# End-To-End Learning of Classical Interatomic Potentials for Benchmarking Anion Polarization Effects in Lithium Polymer Electrolytes

Pablo A. León,[†] Avni Singhal,[†] Jurģis Ruža,[†] Jeremiah Johnson,[‡] Yang Shao-Horn,[¶,†,§] and Rafael Gómez-Bombarelli*[,†]

†*Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA*

‡*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA*

¶*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA*

§*Research Laboratory of Electronics,Massachusetts Institute of Technology, Cambridge,Massachusetts 02139, United States*

E-mail: rafagb@mit.edu

**Abstract**

Solid polymer electrolytes are an exciting solution for safe and stable solid lithium electrode battery systems but are hindered by low ionic conductivity and low lithium transference. All-atom molecular dynamics simulation has become an invaluable tool to probe lithium diffusion mechanisms and accelerate the discovery of promising polymer chemistries. Because of their low computational cost and despite their approximate nature, only classical interatomic potentials can access the time and length scales for appropriate statistics of polymer kinetics. Machine learning (ML) potentials trained

end-to-end on *ab initio* data have proven more accurate but cannot be scaled to the necessary time- and length- scales yet. Historical approaches to parameterize classical force fields have been incremental, reliant on a manual combination of top-down and bottom-up fitting, and are often paywalled and hard to reproduce. We introduce a computational learning workflow to predict classical interatomic potential parameters using quantum mechanical computations as training data that combines the automation and end-to-end fitting of ML with traditional class 1 and class 2 functional forms. The fitting strategy produced potentials whose simulations improved the accuracy of lithium coordination environments, diffusivities, and conductivities relative to experimental approaches when compared to both naive and hand-tuned parameters for liquid and solid organic electrolyte systems. We show that chemistry-informed regularization is necessary to constrain predicted parameters in order to reproduce experimental solvation and kinetic properties. Finally, we explore the limitations of non-polarizable, fixed point-charge schemes in describing electrolyte anions and compare the effects of two alternative schemes to fit point-charge distributions. The two strategies result in distinct lithium coordination mechanisms and highlight that closest parity to DFT forces and energies does not correlate to correct trends with lithium salt concentration in kinetic and solvation properties for fixed-point-charge classical interatomic potentials.

# Introduction

Safe batteries with high charge-density and high power are a critical part of the strategy towards net zero carbon objectives. Lithium ion has emerged as the dominant battery chemistry thanks to its decreasing costs, energy density, efficiency, and both cycling and calendar stability. A key area for further improvement of Li ion (or post-lithium) battery chemistry is electrolyte design. Commercial Li ion batteries utilize liquid electrolytes whose flammability and poor mechanical properties can result in catastrophic failure. Solid polymer

electrolytes (SPEs) have garnered much attention as safe, lightweight, and easily processable alternatives to the conventionally-used, liquid carbonate-based electrolytes for lithium ion batteries. However, SPEs exhibit low lithium ion conductivity due to slower polymer chain motion relative to liquid motion and due to incomplete solvation of the lithium salt, leading to reduced lithium transference number, lower power output, and slower charging.[1–3]

Molecular dynamics simulations are often used to determine the atomistic mechanisms of lithium diffusion in electrolyte systems. Recent computational and experimental studies have determined the charge diffusion of the most well-known polymer electrolyte system, polyethylene oxide (PEO) with lithium bis(triuoromethane-sulfonyl)imide (LiTFSI), to be driven by net-negative ion clusters of mostly anions.[1,4,5] Simulations uncovered that this preference towards net-negative ion clusters is due to the lithium ions coordinating too strongly to the oxygens in PEO, effectively removing them as charge carriers and leaving an excess of anions to diffuse. Even so, there are unexplored disparities in the exact solvation structure and kinetic properties emergent across different PEO simulation studies.

The results of molecular dynamics simulations are highly dependent on the interatomic potential (IP) used. Ab-initio molecular dynamics (AIMD) and neural network-based (NNIP) IPs have been very successful at reproducing structural properties of various material and molecular systems.[6,7] However, these are too computationally expensive to access the time- and length- scales needed to ergodically investigate amorphous solid polymer electrolyte systems, which are in the hundreds of nanoseconds and tens of thousands of atoms.[2]

Thus, classical IPs which incorporate nonreactive covalent interactions, Lennard-Jones dispersion interactions, and nonpolarizable point charge Coulomb interactions are the most accessible approach for modeling these systems. However, the most prominent classical IPs such as OPLS,[8,9] PCFF+,[10] and COMPASS[11] were parameterized through decades of hand-tuning through a combination of bottom up fitting on computational chemistry and top-down fitting to reproduce experimental values. Additionally, these parameters are often obfuscated by commercial licenses. Furthermore, it is common practice to independently adjust these IP

parameters (especially the partial charges assigned to each atom) for individual studies. For example, France-Lanord[12] and Molinari[4] both report report using the proprietary MedeA package from Materials Design to obtain the interatomic potential yet still use different partial charges on the TFSI atoms. A methodical and reproducible parameterization scheme to benchmark existing literature is needed.

A variety of studies have begun exploring the use of NNIP-based workflows to predict classical interatomic potential parameters such as Espaloma,[13,14] BespokeFit,[15] the graph neural network for intermolecular forces proposed by Thürlemann et al,[16] ByteFF,[17] and ForceBalance.[18] However, these studies focus on small molecule systems for biochemical applications, and no workflow has been applied to determining kinetic properties such as diffusivity, conductivity, and lithium-ion transference number of condensed polymer systems.

In this work, we introduce a new machine learning workflow named AutoBADDIE (Automatic Bonds, Angles, Dihedrals, Dispersion, Impropers, and Electrostatics) which predicts classical interatomic potential parameters for condensed systems by recreating quantum mechanical-derived forces and point charges. AutoBADDIE determines unique chemical environments in input molecules using a message passing algorithm, allowing for autonomous enumeration of discrete environments within a certain graph depth and end-to-end optimization of all the arising interatomic potential parameters. Optimized classical IP parameters are then directly determined through gradient-based optimization from quantum mechanical simulations of molecular clusters.

We used this workflow to rigorously benchmark the effects of interatomic potential parameterization on simulation results, in particular the point charge distribution of the TFSI anion in simulations of the prototypical SPE system. We first showed that AutoBADDIE is able to predict purely harmonic class 1 (OPLS-like) IP parameters that reproduce solvation and kinetic behavior of liquid carbonate and PEO-based SPE systems with good agreement to experiments and prior simulation literature. Then, we used AutoBADDIE to predict class 2 IP parameters (anharmonic covalent interactions as well as covalent cross terms and

a softer Lennard-Jones stopping potential, like PCFF+ and COMPASS) for PEO to explore the effect of different nonpolarizable atomic charge partitions of TFSI geometries on resulting solvation and kinetic properties.

In addition to demonstrating the ability of AutoBADDIE to fit IPs, we found that the agreement between DFT-computed and classical IP-computed forces and energies for gas-phase clusters of component molecules does not necessarily correlate with accurate solvation and kinetic properties of condensed phase simulations of polymers. Our findings highlight the complexities of using methodical machine learning-based approaches to predict classical IPs due to the limited expressivity of nonpolarizable charge interactions to model charge-dependent systems.

