From Multiple Scale Modeling to Multiscale-Modeling

Fabian Årén

Department of Physics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021
From Multiple Scale Modeling to Multiscale-Modeling
Fabian Årén

© Fabian Årén, 2021.

Department of Physics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
A deep learning rendering of first a schematic 1st solvation shell, then a mesoscopic structure consisting of many solvation shells, and finally a bulk electrolyte.

Chalmers Digitaltryck
Gothenburg, Sweden 2021
Abstract

To power a sustainable future, interest in battery research and technology is at an all time high. In order to enable a transition to green tech many applications, such as the automotive industry, is in need of higher power densities, energy densities, longer life-times, and safer batteries.

One crucial component of batteries is the electrolyte, which for lithium-ion batteries (LIBs) have not developed as much as one would expect since its introduction in the 1990s. Through the use of novel electrolyte concepts such as highly concentrated electrolytes (HCE) and localized highly concentrated electrolytes (LHCE) desired qualities such as an increased energy density could be achieved. The effects of local properties on macroscopic behaviour within these systems are much more striking than conventional LIB electrolytes, constraining the use of common simulation techniques used in battery research.

This thesis studies these novel electrolyte concepts using an array of different computational methods, such as DFT, AIMD, and classical MD. Based on these techniques, as well as on the CHAMPION method, the work done in this thesis attempts to develop a method for tying together understanding of materials physics at the different scales represented by AIMD and classical MD through force sampling. This force sampling is presented as an alternative to commonplace MD force fields such as AMBER, CHARMM and GROMACS.

Finding the local structure important for explaining global transport phenomenon by showing that local HCE structure is retained when going from HCE to LHCE as well as showing the possibility for these new types of FFs, even though more work is needed on the accuracy of these FFs.

Keywords: lithium-ion batteries, machine learning, multi-scale modeling, electrolytes, DFT, AIMD, MD
# Table of Contents

List of Tables and Figures ......................................................... ix
List of Publications ...................................................................... xi
List of Abbreviations .................................................................... xiii

1 Introduction .................................................................................. 1

2 Batteries .................................................................................... 3
   2.1 Electrochemical cells ................................................................. 3
       2.1.1 Li-ion Batteries ................................................................. 3
       2.1.2 Solid-Electrolyte Interphase .............................................. 5
       2.1.3 Calcium Batteries .............................................................. 6
   2.2 Battery Properties ................................................................. 6

3 Electrolytes .................................................................................. 9
   3.1 Highly Concentrated Electrolytes .............................................. 9
   3.2 Localised Highly Concentrated Electrolytes .............................. 10
   3.3 Ca$^{2+}$ conducting electrolytes ............................................... 11
   3.4 Ion Transport in Liquid Electrolytes ......................................... 11

4 Methods & Modelling of Battery Electrolytes ................................. 13
   4.1 DFT .................................................................................... 13
   4.2 AIMD ................................................................................ 15
   4.3 Classical MD ..................................................................... 17
   4.4 Machine Learning Methods for Electrolyte Simulations .......... 18
       4.4.1 Gaussian Process Regression ........................................... 19
   4.5 CHAMPION ..................................................................... 19
   4.6 Automated Force Field Finder ............................................... 20

5 Results & Discussion .................................................................. 23
   5.1 Ca$^{2+}$ 1st Solvation Shells ................................................... 23
   5.2 Local to Global Structure ....................................................... 25
   5.3 Generated Force Fields .......................................................... 26

6 Conclusion & Outlook .................................................................. 31

References .................................................................................... 35
Table of Contents
## List of Tables and Figures

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Schematic of a typical Li-ion battery cell with a Cu current collector, a graphite anode, a 1 M LiPF$_6$ in EC/DMC electrolyte, a LCO cathode, and a Al current collector.</td>
<td>4</td>
</tr>
<tr>
<td>2.2</td>
<td>Schematic of an ESW.</td>
<td>5</td>
</tr>
<tr>
<td>2.3</td>
<td>Schematic of an SEI (▲.■.●) forming on the anode, protecting it from direct contact with the electrolyte whilst still allowing ionic transport.</td>
<td>6</td>
</tr>
<tr>
<td>3.1</td>
<td>Different charge transfer mechanisms at intermediate and high salt concentrations.</td>
<td>10</td>
</tr>
<tr>
<td>4.1</td>
<td>Multiple scale modeling vs multi-scale modeling.</td>
<td>13</td>
</tr>
<tr>
<td>4.2</td>
<td>Schematic of bonded and non-bonded parameters making up a classical FF.</td>
<td>18</td>
</tr>
<tr>
<td>4.3</td>
<td>Schematic Ávall plot linking the probability distribution of generalised forces to a generalised coordinate describing a interaction. Orange: Occupied bins. Blue: Empty bins.</td>
<td>20</td>
</tr>
<tr>
<td>5.1</td>
<td>Use of DFT to elucidate the local structure in bulk electrolytes.</td>
<td>24</td>
</tr>
<tr>
<td>5.2</td>
<td>The most common topologies around a Li cation in order of probability.</td>
<td>25</td>
</tr>
<tr>
<td>5.3</td>
<td>Snapshots of the periodic simulation cell, highlighting the percolating network in a sea of solvent.</td>
<td>26</td>
</tr>
<tr>
<td>5.4</td>
<td>Comparing the quality of the generalised Ávall FF depend on number of bins.</td>
<td>27</td>
</tr>
<tr>
<td>5.5</td>
<td>The form of the Ávall FF after using the raw data as input to a Gaussian Process regression. The same physics is captured when studying the 1$^{st}$ half of a simulation as when studying the 2$^{nd}$ half.</td>
<td>28</td>
</tr>
<tr>
<td>5.6</td>
<td>The statistical inefficiency $s$ dependence on $\Delta t$, the number of time steps $N_{timesteps}$, and the graph radius $d$.</td>
<td>29</td>
</tr>
<tr>
<td>5.7</td>
<td>The form of the Ávall FF after using the raw data as input to a Gaussian Process regression. Data with a statistical inefficiency $s = 21.1608$.</td>
<td>29</td>
</tr>
</tbody>
</table>
List of Publications

This thesis is based on the following Publications [1, 2, 3]. In the text they are referred to by their Roman numeral:

I. P.-A. Martin, F. Årén, and P. Johansson. Localized Highly Concentrated Electrolytes for Calcium Batteries. In manuscript


III. F. Årén, R. Andersson, and P. Johansson. Learning Force Fields from *Ab initio* Data. In manuscript
List of Abbreviations

ACN       Acetonitrile
AIMD      *ab initio* Molecular Dynamics
AM        Active Material
BOMD      Born-Oppenheimer Molecular Dynamics
CEI       Cathode Electrolyte Interphase
CN        Coordination Number
DFT       Density Functional Theory
DOD       Depth of Discharge
DSD       Dynamic Structure Discovery
ESW       Electrochemical Stability Window
FF        Force Field
GP        Gaussian Process
HCE       Highly Concentrated Electrolyte
HOMO      Highest Occupied Molecular Orbital
KPI       Key Performance Indicator
LCO       LiCoO$_2$
LE        Liquid Electrolyte
LHCE      Localized Highly Concentrated Electrolyte
LIB       Lithium-Ion Battery
LiTFSI    Lithium bis(trifluoromethanesulfonyl)imide
LP30      1 M LiPF$_6$ in EC/DMC
LUMO      Lowest Unoccupied Molecular Orbital
MD        Molecular Dynamics
ML        Machine Learning
NN        Neural Network
SCF       Self-Consistent Field
SEI       Solid-Electrolyte Interphase
SF        Symmetry Function
SHE       Standard Hydrogen Electrode
SN        Solvation Number
SOAP      Smooth Overlap of Atomic Position
SOC       State of Charge
List of Abbreviations
1
Introduction

Materials are all around us. From the components in the electrical device you might be reading this text on to the paper of the printed version. Our modern world is built on the back of materials, and is propelled by humanity’s increased understanding of the mechanisms and structures that give rise to all the unique and useful properties of different materials enabling the level of civilisation we have today.

Within the foreseeable future one of the most important research fields within materials science is the development of energy storage materials, such as the materials used to make batteries, which is a keystone in solving the climate issues [4] the human species will be facing within the coming century. Currently the state-of-the-art battery technology, the lithium-ion battery (LIB) has come a long way since its introduction in the 1990’s by Sony [5]. Having an energy density of 200 Wh/l or 80 Wh/kg enabled the widespread adoption of handheld electronics. Today the energy densities have more than tripled.

One of the major problems of battery material research is the trial and error process used to discover new materials. Looking at trends in other fields [6, 7], it is clear that computational approaches have a great potential to be a solution. However, material simulations at the molecular level are difficult if directly trying to solve the Schrödinger equation, as it is a complex many body problem. One of the most effective techniques to simulate systems at this scale is density functional theory (DFT) [8]. Being able to simulate molecules accurately to achieve understanding quickly breaks down as the system size increases due to DFT lacking the ability to simulate dynamics, which becomes more important for explaining properties emerging at larger system sizes. Hence a conundrum appears; how to describe and simulate systems of any composition at a macroscopic scale with quantum accuracy. Paper I used DFT to identify the origins of spectroscopic observations and showcase the value of this method on a local scale.

Studying systems larger than a few molecules however requires a different tool set. Molecular dynamics (MD) simulations have been a functional tool since the 1950’s [9, 10] and the method was the grounds for the 2013 Nobel Prize in Chemistry [11]. However, even though MD simulations have been shown numerous times to aid materials scientists in their work [12, 13, 14], the method has several drawbacks, such as the method used to evaluate forces at each time step being non-trivial, as well as computationally expensive long range interactions. The umbrella term MD should be separated into two distinct methods, ab initio MD (AIMD), and classical
1. Introduction

MD (from now on simply referred to as MD), both with their own distinct problems and benefits. Paper II display a proof-of-concept for a new method of analysing MD trajectories. Specifically AIMD is used to propagate a system, generating realistic structures, showcasing what AIMD in particular, and MD in general can be used for when studying batteries.

In order to enable MD simulations of non-conventional electrolytes on a nano-scale, a system for generating accurate force fields (FF’s) have to be developed. Many attempts to circumvent this need have been made within the field of materials science in general, quite successfully, using a number of machine learning (ML) approaches, mainly using neural network (NN) approaches [15, 16, 17]. Even though ML techniques have gigantic upsides when it comes to computing hard-to-model problems there are certain drawbacks, such as the large data-sets needed for training, and the long times needed for learning, as well as the lack of understanding of why a model makes the predictions it does, colloquially called a "black box" [18]. Preferably a method to easily generate accurate FFs on the go for specific systems, without the drawbacks of conventional ML techniques, is needed. In paper III such a method is suggested through generalising the methods developed by Ávall and Johansson to extract effective pair-wise forces and interaction energies from AIMD [19].

This thesis first shows how DFT and AIMD can be used to model electrolytes, at multiple scales, but foremost it proposes a novel method to generate FFs. All in order to study the rapidly changing landscape of modern battery electrolytes, in order to obtain properties such as transport numbers, current density, etc., or to understand the structure of these materials, which can enable complementary and comparative information to experimental insights. The background for this thesis comes from studies of electrolytes using conventional computational methods, and discovering a gap in our ability to simulate complex systems at nano-scale, motivating the move from multiple scales to multi-scale modeling.
2

Batteries

2.1 Electrochemical cells

Electrochemical cells store chemical energy, which can be released as electrical energy through redox reactions. There are some miscommunication regarding the nomenclature of what a battery is, where battery can be used both to reference the electrochemical cell, as well as a battery pack. However in this thesis the word "battery" will refer to a battery pack, an assembly of cells and controlling electronics, as well as housing, fully capable of storing energy for large scale applications and in electric vehicles. Battery cells on the other hand are small single unit electrochemical cells, containing the reaction.

The cell is made up of four main components; 1) the anode, where oxidation occurs during discharge and electrons flow from, 2) the cathode, where reduction occurs during discharge, and electrons flow towards 3) the electrolyte, which is the charge carrying medium between the two electrodes, and 4) a separator which physically separates the electrodes to prevent a short circuit (Fig. 2.1). In addition to these components there are also current collectors, which are materials with high electronic conductivity, e.g. metal foil, connecting the cell to an external circuit, as well as some sort of housing.

2.1.1 Li-ion Batteries

The most advanced battery chemistry in common use today are based on Li-ion technology. Li-ion batteries are popular due to both having high specific power and energy ($\sim 300$ Wkg$^{-1}$ and $\sim 150$ Whkg$^{-1}$ respectively at cell level for electric vehicle applications [20]) as well as a high voltage of $\sim 3$ V against a standard hydrogen electrode (SHE) which is three times that of a typical Ni-Cd battery [21].

One of the most common traditional Li-ion battery chemistries is graphite and LiCoO$_2$ (LCO) used together with the LiPF$_6$ in EC/DMC (LP30) electrolyte. To understand contemporary battery technology this system is a good starting point. The ideal discharge reaction would be:

- Oxidation at anode: \[ \text{LiC}_6 \rightarrow \text{C}_6 + \text{Li}^+ + \text{e}^- \]
- Reduction at cathode: \[ \text{Li}^+ + \text{e}^- + \text{CoO}_2 \rightarrow \text{LiCoO}_2 \]
- Total reaction: \[ \text{LiC}_6 + \text{CoO}_2 \rightarrow \text{C}_6 + \text{LiCoO}_2 \]
2. Batteries

**Figure 2.1:** Schematic of a typical Li-ion battery cell with a Cu current collector, a graphite anode, a 1 M LiPF$_6$ in EC/DMC electrolyte, a LCO cathode, and a Al current collector.

By physically separating the oxidation and reduction reaction and connecting them through an external circuit, cycling can be performed by transferring the electrons outside the cell, and the Li$^+$ through an electronically insulating medium, the electrolyte. From these reactions it is clear that the limiting factor of the amount of Li-ions that can be transported, and hence the energy content of the cell, is the Li$^+$, or an alternative charge carrier, storage capacity of both electrodes. This amount is given by the capacity $Q$, although since $Q$ can be increased by simply increasing the electrode mass, which is not useful in most applications a more useful property to keep track of is the specific capacity $C$. Using Faraday’s law the theoretical specific capacity, which puts an upper limit on the actual capacity can be calculated:

$$C_{th} = \frac{nF_0}{M_W}$$

where $n$ is the number of electrons transferred during the reduction or oxidation, $F_0$ is Faraday’s constant, and $M_W$ is the molecular weight of the active material (AM).

The theoretical voltage $V_{th}$ is determined by the redox reactions taking place at the electrode. Given knowledge about the free energy of of the redox reaction, the theoretical voltage can be calculated as:

$$V_{th} = -\frac{\Delta G}{nF_0} = -\frac{\Delta G}{C_{th}M_W}$$

$V_{th}$ gives an upper limit to the voltage. The actual cell voltage is given by, $V = E^+ - E^-$ where $E^+$ is the electrochemical potential at the positive electrode, and $E^-$ is the electrochemical potential at the negative electrode, compared to a reference
2. Batteries

Figure 2.2: Schematic of an ESW.

Electrode. Knowing the voltage at a charge \( Q \), \( V(Q) \) allows to calculate the electrochemical energy of the cell, given by:

\[
E_{\text{chem}} = \int_{0}^{C} V(Q) dQ
\]

where \( C \) is the full capacity of the cell, the total amount of charge that can be reversibly transferred [22]. However when comparing batteries the more useful quantities are the volumetric and gravimetric energy density, which is calculated by dividing the total electrochemical energy by the volume, or the mass of the battery pack, the battery cell, or the AM, meaning batteries can be compared at three different levels.

2.1.2 Solid-Electrolyte Interphase

In order to keep a cell cycling the operating voltage needs to be less than the electrochemical stability window (ESW) (Fig. 2.2) of the electrolyte [23]. The ESW is defined as the potential difference between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). However battery systems such as LIBs operate with a voltage window outside the ESW. These systems continue operating through the decomposition of the electrolyte forming a passivating, yet still ionically conductive film on the anode (Fig. 2.3), called a solid-electrolyte interphase (SEI). The SEI is typically formed during the first couple of cycles on the anode by the reduction of the electrolyte. The oxidation products form the analogous cathode electrolyte interphase (CEI) on the cathode.
2. Batteries

Figure 2.3: Schematic of an SEI (▲, □, ●) forming on the anode, protecting it from direct contact with the electrolyte whilst still allowing ionic transport.

2.1.3 Calcium Batteries

State-of-the-art battery technologies, aka Li-ion based battery technologies, utilise a monovalent charge carrier. However this is not a fundamental limitation for a working battery concept. In recent years the interest for multivalent battery technologies, and especially calcium (Ca) batteries have risen [24, 25, 26]. The ability to drive multiple electron exchanges for every carrier transfer puts technologies such as these at comparable theoretical energy densities to Li-ion batteries [24], even though Ca\(^{2+}\) is a much larger ion than Li\(^{+}\). This makes Ca-metal batteries a viable option for future more sustainable and cheaper batteries. The Ca metallic anode has a low redox potential of \(-2.87\) V vs a SHE, meaning that all components in the electrolyte are at risk of reduction, forming different passivating layers. The problem hence is finding a solution allowing for both plating and stripping of Ca\(^{2+}\).

Similarly cathode materials have to be developed also. A cathode needs to allow for high ionic mobility of the intercalating species. This can be achieved through using a host lattice where the diffusing ion is not coordinated to the lattice in its most thermodynamically stable configuration, as well as using a host structure with minor changes in ion coordination along the diffusion pathway [25]. Even though reversible Ca electrodeposition was demonstrated in 2016 [27] there are still a few hurdles to overcome before Ca batteries will become ready for market, e.g. the need to adapt the manufacturing process to be able to handle producing both Ca metal anodes and cells in an inert atmosphere, alternatively develop pre-passivated anodes.

2.2 Battery Properties

There are certain key performance indicators (KPI) important for comparing systems to keep track of when studying batteries. Important to all applications is knowing the energy content or battery capacity of the pack, as well as power, and the nominal cell voltage. Generally each unique battery system is designed to either optimise for
power or energy content depending on the intended use. Hence knowing a battery’s power-to-energy-ratio can indicate what a battery system can be used for, however other aspects also play a role. The nominal cell voltage is the average voltage output of the cell, which drops as the system discharges, for current LIBs this can drop from 4.6V to 2.5V [28].

Hence to understand the live operation of a system the state-of-charge (SOC) defined as

$$\text{SOC}[%] = \frac{Q}{Q_{\text{max}}}$$

is needed, where the SOC represents the available charge in the system. Since the SOC tend to drop when cycling, this property is important for understanding the deterioration of the system. To understand the health of a battery cell the number of cycles also have to be known. To make sense of the cycling term it is easier to talk about equivalent full cycles. However this measurement lacks the ability to distinguish between one cycle of 100% of depth of discharge (DOD), two cycles of 50% of DOD, or ten cycles of 10% DOD. DOD is one of the largest contributors of battery degradation, determining the lifetime of the system. For example, Li-ion batteries undergo ten times more degradation when operated near 100% DOD compared to when operated at 10% DOD.

An almost equally important property for a cell is the Coulombic efficiency that quantifies the share of charges that are returned by the storage system, defined by:

$$\eta_{\text{Coulombic}} = \frac{\int_{\text{period}} I(t_{\text{discharge}})dt}{\int_{\text{period}} I(t_{\text{charge}})dt}$$

The Coulombic efficiency tends to decline when the battery is cycled, being one of the main reasons why batteries have a limited life span. One reason for such a degradation can be a parasitic reduction of electrode and electrolyte [29].

