method, the EFMO method employs overlap-based damping for the electrostatic term, Gaussian damping for the polarization term, and overlap-based damping for the dispersion term.

In the SIBFA method, the monopole-monopole, monopole-dipole, and monopole-quadrupole terms of the

TABLE I. Damping functions of *ab initio* force field methods discussed in this perspective.

	Electrostatic	Polarization	Dispersion
EFP	Exponential $$P^{AB}(\alpha,\beta)$$	Gaussian $$f(\alpha,\beta,R) = 1 - e^{-\sqrt{\alpha\beta}R^2}\left(1+\sqrt{\alpha\beta}R^2\right)$$	Overlap-based $$f_S^N = 1 - S_{ij}^2 \sum_{n=1}^{N} \frac{\left(-2\ln\left|S_{ij}\right|\right)^{n/2}}{n!}$$
	Overlap-based $$E^{pen} = -2\left(\frac{1}{-2\ln\left|S_{ij}\right|}\right)^{1/2}\frac{S_{ij}^2}{R_{ij}}$$	Exponential $$P^{AB}(\alpha,\beta)$$	Tang-Toennies $$f_{TT} = 1 - e^{-\beta R}\sum_{n=1}^{N}\frac{(\beta R)^n}{n!}$$
SIBFA	Exponential For monopole-monopole, monopole-dipole, and monopole-quadrupole terms $$1 - \exp(\alpha R)$$ [Eq. (15), (16), and (18)]	Gaussian $$1 - q_Q E \exp\left(\frac{-F\left(R_{PQ}^2\right)}{V_P + V_Q}\right) \text{[Eq. (20)]}$$	Exponential $$E_{damp(n)} = \left(1/Z_{AB}^n\right)L_{AB}\exp\left(-a_n\left(\frac{(W_A+W_B)\,b_n}{R_{AB}}-1\right)\right)$$ [Eq. (26)]
EFMO	Overlap-based $$E^{pen} = -2\left(\frac{1}{-2\ln\left|S_{ij}\right|}\right)^{1/2}\frac{S_{ij}^2}{R_{ij}}$$	Gaussian $$f(\alpha,\beta,R) = 1 - e^{-\sqrt{\alpha\beta}R^2}\left(1+\sqrt{\alpha\beta}R^2\right)$$	Overlap-based $$f_S^n = 1 - S_{ij}^2\sum_{n=1}^{N/2}\frac{\left(-2\ln\left|S_{ij}\right|\right)^n}{n!}$$
HMBI	n/a	Tang-Toennies $$f_n(R,\beta) = 1 - e^{-\beta R}\sum_{k=0}^{n}\left(\frac{(\beta R)^k}{k!}\right)$$	Tang-Toennies $$f_n(R,\beta) = 1 - e^{-\beta R}\sum_{k=0}^{n}\left(\frac{(\beta R)^k}{k!}\right)$$

electrostatic energy are damped with a distance-dependent exponential function, as shown in Eqs. (15), (16), and (18). The electrostatic field that is computed to evaluate the polarization energy and the R^{-6}, R^{-8}, and R^{-10} dispersion energy components are also damped with an exponential function that decays with decreasing R [Eqs. (20) and (26)].

In the HMBI method, Tang-Toennies damping is applied when calculating the induced multipole moments for polarization.[110] The parameter β in Eq. (46) is determined by fitting the HMBI polarization energy to the MP2 polarization energy for a set of 10 trimers; therefore, β needs to be optimized separately for each different system.[48] For the two-body and three-body dispersion energy in HMBI, the damping parameter β is determined empirically from atomic van der Waals radii.[127]

E. Availability

This section summarizes the availability of the various *ab initio* force field methods discussed in this perspective (Table II). The EFP and EFMO methods are implemented and continuously developed and improved in the freely available

TABLE II. Summary of availabilities of the *ab initio* force field methods.