# Methods

AutoBADDIE is a modular workflow to parameterize classical interatomic potentials for molecular systems. Atomic interactions of different functional forms can be added or excluded from the learned potential through the choice of terms included in the energy calculation. AutoBADDIE parameterizes an IP in three steps: 1) Autonomously determining the discrete atomic environments in the system; 2) enumerating bond, angle, dihedral, improper and pairwise interactions necessary to fully describe the system; 3) optimizing the enumerated parameters through end-to-end optimization to match the force and partial charges from quantum chemical training data.

## Atomic environment enumeration through message-passing algorithm

As described in Figure 1, the workflow uses a message passing-based algorithm to determine unique atomic environments and classical potential parameters needed to describe the system. Molecules are represented as graphs, where each atom is a node whose feature is a

one-hot encoding representing the element of the atom and each bond is an edge connecting two nodes. A message passing algorithm is used with a depth of 3 nearest neighbors to further differentiate possible node types. These unique node types are used to enumerate distinct types for higher-order atom interactions (angles, dihedrals, and impropers) based on the unique permutations of atom types contained within each topology. Further explanation of this autonomous topology determination can be found in Supporting Information Section S1.1.

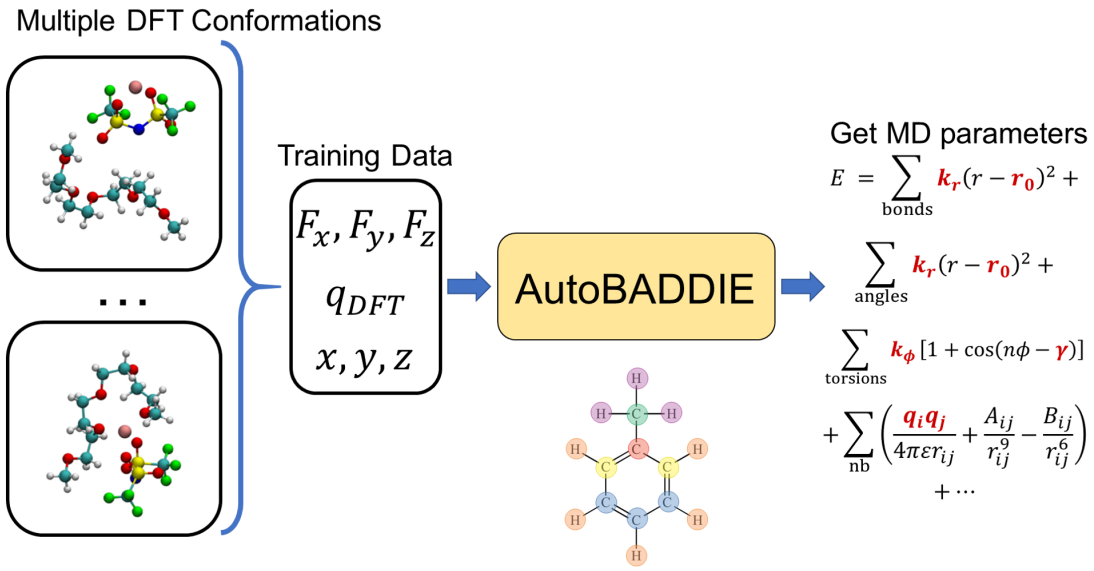


Figure 1: AutoBADDIE learns classical interatomic potential parameters (red variables) by matching quantum chemical calculations of force and partial charges. The colors of each atom in the center graphic correspond to distinct atomic environments autonomously determined through a message-passing algorithm.

These parameters are then fit to recover quantum mechanical forces and partial charges across a range of configurational space for many different combinations of both single-molecules and molecular clusters encompassing the atomic environments needed for condensed-phase simulations. These geometries are created through quantum-chemical constrained-optimization and through classical MD using learned IP in an active learning loop. The predicted parameters are output as LAMMPS data files for use in molecular dynamics simulations. Additional information about the training scheme and training data generation can

be found in Supporting Information Section S1.

## Classical molecular dynamics potentials

Class 1 IPs such as OPLS[9] contain only harmonic bonded interactions with periodic torsions (Equation 1) and have mainly been parameterized for small organic molecules. Class 2 IPs such as PCFF+[10] and COMPASS[11] add additional interactions (e.g. anharmonicity and cross-terms for covalent interactions as well as a softer stopping potential for the Lennard-Jones interaction) and have been typically applied polymer systems.

$$\begin{aligned} E_i = & \sum_{bond} K_r(r - r_0)^2 + \sum_{angle} K_\theta(\theta - \theta_0)^2 + \\ & \sum_{tors} V_1(1 + \cos\phi)/2 + V_2(1 - \cos 2\phi)/2 + V_3(1 + \cos 3\phi)/2 + \\ & \sum_{pair} f_{ij} \left[ \frac{q_i q_j e^2}{r_{ij}} + 4\epsilon_{ij} \left( (\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^6 \right) \right] \end{aligned} \tag{1}$$

where $K_r$, $K_\theta$, $r_0$, $\theta_0$, $V_i$, $q_i$, $q_j$, $\epsilon_{ij}$, and $\sigma_{ij}$ are material- and chemistry-specific parameters.

We parameterize both Class 1 and Class 2 IPs using our workflow, on DFT-level quantum chemical training data. The Lennard-Jones parameters were set from prior literature,[12,19,20] which are typically fitted from experimental data in condensed phase since (i) they show lower sensitivity to the chemical environment, (ii) act a longer length scales than the bonded terms and require larger clusters to train on, and (iii) are not well-captured with DFT and typically require accurate wavefunction methods like CCSD(T).

## Optimization of enumerated interatomic potential parameters

In contrast to traditional, piece-wise parameterizations of classical potentials, AutoBADDIE operates using an ADAM optimizer[21] to directly learn all classical IP parameters based on the mean squared error (MSE) between DFT and classical IP forces and energies using the PyTorch package in Python, building on the well established paradigm of force-matching.

AutoBADDIE is able include modular energy terms into the IP, including all the bonded terms in Class 1 and Class 2 IPs.

Additionally, we introduce two chemistry-informed regularization terms in the loss function to improve the prediction of dihedrals and partial charges (Equation 2). The first regularization term applies L1 regularization to the periodic dihedral stiffnesses in order to sparsify to few torsional periodicties. The improvement of explored configurational space for flexible molecules due to the inclusion of this regularization is further explained in Supporting Information Section 2. The charge regularization ($L_q$) is composed of two terms (Equation 3), the first of which regularizes the partial charges towards the average DFT electron density-partitioned point charges[22] per atom type in the training data geometries. The second charge regularization term enforces integer net charges.

$$L = \frac{\sum(F_{learned} - F_{DFT})^2}{3 * N_{atoms}} + \lambda_1 \frac{\sum(E_{learned} - E_{DFT})^2}{N_{sys}} + \lambda_2 \mid k_{dihed} \mid + L_q \tag{2}$$

$$L_q = \lambda_3 \frac{\sum(q_{learned} - q_{DFT}^{avg})^2}{N_{atoms}} + \lambda_4 \frac{\sum \left( q_{molecule} - \begin{cases} -1 & \text{if anion} \\ 0 & \text{if solvent} \\ 1 & \text{if cation} \end{cases} \right)^2}{N_{molecules}} \tag{3}$$

# Results and Discussion

To benchmark the ability of AutoBADDIE to predict physically-correct force field parameterizations, we use the AutoBADDIE workflow to predict class 1 IP parameters from quantum-chemical DFT simulations for both liquid and solid polymer electrolyte systems and compare the resulting MD simulations to available IPs and simulation literature. Molecular dynamics simulation details are described in the Supporting Information Section 3.

