# The Description of Many-Body Energies and Forces for Water and Ions

Joseph P. Heindel,[†,‡] Selim Sami,[†] and Teresa Head-Gordon[*,†,‡,¶]

†*Kenneth S. Pitzer Theory Center and Department of Chemistry, University of California, Berkeley, California 94720, United States*

‡*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

¶*Departments of Bioengineering and Chemical and Biomolecular Engineering, University of California, Berkeley, California 94720, United States*

E-mail: thg@berkeley.edu

**Abstract**

Solvated atomic ions such as alkaline metals and halides play critical roles in the regulation of biological systems, environmental chemistry, and energy materials. This work describes an advanced force field for water and atomic ions which analyzes the many-body energy for non-additivity and to quantitatively reproduce each term of an energy decomposition analysis to achieve accuracy and transferability. This is achieved through the introduction of new many-body models for polarization and charge transfer in the formulation of a very high quality water force field, which generalizes to provide accurate dimer surfaces and three-body polarization and charge transfer for solvated alkali metal cations and halide anions. We also utilize a new one-body potential that accounts for intramolecular polarization, and includes an electric field-dependent correction to the Morse potential, that dramatically improves the accuracy of forces as well as frequencies for hydrogen-bonded O−H stretches.

# Introduction

Force fields (FFs) are approximations to the quantum mechanical potential energy surface (PES), in which the model design goal is to predict structure, dynamics, and thermodynamics of any molecular system. Although pairwise additive FFs remain popular due to their insight and computational efficiency, they are inherently limited by their inability to describe many-body contributions, thereby reducing their accuracy in property predictions and transferability to new chemistry. For example, it has long been appreciated that non-additive interactions are critical for describing hydrogen-bonded interactions in water,[1] and the many-body energy contributions for water interacting with ions is quite large, typically around 15% of the non-bonded energy, and the cooperativity effects can be either stabilizing or destabilizing depending on ion type.[2,3] And yet many simulations involving ions or ionized molecules are carried out by treating the ion as a scaled point charge with simple functional forms for dispersion and repulsion that do not describe such complexity.[4,5]

Recently there has been a paradigm shift in non-reactive many-body FF development by combining energy decomposition analysis (EDA)[?] with the many-body expansion (MBE)[6] to better control the accuracy and transferability of advanced FF models. The MBE of the non-bonded energy[6] allows one to gain insights into how much non-additivity is present in the quantum mechanical energy (and forces), while the EDA decomposes the total energy (and forces) into physically motivated contributions from pairwise additive terms such as electrostatics with charge penetration and Pauli repulsion, and many-body contributions that arise from polarization, charge transfer, and dispersion. For example, our recent (many-body) MB-UCB force field for water[7] and extensions to monovalent and divalent alkali metal and halide ions describe ion-water interactions through a systematic buildup of 2- and 3-body interactions whose molecular interactions are formulated using the ALMO-EDA energy decomposition scheme.[8] By reproducing the EDA term-by-term, we can ensure that the force field will be transferable across the phase diagram of a homogeneous system and, ideally, to new heterogenous systems. They also provide chemical insights in the relative weights of

particular energy terms, for instance, charge transfer versus polarization, and better separation of the Pauli repulsion and dispersion energies. Similar many-body potentials for water and ions include AMOEBA+, HIPPO,[9] MB-Pol, and q-Aqua that strive for higher accuracy but differ in their use of EDA, MBE, and most critically, their functional forms used for the many-body energy and forces.

Our goal in this paper is to develop a new class of many-body force fields which is able to quantitatively reproduce all of the terms in an energy decomposition,[8,10] with better functional forms for the many-body polarization and charge transfer contributions. Historically, there have been two main approaches to including polarization in force fields: fluctuating charges[11] and induced dipoles.[12] There have also been attempts to unify these approaches allowing for both charge rearrangements and induced dipoles[13] in order to also capture the charge transfer interaction.[11] With the advent of variational EDA techniques,[8,10] it is now clear that the charge transfer energy scales exponentially and hence cannot be modelled by fluctuating charges alone.

This work takes a large step forwards in the construction of many-body force fields capable of handling very strong interactions such as those between water and ions, which has remained elusive and is an open research problem dating back decades. We use the ALMO-EDA method[? ?] to separate the total non-bonded interaction energy into individual contributions

$$E_{\text{int}} = E_{\text{elec}} + E_{\text{Pauli}} + E_{\text{disp}} + E_{\text{pol}} + E_{\text{CT}} \tag{1}$$

where $E_{\text{elec}}$, $E_{\text{Pauli}}$, $E_{\text{disp}}$, $E_{\text{pol}}$, and $E_{\text{CT}}$ correspond to the contributions from the permanent electrostatics, Pauli repulsion, dispersion, polarization and charge transfer, respectively. We introduce a new approach to polarization that allows for both intramolecular charge fluctuations and induced dipoles, as well as a new approach for modelling charge transfer to allow for explicit movement of charge between molecules, and which naturally describes many-body charge transfer by coupling into the polarization equations. We also eliminate the need to treat intramolecular and intermolecular polarization separately through a recently

reported one-body potential[?] but modified with a field-dependent correction to the Morse potential that requires zero free parameters, thereby dramatically improving the accuracy of polarization forces. The resulting FQCT (fluctuating charges and charge transfer) model shows excellent accuracy against EDA data, and additional validation data for water clusters and ion-water energies, and reproduces the structure-frequency correspondence expected for hydrogen-bonded O−H stretches.

# Theory and Methods

The FQCT force field described in this work will model each of the terms in Eq. 1. ALMO-EDA is described elsewhere,[? ?] but we provide two important clarifications here. First the most appropriate choice for $E_{\text{elec}}$ is the quasi-classical expression, which depends only on the geometry of individual monomers,[?] and we use the fragment electric response function approach (at the dipole+quadrupole level) to evaluate the polarization, ensuring a well-defined basis set limit.[?] All the ALMO-EDA calculations are performed at the level of $\omega$B97X-V DFT functional[?] and def2-QZVPPD basis set[?] using the Q-Chem software package.[?] Note that we will use a convention of referring to all energy terms in the force field with a $V$ and all energy terms from electronic structure with an $E$.

Our approach adopts ideas from the density overlap hypothesis[14–18] which states that the short-range contributions to intermolecular interactions is proportional to the electron density overlap. In order for this idea to be amenable to force fields, one must use overlaps of atom-centered densities. The form of the charge density used in our model is,

$$\rho(r) = \frac{Qb^3}{8\pi} e^{-br} + Z\delta(r) \tag{2}$$

where $Q$ is the charge associated with the model electron density, $Z$ is the effective nuclear charge of the atom, and $b$ defines the width of the Slater density. The delta function, $\delta(r)$, means the core is treated as a point particle.

One can show that the overlap, $S^{\rho}_{ii}$, of two identical Slater-like atomic densities at different locations, $\rho_i(\boldsymbol{r}_i)$ and $\rho_i(\boldsymbol{r}_j)$, is,

$$S^{\rho}_{ii} = \frac{\pi D^2}{b^3_{ii}} P(b_{ii} r_{ij}) \exp(-b_{ii} r_{ij}) \tag{3}$$

The above overlap expression is only strictly true for the exponential tail of the Slater density and for identical atoms. The overlap between atoms with different densities, $S^{ij}_{\rho}$, has a more complicated form, but it has been shown that setting $b_{ij} = \sqrt{b_i b_j}$ allows the expression for $S^{\rho}_{ii}$ to be used for different atom types to a good approximation.[17] The polynomial prefactor in the overlap is,

$$P(b_{ij} r_{ij}) = \frac{1}{3}(b_{ij} r_{ij})^2 + b_{ij} r_{ij} + 1 \tag{4}$$

where, again, we will use the combination rule $b_{ij} = \sqrt{b_i b_j}$.

Approaches to density overlap have been developed by Misquitta and others[19,20] based on iterated stockholder atoms which can be used to define Slater-like densities for atoms in molecules, as well as by van Vleet *et al.* in the MASTIFF force field.[17,18] Rackers *et al.* utilize a similar idea in the HIPPO model[9] but rather than relying on density overlap, they treat the Slater function as an orbital and are able to derive models of Pauli repulsion, charge penetration, and even dispersion. Because HIPPO is derived from a model orbital, the damping functions which prevent singularities in various short-range energetic contributions arise naturally. We find the HIPPO approach to both Pauli repulsion and electrostatics to be physically principled and utilize them here without significant modification. In addition, we utilize both Slater density overlap and Slater orbital overlap in the other energy terms within the FQCT model which distinguishes the degree of what is "short-range" for a pair of atoms.

## Electrostatics, Pauli, and Dispersion Energies

**Permanent Electrostatics.** Our description of electrostatics comes from a traditional point multipole approach up to the quadrupoles, and a charge penetration (CP) contribution that modifies the short-range electrostatic energy to be more attractive than the point multipole expansion alone. We isolate the CP energy by taking the total classical electrostatic energy from EDA minus the point multipole interaction energy when using Stone's distributed multipole analysis (DMA)[21,22] out to hexadecapoles on all atoms.

$$E^{CP} = E^{elec}_{EDA} - E^{elec}_{DMA} \tag{5}$$

The advantage of this approach is it allows us to ensure that our multipoles are not biased to compensate for error in the description of charge penetration, and vice versa, which is essential to reproduce the classical electrostatic energy in EDA.

CP is described by treating each atom as having both a positively charged core and negatively charged shell. Considering the interactions of the collection of cores and shells, which are expanded in multipoles, results in the following electrostatic energy expression:

$$V_{elec} = \sum_{i<j} Z_i T_{ij} Z_j + Z_i \boldsymbol{T}^{damp}_{ij} \boldsymbol{M}_j + Z_j \boldsymbol{T}^{damp}_{ji} \boldsymbol{M}_i + \boldsymbol{M}_i \boldsymbol{T}^{overlap}_{ij} \boldsymbol{M}_j \tag{6}$$

where the first term represents repulsive core-core interactions where $T_{ij} = 1/r_{ij}$ with $Z_i$ the core charge on the $i$th atom; note that this is not the nuclear charge but an effective nuclear charge. The second and third terms describe attractive core-shell interactions where $\boldsymbol{M}_i$ is a vector whose entries are the components of the multipoles located on that atom. The final term corresponds to the shell-shell interactions.

The core-shell interactions are damped according to

$$\boldsymbol{T}^{damp}_{ij} = \begin{bmatrix} 1 & \nabla & \nabla^2 \end{bmatrix} \cdot \left( \frac{1}{r_{ij}} f^{damp}_{ij}(r_{ij}) \right) \tag{7}$$

while the corresponding interaction tensor for shell-shell damping is written as:

$$\boldsymbol{T}_{ij}^{overlap} = \begin{bmatrix} 1 & \nabla & \nabla^2 \\ \nabla & \nabla^2 & \nabla^3 \\ \nabla^2 & \nabla^3 & \nabla^4 \end{bmatrix} \cdot \left( \frac{1}{r_{ij}} f_{ij}^{overlap}(r_{ij}) \right) \tag{8}$$

The damping functions $f_{ij}^{damp}(r_{ij})$ and $f_{ij}^{overlap}(r_{ij})$ themselves take the following forms.

$$f_{ij}^{damp}(r_{ij}) = 1 - \left( 1 + \frac{1}{2} b_j r_{ij} \right) e^{-b_j r_{ij}} \tag{9a}$$

$$f_{ij}^{overlap}(r_{ij}) = 1 - \left( 1 + \frac{11}{16} b_{ij} r_{ij} + \frac{3}{16} (b_{ij} r_{ij})^2 + \frac{1}{48} (b_{ij} r_{ij})^3 \right) e^{-b_{ij} r_{ij}} \tag{9b}$$

The damping function in Eq. 9a can be derived directly from the form of the Slater density in Eq. 2 by computing its electrostatic potential. The damping function in Eq. 9b can be derived from a symmetrized coulomb integral where each density interacts with the damped potential generated by the other density.[9] Finally, it is important to note that these damping functions are the ones which apply to charge-charge interactions and that as higher-order multipoles are considered, new damping functions are generated alongside the gradients of $1/r_{ij}$. In other words, every interaction tensor in the multipole expansion will be damped differently. We use the same set of damping functions as derived by Rackers and Ponder which are appropriate for the Slater density in Eq. 2.[9] Very similar models have been applied successfully in MB-UCB[7] based on functional forms proposed by Piquemal[23] and others.[24,25]

**Pauli Repulsion.** The original aim of the density overlap model was to model the Pauli repulsion energy formally as an exponential repulsion at short-range.[15,26,27] However, Rackers and Ponder have made a convincing argument that the appropriate functional form for Pauli repulsion is $V^{exch} \propto e^{-b_{ij} r_{ij}} / r_{ij}$.[9,28] While the exponential is the dominant contribution, the factor of $1/r_{ij}$ becomes important at short distance and allows for the Pauli repulsion energy to be expressed as a multipole expansion. Their basic idea is that the Pauli repulsion energy

between a pair of atoms is proportional to $S^2/r_{ij}$ where $S$ is the overlap between pseudo-orbitals, defined as $\sqrt{\rho}$ where $\rho$ is the density in Eq. 2.

Therefore, the Pauli repulsion energy can be written as

$$V_{Pauli} = \sum_{i<j} \frac{K_{ij}^{q} S_q^2 + K_{ij}^{\mu} S_\mu^2 + K_{ij}^{\Theta} S_\Theta^2}{r_{ij}} \tag{10}$$

where $S^2$ contains contributions from charges, dipoles, and quadrupoles as it is found that Pauli repulsion is highly anisotropic. However, producing parameters for a complete multipole expansion tends to result in overfitting when there is not a way to derive the initial multipoles from electronic structure. Therefore, the proportionality constants $K_{ij} = K_i K_j$ are fit instead.

Since $S^2$ takes the form of a damped multipole expansion,[28] these proportionality constants mean multipoles which handle repulsion are proportional to the actual electrostatic multipoles. Hence the calculation of electrostatics and multipolar Pauli repulsion differs only in the choice of damping function thereby reducing computational cost of the two EDA terms.

The expansion of Pauli repulsion in terms of multipoles has an interesting physical interpretation. Namely, as two electron densities begin to overlap, the electrons will be expelled from the internuclear region in order to keep the total system wavefunction antisymmetric. This results in a "hole" in the electron density where nuclei are exposed to one another. In a sense, then, these multipoles describe the magnitude and shape of the depletion of electron density between two atoms which are near one another.

**Dispersion.** The dispersion energy is the simplest term in the model. We use a damped polynomial interaction given by,

$$V_{disp} = \sum_{i<j} f_6^{TT}(x_{ij}) \frac{C_{6,ij}}{r_{ij}^6} \tag{11}$$

where $C_{6,ij}$ is the dispersion coefficient between atoms $i$ and $j$ which is determined as $C_{6,ij} =$

$\sqrt{C_{6,i}C_{6,j}}$, and $C_{6,i}$ is a parameter fit to the EDA dispersion energy. $f_6^{TT}(x_{ij})$ is the sixth-order Tang-Toennies damping function[29] which was originally derived to damp short-range dispersion,

$$f_n^{TT}(x_{ij}) = 1 - e^{-x_{ij}} \sum_{k=0}^{n} \frac{x_{ij}^k}{k!} \tag{12}$$

The appropriate form of $x$ for the tail of a Slater electron density has been derived before[17] and takes the form,

$$x_{ij} = b_{ij} r_{ij} - \frac{2b_{ij}^2 r_{ij}^2 + 3b_{ij} r_{ij}}{b_{ij}^2 r_{ij}^2 + 3b_{ij} r_{ij} + 3} \tag{13}$$

Note that the TT damping functions, Eq. 12, depend parametrically on the choice of integer $n$. In their original work, Tang and Toennies show that the appropriate choice of $n$ for dispersion is $n = 6$. This makes the damping function an exponential multiplied by a sixth order polynomial. This polynomial is able to control the $r^{-6}$ scaling of dispersion, while the exponential ensures no damping at long distances. As an aside, one could also use TT damping functions of different orders to control mutual polarization. We have tested this and it works just as well as the procedure we described of increasing the order of $f^{overlap}$.

## Polarization, Charge Transfer, and the One-Body Potential

This work introduces new physics for polarization, charge transfer, and the one-body potential. In particular, we introduce a combined fluctuating charge (FQ) and induced dipole model of electronic polarization that couples to our model for many-body charge transfer that we will show better reproduces all terms from EDA, and allows for explicit transfer of charge between molecules. When combined with our new model for the one-body potential, we also ensure that the force field reproduces all physically relevant monomer properties including the dipole moment, dipole derivatives, molecular polarizability, and polarizability derivatives.

**Polarization**. While atomic polarizabilities naturally contain both charge-flow and induced dipole contributions,[30] typically the charge-flow contributions are contained through localization.[31] Our approach allows for charge flow polarizabilities using a modification of the electronegativity equalization model (EEM) of polarization.[32] In EEM, the energy of a molecule is expanded to second-order as a function of charge while allowing all charges to interact

$$V(\boldsymbol{q}) = \sum_i \chi_i q_i + \frac{1}{2}\sum_i \eta_i q_i^2 + \sum_{i<j} \frac{q_i q_j}{r_{ij}} \tag{14}$$

where $\chi_i$ represents the electronegativity of atom $i$ and $\eta_i$ is the atomic hardness of atom $i$. By requiring the electronegativity of all atoms to become become equal, new atomic charges are determined by solving a system of linear equations.