By studying materials at the cell level we can hope to develop systems that optimise these KPIs; e.g. increased energy and power content, as well as slower degradation. These indicators are tightly knit to the system dynamics, studied in paper I - III. The voltage of the system on the other hand is heavily dependent on the choice of chemistry, making the studies of different mono, and multivalent systems interesting.
2. Batteries
3

Electrolytes

In a LIB the electrolyte, *e.g.* LP30, consists of solvents, and a solute, with the main purpose to transport charge carriers from one electrode to the other whilst remaining electronically insulating. Beyond ion transport the electrolyte has to adhere to many other criteria, such as the ability to wet the separator and electrodes, having a low flammability, low toxicity, and low environmental impact. In order to understand the charge transport in electrolytes it is important to study the mechanisms that facilitate transport, and what structures enable transport. The net transport of charges is driven by the difference in electrochemical potential between the electrodes. Even though this is the driving force on a macroscopic scale it is important to keep in mind the importance of both time and length scale. Herein we will discuss the importance of understanding the electrolyte structure at different scales in order to explain behaviour.

3.1 Highly Concentrated Electrolytes

Conventional electrolytes in modern batteries, such as the aforementioned LP30, are most often found at concentrations around 1 M. However since the early 2010’s the interest for much higher salt concentrations in battery electrolytes have increased dramatically. The increased interest is due to some interesting properties displayed such as a lower solubility of transition metals dissolving from cathodes [30], higher rate capabilities [31, 32], and a widened ESW [31, 33]. The addition of more salt causes the electrolyte to have high ion density more akin to solvent in salt, than salt in solvent, making the electrolyte take on behaviour similar to ionic liquids. This also means that there is a low amount of free solvent in the liquid, which is one of the major identifiers of highly concentrated electrolytes (HCEs). Furthermore this leads to a higher density, higher viscosity, and a lower total ionic conductivity. Using a combination of experimental techniques Nilson et al. have shown that ionic conductivity depends more strongly on ionicity than viscosity [34], which is in agreement with the argument by Seo et al. [35]. In paper II we describe the mechanism behind the lower ionic conductivity for lithium bis(trifluoromethanesulfonyl)imide (LiTFSI) in acetonitrile (ACN) at a 1 : 2 molar ratio and show that a percolating network form from the anions and cations at such a high salt concentration whilst the solvent remains free. The lower ionic conductivity is explained by a decreased amount of transportation being conducted through a vehicular mechanism, which is defined as a cation in a shell of coordinated solvent molecules moving freely through the electrolyte (Fig. 3.1a), contrary to a
3. Electrolytes

(a) How a solvent molecule shell forms a "vehicle" surrounding the cation and enabling transport through the solvent

(b) Rotating a structure containing cations can lead to a net charge transfer. This is an example of non-vehicular charge transfer.

Figure 3.1: Different charge transfer mechanisms at intermediate and high salt concentrations.

higher amount of non-vehicular transport, which is all type of transport that is not vehicular. This could for example be structural deformation, structural rotation, or jumping (Fig. 3.1b).

3.2 Localised Highly Concentrated Electrolytes

Localised highly concentrated electrolytes (LHCE) are designed to retain the positive qualities of HCEs, whilst improving the reduced ionic conductivity, that comes with an increased salt concentration. This is done by introducing a non-solvent, diluent, keeping the 1st solvent shell of the cation of the HCE intact. Hence it is globally similar to a liquid electrolyte (LE), but locally a HCE. The diluent oftentimes are some kind of ether, particularly fluoroethers as seen in the works of Qian et al [36], and Wang et al. [37, 38]. Given a shared local structure with HCEs the SEI forming abilities remain [39]. As recently seen the SEI formed from LHCE system are composed mainly of species coming from the anion decomposition and (near the surface) from the decomposition of diluent as well [40, 41]. Hence decomposition will occur using these systems, and the diluent has to be chosen accordingly, increasing the complexity of choice for electrolytes. All choices can be aided using computational means. For example: as preparatory work for paper I we screened many potential fluoroethers using DFT for their interaction energy against some standard acid in order to find out how inert their interaction is. Through this method a set of suitable candidates could be preselected and the scope of testing narrowed down. Similarly, MD is an additional method that could be used for screening [42]. The importance of computational methods for investigating LHCEs are even higher than for other electrolyte types due to properties of the diluent, since many diluents contain high amounts of flourine and the F - F interaction are important for many properties of these types of liquids. These interactions are hard to probe spectroscopically, leaving computational methods as a more effective choice.
3. Electrolytes

3.3 Ca\(^{2+}\) conducting electrolytes

The work done in this thesis touches upon the problem of finding a good electrolyte for these types of systems. Currently Ca(BF\(_4\))\(_2\) salt and carbonate solvent electrolytes show a large ESW. However these system require operating temperatures above 100 °C in order to be reversible [27], severely limiting the commercial applications as well as the energy density of the system. At room-like operating temperatures THF based electrolytes can used to produce operating Ca cells. These types of electrolyte come with the drawbacks that the anodic stability is low (\(\sim 3\) V vs Ca\(^{2+}\)/Ca) which limits the selection of high voltage cathode materials [43], in turn lowering the cell energy density. Hence a viable middle ground has to be developed in order to make Ca-metal batteries a commercial reality. In paper I we investigate some possible special types of Ca electrolytes, and explore their structure using a combination of computational and experimental tools.

Ca-ion electrolytes are more affected by a low redox potential than Li-ion systems are. Since the Li-ion is \(\sim 80\%\) the size of a Ca-ion and only carries half the charge it moves more freely though the passive layer formed on the electrode, and compared to other multivalent ions, Ca has a lower redox potential meaning other multivalent systems do not induce a SEI layer mitigating ion transport. Aurbach et al. showed that a wide range of Ca electrolytes form thick passivating films, making calcium plating impossible on both Ca and noble electrodes making Ca\(^{2+}\) mobility impossible [44]. Hence one of the main goals when studying calcium electrolytes is to develop a system allowing for both calcium plating and stripping at the anode.

3.4 Ion Transport in Liquid Electrolytes

A common cell works with a cell voltage of the order of magnitude of \(< 5\) V, meaning that a Li\(^+\) experiences an energy contribution of 5 eV. Operating at room temperature the thermal energy of the surroundings is \(k_B T \approx 25\) meV. Hence, on a global scale bulk transport is driven completely by the voltage difference. Focusing on a local scale, assuming a thin separator (\(\sim 20 \mu m\)) we see that linearly approximating cell voltage over separator thickness \(\approx 2.5 \cdot 10^5\) Vm\(^{-1}\) over a typical distance of inter-molecular interaction (\(\sim 5\) Å), \(E = 0.125\) meV. As a consequence the local dynamics is assumed to display motion uniformly distributed over all directions. In practice however the voltage profile is extremely steep near the electrode interfaces, making the importance of this potential more substantial there, and even less important than assumed here within the bulk, but that is outside the scope of this thesis.

The most important property for transport performance in a battery cell is the ion conductivity \(\sigma\) of the charge carrier, defined by:

\[
\vec{J}^+ = \sigma^+ \vec{E}
\]

relating the current density \(\vec{J}\) of the charge carrier to the applied electric field \(\vec{E}\). Generally \(\sigma\) is a tensor. In isotropic media such as liquid electrolytes however it is
3. Electrolytes

reduced to a scalar. Through the Nernst-Einstein equation, the ionic conductivity can be related to the diffusivity $D^+$, which for monovalent electrolytes (which will be used for demonstrative purposes here) takes the form:

$$\sigma^+ = \frac{cF}{RT}D^+,$$

where $c$ is the salt concentration, $F$ is Faraday’s constant, $R$ is the gas constant, and $T$ is the temperature. Diffusivities tend to be easier to compute practically when simulating electrolytes, making this a more useful metric. The diffusivity can be approximated through the Stokes-Einstein equation:

$$D_i = \frac{k_B T}{6 \pi \eta r_i},$$

where $\eta$ is the dynamic viscosity of the electrolyte and $r_i$ is the hydrodynamic radius of species $i$, typically on the order of $1 < r_i < 10$ Å in typical battery electrolytes [45]. In practice however the diffusion is calculated through a method such as mean squared displacement, or a Green-Kubo equation.

From the diffusivity two related, but distinct, and oftentimes confused concepts can be defined: the transport number, and the transference number. The transport number is defined as the fraction of the total current that is carried by the cation, e.g. Li$^+$, assuming no ion aggregation. Conversely the transference number is defined as the fraction of the migration current excluding currents due to concentration gradients, that is carried by the cations regardless of the speciation. The migration current is defined as the current driven by an electric field. What can be obtained from MD simulations however is neither of these concepts, but rather $t^+$, which is oftentimes also called the transport number defined as

$$t^+ = \frac{D^+}{D^+ + D^-}.$$

Even though the confusion in definitions of transport properties cause problems when comparing studies with each other, concentration trends, etc. seem to be consistent enough across methods. In paper II we study an array of transport properties using the newly developed CHAMPION software [46].
This thesis focuses on the use of computational methods at multiple scales, ranging from molecular level methods such as DFT, to methods such as classical MD simulations, with AIMD in between (Fig. 4.1a). These methods provide vital information about the system of study, at the scale of the method. However for electrolytes it is important to also study the interconnectivity between the physics ranging from molecules and up. This concept is called multi-scale modeling, a field of science dedicated to studying and solving problems which have important features at multiple temporal and spacial scales [47](Fig. 4.1b). In papers II and III an attempt is made to marry the different scales and here a deeper explanation of the methods used will be given.

(a) Different computational techniques used at different scales.
(b) Local clusters in an electrolyte make up a global network.

Figure 4.1: Multiple scale modeling vs multi-scale modeling.

4.1 DFT

In paper I the local structures in CaB electrolytes were studied at the molecular level using DFT calculations, in order to supplement and explain experimental data. At the centre of DFT and other non relativistic quantum chemistry methods is solving the Schrödinger equation:

\[ H \Psi(r, t, \sigma) = i\hbar \frac{\partial \Psi(r, t, \sigma)}{\partial t} \]
where $\Psi$ is the wavefunction describing the system, $H$ is the system’s Hamiltonian, $\vec{r} = (x, y, z)$ all coordinates describing the system, and $\sigma$ is the spin of the system. Neatly packaged like this it is easy to think that all physics is solved, but as Paul Dirac puts it [48]:

*The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.*

This quote summarises the problem with modern materials science and chemistry, where the many-body-interactions necessary to describe the system quickly become too complex to be solved for systems larger than the hydrogen atom. Hence approximations of both the Hamiltonian and the wavefunction are needed to perform these calculations. Since a true representation of the wavefunction cannot be replicated on a classical computer [49], the first hurdle to overcome is to find an accurate representation of the wavefunction. Oftentimes single-electron wavefunctions, and molecular orbitals $\Phi_n$ are written using a basis set of functions $\varphi_\mu$ centred on the nuclei. The basis set functions are usually a linear combination of Gaussian functions

$$\varphi_\mu = \sum \mu d_\mu g(\xi, r)$$

where $d_\mu$ is a scale factor and

$$g(\xi, r) = C x^{n_1} y^{m_1} z^{l_1} e^{-\xi r^2}$$

where $C$ is a normalisation constant, $n_1$, $m_1$, $l_1$ determines the type of orbital the function represent (s, p, d, etc.). Following the creation of a basis set ($\{\varphi_\mu\}$) molecular orbitals can be defined as:

$$\Phi_k = \sum_\mu c_{k,\mu} \varphi_\mu$$

Solving the Schrödinger equation is done through systematically making better guesses for the wavefunction $\Psi(r_1, r_2, \ldots, r_n; R_1, R_2, \ldots, R_N) = \sum_k a_k \Phi_k$. What is solved for are the values of all coefficients $a_k$, enabling the energy $\epsilon = \langle \Psi | H | \Psi \rangle$ to be calculated. Since the variational principle states that $E_0 < \epsilon$, where $E_0$ is the system’s ground state energy, it is possible to iterate systematically finding better and better solutions to the wavefunction using an array of methods.

In this thesis DFT has been used as the method to find the ground state of a number of systems. Through two theorems postulated by Hohenberg and Kohn, DFT is used to identify a system’s ground state through the electron density rather than its wavefunction, reducing the number of coordinates needed to describe the system from $3(n + N) - 6$ to $3$.

The two Hohenberg-Kohn theorems are:
1. The ground state electron density uniquely determines the external potential of the system, and thus the whole Hamiltonian.
2. A universal functional, valid for any external potential, can be defined in terms of only the electron density.

Using these two theorems the energy of the ground state of the system can be calculated from the electron density:

\[ E_0 = T[n(r)] + \int_{R^3} V_{\text{nuclei}}(r)n(r)dr + \frac{1}{2} \frac{1}{4\pi\varepsilon_0} \int_{R^3} \frac{n(r)n(r')}{|r-r'|} dr dr' + E_{\text{xc}}[n(r)] \]

where \( T[n(r)] \) is the kinetic energy functional of a non-interacting electron gas, \( V_{\text{nuclei}} \) is the potential caused by the nuclei, and \( E_{\text{xc}}[n(r)] \) is the exchange correlation functional containing the remaining energy in the interaction \([50]\). Using the Kohn-Sham equation:

\[
\left(-\frac{\hbar^2}{2m_e} \nabla^2 + V_{\text{eff}}(r)\right) \Psi_i(r) = \epsilon_i \Psi_i(r)
\]

where

\[
V_{\text{eff}} = V_{\text{nuclei}}(r) + \frac{1}{2} \frac{1}{4\pi\varepsilon_0} \int_{R^3} \frac{n(r')}{|r-r'|} dr' + \frac{\delta E_{\text{xc}}[n(r)]}{\delta n(r)}
\]

enables the electron density to be determined using a self-consistent field (SCF) approach. The term \( \frac{\delta E_{\text{xc}}[n(r)]}{\delta n(r)} \), called the exchange-correlation potential \( V_{\text{xc}}[n(r)] \) is approximated differently by the various DFT functionals. Generally \( V_{\text{xc}} \) is approximated by it being expanded in terms of electron density. There are functionals ranging from simple and computationally inexpensive to very accurately and expensive ones. In paper I the Minnesota functional M06-2X \([51]\) has been used for the most part, with some simulations done using the B3LVP functional \([52]\). The Minnesota functionals are a group of parametrised exchange correlation energy functionals, based on the meta-GGA approximation meaning they include terms including the energy density, as well as it’s first and second derivatives. B3LVP on the other hand is a less complex, GGA correlation function, simply containing information of the energy density and the first derivative. However even though DFT is considered accurate the computational cost scales cubically with the number of atoms in the system, making simulations of a single solvation shell the most common in the field of electrolytes. These typically target properties such as HOMO/LUMO levels, vibrational modes and frequencies, and their associated IR intensities as well as Raman activities etc.

### 4.2 AIMD

In paper II larger systems consisting of several solvation shells forming larger structures are studied using AIMD. In order to update the atomic positions the force acting on the atom has to be identified. The Hellmann-Feynman theorem shows that the force acting on a particle \( i \) is simply determined by the electron density as well as
4. Methods & Modelling of Battery Electrolytes

the Coulomb interaction between the nuclei:

\[
F = -\frac{\partial E}{\partial R_i} = -\left\langle \Psi \left| \frac{\partial H}{\partial R_i} \right| \Psi \right\rangle = Z_i \left( \int n(r) \frac{r - R_i}{|r - R_i|^3} dr - \sum_{j \neq i}^N \frac{Z_j (R_j - R_i)}{|R_j - R_i|^3} \right).
\]

Therefore using similar methods for identifying the electron density as in DFT all forces acting within the simulation system can be identified. Knowing all the forces, the system can be propagated through time by solving Newton’s second law, and updating all particles position. One of the most common algorithms for updating positions in MD methods is the Velocity-Verlet algorithm \[50\]. Updating positions with a half-step the algorithm looks as follows:

1. \( \dot{\bar{v}}(t + \frac{1}{2} \Delta t) = \bar{v}(t) + \frac{1}{2} \ddot{\bar{v}}(t) \Delta t \)
2. \( \bar{x}(t + \Delta t) = \bar{x}(t) + \bar{v}(t + \frac{1}{2} \Delta t) \Delta t \)
3. Derive \( \ddot{\bar{a}}(t + \Delta t) \) for the updated positions \( \bar{x}(t + \Delta t) \)
4. \( \bar{v}(t + \Delta t) = \bar{v}(t + \frac{1}{2} \Delta t) + \frac{1}{2} \ddot{\bar{a}}(t) \Delta t \)

Where step 3 is non-trivial. Deriving new accelerations at each iteration means solving a conventional matrix diagonalisation for an updated electron density with each updated position, as is done when using Born-Oppenheimer molecular dynamics (BOMD). This is computationally expensive when done at every time step, making this approach risible for even moderate scale systems \[50\].

An alternative approach to combine the quantum mechanical way to derive the forces acting in the system with classical MD was done by Car and Parrinello in 1985 \[53\]. By treating the total energy in the system as a functional of the electronic wave function and nucleiic positions:

\[
E_{tot} = E_{tot}[[\{\psi_i\}, \{\tilde{R}_j\}]]
\]

in combination with the fact that the electronic wavefunctions are orthonormal \( \langle \psi_i | \psi_j \rangle = \delta_{ij} \) Car and Parrinello were able to compute the total energy. They did this by using the variational principle, minimizing \( E_{tot}[[\{\psi_i\}, \{\tilde{R}_j\}]] \), instead of computing the electronic structure at every step of the MD simulation, as the nuclei positions are varied simultaneously as the electronic orbitals. In order to propagate the system in time and give the electrons a kinetic energy, a fictitious time dependence is given to the wavefunction, which allows the following classical Lagrangian to be constructed:

\[
L(\{\psi_i\}, \{\tilde{R}_j\}) = \frac{\mu}{2} \sum_{i=1}^n \dot{\psi}_i^2 + \frac{1}{2} \sum_{j=1}^N M_j \dot{\tilde{R}}_j^2 + E_{tot}[[\{\psi_i\}, \{\tilde{R}_j\}]] + \sum_{kl} \Lambda_{kl} \langle \psi_k | \psi_l \rangle
\]

where \( \mu \) is a fictitious electron mass. The fictitious mass should be small enough such that the Born-Oppenheimer approximation still holds and the electronic wavefunction adapts to the position of the nuclei, whilst at the same time being large enough to allow for relatively large time steps. A typical choice for \( \mu = 400 \ m_e \), which is what we have used. Beyond the introduction of a fictitious mass, Lagrangian multipliers \( \Lambda_{kl} \) are introduced as to fulfil any external constraints, e.g retaining orthonormality.
of the Kohn-Sham orbitals at each time step. Solving for the classical equations of motion yield:

$$\mu \ddot{\psi}_i = -\frac{\delta E_{\text{tot}}}{\delta \psi_i} + 2 \sum_j \Lambda_{ij} \psi_j$$  \hspace{1cm} (4.1)$$

$$M_j \ddot{R}_j = -\frac{\delta E_{\text{tot}}}{\delta R_j} + \sum_{kl} \Lambda_{kl} \frac{\delta \langle \psi_k | \psi_i \rangle}{\delta R_j}$$  \hspace{1cm} (4.2)$$

which are implemented in CPMD [54], the software used to generate data for analysis in paper II, and to generate training data in paper III.

