Thesis for the degree of Licentiate of Engineering

Novel Multi-Scale Modeling Framework for Structure and Transport in Complex Battery Electrolytes

Rasmus Andersson

Department of Physics
Chalmers University of Technology
Gothenburg, Sweden 2018

Novel Multi-Scale Modeling Framework for Structure and Transport in Complex Battery Electrolytes

RASMUS ANDERSSON

© RASMUS ANDERSSON, 2018

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

This work has received funding through the HELIS project (European Union's Horizon 2020 research and innovation program under Grant Agreement No. 666221), the Swedish Energy Agency (#P39909-1), and Chalmers Area of Advance: Materials Science, Theory and Modelling scheme of Advanced User Support.

Chalmers, Reproservice
Gothenburg, Sweden 2018

Abstract

Affordable high energy rechargeable batteries are crucial for further electrification of the transport sector, which is necessary in order to contribute to limit our CO_2 emissions to acceptable levels. While today's lithium-ion batteries (LIBs) have indeed initiated the electrification of the transportation section successfully, electric vehicles are still expensive and typically have ranges limited to *ca.* 100-500 km depending on price class. There are also safety concerns with LIBs and limited abundance of necessary materials why new chemistries, and especially new electrolytes, need to be explored. Emerging classes of electrolytes, such as highly concentrated electrolytes, have more complex structures than conventional electrolytes, with implications for the ion transport mechanism. This complexity necessitates a multi-scale modeling approach starting at the atomic level to gain further fundamental understanding.

This thesis outlines a framework where *ab initio* molecular dynamics is initially used to simulate small periodic systems (∼100 - 1000 atoms) over relatively short time spans (∼1 ps) to obtain trajectories that are subsequently used to train the parameters of a classical force field by machine learning of all parameters simultaneously by a genetic algorithm (GA). The force fields developed are then used to simulate larger systems (∼1000 - 100 000 atoms) classically over longer time scales (∼1 ns - 1µs). The resulting trajectories are used to collect statistics for a hierarchical analysis, which resolves the structure in terms of dynamic clusters, and quantifies the life-time distribution, population dynamics, and transport properties of identified clusters and non-covalent bonds. The method is ultimately to be of general use to both qualitatively and quantitatively elucidate the ion transport mechanism in novel types of electrolytes as a function of composition.

Keywords: Lithium-ion batteries, electrolytes, non-vehicular transport, genetic algorithms, molecular dynamics, force field development, hierarchical analysis, multi-scale method

List of Papers

This thesis is based on the following paper:

I *Simultaneous All-Parameter Fitting of Force Fields for Electrolytes with Non-Vehicular Ion Transport*
R. Andersson, F. Årén, A. A. Franco and P. Johansson
Manuscript

The thesis is also to a substantial degree based on the development of the following software library, of which a short description is appended:

CHAMPION: Chalmers Hierarchical Atomic, Molecular, Polymeric & Ionic simulation and analysis toolkit.

Table of Contents

1 Introduction **1**
  1.1 Lithium-Ion Batteries 1
  1.2 Lithium-Sulfur (Li-S) Batteries 3
  1.3 Electrolytes 4
    1.3.1 Conventional LIB electrolytes 5
    1.3.2 Highly Concentrated Electrolytes 6
    1.3.3 Electrolytes for Li-S batteries 7
  1.4 Challenges in Modeling Ion Transport 7

2 Computational Background **11**
  2.1 Density Functional Theory 11
    2.1.1 Exchange-Correlation Functionals 13
  2.2 Molecular Dynamics 13
    2.2.1 *Ab initio* Molecular Dynamics 14
    2.2.2 Classical Molecular Dynamics 15
  2.3 Optimization Methods 16
    2.3.1 Gradient-Based Optimization 18
    2.3.2 Genetic Algorithms 19

3 Novel Multi-Scale Modeling Framework **21**
  3.1 Outline 21
  3.2 AIMD Simulations 21
  3.3 Force Field Optimization 22
  3.4 Classical MD Simulations 23
  3.5 Hierarchical Analysis of Trajectories 24

4 Results and Discussion **27**
  4.1 Water – Proof-of-Concept System 27
  4.2 Highly Concentrated Electrolytes 30

5 Future Research **31**

6 Acknowledgments **33**

References **35**

Chapter 1

Introduction

High energy rechargeable batteries can contribute to the transition to a CO_2 neutral economy, not the least in hybrid and fully electric vehicles. Electrifying transport is crucial for climate change mitigation since the transport sector stands for about 15% of global CO_2 emissions [1].

Since their commercialization in the early 1990's, lithium-ion batteries (LIBs) have revolutionized portable electronics, and their volumetric capacity has more than tripled, from 200 Wh/l for the first commercial LIB to 650 Wh/l for energy optimized cells today [2]. Due to high energy and power densities, as well as long cycle and calendar lives, LIBs have become the most prevalent energy storage technology for fully electric vehicles [3].

The electrolytes of today's state-of-the-art LIBs, however, are volatile and flammable, which makes LIBs prone to thermal runaway under abuse or extraordinary conditions [4]. These safety issues are compounded by the large amounts of energy contained in electric vehicle battery packs compared to handheld electronics. The electrolytes are also not electrochemically stable towards novel high voltage positive electrodes [5], while highly optimized in their formulation and not substantially variable without losing some parts of their functionality [4]. Creating LIB cells that combine both higher energy density and improved safety thus require more or less entirely new electrolyte designs.

1.1 Lithium-Ion Batteries

An LIB cell consists mainly of three parts: a positive electrode, a negative electrode and a separator wet by a liquid electrolyte (Fig. 1.1). The electrodes are sometimes referred to as anode (negative) and cathode (positive), but this designation is strictly true only when discharging the cell, why henceforth they will be referred to as positive and negative electrodes [6].

Both electrodes are porous and consist of particles of ion- and electron-

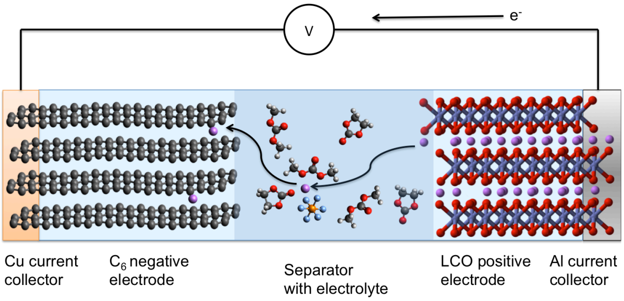


Figure 1.1: Schematic of a LIB in the beginning of the charging process.

conducting active materials (AMs) in a porous structure held together by a binder material, enhancing the mechanical properties and most often an additive enhancing the electron-conduction – often carbon-based. Each electrode is cast on a current collector that connects the electrode AM particles to each other and to the external circuit and give additional mechanical stability. For the positive electrode Al is used, while for the negative electrode, heavier and more expensive Cu must be used as Li alloys with Al at low potentials [7].

Both the AMs in LIBs are intercalation compounds, *i.e.*, they allow Li^+ ions to be inserted at specific sites in the electrode matrix without significantly changing the structure of the host material – intercalated. The negative electrode AM is usually graphite, which can host Li ions between its graphene layers, while the positive electrode AMs are more diverse, but in general transition metal oxides, *e.g.* $LiCoO_2$ (LCO), $LiFePO_4$ (LFP) or $LiNi_xMn_yCo_zO_2$ where $x + y + z = 1$ (NMC), all with layers or channels allowing Li^+ intercalation and transport [7].

The two electrodes are sandwiched around a separator, which is a microporous, electronically insulating, often polymeric material, whose function is to avoid electronic short circuit of the cell. The separator is wet by a liquid electrolyte, which conducts ions in order to balance the electron transport in the external circuit. [7]

In the discharged state, the positive electrode is lithiated and the negative electrode delithiated. Upon charging, the applied voltage oxidizes the transition metal atoms in the positive electrode and releases electrons to the external circuit. The change in oxidation state leads to release of intercalated Li-ions to the electrolyte, while the electrons migrate along the external circuit to the negative electrode which is reduced and accepts Li-ions (Fig. 1.1).

Table 1.1: Average potentials and theoretical specific capacities for some common AMs in LIBs [8] and Li-S batteries [9]

Active Material	Average Potential *vs.* Li^+/Li^o (V)	Specific Capacity (mAh/g)
Li	0.0	3860
Graphite	0.1	372
S	2.3	1670
LFP	3.4	170
$LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$	3.7	278
LCO	4.0	274

The energy content of an electrochemical cell is determined by the voltage between the electrodes, $V = E^+ - E^-$ where E^+ and E^- are the potentials of the positive and the negative electrodes, respectively, *vs.* some reference potential, and the amount of charge transferred between them during a full discharge. The total energy is given by

$$E = \int\limits_0^C V(Q)dQ,$$

where $V(Q)$ is the voltage after Q amount of charge has passed from the negative to the positive electrode and the integral runs from 0 to the full capacity of the cell, C, defined as the total amount of charge that can be reversibly transferred [6]. The (volumetric) energy density of the cell is given by the total energy per volume and the specific energy, or gravimetric energy density by the energy per mass, where the mass or energy can be calculated on the AM, the cell or the battery level.