Method	Software package
EFP	GAMESS (free)
libefp	Q-Chem (commercial), PSI4 (free, open-source), NWChem (free, open-source)
SIBFA	Home-made, TINKER (contact the developers)
EFMO	GAMESS
HMBI	Home-made (contact the developers)

quantum chemistry package GAMESS. Some components of the EFP method have been implemented in *libefp*, an open-source portable software library.[128] EFPs of many common solvent molecules are part of the *libefp* library, in the same format as the potentials generated by GAMESS. This means that *libefp* can also recognize the potentials generated by GAMESS. A program called *efpmd*, built on the top of *libefp*, is capable of single-point energy and gradient, seminumerical Hessian, normal mode analysis, and molecular dynamics in microcanonical ensemble (NVE), canonical ensemble (NVT), and isobaric-isothermal ensemble (NPT). Currently *libefp* is interfaced with GAMESS, the *PSI*4 package, the *Q-chem* commercial software, and the *NWChem* package. Note that currently, *libefp* does not contain the charge transfer interaction implemented in GAMESS, and the QM/EFP exchange repulsion[129] and dispersion[114] interactions are different from those implemented in GAMESS.

The HMBI method is implemented in a homemade software package by the developers. This software generates input for QM calculations using Q-Chem. MM parameters are generated using CamCASP.[130] SIBFA is also developed as an in-house package (Table II).

V. TYPES OF APPLICATIONS AND ACCURACY

Given the varied interest of developers, different *ab initio* force fields possess different features pertaining to the systems of interest to the developers. The EFP method was initiated as an accurate model potential to capture aqueous solvent effects; later it was expanded to be applicable to any materials where intermolecular interactions are crucial. The EFMO method, as it combines the FMO and EFP, is designed

to describe large molecular systems, such as ions solvated by water, ionic liquids, protein structure (in the presence of a solvent, such as water), and heterogeneous catalysts. The SIBFA method has been used in many applications[131–135] for biologically and pharmacologically relevant molecules involving metals. The HMBI method, with its implementation of multipolar Ewald summation and space-group symmetry, is a robust method for predicting molecular crystals. In this section, some applications of each *ab initio* FF method are mentioned.

A. EFP

A recent EFP study by Slipchenko and co-workers demonstrated excellent accuracy against CCSD(T)/CBS interaction energies for both S22 and S66 data sets, with the mean unsigned error of EFP interaction energies being 0.9 and 0.6 kcal/mol for S22 and S66, respectively.[136] Generally, the EFP achieves the accuracy of MP2 and SCS-MP2 for intermolecular interactions.[136] In Ref. 136, the charge transfer interaction was omitted due to the relative expense of that term. Recently, the computational cost of charge transfer has been significantly reduced.[66] A new study has included the CT term and achieved a mean unsigned error of 0.72 kcal/mol for the S22 data set.[137]

The EFP method has been used to study sections of the potential energy surface of styrene dimers,[138] benzene dimers and substituted benzene dimers,[139] benzene-pyridine,[140] and DNA bases.[141] Combined with Monte Carlo or molecular dynamics simulations, the EFP has also been used to study several molecular clusters such as methanol-water mixtures,[142] water-benzene complexes,[143] solvated alanine,[144] solvated hydronium ions,[145] and solvated nitrate ions.[146] Over the years, the EFP has repeatedly demonstrated that it predicts structures and energetics in good agreement with MP2 and in some cases has achieved CCSD(T) quality.[136,139,140]

B. SIBFA

The development of SIBFA aims at biologically relevant systems, in particular metallic complexes and metalloproteins.[131–135,147] SIBFA has also been used to model lanthanide and actinide complexes.[148] The developers tested the SIBFA method and its derivatives described in Secs. II B and II C (GEM, S/G-1) by comparing individual energy contributions [Eq. (14)] to reference *ab initio* computations. In particular, E_{MTP} and E_{rep} as well as E_{pol} and E_{ct} were compared to the values obtained with an energy decomposition analysis (RVS, for example) performed at the Hartree-Fock level. The dispersion energy was compared to the correlation energy obtained at a higher level such as MP2. Overall, the method has good accuracy for systems containing metallic cations at a low computational cost. In particular, the S/G-1 method presents a promising avenue to model heavy metal ions like mercury, where relativistic effects are important.[74]

C. EFMO

The early development of EFMO only included EFP Coulomb and polarization interaction. The utility of the early version was demonstrated through an application to chorismate mutase.[149] The current version of EFMO contains all five types of intermolecular interactions used in the EFP, as

explained in Sec. IV.[107] EFMO performs better than FMO2 and in some cases better than FMO3.[107] The recent implementation of the EFMO analytic gradient shows significant speed up compared to the FMO2 computation.[150] EFMO is expected to provide MP2-quality results for large biomolecules in MD simulations.