### 4.3 Classical MD

Unlike AIMD where the forces are derived from the electron density and Coulombic interactions between nuclei, classical MD calculate the interatomic forces using a FF, which is a set of parameters and functions, together forming a potential energy surface where the total system energy is determined by the atomic positions \(\{\vec{R}_j\}\). The energy is most often given by, but not limited to the form:

$$E_{\text{tot}} = 4 \sum_{ij} \epsilon_{ij} \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} + \frac{1}{4\pi\varepsilon_0} \sum_{ij} \frac{q_i q_j}{r_{ij}}$$

$$+ \sum_{\text{bonds}} \frac{k_b (l - l_{0,b})^2}{2} + \sum_{\text{bond angles}} \frac{k_a (\theta - \theta_{0,a})^2}{2} + \sum_{\text{dihedrals}} \sum_n k_{n,d} \cos(n\phi + \phi_{0,n}).$$

where the first term is the Lennard-Jones approximation describing the Pauli exclusion principle and the van der Waals interactions, the second term is the Coulomb interaction between all pairs of atoms. Term three to five are all bonded interactions modelled as Hooke’s law for two body objects, an angle potential defined by the angle between three bound bodies, and a proper/improper torsion potential defined by the angle between the planes formed between four bonded atoms. This form for the energy term (Fig. 4.2) is used in common FFs such as Amber [55], CHARMM [56], and GROMACS [57]. From this form the FF parameters are the set of \(\{\epsilon_i, \sigma_i, q_i, k_{b,i}, l_{0,b,i}, k_{a,i}, \theta_{0,a}, k_{n,d,i}, \phi_{0,n,d,i}\}\) where \(\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}\), \(\epsilon_{ij} = \sqrt{\epsilon_{ii} \epsilon_{jj}}\), and \(\sigma_{ii}\) is the distance at which the particle-particle potential energy is zero, \(\epsilon_{ii}\) is the depth of the potential well. Conventionally FF parameters are developed using a combination of quantum chemistry computations of single molecules in vacuum, and experimental data, often of thermodynamic nature. However, since FF parameters are often validated by their ability to predict thermodynamical properties through experimentally obtained values it is not guaranteed that a given FF accurately represents local dynamics, even if the parameter set accurately recreates experimental values.
4. Methods & Modelling of Battery Electrolytes

4.4 Machine Learning Methods for Electrolyte Simulations

There is a strong push to understand complex materials such as new electrolytes. One method used to tackle this problem has been machine learning (ML), with a recent surge in methods bridging the realm between AIMD and classical MD simulations. ML used to describe materials structure began in 2007 when Behler and Parinello [58] developed symmetry functions (SF), creating transferable neural network (NN) potentials through the chemical locality around an element. Similarly in 2013, Bartók et. al. developed the Smooth Overlap of Atomic Position (SOAP) directly defining the similarity between any two atomic neighbourhood environments, enabling the development of interatomic potentials through the GAP framework [59, 60]. Other alternatives to develop interatomic potentials have been through different NN approaches, such as SchNet [61]. More concretely relevant for electrolyte applications, Wang et. al. for example have developed a NN based on SchNet, learning chemical embeddings for elements in ionic liquids and new electrolytes [62]. Alternatively NN such as PiNet [63] can be used to learn electronic multipoles of atomistic simulated liquids [64].

Common for almost all ML approaches, especially those based on NN is the need for extensive amounts of data for training and testing of models, making the need for available data more important than ever. This need can be seen in contemporary projects such as BIG-MAP [65] as well as the Materials Genome Initiative [66] trying to standardise data presentation as well as making it available to a wider audience.
4. Methods & Modelling of Battery Electrolytes

4.4.1 Gaussian Process Regression

A method for interpolation of data given a sparse set of data points is Gaussian Process regression (GP), commonly known as Kriging. GP regression is used to predict a function value at a given point by computing the weighted average of the known values of the function in the neighbourhood of the point. This means that a GP is completely specified by a mean function and a positive definite covariance function. Given a set of inputs $x^{(1)} \cdots x^{(n)}$, a mean function $\mu(x) = 0$, and a covariance function $K_{p,q} = \text{Cov} \left( f(x^{(p)}), f(x^{(q)}) \right) = K(x^{(p)}, x^{(q)})$, a joint distribution may be defined

$$f(x^{(1)}) \cdots f(x^{(n)}) \sim \mathcal{N}(0, K).$$

Given this knowledge a GP can be obtained using Bayesian linear regression:

$$f(x) = x^T w$$

where the weights $w \sim \mathcal{N}(0, \Sigma_p)$. Hence the mean function is given by:

$$\mathbb{E}[f(x)] = x^T \mathbb{E}[w] = 0$$

and the covariance function is given by:

$$\mathbb{E}[f(x)f(x')] = x^T \mathbb{E}[ww^T]x' = x^T \Sigma_p x'.$$

The Bayesian linear regression is based on Bayes theorem:

$$P(y|X, \theta) \propto P(\theta)P(\theta|X, y)$$

Where $P(\theta)$ is known as the prior, representing the assumption of the probability of a set of parameters $\theta$ prior to knowledge of data, and $P(\theta|X, y)$ is known as the likelihood, representing the probability of observing the parameters $\theta$ given knowledge about the data $X, y$, and $P(y|X, \theta)$ is the posterior of the given hypothesis explaining the data. Commonly $\theta$ is chosen by optimising the marginal log-likelihood:

$$\log P(y|X, \theta) = -\frac{1}{2} \log |K(X, X) + \sigma^2| - \frac{1}{2} y^T (K(X, X) + \sigma^2)^{-1} y$$

through sampling the probability space of possible $\theta$.

4.5 CHAMPION

In paper II as well as Andersson et. al. [46] an algorithm for structure detection has been created. Through dynamic structure discovery (DSD), finding and analysing what moves together within a simulation, the CHAMPION software is capable of detecting a global bond graph for the system that uniquely determines its bond topology. This bond graph is later able to be subdivided into components such as ionic aggregates and molecules, or into solvation shells. Through knowing all bonds within a simulation, as well as all solvation shells and other structure, we are capable of of sorting all data provided in the simulation in new unique ways providing us with understanding of emergent structures within materials as well as giving insights into the dynamics of the system.
4. Methods & Modelling of Battery Electrolytes

Figure 4.3: Schematic Ávall plot linking the probability distribution of generalised forces to a generalised coordinate describing a interaction. Orange: Occupied bins. Blue: Empty bins

4.6 Automated Force Field Finder

In paper III we propose a new light-weight method for generating system specific FFs. Being inspired from the work of Ávall and Johansson (2020) [19] where they investigated the force distribution between a solvation shell and a central atom in terms of centre-of-mass coordinates, we herein develop a similar method by generalising the Ávall method of binning central forces depending on distance to binning the interaction strength related to each interaction type. The interaction types used are the same as in conventional FFs, with 2-, 3-, and 4-body bonded interactions, as well as an electrostatic term and a Lennard-Jones like term. In this method however the long range electrostatic and dispersion interactions are treated together and will henceforth be referred to as the pairwise interaction. By first running small first principle MD simulations, the forces acting between all particles of a system type can be found. Then projecting the generalised force acting on all particles partaking in an interaction type based on a CHAMPION bond graph, against a generalised coordinate $q$ describing said interaction, a 2D histogram of distributions (Fig. 4.3) are formed.

These 2D histograms contain all information about said effective interaction, e.g. between two carbon atoms, where effective interaction denotes the distributions of interactions in the presence of environmental noise assumed to be normally distributed with zero mean. Having this method based upon the CHAMPION bond-graph identifying method [46] also enables a FF that treats different bond types separately, e.g. separating linearly bonded carbon from cyclically bonded carbon. Such distinctions have been shown to be useful in other, more conventional, FFs such as AMBER and GROMACS [55, 57].
From the positions of and forces acting on each atom, both the generalised coordinates and forces can be computed and sampled. The statistical distributions of generalised forces as functions of the corresponding generalised coordinates result from a combination of the effective interaction between the atoms involved, and effects of the background. The approach taken here is based on assuming the background effect to be unbiased noise, so that the mean of the sampled distributions estimates the true effective interaction.

Knowing the positions and forces acting on each particle within a trajectory, generalised coordinates and generalised forces can be computed and sampled. The distributions of generalised forces as function of the corresponding generalised coordinates give rise to a 2D histogram similar to an Avall plot. These histograms contain the information about the effective interaction between all particles involved, as well as the effective background.

Given a force distribution $F|_q$ of generalised forces $\{F_i(q)\}$ as seen in Fig. 4.3, where $i \in \{0, \text{number of samples of interaction type}\}$ one can show that the specific force $F_i(q)$ can be written as:

$$F_i(q) = F(q) + \Delta f$$

where $F(q)$ is the true generalised interaction strength and $\Delta f$ is a stochastic background noise. Assuming that the noise is Gaussian:

$$\sum_i^N F(q) = \sum_i^N F_i(q) - \sum_i^N \Delta f$$

$$\Rightarrow$$

$$NF(q) = N \langle F_i(q) \rangle - N \langle \Delta f \rangle$$

$$\Leftrightarrow$$

$$F(q) = \langle F_i(q) \rangle .$$

Hence any interaction described using the generalised Avall method should be reproducible given the mean force value at any point along the generalised coordinate axis.

Given the sparse but spread out nature of the data produced this way, a smooth function filling the space between data points can be generated using GP regression. This pairs well with the reactive capability provided by having the method based upon the CHAMPION method. The FF identified can be made reactive through computing bond likelihood functions as a function of distance. Hence it is possible to determine on a snapshot-by-snapshot basis which atoms in the system are bonded in, one, two, three, or even four of the four bond types previously discussed. This possibility enables the look-up tables to be dynamically chosen during the simulation, enabling a great amount of customiseability to the interactions.
4. Methods & Modelling of Battery Electrolytes
Results & Discussion

Here the results of the appended papers I - III are briefly presented and discussed. These move from the molecular level using DFT methods to determine the Raman spectra of solvation shells (paper I), to the microscopic scale where AIMD is used to equilibrate what structures form in a HCE (paper II), and these are finally tied together through the use of the newly developed CHAMPION \[46\] bond graph discovery algorithm and the generalised \textit{Avall} method (paper III).

5.1 Ca$^{2+}$ 1st Solvation Shells

In order to explain experimental results depending on local structure computational methods are useful. In paper I we calculate the Raman spectra of a wide set of Ca$^{2+}$ 1st solvation shell structures in order to elucidate the local structure within Ca$^{2+}$ HCEs and LHCEs. Artificial Raman spectra are calculated from which it is clear that free PC, [Ca(PC)$_4$]$^{2+}$, free TFSI, and [Ca(TFSI)$_2$] all play an important role (Fig. 5.1). These structures are limited in complexity by both the capabilities of the DFT method as well as the researchers’ intuition, which puts limitations to the predictions possible to make for complex materials.

By investigating a wide suit of structures of the form [Ca((TFSI)$_N$PC)$_M$]$^{(N-2)-}$, where $N$ is the number of TFSI ions surrounding a central Ca ion, and $M$ is the number of PC molecules surrounding said Ca ion, in paper I the main peaks in the region of interest could be identified as the ones mentioned (Fig. 5.1c). From the combination of experimental Raman data and computational DFT data we could confirm that the first solvation shell is retained when diluting a HCE, creating a LHCE, making this a viable approach when searching for Ca-electrolytes that have the potential for SEI formation, and cycle at room temperature. By understanding the local structure the design of electrolytes can be aided. DFT works well as a complimentary method to Raman spectroscopy, however to make predictions about structure DFT is not the most optimal method since the input structure is what gets tested, making it a cumbersome approach to make predictions based on the full structure space. Hence in order to understand fully unknown materials, methods such as MD can provide a better fit.
5. Results & Discussion

(a) Band assignments of HCE and LE through the use of DFT (symbol key subFig. 5.1c).

(b) Band assignment when diluting HCE from 3.256 m to a nominal 0.45 m LHCE (symbol key subFig. 5.1c).

(c) A symbol key to 5.1a and 5.1b: Free PC (○), [Ca(PC)₄]²⁺ (●), free TFSI (○), [Ca(TFSI)₂] (★).

Figure 5.1: Use of DFT to elucidate the local structure in bulk electrolytes.
5. Results & Discussion

Figure 5.2: The most common topologies around a Li cation in order of probability. Element colors: purple: Li, red: O, blue: N, grey: C, white: H, yellow: S, green: F.

5.2 Local to Global Structure

In paper II we identify how to use the newly developed CHAMPION method together with AIMD simulations to gain knowledge of both the local and global electrolyte structure. AIMD enables simulations where the local interactions are replicated accurately, as can be observed with the coordination number (CN) for $\text{Li}^+ \approx 4$, which is common for LIB electrolytes based on small organic molecules, regardless of composition [67, 68]. In many electrolytes CN is similar to the solvation number (SN) since almost all coordinations are monodentate. This is, however, not true for LiTFSI in ACN at higher concentrations where the SN is closer to 3, even though the CN remains close to 4 due to more bidentate TFSI coordination by Li$^+$, which concurs with other studies [69]. These results are reflected in the common topologies (Fig. 5.2) found in the simulation.

On a global scale we see that these structures form a percolating network, in a sea of free solvent (Fig. 5.3), which is in sharp contrast to the common conception about HCEs, where its unique behaviour is believed to stem from a lack of free solvent [70].

The accuracy of this analysis scales with the number of ions and even for a concentrated system such as this, system size and trajectory length are both on the smaller scale in order to say something with statistical accuracy. Hence less concentrated systems, such as LHCEs for example, require much larger simulations to enable the same accuracy level as the analysis results of a HCE which more or less disqualify AIMD for such systems.
5. Results & Discussion

Figure 5.3: Snapshots of the periodic simulation cell, highlighting the percolating network in a sea of solvent.

5.3 Generated Force Fields

To use the CHAMPION method to the full extent large simulations must be utilised. For systems that would benefit the most from being studied by the CHAMPION method (HCE, LHCE, etc.) FFs usually do not capture the correct dynamics, hence not allowing for these systems to be studied through MD simulations [13]. Paper III focuses on extending the functionality of the CHAMPION method by combining it with the insights gained in Ávall & Johansson 2020 [19]. The quality of the FF generated depends heavily on the number of bins $n$ (Fig. 5.4). To overcome this hurdle a large enough number of bins was chosen after converging such that it was unlikely that the bin size would affect the mean force values to a larger extent. Then the data was used to generate a GP regression model. To ensure that the correct interactions were captured the data was divided in two halves, where the first half start from an unequilibrated state, and the second half starts where the first ends (Fig. 5.5). As seen the method capture the same physics at both long and short range, indicating that correct electrostatics as well as dispersion is found.

However even though we see that the GP regression take the same form for both the equilibrated and unequilibrated data (Fig. 5.5) using these FFs deteriorate a simulation after only a few time steps. These FF curves were generated by sampling the data every 10 fs. This yields an incredibly low statistical inefficiency $s$ (table 5.6), hence the number of time steps for which a force correlations effectively persist is on the scale of $\sim 1$. Sampling the data every 0.1 fs yields a higher correlation of the forces, hence, when calculating forces using finite difference, the forces should be more accurate when compared to ones yielded using a longer time step, and in turn indicate that the simulation will not melt or explode for a larger number of time steps.

No parameter besides $\Delta t$ should affect $s$, however since this quality is measured through a curve fit there is not an exact measurement, hence the order of magnitude
5. Results & Discussion

(a) How the C-H force field depends on the number of bins $n$.

(b) How the Li-O force field depends on the number of bins $n$.

**Figure 5.4:** Comparing the quality of the generalised Åvall FF depend on number of bins.

is more important than the specific number and the measured values in table 5.6 should be seen as a rough estimate. When sampling every 0.1 fs for finite difference calculations of forces (Fig. 5.7) the magnitude of forces are more similar to literature [19] than when sampling every 10.0 fs. However there is a large discrepancy of one to two orders of magnitude. This discrepancy makes it apparent that finite difference used still does not capture the true dynamics, which is also apparent since MD simulations using these FFs melt. Note though that at present these results are under development, which underpins the discussion here.

Moving forward the data collection has to be changed to solve the issue caused by using a finite difference method to obtain forces. Worth noting is that the error bars (Fig 5.7) are small, indicating that the data could be separated in such a manner that e.g. linearly bonded C-C and cyclically bonded C-C can obtain their own FF term. The fact that different geometries does not at the moment obtain their own term can of course also contribute to the erroneous results, especially since this could explain the multi-peak nature of the FFs. Nonetheless more work is needed to tune this method.
5. Results & Discussion

(a) A force field generated for H-H interaction using Gaussian Process regression for the 1st and 2nd half of a simulation trajectory.

(b) A force field generated for C-H interaction using Gaussian Process regression for the 1st and 2nd half of a simulation trajectory.

(c) A force field generated for Li-O interaction using Gaussian Process regression for the 1st and 2nd half of a simulation trajectory.

Figure 5.5: The form of the Avall FF after using the raw data as input to a Gaussian Process regression. The same physics is captured when studying the 1st half of a simulation as when studying the 2nd half.
5. Results & Discussion

Table 5.6: The statistical inefficiency $s$ dependence on $\Delta t$, the number of time steps $N_{\text{timesteps}}$, and the graph radius $d$.

<table>
<thead>
<tr>
<th>$\Delta t$ [fs]</th>
<th>$N_{\text{time step}}$</th>
<th>$d$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0</td>
<td>60218</td>
<td>3</td>
<td>1.869</td>
</tr>
<tr>
<td>0.1</td>
<td>1000</td>
<td>3</td>
<td>22.7221</td>
</tr>
<tr>
<td>0.1</td>
<td>10000</td>
<td>3</td>
<td>72.0633</td>
</tr>
<tr>
<td>0.1</td>
<td>1000</td>
<td>1</td>
<td>35.3487</td>
</tr>
<tr>
<td>0.1</td>
<td>5000</td>
<td>1</td>
<td>21.1608</td>
</tr>
</tbody>
</table>

(a) A force field generated for H-H interaction using Gaussian Process regression based on data with a statistical inefficiency $s = 21.1608$.

(b) A force field generated for C-H interaction using Gaussian Process regression based on data with a statistical inefficiency $s = 21.1608$.

(c) A force field generated for Li-O interaction using Gaussian Process regression based on data with a statistical inefficiency $s = 21.1608$.

Figure 5.7: The form of the Åvall FF after using the raw data as input to a Gaussian Process regression. Data with a statistical inefficiency $s = 21.1608$. 

29
5. Results & Discussion
Conclusion & Outlook

This thesis has studied battery electrolytes using computational means at different scales, and tied together these scales through the development of a new method. In paper I and II the importance of studying one scale to understand another is showcased. In paper I through seeing that the local structure of HCEs are retained in LHCEs, and from this being able to create hypotheses regarding macroscopic properties. Similarly in paper II through understanding how the local structure gives rise to the global structure, and how the global structure can be used to explain transport properties. These two papers exemplify how computational methods open up a future where materials science can be made more effective and predictive.

