Polarizable Multipolar Molecular Dynamics Using Distributed Point Charges

Mike Devereux, Marco Pezzella, Shampa Raghunathan and Markus Meuwly∗

Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland

E-mail: m.meuwly@unibas.ch

ABSTRACT

Distributed point charge models (DCM) and their minimal variants (MDCM) have been integrated with tools widely used for condensed-phase simulations, including a virial-based barostat and a slow-growth algorithm for thermodynamic integration. Minimal DCM is further developed with a systematic approach to reduce fitting errors in the electrostatic interaction energy and a new fragment-based approach offers considerable speedup of the MDCM fitting process for larger molecules with increased numbers of off-centered charged sites. Finally, polarizable (M)DCM is also introduced in the present work. The developments are used in condensed-phase simulations of popular force fields with commonly applied simulation conditions. (M)DCM equivalents for a range of widely used water force fields and for fluorobenzene (PhF) are developed and applied along with the original models to evaluate the impact of reformulating the electrostatic term. Comparisons of the MEP, electrostatic interaction energies, and bulk properties from molecular dynamics simulations for a range of models from simple TIPnP (n = 3–5) to the polarizable, multipolar iAMOEBA models for water and an existing quadrupolar model for PhF confirm that DCMs retain the accuracy of the original models, providing a homogeneous, efficient, and generic point
charge alternative to a multipolar electrostatic model for force field development and multilevel simulations.

1. INTRODUCTION

Empirical force fields (FFs) are routinely used for simulating a multitude of chemical and biochemical phenomena.\textsuperscript{1–5} Commonly employed FFs divide interactions into intra- and intermolecular terms that include point charges (PCs) for Coulomb interactions. While interactions between nuclear-centered point charges are rapid to evaluate which allows application to large condensed-phase systems and relatively long timescales, their accuracy is compromised as they do not correctly describe charge density anisotropy.\textsuperscript{6} A new generation of FFs aims to overcome these drawbacks, either using higher-order multipolar electrostatics\textsuperscript{7,8} in methods such as AMOEBA (atomic multipole optimized energetics for biomolecular applications),\textsuperscript{9–12} SIBFA (sum of interaction between fragments ab initio computed)\textsuperscript{13–17} and QCTFF (Quantum Chemical Topological Force Field),\textsuperscript{18} or using Gaussian functions to directly describe the underlying charge density in methods such as EFP (effective fragment potential),\textsuperscript{19} GEM (gaussian electrostatic model),\textsuperscript{20,21} and NEMO (nonempirical molecular orbital).\textsuperscript{22} On the other hand, the use of higher-order atomic multipoles, while resulting in improved accuracy, introduces non-negligible computational overhead due to the additional complexity and increased number of terms that need to be evaluated.\textsuperscript{23–25}

An alternate tractable approach is to represent the MEP as a truncated multipole expansion transformed into a set of appropriately distributed point charges. Charges can be placed in fixed arrangements relative to the nuclei,\textsuperscript{25} or machine learning can be used to replace fixed arrangements with a minimal set of optimally positioned off-center charges.\textsuperscript{26} Recently it was demonstrated that these Distributed Charge Models (DCMs)\textsuperscript{25} and Minimal Distributed Charge Models (MDCMs)\textsuperscript{26} can be implemented into widely used molecular dy-
dynamics software as an alternative to conventional, PC-based energy terms. The use of point charges for representing charge density anisotropy reduces the complexity of Coulomb terms relative to a traditional multipolar formalism considerably, allowing efficient MD simulations while maintaining the accuracy of a truncated multipole expansion.

The versatility of (M)DCMs (i.e. distributed charge models with and without machine learning optimization) additionally yields a homogeneous implementation of different types of electrostatic models (nuclear-centered charges, off-center charges and multipolar electrostatics) using a single routine, with combinations of models of different complexities in a single simulation – so-called ‘multilevel’ simulations. Its compatibility with remaining standard bonded and non-bonded FF terms promises to make adaptation of next-generation (multipolar) FF electrostatics straightforward in widely used simulation packages.\textsuperscript{1,3,27,28}

The ability to generate models of increasing accuracy by increasing the number of charges in an MDCM fit offers an important tool to force field developers to carefully balance the accuracy of a model with the computational cost incurred from adding each additional charge. Models for the moiety or moieties of interest, such as a solute or protein ligand and immediate environment, can be created at the highest level of detail, while remaining interacting species can be optimally fitted to balance accuracy in the potential energy surface with computational efficiency to reach the system sizes and timescales required to sufficiently sample the relevant phase space. This multilevel approach is akin to the more familiar mixed quantum mechanical/molecular mechanical (QM/MM) treatments which also employ methods of different accuracy for different parts of a simulation system.\textsuperscript{29} Similarly, promising new models can equally be incorporated from one force field into another by transferring their parameters, and by refitting parameters of interacting neighbors at a level of detail that works optimally with that model, rather than combining existing models of different complexities that may not be compatible.
In this work recent advances in the implementation of (M)DCM are exploited and applied to condensed-phase simulations, including an intermediate fragment fitting step to improve efficiency of the MDCM fitting process for larger systems, improved error handling based on analysis of the relationship between errors in the fitted MEP and errors in the electrostatic interaction energy, integration with barostats for simulations in the isothermal-isobaric ensemble and slow-growth routines for thermodynamic integration calculations. Different (M)DCM models are developed to replace the electrostatic terms in several water force fields commonly used in chemical and biomolecular simulations and in a multipolar force field for PhF. In the first section of the results (M)DCM representations are generated for water models of increasing complexity ranging from the simple but widely used “TIP3P”\textsuperscript{30} to the multipolar, polarizable “iAMOEBA” potential.\textsuperscript{31} All terms of each original force field other than the electrostatic term are left untouched, requiring particularly close agreement with the original charge model to avoid reparametrization. Then, a similar approach is taken for the PhF molecule, demonstrating the applicability to a solute molecule in a condensed-phase aqueous environment. Comparisons of energies and bulk properties in each case are used to demonstrate the accuracy of a distributed charge approach to a more computationally complex multipolar description of molecular electrostatics.

2. Background

2.1 DCM

Multipolar force fields are based on the fact that any charge distribution can be represented as a series expansion, where the successive terms are multipole moments of increasing rank.\textsuperscript{6,32} Nuclear centers are typically used as convenient origins to locate ‘atomic multipoles’, as is the case in the “distributed multipole analysis” (DMA),\textsuperscript{33,34} “Atoms in Molecules” (AIM),\textsuperscript{8}
and AMOEBA,\textsuperscript{10} to improve convergence of the multipole expansion as opposed to using a single molecular origin. As the rank of these terms increases from dipole to quadrupole to octupole and beyond, the contribution that they make to the electric field in regions beyond the extent of the original charge distribution decays increasingly rapidly with distance. It is therefore often possible to truncate the series expansion at the atomic quadrupole moment\textsuperscript{35,36} while maintaining accuracy of the electrostatic potential outside the molecular surface. This is especially true if the terms of the expansion are fitted to the electrostatic potential rather than derived directly from the electron density.\textsuperscript{24,37}

For a discrete distribution of \( n \) charges a multipole expansion truncated at the quadrupole moment can be expressed using spherical harmonics as:

\[
Q_{00} = \sum_{i=1}^{n} q_i \\
Q_{10} = \sum_{i=1}^{n} q_i r_{z,i} \\
Q_{11c} = \sum_{i=1}^{n} q_i r_{x,i} \\
Q_{11s} = \sum_{i=1}^{n} q_i r_{y,i} \\
Q_{20} = \sum_{i=1}^{n} \frac{1}{2} q_i (3r_{z,i}^2 - r_{x,i}^2 - r_{y,i}^2) \\
Q_{22c} = \sum_{i=1}^{n} \sqrt{\frac{3}{4}} q_i (r_{x,i}^2 - r_{y,i}^2) 
\]

where \( q_i \) is point charge \( i \), \( r_{x,i} \) is the \( x \)-coordinate of \( i \) and \( Q_{lm} \) is the total atomic multipole moment of rank \((l, m)\).\textsuperscript{6} For a continuous charge density an analogous volume integral replaces the summation.

DCM is based on the fact that the converse is also true, i.e. any truncated multipole expansion, even one derived from a continuous charge density, can be represented by a suitable arrangement of discrete point charges.\textsuperscript{25,38} An illustrative but general example is an octahedral charge arrangement, where the magnitude of the charge at each vertex of the octahedron is analytically determined to exactly reproduce all multipole moments up to
quadrupole according to:

\[
q(d_q,0,0) = \frac{Q_{00}}{6} + \frac{Q_{11c}}{2d_q} - \frac{Q_{20}}{6d_q^2} + \frac{Q_{22c}}{2\sqrt{3}d_q^2}
\]

\[
q(-d_q,0,0) = \frac{Q_{00}}{6} - \frac{Q_{11c}}{2d_q} - \frac{Q_{20}}{6d_q^2} + \frac{Q_{22c}}{2\sqrt{3}d_q^2}
\]

\[
q(0,d_q,0) = \frac{Q_{00}}{6} + \frac{Q_{11s}}{2d_q} - \frac{Q_{20}}{6d_q^2} - \frac{Q_{22s}}{2\sqrt{3}d_q^2}
\]

\[
q(0,0,d_q) = \frac{Q_{00}}{6} + \frac{Q_{10}}{2d_q} + \frac{Q_{20}}{3d_q^2}
\]

\[
q(0,0,-d_q) = \frac{Q_{00}}{6} - \frac{Q_{10}}{2d_q} + \frac{Q_{20}}{3d_q^2}
\]

(2)

Here, \(d_q\) is the fixed distance of the charges from the nuclear coordinate of an atom. Note that the \(Q_{21c}, Q_{21s}\) and \(Q_{22s}\) quadrupole moment components vanish if the correct local axis system for an atom is chosen.25

It is therefore possible to replace all 6 nonzero multipole moments by 6 point charges. The total multipole moments of the charge distribution will exactly match the multipole expansion up to truncation rank, so the two will differ only by their higher order terms, i.e. octupole and beyond. These terms can either be kept small by reducing the distance \(d_q\), or deliberately enhanced to potentially provide accuracy beyond the truncated multipole expansion by fixing \(d_q\) to reproduce some of the higher order multipole moments of the reference atom. The main advantage of such an approach is a considerable reduction in complexity of the terms, as shown explicitly for the quadrupole–quadrupole interaction in section 1 of the SI. Thus, in DCM no fitting is required. Rather, the multipole moments are converted into a distributed charge arrangement based on analytical formulae.

For both multipolar and DCM approaches torques are generated by the off-centered charges or multipole moments that need to be distributed across the surrounding nuclei. As is generally the case in multipolar force fields, the torques of DCM models are applied to the nuclei that define the local (atomic) axis system of each charge (Figure 1), as described elsewhere.25
Figure 1: Local axis systems required to convert between the various electrostatic models presented in the text and to implement them into CHARMM and OpenMM. H-atom axes are equivalent except for the case of DCM, where one H-atom shares the same local axis system as its bonded O-atom neighbor.

### 2.2 MDCM

The fixed charge arrangements of DCM can be further optimized by releasing constraints on charge positions to exploit spatial degrees of freedom, creating minimal distributed charge models.\(^{26}\) For this, MDCM employs machine-learning to determine charge positions and magnitudes such that a predefined accuracy in the MEP is attained using a minimal number of off-centered charges. Differential Evolution\(^{39}\) was found to be effective in this context.\(^{26}\) Unlike DCMs, MDCMs are not constrained to maintain the same atomic multipole moments as a multipolar reference model, but are fitted directly to describe the MEP around a molecule. Increasing numbers of charges can be added until an MDCM representation reaches a required level of accuracy, with the possibility to even improve beyond what is possible with a multipolar model truncated at quadrupole.\(^{26}\) After fitting, each MDCM charge is assigned to a nucleus and the MDCM arrangements are implemented in MD simulations using the same framework (local axis systems, electrostatic cut-offs etc.) as a DCM model.
2.3 Polarization

For improved accuracy and physical rigour and realism, polarization interactions were also included. This makes (M)DCM models viable for emulating the iAMOEBA water model for which polarizable sites were assigned to the atomic nuclei. Here, the ‘direct’ (non-self-consistent) approach was employed\textsuperscript{40} which allows direct comparison with the original iAMOEBA model.\textsuperscript{31} In this approach, isotropic polarizabilities at nuclear sites are used to add induced dipoles to atoms as a function of the electric field generated by static multipole moments of surrounding atoms only (the field generated by other induced dipoles is ignored). For (M)DCM, this means that the electric field is generated by the point charges of surrounding atoms only. The total polarization energy is therefore:

\begin{equation}
V_{\text{pol}} = \sum_{i=1}^{N} \alpha_i E(r_i)^2
\end{equation}

\begin{equation}
E(r_i) = \sum_{j=1}^{N_{bi}} \sum_{l=1}^{N_{qj}} \lambda_{ij} q_{lj} \hat{r}_{il} \frac{1}{R_{il}^2}
\end{equation}

where the polarization energy \( V_{\text{pol}} \) is determined by a sum over all \( N \) atoms of their scalar (isotropic) polarizabilities \( \alpha_i \) multiplied by the square of the electric field \( E \) at the atom’s nuclear position \( r_i \). The electric field at the nuclear coordinate of atom \( i \) is evaluated by summing over each DCM charge \( q_l \) of the \( N_q \) DCM charges belonging to atom \( j \), for each atom in the list of \( Nb \) nonbonded partners of atom \( i \) within simulation cut-offs. \( \hat{r}_{il} \) is a unit vector in the direction of charge \( l \) from polarizable center \( i \), \( R_{il} \) is the distance from the nucleus of atom \( i \) to charge \( l \). The damping function \( \lambda_{ij} \) used in AMOEBA\textsuperscript{9} is also adopted here, with functional form:

\begin{equation}
\lambda_{ij} = 1 - \exp \left( -a \left( \frac{R_{il}}{(\alpha_i \alpha_j)^{1/6}} \right)^3 \right)
\end{equation}
where \( a = 0.23616 \ \text{Å}^{-1} \). Although the purpose of the damping function in the original AMOEBA force field was to prevent artifacts at close range (the so-called ‘polarization catastrophe’), in the non-iterative case these artifacts should not exist, so \( \lambda \) should be interpreted as a fitted short-range correction to the polarization energy.