D. HMBI

For small to medium sized molecular clusters (≤20 molecules) such as water, formamide, hydrogen fluoride, and mixed glycine-water, the errors relative to RI-MP2 are all less than 0.24 kcal/mol, with an optimized damping parameter.[48] Such accuracy is achieved without implementing many-body dispersion. Adding the Axilrod-Teller-Muto 3-body dispersion, HMBI demonstrated excellent agreement with experiments for crystal lattice energies of several molecular crystals.[45] Polymorphism in molecular crystals is common, and being able to predict the correct energy ordering of polymorphic molecular crystals has extremely important consequences in fields such as pharmaceutical industry.[151] HMBI has shown its usefulness in such applications. One example is crystalline aspirin.[152] Contrary to DFT calculations, which predict that form II is 2-2.5 kJ/mol more stable than form I, HMBI predicted "accidental" degeneracy of forms I and II, which is consistent with experimentally observed intergrowth structures. Another example is that while DFT predicted an incorrect (when compared to experiment) space group for the phase XV of ice, HMBI predicted structures that agreed with experiments.[153]

VI. FORMAL COMPUTATIONAL COST OF THE METHODS

The formal computational cost of the EFP for fragment-fragment calculations is N^2, where N is the number of fragments.[86] The Coulomb and dispersion terms between fragments are the least expensive to compute. The computational cost of the polarization and exchange-repulsion terms is twice and five times as high as these two, respectively. The many-body polarization in SIBFA and EFP is computed using the induced dipole approach, which can be solved directly or iteratively. The computational cost of iterative methods is the cost of one iteration times the number of iterations. These methods are preferable since the computational cost of the direct methods grows with N^3. Iterative methods include Jacobi (used in EFP) and conjugate gradient methods.[154] These methods require tight convergence criteria in order to maintain energy conservation during MD simulations. The truncated preconditioned conjugate gradient (TPCG) method developed by Piquemal *et al.* provides a robust way to avoid this pitfall at virtually no cost.[155] The bottleneck in EFP calculations is the charge-transfer energy due to the computation of a large number of kinetic and potential energy integrals between MOs, which is ultimately caused by the large number of virtual MOs. This term costs 50 times more than the Coulomb and dispersion terms when using canonical virtual MOs. Computational time-savings of 50% or more were obtained by using valence virtual orbitals (VVOs) instead of canonical

virtual MOs.[66] For systems where charge transfer interaction is small, this term is sometimes omitted altogether.[136] In contrast, in SIBFA, charge transfer interactions are only computed between hydrogen-bonded fragments and an empirically parametrized function is used. This term is therefore much less expensive to compute in SIBFA. On the other hand, the electrostatic and exchange-repulsion terms computed using density fitting as in GEM require the evaluation of integrals between fitted densities. As a result, the evaluation of these terms using GEM is noticeably less efficient than that in the original SIBFA method, which uses the Vigné-Maeder and Claverie scheme described in Sec. III B. In order to accelerate the computation of these energy terms, the SIBFA developers compute part of these integrals in the reciprocal space using the particle-mesh Ewald (PME)[156] and fast Fourier-Poisson (FFP) methods,[157] reducing the cost from N^2 in the direct space to $N \log N$.[81] Subsequently, the Smooth Particle Mesh Ewald (SPME)[158] approach was implemented to generate distributed multipoles to compute the electrostatic term in the SBIFA method.[159]

As explained previously, EFMO interfaces the effective fragment potential method with the fragment molecular orbital (FMO) method. The FMO2 method formally scales as N^2, without using any approximations, which is 1000 times faster than the N^3 scaling of the self-consistent field (SCF) procedure.[160] Since the most recent FMO implementation essentially eliminates the I/O overhead,[161] the actual FMO scaling is essentially linear. The cost of the EFP part of the computation in EFMO is similar to the formal EFP method, discussed previously, in addition to the cost of generating the potential. Because of the ability of fragmentation methods to take advantage of multi-level parallelism, *ab initio* methods such as FMO and EFMO can make use of hundreds of thousands of cores. Consequently, these methods can take advantage of petascale, and eventually exascale, computers and therefore be applied to systems with thousands of atoms.