There are several known shortcomings of EEM for non-reactive FFs including allowing for long-range transfer of charge between molecules, which is unphysical,[33,34] as well as a change in charge of atoms in a molecule that interferes with the definition of the permanent electrostatics. Our solution to the first problem is to only allow charge rearrangements within a molecule and not between molecules, a constraint that introduces the use of Lagrange multipliers. For the second problem, we drop the linear term in Eq. 14 and focus only on the fluctuation of charge around the reference charge used for the permanent electrostatics. Thus we are equalizing electronegativity around an "already equalized" state, and the change in electronegativity at each atom due to an environment is simply the electric potential at that atom. We can then write the FQ contribution in our model as,

$$V(\delta\boldsymbol{q}) = \frac{1}{2}\sum_i \eta_i \delta q_i^2 + \sum_i \delta q_i V_i + \sum_{i<j} \frac{\delta q_i \delta q_j}{r_{ij}} + \sum_\alpha \lambda_\alpha \sum_{i\in\alpha} \delta q_i \tag{15}$$

We also allow electric fields due to the environment to induce dipoles on all atoms as done previously for models such as AMOEBA, HIPPO, MB-UCB, MB-Pol and q-Aqua. The energy of an induced dipole $\boldsymbol{\mu}_i^{ind}$ in an electric field, $\boldsymbol{E}$, including mutual polarization is,

$$V(\boldsymbol{\mu}^{ind}) = -\frac{1}{2}\sum_i \boldsymbol{\mu}_i^{ind} \cdot \boldsymbol{E}_i^{damp} + \sum_{i<j} \boldsymbol{\mu}_i^{ind} \boldsymbol{T}_{ij}^{\mu\mu} \boldsymbol{\mu}_j^{ind} \tag{16}$$

The field $\boldsymbol{E}_i^{damp}$ is the damped electric field generated by a Slater density and $\boldsymbol{T}_{ij}^{\mu\mu}$ is the damped dipole-dipole interaction tensor which is derived from appropriate gradients of $f_{ij}^{overlap}/r_{ij}$.

The form of the $ij$ entries of the multipole interaction tensors are as follows:

$$T_{ij}^{qq} = f_1^{overlap} \frac{1}{r_{ij}} \tag{17a}$$

$$\boldsymbol{T}_{ij}^{q\mu} = f_3^{overlap} \frac{-\boldsymbol{r}_{ij}}{r_{ij}^3} \tag{17b}$$

$$\boldsymbol{T}_{ij}^{\mu\mu} = \left( f_5^{overlap} \frac{\boldsymbol{r}_{ij} \otimes \boldsymbol{r}_{ij}}{r_{ij}^5} - f_3^{overlap} \frac{\mathbf{1}}{r_{ij}^3} \right) \tag{17c}$$

The interaction tensors in Eq. 17 are the usual Cartesian multipole interaction tensors, generated by successive gradients of $1/r_{ij}$ where $r_{ij}$ is the distance between two atoms. These tensors are multiplied by the overlap damping function derived from the overlap of two slater pseudo-orbitals.[9] As an aside, the damping functions generated by the usual Thole damping procedure,[35] are exponentials multiplied by first-, second-, and third-order polynomials for charge-charge, charge-dipole, and dipole-dipole interactions, respectively.[35] These polynomials are smaller than the scaling of mutual polarization, which is roughly the square of a similar permanent multipole interaction. That is, the field due to a permanent dipole decays as $r_{ij}^{-3}$. The dipole induced by this field interacts with other induced dipoles with an energy that is proportional to the magnitude of the induced dipole and via a field that decays as $r_{ij}^{-3}$. Since the magnitude of an induced dipole is proportional to the permanent applied field, this means mutual polarization of induced dipoles decays as $\sim r_{ij}^{-6}$.

The damping functions generated from the overlap of slater orbitals are third-, fourth-,

and fifth-order polynmials multiplied by an exponential for charge-charge, charge-dipole, and dipole-dipole interactions, respectively. These damping functions are therefore better able to control the polarization energy than the Thole damping functions used in other force fields. In fact, the small orders of the polynomials in Thole damping likely explains much of the historic difficulty of controlling polarization between ions and water.[36,37]

Normally, the dipole polarizability is treated as a constant in polarizable force fields, but Chung *et al.* have pointed out that polarizabilities can be significantly diminished in the aqueous phase.[38] This effect is not exclusive to ions, but is simply more noticable for diffuse anions. We adopt a slightly simplified version of the scheme suggested by Chung *et al.* for making the polarizability dependent on the local environment. We damp the inverse polarizability, $\boldsymbol{\alpha}^{-1}$, as follows

$$\boldsymbol{\alpha}_i^{-1} = R_i \begin{pmatrix} \alpha_{xx,i}^{-1} & 0 & 0 \\ 0 & \alpha_{yy,i}^{-1} & 0 \\ 0 & 0 & \alpha_{zz,i}^{-1} \end{pmatrix} R_i^T + \mathbb{1} \sum_j k_{ij}^{damp} S_{ij}^{\rho} \tag{18}$$

The first term in Eq. 18 is a typical expression of the dipole polarizability in the local axis frame of that atom. $\alpha_{xx,i}$ is the $xx$ component of the dipole polarizability with other entries defined analogously. $R_i$ is the rotation matrix that transforms the local axis system of atom $i$ to the global axis system. The second term defines an environment-dependent isotropic damping of the polarizability. $k_{ij}^{damp}$ is a pair-specific parameter which modulates the increase of the inverse polarizability which is proportional the density overlap, $S_{ij}^{\rho}$.

Note that this modification of the polarizability describes a completely different effect from the damping of induced electrostatics. In the case of multipolar interactions, the damping arises from the fact that real charge densities have a finite width. The effect modelled in Eq. 18 is the shrinking of atoms which occurs due to antisymmetrization of the wavefunction. The effect is most important for very diffuse atoms, such as $I^-$, or for very close contacts such as the interaction of $Li^+$ with $H_2O$.

What now remains is to determine the values of $\delta\boldsymbol{q}$ and $\boldsymbol{\mu}^{ind}$ which minimize the total energy of the system. In order to do this, we take the derivative with respect to each $\delta q_i$ and each component of each $\boldsymbol{\mu}_i^{ind}$ and set them all equal to zero. This results in a system of linear equations which can be written succinctly as follows:

$$\begin{pmatrix} \boldsymbol{T}^{qq} & \mathbf{1}_\lambda & \boldsymbol{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & 0 & 0 \\ -\boldsymbol{T}^{\mu q} & 0 & \boldsymbol{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{V} \\ \boldsymbol{Q} \\ \boldsymbol{E} \end{pmatrix} \tag{19}$$

The solution vector in Eq. 19 contains the electric potential, $\boldsymbol{V}$, the total charges of each molecule, $\boldsymbol{Q}$, and the electric field on each atom $\boldsymbol{E}$. The matrix has several blocks containing the charge-charge ($\boldsymbol{T}^{qq}$), charge-dipole ($\boldsymbol{T}^{q\mu}$), dipole-charge ($\boldsymbol{T}^{\mu q}$), and dipole-dipole interaction tensors ($\boldsymbol{T}^{\mu\mu}$). Note that the diagonal elements of $\boldsymbol{T}^{qq}$ are the atomic hardness $\eta$ and the $3 \times 3$ diagonal blocks of $\boldsymbol{T}^{\mu\mu}$ are the inverse polarizability tensor $\boldsymbol{\alpha}_i^{-1}$. The block $\mathbf{1}_\lambda$ has a column for each molecule in the system. An entry in that column is 1 if the $i$th atom is in that molecule and zero otherwise. These blocks enforce the charge-conservation constraints for each molecule. Finally, $\delta\boldsymbol{q}$ contains the optimally rearranged charges, $\boldsymbol{\lambda}$ contains the Lagrange multipliers which enforce charge conservation, and $\boldsymbol{\mu}$ are the induced dipoles.

Finally, there is one more term in our polarization model which is designed to only contribute at very short range. Specifically, we introduce another term proportional to density overlap,

$$V_{pol,sr} = \sum_{i<j} a_{ij}^{sr} f_4^{TT}(x_{ij}) \left( \frac{1}{2} \frac{\bar{\alpha}_i + \bar{\alpha}_j}{r_{ij}^3} \right)^{4/3} S_{ij}^{\rho} \tag{20}$$

where $\bar{\alpha}_i$ is the mean dipole polarizability of atom $i$, $f_4^{TT}(x_{ij})$ is the fourth-order Tang-Toennies damping function, defined later in Eq. 12, and $a_{ij}^{sr} = a_i^{sr} a_j^{sr}$ is the pairwise parameter fit for this term. While Eq. 20 is more empirical than the rest of the FQCT force field, it captures the effect of quadrupole polarization, which is known to be important in water[39] (which is very computationally expensive to account for compared to the magnitude

of the effect). In fact the last term in Eq. 20 tends to have a magnitude for water that is only a bit larger than what would one expect for quadrupole polarization of around 20%. Thus even though Eq. 20 is clearly not a true model for quadrupole polarization, since it is pairwise-additive while quadrupole polarization is a many-body contributions, it works quite well. Similar terms have been suggested before for capturing the short-range polarization of ions.[40]

**Charge Transfer.** Charge transfer is the most difficult of the terms in EDA to model since there is no classical analogue to th QM charge transfer process involving electron flow. One common approach to capturing charge transfer is to use a simple exponential dependent on the distance between atoms,[9] which captures the main effect of short-range exponential stabilization due to charge delocalization. But many-body charge transfer is non-negligible and this effect will be completely missed when using just exponentials. Another idea is to treat charge transfer the same way as polarization and solve a set of induced dipole equations as was done in MB-UCB.[7,41] While it has the benefit of capturing many-body charge transfer enrgies, it does not actually allow for charge to flow between molecules and therefore misses some of the salient physics. It is also ambiguous if the induced dipoles relevant to CT should be treated as real dipoles and allowed to interact with permanent and induced multipoles. These unsatisfactory approaches are plagued by a bigger problem in that charge transfer can be an even larger contribution than polarization, especially at short range. This means the charge transfer energy would be even more susceptible to polarization catastrophes than ordinary polarization, and thus requiring new damping schemes that obviate thir inclusion at all.

For all of these reasons, we introduce a new approach to describing charge transfer which is enabled by the fact we allow for explicit charge rearrangements in our description of polarization. Our charge transfer model includes both direct and indirect energy contributions. The direct contributions allow for energetic stabilization associated with both forward and backward charge transfer

$$V_{i\to j}^{CT} = a_{i\to j}^{CT} S_{ij}^{\rho} \tag{21a}$$

$$V_{j\to i}^{CT} = a_{j\to i}^{CT} S_{ij}^{\rho} \tag{21b}$$

$$V_{direct}^{CT} = \sum_{i<j} V_{i\to j}^{CT} + V_{j\to i}^{CT} \tag{21c}$$

in which the forward and backward contributions to charge transfer are proportional to the density overlap. We take inspiration from perturbation theory which shows, approximately, that the amount of charge transferred between two molecules is proportional to the energy associated with forward and backward charge transfer.[42–44] Therefore, we define the amount of charge transferred from $i$ to $j$, $\Delta Q_{i\to j}^{CT}$, and from $j$ to $i$, $\Delta Q_{j\to i}^{CT}$, as

$$\Delta Q_{i\to j}^{CT} = \frac{V_{i\to j}^{CT}}{\epsilon_{i\to j}} \tag{22a}$$

$$\Delta Q_{j\to i}^{CT} = \frac{V_{j\to i}^{CT}}{\epsilon_{j\to i}} \tag{22b}$$

The proportionality constant between direct charge transfer energy and the amount of transferred charge is written as $\epsilon_{i\to j}$ to emphasize that this proportionality is related to the difference in energy of an occupied orbital on $i$ and an unoccupied orbital on $j$.[42] In our model, we choose this to be a pair-specific parameter since it decreases the total number of parameters and avoids having to choose an arbitrary combination rule. The reason making $\epsilon_{i\to j}$ pair-specific reduces the number of parameters is that many pairs do not exchange appreciable amounts of charge, so this energy gap is effectively infinite. For instance, in water, only the oxygen to hydrogen parameter is relevant.

This approach is novel by allowing charge to explicitly move between fragments. This is achieved by modifying the molecular charge constraints used in Eq. 19. The charge

constraint for a fragment $A$ will then take the form,

$$Q_A^{CT} = Q_A + \sum_{i \in A} \sum_{j \notin A} \Delta Q_{j \to i}^{CT} - \Delta Q_{i \to j}^{CT} \tag{23}$$

The charge constraint including charge transfer, $Q_A^{CT}$, is simply the difference in charge transferred to atom $i$ (in $A$) and charge transferred from atom $i$, summed over all atoms in molecule $A$. These charges will not be optimally distributed, so they will be allowed to relax during the polarization process. This allows us to capture the so-called "re-polarization" [42] effect in which orbitals relax after allowing for occupied-virtual mixing. For example, when charge is transferred from oxygen to hydrogen in a water dimer, the final excess charge will mostly come to rest on the oxygen in the water with net-negative charge.

Because the charge transferred between fragments is proportional to the direct CT contributions, the charge constraints depend on the distance between atoms. This means there is a gradient contribution which multiplies the lagrange multipliers with the gradient of $\Delta Q_{i \to j}^{CT}$ and $\Delta Q_{j \to i}^{CT}$. This is not difficult or expensive to evaluate, but is an unusual gradient term which must be accounted for in software implementations.

**The One-Body Potential.** The deformation energy for a single water molecule is constructed following a protocol we have recently published.[?] The one-body potential consists of a Morse potential, cosine angle potential, a bond-bond coupling term, and bond-angle coupling term.

$$V_{bond} = D_{\mathrm{OH}} \left[1 - \exp(-\alpha(R - R_e))\right]^2 \tag{24}$$

$$V_{bb} = k_{bb}(R_1 - R_e)(R_2 - R_e) \tag{25}$$

$$V_{angle} = \frac{k_a}{2}(\cos\theta - \cos\theta_e)^2 \tag{26}$$

$$V_{ba} = k_{ba}(R - R_e)(\cos\theta - \cos\theta_e) \tag{27}$$

where $D_{\mathrm{OH}}$ is the dissociation energy of the O−H bond in water, $R_e$ is the equilibrium bond

length in water, and $\alpha = \sqrt{k_e/2D}$ determines the curvature of the potential as is evident from the fact it is written in terms of the harmonic force constant, $k_e$. The two O−H stretches in water are coupled linearly in Eq. 25 via a single bond-coupling parameter, $k_{bb}$. The angle potential is harmonic in $\cos\theta$ where $\theta$ is the HOH angle and $\theta_e$ is the equilbrium angle in water, as seen in Eq. 26. Eq. 27 shows that the angle and bond potentials are linearly coupled by a single parameter, $k_{ba}$. The parameters are fit to reproduce the CCSD(T)/aug-cc-pV5Z Hessian at the corresponding equilibrium geometry. Note that this is the only term for which we do not use ωB97X-V/def2-QZVPPD as a reference, simply because CCSD(T)/aug-cc-pV5Z is a bit closer to the experimental water monomer geometry.

As will be shown, the polarizability derivatives of the water monomer are not possible to reproduce using just atomic dipole polarizabilities. Our model, however, includes fluctuating charges which improve the polarizability derivatives considerably. Even so, the agreement between polarizability derivates computed from electronic structure and with fluctuating charges leaves something to be desired. In the same way the dipole derivatives can be reproduced accurately by including charge flux in a model,[45] we have implemented geometry-dependent atomic hardness parameters, $\eta$.

$$\eta_{\mathrm{H1}} = \eta_{\mathrm{H}} \left(\frac{R_e}{R_{\mathrm{OH},1}}\right)^{k^\eta} \left(\frac{R_e}{R_{\mathrm{OH},2}}\right)^{k^\eta_{bb}} + k^\eta_a(\theta - \theta_e) \tag{28}$$

In Eq. 28, the atomic hardness of a particular hydrogen, $\eta_{\mathrm{H1}}$, is modified based on the length of both O−H bonds, $R_{\mathrm{OH},1}$ and $R_{\mathrm{OH},2}$, and the angle $\theta$. The parameters $k^\eta$, $k^\eta_{bb}$, and $k^\eta_a$ describe the magnitude of change in atomic hardness and are fit to reproduce the polarizability derivatives computed from electronic structure. This particular functional form was chosen to be well-behaved at long bond lengths. That is, when either bond is elongated, the hardness of atom H1 is decreased, hence increasing the polarizability along that bond. This is a source of so-called electrical anharmonicity and hence contributes to the large, positive second dipole derivative associated with hydrogen-bonded water molecules.[46] Note that this term adds negligible cost to the force field evaulation since we already compute the

derivatives with respect to each internal coordinate when computing the deformation energy.

Like electrostatics, polarization parameters need to be constrained to give physically meaningful parameters. Specifically, in addition to EDA energies, we include the polarizability and polarizability derivatives at the ωB97X-V/def2-QZVPPD equilibrium geometry of water in the fitting process. The loss function we minimize against is,

$$L_{pol} = \sqrt{\frac{\sum_{i=1}^{N}(E_i^{FF} - E_i^{EDA})^2}{N}} + w_1||\boldsymbol{\alpha}^{FF} - \boldsymbol{\alpha}^{EDA}|| + w_2||\frac{\partial \boldsymbol{\alpha}^{FF}}{\partial \boldsymbol{r}} - \frac{\partial \boldsymbol{\alpha}^{EDA}}{\partial \boldsymbol{r}}|| \tag{29}$$

In the above, the first term is the RMSD of the predicted energies, $E_i^{FF}$, from the EDA energies $E_i^{FF}$. The second term is the frobenius norm of the difference between the computed and predicted molecular polarizabilities, $\boldsymbol{\alpha}$. The third term is the same as the second but for the polarizability derivatives. The weights, $w_1$ and $w_2$ are set to 1.0 and 0.5 respectively. This, in essence, forces the molecular polarizability to be reproduced exactly while allowing for some error in the polarizability derivatives which are much more difficult to reproduce.