### 2.4 Water Models

**TIP\( n \)P:** The parameters for these models are summarized in Table 1. In all TIP\( n \)P models, the OH bond length, \( r_{OH} \), and HOH bond angle, \( \angle \text{HOH} \), are the gas-phase experimental values, \( i.e. \), 0.9572 Å and 104.52°, respectively. There is no charge at the O center in both the TIP4P and TIP5P models. The potential energy of the TIP4P and TIP5P models between two water molecules, \( a \) and \( b \), is given by Eq. 6, where \( i \) and \( j \) are the charged sites on molecules \( a \) and \( b \), respectively, and \( r_{OaOb} \) is the oxygen-oxygen distance.

\[
E_{ab} = 4\epsilon_{OO} \left[ \left( \frac{\sigma_{OO}}{r_{OaOb}} \right)^{12} - \left( \frac{\sigma_{OO}}{r_{OaOb}} \right)^{6} \right] + \sum_{ij} \frac{q_i q_j}{r_{ij}}. \tag{6}
\]

Eq. 6 is equally applicable for TIP3P water when including an additional L-J interaction term on hydrogen sites. As the DCM approach uses off-centered charges to describe multipole moments, no modification is necessary to implement the TIP\( n \)P models, which are equivalent to MDCM distributions with 1 or 2 charges per atom. In some respects the TIP\( n \)P models can be viewed as MDCMs with hand-fitted charge positions and magnitudes, as displayed in Figure 2.

**iAMOEBA:** Inexpensive AMOEBA\(^{31} \) was originally conceived as a computationally efficient (fewer parameters and non-iterative polarization) and robust alternative to the existing AMOEBA water model. The requisite Halgren buffered 14-7 potential\(^{41} \) and anharmonic bonded terms of the iAMOEBA model are available in OpenMM,\(^{28} \) which also contains basic “dummy atom” functionality for sites with zero mass that can be used to run sim-
Figure 2: Charge locations in TIP\(_n\)P (\(n = 3, 4, 5\)) water models (top), in an octahedral distributed charge model (DCM, bottom-left) and a 10-charge minimal distributed charge model (MDCM, bottom-right). In all TIP\(_n\)P models positive charges are located at the \(H\) centers. The corresponding negative charges are located either at the \(O\) atom (\(q_O\) in TIP3P), along an axis connecting \(O\) and the center of mass of \(H_2O\) (\(q_M\) in TIP4P), or at two points approximating the \(O\) lone pairs (\(q_L\) in TIP5P). In the octahedral DCM model there are 6 charge sites for the \(O\) atom and for each \(H\) atom, describing the multipole moments of the iAMOEBA model. In the MDCM model there are 4 charge sites per \(H\) atom and 2 charge sites for \(O\).

Simulations with (M)DCMs. iAMOEBA in OpenMM thus provides a suitable test case to compare performance of explicit atomic multipole moments versus distributed charges in polarizable, condensed phase simulations. It is also possible to demonstrate the ease with which (M)DCMs can be made available in existing MD software packages without the need to change the source code. Such an “emulation” should offer advantages in both computational efficiency and ease of implementation.
Table 1: Parameters for TIP$n$P and iAMOEBA potential functions.

|                        | TIP3P | TIP4P | TIP5P | iAMOEBA (M)DCM |
|------------------------|-------|-------|-------|-----------------|
| **Geometry**           |       |       |       |                 |
| $\angle$HOH$^\circ$    | 104.52| 104.52| 104.52| 106.48          |
| $\angle$LOL$^\circ$    |       |       |       |                 |
| $r_{OH}$/Å              | 0.9572| 0.9572| 0.9572| 0.9584          |
| $r_{OL}$/Å              |       |       |       |                 |
| $r_{OM}$/Å              |       |       | 0.70  |                 |
|                        |       |       | 0.15  |                 |
| **Charges**            |       |       |       |                 |
| $q_H$/e                | 0.417 | 0.52  | 0.241 |                 |
| $q_O$/e                | -0.834|       |       |                 |
| $q_L$/e                |       | -0.241|       |                 |
| $q_M$/e                |       | -1.04 |       |                 |
| **L-J parameters**     |       |       |       |                 |
| $\epsilon_{OO}$/kcal/mol| -0.1521| -0.155| -0.16  | 0.19682$^a$     |
| $\epsilon_{HH}$/kcal/mol| -0.0460|       |       |                 |
| $\sigma_{OO}$/Å         | 3.15061| 3.15365| 3.12  | 3.6453$^a$      |
| $\sigma_{HH}$/Å         | 0.4   |       |       |                 |
| **Polarizabilities**   |       |       |       |                 |
| $\alpha_O$/Å$^3$        |       | 0.80636|       |                 |
| $\alpha_H$/Å$^3$        |       | 0.50484|       |                 |
| $\alpha_{LM}$/Å$^{-1}$  |       | 0.23616$^b$|       |                 |

$^\dagger$ (M)DCM charge sites for H atoms, and $^\ddagger$ (M)DCM charge sites for O atoms – refer to section 2 of the SI.

$^a$ iAMOEBA L-J parameters for use with a Halgren 14-7 potential$^{41}$

$^b$ polarization damping factor

3. Computational Details

3.1 DCM Representations

TIP$n$P models were implemented for use with the DCM module in CHARMM by describing the positions of off-centered charges in the standard DCM local axis system.$^{25}$ For multipolar models, in-house scripts were used to convert from the original local axis system, as defined here by the AMOEBA force field for the iAMOEBA model and the multipole module of the
CHARMM force field for the multipolar PhF model, to the final DCM axis system. Also, the necessary diagonalization of the Cartesian quadrupole matrix to obtain a minimal number of non-zero quadrupole components and to calculate the corresponding charge magnitudes of the DCM arrangement were computed using these scripts. An octahedral charge arrangement was used to describe the multipole expansion for all models, as described previously. For the polarization term the standard iAMOEBA polarizabilities and damping parameters were used (Table 1). Interatomic distances $R_{ij}$ in Eq. 4 were the distance between the nuclei to which charges were assigned, rather than the distance between DCM charges, to most closely match the AMOEBA multipolar force field term and avoid refitting polarization damping parameters. A sample input file is provided in section 3 of the SI.

### 3.2 MDCM fitting

The Differential Evolution (DE) fitting code was implemented into the “Fitting Wizard” (FW) tool previously developed to fit multipole moments and L-J parameters to bulk properties for multipolar force fields. Charge positions were constrained to remain within one third of the van der Waals radius of each atom. Atomic multipoles up to rank $l = 5$ were fitted to the MEP across a grid using a least squares fit as before, where the grid used here is generated by the target multipolar model of interest rather than quantum chemical reference data. Grid points between the 0.001 a.u. and 0.0003 a.u. isodensity surfaces were used for fitting, as points outside the outer 0.0003 a.u. surface with lower electron density were found to be far enough away to be generally well described and have a smaller impact on fitting quality. Excluded points outside the outer surface were subsequently used to validate the long-range performance of the model via mean and maximum absolute errors.

Atomic charge models with up to 4 charges per atom were fitted to the atomic multipoles and used to generate initial populations for subsequent DE fitting of the larger systems, as
described elsewhere. MDCM was algorithmically improved by introducing an intermediate fragmentation strategy to increase computational efficiency for larger molecules. After fitting atomic charge models to the ESP of the atomic multipoles, the molecule is divided into fragments (here into 2 arbitrary fragments of roughly equal size for PhF). Next, each fragment is fitted separately to a reference fragment MEP that does not include the ESP contribution of the multipoles of the other fragment(s). This is achieved using the high-rank \( l = 5 \) fitted atomic multipoles already obtained to fit the atomic charge models:

\[
V^{\text{frag}}_{\text{ref}}(r) = V^{\text{mol}}_{\text{ref}}(r) - \sum_{i=1}^{N_{\text{frag,fix}}} \sum_{j=1}^{N_{\text{atom},i}} V^{\text{mtp}}_{i,j}(r)
\]  

(7)

where the fragment reference ESP \( V^{\text{frag}}_{\text{ref}} \) at point \( r \) is equal to the original reference MEP, \( V^{\text{mol}}_{\text{ref}}(r) \), minus the ESP \( V^{\text{mtp}}_{i,j}(r) \) from the fitted multipoles of all \( j \) atoms of all \( N_{\text{frag,fix}} \) fragments that are not included in the current fragment fit. As each fragment contains fewer charges than the full molecule, and all fragments can be fitted independently a considerable speedup of the fitting process is possible and the approach scales favourably for larger systems.

Fragments are fitted with increasing numbers of charges until a predefined/desired accuracy has been obtained. Here, an average of between 1 and 3 charges per atom are used and charges are free to move. Hence, the number of charges for each atom within the fragment may change during fitting while the total number of charges for the fragment remains fixed. 10 models were fitted for each given number of charges for each fragment. The fragment models with lowest mean absolute errors across the ESP grid were combined to build molecular charge models with the number of charges corresponding to between on average 1 and 3 charges per atom (i.e. between 12 and 36 charges for the full PhF molecule). Each molecular model was subjected to a final DE refinement step.
A further improvement to the original approach\textsuperscript{26} was to introduce additional constraints. Constraints are important both to maintain stability of MD simulations and to maintain accuracy of subsequent electrostatic interaction energy calculations. Simulation stability is maintained by constraining charge positions to remain within $r_{\text{atom}}/3$, one third of the van der Waals radius of the atom. If charges are placed too far from nuclear positions they are able to approach one another during MD simulations, overcoming repulsive barriers and causing simulations to collapse due to numerical instability. To improve the accuracy of electrostatic interaction energies, tighter constraints of maximally $1e$ for each point charge were applied to all charge magnitudes. The grounds for constraining charge magnitudes is based on analysis of the error in the interaction energy using MDCMs (described in section 4 of the SI), as larger charge magnitudes often reduce the error in the MEP at the expense of increasing the error in the electrostatic interaction energy in subsequent simulations through error multiplication.

### 3.3 (M)DCM in OpenMM

For iAMOEBA simulations, a single simulation engine (OpenMM 7.1.0\textsuperscript{43}) was used to keep all simulation details and force field terms unchanged apart from the modified electrostatics. As OpenMM lacks native DCM support, the existing “dummy atom” functionality was exploited to run (M)DCM simulations, highlighting the possibility to run (M)DCM simulations in simulation packages that support dummy atoms or equivalent features. (M)DCM charges were placed relative to atoms by converting from local DCM axes to those defined in OpenMM for dummy atoms (Figure 1).

For consistency, the polarization term in OpenMM had to be adapted for use with distributed charges. The polarization energy damping term implemented in OpenMM assumes that charged and polarizable sites will coincide, as the distance $R_{il}$ of the damping term in Eq. 4
is evaluated between nuclear sites. This is not the case in (M)DCM, where charge sites are
shifted from nuclear positions. If every (M)DCM charge site were assigned the polarizability
of the parent atom, the total polarization energy would be significantly overestimated, in
accordance with Eq. 3. Hence, \( \alpha'_i = \alpha_i \cdot 10^{-4} \) was used for charge positions and the damping
factor \( \lambda_{ij} \) was changed to

\[
\lambda_{ij} = 1 - \exp \left( -\frac{a'}{100} \left( \frac{r_{il}}{(\alpha'_i \alpha'_j)^{1/6}} \right)^3 \right)
\]

with \( a' = 0.0023616 \text{ Å}^{-1} \) and \( \alpha'_i = \alpha_i \) for the nuclei. As all nuclei carry zero charge in
(M)DCMs the nuclear–nuclear interactions yield zero polarization energy, in accordance with
Eq. 4. As all charges carry polarizabilities scaled by a factor \( 10^{-4} \), polarization energies be-
tween (M)DCM charges are negligible which is consistent with Eq. 3. For charge–nuclear
site interactions, the nuclear site carries the standard polarizability, so Eq. 3 is unchanged,
and the charge site polarizability is scaled by \( 10^{-4} \), which is counteracted by the factor \( 10^{-2} \)
applied to the damping parameter \( a \) in Eq. 5, recovering the polarization and damping of
the multipolar force field without refitting any parameters. Note that the small remaining
difference in the \( R_{il} \) term from using shifted charge sites was found to not significantly affect
the results, but should be considered a potential source from which slight differences can arise.