VII. SUMMARY AND OUTLOOK

Ab initio force fields have been applied to a wide range of systems, from molecular clusters bound by weak dispersion and hydrogen-bonded networks to molecular crystals, to covalently linked peptides and DNA chains. Compared to non-*ab initio* force field methods, *ab initio* force fields provide more fundamental and sophisticated potentials that are more generally applicable. The fragmentation approach allows much larger systems to be studied compared to traditional *ab initio* methods. With regard to accuracy, *ab initio* force fields generally achieve at least MP2-level results, and some even compare well with coupled cluster results. In addition, insightful energy decomposition is a common feature for these *ab initio* force fields.

One expects new developments that enable dynamically defining fragments to be a highly desirable feature. This feature will allow processes such as crack propagation, surface adsorption, and rearrangement, and many other chemical reactions to be accessible by *ab initio* force fields. These types of applications will also benefit greatly from another desirable feature of *ab initio* force fields: being able to deal with systems that have periodic and non-periodic components simultaneously.

[1] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).

[2] L. Nilsson and M. Karplus, J. Comput. Chem. **7**, 591 (1986).

[3] P. K. Weiner and P. A. Kollman, J. Comput. Chem. **2**, 287 (1981).

[4] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner, J. Am. Chem. Soc. **106**, 765 (1984).

[5] S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, J. Comput. Chem. **7**, 230 (1986).

[6] W. L. Jorgensen and C. J. Swenson, J. Am. Chem. Soc. **107**, 569 (1985).

[7] W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657 (1988).

[8] W. L. Jorgensen, *Encyclopedia of Computational Chemistry* (John Wiley & Sons, Ltd., 2002).

[9] J.-H. Lii and N. L. Allinger, J. Comput. Chem. **12**, 186 (1991).

[10] N. L. Allinger, Y. H. Yuh, and J. H. Lii, J. Am. Chem. Soc. **111**, 8551 (1989).

[11] J. H. Lii and N. L. Allinger, J. Am. Chem. Soc. **111**, 8566 (1989).

[12] J. H. Lii and N. L. Allinger, J. Am. Chem. Soc. **111**, 8576 (1989).

[13] M. J. Hwang, T. P. Stockfisch, and A. T. Hagler, J. Am. Chem. Soc. **116**, 2515 (1994).

[14] J. R. Maple, M. J. Hwang, T. P. Stockfisch, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler, J. Comput. Chem. **15**, 162 (1994).

[15] C. M. Baker, V. M. Anisimov, and A. D. MacKerell, J. Phys. Chem. B **115**, 580 (2011).

[16] S. Patel and C. L. Brooks, J. Comput. Chem. **25**, 1 (2004).

[17] S. Patel, A. D. Mackerell, and C. L. Brooks, J. Comput. Chem. **25**, 1504 (2004).

[18] P. Cieplak, J. Caldwell, and P. Kollman, J. Comput. Chem. **22**, 1048 (2001).

[19] Z. X. Wang, W. Zhang, C. Wu, H. Lei, P. Cieplak, and Y. Duan, J. Comput. Chem. **27**, 781 (2006).

[20] J. L. Banks, G. A. Kaminski, R. Zhou, D. T. Mainz, B. J. Berne, and R. A. Friesner, J. Chem. Phys. **110**, 741 (1999).

[21] P. Ren and J. W. Ponder, J. Phys. Chem. B **107**, 5933 (2003).

[22] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, J. Chem. Theory Comput. **9**, 4046 (2013).

[23] T. A. Halgren and W. Damm, Curr. Opin. Struct. Biol. **11**, 236 (2001).

[24] H. S. Antila and E. Salonen, in *Biomolecular Simulations: Methods and Protocols*, edited by L. Monticelli and E. Salonen (Humana Press, Totowa, NJ, 2013), p. 215.

[25] A. Warshel, M. Kato, and A. V. Pisliakov, J. Chem. Theory Comput. **3**, 2034 (2007).

[26] P. Xu, J. Wang, Y. Xu, H. Chu, J. Liu, M. Zhao, D. Zhang, Y. Mao, B. Li, Y. Ding *et al.*, in *Advance in Structural Bioinformatics*, edited by D. Wei *et al.* (Springer Netherlands, Dordrecht, 2015), p. 19.

[27] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, J. Am. Chem. Soc. **114**, 10024 (1992).