In order to enable such predictions using the CHAMPION [46] framework large scale MD simulations have to be easily available. In paper III the development of a method trying to tackle this problem was initiated, however much work is still needed on this method. Moving forward it would be interesting to fine tune the FF by adding on bonded terms, ensuring that molecular shape is retained more accurately. Besides adding bonded terms the immediate calls to action to improve the method presented in this thesis are:

- Use AIMD forces instead of finite difference generated forces.
- Separate interactions based on local structure rather than solely on participating elements.
- Enable a reactive FF for bonded terms.

Given all these pieces it would be interesting to study large scale simulations of LHCEs and studying how these systems behave and what structures form on a global scale. Especially the cumulative effects when large scale MD simulations for any material quickly and cost effectively is combined with the CHAMPION analysis. Through such endeavours it will be most interesting, seeing what physical properties can be explained this way. These types of studies should enable probing the phase separation between the diluent and the HCE structures, making it possible to study cluster formation, and cluster size, transport phenomena etc. Similarly the combination of the methods presented in this thesis could provide an avenue to study electrolyte-electrode interface and interphase interactions. From there on it is up to the imagination of the reader to find interesting problems where these methods can be applied.
I would like to start by thanking my supervisor Professor Patrik Johansson for giving me the freedom to pursue research freely, and for all the support in writing this thesis. Secondly Professor Aleksandar Matic, my examiner and closest boss, thank you for keeping spirits up around MF both online and offline during these trying times. Thank you Dr. Rasmus Andersson for your continuous support, friendship, and rubber ducking. Without the discussions you and I have had throughout the years we would not be where we are now. I would also like to thank all the colleagues at MF that have come and gone throughout the years, you make the PhD candidate life easier and I hope we can start hanging out face to face again soon. In addition I would like to send a big thank you to Victor Westergård and Peter Deaking at Chalmers Innovation office for all help with patenting our insights. Lastly I would like to thank Emil Krutmeijer, Johannes Henriksson, and Rasmus yet again, my teammates at Compular, the ball is rolling let’s keep it up! Also thank you Bo, hopefully you’ll be back soon! I would also like to show my appreciation to my funders, Energimyndigheten and the CARBAT project for enabling it all as well as the Swedish National Infrastructure for Computing (SNIC) for providing my work with computational resources.

Fabian Årén, Gothenburg, August 2021
Bibliography


[14] Bharath Ravikumar, Mahesh Mynam, and Beena Rai. Effect of salt concentration on properties of lithium ion battery electrolytes: a molecular


[41] Xia Cao, Xiaodi Ren, Lianfeng Zou, Mark H Engelhard, William Huang, Hansen Wang, Bethany E Matthews, Hongkyung Lee, Chaojiang Niu, Bruce W Arey, et al. Monolithic solid–electrolyte interphases formed in fluorinated...


[65] BIG-MAP.

[66] Materials Genome Initiative | WWW.MGI.GOV.


Paper I
(Localized) Highly Concentrated Electrolytes for Calcium Batteries

P-A. Martin¹, F. Årén¹ and P. Johansson¹,²*

¹Department of Physics, Chalmers University of Technology, 412 96, Gothenburg, Sweden
²ALISTORE-European Research Institute, CNRS FR 3104, Hub de l’Energie, Rue Baudelocque, 80039 Amiens, France

*Corresponding author: patrik.johansson@chalmers.se

Abstract
Recently the development of the concept of non-aqueous highly concentrated electrolytes (HCEs) for modern rechargeable batteries has evolved further by also adding a non-coordinating solvent, i.e. a diluent, to create localized HCEs (LHCEs). LHCEs rely on a molecular level design of charge carriers similar to HCEs in synergy with tailoring macroscopic properties, such as reduced electrolyte viscosity. So far LHCEs have mainly been investigated for Li and Na based battery systems, but here we explore the concept of HCEs and LHCEs for divalent Ca²⁺ conducting systems. We do this by systematically elucidating the specific limits and molecular origins of macroscopic features of the Ca(TFSI)₂-PC and Ca(TFSI)₂-PC-TTE electrolytes as function of composition. For these systems some unique HCE properties arise at ca. 2.0 m and the local structure integrity can indeed be maintained even at a nominal 0.45 m LHCE starting from a 3.256 m parent HCE. The collective observations of molecular and macro level features made herein should pave the way for further optimization of the physico-chemical properties, design of electrochemical investigations, and eventually a better understanding of how to best improve the desolvation kinetics at the electrolyte/electrode interfaces e.g. at a Ca metal anode.

Keywords: Calcium battery, cation solvation, ion transport, electrolytes, LHCE
**Introduction**

Sustainable and affordable energy for all, as for instance outlined in the UN SDG #7 [1], heavily implies better ways, in a very wide sense, for storing renewable energy generated from *e.g.* solar and wind power. Today there is fast progress in the area of large-scale electrochemical energy storage wherein foremost lithium-ion batteries (LIBs) are utilized at MW-GW scale and moving towards GW-TW scale [2]. There are, however, long-term sustainability issues and large price fluctuations of the elements and materials used, foremost cobalt (Co) and nickel (Ni) used in the NMC/NCA cathodes, but also for LIBs based on LiFePO₄ cathodes, void of both Co and Ni, lithium itself as well as the natural graphite used as active material in the anodes are problematic. This is recognized by the battery R&D community pushing for various next generation battery (NGB) technologies in an agnostic fashion for stationary and mobile applications.

Within the portfolio of NGBs, multivalent batteries based on Zn²⁺, Mg²⁺, Ca²⁺, and Al¹⁺ are a multifaceted family fundamentally very promising, both with respect to performance measures and from a sustainability perspective [3-5]. Very recently calcium batteries (CaBs) have been highlighted as an especially interesting option as: *i*) calcium (Ca) itself is the 5th most abundant element in the Earth’s crust, and thus holds promise for long term sustainability, *ii*) Ca has a low standard reduction potential of -2.87 V vs. SHE *i.e.* close to that of Li (-3.04 V), thus promise of creating high-voltage cells – in stark difference to both Mg and Al based batteries, and *iii*) the Ca²⁺ cation should exhibit faster kinetics due to its less polarizing character than *e.g.* the more popular Mg²⁺ in Mg based batteries [6,7].

From a performance perspective, CaBs based on calcium metal anodes theoretically offers high both volumetric and gravimetric capacities as shown by “real” cell level simulations [8]: up to *ca.* 400 Wh/kg and 1000 Wh/l at cell level when using an (imaginary) 200 mA/h/g cathode providing a rather modest cell voltage of 3.5 V. As of yet, however, even if fast progress is made at present [7], reliable operation of CaBs is both sluggish, leading to low C-rates and elevated temperature to be used, and hard to enable at all. Much of these problematics, albeit often not very substantiated, have their origin in / are blamed on the electrolytes used and foremost on the slow kinetics of desolvation of Ca²⁺ at the electrolyte/electrode interfaces.

For instance, the seminal work by Ponrouch *et al.* used both a lower salt concentration, 0.45 M of Ca(BF₄)₂, than the 1 M conventionally used for LIB electrolytes and an elevated temperature, 100°C, to achieve reversible electroplating [9]. Many similar, mainly electrochemical studies employing both 2- and 3-electrode cells, have followed in recent years, with other Ca-salts employed [10-13], some with very non-standard anions employed [11-13], but no silver bullet electrolyte has emerged. In parallel, there has been both experimental and computational efforts made to in detail study and connect the molecular level interactions, incl. the Ca²⁺ solvation as function of both salt and solvent(s), to the
observed physico-chemical properties, primarily conductivity, density, and viscosity, but extended also to ionicities and electrochemical stability windows (ESWs) [14,15].

However, all of this development mainly concerns CaBs using standard “salt-in-solvent” liquid electrolytes (LEs). In contrast both Li and Na based batteries R&D have seen an increased interest in applying conceptually different electrolytes, one of these being the class of highly concentrated electrolytes (HCEs) also known as “solvent-in-salt” electrolytes [16-19]. The interest in HCEs stems from that they display wider ESWs, reduced volatilities, and increased liquidus ranges – simply by the increased salt concentration using the same chemistry of salts(s) and solvent(s) as in conventional LEs. This is either directly or indirectly related to the electrolyte speciation, moving as a function of salt concentration from being rich in free (non-coordinated) solvent and with well-separated ions as charge carriers in LEs, to more extensive ion-pairing and less free solvent in HCEs (Scheme 1, left and middle), and for some HCEs finally to ionic (percolating) networks with little or no free solvent [19]. This is also concomitant with changed modes of (cat)ion transport as can be derived by different modelling approaches especially developed for these salt-rich systems, a topic explored heavily at present [20-23]. Thus the basic idea is to create cation, here Ca$^{2+}$, first solvation shells rich in anions, causing ion-pairs and aggregates to dominate the speciation, and to reduce/eliminate the “free” solvent in the electrolyte. However, as a result HCEs often exhibit rather poor (total) ionic conductivities and very high viscosities, causing both severe transport limitations as well as poor wettability of electrodes (and sometimes separators), as compared to their analogous LEs. In addition, the electrolyte cost can be expected to be more or less proportional to the (expensive) salt concentration.

![Scheme 1: The conceptual differences between LEs, HCEs, and LHCEs. Inspired by [24].](image)

In order to tackle these new challenges, the concept of localized HCEs (LHCEs) was launched [24-26]. With the idea being that the charge carrier speciation created in HCEs is beneficial for physical-chemical-electrochemical stability and should be kept, but that the macroscopic dynamic electrolyte properties must be altered – especially the viscosity lowered, the HCE is in a LHCE complemented by a (to all species) non-coordinating solvent, e.g. a diluent (Scheme 1, right). All else matters aside, this should even for relatively expensive diluents lower the electrolyte cost for a given volume of electrolyte.
by reducing the amount of salt needed. To date, the diluents used to create LHCEs are typically heavily fluorinated ethers such as bis(2,2,2-trifluoroethyl) ether (BTFE) or 1,1,2,2-tetrafluoroethyl 2,2,3,3-tetrafluoropropyl ether (TTE) [24]. This is due to a combination of low dielectric constants and low donor numbers, while both their cost and sustainability are questionable.

To the best of our knowledge neither HCEs nor LHCEs have previously been reported for CaB electrolytes, excepted the 1.5 M electrolytes composed of Ca bis(trifluoromethanesulfonyl)imide, Ca(TFSI)$_2$, in various carbonate solvents reported briefly on by one of us [14]. Herein we show proof-of-concept and study in detail and systematically physico-chemical properties of a wide range of HCEs and LHCEs based on Ca(TFSI)$_2$ dissolved in propylene carbonate (PC) as solvent and TTE as diluent.
Experimental and Computational

Materials and Electrolyte preparation

All HCEs were prepared by direct mixing of salt (Ca(TFSI)$_2$, Solvionic, 99.5%), and solvent (PC, Aldrich, anhydrous, 99.7%) by stirring overnight on a hotplate at 50°C. In total 10 different HCEs were prepared and the maximum salt concentration obtained was 3.256 m. This odd concentration value in molal corresponds to a 1:3 Ca:PC molar ratio. All HCEs are marked on the left side of the triangle of the ternary composition diagram (Figure 1) (3.7 m resulted in an oversaturated electrolyte).

![Figure 1: Ternary composition diagram of the HCEs and LHCEs studied.](image)

All LHCEs were prepared similarly as the HCEs with the addition of the diluent component (TTE, Apollo scientific, 99%). (The alternative procedure of first preparing the HCE and then adding the diluent did not result in any notable differences as determined by Raman spectroscopy). Four sets of LHCEs were made, summing up to in total 37 different electrolytes (including the HCEs) – each represented by a sloping line right-to-left in the ternary composition diagram: i) 8 in the range 0.45 – 2.5 m, starting from the 3.256 m HCE composition, ii) 7 in the range 0.45 – 2.0 m from the 2.5 m HCE, iii) 5 in the range 0.45 -1.75 m) from the 2 m HCE, and iv) 4 in the range 0.45 -1.25 m from the 1.5 m HCE. Thus 24 of the 37 electrolytes are LHCEs. All materials and electrolytes were prepared in a dry (<1 ppm O$_2$, <1 ppm H$_2$O) Ar glove box and kept therein until characterization. The water content was <100 ppm for all salts, solvents and electrolytes as measured by Karl-Fisher coulometry (Metrohm 831 Coulometer).
**Physico-chemical characterization**

The ionic conductivities were measured from 20 to 70°C in 10°C steps, with an equilibration time of 20 min at each temperature, by broadband dielectric spectroscopy using a Novocontrol Concept 80 equipment. The DC conductivities were extracted from the AC high frequency plateaus. Coin-cells with stainless steel (SS) electrodes were filled with 100 µl of electrolyte in controlled atmosphere, using a PTFE spacer with a 5 mm inner ring diameter and a thickness of 1 mm. Three coin-cells were prepared for each of the electrolytes to assure repeatability.

The densities and the viscosities of were likewise measured from 20 to 70°C in 10°C steps, using an Anton Paar DMA 4500M density meter equipped with a Lovis 2000M rolling ball viscometer module.

Raman spectra were acquired at room-temperature using a Bruker MultiRAM FT-Raman spectrometer for 1000 scans with a spectral resolution of 2 cm$^{-1}$ using a Nd:YAG laser (1064 nm, 500 mW) as the excitation source. Fitting and deconvolution of the spectra were performed in selected regions to address the TFSI and PC speciation separately using the PeakFit™ using Voigt functions.

**Density functional theory (DFT) calculations**

To assist in the interpretation of the Raman spectra in terms of TFSI and PC speciation, DFT calculations were performed using the B3LYP and the M06-2X functionals and the 6-311G(d,p) basis set employing an implicit solvent (water) via the PCM/SCRF methodology [27,28]. All calculations were made in the Gaussian16 software [29] with the starting geometries created in Avogadro [30]. For the most stable geometries obtained for the “free” TFSI anion, Ca$^{2+}$-TFSI ion-pairs and neutral triplets, pure PC, and each of the [Ca(PC)$_x$]$^{2+}$ systems (x=1-8), the vibrational frequencies and the corresponding Raman activities were calculated analytically by 2nd and partial 3rd derivatives of the energy, respectively. This allows us to qualitatively identify the Ca$^{2+}$ induced shifts in the TFSI and PC modes chosen for analysis and to assist in the semi-quantitative analysis of the speciation.
**Results and Discussions**

Herein we look at the HCEs, starting from the LEs and increasing the salt concentration, and subsequently the LHCEs the same way; first different physico-chemical properties as function of salt concentration are targeted and then we connect those observations with the various local speciation changes as derived from the Raman spectroscopy analysis leveraged by the DFT calculations. Furthermore, the ionicities of the electrolytes as obtained from Walden analyses are handled as a level connecting the local origins and the global behaviour. Some implications and paths forward for practical CaB usage are discussed.

**LEs and HCEs**

Starting with the ion conductivity of the electrolytes, the build-up curve as a function of salt concentration at 50°C shows a rather clear maximum at ca. 0.75 m (Figure 2), but show also very similar ion conductivities for 0.5 and 1.0 m.

![Figure 2: Ion conductivity as function of salt concentration at 50°C for the Ca(TFSI)₂ in PC system.](image)

This is more or less in agreement with the literature wherein Ca(TFSI)₂ in EC, PC and DMF as single solvent electrolytes are quoted to show maxima at 0.42-0.57 M salt concentration at room-temperature [14]. The maximum here is very close to where most conventional LEs are made for CaB studies, but these most often additionally have a linear carbonate solvent such as DMC in the composition which further raises the absolute ion conductivity. It is also somewhat lower in concentration than the corresponding maximum for Li and Na based electrolytes, showing on the role of the increased ion-ion [31] and ion-solvent [32] interactions by the larger charge/radius ratio of divalent cations such as Ca²⁺. Note, however, that the simple measure of total ion conductivity as function of salt concentration does
not at all reflect any changes in the cation vs. anion contributions nor in the mode(s) of ion transport. All build-up curves, \textit{i.e.} for other temperature, have a similar general appearance.

Turning to the viscosity of the electrolytes, this is a seemingly a property having a very simple monotonic and close to linear (\textit{i.e.} Arrhenius behaviour) temperature dependence for all concentrations (Figure 3). Note the inverted logarithmic y-axis, which also means that the less fluid electrolytes are to be found at the bottom of the graph.

![Figure 3: Viscosity as function of temperature and salt concentration for the Ca(TFSI)$_2$ in PC system.](image)

The least concentrated electrolyte at 0.45 m is still more concentrated than the corresponding 0.1 M electrolyte reported in the literature \cite{14} and accordingly and assuring also the viscosities are significantly higher. However, there are quite a few extra features noticeable from the viscosity data and trends. First, the absolute differences in viscosities between the electrolytes decrease at elevated temperatures, \textit{i.e.} the dependence on temperature are somewhat different. Second, overall the differences as function of salt concentration decrease the more concentrated the electrolytes are, \textit{e.g.} the very small difference between the 2.0 m and the 2.5 m electrolytes ($\Delta = 0.5$ m) as compared to between the 0.45 m and 0.75 m ($\Delta = 0.3$ m). Though, this observation is perhaps better thought of in terms of relative salt concentration increases, \textit{i.e.} +25$\%$ and +66$,\%$, respectively. Finally, there is a clear step function for the 3.256 m electrolyte, having a significantly higher viscosity – more akin to an ionic liquid based electrolyte than a traditional LE or the other HCEs. This electrolyte is created such that there no longer is enough PC available to create any Ca$^{2+}$ first solvation shells composed only or even predominantly of PC, this as the Ca:PC molar ratio is 1:3 and the preferred total coordination number (CN) for Ca$^{2+}$ is closer to 5-6 in dilute 0.1 M PC solutions \cite{14}. Hence, the local structure and thereby the dynamics changes drastically as communal solvation becomes important, as recently observed for Li conducting
PC based HCEs [33], and furthermore also a percolation network might be formed at these concentrations [21,22].

Overall the trends of the two electrolyte properties above as function of salt concentration are often argued, for LEs as well as for HCEs, from a Nernst-Einstein basis to be the expected behaviour as the (beneficial) increased amount of charge carriers is counteracted by the (unfavourable) increased viscosity, i.e. slower ion diffusion/migration, and/or increased ion-pairing and aggregation (reducing the number of effective charge carriers). The importance of the latter for both LEs and HCEs can be approached in a more profound manner by creating and analysing a common Walden plot (Figure 4) [34].

![Walden plot connecting the ion conductivities and viscosities.](image)

*Figure 4: Walden plot connecting the ion conductivities and viscosities.*

In the Walden plot the impact of the salt concentration on both the (molar) ion conductivity and the viscosity at different temperatures can be seen. Even if the mix of divalent cations and monovalent anions make the values harder to interpret in any absolute manner and they cannot be compared to literature values for monovalent, i.e. Li or Na, conductors, it is still so that the further below the 1 M KCl reference line the data points are the less ionic are the electrolytes. Clearly the “ionicity” decreases by formation of ion-pairs and aggregates, and as can be expected this is a strong function of salt concentration. Overall, the ion conductivities and the viscosities are affected in a similar manner until the salt concentration reaches approximately 2.0 – 2.5 m. Beyond 2.0 m the viscosity-ion conductivity relation has the same activation energy, but clearly the much higher viscosity of the 3.256 m electrolyte (Figure 3) was not reflected in a similar decrease in the ion conductivity (Figure 2) and hence these properties are decoupled i.e. the ion transport is not a simple vehicular transport limited by the
viscosity/fluidity of the electrolyte. While this in principle is one of the aims of creating HCEs, the decoupling and possibly drastically improved cation transference numbers [18], the very low total ion conductivities at which this occurs means that even if all of the total ion conductivity would be cation based it would still likely be too low for any practical application. Overall this confirms the notions about limitations of HCEs and the urge to create LHCEs or other modifications to the HCEs.