Finally, fitted MDCMs were converted to the standard local axis systems used for dummy
atoms in OpenMM. A sample parameter file is provided in section 5 of the SI. It was noted
when running these simulations that although (M)DCM simulations ran relatively efficiently
in OpenMM, and for MDCM in particular multi-ns simulations ran on a similar timescale to
the standard multipolar iAMOEBA implementation, code did not appear to be optimized for
simulations with multiple off-centered charge sites. Significant slow-down was observed even
if all off-centered sites were defined with zeroed non-bonded parameters. Benchmarking re-
results are therefore not meaningful here, and results with optimized code will be addressed in
future work. It is clear, however, that simulation efficiency will benefit from purpose-written code, and that the functionality required for off-centered charge sites is unlikely to be as highly optimized as more commonly encountered atom-centered terms in current simulation packages.

3.4 MD Simulations and Property Computation

Aside from iAMOEBA simulations, which were run with OpenMM as described below, remaining MD simulations were run with CHARMM\textsuperscript{27} version 45a2 which includes provisions for DCM.\textsuperscript{25} A 1 fs time step was used with SHAKE\textsuperscript{44} to constrain angles and bonds involving hydrogen atoms in an isothermal-isobaric (constant $NPT$) ensemble using a pressure bath at 1 atm coupled to a Nosé-Hoover temperature bath.\textsuperscript{45–49} The simulation system was a cubic box with 500 water molecules employing periodic boundary conditions. For every value of $T$, a simulation of at least 3 nanoseconds (ns) was performed. For TIP5P water and temperatures below freezing point (273 K) the simulations were extended by an additional 3 ns for improved estimates of thermodynamic properties using fluctuation formulae (see below). This strategy has been suggested previously to obtain converged results for modeling bulk water at low temperatures.\textsuperscript{50–53} All simulations were performed with SHIFT and SWITCH cutoff functions for non-bonded electrostatics and van der Waals interactions, respectively. The switching-function parameters are $R_{\text{on}}$ and $R_{\text{off}}$ with values 10.0 and 12.0 Å, respectively, for non-bonded van der Waals interactions. A 12.0 Å cutoff was applied for the shifted non-bonded electrostatics. The implementation of DCM was validated by comparing results of TIP4P and TIP5P simulations with the results of similar simulations with standard routines.\textsuperscript{54}

OpenMM simulations of a cubic box with 500 water molecules were run without SHAKE constraints for compatibility with iAMOEBA, and with a 0.5 fs time step. A Monte Carlo
barostat maintained simulation pressure at 1 atm, after 150 ps equilibration simulations were run for 10 ns at each $T$ used for the CHARMM simulations to facilitate direct comparison between CHARMM and OpenMM data. Particle Mesh Ewald (PME) was used with a real-space cutoff of 7.0 Å and a van der Waals cutoff of 9.0 Å.

**Solvation Free Energies**, $\Delta G$, were calculated using slow-growth thermodynamic integration$^{55,56}$ (TI) for discrete windows of $\lambda$ from 0 to 1, with fixed $\Delta \lambda = 0.1$ and $\Delta t = 1$ fs. The simulation for each value of $\lambda$ (window) was equilibrated for 150 ps, and then sampled for another 200 ps, i.e., cumulatively $11 \times 350$ ps for each water model. Here, $\lambda = 0$ is the state in which the water molecule has full interaction with the system, $\lambda = 1$ is the state in which the water molecule is “decoupled” and in the gas phase. The free energy change for each simulation was calculated by varying $\lambda$ forward from 0 to 1 and then backward from 1 to 0.

**Bulk-density** $\rho$ was computed from the ratio between total mass, $M$, and the time-averaged volume of the simulation box, $< V >$ according to $\rho = \frac{M}{< V >}$.

**Self-diffusion coefficient** $D$ was computed from the mean squared displacement (MSD) of all oxygen atoms using the Einstein relation

$$D = \lim_{t \to \infty} \frac{1}{6t} < |r(t) - r(0)|^2 >,$$  

where $r(t)$ is the position of the oxygen atom of a water molecule at time $t$, and averaged over all water molecules.$^{57}$ OpenMM trajectories were analyzed in CHARMM and $D$ was computed in the same way.
**Enthalpy of vaporization** $\Delta H_{\text{vap}}$ can be obtained from

$$
\Delta H_{\text{vap}} = <E_{\text{gas}} > - <E_{\text{liq}}>/N + RT, \quad (10)
$$

where $E_{\text{liq}}$ is the potential energy of the liquid containing $N$ molecules and $R$ is the ideal gas constant. \(^{50,51,58}\)

**Heat capacity** $C_p$, **isothermal compressibility** $\kappa$, **coefficient of thermal expansion** $\alpha$ can be calculated from standard fluctuation formulae (Eqs. 11 to 13). \(^{50,51,58}\)

$$
C_p = \left( \frac{\partial H}{\partial T} \right)_{N,P} = \frac{1}{Nk_B T^2} (<H^2> - <H>^2) + 3R \quad (11)
$$

$$
\kappa = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{N,T} = \frac{1}{k_B T} <V>(<V^2> - <V>^2) \quad (12)
$$

$$
\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_{N,P} = \frac{1}{k_B T^2} <V>(<VH> - <V><H>) \quad (13)
$$

Here, $C_p$ and $\alpha$ were computed using the central difference formula for estimating derivatives, except at extremes where right and left differences were used, \(^{59}\)

$$
C_p \approx \frac{<H_2> - <H_1>}{T_2 - T_1}, \quad \text{and} \quad \alpha \approx \frac{\ln <\rho_2> - \ln <\rho_1>}{T_2 - T_1} \quad (14)
$$

and $\kappa$ values were calculated from the fluctuations (Eq. 12).
4. Results and Discussion

This section is structured as follows. First, the performance of the (M)DCM iAMOEBA model is evaluated for single-point energy calculations of water clusters, in order to validate the (M)DCM models against reference data from the original multipolar iAMOEBA implementation. Implementations for the TIP$n$P models are more straightforward and hence not shown here. Then the performance of all models for different bulk properties between $T = 235.5$ K and $T = 350$ K is assessed to explore the performance of the DCM framework in describing models of increasing complexity under realistic simulation conditions. For iAMOEBA the impact on simulations of replacing the multipole term with distributed charges is the main focus. Finally, the methods developed are further used for MD and free energy simulations of hydrated PhF.

4.1 MDCM Fitting for H$_2$O

Models with increasing numbers of charges ranging from 6 to 10 per molecule were fitted to the MEP, with 10 independent models fitted for each number of charges. The models that performed best as quantified by mean absolute error in the MEP were then used in dimer energy calculations (next section). As no refitting of other force field parameters was desirable a tight agreement of $\sim 0.2$ kcal/mol (1 kJ/mol) in dimer energies was the threshold to select the MDCM to be used in subsequent calculations. It was found that 10 charges were required to reach this threshold, more than might typically be necessary for such a model. However, this is justified by the challenging nature of the task, and still compares favorably with the 12 non-zero multipole components of the iAMOEBA model and without the additional complexity that these terms will incur in simulations. In a more typical application an error in dimer electrostatic interaction energies of 1 kcal/mol with respect to $ab$ initio reference data might be acceptable, requiring fewer charges, as errors are typically
compensated by remaining force field terms. The MAE across the grid between the two isodensity surfaces for the selected model was 0.023 kcal/mol and the maximum absolute error was 2.10 kcal/mol.

The selected model is shown at the bottom-right of Figure 2. No symmetry constraints were applied, and while such constraints could be introduced as an extension to the current fitting code for symmetric systems, the asymmetric charge distributions obtained during MDCM fitting still accurately describe the symmetry of the underlying MEP if the fitting criteria are sufficiently tight.

Figure 3: Top: comparison of MEP generated by 10-charge MDCM (left) with MEP generated by iAMOEBA multipole moments (right), mapped onto the molecular 0.001 a.u. isodensity surface. Colors range from -38 (red) to +38 (blue) kcal/mol. Bottom: 2D slices of the MEP in different molecular planes (rows) in the region used for fitting. The “True” ESP refers to the iAMOEBA multipolar reference, the fitted ESP refers to the MDCM 10-charge model. The absolute error is plotted in the right-hand column and ranges from -1 kcal/mol (red) to +1 kcal/mol (blue).
4.2 iAMOEBA and its (M)DCM Representation

The quality of the octahedral DCM and 10-charge MDCM descriptions of the iAMOEBA multipolar electrostatics were examined using a series of water dimers, originally proposed by Tschumper et al.\textsuperscript{60} (Figure 4) and a set of larger water clusters up to decamer\textsuperscript{61} to check for error accumulation with cluster size.

![Figure 4: Water dimers from Ref.\textsuperscript{60} used to validate (M)DCM iAMOEBA parameters.](image1)

![Figure 5: Water clusters used to check for error accumulation in (M)DCM iAMOEBA model.](image2)
Table 2: Comparison between multipolar and DCM iAMOEBA electrostatic energies, including polarization. In the top part of the table, energies for the 10 dimer structures (see Figure 4) are shown, while in the bottom part energies for various oligomers (see Figure 5) are presented. The mean absolute error for the dataset is within the chemical accuracy (MAE=0.041 kcal/mol)

| Dimer | iAMOEBA (kcal/mol) | DCM-iAMOEBA (kcal/mol) | ΔE (kcal/mol) |
|-------|--------------------|------------------------|---------------|
| 1     | -5.113             | -5.092                 | 0.021         |
| 2     | -4.505             | -4.532                 | -0.027        |
| 3     | -4.504             | -4.554                 | -0.049        |
| 4     | -3.824             | -3.769                 | 0.056         |
| 5     | -3.269             | -3.217                 | 0.053         |
| 6     | -2.972             | -2.967                 | 0.005         |
| 7     | -3.199             | -3.173                 | 0.027         |
| 8     | -1.572             | -1.579                 | -0.007        |
| 9     | -3.794             | -3.778                 | 0.016         |
| 10    | -3.016             | -3.027                 | -0.011        |
| Oligomer | iAMOEBA (kcal/mol) | DCM-iAMOEBA (kcal/mol) | ΔE (kcal/mol) |
| trimer | -13.770            | -13.691                | 0.079         |
| tetramer | -24.529           | -24.495                | 0.034         |
| pentamer | -32.321           | -32.279                | 0.067         |
| hexamer prism | -41.459   | -41.421                | 0.038         |
| heptamer | -51.299           | -51.193                | 0.106         |
| octamer | -64.672           | -64.495                | 0.178         |
| nonamer | -72.896           | -72.989                | -0.092        |
| decamer | -82.858           | -82.772                | 0.086         |
Table 3: Comparison between multipolar and MDCM iAMOEBA electrostatic energies, including polarization. In the top part of the table, energies for the 10 dimer structures are shown, while in the bottom part energies for various oligomers are presented. The mean absolute error for the dataset is within chemical accuracy (MAE=0.108 kcal/mol)

| Dimer  | iAMOEBA (kcal/mol) | MDCM-iAMOEBA (kcal/mol) | ΔE (kcal/mol) |
|--------|--------------------|--------------------------|---------------|
| 1      | -5.113             | -5.125                   | -0.011        |
| 2      | -4.505             | -4.465                   | 0.040         |
| 3      | -4.504             | -4.471                   | 0.033         |
| 4      | -3.824             | -3.836                   | -0.011        |
| 5      | -3.269             | -3.180                   | 0.089         |
| 6      | -2.972             | -2.863                   | 0.109         |
| 7      | -3.199             | -3.196                   | 0.003         |
| 8      | -1.572             | -1.587                   | -0.014        |
| 9      | -3.794             | -3.770                   | 0.024         |
| 10     | -3.016             | -3.026                   | -0.010        |

| Oligomer | iAMOEBA (kcal/mol) | MDCM-iAMOEBA (kcal/mol) | ΔE (kcal/mol) |
|----------|--------------------|--------------------------|---------------|
| trimer   | -13.770            | -13.782                  | -0.012        |
| tetramer | -24.529            | -24.436                  | 0.093         |
| pentamer | -32.321            | -32.266                  | 0.055         |
| hexamer prism | -41.459 | -41.348 | 0.111 |
| heptamer | -51.299            | -51.129                  | 0.171         |
| octamer  | -64.672            | -64.238                  | 0.434         |
| nonamer  | -72.896            | -72.478                  | 0.419         |
| decamer  | -82.858            | -82.402                  | 0.456         |
As seen in Table 2, very close agreement was obtained between the original multipolar iAMOEBA implementation in OpenMM and its DCM representation developed here. The difference in electrostatic interaction energy (including polarization) is of the order of $10^{-2}$ kcal/mol for all dimers. Close agreement was also obtained for the water clusters, with error accumulation remaining remarkably small up to decamer, and the largest total error of 0.18 kcal/mol still well within chemical accuracy.

Table 3 reveals a similar trend for the fitted MDCM. While dimer errors are slightly larger than for the DCM, the largest error of 0.109 kcal/mol is still very close to the iAMOEBA multipolar energy. For the larger clusters results are again encouraging, although some error accumulation is visible as errors increase to almost 0.5 kcal/mol for the decamer. It should be emphasized that this remains well within chemical accuracy even for these larger clusters, with a percentage error of 0.5% for the decamer and MAE of 0.11 kcal/mol for the whole set of dimers and clusters, despite requiring little more than half the number of charges used in the octahedral DCM.