[28] M. A. Addicoat, N. Vankova, I. F. Akter, and T. Heine, J. Chem. Theory Comput. **10**, 880 (2014).

[29] M. J. S. Phipps, T. Fox, C. S. Tautermann, and C.-K. Skylaris, Chem. Soc. Rev. **44**, 3177 (2015).

[30] K. Morokuma and K. Kitaura, in *Chemical Applications of Atomic and Molecular Electrostatic Potentials: Reactivity, Structure, Scattering, and Energetics of Organic, Inorganic, and Biological Systems*, edited by P. Politzer and D. G. Truhlar (Springer US, Boston, MA, 1981), p. 215.

[31] Y. Mo, J. Gao, and S. D. Peyerimhoff, J. Chem. Phys. **112**, 5530 (2000).

[32] R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell, and M. Head-Gordon, J. Phys. Chem. A **111**, 8753 (2007).

[33] E. D. Glendening and A. Streitwieser, J. Chem. Phys. **100**, 2900 (1994).

[34] P. S. Bagus, K. Hermann, and C. W. Bauschlicher, J. Chem. Phys. **80**, 4378 (1984).

[35] W. J. Stevens and W. H. Fink, Chem. Phys. Lett. **139**, 15 (1987).

[36] P. Su and H. Li, J. Chem. Phys. **131**, 014102 (2009).

[37] P. Jurečka, J. Šponer, J. Černy, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).

[38] M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, Chem. Rev. **112**, 632 (2012).

[39]S. R. Pruitt, C. Bertoni, K. R. Brorsen, and M. S. Gordon, Acc. Chem. Res. **47**, 2786 (2014).

[40]M. A. Collins and R. P. A. Bettens, Chem. Rev. **115**, 5607 (2015).

[41]P. N. Day, R. Pachter, M. S. Gordon, and G. N. Merrill, J. Chem. Phys. **112**, 2063 (2000).

[42]N. Gresh, P. Claverie, and A. Pullman, Theor. Chim. Acta **66**, 1 (1984).

[43]J.-P. Piquemal, H. Chevreau, and N. Gresh, J. Chem. Theory Comput. **3**, 824 (2007).

[44]C. Steinmann, D. G. Fedorov, and J. H. Jensen, J. Phys. Chem. A **114**, 8705 (2010).

[45]S. Wen and G. J. O. Beran, J. Chem. Theory Comput. **7**, 3733 (2011).

[46]G. J. O. Beran and K. Nanda, J. Phys. Chem. Lett. **1**, 3480 (2010).

[47]G. J. O. Beran, J. Chem. Phys. **130**, 164115 (2009).

[48]A. Sebetci and G. J. O. Beran, J. Chem. Theory Comput. **6**, 155 (2010).

[49]S. Wen, K. Nanda, Y. Huang, and G. J. O. Beran, Phys. Chem. Chem. Phys. **14**, 7578 (2012).

[50]A. Stone, *The Theory of Intermolecular Forces*, 2nd ed. (Oxford University Press, Oxford, 2013).

[51]P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, and D. Cohen, J. Chem. Phys. **105**, 1968 (1996).

[52]I. Adamovic, M. A. Freitag, and M. S. Gordon, J. Chem. Phys. **118**, 6725 (2003).

[53]M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, J. Phys. Chem. A **105**, 293 (2000).

[54]J. H. Jensen, J. Chem. Phys. **114**, 8775 (2001).

[55]A. J. Stone, Chem. Phys. Lett. **83**, 233 (1981).

[56]A. J. Stone and M. Alderton, Mol. Phys. **56**, 1047 (1985).

[57]H. Li, H. M. Netzloff, and M. S. Gordon, J. Chem. Phys. **125**, 194103 (2006).

[58]I. Adamovic and M. S. Gordon, Mol. Phys. **103**, 379 (2005).

[59]P. Xu, F. Zahariev, and M. S. Gordon, J. Chem. Theory Comput. **10**, 1576 (2014).

[60]E. B. Guidez, P. Xu, and M. S. Gordon, J. Phys. Chem. A **120**, 639 (2016).

[61]D. R. Williams, L. J. Schaad, and J. N. Murrell, J. Chem. Phys. **47**, 4916 (1967).

[62]J. N. Murrell, M. Randic, and D. R. Williams, Proc. R. Soc. London, Ser. A **284**, 566 (1965).