An important property of an AM is thus its specific capacity, *i.e.*, its capacity per unit mass. The theoretical specific capacity is given by an amount of charge divided by the mass of AM required to hold that charge. Table 1.1 shows average potentials and theoretical specific capacities for a number of common electrode materials in LIBs and Li-sulfur (Li-S) batteries.

1.2 Lithium-Sulfur (Li-S) Batteries

Li-S batteries are one of the more researched next generation battery (NGB) concepts, in its simplest design based on a S positive and a Li metal negative electrode [10], [11]. The Li-S battery concept is attractive due to the (very) low cost of S as an AM and the very high theoretical capacities of both AMs compared to those used in LIBs (Table 1.1). However, the voltages of Li-S cells tend to be considerably lower than for LIBs (Table 1.1, [10]).

Sulfur exists in several allotropes, the most common is S_8 rings [12]. Sulfur can in a Li-S cell be reduced in a number of steps, resulting in various lithium-polysulfides (PSs) [13]:

$$\mathrm{S_8} \rightarrow \mathrm{Li_2S_6^{2-}} \rightarrow \ldots \rightarrow \mathrm{Li_2}S. \tag{1.1}$$

The most obvious problem with using sulfur as an electrode is its negligible electronic conductivity. Therefore carbon/sulfur (C/S) composites with relatively high concentrations of carbon, up to 30 wt%, are used, but this then lowers the practical cell energy density achievable significantly [10].

During discharge electrons travel from the Li metal, which is oxidized and therefore stripped of Li, to the C/S composite, where S_8 rings are reduced and react to form various Li-PSs. At decreasing state-of-charge (SOC), the average PS chain length decreases until short-chain PSs are finally reduced to Li_2S. During charge, the reverse process happens.

One of the main problems of realizing practical Li-S batteries follows from the solubility of elemental sulfur and the intermediate long and medium chain length PSs into the electrolyte [10], [13]. When S and PSs are transported away from the electrode, AM is lost and they might eventually end up at the surface of the negative electrode and partake in parasitic reactions, all leading to irreversible capacity losses. One way to counter this PS shuttle mechanism is to engineer electrolytes to have low PS solubility, while maintaining good ionic transport [13]. Rational design of Li-S electrolytes to this end requires a good understanding of the transport mechanism.

1.3 Electrolytes

Typical electrolytes for Li batteries, both LIBs and Li-S batteries, consist of a lithium salt dissolved in a polar solvent [4]. Recent guidelines by Flamme et al. lists the following requirements for electrolytes in high energy density LIBs [14]:

- ionic conductivity above a few mS/cm,
- electrochemical stability window (ESW) ≥ 4.5 V *vs.* Li^+/Li°,
- liquid range at least between -20°C to 180°C,

- relative permittivity ≥ 20,
- chemically inert against all cell components to avoid unwanted side reactions,
- good wettability towards electrodes and separator.

Electrolytes should also preferably have high flash point, low toxicity, low cost and be environmentally benign.

1.3.1 Conventional LIB electrolytes

The large potential gap between the electrodes in LIBs prohibits aqueous electrolytes from being used, as water has a quite narrow (1.23 V [15]) ESW. Instead, mixtures of aprotic, organic solvents, typically cyclic and linear alkyl carbonates are used, due mainly to their high oxidative stability, *ca.* 4.5-5.0 V *vs.* Li^+/Li^o [16], where the cyclic carbonates provide high dielectric constants needed for high salt solubility while the linear carbonates lower the viscosity, thus increasing ionic conductivity [4].

In fact, the commonly used electrolytes, while having significantly wider ESWs than aqueous electrolytes, are not thermodynamically stable towards reduction against commonly used LIB negative electrodes. Modern LIBs therefore depend on reduction of solvent molecules and anions during the first cycle(s) in order to form a layer, the solid electrolyte interphase (SEI), between the electrode and the electrolyte, protecting the latter from continual decomposition, while allowing diffusion of Li ions through [17], [18].

The most used electrolyte formulation for LIBs consists of 1 M $LiPF_6$ dissolved in an equimolar (1:1) mixture of the cyclic ethyl carbonate (EC) and the linear dimethyl carbonate (DMC) [4] (Fig. 1.2). The 1 M salt concentration is chosen because it gives the maximum ionic conductivity; lower salt concentration leads to fewer charge carriers, while higher salt concentration leads to higher cation-anion aggregation neutralizing a fraction of the charge carriers and furthermore also increases the viscosity. DMC can be replaced by similar carbonates, such as ethyl methyl carbonate (EMC), and the relative concentrations of linear and cyclic carbonates can be varied to make different trade-offs between properties, but so far both EC and PF_6^- have been found hard to replace. EC forms stable SEIs without co-intercalating into graphite and causing exfoliation of graphitic sheets (as propylene carbonate (PC) does) while PF_6^- has a good balance of properties not matched by any other known anion, although it is thermally unstable and very reactive towards traces of water [4]. Most importantly its decomposition products stabilize the Al current

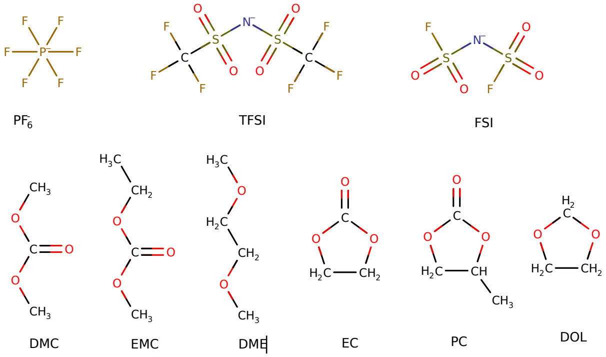


Figure 1.2: Common components of electrolytes for lithium batteries.

collector by creating a thin AlO_xF_y layer [4]. For the standard electrolyte the main available degree of freedom in composition is the addition of small (typically ≤ 5 wt%) concentrations of functional additives, *e.g.* flame retardants, overcharge protectors and SEI-formers [19].

1.3.2 Highly Concentrated Electrolytes

The idea behind highly concentrated electrolytes is to radically increase the salt concentration beyond the conventional 1 M to the point where very few ”free” (*i.e.* uncoordinated to lithium-ions) solvent molecules and/or anions remain. This lowers the vapor pressure and flammability and alters the electrochemical stability of both the anions and the solvents, but also increases the viscosity [20].

Highly concentrated electrolytes have been shown to exhibit a number of interesting properties: Suo et al. managed to suppress solvent co-intercalation into graphite in EC-free electrolytes [15], while Matsumoto et al. and Yamada et al. found a suppressed corrosion of the Al current collector by the anions bis(trifluoromethane sulfonyl) imide (TFSI) and bis(fluorosulfonyl) imide (FSI) [21], [22] (Fig. 1.2)(although this effect may not be strong enough in practice [23]). The latter is especially interesting as these anions have good overall properties and superior thermal stability to PF_6^-, but have not found widespread use in LIBs because they normally, at 1 M Li-salt concentration, cause Al corrosion above 3.5 V *vs.* Li^+/Li^o [24]. Finally, Suo et al. and Lundgren et al. both found increased Li transport numbers, t^+, i.e. the fraction of ion transport due to Li ions, in highly concentrated electrolytes [15], [25], which somewhat compensates for the increased viscosity.

1.3.3 Electrolytes for Li-S batteries

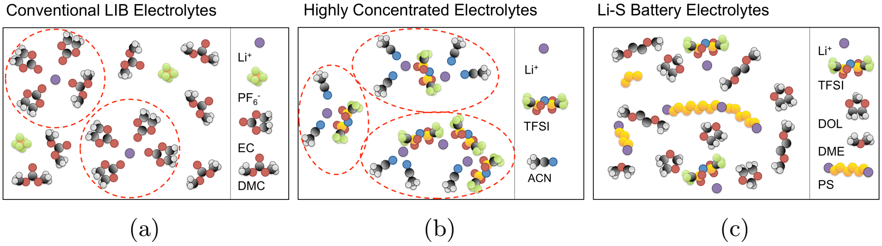
Sulfur as an electrode has a much lower potential (2.3 V *vs.* L^+/Li^o as compared to the AMs used in LIBs, and hence the selection of electrolytes for Li-S batteries is not as constrained w.r.t. oxidative stability [13]. On the other hand, the shuttling of PSs from the C/S composite to the Li electrode and subsequent precipitation of unsoluble PSs leads to irreversible AM loss unless the Li surface is passivated [9]. Therefore, the electrolytes for Li-S batteries should be designed to limit PS solubility, unless a catholyte concept is used [10], [13].