### 4.3 H$_2$O Thermodynamic Properties at 298 K from all Models

After establishing that both the octahedral DCM and 10-charge MDCM yield accurate interaction energies, the performance of the models for bulk properties in condensed phase MD simulations was assessed. (M)DCM representations for models ranging from TIP3P to iAMOEBA were chosen to demonstrate the versatility of the approach. In all cases, simulations are carried out for both the original model and its (M)DCM representation.

Atom-atom pair correlation functions (radial distribution functions, RDFs) $g_{OO}(r)$, $g_{OH}(r)$ and $g_{HH}(r)$ of water characterize the microscopic structure of liquid water. One of the critical tests for water models is accurate reproduction of the experimental (X-ray scattering and
neutron diffraction) OO, OH and HH RDFs, although it should be noted that there remains some uncertainty over reference values in the literature. For example, the reported height ($g_1$) and position ($r_1$) of the first intermolecular peak in $g_{OO}(r)$, which is the characteristic feature of liquid water, varies over the range $g_1 = 2.2–3.0$ and $r_1 = 2.76–2.82$ Å depending on the type of experiment.\textsuperscript{62–71}

Figure 6 shows $g_{OO}(r)$ for the various water models along with experimental neutron diffraction data.\textsuperscript{62} A more detailed comparison of $g_{OO}(r)$ along with $g_{OH}(r)$ and $g_{HH}(r)$ is presented in sections 6–8 of the SI. One of the most commonly used models for biomolecular simulations, TIP3P, is known to inadequately describe the water structure beyond the first peak of $g_{OO}(r)$, which is confirmed here beyond $\sim 3$ Å, see Figure 6. For TIP4P, TIP5P and iAMOEBA improved $g_{OO}(r)$ are found across the whole range of distances. Of particular interest here is the close agreement firstly between TIP$n$P results and their DCM equivalents. This demonstrates that the integration of (M)DCMs with commonly used simulation tools such as barostats and electrostatic cut-offs yields dynamics in condensed phase simulations in line with more widely used approaches for these electrostatic models. Secondly, there is very good agreement for the polarizable, multipolar iAMOEBA results with (M)DCM, demonstrating that the close agreement in MEP and interaction energies yields correspondingly close $g(r)$ in simulations without the need for explicit multipoles, and despite differences in implementation such as the definition of local axis systems and alterations to non-bonded cut-off schemes to accommodate off-centered charges.

Liquid density, heat of vaporization, isobaric heat capacity, isothermal compressibility, thermal expansion coefficient and self-diffusion coefficient at 298 K and 1 atm are summarized in Table 4. Again results for TIP$n$P models are compared with their DCM-representations and those for multipolar iAMOEBA are compared with those from (M)DCM and with experiment.
The calculated \( \rho \), \( \Delta H_{\text{vap}} \) and \( D \) of TIPnP waters are essentially identical from both sets of simulations, further validating the DCM implementation. \( C_p \), \( \kappa \) and \( \alpha \) values of water from the various models vary within \( \pm 1 \) unit, although for TIP4P \( \alpha \) is somewhat overestimated, by about 30 %. As shown in the next section, this discrepancy seems to originate from noise in the data rather than a physical effect, as performance across a range of temperatures shows closer agreement and the other models also show local discrepancies at certain temperatures (Figure 7b). Agreement for the simulations with iAMOEBA and its (M)DCM representations is also very good, with the largest discrepancy being the slightly larger value for \( \kappa \) with the DCM representation.

The range of dynamic properties examined provides confidence that the microscopic structure and dynamics of the solvent are preserved when moving from a full multipolar description to the simpler distributed charge models, also after reducing the number of charges in the MDCM. All iAMOEBA results additionally agree well with experiment, consistent with earlier findings.\(^{31}\) It should also be noted that the data presented are for the simulation conditions described above, and hence certain deviations from previously published results are to be expected. For example \( \Delta H_{\text{vap}} \) is 6 % larger than the originally published Monte Carlo (MC) data,\(^{30}\) \( \kappa \) values are roughly 30 \( (10^6 \ \text{atm}^{-1}) \) smaller for TIP3P and TIP4P compared to original MC data,\(^{30,50}\) and \( D \) for both TIP3P and TIP4P is roughly 40 % larger than previously published results.\(^{72}\) As has been highlighted previously,\(^{57}\) small differences in simulation conditions can have a significant impact on the results from bulk simulations.

### 4.4 Temperature Dependence of H\(_2\)O Thermodynamic Properties

As a more exacting test of the various water models and their (M)DCM representations, the density \( \rho \), enthalpy of vaporization \( \Delta H_{\text{vap}} \), isobaric heat capacity \( C_p \), isothermal compress-
Table 4: Bulk properties of liquid water at 298 K and 1 atm; density $\rho$ (g cm$^{-3}$), enthalpy of vaporization $\Delta H_{vap}$ (kcal/mol), isobaric heat capacity $C_p$ (cal mol$^{-1}$K$^{-1}$), isothermal compressibility $\kappa$ (10$^6$ atm$^{-1}$), thermal expansion coefficient $\alpha$ (10$^5$ K$^{-1}$) and self-diffusion coefficient $D$ (10$^{-5}$ cm$^2$ s$^{-1}$). “DCM” denotes the new DCM code and framework were used in place of standard routines.

| Model               | $\rho$ | $\Delta H_{vap}$ | $C_p$ | $\kappa$ | $\alpha$ | $D$  |
|---------------------|--------|------------------|-------|----------|----------|------|
| TIP3P$^c$           | 1.0266 | 11.04            | 12.9  | 23.1     | 76.8     | 3.9  |
| TIP3P/DCM           | 1.0264 | 11.04            | 12.9  | 22.4     | 75.6     | 3.9  |
| TIP4P               | 1.0090 | 11.15            | 15.6  | 22.7     | 35.3     | 2.3  |
| TIP4P/DCM           | 1.0082 | 11.16            | 15.5  | 23.3     | 44.9     | 2.3  |
| TIP5P               | 0.9848 | 10.73            | 24.2  | 28.8     | 30.9     | 2.2  |
| TIP5P/DCM           | 0.9842 | 10.73            | 24.8  | 30.0     | 30.0     | 2.2  |
| iAMOEBA/OpenMM      | 0.9977 | 10.91            | 17.8  | 40.7     | 23.3     | 1.9  |
| iAMOEBA/DCM         | 0.9916 | 10.85            | 18.3  | 51.0     | 25.8     | 2.0  |
| iAMOEBA/MDCM10      | 0.9964 | 10.76            | 17.9  | 41.5     | 28.9     | 2.2  |
| Exp.$^a$            | 0.9965 | 10.51            | 18.0  | 45.8     | 25.7     | 2.3$^b$|

$^a$ Ref.;$^{73}$ $^b$ Ref.;$^{74}$ $^c$ using conventional TIP3P in CHARMM;

iability $\kappa$, and thermal expansion coefficient $\alpha$ were studied as a function of temperature $T$ between 235.5 K and 350 K. Corresponding property vs. $T$ profiles are presented alongside reference experimental data in Figure 7. In all cases the TIP$n$P models and their DCM representations agree very favourably. The same applies to iAMOEBA with the exception of $\kappa$ using the DCM representation for which a small shift is visible.

**Liquid density:** The dependence of bulk density $\rho$ on $T$ is shown in Figure 7a. In all cases there is close agreement between the original models and their (M)DCM representations, and much closer than the agreement between the different models. The additional detail in the iAMOEBA force field yields a visibly closer agreement with experiment than the simpler non-polarizable TIP$n$P models, with both the multipolar and MDCM descriptions almost indistinguishable from the experimental curve, while the DCM description is also very close.

**Isothermal compressibility and thermal expansion coefficient:** Plots of isothermal
Figure 6: Radial distribution functions (RDF) $g_{OO}(r)$ of iAMOEBA and TIP$n$P H$_2$O at 298 K and 1 atm using standard routines (LP) and DCM functionality in CHARMM (TIP$n$P), and multipolar (MTP) iAMOEBA and (M)DCM implementations of iAMOEBA in OpenMM. Experimental neutron diffraction data from Ref. $^{62}$ Successive curves are offset 1 unit along the y-axis for clarity.

Compressibility $\kappa(T)$, and thermal expansion coefficient $\alpha(T)$ as a function of temperature are also included in Figures 7b and c. (M)DCM representations of TIP$n$P again agree well with reference data, and more closely than the models agree with one another. The same is true for iAMOEBA, although some shift is visible in $\kappa$ for the DCM representation.

Relative to the experimental results, performance is again varied for the TIP$n$P models with no single model performing well across all properties, suggesting that they lack the versatility to describe all properties simultaneously across the full range of $T$. The additional detail in the iAMOEBA model, well encapsulated using distributed charges, affords a more consistent
Figure 7: Experimental and computed densities $\rho$, diffusion coefficients $D$, isothermal compressibilities $\kappa$, enthalpies of vaporization $\Delta H_{\text{vap}}$, thermal expansion coefficients $\alpha$ and heat capacities $C_p$ of liquid water using TIP$n$P and iAMOEBA water models as a function of temperature at 1 atm. Experimental $\rho$ data are obscured by the iAMOEBA/MTP (multipolar iAMOEBA) data. A vertical line at 273 K indicates zero Celsius.

performance across the full range of $T$.

Diffusivity: The self-diffusion coefficient is one of the most frequently examined transport properties of water in MD studies. It measures the mobility of water molecules in the H-bonded liquid water network, and is thus taken as an indication of the accuracy of the
water interaction potential. The simulated $D$, over the range of temperatures studied here, is presented in Figure 7d. While TIP3P overestimates the diffusivity, suggesting that hydrogen bonding is too weak in this water model (consistent with loss of structure beyond the first coordination shell, seen in Figure 6), the remaining models describe this property well, although with some level of divergence at higher $T$.

**Enthalpy of vaporization:** The variation of vaporization enthalpy $\Delta H_{\text{vap}}$ with $T$ (Figure 7e) shows significant differences between the different models, but in each case there is good agreement between the DCM representation and the existing models, and between the multipolar, polarizable approach and (M)DCM results. Interestingly, TIP3P performs better here than the other TIP$n$P models across the full range of temperatures, with TIP5P tuned to perform better under ambient conditions. iAMOEBA also performs well across the range of temperatures, outperforming TIP3P at higher $T$ but overestimating slightly the gradient with respect to $T$.

**Isobaric heat capacity:** For $C_p(T)$, again iAMOEBA and its (M)DCM representations agree well with one another across a broad range of temperatures. The TIP3P and TIP4P DCM models also agree well with existing implementations, whereas slightly more deviation is visible for the DCM implementation of the TIP5P model. This model represents a significant outlier, though, in its agreement with experiment and the remaining models other than at high $T$, which may lead to increased sensitivity in this property.

### 4.5 Free Energies of H$_2$O Self-Solvation

The calculated TIP$n$P free energy changes, $\Delta G$, of decoupling one water molecule from bulk liquid using different water models (see Table 5) in the forward and backward calculations are 0.01 to 0.20 kcal/mol to one another, indicative of good convergence in the simulations. The various TIP$n$P/DCM simulations give nearly the same $\Delta G$ values as standard TIP$n$P,
validating the DCM implementation for thermodynamic integration. The free energy change associated with removal of a water molecule from bulk water to the gas phase ranges from 5.82 to 6.94 kcal/mol (Table 5), close to the experimentally measured value of 6.3 kcal/mol\textsuperscript{78} as well as previously calculated values of 6.0\textsuperscript{79} and 6.1\textsuperscript{80} kcal/mol. These simulations were not carried out for iAMOEBA as the different TI approaches in CHARMM and OpenMM hinder meaningful comparison.

Table 5: Calculated free energy $\Delta G$ (kcal/mol) change (sum of the electrostatic component and the van der Waals interaction component) for removal of a water molecule from bulk water to gas-phase using TIP$n$/DCM and existing TIP$n$P implementations.

|                      | Forward $\Delta G$ | Backward $\Delta G$ |
|----------------------|-------------------|---------------------|
|                      | $E_{\text{elec}}$ | $E_{\text{vdW}}$ | Total |
| Exp.\textsuperscript{a} | 6.30              |                     |       |
| TIP3P\textsuperscript{b} | 8.40              | -2.40              | 6.00  |
| TIP3P/DCM            | 8.95              | -2.01              | 6.94  |
| TIP3P                | 8.99              | -2.01              | 6.98  |
| TIP4P/DCM            | 8.94              | -2.46              | 6.48  |
| TIP4P                | 9.05              | -2.46              | 6.59  |
| TIP5P/DCM            | 8.14              | -2.25              | 5.89  |
| TIP5P                | 8.07              | -2.25              | 5.82  |

\textsuperscript{a} Ref.; \textsuperscript{b} Ref.\textsuperscript{81}

### 4.6 MDCM Model for PhF

Next, the (M)DCM parametrization is extended to solvated systems. As an example, fluorobenzene (PhF) was chosen. In the past it has been demonstrated that for halogenated benzenes including detailed electrostatics is mandatory for quantitative simulations.\textsuperscript{82–84}

The presence of a halogen atom in PhF with a weak “sigma hole”, combined with the availability of an existing multipolar model\textsuperscript{82} make it another suitable choice to evaluate the impact of replacing multipolar terms with distributed charges. In this case previously published experimental\textsuperscript{85–87} and computed\textsuperscript{82} solvation enthalpies, vaporization enthalpies of the
pure liquid and heat capacities were available for comparison.