Staying with the HCEs, however, we do want to verify and enable a connection between the macroscopic limitations found with a more local picture and therefore to elucidate the Ca^{2+} cation solvation – as this molecular level feature directly underpins in principle all those macroscopic properties. By detailed Raman spectroscopy analyses we set out to probe the cation solvation both by the PC solvent and the TFSI anion in a specific region of the Raman spectra, 700-760 cm^{-1}, as shown below (Figure 5).

![Raman spectra for a few selected electrolyte salt concentrations in a region sensitive to both cation-anion and cation-solvent interactions.](image)

In this region the γ(C=O) vibrational mode of both “free”, i.e. non-coordinated, and Ca^{2+}-coordinated PC are noticeable at 713 (circle) and 728 cm^{-1} (grand star), respectively, alongside the TFSI breathing mode
[35] of both “free” (pentagon) and Ca$^{2+}$-coordinated TFSI (star) are present at 741 and 750 cm$^{-1}$, respectively [14]. As HCEs are characterized, and indeed sometimes defined by the absence of free solvent, clear signs of ion-pairing and aggregate formation at the expense of free PC is an utmost important feature to study. Indeed, by the eye it is clearly noticeable that upon moving from pure PC and the salt concentrations clearly being LEs (top) to those approaching HCEs (middle and bottom) the free PC band at 713 cm$^{-1}$ rapidly decreases, while the Ca$^{2+}$-PC band at 728 cm$^{-1}$ increases. A similar behaviour is observable for the all electrolytes for the free TFSI and Ca$^{2+}$-TFSI bands. Ultimately, for the 3.256 m electrolyte both the free PC and free TFSI bands have almost totally disappeared, and while it somewhat depends on the criteria used to define an electrolyte as an HCE, for sure this electrolyte is a strong candidate to be an HCE. From the Raman spectra as function of salt concentration we can also obtain semi-quantitative information on the cation solvation all the way from LEs to HCEs by combining data from DFT calculations for free PC and free TFSI together with different [Ca(PC)$_x$]$^{2+}$, x=1-8, clusters and Ca$^{2+}$-TFSI ion-pairs and aggregates and a proper deconvolution of the series of spectra (Figure 6).

![Raman Spectrum Example](image-url)

**Figure 6:** Example of fit and deconvolution performed on the Raman spectra in the selected region.

In the fitting two peaks were used for free PC, corresponding to two unique vibrations, while free and Ca$^{2+}$-coordinated TFSI both were fitted using only a single peak despite that the TFSI anion exists in two different conformers [35,36], but indeed the Raman shift between them is very minor for more or less all ionic systems [37]. The situation for the Ca$^{2+}$-coordinated PC band, however, is much more complex [33] and here no less than three peaks were used, in order to achieve a reasonable fit as function of salt concentration *i.e.* to mimic/represent the differences in CN in the cation first solvation shells. Overall this process follows nicely what has been seen and used for both Li, Na and Ca based HCEs in the literature [14,33], including the obvious differences in CN and speciation. We stress that the DFT calculations for both the TFSI and the PC containing species show the Raman activities/cross-sections...
to be more or less linear functions of the number of relevant oscillators (TFSI and PC) in the various species. Hence the peak areas from the fitting process can be used as a direct semi-quantitative measure of the (relative) populations of each species. From this and by the resulting percentages the cross-over concentrations as function of salt concentration, *i.e.* where more solvent and anions are coordinated than not, are obtained at 1.5 m for the TFSI anions and 0.75 m for PC (Figure 7). The latter, which is the most conductive electrolyte (Figure 2), thus rather surprisingly only has 50% free PC, but also has rather few ion-pairs (ca. 20%).

![Figure 7: The ion-pairing (left) and cation-solvent interactions (right) as function of electrolyte salt concentration.](image)

Moving to the other end of the salt concentration to what we above has declared the most obvious HCE, the 3.256 m electrolyte with its Ca:PC 1:3 molar ratio, only has 13% of its TFSI anions free. However, more interesting is the PC behaviour as function of salt concentration. With the caveat of how communal solvation by PC would show up in the Raman spectra [33] and how this affects the analysis in terms of speciation, it seems like that from *ca.* the 2 m electrolyte there is a constant (relative) population of free and Ca$^{2+}$-coordinated PC, while less and more, respectively, in absolute numbers. This could also be chosen as a defining point for when a HCE has been formed. Based on simulations performed on Li based systems [21,22] we can also speculate that already then, at 2 m concentration, some kind of cationic percolation network has been formed with some TFSI anions free to move. This is on the other hand not really supported by the viscosity data, where as outlined above the property step change occurs first for the 3.256 m electrolyte.

**LHCEs**

Given the problematic features of HCEs, both quoted in general in the literature and confirmed also for the Ca$^{2+}$ based system above, we now turn to the LHCEs as a possible remedy/solution. Using the system above as the natural starting point and adding the most popular non-solvent/diluent TTE a direct
comparison of the ion conductivities at the same absolute concentrations, clearly shows that the dilution reduces the ion conductivity e.g. at 0.75 m (Figure 8).

![Figure 8: Ion conductivity as function of salt concentration for the Ca(TFSI)_2 in PC and TTE system.](image)

This does, however, not constitute a very informative way of looking at the data. By instead following e.g. the LHCE series starting from the presumed HCE at 2.0 m (blue) and then the LHCE data points (green), a second maximum is present at 1.0 m salt concentration, i.e. at a much lower salt concentration, but perhaps with an intact HCE kind of Ca^{2+} first solvation shell, to be proven below. The same is true for all the LHCE concentration series, with the highest absolute ion conductivity obtained for the 1.0 m LHCE derived from the 1.5 m HCE. It is quite obvious that for a constant absolute concentration of the LHCEs, the ion conductivity is higher the lower the starting LE or HCE salt concentration is; perhaps not surprising, but for the first time quantified and visualized. Again, however, it does not tell about the relative cation and anion contributions to the total ion conductivities.

By using the viscosity data of all LHCEs (Table S1), the large decrease in viscosity with the addition of the TTE diluent is obvious for all electrolytes. Summarizing these effects in a Walden plot (Figure 9) provides the trends for the four LHCE families w r t also the parent LEs and HCEs.
Figure 9: Walden plot connecting the conductivities and viscosities for LEs, HCEs and LHCEs.

All four LHCEs families move further away from the KCl reference line upon dilution. The behaviour is a markedly more prominent for the LHCEs starting from the more concentrated electrolytes i.e. presumably HCEs. This reflects that the dilution reduces the ion conductivity by a similar factor for almost all salt concentrations, while decreasing the viscosity more and more. As can be seen directly in the ion conductivity data (Figure 8) within a LHCE family the dilution does not impact the ion conductivity much.

Thus these LHCEs, by virtue of the lower viscosities, overall hold promise for the practical aspects of wettability and porous electrode penetration, but is this accompanied by any synergy with the advantages of creating a HCE first place? The Raman spectra analysis, performed the same way as for the LEs and HCEs, are instrumental in order to elucidate this. To show the principle and indeed also that the main idea of the LHCEs having an intact cation first solvation shell of the parent HCE intact holds, we compare the neat PC, the 0.45 m LE, and the 3.256 m HCE spectra with the spectrum of the 0.45 m LHCE “derived” by dilution from the latter (Figure 10).
Figure 10: Raman spectra showing the principle of intact HCE local structure in an LHCE.

Indeed, even after this very large dilution, from 3.256 m to 0.45 m, *i.e.* a factor >7, the profiles of the Ca$^{2+}$-PC and Ca$^{2+}$-TFSI features of the LHCE do not change significantly as compared to the parent 3.256 m HCE. The lowered overall Raman intensity in this region is simply a dilution effect in terms of number density. Furthermore, no significant free PC or free TFSI peaks appear. In total the overall spectral profile is nowhere near that of the 0.45 m LE, why we can conclude that an LHCE is truly obtained with the local cation first solvation shell of the parent HCE conserved. To further substantiate this to be a general LHCE feature, a similar Raman analysis of ion-pairing and cation solvation as for the LEs and HCEs was performed for the different starting concentrations / LHCE families (Figure 11).
Looking at the TFSI data, the addition of diluent in fact causes minor increases in the Ca$^{2+}$-TFSI interactions. Up to 93% of the TFSI can be made Ca$^{2+}$-coordinated by creating a 0.75 m LHCE starting from a parent 3.256 m HCE. We can also tailor the (relative) amount of these cation-anion interactions rather freely by virtue of the ternary system, e.g. the 0.45 m LHCE from the 1.5 m HCE has 68% of the TFSI anion coordinated, which is more than for the 2.0 m HCE. That the dilution effect is (relatively) stronger for the less concentrated electrolytes as starting point is most likely an effect of the relatively larger concentration of TTE causing the overall permittivity of the electrolyte to drop. For the Ca$^{2+}$-PC interactions the tendency is roughly the reverse, but also much less pronounced (Figure 12).

**Figure 11:** The cation-anion interaction as function of electrolyte salt concentration and base HCE.
Figure 12: The cation-solvent interactions as function of electrolyte salt concentration and base HCE. Indeed, the relative amount of free PC increases somewhat upon dilution, but as the local Ca\textsuperscript{2+} first solvation shell is maintained, the relative amount of coordinated PC is overall still higher – and never goes below 55% (0.5 m). This is to be compared to the cross-over at 0.75 m for the HCEs (Figure 7). Furthermore, the changes are overall less pronounced for the LHCE families starting from the more concentrated HCEs (as above). It is not easy to directly reveal the molecular level “integrity” from these data, but they point to that the more concentrated the parent HCE is the more intact solvation shell.

Concluding Remarks

By combining several techniques, we have shown how the macroscopic performance relates to specific underlying molecular features and how the very transition from a conventional LE to a HCE, by different definitions and measures, can be tracked. In the future it would be intriguing to see how this affects the exact properties and prospects for different cell set-ups and cycling conditions using these types of electrolytes in practical CaB cells. In absolute numbers the ion conductivities at 0.75 m concentration should be a good target for Ca\textsuperscript{3+}-conducting electrolytes similarly made, i.e. using a Ca-salt with a weakly coordinating anion such as TFSI and a strong permittivity solvent such as PC. At the same time, we confirm that the HCEs have very high viscosities and that the on-set of these properties appear at lower concentrations than for monovalent cation based systems, and this might very well be a practical obstacle for any cell creation as well as for the ion transport, even if the latter perhaps can be remedied by the proven decoupling.

Taking on that problem, the “transfer” from HCE to LHCE was here shown to work also for Ca-based electrolytes as the integrity of the cation first solvation shell can be kept more or less intact at the same
time as the macroscopic properties can be varied – thus a second type of decoupling. While it is not easy to directly reveal the degree of the molecular level charge carrier integrity or the nature of charge carriers for neither HCEs nor LHCEs, it seems that the more concentrated the parent HCE is, the more intact the cation first solvation shell is.

While none of the electrolytes has the ion conductivity often stated to be needed for practical application, the altered charge carrier nature as compared to LEs, but similar to the HCEs, may render the LHCEs a better charge transfer at the electrolyte/electrode interfaces. Finally, it would be highly advantageous to be able to create LHCEs using diluents being non-fluorinated and less expensive if possible as e.g. TTE comes at ca. 45 $ for 5 g.

**Acknowledgements**

The funding from the European Union’s Horizon 2020 research and innovation programme H2020 FETOPEN-1-2016-2017 (CARBAT, grant agreement #766617) is gratefully acknowledged. FÅ and PJ are also grateful for the financial support provided from the Swedish Energy Agency by the project within “Batterifondsprogrammet”: “Highly Concentrated Electrolytes” (#39909-1). The continuous support to PJ from several of Chalmers Areas of Advance: Materials Science, Energy, and Transport, is also gratefully acknowledged. FÅ and PJ are in addition grateful for the computational resources provided by the Swedish National Infrastructure for Computing (SNIC) at Chalmers Centre for Computational Science and Engineering (C³SE).

**References**

2. https://e360.yale.edu/features/in-boost-for-renewables-grid-scale-battery-storage-is-on-the-rise


Paper II
Ion Transport Mechanisms via Time-Dependent Local Structure and Dynamics in Highly Concentrated Electrolytes

Rasmus Andersson,1,2 Fabian Årén,1 Alejandro A. Franco,2,3,4,5 and Patrik Johansson1,4,z

1Department of Physics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
2Laboratoire de Réactivité et Chimie des Solides (LRCs), Université de Picardie Jules Verne UMR CNRS 7314, Hub de l’Energie, 80039 Amiens, France
3Réseau sur le Stockage Electrochimique de l’Energie (RS2E), FR CNRS 3459, Hub de l’Energie, 80039 Amiens, France
4Alistore-ERI European Research Institute, CNRS FR 3104, Hub de l’Energie, 80039 Amiens, France
5Institut Universitaire de France, 75005 Paris, France

Highly concentrated electrolytes (HCEs) are attracting interest as safer and more stable alternatives to current lithium-ion battery electrolytes, but their structure, solvation dynamics and ion transport mechanisms are arguably more complex. We here present a novel general method for analyzing both the structure and the dynamics, and ultimately the ion transport mechanism(s), of electrolytes including HCEs. This is based on automated detection of bonds, both covalent and coordination bonds, including how they dynamically change, in molecular dynamics (MD) simulation trajectories. We thereby classify distinct local structures by their bond topology and characterize their physicochemical properties by statistical mechanics, giving both a qualitative and quantitative description of the structure, solvation and coordination dynamics, and ion transport mechanism(s). We demonstrate the method by in detail analyzing an ab initio MD simulation trajectory of an HCE consisting of the LiTFSI salt dissolved in acetonitrile at a 1:2 molar ratio. We find this electrolyte to form a flexible percolating network which limits vehicular ion transport but enables the Li$^+$ ions to move between different TFSI coordination sites along with their first solvation shells. In contrast, the TFSI anions are immobilized in the network, but often free to rotate which further facilitates the Li$^+$ hopping mechanism.

© 2020 The Author(s). Published by The Electrochemical Society. This is an open access article distributed under the terms of the Creative Commons Attribution Non-commercial No Derivatives 4.0 License (CC BY-NC-ND), which permits non-commercial reuse, distribution, and reproduction in any medium, provided the original work is not changed in any way and is properly cited. For permission for commercial reuse, please email: permissions@ioppublishing.org. [DOI: 10.1149/1945-7111/abc657]

Manuscript submitted June 22, 2020; revised manuscript received October 21, 2020. Published November 11, 2020.

*E-mail: rasmus.andersson@chalmers.se; patrik.johansson@chalmers.se

Lithium-ion batteries (LIBs) are today used for powering both mobile electronics and electric vehicles and may play an increasingly important role also in large-scale energy storage as our energy production systems are globally shifting towards variable renewable energy sources such as solar and wind power. For all these applications, electrolyte improvements are, even if not always recognized, in high demand due to both safety concerns and additional cost arising from their flammability and high vapor pressure.1 In addition, the electrolytes are not especially thermally stable, primarily as the ubiquitous PF6$^-$ anion auto-decomposes into toxic compounds.2 They are also not as electrochemically stable as needed at high voltages; today the aim of LIBs is set for ca. 5.0 V vs Li$^+/Li$.3

Unfortunately, the many requirements on LIB electrolytes means that their composition is hard to vary without impeding their functionality. Indeed, apart from the quite many different additives employed, the electrolyte composition has more or less remained unchanged since a few years after the commercialization of LIBs: 1 M LiPF6 dissolved in a solvent mixture of a cyclic carbonate, often ethylene carbonate (EC), and a non-cyclic carbonate, most often dimethyl carbonate (DMC).4 While these electrolytes have a balanced set of properties, they have been optimized to the point where minor changes cannot substantially improve their safety and stability properties. Thus, any novel high energy density LIB chemistry likely requires a conceptually different electrolyte. In this respect, while solid electrolytes are attracting a lot of interest at present due to their potential to eliminate flammability and enable higher energy density,5 many problems remain to be solved for these to be practical; low ionic conductivity, structural integrity, electrolyte-electrode contact, etc. At present, liquid electrolytes is still the main-track in practice.

One emerging liquid electrolyte concept that potentially can improve safety and performance of LIB electrolytes is highly concentrated electrolytes (HCEs), a.k.a. super-concentrated or solvent-in-salt electrolytes.6–8 HCEs are a class of electrolytes where the concentration of salt is comparable to the concentration of solvent in terms of molar ratio, resulting in ca. 3–5 M electrolytes. This change as compared to ca. 1 M traditional electrolytes, where the molar ratio typically is ≥10, has profound implications on the local structure in the electrolyte, as there is not enough solvent present to fully solvate the Li$^+$ ions. This also affects the electrochemical stability, as the local structure determines the vulnerability to oxidation/reduction, and thus determines both the size of the electrochemical stability window (ESW) as well as the electrolyte decomposition products,9 which in turn affects the ability to passivate the electrodes. Increasing the electrolyte salt concentration has been demonstrated to enable reversible cycling at both lower10,12 and higher13–15 potentials than the corresponding less concentrated electrolytes. First proposed for LIBs by McKimmon and Dahn as early as 1985,10 HCEs have gained much renewed interest in more recent years for both LIBs11,12,16–18 and lithium metal based batteries.19,20 HCEs can also be varied almost endlessly in composition,6,7 enabling their usage also for battery chemistries based on other cations than Li$^+$, e.g. Na$^+$ for sodium-ion batteries (SIBs).21–25

The mechanism(s) of cation transport in HCEs has been subject of much discussion, but surprisingly limited computational investigation.26–28 The higher degree of ion-ion interaction29,30 means that the vehicular ion transport mechanism3 of traditional electrolytes, in which fully solvated cations migrate with their first solvation shell at a rate limited by the electrolyte viscosity, i.e. following a Stokes-Einstein limited behavior, should not at all be dominant, or perhaps not even possible. This hypothesis is supported by experimental findings of higher transport numbers for Li$^+$ in HCEs as compared to traditional LIB electrolytes.31,32 Instead, ligand-exchange mediated mechanisms have been suggested by Okoshi et al.26 and Seo et al. have argued that the ionicity, i.e. the degree of cation-anion coordination, as well as the solvent and anion residence times should be as important as viscosity in explaining HCE transport mechanisms.27 Nilsson et al. have shown by a combination of experimental techniques that while species...
diffusivity tends to increase with decreased viscosity, ionic conductivity depends more strongly on ionicity, and also seems to be more predictive of rate capability in electrochemical cells. From a computational point of view, Åvall et al. have found that increasing salt concentration leads to larger, more varied solvation shells with lower energy barriers, which in turn enables faster kinetics, and probably ion transport. Altogether it is very clear that a more direct molecular-level understanding of the ion transport mechanism (s) in HCEs is needed, and this presupposes knowledge of both the structure and the dynamics of such electrolytes.