The most common electrolyte is a 1 M LiTFSI in an equimolar (1:1) solution of 1,3-dioxolane (DOL) and 1,2-dimethoxyethane (DME), with $LiNO_3$ as an additive or co-solvent (*ca.* 0.5 M) to passivate the Li surface [13] (Fig. 1.2). Both the LiTFSI and LiFSI salts can be used in Li-S batteries without any risk of Al corrosion due to the relatively low potential of sulfur.

The main qualitative difference to LIB electrolytes is that Li-S battery electrolytes contain dissolved PS species which vary in type and concentration, as functions of cycling and SOC. Hence also the physico-chemical properties of the electrolytes vary [13]. Furthermore, the extent to which the PSs are coordinated to Li ions and the extent to which they partake in Li ion ligand exchange, *i.e.* both the structure and the dynamics, are complex, poorly understood phenomena.

1.4 Challenges in Modeling Ion Transport

Modeling of the ion transport in LIB electrolytes has most often been based on Newman's concentrated electrolyte theory [26]–[28], which builds on the assumption that the ions have individually well-defined transport properties and ion pairing is accounted for only through thermodynamic coefficients. This works well if the Li^+ ions are predominantly transported in long-lived solvation shells of a well-defined size – vehicular transport – but lacks explanatory power if ion hopping or solvent exchange plays important roles, which both experimental [29] and computational [30] studies suggest. Vehicular transport leads to more mobile anions than Li^+ ions as the latter drag along their first solvation shell while a substantial fraction of the anions move independently (Fig. 1.3a). Although considerable ion-pairing seems to be the norm [29],


Conventional LIB Electrolytes
Li⁺
PF₆⁻
EC
DMC
Highly Concentrated Electrolytes
Li⁺
TFSI
ACN
Li-S Battery Electrolytes
Li⁺
TFSI
DOL
DME
PS

(a) (b) (c)

Figure 1.3: Schematic illustration of the local structures in conventional, highly concentrated and Li-S battery electrolytes.

[31], many electrolytes with 1 M salt concentration (*ca.* 1:20 salt:solvent) have been found to have low variance in coordination numbers compared to highly concentrated electrolytes [32], and thus the former are more amenable to Newman style modeling than the latter.

With much fewer solvent molecules available per Li^+ ion than in a conventional 1 M electrolyte, the first solvation shell of the cation will be different, no longer ruled mainly by the cation-solvent interaction. They will hence be both more complex and more varied, especially as the dearth of solvent molecules leads to a less effective Coulombic screening and in turn, to more extensive ion-ion coordination [29], [31], [32], and contain a variable number of solvent molecules and anions. The latter can also be further coordinated to other cations, creating clusters and/or networks, and hence the possible number of local structures increases drastically (Fig. 1.3b). At the same time, the ligand exchange rate changes with concentration, why different transport mechanisms may become important [33]. This may explain why t^+ increases for highly concentrated electrolytes, but there is yet no proper understanding of the mechanism. In contrast to the vehicular transport mechanism often assumed for conventional 1 M electrolytes, a Grotthuss-like hopping mechanism can here contribute, in which individual lithium-ions hop from cluster to cluster [33], akin to the hopping of excess protons between water molecules through hydrogen bonds in aqueous electrolytes. However, many more complex mechanisms can easily be imagined.

In Li-S batttery electrolytes dissolved S and PSs can both coordinate to lithium-ions and affect the overall viscosity, and thereby further increase the complexity of the ion transport – both structurally and dynamically (Fig. 1.3c). To be capable of explaining ion transport in highly concentrated and Li-S battery electrolytes any method applied must take the stance of avoiding unwarranted assumptions, why it almost must be based on modeling on the molecular level. Second, because of the structural complexity, it must be based on large enough systems to enable statistically reliable conclusions to be made regarding the ion transport mechanism(s). Third, any system modeled needs to be simulated for a long enough time (or equivalent) so that the dynamics are adequately sampled. Altogether, this combination of requirements is hard to fulfill with a single simulation method, which motivates multi-scale methodology to be applied. *Ab initio* molecular dynamics (AIMD) based on density functional theory (DFT) gives high accuracy but poor scaling with system size [34]. Classical MD on the other hand, is highly sensitive to the quality of the parameterization used for the interactions (*i.e.* the force field) [35], [36]. Unfortunately, well-known transferable force fields for organic chemistry-based systems have been found unsatisfactory in reproducing experimental transport properties in systems with high degrees of ionic association [37].

In this thesis, we use a combination of AIMD and classical MD simulations to attack the transport in highly concentrated electrolytes using multi-scale modeling (MSM). The former simulations are used to train classical MD force fields by using machine learning (ML). The force fields are optimized on a system-to-system basis to reproduce *ab initio* dynamics. The optimized force fields are subsequently used to run the classical MD simulations on scales sufficient for a detailed, statistically significant analysis of the ion transport mechanism(s).

Chapter 2

Computational Background

2.1 Density Functional Theory

Applying computational quantum chemistry to materials is usually reduced to predicting the ground state electronic structure of the material for a given configuration of the atomic nuclei by applying the Born-Oppenheimer approximation [38]. This allows decoupling the dynamics of the electrons from the dynamics of the nuclei, based on their vastly different time-scales due to their disparity in mass. Furthermore, it usually suffices to consider the nuclei, and often also tightly bound electrons, as classical point particles. As only the ground state electron structure is sought, it suffices to solve the time-independent Schrödinger equation (TISE). The remaining many-body problem is still very challenging, however, as this TISE is a non-separable equation for an n-dimensional wave function [39]:

$$\left[-\sum_{i}^{n}\frac{\hbar^2}{2m_e}\nabla^2-\sum_{i}^{n}eV_{\text{ions}}(\boldsymbol{r}_i)+\frac{1}{4\pi\epsilon_0}\sum_{(i,j)}^{n}\frac{e^2}{|\boldsymbol{r}_j-\boldsymbol{r}_i|}\right]\psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n) \\ = E\psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n). \quad (2.1)$$

Hohenberg and Kohn proved in 1964 that all ground state properties of a many-electron system are uniquely determined by the electron density [40],

$$n(\boldsymbol{r})=\sum_{i=1}^{N}\int_{\boldsymbol{R}^3}\delta(\boldsymbol{r}_i-\boldsymbol{r})\psi^*(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)\psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)d\boldsymbol{r}_1d\ldots\boldsymbol{r}_n.$$

This means that there exists a functional $E[n]$ which determines the total ground state energy of the system as a function of the density. They further showed that the correct ground state electron density minimizes the total energy under the constraint of conserving the number of electrons [40],

$$\int_{\boldsymbol{R}^3} n(\boldsymbol{r}) d\boldsymbol{r} = N.$$

This is a very desirable proposition since the electron density is a function only of three spatial coordinates rather than the full $3N$-dimensional configuration space, as in Eq. (2.1). This can be expressed mathematically as

$$E_0 = \int_{\boldsymbol{R}^3} eV_{\text{ions}}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} + \frac{1}{2}\frac{1}{4\pi\epsilon_0} \int_{\boldsymbol{R}^3} \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' d\boldsymbol{r} + G[n], \tag{2.2}$$

where the first term is the energy from the potential of the static ions, the second term is the electron-electron repulsion and $G[n]$ is a universal functional of the electron density that captures the remaining energy contributions. However, the Hohenberg-Kohn theorems only show that such a functional exists and say nothing about how to find it.

Density functional theory (DFT) became practical due to later work by Kohn and Sham in 1965 [41]. Their method starts by splitting $G[n]$ into two terms, $G[n] = T[n] + E_{\text{xc}}[n]$, where $T[n]$ is the kinetic energy functional for a system of non-interacting electrons and $E_{xc}[n]$ collects everything not captured in previous terms. This latter functional is known as the *exchange and correlation functional* and is in general not known exactly for practically interesting systems. In practice there exists a hierarchy of methods for how to approximate this functional, (see 2.1.1).

The second, and most crucial, innovation by Kohn and Sham was to consider a fictional system of non-interacting particles, governed by the Schrödinger-like equation:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\boldsymbol{r})\right]\phi_i(\boldsymbol{r}) = \epsilon_i \phi_i(\boldsymbol{r}), \tag{2.3}$$

where $V_{\text{eff}} = V_{\text{ions}} + V_H + V_{\text{xc}}$ includes the ionic potential, V_{ions}, the Hartree potential, V_H, expressing electron-electron interaction in a mean field sense and the remaining potential, V_{xc} due to Pauli exclusion and Coulombic interaction beyond the mean-field approximation. Eq. (2.3) is written in terms of fictitious non-interacting spin-orbitals which have no simple physical interpretation but that reproduce the correct electron density for the corresponding interacting system, so that

$$n(\boldsymbol{r}) = \sum_i \int_{\boldsymbol{R}^3} \phi_i(\boldsymbol{r})^* \phi_i(\boldsymbol{r}) d^3\boldsymbol{r}. \tag{2.4}$$

Given an electron density, the effective potential V_{eff} can be evaluated as it is a functional of the density, whereafter Eq. (2.3) can be evaluated to give the orbitals ϕ_i and the eigenvalues ϵ_i. Given the full set of orbitals an updated density can be computed using Eq. (2.4), and the cycle is repeated until self-consistency among the equations is reached, *i.e.* when the density, potentials and orbitals have converged to stable values.