As for water, the first step was to obtain suitable MDCMs fitted to a grid of MEP points generated by the pre-existing multipolar model. Models were fitted with between 12 and 36 charges, i.e. with an average of between 1 and 3 charges per atom, see Figure 8. As the underlying multipolar model contained 42 non-zero multipolar terms, all charge models again offer a notable decrease in computational complexity for subsequent simulations. While the visible noise in Figure 8 with increasing number of charges in the fit shows that further refinement of the fitting procedure is possible, for example by increasing the number of DE fitting generations or the number of fits performed, the generally systematic improvement is encouraging and offers the possibility to select an MDCM based on an optimal compromise between computational cost of simulations due to increased number of charges, and improved accuracy in the electrostatics.

It is also encouraging to see that with 13 charges, i.e. an average of 1.1 charges per atom, the RMSE has already dropped to 0.19 kcal/mol with a maximum absolute error across the grid of 1.14 kcal/mol. With 18 charges the RMSE is 0.08 kcal/mol and the maximum absolute error is 0.38 kcal/mol and with an average of 3 charges per atom the accuracy reaches an RMSE of 0.02 kcal/mol and the maximum absolute error is 0.12 kcal/mol.

4.7 Thermodynamic Properties of PhF

After fitting MDCMs for PhF, the performance of the models in condensed phase simulations with respect to the performance of the original multipolar model was investigated. For this, the 22-charge MDCM was selected as it offered good accuracy (an RMSE of 0.06 kcal/mol across the electrostatic grid used for fitting and a maximum absolute error of 0.53 kcal/mol) at modest computational cost with fewer than 2 charges on average per atom (Figure 8).
Figure 8: Left: RMSE and maximum absolute errors for fits of MDCMs to MEP reference data as a function of increasing number of charges. Right: 22-charge model chosen for condensed phase simulations.

Again, as no symmetry constraints were applied the charge positions are spatially asymmetric, but the close agreement in MEP with the symmetric underlying multipolar model demonstrates that the symmetry of the MEP is conserved.

The thermodynamic properties chosen for comparison with available experimental data were the density of pure liquid PhF, the vaporization enthalpy of liquid PhF and the solvation free energy of a single PhF molecule in liquid water. As simulation conditions differed slightly from those used to fit the multipolar model\(^{82}\) a scaling factor of 1.1 was applied to the Lennard-Jones ‘\(\sigma\)’ and ‘\(\epsilon\)’ parameters of heavy atoms to recover the original performance of the multipolar model in describing these properties. The same scaling factor was applied to L-J parameters in MDCM simulations to allow direct comparison.

Results averaged over 8 independent simulations are shown in Table 6. While both models agree quite well with experiment, there is again very good agreement between the multipolar and MDCM condensed-phase simulation results. That this agreement is again possible without refitting any other force-field terms and by fitting only to MEP grid-data highlights the equivalent performance of a modest number of distributed charges to a full multipole-
expansion truncated at quadrupole when describing the dynamics of these systems.

Table 6: Calculated free energy $\Delta G$ (kcal/mol) of solvation of a single molecule of PhF in a TIP3P solvent box (averaged over 8 thermodynamic integration runs, with contributions from electrostatic and vdW terms), $\Delta H$ of vaporization (kcal/mol) and density at 300 K of pure liquid PhF. Results using an existing multipolar model for PhF are compared with results from a fitted MDCM with 22 charges. The standard deviation for each computed value is reported next to the number.

|                | $\rho$  | $\Delta H_{\text{vap}}$ | $\Delta G_{\text{solv}}^{\text{vdw}}$ | $\Delta G_{\text{solv,elec}}$ | $\Delta G_{\text{solv,tot}}$ |
|----------------|---------|------------------------|-------------------------------|-----------------|-------------------|
| Multipolar$^a$ | 0.90±0.01 | 9.37±0.05              | 2.21±0.13                       | -2.32 ± 0.04    | -0.10 ± 0.14      |
| MDCM           | 0.90±0.01 | 9.39±0.06              | 2.21±0.13                       | -2.92 ± 0.13    | -0.71 ± 0.21      |
| Exp.           | 1.02$^b$ | 8.26$^c$               |                               | -               | -0.80$^d$         |

$^a$ Ref. 82; $^b$ Ref. 85; $^c$ Ref. 86; $^d$ Ref. 87.

† The same 8 trajectories were used to evaluate Multipolar and MDCM vdW solvation energy contributions.

5. Conclusions

(M)DCM has been improved in accuracy, generalized and extended to include polarizability and applied to condensed phase simulations. Its performance to capture multipolar interactions within a single point charge-based implementation has been demonstrated by comparing various condensed phase properties of water and solvated PhF with established multipolar representations. Specifically, the distributed charge framework (including distribution of torques, local axis system description, definition of cut-offs) maintains the original bulk simulation properties for traditional force field models such as TIP3P, models with one or more off-centered charges such as TIP4P and TIP5P, and multipolar models such as the recently developed iAMOEBA and a multipolar description of PhF. Although there are different possible ways to define axis systems (Figure 1) and to distribute the torques, the close agreement and consistency of the results from the MD simulations for eight condensed phase properties using the different models indicates that these choices are not critical to
describing the dynamics of the system. This is not self-evident and a gratifying aspect of the present work.

The use of point charges in place of atom-centered multipoles removes the need for computationally expensive higher-order multipole terms to yield accurate and efficient dynamics. Furthermore, machine-learning reduces the number of charges necessary to a minimal set and the method has been combined with isotropic polarization and integrated with familiar tools such as barostats and thermodynamic integration routines to enable straightforward use of the models under widely encountered condensed-phase simulation conditions. The compatibility of the approach with an unmodified release of OpenMM and with all of the force field models tested here shows that distributed charges are a versatile tool for the development of next-generation force fields and multi-level approaches, and also offer an alternative to explicit multipolar terms of existing force fields. Beyond the simulation packages presented already, widely used codes such as Amber\(^1\) also offer functionality for off-centered charges that could be combined with (M)DCM without further modification.

While the current work focused on directly comparing simulation results from (M)DCM with multipolar representations \textit{without refitting} remaining parameters for direct comparison of the implementations, a generic (M)DCM force field, fitted to condensed phase experiments offers additional opportunities for quantitative simulations. This can be envisaged within the framework of an existing fitting environment\(^4\) and will be of particular interest when balancing accuracy (i.e. the number of charges per atom) and speed for specific applications. As was shown here it is possible by increasing the number of charges used in fitting to systematically improve the accuracy of the electrostatic interactions, allowing fine tuning of the computational cost of adding more charges and a corresponding improvement in the description of the electrostatic interaction. In this way, the present work opens up the possibility for custom-made MDCM force fields with calibrated accuracy to encapsulate the
physics of a given application.

6. Supplementary Material

See supplementary material for an illustration of the complexity of multipolar versus distributed charge interaction terms, for parameter files with (M)DCM charge positions and magnitudes, for sample input files used to run (M)DCM in OpenMM, for radial distribution functions from different water model simulations and for a discussion of the relationship between errors in the electrostatic interaction energy and fitting errors in the MEP. The code used to fit the MDCMs presented here is available freely at https://github.com/MMunibas/MDCM.

ACKNOWLEDGEMENT

The authors acknowledge financial support from the Swiss National Foundation for Research for their support within the NCCR MUST program and project 200021-117810.

REFERENCES

(1) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* 2005, 26, 1668–1688.

(2) Jorgensen, W. L.; Tirado-Rives, J. Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. *J. Comput. Chem.* 2005, 26, 1689–1700.

(3) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 2005, 26, 1701–1718.
(4) Christen, M.; Hünenberger, P. H.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; Gunsteren, W. F. V. The GROMOS software for biomolecular simulation: GROMOS05. *J. Comput. Chem.* 2005, 26, 1719–1751.

(5) MacKerell, Jr., A. D.; Brooks, B.; Brooks, III, C. L.; Roux, B. N.; Won, Y.; Karplus, M. In *Encyclopedia of Computational Chemistry*, 1st ed.; Schleyer, P. v. R., Ed.; John Wiley and Sons, Ltd.: New York, NY, 1998; pp 271–277.

(6) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, U.K., 2000.

(7) Plattner, N.; Meuwly, M. The Role of Higher CO-Multipole Moments in Understanding the Dynamics of Photodissociated Carbonmonoxide in Myoglobin. *Biophys. J.* 2008, 94, 2505–2515.

(8) Cardamone, S.; Hughes, T. J.; Popelier, P. L. Multipolar electrostatics. *Phys. Chem. Chem. Phys.* 2014, 16, 10367–10387.

(9) Ren, P.; Ponder, J. W. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* 2003, 107, 5933–5947.

(10) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio Jr, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* 2010, 114, 2549–2564.

(11) Ren, P.; Wu, C.; Ponder, J. W. Polarizable atomic multipole-based molecular mechanics for organic molecules. *J. Chem. Theory Comput.* 2011, 7, 3143–3161.

(12) Laury, M. L.; Wang, L.-P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised
parameters for the AMOEBA polarizable atomic multipole water model. *J. Phys. Chem. B* 2015,

(13) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Anisotropic, polarizable molecular mechanics studies of inter- and intramolecular interactions and ligand-macromolecule complexes. A bottom-up strategy. *J. Chem. Theory Comput.* 2007, 3, 1960–1986.

(14) Piquemal, J.-P.; Williams-Hubbard, B.; Fey, N.; Deeth, R. J.; Gresh, N.; Giessner-Prettre, C. Inclusion of the ligand field contribution in a polarizable molecular mechanics: SIBFA-LF. *J. Comput. Chem.* 2003, 24, 1963–1970.

(15) Gresh, N.; Claverie, P.; Pullman, A. Theoretical studies of molecular conformation. Derivation of an additive procedure for the computation of intramolecular interaction energies. Comparison with ab initio SCF computations. *Theor. Chim. Acta* 1984, 66, 1–20.

(16) Gresh, N. Energetics of Zn2+ binding to a series of biologically relevant ligands: A molecular mechanics investigation grounded on ab initio self-consistent field supermolecular computations. *J. Comput. Chem.* 1995, 16, 856–882.

(17) 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. *J. Phys. Chem. A* 2003, 107, 10353–10359.

(18) Popelier, P. L. QCTFF: On the construction of a novel protein force field. *Int. J. Quantum Chem.* 2015, 115, 1005–1011.

(19) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: a route to accurate calculations on large systems. *Chem. Rev.* 2012, 112, 632–672.
(20) Cisneros, G. A. Application of gaussian electrostatic model (GEM) distributed multipoles in the AMOEBA force field. *J. Chem. Theory Comput.* **2012**, *8*, 5072–5080.

(21) Duke, R. E.; Starovoytov, O. N.; Piquemal, J.-P.; Cisneros, G. A. GEM*: A molecular electronic density-based force field for molecular dynamics simulations. *J. Chem. Theory Comput.* **2014**, *10*, 1361–1365.

(22) Engkvist, O.; Åstrand, P.-O.; Karlström, G. Accurate intermolecular potentials obtained from molecular wave functions: Bridging the gap between quantum chemistry and molecular simulations. *Chem. Rev.* **2000**, *100*, 4087–4108.

(23) Handley, C. M.; Hawe, G. I.; Kell, D. B.; Popelier, P. L. Optimal construction of a fast and accurate polarisable water potential based on multipole moments trained by machine learning. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6365–6376.

(24) Kramer, C.; Gedeck, P.; Meuwly, M. Multipole-based force fields from ab initio interaction energies and the need for jointly refitting all intermolecular parameters. *J. Chem. Theory Comput.* **2013**, *9*, 1499–1511.

(25) Devereux, M.; Raghunathan, S.; Fedorov, D. G.; Meuwly, M. A Novel, computationally efficient multipolar model employing distributed charges for molecular dynamics simulations. *J. Chem. Theory Comput.* **2014**, *10*, 4229–4241.

(26) Unke, O. T.; Devereux, M.; Meuwly, M. Minimal distributed charges: Multipolar quality at the cost of point charge electrostatics. *J. Chem. Phys.* **2017**, *147*, 161712.

(27) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.

(28) Eastman, P. et al. OpenMM 4: A Reusable, Extensible, Hardware Independent Library
for High Performance Molecular Simulation. *J. Chem. Theory Comput.* 2013, 9, 461–469.

(29) Merz, K. M. Using Quantum Mechanical Approaches to Study Biological Systems. *Acc. Chem. Res.* 2014, 47, 2804–2811.

(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 1983, 79, 926–935.

(31) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic improvement of a classical molecular model of water. *J. Phys. Chem. B* 2013, 117, 9956–9972.

(32) Price, S. S. L. Quantifying intermolecular interactions and their use in computational crystal structure prediction. *Cryst. Eng. Comm.* 2004, 6, 344–353.

(33) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* 2005, 1, 1128–1132.

