Construction of Accurate Force Fields from Energy Decomposition Analysis: 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, California94720, United States*

‡*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California94720, United States*

¶*Departments of Bioengineering and Chemical and Biomolecular EngineeringUniversity of California, Berkeley, California94720, United States*

E-mail: thg@berkeley.edu

Abstract

This work describes a force field for water and atomic ions which aims to quantitaively reproduce each term of an energy decomposition analysis. In the model, polarization is handled in a manner that allows for charge fluctuations and induced dipoles. The model contains a new approach to modelling charge transfer which allows for explicit movement of charge between molecules. We show that this approach naturally describes many-body charge transfer by coupling into the polarization equations. Additionally, we highlight the fact that many-body charge transfer is non-negligible in aqueous systems. More here...

Introduction

Historically, there have been two main approaches to including polarization in force fields: fluctuating charges[1] or induced dipoles.[2] There have also been attempts to unify these approaches allowing for both charge rearrangements and induced dipoles.[3]

Our goal in this paper is to develop a new class of polarizable force field which is able to quantitatively reproduce all of the terms in an energy decomposition analysis (EDA).[4,5] The reason to do this is that 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. Additionally, by reproducing the EDA term-by-term, the force field is able to provide insights which many other models simply cannot because they do not include particular terms. For instance, many force fields do not include charge transfer or charge penetrations terms. Many force fields also use terms which package the Pauli repulsion and dispersion energies together. We will discuss this further in the results section, but by neglecting these terms, one limits the interpretability of the energies and forces predicted by a force field. 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. 1). This correlation is not guaranteed to be consistent between systems, however, which may explain part of the difficulty in producing water models which generalize to heterogeneous systems.

The model described here borrows ideas from many other force fields which have proven to be successful. Foremost among these is the use of a model electron density such as that employed in both the HIPPO[6] and the MASTIFF models.[7,8] Specifically, treating short-range interactions as being related to the overlap of a Slater density or Slater orbital is extremely useful. Especially since many of these short-range interactions need to be damped to prevent divergences, and having a model density naturally generates the appropriate damping functions. We argue, however, that polarization should be damped more strongly than the overlap of slater orbitals implies. This follows from the observation that most

damping functions take the form of a polynomial multiplied by an exponential. Ideally, force fields should be free of singularities, which means the polynomial in a damping function should be large enough to control the polynomial scaling of the interaction in question. This should probably go somewhere else. Near the polarization stuff I guess...

Theory

EDA splits the interaction energy into five components: Pauli repulsion, electrostatics, dispersion, polarization, and charge transfer. The force field described in this work will model each of these term by term. Note we will use a convention of referring to all energy terms in the force field with a V and all energy terms from electronic structre with an E.

Our approach borrows ideas from the density overlap hypothesis[7,9–11] 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 atom-centered density overlaps. One way of doing this was developed thoroughly by Misquitta and others[12,13] based on iterated stockholder atoms which can be used to define Slater-like densities for atoms in molecules. Very similar ideas have been developed by Rackers *et al.*[6] and we will make reference to the differences between the various approaches as relevant. In any case, since this approach has been discussed extensivly, we will only summarize the salient points.

The form of the charge density used in the model is,

$$\rho(r) = \frac{Qb^3}{8\pi} e^{-br} + Z\delta(r) \tag{1}$$

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_{ii}^{ρ}, of two identical Slater-like atomic densities at different

locations, $\rho_i(\boldsymbol{r}_i)$ and $\rho_i(\boldsymbol{r}_j)$, is,

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

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_{ρ}^{ij}, has a more complicated form, but it has been shown that setting $b_{ij} = \sqrt{b_i b_j}$ allows the expression for S_{ii}^{ρ} to be used for different atom types to a good approximation.[7] The polynomial prefactor in the overlap is,

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

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

Rackers *et al.* utilize a similar idea in the HIPPO model[6] 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.

We utilize both Slater density overlap and Slater orbital overlap in different terms within the model. It is important to realize that either approach really just provides a physically meaningful indication of what is "short-range" for a pair of atoms. Therefore, combining the idea of overlapping Slater densities with overlapping Slater orbitals seems natural to us.