[63]P. W. Fowler and A. J. Sadlej, Mol. Phys. **77**, 709 (1992).

[64]H. Li, M. S. Gordon, and J. H. Jensen, J. Chem. Phys. **124**, 214108 (2006).

[65]J. H. Jensen, J. Chem. Phys. **104**, 7795 (1996).

[66]P. Xu and M. S. Gordon, J. Chem. Phys. **139**, 194104 (2013).

[67]W. C. Lu, C. Z. Wang, M. W. Schmidt, L. Bytautas, K. M. Ho, and K. Ruedenberg, J. Chem. Phys. **120**, 2629 (2004).

[68]N. Gresh, P. Claverie, and A. Pullman, Int. J. Quantum Chem. **22**, 199 (1982).

[69]N. Gresh, A. Pullman, and P. Claverie, Int. J. Quantum Chem. **28**, 757 (1985).

[70]N. Gresh, P. Claverie, and A. Pullman, Int. J. Quantum Chem. **16**, 243 (1979).

[71]F. Vigné-Maeder and P. Claverie, J. Chem. Phys. **88**, 4934 (1988).

[72]J.-P. Piquemal, N. Gresh, and C. Giessner-Prettre, J. Phys. Chem. A **107**, 10353 (2003).

[73]N. Gresh, J. Comput. Chem. **16**, 856 (1995).

[74]R. Chaudret, N. Gresh, C. Narth, L. Lagardère, T. A. Darden, G. A. Cisneros, and J.-P. Piquemal, J. Phys. Chem. A **118**, 7598 (2014).

[75]N. Gresh, J.-P. Piquemal, and M. Krauss, J. Comput. Chem. **26**, 1113 (2005).

[76]N. Gresh, P. Claverie, and A. Pullman, Int. J. Quantum Chem. **29**, 101 (1986).

[77]N. Gresh, J. Phys. Chem. A **101**, 8680 (1997).

[78]S. Creuzet, J. Langlet, and N. Gresh, J. Chim. Phys. **88**, 2399 (1991).

[79]D. J. Willock, S. L. Price, M. Leslie, and C. R. A. Catlow, J. Comput. Chem. **16**, 628 (1995).

[80]G. A. Cisneros, J.-P. Piquemal, and T. A. Darden, J. Chem. Phys. **123**, 044109 (2005).

[81]G. A. Cisneros, J.-P. Piquemal, and T. A. Darden, J. Chem. Phys. **125**, 184101 (2006).

[82]J.-P. Piquemal, G. A. Cisneros, P. Reinhardt, N. Gresh, and T. A. Darden, J. Chem. Phys. **124**, 104101 (2006).

[83]R. E. Duke, O. N. Starovoytov, J.-P. Piquemal, and G. A. Cisneros, J. Chem. Theory Comput. **10**, 1361 (2014).

[84]W. Xie and J. Gao, J. Chem. Theory Comput. **3**, 1890 (2007).

[85]W. Xie, L. Song, D. G. Truhlar, and J. Gao, J. Phys. Chem. B **112**, 14124 (2008).

[86]M. S. Gordon, Q. A. Smith, P. Xu, and L. V. Slipchenko, Annu. Rev. Phys. Chem. **64**, 553 (2013).

[87]L. D. Jacobson and J. M. Herbert, J. Chem. Phys. **134**, 094118 (2011).

[88]J. M. Herbert, L. D. Jacobson, K. Un Lao, and M. A. Rohrdanz, Phys. Chem. Chem. Phys. **14**, 7679 (2012).

[89]J. R. Schmidt, K. Yu, and J. G. McDaniel, Acc. Chem. Res. **48**, 548 (2015).

[90]J. G. McDaniel and J. R. Schmidt, J. Phys. Chem. B **118**, 8042 (2014).

[91]J. G. McDaniel and J. R. Schmidt, J. Phys. Chem. A **117**, 2053 (2013).

[92]M. J. Van Vleet, A. J. Misquitta, A. J. Stone, and J. R. Schmidt, J. Chem. Theory Comput. **12**, 3851 (2016).

[93]A. J. Misquitta and A. J. Stone, J. Chem. Theory Comput. **12**, 4184 (2016).