(34) Stone, A.; Alderton, M. Distributed multipole analysis. *Mol. Phys.* 1985, 56, 1047–1064.

(35) Spackman, M. A. A simple quantitative model of hydrogen bonding. *J. Chem. Phys.* 1986, 85, 6587–6601.

(36) Misquitta, A. J.; Stone, A. J.; Fazeli, F. Distributed Multipoles from a Robust Basis-Space Implementation of the Iterated Stockholder Atoms Procedure. *J. Chem. Theory Comput.* 2014, 10, 5405–5418.

(37) Devereux, M.; Gresh, N.; Piquemal, J.-P.; Meuwly, M. A Supervised Fitting Approach to Force Field Parametrization with Application to the SIBFA Polarizable Force Field. *J. Comput. Chem.* 2014, 35, 1577–1591.
(38) Gao, Q.; Yokojima, S.; Fedorov, D. G.; Kitaura, K.; Sakurai, M.; Nakamura, S. Octahedral point-charge model and its application to fragment molecular orbital calculations of chemical shifts. *Chem. Phys. Lett.* **2014**, *593*, 165–173.

(39) Storn, R.; Price, K. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *J. Glob. Optim.* **1997**, *11*, 341–359.

(40) Straatsma, T. P.; McCammon, J. A. Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water. *Molecular Simulation* **1990**, *5*, 181–192.

(41) Halgren, T. A. The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *J. Am. Chem. Soc.* **1992**, *114*, 7827–7843.

(42) Hedin, F.; El Hage, K.; Meuwly, M. A Toolkit to Fit Nonbonded Parameters from and for Condensed Phase Simulations. *J. Chem. Inf. Model.* **2016**, *56*, 1479–1489.

(43) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Comp. Biol.* **2017**, *13*, 1–17.

(44) van Gunsteren, W.; Berendsen, H. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* **1977**, *34*, 1311–1327.

(45) Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393.

(46) Melchionna, S.; Ciccotti, G.; Lee Holian, B. Hoover NPT dynamics for systems varying in shape and size. *Mol. Phys.* **1993**, *78*, 533–544.
(47) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

(48) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.

(49) Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695.

(50) Jorgensen, W. L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* **1998**, *19*, 1179–1186.

(51) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(52) Baez, L. A.; Clancy, P. Existence of a density maximum in extended simple point charge water. *J. Chem. Phys.* **1994**, *101*, 9837–9840.

(53) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6665–6670.

(54) Brooks, B. R. Computer code implemented in CHARMM. 1997.

(55) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313.

(56) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. An approach to the application of free energy perturbation methods using molecular dynamics: applications to the transformations of methanol to ethane, oxonium to ammonium, glycine to alanine, and
alanine to phenylalanine in aqueous solution and to H3O+(H2O)3 to NH4+(H2O)3 in the gas phase. *J. Am. Chem. Soc.* **1987**, *109*, 1607–1614.

(57) Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.

(58) Rick, S. W. A reoptimization of the five-site water potential (TIP5P) for use with Ewald sums. *J. Chem. Phys.* **2004**, *120*, 6085–6093.

(59) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(60) Anderson, J. A.; Tschumper, G. S. Characterizing the potential energy surface of the water dimer with DFT: failures of some popular functionals for hydrogen bonding. *J. Phys. Chem. A* **2006**, *110*, 7268–7271.

(61) Maheshwary, S.; Patel, N.; Sathyamurthy, N.; Kulkarni, A. D.; Gadre, S. R. Structure and stability of water clusters (H2O) n, n= 8-20: An ab initio investigation. *J. Phys. Chem. A* **2001**, *105*, 10525–10537.

(62) Soper, A. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *Chem. Phys.* **2000**, *258*, 121–137.

(63) Hura, G.; Sorenson, J. M.; Glaeser, R. M.; Head-Gordon, T. A high-quality X-ray scattering experiment on liquid water at ambient conditions. *J. Chem. Phys.* **2000**, *113*, 9140–9148.

(64) Fu, L.; Bienenstock, A.; Brennan, S. X-ray study of the structure of liquid water. *J. Chem. Phys.* **2009**, *131*, 234702.

(65) Leetmaa, M.; Wikfeldt, K. T.; Ljungberg, M. P.; Odelius, M.; Swenson, J.; Nilsson, A.;
Pettersson, L. G. Diffraction and IR/Raman data do not prove tetrahedral water. *J. Chem. Phys.* **2008**, *129*, 084502.

(66) Wikfeldt, K. T.; Leetmaa, M.; Mace, A.; Nilsson, A.; Pettersson, L. G. Oxygen-oxygen correlations in liquid water: Addressing the discrepancy between diffraction and extended x-ray absorption fine-structure using a novel multiple-data set fitting technique. *J. Chem. Phys.* **2010**, *132*, 104513.

(67) Petkov, V.; Ren, Y.; Suchomel, M. Molecular arrangement in water: random but not quite. *J. Phys.: Condens. Matter* **2012**, *24*, 155102.

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

(69) Head-Gordon, T.; Hura, G. Water structure from scattering experiments and simulation. *Chem. Rev.* **2002**, *102*, 2651–2670.

(70) Soper, A. Joint structure refinement of x-ray and neutron diffraction data on disordered materials: application to liquid water. *J. Phys.: Condens. Matter* **2007**, *19*, 335206.

(71) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A.; Benmore, C. J. Benchmark oxygen-oxygen pair-distribution function of ambient water from X-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.* **2013**, *138*, 074506.

(72) Mahoney, M. W.; Jorgensen, W. L. Diffusion constant of the TIP5P model of liquid water. *J. Chem. Phys.* **2001**, *114*, 363–366.

(73) Wagner, W.; Prüß, A. The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387–535.
(74) Mills, R. Self-diffusion in normal and heavy water in the range 1-45. deg. *J. Phys. Chem.* 1973, 77, 685–688.

(75) Archer, D. G.; Carter, R. W. Thermodynamic properties of the NaCl+ H2O system. 4. heat capacities of h2o and nacl (aq) in cold-stable and supercooled states. *J. Phys. Chem. B* 2000, 104, 8563–8584.

(76) Kell, G. S. Density, thermal expansivity, and compressibility of liquid water from 0. deg. to 150. deg.. Correlations and tables for atmospheric pressure and saturation reviewed and expressed on 1968 temperature scale. *J. Chem. Eng. Data* 1975, 20, 97–105.

(77) Gillen, K. T.; Douglass, D.; Hoch, M. Self-Diffusion in Liquid Water to $-31^\circ$ C. *J. Chem. Phys.* 1972, 57, 5117–5119.

(78) Ben-Naim, A.; Marcus, Y. Solvation thermodynamics of nonionic solutes. *J. Chem. Phys.* 1984, 81, 2016–2027.

(79) Hamelberg, D.; McCammon, J. A. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: double-decoupling method. *J. Am. Chem. Soc.* 2004, 126, 7683–7689.

(80) Jorgensen, W. L.; Blake, J. F.; Buckner, J. K. Free energy of TIP4P water and the free energies of hydration of CH$_4$ and Cl$^{-1}$ from statistical perturbation theory. *Chem. Phys.* 1989, 129, 193–200.

(81) Lu, Y.; Yang, C.-Y.; Wang, S. Binding free energy contributions of interfacial waters in HIV-1 protease/inhibitor complexes. *J. Am. Chem. Soc.* 2006, 128, 11830–11839.

(82) Bereau, T.; Kramer, C.; Meuwly, M. Leveraging Symmetries of Static Atomic Multipole Electrostatics in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* 2013, 9, 5450–5459.
(83) Jorgensen, W. L.; Schyman, P. Treatment of Halogen Bonding in the OPLS-AA Force Field: Application to Potent Anti-HIV Agents. *J. Chem. Theory Comput.* **2012**, *8*, 3895–3901, PMID: 23329896.

(84) Politzer, P.; Murray, J. S.; Clark, T. Halogen bonding: an electrostatically-driven highly directional noncovalent interaction. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7748–7757.

(85) Wang, J.; Hou, T. Application of Molecular Dynamics Simulations in Molecular Property Prediction. 1. Density and Heat of Vaporization. *J. Chem. Theory Comput.* **2011**, *7*, 2151–2165, PMID: 21857814.

(86) Majer, V.; Svoboda, V. *Enthalpies of Vaporization of Organic Compounds: A Critical Review and Data Compilation*; Blackwell Scientific Publications: Oxford, U.K., 1985.

(87) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 350–358, PMID: 20150953.
Supporting information for:

Application of Distributed Charges to Replace Atom-Centered, Off-Centered and Multipolar Polarizable Electrostatic Terms in Molecular Dynamics Simulations

Marco Pezzella, Mike Devereux, Shampa Raghunathan and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland

E-mail: m.meuwly@unibas.ch
Contents

1 Reduction in complexity using distributed charges S3

2 (M)DCM Charge Models S4

3 iAMOEBA DCM OpenMM force field definition file S6

4 Error Reduction in MDCMs S17

5 iAMOEBA MDCM OpenMM force field definition file S19

6 O...O Radial Distribution Functions S26

7 O...H Radial Distribution Functions S27

8 H...H Radial Distribution Functions S28

References S29
1 Reduction in complexity using distributed charges

The main benefit of using distributed charges over a truncated multipole expansion is that the electrostatic interaction energy between two multipolar atoms can be evaluated using a homogeneous set of charge-charge interaction terms in place of the heterogeneous terms between multipole moments of different ranks. I.e. all energy terms are of the form:

\[ V_{ab} = \frac{q_a q_b}{R_{ab}} \] (1)

where \( R_{ab} \) is the distance between atoms \( a \) and \( b \). By contrast, multipolar interaction terms are heterogeneous and especially higher rank terms are substantially more complex. The following example describes the interaction between the \( Q_{22c} \) terms on atoms \( a \) and \( b \):

\[
V_{Q_{22c},aQ_{22c},b} = \frac{1}{4} Q_{22c,a} Q_{22c,b} \left( 35 r_{x}^{a} r_{x}^{b} - 35 r_{x}^{a} r_{y}^{b} r_{y}^{b} - 35 r_{y}^{a} r_{x}^{b} r_{y}^{b} + 35 r_{y}^{a} r_{y}^{b} r_{y}^{b} \right) \\
+ 20 r_{x}^{a} r_{y}^{b} c_{xA} c_{xB} - 20 r_{x}^{a} r_{y}^{b} c_{xA} c_{yB} - 20 r_{y}^{a} r_{x}^{b} c_{yA} c_{xB} + 20 r_{y}^{a} r_{y}^{b} c_{yA} c_{yB} \\
+ 2 c_{xA}^{2} c_{xB} - 2 c_{xA}^{2} c_{yB} - 2 c_{yA}^{2} c_{xB} + 2 c_{yA}^{2} c_{yB} \right) R_{ab}^{-5} \] (2)

Here \( r_{\alpha}^{a} \) is the scalar product \( \hat{e}_{\alpha}^{a} \cdot \hat{e}_{ab} \) where \( \hat{e}_{\alpha}^{a} \) is a unit vector along the \( \alpha \) axis of atom \( a \) \( (\alpha = x, y, z) \). \( \hat{e}_{ab} \) is a unit vector in the direction from \( a \) to \( b \). \( \beta = x, y, z \) for atom \( b \), \( c_{\alpha,\beta} \) is the scalar product \( \hat{e}_{\alpha}^{a} \cdot \hat{e}_{\beta}^{b} \). In order to run molecular dynamics simulations the gradients of each of these terms must also be evaluated for every nonbonded atom pair at each simulation time step.
### 2 (M)DCM Charge Models

Table S1: Octahedrally arranged DCM charges (a.u.) and their coordinates (Å) for H₂O used in iAMOEBA model; atomic polarizabilities \( \alpha \) in Å³. Coordinates are relative to local atomic axes.\(^{S2}\)

|    | \( X \)       | \( Y \)       | \( Z \)       | \( q \)       | \( \alpha \) |
|----|---------------|---------------|---------------|---------------|-------------|
| H  | -0.0575896814 | 0.0000000000  | -0.0800056071 | 2.0122163666  | 0.50484     |
| \( q_1 \) | -0.0956010671 | 0.0000000000  | 0.090474640   | 1.6134817638  |             |
| \( q_2 \) | 0.0000000000  | -0.0959947647 | 0.0000000000  | -2.2882340867 |             |
| \( q_3 \) | 0.0000000000  | 0.0959947647  | 0.0000000000  | 0.2524281689  |             |
| \( q_4 \) | -0.0086850918 | 0.0000000000  | -0.0995898760 | -3.1087375276 |             |
| \( q_5 \) | 0.0086850918  | 0.0000000000  | 0.0995898760  | 1.0336186790  |             |

| O  | -0.0768011943 | 0.0000000000  | 0.0599925231  | 1.0336186790  | 0.80636     |
| \( q_1 \) | 0.0000000000  | -0.0959947647 | 0.0000000000  | -3.1087375276 |             |
| \( q_2 \) | 0.0575896814  | 0.0000000000  | 0.0800056071  | 2.0122163666  |             |
| \( q_3 \) | -0.0768011943 | 0.0000000000  | 0.0599925231  | 1.0336186790  |             |
| \( q_4 \) | 0.0000000000  | 0.0959947647  | 0.0000000000  | -3.1087375276 |             |
| \( q_5 \) | 0.0768011943  | 0.0000000000  | -0.0599925231 | 0.5654036430  |             |

| H  | 0.0956010671  | 0.0000000000  | 0.090474640   | 1.6134817638  |             |
| \( q_1 \) | 0.0000000000  | -0.0959947647 | 0.0000000000  | -2.2882340867 |             |
| \( q_2 \) | -0.0956010671 | 0.0000000000  | -0.090474640  | 2.0482594186  |             |
| \( q_3 \) | 0.0086850918  | 0.0000000000  | -0.0995898760 | -2.2882340867 |             |
| \( q_4 \) | 0.0000000000  | 0.0959947647  | 0.0000000000  | 0.2524281689  |             |
| \( q_5 \) | -0.0086850918 | 0.0000000000  | 0.0995898760  | -1.5813534334 |             |