Electrostatics

Our description of electrostatics comes from a traditional point multipole approach and a charge penetration contribution. We will refer to these contributions as the DMA energy and CP energy, respectively. Our working definition of charge penetration comes by taking

the classical electrostatic energy from EDA minus the interaction energy when using Stone's distributed multipole analysis[14,15] out to hexadecapoles on all atoms.

$$E^{CP} = E^{elec}_{EDA} - E^{elec}_{DMA} \tag{4}$$

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, which is essential to reproduce the classical electrostatic energy in EDA. All distributed multipole calculations were carried out in Stone's Orient program.[16]

Charge penetration is described by treating each atom as having both a positively charged core and negatively charged shell. We use the same set of damping functions as derived by Rackers and Ponder which are appropriate for the Slater density in Eq. 1.[6] Very similar models have been applied successfully in MB-UCB[17] based on functional forms proposed by Piquemal[18] and others.[19,20]

Considering the interactions of the collection of cores and shells, which are expanded in multipoles, results in the following electrostatic energy expression:

$$V_{elec} = \sum_{i<j} Z_i T_{ij} Z_j + Z_i \boldsymbol{T}^{damp}_{ij} \boldsymbol{M}_j + Z_j \boldsymbol{T}^{damp}_{ji} \boldsymbol{M}_i + \boldsymbol{M}_i \boldsymbol{T}^{overlap}_{ij} \boldsymbol{M}_j \tag{5}$$

In eq. 5, the first term represents repulsive core-core interactions where $T_{ij} = 1/r_{ij}$ with Z_i the core charge on the ith 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. We allow all atoms to have multipoles up to the quadrupole. Note that the core-shell interactions are damped according to

$$\boldsymbol{T}^{damp}_{ij} = \begin{bmatrix} 1 & \nabla & \nabla^2 \end{bmatrix} \cdot \left(\frac{1}{r_{ij}} f^{damp}_{ij}(r_{ij}) \right) \tag{6}$$

The shell-shell interactions are damped, albeit with a different damping function, and the corresponding interaction tensor is written as:

$$\boldsymbol{T}_{ij}^{overlap} = \begin{bmatrix} 1 & \nabla & \nabla^2 \\ \nabla & \nabla^2 & \nabla^3 \\ \nabla^2 & \nabla^3 & \nabla^4 \end{bmatrix} \cdot \left(\frac{1}{r_{ij}} f_{ij}^{overlap}(r_{ij}) \right) \tag{7}$$

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

$$f_{ij}^{damp}(r_{ij}) = 1 - \left(1 + \frac{1}{2} b_j r_{ij} \right) e^{-b_j r_{ij}} \tag{8a}$$

$$f_{ij}^{overlap}(r_{ij}) = 1 - \left(1 + \frac{11}{16} b_{ij} r_{ij} + \frac{3}{16} (b_{ij} r_{ij})^2 + \frac{1}{48} (b_{ij} r_{ij})^3 \right) e^{-b_{ij} r_{ij}} \tag{8b}$$

Simply comparing the forms of Eq. 8a with Eq. 8b makes it clear why the core-shell model of charge penetration works. Namely, as long as b_j is similar to b_{ij}, which it almost always will be, then the polynomial in Eq. 8b will be larger than that in Eq. 8a. This means the overlap damping function is stronger and hence the shell-shell repulsions will be damped more strongly than the core-shell attractions. Therefore, at short-range the electrostatic energy will be more attractive than a point multipole expansion of the density would suggest, which is exactly what is meant by charge penetration.

The damping function in Eq. 8a can be derived directly from the form of the Slater density in Eq. 1 by computing its electrostatic potential. The damping function in Eq. 8b can be derived from a symmetrized coulomb integral where each density interacts with the damped potential generated by the other density.[6] 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. The form of all relevant damping functions have been derived and summarized elsewhere.[6]

Polarization

We introduce a combined fluctuating charge (FQ) and induced dipole model of electronic polarization. There are two reasons we have pursued this combined polarization approach. The first is quite simple: we want to reproduce all terms from EDA, one of which is charge transfer. If charges are not allowed to vary, then one cannot model the explicit transfer of charge between molecules. The second reason is that atomic polarizabilities naturally contain both charge-flow and induced dipole contributions.[21] Typically, the charge-flow contributions are localized away,[22] but our approach does not require nonlocal charge flow polarizabilities.

The FQ contribution to our polarization model is a modification of the electronegativity equalization model (EEM) of polarization.[23] In EEM, the energy of a molecule is expanded to second-order as a function of charge, then, these charges are allowed to interact. Mathematically, this takes the following form,

$$V(\boldsymbol{q}) = \sum_i \chi_i q_i + \frac{1}{2}\sum_i \eta_i q_i^2 + \sum_{i<j} \frac{q_i q_j}{r_{ij}} \tag{9}$$