[94]A. J. Stone and A. J. Misquitta, Int. Rev. Phys. Chem. **26**, 193 (2007).

[95]S. Grimme, J. Chem. Theory Comput. **10**, 4497 (2014).

[96]S. Grimme, C. Bannwarth, E. Caldeweyher, J. Pisarek, and A. Hansen, J. Chem. Phys. **147**, 161708 (2017).

[97]K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999).

[98]T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, and K. Kitaura, Chem. Phys. Lett. **351**, 475 (2002).

[99]V. Deev and M. A. Collins, J. Chem. Phys. **122**, 154102 (2005).

[100]M. A. Collins and V. A. Deev, J. Chem. Phys. **125**, 104104 (2006).

[101]M. A. Addicoat and M. A. Collins, J. Chem. Phys. **131**, 104103 (2009).

[102]D. W. Zhang and J. Z. H. Zhang, J. Chem. Phys. **119**, 3599 (2003).

[103]X. H. Chen and J. Z. H. Zhang, J. Chem. Phys. **120**, 11386 (2004).

[104]S. R. Gadre, R. N. Shirsat, and A. C. Limaye, J. Phys. Chem. **98**, 9165 (1994).

[105]K. Babu and S. R. Gadre, J. Comput. Chem. **24**, 484 (2003).

[106]D. G. Fedorov and K. Kitaura, Chem. Phys. Lett. **433**, 182 (2006).

[107]S. R. Pruitt, C. Steinmann, J. H. Jensen, and M. S. Gordon, J. Chem. Theory Comput. **9**, 2235 (2013).

[108]Y. Muto, Proc. Phys. Math. Soc. Jpn. **17**, 629 (1943).

[109]B. M. Axilrod and E. Teller, J. Chem. Phys. **11**, 299 (1943).

[110]K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).

[111]Y. Heit and G. J. O. Beran, J. Comput. Chem. **35**, 2205 (2014).

[112]A. W. Lange and G. A. Voth, J. Chem. Theory Comput. **9**, 4018 (2013).

[113]Q. A. Smith, K. Ruedenberg, M. S. Gordon, and L. V. Slipchenko, J. Chem. Phys. **136**, 244107 (2012).

[114]L. V. Slipchenko, M. S. Gordon, and K. Ruedenberg, J. Phys. Chem. A **121**, 9495 (2017).

[115]J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).

[116]S. Miertuš, E. Scrocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).

[117]S. Miertuš and J. Tomasi, Chem. Phys. **65**, 239 (1982).

[118]M. A. Aguilar, F. J. O. den Valle, and J. Tomasi, J. Chem. Phys. **98**, 7375 (1993).

[119]J. L. Pascual-Ahuir, E. Silla, J. Tomasi, and R. Bonaccorsi, J. Comput. Chem. **8**, 778 (1987).

[120]P. Bandyopadhyay and M. S. Gordon, J. Chem. Phys. **113**, 1104 (2000).

[121]P. Bandyopadhyay, M. S. Gordon, B. Mennucci, and J. Tomasi, J. Chem. Phys. **116**, 5023 (2002).

[122]H. Li, C. S. Pomelli, and J. H. Jensen, Theor. Chem. Acc. **109**, 71 (2003).

[123]J. Langlet, P. Claverie, J. Caillet, and A. Pullman, J. Chem. Phys. **92**, 1617 (1988).

[124]J. Langlet, N. Gresh, and C. Giessner-Prettre, Biopolymers **36**, 765 (1995).

[125]N. Gresh, G. Tiraboschi, and D. R. Salahub, Biopolymers **45**, 405 (1998).

[126]L. V. Slipchenko and M. S. Gordon, Mol. Phys. **107**, 999 (2009).

[127]O. Anatole von Lilienfeld and A. Tkatchenko, J. Chem. Phys. **132**, 234109 (2010).

[128]I. A. Kaliman and L. V. Slipchenko, J. Comput. Chem. **34**, 2284 (2013).

[129]C. Viques-Rojas, J. A. Fine, and L. V. Slipchenko, "Exchange repulsion in QM/EFP" (unpublished).

[130]A. J. Misquitta and A. J. Stone, http://www-stone.ch.cam.ac.uk/programs.html#CamCASP.

[131]N. Gresh, C. Policar, and C. Giessner-Prettre, J. Phys. Chem. A **106**, 5660 (2002).