S4
Table S2: MDCM charges (a.u.) and their positions (Å) for 10-charge H$_2$O model used for iAMOEBA; atomic polarizabilities $\alpha$ in Å$^3$. Coordinates are relative to local atomic axes.$^S_2$

|      | $X$        | $Y$        | $Z$        | $q$        | $\alpha$ |
|------|------------|------------|------------|------------|-----------|
| H    |            |            |            |            | 0.50484   |
| $q_1$| 0.0209144768 | 0.0000004420 | 0.3473443817 | -0.1066619461 |
| $q_2$| -0.0452396324 | 0.0000004981 | -0.1651346312 | -0.9725811260 |
| $q_3$| -0.1905254480 | 0.0000021077 | -0.0528319716 | 0.5997075639 |
| $q_4$| 0.1496075438 | -0.0018910650 | -0.0565055370 | 0.8289103944 |
| O    |            |            |            |            | 0.80636   |
| $q_1$| 0.1009343674 | 0.4067419560 | -0.0294610237 | -0.2369290661 |
| $q_2$| 0.0881957994 | -0.3169912077 | -0.0203134193 | -0.3142376636 |
| H    |            |            |            |            | 0.50484   |
| $q_1$| 0.1659300202 | -0.0008225336 | 0.0275557737 | 0.5465716486 |
| $q_2$| -0.1837974669 | -0.0020877351 | -0.0477859191 | 0.6204782536 |
| $q_3$| 0.0519918506 | -0.0000018568 | -0.3764240427 | -0.6244927174 |
| $q_4$| 0.0355028536 | -0.0000020937 | 0.2009935462 | -0.3407653412 |
3 iAMOEBA DCM OpenMM force field definition file

The following "XML" format force field definition file defines the iAMOEBA/DCM implementation for use in OpenMM. While dummy atom types could potentially be consolidated (all dummy atoms share the same mass and VDW, sites belonging to the same atom-type also have the same polarizability), the exhaustive approach taken here is convenient for development purposes.

```
<ForceField>
  <AtomTypes>
    <Type name="380" class="73" element="O" mass="15.999"/>
    <Type name="381" class="74" element="H" mass="1.008"/>
    <Type name="382" class="75" mass="0"/>
    <Type name="383" class="76" mass="0"/>
    <Type name="384" class="77" mass="0"/>
    <Type name="385" class="78" mass="0"/>
    <Type name="386" class="79" mass="0"/>
    <Type name="387" class="80" mass="0"/>
    <Type name="388" class="81" mass="0"/>
    <Type name="389" class="82" mass="0"/>
    <Type name="390" class="83" mass="0"/>
    <Type name="391" class="84" mass="0"/>
    <Type name="392" class="85" mass="0"/>
    <Type name="393" class="86" mass="0"/>
    <Type name="394" class="87" mass="0"/>
    <Type name="395" class="88" mass="0"/>
    <Type name="396" class="89" mass="0"/>
    <Type name="397" class="90" mass="0"/>
    <Type name="398" class="91" mass="0"/>
  </AtomTypes>
</ForceField>
```
<Type name="399" class="92" mass="0"/>
</AtomTypes>

<Residues>

<Residue name="HOH">
  <Atom name="O" type="380"/>
  <Atom name="H1" type="381"/>
  <Atom name="H2" type="381"/>
  <Atom name="M1" type="382"/>
  <Atom name="M2" type="383"/>
  <Atom name="M3" type="384"/>
  <Atom name="M4" type="385"/>
  <Atom name="M5" type="386"/>
  <Atom name="M6" type="387"/>
  <Atom name="M7" type="388"/>
  <Atom name="M8" type="389"/>
  <Atom name="M9" type="390"/>
  <Atom name="M10" type="391"/>
  <Atom name="M11" type="392"/>
  <Atom name="M12" type="393"/>
  <Atom name="M13" type="394"/>
  <Atom name="M14" type="395"/>
  <Atom name="M15" type="396"/>
  <Atom name="M16" type="397"/>
  <Atom name="M17" type="398"/>
  <Atom name="M18" type="399"/>
  <Bond from="0" to="2"/>
  <Bond from="0" to="1"/>
</Residue>

S7
<!-- H1−atom charges and positions -->
<VirtualSite type="localCoords" index="3" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="0.000000" p2="0.009551" p3="0.000000" />
<VirtualSite type="localCoords" index="4" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="0.000000" p2="−0.000000" p3="0.009591" />
<VirtualSite type="localCoords" index="5" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="−0.009959" p2="−0.000868" p3="0.000000" />
<VirtualSite type="localCoords" index="6" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="−0.009959" p2="0.000868" p3="0.000000" />
<VirtualSite type="localCoords" index="7" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="−0.000000" p2="−0.000000" p3="−0.009591" />
<VirtualSite type="localCoords" index="8" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="0.009959" p2="−0.000868" p3="0.000000" />
<!-- O−atom charges and positions -->
<VirtualSite type="localCoords" index="9" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="−0.008010" p2="0.005742" p3="0.000000" />
<VirtualSite type="localCoords" index="10" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="0.000000" p2="−0.000000" p3="−0.009591" />
<VirtualSite type="localCode" index="11" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="0.008010" p2="-0.005742" p3="0.000000" />
<VirtualSite type="localCode" index="12" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="0.005987" p2="0.007682" p3="0.000000" />
<VirtualSite type="localCode" index="13" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="0.000000" p2="-0.000000" p3="-0.009591" />
<VirtualSite type="localCode" index="14" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="0.005987" p2="-0.007682" p3="0.000000" />
<!-- H2 atom charges and positions -->
<VirtualSite type="localCode" index="15" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="-1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="-1.0" wy3="0.0" p1="0.009050" p2="-0.009551" p3="0.000000" />
<VirtualSite type="localCode" index="16" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="-1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="-1.0" wy3="0.0" p1="0.000000" p2="-0.009551" p3="-0.009591" />
<VirtualSite type="localCode" index="17" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="-1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="-1.0" wy3="0.0" p1="-0.009050" p2="-0.009551" p3="0.000000" />
<VirtualSite type="localCode" index="18" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="-1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="-1.0" wy3="0.0" p1="-0.009050" p2="-0.009551" p3="0.000000" />
<VirtualSite type="localCode" index="19" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="-1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="-1.0" wy3="0.0" p1="-0.009050" p2="-0.009551" p3="0.000000" />

S9
<VirtualSite type="localCoords" index="20" atom1="0" atom2="1" atom3="2" wo1 ="0.0" wo2="0.0" wo3="1.0" wx1="−1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2 ="−1.0" wy3="0.0" p1="0.009959" p2="0.000868" p3="0.000000" />

<Residue>

<AmoebaBondForce bond−cubic="−25.5" bond−quartic="379.3125">

<Bond class1="73" class2="74" length="9.584047e−02" k="2.3331232e+05" />

</AmoebaBondForce>

<AmoebaAngleForce angle−cubic="−0.014" angle−quartic="5.6e−05" angle−pentic="−7e−07" angle−sextic="2.2e−08">

<Angle class1="74" class2="73" class3="74" k="6.359379296918e−02" angle1="1.064826e+02" />

</AmoebaAngleForce>

<AmoebaOutOfPlaneBendForce type="ALLINGER" opbend−cubic="−0.014" opbend−quartic="5.6e−05" opbend−pentic="−7e−07" opbend−sextic="2.2e−08">

<!−− LPW: Mark's force field parsing code requires AmoebaOutOfPlaneBendForce in order to read AmoebaAngleForce, even if the clause is empty −−>

</AmoebaOutOfPlaneBendForce>

<AmoebaVdwForce type="BUFFERED−14−7" radiusrule="CUBIC−MEAN" radiustype="R−MIN" radiussize="DIAMETER" epsilonrule="HHG" vdw−13−scale ="0.0" vdw−14−scale="1.0" vdw−15−scale="1.0">

<Vdw class="73" sigma="3.645297e−01" epsilon="8.2348e−01" reduction="1.0" />

<Vdw class="74" sigma="0.0" epsilon="0.0" reduction="1.0" />

<Vdw class="75" sigma="0.0" epsilon="0.0" reduction="1.0" />

<Vdw class="76" sigma="0.0" epsilon="0.0" reduction="1.0" />

<Vdw class="77" sigma="0.0" epsilon="0.0" reduction="1.0" />

</AmoebaVdwForce>

</Residues>
<AmoebaVdwForce>

<AmoebaMultipoleForce direct11Scale="0.0" direct12Scale="1.0" direct13Scale="1.0" direct14Scale="1.0" mpole12Scale="0.0" mpole13Scale="0.0" mpole14Scale="0.4" mpole15Scale="0.8" mutual11Scale="1.0" mutual12Scale="1.0" mutual13Scale="1.0" mutual14Scale="0.0" mutual15Scale="0.0" polar12Scale="0.0" polar13Scale="0.0" polar14IntraScale="0.5" polar15Scale="1.0">

<Multipole type="380" c0="0.0" d1="0.0" d2="0.0" d3="0.0" q11="0.0" q21="0.0" q22="0.0" q31="0.0" q32="0.0" q33="0.0"/>

<Multipole type="381" c0="0.0" d1="0.0" d2="0.0" d3="0.0" q11="0.0" q21="0.0" q22="0.0" q31="0.0" q32="0.0" q33="0.0"/>

<Multipole type="382" c0="1.613482" d1="0.0" d2="0.0" d3="0.0" q11="0.0" q21="0.0" q22="0.0" q31="0.0" q32="0.0" q33="0.0"/>

</AmoebaMultipoleForce>

</AmoebaVdwForce>
<Multipole type="397" c0="-2.288234" d1="0.0" d2="0.0" d3="0.0" q11="0.0" q21="0.0" q22="0.0" q31="0.0" q32="0.0" q33="0.0"/>

<Multipole type="398" c0="0.252428" d1="0.0" d2="0.0" d3="0.0" q11="0.0" q21="0.0" q22="0.0" q31="0.0" q32="0.0" q33="0.0"/>

<Multipole type="399" c0="-1.581353" d1="0.0" d2="0.0" d3="0.0" q11="0.0" q21="0.0" q22="0.0" q31="0.0" q32="0.0" q33="0.0"/>

<Polarize type="381" polarizability="5.048434386104e-04" thole="2.36164e-03" pgrp1="380" pgrp2="382" pgrp3="383" pgrp4="384" pgrp5="385" pgrp6="386" pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12="392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399"/>

<Polarize type="382" polarizability="5.048434386104e-08" thole="2.36164e-03" pgrp1="380" pgrp2="381" pgrp3="383" pgrp4="384" pgrp5="385" pgrp6="386" pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12="392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399"/>

<Polarize type="383" polarizability="5.048434386104e-08" thole="2.36164e-03" pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="384" pgrp5="385" pgrp6="386" pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12="392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399"/>

<Polarize type="384" polarizability="5.048434386104e-08" thole="2.36164e-03" pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="385" pgrp6="386" pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12="392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399"/>
<Polarize type="385" polarizability="5.048434386104e−08" thole="2.36164e−03">
  pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
  pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12=">
  "392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399" />
</Polarize>

<Polarize type="386" polarizability="5.048434386104e−08" thole="2.36164e−03">
  pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
  pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12=">
  "392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399" />
</Polarize>

<Polarize type="387" polarizability="5.048434386104e−08" thole="2.36164e−03">
  pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
  pgrp7="386" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12=">
  "392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399" />
</Polarize>

<Polarize type="388" polarizability="5.048434386104e−08" thole="2.36164e−03">
  pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
  pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12=">
  "392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399" />
</Polarize>

<Polarize type="389" polarizability="5.048434386104e−08" thole="2.36164e−03">
  pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
  pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12=">
  "392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399" />
</Polarize>

<Polarize type="390" polarizability="5.048434386104e−08" thole="2.36164e−03">
  pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
  pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391" pgrp12=">
  "392" pgrp13="393" pgrp14="394" pgrp15="395" pgrp16="396" pgrp17="397" pgrp18="398" pgrp19="399" />
</Polarize>
4 Error Reduction in MDCMs

In this section it is demonstrated that applying constraints to charge magnitudes during MDCM fitting can reduce the corresponding error in electrostatic interaction energy between molecules of arbitrary types and relative spatial orientations.