In Eq. 9, χ_i represents the electronegativity of atom i and η_i is the atomic hardness of atom i. The principle of electronegativity equalization states that at equilibrium, the electronegativity of all atoms will become equal. This allows the charges in an atom to be determined by solving a system of linear equations. There are several known shortcomings of EEM. The first is the long-range transfer of charge between molecules, even at infinite distance, which is unrealistic for the dielectric systems studied here.[24,25] Our solution to this problem is to only allow charge rearrangements within a molecule and not between molecules. This constraint can be introduced using Lagrange multipliers. A second limitation of EEM is that the approach in Eq. 9 gives back the total charge of atoms in a molecule. This is somewhat awkward because we want to separate the multipolar electrostatics, which depends on the total charge, from the polarization contribution which should only depend on the external potential and field experienced by each atom. Hence, we drop the linear term

and focus only on the fluctuation of charges around the reference charge used for permanent electrostatics. Another way of viewing this is that we are equalizing electronegativity around an "already equalized" state. 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,

$$V(\delta \boldsymbol{q}) = \frac{1}{2}\sum_i \eta_i \delta q_i^2 + \sum_i \delta q_i V_i + \sum_{i<j} \frac{\delta q_i \delta q_j}{r_{ij}} + \sum_\alpha \lambda_\alpha \sum_{i\in\alpha} \delta q_i \tag{10}$$

To summarize, Eq. 10 allows charges to rearrange, δq_i, in response to an external potential, V_i, with a quadratic penalty determined by the atomic hardness, η_i. The charge rearrangements are constrained to only occur between atoms such that rearranged charges in a molecule sum to the total charge of a molecule. Therefore, the model has N lagrange multipliers, λ_α, where N is the number of molecules in the system. These molecules could be water or an ion in this work. Note that all fluctuating charges (and induced dipoles) are allowed to interact whether they are in the same molecule or not.

We also allow electric fields due to the environment to induce dipoles on all atoms. The energy of an induced dipole in an electric field, $\boldsymbol{E}$, including mutual polarization is,

$$V(\boldsymbol{\mu}^{ind}) = -\frac{1}{2}\sum_i \boldsymbol{\mu}_i^{ind} \cdot \boldsymbol{E}_i^{damp} + \sum_{i<j} \boldsymbol{\mu}_i^{ind} \boldsymbol{T}_{ij}^{\mu\mu} \boldsymbol{\mu}_j^{ind} \tag{11}$$

In Eq. 11, the induced dipoles at each polarizable site, $\boldsymbol{\mu}_i^{ind}$, interact with an external electric field, $\boldsymbol{E}_i^{damp}$, and with each other via the dipole interaction tensor. 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 $f_{ij}^{overlap}/r_{ij}$.

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 δq_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 succintly as follows:

$$\begin{pmatrix} \boldsymbol{T}^{qq} & \mathbf{1}_\lambda & \boldsymbol{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & 0 & 0 \\ -\boldsymbol{T}^{\mu q} & 0 & \boldsymbol{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{V} \\ \boldsymbol{Q} \\ \boldsymbol{E} \end{pmatrix} \tag{12}$$

The solution vector in Eq. 12 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 η and the 3×3 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 ith 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.

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

$$T_{ij}^{qq} = f_3^{overlap} \frac{1}{r_{ij}} \tag{13a}$$

$$\boldsymbol{T}_{ij}^{q\mu} = f_5^{overlap} \frac{-\boldsymbol{r}_{ij}}{r_{ij}^3} \tag{13b}$$

$$\boldsymbol{T}_{ij}^{\mu\mu} = \left(f_7^{overlap} \frac{\boldsymbol{r}_{ij} \otimes \boldsymbol{r}_{ij}}{r_{ij}^5} - f_5^{overlap} \frac{\mathbf{1}}{r_{ij}^3} \right) \tag{13c}$$

The interaction tensors in Eq. 13 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 used in the permanent electrostatic interactions. Note, however, that we have done something unusal. Naturally, T_{ij}^{qq} would be damped with $f_1^{overlap}$ since this is the correct damping function for the interaction of two

Slater charge densities. This would mean that each damping function is lowered by one order in $\boldsymbol{T}_{ij}^{q\mu}$ and $\boldsymbol{T}_{ij}^{\mu\mu}$.

We have found that if we do not increase the damping orders used in mutual polarization, then the polarization energy tends to be systematically too large. This effect is especially large for ions which are polarizable and highly charged and hence strongly induce polarization. This modification is not simply empirical though. Let us consider the induced dipoles, which will be roughly proportional to the field experienced at each atom. Therefore, the scaling of the interaction between two induced dipoles is $1/r_{ij}^3$ (the same scaling as for permanent dipoles) multiplied by the rate at which the polarizing field decays. For instance, the field due to a dipole decays as $1/r_{ij}^3$ which means the mutual induced dipole interactions will decay as $1/r_{ij}^6$ under that field. The damping function $f_5^{overlap}$ turns out to be a fifth-order polynomial multiplied by an exponential. Ideally, this would be at least sixth-order to control the significant contribution of dipole fields to the magnitude of induced dipoles. This modification of mutual polarization is very important as systems get larger such that some molecules have unusually large induced multipoles. As an aside, the damping functions generated by the usual Thole damping procedure,[26] are exponentials multiplied by first-, second-, and third-order polynomials for charge-charge, charge-dipole, and dipole-dipole interactions, respectively.[26] The small orders of these polynomials certainly contribute to the historic difficulty of controlling polarization between ions and water.[27,28]