While continuum models of ion transport, e.g. the family of models originating in the work of Newman and co-workers, have been very useful in understanding the overall interplay between cell components and the effects of physicochemical properties on performance, they cannot be used to gain understanding of structure or dynamics on the molecular level. Thermodynamic lattice models, do have the potential to advance the fundamental understanding of HCEs, but they need to be supplemented with molecular level data to be accurately applied to specific electrolytes. Molecular dynamics (MD) simulations, on the other hand, is an indispensable tool for a better understanding of both structure and dynamics in liquids in general as well as for LIB electrolytes. In general, MD simulations provide large amounts of detailed data for a specific system, or family of systems, that can be analyzed statistically to e.g. elucidate solvation structure and dynamics as well as ion transport properties and mechanisms. Given that accurate MD simulations are performed long enough to capture the relevant dynamics, statistical analysis of MD trajectories can be used either to predict macroscopic properties, or conversely explain their molecular level origins. For both LIB and SIB electrolytes, MD simulations have been used to elucidate the structure of the first solvation shell of the cation and ligand angle distributions, as well as ionic diffusivities.

These kinds of analyses for standard LIB/SIB electrolytes are straightforward since the molecular structure is known a priori. However, the structure of the first solvation shell of the cation is less explanatory/useful when the cation is embedded in larger aggregates —as is arguably the case for all HCEs. Similarly, describing transport exclusively in terms of total ion diffusivity loses explanatory power when the transport is to a substantial extent mediated by changes in the local environments of the ions. When structures not known in detail and which furthermore change on time-scales relevant for transport processes determine the behavior, there is a need for more advanced post-analysis of MD simulation trajectories in order to elucidate the ion transport mechanism(s).

We here present a novel and general method for automatically detecting and characterizing the time-dependent structure, dynamics, and ion transport of HCEs (or in fact more or less any type of electrolyte), based on post-processing MD simulation trajectories. Our method is based on identifying bonds, including both covalent and coordination bonds, between atoms based on their dynamics, and subsequently classifying different local structures by their bond topology as subsets of the global bond graph. These topologies are finally characterized using statistical physics to evaluate their dynamical properties, e.g. mean lifetimes and contributions to Li\(^+\) ion diffusivity. The local bulk electrolyte processes we target here are essentially the same regardless of whether an external field is present or not, i.e. migration is approximated. This is a well-founded approximation as the fields created by an external cell potential would be several orders of magnitude weaker than the electrostatic interactions and activation energies/barriers controlling local ion transport events. Our method, while clearly a molecular level description, also starts to bridge the gap between molecular and macroscopic models, as urged by providing mechanistic understanding of the local structure and dynamical processes it can guide model design and clarify the realism of assumptions underlying macroscopic models.

We demonstrate our method by applying it to the well-studied HCE lithium bis(trifluoromethanesulfonyl)imide (LiTFSI) dissolved in acetonitrile (ACN) at 1:2 molar ratio of salt:solvent. Brouillette et al. and Seo et al. have both studied the structure of this electrolyte by combining differential scanning calorimetry with Raman spectroscopy and found it to be amorphous at all temperatures. However, the overlapping Raman bands made a full spectral deconvolution impossible and hence limited the information on the ion-ion and ion-solvent interactions. Both studies found roughly half of the TFSI anions to be coordinated to Li\(^+\). In contrast, Yamada et al. combining Raman spectroscopy and ab initio MD (AIMD), found almost all TFSI ions to be coordinated to Li\(^+\), giving rise to a liquid percolating network. Seo et al. also employed ab initio and density functional theory (DFT) calculations to determine solvation shell formation energies, as well as classical MD simulations, and found the TFSI anions to predominantly be monodentately coordinated to the Li\(^+\) cations. In a follow-up study, wherein they compared the ion transport properties of different electrolytes as function of salt concentrations, these did not clearly correlate negatively with the viscosity, strongly suggesting contributions from some kind of non-vehicular ion transport mechanism. They also evaluated the residence times of the anions and the solvent molecules in the first solvation shell of the Li\(^+\) cation based on MD simulation trajectories and these were typically on the order of hundreds of ps. Lundgren et al. later used the same HCE and by a combination of Newman style modelling and electrochemical characterization they found considerably higher cation transport numbers than for traditional LIB electrolytes.

In terms of increased electrochemical stability, Matsumoto et al. found the cathode Al current collector corrosion induced by conventionally concentrated LiTFSI based electrolytes to be suppressed at higher concentrations. They attributed this to preferential formation of passivating inorganic LiF on the Al surface, but this was questioned by McOwen et al., suggesting that the high TFSI concentration together with a low concentration of “free” solvent allows the positively charged surface to be essentially covered by passivating anions. This is somewhat reminiscent to the Chazalviel theory of high salt concentration passivation. On the anode side, Yamada et al. observed reversible Li\(^+\) intercalation into graphite, even without any solid electrolyte interphase (SEI) forming EC present in the electrolyte composition. They suggested the liquid percolating network to shift the LUMO from being ACN centered to TFSI anion and ligand angle distribution, as well as ionic diffusivities.

In what follows, we first describe our novel method for detecting, classifying and characterizing the local structures and dynamics in electrolytes. We then provide the computational details of the AIMD simulation that serves as input data for our analysis. We thereafter present and discuss our results in terms of structure and dynamics with a special emphasis on the Li\(^+\) ion transport.

Methods

Structure detection and dynamical characterization.—All chemical bonds, including also ion-ion and ion-solvent coordination bonds, in the AIMD simulation trajectory were identified using our recently developed software CHAMPION (Chalmers Atomic, Molecular, Polymeric & Ionic analysis toolkit), based on the dynamics of pairs of neighboring atoms. For a pair of atoms to be identified as bound, they need to oscillate about an equilibrium distance of each other over several vibrational cycles (Fig. 1, center). For coordination bonds, the times of formation and breakage were also identified where applicable. Subsequently, a time-dependent bond graph is constructed: an undirected colored graph where the nodes represent atoms and the edges represent bonds (Fig. 1, right).
The global structure is characterized by the connected components of the time-dependent bond graph (dashed ellipses around components in Fig. 1). Local environments of ions and molecules are identified by the subgraphs consisting of all atoms and bonds up to a selected number of bonds away from a central atom, e.g. Li⁺. The distinct types of connected components and local environments are classified by their subgraph topology, with each topology characterized statistically by sampling:

- The probability of atoms of the considered element to be in each topology at any given time,
- Their lifetime distributions, estimated by fitting a stretched exponential function for the survival probability as a function of lifetime,
- Transition rates, and
- The contributions to diffusive transport.

While we here focus on the Li⁺ cation, we also performed the same analysis for both TFSI and ACN (SI).

The total diffusivities of Li⁺, TFSI and ACN were determined using

\[ D_X = \frac{1}{6} \left( \frac{d}{dt} \Delta^2(t, t + \tau)^2 \right) \]  

where \( X \in \{ \text{Li}^+, \text{TFSI}, \text{ACN} \} \) denotes the species and \( k \) enumerates exemplars of \( X \), \( \Delta^2(t, t + \tau)^2 = (\vec{r}^2(t + \tau) - \vec{r}^2(t))^2 \) is the squared displacement of \( k \) between times \( t \) and \( t + \tau \), and the average goes over all exemplars, all times between 0 and \( T - \tau \) for a trajectory of length \( T \), and \( \tau \) goes between the diffusion onset time, defined as the interval beyond which the mean squared displacement (MSD) grows linearly as opposed to quadratically, and \( T \).

To study transport mechanism(s) we decompose the total diffusivities of the ions and solvent into additive contributions from each different local environment topology and the transitions between them, corresponding to vehicular and non-vehicular transport mechanisms broadly construed, as moving either along with or between them, corresponding to vehicular and non-vehicular transitions. We define

\[ \overline{v}_\delta^k(t) = \chi^j_k(t) v_{\text{com}}(t) \]  

and

\[ \overline{v}_\gamma^k(t) = \xi^j_k(t) (v^k(t) - \overline{v}_\delta^k(t)) \]  

as the diagonal and non-diagonal contributions to instantaneous velocity w.r.t. topologies, where \( \chi^j_k(t) \) is an indicator function active when \( k \) has local environment topology \( i \) and \( v_{\text{com}}^k(t) \) is the center-of-mass velocity of the local environment, and \( \xi^j_k(t) \) is an indicator function active between the midpoints of the lifetimes of two consecutive local environments of topologies \( i \) and \( j \) where \( j \) transitions to \( i \). By this construction, \( \overline{v}_\delta^k(t) \) and \( \overline{v}_\gamma^k(t) \) add up to the total velocity:

\[ v^k(t) = \sum_i \overline{v}_\delta^i(t) + \sum_i \overline{v}_\gamma^i(t) = \sum_i \overline{v}_\gamma^i(t) \]  

where one diagonal and one non-diagonal term of the right-hand side is non-zero at each value of \( t \).

The switching of \( \xi^j_k(t) \) at midpoints might seem arbitrary, but note that most of the non-vehicular contributions to squared displacement are expected to occur near transitions between environments, as the effects of e.g. rotational and vibrational motion within a topology tends to cancel out over time, whereas leading up to or following a structural rearrangement they may well contribute meaningfully to transport if they occur frequently, consider e.g. ion hopping mechanisms. Switching at midpoints minimizes the effects of the switching on assignment as the midpoints are the furthest removed from the transitions.

By this velocity decomposition, the contribution to \( X \) diffusivity due to transition from \( j \) to \( i \) (or due to vehicular motion of topology \( i \) if \( i = j \)) is given by

\[ D_{X,ij} = \frac{1}{3} (\Delta^2(t, t + \tau)^2 \cdot \overline{v}_\gamma^i(t + \tau))_{k,i,t,\tau} \]  

which is arrived at by expanding one of the terms in the square in Eq. 1 and carrying out the differentiation w.r.t. \( \tau \).

Since \( D_{X,ij} \) is an additive contribution to the total diffusivity, it is proportional to \( P_{ij} \), the probability of \( X \) having topology \( j \). Dividing by \( P_{ij} \) gives a measure of effective diffusivity, which allows mechanisms to be compared. For diagonal, i.e. vehicular, mechanisms, \( D_{X,ij}^{\text{veh}} = D_{X,ij}/P_{X,i} \) is a measure of how mobile the topology is in the system. For non-diagonal, i.e. non-vehicular, mechanisms, diffusivity contributions can be written as \( D_{X,ij} = P_{X,j} Q_{X,j} D_{X,j}^{\text{veh}} \) where \( Q_{X,j} \) is the (left-stochastic) transition rate matrix for
topologies of $X$ and $d_{X,j}$ is an effective distance traveled due to the transition, defined by this relation. Since $Q_{X,j}$ can be measured independently, $d_{X,j}$ can be determined using $d_{X,j} = D_{X,j}^{\infty}/Q_{X,j}$. Non-vehicular transport mechanisms can thus be studied in terms of how the composition of local environments affect $Q_{X,j}$ and $d_{X,j}$.

From a local structure point-of-view, in addition to the detailed analysis in terms of first solvation shell composition, we also analyzed, in a more coarse-grained fashion, the Li$^+$ solvation numbers (SNs): the total number of ligands, and coordination numbers (CNs): the total number of coordination bonds, i.e. taking into account the possibility of e.g. bidentate coordination. Similarly, but for the dynamics, we extracted the transport number $t^+$, i.e. the fraction of the Li$^+$ ion diffusivity to the total ionic diffusivity $D_{Li^+}/(D_{Li^+} + D_{TFSI})$, from the AIMD simulation trajectory.

**AIMD simulations.**—An AIMD simulation within the NVT ensemble was performed at 293 K using a Nosé-Hoover thermostat with $\omega_{ions} = 2250 \text{ cm}^{-1}$ and $\omega_{electrons} = 4500 \text{ cm}^{-1}$ using the Car-Parrinello MD (CPMD)$^{32}$ software and the Perdew–Burke–Ernzerhof (PBE)$^{33}$ functional. A system of 20 formula units (684 atoms in total) with a density of 1.48 g cm$^{-3}$ was simulated in a cubic periodic box with side 21.57 Å using a fixed time step of 0.1 fs for 6.33 ps, after first having equilibrated for 1.0 ps. The time needed for equilibration was found by considering the decay of the sum of pointwise squared errors of all partial radial distribution functions (RDFs) compared with their time average (Fig. S1 is available online at stacks.iop.org/JES/167/140537/mmedia). This deviation rapidly decreased for the first 0.7 ps, followed by fluctuations about a low average value $<10\%$ of the original deviation, with no observable trend towards further decrease. During the equilibration time, each ion and molecule settles into its free energy minimum with regards to its nearest neighbors. Based on this, the equilibration time was conservatively set to 1.0 ps. While we cannot exclude the possibility of further equilibration past the timescale of the simulation, the lack of a trend past 0.7 ps indicates any such processes to be orders of magnitude slower than the timescale of the simulation, and hence unreachable by AIMD simulations. To the extent that large-scale structural reorganization would be required to reach the global free energy minimum structure, we may expect our results to be somewhat biased towards entropy maximization rather than energy minimization, and consequently towards more rapid dynamics because of lower energy barriers to (de)complexation. While this bias may affect the relative proportion of the different structural species, and the absolute time-scale of dynamics and ion transport, it should not to first order affect the relative trends for the physicochemical properties among the structures as functions of their composition. The initial geometry was generated by randomizing and relaxing the positions and orientations of ions and molecules, by a conjugate gradient descent method w.r.t. a cost function designed to avoid overlap of atoms. The TFSI ions were initialized to 50% cis and 50% trans conformations, and pre-optimized using the Gaussian 16 software$^{55}$ at the M06–2X/6–311+G(d, p) level of theory.$^{56}$

**Results and Discussion**

We first describe the distribution of coordination bonds for Li$^+$ and TFSI ions and the global electrolyte structure this gives rise to, before giving a more detailed description of the distinct Li$^+$ first solvation shell topologies and their solvation and ion coordination dynamics. We then move on to the transport properties, assessing and comparing the vehicular and non-vehicular contributions to diffusive transport of cations, anions and solvent. From this we propose an overall qualitative view of the Li$^+$ ion dynamics and transport mechanism, which we then study more closely. Finally, we present in rank order the most important contributions to the ion transport in terms of Li$^+$ CNs and transitions between them and

![Figure 2. Distribution of (a) TFSI ions and ACN molecules coordinating to Li$^+$ and (b) Li$^+$ ions coordinating to TFSI.](image-url)
analyze correlations between solvation shell composition and Li\(^{+}\) diffusivity.

**Electrolyte structure.**—First of all, we consider the distribution of ligands around the Li\(^{+}\) ions and the distribution of Li\(^{+}\) ions coordinating the TFSI ions. The Li\(^{+}\) ions on average coordinate with 2.0 TFSI ions and 1.2 ACN molecules, thus giving a total average SN of 3.2 (Fig. 2a). While Seo et al. found a similar partial ACN SN (1.0) from their Raman data,\(^{30}\) their classical MD data resulted in SNs of 2.42 and 1.45, for TFSI and ACN, respectively. While much higher in absolute numbers, likely due to using different cut-off distances, our and their ratios of TFSI/ACN are however almost identical, both \(\approx 1.67\)—hence similarly composed local structure around the Li\(^{+}\) ions. Due to the possibility of bidentate coordination by the TFSI anions, the average Li\(^{+}\) CN is slightly greater than the SN, 3.5, and furthermore CN=4, typical for Li\(^{+}\) in organic solvents,\(^{4}\) holds for slightly more than half of the Li\(^{+}\) ions—to be explored together with the topologies and transitions further below (Fig. 7a).

Looking at the distribution of ligands, all but a very negligible fraction (0.4%) of the Li\(^{+}\) ions coordinate to at least one TFSI ion: roughly half of them to two anions, and a quarter of them each to one and three anions, respectively. While a similarly very tiny fraction (0.5%) coordinate to four different TFSI ions, there are thus very few “free” Li\(^{+}\) ions present at any stage—almost all are to some extent subject to ion-ion interactions. The first solvation shells are completed by ACN molecules for some Li\(^{+}\) ions, and these are more or less evenly distributed between zero and two ligands, ca. 30% each, with a small fraction (6.5%) solvated by three ACN molecules. From the perspective of the TFSI anion, the distribution is centered at 1.9 coordinating Li\(^{+}\) ions on average with a clear majority of TFSI anions coordinating to more than one Li\(^{+}\) ion, roughly a quarter with one and a negligible fraction is “free” (Fig. 2b).

The high degree of ion-ion interaction and coordination, i.e. each ion coordinates ca. two ions of the opposite charge, gives rise to a percolating network structure of alternating cations and anions in which ca. 90% of the ions participate on average (Fig. 3). The remaining ca. 10% exist in ion-pairs and finite clusters; most of the latter consist of one or two ions of each kind and none contain more than ten ions (Fig. 4). Notably, about three quarters of the finite clusters are charge neutral. While this overall shows a significantly higher degree of ion-ion interaction and coordination than reported by both Brouillette et al.\(^{29}\) and Seo et al.\(^{30}\) it is in good agreement with Yamada et al.\(^{9}\) The discrepancies may, at least in part, be explained by the reported difficulties encountered when trying to distinguish contact ion-pairs (CIPs) from solvent-separated ion-pairs (SSIPs) using Raman data.\(^{27,30}\)

Turning to the percolating network structure it can be understood in terms of a branching process. A Markovian process starting from a single Li\(^{+}\) ion and branching out according to the probability distribution of coordinating TFSI ions would in the first step equate the distribution shown in Fig. 2. A continued branching from each of the TFSI ions will give them a probability distribution of additional Li\(^{+}\) ion ligands with the condition of having at least one, at an expectation value of 0.989. In the next step, each Li\(^{+}\) ion coordinates to similarly conditioned TFSI ions with an expectation value of 1.025. Due to the Markovian assumption, pairs of consecutive branching steps can be combined into a single one with a branching factor of 0.989 \(\cdot\) 1.025 = 1.014, which then corresponds to the expectation value for the other Li\(^{+}\) ions sharing a TFSI ion with a given Li\(^{+}\) ion. Except for the first step, this is a Galton-Watson process, for which a well-known result is that there is a finite probability of indefinite survival, corresponding to forming a percolating network, when the branching factor is greater than 1. Our electrolyte is thus poised just above the percolation threshold.
and most likely a modest decrease in salt concentration would lead to a non-percolating structure, and hence quite different properties. The ACN molecules have little effect on the network structure apart from competitively coordinating Li\(^+\) ions; they are roughly evenly distributed between non-coordinating and coordinating a single Li\(^+\) ion, with very few (<2%) being shared between two Li\(^+\) ions (Fig. S5).

We further analyzed the electrolyte local structure in more detail by considering all Li\(^+\) first solvation shells; the fifteen most common topologies are shown in Fig. 5, while their probabilities, relative contributions to the Li\(^+\) diffusivity, effective diffusivity and mean lifetimes are presented in Table I. In the latter all values are presented as averages with an uncertainty (one standard deviation) also taking into account the statistical inefficiency of the sampling. The exact numbers and ordering should in general be interpreted with some caution due to the possible bias towards entropically favored structures and rapid dynamics discussed above, as well as the risk that the trajectory does not fully sample the ensemble.