Then, we use AutoBADDIE to study the effects of different static polarizations of TFSI on resulting structural and kinetic properties of the simulated solid polymer electrolyte sys-

tem. For this, we predict class 2 IP parameters for the polymer and use literature TFSI parameters[12] with different static point charges derived from quantum mechanical calculations. Finally, we compare differences in lithium coordination environments, diffusion, and transference numbers between simulations driven by these distinct interatomic potentials.

## OPLS Liquid Carbonate Electrolyte

To benchmark AutoBADDIE on standard condensed phase lithium electrolyte systems, four liquid carbonate system containing combinations of ethylene carbonate (EC), fluoroethylene carbonate (FEC), and ethyl-methyl carbonate (EMC) with hexafluorophosphate ($PF_6^-$) anions were investigated by optimizing a class 1 IP containing all five molecule types using AutoBADDIE and comparing to Huo et al.[23] EC-Base contains only EC as the solvent, while ECF contains EC and FEC. Gen2 contains EC and the linear EMC while GenF contains all three solvent molecules. Full analysis of liquid carbonate IP training and lithium ion solvation and kinetic properties using AutoBADDIE can be found in Supporting Document #2.

As seen by Figure 2a, AuTopology is able to predict the conductivity of all four systems within half an order of magnitude of experimental expectations, closer than the OPLS-AA+Lopes+Jensen parameters used by Huo et al. Importantly, the differences in conductivities reported by Huo et al. for all four systems are statistically indistinguishable while the EMC-containing systems (Gen2 and GenF) have lower conductivity when using the AutoBADDIE-derived IP. Experimental results reported by Huo et al. show an increase in conductivity with the addition of EMC.

AutoBADDIE simulations with EMC show a lower lithium ion coordination to EMC oxygens and higher coordination to the slower EC oxygens relative to simulations by Huo et al. (the heights of the light blue bars in Figure 2b). This preference of lithium ions to interact with the slower EC molecules rather than the faster EMC molecules, as confirmed by computed lithium interaction energies with all other molecules using both IP and DFT seen

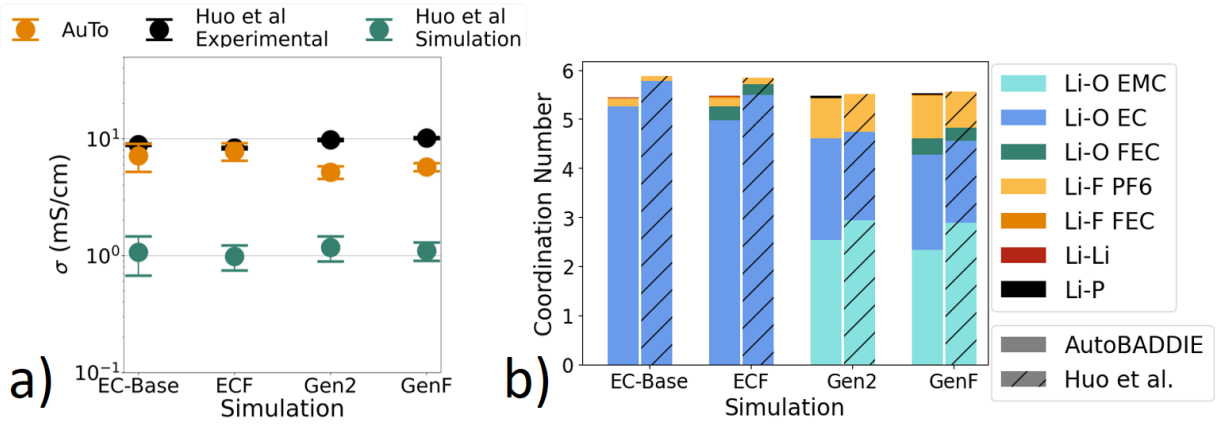


Figure 2: a) Wheeler-Newman conductivity results for different liquid carbonate systems compared to Huo et al[23] b) Lithium coordination number to different solvation and anion atoms.

in Supporting Document #2 Figure S1, could explain this decrease in conductivity seen in AutoBADDIE simulations. Additional training data, such as binary clusters containing EMC and the other carbonates, may help AutoBADDIE improve prediction of EMC parameters to improve the quality of the learned IP.

## OPLS Solid Polymer Electrolyte

In order to test the accuracy and lithium diffusion mechanisms of predicted AutoBADDIE parameters on solid polymer electrolyte systems which depend strongly on by dihedral interactions of polymer backbones, we parameterized a class 1 interatomic potential for the polyethylene oxide (PEO)-lithium-TFSI electrolyte system. AutoBADDIE was benchmarked against simulations using class 1 interatomic potential parameters available in literature, namely, OPLS-AA parameters for the glyme parameters through LigParGen[9,19,24] combined with lithium and TFSI parameters from Doherty et al.[20] (OPLS-AA+D). AutoBADDIE was also compared to simulations conducted by Brooks et al.[25] using the OPLS2005 parameters,[26] which are proprietary and not available in published literature. We show that AutoBADDIE, with the inclusion of a chemistry-informed dihedral regularization term, performs better than naive open-source simulation workflows (OPLS-AA+D) and similarly to

highly-tuned proprietary workflows (Brooks et al.).

## Training

DFT forces and energies of neural network molecular dynamics (NNIPMD) poses of tetraglyme, lithium, and TFSI clusters from a previous study[27] were used as training data to parameterize this class 1 IP. More details of the training data can be found in Supporting Information Section S1.2. Figure 3 shows the mean absolute error (MAE) between DFT-computed forces and energies and forces and energies computed through various class 1 IPs.