[132]D. R. Garmer, N. Gresh, and B.-P. Roques, Proteins: Struct., Funct., Bioinf. **31**, 42 (1998).

[133]F. Rogalewicz, G. Ohanessian, and N. Gresh, J. Comput. Chem. **21**, 963 (2000).

[134]B. de Courcy, J.-P. Dognon, C. Clavaguéra, N. Gresh, and J.-P. Piquemal, Int. J. Quantum Chem. **111**, 1213 (2011).

[135]L. M. M. Jenkins, T. Hara, S. R. Durell, R. Hayashi, J. K. Inman, J.-P. Piquemal, N. Gresh, and E. Appella, J. Am. Chem. Soc. **129**, 11067 (2007).

[136]J. C. Flick, D. Kosenkov, E. G. Hohenstein, C. D. Sherrill, and L. V. Slipchenko, J. Chem. Theory Comput. **8**, 2835 (2012).

[137]C. Bertoni, L. V. Slipchenko, A. J. Misquitta, and M. S. Gordon, J. Phys. Chem. A **121**, 2056 (2017).

[138]I. Adamovic, H. Li, M. H. Lamm, and M. S. Gordon, J. Phys. Chem. A **110**, 519 (2006).

[139]T. Smith, L. V. Slipchenko, and M. S. Gordon, J. Phys. Chem. A **112**, 5286 (2008).

[140]Q. A. Smith, M. S. Gordon, and L. V. Slipchenko, J. Phys. Chem. A **115**, 4598 (2011).

[141]Q. A. Smith, M. S. Gordon, and L. V. Slipchenko, J. Phys. Chem. A **115**, 11269 (2011).

[142]I. Adamovic and M. S. Gordon, J. Phys. Chem. A **110**, 10267 (2006).

[143]L. V. Slipchenko and M. S. Gordon, J. Phys. Chem. A **113**, 2092 (2009).

[144]J. M. Mullin and M. S. Gordon, J. Phys. Chem. B **113**, 8657 (2009).

[145]K. R. Brorsen, S. R. Pruitt, and M. S. Gordon, J. Phys. Chem. B **118**, 14382 (2014).

[146]S. R. Pruitt, K. R. Brorsen, and M. S. Gordon, Phys. Chem. Chem. Phys. **17**, 27027 (2015).

[147]N. Gresh and J. Šponer, J. Phys. Chem. B **103**, 11415 (1999).

[148]A. Marjolin, C. Gourlaouen, C. Clavaguéra, P. Y. Ren, J. C. Wu, N. Gresh, J.-P. Dognon, and J.-P. Piquemal, Theor. Chem. Acc. **131**, 1198 (2012).

[149]C. Steinmann, D. G. Fedorov, and J. H. Jensen, PLoS One **8**, e60602 (2013).

[150]C. Bertoni and M. S. Gordon, J. Chem. Theory Comput. **12**, 4743 (2016).

[151]G. J. O. Beran, Chem. Rev. **116**, 5567 (2016).

[152]S. Wen and G. J. O. Beran, Cryst. Growth Des. **12**, 2169 (2012).

[153]K. D. Nanda and G. J. O. Beran, J. Phys. Chem. Lett. **4**, 3165 (2013).

[154]Y. Saad, *Iterative Methods for Sparse Linear Systems* (SIAM, 2003).

[155]F. Aviat, A. Levitt, B. Stamm, Y. Maday, P. Ren, J. W. Ponder, L. Lagardère, and J.-P. Piquemal, J. Chem. Theory Comput. **13**, 180 (2017).

[156]T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).

[157]D. York and W. Yang, J. Chem. Phys. **101**, 3298 (1994).

[158]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).

[159]C. Narth, L. Lagardère, É. Polack, N. Gresh, Q. Wang, D. R. Bell, J. A. Rackers, J. W. Ponder, P. Y. Ren, and J.-P. Piquemal, J. Comput. Chem. **37**, 494 (2016).

[160]D. Fedorov and K. Kitaura, *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems* (CRC Press, 2009).

[161]S. R. Pruitt, H. Nakata, T. Nagata, M. Mayes, Y. Alexeev, G. Fletcher, D. G. Fedorov, K. Kitaura, and M. S. Gordon, J. Chem. Theory Comput. **12**, 1423 (2016).