2.1.1 Exchange-Correlation Functionals

Exchange and correlation functionals in DFT exist on a "Jakob's ladder" from simple and cheap to more accurate and computationally costly [42]. The need to choose a suitable level of approximation stems from the non-locality of the interactions contained in this term, where the potential at one point depends in principle on all other points in space. In practice, however, the dependence is mostly restricted to a neighborhood of the point, which motivates the general approach of expanding the potential in terms of the local electron density and low order spatial derivatives evaluated at the point in question:

$$V_{\text{xc}}[n(\boldsymbol{r})] = V_{\text{xc}}[n(\boldsymbol{r}), \nabla n(\boldsymbol{r}), \nabla^2 n(\boldsymbol{r}), \dots]$$

The simplest type of functionals are based on the *local density approximation* (LDA), which only depend on the density at the selected point and thus disregards any non-locality. In practise this approach works well when the electron density is close to homogeneous.

The next rung on the ladder also includes the gradient, and functionals on this level are known as generalised gradient approximations (GGA). Even higher order derivatives are considered in meta-GGA functionals.

The AIMD simulations done in this thesis use the exchange-correlation functional of Perdew, Burke and Ernzerhof (PBE), which is a GGA functional [43].

2.2 Molecular Dynamics

Molecular dynamics (MD) is a family of methods for simulating systems of atoms, molecules and/or macromolecules over time [38]. With few exceptions, the atoms are treated as classical bodies, whose dynamics are determined by the forces acting on them, based on Newtonian physics:

$$\ddot{\boldsymbol{r}} = \frac{\boldsymbol{F}}{m}.$$

The differences between different MD methods lie in how these forces are calculated. In *ab initio* MD (AIMD) the electronic structure is computed using quantum chemical methods, while for classical MD, expressions for the interatomic forces need to be specified.

2.2.1 *Ab initio* Molecular Dynamics

Ab initio (lat. from first principles) in the context of MD simulations reflects that no parameters are needed to simulate the dynamics of the system. However, the most common method for evaluating forces is DFT (which is not considered *ab initio* in the context of quantum chemistry).

The most naive implementation of AIMD based on the TISE is the Born-Oppenheimer MD (BOMD) method [34]. It starts by initializing the atomic positions and momenta, whereafter the electron structure is computed by DFT. Based on the converged electron density and total energy, the forces acting on each atom are computed from the gradient of the energy upon perturbing the position of each atom in turn. The force on atom i is given by

$$\boldsymbol{F}_i = -\frac{\partial E}{\partial \boldsymbol{R}_i}$$

in the absence of magnetic or time-dependent external forces. Given the positions, momenta (and thereby velocities) and forces (and thereby accelerations) of all atoms, the system can be propagated to the next time-step where a new electron density needs to be calculated. However, this is typically much cheaper than the initialization since the first guess can be based on the converged electron structure from the previous time-step. The time-steps therefore need to be small enough to limit changes between consecutive snapshots in the electronic structure.

Even so, to update the electronic structure quantum mechanically for every snapshot is computationally expensive. Car and Parrinello therefore proposed a method in 1985 based on a hybrid quantum/classical approach [44]. The problem is framed in terms of a total interaction energy functional dependent on both Kohn-Sham orbitals, ϕ_i, atomic positions, R_I and Lagrange multipliers α_ν, to fulfill any external constraints on the system (*e.g.* volume). The orbitals and Lagrange multipliers are assigned fictitious masses, μ and μ_ν in order to give them classical dynamics. This procedure gives the Lagrangian

$$L = \sum_i \frac{\mu}{2} \int_{\boldsymbol{R}^3} |\dot{\phi}_i|^2 d\boldsymbol{r} + \sum_I \frac{M_I R_I^2}{2} + \sum_\nu \frac{\mu_\nu \alpha_\nu^2}{2} - E[\{\phi_i\}, \{R_I\}, \{\alpha_\nu\}], \quad (2.5)$$

giving rise to the classical equations of motion:

$$\begin{aligned} \mu\ddot{\phi}_i(\boldsymbol{r}, t) &= -\frac{\delta E}{\delta \phi_i^*(\boldsymbol{r}, t)} + \sum_k \Lambda_{ik} \phi_k(\boldsymbol{r}, t), && (2.6) \\ M\ddot{R}_I &= -\nabla_{R_I} E, && (2.7) \\ \mu_\nu \ddot{\alpha_\nu} &= -\frac{\partial E}{\partial \alpha_\nu}, && (2.8) \end{aligned}$$

where Λ_{ik} are Lagrange multipliers to maintain orthonormality of Kohn-Sham orbitals.

The fictitious masses of the electrons need to be selected to be sufficiently lighter than the nuclear masses so that no energy is exchanged between these degrees of freedom in the simulated dynamics. However, the electron mass should be chosen much heavier than the physical value in order to allow longer simulation time-steps than possible in BOMD simulations. A typical choice is 400 times the electron rest mass.

2.2.2 Classical Molecular Dynamics

In classical MD the forces are computed classically, and therefore a simulation is computationally much cheaper than AIMD simulation, but since it is dependent on (a quite large number of) empirical parameters, the quality is very much dependent on the parameters used [35], [36].

In materials science, such a parameter set is usually known as an interatomic potential. Since the materials, mostly crystalline solids, often are quite simple with a small number of distinct interactions, relatively few parameters are needed.

Computational chemistry on the other hand, tends to deal with molecules, which are held together by a multitude of different bonds, whose character and physical origins differ, *e.g.* covalent bonds of different orders, ionic bonds, hydrogen bonds, and additionally also van der Waals interactions. In order to describe all of these accurately by classical models, it is usually not enough to only consider which pairs of elements are bound, but also their respective local environments – a carbonyl oxygen atom is different from an ether oxygen atom, etc. This can lead to a very large number of parameters in practice. The set of potentials in computational chemistry are conventionally called *force fields* [36].

There is much experience in using classical MD to study organic chemistry based systems, which over time has resulted in more or less standardized functional forms for the terms used in force fields [36], [45], [46]:

$$\begin{aligned} E = 4\sum_{(i,j)} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{1}{4\pi\epsilon_0} \sum_{(i,j)} \frac{q_i q_j}{r_{ij}} + \sum_{\text{bonds}} \frac{k_b (l - l_{0,b})^2}{2} \\ + \sum_{\text{bond angles}} \frac{\kappa_a (\theta - \theta_{0,a})^2}{2} + \sum_{\text{dihedrals}} \sum_{n} a_{n,d} \cos(n\phi + \phi_{0,d}), \end{aligned} \tag{2.9}$$

where the first term is the Lennard-Jones approximation for the Pauli exclusion principle and van der Waals interactions, the second term is the Coulomb interactions between all pairs of atoms, the third term describes harmonic bond potentials, the fourth harmonic potential for bond angles and the last a Fourier series for the energy of proper and improper dihedrals, *i.e.*, torsions and out-of-plane bending of groups of four bonded atoms (Fig. 2.1).

The parameters of a force field of the form (2.9) are thus a set:

$$\mathrm{FF} = \{\epsilon_i, \sigma_i, q_i, k_b, l_{0,b}, \kappa_a, \theta_{0,a}, a_{n,d}, \phi_{0,d}\}_{i,b,a,d}\,,$$

where $\sigma_{ij} = (\sigma_i + \sigma_j)/2$ and $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$.

The non-bonded, *i.e.* Lennard-Jones and Coulomb terms, are typically excluded for directly bonded atoms and next-nearest neighbors, and scaled down to a number between 0 and 1 for third neighbors.

2.3 Optimization Methods

Optimization is used in this work to a) generate starting geometries for MD simulations, b) fit the parameters of a force field, and c) find the non-covalent bonds holding together hierarchical structures.