The dipole derivatives needed to evaluate the field-dependent Morse potential, Eqs. 30 and 31 are computed from electronic structure by scanning along the O−H bond length of a water monomer. As will be shown, Pauli repulsion suffers from the same systematic error in forces as the other terms. We do not know of a way to compute the dipole derivatives relevant to Pauli repulsion from electronic structure, so we treat these as fitting parameters. We have already mentioned that our Pauli repulsion model has a natural electrostatic interpretation which provides justification for having a field-dependent contribution to Pauli repulsion. Note that the field used in the case of Pauli repulsion is not the actual electric field, but the field produced by the Pauli multipoles used in evaluating Eq. 10.

**Reference Data**. Our model is parameterized using water clusters of size $(H_2O)_n$ with n=2-5. We use 2400 dimers, trimers, tetramers, and pentamers extracted from various minimized cluster geometries. We additionally generated 4800 psuedo-random water dimers

based on a Sobol sequence. We follow exactly the same procedure as described elsewhere.[47] Using the same procedure we generated 4800 ion-water dimer geometries for all ion species considered in this study, namely $F^-$, $Cl^-$, $Br^-$, $I^-$, $Li^+$, $Na^+$, $K^+$, $Rb^+$, and $Cs^+$. For all ions, we also ran a 10ps *ab initio* molecular dynamics simulation at 500K with $\omega$B97X-V/def2-TZVPPD to generate more probable ion-water configurations. We then sampled 2400 evenly spaced configurations from this trajectory to be used for parameterization.

Some larger ion-water clusters were also generated by the following procedure. We used the Crest software package[48] which uses the semi-empirical GFN2-XTB[49] method to search for global minima on a potential energy surface. We carried out the Crest global minimum search with five different seed structures generated by taking water clusters, $(H_2O)_n$, n=6-17, from a water cluster database[50] and replacing one water randomly with one of the ions mentioned. We then took the structures of up to the ten lowest energy minima which had different hydrogen-bond networks and optimized them at the $\omega$B97X-V/def2-TZVPPD level of theory. This resulted in a total of 1044 unique ion-water clusters. These full clusters are used to characterize the ion-water potentials, but we also extracted all possible dimers and trimers from these clusters to be used in fitting of the ion force field parameters.

We then generated clusters of the form $X^-Y^+(H_2O)_n$ for n=4-10 where $X^-$ is any of the halide anions and $Y^+$ is any of the alkali metals cations. We followed the same protocol as above in generating these clusters except we started from the clusters containing a cation sampled in the previous step. A large number (**insert when finalized**) of trimers and tetramers of the ion pairs, $X^-Y^+(H_2O)_{1,2}$, were then extracted from these clusters for analyis of many-body energies involving pairs of ions rather than just a single ion. All sampled clusters are available with the paper.

All energies used in fitting parameters of the force field are computed at the $\omega$B97X-V/def2-QZVPPD level of theory. In the cases where clusters are optimized at $\omega$B97X-V/def2-TZVPPD, we recompute the energies of those clusters and any derived sub-structures with $\omega$B97X-V/def2-QZVPPD.

**Parameterization strategy.** We fit each term against only the EDA contribution to that particular energy component. Optimization of parameters is done using simple gradient descent against the root mean-square deviation (RMSD) of predicted and EDA energies. For electrostatics and Pauli repulsion, we only use dimers in the fitting process since electrostatics is strictly pairwise-additive and Pauli repulsion is nearly so. For these terms 200 random water dimers from the datasets described above are used in fitting whereas for other many-body terms we use 200 random water dimers, trimers, tetramers, and pentamers from the datasets described above.

When parameterizing electrostatics, we optimize against two objectives. First, we ensure that the dipole derivatives at the equilibrium geometry of water are correct (this can be achieved nearly exactly). Second, we optimize against the distributed multipole electrostatic energy described near Eq. 5. We also include 200 random dimers of $Cl^-(H_2O)$ and $K^+(H_2O)$ when fitting against the distributed multipole electrostatics. This seems to help with optimizing to physically meaningful multipoles. We then freeze the total charges and dipoles on each atom so that the dipole derivatives will remain correct. Next, we fit the value of the core charges, $Z$ and electrostatic exponents, $b_{elec}$, on each atom with respect to the total electrostatic energy from EDA. We also allow the quadrupoles to relax against the total electrostatic energy as a form of compensation for the lack of higher-order multipoles.

The Pauli repulsion term is first fit against the RMSD of the corresponding EDA energy. The repulsion parameters are then allowed to relax against the total interaction energy and interaction forces for only dimers. Using the forces is essential to get meaningful values of the dipole derivatives used in the field-dependent Morse potential for Pauli repulsion. This procedure essentially results in improved error cancellation which we find is still necessary for a robust force field. We will have more to say about the necessity of error cancellation in force fields later. It should be noted that we only allow the Pauli repulsion to optimize against dimers so that it cannot correct errors in the many-body contributions. Furthermore, we will see that the Pauli repulsion energy still ends up providing an unbiased estimate of

the EDA Pauli repulsion energy.

The charge transfer energy and dispersion energies are simply fit against the RMSD from their EDA energies. Dispersion, on the other hand, has a large enough many-body contribution that if only dimers are used in the fitting, one will systematically over-estimate the dispersion energy since many-body dispersion is usually repulsive. There are methods for modeling many-body dispersion, but we have not included such terms in this model.[18,51]

# Results

Before describing the outcomes of applying the FQCT model to various water and ion-water cluster energies outside the training data, let us discuss two important problems which we set out to solve in the construction of this advanced force field. First is the benefit of a new one-body potential that accounts for intramolecular polarization that is more physical than models such as AMOEBA. Second, is the inclusion of non-additivite interactions beyond classical mutual polarization, namely charge transfer.[39] Note that charge transfer can be understood as energetic stabilization associated with tunneling of electrons between molecules and hence grows at least exponentially at short distances.[52] We will focus on the EDA definition of charge transfer energy in this work, though other definitions are possible based on symmetry-adapted perturbation theory.[52,53]

## Intramolecular Polarization using a One-Body Potential

Most polarizable force fields include intramolecular polarization by allowing the induced dipoles to interact regardless of whether they are in the same molecule or not. In our experience with EDA, this tends to result in worse agreement between the model and reference polarization energies. This is not surprising given that multipoles located within a couple bond lengths are too close for electrostatics to give a good representation of the relevant interactions. There is no doubt, however, that intramolecular polarization should be considered

in some way. In the FQCT model the intramolecular polarization is described by coupling the bonding potential to the environment through the electric field. Describing intramolecular polarization by coupling it to the bond potential makes the polarization energies more accurate and dramatically improves the underlying forces.

Furthermore, this term enables us to more accurately reproduce the well-known structure-frequency correspondence in water.[54] We do this by modifying Eq. 24 to be dependent on the environmental electric field. Specifically, the bond force constant, $k_e$, and equilibrium bond length, $R_e$, are coupled to the electric field projected along the bond, $\mathrm{E_{OH}}$, via the first and second dipole derivatives, $\mu^{(1)}$ and $\mu^{(2)}$. These derivatives are treated as parameters which we compute from electronic structure. The equilibrium bond length, $R_e$, becomes

$$R_e(\mathrm{E_{OH}}) = R_e^0 + \frac{\mathrm{E_{OH}}\mu^{(1)}}{k_e^0 - \mathrm{E_{OH}}\mu^{(2)}} \tag{30}$$

where $R_e^0$ is the equilibrium bond length and $k_e^0$ is the force constant under zero field. The force constant under a nonzero field, $k_e(\mathrm{E_{OH}})$, is then

$$k_e(\mathrm{E_{OH}}) = k_e^0 - 3k_e^0\alpha\left(R_e(\mathrm{E_{OH}}) - R_e^0\right) - \mathrm{E_{OH}}\mu^{(2)} \tag{31}$$

These equations can be derived by analyzing the behavior of a Morse potential in an electric field.[54] We see in Eq. 31 that the field-dependent Morse potential guarantees the structure-frequency correlation will be respected at least approximately.

Based on the electrostatic interpretation of Pauli repulsion, we also include a contribution to the Pauli energy based on the field-dependent Morse potential, Eqs. 30 and 31. In fact, we find it interesting to observe that when permanent electrostatics are used, the shifts in bond length and frequency are unrealistically large. Allowing Pauli repulsion to also contribute to the field-dependent Morse potential therefore acts to mostly cancel out the effect of permanent electrostatics in shifting the bond length and frequency. This makes physical sense if one considers how each term affects the electron density of another molecule.

For example, in the water dimer, electrostatics will tend to pull electron density into the internuclear region, hence weakening the hydrogen-bonded O−H bond. Pauli repulsion, however, arises from antisymmetrization of the total system wavefunction which will tend to expel electron density from the internuclear region and therefore shorten the hydrogen-bonded O−H bond.

Using the above equations, in our experience, will result in approximately the correct slope of the structure-frequency correlation but will tend to understimate the actual bond length and frequency shifts. The final step to reproduce the bond lengths and vibrational frequencies is to include a contribution from charge transfer. This is motivated by adiabatic EDA calculations, where the largest contribution to bond elongation and red-shifting occurs on the charge transfer surface.[55] Specifically, we allow both the bond length and force constants to be modified according to the amount of charge transferred into a hydrogen atom, as computed with Eq. 22a. This results in the final expressions used for the bond length and force constant in our environment-dependent bonding potential,

$$R_e(\mathrm{E_{OH}}) = R_e^0 + \frac{\mathrm{E_{OH}}\mu^{(1)}}{k_e^0 - \mathrm{E_{OH}}\mu^{(2)}} + k_{ct}^{(1)}\Delta Q_{CT}^{\mathrm{H}} \tag{32}$$

$$k_e(\mathrm{E_{OH}}) = k_e^0 - 3k_e^0\alpha\left(R_e(\mathrm{E_{OH}}) - R_e^0\right) - \mathrm{E_{OH}}\mu^{(2)} + k_{ct}^{(2)}\Delta Q_{CT}^{\mathrm{H}} \tag{33}$$

The rationale for using the amount of charge accepted by a hydrogen atom, $\Delta Q_{CT}^{\mathrm{H}}$, in Eqs. 32 and 33 is that this charge is transferred into an anti-bonding orbital and hence should have a large effect on the O−H bond in question. Similarly, we ignore the charge transferred out of the oxygen atom since these electrons come from non-bonding orbitals and should therefore minimally affect the O−H bonds in that water. This introduces two additional parameters, $k_{ct}^{(1)}$ and $k_{ct}^{(2)}$, which determine the sensitivity of the bond-length and force constant to charge transfer. It is very important to note that these parameters have a very small effect on the charge transfer energy, but a sizable effect on the charge transfer

forces.

To summarize the large effect this environment-dependent bonding potential has on the forces within FQCT, let us consider a water dimer where each monomer is fixed at the $\omega$B97X-V/def2-QZVPPD geometry. This simply eliminates the geometric distortion forces which are not relevant here. We now compute the forces due to each term in EDA, a so-called force decomposition analysis,[56] and compare them against the same forces predicted by FQCT with and without Eqs. 32 and 33.

Table 1: Comparison of the forces projected along the O−H bond of a water dimer due to each component in the EDA. Forces predicted by FQCT are computed with and without the field-dependent morse potential described in Eqs. 32 and 33. The first block of entries is FQCT with no field-dependent Morse potential, the second block is FQCT, and the third is $\omega$B97X-V/def2-QZVPPD. All forces in KJ/mol/Å. $O_{don.}$ and $H_{don.}$ are the oxygen and hydrogen atoms donating the hydrogen bond and $O_{acc.}$ is the oxygen accepting the hydrogen bond. All forces are projected along the O−H bond vector. The numbers in parentheses are what we forces before error fitting with Pauli repulsion.

Water Dimer Force Decomposition Analysis Along O−H Bond

| Method | Atom | Mod. Pauli | Cls. Elec. | Disp. | Pol. | CT | Total |
|---|---|---|---|---|---|---|---|
| FQCT | $O_{don.}$ | -30.1 | 1.4 | 6.0 | -3.4 | -1.5 | -27.5 |
| No F.D. Morse | $H_{don.}$ | -73.7 | 54.5 | 5.3 | 15.4 | 24.9 | 26.4 |
| | $O_{acc.}$ | 96.4 | -48.9 | -10.1 | -12.7 | -22.3 | 2.4 |
| FQCT | $O_{don.}$ | -15.5 (7.50) | -13.7 | 6.0 | -5.6 | -8.7 | -37.5 |
| | $H_{don.}$ | -88.5 (-111.9) | 70.0 | 5.3 | 17.7 | 31.9 | 36.4 |
| | $O_{acc.}$ | 96.4 (95.5) | -51.7 | -10.1 | -13.0 | -22.0 | -0.4 |
| $\omega$B97X-V | $O_{don.}$ | 3.92 | -15.0 | 0.75 | -8.4 | -15.6 | -34.3 |
| | $H_{don.}$ | -105.7 | 72.4 | 10.8 | 18.9 | 39.3 | 35.8 |
| | $O_{acc.}$ | 92.3 | -52.3 | -10.3 | -10.4 | -19.6 | -0.50 |

The net effect of the field-dependent Morse potential and charge transfer correction is to shift force from the oxygen atom to the hydrogen atom participating in a hydrogen bond. The improvement in accuracy in electrostatics and polarization forces are particularly notable since this correction involves no free parameters. The improvement in Pauli repulsion is also quite large correcting about half the force error at negligible computational cost. The charge transfer forces also remove about half the force error along the O−H bond. Note that the Pauli repulsion forces can be quite a bit more accurate with the present model, as shown in

the parentheses in Table 1. Since we use the Pauli repulsion to correlate errors in the model, as shown in Figure 5, the Pauli forces compensate for remaining errors in the dispersion and charge transfer forces.

In order for a force field to be useful theoretical spectroscopy, it must respect the relationships between structure and vibrational frequencies. For water this manifests as a linear relationship between the change in bond length and change in O−H stretching frequency,[54] colloquially referred to as Badger's rule.[57] It is expected that an accurate FF should satisfy the expected slope $\Delta\omega$ vs $\Delta R_e$ of $\approx -19$ cm$^{-1}$/.001Å.[54] Without a field-dependent bond coupling, we find that we obtain a slope of $\approx -10$ cm$^{-1}$/.001Å (Fig. 7), consistent with a field-independent Morse oscillator analyzed by Boyer *et al.* with parameters appropriate to water.[54] This observation is what motivated us to couple our bonding potential parameters to the field along that bond as shown in Eqs. 32 and 33. The result of doing this is shown in Figure 7, accurately reproducing the structure-frequency correlation in water. When we compute the dipole derivatives from a simple O−H scan we find the QM parameters are $\mu^{(1)} = 0.1654$ and $\mu^{(2)} = -0.01246$, while the same calculation with FQCT yields $\mu^{(1)} = 0.1658$ and $\mu^{(2)} = -0.0104$. This indicates that as long as a force field has an accurate dipole surface, the dipole derivatives needed to compute the field-dependence of a Morse potential can be extracted from the force field itself.

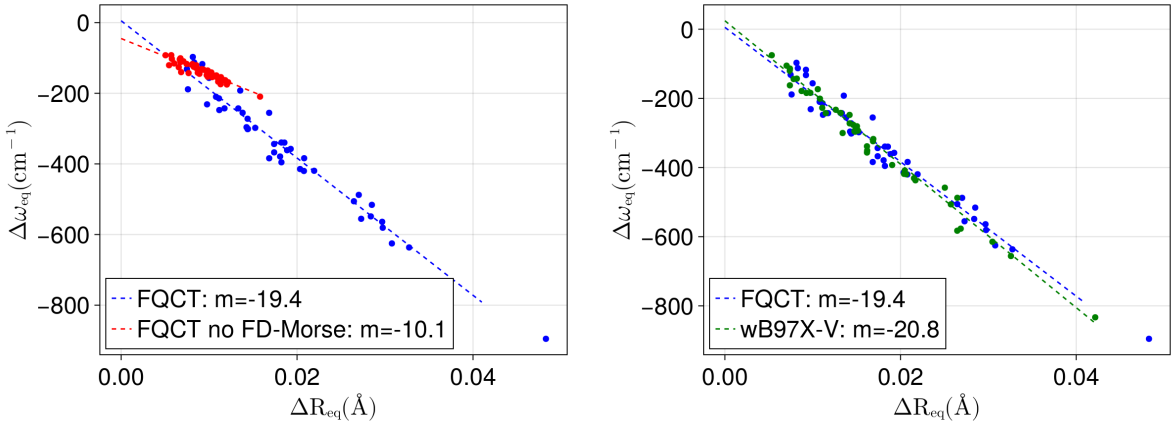
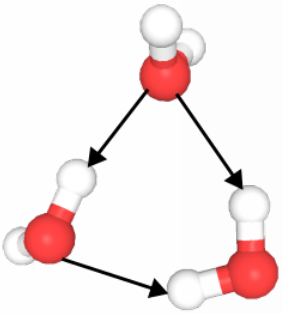


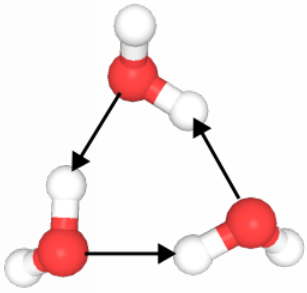

Figure 1: TBA See text for discussion.