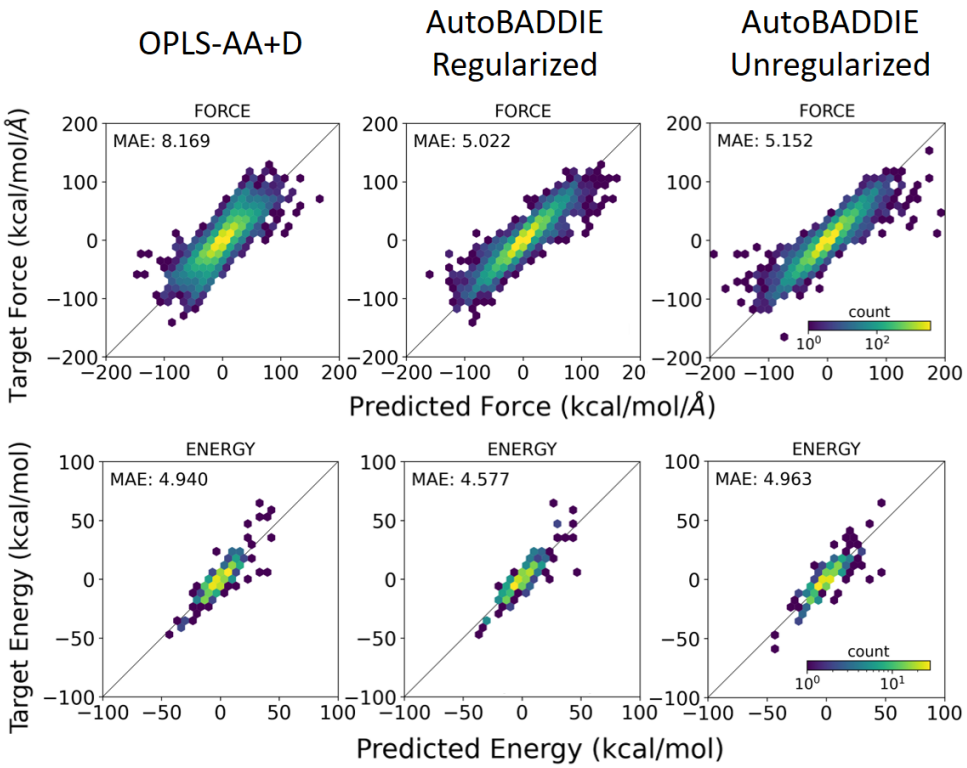


Figure 3: Force and energy parity of different class 1 interatomic potentials on neural network molecular dynamics poses acquired from Wang et al.[27] a) DFT force and energy parity computed with the combined class 1 forcefield parameters of OPLS-AA (glyme) and Doherty (lithium and TFSI).[20] bc) DFT force and energy parity from parameters predicted by AutoBADDIE with and without dihedral regularization, respectively.

AutoBADDIE is able to better predict the DFT forces (5.022 vs 8.169 kcal/mol Å) and energies (4.577 vs 4.940 kcal/mol) than the OPLS-AA+D interatomic potential, as seen by the mean absolute errors presented in Figure 3. The inclusion of an L1 loss as described in

Supporting Information Section 2 counter-intuitively allows for further reduction in MAE of both forces and energies with respect to DFT. Since the OPLS2005 parameters used by Brooks et al. are not available in published literature, the parity with respect to DFT for this interatomic potential could not be determined.

### Kinetic analysis

Figure 4 shows the computed Wheeler-Newmann conductivities and the diffusivities of the cation and anions as a function of lithium salt concentration for simulations driven by different class 1 interatomic potentials. Although the regularized and unregularized AutoBADDIE simulations showed similar force and energy parities with respect to DFT, the simulations using regularized parameters show closer agreement with published experimental literature conductivities and are also able to reproduce the concentration at which the conductivity is maximum to be r=0.075 $Li^+$/EO. Meanwhile, simulations using the unregularized parameters show the maximum conductivity to occur at a lower concentration.

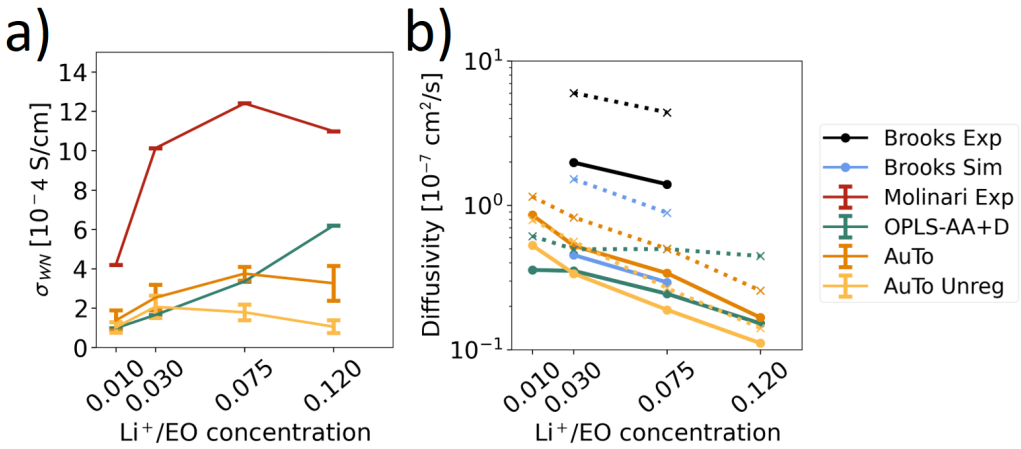


Figure 4: a) Correlation-based conductivities from OPLS-AA+D and AutoBADDIE simulations. Experimental conductivities are from Molinari et al.[4] b) Diffusivities of lithium (solid) and TFSI (dotted) reported by Brooks et al.[25] and simulated with OPLS-AA+D and AutoBADDIE.

.

Published parameters, OPLS-AA+D, are unable to recover the correct trends for either kinetic property. Brooks et al. do not report conductivities for their simulations and therefore

are not included in the conductivity comparisons. However, the diffusivities reported by Brooks et al. are in better agreement to experimental results than from simulations using parameters from AuToBADDIE, as seen in Figure 4b.

### Solvation

To further benchmark the performance of the parameters predicted by AutoBADDIE, the radial distribution functions and coordination numbers between lithium ions and TFSI oxygens were computed for the OPLS-AA+D, regularized AutoBADDIE, and unregularized AutoBADDIE simulations as seen in Figure 5. As seen in plot a), the lithium-TFSI coordination for OPLS-AA+D is much less structured and displays a lower coordination number relative to the simulations driven by AutoBADDIE parameters, especially in higher salt concentration simulations. This reduced interaction using OPLS-AA+D, as seen by the lower peak at 2.5 Å and lack of peak at around 4 Å in Figure 5a compared to 5b and 5c, could explain why the conductivity and diffusivity do not decrease as as lithium salt concentration increases in the OPLS-AA+D simulations in Figure 4.

The regularized AutoBADDIE parameters (panel b) led to a less structured Li-TFSI solvation than the unregularized parameters (panel c) as seen by the difference in magnitude of peaks at around 2.5 and 4 Å and the overall lower coordination number of first solvation shell (plateau in dashed lines) in Figure 5. However, both are similar relative to the OPLS-AA+D simulations. Therefore, differences in both kinetic properties (conductivity and diffusivity) and solvation properties (RDF and coordination numbers) are seen between the regularized and unregularized sets of parameters optimized by AutoBADDIE. The torsional-based regularization leads to polymer conformations that better match NNIPMD, as seen in Supporting Information Section 2, which seems to lead to conductivities and solvation environments closer to experimental expectations.