Normally, the dipole polarizability is treated as a constant in polarizable force fields, but Chung *et al.* have pointed out that the polarizability of ions is significantly diminished in the aqueous phase.[29] 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

$$\boldsymbol{\alpha}_i^{-1} = R_i \begin{pmatrix} \alpha_{xx,i}^{-1} & 0 & 0 \\ 0 & \alpha_{yy,i}^{-1} & 0 \\ 0 & 0 & \alpha_{zz,i}^{-1} \end{pmatrix} R_i^T + \mathbb{1} \sum_j k_i^{damp} \left(\frac{\bar{\alpha}_i \bar{\alpha}_j}{r_{ij}^6} \right) \tag{14}$$

The first term in Eq. 14 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 analogoulsy. 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. We use the dimensionless length scale, $\frac{\bar{\alpha}_i \bar{\alpha}_j}{r_{ij}^6}$, to indicate when damping should become important. $\bar{\alpha}_i$ is the mean dipole polarizability of atom i. k_i^{damp} is an atom-specific parameter which modulates the increase of the inverse polarizability.

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. 14 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.

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,

$$V_{pol,sr} = \sum_{i<j} a_{ij}^{sr} f_4^{TT}(x_{ij}) \left(\frac{1}{2} \frac{\bar{\alpha}_i + \bar{\alpha}_j}{r_{ij}^3} \right)^{4/3} S_{ij}^{\rho} \tag{15}$$

While Eq. 15 is much more empirical than the rest of the force field, it seems capture the effect of quadrupole polarization, which is known to be important in water[30] but is very computationally expensive for the magnitude of the effect. In Eq. 15, $\bar{\alpha}_i$ 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. 21, and $a_{ij}^{sr} = a_i^{sr} a_j^{sr}$ is the pairwise parameter fit for this term. This term tends to have a magnitude for water that is only a bit larger than what would one expect for quadrupole polarization (around 20%). Eq. 15 is clearly not a true model for quadrupole polarization since it is pairwise-additive while quadrupole polarization would make small, nonzero many-body contributions. Similar terms have been suggested before for capturing the short-range polarization of ions.[31]

Charge Transfer

Charge transfer is the most difficult of the terms in EDA to model. This is because there is no classical analogue for charge transfer. One common approach to capturing charge transfer is to use a simple exponential dependent on the distance between atoms.[6] This captures the main effect which is the short-range exponential stabilization due to charge delocalization. Unfortunately, many-body charge transfer is non-negligible and this effect will be completely missed when using just exponentials. Another attempt is to essentially treat charge transfer the same way as polarization and solve a set of induced dipole equations.[17,32] This has the benefit that it can capture many-body charge transfer. One drawback of this approach is that 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. Perhaps a bigger problem is 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.

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.

$$V_{i\to j}^{CT} = a_{i\to j}^{CT} S_{ij}^{\rho} \tag{16a}$$

$$V_{j\to i}^{CT} = a_{j\to i}^{CT} S_{ij}^{\rho} \tag{16b}$$

$$V_{direct}^{CT} = \sum_{i<j} V_{i\to j}^{CT} + V_{j\to i}^{CT} \tag{16c}$$

As shown in Eq. 16, the forward and backward contributions to charge transfer are directly proportional to the density overlap. We take inspiration from perturbation theory which shows, roughly, that the amount of charge transferred between two molecules is proportional to the energy associated with forward and backward charge transfer.[33–35] 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

$$\Delta Q_{i\to j}^{CT} = \frac{V_{i\to j}^{CT}}{\epsilon_{i\to j}} \tag{17a}$$

$$\Delta Q_{j\to i}^{CT} = \frac{V_{j\to i}^{CT}}{\epsilon_{j\to i}} \tag{17b}$$

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.[33]

This approach is novel by allowing charge to explicitly move between fragments. This is achieved by modifying the molecular charge constraints used in Eq. 12. The charge constraint for a fragment A will then take the form,

$$Q_A^{CT} = Q_A + \sum_{i\in A} \sum_{j\notin A} \Delta Q_{j\to i}^{CT} - \Delta Q_{i\to j}^{CT} \tag{18}$$

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"[33] 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.

There is one technical point worth noting about this model. 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 $\Delta Q^{CT}_{i\to j}$ and $\Delta Q^{CT}_{j\to i}$. This is not difficult or expensive to evaluate, but because it is a rather unusual gradient term, we wanted to point this out clearly.