## Motivation for Modelling Many-Body Charge Transfer

If we consider that charge transfer involves the transfer of actual charge between molecules, which is certainly the case in EDA,[58] then it should be clear that there are both attractive and repulsive contributions to charge transfer. The attractive part is the energy lowering associated with delocalizing the electron density. The repulsive part arises from any molecule having a non-integer total charge.



| | 2B | 3B | Total |
|---|---|---|---|
| FRZ | -5.66 | 0.0 | -5.66 |
| POL | -2.01 | 0.14 | -1.87 |
| CT | -2.77 | 0.10 | -2.67 |
| Total | -10.44 | 0.24 | -10.20 |



| | 2B | 3B | Total |
|---|---|---|---|
| FRZ | -3.92 | 0.0 | -3.92 |
| POL | -3.57 | -1.66 | -5.23 |
| CT | -6.26 | -0.76 | -7.02 |
| Total | -13.75 | -2.42 | -16.17 |

Figure 2: Two water trimers illustrating the importance of both polarization and charge transfer for the stability of water molecules. FRZ corresponds to the frozen contribution: the sum of Pauli, electrostatic, and dispersion enegies. POL is the polarization energy and CT is the charge transfer energy. Energies are computed at the $\omega$B97X-V/def2-QZVPPD level of theory. Total energies are interaction energies and not binding energies. See text for discussion.

The existence of a penalty for moving charge between molecules makes it clear why many-body charge transfer is non-negligible. Figure 2 illustrates how hydrogen bond networks benefit from many-body charge transfer. Let us imagine that an equal amount of charge

is transferred along each hydrogen bond, represented by arrows. In the trimer on the left, the hydrogen bonds are organized in such a way that some molecules donate more hydrogen bonds than they receive and vice versa. This means that some molecules in the trimer on the left will end up with nonzero total charges. This is consistent with the known instability of double acceptor water molecules which do not donate any hydrogen bonds.[59]

In the trimer on the right, each molecule donates and receives the same number of hydrogen bonds. This means each molecule will have a nearly net-zero charge while still benefitting from charge delocalization along each hydrogen bond. Now, if we imagine removing any water molecule from the trimer on the right, the two remaining molecules would have nonzero total charges and hence incur a penalty. The elimination of this penalty when a third molecule is added to the network is exactly what gives rise to a many-body charge transfer stabilization.

The ability of hydrogen bond networks to delocalize charge while keeping each fragment very nearly neutral seems to us an under-appreciated aspect of hydrogen bond cooperativity. Indeed, the trimer on the right is a so-called homodromic ring, which results in enhanced cooperativity of the induced dipoles on each molecule.[1] This can be seen by the large 3-body polarization contribution of $-1.66$ kcal/mol. We highlight, here, that there is an additional non-negligible 3-body contribution due to charge transfer of $-0.76$ kcal/mol. The fact that water can receive and donate two hydrogen bonds simultaneously makes it uniquely capable of passing charge between molecules while keeping the total charge of each molecule neutral.

It is interesting to note that the frozen term, i.e. the sum of electrostatics, dispersion, and Pauli repulsion can end up being smaller than either the polarization or charge transfer contributions. Additionally, although not shown explicitly here, the charge penetration contribution, Eq. 5, to the electrostatic energy of a hydrogen bond is typically about equal to the point-electrostatic contribution. This begs the question of how so many force fields have been even qualitatively successful while neglecting such large components of the intermolecular energy?

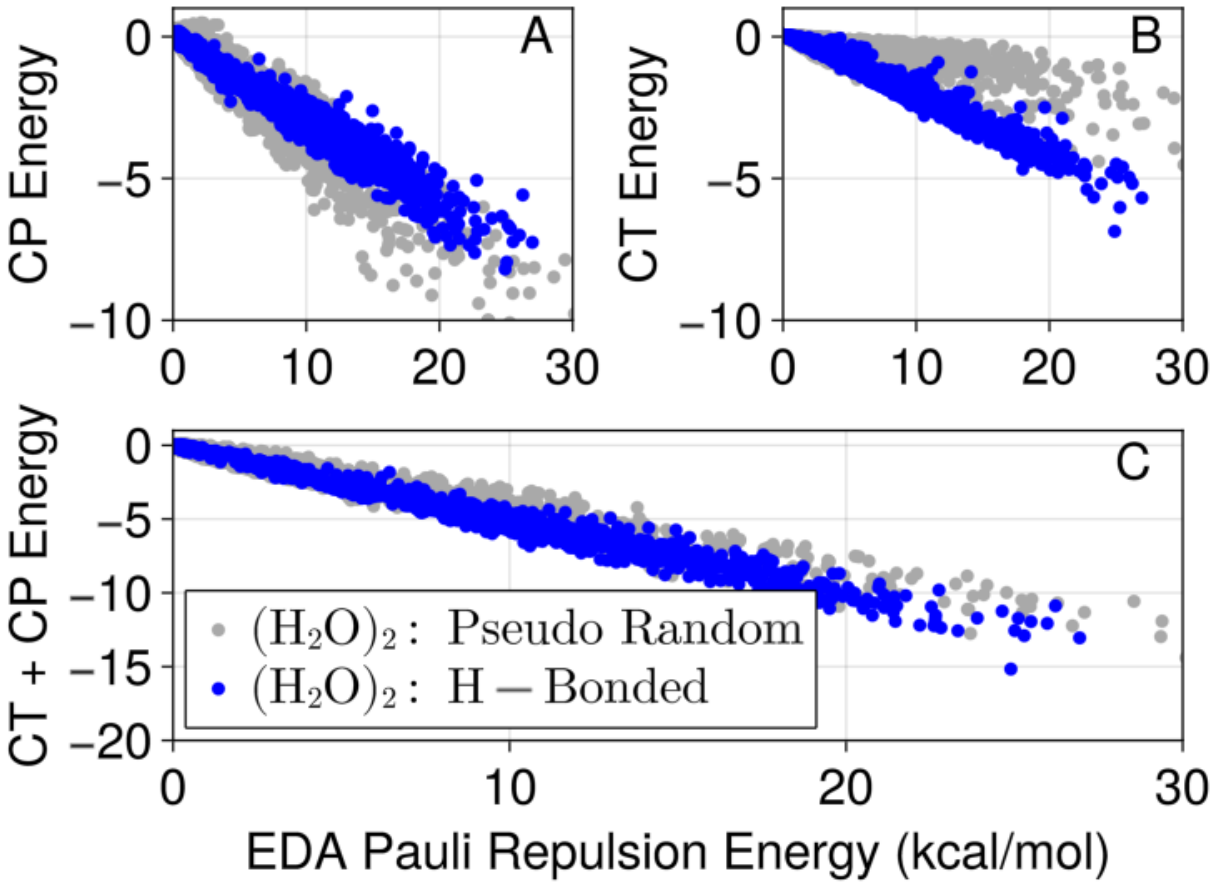



Figure 3: Correlation between (A) charge penetration energy (CP), (B) charge transfer energy (CT), and (C) their sum against Pauli repulsion energy. Energies are computed at the $\omega$B97X-V/def2-QZVPPD level of theory. See text for discussion.

Figure 3 demonstrates that force fields have succeeded despite neglecting charge transfer (CT) and charge penetration (CP) simply because CT and CP are exceptionally strongly correlated to Pauli repulsion. This observation is extremely robust for water. Clearly, most force fields are implicitly describing CT and CP by having too soft of a repulsive wall when compared against an *ab initio* calculation of Pauli repulsion such as that from EDA. It is rather easy to see from Figure 3 that the sum of CT and CP is generally better correlated to Pauli repulsion than either term individually. Whether or not this correlation holds as robustly for other systems is beyond the scope of this work. Figure 3 does raise the alarming possibility that adding charge penetration to a force field without adding charge transfer can

make the force field worse by destroying a fortuitous cancellation of errors. We consider that Figures 2 and 3 provide ample motivation for the development of a better model for charge transfer, which we present here.

## Water Monomer Properties

In the construction of this model, we have gone to great lengths to ensure that the model reproduces as many properties of the water monomer as possible. For instance, in Figure 4 we make a comparison between the dipole surface of this model and the reference surface computed with $\omega$B97X-V/def2-QZVPPD. We also show two references of possible interest. The blue points correspond to the dipole surface of a force field with fixed charges and dipoles that optimally reproduce the electrostatic dimer EDA data described earlier. Clearly, fixed charge force fields completely fail to reproduce the dipole surface of water.

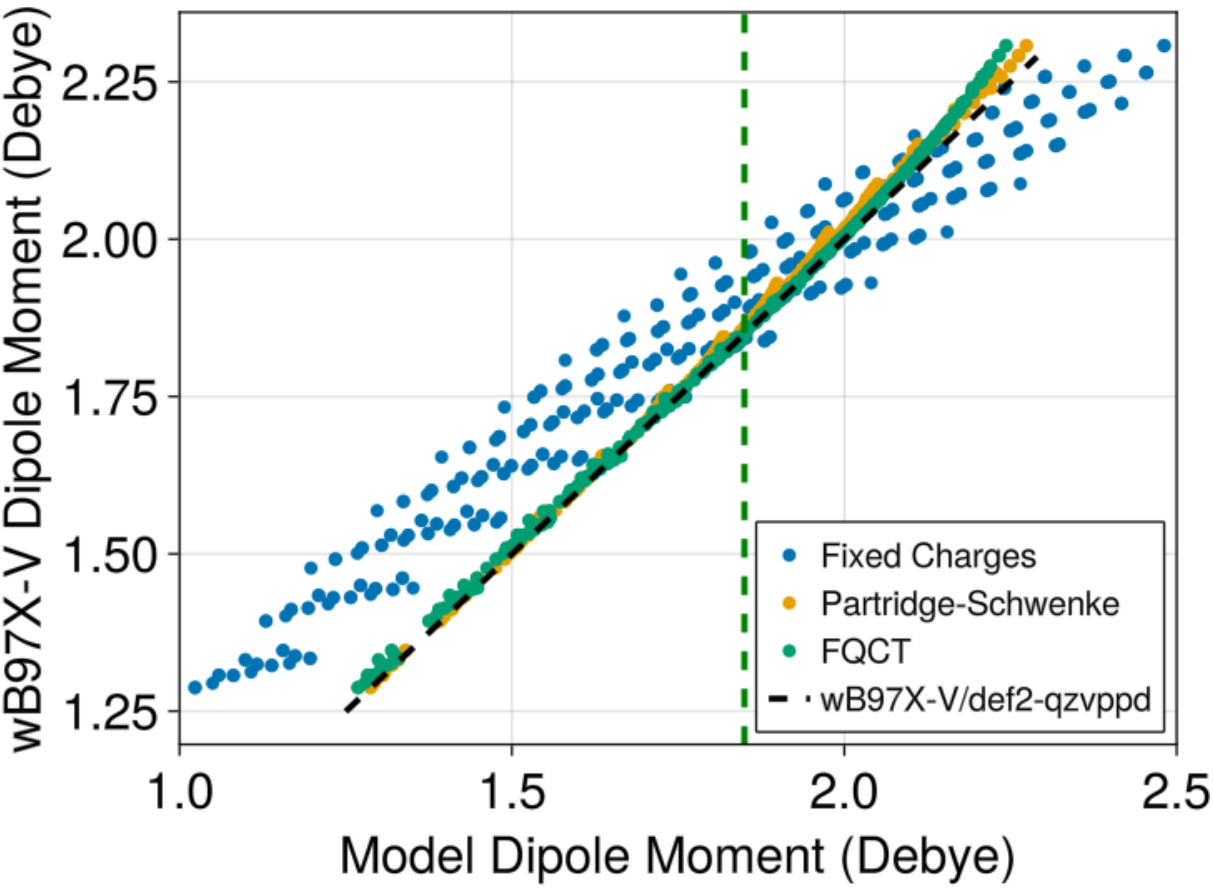



Figure 4: The dipole surface of water for all structures taken from a 3-D scan of water internal coordinates with a deformation energy less than 20 kcal/mol. The black dashed line shows the values computed with $\omega$B97X-V/def2-QZVPPD. The green dashed line corresponds to the experimental gas-phase dipole moment of water of 1.85 Debye.

The orange points in Figure 4 are the dipole surface associated with the Partridge-Schwenke (PS) water monomer surface.[60] The PS potential and dipole surface are used by many water models such as the TTM models,[61,62] MB-Pol,[63,64] and q-aqua.[65] Clearly, the PS dipole surface is very accurate over a wide range of energies. We chose not to use the PS dipole surface since we intend to construct force fields for molecules other than water. Clearly, we have not sacrificed much since the FQCT dipole surface essentially parallels the PS dipole surface.

It was first pointed out by Fanourgakis and Xantheas that reproducing the dipole surface of water is essential for capturing the opening of the bend angle of water in the condensed

phase.[62] Essentially, a model needs to reproduce the dipole derivatives of water to correctly predict the opening of the bend angle as water clusters become larger. Our model is able to reproduce the dipole derivatives of water at its equilibrium geometry to five decimal places in atomic units.

Another important property for accurate energetics and transferability, especially for interactions with ions, is the molecular polarizability. Because our model contains anisotropic dipole polarizabilities and fluctuating charges, there is more than enough flexibility in parameter space to exactly reproduce the molecular polarizability. Our model aims to reproduce the $\omega$B97X-V/def2-QZVPPD molecular polarizability with $\alpha_{xx} = 10.0321$, $\alpha_{yy} = 9.65958$, $\alpha_{zz} = 9.40921$, which it manages to do up to three decimal places in bohr$^3$. The quoted polarizability values correspond to a water molecule with the y-axis as the bisector of the HOH angle and the z-axis normal to the plane of the water molecule.

The last monomer property we explicitly aimed to reproduce are the polarizability derivatives of gas-phase water at its equilibrium geometry. The polarizability derivatives control the intensity of peaks measured with raman spectroscopy but have rarely been discussed in the construction of water models to our knowledge. Additionally, since water distorts when interacting with other molecules, reproducing the polarizability derivatives indicates how well the molecular polarizability at distorted geometries will be reproduced. In Table 2, we report the polarizability derivatives of our final model as well as the polarizability derivatives for an identically parameterized model which does not include fluctuating charges. It is important to note that one can achieve better agreement with reference polarizability derivatives if that is the only quantity one aims to reproduce. These models, however, tend to result in inadequate polarization energies.

Certain polarizability derivatives in Table 2 are reproduced very accurately, while others are not at all. Most notably, our model does not have any $zz$ polarizability derivatives while these should be rather large. The reason our model has no $zz$ polarizability derivatives is that we do not include intramolecular polarization. We tested what happens when intramolecular

Table 2: Polarizability derivatives of water computed at the ωB97X-V/def2-QZVPPD equilibrium geometry in three different ways. The first entry is computed with ωB97X-V/def2-QZVPPD, the second with FQCT, and the third with the same model but using parameters optimized without fluctuating charges. The Atom column tells both the atom and component of its position we take the derivative with respect to. The water monomer has its bisector aligned with the $y$-axis and the $z$-axis is normal to the water molecule plane. Note that the derivatives for the second hydrogen are identical to the first but with opposite sign. The $xz$ and $yz$ entries are omitted since they are small and reproduced to three decimal places by both models.

Polarizability Derivatives of Water (bohr$^2$)

| Atom | $xx$ | $xy$ | $yy$ | $zz$ |
|---|---|---|---|---|
| O$x$ | - | 4.04/3.99/-0.13 | - | - |
| O$y$ | 5.15/6.07/2.04 | - | 4.45/4.85/-2.03 | 1.50/0.0/0.0 |
| O$z$ | - | - | - | - |
| H$x$ | -4.61/-4.80/0.78 | -2.02/-1.99/0.06 | -2.53/-1.68/-0.78 | -1.39/0.0/0.0 |
| H$y$ | -2.57/-3.04/-1.02 | -1.68/-1.66/-0.08 | -2.22/-2.43/1.02 | -0.75/0.0/0.0 |
| H$z$ | - | - | - | - |

polarization is turned on and found that the $zz$ polarizability derivatives are still reproduced very poorly and the polarization energies generally became less accurate. It is of some interest that when intramolecular polarization is turned on, the polarizability derivatives are quite sensitive to the choice and strength of damping function. This may enable the tuning of damping functions based on monomers alone.

We find it very interesting that the model which includes fluctuating charges in the polarization calculation gives much better polarizability derivatives than one which just uses anisotropic dipole polarizabilities. Note that the polarizability derivatives in Table 2 are improved considerably by the geometry-dependent atomic hardness described in Eq. 28. Even without variable atomic hardness, including fluctuating charges dramtically improves polarizability derivatives compared to models with only dipole polarizablities. This indicates that one of the main reasons water models have historically predicted Raman intensities very poorly[66] is the lack of fluctuating charges in the polarization process. Furthermore, we are able to achieve higher accuracy in the polarization energies using fluctuating charges than without.

The $zz$ polarizability derivatives of water are an interesting case since they can be reproduced accurately when using intramolecular polarization, but in our experience, the strength of damping to enable this is too weak to have a stable water model. We already include variable polarizability in this model via Eq. 28. Charge fluctuations only contribute to in-plane polarization. We could, undoubdetdly, reproduce the $zz$ polarizability derivatives by allowing the z-component of atomic polarizabilities to be geometry-dependent. We decided not to do this since it adds additonal complexity that with a fairly small benefit. One can imagine, however, for a system like benzene which predominantly interacts via $\pi$ interactions that accurate out-of-plane polarizability derivatives could be quite important.

It is difficult to compare the accuracy of the polarizability derivatives of this model to others since very few models report this quantity. At least one water model has been constructed for the express purpose of reproducing Raman spectra.[67] The polarizability derivatives were not reported. The polarization is described using anisotropic polarizabilities but no fluctuating charges which, based on the contents of Table 2, lead us to believe the present model likely has more accurate polarizability derivatives.

## Water Model

Having characterized the properties of the monomer surface used in this model, let us turn our attention to its performance on intermolecular properties. To that end, Table 3 presents the mean absolute errors (MAEs) over each cluster size for each energy category in the EDA. Focusing on the dimers first, it is clear that the term in error by the most is Pauli repulsion, but even this is in error by less than 0.25 kcal/mol on average. Note that this MAE includes the re-parameterization to maximize error cancellation. Without including error cancellation, the Pauli MAE for dimers is 0.20 kcal/mol and all other terms, besides the interaction energy, remain the same. In general, the MAEs obtained by this model are excellent.