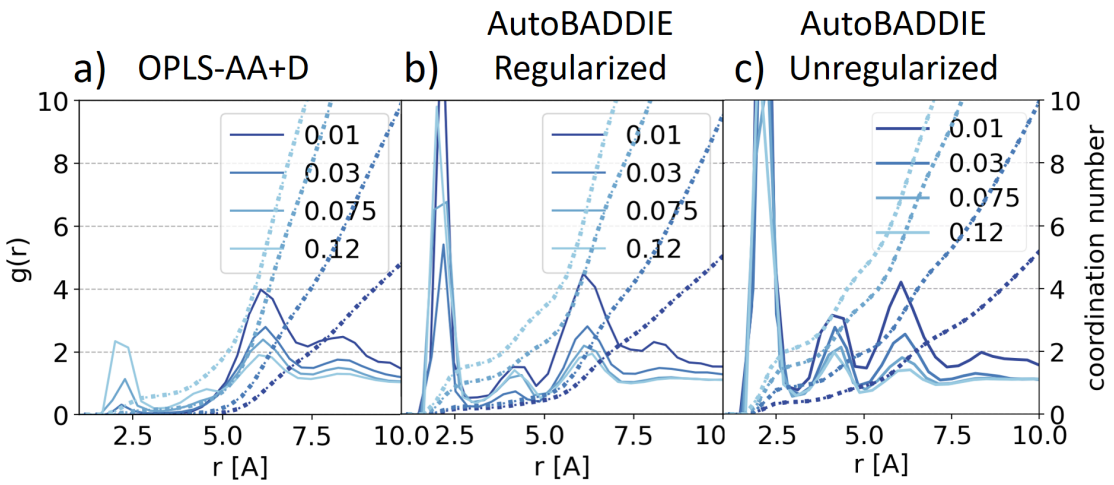


Figure 5: Lithium solvation analysis for lithium solid polymer electrolyte using different class 1 interatomic potentials. Lithium-TFSI (oxygen) radial distribution function (solid lines) for a) OPLS-AA+D, b) AutoBADDIE regularized, c) AutoBADDIE unregularized simulations, at four distinct lithium salt to monomer concentration ratios [r=Li/EO]. Note: The coordination number corresponds to the value of the plateau in coordination (dotted lines) around 3 Å.

## TFSI Charge Partitioning Effect on Nonpolarizable Simulations

After AutoBADDIE was benchmarked against prior simulation literature, we used AutoBADDIE to investigate the effect of different partial charge distributions of TFSI on structural and kinetic properties of the simulated polyelectrolyte systems, comparing to Molinari et al.[4] and France-Lanord et al.[12]

### Nonpolarizeable point charges for polyethylene oxide Lithium TFSI

Because of the nonpolarizable nature of class 1 and class 2 IPs, the per-atom partitioning of the quantum mechanical electron density significantly influences simulation results. There are large differences in the relative polarizations of the oxygen and nitrogen atoms in different IPs from simulation literature of PEO-LiTFSI, with no clear justification for variations across studies.[4,12,20,25] To identify a chemical explanation for these vastly different polarization states of TFSI, we performed DBSCAN clustering[28,29] using SciKitLearn[30] on all of the LiTFSI geometries obtained from trajectories obtained using a previously-trained ether ionic

liquid Neural-Network Molecular Dynamics (NNIPMD)[31] model according to the lithium to anion negative atom (O, N, F) distances. The results of this clustering are seen in Figure 6, where each cluster of geometries was found to represent a different Li-TFSI coordination mechanism as noted by the molecular visualizations in Figure 6b. The AuTo+O and AuTo+N charges are the average charges of all NNIPMD geometries corresponding to each respective coordination mechanism.

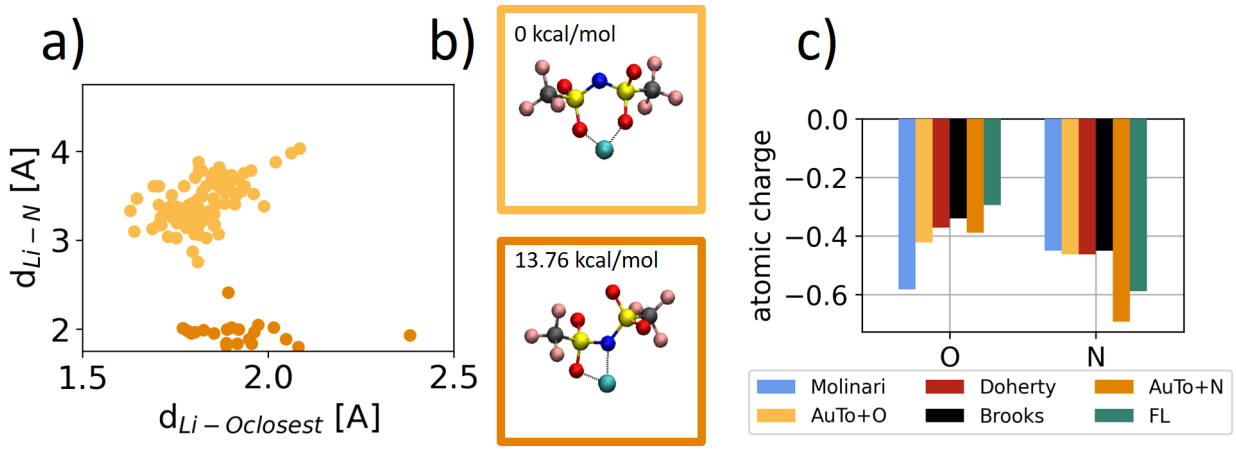


Figure 6: Distribution of Li-TFSI geometries from neural network force field molecular dynamics.[31] a) Partitioning of geometries into two clusters (gold and orange), plotted according to the Li-N distance and the distance between the Li and the closest oxygen in TFSI. b) Representative geometry for oxygen-coordinating mechanism (AuTo+O) and the nitrogen-coordinating mechanism (AuTo+N). DFT studies show the lowest-energy nitrogen-coordinating geometry is 13.76 kcal/mol higher in energy than the lowest oxygen-coordinating geometry. c) Plot of charge partitioned onto the nitrogen and oxygen in TFSI for the investigated interatomic potentials. Note that all charges have been scaled such that the TFSI anion has a net charge of 0.7.

These different coordination mechanisms lead to two distinct polarizations of TFSI. In oxygen-coordinating clusters (Li-N distance above 2.5 Å), the nitrogen and oxygen charges are close to equal, while in the nitrogen-coordinating clusters (Li-N distance below 2.5 Å), the lithium pulls negative charge towards the nitrogen, causing the nitrogen partial charge to become more negative as seen in Figure 6c. The oxygen to nitrogen charge ratios are 0.91 and 0.51 for the oxygen-coordinating (AuTo+O) and nitrogen-coordinating (AuTo+N) clusters, respectively. The seemingly different parameterizations of static atomic charge on TFSI

seen in literature can also be grouped into these two categories, with Molinari et al. (oxygen to nitrogen charge ratio of 1.29)[4] and Doherty et al. (0.80)[20] being oxygen-coordinating; France-Lanord et al. (0.50)[12] being nitrogen-coordinating; and Brooks et al. (0.76)[25] being in between both. The full TFSI partial charges used are shown in Supporting Information Section 6.

## Training

In class 2 potentials, the cross terms lead to a more convoluted optimization surface due to the cross terms. As such, the glyme parameters were trained independently. Details of the class 2 glyme training can be found in Supporting Information Section S1.4. These glyme parameters were then combined with TFSI covalent parameters reported by France-Lanord et al.[12]

The four compared IPs are a) FL, which is exactly the IP reported by France-Lanord et al.;[12] b) AuTo+FL, which uses parameters learned using AutoBADDIE for PEO and the IP for TFSI reported by France-Lanord et al; c) AuTo+N, which is the same as AuTo+FL but changes the TFSI point charge distribution to the average of the nitrogen-coordinating NNIPMD poses; d) AuTo+O, which is the same as AuTo+FL but changes the TFSI point charge distribution to the average of the oxygen-coordinating NNIPMD poses. The final force and energy parities relative to DFT for these four different PEO-LiTFSI parameterizations are shown in Figure 7.