The most common first solvation shell topology, accounting for roughly a fifth of the Li\(^+\) first solvation shells, consists of two TFSI ions and two ACN molecules, with both TFSI ions coordinating monodentately. The next two most common topologies consist of three TFSI ions and three TFSI ions and one ACN molecule, respectively, clearly illustrating the preferential TFSI coordination. Most of the coordination is monodentate by the sulfonyl oxygen atoms, in agreement with Seo et al.\(^{30}\) but similar bidentate coordination is also relatively common, while bidentate coordination by one oxygen atom and one nitrogen atom occurs less frequently, as does coordination only to the nitrogen atom or to a fluorine atom. The latter are also considerably shorter-lived. ACN always coordinates monodentately to Li\(^+\), as expected.

The Li\(^+\) first solvation shells have quite rapid solvation dynamics, with lifetimes mostly ranging between 0.2–0.9 ps. The Li(TFSI)(ACN)\(_2\) topology (\(i = 4\)) is an outlier, with a mean lifetime of >2 ps, but note that the uncertainty is very large. For individual species, the residence time of TFSI is on average ca. 6 ps and that of ACN is only slightly shorter (Table I). While these residence times are orders of magnitude shorter than those of Seo et al.,\(^{27}\) we attribute these differences in part to their very inclusive choice of cut-off distance, and in part to the possible bias towards rapid dynamics discussed above. Our results are, however, of the same order of magnitude as reported elsewhere for other HCEs.\(^{23}\)

The Li\(^+\) first solvation shell lifetime exhibits a moderate correlation (\(r = 0.43\)) with the total CN, reflecting the stability of CNs 3 and 4 (higher CNs are rarely encountered, Fig. 7a). Most of the variance can be explained by the variation in partial ACN CN (\(r = 0.38\)), reflecting its greater variance compared to the partial TFSI CN (Fig. 2a).

**Transport properties and mechanisms.**—An overall picture of the ionic and molecular transport is arrived at by considering the total diffusivities of the respective species, and the degree to which they are vehicular vs non-vehicular in nature (Fig. 6). The vehicular diffusivities of Li\(^+\) and TFSI are quite similar, which is expected since most ions are part of the percolating network, and three quarters of the finite clusters have equal numbers of cations and anions (Fig. S3). The slightly higher vehicular diffusivity of TFSI is likely due to the non-negligible fraction of “free” TFSI ions, ca. 5% (Fig. 2b), being more mobile than the Li\(^+\) ions (Tables I, SI), though the difference is within the margin of error—a comparable phenomenon as observed in ionic liquid based electrolytes by e.g. IR, Raman and NMR spectroscopies.\(^{37,38}\) However, the large fraction of non-vehicular transport, from fast solvation dynamics and high degree of ion association, affects Li\(^+\) and TFSI differently. The overall effect is that Li\(^+\) has a higher diffusivity and \(r^+ = 0.63\), in reasonable agreement with Lundgren et al. (ca. 0.72).\(^{31}\) In stark contrast, but as expected, ACN has a higher diffusivity that is almost exclusively due to vehicular transport, reflecting that a substantial fraction of the ACN molecules are free and much more mobile than the percolating network.

By observing animations of the trajectory, we can start to also qualitatively understand the reasons why the network structure facilitates a higher non-vehicular diffusivity for Li\(^+\). It is clear that the TFSI ions form the backbone of the percolating network structure, with the Li\(^+\) ions contributing with dynamic cross-links, so as to create a structure whose geometry and segmental motion are largely determined by the nature of the TFSI anions. The hard, in a HSAB sense,\(^{39}\) Li\(^+\) ion coordinates strongly to all its ligands and oscillates with large amplitude inside its solvation cage. Due to the large internal flexibility of the TFSI anion\(^{40}\) the network structure

---

**Figure 5.** The most common topologies around Li\(^+\) in order of probability. Element colors: purple: Li; red: O; blue: N; grey: C; white: H; yellow: S; green: F.
Table I. Properties of the most common topologies around Li$^+$. 

<table>
<thead>
<tr>
<th>i</th>
<th>Topology</th>
<th>Probability</th>
<th>$D_{Li^+}^{eff}/D_{Li^+}$</th>
<th>$D_{Li^+}^{eff}[10^{-3}m^2/s]$</th>
<th>Lifetime [ps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[Li(TFSI)$_2$(ACN)$_2$]$^-$</td>
<td>20.8 ± 0.6%</td>
<td>5.9 ± 0.4%</td>
<td>5.00 ± 0.16</td>
<td>0.82 ± 0.09</td>
</tr>
<tr>
<td>2</td>
<td>[Li(TFSI)$_3$]$^{2-}$</td>
<td>10.9 ± 0.5%</td>
<td>2.2 ± 0.2%</td>
<td>3.56 ± 0.17</td>
<td>0.44 ± 0.06</td>
</tr>
<tr>
<td>3</td>
<td>[Li(TFSI)$_2$(ACN)]$^+$</td>
<td>8.3 ± 0.3%</td>
<td>1.7 ± 0.1%</td>
<td>3.53 ± 0.14</td>
<td>0.28 ± 0.07</td>
</tr>
<tr>
<td>4</td>
<td>Li(TFSI)(ACN)$_3$</td>
<td>6.0 ± 0.4%</td>
<td>1.7 ± 0.2%</td>
<td>4.88 ± 0.34</td>
<td>2.18 ± 0.56</td>
</tr>
<tr>
<td>5</td>
<td>Li(TFSI)$_2$(ACN)$_2$</td>
<td>5.9 ± 0.4%</td>
<td>3.0 ± 0.2%</td>
<td>8.85 ± 0.62</td>
<td>0.87 ± 0.23</td>
</tr>
<tr>
<td>6</td>
<td>Li(TFSI)(ACN)$_2$</td>
<td>5.2 ± 0.2%</td>
<td>4.2 ± 0.3%</td>
<td>14.18 ± 0.60</td>
<td>0.18 ± 0.04</td>
</tr>
<tr>
<td>7</td>
<td>Li(TFSI)$_2$(ACN)</td>
<td>4.9 ± 0.4%</td>
<td>0.4 ± 0.1%</td>
<td>1.48 ± 0.14</td>
<td>0.49 ± 0.04</td>
</tr>
<tr>
<td>8</td>
<td>[Li(TFSI)$<em>2$(TFSI$</em>{2}$O$_2$)]$^-$</td>
<td>4.6 ± 0.3%</td>
<td>0.8 ± 0.1%</td>
<td>3.18 ± 0.22</td>
<td>0.59 ± 0.14</td>
</tr>
<tr>
<td>9</td>
<td>[Li(TFSI)$_3$]</td>
<td>4.2 ± 0.3%</td>
<td>0.5 ± 0.1%</td>
<td>2.06 ± 0.17</td>
<td>0.23 ± 0.05</td>
</tr>
<tr>
<td>10</td>
<td>[Li(TFSI)$<em>2$(TFSI$</em>{2}$O$_2$)(ACN)]$^-$</td>
<td>3.5 ± 0.3%</td>
<td>0.2 ± 0.04%</td>
<td>0.76 ± 0.08</td>
<td>0.34 ± 0.06</td>
</tr>
<tr>
<td>11</td>
<td><a href="TFSI$_%7B2%7D$O$_2$">Li(TFSI)$_3$</a>(ACN)</td>
<td>2.2 ± 0.2%</td>
<td>1.2 ± 0.1%</td>
<td>9.30 ± 17.2</td>
<td>0.19 ± 0.03</td>
</tr>
<tr>
<td>12</td>
<td>[Li(TFSI)$<em>2$(TFSI$</em>{2}$O$_2$)]$^-$</td>
<td>2.1 ± 0.2%</td>
<td>0.7 ± 0.02%</td>
<td>5.54 ± 0.56</td>
<td>0.74 ± 2e-4</td>
</tr>
<tr>
<td>13</td>
<td>[Li(TFSI)$<em>2$(TFSI$</em>{2}$O$<em>2$)(TFSI$</em>{2}$O)]$^-$</td>
<td>2.0 ± 0.2%</td>
<td>0.04 ± 0.09%</td>
<td>0.35 ± 0.19</td>
<td>0.33 ± 1e-8</td>
</tr>
<tr>
<td>14</td>
<td>[Li(TFSI)$<em>2$(TFSI$</em>{2}$O$_2$)(ACN)]$^-$</td>
<td>1.7 ± 0.2%</td>
<td>0.04 ± 0.07%</td>
<td>0.42 ± 0.18</td>
<td>0.23 ± 0.03</td>
</tr>
<tr>
<td>15</td>
<td>Li(TFSI)(ACN)</td>
<td>1.6 ± 0.1%</td>
<td>0.8 ± 0.2%</td>
<td>9.04 ± 0.74</td>
<td>0.34 ± 2e-9</td>
</tr>
</tbody>
</table>

Figure 6. Diffusivities of electrolyte species partitioned into vehicular and non-vehicular contributions. The error bars denote 95% confidence intervals for the total diffusivities.

Continually deforms, leading to some drift in the absolute position of the Li$^+$ ion. This contributes to the vehicular transport that affects Li$^+$ and TFSI equally and occasionally facilitates the formation of new coordination bonds. Sometimes, the electrostatic environment of a Li$^+$ coordination bond weakens it enough for a ligand to drift away without giving rise to a strong recoil and thereby the topology, CN and/or SN can be altered, additionally by new ligand(s): nearby TFSI anions and/or free ACN. A new strong Li$^+$ first solvation shell is formed along with a new equilibrium position. When the new ligand is a TFSI ion, most already coordinated to one or more Li$^+$ ions in the percolating network and thus immobilized, the event is more pronounced. The large and flexible TFSI, which usually is constrained by coordination bonds, rotates and deforms to bring the sulfonyl oxygen atoms closer to Li$^+$, which on the other hand, only has a single translative degree of freedom. Overall, all this sums up as a structural diffusion event. This mechanism is facilitated by the large number and angular diversity of possible coordination sites of the TFSI anion (due to its highly delocalized negative charge and extensive conjugation), and furthermore its internal flexibility enables new coordination bonds to form frequently. This is supported by Åvall et al.$^{22,23}$, showing that a higher CN variance is directly associated with faster ligand exchange.

We next consider the distribution of Li$^+$ CNs, as well as the transitions between them, both the rates and the distances moved (Fig. 7). More than half of the Li$^+$ ions have CN = 4, a third has CN = 3, and other CNs are much less prevalent, consistent with the former being more stable in terms of less negative transition rates along the diagonal. For the non-diagonal transition rates, there is a clear bias of probability flux towards the most stable CNs 3 and 4 compared to the fluxes from each CN in the other direction. The transition rates towards CNs 3 and 4 also increase with the distance from the most stable CNs.

Turning instead to the contributions to Li$^+$ diffusivity, over 80% of the transport is accounted for by the vehicular transport with CNs 3 and 4, as well as the structural transport due to the transitions between them, with most of the rest being due to transitions between CNs 3 and 2 (Fig. 8). This can be understood by considering the individual factors giving rise to $D_{Li^+}^{eff}$ (Fig. 7). Recalling that $D_{Li^+}^{eff} = P_{X}D_{X}A_{X}q_{X}d_{X}$ explains why CN 4 makes a large contribution, due to its high probability and the same applies to CN 3, but to a lesser extent. This is the main reason why both vehicular and non-vehicular transport of CNs 3 and 4 dominate over other Li$^+$ ion transport mechanisms. In addition, the transitions between these CNs are also connected to the largest distances moved (Fig. 7c), though this effect is less pronounced.

Finally, we assess the ion transport mechanisms in more detail by considering them in terms of the full structure of the first Li$^+$ solvation shell. Starting with an overview, the 13 most important transport events/mecanisms (having contributions $>2\%$ to the total) are ranked by their contribution to the Li$^+$ diffusivity (Fig. 9). Although these in total account for less than one-third of all Li$^+$ transport, they serve to illustrate some typical events/mechanisms and underscore the diversity of complex electrolytes. There are representatives of both vehicular and structural transport, and among the structural both exchanges of TFSI ions and ACN molecules as well as changes in the TFSI coordination.

Looking more closely at the vehicular transport mechanisms, Table I, we find a positive correlation coefficient $r = 0.36$ between the effective diffusivity $D_{Li^+}^{eff}$ and the number of ACN molecules in the first Li$^+$ solvation shell, and a negative correlation coefficient $r = -0.42$ between $D_{Li^+}^{eff}$ and the number of TFSI ions. Since these correlations almost cancel out, there is no strong correlation between effective diffusivity and total SN. From this we hypothesize that the
negative correlation is due to the TFSI immobilization in the percolating network structure, whereas the ACN blocks TFSI coordination. Furthermore, looking at the animated trajectory, the coordinated ACN tends to slightly change its orientation as a response to external steric hindrance during vehicular diffusion, thus limiting the slowing down else affecting all of the Li\textsuperscript{+} first solvation shell.

The network structure itself makes the vehicular Li\textsuperscript{+} transport slow, but as very local fluctuations cause coordination bonds to occasionally break, this enables Li\textsuperscript{+} ions to re-coordinate to e.g. a nearby TFSI ion, which while being largely immobilized in the network can rotate and deform. This way the network structure promotes diffusive transport of Li\textsuperscript{+} to a much larger extent than for TFSI, leading to an observable increase in the Li\textsuperscript{+} transport number—a common feature of HCEs that we thus here can address at the molecular level.

Based on this understanding, to maximize the Li\textsuperscript{+} transport number, an HCE should be composed of large and flexible anions, such as TFSI, with many angularly distributed available coordination sites and in addition solvent molecules with similar ion-ion and ion-solvent interaction strengths with Li\textsuperscript{+} to enable frequent ligand exchange. However, it is far from obvious how to improve these features beyond the TFSI anion—as its internal flexibility, negative charge delocalization\textsuperscript{60} and versatile coordination geometry are all truly outstanding.

In a real operating battery, with electric fields and concentration gradients, the situation might differ somewhat. While these simulations and analysis do not account directly for migration, as in a true battery cell, the local events determining the ion transport in the bulk of the electrolyte are essentially the very same—but without a directional bias. Likewise, the situation at the two electrolyte/electrode interfaces/interphases likely differs from what we here can conclude from our bulk HCE steady-state AIMD simulations.

Conclusions

We found that our model HCE forms a flexible percolating network structure of alternating cations and anions including most of the ions, in which the Li\textsuperscript{+} ions are partially solvated by ACN, but preferentially coordinate to the TFSI anions. The nature of the TFSI anion largely determines the overall structure and flexibility of the network, why these properties may change drastically for e.g. an HCE based on LiPF\textsubscript{6}. In light of the Galton-Watson theory for branched processes, our model system is just above the percolation threshold and thus slightly lowering the salt concentration would likely suffice to radically alter the global electrolyte structure and in turn the ion transport mechanism(s).

The network structure itself makes the vehicular Li\textsuperscript{+} transport slow, but as very local fluctuations cause coordination bonds to occasionally break, this enables Li\textsuperscript{+} ions to re-coordinate to e.g. a nearby TFSI ion, which while being largely immobilized in the network can rotate and deform. This way the network structure promotes diffusive transport of Li\textsuperscript{+} to a much larger extent than for TFSI, leading to an observable increase in the Li\textsuperscript{+} transport number—a common feature of HCEs that we thus here can address at the molecular level.

Based on this understanding, to maximize the Li\textsuperscript{+} transport number, an HCE should be composed of large and flexible anions, such as TFSI, with many angularly distributed available coordination sites and in addition solvent molecules with similar ion-ion and ion-solvent interaction strengths with Li\textsuperscript{+} to enable frequent ligand exchange. However, it is far from obvious how to improve these features beyond the TFSI anion—as its internal flexibility, negative charge delocalization\textsuperscript{60} and versatile coordination geometry are all truly outstanding.

In a real operating battery, with electric fields and concentration gradients, the situation might differ somewhat. While these simulations and analysis do not account directly for migration, as in a true battery cell, the local events determining the ion transport in the bulk of the electrolyte are essentially the very same—but without a directional bias. Likewise, the situation at the two electrolyte/electrode interfaces/interphases likely differs from what we here can conclude from our bulk HCE steady-state AIMD simulations.
Finally, the system size and trajectory length simulated are both too limited to ascertain the global structure and detailed statistics to be representative of the real bulk system, and larger and longer time-scale simulations would clearly be beneficial. Nevertheless, based on the fast dynamics of the first solvation shell (Table I), and the observed rapid equilibration of partial RDFs (Fig. S1), we are confident that the local structure and dynamics presented here are strongly indicative of the real system structure and dynamics.

In summary, our approach to understanding complex electrolytes, such as HCEs, by automatically detecting dynamic bonds and characterizing the structure and dynamics in terms of bond graph topology and statistical physics is a promising way forward towards rational design of electrolytes. The analysis method that we present here could also rather easily be extended to cover both longer simulation times and cover wider length-scales, to study global ion transport phenomena with better statistics. The most straightforward approach would be to replace our AIMD with classical MD. Another route would be to coarse-grain the electrolyte and apply simulation methods targeting the mesoscale such as kinetic Monte-Carlo.61 Indeed, our method should be well suited as one of the bottom layers in a comprehensive multi-scale approach for modelling electrochemical cells. We also expect that the methodology should generalize well to both similar application areas and other types of materials.

Acknowledgments

The research presented has received funding through the HELIS project (European Union’s Horizon 2020 research and innovation program under Grant Agreement No. 666221) and the Swedish Energy Agency grants (#P43525-1 and #P39909-1). P. J. would also like to acknowledge several of Chalmers Areas of Advance: Materials Science, Energy, and Transport, for continuous support, and specifically the Theory and Modelling scheme of Advanced User Support to R.A. A.A.F. acknowledges the Institut Universitaire de France for the support.

ORCID

Rasmus Andersson @ https://orcid.org/0000-0001-9359-1994
Patrik Johansson @ https://orcid.org/0000-0002-9907-117X

References


Figure 9. Most important Li⁺ transport events/mechanisms in rank order. Element colors: purple: Li, red: O, blue: N, grey: C, white: H, yellow: S, green: F.

Table II. Properties of the most common transitions between local environments of Li⁺.