In addition to the MAEs for each category of the EDA, we also report the skewness of

the error distribution. The skewness is the standardized third moment of a distribution and is a simple measure of asymmetry. Ideally, this number will be zero indicating that the errors are symmetrically distributed. For reference, a skewness between -0.5 and 0.5 is generally interpreted as indicating an approximately symmetric distribution. Therefore, every individual term in our model is very nearly unbiased in the sense of having approximately symmetric errors. The only term of some concern is the interaction energy of dimers which has a skewness of 1.09. Note, however, that the MAE is 0.093 kcal/mol and this asymmtry in the error shrinks as larger clusters are considered. One contributor to the asymmetry in the dimer interaction energies is the lack of many-body dispersion. That is, we have not included many-body dispersion in this model and, as a result, we have to fit the dispersion to implicitly account for this error. This is why the distribution of dispersion errors starts out positively skewed (i.e. too weakly attractive) and ends up negatively skewed (i.e. too attractive). This follows from the fact that many-body dispersion is repulsive on average. It is quite interesting to note that the dimer interaction energy is in error by less than every category besides polarization and dispersion. This immediately indicates that there must be rather significant error cancellation.

Table 3: Comparison of the mean absolute error (MAE) in kcal/mol of all terms in the EDA against predictions of our water model for hydrogen-bonded water dimers, trimers, tetramers, and pentamers. The numbers in parentheses are the skewness of the error distribution. In total, there are 2400 each of dimers, trimers, tetramers, and pentamers. The first row shows the Pauli repulsion energy without inclusion of error fitting while the second row is the Pauli repulsion used in the final model which is calibrated to maximize error cancellation.

MAE of Force Field EDA Terms (kcal/mol)

| | $(H_2O)_2$ | $(H_2O)_3$ | $(H_2O)_4$ | $(H_2O)_5$ |
|---|---|---|---|---|
| Pauli (no error fit) | 0.134 (1.513) | 0.209 (0.538) | 0.277 (0.491) | 0.329 (0.400) |
| Pauli | 0.184 (0.335) | 0.297 (0.090) | 0.387 (0.193) | 0.506 (-0.059) |
| Electrostatics | 0.125 (-0.209) | 0.205 (0.234) | 0.276 (-0.104) | 0.342 (-0.062) |
| Dispersion | 0.069 (0.090) | 0.092 (-0.124) | 0.109 (-0.407) | 0.149 (-0.420) |
| Polarization | 0.046 (-0.219) | 0.087 (-0.037) | 0.121 (0.437) | 0.154 (0.496) |
| Charge Transfer | 0.112 (-0.362) | 0.168 (-0.342) | 0.229 (-0.116) | 0.269 (-0.118) |
| Interaction | 0.093 (1.09) | 0.172 (0.940) | 0.237 (0.330) | 0.298 (0.399) |

Figure 5 shows the correlation of errors in all attractive terms (i.e. not Pauli) and

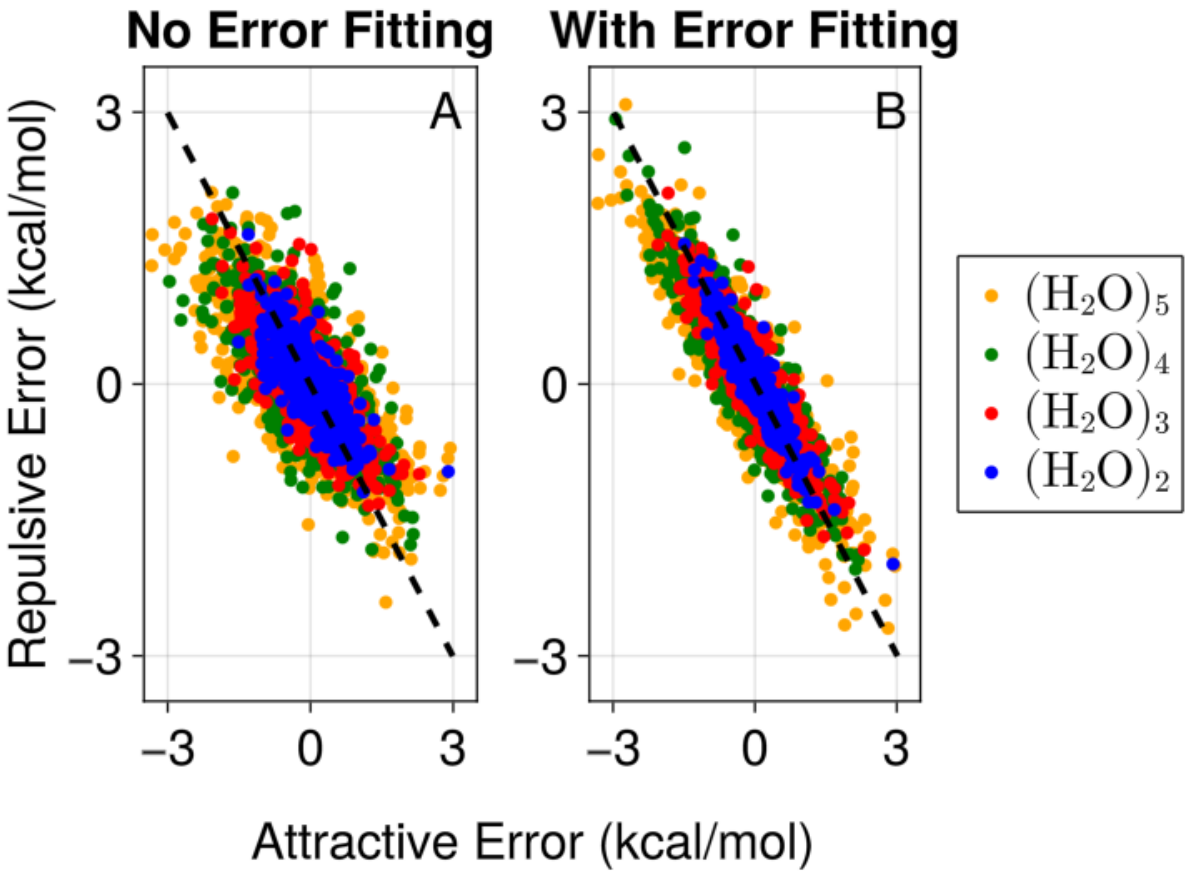



Figure 5: Correlation of the attractive errors (i.e. all terms besides Pauli) and error in Pauli repulsion. Figure A shows the error correlation without any error fitting. That is, each EDA term is fit completely independently. Figure B shows the same plot after allowing the Pauli repulsion to relax against the interaction energy for dimers only. See section on parameterization for more details.

Pauli repulsion both with and without error fitting. As a reminder, we only allow the Pauli repulsion parameters to adjust against the interaction energies and forces and we only use water dimers when fitting for error cancellation. This makes it clear that even for larger clusters, the predominant source of error is in the dimers. Virtually all force fields take advantage of some kind of error cancellation during parameterization. While it would be preferable if no error cancellation was required to construct quantitative force fields, we find it highly unlikely that this will ever be the case. This is because making quantitative force fields requires an accurate description of short-range interactions which

requires modelling quantum mechanics based on simple functional forms. As will be seen, we think this force field is a major step forward in modelling charge transfer, which is a purely quantum mechanical phenomenon, but even so, tuning the error cancellation was necessary to quantitatively reproduce cluster energetics and structures.

One might wonder why this model is so effective at taking advantage of error cancellation. We believe this stems from two facts. First, as discussed, all of the terms in this force field are very nearly unbiased. That is, each term is nearly equally likely to be in error by a positive or negative energy. This means the total energy is also unbiased and small tweaks to the Pauli parameters will tend to rapidly correlate the errors in each term. Second, error cancellation is to be expected from the simple fact that Pauli repulsion and all other terms are naturally correlated. This is easily seen in Figure 3 where the sum of charge penetration and charge transfer are linearly correlated to Pauli repulsion over a very wide energy range. Also, Figure 5A shows that just by fitting each EDA term, our model naturally correlates attractive and repulsive errors. To put it simply, when repulsion gets stronger, all of the attractive terms also get stronger, so some amount of error cancellation is basically guaranteed.

One of the major goals of this model is to quantitiatively reproduce the many-body contributions to both polarization and charge transfer. To assess how well we have achieved this, we computed the three-body contribution to both polarization and charge transfer for many trimers extracted from the ion-water clusters used in parameterizing the ions. That is, these water trimers were not used in parameterizing the water model. The three-body polarization and charge transfer energy are defined as, $E_{3B}(123) = E(123) - E(12) - E(13) - E(23)$. That is, we take any EDA energy component for a trimer, $E_{3B}(123)$, and subtract off the same EDA energy component for the three dimers forming that trimer. There is no one-body contribution to this energy since EDA only deals with intermolecular forces.

Figure 6 shows what we consider to be excellent agreement with electronic structure for the major three-body contributions to the energy. Indeed, the model manages to capture both repulsive and attractive three-body contributions to both polarization and charge

Table 4: Comparison of various advanced force fields and ωB97X-V/def2-QZVPPD against benchmark cluster energies.[68] The reference energies are mostly CCSD(T)/CBS values but some are MP2/CBS values. See original paper for further details on structures.[68] Any energies which could not be found in the literature are left blank. The bottom row shows the mean absolute error per molecule for all available energies.

Comparison of Methods on Benchmark Water Cluster Binding Energies

| $(H_2O)_n$ | Isomer | q-AQUA | MB-Pol | FQCT | ωB97X-V | HIPPO | TTM2.1-F | Ref. |
|---|---|---|---|---|---|---|---|---|
| $(H_2O)_2$ | | -4.97 | -4.96 | -4.93 | -5.00 | -4.96 | -5.03 | -4.99 |
| $(H_2O)_3$ | | -15.73 | 15.69 | -15.38 | -15.77 | -15.77 | -15.94 | -15.77 |
| $(H_2O)_4$ | | -27.35 | -27.18 | -27.30 | -27.75 | -26.69 | -27.62 | -27.39 |
| $(H_2O)_5$ | | -35.71 | -35.55 | -36.19 | -36.51 | -34.58 | -36.81 | -35.9 |
| $(H_2O)_6$ | Prism | -46.21 | -45.94 | -46.10 | -46.53 | -46.15 | -45.91 | -46.2 |
| $(H_2O)_6$ | Cage | -45.94 | -45.66 | -45.76 | -46.30 | -45.39 | -46.51 | -45.9 |
| $(H_2O)_6$ | Book | -45.21 | -44.89 | -45.47 | -45.95 | -44.25 | -46.09 | -45.4 |
| $(H_2O)_6$ | Ring | -43.71 | -43.66 | -45.01 | -45.07 | -42.54 | -45.17 | -44.3 |
| $(H_2O)_7$ | | -57.71 | -57.11 | -57.68 | -58.08 | - | -57.83 | -57.4 |
| $(H_2O)_8$ | $D_2d$ | -73.32 | -72.38 | -72.66 | -73.58 | -71.55 | -73.3 | -73.0 |
| $(H_2O)_8$ | $S_4$ | -72.93 | -72.06 | -72.74 | -73.55 | -71.56 | -73.33 | -72.9 |
| $(H_2O)_9$ | $D_2dDD$ | -82.87 | -81.40 | -82.13 | -83.00 | - | -83.4 | -83.0 |
| $(H_2O)_{10}$ | | -94.72 | -92.53 | -93.59 | -94.50 | - | -94.66 | -94.6 |
| $(H_2O)_{11}$ | 43'4 | -104.23 | -103.93 | -102.76 | -103.77 | -100.23 | -104.14 | -104.6 |
| $(H_2O)_{16}$ | Antiboat | -164.87 | -160.89 | -162.95 | -164.20 | -159.63 | -165.99 | -164.6 |
| $(H_2O)_{16}$ | 4444-a | -163.10 | -162.28 | -163.67 | -164.28 | -161.84 | -167.25 | -164.2 |
| $(H_2O)_{16}$ | 4444-b | -162.54 | -161.08 | -163.56 | -163.84 | -161.56 | -167.11 | -164.1 |
| $(H_2O)_{16}$ | Boat a | -164.53 | -161.43 | -162.71 | -164.51 | -159.36 | -165.71 | -164.4 |
| $(H_2O)_{16}$ | Boat b | -164.31 | -160.86 | -162.82 | -164.35 | -159.43 | -165.82 | -164.2 |
| $(H_2O)_{17}$ | Sphere | -177.56 | -171.75 | -175.01 | -175.78 | -170.68 | -178.60 | -175.7 |
| $(H_2O)_{20}$ | ES Prism | -212.49 | -207.84 | -211.12 | -211.98 | - | -216.46 | -214.2 |
| $(H_2O)_{20}$ | FS Prism | -210.63 | -206.96 | -209.32 | -210.12 | - | -214.07 | -211.9 |
| $(H_2O)_{20}$ | Fused Cubes | -208.07 | -207.45 | -209.39 | -209.90 | - | -214.34 | -210.6 |
| $(H_2O)_{20}$ | Pentag. Dodec. | -199.79 | -195.22 | -199.29 | -201.22 | - | -202.22 | -200.8 |
| $(H_2O)_{25}$ | Isomer 2 | -276.50 | -266.04 | -269.60 | -272.02 | - | -277.06 | -276.3 |
| MAE/n | | **0.040** | **0.156** | **0.079** | **0.051** | **0.194** | **0.092** | - |

transfer despite repulsive three-body contributions to either of those quantities being absent in ordinary water clusters.[6] In fact, ion-water clusters have been shown to have repulsive three-body contributions in many cases.[2,3] We consider Figure 6 a strong validation that our models of polarization and charge transfer capture the salient physics that generates these many-body energies.

We have also computed the energy at the force-field optimized geometries of several

Table 5: Comparison of various advanced force fields and ωB97X-V/def2-QZVPPD against benchmark cluster structures.[68] The reference structures are optimized at either CCSD(T)/aug-cc-pVDZ or MP2/aug-cc-pVTZ. See original paper for further details on structures.[68] The bottom row shows the root mean-squared deviation (RMSD) in angstrom for all available structures.

Comparison of Methods on Benchmark Water Cluster Structures

| $(H_2O)_n$ | Isomer | q-AQUA | MB-Pol | FQCT | ωB97X-V | TTM2.1-F |
|---|---|---|---|---|---|---|
| $(H_2O)_2$ | | 0.005 | 0.008 | 0.017 | 0.005 | 0.059 |
| $(H_2O)_3$ | | 0.01 | 0.014 | 0.020 | 0.008 | 0.077 |
| $(H_2O)_4$ | | 0.008 | 0.024 | 0.015 | 0.006 | 0.102 |
| $(H_2O)_5$ | | 0.013 | 0.059 | 0.046 | 0.008 | 0.156 |
| $(H_2O)_6$ | Prism | 0.01 | 0.035 | 0.019 | 0.009 | 0.094 |
| $(H_2O)_6$ | Cage | 0.013 | 0.027 | 0.023 | 0.018 | 0.134 |
| $(H_2O)_6$ | Book | 0.01 | 0.029 | 0.040 | 0.009 | 0.142 |
| $(H_2O)_6$ | Ring | 0.013 | 0.043 | 0.025 | 0.010 | 0.158 |
| $(H_2O)_7$ | | 0.016 | 0.041 | 0.027 | 0.025 | 0.122 |
| $(H_2O)_8$ | $D_2d$ | 0.006 | 0.041 | 0.017 | 0.004 | 0.068 |
| $(H_2O)_8$ | $S_4$ | 0.007 | 0.019 | 0.016 | 0.005 | 0.064 |
| $(H_2O)_9$ | $D_2dDD$ | 0.089 | 0.116 | 0.035 | 0.052 | 0.207 |
| $(H_2O)_{10}$ | | 0.012 | 0.049 | 0.022 | 0.010 | 0.085 |
| $(H_2O)_{11}$ | 43'4 | 0.034 | 0.065 | 0.024 | 0.017 | 0.102 |
| $(H_2O)_{16}$ | Antiboat | 0.023 | 0.064 | 0.035 | 0.017 | 0.094 |
| $(H_2O)_{16}$ | 4444-a | 0.039 | 0.038 | 0.028 | 0.015 | 0.074 |
| $(H_2O)_{16}$ | 4444-b | 0.04 | 0.049 | 0.029 | 0.029 | 0.071 |
| $(H_2O)_{16}$ | Boat a | 0.023 | 0.038 | 0.027 | 0.016 | 0.076 |
| $(H_2O)_{16}$ | Boat b | 0.028 | 0.057 | 0.048 | 0.016 | 0.102 |
| $(H_2O)_{17}$ | Sphere | 0.039 | 0.063 | 0.035 | 0.022 | 0.086 |
| $(H_2O)_{20}$ | ES Prism | 0.042 | 0.056 | 0.041 | 0.024 | 0.076 |
| $(H_2O)_{20}$ | FS Prism | 0.047 | 0.05 | 0.030 | 0.023 | 0.090 |
| $(H_2O)_{20}$ | Fused Cubes | 0.067 | 0.05 | 0.029 | 0.029 | 0.073 |
| $(H_2O)_{20}$ | Pentag. Dodec. | 0.034 | 0.066 | 0.033 | 0.018 | 0.149 |
| $(H_2O)_{25}$ | Isomer 2 | 0.029 | 0.049 | 0.037 | 0.023 | 0.109 |
| RMSD (Å) | | **0.026** | **0.046** | **0.029** | **0.017** | **0.132** |

clusters for which high-quality benchmark energies are available.[68] Table 4 shows the energies predicted by our model and several other force fields for which data is available. It is worth noting that both q-AQUA and MB-Pol include fits of the 2-body, 3-body, and in the case of q-AQUA, 4-body contributions to the energy. Therefore, the fact that we are able to achieve an MAE over all clusters intermediate between these models is a great result. Interestingly, MB-Pol seems to have a tendency to underestimate the binding energy of clusters as they get