Changing the PEO parameters from those of France-Lanord et al. to learned AutoBADDIE PEO parameters while maintaining the TFSI parameterization constant (Figure 7 (a) vs (b)) leads to a significant reduction in mean absolute error for both force and energy with respect to DFT. Interestingly, changing the partial charges between N-coordinating (Figure 7c) and O-coordinating (Figure 7d) polarizations has almost no effect on either force or energy mean absolute error with respect to DFT for the test set of NNIPMD geometries. The N- and O-coordinating IP show similar parity to DFT for poses generated through classical

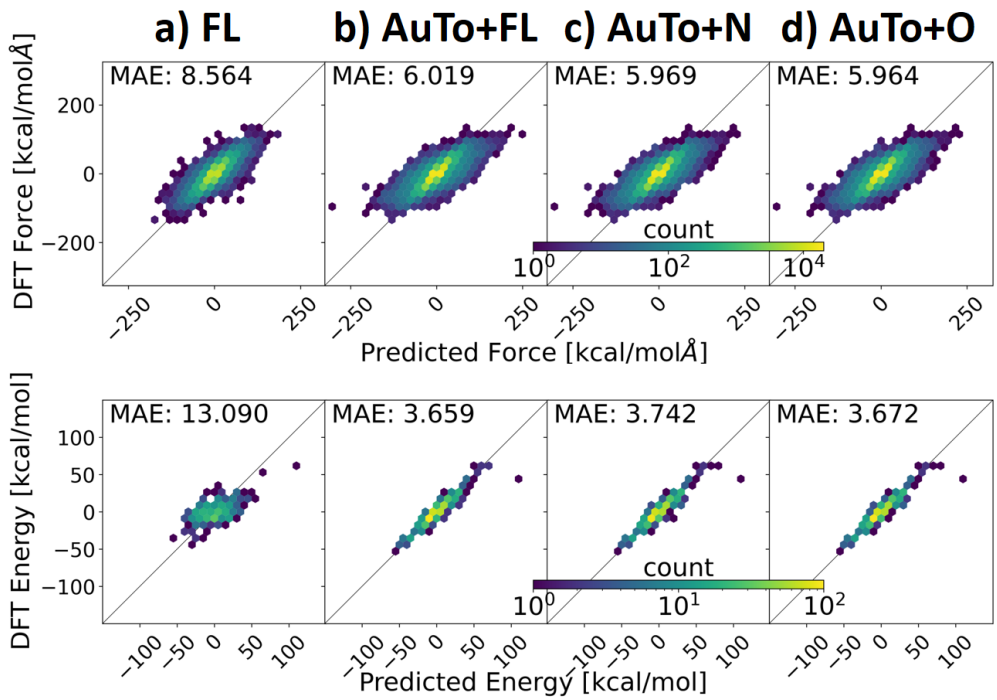


Figure 7: PEO force and energy parities with NNIPMD training data for a) France-Lanord et al.,[12] while the next three columns, b-d, all have AutoBADDIE-predicted glyme parameters with TFSI covalent interactions from France-Lanord et al.[12] The charge of TFSI are from b) France-Lanord et al.,[12] c) nitrogen-coordinated clusters, d) Oxygen-coordinated clusters.

IP MD as well, as seen in Supporting Information Section 4. Thus, force and energy parity with respect to DFT cannot be used to explain the performance of simulations using these IP parameterizations to compute solvation and kinetic properties, in agreement with prior literature.[32]

### Kinetic analysis

Similar to class 1 findings, the trends of conductivity and lithium transference number as a function of salt concentrations are different between all compared IPs in spite of the DFT force and energy MAEs being statistically indistinguishable for all simulations with AutoBADDIE-derived PEO parameters.

AuTo+FL interatomic potential has the lowest average mean absolute error compared to DFT forces and energies across all geometry sets yet shows a monotonically increasing con-

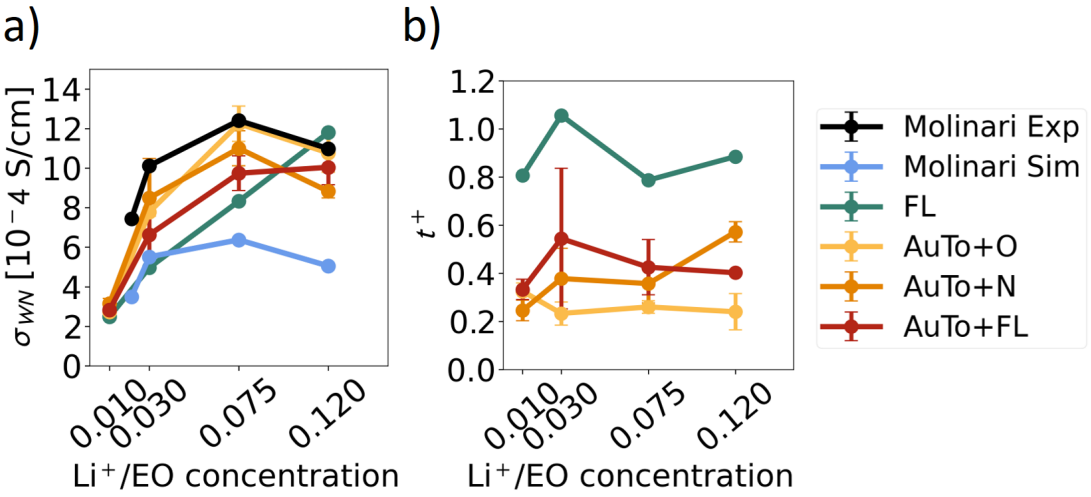


Figure 8: a) Correlation-based conductivities as a function of lithium salt to PEO monomer concentration reported by Molinari et al.[4] (blue) and simulated using different class 2 interatomic potentials. b) Lithium transference number using different class 2 interatomic potentials. Molinari et al. does not report their full IP or their derived lithium transference and therefore cannot be plotted on b.

.

ductivity with increasing lithium salt concentration, as seen in Figure 8a. On the other hand, simulations using AutoBADDIE-derived PEO parameters and DFT-derived TFSI point charges, AuTo+O and AuTo+N, both show a maximum of ionic conductivity at r=0.075 Li$^+$/EO, as does the Molinari IP. The published France-Lanord et al. parameters have the worst force and energy parity with respect to DFT and show monotonically increasing conductivity with respect to ionic salt concentration.

AuTo+O (yellow in Figure 8b) is the only studied IP that shows an overall decreasing lithium transference with increasing lithium salt concentration as expected from experiments,[1] while all nitrogen-coordinating IP (FL, AuTo+N, AuTo+FL) show increasing transference with increasing lithium salt concentration. Molinari et al. did not report their calculated transference numbers and therefore are not included in the comparison.

### Solvation analysis of class 2 simulations

Analysis on the first solvation shell of lithium ions under the different IPs elucidates further differences due to the charge distribution of TFSI. First, lithium under the France-Lanord parameterization has an unstructured interaction with TFSI as seen in Figure 9a and also has a lower overall oxygen coordination number relative to the simulations using AutoBADDIE-derived PEO parameters as seen in Figure 9e. Additionally, the amount that lithium ions coordinate to TFSI remains relatively constant as salt concentration increases in the FL simulations.