The MDCM optimal fitted solution without applied constraints minimizes errors in the MEP, rather than the electrostatic interaction energy relevant to MD simulations. Fitting to the electrostatic interaction energy directly would avoid this issue, but also lead to biased sampling of the electric field around the molecule during fitting by focusing on those regions relevant to the electrostatic interactions included in the finite training set. It is more difficult to generate sufficiently large, heterogeneous training sets of reference electrostatic interaction energies between non-polarized monomers than to generate the more straightforward MEP across a spatial grid. A good solution is therefore to solve the more straightforward MEP fitting problem with constraints introduced to ensure that the resulting charge models will also perform well for the closely related electrostatic interaction energy between molecules.

To relate the error in the MEP to the resulting error in the electrostatic interaction between molecules, we start with a description of the residual error in the MEP of an MDCM for molecule ‘a’ after fitting:

\[
V_a(r) = \int_{\Omega_a} \rho_a(r_a) \frac{dr_a}{|r - r_a|} = \sum_{n=1}^{n_{q,a}} \frac{q_{a,n}}{|r - r_{a,n}|} + \delta V_a(r) \tag{3}
\]

for \( r \) sampled outside the molecular surface used for fitting. \( \rho_a(r_a) \) is the reference electron density of molecule ‘a’ at point \( r_a \) in the molecular volume \( \Omega_a \). \( q_{a,n} \) is charge \( n \) of the fitted MDCM model with \( r_{a,n} \) the position of the charge, and \( \delta V_a(r) \) is the residual error in the MEP after fitting the MDCM.

This can be related to the error in interaction with a single external charge, \( q_b \) at \( r_b \):

\[
U_{ab} = q_b \int_{\Omega_a} \frac{\rho_a(r_a)}{|r_b - r_a|} dr_a = q_b \sum_{n=1}^{n_{q,a}} \frac{q_{a,n}}{|r_b - r_{a,n}|} + q_b \delta V_a(r_b) \tag{4}
\]
Now using the exact interaction between two molecules, a double integral over the molecular volumes of \( a \) and \( b \):

\[
U_{ab} = \int_{\Omega_a} \int_{\Omega_b} \frac{\rho_a(r_a)\rho_b(r_b)}{|r_b - r_a|} dr_a dr_b
\]  

(5)

and inserting (3) and rearranging:

\[
U_{ab} = \sum_{n=1}^{n_{q,a}} q_{a,n} \int_{\Omega_b} \frac{\rho_b(r_b)}{|r_b - r_{a,n}|} + \int_{\Omega_b} \rho_b(r_b)\delta V_a(r_b) dr_b
\]

(6)

then using (3) again for molecule \( b \):

\[
U_{ab} = \sum_{n=1}^{n_{q,a}} q_{a,n} \sum_{m=1}^{n_{q,b}} q_{b,m} \frac{|r_{b,m} - r_{a,n}|}{|r_{b,m} - r_{a,n}|} + \sum_{n=1}^{n_{q,a}} q_{a,n} \delta V_b(r_{a,n}) + \int_{\Omega_b} \rho_b(r_b)\delta V_a(r_b) dr_b
\]

(7)

or, equivalently:

\[
U_{ab} = \sum_{n=1}^{n_{q,b}} q_{b,n} \sum_{m=1}^{n_{q,a}} \frac{q_{a,m}}{|r_{a,m} - r_{b,n}|} + \sum_{n=1}^{n_{q,b}} q_{b,n} \delta V_a(r_{b,n}) + \int_{\Omega_a} \rho_a(r_a)\delta V_b(r_a) dr_a
\]

(8)

Minimizing the \( \delta V_a(r) \) error term in (3) requires simply a charge model that describes the MEP as closely as possible to the corresponding Coulomb integral over the molecular charge density. In practice this often results in large fitted charges, of the order of several a.u. The error terms in (7) and (8) for the electrostatic interaction energy, however, show that the error in the MEP is multiplied by the magnitude of each interacting charge. While error cancellation is possible it is not guaranteed, so the presence of larger charges can lead to an amplification of the error in the MEP. For a typical fitting problem it has been observed that for a given MDCM model containing large charges, the error in the electrostatic interaction energy is amplified in some regions but remains small due to error cancellation in others. It is therefore safest to constrain the magnitude of the fitted charges to remain as low as possible while still achieving an acceptable error in the fitted MEP, i.e. to find a compromise solution with low errors in the fitted MEP and the smallest possible charges.
5 iAMOEBA MDCM OpenMM force field definition file

The following "XML" format force field definition file defines the iAMOEBA/MDCM implementation for use in OpenMM. Dummy atom types could again potentially be consolidated as all MDCM dummy atoms also share the same mass and VDW, and sites belonging to the same atom-type also have the same polarizability.

```xml
<ForceField>
  <AtomTypes>
    <Type name="380" class="73" element="O" mass="15.999"/>
    <Type name="381" class="74" element="H" mass="1.008"/>
    <Type name="382" class="75" mass="0"/>
    <Type name="383" class="76" mass="0"/>
    <Type name="384" class="77" mass="0"/>
    <Type name="385" class="78" mass="0"/>
    <Type name="386" class="79" mass="0"/>
    <Type name="387" class="80" mass="0"/>
    <Type name="388" class="81" mass="0"/>
    <Type name="389" class="82" mass="0"/>
    <Type name="390" class="83" mass="0"/>
    <Type name="391" class="84" mass="0"/>
  </AtomTypes>
  <Residues>
    <Residue name="HOH">
      <Atom name="O" type="380"/>
      <Atom name="H1" type="381"/>
      <Atom name="H2" type="381"/>
    </Residue>
  </Residues>
</ForceField>
```
<Atom name="M1" type="382"/>
<Atom name="M2" type="383"/>
<Atom name="M3" type="384"/>
<Atom name="M4" type="385"/>
<Atom name="M5" type="386"/>
<Atom name="M6" type="387"/>
<Atom name="M7" type="388"/>
<Atom name="M8" type="389"/>
<Atom name="M9" type="390"/>
<Atom name="M10" type="391"/>

<Bond from="0" to="2"/>
<Bond from="0" to="1"/>

<!-- H1--atom charges and positions-->
<VirtualSite type="localCoords" index="3" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="0.034734" p2="-0.002091" p3="0.000000" />
<VirtualSite type="localCoords" index="4" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="-0.016513" p2="0.004524" p3="0.000000" />
<VirtualSite type="localCoords" index="5" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="-0.005283" p2="-0.019053" p3="0.000000" />
<VirtualSite type="localCoords" index="6" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="1.0" wo3="0.0" wx1="-1.0" wx2="1.0" wx3="0.0" wy1="-1.0" wy2="0.0" wy3="1.0" p1="-0.005651" p2="-0.014961" p3="-0.000189" />

<!-- O--atom charges and positions-->

S20
<VirtualSite type="localCoords" index="7" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="−0.002946" p2="−0.010093" p3="0.040674" />
<VirtualSite type="localCoords" index="8" atom1="0" atom2="1" atom3="2" wo1="1.0" wo2="0.0" wo3="0.0" wx1="−1.0" wx2="1.0" wx3="0.0" wy1="−1.0" wy2="0.0" wy3="1.0" p1="−0.002031" p2="−0.008820" p3="−0.031699" />
<!-- H2--atom charges and positions -->
<VirtualSite type="localCoords" index="9" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="−1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="−1.0" wy3="0.0" p1="0.002756" p2="−0.016593" p3="−0.000082" />
<VirtualSite type="localCoords" index="10" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="−1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="−1.0" wy3="0.0" p1="−0.004779" p2="−0.018380" p3="−0.000209" />
<VirtualSite type="localCoords" index="11" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="−1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="−1.0" wy3="0.0" p1="−0.037642" p2="−0.005199" p3="−0.000000" />
<VirtualSite type="localCoords" index="12" atom1="0" atom2="1" atom3="2" wo1="0.0" wo2="0.0" wo3="1.0" wx1="−1.0" wx2="0.0" wx3="1.0" wy1="1.0" wy2="−1.0" wy3="0.0" p1="0.020099" p2="−0.003550" p3="−0.000000" />
</Residue>
</Residues>

<AmoebaBondForce bond−cubic="−25.5" bond−quartic="379.3125">
<Bond class1="73" class2="74" length="9.584047e−02" k="2.3331232e+05" />
</AmoebaBondForce>

<AmoebaAngleForce angle−cubic="−0.014" angle−quartic="5.6e−05" angle−pentic="−7e−07" angle−sextic="2.2e−08">
<Angle class1="74" class2="73" class3="74" k="6.359379296918e−02" angle1="
1.064826e+02" />
</AmoebaAngleForce>
<AmoebaOutOfPlaneBendForce type="ALLINGER" opbend−cubic="−0.014" opbend
−quartic="5.6e−05" opbend−pentic="−7e−07" opbend−sextic="2.2e−08" >
<!−− LPW: Mark’s force field parsing code requires AmoebaOutOfPlaneBendForce, even if the clause is empty−−>
</AmoebaOutOfPlaneBendForce>
<AmoebaVdwForce type="BUFFERED−14−7" radiusrule="CUBIC−MEAN" ∞
radiustype="R−MIN" radiussize="DIAMETER" epsilonrule="HHG" vdw−13−scale
=“0.0” vdw−14−scale="1.0” vdw−15−scale="1.0” >
   <Vdw class="73” sigma="3.645297e−01" epsilon="8.2348e−01” reduction="1.0” />
   <Vdw class="74” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="75” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="76” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="77” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="78” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="79” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="80” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="81” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="82” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="83” sigma="0.0” epsilon="0.0” reduction="1.0” />
   <Vdw class="84” sigma="0.0” epsilon="0.0” reduction="1.0” />
</AmoebaVdwForce>
<AmoebaMultipoleForce direct11Scale="0.0” direct12Scale="1.0” direct13Scale="1.0”
direct14Scale="1.0” mpole12Scale="0.0” mpole13Scale="0.0” mpole14Scale="0.4”
mpole15Scale="0.8” mutual11Scale="1.0” mutual12Scale="1.0” mutual13Scale
=“1.0” mutual14Scale="1.0” polar12Scale="0.0” polar13Scale="0.0” polar14Intra
S22
| Multipole type | c0       | d1 | d2 | d3 | q11  | q21  | q22  | q31  | q32  | q33  |
|----------------|----------|----|----|----|------|------|------|------|------|------|
| 380            | 0.0      | 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 381            | 0.0      | 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 382            | -0.106662| 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 383            | -0.972581| 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 384            | 0.599708 | 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 385            | 0.828910 | 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 386            | -0.236929| 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 387            | -0.314238| 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 388            | 0.546572 | 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 389            | 0.620478 | 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 390            | -0.624493| 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 391            | -0.340765| 0  | 0  | 0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Polarize       |          |    |    |    |      |      |      |      |      | 8.063631227791e-04 |

Polarizability: 8.063631227791e-04, thole: 2.36164e-03
<Polarize type="381" polarizability="5.048434386104e−04" thole="2.36164e−03">
    pgrp1="380" pgrp2="382" pgrp3="383" pgrp4="384" pgrp5="385" pgrp6="386"
pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="382" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="384" pgrp5="385" pgrp6="386"
pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="383" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="384" pgrp5="385" pgrp6="386"
pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="384" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="385" pgrp6="386"
pgrp7="387" pgrp8="388" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="385" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="386" polarizability="8.063631227791e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="387" polarizability="8.063631227791e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="388" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="389" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="390" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>

<Polarize type="391" polarizability="5.048434386104e−08" thole="2.36164e−03">
    pgrp1="380" pgrp2="381" pgrp3="382" pgrp4="383" pgrp5="384" pgrp6="385"
pgrp7="386" pgrp8="387" pgrp9="389" pgrp10="390" pgrp11="391"
</Polarize>
<Polarize type="390" polarizability="5.048434386104e-08" thole="2.36164e-03"/>

<Polarize type="391" polarizability="5.048434386104e-08" thole="2.36164e-03"/>

<AmoebaMultipoleForce>

<AmoebaUreyBradleyForce cubic="0.0" quartic="0.0">
  <UreyBradley class1="74" class2="73" class3="74" k="-4.31294e+03" d="1.535676676685e-01"/>
</AmoebaUreyBradleyForce>

</ForceField>
6 O...O Radial Distribution Functions

Figure S1: Oxygen–oxygen radial distribution functions for liquid water at 298.15 K and 1 atm. DCM and MDCM implementations are compared with results using the original, multipolar water iAMOEBA water model and with experimental neutron diffraction data.\textsuperscript{S3}
7 O...H Radial Distribution Functions

Figure S2: Oxygen–hydogen radial distribution functions for liquid water at 298.15 K and 1 atm. Various simulation methods are compared against experimental Neutron diffraction data. Successive curves are offset 1 unit along the $y$-axis for clarity.
8 H...H Radial Distribution Functions

Figure S3: Hydrogen–hydrogen, radial distribution functions for liquid water at 298.15 K and 1 atm. Various simulation methods are compared against experimental Neutron diffraction data. Successive curves are offset 1 unit along the y-axis for clarity.
References

(S1) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, U.K., 2000.

(S2) Devereux, M.; Raghunathan, S.; Fedorov, D. G.; Meuwly, M. A Novel, computationally efficient multipolar model employing distributed charges for molecular dynamics simulations. *J. Chem. Theory Comput.* 2014, 10, 4229–4241.

(S3) Soper, A. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *Chem. Phys.* 2000, 258, 121–137.