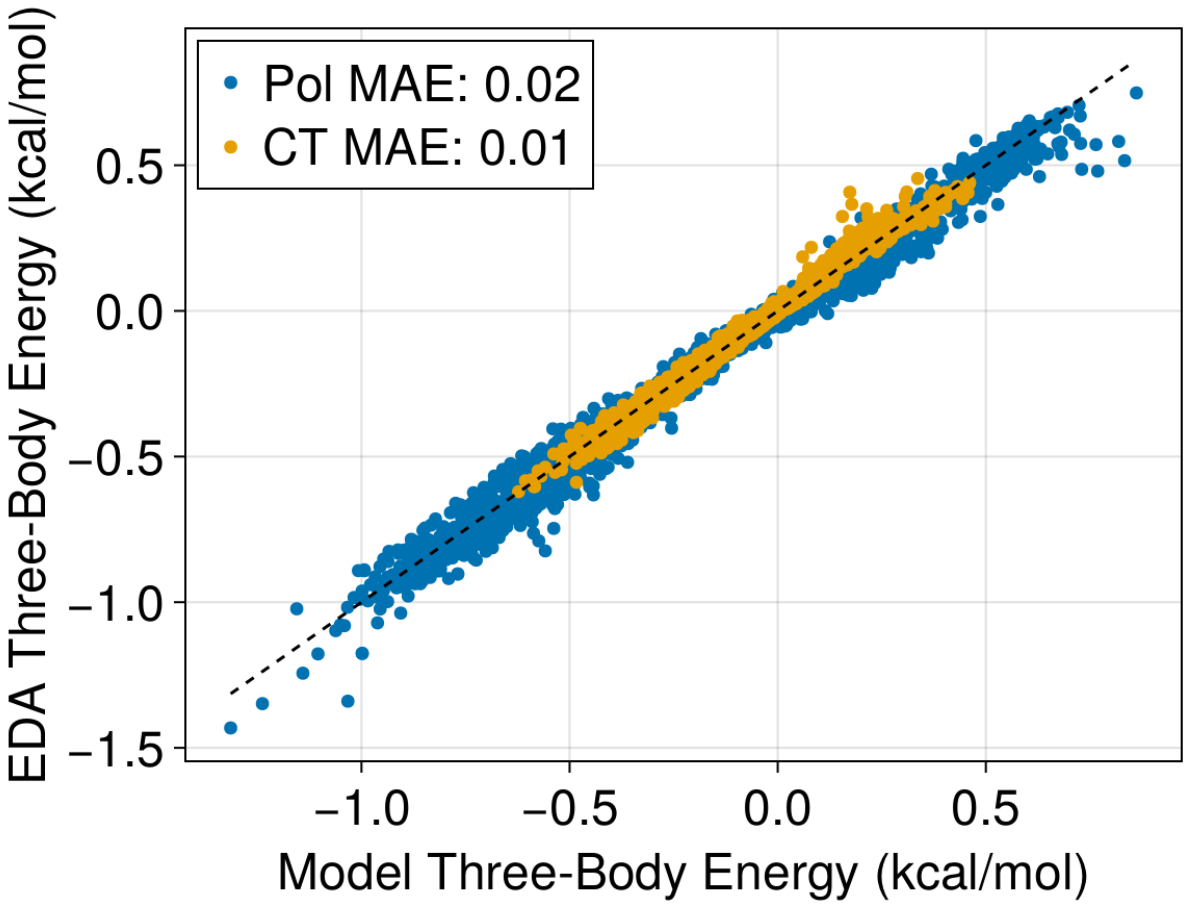



Figure 6: Correlation plot of the three-body contribution to polarization and charge transfer as computed by our model and with ωB97X-V/def2-QZVPPD. All water trimers in this plot are drawn from ion-water clusters and therefore span a wide range of configurations, some of which are atypical of liquid water. Since trimers drawn from a cluster may be disonnected, we enforce that the trimer have an absolute three-body contribution of at least 0.02 kcal/mol.

larger. Presumably this shortcoming has been corrected in the updated version of MB-Pol[69] which is reportedly at least as accurate as q-AQUA.

As highlighted in the methods section, the electrostatics and pauli repulsion used in this model are the same as those used in HIPPO.[9] It is very encouraging, therefore, to see that our model improves on the MAEs of the HIPPO model rather substantially. One of the main reasons we are able to achieve better MAEs is that our many-body energies, as shown in Figure 6, are both accurate and nearly unbiased. The HIPPO model systematically underestimates the many-body contributions to water cluster energies,[9] which is likely attributable

to the absence of many-body charge transfer.

Finally, note that we have included ωB97X-V/def2-QZVPPD in Table 4 since this is the reference method for our force field. It is clear that this level of theory compares very favorably to the CCSD(T)/CBS and MP2/CBS numbers used as reference. The MAE for ωB97X-V is also a useful benchmark since the best our force field can possibly perform is to reproduce these numbers. Interestingly, for larger clusters our model begins to slightly underestimate the energies compared to the references, but so does ωB97X-V/def2-QZVPPD.

In addition to accurate energetics, polarizable force fields should produce accurate geometries compared to electronic structure. To this end, Table 5 shows the root mean-squared deviation (RMSD) of water cluster structures optimized with various force fields and ωB97X-V/def2-QZVPPD compared to previously reported structures optimized with either CCSD(T) or MP2.[68] The average RMSD for our model manages to outperform MB-Pol and is comparable even to q-AQUA. Notably, ωB97X-V/def2-QZVPPD is even more accurate on this benchmark than q-AQUA which provides further evidence of the high accuracy of this functional.

Unfortunately, TTM2.1-F struggles at accurately reproducing the structures of water clusters. It is important to realize, however, that TTM2.1-F was published in 2006[62] and is itself a minor modification of TTM2-F which was published in 2002.[70] TTM2.1-F is quite simple compared to the present model and MB-Pol or q-AQUA, so it is not surprising that it falls short on certain benchmarks. We bring this up not to criticize TTM2.1-F, which was highly innovative and an inspiration in this work, but to emphasize that it is possible to have accurate energies but sub-optimal structures in the same model. Therefore, we suggest that any future water models should benchmark both energies and structures.

In order for a force field to be useful for theoretical spectroscopy, it must respect the relationships between structure and vibrational frequncies. In the case of water, this manifests as a linear relationship between the change in bond length and change in O−H stretching frequency.[54] This is sometimes colloquially referred to as Badger's rule since Richard Badger

famously pointed out a correspondence between equilibrium bond lengths and bond force constants.[57]

In the course of constructing this force field, we tested if it satisfied the expected slope $\Delta\omega$ vs $\Delta R_e$ of $\approx -19$ cm$^{-1}$/.001Å.[54] We found, instead, that the slope was around $\approx -10$ cm$^{-1}$/.001Å(See Fig. S2). This is quite an interesting slope since Boyer *et al.* show that a field-independent Morse oscillator, with parameters appropriate to water, will give a slope of $\approx -11$ cm$^{-1}$/.001Å. This was a strong indication that we should allow our bonding potential parameters to be modulated by the field along that bond. How this is done is shown in Eqs. 30 and 31.

The result of doing this is shown in Figure 7 we can more accurately reproduce the structure-frequency correlation in water which strongly indicates our model will be useful for spectroscopy. The necessity of coupling the bonding potential to the environment to accurately reproduce structure-frequency relationships in a force field is a new observation as far as we are aware. We consider this an excellent result given the simplicity of the field-dependent morse potential, especially since it requires no free parameters. We computed the necessary dipole derivatives from a simple O−H scan and found the parameters $\mu^{(1)} = 0.1654$ and $\mu^{(2)} = -0.01246$. If we do the same calculation with FQCT, we get $\mu^{(1)} = 0.1658$ and $\mu^{(2)} = -0.0104$. This indicates that as long as a force field has an accurate dipole surface, the dipole derivatives needed to compute the field-dependence of a morse potential can be computed directly from the force field.

This approach is easily extensible to other force fields and should immediately improve spectroscopic predictions. Additionally, the ability of FQCT to generate accurate structures is useful for cases where the force field is used to generate configurations for further analysis with electronic structure. As an aside, many force fields use harmonic bond potentials and the field-perturbed quantum harmonic oscillator is exactly solvable, so a similar modification can be made in that case.

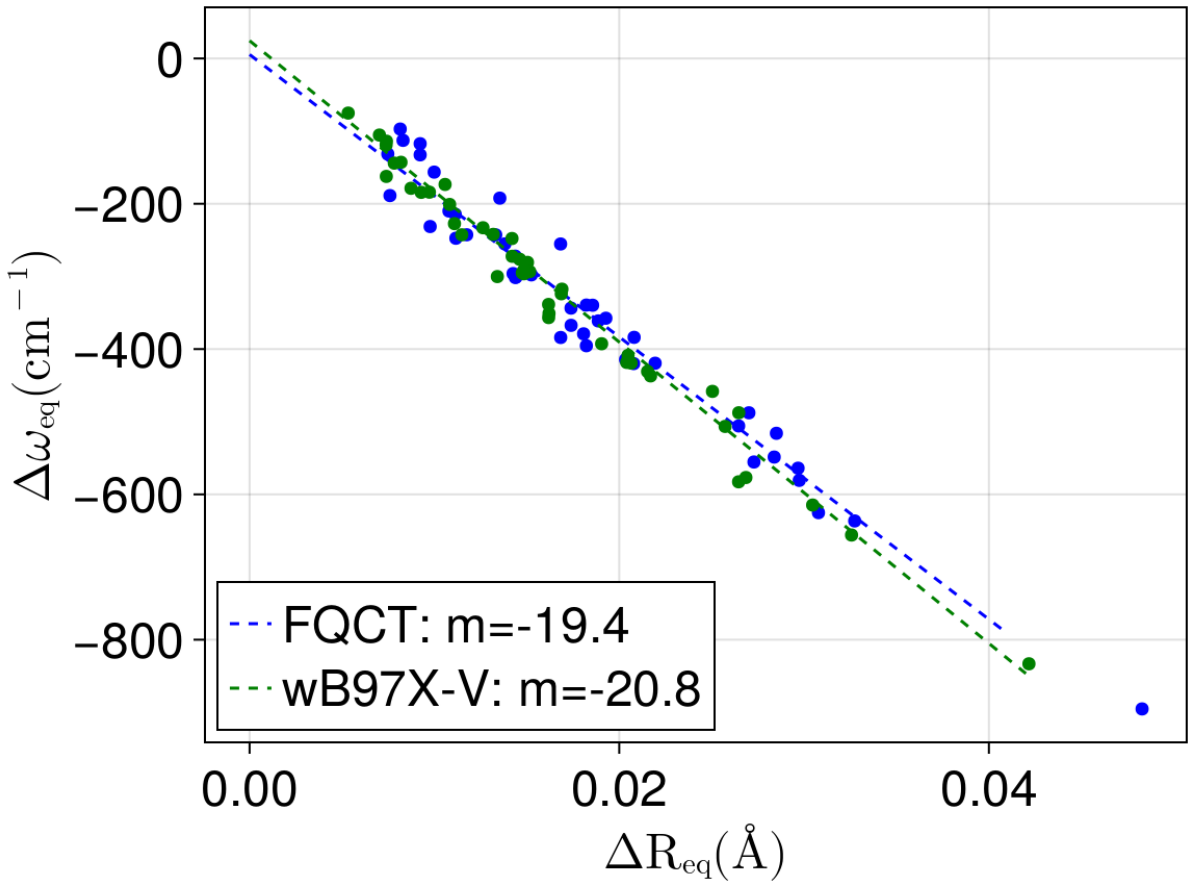



Figure 7: Correlation of $\Delta\omega$ vs $\Delta R_e$ over a collection of small water clusters using FQCT and $\omega$B97X-V/def2-QZVPPD. The water clusters are low-energy structures of $(H_2O)_{2\text{-}6}$ available with the paper. The linear fits are not constrained to pass through zero which explains the slightly large slopes compared to previous work.[54]

## Ion-Water Model

## Ion-Water Model

We now turn our attention to the performance of this model on ion-water clusters. First, we will look at simple scans for each of the ion-water dimers considered in this study. The scans in Figures 8 and 9 show the total energy curve in a solid colored line and the reference energy curve in a dashed colored line. In the SI, we include a variation on these plots where we show the error in each individual component of the scan.

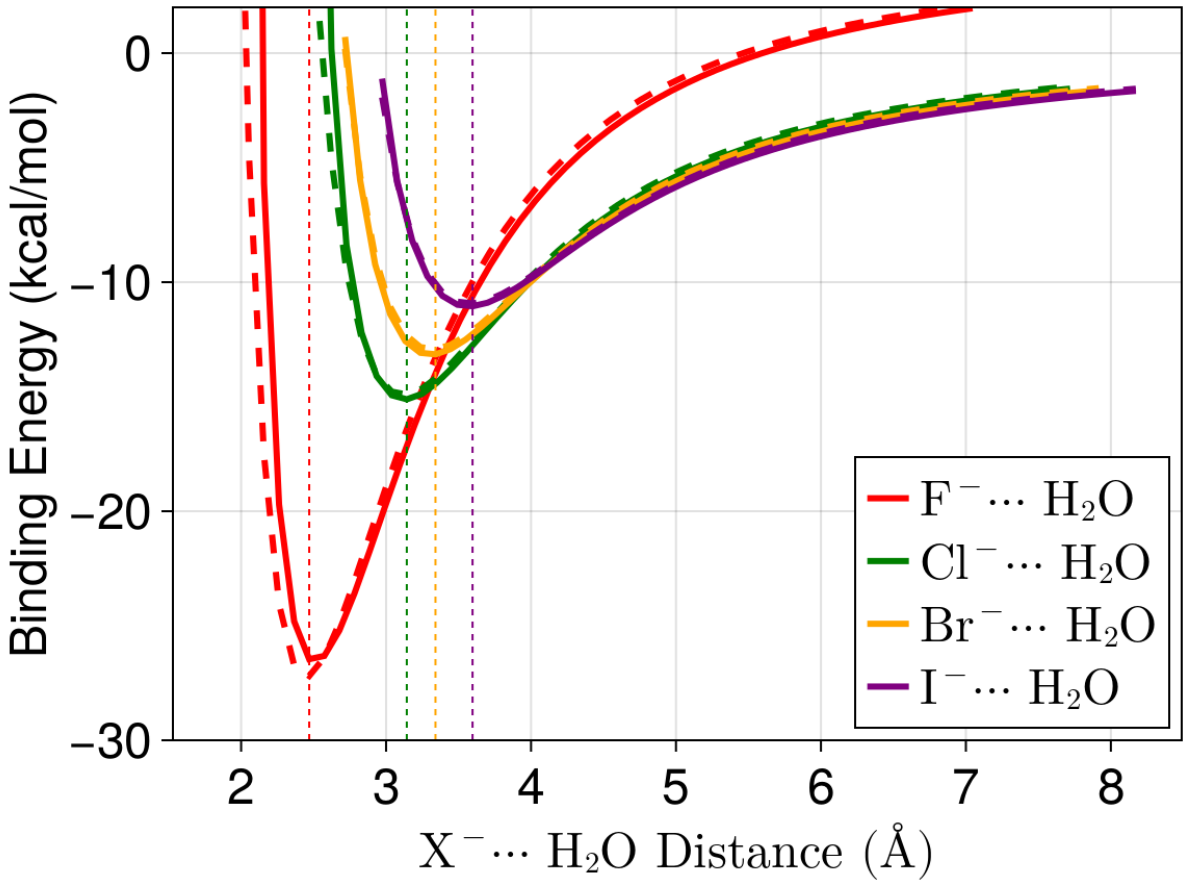



Figure 8: Scans of the $X^-(H_2O)$ ($X^-$=$F^-$, $Cl^-$, $Br^-$, and $I^-$) potentials where the geometry is held fixed at the equilibrium structure of each dimer. The solid curve corresponds to FQCT and the dashed curve corresponds to $\omega$B97X-V/def2-QZVPPD. Vertical dashed lines show the positions of each minimum with $\omega$B97X-V/def2-QZVPPD. Binding energies and harmonic frequencies are reported in Table 6.

Perhaps the most important of the dimer potentials in Figures 8 and 9 is the long-range attractive part of the potential which we are able to capture very accurately in all cases. Unsurprisingly, the largest errors are found for $F^-$ and $Li^+$ which are the most challening ions considered in this work due to their extremely short-range interactions. Another important feature of a dimer potential is the repulsive wall since in the condensed phase, many-body stabilization shortens the oxygen-ion distance beyond the dimer equilibrium distance. Looking at the errors of each term inside the equilbrium distance in Figures **??** and **??**, it is clear that the Pauli repulsion tends to be underestimated on the repulsive wall as is charge trans-

fer. On the other hand, electrostatics and polarization are overestimated on the repulsive wall. These effects are largest for smaller ions but in each case. Achieving accurate energies high up the repulsive wall is not particularly important beyond ensuring that the repulsive wall does not have a short-range singularity.