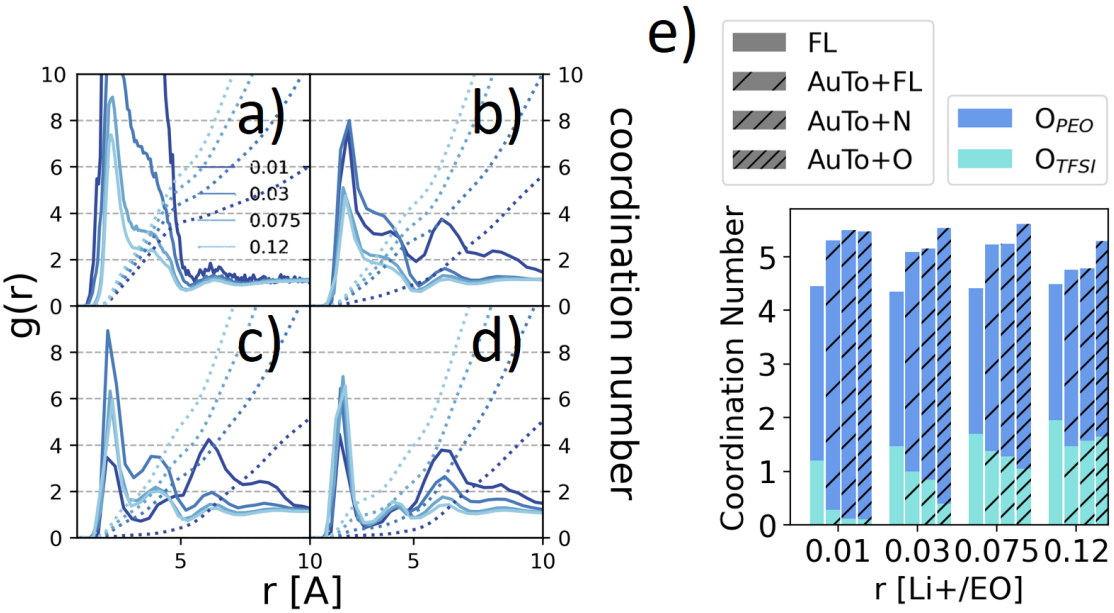


Figure 9: Lithium solvation analysis for lithium solid polymer electrolyte using different class 2 interatomic potentials. Lithium-TFSI (oxygen) radial distribution functions for a) France-Lanord, b) AuTo+FL, c) AuTo+N, d) AuTo+O interatomic potentials, respectively. e) Lithium coordination number for France-Lanord et al[12] (solid), then AutoBADDIE-predicted parameters for PEO with different TFSI charges (hashed). Dark blue refers to the coordination number of Li to oxygens belonging to PEO oxygens while light blue refers to the lithium coordination to oxygens in TFSI. The total height of each bar indicates the total lithium oxygen coordination number for a cutoff of 3.4 Å.

AutoBADDIE-derived PEO parameters result in a more structured Li-TFSI coordination as compared to those from France-Lanord et al., as seen by the tighter peaks of plot (9b)

compared to plot (9a). From the France-Lanord charges (9b, and single hash), to the DFT nitrogen-coordinating (9c, and double hash) and the DFT oxygen-coordinating charges (9d, and triple hash), the Li-TFSI coordination is further increased. The higher coordination of Li to both PEO and TFSI for AuTo+O relative to the other IP could explain why AuTo+O is the only IP to show the overall trend of decreasing transference number as a function of salt concentration.

Molinari et al. report a total lithium-oxygen coordination number of 6-7 in their simulations, which is higher than all other interatomic potentials. Molinari et al. use an ion charge of $0.75|e|$ while all others use a $0.7|e|$ charge, so the higher Coulombic attraction between the lithium and oyxgens might explain their reported lithium-oxygen coordination number of 6-7.[4]

Polarizable simulations conducted by Borodin et al.[33,34] at Li/EO=0.133 at 393K (which most closely matches the rightmost bars in Figure 9 e)) report a total oxygen coordination of 4.6, matching the total coordination AutoBADDIE-predicted PEO parameters and nitrogen-coordinating TFSI (double-hashed bars). The experimental comparison from Borodin et al.[33] at r=0.133 showed a total Li-O coordination number of 4.9, in good agreement with simulations using predicted PEO parameters and oxygen-coordinating TFSI (AuTo+O, as seen by the triple-hashed bar) at the highest simulated salt concentration (r=0.12).

# Conclusions

We showed that AutoBADDIE is able to predict interatomic parameters that can drive MD simulations with performance matching those of parameterizations developed over decades of hand-tuning and scientific research such as OPLS and PCFF+. Across all systems investigated, including for those identical aside from different charge distributions on the TFSI, neither the force nor energy parity with respect to DFT were seen to be solely descriptive of an interatomic potential parameterization's ability to provide solvation or kinetic properties

that match experimental measurements. The correct trend in conductivity across different salt concentrations was seen for the nitrogen-coordinating TFSI charges from DFT and the oxygen-coordinating TFSI charge distributions from DFT and by Molinari et al.[4]

This raises questions on to what extent we can trust solvation analysis based on these static point charge-based classical simulations and warrants further study into what are the expected coordination mechanisms of lithium TFSI in polymer electrolyte systems. Additionally, the diffusivity and conductivity predictions of polarizable simulations are still closer to experimental than any of the nonpolarizable simulations explored.[33] Further investigation into the charge states of TFSI and Li during polarizable simulations could shed light onto these discrepancies. In this paper, we demonstrated AutoBADDIE can successfully parameterize nonpolarizable models. However this same workflow could be extended for polarizable interatomic potentials which could provide better agreement with experiments.

## Acknowledgement

The authors would like to acknowledge the MIT Engaging and MIT Lincoln Lab Supercloud clusters. P.L. and J.R. acknowledge support from Toyota Research Institute. P.L. and A.S. also acknowledge support from the National Defense Science and Engineering Graduate Fellowship.

We also gratefully acknowledge useful discussions with Nicola Molinari.

## Supporting Information Available

The computational model and all learned interatomic potential parameters reported in this work are available under the MIT license at https://github.com/learningmatter-mit/AutoBADDIE.

The following files are available free of charge.

- Supporting Info: Additional training considerations, dihedral regularization conforma-

tion exploration, molecular dynamics simulation details, class 2 interatomic potential performance for different sets of geometries, class 2 diffusivities, all TFSI point charges, carbonate additional information and radial distribution functions.

- Liquid carbonate benchmark results: Training, kinetic property, and solvation benchmarking results for AutoBADDIE-derived liquid carbonate interatomic potential parameters.

# References


(1) Pesko, D. M.; Timachova, K.; Bhattacharya, R.; Smith, M. C.; Villaluenga, I.; Newman, J.; Balsara, N. P. Negative Transference Numbers in Poly(ethylene oxide)-Based Electrolytes. *Journal of The Electrochemical Society* **2017**, *164*, E3569–E3575.

(2) Choo, Y.; Halat, D. M.; Villaluenga, I.; Timachova, K.; Balsara, N. P. Diffusion and migration in polymer electrolytes. *Progress in Polymer Science* **2020**, *103*, 101220.

(3) Chen, X. C.; Sacci, R. L.; Osti, N. C.; Tyagi, M.; Wang, Y.; Keum, J. K.; Dudney, N. J. Study of the Segmental Dynamics and Ion Transport of Solid Polymer Electrolytes in the Semi-crystalline State. *Frontiers in Chemistry* **2020**, *8*, Publisher: Frontiers Media SA.