Optimization is the process of finding a minimum value of a *cost function* (a.k.a.*loss function*), here denoted by $C(\boldsymbol{x})$, associated with a given point $\boldsymbol{x}$ in *design space*, the space of independent coordinates that may be chosen, and where each point is assigned a cost value. An example would be an actual landscape where the optimization problem is to find the lowest point. Design

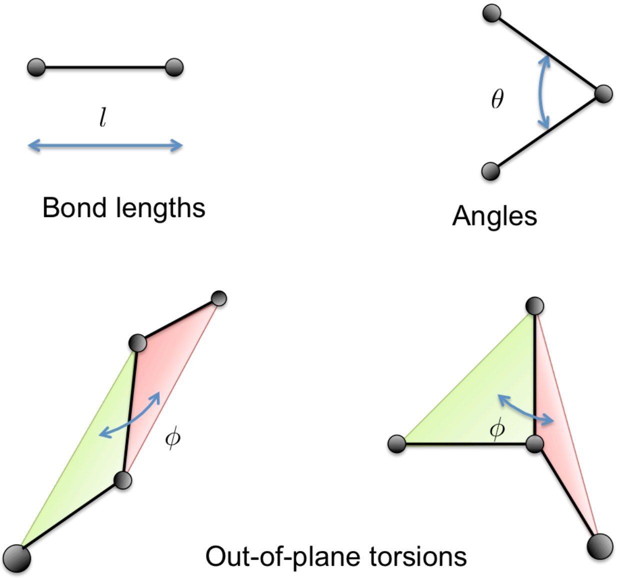

l
Bond lengths
θ
Angles
ϕ
ϕ
Out-of-plane torsions

Figure 2.1: Schematic illustration of the types of bonded interactions included in force fields for organic systems.

space is then the real 2-dimensional space along which you can move, and the cost function is the height above some arbitrary level (*e.g.* sea level).

Any found minimum might be local or global, and in general it is not possible to know which, but this is also strongly method dependent. Genetic algorithms (GAs, see 2.3.2) are designed specifically to avoid getting stuck in local minima, while for example most gradient-based methods (2.3.1) have a very strong tendency to only locate the closest local minimum. Yet, when a local optimum suffices, or the solution domain is known to be convex, gradient-based methods are practical due to their simplicity.

2.3.1 Gradient-Based Optimization

Gradient-based optimization is a class of (usually) deterministic methods where the aim is to find the closest local minimum by taking into account the slope, and possibly higher order derivatives of the current best guess in order to select the next point in design space to evaluate. Gradient-based methods work best for relatively low-dimensional design spaces and a requirement is that it is possible to evaluate the gradients

$$\nabla C = \frac{\partial C}{\partial \boldsymbol{x}} \tag{2.10}$$

of the cost function with respect to small movements in design space. Such evaluations may be possible to do analytically or only numerically by finite differences – for the latter it is intuitive from Eq. (2.10) that the computational cost will be proportional to the number of dimensions along which $\boldsymbol{x}$ needs to be perturbed to get the full gradient vector.

Steepest descent is the most straight-forward gradient-based optimization method and always moves in the direction in which the cost function declines most steeply, *i.e.*, along the negative gradient. The step length is determined by a line search in the search direction, which finds the minimum using one-dimensional convex methods, such as Newton's method:

$$\Delta x = -\frac{f'(x)}{f''(x)}.$$

While the steepest descent method is simple to both understand and implement, it has a tendency to zigzag its way to the solution as the next gradient will always be orthogonal to the previous search direction, making convergence suboptimal. This can be remedied by taking into account previous search directions; the conjugate gradient method, where the first search direction is still chosen to be along the negative gradient, but subsequent directions are conjugated with previous search directions based on the shape of the cost function landscape. Conjugate gradient optimization is here used for geometry generation and to find the rotation of rigid bodies between trajectory snapshots.

2.3.2 Genetic Algorithms

Genetic algorithms (GAs), a.k.a. evolutionary algorithms, are a family of stochastic optimization heuristics based on simulating the process of Darwinian evolution [47]. The objective is for successful solutions to arise over many iterations of an evolutionary process and due to this background, GA terms are often borrowed from biology. The aim of the optimization is to find a *genotype*, *e.g.* a set of force field parameters, that produces the fittest *phenotypes*, *e.g.* accurately reproduces atomic forces as a function of configuration. Genotypes may also be called *chromosomes*, and sometimes *individual* is used for both a genotype and its associated phenotype. A *generation* is simply an iteration of the algorithm, where each generation consists of a distinct population of individuals.

The chromosomes, or genotypes, need to be encoded in some manner. In this work, where genetic algorithms are used for *e.g.* force field optimization, the goal is to find real values for all the parameters, and thus real-valued encoding is used. GAs are also used to find coordination distances holding together ionic aggregatees – where real-valued encoding is also natural. In addition to being relatively resistant against local minima, GAs are practical when evaluating gradients is either computationally costly or the cost function is not smooth.

The basic structure of a GA is presented in Algorithm 1; A population is generated randomly with broad starting guesses for all parameters reflecting some knowledge of the reasonable values. They are thereafter evaluated, which underlies the selection of parents for the next generation, out of a softmax distribution (analogous to a Boltzmann distribution) premiering low-cost solutions:

$$P_i = \frac{e^{-\beta C_i}}{\sum_j e^{-\beta C_j}}.$$

For each new genotype to be constructed, two parents are selected uniformly from the list of parents and recombined. Each new genotype is mutated, *i.e.*, its parameters are perturbed. The best few individuals survive unaltered to the next generation.

Algorithm 1: Basic structure of a GA

```
Initialize a population of individuals from broad starting guesses;
while not converged do
    Evaluate costs $C_i$;
    for $i = 1$, $i \leq$ number of elites, $i{+}{+}$ do
        Save the $i$th best individual to next generation
    end
    for each parental slot do
        Select parent, $P_i = e^{-\beta C_i} / \sum_j e^{-\beta C_j}$;
    end
    for each new individual do
        Choose parent $a$;
        Choose parent $b$;
        Cross-over$(a, b)$;
        Mutate;
    end
end
```

Chapter 3

Novel Multi-Scale Modeling Framework

3.1 Outline

Our MSM methodology (Fig. 3.1) starts from AIMD simulations, which provide training data for a GA to find accurate force field parameters to reproduce the AIMD forces. Using this optimized force field, classical MD simulations are run for longer times and larger systems in order to capture the structures and dynamics relevant for ion transport, which are analyzed hierarchically from their trajectories in terms of whatever clusters/aggregates of ions and molecules are found in the MD trajectories. Statistical physics provides predictions of macroscopic quantities such as density, ionic conductivity and viscosity which can be readily validated against experiments. The hierarchical analysis aims to explain the emergence of the macroscopic transport properties in terms of microscopic/molecular level phenomena.

3.2 AIMD Simulations

The starting geometries for AIMD simulations are generated in two steps: first a randomized molecular geometry with the desired density is spread out in a periodic box by minimizing the cost function

$$C_i = \sum_j \frac{1}{1 + \frac{d_{ij}}{R_i + R_j}}, \tag{3.1}$$

where the sum is over all atoms, d_{ij} denotes the distance between atoms i and j, and R_i is the van der Waals radius of atom i. The cost is minimized w.r.t. translation and rotation of each molecule or ion in turn, using the conjugate gradient method to find the closest minimum given the present positions of

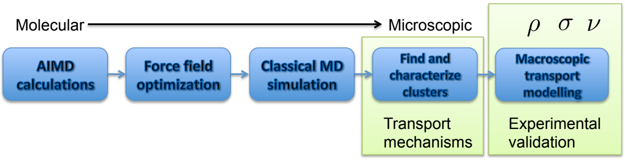

Molecular
Microscopic
$\rho \quad \sigma \quad \nu$
AIMD calculations
Force field optimization
Classical MD simulation
Find and characterize clusters
Macroscopic transport modelling
Transport mechanisms
Experimental validation

Figure 3.1: Overview of our novel MSM framework.

all other molecules. Thereafter a classical MD simulation is run in the NVT ensemble to pre-equilibrate the system, using a conventional force field, *e.g.* AMBER [48] or OPLS [46].

An AIMD simulation is then performed in the NVT ensemble. The first part of the AIMD trajectory is not used in any subsequent analysis as the system should be allowed to equilibrate. The part of the trajectory after this initial equilibration time is the production data that will subsequently be used for the force field optimization.

3.3 Force Field Optimization

The functional form of a system-specific force field is selected in advance, whereafter the parameters of the force field are optimized by GA. A population of individuals is generated from normal distributions with variances proportional to the *a priori* uncertainty in the correct values. The number of individuals should be at least on the same order as the number of parameters, and the number of parents should be approximately a third of the population size, based on trial and error. Individuals are assessed based on their ability to reproduce the forces predicted by AIMD. In each training step a small number of snapshots from the AIMD trajectory are selected and the forces on each atom computed using the trial force fields. The choice of using the forces for training, rather than the energy, is motivated by the fact that the dynamics is uniquely determined by the forces, and that there is a force vector acting on each atom in contrast to just one global value of the energy.