Pauli Repulsion

The original aim of the density overlap model was to model the Pauli repulsion energy.[10,36,37] The density overlap model of Pauli repulsion results in a formally exponential repulsion at short-range. However, Rackers and Ponder have made a convinving argument that the appropriate scaling of Pauli repulsion is actually of the form $E^{exch} \propto e^{-b_{ij}r_{ij}}/r_{ij}$.[6,38] While the exponential is the more important 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.

The multipolar pauli repulsion model has been discussed extensively elsewhere[38] so we only recapitulate the main points. The basic idea is that Pauli repulsion energy between a pair of atoms is proportional to S^2/r_{ij} where S is the pseudo-orbital overlap. The pseudo-orbital is defined as $\sqrt{\rho}$ where ρ is the density in Eq. 1.

Therefore, the Pauli repulsion energy can be written as

$$V_{Pauli} = \sum_{i<j} \frac{K^q_{ij} S^2_q + K^\mu_{ij} S^2_\mu + K^\Theta_{ij} S^2_\Theta}{r_{ij}} \tag{19}$$

In Eq. 19, S^2 represents the orbital overlap squared with different contributions from charges, dipoles, and quadrupoles. This is critically important since Pauli repulsion turns out to be 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,[38] these proportionality constants mean multipoles which handle repulsion are proportional to the actual electrostatic multipoles. One nice thing about this approach is that the calculation of electrostatics and multipolar Pauli repulsion differs only in the choice of damping function. Therefore, significant computation can be shared between the two 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,

$$V_{disp} = \sum_{i<j} f_6^{TT}(x_{ij}) \frac{C_{6,ij}}{r_{ij}^6} \tag{20}$$

In Eq. 20, $C_{6,ij}$ 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[39] which was originally derived to damp short-range dispersion,

$$f_n^{TT}(x_{ij}) = 1 - e^{-x_{ij}} \sum_{k=0}^{n} \frac{x_{ij}^k}{k!} \tag{21}$$

The appropriate form of x for the tail of a Slater electron density has been derived before[7] and takes the form,

$$x_{ij} = b_{ij} r_{ij} - \frac{2b_{ij}^2 r_{ij}^2 + 3b_{ij} r_{ij}}{b_{ij}^2 r_{ij}^2 + 3b_{ij} r_{ij} + 3} \tag{22}$$

Note that the TT damping functions, Eq. 21, 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 $f^{overlap}$.

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.[40] 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^+, Cs^+, Mg^{2+}, and Ca^{2+}. For all ions, we also ran a 10ps *ab initio* molecular dynamics simulation at 500K with ωB97X-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[41] which uses the semi-empirical GFN2-XTB[42] 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[43] 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 ωB97X-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.

All energies and forces that are used in fitting parameters of the force field are computed at the ωB97X-V/def2-QZVPPD level of theory. In the cases where clusters are optimized at ωB97X-V/def2-TZVPPD, we recompute the energies of those clusters and any derived sub-structures with ωB97X-V/def2-QZVPPD.

Parameterization

When parameterizing this force field, we fit each term against only the EDA contribution to that particular energy. We also try to ensure that the force field reproduces all phsyically meaningful monomer properties including the dipole moment, dipole derivatives, molecular polarizability, and polarizability derivatives. Optimization of parameters is done using simple gradient descent against the root mean-square deviation (RMSD) of predicted and EDA energies. For Pauli repulsion and electrostatics 200 random water dimers from the datasets described above are used in fitting. For dispersion, polarization, and charge transfer, 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. 4. Additionally, 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. At this point, we 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.

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,

$$L_{pol} = \sqrt{\frac{\sum_{i=1}^{N}(E_i^{FF} - E_i^{EDA})^2}{N}} + w_1||\boldsymbol{\alpha}^{FF} - \boldsymbol{\alpha}^{EDA}|| + w_2||\frac{\partial \boldsymbol{\alpha}^{FF}}{\partial \boldsymbol{r}} - \frac{\partial \boldsymbol{\alpha}^{EDA}}{\partial \boldsymbol{r}}|| \tag{23}$$

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 0.01 respectively. This, in essence, forces the molecular polarizability to be reproduced exactly while allowing for error in the polarizability derivatives which are much more difficult to reproduce.

The charge transfer energy and dispersion energies are simply fit against the RMSD from their EDA energies. Note that 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. 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.[8,44]

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 for only dimers. This procedure essentially results in improved error cancellation which we find is still necessary for a robust force field despite the effort to create a physically meaningful 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.