(4) Molinari, N.; Mailoa, J. P.; Kozinsky, B. Effect of Salt Concentration on Ion Clustering and Transport in Polymer Solid Electrolytes: A Molecular Dynamics Study of PEO-LiTFSI. *Chemistry of Materials* **2018**, *30*, 6298–6306, Publisher: American Chemical Society.

(5) France-Lanord, A.; Grossman, J. C. Correlations from Ion Pairing and the Nernst-Einstein Equation. *Physical Review Letters* **2019**, *122*, 136001, arXiv: 1812.04772 Publisher: American Physical Society.

(6) Gong, S.; Zhang, Y.; Mu, Z.; Pu, Z.; Wang, H.; Yu, Z.; Chen, M.; Zheng, T.; Wang, Z.; Chen, L.; Wu, X.; Shi, S.; Gao, W.; Yan, W.; Xiang, L. BAMBOO: a predictive and transferable machine learning force field framework for liquid electrolyte development. 2024; `http://arxiv.org/abs/2404.07181`, arXiv:2404.07181 [cond-mat, physics:physics].

(7) Winter, G.; Gómez-Bombarelli, R. Simulations with machine learning potentials identify the ion conduction mechanism mediating non-Arrhenius behavior in LGPS. *Journal of Physics: Energy* **2023**, *5*, 024004, Publisher: IOP Publishing.

(8) Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G. A.; Dahlgren, M. K.; Russell, E.; Von Bargen, C. D.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *Journal of Chemical Theory and Computation* **2021**, *17*, 4291–4300, Publisher: American Chemical Society.

(9) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* **1996**, *118*, 11225–11236, Publisher: American Chemical Society.

(10) Sun, H.; Mumby, S. J.; Maple, J. R.; Hagler, A. T. An ab Initio CFF93 All-Atom Force Field for Polycarbonates. *Journal of the American Chemical Society* **1994**, *116*, 2978–2987, Publisher: American Chemical Society.

(11) Sun, H. COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase ApplicationsOverview with Details on Alkane and Benzene Compounds. *Journal of Physical Chemistry B* **1998**, *102*, 7338–7364, Publisher: American Chemical Society.

(12) France-Lanord, A.; Wang, Y.; Xie, T.; Johnson, J. A.; Shao-Horn, Y.; Grossman, J. C. Effect of Chemical Variations in the Structure of Poly(ethylene oxide)-Based Polymers

on Lithium Transport in Concentrated Electrolytes. *Chemistry of Materials* **2020**, *32*, 121–126.

(13) Wang, W.; Wu, Z.; Gómez-Bombarelli, R. Learning Pair Potentials using Differentiable Simulations. **2022**, arXiv: 2209.07679.

(14) Takaba, K.; Pulido, I.; Behara, P. K.; Cavender, C. E.; Friedman, A. J.; Henry, M. M.; Opeskin, H. M.; Iacovella, C. R.; Nagle, A. M.; Payne, A. M.; Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Wang, Y. Machine-learned molecular mechanics force field for the simulation of protein-ligand systems and beyond. 2023; `https://arxiv.org/abs/2307.07085v4`.

(15) Horton, J. T.; Boothroyd, S.; Wagner, J.; Mitchell, J. A.; Gokey, T.; Dotson, D. L.; Behara, P. K.; Ramaswamy, V. K.; Mackey, M.; Chodera, J. D.; Anwar, J.; Mobley, D. L.; Cole, D. J. Open Force Field BespokeFit: Automating Bespoke Torsion Parametrization at Scale. *Journal of Chemical Information and Modeling* **2022**, *62*, 5622–5633, Publisher: American Chemical Society.

(16) Thürlemann, M.; Böselt, L.; Riniker, S. Regularized by Physics: Graph Neural Network Parametrized Potentials for the Description of Intermolecular Interactions. *Journal of Chemical Theory and Computation* **2023**, *19*, 562–579, Publisher: American Chemical Society.

(17) Zheng, T.; Wang, A.; Han, X.; Xia, Y.; Xu, X.; Zhan, J.; Liu, Y.; Chen, Y.; Wang, Z.; Wu, X.; Gong, S.; Yan, W. Data-Driven Parametrization of Molecular Mechanics Force Fields for Expansive Chemical Space Coverage. 2024; `http://arxiv.org/abs/2408.12817`, arXiv:2408.12817 [physics].

(18) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *The Journal of Physical Chemistry Letters* **2014**, *5*, 1885–1891, Publisher: American Chemical Society.

(19) Dodda, L. S.; CabezaădeăVaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Research* **2017**, *45*, W331–W336.

(20) Doherty, B.; Zhong, X.; Gathiaka, S.; Li, B.; Acevedo, O. Revisiting OPLS Force Field Parameters for Ionic Liquid Simulations. **2017**,

(21) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. 2017; `http://arxiv.org/abs/1412.6980`, arXiv:1412.6980 [cs].

(22) Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry* **1990**, *11*, 361–373, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.540110311.

(23) Hou, T.; Fong, K. D.; Wang, J.; Persson, K. A. The solvation structure, transport properties and reduction behavior of carbonate-based electrolytes of lithium-ion batteries. *Chemical Science* **2021**, *12*, 14740–14751, Publisher: The Royal Society of Chemistry.

(24) Dodda, L. S.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *The Journal of Physical Chemistry B* **2017**, *121*, 3864–3870, Publisher: American Chemical Society.

(25) Brooks, D. J.; Merinov, B. V.; Goddard, W. A.; Kozinsky, B.; Mailoa, J. Atomistic Description of Ionic Diffusion in PEO-LiTFSI: Effect of Temperature, Molecular Weight, and Ionic Concentration. *Macromolecules* **2018**, *51*, 8987–8995, Publisher: American Chemical Society.

(26) Banks, J. L. et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *Journal of Computational Chemistry* **2005**, *26*, 1752–1780, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.20292.

(27) Wang, W.; Yang, T.; H.ăHarris, W.; Gómez-Bombarelli, R. Active learning and neural network potentials accelerate molecular screening of ether-based solvate ionic liquids. *Chemical Communications* **2020**, *56*, 8920–8923, Publisher: Royal Society of Chemistry.

(28) Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.

(29) Schubert, E.; Sander, J.; Ester, M.; Kriegel, H. P.; Xu, X. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Transactions on Database Systems* **2017**, *42*, 19:1–19:21.

(30) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D. Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*

(31) Wang, Y.; Fass, J.; Chodera, J. D. End-to-end differentiable molecular mechanics force field construction. *arXiv* **2020**, arXiv: 2010.01196.

(32) Fu, X.; Wu, Z.; Wang, W.; Xie, T.; Keten, S.; Gomez-Bombarelli, R.; Jaakkola, T. S. Forces are not Enough: Benchmark and Critical Evaluation for Machine Learning Force Fields with Molecular Simulations. *Transactions on Machine Learning Research* **2023**,

(33) Borodin, O.; Smith, G. D. Mechanism of Ion Transport in Amorphous Poly(ethylene oxide)/LiTFSI from Molecular Dynamics Simulations. *Macromolecules* **2006**, *39*, 1620–1629, Publisher: American Chemical Society.

(34) Borodin, O.; Smith, G. D. Li+Transport Mechanism in Oligo(Ethylene Oxide)s Compared to Carbonates. *Journal of Solution Chemistry* **2007**, *36*, 803–813.