The cost function used is the sum of a direction cost and a magnitude cost,

$$C = \underbrace{A \tan\left(\frac{\pi}{4}\left(1 - \hat{\boldsymbol{F}}_{\text{ref}} \cdot \hat{\boldsymbol{F}}_{\text{FF}}\right)\right)}_{\text{Direction cost}} + \underbrace{B \log^2\left(\frac{\boldsymbol{F}_{\text{ref}}^2}{\boldsymbol{F}_{\text{FF}}^2}\right)}_{\text{Magnitude cost}},$$

where $\boldsymbol{F}_{\text{ref}}$ and $\boldsymbol{F}_{\text{FF}}$ are the AIMD and force field forces, respectively, $\hat{\boldsymbol{F}}_{\text{ref}}$

and $\hat{\boldsymbol{F}}_{\mathrm{FF}}$ are the corresponding unit vectors and A and B are parameters setting the relative importance of each term. The direction cost is constructed so that it gives a cost of 0 to parallel force vectors and approaches infinity for anti-parallel forces, while the magnitude cost function is 0 if the magnitudes are equal and grows symmetrically in the order of magnitude discrepancy between $\boldsymbol{F}_{\mathrm{ref}}$ and $\boldsymbol{F}_{\mathrm{FF}}$. Such a partition of the cost function was found necessary to avoid the tendency of force constants to decay to 0 early in the optimization. In order to track the progress of the optimization, the relative force error, defined by

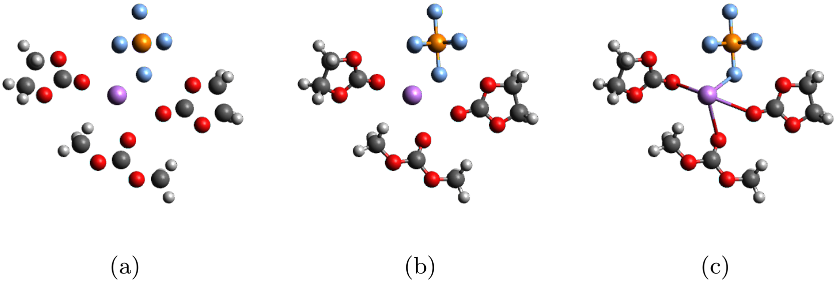
$$E_{\mathrm{rel}} = \langle \frac{|\boldsymbol{F}_{\mathrm{FF}} - \boldsymbol{F}_{\mathrm{ref}}|}{|\boldsymbol{F}_{\mathrm{ref}}|} \rangle,$$

where the average is over evaluated forces, is used as a more intuitive measure of error.

Most of the production data is used in training the force field, but a small part, *ca.* 1000 snapshots, is reserved for out-of-sample testing, to ensure that the learned force field is not only optimized for the specific configurations (snapshots) on which it has been trained (*i.e.* overfitting). The training and testing data should preferably be uncorrelated, which is most easily achieved by discarding part of the trajectory between the training and testing data. The training is successfully completed when the relative force error on the testing data is lower than the chosen tolerance.

3.4 Classical MD Simulations

Classical MD simulations can be performed using the optimized force fields either in NVT or NPT ensembles. The starting geometries are generated using the same routine described in 3.2, whereafter an equilibration needs to be performed before using the subsequent trajectory as production data to be analyzed. NPT simulations are needed if density is used to validate the force field, while NVT simulations can be used to compare transport and structural properties. The latter is thus useful to compare with the AIMD simulations used in the training, in order to assess whether the optimized force field accurately reproduces the structure and dynamics from AIMD (whether or not the AIMD simulations agree with experimental data).



(a) (b) (c)

Figure 3.2: A Li^+ first solvation shell considered on (a) level 0, (b) level 1 and (c) level 2.

3.5 Hierarchical Analysis of Trajectories

The MD trajectoryies are analyzed in terms of mechanical bodies on different hierarchical levels, where a snapshot from a trajectory can be considered on any level. Level 0 is the atomic level, and a snapshot viewed at level 0 is simply a collection of atoms and monoatomic ions (Fig. 3.2a). Level 1 considers the bodies built up by covalent bonds, *i.e.* molecules and molecular ions, but atoms and ions with no covalent bonds to other atoms also have a level 1 representation consisting of that single atom or ion (Fig. 3.2b). A level 2 body is any aggregate of ions and molecules that moves as a dynamical unit due to non-covalent interactions, *e.g.* dispersion, electrostatic interaction or steric hindrance (Fig. 3.2c), *e.g.* first solvation shells, ion pairs or ionic aggregates, and any level 1 body moving freely will have a corresponding level 2 representation consisting of that single level 1 body, *e.g.* a freely moving anion or solvent molecule.

The analysis aims to quantify the transport properties of bodies on the different hierarchical levels. Each body type is characterized by the types of the lower level bodies of which it is composed and the connection graph of these bodies. The only independent transport property of level 0 bodies is a scalar diffusivity. Higher level bodies, on the other hand, have internal structure, and have therefore both translational and rotational transport properties. Their structures also enable the use of a body-fixed coordinate system so that the translational and rotational diffusivities are both vector quantities defined relative to these coordinate systems. In addition, as the bodies of level 2 emerge due to non-bonded interactions and may break up or recombine during

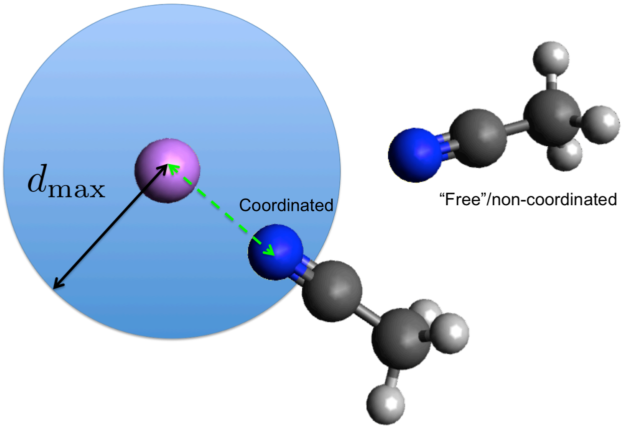


Figure 3.3: The coordination distance between each pair of atom types (level 0 body types) determine which non-covalent aggregates (level 2 bodies) are formed.

an MD simulation, their population dynamics can also be quantified in terms of life-time distributions and ”reaction graphs”.

Since the starting geometries for the MD simulations are known both in terms of level 0 and level 1 bodies, and these do not change during a simulation, the only bodies that need to be ”discovered” during the analysis are those of level 2, which are found using a GA, where the genotypes to be found are pairwise coordination distances between each pair of level 0 body types. The GA evaluates the quality of a set of coordination distances based on their ability to predict the positions of all level 0 bodies in the next time-step based on level 2 rigid body dynamics, where the level 2 bodies are the aggregates of level 1 bodies with pairs of level 0 bodies within their respective coordination distances (Fig. 3.3).

The ability of the framework to switch between hierarchical levels enables for example to find the size and charge distributions of ionic aggregates, to determine the extent to which each contributes to Li^+ diffusion and migration, and to what extent this transport is composed of translation *vs.* rotation of such aggregates. It also shines light on the manner and frequency of ligand exchange. In total the questions that can be answered using this framework should enable a detailed quantitative and qualitative understanding of the transport mechanism present in complex liquid electrolytes.

Chapter 4

Results and Discussion

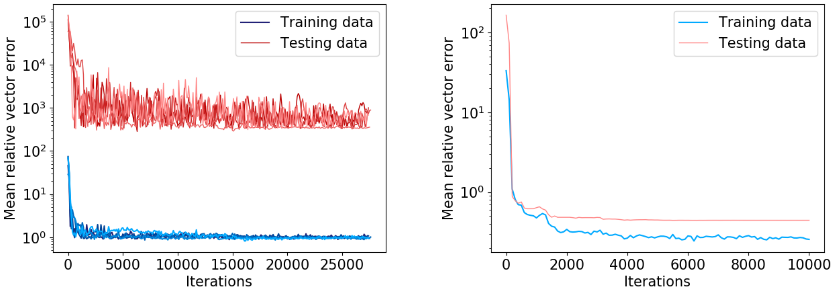
Starting with pure water as a well-studied proof-of-concept system requiring only 9 parameters to be fit, and then moving on to the model highly concentrated electrolyte LiTFSI in ACN (1:2 molar fraction), with *ca.* 200 parameters, preliminary results are discussed as well as likely routes to improving them.

4.1 Water – Proof-of-Concept System

A proof-of-concept of the force field optimization method presented in 3.2-3.3 has been devised and attempted, though not yet successfully achieved (Paper I). For this, pure water was chosen to be simulated to ascertain the validity of the overall approach for a well-investigated system. Here 66 water molecules were put in a cubic periodic box with a side length of 12.55 Å and were simulated. The initial geometry was generated as described in 3.2. The system was thereafter classically pre-equilibrated for 1.0 ns using the well-known water model TIP3P in the OpenMM molecular simulation toolkit, using a Langevin thermostat, with friction coefficient 1 ps^{-1} at 300 K.