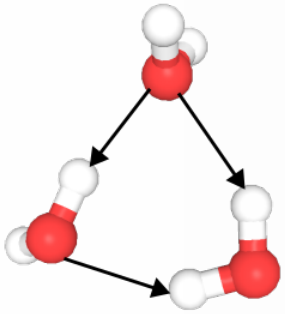
Results and discussion

Fortuitous Cancellation of Charge Transfer and Charge Penetration

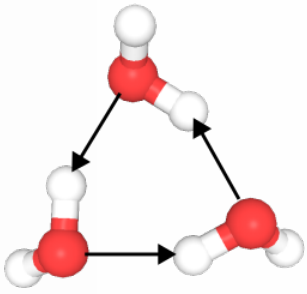
Before describing the outcomes of applying our model to various water clusters and ion-water clusters, let us discuss two important problems which we set out to solve in the construction of this model. First, it is broadly understood that there are contributions to the non-additive component of molecular interactions other than classical mutual polarization. By classical mutual polarization, we mean what some authors call induction.[30] Induction aims to describe how multipoles are induced under an applied field, field gradient, etc. and how these induced multipoles interact further. The reason this is classical is that it does not describe the effect of charge delocalization which we refer to as charge transfer. Note that charge transfer can be understood as energetic stabilization associated with tunneling of electrons between molecules and hence grows exponentially at short distances.[45]

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.[45,46] If we consider that charge transfer involves the transfer of actual charge between molecules, which is cer-

tainly the case in EDA,[47] 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. It can be seen that there must be a penalty for having non-integer charges since the energy of a system with a fractional number of electrons is exactly a linear interpolation of the energy of the two nearest states with integer charges.[48] Therefore, if there were no penalty for moving a fraction of an electron from one molecule to another, then the system would not be in the ground state electronic arrangement, which we assume to be true in this model.



	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 1: 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 ωB97X-V/def2-QZVPPD level of theory. 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 1 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.[49]

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 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 optimal cooperativity of the induced dipoles on each molecule.[50] 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. 4, 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?