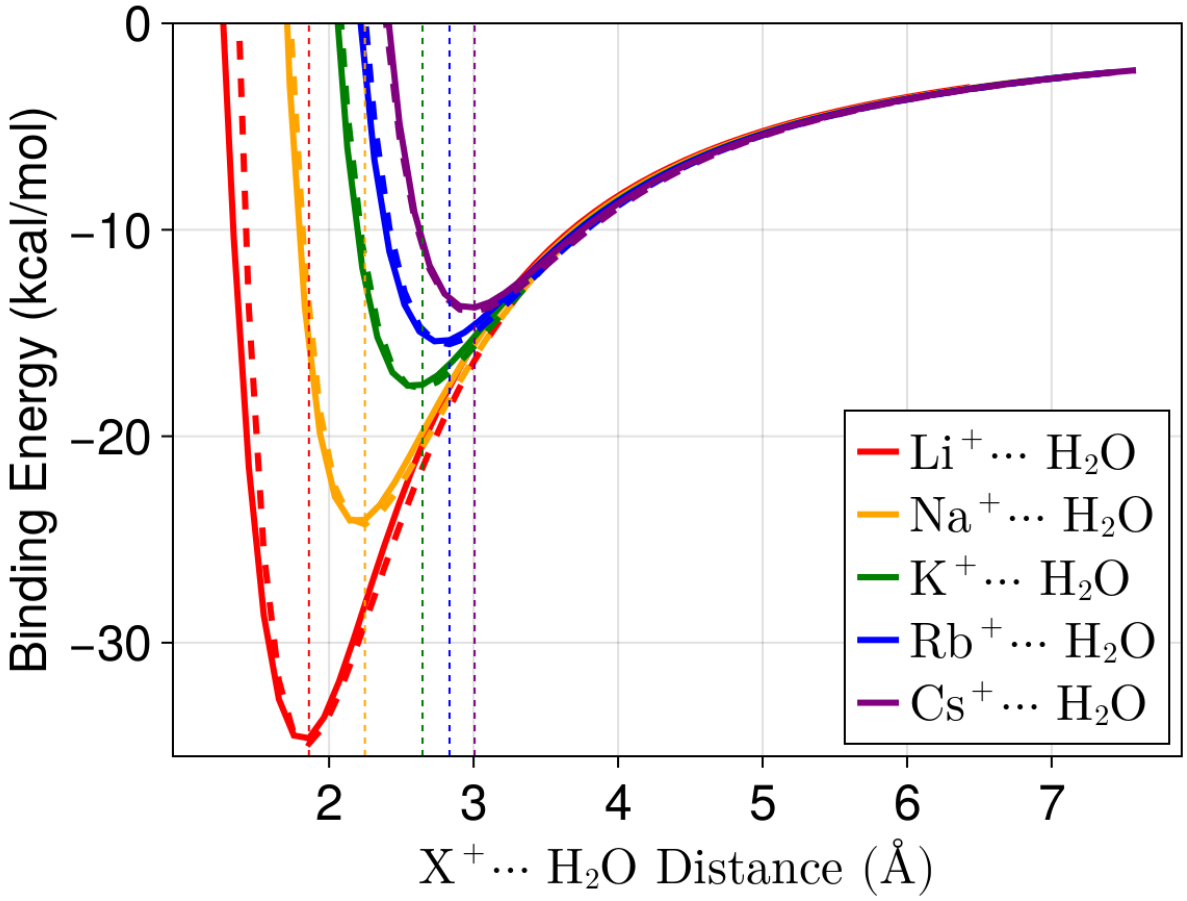



Figure 9: Scans of the $X^+(H_2O)$ ($X^+$=$Li^+$, $Na^+$, $K^+$, $Rb^+$, and $Cs^+$) potentials where the geometry is held fixed at the equilibrium structure of each dimer. The solid curve corresponds to FQCT and the dashed curve corresponds to $\omega$B97X-V/def2-QZVPPD. Vertical dashed lines show the positions of each minimum with $\omega$B97X-V/def2-QZVPPD. Binding energies and harmonic frequencies are reported in Table 6.

Since ion-water interactions are strong, they can result in large red-shifts of the underlying vibrational frequencies. To that end, we show the frequencies and binding energies of FQCT compared to $\omega$B97X-V in Table 6.

To begin, the binding energies of FQCT are extremely accurate compared to its reference

Table 6: Comparison of FQCT frequencies ($cm^{-1}$) and binding energies (kcal/mol) against $\omega$B97X-V/def2-QZVPPD for all ion-water dimers discussed in this paper.

Ion-Water Dimer Vibrational Frequencies

| $X^{+/-}(H_2O)$ | Method | NM1 | NM2 | NM3 | NM4 | NM5 | NM6 | $D_e$ |
|---|---|---|---|---|---|---|---|---|
| $F^-(H_2O)$ | FQCT | 365 | 723 | 1328 | 1806 | 2156 | 3842 | -26.7 |
| | $\omega$B97X-V | 384 | 569 | 1144 | 1702 | 2233 | 3916 | -27.2 |
| $Cl^-(H_2O)$ | FQCT | 197 | 395 | 767 | 1719 | 3427 | 3893 | -15.1 |
| | $\omega$B97X-V | 193 | 341 | 726 | 1678 | 3417 | 3919 | -14.9 |
| $Br^-(H_2O)$ | FQCT | 157 | 360 | 718 | 1713 | 3507 | 3894 | -13.2 |
| | $\omega$B97X-V | 151 | 288 | 657 | 1673 | 3521 | 3916 | -12.9 |
| $I^-(H_2O)$ | FQCT | 129 | 303 | 643 | 1701 | 3591 | 3895 | -11.1 |
| | $\omega$B97X-V | 118 | 220 | 579 | 1668 | 3623 | 3911 | -10.9 |
| $Li^+(H_2O)$ | FQCT | 301 | 512 | 535 | 1669 | 3677 | 3801 | -34.9 |
| | $\omega$B97X-V | 392 | 524 | 554 | 1681 | 3815 | 3882 | -34.9 |
| $Na^+(H_2O)$ | FQCT | 268 | 316 | 434 | 1679 | 3737 | 3848 | -24.2 |
| | $\omega$B97X-V | 307 | 367 | 437 | 1677 | 3830 | 3902 | -24.3 |
| $K^+(H_2O)$ | FQCT | 220 | 306 | 367 | 1680 | 3767 | 3873 | -17.6 |
| | $\omega$B97X-V | 213 | 359 | 369 | 1673 | 3833 | 3910 | -17.7 |
| $Rb^+(H_2O)$ | FQCT | 180 | 304 | 342 | 1680 | 3776 | 3881 | -15.4 |
| | $\omega$B97X-V | 178 | 347 | 351 | 1671 | 3836 | 3914 | -15.6 |
| $Cs^+(H_2O)$ | FQCT | 160 | 306 | 316 | 1679 | 3785 | 3889 | -13.8 |
| | $\omega$B97X-V | 157 | 327 | 339 | 1668 | 3836 | 3916 | -14.0 |

method. Indeed, the largest error is 0.7 kcal/mol for $F^-(H_2O)$, which is by far the most difficult case. The rest of the ions have binding energies accurate to wi thin a few tenths of a kcal/mol.

The harmonic frequencies in Table 6 are generally quite accurate. The O−H stretching modes, NM5, are fairly accurate for all anions especially considering the strength of interactions involved. This highlights that reproducing the structure-frequency correspondence shown in Figure 7 transfers to interactions with anions. Interestingly, the bending mode for anion-water dimers, NM4, is systematically blue-shifted. This is almost certainly attributable to the lack of a field-dependent contribution to the bending potential. Essentially, the anion attracts the other hydrogen atom causing the HOH angle to close more than it should. This effect would be counter-acted by a field-dependent angle potential which would tend to result in an opening of the HOH angle. In the future, we may explore the addition

of a field-dependent contribution to the bending potential. The low-frequency modes for the anions are also rather accurate but a bit overestimated.

Interestingly, the low-frequency modes for cation-water dimers are a bit more accurate than those for the anions except the frequencies tend to be slightly underestimated. The problem with the bending frequencies seen for the anions does not occur for the cations. On the other hand, the O−H stretching modes for cation-water dimers are red-shifted too much compared to $\omega$B97X-V.

Just as with water, many-body contributions to both polarization and charge transfer are important for ion-water systems. On a relative basis, both are less important than in water clusters simply because electrostatic interactions with ions are so large but that does not mean they can be neglected. In Figures 10A and 10B, we show the correlation of three-body polarization energies computed from FQCT and EDA. These energies are computed from 400 trimers drawn from ion-water clusters which we generated for this work.

Figures 10A and 10B demonstrate that our models of polarization and charge transfer are transferable even to very strong interactions. The attractive contributions to many-body polarization are captured exceptionally well, while some additional scatter arises for repulsive three-body polarization. The accuracy of many-body charge transfer with halides is similar regardless of whether the energy is attractive or repulsive.

Figures 11A and 11B show the same correlations we just discussed for halide-water trimers but for the alkali metal cations. The description of many-body polarization is somewhat worse than for halides, but is still quite accurate by consistently being around an MAE of 0.1 kcal/mol. Our experience is that modeling polarization for cations is generally more difficult than for the halides. This is likely the case since the the oxygen atom in water is very polarizable and cations interact directly with the oxygen.

Figure 11B is another illustration of the transferability and general accuracy of our charge transfer model. Interestingly, the errors in three-body charge transfer decrease going from $Li^+$ to $Na^+$ to $K^+$, which is expected since the magnitude of the energy also decreases. The

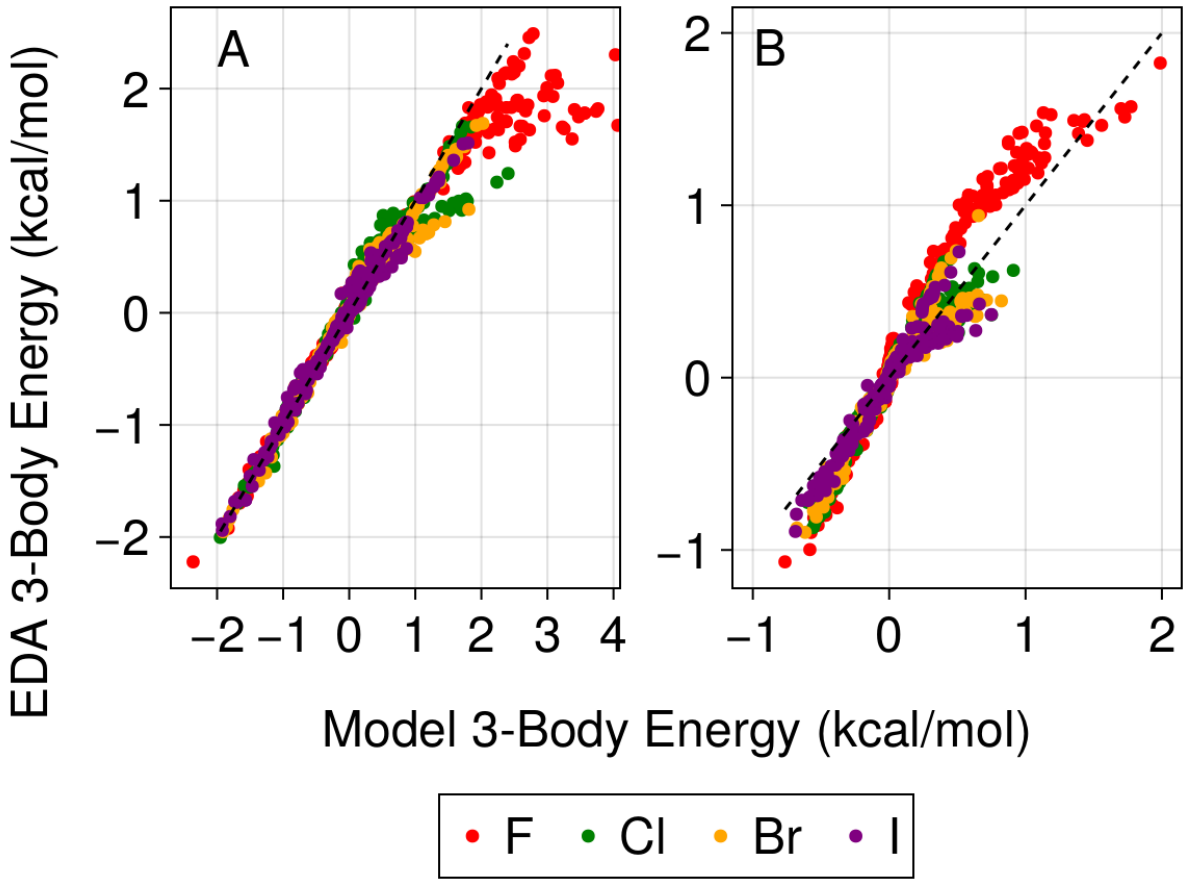



Figure 10: Correlation of three-body polarization (A) and charge transfer energies (B) between FQCT and EDA. All trimers are drawn from ion-water clusters optimized with $\omega$B97X-V/def2-TZVPPD and energies are computed with the def2-QZVPPD basis set. The 3-body polarization MAEs over the trimers in panel A are 0.208, 0.072, 0.060, and 0.046 kcal/mol for $F^-$, $Cl^-$, $Br^-$, and $I^-$, respectively. The 3-body charge transfer MAEs over the trimers in panel B are 0.129, 0.068, 0.056, and 0.044 kcal/mol for $F^-$, $Cl^-$, $Br^-$, and $I^-$, respectively.

MAEs for $Rb^+$ and $Cs^+$ increase to about 0.033 kcal/mol and the correlation is slightly skewed. In any case, many-body charge transfer is rather small for these ions and FQCT is still able to capture the trend accurately.

Overall, we are very pleased with the transferability of our polarization and charge transfer models from water to ions. Accurately reproducing many-body energies is a very stringent test of any model, but especially for ion-water systems where polarizable force fields have historically struggled.

# Ion-Ion Potentials

# Discussion

The necessity of coupling the bonding potential to the environment to accurately reproduce structure-frequency relationships in a force field is a new observation as far as we are aware. We consider this an excellent result given the simplicity of the field-dependent morse potential, especially since it requires no free parameters.

# Conclusions

We have described new approaches to modeling polarization and charge transfer which we parameterize against a very accurate energy decomposition analysis. In doing so, we have highlighted the importance of many-body charge transfer for aqueous systems. Our new model of charge transfer is able to quantitatively capture many-body charge transfer for both water and ion-water clusters. This model of charge transfer is enabled by the use of fluctuating charges in the polarization model.

We also show that fluctuating charges greatly improve the accuracy of polarizability derivatives. Accurate polarizability derivatives are essential for computing Raman spectra, which polarizable force fields have historically modeled very poorly. We therefore expect this model to be useful for theoretical spectroscopy among many other uses. To that end, we also show that our model can reproduce the structure-frequency correspondence central to hydrogen-bonded vibrations. Specifically, we found that accurately predicting the red-shift and bond elongations associated with hydrogen bonds required a field-dependent contribution to the bonding potential. The dipole derivatives which control the field-dependent bond potential can be computed from electronic structure or directly from the force field as long as the force field has an accurate dipole surface.

By using physically motivated damping functions, we are able to control the polarization

catastrophes which have otherwise inhibited accurate polarizable force fields for ions being developed. Indeed, we expect that this force field represents a turning point in the ability to model strong, short-range interactions with physical models. One can, of course, fit error corrections on top of this force field. We expect that achieving such high accuracy with a physical model will enable systems that are otherwise out of reach to be studied. These might include concentrated ionic solutions and divalent ions.

The most recent progress, arguably, has been made by explicitly fitting the terms of the MBE for ion-water interactions.[71,72] This approach has the drawback that generalizing the procedure to multi-component systems, such as a solution containing multiple types of ions, is made difficult by the combinatorial explosion of terms which need to be fit.

Even worse, one can only exclude charge transfer and charge penetration from force fields because these energies are strongly correlated to the Pauli repulsion (see Fig. 3). This correlation is not guaranteed to be consistent between systems, however, which may explain part of the historical difficulty in producing water models which generalize to heterogeneous systems.

# Methods

All distributed multipole calculations were carried out in the Orient program.[73]

# Acknowledgement

We acknowledge support from the U.S. National Science Foundation through Grant No. CHE-2313791. Computational resources were provided by the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract DE-AC02-05CH11231.

# Supporting Information Available

TODO

# References


(1) Xantheas, S. S. Cooperativity and hydrogen bonding network in water clusters. *Chemical Physics* **2000**, *258*, 225–231.

(2) Heindel, J. P.; Xantheas, S. S. The many-body expansion for aqueous systems revisited: II. Alkali metal and halide ion–water interactions. *Journal of chemical theory and computation* **2021**, *17*, 2200–2216.

(3) Herman, K. M.; Heindel, J. P.; Xantheas, S. S. The many-body expansion for aqueous systems revisited: III. Hofmeister ion–water interactions. *Physical Chemistry Chemical Physics* **2021**, *23*, 11196–11210.

(4) Leontyev, I.; Stuchebrukhov, A. Accounting for electronic polarization in non-polarizable force fields. *Physical Chemistry Chemical Physics* **2011**, *13*, 2613–2626.

(5) Bedrov, D.; Piquemal, J.-P.; Borodin, O.; MacKerell Jr, A. D.; Roux, B.; Schroeder, C. Molecular dynamics simulations of ionic liquids and electrolytes using polarizable force fields. *Chemical reviews* **2019**, *119*, 7940–7995.

(6) Heindel, J. P.; Xantheas, S. S. The many-body expansion for aqueous systems revisited: I. Water–water interactions. *Journal of Chemical Theory and Computation* **2020**, *16*, 6843–6855.

(7) Das, A. K.; Urban, L.; Leven, I.; Loipersberger, M.; Aldossary, A.; Head-Gordon, M.; Head-Gordon, T. Development of an advanced force field for water using variational energy decomposition analysis. *Journal of chemical theory and computation* **2019**, *15*, 5001–5013.

(8) Mao, Y.; Loipersberger, M.; Horn, P. R.; Das, A.; Demerdash, O.; Levine, D. S.; Prasad Veccham, S.; Head-Gordon, T.; Head-Gordon, M. From intermolecular interaction energies and observable shifts to component contributions and back again: A tale of variational energy decomposition analysis. *Annual review of physical chemistry* **2021**, *72*, 641–666.

(9) Rackers, J. A.; Silva, R. R.; Wang, Z.; Ponder, J. W. Polarizable water potential derived from a model electron density. *Journal of chemical theory and computation* **2021**, *17*, 7056–7084.