<table>
<thead>
<tr>
<th>ID</th>
<th>j i</th>
<th>Description</th>
<th>( D_{j,i}^+/D_{i,j}^+ )</th>
<th>( Q_j [\text{ps}^{-1}] )</th>
<th>( d_j [\text{Å}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3 1</td>
<td>([\text{Li(TFSI)}_2(\text{ACN})]^- + \text{ACN} \rightarrow [\text{Li(TFSI)}_3(\text{ACN})]^-)</td>
<td>5.9 ± 0.6%</td>
<td>1.44 ± 0.55</td>
<td>0.47 ± 0.10</td>
</tr>
<tr>
<td>2</td>
<td>2 9</td>
<td>([\text{Li(TFSI)}_3]^- \rightarrow [\text{Li(TFSI)}_2]^- + \text{TFSI}^-)</td>
<td>3.8 ± 0.7%</td>
<td>0.63 ± 0.16</td>
<td>0.23 ± 0.11</td>
</tr>
<tr>
<td>3</td>
<td>6 1</td>
<td>([\text{Li(TFSI)}(\text{ACN})_2 + \text{TFSI}^- \rightarrow [\text{Li(TFSI)}_2(\text{ACN})]^-)</td>
<td>3.7 ± 0.5%</td>
<td>3.12 ± 0.90</td>
<td>0.34 ± 0.06</td>
</tr>
<tr>
<td>4</td>
<td>1 6</td>
<td>([\text{Li(TFSI)}_2(\text{ACN})_2]^- \rightarrow [\text{Li(TFSI)}(\text{ACN})_2 + \text{TFSI}^-)</td>
<td>3.3 ± 0.6%</td>
<td>0.39 ± 0.08</td>
<td>0.54 ± 0.09</td>
</tr>
<tr>
<td>5</td>
<td>6 5</td>
<td>([\text{Li(TFSI)}(\text{ACN})_2 \rightarrow \text{Li(TFSI)(ACN)}_2)</td>
<td>2.6 ± 0.6%</td>
<td>0.89 ± 0.48</td>
<td>0.51 ± 0.21</td>
</tr>
<tr>
<td>6</td>
<td>6 4</td>
<td>([\text{Li(TFSI)}(\text{ACN})_2 + \text{ACN} \rightarrow \text{Li(TFSI)(ACN)}_2)</td>
<td>2.6 ± 0.5%</td>
<td>0.89 ± 0.48</td>
<td>0.51 ± 0.20</td>
</tr>
<tr>
<td>7</td>
<td>9 2</td>
<td>([\text{Li(TFSI)}_2]^- + \text{TFSI}^- \rightarrow [\text{Li(TFSI)}_3]^-)</td>
<td>2.2 ± 0.7%</td>
<td>2.29 ± 0.71</td>
<td>0.27 ± 0.09</td>
</tr>
<tr>
<td>8</td>
<td>15 3</td>
<td>([\text{Li(TFSI)}(\text{ACN}) + \text{TFSI}^- \rightarrow [\text{Li(TFSI)}_2(\text{ACN})]^-)</td>
<td>2.1 ± 1.6%</td>
<td>2.30 ± 0.29</td>
<td>0.37 ± 0.17</td>
</tr>
<tr>
<td>9</td>
<td>1 3</td>
<td>([\text{Li(TFSI)}_2(\text{ACN})]^- \rightarrow [\text{Li(TFSI)}_3(\text{ACN})]^- + \text{ACN})</td>
<td>1.9 ± 0.5%</td>
<td>0.39 ± 0.08</td>
<td>0.38 ± 0.07</td>
</tr>
</tbody>
</table>
Paper III
Learning Force Fields from *Ab initio* Data

Fabian Årén¹², Rasmus Andersson¹² and Patrik Johansson¹²³

Correspondence to: Fabian Årén (E-mail: fabian.aaren@chalmers.se)

¹ Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden
² Compular AB, SE-412 96 Göteborg, Sweden
³ ALISTORE-European Research Institute, CNRS FR 3104, Hub de l’Energie, 15 Rue Baudelocque, 80039 Amiens, France

ABSTRACT

The research and development of modern battery electrolytes, with their vast compositional space and intricate ionic interactions, making structure-property predictions difficult, would truly benefit from fast and reliable modelling approaches. One such recent development is machine learning (ML) technologies, but those generally require vast amounts of data. Here a new, light-weight ML scheme is presented using look-up tables to substitute the classical force fields used in molecular dynamics simulations. Together with the CHAMPION bond graph identifying method, this enables the creation of reactive force fields, which we apply to the most common lithium-ion battery electrolyte. We highlight herein both the basics and development of the method as such, including the caveats and the problems we have at present.

Introduction

The challenge of climate change pushes for more renewable energy to be installed and alongside this, climate neutral energy storage solutions is a must¹² in order to mitigate the variable nature of most renewable energy sources. One of the most versatile and efficient large-scale energy storage solutions is lithium-ion batteries (LIBs)³, but for most beneficial impact these would need to be further optimized in terms of cost, environmental impact, and resource abundancy⁴⁵. There are also emerging a flora of next generation batteries (NGBs) aiming to either complement or replace LIBs, depending on the application niche⁶–⁸. Regardless of LIBs or NGBs, more or less all parts of the batteries and especially the electrochemical cells must be improved, e.g. with regard to energy density, environmental friendliness in order to allow and power an environmentally sustainable society. While much of the focus has been on cobalt mining⁹ and emergent risks of lacking both lithium and natural graphite¹⁰, less attention has been paid to the electrolytes employed. Nevertheless, being one of the main battery components, the electrolyte has severe impact both on environmental friendliness directly, through production processing steps needed and resource usage, and indirectly through battery life-length and total energy throughput. It also affects the technological application via e.g. the energy density by its electrochemical stability window (ESW), the power outtake possible by its mass transport limits, the operating temperature by its (in)stability, and safety by its flammability and toxicity. Today, much of the electrolyte development is still trial-and-error based, even if there are efforts to alter this paradigm, such as the BIG-MAP¹¹ and Materials Genome¹² initiatives.

Ample use of advanced modelling can improve this workflow; e.g. methods such as density functional theory (DFT) calculations and molecular dynamics (MD) simulations enable more efficient studies of molecular level properties to approximate and predict relevant macroscopic properties compared to only applying experimental techniques, reducing the need for costly experimental studies and increasing the return on investment (ROI) per experiment. Yet, for modern battery electrolytes DFT calculations are too computationally costly, while classical MD
simulations are laden by the fundamental drawback that the accuracy is heavily dependent on the force field (FF) used. Hence neither method is able to capture the complexity in structure stemming from electronic multipole interactions playing a key role in system dynamics. As an alternative to FFs, first principle methods can be used to calculate the forces needed to propagate the system e.g. ab initio MD (AIMD). Such methods are temporally limited to ca 100 ps, spatially limited to a couple of nm, and numerically limited to thousands of particles\(^{13}\). Hence, even though more accurate, this altogether puts pragmatic limits to the simulation sizes attainable. The limited simulation size in turn restricts the sampling of the phase space possible at reasonable computational costs using state of the art methods, leaving these methods inadequate for many situations. There exists a few special FFs and methods highly useful both to simulate battery electrolytes by classical MD, such as Borodin’s APPLE\&P\(^{14}\), and to use effective atomic charges with a conventional FF, such as OPLS-AA\(^{15}\). Yet, these are non-general and using these approaches for any new (types of) electrolytes will always require tailor made FFs. This is especially true for electrolyte concepts such as highly concentrated electrolytes (HCEs)\(^{16}\) wherein electrostatic interactions to a larger extent determine the properties. For such systems, e.g. higher order multipoles, not handled well by traditional FFs, are crucial to handle properly\(^{14,17}\).

With the recent popularity of machine learning (ML) methods, also a number of alternative ways to develop FFs for classical MD simulations have emerged; the neural network (NN) potentials developed by W. Wang et al.\(^{18}\), based on the SchNet deep learning networks enabling learning chemically plausible embeddings for elements\(^{19}\), and other similar approaches\(^{20–22}\) serve to substitute or recreate FFs. Similar approaches have indeed been made since the early 2000’s\(^{23}\). Due in large part to the extensive flexibility in terms of molecular structure of liquid electrolytes, NN and similar methods have a hard time extracting structural, electronic, and energetic information from these systems as compared to more structured systems such as solid-state electrolytes (SSEs), yielding few examples of these methods used for such purposes\(^{24}\). In contrast, for SSEs there is a plethora of examples using NNs, and ML in general, to predict dynamical, structural, electronic, and energetic properties\(^{22,25,26}\). Nevertheless, Knijff and Zhang\(^{27}\) recently showed how the molecular dipole moments of an atomistically simulated liquid can be learned by a NN method, more specifically PiNet\(^{28}\), being an inherently permutation invariant model. This suggests that ML methods can be applied to also capture the complex multipole interactions seen in HCEs. Hence, properly applied, these methods can recreate the dynamics, material properties, and statistical mechanics of the systems they are trained on. Some other successful approaches to using ML methods for material simulations are the Behler and Parinello (BP) symmetry functions (SF) taking advantage of chemical locality to make transferable NN potentials\(^{29}\) and the smooth overlap of atomic positions (SOAP) developed by Bartók et al.\(^{30,31}\) defining the similarity between two neighborhood environments. A common feature, however, for most ML approaches, is that they require large amounts of data for training and testing.

Herein we propose a novel and light-weight method for calculating the forces acting upon particles in a MD simulation and replace the need for any conventional FF like those used in AMBER\(^{32}\) and GROMACS\(^{33}\). We do this by generalising the method developed by Åvall and Johansson\(^{34}\), wherein the ion-solvent and ion-ion interactions are expressed as heat maps of radial forces. Using a set of generalised coordinates, such as found in common FFs, describing 2-, 3-, and 4-body as the basis allow us to create an effective alternative. We start by simulating a small electrolyte system using AIMD and bin the generalised forces associated with a set of generalised coordinates into a look-up table. This, ideally, renders the computational cost of an MD simulation several orders of magnitudes below even classical MD simulation as using look-up tables are
computationally less costly than even the bonded terms of conventional FFs, but also as all electrostatic long-range interactions should be captured by the look-up tables, alleviating the need for e.g., the computationally costly Ewald summation and variations thereof\textsuperscript{35}. We also high-light the caveats of the approach and the mitigation strategies applied.

**Methods**

Our approach starts from locating and enumerating all types of bonds, bond-angles, proper and improper dihedrals by the CHAMPION software\textsuperscript{36}, a codebase to dynamically determine the global time dependent bond graph of atomistic trajectories, also including e.g. ion-solvent and ion-ion interactions. This enables Åvall plots (Fig. 1) to be created for all individual interactions. Such an interaction could e.g. be a pairwise C-C atom-atom interaction. The Åvall plot will then contain information on the distribution of forces acting on the central C atom, as a function of the distance to the set of all other C atoms. This can then be generalised to bonded 2-, 3-, and 4-body interactions of the kinds commonly treated by conventional FFs.

![Figure 1](image.png)

Figure 1. a) A true heat-map based Åvall plot\textsuperscript{34}. Reproduced with the permission of AIP Publishing. b) A schematic Åvall plot linking the probability distribution of generalised forces to a generalised coordinate.

For 2-body interactions the independent variable, constituting the generalised coordinate, is the bond distance, for 3-body interactions the angle formed around the central atom in the chain, and for the 4-body interactions the angle formed between the planes spanned by the chain or improper connections to the central atom. Generalising this method further, the interaction can be defined from CHAMPION defined force centre types. A force centre type is defined by the local environment around the central atom, hence a distinction is made between e.g. linearly bonded C-C atoms and circularly bonded C-C atoms.

The overall hypothesis is that the system dynamics can be recreated only using the mean values of the generalised force distributions, at a given independent variable value. From the positions of and forces acting on each atom, both the generalised coordinates and forces can be computed and sampled. The generalised force distribution as functions of the corresponding generalised coordinates result from a combination of the effective interaction between the atoms involved and effects of the background. The approach taken here is based on assuming the background effect to be unbiased noise, so that the mean of the sampled distributions estimates the true effective interaction. This is based on the following: Given a force distribution $F|_q$ of generalised
forces $F_i(q)$ at a given value of the generalised coordinate $q$ evaluated based on projecting the forces acting on the involved atoms onto the respective generalised coordinates, $F_i(q)$ can be written as $F_i(q) = F(q) + \Delta f$, where $F(q)$ is the true generalised force of the interaction and $\Delta f$ is a stochastic noise due to the background. Assuming a normal and unbiased background noise

$$\sum_i F(q) = \sum_i F_i(q) + \sum_i \Delta f \Rightarrow F(q) = \langle F_i(q) \rangle$$

where $N$ is the total number of forces, and given a uniform error, the mean is equal to the mode. Given a non-uniform distribution we would still expect $\bar{F}(\varphi)$ to be the dominant feature and hence being the mean value of this distribution. If the error is not Gaussian, the underlying data is unphysical and this method is moot, hence this case will not be considered.

The above analysis started from AIMD simulations produced using CPMD\textsuperscript{37}. A random 1 lithium-bis(trifluoromethane)sulfonimide (LiTFSI) dissolved in 16 acetonitrile (ACN) system was set up using the CHAMPION system builder. The system was equilibrated for 0.83 ps and then a production run was obtained at 293 K using the PDB functional for 5.0 ps. Using the output trajectories and forces as input to the CHAMPION analysis, all bonds, bond-angles, proper and improper dihedrals were identified and arranged according to their constituent atom types. Using these ledgers of bond-types, look-up tables of the average forces as well as standard deviations were produced. From this 2D histograms of the generalised forces as function of the generalised coordinate were created accordingly, defined and projected for the given bond-type as described below.

Bonds between two atoms are projected by $\ell_{ij} = |\vec{r}_j - \vec{r}_i|$ and the generalised force is defined according to

$$Q_j = \sum_{i=1}^{n} F_i \cdot \frac{\partial \vec{r}_i}{\partial q_j}$$

Similarly, bond-angles are projected as $\cos \theta_{ijk} = \frac{(\vec{r}_{ij} \cdot \vec{r}_{jk})}{m}$, proper, and improper dihedrals coordinates are projected by creating the intermediate vectors spanning the planes $\vec{m} = \vec{r}_{ij} \times \vec{r}_{jk}$, and $\vec{n} = \vec{r}_{ik} \times \vec{r}_{jk}$, then finding the angle between them through $\cos \phi_{ijkl} = \frac{\vec{m} \cdot \vec{n}}{mn}$, $\sin \phi_{ijkl} = \frac{(\vec{m} \times \vec{n}) \cdot \vec{r}_{jk}}{mnr_{jk}}$, then $\phi_{ijkl} = -\arctan\left(\frac{\sin\phi_{ijkl}}{\cos\phi_{ijkl}}\right)$. Given these 2D histograms it is straight-forward to create look up tables of the mean force acting on particle $i$ at a given coordinate.

To calculate the force acting on each particle partaking in an interaction the general force identified using the tables is then returned to Cartesian coordinates using virtual work.

$$\delta W = \sum_{i=1}^{N} F_i \cdot \delta r_i = \sum_{j=1}^{m} f_j \delta q_j$$

Using bond pairs as an example:

$$r = \left((x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2\right)^{1/2} = \frac{x_2 - x_1}{r}$$

$$\frac{\partial r}{\partial x_1} = \frac{1}{2} \left[ (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 \right]^{-1/2} \frac{2(x_2 - x_1)}{r}$$

in a similar manner:

$$\frac{\partial r}{\partial x_2} = \frac{x_2 - x_1}{r} , \frac{\partial r}{\partial y_2} = \frac{y_2 - y_1}{r} , \frac{\partial r}{\partial z_2} = \frac{z_2 - z_1}{r}$$
\[ \delta W = F_1 \cdot (\delta x_1, \delta y_1, \delta z_1) + F_2 \cdot (\delta x_2, \delta y_2, \delta z_2) \]
\[ = F_1 \cdot \left( \frac{x_2 - x_1}{r} \left( \frac{y_2 - y_1}{r} + \frac{z_2 - z_1}{r} \right) \right) + F_2 \cdot \left( \frac{x_2 - x_1}{r} \left( \frac{y_2 - y_1}{r} + \frac{z_2 - z_1}{r} \right) \right) \]
\[ = \frac{1}{r} \left( (x_2 - x_1, y_2 - y_1, z_2 - z_1) \right) \left( F_2 - F_1 \right) \]

Similarly, this method is used to obtain force transformations for 3- and 4-body bonded terms.

Another route is to generate the forces acting on a bond type by generating the force using a numerical differentiation of the second order of the transformed coordinates. This method is effective if only coordinates are available from simulations, enabling for an ease of use when sharing data e.g. Here \( Q_j = \mu \ddot{q} \) where \( \mu \) is the reduced mass \( \mu = \left( \frac{1}{\sum \frac{1}{m_i}} \right) \), and \( \ddot{q} \) is the second time derivative of any of the coordinates belonging to the set \( q \in \{ r, \theta, \phi_{prop}, \phi_{improp} \} \).

All the results presented below comes from this second route, while the former route still is being evaluated and implemented.

### Results & Discussion

Note: At present the results contained and presented herein are still under development. This also underpins the discussion being made – that more focusses on the problems remaining than the (partial) solutions obtained and observations possible to make for the model systems at hand.

We start by creating a baseline FF by identifying the pair-wise interactions between all atom types, completely analogous to the Åvall method. From this starting point it is easy to build upon and add more complex interaction terms.

Table 1. A list of how the sample time \( \Delta t \), the number of time step \( N_{\text{time step}} \), the graph radius \( d \), affect the statistical inefficiency \( s \).

<table>
<thead>
<tr>
<th>( \Delta t [\text{fs}] )</th>
<th>( N_{\text{time step}} )</th>
<th>( d )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0</td>
<td>60218</td>
<td>3</td>
<td>1.869</td>
</tr>
<tr>
<td>0.1</td>
<td>1000</td>
<td>3</td>
<td>22.7221</td>
</tr>
<tr>
<td>0.1</td>
<td>10000</td>
<td>3</td>
<td>72.0633</td>
</tr>
<tr>
<td>0.1</td>
<td>10000</td>
<td>1</td>
<td>35.3487</td>
</tr>
<tr>
<td>0.1</td>
<td>5000</td>
<td>1</td>
<td>21.1608</td>
</tr>
</tbody>
</table>

When discussing the quality of these FFs the statistical inefficiency \( s \) is an important factor to consider and is a measure of how many time steps a certain correlation exists (Fig. 1).

Especially, a low value of the statistical inefficiency indicates that the finite difference method is inadequate to use, while a high \( s \) allows all interactions to be captured by this method. How the sample time \( \Delta t \), the number of time step \( N_{\text{time step}} \), and the graph radius \( d \) all affect the statistical inefficiency \( s \), where \( \Delta t \) is a measurement on how often the data set is sampled, \( N_{\text{time step}} \) is the total number of timesteps sampled from, and \( d \) is how deep into the graph network the CHAMPION algorithm will identify connections is all summarized in Table 1. From this, no parameter apart from \( \Delta t \) should affect \( s \), but since this quality obtained from a curve fit only the order of magnitude is relevant/important.

---

1 This method is merely suited for incomplete data as its numerical accuracy is lower. However, if the data comes from AIMD simulations where \( \Delta t \) is small enough, this method will be excellent and provide accurate forces.
Figure 2. The mean force curves generated from: a) H-H interactions, b) C-H interactions, and c) Li-O interactions. Each graph was produced from 100 bins.

When comparing a FF generated from data sampled with $\Delta t = 0.1$ fs (Fig. 2) to the forces in the literature, there is a large discrepancy, two orders of magnitude, between the two. Clearly the finite difference method underestimates the interactions and indeed a statistical inefficiency of ~20 indicates that many extreme forces will be smoothed out. One indication that the retractive forces here are much too small can be seen in Fig. 3, where the kinetic energy continuously increases, and the system eventually melts.
Conclusion remarks
Herein we report on how we have initiated the development of an alternative method to generate FFs for any system based upon the CHAMPION suite of neighborhood identifying algorithms. The preliminary obtained forces are yet several orders of magnitude too small, which is due to them being generated through a finite difference method. This has to be remedied, as well as incorporated with bonded terms as well, to give a better and fuller description of the interatomic interactions.

Acknowledgements
The research presented has received funding through the Energimyndigheten grants #P43525-1 and #P39909-1.

The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

Keywords: Force Field, Machine Learning, Molecular Dynamics, Electrolyte

References and Notes
11. BIG-MAP. [https://www.big-map.eu](https://www.big-map.eu) [https://www.big-map.eu/big-map].
12. Materials Genome Initiative | [WWW.MGI.GOV.](https://www.mgi.gov/).
26. Beal, M. S. *et al.* High Throughput Methodology for Synthesis, Screening, and Optimization of