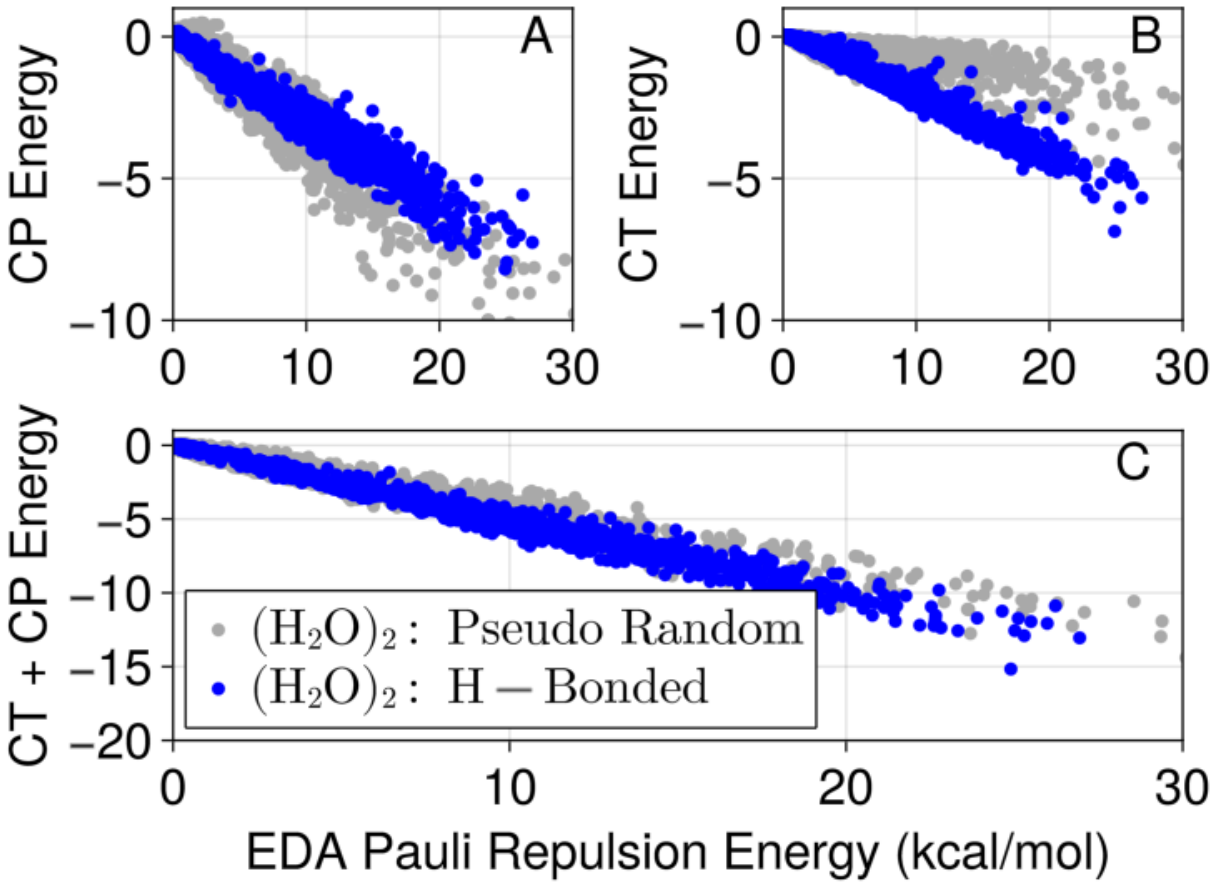


Figure 2: Correlation between (A) charge penetration energy (CP), (B) charge transfer energy (CT), and (C) their sum against Pauli repulsion energy. Energies are computed at the ωB97X-V/def2-QZVPPD level of theory. See text for discussion.

Figure 2 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 2 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 2 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 1 and 2 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 3 we make a comparison between the dipole surface of this model and the reference surface computed with ωB97X-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.

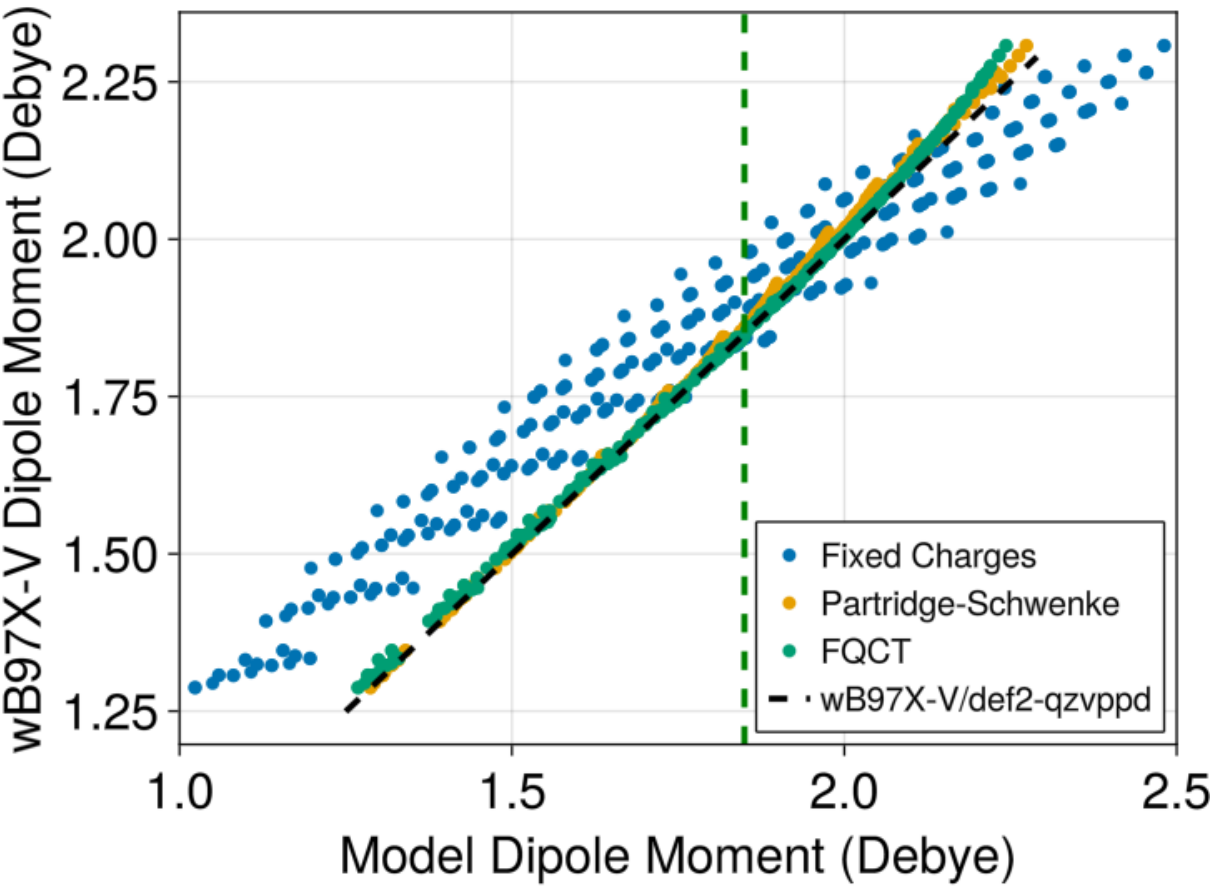


Figure 3: The dipole surface of water for all structures taken from a 3-D scan of water internal coordinates with a deformation energy less than 20 kcal/mol. The black dashed line shows the values computed with ωB97X-V/def2-QZVPPD. The green dashed line corresponds to the experimental gas-phase dipole moment of water of 1.85 Debye.