Subsequently, an AIMD simulation was run for 2.3 ps, where the first 0.1 ps were used for equilibration and the rest constituted the production run used for the ML optimization of the force field. The AIMD simulation used the Car-Parinello molecular dynamics (CPMD) variant, of AIMD with a fictional electron mass of 400 m_e , and a time-step of 0.1 fs, with the PBE GGA functional, and a 70 Ry plane wave cut-off. The simulation was done in the NVT ensemble at 300 K, using Nosé-Hoover thermostat chains of length four for both electrons and ion cores with frequencies 10 000 cm^{-1} and 3 000^{-1} for the electronic and ionic systems, respectively.

The first 2.0 ps of the AIMD production run was used for training and the last 0.1 ps were reserved for testing. The remaining 0.1 ps were discarded in order for the training and testing data to be uncorrelated. To optimize



(a)

(b)

Figure 4.1: Evolution of the mean relative vector error in force predictions for the best force field in each iteration (lowest training cost) for the water test case. The optimizations aimed to reproduce the forces of a) the AIMD simulation over five separate optimizations, b) an unknown force field (single optimization).

the force field the GA was run for $1{\cdot}10^4$ iterations using 60 individuals in each generation, out of which 20 were selected as parents of the next generation with softmax parameter $\beta = 1 \cdot 10^{-3}$ and exponential decay of mutation amplitudes with $\alpha = 1 \cdot 10^{-3}$, with initial permutation variance equal to 1/10 of the initial guess distributions. Forces were evaluated for ten randomly selected snapshots in each generation.

Five optimizations were performed in parallel to test the robustness of the method. The training error converged to relative errors about 100%, though not reaching the target range of 10%. The testing error, on the other hand, kept oscillating around several orders of magnitude greater error (4.1a). This seems to suggest that no proper training has been achieved, and that the better fit on the training data is largely due to overtraining. Nevertheless, while all parameters were initialized with starting guess widths of 30% of the original guesses, the spread of different parameters in the optimized force fields are very different (Table 4.1). Notably, the Lennard-Jones energy parameters have very large standard deviations while all other parameters have relative standard deviations (RSDs) below 50%, and all equilibrium distances and angles at most 20%.

To verify the implementation of the algorithm under more controlled conditions, another optimization was performed, where the reference forces from

Table 4.1: Mean values, standard deviations (std.) and relative standard deviations (RSD) of the final parameter values in five independent optimizations with the aim of reproducing the AIMD forces, *cf.* Eq. (2.9) for symbol meanings.

	q_{H} [e]	ϵ_{H} [eV]	σ_{H} [Å]	ϵ_{O} [eV]	σ_{O} [Å]	l_0 [Å]	k_l [eVÅ^{-2}]	θ_0	k_θ [eV rad^{-2}]
Mean	0.17	4.5e-5	2.6	1.3e-5	2.6	0.99	4.0e5	97°	11
Std.	0.029	1.2e-4	0.27	6.4e-5	0.19	0.026	5.0e4	5.2°	5.0
RSD	17%	280%	20%	510%	7.5%	0.27%	13%	5.4%	47%

the AIMD simulations were replaced by the forces predicted by the TIP3P water model (Fig. 4.1b). Under these conditions, a perfect fit is known to be possible as the reference forces are generated by a force field of the same functional form as the ones being optimized. Here the error evaluated on both the test and the training snapshots converged to well within 100% error, demonstrating that the algorithm works, although a systematic survey of the hyperparameters controlling the GA should be performed in order to maximize convergence. Our interpretation of the difference between these outcomes is that a good fit of the AIMD forces likely does not exist within the space of force fields of the form of Eq. (2.9). The most likely reason for this is that the force fields do not consider induced polarization of water molecules, which is expected to be important due to the very polar nature of water.

4.2 Highly Concentrated Electrolytes

The layout of this ongoing study will by necessity differ to some extent from the proof-of-concept. These differences and implications are provided here.

First, as compared to the force field for pure water, a highly concentrated electrolyte system like LiTFSI in ACN (1:2) would require more than an order of magnitude more parameters, *ca.* 200 as compared to 9. This is likely a strictly more difficult problem. As a successful force field optimization for pure water has not yet been achieved, the exact nature of the optimization of force fields for LiTFSI in ACN has been postponed, but we envisage the workflow to have the same basic layout. However, it will still likely differ in the following aspects: i) longer equilibration times will be needed as this system is more viscous than water, ii) longer AIMD runs may be needed for the force field training as more parameters need to be optimized, iii) more iterations will probably be needed for the GA to fit the greater number of parameters, and iv) larger population size may be required in the GA to maintain sufficient genetic diversity in the population to avoid getting stuck in broad local minima. In addition, it is very likely that polarization effects need to be considered also for electrolytes.

The above is clearly not results for the materials at hand nor for any transport, but for the methodology development of the framework – which is indeed a substantial part of the goals of this thesis work.

Chapter 5

Future Research

Further work primarily includes honing the method so that better fit between force fields and AIMD data are obtained both for pure water and for highly concentrated electrolytes. Upon successful generation of force fields for highly concentrated electrolytes, hierarchical analysis, as described in 3.5, will be pursued. Once the full method presented in this thesis has been validated, it will also be applied to study the ion transport in electrolytes for Li-S batteries, containing dissolved PSs.

The source code developed within this thesis will be released to the public upon reaching sufficient maturity.

Chapter 6

Acknowledgments

First of all, I would like to thank my funders: the Swedish Energy Agency, the EU H2020 through HELIS (Grant Agreement No. 666221), and the Chalmers Area of Advance Materials Science/Theory and Modelling for allowing me to do this very rewarding work. Thanks to my main supervisor Patrik Johansson for giving me this opportunity, and granting me the trust and freedom to make this project very much my own, even when I have taken it on different paths than you may have originally envisioned. This would also not have worked if you hadn't guided me back on track whenever my visions have taken me too far astray. I would also like to thank my co-supervisor Alejandro Franco, for very inspiring discussions about modeling, especially MSM.

Thanks to Fabian Årén, who implemented the GA used in this thesis as his MSc thesis project. I could not have wished for a better master student to help in the development of these methods and with whom to discuss the problems we faced in the process. You very much deserve your new position as my fellow PhD student, and I hope there will be some points of contact between our projects also in the future.

Thanks to all the colleagues with whom I have had fruitful discussions about my project. Special thanks in this regard go to Gustav Åvall, Viktor Nilsson, Piotr Jankowski, Rafael Barros Neves de Araújo, Alexandr Nasedkin, Christoffer Olsson and Joachim Wallenstein.

Thanks also to all my wonderful colleagues who make the social life at KMF such a blast, at and out of work (even as I regrettably miss too much of that these days).

A very, very special thanks to my wife, Emily, who lets me periodically become completely absorbed with work for extended periods of time. Fulfilling dreams is always a matter of sacrifice, and not just one's own. I will always be grateful for your support of my endeavors.

Finally, to my son Elliot, who is 6 months old at the time of writing: thanks for always waking me up with a smile that encourages me to face the day even when stress levels would otherwise be intolerably high.

References

[1] H. Ritchie and M. Roser. (2018). CO_2 and other Greenhouse Gas Emissions, [Online]. Available: `https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions` (visited on 07/24/2018).

[2] Y. Nishi, "The development of lithium ion secondary batteries", *The Chemical Record*, vol. 1, no. 5, pp. 406–413,

[3] B. Schott, A. Püttner, and M. Müller, "3 - The market for battery electric vehicles", in *Advances in Battery Technologies for Electric Vehicles*, ser. Woodhead Publishing Series in Energy, B. Scrosati, J. Garche, and W. Tillmetz, Eds., Woodhead Publishing, Jan. 2015, pp. 35–54.

[4] K. Xu, "Nonaqueous Liquid Electrolytes for Lithium-Based Rechargeable Batteries", *Chem. Rev.*, vol. 104, no. 10, pp. 4303–4418, Oct. 2004.

[5] R. Petibon, J. Xia, L. Ma, M. K. G. Bauer, K. J. Nelson, and J. R. Dahn, "Electrolyte System for High Voltage Li-Ion Cells", *J. Electrochem. Soc.*, vol. 163, no. 13, A2571–A2578, Jan. 2016.

[6] H. Berg, *Batteries for Electric Vehicles: Materials and Electrochemistry.* Cambridge university press, 2015.

[7] Y. Wu, *Lithium-Ion Batteries: Fundamentals and Applications.* CRC Press, 2015, vol. 4.

[8] N. Nitta, F. Wu, J. T. Lee, and G. Yushin, "Li-ion battery materials: Present and future", *Materials Today*, vol. 18, no. 5, pp. 252–264, Jun. 2015.