(10) Horn, P. R.; Mao, Y.; Head-Gordon, M. Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals. *Physical Chemistry Chemical Physics* **2016**, *18*, 23067–23079.

(11) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *The Journal of chemical physics* **1994**, *101*, 6141–6156.

(12) Applequist, J. A multipole interaction theory of electric polarization of atomic and molecular assemblies. *The Journal of chemical physics* **1985**, *83*, 809–826.

(13) Stern, H. A.; Rittner, F.; Berne, B.; Friesner, R. A. Combined fluctuating charge and polarizable dipole models: Application to a five-site water potential function. *The Journal of chemical physics* **2001**, *115*, 2237–2251.

(14) Kim, Y. S.; Kim, S. K.; Lee, W. D. Dependence of the closed-shell repulsive interaction on the overlap of the electron densities. *Chemical Physics Letters* **1981**, *80*, 574–575.

(15) Wheatley, R. J.; Price, S. L. An overlap model for estimating the anisotropy of repulsion. *Molecular Physics* **1990**, *69*, 507–533.

(16) Gavezzotti, A. Calculation of intermolecular interaction energies by direct numerical in-

tegration over electron densities. I. Electrostatic and polarization energies in molecular crystals. *The Journal of Physical Chemistry B* **2002**, *106*, 4145–4154.

(17) Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. Beyond Born–Mayer: Improved models for short-range repulsion in ab initio force fields. *Journal of chemical theory and computation* **2016**, *12*, 3851–3870.

(18) Van Vleet, M. J.; Misquitta, A. J.; Schmidt, J. New angles on standard force fields: Toward a general approach for treating atomic-level anisotropy. *Journal of Chemical Theory and Computation* **2018**, *14*, 739–758.

(19) Misquitta, A. J.; Stone, A. J.; Fazeli, F. Distributed multipoles from a robust basis-space implementation of the iterated stockholder atoms procedure. *Journal of Chemical Theory and Computation* **2014**, *10*, 5405–5418.

(20) Misquitta, A. J.; Stone, A. J. ISA-Pol: distributed polarizabilities and dispersion models from a basis-space implementation of the iterated stockholder atoms procedure. *Theoretical Chemistry Accounts* **2018**, *137*, 1–20.

(21) Stone, A. J. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chemical Physics Letters* **1981**, *83*, 233–239.

(22) Stone, A. J.; Alderton, M. Distributed multipole analysis: methods and applications. *Molecular Physics* **1985**, *56*, 1047–1064.

(23) Piquemal, J.-P.; Gresh, N.; Giessner-Prettre, C. Improved formulas for the calculation of the electrostatic contribution to the intermolecular interaction energy from multipolar expansion of the electronic distribution. *The Journal of Physical Chemistry A* **2003**, *107*, 10353–10359.

(24) Wang, Q.; Rackers, J. A.; He, C.; Qi, R.; Narth, C.; Lagardere, L.; Gresh, N.; Ponder, J. W.; Piquemal, J.-P.; Ren, P. General model for treating short-range electrostatic

penetration in a molecular mechanics force field. *Journal of chemical theory and computation* **2015**, *11*, 2609–2618.

(25) Rackers, J. A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. An optimized charge penetration model for use with the AMOEBA force field. *Physical Chemistry Chemical Physics* **2017**, *19*, 276–291.

(26) Wallqvist, A.; Karlström, G. A new non-empirical force field for computer simulations. *Chem. Scr. A* **1989**, *29*, 1989.

(27) Gordon, J. H. J. M. S. An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. *molecular physics* **1996**, *89*, 1313–1325.

(28) Rackers, J. A.; Ponder, J. W. Classical Pauli repulsion: An anisotropic, atomic multipole model. *The Journal of chemical physics* **2019**, *150*.

(29) Tang, K.; Toennies, J. P. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *The Journal of chemical physics* **1984**, *80*, 3726–3741.

(30) Stone, A. Distributed polarizabilities. *Molecular Physics* **1985**, *56*, 1065–1082.

(31) Ruth Le Sueur, C.; Stone, A. J. Localization methods for distributed polarizabilities. *Molecular Physics* **1994**, *83*, 293–307.

(32) Mortier, W. J.; Ghosh, S. K.; Shankar, S. Electronegativity-equalization method for the calculation of atomic charges in molecules. *Journal of the American Chemical Society* **1986**, *108*, 4315–4320.

(33) Chen, J.; Martínez, T. J. QTPIE: Charge transfer with polarization current equalization. A fluctuating charge model with correct asymptotics. *Chemical physics letters* **2007**, *438*, 315–320.

(34) Chen, J.; Hundertmark, D.; Martínez, T. J. A unified theoretical framework for fluctuating-charge models in atom-space and in bond-space. *The Journal of chemical physics* **2008**, *129*.

(35) Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chemical Physics* **1981**, *59*, 341–350.

(36) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. Simulation of Ca2+ and Mg2+ solvation using polarizable atomic multipole potential. *The journal of physical chemistry B* **2006**, *110*, 18553–18559.

(37) Mason, P. E.; Wernersson, E.; Jungwirth, P. Accurate description of aqueous carbonate ions: an effective polarization model verified by neutron scattering. *The Journal of Physical Chemistry B* **2012**, *116*, 8145–8153.

(38) Chung, M. K.; Wang, Z.; Rackers, J. A.; Ponder, J. W. Classical Exchange Polarization: An Anisotropic Variable Polarizability Model. *The Journal of Physical Chemistry B* **2022**, *126*, 7579–7594.

(39) Herman, K. M.; Stone, A. J.; Xantheas, S. S. Accurate calculation of many-body energies in water clusters using a classical geometry-dependent induction model. **2023**,

(40) Sheng, X.; Tang, K. T.; Toennies, J. P. A semiempirical potential for alkali halide diatoms with damped interactions I. Rittner potential. *Physical Chemistry Chemical Physics* **2022**, *24*, 24823–24833.

(41) Wang, W.; Yan, D.; Cai, Y.; Xu, D.; Ma, J.; Wang, Q. General Charge Transfer Dipole Model for AMOEBA-Like Force Fields. *Journal of Chemical Theory and Computation* **2023**, *19*, 2518–2534.

(42) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling

the origin of intermolecular interactions using absolutely localized molecular orbitals. *The Journal of Physical Chemistry A* **2007**, *111*, 8753–8765.

(43) Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Analysis of charge transfer effects in molecular complexes based on absolutely localized molecular orbitals. *The Journal of chemical physics* **2008**, *128*.

(44) Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Electron donation in the water–water hydrogen bond. *Chemistry–A European Journal* **2009**, *15*, 851–855.

(45) Liu, C.; Piquemal, J.-P.; Ren, P. Implementation of geometry-dependent charge flux into the polarizable AMOEBA+ potential. *The journal of physical chemistry letters* **2019**, *11*, 419–426.

(46) McCoy, A. B. The role of electrical anharmonicity in the association band in the water spectrum. *The Journal of Physical Chemistry B* **2014**, *118*, 8286–8294.

(47) Misquitta, A. J.; Welch, G. W.; Stone, A. J.; Price, S. L. A first principles prediction of the crystal structure of C6Br2ClFH2. *Chemical Physics Letters* **2008**, *456*, 105–109.

(48) Pracht, P.; Bohle, F.; Grimme, S. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Physical Chemistry Chemical Physics* **2020**, *22*, 7169–7192.

(49) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *Journal of chemical theory and computation* **2019**, *15*, 1652–1671.

(50) Rakshit, A.; Bandyopadhyay, P.; Heindel, J. P.; Xantheas, S. S. Atlas of putative minima and low-lying energy networks of water clusters n= 3–25. *The Journal of chemical physics* **2019**, *151*.

(51) Anatole von Lilienfeld, O.; Tkatchenko, A. Two-and three-body interatomic dispersion energy contributions to binding in molecules and solids. *The Journal of chemical physics* **2010**, *132*.

(52) Misquitta, A. J. Charge transfer from regularized symmetry-adapted perturbation theory. *Journal of chemical theory and computation* **2013**, *9*, 5313–5326.

(53) Stone, A. J.; Misquitta, A. J. Charge-transfer in symmetry-adapted perturbation theory. *Chemical Physics Letters* **2009**, *473*, 201–205.

(54) Boyer, M. A.; Marsalek, O.; Heindel, J. P.; Markland, T. E.; McCoy, A. B.; Xantheas, S. S. Beyond Badger's rule: The origins and generality of the structure–spectra relationship of aqueous hydrogen bonds. *The Journal of Physical Chemistry Letters* **2019**, *10*, 918–924.

(55) Mao, Y.; Horn, P. R.; Head-Gordon, M. Energy decomposition analysis in an adiabatic picture. *Physical Chemistry Chemical Physics* **2017**, *19*, 5944–5958.

(56) Aldossary, A.; Gimferrer, M.; Mao, Y.; Hao, H.; Das, A. K.; Salvador, P.; Head-Gordon, T.; Head-Gordon, M. Force Decomposition Analysis: A method to decompose intermolecular forces into physically relevant component contributions. *The Journal of Physical Chemistry A* **2023**, *127*, 1760–1774.

(57) Badger, R. M. A relation between internuclear distances and bond force constants. *The Journal of Chemical Physics* **1934**, *2*, 128–131.

(58) Thirman, J.; Engelage, E.; Huber, S. M.; Head-Gordon, M. Characterizing the interplay of Pauli repulsion, electrostatics, dispersion and charge transfer in halogen bonding with energy decomposition analysis. *Physical Chemistry Chemical Physics* **2018**, *20*, 905–915.

(59) Kirov, M. V.; Fanourgakis, G. S.; Xantheas, S. S. Identifying the most stable networks in polyhedral water clusters. *Chemical Physics Letters* **2008**, *461*, 180–188.

(60) Partridge, H.; Schwenke, D. W. The determination of an accurate isotope dependent potential energy surface for water from extensive ab initio calculations and experimental data. *The Journal of Chemical Physics* **1997**, *106*, 4618–4639.

(61) Burnham, C. J.; Li, J.; Xantheas, S. S.; Leslie, M. The parametrization of a Thole-type all-atom polarizable water model from first principles and its application to the study of water clusters (n= 2–21) and the phonon spectrum of ice Ih. *The Journal of chemical physics* **1999**, *110*, 4566–4581.

(62) Fanourgakis, G. S.; Xantheas, S. S. The flexible, polarizable, thole-type interaction potential for water (TTM2-F) revisited. *The Journal of Physical Chemistry A* **2006**, *110*, 4100–4106.

(63) Babin, V.; Leforestier, C.; Paesani, F. Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface, VRT spectrum, and second virial coefficient. *Journal of chemical theory and computation* **2013**, *9*, 5395–5403.

(64) Babin, V.; Medders, G. R.; Paesani, F. Development of a "first principles" water potential with flexible monomers. II: Trimer potential energy surface, third virial coefficient, and small clusters. *Journal of chemical theory and computation* **2014**, *10*, 1599–1607.

(65) Yu, Q.; Qu, C.; Houston, P. L.; Conte, R.; Nandi, A.; Bowman, J. M. q-AQUA: A many-body CCSD (T) water potential, including four-body interactions, demonstrates the quantum nature of water from clusters to the liquid phase. *The Journal of Physical Chemistry Letters* **2022**, *13*, 5068–5074.

(66) Hamm, P. 2D-Raman-THz spectroscopy: A sensitive test of polarizable water models. *The Journal of Chemical Physics* **2014**, *141*.

(67) Sidler, D.; Meuwly, M.; Hamm, P. An efficient water force field calibrated against intermolecular THz and Raman spectra. *The Journal of chemical physics* **2018**, *148*.

(68) Herman, K. M.; Xantheas, S. S. An extensive assessment of the performance of pairwise and many-body interaction potentials in reproducing ab initio benchmark binding energies for water clusters n= 2–25. *Physical Chemistry Chemical Physics* **2023**, *25*, 7120–7143.

(69) Zhu, X.; Riera, M.; Bull-Vulpe, E. F.; Paesani, F. MB-pol (2023): Sub-chemical accuracy for water simulations from the gas to the liquid phase. *Journal of Chemical Theory and Computation* **2023**, *19*, 3551–3566.

(70) Burnham, C. J.; Xantheas, S. S. Development of transferable interaction models for water. IV. A flexible, all-atom polarizable potential (TTM2-F) based on geometry dependent charges derived from an ab initio monomer dipole moment surface. *The Journal of chemical physics* **2002**, *116*, 5115–5124.

(71) Zhuang, D.; Riera, M.; Zhou, R.; Deary, A.; Paesani, F. Hydration structure of Na+ and K+ ions in solution predicted by data-driven many-body potentials. *The Journal of Physical Chemistry B* **2022**, *126*, 9349–9360.

(72) Caruso, A.; Zhu, X.; Fulton, J. L.; Paesani, F. Accurate modeling of bromide and iodide hydration with data-driven many-body potentials. *The Journal of Physical Chemistry B* **2022**, *126*, 8266–8278.

(73) Stone, A. J.; Dullweber, A.; Engkvist, O.; Fraschini, E.; Hodges, M. P.; Meredith, A.; Nutt, D.; Popelier, P.; Wales, D. ORIENT, version 4.6. *University of Cambridge, England* **2002**,

# TOC Graphic

Some journals require a graphical entry for the Table of Contents. This should be laid out "print ready" so that the sizing of the text is correct.
Inside the `tocentry` environment, the font used is Helvetica 8 pt, as required by *Journal of the American Chemical Society*.
The surrounding frame is 9 cm by 3.5 cm, which is the maximum permitted for *Journal of the American Chemical Society* graphical table of content entries. The box will not resize if the content is too big: instead it will overflow the edge of the box.
This box and the associated title will always be printed on a separate page at the end of the document.

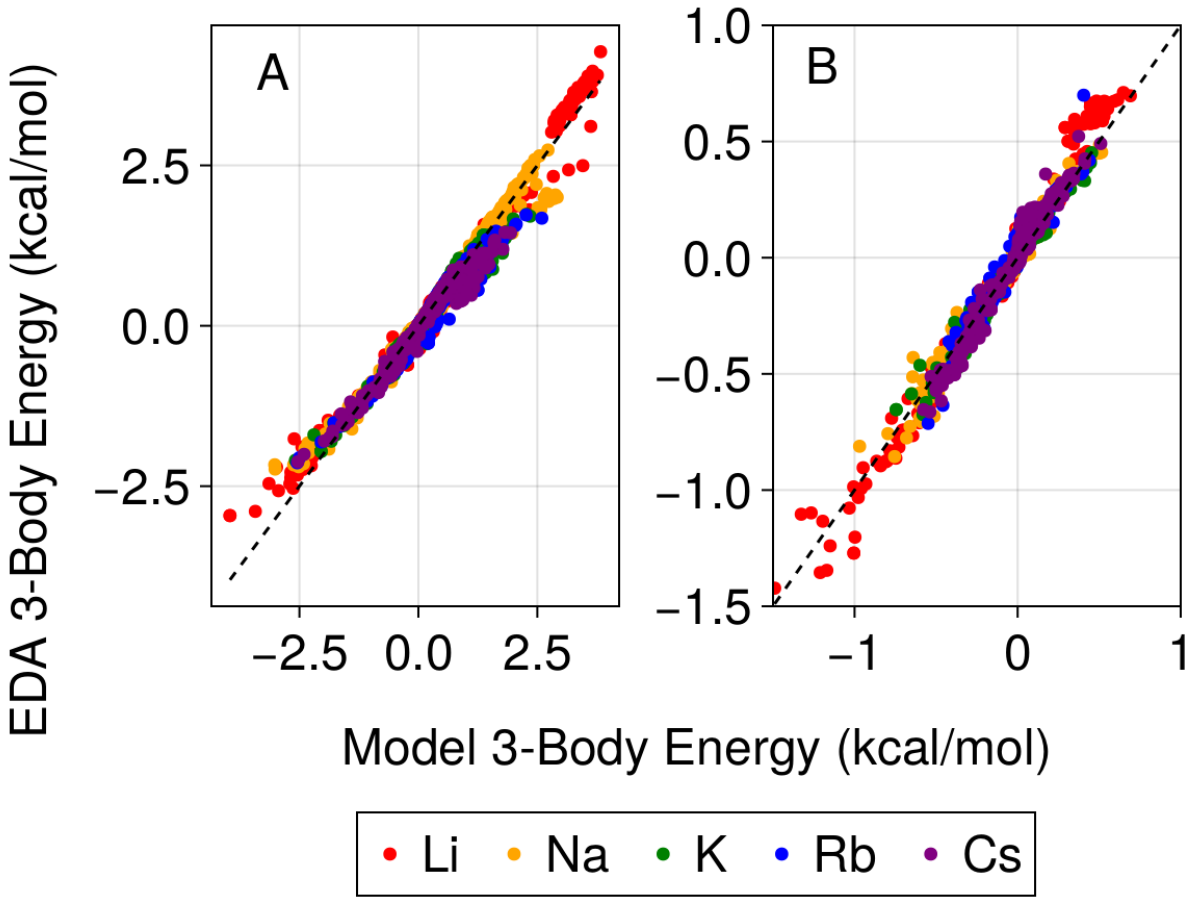



Figure 11: Correlation of three-body polarization (A) and charge transfer energies (B) between FQCT and EDA. All trimers are drawn from ion-water clusters optimized with $\omega$B97X-V/def2-TZVPPD and energies are computed with the def2-QZVPPD basis set. The 3-body polarization MAEs over the trimers in panel A are 0.122, 0.105, 0.080, 0.123, and 0.078 kcal/mol for $Li^+$, $Na^+$, $K^+$, $Rb^+$, and $Cs^+$, respectively. The 3-body charge transfer MAEs over the trimers in panel B are 0.050, 0.024, 0.021, 0.033, and 0.029 kcal/mol for $Li^+$, $Na^+$, $K^+$, $Rb^+$, and $Cs^+$, respectively.