The orange points in Figure 3 are the dipole surface associated with the Partridge-Schwenke (PS) water monomer surface.[51] The PS potential and dipole surface are used by many water models such as the TTM models,[52,53] MB-Pol,[54,55] and q-aqua.[56] 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.[53] 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 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 ωB97X-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 five decimal places in bohr^3. The quoted polarizability values correspond to a water molecule with the x-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 1, 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 polarizbility derivatives in Table 1 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 polarization is turned on and found that the zz polarizability derivatives are still reproduced

Table 1: 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 x-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
Ox	-	4.04/2.92/-0.13	-	-
Oy	5.15/1.58/2.04	-	4.45/3.24/-2.03	1.50/0.0/0.0
Oz	-	-	-	-
Hx	-4.61/-2.57/0.78	-2.02/-1.46/0.06	-2.53/-0.60/-0.78	-1.39/0.0/0.0
Hy	-2.57/-0.79/-1.02	-1.68/-1.25/-0.08	-2.22/-1.62/1.02	-0.75/0.0/0.0
Hz	-	-	-	-

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. This indicates that one of the main reasons water models have historically predicted Raman intensities very poorly[57] 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. This indicates that an interesting future direction for advanced force fields is the construcion of geometry-dependent polarizability surfaces. In much the same way that recent force fields employ charge flux

to ensure accurate dipole surfaces, future models will benefit from accurate polarizability surfaces. Indeed, some recent work indicates that simple linear interpolation based on internal coordinates may be sufficient for water.[30] Linear models of charge flux are accurate for organic molecules,[58,59] but the same remains to be seen for the polarizability surface of organic molecules.

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.[60] The polarizability derivatives were not reported. The polarization is described using anisotropic polarizabilities but no fluctuating charges which, based on the contents of Table 1, lead us to believe this model likely has more accurate polarizability derivatives.

Water Model

Table 2: Comparison of the MAE of all terms in the EDA against predictions of our water model for water dimers, trimers, tetramers, and pentamers.

	MAE of EDA Terms (kcal/mol)			
	$(H_2O)_2$	$(H_2O)_3$	$(H_2O)_4$	$(H_2O)_5$
Deform.				
Pauli				
Disp.				
Elec.				
Pol.				
CT				
Total				

INSERT TABLE WITH MAES AND RMSDS FOR REFERENCE CLUSTERS WITH COMPS TO OTHER FFS

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 all trimers extracted from the ion-water clusters used in parameterizing this 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.

INSERT FIGURE OF 3B CT and POL

Acknowledgement

TODO

Supporting Information Available

TODO

References

(1) 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.

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

(3) 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.

(4) 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.

(5) 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.

(6) 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.

(7) 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.

(8) 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.

(9) 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.

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

(11) Gavezzotti, A. Calculation of intermolecular interaction energies by direct numerical integration over electron densities. I. Electrostatic and polarization energies in molecular crystals. *The Journal of Physical Chemistry B* **2002**, *106*, 4145–4154.

(12) 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.

(13) 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.

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

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

(16) 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**,

(17) 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.

(18) 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.

(19) 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.

(20) 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.

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

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

(23) 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.

(24) 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.

(25) 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*.

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

(27) 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.

(28) 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.

(29) 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.

(30) 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**,

(31) 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.

(32) 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.

(33) 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.

(34) 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*.

(35) 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.

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

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

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

(39) 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.

(40) 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.

(41) 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.

(42) 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.

(43) 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*.

(44) 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*.

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

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

(47) 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.

(48) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz Jr, J. L. Density-functional theory for fractional particle number: Derivative discontinuities of the energy. *Physical Review Letters* **1982**, *49*, 1691.

(49) 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.

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

(51) 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.

(52) 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.

(53) 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.

(54) 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.

(55) Babin, V.; Medders, G. R.; Paesani, F. Development of a "first principles" water poten-

tial with flexible monomers. II: Trimer potential energy surface, third virial coefficient, and small clusters. *Journal of chemical theory and computation* **2014**, *10*, 1599–1607.

(56) 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.

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

(58) Sedghamiz, E.; Nagy, B.; Jensen, F. Probing the importance of charge flux in force field modeling. *Journal of chemical theory and computation* **2017**, *13*, 3715–3721.

(59) Yang, X.; Liu, C.; Walker, B. D.; Ren, P. Accurate description of molecular dipole surface with charge flux implemented for molecular mechanics. *The Journal of Chemical Physics* **2020**, *153*.

(60) 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*.

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.