[9] X. Ji and L. F. Nazar, "Advances in LiS batteries", *J. Mater. Chem.*, vol. 20, no. 44, pp. 9821–9826, Nov. 2010.

[10] L. F. Nazar, M. Cuisinier, and Q. Pang, "Lithium-sulfur batteries", *MRS Bulletin*, vol. 39, no. 5, pp. 436–442, May 2014.

[11] J. W. Choi and D. Aurbach, "Promise and reality of post-lithium-ion batteries with high energy densities", *Nature Reviews Materials*, vol. 1, p. 16 013, Mar. 2016.

[12] N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements.* Elsevier, 2012.

[13] J. Scheers, S. Fantini, and P. Johansson, “A review of electrolytes for lithium-sulphur batteries”, *Journal of Power Sources*, vol. 255, pp. 204–218, Jun. 2014.

[14] B. Flamme, G. R. Garcia, M. Weil, M. Haddad, P. Phansavath, V. Ratovelomanana-Vidal, and A. Chagnes, “Guidelines to design organic electrolytes for lithium-ion batteries: Environmental impact, physicochemical and electrochemical properties”, *Green Chemistry*, vol. 19, no. 8, pp. 1828–1849, 2017.

[15] L. Suo, O. Borodin, T. Gao, M. Olguin, J. Ho, X. Fan, C. Luo, C. Wang, and K. Xu, “Water-in-salt electrolyte enables high-voltage aqueous lithium-ion chemistries”, *Science*, vol. 350, no. 6263, pp. 938–943, Nov. 2015.

[16] V. Etacheri, R. Marom, R. Elazari, G. Salitra, and D. Aurbach, “Challenges in the development of advanced Li-ion batteries: A review”, *Energy & Environmental Science*, vol. 4, no. 9, pp. 3243–3262, 2011.

[17] E. Peled, “The Electrochemical Behavior of Alkali and Alkaline Earth Metals in Nonaqueous Battery SystemsThe Solid Electrolyte Interphase Model”, *J. Electrochem. Soc.*, vol. 126, no. 12, pp. 2047–2051, Jan. 1979.

[18] E. Peled and S. Menkin, “Review – SEI: Past, Present and Future”, *J. Electrochem. Soc.*, vol. 164, no. 7, A1703–A1719, Jan. 2017.

[19] S. S. Zhang, “A review on electrolyte additives for lithium-ion batteries”, *Journal of Power Sources*, Special issue including selected papers from the International Power Sources Symposium 2005 together with regular papers, vol. 162, no. 2, pp. 1379–1394, Nov. 2006.

[20] Y. Yamada and A. Yamada, “Superconcentrated electrolytes for lithium batteries”, *Journal of The Electrochemical Society*, vol. 162, no. 14, A2406–A2423, 2015.

[21] K. Matsumoto, K. Inoue, K. Nakahara, R. Yuge, T. Noguchi, and K. Utsugi, “Suppression of aluminum corrosion by using high concentration LiTFSI electrolyte”, *Journal of Power Sources*, vol. 231, pp. 234–238, Jun. 2013.

[22] Y. Yamada, C. H. Chiang, K. Sodeyama, J. Wang, Y. Tateyama, and A. Yamada, “Corrosion Prevention Mechanism of Aluminum Metal in Superconcentrated Electrolytes”, *ChemElectroChem*, vol. 2, no. 11, pp. 1687–1694, Nov. 2015.

[23] V. Nilsson, R. Younesi, D. Brandell, K. Edström, and P. Johansson, “Critical evaluation of the stability of highly concentrated LiTFSI - Acetonitrile electrolytes vs. graphite, lithium metal and $LiFePO_4$ electrodes”, *Journal of Power Sources*, vol. 384, pp. 334–341, Apr. 2018.

[24] X. Wang, E. Yasukawa, and S. Mori, “Inhibition of anodic corrosion of aluminum cathode current collector on recharging in lithium imide electrolytes”, *Electrochimica Acta*, vol. 45, no. 17, pp. 2677–2684, May 2000.

[25] H. Lundgren, J. Scheers, M. Behm, and G. Lindbergh, “Characterization of the mass-transport phenomena in a superconcentrated litfsi: Acetonitrile electrolyte”, *Journal of The Electrochemical Society*, vol. 162, no. 7, A1334–A1340, 2015.

[26] J. S. Newman and T.-A. K. E, *Electrochemical Systems.* John Wiley & Sons, 2004.

[27] M. Doyle and J. Newman, “The use of mathematical modeling in the design of lithium/polymer battery systems”, *Electrochimica Acta*, International symposium on polymer electrolytes, vol. 40, no. 13, pp. 2191–2196, Oct. 1995.

[28] W. Lai and F. Ciucci, “Mathematical modeling of porous battery electrodes – Revisit of Newman’s model”, *Electrochimica Acta*, vol. 56, no. 11, pp. 4369–4377, Apr. 2011.

[29] S. A. Krachkovskiy, J. D. Bazak, S. Fraser, I. C. Halalay, and G. R. Goward, “Determination of Mass Transfer Parameters and Ionic Association of LiPF6: Organic Carbonates Solutions”, *J. Electrochem. Soc.*, vol. 164, no. 4, A912–A916, Jan. 2017.

[30] O. Borodin, M. Olguin, P. Ganesh, P. R. C. Kent, J. L. Allen, and W. A. Henderson, “Competitive lithium solvation of linear and cyclic carbonates from quantum chemistry”, *Phys. Chem. Chem. Phys.*, vol. 18, no. 1, pp. 164–175, Dec. 2015.

[31] D. M. Seo, O. Borodin, S.-D. Han, P. D. Boyle, and W. A. Henderson, “Electrolyte Solvation and Ionic Association II. Acetonitrile-Lithium Salt Mixtures: Highly Dissociated Salts”, *Journal of The Electrochemical Society*, vol. 159, no. 9, A1489–A1500, 2012.

[32] E. Flores, G. Åvall, S. Jeschke, and P. Johansson, “Solvation structure in dilute to highly concentrated electrolytes for lithium-ion and sodium-ion batteries”, *Electrochimica Acta*, vol. 233, pp. 134–141, Apr. 2017.

[33] D. M. Seo, O. Borodin, D. Balogh, M. O'Connell, Q. Ly, S.-D. Han, S. Passerini, and W. A. Henderson, "Electrolyte Solvation and Ionic Association III. Acetonitrile-Lithium Salt Mixtures – Transport Properties", *Journal of The Electrochemical Society*, vol. 160, no. 8, A1061–A1070, 2013.

[34] D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press, Apr. 2009.

[35] G. Kaminski and W. L. Jorgensen, "Performance of the AMBER94, MMFF94, and OPLS-AA Force Fields for Modeling Organic Liquids", *J. Phys. Chem.*, vol. 100, no. 46, pp. 18 010–18 013, Jan. 1996.

[36] J. W. Ponder and D. A. Case, "Force Fields for Protein Simulations", in *Advances in Protein Chemistry*, ser. Protein Simulations, vol. 66, Academic Press, Jan. 2003.

[37] D. Bedrov, O. Borodin, Z. Li, and G. D. Smith, "Influence of Polarization on Structural, Thermodynamic, and Dynamic Properties of Ionic Liquids Obtained from Molecular Dynamics Simulations", *J. Phys. Chem. B*, vol. 114, no. 15, pp. 4984–4997, Apr. 2010.

[38] J. M. Thijssen, *Computational Physics*. Cambridge University Press, Mar. 2007.

[39] W. Nolting, *Fundamentals of Many-Body Physics: Principles and Methods*. Berlin Heidelberg: Springer-Verlag, 2009.

[40] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", *Phys. Rev.*, vol. 136, B864–B871, 3B Nov. 1964.

[41] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", *Phys. Rev.*, vol. 140, A1133–A1138, 4A Nov. 1965.

[42] J. P. Perdew and K. Schmidt, "Jacobs ladder of density functional approximations for the exchange-correlation energy", *AIP Conference Proceedings*, vol. 577, no. 1, pp. 1–20, Jul. 2001.

[43] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", *Phys. Rev. Lett.*, vol. 77, no. 18, pp. 3865–3868, Oct. 1996.

[44] R. Car and M. Parrinello, "Unified Approach for Molecular Dynamics and Density-Functional Theory", *Phys. Rev. Lett.*, vol. 55, no. 22, pp. 2471–2474, Nov. 1985.

[45] S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, “An all atom force field for simulations of proteins and nucleic acids”, *Journal of Computational Chemistry*, vol. 7, no. 2, pp. 230–252,

[46] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, “Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids”, *J. Am. Chem. Soc.*, vol. 118, no. 45, pp. 11 225–11 236, Jan. 1996.

[47] O. Kramer, *Genetic Algorithm Essentials*, ser. Studies in Computational Intelligence. Springer International Publishing, 2017.

[48] D. Case *et al.*, “AMBER 10. 2008”, *University of California, San Francisco*, 2008.